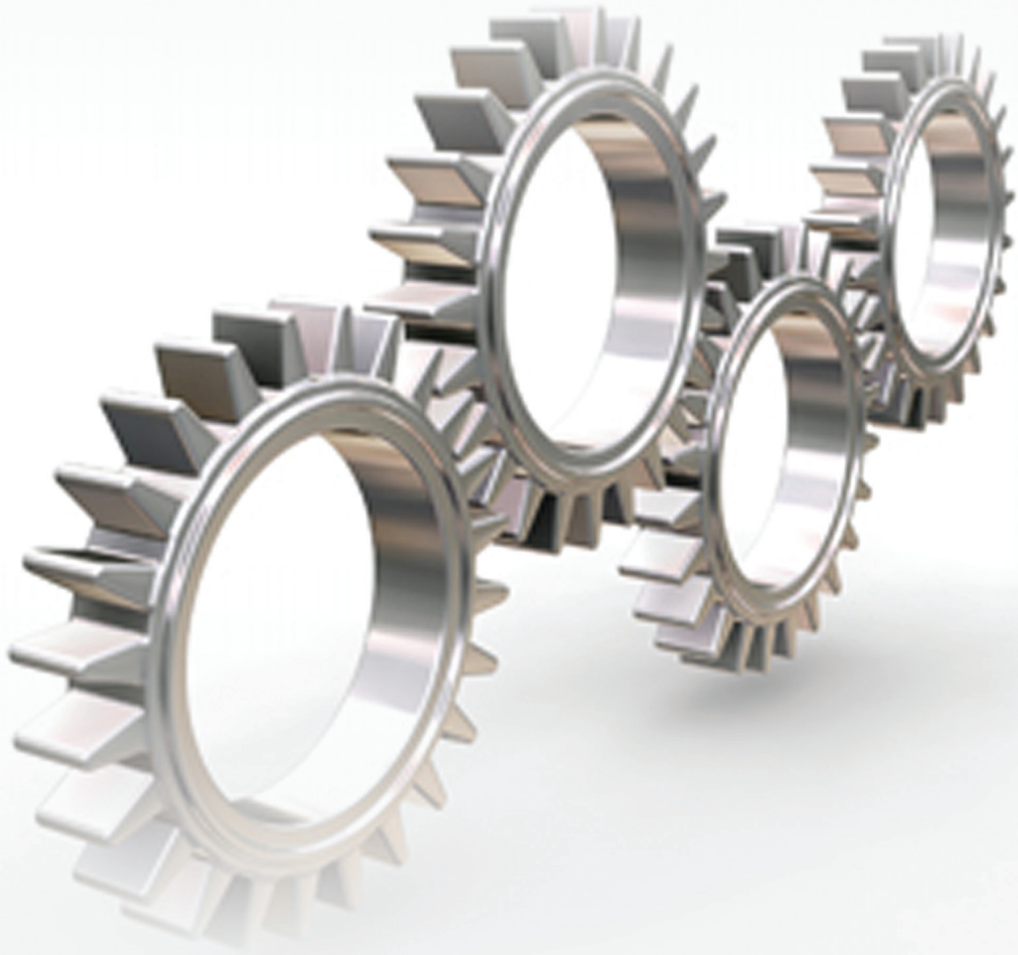
EUROPEAN COMMISSION
EUROSTAT

Handbook on Data Quality Assessment Methods and Tools



Mats Bergdahl, Manfred Ehling, Eva Elvers, Erika Földesi,
Thomas Körner, Andrea Kron, Peter Lohauß, Kornelia Mag,
Vera Morais, Anja Nimmergut, Hans Viggo Sæbø,
Ulrike Timm, Maria João Zilhão

Manfred Ehling and Thomas Körner (eds)

Cover design: Siri Boquist Photo: Crestock

EUROPEAN COMMISSION
EUROSTAT

Handbook on Data Quality Assessment Methods and Tools

Mats Bergdahl, Manfred Ehling, Eva Elvers, Erika Földesi,
Thomas Körner, Andrea Kron, Peter Lohauß, Kornelia Mag,
Vera Morais, Anja Nimmergut, Hans Viggo Sæbø,
Ulrike Timm, Maria João Zilhão

Manfred Ehling and Thomas Körner (eds)

Contributors to the handbook:

Manfred Ehling (chair), Federal Statistical Office Germany
Thomas Körner (chair till 6.2.2007), Federal Statistical Office Germany
Mats Bergdahl, Statistics Sweden
Eva Elvers, Statistics Sweden
Erika Földesi, Hungarian Central Statistical Office
Andrea Kron, Federal Statistical Office Germany
Peter Lohauß, State Statistical Institute Berlin-Brandenburg
Kornelia Mag, Hungarian Central Statistical Office
Vera Morais, National Statistical Institute of Portugal
Anja Nimmergut, Federal Statistical Office Germany
Hans Viggo Sæbø, Statistics Norway
Katalin Szép, Hungarian Central Statistical Office
Ulrike Timm, Federal Statistical Office Germany
Maria João Zilhão, National Statistical Institute of Portugal

Wiesbaden, 2007

Contents

1 Introduction

Production of high quality statistics depends on the assessment of data quality. Without a systematic assessment of data quality, the statistical office will risk to lose control of the various statistical processes such as data collection, editing or weighting. Doing without data quality assessment would result in assuming that the processes can not be further improved and that problems will always be detected without systematic analysis. At the same time, data quality assessment is a precondition for informing the users about the possible uses of the data, or which results could be published with or without a warning. Indeed, without good approaches for data quality assessment statistical institutes are working in the blind and can make no justified claim of being professional and of delivering quality in the first place.

Assessing data quality is therefore one of the core aspects of a statistical institute's work. Consequently, the European Statistics Code of Practice highlights the importance of data quality assessment in several instances. Its principles require an assessment of the various product quality components like relevance, accuracy (sampling and non-sampling errors), timeliness and punctuality, accessibility and clarity as well as comparability and coherence. The code at the same time requires systematic assessments of the processes, including the operations in place for data collection, editing, imputation and weighting as well as the dissemination of statistics.

Several efforts of implementation of data quality assessment methods have been undertaken in recent years. In succession of the work of Leadership Expert Group (LEG) on Quality some development projects have been carried out concerning assessment methods like self-assessment, auditing, user satisfaction surveys etc. (Karlberg and Probst 2004). Also a number of National Statistical Institutes (NSIs) have developed national approaches (see, e.g., Bergdahl and Lyberg 2004). Nevertheless and despite the importance of the topic being generally agreed, there is no coherent system for data quality assessment in the European Statistical System (ESS). The report on the ESS self-assessment against the European Statistics Code of Practice points in this direction and suggests that quality control and quality assurance in the production processes are not very well developed in most NSIs (Eurostat 2006c).

This Handbook on Data Quality Assessment Methods and Tools (DatQAM) aims at facilitating a systematic implementation of data quality assessment in the ESS. It presents the most important assessment methods: Quality reports, quality indicators, measurement of process variables, user surveys, self-assessment and auditing, as well as the approaches labelling and certification. The handbook provides a concise description of the data quality assessment methods currently in use. Furthermore, it gives recommendations on how these methods and tools should be implemented and how they should reasonably be combined: An efficient and cost-effective use of the methods requires that they are used in combination with each other. E.g. quality reports could be the basis for audits and user feedback. The handbook presents numerous successful examples of such combinations. Via the recommendations provided, the handbook at the same time aims at a further harmonisation of data quality assessment in the ESS and at a coherent implementation of the European Statistics Code of Practice.

The handbook is primarily targeted towards quality managers in the ESS. It shall enable them to introduce, systematise and improve the work carried out in the field of data quality management in the light of the experiences of colleagues from other statistical institutes within the ESS. The handbook shall also help to avoid overburdening the subject matter statisticians with assessment work and making data quality assessment an effective support for their work. Finally, the handbook should support top management in their managerial planning in the quality field.

After a short presentation of the basic quality components for products, processes and user perception, chapters 2 and 3 give concise descriptions of each of the methods. The presentation focuses on the practical implementation of the methods and, if applicable, their interlinkages among each other. The handbook also names up-to-date examples from statistical

institutes (see ANNEX B). In order to facilitate the use of the handbook, the chapters presenting the methods are following a standardised structure covering the following items:

- Definition and objectives of the method(s)
- Description of the method(s)
- Experiences in statistical institutes
- Recommendations for implementation
- Interlinkages with other methods (where applicable)
- Recommended readings

Chapter 4 proposes a strategy for the implementation of the methods in different contexts. The handbook recommends a sequential implementation of the methods, identifying three packages with increasing level of ambition. But of course a particular NSI may apply methods and tools from different packages at the same time given the particular circumstances in which they function.

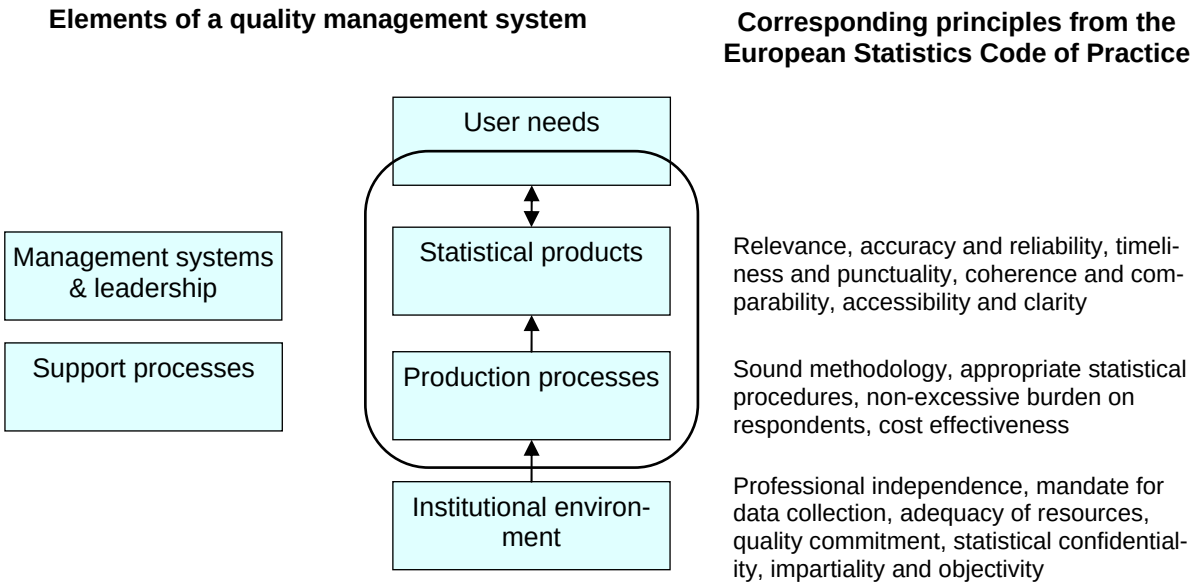
The number of pages of the handbook being heavily restricted, the handbook can not go very much into detail. Especially in order to be able to present more examples and to elaborate certain aspects in more detail, a comprehensive annex is provided together with the handbook. First it includes a background paper on the position of data quality assessment in the general framework of quality management (ANNEX A). ANNEX B presents good practice examples in some more detail. Furthermore, the annex provides a systematic presentation of basic quality tools (ANNEX C) and a glossary (ANNEX D).

1.1 Scope of the Handbook

Data quality assessment is an important part of the overall quality management system of a statistical agency (see ANNEX A for more details). However, its scope is limited to the statistical products and certain aspects of the processes leading to their production. Thus, the handbook does not cover areas like the support processes, management systems or leadership. Neither does it cover the institutional environment of statistics production.

Figure 1 shows the issues of DatQAM within the context of quality management. It also refers to the relevant principles in the European Statistics Code of Practice.

Figure 1: Scope of the handbook within the context of quality management



The methods and tools presented in this handbook facilitate an assessment of statistical products, statistics production processes, as well as the user perception of statistical products.

Before discussing methods and tools it should be clarified what is meant by method and what is meant by tool. In the context of this handbook the term assessment method refers to the approach of evaluation, e.g. documenting/reporting, calculating (indicators), auditing, self-assessing, questioning the user. The term assessment tool refers to the concrete form how the method is implemented, e.g. producing a quality report, calculating key indicators, an auditing procedure, a checklist or a user survey.

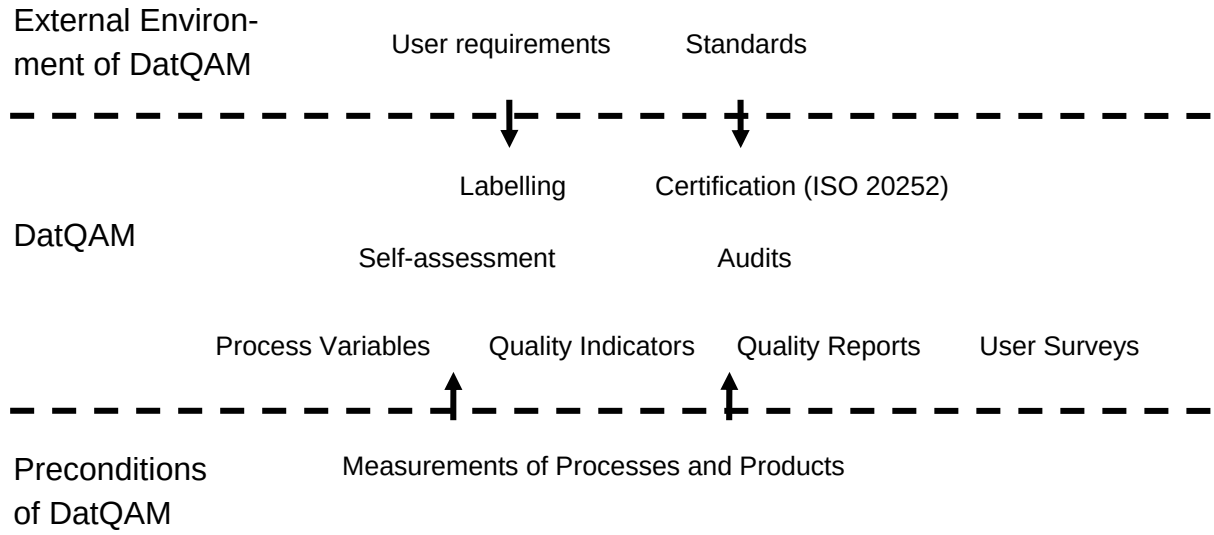
To a certain degree, the methods are relying on a number of preconditions. On the one hand, the application of data quality assessment methods always requires some basic information on the products and processes under consideration. For this reason, at least a basic systematic quality measurement regarding processes and products should be in place. There also has to be some documentation system giving access to key characteristics of the products and processes. On the other hand, data quality assessment methods require an (internal or external) reference against which the assessment can be carried out. Such reference can be provided in the form of general quality guidelines, policies, minimum standards, ISO (International Organization for Standardization) standards or as process specific guidelines (e.g. for questionnaire testing or editing). Similarly, the user requirements are a further key input to data quality assessment.

As figure 2 shows, different data quality assessment methods tend to be either closer to “measurement” or closer to “standards and user requirements”. On the way from the “measurement” to “standards and user requirements”, information is being more and more condensed and hence more appropriate for the information of managers, users or interested parties. The arrows in the figure indicate that certain methods are normally based on the output from other methods. For example, quality indicators are relying on the results from quality measurement. This handbook distinguishes three levels of data quality assessment.

On the first level, the complex information obtained from measurement of processes and products has to be selected and structured in order to become meaningful for data quality assessment. For this purpose, methods like use of key process variables, quality indicators, quality reports, and user surveys are being used. Key process variables start from the information obtained from the processes, sometimes referred to as paradata. Using statistical methods this information is further condensed and thus made available for assessment purposes. For example, data gathered from interviewer control processed, using e.g. control charts, in order to identify possible problems in data collection. Quality indicators select information from data analysis, giving more insight e.g. into non-response bias etc. The user surveys are less based on information from documentation (maybe in the case of a complaint management system), but still measure directly the user perception of a given statistical product.

On the second level, based on the information compiled on the first level, the conformity of the individual statistics is assessed against (internal or external) standards. In a self-assessment, the assessment is carried out by the person or team in charge of the survey or process in question him or herself (or the team). Audits or peer reviews do introduce a neutral (and sometimes external) expert. Self-assessment, audits and peer reviews are based on information from quality indicators and reports, process variables, and user surveys. Furthermore, they sometimes might use specifically designed checklists in order to present the information needed in a more structured and accessible way. In general, self-assessments and audits show a broader picture with less details then process variables, quality indicators, quality reports and user surveys.

Figure 2: The DatQAM map



Self-assessments and audits provide an overall picture of the conformity of a given statistics with the standards. At the same time, they still provide information on various quality components or main processes. For instance, for the orientation of users and the general public it is useful to further condense this information.

On the third level the approaches of labelling and certification are coming into operation to facilitate this task. Both methods have in common that compliance is certified with a whole set of defined requirements. As the most important ISO standard in the field of data quality assessment, this handbook presents the requirements for a certification according to the international standard ISO 20252 “Market, opinion and social research – Vocabulary and service requirements”. Put into practice only very recently, the standard might be of growing importance also for statistical offices. Labelling – as the term indicates – consists of providing any kind of label to statistical products or processes that meet specific quality requirements. This is one way of showing the users which products fulfil some quality criteria and also possibly which ones are not based on an agreed and established standard. The label thus provides a type of highly condensed information regarding the statistics released. Furthermore it can help to support trust and credibility in official statistics.

The distinction of the three levels has been made for the purpose of a systematic presentation. In practice, the methods will mostly be implemented in combination, e.g. quality reports together with quality indicators and audits. There are also connections with activities that are not directly data quality assessment. For example, quality reports and quality indicators should (ideally) be included in a comprehensive metadata system combining the technical documentation of the data with a documentation of data quality.

It should be noted that the handbook focuses on the use of the methods in the field of surveys based on primary data collection. The methods may however also be used in other areas of statistical activity, such as administrative registers and national accounts. In fact audits are quite common in national accounts for many years, although in comparison with primary surveys slightly different items are being covered. In contrast, in the field of administrative data and administrative registers only very limited experience is available so far.

1.2 Aspects of Data Quality

In the context of this handbook, the notion of data quality refers to three aspects: (1) the characteristics of the statistical product (2) the perception of the statistical product by the user and (3) some characteristics of the statistical production process. The three aspects are closely interrelated. The product quality is achieved through the production process. Different process designs will give priority to different product quality components. A process will never maximise all product quality components at a time (e.g. the trade-off between accuracy and timeliness). The way the product (and the process) is perceived by the user will often deviate from the way it is perceived by the producer. Even though user perception is heavily dependent on the product quality achieved, users will have their own perspective and priorities. Therefore we argue that the user will (sometimes) see the statistical product with different eyes than the statistician. For example, the user might not always have a full overview on the entire set of quality components. He or she might also give priority to other quality components (e.g. the famous “timeliness instead of accuracy”), or have difficulties to assess the certain quality components without expert support (like accuracy). For this reason it is vital that data quality assessment also covers the question how the users actually perceive the quality of a statistical product.

Data quality assessment has to take care of all three quality aspects. Focussing only on the product quality (or the process quality or the user perception respectively) will not be a sufficient solution. Some of the methods shown in figure 2 are specifically targeted towards one of the three aspects: Key process variables monitor the production process, quality reports and quality indicators contain mostly information regarding product quality, and user surveys facilitate the assessment of the user perception. The other methods (self-assessments, audits, labelling, certification) can cover all three aspects. Nevertheless, according to the specific context, priority is often given to one of the aspects.

1.2.1 Product quality components

In the European Statistical System (ESS), product quality of statistics is assessed according to the ESS quality components.[1] These components are central for any assessment of product data quality in statistics (but may also be used as a base to assess user perception). The product quality components are defined by Eurostat (2003d), and are identical with the European Statistics Code of Practice principles covering Statistical Output (with minor wording differences), given in figure 1. A brief description of these components follows (adapted from the Eurostat definition of quality of statistics).

Relevance

Relevance is the degree to which statistics meet current and potential user needs. It refers to whether all statistics that are needed are produced and the extent to which concepts (definitions, classifications etc.) reflect user needs.

Accuracy

Accuracy in the general statistical sense denotes the closeness of computations or estimates to the (unknown) exact or true values. Statistics are never identical with the true values because of variability (the statistics change from implementation to implementation of the survey due to random effects) and bias (the average of the estimates from each implementation is not equal to the true value due to systematic effects). A basic distinction is between sampling and non-sampling errors, which are both subject to variability as well as bias.

Timeliness and punctuality

Timeliness of information reflects the length of time between its availability and the event or phenomenon it describes. Punctuality refers to the time lag between the release date of data

[1] Quality components are commonly also referred to as “criteria” or “dimensions”. These terms are normally used synonymously. Nevertheless, all three terms may be used also in other quality respects.

and the target date when it should have been delivered, for instance, with reference to dates announced in some official release calendar, laid down by regulations or previously agreed among partners.

Comparability

Comparability aims at measuring the impact of differences in applied statistical concepts and measurement tools/procedures when statistics are compared between geographical areas, non-geographical domains, or over time. It is the extent to which differences between statistics are attributed to differences between the true values of the statistical characteristic.

There are three main approaches under which comparability of statistics is normally addressed: comparability over time, between geographical areas, and between domains.

Coherence

Coherence of statistics is their adequacy to be reliably combined in different ways and for various uses. When originating from different sources, and in particular from statistical surveys of different nature and/or frequencies, statistics may not be completely coherent in the sense that they may be based on different approaches, classifications and methodological standards.

Accessibility and clarity

Accessibility refers to the physical conditions under which users can obtain data: where to go, how to order, delivery time, clear pricing policy, convenient marketing conditions (copyright, etc.), availability of micro or macro data, various formats (paper, files, CD-ROM, Internet etc.) etc.

Clarity refers to the data's information environment whether data are accompanied with appropriate documentation and metadata, illustrations such as graphs and maps, whether information on their quality is also available (including limitation in use etc.) and the extent to which additional assistance is provided by the NSI.

Quality assessment helps to solve the problem of trade-offs between different components of quality. It is becoming more and more important to analyse interactions between the different quality components (e.g. accuracy – timeliness; relevance – comparability over time etc.) and therefore it is necessary to have the right quality assessment methods in place. Then it is possible to analyse the influence of the different dimensions on the achieved total data quality.

1.2.2 User perception of data quality

The product quality components could also be used as a framework for the assessment of the user perception of a statistical product. It should be noted that the quality components are the same, but users will in many cases perceive product quality differently than the NSI. Furthermore, some of the quality components are difficult to assess for the user. For example, an assessment of accuracy of a given statistics requires at least some basic knowledge of statistical methodology. For the same reason, it will usually not be easy for non-expert users with limited knowledge of statistics to clearly define their quality requirements. Other quality components, like accessibility or timeliness, are more obvious and users are in a better position to clearly formulate their needs.

1.2.3 Process Quality

Process quality is less straightforward in its definition, and there is no ESS standard definition as for product quality. The European Statistics Code of Practice principles linked to the statistical processes are more general: Sound methodology, appropriate statistical procedures, non-excessive burden on respondents and cost effectiveness. Key process variables are usually referred to as those variables with the largest effect on product characteristics

such as the product quality components mentioned (Jones and Lewis 2003). They will vary by product quality component and by type of process. Typical process variables are resources and time used, response rates and burden as well as error rates (in editing). Processes can also be characterised by stability and capability, concepts introduced by Morganstein and Marker (1997). The chapter on measurements of process variables (2.2) discusses process quality with emphasis on its relevance for data quality assessment.

Recommended readings

Biemer, P. and Lyberg, L. (2003): Introduction to Survey Quality. Hoboken, N.J.: Wiley.

Brackstone, G. (1999): Managing Data Quality in a Statistical Agency. In: Survey Methodology 25, pp. 139-149.

Eurostat (2003d): Standard Quality Report, Methodological Documents, Working Group "Assessment of quality in statistics", Luxembourg, 2-3 October 2003. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143233/STANDARD_QUALITY_REPORT_0.PDF#search=%22Standard%20Quality%20report%202003%20Eurostat%22

Hahn, M. and Lindén, H. (2006): The European Statistics Code of Practice for a High Quality European Statistical System. Paper presented at the European Conference on Quality in Survey Statistics (Q2006), Cardiff, United Kingdom, 24-26 April 2006.

Lyberg, L. et al. (2001): Summary Report from the Leadership Group (LEG) on Quality, Proceedings of the International Conference on Quality in Official Statistics, Stockholm, Sweden, 14-15 May 2001.

2 Data Quality Assessment Methods and Tools

2.1 Quality Reports and Indicators

In order to assess data quality, first of all a clear picture of data quality is needed. Definition and components – discussed in the previous chapter are preconditions. Secondly a report on data quality is indispensable, which reflects data characteristics by quality components and presents data features according to data quality requirements.

Quality reports are important for users and producers of official statistics. Users of official statistics need to have access to a range of relevant quality measures and indicators to understand the strengths and limits of statistics and to know how to use them properly (user-oriented quality report). Producers need to have a picture on product quality in order to see the results of earlier production developments and to identify the points of further improvements (producer-oriented quality report). For this end they need the most detailed quality reports and a number of indicators, involving the processes behind.

Whoever the user is, the data producer has to compile a quality report to characterise quality components, and search for indicators to illustrate these features, taking into account that different users have different needs concerning quality information.

Quality reports and indicators provide documentation of the quality features of statistical products. They are the key reference documents for quality assessment. For this reason they form an important input for auditing and self-assessment.

The next subchapters will cover the definition and the description of quality reports and of quality indicators. Due to the close interrelationship between reports and indicators it seems logical to discuss the national experiences on them as a whole, to compile one common set of recommendations and to explain the interlinkages between methods used for quality assessment.

2.1.1 Quality Reports

Definition and objectives

Quality reporting is the preparation and dissemination, on a regular or irregular basis, of reports conveying information about the quality of a statistical product or survey (Eurostat 2003b).

A quality report provides information on the main quality characteristics of a product so that the user should be able to assess product quality. In the optimal case quality reports are based on quality indicators.

Description

The main potential target group of a quality report is composed of the users of statistical products outside or inside the statistical institute. Moreover, quality reports on statistical products are important to the producers and the management, too, for monitoring purposes. Although the quality requirements and statistical expertise of user groups are different[2] and therefore a single quality report would not satisfy them all, a standard structure is preferable. The standard structure makes it easier to find the use- and user-specific relevant information, and facilitates comparability over time or among products (see examples of the many different quality reports' structures in ANNEX B).

In order to assist member states to compile quality reports, Eurostat as a user of member states data and data producer at European level, developed a guide "Standard Quality Report" (Eurostat 2003d) and a comprehensive handbook "How to make a Quality Report" (Eurostat 2003b).

[2] For example, some users are more interested in timeliness when statistics are used as indicators for decision-making, other users in accuracy when data are used for econometric analysis. Naturally their statistical knowledge is different, they need and understand different information of quality.

In order to meet the requirements of the "Standard Quality Report" (Eurostat 2003d), the quality report should be in line with at least the Eurostat definition and the components of quality in statistics (Eurostat 2003a). It means that for each quality component a checklist should be completed.

Briefly, the quality report should give a description and an assessment of quality on the following:

- user satisfaction concerning relevance,
- sampling and non-sampling (measurement, data processing) errors, concerning accuracy,
  - indicator examples: coefficient of variation, unit response rate, item response rate, average size of revisions
- key production dates concerning timeliness and punctuality,
  - indicator examples: punctuality of time schedule of effective publication, time lag between the end of reference period and the date of the first/final results
- forms for dissemination, availability of metadata and documentation, concerning accessibility and clarity,
  - indicator examples: number of publications disseminated or sold, number of accesses to databases
- changes over time, differences between national and European concepts, concerning comparability,
  - indicator examples: number and length of comparable time series, asymmetries for statistics mirror flows
- coherence with other sources.

In addition – although they are not quality components – it is advisable to add a general description of

- the survey and methods used, and
- cost and burden in order to support quality assessment.

On this basis a very detailed "full" quality report is obtained, which addresses the most qualified users in statistics.

In practice, several simple forms are in use as users do not need for their purposes abundant information. They prefer to receive relevant, transparent, easily understandable information in a convenient form, and in a timely manner. The "full" quality report can be considered as a broad framework. For special types of users or for special purposes the relevant parts can be selected to compile a user oriented "summary" or a "basic" quality report (Jenkinson 2006). Examples are the quality reports required by Eurostat in different subject matter areas, or the quality profile for structural indicators available from Eurostat quality homepage. If there are no quality reports the description of the production process or process tables (i.e. quality reports on regional national accounts) can be used as a source of quality information.

Beyond the general Standard Quality Report, some subject matter areas at European level and some National Statistical Institutes (NSIs) at national level have developed their own standard quality report for specific topics. Recently special quality reports were developed for annual national accounts (Gross National Income (GNI) quality report), for regional accounts and quarterly national accounts.

Quality profile[3] is a special type of quality report in the European Statistical System (ESS); it is a user-oriented summary of the main quality features of policy-relevant indicators, as a

[3] Quality profile is used in more general meaning as well, like a collection of all available information on the quality of a system. An example is the survey quality profile in U.S. education statistics where information on errors associated with each phase of survey operations is presented (Kalton et al. 2000).

special type of data. It aims at guiding the use and analysis of the existing indicators as well as at providing input to political selection process of indicators. The Statistical Programme Committee at its 51st meeting in 2003 broadly supported the quality profile for structural indicators[4]. The quality profile adopted is released on the Eurostat website as a central entry point to communicate quality issues related to structural indicators.

The standard quality report served as a guideline to assess the quality of Euro-IND[5] – European and national short term indicators – database completed recently (Eurostat 2006a).

International data users expect to be informed on data quality assessment of each member state's data and expect an overall assessment from the Eurostat perspective as well. Quality characteristics are more and more frequently published together with the data, and quality reports form part of the documentation (Elvers 2004). There are efforts to develop standard metadata formats, to cover quality reports as well (Statistical Data and Metadata Exchange, http://www.sdmx.org/).

The increasing use of administrative data for statistical purposes underlines the need for reports on their quality. Generally statistical agencies assess the quality of administrative data based on the metadata available and their own examinations. Two types of internal quality reports are used:

- A source-specific internal quality report gives a general description of the administrative data source and thus follows the structure of the standard quality report and provides general quality information for potential internal users. Administrative data are used for very different statistical purposes, like registers, sampling frames, background information on population for stratification, as data sources for units in the survey, data editing, imputation, verification etc., each purpose requiring different features.
- A product-specific internal quality report gives a specific picture on the extraction or combination of administrative data sets for the specific use or intended product. Depending on the type and use of data the content of the quality report should reflect from among the following characteristics the most relevant relevant features: administrative concepts, clarity based on metadata, envisaged use of the data, coverage, errors in data, reference time, data up-to-dateness, record matching ability, confidentiality and privacy protection, compatibility between file formats, comparability of administrative datasets.

The Eurostat paper "Quality assessment of administrative data for statistical purposes" (Eurostat 2003h) gives a more detailed overview.

The preparation and updating of quality reports depend on the survey frequency and on the stability of the quality characteristics, where a balance should be sought between the need for recent information and report compiling burden. If necessary, the quality report should be updated as frequently as the survey is carried out. However, if the characteristics are stable, the inclusion of the quality indicators on the newest survey results could be enough to update the report. Another solution is to provide a detailed quality report less frequently, and a shorter one at each survey, covering only the updated characteristics, like some accuracy-related indicators. This is the practice required by Eurostat in the case of the Labour Force Survey. Further developments are expected in the elaboration of more use- and user-oriented quality reports. The systematic involvement of a wide range of users is a precondition.

---

4 Structural indicators are tools for assessing progress on Lisbon strategy.

5 The objective of Euro-IND is to provide a comprehensive and detailed portrait of the short-term economic situation in: the Euro-Zone, the EU, the Member States, and other countries.

2.1.2 Quality indicators

Definition and objectives

Quality indicators are specific and measurable elements of statistical practice that can be used to characterise the quality of statistics.

The use of indicators in describing and monitoring features or facts can be considered as a general method. The indicators as simplified and generally quantified measures – calculated according to clear rules – intend to characterise a complex phenomenon, in this case the many different quality features of data.

Sometimes the indicators do not measure quality directly but they can give enough information to provide an insight into quality (Guidelines for measuring statistical quality, Office for National Statistics (ONS), UK 2006).

The quality indicators measure the quality of statistical products or processes from several aspects. Some product quality indicators are derived from processes, therefore they are called process variables as well (chapter 2.2 on process variables).

Eurostat proposes standard quality indicators (Eurostat 2003f; Eurostat 2005d) that can be used, from the producers' point of view, for summarising the quality of statistical products in various statistical domains.

It is not always possible to measure exactly the quality of a product with indicators. In these cases quality report provides further information for supporting the assessment of the quality components.

Description

Quality indicators are the most widely used tools to measure the quality of statistics. Indicators are developed to change the measurability level from nominal to ordinal or interval scale, and to find indicators more closely related to the phenomenon, namely to move from indirect to direct indicators.

Quality indicators make the description of a product by quality components more informative and increase transparency. The statistician or the user can assess the quality of different surveys or the same data in different periods by using the quality indicators. Indicators always simplify reality. However, there is a danger of false interpretation of quality indicators if the background information is not taken into account as well. When quality indicators are used to inform users on the quality of statistics, it is recommended to include qualitative statements helping to interprete quality information and to summarise the main effects on the usability of the statistics (Eurostat 2005d).

Quality indicators can be grouped according to the following aspects:

- Orientation:

  Quality can be measured, from the producers' point of view, with producer-oriented quality indicators, while from the aspect of users, with user-oriented quality indicators.
- Reference:
  - Product quality indicators measuring the quality of a statistical product, e.g. statistical data.
  - Process quality variables measuring the quality of a statistical process, e.g. data collection.
- Quality components:
  - E.g. relevance, accuracy, timeliness and punctuality, accessibility and clarity, comparability, coherence.

Some quality indicators should be produced for each output in line with the frequency of production or publication (for example, standard errors should be calculated related to each new estimate). However, some quality indicators should be produced once for longer periods, and

should only be rewritten when major changes occur (e.g. time lag between the end of the reference period and the date of first results) (ONS 2006). The calculation frequency of the indicators depends on the purpose of quality indicators (e.g. monitoring the quality over time) or on the survey or publication frequency.

In order to improve the comparability of data quality, the Eurostat working group "Assessment of quality in statistics" proposes Standard Quality Indicators (Eurostat 2005d) for the NSIs. These indicators are producer-oriented quality indicators from Eurostat's point of view. The indicators measure all quality components identified in the quality definition of Eurostat. The above-mentioned paper contains the methodology for the computation of indicators and the metadata needed.

Table 1: List of Standard Quality Indicators (Eurostat 2005d)

Quality component	**Indicator**	**1=Key 2=Supportive 3=Advanced**
Relevance	R1. User satisfaction index	3
	R2. Rate of available statistics	1
Accuracy	A1. Coefficient of variation	1
	A2. Unit response rate (un-weighted/weighted)	2
	A3. Item response rate (un-weighted/weighted)	2
	A4. Imputation rate and ratio	2
	A5. Over-coverage and misclassification rates	2
	A6. Geographical under-coverage ratio	1
	A7. Average size of revisions	1
Timeliness and Punctuality	T1. Punctuality of time schedule of effective publication	1
	T2. Time lag between the end of reference period and the date of first results	1
	T3. Time lag between the end of reference period and the date of the final results	1
Accessibility and Clarity	AC1. Number of publications disseminated and/ or sold	1
	AC2. Number of accesses to databases	1
	AC3. Rate of completeness of metadata information for released statistics.	3
Comparability	C1. Length of comparable time-series	1
	C2. Number of comparable time-series	1
	C3. Rate of differences in concepts and measurement from European norms	3
	C4. Asymmetries for statistics mirror flows	1
Coherence	CH1. Rate of statistics that satisfies the requirements for the main secondary use	3

The main purposes of the "Quality Indicator" Task Force developing these indicators can be summarised as follows:

- The standard quality indicators should be representative for the respective quality components,
- The methodology for their computation should be well established, and
- The indicators should be easy to interpret.

According to the different levels in achieving these purposes, the indicators are classified as key indicators, supportive indicators and indicators for further experiences.

The key indicators should meet all the above-mentioned requirements. The supportive indicators are important as indirect measures of the data quality. The third category of indicators should be subject to further examination and discussion across statistical agencies (Eurostat 2005d).

Some of the proposed indicators are intended to characterise the European aggregates; these indicators were created for Eurostat's purposes. For example A6: Geographical undercoverage ratio. But others could be used by the NSIs at country level as a basis of their own quality indicators. In this case the standard quality indicators could be adjusted to the national practice.

Further methodological developments are necessary in the following fields:

- Developing the standard quality indicators in the second and third categories (supportive indicators and indicators for further experiences),
- Improving the comparability of quality indicators for example among different periods, and
- Developing new quality indicators to improve measuring the quality components.

The applicability and importance of the different quality indicators and quality components depend on the data and users:

- In case of the quality of flash estimates of Gross Domestic Product (GDP) timeliness has a priority over accuracy which can be characterised by the average size of revision.
- The quality profile of structural indicators provides an overall technical assessment of each indicator according to three quality grades as an outline for non-specialist users. These grades reflect only the accuracy and comparability components (Commission of the European Communities 2004).
- In case of the statistical use of administrative data the quality indicators should reflect not only the quality of the statistical product but the quality of the source as well, because the statistical institutions have minor control over the quality of used administrative data (Birch 2005).

Critical values can help to assess the quality via quality indicators. However, the critical value of an indicator can be different depending on the respective statistics. It can also be a precondition of the publication as it is the practice for example in the U.S. Census Bureau. Suppression occurs when one or more of the following criteria are met: the Relative Standard Error is greater than 20 percent; the imputation rate is greater than 50 percent; etc.

There were several attempts to compile one composite indicator to summarise data quality features and give one overall quality measure. A composite indicator should be based on a theoretical concept which would allow individual indicators to be selected, combined and weighted by reflecting the components of the quality. These attempts failed since the set, trade-offs and importance of indicators/components are product- and user-specific. Up to now it was not possible to find a general solution. However, for special well-defined purposes – i.e. assessing the overall quality of statistics disseminated in the EU for a special user group – such a composite indicator can be developed from a conceptual and practical point of view. This needs further developments to define a single meaningful indicator for each quality component, and achieve agreement on the weighting procedures. Even if we can develop a composite indicator, its comparability over time is endangered when serious changes occur in key elements like measurement in different countries.

2.1.3 Experiences in statistical institutes

NSIs produce quality reports required by international organisations like Eurostat, Organization for Economic Cooperation and Development (OECD), International Monetary Fund (IMF), as well as Food and Agriculture Organization (FAO), and deliver them together with the data. Moreover, some NSIs at national level have developed their own standard quality reports for their own purposes or for their users.

Statistics Sweden, Statistics Norway and CBS (Netherlands) have long traditions in preparing quality reports for their statistical products.

Among national practices the Austrian quality concept, the development of their in-house quality report system and their standard documentation system is available on the internet (Burg 2005), the main features are outlined in the example in ANNEX B.

In 2001 the ONS launched a project (Full and Jones 2002) dealing with the standardisation of quality measurement. The project aimed at identifying and developing suitable quality measures and indicators for survey and non-survey outputs and at incorporating them in the Guidelines for Measuring Statistical Quality (ONS 2006). This project covers the identification of key quality measures – the smallest set of measures and indicators which provide a summary of the overall quality of an output. The project team investigates a quality grading scheme for the evaluation of the quality of a product.

In 2006 at the European Conference on Quality in Survey Statistics, Jenkinson (2006) presented the current UK quality reporting system and its application to the National Accounts. Within the ONS National Accounts, three levels of quality report have been set up for GDP:

- Summary Quality Report,
- Basic Quality Information, and
- Full Quality Report.

The Summary Quality Report is a static, web-based summary of quality information that can be compiled once and will apply to all releases of the GDP. Basic Quality Information (BQI) consists of dynamic, release-specific quality measures and the most important quality information. The BQI is incorporated into the background notes of the First Release for the GDP. The Full Quality Report will consist of all quality measures and indicators produced for the GDP (it will be available at the end of 2007).

It should be noted that in February 2006 Eurostat submitted a similar proposal for a three-level quality reporting system: one report for the producers and two different for the users (Eurostat 2006b).

The first evaluation of NSI's self-assessment against the European Statistics Code of Practice Questionnaire (Eurostat 2005c) provided a comprehensive overview on the national practices in applying quality indicators. According to this report, there were large differences in the application of the quality indicators (list of indicators, statistical products characterised by indicators, etc.). All NSIs regularly assessed at least the accuracy of data. But less than half (13 out of 29) NSIs provided coefficient of variation for more than 75% of their statistics.

In the Istat (Italian Statistical Office) an Information System for Survey Documentation (SIDI) has been developed to support the survey managers in their quality control activity (Brancato et al. 2004). The results of the new strategy for a compulsory implementation of the standard quality indicators within SIDI was presented in 2004 at the European Conference on Quality. The main purposes of the Italian system are:

- to allow the survey managers to monitor their production processes, to evaluate the quality over time and to compare the indicators with appropriate average values;
- to allow the users to analyse survey characteristics and compare the quality of different surveys;
- to provide the top management with qualitative and quantitative information for decision making purposes; and

- to collect and disseminate documents on quality-related issues.

About the Italian system see ANNEX B.

In the Slovenian Statistical Office a convenient way of the measurement of data quality of short term statistics has been developed (ANNEX B). The calculation of monthly indicators is incorporated into the data processing system and done automatically every month. The quality reports are generated in an automatic way based on the databases of the indicators and the textual information.

2.1.4 Recommendations for implementation

Recommendation on the development of quality reports and indicators (figure 3, Logical process)

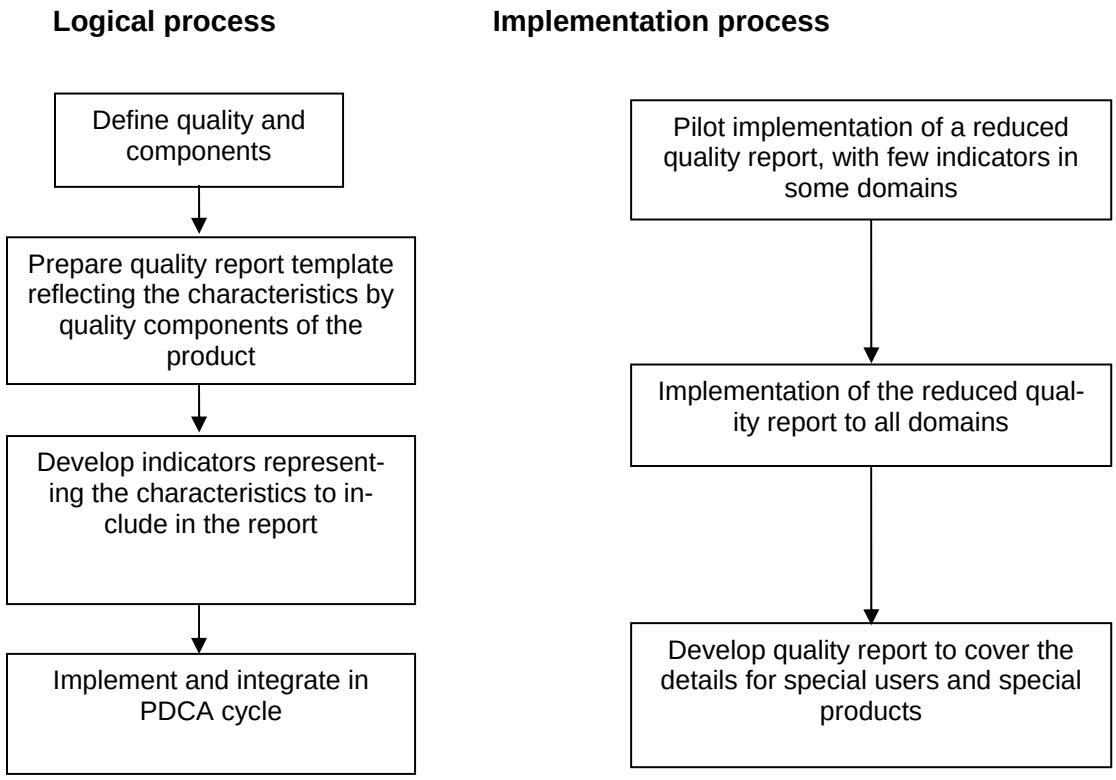
- The prerequisite is to have a definition of quality and its components as well as the knowledge of the selected target group's needs. The use of Eurostat quality concept is recommended as it enables comparability and meets the need of one of the most important users.
- The first steps are building up some prototypes of the quality report templates which present the quality characteristics of the product by components, preferably involving indicators of the different quality components. It is not advisable to describe quality solely by a set of quantitative indicators, appropriate textual information should also be provided.
- Quality reports should be tested by the users to check whether they find and understand the needed information.
- When using indicators it should be kept in mind that these indicators are not identical with the phenomena. Indicators may be misleading or focus only on a part of the phenomena; they may overemphasise the given (measurable) aspect. For example concerning accuracy, generally much more attention is given to sampling error than non-sampling error; however, the latter may have a dominant role, even if it is difficult to measure it. It is vital to improve and enhance the measurability of indicators on products and processes.
- The process of measuring quality should be an inherent part of any statistical production process and should not be a separate activity carried out after the statistics are produced or when somebody needs it. This is not only for cost and time efficiency, but for the purposes of quality improvement as well.
- The way of the dissemination of quality reports will have to be further integrated in the dissemination policy of the statistical organisation. In the short-term, when a detailed quality report is available for internal use, it might be reasonable to extract information that are possibly useful for external users and to prepare standard explanation notes how to exploit this information.
- Any quality related work on the product or the process should be integrated into a quality assessment framework. Introducing quality reporting should be accompanied by an action plan on making them available for users with the possibility of feedback, and on the use of quality reports in reviews to avoid that important reports are made just for being archived.
- Who should be involved in the development? Special expertise is needed to develop tools required by quality assessment. Survey managers and specialists should work together in developing the quality reports and indicators. Moreover users should participate in the planning procedures. In this way all the different aspects can be covered.

How to proceed when implementing quality reports and indicators

It is advisable to proceed towards the aim step by step in order to make the quality report available to all users in all statistical domains (figure 3. Implementation process).

- Quality measurement burden should remain at an acceptable level both in terms of expenses and time use. A good decision is to start with some components and indicators in some main domains. It is advisable to start with domains where some kinds of quality reports are already produced and with indicators which are calculated already or can easily be done. Derivation of metadata elements for quality reporting can help to avoid double work.
- Later on the use of this simplified quality report can be extended to all domains.
- Finally the level of detail and content of the report has to be improved to meet the special needs of the users of the given products. Preparation of quality reports needs time and efforts, they are efficient only if they are used in a proper way: their level of detail, structure and form must be suitable for the targeted users (not too long for managers, easy to understand for public users, importance of a component should be weighted up by considering which users will directly or indirectly require information on it, etc.).
- The use of the same quality report template with a limited set of standard indicators for different products will support transparency and users' understanding. Furthermore, taking into account the users' feedback, the selection of the type and frequency of quality assessment will be more effective.

Figure 3: Introduction of quality reporting



Rough indication of time and cost for compiling quality reports and indicators

The preparation time and cost of indicators and quality report depend on the circumstances: the documentation system in use, the integration level in the data processing, and the Infor-

mation Technology (IT) support of data processing. Developed documentation and metadata system or integrated data processing with IT support can reduce the burden. The estimation of time and financial resources needed can be based on the experiences of the first steps of gradual implementation.

2.1.5 Interlinkages between methods

The quality report summarises the most important information on quality. The measurable aspects of the quality can be characterised by indicators and the textual information helps to understand the limitation of a given product.

Product quality indicators give a sufficient basis for the measurement of statistical product quality. Most of these indicators are rooted in the statistical production process. For this reason, there is usually a very strong relation between the product quality indicators and the process variables. The extra information coming from the process variables can be very useful both for the users and the producers.

Labelling statistical products (chapter 3.1) may need a minimum request on quality indicators, too.

The assessment of the data quality should be based on the information on quality and on the requirements. Quality indicators and quality reports include the most important information on the quality; that is why they are used very often as the basic documents for self-assessments or audits. There are some interesting solutions where the different methods and tools are linked to each other (e.g. the linkage of the Development of a Self Assessment Programme (DESAP) checklist and the quality indicators, assessment questions that are added to the quality reports, etc.).

Recommended readings

Eurostat (2003b): Handbook “How to make a Quality Report”. Methodological Documents, Working Group “Assessment of quality in statistics”, Luxembourg, 2-3 October 2003.

Eurostat (2003d): Standard Quality Report, Methodological Document, Working Group “Assessment of quality in statistics”, Luxembourg, 2-3 October 2003.

Eurostat (2005d): Standard Quality Indicators, Working Group “Quality in statistics”, Luxembourg, 23-24 May 2005.
http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143233/STANDARD%20QUALITY%20INDICATORS.PDF#search=%22standard%20quality%20indicators%20Eurostat%22

Jenkinson, G. (2006): Standard Quality Reports and their Applications for the National Accounts at the ONS, European Conference on Quality in Survey Statistics (Q2006), Cardiff, United Kingdom, 24-26 April 2006.

2.2 Measurement of Process Variables

Data quality is normally communicated to the user according to the product quality components. Process characteristics or variables are often used as proxies for data quality and as quality indicators. A paragraph on process variables is therefore included in the handbook. The study of process variables is a prerequisite for improving processes and quality. Even if improvement techniques are beyond the scope of this handbook, a general procedure for quality improvement is briefly described.

In the European Statistics Code of Practice, main headings for indicators linked to the statistical processes are sound methodology, appropriate statistical procedures, non-excessive burden on respondents and cost effectiveness. Measurement and study of process variables are crucial for monitoring these indicators.

Definition and objectives

Measuring key process variables is a basis for data quality assessments.

"Key process variables are those factors that can vary with each repetition of the process and have the largest effect on critical product characteristics, i.e. those characteristics that best indicate the quality of the product" (Jones and Lewis 2003).

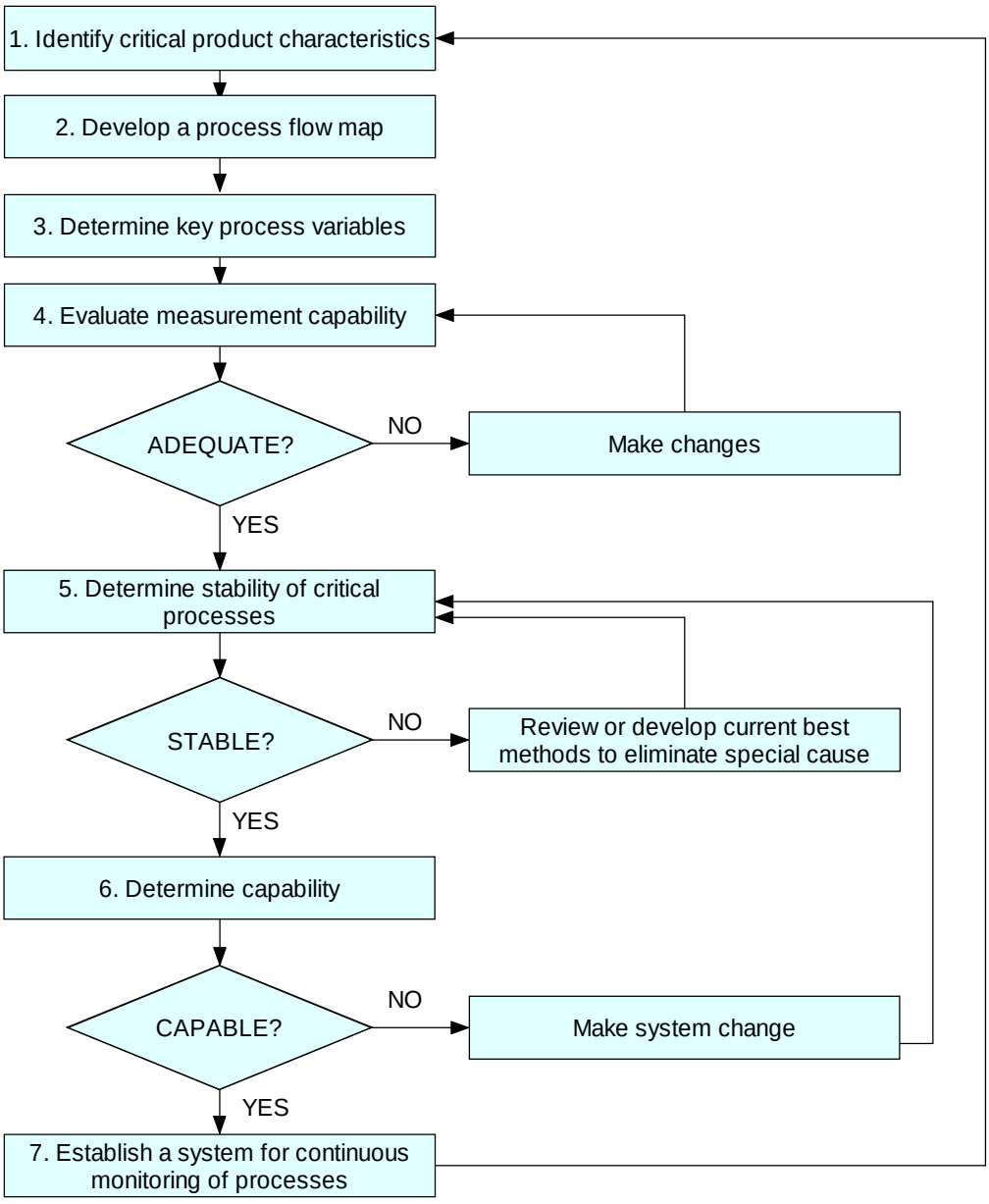
Measurement of key process variables is the basis for process management and continuous quality improvement, but also provides input to quality indicators and quality reports. Furthermore, a selection of key process variables will assume an important role in self-assessments and audits as well as in labelling and certification according to the ISO 20252. For example, the Development of a Self Assessment Programme (DESAP) self-assessment scheme (see chapter 2.4) includes measures of many variables linked to statistical production processes, from planning and data collection to processing and analyses of statistics.

Description

The method consists of using quantitative indicators in order to monitor and to assess processes over time and detect sources of error to improve existing processes. It should therefore be used for established, i.e. repetitive processes that are important for data quality. In this case, measurement series for process variables might exist and the described method with all its parts (see below) can be applied. Often (repetitive) measurements do not exist. This is the case when new processes are developed. However, in order to establish a basis for data quality assessments and future improvements it is important to start identifying and measuring key process variables anyway. The method will be applied most effectively if processes are standardised in the statistical institute at least to some degree.

Statistical institutes have always measured some process variables. Examples are measurements of non-response of different types, interviewer performance, costs and use of time for different processes. However, what has often been lacking is a systematic approach to identifying and measuring such variables. A method that is based on such measurements can be described by a flow chart for the continuous quality improvement process, introduced in the "classical" paper in this area by Morganstein and Marker (1997), which has also been used in a Eurostat handbook on process variables (Jones and Lewis 2003) (figure 4).

Figure 4: A plan for continuous quality improvement (Morganstein and Marker 1997)



The steps of this flowchart are briefly summarised in the following.

1. Identify critical product characteristics

User needs are the point of departure for this step. For statistics in general, these variables correspond to the Eurostat quality components (Eurostat 2003a), or the “output” components of the European Statistics Code of Practice (Eurostat 2005b; chapter 1 of this handbook). However, in a concrete quality improvement case these components may have different priorities, and other product characteristics may apply as well.

2. Develop a process flow map

The process leading to the product in question has to be mapped. A tool for this is the process flow map or chart. Figure 4 is an example of such a map for a process monitoring system. There is no standard for such charts, but symbols used in MS Office and some other relevant software systems are given in ANNEX C on basic quality tools. The process map normally distinguishes between input, processes, decisions and output. Responsibilities may be assigned to each step in the map.

3. Determine key process variables

It is important to determine which variables are critical given the product requirements. Process variables should not be restricted to those aspects which are easily measurable or available, which might provide a partial or biased picture of the situation. The systematic identification of the critical process variables is a precondition to the use of process variables. Tools for this task comprise cause and effect diagrams and Pareto charts (ANNEX C).

Examples of process variables are given below.

4. Evaluate measurement capability

This refers to the possibility to measure the key process variables and to do this with sufficient accuracy.

5. Determine stability of critical processes

Any process will vary. Stability means that the process and the key process variables vary within limits that in principle can be established by knowing the nature of the components behind the variation (for example random sampling), or by studying the variation pattern.

Stability does not necessarily mean that the process is good enough, measured by the average level of the process variables, or by their variation. But only a stable process can be monitored, e.g. compared before and after changes have been made.

Process variation is commonly classified into two types: Special cause and common cause. Special causes are irregular in occurrence, and can be dealt with without changing the process as such. This is normally possible since they are usually already identified. Common causes refer to inherent variation, produced by minor and natural perturbations in the system, affecting all its output.

Examples of special causes are (significant) human errors and process stoppage (for example due to data virus). Common causes in production of statistics are often linked to sampling, interviewer variability and minor variations in routines, for example due to “normal” short-term sick leave. However, long-term sick leave by key persons would be a typical special cause if not dealt with. This example reveals that there is not always a clear difference between common cause and special cause. Variability due to interviewer errors could also be inherent in the process (e.g. deficient interviewer instructions or training). A cause that repeats itself can be regarded as common, but it will increase the variation of the process. Common causes are often linked to the process itself, while special causes are normally linked to the operation of the process.

First of all, those causes that make the process unpredictable have to be eliminated, i.e. the process should be kept between limits based on variation due to what we regard as common causes. When established, these limits will serve as control limits for discovering future special causes. Reducing common cause variation is the second step in quality improvement, as it helps to improve the inherent quality (or capability) of the processes, e.g. by reducing unnecessary interviewer variability.

The main tool for determining stability and later control is control charts. There are different types of such charts. One of the most common types is a graph showing the development of some (subgroup) mean of a key process variable and its variation (range or standard deviation). A general setup is shown in ANNEX C. A few examples are mentioned in the following,

and one is given in a background chapter. More detailed considerations on control charts are given in tutorials in statistical process control, for example Bissel (1996).

6. Determine capability

As mentioned, stability does not mean that the process is good enough, i.e. that the key process variables vary within acceptable borders around a level that is considered as good enough.

For processes that are reasonably stable, one has to evaluate whether the process meets the specifications. Examples of process specifications are production deadlines, minimum response rates and maximum coefficients of variation representing for example sampling errors. Deriving process specifications from product quality requirements is not straightforward due to the complex interaction of product quality and many different processes in statistics production. Again, it is important to focus on those factors which are critical for the final product.

If the process does not meet the desired specifications, the process itself or the system behind has to be changed. A simple example is a survey with unacceptable high non-response rate. System changes may comprise changes in questionnaires or in the time schedule for fieldwork. A more fundamental change would be to base the survey on other data sources.

7. Establish a system for continuous monitoring of processes

A continuous monitoring system is needed to be sure that system capability remains high, and there should always be an objective to improve over time. A process control system that works should not be used as an excuse for not looking for changes, even if the process meets the present specifications and no new special causes occur. The system should follow the Plan-Do-Check-Act (PDCA) cycle, illustrated in ANNEX C on basic quality tools.

Experiences in statistical institutes

Even if process variables have often been measured, there are surprisingly few papers on measuring and analysing process variables in statistical institutes, despite the fact that there has been a focus on the study of processes as a basis for quality improvements during the last 10 years.

The main types used in practice are:

- Interviewer performance,
- Resources used,
- Time used,
- Response rates,
- Response burden,
- Complaints, and
- Error rates (for example in interviewing, editing and coding, disclosure control and dissemination).

Each type usually consists of several variables which are broken down by different characteristics, such as type (of non-response or error), size (for example non-response in businesses) and user group (complaints).

The handbook of Jones and Lewis (2003) is a comprehensive overview of the state-of-the-art of analysis of process variables, and gives guidance on how to identify, measure and analyse such variables. Jones and Lewis (2004) give a summary of the handbook and note that National Statistical Institutes (NSIs) have generally not collected and analysed process data in a systematic way. Sæbø (2006) also mentions that there seems to be a gap between theory and practice within this field. This might be a problem of management that (with a few exceptions) does not demand figures, but also an indication that identifying, measuring and

applying process variables in quality improvement is more difficult than what is taught in courses.

One recent example of detailed studies of process variables is found in a report on coping with decreasing response rates in Statistics Norway (Thomsen et al. 2006). This report uses Jones and Lewis (2003) as a point of departure, and studies total response rates, refusal rates, non contact rates and number of days used for data collection applying several of the techniques presented here. An example from this report is presented in ANNEX B.

Istat has used control charts for several purposes linked to monitoring interviewer's performance in Computer Assisted Telephone Interview (CATI) surveys. Murgia and Simeoni (2005) have studied the process of assisted coding of the variable "occupation" during data collection. In this case one key process variable was the percentage of the total codes assigned with the requested accuracy (number of 4 digits codes, compared to number of codes with fewer digits).

In an earlier paper, Biemer and Caspar (1994) discuss an application of measuring process variables linked to coding operations.

A report on perceived response burden by Hedlin et al. (2005) should also be mentioned. This report gives examples of process variables linked to the response burden, which is one of the issues highlighted in the European Statistics Code of Practice for official statistics. In addition to time measurements, this concerns subjective measures for response burden.

A book of Biemer and Lyberg (2003) on survey statistics contains a series of examples of analysing process data linked to the data collection process.

Recommendations for implementation

Promote measuring process variables, but start carefully (e.g. by focussing on a specific process).

Be selective, i.e. choose variables that are really important (key process variables) when starting measuring.

Try not to ignore process variables which are difficult to measure, but critical for data quality.

Suitable areas for starting the use of key process variables include fieldwork, non-response, data entry or coding, for which a number of examples are already available.

Management should demand measurements and use these or indicators based on these for their planning and follow-up.

The recommendation to start carefully and be selective is based on the experience that the area of process variables is a difficult one, where often has been a difference between theory and practice.

The recommendation on management support is crucial for ensuring continuous measurements and use of process variables.

Time and costs linked to measuring and analysing process variables may vary. In order to collect the "low hanging fruits" it is important to store what is already measured.

Measurements can be facilitated by convenient software packages that are also used for other purposes (for example planning and reporting systems, accounting systems, help-desk systems etc.).

Recommended readings

Biemer, P. and Lyberg, L. (2003): Introduction to survey quality. Hoboken, Wiley 2003.

Jones, N. and Lewis, D. (eds, with Aitken, A.; Hörngren, J. and Zilhão M. J.) (2003): Handbook on improving quality by analysis of process variables. Final Report, Eurostat.

Morganstein, D. and Marker, D. (1997): Continuous Quality Improvement in Statistical Organization, In: Survey Measurement and process Quality, Lyberg, L.; Biemer, P.;Collins, M.; de Leeuw, E.; Dippo, C.; Schwartz, N. and D. Trewin (eds), New York: Wiley, pp. 475 – 500.

Thomsen, I.; Kleven, Ø.; Zhang, L. and Wang, J. H. (2006): Coping with decreasing response rates in Statistics Norway. - Recommended practice for reducing the effect of nonresponse. Reports 2006/29, Statistics Norway 2006.

2.3 User Surveys

Definitions and objectives

Assessing the quality of data from the users' perspective is in line with the view that quality is to be decided by the user and in relation to the stated and implied needs of the user. Both the European Statistics Code of Practice and the Leadership Expert Group (LEG) on Quality Report highlight user orientation and carrying out user surveys as a way of assessing how the users perceive the quality of what they receive.

The main objective of a user survey is normally to get information on the users' perception as a basis for improvement actions. Thus they provide a valuable input to self-assessment and auditing activities which are covered in section 2.4 of this handbook.

National Statistical Institutes (NSIs) are concerned with a great number of different users, all having different expectations concerning the quality of products and services. In order to get information on the expectations and satisfaction of the different users, diverse types of user surveys can be carried out. These include:

- General ("traditional") user surveys, directed to diverse known users of products/services, as well as
- Image studies ("surveys of confidence") directed to unknown members of the general public and
- Target groups and product/service specific survey forms, respectively, like
  - Interviews with key users (important stakeholders),
  - Questionnaires added to printed publications,
  - Web questionnaires for web-users,
  - Special questionnaires for recipients of press releases, press conferences, expert meetings, training courses,
  - Special questionnaires for users of the data shop/library, etc.

Description

Two projects within the area of user surveys have been partly financed by Eurostat during the last few years:

- State-of-the-art regarding planning and carrying out Customer/User Satisfaction Surveys in NSIs (Cassel et al. 2003) and
- Measuring Customer Satisfaction, a methodological guidance (Cassel 2006).

Both these reports are natural starting points for NSIs that plan to implement user surveys or that are looking to improve already implemented surveys. In this chapter we provide some of the major findings in these reports along with information and experiences made elsewhere with the particular focus on Data quality assessment.

User surveys can be of many different kinds with different purposes, and of course the approach must be chosen with the intended purpose in mind. In table 2 eleven different types of user surveys are listed. Most of these types can be used to assess Data Quality. However, it is evident that at present the tool of user surveys is only utilised for this purpose to a very minor degree.

Cassel (2006) states that: "Satisfaction is a somewhat vague concept. A customer can be more or less satisfied with the quality of a service. Satisfaction should be seen as a continuous variable rating from "not satisfied at all" to "completely satisfied". To measure satisfaction, scales with fixed endpoints are often used. The lowest point on the scale represents the situation when a customer is not satisfied at all and the highest point the situation when a customer is completely satisfied."

This implies that one needs to recognize that we are not measuring data quality, but rather we are assessing data quality based on the perception of the user. It is therefore not enough

to know that the satisfaction level is 8 on a 10-point scale, we need to relate the perceived level to something. Given the fact that comparisons between surveys done by different NSIs at the moment are not relevant, due to the lack of common design and common questions (see ANNEX B), we need to look at the comparisons which can be done within the particular user survey or within a "package" of such surveys.

This leads to the conclusion that we can get information on how the user perceives different quality characteristics in relation to each other and, if the survey is repeated, we can see how the perception of the data quality changes over time.

The report by Cassel et al. (2003) also distinguishes between **latent variables** and **manifest variables**, where the latent variables are not directly measurable, but will have to be measured by combining measures of manifest variables. Applying this to quality of data would lead to the conclusion that Data quality is a latent variable, whereas the various components of the Quality definition are manifest variables. Hence, questions need to be formulated targeting components of Data Quality.

Table 2: Types of user surveys

Number 1	General ("Traditional") user surveys: (partly) standardized questionnaires directed to known users of products/services (including surveys that measure user satisfaction with a structural model)
Number 2	Image studies, "Surveys of confidence" (citizens, general public: opinion research, analysis of mass media)
Number 3	Registration of users' contacts/questions/complaints, by telephone, letter, fax and e-mail
Number 4	(Qualitative) Interviews (face-to-face, by post/e-mail) with key users (important stakeholders)
Number 5	Questionnaires added to printed publications (to collect data on the user satisfaction with a particular publication)
Number 6	Web questionnaires for web-users
Number 7	Special questionnaires for recipients of press releases (by post/e-mail; to collect information on the user satisfaction with press releases)
Number 8	Special questionnaires for participants in press conferences, expert meetings, training courses of the NSIs (distribution of feedback sheets during the meeting; to collect information on the user satisfaction with the respective meeting)
Number 9	Special questionnaires for users of the statistical data-bank (by post/web-based/by e-mail; to collect information on the user satisfaction with the data bank)
Number 10	Special questionnaires for users of the data shop/library (distribution to users; to collect information on the users satisfaction with supply and service)
Number 11	Special surveys focusing on "paying" users (addresses/information from "sales statistics")

All of these types of surveys can be used to assess data quality from the user perspective, perhaps with the exception of numbers 8 and 10.

Experiences in statistical institutes

According to the European Statistics Code of Practice questionnaire 62 % of NSIs in the European Statistical System (ESS) regularly carry out user surveys (see ANNEX B for more details). Thus it seems like there is a lot of experience in this field within the ESS. However, the vast majority of these surveys seem to target aspects like profile of the users, how they interact with the NSI and how they assess the personal services rendered by the NSI. Very few examples exist where user surveys have been used to assess data quality. In ANNEX B we have gathered the examples included in the report by Cassel et al. (2003). These exam-

ples are thin to say the least. Either it is difficult to use these kinds of surveys to assess data quality or the NSIs haven't put priority on doing this, but rather on other aspects of their relationship with the users. Though there might be some truth to this, it is our conclusion that this is a difficult area that requires further attention within the NSIs and the ESS as a whole.

Recommendations for implementation

Given the limited experience in assessing data quality through user surveys, strong recommendations for implementation are difficult to provide. There are, however, some issues we would like to highlight:

- Categorise users so that they correspond to categories that are already used for other purposes and that are communicated with. Preferably, the categories should correspond to clear areas of responsibility within the organisation to facilitate "ownership" of the results. For example ministries, researchers, private enterprises, media and the general public.
- Try to identify individuals as respondents – those who really use statistics – rather than representatives of organisations/functions. One should try to get a personal assessment rather than some kind of "official" view of their organisations.
- If necessary tailor the questions to the categories, taking into account that different categories normally can only reply to certain aspects of the quality concept.
  - Less experienced users might only be able to provide meaningful feedback on data quality aspects like timeliness and punctuality as well as accessibility and clarity.
  - More experienced users can be expected to also provide feedback on accuracy, comparability and coherence, while relevance seems too difficult to capture using traditional methods for user satisfaction surveys.
- Take extra care to formulate questions targeting assessment of data quality (in the broad perspective of the quality concept), given the obvious problems this has posed up until today.
- Realise that you are measuring perception, which is influenced by many factors, not some absolute value.
- Test the questionnaire for suitability.
- Choose the survey method in relation to the target group:
  - Some forms of questionnaires approach for the large number of less experienced/less frequent users,
  - Adding deep interviews as an approach for the smaller number of more experienced/more frequent users, and
  - Consider using focus groups to explore the users' assessment of data quality in more detail, relying on the interaction between the people participating to provide insight. This could provide an opportunity also for less experienced users to have an informed opinion on all aspects of data quality.
- Apply survey expertise in designing the process, but take into account that we are dealing with a very specific type of survey.
- Do not draw conclusions from the survey results beyond what is relevant from how the survey has been carried out.
- Communicate the results of the data analysis, both to respondents and internally in the NSI.
- Make sure to really use the results to take actions that will improve "quality".
- Don't ask the users for a repeated assessment before actions based upon the previous results have been taken.

The particular features of a user survey make some aspects of carrying out a “normal” survey somewhat less important:

- Having a complete frame to draw the sample from is not vital since we are not measuring totals. This is also normally not feasible from practical points of view.
- A low non-response rate is certainly better than a high one, but normally questions about data quality in a user survey are not very susceptible to bias. Especially since we usually relate them to other questions in the same survey or to the same question in earlier surveys when doing the analysis, hence, bias that is present tend to cancel itself out.
- Expect low response rates compared to the regular surveys of the NSI, but also consider every reply as unique and important.

Recommended readings

Cassel, C.; Järnbert, M.; Holzer, W.; Barbieri, G. and Tiddi, A. (2003): State-of-the-art regarding planning and carrying out Customer/User Satisfaction Surveys in NSIs. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143266/2ESS_CUSTOMER_SATISFACTION_SURVEY_SE-AT-IT_2003_EN_0_1.PDF

Cassel, C. (2006): Measuring Customer Satisfaction, a methodological guidance. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143266/CUSTOMER%20SATISFACTION%20SURVEYS_SE_2006_EN_1.PDF

Lyberg, L. et al. (2001): Summary Report from the Leadership Group (LEG) on Quality, (Chapter 4 of the Main report and chapter 3 of the Background paper). http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143266/ESS_QUALITY_RECOMMENDATIONS_2002_EN_0_1.PDF

2.4 Self-assessment and Auditing

In this chapter we will describe very powerful approaches that allow organisations to check and review their processes/products. Depending on the key references of this check and review activity, its sponsor and its organisational aspects, they may be called self-assessment, audits, or even peer reviews. It can be undertaken by internal or external experts and the timeframe can vary from days to months, depending on the scope, however the results are fairly identical – the identification of improvement opportunities in processes/products. As will be seen, these approaches constitute an important element of the Plan-Do-Check-Act (PDCA) cycle.

Self-assessment and auditing may be implemented in an organization on different stages. The easiest approach to undertake, and as a first step, may be a self-assessment exercise, where an analysis of a process/product is made against a reference model, in a more general way; a sequent approach might be internal audits, where an independent party inside the organization will review processes/products against procedures or specifications of the organization in a more in-depth assessment; and finally, audits undertaken by means of an external organization – external audits – are seen as a more formal and exigent approach, for example, made against the ISO Norms (International Organization for Standardization), with the aim to get a certification. The latter can be seen as a more final and developed approach.

The LEG on Quality Implementation Group decided that the methods for auditing and self-assessment would be part of one of the recommendations to carry out during its work in 2003, and therefore, a state of the art project was undertaken. Different types of audit activities were identified and several experiences were shared, along with the respective documentation. However, no guidance was given to the European Statistical System (ESS) concerning the auditing activity.

In this handbook, the state of the art project was taken as a basis, and the following approaches – self-assessment, auditing in general, and peer reviews as a special case of external audits – were identified as important tools for the National Statistical Institutes (NSIs). Some recommendations are provided when undertaking these activities, taking into account the recent experiences and documentations that were found.

2.4.1 Self-assessment

Definition and objectives

Self-assessment is a comprehensive, systematic and regular review of an organisation's activities and results referenced against a model/framework[6].

Self-assessment is obviously a do it yourself evaluation against a model that whoever decided to take the exercise believes that it is a good framework. The difference when compared to auditing approaches is that audits are done by a third party (internal or external to the organisation).

Description

The choice of the self-assessment tool is a strategic decision. It should be applicable to the environment of the organisation and to the processes to be assessed. Some aspects are important when undertaking these exercises:

- Team organisation – It is important that the team that will do the exercise of self-assessment should be clearly informed about the model to use and the concepts behind it. It might be important to undertake a few preparatory meetings or even some training on the model. A good communication and transparent environment is needed.

[6] Adapted from European Foundation for Quality Management (1999).

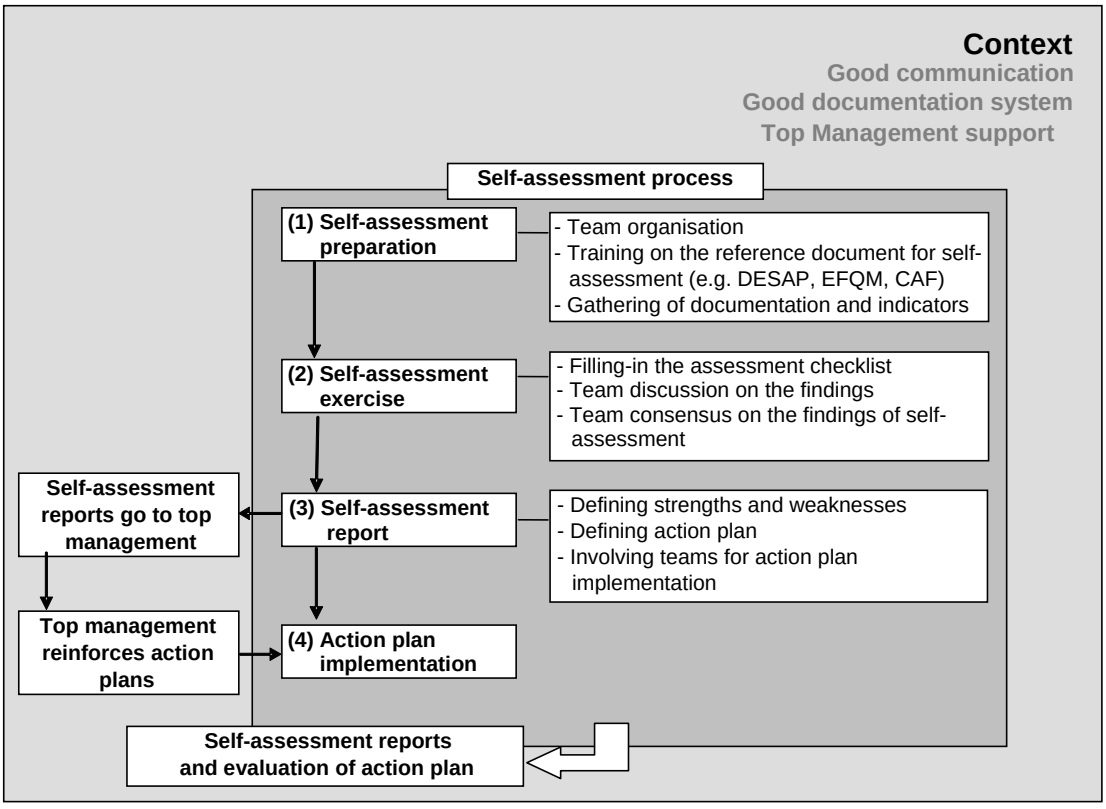
It is also important that the team believes in the model and that the effort will turn into an effective action plan;

- Preparation of the Self-assessment exercise – Self-assessment is based on evidence, the team should be organised in a way that its members will gather documentation and indicators related to the assessed issues;
- Self-assessment exercise – Once the evidences are gathered the team should discuss the state of the art of the processes that are being assessed by the model. Strengths and weaknesses are identified, scores will be marked (according to the model in use) and the team should try to reach a consensus on the main findings of the self-assessment exercise;
- Self-assessment report – The report should be structured according to the model in use for the exercise. It should stress strengths and weaknesses and should be very clear on an action plan. Teams should be involved in the implementation of the action plan, and top management should also be supportive, making available extra resources if necessary.

Experiences in statistical institutes

- Some NSIs have made the exercise of self-assessment based on the DESAP – Development of a Self Assessment Programme, The European Self Assessment Checklist for Survey Managers, 2003, and they have proved to be very efficient (ANNEX B).
- Self-assessment based on European Statistics Code of Practice Questionnaire.
- Other important tools, but at Institutional level, are the EFQM model (European Foundation for Quality Management) and the CAF (Common Assessment Framework), a brief explanation can also be seen in the background chapter.
- Self-assessment can also be made against external standards, like the ISO Norms ISO 9000 or ISO 20252.

Figure 5: Self-assessment process



2.4.2 Auditing

Definition and objectives

"An audit is a systematic, independent and documented process for obtaining audit evidence[7] and evaluating it objectively to determine the extent to which the audit criteria[8] are fulfilled"[9].

The international norm ISO 9000, from which the definition was taken, provides guidance on the audit principles; on how to manage audit programmes, and how to conduct audits to quality management and environmental management systems, as well as guidance on the competence of the respective auditors.

It assumes that auditing is a powerful tool that supports policies and management control actions by providing important information that enables an organisation to improve its performance.

This reference is also applicable to any organisation that wishes to conduct internal or external audits to its quality management and/or environmental management systems or to the simple organisation, planning and management of an audit programme.

Even though the organisation has not developed a quality management system it can also be a reference for organising audits in any organisation. As such this approach can be applicable to all processes in an organisation, and to any type of organisation, a statistical organisation is no exception.

Audits can also be conducted both internally and externally:

- Internal audits are conducted with the purpose of reviewing the quality system (policies, standards, procedures and methods) and internal objectives, and are led by a team of internal quality auditors not in charge of the process/product under review;
- External audits are conducted either by stakeholders or other parties that have interest in the organisation; by an external and independent auditing organisation, or simply by an expert concerning the process/product that will be audited (that certifies that the quality and/or environmental systems are according to the prerequisites of the ISO 9001 and ISO 14001, respectively – in this last case it might mean that the organisation is applying for a quality certification according to the ISO standards).

Both approaches have the function to control conformity with given policies, standards, procedures and methods – audit reference documents – in a systematic way and carried out regularly.

Description

The norm ISO 19011 may help organisations to organise the audit activity. But, in any case (even if the organisation doesn't take this reference into account) some issues have to be bared in mind when undertaking audits:

- Auditors (whether they are internal or external auditors) have to be recognised by the audited teams, as having the proper knowledge to undertake an audit. This means that auditors might have a specific knowledge concerning auditing, as well as concerning the activities that will be audited. So, the selection of auditors, especially when it comes to internal auditors, must be very careful, and training should be provided;
- Audits are conducted against audit reference documents, related to processes (procedures and/or working instructions) or products (specifications). Therefore, in order to perform an audit this type of documentation should be produced and

[7] Records, statements of fact or other information, which are relevant to the audit criteria and verifiable.

[8] Set of policies, procedures or requirements.

[9] In: International Organization for Standardization (2002).

implemented in the organisation, and everyone, audited teams and auditors, should be aware of them. They constitute the basis upon which audit criteria are defined;

- When conducting internal audits, the organisation should produce an internal procedure where the organisational aspects are taken into account. This procedure should make clear to everyone involved in the review process the following:
  - Audit reference documents;
  - Concepts/Definitions;
  - Audit aims;
  - Audit planning requirements;
  - Responsibilities of every part involved;
  - Audits phases;
  - Auditors selection/recruitment;
  - Auditors teams; and
  - Support documentation: audit plan; non-conformity/observation form; audit report; corrective/prevention action form; follow-up forms; checklists.
- The ultimate “client” of an audit is top management, this means that top management should define the processes/products that will be audited in a certain period of time (e.g. in an audit annual plan);
- The preparation of an audit is one of the most important issues, for both auditors and audited teams. The audited teams should know in advance that their processes/products will be reviewed by “others”, in order to prepare documentation and everything that will facilitate the audit. Good communication in the auditing process is very important, as such auditing should also be seen as an involvement activity;
- While conducting an audit, auditors should promote the discussions with the audited teams because that will allow the identification of improvement opportunities and quality improvement as a whole;
- The conclusions of an audit should be clearly summarized in an audit report, high-lighting the improvement opportunities (that sometimes are related with bad results found in the audit), but also the strong points recognised in the process/product under evaluation. The report must be accepted by the audited teams, and top management should consider it as a basis for an action plan. Audit reports should be made available to the whole team involved in the audit. If not the whole report at least a summary of the main findings should be publicly available to the whole organisation. It is also important to state that the actions to undertake are linked to the report;
- The results of an audit should be “transformed” into an action plan. Once the actions are implemented it is important not to neglect a follow-up audit, in order to assure that the planned actions are effectively implemented and that the improvement opportunities and actions for corrections are duly met;
- The auditing annual plan should also be evaluated in terms of its effectiveness. Activities must be audited on a regular basis. These two factors must be taken into account when preparing the following annual audit plan;
- Auditing is a time consuming activity. Depending on the scope of an audit, its length can vary considerably. The experiences show that in the case of internal audits, where processes are well documented, and the review does not go very deep in methodological issues, an audit can take 5 working days. If documentation is scarce and the audited teams do not prepare the audit properly, it will take longer. If an audit will take methodological issues into account, it can take months. So it depends a lot on the scope of an audit; and

- The experiences on internal auditing also show that an audit is undertaken by a team of two or three auditors, and that internal auditors do not perform audits as a full time job, normally they work in other areas of the organisation. Whilst external auditors may be professional auditors or specialists in the processes/products that are audited.

Experiences in statistical institutes[10]

Experiences found at the following NSIs concerning internal audits and/or external audits:

- Statistics Sweden;
- National Statistics Quality reviews – Office for National Statistics, UK (ONS);
- INE-PT (National Statistical Institute Portugal) Internal quality audits;
- Statistics Canada;
- U.S. Bureau of Census;
- Statistics Netherlands;
- Audits performed under the Data Quality Assessment Framework – IMF (International Monetary Fund);
- The so called "peer reviews" against the European Statistics Code of Practice, at European level. Although they are conducted by peers this approach is a case of an external audit taking as a basis the principles and indicators of the European Statistics Code of Practice; and
- Another example is the external audit against National accounts performed by external organisations in many NSIs.

Peer reviews as a special case of external audit

One possible approach of an external audit is "Peer reviews", that are often more informal, less structured and have the aim to assess at a higher level and not to control conformity with requirements item by item from a detailed checklist. Normally, peer reviews often do not concern specific aspects of data quality, but broader organisational and strategic questions.

Definition and objectives

"Peer review can be described as the systematic examination and assessment of the performance of a State by other States, with the ultimate goal of helping the reviewed State improve its policy making, adopt best practices, and comply with established standards and principles. The examination is conducted on a non-adversarial basis, and it relies heavily on mutual trust among the States involved in the review, as well as their shared confidence in the process. When peer review is undertaken in the framework of an international organisation – as is usually the case – the Secretariat of the organisation also plays an important role in supporting and stimulating the process. With these elements in place, peer review tends to create, through this reciprocal evaluation process, a system of mutual accountability"[11].

Description

Although not so formal, the methods to perform a peer review are very similar to what was already presented concerning audits in general. There are special concerns to the scope of the peer review, the peers themselves and the recommendations:

- Scope – The scope of the peer review should be clearly defined, and prepared in advance by both parties;

[10] For practical examples see the ANNEX B.

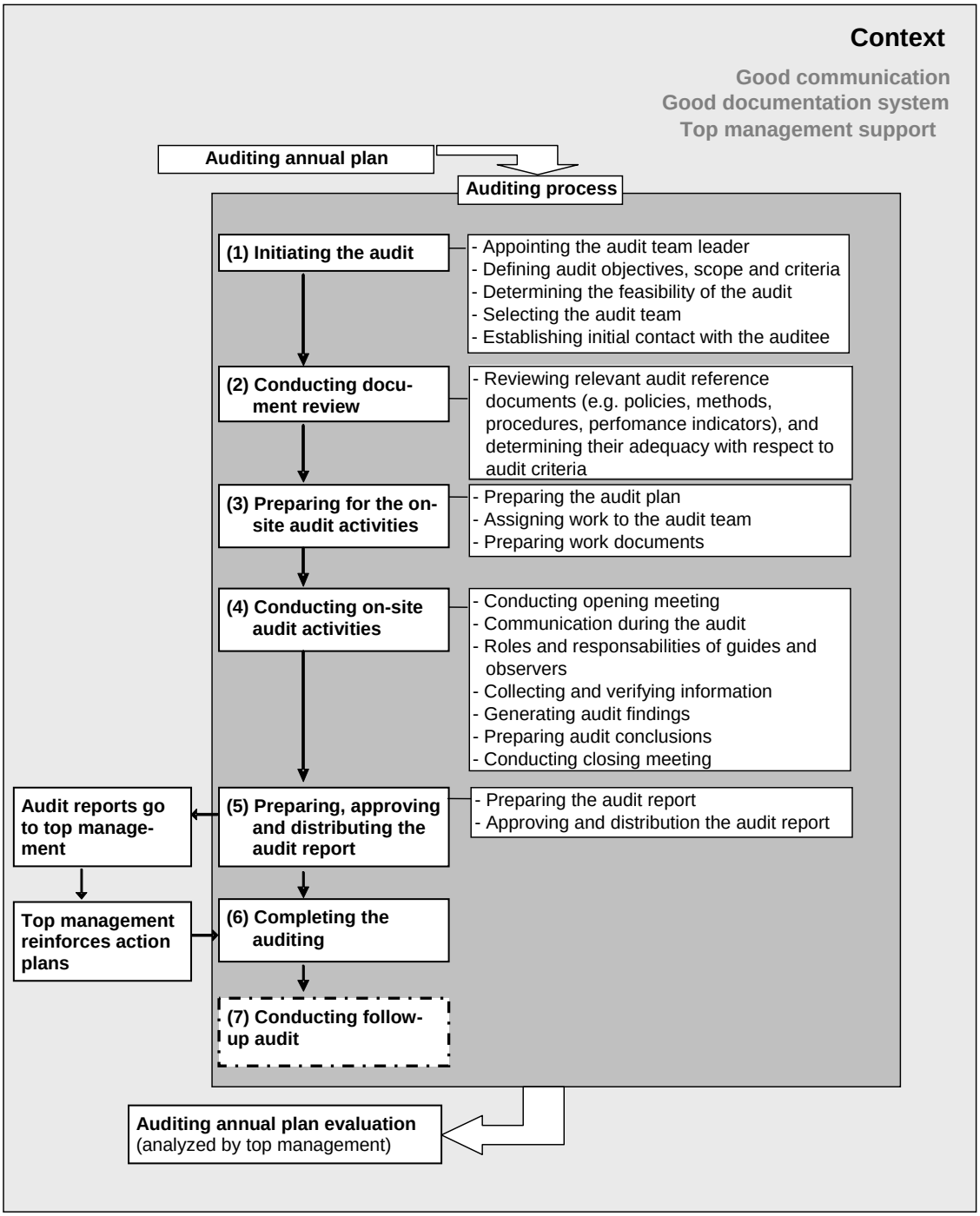
[11] In: Pagani, F. (2002).

- Peers – Peers are chosen because of their effectiveness and quality in the process to be reviewed. Therefore they must represent an excellent reference for the audited teams; and
- Recommendations – The recommendations are highly important and should be made available in the organisation and to the public as much as possible. They should represent an organisation's commitment to future actions.

Experiences in statistical institutes

- There are some examples of peer reviews that have occurred at Institutional level (INE-PT and Swiss NSI by Statistics Canada, etc.).

Figure 6: Auditing process – based on the ISO 19011:2002



Recommendations for implementation

- One of the most important preconditions is **top management support**. Top management has to promote these activities and express, very clearly, that they recognise these approaches as powerful tools to identify improvement opportunities and to continuously improve quality of processes/products. Therefore, this leadership commitment has to be very clear within an organisation when preparing any type of audit or review activity;
- **Documentation** is also an important issue, since reviews are based on evidence. In any type of review, internal or external auditors have to base their evaluation on records (e.g. performance indicators) and documents that "prove" the methods and performance of the processes/product under evaluation;
- All staff related to process/product under evaluation has to be clearly informed about the objectives and methods of the review itself. A **clear communication and the involvement of the staff are also needed**;
- And finally, the results of the review must be clearly used by the organisation, in order to show how effective the reviews can be, therefore **the results of reviews have to be consequent.**

The two approaches – self-assessment and audits – are not independent activities. Although their methods are different, they can be combined efficiently. For example, an audit can be prepared by a self-assessment exercise some months before, using DESAP.

Interlinkages with other methods

When undertaking self-assessment and audits one makes use of evidences that are related to the performance of processes/products under review. It is often the case to make use of quality indicators, quality reports, measurement of process variables and information of user surveys.

Recommended readings

Brancato, G. (2007): Review on IT and Statistical Auditing Procedures, ISTAT Italy.

Eurostat (2003i): DESAP – Development of a Self Assessment Programme. The European Self Assessment Checklist for Survey Managers.

International Monetary Fund (2003): DQAF - Data Quality Assessment Framework.

International Organization for Standardization (2002): ISO 19011:2002 - Guidelines for quality and/or environmental management systems auditing.

Zilhão, *M. J. et al. (2003)*: State of the Art Concerning Auditing Activity in National Statistical Institutes. Final Report, Eurostat, Luxembourg. http://europa.eu.int/comm/eurostat/ (product ID: G0-LEG-20030930).

3 Labelling and Certification

3.1 Labelling

Definition and objectives

As the term indicates, labelling means that a label with a message is attached to something. In this handbook context, (i) the label message is related to quality and quality assessment, and (ii) the label is attached to statistics or a provider/producer of statistics.

Fulfilment of a set of quality standards is an illustrative example of a label message. In that case the label gives the users some information about internal standards in the production of the statistics. Introduction of such a label may lead to user quality assessments that are more accurate than previously due to more information. A further possibility is that the user has a higher trust in the statistics. A credibility gain may be among the objectives of a label. Furthermore, fulfilment of quality standards has internal effects, and the labelling procedure can increase such effects. That may also be an objective.

The illustration shows that a label can have several effects: planned primary effects, and possibly also secondary effects that are more or less planned. Further labelling usages and details are described below.

Description

Labelling involves several choices, in particular the message and the subject to which the message is assigned. The wide range of choices may give the impression that the method is flexible. There are, however, several restrictions in practice. Above all, a fairly small number of labels should be used at the same time in order to avoid confusion among the users; to use many labels would be contra-productive. This implies that choices should be made with a long-term perspective in mind.

The attachment of a label needs a procedure to guarantee that the message is appropriate and true. There are two approaches with different time-scales in relationship to the labelling: ex ante and ex post. In the former case a procedure is in place first, before the statistics are produced, for example to assure that the quality standards of the label message are fulfilled. In the latter case there is a check when the statistics have been produced that they fulfil the message, which could be related to the quality level of some quality components.

The label may be brief in itself, like “official statistics”. If that is the case, explanations need to be given, for instance on the webpage, about the interpretation. Some examples are provided below. More details and further examples are given in ANNEX B.

The Office for National Statistics (ONS) is the principal provider of official statistics about the United Kingdom (UK). The UK also has “National Statistics” as an important concept. There is a description, for example on the webpage:

“The term 'National Statistics' stands for qualities such as relevance, integrity, quality, accessibility and freedom from political influence. Data deemed to be 'National Statistics' provide an up-to-date, comprehensive and meaningful description of the UK's economy and society. They are produced in accordance with the arrangements set out in the Framework for National Statistics and they comply with the professional principles and standards set out in the National Statistics Code of Practice.”

As the text shows, this is an ex ante type of procedure. The UK National Statistics Code of Practice is summarised in a Statement of Principles. There are protocols on specific topics; twelve in number now and since 2004. There is a list of producers and a list of products that comply with the national code of practice and its supporting protocols.

Sweden provides a further and somewhat different example. The Swedish Statistical System is decentralised, and 25 government authorities have responsibility for the official statistics of Sweden. These authorities decide on the content and scope within the statistical area(s) for which they are responsible. In accordance with the Official Statistics Act, the official statistics

shall be accompanied by the phrase “Official Statistics of Sweden” or the symbol shown below.

The concept “sufficient quality for official statistics” has been launched in a guideline of the system and was later clarified together with a set of criteria, as described in ANNEX B.

There is also an interest in labelling in international organisations (see Hahn and Willeke (2006) for a discussion and a summary of different aspects on labelling in general). The European Statistics Code of Practice has been mentioned together with labelling; some thoughts and initiatives are currently discussed.

An important question concerns a possible failure to fulfil the label message. This possibility should obviously be considered already when forming the labelling system. A timeliness commitment provides a simple example. A strong commitment leads to a mixture of labelled and non-labelled values in a time series, which would be unfortunate and confusing. A label should only be introduced if it can be expected to remain. It may become necessary to stop some labelling, for example to change some statistics from official to non-official statistics.

“Experimental Statistics” can also be considered as a way of labelling. These statistics are considered good enough to be published, and the publication may lead to feedback from users. They are, however, not as good as published statistics normally are. They are typically less accurate, and they may have a limited coverage. The label points out the experimental statistics as such. It may also have the further aim to protect the ordinary brand name.

The description given by New Zealand (2004) shows many of the important ingredients when working with a quality label. It discusses official statistics in a government context, differentiating and branding, a system of official statistics, an advisory committee, users and respondents etc. See also ANNEX B.

Experiences in statistical institutes

This is an advanced method, which means that the experience is somewhat limited. There have been recent discussions in relationship to the European Statistics Code of Practice. Some national institutes have introduced procedures similar to labelling but perhaps without that term. Some examples are indicated above and in ANNEX B, e.g. the UK, Sweden, and New Zealand.

Finland is here taken as an illustrative example for activities and experiences. The Advisory Board of Official Statistics of Finland has recently updated the criteria that should be fulfilled by statistics in the Official Statistics of Finland (OSF) series (Finland 2006). There are three basic criteria and five quality criteria. Moreover, the producers of OSF statistics must regularly evaluate the quality of the statistics they produce against the five quality criteria. In addition, they must assess and monitor the needs and the satisfaction of their customers, as well as the expediency of their statistics production process. Now, when the rules are introduced, two types of problems are seen. Firstly, how should “old statistics” be handled? Even if they have been checked previously it was not according to the new systematic procedure. Secondly, how should products that do currently not fulfil the new criteria be handled? Some time will be needed in some cases to achieve the new and higher ambition. Statistics Finland offers training and consultation for other producers.

A label may be related to products and/or producers. There are some similarities to certification, see the section on ISO – International Organization for Standardization – certification (chapter 3.2).

Recommendations for implementation

The method should be decided on a high organisational level: on the national level when there are several institutes or agencies involved or for the National Statistical Institute (NSI). The choice has to be careful with a long-term perspective, e.g. since many different labels

would be counter-productive. Moreover, when a label has been introduced it should be kept for some time. The message should, of course, be easy to understand for the users.

As mentioned, there is a methodological choice between commitment-in-advance and attachment-after-checking, which can briefly be called ex ante and ex post, respectively. In either case a procedure for follow-up has to be included, e.g. a checklist related to the quality standards to be followed.

It is important that those who are involved are well acquainted with the label message and well prepared for the procedures to be used.

Recommended readings

Hahn, M. and Willeke, C. (2006): Labelling of statistical information: Some introductory considerations. Presented at the Conference on Data Quality for International Organizations, Newport, Wales, United Kingdom, 27-28 April 2006.

New Zealand (2004): Official Statistics – A Recognisable and Enduring National Resource. UN/ECE CES/2004/18. This presentation by Brian Pink has been published in 2004 in the Statistical Journal of the United Nations Economic Commission for Europe.

3.2 Certification to the International Standard on Market, Opinion and Social Research (ISO 20252:2006)

Introduction

The certification to an international standard is an elaborated method of an external audit and combines this with a kind of “label” because the standard is internationally recognized as a guaranteed level of quality. Since the International Organization for Standardization (ISO) has published “ISO 20252:2006 (E) Market, opinion and social research – Vocabulary and service requirements” there is for the first time set up an international standard in the field of data quality.

The principal objective of international standardization is to facilitate business development and growth particularly between different national and regional markets. The intention of ISO 20252:2006 is to apply the principles of international quality standards to market, opinion and social research and to harmonize other national standards already available. Subsidiary objectives are to define the level of requirement for service provision and common work procedures to be applied in processes, including such across different countries.

The decision to apply ISO 20252:2006 to official statistics has to take into account the different conditions of business enterprise and public service as a survey research supplier in regard to the respective national statistical system. Official statistics complies with most of the requirements of the standards, but it may require numerous adjustments and improvements to comply with all.

The benefits of ISO 20252:2006 are:

- An external audit by an independent auditor can contribute to strengthening trust in official statistics;
- The system of ISO is transparent because the specification is openly drafted; there are no barriers to participation, it is published and required to be regularly updated;
- The audit requirement is clear and the assessment procedure specification enables an application that provides consistency and comparability for participants;
- It provides a set of standards that have an audit trail for their verification by an approved and independent third party;
- Within a statistical office the need to clearly define processes and their ownership achieves a higher level of staff involvement and comprehension of the processes;
- Quality standards, by identifying where, when, what, how and why errors occur, improve efficiency and reduce costs. Good quality proves cheaper; and
- The increased awareness of process quality brought about by standard compliance, enables staff to become more pro-active in seeking quality improvements.

The costs of ISO 20252:2006 are also to be considered:

- The introduction of some kind of Quality Management System (such as European Foundation for Quality Management (EFQM), Common Assessment Framework (CAF), Total Quality Management (TQM), Balanced Score Card) is a prerequisite;
- The certification of quality management to ISO 9001 is not necessary, but helpful;
- The costs of documentation and keeping the documentation up to date are significant; and
- It may be required to invest in changing/improving processes. The certification process itself requires some resources.

Definitions and objectives

This first document produced by TC 225[12] contains extensive terms and definitions. Participating countries were Spain, United Kingdom, Germany, Italy, France, Netherlands, Sweden, Bulgaria, and 21 non-EU countries as well as the liaison organisations European Federation of Market Research Organisations (EFAMRO), European Society of Opinion and Market Researchers (ESOMAR) and World Association of Public Opinion Researchers (WAPOR). Local standards in Germany, UK, Spain and other countries were or are going to be substituted by ISO 20252:2006. For a comparison of European Statistics Code of Practice and ISO 20252:2006 see ANNEX B.

Description

ISO 20252:2006 is a process quality standard. The full standard in translation is available from local national standards agencies. The core content is sections on:

- Quality management system requirements;
- Managing the executive elements of research;
- Data collection;
- Data management and processing; and
- Project documentation.

Experiences in statistical institutes

Having been adopted in spring 2006, the ISO 20252:2006 clearly is too new to have been object of experience of National Statistical Offices in 2006. Nevertheless the future approach can rely on multiple experiences with the equivalent local standards, which date back to the mid-nineties.

An international service industry quality standard raises a number of major implementation issues. It is critical to the credibility and success of the standard that it is implemented in a manner that results in consistent and comparable assessment and certification between countries, assessment bodies, applicant enterprises and locations. A number of alternative approaches might be pursued (for a description of one possible approach from the UK see Blyth (2006)).

Conceptually, such an approach could be adopted internationally. It would need to be outside the framework of ISO itself, but the constituent national members of the working party that drafted the standard could oversee such a process with the support of one of the liaison bodies such as ESOMAR, EFAMRO or WAPOR who support secretariats with relevant experience. Assessment bodies based in the UK who wish to carry out international inspections have been provided with a specification for inspecting to ISO 20252:2006 that has already been approved by United Kingdom Accreditation Service (UKAS). This specification has been drafted by the UK industry body and is based on the previous British standard that has already proven to be workable and reliable through time. Within the UK the ISO standard will replace the British standard completely by 2008. Given that a number of UK based companies are amongst the largest global survey research organisations it is therefore highly probable that this assessment specification will become the norm for many of the earliest accreditations around the world. This should ensure that the basis for the future growth of the standard internationally has foundations which are solid and proven.

Recommendations for implementation

Certification to ISO Standards is an advanced method/tool of process quality management. It requires documentation, quality reports, quality indicators, self-assessment and audit as mentioned elsewhere in this handbook.

---

[12] International Organization for Standardization, Technical Committee (TC) 225 for market, opinion or social research.

It is helpful, but not required to certify according to ISO 9001. On the other hand, ISO 9001 presents a general standard suitable for any type of organisation, recommended for creation of a proper process basis. An elaborated approach is: to implement ISO 9001 and to enhance it by implementation of specific requirements of other systems (ISO 20252:2006, European Statistical System (ESS) quality requirements, EFQM, etc.) (ANNEX B: Example of the Statistical Office of the Slovak Republic).

To make changes we recommend two steps:

Step I

- Review existing quality systems procedures against ISO 20252:2006;
- Identify changes needed in working practice and quality system documentation to meet new standard;
- Amend quality system documentation and implement required changes in working practice; and
- Do internal audits to ensure effectiveness of implementation.

Up to now it is sufficient to proceed up to this point, because an international coordinated assessment process is just developing.

Step II

- Make arrangements with an external assessment body to assess change.

To become ISO certified, an agency has to ask an accredited auditor to assess whether it meets all requirements, such as having a quality management system in place, keeping proper records, training of interviewers, giving feed back to employees, reporting to clients on project progress and key decisions and all other quality procedures.

Enrique Domingo, Coordinator of the international adoption of ISO 20252:2006 pointed out in an interview: “If senior management is genuinely committed to delivering good quality and service, implementation can be very fast. Still, it may take several years for a committed company starting from zero to become fully ISO certified.”

Interlinkage with other methods

Certification to ISO 20252:2006 is an advanced method which requires the implementation of all basic and intermediate methods (see chapters 2.1-2.4 and 3.1). It is a kind of external audit and all the content depicted in the respective chapter apply to it.

Recommended readings

Blyth, B. (2006): Independent, Transparent, Externally Audited: The ISO Approach to Survey Process Quality Control, European Conference on Quality in Survey Statistics (Q2006), Cardiff, United Kingdom, 24-26 April 2006.
http://www.statistics.gov.uk/events/q2006/downloads/W15_Blyth.doc

International Organization for Standardization: TC 225.
http://www.iso.ch/iso/en/stdsdevelopment/tc/tclist/TechnicalCommitteeDetailPage.TechnicalCommitteeDetail?COMMID=5613

Jackson, P. (2006): Assessment of ISO 20252.
http://www.mrs.org.uk/standards/downloads/mrqsa_conf_2006.pdf

4 Towards a Strategy for the Implementation of Data Quality Assessment

In an ideal situation, all data quality assessment methods covered by this handbook should be implemented in a statistical institute, at least to some extent. One could even argue that for full compliance with the European Statistics Code of Practice it is advisable to use as many of the assessment methods from the DatQAM toolbox as possible. At the same time, it is obvious that implementations efforts can normally only be taken step by step.

However, the various methods and approaches presented in this handbook should not be regarded isolatedly: The objective should be to build up a system in which all methods are linked to each other, and finally should fit together as a system, in which the individual components reinforce each other. Many resources can be saved by making use of the interlinkages existing between the methods. For example, experiences show that quality reports can partly serve the purpose of internal documentation, but at the same time constitute an important basis for auditing and feedback talks with users.

Thus, it is important to note that there are many interrelations between the methods. When a method is introduced one should always use elements of other methods which are already in place. For example, there is a strong link between quality reports, quality indicators and process variables on the one hand and self-assessment and auditing on the other. Thus, the implementation of DatQAM needs strategic planning. This chapter suggests a strategy for the implementation process.

In almost every statistical institute there is something already in place from which one could start. There might be some type of quality report, a few indicators or process variables in place which could be used without large efforts. The implementation of DatQAM should start from such elements available and aim at step by step standardising the efforts within a systematic and regular approach.

Implementing systematic data quality assessment is a big effort for every statistical institute. It requires a strong coordination, for example by a special unit and it is sometimes demanding, perhaps especially for subject matter statisticians and survey managers. It will only be successful if managers commit to a leadership culture which is coherent with the approach chosen.

The methods and approaches presented in chapters 2 and 3 are different regarding their complexity and preconditions for implementation. Some methods require that others have already been implemented successfully. Therefore, there is a reasonable sequence regarding implementation which should be respected. The chapter presents this sequential structure as “packages”. The “packages” aim at making sure that the different levels of data quality assessment are being taken care of as early as possible in the implementation process. Thus, already in an early stage the focus should not exclusively be on quality reporting, but there should be some kind of check according to the conformity with given standards or policies.

We start with discussing some preconditions of data quality assessment. Subsequently, we discuss three packages which are increasingly ambitious regarding both the resources required and the methodological complexity. Furthermore they symbolise a stepwise approach with increasing use of and information from the methods and approaches.

Data quality assessment, as defined in this handbook, is based upon some preconditions. These preconditions concern the (1) standards of what has to be assessed and (2) the information which is easily accessible in the assessment.

Regarding the standards, a statistical institute starting implementation should have a clear view of its strategy and objectives. It should also define what is regarded as “good” or “acceptable” quality within the statistical institute. In the beginning, this could be done in a fairly general way, e.g. in the form of quality guidelines. These quality guidelines should start from the requirements of the European Statistics Code of Practice and describe the processes by the use of which the statistical institute intends to guarantee that the work is carried out in

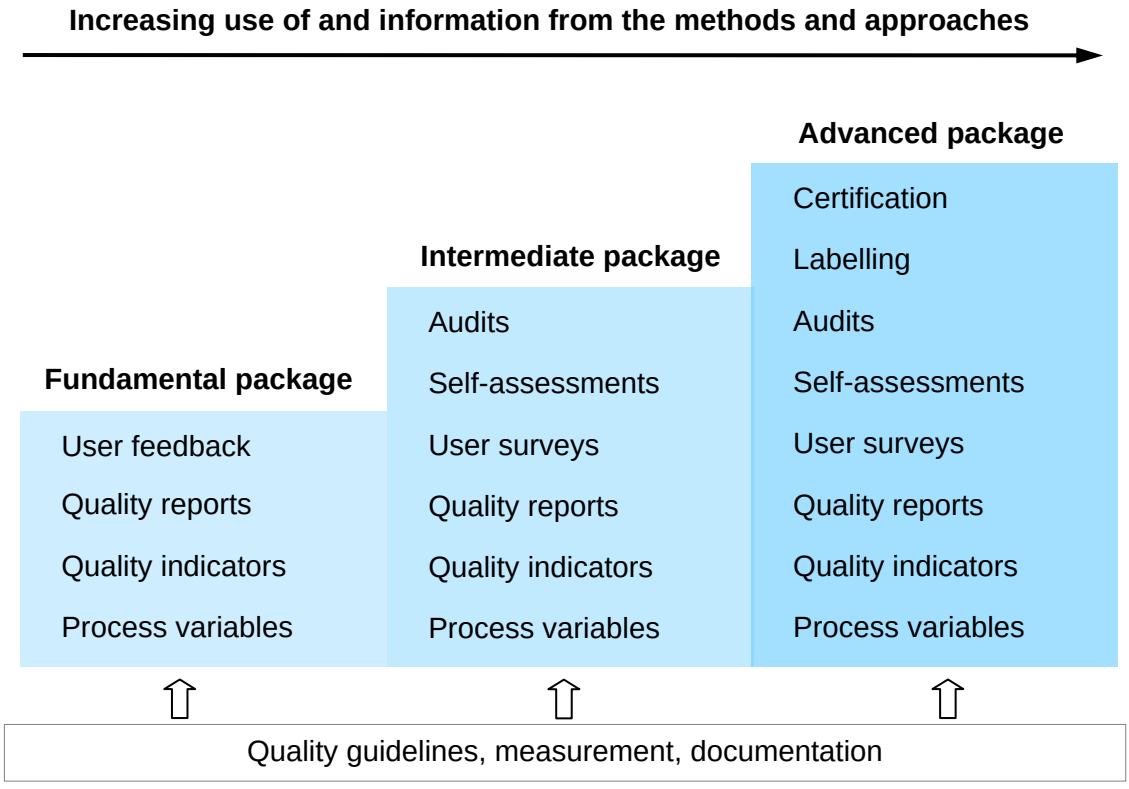
accordance with the code and meet the main user requirements. Data quality assessment will (partly) take these guidelines as a reference against which statistical products and processes can be checked.

Secondly, an assessment implies that some information is documented, which can be used for an objective control of whether there is a compliance with the standards laid down in the quality guidelines. Consequently, data quality assessment to some extent has to start from available results of quality measurement and documentation. It could, in the beginning, be largely based on expert knowledge and subjective evaluation. However, at least for some areas also objective information should be made available. Improving data quality assessment in the other packages implies both improved and successively more comprehensive measurement. At the same time some information should be available in every statistical office.

Examples include timeliness, coefficients of variations and response rates. Before starting the implementation of systematic data quality assessment, an inventory should be compiled showing which information is already available for which type of survey.

Similarly, documentation should be built up as a basis for data quality assessment. Also the documentation system should be continuously improved and completed together with the data quality management system. Similarly, metadata systems should be linked as much as possible to data quality assessment methods in order to establish an efficient system.

Figure 7: The packages of DatQAM implementation



It should also be observed that most of the methods are of a kind that they can be used on differently ambitious levels. A typical example is provided by process variables, where the number, the coverage of the production process, the information depth, and the analysis can vary considerably, as described previously.

Figure 7 shows, as mentioned before, that quality guidelines, measurement and documentation are preconditions for quality assessment. It furthermore suggests a strategy for the implementation of quality assessment methods. The successive development is symbolised by the three different packages and the arrow above them. The fundamental package includes process variables, quality indicators, quality reports and user feedbacks on a basic level. The following intermediate package contains self-assessment and audit as new methods. Furthermore the methods already introduced in the fundamental package are now more developed and more used. This development continues and leads to the advanced package. At this stage the NSIs also may use labelling and certification to communicate their quality.

4.1 The Fundamental Package

We recommend to start with basic standardised quality reports. In order to facilitate the implementation, the reports should follow a standardised structure but be limited in contents and level of detail. The primary objective should be to inform users on data quality and methodology which indicates that the reports are never just for internal use. More detailed quality reports for documentation purposes should follow at a later stage, if no other appropriate documentation system is in place. Concerning the contents, quality reports should include a brief description of use and users of the survey, the survey concepts, the key methodological features (like sampling, data collection techniques, weighting) and a quality assessment according to the European Statistical System (ESS) quality components. The quality assessment should use quantitative indicators wherever available, but could also be based on a general expert appraisal in order not to make the implementation of the reports too burdensome. Such basic quality reports could be expanded step-by-step by using more quantitative quality indicators, successively more process variables, and technical details which might then be of greater interest internally. Examples for quality reports can be found in Slovenia, Finland, Austria, Germany and Norway. The information contained in quality reports stem from quality indicators and process variables. Although they are neither fully developed nor very many in number there should be some quality indicators and process variables.

In the fundamental package, the statistical institute should make use of the process variables already in place or easily to be made use of. As the available experiences show, in some areas, such as data collection and fieldwork, process variables can be used rather easily. Despite the fact that a full system of process variables is difficult to achieve, it is nevertheless important to start using these possibilities at an early stage as well as on a basic level.

Finally, the fundamental package should include at least some kind of user feedback. It will often be hardly feasible to carry out user surveys covering the data quality aspects of various statistical products at this early stage. However, use could be made of a simple feedback obtained from statistical advisory committees or in-depth interviews with key users. Such approach could later be expanded and supplemented by regular user surveys.

Although the fundamental package is limited in scope at first sight, it provides the statistical institutes with key elements of data quality assessment and covers the most relevant elements of data quality assessment. It should be noted that, despite the characterisation of the package as “fundamental”, the conceptual effort and the staff resources needed for implementing this package should not be underestimated.

4.2 The Intermediate Package

The methods assembled in the fundamental package already cover the main parts of data quality assessment. Although these elements are on a low state of development they enable statistical institutes to get a structured and systematic overview regarding the entire set of quality components for all statistical products. The methods of the intermediate package are necessary in order to validate the assessment (largely based on expert appraisal in the fundamental package) with the help of objective information as well as an evaluation by (external or internal) experts in audits and by users in user surveys. Besides the evaluation, the methods introduced in the fundamental package should be expanded and further refined. Quality reports should become more and more detailed. Further quality indicators should be introduced as a standard, and a more extensive use of an expanded set of process variables should be made. Figure 7 indicates the expansions by the arrow, and it shows the additional methods in comparison with the previous package.

Despite quality assessments for the various quality components, quality reports are largely descriptive in character. Therefore elements of an evaluation are being needed for a validation. The appropriate methods are self-assessment and auditing.

Self-assessment is the appropriate method to start implementation: Self-assessments are easy to implement and have a low burden for the survey statisticians. Furthermore, with the DESAP – Development of a Self Assessment Programme – checklist a standardised tool is available in the European Statistical System (ESS). Self-assessment can serve several objectives. First of all, the survey manager gets a structured overview regarding the survey. The self-assessment provides a simple “quality profile” of his or her survey and he or she can systematically identify improvement actions. Secondly, the results from the self-assessment could be stored centrally in a data base, providing the statistical institute with a (rough) basic data quality documentation (which could e.g. facilitate future self-assessment in the context of the European Statistics Code of Practice). Thirdly, self-assessment can easily be supplemented with elements of internal or external feedback, thus taking already a step in the direction of quality audits. Feedback could be introduced by discussing the assessment with experts form other units or even reaching a common view on the assessment results with experts from other units or the quality management unit.

The results from the self-assessment should have an external validation as well. For this purpose, it is recommended to introduce audits, which should be based on the results from the self-assessment. The effort associated with auditing in terms of coordination and staff resources should not be underestimated. Even more, in contrast with the other methods discussed so far, auditing might require cultural changes. The selection, training and deployment of the internal auditors should be done very carefully and be adapted to the organisational culture of the statistical institute. Again, the implementation should start with two or three pilot audits, after which the methodology should be fine-tuned. In contrast to self-assessment, it is not realistic to introduce auditing for all statistics at the same time. Even with a well equipped audit secretariat and motivated team of auditors, it is difficult to carry out more than 5 to 10 audits per year and team. Therefore, we recommend a rolling approach which covers different statistics every year. A good practice example could be found at Statistics Sweden (see ANNEX B). Besides, even the fact that an audit could be carried out will contribute to an improvement of the results from the self-assessment. Introducing external auditors increases the effort needed considerably. Except for justified individual cases, external auditors should not be used in an early stage of the implementation.

Another aspect of external validation is to measure the user perception of the statistical product. Information on the user perception can never be substituted by expert (producer) appraisal, but requires asking the users themselves for their opinion. After the more informal user feedback in the fundamental package, now simple, but systematic user surveys should be introduced. Unfortunately, the experiences available with user surveys do only cover data quality assessment to a minor degree. So far many user perception surveys are not tailored to the needs of a specific survey. Furthermore, some of the quality components are difficult

to assess in a user survey, as a reliable assessment requires a certain degree of expert knowledge (e.g. accuracy, coherence, whereas others could also be addressed to non-expert users like timeliness). A suitable starting point might be to address some more general quality aspects to a larger group of (also non-expert) users. More technical aspects are still more likely to be discussed with more experienced users, e.g. the representatives in sub-committees of a statistical council in charge of specific areas of surveys. For those groups, qualitative expert interviews as well as standardised questionnaires could be applied. However, a standard approach is not available so far.

In an early stage of implementation, quality reports can in many parts rely upon a subjective appraisal of the survey manager in charge of a survey. Such information will in many cases be more understandable for external users than quantitative indicators. Nevertheless, such a subjective assessment risks to be biased and can hardly be standardised completely. Therefore quality reports need to be supplemented by explicitly defined (quantitative) quality indicators. In a first step, those indicators could be included which are already available. In the intermediate package, the objective is to standardise the use of quality indicators and implement them for all surveys. In the ESS, a set of standard quality indicators has been developed which should serve as a guideline. According to the specific context further country-specific or survey-specific indicators should be added. At least for some of the indicators, the implementation will face some obstacles. The information required might not be directly available from the production process (e.g. in case of the editing and imputation rates) or the standard indicators themselves are still labelled “for further development”, which indicates that the implementation is not straightforward, at least not for all various types of surveys. It is therefore recommended, to start with a pilot implementation in a representative sample of surveys. Such a pilot phase could be used e.g. to identify necessary changes in the production process with regard to the calculation of the indicators.

4.3 The Advanced Package

The advanced package brings the methods previously introduced into a high stage of development. It furthermore introduces labelling and certification, which are different in character from the previous methods. Again, figure 7 illustrates the improvements.

Any substantial quality improvement will necessitate changes in the production processes. Data quality assessment should therefore also take into account the processes. Process quality is normally at least in part covered by self-assessments and audits. Continuous process improvement in addition requires comprehensive measurement of the performance of the processes. This handbook recommends to use key process variables for an assessment of process quality. Key process variables should be conceived together with the quality indicators. The difference is that key process variables go much more into the detail of the processes whereas quality indicators are more directed to towards product quality. The systematic implementation of process variables is an ambitious project and depends on a number of preconditions. For this reason the systematic measurement and analysis of process variables is part of the advanced package.

In order to make systematic use of process variables, the process flow has to be documented in flowcharts, and meaningful variables have to be identified. Processes and Information Technology (IT) systems often have to be modified in order to facilitate the measurement of the variables. If a standard set of key process variables has to be used in the statistical institute at least a certain degree of process standardisation is necessary. Many of these aspects might interfere quite fundamentally with the habits and the culture in an organisation. The systematic use of key process variables, for this reason, should be envisaged after a successful implementation of the methods of the fundamental and the intermediate package. Furthermore, implementation should start in pilot projects and be restricted to selected processes for which experiences and practical examples are already available (like fieldwork, coding or data entry).

Labelling and certification according to ISO 20252 are no data quality assessment in the strict sense, but special ways of communicating quality standards to the users. Their implementation can only start if the methods presented in chapter 2 of this handbook are in place and are being used as a standard. Consequently, the focus regarding labelling and certification is on proving that the statistical institute meets certain standards. In case of labelling, this standard will normally be (co-)defined by the statistical institute; in case of certification according to ISO 20252, standard is given externally and compliance has to be proven to an external auditor.

4.4 Recommendations

Data quality assessment should be implemented in a step-wise approach. The packages identified above suggest a general implementation strategy. However, the implementation has to be tailored according to the institutional and cultural context of the statistical institute. Another aspect is, of course, the methods already implemented in a statistical institute. This can also necessitate to choose a different sequence of the implementation of the methods. The following general recommendations should be kept in mind in the adaptation to the requirements of a specific statistical institute.

- Top management commitment is vital. This commitment to the implementation should be explicit and unmistakable. The consequences of the decision to use certain methods have to be communicated clearly to top management. Systematic data quality assessment can never be achieved without full top management support.
- The most effective way of top management support is that managers demand results from the assessment and use these results for their decisions.
- The role of the middle management should not be underestimated. If the middle managers are not committed, a successful implementation is very difficult to achieve. Special incentives are therefore needed in order to reach middle managers.
- Data quality assessment is a long term project. The main problem is not to start systematic data quality assessment, but to keep the process going on. Regularly, new incentives are needed for sustained motivation. Quick success should be promoted internally as well as externally, and the staff should be made aware of the progress reached so far.
- Start as simple as possible and progress step-by-step. Be aware of the organisational culture in place. Start as early as possible with concrete actions and the data quality assessment sufficiently engaging.
- Most methods should be implemented and fine-tuned in pilot projects. One option is to start with a sub-sample of surveys or thematic areas only. The sub-sample should be selected in a way that makes the start easy and makes it probable that quick successes could be achieved which motivate both the quality team and the subject matter statisticians for further developments.
- For some methods a partial application makes the start of the implementation easier. For example, key process variables should be developed first for a limited set of processes, like coding or fieldwork, before possibly extending the methods to a larger set of processes.
- Standardise the use of the methods (e.g. standard quality indicators, standard process variables which could be supplemented by additional information). This handbook names standard tools at least for some of the methods discussed.
- From the very beginning clear responsibilities and authorities should be established. The quality manager should have sufficient resources at his or her disposition.

ANNEX A:
General Framework of Data Quality Assessment

General Framework of Data Quality Assessment

1 Introduction

In the European Statistical System (ESS), in recent years, much work has been going on in the field of quality management. Various quality management models and frameworks (like the EFQM model (European Foundation for Quality Management), ISO 9000, the Data Quality Assessment Framework (DQAF) or the European Statistics Code of Practice) have been advocated for the use in ESS member institutions. Remarkable implementation efforts have been undertaken. However, an efficient implementation has been partly impeded by a number of problems:

- The existing approaches are partly overlapping which created confusion among managers and staff and led to parallel implementations.
- The models set requirements on a fairly general level and thus lack guidance on how to implement them in the concrete processes. This contributed to a general feeling that there is a gap between theory and practice (Sæbø 2006; Statistics Sweden 2006a).
- The models cover a multitude of different topics. This might lead to the impression that everything has to be launched at the same time. A further risk is that under a broad Total Quality Management (TQM) perspective the core of the statistical production processes gets overlooked and does not receive the appropriate attention. The report on the ESS self-assessment against the European Statistics Code of Practice points in this direction and suggests that quality assurance in the production processes is not very well developed in most National Statistical Institutes (NSIs) (Eurostat 2006c).

The situation is similar in the field of data quality assessment. Stimulated by the work of the Leadership Expert Group (LEG) on Quality, much development work has been done. However, concrete implementations of these methods are lacking in many NSIs.

2 The context of Data Quality Assessment

Data quality assessment methods and tools have to be integrated into a quality management system. Thus, data quality assessment is one element in the broader context of TQM. Accordingly, before the methods and tools are described, the general quality management context is outlined in which these methods and tools are being applied. Generally, it can be noticed that notions like TQM or systematic quality work themselves risk to be used as buzzwords. To overcome this problem, first of all this chapter tries to prepare some conceptual common ground in the quality field.

Is there a “coherent quality systematics” that one can refer to? Probably not. When reviewing current texts on quality issues on a broad basis, one notices a lot of key words for systematic approaches such as ‘standard’, ‘framework’, ‘model’, ‘concept’, ‘code’, ‘guideline’, ‘system’ in which ‘principles’, ‘recommendations’, ‘dimensions’ and ‘criteria’ are described. These key words are often synonymous in use or there are several and overlapping meanings for one and the same term.

As a basis for the presentation of the data quality assessment methods and tools, the DatQAM handbook should, in a first step, aim at reaching a consensus regarding the conceptual framework of quality work in statistical agencies. In order to implement data quality assessment methods one needs to understand their position in a broader framework of systematic quality work.

Recent literature criticises the “multitude of overlapping quality frameworks” (Statistics Sweden 2006a) and notes that the complexity of multidimensional quality approaches can have serious drawbacks for the improvement of quality work (Sæbø 2006). Accordingly, a kind of “quality taxonomy” is being developed (chapter 2 of this annex, where the key words given above are highlighted to help the reader to associate them in the “taxonomy”) to give guidance not only on the methods and tools themselves (chapter 3 of this annex) but also to define their use in the various approaches of quality improvement.

The frame of this “quality taxonomy” is considered here to be the approach of TQM. Although TQM will not be treated in-depth in this handbook, it is important as a context for data quality assessment. It should be noted that several NSIs, recognising a lack of acceptance for the management jargon switched from using the term TQM to the term “systematic quality work”, which is however used largely synonymously (Bergdahl and Lyberg 2004).

Besides TQM, there are a number of institutional frameworks which constitute an important background for data quality assessment. In the ESS, the European Statistics Code of Practice is of particular importance and stresses that systematic data quality assessment is vital for providing quality statistics.

2.1 TQM Models

TQM is a management philosophy that addresses processes in a systematic way for organisations as a whole (figure 8). As defined by the Deming Prize Committee of the Union of Japanese Scientists and Engineers (JUSE), TQM is “a set of systematic activities carried out by the entire organisation to effectively and efficiently achieve company objectives so as to provide products and services with a level of quality that satisfies customers, at the appropriate time and price” (The Deming Prize Committee 2006).

TQM in Japan comprises four process steps, namely:

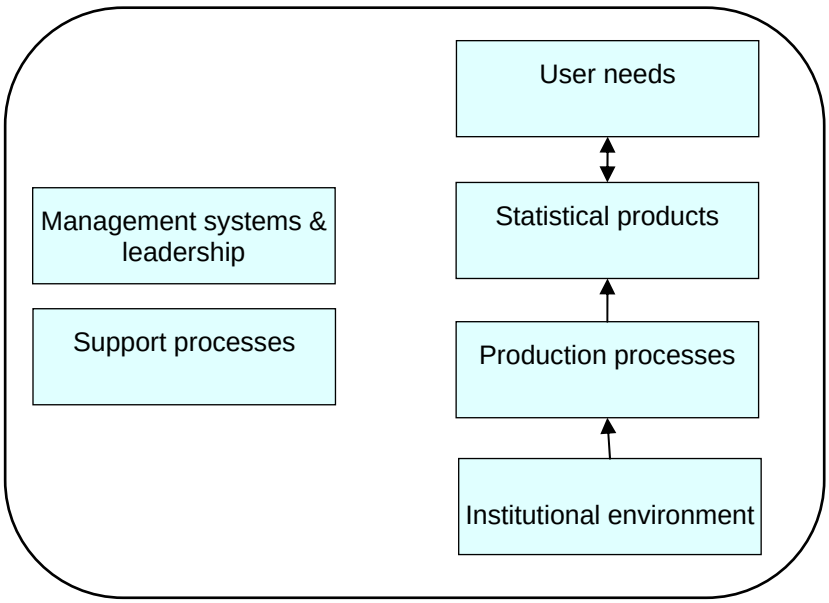
- Kaizen – Continuous Process Improvement, making processes visible, repeatable and measurable.
- Atarimae Hinshitsu – Focusing on intangible effects on processes and ways to optimise and reduce their effects.
- Kansei – Examining the way the user applies the product leads to improvement in the product itself.
- Miryokuteki Hinshisu – Broadens management concern beyond the immediate product.

TQM requires that the company maintains this quality philosophy in all aspects of its activity, including not only production processes, but also management and support processes.

TQM models (sometimes also referred to as QM systems or QM mainframes) facilitate a systematic application of the TQM philosophy by defining areas of activity and criteria for the practical application. Such models comprise the EFQM model, the international standard ISO 9001, the Malcolm Baldrige National Quality Award, the Deming Prize, Balanced Scorecard etc. (Eurostat 2002; Tague 2005). The most important elements of TQM models have been applied to the ESS in the Quality Declaration of the ESS (Eurostat 2001).

Such models define areas which have to be taken into consideration in quality management. Figure 8 summarises the scope of TQM in a simplified model. The most important point of reference is the use made of the final product (user needs). Product characteristics and the design of the production processes have to be streamlined according to these requirements (in terms of product quality, time, and price). TQM models also have a systematic look at factors which determine products and processes more indirectly: Leadership (including policy and also cultural aspects), management systems (e.g. corporate planning) and support processes (partnerships, financial management, human resource management etc.). Whereas all these elements are extensively covered in the focus of TQM models, institutional aspects (like the political and legal framework) are normally regarded as external constraints, given that they are not under direct control of the organisation. They are, however, very important in the context of official statistics as they provide the fundament for official statistics (professional independence, legal mandate for data collection, data protection etc.).

Figure 8: Scope of TQM/systematic quality work



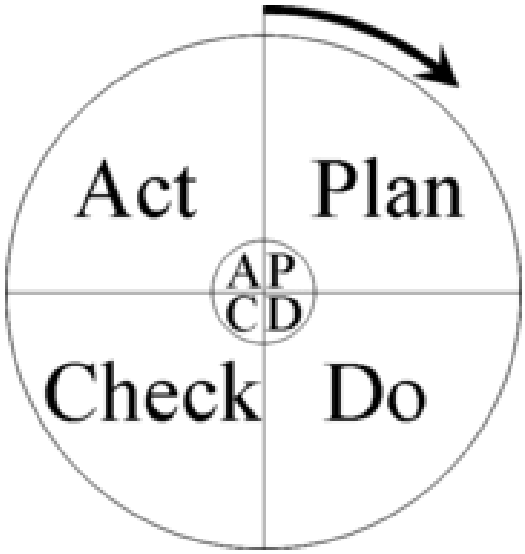
The strategic core of all major TQM models is continuous improvement, often illustrated with reference to the so-called Plan-Do-Check-Act (PDCA) cycle made popular by Deming (figure 9) or a similar concept[13]. This cycle is a four-step process which has to guide all changes for continuous improvement. It implies that any improvement needs to be planned (P) and needs to be tested subsequently (D). However, it is inevitable for the improvement that the test is then evaluated or assessed (C). Without this assessment it would remain unknown if the respective change actually improves a process or whether not. Only on the basis of this assessment it is finally decided if and how process steps are changed (A). As Bergdahl and Lyberg (2004) note, the PDCA cycle is "often degenerated so that it consists of P[lanning] and D[oing] only". If handled in such a way, processes continue until problems become highly visible which in turn leads to a highly inefficient trial-and-error in process design.

The cycle implies that the improvement process is not timely restricted but needs to be followed-up continuously. If the assessment of any test is not positive another plan needs to be prepared but even if the process is improved there is continuous change in the requirements (e.g. from other processes, change in the institutional stetting, according to research progresses) according to which the process needs to be continuously adjusted.

With respect to the aim of this handbook it is clear that the "C" the evaluation or assessment part of the cycle will be discussed. It should be noted that – looking at figure 9 – there is not only one PDCA cycle but there are numerous PDCA cycles for any of the TQM areas and the institutional aspects.

[13] In Six Sigma programs, e.g. this cycle is referred to as "Define, Measure, Analyze, Improve, Control" (DMAIC). In the EFQM context, the idea behind the so called RADAR logic (Results, Approach, Deployment, Assessment, Review) is much similar to the PDCA cycle.

Figure 9: The PDCA cycle (or Shewart cycle)



Although quality assessment plays a major role in all aspects of TQM, this handbook focuses on the methods and tools for assessing the quality of statistical products and the respective production processes. A tentative definition of quality assessment that holds for all three levels would be:

Data quality assessment is an important part of the overall quality management system of a statistical agency. However, its scope is limited to the statistical products and the processes leading to their production.

With respect to quality work in official statistics TQM models introduce the idea of systematic, holistic approaches to assess processes. Comparing contents and scope of TQM models with the tasks of data quality assessment – which is the quality assessment of official statistics – there are differences in a number of respects:

- the concern of TQM is not only on product characteristics and production processes, but also on support and management processes
- TQM stresses the importance of cultural (e.g. corporate identity) and motivational (e.g. leadership) aspects of business performance
- TQM is strictly user oriented (whereas data quality assessment in many cases assumes a more engineer-centred perspective)

Compared to TQM, data quality assessments focus less on the user requirements, management systems as well as on cultural factors.

TQM models concern the implementation of quality management on a meta-level. For example, the EFQM model focusses very much on the management systems within an organisation but gives only little concrete guidelines on how to improve a specific process or product. Similarly, ISO 9001 describes quality management requirements in a quite abstract way which fits in many different fields of economic activity. Accordingly, in the context of official statistics a further specification of these models is needed. Quality improvement at the organisational level takes into account the whole organisation and is reflected in so called institutional frameworks, while quality assurance frameworks provide guidelines for improvements on the product and process level.

2.2 Institutional Frameworks

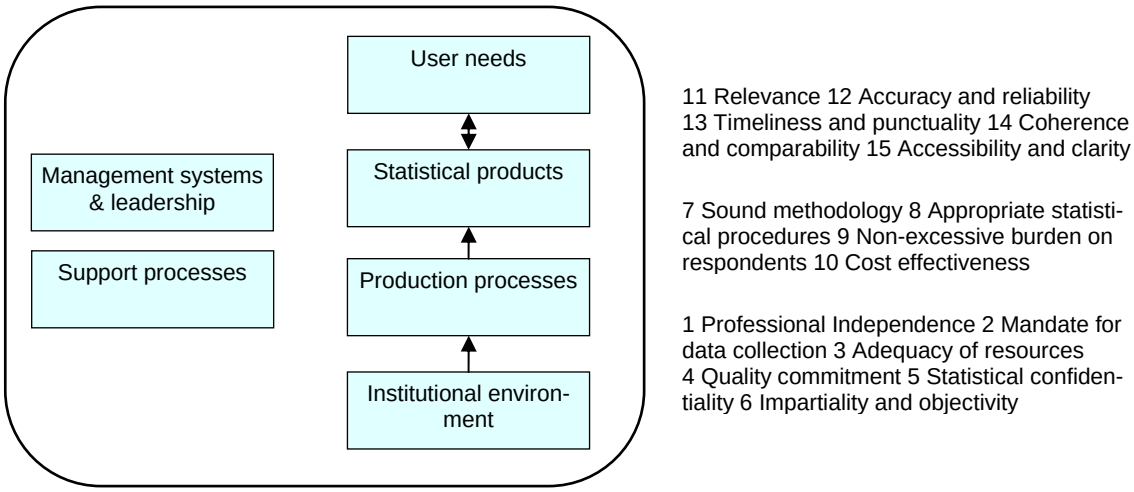
Partly overlapping with TQM models, a number of institutional frameworks have been developed in recent years. The major objective of these institutional frameworks is to guarantee a certain number of minimum requirements, mostly in an international context. These minimum requirements in first instance concern basic institutional features, like the professional independence, the legal mandate for data collection or the measures taken to guarantee statistical confidentiality. Besides such institutional aspects, often further aspects concerning statistical products and statistical processes are dealt with in some detail. Examples of such institutional frameworks are the DQAF of the International Monetary Fund (IMF), the United Nations Fundamental Principles, the Principles Governing International Statistical Activities (UN Statistics Commission 2005) and the European Statistics Code of Practice.

In the ESS, the European Statistics Code of Practice is the most important institutional framework, which has quite detailed principles and indicators in the field of data quality assessment.

Although institutional frameworks focus on minimum standards regarding institutional aspects and (partly) regarding statistical processes and products, inspiration via TQM is also obvious. Nevertheless institutional frameworks do not cover the full range of TQM models. In contrast, they deal with the institutional aspects quite extensively (e.g. the European Statistics Code of Practice is very much influenced by the original idea of dealing with institutional aspects) and in focussing on product and process quality from a less management oriented, more institutional oriented point of view. On the other hand, institutional frameworks differ from the product and process level in covering various statistics and in focussing on the quality of statistical systems as a whole; less on individual statistics which the product and process level do. Figure 10 summarises the scope of institutional frameworks and shows the relating principles of the European Statistics Code of Practice.

The European Statistics Code of Practice consists of 15 principles and 77 indicators. These principles are partly inspired by TQM and partly cover the nine EFQM quality criteria (e.g. leadership, policy and strategy, people, people results and society results). They also overlap with the ESS quality declaration but the Code of Practice describes a minimum standard whereas the quality declaration is rather a vision. The indicators have been developed for a periodic review of implementation in the ESS, based on a sequence of self-assessments, peer reviews and reports.

Figure 10: Scope of institutional frameworks and the principles of the European Statistics Code of Practice



Generally, the principles in institutional frameworks, like the European Statistics Code of Practice, can be seen as a general superstructure to all other measures which are later de-

scribed on the product and process level. They aim at supporting improvement of quality in the organisations as well as enhancing the credibility of the outputs via defining and assessing performance indicators. Special emphasis lies on the assessment of statistical systems and their positive development for international (cooperative) purposes.

The principles integrate a considerable number of (if not all) indicators of statistical products and processes and thus cover the aspects of product and process quality nearly completely. In theory, all products and processes can be synthesised to a general picture of the respective organisation or statistical system. Therefore it should be possible to assign nearly all aspects of process and product quality to one or more principles in the institutional framework, although they are not always mentioned explicitly.

Data quality assessment is mentioned in various principles and indicators of the European Statistics Code of Practice. Numerous principles and indicators are only conceivable under the precondition that coherent and well implemented approaches towards data quality assessment are in place. For example, according to principle 7, “adequate tools, procedures and expertise” have to be applied in order to achieve “sound methodology”. In order to figure out whether a statistical procedure could be referred to as “adequate”, generally accepted assessment methods and tools are required. According to the European Statistics Code of Practice, data quality assessment methods and tools have to provide a complete picture including all steps in the survey process:

- According to principle 4, the complete production process as well as the entirety of product quality dimensions have to be regularly monitored: it is required that “product quality is regularly monitored”, “processes are in place to monitor the quality of the collection, processing and dissemination of statistics” and that there is a “regular and thorough review of the key statistical outputs”.
- Monitoring and reviewing survey and sample design as well as fieldwork, data entry, coding, editing and imputation constitute the core of principle 8.
- Principle 11 requires monitoring relevance.
- Principle 12 requires an assessment and validation of the source data as well as statistical outputs, and an assessment of sampling and non-sampling errors.

While there should be only one figure per indicator when conducting an assessment on the institutional framework level (assessing the institution as a whole), there are several figures per indicator when assessing data quality on the product and process level (at least one number per indicator for every statistic).

This handbook does, however, not focus on quality assessments of statistical systems. The main reason is that any assessment on the organisational level requires a synthesis that goes far beyond the scope of recommending data quality assessment methods. Nevertheless, the area of assessments on the organisational level might be touched in the context of assessing coherence of individual surveys.

2.3 Quality Assurance Frameworks

Quality assurance frameworks (or frameworks for statistics production) have the objective to establish, in a specific statistical organisation, a system of coordinated methods and tools guaranteeing the adherence to minimum requirements concerning the statistical processes and products. Similarly to institutional frameworks, this includes some kind of assessment. The difference between quality assurance frameworks and institutional frameworks is, however, that quality assurance frameworks are mostly applied at the level of individual statistics and that there is a clear concentration on the quality assessment of individual data sets and processes that lead to their generation (not a synthesis!). In contrast, institutional frameworks tend to focus on the quality of statistical systems as a whole.

According to a general definition, quality assurance is referred to as “the part of quality management focused on providing confidence that quality requirements will be fulfilled” (ISO 9000:2005). Consequently, quality assurance comprises all measures that make sure that

- Product quality requirements are being explicitly documented
- Processes are defined and made known to all staff
- The correct implementation of the processes is monitored on a regular basis
- Product and process quality are continuously monitored and documented
- Users are being informed on the quality of the products and possible deficits
- A procedure is implemented that guarantees that the necessary improvement measures are being planned, implemented and evaluated.

A quality assurance framework defines a set of concrete measures (e.g. periodic reviews, audits, quality documentation etc.) and how these general objectives should be achieved within a given organisation.

Quality assurance should not be confused with quality control, which is limited to controlling whether the products meet the quality requirements. Quality assurance, in contrast, means a regular overall evaluation of the production performance: “Under quality control, the prime purpose is to serve those who are directly responsible for conducting operations – to help them regulate current operations. Under quality assurance, the prime purpose is to serve those who are not directly responsible for conducting operations but who have a need to know – to be informed as to the state of affairs, and hopefully, to be assured that all is well” (Juran 1999).

It is worth noting that in the literature terms like model or system are sometimes used synonymously to the term quality framework or quality assurance frameworks. The use of several terms in parallel might already indicate that “quality assurance framework” is difficult to define and indeed includes many (compatible[14]) components, which might not always be restricted clearly to the product and process level but tackle the organisational level as well, e.g. strategies and systems for measuring and reporting product quality, corporate planning, identification of current best methods, developing user-producer dialogue, standardised processes, review approaches, training and staff perception studies.

Quality assurance frameworks have already been developed by several international organisations as well as NSIs. Examples include Statistics Canada (2002), Statistics Sweden (Bergdahl and Lyberg 2004), the Office for National Statistics of the United Kingdom (ONS), Statistics Norway (Sæbø 2003), the U.S. Census Bureau (Bushery 2004), Eurostat, the Organization of Economic Cooperation and Development (OECD 2003), and the Australian Bureau of Statistics.

As noted by Lindén (2006), quality assurance frameworks encompass the definition of quality requirements for statistical processes and products (as well as some institutional features) and the related quality assessment methods and tools which are to be applied in the organisation. In some cases the concrete quality requirements for products and processes are additionally documented in separate quality guidelines (e.g. Statistics Canada 2003, Statistics Finland 2003, Statistisches Bundesamt 2006). The data quality assessment methods and tools to be covered in this handbook are therefore an important feature of a data quality assurance framework, but only part of it.

The International Standard ISO 20252 can be regarded as a general quality assurance framework for market and opinion research (which does however not cover the full range of activities of official statistics).

To illustrate the scope of quality assurance frameworks, two examples are given here: the Statistics Canada Quality Assurance Framework and the Statistics Sweden Quality Assurance Framework.

[14] Although most components for a quality framework might be compatible there is the danger of overlaps as it will be shown below.

The Statistics Canada Quality Assurance Framework is structured according to the Statistics Canada quality criteria. It includes the following methods and tools, each further defined in specific documents:

- Managing Relevance
  - Client and Stakeholder Feedback Mechanisms
  - Programme Review
  - Data Analysis
  - Long-term and annual planning processes

- Managing Accuracy
  - Programme Design
  - Implementation (mainly monitoring of production processes)
  - Accuracy Assessment
  - Various Independent Reviews
    - Internal technical review committees for major programmes;
    - Referral of issues of technical standards, or general methods or approaches to the Methods and Standards Committee;
    - Referral of technical issues to the Advisory Committee on Statistical Methods (or to other advisory committees on specific programmes);
    - Review of the practices of other national statistical agencies and the exchange of experiences with them;
    - Participation in working groups of multilateral international organisations
    - Addressing particular technical problems;
    - Presentation of technical issues and proposed solutions for review at Symposia and other professional meetings; and
    - Use of Work-in-Progress reviews subject to the procedures laid out in the Policy on Statistics Canada's Daily (Policy 3.3).
- Managing Timeliness
  - Monitoring of timeliness
  - Early warnings in case of deterioration of timeliness
- Managing Accessibility
  - Product definition and design (according to user profiles and findings from market research)
  - Dissemination
  - Needs of analysts (e.g. microdata access)
  - Efficient search mechanisms to help users find what they need
- Managing Interpretability
  - Informing users on the concepts and classifications that underlie the data
  - Informing users on the methodology used to collect and compile the data
  - Informing users on measures of accuracy of the data (Policy on Informing Users on Data Quality and Methodology)
  - Provision of metadata
- Managing Coherence
  - Application of Standards (Policy on Standards)
  - Measurement of Inconsistencies between Data Sources

- Environmental Factors
  - Partnership with suppliers
    - Respondent relations programme;
    - Response burden management programme;
    - Bilateral committees (e.g., with Canada Customs and Revenue Agency);
    - Engagement with the small business community;
    - Small business ombudsman;
    - Electronic reporting initiatives; and
    - Recognition of respondents in publications.
  - Recruitment and Training
  - Data Analysis and Research and Development

Elements of the Statistics Sweden Quality Assurance Framework (Bergdahl and Lyberg 2004):

- Checklists as tools for the stabilisation of processes
- Current Best Methods (CBMs) as coordination tools for about 10 areas of strategic importance (primarily in various fields of statistical methodology)
- Management information for each survey on the status of finances, operations, users as well as staff
- Annual internal quality survey on product quality (covering also some aspects regarding process quality)
- Annual staff perception survey
- User surveys
  - User Satisfaction Index Survey
  - Delivery Survey
  - Image Survey
- Auditing approach to have a systematic and continuing review of all surveys and improve quality

Future tasks:

- Strategic planning and controlling
- Embedded experiments to evaluate the effects of changes in processes
- Analysis of process data
- Implementation of a uniform project model
- Coordination of processes

2.4 The Tasks of Data Quality Assessment

As noted above, quality assurance frameworks do not only encompass methods for data quality assessment, but further aspects like:

- Documentation
- Process management
- Standardisation of processes and statistical methods
- Strategic planning and controlling
- Enhancement of improvements
- Quality measurement
- etc.

Effective methods and procedures for data quality assessment of all of these aspects are a key success factor for every quality assurance framework. Furthermore, data quality assessment methods and tools in use have to be fully integrated into the quality assurance framework. Any quality assurance framework builds on the results from product quality measurement and provides inputs to the strategic planning system and improvement projects. Data quality assessment needs, as a frame of reference, some definition of minimum requirements, guidelines or recommendations. Therefore, a standardisation of production processes largely facilitates an effective data quality assessment.

Data quality assessment on the product and process level

Data quality assessment methods, based on the results of quality measurement and documentation of processes and products, provide information that enable the managers to systematically control data quality on the level of each individual statistics. The results of quality assessment are the main input to improvement actions. Data quality assessment tools in the context of this handbook include those procedures which are available to report and assess data quality, e.g. quality reports, self-assessment checklists, or auditing schemes. Such tools are typically provided as checklists, templates or organisational procedures but can also include software package that facilitate their application.

Any assessment work on the product or on the process level should be embedded into a quality assurance framework. Only with such an approach can quality improvement be achieved in a systematic way including comparability between indicators of different statistics and of different times. Without a systematic approach many indicators will lack objectivity and it will only be possible to create a quality culture in an organisation if there is a common endorsement of all hierarchies to a quality framework.

When planning data quality assessment methods, an organisation has to make a selection from a set of assessment methods. These methods include audits, self-assessments, peer reviews, quality reporting, quality indicators and process variables. The practical implementation of these methods is, in turn, much facilitated by a large number of tools. For example, auditing and self-assessment often rely on checklists (like the DESAP – Development of a Self Assessment Programme – checklist) or a sequence of workshops. The application of quality indicators requires the use of certain types of graphs and charts. Tague (2005) distinguishes nearby 150 such quality tools in the categories “project planning and implementation tools”, “idea creation tools”, “process analysis tools”, “data collection and analysis tools”, “cause analysis tools”, as well as “evaluation and decision making tools”. It is however obvious that only part of these tools is relevant for data quality assessment.

A – probably central – issue quality work is currently facing prevents a more aimed implementation of TQM models and quality assurance frameworks: the difference between theory and practice on the level of products and processes (Sæbø 2006). This means that in theory the NSIs implemented central quality improvement components and that they do report so while the actual progress in using the component or having it readily available in practice varies from being incomplete to absent. Anyway, many NSIs seem to have severe problems to find a link between general quality frameworks and their core processes. Thus, although having implemented a TQM model on a very general level, only little concrete improvements in the daily business of the NSI might be achieved. Apart from the fact that this link is the most difficult task of TQM implementation, one reason might also be, that measuring the indicators (of product as well as process quality) is far more difficult than originally thought (Sæbø 2006).

To achieve progress in quality work it is therefore vital to improve and enhance the measurability of indicators on product and process level and accordingly to systemise existing assessment approaches – the central issue of this handbook. There are multiple approaches to data quality assessment, as will be shown. However to proceed with the idea of a "quality taxonomy" first of all different quality dimensions are introduced to systemise the product and process level.

3 Data Quality Assessment Tools and Data Quality Dimensions

3.1 Product, process quality and user perception in the light of the ESS quality criteria

In the context of this handbook (and in analogy with the European Statistics Code of Practice), the notion of data quality refers to the characteristics of the statistical product and of the statistical production process as well as the user perception of quality of statistics. All three can be thought of as "dimensions" of data quality. However assessment is different. Along the product dimension data quality is assessed via the ESS quality criteria[15] relevance, accuracy, timeliness and punctuality, accessibility and clarity, comparability and coherence. These components are central for any assessment of product data quality in statistics but maybe also used to assess user perception.

The second dimension defiantly is the process level as it was noticed before that processes have a vital impact on the quality of any product. However the ESS quality criteria are only assessing product quality and other criteria for assessing process quality are needed. Their identification is not so easy, as it will be discussed in the respective chapter "Measurement of process variables".

It could be discussed whether user perception of the data should equally be treated as "data quality" in the context of this handbook due to the different view on the quality criteria and also because dealing with them might create a volume on its own. As a compromise, user perception is not totally left out but is discussed under "other methods". The reader should be aware that this handbook can not cover all issues of user perception. Anyway it is important to notice that user perception can be regarded as a third dimension of data quality along the quality criteria (not to be confused with the product characteristics in a technical sense) and can equally be assessed according to the ESS quality criteria.

The basic idea behind the distinction of the three dimensions is that a balanced quality framework should strive for an equal assessment on each dimension to achieve a comprehensive picture on the data quality of each statistics according to the ESS quality criteria. Hence each data assessment method needs to be described in the view of the two (three) dimensions to reveal which data assessment method is suitable to measure which dimension.

In the text, until now a rather theoretic view point on data quality has been taken up with the aim to provide a systematic fundament ("taxonomy") of existing quality approaches. This theoretical view is now left to achieve an overview of the assessment methods available and to examine how each method relates to the different levels and dimensions, before they are described in detail within the respective chapter.

Process stability – as the first process quality criterion – is a precondition to any controlled improvement. Any process will show some variation (e.g. the response rates obtained by different fieldwork organisations). A process can only be considered stable if its variation consists entirely of random effects; in other words, the outcome of the process has to be reasonably predictable within certain limits. Only if statistical processes are reasonably stable, key process variables could provide a basis for an evaluation of process changes. Process

[15] The term quality components is used in the literature synonymously to the term criteria. One can also find the synonymous use of the term dimension in the ESS literature, however, in the context of this handbook the term dimension is rather used to describe the different views or vectors to each quality criteria.

stability does not necessarily imply that the process is operating well enough to meet user requirements. Stability (or control) only tells us that no unusual process variations are being observed.

The second process quality criterion, capability, can only be determined for stable processes. In general terms, a process is capable if it predictably meets the process specifications. Process specifications have to be defined in accordance with the users' quality requirements for the final product. They therefore have to be defined for all product quality criteria, which is a highly complex task. Process specifications are not always inherent in the product characteristics, but have to be derived from the product quality requirements. Furthermore, in statistics, the final product quality is not only multidimensional, but is achieved through a package of interrelated process specification. In some cases, a product quality requirement could be achieved with different "packages". Examples of process specifications are minimum response rates, production deadlines, and maximum coefficients of variation.

3.2 Systematic overview of Data Quality Assessment Methods and Tools

Following the tentative definition of data quality assessment on the product and process level given above "Data quality assessment is an important part of the overall quality management (or TQM) system of a statistical agency. However, its scope is limited to the statistical products and the processes leading to their production". The first precondition accordingly is measurement and the immediate question arising is: Can we measure data quality along every quality criteria and along each dimension? The answer is given in the final LEG on Quality report (Eurostat 2002) stating that the quality dimensions are not always measurable in an objective and direct way. Accordingly it is known at least since 2002 that it will often be proxy measures or qualitative appraisals that need to be used because most of the criteria are qualitative in nature. The LEG on Quality in 2001 concluded that the "current level of measurement capability is low" and it is recommended that measurability should be improved. A more recent document states that it is still a "general scarcity of measurements of various quality dimensions" and that "it seems as if there is greater need for measures of error magnitudes and less need for revised or new frameworks" (Statistics Sweden 2006a).

The "quality taxonomy" which has been developed according to the three levels organisation, product and process can be summarised as following: most existing quality approaches should – to achieve more clarity – be distinguished according to the perspective (level) they inhere to be able to develop the appropriate measures for improving quality. Still all measures taken to assess and improve quality need to be in line with each other meaning that interfaces between the different levels should be developed wherever possible. Accordingly, nearly any quality approach can be assigned unequivocally to a single level but parts of each level should always be detected.

3.3 Current Status of Data Quality Assessment Methods and Tools in the ESS

It has been mentioned earlier in this chapter that there is discrepancy between the actual status of the implementation of central quality improvement components and the reports on their implementation. To get an overview of how this handbook relates to the ESS quality framework with respect to the product and process level and on the progress that has been achieved on the assessment methods and tools, the recommendations in 'Quality in the European statistical system - The way forward' (2002) which are most relevant to data quality assessments are selected (table 3) and the current status of their implementation according to the 2004 LEG implementation status report is reviewed.

Table 3: LEG on quality recommendations referring directly to data quality assessment and status of their overall implementation (recommendations on user relationships are provisionally left out)

Recommendations	Current implementation Status
Recommendation no. 1: Each NSI should report product quality according to the ESS quality dimensions and sub-dimensions.	Quality reports are widely implemented however there are still differences in their degree of detailedness and in the amount of reporting quality indicators.
Recommendation no. 2: The measurability of each ESS quality dimension and sub-dimension should be improved.	This recommendation might be partly met, however it was found there is need to define more precisely when it will be met.
Recommendation no. 3: Process measurements are vital for all improvement work. A handbook on the identification of key process variables, their measurement, and measurement analysis should be developed.	The handbook has been provided but measurement and implementation of process variables need further development.
Recommendation no. 15: A generic checklist should be developed for a simple self-assessment programme for survey managers in the ESS.	A checklist (DESAP) has been provided but there is potential for further development and its implementation in the NSIs.
Recommendation no. 16: The methods for auditing on different levels and for different purposes such as internal, external, one point in time, continuing or rolling, rapid, and more extensive (such as EFQM assessment) should be reviewed and recommendations should be provided to the ESS.	This recommendation has been partly met, however additional development will be required.

This review reveals that currently there still a lot of work to do with regard to the implementation of data quality assessment methods. Accordingly the handbook will not only provide an overview of current knowledge but will also focus on possible gaps which need to be filled to achieve a complementary picture of data quality. Experience in the NSIs on the implementation of data quality assessments varies as it can be deduced from the first self-assessment that was conducted as a basis for the implementation of the European Statistics Code of Practice.

The results of this self-assessment reveals that a TQM system is introduced by 10 NSIs and that further 10 NSIs are planning to implement a TQM system. There are 21 NSIs that have established a long-term Strategic Plan or Business Plan.

With respect to data quality assessment the results on monitoring product and process quality are of special interest. It became obvious that only 7 out of the 29 NSIs regularly monitor quality according to the ESS quality components for more than 75% of their statistical output and that 4 do not monitor output quality at all. Further on, less then 25% to none of the key statistical outputs are subject to comprehensive regular reviews in most NSIs and only 3 NSIs report to review of more than 75% of its key output regularly.

Some kind of quality monitoring of the various production processes is reported by all but 2 NSIs for data collection, capturing and processing, analysis and dissemination. Survey design is monitored in all but 4 NSIs. Methods employed for monitoring range from internal audit (15), self-assessment (21), quality reports (26), quality indicators (21) or other (7), like more sophisticated monitoring systems and all but 1 statistical authority report they use more than one procedure.

Already these results reveal that there is no one single way how to embed data quality assessments in the quality frameworks. Accordingly this handbook aims to provide an informative collection of good ideas, which until now have proven successful and hints where future developments might lead. However it can not be a strict guideline how to implement data quality assessments in the different NSIs as this needs individual solutions.

ANNEX B:
Examples

Examples for Chapter 2.1: Quality Reports and Indicators

1 Examples for different quality report structures

Detailed quality report for producers and expert users – Recommendations:

- Administrative information
  - The name, the reference period and the periodicity of the survey.
- General Description
  - The design and methods used for the survey
    - A description of the methods used during the survey process (classification, sampling design, data collection process, etc.)
- Relevance
  - A description and the classification of the users.
  - A description of the variety of the users' needs.
  - Main results regarding the satisfaction of users.
- Accuracy
  - Sampling errors
    - Order or magnitude (or at least sign) of the bias of the main variables.
    - Estimated coefficients of variation (CV) for the statistics.
    - Methodologies applied for variance estimation.
  - Coverage errors
    - Type and size of coverage errors.
    - Information about the frame: reference period, updating actions, quality review actions.
  - Measurement errors
    - The measurement errors identified and their extent.
    - Indications about the causes of measurement errors.
  - Processing errors
    - A summary of the processing the data are subjected between collection and production statistics.
    - Processing errors identified and their extent.
  - Non-response errors
    - Non-response; unit and item non-response rates for the main variables, both unweighted and weighted.
    - Imputation methods used (if any).
    - Indications about the causes of non-response.
- Timeliness and Punctuality
  - The average timeliness of data
  - The data frequency and average data freshness
  - The reasons for late delivery
- Accessibility and Clarity
  - A summary description of the conditions of access to data: media, support, marketing conditions, existing service-level agreement, etc.
  - A summary description of the information accompanying the statistics (documentation, explanation, etc.)

- Comparability
  - Comparability over time
    - The reference period of the survey where the break occurred.
    - The differences in concepts and methods of measurement before and after the break
- Coherence
  - Coherence of statistics in same domain
    - Summaries of the mirror statistics.
  - Coherence with National Accounts
    - A summary of the comparison.
- Cost and Burden
    - Cost supported by National Statistical Institute (NSI).
    - Response burden.

(*Based on*: How to make a quality report, Standard quality report for Labour Force Survey, Eurostat)

User oriented quality report for non-expert users – Recommendations:

- Introduction
    - Brief summary on the quality reporting programme of the NSI.
- Summary of quality
    - Information for the output around the six European Statistical System (ESS) dimensions of quality.
    - Brief textual summary on each quality dimensions.
    - Links to quality indicators or detailed descriptions.
  - Relevance
  - Accuracy
  - Timeliness and Punctuality
  - Accessibility and Clarity
  - Comparability
  - Coherence

Summary of methods used to compile the output

- A high level summary with a number of links to more detailed information for the more expert user.

(*Based on*: Summary quality report for Internal Migration; Office for National Statistics (ONS))

2 Austria

Statistics Austria started to install some projects to create new innovative ideas for the statistical production processes in 1998. One of these projects was called SYSQUAST (System of Quality Measurement for Statistical Products).

The aim of this project was to develop a prototype of the standard quality report for all statistical products in Statistics Austria. The first recommendation of the project was to select indicators which should be built in the Austrian standard quality report to characterize the Eurostat quality components. There were good reasons to use the concepts of Eurostat. First it was a well considered concept and the definition of quality also fit to the Austrian Statistical system. The obligation for compiling quality reports for the Eurostat arose in some statistical areas parallel with the implementation of the Austrian quality report system.

The need of assessing, assuring and reporting quality of statistical products was not only desired by the management and users but also a mandate by the Federal Statistical Act of the year 2000. As a follow up of SYSQUAST it was planned to implement a full in-house quality report system as a project of the TQM-Board – Total Quality Management-Board.

To implement the quality report an MS Access database was developed (called Quality Report Database (QRD)) where the responsible person for a survey could compile the report. Along with the QRD a manual was developed which should guideline the statisticians in compiling the QRD. The results of the testing showed that the filling in of the database was too inconvenient and the controlling of the completeness was too demanding.

In the year 2000 because of the users needs it turned out to be necessary to give supplement to the indicators of the QRD and to prepare detailed quality reports. They selected a set of important products for which it would be profitable to have such detailed quality reports.

As a consequence of the receiving of detailed quality reports it was necessary to think over the position of the reports within the organisation.

The experiences with compiling detailed quality reports showed also some problematic aspects. Compiling a quality report requests a large amount of information. The experts have to collect a lot of information which is not easy at the first time. The problem is that the survey expert often does not know which information is needed for compiling the report and/or sometimes does not know from where the information comes from.

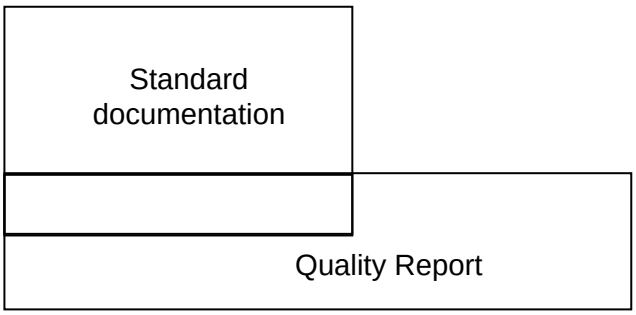
Otherwise the quality report is not a control instrument but a good opportunity to get to a sufficient documentation.

A very huge wave of complaints by the statisticians arose due to fact that they do not like to have double works. On the one hand they had to compile a quality report. But on the other hand a documentation was needed which should be delivered with standard publications. Basically the detailed quality reports followed a pattern very strictly along the definitional concepts of Eurostat. The quality report had the following chapters:

- Introduction
- Relevance
- Accuracy
- Timeliness and Punctuality
- Accessibility and Clarity
- Comparability
- Coherence
- Additional aspects effecting quality

Some special aspects which should be contained in a standard documentation are not part of these quality aspects. On the other hand there are aspects which should be contained in both documents.

Figure 11: Standard documentation and quality report



To find a compromise between avoiding double work and fulfilling both documentation and quality reporting requirements an other structure was created after which survey managers can compile the document in a standardized way. This structure contains the following elements:

- Important hints
- General information
- Statistical concepts, methods
- Statistical processing, quality insurance
- Publication (Accessibility)
- Quality
  - Relevance
  - Accuracy
  - Timeliness and Actuality
  - Comparability
  - Coherence

The quality part has been moved to the end of the document. Besides the quality information a lot of metadata are documented which are helpful for understanding and interpretation of the results.

In the first chapter, *Important hints* you can find some important issues which should make clear to the user what he can expect from the statistical product. If there are special conceptual differences to some other agreed standards it should be noted here.

Under *General information* you can find some organisational aspects such as type of statistics (primary, secondary, etc.), responsible person (or unit), periodicity, legal basis and so on as well as historical aspects and purposes of the product.

Statistical processing, Quality insurance describes all the methods which are used during the several processing steps as data capture, coding, editing and imputation, estimation methods, other measures for quality measurement etc.

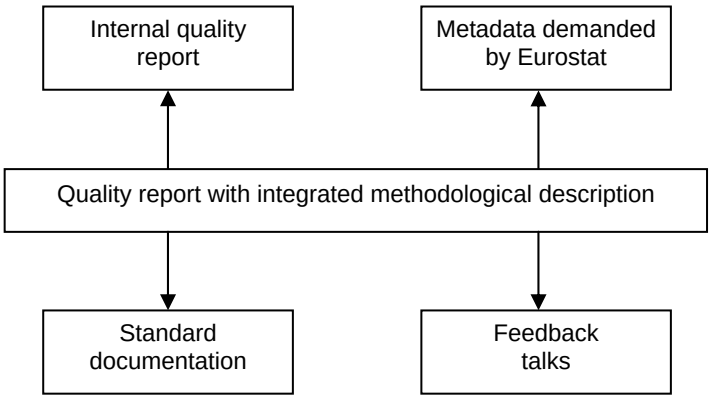
Finally *Publication (Accessibility)* refers to the aspects of publication availability. The matter of confidentiality is also covered in this chapter. The quality components Accessibility and Clarity are moved out from the other quality components.

Such a document described above is called *Quality report with integrated methodological description***.**

The new concept of quality report gave the possibility to discuss the main characteristics of the report with broader audience. The aim of the so called Feedback talks is to discuss and identify the potential sources of the improvements with the most important users. During the Feedback talks the needs of the users, the user satisfaction with the documentation and quality are identified.

This type of quality report is appropriate to serve different purposes.

Figure 12: Uses of quality reports



A strategic and political decision is which parts of a detailed quality report should be published. To put the whole on the internet can be problematic because the detailed quality reports contain self-critical statements of the product which are only adequate for internal purposes. The solution for this problem was to take the document (Quality report with integrated methodological description) after removing the critical statements (especially from the chapters “Statistical concepts, methods”, “Statistical processing, quality insurance” and “Quality”) as standard documentation on the internet.

Recommended Readings

Burg, T. (2004): Quality Reports at Statistic Austria. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.
http://www.statistik.at/

Burg, T. (2005): Quality Measurement at Statistics Austria, “Managing Modern National Statistical Systems in Democratic Societies”, Tacis High Level-Seminar, Vienna, Austria, 31 August-2 September 2005.

3 Italy

Under the mandate of Italian Statistical Office (Istat) top management, a working group in charge of defining a minimum set of quality indicators for the external users carried out its activities during 2006.

The working group selected a set of standard quality indicators having the following characteristics:

- To be relevant of the users;
- To be available from the Information System for Survey Documentation (SIDI) in use at Istat;
- To be coherent with the standard quality indicators defined at Eurostat;
- To be feasible to compute.

The complete group of indicators included the following list:

Accuracy
- Coefficient of variation (CV) or Confidence Interval
- Response Rate
- Imputation Rates
- Indicators on the Revision policy

Timeliness
- Timeliness for provisional data release
- Timeliness for definitive data release
- Timeliness for the external sources of data (where appropriate)

Comparability
- Length of the homogeneous time series

Coherence
- Difference or Relative Difference between definitive and provisional estimates

It was then established that:

- The quality indicators should be presented to the external users in the framework of a qualitative Quality Report, supporting their interpretation;
- The Quality Reports should be disseminated within the SIQual system, i.e. the System on the Quality available on Istat website;
- Each process could compute and present a subset of the quality indicators, appropriately chosen within the wider group.

The Quality Report was defined as to follow Eurostat quality components and the typical error classification. A paragraph on the methodological context was considered relevant to be added.

The activities of the working group included also an experimental phase. Indeed, the indicators and the quality reports where prepared by the survey managers of 21 statistical processes (13 direct surveys and 5 secondary studies) chosen across the Institute as to represent different areas and different level of complexity of the surveys.

The survey managers of the processes involved in the pilot test, were provided with: i) guidelines for preparing the Quality Reports; ii) explanations on the aims of the reports and the meanings and formulas of the indicators to be included; iii) a tool for producing and storing the quality reports in the SIDI documentation system. However, they were left rather free to prepare the quality reports as they considered more suitable for the external users.

The results from the pilot experience highlighted a great heterogeneity concerning both the length and the content of the quality reports, ranging from 2 to 12 pages. With regard to the quality indicators, the coefficient of variation was preferred to the confidence interval as an indicator of the sampling error. Furthermore, for the surveys producing a wide number of estimates and having a large number of domains it has been necessary to report summary measures, providing the link to documents with the complete series of indicators. One of the quality reports did not report the values of the sampling indicators, but some symbols associated to ranges of values, as required by Eurostat.

Concerning the response rates, no criticisms were found in the computation. It was decided that for the Short Terms Businesses Statistics, characterised to have a monthly data collection process, was better to provide the external users with year average rates, rather than with monthly ones.

Among the many different indicators that were proposed to describe the revision policy, those that were preferred by the survey managers were the Index of Mean Revision, and the Index of Absolute Mean Revision.

Concerning the timeliness, it was decided to integrate the final data timeliness indicator with the timeliness for external sources in order to better reflect some uncompressible times typical of some surveys relying on administrative sources.

Depending on the type of released statistics (variation vs. level estimates), the difference or the relative difference percentage was chosen as coherence indicator.

The activities of the working group have been presented to the top management of the Institute and a mandate to continue the working group was assigned. The aim is now to analyse in deeper detail all the quality reports, also with respect to the language used and with the support of experts from the Communication and Press Unit, in order to obtain a set of prototype of quality reports. On this basis, the guidelines for preparing the quality reports will be finalised, the quality reports will be submitted for final validation to some external users and a procedure for their release outlined. The first quality reports should be provided to the external users within few months.

Recommended Readings

Brancato, G. et al. (2004): Standardising, Evaluating and Documenting Quality: the Implementation of ISTAT Information System for Survey Documentation – SIDI. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.

Di Fonzo, T. (2005): The OECD Project on revision analysis, first elements for discussion. Technical Report, OECD Short-term Economic Statistics Expert Group (STESEG).

Eurostat (2003d): Standard Quality Report. Methodological Documents, Working Group "Assessment of quality in statistics", Luxembourg, 2-3 October 2003. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143233/STANDARD_QUALITY_REPORT_0.PDF#search=%22Standard%20Quality%20report%202003%20Eurostat%22

Eurostat (2005d): Standard Quality Indicators. Working Group "Quality in statistics", Luxembourg, 23-24 May 2005. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143233/STANDARD%20QUALITY%20INDICATORS.PDF#search=%22standard%20quality%20indicators%20Eurostat%22

4 Slovenia

The procedure developed in the Statistical Office of the Republic of Slovenia (SORS) serve as an example

- To provide indicators in a standardised, quick and automatic way with simple tools;
- To compile standard quality reports from databases;
- To improve comparability of indicators in range and orientation with simple transformation.

The main experienced advantages of the system:

- The system enables methodologist to have quick access to the information on data quality, in addition many times enables to detect some errors in the processing system or even in reporting data.
- Standardized form of the Summary Quality Report makes the information more readable and comparable.
- Keeping quality indicators in the same database should enable easy and effective control of the attained degree of quality for particular survey over time.

The process of quality control has been firstly developed for the purposes of some short-term business surveys, where growing demand for quick results of the survey was experienced.

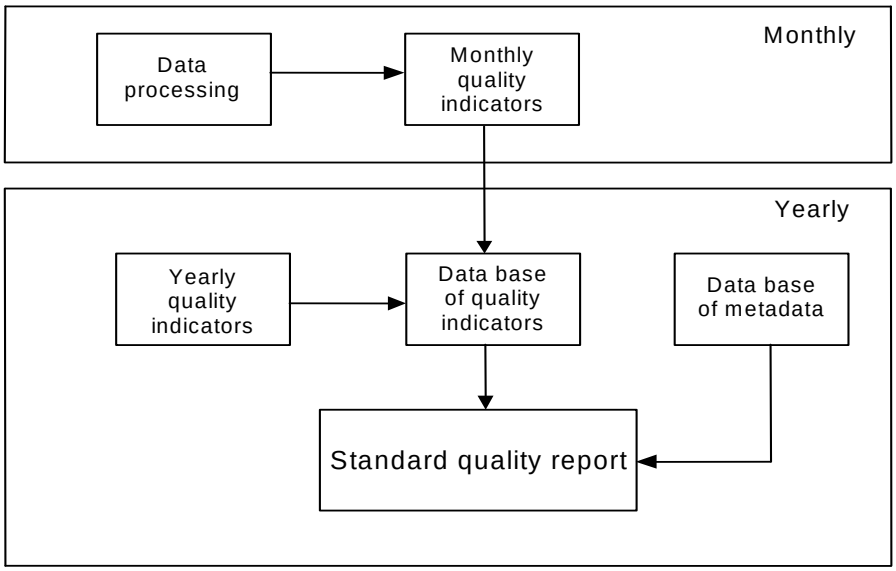
In year 2003 all the main parts of the methodology of the three monthly surveys in business statistics (on Retail Trade, Hotels and Restaurants, Service Sector) have been standardized.

The data of the surveys are stored in two databases. The first database contains information on the units which are defined at the time when the sample was drawn (i.e. size class, NACE code, last year's turnover, address, ...). The second database contains information that they get monthly as the result of their survey (raw data). The data in this database are of course keyed every month. The corrections of reported data are allowed for current and previous year.

The data processing including imputation, weighting and tabulation is fully automatized through a set of SAS macros (Statistical Analysis System) which are run by using tailor-made MS Access graphical interface. They minimize the time used for this part of the process to approximately 15 minutes and consequently gain pretty large amount of time that they can use for other parts (i.e. data cleaning, data analyzing,...) of the process.

Because of the mentioned demand for quick results there is a risk of publishing results of poor quality. To avoid such publications and to assure quick and effective control of quality of produced results they set up the system of quality control which will be presented in the next sections. Their main goal was to incorporate this system as much as possible in the automated process shortly explained above. The basic concept is presented by following picture:

Figure 13: Automated process for Standard Quality Report



The system could be shortly summarized as follows:

- Monthly control of produced results is based on the set of quality indicators which are subset of the complete list of quality indicators defined for the purposes of the Standard Quality Report (SQR). They will refer to these indicators as monthly indicators. The calculation of monthly indicators is incorporated into the data processing system and is done automatically every month.
- At the end of the year they calculate “yearly quality indicators” by prescribed methodology. Some of these indicators are calculated by using monthly indicators and some of them are calculated independently. The whole set of quality indicators is stored in the database that contains indicators for different surveys and for different reference years.
- Some additional textual data information (i.e. information on sampling frame, sampling design, media used for publishing etc.) that should be included in SQR are also stored

in special database which contains these information for different surveys and for different reference years.

- Information from both databases are then merged together into the prescribed and standardized form of the SQR. The standardized form of SQR is defined by Word's template which is directly linked with both databases and enables quick and user friendly procedure of producing final version of SQR.

Standard indicators

The complete list of quality indicators which should be included in the SQR has been determined on the basis of the list proposed by Eurostat Task Force on Quality Indicators. The proposed list was studied and discussed within the special working group of the SORS. The result of these discussions was the list of 18 indicators defined for the need of the SQR produced in the SORS. The methodology of calculation of these indicators is based on the following rules:

- The values of the indicators should be on the interval $[0,1]$. The additional values $\{-2,-1\}$ are also allowed where –1 stands for the case when the value of the indicator is not available and –2 for the case when indicator is not applicable.
- The indicators should be defined in the way that smaller value of indicator means higher degree of quality. Thus value 0 should refer to the ideal degree of quality.

In some cases limitation of values to the interval $[0,1]$ follows directly from the definition of the indicator, while sometimes some additional calculation to assure the prescribed range is needed. For the latter case it is recommended to publish both original value (it shall be called nominal value) as well as recalculated value in the interval $[0,1]$ (it shall be called standardized value).

The monthly indicators are automatically calculated by SAS macro procedures and then exported into several EXCEL files but the methodologist can see them in just one EXCEL file which is linked with all the produced EXCEL workbooks. Calculation of indicators is always done for the results of every month of the current year. This enables methodologist not only to see the indicators for the results that should be published but also compare them to the indicators for the results of previous months. Indicators could be seen in the form of an EXCEL table and could be also graphically presented by an EXCEL chart.

Standard Quality Report

Described list of monthly indicators is calculated each month but the Standard Quality Report should be prepared just once a year. Therefore when all results for a particular year were published as a last step all the monthly indicators should be summarized (usually the average of monthly indicators is calculated). Together with the rest of indicators which could be calculated only after the results for complete 12 months, they should be inserted into the database of quality indicators. In the case where the indicator refers to a particular variable, indicators for all key variables are stored.

Indicators are inserted into database through MS Access interface which also contains some basic logical controls (i.e. the entered value could only be values from interval $[0,1]$) The interface offers the user possibility of looking over, editing and inserting new indicators.

The second database that should be filled after the completion of the particular reference year is the database of textual information on the survey. The database is organized in the similar way as the database of quality indicators meaning that all the information for particular year is stored in one record of the database. The identification number of the record is the code of the survey together with reference year. Also similar interface as in the case of the quality indicators database is on disposal.

There is quite large amount of textual information that should be provided for each survey. Examples of the required information:

- Users of the survey.
- Target population and procedure for construction of sampling frame.
- Sample design.
- Weighting system and imputation method used.
- Reasons for possible delays of the first release.
- Detailed list of types of dissemination.
- Procedure of disclosure control.
- Reasons for larger deviations of final results from the first results.

The final step of the procedure of preparing SQR for particular survey and for particular reference year is merging the data from both described databases into prescribed, standardized document form. The standardized part of the document is prepared in the form of MS-Word template. Fields in the template are directly linked to the variables in the databases. To prepare the final SQR just the right record in the database has to be selected and saved as the SQR.

The main advantages of the system experienced after the testing period were as expected the increasing efficiency of the statistical work.

Recommended reading

Seljak, R. and Zalatel, M. (2004): Measurement of Data Quality in the case of Short Term Statistics. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004. http://www.oecd.org/dataoecd/6/8/35238995.pdf

Examples for Chapter 2.2: Measurement of Process Variables

The Eurostat handbook on process variables (Jones and Lewis 2003) contains examples of usage of different tools for the identification and analysis of key process variables: Cause and effect diagrams, pareto charts, flow diagrams and control charts.

Figure 14 shows a cause and effect diagram used to identify key process variables in quality of accessing administrative data sources, and hence in the quality of administrative data (AD) as such. The purpose of this work was to develop and improve the accessing of administrative data in INE Portugal. Data quality of administrative sources (AS) includes some of the same components as those of statistics, i.e. accuracy, timeliness, punctuality and accessibility. Accessibility depends on both technical and organisational factors. It is crucial that concepts and classifications in administrative data sources correspond with what is needed for statistics.

Figure 14: Cause and effect diagram for determining key process variables for quality improvement in accessing administrative data sources

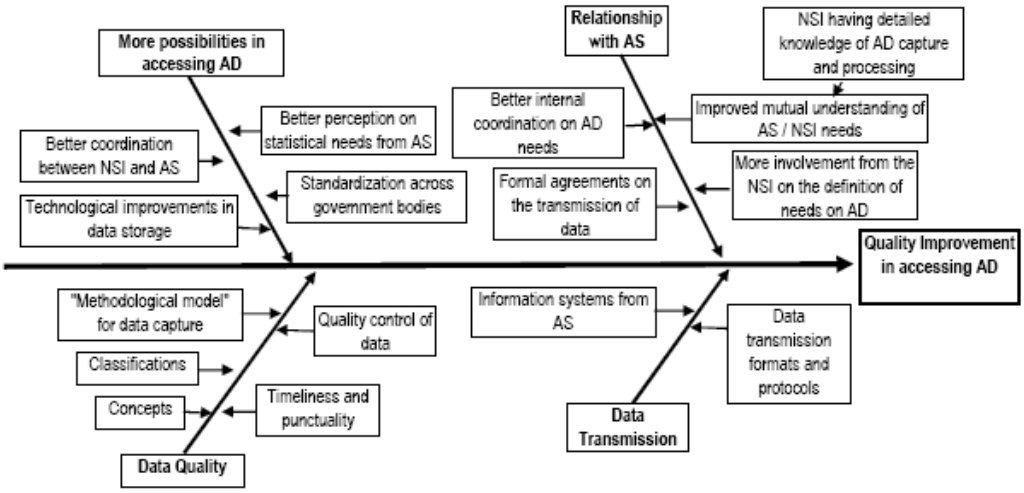
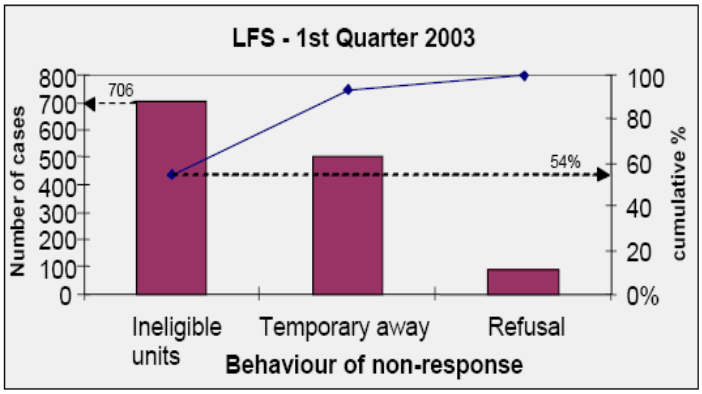


Figure 15 also originates from Jones and Lewis (2003). It shows an example of the use of Pareto chart in the Portuguese Labour Force Study to identify the main variables behind the non-response rate. It reveals that ineligible units (of respondents) were the key variable explaining non-response in this case.

Figure 15: Pareto diagram for studying non-response

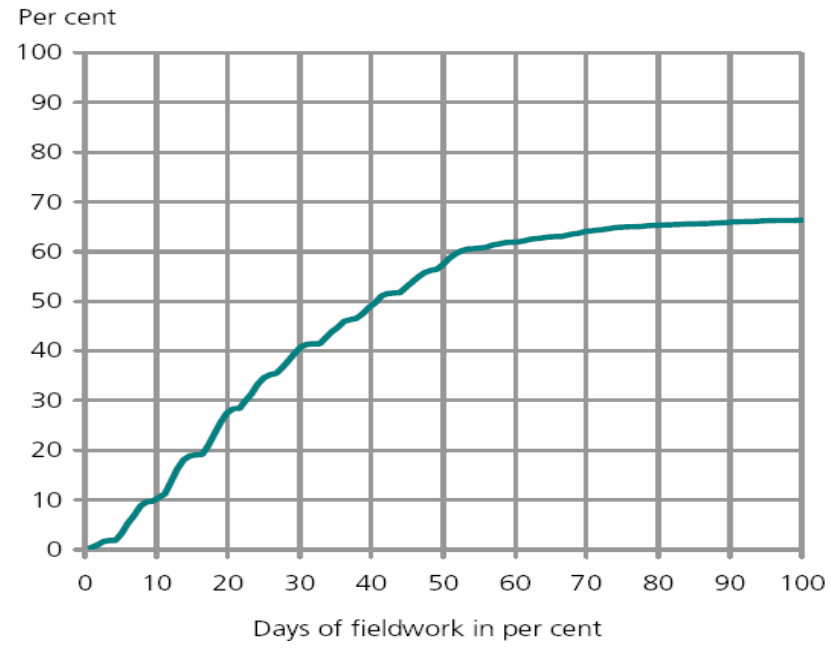


The techniques described by Jones and Lewis (2003) are the point of departure for the work of Thomsen et al. (2006). The following key process variables have been identified and measured:

- Response rate (by time and reminders)
  - Total
  - Refusal rate
  - Non-contact rate
- Days in field for data collection

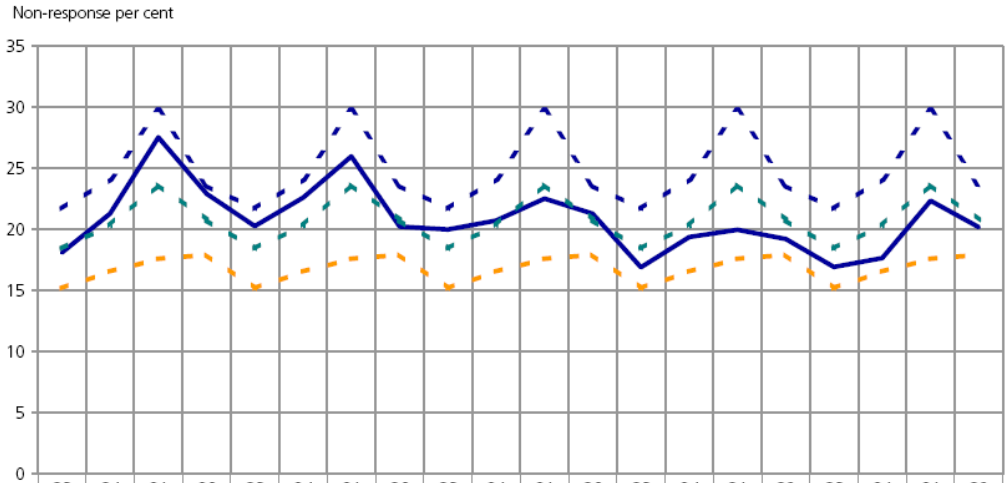
The study considers both household surveys and business surveys, both with interviewers, postal and electronic questionnaires, and mandatory and voluntary surveys. It shows the impact of non-response on product quality, i.e. accuracy vs. timeliness. It contains several tools and figures, such as cause-effect diagram, flow charts and control charts. One of the studies considered is the European Social Survey. The first example below is taken from this survey (figure 16). It shows the typical response rate development by days of fieldwork. This alone might give some ideas of the relationship between accuracy and timeliness.

Figure 16: A typical response rate development. European Social Survey 2004



The second example is a control chart for non-response in the Norwegian Quarterly Investment Statistics (QIS) survey (figure 17). It defines the outer boundaries for the accepted variation in a process. If a process variable is measured outside one of these boundaries an investigation should be undertaken to disclose the source of the abnormal variation (special cause). Typical reasons for this could be human mistakes such as forgetting to enclose pre-paid reply envelopes or an error in an electronic questionnaire making it impossible to send the data via the Internet.

Figure 17: Control chart for non-response in QIS for 1st reminder



	Q3 2000	Q4 2000	Q1 2001	Q2 2001	Q3 2001	Q4 2001	Q1 2002	Q2 2002	Q3 2002	Q4 2002	Q1 2003	Q2 2003	Q3 2003	Q4 2003	Q1 2004	Q2 2004	Q3 2004	Q4 2004	Q1 2005	Q2 2005
1st Reminder	18,1	21,3	27,5	22,9	20,3	22,6	26,0	20,2	20,0	20,7	22,5	21,3	16,9	19,4	20,0	19,2	16,9	17,6	22,3	20,2
Average + 2*Stdv	21,7	24,1	29,8	23,6	21,7	24,1	29,8	23,6	21,7	24,1	29,8	23,6	21,7	24,1	29,8	23,6	21,7	24,1	29,8	23,6
Quarterly average	18,4	20,3	23,7	20,8	18,4	20,3	23,7	20,8	18,4	20,3	23,7	20,8	18,4	20,3	23,7	20,8	18,4	20,3	23,7	20,8
Average - 2*Stdv	15,2	16,5	17,6	17,9	15,2	16,5	17,6	17,9	15,2	16,5	17,6	17,9	15,2	16,5	17,6	17,9	15,2	16,5	17,6	17,9

The figure shows that there have not been any periods with non-response rates breaking the outer boundary for what is tolerated. What could also be noticed is that the non-response is below the quarterly average in the latter quarters. If we believe that the lower non-response has stabilised in the later quarters and that this level constitutes the new lower level we should adjust the boundaries in the control chart by only using non-response data from 2003 and onwards in the calculations. The only data needed to establish this tool is the non-response rate at critical dates.

Examples for Chapter 2.3: User Surveys

The European Statistics Code of Practice has listed as one of the indicators under principle 11, Relevance, that ”User satisfaction surveys are undertaken periodically.” The Indicator has been further defined through a set of questions in the self-assessment questionnaire and some findings from the summary report are listed below.

The Leadership Expert Group (LEG) on Quality, conducted between 1999 and 2001, in its final report clearly recognized that one of the key principles of quality management in official statistics is user orientation. User surveys were mentioned as one tool for assessing user needs, but it was stated that the methodology for such surveys was still in its infancy. The LEG therefore provided the following recommendation:

Recommendation no. 7: A development project regarding the design, implementation and analysis of customer satisfaction surveys should be initiated.

This recommendation resulted in the two projects carried out by Cassel et al. (2003) and Cassel (2006).

Table 4: Information from the European Statistics Code of Practice Questionnaire 2005

NSI	User survey	Frequency
Austria		
Belgium	X	Information not available
Bulgaria	X	Bi-annually
Cyprus		
Czech Republic	X	Annually
Denmark		
Estonia	X	Twice annually
Finland	X	Bi-annually
France	X	Annually
Germany*	X	Annually
Greece	X	Every six months
Hungary		
Iceland		
Ireland	X	Every 5 years
Italy		
Latvia	X	Bi-annually
Liechtenstein		
Lithuania	X	Annually
Malta		
Netherlands	X	Annually
Norway	X	Twice annually
Poland		

(table 4 continued)

NSI	User Survey	Frequency
Portugal*	X	Occasionally
Slovakia	X	Bi-annually
Slovenia	X	Bi-annually
Spain		
Sweden	X	Annually
United Kingdom	X	Annually

* Updated after the survey was conducted

In the following some examples from the report by Cassel et al. (2003) are presented (translation according to the report; note that changes to the practices below have been made in many cases):

Germany (concerning the Eurostat Data Shop)

Please evaluate the PRODUCTS of Eurostat after their Importance (A to F) and how satisfied (1 to 6) you are with them.

Clarity of data tables

Data completeness

Data quality

Several other questions were posed but they did not relate to the European Quality Concept.

The scale used was:

A – extremely important	1 – extremely satisfied
B – very important	2 – very satisfied
C – important	3 – satisfied
D – less important	4 – less satisfied
E – unimportant	5 – dissatisfied
F – not applicable	6 – not applicable

Ireland

Please think about the Central Statistical Office's products IN GENERAL, and rate your level of satisfaction with these products in terms of the following points. Please supply a reason for your response (1 means the product is very good and 7 means the product is very poor).

Level of detail provided

Timeliness of the data

Accuracy

Relevance

Several other questions were posed but they did not relate to the European Quality Concept.

Shortcomings of existing statistics – Specific Areas

The users were asked to identify and list the shortcomings that they experienced and rank them according to level of importance using the outline below.

Specify Area and Shortcoming/Rate Importance of improving the shortcoming (1=Extremely important, 7=Not important), Rank, shortcomings in order of importance.

Finland

Using school grades (4 to 10).

A – What personal experience do you have of Statistics Finland?

B – What could reasonably be expected of Statistics Finland?

These questions were posed for the following aspects (among others):

- Basic data;
- No errors in basic data;
- Up-to-dateness and freshness of data;
- Comparability with earlier data;
- Analysed data; and
- Data meet the demands of my work.

Sweden

Application

Let us now focus on the application of the statistics produced by SCB. How do you grade the statistics (1 to 10) as regards to:

- Reliability;
- Timeliness;
- Periodicity;
- Grade of detail;
- Comparability over time; and
- Comparability between the statistical products of Statistics Sweden.

How well the contents cover your needs

The scale used is: 1 – Lowest grade, 10 – Highest grade

Portugal

Evaluation of the statistical information published by INE (National Statistics Institute, Portugal) (Paper publications and INE's internet site) according to level of importance (1 to 4) and level of satisfaction (1 to 4).

- Timeliness of information;
- Punctuality of dissemination calendars;
- Coherence of statistical information;
- Geographical disaggregation;
- Relevance of published data;
- Clarity of information; and
- Access to metadata.

The scale used is: 1 – without importance/bad, 2 – of few importance/satisfactory, 3 – important/good, 4 – very important/very good.

The same questions were also asked for the "Evaluation of non-published statistical information supplied by INE under request".

Portugal – Recent experience (2006)

User Surveys has been recognised as an important instrument to assess quality perceived by its users at INE-PT, and a new approach was developed taking into account recent experiences.

Quality dimensions were discussed and defined according to the following structure:

1. Quality perception, concerning
   1.1. Statistical information
   1.2. Products
   1.3. Services
2. Value for money
3. Global image
4. Expectations
5. Loyalty/fidelity

Examples of questions that targets data quality according to the structure referred above:

Table 5: Quality perception, concerning statistical information

I	Evaluation of the statistical information produced by the Portuguese National Statistical Institute					
	Reference scale					
	1 - Very unsatisfied	**2**	**3**	**4**	**5**	**6** - Very satisfied

		Satisfaction level
1	Statistical information pertinence (it follows evolution of the country's reality)	
2	Data credibility	
3	Timeliness of data / recent data (lag between the reference period of data and its dissemination)	
4	Easiness to access statistical information, through INE-PT website	
5	Metadata accessibility (concepts, nomenclatures, methodologies, statistical sources)	
6	Data analysis made in different statistical products	
7	Data geographical disaggregation	

The questionnaires are built in an identical way, that is, they always have the same structure. However, taking into account the different segments that are to be inquired, questions may vary (although they are all linked to the dimensions and its concepts).

Examples for Chapter 2.4: Self-assessment and Auditing

In this chapter we will present some of the information that was found related to self-assessment tools and audits. Some documentation is referred to with the aim of giving the reader a better understanding on the tools that are available to implement these activities; we wrote down some experiences on the organisational aspects of some cases that are considered good practices; and some documents on public reports are also delivered to highlight the type of recommendations and findings of these activities.

We have learnt that the experiences may vary a lot depending on the context aspects that are referred to in the main chapter (good communication, good documentation system, top management support). The experiences that we share are those where the key aspects for developing the activities were taken into account.

Self-assessment tools

1 Development of a Self Assessment Programme – The DESAP Checklist

"What is DESAP?

DESAP is the generic checklist for a systematic quality assessment of surveys in the ESS. It has been designed as a tool for survey managers and should support them in assessing the quality of their statistics and considering improvement measures. During its development, the checklist has been tested in a pilot study covering a large variety of survey areas. It is fully compliant with the ESS quality criteria and comprises the main aspects relevant to the quality of statistical data. The checklist is generic in the sense that it applies to all individual statistics collecting micro data, irrespective of the subject matter area and the specific survey methodology.

Who should fill in the Checklist?

The checklist has been designed to be filled in by the survey manager, i. e. the person responsible for the survey (or the main parts of it) without time consuming preparation. However, depending on the specific organisation of a statistical agency, it might be useful to involve other experts in the assessment process (e.g. the methodologist). Another option for carrying out the self-assessment is to fill it in together with the team responsible for the different aspects of the survey.

The objectives of DESAP

The DESAP Checklist has been designed in order to fulfil a number of quite different functions. First of all it is an assessment tool: It should be used to assess the overall quality of a survey and to compile a quality profile covering the ESS quality components. But this is not all. The checklist also provides guidance in the consideration of improvement measures and could facilitate a basic appraisal of the risk of potential quality problems. Additionally it provides a means for simple comparisons of the level of quality over time and across domains (given that data are stored centrally in the statistical agency). It could also provide support for resource allocation within statistical offices or be helpful for the training of new staff."[16]

Processes and process elements of DESAP

I Decision to undertake a survey

- Analysis of user needs
- Survey objectives

[16] Eurostat (2003i): DESAP – Development of a European Self Assessment Program. The European Self Assessment Checklist for Survey Managers, Introduction.

II Survey design

- Survey concepts
- Sample design (for sample surveys only)
- Development and testing of the measurement instruments

III Data collection

- Data sources
- Interviewer deployment and training
- Reducing non-response
- Fieldwork

IV Data capture and data processing

- Data capture
- Editing procedures
- Imputation procedures

V Data Analysis and output quality

- Relevance
- Accuracy
- Timeliness and punctuality
- Comparability
- Coherence
- Data analysis
- Disclosure control

VI Documentation and dissemination

- Metadata documentation
- Dissemination strategies
- Data management

VII Improvement cycle

- Adaptability/flexibility
- Expertise in relevant areas
- Quality management

DESAP related documents are available on the webpage of the European Commission, Eurostat:
http://epp.eurostat.ec.europa.eu/portal/page?_pageid=2273,1,2273_47143234&_dad=portal&_schema=PORTAL

2 European Statistics Code of Practice

The European Statistics Code of Practice and all related documentation are available at:
http://epp.eurostat.ec.europa.eu/portal/page?_pageid=2273,1,2273_47141302&_dad=portal&_schema=PORTAL
(e.g. the European Statistics Code of Practice, questionnaire for self-assessment, and other documentation)

3 European Foundation for Quality Management Excellence Model – EFQM

"The EFQM - European Foundation for Quality Management Model is a non-prescriptive framework that recognises there are many approaches to achieving sustainable excellence. Within this non-prescriptive approach there are some fundamental concepts which underpin the EFQM model: Results Orientation; Customer Focus; Leadership & Constancy of Purpose; Management by Processes & facts; People Development & Involvement; Continuous Learning, Innovation & Improvement; Partnership Development; Public Responsibility.

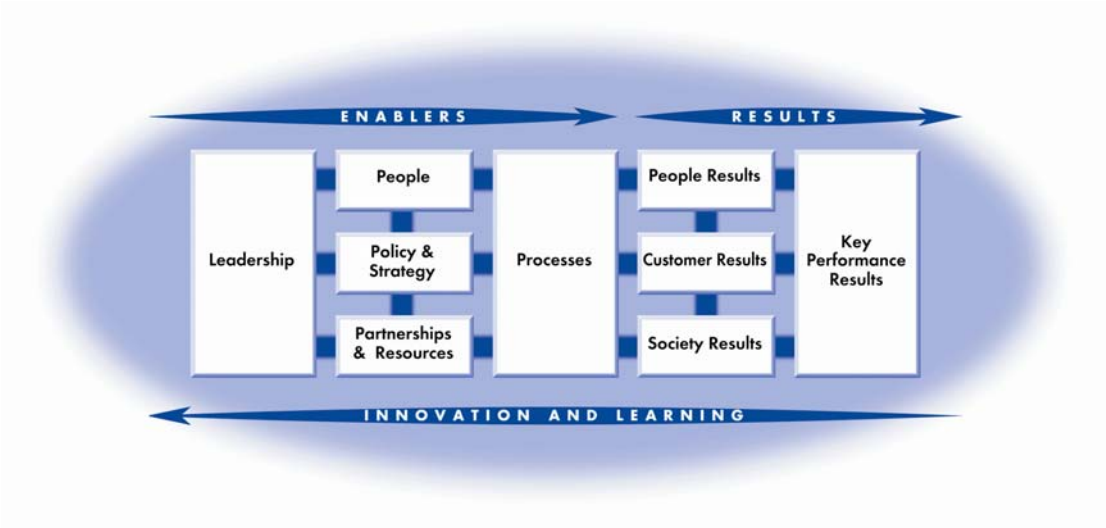
The EFQM Model can be used for self-assessment in any organization, and it's also the basis for European and National Quality awards."

"The EFQM model may be applied by an organization as an exercise of self-assessment, made by an internal team; or in case the organization wants to apply to the European Excellence Award. It can be applied to the whole organization, or to a small part like a single Department.

Either the self-assessment exercise or the external audits must be undertaken by teams of auditors or consultants that have good knowledge of the model. The model itself has checklists and assessors score books that lead auditing teams through the model. The audit report should go to top management and improvement actions should be oriented also by top management."[17]

The EFQM model and all related documents are available at: http://www.efqm.org.

Figure 18: EFQM model



4 Common Assessment Framework – CAF

The CAF is a self-assessment tool based on the EFQM model and adapted to Public Administration environment.

"It has been developed under the authority of the Directors-General of Public Administration of the EU member states in order to support the introduction of the idea and the principles of TQM in the public sector across the EU and beyond".

The CAF is designed to be an easy-to-use tool for quality management in the public sector, allowing public organisations to identify their strengths and their areas for improvement, and paving the way for application of the fundamental concepts of quality. The use of the Common Assessment Framework is free, and it is open to all types of public sector organisations in all countries.

[17] Zilhão, *M. J. et al. (2003):* State of the art concerning auditing activity in NSIs. Final Report, Eurostat, Luxembourg, p. 38 and p. 12.
http://epp.eurostat.ec.europa.eu/cache/ITY_PUBLIC/G0-LEG-20030930/DE/G0-LEG-20030930-DE.PDF

The Common Assessment Framework has been developed by exploiting the concepts, models and experiences of the "European Excellence Model" promoted by the EFQM, various national quality management tools derived from the Excellence Model and the Speyer Quality Award Scheme widely used in German-speaking European countries. One of the aims of the CAF is therefore to serve as a bridge between different quality management instruments and quality management cultures in Europe. It is not in competition with these instruments, but tries to complement them and to bring together their most important elements. The CAF has also been developed specifically as a quality management instrument for the public sector, taking into account the specific features of the mission and working environment or organisations in the public sector.

Self-assessment with the Common Assessment Framework is a less rigorous process than for instance a full assessment against the EFQM Excellence Model. The advantages of using the CAF as an introductory tool are however considerable: it is relatively easy to handle, it clearly takes less time, it demands fewer resources and introduces a common language to organisations wishing to benchmark using CAF. Being a self-assessment tool, it also has the advantage of involving the people of an organisation and of bringing forward their own views", in "Introduction to the CAF", CAF website.

The CAF brochure and other related documents available at: http://www.eipa.eu/en/home/.

Auditing

5 Auditing at Statistics Sweden – organisational aspects

Approximately 90 audits have been carried out over the four years the approach has been in place. The approach is being managed within the Audit Office in the Research and Development (R&D) Department. The office consists of three people but none work full time with the audits and in total the resources used are less than 1,5 full time equivalents. The actual audits run full time for a calendar week for a team of three people (methodologist, IT-expert, survey manager/production statistician). The auditors are chosen from a pool that at present contains about 60 people that have volunteered to become auditors and which have received a one-day introductory training. Before being admitted into the pool, the applicants are screened to assure to get the right people.

The auditing is first and foremost a supporting activity. Therefore the respective departments are asked which surveys they would like to have audited each year. To further strengthen the support aspect, the work of the audit team is financed through the survey's budget. The premise is that all surveys will go through the process, but they can influence when it will take place. The Audit Office then makes a schedule for the year matching surveys with auditors from the pool, and making sure that new auditors are teamed with more experienced. It is a principle that one should not audit a survey to which one has any connection. One member of the team is chosen to be the leading auditor and handles administrative issues and coordinates the contacts with the survey.

Before the audit takes place the survey team shall complete a self-assessment questionnaire and this, together with all relevant documentation, are used by the audit team to guide their work to the areas of the process where they can be most useful. During the audit week active discussions are held with the staff of the survey to clarify what is being done and the reasons for the present approach as well as all experiences the staff has. The auditors then compile a report with their main recommendations and any good examples that have been identified and which they think might be useful elsewhere in the organisation. The report is discussed with the survey staff and the idea is to have a consensus view on the recommendations (this is also almost always the case), which will facilitate their implementation.

The subject matter statisticians are then obliged to complete an action plan based on the report and the resources available to tackle them. The Audit Office makes regular follow-ups of what has been done to the recommendations afterwards.

Note: since late 2006 the approach is on hold to focus the organisation's resources on a major restructuring project. This project will also, when completed at the end of 2007, recommend a revised auditing approach that will be tailored to the changing needs.

6 National Statistics Quality Reviews – ONS – UK

The programme of quality reviews is an important way of ensuring that National Statistics and other official statistics are fit for purpose and that ONS and other National Statistics Producers are continuing to improve the quality and value of these outputs.

The National Statistics Code of Practice sets out the key principles and standards which official statisticians are expected to follow and uphold. It is supported by twelve protocols which describe how those principles and standards are to be implemented in practice.

The whole review process is about to change, coming to new legislation for statistics, legal Code of Practice and a legal Assessment (and reassessment) to determine if statistics meet appropriate standards.

All documentation from National Statistics Quality reviews are available at:

http://www.statistics.gov.uk/about/data/methodology/quality/reviews/default.asp

http://www.statistics.gov.uk/about/national_statistics/cop/default.asp

http://www.statistics.gov.uk/about/data/methodology/quality/projects/survey_control.asp

7 INE-PT experience in Internal Quality Auditing

The process

INE-PT started its process of internal auditing in 2000, focused on the statistical process. The reference of these audits is the Statistical Production Procedures Handbook. However, the whole process of internal quality audits was organised taking into account the international standards – ISO Norms (International Organization for Standardization).

The aim of these audits is to verify if statisticians are working according to the procedures that are settled in the organisation for this process and to improve the process description as well. Therefore, the audits work as a tool for continuous quality improvement.

According to the above mentioned norms, an internal procedure was settled for the Internal Quality Audits process, which comprises the following information for the execution of quality audits:

- Quality audits reference documents,
- Concepts/definitions,
- Quality audits aims,
- Quality audits planning,
- Responsibilities,
- Quality audits phases,
- Auditors selection/recruitment,
- Auditors registration,
- Quality auditors teams, and
- Support documentation: audit plan; non-conformity/observation form; audit report; corrective/prevention action form.

Apart from this procedure, documentation to help the auditors do their work was also provided – mainly checklists based on the procedures. The audit is based on observation of evidence and on minimum documentation that every project leader has to organize for its "survey". This minimum documentation is compulsory by means of the Statistical Production Procedures Handbook that establishes a set of common internal procedures for all statistical processes, mainly for statistics based on surveys. The idea behind this handbook is that if we define a process in a way we believe it is best then our product will have quality. This hand-

book defines compulsory minimum documentation, and internal procedures among units that somehow are involved in the process.

It is important to mention that training on “Internal Quality Auditing” was organised when launching this process. The whole team of 14 auditors participated in this training programme – behaviour and techniques in auditing – based on ISO norms.

The whole process took a year to prepare and all of the basic documentation was discussed and prepared by the team (this aspect was important for their involvement and motivation).

Some rules

- No one on the team is a full time auditor. The team is composed of statisticians from different areas.
- Another important aspect is that each audit takes, in average, 4 to 5 days, and each auditor undertakes 2 to 3 audits per year (each team has 2 or 3 auditors).
- Top management defines the quality audits annual plan.
- No one should audit his own activity.
- Improvement actions must be carried out by the audited teams.
- A follow-up audit is settled whenever there are non-conformities and corrective actions to be carried out, resulting from an audit. The date to undertake a follow-up audit is settled between auditors and audited teams.
- Improvement actions must be published and followed by the audit teams.
- Each activity should be audited on a regularly basis.

The results

- The identification of improvement opportunities in the audited process (particular surveys that were audited).
- The implementation of corrective and/or preventive actions in the audited process.
- The improvement of process definition, and basic documentation (internal procedures).
- The auditing activity involves people in the quality programme in the organisation, both auditors and audited team.
- Auditor is a “consultant”.
- Our handbook (1997 edition) had to be revised in order to define some of the sub-processes.
- It gives input for “quality improvement projects” to be developed by quality teams across the organisation.

Other approaches – INE-PT has decided to undertake External Audits, by international experts on an occasional basis.

Self-assessment using the DESAP – is also considered to be a very productive activity that is being introduced.

The whole process of internal auditing is under reformulation taking new internal procedures for statistical production.

8 Statistics Canada – Example of a policy for internal auditing

Approved June 21, 1983

Internal Audit Policy

Definition

Internal Audit will comprise the systematic, independent review and appraisal of all operations for purposes of advising Statistics Canada's management on the efficiency, effectiveness and economy of management.

Organization and relationships

The Chief Statistician is the client for Internal Audit in Statistics Canada. The Director of the Internal Audit Division will have direct access to the Chief Statistician when required. The Internal Audit Division will have access to any information necessary for the fulfilment of the audit mandate.

The Chief Statistician, with the support of an Internal Audit Committee which he chairs, approves and promulgates audit policy, approves the scope and frequency of audit, reviews audit reports and approves action plans. The Internal Audit Committee will consist of members of the Policy Committee; staff support will be provided by the Internal Audit Division.

Internal Audit will be coordinated with the function of program evaluation, which focuses on program structure and results rather than on program systems and management controls.

The Internal Audit Division will exchange schedules and audit information with the Office of the Auditor General, in order to minimize duplication of effort and disruption of departmental staff, and will establish effective communication with the central agencies, so that they are informed on the level of compliance with central agency directives and guidelines.

Plans and reports

The Internal Audit Division will develop, maintain and follow:

(a) A Long-term Audit Plan to provide for coverage of Statistics Canada's operations over the three-to-five year period.

(b) An Annual Audit Schedule to describe the specific audits to be carried out over a 12-month period.

Both plans will allow for flexibility in the use of resources to take advantage of opportunities as they present themselves and to react to significant changes in the environment within which Statistics Canada functions.

An annual report will be submitted to the Chief Statistician on audit coverage and major findings. The report will include an assessment of the impact of the function upon Statistics Canada and an estimate of benefits realized as against cost incurred.

Process

Audits will be conducted by the staff of the Internal Audit Division supplemented by individuals or organizations hired under contract by the department. Responsibility for audit assignments will remain with the staff of the Internal Audit Division.

Audits will be conducted openly. Audit findings will be disclosed to the auditee on an ongoing basis during the course of the audit. Full discussion of findings and proposed recommendations will be held with all involved levels of management.

Auditees will actively participate in the audit process. They will assist in setting audit objectives, provide necessary information, discuss preliminary findings with the audit team and participate in their modification, review and comment upon the draft final report and prepare an action plan in response to agreed-upon recommendations for submission to the Chief Statistician and the Audit Committee.

Every audit assignment will result in a written report. Final audit reports will incorporate comments and proposed actions of the auditees and will be available as appropriate.

Evaluation

This policy will be reviewed one year after its approval.

9 U.S. Bureau of Census

For information about quality audits at the U.S. Bureau of Census see:

Bushery, J. M.; McGovern, P.D.; Marck, P.S. and Hogan, H. R. (2006): Quality Audits at the U.S. Census Bureau. Paper presented at the European Conference on Quality in Survey Statistics (Q2006), Cardiff, United Kingdom, 24-26 April 2006. http://www.statistics.gov.uk/events/q2006/downloads/T08_Bushery.doc

10 Statistics Netherlands

The purpose of statistical audits at Statistics Netherlands is twofold: (1) to present an objective, internal quality check of statistical processes and products to inform the Director-General; (2) to conceive improvements and advice to support the management concerned.

The Board of Directors has stipulated that audits are not intended to generate advice about processes of which it is known that they require support. Audits should perform an objective, internal quality check.

The operating procedures in the area of statistical audits and the collected instruments are distributed widely in Statistics Netherlands, for example via intranet. Therefore each manager can organize a self-assessment. He can also prepare his unit for a statistical audit. If necessary the Audit bureau gives help and advice.

The management of a statistical process receives a judgement about the statistical process or the statistical products. On the basis of that judgement the management is provided with recommendations how to improve process or products. In the end all this should contribute to risk reduction.

Defined preconditions should be fulfilled to come to an effective and efficient auditing operation. It should be clear what the wanted goal of the audit is, how the operation will be performed, what the expected results are and what resources are available.

The independence and neutrality of the auditors, the audit team, the audit bureau and the methodology and quality division should be assured.

The audits are coordinated by an audit organisation that is independent and neutral. Not even the principal is able to influence the content of audits.

The audit bureau starts by recruiting an audit team leader. Subsequently, a draft audit assignment is prepared. After that the audit team leader drafts a project plan and adopts an available framework of standards if it is not already indicated in the assignment.

The audit team leader is responsible for the realisation of the statistical audit on the basis of the project plan. He reports on progress and content to the Audit bureau.

The audit team leader sends a draft of the final report to the Audit bureau that's responsible for further completion.

The methodology and quality division delivers the final report to the Director-General and to other relevant parties as indicated by the Director-General. The division is responsible for the publication of the audit report. For this whole process, arranging and administering the audit file is valid.[18]

[18] Policy and application of statistical audits at Statistics Netherlands (2007).

11 Statistics Austria – Feedback Talks

The "feedback talks" are an important and innovative approach to the assessment and evaluation of the quality of Statistics Austria's statistical work. They involve both internal (including the production domain, a peer domain (quality management unit, TQM-Board) and national accountants) and external experts (such as main users and methodologists).

Feedback talks were introduced in mid 2003. They have been carried out - as far as possible on a monthly basis - in co-operation with the Quality Committee of the Statistical Council with the objective of having all statistics reviewed by end of 2008. The Talks represent a statistical audit function as they aim:

- To review statistical methods and the quality of processes in detail;
- To identify quality improvement potential on the basis of external and internal users' views;
- To improve the quality reports; and
- To define recommendations and actions.

The first step was to develop a 5-year plan in which all statistical products of Statistics Austria are scheduled for a feedback talk. The detailed planning phase is performed during every 2nd month of a quarter for the subsequent quarter in a way that for every month (as far as possible) of the next quarter one product – or a set of connected products - is selected for a feedback talk. Once a product is appointed - which is done in agreement with the corresponding statistical unit - a date for the delivery of the provisional quality report is fixed. This report is evaluated by the quality management and the methodological unit. Together they check the report for things which should be added, removed or changed. About 4 weeks before the feedback talks these suggestions are preliminarily discussed with the respective statistical unit. As a result the quality report is revised according to the agreements found at the preliminary discussion and 2 weeks before the feedback talk, at the latest, a final version of the quality report is sent out to all the participants. Parallel to this, the quality management unit is informed by the statistical unit and by the Quality Committee of the Statistical Council which key users should be invited to the feedback talks. Additionally every member of the corresponding external expert group (members of advisory subcommittees of the Central Statistical Commission) which is related to the statistical product is asked if he or she is interested in taking part in these talks. The quality management unit, the members of the TQM-Board, the corresponding experts from the statistical unit of Statistics Austria and the members of the Quality Committee of the Statistical Council are standard participants.

Together with the quality report every participant receives a list of key questions in advance to the feedback talk. These central questions are:

- Which are the demands/expectations of the user / client/expert with regard to the quality of the product?
- Are these demands/expectations fulfilled?
- Are the used inputs (data sources), processes and applied methods adequate, of sufficient quality and documentation?
- Are there guidelines for quality assurance, are they sufficiently documented and observed?
- Are “good/best practices” considered in the production of statistics? (e.g., approved/successful methods/trouble shooting in other comparable areas/problems, in other countries)
- Are the burdens imposed on the data providers justified and as light as possible?
- Which are the strengths of the product that should be retained at any rate?
- Where should improvements be undertaken?

- Is it possible to improve information on the quality of the statistical product for the user /client?

Provided with this information every participant can enter into the feedback talks well prepared. At the feedback talks, the product is discussed following the quality report step by step. In the minutes all proposals agreed for improvements of the specified product and the quality report are included. The implementation of improvement actions is documented and is itself subject to review.

12 Peer reviews

Several NSIs took part in a Peer Review. To get information about the practices and the results see:

European Commission, Eurostat Quality Website:

http://epp.eurostat.ec.europa.eu/portal/page?_pageid=2273,61904978&_dad=portal&_schema=PORTAL

A Peer Review of the Swiss Statistical System:

http://www.bfs.admin.ch/bfs/portal/en/index/institutionen/oeffentliche_statistik/bundesstatistik/entwicklungen__trends/peer_review.ContentPar.0005.DownloadFile.tmp/peer_review_e.pdf

13 Audits performed under the Data Quality Assessment Framework – IMF

The Data Quality Assessment Framework from the International Monetary Fund (IMF) provides an integrated and flexible framework in which data quality is assessed using a structure that spans institutional environments, statistical processes, and characteristics of the statistical products. This methodological framework was developed by the Fund's Statistical Department in consultation with national statistical offices, international organisations, and data users outside the fund. It brings together best practices and internationally accepted concepts and definition in statistics and covers multifarious dimensions of data quality, such as integrity, methodological soundness, accuracy and reliability, serviceability, and accessibility, as well as the related institutional prerequisites.

International Monetary Fund (2003): Data Quality Assessment Framework and Data Quality Program.
http://www.imf.org/external/np/sta/dsbb/2003/eng/dqaf.htm

International Monetary Fund (2005): Sixth Review of the Fund's Data Standards Initiatives.
http://www.imf.org/external/np/pp/eng/2005/070105s.pdf

Examples for Chapter 3.1: Labelling

South Africa

A Draft Data Quality Framework provides an example here (Statistics South Africa 2006). There are four levels of ranking called (4) official statistics, (3) acceptable statistics, (2) poor statistics, and (1) questionable statistics. There is a ranking scheme for 62 indicators presented in nine sets most of which are quality dimensions. The sets include: prerequisites of quality, relevance, accuracy, timeliness, accessibility, interpretability, coherence, methodological soundness, and integrity. For example, one indicator within accessibility is "Statistics are made available to all users at the same time", and the rankings are (4) no exception to the rule, (3) occasional exceptions to the rule, (2) several exceptions to the rule, and (1) statistics are not released simultaneously to all interested parties. The framework is included here, in spite of its being a draft, because it is unusually detailed in its descriptions.

France

The term "quality label" exists in France for a long time. The French National Council for Statistical Information (CNIS) has a Quality Label Committee that was tested in 1994 and established in 1997. The Committee has four configurations: enterprises, household, agriculture, and local government. It issues (or withholds) a "public interest and quality label". The decision is based on rules of compliance with sampling methods, response burden etc. The survey may be compulsory or not. Desrosières (2000) includes a description of this French system and an Official statistics brand in a more general setting of social statistics in a European perspective.

New Zealand

New Zealand selects a subset of its official statistics and calls these key statistics Tier 1; see further New Zealand (2007). These Tier 1 statistics:

- are essential to central government decision making;
- are of high public interest;
- meet public expectations of impartiality and statistical quality;
- require long-term continuity of the data; and
- provide international comparability in a global environment.

The development of principles and protocols is coordinated by Statistics New Zealand in consultation with other Tier 1 producing agencies. The statement of principles embodies the key aspects of the Statistics Act from 1975 and the United Nations fundamental principles. There is an Advisory Committee on Official Statistics that represents the interests of the wider statistical community, including users, producers, and suppliers of statistical data to ensure that official statistics are developed to meet the needs of users while taking account of the costs to the producers.

Sweden

The decentralised Swedish system with its label for official statistics was briefly outlined in the main text. When guidelines were set up, there was a statement about sufficient quality that has then been clarified: "Official statistics shall have sufficient quality in relation to how they are intended to be used". This is a clear objective that refers to the characteristics of the statistical product and has implications for the responsible authority; which statistics shall be official and what level of quality is sufficient for the intended uses. The authority must study, assess and prioritise statistical needs. The production process must be planned and implemented so that the obtained quality is at least equal to the sufficient quality. This requires active follow-up. There is a set of criteria that serve as a clarification and as a tool for those responsible for and/or producing official statistics. The criteria are not mandatory, but there is a planned procedure with a commitment that a statistical authority can make about its official

statistics and fulfilment of the criteria (Statistics Sweden 2006b). The criteria are structured in three groups:

- Laws, ordinances and regulations that steer the official statistics,
- Contacts with users, and
- Planning – Implementation – Follow-up.

The last group has twelve criteria dealing for instance with motivation of data providers including reduction of their burden and with the relationship between quality and the production process with sub-processes. Hence, the criteria have features of *ex ante* and *ex post*.

Examples for Chapter 3.2: Certification to the International Standard on market, opinion and social research (ISO 20252:2006)

1 Introduction

1.1 European Statistics Code of Practice, statistical production process and ISO 20252:2006

Since the International Organization for Standardization (ISO) has published "ISO 20252:2006 (E) Market, opinion and social research – Vocabulary and service requirements" in May 2006 for the first time an international standard in the field of data quality is set up.

The principal objective of international standardization is to facilitate business development and growth particularly between different national and regional markets. The intention of ISO 20252:2006 is to apply the principles of international quality standards to market, opinion and social research and to harmonise other national standards already available. Subsidiary objectives are to define the level of requirement for service provision and common work procedures to be applied in processes, including across different countries.

The principles of official statistic in the ESS are laid down in the European Statistics Code of Practice. Quality reports describe the data quality. Numerous regulations contain specific procedures to obtain a defined quality level.

A comparison between the official and the market orientated approach shows, that ISO 20252:2006 focuses on process quality, while the European Statistics Code of Practice refers to the quality of institutional environment, statistical processes and statistical output.

The table below shows that most of the institutional environment of official statistics cannot be compared to market research. Professional independence, impartiality and objectivity, and mandate for data collection are typical requirements for public services. Adequate funding and cost effectiveness are requirements only for non-market institutions. Non-excessive burden on respondents as well as coherence and comparability are required, because most official statistics stem from mandatory surveys and contribute to comprehensive statistical systems.

Table 6: Principles of Official Statistics in comparison to ISO 20252:2006

Principle specific for Official Statistics	Principle covers or partly covers ISO 20252:2006 requirements
1 Professional Independence	4 Quality Commitment
2 Mandate for Data Collection	5 Statistical Confidentiality
3 Adequacy of Resources	7 Sound Methodology
6 Impartiality and Objectivity	8 Appropriate statistical procedures
9 Non-excessive burden on respondents	11 Relevance
14 Coherence and Comparability	12 Accuracy and reliability
10 Cost effectiveness	13 Timeliness and Punctuality
	15 Accessibility and Clarity

The decision to apply ISO 20252:2006 to official statistics has to take into account the different conditions of business enterprise and public service as a survey research supplier with regard to the respective national statistical system. Official statistics complies with most of the requirements of the standards, but it may require numerous adjustments and improvements of processes to comply with all. On the other hand ISO 20252:2006 gives a detailed

guidance following the stages of the statistical production process and this is very much the same for statistics of any type.

As an example one could refer to the process model of the "European Self Assessment Checklist for Survey Managers" (DESAP Checklist), a checklist which allows a step-by-step assessment of seven production processes. This process structure is roughly oriented according to the "Statistical Value Chain" (SVC) as developed by the ONS and the standard process model of the Federal Statistical Office of Germany. The following table gives an overview with chapters of the ISO norm correspond to the process steps of the DESAP model.

Table 7: DESAP Process Model and corresponding chapters of ISO 20252:2006

DESAP Process Model	corresponding chapters of ISO 20252:2006
(1) Decision to undertake a survey	4.1 Responding to research requests
(2) Survey design	4.2 Project schedule 4.4 Questionnaires and discussion guides 4.5 Managing sampling, data collection and data analysis 4.6 Monitoring the execution of research
(3) Data collection	5 Data collection
(4) Data capture and data processing	6 Data management & processing
(5) Data analysis and output quality	7 Report on research projects
(6) Documentation and dissemination	7 Report on research projects
(7) Improving statistics	-

Referring to ISO 20252:2006 could give guidance for quality guidelines and assessing the quality of the statistical production process chain. It is to be noted, that ISO 20252:2006 focuses on the process of statistical production and does not intent to give technical specifications. Therefore it lacks precise quality indicators for products as are described in ESS Quality Indicators, but merely states that surveys should comply with agreed standards. Auditing according to ISO 20252:2006 would indicate, that an organisation has the framework to comply with the standards of the ESS, but would not assess the ESS-standards themselves.

1.2. The content of ISO 20252:2006

The core is sections on:

(1) Quality management system requirements

Satisfactory evidence of compliance to an ISO service industry standard requires elements of a quality assurance procedure. Where organisations have already undertaken ISO 9001 certification such evidence is in place. However, not all organisations wish to undertake ISO 9001 as it may be inappropriate to some components of activity. This section enables organisations wishing to undertake ISO 20252:2006 to do so without the need for an additional ISO 9001 certification as well.

Thus this section deals with organisation and responsibilities, documentation, record keeping, training and competency, the use of sub-contractors, error resolution and analysis, and problem resolution.

(2) Managing the executive elements of research

This section itemises the main stages in the survey research process. It describes the procedures that need to be put in place to ensure that the requirement is understood; that the methods by which this will be achieved are clearly defined; that the inputs from both sides are specified together with their associated schedules and ownership; that all the relevant

resources, expertise and procedures are in place at the appropriate time to achieve satisfactory performance of each component of the survey process.

Thus this section begins with 'responding to research requests' and via sections such as 'contents of research proposals', 'project schedule', 'assistance by and cooperation with clients/users', 'managing sampling and data processing, 'monitoring the execution of research' proceeds to 'research reporting' and 'research records'. Clear description of the key components of the research task aims to result in comprehensive planning and clearly defined ownership of execution.

(3) Data collection

This section defines the requirements for all aspects of data collection with the emphasis on fieldwork. Rules are laid down for the recruitment and training of different types of interviewers. These training rules include content, duration and documentation. Further guidelines describe the needs of ongoing training and appraisal. The needs for adequate identification and the issue of data protection and confidentiality are covered here. Validation is covered in some detail with the specification of minimum levels – typically via back-checking, or in the case of telephone monitoring. Guidelines are set for the recording of all validation activity as part of the project record.

Example:

"*Validation levels*

Validation shall be carried out to at least the levels below (depending on the type of validation). Validation levels shall be calculated on the basis of achieved interviews and achieved validations.

Where checking of data records is the only method of validation that can be used (e.g. observational research, retail audits) the required level shall be 100%. No specific level is specified where checking of data records is used in conjunction with other methods.

For all projects where data collection is by interview, validation shall be by either re-contact or monitoring to the levels shown below, whether or not data records are also checked.

The minimum re-contact validation level shall be 10% of the interviews/cases.

The minimum monitoring validation level shall be 5% of the interviews/cases with at least 75% of the whole interview monitored/listened to.

Every fieldworker working on a project should be validated or monitored.

In exceptional cases it can be organizationally impossible to carry out re-contact or monitoring to the required level, or at all, or it may be considered contrary to respondents' interest. In such cases project records shall explain why this is the case and what other steps (e.g. checking data records) have been taken to validate data collection" (ISO 20252:2006).

(4) Data management and processing

This section deals with all activities from data entry to delivery of analysis. Separate sections are devoted to data entry, coding, editing, file management, analysis, delivery and back-up, retention and data security. Clear instruction, verification, documentation and error resolution are the common theme throughout this section. As with data collection mandatory levels of verification are required for some aspects such as data entry and coding.

Example:

„*Data entry verification for paper documents*

A systematic method of verifying a minimum percentage of data entry work shall be carried out on a project or stage/wave. In the case of logic data entry, the minimum total percentage verification per project shall be 5% of entries and for simple data entry, it shall be 10% of entries. Procedures shall ensure that there is a systematic method of verification of each operator's work and the verification shall be undertaken by a second person.

If an individual operator's work contains frequent errors, that individual's work (on the project) shall be 100% verified/re-worked. If necessary, appropriate retraining shall be given to that operator until error rates are acceptable. The effectiveness of the retraining shall be reviewed and documented.

The research service provider shall define the meaning of frequent errors and document that definition." (ISO 20252:2006).

(5) Project documentation

This section lays down the requirements for the complete record of the project.

2 Experiences in statistical institutes

Great Britain

In the UK the consistency issue was addressed by providing detailed definitions and procedures for companies that wished to be assessed to BSI 7911 (British Standard Institution). This consisted of detailed expansion of the content of each section and described how practically these procedures could be put into place and managed. Where they wished to, companies seeking certification had available the services of a number of consultants with experience of both the industry and assessment procedures who could provide support to this process. In parallel detailed guidelines were drawn up to specify procedures for inspection agencies. These were drawn up in consultation with the principal participants in this area and they served as the template for all inspections in the UK. In the UK a separate body, the United Kingdom Accreditation Service (UKAS), oversees the inspection process and their support and guidance was *sought* for this process. Their role is to ensure that the very concerns of consistency and comparability that the industry had identified are met, so their involvement was particularly appropriate. In turn the whole effort was coordinated by an all industry body. This same body also serves as the Technical Committee charged with drafting and updating the standard, and who also provides general support, publicity and information, as well as being the contact with the ISO procedure.

The UK National Audit Office report in 2005 on National Statistics Quality Assurance recommended the use of an audit based approach and this has been endorsed by the UK Statistics Commission Report, Managing the Quality of Official Statistics. In particular the need for independence was stressed: "One of the principles of audit is the requisite independence of the auditor from the area being audited. How this is achieved in relation to statistics is likely to vary with the circumstance". Clearly the Commission is looking at a much wider canvas than the narrow constraints of process quality, but the principle of auditing and independence as being appropriate are clearly stated. For a global standard the necessity and benefits of implementing these principles can be clearly described.

An International approach: Conceptually such an approach could be adopted internationally. It would need to be outside the framework of ISO itself, but the constituent national members of the working party that drafted the standard could oversee such a process with the support of one of the liaison bodies such as European Society of Opinion and Market Researchers (ESOMAR), European Federation of Market Research Organisations (EFAMRO) or World Association of Public Opinion Researchers (WAPOR) who support secretariats with relevant experience. Assessment bodies based in the UK who wish to carry out international inspections have been provided with a specification for inspecting to ISO 20252:2006 that has already been approved by UKAS. This specification has been drafted by the UK industry body and is based on the previous British standard that has already proven to be workable and reliable through time. Within the UK the ISO standard will replace the British standard completely by 2008. Given that a number of UK based companies are amongst the largest global survey research operations it is therefore highly probable that this assessment specification will become the norm for many of the earliest accreditations around the world. This should ensure that the basis for the future growth of the standard internationally has foundations which are solid and proven.

Slovakia

The Statistical office of the Slovak Republic has published a declaration on quality policy and quality management system which is documented here in excerpts:

"Declaration of the Statistical Office of the Slovak Republic on Quality Policy and Quality Management System

The Strategy of the SO SR management for ensuring the quality of statistics

The SO SR management commits to execute the policy of quality, which is based on requests of statistics users and set by rules, principles, recommendations and conclusions of the standard STN EN ISO 9001:2000 and by further relevant documents.

The following tasks ensure successful realisation of the commitment by the SO SR management:

- to set-up and maintain defined quality objectives of the SO SR
- to ensure elaboration, implementation and maintaining of efficient and effective Quality Management System to reach its objectives
- to focus attention on requests of statistics users in all departments of the SO SR
- to ensure implementation of clearly defined processes, which facilitate fulfilment of requests of statistics users and other involved parties and reaching the objectives of quality
- to ensure availability of all necessary resources
- to ensure performance of the Quality Management System by periodical reviews
- to select activities related to quality policy and quality objectives
- to select activities leading to improvement of the Quality Management System

Quality criteria and control

- Criteria of quality of statistics
- Control of quality criteria

Quality of statistics creation process

Quality of statistics is determined by quality of its creation. In the process of creation of statistical data only activities based on following criteria are applied:

- Orientation to user
- Efficiency
- Flexibility
- Transparency
- Integration
- Harmonisation

The management of the SO SR regularly monitors the efficiency of all realised processes and activities leading to accomplishment of declared criteria of statistics quality.

Successful fulfilment of these criteria is possible only if activities and related resources are managed like a process. Therefore the SO SR management develops the Quality Management System based on process approach according to the STN EN ISO 9001:2000 standard.

Implementation of the approach ensures:

- Systematic definition of all necessary (key) activities to achieve desired results
- Clearly specified responsibilities, competencies and duties in relation to the management of key activities
- Measurement and analysis of capacity of key activities
- Identification of interconnection of key activities with/or among individual components of organisation
- Focus on factors like resources, methods and materials, which improve key activities of organisation
- Evaluation of risks, consequences and impacts of activities in relation to users, reporting units and other involved parties.

Conclusion

The above-described Declaration of the SO SR on Policy of Statistics Quality and the Quality Management System expresses the basic tasks of the SO SR management in the field of creation, management and improvement of quality. It also represents the first level of documentation according to the standard STN EN ISO 9001:2000. This basic document, together with other first level documents, is binding for the complete process documentation, obligatory documented procedures, other documentation relevant to definition of activities on the level of individual organisational units of the SO SR, working instructions as well as for other documentation related to assurance of the quality of statistics."

The full text is available at: http://www.statistics.sk/webdata/english/office/deklar/deklar.htm.

Recommended readings

Blyth, B. (2006): Independent, Transparent, Externally Audited: The ISO Approach to Survey Process Quality Control, European Conference on Quality in Survey Statistics (Q2006), Cardiff, United Kingdom, 24-26 April 2006. http://www.statistics.gov.uk/events/q2006/downloads/W15_Blyth.doc

Havermans, J. (2006): Who's s ready for the ISO Norm? http://www.esomar.org/index.php/who-is-ready-for-the-iso-normes.html

International Organization for Standardization (ISO) (2006): Market, opinion and social research – Vocabulary and service requirements (ISO 20256:2006).

Jackson, P. (2006): Assessment of ISO 20252. http://www.mrs.org.uk/standards/downloads/mrqsa_conf_2006.pdf

Statistical Office of the Slovak Republic: Declaration of the SO SR in the field of statistics quality. http://www.statistics.sk/webdata/english/office/deklar/deklar.htm

ANNEX C:
Basic Quality Tools

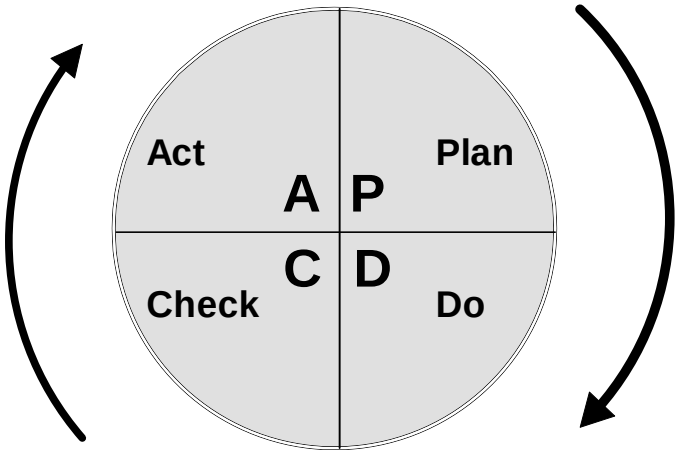
Basic quality tools

This annex briefly describes some basic quality tools relevant for data quality assessments. Many other tools exist and are described in Tague (2005).

1 The Plan-Do-Check-Act cycle – PDCA cycle

Data quality assessment plays an important part in studying products and processes aiming at improving data quality and processes. The PDCA cycle illustrates a way of thinking that can be used in management of all processes and activities.

Figure 19: The PDCA cycle



The cycle starts with planning of an activity that may or may not be a new development (P).The activity is carried out or the process is run (D), then the result is checked (C) by measuring and checking data quality or process variables. As a result, measures to change the process may be taken (A). The cycle starts over next time the process is run.

In the case of the production of a statistics, regular data quality assessment is a part of the checking phase, which also may be denoted the studying (S) phase, underlining the need for analysing quality data. Therefore, one may see that “PDCA” is denoted “PDSA”.

2 Process flow chart

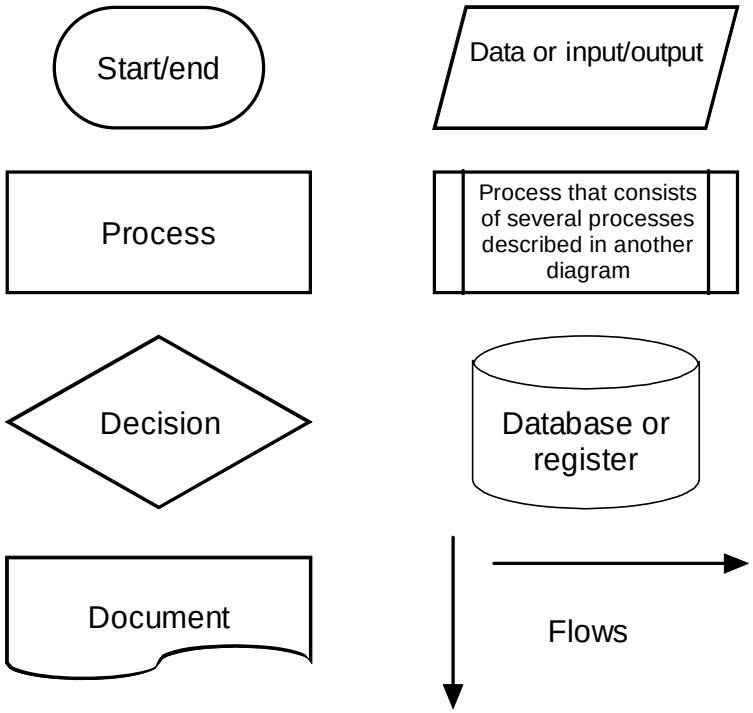
A process flow chart is a tool to map and document processes, showing the dependences between the processes and the respective responsibilities. It can be used for mapping and analysing the processes with possible bottlenecks, unclear responsibilities or to detect redundancies in workflow. It provides a basis for understanding the processes and for identifying key process variables.

There is no internationally agreed standard for process flow charts. However, there is a standard set of symbols that is implemented in MS Office, and hence available through the drawing menus in Word and Power Point. These symbols are also used in software more specifically designed for quality management. The most common symbols for process mapping are shown in figure 20.

These symbols are used in system development as well. Here "program" will often correspond to "process".

The flow chart can be extended with names of people responsible and others linked to each process, with milestones etc.

Figure 20: Some symbols used in process flow charts



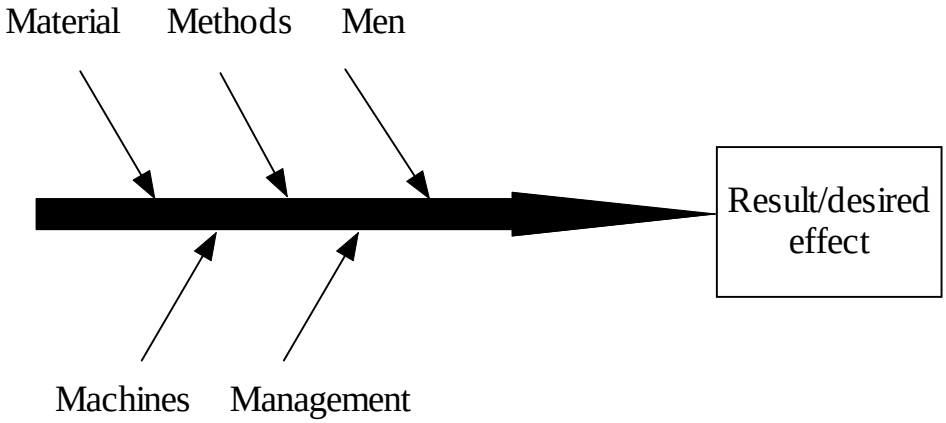
3 Cause and effect diagram

The cause and effect diagram (Ishikawa-diagram or ”fishbone”-diagram) is used during brainstorming for structuring a task with objectives or results that depend on different factors. It is a convenient tool for identifying key process variables.

Often the variables are linked to what we denote the five "M’s”: Material, Methods, Men (=people), Machines and Management, as shown in the general example in figure 21.

There are several other possible headings in the diagram, among these Mother Nature (Environment). For service industries another example of key words is the four “S’s”: Surroundings, Suppliers, Systems and Skills.

Figure 21: Cause and effect diagram



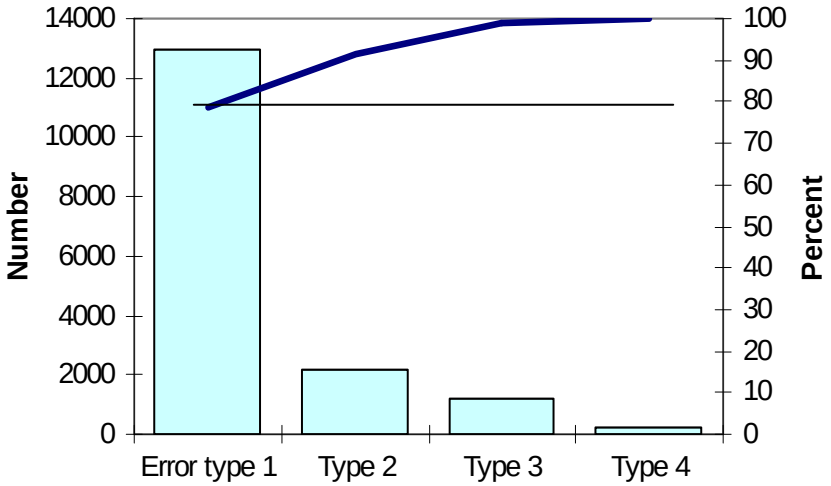
Jones and Lewis (2003) contain several examples of use of cause and effect diagrams for identifying key process variables.

4 Pareto chart

The pareto chart is a graph showing different sources' or groups' contribution to a total effect or error. The sources are grouped by significance. It is often said that 20 percent of the causes contribute to 80 percent of the errors (or the effects). This will of course vary, but the message reflects an important point: to distinguish the vital few from the many small contributions. This is what the identification of key process variables is about!

There are many examples of use of Pareto diagram in quality assurance work, such as in editing of statistics. Here often the number of errors is registered and grouped by cause. Figure 22 shows a general Pareto chart.

Figure 22: Example of typical Pareto chart for edit failures



Jones and Lewis (2003) give several examples of use of Pareto charts.

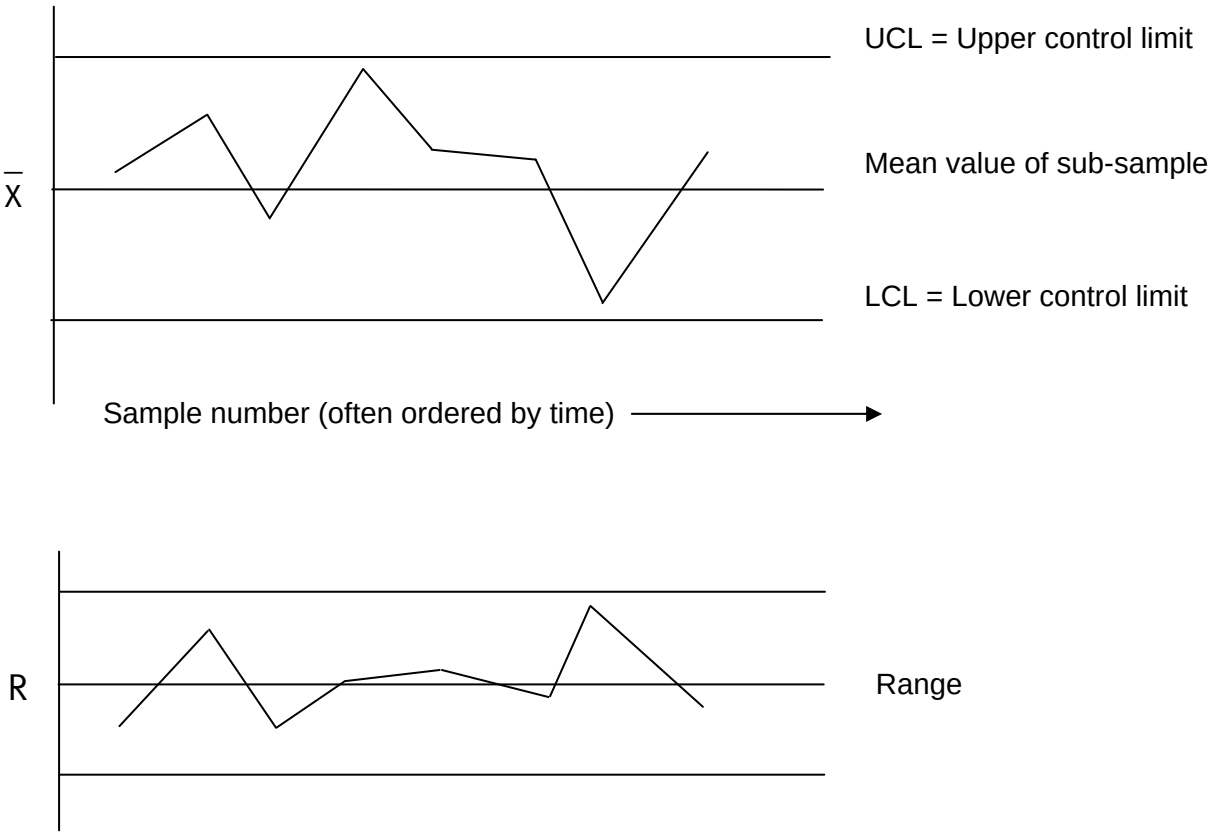
5 Control chart

A control chart is a tool to establish a stable process varying within known limits, and for monitoring the process to detect and eventually avoid unexpected special causes. It can also be used to compare the capabilities (level and variation of key process variables) of a new process with the old one after changes have been done.

The control chart illustrates data quality development, even if work with quality improvements is beyond the scope of data quality assessments.

There are different types of control chart, see for example Bissel (1996). Figure 23 illustrates a general but commonly used one, showing the mean and range of a key process variable (clusters of the variable containing a few, typically 5, observations produced by similar system characteristics, for example close in time).

Figure 23: Variables control chart



Jones and Lewis (2003) and Thomsen et al. (2006) give several examples of use of control charts.

ANNEX D:
Glossary

Glossary

This glossary comprises the most important terms that are used in this handbook as well as aspects related to data quality assessment issues in general. However, it is not exhaustive.

Accessibility	Accessibility refers to the physical conditions in which users can obtain data: where to go, how to order, delivery time, clear pricing policy, convenient marketing conditions (copyright, etc.), availability of micro or macro data, various formats (paper, files, CD-ROM, Internet etc.), etc.
Accuracy	Accuracy in the general statistical sense denotes the closeness of computations or estimates to the exact or true values.
Administrative data	Data originally collected for non-statistical purpose. Complete coverage is the aim. Control of the methods by which the administrative data are collected and processed rests with the administrative agency. In most cases the administrative agency will be a government unit.
Audit	An audit is a systematic, independent and documented process for obtaining audit evidence (records, statements of fact or other information, which are relevant to the audit criteria and verifiable) and evaluating it objectively to determine the extent to which the audit criteria (set of policies, procedures or requirements) are fulfilled.
Clarity	Clarity refers to the data's information environment whether data are accompanied with appropriate metadata, illustrations such as graphs and maps, whether information on their quality is also available (including limitation in use etc.) and the extent to which additional assistance is provided by the National Statistical Institute (NSI).
Coherence	Coherence of statistics is their adequacy to be reliably combined in different ways and for various uses.
Comparability	Comparability is the extent to which differences between statistics from different geographical areas, non-geographical domains, or over time, can be attributed to differences between the true values of the statistics.
Completeness	Completeness is the extent to which all statistics that are needed are available. It is usually described as a measure of the amount of available data from a statistical system compared to the amount that was expected to be obtained.
Editing, Edits	Data editing is the application of checks that identify missing, invalid or inconsistent entries or that point to data records that are potentially in error.
Error	In general, a mistake or error in the colloquial sense. There may, for example, be a gross error or avoidable mistake; an error of reference, when data concerning one phenomenon are attributed to another; copying errors; an error of interpretation. In a more limited sense the word error is used in statistics to denote the difference between an occurring value and its true or expected value. There is here no imputation of mistakes on the part of a human agent; the deviation is a chance effect. In this sense we have, for example, errors of observations, errors in equations, errors of the first and second kinds in the testing hypothesis, and the error band surrounding an estimate; and also the Normal curve of errors itself.

Focus groups interviews	Interviewers and correspondents meet in a group which is moderated to identify any problems or successes of the survey and to suggest how potential problems could be solved.
Frame	The frame consists of previously available descriptions of the objects or material related to the physical field in the form of maps, lists, directories, etc., from which sampling units may be constructed and a set of sampling units selected; and also information on communications, transport, etc., which may be of value in improving the design for the choice of sampling units, and in the formation of strata, etc.
In-depth interviews	A conversation conducted by trained staff that usually collects specific information when not much is known about a population to get preliminary ideas from the participants.
Interviewer error	Interviewer errors are associated with effects on respondents' answers stemming from the different ways that interviewers administer the same survey. Examples of these errors include the failure to read the question correctly (leading to response errors by the respondent), delivery of the question with an intonation that influences the respondent's choice of answer, and failure to record the respondent's answer correctly.
Item non-response	Item non-response occurs when a respondent provides some, but not all, of the requested information, or if the reported information is not usable.
Item response rate	The item response rate is the ratio of the number of eligible units responding to an item to the number of responding units eligible to have responded to the item.
Key process variables	They can vary with each repetition of the process and have the largest effect on critical product characteristics, i.e. those characteristics that best indicate the quality of the product.
Labelling	The labelling method means that a label is attached to some statistics. The label has a message about these statistics, a message that in this handbook context is related to quality and quality assessment
Measurement error	Measurement error refers to errors in survey responses arising from the method of data collection, the respondent, or the questionnaire (or other instruments). It includes the error in a survey response as a result of respondent confusion, ignorance, carelessness, or dishonesty; the error attributable to the interviewer, perhaps as a consequence of poor or inadequate training, prior expectations regarding respondents' responses, or deliberate errors; and error attributable to the wording of the questions in the questionnaire, the order or context in which the questions are presented, and the method used to obtain the responses.
Non-response	Non-response is a form of non observation present in most surveys. Non-response means failure to obtain a measurement on one or more study variables for one or more elements k selected for the survey. The term encompasses a wide variety of reasons for non observation: "impossible to contact", "not at home", "incapacity", "hard core refusal", "inaccessible", "unreturned questionnaire", and others. In the first two cases contact with the selected element is never established.
Out of scope units	Units that should not be included in the sampling frame because they do not belong to the target population in the reference period. If enumerated, they cause over-coverage.

Outlier	In a sample of n observations it is possible for a limited number to be so far separated in value from the remainder that they give rise to the question whether they are not from a different population, or that the sampling technique is a fault. Such values are called outliers.
PDCA cycle (also: PDSA cycle)	Tool for quality assessment. The cycle starts with planning of an activity that may or may not be a new development (P).The activity is carried out or the process is run (D), then the result is checked (C) by measuring and checking data quality or process variables. As a result, measures to change the process may be taken (A). The cycle starts over next time the process is run.
Peer review	A special kind of external audit, carried out e.g. by a NSI for another NSI (=peers). In general, it is less formal than an audit. It aims rather at assessing the general quality than at controlling the conformity with an external quality standard.
Punctuality	Punctuality refers to the time lag existing between the actual delivery date of data and the target date when it should have been delivered, for instance, with reference to dates announced in some official release calendar, laid down by regulations or previously agreed among partners.
Quality control survey	A replicated survey carried out on a small scale by very experienced staff in order to obtain some “zero-default” results with which the actual results of the survey can be compared.
Quality index	A one-dimension synthetic information on quality, possibly calculated as a weighted mean of all available quality indicators.
Quality indicator	Quality indicator is a specific and measurable element that can be used to characterise the quality of statistics.
Quality report	Quality report is a report conveying information about the quality of a statistical product or survey.
Register	(Administrative) Registers are a sub-group of → administrative records. If an administrative record consists of unit-level data, it can be called a register. Administrative registers come from administrative sources and become statistical registers after passing through statistical processing in order to make it fit for statistical purposes (production of register based statistics, frame creation, etc.).
Relevance	Relevance is the degree to which statistics meet current and potential users’ needs. It refers to whether all statistics that are needed are produced and the extent to which concepts used (definitions, classifications etc.) reflect user needs.
Self-assessment	Self-assessment is a comprehensive, systematic and regular review of an organisation’s activities and results referenced against a model/framework, carried out by the organisation itself.
Timeliness	Timeliness of information reflects the length of time between its availability and the event or phenomenon it describes.

Unit response rate	The ratio, expressed in percentage of the number of interviews to the number of eligible units in the sample. The weighted response rate calculates the ratio using the inverse probability of inclusion in the sample as a weight for each unit. In some occasions a value that reflects the importance of the unit is also used as a weighting factor (like size of workforce for establishments).
User Surveys	A survey aiming at assessing the satisfaction or the perception of the users, normally as a basis for improvement actions.

Abbreviations

Abbreviations

AD	Administrative Data
AS	Administrative Source
BSI	British Standards Institution
BQI	Basic Quality Information
CAF	Common Assessment Framework
CATI	Computer Assisted Telephone Interview
CBM	Current Best Method
CNIS	National Council for Statistical Information, France (Abbreviation of French Commission National pour la Société de l'Information)
CV	Coefficient of variation
DatQAM	Data Quality Assessment Methods and Tools
DESAP	Development of a Self Assessment Programme
DMAIC	Define, Measure, Analyze, Improve, Control
DQAF	Data Quality Assessment Framework
EFAMRO	European Federation of Market Research Organisations
EFQM	European Foundation for Quality Management
e.g.	for example (Abbreviation of Latin exempli gratia)
ESOMAR	European Society of Opinion and Market Researchers
ESS	European Statistical System
etc.	and so on (Abbreviation of Latin et cetera)
Eurostat	Statistical Office of the European Communities
Euro-IND	European and national short term indicators
FAO	Food and Agriculture Organization
GDP	Gross Domestic Product
GNI	Gross National Income
i.e.	that is (Abbreviation of Latin id est)
IMF	International Monetary Fund
INE-PT	National Statistics Institute Portugal (Abbreviation of Portuguese Instituto Nacional de Estadística)
ISO	International Organization for Standardization
Istat	Italian Statistical Office (Abbreviation of Italian Istituto nazionale di statistica)
IT	Information Technology
JUSE	Union of Japanese Scientists and Engineers
LEG	Leadership Expert Group
NACE	Nomenclature of economic activities (Abbreviation of French Nomenclature générale des activités économiques dans les Communautés Européennes)
NSI	National Statistical Institute
OECD	Organization for Economic Cooperation and Development
ONS	Office for National Statistics, United Kingdom
PDCA also PDSP	Plan-Do-Check-Act; Plan-Do-Study-Act
QIS	Quarterly Investment Statistics
QM	Quality Management
QRD	Quality Report Database
RADAR	Results, Approach, Deployment, Assessment, Review
R&D	Research and Development
SAS	Statistical Analysis System

SIDI	Information System for Survey Documentation (Abbreviation of Italian sistema informativo di documentazione)
SORS	Statistical Office of the Republic of Slovenia
SO SR	Statistical Office of the Slovak Republic
SQR	Standard Quality Report
SVC	Statistical Value Chain
SYSQUAST	System of Quality Measurement for Statistical Products
TQM	Total Quality Management
UK	United Kingdom
UKAS	United Kingdom Accreditation Service
UN	United Nations
U.S.	United States
WAPOR	World Association of Public Opinion Researchers

List of Figures and Tables

List of Figures

List of Tables

References

References

Bergdahl, M. and Lyberg, L. (2004): Quality Management at Statistics Sweden. Current Work and the Future. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.

Biemer, P. and Caspar, R. (1994): Continuous Quality Improvement for Survey Operations: Some general principles and applications. In: Journal of Official Statistics, 10, pp. 307-326.

Biemer, P. and Lyberg, L. (2003): Introduction to Survey Quality. Hoboken, N.J.: Wiley.

Birch, J. (2005): The Quality Measurement and Reporting Project at the Office for National Statistics, GSS Methodology Conference, London, United Kingdom, 27 June 2005. http://www.statistics.gov.uk/events/gss2005/downloads/PaperC1.doc

Bissel, D. (1996): Statistical Methods for SPC and TQM. Chapman & Hall 1996.

Blanc, M.; Radermacher, W. and Körner, T. (2002): Quality and Users. In: Eurostat (ed.): Quality in the European Statistical System – The way forward, Luxembourg, pp. 45-60.

Blyth, B. (2006): Independent, Transparent, Externally Audited: The ISO Approach to Survey Process Quality Control, European Conference on Quality in Survey Statistics (Q2006), Cardiff, United Kingdom, 24-26 April 2006. http://www.statistics.gov.uk/events/q2006/downloads/W15_Blyth.doc

Brackstone, G. (1999): Managing Data Quality in a Statistical Agency. In: Survey Methodology 25, pp. 139-149.

Brancato, G. et al. (2004): Standardising, Evaluating and Documenting Quality: the Implementation of ISTAT Information System for Survey Documentation – SIDI. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.

Brancato, G. et al. (2006): Assessing Quality through the Collection and Analysis of Standard Quality Indicators: The ISTAT Experience, European Conference on Quality in Survey Statistics (Q2006), Cardiff, United Kingdom, 24-26 April 2006. http://www.statistics.gov.uk/events/q2006/downloads/T16_Signore.doc

Brancato, G. (2007): Review on IT and Statistical Auditing Procedures, ISTAT Italy.

Brion, P. (2004): The Management of Quality in French Business Surveys. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.

Burg, T. (2004): Quality Reports at Statistics Austria. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004. http://www.statistik.at/

Burg, T. (2005): Quality Measurement at Statistics Austria, “Managing Modern National Statistical Systems in Democratic Societies”, Tacis High Level-Seminar, Vienna, Austria, 31 August-2 September 2005.

Bushery, J. M. (2004): How the U.S. Census Bureau Uses the Capability Maturity Model ® as a Roadmap to Process Improvement. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.

Bushery, J. M.; McGovern, P.D.; Marck, P.S. and Hogan, H. R. (2006): Quality Audits at the U.S. Census Bureau. Paper presented at the European Conference on Quality in Survey Statistics (Q2006), Cardiff, United Kingdom, 24-26 April 2006.

Cassel, C. (2006): Measuring Customer Satisfaction, a methodological guidance. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143266/CUSTOMER%20SATISFACTION%20SURVEYS_SE_2006_EN_1.PDF

Cassel, C.; Järnbert, M.; Holzer, W.; Barbieri, G. and Tiddi, A. (2003): State-of-the-art regarding planning and carrying out Customer/User Satisfaction Surveys in NSIs. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143266/2ESS_CUSTOMER_SATISFACTION_SURVEY_SE-AT-IT_2003_EN_0_1.PDF

Colledge, M. (2006): Quality Frameworks. Implementation and Impact. Paper presented at the International Conference on Data Quality in International Organisations, Newport, Wales, 27-28 April 2006.

Colledge, M. and Dalén, J. (2006): Improving Statistical Quality Assessment in Ten Counties. Paper presented at the European Conference on Quality in Survey Statistics (Q2006), Cardiff, United Kingdom, 24-26 April 2006. http://www.statistics.gov.uk/events/q2006/downloads/W08_Colledge.doc

Commission of the European Communities (2003): Structural indicators, Communication from the Commission, COM (2003) 585, Brussels. http://europa.eu/eur-lex/en/com/cnc/2003/com2003_0585en01.pdf

Commission of the European Communities (2004): Eurostat Quality Profile for Structural Indicators. Commission staff working paper. Brussels, 20.10.2004. SEC (2004) 1286, p. 33.

Cornish, J. and Templeton, R. (2001): Production and Presentation of Official Statistics: Strategies for Managing Quality. Paper presented at the 2001 ISI Conference in Seoul, Republic of Korea.

Deming Prize Committee (2006): The Deming Prize Guide 2006 for Overseas. Tokyo, The Union of Japanese Scientists and Engineers (JUSE).

Deming, W. E. (1986): Out of the Crisis. Cambridge: Cambridge University Press.

Desrosières, A. (2000): Measurement and its Uses: Harmonisation and Quality in Social Statistics. In: International Statistical Review, 68, pp. 173-187.

Di Fonzo, T. (2005): The OECD Project on revision analysis, first elements for discussion. Technical Report, OECD Short-term Economic Statistics Expert Group (STESEG).

Eiderbrant-Nilsson, G. (2004): Quality Audits at Statistics Sweden. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.

Elvers, E. and Rosén, B. (1999): Quality Concept for Official Statistics. In: Encyclopedia of Statistical Sciences, Update Volume 3, A Wiley-Interscience Publication, John Wiles & Sons, Inc. New York – Chiclster – Weinheim – Brisbane – Singapore – Toronto, pp. 621-629.

Elvers, E. (2004): Quality and Metadata – Some Connections and Reflections. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.

European Foundation for Quality Management (1999): Assessing for Excellence – A practical guide for self-assessment, Brussels.

Eurostat (2001a): Quality Declaration of the European Statistical System. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47141301/DECLARATIONS.PDF

Eurostat (2002): Quality in the European Statistical System – The Way Forward. European Commission and Eurostat, Luxembourg. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143266/ESS_QUALITY_RECOMMENDATIONS_2002_EN_0_1.PDF

Eurostat (2003a): Definition of Quality in Statistics. Methodological Documents, Working Group "Assessment of quality in statistics", Luxembourg, 2-3 October 2003. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47141301/DEFINITION_2.PDF

Eurostat (2003b): Handbook "How to make a Quality Report". Methodological Documents, Working Group "Assessment of quality in statistics", Luxembourg, 2-3 October 2003. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143233/HOWTOMAKEQUALITYREPORT_0.PDF#search=%22Handbook%20%E2%80%9CHow%20to%20make%20a%20Quality%20Report%E2%80%9D.%20%22

Eurostat (2003c): Glossary "Quality in Statistics". Methodological Documents, Working Group "Assessment of quality in statistics", Luxembourg, 2-3 October 2003. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47141301/GLOSSARY_1.PDF

Eurostat (2003d): Standard Quality Report. Methodological Documents, Working Group "Assessment of quality in statistics", Luxembourg, 2-3 October 2003. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143233/STANDARD_QUALITY_REPORT_0.PDF#search=%22Standard%20Quality%20report%202003%20Eurostat%22

Eurostat (2003e): Quality Report on Labour Force Statistics. Working Group "Assessment of quality in statistics", Luxembourg, 2-3 October 2003.

Eurostat (2003f): Standard Quality Indicators, Producer-Oriented. Working Group "Assessment of quality in statistics", Luxembourg, 2-3 October 2003.

Eurostat (2003g): Structural Indicators – Quality Profile and Long-term Assessment. Working Group "Assessment of quality in statistics", Luxembourg, 2-3 October 2003.

Eurostat (2003h): Quality assessment of administrative data for statistical purposes. Working Group "Assessment of quality in statistics", Luxembourg, 2-3 October 2003. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143266/QUALITY%20ASSESSMENT%20OF%20ADMINISTRATIVE%20DATA%20FOR%20STATISTICAL%20PURPOSES.PDF

Eurostat (2003i): DESAP – Development of a Self Assessment Programme, The European Self Assessment Checklist for Survey Managers.

Eurostat (2004): Principal European Economic Indicator Pilot Monitoring System Proposal. Working Group Document Doc: 2004/FROCH/070.

Eurostat (2005a): 2004 LEG Implementation Status Report. Working Group „Quality in Statistics", Seventh meeting, Eurostat, Luxembourg, 23-24 May 2005.

Eurostat (2005b): European Statistics Code of Practice. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47141301/VERSIONE_INGLESE_WEB.PDF

Eurostat (2005c): European Statistics Code of Practice, Self Assessment Questionnaire. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/PGE_DS_QUALITY_05/CoP%20Questionnaire%20FINAL%20EN1.pdf

Eurostat (2005d): Standard Quality Indicators. Working Group "Quality in statistics", Luxembourg, 23-24 May 2005. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143233/STANDARD%20QUALITY%20INDICATORS.PDF#search=%22standard%20quality%20indicators%20Eurostat%22

Eurostat (2006a): Euro-IND State of affairs, July 2006. http://epp.eurostat.ec.europa.eu/cache/ITY_PUBLIC/QR-072006/EN/QR-072006-EN.PDF

Eurostat (2006b): Quality Reporting: Quality Barometer. Working Group "Quality in Statistics", Luxembourg, 27-28 February 2006.

Eurostat (2006c): Report on the results of the first self-assessments carried out by the statistical authorities of the European Statistical System against the principles and indicators of the European Statistics Code of Practice, May 2006. http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/PGE_DS_QUALITY_05/EFC%20REPORT%20TO%20EFC%20ESS%20SELF-ASSESSMENTS.PDF

Finland (2006): Basic and Quality Criteria for Official Statistics of Finland. Advisory Board of Official Statistics of Finland. TK-41-1227-06.

Full, S. and Jones, N. (2002): Towards a Framework for Quality Measurement and Reporting within the Office for National Statistics (ONS). GSS Methodology Conference, 8 July 2002. http://www.statistics.gov.uk/methods_quality/gss_method_conf_2002/downloads/Jones_Full.doc

Garonna, P. and Luige, T (2001): Quality of the European Statistical System. The Statistical Dividend of European Integration. Paper presented at the International Conference on Quality in Official Statistics (Q2001), Stockholm, Sweden, 14-15 May 2001.

Grais, B. (1999): Statistical Harmonisation and Quality. In The Future of European Social Statistics, Eurostat, Luxembourg.

Grünewald, W. and Körner, T. (2005): Quality on its Way to Maturity. Results of the European Conference on Quality and Methodology in Official Statistics (Q2004). In: Journal of Official Statistics, 21, pp. 747-759.

Grünewald, W. and Lindén, H. (2001). Quality measurement - Eurostat experiences. Paper presented at the Statistics Canada Symposium 2001, Hull, Canada, 17-19 October 2001.

Hahn, M. (2004): Quality Profile for Structural Indicators. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.

Hahn, M. and Lindén, H. (2006): The European Statistics Code of Practice for a High Quality European Statistical System. Paper presented at the European Conference on Quality in Survey Statistics (Q2006), Cardiff, United Kingdom, 24-26 April 2006.

Hahn, M. and Willeke, C. (2006): Labelling of statistical information: Some introductory considerations. Presented at the Conference on Data Quality for International Organizations, Newport, Wales, United Kingdom, 27-28 April 2006.

Havermans, J. (2006): Who's s ready for the ISO Norm?
http://www.esomar.org/index.php/who-is-ready-for-the-iso-normes.html

Hedlin, D.; Dale, T.; Haraldsen, G. and Jones, J. (2005): Developing Methods for Assessing Perceived Response Burden. Final Report. Eurostat.

Holzer, W. (2004): Planning and Carrying out User/Customer Satisfaction Surveys by NSIs in EU and EEA Countries. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.

International Monetary Fund (2003): DQAF - Data Quality Assessment Framework.

International Monetary Fund (2003): Data Quality Assessment Framework and Data Quality Program.
http://www.imf.org/external/np/sta/dsbb/2003/eng/dqaf.htm

International Monetary Fund (2005): Sixth Review of the Fund's Data Standards Initiatives.
http://www.imf.org/external/np/pp/eng/2005/070105s.pdf

International Organization for Standardization (2000a): Quality management systems – Requirements (ISO 9001:2000).

International Organization for Standardization (2000b): Quality management systems – Guidelines for performance improvements (ISO 9004:2000).

International Organization for Standardization (2002): Guidelines for quality and/or environmental management systems auditing (ISO 19011:2002).

International Organization for Standardization (2003a): Quality management systems – Guidelines for configuration management (ISO 10007:2003).

International Organization for Standardization (2003b): Measurement management sys tems – Requirements for measurement processes and measuring equipment (ISO 10012:2003).

International Organization for Standardization (2004): Quality management – Customer satisfaction – Guidelines for complaints handling in organizations (ISO 10002:2004).

International Organization for Standardization (2005): Quality management systems – Fundamentals and vocabulary (ISO 9000:2005).

International Organization for Standardization (ISO) (2006): Market, opinion and social research – Vocabulary and service requirements (ISO 20256:2006).

International Organization for Standardization: TC 225. http://www.iso.ch/iso/en/stdsdevelopment/tc/tclist/TechnicalCommitteeDetailPage.TechnicalCommitteeDetail?COMMID=5613

Jackson, P. (2006): Assessment of ISO 20252. http://www.mrs.org.uk/standards/downloads/mrqsa_conf_2006.pdf

Jenkinson, G. (2006): Standard Quality Reports and their Applications for the National Accounts at the ONS, paper presented at the European Conference on Quality in Survey Statistics (Q2006), Cardiff, United Kingdom, 24-26 April 2006. http://www.statistics.gov.uk/events/q2006/downloads/T16_Jenkinson.doc

Jones, N. and Lewis, D. (eds, with Aitken, A.; Hörngren, J. and Zilhão M. J.) (2003): Handbook on improving quality by analysis of process variables. Final Report, Eurostat.

Jones, N. and Lewis, D. (2004): Developing a Handbook on Improving Quality by Analysis of Process Variables. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.

Juran, J. M. and Godfrey, A. B. (1999): Juran's Quality Handbook, 5th, New York, McGraw-Hill, 1999.

Kalton, G.; Winglee, M.; Krawchuk, S. and Levine, D. (2000): Quality Profile for SASS Rounds 1-3: 1987-1995. U.S. Department for Education National Center for Education Statistics. http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2000308

Karlberg, M. and Probst, L. (2004): The Road to Quality. The Implementation of the Recommendations of the LEG on Quality in the ESS. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.

Körner, T. (2004): Data Quality Management in the European Statistical System. Paper presented at the International Seminar on Non-Sampling Errors in the Surveys and Censuses of Official Statistics, Seoul, Republic of Korea, 22 October 2004. http://iscd.nso.go.kr/data/workdata/311ThomasKoerner_1.pdf

Körner, T. and Nimmergut, A. (2003): Development of a Self Assessment Programme. Final report. Wiesbaden, Federal Statistical Office Germany.

Laiho, J. and Nimmergut, A. (2004): Using Self-Assessments for Data Quality Management – DESAP Experiences. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.

Lindén, H. (2006): First results from the in-depth surveys on quality assurance frameworks and quality reporting. Paper presented at the Conference on Data Quality for International Organisations, Newport, Wales, 27-28 April 2006.

Lindén, H. and Papageorgiou, H. (2004): Standard Quality Indicators. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004.

Lyberg, L. et al. (2001): Summary Report from the Leadership Group (LEG) on Quality, Proceedings of the International Conference on Quality in Official Statistics, Stockholm, Sweden, 14-15 May 2001.
http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143266/ESS_QUALITY_RECOMMENDATIONS_2002_EN_0_1.PDF

Morganstein, D. and Marker, D. (1997): Continuous Quality Improvement in Statistical Organization, In: Lyberg, L.; Biemer, P.; Collins, M.; de Leeuw, E.; Dippo, C.; Schwartz, N. and Trewin, D. (eds): Survey Measurement and Process Quality, New York: Wiley, pp. 475 – 500.

Murgia, M. and Simeoni, G. (2005): Improving the Quality of the Assisted Coding of Occupation in CATI Surveys through Control Charts, In: Zani, S. and Cerioli, A. (eds): Classification and Data Analysis 2005. Book of short papers. Meeting of the Classification and Data Analysis Group of the Italian Statistical Society, Parma, Italy, 6-8 June 2005, pp. 445-448.

New Zealand (2004): Official Statistics – A Recognisable and Enduring National Resource. UN/ECE CES/2004/18. This presentation by Brian Pink has been published in 2004 in the Statistical Journal of the United Nations Economic Commission for Europe, 21, pp. 231-239.

New Zealand (2007): Statement of Principles for Tier 1 Statistics Producers.
http://www.statisphere.govt.nz/default.htm

OECD (2003): Quality Framework and Guidelines for OECD Statistical Activities. Paris, OECD Document STD/QFS(2003)1.

ONS (2004): Guidelines for Measuring Statistical Quality.

ONS (2006): Guidelines for Measuring Statistical Quality, ONS, London, 2006.
http://www.statistics.gov.uk/downloads/theme_other/Guidelines_Subject.pdf

ONS (2007): Summary quality report for Internal Migration Releases.
http://www.statistics.gov.uk/about/data/methodology/quality/downloads/SQR_InternationalMigration.pdf

ONS et al. (1997): Model Quality Report in Business Statistics. 4 vols., Eurostat, Luxembourg.

Pagani, F. (2002): Peer Review: a tool for co-operation and change, an analysis of an OECD working method. OECD.

Sæbø, H. V. (2003): Quality Issues at Statistics Norway. In: Journal of Official Statistics, 19, pp. 287-303.

Sæbø, H. V. (2006): Systematic Quality Work in Official Statistics – Theory and Practice. Paper presented at the European Conference on Quality in Survey Statistics (Q2006), Cardiff, United Kingdom, 24-26 April 2006.

Seljak, R. and Zalatel, M. (2004): Measurement of Data Quality in the case of Short Term Statistics. Paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), Mainz, Germany, 24-26 May 2004. http://www.oecd.org/dataoecd/6/8/35238995.pdf

Statistical Data and Metadata Exchange (2006): SDMX Content-Oriented Guidelines: Framework. http://www.sdmx.org/

Statistics Canada (1985): Policy Manual 1.2, Policy on Responses to Erroneous Statements in the Media, December 1985.

Statistics Canada (1986): Policy Manual 2.1, Policy on Highlights of Publications, October 1986.

Statistics Canada (1986): Policy Manual 2.2, Policy on Estimates with Future Reference Dates, March 1986.

Statistics Canada (1992): Policy Manual 3.6, Policy on Fees and Charges for Products and Services, November 1992.

Statistics Canada (1993): Policy Manual 3.3, Policy on Statistics Canada's Daily (Official Release), December 1993.

Statistics Canada (1994): Policy Manual 2.8, Policy on the Development, Testing and Evaluation of Questionnaires, January 1994.

Statistics Canada (1994): Policy Manual 3.10, Policy on the Registration of Information Products and Services, May 1994.

Statistics Canada (1994): Policy Manual 5.12, Policy on Use of Software Products, August 1994.

Statistics Canada (1995): Policy Manual 2.5, Policy on the Review of Information Products, January 1995.

Statistics Canada (1998): Policy Manual 1.1, Policy on Informing Survey Respondents, September 1998.

Statistics Canada (1998): Policy Manual 2.10, Policy on Standards, July 1998.

Statistics Canada (2000): Policy Manual 2.3, Policy on Informing Users of Data Quality and Methodology, March 2000.

Statistics Canada (2002): Statistics Canada's Quality Assurance Framework. Ottawa: Statistics Canada.

Statistics Canada (2003): Statistics Canada Quality Guidelines. 3rd edition, Ottawa: Statistics Canada.

Statistical Office of the Slovak Republic: Declaration of the SO SR in the field of statistics quality.
http://www.statistics.sk/webdata/english/office/deklar/deklar.htm

Statistics South Africa (2006): Draft Data Quality Framework 0001: South African Statistical Quality Assessment Framework (SASQAF), Statistics South Africa, January 2006.

Statistics Sweden (2006a): Rapporteur Report on Quality Frameworks and the Link with Institutional Frameworks and Principles. Paper presented at the 54th session of the Conference of European Statisticians, Paris, France, 13-15 June 2006.

Statistics Sweden (2006b): Sufficient quality and criteria for official statistics. Guidelines issued by the Council for Official Statistics 2006:1. Statistics Sweden, 2006.

Statistisches Bundesamt (2006): Qualitätsstandards der amtlichen Statistik [Quality Guidelines of Official Statistics]. 2. Auflage, Wiesbaden, Statistische Ämter des Bundes und der Länder.

Tague, N. R. (2005): The Quality Toolbox. 2nd edition. Milwaukee: Quality Press.

Thomsen, I.; Kleven, Ø.; Zhang, L. and Wang, J. H. (2006): Coping with decreasing response rates in Statistics Norway – Recommended practice for reducing the effect of nonresponse. Reports 2006/29, Statistics Norway 2006.

U.S. Census Bureau (2002): 2002 Economic Census: Business Expenses
Appendix C. Methodology.
http://www.census.gov/econ/census02/text/bes/besmethodology.htm

United Kingdom (2007): The National Statistics standard.
http://www.statistics.gov.uk/about/national_statistics/default.asp
See the Introduction to National Statistics (2004): *and the Code of Practice.*

UN (2004): Fundamental Principles of Official Statistics.
http://unstats.un.org/unsd/statcom/doc04/2004-21e.pdf

UN Statistical Commission (2005): Principles Governing International Statistical Activities.
http://unstats.un.org/unsd/methods/statorg/Principles_stat_activities/principles_stat_activities.pdf

Zilhão, *M. J. et al. (2003):* State of the Art Concerning Auditing Activity in National Statistical Institutes. Final Report, Eurostat, Luxembourg.
http://epp.eurostat.ec.europa.eu/cache/ITY_PUBLIC/G0-LEG-20030930/DE/G0-LEG-20030930-DE.PDF