Principles and Recommendations for Population and Housing Censuses

Revision 4

United Nations

PART ONE. ESSENTIAL FEATURES AND CENSUS METHODOLOGY 1
I. Essential roles of the census 1
II. Definitions and key features 2
A. Definitions 2
1. Population census 2
2. Housing census 3
B. Key features 3
1. Individual enumeration 3
2. Universality within a defined territory 3
3. Simultaneity 3
4. Defined periodicity 4
5. Capacity to produce small-area statistics 4
III. Uses of population and housing censuses 4
A. Uses of population censuses 6
1. Uses for policymaking, planning and administrative purposes 6
2. Uses for research purposes 6
3. Uses for business, industry and labour 7
4. Uses for boundary delimitation 7
5. Use as a sampling frame for surveys 7
6. Use in humanitarian settings 8
7. Use for monitoring internationally agreed development goals and international comparisons 9
8. Use for population estimates and projections 10
9. Use in small area estimation 10
10. Use in integrated population data systems 11
B. Uses of housing censuses 12
1. Uses for development of benchmark housing statistics 12
2. Uses for the formulation of housing policy and programmes 12
3. Assessment of the quality of housing 13
C. Relationship between the population census and the housing census 13
D. Relationship of population and housing censuses to intercensal sample surveys 14
E. Relationship of population and/or housing censuses to other types of censuses and other statistical investigations 15
1. Census of agriculture 15
2. Census of establishments 16
3. Census of buildings 17
4. System of current housing statistics 17
F. Relationship between the population census and/or housing census and administrative data sources 18
1. Civil registration and vital statistics 18
2. Population register 19
3. Building/dwelling or address register 20
4. Business register 20
5. Other administrative data sets 21
G. Integration of census data with geospatial information and other data sources 22
IV. Census methodology 26

A. Full field enumeration 29
1. Types of censuses conducted with full-field enumeration 30
2. Innovations in conducting field-based censuses 34
B. Register-based censuses 36
1. Necessary pre-conditions for the use of administrative registers for censuses 37
2. Types of registers 41
3. Establishing a statistical population register 43
4. Constructing integrated statistical registers 44
5. Transitioning from field-based censuses to register-based approaches 45
6. Types of register-based censuses 46
7. Advantages and challenges 47
C. Combined census methodologies 49
1. Types of combined census methodologies 51
2. Advantages and challenges 52
V. Methods for estimating population for areas not enumerated in the census 54
A. Requirements for estimating population 54
B. Methodological approaches 55
1. Use of administrative data sources 55
2. Use of sample surveys 56
3. Modelled methods 56
VI. Strategies for producing more frequent population and housing statistics 57
A. The need for more frequent data on population and housing characteristics 57
B. Use of administrative registers 58
C. Intercensal surveys 58
D. Annual continuous surveys 59
E. Use of mobile phone data 59

PART TWO. PLANNING, ORGANIZATION AND MANAGEMENT 61
I. Introduction 61
II. Census planning 61
A. Overall census planning 61
B. Planning for both a population and housing census 63
C. Special considerations when planning a combined or register-based census 63
III. Strategic objectives and management 64
A. Strategic objectives 64
B. Strategic management 66
C. Avoiding gender biases and biases affecting indigenous peoples and minorities 68
IV. Units, place and time of enumeration 69
A. Units of enumeration 69
1. Person 70
2. Household 70
3. Population in collective living quarters 72
4. Building 72
5. Living quarters 73
B. Place of enumeration 73
1. Concepts relating to the place of enumeration 73
2. Operational issues relating to the place of residence and the place of enumeration 75

C. Enumeration point of time 76
D. Time reference period for data on the characteristics of the population and of living quarters 77
V. Legal basis 77
VI. Financial management 79
A. Financial basis for censuses 79
B. Budget and cost control 81
VII. Resource mobilization 83
VIII. Project management 84
A. Development of workplans 84
B. Emergency preparedness, risk management and contingency planning 85
1. Documentation of risks and mitigation plans 85
2. Emergency management, governance and decision-making 86
3. Disinformation and misinformation threats to the census 88
4. Crisis response training for census staff 88
5. Ensuring fast workflow during an emergency 89
6. Contingency planning for main plan failure 89
C. Change management 90
IX. Administrative organization 92
A. Overall overview 92
B. Roles and responsibilities of key stakeholders in the census process 93
C. Statistical leadership 95
X. User consultation, communication and publicity 95
XI. Census calendar 103
XII. Human resources management 105
XIII. Logistics management 107
A. Procurement management 108
B. Forward and reverse logistics 109
XIV. Contracting out 110
XV. Use of technology 114
XVI. Innovation to optimize census processes 118

PART THREE. CENSUS OPERATION ACTIVITIES 120
I. Introduction 120
II. Census questionnaires: content and design 120
a) Selection of census topics 121
b) Use of short and long questionnaires 121
c) Sections of the questionnaire 123
d) Questionnaire design 123
e) Design of questionnaires for electronic data collection (CAPI, CAWI, CATI) 126
III. Building census infrastructure 126
IV. Geospatial information and mapping 128
A. Strategic planning 128
B. The role of geospatial information in the census 131
C. Geospatial information technology for census 134
D. Geographic information systems 136
E. Census geography 139
F. Implementation of a census geospatial information programme 144

G. Contracting out and partnerships for census geospatial operations 148
V. Census tests 149
VI. Establishing the census frame 152
VII. Field enumeration 153
A. Method of enumeration 154
B. Multi-mode data collection 158
C. Use of administrative data during field-based enumeration 161
D. Timing and length of the enumeration period 162
E. Management and supervision 163
F. Security during data collection 166
G. Use of technology 167
VIII. Data processing 171
A. Method of processing 174
B. Preparation for data capture 175
C. Data capture 176
D. Coding 178
E. Data editing 180
F. Validation 181
G. Processing control 182
H. Master file 183
I. Methods of tabulation 184
J. Use of administrative data during data processing (for field-based censuses) 185
K. Security during data processing 186
IX. Census products, dissemination and utilization 187
A. Introduction 187
B. Plans for census products and data dissemination 188
1. Developing a dissemination strategy 188
2. Tabulation programme 193
3. Geospatial information for analysis and dissemination 193
4. Mode of dissemination of outputs 197
5. Confidentiality and privacy 199
6. Metadata 201
7. Promotion of, and training on, uses of census data 203
C. Census data dissemination: products and services 204
1. Provisional and final results 204
2. Census reports 205
3. Databases 210
4. Geographic products 214
5. Interactive electronic outputs 219
6. Microdata dissemination 220
7. Customized products 225
8. General interest products and special audience products 226
D. Census data utilization 229
1. General uses of population and housing censuses 229
2. Analysis of the results 232
3. Intercensal and post-census population estimates 232
4. Cross-cutting and emerging social issues 233
5. Development indicators 239

X. Documentation of census experience 240
XI. Archiving individual records 241
A. Purpose of archiving individual records 241
B. Procedures for archiving 241
C. Archiving individual records and microdata 242

PART FOUR. QUALITY ASSURANCE AND EVALUATION 244
I. Quality assurance 244
A. Plans for quality assurance 244
B. Quality assurance components 244
1. Managing the census within the national statistical system 245
2. Managing the institutional environment 245
3. Managing census processes 245
4. Managing census outputs 246
C. Other international standards, tools and references 247
D. The role of managers 248
E. Quality improvement and the census 250
1. Topic selection 251
2. Questionnaire design and development 252
3. Geospatial information and mapping 252
4. Development of the data capture system 254
5. Field operations and pilot testing 256
6. Processing 258
7. Analysis 261
8. Dissemination 262
9. Evaluation 263
10. Documentation 264
F. Measuring the quality of data derived from administrative sources 264
II. Evaluation of the results 266
A. Purpose of census evaluation 267
B. Methods of census evaluation 269
C. Post-enumeration survey 270
D. Demographic analysis for census evaluation 271
E. Comparison with administrative sources 273
F. Adjusting census data 274
1. Adjustments for full field enumeration censuses 275
2. Adjustments for register-based censuses 275
3. Adjustments for combined censuses 276
4. The decision to adjust, or not to adjust, census figures 277
G. Acceptance of results 277
H. User satisfaction 278
III. Overall evaluation of the census 279
A. Importance of evaluations 279
B. Planning for the evaluation 280

PART FIVE. POPULATION AND HOUSING CENSUS TOPICS 282
Chapter I: Population census topics 282
I. Factors determining the selection of topics 282

A. National priority 283
B. International comparability 283
C. Suitability 284
D. Resources 284
E. Alternative sources 284
II. List of topics 285
III. Population count 288
A. Population present count 289
B. Usual resident population count 290
C. Supplementary population counts 292
1. Service population count 292
2. Counts of population subgroups 293
D. Difficult-to-enumerate groups or difficult-to-reach groups 293
IV. Definitions and specifications of topics 297
A. Geographic and internal migration characteristics 297
1. Place of usual residence (core topic) 297
2. Place where present at time of census (core topic) 298
3. Place of birth (core topic) 299
4. Duration of residence (core topic) 300
5. Place of previous residence (core topic) 300
6. Place of residence at a specified date in the past (core topic) 300
7. Reason for change of residence 301
8. Total population (core topic) 302
9. Locality (core topic) 303
10. Urban and rural (core topic) 304
B. International migration characteristics 306
1. Country of birth (core topic) 306
2. Country of citizenship (core topic) 307
3. Acquisition of citizenship 309
4. Year or period of arrival in the country (core topic) 309
5. Ever resided abroad 309
6. Reason for international migration 310
C. Household and family characteristics 311
1. Relationship to the reference person of household (core topic) 313
2. Household and family composition (core topic) 315
3. Household and family status 317
D. Demographic and social characteristics 318
1. Sex (core topic) 319
2. Age (core topic) 319
3. Marital status (core topic) 321
4. Ethnocultural characteristics 323
5. Religion 323
6. Language 324
7. Ethnicity 324
8. Indigenous peoples 325
9. Disability characteristics 326
E. Fertility and mortality 331
1. Children ever born alive (core topic) 334

2. Children living (core topic) 335
3. Date of birth and sex of last child born alive (core topic) 336
4. Births in the past 12 months 337
5. Deaths among children born in the past 12 months 337
6. Age, date or duration of first marriage or union 338
7. Age of mother at birth of (date or time since) first child born alive 338
8. Household deaths in the past 12 months (core topic) 339
9. Maternal or paternal orphanhood 340
F. Registration of vital events and assignment of legal identity 341
G. Educational characteristics 344
1. Literacy (core topic) 344
2. School attendance (core topic) 345
3. Educational attainment (core topic) 346
4. Field of education and educational qualifications 348
H. Individual use of information and communication technology 349
1. Individuals using the Internet 350
2. Individuals owning mobile cellular phones 351
I. Economic characteristics 351
1. Introduction 351
2. Conceptual framework for work statistics 353
3. Labour force status (core topic) 356
4. Characteristics of jobs and establishments 363
5. Status in employment (core topic) 364
6. Occupation (core topic) 368
7. Industry (core topic) 369
8. Place of work 370
9. Institutional sector of employment 371
10. Working time 372
11. Participation in own use production of goods (core topic) 373
12. Income 374
13. Informal employment 375
J. Agriculture 377
1. Introduction 377
2. Own-account agriculture production 377
3. Size of farm and number of livestock 377
Chapter II. Housing census topics 379
I. Factors determining the selection of topics 379
A. National priority 380
B. International comparability 381
C. Suitability 381
D. Resources 381
E. Alternative sources 381
II. List of topics 381
III. Definitions and specifications of topics 384
1. Living quarters – type of (core topic) 384
2. Location of living quarters (core topic) 393
3. Occupancy status (core topic) 394
4. Ownership – type of (core topic) 395

5. Rooms – number of (core topic) 396
6. Bedrooms – number of 397
7. Useful floor space 397
8. Water supply system (core topic) 397
9. Drinking water – main source of (core topic) 398
10. Toilet – type of (core topic) 399
11. Sewage disposal (core topic) 400
12. Solid waste disposal – main type of (core topic) 401
13. Bathing facilities (core topic) 401
14. Kitchen – availability of (core topic) 402
15. Fuel used for cooking (core topic) 403
16. Energy used for lighting – type of (core topic) 403
17. Heating – type and energy used 404
18. Hot water – availability of 404
19. Piped gas – availability of 404
20. Use of the housing unit 404
21. Occupancy by one or more households (core topic) 405
22. Occupants – number of (core topic) 405
23. Building – type of (core topic) 405
24. Year or period of construction 407
25. Dwellings in the building – number of 408
26. Position of dwelling in the building 408
27. Accessibility to dwelling 409
28. Construction material of outer walls (core topic) 409
29. Construction material of floor and roof 410
30. Elevator – availability of 410
31. Farm building 411
32. State of repair 411
33. Age and sex of the reference person of household (core topic) 411
34. Tenure (core topic) 411
35. Rental and housing costs 412
36. Furnished/unfurnished 413
37. Information and communication technology devices – availability of (core topic) 413
38. Cars – number of available 414
39. Durable household appliances – availability of 414
40. Outdoor space – access to 414

PART ONE. ESSENTIAL FEATURES AND CENSUS METHODOLOGY

I. Essential roles of the census

1.1. Evidence-based decision-making is a universally recognized paradigm of efficient management of economic and social affairs and of overall effective governing of societies today. Making use of relevant, accurate and timely statistics is essential to this model; producing detailed statistics for small areas and small population groups is its foundation. The role of the population and housing census is to produce and disseminate such detailed small-area statistics on population, its composition, characteristics, spatial distribution and organization into families, households and dwellings. Censuses have been promoted internationally since the end of the nineteenth century, when the International Statistical Congress recommended that all countries in the world conduct them.[1] Since 1958, the United Nations has also been actively promoting the population and housing census by compiling the principles and recommendations for population and housing censuses and launching regular decennial worldwide programmes on population and housing censuses. In fact, the majority of the countries in the world conduct a population and housing census periodically. In order to increase the relevance of the data produced by censuses, several countries are now moving towards using alternative census methodologies to produce census-like statistics with a more frequent periodicity than the traditional decennial or quinquennial one.

1.2. While the roles of the population and housing census are many and will be elaborated in detail throughout the present revision of these *Principles and Recommendations for Population and Housing Censuses*, several of the essential roles are summarised here:

(a) **The population and housing census plays an essential role in public administration**. The results of a census are used as a critical reference to ensure equity in distribution of wealth, government services and representation nationwide by informing the distribution and allocation of government funds among various regions and districts for education, health services, delineating electoral districts at the national and local levels, and measuring the impact of industrial development, to name a few. Establishing a public consensus on priorities would be almost impossible to achieve if it were not built on census counts. A wide range of other users, including the private sector, academia, civil society and individuals, make use of census outputs.

(b) **The census also plays an essential role in all elements of the national statistical system**, including the economic and social components. Census statistics are used as benchmarks for statistical compilation, definitions and classifications, and as a sampling frame for sample surveys. Today, the national statistical system of almost every country relies on sample surveys for efficient and reliable data collection, notwithstanding the increasing availability of non-traditional data sources. Without the sampling frame and population benchmarks derived from the population and housing census, the national statistical system would face difficulties in providing reliable official statistics for use by the government and the general public. For many countries with less developed statistical systems, a

[1] *Report of the Proceedings of the Fourth Session of the International Statistics Congress, held in London July 16th, 1860, and the Five following Days*, printed by George Edward Eyre and William Spottiswoode, London, 1861.

census is also an essential source providing data to integrate with survey data for producing model-based small-area estimations.

(c) **A basic function of a census is to generate statistics on small areas and small population groups**. While statistics on small areas are useful per se, they are important because they can be used to produce statistics on any geographic unit with arbitrary boundaries, subject to any statistical disclosure control restrictions for overlapping boundaries. For example, in planning the location of a school, it is necessary to have the data on the distribution of school-age children by school area, which may not necessarily correspond to the administrative area units. Similarly, small-area data from the census can be combined to approximate natural regions (for example, water catchments or vegetation zones) that do not follow administrative boundaries. Since census data can be tabulated for any geographic unit, it is possible to provide the required statistics in a remarkably flexible manner. This versatile feature of the census is also invaluable for use in the private sector for applications such as business planning and market analyses.

(d) **Census data are used as a benchmark for research and analysis**. They are used to study population dynamics. Population projections are one of the most important analytical outputs based on census data; future population projections are crucial for all segments of the public and private sectors.

1.3. It is critically important to produce detailed statistics for small areas and small population groups **as a building block for efficient governance at all levels**. For many countries the method for assembling this building block will be by conducting a population and housing census through universal and simultaneous individual enumeration of each set of living quarters and of each individual within the country's boundaries. Some countries will adopt alternative approaches; yet, all of these methods should produce detailed, timely statistics for small areas and small population groups at a well-defined point or period in time.

II. Definitions and key features

A. Definitions

1. Population census

1.4. A population census is the total process of planning, collecting, processing, analysing, disseminating and evaluating demographic, economic and social data at the smallest geographic level pertaining, at a specified time, to all persons in a country or in a well-delimited part of a country.

1.5. Population is basic to the production and distribution of material wealth. In order to plan for, and implement, economic and social development, administrative activity or scientific research, it is necessary to have reliable and detailed data on the size, distribution and composition of population. The population census is a primary source of these basic benchmark statistics, covering not only the settled population but also homeless persons and nomadic groups. Data from population censuses should allow presentation and analysis of statistics on persons and households and for a wide variety of geographic units, ranging from the country as a whole to individual small localities or city blocks.

2. Housing census

1.6. A housing census is the total process of planning, collecting, processing, analysing, disseminating and evaluating statistical data relating to the number and condition of housing units and facilities as available to the households pertaining, at a specified time, to all living quarters[2] and occupants thereof in a country or in a well-delimited part of a country.

1.7. The census must provide information on the supply of housing units together with information on the structural characteristics and facilities that have a bearing upon the maintenance of privacy and health and living conditions. Sufficient demographic, social and economic data concerning the occupants must be collected to furnish a description of housing conditions and also to provide basic data for analysing the causes of housing deficiencies and for studying possibilities for remedial action. In this connection, data obtained as part of the population census, including data on homeless persons,[3] are often used in the presentation and analysis of the results of the housing census, if both operations are conducted together or there is a link between them.

B. Key features

1.8. The key features of population and housing censuses are individual enumeration, universality within a defined territory, simultaneity, defined periodicity and small-area statistics.

1. Individual enumeration

1.9. The term "census" implies that each individual and each set of living quarters is enumerated separately and that the characteristics thereof are separately recorded. Only by this procedure can the data on the various characteristics be cross-classified. The requirement of individual enumeration can be met by the collection of information in the field, by the use of information contained in an appropriate administrative register or set of registers, or by a combination of these methods.

2. Universality within a defined territory

1.10. The census should cover a precisely defined territory (for example, the entire country or a well-delimited part of it). The population census should include every person present and/or residing within its scope, depending upon the type of population count required, and should include them only once. The housing census should include every set of living quarters irrespective of type, and should include each one only once. This does not preclude the use of sampling techniques for obtaining data on specified characteristics, provided that the sample design is consistent with the size of the areas for which the data are to be tabulated and the degree of detail in the cross-tabulations to be made.

3. Simultaneity

1.11. Each person and each set of living quarters should be enumerated at the same well-defined point in time and the data collected should refer to a well-defined reference period. The key benefit of referring all data to a common date or reference period is to minimise over-and under-counting. The time reference

[2] For the definition of "living quarters", see paragraph 5.485.
[3] For the definition of “homeless persons”, see paragraph 2.39.

period need not, however, be identical for all of the data collected. For most of the data, it will be the day of the census; in some instances, it may be a period prior to the census.[4] However, it is recognised that not all census methodologies satisfy this criterion. [5]

4. Defined periodicity

1.12. Censuses should, ideally, be taken at regular intervals so that comparable information is made available in a fixed sequence. A series of censuses makes it possible to appraise the past, accurately describe the present and estimate the future. It is recommended that a national census be taken at least every 10 years. Some countries may find it necessary to carry out censuses more frequently because of the rapidity of major changes in their population and/or its housing circumstances. The use of data derived from administrative sources may provide the opportunity for more frequently produced statistical outputs.

1.13. The census data of any country are of greater value nationally, regionally and internationally if they can be compared with the results of censuses of other countries that were taken at approximately the same time. Therefore, countries should make all efforts to undertake a census in years ending in "0" or at a time as near to those years as possible[6]. It is obvious, however, that legal, administrative, financial and other considerations often make it inadvisable for a country to adhere to a standard international pattern in the timing of its censuses. Some regional groupings of countries, such as the European Union, have their own requirements regulating the dates of censuses among their members. In fixing a census date, therefore, such national and regional factors could be given greater weight than the desirability of international simultaneity.

5. Capacity to produce small-area statistics

1.14. The census should produce data on the number and characteristics of the population and housing units down to the lowest appropriate geographic level, compatible with national circumstances, and for small population groups, while protecting confidentiality of personal information on each individual.

III. Uses of population and housing censuses

1.15. Population and housing censuses are a principal means of collecting basic population and housing statistics. They form the core of an integrated programme of data collection and compilation aimed at providing a comprehensive source of statistical information for economic and social development planning, administration, assessing conditions in human settlements, research and commercial and other uses. Population and housing censuses are often a basis for sample surveys and provide insights and updates into buildings, household and population characteristics as well as increasingly being used as a source of environmental and sustainable development statistics and geospatial information – addresses, coordinates, population distribution, buildings and settlements, etc.

[4] For example, collecting information on the core topic of household deaths in the past 12 months (see paragraphs 5.272-5.276).

[5] Simultaneity is an essential feature of censuses that rely on full field enumeration (the so-called traditional approach) (see paragraphs 1.134) as well as of some types of register-based censuses (see paragraph 1.171).

[6] For EU countries it is mandatory to have censuses in years ending in "1".

1.16. The value of population and housing censuses is increased if the results can be employed together with other data sources, as in the use of the census data as a basis or benchmark for current statistics, and if it can furnish the information needed for conducting other statistical data collection. It can, for example, provide a statistical frame for other censuses or sample surveys. The population census is also important in developing the population estimates needed to calculate vital rates in combination with civil registration data (see paragraphs 1.84–1.89). In addition, these censuses are a major source of data used in official compilations of social indicators, particularly on topics that usually change slowly over time. It is important, therefore, to take into account the requirements of a continuous, coordinated programme of data collection and compilation when planning a census. When the relationships between the population census, the housing census and other statistical investigations are considered from the outset, their shared purposes can best be served and their benefits best harnessed. The use of consistent concepts and definitions throughout an integrated programme of data collection and compilation is essential if the advantages of these relationships are to be fully realized. Of course, census-type information can also be derived from population registers and can also be estimated from sample surveys without undertaking a complete enumeration. These alternative data sources are discussed under "Census methodology" in paragraphs 1.128–1.255.

1.17. Data to measure and monitor indicators are required by countries to track the progress towards internationally agreed development agendas, including the 2030 Agenda and its Sustainable Development Goals. As such, efforts must be made by census offices to produce the relevant data to meet these commitments.

1.18. Increasingly, of population and housing censuses are a key source of data in the compilation of environment and sustainable development statistics. They provide environmental agencies with a range of information to analyze energy consumption, identify conservation opportunities, and forecast energy needs. Census data help national and local government and relief agencies in planning relief operations to assist populations and areas affected by natural disasters such as floods, hurricanes, tornadoes, and earthquakes. It also provides a rich source of data for public health authorities in identifying prevalence levels of disease, responses to pandemics and measurement of outcomes of public health initiatives. Moreover, census data is frequently integrated with other thematic data to derive essential indicators for comprehensive analyses.

1.19. A population and housing census also serves as the logical starting point for work on the organization and construction of statistical products to serve continuing national and local needs for data in the intercensal period.[7]

1.20. In addition to the statistical value obtained directly from the census results themselves, there are further, indirect benefits from taking a census, particularly to the organization responsible for the census, or the national statistical office. These benefits include:

(i) **Improved skills and experience:** varied sets of skills are often required for administering a census that are not necessarily prominent in other parts of the organization, such as project management, procurement and commercial, communication, human resources, and geospatial (GIS) and information technology (IT) skills.

(ii) **Technological advancement:** often a census requires new technology to support complex data collection and processing requirements. These developments may be reused for other exercises within the national statistical office or lead to new technological developments.

[7] See Chapter X, Part Three of this publication.

(iii) **New methods:** the development of methods for enumerating the whole population, or statistical methods (such as edit and imputation) developed for processing census results, can often be reused for other statistical exercises within the national statistical office such as stratification of the population, estimation and modelling population characteristics, or the development of indicators.

(iv) **Halo effect:** the extensive promotion of the census may also have a positive effect, the "halo effect", on other surveys, resulting in increases in response rates.[8]

A. Uses of population censuses

1. Uses for policymaking, planning and administrative purposes

1.21. The fundamental purpose of the population census is to provide the facts essential to national policymaking, planning and administration. Information on the size, distribution and characteristics of a country's population is essential for describing and assessing its economic, social and demographic circumstances and for developing sound policies and programmes aimed at fostering the welfare of a country and its population. The population census, by providing comparable basic statistics for a country as a whole and for each administrative unit, locality and small area therein, can make an important contribution to the overall planning process and the management of national affairs. Counts of the population overall, or of subgroups within the population, by geographic region are often used for the distribution of government funding, development programmes and services. Population censuses in many countries represent the very foundation of their national statistical systems, with census data providing important baseline data for policy development and planning, for managing and evaluating programme activities across a broad range of sectoral applications, and for monitoring overall development progress. An emerging use for census data is the assessment of good governance by civil society groups and measuring the achievement of Sustainable Development Goals. The performance of a democratically elected government in improving the welfare of its citizens can be monitored from one census to the other by ordinary citizens through the widespread and timely dissemination of census results.

1.22. Geo-referenced statistics, in particular at subnational and local levels, greatly increase the relevance of statistical information by providing the geographic context of the characteristics that censuses or surveys are portraying. This allows policymakers and researchers to more easily understand and analyse this geographic relationship, leading to the development of more targeted, locally relevant and actionable plans, such as access to physical infrastructure, urban-regional planning, and planning to deal with climate change, disasters and pandemics.

2. Uses for research purposes

1.23. In addition to serving specific governmental policy purposes, the population census provides indispensable data for the scientific analysis and appraisal of the composition, distribution and past and

[8] The halo effect is a cognitive bias in which an observer's overall impression of a person, company, brand or product influences the observer's feelings and thoughts about that entity's character or properties. It was named by psychologist Edward Thorndike in reference to a person being perceived as having a halo or aureole. Subsequent researchers have studied it in relation to attractiveness and its bearing on the judicial and educational systems. The halo effect is a specific type of confirmation bias, wherein positive feelings in one area cause ambiguous or neutral traits to be viewed positively. Edward Thorndike originally coined the term referring only to people; however, its use has been greatly expanded, especially in the area of brand marketing.

prospective growth of the population. The changing patterns of urban–rural concentration, the development of urbanized areas, the geographic distribution of the population according to such variables as occupation and education, the changes in the sex and age structure of the population, household composition, and the mortality, fertility and migration differentials for various population groups, as well as the economic and social characteristics of the population and the labour force, are questions of scientific interest that are of importance both to research and for solving practical problems of industrial and commercial growth and understanding population dynamics.

3. Uses for business, industry and labour

1.24. In addition to those uses given above, the census has many important uses for individuals and institutions in business, industry and labour. Reliable estimates of consumer demand for an ever-expanding variety of goods and services depend on accurate information on the size of the population in subnational areas and its distribution at least by sex and age, since these characteristics heavily influence the demand for housing, transport, furnishings, food, clothing, recreational facilities, medical supplies and so forth. Furthermore, the census can be used to generate statistics on the size and characteristics of the supply and demand of labour needed for the production and distribution of such commodities and services in conformity with International Labour Organization statistical standards.[9] Such statistics on the local availability of labour may be important in determining the location and organization of enterprises.

4. Uses for boundary delimitation

1.25. One of the basic administrative uses of census data is to support political and administrative mapping. Detailed information on the geographic distribution of the population is indispensable for this purpose. Certain aspects of the legal or administrative status of territorial divisions may also depend on the size and characteristics of their populations, for example, whether a previously rural area is now to be declared as urban.

1.26. A compelling use of census data is in the redrawing of electoral constituency boundaries in most countries. This is often enshrined in the country's constitution and provides a legal basis for census taking. The current distribution of a country's population is thereby used to assign the number of elected officials who will represent people in the country's legislature.

5. Use as a sampling frame for surveys

1.27. For countries without comprehensive population registers, population censuses can serve as a foundation for construction of a sampling frame for surveys during the intercensal years on many topics, such as the labour force, poverty, fertility, mortality, health, time-use, gender-based violence, disability and international migration.

1.28. Censuses provide an area frame that contains hierarchical geographical areas from the largest area (at the national level) to the smallest geographic division, usually called enumeration areas (EAs) and a list frame that contains the list of households located within each EA. For countries with appropriate resources, address-based sampling frames have been gaining popularity given their efficiency and quality. The addresses are updated regularly and important auxiliary variables are available on the frame to help

[9] See Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (October 2013).

improve sampling efficiency. For countries that do not have the resources or capacity to maintain a comprehensive list of addresses, a master sample frame is often used. With a master sample frame, the address list is updated only for selected enumeration areas.

1.29. It is important to give careful consideration to the construction of a census frame for subsequent use as a survey sample frame when the census is in the planning stage. The census frame should be complete in coverage. This means, for example, that care must be taken to ensure that the entire country is divided into enumeration areas without any omissions or overlaps, i.e. all land area belongs to one and only one enumeration area. The coverage of population within the enumeration areas should be complete, regardless whether they reside in private households or not.

1.30. Population and household counts for the enumeration areas, taken from the census, are also a highly useful ingredient for post-census sample survey design planning. This information is often used to establish measures of size for the selection of first- or second-stage sampling units, or to help in various stratification schemes. Characteristics of living quarters, the households and individuals collected in the census are also useful as stratification variables to support efficient sampling design for surveys. For example, these may include the degree of urbanization, age, sex, education attainment, labour force status, nationality/statelessness status, forced displacement status, type of living quarter, as well as other characteristics that may be used to support design of specialized surveys targeting rare and/or difficult-to-reach populations such as international migrants, persons with disabilities, indigenous population, etc.

1.31. Efforts should be made to ensure census records are geospatially enabled, i.e., geocoded to a specific location – this can facilitate selecting samples for household surveys as well as data integration, including the integration of census results with administrative data.

1.32. Many countries use telephone to follow up with respondents to address missing values and nonresponse.[10] For countries without a good telephone frame for phone surveys, phone numbers collected during censuses can be used for subsequent surveys. However, this requires the following of strict protocols and consent from respondents.

6. Use in humanitarian settings

1.33. In humanitarian settings, comprehensive and sound census outputs can be used to support the identification of crisis affected populations and specifically vulnerable groups.

1.34. Census data can be used in humanitarian settings in a number of ways, including:

- *To assess the needs of the population*: Census data can be used to identify the number and location of people who have been affected by a humanitarian crisis. This information can be used to plan and deliver relief and recovery efforts. For example, census data can be used to identify the potential number of people who may need food, water, shelter and medical assistance.
  To identify, protect and assist vulnerable groups: Census data can be used to identify population groups facing increased vulnerabilities. This information can be used to target assistance to those who need it most such as women, children, older persons and persons with disabilities. By

[10] United Nations Statistical Commission. 51st session, item 3(j), background document. Report on the Results of the UNSD Survey on 2020 round population and housing censuses. 2020. Available from: https://unstats.un.org/unsd/statcom/51st-session/documents/BG-Item3j-Survey-E.pdf

identifying vulnerable populations in areas like refugee camps, census data helps prevent violence, exploitation, and ensures their safety and well-being.

1.35. In addition to these general uses, census data can also be used to address specific humanitarian challenges. For example, census data can be used to:

- *Track the spread of disease*: Census data plays a vital role in tracking the spread of disease in several ways, including by helping to identify demographics more susceptible to certain diseases based on age, socioeconomic status, or housing density. Researchers can use census data to create simulations of how a disease might spread through a population, helping to predict potential outbreaks and allowing for targeted interventions before a disease becomes widespread. By comparing data from different censuses, public health officials can track changes in disease prevalence over time. This helps them understand the effectiveness of existing interventions and identify emerging threats.
- *Plan for long-term recovery*: Census data informs long-term recovery plans following a crisis. By identifying areas needing new schools or hospitals, it supports infrastructure rebuilding and economic development.

1.36. The use of census data is an essential tool for humanitarian organizations. By providing accurate and timely information on the population, census data can help humanitarian organizations to save lives and to build a better future for people affected by crisis.

7. Use for monitoring internationally agreed development goals and international comparisons

1.37. Census data is essential for monitoring international development goals. The Sustainable Development Goals (SDGs), adopted by all United Nations Member States in 2015, provide a blueprint for a better and more sustainable future for all. They address the global challenges we face, including poverty, inequality, climate change, environmental degradation, peace and justice.

1.38. Census data can be used to track progress towards the SDGs in a number of ways, including:

- Measuring progress towards SDG indicators: Census data can be used to measure progress towards many of the SDG indicators. For example, census data can be used to measure SDG 1 on poverty, SDG 3 on health, and SDG 4 on education.
- Identifying disparities: Census data can be used to identify disparities in progress towards the SDGs. For example, census data can be used to identify disparities in access to education and healthcare between different groups of people, such as women and men, rural and urban populations, and among ethnic groups, migrants, forcibly displaced or stateless populations.
- Monitoring the impact of interventions: Census data can be used to monitor the impact of interventions designed to achieve the SDGs. For example, census data can be used to monitor the impact of a program to reduce poverty or improve educational attainment.
- Geographic disaggregation and geo-statistical integration: Census is often the only source of information that allows analysis for small geographic areas and specific population groups. There are indicators that can only be generated with the integration of census information with geospatial information.

1.39. Census data is a valuable tool for monitoring progress towards the SDGs. By providing accurate and timely information on the population, census data can help to ensure that no one is left behind.

1.40. As seen previously, census results are essential to demonstrate the changing patterns of urban–rural concentration, the development and expansion of urbanized areas and the population distribution. This also applies to monitoring urbanization at a global level, allowing international comparison. The Degree of Urbanisation (DEGURBA)[11] combines population size and population density thresholds to classify the territory of a country on an urban-rural continuum and delineate cities, towns, suburbs, and rural areas.

1.41. Censuses are a key information source to harmonise the definition of urban and rural areas for international statistical comparison (Degree of Urbanisation) as several indicators require urban-rural disaggregation or are sensitive to how urban and rural areas are defined.

8. Use for population estimates and projections

1.42. Census data plays a vital role in estimating and projecting future population trends. These estimates and projections are crucial for analytical and planning purposes.

1.43. Intercensal and post-censal population estimates are calculated using a combination of data sources and statistical methods, which can vary by country. These estimates typically include: the base population from the most recent census; estimates of births and deaths, both internal and international migration; and adjustments for specific subgroups, such as age groups. Results from the census evaluation are crucial for improving the accuracy of intercensal and post-census estimates. This evaluation helps identify and correct any deficiencies in the data collection phase, particularly regarding the base population figures.

1.44. It is important to acknowledge that population estimates and projections are not exact. They rely on the accuracy of the baseline population (including by age and sex) as well as on assumptions about future trends, which can change and lead to inaccuracies. Nevertheless, they remain the best tools we have for understanding the dynamics of future populations. Regularly re-basing these estimates with a population census helps minimize these inaccuracies.

9. Use in small area estimation

1.45. Small Area Estimation (SAE) offers a means of estimating indicators when critical data are not available at lower levels of geography (i.e. district or municipality levels) or for domains of study in which the sample sizes are too small. Small area estimation methods let two sources borrow strength from each other. In producing estimates on poverty, for example, a household survey that collects detailed information on household expenditure or consumption is required. Given the sample size constraint, direct estimates of poverty typically do not reach a user-specified level of precision for smaller geographical areas. Population censuses, on the other hand, have comprehensive coverage of a country's population. Small area estimation method brings these two sources together to produce poverty estimates for smaller geographical areas. Furthermore, small area estimators can provide predictions for domains where no sample information is available. Other data sources that have been used to integrate with household surveys include administrative data sources and geospatial data.

[11] The Degree of Urbanisation (DEGURBA)[11] was adopted by the 51st Statistical Commission of the United Nations in March 2020.

1.46. SAE has several advantages especially when dealing with development needs. SAE allows better use of existing data within a national statistical system to identify smaller geographic areas or specific population groups with the greatest development needs. With the ability to produce estimates at lower administrative levels or for population groups that are difficult to measure through household surveys, SAE also helps improve targeting of intervention and resource allocation.

10. Use in integrated population data systems

1.47. Population data systems generally refer to integrated frameworks for the collection, management, analysis, and dissemination of population-related data. They draw from diverse data sources, including but not limited to, censuses, surveys, administrative records, geospatial data, and alternative data sources. These systems play a pivotal role in guiding policy and decision-making processes, tracking population trends, highlighting disparities, assessing the effectiveness of programs, and fostering research and innovation. With rapid advancements in technology and methodology, as well as increasing demands for frequent, granular data, there is a need to envision the future of these systems.

1.48. Integrated population data systems rely heavily on the population and housing census as their foundational element. Population and housing censuses, typically conducted every 10 years, provide the granular data that forms the backbone of the system. The census captures population size, characteristics, and household information, offering a detailed snapshot at a specific point in time.

1.49. Census data is crucial for generating intercensal estimates, which bridge the gap between censuses and provide a more continuous picture of population dynamics. However, traditional census methods have limitations. The 10-year interval can leave significant gaps in data for users who require more frequent updates. Additionally, the accuracy of estimates can deteriorate over time, especially in areas with rapid population shifts.

1.50. While these limitations exist, efforts should be made to make census data more interoperable by design. This means designing and collecting census data with the specific goal of integrating it smoothly with data from other sources. Some key considerations for achieving this include:
i) Thoroughly documenting the coverage and data collection methodologies used in the census makes this information readily available to data users. This transparency fosters trust and facilitates accurate interpretation of the data.
ii) Ensuring that variables collected in both the census and other surveys use consistent concepts, definitions, and classifications is crucial. This allows for seamless comparisons and avoids misinterpretations when combining data from different sources.
iii) Assigning geographic codes (geocoding) to census data allows for spatial analysis and integration with other geospatial datasets. This enables us to visualize population trends across geographic areas and identify local variations.
iv) Collecting in the census information on unique ID numbers of people for countries that utilize them.

1.51. Implementing these strategies can help unlock the full potential of the population and housing census within integrated population data systems. A more interoperable census becomes a stronger foundation, allowing us to layer and analyze data from diverse sources to create a richer and more dynamic understanding of population trends.

B. Uses of housing censuses

1. Uses for development of benchmark housing statistics

1.52. The housing census produces benchmark statistics on the current housing situation and is vital for developing national housing and human settlements programmes and policies. The housing census is also valuable for providing the sampling frame for special housing and related surveys during the intercensal years.

1.53. Housing benchmark statistics are also critical for emergency planning for response to natural hazards (such as destructive storms, earthquakes, tsunami and fires), or post-conflict situations. Following such situations, these statistics can be used to estimate the numbers of people and structures affected, the need for emergency response and reconstruction requirements.

1.54. National statistical authorities would need to develop, from housing censuses, the sort of benchmark statistics in housing that could be supplemented by current building and construction statistics and would provide continuous up-to-date information of the housing position needed for the consideration of housing programmes. Since not all the basic information required to assess housing needs or to formulate housing programmes can be obtained through a housing census, additional requisite data must be obtained through the population census, special housing surveys and environmental surveys, and from vital statistics, economic statistics and so forth; but data obtained from the housing census will constitute the basic framework within which the estimates are made, indices computed and further statistical enquiries planned.

1.55. When population and housing censuses are carried out as a single operation or independently but in a well-coordinated fashion, the combined information provided is of much higher value than when conducted in isolation, since the essential features of both censuses are interrelated. The information on housing may be analyzed in association with the spatial, demographic and socioeconomic conditions of the occupants and vice versa.

2. Uses for the formulation of housing policy and programmes

1.56. Housing census data plays a critical role in formulating housing policies and programs. Objective data on the housing situation provides a foundation for policymakers to make informed decisions, considering social, economic, and political factors.

1.57. Additionally, this data helps identify the housing needs of various population groups (number of households, family size, housing conditions). By understanding these demographics, policymakers can design targeted housing programs at local and national levels. Furthermore, census data provides a baseline to monitor the impact of these programs over time, enabling continuous improvement with geographically relevant and comparable data.

1.58. In most countries, housing programmes encompass both governmental and private activity. The data derived from a housing census are used by governmental authorities for making an analysis or diagnosis of the housing situation. Housing conditions are analysed in quantitative and qualitative terms and data from previous censuses are used to indicate the changes in the housing situation that have occurred during the intercensal periods; the housing stock and future housing requirements are estimated

and compared with the rates of dwelling production being attained; and the characteristics of the households in need of housing are considered in relation to the availability and cost of housing. As part of overall development plans, such an analysis is necessary for the formulation of national housing programmes and for their execution.

1.59. Commercial users also study housing census data. Those engaged in the construction industry, financing institutions, and manufacturers of housing fixtures and equipment and household appliances use housing census data, in combination with population census data, to assess the possible demand for these goods and services.

3. Assessment of the quality of housing

1.60. The materials used for the construction of housing units (roof, walls, floors) are a significant indication of the quality of life experienced by the occupants and, viewed in aggregate across multiple dwellings in an area and alongside other information such as weather conditions and date of construction, can provide an indicator of disaster risk. Trends indicated by census data in the type of housing materials used can show improvements in the welfare of the citizenry, such as where the percentage of poor-quality or slum-like housing facilities has decreased, or differences in housing conditions between women-headed households and men-headed households. Additionally, measures of water, sanitation and hygiene (WASH), as well as energy usage and telecommunication data are valuable for the planning and development of future infrastructure needs.

C. Relationship between the population census and the housing census

1.61. There is an especially close association between population censuses and housing censuses. The two censuses may constitute one statistical operation or they may be two separate but well-coordinated activities, but in either case they should never be considered completely independently of each other because essential elements of each census are common to both. For example, an essential feature of a population census is determining the location of each occupied set of living quarters and of the persons living therein, and an essential feature of a housing census is the collection of information on the characteristics of each set of living quarters in association with the number and characteristics of its occupants.

1.62. In many countries, the population and housing censuses are taken concurrently, often with the use of a single schedule. In this way, the information on population and living quarters can be more readily matched and processed. This also makes it possible to relate the information on demographic and economic characteristics of each household member that is routinely collected in the population census to the housing census data, thereby facilitating a wide range of multivariate analyses.

1.63. In censuses that are conducted with field enumeration, the advantages of simultaneous investigation may be offset to some extent by the additional burden on the respondent and the enumerator resulting from the increased amount of information that must be collected at one time. In countries where this is likely to be a serious problem, consideration might be given to collecting data for a limited number of topics on the basis of a complete enumeration in the population and housing census, with more complex data in both fields being collected on a sample basis only, either concurrently with or immediately following the full enumeration. Alternatively, consideration might be given to carrying out the housing census as part of the advance household listing operations of the population census.

1.64. The relationship between the population census and the housing census will affect the means by which data on homeless persons are obtained. In the case of simultaneous censuses of population and housing, data on homeless persons will be obtained as part of the population census. Where the housing census is carried out independently of the population census, it may be necessary to assign resources to enumerate homeless persons in the housing census. Information collected from enumerating homeless persons may reflect, among other things, the magnitude of the housing problem in a given locality.

1.65. Beyond the operational side, population and housing censuses are also fundamental statistical exercises that offer invaluable insights into a country's demographics, living conditions, and socioeconomic status. These intrinsically linked operations inform each other, painting a deeper picture of the social, infrastructural, and economic landscape. By exploring these synergies, policymakers can make informed decisions that address the multifaceted needs of the population.

1.66. For instance, high population density revealed by the population census can be combined with housing data to understand the availability and quality of housing in that area. This combined insight allows for targeted infrastructure development, such as building affordable housing or improving public transportation, to better serve the local population.

1.67. Additionally, spatial analysis of combined population and housing data helps identify patterns, hotspots, and disparities in areas such as housing affordability, access to basic services, and urban-rural divides. This information guides resource allocation and targeted policies to address inequalities and improve overall living standards.

1.68. Joint assessments of population distribution, housing characteristics, and homelessness inform disaster risk management and emergency response planning. This combined data identifies vulnerable communities, allocates resources for mitigation, and shapes resilient housing strategies that protect lives and livelihoods.

D. Relationship of population and housing censuses to intercensal sample surveys

1.69. The rapidity of current changes in the size and other characteristics of populations, and the demand for additional detailed data on social and economic characteristics of population and housing characteristics that are not appropriate for collection in a full-scale census, have brought about the need for continuing programmes of intercensal household sampling surveys to collect current and detailed information on many topics.[12]

1.70. The population and housing census can provide the frame for scientific sample design in connection with such surveys (see paragraphs 1.27–1.32); at the same time, it provides benchmark data for evaluating the reasonableness of the overall survey results as well as a base against which changes in the characteristics investigated in both enquiries can be measured. To allow for the comparison of census and survey results, the definitions and classifications employed should be either identical or harmonized, while remaining consistent with the aims of each investigation. Because of the relative permanence of living

[12] *Designing Household Survey Samples: Practical Guidelines No. 98* (United Nations publication, Sales No. E.06.XVII.13).

quarters, the lists available from the housing census (with suitable updating) may also provide a convenient frame for carrying out enquiries dealing with topics other than population and housing.

E. Relationship of population and/or housing censuses to other types of censuses and other statistical investigations

1. Census of agriculture

1.71. While the population and housing censuses have a close relationship, their relationship with the agricultural census is often limited to the coordination of data collection in households involved in agricultural production. As the result of increasing integration within programmes of data collection, the relationship between the population and housing census and the agricultural census is now far closer than in the past and countries are increasingly looking at new ways to strengthen this relationship. This relationship is particularly relevant in countries where most agricultural production activities are in the household sector, which is the case in many developing countries.

1.72. One conceptual issue in relating the two censuses is that they use different units of enumeration. The unit of enumeration in the agricultural census is the agricultural holding,[13] which is the economic unit of agricultural production, while the units of enumeration in the population census are the household and the individuals within the household. The unit common to the two censuses is therefore the household engaged in own-account agricultural activities.[14] In many developing countries, there is usually a one-to-one relationship between households and agricultural holdings. In these cases the same unit is enumerated in both types of censuses.

1.73. Another conceptual issue is the reference periods used by the two censuses. The population census normally collects information about a person's main job or work activity during a short reference period (usually one week), while the census of agriculture collects information on work on agriculture during the year preceding the census because of the seasonality of many agricultural activities. In consequence, users may find some agricultural activity data from the agricultural census more comprehensive than from the population census because the latter may not identify persons connected with agricultural activity on a seasonal basis or as a secondary activity. On the other hand, the population census provides data on all persons working in agriculture, including as paid employees. Such information is not available from the agricultural census. Both censuses cover households with own-account agricultural production. To get a complete picture, agricultural data users will need both agricultural census data and population census data.

1.74. In planning the population and housing census, every opportunity for developing the relationship between this census and the agricultural census should be explored. This can take several forms. Definitions used in the population and housing censuses should be compatible with those used in the agricultural census so that meaningful comparisons can be made between the two data sets. The population and housing census can also be of use in the preparation of the agricultural census, such as in

[13] An agricultural holding is an economic unit of agricultural production under single management comprising all livestock kept and all land used wholly or partly for agricultural production purposes, without regard to title, legal form or size.

[14] That is, households with members engaged in agricultural activities as self-employed workers or as own-use producers.

the demarcation of enumeration areas, the preparation of the frame for the agricultural census or, if applicable, the sample design for agricultural surveys.

1.75. In planning the national census programme, serious consideration should be given to the possibility of collecting key agricultural information as part of the population and housing census exercise that would facilitate the preparation of an adequate frame of agricultural holdings in the household sector for a subsequent agricultural census. This integration can be achieved through pre-census cartographic work and/or listing exercise, or by adding two non-core agricultural topics to the census questionnaire (as detailed in paragraphs 5.453–5.461). Regardless of the chosen method (cartographic/listing or questionnaire), countries aiming to establish a frame for the agricultural census through the population and housing census should ensure synchronization. Conducting the agricultural census as soon as possible after the population and housing census leverages the most up-to-date and relevant frame data. This approach is detailed in the World Programme for the Census of Agriculture 2030[15] and *Guidelines for Linking Population and Housing Censuses with Agricultural Censuses*.[16]

1.76. Linking population and agricultural census data in countries where most agricultural production activities are in the household sector can bring many benefits. This could add considerable analytical value to data sets from both censuses and save on data collection costs. Many of the demographic and activity status data collected in the population census are also collected in the agricultural census. If data from the two censuses could be linked, it would no longer be necessary to collect these data again in the agricultural census, while still allowing for comprehensive cross-tabulations.

1.77. A few countries conduct the data collection for the population and agricultural censuses as a joint field operation. This is feasible only in countries with extensive methodological and field experience. Normally, each census retains its separate identity and uses its own questionnaire, but field operations are synchronized so that the two data collections can be done at the same time by the same enumerators. Occasionally, the two censuses are merged into one, particularly in countries where fieldwork costs are high and logistical challenges make costly conducting the two censuses separately. This may have a number of advantages; however, as this is an increasingly complex operation, its impact on field operations and data quality needs to be carefully considered. Errors or shortcomings in the design and execution would impact equally on both censuses.

2. Census of establishments

1.78. Although the collection of information on industrial and commercial establishments does not constitute a part of the population census, the information that is collected from employers and own-account workers regarding the economic units they operate, such as the kind of economic activity and the size of the unit, can be used for preparing listings of the proprietors of such establishments. The information can also be used to prepare listings of the establishments themselves, if information is requested on their location or when the establishments are located within the living quarters (or dwellings). Experience shows that these listings can be used in a subsequent census of establishments or for supplementing the registers of establishments maintained by most countries and utilized as a list-

[15] Food and Agriculture Organization of the United Nations, World Programme for the Census of Agriculture 2030, Statistical Development Series No. 20 (Rome, 2026).

[16] *Guidelines for Linking Population and Housing Censuses with Agricultural Censuses with Selected Country Practices*, Food and Agriculture Organization of the United Nations and the United Nations Population Fund, Rome, 2012.

based or area-based sampling frame for their establishment surveys. In these cases, it is important that the information is geocoded or geospatially enabled (coordinates and/or addresses) to allow integration with other databases, including administrative records.

1.79. Many business registers cover only establishments with fixed visible premises in which more than some minimum of persons (usually 5 or 10) are employed. In these cases, the population census can be used to collect basic information (such as kind of activity and size) on those establishments with employment below this minimum number of persons. However, special care should be taken in the choice of the unit of enumeration to ensure that there is no double counting of establishments.

1.80. When the information from the population census is to be used to construct a list-based sample frame, it is essential that the information from the population census be available and used shortly after the enumeration is carried out because this information can quickly become outdated. This requirement is less imperative when the information is to be used to construct an area-based sample frame.

1.81. The population census information needed for these purposes is the status in employment, in order to identify employers and own-account workers. For this subset of workers, information needs to be collected on the number of establishments operated, and for each of these, the kind of economic activity, the name and address of the establishment (if any), the number of workers engaged (including contributing family workers and employees) and whether the establishment is operated in partnership with other persons. If all of this information appears in the census questionnaire (or can be derived from other information collected), the number of small establishments can be extracted from the schedule or from the processing documents after the enumeration.

3. Census of buildings

1.82. In certain circumstances, it may be necessary, as part of the housing census operations, to enquire whether buildings (both residential and non-residential) are occupied. Thus, it may be convenient to record basic information for all buildings at the time of the housing census, even though detailed data may be collected only for those in which housing units or other sets of living quarters are located. The comprehensive list thus obtained sometimes provides the basis for a census of buildings, carried out concurrently with, or subsequent to, the housing census, or it may provide for the identification of special types of buildings significant for other enquiries, such as the census of establishments or the census of schools. In these cases, it is important that the information is geocoded or geospatially enabled (coordinates and/or addresses) to allow integration with other databases, including administrative records. If a listing of households is to be carried out before the actual enumeration, this would be most ideal for collecting basic information about buildings.

4. System of current housing statistics

1.83. Current housing statistics refer to latest data and trends in the housing market. They reflect the number of dwellings constructed and certain related information such as value, number of rooms, floor space, and so forth, as well as number of dwellings destroyed or demolished. These data are usually obtained from a system of data collection based on the administrative procedures required in connection with activities in the housing market. For example, construction statistics may be derived from permits issued for the construction of dwellings, from records of dwelling starts or completions, or from certificates of occupancy. Statistics on dwellings destroyed may be obtained from the records maintained for the levying of rates and the collection of taxes. Compiled monthly or quarterly, current housing

statistics reflect changes in the housing inventory and, although they may serve other purposes, they are also used to update the benchmark data obtained from housing censuses.

F. Relationship between the population census and/or housing census and administrative data sources

1. Civil registration and vital statistics

1.84. Population censuses provide a fundamental baseline for the Civil Registration and Vital Statistics (CRVS) systems. By comparing census records with civil registration records, gaps in registration can be identified. In addition, census data can:

a. help identify population groups facing increased vulnerabilities such as refugees, internally displaced persons, stateless persons, or marginalized communities, who may face challenges in accessing civil registration services;
b. be used to estimate mortality rates (e.g., age-specific death rates) and fertility rates, which are essential components of vital statistics.
c. be used as a benchmark for evaluating the performance of CRVS systems.

1.85. The United Nations Statistics Division (UNSD) in its *Handbook on Civil Registration and Vital statistics Systems* defines civil registration as "the continuous, permanent, compulsory, and universal recording of the occurrence and characteristics of vital events pertaining to the population, as provided through decree or regulation in accordance with the legal requirements in each country". A well-functioning CRVS system helps ensure that every person has a legal identity and all vital events are registered, facilitating access to public benefits, social protections, and human rights. It also produces vital statistics (statistics on births, deaths, marriages and divorces) and serves as an optimal source of data on the prevalence, distribution, and cause-specific mortality for evidence-based decision policy, planning and programme implementation.

1.86. Additionally, the 2030 Agenda for Sustainable Development highlights the importance of CRVS system, as numerous targets and indicators rely on data from this system. Complete, reliable, and timely data from CRVS systems can provide cause-specific mortality to monitor many targets and indicators listed in the SDGs. Furthermore, a complete CRVS system facilitates the calculation of birth rates, age-specific fertility rates, and mortality rates.

1.87. Population census data serve as denominators for the computation of vital rates, especially rates specific for characteristics normally investigated only at the time of the census. Conversely, census results, time-adjusted by vital and migration statistics, can provide estimates of the future size, distribution and other characteristics of the population of the total country and subnational areas. Furthermore, census data on fertility can provide a benchmark check on the reliability of current birth statistics, and vice versa. It is consequently desirable that procedures for the collection of population census data, vital statistics and migration statistics be closely coordinated with regard to coverage, concepts, definitions, classifications and tabulations.

1.88. Some countries have linked individual census returns for infants less than 1 year of age with birth registration reports for the year preceding the census date as a means of checking on the completeness of one or the other type of investigation. Linkage of death reports with census returns has been used to

compare the information on characteristics of the deceased as reported in the two sources.[17] While the many problems posed in the past by the one-to-one matching of two types of records have not been entirely solved, their severity has been mitigated by developments in computer technology. Before undertaking either of the procedures, however, countries should consider carefully the possible advantages of using household sample survey returns rather than census returns in the operation. Moreover, such operations have to be carried out in complete accord with national laws and policies governing the confidentiality of information obtained in the census if public confidence in the census is to be maintained.

1.89. In the establishment of a civil registration system, census results on the geographic distribution of the population can be useful for determining appropriate locations for registration offices.

2. Population register

1.90. The term "population register" was originally defined in 1969, in the publication entitled Methodology and Evaluation of Population Registers and Similar Systems (United Nations, 1969), as "an individualized data system, that is, a mechanism of continuous recording, and/or of coordinated linkage, of selected information pertaining to each member of the resident population of a country in such a way to provide the possibility of determining up-to-date information concerning the size and characteristics of that population at selected time intervals" (chap. I.A). According to the UN Handbook on Registers-Based Population and Housing Censuses (2022), a population register is a system that "records all residents in the country, together with their basic demographic information, and updates for births, deaths, changes of marital status, international migration (both into and out of the country), and internal migration (by recording changes of address within the country)." Thus, the population register is the product of a continuous process, in which notifications of certain events, which may have been recorded originally in different administrative systems, are automatically linked to it on a current basis. The method and sources of updating should cover all changes so that the characteristics of individuals in the register remain current. Because of the nature of a population register, its organization, as well as its operation, should have a legal basis.

1.91. Basic characteristics that may be included in a population register are date and place of birth, sex, date and place of death, date of arrival/departure, citizenship(s), marital status, and place of residence. Depending on the possibility of proper linking with other registers, much additional information may be added to the single record, such as language(s), ethnicity, educational attainment, parity, activity status and occupation. In order to be useful, any additional information must be kept up to date. If complete and up-to-date, population registers can produce data on both internal and international migration through the recording of changes of residence as well as the recording of international arrivals and departures.

1.92. There is no universally agreed-upon definition for a population register. Existing definitions vary in terms of how they describe the population being registered and the system used for registration. However, a population register may be taken to mean an administrative register primarily used in an administrative information system by public agencies for their own specific purposes such as establishing

[17] An elaboration of comparison between census and vital statistics is provided in *Principles and Recommendations for a Vital Statistics System, Revision 3*, United Nations publication, Sales Number E.13.XVII.10, United Nations, New York, 2013 (see paragraphs 595–597).

personal identification, voting, education and military service, social insurance and welfare, and for police and court reference.

1.93. A key requirement is that each unit in the population register should be always uniquely identifiable. This is best achieved by using a system of identification codes (keys). Identification is however also possible without such a code being used, if sufficient information is available on the units (for persons: name, address, date of birth, etc.).

3. Building/dwelling or address register

1.94. An address register is an up-to-date comprehensive list of all known physical addresses within a country. In order for it to be a key piece of statistical infrastructure that can be used for a census it should preferably contain key pieces of information such as address text, address use, a latitude and longitude (x- and y-coordinates), dwelling structure for residential addresses and non-private dwelling information (i.e. hotels and caravan parks).

1.95. An address register can enable the running of a e-census by providing a mail out frame which is a dataset of residential and non-private dwellings (such as hotels and caravan parks) and is used to support the mail out of census materials. The latitude and longitude and dwelling structure is particularly useful for helping to locate dwellings in remote areas.

1.96. An address register can also be used to facilitate a register-based or combined census by using common variables (i.e. address text) to link to other government data sources in order to add further value to the current information, for example, employment and income data. Utilities data could also be linked to an address register to help determine if a residential address is being actively lived in or whether it is only used at certain times across the year.

1.97. As there are certain fields (particularly in a register-based or combined census) that exist as inputs from an administrative register and as outputs of the census, it is important to ensure that the descriptions and concepts of these fields are harmonised to make it fit for purpose. For instance, non-private dwellings and dwelling structure both appear as an input from an address register but also as an output in a census, in order to ensure the outputs are accurate the definitions of the different non-private dwellings and dwelling structures need to be consistent across the address register and the census.

4. Business register

1.98. Statistical Business Registers (SBRs) provide the infrastructure to support the collection of high quality, reliable and consistent economic data and the production of economic statistics. SBRs are maintained by National Statistical Offices and used for statistical purposes.

1.99. To ensure effective and complete coverage, SBRs must be regularly maintained and updated. Administrative data sources such as legal or taxation business registers including unique identifiers are used to update less significant units with simpler structures. Large, complex and significant businesses are updated through profiling activity to transform these businesses from administrative data to statistical units. Other sources of updates may include economic censuses, survey feedback, big data, web scraping or data mining.

1.100. SBRs provide a coherent set of statistical units and classifications to collect and assemble data. SBRs include units represented in the economic units model, which describes the characteristics of businesses and the structural relationships between related businesses.

1.101. SBRs are used to create regular point-in-time frames (frozen or common frames) from which survey or sampling frames (subsets) for most economic surveys are drawn, ultimately feeding into the System of National Accounts. SBRs are also used to analyse and publish business demography statistics. Business demography includes events such as creation and cessation of units, survival information, and changes in business characteristics over time including type of activity, size measures (employment, turnover), location, institutional sector, employing versus non-employing units.

1.102. Business registers are valuable for population and housing censuses. They provide a solid foundation for sampling, enhance the collection of economic data, and contribute to a more accurate and comprehensive picture of the population. Business registers can play a vital role in strengthening censuses and could contribute towards:

i) **Enhanced economic data collection**: Information within business registers, such as industry classification codes, can be directly linked to corresponding economic activity questions in the census. This allows for efficient coding and analysis of economic characteristics like industry of employment or type of business.

ii) **Accurate place of work data**: Business register data, including business locations, can be used to improve the accuracy of "place of work" information collected during the census. This fosters a better understanding of commuting patterns and workforce distribution.

iii) **Improved sampling frame**: Business registers provide a comprehensive list of businesses operating within a specific geographic area. This data serves as a robust sampling frame for establishment censuses, ensuring a more representative selection of establishments to be surveyed.

5. Other administrative data sets

1.103. Other administrative data may include such as records of foreign citizens and foreign workers, people living in institutional places, enrolment and graduation records, social insurance, employment records, etc. The availability and use of these sources is very country specific.

1.104. There is an increasing availability of a range of administrative data sources (held either by governmental agencies or private sector) containing information on the persons or the households within a country. The utility of these data sets for statistical analysis may be limited, in some countries, by their lack of population coverage, data accuracy or range of characteristics, especially in the case of newly-developed sources and those designed for purposes in which completeness and accuracy are not key quality criteria. The linkage of these data sets with the census file, with its complete coverage of the persons and households, can provide the ability to create new insights and new statistical products to leverage more value from the census.

1.105. Administrative data may replace data obtained by field enumeration – for example, in some countries income data from the taxation or revenue department can replace the need to directly collect those data in the census. Administrative data can also extend census data – for example, census data can be linked with visa information or health information to extend the census data set into areas that may be too sensitive to collect on the census form, or with past education data to analyse longitudinally the impact of education on labour force outcomes. Administrative data can also replace missing data – for

example, health records have been used in certain countries to estimate the count and characteristics of usual residents who did not respond during the census enumeration period.

1.106. As described above in paragraph 1.88, linkage operations should be undertaken with caution, ensuring not only that all national laws are met but also that the trust of the public in the census and the statistical systems is maintained.

G. Integration of census data with geospatial information and other data sources

(i) Integration of census data with other data sources

1.107. Data integration is the combination of technical and business processes used to combine data from disparate sources, either at the unit level (for example, persons, dwellings, businesses) or at the aggregated level. Data sources may include censuses, sample surveys, administrative data, or new data sources including geospatial information or big data. It should be noted that if some data sources not controlled by the census office, legislative authority maybe need to conduct data integration.

1.108. Data integration has the potential to improve the cost-effectiveness and relevance compared to single data collection activity. It also has the potential to enhance the coverage, granularity, timeliness and richness of available data, allowing more complex questions to be analysed and more insights to be drawn than would be possible from a single source. Many integrated data assets are also longitudinal in nature, meaning they allow changes and patterns in population, economy and environment to be better understood over time. Integrating data can also serve to reduce respondent burden by reducing the number of questions that need to be asked in censuses or surveys, as shown in some countries when quality administrative data sources are used with proper consent from respondents. Censuses are often used as a spine[18] for linking population data from multiple sources.

1.109. To fully realise the potential of data integration, institutional and technical capacities must be in place. Institutional capacity mainly refers to a) the availability and accessibility of relevant data sources, b) the existence of a strong legal basis supporting the use of and access to data and metadata for statistical purposes and legislation ensuring privacy and confidentiality, c) appropriate mechanisms for collaboration with data holders (either from the public or private sector), d) human resources and ICT Infrastructure. At the technical level, expertise and skill are required for (a) dealing with the lack of interoperability, in terms of lack of unique identifiers, differences in concepts, classifications, coverage, data formats, reference periods, etc.; (b) handling data quality issues, such as missing data, erroneous values, and inconsistencies within and across data sources; as well as (c) linking records across data sources.

1.110. Facilitation the integration of census with other data sources requires careful planning and designing of the census data collection, so it is "interoperable" by design. This may include using standard concepts, definitions, and classification as well as collecting unique identifiers, like national IDs, in the census. Georeferencing censuses is another important element to ensure that census data can be integrated with other sources through geospatial information.

[18] A **population spine** is a concept related to a reference population that is used to link together and manage different data sets about individuals. It is often a term used in relation to register-based censuses, which makes it easier to track changes in the population over time.

1.111. Data integration at unit level is made possible with the use of key linking variables. In the case of people, this may be unique personal identification numbers (PIN), such as national IDs, social security numbers, etc. or a combination of identifying variables like name, sex, date of birth, addresses, etc., if unique identifiers are not available across the data sources.

1.112. Data integration at the aggregated level required is carried out through modelling, for two different sources borrowing strength from each other. One example is the small area estimation methods that have been traditionally used for poverty mapping for lower geographical areas but have now been extended to cover a wide variety of outcome indicators including employment and social welfare. The method is also being used by countries to produce estimates for population groups facing increased vulnerabilities such as persons with disabilities.

1.113. For countries that do not have access to good administrative data, population censuses have been an important source to be integrated with household surveys, for two reasons. First, population censuses tend to be more comprehensive in coverage compared to household surveys. Second, population censuses collect a number of variables that can be used as auxiliary variables for the data integration. On the other hand, household surveys collected a lot of more in-depth information with limitations on coverage. Combining the two sources, through methods like small are estimation, can produce data that are more granular and timelier.

1.114. Successful integration of census with other data sources hinges on careful planning and design that prioritizes interoperability. This might involve using consistent concepts and definitions for key socio-demographic variables across all data sources. Additionally, georeferencing censuses plays a crucial role in enabling the integration of census data with other data sources. This linkage, ideally done through consistent, unambiguous and persistent identifiers (PIDs), can also allow the dataset to use various additional geospatial information within the address register, geocode database and/or coordinates. In the absence of such a system, organisations may rely on other ways to reference location (e.g. geographical names), alternative sources (e.g. address registry from utility provider) or higher-level geography (e.g. large geographical area).

1.115. Successful census integration with other data sources hinges on careful planning and design that prioritizes interoperability. This might involve using consistent concepts and definitions for key socio-demographic variables across all data sources. Additionally, georeferencing censuses plays a crucial role in enabling the integration of census data with other data sources.

(ii) Integration of census data with geospatial information

1.116. Geospatial information (e.g. address, x- and y-coordinates, geographical name, or unique identifiers related to geospatial features) can play an important role in bringing together information from various domains by enabling integration of datasets from different sources using the common location information as a matching and universal key variable (e.g. integrating administrative data with survey data using address or postal code that exists in both datasets). Geospatially enabled statistics information can provide critical knowledge by the integration with other data produced by various stakeholders in the data ecosystem to understand multi-faceted issues that the society currently faces such as sustainable development, urbanization, and climate change. To support these efforts, the Global Statistical Geospatial

Framework (GSGF) [19] details some of the very basic conceptual phases needed to create a flexible, interoperable national geospatial structure, fundamentally integrated with the statistical data collected during enumeration.

1.117. The GSGF enables a range of data to be integrated from both statistical and geospatial communities and, through the application of its five principles and supporting key elements, permits the production of harmonised and standardised geospatially enabled statistical data. The resulting data can then be integrated with statistical, geospatial, and other information to inform and facilitate data-driven and evidence-based decision making to support local, sub-national, national, regional, and global development priorities and agendas, such as the censuses.

1.118. The five principles of the GSGF outline the broad processes by which a range of geospatial and statistical infrastructures and processes are applied to input data to enable integration. Firstly, the statistical data are geospatially enabled to the finest level possible. Then, geospatial tools and methods, such as common geographies and common standards of good practice, are used to ensure the data are interoperable, accessible, and usable. The five principles are:

1. Use of fundamental geospatial infrastructure and geocoding;
2. Geocoded unit record data in a data management environment;
3. Common geographies for the dissemination of statistics;
4. Statistical and geospatial interoperability; and
5. Accessible and usable geospatially enabled statistics.

1.119. In addition to the five principles, four key elements play a fundamental role in enabling data to be obtained from different sources and the GSGF Principles to be applied. These elements comprise:

a. Standards and Good Practice;
b. National Laws and Policies;
c. Technical Infrastructure; and
d. Institutional collaboration.

1.120. To ensure the quality of integration, standardising the geospatial information in the different datasets is critical. This standardisation should normally take place before the integration of datasets and can be done through, for example, matching location information in the datasets with a centralised standard system (e.g. address matching, coordinates, geocoding) which should be part of the National Spatial Data Infrastructure (NSDI).

1.121. The National Spatial Data Infrastructure (NSDI) is a framework of policies, institutional arrangements, technologies, data, and people that enables the sharing and effective usage of geographic information by standardizing formats and protocols for access and interoperability. The development of a NSDI is a multi-agency and multi-partner activity. All governmental authorities, private sector and public users can view NSDI data set and map layers based on user rules and privilege to download, print, geoprocessing services and dynamic queries.

1.122. The benefits of using NSDI:

- Providing a unified standard base map with different scales
- Reduce duplication of efforts among governmental authorities

[19] *The Global Statistical Geospatial Framework*, United Nations Statistical Division, New York, 2019.

- Increase the benefits of using available spatial data
- Lower costs related to geographic information while making geographic data more accessible
- Legal arrangements and governance
- Establish key partnerships between Gov. Authorities, states, cities, academic, and the private sector

1.123. The advantages of NSDI

- Gains in time for searching and accessing the necessary spatial data
- Access to the latest data and complete metadata
- Flexibility to increase
- Contributes to openness & transparency
- Improved and updated continuously
- NSDI web mapping governmental portal

1.124. Census is an opportunity to bring the national statistical office (NSO) and the national geospatial information agency (NGIA)[20] closer together in mapping initiatives and information/data exchange. This approach can optimize the use of resources both in the census operation and in the production of national mapping, also involving the contracting of services (satellite images, location collection equipment, software, etc.). It can be an input for the NSDI and for the NSO to integrate efforts to manage national geospatial information according to the Integrated Geospatial Information Framework (UN-IGIF).

1.125. The use of fundamental geospatial infrastructure and geocoding[21] should be used in censuses to obtain standardized, high-quality location references and cover precise coordinates or a small geographic area or standard grid reference for each population or household unit. Preferably, the location should be recorded through direct or indirect capture of x and y coordinates and/or address point. When this level of accuracy is not possible using a country's current geospatial and statistical infrastructure, adaptations based on more general locations or larger geographies will be necessary.

1.126. The census should produce geospatially enabled data on the number and characteristics of the population and housing units that allow aggregation to the lowest appropriate geographic level, compatible with national circumstances, and for small population groups [22], awhile protecting confidentiality of personal information on each individual. This will also allow the definition of user-defined flexible areas or user-centered geographies as long as confidentiality issues are observed. Not all census methodologies satisfy this criterion.

[20] National Geospatial Information Agency (NGIA) is used as an encompassing term to cover National Mapping, National Cartographic, National Geospatial Information Agencies and Authorities.

[21] Geocoding is generally defined as the process of geospatially enabling statistical unit records so that they can be used in geospatial analysis. More specifically, geocoding is the process of linking unreferenced location information (e.g. an address), that is associated with a statistical unit, to a geocode (i.e. a geospatially referenced object). Alternatively, the geocode can be directly incorporated into the statistical unit record. Geocodes are, preferably, fine scale geospatially referenced objects that are stored as a geometry data type, such as location coordinates (i.e. x, y, z coordinates), and/or small area geographies (e.g. mesh blocks, block faces or similar small building block geographies). Larger geographic units, such as enumeration geographies, can be used as geocodes where finer scale geospatial units are not available. The linkage of a geocode to a statistical unit record can occur through use of standard geographic coding systems, a Uniform Resource Identifier (URI) or through other computer-based linkage mechanisms.

[22] One example is the application 'Build a custom area profile' from the United Kingdom Office for National Statistics. The application a simple tool that allows you to create a Census 2021 data profile covering any area in England and Wales, including one that you draw for yourself on a map.

1.127. The resulting geospatially enabled statistics and integrated data will support informed and evidence-based decisions, development plans at all levels within a country and provide valuable information on the underpinning pillars of a country's society, economy, and environment, as well as informing regional and global action.

IV. Census methodology

1.128. There are three basic approaches to conducting a census based on different methods of data collection; these are: the *full-field enumeration* in which data is collected directly from the public in a field census; the use of data sources solely based on administrative records (commonly referred to as a *register-based census*); or combining data derived from administrative records with data collected from the field (the so-called *combined census* approach).

1.129. Taking account of the experiences of the previous rounds of population and housing censuses, it has become evident that an increasing number of countries have been exploring the use of alternative methodologies to the conventional, field-enumeration based approach for collecting census information. The use of registers – primarily population registers – in combination with other sources is being considered in a number of countries for the purpose of producing detailed small-area statistics on population and housing, as well as the application of continuous survey methodology for the same purpose. Furthermore, these alternatives to the traditional method of conducting population and housing censuses are becoming more diverse in terms of developing combinations of various data collection methods (see paragraph 1.97), and it is thus a challenge to summarize and categorize them using generally accepted data source methodologies. Many countries have developed combined systems, making use of the information available in the registers or from administrative systems to complement information collected through field operations (e.g. in order to ensure the exhaustivity of the count) or, vice versa, making use of survey data, in order to complement information available in registers. Despite this complexity, a three-folded classification (field enumeration approached, register-based approaches, combined approaches) appears as a useful instrument in order to describe the main features of the different approaches and thus provide recommendations according to the main issues and challenges.

1.130. It should be noted that many countries are expected to continue using the full-field enumeration census approach – collecting information from each household in a country – in the 2030 round of censuses,[23] while at the same time it is anticipated that increasing numbers of countries will intend to use alternative methodologies. There are a number of reasons for exploring alternative approaches, principally:
(a) the need to produce more frequent and timely statistics;
(b) budgetary limitations for census taking;
(c) reluctance of the population to participate in the census;
(d) increased technical capacities to utilize data sources; and,
(e) utilizing data available on more variables from multiple sources.

1.131. This section aims to briefly elaborate on possible methodologies for conducting censuses based on the recent experiences of countries. The section also describes the necessary conditions for using a specific methodology, its advantages and challenges, and its implications for the content and

[23] The 2030 round of censuses covers the decade 2025–2034.

administration of the census. It should be kept in mind that countries using a specific census methodology might have significant differences in implementation of the methodology, arising from differing national conditions and expectations. Regardless of the approach, the crucial principle of providing detailed statistics at the lowest geographic level (and small population groups) remains of paramount importance, reinforcing the importance of geocoded information.

1.132. The various census methodologies are represented in a matrix in Table 1, where the rows describe data collection through field enumeration and the columns represent use of administrative or population registers as census data sources. The matrix presents only those options that either have been used or are likely to be used by countries and does not present all possible combinations.

1.133. The different approaches are summarized in Table 1[24] and then described in sections A to C. First, the full-field enumeration approaches (section A) and the register-based census (section B) are presented; then the combined methodologies are described (section C). Alternative approaches have been adopted in different ways by different countries, depending on national contexts, preferences and practices and the availability of appropriate data sources.

[24] For more details see *Guidelines on the use of registers and administrative data for population and housing census,* UNECE, 2018; *Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices*, United Nations Economic Commission for Europe, ECE/CES/13, Geneva, 2011; and *Register-based statistics in the Nordic countries: Review of best practices with focus on population and social statistics*, United Nations publication, Sales Number E.07.II.E.11, United Nations, Geneva, 2007.

Table 1. Main features of census approaches

Methodology	Approach	Description
Field enumeration (see A)	Full field enumeration (traditional census) or full field enumeration with yearly updates of characteristics (see A.1-A.3)	Information on census topics concerning individuals and households is collected by census questionnaire directly from respondents using enumerators or other modes of data collection (e.g. telephone interview, mail out/mail back, Internet), or by applying a combination of different modes of data collection. In a traditional census, full field enumeration may include an in-built ad hoc survey by use of the long form. It is common that a short form is used together with a long form. Both forms are utilized during the same time frame of the census. While the long form estimates are not based on full coverage, they are regarded as census output. Through modelling, the full field enumeration with the continuous cumulative survey is used to generate yearly (or other interval) estimates of detailed characteristics for different geographic levels. With this approach, the sample can be cumulated over time to produce statistics at the lowest levels of geographic detail to provide more frequent and relevant data.
	Rolling census (see A.4)	Information on individuals and households is collected through a continuous cumulative survey covering the whole country over a period of time (generally years), rather than a particular day or short period of enumeration. The two main parameters of a rolling census are the length of the period of enumeration (which is linked to the frequency of updates required) and the sampling rate (which depends on the geographic levels required for dissemination purposes).
Register-based (see B)	Base register and existing sample surveys or integrated administrative sources and existing sample surveys (see B.6)	Information on individuals and households is collected from existing administrative sources, namely different kinds of registers, of which the following are of primary importance: individuals, households and dwellings. These are linked at the individual level with information from existing sample surveys. No field ad hoc data collection will take place. Existing sample surveys include intercensal sample surveys on different topics, such as the labour force survey and the living standards survey.
	Fully register based (see B.6)	Information on individuals and households is collected from existing administrative sources, namely different types of registers, of which the following are of primary importance: individuals, households and dwellings. These are linked at the individual level with information taken from other

	administrative or statistical sources, such as business, tax, education, employment and other relevant registers.
Combined methodology (see C)	Information on individuals and households is collected by combining data collected from one or more surveys or full field enumeration with administrative data sources or statistical registers. Data from administrative sources or registers are employed not only as a frame or to support field operations, but directly as a data source for some census information. In some cases, register data are used to prefill the questionnaires to be verified or corrected during data collection. Ad hoc sample surveys are used to provide information on census topics not available from administrative sources or to adjust data that are of poor quality in registers.

A. Full field enumeration

1.134. The full field enumeration or traditional approach comprises a complex operation of actively collecting information from individuals and households on a range of topics at a specified time, accompanied by the compilation, evaluation, analysis and dissemination of demographic, economic and social data pertaining to a country or a well-delimited part of the country. Members of the public respond to a census questionnaire, or interviewers are deployed to collect information from respondents. For interviewer-based censuses, enumerators assigned to different enumeration areas cover all households and persons in the enumeration area during a specified and usually short period of time in order to meet the requirements of universality and simultaneity.

1.135. Various methods can be used for collecting the data, including a mailed or dropped-off questionnaire, the telephone, the Internet, personal visit, or a combination of such methods, countries employing the so-called traditional approach may utilize very different collection approaches in doing so.

1.136. The full field enumeration census has merit in providing a snapshot of the entire population at a specified period and data for small geographic areas and small population groups. In that sense, the traditional census is perhaps unique in nature. This approach is particularly suitable for countries requiring population numbers by various social and economic characteristics simultaneously for all geographic levels to meet the needs of planning and the allocation of funds. The delimitation of electoral boundaries requires simultaneity, and for that reason also the traditional approach may be more appropriate. But at the same time, traditional censuses have been singled out as the most elaborate, complex and costly data collection activity that national statistical authorities undertake. In addition to costs, this complex task requires full awareness and agreement of the public to participate in it.

Necessary conditions

1.137. It is essential to have national legislation for conducting the population and housing census to ensure confidentiality, transparency and the cooperation of the population. A permanent central census organization, which may or may not be part of the statistical office, needs to exist in the country, which can be expanded during the time of the census. Since a traditional census requires substantial resources,

sufficient funding for a field operation covering the entire country and subsequent data processing needs to be ensured. Other conditions necessary for this approach are the support of and general acceptance by the public to participate in the enumeration, and trust in the integrity of the statistical office or census agency.

Some considerations for census taking and content

1.138. Very precise planning is required for every stage of census taking in the traditional approach, due to the sheer volume of work and overlapping time frames. The recruitment and training of a large number of census takers adds to cost and complexity. Involvement of administrative machinery at the central, provincial and local levels is essential for successful field operations.

1.139. Since the data obtained in the full field enumeration method are respondent- and enumerator-based, there is risk of error in canvassing the questions and in the quality of response. This, however, can be minimized through a well-considered design and testing of the questionnaire, effective training and effective publicity.

1.140. Data can be provided for every administrative level subject to privacy and confidentiality considerations, which may not always be possible with other methods if some parts of the data collection are based on sample surveys.

1.141. Essential features of a population and housing census are fully satisfied with the traditional census method.

Advantages and challenges

1.142. The two most significant advantages of a traditional census are comprehensiveness of coverage and simultaneity. Another major advantage is the flexibility in deciding the topics to be covered and design of the questionnaire. There is lesser need for complex data adjustment since processing of raw data provides all inputs. The census frame becomes the base for all subsequent sampling frames. Finally, the focused and time-bound nature of the field operation implies that the data collection is finished in a short period and does not require long-term or constant monitoring.

1.143. One of the biggest disadvantages of a full field enumeration census is its cost and administrative complexity. Another disadvantage is that it has a very long processing time. Also, since it can be conducted only after 5- or 10-year intervals, data tend to become quickly outdated. The burden on respondents can be considered a disadvantage, especially in countries where participation in the enumeration is declining. Finally, many countries experience increasing difficulties in enumerating specific population groups, such as persons with high mobility or with multiple residences, or who are difficult to reach for other reasons. Unexpected emergencies such as natural disasters and public health incidents could also significantly disrupt the traditional census.

1. Types of censuses conducted with full-field enumeration

1.144. There are two main types of the full field enumeration census: the census with a single questionnaire and the census with both short and long questionnaires. In the latter case, the short form contains only questions intended for universal coverage, while the long form is used to collect information

from only a sample of households and population. In addition to these two main types, there are also a number of hybrid census approaches that combine elements of both types of questionnaire.

1.145. Overall, the decision of whether to use a traditional census with both short and long questionnaires, a single-form census, or an alternative method depends on a number of factors, such as the budget, the resources available, and the specific needs of the country or region.

(i) Full field enumeration census with a single questionnaire

1.146. A full field enumeration census with a single questionnaire is a census in which all households receive the same questionnaire. This type of census is conducted using door-to-door enumeration or self-enumeration by mail or online or enumeration through a call center established at the census office.

Advantages and disadvantages

1.147. There are several advantages to using a full field enumeration census with a single questionnaire. First, it is a proven method that has been used successfully for many years. Second, the use of a single form can reduce the complexity of the census (it is relatively simple to administer and process). Third, it can help to ensure that all households are counted and weighted equally. However, there are also some disadvantages to using a traditional census with a single questionnaire. One is that it can be expensive and time-consuming. Furthermore, the single-form census may not allow to collect enough data to meet the needs of policymakers and researchers.

1.148. Full field enumeration censuses with single questionnaires are still the most common method of conducting a census today, but countries are increasingly looking for ways to improve the efficiency and effectiveness of such censuses. They are also exploring alternative sources of data, such as administrative records.

(ii) Full field enumeration census with both short and long questionnaires

1.149. A full field enumeration census with both short and long questionnaires is a census in which some households receive a short questionnaire, and other households receive a long questionnaire with more questions. The short questionnaire, meant for universal coverage, collects basic demographic information, such as name, age, sex, marital status and ethnicity. The long questionnaire, on the other hand, collects more detailed information on a variety of socioeconomic and housing topics, such as education, employment, income, and housing. Both forms are utilized during the same time frame of the census. While the data collected on the long form provide estimates that are not based on full coverage, they are, nevertheless, still regarded as census outputs (as the figures disseminated cover the whole population).

Advantages and disadvantages

1.150. The use of two questionnaires allows the census to collect a broad range of data without burdening all households with the long questionnaire. However, this type of census can be more complex and expensive to administer than a census with a single questionnaire. Some of the advantages of using a full field enumeration census with both short and long questionnaires are:

- collects a broad range of data;

- reduces the burden on households; and
- can be used to estimate population characteristics for small geographic areas with the expansion of the sample data (usually an aggregate of enumeration areas).

But there are some disadvantages as well:

- more complex and expensive to administer than a census with a single questionnaire;
- may be subject to sampling error;
- may be difficult to reach certain populations.

1.151. The use of a single form can reduce both the cost and complexity of the census, and can ensure that all households are counted equally. However, the single-form census may not collect enough data to meet the needs of policymakers and researchers. The use of two-form census, or using alternative methods, such as administrative records and sample surveys, could help to collect more detailed data. The decision of whether to use a traditional census with both short and long questionnaires, a single-form census, or an alternative method depends on a number of factors, such as the budget, the resources available, and the specific needs of the country or region.

(iii) Full field enumeration census with yearly updates of characteristics

1.152. A full field enumeration census with yearly updates of characteristics is a variation of the traditional census design combining a full enumeration with a yearly sample survey. This approach focuses on counting the population and collecting only basic data in the census year while detailed population and housing data are collected throughout the intercensal period. In a traditional census, typically conducted once every 10 years, all households are counted at a specific point in time. This means that the data collected from the traditional census can become outdated over time. In a continuous survey, a household sample is interviewed regularly, such as quarterly or annually. This means that the data collected from the annual continuous survey is more up-to-date than the data collected from the decennial enumeration. However, the annual continuous survey only collects data from a sample of the population, so it's coverage is not as complete as the traditional census. By combining the two methods, a traditional census with annual continuous survey can provide both timely and accurate data on the population. This approach generates yearly (or at other intervals) estimates of detailed characteristics of population and housing. This can be useful for planning government programs and services, allocating resources, and conducting research.

Advantages and disadvantages

1.153. The primary advantage of this approach is to provide more frequent and relevant data on population and housing than would be available when a census is conducted only once a decade. However, such a programme might be costly and technically difficult, as it requires a continuous programme of comprehensive planning, development and testing.

1.154. Some of the advantages of adopting this approach are:

- produces more frequent and up-to-date population estimates;
- can be used to estimate population characteristics for small geographic areas with the expansion of the sample data (usually an aggregate of enumeration areas); and
- reduces the burden on households.

But there are some disadvantages as well:

- more complex and expensive to administer than a traditional census alone;

- may be subject to sampling error; and
- may be difficult to capture information on certain populations.

(iv) Rolling census

1.155. The rolling census is a variation of the traditional census whereby the whole population of a country is covered over an extended period of time rather than on one particular day as is the case in other types of censuses. The rolling census involves conducting a survey among a sample of households each year, and then combining the data from that sample survey with data from previous years to produce annual population estimates.

1.156. In a rolling census, information is collected on individuals, households and dwellings by a continuous cumulative survey covering the whole country over a period of time (generally years), rather than a particular day or short period of enumeration.[25] The two main parameters of a rolling census are the length of the period of enumeration and the sampling rate (which depends on the geographic levels required for dissemination purposes). For example, it is possible to build a sample framework in order to produce national results with one annual survey, regional results by cumulating three annual surveys, and small-area results by cumulating data over five years. Annual surveys may be conducted over the full course of the year or in a particular month or other shorter time frame.

Necessary conditions

1.157. Implementation of such an approach requires highly complex sampling and modelling techniques; a high-quality sampling frame in order to allow sampling at very low levels of geography; and successful consultation to gain acceptance of the approach with major stakeholders, including national and local governments and the user community.

1.158. The necessary conditions partly depend on the complexity of the sample framework. If the sampling units are addresses, a master address file is to be built first. But if the sampling units are larger, for example municipalities, it is only necessary to have enough information to spread the municipalities over the different years. It will be necessary to explain to statistics users the impact of the rolling sample on the use and interpretation of data, as many users are more used to snapshot data rather than period data.

Advantages and disadvantages

1.159. The main advantage of the rolling census approach is the potential for higher frequency for updating data: a traditional census provides benchmarks every five or, more commonly, ten years. In contrast, the rolling census provides annual updates. The high peak costs and labour requirements of a traditional census are instead spread over a longer period. Furthermore, it is possible to improve the census process over time, and to test methodological refinements and new technologies as they emerge.

1.160. The rolling census offers the opportunity to possibility introducing new topics as and when data requirements change thus making the census more relevant. Adopting new technologies as they emerge, also allows for the deployment of permanent teams that can focus on the continuous

[25] In the 2010 round of censuses France was the only country applying this concept.

evaluation of data quality and the training of field staff. However, introducing a new topic in a rolling census only gives an unbiased picture after information about this new topic has been collected in all regions and that may take a few years.

1.161. The disadvantage is that the rolling census approach no longer provides a snapshot of the whole population, complicating comparisons between areas due to significantly different enumeration times. In addition, as the rolling census covers the whole country over an extended period of time, some respondents will move. Thus some people may be surveyed several times and others may not be surveyed at all. As a result, universality might not be ensured unless careful methodological adjustments are made.

Some considerations for census taking and content

1.162. It is better to initiate a rolling census just after a full traditional census, in order to exploit the recent census information to build the sample framework. As the operation is annual, the process must be very carefully prepared, since any delay can be problematic for the subsequent stages.

1.163. A rolling census is able to include all usual census topics. There is also the possibility of changing the questions more regularly than in a decennial cycle. This enables the census to be more reactive to changes in the needs of users, even if comparability over time should in principle be preserved. However, only if the questions are stable over a number of years can a rolling census produce statistics at the same level of detail. Depending on the census organization and procedures, it may be possible to add some thematic surveys if required

2. Innovations in conducting field-based censuses

1.164. There are a number of innovations that NSOs can pursue to improve the efficiency, accuracy, and accessibility of field-based censuses. Some of these innovations include:

a) **Technology**. Technology is playing an increasingly important role in conducting traditional censuses. For example, many countries are now using mobile devices to collect census and geographic coordinates data in the field. This can help to improve the efficiency and accuracy of data collection and the territorial coverage of the census operation. Technology can also facilitate the real time monitoring of field operations and response rates through digital field management systems. Some countries are experimenting with using machine learning to develop new methods for estimating the population of small geographic areas, and to improve the accuracy of annual population estimates.

b) **Administrative data**. Administrative data is providing useful information to improve the preparations for, and the collection of information in, a census. Some examples of this include:
   - Creation of address register
   - Inform on planning and collection operations (for example, determining the size and boundaries of enumeration areas)
   - Dwelling occupancy verification
   - Informing on areas likely to respond using paper forms
   - Informing on likely areas of hard-to-reach populations
   - Geocoding special locations (e.g., places of education)
   - Determining non-response follow up contacts

- Quality assurance of online responses

c) **Geospatial information.** Geospatial information must be considered an integral part of the census methodology and operations. Geospatially relevant methods, tools and activities – such as the Geographic Information System (GIS) – must be integrated into the regular production processes of NSOs, so that the design and production of geospatially enabled statistics can be conducted in a systematic and consistent manner. Some examples of this include:
   - Satellite images and aerial photography
   - Global Navigation Satellite Systems (GNSS)[26]
   - Geospatial intelligence (for efficient placement of field offices)
   - Georeferenced address registry
   - Geocoding buildings and households
   - Map of enumeration areas
   - Geodatabase
   - Digital census cartography
   - Geocoded administrative data (e.g., electricity agencies or distribution companies)
   - Real-time monitoring of fieldwork with geoportals and dashboards
   - Route planning and tracking enumerators
   - Geopackages for handheld devices
   - Geofencing[27]
   - Heatmaps with concentration of refusals and vacant households
   - Geographic population distribution (e.g., urban, rural, remote areas)
   - Spatially enabled topics/questions[28]
   - Map services and geoportals to disseminate census results
   - Degree of Urbanzation (DEGURBA)
   - Grids
   - Common Geographies and Functional Areas

d) **Data integration.** Countries are also increasingly integrating data from different sources to produce more complete and accurate census data. For example, some countries are using administrative records, such as birth and death records, to update their population data, while others are using satellite imagery to estimate the population of remote areas. Others are using a combination of administrative records and sample surveys and geospatial data to produce more frequent population statistics. (See the topic Integration of census data with geospatial information and other data sources).

[26] Global Navigation Satellite System - GNSS is the standard generic term for satellite navigation systems that provide autonomous geo-spatial positioning with global coverage. This term includes e.g. the GPS, GLONASS, Galileo, Beidou and other regional systems.

[27] Geofencing is a resource that uses technologies to create a virtual geographic perimeter, by obtaining data through Wi-Fi networks, GNSS, radio frequency identification (RFID) and other geolocation tools.

[28] Applicable for opening topics or questions for specific populations (e.g. native or indigenous population, informal settlement dwellers, quilombolas) in locations that are not defined in enumeration areas or where the population is dispersed. These can configure areas of operational interest (AOIs) that will define the application of the question through on-site geospatial verification via geolocation at the time of the interview.

e) **Using alternative population bases**: Generating population statistics using alternative population bases. Alternative population bases (such as service or daytime populations) may offer researchers and policymakers the most accurate and relevant information for specific analysis needs.

f) **Public engagement**: To ensure a successful and accurate census, innovative public engagement strategies are crucial. This includes utilizing interactive and accessible outreach through social media campaigns, user-friendly websites and mobile apps. Additionally, multilingual communication and engaging content like infographics and videos can further bridge the gap. Incentives for participation and gamification elements can also increase engagement. Transparency and building trust through clear data privacy assurances, open data platforms, and community feedback mechanisms are equally important. Promoting civic engagement with school and intergenerational engagement activities programs can ensure everyone understands the importance of being counted. Through these innovative approaches, census operations can foster a more engaged and informed public, leading to higher participation rates and ultimately, more accurate and complete data.

1.165. These are just a few examples of the many innovations that are being used to improve field-based censuses. By using these innovations, countries can produce more accurate, timely, and accessible census data. In addition to the above innovations, countries are also exploring the use of new technologies, such as artificial intelligence and machine learning, to produce population statistics. These technologies have the potential to revolutionize the way that population statistics are collected and produced.

B. Register-based censuses

1.166. The concept of producing census-like results based on registers was developed in the 2000 round of censuses, although it has been debated and tested to various degrees since the 1970s, and several countries succeeded in using this approach to generate census data in the 1990 round of censuses. The philosophy underlying this concept is to take advantage of existing administrative sources, namely different kinds of registers on individuals, households and dwellings. These registers are linked at the individual record level with information held on business, tax, education, employment and other relevant registers. While it is theoretically possible to link records on the basis of the name and other unique details of the individuals, the existence of a unique identification number for each individual, household and dwelling allows a much more effective and reliable linkage of records from different registers.

1.167. Administrative data sources can, depending on their content and quality, be used in all phases of census taking. In principle, where greater amounts of information can be obtained from administrative sources, the production of census-type statistics will be faster, cheaper and more complete. The most complete use of registers will be where all core, and in some countries non-core, census topics can be based on register information. It is possible to improve the quality of data collected through the administrative register by measuring/assessing the quality through a sample survey or surveys and also to collect information not available. The sample survey(s) may either use the register as a sampling frame, or else be completely independent of the register.

1.168. To provide specific demographic data and socio-economic characteristics, the register-based census relies on the collection of available administrative data and information from government and private entities, as well as other service providers. The adoption of a register-based approach will enable

the provision of systematic, wide-ranging and timely statistical products that support informed decision-making and anticipate countries' future.

1.169. The use of administrative data sources may reduce the flexibility of the census exercise in terms of the variables that are available and their definitions. It may be difficult to change the variables as these are defined to fulfill the various administrative purposes for which the registers are created. A significant potential risk to the success of the census exercise is that the administrative source will usually be beyond the control of the statistical authority. The influence of the statistical authority over the administrative source can often be very limited. The content and availability of the administrative source may change at relatively short notice and without reference to statistical needs. For example, a change in taxation legislation may mean that a key administrative register may no longer collect information needed for the census. This risk can be minimized by establishing close and regular communication and cooperation between the statistical authority and the owners of the administrative sources.

1.170. Most of the time, only persons legally present in the country would get covered through this approach. There is a risk of certain population groups being absent in administrative data sources. For example, irregular or undocumented immigrants, homeless persons, nomadic or floating populations, persons involved in illegal employment or activities, etc. These groups are not likely to be recorded on certain occasions in any such administrative registers. On the other hand, registers may include persons who are no longer living in the country (for example persons who emigrated but were not cancelled from the registers), or, indeed, no longer living at all (where, for example, their deaths have not yet been recorded).

1.171. Subject to the caveats mentioned in the paragraphs above, the feature of “individual enumeration” is satisfied in this approach as separate information is collected regarding the characteristics of each individual. As regards “universality within a defined territory”, this criterion is satisfied as the enumeration base is taken from a population register. With regard to “simultaneity”, the timing of the census extraction may require careful thought where register update cycles vary. With respect to “periodicity”, this approach allows extraction at desired frequencies, including “at least once in 10 years”, noting again the need to manage the updating cycles for the registers. Finally, in most cases, the requirement in terms of producing small-area statistics is largely met, as the information in the registers generally allows for such aggregates to be generated, if information in the register have sufficient information based on which geographic or geocoded/georeferenced at the desired level.

1.172. It is possible that countries with larger populations may have greater difficulties when making the transition to a register-based census, since the volume of information that must be analyzed and integrated is higher. On the other hand, it is also possible that larger countries, instead of having a single register at national level, have several registers at regional and local levels, which would increase the number of data sources to be processed and complicate the transition to a register-based census.

1. Necessary pre-conditions for the use of administrative registers for censuses

1.173. Before administrative data can be effectively utilized for censuses, several key conditions must be met. These prerequisites are crucial to ensure the successful integration of administrative data into the census process, ultimately leading to the cost-effective production of more accurate, efficient and timely population statistics. A well-established central population register, or systems of registers, must exist with the infrastructure to facilitate access to the data it houses. The data within the registers must be of high quality, ensuring its accuracy, reliability, and relevance for statistical analysis, as well as have

comprehensive geographic coverage, encompassing the entire population of the country. The registers must be updated continuously to ensure they reflect the most recent demographic changes. Furthermore, a consistent set of concepts and definitions must be utilized across all administrative registers to facilitate data integration and analysis. Having common identifiers across all registers also ensures an effective linkage mechanism.

1.174. Some of the essential pre-conditions that a country should have to be able to conduct a register-based census are:

a) Legal framework

1.175. If administrative data are used for census purposes, statistical authorities should have a clear legal mandate to collect administrative data for statistical purposes. Individual historical, cultural and political factors of each country can lead to highly diverse legal frameworks, but ideally such legislation should encompass:

Data access

1.176. A legal basis should enable the statistical authority to collect administrative data. The statistics act should grant the National Statistical Office (NSO) the authority to access relevant administrative data held by other public authorities for the purpose of producing official statistics. This access should be clearly defined in the legislation, outlining any limitations related to data confidentiality or duration of access.

1.177. While specifying data sources in primary legislation might be inflexible, the act could acknowledge the NSO's right to access all relevant administrative data, with the understanding that specific sources may evolve based on user needs and administrative changes. Additionally, the legislation should explore avenues for the NSO to collaborate with data providers. This might involve some influence over the creation, revision, or deletion of administrative data elements crucial for statistical purposes, although achieving this might require further exploration and potentially separate agreements.

Privacy, integrity and security

1.178. To ensure the safe and secure handling of data and foster public trust, a robust legal framework is essential. This framework should comprise several key acts, including a statistics act, a privacy act, and potentially a data act. These regulations should govern all aspects of data handling within the National Statistical Office (NSO) and its interactions with other entities.

1.179. The legislation should emphasize the NSO's legal obligation to protect the confidentiality of administrative data. A "one-way traffic" principle should be established, where data flows from data providers to the NSO for statistical purposes but generally not the other way around. Specific circumstances permitting exceptions to this principle could be outlined in the legislation.

1.180. Regulations should clearly define how data can be transferred, handled, delivered, and used within the NSO. Additionally, the legislation should specify how data can be shared with other departments, organizations, and users, ensuring such access is strictly for statistical purposes and adheres to robust privacy-preserving techniques.

Data use

1.181. The list of variables on which data is to be collected may vary from census to census. Thus, these should not be included in the 'framework' legislation; but can be stipulated in the secondary regulations relating to each particular census. Variables, with metadata, including identifiers of administrative data sources, should be listed completely and described clearly. Limitations to the data use (for example duration of use, deletion of microdata, recoding of national unique identifiers, etc.) should be described. Furthermore, it should be clearly defined that data compiled for statistical purposes will by no means be retransmitted to the data-providing organization or other governmental authorities.

Transparency

1.182. Transparency serves as a foundational principle of a register-based census, fostering public trust, accountability, accuracy, understanding, privacy protection, and international comparability. By promoting openness and clarity throughout the census process, transparency contributes to the credibility and legitimacy of census results and supports evidence-based.

b) Political support and cooperation

1.183. A joint effort towards register-based statistics production requires firm and explicit commitment at the highest possible political level, as well as close collaboration among relevant authorities. Cooperation between statistical and administrative authorities generates a mutual and deeper understanding of the primary purpose of the registers and the needs of the statistical authority. It is recommended that the collaboration between the administrative authorities and the statistical agency be formalized.

c) Confidentiality and public approval

1.184. In the context of a census, the most important principle for the population is the confidential use of individual information, as stated in the sixth Fundamental Principle of Official Statistics, which requires that the use by statistical authorities of individual data, whether they refer to natural or legal persons, be strictly confidential and used exclusively for statistical purposes.

1.185. The political decision concerning the use of administrative data in a census can be highly influenced by public approval or refusal. In the run-up to implementing a new or modified census methodology it is helpful to inform the public about the project. It can be expected that people will become increasingly sensitive towards the collection and analysis of personal data by governmental authorities. In addition to outlining the general benefits and risks of the use of administrative data, information to the public should focus on the confidentiality of personal microdata. Clear limits and rules regarding the use of administrative data provide a common understanding that individual data collected for statistical purposes will not be passed on to other governmental authorities.

1.186. User acceptance is crucial for leveraging administrative data sources in censuses. Transparency about data quality, limitations, and potential biases is essential to ensure users trust and utilize the resulting statistics effectively. Building trust in the use of administrative data requires clear

communication with stakeholders and data users. Addressing concerns about privacy and data security is paramount for achieving widespread user acceptance.

d) Administrative routines

1.187. The decision about the use of register data for statistical purposes largely depends on the nature of the register itself, including qualities such as integrity, reliability and lifespan. Normally, data that are important for the administrative agencies responsible for the registers are recorded with a high degree of accuracy. There should be confidence in the administrative authority's capacity to be a reliable partner and data supplier. This implies the presence of administrative routines and safeguards. Does the administrative authority have extensive experience with the collection of the data that may be used for statistical purposes? Is the administrative authority well organized and is it anticipated that the necessary data collection will continue into the future? Are there existing quality guidelines for the administrative authority that guarantee long-term data quality? These are some of the questions that need to be answered in terms of assessing the feasibility of exploiting administrative registers for statistical purposes in general, and for the purposes of generating census statistics in particular.

e) Identifiers

1.188. Regardless of the census methodology adopted, it is extremely important that a unique unit identifier is used in all the data sources. The use of a unique identifier is essential in order to link unit record information from different sources successfully. This primary key may already exist in the country – for example, a national personal identification number (PIN). Where it does not exist, or exists but with poor quality (for example, too many missing), it can be artificially created for statistical purposes. A statistical linkage key can be built from unchanging variables for persons, such as "family name at birth", "first name", "date of birth", "sex" and "place of birth". Care needs to be taken with alternative spellings, for example, incorrect or incomplete registry entries, transcription errors, and the varying transcription of foreign alphabets, names or place names. In any case, linkage between different registers can be made both by deterministic criteria (complete equality of one or more variables) or by probabilistic criteria (high similarity of one or more variables).

1.189. Unique identifiers assist in the detection (and correction as necessary) of identical statistical units (duplicates). Duplicate records most often arise when collecting data from more than one decentralized register, but are also possible within one centralized register. The problem with duplicate data entries is the risk of multiple counting of identical statistical units.

1.190. In the case of fully register-based censuses, information from different registers should be matched using good-quality identifiers. It is important to define how often information from different sources is updated and the reference date of the information stored in the different registers. When two or more data deliveries with the same content from the same administrative authority are planned, a linkage key will enable validation of data quality with regard to the statistical reference period.

1.191. Successful data linkage may be compromised by poor quality of the source data. Information stored or provided by data owners may have errors resulting in non-linkage or multi-linkage of records. In these situations, probabilistic approaches that choose the closest candidate, or the use of geographic information (starting from the lowest detail level and gradually increasing), can help to improve the linkage process.

2. Types of registers

1.192. Administrative registers are produced on the basis of administrative processes to collect information on units (persons and housing units) and variables that are defined by administrative rules and demands in a country. Although the content and process of registers may differ from one country to another, the types of registers are usually very similar. The following provides definitions of the main concepts used in the system of administrative registers.[29]

a) A *register* is defined as systematic collection of unit-level data organized in such a way that *updating* is possible. Updating is the processing of identifiable information with the purpose of establishing, updating, correcting or extending the register.

b) *Administrative registers* are registers primarily used in an administrative information system. This means that the registers are used in the production of goods and services in public or private institutions or companies, or that the information is a result of such production.

c) *Base registers* are kept as a basic resource for public administration. The function is to keep stock of the population and to maintain identification information. It may also involve fundamental geospatial data such as addresses, geocoded buildings, establishments and dwellings, functional areas, geographical names etc. *Statistical base registers* are based on the corresponding administrative registers.

d) *Specialized registers* focus on a specific theme or domain, unlike base registers which serve a broad foundational purpose. Specialized registers often receive core demographic data (like name, address, date of birth) from a base population register, but collect additional information relevant to their specific purpose. Examples include registers of vehicles, social security, education, business, health, etc.

e) *Statistical registers* are created by processing data from administrative registers for statistical purposes. A statistical register could be based on one or several administrative registers.

1.193. As noted above, administrative registers are maintained primarily for administrative purposes. Units and variables of administrative data are described according to administrative rules and demands. Before a register is used for census purposes, the suitability of its data in terms of definitions, concepts, content, reference date, accuracy and other criteria should be statistically tested by comparing them with previous census and survey results, and conducting quality and compatibility surveys. A pilot census may be used for this purpose.

1.194. All persons in the register that are included within the definition of the census's population base are enumerated. The timing of the census extraction may require careful thought where register update cycles vary. Registration delays and administrative delays in updating between regional and national databases can otherwise have a serious impact on the quality of the output.

(i) Base registers

[29] *Register-based statistics in the Nordic countries: Review of best practices with focus on population and social statistics*, United Nations publication, Sales Number E.07.II.E.11, United Nations, Geneva, 2007.

1.195. Base registers are those registers that hold the basic information relating to the stock of the entire set of 'population' units that are being counted in the census – typically persons, dwellings and buildings. Base registers serve as the foundation, and data from specialized registers is linked at the unit record level to create the statistical register. This linkage allows for the derivation of additional characteristics needed for the census. By combining the base register data with information from specialized registers, the requisite census attribute variables for the population and dwellings are created, ultimately enabling the production of census statistics.

1.196. The base registers that are most commonly used for the purpose of the census are:

(i) **Population register** – records all residents in the country, together with their basic demographic information (such as age, sex and marital status), and updates for births, deaths, changes of marital status, international migration (both into and out of the country), and internal migration (by recording changes of address within the country).
(ii) **Building and/or dwelling register** - usually held by land and property valuation agencies or by local or municipal authorities responsible for the development of housing policies and urban planning. They can provide information on the location of all places where people might live (by distinguishing between wholly residential, partially residential and non-residential buildings) as well as some of the core census housing topics such as ownership, type of housing (for example whether a detached house, terrace house, or apartment), whether or not the housing is rented, living area, number of rooms, condition concerning availability of electricity, bathroom, kitchen, and toilet, and year of construction. With respect to a register-based population census such registers are of fundamental importance in the harmonisation with the population register, in that every person that is allocated a dwelling number can be associated with a particular housing unit, and (in some cases) a distinct household. Ideally, it is important that these records are geocoded with coordinates and/or addresses to facilitate their integration with other statistical and geospatial data.
(iii) **Establishment (or business) register** - usually contains information about business establishments (including their industry, institutional sector, size of workforce, and location). Ideally, it is important that these records are geocoded with coordinates and/or addresses to facilitate their integration with other statistical and geospatial data.
(iv) **Address register** – records all addresses, by location, and identifies whether the address is used for accommodation, commercial (business) or office purposes. The address register may also include links to the building (such as an apartment building) and/or the dwelling. Ideally, addresses should have associated coordinates to facilitate checking for duplicate records and integration with other databases.

1.197. Without these base registers, it is very difficult to conduct a fully register-based census. However, even without a population register, it may still be possible to conduct a combined census, using data from other registers and information collected in a field enumeration (see paragraphs 1.231-1.255 below).

(ii) Specialized registers

1.198. In addition to the base registers described above, there will be other specialised (or supplementary/subject-matter) administrative registers available to the NSO that contain records relating to those characteristics for which census information is required and from which the requisite variables for core and non-core topics can be derived using the PIN for record linkage. Such registers may include the following:

a) Social security or pension registers;
b) Tax registers;
c) Employment, unemployment and jobseeker registers;
d) Education and student registers;
e) Health registers;
f) Border control; and
g) Other supplementary registers.

1.199. NSOs may also be able to access data from a range of other administrative sources either to provide topic-related census variables or to identify particular population sub-groups that are traditionally hard to count accurately (and for whom full information may thus not be recorded in the population register), such as the institutional population. For example (and this is by no means a complete list): electoral registers; registers of motor vehicles; registers of foreign nationals or of residence permits; lists of military service personnel; registers of persons living in institutional places; registers of persons with disabilities; registers of recipients of social programmes; registers of building permits; farm registers; and, registers held by public facility service providers (such as for piped water supply, electricity and/or piped gas, and sewage and waste disposal).

1.200. Some specialized registers may help to identify issues of coverage by, in particular, identifying 'signs of life' activities. The basis of this methodology is that over time, a person living in a country inevitably leaves certain traces or markers of administrative activity in the form of records in different registers. This means it is possible to verify the person's presence in the country, as well as connections between persons and their locations, on an annual basis. Such verification is based on signs of life, signs of partnership that are recorded in registers every year. For example: persons in the age 0-3 years old can be captured in health registers. School age population 4-17 years old can be captured in education enrolment registers. The remaining population may be captured in labour force and other registers.

3. Establishing a statistical population register

1.201. The statistical population register is a systematized and indexed collection of individual records for every resident (including nationals and foreign citizens) of the country. Where developed and functioning, it represents the backbone of the compilation of official statistics and represents a master population frame as well for designing and running statistical surveys. Connected on a regular basis and under the provisions of the law with administrative registers, starting with the administrative population register, it should be regularly updated with pertinent information. The legal framework for establishing, maintaining and exploiting the national statistical population register must ensure that it is used solely for compiling aggregate statistics and that it cannot be accessed for any other purposes either within or outside of the national statistical authorities except for approved research, if lawfully authorized and strictly controlled and restricted only to anonymised records.

1.202. Creating a Statistical Population Register (SPR) is a crucial step towards achieving the use of administrative data for census purposes. It should be noted that it is important to distinguish between the registration of an individual unit (person) in the population register and the inclusion of that record in the computation of internationally comparable statistics on the stock of population and its basic characteristics, such as age and sex. While the listing of those persons in the register will be in accordance with national legislation for the requisite administrative purpose, for statistical purposes only those persons meeting the requirements of being residents of the country should be included. For instance, persons who have left the country but are still holding a permanent residence permit

should be still listed in the population register, but they should be excluded from the census population of usual residents. Particular care should then be applied to properly classifying registered persons who have emigrated, especially in those countries where there is no incentive to declare the emigration to the authorities.

1.203. An SPR is usually generated from the existing administrative population register by the NSO. In countries that have a reliable population register (for both nationals and foreign citizens), a statistical population register is routinely generated and updated from the existing population register.

1.204. On the other hand, if there is no centralized population register, a statistical population register can be constructed by integrating local population registers and linking them with other existing registers.

1.205. Thus, establishing a national statistical population register may involve some investment in collecting records from different existing administrative registers and/or administrative data sources, harmonizing them, linking information from different registers and running complex editing procedures to ensure the consistency and quality of individual records. Once established, and under the protocols spelled out in legal provisions, the national statistical population register is updated from administrative registers on a regular basis.

4. Constructing integrated statistical registers

1.206. A statistical register contains the data integrated from different sources including surveys and administrative data. The statistical registers are usually housed in the NSOs. The statistical agency usually can transform the data from administrative sources into the statistical data by editing, correcting the incorrect entries, remove inconsistencies, removing duplicates and allocating new statistical identifications. Among the statistical registers, the population register is the backbone for the register based statistical system.

1.207. A statistical register typically plays the role of a data coordination tool, integrating data from several sources, both statistical and administrative. This may be done by linking records using common identifiers or by using matching techniques. It may sometimes be easier to use data from a single source, but in such cases it is often difficult to check the accuracy of that source. When several sources are used and integrated within a statistical register it is possible to have a much better view of the accuracy of the data. Unfortunately, the negative side of this is that it becomes necessary to have a strategy for dealing with conflicting data from different sources. However, if variables in statistical registers are stored with source codes and dates, automated algorithms can be used to prioritise sources and resolve most data conflicts.

1.208. As well as integrating data from different sources, a statistical register may also provide the possibility of deriving new (census-related) variables not otherwise available from any one single register (such as industry of occupation, level of overcrowding or a housing quality index.)

1.209. Traditionally, statistical registers have been used as sampling frames for surveys, but as has been already noted, a statistical population register is at the very heart of a register-based census. A population register may be established inside a statistical institute for statistical purposes. However, it is more common (and far simpler) for the statistical agency to transform the administrative register into a statistical register owned and managed by the NSO, if permitted by law and the NSO has the resources to do so. This transformation process is usually a relatively simple task and may include editing, correcting,

removal of duplicates, and translating the administrative ID into a statistical ID. The NSO may also then integrate data that the administrative register does not have with authenticated evidence of data from other sources (such as, for example, country of birth for people immigrating before the administrative register was established). The statistical population register should be updated regularly.

1.210. In the process of transforming administrative records into statistical records, NSOs must consider geocoding as a fundamental process in building integrated statistical records. Location, in many cases an identifier (coordinates and addresses) that will facilitate integration, should be incorporated into the regular production processes of statistical organizations so that the planning and production of geospatially enabled statistical records can be conducted in a systematic and consistent manner. Although official geospatial information is mostly produced by the NGIA, all geo-located data is a constituent component of the data ecosystem, which national statistical organizations often see themselves as custodians. In this sense, institutional integration and partnership must be fostered, taking advantage of the NSDI and global frameworks such as the GSGF and the UN-IGIF.

1.211. In addition to identifying the data sources to be used for any particular statistical operation, some of the key processes involved in the construction of a statistical register are:

- data quality management;
- linking the unit records from these sources;
- dealing with duplications;
- resolution of conflicting information relating to the same data items in different linked sources;
- updating; and
- editing and imputation.

5. Transitioning from field-based censuses to register-based approaches

1.212. Moving forward to a register-based census from the traditional census has, as already noted, certain advantages such as reduced cost in the long run, improvement in the quality of data and ability to produce more frequent information at lower geographic level. However, countries may consider their country context when deciding to benefit from available administrative data for census purposes and not all countries may wish or need to make transition to a register-based census. For example, in some countries, privacy concerns discourage the use of administrative data for census or for some countries the data on variables essential to policymaking (e.g., religion or ethnocultural characteristics) are not available in administrative data sources.

1.213. Country practices indicate that changing census methodology from a full field enumeration census to a fully register-based census usually takes several decades. There is no ideal method or duration for this transition and therefore each country should plan this process based on the initial assessment of the availability of administrative registers as well as assessment of preconditions that have to be met before changing the census methodology. It should be noted that direct transition from a full field enumeration census to fully register-based census is, generally, not realistic, and that this transition should be planned gradually, introducing more variables each time, providing that the data from registers have been proven to be of good quality.

1.214. If a country decides to move to a combined or register-based census, the transition is usually done in several stages, often by first adopting one form of a combined census approach (discussed in Section C below), such as registers with full-field enumeration or registers with ad-hoc surveys, before moving to a fully register-based census. In this case, dependency on administrative data sources can be gradually

increased and the cost of the change can then be spread over two or three census cycles. At the beginning of this process, key registers such as a population register might be incomplete or of insufficient quality to be used as source for the census. In such cases, it should be made clear from the beginning that a country should not attempt a move immediately and should continue conducting some form of field-based census. However, even when countries continue to do so, innovations making greater use of administrative data could help the NSO work more efficiently (as discussed in Section A2 above). It is beneficial if the relevant public authorities make administrative data sources available to the statistical institutes to produce proto-type register-based statistics. The government itself can help a great deal both by removing any legal barriers to data sharing and by subsidising the transition.

1.215. It should be emphasised that there is no unique way of moving to register-based approaches; in fact, actually national practices show that there are many different ways of designing this transition, especially for countries that are planning to use administrative data sources for producing census information for the first time. Therefore, these recommendations provide only a general framework for moving to a combined or register-based censuses. Four key phases of the transition include:

(i) **Establishing/maintaining necessary pre-conditions**: For a successful transition from a full field enumeration to a combined or register-based census, it is necessary to establish legal framework for being able to use individual data for statistical purposes. Institutional capacity and establishing effective and continuous communication and cooperation with all stakeholders are also very critical.

(ii) **Identifying administrative data sources and other sources**: This phase needs to be carefully planned and managed in collaboration with register-holders and users. NSO should build up a wide-ranging knowledge of the existing administrative data sources before making any decision on changing the census methodology.

(iii) **Transforming administrative data into statistical registers**. This phase includes a series of activities that are implemented for transforming administrative data into statistical data from which the census outputs can be produced. The design of this phase will differ from one country to another depending on the quality of existing population register and other administrative registers. Countries with good quality of population register can more easily construct statistical population register. However, countries that do not have reliable population register, may have to conduct a full-field enumeration or ad-hoc surveys in order to be able to construct the statistical population register.

(iv) **Quality management and assessment**: Quality management and assessment is an overarching process implemented by NSOs throughout whole phases of this transition. For a successful design of this phase, it is important to develop a system for continuously monitoring the quality of activities and assessing the quality of each phase so as to be able to understand the effects of the quality on the next phase and, finally, on the census outputs[30].

6. Types of register-based censuses

(i) Fully register-based census

1.216. A census is said to be fully register-based when information on all variables is obtained solely from different registers. This option allows the countries that have it implemented, to obtain information on

30 See UNECE Guidelines for assessing the quality of administrative sources for use in censuses, https://unece.org/statistics/publications/CensusAdminQuality

the different census variables as frequently as they deem appropriate. Due to the advantages that this alternative provides, many countries have the medium or long-term objective of carrying out a census using this methodology.

1.217. For a successful register-based census, it must be verified that each and every one of the census characteristics can be obtained from registers with sufficient quality to meet users' needs. Therefore, a detailed study must be made of the information provided by these sources and contrasted with other results to verify its reliability. On the other hand, its level of coverage - which should encompass the entire population - must also be checked. Additionally, the periodic updating of information held in the registers must be carefully assessed.

(ii) Census with registers and existing sample surveys

1.218. A special case of register-based census is when population and/or integrated administrative registers are combined with existing sample survey(s). In this approach, no field-based data collection occurs as part of the census exercise, rather existing survey data are utilized. Different data sources are integrated as part of a "virtual census"[31] process. The data for the census exercise are derived from many types of registers and surveys, covering different population groups throughout the country and its subregions. Compared to field-based census methods, this process is lower in cost and staff requirements, and uses more frequently updated data sources. More significantly, by combining data that are already available from other sources, the virtual census makes no additional respondent burden, increasing its public acceptability.

1.219. A weakness of the use of sample data within the census exercise is that for the variables derived from the sample surveys the sample size may not be sufficient to offer the geographically detailed outputs, or thematically disaggregated information with multiple variables, that are an essential feature of the census. Importantly though, the use of data from existing regular surveys will often mean that time series of data are available. By combining samples for several survey exercises, it may be possible to produce reliable estimates for small geographic areas (although there would be the problem that the reference date would not be homogeneous for all the individuals).

1.220. However, the use of existing surveys does present a number of problems compared to the use of ad hoc surveys. The timing, content, statistical definitions and sampling approaches used in an existing survey may not be appropriate to allow the data to be readily combined with data from the administrative source(s). For example, many major household surveys are not designed to cover persons living in institutional households (such as student accommodation, hospitals, prisons and military establishments), meaning that an additional source of information is needed for these persons.

7. Advantages and challenges

1.221. The primary advantages of a register-based approach are: reduced long-term costs resulting from there being no need to conduct an extensive field operation; the opportunity for greater frequency of data, potentially on an annual basis rather than the decennial or quinquennial cycle; reducing non-response; and eliminating respondent burden. However, it should be recognised that there is always a

[31] The term "virtual census" originated in the Netherlands, where it was first developed, and here it refers to that specific methodology.

considerable initial cost involved in setting up systems for moving from the full field enumeration to the registered-based census.

1.222. There are, however, a number of potential drawbacks and challenges with the use of administrative data sources that need to be taken into account. One limitation is that the scope of statistical topics, key definitions and, indeed, the population base depend on the information that can be compiled from the available registers. These, in turn, will be based on the underlying administrative purpose and procedures of the registers. This may impose restrictions with respect to the characteristics that are available for description and may also undermine comparability both in comparison to data collected through full-field enumeration and in international comparisons. When a registered data item is changed, new or updated information is not always registered immediately. In certain cases, new or updated information may not be registered at all. Where this occurs, the register information will not accurately reflect real circumstances. Also, the choice of new and emerging topics is limited in registered-based approach as compared to full-field enumeration which facilitates a lot more flexibility. This also applies to several non-core topics that can be difficult to derive using administrative sources only and to difficult-to-reach populations.

1.223. Furthermore, when using different data sources based on registers, not all of them have the same reference date: there are those that are updated continuously, weekly, monthly, etc. This can be a problem, because in theory, all the census variables should refer to the same reference date.

1.224. Depending on the amount of information available from different data sources and its degree of integration, data processing can be more complex with these census methods than with a field-based census, although good-quality results can, nevertheless, still be obtained.

1.225. The use of registers also imposes on the statistical agency a dependency on the authorities responsible for holding and maintaining the registers as well as on any changes in legislation and administrative policy and practices. The decision to use administrative sources in the statistical production process, therefore, requires close collaboration between the administrative authorities and the national statistical offices. During preparations for data delivery, all parties concerned must agree on date(s) of delivery and the content of the data. This implies a bilateral agreement at a high hierarchical level on a detailed data set description, scheduled delivery dates and the statistical reference period. Test data deliveries help to solve or minimize problems with the subsequent data processing by the national statistical authority. Validation techniques appropriate to administrative data should be applied, including checks on the plausibility, completeness and reference periods.

1.226. Register information may contain errors (for example, records showing people as being implausibly old, invalid occupations, information about migration that is not consistent with other data). Edit rules may be defined to highlight inconsistent or implausible information. If a data source based in registers has errors in a significant percentage of cases, then that data source should be discarded. Correction or imputation of records with errors can be attempted in different ways: first, if possible, using another data source (register) that also has information about that specific record and topic; or second, carrying out probabilistic imputation based on available information that is considered to be reliable.

1.227. Households are an important unit in every census, however, identifying households from available administrative data sources in most cases is challenging, especially when the “housekeeping” concept rather than the “household dwelling” concept is used. Household relationship information is not typically available in registers but could be derived by combining information from different registers. For instance,

to determine cohabiting couples, information such as whether two persons living at the same address have a child in common, if the persons are relatives or not, the mutual age difference, and the date of moving into the dwelling can be used. This method makes it possible to identify cohabiting couples of opposite sex reliably, but there can still be challenges in some cases to detect cohabiting couples without common children.

1.228. Some socio-economic variables such as religion and ethnicity are often difficult to gather by using administrative data sources. In many countries such information regarded as sensitive and not suitable to store in an administrative register.

1.229. Another challenge when conducting a register-based census can be the lack of data describing the situation before the register was established. In the census this is associated with topics like ever resided abroad and place of birth. There are also challenging to derive topics related to human behaviour and attitudes, examples being mode of transport to work/school and type of energy used for heating.

1.230. To conduct a full field enumeration census is a massive operation that often requires significant investment in technology. These investments also benefit the entire statistical institute. When a fully register-based census system is established, the census cannot serve as a 'tool of investment' to the same extent as in a census using full-field enumeration.

C. Combined census methodologies

1.231. The combined census is a methodological approach whereby some information is obtained from administrative sources such as population registers, while other information on selected census topics are collected through field operations, either by using census forms that are completed by all units, in a way similar to the traditional approach (full field enumeration), or by means of ad hoc sample surveys.

1.232. The combined census is a frequently used as a stepping stone strategy of NSOs aiming to move from the full field enumeration census to a fully register-based census. Compared to a register-based census, the combined census method is more expensive, and imposes more response burden on the public. On the other hand, it may help to increase the coverage and quality of statistical registers for variables collected in the field.

1.233. In recent years, it has been observed in a number of countries that:

(a) The quality of the administrative registers is relatively good (at least for certain key census variables);

(b) Administrative registers, especially base registers, are being extensively geocoded with coordinates and/or addresses.

(c) Information for some census topics is not available in the administrative registers or the quality is not sufficiently high;

(d) The population generally, and certain population groups (in particular, those population groups that are difficult to enumerate), are becoming more sensitive to the handling of personal information, and possibly more reluctant to cooperate with the statistical office or more difficult to enumerate due to their high mobility or other reasons.

(e) Administrative registers exist but the information on some topics cannot be shared with NSOs due to lack of coordination and/or legal constraints.

1.234. In these cases, a combined census that uses register(s) and questionnaire(s) could be an option. Essentially, the combined methodology makes use of registers holding data that is relevant to a census, complemented by surveys or a complete field enumeration. The use of survey and enumeration data is intended to:

(a) Improve the accuracy of the population counts;

(b) Provide information from population groups or regions where registration information may be more fragile or non-existent;

(c) Provide information for census variables that cannot be reliably based on administrative data;

(d) Check, update and improve the quality of census data derived from administrative sources;

(e) Add additional variables to the census;

(f) Provide a frame to link together unit record data from different sources.

(g) To pave a way forward towards a fully register-based census by analysing the accuracy of available multiple data sources. The combined census methodology can be considered as a pilot test for the fully register-based census. It provides the opportunity to assess the accuracy, completeness of the available administrative data sources and hence, if the conditions are right, to plan for a fully register-based approach in subsequent censuses.

1.235. In short, when a country is in a state of transition towards a register-based census, due to having a limited set of records with partial information or geographic coverage, the combined census may be an appropriate option.

1.236. Information on individuals, households and dwellings is collected by combining data from registers with data collected from one or more surveys. Data collection may be based on full field enumeration, an ad hoc sample and rolling survey methods. Data from registers are employed not only as a sampling frame or to support field operations, but also directly as the data source for some census information. In a case where registers are used along with total enumeration, data from registers may be prefilled in questionnaires, and respondents may be asked to check, update and confirm their details. Other questions relating to fields not available in the registers may also be canvassed during this exercise. In the case of new individuals, households or dwellings that do not feature in the register, all fields of information that are required for the register and the census are to be canvassed afresh. When registers are used along with sample surveys (ad hoc sample or rolling surveys), some census tables may be produced entirely from the information available on the register(s), while for other census tables, information from the survey(s), duly weighted to the population totals, could be used. Some census tables may also be generated from model-based approaches where survey data and register data are used jointly.

1.237. It is worth clarifying that the results obtained from the registers will not always be consistent with the results obtained from the sample surveys, so it would be desirable to use weighting so that the information produced is as coherent as possible. The surveys would also serve to evaluate the accuracy of the register counts.

1.238. The option of a register-based census with existing sample surveys can be adopted only if all necessary census information is available from the various administrative or survey sources, and it is possible to link the information from the different sources at the unit record level. During the process of integrating individual records, care should be taken to check the accuracy of the data and remove inconsistencies prior to the production of statistical outputs.

1.239. The data sources should include verified and accurate personal information (name, ID number, date of birth, sex, marital status**,** family structure). In an ideal situation, a "base" register can be envisaged, to include unified identity codes for both people and address components in order to link more efficiently the related register and survey data. The link between persons and their dwellings is equally important, giving the household unit. Other administrative data sources can include tax files, social security files, public records of unemployed and registers of educational qualifications (as noted in paragraph 1.164 above). It is preferable to have a centralized base register. If this is not available, regional registers will need to be consolidated. The data sharing protocols/procedures be in place among the holders of administrative registers and the NSOs. These procedures should be subject to the provisions prescribed in the legal framework and ensure the confidentiality of individual records. The census reference period should be aligned with the regular updating cycle of the registers that has to be used for producing census outputs.

1.240. The different data sources of a combined census may have overlaps in some variables and the information contained may be different in certain cases. If this happens, it should be prioritized which data source is used in that situation.

1. Types of combined census methodologies

1.241. There are many different approaches to the implementation of a combined census methodology. Two prominent approaches combine base registers (like population or dwelling registers) and/or integrated administrative sources with either full enumeration or ad hoc sample surveys.

(i) Base register[32] and/or integrated administrative sources[33] with full field enumeration

1.242. One approach can be to combine the full enumeration with a base register or registers.[34] The questionnaire used in the total enumeration may then contain fewer questions compared to a full field enumeration census questionnaire, but still covers the whole population of individuals, households and dwellings. Over time, countries may decide to adopt this model, increasing in successive censuses the use of integrated administrative registers and reducing the number of questions in the questionnaire in successive censuses. In any case, to carry out a census that is as efficient as possible, NSOs should consider the fact that the variables that can be obtained with sufficient quality from the administrative registers need then not be part of the full field enumeration census questionnaire. The registers can be used to prefill such information as name, address, family composition, education, occupation and dwelling characteristics on the questionnaire. These prefills can then be used to ask the respondents if the information is still valid or needs to be changed.

(ii) Base register and/or integrated administrative sources with ad hoc sample survey(s)

[32] Examples of base registers could be population register and/or dwelling register; and address register and/or business register.

[33] Examples of integrated administrative sources could be administrative or statistical registers of education and/or occupation.

[34] In Europe in the 2010 round of censuses, four countries applied this approach: Estonia, Latvia, Liechtenstein and Lithuania. For more details, see *Measuring population and housing: practices of UNECE countries in the 2010 round of censuses,* United Nations Economic Commission for Europe, United Nations, New York and Geneva, 2014, page 12.

1.243. Another model involves the use of an ad hoc sample survey instead of a full enumeration.[35] The backbone register could then be used as a sample frame and also to prefill some information such as name and address on the questionnaire. The sample can be sized and stratified in such a way that data are available for small groups and geographic areas.

1.244. This model requires far fewer enumerators than a full enumeration, so a more specific training operation with skilled and prepared professional interviewers can be carried out. The follow-up of the operation is also simpler. There is no need to obtain information from each member of the population, giving a clear reduction in response burden. Non-response can be corrected in the sample by the use of statistical techniques to ensure information is still representative of the total population. In any case, it is important to pay attention to ensuring that all traditionally difficult-to-count groups are also correctly represented in the survey responses.

1.245. By using a sample and a much smaller number of enumerators, a significant reduction in the cost of the census is possible, in particular when compared with the full field enumeration approach.

1.246. The larger the size of the survey, the greater the geographical and thematic detail of the information that can be produced, but the higher the costs of the operation will be. The extent to which this model can produce detailed statistics (in particular for municipalities or smaller geographic areas) will depend on the size of the sample survey. A larger sample should allow more geographically detailed statistics to be produced but this will correspondingly increase the financial cost of the exercise. Even with a large sample, the availability of detailed information is likely to be lower than if a comprehensive approach were taken – such as with a full field enumeration census or fully register-based exercise.

2. Advantages and challenges

1.247. The main advantage with using the full enumeration with a base register or registers model is that it will reduce the response burden for respondents and reduce the cost of the census (if we compare it with a full field enumeration census). The model also allows the preparation of small-area statistics as all variables are collected as total counts. However, this model will still involve a large data collection exercise with the use of enumerators. Mail out/mail back data collection may sometimes be used, but a significant proportion of respondents may still require enumerators' follow-up and assistance in completing the questionnaire. A well-equipped call center may be established at the census office to help respondents fill the questionnaire and also filling up missing data using this facility. This model will continue to necessitate great planning in terms of data collection as it requires the response of the entire population.

1.248. The register-based census combined with ad-hoc surveys(s) offers several advantages:

(a) It can be much cheaper than a full field enumeration census collecting all census items from the whole population;

(b) It will reduce the burden of enumerators and respondents;

(c) It will reduce non-response in case information is obtained from registers;

[35] In the 2010 round of censuses, countries that applied this approach included Germany, Israel, Poland, Spain, Switzerland and Turkey. For more details, see *Measuring population and housing: practices of UNECE countries in the 2010 round of censuses,* United Nations Economic Commission for Europe, United Nations, New York and Geneva, 2014, page 12.

(d) It should be possible to correct the survey data for differing levels of non-response in different population groups.

(e) It will provide greater flexibility on introducing questions on new and emerging topics.

1.249. Micro-integrated data might be expected to provide very reliable results, because they are based on a maximum amount of information. The coverage of subpopulations may be more reliable because when data are missing in one source, another source may be used. Another advantage of micro-integration is that there will be less reason for confusion among statistics users. For example, there will be one figure on each socioeconomic phenomenon, instead of several different figures depending on which sources have been used.

1.250. A major challenge is that it involves more work to produce the tables from the sample survey microdata, as weighting problems may arise. As the combined census may lack the high public profile and publicity of a full field enumeration census, there may be less interest in and use of the census results, as there is no longer a single census event to attract public attention.

1.251. Finally, this type of census cannot have the same level of thematic and geographic detail as a full field enumeration census, and this may require giving explanations to the different census users. Another disadvantage arises when the register data is combined with the pre-existing sample surveys. The timing of the existing survey can be different from the census reference period, also the objectives, target population and sampling approaches, which makes it difficult to combine the data with administrative registers in an efficient way.

1.252. Data validation, processing and dissemination may be more complicated, as this approach involves both total counts based on the register and sample data from surveys. In addition, as some variables are based only on sample data, it may be impossible to meet the level of statistical and geographic detail required in some tables. It is advisable to be very didactic with census users: explain to them in a transparent way how the census results are obtained and the limitations they may have. On the other hand, the possibilities of reducing cost and response burden provide a very strong reason to adopt this approach.

1.253. Some of the required variables will need to be constructed from different sources. The census results obtained may differ to some extent from those that would be obtained from a full enumeration covering all census topics. This may have a negative impact on the comparability of results between countries and over time. An advantage of registers is that, in effect, they offer complete coverage subject to the quality of the data contained (see also paragraph 1.93). It is preferable that statistical authorities make full use of the register data that are available.

1.254. For the combined census method (as with the full field enumeration approach), a number of different modes can be employed to collect information, including paper and/or online questionnaires, handheld devices and telephone interviews and emails. Electronic devices have important advantages that influence the quality of the information obtained: validation controls can be included in the different questions, time taken to respond to the questions is reduced and the analysis and dissemination of information is faster. Where data are collected via different routes (such as Internet collection in parallel with face-to-face interviews), controls are needed to avoid duplication of information. Employing different modes to collect information from different households may introduce some bias in the responses.

1.255. For the data collection component of combined census methodologies, it is important to store control information and indicators at the lowest geographic level available in the central database, covering issues related to the census operation, such as progress with the fieldwork, response rates and comparisons with information in registers. This control information – normally based on web reports and analysed daily by project managers or regional offices – can be used to detect problems that appear during the fieldwork and to plan necessary actions to overcome these problems. If data are analysed on a daily basis, close monitoring of the continuing field operation and data entry is possible.

V. Methods for estimating population for areas not enumerated in the census

1.256. Conducting a census can be disrupted by unforeseen circumstances like natural disasters or security risks, hindering field operations in specific areas. In such scenarios, various approaches can estimate population numbers and densities for these inaccessible regions. These methods, including administrative data sources and sample surveys, involve some degree of error. Therefore, accuracy requirements should be carefully considered when selecting an estimation method.

1.257. Beyond traditional methods, innovative techniques are emerging for population estimation in inaccessible areas. One such example is the use of mobile positioning data to estimate daytime population distribution. As technology advances, we can expect even more precise and efficient methods for handling these situations.

1.258. While complete coverage during a census is an ideal goal, achieving it can be challenging. Factors like low response rates in self-enumeration methods, limited administrative data coverage, or security issues during field enumeration can all contribute to incomplete data in specific areas. In such cases, estimation approaches can be employed to fill these data gaps, but it's crucial to acknowledge them as estimates and distinct from actual counts. The inherent error associated with estimation methods needs to be considered when making this choice.

1.259. Estimation methods are often used to address data gaps at smaller geographical levels like districts, municipalities, or villages. It is important to remember that these "small area estimates" are not replacements for actual data and may differ from reality. Therefore, accuracy requirements must be a key factor when selecting an appropriate estimation method.

A. Requirements for estimating population

1.260. In order to generate reasonably accurate population data for uncounted areas during a census, careful consideration is needed to select the appropriate estimation method. Understanding the acceptable level of error is crucial. The methodology used for estimation should be clearly documented and accessible to users. The chosen method(s) should rely on data sources that are readily available, reliable, and of sufficient quality. Balancing accuracy needs with resource constraints is important.

1.261. Some key requirements for using estimation methods for parts of a country not possible to enumerate during a census include (based on the specific requirements by method):

i) Sample surveys: A high-quality sampling frame, ideally derived from a previous census or administrative data, is essential for selecting a representative sample. The chosen sampling

method (e.g., random sampling, stratified sampling) should be appropriate for the target population and context. The survey questionnaire should be well-designed and address the key data needs for population estimation, and achieve a high response rate.

ii) Administrative data: The administrative data sources used should be accurate, complete, and up-to-date and should cover the target population in the inaccessible area as comprehensively as possible. The ability to link administrative data records to geographic locations is also essential for spatial analysis.

iii) Combination methods: The chosen data sources should be compatible and allow for meaningful integration, while robust statistical methods are required to combine data from different sources and produce reliable estimates.

iv) Geospatial information: A high-quality geographic framework, such as digital maps and boundaries, satellite imagery, and modeling techniques, is needed for spatial analysis to support population estimation, particularly in geographically remote areas.

B. Methodological approaches

1. Use of administrative data sources

1.262. Administrative data can be used to estimate the population not enumerated in a census in a number of ways, for example through data from birth records, death records, change of address, and school enrolment records. Migration is likely to account for a significant element of population change, therefore, the appropriate administrative sources, such as a register of foreigners, should be used to provide relevant data.

1.263. Administrative records can be used to estimate total number of housing units, and the proportion of housing units which are occupied, in areas where individual enumeration is not possible. Administrative records to consider for this approach may include electricity and water consumption, telecom services, property registrations and cadastral information.

1.264. Estimation of the total resident population in parts of the territory not covered by the main census can be achieved with the Housing-Unit Ratio Method[36] or one of its variants. This method requires separate estimates of the total number of housing units, the proportion of housing units which are occupied, and the average number of residents per occupied housing unit. Administrative data can be used to estimate the first two components, while estimates of the latter is likely to be sourced elsewhere. The average number of residents per occupied housing unit is the average household size when using the "household dwelling" concept, but it is not when using the "housekeeping" concept, to identify households.

1.265. It is important to note that administrative data of not sufficient coverage may not be a perfect substitute for a census. Administrative data sets may not be complete or accurate, and they may not cover all members of the population. However, administrative data can be used to improve the accuracy of population estimates, and to identify populations that are at risk of being undercounted in the census. Despite these challenges, administrative data can be a valuable tool for estimating the population not enumerated in a census. By carefully considering the strengths and weaknesses of administrative data, researchers can develop accurate and reliable estimates of the total population.

[36] Methods and Materials of Demography 2nd ed, Siegel and Swanson 2010, pp 550-551.

2. Use of sample surveys

1.266. When census enumeration cannot reach parts of a country due to natural disasters, security concerns, or other challenges, sample surveys can become a vital tool for estimating population and its characteristics in those inaccessible areas.

1.267. The sample survey approach entails carefully selecting a sample from the target population in the inaccessible area that can represent the larger population's demographics. This selection process ensures the sample is not biased towards a specific group, leading to more reliable estimates. The survey questions are designed to gather key information about the population, similar to the questions asked during a full field enumeration census. This data might be collected through phone interviews, online surveys, or even in-person interviews in accessible areas bordering the inaccessible region. Once data from the sample is collected, statistical techniques can be used to "blow up" the findings to represent the entire inaccessible population. This involves applying appropriate weights to the sample data based on its representativeness.

1.268. Sample surveys are significantly cheaper and faster to conduct than a full field enumeration census in inaccessible areas. Surveys can be tailored to gather specific information relevant to the challenges faced in the inaccessible area. Different sampling methods can be used depending on the available resources and context.

1.269. Estimates derived from samples inherently have a margin of error, meaning they might not perfectly reflect the true population characteristics. The smaller the sample size, the larger the potential error. The accuracy of the estimates hinges on the effectiveness of the sampling method in capturing the diversity of the inaccessible population. By acknowledging the limitations and employing robust sampling techniques, statisticians can generate reasonably accurate population data.

3. Modelled methods

1.270. Within the 2030 Agenda for Sustainable Development, the growing requirement for spatially disaggregated population data has triggered the exploration of new data sources at different geographical scales, especially in highly stressed countries and countries without a recent population census. Advances in the availability of detailed satellite imagery, geo-positioning tools for field surveys, statistical methods and computational power are enabling the development and application of modelled approaches that can estimate population distributions at fine spatial scales across the entire country, in cases where population and housing censuses cannot be conducted.

1.271. In contexts where a population and housing census cannot be fully executed in all locations of a given territory, for instance because of security concerns, a hybrid approach can be adopted to produce population estimates for small areas or for uniform and detailed grids. Such an approach relies on complete population counts within small, defined areas, through a partial census, a census-like population survey or specifically designed “micro-census surveys” undertaken where it is possible to conduct field work. Statistical models can then be used to link these micro-census data to spatial data with full coverage over the regions of interest to predict population numbers in locations where no survey was conducted. By aggregating these high-resolution predictions, population totals can also be produced for administrative units or for the national level if required.

1.272. The success of this approach relies on the identification of relevant geospatial covariates and the definition of functional relationships between geospatial datasets and population density to predict population numbers in un-surveyed locations. It also requires robust geo-referencing of the geographical areas where population data are collected, and a design that captures as best as possible the range of population densities, demographics and environmental dimensions that exist across the regions of interest. The covariates need to be strongly correlated to population density and available consistently across all areas where population estimation is required. Although access to high-quality and spatially comprehensive datasets has traditionally been difficult in resource-poor settings, the increasing availability of very high-resolution satellite imagery makes the production of high-quality covariates increasingly feasible for many settings.

1.273. Although the production of such modelled population estimates can never replace the rich production of data on the individual, family, household or community generated by a full-scale population and housing census, their high level of geographical disaggregation makes them useful in contexts where no other population counts are available.

VI. Strategies for producing more frequent population and housing statistics

1.274. There are a number of alternative approaches that could be used for producing more frequent population statistics such as intercensal or annual surveys for countries conducting a census with the full field enumeration method, or use of other sample surveys; administrative data sources; geospatial information; or combined approaches, for example, administrative records combined with sample surveys. The best approach for producing more frequent population statistics will vary depending on the specific needs of the country or region. However, all of the approaches discussed above have the potential to improve the timeliness and accuracy of population statistics. In addition to the above approaches, countries are also increasingly exploring the use of new technologies, and alternative data sources, to produce population statistics. These sources can revolutionize the way population statistics are collected and produced, and make population statistics more frequent, accurate, and accessible to everyone.

A. The need for more frequent data on population and housing characteristics

1.275. There are a number of reasons why there is a need for more frequent data on population and housing characteristics. Some of these reasons include:

(i) **Planning government programs and services**: Timely and accurate data on population and housing characteristics allows governments to plan and deliver effective programs, such as new schools and hospitals based on population growth and distribution, or rental assistance based on housing affordability.

(ii) **Allocating resources**: Population and housing data guide fair and efficient resource allocation, like directing funds to low-income communities based on poverty rates or to schools serving those communities based on educational attainment.

(iii) **Conducting research**: Population and housing data are crucial for researchers studying economic trends, social inequality, and public health; for instance, income and employment data analyze policy impacts, while race and ethnicity data examine discrimination.

1.276. In addition to these general reasons, there are also a number of specific reasons why more frequent data on population and housing characteristics is needed. For example, the COVID-19 pandemic has highlighted the need for more timely data on population health and housing conditions. The pandemic has also shown how quickly population characteristics can change, making it more important than ever to collect data on a regular basis. More frequent data can be used to respond to unexpected events, such as natural disasters or economic crises. For example, data on population distribution can be used to target resources to areas that have been affected by a flooding or earthquakes.

1.277. More frequent data can also be used to identify emerging trends in population and housing characteristics. For example, data on housing affordability can be used to identify areas where housing prices are rising rapidly and people are at risk of displacement and homelessness.

1.278. Overall, there is a strong need for more frequent data on population and housing characteristics that are essential tool for making informed decisions about the future. This data is essential for planning and delivering government programmes and services, allocating resources fairly and efficiently, and conducting research on a wide range of topics. By investing in more frequent data collection, governments and researchers can ensure that they have the information they need to make informed decisions that benefit everyone.

B. Use of administrative registers

1.279. As mentioned in previous sections, administrative data sources, offer a promising avenue for producing more frequent population and housing statistics. However, there are preconditions that should be met. First and foremost is the availability of quality administrative data source. There should be mechanisms in place that guarantee long term access to and use of administrative data sources by the NSO for statistical purposes. The use of administrative data can be as limited as the use of birth, death and migration data to update the total population count at geographical levels. Or can be more extensive in case quality administrative data sources on other topics like education, employment, etc are also available.

C. Intercensal surveys

1.280. Intercensal surveys can be used to produce more frequent population and housing statistics by collecting data from a sample of the population between censuses.

1.281. Intercensal surveys have a number of advantages over full field enumeration censuses. First, they are less expensive and time-consuming to conduct. Second, they can be used to collect data on a wider range of topics than full field enumeration censuses. Third, they can be used to produce more timely data on the population.

1.282. However, intercensal surveys also have some disadvantages. First, they are subject to sampling error, which means that the estimates produced from the survey may not be as accurate as the estimates produced from a census. Second, intercensal surveys may not be able to reach all segments of the population, which can lead to bias in the results. Furthermore, these surveys may not be able to provide accurate estimates at lower geographic levels.

1.283. Despite these disadvantages, intercensal surveys can be a valuable tool for producing more frequent population and housing statistics. By using intercensal surveys, countries can produce more

timely and accurate data on the population, which can be used to inform decision-making on a range of issues. The intercensal survey combined with administrative data on births, deaths and migration can be used to update the overall picture (population snapshot) provided by the census, even if this picture has a lower resolution than that of the census.

D. Annual continuous surveys

1.284. The decennial census offers a valuable snapshot of a population at a specific point in time. However, socio-economic phenomena are constantly evolving, and relying solely on these snapshots can leave decision-makers with outdated information. Annual continuous surveys, on the other hand, can be used to produce more frequent population and housing statistics, thus providing a more dynamic picture.

1.285. Continuous annual surveys involve collecting data from a representative sample of the population on a year-round basis, continuously collecting data on demographics and other characteristics. By collecting data throughout the year, such as approach to data collection allow better tracking of trends, identification of emerging issues, and measurement of the impact of policies in a timelier manner.

1.286. Continuous annual surveys offer several advantages over full field enumeration census methods. First, they provide a significant enhancement in timeliness. By collecting data throughout the year, they reflect the most recent changes in the population, making them ideal for tracking rapidly evolving social and economic trends. Second, continuous surveys enable a more granular analysis of trends. Data collected over shorter time periods allows us to not only see the overall direction of change, but also the pace at which things are evolving. This can be crucial for identifying emerging issues early on. Finally, continuous surveys offer valuable insights for policy evaluation. Since they collect data in the years between full censuses, they can help us assess the effectiveness of recently implemented policies and guide any necessary adjustments.

1.287. However, continuous annual surveys also come with their own set of challenges. Maintaining a representative sample throughout the year can be more complex and potentially more expensive than conducting a single large-scale census survey. Additionally, continuous surveys might face lower response rates compared to high-profile decennial censuses. Survey fatigue and a perception of less importance can contribute to this challenge. Finally, ensuring data comparability across different years is crucial for valid trend analysis. Even minor changes in how questions are asked, or data is collected can make it difficult to accurately compare results over time.

1.288. Despite the challenges, continuous annual surveys offer a valuable approach to data collection, providing a more dynamic and up-to-date understanding of populations. By combining this method with traditional census data, we can gain a richer and more nuanced picture of the ever-changing world around us.

E. Use of mobile phone data

1.289. In recent years, mobile phone data has emerged as a potentially promising solution for generating more frequent and timely population statistics to fill the gaps in the inter-censal period. The advantages of these data are that often there is substantial coverage, potential for high accuracy, and minimal effort needed on the part of individuals to share the data.

1.290. Mobile phone data, particularly call detail records (CDRs) and location data, offer unique insights into population mobility and daytime population statistics. By analysing anonymized and aggregated data from mobile networks, national statistical offices can produce granular data on population movements, thereby identifying patterns in commuting, international temporary mobility, and other mobility trends. This data can be used for urban planning, emergency response, and transportation management, providing a dynamic picture of population distribution that static census data cannot offer.

1.291. The integration of mobile data with traditional data sources can improve the accuracy of resident population estimates. Mobile phone data can serve as a secondary data source in "sign of life" methods, which are used to validate and update register-based population data. These methods help identify active users within a geographic area, providing an additional layer of verification and enhancing the reliability of population registers.

1.292. Mobile phone data can also be used to validate data from other sources and collection exercises. For instance, during household surveys or other field operations, discrepancies between collected data and mobile phone data can highlight areas that require further investigation. This triangulation of data sources can enhance the overall quality and robustness of population statistics.

1.293. Despite its potential, the use of mobile phone data in official statistics is still in the experimental and innovative stages and comes with its own challenges. Access to mobile phone data is a significant barrier, often requiring negotiations with telecom operators and adherence to strict privacy regulations. Additionally, analysing mobile phone data requires specialized skills and technology that may not be readily available in many statistical offices. Investments in technological infrastructure and capacity building are essential to establish the necessary analytical capabilities.

1.294. Privacy and ethical considerations also play a critical role in the use of mobile phone data. Ensuring that data is anonymized and aggregated to protect individual privacy is paramount. National statistical offices must establish robust data governance frameworks to manage these concerns while maintaining public trust.

1.295. It is also important to keep in mind the limitations of mobile phone data with regard to representativity, particularly in regions with low mobile phone penetration or among demographic groups less likely to use mobile phones. Addressing these biases is crucial for ensuring the accuracy and reliability of the derived statistics.

PART TWO. PLANNING, ORGANIZATION AND MANAGEMENT

I. Introduction

2.1. Part Two of these *Principles and Recommendations* elaborates on planning, organization and management of the population and housing census. Regardless of whether countries adopt a full field enumeration census or alternative approaches involving the use of administrative registers, this is a unique and most complex statistical exercise requiring a multi-dimensional approach during preparation and management.

II. Census planning

A. Overall census planning

2.2. A population and housing census (or a population census by itself) consists of a complex series of interrelated activities, and constitutes perhaps the single most extensive, complex, and expensive operation that a country undertakes. Some of these activities, for example the printing of the census questionnaires, may be massive in scale, while other activities, for example the training of the supervisory staff, must be carried out in a uniform manner in all parts of the country. For register-based censuses or combined censuses that incorporate administrative data and field-based enumeration, there is also extensive work involved with the intake and integration of administrative data sources, particularly during periods of transition from full field enumeration census-taking to alternative methods involving administrative registers. Also, since censuses generally take place at five to ten yearly intervals, the planning and preparation for each new census round must account for changes in external conditions, census methodology and topics, technological innovations such as the introduction of web-based enumeration, stakeholders' requirements, personnel and societal conditions as well as changing administrative frameworks.

2.3. To ensure that the diverse operations occur in their proper sequence and in a timely manner, the entire census and its various component steps must be planned for carefully in advance. An apparently minor oversight in planning may lead to serious defects in the census results and to costly inefficiencies in the census operations. Careful planning is therefore critically important to a successful census, not only in countries with limited statistical experience but also in those with a well-developed system of statistics. Coupled with the need for careful planning is the need for appropriate organizational and administrative arrangements and procedures. Such arrangements and procedures are necessary to ensure both that the extensive human and material resources mobilized for the census are effectively and efficiently used and that its very tight time schedules and massive logistic requirements are met. For register-based censuses, which normally require fewer human resources, the management of the scaling-up of a census may not constitute a major risk, but other challenges must be overcome through appropriate administration, such as ensuring institutional memory of a still-complex set of processes.

2.4. It must be stressed, however, that at each stage of census planning and implementation, the various administrative arrangements developed will need to be guided by sound technical considerations. Ensuring the quality of census data in terms of accuracy and timeliness hinges on comprehensive consideration of a diverse range of subject matters and statistical requirements at every stage of the

operation. This is especially valid in the case of cross-cutting issues, such as information technology that are present throughout many essential phases of the census. It is for this reason that the management of any large statistical operation, but especially a population and housing census, cannot be considered a routine administrative assignment.[37]

2.5. Not all censuses follow a uniform pattern but there are certain major elements that must be taken into account in every one of them. In general, census operations can be divided into all eight of the phases within the Generalized Statistical Business Process Model (GSBPM)[38] as in Figure 2.1.:

(a) the specification of needs when consultations occur and objectives are set;

(b) the design phase, where outputs, collection and/or data intake as well as processing and production systems are designed and specified;

(c) the "build" stage when census systems and processes are developed and then tested;

(d) the collection phase when field operations occur and/or administrative data are received;

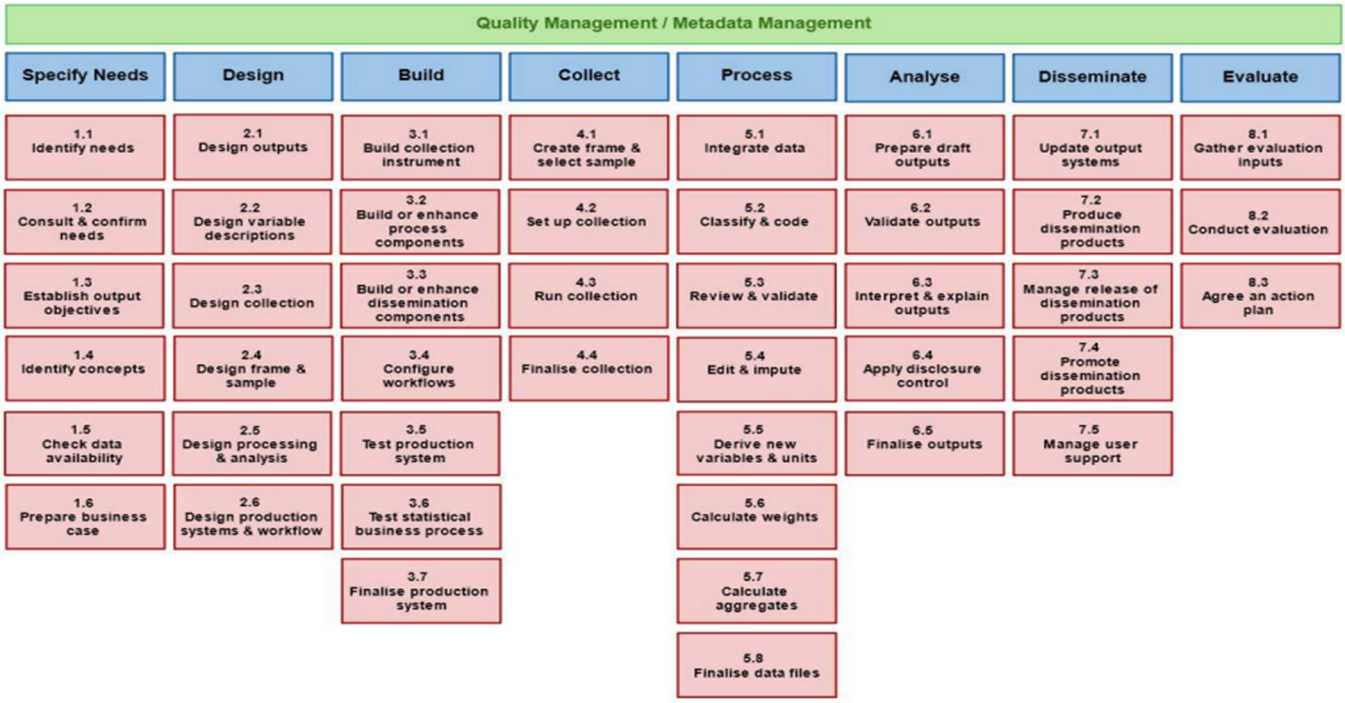
(e) the data processing stage when data are integrated, classified, edited and imputation occurs and data are weighted for the calculation of aggregate data files;

(f) the analysis stage when outputs are studied and validated as analytical outputs are produced and where disclosure control is applied;

(g) the dissemination phase when results are produced and made public and where user support is applied; and finally

(h) the evaluation phase when lessons learned are assembled and documented for action at the next census.

Figure 2.1. Generalized Statistical Business Process Model



[37] For a discussion of statistical management generally, see Handbook on Management and Organization of National Statistical Systems, Fourth Edition, https://unstats.un.org/capacity-development/handbook/Handbook_20230417.pdf

[38] United Nations Economic Commission for Europe (UNECE), on behalf of the International Statistical Community, *Generic Statistical Business Process Model (GSBPM)*, Version 5.1 (January 2019).

2.6. Throughout these phases, quality assurance measures must be applied to ensure the outputs that census data users expect. Also throughout, census-takers should be evaluating and recording the experience so that a thorough set of lessons learned can be discussed at the "evaluation" phase. Finally, these phases may not be sequential or mutually exclusive. For example, some census results might be released before all data-processing activities are completed; the analysis and the dissemination of census results overlap quite extensively; and the systematic evaluation of the census experience might be done in stages, between other GSBPM stages. Furthermore, certain elements that are discussed below, such as the initial budget and staff profile, may have to be amended according to the circumstances arising at a later stage of operations. The elements of each of these phases are discussed below in terms of their implications for sound census management.

B. Planning for both a population and housing census

2.7. A housing census may be conducted separately from a population census, or the two can be integrated into a single operation. A joint population and housing census will be more costly and complex than each census considered by itself, but less expensive than the total operation of carrying out both censuses independently, even if those independent censuses are managed to minimize extra expense. If conducted separately, there may be added expenses and challenges for training, resource mobilization, logistics, procurement, communication and data processing, among other operations. In addition to the financial and logistical advantages, amalgamating both a population and housing census will provide a greater wealth of data than if each census is carried out independently. In assessing whether to run population and housing censuses as one integrated program, or as separate censuses, each country will have to decide on the trade-offs involved in the light of its own needs and circumstances (see also paragraphs 1.61–1.68).

C. Special considerations when planning a combined or register-based census

2.8. As described in Part I, Chapter IV, some countries continue to use full field enumeration census methods, while others are shifting towards combined approaches (integrated administrative sources and field enumeration and/or sample surveying) or fully using base registers of individuals, households and dwellings. When a census is conducted to any degree using administrative registers, a comprehensive study should be conducted of the benefits and issues, and whether a move away from full field enumeration is feasible. Such a move is complex and can be high-risk so all aspects of the transition should be thoroughly studied, using assessment criteria established from international experience, as is provided through UNSD's assessment advice in the *Handbook on Registers-based Population and Housing Censuses*[39].

2.9. For any base register or administrative data source used by a census, it is advisable that the registers chosen for a census be those that were prepared as a legal requirement in the country. Countries that are planning for a register-based or combined census should assess their legal mandates to ensure they can use government and private data sets, some containing considerable demographic, social, cultural or statistical information, to enumerate the population and provide characteristic data.

[39] https://unstats.un.org/unsd/demographic-social/publication/handbook-registers-phc.pdf

2.10. Before a register is used for census purposes, the suitability of data in terms of definitions, concepts, content, coverage, reference date and other criteria should be ensured. Birth and death data are among the most important sources for the creation and maintenance of registers- based census and so these and all other data used for updating base data should be available in digital format. For linkage of records from different registers, the existence or creation of a unique identification number for each individual, household and/or dwelling will be required. In addition, modelling techniques may be required in order to integrate and select the best available information from the various administrative data sources. A well-defined methodology to link different registers and its maintenance must be prescribed, which should include a plan for the periodicity of updates. Transparency in these methodologies is crucial, and clear communication to users of the census data about the data sources, linkage processes, and any modelling techniques employed is essential to build trust and understanding.

2.11. The analysis of a country's readiness for a move to a combined or register-based census should include a thorough assessment of the public acceptability within the population of the large-scale use of their administrative data. Some portion of the population may be concerned with the privacy implications of using administrative data or registers to enumerate and characterize the population, so the national statistical office should ensure that it has the legal and technical capacity to use large scale administrative data in a way that guarantees the safety and security of the data it may have in its holdings.

III. Strategic objectives and management

A. Strategic objectives

2.12. The development of plans for a census should include the early preparation of a set of strategic aims and objectives that may be used to guide the implementation of the plans, set standards and form a set of benchmarks against which outcomes can be assessed to help determine the success of the census. Ideally, the starting point for developing these objectives would lie in combining information derived from evaluating previous census experience, from understanding stakeholders' requirements for information from the census, and from assessing changes in both society and technology. In practice, some of this information is difficult to obtain and often provides conflicting guidance. Nevertheless, such objectives can be used to assist in planning major elements of the process. Although the strategic objectives of the census will be specific to individual countries and will differ according to local circumstances and priorities, they can be described under the following headings: (a) census content; (b) impact on the public and census staff; (c) production of census results; and (d) cost-effectiveness. The cost, benefit and the ability to harness technology must be in line with the general situation of each country.

2.13. *Census content.* The aim is to ensure a relevant census, such that the topics are appropriate to meet the clear requirements of stakeholders (see para 2.153 for key stakeholders in the census process), while taking into account the considerations of cost-effectiveness, human resources, time availability and respondent burden. Census content should be the result of:

(a) appropriate consultation with current and potential stakeholders at all stages;

(b) the establishment of measurable standards of quality incorporating stakeholders views on priorities;

(c) adequate testing of new topics to ensure successful collection and production of reliable census products;

(d) ensuring that outputs are comparable internationally, nationally and regionally; and

(e) the suitability of the census as the source of the data, relative to the merits of using their vehicles. The census may form the backbone of a program of socio-economic statistics and content objectives for the census must consider similar data from other programmes and where content is best placed within the overall socio-economic statistics of a nation.

Furthermore, should any administrative-sourced data be conceptually different to that defined for census purposes, these conceptual limitations in the administrative data should be considered when designing any household surveys that are associated with the census or within the national statistical organization. It will also be important for the statistical organization to develop conceptual standards jointly between the institutions providing the administrative data. These standards should include agreed upon definitions, classifications and units of measure.

2.14. *Impact on the public and on census staff.* The aim is to ensure that all the aspects of the field enumeration operations, administrative data intake and use, as well as the dissemination of results are acceptable to the public and fully comply with legal and ethical standards for protecting the confidentiality of individual information. The public should be fully informed about census objectives, contents and methods, as well as about their rights and obligations with respect to the census. Similarly, all census staff must be fully aware of their responsibilities. Subsidiary objectives include such issues as:

(a) keeping completed forms, accessed administrative records and other records containing personal information secure and confidential;

(b) ensuring that public support for all aspects of the census is as strong as possible, whether field-based or register-based; and

(c) producing requested customized output in a manner consistent with preventing disclosure of personal information, adhering to established quality and privacy standards for the release of data, and implementing policies designed to safeguard the access to all census products by the various stakeholders.

2.15. *Production of census results.* The aim is to deliver census products and services and to meet legal obligations and stakeholder needs within stated quality standards and a predetermined and acceptable timetable. Subsidiary objectives include:

(a) producing timely outputs with minimal error so that the data are suitable for the purposes for which they are to be used;

(b) providing standard outputs for the main results and services for customized output;

(c) providing access to outputs;

(d) using geographic bases appropriate for collecting and referencing data for output;

(e) improving methods of field-based or register-based enumeration, particularly in difficult areas, so as to reduce levels of net undercoverage or net overcoverage and response error; this may involve a variety of data collection methods, such as in-person, self-response paper or web, to allow maximum flexibility to address and regional or other challenges in field-based census collection;

(f) improving methods of evaluation and the means to convey findings to users; and

(g) developing a measure of coverage, quality and targets.

2.16. *Cost-effectiveness*. The aim is to achieve the main strategic objectives while minimizing the expense of doing so. Subsidiary objectives relate to minimizing costs by:

(a) adopting more efficient data collection, data capture and data-processing approaches and related technological innovations;

(b) contracting out appropriate parts of the operation;

(c) exploring possible sources of alternative funding and, if appropriate, developing proposals for cost recovery and income generation;
(d) international and in-country collaboration and reuse of systems and other resources;
(e) encouraging the public to self-complete forms online or on paper where required;
(f) replacing direct collection of data with use of administrative data;
(g) integrating and maintaining linkages of different administrative data sources; and
(h) exploring shorter periodic production of census results (for register-based censuses).

2.17. *Cost–benefit.* The aim is to increase the value or benefit generated from the census while also managing the overall cost. A cost-benefit analysis may also outline the costs and consequences of not having the necessary information. This requires clearly articulated benefits that follow logically from census activities and outputs, along with plans to realise these benefits. Some of these benefits can be quantified, for example benefits generated through the use of census data. Other benefits can be more difficult to measure, but are nonetheless important and should still be recorded. Subsidiary objectives include illustrating the value of the census as an educational tool and framework, for comparative purposes at national and international level and as a cornerstone of the national statistical system. The results of this cost-benefit analysis are crucial for effectively advocating for the census, mobilizing resources, and engaging respondents. By demonstrating the clear value and return on investment of the census, it is possible to secure necessary support and funding while also building public trust and participation.

2.18. These objectives can be used as benchmarks to assess user requirements and may also be built into appraisal systems that, with suitable weighting, can be used to compare and review options. In general, strategic objectives of the population and housing census need to be clearly emphasized throughout the process of preparing, conducting and exploiting census data.

B. Strategic management

2.19. The primary value of strategic management is to assist census organizations to operate successfully in a dynamic, complex environment. The strategy drives key decisions and choices over the term of the census undertaking in response to external and internal forces. The crafting of a census management strategy is critical for a successful execution, and is dependent on a variety of knowledge and skills in different areas. While the development of a strategy is an art, it is important that it is governed by a systematic process to ensure careful examination and consideration of all issues that might have an influence on the future state of the programme. Census organizations may consider adopting the following strategic management process that will guide census operations in all its phases.

2.20. The strategic management process consists of four phases, namely:
(a) Strategy analysis;
(b) Strategy formulation;
(c) Strategy implementation; and,
(d) Strategy monitoring and review.

2.21. *Strategy analysis* is about ascertaining the issues that need to be addressed to take corrective action or to chart a new direction. It seeks to change the organizational set-up to one that is systemic, holistic, comprehensive, and coordinated. Strategy analysis consists of:

(a) Setting the direction for the census operation, with the objective of reaffirming its purpose and the way it should conduct its business;

(b) Setting strategic goals for the census operation by way of defining what the system aims to achieve in terms of defining its highest goals and strategic outcomes. Strategic outcomes and goals must be aligned to what the users can expect and must therefore address user needs and requirements. A strategy driven by outcomes means "planning backwards" from the desired outcome through how best to achieve it; and,

(c) Strategic analysis of the census operations by defining its current and previous status or situation and identifying the key issues that need to be addressed. Analysing the internal and external environment provides the evidence base to inform the development of the strategy. This analysis of the situation forms the basis of the strategy and its objectives to be defined. Census organizations may consider using a SWOT (strengths, weaknesses, opportunities, threats) analysis to do a strategic analysis of the internal and external environment.

2.22. *Strategy formulation* is about defining how and where the census organization must respond. The key steps involved in this phase include (a) develop and review the value chain; (b) develop strategic objectives (discussed below) and sub-objectives; (c) compile a strategy map; (d) define the strategic intent; (e) identify critical success factors; (f) identify strategic risks; and (g) identify or develop strategic interventions.

2.23. *Strategy implementation* is the process that turns strategies and plans into actions in order to accomplish strategic goals, outcomes and objectives. Implementing the strategic plan is as important as, or even more important than, developing the strategy. The key steps involved in this phase include:

(a) Compiling a census strategic plan, work programme and operational plan with the following elements:

i. Work planning: Compilation of the operational plan by defining the inputs, processes and outputs;

ii. Resource planning: Identifying what human and other resources are required and how these resources should be deployed and developed to create the competencies needed to deliver the strategy; For a full field enumeration census this might focus on determining the large numbers of enumerators and supervisors and their skills development. For register-based censuses with fewer individuals, the focus here might be on institutional knowledge retention and succession planning;

iii. Financial planning: Compilation of the budget outlining the funding required to implement the strategy;

iv. Contingency planning: Identifying the operational risks and developing control and mitigating actions in response to those risks.

(b) Translating the strategy into action through:

i. Defining the body of work (identifying key performance indicators, outputs to be delivered, targets and milestones that will deliver the strategy);

ii. Defining the method of work (developing a value chain at conceptual and operational levels, including the quality management process that defines how data users' value will be delivered);

iii. Defining the organization of work (designing an organizational structure that implement the strategy).

(c) Managing the implementation and strategic change.

2.24. Strategy *monitoring and review* is about monitoring and reporting on the progress, achievements and challenges in the programme; taking corrective action where required; and evaluating the impact of the changes and improvements. Implementation must be monitored to be successful. Due to constantly changing external and internal conditions, census managers must continuously review environments, new strengths, weaknesses, opportunities and threats may arise. The key element in strategy monitoring is to get relevant and timely information to take corrective actions where required.

C. Avoiding gender biases and biases affecting indigenous peoples and minorities

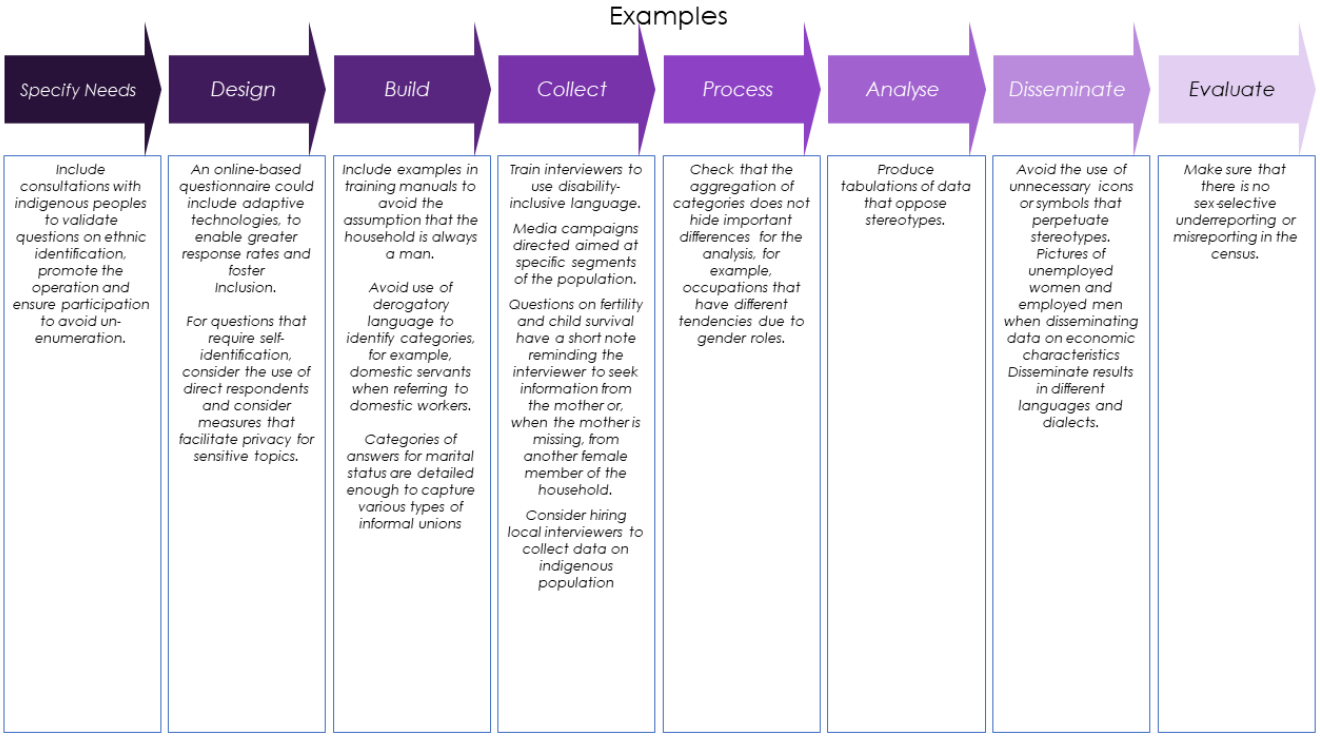
2.25. Censuses are the most important source of information on small population groups and at low levels of geography. While there is little to no sampling error in census data, bias can affect census data quality. Gender, indigenous peoples, or minority-based stereotypes have the potential to introduce biases in census data and the conclusions drawn from these data. These biases are discussed in more detail in Part Five (see for example paragraphs 5.142–5.153 and 5.336–5.452 relating to household relationships and economic characteristics, respectively). There is much, however, that can be done in the various stages of the census to help minimize gender-based and other population group biases (Figure 2.2.).

2.26. Issues of census content, including what information is sought and how, the definitions and classifications used, and the manner in which databases and tabulations are specified and disseminated, are important in generating data needed to examine questions of gender or population group equality. In addressing these content issues, census planners and users will need to be alert to prevailing stereotypes so as to develop a census that both minimizes the influence of the stereotypes that respondents and enumerators, where applicable, may hold and avoids further perpetuation of these stereotypes.

2.27. With regards to census operations, particular attention will need to be given to the selection, training and supervision of the field staff. This involves ensuring that women and men from various population groups and communities are recruited to the field staff (both as interviewers and supervisors) and that manuals and training materials cover gender or other bias issues just as they do other important sources of error. Consultations with women's groups and other civil society organisations concerned with specific population groups can help in addressing both content and operational issues. See paragraphs 3.530-3.532 for discussion of biases in the collection, processing, compilation, and presentation of data.

2.28. Gender-related stereotypes and biases are concerns that have relevance for all countries. Census authorities in a number of countries must also be alert to the possibility of stereotypes and biases affecting data on other population groups. Such groups may include persons belonging to national or ethnic, religious, linguistic, racial, or gender identity minorities. Persons with disabilities, migrants, and nomadic populations may often be subject to similar bias. As with gender issues, the problem will need to be addressed in terms of both census content and census operations. Representatives of these groups can often provide census planners with important information and insights relevant to both census content and operations (for detailed information about difficult-to-enumerate groups, see paragraphs 3.162–3.173 and 5.47). Thus, special efforts should be made to consult with them when planning the census. In the case of indigenous and populations living in isolated settlements or enclaves, such consultations are often critical for minimizing under-enumeration among these populations.

Figure 2.2. Examples of best practices on how to minimize population group biases



IV. Units, place and time of enumeration

A. Units of enumeration

2.29. Since individual enumeration is an essential feature of a population and housing census, clarity about the unit of enumeration is an essential element of census planning. In the case of the population census, the primary unit of enumeration is the person. There are two general frameworks within which individuals are identified: (a) households and (b) institutions as a subset of collective living quarters. The household is a general framework within which most individuals are identified. Since the majority of the population live in households, these are also considered units of enumeration. Therefore, particularly for full field enumeration censuses, careful identification as a preliminary step in the enumeration can facilitate the efficient collection of the data and the control of its completeness. For register-based censuses, the source data are often at the individual level, but data requirements from stakeholders may necessitate the construction of households from the individual-level source data.

2.30. As mentioned in the previous paragraph, the second framework within which individuals are identified comprises “institutions” as a subset of collective living quarters. In addition to persons identified within households, there are persons living in institutions who are not members of a household. This group constitutes the “institutional population”, which is also relevant to population censuses.

2.31. For the housing census, the household is one of the three units of enumeration; the other two units are living quarters (in other words, housing units and collective living quarters) and buildings. It is important to bear in mind that, in conceptual terms, these three units are clearly distinguishable. There is

not necessarily an identity or exact correspondence among these concepts nor are the terms themselves interchangeable. Several households may live together in one set of living quarters and one household may occupy more than one set of living quarters. Similarly, several sets of living quarters may together occupy one building and one set of living quarters may occupy more than one building.

2.32. It is recognized that there may be difficulty in some countries in maintaining independent concepts of "household" and of "housing unit".[40] However, the advantages in terms of the usefulness of the data that result from preserving separate concepts usually outweigh the additional effort required in maintaining them.

2.33. In carrying out a census, it is essential that the units of enumeration be clearly defined and that the definitions be included in manuals of instruction for the enumeration and, to provide appropriate guidance for users of the resulting statistical information, in census reports. In order to reduce the possibility of difficulties in applying the definitions recommended below, countries may find it necessary to expand the definitions and to illustrate them in terms of national conditions and circumstances. Post-enumeration field checks can provide a useful means of determining to what extent the national definitions of the units of enumeration have been applied in the field and the consequent effects on census results.

1. Person

2.34. For census purposes, the term "person" denotes each individual falling within the scope of the census. As emphasized above (paragraph 2.29), a person can be identified as belonging to the household population (that is to say, the population living in households) or to the institutional population (that is to say, the population living in collective living quarters), as defined in paragraph 2.41 below. Although each person must be included in the count of the population, there will be some variation in regard to the persons for whom information is collected on different topics. The variations usually depend on the person's:

(a) age (for example, questions relating to economic activity, in which case the age boundary may be driven by national legislation);
(b) sex (for example, questions relating to children born), or
(c) relationship to the reference member (or head) of the household.

It may be recommended that information on a particular topic need not be investigated for the entire population, and the group of persons for which a given topic is investigated is indicated under the definitions and specifications of such topics presented in Part Five, Chapter I, section D. Similarly, the scope of the census should clearly indicate the persons to be covered and those to be left out.

2. Household

2.35. The concept of "household" is based on the arrangements made by persons, individually or in groups, for providing themselves with food and other essentials for living. A household may be either:

(a) a one-person household, that is to say, a person who makes provision for their own food and other essentials for living without combining with any other person to form a multi-person household; or

[40] For further discussion on the concept of households, see paragraphs 2.35-2.40; also, for the definition of housing unit, see paragraph 5.491.

(b) a multi-person household, that is to say, a group of two or more persons living together who make common provision for food and other essentials for living. The persons in such a group may pool their resources and may have a common budget; and they may be related and/or unrelated to each other.

2.36. The concept of a "household" as described in paragraph 2.35 is known as the "housekeeping concept". It does not assume that the number of households and housing units are, or should, be equal. A housing unit, as defined in paragraph 5.491, is a separate and independent place that is intended for habitation by one household, but that may be occupied by more than one household or by a part of a household (for example, two nuclear households that share one housing unit for economic reasons or one household in a polygamous society routinely occupying two or more housing units).

2.37. However, some countries use a concept different from the housekeeping concept described in the previous paragraph, namely, the "household dwelling" concept, which regards all persons living in a housing unit as belonging to the same household. According to this concept, there is one household per occupied housing unit. Consequently, the number of occupied housing units and the number of households occupying should always be equal and the locations of the housing units and households are identical. However, this concept can obscure some information on living arrangements, such as multi-household or shared dwellings, that is relevant for evaluating housing needs.

2.38. Households usually occupy the whole or a part of one housing unit, but they may also be found in camps, boarding houses or hotels, or as administrative personnel in institutions, or they may be homeless. "Families" on the other hand may occupy more than one housing unit, if for example there are polygamous unions, or other second homes. For more discussion of household occupancy, see paragraphs 5.535–5.539 and of household and family characteristics, see paragraphs 5.134-5.141.

2.39. A household may also consist of one or more homeless people. The definition of the homeless can vary from country to country because homelessness is essentially a cultural definition based on concepts such as "adequate housing", "minimum community housing standard", or "security of tenure",[41] which can be perceived in different ways by different communities. The following two categories or degrees of homelessness are recommended:

(a) Primary homelessness (or rooflessness): This category includes persons living in streets or without a shelter that would fall within the scope of living quarters;

(b) Secondary homelessness (or rootlessness): This category may include the following groups:

i. Persons with no place of usual residence who move frequently between various types of accommodation (including dwellings, shelters or other living quarters);

ii. Persons usually resident in long-term (also called "transitional") shelters or similar arrangements for the homeless.

2.40. For some topics investigated in housing censuses, the household may serve more efficiently than living quarters as the unit of enumeration. For example, topics such as tenure and rent, if investigated in the census, should be collected with reference to households rather than living quarters, whereas type of ownership is a characteristic more relevant to the living quarters. Similarly, information about household possessions that are normally included as part of the equipment of living quarters (radio, television receivers and laptops, for example) should be collected with reference to households.

[41] For the definition of tenure, see paragraphs 5.619-5.622.

3. Population in collective living quarters

2.41. As emphasized in paragraph 2.29, institutions represent the second general framework within which persons, as major units of enumeration, are identified. The institutional population comprises persons who are not members of households. These include persons living in military installations, correctional and penal institutions, dormitories of schools and universities, religious institutions, hospitals, refugee camps and so forth.[42] Personnel responsible for the running of an institution and not living in dormitories or similar accommodations should be excluded from the institutional population. Similarly, some buildings housing seniors may have different levels of care, with independent living quarters that should not be considered as collective living quarters like some other quarters with higher levels of care. Although the members residing in the institutional households constitute a single unit, keeping in mind that in future more and more people would submit the responses in census through self-enumeration and past data may also be used in future censuses, each member may be enumerated as single member household.

2.42. Persons living in hotels or boarding houses are not part of the institutional population and should be distinguished as members of one-person or multi-person households, on the basis of the arrangements that they make for providing themselves with the essentials for living.

4. Building

2.43. The building is regarded as an indirect but important unit of enumeration for housing censuses since the information concerning the building (building type, material of construction and certain other characteristics) is required for proper description of the living quarters located within the building and for the formulation of housing programmes. In a housing census, the questions on building characteristics are normally framed in terms of the building in which the living quarters enumerated are located, and the information is recorded for each of the housing units or other living quarters located within it.

2.44. A building is any independent free-standing structure comprising one or more rooms[43] or other spaces, covered by a roof and usually enclosed within external walls or dividing walls[44] that extend from the foundations to the roof. However, in tropical areas, a building may consist of a roof with supports only, that is to say, one without constructed walls; in some cases, a roofless structure consisting of a space enclosed by walls may be considered a building.[45]

2.45. In some countries, it may be appropriate to use the "compound" as a unit of enumeration, either in addition to the building or as a substitute for it. In some areas of the world, living quarters are traditionally located within compounds and the grouping of living quarters in this way may have certain economic and social implications that it would be useful to study. In such cases it may be appropriate, during the census, to identify compounds and to record information suitable for linking them to the living quarters located within them.

[42] For more detailed definition and specifications of institutions as a subset of collective living quarters, see paragraphs 5.517-5.526.
[43] For the definition of rooms, see paragraph 5.546.
[44] The term dividing walls refers to the walls of adjoining buildings (for example, of row houses) that have been constructed so as to be contiguous.
[45] For a more detailed discussion of the definition of building and related concepts, see paragraphs 5.589-5.592.

5. Living quarters

2.46. The principal units of enumeration in a census of housing are living quarters. Only by precise recognition of these identities can data be obtained that will provide a meaningful description of the housing situation and a suitable basis for the formulation of housing programmes and policies.

2.47. Living quarters are structurally separate and independent places of abode. They may:

(a) have been constructed, built, converted or arranged for human habitation, provided that they are not at the time of the census used wholly for other purposes and that, in the case of improvised housing units and collective living quarters, they are occupied; or

(b) although not intended for habitation, actually be in use for such a purpose at the time of the census.[46]

B. Place of enumeration

1. Concepts relating to the place of enumeration

2.48. In the context of the population census, a country may wish to enumerate all persons present in the territory and/or potentially belonging to the population of interest. The *population to be enumerated* is the group of persons whom the country decides should be covered by the census, regardless of their later inclusion in population count. In other words, it is the group of people that form the target population of the data collection activities (be that direct enumeration or compilation of data from registers or administrative sources). A country might collect data on everyone present at a point in time, but then only produce official statistics referring to those who are usually resident, for example.

2.49. The place of enumeration is either the place where the person is found, or the place of usual residence of the person, at the census reference moment. It should be ensured that each person has only one place of enumeration, to avoid double-counting. Countries should document the definition of place of enumeration adopted for their census. In those countries which use direct enumeration, they should also provide explicit instructions to enumerators on how this definition should be applied during an interview, or to respondents when filling in self-administered questionnaires.

2.50. In general, *usual residence* is defined for census purposes as the place at which the person lives at the time of the census, and has been there for some time or intends to stay there for some time (i.e., the geographic place where the person usually spends their daily rest, assessed over a defined period of time including the census reference time). Most individuals will have not changed residence for some time and thus defining their place of usual residence is unambiguous. For others, however, the application of the definition can lead to many interpretations, particularly if the person moves or has moved residence often.

2.51. It is recommended that countries apply a threshold of 12 months when considering place of usual residence according to one of the following two criteria:

[46] For a more detailed discussion of the definition of living quarters and of the concepts of separateness and independence as used in the definition, see paragraphs 5.485-5.486.

(a) The place at which the person has lived continuously for most of the 12 months before census reference day (that is, for at least six months and one day, not including temporary absences for holidays or work assignments) or intends to live for at least six months;

(b) The place at which the person has lived continuously for at least 12 months before census reference day (not including temporary absences for holidays or work assignments) or intends to live for at least 12 months.[47]

2.52. For register-based censuses, the place of usual residence could be taken to mean the place of legal or registered residence, defined with reference to a qualification period of 12 months or assessed with reference to the 12-month criterion.

2.53. Persons who move frequently and do not have a place of usual residence should be enumerated at the place where they are found at the time of the census.

2.54. Regardless of the criteria used to define the 12-month period, countries should ensure that each person has one and only one place of usual residence (or registered place of residence for census purposes).

2.55. There are various population groups for whom some uncertainty may arise about their inclusion in the usual resident population. The following persons should generally be considered as part of the usually resident population:

(a) Persons found at the moment of enumeration who cannot identify a place of usual residence, such as those who change residence often;

(b) National military, naval and diplomatic personnel and their families, located outside the country, irrespective of their duration of stay abroad;

(c) Foreign citizens working within the country for international businesses or organizations (but not including foreign diplomats or military forces) and their families, provided that they meet the criteria for the usual residence in the country;

(d) Merchant seafarers and fishers usually resident in the country but at sea at the census reference time (including those who have no place of residence other than their quarters aboard ship);

(e) Persons who are irregular or undocumented migrants, as well as asylum seekers and persons who have applied for, or been granted, refugee status or similar types of international protections, provided that they meet the criteria for usual residence in the country (the intention is not to distinguish these persons separately, but rather to ensure that they are not missed from the enumeration);

(f) Persons who cross a frontier daily or weekly to work or study in another country, provided that they meet the criteria for usual residence in the enumeration country;

(g) Children born in the 12 months before the census reference time and are usually resident in the country at the census reference time; and

(h) Persons who regularly live in more than one country during a year, if they reside in the country conducting the census most of the time, regardless of whether they are physically present in that country at the census reference time.

[47] Conceptual frameworks and Statistical Concepts and Definitions on International Migration (2021). https://unstats.un.org/unsd/demographic-social/migration-expert-group/task-forces/TF2-ConceptualFramework-Final.pdf

On the other hand, the following group of persons should generally be excluded from the usual resident population:

(a) Foreign military, naval and diplomatic personnel and their families, located in the country, regardless of their duration of stay and/or place of usual residence;

(b) Third-level students who are or intend to be absent from the country for one year or more;

(c) Persons who regularly live in more than one country during a year, if they reside in the country conducting the census the least amount of time, regardless of whether they are physically present in that country at the census reference time.

2.56. The concept of usual residence is often referred to as though it is synonymous with the concept of de jure residence. In certain circumstances, however, the term “de jure” may carry with it a requirement that the person’s residence at that place has a basis in the legal system applicable to that specific place. In turn this implies that people without such a legal basis should not be enumerated in that area. It is not recommended that censuses of population and housing limit their enumeration only to those people with a legal right to be in a place but rather, as described in section 2 below, they should include either all those present at the census reference time or all those whose usual residence at the census reference time was at the place of enumeration.

2. Operational issues relating to the place of residence and the place of enumeration

2.57. In a field enumeration-based population census, information about each person can be collected and entered in the census questionnaire either where he or she is (or was) present on the day of the census or at his or her usual residence.

2.58. In compiling the census results by geographic areas, however, each person who is part of a household can be included in either:

(a) the household (and hence the geographic area) where the person was present on the day of the census; or

(b) the household (and the geographic area) where he or she usually resides.

The same also applies for the institutional population. This allocation is not necessarily dependent upon the place at which information was collected for the individual but it can be simplified by the proper choice of a place of enumeration.

2.59. If a "present-in-area" population distribution is wanted, it is logical to enumerate persons at the place where they are (or were) present at the time of the census. If a distribution by usual residence only is required, it is more satisfactory to collect the information about each person at the person's place of usual residence. It should be noted, however, that it is not always possible to collect information about each individual at their usual residence, as, for example, when an entire household is away from its usual residence at the time of the census. Some provision must therefore be made for collecting information about such persons at the place where they are found at the time of the census.

2.60. If the objective is to obtain information on *both* the usually resident population and the present-in-area population, then either each person present in each household or institution on the census day, or each person present and each usual resident temporarily absent can be enumerated at the appropriate household or institution. A clear distinction must then be made in the questionnaire, as applicable,

between (a) persons usually resident and present on the day of the census; (b) persons usually resident but temporarily absent on the day of the census; and (c) persons not usually resident but temporarily present on the day of the census.

2.61. Depending on the categories of persons enumerated at any given place, information may then be collected both on the usual residence (address) of those only temporarily present, and on the place (address) at which each temporarily absent person can be found. This information can be used for the purpose of allocating persons to the household (or institution) and geographic location at which they are to be counted and for checking that no person is counted twice (namely, at both the usual residence and the place where present). The procedures to be followed during the enumeration and through the subsequent allocation of persons must, however, be very carefully planned and strictly adhered to if the allocation is to be accurate.

2.62. With the exception of mobile housing units (see discussion in paragraph 2.64), living quarters and buildings have a fixed location. The place where they are to be enumerated does not have, therefore, to be considered in taking a housing census. Information on households, however, and on the persons in households, can be collected and entered in the housing census questionnaire (for census with direct enumeration) either where they are (or were) present on the day of the census or at the usual residence. The procedure followed in the housing census should be governed by that adopted in carrying out the population census if the two censuses are carried out simultaneously. If the housing census is an independent operation, however, the procedure to be followed should be carefully considered since it may have a significant effect on the validity of the results of the housing census.

2.63. Where persons and households are allocated to the place of usual residence, they should also be allocated to the living quarters that they usually occupy. The living quarters that they are actually occupying at the time of the census should be counted as vacant if they are conventional dwellings or should be excluded from the census if they are non-conventional dwellings.[48]

2.64. Mobile housing units represent a special case as far as the place of enumeration is concerned. They should be enumerated where they are found on the day of the census. However, in accordance with the procedure adopted for the allocation of the population, mobile housing units may also be allocated to the area where the occupants usually reside, provided that they are the usual living quarters of the occupants in the area of usual residence. Where they are not the usual living quarters of the occupants in the area of usual residence, the occupants will be allocated to their usual living quarters and the mobile housing unit will be excluded from the census.

C. Enumeration point of time

2.65. One of the essential features of population and housing censuses is that each person and each set of living quarters must be enumerated at the same, well-defined point of time. This is usually accomplished by fixing a census “moment” at midnight at the beginning of the date selected as the “census day”. This moment is the “census reference moment”, a particular point in time to which the census is referring.

[48] To be considered as living quarters, non-conventional housing units and collective living quarters are required to be occupied in order to be included in the census.

2.66. For the population census, each living person who is part of the target universe at the census moment should be included, even if the process of completing the schedule or inputting the administrative data takes place after the census day. Therefore, infants born after the census moment should not be included, even if they are living at the time of the enumeration. Likewise, persons living on the census day but who have deceased before the enumeration should be included. In both field-based and register-based censuses, a rigorous application of the concept of a census day is important to produce high-quality data, especially given the fact that the enumeration process in a field-based census can extend for a period after the census day. With a register-based census with unclear or multiple reference dates, modelling techniques may be required, to ensure the administrative enumeration is focused on the census day.

2.67. For the housing census, each set of living quarters that has reached an established stage of completion and is not scheduled for, or in the process of, demolition should be included in a census schedule and counted as a part of the housing inventory even though the process of completing the schedule does not take place until after the census moment or even after the census day, and the living quarters may have been scheduled for demolition in the interim. Living quarters that have attained the prescribed state of completion after the census moment are not to be entered in a schedule (unless special instructions are issued for recording living quarters under construction), nor should they be included in the total number of sets of living quarters.

D. Time reference period for data on the characteristics of the population and of living quarters

2.68. The data collected about the characteristics of the population and of living quarters should refer to a well-defined reference period. The time reference period need not, however, be the same for all of the data collected. For most of the data, it will be the census moment or the census day; however, in some instances (as is the case for current economic characteristics and rental arrangements), it may be a brief period just prior to the census or (as is the case for fertility questions, usual economic activity and information on the period of construction of the building in which living quarters are located) a longer period of time.

V. Legal basis

2.69. Legal authority for conducting a full field enumeration or administrative-based census is required for:

- regulating primary administrative responsibility;
- obtaining the necessary resources;
- determining the general scope and timing of the census;
- placing a legal obligation upon the public to cooperate and provide truthful answers
- providing a mandate to collect and use administrative data for the enumeration;
- establishing the obligations of register-holders to provide the requisite data and to inform about any changes to the content and/or structure of the relevant registers and
- imposing a legal obligation upon the enumerator and census officials to record the responses faithfully, and to perform specific responsibilities upon other census field personnel at various supervisory levels.

2.70. In some countries, the legal basis for a census may be specified in legislation that is separate from that for the general statistical authority. In this case, linkages should be made between census and other statistical legal frameworks.

2.71. In addition, the confidentiality of the individual information should be strongly and clearly established in the census legislation and guaranteed by adequate sanctions so as to provide a basis for the confident cooperation of the public, or assurance to the public that available administrative data is managed appropriately.

2.72. In countries that lack permanent legal authority for the taking of periodic censuses, it is important to act early to establish ad hoc legal authority or, preferably, legislation calling for a system of periodic censuses. In case of an ad hoc legal authority, attempts should be made to ensure the independency of the implementing agency is safeguarded and powers to make technical and administrative actions are delegated to the said agency.

2.73. The principle of conceptual and organizational flexibility should be observed in drafting the census legislation. The legislative provisions should ensure data security and confidentiality. However, the inclusion of provisions that may be too rigid regarding the type of data to be collected or the structure and relationships of the various parts of the census organization is undesirable. Rather, necessary details should be contained in secondary census regulations promulgated by the census authorities. Moreover, provision may have to be made, in either the legislation or the regulations, for sanctioning the use of simplified administrative procedures, including the appropriate delegations of authority for the procurement of equipment and supplies and the recruitment of personnel during the operational phase of the census.

2.74. While the content of the census legislation will inevitably depend on national legal practices and procedures, as well as on the organization of the national civil service, the following components are usually represented:

- the purpose of the law;
- the coverage of the census;
- assigning the mandate for conducting a census to a specific institution;
- purpose of the census;
- obligation and rights of the citizens;
- modes of financing the census;
- organization of the census;
- the bodies responsible for administering the census;
- the mode of data collection;
- access to, and use of, record-level administrative data;
- the technology to be used;
- rights and obligations of enumerators and supervisors;
- census data dissemination and exploitation;
- treatment of individual data;
- confidentiality and privacy of respondents and their data;
- disposal of the census equipment and materials; and
- archiving.

2.75. The legal basis for a census should be continuously improving, based on past experiences. At each evaluation of a census cycle, adaptations to legislation should be considered, in order to strengthen

protections to the population, to foster the transparency and accountability of the census-taker, to improve public engagement and to ensure high quality of outputs.

2.76. A comprehensive and well-timed legislative framework is of utmost importance for ensuring the legality and authority for conducting the census itself. In the case of an ad hoc approach to the census legislation, that is, in the case where it is done before each census, it often also contains the census topics clearly spelled out, thus providing additional legal weight to the composition of the questionnaire and the content of the census.

2.77. In addition to legislation that specifically provides for the necessary authority and parameters of a census, there may be legal requirements for the various data elements that a census provides. That is, there may be laws that establish the census as the required source of data for particular purposes, such as the delineation of electoral boundaries or transfers of funds between central and regional governments within a country.

VI. Financial management

A. Financial basis for censuses

2.78. A census is the primary source of data about the size and characteristics of the population disaggregated into lower geographic levels; it provides a demographic profile of a country and is the basis for developing area sampling frames for use in surveys. A census, however, is usually one of the largest and costliest statistical activities that governments and their national statistical offices undertake. As a result, some countries have been forced to delay or even cancel a census owing to funding constraints. Countries that have been able to secure partial funds or secure funds but at a late stage of their census preparation have been forced to compromise their data collection, data processing, and dissemination of census results. It is therefore recommended that all census operations, including planning, cartography, enumeration, processing, analysis and dissemination, be fully budgeted from the beginning and efforts made to mobilize the required resources, for the entirety of the census cycle (i.e., for all fiscal years). Inflation should be taken into account, keeping in mind that duration has an impact on cost.

2.79. Hence, there is growing pressure to look into the solutions to census funding, taking into account the role of key stakeholders, namely governments and their statistical agencies, and the greater involvement of national and international development partners and stakeholders including the private sector. Concurrently, cost-effective strategies need to be put in place that would reduce census costs without compromising the quality of census data.

2.80. It should be emphasized, however, that censuses cannot be carried out merely by national statistical and census offices alone. Rather, conducting a census should be seen as a national task involving many stakeholders. Thus, government departments, non-governmental organizations and private sector end users should be consulted at all stages to ensure the legitimacy of, and need for, conducting the census and, at the same time, to improve advocacy for sufficient funding. Although conducting a census is principally financed by the government, the census must be designed in partnership with all political actors so as to obtain their support and involvement in the census process. A high-level committee consisting of the government, the private sector and civil society, including non-governmental organizations, communities and donors, could be formed to discuss issues related to the funding of, and resourcing for, the census.

2.81. National statistical and census offices need to advocate the importance of investing in censuses within their own governments. It is also important for the national statistical and census authorities to ensure continuous feedback and promote the use of statistical data from previous censuses, in order for users to recognize and appreciate the importance of the population census as a source of crucial statistical data and give their support. The possibility of reducing the financial costs through collaborations with other government departments and local authorities with existing devolved infrastructure, such as education and health ministries, should be further explored. These institutions could be supportive in providing administrative data and/or logistics arrangements for the census, such as the use of existing infrastructure, transportation, and communications facilities, and sharing of employees of other government departments.

2.82. Excellence in planning is an essential prerequisite not only for achieving a cost-effective census but also for securing comprehensive financial support for its funding. Technologies and methods that will be used in mapping, data collection, administrative data intake and management, processing, questionnaire design, and other activities must be decided upon in advance, as these have an influence on costs. Census planning must bring out the links between the various components, which will include types of resources (such as personnel, cost of equipment, stationery or printing) as well as tasks (including data collection and capture, data processing, data management, dissemination and archiving). Cost tags must be attached to each of these components together with a justification. Experience and lessons from past censuses or similar activities must be considered when estimating costs for the next census. Where multiple modes of data collection and new technologies are being used for the first time, these must be tested for data quality and cost implications.

2.83. For each stage of the census process the costs must be optimized. A careful choice of the appropriate technology will greatly assist in this. Recent advances in technologies throughout the field-based census process, such as digital mapping, computer-assisted or Internet data collection, scanning, data processing (including machine learning technologies for coding) and data management and archiving, and census data analysis and dissemination, may be of assistance in achieving significant efficiencies, allowing more cost reductions or increased productivity. In addition, the proper selection and use of such technologies will speed up the computation of results and enhance their preservation. However, the choice of technology should be made only after carefully evaluating the costs and benefits of possible options. Some potential risks to canvass include the following:

- some approaches only become cost-effective for large operations;
- some are dependent on expensive and scarce inputs (for example very high-quality satellite images or paper for scanning);
- some are dependent on services that may not be available throughout the country (for example Internet access); and
- others require significant investments in high-quality computers and upfront investment in human resources.

The options examined in the cost–benefit analysis could incorporate consideration of leasing (rather than purchasing) equipment or sharing/borrowing it between countries that are undertaking censuses at convenient times.

2.84. Outsourcing or collaborating with the private sector could be considered as another potential cost-saving option, particularly in the context of publicity and advocacy or for systems development for data collection, transmission, processing, dissemination and archiving. Outsourcing can contribute

technical expertise or resources not readily available within the national statistical office or census implementing agency.

2.85. It is anticipated that international donors will continue to play a pivotal role in supporting funding census costs in many countries. Technical cooperation and assistance from international agencies have also contributed greatly to the success of censuses in many countries. It is worth noting that a population and housing census has some intangible positive values. It is an opportunity for mobilizing the whole country and reaching even the most remote corners of it. In the life of many citizens, a regular census is often the only time that the State reaches out to them and asks them some questions. Successfully conducting a census is a matter of pride in many countries and a welcome opportunity to recruit a massive labour force and generate jobs and train people in valuable tasks (such as data entry) or in other ways to add to the national infrastructure.

2.86. There is opportunity for cost savings through the use of administrative data to replace or support field-based census operations, usually the most expensive phase of census-taking. As countries increasingly integrate administrative data into their censuses, this can be an important driver of change. However, considerations should be made for the added expense associated with the acquisition, intake, processing and other management (e.g. modelling) of administrative data used for census-taking. Also, if administrative data cannot entirely replace what is collected traditionally, good financial planning will be required to supplement any continued surveying that must continue, post-transition to administrative data.

2.87. In general, population and housing censuses are exclusively the responsibility of national governments and structures; this is particularly true for funding the census. Thus, all activities related to funding need to be elaborated, documented, justified and presented to all stakeholders in a transparent and comprehensive manner.

B. Budget and cost control

2.88. While no universal system of census budgeting and cost control can be suggested - since financial practices vary greatly among countries - a few generally accepted principles can be noted. First and foremost, effective planning and control of the various census operations is not possible without a very careful financial estimate of the cost of each census operation, including all of its components, no matter how small. It is recommended to draft a detailed list of activities related to censuses with timelines and, as far as it is possible, to draft the budget in such a way that it corresponds to this list of activities. Secondly, it is critical for this census plan and budget to be presented by national statistical and census agencies to their respective governments with adequate lead time, to facilitate the timely appropriation of sufficient resources from national budgets or, where required, from the international development community. Moreover, funding of the census must be accompanied and developed on a sound and adequate legal basis if effective national census operations are to be enabled.

2.89. Information on expenditures from the previous census classified by census phases, starting with the expenditure for different elements of the preparatory work and ending with expenditure for the dissemination of the census results, provides an important basis for estimating the budget of the census. Figures from the previous census will of course have to be reviewed and updated in order to take into account:

- quantitative and qualitative change in hardware and software;

- changes in wage rates;
- the costs of equipment, supplies and other materials;
- planned changes in census content, methods and procedures; and
- anticipated changes in the population itself (for example, total size, percentage urban, and average household size).

2.90. All of the above may affect the cost structure of the census. In most countries, several cost elements tend to increase (for example, wage rates and the size of population) so that there is considerable pressure to achieve economies in other items of the census budget. In addition, time must be spent to identify new features of the next census, for which a past expenditure cannot provide the basis of a cost estimate. For example, future censuses might need to consider a new budget for cybersecurity, and/or the acquisition and use of administrative data.

2.91. The census offices need to implement transparent accounting procedures and financial management systems to ensure speedy disbursement of funds, proper receipting of expenditures and an efficient audit. This would enable prompt release of periodic allocations of census funds by national governments. A clean outcome from a financial audit adds credibility to the census process so that the government and civil society are more likely to accept the final results.

2.92. In the case of external or donor funds, the required conditions should be established well in advance by discussion between the donor and the national statistical or census office. This will avoid delay in the release of such funds for census operations.

2.93. Control measures and monitoring systems must be developed for cost-effectiveness. Activities to be outsourced must be clearly defined and contracts for outsourcing should be well prepared with clear deliverables and timelines.

2.94. For planning the costs of a census, detailed and precise data will be required on the following:

(a) number and cost of census staff classified by function and manner of payment;

(b) type of equipment and material used for the census, manner of acquisition (purchase or rental) and cost;

(c) office space (surface measurement), classified by use and type of cost (that is, for construction or for rent); and

(d) type of services used for census operations.

2.95. The usefulness of the above information would be enhanced if the information could be recorded by source of funding, in other words, in terms of whether the expenditure has come from:

(a) the official census budget;

(b) other funds of the census office (for example, a regular annual budget not specifically intended for census purposes, or general funds of the governmental agency or department of which the census office is a part);

(c) other parts of the government;

(d) non-governmental organizations; or

(e) international donors.

2.96. This information is needed not only for fiscal planning and control but also in order to examine the trade-offs in terms of costs and benefits among alternative ways of carrying out various census operations. Although cost experience from a previous census in a country may provide useful experience

for planning the next census, much more caution should be exercised in using the cost parameters from other countries. Differences in census content, organization and operations, as well as in cost accounting, can introduce serious incompatibilities into such country-to-country cost comparisons.[49]

2.97. It is important that the persons at the administrative and supervisory levels who will be responsible for the execution of each operation should participate in estimating the budget items. Such an organization of the work pre-supposes detailed advance planning and "cost consciousness" on the part of those responsible for a census.

2.98. The census plan as executed will certainly change in a number of respects after the making of the original calculations. Consequently, a perfect correspondence between the estimates and the final costs is not to be expected. Changes in the prices of major components of census costs should be monitored on a regular basis with either the census budget adjusted accordingly or the census plans modified. Indeed, the development of the census budget is usually an incremental process in which rough initial estimates are replaced by more detailed and precise statements of resource requirements. Throughout the period of census taking and compilation of census results, the budget will have to be re-examined and performance compared with plans. With detailed information on expenditure, the governmental and census authorities will be better able to exercise control over keeping the development of census operations within the census budget and to assess and control the effectiveness and efficiency of these operations. This information is also very useful for studying possible improvements in census techniques and census methodology.

2.99. As with any project, particularly ones as large and complex as the census, it will not, as a general rule, go according to plan and there will be difficulties and deviations. Accepting this at the outset and making sure there are arrangements in place to deal with delays, changes or other unforeseen issues is essential. In particular, there must be resources set aside to enable such issues to be dealt with quickly. Therefore, some contingency funding should be included within the overall costs of the census, and some controls put in place to monitor and allocate the contingency pot. Different methods exist for estimating the cost of the contingency budget, such as risk modelling, but a good starting point might be to allocate a percentage of the annual budget (say 15 per cent) each year for such contingency.

VII. Resource mobilization

2.100. A census is a massive undertaking that requires huge resources for its execution, particularly for countries with full field enumeration censuses. The resources range from financial resources to fund the costs of various activities and equipment, huge number of personnel with various skills and qualifications to implement various activities, and materials and equipment. There is need therefore for the implementing agencies to develop a comprehensive resource mobilization strategy for a census prior to commencement of the exercise. The strategy should detail the required resources, timelines, and proposed mechanisms or strategy for mobilizing the resources.

2.101. A resource mobilization committee for the census may be required and should be established well in time to develop, implement and monitor implementation of the resource mobilization strategy for the census. The strategy should have a detailed workplan and implementation plan to be monitored.

[49] See United Nations Economic Commission for Europe, Measuring Population and Housing: Practices of UNECE Countries in the 2010 Round of Censuses (New York and Geneva, 2014).

2.102. As censuses are repeated over time, the resource strategy should be sustained and be adapted with each cycle. In other words, resource mobilization should not be a one-time action but instead should seek sustainable and impactful longer-term outcomes. For example, in countries where external census funding is required, ensuring the diversification of grants and partnerships is a best practice that minimizes any impact from a change in funding. In addition, staff capacity building through training will strengthen the census' capacity to mobilize at the next cycle.

VIII. Project management

A. Development of workplans

2.103. The Project Management Institute defines projects as "temporary efforts to create value through unique products, services, and processes". Censuses – whether field-based or register-based - are major projects which require very careful, advanced planning and governance. Their complexity necessitates a large number of people working together, with various skill types and levels. At the centre, orchestrating it all, should be a census manager and governance that pulls the necessary team members together to make planning and execution decisions.

2.104. During the planning phases of a census, the team managing a census will need to detail their workplans. Using previous censuses as a benchmark, the team should consider what is required for the next cycle or iteration of the census, and articulate a vision. This high-level plan for the next census will require adaptations or new development of multiple workplans. These workplans should describe what is planned for the next census, the resources (human, financial, technical and other) required to achieve this outcome, and the timelines for planning, developing and implementing it.

2.105. Multiple workplans will be required. Each should be prepared by an assigned team member with the greatest expertise, after they are fully articulated with the appropriate inputs from as many elements of the census program as necessary. The assigned team member who prepares the workplan should also be the same person who presents updates on how plans are developing. In other words, it is the person accountable for planning and executing on the workplans.

2.106. To ensure clarity, census project teams developing their workplans should have "charters" that describe their function within the census program. These charters should describe roles and responsibilities that make it sufficiently clear as who should be the responsible team member taking the lead on the development of each workplan. The charter should also integrate with the overall census project charter, a document that provides the basic parameters for the next census, developed from previous cycles and adapted to incorporate the vision for the next.

2.107. A key product of the workplans should be a "critical path" of milestone dates for the completion of work to prepare for the next census. This critical path represents key dates when if one activity is not completed, it would have downstream impacts on other parts of the census plan. Other dates should be tracked as well, but the critical path defines the most important dates that the census team should be reviewing regularly and with greater frequency at key points during the census cycle (for example, in the months leading up to field operations of a traditional census).

2.108. It is recommended that all census offices have, as a member of staff, at least one certified project management expert, who can develop project plans and other elements crucial to the successful completion of the census.

B. Emergency preparedness, risk management and contingency planning

2.109. The proper conduct of a population and/or housing census requires thorough planning that often spans years of work by dedicated and knowledgeable staff. However, those plans can be interrupted by unexpected and very serious events, such as natural disasters, pandemics, political and civil unrest, insecurity or war. These serious events can lead to the failure, delay or cancellation of a census, in particular for full field enumeration, if there is no time or ability within a country to respond properly. But it is not just these serious situations that can cause census-takers great grief. Given the enormous complexity of censuses, less catastrophic incidents, such as labour issues or technical challenges, can also cause considerable problems for census managers.

2.110. Unfortunately, the context for census-taking is getting increasingly complex and risk-prone. Censuses are more complicated and expensive than ever, with new technologies and techniques being adopted to enumerate populations more effectively. In addition, it seems that these more complex censuses are being run in a world with increasing risks. For example, in the 2020 round of censuses, countries were faced with a world-wide pandemic, numerous natural disasters (some associated with climate change), new cybersecurity threats and tight labour markets. While some countries had time to manage and adapt to this rapidly changing context, others were forced to adapt "on the fly", with varying levels of capability to do so.

2.111. In response, countries require new support managing issues and risks related to censuses. Census-takers in all phases of a census require added support, but especially during high-risk periods such as during the field operations of a traditional census.

2.112. Emergency preparedness and contingency planning are broad topics which can mean different things to different people. The following guidance demonstrates how to avoid issues in the first place by effectively using the planning phase for risk mitigation, but also to support census takers when the inevitable happens during the execution phase. Also, the following is designed to provide support and ideas for when, despite good planning and emergency management, a census has a major break-down of its intended plan.

1. Documentation of risks and mitigation plans

2.113. To avoid or manage risk, one must first identify it. Census takers should be aware of their most important risks and set up processes to document them by creating a risk register. As defined by the International Standards Organization (ISO) in their risk management standard ISO 31000, a risk is "…anything that generates uncertainty or creates a deviation from the expected".[50]

2.114. In a census context, these are issues that may threaten the proper execution of a census, the quality of the data, the security of the information collected, or the availability of human or financial

[50] International Standards Organization. (2021, December 10). Effective risk management. New international guidance on how to get it right. https://www.iso.org/news/ref2773.html

resources. As a country plans to conduct a census – either field-based, register-based or combined – time should be dedicated to on-going discussions with census management teams, in order to identify and manage risks. When attempting to identify risks, census takers should build upon previous experiences by asking themselves such questions as: "What are the threats to the census"?; "What happened the last time?"; and "Can it happen again?". However, it is crucial to recognize that some incidents cannot be anticipated based on previous experience. This underscores the importance of considering the current context, which is constantly evolving. Unforeseen challenges, such as pandemics or natural disasters, may emerge. Therefore, a proactive approach incorporating scenario planning and contingency measures is essential.

2.115. Risks can be further assessed using categories such as "low", "moderate", "high" and "extreme" to characterize their inherent likelihood of occurrence and the level of impact if the risks were to materialize. With this approach, national statistical and census organizations should be better aware of the vulnerability of the census to various threats.

2.116. Census managers and experts should then discuss mitigations, or ways to lower the likelihood of occurrence and/or the impact of any threat to a census. Management must often prioritize this work, focussing on mitigations for the most significant risks, and/or those risks which are most likely to affect the census in the near term.

2.117. The best way to deal with a risk is to avoid it altogether through mitigating steps, but not all risks are fully within the control of the statistical agency. It may be important to discuss risk mitigations with external partners, such as IT contractors, postal agencies, printing firms, labour providers, administrative data providers, or any other key stakeholder in the proper conduct of a census.

2. Emergency management, governance and decision-making

2.118. A census will have a large number of "moving parts" subject to events not controlled or even anticipated by the census taker. Natural disasters are a good example. Although census managers understand that natural disasters can affect their operations, there is no predicting when, where or to what extent. Even with the best risk management and mitigation plan in place, a census is too complex to anticipate every potential issue. As a result, the statistical agency must be prepared to respond to the unexpected (such as the COVID-19 pandemic during the 2020 round). With proper emergency management, governance and decision-making, censuses can weather difficult storms effectively.

2.119. Every census should have a process in place to identify, classify and respond to an inevitable emergency. Key members of the census team should know how to respond to emergencies during crucial phases of their census. They need to be able to determine if an emergency is underway, how to classify the severity of the emergency, who to contact when it occurs and how contact can be made, so that solutions can be put into place.

2.120. Census team members must be equipped with the information necessary to detect new issues and to respond effectively using communication lines to other team members, senior officials and perhaps the general public. For example, field-based censuses are making greater use of internet questionnaires and real time metrics which can be monitored to ensure the proper functionality of that service to citizens. As well, for both field-based and register-based census, massive databases require constant monitoring to detect anomalies and cyberthreats. Should there be an issue detected during such monitoring, the

person viewing that information must be prepared and competent enough to interpret what is happening, classify the degree of the emergency and communicate to others to establish a resolution plan.

2.121. When any emergency is detected, it should be assessed for its severity. Census team members responsible for the component of the census impacted by the event should be trained on various severity definitions during the planning phase of the census. For example, when an event impacting the internet response service is detected, the information should be shared with the census manager who is responsible for that service, and who is trained to determine the emergency's severity level by deciding, for example, whether the event being detected is creating a slowdown of the service that might be of minor annoyance to the user, or is a broader interruption of internet response, such as a cyberattack.

2.122. The severity level of the emergency as determined by the responsible manager should trigger a pre-determined response. For example, a high severity event could require a meeting of the full census management team within a short period (no later than one hour for example) after detection, and require faster resolution, relative to a low severity event. At that meeting, the situation should be described by the responsible manager, who should then present information on the cause of the emergency and any identified remedies. If possible, there should be an estimate of when impacted census operations will be fully operational again, which will be particularly valuable for communications and any down-stream activity that may be subject to a subsequent impact.

2.123. To detect and respond to an emergency, a census should have an Operations Centre where employees have access to tools that track relevant operational data. When a census is underway, a number of operational metrics are available to staff, who should be presented with those data in real-time and with an historical archive to reference as well. Prior to the execution of a census, those data can be specified in advance, using expertise from previous cycles. The resulting "Management Information System" (MIS) will gather information from various systems and present them in a way that allows managers to detect and respond to incidents.

2.124. In the event of an emergency, census managers should be prepared to communicate within their organizations. As mentioned earlier, priority should be taken to communicate first with the census team, to ensure a timely response. However, internal communication to senior leaders in the organization and political staff may be required, depending upon the severity of the incident. Also, should they be implicated in the response, or impacted by the emergency, external contractors may need to be notified.

2.125. In addition to providing tools for employees tracking MIS systems for signals of trouble, the physical Operations Centre can also be a place to meet to discuss the issues and respond. The Operations Centre should be a place where communication crises are managed and should provide access to facilities and equipment to record video and develop other content to be broadcast on the internet and other media. To monitor any ongoing situations and detect any new ones, the facility should have access to television, radio and internet broadcasts.

2.126. It is recommended that the Operations Centre be both a physical and virtual space. Should a physical space not be available (which was the case for many countries during the COVID-19 emergency), staff should have all the same tools that exist in the Operations Centre, available to them while working remotely. From a location away from the census office, the fundamental requirements for staff should be to have access to MIS and other reports and have an ability to communicate via phone and video calls. Any capacity to respond to an emergency should not be impeded by the closure of a census management office location.

3. Disinformation and misinformation threats to the census

2.127. Censuses face a new threat of dis/misinformation, which is unlikely to go away anytime soon. Exploiting weakness in Internet literacy and mistrust in government, sophisticated actors can launch campaigns to divert attention away from a census or embarrass governments by disrupting operations. Censuses now need to invest in tools to monitor social media, the most common way that dis/misinformation campaigns are conducted.

2.128. Specialized "social listening" tools can comb through large databases of publicly available social media posts to find incorrect information being shared about censuses. When detected, a statistical agency can decide to intervene as appropriate with responsive communication tactics, to correct the record and re-establish the facts regarding the census. Not all situations will be the same, and some responses will require close working relationships with major social media companies. If a major dis/misinformation campaign is detected, the statistical office should be able to quickly contact the platform(s) on which the incorrect information is circulating, and request that posts be removed before more damage is inflicted on the census. Other interventions may be necessary, such as countering the offending posts with information in the social media accounts of the statistical agency.

2.129. While powerful and effective, social listening tools must be used carefully, however. While reviewing only publicly available information, there is a risk that government census takers could be perceived as invading privacy. As a result, before these tools are used, privacy experts should be consulted and offered an opportunity to approve or reject the proposed social listening.

4. Crisis response training for census staff

2.130. While it is impossible to anticipate all possible scenarios, key census staff should be prepared for all significant risks. After establishing a risk register and emergency management protocols, prior to the enumeration of a census, census teams should practice what they might encounter and how the issue would be mitigated. These practice sessions can take various forms and can be adapted for different types or levels of census staff. For example, one might be targeted toward IT staff, while another might be for the management team.

2.131. These exercises are an effective method to condition census staff on emergency management. The exercise would involve staff members who should be made aware of all (generic) emergency management protocols already developed. With that knowledge in hand, participants to the exercise will be introduced to various scenarios that should be based on potential issues identified in the risk register. When the scenarios are presented, staff should reflect on the best next steps that should be taken. Everyone at the sessions should be expected to contribute and provide feedback on the reflections of other team members to the scenarios.

2.132. During the conduct of practice exercises, team members may find themselves uncomfortable. This is part of the nature of the exercises – the whole point should be to get used to very difficult situations. As a result, participants will find themselves unsure of what to suggest as solutions to the situations they are practicing. Furthermore, these practices sessions should include scenarios which are progressively more difficult to deal with. It will be impossible to imagine and practice all scenarios that census managers might find themselves in. However, they should be realistic and designed by

knowledgeable individuals who use the previously developed risk register. The sessions should not seek to develop an ability within the team to know all steps to follow in all emergencies, but instead the sessions should develop a more general ability within the census team to respond quickly and effectively during the execution of a census.

5. Ensuring fast workflow during an emergency

2.133. Census teams understand the importance of their work and respond effectively during emergencies. However, without proper structure and management, some census teams may feel compelled to do "whatever it takes" to ensure success despite setbacks. It is important to note that techniques beyond hard work and extended hours can be employed in emergency situations to ensure the completion of the work without causing unnecessary negative effects on the people involved.

2.134. One likely requirement of a census emergency will be fast workflow and strong communication to ensure the responsible teams are aware of what is changing, so they can adapt as necessary. During a census emergency, solutions must be found quickly and acted upon appropriately. As described above, once an issue is detected a quick response will be required.

2.135. The census manager in the statistical organization has an obligation to orchestrate the census team such that the resolution to the problem is efficient and effective. Management techniques can be applied which ensure fast workflow, such as Kanban boards[51] and quick, frequent "stand up" or "sprint" meetings. Management techniques such as these should be applied in order to make sure the full census team is working on the correct priorities and that these new priorities are communicated to the affected team members.

2.136. One way that census managers can ensure fast workflow and project visibility, is to create virtual and real Kanban boards during an emergency. The Kanban technique will have the team develop and act upon problem statements describing the response to an emergency. For example, if there has been an attempted cyber-attack a problem statement might declare an issue to be resolved, such as "Re-open internet portal" and a team is assigned to resolve the issue. There might be related issues created by the cyberattack, for the communication team or the downstream operations. After the creation of the various issue statements, they can be placed into a "to-do" list or Kanban board, which provides a visual representation on the work created by the emergency. The census manager can meet with the team regularly, perhaps in the form of quick sprint meetings, to review the Kanban board and make sure the work to re-establish regular operations is moving at a fast pace. As well, the Kanban board is an effective way for other members of the team to see if there are any implications of the work of another team, on their area of responsibility.

6. Contingency planning for main plan failure

2.137. Despite the NSO's best efforts, censuses can fail to reach their full objectives. Disruptions to the execution of a census are inevitable and may severely impact the expected data quality. When risks are looming, a census taker can respond in a number of ways. The best scenario is that the risk is fully mitigated such that it does not develop into an issue. Another situation might require a census taker to

[51] Kanban boards are a visual project management tool used to track and manage work. They are based on the Kanban method, which originated in Toyota's production system.

cancel or pause a census, as was the case for many countries in the 2020 round, impacted by the COVID-19 pandemic. However, if a census risk cannot be fully mitigated or avoided, the issues that are created may not always be manageable. So much can be beyond the control of the census taker, and with censuses becoming increasingly complex (as, for example, with combined censuses or those using large amounts of administrative data from multiple sources), there is always the risk that the main plan fails. What does a statistical agency do in that case?

2.138. Unfortunately, the options may be limited. For countries conducting field-based censuses, there may be an option to accelerate plans to use administrative data. While the available administrative data may not be suitable to replace all planned data variables and on the original timelines, national statistical agencies may be able to perform the basic function of a census with an administrative population count. If the risk of main plan failure is high enough, a national statistical agency might consider a parallel administrative count, to ensure the delivery of population counts in any scenario.

2.139. While it may not be considered as such, a robust administrative data plan for field-based census takers can be an effective "insurance plan" and therefore an incentive for investment in administrative data, technology to manage it, and expertise within the census team on how to use administrative data properly for census taking.

2.140. For register-based censuses, a disruption of the census would likely involve a communication plan and an adjustment of the timeline of the census, to allow a re-start of the process when the disrupting event has passed. If this is the case, there would be an impact on other programs dependent upon the census, such as an extension or temporary introduction of intercensal population estimates.

C. Change management

2.141. In each census round, when planning for a new population and housing census begins, various innovations derived from technological developments, lessons learned from the previous census, and the ongoing need to improve the balance between budget and data quality are evaluated. When a new census involves significant conceptual changes, from de jure to de facto enumeration for example, or technological changes, such as transitioning from paper questionnaires to electronic forms, or to a combination of collection methods, or methodological changes from a field-based to a register-based census, change management becomes essential to ensure that these innovations are implemented effectively, and the desired results are achieved.

2.142. Change management is grounded in a set of principles, approaches, and strategies designed to facilitate the successful transition of an organization or project from its current state to a desired one. There are various models of change management and its components. One option that a census might consider, is the model presented by the Harvard Business School:[52]

- Prepare the organization for change;
- Develop a vision and associated plan;
- Implement the vision and plan;
- Embed the changes within the culture of the organization; and
- Review progress and analyse results.

[52] Harvard Business School Online (2020, March 19), 5 Critical Steps in the Change Management Process. https://online.hbs.edu/blog/post/change-management-process

2.143. To achieve this, it is crucial to understand the reasons for the change and assess the prevailing conditions, challenges, and opportunities, to properly prepare the organization for change and to develop a vision and plan that is achievable. For instance, when transitioning from a paper-based census to one that is more electronic, it is imperative that staff appreciate the comprehensive benefits and impacts on the census process. Therefore, it is essential that senior management actively supports and effectively communicates the change vision and is actively engaged in the implementation process. A high-level "vision" paper prepared by census management, describing the proposed changes and their reasons can be an effective reference for the census team executing upon the vision. The vision should be circulated and discussed, so that employee insights and concerns should be taken into consideration, and they should be provided with training in the new skills and knowledge necessary to adapt to the change.

2.144. In order to implement the envisioned changes, detailed planning is essential, encompassing clear objectives, timelines, responsibilities, allocated resources, and risk mitigation strategies, as well as other elements described in the previous sections on resource mobilization, project management and emergency preparedness/contingency planning. This plan should also maintain flexibility to address any deviations that may occur.

2.145. Ideally, the need for change from one census cycle to the next is sufficiently well communicated to ensure that it becomes part of the culture of an organization. That is, the driver for change is so well understood by the census team that they put complete focus on the new change element, and it changes the way they think and behave. For example, imagine a field-based census whose past cycles have provided insufficient respondent support and burden management. The change may involve considerable new investment in respondent materials or services, and changes to questionnaire content. With a major re-focus towards making a census friendlier for the population, the result may be a major change in employee perspective, shifting it from an NSO-centric and results-oriented organization towards one which *also* expects a high level of service to the population and where staff are highly focused on the new service elements of the next census.

2.146. During the execution of a census, it is essential to stay on-scope with the original vision. An important part of change management is to ensure that the census team does not deviate from the original plan. To do so, census management must keep tight control over the implemented change. The original vision should be associated with a series of "baselined" assumptions and activities, any deviations from which there should be a required "change request". For example, if in the previous paragraph's scenario there was a baseline assumption that there be a reduction in the amount of time it takes to complete a questionnaire, any deviation from this plan that holds constant or increases questionnaire time, can only occur if a request for change is approved by census management. Not only does this approach manage scope creep[53], but it can also reinforce the staff cultural elements described above.

2.147. Finally, when important change is executed for a census, time should be taken to re-evaluate how the change management process worked. That evaluation is best considered at the end of a census cycle, when other elements are being reviewed in preparation for the next cycle. Improvements to the change management approach should be treated like other lessons learned of a cycle.

[53] Scope creep in project management is continuous or uncontrolled growth in a project's scope, generally experienced after the project begins.

IX. Administrative organization

A. Overall overview

2.148. In planning the organization and administration of a census, particularly for a field-based census, it is important to consider the role and relationship of the various executive and advisory organs. National, subnational and local commissions and committees are frequently useful in the planning and preparations of a census. Such bodies may be composed of representatives of governmental agencies, community leaders with due representation to all sections of society, and non-governmental users of the census data, particularly those involved in policy-oriented analysis of census results and analytical studies of the social, economic and demographic situation of the country. This ensures broad-based and complete participation of the people to enable proper canvassing of sensitive issues such as ethnicity, gender, disability, migration, and marginalized groups. It is important, however, that their advisory and promotional functions be clearly defined and that the final responsibility for planning and execution rest with the executive agency.

2.149. There are well documented and proven advantages in having an office continuously responsible for census work established as an integral part of the statistical system of a country. Such an office assures continuity in census work and is the principal centre for the formulation of the programme and the initiation of preparatory work for the next census. Its permanence permits the development of specialized and experienced personnel and the maintenance of statistical and cartographic information, including cross-cutting issues such as information technology, essential for planning the next census.

2.150. At the pre-enumeration stage, the census office will need to be expanded to form the nucleus of the full census organization, which must be capable of directing the field organization during the preparatory work as well as during the enumeration and processing. In order to provide immediate supervision in each area, field offices at various levels are needed for the later part of the preparatory work, including staff recruitment and training, as well as for the enumeration period. Supervisory personnel in such offices should be persons who, being familiar with the particular area and the local language, are able to deal with local problems. This does not mean, however, that all supervisory positions need necessarily be filled by persons from the area. Personnel may be transferred from the central office or from other areas as the need arises. Prior training of all such available personnel is necessary, so that they have a working knowledge of all aspects of the current census programme. An essential part of the preparatory work is local administrative planning, which would set out the likely problems and challenges specific to the local area and how they are to be dealt with. Adequate coordination with local public authorities is always important so that the enumeration work is not interrupted due to other activities.

2.151. Subsequent to the enumeration, the census organization is usually readjusted to meet the needs involved in compiling, evaluating, analysing and publishing the results and to provide the continuity desirable for promoting the continued use of census materials. Census organizations need to pay special attention to continuity of knowledge and skills from one census to the next, since the intervening gap, which is usually a whole decade, is likely to cause loss of institutional memory and attrition of qualified personnel. Comprehensive documentation of census activities while they are being carried out is essential, as is training of younger personnel to create a pool of knowledgeable and experienced persons by the time the next census comes.

B. Roles and responsibilities of key stakeholders in the census process

2.152. Population and housing censuses are complex programs that require the involvement of several stakeholders. The roles and responsibilities of key stakeholders in the various phases of the census should be clarified and adapted to the local situation. For example, while roles and responsibilities of the government are at the central level, these can also be defined for sub-national governmental authorities.

2.153. The role of each of the key stakeholders can be summarised as follows:

Government

- Issue a proclamation or decree for the census with an adequate lead time for preparatory activities.
- Ensure allocation of adequate resources (financial, human, physical, etc) for the census implementation.
- Assure a conducive environment for the census undertaking.
- Reinforce the statistical purpose of the census and assure against politicization or other influences that can compromise the census.
- Contribute to the identification and prioritization of data gaps to be addressed through the census.
- Ensure census data serve a public good for improving population wellbeing.
- Provide administrative data for use in producing or supporting the census.
- Provide security services to protect data during collection, transfer, and storage. Additionally, implement security measures to safeguard personnel and physical assets involved in field operations.

National statistical organizations

- Ensure effective intragovernmental coordination for the successful implementation of the census, particularly if conducting a register-based census.
- Ensure consultations and effective engagement with various census stakeholders throughout the census process.
- Lead the development of technical resources for strengthening capacity of the census personnel in all phases of the census.
- Ensure high quality, international standards for the census, including the selection of the most appropriate enumeration method(s) for the country context.
- Promote effective dissemination and use of the census data, including public data access.
- Ensure the safety and security of census data, as well as for employees.

The United Nations and other international organizations

- Coordinate the development and promote adoption of international principles and standards for the census.
- Provide technical assistance for the census through National Statistical Offices.
- Promote the census as a public good for improving population wellbeing.
- Contribute to the identification and prioritization of data gaps to be addressed through the census.

Donors

- Provide financial and technical assistance for the census within the scope agreed with the government.
- Promote learning and capacity strengthening for census undertaking, including through South-to-South-and Triangular Cooperation.
- Contribute to the identification and prioritization of data gaps to be addressed through the census.

The private sector

- Provide financial and technical assistance for the census within the scope agreed with the government.
- Promote the adoption of technology for the census process in accordance with industry best standards.
- Contribute to the identification and prioritization of data gaps to be addressed through the census.
- Provide support services and goods for the conduct of a census.
- Provide administrative data for use in producing or supporting the census.

Academic and research institutions

- Contribute to the identification and prioritization of data gaps to be addressed through the census.
- Provide administrative data for use in producing or supporting the census.
- Promote the dissemination and use of census data.

The media

- Contribute to the identification and prioritization of data gaps to be addressed through the census.
- Promote public education on the value and purpose of the census.
- Promote the dissemination and use of census, including through thematic analyses on dimensions of sustainable development.
- Support public mobilization for participation in the census.

Civil society and other interest-group organizations

- Contribute to the identification and prioritization of data gaps to be addressed through the census.
- Promote public education on the value and purpose of the census.
- Advocate for the responsible and effective use of census data to enhance population well-being. Support public mobilization for participation in the census.
- Provide administrative data for use in producing or supporting the census.

The general public

- Contribute to the identification and prioritization of data gaps to be addressed through the census.
- Provide complete and accurate responses to the census questionnaire(s).
- Demand accountability for the use of census data to improve population wellbeing.
- Provide feedback to NSOs on the acceptability of the content and perceived trust in the organization for the management of their data.

C. Statistical leadership

2.154. The period of preparation for the population and housing censuses represents a unique opportunity to exercise leadership in promoting the use of statistics in overall development of societies, with the focus on improving service delivery and policy development. Statistical leadership needs to be built and should rely on international standards and guidelines, such as the Fundamental Principles of Official Statistics[54], on national statistical legislation and on the national statistical code of ethics.

2.155. An important role of the head of the national statistical office or statistical agency responsible for census undertaking is to ensure a successful census programme that delivers results for use in evidence-based policy development, evaluation and research, and decision-making. It is therefore recommended that heads of national statistical offices or statistical agencies responsible for census undertaking drive the following activities:

(a) Facilitate the establishment of the statistical legislation that mandates the undertaking of a census;

(b) Setting policy and strategy by defining targeted outputs and outcomes for the programme;

(c) Strategic engagement with stakeholders by mobilizing participation across government, business and the public at large;

(d) Raising the profile of and commitment to use statistical information, providing an opportunity for engagement on key policy issues and strengthening relationships between information providers, policymakers and opinion leaders;

(e) Adhering to the Fundamental Principles of Official Statistics and ensuring best practice is embedded in statistical procedures;

(f) Aligning to international standards and frameworks;

(g) Establishing statistical infrastructure and resources for undertaking the census;

(h) Setting up the census management project structure.

2.156. Preparing and conducting a population and housing census provides a unique opportunity to exercise statistical leadership, promoting the use of official statistics for development at all levels of society. Given the decennial nature of most censuses, it is essential to maximize this opportunity to build statistical leadership, grounded in international standards and national legislation.

X. User consultation, communication and publicity

2.157. The extent of public consultation, communication and publicity can vary enormously according to the type of census being undertaken. Most of this section relates to activities associated with a full field enumeration or combined census with a field operations component. Countries conducting fully register-based censuses that require little or no direct participation by the general public should implement a communication and publicity strategy with messaging that aligns with the register-based approach.[55]

2.158. A comprehensive programme of communications and engagement for a population and housing census covers three distinct audiences: (a) major users of census data, (b) persons and institutions

[54] United Nations General Assembly. (2014). Fundamental Principles of Official Statistics. Resolution 68/261. New York: United Nations.

[55] Link to UNECE guidance on Communications for RBC countries.

participating in the census operations and (c) the general public, including particular population sub-groups. Since the field-based census is a national activity that is completely dependent for its success upon the cooperation and support of the general public and many governmental and local organizations, the entire communications and engagement effort should be developed as a coordinated activity in close coordination with the other substantive preparations for the census. These communications and engagement activities are valuable not only for informing stakeholders about the census but also for providing census authorities with early and continuing information about the reactions to census plans and activities of the general public in various parts of the country and of key population groups and institutions. Many national statistical offices use design thinking approaches to create better respondent experiences. This involves developing user personas to segment population groups and better understand differing needs, expectations and behaviours.

Consultation with data users

2.159. Extensive consultation with users and advocacy groups with an interest in census data on topics, on definitions and, particularly, on planned tabulations and other outputs and the development of the census database is an indispensable step in the preparations for the census, which should be conducted early. These consultations will assist the census authorities in planning for a census that, within the resources available, is as responsive as possible to user needs in terms of the collection, processing, tabulation, storage and availability of meaningful statistics. Such consultations can also serve to foster a wider and more informed understanding of, and support for, census plans and activities. The users to be consulted should be from governmental departments, ministries, universities and other research institutions, the private sector, and other organizations (or individuals) representing the economic, social, educational, commercial and cultural life of a country. Many countries will want to include among the groups to be consulted organizations or bodies representing culturally diverse communities, ethnic and indigenous communities, refugee and internally displaced community groups, and religious groups, those with accessibility requirements and people with disabilities, housing associations and agencies supporting homeless people. Other key stakeholders may include partners with whom the census office collaborates for the provision of specialist services, and donors who may help fund elements of the census operation.

2.160. Taking into account the importance of the census in providing data for local planning and administration, it is also often advisable to have consultations with users in provincial and local governments and institutions in various parts of the country, particularly in remote rural areas. In large countries or countries where the provincial or local governments have a comparatively high degree of autonomy, consultation with users at the subnational level is essential if the full potential of the census is to be achieved. Strategies should be chosen according to the target group.

2.161. The consultation process can take many forms. If done in the form of meetings, it is often more useful to hold separate consultations with different types of users with common interests, such as administrators, policymakers, planners, demographers, academic researchers, users in the business community and so forth, rather than conduct a simultaneous consultation with all data users. Consultations involving different types of stakeholders in the same setting frequently prove frustrating to participants because there are substantial opposing differences among users in their data requirements, technical background and in their concern with the details of census content and operations.

2.162. Meeting data users in person is very informative but imposes physical and budgetary limitations. Broad consultation can be conducted online and via social media, e-newsletters, and digital stakeholder outreach. The strategy can be used both to identify users' requirements and also to provide transparency

in the census preparation activities. Other forms of technology can be considered for holding decentralized or remote consultations. Users may be sent an electronic questionnaire to collect their priority data requirements, or invited to complete the questions online. Digital channels, including social media engagement, offer the opportunity to reach a substantial number of users if accounts have adequate reach. Video content may be employed to support or promote consultation materials, and webinars can be an effective way to meet with large numbers of users remotely.

Communication with stakeholders

2.163. In the context of conducting a census, it is essential to establish two-way relationships with key stakeholders at the lowest territorial units such as municipalities. These relationships will facilitate continuous communication and awareness throughout the entire implementation and development process of the census. Additionally, a meticulously crafted message tailored to the social context of the region should be developed to effectively introduce the census to various interest groups. This lays the foundation for ensuring the entry of field personnel into information sources, which is essential for successful awareness-building and data collection. This tailored message is crucial for effectively introducing the census to various interest groups, and ultimately enhancing awareness and data collection efforts.

2.164. Once these relationships have been established and the census operation has been effectively presented, it becomes necessary to focus on structured dissemination of census information. This becomes a pivotal element in mobilizing society towards the production, understanding, and utilization of statistical information that will benefit the country. To achieve this purpose, efforts to raise awareness must be strengthened, aiming to prevent disinterest and disengagement and ensuring the acquisition of information with the required coverage and quality. In this way, the objectives of collecting complete, accurate and valuable data for the benefit of all can be achieved.

2.165. In order to complete the preparatory work for a census involving a field enumeration and to carry out the enumeration itself, the census office will have to expand its staff substantially. In addition, numerous governmental and non-governmental organizations outside the census office may be called upon to provide personnel, expertise, equipment, supplies, workspace, transportation or communications facilities and so on to help in the census work. As a result, large numbers of temporary personnel will have to be recruited and trained (see paragraphs 2.192–2.202) and the contributions of a diverse group of national and local organizations will have to be effectively mobilized. A well-planned recruitment campaign and programme of engagement with external organizations can contribute to both efforts.

Communication with field personnel

2.166. The training and capacity development of field personnel play a pivotal role in the census process and necessitate effective communication, much like the communication with various data users. To achieve this, it is imperative that data collectors comprehend three key aspects. Firstly, they must grasp the significance of statistical information production for the nation and the importance of Censuses, along with a comprehensive understanding of the context in which field operations will take place. Secondly, those responsible for data collection must receive guidance, conveyed through communication, emphasizing the importance of a commitment to data collection quality. Lastly, communication must ensure that census enumerators are well-versed in (a) the content of the census questionnaires, (b) general statistical operation concepts, and, specifically, those related to cartography, and (c) the appropriate use of mobile data capture devices in countries employing such data collection technology.

Therefore, it is necessary to prepare a field data collection guidebook with sufficient teaching materials and targeted teaching strategies.

Public outreach and information campaigns

2.167. An effective communications strategy, together with far-reaching publicity and information campaigns, play an essential role in ensuring the success of the census. This is especially so for those countries adopting a field enumeration methodology, either wholly or in part, where the general public is expected or required to participate actively in the census activities as respondents. Particularly in the case of countries that undertake a significant field operation, public acceptance and cooperation is essential to ensure the success of the census. It is crucial to underscore the vital importance of securing active participation from each data source or qualified informant in the census process. This translates into an imperative need to avoid definitive refusals to ensure the acquisition of information with the required coverage and quality. In specific cases where conventional communication channels may not be accessible, establishing a more personalized form of doorstep communication with the individuals to be enumerated becomes essential. The aim is to ensure the presence of qualified informants when delivering the questionnaire. It is important to differentiate between a recruitment campaign to hire staff for both office-based and field-based census roles, and a general campaign to encourage participation.

2.168. As part of the communications strategy, countries should consider both proactive and reactive media management as well as planning engagements across multiple channels. The use of a call centre to deliver help and support via telephone and providing chat support for online enquires should be considered. Consideration should also be given to the appointment and training of census spokespeople to communicate key messages consistently in local media and through community engagement activities.

2.169. While there may be occasions when door-to-door communication is necessary, a large-scale publicity and information campaign is recommended to inform the population about the census and to explain its purpose. Opportunities may be taken throughout the campaign to monitor feedback from the general public, particularly through social media, both to help shape future messaging and where applicable to inform the census operation. Designing and implementing the publicity programme is best undertaken by professional experts in the field of public relations, advertising and sociology. Such expertise is frequently not found within the national statistical office itself, and it may therefore be appropriate to outsource some or all of this work. The publicity programme may include:

(a) The development of a census slogan and logo
(b) A public relations campaign;
(c) A community liaison (or outreach) programme;
(d) An advertising campaign;
(e) Monitoring of public opinion;
(f) Media relations, including monitoring of the mass media;
(g) A social media strategy.

Key messages for public campaigns

2.170. There are several main messages that census agencies will need to communicate to the public in order to optimize outcomes for the census. Census publicity campaigns should encompass a wider set of messages, whose components might include:

(a) making the public aware of the census;

(b) educating the public about the benefits (to them and to the country) of the census, in order to improve sentiment towards participation;
(c) reminding people about their legal obligation and duty to take part in the census and the penalties for refusing to do so;
(d) mitigate the risks of digital exclusion by explaining to the public what to do and when, and providing information on the practical support available to those who may face barriers to responding;
(e) informing and reassuring the public that privacy and confidentiality will be protected; and,
(f) expressing thanks to the public for taking part in the census.

2.171. Care is necessary in finding the correct balance between these different messages. For example, an overemphasis on the obligatory nature of the census may serve to reinforce negative perceptions that the census is an imposition by the State on the population, rather than an activity for the common good.

Objectives of public campaigns

2.172. Publicity for a census operation entails an educational campaign, the purpose of which is to ensure the awareness, interest and cooperation of the general public, particularly population groups that may be reluctant to be enumerated. The aims, as a general rule, are not only to dispel any anxiety regarding the purposes of the census but also to explain the reasons for the various questions in the questionnaire, including the value census information provides by contributing to an informed data driven society, and to offer some guidance on how questions should be answered. The publicity campaign may also be an important tool for increasing the completeness of census coverage, particularly among groups with lower coverage rates in previous censuses. Planning for the general publicity campaign should start as soon as the census is authorized. The campaign itself should be closely synchronized with other census activities and full-scale publicity should not begin too far in advance of the date on which enumeration is scheduled to start. However, plans for the publicity programme should be formulated early enough to take effect for the purposes of any census tests. In addition, the programme can use these tests to study the impact of alternative publicity materials and methods. If either the cartographic or house-listing operations require extensive fieldwork and widespread contacts with the public, it should be recognized that personnel involved in these activities often provide the public with its first impression of the census. Training and publicity programmes should take this into account.

Strategies for reaching specific population groups

2.173. To the extent that the census communications budget allows, the general campaign should be directed to all sections of the country and all segments of the population through the use of all available publicity media, with special emphasis on the use of digital media channels. The general campaign may be supplemented by a number of specialized campaigns aimed at specific segments of the population to raise awareness on specific subjects, such as gender, migration, ethnicity and disability, in which the quality of response may depend on the level of prior awareness among the public at large or among the specific groups concerned. In multilingual countries, creating campaigns in the various languages used within the country (both official and non-official) is crucial to ensure all communities have the opportunity to understand and participate in the census. Publicity can be secured at low or no cost through social media and through public relations with national and local mass media, such as newspapers, television and radio. It is recommended that a range of spokespeople from the census organization are trained and made available to speak to national and local media, and as well as for community engagement activities. Paid-for advertising through various forms of media can also be effective. Interactive media may also be

used such as a toll-free helpline and text messaging. Local events, conducted in local languages, where the public can participate can also improve public awareness and build trust.

2.174. Disseminating information about the rationale of the census and its utility helps allay possible misconceptions among the general public, thus increasing participation and coverage. Outreach campaigns involving a range of organizations and enlisting the support of local leaders and opinion makers to spread the word about the census in their area of influence are also good strategies. Maintaining ongoing relationships with community organisations and representatives in the period between censuses can be of mutual benefit, with the potential to improve engagement not only during the census operation but also during the consultation and dissemination phases. In addition to recruiting such organizations and leaders as partners, the census organization may develop key messages, web content and links, printed material and other material to support them in their activities. The use of publicity may also be considered to support the recruitment of field personnel.

Branding and inclusiveness

2.175. For countries running a field-based census operation, it is recommended that a census "brand" be established, including a logo and tag line. Census branding is important for establishing awareness and trust. A simple but effective slogan and distinct logo can be used in all national and local advertising campaigns and in all types of media, booklets, posters, brochures and other promotional materials. A slogan and logo that are well recognized from the initial stages of the publicity campaign may serve to improve "brand recognition" for the census[56]. The aim should be to encourage the respondent to feel more reassured that the census is an inclusive and beneficial activity.

2.176. Special attention should be given to making communications as inclusive as possible in order to improve the quality of responses across the country and across different population groups, as well as to bolster trust in the resulting data. In essence, the aim of the communication approach is to ensure members of the public are made aware of the census, their obligation to take part, and how to get support if required. Inclusivity should be considered in all aspects of the development and delivery of the communications approach. An important first step is to identify population groups or communities that may be less likely to take part in the census without additional interactions. These population groups will vary from country to country, but examples could include transient groups, recent migrants, young people, older people, people living with disabilities, the digitally excluded, the homeless, people living with literacy and language difficulties, and inhabitants of high-rise apartments and dense urban areas. Having identified these population groups, census organizations should consider research and insight-gathering to understand the reasons that may prevent communities from taking part and design strategies to address this. An approach that has been successful in many countries is for census organizations to speak to organizations and charities that support the interests of these communities. The reasons are likely to be able to be classified as motivational or practical.

2.177. In rural areas, particularly where digital channels are limited, weekly markets, fairs and public festivals are a good opportunity to publicize the census message among people who may not have much exposure to mass media. An excellent opportunity exists to create widespread awareness of the census through a campaign targeted at schools. These campaigns can be targeted to areas where high

[56] United Nations Economic Commission for Europe. Conference of European Statisticians Recommendations for the 2010 Censuses of Population and Housing. New York and Geneva, 2006.

proportions of people face barriers to taking part. Census organizations can develop free education resources (in multiple languages, as appropriate) which focus on educating pupils on the census, the use of census data to shape services locally and nationally, and the importance of everyone taking part. They can also include take-home materials and activities providing children with the opportunity to tell family and guardians what they have learned. These campaigns have proven to increase awareness of the census among parents in some countries. Other kinds of local-level publicity, such as wall writing and village announcements, can be planned according to local circumstances.

Communication for digital-first census

2.178. For countries taking a digital-first approach to collection, three additional priorities will inform communications and engagement planning across the census operation. Firstly, a website must be built and tested. Design of the website should be informed by user research to ensure users have a straightforward experience in accessing and submitting the electronic questionnaire. The website must be equipped to securely host very large numbers of users concurrently. Secondly, codes permitting access to the electronic questionnaire must be supplied to all households and other residential establishments, with accompanying messaging to facilitate understanding of when and how to complete the electronic questionnaire, and to provide reassurance about the security and privacy of their responses. The general campaign should support these messages.

Communication for changing census processes

2.179. Where there are significant process changes to the census operation such as the introduction of an online response, the communications campaign should be tailored to reflect such changes. The changes to a census should be communicated to the public well in advance of the census campaign through a variety of traditional and social media channels in an accessible manner that reaches all sectors of society. The communication should address a range of issues including the benefits of moving to an online census for both the respondent experience and the quality of the data; the reduced impact on the environment; increased the efficiency of the data collection operation. Clear information on the sequencing of collection activities from initial communication (e.g., via publicity, mailout, enumerator) of details on how to access an online or paper form, clear instructions on how to complete an online return, to receive support in completing an online return or how to request a paper census form. The public should also receive assurances about the security of their data completed using the online mode. Additionally, it is essential to provide clear guidance on how to handle technical difficulties that may arise during the online census process, including information on available support channels and alternative methods for data collection. Communication strategies should also address changes beyond data collection methods, such as the introduction of new census topics or modifications to the questionnaire due to resource constraints.

Support for completing the questionnaire

2.180. Support should be made available to anyone who may face barriers to completing the questionnaire, regardless of the reason. Providing detailed explanations on how to fill out the questionnaire is crucial for both field-based and online respondents. This can help prepare respondents for the enumeration process and reduce potential difficulties. Materials supplied to households should clarify the various routes available to individuals (e.g., completion online, by telephone, or the option to

request a paper questionnaire). Websites containing frequently asked questions, or chatbot[57] services, can be a cost-effective way to answer common questions from respondents in the first instance. Additionally, census organizations may wish to consider providing tailored support such as webchat services, social media, SMS, telephone advisors, or email. Consideration should be given, where appropriate, to alternative language options for all such resources. In countries where censuses are primarily conducted online, in-person support centres may assist people without internet access. All guidance should be updated in response to emerging themes and decisions before and during the live census operation. Real-time management information (MI) from online responses (and slower MI from paper questionnaires) can be used to deploy the field force more flexibly and responsively towards non-responding households. It is advised that the field force be provided with guidance in responding to common questions and issues, to support households to respond.

Managing misinformation and incidents

2.181. Census organizations should develop a framework for managing misinformation and other incidents that may impact on census operations (see paras 2.127-129 for disinformation and misinformation threats to the census). Communications staff should be included in strategic discussions during operations to support the organizational response to issues as they arise. During operations public opinion and mass media should be monitored to assess the effectiveness of publicity campaigns and identify such issues. Public opinion could be monitored through surveys that can provide information on public attitudes to the census, or through interaction on digital channels including social media. Media monitoring refers to tracking, analysing, and measuring various media channels. These channels can include both traditional and digital media. Analysing the data from these sources helps NSOs understand public perception and identify emerging challenges related to the census. Such monitoring implies an ongoing accumulation of information, detection and prevention of the development of negative published comments on the census, and preparation of adequate responses to negative reports and information. Increasingly the media has a significant influence on people's behaviour and even minor distractions and mistruths can have a detrimental effect on the outcome of the census. Therefore, in developing their publicity campaigns, NSOs should give particular attention to preparing for unexpected events (such as negative attitudes, malicious lobbying, technical difficulties, delays and misleading information).

2.182. 'Lines to take' should be developed and updated to respond to emerging issues, and may be used on social media directly in response to misinformation on such platforms. Direct communication with social media companies and establishing procedure to work with these companies, as well as points of contact, is also encouraged in advance of a census operation, to develop processes for identifying and managing misinformation. Public relations with mass media can also be used to challenge or alert the public to misinformation. It is also recommended that all official participants involved in census operations know their roles in the communication process both with the media and with the public at large.

Post-census communication

[57] A chatbot is an advanced computer program that offers automated responses to frequently asked questions, while the live chat connects users with an agent for assistance when the chatbot cannot address their specific query. These features aim to improve and supplement the existing services provided to citizens by creating a positive, secure online experience that assist respondents with their census questions. The chatbot provides faster response times by answering simple FAQs quickly and automatically while reducing the burden on help desk staff. Extensive usability tests should be conducted to assess the quality and effectiveness of the chatbot and live chat options.

2.183. An integral part of census communication and publicity is informing key census data users and the general public about the availability of the census results and their utility (see paragraphs 1.21–1.60). Awareness about the available census data and products should be raised during the intercensal period before the commencement of the next census. This is to make sure that the public recognizes the importance of the census and appreciate statistics that are generated from it. It is critical that such communication strategies be developed as an integral part of census planning and not left as an optional add-on. It has been the experience of quite a few countries that the engagement of professional media and communication personnel adds value to the campaign.

XI. Census calendar

2.184. An indispensable element in the planning of a census is a calendar or timetable indicating the sequence and estimated duration of each of the component operations of the census. At the early stages of census planning, a provisional calendar of selected key dates should be prepared as an overall framework for the census. The calendar should be shared with stakeholders in advance for advice and support. The calendar will, inevitably, be revised and made more detailed as planning proceeds, with the aim of establishing final dates as soon as practicable. However, adherence to major milestones and deadlines is crucial, and any necessary changes in plans must be communicated transparently to the public, including clear explanations for the adjustments.

2.185. Such calendars are essential, since they indicate the dates on which each of the numerous tasks and activities that make up a census are to be started and completed, and they serve as a guide for measuring the progress of each stage of the census operation. Some census program managers may refer to a census calendar's main elements as the "Census Critical Path", or a series of milestone dates are essential to meet and which significantly impact other elements of the program. Serious delays in work, or errors in time estimates, can be detected by comparing the calendar target dates with the actual dates of each activity. A census calendar is a very efficient instrument not only in the timing of each census phase but also in the control of all those census operations that are interdependent. Therefore, when modifications in the census timetable become necessary, all inter-related operations should be taken into consideration in order to avoid disruptions to the whole census programme. Obviously, the time schedule will differ for each national census depending upon the general census plan and the resources that are available.

2.186. For censuses conducted with a field-based enumeration, the census calendar usually shows the various operations grouped into three broad sectors: (a) pre-enumeration, (b) enumeration and (c) post-enumeration. The last-named sector includes evaluation and analysis as well as the data processing and dissemination activities. The basic date on which the census calendar and the scheduling of all other operations hinge is the starting date for the general enumeration of the population. For purposes of control, many operations that in fact overlap are shown separately in the calendar. Census calendars sometimes take the form of a chart or graph, in addition to a detailed checklist of operations. Project management software may help in the preparation of the census calendar.

2.187. In establishing the census calendar, it is necessary to consider the relationship of the population and housing censuses to one another as well as to other statistical projects or other large-scale national activities. Although a joint population and housing census operation is likely to constitute, for the period of its duration, the major statistical undertaking of the government, care should be taken that it does not

interfere unduly with the other regular national statistical activities that may be going on at the same time. A balanced statistical programme should avoid having too many simultaneous competing enquiries, which might place too heavy a burden on the statistical services and on the public, with a possible resultant loss of both administrative efficiency and public cooperation. Other conflicting events that should be avoided include national, local, or regional elections, public holidays, and any other major events in the country that could distract the general population.

2.188. It is often useful to draw up a comprehensive diagram showing the sequence, interrelationship and timing of all the various steps in the census programme – a Gantt chart and a critical path would be a good example.[58] This type of analysis often reveals the consequences of a delay at one step in terms of delays at other steps in the programme. It can therefore be a useful instrument against which the actual progress of the census preparations may be compared. Indeed, some countries have attempted to use such critical path analyses not only as an aid to census planning but also as a tool for the ongoing management of their census operations. In these instances, it is essential to establish procedures for revising the critical path analysis in response to actual progress. It should be stressed, moreover, that the usefulness of such devices depends on how soundly they are designed, applied and understood.

2.189. Project management software can be useful in linking the diagrammatic structure of census operations with information about nodes or centres of responsibility for individual broad or detailed operations so as to control the chain of responsibility. Alternatively, event calendars can provide a broad view of the steps of the census programme and allow follow-up. Different tools can be found on the Internet for download or online use. Online versions allow immediate update and make it easier for staff to work in a group but are dependent on Internet access. Other tools, commonly referred to as groupware and collaboration software, as well as Internet and social media forums, can support census operations by providing an environment for exchange of information, files, and data among dispersed teams. A clear monitoring and evaluation plan of the calendar should be in place and regularly reviewed.

2.190. For countries transitioning from a field-based census with paper questionnaires to one with electronic questionnaires, the census calendar should take into account the need for a sufficient period in the pre-census stage of system development. This is because the process of designing and testing the collection systems is more complex and lengthier than that of paper-based questionnaires. Therefore, it is important to allocate enough time for this stage to ensure a smooth transition to electronic questionnaires. In addition to the added time needed for electronic questionnaire design and testing, there are other aspects of electronic data collection that require time, such as the implementation of robust data security measures and the evaluation, selection and procurement of the necessary technology. Finally, electronic questionnaire use may require legal evaluation, added communication to the public and increased training for staff also using the platform.

2.191. The census calendar must also allocate sufficient time for the preparation of geospatial information that will support the census operation. The preparation of up-to-date census maps is a critical and time-consuming process that requires careful planning and execution. Sufficient time allocated for map preparation ensures the accuracy, completeness, and effectiveness of the census process, leading to reliable and actionable data for decision-making.

[58] The Gantt chart was developed around 1910 by Henry Gantt of the United States, based on the work of Karol Adamiecki of Poland. It is a type of bar chart that illustrates a project schedule and is available in a number of office software packages.

XII. Human resources management

2.192. Field-based census taking requires a large number of people to function properly. Early arrangements and good human resource information systems are necessary to secure the proper number and type of personnel required for each of the various census operations. While the preparatory and processing work generally calls for office employees possessing, or able to learn, certain specialized skills (cartographers, coders, data entry operators, programmers and so on), in a field-based census the enumeration stage usually demands a large number of persons capable of going to their assigned urban or rural enumeration areas and collecting the information according to specific definitions and instructions. The number of enumerators required being quite high compared to normal staff strengths, and the period for which their services are needed being rather short, demands that the method of recruiting them needs to be worked out carefully in advance to facilitate quick, simultaneous and transparent recruiting, and subsequently remunerating them and relieving them of their duties promptly and efficiently. Good human resource management of temporary census staff can benefit other programs within the statistical organization or subsequent censuses, as they are already trained and can sometimes be retained for non-census programs and may form part of the next census team.

2.193. Human resource information systems can greatly assist the management of the large temporary workforce of a full field enumeration census. These systems can help streamline and optimize recruitment and deployment of new staff and can help identify personnel with relevant skills who can be cross trained for additional tasks, or for redeployment for other non-census activities.

2.194. Consideration should be given to computer skills if electronic means of enumeration are going to be used. It is essential that the enumerators and, to the extent possible, their immediate supervisors be conversant with the languages or dialects of the area in which they will be working, and the knowledge of these languages should extend beyond any official languages in which the nation provides services. In addition, attention should be paid to efficiency, economy, experience, ability to read maps and communication skills in general. It is only prudent to recruit and train sufficient reserves of people for a variety of temporary census positions, to take care of any attrition that may occur in the process.

2.195. Once the cartographic preparations are substantially complete and the questionnaire has been sent for printing, if applicable, perhaps the single most important means that the census authorities have for influencing the success of the census is its training programme. The contribution that a well-planned and executed training programme can make to the quality of the census results cannot be stressed too strongly. Such a training programme must of course focus on the widely dispersed and difficult-to-supervise field staff (namely, the enumerators and their immediate supervisors) but it must also cover others (for example, the higher-level supervisors, editors, coders and computer operators).

2.196. Giving office employees who are working with the census preparations a brief, uniform basic training of all aspects of the census has two prime advantages: firstly, all personnel can understand the importance and the context of their part of the task; and secondly, since they are conversant with the basics, they can be more efficiently deployed in the field for supervision or coordination during the actual census operations whenever and wherever needed. The nature of the training can be traditional (classroom style), e-learning (self-training or live training via internet) or combination of these approaches.

2.197. The entire census training programme should comprise different courses that are designed to cover each phase of the work and provide an efficient and consistent means of effectively equipping large numbers of fresh employees with the necessary skills. The programme will need to correspond closely to the needs of the various operations and, where appropriate, may include both theoretical and practical instruction, with emphasis on the latter. In the case of the enumerators and their immediate supervisors, the training is most effective if it includes several opportunities for the trainees to participate in practice interviews and role-playing exercises, including the use of adopted IT solutions, if any. In countries in which multiple languages are used, the method and content of the enumerator training programme will need to be suitably adjusted. For example, if the questionnaire is printed in another language, provision will have to be made for instructing enumerators on the correct formulation of the census questions in the vernacular. In countries with large geographies or a variety of regions, it may be necessary to prepare training material for local areas, which is specific to the issues that are faced in those regions.

2.198. Enumerators and supervisors should be trained as close to the field operations as possible so as to avoid recall lapses. This leaves very limited time for conducting the training. Therefore, the logistics need to be worked out carefully in advance. The training programme for editors, coders, operators of data recording equipment and so forth should also provide opportunities for the trainees to practise under the supervision of the trainers. The intermediate- and higher-level technical staff, such as programmers and system analysts, should also be given special training with emphasis on recent technical developments of relevance to the forthcoming census and on the interrelationships among the various aspects of census plans and operations. Thorough training in census practices is an extremely important component of quality assurance. Detailed and clear documentation of instructions with appropriate illustrations is a basic requirement in this regard. A proper training methodology and a variety of training aids will go a long way in enhancing the training effort.

2.199. The organization and conduct of training courses should be entrusted to those having the necessary qualifications to carry out this task successfully, taking into account not only their professional abilities but also their ability to teach. This means that staff in charge of training should have certain qualifications that will enable them to stimulate the interest of trainees and to transfer the required knowledge, since otherwise well-qualified technical personnel who are unable to transfer their knowledge to the trainees in a satisfactory manner will be unsuitable as instructors for group training activities. This must be taken into consideration when selecting instructors and it is recommended that objective criteria should be used. In practice, however, it is difficult to find the necessary number of instructors who have both the professional and the teaching qualifications; for this reason, any outsourced instructors selected should themselves undergo training in how to organize and conduct training courses. The use of professionally designed training guides can add immense value to the training effort. The involvement of experienced professional experts in the design and delivery of training programmes is also very useful. It should however be noted that the content should be the responsibility of the census authorities and not that of outsourced experts.

2.200. It is important that training manuals for each training programme are made available to the census organizers and training instructors. Such a manual would be a valuable guide and would help considerably in the efficient training of census staff. It would also contribute to the uniformity of training, which is an essential factor for a successful enumeration, taking into account the great number of census instructors who will be engaged in training. Simple audiovisual aids (for example videos, posters, audio recordings) can also be used to help make the training more effective and uniform throughout the country. If available, new multimedia technologies can facilitate the provision of training at distant

locations and be effective and efficient supplementary tools for training. Standardized training may also be provided in e-learning format on the Internet and on handheld devices.

2.201. It is very important to determine the time required to train staff for the various aspects of the census. This depends on several factors, such as the task for which they are being trained, the complexity of the content, the educational level of trainees, the number of instructors available and the funds available. Apart from fixing the number of days for training, it is also important to allocate appropriate time for each subject. A balance is to be made between the allocation of sufficient time and the cost of training. Drawing up lesson plans for each session of training is an effective way of ensuring that all subjects are covered, with the right amount of time being devoted to each.

2.202. Register-based censuses may have a human resource challenge very different from the scale-related issues of managing a large workforce needed in field-based censuses. Register-based censuses can involve relatively few people with specialized skills. Therefore, knowledge retention and succession planning, as well as the need for good documentation of processes, as well as specialized training, may be a greater challenge for countries with this type of census. Enumerator training should also include how to handle distrustful respondents, data confidentiality and privacy.

XIII. Logistics management

2.203. A field-based population and housing census differs in many respects from other statistical operations. It requires efficient communication between many different components, including the procurement operation and storage of a large variety of items, most of which have to be distributed to all geographic areas of the country and then recollected.

2.204. Logistics management is a process of planning, implementing and controlling the flow of census materials and equipment needed for implementation of census operations. Logistics planning requires careful coordination between different phases of the census operation, such as mapping, training, field enumeration, data processing and dissemination. The scope of the logistics programme usually differs from one country to another, but mostly covers the following activities: (a) renting central and field offices; (b) installing furniture and equipment; (c) providing help desk support; and (d) delivering and collecting all census materials, including manuals, questionnaires and publicity materials.

2.205. National statistical offices may need to establish a special team for planning, implementation and controlling the logistics programme. The functions of this team should be clearly determined in order to avoid overlap or omission of any activity. During the planning phase of the logistics programme, outsourcing of some activities should be carefully examined as an option.[59] In the context of census logistics, procurement plays a particularly important role throughout the whole exercise.

2.206. The complexity of logistics management can be facilitated by the introduction of a logistics management information technology system. Such a system can be used to manage and monitor logistics and provide necessary information to the logistics management team on the availability of supplies, their distribution to field operators and the return of questionnaires and materials after operations have concluded.

[59] For a detailed discussion on outsourcing of census activities, see paragraphs 2.220-2.235.

A. Procurement management

2.207. Developing a strategic approach to procurement is another particular element for successful implementation of a field-based census operation. Taking into consideration the complexity of the process, procurement planning requires logistical coordination with multiple census activities and counterparts. Proper planning contributes to efficient procurement processes, and reduces the risk of confronting problems that may lead to additional costs and delays. Procurement planning is a complicated exercise and emergency work and last-minute operations are usually unavoidable. However, the benefits of procurement planning early in the census operation usually outweigh the disadvantages.

2.208. In the context of censuses, procurement planning entails the process of assessing and projecting the procurement needs of census operations. Needs assessment, cost estimation and requirement definition are the first steps in the procurement process, and are essential components in procurement planning. The purpose of the requirement definition is to identify the precise needs of the census operation and to search for the best solution to meet those needs. The needs must be described in the requirement definition in a way that will facilitate the procurement process. The requirement definition is often done in parallel with supplier sourcing and market research in order to let information from the market research influence the requirement definition.

2.209. Procurement practices vary greatly among countries; therefore, no universal system of procurement management can be suggested. However, a few generally accepted procurement principles can be noted. First is the principle of *best value for money*. This represents an optimal combination of technical and financial attributes – that is, the balance between price and performance that provides greatest overall benefit under the specified selection criteria. This does not necessarily mean selecting the lowest initial price option, but rather represents the best return on the investment, following a proper evaluation of offers under appropriate criteria contained in the solicitation documents. It requires an integrated assessment of technical, commercial, organizational and pricing factors in light of their relative importance. Best value for money can include non-cost factors such as fitness for purpose, quality, service and support, as well as cost-related factors such as price, life cycle costs and transaction costs associated with acquiring, using, holding, maintaining and disposing of the goods or services. The principle of best value for money should be applied throughout the procurement process in order to attract the offer that most effectively meets the stated requirements of the census operation.

2.210. Second is the principle of *effective competition*. Effective competition is best explained as a situation in which at least three independent contractors acting on their own (that is, not in collusion with each other) effectively compete for the same business opportunity and each submit a responsive bid. The procurement processes should foster effective competition as a means of ensuring fairness, integrity, transparency and achieving best value for money. The competitive process should, as necessary, include: (a) procurement planning for developing an overall procurement strategy; (b) market research for identifying potential suppliers; (c) consideration of prudent commercial practices and applicable national regulations, rules and procedures relating to procurement; and (d) formal methods of solicitation, utilizing invitations to bid or requests for proposals on the basis of advertisement or direct solicitation of invited suppliers; or informal methods of solicitation, such as requests for quotations.

2.211. Another important principle is *fairness*. The manner in which the procurement process is carried out must give all stakeholders the assurance that the process is fair. The concept of fairness includes that the procurement process should be free from favouritism, self-interest or preference in judgment. The

assurance of a fair process promotes transparency, a principle that ensures that timely information about existing conditions, decisions and actions relating to procurement activities and about procurement policies, procedures, opportunities and processes are clearly defined and made known simultaneously to all interested parties. A transparent system has clear rules and mechanisms to ensure compliance with those rules. A transparent system of procurement further ensures that procurement records are open, as appropriate, to inspection by auditors. In line with the procurement principles of transparency, every step in the procurement process should be documented and kept on file, preferably electronically as well as in hard copy.

2.212. For register-based or combined censuses, the procurement and intake of administrative data will be required. Relationships with administrative data providers will need to be established and formalized through agreements. Often, administrative data providers will require compensation for expenses incurred during the process of providing the data to the national statistical organization and therefore financial agreements and arrangements to settle funds will be required. While the large majority of administrative data will be help by public sector organizations, it is possible that some administrative data may be held by multiple organizations within the private sector and (depending upon national procurement rules) require the national statistical organization to tender bids to ensure a fair price is arranged, and the fairness principle is applied.

B. Forward and reverse logistics

2.213. The type of census materials would differ depending on census methodologies and technologies used for enumeration and data processing. However, in a full field enumeration census, or for the field enumeration element of a combined census, any kind of materials related to the fieldwork has to be supplied to the field staff and returned. Strategies for distribution and return of the materials should be carefully planned according to the type of materials, volume and final destination of delivery.

2.214. As a first stage in this process, decisions must be made concerning the nature and responsibilities of the centrally controlled distribution and return operation. For example, a decision should be made about the geographic levels to which the materials will be distributed – regional office, local census committee, supervisors or other. These decisions must be made keeping in mind the impact of the amounts of material to be transported, the transport facilities available to field staff and the condition or existence of roads or other means of transport. Once these decisions have been made, the key inputs to the dispatch and return of materials are as follows:

(a) Workload estimates from the mapping programme to establish packing volumes for transport requirements; and

(b) Name and address details for delivery and pickup points.

2.215. During the design of enumeration areas and mapping activities, an estimate will be made of the number of enumeration areas, and the amount of work in each. This information can be used to calculate how much material will be needed by each enumerator, supervisor, manager and so on. This method should provide a more accurate estimate of the total volume of all the materials.

2.216. The majority of these tasks are usually carried out under contract by a government transport service or commercial operator. The contractor will use specifications and consignment details provided by the national statistical offices. The postal service may also be a feasible method of distribution.

2.217. A significant task in planning field operations is establishing the specifications for the packing and transport of materials. These specifications need to be developed regardless of whether these activities are carried out by the census agency itself or contracted out to another government agency or private company.

2.218. The role of the national statistical authority with regard to dispatch and return tasks is primarily one of liaison and monitoring. For the most part, the contractor will contact regional managers and supervisors directly about the delivery or pickup of material. The national statistical authority can expect to be involved as a liaison between the contractor and field staff in the early stages of the operation or if there are any particular problems experienced by either field staff or the contractor. The national statistical authority management staff should meet frequently with the contractor to discuss the operation and liaison arrangements. Part of the planning of the operation will include arrangements to enable the national statistical authority's management staff to monitor the delivery and return of materials. In particular, when material is picked up from supervisors, the national statistical authority's management staff should maintain a close watch over what is taking place in the field as the transport of completed census forms is involved.

2.219. Distribution and collection of census materials can be monitored through use of a management information system (see paragraphs 3.203–3.207). Timely information can be produced about the progress in delivering the census materials and questionnaires and collecting the completed questionnaires and the other return materials. Depending on the logistics programme, the information needed for monitoring the logistics activities will differ. In general, the following information would be needed for both forward and reverse delivery: (a) type of materials; (b) timing of delivery; (c) number of delivered materials; and (d) name of the persons involved in the delivery. The periodic reports produced from the management information system will be crucial to ensure the timeliness of the fieldwork across the country by giving an alert if there is any delay or any other problems regarding the delivery of census materials.

XIV. Contracting out

2.220. It is a contemporary practice in many countries to contract out some of the tasks or activities of the population and housing census as a way of increasing efficiency by utilizing the advanced methods and technologies not necessarily available within the national statistical office or other public sector organisation responsible for conducting the census. At the same time, costs reductions may be achieved through a competitive selection process. However, not all census tasks are appropriate for outsourcing or contracting out, and doing so will not necessarily bring the desired benefit of strengthening national capacities. Census activities may be broadly classified as core and non-core activities, which some may define as those outside of the critical path of the census. As a general rule of thumb, core activities should not be contracted out. If for some reason core activities need to be contracted out, then it is essential that the strategic control of such activities should remain firmly within the census authorities at all times.

2.221. In the context of contracting out components of census operations, the national statistical authority would need to build the capacity to ensure proper outsourcing. This is of primary importance at the preparatory stages, as outsourcing requires a solid and comprehensive knowledge of contemporary technologies and their advantages and disadvantages, as well as past experiences at home or in other countries. Consequently, the national statistical authority would need to plan and develop a particular

unit for the purpose of ensuring adequate and efficient outsourcing well in advance of the census itself, as there would need to be extensive testing of the products and services that were contracted out.

2.222. The terms of engagement (scope of work), the deliverables and the timelines should be clearly laid down with definite dispute redresser mechanisms. Illustrative examples of items of work that may be contracted out in the context of a full field enumeration census are as follows:

(a) Layout and printing of census questionnaires;
(b) Development of electronic questionnaire systems;
(c) Packaging of census questionnaires;
(d) Dispatch and delivery of census material;
(e) Census mapping;
(f) Publicity and public relations;
(g) Training;
(h) Return collection of census questionnaires and other material;
(i) Inventory and storage of filled-in questionnaires;
(j) Scanning and data entry;
(k) Data processing and tabulation (also relevant for administrative censuses); and
(l) Publication and dissemination (also relevant for administrative censuses).

2.223. Countries may find it advantageous to have the same organization take care of more than one of the steps above. That is, the list is not intended to suggest that each element to be contracted to a single, separate organization.

2.224. Time is of the essence in all these activities and it is vital that adequate time is allocated. At the same time, backup plans should be in place in order to deal with any failure on the part of the vendors. Fundamentally, census operations are time critical and commercial compensation is secondary. Depending on whether an activity is on the critical path or not, adequate flags should be provided. Milestones and timelines are also essential. The moment there is a failure in achieving any milestone, alerts should automatically be raised. Risk assessment represents a critical component for outsourcing; the risk of failure, and the costs involved in developing contingencies in case of failure, require particular consideration.

2.225. The appropriateness of contracting out should be determined step by step and after subdividing the overall census tasks into stages. In the context of quality management, the outsourcing of components of census operations still requires the national statistical office to take full responsibility for, and manage the quality of, the census data. Throughout the overall process, activities should be conducted by a method (considering accuracy and timeliness of the results) that can best satisfy the general public. No part of the work tasks should be done by a method that may result in loss of trust of the general public. When outsourcing, the statistical office needs to ensure that it continues to be in a position to understand and manage elements that contribute to final data quality. So, in judging the propriety of contracting out, it is recommended that national statistical office should carefully consider the following criteria:

(a) Strict protection of data confidentiality;
(b) Method of confidentiality assurance that satisfies the general public;
(c) Guaranteed measures of quality assurance;
(d) Ability to manage and monitor the outsourced census tasks or activities;
(e) Ability of the contractor to stay within the planned budget; and
(f) Control over the core competence of the national statistical office, and appropriateness of judgement, considering the specific situation of each country.

2.226. Confidentiality assurance is the first and most important issue that should be considered by national statistical offices. National statistical offices are responsible for data confidentiality, in terms of both perception and reality. It is extremely high risk for national statistical offices to have to manage leakage or misuse of confidential information. Consequently, contracting out of tasks that have the risk of such an incidence should be avoided. For example, in the phase of data collection, it is highly recommended that contracting out should be avoided because the task is closely related to the earning of trust from citizens and the strict protection of confidentiality. Where temporary enumeration staff are engaged under contract, this should be done in such a way that they are subject to strict measures of monitoring and control by the national statistical office. These enumeration staff should be engaged in such a way that their activities are governed by the relevant statistical legislation to preserve the confidentiality of the data they collect.

2.227. The second important and related issue that should be considered carefully is conveying confidentiality assurance to the general public. As described in the "Essential roles of the census" (see paragraphs 1.1–1.3), a census should be undertaken by the method that can produce the most reliable results and in a manner that ensures the trust of the general public in terms of both perception and reality. If either one of these attributes is not met, then the method used as well as the results obtained may not meet the approval of users and the general public and could result in the census itself being questioned. Thus, protecting data confidentiality refers not just to the actual protection of confidential data, but also to protecting the perception of confidentiality among the general public and providing a sense of corporate security.

2.228. The third significant issue to be considered in outsourcing is the guarantee of quality assurance in the outsourcing environment. The key point is that the national statistical office should be satisfied that the goods or services paid for are provided to the requisite standard. Cost should not be the first priority in considering and judging the successful bidder in this respect unless prescribed by procurement rules. Although it is desirable to engage in fair competition among several companies to reduce costs, it is worth mentioning that merely considering low-price bidding as a determinant factor may adversely affect the quality of the job to be done by the successful bidder. Low-quality work could cause a significant loss of trust among the general public. To assess the quality of work, as part of the contract allocation process, potential contractors should be required to provide samples of their work (for example, for printing, manufacturing satchels, and other work), or if this is not possible, to list referees who could be contacted to verify their claims or allow sites at which previous work has been carried out to be inspected. The contracting process should state all the key requirements for the services sought and bidders should be measured against these. Although not a guarantee of quality, it will minimize surprises. Once the contract has been awarded, continuous monitoring of the progress of work entrusted to the selected company is necessary and the national statistical office should ensure that a system for monitoring quality is built into the contract. Consequently, in considering the proper contracting-out procedures, national statistical offices should also take into account the costs for constructing a system of surveillance for monitoring progress of the work being contracted out.

2.229. In addition to monitoring the providers of goods and services, national statistical offices need to plan for continuous interaction with vendors. This implies an additional step to monitoring and amounts to a necessity to work side by side on a regular basis in order to ensure the best quality of the products and services and to meet the standards and needs of the census operations. This coordinated work refers to providing technical and technological advice, as well as following the development of the services and applications from the substantive point of view. While the national statistical office may not have the full

capability to develop certain products or applications, it certainly possesses considerable technical experience and understanding of producing statistics on a regular basis. Therefore, planning and implementing for a regular and continuous interface with providers when parts of the operations are outsourced needs to be incorporated in the overall planning from the beginning.

2.230. The fourth major issue in outsourcing census activities is the procedure of assessment and evaluation of the capabilities of the candidate providers. A quality assurance framework (for a detailed discussion on quality assurance, see Part IV) and implementation should be established in a first phase of outsourcing. Through this procedure national statistical offices should fully assess both the capabilities and the limitations of companies in order to select the winner to which the activities in question are to be outsourced. It is highly recommended that practical and financial particularities of providers should be considered after the assessment of their capabilities. Any private company has a potential risk of bankruptcy or of changing the field of its activity. It should be kept in mind that if a selected company is unable to fulfil the assigned tasks, any consequent problems might not be resolved by applying penalties. However, a very significant problem that could occur is that users might not be able to make use of accurate and timely census results. In such a case, national statistical offices might lose the trust of stakeholders in the census and even in future censuses or other routine statistical projects conducted by the statistical office.

2.231. Some approaches to outsourcing put an emphasis on a “turnkey” arrangement, by which contractors deliver the system according to a set of predetermined client specifications with the expectation that the client focuses solely on the outputs and not the internal working of the system. This assumes that the national statistical office completely understands and can fully anticipate all data quality issues that might arise during the census and has included these in the specifications. The client is not expected to have any understanding of how these systems work or how they might contribute to the final outputs. Any changes to the system typically require cumbersome processes to determine contractual responsibilities and heavy financial costs. This sort of approach is not recommended as it effectively hands over the quality of the census data to the contractor, while the risks associated with intervention remain with the census agency. It removes any flexibility and greatly restricts the ability of the census agency to react to quality problems that emerge during processing.

2.232. Suppliers should be made fully aware of the quality targets at the outset of the census programme, and the quality requirements of the outsourced components that enable the overall census quality target to be achieved. Operational quality control should apply to outsourced services in the same way as those that are not outsourced. For some contracted-out operations (such as the printing of questionnaires), national statistical offices may want to consider having staff on-site, to quickly judge and correct any quality issues.

2.233. In addition to managing outsourced activities or tasks, the ability to adapt to sudden or unforeseen changes is crucial. While outsourcing can offer potential benefits, it's essential to recognize that it doesn't automatically lead to cost savings. The overhead associated with monitoring contractors, addressing unforeseen challenges, and managing other complexities can offset expected cost reductions. As a result, it is recommended that national statistical offices should carefully assess whether outsourcing specific census activities aligns with their overall goals and capacity. This evaluation should consider factors such as cost-effectiveness, control, and the ability to manage potential risks.

2.234. It is also recommended that for critical activities, such as the coding of education, occupation and industrial classification, special care should be taken to ensure adequate training of the personnel to

undertake the task, particularly when it is contracted out. The same amount of care and training is required even when the task is performed by the national statistical office's own staff. This is due to the fact that the coding depends on the minor differentiation and level of coding (general and detailed classifications according to different coding standards), as well as the coding manual and the educational level of the coders as well as the effectiveness of the training provided. However difficult, complete coding manuals should be prepared and provided to the contractor in advance. If coding is automated, working in advance with the vendor may be necessary to ensure that automated coding methods are thoroughly tested in advance of the conduct of census field operations.

2.235. Censuses are large operations that collect massive quantities of data that require coding and editing. To reduce the staff resources required and to improve timeliness, uniformity and accuracy, automated coding procedures may be employed. Some countries have already implemented automated coding procedures for recording addresses, countries, education levels, occupation and industry. The development of the application software could be contracted out although the rules to be followed must be carefully specified by the national statistical office, which should retain responsibility for implementing the system. The software application can often be used for other statistical collections undertaken by the national statistical office. When outsourcing, the staff of the national statistical office should be able to modify the parameters of such operations themselves at little cost and in a timely manner. By having this ability, the national statistical office can manage the appropriate balance between data quality, cost and timeliness.

XV. Use of technology

2.236. Technological advances and expanded access to modern technology have enabled significant streamlining of the way in which the business of a population census is conducted in all its phases. With modern tools, every stage of the census—from planning and monitoring to implementation and evaluation—can be seamlessly integrated. Many facets of census activities can benefit from the use of technology. For example, remote sensing and imaging technology can generate maps with coordinates against which enumeration activities can be tracked. Fieldworkers can be paid their salaries and stipends using mobile device technology. Handheld devices equipped with editing functions and global navigation satellite system can improve the consistency of responses while tracking geo-activities, including those that can report on spatial coverage. Given the breadth and extensive capabilities of today's technology, it is crucial to select carefully at the planning stage the technological innovation elements that will be adopted in the census value chain.

2.237. The comprehensive integration of geospatial technology at every relevant stage of the census process has the potential to bring about a transformative shift in census planning, data collection, analysis and dissemination. In the planning phase, geographical information systems (GIS) can aid in the delineation of enumeration areas, ensuring that no areas are overlooked or double-counted. This precise demarcation facilitates the efficient allocation of resources and fieldworkers. During data collection through direct enumeration in the field, geospatial technology enables real-time tracking of enumerators, ensuring their safety and optimizing their routes for efficiency. The geolocation data can be cross-referenced with responses to validate the data collected and to ensure full coverage, thereby enhancing the accuracy of the data and aiding in real-time monitoring and evaluation of census activities. In the data analysis phase, the overlay of census data onto geospatial maps can allow for a deeper understanding of demographic patterns, population densities, and other socio-economic indicators in a spatial context,

allowing integration with other geo-coded datasets even from other sectors and providing richer dissemination products for data users.

2.238. Rapid innovation has led to both exponential growth and reduction in costs of technology. These changes have seen an increased adoption of technology across many aspects of the census in a variety of different ways. While on the one hand this is providing considerable benefit, on the other it is increasing dependence on technology providers and introducing new challenges and risks, such as the use of safe, secure and trustworthy artificial intelligence. The key to the successful use of technology in a census is to clearly understand the rationale or the objective of the technology introduction and consider a range of key success factors for technology adoption, which may include suitability, security, scalability, stability, safety and skills.

2.239. Understanding of the true value of the technology is critical to formulating a business case and assessing whether or not to proceed with the project, and if so what technology choices to make. As the introduction of technology can be an expensive and risky exercise, it is important to ensure that there is sufficient value in its introduction for each specific census, and that the most suitable solution is selected.

2.240. The most common reasons for the introduction of technology in censuses operations are:

- **Efficiency and reduction of costs.** Technology provides an opportunity to reduce the number of personnel involved in different aspects of the census. For example, scanning and character recognition can reduce manual data entry and increase data accuracy; satellite imagery can reduce manual mapping; and Internet self-response can reduce fieldwork. It can potentially reduce respondent burden, shorten interview times and enhance communication with respondents. Technology can also reduce other non-labour expenses such as printing, freight and travel. In some cases, technology can simplify business processes and thus reduce cost or risk. For example, fieldworker workflows can be simplified through the use of secure mobile technology that allows census takers to submit their availability, receive assignments and navigation instructions, collect and transmit data, and communicate with their supervisors via their mobile devices.
- **Data quality and consistency.** Technology, and in particular the automation of processes, can increase the consistency of census data and reduce data errors – for example, scanning and character recognition reduces data entry errors, and automated validation rules or edits ensure that data are checked and changed in consistent ways rather than relying on manual, dispersed field processes.
- **Timeliness.** Technology can reduce the time needed to carry out a field-based census enumeration, data processing, data analysis and preparation of results for publishing. The faster the census data are released, the more valuable the data are to census users, and thus the use of technology for data scanning, recognition, processing and publishing should be considered for its ability to advance publish dates.
- **Public expectations and confidentiality.** The census relies on the general public to provide data. Both public and user expectations, and in some cases legal requirements, may make it advisable or necessary to implement specific technologies to support and ensure their convenience of use and confidentiality. protection features. This can manifest in various ways, such as offering online census questionnaires, facilitating online job applications for census roles, or providing census results digitally. Notably, there is some evidence that the provision of online questionnaires has had a positive impact on census response rates.
- **Data integration and dissemination.** For census statistics to hold value, they must be accessible and usable by a diverse group of data users, and, wherever possible, meet the

principles of findability, accessibility, interoperability, and reusability (FAIR). This includes everyone from the general public, private sector and media to government officials, civil society and researchers, spanning a variety of levels of expertise and data needs. By leveraging various technologies, including GIS, the collection and dissemination processes can be enhanced. This allows for seamless integration of census data with other sectors, offers data aggregation and visualization capabilities, and ensures the availability of data in diverse formats tailored to user needs.

- **Event management.** The coordination and monitoring of a census enumeration across the complete nation or area has always been challenging and has relied on dispersed accountability and manual processes. Technology provides better visibility, oversight and capacity to monitor performance indicators and respond to enumeration events.
- **Data retention and utility.** The costs of conducting a census are significant and thus attempts should be made to optimize the value achieved from the census results. The use of technology enables census data to be safely managed, secured and retained, as well as opening up opportunities for statistical data integration and data reuse, such as for time series analyses and other analytical needs.
- **Assurance and anticorruption.** The adoption of technology can reduce the risk of fraud or corruption through providing more standardized, controlled and auditable records of actions, for example for the records of financial expenditure.

2.241. Considering the reasons listed above in a business case for a technology project often reveals inherent tensions and necessitates trade-offs that should be considered openly and transparently. For example, a solution that aims to fulfil a broader range of user expectations might entail higher costs. While the approach may benefit user satisfaction, it may negatively impact overall programme efficiency. It is possible to consider various tiers of technology use and functionalities and adjust the scope of technology to meet as many goals as appropriate (step approach).

2.242. While the conduct of significant technology projects is becoming increasingly commonplace across government, there is limited evidence to suggest that the mere introduction of new technology alone would deliver the projects on time and within budget. Careful planning and management are critical to the success of these projects. The unique size, timing and nature of a census throw up unique challenges, which need careful consideration. The following factors should be taken into consideration in the context of optimizing the use of contemporary technology for census operations:

- **Suitability.** The functions and benefits of any new technology need to be assessed thoroughly and objectively. Consideration of whether or not the technology has adequate user-friendliness for the intended user is another important point.
- **Scalability.** Technology solution should be designed, implemented and tested to be able to handle the number of users and the number of data items. Hardware and software used in delivering technology must be scalable to the actual load of each business process during the census project.
- **Security.** The security of census data is critical for guaranteeing the confidentiality of respondents' personal information. Census systems, especially an online census, must be secured to protect privacy of respondent data and to ensure respondent confidence in the system. Systems need to preserve confidentiality, integrity and availability of census information. IT security is especially important in a contemporary census, as cybersecurity risks have greatly increased in recent years. National statistical offices with an existing or planned digital presence should be prepared to invest properly, to minimize all risk of data loss to cyberthreats.

- **Stability.** A field-based census relies on the successful conduct of an enumeration in a tight time period. Key pieces of technology – such as the census call centre, self-response portal or self-help website – may be designed in such a way that they do not have an unrecoverable impact on data response rates, particularly in countries using self-response specially during close to census day.
- **Skill.** The census organization, whether planning to outsource or deliver the solution internally, must ensure that it has the adequate skill, knowledge and capacity to lead and manage this project.

2.243. Census technology projects require strong project management expertise and thus adequately competent, experienced, motivated and knowledgeable staff need to be assigned to this role.

2.244. Detailed business requirements (meaning what the technology needs to do and how it needs to do it) need to be developed by a suitably skilled business analyst in order to ensure that the technical specifications and implementation are aligned with the census need.

2.245. Rigorous testing and piloting are essential for the statistical or census office when integrating new technology into its processes. It is crucial to understand the various stages impacted by technological changes. Separate tests should be conducted to prove the efficacy of new technologies and to identify potential problems linked to their implementation. Depending on the extent and characteristics of IT, these tests should include all information technology components related to the fieldwork, data transfer or entry and processing well ahead of the actual census. Testing and piloting should evaluate the application systems (software) and the equipment (hardware) under real-world conditions, so as to identify and mitigate any potential malfunctions. Continuity planning should also be integrated into the testing phase, ensuring that there are fallback measures in place to counteract any equipment or system failures that could disrupt a smooth census operation.

2.246. A pilot census ensures that technology can be tested as part of a complete, end-to-end business process and highlight major implementations with functionality or capability. The pilot test should be accompanied by activities to test whether technology solutions are scalable, secure, accessible and robust enough for the real census. Stress tests should be conducted during the testing phase to ensure that the use of such technology can handle the maximum load of each module or business case and at the same time sustain data integrity.

2.247. Once the implementation of technology for the census is completed, a comprehensive evaluation should be conducted. This evaluation serves a dual purpose: to identify lessons learned and define areas for improvement to enhance future census operations and maximize the value of the technology. During this phase, the results will be analysed, and the efficiency and accuracy of the technology will be assessed, taking into account the costs and benefits associated with the implementation. Feedback from enumerators and citizens who participated in the process will be collected and considered. This will help identify specific issues that require solutions and contribute to the continuous improvement of census operations in the future.

2.248. There is an increasing adoption of technology in every national statistical office across the world. A number of nations have been developing their own technology solutions to support different aspects of the census, with some of these products being made available for statistical organizations in other countries. When making decisions around technology selection or development approaches there should be strong, favourable consideration of existing products within the international census community, and

where products do not already exist the preference should be to undertake co-development with other members of the statistics community. The exchange of best practices and lessons learned is key to streamlining these processes.

2.249. There are, however, a number of risks or drawbacks associated with the introduction of technology that should be considered and managed, otherwise these could lead to increases in expenditure, delays to the census timetable or impacts on the quality of the census data. These include:

- Incompatibility or other integration issues between different hardware and software applications;
- Solution outage or failure (which could be for a variety of reasons – lack of connectivity, hardware failure, battery life, GPS black spots, software bugs, device theft);
- Lack of skills or knowledge among system users, particularly temporary census staff;
- Insufficient or inadequate communication between technology staff and business staff, particularly leading to misunderstanding of requirements and specifications;
- Hacking, online attack or other information technology security events;
- Maintaining, upgrading or decommissioning old or legacy systems;
- Lack of documentation or over-reliance on a small number of key people;
- Huge amount of digital data available, creating a potential distraction for staff.

2.250. A wide range of technologies, covering all stages of the census from planning to data dissemination, is presented throughout all the chapters of these principles and recommendations. However, the integration of technologies in census operation requires taking into account various specific national needs and the value each technology would bring to a specific aspect of the census.

2.251. The utilization of technology is expected to vary considerably in statistical offices across the world, considering the need and the resources that are available to them. Ultimately, technology is an enabler that can enhance efficiency, accuracy, speed and transparency of census operations. It can also lead to optimization of costs, depending on the circumstances of each nation. It is to be reiterated, however, that the choice of technology and its level of deployment should be diligently assessed by each country before introduction.

XVI. Innovation to optimize census processes

2.252. Innovation is an important investment for any major project, and censuses are no exception. The world around us is constantly changing, and organisations must innovate to maintain a modern approach to census activities and interactions with the public and to meet ever-changing user needs. Changes in society bring new public expectations, and advances in technology lead to new opportunities. An ongoing program of innovation provides the platform necessary to keep up with new demands and opportunities.

2.253. Innovations can differ in scale. Some are adaptive, as censuses adjust to new realities of public expectations, data user needs, upgraded technology and other factors that lead to an evolution of current census practices. Other innovations can be transformative with important implications for change to a census. For example, the adoption of web-based censuses or a move from field-based to an administrative or combined census would be transformative.

2.254. It is important that censuses be innovative. While the large scale and complexity of a census may be a barrier to innovation, censuses have the advantage of a longer planning period than most other

statistical programs, during which there are pauses in operations that allow for the adoption of new innovations. Those innovations can be identified for implementation during the post-operations evaluation phase, or through a process of ongoing identification, documentation, and action of new innovative ideas.

2.255. Census innovation can have an important effect on the statistical organization. Census programs are often well-funded, relative to their non-census counterparts. This can allow for a first adoption of an innovation by the census, which can be secondarily applied to the non-census program. Leveraging a census innovation across the non-census program can be efficient and can lead to cost savings across the organization, if managed properly. It is therefore recommended that the census innovations be shared within the national statistical organization and that good lines of communication are established to also allow applicable innovations from outside the census to inform census plans.

2.256. Innovation should not take place in isolation. Those conducting a census should look to international counterparts to share innovative ideas and learnings from testing and operational experience. Domestically, it is worth exploring how other organisations undertake similar activities, such as large-scale operations or working with big datasets. Establishing a local culture of innovation, and investing in long-term research activities, is important to ensure the ongoing modernisation and future-proofing of the census.

PART THREE. CENSUS OPERATION ACTIVITIES

I. Introduction

3.1. Part Three of these *Principles and Recommendations* focuses on elaboration of census operational activities pertaining to the census methods described in Part One. While the population and housing census is in essence a statistical data collection exercise, it still incorporates components that are not present in a routine statistical survey conducted within the frame of the national statistical system. In consequence, the layout of this part follows the frame and the logic of the Generic Statistical Business Process Model (see figure 2.1.);[60] it also provides particular methodological and operational guidelines relevant to the population and housing census. This part of the *Principles and Recommendations*, therefore, starts with a discussion on developing the census questionnaire; building census infrastructure; mapping; testing; living quarters and household listings; field enumeration; data processing; analysis; evaluation of the results; dissemination; evaluation of overall census operations; documentation; and archiving.

II. Census questionnaires: content and design

3.2. The preparation of the census questionnaire to be used in a full field enumeration or combined census refers to a well-designed process that should start at a very early stage of the census operation. This process includes developing various methods and tools that will respond to national priorities and data user needs from different sectors, such as governmental agencies, research institutions, the private sector, the general public, development partners, civil society and non-governmental organizations. In addition to communicating with users to determine data needs, other factors that would have an impact on the selection of census topics, include quality of data collected through the previous census, timeliness of data production and dissemination, national sensitivity for specific topics, the burden on the responding public and available resources. Another factor to consider is the feasibility of collecting reliable data on a particular topic through the census.

3.3. Taking into consideration that the basic principle is to meet user needs and to make census statistics as useful as possible, the content of the census questionnaire, therefore, would be determined through consultations with census data users. The consolidation of the results of the consultations with all stakeholders needs to be balanced by factors such as the response burden, respecting respondents' privacy and confidentiality of their information, as well as other critical considerations such as availability of resources (financial, personnel and equipment).

3.4. The final content and design of the questionnaire will also be the result of rigorous and meticulous testing. It is necessary that all aspects of the census questionnaire, such as wording, structure of the questions and design, be carefully tested[61] to ensure the successful application of the questionnaire in the field. Questionnaire content should also be examined with respect to the proposed mode(s) of data collection to ensure that formatting and wording are appropriate to the mode(s).

[60] United Nations Economic Commission for Europe (UNECE), on behalf of the International Statistical Community, *Generic Statistical Business Process Model (GSBPM*), Version 5.1 (January 2019).
[61] For more information about census tests, see Chapter V.

3.5. Preparation of the census questionnaire(s) requires the most careful consideration, since the consequences of a poorly designed questionnaire(s), and content, cannot be overcome during and after enumeration. To ensure inclusivity, census questionnaire design must prioritize accessibility and cultural sensitivity. This involves offering questionnaires in multiple languages to accommodate diverse linguistic communities and implementing design features to meet the needs of individuals with disabilities. Successful implementation of this process will have significant impacts on the quality of collected data and census outputs. The value of evaluating the efficacy of the questionnaire during a pilot census cannot be over-stressed.

a) Selection of census topics

3.6. As a first step in determining the content of the census questionnaires, the topics that will be covered in the census should be selected taking into consideration the priorities of national needs, regional and international recommendations,[62] historical comparisons, regional and international comparability, suitability of topics for collecting reliable information and resources available for the census undertaking. Census takers should evaluate the national needs in the light of possible new topics and the needs for continuation of the topics covered in the past censuses and/or surveys. Topics that are no longer necessary in the current census due to evolving data needs and availability of alternative data sources, as well as new topics for which data is needed, should be carefully reviewed in this process.

3.7. During the process of selecting census topics, data users and interested parties should be consulted on the type and extent of socioeconomic information they believe should be provided by the census. The results from consultations should be reviewed in conjunction with available resources for the census, the suitability of the topic(s) for inclusion in the census, and the burden to respondents. In deciding what topics should be covered in the census, the length of the questionnaires and complexity of the questions should be carefully examined. More information on the factors determining the selection of census topics is given in Chapters I and II of Part Five.

3.8. Another factor that should be considered in the process of selecting census topics is related to the decision on whether to use a single census questionnaire for all respondents (all households and individuals therein) or adopt a two- questionnaire approach – a combination of a short-form and a long-form questionnaire. Using a single questionnaire consists of a standard set of questions for all individuals and housing units covered in the census. In the latter approach, countries use a short-form questionnaire with a limited number of basic questions for enumerating most of the population, while a long-form questionnaire is applied to a sample of the population for collecting more detailed information (on the same topics as contained on the short form and on additional select topics). The following paragraphs provide a more in-depth elaboration on this approach.

b) Use of short and long questionnaires

3.9. With each new census and the advancement in processing and utilising census data for informed decision making, there is an increased interest in adding topics to the population and housing census, in addition to those historically covered. Because of additional costs, the required time to collect

[62] Part Five of these *Principles and Recommendations* focuses exclusively on core and non-core topics for population and housing censuses.

and process the information, and the attendant burden on the respondents, imposing a long questionnaire on the total population, in many cases, may not seem to be appropriate or necessary. Hence, countries often decide to broaden the scope of the census by covering additional topics through the use of a sampling methodology. In this approach, two questionnaires are used: (a) a short questionnaire containing only those questions intended for universal coverage; and (b) a long questionnaire typically including all the questions from the short questionnaire plus additional questions to gather more detailed information from a sample.

3.10. Using a sampling methodology in conjunction with full enumeration requires careful planning for two key aspects:

(i) Defining the additional topics for the long questionnaire.

(ii) Selecting the sample of households/families to complete the long questionnaire: This involves determining the sampling method and sample size to ensure the data collected from the sample is representative of the entire population.

While this approach can be cost-effective by reducing the overall data collection time, it does introduce some complexity in field organization. Careful planning is needed to manage the logistics of deploying two questionnaires and ensuring proper completion by the selected sample.

3.11. The ever-expanding needs in most countries for extensive and reliable data have made the use of sampling a cost-effective part of census taking. Sampling is increasingly being used to broaden the scope of the census through asking additional questions of only a sample of the population and households. The use of sampling makes it feasible to obtain urgently needed data of acceptable precision when factors of timing, respondent burden, and cost would have made it impractical to obtain such data on a complete count basis.

3.12. The suitability of particular questions for a sample enumeration depends on the precision with which results are needed for small areas, and small population groupsI, and on the enumeration costs involved. Collecting data on particular questions for smaller areas or groups requires careful consideration of sample size. While moving the questions to full enumeration may initially appear as the only solution, exploring the option of increasing the sample size can offer advantages. A thorough cost-benefit analysis will ultimately determine the most efficient and effective approach.

3.13. It is important to bear in mind, however, that national legal requirements may make it mandatory to collect certain information on a complete count basis. Legislation in many countries prescribes complete population enumerations at particular times or makes certain political or administrative dispositions dependent on particular results from a complete enumeration. For example, the apportionment of seats in the legislature among the civil divisions of a country often depends on the number of persons actually enumerated in each division. The data needed for this and similar purposes may not be obtained using a sampling methodology.

3.14. Census information that is collected for only a sample of the population or housing units is usually obtained by one of two different methods. The first predesignates a systematic subset of census households to receive a so-called “long” questionnaire, or the census form that contains the questions on the short form as well as some additional questions on select topics. Depending on the sample requirements, which in turn take account of considerations of cost and precision, the systematic subset that is designated for the long questionnaire may represent, for example, 1 in 4, or 1 in 5, or 1 in 10 of the census households. Under such a sampling scheme, all other households in the census will receive a short questionnaire containing only those questions intended for universal coverage. If countries choose

this option, it is recommended that the predesignation of the sample households that are to receive the long questionnaire be carried out at a central location by supervisory statistical staff, since it has been shown that when the enumerators themselves actually identify the sample households in the field, the results are often biased. This can be done after the pre-enumeration household listing is completed in an enumeration area.

3.15. The second method of sampling often used involves designating a sample of enumeration areas in which the long questionnaire will be administered. In this approach, the long questionnaire will be administered in all households in the designated enumeration areas and the short questionnaire will be administered in all households in the remaining enumeration areas.

3.16. The advantage of the first method over the second is that the sampling precision of the census results is greater in the former because clustering effects increase the sampling variance when whole enumeration areas are used as sampling units. On the other hand, the advantage of the second method is that different enumerator staff may be trained more easily, since one set of enumerators need be trained only for the long questionnaire and the other set only for the short questionnaire. Sampling the entire universe often requires a significant sample size if there are questions on the long-form questionnaire that target a small population group. To reach these populations, NSOs must weigh the costs of increasing sample sizes against the increased variance of a clustered sample. However, the clustered sample would only be effective if the rare population is known to be clustered in the enumeration areas, which is often not the case.

3.17. It is important to make certain that the inclusion of questions intended only for specific subsets of the population does not create legal, administrative, or political challenges. Census data often carries legal obligations and potential penalties for non-response, necessitating careful consideration of question design to avoid complications.

c) Sections of the questionnaire

3.18. Clearly distinguishing between the units of enumeration is an essential element when designing the census questionnaire, as each question aims to collect data for a specific unit of enumeration. The census questionnaire can be clearly and intuitively structured if it is done based on the unit of enumeration. This structure has a positive impact in better understanding the meaning of the questions and in conducting the interview smoothly. This also facilitates successful completion of the questionnaire for self-enumeration (see chapter IV, section A in Part Two of these Recommendations).

d) Questionnaire design

3.19. Although the majority of countries are using face-to-face interviews with electronic questionnaires, many countries are using other methods of gathering or otherwise generating the required information, including the use of multi-mode methods. Some of these methods are face-to-face interview with electronic questionnaire, telephone interviewing, self-enumeration with a paper questionnaire collected by enumerators or returned by mail, self-enumeration via the Internet, as well as the use of population registers-and other pre-existing administrative records, either alone or in combination with other sources to generate data for the census. The methods of enumeration and technology used for data capture are among the main factors affecting the design of the questionnaires. For example, the structure of questionnaires that will be used for the face-to-face interview and for self-enumeration will differ, as the former will be administered by enumerators while the latter will be

completed solely by the respondents. Whatever methods are chosen, these need to be tested and assessed in advance for data quality and feasibility. It is also important that data security and confidentiality is maintained whatever modes or approaches are used.

3.20. The design of the questionnaire should be based on the type of data collection mode and approach used. Questionnaire design should also be based on the technology being adopted for data processing, for example whether data processing will be done through scanning, manual entry or electronic transmission of data to a database. It should be noted, however, that regardless of the mode of enumeration used, it is advisable that there is compatibility in the design among the modes in order to enhance data comparability (for more information, see chapter VII, section B on designing multi-mode data collection).

3.21. Where countries utilize the Internet or handheld devices to collect their census information, or a portion of that, the layout and organization of the data collection instrument may differ from that of the paper questionnaire, although it is important that the content be the same for all the modes of data collection used. Comparability of content is important to ensure that the same information is collected and that entries are similarly checked across modes of data collection in order to enhance comparability of data among the modes used. It is important to note, however, that most often, adopting an Internet-based data collection approach also means moving from an enumerator-based approach to a self-completion approach. Consequently, the questions must be designed to be completed by the respondent without the assistance of an interviewer. In this case, care should be given to the design of the questions so that respondents can easily provide the required information without assistance. Therefore, it is imperative that census management involve the information technology team, as well as cognitive testing experts right from the questionnaire preparation stage. While many of the same principles (for example clarity of wording, omission of unnecessary material) will apply also to an Internet-based or handheld device-based collection of information, specialized advice should be sought and provided regarding such issues as (a) the technology employed to present the questions to the respondent; (b) the method of capturing the response; and (c) quality assurance checks employed during data capture process. It is also important to evaluate mode bias resulting from differences in responses to the same question being collected using different modes (e.g., Internet-based vs. in-person).

3.22. A crucial principle is that questionnaire design must be regarded as part of an integrated process of satisfying user demands by collecting, processing and disseminating high quality information provided by respondents.

3.23. It is important also that questions and response options are free from ambiguity. Moreover, questions should not be offensive nor biased and should use inclusive language to the extent possible; in many cases this can be avoided by excluding extremely sensitive topics from the census questionnaire, but care must always be taken to consider the reaction of respondents when designing questions. In addition, it should be noted that the quality of information collected in a census will be reduced if the questionnaire is excessively long. These issues should be carefully assessed during the testing programme, including any cognitive tests and a pilot census (see paragraphs 3.144–3.145) since poorly worded questions will not only collect poor quality data, but, by confusing respondents and/or enumerators, may also impact on subsequent questions in the questionnaire.

3.24. Skip patterns and the form and format of the response options require careful consideration. A question can be perfectly worded, but if the response options are not presented in a way that is meaningful, unambiguous, and easily understood, resulting data may be of poor quality.

3.25. Special provisions will have to be made if two or more languages are used in the country. Several methods have been used to deal with this situation, such as (a) a single, multilingual questionnaire; or (b) one version of the questionnaire for each major language; or (c) translations of the questionnaire in the various languages available in the enumerators' manual or on the Internet site for the census. In order to ensure consistency and to avoid interviewer bias, it is important that translation of the questionnaire into various languages that are used in the country be centralized and not be undertaken in the field by interviewers. Information on the distribution of languages in the country is important for sound census planning and, if not available, will have to be collected at some stage of the census preparations. Staff recruitment and training procedures (see paragraphs 2.192–2.202) will also have to take language issues into account.

3.26. If the housing census and the population census are to be carried out concurrently, it is necessary to consider whether a single questionnaire should be utilized to collect information on both population and housing topics. If separate questionnaires are used, they should be uniquely identified in a way that links the component forms so as to permit subsequent matching, both physical and automated, of the data for each set of living quarters with the data that refer to the occupants thereof. This consideration is particularly important when employing separate personal forms for each individual in the household.

3.27. When paper questionnaires are used for data collection, the use of processing techniques, such as optical mark reader or intelligent character recognition, will have a significant effect upon the questionnaire design (see paragraphs 3.262–3.264). In the case of optical mark reader, it is necessary both to provide spacing of response areas and to ensure printing is undertaken to precise tolerances so that the data capture software is able to capture all required data but not any of the material around the designated response areas. With regard to intelligent character recognition (ICR), it is crucial to allow sufficient room for response areas and to ensure that these are designed according to the requirements of the processing system so that each response box contains only one character. Enumerators should be trained to print neatly and correctly (usually in upper case). To better handle the entries by self-respondents, the ICR technology has to be built to interpret various forms of manual entries. As noted in paragraph 3.26, where the scanning process requires that a booklet questionnaire is separated into component pages, it is important that some form of linking (for example by serial numbers or barcodes) is employed to ensure that the correct information is amalgamated in the computer records.

3.28. The final questionnaire must be drafted in time to allow for: a) printing and developing the data collection application, in the case of using electronic questionnaire (making allowance for the many contingencies, such as industrial action, breakdown of printing equipment or delay of programming activities, that can arise in these processes); and b) undertaking quality assurance checks to ensure the printing is of sufficient quality to be used in the data capture regime and to ensure that the data collection application is running correctly under the data entry rules.

3.29. As some countries are also utilizing Internet portals and handheld devices for conducting the census, sufficient time must be given to design, develop, test and implement bug-free e-questionnaires

and related software systems. Last-minute inclusion or changes in the questionnaire may affect the overall quality of the programs, and in turn the census results.

3.30. In view of the many issues to be addressed in designing a census questionnaire, it is not feasible to suggest specific model questions for the census topics that are covered in Part Five of these Recommendations. However, images of all census questionnaires that have been made available to the United Nations Statistics Division have been placed on the Division's website[63] together with research papers relating to questionnaires used to collect information on the various topics recommended for collection and also using the different technologies (Internet, handheld devices, programs and scripts used for quality control).

e) Design of questionnaires for electronic data collection (CAPI, CAWI, CATI)

3.31. The design of electronic questionnaires for data collection via CAPI, CAWI and CATI methods require additional considerations to make the data entry process intuitive for the enumerator or respondent. Some essential functional features that should be used in the design of the electronic questionnaires include:

(i) **Questionnaire navigation** should allow enumerators/respondents to move sequentially through the questionnaire in order to enter responses in the most effective way, giving the ability to pause and resume at the last answered question with a "save and continue later" functionality. On the other hand, the design should impose some restrictions on navigation, for example, by preventing enumerators/respondents from entering certain questions without having first obtained responses from other, earlier, questions.

(ii) **Skipping/automated routing** is one of the most important error reducing features in electronic questionnaires. It obviates the need to respond to questions that should be skipped. It also avoids the converse - skipping questions that should be asked, thus minimizing the need to impute for missing responses. Basic skips allow the response to a particular question to determine whether or not the next question is relevant, while complex skips are those that either use responses from several previous questions to determine whether the next question is relevant.

(iii) **Pre-coding** allows relevant questions to be answered from precoded drop-down menus. In some cases, drop-down menus could be altered dynamically, depending on previous responses, so that the interviewer is never presented with an impossible response code.

(iv) **Validation.** Real-time data validation checks can correct invalid or inconsistent responses that could be the result of either interviewer or respondent error, thus reducing the amount of post-enumeration data edits.

(v) **Support** should be available for different types of issues such as: technical problems; and the actual content of the census questionnaire – in case respondents have difficulties understanding what is being asked of them. Support can be provided in a number of ways including through help buttons and additional instructions and help materials.

III. Building census infrastructure

[63] See http://unstats.un.org/unsd/demographic/sources/census/censusquest.htm.

3.32. The population and housing census requires a concentrated effort in building an intricate and complex infrastructure that is suitable for conducting massive activities in a simultaneous manner. This infrastructure refers to: the development of the census instruments/questionnaires; logistic schemes for disseminating and collecting returns; facilities and systems for capturing, processing, editing, and validating response data; and the systems and tools for dissemination of statistics and data products. It also requires well-defined workflows and testing of the production system as well as of the whole statistical business process. As most censuses are periodic with as many as ten years between them, the infrastructure is also likely to be temporary, so care must be taken to ensure the appropriate balance between what is needed for a given census and what is a reasonable expenditure of resources for a temporary infrastructure.

3.33. The necessary infrastructure for a census will vary significantly based on the design of the census and the data collection mode(s) being employed. For a field-based paper-based census, the infrastructure will need to support the paper-based activities such as printing questionnaires and materials, space for storage of paper materials, mailing and distribution plans for distributing paper materials to field operational areas and returning completed questionnaires for processing, systems and equipment for capturing data from paper questionnaires, and space/procedures for long-term storage of paper materials based on the organization's records retention policy.In addition, special procedures must be developed and distributed to all office and field staff on the appropriate security measures necessary to ensure the safety and integrity of the paper materials including sensitive information such as addresses and personal data.

3.34. For electronic data collection, on the other hand, the infrastructure will need to include facilities, equipment, software, and information technology security measures involved in the electronic data entry, transmission, and storage. In this case, cloud infrastructure can be involved to achieve the necessary flexibility with the performance and reliability of the provisioned resources. The infrastructure for a register-based census will need to also include extensive storage for data files and tools and systems to facilitate the necessary matching and processing of the administrative data. Censuses that employ a hybrid approach using a combination of two or more of these approaches can require a very complicated infrastructure that must support all of the data collection modes being used.

3.35. A census infrastructure must be put in place long before the data collection exercise itself, as all of the components need to be extensively tested in circumstances that are as real as possible and require reasonable time for any corrections needed as the testing may indicate. The importance of testing cannot be overestimated, especially in the context of the need to build process components for processing and analysing data. The preparation of the infrastructure does not usually translate into immediate procurement decisions, as the technological process results in better performance, richer functionality, and lower cost of the equipment with time. Thus, purchasing equipment or devices (such as mobile devices for data collection) several years in advance of the fieldwork brings the risks that the equipment/devices will become obsolete or unsupported by the time of the census. Instead, a limited sample of the equipment can be acquired at the time of building the infrastructure to facilitate testing, with the bulk being acquired later, when confirmed suitable by the tests, and to take advantage of the most recent technology available.

3.36. Building census infrastructure should be based primarily on the experiences and lessons learned from the previous census. If the previous census took place a long time ago and the documentation and institutional memory are not sufficient, putting together census components should rely on statistical practice in conducting large surveys and on the schemes for developing and implementing a statistical

business process. Lessons learned from conducting similarly demanding operations (for example, an agricultural census, or census of establishments) can also be considered. Collaboration and consulting with other statistical organizations that have more recently conducted a large-scale survey effort or a census (for example, a neighbouring country) can also be a valuable source of information and guidance.

3.37. Particular components of the census infrastructure may already be available within the national statistical office – for example, the statistical network of field offices covering the country's territory and staff experienced in collecting and producing statistics. Other components will need to be established from the beginning, such as the training synopsis and schedules, recruitment procedures and logistical arrangements. Methodical planning is critical in this process of building census infrastructure and developing detailed project schedules early would significantly facilitate managing the process in a holistic manner.

3.38. In that context, clear and unambiguous delineation of management responsibilities is equally important to the building process. Managers and supervisors need to have a comprehensive assignment from the very beginning and at all levels of managing. Along with roles and responsibilities, methods and channels of communication must also be clearly defined to ensure the different organizational units are collaborating and coordinating their assignments. While improvising may be necessary in certain circumstances, all efforts should be put into anticipating risks and developing and testing strategies to allow managers to mitigate them and have prescribed alternatives in place when needed.

3.39. In building new components of the census infrastructure – for example, outsourcing – it is necessary to liaise with other parts of the government that may have more extensive experience in that respect. In most cases, other government departments may not only have extensive experiences – such as subcontracting for defence purposes – but also might have the capacity to provide the goods and services that are needed. Therefore, the process of building infrastructure needs to first assess existing government capacities, and then extend beyond them.

3.40. The process of building the census infrastructure has to be completely finished before the pilot census takes place. Certainly, parts or perhaps even whole components of the infrastructure might need to be adjusted and altered as a result of the pilot census. Taking this process into account, the schedule for the preparations of the census must allocate sufficient time for the modification of systems necessary at the end of the pilot test.

3.41. Once the census takes place, all the steps in building the census infrastructure and its testing need to be documented in detail and stored for future use. Certain components will certainly remain permanently as part of the broader national statistical infrastructure, such as computers, networks and servers. However, documenting in a comprehensive manner is of paramount importance, not only for future censuses, but also for a better understanding of the one that has just taken place.

IV. Geospatial information and mapping

A. Strategic planning

3.42. Everything happens somewhere, at a given point in time. Geospatial information provides one of the most crucial aspects of census taking – location, essential at all its stages. The fundamental role of geospatial information in census operation activities is to support enumeration and to present

aggregate census results in geographic and cartographic form. Very few enumerations during the last several census rounds were executed without extensive geospatial support.

3.43. There is widespread recognition of the importance for national statistical agencies to develop a long-term geospatial data strategy and build the necessary infrastructure and human resources to support the nation's data needs. Leveraging geospatial capabilities can significantly enhance population and housing censuses and other statistical activities. A dedicated mapping function within the NSO can contribute to census analysis, data visualization, and a range of geospatial applications and services. This rich data source can serve as the foundation for a robust National Spatial Data Infrastructure (NSDI), promoting standardized geographic units and facilitating location-based services across government levels. While building internal geospatial capacity is essential, NSOs should also explore opportunities to collaborate with existing national geospatial information agencies (NGIAs) or research institutions. Partnerships and knowledge-sharing can optimize resource utilization and accelerate geospatial integration within the statistical system.

3.44. Geospatially enabled census data is indispensable for effective decision-making. Location-specific information provides unparalleled insights compared to aggregated national data. By integrating census data with other statistical information through geographic units, policymakers can access actionable knowledge to inform policy and implementation.

3.45. A census limited to aggregated data offers restricted value. Understanding population distribution at a granular level is essential for comprehensive analysis and evidence-based decision-making. Strategic planning for the integration of census statistics and geospatial information yields interoperable datasets that can be analyzed at various spatial scales, ensuring data relevance for diverse policy and decision-making needs.

3.46. Noting the growing relevance and need for the integration of statistical and geospatial information as an important bridge to enable the production of harmonized and standardized information and integrated geospatially enabled statistical data to facilitate data-driven decision-making, in particular with regards to population censuses and monitoring of internationally agreed development agendas, the United Nations Statistical Commission, at its fifty-first session, in March 2020, approved the Global Statistical Geospatial Framework (GSGF).

3.47. The GSGF should be considered a critical component of census planning and operations, aligned with the broader context of the GSBPM. The GSBPM provides a structured approach for integrating geospatial considerations into the overall statistical production process. The GSBPM sets out typical activities and stages that statistical organizations undertake when producing statistics and this provides a common thread for documenting activities related to the geospatial domain so that relevant actions are taken at the correct stage of the census. Moreover, geospatial information, including geographic information system (GIS) and mapping, can be applied in a census structured according to the GSBPM.

3.48. In summary, the GSBPM geospatial view[64] will enable the description of geospatial activities, essential to producing geospatially enabled statistics, using the GSBPM framework. By considering the GSGF Principles, these geospatial activities will contribute to increased data standardization, flexibility, and integration.

[64] Geospatial View of Generic Statistical Business Process Model - GeoGSBPM.

3.49. The use of geospatial information and mapping methods can range from printed maps to Global Navigation Satellite System (GNSS)-equipped handheld devices and geoportals according to the technological and financial capacity of each country. Adding geospatial information and technology to the census may, at first, seem costly for some NSOs, however experiences in many countries have been gains in efficiency, increased coverage and accuracy and return on investment into the census activities.

3.50. Major technological advances include the widespread availability of handheld devices, global navigation satellite system (GNSS), geographic information system (GIS) software, low-cost aerial, drone, and satellite acquired imagery, and open and free earth observation data[65]. These advances have put new tools in the hands of national statistical organizations to collect more accurate and timely information about their populations. Adopting such new methods can require significant planning and lead times for building new enterprise capacity and organizational restructuring but this allows for the development of an incredibly powerful geospatial data system that should then be leveraged to support a variety of services and applications beyond census activities.

3.51. There are various approaches to geospatial support services for census activities. For example, traditional mapping techniques have been used successfully for many countries and are still relevant in some countries or at least some parts of many countries, particularly in remote areas. However, with recent developments in technology, countries are now leveraging new technologies to develop digital mapping techniques and improve the quality of census operations. The application of new technologies requires more careful and long-term operational and managerial plans based on a realistic assessment of costs and human resources required, but costs have reduced rapidly for hardware and software solutions. The proliferation of handheld devices, software vendors with off-the-shelf solutions, and the rise of free and open-source software have driven down the costs of the technical infrastructure. Outsourcing can be considered as a viable option when internal capacity for geospatial support is limited. However, a thorough evaluation of costs, benefits, and potential risks should precede any outsourcing decision.

3.52. It is recommended that all NSOs prioritize digital collection of information for full field enumeration censuses. Additionally, location data through coordinates and/or addresses should be considered core census information. For countries where printed maps remain necessary, greater emphasis should be placed on registering addresses and locational information (e.g., geographical names). This will allow for better supervision of the operation, data disaggregation, and the creation of a solid location database for future use. In many countries, a combined solution using paper and digital maps can be useful, especially during the information collection stage.

3.53. In the case of register-based or combined census, location must also be considered core, especially due to the ability to allow cross-referencing with other georeferenced statistical databases. It is also an ally to assess the territorial coverage of administrative records and the need for field collection. In these types of censuses, addresses tend to be more frequent than coordinates, so attention must be paid to situations involving population in the remote areas, indigenous populations, and residents of informal settlements whose addresses may not be clearly defined or may be unknown. The same applies to countries that use mixed-mode data collection, making sure that in all the modes used it is possible to collect precise locational information.

[65] Open and free data from programmes such as Landsat and Copernicus, are relevant for several applications including the census.

3.54. A range of techniques and technologies exist for incorporating geospatial information into census-mapping exercises. However, the financial and technological realities, technical capacity, and team experience of each national statistical office must be considered when selecting techniques and technologies. These recommendations aim to present and develop the essential principles behind geocoding of census information. They emphasize the importance of developing and improving the utilization of geospatial information and its integrating with census statistics.

B. The role of geospatial information in the census

3.55. Mapping has been an integral part of census taking for a long time. The fundamental act of taking a census is inherently a geospatial endeavour; enumerating a population in each area, at a particular location, in a specific dwelling, is fundamentally geographic in nature. Over the years, census maps have played a critical role in all processes from preparation to dissemination of the census results. More recently, the use of geospatial information has been increasing in various stages of censuses around the world.

3.56. In general, mapping, including its digital version, serves several purposes in the census process, as follows:

(a) Maps ensure coverage and facilitate census operations (pre-enumeration). The census office needs to ensure that every household and person in the country is counted and counted only once. For this purpose, census geographers partition the national territory into contiguous small data collection units. Maps showing enumeration areas thus provide an essential control device to ensure full coverage of the census.

(b) Maps support data collection and can help supervise census activities (during enumeration). During field-based census, maps ensure that enumerators can easily identify their assigned geographic areas, in which they will enumerate households. In a more technologically mature census, it is possible to load these maps onto handheld collection devices equipped with GNSS and issue alerts if the enumerator leaves his work area and invades another's work area (geofencing).

(c) Maps are also issued to the census supervisors assigned to enumerators to support census operation activities planning and control tasks. Maps can thus also play a role in supervising the progress of census operations. This allows supervisors to strategically plan, make assignments, identify problem areas, and implement remedial action quickly.

(d) More recently, digital maps or geoportals can be used to display collection progress in real time. In addition to showing the progress of the collection on a daily basis, these maps, when used with handheld equipments with GNSS, will be able to accurately display the houses visited with coordinates. It is possible to view this data at different scales, such as territories, regions and enumeration areas. Base layers such as satellite images and Open Street Map can be incorporated to help identify areas not covered or even not previously mapped.

(e) Digital maps can also be used to verify spatial patterns of refused, absent or unoccupied housing units, directing surveillance efforts or even localized advertising campaigns to reduce non-response rates in census operations.

(f) Maps are essential tools for presenting, analyzing, and disseminating census results, offering a visual representation of demographic and social patterns. Geospatial information plays a crucial role in enabling the production of geocoded census outputs, including grid-based

data, which enhances the ability to analyze and visualize census results at various spatial scales. This supports evidence-based decision-making across the public and private sectors.

(g) Digital maps loaded with high-resolution satellite images, household coordinates and aggregated demographic data have also proven useful for disaster response action. The confidentiality and granularity or geographic disaggregation of this data must always be considered.

(h) Digital maps loaded with recent high-resolution satellite images can also be useful to verify the migratory dynamics of nomadic populations or the expansion of informal settlements, ensuring the count of these population groups.

3.57. The census enumeration team needs to have a set of unique maps covering the entire country that accurately defines the boundaries within which each enumerator will work during the enumeration phase. Therefore, the quality of maps used in the census has a major influence on the quality and reliability of census data.

3.58. The types of maps required for census management include the following: (a) small-scale reference maps for use in the national statistical office or census agency to manage the overall operation; (b) large-scale topographical maps for use by enumerators; and (c) maps of the subregions or administrative areas, for the use of managers, showing the location of small population settlements and dominant physical features, such as roads, rivers, bridges and the type of terrain. The use of satellite images and earth observation (EO) data to generate these maps is now increasingly common in many countries. NSOs with fewer resources can implement using readily available free and open global human settlement maps. While such maps offer a valuable starting point for resource constrained NSOs, it's crucial to recognize their potential limitations in terms of currency and accuracy. Regular updates and ground-truthing are essential to ensure map reliability. Collaborating with national geospatial information authorities (NGIAs) to develop and maintain accurate cadastres is vital for supporting census operations and broader geospatial data needs.

3.59. Careful consideration should be given to organization and management of geospatial support activities during the census planning and preparation phases. The lead time necessary for the implementation of a mapping and geospatial support solution for a country will be determined by a wide range of factors, including the number of maps to be produced, the technology available to produce and assign them, the availability of funds to acquire additional resources and the time required for distribution of the maps to field staff.

3.60. It is still the case that in many countries there are only a limited range of maps available and these often do not show sufficient detail to enable the boundaries of small areas to be clearly defined. This is particularly likely to apply in areas of unplanned or informal settlement. However, with high resolution satellite imagery of sufficient quality and other current earth observation (EO) data sources this is no longer the case. Unplanned areas in some countries can be mapped precisely allowing for efficient field activity planning. These spatial data sources can be further enriched with other material, such as:

(a) lists of households, preferably compiled by statistical agency staff as part of the process of delineating enumeration areas, but on occasion provided by local leaders (see paragraphs 3.147–3.152); or

(b) a textual description of the boundary, including roads, railway lines, power lines, rivers and other physical features. This description may also include obvious landmarks on the boundary (school buildings, water points and other reference features).

3.61. This can create the base of a rich geographic framework for use as a reference moving forward. In this way, census activities can be used to build out or update the national map or provide an important source for the update of an existing national geospatial database. The geospatial data collected during planning and field activities should be treated as an extremely important and rich data source for further geospatial processing.

3.62. Previous experience has shown that relying entirely on a list of households, written or verbal descriptions and directions or local knowledge of the area boundaries will often lead to confusion and error because people tend to have mental images of places that may not align with the area as it is really reflected in the design of the enumeration area. Similarly, the supervisor's mental map of an enumeration area may differ markedly from that of an enumerator, and from enumerator to enumerator. To overcome such problems, it is important that the best possible quality maps be the basis for census enumeration operations and that the collection staff receive comprehensive training in the correct use of use of the maps, digital maps, GNSS handheld devices and any associated documentation provided.

3.63. Despite the great value that paper maps had in censuses, technology today allows many of the limitations of this instrument to be overcome. The widespread use of handheld devices equipped with GNSS positioning technology allows census takers to locate themselves, within their responsible enumeration area by receiving alerts in case of invasion of other EAs, among other functionalities. Furthermore, collecting coordinates of the units investigated in the census allows for a significant improvement in supervision activities.

3.64. Where a digital base map is prepared, this may be used in conjunction with a series of geospatial layers as the basis for coding information collected in the census. This could apply to address of usual residence now and/or in the past, or place of work or education.

3.65. The use of paper maps is still very useful, especially in areas not or inadequately covered by GNSS or where the use of technological equipment poses risks, such as theft or damage. Many countries have successfully implemented a hybrid approach, combining paper and digital maps to optimize census operations. Paper maps offer flexibility for on-the-ground adjustments and data recording, while digital maps provide precise location information and enable advanced spatial analysis. A hybrid approach, combining both formats, can optimize data collection and management efficiency.

3.66. Statistical agencies should avoid a paper-only solution. This approach greatly limits exploring the potential of location data at all stages of the census and reduces the possibility of integrating statistical and geospatial data. The subsequent costs in time and financial resources to standardize and process paper-collected data are unlikely to be justified. A paper-only solution should only be considered as a last resort and in very limited circumstances.

3.67. With the rapid growth in freely accessible high and medium resolution earth observation data, specifically satellite and aerial imagery, there are increasingly few, if any circumstances where hand-drawn maps or hand-collected data should be incorporated as an acceptable standard solution for a modern census. A paper-based collection solution is not subject to the same level of quality assurance and quality control that can be built into a digital capture solution. In the case of capturing geographic data (such as the location of a housing unit) on paper, the basic geospatial reference data is not

captured, and may require additional time and research in the central office or necessitate additional field visits.

3.68. The next two sections cover geospatial information technology and geographic information system (GIS), fundamental elements for building a geolocated database in a census. These recommendations speak directly to Principle 1: Use of fundamental geospatial infrastructure and geocoding of the Global Statistical Geospatial Framework (GSGF) whose objective is to obtain high quality, standardized location references such as physical addresses, property or building identifiers, or other location descriptions, in order to assign accurate coordinates, and/or a small geographic area or standard grid reference, to each statistical unit at the microdata/unit record level.

3.69. Time and date stamping locations is crucial to clearly place the statistical unit both in time and in space. By grounding census data in a strong geospatial foundation, NSOs can more effectively develop additional statistical products, integrate administrative records, and disseminate census results at granular levels. This enables a wide range of data-driven applications and insights.

C. Geospatial information technology for census

3.70. Before census mapping commences, the census agency needs to determine the appropriate technology for doing so. Countries need to choose technologies to improve efficiency of census operations, data quality and timeliness, balancing these with cost factors within the context of their national needs and circumstances. The application of technology must also ensure that confidentiality of data is maintained.

3.71. In general, countries need to approach the use of technology for mapping as a continuous process rather than merely a sequence of mapping and dissemination operations. Use and application of geospatial technologies and geographic databases are very beneficial in improving the overall quality of census activities at all the stages of the census. Major technological advances include the widespread availability of personal computers, handheld devices, and personal digital assistants, GNSS and GIS software, and low-cost or open-source aerial and satellite imagery. These advances would be of interest to national statistical organizations in collecting more accurate data in a timely manner.

3.72. Where accurate and current maps at relevant scales are not available for a country, or part of a country, the technological alternatives described in the following paragraphs could be employed subject to consideration of the constraining factors described in paragraphs 3.73–3.77:

a) ***Satellite images*****.** A satellite image typically covers a large area and can be cost-effective compared to other sources. Imagery should be pre-processed by the supplier so that it is rectified and georeferenced (a known scale and orientation, with some latitudes and longitudes, is printed on the face of the image). Satellite image data have gained in volume, popularity, and ease of use. Satellite imagery, if used pragmatically, can save countless person–hours by focusing attention on critical areas. Remote sensing data can be used as an independent check on the field verification process.

b) ***Aerial photography*****.** Acquisition of aerial photographs for large areas of a country may be expensive. However, existing archives of photographs can be an excellent resource for preliminary counts of dwellings and as a base for basic maps. In many cases digital images captured from an

aircraft or drone-based platforms are a cost-effective way of incorporating components of a GIS. However, NSOs should not equate the use of satellite images and aerial photography as an end product. The information contained in them is raw and has to be processed – thereby adding to the burden/cost of their use. Sometimes, where there is no capacity for additional field operations, or no current digital information is available, open data of human settlements can be a rather good asset to fill geospatial data gaps (see point (e) below).

c) ***Other Earth Observation (EO) Data.*** Earth Observation (EO) is the process of gathering information about the Earth's surface, waters, and atmosphere via ground-based, airborne and/or satellite remote sensing platforms. The acquired data, usually in the form of digital imagery, are processed and analysed to extract different types of information that can be used to monitor and assess the status of – and changes in – both the natural and human-made environments. For situations where there is no digital map available, or specific acquisition and processing of satellite imagery or aerial photography is unpractical, open and free global human settlement maps can fill data gaps and serve as information baseline for census activities. These products delineate built-up areas and human settlements based on open satellite imagery data. These resources are well documented and widely used to delineate human settlements. For some regions of the globe there are also building footprints which increasingly available for free and which can provide an initial georeferenced database of buildings. Digital Surface Models (DSMs), which represent the Earth's surface with all the objects on it, and Digital Terrain Models (DTMs), which represent the bare surface of the ground without any objects like plants and buildings can also allow EO information to be incorporated into census mapping.

d) ***Global Navigation Satellite System (GNSS).*** The standard generic term for satellite navigation systems that provide autonomous geospatial positioning with global coverage. This term includes e.g., the Global Positioning System (GPS), Global Navigational Satellite System (GLONASS), Galileo, Beidou and other regional systems. As laptop computers and other portable computing devices equipped with GNSS are becoming less expensive, integrated field mapping systems are becoming a viable option for collecting accurate field data (latitude and longitude) for census purposes. Advances in technology, including GNSS, wireless communication, and computer miniaturization, have made numerous new applications for portable GIS possible, particularly the development of specialized software for census fieldwork. The use of this equipment combined with digital maps or geoportals will bring significant gains to the supervision of activities, including the possibility of recording enumerators' routes.

e) ***Georeferenced address register.*** A high-quality, comprehensive, updated, and georeferenced address register of each building and dwelling can give great support in planning and organizing a census. A georeferenced list of addresses can play a central role in many fieldwork operations and will provide the key to accurate delivery, collection, and follow-up of questionnaires. The best way to associate each address with a location in physical space of a map is to specify its coordinates in a proper geographic reference system. With geographic coordinates addresses can be entered in available maps. Ideally, addresses and coordinates should be collected in combination, as they are complementary information. In rural or remote areas, it may not be possible to collect the address, making it important to collect the coordinates.

3.73. The implementation of strategies using such technologies must be thoroughly planned with the guidance of qualified staff or external experts with formal qualifications in the use of advanced mapping technology. It is particularly important that the cost of acquiring and maintaining the hardware required

to use this technology is factored into the budget (and a sound cost–benefit analysis undertaken to support such changes), and adequate plans are made to ensure the availability of sufficient quantities of hardware in time for the census.

3.74. It should be noted that there might be additional risks due to the need for equipment to be operated in suboptimal conditions, including poor weather, dusty conditions, or poor lighting. Despite its versatility, GNSS may not be able to differentiate the coordinates of overlapping or closely located dwellings in multistorey buildings and in this circumstance should only be regarded as providing coordinates for the building rather than the dwelling units within it.

3.75.Software and applications that can read both static and interactive maps should be considered as they will greatly help in deciding the type of maps to produce. Interactive maps are very useful in the identification of EAs in countries where structures have no addresses. Software and application selection should consider a balance of cost-effectiveness, compatibility, and usability. In addition to the geospatial team's expertise and the need of enumerators, input from other relevant agencies, such as national cartography and geospatial information authorities, is crucial to align technology choices with broader national geospatial strategies.

3.76. It is important to ensure that where such systems are employed, they are clearly understood by enumeration staff. This should be achieved by ensuring that the staff, whether at the mapping update (pre-enumeration) stage or enumeration stage, are given adequate training in the interpretation of the maps. Should the maps be incorporated into digital devices such as personal data assistants, the staff should be trained in the use of both the hardware and the software.

3.77. As with all other significant changes to census procedures, it is crucial that census geographic and mapping processes are successfully included in tests prior to being used in the main operation. This is particularly the case where a change in level of technology is being considered.

D. **Geographic information systems**

3.78. A geographic information system (GIS) can be seen as a system of hardware, software (commercial and/or open source) and procedures designed to support the capture, management, processing, analysis, modelling, and display of spatially referenced data. In practical terms, a user’s interaction with such a system may range from a simple hand-held or desktop mapping facility or a web-based mapping and digital visualization application.

3.79. Careful planning should be undertaken to plan for a complete GIS system that can solve complex census operational planning, providing solutions for managing operations, identifying, and resolving enumeration coverage problems, monitor field operations, enforcing quality controls, producing detailed georeferenced inventories, and supporting a variety of dissemination platforms, products, and formats. The ability to use geographic space to integrate and manipulate data sets from heterogeneous sources makes a modern geospatial system the critical and the fundamental tool for planning and managing the modern census.

3.80. For example, a GIS provides functions for the aerial interpolation of statistical data in cases where the administrative unit boundaries have changed between censuses. While the development and implementation of such a repository of georeferenced data can be daunting and take extensive

planning, simple desktop mapping systems that generate thematic maps from a database of base maps and indicators will satisfy the needs of most census organizations, the census should be seen as an opportunity to make strides and build upon previous geospatial operations. Once collected, a plan for geospatial data maintenance can be developed, and each census can provide the opportunity for an update and advancing development of the NSDI.

3.81. GIS technology should be considered only at a level appropriate to the skills and resources available and should constitute an integral part of the overall work of the organization. Cooperative arrangements with other agencies – especially with the national geospatial information agency or national mapping agency – should be pursued particularly regarding the acquisition and maintenance of base map data, which usually is not the responsibility of the statistical organization. The United Nations Integrated Geospatial Information Framework (UN-IGIF) provides a basis and guide for developing, integrating, strengthening, and maximizing geospatial information management and related resources supporting NSOs in seeing themselves in the national geospatial data production ecosystem.

3.82. Given the great potential of geospatial information for censuses, statistical organizations should pursue the development or implementation of GIS and integration of geospatial information whenever possible, including for intercensal surveys. The integration of statistical and geospatial information must be an objective of NSOs, understanding themselves as important institutions that produce or consume geospatial information. Therefore, NSOs must pay attention to the 5 GSGF Principles for internal structuring.[66]

3.83. Statistical offices must also understand GIS as an important tool for spatial analysis and dissemination of census data. The use of interactive maps, geoportals, and user-centred geography construction tools[67] tend to be more direct and simple ways to interpret information than tabular data, especially for the general public. NSOs provide vital information about current demographic conditions and future trends to policymakers in a range of sectors, such as health care, education, infrastructure planning, agriculture, and natural resource management; and the availability of spatially referenced census databases is an essential prerequisite to facilitate the use of demographic data.

3.84. In this regard, it should be noted that the GIS should be capable of generating additional geographic delimitations beyond those used in the census, such as school districts, water catchment areas, or power service units. These entities will have to be constructed from the smallest geographically identified units available in the census (for example block faces, grid squares or enumeration areas). If (as is the case in most developing countries) enumeration areas are the smallest unit, this will have important implications for the establishment of enumeration area boundaries. Cooperation with the authorities responsible for these geographic entities before the boundaries of enumeration areas are drawn can reduce later problems.

3.85. With the rise of free and open-source software, the cost-benefit analysis for rapid deployment and implementation of a GIS programme in a national census is irrefutable. However, in some developing countries with limited capacity and resources, GIS may need to be introduced carefully and gradually, given that is a rather complex technology and a resource-consuming one. As an alternative to

[66] See https://unstats.un.org/unsd/statcom/51st-session/documents/The_GSGF-E.pdf

[67] See https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/buildacustomareaprofile/2023-01-17.

immediately launching full-scale GIS applications, countries may start with a simple and robust design that is likely to be understood and maintained by a wide array of users, transferable to a wide range of software packages and independent of any hardware platform.

3.86. GIS implementation in a developing country may follow a hierarchical strategy, with the national statistical office employing a high-end commercial GIS with extensive capabilities for handling and analysing large amounts of spatial data. Widespread dissemination of databases can then be achieved by creating a version of the finished databases using a low-end mapping software format for distribution at low cost and through web dissemination of macro information in an online GIS.

3.87. In recent years, many countries have adopted GIS to facilitate census mapping in the production of enumeration maps and outreach products. As the cost is decreasing, geospatial technologies and inputs are being made available free of charge or openly, and the basic technology is now well established, we expect this to continue to expand in future census rounds. It is likely that the census could help to increase the capacity of the statistical service (or the country as a whole). The adoption of GIS should therefore be seen as an important strategic decision, with impacts that go beyond the operation of the census, and many issues need to be considered.

3.88. A geospatial database at the enumeration area or grid level is a cornerstone for the national statistical system, providing integrated statistical and geospatial data essential for managing, analyzing, and disseminating census information and monitoring changes over time. It also serves as a fundamental building block for a national spatial data infrastructure, enabling the integration of socio-economic and environmental data for evidence-based decision-making. A prerequisite for building a geographic database at enumeration area level or in grid format is the development of a geocoding scheme, in which each enumeration area (EA) has a unique code, an administrative identifier which can be used to link geographic features to the attributes registered for them.

3.89. The potential benefits and challenges of GIS are summarized as follows:[68]

(a) Benefits
    (i) Efficiencies and insights for field preparation for data collection
    (ii) Closer linkage between maps for enumerators and map-based products for users;
    (iii) Enriched dissemination of census data as they can be visualized in geographic areas for easy understanding by users;
    (iv) The cost of intercensal updating of the base map will be less with a digital base map, enabling among other things the construction and updating of sampling frames;
    (v) Producing duplicate maps may be less expensive with a GIS solution;
    (vi) GIS will have increased ability to undertake quality assurance of geographic boundaries;
    (vii) The census agency will have a greater ability to perform spatial queries and advanced analysis under GIS, including the application of the degree of urbanisation method; and
    (viii) Space needed to store input maps for digital purposes will be far less.
(b) Challenges
    (i) GIS requires additional technical expertise;
    (ii) GIS will require a higher level of computing infrastructure;
    (iii) A clerical census system can proceed on the basis of basic maps. However, use of GIS in this task requires that a digital map base exists. If it is necessary to create the digital map

base, significant lead times are required as well as significant funding. In both cases, more experienced technical staff are required;

(iv) In most cases, the preparation of maps or GIS will not be the core business of the statistical agency.

E. Census geography

3.90. Prior to developing a geospatial support strategy for the census, consideration needs to be given to the geographic classification to be used and the geospatial infrastructure necessary to carry out the required tasks. As the geographic unit used for basic census collection usually[69] determine the minimum geography on which the census data can be disseminated, a geographic classification should be devised in conjunction with the development of census mapping.

3.91. The publication and utilization of the geographic classification system developed for the census across the statistical system, national spatial data infrastructure, and administrative domains will enhance the value and utility of census data by improving its compatibility and integration with other information sources. The details of designing a general geographic classification, including the definition of the various areas of the geographic classification and their relationship to one another, are more complex than those involved in census geospatial activities and will not be covered further in this chapter.[70] However, the design of enumeration areas and the careful planning of their relationship to other census management areas is of crucial importance for the census and is outlined in the following paragraphs.

3.92. It is of critical importance to ensure that the boundaries of various administrative units are available in a geospatial format (i.e. shapefile), and frozen at least six months in advance of the census reference date so that no further jurisdictional changes are affected until the enumeration is over. This would be of considerable help in delimiting enumeration areas and minimizing chances of omission or duplication. To this end, the geospatial data infrastructure must be organized with nested hierarchical levels, while the more detailed geometry is a subdivision of the higher-level geometry without gaps or overlaps (e.g., state, region, province or municipality, district, neighbourhood, enumeration area). To be useful, it must be stored as geographies in geospatial datasets containing identifiers (geocodes) and their attributes.

3.93. If point-based geospatial data is collected in conjunction with each individual enumeration, then the dissemination unit can be easily conformed to different geographic layers. To say it another way, the availability of a precise coordinate for the location of each household can tie each individual enumeration response to a precise location where it was collected. Then there is a much greater flexibility in the eventual dissemination units, and those point locations can be aggregate up and associated data to any geography.

3.94. Although in this chapter we are dealing more directly with census geographies for census planning, organization, and management, they are not disconnected from census dissemination geographies, as many serve both purposes. In this sense, it is essential that there is a holistic view that

[69] In some cases, the geographic unit for collection slightly differs from the tabulation/dissemination geographical unit.

[70] For a full elaboration and details, please see *Handbook on Geospatial Infrastructure in Support of Census Activities,* United Nations publication, Sales No. E.09.XVIII.8, United Nations, 2009.

see geography as a tool for integrating data. It is important to consider that, in many countries, census data generate functional areas[71] (e.g., electoral districts, commuting areas, etc.) and that census geographies themselves are geostatistical functional areas. Functional areas form the link between data collection and implementation of actions. They can be used to visualize data, but also for analytical purposes and, if stable, trends over time.

3.95. Point-based (geolocated) information must be considered with the enumeration, dissemination, administrative and functional geographies so that aspects of spatial disaggregation, sample expansion and confidentiality issues such as spatial differentiation can be assessed.

3.96. A common set of geographies ensures that statistical data is geospatially enabled in a consistent manner and is capable of being integrated at the aggregate level; and ensures that users can discover, access, integrate, analyse, and visualise statistical information seamlessly into geographies of interest. This speaks directly to the GSGF Principle 3: Common geographies for dissemination of statistics. The goal of Principle 3 is to support the provision of a common set of geographies that ensure the consistent geospatial aggregation and dissemination of statistical data, irrespective of whether they are in gridded or administrative boundaries.

(i) Administrative hierarchy

3.97. One of the earliest decisions in census planning pertains to the administrative areas for which census data will be reported. Linking data to administrative geographies provides the spatial component that further enables data integration and broader comparability. Administrative areas can be any special geographic unit, but mainly they are units of administration, that is, some governmental authority has jurisdiction over the territory. Census preparation involves creating a list of all administrative and statistical reporting units in the country. The relationships between all types of administrative and reporting unit boundaries should be defined. Every country has its own specific administrative hierarchy, that is, a system by which the country and each lower-level set of administrative units (except the lowest) are subdivided to form the next lower level. Census operations provide an opportunity to verify and update these boundaries.

3.98. Only some of these hierarchical levels may have actual administrative roles; for example, the province, district and locality levels may have capitals with local government offices that are responsible for those regions. Other units may have statistical roles alone; that is, they are designed for the display of data and not for administering the territory. In some instances, however, administrative units may not be completely nested. Especially when considering both administrative and other statistical reporting units, the census office may need to deal with a very complex system of geographic regions.

3.99. In some regions, the establishment of a definitive list of units is a major operation because of difficulties arising from the frequent fragmentation, disappearance, or combination of small localities, and from changes in name, variations in spelling, the existence of more than one name for the same place or the use of identical names for different places. This listing should be held as a formal database or as an integral component of the databases forming part of a census fundamental geospatial

[71] Functional areas relate to, and support, the organisation and management of people, communities, society, and their activities in geographic space. For a full elaboration and details, please see The Global Fundamental Geospatial Data Themes, United Nations, 2019.

infrastructure.[72] This should be incorporated into basic census planning to allow for enough advance planning for the geospatial activities.

(ii) *Delineation of enumeration areas*

3.100. The design of enumeration areas should take various criteria into account. Correctly delineated, enumeration areas will:

a. Be mutually exclusive (non-overlapping) and exhaustive (cover the entire country);
b. Have boundaries that are easily identifiable on the ground;
c. Be coterminous with the administrative hierarchy;
d. Be compact and have no pockets or disjoined sections;
e. Have populations of approximately equal size;
f. Be small and accessible enough to be covered by an enumerator within the enumeration period;
g. Be homogeneous in aspects of morphology and function (surface area, land cover and use, rural and urban etc.);
h. Be planned as a (potential) common geography for dissemination
i. Be appropriate for other types of censuses and data collection activities as well.

3.101. For the delineation of enumeration areas, population size and number of households are generally the most important criterion, but surface area and accessibility must also be taken into account to ensure that an enumerator can service an enumeration area within the time allotted. Another factor that is sometimes taken into consideration is a hard-to-count index which assigns a score to each EA taking account certain population characteristics such as the proportion of immigrant populations, multi-occupied dwellings and other factors which make the enumeration difficult. Testing should be conducted to determine enumeration capabilities and capacity. Population density will vary between rural and urban areas and have an impact on the size of enumeration areas, since enumeration can proceed more quickly in towns and cities (densely populated areas) than in the countryside (sparsely populated areas). Given these considerations, enumeration areas that are significantly larger or smaller than the national average may have to be defined.

3.102. As a basic reference, before the delineation of enumeration area boundaries, the number of persons living in an area and their geographic distribution need to be estimated. Unless there is information from an exhaustive registration system or some other accurate and recent information source (administrative records, georeferenced address register etc.) these numbers need to be determined by counting the housing units, the extent of residential built-up surfaces, the number and the size of households. This is usually done through fieldwork, cooperation with sub-national governments or, more recently with the extrapolation from previous census results by means of aerial photographs, satellite imagery or derived products such as building footprint and settlement data. There are also freely available high-resolution population estimates produced using geospatial data and population from sub-national projections or surveys to approximate current population distribution. But this data should be used only if the accuracy is of acceptable quality.

[72] For further details on GIS mapping, see *Handbook on Geospatial Infrastructure in Support of Census Activities*, United Nations publication, Sales No. E.09.XVIII.8, United Nations, 2009.

3.103. Enumeration area boundaries need to be clearly observable on the ground. Even if they do not have considerable geographic training, all enumerators need to be aware of the boundary limits of the area for which they are responsible. Thus, population sizes between enumeration areas may be varied to produce an easily identifiable delineation. Natural features that can be used for this purpose are fences, landmarks, roads, railroads, creeks and rivers, lakes, boundaries, or any other feature that defines an observable area. If handheld equipment with embedded GNSS technology is used in enumeration, alerts through geofencing must be used.

3.104. Procedures for the delineation of enumeration areas should be developed that will allow, as far as possible, comparability of areas from one census to the next. Change analysis at the local level is greatly facilitated if the units of enumeration remain compatible between censuses. A unique code should be assigned to each enumeration area and the changes (for example in the case of splitting areas with a high level of growth) should be tracked. The statistical office is often the custodian of coding schemes in the country and should also be the focal point for the design of the census-mapping codes. In cases where this is not possible, the criteria can outline design principles that will allow users to easily compare enumeration area-based data across censuses. When the population and housing census are conducted separately, effort should be made to use the same enumeration area frame for both censuses as far as possible.

3.105. In addition to the limits and physical-natural features to support localization, the delimitation of enumeration areas must pay special attention to building blocks and block-faces, as these will be essential for the census taker to move within the EA and find the households. These spatial elements must also have their unique geocode following the spatial hierarchy. In censuses using handheld GNSS equipment, census takers' routes can be tracked to improve and update the internal mapping of enumeration areas, with potential use for national mapping.

3.106. In censuses that simultaneously use short and long questionnaires, special attention must be paid to the distribution of the statistical weights of the sample in the enumeration areas to guarantee adequate territorial coverage of the sample and not burden census takers. It is also necessary to ensure that the areas of dissemination of sample results, generally aggregations of enumeration areas, include all EAs without gaps or overlaps to guarantee coverage and expansion of the sample throughout the territory.

(iii) Delineation of supervision areas

3.107. Supervisory areas provide the means for a group of enumerators to be effectively managed. After delineation of enumeration areas, the design of supervisory maps is usually straightforward. Supervision areas consist of groups of contiguous enumeration areas that share some of the same characteristics as enumeration areas. The enumeration areas assigned to the same supervisor should be compact, to minimize travel times, and of approximately equal size. They should be included in the same field office area, which usually is defined according to administrative units.

(iv) Delineation of census management areas

3.108. Census management areas will consist of aggregations of supervisory areas brought together for ease of managing the enumeration staff. Where existing government staff and structure are used for enumeration purposes, the census management areas may be the same as the administrative regions. It

should be noted that this may be a matter of administrative convenience and the hierarchy (or way of combining enumeration areas into larger areas) for this purpose need not necessarily be the same as that for the dissemination phase, which must be driven by the needs of users. However, the greater the congruence between enumeration areas and pre-existing administrative boundaries the easier the task of conducting the census.

(v) Other census geographies

3.109. Progressively and with the use of geospatial technology and GIS, NSOs have been incorporating other geographies into the collection of census geographies, taking advantage of the ability to overlay spatial layers. These geographies are often used for tabulating results and presenting data, but their potential has increasingly been used for the planning and enumeration phases.

3.110. Many of these geographies arise from the need for the dissemination of results for differentiated population groups such as indigenous and native populations. These geographies are used for populations that do not have their territories recognized or delimited or are spatially distributed. The primary objective is to obtain significant statistical data for a recognized population within a specific geographic area encompassing a substantial concentration of members of this population group. Although, there are already experiences of using these geographies and GNSS positioning technology to geo-enable topics, avoiding questionnaires in which questions planned for specific population groups are incorporated into the basic or expanded questionnaire.

3.111. In other words, areas with characteristics of specific population groups are previously mapped and defined in office. It can be either incorporated areas – legally incorporated under state law, have a legally defined boundary – or census designated places – statistical geography representing closely settled, unincorporated communities that are locally recognized and identified by name. These areas may be composed of aggregations of enumeration areas or be contained within enumeration areas. The definition of the households that will respond to these specific questions will take place based on the assessment of the household's positioning in relation to the areas of interest for opening these topics[73].

(vi) *Geocoding*[74]

3.112. The coding scheme needs to be determined on a country-by-country basis, ensuring codes are unique, unambiguous, and should preferably be designed in collaboration with the national geospatial information agency. The most important principles in the design of a coding scheme are flexibility, expandability, and compatibility with other coding schemes in use in the country. As set out in Principle 2 of the GSGF, this includes the use of data management tools, techniques, standards, and good practices to facilitate the linking and management of geocodes within statistical datasets. This also serves to ensure that privacy and confidentiality requirements are correctly managed for the released data.

[73] Brazil has applied it for two population groups of the Brazilian Traditional Peoples and Communities, the indigenous people and the quilombolas in the 2022 census.

[74] This section draws heavily upon material in *Handbook on Geospatial Infrastructure in Support of Census Activities*, United Nations publication, Sales No. E.09.XVIII.8, United Nations, 2009 and in the *Global Statistical Geospatial Framework,* United Nations, 2019.

3.113. The coding of each housing or population unit to a specific longitude and latitude, or to the smallest geographic unit, preferably smaller than the enumeration area, allows for a flexible production of different geographic tabulations of outputs, and production of comparable area-based geography over time. Two somewhat different approaches are available for coding the location of housing or population units. The first approach is to code all units to the lowest-level enumeration area. The second approach, which permits finer geographic specificity, is usually based on some coordinate, address, or grid system. In either situation, careful planning should allow for maximum flexibility in future aggregation and disaggregation of the data and geographic units.

3.114. A digital geographic database organised in geospatial layers[75] in vector format consists of a structured set of points, lines and polygons, or grid cells. Each geographic feature — point, line, or area — has a unique identifier that is used by the system internally. This internal identifier is not usually accessible by the user and should not be modified externally. A more meaningful identifier is needed that can be used to link the geographic features to the attributes recorded for them. For the enumeration areas and administrative units, this link is the unique enumeration area or administrative identifier, which is listed in the master file of all geographic areas relevant in the census.

3.115. Indeed, a unique code needs to be assigned to each enumeration area. This code is used in data processing to compile enumerated information for households in each enumeration area and to aggregate this information for administrative or statistical zones for publication. This is the numerical code that provides the link between the aggregated census data and the digital enumeration area boundary database stored in the case of using GIS.

F. **Implementation of a census geospatial information programme**

3.116. The main goal of a census geospatial information programme is to structure a location-based conduct of census operations and activities and to explore the possibility of integrating statistical and geospatial information. The development within the census agency requires the coordination of a series of complex tasks with relatively long lead times. It is important that project plans are established to manage this process. In the implementation of a census geospatial information programme, NSOs must consider standardisation and interoperability of data which will lead to improved efficiency and simplification in the creation, discovery, integration, and use of geospatially enabled statistics and geospatial data. This will increase the potential application of a larger range of data and technologies, and thereby enable a wider range of information to be available and accessible for use in decision making. It will also facilitate better cooperation between all stakeholders producing and using statistical and geospatial information.

3.117. NSO must consider a long-term investment in geospatial information technology that can be fully applied in the statistical business processes. Principle 4: Statistical and geospatial interoperability of the GSGF defines the preconditions for statistical and geospatial data to work as a data ecosystem, in which those involved interact with each other to exchange, produce, and consume data. Interoperability between statistical and geospatial data and metadata standards, is needed to overcome structural, semantic, and syntactic barriers between data and metadata from different communities and providers. Furthermore, it is necessary to enhance the efficiency of discovery, access, and use of geospatially enabled data. Often, full interoperability of data first requires the removal of obstacles in country-level laws, policies, and organisations that hamper cooperation between stakeholders and create barriers

[75] Geospatial layers can also be in raster format, with regular grids of defined resolution.

between producers and end users. The UN-IGIF offers tools and pathways that can support NSOs in this regard.

3.118. The main activities to be reflected in implementing a census geospatial information programme are discussed below:

(a) Establish a geospatial unit. Geospatial information programme requires a specialized project team. With each round of censuses, these activities go beyond census mapping, including the collection of coordinates and addresses, spatial supervision of collection, the development of geoportals for dissemination, among other activities. Therefore, it is important to have an interdisciplinary team of geographers, cartographers, data scientists, statisticians and other specialists that allow for an integrated vision. When geospatial information incorporation and mapping activities are outsourced (see Section G below), this team will be responsible for specifying census requirements for products and coordinating agreements with the service provider.

(b) ***Developing a timetable*.** The critical date is the date on which geospatial inputs and maps must be delivered or made available to teams in the field. The geospatial information program should begin early in the census cycle to allow sufficient time to produce national mapping coverage well in advance of the census date and before field staff are trained. In censuses where handheld equipment and geoportals are used, sufficient time must be given for IT development. Hiring an updated high-resolution satellite imagery service is also an activity that must take considerable time to complete.

(c) Sourcing of basic mapping and digital geographic data. A major step in the mapping project is establishing a base map of the country, including digital mapping data if required. If a census-mapping project already exists, the agency may still require updates to their existing map holdings. In this context, it is important that NSOs see themselves within the geospatial data production ecosystem as an important stakeholder and seek exchanges and partnerships in the context of NSDI and the implementation of the UN-IGIF national action plan.

(d) Integrating statistical and geospatial information: The census geospatial information programme must be guided by the five principles and key elements of the Global Statistical Geospatial Framework (GSGF) to ensure the production of harmonised and standardised geospatially enabled statistical data and that the census results can be integrated with statistical, geospatial, and other information to inform and facilitate data-driven and evidence-based decision making for support local, subnational, national, regional, and global development priorities and agendas.

(e) **Disseminating information geographically:** It is essential that the census geospatial information programme includes geographic, cartographic, and geospatial products, platforms and applications resulting from the census. This topic will be discussed later, but planning must be carried out in an integrated manner. In general, maps and geoportals are powerful, simple language tools for interpreting data and information.

(f) **Continuing capacity development plan.** The NSO must keep an ongoing capacity development plan for census teams conducting a series of training in the use of geospatial information technology, GIS, integration of statistical and geospatial information, developing projects and participation in local and international forums and meetings.

(i) Sources and types of hard copy maps

3.119. Where a hard copy base map is to be used, official published maps may be available from national or provincial government mapping agencies, the local government, or municipal bodies. Other sources of maps may be other government agencies or private companies. Where the maps are obtained from sources outside the census agency, permission to use the maps collected must first be sought from the original source and any copyright issues addressed.

(ii) Digital mapping data

3.120. When establishing a digital geographic database, a major consideration is the determination by the census agency of data requirements. With increasing amounts of digital spatial data becoming available, it is also important that standards and a common data specification be produced to ensure data validity and consistency.

3.121. The key rules to be followed in selecting data items for inclusion are to question whether (a) the data item will be useful to enumerators in navigating their way around their enumeration area; and (b) the data item is relevant to users. Assessing the utility of data items to users in a census-mapping context must place significant emphasis on the user needs for small or customized areas. Data items that meet neither of those criteria should not be included in the database.

(iii) Updating maps or digital mapping data

3.122. Preparing or updating base maps, or the base map digital data, requires substantial resources. The final content of base maps will have a major bearing on the accuracy and completeness of enumeration area maps and, subsequently, the effectiveness of census enumeration. The updating of base maps should be scheduled according to priorities, based on areas in which changes to the number or characteristics of the people require the maps to be updated. Important features to be updated include (a) accurately named and presented roads and waterways; (b) administrative boundaries; and (c) landmark features, such as schools, place of worship, post offices, parks, and large buildings.

(iv) Operational design for enumeration and supervisory areas

3.123. Whether a hard copy or digital base is employed, an enumeration area design manual should be produced that contains the design criteria and the procedures to be followed when designing the enumeration area. The manual can be used as a basis of training for those involved in the design process.

3.124. If possible, enumeration area design should be conducted by regional statistical office staff who are primarily responsible for the enumeration areas in their province or region. This ensures that local knowledge can be utilized in the design process. A considerable part of the process is the gathering of information on where population and boundary variations have occurred to determine the best way to design and update enumeration areas. As an output of enumeration area design, a list should be produced that provides the enumeration phase with all relevant field data for each enumeration area, and the dissemination phase with relevant geographic data.

3.125. The design of field supervisor and management area boundaries can be determined at the completion of the process through the aggregation of enumeration areas, and the allocation of geocodes.

3.126. Quality assurance measures should be implemented to ensure that data are correct to a minimum standard, both for field navigation and for technical correctness in cases where a digital base is to be used as an output medium.

(v) Printing and content of field maps

3.127. Careful consideration should be given to the (considerable) time required for printing maps when establishing the project plan for census mapping. Digital options should be given primary consideration and accepted as the default format for any static reference maps. In this case, attention must be paid to IT development requirements and time.

3.128. Maps should be provided to every level of field staff. If paper maps are used, at least one map must be printed for every enumeration area in the country. It is recommended that two copies of the map be produced, one copy to be used by the enumerator and the other by the field supervisor for training and reference purposes. Both copies must be retained by the statistics office so that the information aggregated by the field team can be incorporated into the territorial base for the regular surveys and the following census cycle. The use of computer technology for data collection, such as tablets, laptops and other handheld devices with GNSS technology should be considered as the default solution and suitable for displaying maps available to field personnel.

3.129. Other considerations for the preparation of enumeration maps, also considering digital versions available on handheld devices, tablets or computers, include the following:

(a) Enumerators may be required to navigate in poor lighting conditions and thus details should be easily read;
(b) The maps must be easily interpreted with text and symbols readily identifiable and correctly placed, along with the information being presented in a standard format compared to other source maps;
(c) All maps must include basic cartographic elements such as title, legend, scale and a compass for orientation and navigation purposes.
(d) Boundaries (such as enumeration area boundaries) overprinted on the maps must be clear and unambiguous;
(e) Enumeration areas must be distinguishable when compared to the surrounding area;
(f) Folding or refolding of large paper maps (larger than A2 in size) is inefficient for staff;
(g) Paper and digital maps need to facilitate the addition of written enumerator comments;
(h) Production of the maps should be cost-effective;
(i) The maps should be suitable for reuse to meet dissemination purposes where this reflects user demands.

3.130. Maps for supervisors or regional managers should be of smaller scale, providing sufficient detail to identify major features. When using paper maps those should not be so large as to be difficult to handle easily in the field. In many cases, the use of inset or supplementary maps may be required if the map is to cover a relatively large area. For all levels of senior field staff, the maps should show the boundaries of all subsidiary units for which they are responsible.

G. **Contracting out and partnerships for census geospatial operations**

3.131. The development of a census geospatial information programme beyond rudimentary clerical systems requires considerable knowledge of mapping, cartography, and geographic systems. Furthermore, the census geospatial information program increasingly requires the integration of the resulting geospatial data with statistical data, as well as the incorporation of big data, machine learning and data sciences. If a census agency cannot draw on such skills internally, it may be required to contract out some or all the elements of the preparation of geospatial input preparation. A partnership with the national geospatial information agency (NGIA) can fully or partially meet the NSO's needs.

3.132. Mapping for field purposes under a contract or agreement basis requires the statistical agency to specify its requirements and prepare clear terms of reference to the contractor. These may include the following:
(a) acquiring the base map data;
(b) creating (or obtaining) the statistical boundaries and aligning them to the base map;
(c) providing a process for enumeration area designers to advise on changes to boundaries (and updates to associated spatial data);
(d) producing hard copy maps or soft copy using handheld devices such as a smartphone or tablet computer as specified for fieldwork; and
(e) ensuring that digital maps are compliant with standards (such as file formats that ensure compatibility, portability, and exchange).

3.133. The statistical agency should undertake the enumeration area design work and validation of the associated geospatial data, as well as take delivery of the hard copy maps for quality assurance checks and subsequent delivery into the field. The statistical agency must also accept full responsibility for the quality standards and delivery of the maps to field staff as required. After the census, any feedback received from enumerators about the base map should be communicated to the mapping agency.

3.134. In general, NSOs have progressively gained autonomy in census mapping, eliminating the need for full contracting of this service. The hiring of updated high-resolution satellite images, however, is a constant need in the production of census mapping and in the generation of input for digital maps and must be considered in census planning from the beginning. The acquisition of handheld GNSS equipment for collecting coordinates and tracking the census takers' route must also involve the participation of the geospatial team to guarantee the quality of the equipment and, consequently, the results.

3.135. The acquisition or incorporation of georeferenced databases such as address register, electricity company databases, mobile phone databases, administrative records, earth observation data can be very useful for the census geospatial information programme. In many countries, these databases are available in government agencies and can be incorporated through partnerships. In some countries, this data is commercialized by these government agencies or is privatized, requiring the expenditure of financial resources. The NSO must observe the need for this data and evaluate the most viable options within the national context.

3.136. There are good options for open-source and commercial GIS programs. In general, it is recommended to adopt open solutions, both due to the cost and the option of developing own solutions, with a vast library in forums, which can be complemented by commercial solutions when necessary. This assessment must be made jointly by the census geospatial information team.

3.137. Mapping for dissemination purposes may be more challenging because the outputs will involve representation of statistical information (with, or as part of, a map) and will often be accompanied by analysis or commentary about the information. Advances in mapping software have made it easier for census agencies to produce a wide variety of standard thematic maps (e.g., point maps; proportional symbol maps, choropleth maps, cartograms, etc.). However, advanced mapping products, as geoportals and dashboards, may require more expertise. In these cases, the NSO may first seek support from the NGIA and the country's National Spatial Data Infrastructure (NSDI) to find solutions. The United Nations Integrated Geospatial Information Framework (UN-IGIF) can be of great use in establishing this partnership between statistical and geospatial institutions. If this partnership is unfeasible, the NSO must evaluate the contracting of a service considering the quality of the geospatial information technology, knowledge transfer, internalization capacity, customization, statistical disclosure control procedures and the confidentiality of the respondents' information.

V. Census tests

3.138. The success of a census operation depends substantially not only on the effort spent in designing and building the census operations and infrastructure, but also on the effort spent on testing and piloting all components of the census design. Thus, having systematic testing procedures in place is vital for data quality and for the efficiency and effectiveness of the data collection operations. In order to identify problems and to suggest appropriate improvements, the use of pre-field tests (testing under laboratory conditions) and field tests (testing under field conditions), including a pilot census, is indispensable.

3.139. The testing of various aspects of a census plan prior to the enumeration is of critical importance for all countries, and an essential one for countries without a long history of census taking and for those in which fundamental changes in census methods or use of new technologies are being considered. Census tests can be designed for different purposes and in different ways. To yield full benefits, tests should be employed for all stages of the census, including enumeration, data processing, and methods for evaluating census results. Testing should be conducted in phases to build up to the full test of the holistic design, normally done with a pilot census. The testing phases should begin with testing individual components of the census design. For example, individual software components can be testing during a unit testing phase and the wording of questions for the census questionnaire may be tested in specialized content testing such as cognitive testing. The testing will build to eventually focus on the integration of all of the components together. Testing of the individual components, such as specific technological solutions (for example, internet data collection or the use of handheld devices), ensures that those solutions are working as intended and allows for efficient identification and resolution of errors specific to the individual solutions. However, the larger scale testing of the integration of all components is of utmost importance to ensure that all of the components work together successfully. Testing phases should be planned and conducted with both pre-field testing and field testing. The work to design all testing efforts, whether pre-field or field, should include the development of test objectives and measures. The test design should reflect those objectives and include the ability to collect the necessary data to demonstrate success, or failure, in meeting the predetermined objectives.

3.140. Much important testing can be conducted without requiring field work – pre-field testing. Technological solutions can be tested using data from previous surveys or censuses to conduct simulated data collection activities. If data processing methods are being tested, for example, those methods can be tested against existing data to determine if the processing generates the expected results. Using existing survey or census data for this purpose is especially helpful as the outcome is already known. In addition, cognitive research can be conducted in a laboratory environment with small numbers of participants to gain insight into how a respondent interacts with a questionnaire or how an interviewer interprets their instructions. Findings from these cognitive testing events are particularly valuable in probing for weaknesses in the questionnaire, in understanding of definitions and concepts, in the instructions, or in enumeration procedures that might affect the quality of the data. Cognitive testing should be designed to ensure that the subjects participating in the testing represent groups and populations of interest to ensure the results are relevant for all expected census respondents.

3.141. Field testing is an important stage to ensuring readiness of all components of the census design. Bringing all of the components together in field test creates the ability to identify problems linked to the integration of the data collection application design and architecture, the data transfer system, and the integrity and security of data transferred. Such tests can also give important information on the adequacy of the field organization, training programme, extent of respondent burden, processing plan, budget, and other important aspects of the census. They can be designed to provide information on the relative efficacy of alternative methods of enumeration and technology, and on the average time required for enumerating a single household or a single set of living quarters. Such information is useful in estimating staff and cost requirements. In addition, census tests serve as practical training for the nuclear staff of supervisors and other officials. If resource limitations make it impossible to test all components of the census design, the testing that is feasible should be prioritized based on those components that present the most risk to the full census, such as major changes or innovations to the census design, new or emergent technologies, or design elements that have not performed well in previous censuses.

3.142. When carrying out census tests, probability samples of geographic areas or units are not always necessary. Since the purpose of the pilot census and pretest is to judge the operational feasibility of a proposed course of action for the main census rather than make population estimates, purposive samples can often be used for such tests. Purposive selection of one or a few geographic areas is generally preferable for such feasibility testing. Purposive samples are also particularly useful when it is necessary to test particular census questions, concepts and methods in areas with particularly difficult conditions, or where population sub-groups or housing conditions of particular interest are predominant. On the other hand, random sampling procedures are often necessary, depending on the objectives of the test. For example, when overall quantitative measures are needed for comparing efficiencies of different procedures (for instance, in examining the anticipated response errors arising from different systems of enumeration), you may need to design the test to ensure adequate representation of the different systems in the resulting data.

3.143. A critically important test to be carried out during census preparations is questionnaire testing. The purpose is to test the suitability of intended census questions, including their formulation and the instructions provided, as well as the suitability of the questionnaire design. Such tests can be particularly helpful in assessing the suitability of the proposed material for enumerating specific population groups, as well as the general public. For example, these tests can determine if the terminology used in a particular question is familiar to potential respondents, or if the flow of questions is easy to follow for respondents of varying levels of literacy. These tests are also used for estimating the time requirements

in enumeration. It is practical to carry out questionnaire tests on a small scale in a laboratory environment or in several purposively selected places, for example, a location that is known to be populated with particular demographic groups whose interaction with the questionnaire is particularly important. Because they are relatively inexpensive, repeated rounds of questionnaire tests may be carried out until a satisfactory questionnaire has been evolved. Questionnaire testing can be conducted with a wide range of sophistication. They can be as simple as conducting cognitive testing in a laboratory environment where potential respondents interact with a paper questionnaire and share their thoughts and reactions to the content. Or, more sophisticated testing, especially with electronic data collection modes, can use leading-edge methods such as eye-tracking technology. The full range of methods is beneficial to design questionnaires more scientifically as reading patterns of respondents can be recognized technically through these types of tests.

3.144. One critical form of field testing is a final pilot census that is a final opportunity to bring together all aspects of the census design on a smaller scale to ensure the design is complete and can be scaled up in readiness for the full census. The results of the pilot census can be used to ensure accurate coverage of the target populations and to identify any gaps in the design that must be addressed before the full census. Such large-scale tests should ideally be designed and managed to thoroughly test all census components and the entire census infrastructure. Effectively, such a test serves as a dress rehearsal for the main event. Essential features of a pilot census are coverage of one or more sizeable administrative divisions and encompassment of the preparatory, enumeration and processing stages of a census, by which it thus tests the adequacy of the entire census plan and of the census organization. In order to best serve this purpose, care should be taken to ensure that conditions in the pilot census are as close to the conditions that would be present during the actual enumeration as possible. For this reason, it is often taken exactly one year before the planned census so as to conform to the expected seasonal patterns of climate and activity and to be conducted as close in time to the production census as possible. It is generally unwise to consider the pilot census a source from which to derive usable substantive data. Apart from the sampling problems involved, such a use inevitably detracts from the central purpose of the pilot, which is to prepare for the main census. Nevertheless, a pilot census can be an ideal opportunity for testing the tabulation and analysis processes necessary to generate statistical data products.

3.145. The area(s) where the pilot census is conducted must not, of course, be excluded from the actual census itself. However, it should be anticipated that when the full census is being conducted participants in the pilot census may be of the opinion that they already responded. So, appropriate publicity notifications should attempt to address this potential issue and include guidance to respondent to the pilot census that their involvement does not exempt their participation in the full census.

3.146. It is critically important to undertake a set of tests of the information and communication technology (ICT) solutions and production systems that are planned to be applied in the census. Depending on the extent and characteristics of ICT, these tests should include all ICT components related to the fieldwork and to data transfer, entry, and processing well ahead of the census itself. This is particularly important if a new technology is being introduced, such as the Internet, handheld devices or other electronic collection tools, and scanning the questionnaires as a means of capturing data. Tests should include the testing of applications, systems and the equipment itself, as well as the underlying circumstances necessary to avoid equipment malfunctioning, such as climate, or significant delays due to inadequate quality of paper causing paper jams or unexpected problems in programming activities. In the context of new approaches using electronic data collection, testing should include data transfers, and the relevant frequency of those data transfers, to the major depository of data and validation of

those data deliveries. Testing the efficiency of data entry rules, coding, editing and tabulation applications should be done based on results collected by the pilot census.

VI. Establishing the census frame

3.147. Prior to implementing a census, it is important to establish the frame or universe for the census. In most cases, this frame will consist of all living quarters and/or structures containing living quarters. The frame can be: (a) constructed before the census using listing activities in the field; (b) developed from existing address and/or person registers; or, (c) built from available information such as aerial imagery and GIS data files. Field listing activities generally require traveling all roads and neighbourhoods within the area to record the address or location of all living quarters and/or structures. This listing activity can be done with or without contact and interviewing of residents or representatives of the living quarters/structures, but contact is recommended to ensure the most accurate information about the living quarters/structure is recorded. Field listing is often conducted using paper address listing tools and paper maps, but can also involve automated tools.

3.148. For automated address listing activities, the enumerator/lister records information for every address or location in the entirety of the enumeration area (EA). It is in digital format, with a georeferenced list of addresses/locations and basic infrastructure/equipment and will be loaded onto the mobile device. Armed with a mobile device containing a pre-loaded list of the EA's addresses/locations, the enumerator begins by identifying the boundaries of his EA and then begins the listing activity within those boundaries. This means updating the list of addresses/locations they contain. Changes in the list are recorded on the map of the EA that has been given provided for this purpose. However, they may also be asked to update the addresses and locations in their EA by specifying their geographical coordinates using the mobile devices' GPS functionality.

3.149. The frame that is available at the start of the census is an instrument for the control of the enumeration, particularly in the absence of adequate and updated maps. The frame is also useful for: (a) estimating the number of enumerators and the number of schedules and quantities of other census materials needed in an area; (b) for estimating the time required for the enumeration; and, (c) for compiling provisional results of the census. It is also very useful for determining the boundaries of the enumeration areas and for establishing the necessary links between population and housing units if the two censuses are carried out separately. Finally, as the frame represents the full universe of living quarters or households that should be enumerated, it can be used as guide to monitor the completeness and quality of the enumeration in a given area.

3.150. Consideration should be given to the importance of permanent identifiers to streets and buildings, which can be used not only for successive censuses but also for other statistical purposes. A listing of sets of living quarters, particularly in densely populated places, cannot be made unless streets have names and buildings have unique numbers. Individual apartments in multi-dwelling buildings need to be numbered or otherwise unambiguously identified. Where these prerequisites do not exist, numbering immediately prior to the census would prove useful, but is likely outside of the authority of the census office.

3.151. Having a complete and accurate address frame ensures the ability to provide the enumerators with assistance in the form of lists of addresses to visit. Having the full list enables the enumerators to

recognize existing structures and identify new structures that may need to be added to the list and enumerated. Address lists will be essential if self-enumeration, where questionnaires are sent to the households by mail, is part of the plan. Some countries have and maintain address or population registers that allow more or less complete address lists to be generated relatively simply. The census can then not only use these lists, but also assist in further improving the population register by reporting any discrepancies found in the field. Where official population registers are not available, or insufficiently complete, it may be possible to obtain additional address lists from postal authorities, utility companies, or the private sector (for example, mail order companies). A definitive list for the enumerators could then be prepared by the census office by merging and rationalising the lists obtained from these various sources.

3.152. Where a functioning population register exists, it may be possible to prefill the household questionnaires with information such as the names of the persons expected to be members of a household that are already recorded in the register. This may reduce the response burden, accelerates the information-gathering process, and helps to pinpoint deviations. On the other hand, it might have a negative psychological effect if respondents believed that the authorities were monitoring them too closely. To combat this negative effect, the census organization should develop proactive messaging and notification to the respondent universe about the intent to use register data and clearly articulate the benefits of such an approach. Using one or several registers as the point of departure for a census that still includes full coverage field enumeration is an approach applied in some countries; differences between the register(s) and the field situation will necessarily come to light, and rules will be required to deal with such differences. Any differences in the data from registers and the data from field data collection can be caused by one of several reasons. They could be due to a difference in reference dates where the register has not been updated for some time prior to the census, or where there are differences in definitions and concepts, such as when the register records only persons who are “legally registered” at an address but do not necessarily correspond to the persons who are actually members of the household at the time of the census. An understanding of the cause of any such difference is critical to making appropriate decisions on which source to trust and use, and for ensuring that the merging of the records data and the field-collected data leads to accurate and meaningful census results.

VII. Field enumeration

3.153. Recent years have witnessed substantial changes in how countries are conducting a field enumeration, resulting mainly from the use of technology in census data collection. While the field-based method of enumerating the population with face-to-face interviews remains widespread, these recent changes show that the enumeration can be effected in different ways, including by the use of a self-completion paper questionnaire, either exclusively or in combination with electronic data collection and/or use of online collection via the Internet. Increasingly, more and more countries have been able to carry out their census electronically, making it possible to capture data automatically during the enumeration. On the other hand, self-enumeration methods can also be applied in different methods using the Internet. The use of new technology during enumeration creates a main challenge for many countries. It should be noted that only countries that have high penetration rates of information technology (including the Internet) have implemented online data collection, and mainly in conjunction with more field-based methods. However, these options may never entirely replace face-to-face enumeration, as even where societies enjoy a high degree of information technology use, the entire population cannot reasonably be expected to comply with such a mode of self-enumeration.

A. Method of enumeration

3.154. There are two major methods of enumeration: (I) the face-to-face (or enumerator) method, whereby information for each individual (in a population census) and for each set of living quarters and the occupants thereof (in a housing census) is collected and entered in the questionnaire (either in paper form or electronically via a tablet) by a census official; and (ii) the self-enumeration method, whereby the major responsibility for entering the information is given to a designated respondent in the unit being enumerated (usually the reference person of the household) and where the questionnaire is usually distributed, collected and checked by an enumerator, or mailed out and mailed back, or completed via the Internet.

3.155. The face-to-face method relied on paper questionnaires in most countries until recently. However, electronic collection via tablets has become the dominant method and the new norm in many developing countries. Paper questionnaires might still be needed in these countries for parts of the population or areas with limited electricity and internet connectivity. For self-enumeration, countries with up-to-date address/population registers and high literacy rates traditionally utilized the approach of mailing out census forms and having the public mail them back completed. Recently, Internet data collection has become a common method for self-enumeration, especially in countries with widespread telecommunication services. Some countries use a combination of Internet data collection and postal distribution of the questionnaire (with or without return postage) for self-enumeration. Self-enumeration approaches, using methods like mailing, telephone, and internet data collection, can also be combined with face-to-face interviews. All these methods can be employed independently or combined with verification by a census official.

3.156. Whatever approach is used, the complete enumeration plan should be prepared and adequately tested and piloted well before the enumeration begins. This involves: (a) the determination of the enumeration method(s) to be used, and whether there are differences by specified geographic area and/or population groups; (b) the basic procedures to be followed in the collection of the data and the control of the enumeration; (c) the procedures for quality control of the data; and (d) an estimation of the number of sets of living quarters and the probable size of the population to be enumerated so that the number of questionnaires, handheld devices and other materials required for the enumeration, and the number of enumerators and supervisors needed, as well as the cost of enumeration can be properly ascertained. If more than one enumeration method is offered, it will have to be decided whether the methods will be offered in phases or simultaneously. A phased approach may be used if the level of self-enumeration is expected to be significant.

3.157. Each method has its own advantages and limitations. The face-to-face method is the only method that can be used in largely illiterate populations or in other population groups that may be unwilling to complete the census forms themselves or find it difficult to do so. On the other hand, in countries where literacy is virtually universal and educational attainment relatively high, the self-enumeration method may often yield more reliable results at substantially lower costs, particularly if Internet data collection or a mail out/mail back procedure can be used. However, the postal services may be used to distribute the census forms only when a comprehensive and up-to-date list of addresses is available or can be prepared.

3.158. Another consideration is the emphasis to be placed in the census on obtaining responses, whenever possible, directly from the person concerned. The self-enumeration method allows for - and its instructions may encourage, at no extra cost to the census organization - consultations among family

members when they complete the census form. In contrast, with the face-to-face method it may be prohibitively time-consuming to encourage enumerators to attempt to interview persons in addition to the designated responsible adult[76] in order to collect the requisite information on the other household members. In the light of these considerations, it may sometimes be desirable to rely on one method for enumerating most of the population and to use another method in certain areas or for special groups of the population.

3.159. With the advance of information technology, the penetration of the Internet has increased in recent decades. In these circumstances, it is recommended that Internet collection methodology should be explored depending on national circumstance. This method can be cost-effective as the expense of printing questionnaires and the field staff wage bill can be greatly reduced. Also, self-enumeration through the Internet can help secure the privacy of respondents. However, a combination of face-to-face and Internet modes can result in duplication during the enumeration. Therefore, careful consideration of the management of the dwelling/household list is essential. Overly complex designs should be avoided, and adequate quality checks introduced to avoid duplications and frauds.

3.160. The decision regarding the method of enumeration to be employed should be taken at an early stage on the basis of thorough testing of the various alternatives in terms of the costs involved, the quality of the data produced and their operational feasibility. Even where a method has been followed traditionally, it is best to periodically reassess its relative advantages in the light of current census needs, changing techniques, and emerging technologies. An early decision is required because the method of enumeration used affects the budget, the organizational structure, the publicity plan, the training programme, the design of the questionnaire and, to some extent, the kind of data that can be collected.

3.161. Challenges that affect or hinder the ability to achieve a response should be fully considered when developing the enumeration design and methods. The design should, as far as possible, reflect particular activities or actions that seek to increase the likelihood of receiving a response. Challenges can be grouped into two types: people and physical challenges. People challenges tend to focus on particular subgroups of the population that tend to be difficult to enumerate for a number of reasons. Physical challenges are those related to the type of environment in which the people live. The two are not necessarily mutually exclusive, but each require different methods to overcome the difficulties that they present.

3.162. The section on population sub-groups that are difficult to enumerate (see paragraph 5.47) gives a thorough overview of the more common examples and how such difficulties might be addressed during the enumeration. Population groups that are typically difficult to enumerate include people with language difficulties, nomads, migrants, students and older persons.

3.163. ***People with language difficulties*.** Not all respondents will speak or understand the language(s) in which the census is being conducted. Therefore, consideration needs to be given to the provision of translation services and materials, taking into account the types and concentrations of languages required. To the extent possible, decisions around languages/translations to support the enumeration should be data driven to avoid any perception of bias or discrimination.

[76] In some countries the census legislation prescribes the person who is to be legally responsible for providing the information.

3.164. ***Nomads*.** To successfully carry out the enumeration of nomads, it is particularly necessary to pay full attention to the preparatory work in order to determine suitable enumeration techniques. It should be pointed out that there is no absolute recommended methodology for the enumeration of nomads since conditions will vary from country to country. The particular method suitable for a country undertaking to enumerate nomads as part of the census should be determined only after a detailed preliminary study and after undertaking field tests.

3.165. Some of the methods used to enumerate nomads and semi-nomads may be classified as follows:
(a) In the ***group assembly approach*** nomads are asked to assemble at particular interview sites on certain fixed dates. This method can be adopted only through the administrative or tribal authorities.
(b) The ***tribal or hierarchical approach*** is a favourite method, since the nomads usually follow what is dictated by the tribal or hierarchical chief. The enumeration can be carried out as a kind of administrative census by contacting the tribal chief and collecting, sometimes from memory and sometimes from a register, all the needed information on the chief's followers. The other approach is to contact those followers with the assistance of the chief or a representative and to collect the necessary data directly from the household. In this case, the unit of enumeration is not areal but tribal.
(c) The ***enumeration area approach*** presupposes creating conventional census enumeration areas and then contacting each nomadic household that happens to be staying in the enumeration area during the census.
(d) In the ***water point approach***, a list of all water points available to the nomads during the period of enumeration is prepared. Since numerous temporary water points are created during the rainy season, a meaningful list of water points may be prepared with reference only to the dry season. The enumerator is given the task of locating and visiting every nomadic household that may be using a certain water point.
(e) In the ***camp approach***~~,~~ a list of camps is prepared together with the approximate location of each within the country, and enumerators are sent to visit all the households in each camp.
Sometimes a combination of two or more of these methods may be appropriate.

3.166. ***Migrants*.** Effective communication is essential for engaging immigrants (particularly new arrivals) in the census. This can be done by providing materials and assistance in multiple languages, working with community organizations that work with migrants, and addressing concerns about immigration status. As part of developing the enumeration design consideration needs to be given to communicating with these groups, particularly about the benefits of the census, to ensure that they are more likely to respond. For example, materials can be translated into the languages that are most commonly spoken by migrants in the area. Community organizations can help to spread the word about the census and encourage migrants to participate. And enumerators can be trained to answer questions about immigration status and to assure migrants that their data will be kept confidential. Flexible enumeration methods, such as online questionnaires, phone interviews, and in-person visits, can also be used to accommodate the diverse needs of migrants. And ensuring that enumeration venues are safe and comfortable will further contribute to a positive experience for migrants.

3.167. ***Refugees and internally displaced persons (IDPs) in camps***. In many censuses there are several questionnaires, including one for conventional/private households and another for institutional or collective living quarters. These institutions often comprise camps for IDPs or refugees. Institutional questionnaires may not be suitable for camps for the displaced, particularly where these facilitate more than short-term arrangements. As the question content is often much shorter in the institutional questionnaire than in the household questionnaire, the information will be insufficient to provide for

the requirements of users interested in displaced people. Wherever possible the camps for refugees and IDPs should be enumerated using the questionnaires designed for private households. The census communication plan should include messages to convey that the census will include displaced people, refugees and IDPs and that their responses are important and to assure the confidentiality of data collected. It is important to ensure that communication is available in languages relevant to displaced populations, as refugees and IDPs may speak a different language than the local population. The Compiler's Manual[77], developed by Expert Group on Refugee, IDP and Statelessness Statistics (EGRISS) primarily aimed at technical personnel in NSS who want to include forcibly displaced populations in official statistics, provides additional insights. Use Case on Censuses discusses scenarios relevant to producing official statistics on refugees and IDPs.

3.168. ***Students*.** Tailored approaches are needed to reach students, such as collaborating closely with schools, colleges, and universities. This can help to ensure that students are aware of their legal obligation to participate in the census and that they have the opportunity to do so. Clear and concise communication materials that resonate with student life and their priorities are essential. For example, materials could be developed that focus on the benefits of the census for students, such as how it can help to improve education funding and opportunities. Appointing student ambassadors can also be an effective way to promote the census within student peer groups. Leveraging online platforms, social media, and mobile applications is crucial for engaging technically aware students, who are more likely to respond through digital channels. For example, the census could create a dedicated website or social media page for students. And students could be encouraged to share information about the census with their friends and classmates through social media.

3.169. **Older persons**. Additional support and tailored strategies are essential for engaging older persons in the census. Effective engagement includes providing supplementary materials and support mechanisms to assist with the census process (for instance, a national events calendar may be provided to assist older persons to recall or estimate their age). It can also involve supporters, such as family members, village elders, or residential home staff, in the enumeration process. For example, family members could help to translate census materials or to answer questions from elderly relatives. Enumeration venues, where used, should be accessible to all, including those with mobility limitations. They should also provide a comfortable environment for older participants. Assistance could be provided through call centres as well. [78], [79]

3.170. During the development of field enumeration methods, several physical challenges need to be considered. These include: access-controlled properties such as gated communities, apartment buildings with secured entrances, or private residences with restricted access which can pose challenges for enumerators trying to reach residents; rural areas that are sparsely populated areas with limited infrastructure, long distances between residences, and difficult terrain can complicate the logistics of enumeration; communal institutions like prisons, hospitals, or military bases often have specific protocols and access restrictions that require special arrangements for enumerating residents; multi-occupied dwellings such as buildings with multiple dwellings (e.g., apartments, duplexes) can make it

[77] https://egrisstats.org/activities/compilers-manual/

78 Comisión Económica para América Latina y el Caribe (CEPAL), "Recomendaciones para los censos de población y vivienda en América Latina. Revisión 2020", Documentos de Proyectos (LC/TS.2021/150), Santiago, 2021.

79 Comisión Económica para América Latina y el Caribe (CEPAL), "Recomendaciones para los censos de población y vivienda en América Latina. Revisión 2020", Documentos de Proyectos (LC/TS.2021/150), Santiago, 2021.

challenging to ensure all residents are counted, especially if there's no easy way to identify separate living units; and, areas that are insecure/violent which may pose a risk to enumerators.

3.171. ***Access-controlled properties*.** Some properties, establishments, communities or compounds have controlled access, presenting difficulties in gaining access to undertake an interview or to deliver or follow-up on a questionnaire. Access control mechanisms may include locked gates with an intercom to each individual dwelling, commercial farms where the gates are unmanned and are some distance from the dwelling units or gates or doors managed by a concierge or security guards. When developing enumeration procedures, advice needs to be given as to how to gain access and actions to take if access proves difficult. Some of the activities may include building a relationship with the owner of the properties to approve access to engage with residents; using the postal service to deliver questionnaires to these properties; and additional communication methods (such as a letter informing residents about the census and how to complete their questionnaire or inviting them to arrange a particular time to complete their questionnaire via interview. The use of mail out/mail back or online questionnaires that do not require the intervention of enumerators can be employed to resolve this problem.

3.172. ***Rurality*.** Understanding the extent of population in rural areas and the associated logistical and management challenges with running a collection exercise in these areas needs careful consideration. Lack of infrastructure such as access to electricity and Internet connectivity may limit the modes of data collection that could be used in these areas.

3.173. The enumeration of other groups, such as the ***homeless*** and ***institutional populations***, also requires special consideration and tailored approaches. Due to their transient nature and lack of fixed addresses, traditional enumeration methods might be ineffective for the homeless. Solutions for the homeless include collaboration with homeless shelters, soup kitchens, and outreach programs to establish trusted points of contact; and deploying enumerators with experience and sensitivity to engage with this population. Solutions for institutionalized populations might include establishing partnerships with institutional administrators to facilitate data collection; utilizing administrative records from these institutions with proper consent and privacy safeguards; and, deploying trained enumerators to conduct interviews within the institutions, following established protocols. These targeted approaches can help ensure a more complete and accurate count.

B. Multi-mode data collection

3.174. Countries are becoming more innovative in how they conduct population and housing censuses in order to improve response rates while also reducing census costs. One of these innovations is the exploration of alternative approaches to the traditional method of face-to-face canvassing of the whole country to enumerate the total population during a census enumeration. The introduction of these alternative approaches entails also use of a combination of modes of enumeration or multi-mode collection of census information. This involves the use of two or more data collection methods to collect census data. These modes include various combinations of (i) interviewer-administered modes, such as face-to-face interviews or interviews over the telephone, and (ii) self-administered modes, whether online or by mail.

3.175. Interviewer-based data collection modes include: the traditional face-to-to-face personal interview with paper and pencil (PAPI), computer assisted personal interviewing (CAPI), and computer assisted telephone interviewing (CATI). On the other hand, self-administered data collection modes include paper questionnaire with self-enumeration (PASI) and computer assisted (online) self-

interviewing (CASI). For more detailed information on these modes of data collections, see Chapter E of the *Guidelines on the use of electronic data collection technologies in population and housing censuses*.[1]

3.176. Application of multi-mode data collection entails the uses of two or more of the modes of data collection listed above, either concurrently or sequentially. In concurrent multi-mode data collection, respondents can choose between one of two or more modes, such as completing a paper questionnaire or providing the information online via the Internet. In sequential mixed-mode data collection, all respondents are first requested to provide information in one mode but then, should they not do so, are offered other modes to increase the response rate. Sequential mixed-mode designs start with the mode that is least costly to implement and progress to more expensive and persuasive modes. Both options aim to improve coverage/response rates and completion rates, both for the general population and for special groups. It should be noted, however, that while mixed mode data collection methods are effective at reducing non-response, they could lead to measurement error due to response bias arising from the use of the various modes.

3.177. Multi-mode data collection offers several benefits over the traditional single-mode of data collection. Firstly, offering different modes of data collection, to different populations, representing a broader cross-section of the population, has the potential to increase coverage/response rate, improve data quality, and reduce the cost of the census. For example, some individuals may prefer to complete the census questionnaire online, while others may prefer a telephone call, or a paper-based questionnaire. Secondly, providing various modes of data collection will enhance the ability to reach respondents who are difficult to enumerate with the use of a single mode of data collection, such as persons living in remote areas or those with no Internet access. This inclusiveness has the potential to improve the coverage and representativeness of the collected data. Also, multi-mode data collection can help to reduce field costs particularly when some of the tasks are automated, such data entry through electronic data collection.

1. Management of multi-mode data collection operations

3.178. There are some potential challenges to the use of multi-mode data collection. It is more complex to implement than it is for single-mode data collection. This is because multi-mode collection requires more intricate planning and coordination than a single mode of data collection. Also, once data has been collected from multiple modes, it needs to be integrated into a single dataset, and this can be challenging, as different modes of data collection may produce data that may not be directly comparable. When merging census data from different collection modes, it is important to be aware of the potential for bias. This can be due to a number of factors, such as differences in response rates by mode of data collection, and differences in the wording or routing of questions. As a result, comparability of data from various modes of data collection may be a challenge due to "mode effect"[80] and has implications for the quality of data that has been collected. Understanding mode effects on coverage, response, and measurement can help develop methods to minimize them, and to minimize mode-associated bias.

3.179. Multi-mode data collection often relies on technology, such as the Internet and mobile devices, and this can be a challenge in areas with limited access to such technology. Use of multiple modes of data collection can be more expensive than single-mode data collection, especially if new technologies

[80] See Chapter E of the Guidelines on the use of electronic data collection technologies in population and housing censuses. https://unstats.un.org/unsd/demographic/standmeth/handbooks/guideline-edct-census-v1.pdf

are used. It is worth noting, also that not everyone has access to the Internet or other technologies needed to participate in online data collection. This can lead to under-coverage of certain groups in the census data. Despite these challenges, multimode data collection is becoming increasingly common. This is because it offers several advantages over single-mode data collection, and offers the possibility to compensate the weaknesses of one mode with the strengths of another. By carefully planning and implementing multimode data collection, census bureaus can improve the accuracy and completeness of collected data.

3.180. The management of multi-mode data collection operations is a difficult task due to the complexity of designing questionnaires for interviews and for self-completion, monitoring of enumeration, as well as the validation and integration of data collected by different modes. The more options that have been employed for data collection, the more complex the operation becomes. Consequently, an integrated management information system should be set up for executing, monitoring, and reporting functions related to operational procedures, including scheduling, contact management, case status management, avoiding duplication of responses, managing switching/transfer of a case between data collection modes, response management, assignment of field staff, and response integration. Furthermore, a dedicated management team is needed to monitor and control the multi-mode collection operation.

2. Designing a census multi-mode data collection

3.181. A crucial part of the implementation of a multi-mode census data collection relates to decisions on the modes of data collection and how they are to be implemented in a national data collection. It is the case that such decisions should be based on national circumstances given the diversity of national experiences and capacities to successfully implement the various options related to census multi-mode data collection. Hence, there is no “one-size-fits-all” option that can be recommended for international use.

3.182. However, regardless of the modes included in the combination for a multi-mode census data collection, there are pertinent factors that should be considered in its design. First is the need to have a clear understanding of the target population as well as its characteristics, and how appropriate the modes under consideration can easily be implemented. In addition to national differences in abilities to implement the various combinations of modes of data collection, it is also the case that even within the same country there may be variations in ability based on the levels of literacy and education of different population groups as well as differences by geographic area of residence of the population. For example, self-enumeration by mail or by use of the Internet may not be appropriate where a large proportion of the population is not literate, as well as for those living in remote areas of a country and in other areas where access to the Internet or to postal services may be very weak or non-existent.

3.183. As an input into the choice of modes to be included in a census multi-mode data collection, the census agency, therefore, will need detailed information about the characteristics of the national population, such as age, education, and location. This would enable the selection, development, and tailoring of multiple modes of data collection to include only those modes that can easily be implemented in the country. A further step would be to develop a mode assignment strategy, or for modes to be concurrently assigned to different sub-groups based on well establish criteria on the ability of the corresponding population groups to easily and accurate provide the information that is being requested. In a sequential design, the census office needs a strategy, not only on the various modes of

data collection to be offered, but also on the order that the different modes will be offered in order to make it convenient for the population to take part in the census.

3.184. Another issue to which attention should be given in designing a multi-mode data collection is ensuring that the questionnaire or data collection instrument is designed in a compatible way across the modes of data collection to be used. For example, question wording and format, response options, as well as the layout and design of the data collection questionnaires should be consistent and comparable across modes to be used. This would enhance the comparability of the collected data across the various modes by minimising response bias.

3.185. Designing multi-mode data collection entails much more than is the case when just a single mode is adopted. A multi-mode data collection strategy may offer respondents the flexibility to switch between modes during the data collection process. This allows individuals who start a questionnaire online (e.g., via the Internet) to complete it using a different mode, such as paper or telephone, if needed (e.g., in cases when online questionnaires are not wholly completed). If this option is offered, the census office should develop a strategy to monitor and manage switching/transfer of a case between different data collection modes during a census. A strategy is also required for handling of cases to ensure completion of entry of information where, for example, a questionnaire may not be completed in one go, particularly when the household has several members.

3.186. It is, therefore, important to test and pilot the modes of data collection process using the intended modes in order to identify any problems and make adjustments to the questionnaire and the data collection process. For example, the pilot census would help ascertain the compatibility of the modes of data collection and the collected data, as well as to assess whether the questionnaires in the different modes are easy to understand and complete.

C. Use of administrative data during field-based enumeration

3.187. The use of administrative data can facilitate the enumeration in a field-based census in a number of ways. Pre-filling electronic census questionnaires is one of the ways in which administrative data could improve enumeration. Such an approach can reduce the workload of enumerators (e.g., reduce data entry for basic demographic information). Leveraging existing administrative data can potentially improve accuracy by reducing the reliance on respondent recall for basic demographics. When data items are taken from administrative sources to pre-fill the questionnaire, the data should be checked or validated by the respondent. It's crucial for the census office to establish clear protocols for instances when there are inconsistencies between administrative data and respondent information. However, this approach may have negative effect on respondents that have privacy concerns.

3.188. In certain cases, necessary information for particular subpopulations, such as the members of institutional establishments (such as students or workers dormitories, old people's homes, assisted-living facilities and welfare institutions, military barracks, correctional and penal institutions, religious institutions) may be available from the administrative data sources. Using these data has the potential to effectively mitigate or substantially decrease the extent and complexity of field operations, especially when direct data collection for a subpopulation is problematic.

3.189. There are also cases where availability of administrative data assists in facilitating the field operation in other ways. For instance, as noted earlier, sometimes prior to the field operation, households that are likely to be missed by enumerators, such as the ones that have recently relocated,

or comprise older persons or migrants, or that are located in remote or hard-to-reach areas, or within indigenous communities, can be identified in order to devise appropriate strategies to cover them. Additionally, administrative data may be used to determine whether an address is occupied or not, which will assist in planning follow-up visits to non-respondents.

3.190. Furthermore, when administrative data sources provide an accurate aggregate estimate of the number of housing units/households/individuals in a given area, this information can be used as benchmark to monitor and quality assure actual performance in the enumeration.

3.191. In the combined census approach, data for specific high-quality variables can be directly obtained from administrative sources, eliminating the need to collect them from individuals during field enumeration. This is particularly beneficial for: sensitive topics for which administrative data can provide information on culturally sensitive topics like infant mortality without requiring direct questioning; and confidential information such as data on income or other confidential matters obtained from sources like tax records, reducing respondent burden and potential disclosure risks.

D. Timing and length of the enumeration period

3.192. The choice of the time of year in which the census is to be taken is of great importance. The main consideration should be to select a period in which the census is likely to be most successful and to yield the most useful data. This may depend on a number of factors. Firstly, it is necessary to avoid those seasons in which it will be difficult to reach all inhabited areas because of rains, flooding, snow and so forth or in which the work will be particularly arduous, as is the case during extremely hot weather. Secondly, a time should be chosen when most people are staying at their place of usual residence; such a choice will simplify the census operations both in a *de jure* and in a *de facto* enumeration, and it can make the results of a *de facto* enumeration more meaningful. Seasons of peak agricultural activity should be avoided because it is difficult to interview persons who work late every day and who may even stay nights on their land if the land is far from home. Great traditional festivals, pilgrimages and fasting periods are also unsuitable times for census work. Since in many developing countries the bulk of the field staff is recruited among schoolteachers and older students, the conduct of the census may be feasible only during school holidays, though, as already indicated, the days of major festivals should be avoided. On the other hand, enumerating tertiary students in vacation time is problematic as many will not be present at their place of usual residence (their term-time address) and are likely to be undercounted. Another criterion to consider in the timing of the field work is to avoid the period when hours of daylight are at their minimum, as interviews should, preferably, not be conducted at night.

3.193. In a country that includes areas of sharply contrasting seasonal patterns of weather or activity or in which potential census personnel are in very short supply, it may be necessary to enumerate different parts of the country at different times or to enumerate the nomads or other special population groups at a different time from that established for the settled population. This, however, is generally not a very desirable solution because nomads cannot always be clearly differentiated from settled population because there may also be mobility among the settled inhabitants. Furthermore, such a solution creates complications with respect to the use of the census data due to the varying currency of the data. The census design should consider these complications and risks and attempt minimize negative impacts to or biases in the data.

3.194. When a census has been undertaken and the census date is found to have been, on the whole, satisfactory, subsequent censuses should be taken using the same time of the year, unless there are

strong overriding reasons for not doing so. A regular census date enhances the comparability of the data. The tradition of a fixed census period within the calendar also provides administrative discipline, motivating and enabling all those involved in the census to make necessary preparations in a timely manner.

3.195. It is desirable to keep the census enumeration period as short as possible in order to avoid double counting and omissions, which can occur in spite of a single reference date. On the other hand, the shorter the enumeration period, the greater the number of field staff that have to be recruited, trained and supervised. This increases the cost and may lower the quality of the data. If the census is using automated methods for in-person enumeration, the increase in costs is further exaggerated by the need for additional devices for the additional staff. How these different considerations should be reconciled depends on the size and nature of the country and on the resources at its disposal. The length of school holidays is sometimes a restricting factor, although governments of several developing countries, recognizing the national importance of a census, have prolonged the school holidays in the census year in order to allow teachers and students to work on the census as long as required.

3.196. While it is desirable to design the census enumeration period to be of optimal length and to conduct the enumeration at the optimal time of the year, the design must also include flexibility and contingency measures to enable adjustment in the case of an unexpected emergency or disaster. If an unexpected event makes it impossible or impractical to conduct the enumeration period as planned, an alternative timing or length must be implemented. In this case, it is imperative to attempt to understand and quantify the possible impacts to data quality and the comparability of the census data over time. Lengthening the enumeration period introduces the risk of recall bias, duplicate enumerations, and/or omissions. Changing the timing of the enumeration period can increase the risk of omissions if some populations (such as migratory workers) are no longer accessible to the field staff. It should be noted that any change to the census date will usually require an amendment to the relevant legislation, which itself can be a lengthy process.

3.197. In recent censuses, most developing countries have allowed about one week to 10 days for the training of enumerators, while the enumeration period has generally varied from a few days to two weeks. In the case of using a self-response method through the Internet, the enumeration period can take longer, and enumerators may require less training if they are simply going to follow up on non-respondents. Short periods are often feasible in small countries while longer periods may be necessary in large countries with poor communications, widely dispersed populations, or where multi-mode data collection strategy is adopted.

3.198. One method sometimes used to allow sufficient time for enumeration and yet make the census simultaneous is first to enumerate the population over a longer period, say a week or more, and then, on one single day, to revisit all households, deleting and adding persons as needed to update the census return. This procedure is, however, not practicable in very sparsely settled areas and can be very costly to deploy a large number of staff on the single re-enumeration day. It's important to note that utilizing a post-enumeration survey can eliminate the need for this second visit for the purpose of assessing coverage errors.

E. Management and supervision

3.199. A comprehensive and elaborate management system is necessary for resource management and providing timely managerial advice to the field staff. There is no unique approach for such a system;

however, it is important to develop a hierarchical and geographically dispersed system to build a direct and effective communication mechanism between two key groups: field staff (this includes enumerators, assistant census officers, and census supervisors) and central management (this refers to the permanent managers within the statistics office). The system should facilitate a clear distinction between these groups while ensuring a smooth flow of information and support.

3.200. Adequate supervision of the enumeration is essential for ensuring the quality of the field enumeration. Many countries use a field supervisor to enumerator ratio ranging from 1:10 to 1:15 for field enumeration. Periodic control of the quantity and quality of the work accomplished by enumerators and other field staff is recommended, in order to facilitate the correction of inefficiencies and to maintain satisfactory progress during the enumeration period. Periodic and systematic assessment should be carefully organized to ensure the quality of the work and also for collecting appropriate information about the the progress of the enumeration for management and supervision of the fieldwork.

3.201. Each staff member involved in the management and supervision system should have a clear role and responsibility and should be fully trained for possible problems occurring during field enumeration and their solutions. For an efficient system, it is important to give clear instructions to the field staff on how to perform their own responsibilities. It is important to note that methods and technologies used during enumeration have a direct impact on the roles of managers and supervisors; consequently, a complete understanding of the characteristics and operational aspects of both the enumeration method and the enumeration technology is a prerequisite for efficient supervision of the enumeration component of the census.

3.202. Depending on the communication facilities and other infrastructure available in a country, different mechanisms for exchanging information among managers and field staff need to be developed. These mechanisms are important for ensuring consistent dispatch of field instructions and also sharing best practices, particularly for finding solutions to unexpected problems during the field operation. The use of mobile phones and accompanying technologies, such as Short Message Service (SMS), Unstructured Supplementary Service Data (USSD) and other enterprise level communication platforms significantly increases communication capabilities, more especially when such communication needs to broadcast simultaneously to allow the field staff getting one message from the source.

a) Management information system

3.203. A management information system for the field operation should be established to collect information needed for timely management and supervision of field operations. To establish this system, the following steps can be considered:

i. Determining information needed for supervising and managing fieldwork;
ii. How and when each piece of information/data should be collected;
iii. How and by whom each piece of information/data will be used.

It is important to collect the amount of information that can reasonably be assimilated with good quality and used effectively, otherwise every additional datum with low priority will affect the cost of collecting the requisite information.

3.204. The advent of real-time data analysis in censuses has resulted from the technological advancements and practices that enable the collection, processing, and dissemination of information on field activities in real time as they happen, or at least as soon as they are recorded. This concept has

gained significant importance of late due to its potential to provide immediate insights, enhance decision-making, and improve overall efficiency.

3.205. The following information can be collected through such a system:

i. Information about particular activities that are implemented before the enumeration, such as establishment of local census commissions and training of census field staff;
ii. Information about the field staff needed for administrative tasks, such as recruitment and hiring field staff, bank account information for payment, work accomplished, dynamic fieldworker payment advices especially when utilising electronic data collection methods;
iii. Progress of enumeration of population and housing units to evaluate whether or not the field operation proceeds according to schedule;
iv. Information about logistics issues, such as the shipment, storage and maintenance of census materials and questionnaires, timing of receiving and sending materials, and quantities and types of materials.

3.206. Census field operations can be made more efficient through the availability of a management information system and use of this system by field staff for administrative tasks and supervision. It is possible to create a quick communication mechanism for key messages and work allocation. This system should be used for producing and submitting regular reports providing information about the progress of field activities and enumeration. There are several ways of collecting such information. Technology-based solutions include the use of SMS, USSD, websites and portals, and mobile or handheld applications.

3.207. Paradata which is data automatically collected about the census field process captured during computer assisted data collection can become a vital source of information for the management information systems. Examples of these include interviewer observations, keystrokes, data time stamps to record responses and various other data captured during the process. This information is increasingly being used in real time for both monitoring purposes and to manage the process of collecting large amounts of data

b) Supervising the enumeration

3.208. A supervision system to monitor the progress of the field operation is important to allow for the correction of any errors and omissions and to make any necessary adjustments in the course of the fieldwork. In countries where the Internet or handheld devices are used in data collection, a computerized online system can be developed, and some automated procedures introduced for the supervision.

3.209. The key to rapid quality control of the enumeration is the fast flow of information from the supervisors in the field to the managers in the local or regional statistical committees and/or to the central census office. The most efficient way of exchanging this information is via the Internet. If field supervisors have Internet access, information can even be submitted through a password-protected database interface (a web-based application).

3.210. Close monitoring during the enumeration phase is essential to ensure coverage, quality and compliance with deadlines. It must be ensured that all staff involved in the data collection have access to up-to-date reports with relevant information. These reports should be made available regularly in

printed or digital form. As noted at paragraph 3.212 below, data from previous censuses or other sources can be utilized to improve monitoring and form a database for management indicators.

3.211. As the enumeration is one of the core and critical census processes, each task performed during the enumeration stage must be carefully planned, executed and supervised to achieve qualitative and quantitative targets. For successful monitoring of the field enumeration, actual performances should be evaluated against the set targets. The following are recommended to achieve the desired outcomes.

i. Using historical data

3.212. Data from previous censuses and other relevant data sources such as household surveys and administrative registers can be used as a benchmark to determine the data needed for monitoring the performance of enumerators against the set goals and targets. Examples of historical information include total population of previous censuses; population growth rate; sex ratio; urbanization rate; proportion of vacant dwellings; and occasional use dwellings in relation to those occupied.

ii. Setting goals and targets

3.213. As a population census is a time-bound project, extension can be considered as failure. Setting goals and targets will be very important to measure whether the series of activities is under control or not. Goals and targets for measuring the quality of enumeration and for systematic monitoring of the enumeration can be set based on the experience of previous censuses and other relevant data collection processes. The following indicators can be used for monitoring the enumeration:

(a) the proportion of occupied and vacant dwellings;
(b) the average number of residents per dwelling;
(c)' response and refusal rates;
(d) population size; and
(e) population growth rate.

Significant deviation between the target values and the enumerated values may indicate a problem in the collection process. Estimation of the number of housing units and population size – if available – based on census maps and the listing of living quarters and households can also be used as information for monitoring the enumeration.

iii. Preparing policies and procedures

3.214. Policies and the procedures to be used as the baseline for monitoring during enumeration should be defined at an early stage of planning the census, with endorsement from the highest levels of decision-making, for proper management support. Therefore, it is important that there is a stage of evaluation of the previous census field operation in order to identify gaps and improve control procedures and execution of work.

F. Security during data collection

3.215. The census is a monumental undertaking which collects sensitive personal information from the population. To maintain public trust, it is important to ensure the data gathered in a census is not only accurate but also secure and protected from unauthorized access and misuse. Census data should be safeguarded throughout the entire lifecycle of the census.

3.216. In the *pre-enumeration phase*, plans need to be drawn up for assuring the security and confidentiality of the information to be collected. These safeguards should aim to protect the privacy of individuals and prevent unauthorized access or disclosure of their personal information. Information security measures include:

(a) robust data encryption methods to secure data both during storage and transmission;
(b) strict access controls and authentication protocols to limit access only to authorized personnel;
(c) regular security audits and assessments to identify vulnerabilities; and,
(d) comprehensive training programs to educate census staff on the importance of upholding data privacy and security practices.

In parallel, it is essential to establish strong physical security measures, such as limiting access to census facilities and using secure data storage systems~~, are put in place~~ to safeguard the physical infrastructure and prevent unauthorized tampering or theft of data.

3.217. During the *enumeration phase*, various information security safeguards should be implemented to ensure the protection and confidentiality of the collected information. First and foremost, robust encryption protocols should be used to secure data transmission between the data collection devices and the central database. Secondly, data accessibility should be limited to census officials, especially in the case where questionnaires are pre-filled with administrative data. Strict confidentiality agreements should be upheld by all personnel involved in data collection, with severe penalties for any unauthorized disclosure.

3.218. Census data collection with electronic questionnaires (whether with handheld/mobile devices or via the Internet) should be subjected to a thorough security review to confirm that the overall design of the system is robust to shocks (connectivity disruptions, power failures, etc) and security vulnerabilities (unauthorized access to data). The review process must concentrate on as many risks as it is possible to formulate in advance and include the common risk points: data at rest on a mobile device; data in transit to the intermediate or central location (via networks or via physical media, or both depending on the system); data at store in the central location; and data held in processing/ collaborating systems.

3.219. In the *post-enumeration phase*, secure physical and/or virtual environments should be used for census data processing, analysis and storage. In this environment strict measures such as encryption, access controls, and firewalls should be put into place to prevent unauthorized access to the data. Additionally, regular security audits and assessments should be conducted to identify and address any potential vulnerabilities. Strict data handling and confidentiality protocols should be followed by trained and authorized personnel, who are required to adhere to ethical guidelines and sign confidentiality agreements. Data anonymization/pseudonymization techniques may also be utilized to further protect the confidentiality of census data.

3.220. The variables for data analysis purposes should be provided on the need-to-know basis, for example, if a report involves construction of the population pyramid, then only the age and sex of enumerated persons should be supplied for such analysis, and variables like location data or contact information should not. Data requirements for each report, table, chart, or analytical product may and should be documented well in advance of the census and accompany the tabulation plan.

G. Use of technology

3.221. Technology, namely the use of computer-assisted methods for census planning, organization, data collection and data processing, has become indispensable for improving the cost-benefit, quality

and efficiency of the population and housing census. Rapidly expanding worldwide mobile and Internet connectivity coupled with advancements in areas such as cloud computing, smart mobile devices, GPS, GIS, natural language processing and AI provide new opportunities for further increasing the quality and speed of census data collection. Furthermore, the disruptive effects of the COVID-19 pandemic precipitated more widespread adoption of online tools and methods, as social distancing measures necessitated a shift towards remote solutions for data collection. The technological tools and instruments described below are well documented in national practices in conducting the population and housing censuses in the most recent census rounds.

a) Electronic questionnaire

3.222. Combined with, or completely replacing, the paper questionnaire, an electronic questionnaire can be used in either the face-to-face or self-enumeration method. Electronic forms can provide improved data quality and operational efficiencies by implementing validation rules on individual questions, cross-validation between questions or with other records, automatic sequencing (leading the operator to the next appropriate question), more options in pull-down lists, capturing more detailed data, providing computer-assisted coding and the ability to ask tailored supplementary questions. Electronic questionnaires can give access to guidelines, illustrations, help material and even videos to provide instruction to the interviewer or household. Electronic questionnaires can also be prefilled with geocodes, which enumerators can validate against GPS information collected from the field and update, if necessary. Pre-filled administrative data could be validated by respondents.

3.223. The use of electronic questionnaires provides a great potential towards audit and quality control: in particular, a confirmation of the visit and contact with the household. This can be done with a combination of capturing an image, recording the interview conversation, capturing the geo-location during the interviewing process, and a range of operational information that can be used to monitor operations and analyse responses (often referred to as paradata – see paragraph 3.207), including the time taken to complete the form, the date and time the form was completed, and the device used to complete the form.

3.224. Electronic questionnaires can also provide census results more quickly by transferring data to a central database immediately or soon after the enumeration. This can be achieved through various methods: real-time connectivity whereby data is transmitted electronically as soon as it's collected; periodic/on-demand synchronization when a connection becomes available; and offline data collection with later uploads to a local centre via cable or Bluetooth connectivity. Electronic questionnaires reduce the amount of material (such as paper questionnaires) to be printed, distributed and returned, and reduce data scanning and capture costs and errors. Yet the use of electronic questionnaires may introduce new expenses such as cost of the electronic devices used for data entry in the field, as well as data transmission costs.

i. Electronic questionnaire: interviewer-administered

3.225. Census interviews can be undertaken using an electronic questionnaire on smart phones, tablets, laptops or other portable devices using a computer-assisted personal interviewing (CAPI) methodology. Electronic questionnaires may also be employed in computer-assisted telephone interviewing (CATI) in which the interviewer records the respondent's answers to questions asked over the telephone. Each enumerator can be assigned one or several enumeration areas so that the records are tagged with the respective enumeration area and enumerator's/supervisor's name. The device may

also be able to capture information on the location of the interview, time of day and other metrics that may be useful for quality control. A CAPI system integrated with GIS and operational management applications which captures geo-coded data together with operational information (such as interview date and interview start and end times) can improve the monitoring of census data collection operations, and the coordination of field operations, logistics, and communications.

3.226. For areas with lack of Internet connectivity, systems with local direct data exchange between the interviewers and supervisors should be considered to facilitate a review of the completed interviews by the teams' supervisors even before such interviews are sent to the central database.

ii. Electronic questionnaire: self-enumeration method

3.227. Computer-assisted self-interviewing (CASI), whereby respondents access and self-administer the electronic questionnaire by means of an Internet browser application, can substantially reduce operating costs in areas where a high percentage of enumeration is achieved using this method. Achieving a good percentage of enumeration using this method can reduce the operation costs substantially. Electronic questionnaires for households should be implemented via a publicly-accessible Internet portal. Households are usually provided with a unique identifier (often delivered by mail, email, by enumerator or by self-registration system) that is used to initiate their questionnaire or resume a partially complete questionnaire via the Internet. Householders may prefer to respond using an electronic rather than a paper questionnaire for its convenience. If the census is collected in a multimodal approach, for example by offering both electronic and paper questionnaire options, a system will be required to track the status of each dwelling (questionnaire) throughout data collection to assess the completeness of coverage and to ensure non-response follow-up is not attempted with households that have already responded.

3.228. The design of a system implementing self-enumeration should demonstrate effective solutions for dealing with the following common problems: (a) duplication of submission of information (for example by different members of the same household); (b) fake or bogus submissions/digital vandalism (for example, the submission of data for a non-existing address or in the name of another household by hackers); and (c) attempts to obtain confidential data by gaining access to the unique identifier used to access the form.

3.229. However, total self-enumeration remains a challenge: access to the Internet is not universal, and there are groups of persons in the population who are not frequent users (older persons in particular) and may have technological barriers; recent immigrants may have language barriers; and persons with disabilities may experience difficulties reading the questions or supplying the answers to them. Thus, this method is often implemented as complementary to, but not entirely replacing, the interview-administered enumeration.

b) Handheld or mobile devices

3.230. Mobile devices such as smartphones or tablets can greatly improve the efficiency of census field operations by facilitating the interview, communication and supervision of fieldwork, data entry and validation, and data transmission. GPS-enabled handheld devices can assist enumerators with navigation by identifying buildings for each enumeration area using GPS coordinates, and capturing the interview location information.

3.231. When deciding to use mobile devices in the census process, it needs to be determined whether the census agency will purchase and provide the device, or whether the field officers will use their own devices. As the availability and proliferation of devices increases there can be financial benefits, as well as reduced training needs, if field officers can utilize their current device rather than be provisioned with a new device. This, however, presents a range of technical, support, security and legal considerations, and for this reason this deployment option should require detailed specifications on the compatibility/ platform of the device, minimal requirements (in terms of the storage, memory, performance, communication capabilities, battery capacity and life as well as screen size and resolution). Another significant consideration is the fact that mobile devices operate on different platforms and developing applications for different platforms increases the costs and efforts, unless such cross-platform portability is already provided by the electronic system chosen for the data collection and census operations management.

c) Geographic information system (GIS) and geospatial information

3.232. Geographic information systems (GIS) may be used for several major purposes during a census:

i. to place the enumerator in the context of the environment, by showing his location, surrounding dwellings (as points or where available as footprint polygons), and their status, whether already enumerated, in progress, or to be enumerated;
ii. routing capabilities for enumerators to navigate to the household (for the purpose of making return visits for example) and/or for supervisors (for audit visits);
iii. to provide controls for the supervisory staff, especially with respect to the coverage control of the enumerated areas;
iv. to produce various dashboards and GIS-based analysis of the progress of the fieldwork for the census managers.

3.233. Such uses rely on the location-enabled mobile devices (GPS or analogous positioning system) and availability of digital maps (sourced from the Internet in the areas with connectivity or stored on the device in areas where the connectivity is absent or sporadic).

3.234. The choice of technology for production of EA-level maps – which is dependent on the choice of data capture software - can greatly affect the time it takes to produce maps. These must be taken into account during the planning stage in order to establish an accurate timeline for the census operation. An extensive elaboration on the use of both GIS and GPS is presented in Chapter IV above.

d) Dashboards

3.235. Dashboards present aggregate information about the progression of the census field operation overall in a concise, visually compelling and efficient manner. The key indicators to be tracked with the dashboards are to be chosen and defined at the planning stage. A dashboard is not a substitute for detailed reports, but is a supplementary tool, which provides supervisors and managers with dynamic, possibly real-time, updated information.

3.236. A layered approach to the construction of the dashboard presents a possibility to visualize different indicators to various members of the management team. While the managers may be more interested in the overall progress of the census data collection, the supervisors in the field may be getting the indicators stemming from the data collected by their specific team.

3.237. Access to the dashboard information should be well-protected and limited to only those team members that need such information for their operational duties.

e) ***Contact centre***

3.238. The contact centre, also known as a call centre, is a crucial component of the field operation. It can be utilized at every stage of the census – but particularly to support fieldwork and ensure that the enumeration is conducted efficiently and accurately. A contact centre can employ various communication channels, including phone calls, chats, and web-based FAQs, to provide assistance to respondents. Interactive voice response (IVR) technology can be used to direct calls to specific agents based on the options selected by the caller. It can also be used to provide common answers to frequently asked questions, thereby reducing the workload of field and management staff. Website call-back and chat features can be implemented to assist respondents while they fill out the online e-questionnaire in the portal. Contact centres can therefore offer support to both the public and census personnel, ensuring that everyone involved in the census process has the resources they need to complete their tasks effectively.

f) Situation room

3.239. A census situation room serves as a centralized operational hub for coordinating and managing the multifaceted operations of a census field operation. Its primary purpose is to ensure the smooth execution of enumeration activities, swiftly addressing challenges and optimizing resources in real-time. In leveraging geospatial technology, the situation room can: (a) provide real-time visualizations of field operations across regions thereby enabling decision-makers to monitor the progress of data collection; and (b) cross-check enumerator locations with predefined areas to predict potential bottlenecks and allocate resources as and where needed and ensure enumerator safety. In the event of unforeseen incidents or challenges, such as natural disasters, civil unrest, or other disruptions, geospatial data can help the situation room quickly assess the impact on census operations and devise strategies to mitigate the challenges.

g) Short messaging service (SMS)

3.240. Short messaging service (SMS) offers a versatile tool for the census project. It can be used throughout various stages to: i) share information among field personnel and respondents (e.g., passwords, guidelines); ii) send alerts, marketing messages, and reminders; and iii) integrate with the central database for critical event notifications (e.g., low coverage, violations). Beyond SMS, other communication media like USSD and media platforms can further facilitate communication. These platforms can be used for sharing information, receiving feedback, and supporting both respondents and field staff.

VIII. Data processing

3.241. No matter how thorough and accurate the census enumeration is, the usefulness, quality and timeliness of the census tabulations will suffer unless the collected data are properly processed. An important element of a successful processing operation is the close and continuing collaboration, at all levels, between the data-processing staff, and the subject matter specialists and general statistical staff.

At a minimum, the subject matter and general statistical staff will need to become familiar with, and take a continuing interest in, the processing plans and operations, while the processing staff will need to become, similarly familiar with, and involved in, the core elements of the census.

3.242. Plans for data processing should be formulated as an integral part of the overall plan of the census, and those responsible for such processing of the census should be involved from the inception of the planning process. Data processing will be required in connection with the results of census tests, the compilation of preliminary results, preparation of the output tabulations, evaluation of census results, analysis of census data, arrangements for storage in and retrieval from a database, identification and correction of errors. In addition, data-processing technologies are playing an increasing role in the planning and control of field operations and other aspects of census administration. Data processing has an impact on almost all aspects of the census operation ranging from the selection of topics and the design of the questionnaire to the analysis of the final results. Therefore, data-processing requirements in terms of personnel skills and knowledge, space, equipment and software (computer programs) need to be considered from the point of view of the census as a whole and at an early stage in the planning.

3.243. When processing census data collected through various methods and modern practices, several crucial considerations come into play. Firstly, integrating data from diverse collection modes into a cohesive dataset is essential, but it can be challenging due to potential differences in formats and structures. Therefore, establishing a robust integration process is vital for accuracy and completeness. Additionally, assessing the quality of data collected through different modes is imperative. For instance, the responses given in self-enumeration questionnaires may be less reliable, necessitating thorough error identification and correction. Furthermore, ensuring the security and confidentiality of the data, especially when collected electronically, is paramount. Implementing encryption and other security measures safeguards against unauthorized access, while anonymization techniques protect individual identities.

3.244. To address these considerations, employing a versatile data management system is advised - one capable of seamlessly integrating and handling data from various modes. Establishing specific data quality checks and validation rules aid in error identification and rectification. Encryption and other security measures are crucial for safeguarding data integrity, and anonymization techniques help preserve confidentiality.

3.245. To successfully integrate new computer hardware or software into the census, it is crucial to expand and enhance the skills of the existing data-processing staff. This is because unfamiliar technology often demands new proficiencies for operation and maintenance. Given the intricate nature of the census process, it is crucial for the data-processing team to be well-versed in the specific requirements of the census for efficient data processing. Early training is essential to integrate the staff into census planning and operations promptly, enabling them to acquaint themselves with the new technology and contribute valuable insights to the process. This early preparation ensures the team is poised to process the collected data promptly. The training program should encompass a comprehensive range of skills necessary for the staff to execute their roles effectively, including proficiency in operating the new hardware and software, processing census data, conducting quality checks, generating census reports, troubleshooting hardware and software issues, and maintaining the technology infrastructure. To cater to diverse learning styles, the training should be delivered through various formats, such as classroom instruction, hands-on practice, and online learning modules. Additionally, it is crucial to evaluate the effectiveness of the training program through methods like surveys, interviews, and observations, using the results for refining future training endeavours.

3.246. There's a potential downside to training census staff on the software too early. Staff might leave before the census even starts due to various reasons like finding other job opportunities, retiring, relocating, or facing personal commitments. This exodus could adversely affect census quality and efficiency as remaining staff may lack adequate training and require additional resources. To mitigate this risk, it is recommended that statistical offices train only committed staff, offer attractive incentives, provide avenues for professional growth, foster a positive work environment, and have a contingency plan for replacement training. Factors to consider include the cost of training, availability of qualified trainers, and the importance of maintaining consistency in the census process. Ultimately, the decision to conduct early training should be tailored to the specific circumstances of each case.

3.247. Decisions will need to be made concerning the location of the various data-processing activities within the country, including the extent to which the processing work is to be decentralized. This decision should be partly based on the ability to recruit the required personnel for the processing operations. Acquisition of both equipment and supplies can require long lead times; estimates of both data capture and computer processing workloads must be made early to enable timely procurement. Closely related to the question of equipment is that of the provision of adequate space. Although the maintenance of most personal computer equipment no longer requires adherence to rigid standards in terms of temperature, humidity, dust and so on, attention to issues related to power supplies is still important. Inevitably, more important is the attention to be devoted to the maintenance of servers (especially heavy-duty servers), where most of the information processing is likely to take place and saved, as well as the data transmission infrastructure.

3.248. Ensuring smooth and uninterrupted Internet and web communications is vital for the success of census operations, as timely and accurate data transmission between different units is crucial. To achieve this, it is advised that census agencies employ diverse Internet connections from different providers, implement load balancing for improved performance, utilize content delivery networks to ease server load, and maintain vigilant network monitoring to promptly address any issues. In addition, for traditional archiving, creating a secure and controlled space to store completed census forms is imperative due to the sensitive personal data they contain. This involves employing physical security measures, logical security measures such as encryption and access controls, environmental controls to protect against damage, and implementing audit trails for accountability.

3.249. When preparing for census data editing and tabulation, decisions regarding the software to be employed are pivotal. This choice hinges on several key factors, encompassing the census's scale and complexity, the proficiency of the data-processing team, budgetary constraints, and the availability of technical support. Various software options are available, falling into three primary categories:

(a) census-specific software tailored explicitly for census processing, such as CSPro;
(b) general-purpose software applicable for various tasks including census processing, such as spreadsheets, databases, and statistical analysis software; and
(c) custom software developed specifically to align with the unique needs of a particular census organization.

3.250. When opting for editing and tabulation software, it is imperative to weigh factors such as:

- user-friendliness for the data-processing team,
- requisite functionality encompassing data entry, editing, imputation, tabulation, and reporting,
- scalability to manage extensive datasets,

- robust security features to safeguard data confidentiality, affordability, and
- the availability of reliable technical support.

Subsequently, the selected software will necessitate customization to align with the precise requirements of the census organization, potentially involving code modification, report creation, or procedure development. Allocating ample time for customization is crucial. Additionally, the data-processing team should undergo comprehensive training on the software, covering all essential features. This training should undergo evaluation to ensure its effectiveness.

3.251. Outsourcing some of the predominantly IT-related operations may be considered. Outsourcing should be implemented in such a way as to bring immediate economic and quality advantages to census operations. However, it is crucial to approach outsourcing in a manner that upholds data confidentiality. To safeguard this, national statistical offices should rigorously select reputable contractors known for their commitment to data confidentiality, and contracts should encompass provisions mandating confidentiality protection. Continuous monitoring of contractors' adherence to these provisions is imperative. Employing encryption and other security measures during data transmission, and limiting access to the data, both internally and externally, are essential steps in preserving data confidentiality. Lastly, it is paramount that the national statistical offices retain control over the data, ensuring that outsourcing operations do not entail any relinquishment of data control.

A. Method of processing

3.252. The choice of an appropriate method of processing is determined by the circumstances of each country. Rapid advances in data-processing technology have greatly increased the speed and reliability of producing detailed tabulation, thereby making computer processing the standard method of processing around the world. Furthermore, an alternative to mainframes, whose computational power was necessary before the advent of lighter and more scalable IT hardware solutions and more recently cloud-based computing, is the use of a client-server environment.

3.253. Cloud-based computing, often referred to as cloud computing, is a technology model that allows individuals and organizations to access and use computing resources over the internet, as opposed to owning and managing physical hardware and software on their local computers or data centres. In cloud computing, these computing resources are hosted and managed by third-party service providers in data centres located around the world. Cloud computing offers several benefits for conducting a census, making the process more efficient, cost-effective, and scalable. Cloud computing offers census agencies the ability to store, process, and analyse data while providing flexibility, scalability, and cost savings compared to traditional on-premises infrastructure efficiently and securely. These advantages can lead to a more accurate and timelier census, which is crucial for policy-making and resource allocation.

3.254. Several lighter tasks, including the editing and tabulation of data files, can very well be done on small-sized desktop systems that can be placed in substantive departments and in field offices. On the server side, most of the heavier computing operations, such as scanning, aggregation and analysis of large sets of microdata, coordination of data transmission, Intranet web hosting and so forth, can be executed more reliably than on microcomputers. However, a client-server environment to handle census data must operate over a robust and secure local area network (LAN) or wide area network (WAN) into the cloud environment. Therefore, computer work is not necessarily dependent on a centralized data-processing facility, provided that a robust LAN or WAN interconnects workstations dispersed over various offices, buildings and different parts of the country.

3.255. In a census office that utilizes a networked computer environment, the central file or database servers allow both data and programme files to be stored in a central location. This system economizes specifications of client computers and removes the need for much physical movement of programs and data on other computer media. Data storage requires frequent backups of the system information to avoid major data loss due to hardware or software faults. Thus, servers have a strategic importance, and their location and administration must be well defined and secure enough to ensure data protection. Also, it is recommended that proper business continuity and security policies duly certified by the competent authorities should be in place.

3.256. In determining the type of equipment to be employed and the advisability of a new machine installation (either complete or partial), or of additions or upgrades to existing equipment, consideration should be given to all the processing requirements of the data collection programme for which the population and housing census is but one part. Only on this basis can a reasonable decision be made. Decisions on the type of data-recording equipment and computer equipment should be made at least one year in advance of the scheduled date of enumeration in order to allow for:

(a) appropriate questionnaire design and proper preparation of instructions to field staff;
(b) the development of coding schemes, specification of data-handling controls and procedures; and
(c) the recruitment and training of data processing personnel.

Rapid processing of a pretest or pilot census that covers end-to-end census operation, including enumeration, initial census result, output dissemination, and handing over and closure procedures, is particularly important for identifying any improvements needed in the census questionnaire, instructions to enumerators, computer systems or whatever other preparations may be needed. It is recommended, therefore, that arrangements for using appropriate equipment and software be made well in advance of such tests. It is also recommended that all systems used to support census operation be thoroughly tested in advance of operations to ensure that they function as intended and that they are secure (that is, they will not lead to loss or unauthorised disclosure of data).

B. Preparation for data capture

3.257. In the case of paper questionnaires, the most common procedure is to have the census documents arrive in the processing centre in batches by enumeration area. Maintenance of these batches throughout the data processing is recommended, since documents for a given enumeration area reflect the work of one enumerator and may contain a series of errors systematic of that person. To ensure the integrity of the batches, the census documents should be stored in a specially designed storage facility. The batch for each enumeration area should first be checked for completeness, geographic identification codes and other characteristics of acceptability before being sent to the next stage of data processing. Transcribing all coded data onto another sheet (for example, the coding form) should be avoided since it may add transcription errors. The same considerations apply to the case of electronic transmission of questionnaires or when the first phase of data processing consists of the scanning, and text or image recognition, of census questionnaires. In the case of questionnaires transmitted electronically (self-enumeration on the Internet or via mobile devices), it is appropriate to set up a metadata model where the enumeration area can be recorded. As far as storage is concerned, if paper questionnaires are scanned, secure media storage for their backup copies as well as the originals should be planned for.

3.258. If the census is conducted in a multimodal approach, for example using self-response by Internet and field follow-up of non-respondents, it will not be possible to batch questionnaires by enumeration

area for processing. A master control system will be required to track the status of each dwelling (questionnaire) throughout collection and processing operations and ensure completeness of coverage. This is closely aligned to any master control system that is being utilised to manage dashboards and monitor the progress of the data collection.

C. Data capture

3.259. Data capture involves converting census data, collected through various methods including paper questionnaires, electronic data collection devices, or online self-enumeration, into a computer-readable format. This process can occur both during and after the enumeration period.

(vii) Data capture during enumeration

3.260. Electronic data collection (including collection by Internet or using handheld or laptop computers) has been widely used by countries. Electronic data collection means the integration of interviewing and the data entry process including data capture, coding and consistency checks. Electronic data collection with handheld devices, Internet or telephone allows the capture of information with relevant codes (there might be some exceptional variables that may require more detailed coding centrally (such as occupation and industry). It also allows the identification of potential errors during the interview with pre-programmed consistency checks. Because the consistency checks are performed in real-time, errors and inconsistencies can be resolved, and corrective action can be taken by the respondent or the enumerator. However, introducing editing rules into the data collection application has to be carefully considered so as (a) not to affect its performance significantly in the field, (b) not to introduce a bias in the responses, and (c) not to affect the quality of the interview for questions that may not be answered properly, especially in case of information provided from proxy respondents.

(viii) Data capture after enumeration

3.261. The traditional methods of processing census data occur after the enumeration phase. They include keyboard data entry, optical mark reading, optical character reading or image-processing techniques, such as intelligent character recognition. Computer-assisted keyboard data entry is usually carried out using personal computer data entry programs with built-in logic controls. Some of the tasks accomplished by the programs are:
(a) verifying that enumeration area codes are valid, and copying them automatically from one record to the next;
(b) assigning a number to each person in a household automatically (and perhaps to each household within an enumeration area);
(c) switching record types automatically if the program's logic requires it;
(d) checking that variable values are always within predetermined ranges;
(e) skipping fields if the logic indicates doing so;
(f) supporting keyboard verification of the information entered earlier; and
(g) generating summary statistics for the operator and the batch.
In order not to delay the data capture task, data entry applications should limit checking to problems that are either very serious (for example, wrong enumeration area code), or likely to be caused by a simple misread or key entry mistake. More sophisticated checking is deferred until the editing stage.

3.262. Optical mark reading (often called optical mark recognition) equipment has been available for many years and has nowadays reached good levels of reliability. Optical mark reading is the simplest of the commonly available form data capture technologies. Owing to relatively stringent requirements for the successful data capture of the paper, countries with very dusty environment or humid climates and poor transport infrastructures are discouraged from using optical mark reading as it demands the necessity to heed special questionnaire design restrictions, to consider the quality of the paper, and to adhere to precise specifications regarding the printing and cutting of the sheets. In some developing countries, this may mean that local production of the questionnaires will be problematic. The need to reserve a relatively large space for marking areas and to adhere to other limitations imposed by optical mark reading equipment sometimes make it difficult to design the best questionnaire from the point of view of the enumeration process.

3.263. Optical mark reading questionnaires can be marked by the respondent or by the enumerator. Marking by respondents is attractive from a cost perspective, but it depends on the presence of a cooperative spirit and relatively universal literacy. A practical problem is that most optical mark reading devices put restrictions on the writing instrument and the colours that can be used in the marking. Assuming that the rules are followed, the rejection rate for marked forms is often low, especially if the forms have been inspected visually before being fed into the readers. Converting a manually completed census questionnaire to an optical mark reading format after it has been received in the census office is inefficient and becomes a source of errors and should therefore be avoided.

3.264. Optical character reading (also called optical character recognition) and intelligent character recognition consist of the use of special equipment to read characters at specific locations in the questionnaire. These two methods involve very similar technological approaches. Specialized sources tend to identify optical character reading with the capability of recognizing printed characters only, whereas intelligent character recognition would extend this capability to handwritten text. There is no agreed definition of intelligent character recognition. In the context of censuses, therefore, this would require that handwritten text in the filled-in questions be as standard as possible so as to enable efficient recognition. In general, recognition of numerals is more efficient in an uncontrolled environment, that is to say, where the machine has not been adapted to the writing style of a particular person. Optical character reading and intelligent character recognition technology has matured considerable with sophisticated recognition algorithms and the use of neural networks for self-learning.

3.265. Imaging techniques and scanner devices, together with optical character reading and intelligent character recognition software, have been used by several countries for data capture. Experience shows that significantly low error rates are achieved at an optimum cost using these techniques. The efficiency is greater in the case of numerical and alphanumerical characters written by trained enumerators. However, alphanumerical characters are prone to higher error rates. Extensive testing must be conducted well in advance to determine the best type of equipment and paper. The use of imaging techniques is also dependent on the availability of local maintenance and support capabilities. Whatever methods of coding and data capture are chosen, it is essential that they be carefully tested before final adoption. Recognition engines can be customized to recognize various sets of characters and scripts, but unless good experience is available at the census office, careful planning and preliminary work are needed in conjunction with the optical character reading or intelligent character recognition system providers. A combination of intelligent character recognition (for numerical characters) with computer-assisted coding (for alpha characters) is also an effective method used by some countries.

3.266. In addition to the benefits of the scanning technology for capturing the information, an important by-product of scanning census questionnaires is that this allows for the possibility of digitally filing and naming the scanned questionnaires. This increases the efficiency of storage and retrieval of the questionnaires for future use, particularly during subsequent data-editing operations.

3.267. The quantity and type of data entry equipment required will depend on the method of data capture selected, the time available, the size of the country, the degree of decentralization of the data capture operations, and a number of other factors, such as the use of digital enumeration approaches. For keyboard data entry, the average input rates usually vary between 5,000 and 10,000 keystrokes per hour. Among the factors that affect operator speed are: (a) the supporting software and program with easily navigable screens, spell checker on the description fields if any, keyboard shortcuts throughout the program, less utilization of the computer mouse, and so forth; (b) the complexity of the operators' tasks; (c) the ergonomic characteristics, reliability and speed of the equipment; (d) the question whether work is always available; (e) the training and aptitude of the recruited staff; and (f) the motivation of the workers.

3.268. Several options are available to help ensure that data entry operations are completed in a timely manner. They include (a) procuring more equipment; (b) increasing the number of working hours by working double or even triple shifts and during weekends; and (c) applying independent verification to varying extents. In the case of keyboard data entry, with the increasing safeguard of data quality by data entry programs, complete verification has become less necessary. Full independent verification may be applied only in the initial stage of data entry and may be reduced when each worker has achieved an acceptable level of quality. After that, a sample verification plan can be applied. Operators may be assigned to sample verification depending on their observed error rate. The work of reliable operators may be verified only for a small sample of the enumeration areas, while more extensive verification is continued for the more error-prone operators.

D. Coding

3.269. Whenever possible, pre-coded responses should be used in census questionnaires with numerical or alphanumerical codes. Since computer editing and tabulation of textual material are not practical, verbal responses will have to be replaced by a code. The incorporation of artificial intelligence (AI) and machine learning (ML) technologies would be ideal. These advanced technologies offer the potential to automate critical tasks, including the classification and coding of open-ended responses. Such integration stands to enhance the overall efficiency and accuracy of the data collection process. While the use of pre-coded responses in census questionnaires is encouraged, there will inevitably be instances where respondents provide written responses necessitating coding. In such cases, dedicated computer programs or skilled coders can be employed to translate these written responses into codes. Although automatic coding is the preferred approach, certain situations may call for manual intervention. While coding directly during interviews can be advantageous, it is often logistically challenging as enumerators may not always possess the requisite training or resources for accurate coding. Therefore, in most cases, responses will be gathered in text format and subsequently coded by proficient coding experts. Given the scale of coding operations in a census, optimizing automatic coding processes is paramount to minimize the need for human intervention (see paragraph 3.186). Techniques like natural language processing (NLP) and deep learning can be employed to facilitate this optimization. Additionally, AI and ML technologies can be harnessed to effectively classify occupations and industries based on text descriptions, a valuable tool for coding open-ended responses pertaining to occupation and industry.

3.270. Automatic or computer-assisted coding will efficiently support the coding activity, reducing coding errors and speeding up the coding process. When required, a coder normally works with one or several codebooks for various items in the questionnaires. This technology expedites coding tasks, liberating coders to concentrate on more intricate responsibilities, such as coding open-ended responses (to the occupation question for example) or addressing coding discrepancies. Additionally, automatic or computer-assisted coding provides invaluable feedback to coders, enhancing the overall accuracy of the coding process by aiding in error identification and correction. By automating laborious coding tasks, this technology alleviates coder fatigue, thereby elevating the overall quality of coding work. Lastly, it ensures the consistency of the coding process by enforcing uniform adherence to coding guidelines among all coders, leading to enhanced data quality.

3.271. Computer-assisted coding uses personal computers to assist the coders. The process requires that all the codes be stored in a database file and be accessed by coders during the coding operation. Computer-assisted coding is based on at least two general approaches. In the first one, coded answers are matched to a set of keywords. Textual information from the census questionnaire is parsed and compared to an indexed list of keywords, and then the likelihood of matching between found keywords and coded answers is measured and scored. If the score results are over a certain (high) threshold and there is no ambiguity, a sorted list of coded answers is presented to the coder, who retains the ultimate decision of accepting or refusing the system's proposed answers. In using this method, it may be advantageous to change the order of activities so that the capture of pre-coded information in the questionnaire occurs first, followed by the capture and computer-assisted coding of the remaining information.

3.272. In the second approach, which is mainly used in image processing of data (intelligent character recognition method) for non-Latin languages or multilingual countries, owing to the difficulty and existing problems in character (alphanumerical string) recognition, the procedure is as follows. After the scanning and during the coding operation phase, the image of the text will be shown on the monitor, and at the same time, a pull-down menu from a coding database will present the coder with the ability to enter as few key entries as possible to get to the full textual and coding content of a specific case. When the coder selects a code, it will be allocated and saved in the database for that specific case. Although this approach is more time consuming and costly in comparison to the first approach, the quality of coding is much higher than in the traditional way of coding.

3.273. On the other hand, both techniques have several similar advantages: (a) capturing the pre-coded information at an early stage leads to some data files becoming rapidly available, which opens up the possibility of generating and releasing preliminary census results; (b) the computer-assisted coding process provides an opportunity for a computer system to alert the operator to problems with data supposedly already captured, for example, missing information for a fully pre-coded variable; (c) the coder works directly on the computer screen; and (d) information from other variables may be helpful in determining applicable codes for write-ins.

3.274. Automatic coding is a process in which the decision about the code to be assigned is delegated to a computer program. The main difference from computer-assisted coding consists in the automatic acceptance of the answer if its score is over a predetermined threshold and relatively higher than possible identified alternatives. Both computer-assisted and automatic coding systems may exploit self-learning capabilities of neural networks to fine-tune their capacity of detection. A human operator becomes involved only in those cases where the software cannot resolve the issue. Computer coding

may use, in addition to the written response for the item in question, other relevant information available in the record or the questionnaire. Therefore, automatic coding is more applicable in cases where the data-capturing process has already been completed, by Internet, handheld devices or other forms of electronic data collection, manually or by some form of automatic reading. Developing computer software for automatic coding is a complex task. Automatic coding methods need to be complemented by computer-assisted or conventional coding methods for unresolved responses.

E. Data editing[81]

3.275. Raw data files contain errors of many kinds, some generated by the respondents and others caused by enumerators who may have misunderstood the respondents' answers. Further mistakes are introduced in the data-processing operations and during coding and data entry, or in the course of the transcriptions that take place. From an operational point of view, such errors are of two types: (a) those that have the potential of blocking further processing (critical errors); and (b) those that introduce distortions into census results without interrupting the logical flow of subsequent processing operations (non-critical errors). Throughout the correction process, backup copies of the original data file should be consistently created. Numerous errors can manifest themselves, including misinterpretation by respondents, enumerator misunderstanding or data entry mistakes, typographical errors during data entry, software-related errors, and transcription errors. The timely identification and correction or errors is imperative to prevent inaccuracies that could profoundly impact policy-making, resource allocation, and crucial decision-making processes. Various error-checking and correction methods, such as range checks, consistency checks, logic checks, and manual review, can be employed to enhance data accuracy, with manual review offering the highest accuracy albeit with increased time investment.

3.276. Since for large censuses manual correction is rarely economically feasible, the conditions for such corrections are usually specified in specially designed computer programs for automatic error scrutiny and imputation based on other information for the person or household or for other persons or households. Whenever imputation is used, a flag should be set so that analysts are able to distinguish between reported information and that imputed by the editing system. In cases where insufficient information hinders error correction, methods such as hot deck imputation can be deployed. This technique uses information obtained from previously processed persons, families or households with similar characteristics as the "best fit" value in replacing missing values or values that have failed processing edits. However, this technique requires careful programming work, considering that the search for appropriate information in the census database would slow down computer program execution.

3.277. A prudent approach in data processing involves isolating out-of-range or blatantly inconsistent values prior to editing or classification, mitigating the risk of introducing statistical biases. But precautionary measures should also be defined and set for the fact that overambitious automatic editing programs may cause the so-called "corrected" data to be significantly flawed. In this respect, it would make sense to have an acceptable cut-off value for error rates at the enumeration area level. If a data scrutiny program finds that more than a certain percentage of the records in a particular batch have one or more serious problems, the whole batch should be rejected and subjected to human or fieldwork verification.

[81] For further details on census data editing, see *Handbook on population and housing census editing*, *Revision 1,* United Nations publication, Sales No. E.09.XVII.11, United Nations, 2010.

3.278. Developing editing and imputation rules should be entrusted to subject matter experts rather than computer programmers. This is because specialists possess a profound comprehension of the data and potential sources of errors, enabling them to identify likely error types and establish effective correction protocols. An error scrutiny and editing plan should be crafted early in the census, delineating specific rules and their implementation procedure. This plan should also incorporate a mechanism for periodic review and adjustment of the rules. These rules must be meticulously documented and shared with programming staff to ensure accurate and consistent implementation, utilizing clear and straightforward language for easy comprehension. While programmers bear the responsibility for preparing the coding rules, subject matter experts should verify the accuracy of the codes. Thorough testing by both subject matter and software testing experts is imperative before applying the editing programs to the census dataset, detecting and rectifying any program flaws under diverse conditions to ensure comprehensive functionality. Adhering to these guidelines yields numerous benefits, including heightened data accuracy, reduced error risks, enhanced transparency in the editing process, improved collaboration between subject matter experts and programmers, and elevated quality of census results. However, challenges persist, encompassing the time-consuming nature of rule development and program testing, and the potential complexities in coordinating the efforts of subject matter experts and programmers, along with the need to mediate disagreements between them.

F. Validation

3.279. The outcome of editing is a set of records that are internally consistent and in which person records relate logically to other person records within the same household. This process does not, however, provide the full range of assurance necessary to accept the data set as the best possible. A range of conditions could create errors that cause the data to be consistently wrong: for example, perhaps a condition in the editing suite itself is set incorrectly; proportions in an imputation program may be set wrongly; or enumerators may complete a collection control panel incorrectly. To identify such consistent errors, it is necessary to critically review some key aggregated tables to isolate outlying numbers and identify the cause of the unusual values. These key tables may be a subset of those intended for output or may be tables specifically designed for this purpose.

3.280. It is recommended that a bottom-up approach be used in this process. That is, the tables should first be examined for a selection of enumeration areas, then for the next level up and so on up to the first set of national tables. There are two reasons for this:

(a) The first enumeration area will complete the processing cycle well before any other geographic level. Thus, commencing at this level gives the earliest possible warning of any problem, enabling corrections to be made before a large amount of reprocessing is required.
(b) It is far simpler to examine a few hundred records within an enumeration area than to attempt to resolve the problem in the millions of records in a national file.

3.281. A crucial stage in the process is designing the analytical tables. One way of approaching this could be to identify a set of variables that are conceptually consistent with those in the previous census (or a major survey) or administrative records from various authorities in the country, such as expatriate visas issued, national ID programme or number of registered establishments. Thus, a set of benchmark values could be constructed before the census operation commences and compared with those from the current enumeration. The content of the benchmark set will depend upon the content of the enumeration and much of this must therefore be determined by each country. However, any census will include the variables age and sex so a comparison of the age pyramid and sex ratio for each 10-year age cohort would be basic elements of such analysis.

3.282. A second component of the analysis is the compilation of a set of data regarding expected changes since the benchmark survey. For example:

(a) It is possible that in the time since the previous collection, improvements in maternal health care programmes have led to an increased survival rate for women, in which case cohort survival ratios for females should be higher for younger women than older ones.

(b) If literacy is included in the analysis and government policy has been to strongly support increased school attendance, an increase in the proportion of literate people could be expected.

3.283. There will be a need for careful judgement when the analytical tables show a significant and unexpected difference from the benchmarks. While it may be found that the difference is due to a problem with the current collection, it could also be due to:

(a) a problem in the collection that has generated the benchmarks;

(b) a genuine and previously undetected social change that is being correctly revealed by the current census data.

3.284. In the latter two cases it would be wrong to make any change to the current data set. However, it is crucial that details of the investigation are made known to users (by preparing suitable metadata) so that they would be able to interpret and analyse the data correctly. If such analysis indicates that there is a problem with the census data, it will also be a matter of judgement on how to react to it. One solution might be to revise the input processing system to prevent the problem from being continued. After applying such changes, and to avoid introducing further problems, it is essential that the input processing system be fully tested and accepted. The second possibility is to decide as to whether to reprocess the records that have already been processed. This decision should be guided by the following:

(a) the significance of the error;

(b) the number of questionnaires that have already been processed;

(c) the time required for the reprocessing;

(d) the impact of such a decision on other subsequent phases of the census (such as tabulation and dissemination); and

(e) the cost implications.

3.285. As previously noted, the elements involved in data processing require joint and integrated work between analytical (demography, statistics, sociology, economics, geography, etc.) and technological (engineers, analysts, programmers) personnel. This is especially so for validation where thematic and comparability (with other sources) indicators are generated apart from tabulations to improve data coherence. Before and after comparations (original and validated data) should be included as part of the validation. In the same way, knowing the data dissemination plan helps organize and prioritize both editing and validation.

G. Processing control

3.286. Careful planning and control are required to ensure an uninterrupted flow of work through the various stages from receipt of the census questionnaires through preparation of the database and final tabulations. For the data from questionnaires completed through hand-held devices and transferred to a server, a clear data extraction strategy must be put in place with consideration to the scale of the census operation. The plan should provide for the computer edit to follow closely the coding, checking and

recording of the data so that errors can be detected while knowledge related to them is fresh and appropriate remedial actions may be taken.

3.287. Countries may wish to establish a computer-based processing management and control system to check individual forms or groups of forms for each enumeration area or for other processing units. Such a system should link the databases for enumeration areas and other geographic entities with the control information.

3.288. The system would check and manage progress from process to process so as to ensure the completeness of records at each stage of the processing operations. As specified earlier, project management software may support the formal description of different processes and provide an environment to control the execution of all operations connected to an individual phase or status of the census. This system should be fed into the overall quality assurance and improvement system, the management of which is elaborated in paragraphs 4.19-4.27. If a computer-based processing system is established, a close and real-time communication between the headquarters, local offices and field enumerators should also be established. This is beneficial for the control of field staff and the management of logistics of enumeration materials. Also, as each problem occurs in the field, a solution can be shared through the bulletin board in the system, which can greatly reduce non-sampling error.

3.289. Countries that intend to integrate census operation and census processing should test the whole cycle before conducting the actual census.

H. Master file

3.290. When data editing and validation are in progress, new files consisting of clean data records for each person are produced; these can be assembled to build a master file for later tabulations and census indicators (often called the microdata file). This master file, like the raw data files, can have a simple rectangular sequential format. There is usually no need for having the master file organized with a database structure with index files (but neither should it be discouraged). However, the master file should usually be maintained in geographic hierarchy, starting with the lowest geographic entity, sorted by housing unit, household or family. Another method commonly used to generate tabulations involving both the person and the family, household or housing unit, is to include in the household reference person's record (or head of household's record) selected characteristics of these latter units. Alternatively, a single hierarchical file can be created involving, for example, person, family and housing unit records. Whatever the chosen structure, the master file must allow for easy checks, controls and computations to be performed. The hierarchical model should be the best shape to generate this master file including all in one.

3.291. One of the most common and problematic errors in census files is that different enumeration areas carry, for one reason or another, the same identification codes, more common now with the mixed-mode collection method. Upon sorting the file, these enumeration areas may have been merged, generating households with abnormal characteristics such as two different reference members (or heads) of household, twice the usual number of members, two or more housing records, and so on. To avoid this problem, the enumeration area geocodes should be checked carefully prior to the editing phase. This is best done by keeping a check file of all expected code combinations and marking a code as "used" once an enumeration area using the code has been processed. A module of this functionality can be part of the editing programme. The check file will serve to flag impossible or double identification

codes, and towards the end will show which enumeration areas were expected but have not been processed.

3.292. Census master data files are usually very large and require powerful servers to process them or specialized software with those capabilities. Well-equipped desktop systems also have higher computational power and are equipped with much bigger and cheaper mass storage devices. Nonetheless, the hardware infrastructure available to many countries is older, thus two strategies are applied to reduce file size and to make data management simpler. The first involves working with the next lowest geographic entity as a basis, processing the data on this level and aggregating later to obtain national results. The second remedy is to apply on-the-fly compression and decompression to the storage medium. Census files can be compressed quite significantly to less than 20 per cent of their original size. Since tabulation programs access the data in sequential order, using the compressed data will result in a faster reading process. New compression and storage technologies can be incorporated as well as encryption methods.

I. Methods of tabulation

3.293. Preparing the tabulation plan is the substantive responsibility of the demographers and other subject matter specialists who have the necessary expertise in interpreting the census results. This will require consultation prior to the census with principal users of the census information (see paragraphs 2.157–2.162). The duties of the data-processing department should be limited to checking the logic of the various aggregations, designing the required programs and producing good quality results within the shortest possible time. It is possible that the need for initially unanticipated tables will become apparent, so the census organization should always be prepared to produce additional aggregations as required. This may involve newly defined classifications for certain variables, new types of cross-classifications, differently defined geographic subdivisions, and so on. If the master file is organized according to the principles of relational databases in a relational database management system, original and additional aggregations can be designed according to relatively easy structured query language statements. Alternatively, if the master file is a list of records with a rectangular structure, online analytical processing (OLAP) tools can be leveraged for creating multidimensional tabulations. It is imperative to ensure that all requisite information for tabulation production is readily available in the master file, as adding it later may prove to be prohibitively expensive. The tabulation plan should meticulously outline the tables to be produced, the variables encompassed in each table, and the cross-classifications to be applied, also specifying the level of detail to be provided in each table. It is essential for the census organization to review and endorse the tabulation plan prior to data processing, with provisions for revisions as necessary throughout the data processing phase. The tabulation plan stands as a pivotal component of the census process, guaranteeing that census data is presented in a format conducive to user needs, while also ensuring the accuracy and reliability of the data.

3.294. Specialized census tabulation software is crucial for census organizations due to its efficiency in processing large datasets, ensuring accuracy through error-checking features, and offering flexibility in generating customized tables and reports. Its user-friendly interface and low cost make it accessible and cost-effective. The software also includes various functionalities like data import/export, data cleaning, table generation, charting/graphing, and report generation, further enhancing its utility. Overall, specialized census tabulation software is an invaluable tool for census organizations worldwide, significantly improving data processing efficiency, accuracy, and cost-effectiveness.

3.295. Tabulation work can also be easily done by software belonging to either one of two other classes: statistical analysis and database software. However, these packages have not been designed with large-scale sequential or geographic processing in mind. They may require substantially more computer time than a specialized census tabulation system. In countries with a limited capacity of powerful computers, this can be an important consideration.

3.296. Other factors that should be taken into consideration when selecting software packages for tabulation work include:

(a) The availability of expertise in the census office. It makes no sense to switch to a software system that is only marginally better when this would require a major retraining effort.

(b) Opting for a software that the census office is already familiar with is often a more prudent choice; this encompasses factors such as the operating system and database software.

(c) Compatibility with existing resources at the census office. Incompatibility may lead to the additional expense of acquiring new hardware or software, escalating the overall implementation cost.

(d) Furthermore, the potential need for software customization to execute advanced functions like random perturbation for confidentiality preservation should be evaluated. Customization can be both resource-intensive and costly, underscoring the importance of assessing this requirement prior to finalizing a software choice. Alongside these considerations, factors such as cost-effectiveness, robust vendor support in terms of technical assistance, documentation, and training, scalability to handle large datasets, and robust security features should also be taken into account in the decision-making process. These factors collectively contribute to ensuring the effectiveness and suitability of the chosen software for census tabulation work.

J. Use of administrative data during data processing (for field-based censuses)

3.297. Administrative data may be leveraged during the data processing of a field-based census, mainly for a) editing of missing or implausible data or non-respondent households or b) validation of results. It may also inform the development of proper edit rules such as checks on the plausibility, or automatic coding systems for specific variables. Before using administrative data, two key conditions should be in place: (a) there must be a legal framework that allows such use and (b) there must be a proven method of linking data between different sources at the unit record level.

3.298. Census data often suffer from undercounting or missing values due to various reasons, such as non-response or errors in data collection. Administrative data can play a crucial role in imputing missing or incomplete census data. When the census form is incomplete, information from administrative sources, such as records of government benefits, healthcare records, or school enrolments, can be used to either replace - when microdata integration is possible - or to impute missing values.

3.299. In cases of non-response, administrative data sources, such as data on electricity consumption from electricity distributors, can be used to identify occupied dwellings. It may also aid in estimating the number of individuals residing in non-respondent dwellings. This will enhance the imputation procedures, resulting in improved census count.

3.300. When administrative data is of high quality, it can serve as an invaluable tool for validating the coverage and content of census data. By cross-referencing the census records with various administrative records such as tax filings, birth and death certificates, or immigration records,

discrepancies and inconsistencies can be identified. This validation process helps ensure that the census data are reliable, providing a more accurate reflection of the population. For example, administrative data can be used to verify that individuals listed in the census are indeed living at their reported address and that demographic information aligns with official records. However, as mentioned earlier, this is contingent upon the quality of administrative data. In cases where the quality of administrative data at the micro-level is not high yet satisfactory at the aggregates, it may be used in the validation of census outputs.

3.301. The use of administrative data in census, either to provide data on specific variables or sub-populations can significantly affect data processing. Among others, it will add a so-called "pre-processing step" that entails cleaning, standardizing and transforming administrative data to assure it suitability for census purposes. Data-processing staff will certainly need to develop skills to deal with challenges involved in the use of administrative data sources. A close and continuing collaboration, between the data-processing staff, the subject matter staff and relevant administrative agency at all levels is essential in mitigating and tackling the challenges.

K. Security during data processing

3.302. Security concerns during census data processing are paramount due to the sensitive and confidential nature of the information collected. Ensuring the confidentiality, integrity, and availability of census data is critical to maintaining public trust and complying with privacy regulations. Some of the key security concerns, including disclosure and confidentiality concerns, during census data processing are noted here:

a) *Data privacy and confidentiality*
   i. Data privacy laws: Census agencies must adhere to data privacy laws and regulations that govern the collection, storage, and processing of personal data. Violations can lead to legal consequences and damage public trust.
   ii. Anonymization: Techniques such as data anonymization and de-identification are crucial to protect individual privacy. Personally identifiable information should be removed or transformed to prevent data from being linked to identifiable individuals.

b) *Data encryption*
   i. Data in transit: Data should be encrypted while it is being transmitted over networks (e.g., from data collection points to data processing centres) to help protect against interception and eavesdropping.
   ii. Data at rest: Data at rest, whether stored in databases or backups, should also be encrypted to prevent unauthorized access in case of data breaches or physical theft.

c) *Access control*
   Role-based access should be implemented to restrict access to census data based on job roles and responsibilities. Only authorized personnel should have access to specific data sets.

d) *Data leakage and disclosure prevention*
   Solutions to monitor and prevent the unauthorized sharing or leakage of sensitive data, either intentionally or unintentionally should be implemented.

e) *Security auditing and monitoring*

i. Logging and auditing: Detailed logs of data access and processing activities should be maintained reviewed regularly for suspicious activities and potential security breaches.
ii. Intrusion detection and prevention systems: Intrusion detection and prevention systems should be deployed to detect and respond to potential threats and attacks in real-time.

f) *Physical security*
i. Data centre security: Census agencies should ensure that data centres housing census data are physically secure with access controls, surveillance, and environmental protections to prevent unauthorized physical access or damage.

g) *Incident response plan*
Preparations for dealing with security incidents should be developed by creating a well-defined incident response plan. This plan should outline the steps to take in the event of a breach and how to communicate with affected parties.

h) *Compliance and audits*
Regular security audits should be conducted and assessed to verify compliance with security policies and regulations. These should include third-party audits and assessments for independent verification.

i) *Data retention and disposal*
Data retention policies should be defined and enforced data to ensure that data is retained only for the required duration and is securely disposed of when no longer needed.

j) *Employee training and awareness*
i. Security training: Regular security awareness and training programs should be undertaken for employees and contractors to educate them about security best practices and the importance of data protection.
ii. Vendor security: Similarly, if third-party vendors are involved in data processing, census agencies should ensure that they adhere to stringent security standards and practices to safeguard census data.

k) *Public communication*
Develop a clear communication strategy for addressing security incidents, ensuring transparency, and maintaining public trust.

3.303. In summary, maintaining security and confidentiality during census data processing requires a comprehensive approach that includes technical, organizational, and procedural measures. It is crucial to prioritize privacy and security to protect sensitive information and maintain the integrity of the census process.

IX. Census products, dissemination and utilization

A. Introduction

3.304. The population and housing census is a statistical operation of exceptional value to every country. It is the primary source of basic national population data for administration and for many aspects of economic and social planning. Consequently, data from national censuses represent a valuable public good that should be widely promoted by national statistical and census offices in order to enhance its utilization by the various users. Thus, the census should not be an end in itself but should be backed by the value of the results, in terms of utilization, by a wide range of data users.

3.305. While the importance of disseminating census results to the fullest extent possible should be stressed, there are, nevertheless, some things that are essential for the national statistical or census office to keep in mind. Among these is the prerequisite to ensure openness and transparency in the way the results are disseminated. It is equally important that national statistical and census offices maintain professionalism and demonstrate neutrality and objectivity in the presentation and interpretation of the results, and are free from real or perceived political interference so that the objectivity and impartiality of the statistics is assured. This in turn will build trust in, and acceptance of, the results. Furthermore, the disseminated census results should be of sufficient quality to meet user's stated needs, and safeguards should be in place to ensure individual information is kept confidential.

B. Plans for census products and data dissemination

3.306. In order to maximize the utilization of results from their population and housing censuses, national statistical and census offices should have a sound dissemination programme whose objective is to promote the benefits and applications of census data. The statistical or census office should develop and implement an effective strategy for producing and disseminating output products and providing related services based on the demonstrated needs of the diverse users of census data. What follows are some salient issues for an effective dissemination programme.

1. Developing a dissemination strategy

3.307. A census is not complete until the information collected is made available to users in a form and scale suited to their needs. In order to fulfil this requirement, it is essential to develop a strategy for producing and disseminating outputs taking into account all potential users of the data. This in turn requires identifying potential users of census data and their demonstrated needs at a very early stage in the census planning programme so that appropriate products and related services can be developed in good time (See section on user consultation communication and publicity).

3.308. An effective dissemination strategy should have a diversified approach for meeting the requirements of different users. Appropriate technologies and media need to be identified for effective and easy dissemination of census data and information. When developing a dissemination strategy for census data, it is important to take into account geospatial information to provide a spatial context to the census data. Geospatial data can be used to create thematic maps that visualize census information to make it more user friendly. Geospatially-enabled statistics, in particular at a high level of resolution, can allow statistical organisations to produce new analytical outputs and conduct a wide range of spatial analysis.

3.309. A number of key elements should be taken into account in the development of a strategy for census data dissemination, including:

a. identifying the diverse categories of users and their data needs and uses through user consultation and stakeholder engagement
b. the media of dissemination
c. metadata to aid in the interpretation of the results
d. legal and ethical consideration, confidentiality and privacy measures
e. assessing the required technologies to meet user needs
f. dissemination policy
g. quality assurance in terms of accuracy and timeliness
h. available financial and human resources
i. the need for continuous improvement along with a feedback mechanism for dissemination products,
j. alignment of census outputs with international standards.
k. census calendar.

These elements are summarized below, or a cross reference to another location in the document is provided.

Identifying the diverse categories of users and their data needs and uses through user consultation and stakeholder engagement

3.310. The demand for, and use of, statistical products and services must drive all census operations. National statistical and census offices should have a sound strategy for developing suitable products and services to respond to the diverse needs of data users so as to promote the utilization of census results. Such strategies should be based on an active dialogue with the users regarding their needs in terms of products and the format of those products.

3.311. The user consultation process in terms of census products is a major factor in the development of a dissemination programme. The type of consultation discussed in this section complements the consultation that is undertaken to determine census content (see paragraphs 2.159–2.162). The work done at this stage of the census is important in achieving the objective that the census is relevant to users, which is a major indicator of the quality of the census. The selection of suitable census data products and related services should be guided by a detailed assessment of user requirements with due consideration given to budgetary and human resource constraints.

3.312. Plans for what and how products will be disseminated should be made early enough in the planning process and shared with potential users in order to get their feedback. Based on this feedback, the national statistical or census office can tailor its data dissemination programme to suit the requirements of the users. Maintaining continuous communication and obtaining feedback from users is also important for making modifications to products and services, including being able to respond to user requests that become known later in the programme.

3.313. Based on the foregoing, it is important to note that the supply of census products and services goes far beyond the first couple of years after the census. It is important, therefore, that budget and human resources are available for many years after the end of the census collection activities.

3.314. A wide range of statistical products can be made available to different users. A detailed plan for producing different census outputs should be guided by early user consultations (see paragraphs 2.159–2.162) to ensure data and information requirements will be met in a format commensurate with user needs and demands; such a plan will also be a useful guide to prioritizing data processing and

tabulations. To help communicate the timeline for dissemination, a calendar with the expected release dates for the preliminary results, final results, as well as the timing of the release of the various census products should be published.

3.315. The definition of geographic output products and services and the scheduling of their release needs to be closely coordinated with the timetable for the overall census project. The tabulation of census data may require information from the census geographic unit, and thematic maps and digital geographic databases can only be completed once census data processing has been completed.

The media of dissemination

3.316. With the rapid development of technology, census data users have an increasing interest in a broad range of products and services from the census organization. The types of output that census offices may produce and disseminate must be current and may include printed products, static electronic products, interactive electronic products, APIs, customized products, user-interactive products and special audience products and services. Partnerships with key stakeholders are encouraged in the development of the various census products.

3.317. Some data users will need specialized products that the census organization is not planning to produce as part of the general census programme. It is recommended that the census organization establish a service to meet such specialized requests. The pricing of such special products and services may be included in any pricing policy.

3.318. Printed publications are becoming less popular as many countries move towards digital publications. For those countries that produce printed publications, target dates for publication should be determined well in advance and processing and printing programmes should be planned accordingly. In addition to traditional methods of printing, there are various methods of reproduction available that are fast, economical and of good quality, and these should be investigated. For an increasing number of users, environmental concerns and a lack of searchability means online electronic data dissemination is preferred over printed paper. Electronic dissemination reduces cost, and unlike weighty paper publications, supports the direct availability of the data for further computer processing. In addition to pre-processed tabulations, more and more countries are offering interactive dissemination tools enabling clients to obtain customisable outputs such as tabulations, maps, etc. Finally, sample data at the unit record level are also provided by some countries for research purposes. In such cases, the sample should be carefully drawn to ensure an adequate level of representation while at the same time ensuring that anonymity is not compromised. Some countries have also adopted very creative techniques for data dissemination, cross-tabulations, infographics, story maps and visualization to allow users to understand relationships and patterns within the data. Census organizations should consider developing visualization products incorporating storytelling—a technique where narratives unfold as the user scrolls down a webpage—since it offers a compelling method to communicate key insights from censuses to non-expert audiences. The development of such dissemination products should be part of the planning process of the census.

3.319. The information stored in the census database allows fast and relatively inexpensive production of additional tables, maps and other data products. Traditionally, countries have often offered on-demand services to provide census information to users who require tables or other outputs not produced, or aggregates not available, through other means. Now that electronic dissemination is the new standard, often including user-friendly interactive dissemination tools, customized tabulations and

applications can often be extracted directly by end users. In this case, the census organization should prepare in advance and then implement an authorization and data security and disclosure control strategy, so that additional disclosure risks entailed by providing customized outputs are accounted for.

Metadata to aid in the interpretation of the results

3.320. See section on metadata.

Legal and ethical consideration, confidentiality and privacy measures

3.321. See section on confidentiality and privacy.

Technology

3.322. Given the widespread availability and use of technology for easier production and access to census products, national statistical and census offices should evaluate which ones are best suited to the needs of their data users taking into account budgetary and human resource constraints. These technologies include use of GIS and thematic map generators, new sophisticated data base systems and interactive web access, including client-customizable and user-friendly table builders. This may be complemented by application programming interfaces (APIs) to enable standardised automated data retrieval for heavy users. Interactive online tools are becoming the main dissemination channel in more and more countries. A high level of client customizability and automation of output retrieval may also increase efficiency, as it reduces the need for staff supported on-demand services. On the other hand, very flexible tools that serve highly customisable user requests by direct queries to the underlying microdata entail specific additional disclosure risks that need to be accounted for by dedicated disclosure control strategies (see paragraph 3.479). When developing websites or web-based applications aimed at disseminating the census data, it is recommended to ensure inclusivity and accessibility for all users by considering the adoption of the Web Content Accessibility Guidelines (WCAG), an internationally recognized ISO standard. Following WCAG guidelines will enhance the usability and navigability of the websites for individuals with disabilities, ensuring equal access to information.

Dissemination policy

3.323. When planning the dissemination programme, a dissemination policy should be established as well. This policy should cover issues such as ways of marketing the census products, which in most cases means mainly how to inform a wide range of potential users about the availability of the products. Where applicable, a clear pricing policy should also be determined in countries where not all outputs are made available free of charge, and a decision made as to the conditions under which external distributors are permitted to further disseminate census data. The dissemination policy should be transparent and cover relevant issues connected with the protection of the confidentiality and privacy of personal data, and the broad measures that will be used for each of the different products. Disseminating census data (other than microdata) under a clear, open license such as CC-BY 4.0 is recommended to ensure that users have a precise understanding of the terms of use, including redistribution and transformation rights. CC-BY 4.0, for instance, mandates attribution, which is crucial for acknowledging the original data source and maintaining transparency. Moreover, it requires users to indicate any modifications or derivatives, preventing misrepresentation. However, if the government

has already established an open data license, it should be utilized. This approach promotes responsible and ethical data practices.

Quality assurance

3.324. Quality refers primarily to user needs and satisfaction. Even if data are accurate, they do not have sufficient quality if they are produced too late to be useful, or cannot be easily accessed, or conflict with other credible data. Therefore, quality is increasingly approached as a multi-dimensional concept. And the output of any statistical exercise should possess the following attributes: accuracy, relevance, reliability, timeliness, punctuality, accessibility, clarity, coherence, comparability and metadata.[82] Management of quality in census dissemination is driven by concerns to (a) deliver relevant products and services while (b) maintaining accuracy of the data, and (c) timeliness and reliability of data release within agreed cost constraints. See chapter on quality assurance for more details.

Budget and human resources

3.325. Two obvious key elements (usually constraints) in the development of strategies for census data dissemination are the budget that can be allocated and the availability of human resources. With the high relevance of new technologies in all the census stages, and in particular for data dissemination, this is a factor that needs to be carefully analysed when deciding about the specific strategy of census data dissemination. For instance, interactive client-customisable dissemination tools may increase efficiency (see 'Technology' above). The alternative to the recruitment of human resources may be contracting out some of the dissemination activities, in particular those connected with the development of more sophisticated systems. However, this solution needs to be carefully considered, and it is extremely important to ensure that the contractor is committed to the census project until its very end, and that at least some of the new abilities remain in the organization for further use.

The need for continuous improvement along with a feedback mechanism for dissemination products

3.326. See the section on Dissemination in the chapter on Quality Assurance

Alignment of census outputs with international standards

3.327. See section on Census data dissemination: products and services.

Census release calendar

3.328. To help communicate the timeline for dissemination, a calendar with the expected release dates for the preliminary results, final results, as well as the timing of the release of the various census products should be published. See section on Census data dissemination: products and services and Census Calendar for a more detailed description.

[82] United Nations Statistics Division (2012): *Guidelines for the template for a generic national quality assurance framework (NQAF)* – available at http://unstats.un.org/unsd/dnss/docs-nqaf/GUIDELINES%208%20Feb%202012.pdf.

2. Tabulation programme

3.329. In most countries, the tabulation programme represents a compromise between the full range of desired tabulations and the limits imposed by practical circumstances. To ensure that this compromise is made transparently and efficiently it is important that, as previously emphasised, planning the census dissemination task is started at the earliest stage of the census development cycle by a round of user consultations. Once the census-testing programme has identified a practicable range of data items to be included in the questionnaire, data users should again be consulted on the specific cross-tabulations required and the relative priority for their production. It is essential that the programme be outlined sufficiently early so that the procedures and costs involved are investigated thoroughly before a final decision is reached. The type of questionnaire and the method of enumeration may limit the kinds and amounts of data that it is possible to collect. Publication time and costs, and the data-processing resources available, will determine the number and complexity of the tabulations that can be produced within a reasonable time. This will enable prospective census data users to make firm plans and the census data processing staff to complete all systems analysis, programming and testing work in a timely manner.

3.330. The tabulations presented on the website of the United Nations Statistics Division are those fulfilling the most essential, or generally required, information[83]. The databases of census information can be used throughout the intercensal period to address the needs of specialist users for whom these tabulations are not adequate.

3.331. It is important to plan the tabulation programme in such a way that final results can be issued within a reasonable period of time after the enumeration and before the information has become out of date for current needs. It is desirable that the details of the tables be prepared, and the order of their preparation be decided early in the planning so that the processing of the data is not delayed.

3. Geospatial information for analysis and dissemination

3.332. **Geospatial** information can unlock the full potential of census data and facilitate its use for territorial analysis, to describe and investigate phenomena, such as demographic trends and spatial inequalities in accessing services (such as schools, health facilities, etc.), that unfold across time and space.

(a) Dissemination geography

3.333. Maps, which are now commonly found in the form of digital products, play an increasingly important role in the dissemination phase of the census. Statistics compiled from census data can be geographically referenced and provide for methods of analysing the geographic characteristics of those statistics. Maps may then be used effectively to relate statistical data to the geographic area to which the census results refer. This makes the statistics easier to understand and more readily usable by both expert users and the general public.

3.334. If a complete digital census geographic database has been created, then statistical databases for administrative or statistical units can be produced through aggregation. For the countries that do not

[83] Link to UNSD website containing the tabulation shells.

use digital techniques for the production of enumeration area[84] maps, options still exist to develop a digital geo-referenced census database at this stage for producing publication-quality maps to accompany census reports, for distribution to outside users who want to analyse census data spatially, or for internal applications. This database can be compiled for a suitable level of the administrative hierarchy or for other aggregated statistical regions. At that level of aggregation, the resources required for producing a digital database are much less than those necessary for a complete digital enumeration area map database.

(i) Linking collection geography to dissemination geography

3.335. An essential feature of the census is its diversity in terms of the geographic level at which data can be disseminated. This is due to the ability of the census to produce statistics that can be disseminated at the lowest geographic level (small area), through a geographic hierarchy up to the country level and by geocoding at the lowest level of collection to support the production of grid-based outputs. Consequently, one of the earliest decisions in census planning relates to the administrative and geographic areas for which census data on diverse socioeconomic characteristics of the population will be reported and disseminated in order to satisfy the needs of the various data users, taking into consideration statistical confidentiality, in particular where non-nested small areas (such as smallest administrative unit and grid) are disseminated.

3.336. In addition to administrative units, most countries will have a number of other sets of areas that are used for different purposes and for which census data will need to be compiled.[85] Such areas, which have special uses, include health regions, electoral districts, urban agglomeration or metropolitan areas, grids, and utility zones (water or electricity supply districts). It should be noted that some of these areas may not fit perfectly into the administrative hierarchy of the country. It is important, therefore, that to the extent possible these reporting units are taken into account when designing enumeration areas in order to facilitate generation of census data for these regions. This draws attention to the fact that when delineating collection geography (enumeration areas), it is essential that dissemination geography and disclosure risks due to overlapping geographies is kept in mind.

3.337. Two somewhat different methods are available to provide the census with a flexible capability for generating tabulations in terms of a wide variety of geographic aggregations, including those needed for public and private sector data uses at the local level. The first method simply extends the traditional hierarchical system for coding all major and minor civil divisions so as to cover at the lowest level the enumeration area, sometimes referred to as the “enumeration district”. The second method, which at greater cost permits finer geographic specificity, is usually based on some coordinate or grid system, such as latitude and longitude. This method is often referred to as a “geocoding system”. Particularly in the absence of a comprehensive system of street names, numbers or geocoded coordinates of housing and building units, the first method, which uses the enumeration area as the key unit to produce small area data, is to be preferred.

[84] The Enumeration Area or District is only relevant in a traditional or combined census methodology. Where data is derived wholly from administrative sources in a register-based census, such areas do not have any meaning and small area geography must be determined by other boundaries.

[85] *Handbook on Geospatial Infrastructure in Support of Census Activities*, Studies in Methods, No. 104 (United Nations publication, Sales No. E.09.XVII.8).

3.338. Where buildings or housing units have been geocoded, these geocodes can be used to directly allocate each household to the correct enumeration area. The fact that census data are available by enumeration area provides for considerable flexibility. Such flexibility can be of value given that the geographic divisions used by various branches of the administration or by other data users do not always coincide and may therefore require different regroupings. Moreover, when changes are planned in administrative boundaries, tabulation of census data by the planned new entities can also be facilitated through the enumeration area approach. However, if these changes cross enumeration area boundaries and it is decided to try to retabulate the census according to the new boundaries, very complex recoding of individual records may be involved. As an alternative, statistical concordances, showing the quantitative relationship between the previous and current classifications, could be used.

(ii) Uses of small geographic data

3.339. Censuses provide data from the highest to the lowest geographic levels of aggregation. Tabulations from census results yield relevant statistics for any reasonable combination of characteristics for the country as a whole, regions or provinces, down to small areas such as localities, and even enumeration areas and geographic grids. This important feature of the census makes the data amenable to the development of estimates of variables of interest for small and local areas in two major ways: directly from the production of tables from the microlevel data for the required characteristics; and indirectly from applying estimation techniques by combining other sources, such as sample surveys and administrative statistics to the population and housing census results.

3.340. Census data are typically aggregations of data for many individual small areas and may commonly be used to study large regions or entire nations. Data for small areas enable the user to obtain statistical information about any number of local areas of interest, in addition to showing variations among small areas in individual parts of the country. The increased capabilities of data processing systems greatly facilitates the utilization of census results for analysing the information for small areas, limited only by issues of collection design and confidentiality and statistical disclosure when cell entries in cross-tabulations become very small. For example, the analysis of whether or not programmes have improved educational attainment for women and girls at a regional level may be carried out by analysing data from the smallest administrative units so as to observe local variations and produce more accurate assessments of cause and effect.

3.341. Implementation of various national social and economic development programmes is a function of the state, province or lower levels of government in many countries. Results of population and housing censuses are useful for planning and monitoring development at the local area, small town level or small area. Small-area data are also important for private businesses in developing their distribution and marketing strategies. For example, information on housing demand from the population and housing census may be used by local authorities, local real estate companies, building and housing development contractors, and manufacturers of construction materials, among others.

3.342. Census data have been traditionally aggregated by various types of administrative units (for example, towns, villages, provinces and electoral units). In addition, other types of small areas are sometimes used in the census that are essentially statistical in nature (for example, census tracts and grid squares that do not change from census to census, and very small units such as city blocks or block faces). There have also been increasing demands for small-area data that cut across the local administrative boundaries. Population and housing censuses provide a powerful tool for assessing the impact of population on the environment, for example on drainage basins and on water resource

management systems. The spatial units for such a study may combine a group of local administrative areas. In this situation the availability of census databases with mapping capability (see paragraphs 3.432–3.435) is of great importance.

3.343. Tabulations for small areas may be prepared on the basis of the resident population of each area or on the basis of the population present in each area at the time of the census. Tabulations relating to the resident population are produced for the apportionment of representation in legislative bodies, the measurement of internal and international migration, the computation of measures of fertility and mortality by place of residence, and the planning and administration of such services as schools and housing, which have relevance only to the resident population. Tabulations based on the population present in the area at the time of the census are useful where this population is considerably larger than the resident population and thus raises the demand for products and services above the level required by the resident population alone. The combined population and housing census may also be used to make comparisons of resident and daytime populations in specific localities, if an item on place of work is included in the population census. It is therefore important that users express their needs for particular data disseminated in a given format, based on the usual residence or place of enumeration, at an early stage of census preparations.

3.344. It was elaborated in Part One, Chapter I, how the population and housing census plays an essential role in the economic, social and environmental components of the national statistical system and also serves as a sampling frame for sample surveys. Another significant way in which the census results complement survey statistics is in small-area estimation, whereby models constructed from survey data are applied to census results for any specified geographic area. This estimation approach may be used for generating such indicators as employment, poverty and other economic indicators, for which measurement is required at the local area level.[86] The application of small-area estimation techniques to poverty measurement and mapping is an important extension of the use of census results.

(b) Spatial analysis

3.345. Spatial analysis on census data is key to presenting census data in an understandable and accessible way as decision makers are able to relate more easily to information presented in maps than with data tables. For example, the SDGs vision of leaving no one and no place behind requires, at a minimum, analyses and presentation of trends at the highest spatial resolution possible – to show existing (spatial) inequalities among places and people. Census data, because their high resolution is able to accurately showcase prevailing inequalities and, in turn, inform decisions and actions to ensure equitable development. In addition, census data are crucial to producing data disaggregated by location (such as city, towns and suburbs, villages, and rural areas), sex, income groups, and persons with disabilities (among others) as required for monitoring global trends. For example, understanding the share of different groups of urban populations who are within a proximal distance of access to public transport (SDG 11.2.1), open spaces (SDG 11.7.1), or the share of rural populations with access to all-season roads (SDG 9.1.1) or the computation of Ratio of land consumption rate to population growth rate (SDG 11.3.1) requires a clear analysis of census data to depict location of populations across the urban-rural continuum.

[86] *Measuring the Economically Active in Population Censuses: A Handbook*, United Nations publication, Sales No. E.09.XVII.7.

4. Mode of dissemination of outputs

3.346. As has already been mentioned, a census is not complete until the information collected is made available to potential users in a format suited to their needs (paragraph 3.307). Consequently, meeting the needs of data users means that the data producer should not only provide data products to the users, but should also provide them in formats that are suitable to the needs of the users. The information in the products may be included in published tables and reports for general distribution, produced as tables in unpublished form for limited distribution or stored in a database and supplied upon request, or disseminated online either as static or interactive products.

3.347. It should be noted, however, that regardless of mode, all dissemination is subject to issues of (a) quality assurance; (b) possible disclosure of information about identifiable respondents; and (c) copyright and ownership. In addition, the issue of cost recovery has become important to many statistical organizations. Each medium of dissemination has its advantages and limitations, and the choice of one or more of them depends on the context, and on the intended categories of users. In most instances, these methods complement each other and can provide effective ways to reach out to the public and private sectors.

3.348. When data are provided in electronic form, special attention should be given to providing users with easy means of data and metadata retrieval. The options for obtaining the relevant metadata and the data should be accessible in standard and contemporary formats.

(a) Publication of printed tables and reports

3.349. Nearly all countries now disseminate their census results online, and printed publications have become a secondary choice for the dissemination of the main census results.

3.350. When printed publications are still used, the choice of how the actual printing is to be done entails in fact a trade-off involving quality, cost and speed. The best results can usually be obtained by sending the documents in computer-readable format to a professional printing plant[87]. This will allow high-quality typesetting and the use of supporting colours. Alternatively, master printouts can be made in the census office and sent to the printer for cheaper duplication or offset printing. There are also affordable high-speed printing systems that can be directly controlled by the computers in the census office. Publications can also be distributed in electronic format (pdfs) as part of the NSO commitment to reducing paper waste.

3.351. Target dates for online and printed publications should be determined well in advance and processing and reproduction programmes should be planned accordingly. In addition to traditional methods of printing, there are various methods of reproduction available that are rapid, economical and legible, and these should be investigated.

(b) Online dissemination

3.352. For most users, online is now the preferred channel of to access census outputs. NSOs are increasingly using cloud-based solutions for the storage, security, management, and dissemination of census data to users. Online dissemination of all kinds of information, including statistical information,

[87] In some countries the printing of census reports is often managed by an official government publications agency that is responsible for maintaining quality standards

has increased with new innovative formats for displaying census data. The advantages of online dissemination are found primarily in terms of speed, flexibility and cost, and making results accessible to a wide range of data users. Online census dissemination should feature a modern interface and API as well as good search and filtering abilities.

3.353. Online dissemination of data was common well before the Internet gained prominence. The same website could be used for both internal and broad community communication, with the granting of access rights in certain areas to privileged users only. Security measures, including passwords and callback procedures, can be used to exclude unauthorized users from reaching these areas. This is however risky, since resourceful hackers may find their way around the barriers and gain access to confidential information. The census office website is probably the first dissemination medium where Internet-connected users would look for census information. It is recommended that microdata should not be stored on a website in direct contact with the public. It is also recommended that a powerful firewall constitute a security layer between the website that is visible to the public and the working network of the census office. Firewalls are hardware or software security systems that limit the exposure of a computer or network to malicious infiltration from an external location. Websites of public administrations are under constant attack from hackers and very sophisticated security measures must be adopted when disseminating interactive census products on the Internet. Cyber security, despite being an issue of a technical nature, has to be mandated, demanded and resourced by the highest levels of management of the census office.

3.354. Flexible table builder software provides the potential for users to access census microdata and create their own multivariate output tables at different levels of geography. The software runs statistical disclosure control while the request is being processed to ensure protection of individual identities in detailed statistics, making sure that no individual can be picked out in the tables. It allows NSIs to release billions of anonymised census statistics far more quickly than ever before. A key design choice with profound implications for data security and confidentiality is whether such flexible table builders query the underlying microdata in the background, or a set of very detailed pre-curated queries. Direct microdata queries may lead to additional disclosure risks even with automatic disclosure control methods running in the software.

3.355. A website can be used not only to make information available as soon as it has been cleared, but also for other forms of communication with users. Possibilities include online ordering of publications and one or more receiving areas for questions that would be answered later through the same medium by appropriate specialists. One such area could be the census forum or “chat room”.

3.356. Internet websites may support “door” or “gateway” applications that allow users to run outside programs on the computer on which the Internet web server operates. Interactive access to census outputs can be offered to most types of databases and census products, including reports, publications, tables, maps and graphs. For example, there may be a database of aggregated census data for small areas or a microdata database that users can access in this way. When the required data are not readily available, users could run an on-the-spot query to obtain and retrieve results that satisfy their needs. This can be done by offering to Internet users census microdata samples and an interactive tabulation system. Users can then select records from these data sets that satisfy certain parameters and compute statistical information, such as two-dimensional cross-tabulations of either original or recoded variables. Program execution by users, however, raises important questions of cost, efficiency and confidentiality, which have to be resolved. For reasons of efficiency, it is recommended that information that is provided or likely to be heavily requested by users accessing the census website be made available in a

static format, which is faster to download. Letting the user run data extraction on online databases, which would be a dynamic way of accessing the census information, is more resource consuming and should be the second choice for those users needing more detailed data than those available through static pages.

3.357. Other media such as social media are useful in disseminating census information targeted at different sectors of the population. More generalist media, such as the radio, television programmes, newspapers and press conferences, offer the possibility of reaching out to sectors of the population not otherwise reachable.

3.358. A hybrid solution for data dissemination that appears to combine the advantages of several approaches is one whereby the statistical or census organization makes basic data available to users online, while additional information may be provided by some other online protocols, such as file transfer protocol sites. This will usually take the form of a package that contains basic data, metadata and data browser software. The basic data may contain existing time series, report files and the like, as well as country and region maps that can be used to generate thematic maps with various indicators. Maps made available to general users need not ensure the same geographic detail as maps used for enumeration areas. Lighter versions of maps at any subnational level may be provided to the general public, and more sophisticated and detailed ones to those fewer users who would actually need an increased level of detail. It is thus important that the website specify the instructions on how to contact officers responsible for special dissemination needs.

3.359. For some users, if the particular statistical information is not yet available on the physical distribution medium, special access may be granted, provided that adequate screening of their credentials and security checks are performed, to protected areas of the Internet site where up-to-date census information becomes available. Since opening up online resources to users has to be planned carefully and a clear policy established in advance (so that criteria for deciding whether or not to grant access are unambiguous), it is not recommended. Instead, provision of an online data tabulation system for expert end users is advised.

5. Confidentiality and privacy

3.360. According to principle 6 of the Fundamental Principles of Official Statistics, “Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes”.[88] Maintaining data confidentiality is an indispensable element for retaining the trust of respondents. If respondents believe or perceive that a national statistical or census office does not protect the confidentiality of their data, they are less likely to cooperate or provide accurate responses to the census questionnaire. This in turn affects the quality of the statistics.

3.361. All the information stored in the census database allows the production of tables both for very small areas (such as enumeration areas or villages) and for all individual units in these areas. Therefore, two key considerations guide the construction of a census database: technical aspects and ensuring confidentiality. Protecting individual privacy is paramount, and it should influence the design of both data collection and data processing programs. Accordingly, microdata, such as name and local address,

[88] http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx.

or the unique characteristics that permit the identification of individual respondents, must be removed from the database or encrypted.

3.362. The same care must be taken if a transcription of information from original questionnaires (that is to say, from a representative sample) is needed for use by qualified agencies and research institutes engaged in special studies beyond the purview of the regular census programme. Such needs have sharply decreased with the almost universal use of computer technology. However, when such a procedure is possible under the census law, individual privacy should be ensured in all cases.

3.363. The ever-increasing demand from users for more data, especially microdata and at lower geographic levels, and also with more technological advancement for data linking, particularly over the Internet, has created more challenges for managing data confidentiality. As a result, national statistical and census offices should examine the data and make modifications, when necessary, prior to dissemination of the data. The objective of the modifications is to prevent identification of individual respondents, and also intentional or inadvertent disclosure of their personal information. This is particularly the case when microdata is disseminated and when data are linked to location, such as with the use of GIS.

3.364. Statistical confidentiality methods range from traditional methods like cell suppression, rounding, Barnardisation[89], or recoding to more modern methods based on the introduction of random noise. Depending on the method, the implementation complexity also ranges from a few simple processing steps to elaborate statistical applications for entire databases[90]. The approaches used to limit disclosure are tailored according to the type of data and the product to be disseminated.

3.365. Different techniques are also employed depending upon the type of data product to be released (e.g. microdata files, tables, interactive data explorers, maps).[91] For example, disclosure methods for **microdata** include (i) suppression of direct identifiers, (ii) recoding or local suppression for quasi-identifiers, (iii) noise addition, and (iv) suppression of sensitive variables (v) suppression of outlying values. More recently, differential privacy and synthetic data have been proposed as safer approaches.

3.366. For **tabular/query data**, more traditional methods aiming to mask confidential information (e.g. cell suppression, recoding) have been widely used over the past decades. However, these methods have meanwhile been found to suffer from additional loss of data security and utility in modern dissemination scenarios with very detailed tables particularly when combined with various other data products. Emerging alternatives are predominantly based on adding some uncertainty to the outputs in a consistent manner (noise injection).

3.367. **Interactive data explorers** are often based on flexible rules for users to submit custom table queries (table builders), where additional disclosure risks are present in implementations where highly flexible queries are evaluated directly from the microdata (as opposed to querying a curated fixed set of detailed tabulations stored in the background).

---

[89] A method of statistical disclosure control for tables of counts. It involves adding +1, 0 or -1 to some or all of the internal non-zero cells in a table.

[90] Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Schulte Nordholt, E., Spicer, K., de Wolf, P. (2012). Statistical Disclosure Control. Germany: Wiley.

[91] de Wolf, V.A. (2003). Issues in accessing and sharing confidential survey and social science data. *Data Science Journal* 2(17): 66–74.

3.368. Finally, **maps** are becoming a standard data product from censuses, more and more often with interactive functions. Specific additional disclosure risks may emerge from very detailed geographical outputs[92] when these are not nested (geographic differencing – for example between very small administrative or enumeration units and grids).

3.369. The disclosure of confidential information in the context of the dissemination of census data integrated with geospatial information requires careful attention. Geocoded census data are becoming available at increasingly finer resolutions, increasing the risk of data disclosure. Geospatial information adds a new dimension to data with which an individual statistical unit can be more easily identified in combination with other information.

3.370. When an interactive mapping application is used, the tool should be configured in a way that users are not allowed to drill down to the spatial resolution level that unit record data might be disclosed. When microdata are released with geographic coordinates, geomasking techniques should be used to anonymize their precise location and prevent identity disclosure. Since aggregated data with spatial attributes provide coarsened location information of the respondents, their privacy can be protected by data suppression on the areas with low values. Geospatial products should be cross-checked with other dissemination products (such as tabular aggregates or micro datasets) before release so that they do not breach confidentiality on their own as well as in combination with other outputs.

6. Metadata

3.371. In order to assist data users to better understand and interpret the data, it is important that there is adequate documentation providing a complete and clear description of the production process, including data sources, concepts, definitions and methods used[93]. This information represents metadata that, it is recommended, should accompany all census products. Providing metadata promotes transparency and credibility of census results. Additionally, dissemination of census products with accompanying metadata ensures harmonization and comparability of census data with other data sets. International metadata standards (such as the DDI Codebook standard for microdata or the ISO19139 for geographic datasets) provide a convenient solution to ensure completeness of the metadata and to make metadata machine readable (which can help make the data more discoverable, visible, and reusable). SDMX is an example of an exchange and data warehouse tool.

(a) Definition and content

3.372. Metadata comprise descriptive and structured information or documentation about data that informs users about the content, quality and condition of the data. In this context, metadata offer guidance on the proper usage or interpretation of data by providing information on the processes of production and describing the structure of data sets, thereby making it easier to retrieve, use or manage the data. Metadata constitute a standardized way of organizing data and can be categorized as follows: (a) reference metadata, which allow understanding and interpretation of the corresponding statistical data by describing the concepts, definitions, methodology and quality of data, production and dissemination processes, data access conditions, etc.; and (b) structural metadata, or “data about data”,

[92] Where maps illustrate percentage distributions rather than absolute numbers, there are far fewer risks of disclosure.
[93] See Annex for an example of a metadata checklist (Census_Metadata_Checlklist.docx)

which provide information about the structure of the data set and act as identifiers and descriptors of the data, making it possible to properly identify, retrieve, browse and further process the data.

(b) Uses of metadata

3.373. Comprehensive and readily accessible metadata foster the comprehension of the statistical data being disseminated, public trust in the data, and informed decision-making. Detailed metadata shed light on underlying concepts, methods, and definitions, thereby fortifying the accurate interpretation of the census results. Rich metadata also bolster data assessment and discoverability, enabling users to navigate and locate pertinent information efficiently, and to evaluate the fitness of the data for their specific purposes. Moreover, by increasing transparency and credibility, meticulous metadata empower both individuals and electronic systems to reliably capture, interpret, and present statistical and geospatial information.

3.374. All tabulations should include the following metadata or references to where this information can be obtained: census questions; reasons why they are asked; conceptual definitions (census dictionary); geographic hierarchies used; changes since the previous census with regard to content, operational methods or geographic boundaries; and quality indicators such as coverage rates and item non-response. Data files must also be accompanied with metadata, including names and codes for common variables, personal files and household files. If a long-form sample is used in the census, metadata should also provide information on the sampling variability of the results. When the census tabulations include suppressed data cells due to small numbers, the metadata should also include a methodological note on the rules and methods of suppression[94]. Metadata should be preserved for future reference. With the increased use of technology, properly designed metadata systems for machine-readability and web-based applications are recommended.

(c) Geospatial metadata

3.375. Geospatially integrating social and demographic data necessitates the inclusion of additional metadata to the statistical record, preferably in the form of x- and y-coordinates but other forms of geographic information, through a "geocode" such as an address or locality can be used. Furthermore, additional metadata including details about the coordinate system, projection, the precision of the geocoded locations, as well as data confidentiality measures, such as any spatial aggregation or masking techniques used to protect sensitive locations should be included to facilitate interoperability, appropriate use, and correct interpretation of the data.

3.376. Compared to statistical metadata, there is limited awareness and understanding in statistical organisations on the metadata associated with geospatial information and services. Geospatial information is in the form of various types, and this great variety as well as methodologies / technologies involved in producing the data adds challenges for statistical organisations to standardise geospatial metadata and systematically manage it.

3.377. Given the wide scope of geospatial metadata, it is important to first investigate crucial metadata elements (e.g. data type (point, line or polygon), time stamp, and coordinate system) needed

[94] The rules and methods of suppression should be described in a manner that protects against the possibility of disclosure by reversing the process and deriving the unsuppressed data.

for different stages of production and determine a core metadata set and standards (for example ISO 19115, ISO 19119, GeoDCAT) to follow at the corporate level. After priorities are defined, a continuous improvement process could be put in place to gradually improve the scope covered by the metadata. It is important to have a corporate strategy in place to build a consistent metadata system to avoid compliance issues with other existing metadata systems. As for the statistical metadata, the geospatial metadata should be managed and updated continuously throughout the production process as the changes affect downstream tasks and influence the final outputs.

3.378. Alignment harmonisation of geospatial metadata concepts with those of statistical metadata in existing metadata systems is critical, and there has been an increasing effort to connect statistical metadata with geospatial metadata (for example technical specification expanded for geospatial metadata in SDMX 3.022).

7. Promotion of, and training on, uses of census data

3.379. The main purpose of a census is to collect, process and disseminate information that will be used as the basis of informed, evidence-based decision-making. The benefits of this approach to decisions are not always apparent to users, especially in situations where other approaches may have been used in the past. It is therefore important to promote such uses of census results among users.

3.380. In other cases, users may be willing to use the information but require additional training to more fully understand the data. Such training may be usefully combined with training in statistical dissemination techniques or uses of more advanced data products. At a very basic level, some users may require training in such mundane issues as how to contact the national statistical office, or how to find the information they require within the systems of that office, or how to use the website and other electronic applications and tools. This training can be delivered via video tutorials, podcasts or e-learning platforms.

3.381. Whichever approach is taken to enhancing promotion of, and training in, the use of statistical data, a number of strategic issues need to be addressed. These include:

a. Ensuring that the needs for training are identified early in the census planning process and that required funds are included in the census budget. In this regard it should be noted that in many cases the courses requested by users will be specific to those users; in such cases it may be desirable to request the user to provide funds to cover the marginal (or full) costs of the course.
b. The proposed courses or materials should be fully integrated into the overall census advocacy or training programme. It is essential that messages about the use of data fully reflect the messages given when initially advocating conducting the census or seeking public cooperation with and participation in the collection phase.
c. If the training facility is itself promoted properly, it is highly likely that the demand for training will far outstrip the ability of the statistical office to deliver it. In this case it will be necessary for the statistical office to have prepared transparent strategies that (a) identify those areas in which the statistical office wishes to participate (for example, dealing with lifeline clients, and topics on which the statistical office has particular knowledge or expertise); (b) establish partnerships with other bodies to provide training; (c) use approaches other than classroom training to provide learning-at-a-distance opportunities (for example, e-learning); and (d) have a pricing regime to cover costs where this is seen as desirable.

3.382. The list of target audiences and topics for such training must clearly be determined by countries. It should be noted, however, that basic training in the use and interpretation of the results of one census is a very strong method of advocating support for future censuses. It is thus recommended that countries consider development of a basic course in (a) potential uses of census data; (b) how to access census data; (c) interpretation of census data at the broadest level, including the interpretation of its completeness and level of accuracy; and (d) spatial analysis. The target audience for such training should be, in the first instance, key decision makers in the political and administrative hierarchy of the country. It should be outlined that the uses of census data at the local or small areas level offer substantial potential for constructive use of census data; the spatial distribution of population by age and sex, for example, provides an ideal framework for local officials to address the most pressing issues of their constituents, such as location of schools, utilities, service delivery and so forth.

3.383. A second group of key importance are members of the mass media, such as print, radio and television journalists. A focus on training such personalities is important because they can carry the message to a wide audience. This will assist in the general raising of awareness in the population at large, as well as in generating an awareness of the census among the governmental, academic and business users who may not have contact with the statistical office on a regular basis. Obviously, such training should be completely integrated with the overall public relations and advocacy work.

3.384. A third group to be targeted are schools - both students and teachers. A focus on training of teachers to use census data in the mathematics, social sciences and geography curricula creates awareness among children about using statistics in decision-making and allows them to develop numerical skills using real data (see section on education material for schools).

3.385. A fourth group to be targeted are geographers, with the aim of integrating census and survey data with GIS shapefiles in order to perform spatial statistical analysis. This training will enable specialists to better present statistics geographically.

3.386. A fifth group to be targeted is the research community and civil society organisations. Researchers will benefit from focused training on the application of various statistical techniques to census data to enhance data utilization. Civil society organizations will require briefings on key census findings relevant to their constituencies to support advocacy and program development.

3.387. A sixth group to be targeted are non-users of census data. A number of stakeholders are unaware of how census data can be used in their area of work to make evidence-based decisions. Countries need to aim to increase the usefulness of census data by identifying non-users.

C. Census data dissemination: products and services

1. Provisional and final results

3.388. Some countries release provisional results very soon after enumeration is completed. Subject to change once the full data-processing and verification operations have been completed, they nevertheless provide a general picture of population trends. Provisional census results may be processed manually or by computer. For reasons of efficiency and quality, the use of computers is always preferable. The ability to verify data quality during the enumeration phase with the help of validation programs, quick indicator reports, data consistency reports, and tabulations greatly increases

the confidence with which provisional results can be announced. Provisional results will normally cover information only on total population by sex and by major division. The number of households and housing units may also be derived easily from this exercise. The provisional result of the census can be reported right after the end of the enumeration by utilizing the summary of household lists without individual data processing. This can be possible as these summaries usually include the total population, households and housing unit in each major division.

3.389. The final census results will be the output of the main tabulation programme. Tabulations may be based on all of the returns or on a sample. If some of the topics are collected on a sample basis only, proper weights will have to be applied in the tabulation stage to produce valid national estimates. In addition, the census office should be prepared to facilitate the production of tables requested by researchers and users (see paragraphs 3.486–3.492).

3.390. Since provisional and final results may differ (for example, the summaries on which provisional results were based might contain errors), it is important that data users be made aware of and warned about the possibility of such differences. Implications of using provisional population counts must be outlined. It is recommended that quality assurance processes be put in place to minimize variances between the provisional and final results.

3.391. The final census results should be published as soon as possible. Countries may aim to publish the basic, essential results within one year of enumeration. The use of technology may reduce the time between the release of the provisional and final results, which may over time render provisional results obsolete. The dissemination of the final census results should be part of a comprehensive dissemination strategy and plan. The schedule and description of upcoming releases of final results and products should be made public early in the process to maintain interest by the public in the census (see also release calendar, paragraphs 2.184–2.191). The releases can be staggered, from simple, descriptive one-page summary fact sheets covering a country's major geographic divisions initially, to more comprehensive tabulations and descriptive reports later on.

2. Census reports

(a) Basic reports

3.392. Every effort should be made to publish the principal results of a population census (such as those on age, sex and geographic distribution of the population) and of a housing census (such as a geographic distribution of sets of living quarters, households and population by type of living quarters) as soon as possible after the enumeration, otherwise their usefulness and the extent of their interest to the public will be diminished. With technological advancements, the time required for processing and tabulating results has been significantly reduced. As a result, collection restrictions, in terms of cost and accuracy of the data, have a greater relative weight in determining the number and complexity of the tabulations that can be produced and disseminated.

3.393. Creating a release schedule for census outputs and making it public can provide census organizations with the opportunity to influence the reception of statistical releases. It ensures that stakeholders know when to expect particular outputs, which may impact their plans for making decisions, producing complementary publications, or updating information they hold. It can also ensure that the media is prepared to publish news on specific topics, and advance press releases for each topic can help messages land through the media with the necessary context and sensitivity. Releasing

univariate outputs at a steady pace, instead of in batches, can improve the ability of census organizations to guide their audiences through the key news and information in census outputs. This can help ensure that the most important stories are given appropriate prominence.

3.394. Consultation and engagement with users should be considered when determining the timing and prioritization of different outputs in the release schedule. Census organizations should be cautious of optimism bias when assembling a release schedule. They should carefully consider how and when to communicate the release schedule to users before making it public. There is a significant risk of failing to meet users' expectations relating to the timeliness or quality of census outputs, which can impact trust. In some cases, it may be necessary to deviate from a planned release schedule to respond to emerging priority user needs for information.

3.395. A successful approach could be to announce the broad periods in which different types of releases will be released, such as the first results, the univariate topic summaries, and the multivariate data and detailed analysis. Detailed plans should be released for the relevant phase at its commencement, or with dates telegraphed on a rolling basis.

3.396. Census organizations should consider which tabulations and analysis should be produced and in which format, such as paper reports or online. "Flexible table builder" technologies can be used to complement fixed multivariate tabulations and analyses that may be produced for priority user needs. Subject to relevant statistical disclosure control measures, flexible table builders enable users to create cross-tabulations across variables and geographies of interest. This can reduce the resource and time requirements for meeting a wider range of needs.

3.397. As with communications during the census operation, accessibility requirements should be considered at the outputs phase, ensuring that outputs can be accessed by a variety of audiences. This could entail producing outputs in a variety of formats.

3.398. The population and housing census tabulations presented and illustrated at the website of the United Nations World Population and Housing Census Programme are intended to provide, in tabular form, the most important census information needed as a basis for programmes of economic and social development and to be used for research purposes. They do not in any way represent all of the tabulations that a given country may publish and certainly not all of the tabulations that may eventually be prepared for special purposes. The tabulations do not take into account the form in which information may be entered into a database, which may be more detailed than that required for these illustrative census tabulations.

3.399. A major goal of these recommendations is to provide a set of tabulations that need to be produced at the lowest geographic level pertaining to the same point in time so that a country or area is able to meet its data needs for evidence-based socioeconomic development planning and monitoring. While the majority of national statistical authorities use a population and housing census as the single most comprehensive vehicle to collect these necessary statistics, others use sample surveys, registers of population and vital events, and other administrative sources or a combination of these methods to derive them.

3.400. Three categories of tabulations are described below: (a) basic or essential, (b) recommended, and (c) optimum tabulations.

(i) Basic or essential tabulations

3.401. These are tabulations that are deemed of top priority for production by countries. They are also regarded as essential for countries in difficult circumstances, such as those that have emerged from a conflict or those that have not carried out a census in a long time, in terms of providing minimum statistics to meet their basic data needs.

(ii) Recommended tabulations

3.402. Recommended tabulations are those that are considered adequate for meeting the essential data needs for evidence-based planning, monitoring and implementation of national policies because of their perceived relevance at both the national and the international levels. These tabulations are also designed with the potential for producing statistics at the lowest geographic level and are expected to be produced by each country.

3.403. The recommended set of tabulations also includes the basic or essential tabulations discussed above. Schematic presentations of all tabulations are presented online at the website of the United Nations World Population and Housing Census Programme[95].

3.404. Associated with the recommended tabulations are the core topics that go into their production. Core topics are therefore the main variables for the recommended tabulations (for a more detailed discussion of the topics, see the section “List of topics” in part V).

3.405. As stated in paragraph 5.1, the aim of the recommended tabulations is to permit national and international comparability of data due to use of common concepts and definitions of the core topics. For each of the recommended tabulations, the core topics that it represents should be listed as part of the metadata. Other metadata that are presented for each of the recommended tabulations should include:

l. the source of statistics, that is to say, whether from a traditional census, register-based census, survey or rolling survey;
m. the type of population count, that is to say, whether a de jure or de facto population or a combination of these; and
n. the definition of urban and rural areas used.

(iii) Optimum tabulations

3.406. The optimum set of tabulations includes the basic or essential and the recommended tabulations discussed above, as well as additional tabulations, and is designed to meet the needs of most of the users at the national and the international levels. This set can be viewed as being equivalent to the complete set of tabulations that could be generated from a population and housing census.

3.407. In order to avoid producing census tabulations that are overly voluminous or that contain a large number of empty cells, some countries may find it necessary to employ a more restricted geographic classification than that suggested in the illustrations. For example, basic facilities such as piped water or

[95] Link to UNSD website containing the tabulation shells.

electricity may be almost completely lacking for large areas of some countries. Under these circumstances, tabulation of the relevant data for small geographic areas would not be appropriate. The geographic classification to be utilized needs to be carefully considered, taking into account the type of information being tabulated, its probable frequency distribution and the uses to which the data are likely to be put. Privacy and confidentiality of individuals and households must always be protected (paragraphs 3.360–3.370, 3.488).

3.408. As well as collecting data on core topics, some countries may also collect data on additional topics in the census questionnaire to address specific concerns, for example, whether or not the birth of an individual is registered, the age at which a woman first marries, or vocational and technical skills. In other cases, detailed tabulations for specific population sub-groups may be required for use in planning or evaluation of programmes. Tabulations for the non-core topics may be done after the basic tabulations are completed. This should be based on user needs. Consequently, consultations with user groups at both the national and local levels may be helpful in determining the most suitable tabulation plan and method of dissemination.

(b) Thematic statistical or analytical reports

3.409. Many countries prepare different types of thematic or analytical reports. These reports should be planned and scheduled during the preparatory phase and published according to the release calendar in order to avoid outdated reports. The reports may range from volumes presenting extensive and detailed statistical tabulations, particularly cross-tabulations, to more analytical reports that combine tabular materials with some interpretative or analytical text. This latter group of reports might include, for example, volumes of regional analysis on such subjects as population or housing conditions of urban areas, major metropolitan areas or big cities, and regional distributions; locality reports on infrastructure; and comparisons of key social indicators such as education, living arrangements, housing conditions, sanitation, disability and economic activities. Other such reports might include community profile analysis of, for example, indigenous populations, and profiles of specific population groups, such as families, children, youths, persons with disabilities, forcibly displaced and older persons. Reports on population growth and distribution that examine changes in the demographic characteristics of the country's population with breakdowns by two or three levels of administrative areas would be very useful. Such reports might focus on the growth, location and mobility of the population at the national and regional levels, and administrative areas. It is important that appropriate language is used to correspond to the target audience for each thematic report. It is recommended that multidisciplinary task teams be established, including line ministries and agencies, for the preparation of thematic and analytical reports in line with agreed guidelines. Partnership and external cooperation with academic institutions and other specialists in subject matter, which can facilitate such work and strengthen collaborations, may be sought whenever possible.

3.410. Thematic and analytical reports must be based on user needs and respond to a country's specific development needs and emerging issues. These reports can also be used to show time series and trend analyses of socioeconomic and demographic indicators and may combine census data with other data sources to provide a more comprehensive and current outlook.

(c) Methodological reports

3.411. Other published reports may include the census methodology, encompassing, if applicable, sampling design and methodology and a census quality and/or evaluation report, which may include estimates of census coverage and the methodology used for their preparation.

3.412. It is important that users of census products be provided on a timely basis with as much relevant information regarding the census as possible. A prospectus that contains information on all types of products that will be available following the census is very useful to users. A brief description of each product should be provided including the estimated timing of release, the level of geographic detail that each product carries and, for products released periodically, the frequency of release. In the case of large census operations, several such documents tailored to the needs of different sets of users (for example, users in education, health or local government) may be useful.

3.413. Many countries publish a census dictionary, which contains comprehensive definitions of terms and concepts and detailed classifications used to present census outputs. Some countries also publish geographic classifications and codes and the definitions of areas used in the census and their relationships with the administrative areas. Explanations of user-defined areas for specific census tabulations and the type of format available (printed or electronic) may be provided.

(d) Administrative report

3.414. One of the most important reports in the publication programme is the administrative report, which is a record of the entire census undertaking, including problems encountered and their solutions (see also paragraphs 3.564–3.568 on systematic recording and documentation of the census experience). The report may include the following topics:

- a brief history of the census in the country and the legal basis for conducting the census
- budget requirements, expenditure, and control
- source and allotment of funding
- census committees and their activities
- stakeholder management
- census organization and personnel structure
- staff management
- quality control procedures
- census calendar
- census cartographic work
- development and design of the questionnaires
- enumeration methodology
- census promotion, publicity, and communication;
- field organization;
- manual editing and coding;
- data-processing development and organization
- data capture
- computer editing and imputation procedure
- capital, equipment and infrastructure management
- computer hardware and software used
- census evaluation
- publication and data dissemination programme
- and archiving.

The census administrative report is very useful both for the users and for the census organization itself. The administrative report is an essential product for the planning of future censuses (see also paragraphs 3.567–3.568).

3.415. With developments in information technology, the census data files and publications have become increasingly available in electronic formats. A description of the procedure in the development of these data files may also be included in the procedural report. Consideration of a separate volume of the procedural report for the processing and dissemination phases may be considered to ensure the completion of the planning and field operations phases immediately after the census enumeration.

3. Databases

3.416. To maximize the utility and lifespan of census data, and as a complement to the standard production of tables, national statistical offices are encouraged to store the census data in various computerized database forms to better satisfy the full range of needs of internal and external data users. Census databases assist data users by providing easy access to a wide range of census data.

3.417. The establishment of such databases can enhance the dissemination of the census results as well as increase their usefulness by combining census data together with related information from other demographic enquiries in a common format. (An important special case is bringing together the data from prior censuses into a single database.) In addition, such databases can improve the coherence of the input and output processing systems.

3.418. Given the wide-ranging needs of users, there is no preferred one-size-fits-all approach to setting up a census or population database. Decisions range from the type of data to provide (microdata, sample microdata, aggregated data or a combination of these options) to whether to try and incorporate the new census results into existing database structures from previous census rounds[96] or establish new ones. In the latter case, the issue to be considered is whether the new database(s) will be exclusively in the form of a census database or constitute instead the nucleus of one or more population databases incorporating data from other sources. Consideration will also have to be given to such issues as:

- identification of the different types of users
- their information requirements
- types of information to be stored in the database
- sources of information
- maintenance and update of information
- processing of user queries
- identification of the appropriate commercial software or
- alternatively, whether it is feasible to develop such software
- and selection of the appropriate hardware capable of supporting the current database and its anticipated growth.

3.419. Since building a census or population database requires careful planning and can be time consuming, such implementation should fit within the global statistical framework of the organisation,

[96] Or in some cases existing survey or population databases used by the NSO for other collections.

and be seen as an ongoing process both complementing the data dissemination strategy and strengthening the statistical capacity of the organization.

(a) *Microdatabases*

3.420. Microdata (records of individual persons and households) collected in the census can be stored either in their raw form, or in their final edited form, or in a file that combines both raw and edited records. Digital technologies for mass storage present two issues for census managers and technicians: (a) the issue of when it will be appropriate to adopt a new technology as the standard; and (b) that of the need to convert materials stored in older media to the new standard or otherwise provide accessibility to the older materials.

3.421. Cloud-based data storage is becoming increasingly pervasive, and while it offers several advantages, including scalability, enhanced accessibility, disaster recovery, automatic updates, and seamless integration with other services, it also presents significant challenges. Data security remains a primary concern, especially with sensitive census information. Additionally, issues related to data sovereignty, potential cost overruns, data transfer bottlenecks, and the complexity of managing cloud configurations can arise. While the cloud provides a flexible and modern storage solution, its application to storing census microdata should be carefully evaluated.

3.422. With technological advances in mass storage devices and media, it is now feasible to store the full census data file as a single large rectangular file. After adding a data dictionary that describes the data format and a tabulation module, one obtains a set that could be described as a census database. The microdatabase requires a cross-tabulation program, which can be either part of the package or external. The software normally used for census tabulation still requires some prior training and may be confusing to inexperienced users. More intuitive tabulation software is available but may be either too slow in processing or too limited in its options to be fully satisfactory.

3.423. The organization of the microdatabase may take several formats, for example the software may allow for reorganizing the data in a transposed format (for example, one separate file per variable). This can substantially reduce the need for storage space and increase the speed of tabulations. However, establishing this kind of database is more complex, technically demanding and time consuming. There would be advantages in storing census microdata with standard commercial databases. This approach has the advantage that many users are already familiar with such software and so it is easier to find programmers and system analysts in the labour market. Even though the storage space required would be comparatively larger, today's market for mass storage has made available very large and fast hard disks at much cheaper prices than a few years ago and the hardware market seems to continue to follow this trend.

3.424. One of the main advantages of a microdatabase is that it permits the retrieval of data, at least in principle, at any level of detail. Since microdata could be used to obtain information on individual persons, families, households or family enterprises, privacy concerns must always be taken into consideration. In most countries, the use of census data to identify individuals is prohibited by law. Moreover, the long-term reputation of the national statistical authority may well be jeopardized if such disclosures occur. Full elaboration of principles and protocols for disseminating microdata is elaborated below in paragraphs 3.466–3.485.

3.425. As presented in this subsection, there are methods (such as sampling, introduction of random disturbances, recoding and aggregation) that can be used to make such microdata available while still protecting individuals' rights to privacy. All have in common the fact that they sacrifice some information in order to eliminate or greatly reduce the risk of disclosure. However, it is important that census organizations interested in disseminating microdata to outside users should take the appropriate precautions to protect privacy and confidentiality.

(b) ***Macrodatabases***

3.426. Aggregated census data can be stored in many formats, either as the results for one census, as a database covering more than one demographic enquiry, or in a broad database of statistical and geospatial information. Whereas microdata are saved to allow aggregations to be made that were not programmed initially, macrodata are stored to preserve earlier aggregations, offering the general public readily usable information and avoiding redundancy for those who may find that the summary data they require have already been produced.

(i) Publication equivalents

3.427. The simplest form of what could be called a database for macrodata is a straight copy of a publication on the website of the census office. A machine-readable publication-equivalent database may have the advantage of being less expensive to prepare than its hard copy counterpart. In addition, electronic or paper copies can be made quickly, copying only the part of the publication that is required. However, accessing this information requires a computer, potentially with specific software to interpret the census information.

(ii) Table-oriented databases

3.428. More advanced users, especially those engaged in spatial analysis, may require a macrodatabase that transcends the scope of the printed publication. They might like to be able to manipulate the tables and shapefiles themselves in various ways in order to obtain views or results that represent their specific requirements more precisely (see section on interactive electronic outputs). Standard formats should be used to provide (partly) customizable outputs, tables or georeferenced information to advanced users.

3.429. Census offices should be aware of the potential offered by extensible markup language (XML). XML is not, as a matter of fact, a language itself but rather a metalanguage system designed to be used on the Internet. With XML, users can define their own "tags" to structure the information within a document. XML thus offers the potential of precisely describing all elements composing a statistical table: title, subtitle, units of measure, indicators, values, the time dimension and footnotes, and in short the metadata. One extension of XML is the Geography Markup Language (GML), which is a modelling language for expressing geographical features. Other solutions, such as EDI/EDIFACT (electronic data interchange for administration, commerce and transport) and SDMX (Statistical Data and Metadata eXchange), are a set of internationally agreed standards, directories and guidelines for the electronic interchange of structured data between independent, computerized information systems.

(iii) Time series and indicator databases

3.430. Databases can also cover more than one demographic enquiry, and census results can be integrated with various other data sets, including the results of earlier censuses. In developing

databases that are aimed at serving a heterogeneous user community, the issue of a number of basic trade-offs will have to be addressed. For example, on the one hand, the number of variables should be kept as small as possible to make the database easy to use; on the other hand, it should be as comprehensive as possible to address the broadest possible requirements. A minimum data set of versatile indicators should consist of those variables that are useful for a wide range of applications and consistently available across space and time, and whose characteristics are clearly defined. In developing such a database, not only storage of the key indicators and variables themselves, but also the inclusion of some basic figures (absolute numbers or basic data) as a way of standardizing the basic statistical framework, is recommended.

3.431. With regard to time series and indicators, standardization and interoperability of data and metadata can be achieved via the Statistical Data and Metadata eXchange (SDMX)[97] international standard. In SDMX, Data Structure Definitions (DSDs) define structures that describe the data for each statistical domain, specifying the dimensions, attributes and measures, together with associated code lists. Metadata is similarly structured using the Metadata Structure Definition. Data and metadata adhering to the SDMX encoding standard may be stored in several formats, including SDMX-ML (an extension of XML) and comma separated values. Data and metadata represented in SDMX and conforming to the relevant DSD can be easily read and processed by any system or tool that supports SDMX, as well as stored in traditional databases, file systems, or other storage solutions.

(c) ***Graphing and mapping databases***

3.432. Databases equipped with graphing and mapping capabilities, especially those powered by GIS, significantly enhance their utility and relevance. GIS databases are specialized systems that store, retrieve, manage, display, and analyse geographic and spatial information. Unlike traditional databases, GIS databases are designed to handle spatial data, which are associated with geographic locations, and can visualize these data in the form of maps.

3.433. A GIS database typically consists of layers, each representing a particular type of data, such as roads, rivers, population density or land use. Each layer contains features, such as particular roads or rivers, which, in turn, have attributes, which are the specific pieces of information associated with that feature, stored in tables. For example, a road feature may have attributes such as its name, type, length, and surface material. Users can query the database, overlaying multiple layers, including geocoded census data, to see relationships between different types of data and perform spatial analysis. Users can also generate custom maps that highlight trends or illustrate spatial relationships.

3.434. Many users want data for relatively small areas concerning such matters as home ownership, educational profiles and the labour market. While the database may be for one census, some historical information can be included to allow users to observe prevailing trends over time.

3.435. Both microdata and macrodata can be at the basis of these dissemination products. However, owing to disclosure problems as well as in order to increase processing speed, some form of prior aggregation is usually applied, for example by using summary data. Such summary data could also be combined with the general-purpose graphing and mapping software. However, this would result in a reduction of the user community to those able to handle rather more complicated processing jobs. Making available a census database with tightly integrated graphing and mapping capabilities (which

[97] https://sdmx.org

usually implies a tabulation function) is an excellent way to improve the effectiveness of census information dissemination.

4. Geographic products

a. *GIS for census data dissemination*

3.436. Geographic information systems (GIS) embody hardware and software configurations designed to support the capture, management, analysis and dissemination of spatially referenced data. Applied to census activities and outputs, such systems facilitate census cartography and data capture, and by linking population data (demographic, social and socio-economic) to geographic areas, GIS provides very powerful data management functionalities in allowing users to explore, analyse, describe and communicate population census information according to their own data and information demands.

3.437. In practical terms, such systems may range from simple desktop mapping facilities to complete GIS systems are capable of solving complex planning and management problems, producing detailed georeferenced inventories and spatial statistical analysis. The ability to use space to integrate and manipulate data sets from heterogeneous sources can make its application relevant to planning and managing the census process itself. For example, GIS provides functions for the aerial interpolation of statistical data in cases where the boundaries of aerial units have changed between censuses.

3.438. Every census office should have organizational competency and staff with competences in geospatial analysis.

3.439. Statistical offices should develop GIS applications with population data and other georeferenced data from other sources for more advanced forms of spatial analysis. The role of the census office should be to supply census data at the right level and in the right format to users. Census offices provide vital information on current demographic conditions and future trends for policymakers in a range of sectors, such as health care, education, infrastructure planning, agriculture and natural resources management; and the provision of spatially referenced census databases is an essential prerequisite of the facilitation of the use of demographic data in these fields.

3.440. To achieve maximum efficiency gains, GIS applications should also be capable of generating additional geographic information beyond those used in the census, such as school and health districts, water and other biophysical catchment areas, and power and utility service units. These entities will have to be constructed from the smallest geographically identified units available in the census, such as census blocks, grid squares (see section on grid-based outputs) or enumeration areas. If, as is the case in most developing countries, enumeration areas are the smallest units, this will have important implications for the establishment of enumeration area boundaries. This requires close collaboration between national statistical organizations and national mapping and survey agencies on the one side, and school, health, water and power authorities on the other, when enumeration area boundaries are drawn or modified, to avoid potential discrepancies later on.

3.441. Apart from providing national statistical organizations with a very effective means/tools to disseminate and increase the utilization of census data, a GIS, more than any other data/information management system, provides easy and user-friendly access to census data in user-relevant formats. This allows analysts and planners to undertake policy analysis, planning and research that can more readily identify thematic and geographic priority areas and thus contribute to evidence-based and

better-informed policy and decision-making at different levels of geography. Some of the spatial statistical analysis includes clustering, Moran's spatial autocorrelation, Anselin outlier analysis, Getis-Ord hotspot analysis, ordinary least squares regression and geographically weighted regression. A GIS allows governments to effectively monitor development progress across different sectors at village, municipality and subregional levels; it raises awareness about the importance of census and other socioeconomic data; and it increases the institutional capacity of national statistical offices and social and economic planning agencies to engage in more in-depth analyses of social and economic data and deliver information products in even more user-friendly formats.

b. ***Basic maps***

3.442. Census offices should take advantage of emerging GIS technologies to make the census results more understandable and easier to use. The purpose of statistical maps is to present the results in terms of their geographic distribution and also to make it easier for the general public to understand census results than when information is presented in the form of only statistical tables. There is special interest in the current pattern of the geographic distribution and also in changes in the patterns that have occurred over time, particularly since the last census. Harmonization of the boundaries between the censuses is essential for comparability of data.

3.443. The provision of maps serves two purposes: firstly, census area identification maps locate and show the boundaries of all administrative areas and units for which data are reported in census publications; and secondly, statistical or thematic maps and graphs present the significant results of the census, thus allowing the general user to visualize the geographic distributions and patterns inherent in the data. Well-designed and attractive maps will interest the users of census reports and may raise questions that send them to the statistical tables for further details.

3.444. There are three major types of area identification maps that are commonly used in most census publications in printed or GIS shapefile formats:

a. national maps showing the boundaries of the first- and second-order geographic divisions and of the major cities or metropolitan areas.
b. maps of each first-order division showing the boundaries of the second- and third-order divisions for which statistical tables will be prepared. And
c. urban or metropolitan and degree of urbanisation maps showing small sub-area boundaries as well as general streets, roads and rivers.

c. ***Thematic maps***

3.445. A comprehensive map publication programme should be developed as part of the overall population and housing census publication programme in order that the needed resources may be provided within the budget at the initial planning stages. In addition to preparing maps for the census tables and reports, many countries have also found it useful to produce a population atlas (dynamic or static) as a census output. Collaboration with other departments and interested agencies might be sought to facilitate the production of an atlas volume. The atlas would include maps depicting population and housing characteristics, as well as other data influencing the growth, composition and distribution pattern of population and housing (see paragraphs 3.78–3.89).

3.446. As regards thematic maps, priority indicators for a population and housing census are total population and its distribution by sub-area, population density, urban and rural population by degree of urbanisation or metropolitan and non-metropolitan population, and percentage changes in the population totals since the last census. Other important indicators include age, sex, fertility, mortality, migration, educational attainment, employment, household size, type of housing, ownership, number of rooms and sanitary facilities, with a growing demand also for data on communication (telephones, television, computers and Internet access), transport (vehicles), a broad range of household amenities, and also, more recently, population-based development indicators such as household access to safe water, household waste management and multiple sources of household incomes, such as the incidence of remittances. This list of indicators is merely an illustration of the type of thematic maps individual countries might find useful to produce. Producing maps using the same set of indicators enables countries to meaningfully compare their results over time and with international or regional norms.

3.447. Maps are an invaluable aid in meaningfully comparing subnational results with national values or with other international and regional norms. Emerging technologies provide great flexibility in composing informative and visually appealing maps. Often several maps can be combined on a single page to show one indicator, for example, for the urban and the rural population by degree of urbanisation to provide a more nuanced representation in the three classes of the methodology (cities, towns and semi-dense areas, and rural areas). Also, combining maps and statistical charts is an effective means of presenting census information. Dashboards with different methods of presenting census result on maps, charts, tables and infographics can provide users with comprehensive but easy to understand message about presented phenomena.

3.448. By having associated graphing and mapping capabilities, databases will greatly increase their usefulness. Ideally users should be able to generate the graphs or maps for their own needs. Several census organizations have developed the facility to produce this kind of product, sometimes in cooperation with a commercial company. However, it is recommended that census offices develop mapping capabilities as a core competence for statistical production. Many users require small-area data concerning such matters as home ownership, educational profiles, and the labour market. While the database may be for one census only, some historical information can be included to allow users to observe prevailing trends over time. As with all time series-type data, it is important to maintain consistency in both definitions and spatial representations to ensure comparability.

3.449. Both microdata and macrodata can be at the basis of these dissemination products. However, owing to the need to maintain confidentiality, and in order to increase processing speed, some form of prior aggregation is usually applied, for example by using summary data. Such summary data could also be combined with the general-purpose graphing and mapping software. Making available a census database with codes and names matching the GIS shapefiles with tightly integrated graphing and mapping capabilities (which usually implies a tabulation function) is an excellent way to improve the effectiveness of census information dissemination.

3.450. The following list presents some suggested topics for census maps. The list is not exhaustive: most topics that appear in the questionnaire, as well as derived topics covered in Part Two, can be presented in cartographic form. In some countries, special themes such as population distribution by ethnic or language group may be appropriate. Conversely, some of the listed maps present information on the same topic in somewhat different form, so that a statistical agency may wish to select the most suitable indicator for the needs of the country.

Illustrative list of thematic census maps

Population dynamics and distribution Percentage population change during intercensal period(s) Average annual growth rate Population density (persons per square kilometre) Urban population as percentage of total population Distribution and size of major cities and towns In-migration, out-migration and net migration rates Percentage born in country and foreign born Percentage born in another division of the country	***Socioeconomic characteristics*** Percentage of children not in primary school Percentage of youths neither in education training nor employment (15-24) NEET Adult literacy rate (aged 15 and over) Mean years of schooling (aged 25 and over) Illiteracy rate of population aged 15 and over Illiterate population aged 15 and over (total number) Educational level of population aged 10 and over Labour force as percentage of total population Women's share of adult labour force Percentage of labour force by economic sector, type of occupation and status in employment Poverty mapping
Demographic characteristics Sex ratio (males per 100 females), possibly by age groups Percentage of population aged 0–14 Percentage of population aged 15–24 Percentage of population aged under 18 Percentage of population aged 15–64 Percentage of population aged 65 and over Percentage female population in childbearing ages 15–49 Total dependency ratio (population aged 0–14, and 65 and over, as percentage of population aged 15–64) Percentage of population by marital status Birth rate Total fertility rate Mean age at first marriage Death rate Infant mortality rate Under-five mortality rate Life expectancy at birth Percentage of persons with disabilities	***Households and housing*** Average number of persons per household Percentage of households headed by women Average number of dwelling rooms per household Percentage of households by Tenure status (owned, rented, and so forth) Percentage of dwellings by type of construction material Percentage of population with access to adequate shelter Percentage of population with access to safe water Percentage of population with access to electricity Percentage of population with access to sanitation Percentage of population with access to health services

3.451. Where appropriate, the indicators can be presented disaggregated by gender as well as by urban or rural area (for example, where the rural population is greater than about 25 per cent of the total population). If information about an indicator is also available from a previous census, it is often very informative to produce maps showing change over time or to present maps for both time periods.

3.452. Also, where appropriate, countries are encouraged to perform spatial statistical analysis by producing maps showing spatial clustering and outlier analysis of the variables of interest, such as electricity and water.

3.453. The development of locality (village, town, city, community, small area) population size maps by region is of particular value. These maps combine two types of information: locality population statistics and locality locations in each region or subnational area. More information can be presented on, for example, the locality location within the district and the region, habitable and non-habitable areas, densely populated localities, areas with no localities, and the proximity of localities. Locality population size maps can also be used as base maps for additional information on locality services and activities,

and on location and distribution of localities without specific services, such as primary schools, dispensaries, piped water, and so forth.

d. ***Grid based census outputs***

3.454. Grid-based[98] census outputs are a type of data presentation where census data is aggregated and organized within a grid system, usually a rectangular grid of equal-sized cells. This system replaces the traditional method of presenting census data based on administrative boundaries, offering several advantages and facilitating new applications.

3.455. Grid-based census outputs are an increasingly valuable tool for researchers, policymakers, and the public alike. They provide a powerful way to analyse, visualize, and understand census data, offering significant benefits for decision-making, planning, and resource allocation across various sectors. As technology advances and data integration techniques improve, census takers can expect grid-based outputs to play an even greater role in future data analysis and research.

3.456. The benefits of grid-based census outputs are:

- **Spatial Analysis**: Grids provide a consistent and easily interpretable spatial framework for analysing census data. This enables researchers and policymakers to identify spatial patterns and trends, visualize data variations across space, and conduct advanced geospatial analyses.
- **Data Integration**: Grids facilitate the integration of census data with other spatial datasets, such as environmental data, infrastructure maps, and land use data. This allows for a more comprehensive understanding of the relationships between population characteristics and other factors in the environment.
- **Data Accessibility and Transparency**: Grid-based outputs can be easily visualized and shared through maps and other interactive tools, making census data more accessible and transparent to the public.
- **Flexibility**: Unlike traditional area-based geographies, the grid system does away with the need for predefined local and national boundaries. This makes it a highly flexible tool for analysing data at any level of detail, from meter-scale grids to grids spanning tens of kilometers.
- **Standardization**: Grid-based outputs often use standardized grid systems, such as the Universal Transverse Mercator (UTM) grid, allowing for data comparison across different regions and time periods and between countries with vastly different administrative geographies.

e. ***Innovative geospatial products***

3.457. Other innovative geographic products for census data which utilize the emerging and still relatively new technologies of artificial intelligence (AI) and machine learning include:

(i) Smart maps and geoportals

3.458. Smarter Maps is a term used to describe maps that are more interactive, intuitive, and informative than traditional maps. Smarter maps may use artificial intelligence (AI) and machine learning to improve the accuracy and usability of traditional maps. They are designed to help users

[98] The grid square system is a method of dividing a geographic area into smaller, regularly shaped units for statistical purposes. These units are typically squares defined by lines of latitude and longitude, forming a grid that covers the entire area of interest.

visualize data in a more meaningful way and make better decisions based on that data. They can be used to analyze large datasets, identify trends and patterns, and create maps that are more informative and engaging.

(ii) GeoAI and GeoBlockchain

3.459. GeoAI is the application of AI technology to enhance the intelligence of GIS software which enables the analysis of e massive amounts of geospatial data (such as satellite imagery), whereas GeoBlockchain is a combination of blockchain technology and geographic information systems (GIS) that can be used to support the analysis of spatial-temporal trends. When combined, GeoAI and GeoBlockchain can be used to create a powerful platform for analysing geospatial data and generating insights that can drive better decision-making.

5. Interactive electronic outputs

3.460. It is of paramount importance that census data and any information produced to support them are widely disseminated and communicated, and that national statistical and census offices involved in this process have a clear customer, client and stakeholder focus. That means that national statistical and census offices should place more emphasis on providing a service and creating partnerships than on merely providing products and should be guided by user-relevance and user-friendliness in all their operations, rather than by simply providing the tables, graphs and reports that they have traditionally produced.

3.461. Given its importance and widespread use, the web has emerged as the primary means of providing general access to census statistics. Many national statistical and census offices have utilized the Internet as the principal channel for data communication, positioning their websites into comprehensive census data repositories, enabling users to have access to all published data online. When developing new census products, and when reviewing existing products, national statistical and census offices should consider all ways and means of making census statistics accessible, giving high priority to dissemination on the web. The advantages of online dissemination are primarily in terms of speed, flexibility and cost, as well as in providing accessibility to census results to a wide range of data users and allowing the delivery of data to be tailored to the level of sophistication of the user.

3.462. Making a census database available online along with integrated searching, tabulating, graphing, mapping and analytic capabilities is an important way to improve the effectiveness of census data dissemination. Most national statistical and census offices provide user access to electronic databases and data files through their websites, satisfying the full range of needs of internal and external data users. This is a valuable service that allows users to access and display census data instantaneously and interactively. The establishment of such databases can enhance the dissemination of census results as well as increase their usefulness by allowing user interaction with census data. User interaction is a key concept whereby users are enabled and empowered to access and explore census data themselves, and build their own customized tables or spatially configure data outputs according to their own requirements.

3.463. Interactive web-based data tools provide a user-friendly entry point to the entire range of census outputs disseminated by national statistical and census offices. Basic design considerations of web-based interactive tools should factor issues such as identification of the different types of users, their information requirements and the types of information to be stored in the database. Content

should be organized so that it can be easily understood and found, with an overview or tutorial to provide orienting information to users about the data products that can be obtained using the tool(s). Context should always be provided to all outputs through metadata, links to related information, and cross-referencing to glossaries, publications and other background material.

3.464. In practical terms, interactive web-based data tools should enable users to access census data themselves, and build their own customized tables or spatially configure data outputs according to varying spatial requirements. The tools should allow users to visualize and explore the data in various standard visualizations such as column charts, line graphs, maps and scatterplots. The table-building functionality should also have the ability to sort and order tabular results, and more easily select census dates and indicators. Tools should also be provided for downloading, conducting analysis or retrieval for use in other software. Design considerations to improve the interactivity of data interfaces should include the provision of user support. It is highly recommended to help users to anticipate, interpret and evaluate results. Support to users should include demonstrations and tutorials intended to describe how to perform the various functions related to the interactive web-based tools.

3.465. Especially for highly flexible interactive online tools, additional disclosure risks entailed, and a corresponding protection strategy must be a key part of the design of such tools from the outset (See section on Confidentiality and privacy). In addition to web-based tools, some interactive electronic products may also be published through other media.

6. Microdata dissemination[99]

3.466. The provision of census microdata, including subsets of the microdata based on a random sample, allows for further analysis and research, especially in the context of analysis of SDGs where further data dissemination by various characteristics is required. Also, longitudinal microdata files linking a sample of persons on census files over time have been produced and are widely used among the research communities in several countries.

(a) Definition of microdata

3.467. In general, when statistical agencies or other data producers conduct censuses, surveys or collect administrative data, they gather information from each unit of observation. Such a unit can, for example, be a household, person, firm or enterprise, agricultural holding, school, or health facility. In the context, microdata are the electronic data files containing the information about each unit of observation. Microdata are thus opposed to macrodata or aggregated data, which provide a summarized version of this information in the form of counts, means, ratios, frequencies or other summary statistics.

3.468. Typically, microdata are organized in data files in which each line (or record) contains information about one unit of observation. This information is stored in variables. Variables can be of different types (for example, numerical or alphanumerical, discrete or continuous). They can be obtained directly from the respondent via a questionnaire or by observation or measurement (for example, by GPS positioning), or imputed or calculated.

[99] The elaboration on the dissemination of microdata is largely based on the *Dissemination of Microdata Files: Principles, Procedures and Practices*, Olivier Dupriez and Ernie Boyko, IHSN Working Paper No. 005, August 2010.

3.469. In the context of the population and housing census, microdata refer to electronic files consisting of individual records on persons, households and housing units. More specifically, microdata would typically be organized in multiple files: one with records on households, another with records on individuals, and yet another with records on housing units.

(b) Core principles for disseminating census microdata[100]

3.470. The United Nations Fundamental Principles of Official Statistics[101] provide unambiguous guidance in administering official statistics at national and international levels. A particular emphasis of these principles is on confidentiality of information collected for statistical purposes. In the context, the sixth principle, governing international statistical activities, states: "Individual data collected by statistical agencies for statistical compilation, whether or not they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes."[102]

3.471. Any principles for microdata access must be consistent with this recommended principle and the principles contained in the legislation pertaining to the national statistical authority. The following principles should be considered for managing the confidentiality of microdata:

Principle 1: Appropriate use of microdata

3.472. It is appropriate for microdata collected for official statistical purposes to be used for statistical analysis to support research as long as confidentiality is protected.

3.473. Making available microdata, or random samples microdata, for research is not in contradiction with the sixth United Nations Fundamental Principle as long as it is not possible for such data to refer to an identifiable individual. Principle 1 does not constitute an obligation to provide microdata. The national statistical office should be the one to decide whether to provide microdata or not. There may be other concerns (for example, quality) that make it inappropriate to provide access to microdata. Or there may be specific persons or institutions to whom it would be inappropriate to provide microdata.

Principle 2: Microdata should only be made available for statistical and research purposes

3.474. For principle 2, a distinction has to be made between statistical or research uses and administrative uses. In the case of statistical or research use, the aim is to derive statistics that refer to a group (be it of persons or legal entities). In the case of administrative use, the aim is to derive information about a particular person or legal entity to make a decision that may bring benefit or harm to the individual. For example, some requests for data may be legal (a court order) but inconsistent with this principle. It is in the interest of public confidence in the official statistical system that these requests are refused. If the use of the microdata is incompatible with statistical or research purposes, then

---

[100] The elaboration of core principles for dissemination of microdata is quoted from *Managing Statistical Confidentiality and Microdata Access: Principles and Guidelines of Good Practice*, United Nations Economic Commission for Europe, Conference of European Statistician, United Nations publication, Sales No. E.07.II.E.7, United Nations, 2007.

[101] Presented at the United Nations Statistics Division website at: http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx.

[102] Ibid.

microdata access should not be provided. Ethics committees or a similar arrangement may assist in situations where there is uncertainty whether to provide access or not.

3.475. Researchers may wish to access microdata for research purposes but to support this research they may need to compile statistical aggregations of various forms, compile statistical distributions, fit statistical models or analyse statistical differences between subpopulations. These uses would be consistent with statistical purposes. To the extent that this is how the microdata are being used, it could also be said to support research purposes.

Principle 3: Provision of microdata should be consistent with legal and other necessary arrangements that ensure that confidentiality of the released microdata is protected

3.476. With respect to principle 3, legal arrangements to protect confidentiality should be in place before any microdata are released. However, the legal arrangements have to be complemented with administrative and technical measures to regulate the access to microdata and to ensure that individual data cannot be disclosed. The existence and visibility of such arrangements (whether in law or supplementary regulations, ordinances, and so forth) are necessary to increase public confidence that microdata will be used appropriately. Legal arrangements are clearly preferable, but in some countries, this may not be possible and some other form of administrative arrangements should be put in place. The legal (or other arrangements) should also be cleared with the privacy authorities of countries where they exist before they are established by law. If such authorities do not exist, there may be non-governmental organizations that have a "watchdog" role on privacy matters. It would be sensible to get their support for any legal or other arrangements, or at least to address any serious concerns they might have. In some countries, authorizing legislation does not exist. At a minimum, release of microdata should be supported by some form of authority. However, an authorizing legislation is a preferable approach.

Principle 4: The procedures for researcher access to microdata, as well as the uses and users of microdata, should be transparent and publicly available

3.477. Principle 4 is important to increase public confidence that microdata is being used appropriately and to show that decisions about microdata release are taken on an objective basis. It is up to the national statistical authority to decide whether, how and to whom microdata can be released. But their decisions should be transparent. The website of the national statistical authority is an effective way of ensuring compliance and also for providing information on how to access research reports based on released microdata.

(c) Microdata anonymization[103]

3.478. When disseminating census microdata files to the public, researchers or other agencies, the national statistical authority faces a conflicting mission. On the one hand, it aims to release microdata files supporting a wide range of statistical analyses; on the other, it must safeguard the confidentiality of

[103] The elaboration of procedures for anonymizing microdata draws extensively from *Handbook on Statistical Disclosure Control, Version 1.0*, Anco Hundepool, Josep Dominho-Ferrer, Luisa Franconi, Sarah Giessing, Rainer Lenz, Jane Longhurst, Eric Schulte Nordholt, Giovanni Seri, Peter-Paul de Wolf, Centre of Excellence for Statistical Disclosure Control, Eurostat project, December 2006.

respondents' identities. Processes aimed at the latter are referred to collectively as statistical disclosure control or anonymization.

3.479. A disclosure occurs when a person or organization recognizes or learns via released data something they did not know about another identifiable person. There are two types of disclosure risk: identity disclosure and attribute disclosure.[104] The former occurs when a respondent's identity is directly associated with a disseminated data record. This can occur easily when the data record includes variables unambiguously identifying the respondent – for instance, the respondent's name, address, passport or identification number, or telephone number. It is essential that such identifying variables be removed from any microdata files before dissemination. Attribute disclosure occurs when attribute values (or estimates thereof) in the disseminated data are associated with a particular respondent.

3.480. A combination of variables in a microdata record that can be applied to reidentify a respondent is referred to as a "key". Reidentification can occur (a) when a certain key value is rare or unique in the population with respect to a specific respondent; and (b) when this key can be used to match a microdata file to other data files that might contain direct or other identifiers such as voter lists, land registers or school records (or even publicly accessible Internet search engines).

3.481. The essential component of dissemination of census microdata files is avoiding both identity and attribute disclosures. In that respect, there is a need to strictly apply statistical disclosure control or anonymization techniques for census microdata files. The first key step in anonymizing a microdata file is to remove (or encrypt) all direct identifiers– variables that unambiguously identify the respondent. Thereafter, a microdata file can be anonymized further by applying statistical disclosure control techniques.[105]

(d) Protocols for dissemination of census microdata

3.482. Disseminating census microdata may be an unprecedented activity for the national statistical authority. In that context, there is a need to develop particular protocols that would comply with the essential principles for disseminating microdata, as described above, and would also spell out the necessary requirements and components of such dissemination. Hence, such a protocol would need to take into account the following components and requirements:[106]

Openness. Access on equal terms for the research community at the lowest possible cost, preferably at no more than the marginal cost of dissemination.

Flexibility. Taking into account the rapid and often unpredictable changes in information technologies, the characteristics of each research field and the diversity of research systems, legal systems and cultures of each member country.

---

[104] *Measures of Disclosure Risk and Harm*, Diane Lambert, Journal of Official Statistics, Volume 9, Number 2, Statistics Sweden, 1993.

[105] For full elaboration of these techniques, refer to Chapter 7 of the *Dissemination of Microdata Files: Principles, Procedures and Practices*, Olivier Dupriez and Ernie Boyko, IHSN Working Paper No. 005, August 2010.

[106] As presented in OECD Principles and Guidelines for Access to Research Data from Public Funding, Organisation for Economic Co-operation and Development, 2007.

Transparency. Information on research data and data-producing organizations, documentation on the data and specifications of conditions attached to the use of these data should be internationally available in a transparent way, ideally through the Internet.

Legal conformity. Data access arrangements should respect the legal rights and legitimate interests of all stakeholders.

Protection of intellectual property. Data access arrangements should consider the applicability of copyright or of other intellectual property laws that may be relevant to publicly funded research databases.

Formal responsibility. Access arrangements should promote explicit, formal institutional practices, such as the development of rules and regulations, regarding the responsibilities of the various parties involved in data-related activities. These practices should pertain to authorship, producer credits, ownership, dissemination, usage restrictions, financial arrangements, ethical rules, licensing terms, liability and sustainable archiving.

Professionalism. Institutional arrangements for the management of research data should be based on the relevant professional standards and values embodied in the codes of conduct of the scientific communities involved.

Interoperability. Technological and semantic interoperability is a key consideration in enabling and promoting international and interdisciplinary access to and use of research data.

Quality. The value and utility of research data depend, to a large extent, on the quality of the data. Data managers, and data collection organizations, should pay particular attention to ensuring compliance with explicit quality standards.

Security. Specific attention should be devoted to supporting the use of techniques and instruments to guarantee the integrity and security of data.

Efficiency. One of the central goals of promoting data access and sharing is to improve the overall efficiency of publicly funded data collection to avoid the expensive and unnecessary duplication of data collection efforts.

Accountability. The performance of data access arrangements should be subject to periodic evaluation by user groups, responsible institutions and … funding agencies.

Sustainability. Due consideration should be given to the sustainability of access to publicly funded research data as a key element of the research infrastructure. This means taking administrative responsibility for the measures to guarantee permanent access to data that have been determined to require long-term retention.

(e) Dissemination of population and housing census microdata in practice

3.483. It can be expected that the 2030 round of population and housing censuses, taking into account the contemporary development of processing technology and power, will put considerable pressure on national statistical authorities in respect of dissemination of population and housing census microdata.

In the context of implementing the principles and protocols described in the preceding paragraphs, national statistical authorities would need to ensure such dissemination in at least two different settings.

3.484. For the purpose of public dissemination, either online or on electronic media, only a representative sample of the individual records should be made available after ensuring the confidentiality or non-disclosure of individual information as elaborated above. The size of the sample would depend on the capacity and resources of the national statistical or census office. When generating public-use or research-use microdata samples from population census data, not only the sample size but also the sample design should be optimized, recognizing that while simple random sampling methods may often suffice, consulting a sampling expert is essential to ensure the appropriateness and effectiveness of the chosen approach.

3.485. As for the access to the complete population and housing census master file, it should be made available to users using the model of the data enclave[107]. This is a facility equipped with computers not linked to the Internet or an external network and from which no information can be downloaded via USB ports or other external drives. Users interested in accessing a data enclave will not necessarily have access to the full census data set – only to the particular data subset they require. They will be asked to complete an application form demonstrating a legitimate need to access these data to fulfil a stated statistical or research purpose and be briefed on the legal responsibility and repercussions related to maintaining the confidentiality of individual information. The outputs generated need to be scrutinized by way of a full disclosure review before release and they can contain only aggregates.

7. Customized products

3.486. The increasing activity in the field of economic and social planning and the attention of such planning to subnational areas are placing new demands on statistical information in general and on population and housing censuses in particular. There is an increasing need for tabulations and mapping not only by major and minor civil divisions and by other units of analysis such as metropolitan areas but even, beyond these, by small local areas.

3.487. Therefore, it is useful to establish an “on request” service for users who require aggregates not available through other means. This will be especially relevant in situations where outsiders cannot obtain census microdatabases. In essence, the service would require that users provide the census office with the details of the tables or other aggregates requested so that the census office could fulfil the request, normally against payment of a cost-recovery customer service fee. Offering and promoting this service, especially online, would place the statistical service in a more favourable proactive position, rather than a static one, and could be a strong catalyst for closer cooperation with census product users.

3.488. The cost of such special purpose tabulations, which require computer programming, could be high, especially for academic institutions and other users who do not have access to a large budget. Some statistical organizations allow the users to do the necessary work using a user-friendly kind of

[107] A data enclave is a secure network through which confidential data, such as identifiable information from census data, can be stored and disseminated. In a virtual data enclave, a researcher can access the data from their own computer but cannot download or remove it from the remote server. Higher security data can be accessed through a physical data enclave where a researcher must access the data from a monitored room where the data is stored on non-network computers.

software. A clearly written manual is required to guide the users in using the software, including the contents of the census data dictionary and other relevant information. The resulting tables are checked for any possible breach of confidentiality, in particular table cells with very small values.

3.489. Many census organizations provide services for special requests for census products, such as thematic databases, tables, and graphic and mapping outputs that can be designed for small, medium and large businesses, communities or special interest groups. These services are normally provided to meet the increasing demand of data users for a wide range of applications, such as monitoring trends, analysing unmet needs, identifying market potential, segmenting markets, identifying service areas and priority zones, determining optimum site locations, and designing and advertising new products and services. Each category of products should also be made available on various media (namely, paper, disk, online) for dissemination according to user requirements.

3.490. Once the databases are created and have served the policy needs, they can serve other data users if they have market value. Since the national statistical organization is normally the only source of many geographic databases related to census data applications, market demand for these products is increasing, particularly in the geographic and population-related areas. In such cases, census products could be governed by a licence. The licence permits the users to use the product without transferring ownership, since the ownership remains with the government agency. Either of two different licensing arrangements may be applied. The first is offered to organizations that use the data for their own needs and the other is offered to organizations that redistribute data or provide analytical services using census data to other persons or organizations for a fee (see section on dissemination policy).

3.491. Customized services of data on computer media are differentiated in terms of the forms of the data. Census products may be distributed in their original form, with or without other related information, or they can be distributed after making certain value-added modifications to meet the need of the users. Examples of such value-added activities include converting the data into another format (for use by other software packages), making the data more useful by creating subsets of the original data sets, merging the data from other sources and bundling with software. It is encouraged to disseminate census content in formats that comply with the requirements assessed by the Open Knowledge Foundation.[108] According to these requirements, the openness of the contents is assessed within the Open Knowledge conceptual framework and concerns the possibility to reuse, revise, remix and redistribute data. In cases where copyright laws protect census data ownership, some royalty fees and data usage fees may be charged to the distributors to ensure a minimum return. However, if prices are too high such charges can also be a barrier to the use of the census data.

3.492. Some countries may assist their users by merging selected variables with the GIS shapefiles as a customized product. This has proven to be beneficial for schoolchildren.

8. General interest products and special audience products

3.493. Information generated by a census is, by definition, of use to a wide range of users with a variety of expertise. With the increase of demand for census products, efforts must be made by census offices to produce a variety of products for various stakeholder groups, including special interest groups. In order to address their various needs, census offices must segment stakeholders and potential

[108] https://okfn.org/opendata/

audiences into groups to better know, understand and respond to their needs. This will form the basis of the various census products. In response to the needs of special audiences, census offices may prepare products for key variables such as: policy summary reports; thematic and analytical reports; key findings reports; fact sheets; posters, brochures and flyers; basic reports; detailed tables and spreadsheets; articles; and video and social media products. Special audience analytical or thematic reports incorporate a high level of very sound analysis undertaken by staff who have a solid foundation in analytical techniques as well as the topic being analysed. In some cases, countries may undertake the analysis in collaboration with academic institutions or other specialists. Criteria used in establishing the topics chosen will have to be set by the country concerned, and may include: particularly interesting facts shown by the census data (perhaps confirming or rebutting conventional theories; confronting census data with material from other sources; or responding to issues raised by the public during user consultations of the collection).

3.494. It should be noted that the following products can only be effective in encouraging the use of census information if they are prepared in a timely and professional manner. This will require specialist skills from people familiar with communicating to the target audiences. These resources are expensive and countries are required to adequately plan and budget for these products and campaigns.

(a) Posters

3.495. One of the most common ways to disseminate census information consists of publishing posters highlighting key facts such as: How many are we? Where do we live? and summarizing a profile for the major civil divisions of a country. Posters might also be prepared addressing issues relevant to special population segments: teenagers, adults, indigenous populations, seniors and women's groups.

3.496. Since the objective of a poster is to catch the eye at a distance, relatively few facts should be presented in a way so that the key message is immediately visible. Posters can be greatly enhanced by the addition of well-designed graphs, infographics and maps to increase the readability and comprehensibility of the key message. Posters and banners are short-term communication products and should be used to communicate key findings.

(b) Brochures and flyers

3.497. Professionally designed brochures and flyers are another way to disseminate basic census data. These brochures should be written in a very easy and comprehensible language indicating the demographic profile of the country illustrated with suitable graphics and explanatory material. In some countries these brochures might be addressed to specific issues on population. They are particularly suitable for preparation as promotional materials for people attending events and exhibitions, such as the launch of more traditional materials, or for inclusion on display racks in libraries of government offices.

(c) Videos and sketches

3.498. In order to create a better understanding among certain interest groups, the use of other communication media are recommended, including videos, and sketches. In order to promote the story behind the numbers and increasing the use of census data, graphics such as charts or maps could be included on videotape, CD or DVD, or memory sticks. These might indicate how census data can assist

policymakers, planners and people in general to understand their societies and how census data can assist in identifying the main problems and assist with evaluation of solutions.

(d) Instructional materials

3.499. Instructional materials in an easy-to-understand form can be prepared for the general public, indicating the advantages and limitations of census data. Such material can often form the basis of information campaigns as part of the advocacy material for the next census.

3.500. A particular implementation of instructional materials can be the preparation of a kit for use in schools. Not only will this provide high-quality information for the students but, by including exposure to the use of statistical materials in the school process, it will encourage the use of evidence-based analysis throughout society. It should be noted that professional assistance should be sought in ensuring that these materials follow sound educational practices and can be accommodated within the appropriate curriculum.

(e) Social media

3.501. Internet-based social media have become an indispensable tool in disseminating information and marketing statistical products. Various social media platforms have been successfully used by countries as part of the dissemination of their census results. Interacting with followers and users on these platforms provides the census organization with an opportunity to disseminate information, build relationships with established and new users, and engage the public on a regular basis.

3.502. Social media platforms such as Facebook, Twitter and online video sites can be used to post all census-related advertisements and other videos.

(f) Education material for schools

3.503. Learning about the census and about how to effectively use census and other statistical data is a key part of our education. It helps students to understand our society and their role as responsible citizens in participating in that society. It helps us them learn some very important skills. For instance learning how data is collected, analysed and reported in a meaningful way not only helps maths, presentational and descriptive skills but it helps to understand the data that is presented to them. It gives them the skills to ask questions to find the facts to support their understanding of what is happening in the world around them in a meaningful way. Students will use these skills in many of their future studies whether these are in the areas of maths, business, economics, history, geography or in conducting scientific projects.

3.504. In addition to equipping students of all ages with these skills, including census in the school curricula can provide an extremely good promotional opportunity to encourage participation in the census. This is particularly useful in promoting the census to hard-to-reach groups. It is advisable to get practicing school teachers involved in developing appropriate census content that aligns with the school curricula. This should include extremely accessible website content featuring videos, infographics and practical exercises[109].

[109] https://www.cso.ie/en/census/censusforschools/censusinschools2022/.

D. Census data utilization

1. General uses of population and housing censuses

3.505. Population censuses are traditionally used for public and private sector policymaking, planning, administrative and research purposes at national and subnational level. One of the most basic of the administrative uses of census data is in the demarcation of constituencies and the allocation of representation on governing bodies, as well as allocation of resources within a country. Certain aspects of the legal or administrative status of territorial divisions may also depend on the size of their populations. Housing censuses are used to develop benchmark housing statistics and to formulate housing policy and programmes, and in the private sector to assist in site selection for industrial, retail and service facilities, as well as for the commercial development of residential housing.

3.506. As has been noted several times in these Principles and Recommendations, information on the size, distribution and characteristics of a country's population is essential to describing and assessing its economic, social and demographic circumstances and to developing sound policies and programmes aimed at fostering the welfare of a country and its population. The population and housing census, by providing comparable basic statistics for a country as a whole and for each administrative unit and locality therein, can make an important contribution to the overall planning process and the management of national development. The availability of information at the lowest levels of administrative unit is valuable for the management and evaluation of such programmes as education and literacy, employment and human resources, reproductive health and family planning, housing and environment, maternal and child health, disability, rural development, transportation and highway planning, urbanization and welfare. Population and housing censuses are also unique sources of data for producing relevant social indicators to monitor the impact of these government policies and programmes (see paragraphs 3.526–3.528).

(a) Uses of population censuses

3.507. The uses of population census results and the associated tabulations described in this volume are listed according to the topics presented in paragraph 5.21.

3.508. The total population and its distribution among major and minor territorial divisions and localities are frequently a fundamental requirement of the census because these results are used for determining the apportionment of representation in legislative bodies, for administrative purposes and for planning the location of economic and social facilities. Internal migration, one of the major sources of population change, frequently affects the trends in population distribution. Data on internal and international migration, together with fertility and mortality, are needed to prepare population estimates for planning purposes and for determining policies on migration and assessing their effectiveness. Data on forcibly displaced status, both for refugee and internally displaced populations allows to measure integration and policy responses in this area[110].

[110] IRRS see paragraph 4, more info https://egrisstats.org/recommendations/international-recommendations-on-refugee-statistics-irrs/. Also, IRIS see paragraph 3, more info on https://egrisstats.org/recommendations/international-recommendations-on-idp-statistics-iris/

3.509. The household, a basic socioeconomic unit in all countries, is often central to the study of social and economic development. The number, size and structure of households and changes in the rate of household formation are useful for planning and for developing special policies formulated for selected groups of the population, such as children, youth, older persons, and persons with disabilities. Therefore, the distribution of individuals within households is used to determine the living arrangements of families, the patterns of family structure observed, the time when new families are formed and changes in family structure due to death, divorce or separation, migration, or the departure of children to form their own households. The relationship among household members can be used to determine family structure and the existence of households composed, partially or completely, of unrelated persons.

3.510. Traditionally defined demographic, social and ethnocultural characteristics collected from the population census include sex, age, marital status, religion, language, nationality, and ethnicity and indigenous identification. Sex and age are fundamental to the majority of the characteristics collected in the census. The census provides more data than any other single source on differences between women and men.

3.511. Depending on national circumstances, cultural diversity may be indicated by language spoken in the home or community, religion and national or ethnic group. For countries that are not homogeneous in terms of one or more of these variables, linguistic, religious and national or ethnic groups provide the basic information for a quantitative assessment of the relative size and age-sex distribution of this diversity.

3.512. Although census data on fertility and mortality cannot serve as a substitute for reliable birth and death statistics from civil registrations, they are particularly valuable for countries where birth or death registration is lacking, or incomplete and vital statistics are therefore unavailable. Even in countries with complete registration of these events, the population census is useful as a supplement to satisfactory registration data because the fertility questions provide data for calculating lifetime fertility of the female population or cohort fertility.

3.513. Education has historically been one of the key factors determining the quality of life, and interest in education continues today in most countries of the world, with emphasis on improving access to education and the quality of education, as well as broadening the scope of basic education.[111] Education is also considered a major tool in closing the gap between women and men in respect of socio-economic opportunities. Benchmark data obtained from national population censuses will therefore be of considerable importance towards fulfilling this objective. Census data reveal the disparity in educational opportunities between the sexes, age cohorts or generations, and urban–rural populations, and provide important indications of the capacity of the nation for economic and social development. They furnish material for the comparison of the present educational attainment of the adult population with the present and anticipated requirements of educated human resources for various types of economic activities. Such a comparison may serve as a guide both for national policy in terms of the development of the educational system, and for the planning of the economic development programmes that it will be feasible to undertake in view of human resource requirements.

[111] Education for All Summit of Nine High-Population Countries, New Delhi, 12–16 December 1993: Final Report (Paris, UNESCO, 1994).

3.514. Census information on the economic characteristics of the population focuses on enumerating the labour force so as to provide benchmark data for current studies of employment and labour underutilization, in particular unemployment and the potential labour force. It provides information on the growth, composition and distribution of the labour force for use in policy formulation and the appraisal of human resource utilization. Economic data from censuses can also provide some input into statistics on the distribution of income, consumption and accumulation of households, and participation in agriculture and non-agricultural activities. Furthermore, the data on the labour force may give an approximate indication of the number of workers who are responsible for the support of dependants.

3.515. Statistics obtained from different sources (for example, labour force surveys, agriculture surveys, establishment surveys and administrative records) rely on the census for sampling frames, and the use of common concepts in the different sources helps in securing comparability when multiple sources for changing patterns of economic activity are being relied upon.

3.516. As interest in the movement of people across national boundaries – in other words, international migration – has grown steadily among countries, census items and tabulations relative to international migration have grown in importance. Such tabulations are designed to assess the impact of migration on receiving countries, to understand patterns of diversity and develop programmes for the adaptation of migrants, and to serve as a source of information on emigration from sending countries. It is worth noting, the census is also an important data source for measuring internal migration within a country. Additionally, census data on refugees and internally displaced persons is essential for understanding the impact of forced displacement and informing policies to support these populations.

3.517. The census is also an important source of data on persons with disabilities. Census data help to monitor the social and living conditions of persons with disabilities in terms of school attendance, educational attainment, employment, marital status and living arrangements. The data also provide a basis for developing policies to meet the needs of persons with disabilities and for evaluating the effectiveness of these policies.

3.518. The census is also an important source of information on household poverty and living conditions. Population censuses also provided historical data for historians, and researchers studying social, economic, and demographic changes over time.

(b) Uses of housing censuses

3.519. The primary uses of information from housing censuses include: development of a basis for planning housing and human settlement programmes and policies, public and private sector studies of urban and other non-agricultural land use, evaluation of the adequacy of housing stock and assessment of the need and market for new housing, measuring access to amenities and facilities, and studies of the living conditions of the homeless and those living in temporary or substandard housing. Information collected on the number of sets, type and characteristics of living quarters and their occupants is crucial from the point of view of monitoring housing conditions and needs of the population. Combined with the information collected by regular annual statistical programmes on housing construction, data from the housing census provide a basis for identifying national, regional and local housing patterns, which are needed for the development of a rational housing market aimed at stimulating various types of housing construction. The type and quality of shelter in which people are housed, that is to say, the space, degree of crowding, facilities, surroundings and available transport, affect their economic activity,

health, social intercourse and general outlook. The supply, characteristics and costs of housing are therefore subjects for which the housing census is an important source of information.

2. Analysis of the results

3.520. In order to ensure the fullest possible utilization of census results by national and local governmental authorities, by academic researchers and a diversity of other users, it is advisable to draw up a comprehensive and coordinated programme of analytical studies, phased over a period of several years. This will help allocate effort and resources in such a way as to ensure that important policy needs are adequately met, undue duplication of research effort is avoided, and priorities are observed as far as possible. In these studies, the data of the current census should be examined not only by themselves but also as complemented by relevant data from other sources and from earlier censuses, in order to obtain a broader context, improve the estimates and establish trends.

3.521. The analytical studies to be included in such a programme will vary according to the needs and circumstances of the country. The programme may include descriptive summaries of results, policy-oriented analyses of census results and detailed analytical studies of one or more aspects of the demographic and social situation of the country. Some of these studies may be undertaken by the census organization itself, but others, particularly the more time-consuming studies, can most effectively be carried out in cooperation with specialists in different subjects having experience in in-depth analytical studies from universities or other research centres. In any case, it is desirable to invite specialists from other governmental offices and experts outside government to take part in drawing up this programme of studies, and it is natural that they would play an important part in the execution of various parts of the analytical programme.

3.522. One important aspect to be considered in establishing a programme of analysis is the possible use of census results in achieving the goals and objectives of population, human settlements or similar policies and strategies at the national and local level, and in applying available resources effectively towards the improvement of conditions in these fields. For this purpose, it will be necessary to analyse the population and housing census results within the framework provided by other available information so as to achieve an integrated approach to the solutions of population, human settlements and similar problems.

3.523. In addition to the studies that are part of the overall census programme, additional analyses carried out on their own initiative by research organizations, universities or other experts should be encouraged.

3. Intercensal and post-census population estimates

3.524. Intercensal population estimates are estimates made for the years between two censuses and are produced once the census has been completed. They replace the postcensal estimates produced prior to the completion of the census.

3.525. Intercensal population estimates can be calculated using a combination of data sources and statistical methods and actual methods used vary by country and organization[112]. Inputs for intercensal and post-census estimates can include the base population from the most recent census, estimates of the number of births and deaths, both international and domestic migration, as well as adjustment for specific sub-groups such as specific age groups[113]. For analytical and planning purposes, results of the census evaluation should be taken into account in preparing the intercensal and post-censual population estimates in order to correct deficiencies in the data collection phase, and in the base population in particular.[114] It is recommended that national statistical institutions in charge of producing intercensal and/or post-census population estimates make available information on the data, methods and adjustment (if any) used to update these population estimates. Such information is important if national official population estimates are to be considered and used in processes at the global level.

4. Cross-cutting and emerging social issues

3.526. Reflecting the concerns and priorities among countries around the world, the United Nations convened a series of global conferences: on children, education, environment and climate change, human rights, population, sustainable development, women and human settlements. Each of these conferences recognized the importance of adequate information for formulating policy and monitoring progress in the achievement of conference goals and called on countries and international organizations to develop and improve the requisite statistics and indicators. These recommendations are reflected for example in the Vienna Declaration and Programme of Action of the World Conference on Human Rights;[115] the Programme of Action of the International Conference on Population and Development;[116] the Copenhagen Declaration on Social Development and the Programme of Action of the World Summit for Social Development;[117] the Platform for Action[118] adopted by the Fourth World Conference on Women; the Sustainable Development Goals adopted by the General Assembly; and the Climate Change Declaration. The programmes of action adopted by these international conferences targeted many interrelated areas of concern and called for improved statistics to monitor progress. In deciding which social groups merit monitoring in regard to measuring the disadvantages suffered by particular groups

[112] Spoorenberg, T. (2020). "Data and methods for the production of national population estimates: An overview and analysis of available metadata" United Nations, DESA, Population Division, Technical Paper No. 2020/01 https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Sep/un_2020_techpaper1.pdf

[113] European Commission (2003). Basic Methodology for the Recalculation of Intercensal Population Estimates. Working Papers and Studies (3/2003/E/no 27), Luxembourg: Office for Official Publications of the European Communities; Statistics Canada (2012). Population and Family Estimation Methods at Statistics Canada. Ottawa: Statistics Canada, Demography Division, Catalogue no. 91-528-X; U.S. Bureau of the Census (2012). Methodology for the Intercensal Population and Housing Unit Estimates: 2000 to 2010. Washington, D.C., U.S. Bureau of the Census, Revised October 2012.

[114] United Nations, Department of Economic and Social Affairs, Population Division (2022). World Population Prospects 2022: Methodology of the United Nations population estimates and projections. UN DESA/POP/2022/TR/NO. 4

[115] A/CONF.157/24 (Part I), Chapter III.

[116] *Report of the International Conference on Population and Development, Cairo, 5–13 September 1994* (United Nations publication, Sales No. E.95.XIII.18), Chapter I, Resolution 1, Annex.

[117] *Report of the World Summit for Social Development, Copenhagen, 6–12 March 1995* (United Nations publication, Sales No. E.96.IV.8), Chapter I, Resolution 1, Annexes I and II.

[118] *Report of the Fourth World Conference on Women, Beijing, 4–15 September 1995* (United Nations publication, Sales No. E.96.IV.13), Chapter I, Resolution I, Annex II.

of people, each country should determine which groups within it need special attention. Some of the common factors leading to social disadvantage are gender, age, physical or mental impairment, race and creed. The disadvantaged are not necessarily small in number; they may constitute the majority of the population.[119]

3.527. To meet the need for statistics on gender equality, many activities have been undertaken during the last two decades at the national and international levels to improve concepts, definitions and classifications for collection of statistics related to women and men. In the present publication, the importance of the population and housing census as a data source has often been stressed. The population and housing census is also the principal or sometimes the only comprehensive national data source with respect to meeting the need for statistics on children, youths, older persons and persons with disabilities in the development of policies and programmes at the national and international levels. Therefore, it is important that countries identify data requirements concerning various population groups of particular interest when planning their censuses and ensure that the definitions and classification to be followed in censuses are appropriate and also consistent with those in use for the entire population.

3.528. Furthermore, the census tabulation plan should ensure in advance the inclusion of all relevant details about special population groups and a range of cross classifications for each group, with a view to analysing its social and economic conditions. Concepts and methods for the census and the tabulation plan should be reviewed with users concerned with statistics for each special population group. In the case of some groups, for example persons with disabilities, a special set of questions is required to identify members of this population group. In the case of others, standard questions, for example on age, are sufficient to identify groups such as children, youths and older persons. In both cases, most variables needed for cross-tabulations are already provided for in the international recommendations and many national censuses. In the census operations, however, attention will often need to be given to improvement of coverage, quality-of-data issues and avoidance of stereotypic treatment. The present section deals with gender issues, a few special population categories such as children and youths, older persons, and persons with disabilities, so as to assist in the preparation of detailed tabulations and databases according to international standards (See section on Planning and Management).

(a) Statistics on gender equality

3.529. The global conferences on women have contributed to an increased awareness of the importance of statistics not only on women and girls but, more broadly, on gender issues. For example, in developing census plans in a number of countries, efforts have been made to review and assess the adequacy of statistics for understanding the diversity of both women's and men's lives. Improvement of statistics and statistical methods related to gender issues should be an important priority in all stages of work on the census, in planning, data collection, analysis and dissemination, and in all topics.

3.530. In addition to the more general problems of the quality of census data, other types of problems that apply particularly to women and stem from gender-based stereotypes and biases have been noted. Similarly, the notion that only men can be heads of the household affects the way questions have been

[119] *Note by the Secretary-General transmitting the report of the Expert Group on the Statistical Implications of Recent Major United Nations Conferences* (E/CN.3/AC.1/1996/R.4), annex, paragraphs 68–69. Presented to the Working Group on International Statistical Programmes and Coordination at its eighteenth session, New York, 16–19 April 1996.

designed and asked in censuses. Such stereotypes also affect the way respondents reply to the questions. A common problem, for example, is classifying women automatically as outside the labour force without asking whether they perform any work for pay or profit, even as a part-time or secondary activity.

3.531. Another problem relates to biases in the collection, processing, compilation, and presentation of data. For example, when census tabulations are prepared for the employed by occupation, they may be prepared either for males only or for both sexes, on the assumption that information on the occupational pattern of women is not of much use. When tabulations of the employed by occupation are disaggregated by sex, main gender differences in occupations may be missed if the occupation data are presented only at the two-digit level of the classification. Rather, special tabulations showing, for example, the 10 or 20 detailed occupational groups with the highest concentrations of women or men would be needed to render visible gender-based occupational segregation.

3.532. During the past few decades, considerable effort has been devoted, on the one hand, to reviewing such bias and its impact on statistics concerning gender equality and, on the other hand, to improving the concepts and methods involved in the collection of data in censuses and surveys. Related improvements in the revised System of National Accounts and recommendations concerning statistics of work, employment and labour underutilization adopted by the International Conference of Labour Statisticians in their 19th session[120] in 2013 and 21st session[121] in 2023 are also of importance to the population census. They are intended to overcome the above-mentioned conceptual deficiencies and to identify and provide measures for all productive activities (that is, all forms of work) performed by women and men, whether paid or unpaid. Similarly, efforts at the national level have been focused, for example, on eliminating biases in concepts, classifications and definitions of head of the household.

3.533. Important statistical series and measures on the status of women can be readily obtained based on the topics and recommended tabulations from censuses. Furthermore, in the case of most topics, the primary unit of classification is the individual and therefore a vast array of indicators may be obtained by devising appropriate additional cross classifications- for the female and male populations separately. For an illustration of census topics and tabulations that are useful for developing comprehensive statistics on women and men, see the Minimum Set of Gender Indicators,[122] *Integrating a Gender Perspective into Statistics,*[123] *Methodological Guidelines for the Gender Analysis of National Population and Housing Census Data.*[124] The household and family status classifications presented in paragraph 5.162 are appropriate for analysing the living situation of women and men, with specific reference to single mothers and fathers and older women and men living alone.

3.534. It should be emphasized that while all data collected at the individual level can be presented by sex, this is not always done. Cross-classifications by sex tend to be suppressed when cross-tabulations become complex with multiple-variable tables. In order to satisfy one basic condition for gender statistics, which is that all statistics on individuals should be presented by sex, sex should be considered the overriding variable in all tables, irrespective of the medium of storage or dissemination. This disaggregation should be provided in all publications, databases and census tables on individuals.

[120] https://webapps.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/publication/wcms_220535.pdf
[121] https://ilostat.ilo.org/about/standards/icls/icls-documents/
[122] Available at: https://gender-data-hub-2-undesa.hub.arcgis.com/.
[123] United Nations ST/ESA/STAT/SER.F/111, Sales No. 13.XVII.9.
[124] UNFPA, 2014.

3.535. Another important consideration is to broaden the target of dissemination and use of census data by popularizing the statistics that are published. One approach to achieving this wide outreach is to present statistics in the form of charts and simplified tables, with a simple and clear interpretation of the data. Countries planning to issue an analytical report might wish to consider using such innovative techniques and formats. The analytical publication could cover the main census topics or alternatively a few areas that are especially important to understanding the relative position of women and men in the country.

(b) Statistics on children and youths

3.536. Extensive data on children and youths are available in censuses but may need improvements in terms of coverage to ensure all children and youths are counted, and quality of information on specific characteristics of these sub-populations, and on the range of tabulations available.

3.537. For statistical purposes, "children" are defined as persons under 15 years of age[125], and "youths" are defined as those aged 15–24. However, it is useful to further divide these special groups by 5-year age groups (or nationally, by groups of specific school ages) because of the rapid changes in characteristics in this age range, such as in school attendance, marital status and activity status. Also, because of differences by sex in the age at marriage or family formation, family or household status and entry into the labour market, data should be classified not only by age but also by sex. To this end, the distribution by single years of age and sex is useful. If single-year age distribution is not feasible for young children under age 5, it would be desirable to distinguish between those under 1 year of age (infants) and those aged 1–4. For youths aged 15–19, it would be desirable to distinguish between those 15–17 years of age and those 18–19 years of age, or to have a distinction corresponding to the age below which the country considers an individual to be a minor.

3.538. For the purpose of developing statistics on children, the principally recommended topics include (a) sex, (b) age, (c) school attendance (for school-age children) and (d) relationship to the reference member (or head) of the household.

3.539. Given the priority on the girl child highlighted by the World Summit for Children (1990), the International Conference on Population and Development (1994) and the Fourth World Conference on Women (1995), special attention needs to be given to improving and disseminating statistics on children by sex. Of particular concern is the situation of the girl child with respect to school attendance, mortality, early marriage, early fertility and other issues of importance. A basic problem with statistics on the girl child is that data on children ever born and children surviving tend not to be disaggregated by sex at either the questionnaire design or the tabulation stage. These data are used for indirect estimates of child mortality.

3.540. The principal topics of investigation identified for children apply also to youths, with the following additions: (a) marital status, (b) literacy, (c) educational attainment, (d) economic activity status, (e) number of children born alive and (f) age at marriage or family formation.

[125] Note that UNICEF uses under 18 to define children based on the Convention of the Rights of the Child. The Convention defines a child as any human being below the age of 18 years unless the law of a specific country determines an earlier age of majority.

3.541. Some of the useful statistics and measures can be readily compiled based on the abovementioned topics, while any additional indicators can also be obtained based on more detailed cross classifications using other recommended census topics or tabulations.

(c) Statistics on older persons

3.542. Older persons are defined by the United Nations as all persons aged 60 years and over. For purposes of classification, depending on the national situation, it is useful to tabulate data by 5year age groups up to age 100, instead of including them in the single broad age category 60 and over.

3.543. For older persons also, extensive data are available in population and housing censuses but may need detailed age-sex classification, as described below.

3.544. For the purpose of developing statistics and indicators on older persons, the principal topics in census recommendations include (a) sex, (b) age, (c) marital status, (d) labour force status, (e) income, (f) type of household, (g) type of living quarters, (h) institutional population and (i) disability.

3.545. The statistics needed for studies of older persons are disparate, depending as they do on national policies and circumstances. Internationally, no illustrative list of indicators is available to ensure appropriate census tabulations.

(d) Statistics on persons with disabilities

3.546. The census can provide a valuable source of information on the frequency and distribution of disability in the population, at national, regional and local levels. Experience shows that although an increasing number of countries ask questions about disability in their censuses, the presentation of disability data has often been limited to tabulations showing the number of specific severe disabilities present in the population. Unfortunately, cross-tabulations with other characteristics are not usually provided.

3.547. A great deal of work on concepts, classifications and development of statistics on persons with disabilities has been undertaken in recent years, particularly through the work of the Washington Group on Disability Statistics[126]. For the third time, recommendations on including disability questions in a population census are included in this document. A brief treatment of this topic is given below to highlight issues involved in preparing detailed census tabulations on persons with disabilities.

3.548. For the purpose of developing statistics on the situation of persons with disabilities the principal topics in census recommendations that would be necessary for the assessment of equality of opportunities include (a) sex, (b) age, (c) place of residence, (d) type of household, (e) marital status, (f) educational attainment and school attendance, (g) labour force status, (h) status in employment, (i) industry and (j) occupation.[127]

[126] For more information on the Washington Group on Disability Statistics, see: http://www.cdc.gov/nchs/citygroup.htm.

[127] The DHS Disability module is a tool to capture such aspects

3.549. Not only should the tabulation plan for the disability data include prevalence rates by sex and age, but it is also very important that tabulations comparing persons with and without disabilities on key social and economic characteristics be presented.

(e) Ethnocultural characteristics

3.550. Receiving information about the ethnic composition of the population allows deeper study of the ethnic background of a country's population, especially with respect to indigenous populations, international migrants, stateless and forcibly displaced populations, and other specific groups of population (for example nomads).

3.551. There are some difficulties in collecting this information since some population groups may name their ethnic identification based on its local meaning, and so, in order to correctly allocate these persons to their particular ethnic group it is necessary to compile a list of ethnic groups, sub-ethnic groups and local definitions of small ethnic population groups. This will allow for obtaining accurate and standardised data about the ethnic composition of population. It would also be useful if scientists and specialists in the field of ethnography, as well as organizations of indigenous peoples, would be involved in creating such a list.

3.552. In order to obtain comprehensive information characterizing ethnic composition of the population, it would be useful to tabulate data by (a) sex, (b) age, (c) place of residence, (d) marital status, (e) births, (f) deaths, (g) educational attainment, (h) labour force status (i) status in employment (j) industry, (k) occupation, and (l) type and size of household.

3.553. It is important to obtain comprehensive information on indigenous populations in order to have statistics on the number as well as the demographic and socio-economic structure of the given population group. These data would provide valuable information to support the development of programmes for social support of indigenous peoples.

3.554. Statistics about the ethnic composition of international migrants together with information about country of birth and citizenship will help to more precisely determine the inward flow and volume of international migration.

3.555. Population censuses are also the sources of information about religious identification of the population. It would be useful to obtain this information by (a) sex, (b) age, (c) ethnic group, (d) place of residence, and (d) place of birth. This information would be useful to study distribution of religious affiliations.

3.556. Information about knowledge of languages is widely used. Countries find it useful to study the official language of the country as well as the use of mother tongues or some other languages. In any case it would be useful to have this information by (a) sex, (b) age, (c) ethnic group, (d) place of residence, and (e) place of birth.

3.557. Information about knowledge of the official language of the country would be very useful for studying the integration of international migrants and may be used, for example, for development of programmes to learn the language.

3.558. Information about knowledge of the mother tongue of indigenous populations is very important. This information could allow obtaining statistics of “indigenous” languages and would be very useful for development programmes to support the development of those languages.

(f) Statistics on poverty

3.559. The census data can provide a valuable source of information on conditions of life of households as a proxy measure of poverty to complement quantitative survey data. Census data can provide a quantitative approach to measuring poverty. In some cases, countries may compile multiple deprivation indices using census data such as tenure, economic activity status, income, household amenities.

5. Development indicators

3.560. Indicators are required by countries to track the progress of various developmental goals and, as such, efforts must be made by census offices to produce relevant indicators to meet this need. In the 2020 round of population and housing censuses, many countries produced indicators based on the Sustainable Development Goals as was recommended. The type of indicators required to meet international and national reporting requirements need to be taken into account early in the planning phase of the census.

3.561. It should be emphasized that both global and national reporting and monitoring require reliable and comparable national data for the compilation of indicators. In this regard, it is of paramount importance that countries have the statistical capacity to produce, analyse and disseminate the requisite data for these indicators. The availability of reliable statistics and the capacity of governments to systematically measure and monitor indicators is a critical success factor for the achievement of development goals. The lack of statistical capabilities in some developing countries makes it difficult to obtain good and reliable data. Many countries do not have a sustainable, coherent programme of household surveys, or administrative data systems that can be used to produce basic statistics routinely. Where basic statistical systems are not available, the global monitoring may have to rely on national and international estimates of widely varying quality and reliability. This may lead to misjudgements regarding progress and may undermine the effectiveness of policy interventions at national and subnational levels.

3.562. The 2030 Agenda for Sustainable Development, adopted by all United Nations Member States in 2015, provides a shared blueprint for peace and prosperity for people and the planet, now and into the future. At its heart are the 17 Sustainable Development Goals (SDGs), which are an urgent call for action by all countries - developed and developing - in a global partnership. They recognize that ending poverty and other deprivations must go hand-in-hand with strategies that improve health and education, reduce inequality, and spur economic growth – all while tackling climate change and working to preserve our oceans and forests.

3.563. The 17 SDGs are divided into 169 accompanying targets and over 230 indicators, many of which are subject to disaggregation, and which pose a major challenge to the data generation capabilities of National Statistical Offices. In some cases, SDG indicators can be directly generated from the population and housing census which allows for disaggregation of the indicators at the sub-national level and for sub-populations that are harder to report on through sample surveys. (Reference and further detail - Technical Report: Measuring Sustainable Development Goals Indicators through Population and Housing Censuses and Civil Registration and Vital Statistics Data – UNSD Oct 2022).

X. Documentation of census experience

3.564. The cumulative experience of previous censuses in a country is very useful in the preparation of a new census. Because of the lapse of time between censuses (generally 10 years) and the likelihood that experienced staff may leave the census office, it is essential that there is a comprehensive record of how the census was planned, organized and conducted.

3.565. The census office should, therefore, plan for and implement a knowledge management system to assemble complete records on plans, activities, and decisions taken during the entire census operation. This would entail documentation and archiving of information related to plans and their implementation, as well as problems encountered and how they were resolved at each stage of the census cycle. It is recommended that documentation of census experience be undertaken at each stage of the census operation and not be left until the end of the census process. This would include plans, decisions and activities related to preparatory activities, the methodology of the census, fieldwork or other data collection activity, data processing, analysis, dissemination, cost and implementation of the census budget, and evaluation of performance of each of these activities. Examples of items to track or monitor include implementation of activities, time taken to complete an activity, resources used and cost. All these should be assessed against set goals so that changes to plans can be recorded, including information on what changed and why. Tracking and systematically recording the census experience should also take into account risks encountered and how these risks were managed. For more information see Part Four on “Quality assurance”.

3.566. Use of knowledge management tools and techniques is thus beneficial for preserving institutional memory in a codified way so that lessons learned from the past may be used for better management of future census planning and execution. Records in the system should be arranged in such a way that information on each aspect of the census operation is found easily.

3.567. Systematic recording of census experience is not an end in itself. It is recommended that every country prepare and, if possible, publish an administrative and methodological report, as a census “historical memory”, based on information that has been recorded in the knowledge management system (see paragraphs 3.414–3.415 on administrative report). Depending on the methodology of the census, the administrative and methodological report should contain information on the manner in which the census was planned, organized and conducted, as well as important methodological and other problems encountered at various stages of the programme. As appropriate, the report should provide specimens of the census questionnaires and forms, instructions for the enumeration, and detailed information on the cost of the census and on the implementation of the census budget, as well as points to be considered in future censuses.

3.568. The structure of the report could be similar to the structure of the project plan. It is important that the report be as comprehensive as possible, covering all stages and aspects of census planning and operations, including fieldwork, processing, analysis, dissemination and evaluation. It is important to note that while such a report would be based on items and information in the knowledge management system, it may not necessarily contain detailed descriptions of all the processes or information as some may be for internal use only. This report would both assist the users of the census results in appraising and interpreting the data and facilitate the proper planning of future data collection programmes, including population and housing censuses.

XI. Archiving individual records

A. Purpose of archiving individual records

3.569. The focus of the following elaboration of archiving is on individual census records irrespective of their format – paper questionnaires or electronic records. The discussion on the complete process and characteristics of archiving all census documents is presented in paragraphs 3.564–3.568 on systematic recording and documentation of census experiences.

3.570. Individual census records refer to either census paper questionnaires where these were used for collecting information from the population or to the digital records on each enumerated person and household if the information was collected electronically or online, including their direct identifiers, such as name and address. In the case of bimodal or multimodal data collection, that is, a combination of paper and non-paper questionnaires as, for example where Internet forms and mail out/mail back paper questionnaires were used, the resulting collection of individual records would also be a combination of paper and digital recordings.

3.571. In the case of digital records, the accompanying documentation becomes an indispensable part of the archiving process. As a number of variables in the digital record are presented as codes, it is necessary to archive all the codebooks and all the other documentation, such as the data collection instrument, that are needed for fully unlocking the value of each of the variables in the record. While this accompanying documentation is also valuable in the case or archiving paper questionnaires, these are by nature visual, thus requiring only reading skills and the knowledge of the language initially used for filling them to understand the content, as long as they are in a good state of preservation.

3.572. The essential purpose of archiving individual census records is to keep them safe for future use, primarily in the domain of genealogical research and longitudinal social and anthropological studies, as well as for use by historians and demographers. The release of archived individual census records is subject to a passage of time prescribed in census legislation and that usually encompasses many decades, thus ensuring that the use of individual information would not endanger the confidentiality and the privacy of the respondents.

3.573. Consequently, the importance of providing detailed guidance on the process of archiving individual census records in the census legislation cannot be overstated. These provisions provide the legal basis for maintaining the archives and procedures related to the release of archived records. The time lag between the data collection and the release of the archived records needs to be clearly indicated – it varies from 72 years (in United States of America for example) to 100 years (in the UK). In some cases, the original questionnaires are only temporarily stored before being fully destroyed as in India, one year before the next census takes place.

B. Procedures for archiving

3.574. Archiving a vast number of records represents a considerable challenge in all circumstances. In the case of individual census records it may be compounded by the sheer number and format. However, in all cases the national statistical authority needs to develop an institutional strategy for archiving, based on three components: organizational infrastructure, technological infrastructure and resources.

3.575. **Organizational infrastructure** refers to the arrangements that need to be put in place within the national statistical office in such a manner as to ensure the efficiency of the archiving and eventual retrieval process. In most cases it is a centralized unit within the office that is put in charge of the archiving, maintenance, secure storage and eventual release[128] of individual records. Once the time lapse mandated by the law for the release of records expires, the actual release to the public is usually implemented by dispatching relevant batches to the libraries covering parts of the country to which the records refer and to a central national library.

3.576. **Technological infrastructure** refers to the actual technology used for archiving. In the contemporary circumstances, storing huge numbers of paper questionnaires would prove not to be cost-effective, as it would require a significant physically secure structure, regulated temperature and humidity, and a host of other requirements, including protection from fire hazards, floods and extreme weather events. Consequently, in most cases the actual questionnaires are scanned and images of them stored in various electronic storage devices.

3.577. The technological infrastructure does not refer only to the actual technology used in the archiving process – it also consists of a series of protocols for archiving and establishing cross-references that enable successful retrieval of records. Therefore, the technology should be built around a well-developed archiving scheme that enables efficient identification and retrieval of the records.

3.578. In the case of some more recent censuses, in countries where digital and paper questionnaires have been available interchangeably to respondents, data from paper questionnaires have been stored digitally and the paper forms destroyed, with the resulting digital census data subsequently kept in secure cloud storage alongside accompanying metadata and documentation. It is best practice for such digital archives to be backed up in physical and cloud formats, with regular checks to ensure the records remain stable and usable over time.

3.579. **Resources** for archiving need to be taken into account at the early stages of planning for the census, in the context of the technological and organizational infrastructure. In assessing the volume of the necessary funds it is necessary to adopt a strategic, long-term approach, as the archiving, maintaining and releasing would essentially constitute a perpetual activity as long as censuses are part of the national statistical systems: there would always be a need to prepare either for the next round of release of records or for archiving the newly acquired one.

C. Archiving individual records and microdata

3.580. Individual census records as described above for archival purposes differ from the census microdata in a most significant manner: they retain the direct identifiers – name, address, enumeration area – as these very identifiers represent the essential information for genealogical, anthropological, historical and longitudinal social studies. In the case of microdata, these identifiers would be removed, as well as any other that can directly or indirectly identify the respondent.

[128] Or the transfer of individual records to the National Archive or Public Records Office for public release.

3.581. Microdata are defined as electronic records pertaining to each unit of observation; in the case of the population and housing censuses, it would be individuals, housing units and households. Electronic records with variables for each observation unit (individuals, housing units, households). While names and addresses are the most obvious identifiers, certain housing unit characteristics, especially in combination with other information, could increase the risk of re-identification. National statistical agencies should carefully consider the level of detail provided in housing unit data within microdata files to ensure confidentiality. This may involve techniques like data aggregation or suppression to minimize the risk of re-identification.

3.582. It is expected that the use of anonymized microdata will continue to be a standard feature of census data dissemination for the 2030 round of censuses. Consequently, paragraphs 3.466–3.485 of these *Principles and Recommendations* present a comprehensive elaboration of principles and protocols for dissemination of microdata files.

PART FOUR. QUALITY ASSURANCE AND EVALUATION

I. Quality assurance

A. Plans for quality assurance

4.1. Regardless of census methodology used – full field enumeration, register-based or combined – the visibility and significance of a census plays a large role in the reputation of a national statistical system's capacity to produce and disseminate high-quality data. All countries use it as the base to project future population for planning purposes. For countries that conduct a full field enumeration or combined census, it also provides the sampling frame for future surveys. Furthermore, a high-quality census increases the trust of data users from all sectors. Thus, the quality of a census has serious implications for long-term data quality, planning and evidence-based policy making.

4.2. However, since most countries conduct population and housing censuses once in 10 years, retaining institutional knowledge from one census to the next becomes challenging. But the experience gained from previous population and housing censuses as well as other censuses, such as agricultural censuses, is very useful in planning for a quality assurance and improvement programme for the current census. It is of critical importance that errors detected from previous censuses or similar activities are documented and used as the basis for developing quality assurance measures for the next census. Each country must have a quality assurance and improvement programme in place to measure the quality of every phase of the census operation. A quality assurance programme should also be viewed as a quality improvement programme as errors detected can then be addressed promptly. If data are of poor quality, decisions based on these data can lead to costly mistakes. Eventually, the credibility of the census and even the national statistics office itself may be called into question.

4.3. Quality assurance should be viewed as an important element of any census, regardless of whether it is a full field enumeration, combined or a register-based census. The quality assurance and improvement programme should be developed as part of the overall census project and integrated with other census plans, schedules and procedures. The programme should be established at all phases of census operations, based on the United Nations National Quality Assurance Framework. Establishing a quality assurance and improvement system and setting quality standards at the planning stage is crucial to the success of the overall census operation.

B. Quality assurance components

4.4. Census quality management should be viewed within the broader context of managing the statistical system and managing the institutional environment. At the operational level, quality management should be comprehensive and should cover all census activities including planning, mapping, questionnaire development, data collection, processing, evaluation, analysis, and dissemination.

4.5. The United Nations National Quality Assurance Frameworks (UNNQAF)[129] for Official Statistics provides a framework for addressing quality within the context of official statistics and contains 19 principles that are relevant for quality management in population and housing censuses. These quality principles and associated requirements consist of four levels, ranging from overarching institutional and cross-institutional management through to the management of statistical production processes and statistical outputs:

1. Managing the statistical system.
2. Managing the institutional environment.
3. Managing statistical processes.
4. Managing statistical outputs.

1. Managing the census within the national statistical system

4.6. A precondition for the efficient production of high-quality census statistics is a national statistical system that supports the coordination of census activities with other parts of that system, the management of relationships with stakeholders, and an enabling legislative environment. Due consideration should be given to the following principles:

i. Coordinating census activities within the broader context of the national statistical system.
ii. Managing relationships with data users, data providers, and other stakeholders.
iii. Managing statistical concepts, definitions, classifications, methods, and procedures.

2. Managing the institutional environment

4.7. Quality will be better supported by sound institutional arrangements that assure the following principles:

iv. *Professional independence*. Census results should be developed, produced, and disseminated without any political or other interference or pressure.
v. *Impartiality and objectivity*. The census should produce and disseminate statistics in a way that is professional, impartial, and unbiased.
vi. *Transparency*. Census policies and management practices are well-documented and publicly available.
vii. *Statistical confidentiality and data security*. Information provided by respondents will be kept confidential and will be used for statistical purposes only.
viii. *Commitment to quality*. Census managers should be dedicated to assuring quality in their work, and the work of their teams, and systematically and regularly identify strengths and weaknesses to continuously improve census processes.
ix. *Adequacy of resources*. The financial, human, and technological resources should be adequately planned for and adequate both in magnitude and quality, and sufficient to meet their needs regarding the development, production, and dissemination of census-based statistics.

3. Managing census processes

4.8. Quality is the outcome of sound statistical processes, and deficiencies in quality (for example, delays in processing or lack of accuracy in the results) are usually the result of deficiencies in the process

[129] https://unstats.un.org/unsd/methodology/dataquality/un-nqaf-manual/

rather than the actions of individuals working in that process. Census processes should be based on the following principles:

x. *Methodological soundness*. This encompasses adherence to professional methods and (internationally) agreed standards, guidelines, and good practices.

xi. *Cost-effectiveness*. Resources should be effectively and efficiently used, and census objectives achieved at a reasonable cost.

xii. *Appropriate statistical procedures*. Effective and efficient statistical procedures should be implemented throughout the census production chain.

xiii. *Managing respondent burden*. Individuals who provide census data are essential to its quality. Minimizing respondent burden is crucial, requiring a careful balance between data needs and questionnaire length. While register-based censuses reduce direct respondent burden, maximizing the use of administrative and other data sources remains essential to optimize data collection efficiency.

4. Managing census outputs

4.9. It is suggested that the outputs of any census should follow all the following principles:

xiv. Relevance and completeness

xv. Accuracy and reliability

xvi. Timeliness and punctuality

xvii. Ready accessibility and clarity

xviii. Coherence and comparability

xix. Interpretability

These principles are expanded upon below in paragraphs 4.11 to 4.17.

4.10. In the census context, some attributes of quality assurance may be emphasized over the rest. The census should produce statistics that are relevant to data users. A census with full field enumeration is a particularly expensive exercise to undertake and creates a burden on respondents. Therefore, it is crucial to ensure that any demand for data is met to a minimum acceptable level and that topics for which there is little demand are not included on the census form. Consulting with users of census data as one of the first steps in designing the census process is a positive public relation undertaking and an efficient, transparent means of determining the demand for potential census topics.

4.11. The relevance of data or of statistical information is a qualitative assessment of the value contributed by these data. Value is characterized by the degree to which the data or information serve to address the purposes for which they are produced and sought by users. Value is further characterized by the merit of those purposes, in terms of the mandate of the agency, and legislated requirements. In the context of a census the concept of fitness for purpose as a measure of relevance is important. If it is only necessary that data are available at the broad level (for example, national or major civil division level of geography; broad demographic level), user requirements could be met more cheaply and effectively through a sample survey or the use of administrative records.

4.12. The completeness of data (defined as the degree to which statistics fully cover the phenomenon or population they are supposed to describe) means not only that statistics should serve user needs, but also that they should serve them as completely as possible, taking restricted resources and respondent burden into account.

4.13. Statistical agencies should develop, produce and disseminate statistics that accurately and reliably portray reality. Accuracy of data or statistical information is the degree to which those data correctly estimate or describe the quantities or characteristics that the statistical activity was designed to measure. Accuracy has many attributes, and in practical terms there is no single aggregate or overall measure of it. Of necessity these attributes are typically measured or described in terms of the error, or the potential significance of error, introduced through individual major sources of error, for example coverage, sampling, non-response, response, processing and dissemination.

4.14. *Timeliness* of information reflects the length of time between its availability and the event or phenomenon it describes, but considered in the context of the time period that permits the information to be of value and still acted upon. It is typically involved in a trade-off with accuracy. *Punctuality* is defined by the degree to which the availability of the data meets pre-announced release dates.

4.15. *Readily accessibility reflects* the availability of information from the holdings of the agency, also taking into account the suitability of the form in which the information is available, the media of dissemination, and whether the user has reasonable opportunity to know they are available and how to access that information. The affordability of that information to users in relation to its value to them is also an aspect of this characteristic. *Clarity* refers to the degree to which the data are understood (regardless of how accessible they are), particularly for non-experts.

4.16. *Coherence* reflects the degree to which the census data can be combined with other statistical information within an integrated framework over time. The use of standard concepts, definitions and classifications promotes coherence. Equally important is *internal coherence of data*, referring to the consistency of information across different topics of the census and census outputs. This coherence is usually addressed through the meticulous development of data edits. Data are most useful when *comparable* across space, such as between countries or between regions within a country, and over time. More and more emphasis is also put on enabling comparison of geography over time, as well as maintaining consistency and comparison of census topics from one census to another. Another aspect of comparability that is sometimes overlooked is the degree to which data referring to the same characteristics derived from different sources (for example, rates of unemployment as measured by the census and the labour force survey) are comparable.

4.17. For the purposes of achieving *interpretability*, *Metadata* should be provided covering the underlying concepts, the variables and classifications used, the methodology of data collection and processing, and indications of the quality of the statistical information to enable the user to fully understand census statistics and their limitations.

C. Other international standards, tools and references

4.18. It should be noted that there are other international standards and quality frameworks and tools that may be valuable for quality assessment work:

a) The Handbook on Management and Organization of National Statistical Systems[130], the 4th edition of a series of publications on the management of statistical organizations, was endorsed by the 52nd session of the UNSC in March 2021 as a non-prescriptive compendium targeting chief statisticians, senior managers, and staff members of statistical organizations. The Handbook provides guidance on how to develop and maintain national statistical capacity that is fit for

[130] https://unstats.un.org/capacity-development/handbook/

purpose and apprises critical issues and topics, including the integration of innovative data sources and technologies in the production of statistics and indicators. The handbook includes a dedicated chapter on quality which describes a number of national quality frameworks, guidelines, and tools for potential use by countries.

b) Also, for reference, UNSD maintains an extensive list by country of quality references and tools[131].

c) For countries that conduct register-based censuses, the Principles and Recommendations for a Vital Statistics System, Revision 3, contains information for assessing quality. It provides recommendations on some of the practices adopted to assess the quality of civil registration data and the quality of vital statistics based on those data.[132] Additionally, UNSD maintains an inventory[133] of resources on the use of administrative data for statistical purposes. The aim of this inventory is to make materials on the use administrative data for statistical purposes more readily available and easy to find. The inventory includes a variety of materials, ranging from recommendations to practical examples in different contexts.

d) Countries conducting a combined census should identify from planning stage the elements to be collected from enumeration, from registers and the methods for linking this information. As a consequence, certain stages and components could benefit from different quality frameworks such as the Guidelines for Assessing the Quality of Administrative Sources for Use in Censuses.[134]

e) The Generic Statistical Business Process Model (GSBPM) is a model for the business processes that statistical offices follow to produce official statistics, conceptualised in terms of phases and associated sub-processes. In order to monitor the quality of the statistical production process for each of these phases, a set of quality indicators have been developed in collaboration by representatives from the national statistical agencies. This work maps quality indicators for surveys to the structure of the GSBPM.[135] (See Part Two, chapter II)

D. The role of managers

4.19. Managers have a vital role in setting quality standards and quality management as a priority in the census (See also Part Two, Chapter III: Strategic objectives and management). The biggest challenge to managers is first to establish a culture within the census agency that has a focus on overall quality management issues and to obtain the commitment of staff to strive to achieve high-quality goals. At the same time, managers need to be aware that to achieve high-quality outcomes they need to give their staff responsibility to achieve these outcomes. Managers who do not delegate responsibility will find it difficult, if not impossible, to establish teams that strive for high-quality outcomes.

4.20. The project manager is responsible for the project work from the initial kick-off through to closure. Only the primary responsibilities are given in the list below, which can be expanded considerably. The responsibilities of the project manager include:

a. Using quality project management tools and techniques;

b. Identifying and managing the project stakeholders, both those inside the National Statistical Office as well as external groups and organizations that can impact the project outcomes;

c. Creating the conditions for good team working by:

(i) Setting the team norms and behaviours;

[131] https://unstats.un.org/unsd/methodology/dataquality/quality-references/

[132] M19Rev3-E.pdf (un.org)

[133] https://unstats.un.org/UNSDWebsite/capacity-development/admin-data/inventory

[134] https://unece.org/sites/default/files/2021-03/03_CensusAdminQuality_forConsultation_0.pdf

[135] Quality Indicators - Generic Statistical Business Process Model - UNECE Statswiki

(ii) Deciding responsibilities and coaching team members in new skills;
(iii) Leading, guiding and directing team members;
(iv) Controlling the work of the team – input and output;
(v) Building trust and respect within the team;
(vi) Encouraging personal growth, development, empowerment and continuous learning of the team members;

d. Establishing a project support office responsible for:
(i) Developing project plans and budgets aligned to the strategy;
(ii) Prioritizing activities within the project;
(iii) Allocating and securing resource (for example financial and technological) commitments;
(iv) Working with established working practices and customs and developing relevant policies, systems and processes to implement objectives and plans;
(v) Developing a project communication plan;
(vi) Coordinating between internal and external stakeholders;
(vii) Developing a governance framework for the project;
(viii) Monitoring and tracking project progress against set objectives and plans and compiling relevant reports, for example progress, variance and status reports;
(ix) Monitoring and controlling expenditure and compiling financial and variance reports;
(x) Solving the problems that interfere with progress;
(xi) Informing internal and external stakeholders of progress and status;

e. Identifying and managing the project risks by developing a risk register, establishing a risk prioritization and risk mitigation plan, and monitoring and managing these risks;
f. Working with the unknown and unpredictable including assuring professional independence of the project staff and demonstrating objectivity through transparency;
g. Implementing improvement and change initiatives to ensure the soundness of methodologies, the cost-effectiveness of resources, the appropriateness of statistical procedures and the proactive management of respondent burden;
h. Delivering the project deliverables and benefits;
i. Leading the project team;
j. Evaluating and closing the project.

4.21. Firstly, managers must ensure that staff understand the philosophy behind the approach to quality. As mentioned above, staff involvement is a vital ingredient to quality improvement. Therefore, an environment needs to be established in which staff contributions are expected.

4.22. The second part of their role is to ensure that users' data requirements are understood, and that these are met (particularly in terms of relevance, accuracy, and timeliness) and built into planning objectives and into the systems that are to deliver them. Managers may have to work with multiple clients at once to satisfy project objectives and this complexity should be considered when building systems and during project planning.

4.23. Thirdly, processes need to be documented and understood by the staff implementing them. Systems and processes for implementing the quality assurance circle also need to be documented and put in place. Questions such as how quality is going to be measured, who is involved in identifying root causes of problems with quality, and how the process improvements are going to be implemented need to be answered. These will vary greatly depending on the nature of the process. Appropriate quality assurance techniques for each phase of the census are summarized below.

4.24. These processes and metrics should also be shared with external clients as appropriate, to ensure their expectations for quality are being met. The greatest test of management commitment to genuine quality improvement lies in how management approaches problem-solving. Staff will monitor management responses closely and adjust their own behaviour accordingly. Staff will act in accordance with how they see managers behave rather than what they hear managers say.

4.25. Managers who always react to problems by seeking out people to blame, or who establish systems that focus disproportionately on the merits or demerits of individuals at the expense of the team, are sending messages that are contrary to the thrust of quality improvement. An environment where the emphasis is on fault finding, rather than on finding solutions to problems, or on excessive competition, will assure that staff ceases to be part of the solution and becomes part of the problem. Managers need to take upon themselves the responsibility for problems, as they are ultimately responsible for the systems or processes that caused the problems. They should not seek to transfer the problems to lower-level staff.

4.26. However, even in the best-managed processes, there are circumstances where individuals can be justifiably responsible for negatively impacting quality. These may be individuals who do not possess adequate skills for performing their duties, and even deliberately flout procedures. These individuals need to be dealt with decisively, first and foremost, primarily by providing additional training and guidance all the way to administering disciplinary measures. Managers must deal promptly with these cases and act in a consistent manner. By doing so, managers will demonstrate to all other staff their commitment to quality.

4.27. To be successful, it is necessary to create a culture in which everyone has the opportunity to contribute to quality improvement. Most of the staff engaged in census operational work undertake routine tasks, and it is up to management to help them see the bigger picture, to motivate them and to encourage them to assume ownership of their work. This can be done by promoting a commitment to quality improvement and by adopting a consistent approach to management.

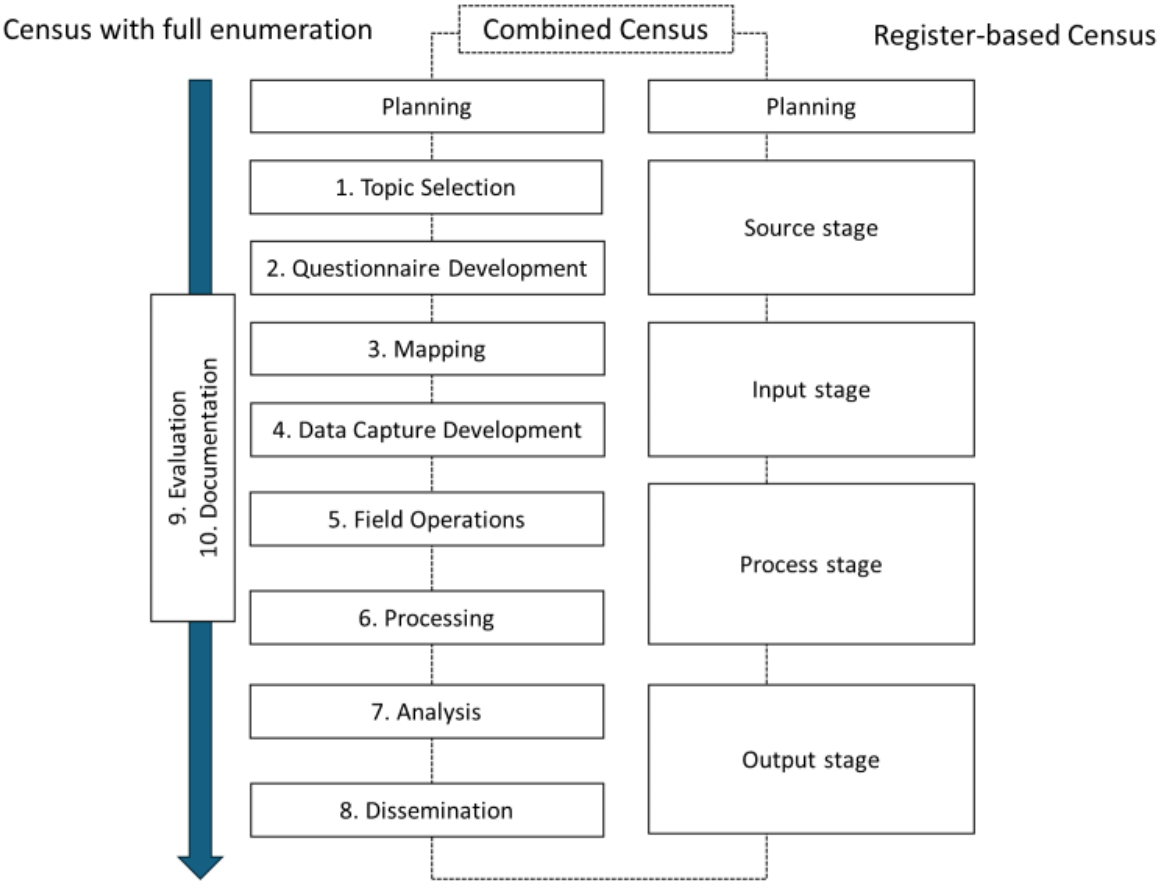
E. Quality improvement and the census

4.28. The following sections outline the way in which the concepts of quality relate to the different phases of census activities. As a basic premise, it is critical to build trust and confidence with stakeholders for documentation on the various quality considerations to be provided as soon as feasible. This can be done on a public website, presented during a presentation, or other common methods used to provide information about the census. Quality considerations are discussed in relation to:

1. Topic selection;
2. Questionnaire design, development, and processing system testing;
3. Geospatial information and mapping;
4. Development of a data capture system; (including CAPI and CAWI)
5. Field operations (including pilot testing);
6. Processing;
7. Analysis;
8. Dissemination;
9. Evaluation.
10. Documentation

4.29. While the quality considerations for these 10 stages are based on a full field enumeration census, some of the stages also apply to combined and register-based censuses. (see figure 4.1). The additional quality considerations related to data derived from administrative sources are covered in section F of this chapter.

Figure 4.1: Overview of Quality Processes Outlined in Part 4, Chapter I, Sections E and F



1. Topic selection

4.30. This section discusses the quality aspects of "topic selection". For further details and a description of the processes involved in topic selection, please refer to Part Three, Chapter II, subtopic a; and Part Five.

4.31. The first step in managing the quality of the product (namely, the statistical outputs to be produced) is to ensure that the product will be relevant to the users' requirements. The key process is extensive consultation with actual and potential users of census information, who may include for example, government ministries and departments, sub-national authorities, universities and other research institutions, the private sector, donors, organizations (or individuals) representing economic, social, cultural or special interest groups, such as religious groups, ethnic groups, and groups representing persons with disabilities and older persons. Potential topics for inclusion in the census should be agreed early in the census planning and questionnaire development phases. The key success factor in this process

is full, frank and open consultation with users, subject matter and classification experts to ensure that census outputs address well-defined user needs. If their needs are not met from the current census products, users may be reluctant to propose their needs for a future census. It is important to note that user needs will change over time, and users should be kept informed of changes to variable definitions and output classifications. A review of the extent to which user needs are met should be seen as an evaluation process feeding into the current cycle, the first step of quality management.

2. Questionnaire design and development

4.32. This section discusses the quality aspects of "questionnaire design and development". For further details and a description of the processes involved in questionnaire design and development, please refer to Part Three, Chapter II.

4.33. The next quality management task concerns the testing of each census question and the testing of the design of the form (paper or electronic version according to instrument(s) used). The development and testing of paper forms and the development and testing of electronic versions (including apps) should be treated as separate, but parallel processes with their own set of quality indicators. The results of each test should be analysed and evaluated before being fed into further design and testing. The following areas are the key internal stakeholders of the form design process and their requirements need to be taken into account:

(a) The analysis and dissemination teams (to ensure that the questions asked will deliver the data to meet the needs of users);
(b) The subject matter specialist team;
(c) The team responsible for development of the data capture system. For example, positioning of text and delineation of response areas may be dependent on the data capture and processing methodology to be adopted. It is critical that there is ongoing coordination between the form design and processing areas;
(d) The field operations team, which is responsible for training the enumeration workforce and printing the form.

4.34. The process of testing the classifications employed in the questionnaire is pivotal to ensure their alignment with both national and international requirements. This alignment guarantees that these classifications can be effectively compared and seamlessly linked to other databases. This meticulous testing procedure is imperative to guarantee the accuracy and reliability of the data. By harmonizing the classifications, the foundation is laid for cross-referencing and integration with external datasets thereby maximising the coherence of the data. Consequently, this endeavour facilitates a comprehensive analysis and provides valuable insights into various quality dimensions of the collected data. Furthermore, the rigorous testing of these classifications ensures the apt design of the data structure for the census outputs. A well-structured data framework enhances the usability, accessibility, and interpretability of the census results. This approach ensures that the data outputs are not only accurate but also optimally organized for analytical purposes, enabling more informed decision-making processes.

3. Geospatial information and mapping

4.35. This section discusses the quality aspects of "geospatial information and mapping". For further details and a description of the processes involved in geospatial information and mapping, please refer to Part Three, Chapter IV.

4.36. Many of the quality assurance components listed in section B are also applicable for mapping operations, geospatial products, and their downstream uses in the census. For example, processes that improve the completeness or accessibility of geospatial data (for example, identifying new housing units or enabling user download of geographical areas) support quality. The dual role of geospatial data in providing the operational framework for pre-census operations and the means to associate statistical observations with recognizable areas during dissemination, leads to some noteworthy quality considerations. These considerations relate specifically to the capacity of geospatial data to:

(a) Deploy and route field staff and support material efficiently;
(b) Define assigned collection and supervisory areas unambiguously; and
(c) Associate tabulated census data with administrative and other useful geographic areas.

4.37. Ensuring the quality of demarcation requires rigorous testing during census planning and design that measures the efficiency of statistical unit design for operational purposes. The exact rules used to define physical features (such as a river, or road) priority order or the target number of households or persons per statistical unit will necessarily vary by country. A high-quality demarcation helps to predict the rates at which enumerators should collect data in different types of terrain during census enumeration. These rates are estimated during census pre-tests and a full census pilot that use planned census modalities within varied geographic areas.

4.38. The quality of spatial attributes associated with geospatial data supports quality by improving interoperability and reducing duplication of effort. The use of harmonized classification and coding systems for qualitative characteristics (such as road quality, river type, urban status) and the administrative geocoding naming scheme (that is, a gazetteer) within the NSO and other agencies with mapping functions supports data coherence in cartographic products used to support enumeration and dissemination.

4.39. The sophistication of the technology suite used during a census does not singularly guarantee the quality of geospatial data produced. Progressively more advanced technologies may support improved quality as they allow geospatial data to achieve improved relational and positional accuracy. However, the technological capacity to associate geographic coordinates with housing units using CAPI does not solely ensure the integrity of the census or that a household frame can be generated from the census. Coordinate capture must be part of a coherent, reproducible census methodology with measured positional accuracy and coverage error rates to meaningfully support census data integrity. NSOs must also have the financial and human resources to maintain the frame during the intercensal period, or the potential quality gains for future statistical processes from increased technological sophistication during the census are greatly diminished.

4.40. Programs that reduce ambiguity when interpreting boundaries contribute to quality by improving accuracy and enhancing reproducibility during data collection (so that, for example, a field worker assigned to a statistical unit based on census geography unambiguously knows which housing units are contained within, or lie outside of, their unit with as little re-canvassing as possible). An ongoing program to harmonize geographic data should coordinate between the NSO, other custodians of geospatial data, local governments, and local statistical offices. Since most NSOs are not the custodians of all administrative boundaries, this coordination ideally takes place within the framework of a national spatial data infrastructure.

4.41. An ongoing and consistent approach to geographic area and boundary maintenance is not always possible in resource-constrained environments. The census may be a rare opportunity to improve the completeness and congruence of administrative/statistical geographies and physical features. Collection and collation of metadata during the census process regarding how cartographic staff and field personnel resolve boundary questions will support the integrity of the census and support data maintenance efforts during the intercensal period. These measures can be as simple as encoding the rules used to prioritize physical features or ensuring that census field notes are digitally captured and saved rather than discarded once an issue is resolved.

4.42. Programs that improve data user familiarity with geographic areas, especially functionally defined areas, improve quality by supporting usability. NSOs may even solicit data user feedback to ensure that the geographic areas used for census tabulation continue to meet their needs.

4. Development of the data capture system

4.43. This section discusses the quality aspects of "development of the data capture system". For further details and a description of the processes involved in development of the data capture system, please refer to Part Three, Chapter VIII.

4.44. The transition from paper to digital data collection introduces a new operation- that of developing a data capture system. In this section, specific quality concerns related to the development of Computer Assisted Personal Interviewing (CAPI) applications and Computer Assisted Web Interviewing (CAWI) applications are discussed. Additional material on quality considerations when capturing data from administrative records is covered in section F (Measuring the quality of data derived from administrative sources).

a) Computer Assisted Personal Interviewing (CAPI)

4.45. The development of a Computer Assisted Personal Interviewing (CAPI) application is frontloaded. If the CAPI application is poorly designed, implemented, and tested it will collect low quality data. Several software development methodologies can be applied to CAPI development. These include, creating a questionnaire specification, managing change control, and establishing a freeze date after which modifications should not be made to the data collection application. By developing a questionnaire specification which includes CAPI text, response values, consistency checks, and routing the questionnaire design and implementation are separated into two independent activities. Establishing a change control mechanism enables documentation of who, why, and when a specific change is made. Revision control software can help facilitate change control. It is important to select a date when all development will stop and after which only bug fixes are allowed. Unlike the traditional paper-based approach, CAPI applications tend to be seen as an ever-evolving piece of software. However, even a simple change can have unintended side effects, be incomplete, or simply be wrong, which in a CAPI environment can lead to a small change having an oversized negative impact. Thus, testing after modifications and retesting after fixing bugs is crucial in the CAPI application development process.

4.46. The following are best practices that will improve the quality of the data collected by the CAPI application:

i. Consideration should be given to the development of an operational control system to handle tasks that are not directly related to the interview. This will include logging in to the system, the

ability to assign and receive assignments, complete and review questionnaires, track progress, and the ability to synchronize data. Once an enumerator is logged into the operational control system and has selected an assignment the questionnaire can be prefilled with data such as the enumerator and geographical codes which will reduce human error. Additionally, it can be helpful to prefill paradata such as device id, interview start time, and software versions.

ii. At a minimum each CAPI question will comprise question text and response options. The question text should include fills, like the name of the household member, customized for each interview to reduce error. Text boxes which allow open-ended responses should be kept to a minimum and used for fields such as name. More often, a discrete number of pre-coded options should be shown as dropdown options. Ideally, the dropdown options include human-readable responses. Numeric codes can be recorded in the background without the enumerator having to be familiar with the numeric codes. Furthermore, response options can be created dynamically, so an enumerator is only presented with valid responses. Also, help and instruction text can be included to assist the interviewer. The help text will include supplemental information like definitions, while the instruction text will include reminders such as probing techniques specific to the question.

iii. The flow of the CAPI application and use of consistency checks are critical considerations that will impact data quality. The CAPI application must be able to enforce the sequencing of questions, so that they cannot be asked out of order. The CAPI application also needs the ability to skip questions that are no longer relevant to the current interview based on previously asked questions. Consistency checks are helpful for identifying grossly inconsistent data while the enumerator is in the field, but should not prohibit the enumerator from entering data in a timely and unobtrusive manner.

b) Computer Assisted Web Interviewing (CAWI)

4.47. When designing a census using Computer-Assisted Web Interviewing (CAWI)[136] it is important to consider factors that will maximize data quality, efficiency for respondents, user satisfaction, and census branding experience. Additionally, it will be equally important to consider factors that minimize respondent burden, measurement error, and design inconsistencies across data collection.

4.48. Improving the quality of data collected in CAWI involves meticulous questionnaire design, robust pre-testing, and proactive data validation to ensure clear, accurate, and reliable responses from participants. Minimising the time taken to complete the questionnaire means optimizing the interface with the respondent and providing user-friendly features to streamline the completion process. Elevating user satisfaction in CAWI can be achieved by prioritizing intuitive design and ensuring transparent data handling practices to enhance respondents' overall experience and trust in the survey process. An additional priority is strengthening the census branding experience in CAWI to align the survey platform with the official branding and messaging of the census. This alignment communicates a consistent and trustworthy identity to respondents.

4.49. Decreasing measurement error in CAWI involves using clear and unambiguous questions, employing thorough pre-testing and validation checks, and ensuring data quality control measures are in place to enhance the accuracy of collected data. To lessen design inconsistencies across data collection in CAWI, consideration should be given to maintaining a standardized survey template, providing detailed

[136] Alternatively described as Computer-Assisted Self-Interview (CASI).

design guidelines, and conducting regular reviews of question order, response scales, and visual elements.

4.50. Reducing respondent burden in CAWI involves designing concise questionnaires that only request essential information and employing skip logic to personalize the questionnaire.

5. Field operations and pilot testing

4.51. This section discusses the quality aspects of "field operations and pilot testing". For further details and a description of the processes involved in field operations and pilot testing, please refer to Part Three, Chapter VII and Part Three, Chapter V.

a) Quality considerations during field operations (see Part Three, Chapter VII)

4.52. The quality management process continues throughout the design of the census field operations. These are tested as far as possible in conjunction with questionnaire designs testing. The key internal client of field operations is processing. However, field operations can also impinge on other areas, such as dissemination and classification and subject matter areas where certain concepts, such as what constitutes a dwelling, are implemented during the field operations phase. Field operations depend on the quality of operations that precede it, such as:

i. Demarcation of enumeration areas;
ii. Map production and/or actualization;
iii. Form printing, where a sample of forms is rigorously tested for adherence to standards.
iv. Development of applications to ensure data quality, such as:
    - *a collection application* with automatic data storage (in the device and online), the necessary control specifications (consistency and completeness checks, automatic skips and alert messages), and GPS coordinates of the household dwelling to check the enumerator's presence;
    - *a control application* for the enumeration supervisor;
    - *a data transfer application* at central level, with all the necessary security features (data encryption, data bandwidth fluidity) to avoid data loss in the field;
    - *a web application/dashboard* for monitoring progress and data quality via real-time indicators for management teams in the field and at central level;
v. Installation of a real-time data back-up server at central level and a second replication server at another site to ensure data backup and collection continuity in the event of malfunction;
vi. Installation of a call-centre with a toll-free number and a team of teleoperators to answer people's questions about the census in real-time, or the implementation of a website with support chat. Teleoperators can collect the various problems encountered by field agents. These problems can be reviewed and analysed to identify improvements to census operations.

4.53. All systems supporting data collection must be thoroughly tested and piloted before collection. This is especially critical when collection technology is used, such as handheld devices or online forms for self-response. All data quality benefits of using such technology could be compromised if problems arise during enumeration.

4.54. Quality monitoring should be established for each of these components and mechanisms put in place to ensure that the outcomes of the monitoring are used to improve processes. It is more difficult to

implement quality improvements during actual enumeration owing to the very tight time constraints. However, this can be achieved by:

i. Clearly establishing the aims of the field operations phase;
ii. Applying thoroughly documented procedures;
iii. Ensuring that the enumerators understand their role through appropriate training;
iv. Providing opportunities for field staff to be observed operating on the job so that feedback can be given;
v. Training must consider the monitoring and evaluation of quality indicators during enumeration.
vi. Establishing communication and feedback loops with the general public through helplines, online forums, social media, etc., so that problems in the field can be detected and corrected.
vii. Establishing communication with supervisory field staff through digital field management systems.

4.55. However, it has to be acknowledged that during the actual execution of the enumeration this approach tends to identify "problem enumerators" rather than systemic or process errors. This means that an evaluation following the data collection process is vital. Such an evaluation should attempt to capture the experiences and suggestions of a range of enumerators and other field staff so that improvements can be made to the subsequent census.

4.56. A general overview of the quality of the enumeration can be obtained through:

i. Use of techniques such as post-enumeration surveys to gauge the level of under-enumeration or over-enumeration of people, households or dwellings;
ii. Overall response from the target population or level of non-response at the individual question level;
iii. Consistency and results of the imputation and editing process at the questionnaire and variable level.
iv. Feedback from field staff;
v. Measures of the quality of any coding undertaken by field staff;
vi. Mechanisms that may be in place to handle queries from the public;
vii. Analysis of administrative data used to assist the field operation and historical trends.

4.57. In certain situations, the enumeration schedule dates will be extended to ensure a complete and accurate census. Using quality monitoring processes and mechanisms that are already in place, the potential impacts on data quality should be monitored carefully during an extended enumeration period. Shifts in schedule dates could potentially impact quality in a number of ways:

i. **Transient populations:** The inability to enumerate people in one place before they move to another place, introduces the risk of missing them in the census altogether or counting them in both places, based on the timing of their move. A longer data collection period increases the risk that people move.
ii. **Recall bias:** Any timing shift for data collection activities increases the risk related to recall bias. For example, a longer enumeration period increases the risk that respondents would misremember who lived at the housing unit on the census reference day.
iii. **Vacant Units:** Any timing shift in data collection activities increases the risk of enumeration errors because of change of occupancy status for housing units, be it vacant, occupied or destroyed.
iv. **Training:** Shifts in training to accommodate schedule changes increase the risk that census workers would be underprepared to do their work and that could result in data quality issues.

Changes in schedule may increase the need to develop further training to avoid potential data quality issues for field staff.

2.281 The effectiveness of the public communication strategy may be assessed by the amount of press coverage (positive and negative) of the census and follow-up surveys specifically designed to gauge public reaction to particular advertising.

b) Pilot testing (see Part Three, Chapter V)

4.58. A pilot census is a comprehensive trial or rehearsal that tests all census procedures including personnel systems and procedures and plays a crucial role in census quality assurance. It provides an opportunity to fine-tune procedures, identify and rectify potential issues, and ultimately ensure the accuracy and reliability of the census data. A pilot census allows census organizers to (1) test all procedures, methodologies, and technologies that will be used in the actual census, including data collection methods, questionnaire design, interview techniques, and data processing systems; (2) identify any potential problems or challenges that may arise during the actual census such as logistical issues, errors in questionnaire wording, or difficulties in data collection in certain areas; (3) assess the effectiveness of the training of census staff and enumerators, and their familiarity with expected roles and responsibilities; (4) testing of sampling methods and procedures as may be relevant.

4.59. Also, the data collected during the pilot census can be compared against known population statistics or other sources of demographic data to validate the accuracy of the census methods. Any discrepancies or errors can then be addressed and corrected before the actual census takes place. Overall, a properly executed pilot census contributes to the improvement of the accuracy and consistency of data collection during the actual census.

4.60. In the case of multimode data collection (field, Internet, and telephone for example), it's required that all applications are interoperable. Controls must be established for integrating information according to the collection mode to ensure the integrity and uniqueness of the data, allowing for proper tracking of the coverage.

6. Processing

4.61. This section discusses the quality aspects of “processing”. For further details and a description of the processes involved in processing, please refer to Part Three, Chapter VIII.

4.62. Within the country’s statistical agency, the key recipients of the processed census data are the teams responsible for: the dissemination of the census results; maintaining standard classifications; and preparing specialist subject matter documentation. Effective dissemination requires the processing team to produce data in an agreed format and compiled to agreed quality standards. This is necessary so that the data can be used in dissemination systems. Additional material on the quality control processes that are necessary when integrating administrative data from different sources is covered in section E (Process stage) and further discussed in the UNECE Guidelines on Assessing the Quality of Administrative Sources for Use in Censuses. [137]

[137] Also see chapter 6 (Process Stage) of The UNECE *Guidelines on Assessing the Quality of Administrative Sources for Use in Censuses.*

4.63. Since the census is part of an overall national statistical system, data from the census are likely to be used in conjunction with data from other sources. Thus, the classification and subject matter specialist teams that are responsible for those other data collections, need to be satisfied that the coding, editing and other data transformation processes are conceptually sound and deliver data of acceptable quality.

4.64. Extensive testing of processing systems should be undertaken in advance of the census. Coding processes and training packages need to be prepared and tested utilising the type of staff likely to be involved in the census itself. The processing phase gives the fullest scope for the use of quality improvement techniques, as many of the processes in this phase are repetitive and take a reasonable amount of time. It is vital that structures are put in place not only to monitor quality but also to involve processing staff in the identification of problems with quality and in proposing solutions.

4.65. Processes such as editing and imputation may redress some of the inconsistencies and omissions within the data, thus improving data quality.

4.66. However, if a strategy of adjusting census data items after collection to eliminate identified errors and inconsistencies is adopted, then the following paragraphs provide examples of how ***data cleansing***, ***auditing***, and ***output consistency checks*** that can be implemented. Information about the nature of the checks undertaken and the proportion of records subjected to editing should be published to facilitate interpretability and to maintain user confidence in census results and the methodologies adopted for such adjustments.

4.67. ***Data cleansing*** is an important step in ensuring the quality and reliability of the census dataset. This process involves several key components, including:

i. a data validity review to confirm that all entered responses are accurate. This review ensures that numeric fields contain only numerical data, and text fields do not contain numerical data or other inappropriate information.
ii. methods for handling missing data to ensure that all variables pertaining to the characteristics of the census units are complete. In cases where data is missing, appropriate imputation techniques can be used to fill in the gaps and maintain data completeness (see section on data editing).
iii. the correction of values outside electronic validation rules. It is important to check that any values outside the scope of the electronic validation rules have been rectified. These values may arise due to defects in the electronic survey during data collection or errors made in the survey process.
iv. The detection of duplicates and a set of rules to support a decision for deleting duplicates.

4.68. ***Auditing*** is an important process that involves the review and validation of various aspects of the data to ensure its quality and reliability. The key components of data auditing include:

i. validation of the coding during the conversion of textual data into numeric codes.
ii. validation of the data collected to ensure that they adhere to international standards and classifications to ensure consistency and accuracy in data representation.
iii. range checks to ensure that all answer values of variables are examined to confirm that they fall within predefined limits or ranges. Range checking can identify any outliers or values that may be outside of the expected bounds.
iv. applying validation rules to review the logical relationships and correlations between variables.
v. dealing with extreme values by applying various reviews and audits to ensure that any illogical values have been addressed and corrected. This involves deciding on the treatment of unrealistic

or unreasonable values and making a determination if values should be edited or replaced with more realistic data points.

vi. verification that all targeted variables and geographic distributions have been thoroughly and comprehensively covered during the census. This means ensuring that all relevant data elements have been collected and included in the dataset to provide a comprehensive and accurate representation of the population or sample.

4.69. ***Output Consistency*** and maintaining the consistency of the main census outputs is essential to ensure the reliability of data. Maintaining consistency in outputs is vital for making informed decisions, conducting meaningful analyses, and ensuring that the data accurately represents the real-world context it is intended to describe. It helps build trust in the data and usability for various applications. This involves considering various aspects, including:

i. *Historical Consistency* to verify that the main outputs align with historical data and trends. Data should be consistent over time to ensure that there are no significant deviations or irregularities that might raise questions about data quality.

ii. *Geographical Consistency* to verify that the main outputs align with location and geographic data and trends data should be consistent and comparable.

iii. *Consistency with Other Sources* in order to cross-reference the initial main outputs with data from other reputable sources so that data aligns with external, and other internal, sources and any disparities are investigated and explained.

4.70. To ensure data accuracy it is important to check and cross-check the collected data with backend data. This involves the following steps:

i. *Establishing data validation procedures.* Define specific validation rules or checks for each type of data collected. For example, we might validate numerical values within a certain range, verify consistency between related fields, or check for missing or incomplete responses.

ii. *Match collected data with backend data.* Compare the collected data with the corresponding data stored in the backend system. This can involve matching participant identifiers, survey responses, or other relevant data points.

iii. *Conduct data reconciliation.* Identify any discrepancies between the collected data and the backend data. Investigate the reasons behind the discrepancies and resolve any inconsistencies or errors.

iv. *Perform data cleaning.* Remove any duplicate, invalid, or inconsistent data from the dataset. This can involve correcting errors, standardizing formats, or removing outliers.

v. *Document data validation and reconciliation process.* Keep a clear record of the steps taken to validate and reconcile the data. This documentation helps ensure transparency and reproducibility of the data analysis process.

vi. *Define data integration processes.* Determine how the collected data will be integrated with the backend data. This can involve data mapping, transformation, or merging procedures to ensure compatibility and consistency between the datasets.

vii. *Automate data validation and reconciliation.* Whenever possible, automate the data validation and reconciliation processes to minimize human error and increase efficiency. This can be achieved through scripting or programming techniques that compare and reconcile the data automatically.

viii. *Perform data analysis and verification.* Once the data has been validated and reconciled, perform thorough data analysis to identify patterns, trends, and insights. Additionally, verify the results against the backend data and ensure consistency between the analysed data and the original source.

ix. *Conduct data reconciliation in real-time.* If possible, perform data reconciliation in real-time or near real-time. This allows for immediate identification and resolution of any discrepancies between the collected data and the backend data. Real-time reconciliation ensures that any errors or inconsistencies are addressed promptly, reducing the risk of relying on inaccurate or outdated data.

By following these practices, the accuracy and reliability of the collected data can be enhanced. These data validation steps also improve confidence in the alignment of the data collected with the other sources of administrative data.

7. Analysis

4.71. This section discusses the quality aspects of "analysis". For further details and a description of the processes involved in analysis, please refer to Part Three, Chapter VIII.

4.72. Processing and analysis can be iterative and parallel. Therefore, some of the quality considerations discussed in this section are also contained in the processing section. Analysis can provide a deeper understanding of the data, which could indicate the need for additional processing. In the analysis step, statistics are produced, validated, interpreted, and made ready for dissemination. Also, during tabulation, statistical disclosure control (SDC) methods are applied to ensure data privacy, and statistical content to support the interpretation of census outputs are prepared.

4.73. Quality assurance should be applied for each of these activities to guarantee interpretability and clarity of the statistical outputs, maintain accuracy, ensure comparability, address coherence with outside sources, guarantee relevance, and ensure privacy.

4.74. The interpretability of the statistical outputs is assessed by checking the existence of metadata, commentaries, technical notes and any other necessary supporting statistical information. In order to interpret and explain census statistics, it is necessary to view them from multiple perspectives and understand the extent to which these statistics meet the needs of the intended audience by conducting in-depth statistical analyses, including consistency and comparability analysis.

4.75. An assessment of statistical confidentiality and data privacy should consider three quality attributes, (i) accuracy, (ii) relevance and (iii) clarity.

i. Accuracy is assessed by analyzing the impact of SDC methods on the quality of the output and by verifying that the statistical disclosure control does not affect significant relationships in the data. SDC methods must inevitably, by intention, affect the accuracy of the data, but not to an unacceptable degree.
ii. Relevance is assessed by analyzing how well the output meets user needs after applying statistical disclosure control methods.
iii. Clarity is assessed by describing the statistical disclosure control method(s) applied and any restrictions on the use of the dataset.

4.76. Methods for Statistical Disclosure Control (SDC) and the importance of confidentiality are discussed in detail in the dissemination section (paras 3.263 to 3.273). The impact of SDC methods on data quality should be considered when determining the methods to use for census outputs. Minimizing information loss due to suppression, perturbation, or other SDC techniques should be prioritized, while

still adhering to the privacy and confidentiality principles outlined in the Fundamental Principles of Official Statistics.

4.77. After assessing the impact of SDC on data quality and determining the appropriate techniques to apply, pre-release discussions and review with appropriate internal subject matter experts with in-depth knowledge of specific statistical domains should be undertaken. In addition, consistency checks should be implemented to ensure relevance and confirm that census outputs are fit for purpose and ready for dissemination. One of the main quality assurance measures is having a thorough review system in place for data products to make sure that the data or indicators being released are of high quality.

8. Dissemination

4.78. This section discusses the quality aspects of "dissemination". For further details and a description of the processes involved in dissemination, please refer to Part Three, Chapter IX.

4.79. Census dissemination can easily be overlooked in the chain of providing a quality outcome for the census as management attention is diverted to the far more costly and risk prone data collection and processing operations. The dissemination area is responsible for the timely delivery of products and services to the census data users. Therefore, insufficient planning and resources for this phase can have the effect of delaying the release of the data and thus adversely impacting on the dimensions of timeliness and punctuality thereby compromising the overall achievement of the census objectives. The dissemination phase should also be regarded as an ongoing process that will service the needs of users over a long period of time.

4.80. Management of quality in census dissemination is driven by concerns to (a) deliver relevant products and services; (b) maintain accuracy of the data; and (c) ensure timeliness and punctuality of data release within agreed cost constraints. Effective management of quality in census dissemination should also aim to prevent miscommunication and misunderstanding.

4.81. The first of these objectives is to provide relevant products and services. This can only be done by reviewing the experiences of the previous census products and services and by user consultation processes with both current and potential users of census data.

4.82. The second objective is to ensure that the data supplied from the processing system are accurately transformed into output products. A quality assurance strategy to ensure that data tabulations and transformations are carried out accurately needs to be documented and followed. Key aspects of a quality assurance strategy are outlined in the previous section on analysis. The quality circle approach to these processes needs to be applied and any gaps identified and corrected through extensive testing prior to the census and ongoing process improvement during the analysis phase before the dissemination of census outputs.

4.83. The third objective is the timely and predictable release of data from the census. While this is the responsibility of all phases of the census programme, the role of dissemination is crucial. The dissemination team needs to be realistic about release dates and ensure that these are communicated to users early so as to manage their expectations. The involvement of staff actually responsible for the dissemination phase in devising these dates is recommended where this is possible. Dissemination

systems and processes need to be available, documented and tested prior to the release of data from the processing and analysis phases.

4.84. A release calendar needs to be prepared to keep the user community informed about the likely month of release so that their use of data can be planned in advance. A mechanism to provide metadata on census indicators and the geographic level at which these are made available needs serious consideration. Every country should assess the requirements and put in place a dedicated team to assist data users. The services of call centres may be used if the number of data seekers cannot be handled in-house.

9. Evaluation

4.85. This section discusses the quality aspects of "evaluation". For further details and a description of the processes involved in evaluation, please refer to Chapters II and III of Part Four.

4.86. Evaluation of the overall census operation is vital for identifying strengths and weaknesses of each of the census phases, including planning, data collection, data processing and dissemination, and also for the purpose of analysing the quality of census statistics. With the quality assurance and improvement programme, the main objective is to ensure that quality assessment is consistently incorporated in all phases of the population and housing census, focusing on efforts in controlling the occurrence of errors and taking actions to ensure the highest quality of both the processes and their outcomes. Yet, errors appear to be inevitable in such a complex undertaking – consequently, there needs to be a mechanism put in place to determine the deficiencies and their quantitative impact on the census results.

4.87. Census evaluation with all dimensions of quality requires a comprehensive evaluation programme for assessing and documenting the outcomes of each process using appropriate and customized methodologies. Methodologies for evaluation should be planned well in advance, in the planning phase of the census. It should be noted that this is a continuous process implemented from the planning to the end of census operations. It is also appropriate to consider it as being the first step in the subsequent census cycle. Similarly, the evaluation of one process within a census cycle could be the first stage in the next process of the same census cycle.

4.88. Evaluation of the accuracy of the census data should also be undertaken, to the extent possible, by conducting a post-enumeration survey for measuring coverage and content errors, by comparing the census results with similar data from other sources and by applying demographic analysis. Such other sources may include surveys and administrative records (in particular, a population register) relating to a similar time frame, tabulations when a population register is available, and previous census results. The purposes of evaluating the accuracy of the data are to inform users of the quality of the current census data and to assist in future improvements. Future improvement may be achieved by: (a) improving processes; and (b) establishing performance benchmarks against which the quality of the data from subsequent censuses can be measured.

4.89. Evaluation of data accuracy may be undertaken in two stages. Preliminary evaluation will enable the identification of any problem areas that have not been previously detected through the quality management processes in earlier phases of the census. More extensive evaluation should be undertaken on data items where problems have been identified or where the data has been collected through new questions or processes have been attempted.

4.90. The census evaluation programme should be undertaken by subject specialists according to the agreed goals and methodologies covering all dimensions of quality. The following are some examples:

i. Identification of the deficiencies and achievements in data capture, coding and editing (through mechanisms developed for checking the quality of process and the work of personnel);
ii. Relevance of census data to user needs and satisfaction of users with dissemination tools and products (based on information collected through user consultation);
iii. Achievements and difficulties in the use of new technologies and methodologies and identification of possible improvements for the next census;

4.91. The results of evaluations of census operations for both operational aspects and the quality of data should be made available to the stakeholders promptly as they become available. A commitment to transparency will ensure confidence in census results and provide a basis for improvements for the next census cycle.

4.92. Depending on the national context, census results by default may be presented without adjustments. But for analytical and planning purposes, results of the census evaluation should be taken into account in preparing intercensal and post-censual population estimates in order to correct deficiencies in the data collection phase.

10. Documentation

4.93. This section discusses the quality aspects of "documentation". For further details and a description of the processes involved in documentation, please refer to Part Three, Chapter X.

4.94. Through all census phases discussed above, documentation is extremely important for improving the quality of future censuses. But that is not enough. It is recommended that once census data are evaluated and disseminated, NSOs should develop various reports to retain institutional knowledge and memory. Administrative and methodological reports should be written to document planning and implementation phases of the data collection, while a procedural history report should cover the product and dissemination phases.[138] It is also helpful to document the SDC measures that have been applied as these too will affect the quality of the data. However, specific details of how SDC methods have been applied should never be disclosed.

F. Measuring the quality of data derived from administrative sources

4.95. There has been a clear trend towards increased use of administrative data in censuses, in line with a more generalized trend towards increased use of administrative data in all statistical domains. Their use can provide more frequent and timely statistics about the population, improvements in accuracy and reliability, and significant reductions in costs and respondent burden. Nevertheless, there are many key quality considerations that must be assessed and evaluated before incorporating the data from any administrative source into a census.

[138] https://www2.census.gov/programs-surveys/international-programs/events/stic/planning/improve-future-data-quality.pdf

4.96. The quality of data used in the process of compiling census statistics strongly affects the quality of the statistical output produced. Thus, the quality of data from administrative registers and other administrative sources is a key element that should be considered in the decision-making process on the use of administrative data in the production of statistics. It is necessary to prepare and implement a standard method of assessing the quality of administrative data as potential sources for the census. The assessment of the quality of administrative data may be used to evaluate the usability of the data source. To this end a set of indicators should be developed by which each source may be analysed. Based on these indicators, the NSO can decide whether or not to use a specific source.

4.97. In general terms, the NSO should identify the administrative sources that may be relevant for the purposes of the census based on the users' requirements for data. It is important to set out what the expected or required outcomes of using the source would be. This could include improvements to the efficiency of the census operation in terms of reductions in cost and respondent burden, improvements to the quality of the census, or the delivery of new or enhanced census outputs.

4.98. The following key quality considerations should inform decisions about the use of administrative data in a census:

i. Administrative data have, in general, not been collected for the purpose of a census. Therefore, the data sources may have adopted different concepts, classifications and definitions from those required by the census, they may refer to different reference periods, be subject to lags in the updating of information, and may have limited coverage of the census population. Additionally, the accuracy and completeness of the data will be highly dependent on the importance of the data to the data supplier's administrative function. The administrative sources may also be subject to changes over time and inconsistencies in the way the data are collected across segments of the population.

ii. The use of administrative data by the National Statistical Offices for purposes other than those for which the data were originally collected raises privacy, security, and legal concerns. An assessment of public acceptability of the use of administrative data for statistical purposes must be performed on a credible legal basis and must be publicly acceptable. It is important that the public and data users alike understand how and why administrative data are being used in the census. The NSO should be transparent about their use and provide a clear justification of the benefits against any risks and costs. This can be achieved through good communication, including the publication of the procedures and policies that support the effective use and protection of the data.

iii. The NSO has limited control over the way the administrative data are collected and processed. Administrative data can be subject to changes in population coverage and population measures over time due, for example, to legal, policy, procedural, or system changes affecting the data and/or their delivery. This limited control and the associated risks to census processes can be mitigated by working with data suppliers on potential or planned changes, by being flexible and responsive to change, and by reducing reliance on any single data source or item where possible.

4.99. Assessing the quality of administrative sources requires mapping the errors which may occur before and after the data is supplied to NSOs and determining how these errors can be mitigated. At the same time, the quality assessment should not be limited to assessing the error. More generally, it should include the assessment of the impact of the integration of administrative data into the census in terms of the extent to which it adds error or uncertainty to the outputs, vis-à-vis of the advantages gained from integration.

4.100. A framework for assessing the quality of administrative sources for use in censuses can be built in stages: the Source stage, the Data stage, the Process stage, and the Output stage. The first two stages mainly deal with the assessment of the quality of single sources, while the last two stages are concerned with the quality of transformed and/or integrated sources. For each stage, the key quality dimensions should be identified as well as the respective tools and indicators for the quality assessment[139,140]. A brief description of each stage is included below.

i. ***Source Stage.*** The quality of administrative data sources to be used in the census production should be assessed both at first acquisition and when they are regularly re-supplied to the NSO. This is because the characteristics of any data set that has previously been supplied might differ due to changes in concepts, classifications, collection methods, and so forth.

ii. ***Input Stage[141].*** The quality assessment at the Input stage refers to the quality assessment of raw administrative data as supplied to the NSO, with reference to the expectations and requirements established through the metadata-based assessment at the Source Stage. The quality of administrative data at the Data Stage is to be assessed against several dimensions including readability and validity, accuracy and reliability, timeliness and punctuality, and linkage ability.

iii. ***Process Stage.*** Once administrative data are received and the quality assessed, the data will require processing to be usable in the census. The administrative data will need to be integrated into the census design and any quality issues will need to be addressed. Some of the most common processes required for using administrative data in the census are: linking records, assessing coverage errors, resolving inconsistencies from different sources, editing and imputation.

iv. ***Output Stage.*** Measuring output quality cannot be reduced to only the estimation of uncertainty (the accuracy of estimates); rather, it should include an assessment across all other quality output dimensions. The introduction of administrative data will likely lead to gains in some dimensions and losses in others. Achieving the right balance across the quality dimensions is the key to best meeting user needs. Updated and new administrative data can be used to improve census statistics by revising previous estimates. However, this needs to be balanced against user needs with respect to revisions.

4.101. The results of the quality assessment and assurance should be recorded and published throughout the census production process. This will support an ongoing dialogue between producers and users of census data and help determine whether the right balance has been achieved across the multiple dimensions of quality.

II. Evaluation of the results

4.102. As noted earlier in Chapter I of Part Four, a census evaluation programme should be developed as part of the overall census programme and integrated with other census activities. The scope and objectives of the evaluation programme should be decided well in advance to determine early enough the adequate resources (both financial and human) needed for the evaluation programme. It is important to establish a dedicated team responsible for the planning, organization and implementation of the evaluation programme. The cost of evaluation should be covered in the census budget as a separate item.

[139] See UNSD Handbook on Registers-based Population and Housing Censuses (2022).

[140] See UNECE, Guidelines for assessing the quality of administrative sources for use in censuses, 2021.

[141] Also referred to as “data stage” in the UNECE Guidelines for assessing the quality of administrative sources for use in censuses (2021).

4.103. The evaluation of the activities of the census implementation program should cover the following elements:

i. The timing of the census operations.
ii. The census date.
iii. The media campaign for the census and the means used to deliver the media messages to the population.
iv. The effectiveness of the media campaign in raising awareness about the census and encouraging participation. This can include analysing the reach and impact of different communication channels such as television, radio, newspapers, social media, and community outreach programs.
v. The level of understanding and knowledge among the population about the purpose and importance of the census. This can be measured through surveys, focus groups, or interviews to assess the public's comprehension of the census process and its significance in decision-making and resource allocation.
vi. The level of cooperation and participation from the population. This can be assessed by analysing response rates, identifying any barriers that hindered participation, and understanding the reasons behind non-participation or incomplete responses.

A. Purpose of census evaluation

4.104. The quality of population and housing census data is very important for many reasons, including building public trust in and understanding of the national statistical system. The purpose of census evaluation is to provide users with a measure of the level of accuracy and confidence when utilizing the data, and to explain any errors in the census result. It is therefore important to choose an appropriate way of sending out these messages to the right group of stakeholders.

4.105. The evaluation methods discussed here are those that apply to field-based censuses. To some extent they also apply to register-based censuses and other census methodologies, but these also present their own particular challenges and solutions[142].

4.106. It is universally accepted that a population census is not perfect and that errors can and do occur at all stages of the census operation, but these errors should be measured. Errors in the census results are classified into two general categories – coverage errors and content errors. *Coverage errors* are the errors that arise due to omissions or duplications of any of the enumeration units – persons, households or housing units – in the census enumeration. The sources of coverage error include incomplete or inaccurate maps or lists of enumeration areas, failure on the part of enumerators to canvass all the units in their assignment areas, duplicate counting for persons who have two or more places of residence, persons who for one reason or another do not allow themselves to be enumerated, erroneous treatment of certain categories of persons such as visitors or non-resident aliens, and loss or destruction of census records after enumeration. Content errors are errors that arise in the incorrect reporting or recording of the characteristics of persons, households and housing units enumerated in the census. *Content errors* may be caused by several factors, including poorly phrased questions or instructions, or enumerator errors in phrasing the census questions; inability or misunderstanding on the part of respondents in

[142] See, for example, *Census Quality Evaluation: Considerations from an international perspective*, note produced by the UNECE Secretariat for the 11th Joint UNECE/Eurostat Meeting on Population and Housing Censuses, Geneva, 13–15 May 2008, http://www.unece.org/stats/documents/2008.05.census.html.

respect of answering specific items; deliberate misreporting; errors due to proxy response; and coding or data entry mistakes.

4.107. Many countries that conduct a field-based census have recognized the need to evaluate the overall quality of their census results and have employed various methods for measuring census coverage as well as certain types of content error. Comprehensive evaluation should also include an assessment of the success of all aspects of the census operation, in each of its phases, including activities such as designing the geographic frame, the census publicity campaign, data collection, data processing, data dissemination and data utilization. Countries should ensure, therefore, that their overall census evaluation effort addresses the census process, as well as the results. The present section is devoted to evaluation of the results. However, the section on the quality improvement and the census (paragraphs 4.28-4.94) provides further recommendations relating to controlling and assessing the quality of census operations.

4.108. Evaluation efforts focused on census results should generally be designed to serve the following objectives: firstly, to provide users with some measures of the quality of census data to help them interpret the results; secondly, to identify as far as is practicable the types and sources of error in order to assist the planning of future censuses; and thirdly, to serve as a basis for constructing a best estimate of census aggregates, such as the total population, or to provide census results adjusted to take into account identified errors at national or subnational levels if some errors, such as coverage error, are substantial and the validity of census results would otherwise be questionable.

4.109. As the decision to adjust census figures is sensitive, it is bound to be decided at the highest levels of the government bureaucracy. There is also some critical statistical considerations that should be very carefully weighed ~~in~~. Consideration must be given to what geographic domains the adjustment would cover, knowing that such adjustments have an effect on demographic distributions. Transparency in the methodology used for such adjustment is critical to maintain trust with all stakeholders but particularly users. In cases where census results are presented without adjustments, the results of the census evaluation should nevertheless be taken into account in the preparation of post-census population estimates and for analytical and planning purposes.

4.110. The final publication of census results should include an estimate of coverage error, together with a full indication of the methods used for evaluating the completeness of the data. The publication should also provide users with some cautions or important notes about the results, in addition to some guidance on how they might use the evaluation results. It is also desirable to provide, as far as possible, an evaluation of the quality of the information on each topic and of the effects of the editing and imputation procedures used.

4.111. The range and quality of the editing of incorrect or inconsistent data and the imputation of missing data items are greatly enhanced by the use of computer edit programs that permit inter-record checks (for example, the replacement of missing values based on one or more items on the basis of reported information for other persons or items in similar households). If any imputation is made, the topics affected, the methods used, and the number of cases affected should be documented and clearly described in the census evaluation report.

4.112. The results of the evaluation of census results should be made available to users with a measure of the quality to help them interpret the results.

4.113. As discussed in the following subsection, a number of methods exist for carrying out a census evaluation. In practice, many countries use a combination of such methods in order to fully serve these objectives.

B. Methods of census evaluation

4.114. The choice of evaluation methods to be used depends upon the evaluation objectives. These, in turn, depend on national census experience in terms of past and anticipated errors, user and public concerns, and the financial and technical resources available for evaluation. The decision needs to be made whether to measure coverage error, content error or a combination of the two. In addition, both gross and net error must be considered in developing the overall evaluation plan. Gross coverage error in a census is defined as the total of all persons omitted, duplicated or erroneously enumerated. Net coverage error takes into account the underestimates due to omissions and the overestimates due to duplications and erroneous inclusions. When omissions exceed the sum of duplications and erroneous inclusions, as is usually the case in most countries, a net undercount is said to exist; otherwise, a net overcount results. Similarly, both gross and net content errors have to be considered in the evaluation design.

4.115. Numerous methods are available to estimate the coverage and content error of censuses. Methods for evaluating the quality of censuses can be divided according to their complexity, the type of measures used, or data sources considered in the evaluation process. One of the most commonly used classifications is based on what sources of information are taken into account in the process of census quality assessment. According to this classification, there are three types of methods for assessing the quality of census data: 1) those based solely on data from the census itself (single source); 2) those in which census results are compared with data from other statistical surveys (multiple data sources); 3) those based on results from additionally conducted surveys. Each of these methods can be used to assess the size of coverage and content errors. These include simple techniques of quality assurance, such as internal consistency checks. Comparisons of results with other data sources, including previous censuses, current household surveys and administrative records, are also useful techniques. Such comparisons may be made in aggregate by comparing the overall estimates from two sources (net error only). Differences in conceptual frameworks and quality assessment for all these sources should be considered when using them as references to assess the coverage and content error of censuses. Alternatively, record checking, whereby individual census records are matched against alternative sources and specific items of information are checked for accuracy, may be used. Both gross and net errors can be estimated in record checks, which may involve field reconciliation of differences, a costly exercise that cannot be overlooked. An important but complicating factor in the use of record checks is the requirement of accurate matching. It is essential to plan carefully for this aspect, since the operation can be tedious and costly. It should be noted that record checks are best employed to study the coverage of certain segments of a population, such as children whose birth records are complete, since these checks are, by definition, limited to subpopulations with complete, accurate records.

4.116. Demographic analysis and post-enumeration surveys[68] are two very important methods for evaluating census data, and these are discussed in further detail in the following two subsections. However, there are many other methods for testing the validity of population counts, such as the Census Coverage Survey (CCS), Reverse Record Check (RRC), Dual System Estimation (DSE), adjustments for Residents Temporarily Overseas (RTO). These are basically post-census surveys, record checks and dual estimation methods to verify the completeness (coverage) of the census estimates.

C. Post-enumeration survey

4.117. The post-enumeration survey (PES) can be defined as the complete re-enumeration of a representative sample of the census population and matching each individual who is enumerated in the PES with information from the main enumeration. The objectives of the post-enumeration survey can be summed up as follows:

a. To assess the degree of coverage during census enumeration;
b. To examine the impacts of coverage deficiencies, if any, on the usefulness of the census data;
c. To examine the characteristics of persons who may have been missed during census enumeration.
d. To obtain information for the design of future censuses and surveys;

4.118. While a post-enumeration survey can be designed to provide a comprehensive evaluation of coverage and content error, especially when supplemented by, and integrated with, detailed demographic analysis of census quality, the methodology of a sound post-enumeration survey is complex, so that countries should accordingly weigh with care the demanding technical requirements and the costs of conducting a successful PES, and elaborate a clear statement of its objectives, before deciding to undertake such a survey.[69] Careful advance planning is crucial. To be valid, a PES has to function within a number of operational and statistical constraints. These include the requirement that the survey be carried out within a few months of the end of the census, changes in transitory conditions (such as employment status) and lapses in respondent recall do not hopelessly complicate the exercise.

4.119. Another basic property of PES design and execution involves matching and reconciliation. Matching the unit records from the PES with the corresponding census records is an operation whose performance must be of very high quality to ensure that inaccuracies in the PES itself do not effectively ruin the estimate of coverage error. Matching is especially difficult in countries where many surnames are identical or where individuals are known under more than one name, and where well-defined street addresses do not exist. Part of the matching operation usually involves a field visit to reconcile differences between the census and the PES with regard to either coverage or content. Reconciliation of course adds another dimension of cost and complexity, since it entails a second PES visit in the field.

4.120. Clearly defining the objectives of a PES is the first and most crucial step in planning the survey. The objectives might include estimation of coverage error at the national level; estimation of coverage error for major subnational domains or population subgroups, each with its own specified level of precision; and measurement of content error for specific census items. A secondary objective maybe to compare the level of correlation between the distributions of the most important characteristics from the census and the post-enumeration survey. For this purpose, some measures of consistency may be used (for instance the concordance correlation coefficient (CCC) or Kappa Cohen coefficient)[143].

4.121. As already noted, the design of a PES is complex and there are various alternatives. A number of references are available that set out highly detailed procedures for designing a PES and the conditions under which they may or should be considered. However, in general, when designing a post-enumeration survey, the following key considerations should be taken into account:

143 Statistics Poland (2023), National Population and Housing Census 2021. Assessment of the data quality, Warsaw.

i. The time between the census and the PES should be minimized to avoid as much recall error as possible and the impact of population changes (births, deaths and migration).
ii. The PES must be independent of the census. Interviewers must not have census information about the areas where they are working. When interviewers have knowledge of census responses, bias in introduced from the tendency to confirm only what the census recorded.
iii. To preserve the independence of the PES, its data collection and processing operations must be completely separate from those of the census.

4.122. The sample design for a PES must be based upon sound probability sampling methods taking account of the measurement objectives of the evaluation study. These usually include the need to estimate census coverage with a certain degree of reliability. In addition, the estimates of coverage may be wanted for geographic areas such as provinces or states and large cities, for urban–rural comparisons and so forth. Such requirements also greatly affect the sample design of a PES, as the necessary sample size is increased substantially when estimates of subnational coverage (or under-coverage) are wanted.

4.123. Sometimes a PES is designed to measure content error only, in which case it is sometimes referred to as a reinterview survey. The advantage of a well-designed reinterview survey is that the results are more accurate than those of the census insofar as the operation is much smaller and can be more effectively controlled. Estimates of relative response bias can be obtained from a reinterview survey, which (rather than the census) is generally taken as the standard in this area on the grounds that the survey, with its better-trained interviewers and more intensive survey procedures, yields superior results.

4.124. As part of the design of some PESs, a sample of the original census enumeration districts, blocks or areas is chosen and recanvassed for the survey. As regards methodology, this constitutes a useful reinterview technique for measuring content error, and such an element in the design is often put into practice because the matching operation between survey and census records is then dramatically simplified.

D. Demographic analysis for census evaluation

4.125. Demographic analysis offers a powerful methodology for evaluating the quality of a census and countries are encouraged to use it as part of their overall census evaluation methodology. A wide variety of demographic techniques have been developed and used, ranging from visual inspection of census data to comparative analysis of two census age distributions. A basic procedure for assessing census quality on age-sex data is graphical analysis of the population pyramid. Age heaping, or the tendency of respondents to report a particular ending digit, is a useful check of the quality of age reporting, as are sex ratios by age and certain summary indices of age-sex data, including the United Nations age-sex accuracy index, which extends age-sex ratio analysis by observing deviations of the observed age-sex ratios from the ones expected for each five-year age group and combining the results into a single score.[144] Other summary indices are Whipple's index and Myer's blended index, used for judging age heaping[145].

[144] See *Methods of Appraisal of Quality of Basic Data for Population Estimates: Manual II* (United Nations publication, Sales No. E.56.XIII.2).
[145] See Johnson, P. and others (2022). *Method protocol for the evaluation of census population data by age and sex*. United Nations, DESA, Population Division, Technical Paper No. 2022/05. https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesa_pd_2022_tp-methodprotocol.pdf

4.126. The methods mentioned above, while useful in providing an overall assessment of census quality, cannot differentiate the sources of census error in terms of the relative contributions from under-coverage (or over-coverage) or content error. Better information about coverage error, through demographic analysis, derives chiefly from comparative analysis of data from successive censuses, in which four methods are used. The four methods are:

i. derivation of an expected population estimate taking account of vital registers of births, deaths and net migrants between censuses, as compared with the latest census[146];
ii. population projections based on the results of the prior census plus data on fertility, mortality and migration from various sources and comparing the projected estimates with the new census results (cohort component method)[147];
iii. comparison of two consecutive census age distributions based on intercensal cohort survival rates and evaluation of the population age distribution for two or three census dates by year of birth of each cohort[148];
iv. and estimates of coverage correction factors using regression methods to make the age results from the two censuses mutually consistent (cohort survival regression method).

It should be noted that the first two methods would probably have to be restricted to evaluation studies of coverage at the national level, especially in countries that do not have good subnational data on internal and international migration.

4.127. A fifth method focuses on the evaluation and potential correction of under-enumeration of young children. The population at younger ages is often presenting a pattern of under-enumeration[149]. The observed population of children at age 0, 1, and/or 2 are lower than expected, in particular in the developing countries with high fertility, where a continuous increase in the population is expected at younger ages. But patterns of under-enumeration of young children is found in more statistically advanced countries as well[150]. Some of these under-enumeration problems also affect the post-enumerations surveys because the information provided on the household composition, and children in particular, is provided by the same proxy respondent.

4.128. Patterns of child under-enumeration can be corrected through demographic analysis by estimating the population at ages 0, 1-4 and 5-9 based on the smoothed adult female population from a census and estimates of fertility and mortality for the various periods preceding the census. In other

[146] Himes, C. L., & Clogg, C. C. (1992). An overview of demographic analysis as a method for evaluating census coverage in the United States. *Population index*, 587-607.
[147] See United Nations, Department of Economic and Social Affairs, Population Division (2022). *World Population Prospects 2022: Methodology of the United Nations population estimates and projections*. UN DESA/POP/2022/TR/NO. 4
[148] see GRPOP-YB workbook from the U.S. Census Bureau Population Analysis System (PAS) (https://www.census.gov/data/software/pas.html) and newer birth Cohorts () R implementation from Demo Tools (Riffe, 2021). Riffe, T. (2021). DemoTools: Standardize, Evaluate, and Adjust Demographic Data. https://rdrr.io/github/timriffe/DemoTools/
[149] Ewbank, D. C. (1981). *Age Misreporting and Age-Selective Underenumeration: Sources, Patterns and Consequences for Demographic Analysis*. Washington, D.C., National Academy Press, Committee on Population and Demography, Report 4.
[150] Special issue on undercount of children in censuses from the Statistical Journal of the IAOS, Volume 33, issue 2, https://content.iospress.com/journals/statistical-journal-of-the-iaos/33/2 and Tomkinson, J. (2023). The Omission of Young Children in the French Census: What Can Linked Census Data Reveal? | Cairn International Edition https://www.cairn-int.info/journal-population-2023-2-page-229.htm

words, the population below age 10 can be reconstructed based on estimates of past fertility and mortality (e.g., from either vital registration or recent surveys).[151]

4.129. Further evaluation and potential adjustments of age misreporting or exaggeration at older ages (e.g., 70 or 80 years and over)[152], especially for centenarians, require the use of additional administrative data sources (see section E below) or registration data (e.g., birth records) with more accurate recording of the true year of birth rather than self or proxy reporting of age, but the availability and accuracy of vital records for older birth cohorts can be limited due to low completeness of birth registration more than 50 years ago in many locations. Several demographic methods can also be used using additional assumptions on expected survival at older ages between censuses, the expected age distribution of the population at older ages based on the underlying mortality[153], or the extinct and almost extinct generations method upon availability and sufficient completeness of registered deaths[154] [155]

E. Comparison with administrative sources

4.130. Comparing census data with administrative data in a country or specific area can provide valuable insights into changes since the last census. By analysing administrative data (including data on development trends), it is possible to assess how well the new census data captures these changes and whether the census accurately reflects the characteristics of the underlying population.

4.131. Administrative data sources, such as population registers, records of birth, death, migration, etc. provide estimates of population changes. By comparing census counts with these data sources at aggregate level, census under- or over-coverages can be assessed. However, in these comparisons, the alignment of geographical boundaries, reference periods, and the target population covered by administrative data and the census should be considered. The coverage of census at the household and dwelling level can also be assessed if relevant administrative data on households (like family registration systems), building or property, etc. exist.

4.132. Availability of high-quality administrative data enables comparisons at the micro-level as well. Such comparisons can be facilitated if national unique identification numbers exist and have been included in the census questionnaire. This will not only assist with locating coverage errors but also enables evaluation of the content. Administrative data sources, such as healthcare records, school enrolments, and tax filings, social welfare programs, and employment record, etc. can be compared to census data to identify content errors related to demographic or socio-economic information. Additional information from utility services, such as electricity, water, sanitation, phone, etc. with sufficient

[151] see BASEPOP and BPSTRNG workbooks from the U.S. Census Bureau Population Analysis System (PAS) and newer basepop_five() R implementation from DemoTools (Riffe, 2021).
[152] Del Popolo, F. (2000). *Los problemas en la declaración de la edad de la población adulta mayor en los censos*. CEPAL.
[153] see OPAG workbook from the U.S. Census Bureau Population Analysis System (PAS) and newer R extension from DemoTools implementation (Riffe, 2021).
[154] Jdanov, D. A., Scholz, R. D., & Shkolnikov, V. M. (2005). Official population statistics and the Human Mortality Database estimates of populations aged 80+ in Germany and nine other European countries. *Demographic Research*, 13, 335-362.
[155] Terblanche W, Wilson T (2015) An Evaluation of Nearly-Extinct Cohort Methods for estimating the Very Elderly Populations of Australia and New Zealand. PLoS ONE 10(4): e0123692.

household coverage can provide complementary data sources useful to evaluate the occupancy of housing units.

4.133. The comparison with administrative data involves the following stages:

i. *Identify the relevant administrative data sources:* Administrative data can include records from government agencies, local authorities, educational institutions, healthcare systems, and other organizations.

ii. *Gather administrative data:* Collect the administrative data that aligns with the relevant timeframe and census topics. This data may include population records, building permits, school enrollment figures, healthcare statistics, or employment data.

iii. *Analyze administrative data trends:* Examine the administrative data to identify any significant changes or trends that have occurred. Look for relevant indicators such as population growth or decline, changes in employment sectors, infrastructure development, or changes in school enrollment.

iv. *Compare census data with administrative data:* Compare the findings from the analysis the administrative data with the census data and evaluate the level of alignment and consistency between these sources. Comparing census data with administrative data can help identify potential discrepancies and provide insights into the accuracy of the census data.

4.134. When comparing census data with administrative data it is important to: assess the quality and reliability of both the census data and administrative data; understand differences in data definitions, concepts and reference dates; consider spatial variations by analyzing the data at different geographic levels; conduct additional research to validate the findings; recognize that both census data and administrative have limitations; seek stakeholder perspectives; document findings; and communicate the results of the evaluation.

F. Adjusting census data

4.135. Achieving full coverage of the population is one of the major challenges in carrying out any census. Nonetheless some coverage error is unavoidable. Hence it is important to measure, analyse and report on coverage error.

4.136. Population under-coverage is regarded as one of the most significant sources of error affecting census data. This results from missing housing units, households, or persons during the census data collection phase. Over coverage, on the other hand, results from duplications (occurring when persons, households or housing units are enumerated more than once) or from erroneous inclusions (occurring where persons, households or housing units are enumerated in the census when they should not have been or when persons are enumerated in the wrong place[156]). These two sources of error (resulting in census counts under-estimating or over-estimating the true population counts) can also distort the distribution of population characteristics estimated from census data when the over-covered and missed persons don't have the same characteristics as those who were enumerated once.

4.137. Both gross and net error must be taken into account in developing the overall evaluation plan (see paragraph 4.114 for definition of gross coverage error and net coverage error). A net undercount

156 Of course, counting people in the wrong place may only seriously affect the coverage at the local area level; at higher levels of geography the effect will be minimised and at the national level the effect will be zero.

occurs when omissions outweigh duplications and erroneous inclusions, while a net overcount results from the opposite. Evaluation efforts should assess both gross and net content errors.

4.138. In some countries the evaluation of census coverage forms an integral part of the census quality assurance and improvement component of the quality management framework, and feeds directly into the published census results. In other countries, it is a separate exercise, used either to measure the census coverage error or to adjust the census results.

1. Adjustments for full field enumeration censuses

4.139. In the case of a census employing a full field-based enumeration, the most common problem is under-coverage[157]. The evaluation is usually done via an independent post-enumeration survey (PES) of a sample of enumerations areas (as described in Section C above) or via a Reverse Record Check methodology[158]. In case of a Reverse Record Check used to estimate the number of persons missed in the census, this estimate is combined with the estimate of the number of persons enumerated more than once from a Census Overcoverage Study in order to calculate net under-coverage[159].

4.140. In both cases the purpose is to measure the accuracy of the census. By independently surveying a sample of the population, the aim is to estimate the proportion of people and housing units potentially missed or counted erroneously in the census. On the basis of net coverage rates, adjustments may be made to population census results. Using results from a carefully designed and implemented PES, under- or over-counts can be converted into adjustment factors, i.e. a correction factor attached to each record, so that the census population is increased or decreased accordingly by such factors. The adjustment is usually done at the national level or for other sufficiently large geographic areas.

2. Adjustments for register-based censuses

4.141. In the case of census data derived mainly or wholly from administrative sources, the most common issue is instead over-coverage. If, for example, people have not officially declared their emigration, their records may be kept unchanged in administrative registers, and consequently be the cause of over-coverage in the census[160]. Over-coverage in registers might be corrected on the basis of ad hoc surveys focused on plausible unregistered emigration or, more generally, through an approach called

[157] Of the 1,758 census populations between 1950 and 2021 used for the production of the estimates of the *World Population Prospects 2022*, 310 had associated PES results available and the mean net error was -3.4%. United Nations, Department of Economic and Social Affairs, Population Division (2022). World Population Prospects 2022

[158] A frame containing all persons who should be enumerated in the current census is built up from the returns of the previous census and intercensal birth and immigration registrations. A random sample is selected from each source and each selected person is traced to his/her current census address (since the address obtained at the time of selection is usually out of date, a tracing operation must be undertaken to determine the address of each selected person as of Census day) census forms are then checked to determine whether or not the selected person was enumerated. See e.g. the Canadian Census Undercoverage Survey (CUS) methodology. (https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3902).

[159] Net coverage error takes into account the underestimates due to omissions and the overestimates due to duplications and erroneous inclusions. When omissions exceed the sum of duplications and erroneous inclusions, a net undercount is said to exist; otherwise, a net over-count results.

[160] UNECE: Guidelines for assessing the quality of administrative sources for use in censuses, United Nations, Geneva (2021), http://www.unece.org/fileadmin/DAM/stats/publications/2018/ECECESSTAT20184.pdf

'dependent interviewing' aimed at verifying administrative records in the field. However, not all countries can carry out dependent interviewing, for both ethical and privacy concerns[161].

4.142. Another possible means of avoiding coverage errors is to create a residency index on the basis of the records held in multiple registers in order to determine a so-called 'signs of life' (SoL) score. The approach is to define for all possible residents the 'sign of life' as a binary score (with a value 0 or 1) for each record in each register. Using these signs of life as explanatory variables, it is possible to build a model forecasting the size of under and over-coverage[162].

4.143. However, even though register-based censuses are usually concerned mainly with over-coverage, in principle they are also affected by under-coverage i.e., there might be a share of the population not included in any register. Another way to account for coverage errors in statistical registers is based on the correction of the population register through estimates obtained from two coverage samples for each area. More precisely, a field (area) sample of people living in the area on census day is used for estimating the register under count (the "U sample"), whereas a sample of people registered in the same area is used for estimating the register over count (the "O sample"). The U sample can also be used for collecting information on census variables not available in registers.

4.144. Some countries are using a combination of the methods above to investigate and refine methodologies for using of administrative sources as a means to understand coverage issues in the census and make adjustments to census data. This is an emerging area of statistical practice. The availability of administrative data sources is changing rapidly as more data is captured and stored electronically. Therefore, approaches differ from country-to-country and often need to be tailored based on national circumstances.

3. Adjustments for combined censuses

4.145. There are several options to check the coverage errors in the case of a combined census. One is to use the methodology commonly adopted for full field enumeration censuses (for the field enumeration component), by organizing a post-enumeration survey and estimating both over-coverage and under-coverage statistically (see section C). For example, the Census Coverage Survey (CCS) is a post-enumeration survey that many countries conduct after the decennial population census. The CCS plays a crucial role in providing accurate population or subpopulation estimates and helps to identify any biases or errors in the census data. The main goal of this survey is to measure the coverage of the census by identifying any undercounts or overcounts of the population. It is often based on dual-system estimation whereby data from the census and the coverage survey are matched to estimate the population size. After combining census and CCS results, the capture-recapture methodology is used to assess the numbers of households/individuals missed by the census enumeration.

161 A discussion of dual- and multisystem estimation methods for tackling coverage errors is provided in Brown J., C. Bycroft, D. Di Cecco, J. Elleouet, G. Powell, V. Račinskij, P. Smith, S.-M. Tam, T. Tuoto, and L.-C. Zhang (2020), Exploring developments in population size estimation, The Survey Statistician 82, 27-39.

162 UNECE: Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses, United Nations, New York and Geneva (2018), http://www.unece.org/fileadmin/DAM/stats/publications/2018/ECECESSTAT20184.pdf.

4.146. The 'signs of life' approach, comparing records from different registers, can then be adopted to provide information on the level of under-coverage and over-coverage of the register-based component of the census.

4. The decision to adjust, or not to adjust, census figures

4.147. After measuring the accuracy of the population count, some countries adjust census figures while others choose to leave census data unadjusted. In fact, many countries do not adjust census figures on the basis of the results of their PES, demographic analysis evaluation or, more generally, on the basis of coverage errors assessment. This is due to various reasons, e.g., adjusting census figures can be impossible due to national circumstances, legal reasons or too politically sensitive (due to the impact of adjustment on budget allocation or the size of electoral districts/constituencies), or coverage errors assessment is performed for quality reasons i.e., for measuring accuracy of the population count rather than in order to adjust census figures.

4.148. Countries may choose to decide about adjustments to the census results only when coverage errors are substantial, and the validity of census results is questionable. Consideration must be made to what level of geography the adjustment would cover knowing that such adjustments have an effect on demographic distributions. Input from stakeholders, including experts in demography, statistics, and affected communities, may as well be taken into account.

4.149. In those countries where the total population figure has been adjusted for under-enumeration or over-enumeration, both the enumerated figure (the population count) and the adjusted population figure (the population estimate) should be published for the sake of transparency and ensure that the method used should be well understood by census users in order not to undermine public trust in the census process. However, when the adjustment is done at the aggregate level, the detailed tabulations may be based only on the population that was actually enumerated[163].

4.150. Whatever decision is made (to adjust or not to adjust census data), it is important that any adjustment or information about the accuracy of census data (e.g., PES results and/or demographic analysis evaluation) is fully documented alongside the census results.

G. Acceptance of results

4.151. In countries with limited prior census experience and without a well-functioning civil registration system, where population data are based largely on estimates, it is important to inform the users, particularly the relevant governmental authorities, that the census results could differ from such estimates and to explain the reasons for these differences. In some cases, there may be doubts expressed about the census results; usually those doubts focus narrowly on the total population of the country, major subdivisions or population subgroups, rather than on the bulk of the census data relating to characteristics of the population or on the data for local areas.

4.152. In this situation, it is advisable to modify the census evaluation programme to better document differences and/or by adding appropriate qualifications to the text of the census reports or in tabular

163 UN: Post enumeration surveys. Operational guidelines, Technical Report, United Nations, New York (2010), https://unstats.un.org/unsd/demographic/standmeth/handbooks/manual_pesen.pdf.

footnotes. The other programmes (like population estimates) might also provide documentation to explain differences with census results and help users find their way across the different sources.

4.153. Nevertheless, the government may proceed with the utilization of the census results for official purposes. In any case, every effort should be made to process and evaluate the complete census operation and to make appropriate use of as many of the census tabulations as possible.

4.154. More generally in order to ensure transparency about data quality and enhance users' confidence in census results, it's important to assess the quality of census data as much as possible prior to their release, and to make this information publicly available, even though, in order to provide detailed coverage results and information on how accurately the census counted certain population groups, it might be necessary to wait for the results of the PES or other coverage surveys.

4.155. One of the primary methods of evaluating the quality of a census is comparing the results to other sources, mainly population benchmarks. While other sources certainly offer views on the quality of the census, there are limitations to the information they provide and the comparisons that can be made according to the sources. Therefore, data users should be guided in exercising caution in how they interpret differences. NSIs should explain which population comparisons are possible, and provide guidance on how to interpret differences between demographic benchmarks and census results. and results from the Labour Force Survey data are generally used for making comparisons about the population labour market (economic activity) status. In this case, users should be helped in interpreting differences due to data collection and questions design.

4.156. Many times doubts will be expressed with regard to how well specific ethnic groups or religious groups were counted in the Census, by comparing census data with data derived from other sources. In such cases as well, it is important to explain the concepts and definitions used in the census, and how these might differ from the one used in other sources (e.g. in administrative sources). To assist with making this comparison, NSIs might reconcile the categories from the census with those that appear in data from the administrative records and explain the methodology used for this.

4.157. On the overall, by providing documentation on the main comparisons possible between census data and other sources, NSOs should train data users in being cautious when they interpret differences and in recognizing that errors in the census are only one possible source of difference between the census and the population estimates or other sources.

H. User satisfaction

4.158. In many countries, National Statistical Institutes (NSIs) conduct user satisfaction surveys on products and services on a regular or occasional basis, in order to identify the gaps between what the institution achieves and what users actually need or expect to receive. In general, the main aim of user satisfaction survey is to gather opinions of NSIs users about the quality of data generated by the NSIs. Some NSIs collect information from users in an anonymous form, according to a continuous flow when accessing the institutional website or dissemination web platform, after consultation or downloading of statistical data. Periodically, this information is processed for internal use and in some cases presented in technical reports available on the web. The most important aspects that are measured through user ratings are timeliness or relevance of the data offered. Other statistics collected in this survey relate to the purpose of using NSIs data, the quality of NSIs data compared to statistics produced by other

organisations, users' trust in NSIs statistics, the frequency of tailored requests and the frequency of accessing the NSIs website. In other countries, specific surveys are occasionally conducted on samples of advanced users from different sectors, including central or local government, academia, businesses, the NGOs sector, to gather feedback from users with high statistical and IT skills who are able to use the data for policy, study and planning purposes.

4.159. Such a survey should address questions such as whether the disseminated census data fully satisfies users' demands with regard to the topics covered and breakdown provided; whether the level of territorial detail allowed by the census methodology and privacy constraints meet the user's needs; whether the dissemination timeliness fully satisfies users' needs; whether census data have been integrated with other sources available to the users' and whether they have been used through the use of GIS technologies; whether they have been used to carry out the studies, research, analyses, planning activities to be conducted within the scope of their institutional mission, and so on. Attention should also be paid to the accessibility, with regard to the suitability of the form or medium through which census information can be accessed both with regard to specialized users and to the public at large, and to interpretability of the census information, i.e., the availability of supplementary information and metadata necessary to interpret and use it, and indications of the accuracy of the information. Furthermore, in case of changes in census methodology that may have affected the scope and geography of disseminated census data, their impact on users should be investigated.

III. Overall evaluation of the census

A. Importance of evaluations

4.160. A population and housing census consists of a complex series of interrelated steps, and constitutes perhaps the single most extensive, complex and expensive statistical operation that a country undertakes. Like any other project, the census needs to be evaluated to ensure that the effort and investment of resources have been worthwhile.[101] Evaluation of the census for coverage and also to assess the quality of the data is covered in Part Four, Chapter II, and quality assurance for the census in Part Four, Chapter I. This section is concerned with the evaluation processes and procedures of the census operation.

4.161. Evaluation of the census is important for a variety of reasons, among which is to provide lessons learned from one census to the other. For this reason, evaluation is generally regarded as the last stage of the census cycle or the first step in the next census cycle. Evaluation assesses the effectiveness of operations, systems and processes and their likely impact on data quality. In this connection, it is particularly important to assess various aspects of the census operation, particularly when changes have been introduced compared to the previous census. Assessments provide valuable information on the strengths and weaknesses of past operational procedures, which should be carefully reviewed prior to the development of the next census.

4.162. In addition to the evaluation of the conduct and operational elements of the census, it is valuable to evaluate the use and value of census data and products. Evaluation of the outputs of the census can be conducted through the conduct of user surveys to gain subjective feedback or through looking at metrics around product use, for example number of visits to web products or the number of publications requested or purchased. Evaluation of census products can be used to measure whether the data and the selected topics are sufficiently accessible, timely, relevant, accurate, coherent, trusted and interpretable

to meet user requirements. The outcomes of the evaluation might provide information that can be immediately acted upon and remedied in the intercensal period, and should certainly feed into the preparations for the subsequent census.

4.163. In recent years, countries have introduced new methodologies and technologies in conducting censuses in order to reduce census costs and also to improve the quality and timeliness of census data. Plans should be made for thorough evaluations of the effectiveness of the new technologies so additional improvements can be implemented for the next cycle. There is use of new methodologies based on administrative registers and combinations of sources to produce census information all phases of the census. Plans should be developed to assess the quality of potential administrative sources. The added value of including these new administrative sources and the potential improvements in quality should be assessed alongside the risks associated with these new sources (for example, the status of the data supplier and the level of commitment of the data suppliers to the dimensions of quality). Other changes relate to the use of multimodal enumeration methods as well as substantial outsourcing of some aspects of census operations. Each new methodology, technology, and administrative source needs to undergo quality assessment so that a comprehensive risk profile and risk management plan can be established for the census design.

4.164. The evolving nature of census methodologies and operations underscores the need for rigorous evaluation to assess the impact of recent innovations. Furthermore, evaluations are necessary so as to provide lessons learned not just for the countries concerned but also for those who want to adopt similar processes for their future censuses.

B. Planning for the evaluation

4.165. Evaluation of census processes should not be undertaken on an ad hoc basis. The evaluation programme should be included in the overall census plan and be appropriately budgeted for. Lack of proper budgeting and planning for activities that come after enumeration, such as the post-enumeration survey, has in some cases led to financial shortages with negative consequences for the activities concerned. The success of the programme of evaluation depends to a large extent on setting, early enough, clear and unambiguous objectives to be served by the evaluation. Having clear objectives would help to design the best way to undertake the evaluation.

4.166. Ascertaining the objectives of the evaluation encompasses other aspects as well. It is important to establish the scope of the evaluation in terms of intended objectives, keeping in mind that the wider the scope the more complex the evaluation is likely to be. It should be noted also that evaluation of census processes and procedures could last several years and may cover different aspects of the census operation, as necessary, thereby resulting in a series of reports. These reports should be shared externally, as appropriate, to maintain transparency about the evaluation results and how those results can be used to improve the future design. As for other census activities, there is a cost associated with the evaluation and the more complex the undertaking or the more processes get evaluated, the higher the cost is likely to be.

4.167. In addition to the financial cost, human resources and skilled staff to undertake the evaluation should also be planned for. Depending on what aspects of the census operation will be evaluated, the national statistical or census office should ensure that they have adequate personnel (in type and quantity) to undertake the evaluation. A related issue that should be taken into account is the feasibility of the staff to undertake the evaluation. Lack of requisite skilled staff may limit the ability of the country

to undertake some or all of the planned components of the evaluation. An alternative is to hire temporary staff with the required technical skills to conduct the evaluation or to collaborate with other agencies such as research institutions.

4.168. As has already been stated, plans for the evaluation of census processes and procedures should be an integral part of the overall census plan and be planned for from the start of census activities. In addition, documentation of the census experience should have the provision of information for evaluation as one of its objectives. As presented in paragraphs 3.466–3.470 on "Documentation of census experience", the census organization should have a knowledge management system to document and archive complete records on plans, activities, and decisions taken during the entire census operation, including on problems faced and how they were resolved. Documented evidence on how the census was undertaken provides valuable input for the evaluation programme. It should be noted, however, that depending on what has to be evaluated, some of the information may not be available until the end of the census operation.

PART FIVE. POPULATION AND HOUSING CENSUS TOPICS

Chapter I: Population census topics

I. Factors determining the selection of topics

5.1. In line with the overall approach to revision 4 of *Principles and Recommendations for Population and Housing Censuses*, the selection of census topics is based on the outputs expected to be produced by the census. Therefore, the first step involves clear identification of user requirements for data; the core and additional topics are then decided on that basis. For each of the core topics there is a recommended tabulation[164]. It is recommended that countries collect data on the core topics and also produce the recommended tabulations, as this would improve the international harmonization and comparability of statistics through the use of common concepts, definitions and classifications. Use of an agreed international approach would also enhance the capacity of countries to generate statistics for monitoring the socioeconomic situation of their populations, including for the provision of data for the internationally agreed development goals.

5.2. The topics to be covered in the census (that is, the subjects regarding which information is to be sought for each individual or household) should, however, be determined upon a balanced consideration of:
(a) **National priority**: It is essential to consider the diverse needs of data users in the country at both the national and local area levels to ensure that the census generates relevant information to effectively address these needs.
(b) **International comparability**: It is crucial to ensure that the data collected is comparable at regional and global levels through aligning with international standards. This would facilitate cross-country analyses and monitoring.
(c) **Suitability**: For information collected through questionnaires, consider the sensitivity of the topics and the respondent burden, i.e., the willingness and ability of the public to provide accurate information on these topics. Also, assess the technical competence of enumerators to obtain information on the topics to be collected in field enumeration, ensuring they are well-trained and capable. Similarly, with respect to information to be collected through administrative registers, consider the sensitivity of the topics and the willingness of the public to have their personal information linked and used, i.e., the acceptability of the use of existing administrative information for statistical purposes. It is also important to consider the technical competence of data scientists in regard to obtaining information on the topics through data linkage.
(d) **Resources**: Consider the total national resources available for conducting the census, including financial, human, and technological assets.
(e) **Alternative sources**: Assess the availability of relevant information from alternative data sources to reduce duplication and improve the efficiency of data collection.

[164] Link to UNSD website containing the tabulation shells.

5.3. Such a balanced consideration will need to take into account the advantages and limitations of alternative methods of obtaining data on a given topic within the context of an integrated national programme for gathering demographic and related socioeconomic statistics (see paragraphs 1.128–1.133).

5.4. In selecting the population topics, careful consideration should also be given to the usefulness of historical continuity, which provides the opportunity for comparison of changes over a period of time. Census offices should avoid, however, collecting information that is no longer required by users. Information should not be collected simply because it was traditionally collected in the past, bearing in mind changes in the socioeconomic circumstances of the country. It becomes necessary, therefore, in consultation with a broad range of users of census data, to review periodically the value of even long-standing topics and to re-evaluate the need for their continued collection, particularly in the light of new data needs and alternative data sources that may have become available for investigating topics hitherto covered in the population census. Each of the five key factors (i.e., national priority, international comparability, suitability, resources, and alternative sources) that need to be taken into account in reaching a final decision on the content are briefly reviewed in the following paragraphs.

A. National priority

5.5. Prime importance should be given to the fact that population censuses should be designed to meet national needs. In defining national data needs for population census data, the full range of national uses (for example, policymaking and monitoring, administration and research) and national users (for example, national and local government agencies, the private sector, and academia) should be considered. The prime consideration is that the census should provide information on those topics that are of greatest value to the country, with questions framed, or appropriate administrative data sources identified, so as to elicit data of maximum utility.

5.6. Each country's decision with regard to the topics to be covered should depend upon a balanced appraisal of how urgently the data are needed and whether the information could be equally well or better obtained from alternative data sources such as administrative records or household surveys. Experience has shown that national needs will best be served if the census includes topics generally recognized as being of basic value and defined in accordance with regional and global standards. International census recommendations can help in this appraisal by providing information about core and additional census topics and their definitions and concepts, emanating from national experiences. It is recognized however that countries that rely more on administrative records as their primary data source may be more limited in the precise detail of the information that can be collected on certain topics.

5.7. Many countries may find it necessary to include topics of national or local interest in their census, in addition to the topics suggested in these recommendations. Labour force surveys and/or other survey data may supplement census data so as to obtain information on topics that cannot be included in the census for whatever reason. It is possible that some countries may omit from their census certain recommended topics either because there is no longer a need to collect the data or because there are legal barriers or particular sensitivities in collecting information on such topics (for example, fertility, ethnicity and religion).

B. International comparability

5.8. The desirability of achieving regional and international comparability should be another major consideration in the selection and formulation of topics to be included in the census. National and international objectives are usually compatible, since international recommendations are based on a broad study of country experiences and practices, the definitions and methods contained in international recommendations have successfully met general national needs in a wide range of circumstances. Furthermore, the analysis of census data for national purposes will often be facilitated if, through the use of international recommendations, it is possible to compare the data with those of other countries on the basis of consistent concepts, definitions and classifications. The 2030 Sustainable Development Agenda, which places increasing demand for expanded data collection, is also another determining factor that countries should take into consideration.

5.9. If the particular circumstances within a country require a departure from international standards, every effort should be made to explain these departures in the census publications and to indicate how the national presentation can be adapted to the international standards.

C. Suitability

5.10. A prerequisite for the inclusion of topics in field-based censuses should be the willingness and ability of respondents to provide accurate information on them, or, for register-based censuses, the willingness of the public to have their personal information linked and used. It is advisable to avoid topics that could increase the burden on respondents and those that are likely to arouse fear, local prejudice or superstition or that might be used to deliberately promote political or sectarian causes as these are likely to have a detrimental effect on response rates and support for the census. In an interview-based census or when the enumerators need to obtain information through observation, consideration needs to also be given to the level of knowledge and skill of the enumerators and whether they can be adequately trained to collect this information accurately. Topics that are too complicated or difficult for the average respondent or enumerator to answer quickly should not be included. The exact phrasing of a question that will obtain the most reliable responses may depend on national circumstances and, as described in Part Three of these recommendations, should be well tested prior to the census (see paragraphs 3.138–3.146).

D. Resources

5.11. The selection of topics should be carefully considered in relation to the total resources available for the census. An efficient collection of accurate data for a limited number of topics, followed by prompt tabulation and publication, is more useful than the collection of data for an overambitious list of topics that cannot be properly processed and disseminated in a timely, reliable and cost-effective manner. In balancing the need for data against resources available, the extent to which questions can be precoded is yet another consideration. Information from studies on the capacity of users and on the measurement of data utilization may also be an important factor in determining whether or not it is economically feasible to include certain topics in the census. In the case of those topics on which data is taken from registers, the additional resources required to transform the administrative records into useable statistical data should also be noted.

E. Alternative sources

5.12. In the selection of topics to be investigated in a census, consideration should be given to whether data are available in alternative sources, such as administrative records or household surveys,

taking into account their relative advantages and limitations. While these sources offer valuable information, they often have limitations. Household surveys, for example, typically rely on samples and might lack the geographic coverage needed for detailed analysis of small areas or certain small population groups. Censuses, on the other hand, can collect comprehensive data through direct interviews. This allows for probing questions and clarification, which may not be possible with self-enumeration or administrative records. Those topics for which no alternative sources exist should be given higher priority for inclusion in a census, while those for which alternative sources are readily available may be accorded a lower priority.

II. List of topics

5.13. The list of topics included in these recommendations for population censuses are based on the global and regional census experiences in the last several decades. Most of the topics included here are, with some revisions, generally the same as those included in the previous United Nations population census recommendations.[165] The concepts and definitions for some of the topics relating to economic characteristics have been substantially revised to reflect the more recent recommendations of the International Conference of Labour Statisticians[166] and other recommendations. Other topics of revision include ICT and agriculture, while new topics on reason for change of residence, ever resided abroad, reason for international migration, and registration of vital events and assignment of legal identity have been introduced.

5.14. It should be emphasised that no country should necessarily attempt to cover all the topics included in the list of population topics (see Table 3). Rather, countries will need to make their selection of topics in the light of the five key considerations discussed in paragraphs 5.1–5.12 above, bearing in mind any regional recommendations currently pertaining to census topics. In using the classifications of different topics presented in this part of the *Principles and Recommendations for Population and Housing Censuses,* it is necessary to outline that all the one- and two-digit classification levels are recommended, while those at the three-digit level are provided for reference and should be applied only when specific data requirements necessitate greater detail.

5.15. Over the past several decades, global and regional census experiences have highlighted a set of topics on which there is considerable agreement on their importance and feasibility for data collection in a census. However, some topics within this set can be excessively costly to collect directly through a census. In such cases, separate surveys of a sample population are a more practical approach. Notably, exceptions exist at both ends of the spectrum: in countries with highly developed statistical systems, adequate data on several of these topics (including some of the core ones) are available from non-census sources such as registers and administrative records. Conversely, among countries with limited data collection opportunities (and where it is felt that advantage must be taken of the possibilities offered by the census), the census becomes crucial for investigating topics that might otherwise be better suited for alternative means. Ultimately, the decision on which topics to include in a census should be based on a careful assessment of data availability, cost-effectiveness, and the specific needs of the country.

---

[165] *Principles and Recommendations for Population and Housing Censuses Revision 2*, Statistical Papers No. 67/Rev2 (United Nations publication, Sales No. E.07.XVII.8).

[166] For more detail see the Resolution concerning statistics of work, employment and labour underutilization, adopted by the 19th International Conference of Labour Statisticians (Geneva, 2013).

5.16. Although the set of topics covered in these recommendations is quite comprehensive in terms of those generally considered suitable for inclusion in a population census, it is also recognized that some countries may find it necessary to include additional topics on which information is of particular national or local importance. However, before the final decision is made to include any such additional topics, their suitability should always be carefully tested, and the quality of the resulting data evaluated beforehand.

5.17. To assist countries in using the present publication and in determining their own priorities, lists of recommended population topics are summarized in paragraph 5.21, with the core topics shown in boldface. These core topics correspond to those that were included as "priority topics" in the majority of the regional recommendations in previous census decades.

5.18. The topics listed in paragraph 5.21 are grouped into eight categories: geographic and internal migration characteristics, international migration characteristics, household and family characteristics, demographic and social characteristics, fertility and mortality, educational characteristics, economic characteristics, and agriculture.

5.19. Within each category, a distinction is made between topics collected directly (either from responses given in a census questionnaire or taken from an administrative register) or derived. Although data for the derived topics also come from information on the questionnaire or the register(s), they do not necessarily come from replies to a specific question. "Total population", for example, is derived from a count of the persons entered on the questionnaires as persons present or resident in each geographic unit. Such derived topics may perhaps be more correctly considered as tabulation components, but they are listed as topics in order to emphasize the fact that the questionnaire must in some way yield this information.

5.20. The paragraph numbers in parentheses after each entry in Table 3 refer either to the paragraphs in which the group of topics as a whole is discussed in section IV below or to the paragraphs in which the definition and specifications of individual topics are discussed.

5.21. In the following list of population census topics, core topics are shown in bold and are represented by ♦ for topics that are collected directly, and by □ for those that are derived. Additional topics are represented by ○, and additional topics derived from a core topic are indicated with Δ.

Table 3. List of population census topics

A. Geographic and internal migration characteristics (paras. 5.50–5.105)		
(1)	**Place of usual residence** (paras. 5.52–5.57)	♦
(2)	**Place where present at time of census** (paras. 5.58–5.63)	♦
(3)	**Place of birth** (paras. 5.64–5.71)	♦
(4)	**Duration of residence** (paras. 5.72–5.74)	♦
(5)	**Place of previous residence** (paras. 5.75–5.76)	♦
(6)	**Place of residence at a specified date in the past** (paras. 5.77–5.81)	♦

(7)	Reason for change of residence (paras. 5.82–5.85)	○
(8)	**Total population** (paras. 5.86–5.93)	□
(9)	**Locality** (paras. 5.94–5.98)	□
(10)	**Urban and rural** (paras. 5.99–5.105)	□
	B. International migration characteristics (paras. 5.106–5.133)	
(1)	**Country of birth** (paras. 5.111–5.115)	♦
(2)	**Country of citizenship** (paras. 5.116–5.121)	♦
(3)	Acquisition of citizenship (paras. 5.122)	○
(4)	**Year or period of arrival** (paras. 5.123–5.126)	♦
(5)	Ever resided abroad (paras. 5.127–5.128)	○
(6)	Reason for international migration (paras. 5.129–5.133)	○
	C. Household and family characteristics (paras. 5.134–5.162)	
(1)	**Relationship to the reference person of household** (paras. 5.142–5.153)	♦
(2)	**Household and family composition** (paras. 5.154–5.161)	□
(3)	Household and family status (para. 5.162)	○
	D. Demographic and social characteristics (paras. 5.163–5.233)	
(1)	**Sex** (para. 5.166 – 5.169)	♦
(2)	**Age** (paras. 5.170–5.181)	♦
(3)	**Marital status** (paras. 5.182–5.190)	♦
(4)	Ethnocultural characteristics (paras. 5.191-5.192)	○
(5)	Religion (paras. 5.193–5.197)	○
(6)	Language (paras. 5.198–5.201)	○
(7)	Ethnicity (paras. 5.202–5.206)	○
(8)	Indigenous peoples (paras. 5.207–5.211)	○
(9)	**Disability status** (paras. 5.212–5.233)	♦
	E. Fertility and mortality (paras. 5.234–5.279)	
(1)	**Children ever born alive** (paras. 5.248–5.255)	♦
(2)	**Children living** (paras. 5.256–5.258)	♦
(3)	**Date of birth and sex of last child born alive** (paras. 5.259–5.262)	♦
(4)	Births in the past 12 months (paras. 5.263–5.265)	Δ
(5)	Deaths among children born in the past 12 months (paras. 5.266–5.268)	Δ
(6)	Age, date or duration of first marriage (para. 5.269–5.270)	○
(7)	Age of mother at birth of (date or time when) first child born alive (para. 5.271)	○
(8)	**Household deaths in the past 12 months** (paras. 5.272–5.276)	♦

	(9)	Maternal or paternal orphanhood (paras. 5.277–5.279)	○
	F. Registration of vital events and assignment of legal identity (paras. 5.280–5.298)		
	G. Educational characteristics (paras. 5.299–5.329)		
	(1)	**Literacy** (paras. 5.299–5.305)	♦
	(2)	**School attendance** (paras. 5.306–5.312)	♦
	(3)	**Educational attainment** (paras. 5.313–5.321)	♦
	(4)	Field of education and training, and educational qualifications (paras. 5.322–5.329)	○
	H. Use of information and communication technology at individual level (paras. 5.330–5.335)		
	(1)	Individuals using the Internet (paras. 5.334)	○
	(2)	Individuals owning mobile cellular phones (paras. 5.335)	○
	I. Economic characteristics (paras. 5.336–5.452)		
	(3)	**Labour force status** (paras. 5.354–5.379)	♦
	(5)	**Status in employment** (paras. 5.385–5.408)	♦
	(6)	**Occupation** (paras. 5.409–5.412)	♦
	(7)	**Industry** (paras. 5.413–5.416)	♦
	(8)	Place of work (paras. 5.417–5.422)	○
	(9)	Institutional sector of employment (paras. 5.423–5.425)	○
	(10)	Working time (paras. 5.426–5.432)	○
	(11)	**Participation in own-use production of goods** (paras. 5.433–5.438)	♦
	(12)	Income (paras. 5.439–5.443)	○
	(13)	Informal employment (paras. 5.444–5.452)	○
	J. Agriculture (paras. 5.453–5.461)		
	(2)	Own-account agriculture production (paras. 5.455–5.456)	○
	(3)	Size of farm and number of livestock (paras. 5.457–5.461)	○
	Legend: ♦ Core topic, collected directly; □ Core topic, derived; ○ Additional topic; and Δ Additional topic, derived from a core topic.		

III. Population count

5.22. One of the main objectives of a population census is to provide an accurate count of the population of a country at a specific point in time. An accurate population count is essential for, inter alia, the efficient planning and delivery of services, distribution of resources, defining boundaries for electoral representation, policy development and a wide range of other administrative, commercial, and

statistical purposes (see paragraphs 1.15-1.60 on the various uses of the population and housing census data).

5.23. A "population count" may refer to the whole, or a subset, of the enumerated population. A country may produce one or more population counts, all derived from the enumerated population.

5.24. Countries are usually most interested in the count and distribution of usual residents because usual residence is generally the best indication of where people will demand and consume services. A count of usual residents is therefore most relevant for many common planning and policy purposes.

5.25. Some countries may supplement the population count from their census with information from other sources, for example on usual residents temporarily outside the country at the time of the census, to produce their population estimates. Alternatively, the census itself may collect data on these groups, resulting in an enumerated population that includes individuals who are usually resident but temporarily absent.

5.26. In countries whose censuses use direct enumeration, information about each person can be collected and entered on the census questionnaire either where the person is (or was) present on the day of the census or at their place of usual residence. Paragraphs 2.48–2.64 describe the basis for deciding the place of enumeration in the census.

5.27. A country may need to produce one or more different counts for different national purposes. Population counts may be required on a population present basis, or for the usually resident population, or for the service population (see paragraphs 5.43-5.47). The choice of population count(s) required will depend on national circumstances; some countries will require more than one. The information collected about each person by the census will need to enable the required population counts to be derived. In many cases, for regional and international comparison purposes, the population count based on the generally accepted concept of usual residence will need to be produced.

5.28. A census aims to achieve a full and unduplicated coverage of the population. In practice, countries may face a range of challenges in enumerating the population at the place they decide (where present on census day or where usually resident), and in producing the population counts they require. Many of these challenges relate to the difficult-to-enumerate/reach groups of the population and persons for whom the concept of "usual residence" is not easily defined. The latter present an increasing problem as populations become more mobile (nationally and globally) and household and family structures more dynamic.

5.29. In developing strategies for enumerating the population and collecting information to support the required population counts, it is important to consider consistency with the standards for international migration statistics described in paragraphs 5.107–5.110.

A. Population present count

5.30. A population present count is the simplest form of population count from a population census. In a questionnaire-based census where no reference is made to usual residence, people are enumerated at the place where they are found, usually the dwelling where they spend census night. Foreign residents who are in the country at the time of the census will be included but usual residents of the country who are absent at that time will be excluded.

5.31. A population present count removes complications associated with the application of the concept of place of usual residence, and can reduce the incidence of double counting or missing people if the enumeration is carried out in a single day or if reference is made to the same census moment for the whole population. Apart from the benefit of simplicity, a population present count offers a cost advantage because the census does not need to collect additional information about usual residents not at their usual place of residence at the time of the census.

5.32. The major disadvantage of a population present count is that it does not enable a full count of usual residents to be derived, and does not provide a true geographic distribution of usual residents for effective planning and policy purposes.

5.33. A population present count may be a good proxy for a count and distribution of usual residents, particularly if nearly all the population will be at their usual residence at the time of the census, or if the characteristics of those persons present are very similar to the characteristics of usual residents. However, in many countries significant numbers of people will not be at their usual residence at the time of the census, and the characteristics of absent usual residents will be systematically different from non-residents present. In such cases, a population present count will not be a good proxy for a count of usual residents. Large seasonal movements of people due to weather changes, employment, education, holidays and other factors can result in these systematic differences between the population present and the usually resident population. The ability to produce accurate information on families and households is also reduced when persons temporarily absent from their usual residence (e.g. seasonal workers or students) are not enumerated with their families or households.

To produce a population present count, information is required on all persons present and the address at which they are enumerated. Additionally, while it is not necessary to produce the population present count, it might be valuable to collect data that identify individuals who are not at their usual residence or who are not usual residents of the country. By collecting this additional data, census agencies can create a more comprehensive and accurate snapshot of the population at a specific point in time, enabling better analysis of population dynamics, migration patterns, and service needs.

5.34. Ideally a population present count should include all the persons present at the census reference moment, regardless of the difficulty of their enumeration. For some of these groups the concept of “at the time of the census” may need to be extended to allow a full enumeration to be completed. When, however, the enumeration is extended over a period of time, the risk of either overcount or undercount may increase. In fact, persons who are at multiple locations during this extended period may be counted at more than one location, or alternatively they may not be counted at any location. Those risks increase further when reference is made to a *census period* rather than to a *census moment*.

B. Usual resident population count

5.35. Countries increasingly prefer a usual resident population count because this count offers better information for planning and policy purposes on service demand, households, families and internal migration.

5.36. A **usual resident population count** is a count of all usual residents of a country at the time of the census. Although countries will determine the definition of a usual resident according to their own

particular circumstances, it is recommended that in defining a usual resident and the place of usual residence, countries apply the definition contained in paragraphs 2.50 and 2.51. Usual residents may or may not have citizenship of the country, and they may also include undocumented persons or irregular migrants, applicants for asylum or refugees. For statistical purposes, usual residents then may include foreign citizens who reside (legally or illegally), or intend to reside, in the country continuously for either most of the last 12 months or for 12 months or more, depending on the definition of place of usual residence that is adopted by the country. Persons who may consider themselves usual residents of a country because of citizenship or family ties, but were absent from the country for either most of the last 12 months, or for 12 months or more, depending on the definition adopted, should be excluded. Conversely, persons who are normally resident in the country but who are temporarily absent should be included in the usually resident population. Countries applying a different definition of a usual resident for national purposes should produce a usual resident population count using the recommended 12-month definition for the purposes of international comparability.

5.37. A usual resident count provides a comprehensive count of the population of a country for long-term planning and policymaking purposes, and for service delivery purposes at subnational geographic levels.

5.38. Whether enumeration takes place on a “place where present” basis or on a “where usually resident” basis (as described in paragraphs 2.57–2.64), a count of usual residents can be produced provided that the necessary information is collected. In which case, information is required on all usual residents and the address of their usual residence, with sufficient detail to be able to record usual residence at the lowest geographic area level required for tabulation. If the census is taken on a population present basis, then the information collected needs to differentiate clearly among persons enumerated at their usual residence, persons usually resident who were elsewhere at the time of the census, and persons present who are usually resident elsewhere. Information should also be collected to identify those persons present on census day who are not usual residents of the country. If, however, the census is taken on a usual residence basis, then information about all usual residents needs to be collected with respect to their usual residence, regardless of whether they are present at the time of the census or not, to ensure full coverage.

5.39. It may be difficult to obtain information on those usual residents who are absent from the country at the time of the census, particularly where no other person is present at the place of usual residence at the time of the census to provide the relevant information. In such cases, imputations, using data from the census itself or other sources, could be utilised regarding the number and characteristics of these usual residents to estimate the census population count.

5.40. There can be challenges in applying the concept of “usual resident” if a person is considered to have more than one residence, within or outside the census country. This is particularly the case for people who may spend parts of the time in communal establishments or institutions, such as schools/dormitories or military camps. There may also be those who do not consider themselves to have a usual residence at all, such as nomadic peoples or persons sleeping rough[167]. In such cases place of usual residence can be considered to be the place where they are enumerated on the census date. Countries will need to develop appropriate operational rules for resolving cases where it is not clear

[167] People ‘sleeping rough' are those living in the streets or public spaces, without a shelter that can be defined as living quarters. The group sleeping rough is a subset of the homeless. Not everyone who is homeless is sleeping rough. There are many other forms of homelessness, including those living in shelters, cars, or couch surfing.

whether a person is a usual resident of the country, or where the usual residence of the person within the country is not clear.

5.41. There are population groups for which some uncertainty may arise in defining their place of usual residence within the country. The recommended conventional treatment of these cases is as follows:

a) For persons who work away from home during the week and return at weekends, their place of usual residence should be their weekend residence.

b) For persons of minor age in primary and secondary education who are away from home during the school term, their place of usual residence should be their family home/residence.

c) For students in tertiary education who are away from home while at college or university, their place of usual residence should be their term-time address regardless of whether this is an institution (such as a boarding school) or a private residence.

d) The institution should be taken as the place of usual residence of all occupants who at the time of the census have spent, or are likely to spend, six months or more in the relevant institution. Examples of occupants of institutions include patients in hospitals or hospices, older persons in nursing homes or convalescent homes, children in orphanages, prisoners and those in juvenile detention centres.

e) Where a person regularly lives in more than one residence within the country during the year, the one where he or she spends the majority of the year before the census should be taken as his or her place of usual residence. These persons are not considered to be persons with no usual residence.

f) For the (national) military, naval and diplomatic personnel and their families located outside the country the following classification rules should be applied:
i) If they are residing abroad for less than 12 months and they intend to return to the place/country of departure, they should be allocated within the country in accordance with the rules for usual residence. In particular, they could be allocated to (by decreasing order of priority):
- The family home address within the country, if any;
- The duty station within the country to which they were attached before leaving.
ii) If they are residing abroad for at least 12 months or if they do not intend to return to the place/country of departure (although returning in/visited the country within a 12-month period), they should be attributed to a “virtual region” (extra-region) of the country of departure.

g) The place of enumeration should be taken as the place of usual residence of homeless or roofless persons, nomads, vagrants and persons with no concept of usual residence.

h) For a child who alternates between two households within the country (for instance when his or her parents are not living together because of divorce, legal separation, or other reasons), the household where he or she spends the majority of the year before the census should be considered as his or her place of usual residence. Where an equal amount of time is spent with both parents, the place of usual residence should be with the parent or household with whom the child is staying at the census reference time.

C. Supplementary population counts

1. Service population count

5.42. Counts of usual resident and/or population present are the most commonly produced census population counts, since in most cases they provide a good indication of the geographic distribution of demand for and patterns of use of services. However, in some countries or specific parts of some countries, the demand for and use of services, as well as other considerations such as exposure to risks and opportunities, is shaped significantly by temporary movements. This is the case, for example, in some agricultural areas which may have very few usual residents but large populations of short-term residents during harvest season; in tourist resorts with large numbers of visitors but relatively few usual residents; or in urban centres where few people live but many are present as workers during the daytime. In such cases a country may wish to produce supplementary population counts.

5.43. Service populations are relevant where a significant proportion of the population providing or using services in an area are not usual residents of that area. Types of service population counts include daytime populations, workplace populations and visitor populations. In some countries there may also be an interest in foreign service populations, consisting of foreign citizens who cross the border regularly to provide or consume services. This is particularly important in the planning and provision of transport services.

5.44. A service population count may include some of the difficult-to-enumerate groups, depending on the type of service population required. For example, daytime service populations may include civilian foreign citizens who cross the border daily to work or consume services in the country.

5.45. To produce a service population count, in addition to an estimate of usual residents, information is required about where people provide or require services. For seasonal populations (holiday, resort), information is needed on the destination and timing of seasonal trips. Some countries will produce service population counts by supplementing the population present count or usual resident population count with information from other sources, such as visitor information from hotels and resorts, to produce visitor population counts. Alternatively, such additional information may be collected by the census itself.

2. Counts of population subgroups

5.46. Accurate population counts, necessary for the efficient planning and delivery of services, distribution of resources, defining of boundaries for electoral representation, policy development and the design and analysis of household surveys, are required for various population subgroups within a country. These subgroups are typically based on geography, age and sex. There may also be a need to identify other populations such as the school population, working population, indigenous population or population groups in vulnerable situations to enable more informed policy formulation and better targeted service provision. A range of characteristics will be required to identify such population subgroups, depending on the services being planned, the resources to be distributed and so on. The need for population counts for particular subgroups will determine some of the questions asked in the census or the information collected from registers.

D. Difficult-to-enumerate groups or difficult-to-reach groups

5.47. Difficult-to-enumerate groups can be defined as those for whom a real or perceived barrier exists to full and representative inclusion in the data collection process[168]. In fact, some populations present special challenges of various sorts that make them harder to enumerate than the general population. The definition of difficult-to-enumerate groups can vary between countries as the reasons why members of a population group are hard-to-reach (or difficult-to-enumerate) can vary according to the context of each national, geographic, or social environment. Nonetheless, countries should ensure that the following difficult-to-reach/enumerate groups are included in the population count:

- **Nomads and persons living in areas to which access is difficult.** Making contact with these groups to enumerate them can be difficult, particularly as part of a point-in-time count. Enumeration may need to be done at a different time, over an extended period, or by using alternative methods to enable contact with these groups. For example, countries might consider asking those who provide services to these groups to assist with their enumeration. Seasonal movements may be identified in advance and this information can be used by collectors to enable contact. There needs to be planning and consultation, particularly with influential members of these groups, prior to the census to organize for their enumeration. Communications that publicize the benefits of the census and engaging appropriate leaders in support of the census may assist coverage. Awareness of cultural issues relevant to specific groups should also be considered in developing enumeration strategies.
- **Civilian residents temporarily absent from the country.** Usual residents absent from the country at the time of the census should be included in the usual resident population count, despite their absence. To produce a usual resident count, countries may collect information on these people from another family or household member present at the time of the census, but where a complete family or household is outside the country at the time of the census, it may not be possible for the census to collect information about these people. Estimates for usual residents temporarily absent from the country based on other sources may be required to produce reliable estimates of usual residents for planning and policy purposes.
- **Civilian foreign citizens who do not cross a frontier daily and are in the country temporarily.** These may include undocumented persons, or transients on ships in harbour at the time of the census. If these groups are present in the country on census day, they should be included in the population present count (they would typically be excluded from the count of usual residents). It is important to include these groups in the population count if their demand for services is to be considered for planning and policy development purposes. However, these groups may prefer not to be counted, either because they fear ramifications from being counted or because they do not identify themselves as part of the population of the country. Language and communication may present challenges. Countries need to develop strategies, appropriate for their context, to include these groups in their enumeration.

[168] Tourangeau R., Brad Edwards B., Johnson T. P., Kirk M. Wolter K.M. Bates N. (eds), Hard-to-Survey Populations, Cambridge University Press (2014), DOI: https://doi.org/10.1017/CBO9781139381635.003; U.S. Census Bureau: Counting the hard to count in a census (2019), Available from: https://www.census.gov/content/dam/Census/library/working-papers/2019/demo/Hard-to-Count-Populations-Brief.pdf

- **Refugees, asylum seekers and internally displaced persons.** Refugees, asylum seekers[169] and internally displaced persons[170] (in and outside camps) should be enumerated in both the population present count, and in the usual resident count if they meet the established duration criteria. To facilitate analysis, these groups should be reported separately, to allow the calculation of a country's population excluding these categories when required for non-demographic purposes.
- **Military, naval and diplomatic personnel and their families located outside the country and foreign military, naval and diplomatic personnel and their families located in the country.** Apart from the difficulties mentioned in (b) and (c) that are common to groups who are absent from their own country, enumeration of these groups is subject to diplomatic protocols. Detailed counts and characteristics of these groups may be considered sensitive on security grounds in some countries. Counts of these groups may be available from administrative records.
- **Civilian foreign citizens who cross a frontier daily to work in the country.** This group should be excluded from a usual resident population count. The practice of counting people where they spend census night removes much ambiguity and reduces possible duplication. The difficulty then is trying to include them in a service population if countries want to consider this group in policy development and in planning service delivery.
- **Civilian national citizens who cross a frontier daily to work in another country.** These persons are usual residents of the country and should be included in the population count.
- **Merchant seafarers and fishermen resident in the country but at sea at the time of the census.** This group includes those who have no place of residence other than their quarters aboard ship. Identifying that the ship will be at sea at the time of the census may be problematic, so countries will need to develop strategies to ensure inclusion of this group in the population count. This may include providing this group with census forms before their ship goes to sea or enumerating the ship before the census date.
- **Homeless or roofless persons, vagrants and persons with no concept of usual residence.** These should be included in the population count, and the census office should work with local government agencies, charities and other supporting bodies that provide support for these population groups to identify the best method of collecting census information from these people.
- **Persons living in buildings with restricted access.** Persons living in gated communities, condominiums and apartment buildings could be difficult to enumerate, particularly as part of a point-in-time count. Enumeration may need to be done at a different time, over an extended period, or by using alternative methods to enable contact with these groups.
- **Stateless persons[171].** These are individuals who are not considered as nationals by any State under the operation of its laws. They are often undocumented and may not wish to be enumerated. However, every effort should be made to include such persons - both those with recognised stateless status and those without a recognised nationality status[172] - in the census. The census office should work with responsible government agencies, non-governmental organizations familiar with this population group and the United Nations High Commissioner for Refugees (UNHCR) to establish the best method for including

[169] https://egrisstats.org/recommendations/international-recommendations-on-refugee-statistics-irrs/
[170] https://egrisstats.org/recommendations/international-recommendations-on-idp-statistics-iris/
[171] https://egrisstats.org/recommendations/international-recommendations-on-statelessness-statistics-iross/
[172] IROSS, see paragraph 172, more info https://egrisstats.org/recommendations/international-recommendations-on-statelessness-statistics-iross/.

stateless persons in the population count. Information on country of citizenship is generally essential for the identification of this group, though the census office should consult with relevant ministries and agencies, including the UNHCR, to determine whether additional information (such as residence history or identity documentation) may be required to establish the status of a stateless person.

- **Persons with an irregular migration situation.** Persons moving outside regular migration channels [173] as well as persons who entered regularly but overstayed the permitted duration.
- **Persons who refuse to participate in the census due to political reasons.** These are individuals who distrust the government and refuse to participate in the census. Census offices should work closely, via a strong communication/publicity team, with all the major political parties to encourage participation of this group.
- **Indigenous peoples**. See paragraphs 5.207-5.211. The indigenous population may tend to live in sparsely populated remote areas which are usually difficult to access and enumerate. In some countries, these areas are managed by a leader and councillors who maintain up to date records on basic demographic and social data (including address) on each family/household. In such cases, the census office/enumerators could access these records and work in collaboration with the respective leaders to mobilize and or locate the actual families/households for enumeration. In cases where the areas fall under a government ministry/agency/department, the census office could access these population with the assistance of the responsible organisation.
- **Seasonal workers**. These are persons who are employed in a temporary employment situation that repeats yearly around the same time. Some of the most common industries for seasonal workers are farming and agriculture, tourism, and construction. This group should be included in the usual resident population count if they meet the 12-month criteria described in paragraph 5.36 and counted where they spend census night. This will remove ambiguities and limit possible duplications.
- **Older persons and those with cognitive impairments**. Older people living in institutions (relevant for both enumeration-based and register-based censuses), and persons living in private households but without the cognitive capability to complete a questionnaire-based census.

5.48. The same difficulties apply to register-based censuses. While administrative data can enhance coverage for certain hard-to-reach populations, it may also miss other 'hidden' groups within different population segments. For instance, undocumented migrants are not automatically included in population registers, and consequently, they are absent from statistics derived from administrative records. However, many other difficult-to-enumerate groups might be included but not easily identifiable within administrative data. It's crucial to recognize these potential limitations when relying primarily on administrative sources.

[173] See: https://www.iom.int/key-migration-terms. Although a universally accepted definition of irregular migration does not exist, the term is generally used to identify persons moving outside regular migration channels. Moreover, categories of migrants who may not have any other choice but to use irregular migration channels can also include refugees, victims of trafficking, or unaccompanied migrant children. Persons who entered a country's territory without authorization as well as persons who arrived in a country legally for work or study and overstayed their permits or visas, as well as persons who received a negative decision in an asylum determination procedure or who lose their residency status because of a conviction for a criminal offence or following a review of their refugee or subsidiary protection status fall under this category.

IV. Definitions and specifications of topics

5.49. This section contains the recommended definitions and specifications of all topics presented in the order in which they appear in Table 3 above. It is important that census data be accompanied by the definitions used in carrying out the census. It is also important that any changes in definitions that have been made since the previous census are reported in the metadata and, if possible, accompanied by an assessment of the effect of such changes on the relevant data, in order to ensure that users will not confuse valid changes over a period of time with increases or decreases resulting from changed definitions.

A. Geographic and internal migration characteristics

5.50. It should be noted that "place of usual residence" and "place where present at time of census" may be considered alternative topics when countries do not have the resources to investigate both topics for general census purposes. Some countries, however, will want to investigate both topics. The relationship between the two topics and their further relationship to the topic of "place of enumeration" are set out in Part Two, Chapter IV (see paragraphs 2.57–2.64).

5.51. It is recommended that countries investigating only "place where present at time of census" for general purposes should also obtain information on "place of usual residence" for all persons who do not usually reside in the household where they were enumerated, to be used in connection with the information on "place of birth", "duration of residence", "place of previous residence" or "place of residence at a specified date in the past" for the purposes of determining internal migration status. If, in the compilation of the population of geographic units, persons are allocated to the place where they were present at the time of the census, information on the four above-mentioned migration characteristics will be irrelevant for persons who were only visiting, or transient in the place at which they were present. Since such persons must, in any case, be identified in the questionnaire as non-residents so that they will not be erroneously classified as recent immigrants, information on place of usual residence should be collected, which will make it possible to include the entire population in the tabulation of internal migration characteristics.

1. Place of usual residence (core topic)

5.52. Information on the number of people usually residing in an area is basic to most informed decision-making about the area, whether it be a country, an urban agglomeration or a civil division. The number of residents determines the levels of most services required in an area.

5.53. The *place of usual residence* may be the same as, or different from, the place where the enumerated person was present at the time of the census or his or her legal residence. For a definition of place of usual residence, see paragraphs 2.50–2.52.

5.54. Although most persons will have no difficulty in stating their place of usual residence, some confusion is bound to arise in a number of situations where persons have more than one residential address. These cases might include persons who maintain two or more residences, students living at school, members of the armed forces living at a military installation but still maintaining private living quarters away from the installation, and persons who sleep away from their homes during the working week but return home for several days at the end of each week (see paragraph 5.41). In some other

circumstances, referring to the person's intentions for the future may assist the determination of the place of usual residence.

5.55. Problems may also arise with persons who have (a) been residing at the place where they are enumerated for some time, but do not consider themselves to be residents of that place because they intend to return to their previous residence at some future time; or (b) left the country temporarily but are expected to return after some time longer than 12 months from the departure. In such instances, clearly stated time limits of presence in or absence from a particular place must be based upon the 12-month limit and used to determine whether the person is usually resident there. The 12-month criterion is necessary for determining whether a person is usually resident in the country. It is also required for international comparability of migration statistics.

5.56. If each person is to be entered in the questionnaire only at his or her place of usual residence, the topic need not be investigated separately for each person, because the information will be available from the location information entered for the questionnaire as a whole.

5.57. Information on the place of usual residence should be collected with enough locational details to enable tabulations to be made for the smallest geographic subdivisions required by the tabulation plan and to meet the requirements of the database within the cost limits and operational procedures required to code to a fine degree of detail.

2. Place where present at time of census (core topic)

5.58. In cases where the census is taken on the basis of "place where counted" (often referred to as a *de facto* count), this topic may fulfil some of the functions of place of usual residence.

5.59. The *place where present at time of census* is, in theory, the geographic place at which each person was present on the day of the census, whether or not this was his or her place of usual residence. In practice, the concept is generally applied to the place where the person was present at the moment of the census, because many persons may not be physically present at the place of enumeration during most of the day.

5.60. The concept is sometimes further extended to apply to the night preceding the day of actual enumeration in cases where the enumeration extends over a long period of time and persons are not likely to be able to supply information relating to a single moment in this period. Other departures from the definition may be necessary to deal with individual cases, such as persons travelling during the entire night or day of the census and persons who work night shifts.

5.61. If each person is to be entered on the questionnaire only at the place where he or she was present at the time of the census, the topic needs to be investigated separately for each person.

5.62. Information on the place where each person was present should be collected with enough locational details to enable tabulation to be made for the smallest geographic subdivisions required by the tabulation plan and to meet the requirements of the database within the cost limits and operational procedures required to code to a fine degree of detail.

5.63. For countries relying on administrative data sources for their census, the concept of “present at the time of the census” may not be relevant. In such cases, and especially when the census population has legal implications, it's recommended to define the population based on usual residence.

3. Place of birth (core topic)

5.64. Information on the place of birth is a major input to the development of policies relating to migration and the related issues of service delivery to migrants. For the purposes of measuring internal migration, migrants are defined as those persons who usually are residing in a civil division of the country at the time of the census, but were previously resident outside that division. That is, movements within the civil division should not be regarded as being migratory.

5.65. The *place of birth* for those persons born within the country is the civil division in which the person was born; for those born in other countries, it is the country of birth. For persons born in the country (the native-born population), the concept of place of birth usually refers to the geographic unit where the mother of the individual resided at the time of the person’s birth and not the hospital/health facility in which the birth occurred. In some countries, however, the place of birth is defined as the geographic unit in which the birth actually occurred. It is recommended to use the place of birth as where the mother of the individual resided at the time of the person’s birth. If this concept is not used, the country should explain the definitions it uses in both the census enumerator instructions and in the census reports to aid the interpretation of the data.

5.66. The collection of information distinguishing between the native-born population and those born elsewhere (foreign-born) is necessary where any enquiry on place of birth is made. Even countries where the proportion of foreign-born population is insignificant, and who may only be interested in information on the place of birth of the native-born population, must first separate the native-born from the foreign-born population. It is therefore recommended that place of birth be asked of all persons. In countries that combine the questions on place of birth and country of birth (where the latter is used to measure international migration), the guidance on the country of birth (see paragraphs 5.111–5.115) should apply.

5.67. Information on the place of birth of the native-born population is usually used primarily for the investigation of internal migration. For countries that have been recently formed from parts of previously separate entities, however, such information may be of use in assessing the relative size of the population segments from each of those entities and their distribution throughout the country.

5.68. Information on whether or not a person is “born in the country” captures the population according to the boundaries at the time of the census. Using the “born in the country” concept would account for individuals who may have been affected by changes to a country’s boundary.

5.69. For the purposes of measuring internal migration, it is usually sufficient to collect information only on the major civil division (state, province or department, for example) in which the place of birth is located. If desired, more detailed information on the subdivision of a specific locality can be collected and used for accurate coding of the major division or for presenting data for smaller areas.

5.70. However, for more detailed studies of internal migration, data on the place of birth of the native-born population even in terms of major civil divisions may not be adequate. For better

understanding of the movements of people since birth it ~~is~~ may be necessary to collect information at the smallest possible geographic level, bearing in mind that:

(a) The boundaries of administrative units such as cities and other civil divisions will change over time, which may give rise to ambiguity in data reported;

(b) The costs of coding the reported data to these smaller units may be prohibitive, especially where there are many units and the population is highly mobile.

To overcome the first problem, to the extent possible, both national and subnational boundaries should refer to the boundaries applying at the time of the census. Countries must address the second problem in the light of their own circumstances, bearing in mind the reduced value of place of birth as a measurement of internal migration in a very mobile population.

5.71. It is recommended that, for the study of internal migration, the data on place of birth be supplemented by information collected on duration of residence (see paragraphs 5.72–5.74) and place of previous residence (see paragraphs 5.75–5.76) or of residence at a specified date in the past (see paragraphs 5.77 and 5.81).

4. Duration of residence (core topic)

5.72. The *duration of residence* is the interval of time up to the date of the census, expressed in complete years, during which each person has lived in (a) the locality that is his or her usual residence at the time of the census; or (b) the major or smaller civil division in which that locality is situated.

5.73. In collecting information on duration of residence, it should be made clear that the interest is in length of residence in the major or smaller civil division, or the locality, but not in the particular housing unit. The concept of duration of residence also relates to the most recent move to the current place of usual residence.

5.74. Data on the duration of residence have only limited value in themselves because they do not provide information on the place of origin of in-migrants. Therefore, when the topic is investigated, the place of previous residence should also be collected so that the data can be cross-classified.

5. Place of previous residence (core topic)

5.75. The *place of previous usual residence* is the major or smaller civil division, or the foreign country, in which the individual resided immediately prior to migrating into the civil division of present usual residence.

5.76. Data on the place of previous residence have only limited value in themselves because they do not provide information on the time of in-migration. Therefore, when the topic is investigated and included in the census, the duration of residence (see paragraphs 5.72–5.74) should also be included so that the data can be cross-classified. Alternatively, countries may choose to include a question on place of residence as a specified date in the past (see paragraphs 5.77–5.81 below).

6. Place of residence at a specified date in the past (core topic)

5.77. The *place of usual residence at a specified date in the past* is the major or smaller division, or the foreign country, in which the individual resided at a specified date preceding the census. The

reference date chosen should be that most useful for national purposes. In most cases, this has been deemed to be one year or five years preceding the census (or both of these time frames in cases where internal migration is of particular importance to users and resources are sufficient to code the data).

5.78. The former reference date provides information for statistics of both recent internal and international migration during a single year, while the latter may be more appropriate for collecting data for longer-term analysis of migration. When selecting the reference date the ability of individuals to recall with accuracy their usual residence one year or five years earlier than the census date should be considered. For countries conducting quinquennial censuses, the date of five years earlier can be readily tied in, for most persons, with the time of the previous census, but it should be noted that a one-year recall is likely to result in more accurate information than a five-year recall.

5.79. Some countries, however, may wish to use a different time reference than either one year or five years preceding the census because these intervals may present recall difficulties. In such circumstances the time reference should be one that can be associated with the occurrence of an important event that most people will remember.

5.80. For foreign-born persons, the collection of information on year of first or last arrival in the country is recommended (see "International migration characteristics", paragraphs 5.123–5.126).

5.81. However, no matter what previous date is used, provision must be made for the treatment of infants and young children who are resident at the time of the census but were not yet born at the earlier date. Tabulations of the data should indicate the nature of the treatment of this group.

7. **Reason for change of residence**

5.82. Some countries may wish to collect information on the *reason for change of residence* of persons who are usual residents in a civil division of the country at the time of the census but were previously resident in a different civil division. The information should refer only to the *main* reason for such a change of residence. If a person had moved from the place of previous residence for the purpose of education, for example, but at some subsequent point in time had obtained employment, the reason for change of residence would be education.

5.83. As different members of the household might change residence for different reasons, the information should be collected for each member of the household.

5.84. There are no recommendations on the specific reasons, and countries must establish their own reasons in accordance with their needs. However, examples of some common reasons[174] for change of residence include:

a) employment (including military service)
b) education and training

[174] Categories in line with EGRISS recommendations, para 155; EGRISS Compiler's Manual (Box A.2, pp. 20); and UNSD 2022 Handbook on Measuring International Migration, para 316 https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/international-migration/2022-UNSD-Handbook-Meas-Intern-Migration-E.pdf

c) marriage, family reunification or family formation,
d) forced displacement (refugees, asylum seekers, temporary protected status, etc.)
e) other

5.85. In collecting information on reason for change of residence, it should be made clear that it is with respect to place of previous residence (see paragraphs 5.75-5.76) or a sub-topic of the item on ever resided abroad (see paragraphs 5.127-5.128).

8. Total population (core topic)

5.86. For census purposes, the *total population* of the country consists of all the persons falling within the scope of the census. In the broadest sense, the total may comprise either all usual residents of the country or all persons present in the country at the time of the census. The total of all usual residents is generally referred to as the de jure population, while the total of all persons present is referred to as the de facto population.

5.87. In practice, however, countries do not usually fully achieve either type of count, because one or more groups of the population are included or excluded, depending on national circumstances. The general term used to describe the total might imply a treatment opposite to the one given to any of these groups. It is recommended, therefore, that each country describe in detail the figure accepted officially as the total, rather than simply label it as “de jure” or “de facto”.

5.88. The description should show clearly whether each group listed below was or was not included in the total. If the group was enumerated and identified as a separate group, its magnitude should be given; if it was not enumerated, an estimate of its size and the method of estimation should be given, if possible. If any of the groups listed below is not represented at all in the population, this fact should be stated and the magnitude of the group should be shown as “zero”. This may occur particularly with groups (a), (b, (d) and (n) described below (see also paragraph 5.47).

5.89. The groups to be considered are:
a) Nomads;
b) Persons living in areas to which access is difficult (e.g., remote areas);
c) Military, naval and diplomatic personnel and their families located outside the country;
d) Merchant seafarers and fishermen resident in the country but at sea at the time of the census (including those who have no place of residence other than their quarters aboard ship);
e) Civilian residents temporarily in another country as seasonal workers;
f) Civilian residents who cross a frontier daily to work in another country;
g) Civilian residents other than those in groups (c), (e) or (f) who are working in another country;
h) Civilian residents other than those in groups (c), (d), (e), (f) or (g) who are temporarily absent from the country;
i) Foreign military, naval and diplomatic personnel and their families located in the country;
j) Civilian foreign citizens temporarily in the country as seasonal workers;
k) Civilian foreign citizens who cross a frontier daily to work in the country;
l) Civilian foreign citizens other than those in groups (i), (j) or (k) who are working in the country;
m) Civilian foreign citizens other than those in groups (i), (j), (k) or (l) who are in the country temporarily;
n) Refugees, asylum seekers and internally displaced persons in camps;
o) Transients on ships in harbour or at airports at the time of the census.

5.90. In the case of groups (h) and (m), it is recommended that an indication be given to describe the criteria used (including with respect to the 12-month rule) in determining the presence in, or absence from, the country as temporary.

5.91. In those countries where the total population figure has been adjusted for under-coverage or over-coverage (or net under-coverage), both the enumerated figure and the estimated adjusted population figure should be shown and described. In such cases, documentation should be provided for users explaining how the total population figure from the census has been adjusted. Ideally, the detailed tabulations should be consistent with the adjusted population figures. However, where this is not possible, due, for example, to the costs of the methodology for undertaking these adjustments being prohibitive, the detailed tabulations will, of necessity, be based only on the actual enumerated population.

5.92. The use of both adjusted and non-adjusted population figures may be confusing for some users, particularly those unfamiliar with the adjustments made. For example, policy planners may find it more helpful to receive only the adjusted population data, while researchers are likely to require both figures for their analysis.

5.93. The population of each geographic unit of the country, like the total population of the country (see paragraph 5.86), may comprise either all usual residents of the unit (see paragraph 5.53) or all persons present in the unit at the time of the census (see paragraphs 5.58–5.59).

9. Locality (core topic)

5.94. For census purposes, a *locality* should be defined as a distinct population cluster (also designated as inhabited place, populated centre, settlement and so forth) in which the inhabitants live in neighbouring or contiguous sets of living quarters and that has a name or a locally recognized status. It thus includes fishing hamlets, mining camps, ranches, farms, market towns, villages, towns, cities and many other population clusters that meet these criteria. Localities can be defined using a national methodology and/or the methodology associated with the Degree of Urbanisation[175], particularly for international statistical comparison purposes. The census reports should explain which methodology has been applied, and if it departs from the concept of a distinct population cluster, provide an explanation to aid in the interpretation of the data.

5.95. Localities as defined above should not be confused with the smallest civil divisions of a country. In some cases, the two may coincide. In others, however, even the smallest civil division may contain two or more localities. On the other hand, some large cities or towns may contain two or more civil divisions, which should be considered as subdivisions of a single locality rather than separate localities.

5.96. A large locality (that is to say, a city or a town) is often part of an urban agglomeration, which may comprise the city or town proper together with a suburban fringe or heavily populated area lying outside, but adjacent to, its boundaries. The urban agglomeration is therefore not coterminous with the

[175] The Degree of Urbanisation is a harmonised methodology to facilitate international statistical comparisons and to classify the entire territory of a country along an urban-rural continuum. The Degree of Urbanisation classification defines cities, towns and semi-dense areas, and rural areas. See https://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/-/ks-02-20-499.

locality but is an additional geographic unit, which may include more than one locality. In some cases, a single large urban agglomeration may comprise several cities or towns and their suburban fringes. The components of such large agglomerations should be specified in the census results.

5.97. To classify localities by size, it is recommended to use a subset of the following size classes and to place an open-ended category at the two extremes:

- 49 inhabitants or less
- 50-99
- 100-249
- 250-499
- 500-999
- 1,000-2,499
- 2,500-4,999
- 5,000-9,999
- 10,000-24,999
- 25,000-49,999
- 50,000-99,999
- 100,000-249,999
- 250,000-499,999
- 500,000-999,999
- 1,000,000-2,499,999
- 2,500,000-4,999,999
- 5,000,000 or more inhabitants

5.98. For each locality, it is recommended to measure the share of population living in slums[176]. A slum may be defined by the Government or based on the UN-Habitat criteria below. Slum households may be defined as one in which the inhabitants lack one or more of the following[177]:

1) Lack of access to clean water;
2) Lack of access to improved sanitation;
3) Lack of sufficient living space;
4) Lack of housing durability; and,
5) Lack of security of tenure.

10. Urban and rural (core topic)

5.99. Urban and rural areas can be defined using a national methodology, which can be established in accordance with a country's own needs. National methodologies for classifying urban and rural areas often incorporate criteria such as population size, density, land use patterns, administrative boundaries, and access to basic services. In addition, for statistical comparisons, urban and rural areas can be delineated using the Degree of Urbanisation (DEGURBA) methodology which was endorsed by the UN

[176] UN-Habitat defines a slum household in operational terms, as lacking one or more of the following indicators: a durable housing structure; access to clean water; access to improved sanitation; sufficient living space; and secure tenure.
https://unhabitat.org/sites/default/files/2020/06/indicator_11.1.1_training_module_adequate_housing_and_slum_upgrading.pdf

[177] The topics of household characteristics, including lack of access to clean water, sanitation, sufficient living space, durability, and security of tenure, are detailed in Part Five, Chapter II of this publication.

Statistical Commission in 2020. For a full elaboration on the Degree of Urbanisation methodology, please see the manual[178].

5.100. A classification of areas as urban or rural should be done at the smallest administrative unit of the country, or the smallest census collection unit. The classification should be made, in addition to the national methodology, first and foremost, on a measure of population density and population size.

5.101. Some of the information required for classification as urban or rural may be provided by the census results themselves, while other information may be obtained from external sources. The use of information provided by the census (as, for example, the size class of the locality or the percentage of the population engaged in agriculture), whether alone or in conjunction with information from other sources, means that the classification will not be available until the relevant census results have been tabulated. If, however, the census plans call for the investigation of a smaller number of topics in rural areas than in urban areas or for a greater use of sampling in rural areas, the classification must be available before the enumeration takes place. In these cases, reliance must be placed on previous censuses and/or external sources of information, even if only to bring up to date any urban–rural classification that was prepared at an earlier date.

5.102. The usefulness of housing census data (for example, the availability of electricity or piped water) collected simultaneously with, or not too long before, the population census should be kept in mind. For assembling information from more than one source, the importance of a well-developed system of geocoding should not be overlooked.

5.103. The traditional distinction between urban and rural areas within a country has been based on the assumption that urban areas, no matter how they are defined, provide a different way of life and usually a higher standard of living than are found in rural areas. In many developed countries this distinction has become blurred, and the principal difference between urban and rural areas in terms of the living standards tends to be the degree of population concentration or density. On the other hand, the differences between urban and rural ways of life and standards of living remain significant in developing countries, but even there rapid urbanization in these countries has created a great need for information related to different types of urban areas.

5.104. Hence, although the traditional urban–rural dichotomy is still needed, a classification by Degree of Urbanisation can usefully supplement the dichotomy as it captures characteristics related to the rural-urban continuum from the most sparsely settled areas to the most densely built-up localities with seven classes. When using the Degree of Urbanisation, it is recommended to use the three classes at level 1 (i.e., with two classes within urban areas i.e. urban and semi-urban).

5.105. The Degree of Urbanisation should be applied first to a population grid and then to small administrative or statistical spatial units. To effectively delineate urban and rural areas, a population grid should be created, ideally using point-based census data. Where such data is unavailable, the DEGURBA methodology can be applied to estimate population distribution within smaller geographic units. It's crucial to recognize the limitations of population grid estimation, as it may not perfectly align with actual

[178] EC, FAO, ILO, OECD, UN-Habitat and the World Bank (2021). Applying the Degree of Urbanisation: A methodological manual to define cities, towns and rural areas for international comparisons — 2021 edition https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-02-20-499

population distribution, especially in areas with significant seasonal variation or specific land use patterns (e.g., non-residential resort areas).

B. International migration characteristics

5.106. Interest in the movement of people across national boundaries, namely, international migration, has steadily grown among countries concomitant with the increase in international migration. The decision to collect and disseminate information on international migration in a census is dependent upon a number of considerations and national circumstances, including, for example, the national needs for such data. Data on international migrants could provide information on the diversity of a population and can serve to identify subgroups of a population. The present section on international migration supplements and expands the topic "geographic and internal migration characteristics", which is covered above.

5.107. The revised conceptual framework on international migration and mobility statistics[179] deals with both international migration flows and international migrant stock, identifies population censuses as being one of the main sources for collecting data on international migrants and their characteristics, and recognizes the value of the population census to produce an estimate of annual flows when a question on place of residence one year before the census is included. It emphasizes consistency and alignment between international migration flows and immigrant populations (stocks) and international temporary mobility and temporary populations for improved accuracy in international migration data.

5.108. This section is concerned chiefly with the topic of international migrant stock as derived from field-based population censuses.

5.109. In the revised conceptual framework on international migration statistics, four subgroups of the resident population critical to understanding immigrant and emigrant populations are defined: native-born citizens, native-born foreign citizens (including stateless or those without citizenship), foreign-born citizens and foreign-born foreign citizens (including stateless or those without citizenship).

5.110. Consequently, for the study of international migration and migrants, identifying all four subgroups of the resident population of the country of enumeration, and producing statistics for all four subgroups, is critical. In order to identify members of those groups, two items must be recorded in the census: (a) the country of birth, and (b) the country of citizenship. In addition, it is important to record the year of arrival in the country of enumeration so as to establish the length of stay of international migrants residing in the country.

1. Country of birth (core topic)

5.111. *Country of birth* is the country in which the person was born. This information is generally collected from the same question used to determine 'place of birth' covered at paragraphs 5.64-5.71 above. The concept of country of birth may refer to either the country in which the mother was usually residing at the time of the person's birth or the country where the physical birth actually took place. Either concept can be used depending on the information needs of the country; but for the purpose of

[179] Conceptual frameworks and Statistical Concepts and Definitions on International Migration (2021). https://unstats.un.org/unsd/demographic-social/migration-expert-group/task-forces/TF2-ConceptualFramework-Final.pdf

measuring international migration, it is recommended to use the country of birth as where the mother of the individual resided at the time of the person's birth. Each country should explain which definition it used in the census. It should be noted that the country of birth of a person is not necessarily the same as his or her country of citizenship, which is a separate census topic dealt with below. It is recommended that country of birth be asked of all persons to distinguish the native-born from the foreign-born population. The collection of this information is necessary even in countries where the proportion of foreign-born population is small. For the foreign-born population, the collection of information on the specific country of birth is recommended so as to permit the classification of the foreign-born population by country of birth. For respondents who are born outside the country of enumeration and cannot identify their country of birth, at least the continent or region where that country is located should be ascertained.

5.112. For purposes of both internal consistency and international comparability, it is recommended that information on the country of birth be recorded according to national boundaries existing at the time of the census. Information on the year of arrival in the country (see paragraph 5.123 below) can be used to identify persons who owe their status of foreign-born to changes in national boundaries. It is essential that the coding of information on the country of birth be done in sufficient detail to allow for the identification of all relevant countries of birth.

5.113. For purposes of coding, it is recommended that countries use the numerical coding system presented in *Standard Country or Area Codes for Statistical Use*.[180] The use of standard codes for classification of the foreign-born population according to the country of birth will enhance the usefulness of such data, including an international exchange of foreign-born population statistics among countries. The option "Unknown" should be added to the codes for country of birth. If countries decide to combine countries into broad groups, it is recommended that the standard regional and subregional classifications identified in the above-mentioned publication be adopted.

5.114. Countries with a significant number of international migrants may wish to collect information on the **country of birth of parents** (both father and mother), in which case the information should be asked of all respondents following the same guidelines given for country of birth. The decision to collect and disseminate information on country of birth of parents in a census is dependent upon a number of considerations and national circumstances, including for example the suitability and sensitivity of asking such a question that relates to persons who may not be in the country in which the census is taking place.

5.115. Information on the country of birth of parents can be used, in combination with information on the country of birth of the enumerated person, to identify native-born children of the foreign-born population (the so-called "second generation") and to study the integration processes and outcomes of migrants and their descendants. Moreover, in countries that have experienced return migration, information from this topic allows the identification of foreign-born children of native-born parents.

2. Country of citizenship (core topic)

5.116. *Country of citizenship* is defined as the country an individual is a citizen of and with which the individual enjoys a particular legal bond, acquired by birth, naturalization, marriage or some other

[180] United Nations, *Standard Country or Area Codes for Statistical Use*, http://unstats.un.org/unsd/methods/m49/m49.htm.

mechanism. A citizen is a legal national of the country of enumeration; a non-citizen may be a foreign-citizen (that is a citizen of another country), or a stateless person or a person with unknown or undetermined citizenship status. Because the country of citizenship is not necessarily identical to the country of birth, both items should be collected in a census.

5.117. Information on the country of citizenship is particularly important for foreign citizens. It is important to record country of citizenship as such and not to use another concept (such as nationality) to indicate citizenship, since some of those concepts may also be used to designate ethnic groups.

5.118. It is essential that the coding of information on country of citizenship be done in sufficient detail to allow for the individual identification of all countries of citizenship that are represented among the foreign population in the country. For purposes of coding, it is recommended that countries use the numerical coding system presented in *Standard Country or Area Codes for Statistical Use*. The use of standard codes for classification of the foreign population by country of citizenship will enhance the usefulness of such data and permit an international exchange of information among countries on their foreign populations. If countries decide to combine countries of citizenship into broad groups, it is recommended that the standard regional and subregional classifications identified in the above-mentioned publication be adopted. Two additional categories should be listed: 1) stateless or no citizenship/nationality status, and 2) unknown citizenship/nationality status. IROSS[181] provides a suitable framework for estimating the combined category of 'stateless persons and those without a recognized nationality status' within censuses. However, clearly differentiating between individuals who are stateless and those whose nationality is unclear presents a significant challenge in most census operations.

5.119. The reliability of reported citizenship may be doubtful in the case of persons whose citizenship has recently changed as a result of territorial changes, or among the population of some newly independent countries where the concept of citizenship may have only recently become important. Clear guidelines issued by the national statistical authority can help improve the quality of the data collected. As an aid to the analysis and interpretation of the results, notes on the likelihood of these and other possible causes of misstatement should accompany tabulations based on citizenship.

5.120. Enumeration and processing instructions should provide clear guidance on the treatment of stateless persons, persons with dual nationality, persons in the process of naturalization and any other groups with ambiguous citizenship. The treatment of these groups should be described in the census reports and included in the metadata for accompanying tabulations.

5.121. In cases where people have more than one citizenship and where this information is useful for decision-making, details may be collected on whether the person holds one or multiple citizenship. If this information is to be published, care should be taken to explain how the possibility of people being included in the table more than once affects the marginal totals on the table. Usually, however, it may be more practicable for tabulations by citizenship to refer to one citizenship only. Thus, persons with multiple citizenships should be allocated to a single “primary” citizenship, for example by giving precedence to the citizenship of the “home” country.

[181] IROSS, see paragraphs 154 and 165, more info https://egrisstats.org/recommendations/international-recommendations-on-statelessness-statistics-iross/

3. Acquisition of citizenship

5.122. In addition to collecting information on citizenship, for countries where the population includes a significant proportion of naturalized citizens it may be important to collect information on the method of acquisition of citizenship so as to enable the classification of the population into (a) citizens by birth; (b) citizens by naturalization whether by declaration, option, marriage or other means; and (c) non-nationals or foreign citizens (including stateless or those without citizenship). In such countries it may also be useful to ask questions on previous citizenship and year of naturalization.

4. Year or period of arrival in the country (core topic)

5.123. *Year or period of arrival in the country* refers to the calendar year and month of arrival to the country of enumeration of any person who has been resident abroad. This information enables the calculation of the number of completed years between the time of arrival in the country and the time of enquiry, usually the census date. Information on the month and year of arrival also provides the flexibility of classifying foreign-born persons by period of arrival in terms of any prespecified period, such as 1995–1999, 2000–2004, 2005-2009 and so forth. It is thus recommended that the period of arrival be shown, in any tabulation in which the variable appears, in terms of the actual year of arrival.

5.124. It is possible to collect information on either the date of first arrival in the country or the date of the most recent arrival in the country. Each has its own advantages and disadvantages. In making the choice of which information to collect, countries should be guided first and foremost by their policy and user needs.

5.125. Information on time since arrival can also be collected by asking how many years have elapsed since the time of arrival, instead of in what calendar year and month the person arrived. However, use of such a question is not recommended because it is likely to yield less accurate information.

5.126. Note that information on the year and month of arrival is focused mainly on persons born outside the country of enumeration, that is to say, persons who must have arrived in that country at some time after their birth. However, it should be noted that the phenomenon of "international return migration" is becoming increasingly common, and countries that have population groups that maintain links to other countries, migrating to or from another country at different life stages (for example, as students or pensioners), may have an interest in collecting information on returning migrants: in this case, the question on year and month of arrival could also be asked of native-born respondents who have ever lived in another country. In addition, it might also be important to collect information on previous country of residence for persons who have ever lived abroad.

5. Ever resided abroad

5.127. The topic "ever resided abroad" focuses on all persons who have ever resided outside the current country of usual residence, regardless of country of birth or citizenship. To collect information on this topic, individuals should be classified depending on whether or not they have ever had a usual residence abroad using the 12-month rule described in paragraph 5.36. Countries may choose to collect information on this topic from (a) citizens only; (b) native born persons only; or (c) all persons. The choice depends on which population subgroup's arrival or return is most relevant to policy, and on

whether a topic on country of birth or country of citizenship has been included in the census.[182] When asked of all persons, the topic allows the identification of all international migrants.[183] When asked of native-born persons, it can serve more specifically to measure the stock of returned international migrants, defined as "persons who previously resided in the country of enumeration who emigrated and subsequently came back to live in the country and stayed or intended to stay for the minimum duration required for residence."[184].

5.128. For those who have ever resided abroad, the year of arrival in the current country of residence should also be collected.[185] Through inclusion of a question on the respondents' most recent year or period of arrival or return, most countries differentiate between migrants arriving or returning recently and those who arrived or returned at a more distant time in the past. The year of arrival or return can also be matched with the date of birth to permit calculation of the age at which the person arrived or returned to live in the census country, information that can be of great value for research and policy.

6. **Reason for international migration**

5.129. The topic on "reason for international migration" refers to the main reason that drove the respondent to undertake the most recent migratory move.[186] When included in the census, it may be most appropriate to include this topic as a sub-topic of the item on ever resided abroad (see paragraphs 5.127-5.128). The question should be asked with explicit reference to the main reason for (last/most recent arrival (or return) to the country, to avoid potential conflation with any subsequent internal migration moves which would be captured when asked in reference to a move from the previous place of usual residence.

5.130. In all cases, it is recommended that only the main reason for migration, as self-declared by the respondent be recorded. The information should be collected at the individual level, for each household member, as the main reason for migration may differ among the members. For example, one member may have moved for employment reasons, while another may have moved as an accompanying family member.

[182] From 2022 UNSD Handbook on Measuring International Migration through Population censuses, para. 313 https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/international-migration/2022-UNSD-Handbook-Meas-Intern-Migration-E.pdf

[183] From UNECE CES 2020 Census recommendations, para 662. https://unece.org/DAM/stats/publications/2015/ECECES41_EN.pdf

[184] From UN Expert Group on Migration Statistics. https://unstats.un.org/unsd/demographic-social/migration-expert-group/task-forces/TF2-ConceptualFramework-Final.pdf

[185] From 2022 UNSD Handbook on Measuring International Migration through Population censuses, para. 313 https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/international-migration/2022-UNSD-Handbook-Meas-Intern-Migration-E.pdf

[186] From UNECE CES 2020 Census recommandations para 670. https://unece.org/DAM/stats/publications/2015/ECECES41_EN.pdf

5.131. There are no internationally agreed recommendations on the main reasons to include as answer categories. Countries considering this topic for inclusion should establish nationally relevant reasons in accordance with their needs. Examples of some of the more common reasons for migration include[187]:
(a) work (including to invest or start a business) or employment (including military assignments)
(b) education and training
(c) marriage, family reunification or family formation
(d) forced displacement (refugees, asylum seekers, temporary protected status, others)
(e) deportation/involuntary return
(g) retirement
(e) other

5.132. While the main reason for migration can be a topic of interest to governments and researchers, this information alone may not be sufficient to assess the various drivers of migration, as persons may migrate for multiple reasons that include both push and pull factors. This should be taken into consideration when using information on the main reason for international migration for assessments of population movements due to complex factors such as economic conditions, climate change or forced displacement.

5.133. When a question on main reason for international migration is included in the census, information on the year of arrival (see paragraphs 5.123-5.126) should also be collected to enable analysis of changes across periods and/or over the life-course.

C. Household and family characteristics

5.134. In considering the topics related to household characteristics, it is important to be aware of the differences between the concepts of a “household” and a “family” as used herein.

5.135. A *household* may be:

(a) A one-person household, that is to say, a person who makes provision for his or her own food or other essentials for living without combining with any other person to form part of a multiperson household; or

(b) A multiperson household, that is to say, a group of two or more persons living together under the same roof who make common provision for food, housing or other essentials for living. The persons in the group may pool their resources and have a common budget; they may be related or unrelated persons or a combination of persons both related and unrelated. This arrangement exemplifies the “housekeeping” concept.

Some countries use a concept different from the housekeeping concept, namely, the “household dwelling” concept, which regards all persons living in a single housing unit as belonging to the same household. According to this concept, there is one household per occupied housing unit, that is the number of occupied housing units and the number of households occupying them are equal and the locations of the housing units and households are identical. On the other hand, in countries that use the “housekeeping” concept, the number of households may be greater than the number of dwelling units.

[187] Categories in line with EGRISS recommendations, para 155; EGRISS Compiler's Manual (Box A.2, pp. 20); UNSD 2022 Handbook on Measuring International Migration, para 316: https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/international-migration/2022-UNSD-Handbook-Meas-Intern-Migration-E.pdf

Countries should specify in their census reports whether they used the “housekeeping” or the “household dwelling” concept of a private household.

In countries that use register-based census methods, the concept of "household dwelling" is generally used.

5.136. A household may be located in a housing unit (see paragraph 5.487) or in a set of collective living quarters such as a boarding house, a residential hotel or a camp, or may comprise places such as construction sites, caravans etc.

5.137. The *family* within the household is defined as those members of the household who are related, to a specified degree, through blood, adoption, or marriage. The degree of relationship used in determining the limits of the family in this sense is dependent upon the uses to which the data are to be put and so cannot be established for worldwide use. See paragraph 5.155 for a definition of the family nucleus.

5.138. Although in practice most households are composed of a single family consisting of a couple with or without children or of one or both parents and their children, it should not be assumed that this identity always exists; census tabulations should therefore clearly indicate whether they relate to households or to families within households.

5.139. From the definitions of “household” and “family”, it is clear that these are different concepts that cannot be used interchangeably in the same census. The differences between the household and the family are that:

a) A household may consist of only one person but a family must contain at least two members;

b) The members of a multiperson household need not be related to each other, while the members of a family must – by definition – be related.

5.140. A household can contain more than one family, or one or more families together with one or more non-related persons, or it can consist entirely of non-related persons. A family typically will not comprise more than one household. However, the existence of polygamous families in some countries, as well as shared child custody and support arrangements provided by family members living in other households (like grandparents) in others, means that individual countries should decide how best to derive and report data on families.

5.141. It is recommended that the household be used as the unit of enumeration (as defined in paragraphs 2.35–2.40) and that the family be a derived topic only. The place of usual residence is recommended as the basis for assigning persons to households where they usually reside. Where the de facto approach is used as the method of enumeration (see paragraphs 2.57–2.64), household lists should, where feasible, also include usual residents temporarily absent. In published reports, countries should indicate whether household information refers to usual residents and also what the time limits are in respect of being included or excluded as a usual resident. For a more detailed discussion on the difficulty of collecting information on place of usual residence, see paragraphs 5.52–5.57. In countries that produce household information based on register data, definitions can be made by setting limits on the minimum ages of the people residing (e.g., in cases where 3 children under the age of 18 live together, additional inquiries should be made).

1. Relationship to the reference person of household (core topic)

5.142. In identifying the members of a household (as defined in paragraphs 5.135–5.136), it is useful to identify first the household reference person and then the remaining members of the household according to their relationship to the reference person. Countries may use the term they deem most appropriate to identify this person (such as, for example, household reference person, head of household, householder, among others) if the person so identified is used solely to determine relationships between household members. It is recommended that each country present, in published reports, the concepts and definitions that are used. It is noted that when administrative data sources are used, deriving the relationship of household members to the reference person may present several challenges.

5.143. With respect to selecting the household reference person, it is important to specify criteria for choosing that person in relation to whom household members would be best distinguished, especially in polygamous, multi-family and other households, such as those composed only of siblings without a parent and those composed entirely of unrelated persons. This information should be included in training materials and instructions to enumerators.

5.144. The traditional notion of head of household assumes that most households are family households and that one person in such family households has primary authority and responsibility for household affairs and is, in the majority of cases, its chief economic support. This person is then designated as the head of household.

5.145. Where more than one member of the household consider themselves to be equal in household authority and responsibility and may share the economic support of the household, the concept of a single head of household may no longer be considered valid even for family households. For the relationship among members of the household to be determined under these circumstances, it is essential that either:

(a) the members of the household designate one among them as a reference member with no implication of headship; or
(b) provision be made for designation of joint headship where desired.

In any case, it is important that clear instructions be provided in the census as to how this situation is to be handled.

5.146. Even in countries where the traditional concept of head of household is still relevant, it is important to recognize that the procedures followed in applying the concept may distort the true picture, particularly with regard to female heads of households. The most common assumption that can distort the facts is that no woman can be the head of any household that also contains an adult male. Enumerators and even respondents may, in some countries, simply take such an assumption for granted.

5.147. This common sex-based stereotype often reflects circumstances that may have been true in the past but are no longer valid, insofar as the household and economic roles of women continue to change. It is therefore important that clear instructions be provided as to who is to be treated as the reference person of the household so as to avoid the complications of enumerator or respondent preconceptions on the subject and the bias that such preconceptions may create. The procedure to follow in identifying

a reference person when the members of the household are unable to do so should be clear and unambiguous and should avoid sex-based bias. Where alternative definitions are used, this should be made explicitly clear in the census questionnaire and in the tabulated census results.

5.148. The selection of one reference person in a household to whom all other persons in the household report, or designate, their relationship requires careful consideration. In the past the person considered to be the "head of the household" was generally used as the reference person, but this concept is no longer considered appropriate in many countries. It has also sometimes been proposed that the person designated as the reference person should be the oldest person in the household or the one who contributes the most income. However, given that the primary purpose of the question is to assign family status and to assign individuals into families, both of these approaches have weaknesses. The automatic selection of the oldest person may be undesirable because in multigenerational households many explicit kin relationships can be reported where the reference person is selected from the middle generation. Similarly, the selection of the person with the highest income may be a person who will not solicit the broadest range of explicit kin relationships.

5.149. Given below is some guidance on the selection of the reference person, which will yield some explicit kin relationships:

(a) Either spouse of a married couple living in the household (preferably from the middle generation in a multigenerational household);
(b) Either partner of a consensual union couple living in the household where there is no married couple present (where applicable);
(c) The parent, where one parent lives with his or her children of any age; or
(d) Where none of the above conditions apply, any adult member of the household may be selected.

Note that these categories are neither comprehensive nor mutually exclusive.

5.150. After identification of the reference member of the household, each of the remaining members of the household should be distinguished in relation to that person, as appropriate, as one of the following:

(a) Spouse,
(b) Partner in consensual union (cohabiting partner), where applicable,
(c) Child,
(d) Spouse of child,
(e) Grandchild or great-grandchild,
(f) Parent (or parent of spouse),
(g) Other relative,
(h) Domestic worker, or
(i) Other person not related to the head or other reference member.

Where this classification is considered too detailed for successful collection of the information, categories (f) and (g) may be consolidated as "Other relative" and (h) and (i) can be consolidated as "Other unrelated person".

5.151. As an aid to the identification of family nuclei (as defined in paragraphs. 5.155–5.157) within the household, it might be helpful if persons were recorded on the census questionnaire to the extent possible in the order of nuclear relationship. Thus, the first person entered after the head or other reference person would be the spouse or cohabiting partner of that person, followed by unmarried children and then by married children, their spouses or cohabiting partners and children. For polygamous households, the order of entry could be such that each wife or cohabiting partner and her unmarried children appeared in succession.

5.152. For estimating fertility by the *own children* method (see paragraph 5.237), the natural mother of each child under 15 years of age should be identified if she appears in the same questionnaire as her child. One way of doing this is to provide the line number of the mother alongside that of the child, if both are living in the same household. The information is not relevant for stepchildren, adopted children or foster children under permanent or temporary care.

5.153. In order to meet increased data needs on households and families, countries may wish, while conducting their population censuses, to collect more detailed information on relationships. In households where the relationship structure is complex, including those with foster children, obtaining accurate information on the relationships between household members may be difficult. Some countries may supplement information on relationship to the reference person of the household with information on direct relationships between household members by, for instance, relating a child to its parents even when neither parent is the reference person of the household. Enumerators should be encouraged to probe for a clear relationship (such as child, niece or aunt). The recording of non-specific responses such as "relative" should be avoided. It is recommended that specific guidance be provided on acceptable responses, that relationships be specified completely in the census questionnaire, and that any precoded categories used should be sufficiently detailed to produce desired outputs.

2. Household and family composition (core topic)

5.154. Household and family composition can be examined from different points of view, but for census purposes it is recommended that the primary aspect considered should be that of the family nucleus.

5.155. A *family nucleus* is of one of the following types (each of which must consist of persons living in the same household):

(a) A married or cohabiting couple without children;
(b) A married or cohabiting couple with one or more unmarried children;
(c) A father with one or more unmarried children;
(d) A mother with one or more unmarried children.[188]

5.156. The concept of family nucleus as defined above limits relationships between children and adults to direct (first-degree) relationships, that is to say, between parents and children. In some countries, numbers of skip-generation households, that is to say, households consisting of grandparent(s) and one or more grandchildren with no parent of those grandchildren present, are considerable. Therefore,

[188] In countries where a different definition of family nucleus is used, it should be clearly stated in the census report.

countries may include such skip-generation households in their family nucleus definition, and the census report should clearly state whether or not skip-generation households are included in the family nucleus definition.

5.157. The family nucleus is identified from the answers to the question on relationship to the reference member of the household, supplemented where necessary by information on name and marital status. The identification of offspring and their mother and the order in which persons are entered in the questionnaire may be of additional assistance in this respect. The identification of family nuclei is likely to be more complete in de jure than in de facto enumerations, because the latter do not take account of temporarily absent household members who may constitute part of a family nucleus.

5.158. For census purposes, a child is any unmarried individual, regardless of age, who lives with his or her parent(s) and has no children in the same household. Consequently, the definition of a child is primarily a function of an individual's relationship to other household members, regardless of age. In accordance with this definition, a household consisting of a married or cohabiting couple with two never-married children, divorced son, and a married daughter and her husband would be considered to be composed of two family nuclei, with the divorced child being regarded as a member of the parents' family and the married daughter and son-in-law as a second family. As used here, the term "child" does not imply dependency, but rather is used to capture household living arrangements of persons who are in a parent-child relationship. Countries need to be clear in their metadata how they treat foster and adopted children.

5.159. The family nucleus does not include all family types, such as brothers or sisters living together without their offspring or parents, or an aunt living with a niece who has no child. It also excludes the case of a related person living with a family nucleus as defined above, for example, a widowed parent living with her married son and his family. The family nucleus approach does not, therefore, provide information on all types of families. Countries may extend the investigation of families beyond that of the family nucleus, in accordance with their own interests.

5.160. Households could be classified by type according to the number of family nuclei they contain and the relationship, if any, between the family nuclei and the other members of the household. The relationship should be through blood, adoption, or marriage, to whatever degree is considered pertinent by the country (see paragraph 5.153). Given the complexity of this item, it is important that information on relationship to the household reference person be properly processed. The types of household to be distinguished could be:

- *(a)* *One-person household*;
- *(b)* *Nuclear household*, defined as a household consisting entirely of a single family nucleus. It may be classified into:
  - *(i)* Married couple family:
    - a. With child(ren);
    - b. Without child(ren);
  - *(ii)* Partners in consensual union (cohabiting partners):
    - *a.* With child(ren);
    - *b.* Without child(ren);
  - *(iii)* Father with child(ren);
  - *(iv)* Mother with child(ren);

(c) *Extended household*, defined as a household containing one or more family nuclei and their relatives: either one or more family nuclei (related to each other), with other persons related to any of the family nuclei in the same household (see paragraph 4.146 in Revision 3 of the Principles and Recommendations). The subdivisions of this category may be modified to suit national circumstances.

(d) *Composite household*, defined as a household consisting of one or more family nuclei (related or unrelated to each other) plus other persons, some of whom are related to at least one of the nuclei and some of whom are not related to any of the nuclei (see paragraph 4.146 in Revision 3 of the Principles and Recommendations). The subdivisions of this category may be modified to suit national circumstances.

(e) Other;

(f) Unknown or not stated.

5.161. In the census tabulations, all countries should at least distinguish between one-person, nuclear, extended and composite households. Where feasible, some or all of the subcategories shown above should also be distinguished, although countries may find it appropriate to modify the classification according to national circumstances. For example, in countries where almost all households contain only one family nucleus at most, the distinction between nuclear, extended and composite households may be applied only to households containing one nucleus or no nucleus; multinuclear households may then be shown as an additional category without any further classification by type. In countries where multinuclear households are comparatively common, further breakdowns of extended and composite households, distinguishing between those with three, four or more family nuclei, may be helpful.

3. Household and family status

5.162. For purposes of determining household and family status and identifying how a person relates to other household or family members, persons may be classified according to their position in the household or family nucleus. Classifying persons according to household and family status has uses in social and demographic research and policy formulation. Census data could be presented according to both household and family status for a variety of purposes. Although status itself is based on information derived from responses to the item on relationship to the head or other reference member of the household and other items, the classification of persons by their household and family status is a different approach from the traditional one of classifying household members solely according to their relationship to the head or reference person. The following household and family status classifications illustrate how such an approach may be used.[189] Care should be taken at the planning stages to relate this item to the classification of households by type as recommended in paragraph 5.160.

Persons living in households are classified by household status as:

1 Person in a household with at least one family nucleus
- 1.1 Married spouse
- 1.2 Partner in consensual union (cohabiting partner)
- 1.3 Lone mother[190]

[189] To date, only the population and housing census recommendations for the Economic Commission for Europe region contain household and family status classifications.

[190] Person living with children, without spouse or partner.

1.4 Lone father[191]
1.5 Child living with both parents
1.6 Child living with lone mother
1.7 Child living with lone father
1.8 Not a member of a family nucleus
1.8.1 Living with relatives
1.8.2 Living with non-relatives
2 Person in a household with no family nucleus
2.1 Living alone
2.2 Living with others[192]
2.2.1 Living with sibling(s) only
2.2.2 Living with other relatives only
2.2.3 Living with non-relatives only
2.2.4 Living with a combination of siblings, other relatives, and non-relatives

Persons are classified by family status as:[193]

1 Spouse
1.1 Husband
1.1.1 With child(ren)
1.1.2 Without a child
1.2 Wife
1.2.1 With child(ren)
1.2.2 Without child
1.3 Cohabiting partner
1.2.1 With child(ren)
1.2.2 Without child
2 Lone parent
2.1 Male
2.2 Female
3 Child
3.1 With both parents
3.2 With lone parent
3.2.1 With lone father
3.2.2 With lone mother
4 Not member of a family nucleus
4.1 Relative of husband or wife
4.1.1 Parent of husband or wife
4.1.2 Sibling of husband or wife
4.1.3 Other relative of husband or wife
4.2 Non-relative

D. Demographic and social characteristics

[191] Person living with children, without spouse or partner.
[192] The subdivisions in this category should be modified to suit national circumstances.
[193] The subdivisions in this category should be modified to suit national circumstances.

5.163. Sex and age are considered to be the most basic of all demographic variables. Of all the topics included in population censuses, *sex* and *age* are more frequently cross-classified with other characteristics of the population than are any other topics. Apart from the importance of the age-sex structure of the population in itself, accurate information on these two topics is fundamental to the great majority of the census tabulations.

5.164. A very important use of census data on the sex and age composition of the population is the evaluation of the data, especially with respect to coverage. The variables are therefore very crucial, and it is important that this information be reported in respect of every person for whom census information has been collected.

5.165. It is therefore recommended that where this information is incomplete it should be imputed for census purposes rather than being reported as "not stated". Possible difficulties in securing accurate age data are often not recognized because the topic appears to be a simple one. The difficulties associated with this topic are therefore highlighted in paragraphs 5.170–5.181 below.

1. Sex (core topic)

5.166. The sex of every individual should be recorded on the census questionnaire for those countries that collect their census information in this way. The disaggregation of data by sex is a fundamental requirement for gender statistics. For many socioeconomic and demographic characteristics that could be collected through a census, such as education, economic activity, marital status, migration, disability and living arrangements, there are generally variations by sex. The successful planning and implementation of gender-sensitive policies and programmes requires the disaggregation of data by sex to reflect problems, issues and questions related to both men and women in society.

5.167. Sex, together with age, represents the most basic type of demographic information collected about individuals in censuses and surveys, as well as through administrative recording systems, and the cross-classification of these data with other characteristics forms the basis of most analyses of the social and demographic characteristics of the population, as it provides the context within which all other information is placed.

5.168. Sex and gender are different concepts, although the two are often conflated or used interchangeability. Sex is a biological attribute whereby persons are traditionally classified as either male or female, while gender is a social construct that refers to the roles, behaviours, activities, and attributes that a particular society associates with being a man or a woman.

5.169. To reflect social and cultural changes and, in some cases, meet legislative requirements, some countries are now collecting data on gender identity in addition to information about the sex of individuals. This allows respondents to express their identity beyond traditional binary options. Including questions on gender identity in a census requires careful consideration and testing, as there are no international standards for measuring this fluid and evolving concept.

2. Age (core topic)

5.170. *Age* is the interval of time between the date of birth and the date of the census, expressed in completed solar years. Every effort should be made to ascertain the precise and accurate age of each person, particularly of children and older persons.

5.171. Information on age may be secured either by obtaining the date (year, month and day) of birth or by asking directly for age at the person's last birthday.

5.172. The first method yields more precise information and should be used whenever circumstances permit. It also allows for the calculation of age at a reference date other than census day for the purposes, for example, of deriving annual census-based mid- or end-year population estimates. If neither the exact date nor the month of birth is known, an indication of the season of the year can be substituted if this information can be easily recorded. The question on date of birth is appropriate wherever people know their birth date, whether in accordance with the solar calendar or a lunar calendar, or whether years are numbered or identified in traditional folk culture by names within a regular cycle. It is extremely important, however, that there should be a clear understanding between the enumerator and the respondent about which calendar system the date of birth is based on. If there is a possibility that some respondents will reply with reference to a calendar system different than that of other respondents, provision must be made in the questionnaire for noting the calendar system that has been used. It is not advisable for the enumerator to attempt to convert the date from one system to another. The necessary conversion can be best carried out as part of the data-editing work.

5.173. Where the information is taken from administrative data sources, date of birth is usually more accurately recorded.

5.174. The direct question on age is likely to yield less accurate responses for a number of reasons. Even if all responses are based on the same method of reckoning age, there is the possibility of a misunderstanding on the part of the respondent as to whether the age wanted is that at the last birthday, the next birthday or the nearest birthday. In addition, asking a direct question on age can result in occurrences, with comparative ease, of rounding to the nearest age ending in zero or five, providing estimates not identified as such and deliberate misstatements. Difficulties may arise in the reporting or in the recording of the information for children under 1 year of age, which may be given erroneously as "1 year of age" rather than "zero years of age". These difficulties may be mitigated by collecting information on the date of birth of all children reported as "1 year of age", while using only the direct age question for the remainder of the population. Another possible approach is to obtain age in completed months for children under 1 year of age. This method, however, can give rise to another type of recording error, that is to say, the substitution of years for months, so that a 3-month-old child, for example, might be entered in the questionnaire as being 3 years of age.

5.175. Some countries have made improvements in the quality of age data by asking both questions on age and date of birth.

5.176. An additional complication may occur with the use of the direct question if more than one method of calculating age is in use in the country. In some countries, certain segments of the population may use an old traditional method whereby persons are considered to be 1 year of age at the time of birth and everyone advances 1 year in age at the same fixed date each year. Other segments of the population in the same countries may use the Western method, in which a person is not regarded as being 1 year of age until 12 months after the date of birth, and advances 1 year in age every succeeding 12 months. If there is a risk of different methods of age calculation being used by respondents, provision must be made to ensure that the method used in each case is clearly indicated in the questionnaire and that the conversion is left to the data-editing stage.

5.177. In spite of its drawbacks, the direct question on age is the only one that should be used when people cannot provide even a birth year. As regards persons for whom information on age is unavailable or appears to be unreliable, an estimated age may have to be recorded. This may occur in isolated cases in societies where knowledge of age is widespread or in general in cultures where there is little awareness of individual age and no interest in it. In the latter circumstances, criteria for making estimates should be provided in the instructions for the enumerators.

5.178. One of the techniques that have been used to aid enumerators consists in providing them with calendars of historic events of national or local significance to be used either in probing questions or in identifying the earliest event the respondent recalls. Another technique consists in pre-identifying locally recognized age cohorts in the population and then asking about membership in the cohorts. Enumerators may also ask if the person in question was born before or after other persons whose ages have been roughly determined. Furthermore, use can be made of age norms for weaning, talking, marriage and so forth. Whatever techniques are used, enumerators should be impressed with the importance of securing age data that are as accurate as possible within the amount of time that they can devote to the topic.

5.179. In view of the possible difficulties in the collection of age data, census tests should be used, as appropriate, to determine the difference in results with the use of a question on age as compared with a question on date of birth, what calendar or method of age reckoning most people use, and in what parts of the country age will have to be estimated for the majority of the population and what techniques to use as an aid in estimation. Testing of the calendar or method of age reckoning that most people use is particularly important where an official change from one calendar or method of reckoning to another calendar or method has taken place recently enough so that the new calendar or method of reckoning may not yet be in popular use among some or all of the population.

5.180. Enumerators who are likely to be called upon to estimate age in a substantial number of cases should be given training in the applicable techniques as part of their general training.

5.181. As noted in paragraph 5.165 it is recommended that where this information is incomplete it should be derived or imputed for census purposes rather than being reported as “not stated”.

3. Marital status (core topic)

5.182. Despite the changing nature of marriage, marital status remains a useful demographic variable. The direct relationship between marriage and fertility is still recognized, as is the indirect relationship with other demographic, social and economic characteristics. Numerous variations exist in many countries, but it is important that marriage be defined in terms of the laws and customs of individual countries.

5.183. *Marital status* is the personal status of each individual in relation to the marriage laws or customs of the country. The categories of marital status to be identified should at least include the following:

(a) Never married;
(b) Married;
(c) Married, but separated;
(d) Widowed and not remarried;
(e) Divorced and not remarried.

5.184. In some countries, category (b) may require a subcategory of persons who are contractually married but not yet living as spouses. In all countries, category (c) should comprise both the legally and the de facto separated, who may be shown as separate subcategories if desired.

5.185. In some countries, it will be necessary to take into account customary unions, such as registered partnerships and consensual unions, which are legal and binding under law. In countries with legal provision for registered or legal partnership (for opposite-sex couples or same-sex couples), or where same-sex couples can legally marry, subcategories may either be included in the category (b) Married or in a legally registered partnership, namely (b)(i) "Opposite-sex marriage/partnership", (b)(ii) "Same-sex marriage/partnership".

5.186. The treatment of persons whose only or latest marriage has been annulled is dependent upon the relative size of this group in the country. Where its size is substantial, the group could constitute an additional category; if its size is insignificant, however, the individuals in the group should be classified according to their marital status before the (annulled) marriage took place.

5.187. At times countries have experienced difficulties in distinguishing between (a) formal marriages and de facto unions and (b) persons legally separated and those legally divorced. If either of these circumstances necessitates a departure from the recommended classification of marital status, the composition of each category shown in the tabulations should be clearly stated.

5.188. If complete information on marital status is needed, then this information should be collected and tabulated for persons of all ages, irrespective of the national minimum legal age, or the customary age for marriage, because the population may include persons who were married in another country with a different minimum marriage age. In most countries, there are also likely to be persons who were permitted to marry below the legal minimum age because of special circumstances. In order to permit international comparisons of data on marital status, however, any tabulations of marital status not cross-classified by detailed age should at least distinguish between persons under 15 years of age and those 15 years of age and over. If circumstances necessitate a departure from the recommended classification of marital status, the composition of each category shown in the tabulations should be clearly stated.

5.189. The collection of additional information related to customs in particular countries (such as concubinage, polygamous or polyandrous marital status, or inheritance of widows) may be useful in meeting national needs. For example, at times countries may wish to collect data on the number of spouses of each married person. Modifications of the tabulations to take account of such information should be made within the framework of the basic classification in order to maintain international comparability as far as possible.

5.190. The concept of marital status and the marital status categories described above should not be confused with the concept of de facto union status, which describes extralegal unions (including some consensual unions) of varying degrees of stability common in some countries. Information on these relationships is very useful in studies of fertility, but it is not possible to provide an international recommendation on this matter because of the different circumstances prevailing among countries. It is suggested, however, that countries wishing to investigate these relationships should consider the possibility of collecting separate data for each person on de facto unions and on the duration of each type of union (see paragraphs 5.269–5.270). Information on these relationships may also be derived

from information collected on the relationship to head or reference person or other persons in the household, in order to distinguish between people who are living in either a consensual union or marriage, and those who are not.

4. Ethnocultural characteristics

5.191. Countries with a culturally diverse population may wish to collect information on the ethnic identity (or composition) of the population, on mother tongue, on the knowledge and practice of languages and on religious communities and denominations. They are all characteristics that allow people the flexibility to express their ethnocultural identity in the way that they choose. Data on such ethnocultural characteristics of the population are of increasing relevance to countries in the context of migration, integration and minority policies.

5.192. Ethnocultural characteristics, such as ethnic identity, mother tongue, and knowledge and practice of languages, generally have a subjective dimension. Moreover, different countries will adopt different concepts. Ethnocultural characteristics can also be politically sensitive and may apply to very small, yet identifiable, population subgroups. The free and open declaration of the respondents is therefore of essential importance. Members of certain minority groups may be particularly vulnerable to discrimination on the grounds of ethnic group or religion. Special care, therefore, may be required in census procedures and outputs relating to ethnic group and religion in order to demonstrate to respondents that appropriate data protection and disclosure control measures are in place. In some cases, countries may even wish to collect such data on a voluntary basis if this is permitted by national legislation.

5. Religion

5.193. Each country that investigates religion in its census should use the definition most appropriate to its needs and should display the definition that has been used as part of the metadata in the census publications and dissemination programme.

5.194. For census purposes, *religion* may be defined as either:

(a) Religious or spiritual belief of preference, regardless of whether or not this belief is represented by an organized group; or

(b) Affiliation with an organized group having specific religious or spiritual tenets.

5.195. The decision whether to collect and disseminate information on religion in a national census is dependent upon a number of considerations and national circumstances, including, for example, the national needs for such data, and the suitability and sensitivity of asking a religion question in a country's census. Owing to the sensitive nature of a question on religion, special care may be required to demonstrate to respondents that appropriate data protection and disclosure control measures are in place. It is important that the responding public be informed of the potential uses and needs for this information.

5.196. The amount of detail collected on this topic is dependent upon the requirements of the country. It may, for example, be sufficient to enquire only about the religion of each person; on the other hand, respondents may be asked to specify, if relevant, the particular denominations to which they adhere within a religion. In countries where a large number of sects or denominations exist there will be implications for space on any census questionnaire and implications for data capture, especially in cases

where “write-in” responses are required. In an effort to ensure international comparability as far as possible, it is recommended that religion or religious affiliation should be measured directly by a question that asks “What is your religion?” rather than use of a filter question that asks for example “Are you religious?” and if so “What is your religion?” Response categories should include “No religion/religious affiliation” together with a “Religious but prefer no to disclose” or “Not stated” category, in effect making responses to such a question voluntary.

5.197. For the benefit of users of the data who may not be familiar with all of the religions or denominations within the country, as well as for purposes of international comparability, the classifications of the data should show each denomination as a subcategory of the religion of which it forms a part. A brief statement of the tenets of religions or denominations that are not likely to be known beyond the country or region would also be helpful.

6. Language

5.198. There are four types of language data that can be collected in a census, namely:

(a) Mother tongue, defined as the language usually spoken in the individual’s home in his or her early childhood;

(b) Main language, defined as the language that the person commands best;

(c) Usual language, defined as the language currently spoken, or most often spoken, by the individual in his or her present home;

(d) Ability to speak one or more designated languages, including the country’s official language(s) and/or indigenous languages.

5.199. Each of these types of information serves a very different analytical purpose. Each country should decide which, if any, of these types of information is applicable to its own needs. International comparability of tabulations is not a major factor in determining the form of the data to be collected on this topic.

5.200. In compiling data on the usual language or on the mother tongue, it is desirable to show each language that is numerically important in the country and not merely the dominant language.

5.201. Information on language (including any sign language) should be collected for all persons. In the tabulated results, the criterion for determining language for children not yet able to speak should be clearly indicated.

7. Ethnicity

5.202. The decision whether to collect and disseminate information on ethnic or national groups of a population in a census is dependent upon a number of considerations and national circumstances, including, for example, the national needs for such data, and the suitability and sensitivity of asking ethnicity questions in a country’s census. Owing to the sensitive nature of questions on ethnicity, special care may be required to demonstrate to respondents that appropriate data protection and disclosure control measures are in place. It is important that the responding public be informed of the potential uses and need for data pertaining to ethnicity, as this improves public support for the census exercise. Data on ethnicity provide information on the diversity of a population and can serve to identify subgroups of a population. Some areas of study that rely on such data include demographic trends, employment practices and opportunities, income distribution, educational levels, migration patterns

and trends, family composition and structure, social support networks, and health conditions of a population.

5.203. Broadly defined, *ethnicity* is based on a shared understanding of history and territorial origins (regional and national) of an ethnic group or community, as well as on particular cultural characteristics such as language or religion. Respondents' understanding or views about ethnicity, awareness of their family background, the number of generations they have spent in a country, and the length of time since immigration are all possible factors affecting the reporting of ethnicity in a census. Ethnicity is multidimensional and is more a process than a static concept, and so ethnic classification should be treated with movable boundaries. This may mean that classifications of ethnic categories will change between censuses which, while mirroring society at any one time, may lead to a degree of non-comparability between one census and another.

5.204. Ethnicity can be measured using a variety of concepts, including ethnic ancestry or origin, ethnic identity, cultural origins, nationality, race, colour, minority status, tribe, language, religion or various combinations of these concepts. Because of the interpretative difficulties that may occur with measuring ethnicity in a census, it is important that, where such an investigation is undertaken, the basic criteria used to measure the concept are clearly explained to respondents and in the dissemination of the resulting data. The method and the format of the question used to measure ethnicity can influence the choices that respondents make regarding their ethnic backgrounds and current ethnic identification. The subjective nature of the term (not to mention increasing intermarriage among various groups in some countries, for example) requires that information on ethnicity be acquired through self-declaration of a respondent and also that respondents have the option of indicating multiple ethnic affiliations. Data on ethnicity should not be derived from information on country of citizenship or country of birth. The classification of ethnic groups also requires the inclusion of the finest levels of ethnic groups, self-perceived groups, regional and local groups, as well as groups that are not usually considered to be ethnic groups, such as religious groups and those based on nationality. Countries collecting data on ethnicity should note that the precoding or preclassification of ethnic groups at the time of data capture may have a tendency to lose detailed information on the diversity of a population unless space to record write-in, free-form responses is provided.

5.205. Respondents should be free to indicate more than one ethnic affiliation or a combination of ethnic affiliations if they wish to do so. Countries should explain in the census instructions and the census documentation how the ethnicity of children from mixed couples is to be reported (for example, explicit instructions to allow respondents to provide multiple responses or to allow for responses such as "Biracial"). Also, to guarantee the free self-declaration of ethnicity, respondents should be allowed to indicate "None" or "Not declared".

5.206. Because the ethnocultural composition of a country can vary widely from country to country and due to the diversity in the approach and the various criteria for establishing ethnicity, it is recognized that there is no single definition or classification that could be recommended that would be applicable to all countries. However, countries should document the basic criteria and classification procedures for ethnicity and inform the data users about the concepts on which they are based.

8. Indigenous peoples

5.207. Facilitating the collection of data on indigenous peoples for national and international needs can serve to improve socioeconomic and active participation of indigenous peoples in the development

process for many countries. The sensitive nature of questions pertaining to the indigenous population requires care in assuring the public that the appropriate disclosure and data protection methods are being enforced. The responding public should be informed on the potential uses and need for such data to improve public support for the census exercise.

5.208. Dissemination of census data pertaining to indigenous peoples contributes to research in areas such as the socioeconomic conditions of the indigenous population, trends, causes for inequities, and the effectiveness of existing policies and programmes. Availability of these data can also assist indigenous peoples in assessing their conditions of living and give them the information they need to participate and advocate in the development of programmes and policies affecting their communities, such as those impacting health systems, models of economic production, environmental management and social organization. In addition, the development of indicators relevant to the indigenous population and the measurement of such indicators in the data collection process can be used to monitor the human development of indigenous peoples.

5.209. Generally, *indigenous peoples* of a particular country are social groups with an identity that is distinct from the social and cultural identity of the dominant society in that country. Questions on indigenous identity should abide by the principle of self-identification. It is important that, where such an investigation is undertaken, multiple criteria are developed to accurately capture identity and socioeconomic conditions of indigenous peoples. Defining the indigenous population can be done in many ways, such as through a question on ethnic origin (that is to say, ancestry) or on indigenous identity. Identifying the indigenous community also requires recognition of the diversity in this subpopulation, including nomadic, semi-nomadic and migrating peoples, peoples in transition, displaced persons, indigenous peoples in urban areas, and particularly vulnerable subgroups. It is important to point out that there is no single term among countries to describe the indigenous population. Consequently, countries tend to use their own national concepts to identify the indigenous population.

5.210. Differing national contexts also imply that enumerating the indigenous population can be done in multiple ways, for example, by way of specific questions on the census form, with specialized questionnaires for the indigenous population, or with follow-up or complementary surveys. In Canada, for example, identification of the indigenous population comes not only from its national census, but also from a post-censal survey. In Australia, in addition to the national census, there is the National Aboriginal and Torres Strait Islander survey, while in Argentina there is a complementary survey after the census targeting indigenous peoples. In addition to a general census, Paraguay also administers a specific census in the same year to identify the indigenous population.

5.211. Involvement of the indigenous peoples in the data development and data collection processes provides the arena for capacity-building and helps to ensure the relevance and accuracy of the data. Using local indigenous languages, employing local indigenous people (as interpreters, for example), and training and building the capacity of local indigenous people in data collection processes can facilitate the collection and dissemination of this information. Non-indigenous professionals and technicians should also be informed of the culture and practices of indigenous peoples.

9. Disability characteristics

5.212. A census can provide valuable information on disability in a country. For countries that do not have regular special population-based disability surveys or disability modules in ongoing surveys, the census may be the only source of information on the prevalence and distribution of disability and

functioning in the population at national, regional and local levels. Countries that have a registration system providing regular data on persons with the most severe types of impairments may use the census to complement these data with information related to selected aspects of the broader concept of disability and functioning based on the International Classification of Functioning, Disability and Health (ICF).[194] Census data can be utilized for general planning programmes and services (prevention and rehabilitation), monitoring selected aspects of disability trends in the country, evaluation of national programmes and services concerning the equalization of opportunities, and international comparison of selected aspects of disability prevalence in countries. While considering measures to determine disability in a census, preferences may need to be given to any legal provisions defining and categorizing disability.

a. Disability status (core topic)

5.213. *Disability status* characterizes the population into those with and those without a disability. Persons with disabilities are defined as those persons who are at greater risk than the general population for experiencing restrictions in performing specific tasks or participating in activities. This group would include persons who experience limitations in basic activity functioning, such as walking or hearing, even if such limitations were ameliorated by the use of assistive devices, a supportive environment or plentiful resources. Such persons may not experience limitations in specifically measured tasks, such as bathing or dressing, or participation activities, such as working or going to church or shopping, because the necessary adaptations have been made at the personal or environmental levels. These persons would still, however, be considered to be at greater risk of restrictions in activities or participation than the general population because of the presence of limitations in basic activity functioning, and because the absence of necessary accommodations would jeopardize their current levels of participation.

5.214. A comprehensive measure to determine disability would include the following six domains of functioning in a way that can be reasonably measured using a census and that would be appropriate for international comparison:

(a) Walking
(b) Seeing
(c) Hearing
(d) Cognition
(e) Self-care
(f) Communication.

5.215. The first four domains (a) to (d) are to be considered essential in determining disability status. The additional domains (e) and (f) comprise a more comprehensive measure for determining disability.

b. Use of the census to measure disability at the aggregate level

5.216. A census format offers only limited space and time for questions on any one topic such as disability. Since ICF offers several dimensions for use to develop a census measure, it is best to focus on a few of those dimensions, leaving the remaining dimensions for use in more extensive household

[194] Adopted in 2001, the ICF is the international standard for describing and measuring health and disability at both the individual and population levels. More information on the ICF framework is available at: http://www.who.int/classifications/icf/en/.

surveys. Short sets of disability questions, which can be included in censuses and extended sets to be recommended for inclusion in population-based surveys, have been developed and tested by the Washington Group on Disability Statistics.[195] The aim of the recommended sets is to improve comparability of disability and functioning data across countries.

5.217. The definition of disability status (see paragraph 5.213) requires that disability be defined in terms of limitations in basic activity functioning that would place a person at greater risk than the general public of restricted performance of or participation in organized activities (such as educational attendance or work participation). Given the complexity of disability definition and measurement and, in certain cultures, the sensitivity attached to people identifying as having a disability, it is recommended that several functional activity domains be defined whereby people can respond to questions about their difficulty in performing those activities rather than enquiring directly whether or not they have a particular disability.

(i) **Essential domains**

5.218. It is suggested that only those domains that have satisfied a set of selection criteria be eligible for inclusion in a short set of questions recommended for use in censuses. Criteria for inclusion include cross-population or cross-cultural comparability, suitability for self-reporting and space on the census form. Other suggested criteria include the importance of the domain in terms of public health problems. Based on these criteria, the Washington Group on Disability Statistics has developed a Short Set on Functioning (WG SS-F) questions in the six domains for the purpose of measuring disability in a census format. The four basic domains are considered to be essential: walking, seeing, hearing and cognition. In addition, if space permits, two other domains have been identified for possible inclusion: self-care and communication.

5.219. *Walking* fulfils the criteria of cross-cultural applicability and space requirements for comparable data since walking is a good indicator of a central physical function and is a major cause of limitation in participation.

5.220. While *seeing* also represents a public health problem, self-reporting of seeing limitation is more problematic, particularly when individuals use glasses to correct visual impairments. Similar difficulties are associated with asking about *hearing* activity.

5.221. Assistive devices, such as glasses and hearing aids, provide almost complete accommodation for a large proportion of those with impaired functioning. It is often argued that asking about seeing without the use of glasses greatly increases the number of persons with disabilities and makes the group too heterogeneous, that is, the group would include persons at very little risk of participation problems along with those at greater risk. As a result, questions on difficulty seeing or hearing should be asked with the use of glasses or hearing aids if they are usually worn.

5.222. Of the four essential domains, *cognition* is the most difficult to operationalize. Cognition includes many functions such as remembering, concentrating, decision-making, understanding spoken

[195] The Washington Group on Disability Statistics, a United Nations City Group that focuses on proposing international measures of disability, has developed these questions. For updates on the question wording and more information supporting the collection and use of data on disability, see www.cdc.gov/nchs/washington_group.htm.

and written language, finding one's way or following a map, doing mathematical calculations, reading and thinking. Deciding on a cross-culturally similar function that would represent even one aspect of cognition is difficult. However, remembering and concentrating or making decisions would probably serve the cultural compatibility aspects the best.

(ii) **Additional domains**

5.223. In additional to the four essential domains, two other have been identified for possible inclusion: self-care and communication. The self-care domain is intended to identify persons who have some problems with taking care of themselves independently. Washing and dressing represent self-care tasks that occur on a daily basis and are considered to be basic activities.

5.224. The purpose of the communication domain is to identify persons who have some problems with talking, listening or understanding speech such that it contributes to difficulty in doing their daily activities. Two aspects of communication are considered: understanding others (receptive communication) and being understood by others (expressive communication). Communicating (understanding and being understood) refers to the exchange of information or ideas between two people through the use of language.

5.225. Beyond the six domains identified above, there are further physical functioning domains that could be included in a set of census questions depending on the space available. The Washington Group Extended Set on Functioning (WG ES-F) includes questions that address functioning in domains such as upper body (functioning of the arms, hands and fingers), affect (anxiety and depression), pain and fatigue.

5.226. A set of questions on child functioning and disability has been developed through a collaboration between the WG and UNICEF. The developed children's module has been undergoing testing in a number of countries[196]. The limitation with the current children's module is that it comprises more than ten questions and thus is not practical for inclusion in censuses. There is a need to review and reduce the questions to be considered for inclusion in censuses.

c. Census question wording

5.227. It is recommended that special attention be paid in designing census questions to measure disability. The wording and the construct of questions greatly affect the precision in identifying persons with disabilities. Each domain should be asked through a separate question.[197] The language used should be clear, unambiguous and simple. Negative terms should always be avoided. The disability questions should be addressed to each single household member and general questions on the presence of persons with disabilities in the household should be avoided. If necessary, a proxy respondent can be used to report for the family member who is incapacitated. The important thing is to account for each family member individually rather than ask a blanket question. Scaled response categories can also

[196] See https://www.un.org/disabilities/documents/events/2014_summary_strengthening_statistics_shildren_with_disabilities.pdf.

[197] When domains are combined, such as asking a question about seeing or hearing, respondents frequently are confused and think they need to have difficulty in both domains in order to answer yes. In addition, having the numbers with specific limitations is useful for both internal planning and for cross-national comparisons.

improve the reporting of disability. The census questions on disability endorsed by the Washington Group include four response categories:

(a) No (meaning no difficulty at all);
(b) Yes – some difficulty;
(c) Yes – a lot of difficulty;
(d) Cannot do [the activity] at all.

Disability prevalence is determined based on any response that is "a lot of difficulty" or "cannot do at all" for any of the questions.

5.228. The information that results from measuring disability status (see paragraph 5.213) is expected to:

(a) Represent a large proportion of, but not all, persons with limitation in basic activity functioning in any one country (only the use of a wider set of domains would potentially cover close to all such persons, but as stated this would not be possible in a census context);
(b) Represent the most commonly occurring basic activity limitations within any country;
(c) Capture persons with similar problems across countries.

5.229. The questions identify the population with limitations in basic activities that have the potential to limit independent participation in society. The intended use of these data would be to compare levels of participation in employment, education, or family life for those with disability as measured by the question set versus those without disability to see if persons with disabilities have achieved social inclusion. In addition, the data could be used to monitor prevalence trends for persons with limitations in the particular basic activity domains selected.

5.230. Because disability is a complex concept, it is necessary to adopt an explicit definition based on the ICF domains used when developing census or survey questions that will be used to identify disability status. The recommended set of questions for censuses is based on such an explicit definition (as described above). It is essential that estimates or tabulations based on the recommended set be accompanied by information on how disability is defined and how the questions are asked. This information should be included as part of the metadata associated with the questions and data set and it should be included as a footnote to tables that include these estimates.

d. Use of census to screen for disability and follow-up with other surveys

5.231. Countries that are planning specialized surveys on disability may want to use the census to develop a sampling frame for these surveys and include a screening instrument to identify persons who will be interviewed subsequently. The main purpose of a screening instrument is to be as inclusive as possible in order to identify the largest group of people who could be further studied. The screening question should be designed so that false negatives[198] are minimized, while false positives[199] should be less of a concern.

[198] Persons who have disabilities but are not identified in the census as having disabilities.
[199] Persons who are identified with disabilities in the census but in reality do not have disabilities (as assessed in the largest instrument used in the follow-up survey).

5.232. The same recommendations highlighted in paragraphs 5.227–5.230 should also be considered when a screening module is designed.

5.233. Before embarking on using the census to develop a frame for a follow-up survey, it is important that the legal implications of using the census data for this purpose are fully considered. Respondents should be informed that the data may be used for follow-up studies and national authorities responsible for ensuring the privacy rights of the population may need to be consulted in order to obtain their approval.

E. Fertility and mortality

5.234. The investigation of fertility and mortality in population censuses is particularly important in countries lacking a timely and reliable system of civil registration and vital statistics because of the opportunity the data provide for estimating vital rates that would not otherwise be available. Even in countries with complete birth and death registration, some of the topics (such as “children born alive”, “children living”, “age at marriage or union” and “age at first birth”) are equally appropriate because they provide data that are not easily available from registration data but are necessary for the computation of cohort and period fertility tables. The census provides an opportunity to collect data for estimating fertility and mortality at national and subnational levels in a cost-effective manner. The inclusion of these topics in population censuses for the purpose of estimating fertility and mortality rates and other related indicators is both prudent and cost-effective, particularly in countries where civil registration and vital statistics systems are weak, and costs of conducting large periodic demographic surveys are high. Nevertheless, it is important to note that census information is a poor substitute for complete and reliable vital registration data. If countries desire accurate and detailed estimates of fertility and mortality, they must establish, and need to maintain, civil registration systems and ensure their universal coverage.

5.235. To obtain data on fertility, information may be collected on “children ever born”, “date of last child born alive” and “age of mother at birth of first child born alive”. In addition, questions on age, date or duration of marriage or union may improve fertility estimates based on children ever born (see paragraph 5.269). For the collection of reliable data, some of the topics may require a series of probing questions that, because they are time consuming, are more suitable for use in sample surveys than in censuses.

5.236. The universe for which data should be collected for each of the topics included in this section consists of women 15 years of age and over regardless of marital status. It may be appropriate in some countries to reduce the lower age limit by a few years. It is acknowledged that in some countries certain cultural sensitivities (for example, regarding the collection of information on childbearing from never-married women) exist towards collecting information from all women aged 15 years of age and over without regard to marital status. In such cases, every effort should be made to collect the requisite information. In countries that do not use the data for women 50 years of age and over, it may be appropriate to limit data collection to women under the age of 50, allowing more concentrated effort on data collection for such women.

5.237. In addition to the characteristics indicated above that are used to estimate fertility, another useful variable that allows the estimation of fertility is the "own children" method[200] and birth history reconstruction.[201] The application of these methods requires the identification of the natural mother of each child in the household when the natural mother appears in the same questionnaire as the child. In cases where it is difficult to ascertain the identity of the natural mother, one may use as a proxy the relationship to the reference person of the household (see paragraphs 5.142–5.153) or children living (see paragraphs 5.256–5.258) to establish the identity of the natural mother. In essence, information on the child's age and the mother's age are used to estimate a series of annual fertility rates for years prior to the census. The reliability of the estimates produced depends, among other things, on the proportion of mothers enumerated in the same questionnaire as their own children, the accuracy of age reporting for both mothers and their children and the accuracy of available estimates of mortality for women and children. In the case of subnational estimates, the fertility rates may be affected by migration patterns of mothers who leave their children with other family members while they pursue work elsewhere and also good and timely birth registration of children born to migrant women. A more parsimonious estimation method of fertility, the reverse survival method, allows estimating fertility from data collected in a census even if no questions have been asked about fertility directly.[202] Based on population data by single age and sex, the method consists in 'reverse surviving' those no longer present in the population of a given age in order to derive the number of births that occurred in the past, usually for a period of 15 years preceding the data collection. The consistency of the fertility estimates obtained from the reverse survival method depends on the errors in the age distribution (e.g., omission of infants and young children and heaping on specific ages).

5.238. Mortality topics include infant and child mortality, obtained from data on children ever born and children living, and adult mortality, obtained from household deaths in the past 12 months and maternal or paternal orphanhood. The extent to which mortality (particularly infant and child) can be adequately measured from population census data is largely dependent on the quality of the training of field staff to minimize non-response to questions on maternal and adult mortality and the reporting of erroneous information. Enumerator manuals should include the measures that are needed to minimize such errors. Accurate responses to the questions described here are often difficult to obtain, thus resulting in faulty data. Nevertheless, it is often possible to derive usable adjusted estimates from this information.

5.239. As far as possible, efforts should be made to obtain information on fertility, child mortality (or survival) and marriage or union directly from the woman or mother involved, because she is more likely

---

[200] For methodological details, see United Nations (2004). *Handbook on the Collection of Fertility and Mortality Data*. Studies in Methods. Series F, No.92, Chapter V, section B (http://unstats.un.org/unsd/publication/SeriesF/SeriesF_92E.pdf); United Nations (1983). *Manual X: Indirect Techniques for Demographic Estimation*, Population Studies, No. 81 (http://www.un.org/esa/population/publications/Manual_X/Manual_X.htm), Chapter VIII, section C; and Cho, L.-J., Retherford, R. D., & Choe, M. K. (1986). *The own-children method of fertility estimation*. Honolulu, HI: Population Institute.

[201] For methodological details, see United Nations (2004). *Handbook on the Collection of Fertility and Mortality Data*. Studies in Methods. Series F, No.92, Chapter V, section C (http://unstats.un.org/unsd/publication/SeriesF/SeriesF_92E.pdf); and Luther, N.Y., Cho, L-J. (1988). Reconstruction of birth histories from census and household survey data. *Population Studies*, 42: 451–472.

[202] For methodological details, see Moultrie, T.A., Dorrington, R.E. , Hill, A.G., Hill, K.H. , Timæus, I.M. , and Zaba, B. (eds.) (2012). *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population; Spoorenberg, T. (2014), "Reverse survival method of fertility estimation: An evaluation", *Demographic Research* 31(9): 217-246.

to recall correctly the details of her fertility, the mortality of her offspring and her marital experiences than any other member of the household. Information on household deaths, by date, sex and age, in the 12-month period prior to the census should be collected from the head of the household (or household reference person). Information on maternal orphanhood and paternal orphanhood should be collected for each person in the household regardless of age. As with fertility, mortality questions may be limited to a survey sample.

5.240. A number of countries have restricted the collection of data from fertility and mortality questions in the census to a sample of enumeration areas,[203] entailing the introduction of more vigorous training and permitting the selection of more suitable field staff. When those items are included in the census, certain precautions to ensure accuracy and completeness should be observed. As mentioned above, every effort should be made to collect all relevant information directly from the woman concerned. To reduce underreporting of events and to improve the accuracy of responses to questions on fertility and mortality, enumerators need to receive specific training on probing questions that highlight common errors and omissions.

5.241. The limitations of the data collected and of the estimates based on them should be made clear in the census reports. Furthermore, since some of the estimation procedures are only suitable for use in certain circumstances, it is important that census data producers consult specialists or carefully evaluate the methodologies for estimating the indicators for their appropriateness in a given situation. In general, the data in the basic tabulations resulting from these questions should not be used for the direct calculation of fertility and mortality rates. Reliable estimation of fertility and mortality levels using census data requires adjustment based on methods of demographic analysis.[204]

5.242. As a general guide, only one of the items related to fertility discussed below ("children ever born") is recommended for inclusion in all situations. Even in countries with reliable civil registration of births, census information on this topic can be useful for assessing the completeness of the registration system and for estimating levels of lifetime fertility for older cohorts.

5.243. In countries where civil registration of births and deaths is incomplete or unreliable, it is recommended that a subset of the remaining items should be included as well. Among these, one item ("date of birth of last child born alive") is useful for the indirect estimation of current fertility levels. Two additional items ("children living" and "household deaths in the past 12 months") are especially important, as they allow for the indirect estimation of mortality levels.

5.244. The three remaining items have lower priority: "age, date or duration of first marriage/union"; "age of mother at birth of first child born alive"; and "maternal or paternal orphanhood". However, in situations where a country has included one of these items in consecutive previous censuses, it may be

[203] For the use of sampling in the enumeration, see Part One, Chapter VI.

[204] *Manual X: Indirect Techniques for Demographic Estimation*, Population Studies, No. 81 (United Nations publication, Sales No. E.83.XIII.2); National Academy of Sciences, Committee on Population and Demography, *Collecting Data for the Estimation of Fertility and Mortality*, Report No.6 (Washington D.C., National Academy Press, 1981), p. 220; *Handbook of Population and Housing Censuses, Part II*, Studies in Methods, No. 54 (United Nations publication, Sales No. E.91.XVII.9), Chapters III and IV; *Step-by-Step Guide to the Estimation of Child Mortality*, Population Studies, No. 107 (United Nations publication, Sales No. E.89.XIII.9); Moultrie T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus and B. Zaba (eds). *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/.

useful to collect comparable information to measure changes over time and because cohort analysis, particularly of the prevalence of orphanhood, can be useful in assessing levels of mortality.

5.245. For countries with low fertility and mortality, and where the civil registration is not reliable, further consideration should be given to the item "age of mother at birth of first child born alive", since it improves the timeliness of estimates of child mortality based upon children ever born and children surviving.

5.246. In countries with high adolescent birth rates and common child or early marriages (with large proportion of women married or in union before age 18), it is recommended that the questions on "age or date of first marriage/union" and "age of mother at birth of first child born alive" are included, since it improves understanding of the spatial distribution of extent of early childbearing and marriage that surveys cannot provide.

5.247. It is worth emphasizing that all estimates of fertility and mortality derived from census data are approximate and subject to various sorts of error. Therefore, in the absence of complete and reliable civil registration data, it may be desirable to have more than one type of census information on each topic (for example, both household deaths in the past 12 months and maternal or paternal orphanhood for the purpose of estimating adult mortality). Lastly, it should also be born in mind that while fertility surveys can provide data on current fertility, they cannot provide the small-area data that the census can. Therefore, a fertility question in the census can still be a priority for many countries.

1. Children ever born alive (core topic)

5.248. Information on number of *children born alive* (lifetime fertility) should include all children born alive (that is to say, excluding foetal deaths) during the lifetime of the woman concerned up to the census date. The number recorded should include all live-born children, whether born in or out of marriage, whether born in the present or a prior marriage, or in a de facto union, or whether living or dead at the time of the census. In the event of multiple births (for example twins), each child should be counted as individual birth.

5.249. Data on the total number of live-born children should preferably be collected for all women 15 years of age and over, regardless of marital status. If, from a cultural standpoint, it is not acceptable in some countries to attempt to obtain the information for single women, it should be collected for all women 15 years of age and over who are or have been married or in a union (in other words, all ever-married or ever cohabiting women), a group that also includes all widowed, divorced and separated women. Depending on national needs and circumstances, data on their live-born children can be collected for girls under 15 years old as well. In either case, the group of women for whom the data have been collected should be clearly described in the census report so as to avoid ambiguity in the analysis of the results.

5.250. In some countries, there can be substantial misreporting of ages or dates in the census, which distorts fertility and mortality estimation based on children ever born and children living cross-tabulated by age or years since the woman's first birth.

5.251. In order to improve the completeness of coverage and to assist the respondent in recalling her children ever born alive, it is recommended that a sequence of questions be included in the following order:

(a) "total number of sons ever born alive during the lifetime of the woman";
(b) "total number of sons living (surviving) at the time of the census"; and
(c) "total number of sons born alive who have died before the census date";
and then
(d) "total number of daughters ever born alive during the lifetime of the woman";
(e) "total number of daughters living (surviving) at the time of the census"; and
(f) "total number of daughters born alive who have died before the census date".

5.252. The responses to topics (b), (c), (e) and (f) allow for a checking of the consistency of the responses to (a) and (d). Inconsistencies in the figures, if any, can sometimes be resolved during the interview. In obtaining this information and ensuring its accuracy, more than one filter question can be asked to serve as a reminder for the respondent. The responses can be confirmed by using additional tables/lists containing the names and birth years of the children she has given birth to.

5.253. The number of sons and daughters should comprise all children ever born alive whether born of the present or a prior marriage or union[205] and should exclude foetal deaths and adopted children. Also, the number of children, male and female, who are alive at the time of the census should include those living with the mother in the household and those living elsewhere, no matter where the latter may reside and regardless of their age and marital status.

5.254. The collection of data on children ever born specified by sex not only improves the accuracy of information but also provides data for the indirect estimation of sex ratio at birth as well as the calculation of infant and child mortality, in combination with data on children living (surviving) by sex (see paragraph 5.256). If the information on "children ever born alive by sex" is collected for only a sample of women, the data on "children living by sex" should also be obtained for the same sample.

5.255. Collecting data on the "total number of children ever born alive by sex" is desirable because it can enhance the quality of demographic analysis. This can be achieved by ensuring that sex ratios of births follow an expected pattern and do not exhibit unusual behaviour. While this point does add value by addressing the importance of sex-specific data collection, it could benefit from providing more context on how anomalies in sex ratios might impact demographic analysis and planning.

2. Children living [206] (core topic)

5.256. Data on *children living*, in conjunction with those on children ever born, are used in indirect estimation of infant and child mortality in situations where there are no reliable data from a civil registration system.

5.257. It is expected that improved coverage and quality of data on the total number of children ever born will be achieved if additional questions about the current residence of children ever born are asked, in terms of the following:

(a) “Total number of sons living in the household”;

[205] As indicated in paragraph 5.160, couples living in consensual unions may, where appropriate, be regarded as married.
[206] For methodological details on the uses of the data, together with data on live-born children, see the publications mentioned in footnote 120.

(b) “Total number of sons living elsewhere”;
(c) “Total number of sons born alive who have died before the census date”;
(d) “Total number of daughters living in the household”;
(e) “Total number of daughters living elsewhere”;
(f) “Total number of daughters born alive who have died before the census date”.

These questions not only give a more complete and accurate reporting of children ever born alive specified by sex but also increase the suitability of the information collected for subsequent analysis.

5.258. The identification of the natural mother of each child under 15 years of age in the same household, to be used in the “own children” method of estimating fertility (see paragraph 5.237), should be made by asking each woman who reports one or more of her children as being born alive and living in the household to identify these children in the census questionnaire. The section of the questionnaire on “relationship to the head of the household or to the reference person in the household” may be used for identifying the natural mother of each child living in the household.

3. Date of birth and sex of last child born alive (core topic)

5.259. Information on date of birth (day, month and year) and sex of the last child born alive is used for estimating current fertility. Data on the sex of the child can also be used to evaluate the sex ratio at birth and to detect potential sex-selective birth omissions, misreporting or coding errors. This information can be useful as a means of deriving both national and subnational fertility estimates. In countries lacking adequate data from civil registration, sample surveys have become a major source of information for estimating national fertility levels, but surveys usually do not permit the derivation of reliable estimates at subnational levels.

5.260. An estimate of the number of live births during the 12 months immediately preceding the census date can be derived at the data-processing stage from the information collected. For estimating current age-specific fertility rates and other fertility measures, the data provided by this approach are more accurate than information that may have been collected in earlier censuses from a question on the number of births to a woman during the 12 months immediately preceding the census.[207] Information on the date of birth of the last child born alive provides the number of women who had at least one live-born child during the 12-month period, not the number of births during the 12-month period. However, generally only a very small proportion of women will have had more than one child in a year and hence that omission will not significantly affect the fertility estimate derived from it.

5.261. The information needs to be collected only for women between 15 and 49 years of age who have reported having at least one live birth during their lifetime. For countries where childbearing starts early the minimum age can be lowered according to the national circumstances and needs. Also, the information should be collected for all the marital or union status categories of women for whom data on children ever born by sex (see paragraph 5.248) are collected. If the data on children ever born are collected only for a sample of women, information on date of birth for the last child born alive should be collected for the same sample.

[207] The approach to calculating fertility rates from these data is described in Moultrie T.A. 2013. Evaluation of data on recent fertility from censuses, in Moultrie T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus and B. Zaba (eds). *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/content/evaluation-data-recent-fertility-censuses (accessed 06/11/2013).

5.262. A census question on birth history should always be paired with a simple follow-up question about whether or not the child is still alive, from which information on deaths of children born in the last 12 months can be rederived (see paragraph 5.266 below), and which yields data that can be used for studying child mortality (see paragraph 5.272).

4. Births in the past 12 months

5.263. Births in the past 12 months refers to the total number of children born alive to the woman concerned during the 12 months immediately preceding the census. The topic is more suitable for investigation in a sample survey than in a census. The information about births in the past 12 months is useful for calculating fertility indicators either directly or indirectly.

5.264. Because of errors and omissions commonly encountered in the reporting of live births within a 12-month retrospective period, this topic by itself cannot generally be relied on to generate accurate estimates of current fertility. It is recommended that an estimate of the number of live births during the 12 months immediately preceding the census date be derived from information on "date of birth of last child born alive" (see paragraph 5.260). For estimating current age-specific fertility rates and other fertility measures, the data provided by this approach are more accurate than information that may have been collected in earlier censuses from a question on the number of births to a woman during the 12 months immediately preceding the census.[208] Only if a country's population is characterized by low levels of date numeracy – that is, there is a strong indication of deficient accuracy of the replies – should a direct question on number of births in the last 12 months be asked. In all other cases, the core question on date of last child born alive should be sufficient.

5.265. If the topic needs to be introduced into the census, information should be collected for all the marital status categories of women for whom data on children born alive are collected. In countries in which current births to mothers at ages below 15 years are statistically significant, the age limit for the investigation of current fertility may be lowered to include mothers at appropriate younger ages.

5. Deaths among children born in the past 12 months

5.266. Deaths among children born in the past 12 months refers to the number of deaths that occurred among the live births within the 12 months preceding the census reported for the woman concerned. Hence, the topic should be investigated only if live births within this period are also investigated.

5.267. It is recommended that information on the topic be derived from a pair of census questions on "date of birth of last child born alive" and the follow-up question about "whether or not the child is still alive" (see paragraph 5.262). Although this pair of questions does not produce a valid estimate of the infant mortality rate (since the numerator excludes infant deaths occurring below age 1 in the past 12 months among children born 1–2 years before the census date), they can provide useful information on differences in child survival by age of mother or other socioeconomic characteristics.

[208] The approach to calculating fertility rates from these data is described in Moultrie T.A. 2013. Evaluation of data on recent fertility from censuses, in Moultrie T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus and B. Zaba (eds). *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/content/evaluation-data-recent-fertility-censuses (accessed 06/11/2013).

5.268. The information needs to be collected only for women for whom at least one live birth during the reference period was reported.

6. Age, date or duration of first marriage or union

5.269. *Date of first marriage* comprises the day, month and year when the first marriage took place. In situations where date of first marriage is difficult to obtain, it is advisable to collect information on age at marriage or on how many years ago the marriage took place (duration of marriage[209]). The information should relate to all types of marriages such as contractual first marriages and de facto unions, customary marriages and religious marriages.

5.270. For women who are widowed, separated or divorced at the time of the census, information on the "date of/age at/number of years since dissolution of first marriage" should be collected. Information on dissolution of first marriage (if pertinent) provides data necessary to calculate "duration of first marriage" as a derived topic at the data-processing stage. In countries in which duration of marriage is reported more reliably than age, tabulations of children ever born by duration of marriage yield better fertility estimates than those based on data on children born alive classified by age of the woman.[210] Data on duration of marriage can be obtained by subtracting the age at marriage from the current age, or directly from the number of years elapsed since the marriage took place.

7. Age of mother at birth of (date or time since) first child born alive[211]

5.271. Date of first birth comprises the day, month and year when the woman's first live birth took place. In situations where date of first birth is difficult to obtain, it is advisable to collect information on age of mother at first birth or on how many years ago the first birth took place (time since first birth). In countries in which time since first birth is reported more reliably than age, tabulations of children ever born and children surviving by time since first birth yield more timely child mortality estimates than those based on data on children born alive classified by age of the woman.[212] If the topic is included in the census, information should be obtained for each woman who has had at least one child born alive.

[209] While this provides some insights into marital history, it doesn't correctly capture the duration of the marriage if it has ended prior to the census date.

[210] See United Nations (1983). *Manual X: Indirect Techniques for Demographic Estimation*, Population Studies, No. 81 (http://www.un.org/esa/population/publications/Manual_X/Manual_X.htm), Chapter II, section D; and United Nations (2004). *Handbook on the Collection of Fertility and Mortality Data*, Studies in Methods Series F, No. 92 (http://unstats.un.org/unsd/publication/SeriesF/SeriesF_92E.pdf), Chapter V, section D.

[211] Ibid., Chapter II, section B.3.

[212] See Hill K. 2013. Indirect estimation of child mortality, in Moultrie T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus and B. Zaba (eds). *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population (http://demographicestimation.iussp.org/content/indirect-estimation-child-mortality); and Rajaratnam, J.K., L.N. Tran, A.D. Lopez, & C.J.L. Murray (2010). Measuring under-five mortality: validation of new low-cost methods. *PLoS Med*, 7(4), e1000253. doi: 10.1371/journal.pmed.1000253http://demographicestimation.iussp.org/content/indirect-estimation-child-mortality (accessed 13/12/2013).

8. Household deaths in the past 12 months[213] (core topic)

5.272. Information on household deaths in a specified period (usually 12 months[214]) prior to the census date classified by sex of deceased and age at death is used to estimate the level and pattern of mortality in countries that lack satisfactory continuous death statistics from civil registration. In order for estimation derived from this item to be reliable, it is important that all deaths to household members occurring during the specified period preceding enumeration be reported as completely and as accurately as possible. Typically, reports of deaths in censuses underestimate the overall number of deaths if only because some deaths result in the disintegration of households so that household survivors, if any, may not report their occurrence (in particular, deaths of persons living alone at the time of death are unlikely to be reported). Nevertheless, provided that there are no serious errors in the reporting of age at death, estimates of completeness of death reporting can be derived via indirect estimation and adequate mortality estimates obtained.[215,216]

5.273. Ideally, information on mortality should be collected for each household in terms of the *total number of deaths in the specified period* prior to the census date. For each deceased person reported, name, age, sex and date (day, month and year) of death should also be collected. Care should be taken to clearly specify the reference period to the respondent so as to avoid errors due to its misinterpretation. For example, a precise reference period could be defined in terms of a festive or historic date for each country. It is also recommended for a country to use specific reference period e.g. between DD/MM/20YY and DD/MM/20YY.

5.274. When information is collected on household deaths in the previous specified reference period, countries may wish to ask a pair of follow-up questions concerning cause of death. After ascertaining the name, age and sex of the deceased person and date of death, two additional optional questions could be asked:

(a) Was the death due to an accident, violence, homicide or suicide?

[213] See United Nations (2004). *Handbook on the Collection of Fertility and Mortality Data*. Studies in Methods, Series F, No.92 (http://unstats.un.org/unsd/publication/SeriesF/SeriesF_92E.pdf); United Nations (2002). *Methods for Estimating Adult Mortality.* ESA/P/WP.175 (http://www.un.org/esa/population/techcoop/DemEst/methods_adultmort/methods_adultmort.html); and Dorrington, Rob E. (2013). The Brass Growth Balance Method and the Preston-Coale Method for one census, and the Generalized Growth Balance Method and Synthetic Extinct Generations Methods upon the availability of deaths from two censuses, in T.A. Moultrie, R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus & B. Zaba (eds.), *Tools for Demographic Estimation.* International Union for the Scientific Study of Population. http://demographicestimation.iussp.org.

[214] The DESA-WHO group, established during the COVID-19 pandemic, on mortality estimates has recommended extending the reference period for household death inquiries from one year to a flexible duration, potentially up to two years or longer in the event of crises like pandemics or natural disasters. This change should be implemented following rigorous testing and evaluation of its impact on data quality.

[216] See Chapter 4 on methods for data evaluation and adjustment in WHO (2013). *WHO guidance for measuring maternal mortality from a census*. World Health Organization, Geneva (http://www.who.int/reproductivehealth/publications/monitoring/9789241506113/en/index.html); and Moultrie, T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus & B. Zaba (eds.) (2013). *Tools for Demographic Estimation.* International Union for the Scientific Study of Population (http://demographicestimation.iussp.org).

(b) If the deceased was a woman aged 15[217] to 49 at the time of death, did the death occur while she was:

b1. Pregnant? Yes/No

b2. During childbirth? Yes/No

b3. During the six weeks after the end of pregnancy? Yes/No

See: WHO, 2013 and US Census Bureau, 2017.

5.275. Data derived from such questions can help to assess trends and levels, and some causes, of adult mortality and maternal mortality. According to the Tenth International Classification of Diseases, (ICD-10, Class XV, codes O00–O99) there are several concepts regarding maternal death, such as:

- Maternal death: a maternal death is the death of a woman while pregnant or within 42 days of the termination of pregnancy, irrespective of the duration and the site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental causes.
- Pregnancy-related death: a pregnancy-related death is the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of cause.
- Late pregnancy-related death: a late pregnancy-related death is the death of a woman while pregnant or within 12 months of termination of pregnancy, irrespective of cause.

5.276. There is no universal agreement about the feasibility of collecting reliable cause of death information as part of a population and housing census. More research is needed on both the feasibility and methods of collecting cause of death information as part of a national census.

9. Maternal or paternal orphanhood[218]

5.277. Some countries may also wish to collect information on maternal or paternal orphanhood in another attempt to ascertain the level and patterns of mortality in the population. Census data from these two topics are intended for indirect estimation of mortality by sex. Estimates are based on the proportion of persons classified by age whose natural mothers or fathers are still alive at the time of the census.

5.278. For the collection of information on orphanhood, two direct questions should be asked, namely:

(a) Whether or not the natural mother of the person enumerated in the household is still alive at the time of the census, regardless of whether or not the mother and father are enumerated in the same household.

(b) Whether or not the natural father of the person enumerated in the household is still alive at the time of the census, regardless of whether or not the mother and father are enumerated in the same household.

[217] It may be appropriate in some countries to reduce the lower age limit by several years.

[218] For methodological details on the uses of the data, see Timæus, I.M. (2013). Indirect estimation of adult mortality from orphanhood, in T.A. Moultrie, R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus & B. Zaba (eds.), *Tools for Demographic Estimation.* International Union for the Scientific Study of Population (http://demographicestimation.iussp.org/content/orphanhood); and United Nations (2002). *Methods for Estimating Adult Mortality*. ESA/P/WP.175 (http://www.un.org/esa/population/techcoop/DemEst/methods_adultmort/methods_adultmort.html).

The investigation should secure information on biological parents. Thus, care should be taken to exclude adopting and fostering parents. It should be kept in mind, however, that overcounting may occur in the case of parents with more than one surviving child among the respondents, particularly in high fertility societies.

5.279. It is preferable for these questions to be collected from every person in the household regardless of age (not just children under 18, which would otherwise make the information useless for estimating adult mortality). Not only is this important for estimating mortality at older ages, but also for estimating the extent of age exaggeration at the older ages. Whenever the context allows, the date of death should be collected to help to improve knowledge of the timing of death, and in other contexts a simple follow-up question about whether the parent was still alive five years ago can help to narrow down the timing of death and to improve adult mortality measurement for recent years by analysing these data as successive cross-sectional enquiries.[219]

F. Registration of vital events and assignment of legal identity

5.280. Including questions on the registration of vital events and the assignment of legal identity in population censuses provides valuable data for assessing civil registration and legal identity management systems. These questions offer insights into the completeness, and quality of these systems across sub-national regions, population sub-groups, and socio-economic segments. This data is crucial for aligning with the principles of universal legal identity and "Leaving No One Behind" as outlined in the 2030 Sustainable Development Agenda. By identifying disparities and inequalities, targeted interventions can be implemented to strengthen civil registration and vital statistics systems and ensure everyone has access to legal identification.

5.281. Census data on these topics can provide valuable insights into societal vulnerabilities. For example, by cross-referencing birth registration data with labor force information, child labor patterns can be identified. Similarly, analyzing marriage registration data can help uncover instances of early marriage. Linking legal identity information with demographic and socioeconomic data can highlight disparities in access to essential services like education, healthcare, and financial systems.

5.282. The decision whether to include questions on registration of vital events and assignment of legal identity requires careful consideration. Factors to weigh include the development stage of the civil registration system, census costs and feasibility, the availability of recent data on the civil registration system's completeness and quality, and sensitivity of the topic. Ultimately, the specific reasons for including these questions will vary depending on each country's unique context and needs.

5.283. *Civil registration* is the permanent, compulsory, and universal recording of the occurrence and characteristics of vital events pertaining to the population in accordance with the law [220]. There are five types of registration of vital events that can be collected in a census, namely:

[219] See Timæus, I.M. (2013). Indirect estimation from orphanhood in multiple inquiries, in T.A. Moultrie, R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus & B. Zaba (eds.), *Tools for Demographic Estimation.* International Union for the Scientific Study of Population (http://demographicestimation.iussp.org/content/orphanhood-multiple-inquiries).

[220] UN Principles and Recommendations for a Vital Statistics, Revision 3 (2014).

i) *registration of birth,* the permanent and official record of a person's existence which provides legal recognition of the person's identity;
ii) *registration of death*, the official recording of the occurrence and characteristics of a death;
iii) *registration of marriage*, the process by which the legal relationship of spouses is constituted;
iv) *registration of divorce,* the legal dissolution of a marriage or registered partnership, according to national laws, conferring on the parties the right to reenter into another partnership or marriage; and,
v) *assignment of legal identity,* the issuance or possession of legal identity documents.

5.284. The registration of vital events is usually a multi-stage process. In most countries it involves three basic stages[221]: notification, registration, and certification[222].

5.285. Some census respondents and enumerators may not be clear on the differences between notification, registration, or certification of a vital event. And some may even conflate official civil registration with (non-official) recognition of a vital event by other important entities, such as religious institutions or community leaders. Hence, the design and wording of census questions should be carefully reviewed by multiple stakeholders (across the government entities involved in civil registration and identity management processes) as well as community institutions involved in ritual and documentation that are often carried out around the time of a birth, marriage or death.

5.286. Testing of the questions prior to finalization is also recommended. To ensure good quality data, it is important for these concepts to be clarified to field workers during training on data collection.

5.287. Ideally countries should strive to assess levels of registration as evidence of permanent recording of the occurrence of vital events.

5.288. Additional questions on date of occurrence of events can be included. Date of notification and date of registration may be asked in countries where notification does not immediately result in registration. This is largely where the country's legislation provides longer time frames, or no legal period for registration (UNICEF, 2017).

5.289. Alternatively, countries may opt to skip registration questions and ask about notification along with possession of a birth/death/marriage certificate. In these instances, possession of a certificate is regarded as proof for registration. In such cases, questions can be added on both date of notification and date of issue of a certificate.

5.290. One of the key considerations is the reference period to which questions relate. Countries can align registration questions to the census reference period for both births and deaths. For censuses with a reference of twelve months for household deaths, questions can be in relation to registrations within

[221] Ibid.

[222] *Notification* is the process of providing all the information on and all the characteristics of an event that is to be legally registered by the registration authority. The informant reports or notifies the vital event to the civil registrar. *Civil Registration* is the permanent, compulsory, and universal recording of the occurrence and characteristics of vital events pertaining to the population in accordance with the law. A certificate is an official document containing all or part of the exact information contained on the original vital record, issued by the registration office, and has the full force and effect of the original vital record. *Certification* is thus a product (output) of registration. It is when the civil registrar issues an official copy containing some or all the vital event information contained in the register.

the last twelve months. This reference period is ideal for countries with a legal framework that stipulates registrations within one to twelve months after occurrence. In the case of birth registration questions, alignment can be with questions on births in the last twelve months.

5.291. The reference population for questions on vital events' registration and the assignment of legal identity documents is crucial. This applies primarily to birth registrations, marriage registrations, and, to a lesser extent, identity card possession. The age range for birth registration questions depends on the purpose:

(i) Assessing recent interventions: Focus on children born within the last twelve months to evaluate the effectiveness of timely birth registration initiatives.
(ii) Tracking progress over time: To assess long-term birth registration trends, consider data spanning several recent years.
(iii) Evaluating policy changes: If a major policy shift or increased investment in Civil Registration and Vital Statistics (CRVS) systems happened five years ago, analyze data for 10-15 years. This allows comparison of birth registration completeness before and after the change.

5.292. If the goal is to assess birth registration at the national level due to efforts to improve registration for all, questions can apply to all individuals in a household. Where legislation mandates birth registration before issuing legal IDs (linked CRVS-ID systems), then careful consideration should be given to the age range for questions. The reference population depends on the legal age for obtaining a national ID card in each country. Countries with separate CRVS and ID systems might assess the linkage extent by asking everyone in a household about birth registration and ID possession.

5.293. The concept of marriage can vary significantly across cultures and societies. To ensure accurate data collection, the census questionnaire design and fieldworker training should clarify whether the question on marriage registration refers to the most recent marriage, or the latest marriage that is still ongoing. Only "married" persons can then answer to marriage registration question.

5.294. This clarification will help fieldworkers accurately record data and avoid confusion for respondents with diverse marital histories. Further, there may be divergence, within and among communities, between the legal concept of marriage, as per the Family Code and/or civil registration laws, and how nuptiality is practiced (sometimes entailing a wide array of marriage types - such as religious marriage, civil marriage, traditional marriages, and polygamous marriages). Assessment of marriage registration completeness via census questions needs to be well-defined as to the specific phenomenon being assessed, the reference population and time period to be considered.

5.295. Regarding the assignment of legal identity, it is important to consider that eligibility for inclusion in identity management systems varies by country. Some countries establish eligibility for a national identity card/credential at birth, whereas others establish it for persons above 16, or 18 years. Countries also have different regulations/processes around naturalization of foreign citizens. These fundamental differences are important considerations, when designing census questions on legal identity.

5.296. Questions on the assignment of legal identity need to consider privacy concerns. Some countries have asked census respondents to disclose the national identity number of individuals or to show proof of a civil registration certificate or national ID card as part of the census enumeration process. The collection of information on personal ID number can help countries to link census data with administrative records, paving the way for a move towards register-based censuses.

5.297. It is recommended that, as with all parts of a census enumeration, the collection of data on registration of vital events and the assignment of legal identity is carried out with careful consideration of privacy and confidentiality of individuals, national data protection standards and the UN Fundamental Principles of Official Statistics.

5.298. Civil registration and legal identity systems are multi-stakeholder in nature. Such systems often are coordinated across multiple ministries and agencies, and this inherent nature of the systems can affect access to and inclusion in public benefit systems such as health, education, and social protection. Therefore, it is recommended, when designing and implementing census questions on civil registration and legal identity systems, that key stakeholders (e.g., line ministries, civil society, religious and local community leaders, etc.) be consulted. Consultation ensures alignment between the chosen measurement approach and the legal framework and operational practices of the civil registration and legal identity systems.

G. Educational characteristics

1. Literacy (core topic)

5.299. *Literacy* has historically been defined as the ability both to read and to write, distinguishing between "literate" and "illiterate" people. A literate person is one who can both read and write, with understanding, a short, simple statement on his or her everyday life. An illiterate person is one who cannot, with understanding, both read and write such a statement. Hence, a person capable of reading and writing only figures and his or her own name should be considered illiterate, as should a person who can read but not write as well as one who can read and write only a ritual phrase that has been memorized. However, a more modern understanding referring to literacy as a continuum of skills, levels, domains of application and functionality is now widely accepted.

5.300. The notion of literacy applies to any language insofar as it exists in written form. In multilingual countries, the census questionnaire may also enquire into the languages in which a person can read and write. Such information can be essential for the determination of education policy. This item would, therefore, be a useful additional subject of enquiry.

5.301. It is recommended that data on literacy be collected for all persons 10 years of age and over. In a number of countries, however, some children may only become literate through school between the ages of 10 and 14 years. The literacy rate for this age group may be misleading. Therefore, in an international comparison of literacy, data on literacy should be tabulated for all persons 15 years of age and over. Where countries collect the data for younger persons, the tabulations on literacy should at least distinguish between persons under 15 years of age and those 15 years of age and over.

5.302. Straightforward operational criteria and instructions for collecting literacy statistics should be clearly established on the basis of the concept given in paragraph 5.299, and applied during census taking.[223] Accordingly, although data on literacy should be collected so as to distinguish between persons who are literate and those who are illiterate, consideration should be given to distinguishing broad levels of literacy skills. Simple questions with response categories that reflect different levels of literacy skills should be used. In addition, since literacy is an applied skill, it needs to be measured in

[223] Depending on the need for small-area data and the circumstances in a country, literacy may best be measured through surveys.

relation to a particular task, such as reading, with understanding, personal letters and newspapers or magazines, or writing a personal letter or message. Respondents may be able to do so easily, with difficulty or not at all, reflecting the different levels of literacy skills. Reading and writing may be measured separately to simplify the questions.

5.303. It is preferable to use standardized questions, harmonized across countries to ensure comparability. The United Nations Educational, Scientific and Cultural Organization (UNESCO) has developed a reference database of model questions. In addition, UNESCO recommends that literacy tests should be administered, in order to verify, as well as improve, the quality of literacy data. Nevertheless, administering a literacy test to all household members in the course of enumeration may prove impractical and affect participation, therefore limiting the utility of the results. Instead, administering such a test to a sample of respondents may be considered either in the census itself or in a targeted post-enumeration follow-up survey. Some countries have regularly used simple self-assessment questions within a census to provide an indication of literacy rates at the small-area level. An evaluation of the quality of statistics should be provided with census statistics on literacy.

5.304. The collection and tabulation of statistics on literacy during the population census should not be based on any assumed linkages between literacy, school attendance and educational attainment. In operational terms, this means systematically enquiring about the literacy status of each household member irrespective of school attendance or highest grade or level completed.

5.305. The literacy question currently varies across countries and, as a result, the data based on it are not always internationally comparable. Literacy should not be derived as an educational attainment proxy because although the two are related, there are substantial differences. For example, there are numerous cases where people leave school with only partial literacy skills, or lose them because of a lack of practice. Therefore, educational attainment is not a good proxy measure of literacy skills. For countries adopting a register-based census, deriving levels of literacy from administrative data sources is filled with difficulties.

2. School attendance (core topic)

5.306. *School attendance* is defined as regular attendance at any regular accredited educational institution or programme, public or private, for organized learning at any level of education at the time of the census or, if the census is taken during the vacation period, at the end of the school year or during the previous school year. According to the International Standard Classification of Education (ISCED), education is taken to comprise all institutionalized, intentional and planned activities designed to meet learning needs. Instruction in particular skills that is not part of the recognized educational structure of the country (for example, in-service training courses in factories) is not normally considered “school attendance” for census purposes.

5.307. Information on school attendance should, in principle, be collected for persons of all ages. It relates in particular to the population of official school age, which ranges in general from 5 to 29 years of age but can vary from country to country depending on the national education structure. In the case where data collection is extended to cover attendance in pre-primary education or other systematic educational and training programmes organized for adults in productive and service enterprises (such as the in-service training courses mentioned in paragraph 5.306), community-based organizations and other non-educational institutions, the age range may be adjusted as appropriate.

5.308. Data on school attendance should be cross-classified with data on educational attainment, according to the person's current level and grade (see paragraph 5.313). This cross-classification can provide useful information on the correspondence between age and level or grade of educational attainment for persons attending school.

5.309. The issue surrounding the number of out-of-school children has grown in importance within recent decades, particularly within the context of the UNESCO Education for All goal with regard to achieving universal primary education. The target year for Education for All was 2015 and new goals for the post-2015 period were not yet defined at the time this document was prepared. The census offers an opportunity to measure the number of "out-of-school" (reciprocal of attendance) or "ever-in-school" children.

5.310. There is a difference between "attending school" and "enrolled in school", thus results from censuses and administrative data may differ. A child can be enrolled in school but not necessarily be attending. It is recommended that these concepts be clearly defined so that countries can determine which variable they wish to collect via the census.

5.311. It is also recommended that Member States consider the need for internationally harmonized questions in order to measure school attendance and school enrolment. Field-based censuses will usually collect data on 'attendance' while administrative data will only refer to 'enrolment'.

5.312. For purposes of international comparison, data on school attendance should be presented by the ISCED-P (or ISCED-Programmes) levels listed below, which are used for the classification of education programmes in ISCED 2011. Correspondence between a national education system and ISCED can be established through mapping of national education programmes to the ISCED classification.[224]

ISCED level 0: Early childhood education
ISCED level 1: Primary education
ISCED level 2: Lower secondary education
ISCED level 3: Upper secondary education
ISCED level 4: Post-secondary non-tertiary education
ISCED level 5: Short-cycle tertiary education
ISCED level 6: Bachelor's or equivalent level
ISCED level 7: Master's or equivalent level
ISCED level 8: Doctoral or equivalent level

3. Educational attainment (core topic)

5.313. The recommendations on "educational attainment" (see paragraph 5.314) and "educational qualifications" (see paragraph 5.328) make use of categories of the 2011 revision of ISCED, issued by UNESCO.[225] In accordance with national conditions and requirements, many countries can continue to apply national classifications of levels and grades of education and of fields of education in collecting and tabulating statistics from population censuses. Special attention needs to be paid to establishing

[224] The UNESCO Institute for Statistics maintains a database with ISCED mappings at http://www.uis.unesco.org/ISCED.
[225] UNESCO Institute for Statistics (UIS). 2012. *International Standard Classification of Education: ISCED 2011*. Montreal: UIS. http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf.

appropriate level or grade equivalence for persons who have received education under a different or foreign educational system. These national classifications, however, should be able to be converted or mapped to the ISCED 2011 classification system, this typically being achieved during post-census processing.

5.314. *Educational attainment* is defined as the highest ISCED level successfully completed by an individual. Educational attainment is usually measured with respect to the highest education programme successfully completed, which is typically certified by a recognized qualification. Some countries may also find it useful to present data on educational attainment in terms of the highest grade completed. For international purposes a "grade" is a specific stage of instruction usually covered in the course of an academic year, though this may vary from country to country. Information on educational attainment should preferably be collected for all persons 5 years of age and over.

5.315. To produce statistics on educational attainment, a classification is needed that indicates the qualifications certifying the successful completion of levels of education. Since the educational structure may have changed over time, it is necessary to make provisions for persons educated at a time when the national educational system differed from that in place at the time of the census. In addition to focusing attention on the collection of educational attainment data, enumerator instructions, coding and data processing need to be designed in a way that will take account of any changes in the educational system of a country over the years and of those educated in another country, as well as those educated in the current system.

5.316. Information collected on the highest level of education successfully completed by each individual, typically certified by a recognized qualification, facilitates flexible regrouping of the data according to various kinds of aggregation. Recognized intermediate qualifications are classified at a lower level than the programme itself. Information on intermediate qualifications or on the highest grade completed can be used to distinguish between persons who did and persons who did not complete each level of education.

5.317. For international comparison, data from the population census are needed for all levels of education defined in ISCED. To the extent possible, countries should classify statistics on educational attainment by the individual ISCED-A (or ISCED-Attainment) levels listed below, which are used for the classification of educational attainment in ISCED 2011 (or by their equivalent as set forth according to the national classification of levels of education):

ISCED level 0: Less than primary education
ISCED level 1: Primary education
ISCED level 2: Lower secondary education
ISCED level 3: Upper secondary education
ISCED level 4: Post-secondary non-tertiary education
ISCED level 5: Short-cycle tertiary education
ISCED level 6: Bachelor's or equivalent level
ISCED level 7: Master's or equivalent level
ISCED level 8: Doctoral or equivalent level

5.318. For the classification of educational attainment, ISCED level 0 has a different meaning in ISCED 2011 than for the classification of education programmes (see paragraph 5.312): it means not having successfully completed ISCED level 1. This includes individuals who have never attended an education

programme, who have attended some early childhood education (defined as ISCED level 0 in the classification of education programmes), or who have attended some primary education but have not successfully completed ISCED level 1. Any differences between national and international definitions and classifications of education should be explained in the census publications in order to facilitate comparison and analysis.

5.319. Countries could consider asking a question that captures levels of education not successfully completed, should this be of interest to policymakers or other users. This could be in the form of a direct question asking if a person has some education at the relevant level or via a question asking the last grade or year completed from any given level of education.

5.320. Data on school attendance, educational attainment and literacy status should be collected and tabulated separately and independently of each other, without (as elaborated in paragraph 5.304) any assumption of linkages between them.

5.321. In order to ensure continued and improved international comparability of census data by level of education, it is recommended that countries continue to ensure that the educational attainment variable can be mapped into the ISCED 2011 classification. This is typically achieved in post-census processing.

4. Field of education and educational qualifications

a. Field of education and training

5.322. Information on persons by field of education and training is important for examining the match between the supply and demand for qualified workers with specific specializations within the labour market. It is equally important for planning and regulating the production capacities of different levels, types and branches of educational institutions and training programmes.

5.323. A question on field of education and training needs to be addressed to persons 15 years of age and over who attended at least one grade in secondary education or who attended other organized educational and training programmes at equivalent levels.

5.324. The ISCED Fields of Education and Training 2013 (ISCED-F 2013) distinguishes between broad fields (two-digit codes), narrow fields (three-digit codes) and detailed fields (four-digit codes) of education and training.[226] The broad fields are listed here:

00 Generic programmes and qualifications
01 Education
02 Arts and humanities
03 Social sciences, journalism and information
04 Business, administration and law

[226] UNESCO Institute for Statistics (UIS). 2014. *ISCED Fields of Education and Training 2013 (ISCED-F 2013).* Montreal: UIS. http://www.uis.unesco.org/Education/Documents/isced-fields-of-education-training-2013.pdf.

05 Natural sciences, mathematics and statistics
06 Information and communication technologies (ICTs)
07 Engineering, manufacturing and construction
08 Agriculture, forestry, fisheries and veterinary
09 Health and welfare
10 Services
99 Unknown

5.325. Countries may wish to consider collecting data on more narrow and detailed fields of education and training, not only on the broad fields. For this, countries should make use of the classification and coding of fields of education and training of ISCED.

5.326. Countries coding field of education and training according to a national classification should establish correspondence with ISCED, either through double coding or through conversion from the detailed national classification to ISCED. A problem may arise in identifying the exact fields of education and training of persons with interdisciplinary or multidisciplinary fields of specialization. In these cases it is recommended that countries follow the procedure of identifying the major or principal field of education and training of those with multidisciplinary specialization.

5.327. In order to ensure continued and improved international comparability of census data by field of education and training, it is recommended that the classification structure for the fields of education and training continue to be based on the most recent version of ISCED.

b. Educational qualifications

5.328. *Educational qualification* is the official confirmation, usually in the form of a document, certifying the successful completion of an education programme or a stage of a programme. Qualifications can be obtained through (a) successful completion of a full education programme; (b) successful completion of a stage of an education programme (intermediate qualifications); or (c) validation of acquired knowledge, skills and competencies, independent of participation in an education programme.

5.329. According to national needs, information on qualifications may be collected from persons who have reached a certain minimum age or level of educational attainment. Such information should refer to the title of the highest certificate, diploma or degree received.

H. Individual use of information and communication technology

5.330. In today's digital world, access to and use of Information and Communication Technologies (ICT) by individuals has become a defining feature of modern life. Measuring this phenomenon goes beyond simple device ownership, encompassing the variety of ways people utilize technology for communication, information access, and participation in the digital economy. Understanding these patterns of individual ICT use is crucial for policymakers. It provides insights into factors influencing digital inclusion, highlights potential inequalities in access and skills, and ultimately informs the development of effective policies to bridge these gaps.

5.331. This data is also essential for monitoring progress towards internationally agreed development goals, such as those outlined in the UN's 2030 Agenda for Sustainable Development (indicators on individual Internet use and mobile phone ownership are each referenced through SDG indicators 17.8.1 and 5.b.1, respectively). Mobile phone ownership, in particular, is important to track gender equality since the mobile phone is a personal device that, if owned and not just shared, provides women with a degree of independence and autonomy, including for professional purposes. Several studies have highlighted the link between mobile phone ownership and empowerment, and productivity growth. Existing data on the proportion of women owning a mobile phone suggest that fewer women than men own a mobile phone. This indicator highlights the importance of mobile phone ownership to track and improve gender equality, and monitoring will help design targeted policies to overcome the gender digital divide. By tracking individual ICT usage, policymakers can assess the effectiveness of initiatives aimed at promoting digital literacy, fostering inclusive growth, and ensuring everyone benefits from the opportunities offered by the digital revolution.

5.332. The 3rd Revision of the Principles and Recommendations recommended collection of information on communication device ownership and internet access at the household level. However, the rapid evolution of information and communication technology (ICT) and the widespread use of mobile devices by individuals means this approach may miss crucial details about a country's digital landscape. Specifically, relying solely on household data overlooks the significant mobile phone penetration and diverse usage patterns within a population. Mobile phones have transcended communication purposes, enabling internet access, streaming services, and other functionalities that may not be captured through traditional household surveys. This proliferation of individual mobile phone use necessitates a shift in data collection.

5.333. The ITU recommends[227] collecting information on access to and use of ICTs at both the household and individual levels as that would provide a more accurate picture. At the individual level, the ITU recommends collecting data on Internet use and mobile phone ownership in population and housing censuses.

1. Individuals using the Internet

5.334. Use of the Internet[228] by individuals is tracked by policymakers and analysts as a key indication of a country's progress towards becoming an information society. The recommended question formulation by the ITU[229] is "Have you used the Internet from any location in the last three months? With response categories of "Yes" or "No". This question is asked of all in-scope individuals, not only those who have used a computer. Access to the Internet can be via a fixed or mobile network.
The reference period for individuals who used the Internet should be in the last three months and include use by individuals from any location. Classificatory variables for individuals using the Internet – such as age, sex, education level or labour force status – can help identify digital divides in individuals

[227] *Manual for measuring ICT access and use by households and individuals, 2020 Edition.* https://www.itu.int/en/ITU-D/Statistics/Documents/publications/manual/ITUManualHouseholds2020_E.pdf
[228] The Internet is a worldwide public computer network. It provides access to a number of communication services including the World Wide Web and carries e-mail, news, entertainment and data files, irrespective of the device used (not assumed to be only via a computer - it may also be by mobile telephone, tablet, PDA, games machine, digital TV etc.).
[229] Manual for Measuring ICT Access and Use by Households and Individuals, 2020 Edition. Link: https://www.itu.int/en/ITU-D/Statistics/Documents/publications/manual/ITUManualHouseholds2020_E.pdf

using the Internet. This information can contribute to the design of targeted policies to overcome those divides.

2. Individuals owning mobile cellular phones

5.335. Mobile cellular telephones have become the predominant method of communication in many countries. Indicators related to mobile telephony are therefore fundamental indicators of the information society. Mobile cellular telephones[230] refer to portable telephones using cellular technology that provides access to Public Switched Telephone Network (PSTN). An individual owns a mobile cellular telephone if he/she has a mobile cellular phone device with at least one active SIM card for personal use. It includes mobile cellular phones supplied by employers that can be used for personal reasons (to make personal calls, access the Internet, etc.) and those who have a mobile phone for personal use that is not registered under his/her name. It excludes individuals who have only active SIM card(s) and not a mobile phone device. Users of both post-paid subscriptions and pre-paid accounts are included. The equipment should be in working condition at the time of the data collection. For countries with this data, further analysis can be conducted across various dimensions: region (urban/rural), sex, age group, educational level, labour force status, and occupation. This detailed breakdown empowers policymakers to assess the effectiveness of initiatives.

I. Economic characteristics

1. Introduction

5.336. Statistics on the economic characteristics of persons are needed from population censuses for many reasons. Information on the productive activities of persons is vital to establish a comprehensive picture of the economic structure of a country, and the work patterns, labour market participation and extent of labour underutilization of its population. This information, when combined with other personal, household and dwelling characteristics collected in the census, enables assessments of the socioeconomic situation of persons and households, which are essential to inform the planning of a wide range of economic and social policies and programmes related to such areas as employment creation, poverty reduction, work–life balance, vocational education and training, provision of social security and other social benefits, gender justice and social inclusion.

5.337. Such statistics can be obtained from other sources, such as household-based surveys or administrative records, but these other sources have certain limitations. Household surveys, especially labour force surveys, are particularly well suited for generating a broad range of statistics on the economic characteristics of the population at aggregate levels, such as national and broad regional groupings. Data obtained from labour force surveys, however, are subject to sampling error and, therefore, rarely provide reliable estimates for small areas, small population groups, or detailed groups of industries and occupations. In contrast, population censuses can provide certain core statistics at the lowest levels of aggregation for such small population groups and for detailed occupation and industry groups. Administrative records may not have the same quality of occupational and industry coding, nor have the same comprehensiveness in population or activity coverage, generally excluding productive activities that are informal or unpaid.

[230] This includes analogue and digital cellular systems and technologies such as IMT-2000 (3G) and IMT-Advanced.

5.338. The population census also provides benchmark information to which statistics from other sources can be related. Population censuses likewise provide the sample frames for most household-based surveys, including labour force surveys. In countries with a limited or infrequent household survey programme, the population census may represent the main or only source of information on the economic characteristics of the population.

5.339. In deciding which topics relating to the economic characteristics of the population to include in the population census, countries will need to assess the existence of other data sources and their complementary uses. The aim should be to cover the core topics needed as benchmark information for the preparation of sample frames, and to provide essential statistics for small areas and small population groups, and for detailed occupation and industry groups, as relevant in the national context.

5.340. International resolutions and guidelines to produce statistics relating to the economic characteristics of the population are adopted by the International Conference of Labour Statisticians and endorsed by the Governing Body of the International Labour Organization.[231] Recommendations on topics amenable for inclusion in population censuses are discussed in general in Box 1. The complete labour statistics standards are available at http://www.ilo.org/global/statistics-and-databases/standards-and-guidelines/lang--en/index.htm.

Box 1. International recommendations concerning statistics of work, employment and labour underutilization

In October 2013, the Nineteenth International Conference of Labour Statisticians (19th ICLS) adopted the *Resolution concerning statistics of work, employment and labour underutilization*.[232] The resolution replaced the previous international recommendations relating to the measurement of the economically active population, employment, unemployment and underemployment dating from 1982 (Thirteenth International Conference of Labour Statisticians- 13th ICLS) and related guidelines.

The standards adopted at the 19th ICLS introduced a number of important revisions, among which are a conceptual framework for work statistics consistent with the System of National Accounts; guidelines for separately measuring different forms of work, including a more targeted definition of employment as *work for pay or profit*, and for expanding the range of measures of labour underutilization beyond the traditional unemployment. New terminology was also introduced, as relevant, and terms considered to be out of date, particularly "economically active/inactive", were replaced with "in the labour force/outside the labour force."

Important elements from the 13th ICLS resolution of the economically active population, employment, unemployment and underemployment essential to the internal consistency of the statistics remain unchanged. The refinements of the definition of employment and introduction of measures of labour underutilization may result, however, in breaks in the historical series of statistics of the economically active population, employment, unemployment and underemployment. In particular, productive

[231] See http://www.ilo.org/stat.

[232] The Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, October 2013), is available at: http://www.ilo.org/global/statistics-and-databases/meetings-and-events/international-conference-of-labour-statisticians/19/WCMS_230304/lang--en/index.htm.

activities carried out without pay, such as those listed below, are no longer included within the scope of employment:

- Production of goods intended for own final use by the household;
- Unpaid work by apprentices, interns and trainees;
- Organization-based volunteer work;
- Direct volunteering to produce goods for other households.

Participation in these activities is now to be measured separately through the defined forms of work identified as: own-use production work, unpaid trainee work and volunteer work, respectively.

Countries are encouraged to develop their statistical system to cover work statistics, including statistics on the labour force, based on their specific national needs and resources. In the case of the measures affected by the Nineteenth International Conference of Labour Statisticians Resolution, the updated international standards would ideally be implemented over time, in a way that is feasible for national statistical systems. It is of utmost importance that the institutions and persons responsible for planning and managing the production of statistics on the economic characteristics of the population develop a strategic and coordinated approach that takes into account all official sources of statistics, including the population census, labour force survey and other household-based surveys and administrative records. Data users will need to be kept well informed of the process, including by widely disseminating the relevant metadata and by maintaining parallel series for a specified period following their implementation.

2. Conceptual framework for work statistics

a. Work

5.341. Measurement of the economic characteristics of the population is based on the conceptual framework for work statistics (see Box 1). In this framework, work is defined for reference purposes as "any activity performed by persons of any sex and age to produce goods or to provide services for use by others or for own use".

5.342. The concept of *work* is aligned with the general production boundary as defined in the System of National Accounts 2008 (2008 SNA), enabling full integration between work statistics and production statistics. All *work* or *productive* activities are thus included, irrespective of their formal or informal character, or whether paid or unpaid. Excluded are activities that do not involve producing goods or services (for example begging and stealing), self-care (for example personal grooming and hygiene), and activities that cannot be performed by another person on one's own behalf (for example sleeping, learning and activities for own recreation).

5.343. *Work* can be performed in any kind of economic unit, including market units (for example corporations, quasi-corporations and household unincorporated market enterprises); non-market units (for example government and non-profit institutions serving households); and households that produce goods or services for own final use by the producers.[233]

[233] United Nations, *System of National Accounts 2008* (New York, 2008).

5.344. The conceptual framework for work statistics identifies five mutually exclusive *forms of work* for separate measurement (see Figure 3):

(a) *Own-use production work*, comprising production of goods and services for own final use;

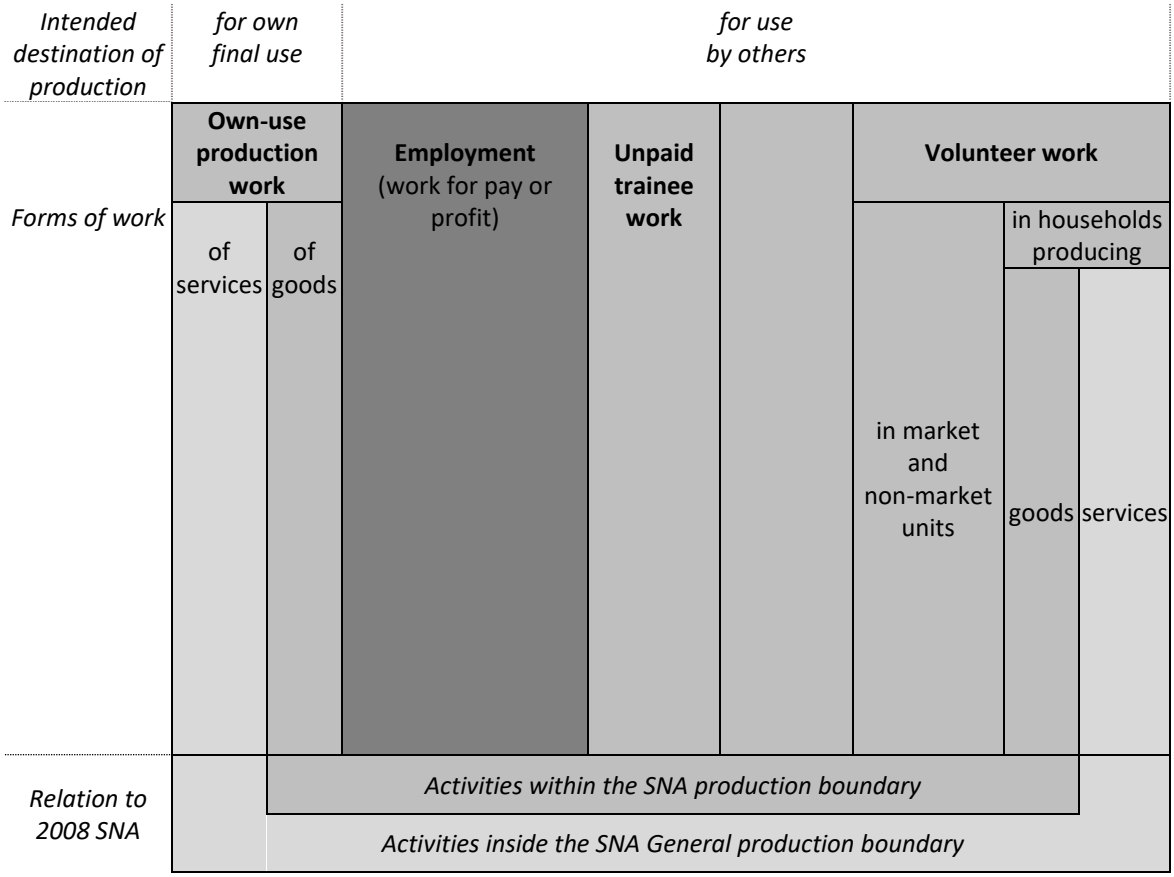
(b) *Employment work*, comprising work performed in exchange for pay or profit;

(c) *Unpaid trainee work*, comprising work performed for others without pay to acquire workplace experience or skills;

(d) *Volunteer work*, comprising non-compulsory work performed for others without pay;

(e) *Other work activities*, including unpaid compulsory work performed for others, such as community service and work by prisoners, when ordered by a court or similar authority, and unpaid military or alternative civilian service.

Figure 3. Forms of work and the System of National Accounts 2008



Note: in figure, in the column next to the unpaid trainee work column, there will be "other work activities".

5.345. During a given reference period, persons may engage in one or more forms of work in parallel or consecutively, that is, persons may be employed, volunteering, doing unpaid trainee work or producing for own final use, in any combination.

5.346. To meet different objectives, countries may measure the economic characteristics of the population with respect to their participation in one or in several forms of work. In particular, in the population census, this may include measurement of the following:

(a) *Persons in employment* is essential as part of the preparation of labour force statistics that include unemployment and other measures of labour underutilization. It is needed to assess the labour market participation of the population and to classify the population according to their labour force status in a short reference period (see paragraphs 5.354–5.379).

(b) *Persons in own-use production of goods* is especially important in countries where particular groups of the population engage in agriculture, fishing or hunting and gathering for own final consumption, including for subsistence (see paragraphs 5.433–5.438), and to enable integration of the population census with the agricultural census (see paragraphs 1.71–1.77).

(c*) Persons in unpaid trainee work* may be advisable where unpaid apprenticeships, internships and traineeships may be a main mechanism of labour market entry for particular groups such as youth or for specific occupations such as mechanics or tailors, given their likely overall small size in the country and limited availability of alternative statistical sources.

5.347. Given the need for detailed probing, measurement of participation in *own-use provision of services, unpaid trainee work* and *volunteer work* is more appropriate through household surveys or, if desired, through the population census by means of a long form applied to a subset of the population.[234]

5.348. Additional information may also be collected in the population census in order to classify the population according to their main form of work based on self-declaration, in a short or long reference period.[235]

b. Working time

5.349. The concept of working time comprises the time associated with productive or work activities and the arrangement of this time during a specified reference period.[236] Working time relates to each form of work.

5.350. The number of persons engaged in a given form of work provides only a very rough estimate of the volume of work performed, particularly when the work is performed on a part-time, casual or occasional basis. Information on working time is necessary to prepare estimates of the volume of work or labour input for complete national production accounts. It is also essential to support the design, monitoring and evaluation of economic, social and labour market policies and programmes targeting

[234] For more details see Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, 2013), paragraphs 22(c) and 37–39.
[235] Ibid, paragraph 17.
[236] Resolution concerning the measurement of working time, adopted by the Eighteenth International Conference of Labour Statisticians (Geneva, 2008).

labour market flexibility, work–life balance and conditions of work, including situations of underemployment due to insufficient working time (that is, time-related underemployment) and of excessive working time.

5.351. The population census can serve to provide information on two measures of working time in particular: *hours usually worked* and *hours actually worked*. Where the census is the only available data source it may as a minimum incorporate a single question on *hours usually worked* for persons in employment (see paragraphs 5.426–5.432) and for persons in own-use production of goods, as relevant (see paragraph 5.438).

c. Population coverage and age limits

5.352. Information on the economic characteristics of the population should in principle cover the entire population, regardless of country of origin or birth, citizenship or geographic location of their place of work. In practice, a lower age limit is usually set in accordance with the conditions in the country. Where national programmes of statistics on the working-age population or on child labour exist, the statistics derived from the population census will serve to complement those bodies of statistics. For purposes of compiling statistics on the working-age population, international standards recommend that countries set the lower age limit taking into consideration the minimum age for employment and exceptions specified in national laws or regulations, or the age of completion of compulsory schooling.[237] For compiling child labour statistics, the relevant international standards identify the target population as all persons in the 5 to 17 years age group.[238] Countries in which many children participate in employment or in other forms of work, including in agriculture, will need to select a lower minimum age than countries where work of children is uncommon. Census tabulations of economic characteristics should at least distinguish between persons under 15 years of age and those 15 years of age and over.

5.353. In general, an upper age limit is not recommended, so as to permit comprehensive coverage of work activities of the adult population and to examine transitions between employment and retirement. Many people continue to be engaged in employment and in other forms of work beyond retirement age and the numbers involved are likely to increase as a result of factors associated with the "ageing" of the population. Countries may, however, wish to balance the cost of collecting and processing information relating to the productive activities of older persons (those aged 75 years or more) and the additional response burden imposed on them against the significance and reliability of the information provided.

3. Labour force status (core topic)

5.354. Persons may be classified in a short reference period according to their *labour force status* as being employed, unemployed, or outside the labour force as defined below in paragraphs 5.359–5.379.

5.355. A classification of persons by their *labour force status* provides important information about their relation to the labour market, in particular to *work for pay or profit*, in a short reference period. The three categories of *labour force status* i.e. employed, unemployed or outside the labour force, are mutually exclusive and exhaustive. While even during a short period persons may be engaged in multiple activities,

[237] Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, 2013), paragraph 65.
[238] Resolution concerning statistics of child labour, adopted by the Eighteenth International Conference of Labour Statisticians (Geneva, 2008).

to establish their *labour force status*, priority is given to employment over other forms of work, and over unemployment; and to unemployment over outside the labour force (see Figure 4). Thus, for example:

- a volunteer worker who also has a part-time employee job should be classified as *employed*,
- a student who is also seeking and available for employment should be classified as *unemployed*, and
- a person who has a part-time job working only a few hours for pay and who is also seeking another job should be classified as *employed*.

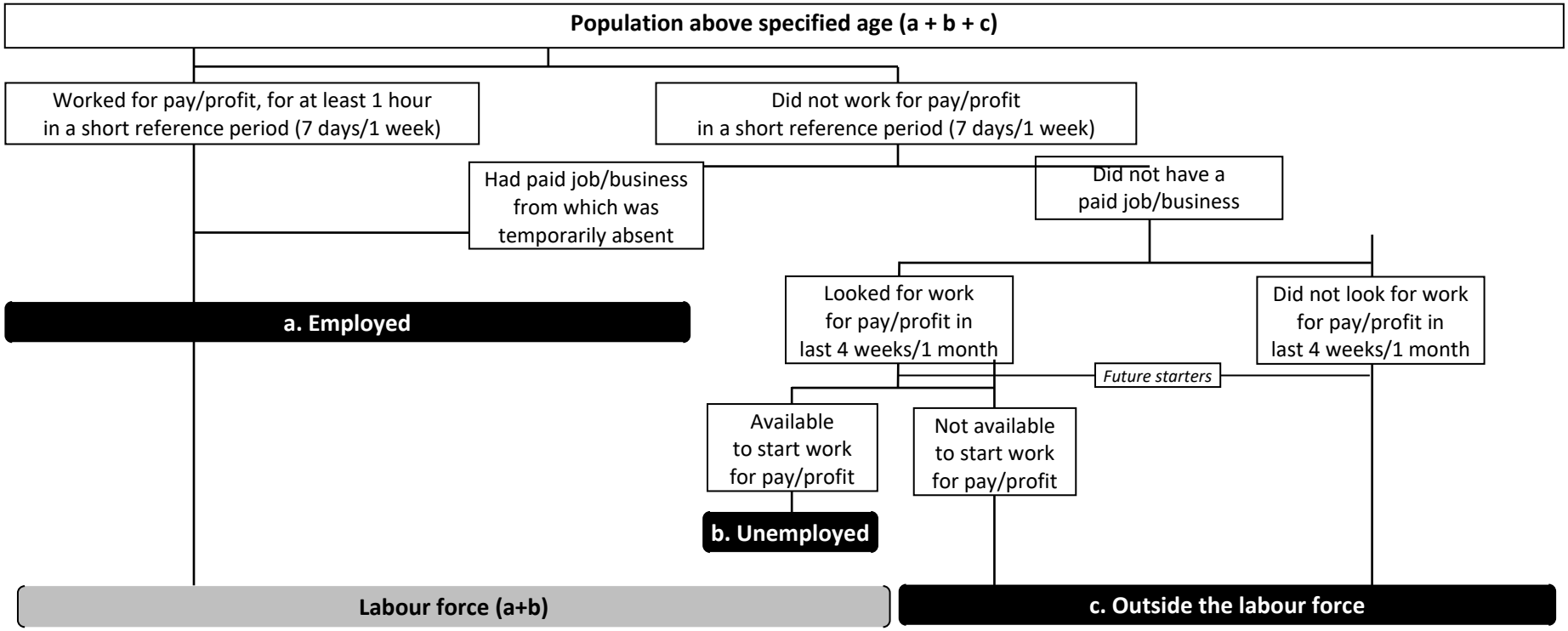
The sum of persons in employment plus persons in unemployment comprises the labour force.

5.356. The *labour force status* of persons is established with regard to a short reference period of seven days or one week, which may be the last seven days prior to the specified census day, the last completed calendar week or a specified recent fixed week. For comparability purposes, it is particularly useful to apply a short reference period with the same total duration (such as a reference week) for the census as for the national labour force survey, if any. This short reference period serves to provide a snapshot picture of labour market participation in the country around the time of the census. As such, the *labour force* (that is, persons in employment plus persons in unemployment) reflects the supply of labour for the production of goods and services in exchange for pay or profit at a specified point in time. Seasonal variations in employment and unemployment levels, which may be significant both in industrialized and in developing economies, will not be captured. Assessments of such temporal variations in work patterns are more adequately captured through sub-annual household surveys (for example monthly, quarterly).

5.357. Depending on the way the relevant parts of the census questionnaire have been constructed, the determination of the labour force status of a person may be influenced by respondents' or enumerators' subjective understanding of the concepts of employment and unemployment. In this regard, particular attention should be given to special groups for which the determination of the labour force status may be difficult. These groups include youths, women and older persons after the normal age of retirement, in particular those working as contributing family members. Their participation in employment and job search activities is frequently

Figure 4. Classification of working age population by labour force status

***Note: Figure 4 will be reformatted*.**



overlooked and needs close attention. In particular, a problem in classifying women as outside the labour force, or cultural perceptions relating to gender roles, can result in serious omissions with respect to measuring women's participation in employment and job search activities. To reduce underreporting, enumerators need to be explicitly instructed, or the questionnaires specifically designed, to ask about the possible jobs, including part-time, casual, temporary and informal jobs, or job search activities of every woman and man above the specified age in the household.

5.358. The addition of probing questions in an interview, or more detailed questions in a self-administered questionnaire, may lengthen the time required to complete the questionnaire and increase the cost of the census. Accordingly, it will be necessary to balance the gains in terms of minimizing response errors when such questions are used against the added costs associated with their inclusion. Given the importance of reliable data on labour force status, however, serious consideration should be given to minimizing classification errors. To this end, the training of enumerators should highlight likely sources of omission or gender bias leading to underestimation of participation in employment.

a. Employed persons

5.359. *Employed persons* are all those above the specified age who during a short reference period of seven days or one week were engaged in any activity to produce goods or provide services for pay or profit. The notion "for pay or profit" refers to work done as part of a transaction in exchange for remuneration payable in the form of wages or salaries for time worked or work done or in the form of profits derived through market transactions from the goods and services produced. It includes remuneration in cash or in kind, whether actually received or not, payable directly to the person performing the work or indirectly to a household or family member.

5.360. Two categories of persons in employment are (a) employed persons "at work", that is, who worked for pay or profit for *at least* one hour; and (b) employed persons "not at work" due to working-time arrangements (such as shift work, flexitime and compensatory leave for overtime) or to "temporary absence" from a job for pay or profit.

5.361. Use of the one-hour criterion serves to ensure coverage of all types of jobs engaged in, including part-time, temporary or casual jobs, thereby supporting identification of all persons in employment and analysis of their working conditions. This criterion is also essential in order to ensure that unemployed persons refer to those without *any* employment, who are seeking and available for work for pay or profit. Moreover, this criterion is a prerequisite for the consistency of employment statistics with national accounts data on production. When information on working time is also collected (see paragraphs 5.426–5.432, it is recommended that employed persons be classified by specified bands of working time so as to enable identification of persons with both short and excessive working time.

5.362. Persons on "temporary absence" from a job, including as employees or self-employed, should be considered as in employment, provided that they were "not at work" for a short duration and maintained a labour attachment during the absence. The existence of a labour attachment should be established on the basis of the reason for the absence and, in the case of certain reasons, the continued receipt of remuneration or the total duration of the absence (in general not greater than three months). Reasons for absence include:

(a) those where labour attachment is generally maintained and thus do not require further probing include sick leave due to own illness or injury (including occupational); public holidays, vacation or annual leave; and periods of maternity or paternity leave as specified by legislation; and

(b) those requiring further assessment of continued receipt of remuneration or total duration including parental leave, educational leave, care for others, other personal absences, strikes or lockouts, reduction in economic activity (for example temporary lay-off, slack work), disorganization or suspension of work (for example due to bad weather, mechanical, electrical or communication breakdown, problems with ICT, and shortage of raw materials or fuels).

Treatment of specific groups

5.363. According to international standards, the following groups of persons should be classified as in employment:

(a) Persons with a job for pay or profit who, during the reference period, were on training or skills enhancement activities required by their job or for another job in the same economic unit;

(b) Apprentices, interns or trainees who receive remuneration in cash or in kind;

(c) Persons who work for pay or profit through employment promotion programmes;

(d) Persons who work in their own economic units to produce goods intended *mainly* for sale or barter, even if part of the output is consumed by the household or family;

(e) Persons with seasonal jobs during the off season, if they continue to perform some tasks and duties of the job, excluding, however, fulfilment of legal or administrative obligations (for example pay taxes);

(f) Regular members of the armed forces and persons on military or alternative civilian service who perform this work for pay in cash or in kind.

(g) Persons who either work in a market unit operated by a family member living in the same or in another household (that is, contributing family workers) or perform tasks or duties of an employee job held by a family member living in the same or in another household should also be classified as in employment.

5.364. These groups of workers are included in employment, regardless of the number of hours actually worked, as they contribute their labour to produce goods and services for pay or profit, payable to the household or family.

5.365. In accordance with the priority rule to establish their *labour force status* (see paragraph 5.355), persons who during the reference period were primarily students, homemakers, pensioners, registered unemployed, or engaged in other forms of work, and at the same time were engaged in employment, as defined above, should be classified as in employment.

5.366. Excluded from employment are:

(a) Apprentices, interns and trainees who work *without* pay in cash or in kind (that is, unpaid trainee work);

(b) Participants in skills training or retraining schemes within employment promotion programmes, when *not* engaged in the production process of an economic unit;

(c) Persons who are required to perform work as a condition of continued receipt of a government social benefit such as unemployment insurance;

(d) Persons with seasonal jobs during the off season, if they *cease* to perform the tasks and duties of the job;

(e) Persons who retain a right to return to the same economic unit but who were absent for reasons specified in paragraph 5.362(b), when the total duration of the absence exceeds the specified threshold or if the test of receipt of remuneration is not fulfilled;

(f) Persons on indefinite lay-off who do not have an assurance of return to employment with the same economic unit;

(g) Persons who work to produce goods intended *mainly* or *exclusively* for consumption or use by the household or family, even if a surplus or part of the output is sold or bartered (that is, own-use production of goods, see paragraphs 5.433-5.438);
(h) Household members who provide *unpaid* services for consumption or use by their household (that is, own-use provision of services);
(i) Persons who work voluntarily and *without* pay to produce goods or services through or for other economic units, including market, non-market units and households (that is, volunteer work).

5.367. Information should be given in the census reports describing how the above-mentioned groups and other relevant groups were treated. Consideration should also be given to the desirability of identifying some of the groups (for example paid apprentices, interns and trainees) separately in tabulations.

b. Unemployed persons

5.368. *Unemployed persons* are all those above the specified age who (a) were not in employment, (b) carried out activities to seek employment during a specified recent period and (c) were currently available to take up employment given a job opportunity.

5.369. To be classified as unemployed, a person must satisfy all of the three criteria, where:

(a) "Not in employment" (that is, not engaged in work for pay or profit) is assessed with respect to the short reference period for the measurement of employment as defined in paragraph 5.359.

(b) To "seek employment" refers to any activity when carried out, during a specified recent period comprising the last four weeks prior to enumeration or calendar month, for the purpose of finding a job or setting up a business or agricultural undertaking. This includes also part-time, informal, temporary, seasonal or casual employment, paid apprenticeships, internships or traineeships, within the national territory or abroad. Examples of such activities are arranging for financial resources; applying for permits or licences; looking for land, premises, machinery, supplies or farming inputs; seeking the assistance of friends, relatives or other types of intermediaries; registering with or contacting public or private employment services; applying to employers directly, or checking at worksites, farms, factory gates, markets or other assembly places; placing or answering newspaper or online job advertisements; and placing or updating résumés on professional or social networking sites online.

(c) "Currently available" serves as a test of readiness to start a job in the present, assessed with respect to the same short reference period that is used to measure employment. Depending on national circumstances, the reference period may be extended to include a short subsequent period not exceeding two weeks in total, so as to ensure adequate coverage of unemployment situations among different population groups.

5.370. Unemployment has been one of the most widely used measures of labour underutilization. However, it only captures persons in situations of complete lack of work for pay or profit, and where opportunities for job search exist. In circumstances where there are few channels for seeking employment or where labour markets are limited in scope, or when labour absorption is inadequate, unemployment will not capture fully all persons with an unmet need for employment, as persons will take any available jobs, create their own jobs, often as own-account workers, become discouraged, or engage in production of goods for own final use (for example, subsistence agriculture or fishing).

Additional measures of underutilized labour include the *potential labour force*,[239] defined in paragraph 5.376, and persons in *time-related underemployment*.[240] Separate identification of these groups of persons supports better assessment of the different types of underutilization affecting labour markets across settings, and for more targeted policymaking.

5.371. It may be useful to distinguish first-time jobseekers, who have never worked before, from other jobseekers in the classification of the unemployed. Such a separation would be useful for policy purposes as well as in improving the international comparability of employment statistics. To do so, however, may require an additional question regarding previous work experience, which may impose too much of a burden for a population census.

Treatment of specific groups

5.372. Also classified as unemployed according to international standards are:

(a) *Future starters*, defined as persons "not in employment" and "currently available" who did not "seek employment" because they had already made arrangements to start a job within a short subsequent period, set according to the general length of waiting time for starting a new job in the national context but generally not greater than three months;

(b) Participants in skills training or retraining schemes within employment promotion programmes, who, on that basis, were "not in employment", not "currently available" and did not "seek employment" because they had a job offer to start within a short subsequent period, generally not greater than three months;

(c) Persons "not in employment" who carried out activities to migrate abroad in order to work for pay or profit but who were still waiting for the opportunity to leave.

(d) In accordance with the priority rule to establish their *labour force status* (see paragraph 5.355), persons who during the reference period were mainly students, homemakers, pensioners, registered unemployed or engaged in forms of work other than employment (for example, own-use production work, volunteer work), and who at the same time were "not in employment", carried out activities to "seek employment" and were "currently available", as defined above, should be classified as in unemployment. Information should be given in the census reports on how persons in these and any other specific groups were treated.

c. Persons outside the labour force

5.373. *Persons outside the labour force* comprise all those in the working-age population who in the short reference period were neither employed nor unemployed as defined above, including persons below the minimum age specified for the collection of economic characteristics.

5.374. Different classifications of persons outside the labour force may be used for analytical purposes. Particularly useful for informing labour market and social policies and programmes are classifications by *degree of labour market attachment* and by *main reason for not entering the labour force*. These alternative classifications can be derived from the same questions used to identify the unemployed and may be used separately or in combination to enable further analysis.

5.375. Persons outside the labour force may be classified by the *degree of labour market attachment* into the following groups:

[239] Replaces the optional relaxation of the "seeking work" criterion in the previous International Conference of Labour Statisticians standards.

[240] See Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, 2013).

(a) *Unavailable jobseekers*, that is, those "seeking employment" but not "currently available";

(b) *Available potential jobseekers*, that is, those not "seeking employment" but "currently available";

(c) *Willing non-jobseekers*, that is, those neither "seeking employment" nor "currently available" but who want employment;

(d) *Others*, that is, persons neither "seeking employment" nor "currently available" who do not want employment.

5.376. The classification of persons outside the labour force by degree of labour market attachment allows identification of the *potential labour force*, computed as the sum of (a) *unavailable jobseekers* plus (b) *available potential jobseekers*. Together with unemployment, the *potential labour force* is a key measure of labour underutilization, relevant both in more and less developed settings, especially where the conventional means of seeking employment are of limited relevance, where the labour market is largely unorganized or of limited scope, when labour absorption is, at the time, inadequate, or where persons are largely self-employed.

5.377. Although not a part of the potential labour force, the group (c) *willing non-jobseekers* represents another group of persons outside the labour force with an expressed interest in employment and is particularly relevant for social and gender analysis in specific contexts.

5.378. Persons outside the labour force may also be classified by their *main activity or reason for not entering the labour market* into the following groups. Some persons may be classifiable in more than one category. In such situations, priority should be given to the possible categories in the following order:

(a) *Attending an educational institution* refers to persons outside the labour force, who attended any regular educational institution, public or private, for systematic instruction at any level of education, or were on temporary absence from the institution for relevant reasons corresponding to those specified for employed persons "not at work".

(b) *Performing unpaid household services* refers to persons outside the labour force engaged in the unpaid provision of services for their own household, such as spouses and other relatives responsible for the care and management of the home, children and older persons. (Domestic and personal services provided by domestic employees working *for pay* in somebody else's home are considered as employed in line with paragraph 5.359 above).

(c) *Retiring on pension or capital income* refers to persons outside the labour force who receive income from property or investments, interests, rents, royalties or pensions from former employment.

(d) *Other reasons* refers to all persons outside the labour force who do not fall into any of the above categories (for example, children not attending school, those receiving public aid or private support and persons with disabilities unable to work).

5.379. Additional reasons for not entering the labour force that are considered particularly important at national or regional level, such as "engaged in own-use production of goods" (for example, foodstuffs from agriculture, fishing, hunting or gathering) should also be taken into account in the classification of population outside the labour force.

4. Characteristics of jobs and establishments

5.380. Once the labour force status of persons has been established, additional important topics regarding the labour market participation of the population relate to the characteristics of their jobs and of the establishments in which they work. These include, in particular, status in employment, occupation, place of work, industry, institutional sector, working time and income.

5.381. A *job* is defined as the set of tasks and duties performed or meant to be performed by one person for a single economic unit. Persons may have one or several jobs during the reference period. The *main* job is that with the longest hours *usually* worked even if the employed person was not at work in the reference period.[241]

5.382. Job-related characteristics are generally collected in reference to the main job for persons in employment, and may also be collected in reference to the *last main job* (if any) for persons not in employment (that is, unemployed or outside the labour force). This allows for classification of the labour force (that is, employed persons and unemployed persons) and of persons outside the labour force by characteristics of their (last) main job. Once the (last) main job is identified, it is essential that all subsequent questions refer to that same job, even if the respondent was not at work in the reference period. The census questionnaire or the census information taken from registers should be designed in a way that will ensure that the variables "status in employment", "occupation", "industry", and "institutional sector" are measured for the same job. This should be a central concern also for countries that rely on the use of administrative registrations for the capturing of the correct values of these variables.

5.383. The collection of data on characteristics of the *last main job* of unemployed persons, especially occupation, industry and status in employment, may be useful in order to inform policies aimed at promoting employability and job creation. To serve this purpose, it is generally recommended to set a time limit for past employment experience (for example, during the last five or ten years) and only collect information on the characteristics of the last main job if it was held within the time limit.

5.384. When secondary jobs held in the reference period are also identified, the questionnaire should be designed so as to enable clear and separate identification of characteristics relating to main and secondary jobs. Identification of secondary jobs is particularly important in countries where multiple job holding is commonplace, particularly in agriculture, and when collecting information on income from employment and working time, in order to support analysis of the relationship between employment, income and poverty.

5. Status in employment (core topic)

5.385. *Status in employment* refers to the type of explicit or implicit contract of employment with other persons or organizations that the employed person has in his or her job. The basic criteria used to define the groups of the classification are: the *type of economic risk* to which the worker may be exposed to the loss of financial or other resources in pursuance of the activity and the unreliability of remuneration; and *the type of authority* over the organization of the work and over the economic unit for which the work is performed. Care should be taken to ensure that an employed person is classified by status in employment on the basis of the same job used for classifying the person by "occupation", "industry" and "sector".

[241] Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, 2013), paragraph 12(b).

5.386. In October 2018, the Twentieth International Conference of Labour Statisticians (20th ICLS) adopted a new standard defining the classification of work relationships. The Resolution concerning statistics on work relationships, replaced the previous Resolution concerning the International Classification of Status in Employment (ICSE-93) set in January 1993 by the Fifteenth International Conference of Labour Statisticians (15th ICLS) regarding the International Classification of Status in Employment.

5.387. The new International Classification of Status in Employment (ICSE-18)[242], defined in the resolution provides improved boundaries and definitions and facilitates more detailed statistics on various types of work relationships, helping to describe the structure and changes in labour markets more accurately.

5.388. ICSE-18 introduces ten detailed categories that apply to activities defined as employment (see paragraph 5.390 below). These categories are based on the:

a) *type of authority* that the worker has to exercise control over the economic unit for which the work is carried out and for the organizing the work, and the extent to which the worker is dependent on another person or economic unit for organization of the work and/or for access to the market; and

b) *type of economic risk* the worker is exposed to in relation to a possible loss of financial or other resources related to the activity and unreliability of receiving remuneration.

5.389. One significant change with ICSE-18 is that these ten categories also can be organized in two ways: according to the type of authority (ICSE-18-A) and according to the type of economic risk (ICSE-18-R).

5.390. When organized by authority, ICSE-18 distinguishes between independent workers and dependent workers, making it useful for various types of labour market analysis such as the impact of economic cycles and government policies. When organized by economic risk, it categorizes workers into those employed for profit and those employed for pay. This hierarchy is typically preferred for national accounts, identifying wage employment, and producing statistics on wages, earnings, and labour costs.

5.391. Correctly identifying the ten detailed categories in ICSE-18 may require additional questions. Accordingly, it will be important to balance the benefits of providing detailed statistics with the extra burden and cost of including these questions in a census. A recommended approach to minimize burden and costs while still obtaining key statistics is to limit the identification of ICSE-18 to selected aggregated categories. This includes identifying employers, independent workers without employees, and contributing family workers, as defined by ICSE-18 based on authority.

5.392. For more detailed categories, such as identifying workers employed for pay, or distinguishing dependent contractors, these can be considered based on the country's objectives, keeping in mind the feasibility and relevance of such data as further discussed in paragraph 5.393-5.408.

5.393. **ICSE-18-A (according to type of authority):**

<u>Independent workers</u>

A. Employers

11. Employers in corporations

[242] The new international classification of status in employment (ICSE-18) adopted in a resolution at the 20th ICLS. https://www.ilo.org/publications/international-classification-status-employment-icse-18-manual

12. Employers in household market enterprises

B. Independent workers without employees

21. Owner-operators of corporations without employees

22. Own-account workers in household market enterprises without employees

Dependent workers

C. Dependent contractors

30. Owner-operators of corporations without employees

D. Employees

41. Permanent employees

42. Fixed-term employees

43. Short-term and casual employees

44. Paid apprentices, trainees and interns

E. Contributing family workers

51. Contributing family workers

5.394. An **employer** is a person who owns and manages their own business activities and has control over its operations. This can be done individually or with one or more partners. Employers hire one or more people as employees on a *regular basis*. Hiring employees on a *regular basis* includes engaging temporary employees (with an agreement of three weeks or more including the reference period), engaging different employees with shorter durations in subsequently weeks and employees that are temporary absent. It excludes the employers themselves, their partners, and contributing family workers.

5.395. An **independent worker without employees** is a person who operates a business alone or with partners but does not hire anyone regularly, except themselves, their partners, and contributing family workers. This category corresponds to "own-account worker" in ICSE-93.

5.396. Employers and independent workers without employees can, if feasible and relevant, be further categorized based on whether they operate a corporation or a household market, leading to the identification of the detailed categories: *employers in corporations*, *employers in household market enterprises*, *owner-operators of corporations without employees*, and *own-account workers in household market enterprises without employees*. This would require determining if the enterprise is incorporated. An incorporated enterprise holds a distinct legal identity, and its owner(s) bear limited liability for example in relation to any debts incurred by the enterprise. Incorporated enterprises can potentially be identified by adding an extra question to the census or by utilizing administrative records. The identification of these detailed categories would enhance the accuracy in differentiating between workers in employment for pay and workers in employment for profit, as defined in ICSE-18-R[243].

5.397. *Dependent contractors* are persons whose activities depend on another entity that exercise operational and/or economic control over the activities. They are not employees of that entity as they have a commercial agreement but are dependent on that unit for organization and execution of the work, income, or for access to the market. By definition, dependent contractors lack regular employees and do not own or operate an incorporated enterprise. Dependent contractors are a new category in ICSE-18, sharing characteristics of both independent workers and employees.

[243] See paragraph 24 in the 20th ICLS resolution concerning statistics on work relationships for the classification of status in employment according to economic risk (ICSE-18-R), adopted by the Twentieth International Conference of Labour Statisticians (Geneva, 2018).

5.398. The separate identification of dependent contractors typically requires additional questions to establish whether there is a dependent relationship and if control is exercised. In a Census, this might not be relevant unless there is a strong national need. If not separately identified dependent contractors will be categorized as independent workers or employees depending on the perception of the person as dependent contractors are to be found among both groups. The impact of this treatment could be estimated using a more suitable statistical source such as a labour force survey by for example cross tabulate ICSE-18 with ICSE-93.

5.399. If deemed relevant a minimized approach in a Census could be used to get a broad estimation of the share of dependent contractors in the country. Restricting the measurement to non-employees with commercial agreements that do not have regular employees and do not own an incorporated enterprise, focusing on identifying whether these persons have a main client or an intermediary client that sets the price for the goods and services produced[244].

5.400. An **employee** is a person who is in employment for pay, on a formal or informal basis, and who does not hold controlling ownership of the unit (enterprise, a non-profit institution, government unit or household) in which they are employed. Employees would typically be remunerated in cash or in kind for time worked but may also be paid for each task or piece of work done or for service provided. Employees would be accountable to a third party within the unit such as a person or board.

5.401. ICSE-18 provides for detailed categories of employees: *Permanent employees*, *fixed-term employees*, *Short-term and causal employees* and *Paid apprentices, trainees and interns*. If feasible and relevant these four detailed groups can be identified by establishing whether the employee has a permanent agreement until retirement (or until further notice) or a fixed-term agreement, and whether the fixed term-agreement is less than three months or three months or more. Paid apprentices, trainees and interns are employees who work to acquire workplace experience or skills in a trade or profession and receive payment in return for work performed. They can potentially be identified separately through a pre-coded category.

5.402. **Contributing family workers** are persons who assist a family member or a household member in a market-oriented enterprise operated by the family or household member. They do not receive regular payments, such as a wage or salary, in return for the work performed, but may receive some irregular payment. They do not make the most important decisions affecting the enterprise or have responsibility for it.

5.403. Member of a producers' cooperative do no longer constitute a status in employment category. In general, members of a producers' cooperatives are to be consider independent workers in ICSE-18 but may under some circumstances be regarded as dependent contractors if they depend significantly on the cooperative in terms of access to marker, organization of pricing of the work and fulfil the criteria to be classified as dependent contractors.

5.404. Seasonal workers, domestic workers, home-based workers, and workers in multi-party work relationships (including agency workers, and employees providing outsourced services) are according to the Resolution concerning statistics on work relationships so called cross-cutting categories as these would be groups that consists of several different status in employment categories. The definition and treatment of these groups can be found in the resolution (paragraph 73-124)[245].

[244] For the recommended approach for identifying dependent contractors see *ICSE-18 manual* at https://www.ilo.org/ilostat-files/Documents/ICSE-18_manual.pdf

[245] The Resolution concerning statistics of work relationships, adopted by the twentieth International Conference of Labour Statisticians (Geneva, October 2018).

5.405. When members of the armed forces paid in cash or in kind are counted among the employed, they should be included in the category of employees. However, because of the wide range of national practices in the treatment of the armed forces, it is recommended that census tabulations and related notes provide an explicit indication of the status in employment category in which they are included.

5.406. The treatment of additional groups such as work gang (crew) members, franchises, sharecroppers, communal resource exploiters, workers in cooperatives, outworkers and crowd workers can be found in the conceptual framework for Statistics on Work relationships as these groups are no longer addressed explicitly in the Resolution.

5.407. In most census questionnaires, the information concerning status in employment will be captured through pre-coded alternatives where only a few words can be used to convey the intended meaning of each category. This may mean that classification of some of the situations on the borderline between two or more categories will be carried out according to the subjective understanding of the respondent rather than according to the intended distinctions. This should be kept in mind in designing the questionnaire and also when presenting the resulting statistics. Countries that rely on the direct use of administrative records for the classification of persons according to status in employment may find that the groups "contributing family workers" and "dependent contractors" cannot be separately identified. Those who would have been classified as being in these groups when using a questionnaire may either be classified as part of one of the other groups or, in the case of contributing family workers, excluded from persons in employment.

5.408. Activities in areas such as agriculture, fishing, hunting, and gathering intended mainly for own consumption by the household are not considered employment and therefore not to be included in ICSE-18. Instead, participation in these productive activities is to be measured through the separate concept of "own-use production of goods" (see paragraphs 5.433–5.438) and categorized according to the International Classification of Status at Work (ICSaW-18) as defined in the Resolution concerning statistics on work s relationships[246].

6. Occupation (core topic)

5.409. *Occupation* refers to the type of work done in a job by the person employed (or the type of work done in the last job held, if the person is unemployed), irrespective of the industry or the status in employment in which the person's job should be classified. Type of work is considered in terms of the main tasks and duties performed in the job.

5.410. For purposes of international comparison, it is recommended that countries make it possible to prepare tabulations involving occupations in accordance with the latest revision available of the International Standard Classification of Occupations (ISCO). At the time the present set of census recommendations was approved, the latest revision was the one adopted by a Tripartite Meeting of Experts in Labour Statistics in 2007 and endorsed by the Governing Body of the International Labour Organization in 2008[247] and generally known as ISCO-08. Countries coding occupation according to a national standard classification should establish a correspondence with ISCO either through double coding or through mapping from the detailed groups of the national classification to ISCO.

[246] The Resolution concerning statistics of work relationships, adopted by the twentieth International Conference of Labour Statisticians (Geneva, October 2018), Paragraph 59.

[247] International Standard Classification of Occupations (ISCO-08), Volume 1, Structure, Group Definitions and Correspondence Tables (Geneva, International Labour Office, 2012). The ISCO companion guide is also available to assist countries using ISCO-08.

5.411. Countries should code the collected occupational responses at the lowest possible level of ISCO or a related national classification supported by the information given in each response. In order to facilitate detailed and accurate coding, it would be useful for the census questionnaire to ask each employed person for both the occupational title and a brief description of the main tasks and duties performed on the job. Information provided in response to the industry questions (see following section) may also be used to assist in the coding of occupation data, where the occupation response on its own is insufficient to assign a detailed occupation classification code.

5.412. In preparation for the coding of the occupation responses, the organization responsible for the census should prepare a coding index reflecting the type of responses that will be given by the respondents. The coding index should be constructed by occupational classification experts on the basis of responses to similar questions in other data collections, such as previous censuses, census tests and labour force surveys, as well as input from job placement officers of the employment service and the content of newspaper advertisements of vacant jobs. The coding index should clearly distinguish between responses belonging to "not elsewhere classified" categories and responses that do not provide enough information to determine an occupational group.

7. Industry (core topic)

5.413. *Industry* (branch of economic activity) refers to the kind of production or activity of the establishment or similar unit in which the job(s) of the employed or unemployed person was located during the time reference period established for data collection on economic characteristics.[248]

5.414. For purposes of international comparison, it is recommended that countries compile information on industry according to the most recent revision of the International Standard Industrial Classification of All Economic Activities (ISIC) available at the time of the census. At the time this present set of census recommendations was approved, the fourth edition of ISIC, adopted by the United Nations Statistical Commission at its thirty-seventh session in 2006, was the latest revision available. Countries coding industry according to a national standard classification should establish correspondence with ISIC either through double coding or through mapping from the detailed groups of the national classification to ISIC.

5.415. Countries should code the collected industry responses at the lowest possible level of ISIC or a related national classification supported by the information given in each response. In order to facilitate detailed and accurate coding, for each job to be coded the census questionnaire should ask for the main products and services produced, or the main functions carried out at the establishment or enterprise in which the person was employed. It is recommended that the name and address of the establishment should also be collected (see also paragraph 5.420). Countries with business registers that are complete and up to date can then use this response as a link to the register in order to obtain the industry code given there to the establishment.

[248] For those persons who are recruited and employed by one enterprise but who actually work at the place of another enterprise (called "agency workers" or "seconded workers" in some countries), there would be user interest in gathering information about the industry of the employer as well as the industry of the place of work. However, the collection of both would be more appropriate in a labour force survey rather than in a population census. The industry of the actual place of work may provide more reliable reporting of the "industry" variable in a population census. Any such choice should, however, be consistent with the treatment of this group in the System of National Accounts.

5.416. In preparation for the coding of the industry responses that cannot be matched to a pre-coded register, the organization responsible for the census should create a coding index that reflects the type of responses that will be given on the census questionnaire. This coding index should be constructed by industry classification experts on the basis of available lists of enterprises, establishments, businesses and so forth, as well as from responses to similar questions in other data collections, including previous censuses, census tests and labour force surveys. The coding index should clearly distinguish between responses belonging to "not elsewhere classified" categories and responses that do not provide enough information to allow for the coding of a detailed industry group.

8. Place of work

5.417. Two main topics related to the place of work of persons in employment are the *type of workplace* and its *geographic location*. The type of workplace refers to the nature of the place where the person performed his or her *main job* and distinguishes between the home and other workplaces, whether fixed or otherwise.

5.418. Three main categories, or a variation thereof necessitated by national circumstances, are recommended for classifying the *type of workplace*:

(a) *Work at home*. This category includes those who perform the tasks and duties of their main job from within the home, such as farmers who work and live on their farms, homeworkers, self-employed persons operating (work)shops or offices inside their own homes, and persons working and living at work camps.

(b) *No fixed place of work*. This category should be restricted to persons who, in performing the tasks and duties of their main job, travel in different areas and who do not report daily in person to a fixed address as a work base, for example, travelling salespersons, long-distance commercial vehicle drivers, seafarers, fishermen and own-account taxi drivers. It also includes ambulant vendors, operators of street or market stalls that are removed at the end of the workday, construction workers working at different sites during the reference period and push-cart operators.

(c) *With a fixed place of work outside the home*. All other persons in employment should be included in this category, including persons who move around in their job but have a fixed-base location to which they report daily, such as bus and taxi drivers (with a base), train and airline staff, and operators of street and market stalls that are not removed at the end of each workday. This group may also include individuals who travel to work, on a regular basis, across the national border to a neighbouring country. Fixed place of work outside the home includes: client's or employer's home; employer's workplace or site; own business premises; or, client's workplace or site.

5.419. It is likely that for some jobs, performance is at more than one location (for example, at home some of the time or season and in a fixed location outside the home at other times) or the category cannot be clearly distinguished. One approach, in the case of the former, would be to select the place where the individual spends or spent a major part of his or her working time. Where the distinction between categories is blurred, as is the case for work done, for example, on a rented plot of land adjacent to one's home, it would be useful to identify borderline cases, according to national circumstances. Specific instructions should be given to the enumerators on how to select between two or three possible responses to classify borderline cases.

5.420. The *geographic location* of the place of work can provide useful information for planning when used together with information on place of residence. To this end, countries may collect, for employed persons with a fixed place of work outside the home, information on the location of the

place of work (or the reporting place) during the reference period. The information collected should relate to the smallest civil division in which the job is performed, for example, in order to establish commuter flows from the place of residence to the place of work. Some countries investigating this topic in the population census have recorded the actual address of the place of work, allowing detailed tabulations and mapping of place of residence by geographic location of place of work. Information on actual address of the place of work can also be useful for industry coding (see paragraph 5.415) in countries where a business register has been developed that shows the industry code of each recorded establishment.

5.421. In some countries there may be concerns about the sensitivity of questions on the address of place of work owing to fears that there may be follow-up to a respondent's employer. In many developing countries, it may not be possible to gather information on actual address of place of work because street addresses do not exist, and for proxy responses, the address may not be known. In those situations, it would be useful to consider collecting information on the village, suburb, or similar low level of geography.

5.422. Additional questions may also be asked on the method of travel to work in order to produce statistics on travel-to-work patterns, valuable as basis for transportation planning. Persons not travelling to work should be classified as "no travel to work."

9. Institutional sector of employment

5.423. The *institutional sector of employment* relates to the legal organization and principal functions, behaviour and objectives of the enterprise with which a job is associated.

5.424. Following the definitions provided in the System of National Accounts, distinction should be made between the following institutional sectors:

(a) *Corporation*, comprising non-financial and financial corporations (in other words incorporated enterprises, private and public companies, joint stock companies, limited liability companies, registered cooperatives, limited liability partnerships, and so forth) and quasi-corporations (that is to say, an unincorporated enterprise that is managed as if it were a corporation, in that a complete set of accounts is kept), as well as non-profit institutions, such as hospitals, schools and colleges that charge fees to cover their current production costs;

(b) *General government*, comprising central, state and local government units together with social security funds imposed or controlled by those units, and non-profit institutions engaged in non-market production controlled and financed by government, or by social security funds;

(c) *Non-profit institutions serving households* (for example, churches, professional societies, sports and cultural clubs, charitable institutions and aid agencies) that provide non-market goods and services for households (that is to say, free or at prices that are not economically significant) and whose main resources are from voluntary contributions;

(d) *Households* (including unincorporated enterprises owned by households) comprising unincorporated enterprises directly owned and controlled by members of private and institutional households (made up of persons staying in hospitals, retirement homes, convents, prisons and so forth, for long periods of time), either individually or in partnership with others. Partners may be members of the same household or from different households.

5.425. In most census questionnaires, the information concerning institutional sector of employment will be captured through precoded alternatives where only a few words can be used to convey the intended meaning of each category. This may mean that classification of some units on the borderline between two or more categories will be carried out according to the subjective understanding of the respondent rather than according to the intended distinctions. This should be kept in mind when presenting the resulting statistics.

10. Working time

5.426. The number of employed persons provides only a very rough estimate of the volume of work performed, especially when such persons have non-standard working hours. Inclusion in the census of an item on time worked helps to ensure a more accurate measurement of the concept by capturing the full contribution of persons who were in and out of the workforce or who worked only for a brief time during the year (for example, women).

5.427. To provide a comprehensive measure of working time in employment that will best inform policy and analytical needs, it may be preferable to collect information about the total hours worked in all jobs rather than to limit the information to hours worked in the main job.

5.428. Information on two distinct concepts of working time can be collected in a population census: *hours actually worked* and *hours usually worked*.

5.429. *Hours actually worked* is defined as the time spent in a job for the performance of activities that contribute to the production of goods and/or services during a specified reference period. It covers the time spent in "direct hours", in "related hours", "down time" and short "resting time". "Direct hours" is the time spent carrying out the tasks and duties of the job – and may be performed in any location. "Related hours", while not leading directly to goods produced or services provided, is the time spent maintaining, facilitating or enhancing productive activities, including upkeep of the workplace, changing time or decontamination of work clothes, purchasing or transporting materials, waiting for business, customers or patients, on-call duties, travelling between work locations, and work training or skills enhancement required by the economic unit. In practice, "down time" includes unavoidable, temporary interruptions to work (for example machinery or Internet breakdown, lack of supplies). "Resting time" is inactive time for short rest or refreshment in the course of performing job-related activities, (for example coffee breaks). Longer breaks for meals, time spent not working because of vacation, holidays, sickness, industrial disputes, etc., commuting to work (if not also performing job tasks or duties) and educational leave even if paid, are excluded from hours actually worked.

5.430. Measurement of hours actually worked in employment, in the context of the population census, is usually collected using one direct question; it is optimally measured using a set of questions, requesting hours separately for each day of the week. For employed persons not at work in the short reference period, it is possible to have a value for hours actually worked of zero (for persons away on leave) or reduced (if a part of the reference period was taken off for sickness, holiday, or other purpose).

5.431. *Hours usually worked* is defined as the typical value of the hours actually worked in a job per short reference period (for example one week) over a long observation period (month, quarter, season, year) that comprises the short reference period itself. This "typical value" of time worked during a normal or typical week may be the modal number of the hours actually worked in the short period as distributed over the long period. This would include overtime hours regularly worked

whether paid or unpaid. Days and hours not usually worked and unusual periods of overtime are not included.

5.432. Measurement of hours usually worked in employment relating to the short reference period of one week can be done with one direct question: how many hours do you usually work per week (in your main job or in all jobs)? For persons with more than one job during the reference week, to record both working time in the main job (for which the other descriptive variables are collected) and total working time (sum of working time in all jobs) the questionnaire would require, at minimum, two questions.

11. Participation in own use production of goods (core topic)

5.433. Countries where production of goods for own final use (such as foodstuffs from agriculture, fishing, hunting and gathering, water, firewood and other household goods) represents an important component of the livelihood of a part of the population, whether as a main or secondary activity, will need to consider collecting information in the population census on the number of persons engaged in this form of work (previously included within the concept of employment). Such information is essential for benchmarking purposes, especially where household surveys are not frequent, for comprehensive sectoral analysis, particularly of work in agriculture, forestry and fishing, and to enable integration of the population census with the agricultural census (see also paragraphs 5.453–5.461).

5.434. *Persons in own-use production of goods* are all those above the specified age who, during a specified reference period, performed "any activity" to produce goods for own final use. The notion "for own final use" is interpreted as production where the intended destination of the output is *mainly* for final use by the producer in the form of capital formation, or final consumption by household members, or by family members living in other households.

5.435. According to international standards, "any activity" to produce goods (within the 2008 System of National Accounts production boundary) covers work performed for at least one hour in the following activities, when the intended destination of the output is *mainly* for own final use, as specified above:

(i) Producing and/or processing for storage agricultural, fishing, hunting and gathering products;
(ii) Collecting and/or processing for storage mining and forestry products, including firewood and other fuels;
(iii) Fetching water from natural and other sources;
(iv) Manufacturing household goods (such as furniture, textiles, clothing, footwear, pottery or other durables, including boats and canoes);
(v) Building, or effecting major repairs to, one's own dwelling, farm buildings, etc.

5.436. For measurement purposes, the intended destination of the output is established in reference to the specific goods produced, based on self-declaration (that is, mainly for own final use). In the case of goods from agriculture, fishing, hunting or gathering intended mainly for own consumption, a part or surplus may nevertheless be sold or bartered.

5.437. Persons may engage in own-use production of goods as a main or secondary activity, throughout the year or on a seasonal basis. To ensure complete coverage, the census questions on participation in own-use production of goods should be applied to all persons above the specified age for collecting information on the economic characteristics of the population, *irrespective* of their labour force status. The reference period may refer to the last 12 months, calendar year, agricultural

year or season, as relevant to national circumstances. Where pertinent, the choice of reference period should promote coherence with the agricultural census (see also paragraphs 1.71–1.77).

5.438. For assessments of the volume of work performed by persons in own-use production of goods, particularly when using a long reference period, it may be useful to include a question on working time, in particular hours usually worked (see paragraph 5.431), or based on broad categories such as part time or full time, part year or full year, number of months, as feasible and relevant to the main uses of the statistics.

12. Income

5.439. Countries may wish to collect information on the amounts of income received by individual persons or households during a specified reference period, from any source. If this topic is included in the census it is recommended that data be obtained for all persons above a specified age, whether they are employed or not. Income may be measured at the household level, or for each individual in the household.

5.440. *Income* may be defined as all receipts whether monetary or in kind (goods and services) that are received by the household or by individual members of the household at annual or more frequent intervals, but excluding windfall gains and other such irregular and typically one-time receipts. Household income covers (a) income from employment (both paid and self-employment); (b) income from the production of goods for own final use; (c) income from the provision of household services for own final use; (d) property income; and (e) current transfers received.[249]

5.441. The collection of reliable data on income, especially income from self-employment and property income, is extremely difficult in general field enquiries, particularly population censuses. The inclusion of non-cash income further compounds the difficulties. Collection of household income data in a census, even when confined to cash income, presents special problems in terms of burden of work, response errors, and so forth. Therefore, this topic is generally considered more suitable in a sample survey of households or from administrative data sources such as tax or social security records. Depending on the national requirements, countries may nonetheless wish to obtain limited information on personal or household income, by covering only some of the income components (such as income from employment), for shorter reference period (such as one month), and cover only cash income. As thus defined, the information collected can provide some input into statistics that have many important uses.

5.442. According to international standards on the subject, the income from employment of employed persons should include wages and salaries of employees, income of members from producers' cooperatives and the mixed income of employers and own-account workers operating business and unincorporated enterprises. In addition to the income from employment of employed household members, the total income of the household should include, for example, the interest, dividends, rent, social security benefits, pensions and life insurance annuity benefits of all its members. The *Handbook on household income statistics*[250] provides further guidance on concepts and methods related to this topic.

[249] See Resolution concerning household income and expenditure statistics, adopted by the Seventeenth International Conference of Labour Statisticians (Geneva, 2003), paragraphs 4–5. Available at http://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_087503.pdf.

[250] *Canberra Group Handbook on Household Income Statistics*, second edition, 2011, available at http://www.unece.org/fileadmin/DAM/stats/groups/cgh/Canbera_Handbook_2011_WEB.pdf.

5.443. The concepts involved in determining income are not simple to grasp and respondents may be unable or unwilling to provide exact information. For example, income should include social security benefits, pension fund contributions and direct taxes withheld from employees' salaries, but some persons will undoubtedly not include these amounts in reporting their salaries. Significant items of total household income may also be excluded or misstated. Despite instructions given to enumerators, the data collected can therefore only be expected to be approximate. Accordingly, in the presentation of results it is usually appropriate to use broad income or earnings size classes. As an aid to the interpretation of the results, tabulations of the data should be accompanied by a description of the items of income assumed to be included and, if possible, an estimate of the accuracy of the figures.

13. Informal employment

5.444. *Informal employment* is defined as any activity of persons to produce goods or provide services for pay or profit that is -in law or in practice- not covered by formal arrangements such as commercial laws, procedures to report economic activities, income taxation, labour legislation and social security laws and regulations. Informal employment includes the activities carried out in relation to an informal job, creating a link between the definition of an informal job and the definition of informal employment. Persons holding informal jobs includes independent workers (employers and own account workers) that own and operate an informal enterprise, and dependent workers whose job do not have a formal status in relation to the legal administrative framework or whose activities are not effectively covered by formal arrangements. The operational definition of informal and formal jobs thus depends on the persons status in employment category, holding the specific job.[251]

5.445. Informal employment is a non-core topic defined in the Resolution concerning statistics on the informal economy, adopted at the twenty first International Conference of Labour Statisticians in 2023. The resolution replaces the two preceding standards, which separately defined the informal sector and informal employment[252]. It encompasses a comprehensive statistical framework for the informal economy, incorporating conceptual and operational definitions for its various statistical components, including the informal sector and informal employment.

5.446. Informal employment constitutes a significant portion of the economies and labour markets across countries, creating employment opportunities, and income generation. However, informality also exposes individuals to heightened vulnerability and precarious circumstances and impacts on the adequacy of earnings, access to social protection and employment benefits, compromised occupational safety and health, and on the overall working conditions.

[251] *Resolution concerning statistics on the informal economy*, adopted by the twenty first International Conference of Labour Statisticians (Geneva, 2021), paragraphs 55–91. Available at: https://ilostat.ilo.org/about/standards/icls/icls-documents/

[252] Respectively: *the Resolution concerning statistics of employment in the informal sector*, fifteenth International Conference of Labour Statisticians, 1993, available at: https://www.ilo.org/global/statistics-and-databases/standards-and-guidelines/resolutions-adopted-by-international-conferences-of-labour-statisticians/WCMS_087484/lang--en/index.htm. and *the Guidelines concerning a statistical definition of informal employment*. Seventeenth International Conference of Labour Statisticians, 2013, available at: https://www.ilo.org/global/statistics-and-databases/standards-and-guidelines/guidelines-adopted-by-international-conferences-of-labour-statisticians/WCMS_087622/lang--en/index.htm

5.447. Including questions to classify those employed according to the formal or informal nature of their job in a census can allow the generation of estimates of informality for small geographical areas as well as for small population groups, and by that potentially complement data on informality collected through other sources such as for example the labour force survey.

5.448. The operational definition of informal employment depends on the person's employment status category. Distinct operational criteria are applied, contingent on whether the person holds a specific job as an **independent worker** (an employer or independent worker without employees) or a dependent worker i.e., an **employee**, **contributing family worker**, or **dependent contractor**). As several criteria are applied for defining informal and formal jobs[253] it would typically be a need to limit the criteria used to further reduce the response burden, if measured in a census.

5.449. For **independent workers** the main criterion of registration of the enterprise can be used to determine the formal status of the enterprise and thereby the informal or formal status of their jobs. If the enterprise is registered then the job held by the independent worker is to be considered a formal job, and if not, then an informal job. The type of registration would depend on the registration system within the country but should be a governmentally established registration used for granting access to benefits such as tax deductions, obtaining a separate legal identity for enterprises and carrying obligations such as paying business tax and keeping accounts. Typically, it would be the national business register, commercial register, tax register or similar, depending on the registration system in the country.

5.450. For **employees** the main criterion of *employer's contribution to statutory social insurance* can be applied to define the informal or formal nature of their jobs. If the employer makes contributions on the behalf of the employee, then the job held by the employee is considered to be formal, if not, informal. If deemed necessary and feasible the additional criteria of *access to paid annual leave* and *access to paid sick leave* could be added to further support the definition[254.] When operationalizing the criterion of *employer's contribution to statutory social insurance* countries need to take the national context and social security laws into account. However, it is essential that it is a social insurance linked to the national social security system that depends on holding the specific job. It therefore excludes universal non-contributory social protection schemes and private insurances. Typically, employer's contribution to a pension fund on behalf of the employee would be relevant for operationalization but other types of job-related statutory insurances could also be of relevance, depending on the national context.

5.451. **Contributing family workers** can, with the new standard, be regarded as having formal jobs if formal arrangements such as social insurance are in place for this group. This requires establishing whether *the family business is formal* (by applying the same criteria as for independent workers) and whether they have access to such formal arrangements i.e., *whether contributions are made to the statutory social insurance scheme*. If no such arrangements are in place in the country, or if there is a need to reduce the respondent burden in the census, *then the jobs held by contributing family workers can be considered informal by default*.

[253] Resolution concerning statistics on the informal economy, adopted by the twenty first International Conference of Labour Statisticians (Geneva, 2021), paragraphs 66–91. Available at: https://ilostat.ilo.org/about/standards/icls/icls-documents/

[254] Access to paid annual leave and paid sick leave can for example be used in case the respondent might not know whether the employer contributes to social insurance. If the respondent has access to both then the job may still be categorized as formal

5.452. If **dependent contractors** are explicitly identified in the census, as outlined in paragraphs 5.397-5.398, the distinction between formal and informal jobs within this category hinges on the operational criteria of enterprise registration, the registration of the dependent contractor for tax on profits, and their contribution to job-related social insurance such as a statutory pension fund. The job held by the dependent contractor is deemed formal when they have a registered enterprise or are registered for tax on profits, coupled with contribution to social insurance. If there is a strong need for reducing the response burden, the operational definition may opt to rely solely on contributions to social insurance as a basis for estimating dependent contractors with formal and informal jobs. Countries shall also pursue investing in methodology to collect detailed data on work using administrative sources.

J. Agriculture

1. Introduction

5.453. This section presents two non-core topics related to agriculture. These topics are helpful for countries aiming to gather information in their population census that can be used to create a frame (sampling list) of agricultural holdings in the household sector for a future agricultural census (see paragraphs 1.71-1.77).

5.454. Some countries use population and housing census data like occupation and employment status to identify individuals involved in agriculture. However, as mentioned in paragraph 1.73, this approach has limitations. The short reference period (often a week) might miss people involved in seasonal agricultural activities or those for whom agriculture isn't their main activity. To address this issue, two non-core topics at the household level are proposed.

2. Own-account agriculture production

5.455. Some countries might use the population census to identify households engaged in "own-account agricultural production." This refers to households raising livestock or using land (wholly or partly) for agricultural purposes. Households with members engaged in agricultural activities only as paid employees wouldn't qualify. The definitions used here align with those outlined in Resolution I of the 19th ICLS, which includes own-use production of goods (see paragraph 5.434). Paragraph 5.379 highlights the importance of capturing information on own-use production, especially "in countries where populations rely on agriculture, fishing, or hunting and gathering for their own consumption." This data is crucial for integrating population and agricultural censuses. Including this topic can facilitate analysis of the population census data related to agriculture and its use as a frame for a subsequent agricultural census or surveys. However, comprehensive data collection would require measuring both employment and own-use production in agriculture.

5.456. The question on own-account agriculture production gathers information at the household level. It asks if any household member participated in own-account agricultural production activities during the year preceding the census, either at their usual residence or elsewhere. As explained earlier, a year-long reference period is necessary to capture seasonal variations, which wouldn't be possible with the shorter reference period used in topics pertaining to economic characteristics.

3. Size of farm and number of livestock

5.457. The second topic collects information at the household level on the size of the farm (land area or number of plots) used for agriculture and the number of livestock owned on the census day. This helps assess the extent of a household's agricultural activities. For some countries, collecting such

detailed information could prove difficult and/or misleading without collecting further details on the types of livestock. In such cases, a question with a dichotomic (yes/no) answer might be more feasible. For countries seeking a more comprehensive picture of agricultural activities within population and housing censuses, gathering data on employment in agriculture and own-use production of goods is recommended.

5.458. "Agricultural activities" encompass all the major agricultural activities relevant to the country (including crops, livestock, and related activities). Information on aquaculture, forestry, and fishing can also be collected if significant for the country.

5.459. Countries aiming for more extensive agricultural data can include an agricultural module with frame data items recommended by the FAO World Programme for the Census of Agriculture 2030 (WCA 2030, FAO 2026) and the FAO/UNFPA Guidelines for Linking Population and Housing Censuses with Agricultural Censuses (FAO and UNFPA, 2012).

5.460. For countries where household-level aquaculture production is important, information can be collected on whether any household member participates in any form of own-account aquacultural production activities.

5.461. Agricultural production activities refer to groups 011, 012, 013, 014, and 015 of ISIC (Rev. 5.0):

- Group 011: Growing of non-perennial crops
- Group 012: Growing of perennial crops
- Group 013: Plant propagation
- Group 014: Animal production
- Group 015: Mixed farming

Aquacultural production activities refer to group 032 of ISIC (Rev. 5.0):

- Group 032: Aquaculture

Chapter II. Housing census topics

I. Factors determining the selection of topics

5.462. In line with the overall approach to revision 4 of *Principles and Recommendations for Population and Housing Censuses*, the selection of housing census topics, as with the population topics described in Chapter I, is based on the outputs expected to be produced. Therefore, the first step involves the clear identification of expected outputs; the core and additional topics are then decided on that basis. It is recommended that countries collect data on the core topics and also produce the recommended tabulations, to improve the international harmonization and comparability of statistics through the use of common concepts, definitions and classifications[255]. Use of an agreed international approach would also enhance the capacity of countries to generate statistics for monitoring the socioeconomic situation of their populations, including for the provision of data for the internationally agreed development goals.

5.463. In reference to the selection of topics to be included in a housing census, limiting statistical enquiries to the collection of data that can be processed and published within a reasonable period of time is very important. Such cautions are especially applicable to a housing census, since it is customary to conduct housing and population censuses simultaneously or as consecutive operations. There is a high probability that the amount of data required from a census may be beyond the capacity of enumerators to collect or the census agencies to process. It may be sufficient in some developing countries, for example, to ascertain only the number of housing units and other sets of living quarters of various types, the number and characteristics of the occupants thereof and the availability of a water supply system.

5.464. In this context, countries should not attempt to collect housing data that are so incomplete that they fail to serve the principal purposes for which they are required. It is important, therefore, for census takers to consult closely with the principal users at an early planning stage in order to identify the data that are of highest priority and the means of supplying them in the most useful formats.

5.465. The topics, therefore, to be covered in a housing census (that is to say, the subjects for which information is to be collected regarding living quarters, households and buildings) should be based on a balanced consideration of:
(a) **National priority**: It is essential to consider the diverse needs of data users in the country at both the national and local area levels to ensure that the census generates relevant information to effectively address these needs.
(b) **International comparability**: It is crucial to ensure that the data collected is comparable at regional and global levels through aligning with international standards. This would facilitate cross-country analyses and monitoring.
(c) **Suitability**: For information collected through questionnaires, consider the sensitivity of the topics and the respondent burden, i.e., the willingness and ability of the public to provide accurate information on these topics. Also, assess the technical competence of enumerators to obtain information on the topics to be collected in field enumeration, ensuring they are well-trained and capable. Similarly, with respect to information to be collected through administrative registers, consider the sensitivity of the topics and the willingness of the public to have their personal information linked and used, i.e., the acceptability of the use of existing administrative information for statistical purposes. It is also important to consider the

[255] Link to UNSD website containing the tabulation shells.

technical competence of data scientists in regard to obtaining information on the topics through data linkage.
(f) **Resources**: Consider the total national resources available for conducting the census, including financial, human, and technological assets.
(g) **Alternative sources**: Assess the availability of relevant information from alternative data sources to reduce duplication and improve the efficiency of data collection.

5.466. Such a balanced consideration will need to take into account the advantages and limitations of alternative methods of obtaining data on a given topic within the context of an integrated national programme for gathering housing statistics.

5.467. In selecting the housing topics, regard should also be given to the usefulness of historical continuity, which provides the opportunity for comparison of changes over a period of time. Census takers should avoid, however, collecting information that is no longer required by users. Information should not be collected simply because it was traditionally collected in the past, bearing in mind changes in the socioeconomic and housing circumstances of the country. It becomes necessary, therefore, in consultation with a broad range of users of census data, to review periodically the value of even long-standing topics and to re-evaluate the need for their continued collection, particularly in the light of new data needs and alternative data sources that may have become available for investigating topics hitherto covered in the population and housing census. Each of five key factors that need to be taken into account in reaching a final decision on census content are briefly reviewed in the following paragraphs.

A. National priority

5.468. The priority of designing a housing census should (as with the design of the population census) be to meet national needs. The prime consideration is that the census should provide information on those topics that are of greatest value to the country, with questions framed so as to elicit data of maximum utility. Each country's decision with regard to the topics to be covered should depend upon a balanced appraisal of how urgently the data are needed and whether the information could be equally well or better obtained from other sources.

5.469. Some countries may omit from the census certain recommended topics because there is not a national need to collect the data. For example, a particular amenity, such as electricity or toilet facilities, might be available virtually everywhere in a country, and, consequently, there may be no need to collect such information in a census at all. Conversely, some topics may not be included in a census because of the almost total absence of certain amenities, particularly in the rural areas of some developing countries.

5.470. The importance of involving stakeholders in the process of identifying priorities and policy needs has to be taken into consideration early in the process of designing the housing census. The topics that are of particular interest to policymakers need to be carefully assessed in terms of applicability, reliability of data and census limitations (number of questions, and so forth). More detailed information on involvement of stakeholders is presented in Part Two, in Chapter X, on "User consultation, communication and publicity" (paragraphs 2.157–2.165), and also in the *Handbook on Census Management for Population and Housing Censuses*.[256]

[256] United Nations publication, Sales No. E.00XVII.15 Rev.1.

B. International comparability

5.471. The desirability of achieving regional and worldwide comparability should be another major consideration in the selection and formulation of topics to be included in the census. National and international objectives are usually compatible, since broad studies of countries' experiences and practices are the basis of international recommendations.

5.472. If particular circumstances within a country necessitate a departure from international standards, every effort should be made to explain these departures in the census publications and to indicate how the national presentation can be adapted to the international standards.

C. Suitability

5.473. A prerequisite for the inclusion of housing topics in the census should be the willingness and ability of respondents to provide accurate information on them, or, for register-based censuses, the willingness of the public to have their housing information linked and used. It is advisable to avoid topics that could increase the burden on respondents and those that are likely to arouse fear, local prejudice or which might be used to deliberately promote political or sectarian causes, as these are likely to have a detrimental effect on response rates and support of the census. In an interview-based census or where the collector needs to obtain information through observation, consideration needs also to be given to the level of knowledge and skills of the interviewers or collector and whether they can be adequately trained to collect this information accurately. Topics that are too complicated or difficult for the average respondent or enumerator to answer quickly should also not be included. The exact phrasing of questions that will obtain the most reliable responses may depend on national circumstances and, as described in Part Three, should be well tested prior to the census (see paragraph 3.139).

D. Resources

5.474. The selection of topics should be carefully considered in relation to the total resources available for the census. An efficient collection of accurate data for a limited number of topics, followed by prompt tabulation and publication, is more useful than the collection of data for an overambitious list of topics that cannot be properly processed and disseminated. In balancing the need for data against resources available, the extent to which questions can be precoded is yet another consideration. This may be an important factor in determining whether or not it is economically feasible to include certain topics in the census. In the case of those topics on which data is taken from registers, the additional resources required to transform the administrative records into useable statistical data should also be noted.

E. Alternative sources

5.475. In the selection of topics to be investigated in a housing census, consideration should be given to whether data are available from other sources, such as surveys and administrative records, taking into account their relative advantages and limitations. Those topics for which no alternative sources exist should be given higher priority while those for which alternative sources are readily available should be accorded lower priority.

II. List of topics

5.476. The units of enumeration for housing censuses are buildings, living quarters, households and occupants. The building is often an indirect but important unit of enumeration for housing censuses since the information concerning the building (building type, material of construction of external walls and certain other characteristics) is required to describe properly the living quarters located within the building and for the formulation of housing programmes. In a housing census, the questions on building characteristics are normally framed in terms of the building in which sets of living quarters being enumerated are located, and the information is recorded for each of the housing units or other sets of living quarters located within it. Administrative registers may contain useful information concerning buildings.

5.477. The principal direct enumeration unit in a housing census is the living quarters. Only by recognizing this as such can data be obtained that will provide a meaningful description of the housing situation and a suitable basis for the formulation of housing programmes.

5.478. The second direct unit of enumeration is the households occupying the living quarters. For each household, it is often useful to collect information on the characteristics of the head or reference person, tenure in the housing unit, and other relevant characteristics.

5.479. The final units of enumeration are the occupants within households. However, the detailed characteristics of each individual household member are collected in a population census and are covered in Chapter I.

5.480. The list presented below (Table 4) is based on the global and regional census experience of the last several decades. The topics included are those on which there is considerable agreement on their importance and feasibility for inclusion in a census for the purpose of measuring and evaluating housing conditions and formulating housing programmes. Those that are likely to present difficulties and require time-consuming questioning can probably best be investigated in a separate housing survey of a sample of living quarters.

5.481. Core topics are those of common interest and value to countries and also of importance in enabling comprehensive comparison of statistics at the international level. Other topics refer to data that need to be collected in order to meet the additional requirements of national users.

5.482. It should be emphasized that the topics or variables on housing contained herein are for tabulation and production of outputs as this is the overall orientation of these guidelines. Issues that pertain to data collection are addressed in other parts of this document and other relevant United Nations handbooks.

Table 4. Housing census topics by unit of enumeration

No.	Topic	Living quarter		Building	Household
		Housing unit	Collective living quarter		
1	**Living quarters – type of** (paras. 5.485– 5.526)	□	♦		
2	**Location of living quarters** (paras. 5.527– 5.534)	□	♦	□	□
3	**Occupancy status** (paras. 5.535– 5.539)	♦			
4	**Ownership – type of** (paras. 5.540– 5.545)	♦			
5	**Rooms – number of** (paras. 5.546– 5.548)	♦			□
6	Bedrooms – number of (paras. 5.549– 5.550)	○			○
7	Useful floor space – (paras. 5.551– 5.552)	○	○		○
8	**Water supply system** (paras. 5.553– 5.556)	♦	○		□
9	**Drinking water – main source of** (paras. 5.557– 5.558)	♦	○		□
10	**Toilet – type of** (paras. 5.559–5.562)	♦	○		□
11	**Sewage disposal** (para. 5.563)	♦			□
12	**Solid waste disposal – main type of** (paras. 5.564– 5.565)	♦			□
13	**Bathing facilities** (paras. 5.566– 5.568)	♦	○		□
14	**Kitchen – availability of** (paras. 5.569– 5.572)	♦	○		□
15	**Fuel used for cooking** (para. 5.573)	♦			□
16	**Energy used for lighting – type of** (paras. 5.574– 5.575)	♦	○		□
17	Heating – type and energy used (paras. 5.576– 5.577)	○			○
18	Hot water – availability of (para. 5.578)	○			○
19	Piped gas – availability of (para. 5.579)	○			○
20	Use of the housing unit (paras. 5.580– 5.581)	○			○
21	**Occupancy by one or more households** (paras. 5.582– 5.586)	□			♦
22	**Occupants – number of** (paras. 5.587– 5.588)	♦	♦		□
23	**Building – type of** (paras. 5.589-5.597)			♦	
24	Year or period of construction (paras. 5.598– 5.602)	○		○	
25	Dwellings in the building – number of (para. 5.603)	○		○	
26	Position of dwelling in the building (paras. 5.604– 5.606)			○	
27	Accessibility to dwelling (para. 5.607)	○			
28	**Construction material of outer walls** (paras. 5.608– 5.610)	♦		♦	
29	Construction material of floor and roof (para. 5.611)	○		○	
30	Elevator – availability of (paras. 5.612–5.613)	○		○	
31	Farm building (para. 5.614)	○		○	
32	State of repair (paras. 5.615– 5.616)	○		○	
33	**Age and sex of the reference person of the household** (paras. 5.617– 5.618)				♦
34	**Tenure** (paras. 5.619– 5.622)				♦
35	Rental and housing costs (paras. 5.623– 5.625)				○
36	Furnished/unfurnished (para. 5.626)	○			○

No.	Topic	Living quarter		Building	Household
		Housing unit	Collective living quarter		
37	**ICT devices – availability of** (paras. 5.627– 5.631)				♦
38	Cars – number of available (para. 5.632)				○
39	Durable household appliances – availability of (para. 5.633)				○
40	Outdoor space – access to (para. 5.634)				○
	Legend: ♦ - Core topic □ - Core topic, derived ○ - Additional topic				

III. Definitions and specifications of topics

5.483. Paragraphs 5.485–5.634 below contain the recommended definitions and specifications of the housing topics.

5.484. It is important that census data be accompanied by the definitions used in carrying out the census. It is also important that any changes in definitions that might have been made since the previous census be indicated and, if possible, accompanied by estimates of the effect of such changes on the relevant data. In this way, users will not confuse valid changes over time with increases or decreases that have occurred as the result of changed definitions.

1. Living quarters – type of (core topic)

i. Definition of living quarters

5.485. *Living quarters* are structurally separate and independent places of abode. They (a) may have been constructed, built, converted or arranged for human habitation, provided that they are not at the time of the census used wholly for other purposes and that, in the case of non-conventional housing units and collective living quarters, they are occupied at the time of the census; or (b) though not intended for habitation, were in use for such a purpose at the time of the census.

5.486. In a census with a field enumeration, instructions should be issued to field staff so that it is clearly understood at what stage of completion living quarters should be in order to be included. Living quarters being demolished or awaiting demolition should normally be excluded. The system used should be consistent with that employed for the system of current housing statistics and should avoid double counting where construction statistics are used to bring the census data up to date. Special instructions will need to be issued concerning “core dwellings” in countries where these are provided within a preliminary phase of dwelling construction (see paragraphs 5.502–5.505).

ii. Classification of living quarters

5.487. Living quarters are either housing units or collective living quarters. Normally, the collection of information concerning buildings and housing units located within buildings is of prime importance in a housing census, since it is in buildings and housing units that the majority of the population permanently lives. Furthermore, housing units are intended for occupancy, or are occupied, by households, and it is

with the provision of accommodation for households that housing programmes and policies are mainly concerned. However, certain types of "collective living quarters" are also of significance with respect to the housing conditions of households; these include hotels, rooming houses and other lodging houses and camps occupied by households. Housing units should be classified so as to distinguish conventional dwellings from other types of housing units. It should be emphasized that without an adequate classification of living quarters, no meaningful analysis of housing conditions based on housing census data is possible.

5.488. The classification outlined below (see also Figure 5) and a system of three-digit codes have been designed to group in broad classes housing units and collective living quarters with similar structural characteristics. The distribution of occupants (population) among the various groups provides valuable information about the housing accommodation available at the time of the census. The classification also affords a useful basis of stratification for sample surveys. The living quarters may be classified into the following categories:

1 Housing units
- 1.1 Conventional dwellings
  - 1.1.1 Has all basic facilities
  - 1.1.2 Does not have all basic facilities
- 1.2 Other housing units
  - 1.2.1 Semi-permanent housing units
  - 1.2.2 Mobile housing units
  - 1.2.3 Informal housing units
  - 1.2.4 Housing units in permanent buildings not intended for human habitation
  - 1.2.5 Other premises not intended for human habitation

2 Collective living quarters
- 2.1 Hotels, rooming houses and other lodging houses
- 2.2 Institutions
  - 2.2.1 Hospitals
  - 2.2.2 Correctional institutions (prisons, penitentiaries)
  - 2.2.3 Military institutions
  - 2.2.4 Religious institutions (monasteries, convents, and so forth)
  - 2.2.5 Retirement homes, homes for older persons
  - 2.2.6 Student dormitories
  - 2.2.7 Staff quarters (for example, hostels and nurses' homes)
  - 2.2.8 Orphanages
  - 2.2.9 Other institutional places
- 2.3 Camps and workers' quarters
  - 2.3.1 Military camps
  - 2.3.2 Worker camps
  - 2.3.3 Refugee camps
  - 2.3.4 Camps for internally displaced people (IDPs)
  - 2.3.5 Other camps and worker's quarters
- 2.4 Other collective living quarters

5.489. Not all the categories in the above classification are of importance under all circumstances. For example, in some countries certain categories may not need to be considered separately, while in others it may be convenient to subdivide them. However, some of the categories are of special significance for

assessing the housing situation and should be distinguished even where a simplified classification is employed. The distinction between conventional and informal housing units is referred to particularly.

iii. Definitions of each type of living quarters

5.490. A description of the categories listed in paragraph 5.488 is given below.

1. Housing units

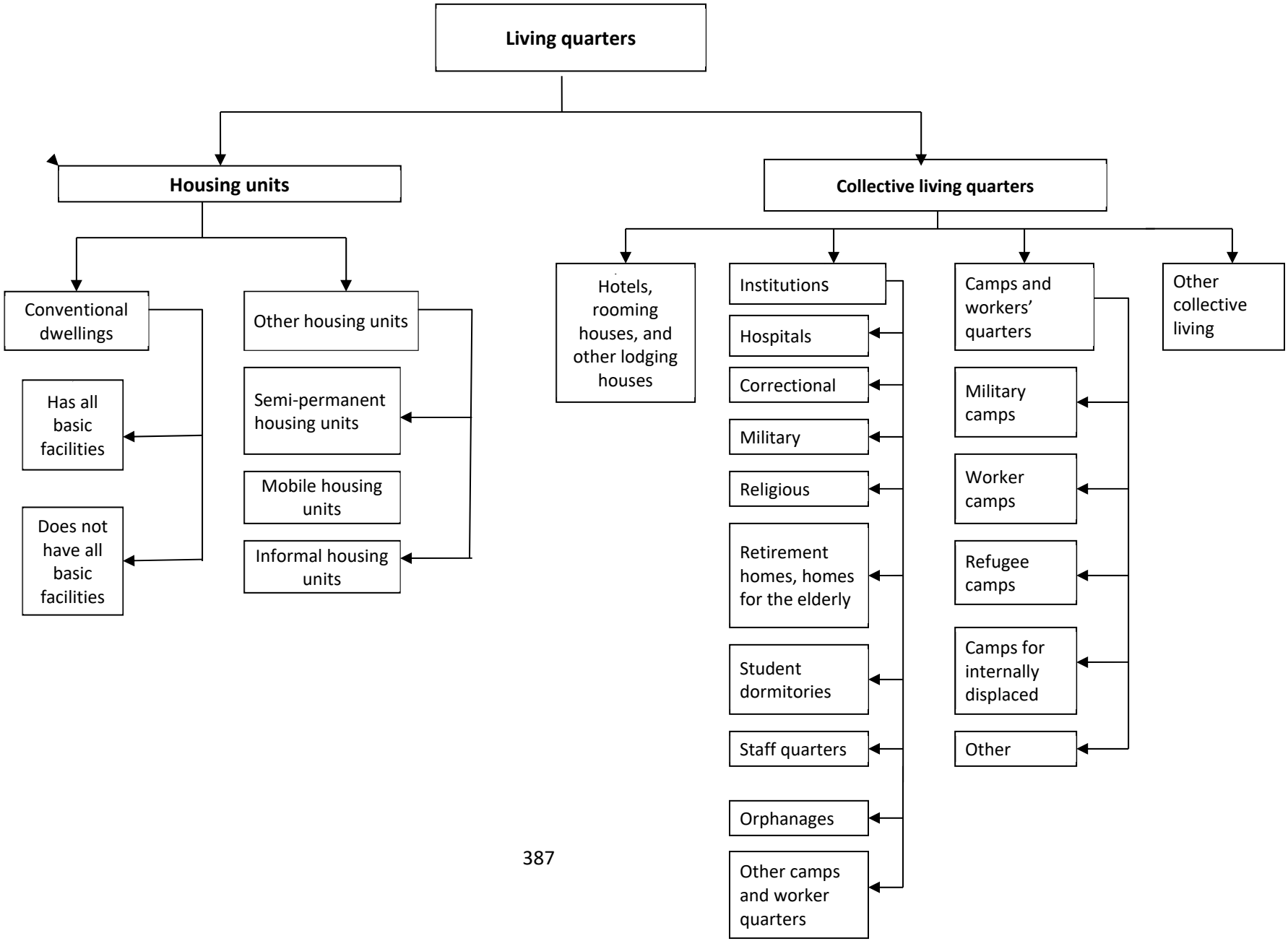
5.491. A *housing unit* is a separate and independent place of abode intended for habitation by a single household,[257] or one not intended for habitation but occupied as living quarters by a household at the time of the census. Thus, it may be an occupied or vacant dwelling, an occupied non-conventional housing unit or any other place occupied as living quarters by a household at the time of the census. This category includes housing of various levels of permanency and acceptability and therefore requires further classification in order to provide for a meaningful assessment of housing conditions.

5.492. The essential features of housing units are separateness and independence. An enclosure may be considered separate if surrounded by walls, fences, and so forth, and whether or not covered by a roof so that a person or group of persons can isolate themselves from other persons in the community for the purposes of sleeping, preparing and taking their meals, and protecting themselves from the hazards of climate and environment. Such an enclosure may be considered independent when it has direct access from the street or from a public or communal staircase, passage, gallery or grounds, in other words, when the occupants can come in and go out of their living quarters without passing through anybody else's premises.

5.493. Attached rooms having an independent entrance, or detached rooms for habitation that clearly have been built or rebuilt or converted for use as part of living quarters, should be counted as part of the housing units. Thus, housing units may comprise rooms or groups of rooms with independent entrances, or separate buildings.

[257] Although intended for habitation by one household, a housing unit may, at the time of the census, be occupied by one or more households or by a part of a household.

Figure 5. Classification of living quarters



5.494. It should be noted that any housing units located on the grounds or within the buildings of an institution, camp, and so forth should be separately identified and counted as such. For example, if, within the grounds of a hospital, there is a separate and independent house intended for the habitation of the director and his or her family, the house should be counted as a housing unit. In the same way, self-contained apartments located within hotel buildings should be counted as housing units if they have direct access to the street or to a common space within the building. Similar cases will need to be identified and described in the instructions for the enumeration.

1.1 Conventional dwellings

5.495. A "conventional dwelling" is a room or suite of rooms and its amenities in a permanent building or structurally separated part thereof which, by the way it has been built, rebuilt or converted, is intended for habitation by one household and is not, at the time of the census, used wholly for other purposes. It should have a separate access to a street (direct or via a garden or grounds) or to a common space within the building (staircase, passage, gallery and so on). Therefore, there are four essential features of a conventional dwelling:

(a) It is a room or suite of rooms;
(b) It is located in a permanent building;
(c) It has separate access to a street or to a common space;
(d) It was intended to be occupied by one household.

5.496. Examples of conventional dwellings are houses, flats, suites of rooms and apartments. Although a conventional dwelling is a housing unit originally intended, that is to say, constructed or converted, for habitation by one household, it may, at the time of the census, be vacant or occupied by one or more households. It may be noted that the terms dwelling, dwelling unit, dwelling house, residential dwelling unit, family dwelling, house, *logement*, *vivienda*, *unidad de vivienda* and so forth have been used indiscriminately to refer to housing units of any type. The referent of the term "dwelling" is here limited to a housing unit located in a permanent building and designed for occupancy by one household.

5.497. A "permanent building" is understood to be a structure that is not intended to be moved and that may be expected to maintain its stability for 15 years or more, depending on the way countries define durability. It is recognized that the criterion of permanency or durability (particularly with respect to the significance of materials and methods of construction) may be difficult to apply either in the field or from information held in administrative records and that its adaptation to local conditions would require considerable study and experimentation by the national offices. In some cases, it may be of greater significance nationally to apply the criteria of construction materials and methods of construction directly in order to establish whether or not the building containing the housing unit is of permanent construction, rather than to translate these criteria into a time period.

1.1.1. Conventional dwelling – has all basic facilities

5.498. A conventional dwelling that has all basic facilities refers to a unit that meets all the needs of the household within its confines, such as protection from elements, cooking and maintaining hygiene. Thus, in addition to the four essential features of a conventional dwelling described in paragraph 5.495, all of the following facilities must be available for a dwelling to fall in this category:

(a) Piped water within dwelling;

(b) Toilet within dwelling;
(c) Fixed bath or shower within dwelling;
(d) Kitchen or other space for cooking within dwelling.

1.1.2. Conventional dwelling – does not have all basic facilities

5.499. The conventional dwellings that fall in this category are dwellings that have the essential features of a conventional dwelling (see paragraph 5.495) and some, but not all, of the basic facilities described in paragraph 5.498.

5.500. With increased urbanization, the need for building low-cost housing units within the city limit has been increasing in many countries. This housing most frequently consists of buildings containing a number of separate rooms whose occupants may share some or all facilities (bathing, toilet or cooking facilities). Those units do not meet all the criteria of a conventional dwelling with all basic facilities available within the dwelling, especially from the point of view of maintaining health standards and privacy. For example, these units are known as *casa de palomar* in Latin America.

1.2 Other housing units

1.2.1. Semi-permanent housing unit

5.501. The term "semi-permanent housing unit" refers to a structure that, by the way it has been built, is not expected to maintain its durability for as long a period of time as a conventional dwelling, but has some of the main features and facilities of a conventional dwelling. As discussed earlier, durability needs to be specifically defined on the basis of national standards and practices. The number of these units in some countries and areas may be substantial. Semi-permanent housing is not to be confused with informal housing units (see paragraph 5.509 below).

5.502. For example, in some countries "core" or "nuclear" dwellings around which a dwelling will eventually be constructed are provided as part of the housing programmes. In others, a significant proportion of the housing inventory is composed of dwellings that are constructed of locally available raw materials and may be less durable than conventional dwellings.

5.503. Many countries with insufficient resources to meet their housing needs have attempted to alleviate the housing conditions of the population living in squatter areas by providing core or nuclear dwellings. Under these programmes, the households move their improvised shacks from the squatter area to a new location, the idea being that gradually, and generally with government assistance, the households with core or nuclear dwellings will keep adding to the nucleus until they can abandon their shacks entirely.

5.504. A core dwelling is sometimes only a sanitary unit containing bathing and toilet facilities, to which may be added, in subsequent phases, the other elements that will finally make up the completed dwelling. Such units do not fall within the definition of a conventional dwelling. However, although the household obviously continues to occupy its original shelter (which would probably be classified as an "improvised housing unit"), its housing situation is a vast improvement over that of households remaining in the squatter areas and the provision of the cores is a significant step towards the alleviation of housing shortages.

5.505. The problem is thus one of reflecting in the statistics the improvements brought about by programmes such as those described above without distorting the data that refer to fully constructed conventional dwellings. It is recommended, therefore, that core dwellings should be counted as dwellings in the census if at least one room[258] in addition to the sanitary facilities, is completed, and also that those dwellings that have not reached this stage of completion should be recorded as cores. Arrangements should be made so that the facilities available in the core can be related during data processing to the households for whose use they have been provided.

5.506. In other countries and areas, the population has developed, over time, a traditional and typical type of housing unit that does not have all the characteristics of conventional dwellings but is considered somewhat suitable from the point of view of climate and tradition. This is especially the case in many tropical and subtropical rural areas where housing units have been constructed or built with locally available raw materials such as bamboo, palm, straw or any similar materials. Such units often have mud walls, thatched roofs and so forth, and may be expected to last only for a limited time (from a few months to several years), although occasionally they may last for longer periods. This category is intended to cover housing units that are typical and traditional in many tropical rural areas. Such units may be known, for example, as cabins, *ranchos* or *bohíos* (Latin America), *barastis* (Bahrain), or *bahay kubo* (the Philippines).

1.2.2. Mobile housing units

5.507. A "mobile housing unit" is any type of living accommodation that has been produced to be transported (such as a tent) or is a moving unit (such as a ship, boat, barge, vessel, railroad car, caravan, trailer or yacht) occupied as living quarters at the time of the census. Trailers and tents used as permanent living quarters may be of special interest in some countries.

5.508. Although mobile housing units are significantly different from other housing units in that they can be readily moved or transported, mobility in itself is not necessarily an indicator of low quality. For the assessment of housing conditions in countries with a substantial number of mobile units, it may be useful to classify them further as tents, wagons, boats, trailers or other unit types.

1.2.3 Informal housing units

5.509. The term "informal housing unit" refers to a unit that does not have many of the features of a conventional dwelling and is generally characterized as unfit for human habitation, but that is used for that purpose at the time of the census. Therefore, it is neither a permanent structure nor one equipped with any of the essential facilities. Depending on national circumstances, countries should develop detailed instructions to distinguish between informal and semi-temporary housing units.

5.510. Informal housing units comprise three subgroups: "improvised housing units", "housing units in permanent buildings not intended for human habitation" and "other informal housing units". These units are characterized by the fact that they are either makeshift shelters constructed of waste materials and generally considered unfit for habitation (squatters' huts, for example) or places that are not intended for human habitation although in use for that purpose at the time of the census (barns, warehouses, natural shelters and so on). Under almost all circumstances, such places of abode represent unacceptable housing and they may be usefully grouped together in order to analyse the housing conditions of the population and to estimate housing needs. Each subgroup is defined below.

[258] For the definition of "room", see paragraph 5.546.

1.2.3.1 Improvised housing units

5.511. An improvised housing unit is an independent, makeshift shelter or structure, built of waste materials and without a predetermined plan for the purpose of habitation by one household, which is being used as living quarters at the time of the census. Included in this category are squatters' huts, *poblaciones callampas* (Chile), *hongos* (Peru), *favelas* (Brazil), *sarifas* (Iraq), *barong barong* (the Philippines) and any similar premises arranged and used as living quarters, though they may not comply with generally accepted standards for habitation, and may not have many of the characteristics of conventional dwellings. This type of housing unit is usually found in urban and suburban areas, particularly at the peripheries of the principal cities.

5.512. There is a wide variation in the procedures and criteria used in classifying these units. There are many borderline cases, and countries will need to make decisions and issue detailed instruction on how to enumerate and classify such housing units.

1.2.3.2 Housing units in permanent buildings not intended for human habitation

5.513. Included in this category are housing units (in permanent buildings) that have not been built, constructed, converted or arranged for human habitation but that are actually in use as living quarters at the time of the census. These include housing units in stables, barns, mills, garages, warehouses, offices, booths and so forth.

5.514. This category also may cover units and their occupants in buildings initially built for human habitation but later abandoned with all services cut because of deterioration. These dilapidated buildings can be found, especially in large cities, still standing, although marked for demolition. They should be included in this category if inhabited.

5.515. Premises that have been converted for human habitation, although not initially designed or constructed for this purpose, should not be included in this category, but classified as "other informal housing units".

1.2.3.3 Other informal housing units

5.516. This category refers to living quarters that are not intended for human habitation or located in permanent buildings but that are nevertheless being used as living quarters at the time of the census. Caves and other natural shelters fall within this category.

2. Collective living quarters

5.517. *Collective living quarters* include structurally separate and independent places of abode intended for habitation by large groups of individuals or several households and occupied at the time of the census. Such quarters usually have certain common facilities, such as cooking and toilet installations, baths, lounge rooms or dormitories, which are shared by the occupants. They may be further classified into hotels, rooming houses and other lodging houses, institutions and camps.

5.518. Housing units located on the grounds or within the buildings of an institution, camp, hotel and so forth should be separately identified and counted as such.

5.519. The criteria established for the identification of collective living quarters are not always easy to apply and it is sometimes difficult for an enumerator to decide whether living quarters should be classified as a housing unit or not. This is particularly true in the case of a building occupied by a number of households. Enumerators should be given clear instructions as to whether the premises occupied by a group of people living together are to be considered a housing unit or as collective living quarters. This may be less of a problem where census information is collected from administrative data sources and where such buildings are registered as being for communal living.

2.1 Hotels, rooming houses and other lodging houses

5.520. This group comprises permanent structures that provide lodging on a fee basis and in which the number of borders or lodgers exceeds five.[259] Where there are less than five, the living quarters should be classified as a housing unit. Hotels, motels, inns, boarding houses, pensions, lodging houses and similar structures fall within this category. If there is any accommodation within a hotel or similar establishment that is occupied by a household and which fulfils the requirement of a conventional dwelling it should be classified as such.

2.2 Institutions

5.521. This group covers any set of premises in a permanent structure or structures designed to house (usually large) groups of persons who are bound by either a common public objective or a common personal interest. Such sets of living quarters usually have certain common facilities shared by the occupants (for example baths, lounges and dormitories). Hospitals, military barracks, boarding schools, convents, prisons and so forth fall within this category (see the categories in paragraph 5.488).

5.522. It may be useful, depending on national needs, to require that an institution be used as the principle usual residence of at least one person at the time of the census.

2.3 Camps

5.523. Camps are sets of premises originally intended for the temporary accommodation of persons with common activities or interests. Included in this category are military camps, refugee camps and camps established for the housing of workers in mining, agriculture, public work, seasonal work~~s~~, other types of enterprises, or victims of disasters like earthquake and floods~~.~~

2.4 Other collective living quarters

5.524. This is a residual category for collective living quarters that may not conform to the definitions of those included in groups 2.1 through 2.3. It should be used only when the number of units in question is small. Where the number is substantial, additional groups of living quarters that have common characteristics and that are of significance for an improved appraisal of housing conditions should be established.

[259] The threshold of five lodgers is the one most used. However, depending on national circumstances, this number might be adjusted accordingly.

5.525. In some countries, it seems that certain types of multi-household living quarters have emerged in response to the particular needs of the population, and that the characteristics of these quarters enable them to be readily identified by an enumerator. It may be useful in these countries to provide a separate subgroup for any such special types.

5.526. It should be stressed that the types of living quarters to be included in this category are those intended for communal habitation by several households, that is to say, constructed or converted for this purpose. Housing units intended for occupancy by one household, but those at the time of the census are occupied by several households, are not to be included as collective living quarters because this obscures the identification of households doubling up in dwellings (an important element in estimating housing needs). It is suggested that, in carrying out the census, a strict distinction be maintained between a housing unit occupied by more than one household and living quarters constructed or converted for communal habitation by several households.

2. Location of living quarters (core topic)

5.527. The definitions of "locality" and "urban and rural" (paragraphs 5.94-5.105) provide valuable information for locating living quarters. Understanding these geographic concepts is crucial for housing censuses, impacting both census execution and data tabulation. When a housing census is combined with, or closely linked to, a population census, careful coordination is essential. The geographic areas used in both censuses should be defined to optimize the value of each operation. For countries conducting separate population and housing censuses, ensuring consistent recording of geographic location for dwellings and resident households is critical. However, in countries with a single combined census, collecting this information twice is unnecessary.

5.528. Information on location should be collected in sufficient detail to enable tabulations to be made for the smallest geographic subdivisions required by the tabulation plan. To satisfy the requirements of the geographic classifications recommended in the tabulations as a virtual part of this publication, information is needed on whether the living quarters are located in an urban or rural area, the major civil division, the minor civil division and, for living quarters located in principal localities, the name of the locality.

5.529. Where a permanent system of house or building numbers does not already exist, it is essential for the census to establish a numbering system so that the location of each set of living quarters can be adequately described. Similarly, in cases where streets do not have names or numbers properly displayed, such identification should be provided as one of the pre-census operations. Adequate identification provides the basis for the preparation of census control lists (see also “living quarters and household listing” in paragraphs 3.147–3.152); it is required in order to monitor and control the enumeration, and to identify living quarters for possible callbacks and post-enumeration evaluation surveys (PES) as well as for other post-censal enquiries that use the census as a sampling frame or other point of departure. Ideally, each building or other inhabited structure should be provided with a number, as should each set of living quarters within buildings or structures. In preparing a census control listing, it is the practice to identify further each household within the living quarters. In register-based censuses, it is very important that the frame used does not include under coverage or over coverage and is deduplicated as much as possible. Therefore, periodic quality controls must be carried out.

5.530. Living quarters that are not located in areas with a conventional pattern of streets, such as those in squatter areas or in some places not intended for habitation, may require special identification. Since

it may not be possible to describe the location of these units in terms of a formal address, it may be necessary to describe them in terms of their proximity to natural or created landmarks of various kinds or in relation to buildings that are located in areas where a formal address is possible. If possible, coding "slums" separately is of critical importance for countries with prominent slum populations.

5.531. The various geographic designations that together define the location of living quarters are discussed below.

i. Address

5.532. Information that describes the place where the living quarters are to be found and distinguishes them from other living quarters in the same locality falls within this category. As a rule, the information includes the name or number of the street and the number of the living quarters; in the case of apartments, the building number and the apartment number are required.

ii. Locality

5.533. For the definition of “locality”, see paragraphs 5.94–5.98.

iii. Urban and rural

5.534. For the definition of “urban and rural”, see paragraphs 5.99–5.105.

3. Occupancy status (core topic)

5.535. *Occupancy status* refers to whether or not a conventional dwelling is occupied at the time of the census. For those dwellings not occupied (because they are vacant or in secondary use), the reason for not being occupied should be classified.

5.536. Information should be obtained for each conventional dwelling to show whether the dwelling is occupied or vacant at the time of the census. For vacant units intended for year-round occupancy, the type of vacancy (for rent or for sale, for example) should be reported. Occupancy status applies only to conventional dwellings, since all other types of living quarters are required by definition to be occupied in order to fall within the scope of the census.

5.537. The enumeration of vacant conventional dwellings is likely to pose difficult problems, but at least a total count should be made for purposes of controlling the enumeration. The type of vacancy is frequently indicated by “for sale” or “for rent” signs posted on the dwelling. Although it may not be feasible to investigate all of the topics included in the census for vacant units, as much information as possible should be collected, including information on whether the living quarters are vacant seasonally or non-seasonally.

5.538. Vacant units intended for seasonal or secondary occupancy may represent a substantial proportion of the housing stock in resort areas and in areas where large numbers of seasonal workers are employed. The separate identification of such categories may be necessary for the correct interpretation of the overall vacancy rate, as well as for an evaluation of the housing situation in the area concerned.

5.539. Whether living quarters whose occupants are temporarily absent or temporarily present are to be recorded as occupied or vacant will need to be considered in relation to whether a de jure or de facto population census is being carried out. In either case, it would seem useful to distinguish as far as possible conventional dwellings that are used as a second residence. This is particularly important if the second residence has markedly different characteristics from the primary residence, as is the case, for example, when agricultural households move during certain seasons of the year from their permanent living quarters in a village to rudimentary structures located on agricultural holdings. The recommended classification of occupancy status for conventional dwellings is as follows:

1 Occupied
2 Vacant / not occupied
- 2.1 Seasonally vacant
  - 2.1.1 Holiday homes
  - 2.1.2 Seasonal workers' quarters
  - 2.1.3 Other
- 2.2 Secondary residences
- 2.3 For rent/sale
- 2.4 For demolition
- 2.5 Other

4. Ownership – type of (core topic)

5.540. This topic refers to the type of ownership of the housing unit itself and not of that of the land on which it stands. Type of ownership should not be confused with tenure, which is a characteristic of the household and is covered in paragraphs 5.619–5.622.

5.541. Information should be obtained to show:

(a) Whether the housing unit is owned by the public sector (central government, local government, public corporations);

(b) Whether the housing unit is privately owned (by households, private corporations, cooperatives, housing associations and so on). The question is sometimes expanded to show whether the housing units are fully paid for, being purchased in instalments or mortgaged. The classification of housing units by type of ownership is as follows:

1 Owner-occupied
2 Non-owner-occupied
- 2.1 Publicly owned
- 2.2 Privately owned
- 2.3 Communally owned
- 2.4 Cooperatively owned
- 2.5 Other

5.542. Housing units are defined as owner-occupied if used wholly or partly for own occupation by the owner. In principle, if a housing unit is being purchased in instalments or mortgaged according to national legal systems and practices, it should be enumerated as being owned. Instructions should also cover other arrangements, such as housing units in cooperatives or housing associations.

5.543. The information on ownership may be classified, as a minimum, into two main groups, namely "private ownership" and "other ownership". Depending upon the prevalence of various types of ownership and their significance with respect to housing conditions and the formulation of housing programmes, it may be useful to dissect the category "other ownership" into the relevant examples of the subgroups shown. The categories used should be consistent with those employed in the system of national accounts of the country concerned and in accordance with the recommendations contained in the *System of National Accounts, 2008*.[260]

5.544. It has been observed that the collection of information on type of ownership in a general census may be hampered by the fact that the occupants might not know who the owner of the property is and that the owners or their representatives may be situated outside the enumeration zone. Furthermore, there are numerous cases of borderline and mixed ownership, which make the topic difficult for nationwide enumeration. This is one of the topics for which more accurate information might be obtained through a housing survey.

5.545. In countries where there is a substantial amount of employer-issued housing, it would be useful to include the subcategories "issued by the employer" and "not issued by the employer" under the category "privately owned" (or publicly owned where the employer is a public sector entity). It is important that such information be known from the point of view of assessing the impact of job loss, in order to gauge the magnitude of the population whose loss of a job would include loss of housing as well.

5. Rooms – number of (core topic)

5.546. A *room* is defined as a space in a housing unit enclosed by walls reaching from the floor to the ceiling or roof covering, or to a height of at least 2 metres, of an area large enough to hold a bed for an adult, that is, at least 4 square metres. Usually only information on rooms in housing units is collected in a census. The total number of types of rooms therefore includes bedrooms, dining rooms, living rooms, studies, habitable attics, domestic workers' rooms, kitchens, rooms used for professional or business purposes, and other separate spaces used or intended for dwelling purposes, so long as they meet the criteria concerning walls and floor space. Passageways, verandas, lobbies, bathrooms and toilet rooms should not be counted as rooms, even if they meet the criteria. Separate information may be collected for national purposes on spaces of less than 4 square metres that conform in other respects to the definition of "room" if it is considered that their number warrants such a procedure.

5.547. Rooms used exclusively for business or professional purposes should be counted separately, as it is desirable to include them when calculating the number of rooms in a dwelling but to exclude them when calculating the number of persons per room. This procedure allows density levels to be studied according to the number of rooms available for living purposes in relation to the number of occupants. In any event, each country should indicate the procedure that has been followed.

5.548. It is recommended that kitchens be included in the count of rooms provided they meet the criteria concerning walls and floor space. Kitchens or kitchenettes that have an area smaller than 4 square metres or that have other characteristics that disqualify them should be excluded. For national purposes,

[260] *System of National Accounts 2008*, United Nations publication, Sales No. E.08.XVII.29.

countries may wish to identify and count kitchens within a separate group that may be analysed with respect to size and utilization, and to consider separately those used exclusively for cooking.

6. Bedrooms – number of

5.549. In addition to enumerating the number of rooms, some countries may wish to collect information on the number of bedrooms in a housing unit in order to provide a measure of overcrowding.

5.550. A bedroom is defined as a room equipped with a bed and used for night rest and excludes makeshift and temporary sleeping quarters. Rooms where sleeping areas are partitioned with curtains or blinds or any similar makeshift arrangements cannot be counted as a bedroom. The count of bedrooms includes spare/guest bedrooms not occupied at the time of the census. As mentioned above (paragraph 5.546), there must be some permanency in relation to the walls enclosing the bedrooms. A room used for other activities by day and sleeping by night is not a bedroom except for one-room dwelling units, in which case, it must be recorded as having one room and one bedroom.

7. Useful floor space

5.551. This topic refers to the useful and liveable floor space in housing units, that is to say, the floor space measured inside the outer walls of housing units, excluding non-habitable cellars and attics. Information on this topic is often collected to supplement that on the number of rooms. In some countries, however, such information is collected in preference. In multiple-dwelling buildings, all common spaces should be excluded.

5.552. Collecting information on the floor space available to occupants of housing units may prove to be difficult; occupants often may not know the exact or even the approximate area of the housing unit they occupy, and training enumerators to calculate the floor space would be complicated and costly, and would result in inaccuracies. In this context, and taking into account the importance of the information concerned, countries should take into consideration developing detailed instructions on proper procedures for assessing these data (for example, a request for information on floor space from the official documents available to the occupants, such as the rental agreement and the title, which are supposed to include such information).

8. Water supply system (core topic)

5.553. Basic information to be obtained in the census is whether housing units have or do not have a piped water installation, in other words, whether or not water is provided to the housing unit by pipes from a community-wide system or a private installation, such as a pressure tank or pump. The unit of enumeration for this topic is a housing unit. It is also recommended that countries indicate whether the unit has tap water inside or, if not, whether it is within a certain distance from the door. The recommended distance is 200 metres, assuming that access to piped water within that distance allows occupants of the housing unit to obtain water for household needs without being subjected to extreme efforts. Besides the location of the tap water relative to the housing unit, the source of water available to households is also of interest. Therefore, the recommended classification of housing unit by water supply system is as follows:

1 Piped water inside the housing unit
  1.1 From the community scheme

1.2 From an individual source
2 Piped water outside the unit but within 200 metres
2.1 From the community scheme
2.1.1 For exclusive use
2.1.2 Shared
2.2 From an individual source
2.2.1 For exclusive use
2.2.2 Shared
3 Other (see category 3 of the classification in paragraph 5.558 for more details)
4 No piped water

5.554. A community scheme is one that is subject to inspection and control by public authorities. Such schemes are generally operated by a public body but, in some cases, they are operated by a cooperative or private enterprise. An individual source of water refers to a source of water that is not part of a community scheme, such as an individual or shared water reservoir.

5.555. As noted above the unit of enumeration for this topic is the housing unit. However, some countries may find it useful to collect information on the availability of piped water for the use of occupants in collective living quarters. Such living quarters are usually equipped with multi-facilities for the use of large groups, and information on the water supply system in relation to the number of occupants would be significant in respect of analysing housing conditions. The water supply system in collective living quarters constitutes an additional topic.

5.556. The most significant information from a health point of view is whether the living quarters have piped water within the premises. However, a category may be added to distinguish cases where the piped water supply is not within the living quarters but rather within the building in which the living quarters are situated. It may also be useful to collect information that would show whether the water supply is for the sole use of the occupants of the living quarters being enumerated or whether it is for the use of the occupants of several sets of living quarters, as indicated in the above classification at the three-digit level. Where there is a large proportion of housing units with no piped water, this category may be expanded to specify sources commonly used in a country.

9. Drinking water – main source of (core topic)

5.557. Having enough water for drinking and personal hygiene is essential, but quantity by itself is not sufficient. The quality of the water is also a crucial health issue. Consequently, one of the targets of the water supply, sanitation and hygiene (WASH) post-2015 recommendations[261] proposed by the WHO/UNICEF Joint Monitoring Programme on Water Supply and Sanitation (JMP) is "universal access to basic drinking water, sanitation and hygiene", assessed in part by having access at home to safely managed drinking water. A safely managed drinking water service is defined as one that reliably delivers water that is sufficient to meet domestic needs and does not represent a significant risk to health. This implies a system that delivers water to the household and includes measures to prevent risks and to verify water quality through compliance monitoring. An improved water source (piped water, public tap or standpost, tubewell or borehole, protected dug well, protected spring, rainwater in collection tank) can

[261] See WHO/UNICEF Joint Monitoring Programme on Water Supply and Sanitation (JMP) report entitled *WASH post-2015: proposed targets and indicators for drinking-water, sanitation and hygiene* (April 2014), accessible at http://www.wssinfo.org/.

be safely managed. Unimproved sources, which by definition are not safely managed, include unprotected dug well, unprotected spring and surface water from a river, stream, dam, lake, pond, canal or irrigation channel. Delivered water (for example, through trucks, carts, sachets or bottles) can potentially be safely managed, but if these are the primary drinking water sources, other improved sources of water must be accessible at the household for other domestic uses (for example, washing, bathing).

5.558. Countries are encouraged to collect the information on the main source of drinking water for the household, particularly where there is considerable difference between sources of water for general household use and for drinking. For those countries wishing to collect this information, the following categories of main source of drinking water are recommended:

1 Piped water inside the unit
- 1.1 From the community scheme
- 1.2 From an individual source

2 Piped water outside the unit but within 200 metres
- 2.1 From the community scheme
  - 2.1.1 For exclusive use
  - 2.1.2 Shared
- 2.2 From an individual source
  - 2.2.1 For exclusive use
  - 2.2.2 Shared

3 Other sources of water
- 3.1 Borehole/tubewell
- 3.2 Protected dug well
- 3.3 Protected spring
- 3.4 Rainwater collection tank
- 3.5 Delivered water – bottled, sachet
- 3.6 Delivered water – tanker trucks, carts
- 3.7 Unprotected dug well/spring/river/stream/lake/pond/dam/canal/irrigation channel

10. Toilet – type of (core topic)[262]

5.559. A *toilet* may be defined as an installation for the disposal of human excreta. A flush toilet is an installation provided with piped water that permits humans to discharge their wastes and from which the wastes are flushed by water. The unit of enumeration for this topic is a housing unit.

5.560. For housing units reported as having a toilet, additional information may be sought to determine whether the toilet is used exclusively by the occupants of the living quarters being enumerated or is shared with the occupants of other living quarters. For living quarters reported as having no toilet, it would be useful to know if the occupants have the use of a communal facility and the type of facility, or if they have the use of the toilet of other living quarters and the type, or if there is no toilet of any kind available for the use of the occupants.

[262] It is also necessary to distinguish between conventional dwellings with all main facilities and other conventional dwellings.

5.561. Some countries have found it useful to expand the classification for non-flush toilets so as to distinguish certain types that are widely used and indicate a certain level of sanitation. The recommended classification of housing unit by toilet facilities is as follows:

1 With toilet within housing unit
- 1.1 Flush/pour flush[263] toilet
- 1.2 Other

2 With toilet outside housing unit
- 2.1 For exclusive use
  - 2.1.1 Flush/pour flush toilet
  - 2.1.2 Ventilated improved pit latrine[264]
  - 2.1.3 Pit latrine without ventilation with covering
  - 2.1.4 Holes or dug pits with temporary coverings or without shelter
  - 2.1.5 Other
- 2.2 Shared
  - 2.2.1 Flush/pour flush toilet
  - 2.2.2 Ventilated improved pit latrine
  - 2.2.3 Pit latrine without ventilation with covering
  - 2.2.4 Holes or dug pits with temporary coverings or without shelter
  - 2.2.5 Other

3 No toilet available
- 3.1 Service or bucket facility (excreta manually removed)
- 3.2 Use of natural environment, for example, bush, river, stream
- 3.3 Other

5.562. As noted above the unit of enumeration for this topic is the housing unit. However, some countries may find it useful to collect information on the availability of toilet facilities for the use of occupants in collective living quarters. Living quarters of this type are usually equipped with multi-facilities for the use of large groups, and information on the number and type of toilets in relation to the number of occupants would be significant in terms of analysing housing conditions. The availability of toilets for collective living quarters represents an additional topic.

11. Sewage disposal (core topic)

5.563. Information on toilets should be combined with the sewage disposal system to which they are connected in order to determine the adequacy of sanitation facilities of the housing unit. To be considered adequate sanitation, toilets or latrines have to be connected to non-clogged sewage disposal systems. The information on housing units by type of sewage disposal system may be classified as follows:

1 Empties into a piped system connected to a public sewage disposal plant

[263] A pour flush toilet uses a water seal, but unlike a flush toilet, a pour flush toilet uses water poured by hand for flushing (no cistern is used).

[264] A ventilated improved pit latrine (VIP) is a dry pit latrine that uses a hole in the ground to collect the excreta and a squatting slab or platform that is firmly supported on all sides, easy to clean and raised above the surrounding ground level to prevent surface water from entering the pit. The platform has a squatting hole, or is fitted with a seat.

2 Empties into a piped system connected to an individual sewage disposal system (septic tank, cesspool)
3 Other – such as toilet empties into an open ditch, a pit, a river, the sea
4 No disposal system

12. Solid waste disposal – main type of (core topic)

5.564. Securing sustainable development and, in this context, the usual manner of treatment of solid waste (garbage) generated by the household, has prompted the incorporation of this topic in a number of national housing censuses.

5.565. This topic refers to the usual manner of collection and disposal of solid waste or garbage generated by occupants of the housing unit, including e-waste[265] disposal. The unit of enumeration is a housing unit. The classification of housing units by type of solid waste disposal is according to the following guidelines:

1 Solid waste collected on a regular basis by authorized collectors
2 Solid waste collected on an irregular basis by authorized collectors
3 Solid waste collected by self-appointed collectors
4 Occupants dispose of solid waste in a local dump supervised by authorities
5 Occupants dispose of solid waste in a local dump not supervised by authorities
6 Occupants burn solid waste
7 Occupants bury solid waste
8 Occupants dispose solid waste into river, sea, creek, pond
9 Occupants compost solid waste
10 Other arrangement

13. Bathing facilities (core topic)

5.566. Information should be obtained on whether or not there is a fixed bath or shower installation within the premises of each set of housing units. The unit of enumeration for this topic is a housing unit. Additional information may be collected to show whether or not the facilities are for the exclusive use of the occupants of the living quarters and where there is a supply of hot water for bathing purposes or cold water only. In some areas of the world the distinction proposed above may not be the most appropriate for national needs. It may be important, for example, to distinguish between whether, in terms of availability, there is a separate room for bathing in the living quarters, a separate room for bathing in the building, an open cubicle for bathing in the building and a public bathhouse. The recommended classification of housing units by availability and type of bathing facilities is as follows:

1 With fixed bath or shower within housing unit
2 Without fixed bath or shower within housing unit
2.1 Fixed bath or shower available outside housing unit

[265] E-waste has become an important environmental and public health issue, not only at global scale but also in low- and middle-income countries as well. Many of Electrical and Electronic Equipments (EEE) contain hazardous chemicals and materials. For this reason, inadequacies in recycling and disposing of electronic waste (e-waste) can cause serious health and environmental pollution problems. Great care must be taken to prevent unsafe exposure in recycling operations and prevent contamination of the environment.

2.1.1 For exclusive use
2.1.2 Shared
2.2 No fixed bath or shower available

5.567. Alternatively, and in line with the elaboration in the preceding paragraph, the following classification may be more appropriate in certain circumstances:

1 Separate room for bath or shower within the housing unit
2 No separate room for bath or shower but bathing space available within the housing unit (for example, in an open area around the well within the housing unit, in the courtyard)
3 Bathing room available but outside the housing unit for exclusive use
4 Shared bathing room outside the housing unit
5 No specific bathing room available

5.568. As noted above the unit of enumeration for this topic is the housing unit. However, as with the type of toilet, some countries may find it useful to collect information, as an additional enquiry, on the availability of a bath or shower for the use of occupants in collective living quarters as well. Living quarters of this type are usually equipped with multi-facilities for the use of large groups, and information on the number of fixed baths or showers in relation to the number of occupants would be significant in terms of analysing housing conditions. The number of fixed baths or showers in collective living quarters would represent an additional enquiry.

14. Kitchen – availability of (core topic)

5.569. Information should be obtained on whether the housing unit has a kitchen, whether some other space is set aside for cooking, such as a kitchenette, or whether there is no special place set aside for cooking. The unit of enumeration for this topic is a housing unit.

5.570. A *kitchen* is defined as a space that conforms in all respects to the criteria for a room, and is equipped for the preparation of the principal meals of the day and intended primarily for that purpose.

5.571. Any other space reserved for cooking, such as a kitchenette, will fall short in respect of possessing the attributes of a room, even though it may be equipped for the preparation of the principal meals of the day and is intended primarily for that purpose. The collection of data on the availability of a kitchen may provide a convenient opportunity to collect information on the kind of equipment that is used for cooking, for example, a stove, hotplate or open fire, and on the availability of a kitchen sink and a space for food storage so as to prevent spoilage. The recommended classification of housing units by availability of a kitchen or other space reserved for cooking within the housing unit is as follows:

1 With kitchen within housing unit
1.1 For exclusive use
1.2 Shared
2 With other space for cooking within housing unit, such as kitchenette
2.1 For exclusive use
2.2 Shared
3 Without kitchen or other space for cooking within housing unit
3.1 Kitchen or other space for cooking available outside housing unit
3.1.1 For exclusive use

3.1.2 Shared
3.2 No kitchen or other space for cooking available

5.572. As noted above the unit of enumeration for this topic is the housing unit. However, some countries may find it useful to collect information on the availability of kitchen facilities for the use of occupants in collective living quarters. Living quarters of this type are usually equipped with multi-facilities for the use of large groups, and information on the number of kitchens or kitchenettes in relation to the number of occupants would be significant in terms of analysing housing conditions. It represents an additional topic.

15. Fuel used for cooking (core topic)

5.573. The proportion of households using solid fuels is one of the indicators used in monitoring internationally agreed development goals. There are important linkages between household solid fuel use, indoor air pollution, deforestation and soil erosion and greenhouse gas emissions. The type of fuel and participation in cooking tasks are important predictors of exposure to indoor air pollution. It is thus recommended to collect information on the fuel used for cooking by each housing unit. Fuel used for cooking refers to the fuel used predominantly for preparation of principal meals. If two fuels (for example, electricity and gas) are used, the one used most often should be enumerated. The classification of fuels used for cooking depends on national circumstances and may include electricity, gas, oil, coal, firewood and animal dung. It would also be useful to collect this information for collective living quarters, especially if the number of sets of collective living quarters in the country is significant. The classification of fuel used for cooking is as follows:

1 Gas
2 Electricity
3 Liquefied petroleum gas (LPG)
4 Kerosene/paraffin (petroleum-based)
5 Oil (including vegetable oils used as fuel)
6 Coal
7 Firewood
8 Charcoal
9 Animal dung
10 Crop residues (for example, cereal straw from maize, wheat, paddy rice, rice hulls, coconut husks, groundnut shells)
11 Other

16. Energy used for lighting – type of (core topic)

5.574. Information should be collected on the main type of energy used for lighting in the housing unit, such as electricity, gas or oil lamp. If the source of energy for lighting is electricity, some countries may wish to collect information showing whether the electricity mainly comes from a community supply, private generating plant or some other source (industrial plant, mine and so on). In addition to the type of lighting, countries may assess the information on the availability of electricity for purposes other than lighting (such as cooking, heating water and heating the premises). If housing conditions in the country allow this information to be derived from the type of lighting, there would be no need for additional enquiry.

5.575. As noted above the unit of enumeration for this topic is the housing unit. However, some countries may find it useful to collect information on the availability of electricity for the use of occupants in collective living quarters. Such living quarters are usually equipped with multi-facilities for the use of large groups, and information on electricity would be significant in terms of analysing housing conditions. The availability of electricity in collective living quarters is defined as an additional topic. No classification is specifically recommended.

17. Heating – type and energy used

5.576. This topic refers to the type of heating of housing units and the energy used for that purpose. The units of enumeration are all housing units. This topic may be less relevant for a number of countries where, owing to their geographic position and climate, there is no need to provide energy for heating.

5.577. *Type of heating* refers to the kind of system used to provide heating for most of the space: it may be central heating serving all the sets of living quarters or serving a set of living quarters, or it may not be central, in which case the heating will be provided separately within the living quarters by a stove, fireplace or some other heating body. As for the energy used for heating, it is closely related to the type of heating and refers to the predominant source of energy, such as solid fuels (coal, lignite and products of coal and lignite, wood), oils, gaseous fuels (natural or liquefied gas), or electricity. No classification is specifically recommended.

18. Hot water – availability of

5.578. This topic refers to the availability of hot water in housing units. Hot water denotes water heated to a certain temperature and conducted through pipes and tap to occupants. The information collected may indicate whether there is hot water available within the housing units, or outside the living quarters for exclusive or shared use, or not at all. No classification is specifically recommended.

19. Piped gas – availability of

5.579. This topic refers to whether piped gas is available in the housing unit or not. Piped gas is usually defined as natural or manufactured gas that is distributed by pipeline and whose consumption is recorded. This topic may be irrelevant for a number of countries where there is either a lack of sources of natural gas or no developed pipeline system. No classification is specifically recommended.

20. Use of the housing unit

5.580. *Use of the housing unit* refers to whether the housing unit is being used wholly for habitation (residential) purposes or not. The housing unit can be used for habitation and for commercial, manufacturing or some other purposes. In a number of countries, houses are used simultaneously for more than one purpose. For example, the lower floor is used as a store or workshop, and the upper floors for habitation.

5.581. The recommended classification of the use of the housing unit is as follows:

1 Used solely for habitation
2 Used for habitation and other purposes
- 2.1 Economic or business purposes
- 2.2 Other purposes

21. Occupancy by one or more households (core topic)

5.582. For the purpose of a housing census, each household must be identified separately. With respect to housing programmes, the use of the separate concepts of "household" and "living quarters" in carrying out housing censuses permits the identification of the persons or groups of persons in need of their own dwellings. If the household is defined as a group of persons occupying a set of living quarters, the number of households in the living quarters and the number of sets of occupied living quarters will always be equal and there will be no apparent housing need as reflected by the number of "sharing" households that require their own living quarters. If living quarters are defined as the space occupied by a household, the number of households in living quarters will again be equal to the number of sets of living quarters, with the added disadvantage that there will be no record of the number of structurally separate living quarters.

5.583. Occupancy by more than one household is a useful topic for assessing the current housing situation and measuring the need for additional housing. For countries relying on the housekeeping concept (see paragraph 2.36), the number of households occupying a housing unit is needed to understand the extent of shared housing. For countries relying on the dwelling unit concept of households (see paragraph 2.37), the household is equivalent to the dwelling unit.

5.584. In countries where it is traditional to count families, the family in the broad sense of the term may be adopted as an additional unit of enumeration; in the great majority of cases the composition of this unit will coincide with that of the household.

5.585. A household and family should be defined in the same way for housing census purposes as for population censuses (see paragraphs 5.134–5.141 and 5.154–5.159).

5.586. For the definitions of "household", "reference person of household" and "persons living in institutions", see paragraphs 5.134–5.162 and 2.41–2.42 in the current revision of the *Principles and Recommendations for Population and Housing Censuses*.

22. Occupants – number of (core topic)

5.587. Each person usually resident in a housing unit or in collective living quarters should be counted as an occupant. Therefore, the units of enumeration for this topic are living quarters. However, since housing censuses are usually carried out simultaneously with population censuses, the applicability of this definition depends upon whether the information collected and recorded for each person in the population census indicates where he or she was on the day of the census or whether it refers to the usual residence (see paragraphs 5.52–5.63). Care should be exercised in distinguishing persons occupying mobile units, such as boats, caravans and trailers, as living quarters from persons using these units merely as a means of transportation.

5.588. Depending on their national requirements for information, some countries may wish to distinguish between those occupants that are usually resident and those that are not usually resident in the living quarters for the purposes of better understanding the housing conditions and living arrangements of non-residents.

23. Building – type of (core topic)

a. Definition of building

5.589. A *building* is any independent free-standing structure comprising one or more rooms[266] or other spaces, covered by a roof and usually enclosed within external walls or dividing walls[267] that extend from the foundations to the roof. However, in tropical areas, a building may consist of a roof with supports only, that is to say, without constructed walls; in some cases, a roofless structure consisting of a space enclosed by walls may be considered a "building" (see also "compound" in paragraph 5.597).

5.590. In defining a "building", particular care should be given to differentiating this from "type of living quarters" (see paragraph 5.485). Type of living quarters refers to structures that are designed for residential habitation or are being used for residential habitation. A building could be a number of living quarters, a commercial premises not meant, or being used, for habitation, or a mix of the two.

5.591. A building may be used or intended for residential, commercial or industrial purposes or for the provision of services. It may therefore be a factory, shop, detached dwelling, apartment building, warehouse, garage, barn and so forth. In some exceptional cases, facilities usually provided by a set of living quarters are located in two or more separate detached structures, as when a kitchen is in a separate structure. In the case of living quarters with detached rooms, these rooms should be considered separate buildings. A building may therefore contain several sets of living quarters, as is the case for an apartment building or duplex; it may be coextensive with single detached living quarters; or it may be only part of the living quarters, as is the case, for example, for living quarters with detached rooms that are clearly intended to be used as part of the living quarters.

5.592. The concept of a building should be clearly defined and, in a census with a field enumeration, the instructions should indicate whether all buildings are to be listed and enumerated or only those used in whole or in part for residential purposes. Instructions should also indicate whether buildings under construction are to be recorded and, if so, at what stage of completion they are to be considered eligible for inclusion. Buildings being demolished or awaiting demolition should normally be excluded.

b. Classification of buildings by type

5.593. The following classification of buildings (or of living quarters) by type of building is recommended:

1.0 Residential buildings
- 1.1 Buildings containing a single housing unit
  - 1.1.1 Detached
  - 1.1.2 Attached
- 1.2 Buildings containing more than one housing unit
  - 2.2.1 Up to 2 floors
  - 2.2.2 From 3 to 4 floors
  - 2.2.3 From 5 to 10 floors
  - 2.2.4 11 floors or more
- 1.3 Buildings for persons living in institutions
- 1.4 Other residential buildings

[266] For the definition of "room", see paragraph 5.546.

[267] The term "dividing walls" refers to the walls of adjoining buildings that have been so constructed as to be contiguous, for example, the dividing walls of "row" houses.

2.0 Non-residential buildings

5.594. It should be noted that, for the purpose of the housing census, the above classification refers to the building in which the sets of enumerated living quarters are located and that usually it will be the living quarters, not buildings, that will be tabulated according to the classification.

5.595. Category 1.1 provides separate subgroupings for "detached" and "attached" buildings because, although most single-unit buildings (suburban homes, villas, and so forth) are detached, in some countries a substantial number may be attached (row or terraced houses, for example) and in such cases it may be useful to identify these separately. According to the definition of "building" in paragraph 5.589 above, a group of, for example, three row or terraced houses that are attached is considered to be three separate buildings if their "external walls or dividing walls" extend from "the foundations to the roof". Buildings containing more than one housing unit (category 1.2) will usually be apartment buildings, but they may also be other types of buildings, for example, buildings that are structurally subdivided so as to contain more than one housing unit. Buildings under the latter category should be subdivided into the following: up to 2 floors, from 3 to 10 floors and 11 floors or more. Category 1.3, "buildings for persons living in institutions", includes hospital buildings, prisons, military establishments, and so on. On the other hand, a structurally separate housing unit (a house or apartment intended for the occupancy of staff of the institution) or one that is either within a building of the institution or detached but within the grounds, belongs in category 1.0; if the housing unit is coextensive with a building, it belongs in category 1.2.

5.596. In addition to the above, and for subsequent analysis of housing conditions, each country will find it useful to provide for separate identification of the special types of buildings that are characteristic of the country concerned. These can be classified as category 1.4.

c. Compound

5.597. In some countries, it may be appropriate to use the "compound" as a unit of enumeration. In some areas of the world, housing units are traditionally located within compounds, and the grouping of sets of housing units in this way has economic and social implications that need to be studied. A compound, in these circumstances, becomes a distinct unit of enumeration, on a par with a housing unit. For purposes of international comparability, a compound should be classified according to the main features and facilities it displays and classified with housing units.

24. Year or period of construction

5.598. This topic refers to the age of the building in which the living quarters are located. It is recommended that the exact year of construction be sought for buildings constructed during the decennial period immediately preceding the census. Exact year of construction refers to the year the building was considered complete and not the year construction started). For buildings constructed before that time, the information should be collected in terms of periods that will provide a useful means of assessing the age of the housing stock. Difficulty may be experienced in collecting data on this topic in a field enumeration because in some cases the occupants may not know the date of construction, particularly for older buildings. However, more accurate information is more likely to be available where countries use housing registers or other administrative data sources for the census.

5.599. The collection of data for single years during the most recent intercensal period is seen as a method for comparing census information with construction statistics to assess under-coverage of the construction statistics or of the census.

5.600. Instead of collecting single years of construction, if this is seen to be too burdensome on the respondent, periods of construction should be collected. The periods could be defined in terms of events that have some special significance in the country concerned, particularly with regard to the effect on the condition of the housing stock; examples would be the period since the Second World War; the period between the First World War and the Second World War; and the period before a major earthquake, flood or fire. Alternatively, the response ranges could be equal to intervals from one census to the next, such as ten- or five-year periods depending on the frequency of census collection. This allows for comparisons across the same periods and across censuses. Shorter periods of construction are most important in the first few decades of a dwelling when the dwelling is undergoing changes, such as foundation setting, or when defects in dwelling systems, such as electrical or plumbing faults, reveal themselves. Afterwards the reference periods could be lengthened but should be as homogeneous as possible to allow for cohort analysis. The period selected will depend upon the materials and methods of construction used in the country concerned and the number of years that buildings normally last.

5.601. Where parts of buildings have been constructed at different times (such as the construction of extensions or annexes), the year or period of construction should refer to the major part. Where living quarters comprise more than one building (living quarters with detached rooms, for example), the age of the building that contains the major part of the living quarters should be recorded.

5.602. In countries where a significant number of households construct their own living quarters (countries with large non-monetary sectors, for example), it may be useful to collect additional information that will distinguish the living quarters according to whether or not they were constructed by the households occupying them. The information should refer only to living quarters constructed during the preceding intercensal or 10-year period, and it should be made clear in formulating the question that it refers to living quarters constructed mainly by households (with or without the help of other households in the community) and not to construction executed by enterprises on behalf of households.

25. Dwellings in the building – number of

5.603. This topic refers to the number of conventional dwellings in the building. This topic is applicable in cases where there is a possibility to have unique identifier for the building itself. If a census established such an identifier (building number, for example, linked to the address) then it would be possible to introduce this topic. This information is best collected during the listing exercise or by the enumerator recording this information from his/her own observations.

26. Position of dwelling in the building

5.604. Some countries may want to collect information on the position of the dwelling or housing unit in the building. This information can be used as an indicator of accessibility to dwellings, possibly in conjunction with information on the accessibility to the dwellings.

5.605. The following classification of dwellings by position in the building is recommended:

1.0 Dwelling on one floor only

1.1 Dwelling below the ground floor
1.2 Dwelling on the ground floor of the building
1.3 Dwelling on the 1st or 2nd floor of the building
1.4 Dwelling on the 3rd or 4th floor of the building
1.5 Dwelling on the 5th floor of the building or higher
2.0 Dwellings on two or more floors
2.1 Dwelling on the ground floor of the building or below ground level
2.2 Dwelling on the 1st or 2nd floor of the building
2.3 Dwelling on the 3rd or 4th floor of the building
2.4 Dwelling on the 5th floor of the building or higher

5.606. For dwellings on two or more floors, information should be provided with reference to the lowest floor level of the dwelling.

27. Accessibility to dwelling

5.607. The following classification of accessibility to the front door of the dwelling or housing unit is recommended, based on the presence or absence of ramps, steps and lifts:

1 Access with no steps or ramp
2 Access by ramp
3 Access by disabled stair lift
4 Access using lift only (though the building may have staircases as well)
5 Access by using only steps
6 Access only by using both lift and steps

Note that these categories are not necessarily mutually exclusive.

28. Construction material of outer walls (core topic)

5.608. This topic refers to the construction material of external (outer) walls of the building in which the living quarters are located. If the walls are constructed of more than one type of material, the predominant type of material should be reported. The types of materials distinguished will depend upon the materials most frequently used in the country concerned and on their significance from the point of view of permanency of construction or assessment of durability. The following classification of construction materials is recommended:

2.2.5 Burnt clay (bricks, blocks, panels), stone, concrete
2.2.6 Unburnt clay (bricks, blocks) mud, earth
2.2.7 Wood
2.2.8 Bamboo, trees, grass
2.2.9 Corrugated sheets
2.2.10 Prefabricated units
2.2.11 Other materials

5.609. In some countries, the material used for the construction of roofs or of floors may be of special significance for the assessment of durability and, in such cases, it may be necessary to collect information on this as well as on the material of the walls. Durability refers to the period of time for which the structure

remains habitable, subject to regular maintenance. A durable structure is one expected to remain sound for a considerable period of time. Countries may wish to define the length of the period, for example, 15 or 20 years. Durability does not depend solely on the materials used in construction, since it is also affected by the way the building was erected, that is to say, whether it was built according to construction standards and regulations. Technological developments in treating traditional building materials, such as bamboo, have extended the durability of those materials for several decades. Construction material of outer walls may be considered an indicator of the building's durability. Therefore, in order to assess quality of the national housing stock, durability may be measured in terms of material used together with adherence to construction standards. Specific instructions for enumerators at the national level should be developed on the basis of national building construction practice.

5.610. While the material of construction is a useful addition to data collected on the type of living quarters, it should not be considered a substitute for the latter topic. Wood, for example, may be the material of both a poorly constructed squatter's hut and a durable and well-constructed dwelling. In these cases, information on the type of living quarters adds significantly to the value of the census in assessing the quality of a country's housing stock.

29. Construction material of floor and roof

5.611. In some cases the material used for the construction of roofs and floors may be of special interest and can be used to further assess the quality of dwellings. This topic refers to the material used for roof and floor (although, depending on the specific needs of a country, it may also refer to other parts of the building in which the housing unit is located, such as the frame or the foundation). Information on the predominant material only should be collected. The following classification of construction materials is recommended:

1. Tile
2. Concrete
3. Metal sheeting
4. Wood
5. Bamboo
6. Palm, straw
7. Mud
8. Plastic sheeting
9. Other materials

Where the materials used on floor and on roof differ, classification specific to each should be used.

30. Elevator – availability of

5.612. This topic refers to the availability of an elevator (or lift) in a multi-storey building (categories 2.2.3–2.2.4 of the classification of type of buildings). It is recommended that the information should be collected on the availability of an elevator that is operational for most of the time, subject to regular maintenance.

5.613. This topic can be useful for providing further information for indicating the accessibility to the building or the housing unit. This is of particular relevance for older persons and persons with disabilities. In this context it could also be useful to collect information on the size of the lift (for the handicapped persons and ambulance transport), if the lift goes to the ground floor, and whether or not the lift stops on the same floor as the dwelling.

31. Farm building

5.614. Some national censuses may collect information to identify if a buildings or dwelling is located on a farm. A farm building may be considered as being one that is part of an agricultural holding whether it is residential or not, that is, whether it is used for agricultural or housing purposes. All the information that is relevant to other buildings and dwellings should also be collected.

32. State of repair

5.615. This topic refers to whether the housing unit or the building in which the housing unit is located is in need of repair and to the kind of repair needed. The following classification is recommended:

1. Repair not needed
2. In need of repair
   - 2.1 Minor repair
   - 2.2 Moderate repair
   - 2.3 Serious repair
3. Irreparable

5.616. Minor repairs refer mostly to the regular maintenance of the building and its component housing units, such as repair of a cracked window. Moderate repairs refer to the correcting of moderate defects such as missing gutters on the roof, large areas of broken plaster or stairways with no secure handrails. Serious repairs are needed in the case of serious structural defects of the building, such as missing shingles or tiles on the roof, cracks and holes in the exterior walls, and missing stairways. The term “irreparable” refers to buildings that are beyond repair, that is to say, with so many serious structural defects that it is deemed more appropriate to demolish the building than to undertake repairs; most usually this term is used for buildings with only the frame left standing or without complete external walls or roof.

33. Age and sex of the reference person of household (core topic)

5.617. From among the topics recommended for inclusion in the population census, age has been selected as being of most significance in relation to housing conditions. For the housing census, the data usually relate only to the housing units or building in which the housing units are located, but some characteristics of households that are related the housing condition can usefully be presented by the age and sex of household head or other reference person in the household.

5.618. While this information will usually be collected in a country’s population censuses and, if the population and housing censuses are conducted simultaneously, as is the practice in the majority countries, then information on age of the head or other reference member of the household will be collected together with other relevant demographic characteristics in the population part of the census. If, however, the housing census is collected independently of the population census, then there should be a separate provision for collecting this information.

34. Tenure (core topic)

5.619. Tenure refers to the arrangements under which the household occupies all or part of a housing unit. The unit of enumeration is a household occupying a housing unit. The classification of households by tenure is as follows:

1 Household owns housing unit
2 Household rents all or a part of housing unit
  2.1 As a main tenant
  2.2 As a subtenant
3 Household occupies housing unit partly free of rent
4 Household occupies housing unit wholly free of rent
5 Household occupies housing unit under some other arrangement

5.620. National circumstances can dictate the need to assess the number of households occupying the housing unit free of rent to further distinguish whether such arrangement is with or without the consent of the owner. However, this information regarding the consent of the owner is subject to special scrutiny in terms of reliability. Furthermore, in countries where communal ownership is significantly represented, this topic on tenure might be further expanded in order to capture tenure arrangements of communally owned housing. Likewise, the category "other arrangements" can be extended to capture forms of tenure specific to some countries.

5.621. The information on tenure collected in the census needs to be clearly distinguished from the information on ownership (see paragraphs 5.540–5.545) and is one that should be asked of all households, otherwise there is a danger that it may be omitted in cases where more than one household occupies a single housing unit. Under some circumstances, it may be useful to indicate separately households that, although not subtenants in the sense that they rent from another occupant who is a main tenant or owner-occupant, rent part of a housing unit from a landlord who lives elsewhere. These households and subtenant households may be of special significance in formulating housing programmes. On the contrary, in countries where subtenancy is not usual, information on subtenants may not be collected in the census or, if collected, may be tabulated only for selected areas.

5.622. In countries where the land and the living quarters are frequently occupied under separate tenure, the topic may be expanded to show separate information for the tenure under which the household or households occupy the living quarters and for the tenure of the land upon which those living quarters are located.

35. Rental and housing costs

5.623. Rent is the amount paid periodically (weekly, monthly, and so forth) for the space occupied by a household. Information may be obtained on the basis of a scale of rents rather than on that of the exact amount paid. The data may be considered in relation either to household characteristics or to the characteristics of the living quarters. In the latter case, where more than one household occupies a single set of living quarters, the rents paid by all the households will need to be summed in order to obtain the total rent for the living quarters. In the case of living quarters that are partly rented and partly owner-occupied, it may be necessary to impute the rent for the owner-occupied portion.

5.624. In countries where rent for the housing unit is paid separately from rent for the land upon which the housing unit stands, separate information may need to be collected reflecting the amount of ground rent paid.

5.625. In addition to the amount of rent paid by renting households, it may be useful to collect information on the housing costs, which could include information on monthly mortgage payments (for owner-occupiers), taxes and cost of utilities.

36. Furnished/unfurnished

5.626. Countries may wish to make some provision for indicating whether the housing units covered by the rent are furnished or unfurnished and whether utilities such as gas, electricity, heat and water are included.

37. Information and communication technology devices – availability of (core topic)

5.627. Information and communication technology (ICT) devices play an increasingly crucial role in modern society. Modern technology, with its ever-evolving array of devices, has fundamentally reshaped how we communicate and consume media. This constant innovation disrupts traditional structures and patterns, creating entirely new ways for societies to engage with information and connect with each other. Population censuses offer a valuable opportunity to assess household access to ICT devices. The chosen categories of devices to be included in the census should provide sufficient insight into the role of ICT within households and support government and private sector planning. This data can help optimize service delivery and understand the societal impact of technology.

5.628. The International Telecommunication Union (ITU) recommends including two topics in population censuses (collected at the household level):

(i) Household having a computer (desktop, laptop, or tablet)
(ii) Household with Internet access

5.629. In designing the questions, census planners should differentiate between two distinct viewpoints, namely (a) the availability of ICT devices to the households; and (b) use of ICT devices by the household members. The distinction is important, since households need not own, but may still have access to computers, mobile phones and the Internet through devices provided by school, university or work to household members. It also means that countries interested in collecting information on ICT use, particularly of the Internet and mobile phones, would need to include a relevant question topic in their census individual form. The rational for adopting either viewpoint, or even a combination of both, is not necessarily only technical, but rests more on the prevailing conditions in the society, and on how the information will be used to characterize the socioeconomic profile of households of a country. Detailed usage statistics, including the intensity (frequency) of use and the full range of activities performed, are preferably obtained using household surveys.

5.630. Computer refers to a desktop computer, a laptop (portable) computer or a tablet (or similar handheld computer). They exclude terminals connected to mainframe computers for data processing, and midrange multi-user systems that are primarily intended for shared use. Devices with some embedded computing abilities, such as smart TV sets, and devices with telephony as their primary function are also excluded. 'Household with a computer' means that the computer is generally available for use by all members of the household at any time, regardless of whether it is actually used. The computer may or may not be owned by the household, but should be considered a household asset.

5.631. Internet access from home refers to the ability of the household to connect to the Internet using TCP/IP protocols. Internet connections may be classified according to the technology employed, devices

used, communication medium, or connection bandwidth (speed). Internet access at home is meant to include both narrowband and broadband connections. Access to the Internet is measured irrespective of the type of access, device used to access the Internet, or the method of payment. 'Households with Internet access' means that the Internet is generally available for use by all members of the household at any time, regardless of whether it is actually used. The connection and devices may or may not be owned by the household but should be considered household assets. If one member of the household has a mobile phone with connection to the Internet and makes it available for all members, then it should be considered that the household has access to the Internet.

38. Cars – number of available

5.632. This topic refers to the number of cars or vans normally available for use by members of the household. The term "normally available" refers to cars and vans that are either owned by occupants or are under some other more or less permanent agreement, such as a lease, and includes those provided by an employer if available for use by the household, but excludes vans used solely for carrying goods or other commercial purposes.

39. Durable household appliances – availability of

5.633. The unit of enumeration is a household occupying a housing unit and information may be collected on the availability, within the housing unit, of durable appliances such as washing machines, dishwashing machines, refrigerators, deep freezes and microwave cookers, depending on national circumstances.

40. Outdoor space – access to

5.634. This topic refers to the reasonable access to an outdoor space intended for the recreational activities of the members of a household occupying a housing unit. The classification can refer to any outdoor space that is available:

1 As part of a housing unit (for example, a garden or backyard)
2 Adjacent to the building (for example, playgrounds placed next to the apartment building)
3 As part of common recreational areas within a walkable distance from the housing unit (for example, parks, lakes, sports centres and similar sites)
4 Beyond a 10-minute walk.