R&D Report

Research – Methods – Development 1997:2

Statistiska centralbyrån Statistics Sweden

Quality Concept for Official Statistics

Entry in the forthcoming update of
the Encyclopedia of Statistical Sciences, Wiley & Sons

Eva Elvers and **Bengt Rosén**

R&D Report
Research - Methods - Development 1997:2

Quality Concept for Official Statistics - Entry in the forthcoming update of the Encyclopedia of Statistical Sciences, Wiley & Sons

Statistics Sweden
1997

Från trycket	September 1997
Ansvarig utgivare	Lars Lyberg
Producent	Statistiska centralbyrån, utvecklingsavdelningen

Förfrågningar	Eva Elvers, ES/SES telefon 08-783 47 14 telefax 08-783 45 76
	Bengt Rosén, ES/LEDN telefon 08-783 44 90 telefax 08-667 77 88

ISSN 0283-8680
Printed in Sweden
SCB-Tryck, Örebro, 1997

INLEDNING

TILL

R & D report : research, methods, development / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1988-2004. – Nr. 1988:1-2004:2.
Häri ingår Abstracts : sammanfattningar av metodrapporter från SCB med egen numrering.

Föregångare:

Metodinformation : preliminär rapport från Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån. – 1984-1986. – Nr 1984:1-1986:8.

U/ADB / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1986-1987. – Nr E24-E26

R & D report : research, methods, development, U/STM / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1987. – Nr 29-41.

Efterföljare:

Research and development : methodology reports from Statistics Sweden. – Stockholm : Statistiska centralbyrån. – 2006-. – Nr 2006:1-.

R & D Report 1997:2. Quality concept for official statistics. Entry in the forthcoming update of the Encyclopedia of statistical sciences, Wiley & Sons / Eva Elvers; Bengt Rosen.
Digitaliserad av Statistiska centralbyrån (SCB) 2016.

urn:nbn:se:scb-1997-X101OP9702

R&D Report

Research – Methods – Development 1997:2

Quality Concept for Official Statistics

Entry in the forthcoming update of
the Encyclopedia of Statistical Sciences, Wiley & Sons

Eva Elvers and **Bengt Rosén**

Quality Concept for Official Statistics

Entry in the forthcoming update of
the Encyclopedia of Statistical Sciences, Wiley & Sons

Eva Elvers and **Bengt Rosén**

Background for and scope of the report

The authors, together with Chris Denell, worked earlier out a Statistics Sweden co-ordination report: "Kvalitetsbegrepp och riktlinjer för kvalitetsdeklaration av officiell statistik, Meddelande i samordningsfrågor 1994:3" (in English: "Quality definition and recommendations for quality declarations of official statistics"), henceforth referred to as MIS 1994:3. That document was translated into English by Eurostat.

The Eurostat version of MIS 1994:3 came under the eyes of prof. Samuel Kotz, editor of the Encyclopedia of Statistical Sciences, Wiley & Sons. Work on updating the Encyclopedia has been in progress since long. Meaning that "quality concept for official statistics" deserves a place in a modern Encyclopedia, prof. Kotz invited us to work out an entry on the topic. So we did, and this report comprises the result.

The relation between the present report and MIS 1994:3 is as follows. The MIS provides the main background for the quality concept presented here. However, we have also been influenced by (i) criticism of some of the MIS 1994:3 ideas, (ii) user requests encountered in Eurostat work on quality issues. As a consequence, the quality concepts in MIS 1994:3 and this paper differ to some extent. Hence, the report should not be read as stating Statistics Sweden's present quality concept, which is that in MIS 1994:3. It conveys the authors' present views on a good quality concept for official statistics. Our main incentives for making the contents available to others are: (i) to contribute to a never-ending discussion of quality issues for official statistics, (ii) to provide material for tutorial purposes.

CONTENTS

Quality Concept for Official Statistics

1 Quality – general view

In everyday language ***quality*** refers to where, on a scale bad - good - excellent, a user places a certain product with regard to its intended use and in comparison with similar products. Sometimes the word "quality" is given a positive value, and is taken as a synonym to "good quality". This makes the notion somewhat difficult to handle, and different definitions have been used.

Even if the definition of quality has varied over time, quality improvement, control, assurance etc. have always concerned producers of goods and services. The presently dominating approach to quality issues is based on the notion ***total quality***, which has the following main ingredients.

(i) The ***user shall be in focus***. A product's quality is determined by (existing and potential) users' opinions of the product and its fitness for their use purposes.

(ii) The quality concept should reflect ***all aspects of a product*** that affect users' views on how well the product meets their needs and expectations.

With this definition quality has a ***descriptive*** meaning for the producer. It lists the aspects of the product that users take into account when judging how well it satisfies their needs and expectations. The producer's quality concept should not take a standpoint on whether the product is of good or bad quality in any absolute sense.

Quality assessment is left to the users, who are entitled to have subjective opinions on whether the quality is good or bad. Their assessments do not depend on the product alone, but on a combination of product and use purpose. A certain product may be judged to be of good quality in one application and bad in another. For the producer it is, of course, essential to learn about users' quality opinions, since they constitute the basis for quality work aiming at higher quality, in the sense "greater user satisfaction".

2 Quality of Official Statistics

In the official statistics context the core part of a "product" consists of ***statistics***, i.e. estimates of ***statistical characteristics***. Such characteristics are numeric values which summarize, via some statistical ***measure*** (total, mean, median, etc.), individual ***variable*** values for the ***units*** (households, enterprises, farms, etc.) in a specific ***group***. The total collection of units of interest is called the ***population***. In most surveys the interest does not only concern statistics for the entire population but also for different subgroups, called ***study domains***.

We speak of "estimates" not only when the statistics emanate from sample surveys, but also when they come from total enumeration surveys. In the latter case one should ideally achieve exactly correct figures, but reality is seldom ideal. Surveys are subject to various kinds of disturbances. Therefore, statistical characteristics are referred to as ***target characteristics***.

Quality considerations of statistics may relate to "statistics products" of different scope, from a single figure in a table cell to the entire outflow from a system of statistics sources, with survey repetitions over time as a vital ingredient. The quality concept to be formulated is meant to be wide enough to cover any type of official statistics product.

Nowadays many producers of official statistics have adopted the total quality approach, in which the notion of "quality of statistics" takes the following form.

(iii) ***Quality of statistics*** refers to all aspects of how well disseminated statistics meet users' needs and expectations of statistical information.

In accordance with (ii) the quality concept should list all aspects of statistics which are implicitly indicated by (iii). When making the concept concrete, it is natural to group the aspects by main ***quality components*** with sub - components. This structure is used in the quality concept formulated below, which comprises five main components; ***Contents***, ***Accuracy***, ***Timeliness***, ***Coherence*** especially ***comparability*** and finally ***Availability*** and ***clarity***. It should be emphasized, however, that even if there is quite wide agreement on which the sub - components should be, there is no world - wide consensus on how to group them under main components. The grouping below is a blend of many views, notably those of Statistics Sweden and Eurostat. The quality components are discussed in more detail in the subsequent section.

The quality concept is used in the following areas.

Quality declarations. To be able to use statistics adequately, users need, and require, information about their properties. For this purpose the producer should provide neutral, descriptive information, commonly called a quality declaration.

Survey planning. For a producer, as well as for a user with influence on the planning of a statistical survey (e.g. by financing it), the quality concept gives a check - list of quality aspects to take into consideration in the planning process.

Productivity evaluation and **quality improvement**. Statistics production processes need, as all production processes, evaluation and revision with regard to costs and benefits of the resource allocation. The quality concept provides a basis for such analyses.

The quality declaration context highlights the descriptive side of the quality concept. In the two other contexts it is important for the producer to know about users' quality assessments and preferences. The vehicle for this task is user - producer dialogue.

Quality concept for official statistics

CONTENTS
- Statistical target characteristics
  - Units and population
  - Variables
  - Statistical measures
  - Study domains
  - Reference time
- Comprehensiveness

ACCURACY
- Overall accuracy
- Sources of inaccuracy
  - Sampling
  - Coverage
  - Measurement
  - Non - response
  - Data processing
  - Model assumptions
- Presentation of accuracy measures

TIMELINESS
- Frequency
- Production time
- Punctuality

COHERENCE especially **COMPARABILITY**
- Comparability over time
- Comparability over space
- Coherence in general

AVAILABILITY and **CLARITY**
- Forms of dissemination
- Presentation
- Documentation
- Access to micro data
- Information services

General comments on the quality concept

a. Producers are well aware that users pay considerable regard to the ***cost*** of a product. Cost is not included as a quality component, however. This is in line with general quality philosophy which states that quality considerations, at least in a first round, should disregard cost. Often quality improvements can be achieved without increased cost. However, quality and cost have to be appropriately balanced in a final round.

b. It is a fact of life that deliberately false official statistics appear sometimes, that the statistics are not ***objective***. If so, this is of course a serious quality defect. Objectivity is not included as an aspect of the quality concept, however, for the following reasons. (i) We believe that deliberately false official statistics are exceptional. (ii) It is difficult to discuss, and assess, this quality aspect openly.

c. Some writers use the term ***relevance*** instead of contents, while others (including the present authors) think that the term lies too much on the assessing side. It should be a user's privilege to judge if some specific statistics are relevant for him/her.

d. Some writers advocate a broader quality concept, which takes into consideration not only the users but also the ***data suppliers***. If so, such aspects as ***response burden***, ***confidentiality*** and ***integrity*** also enter the quality picture. However, these aspects are here left outside the quality concept.

3 Elaboration on the quality components

As already stated, the producer's quality concept should in the first round be descriptive, while the users make ***quality assessments***. It has also been emphasized that it is important for a producer to have good knowledge of users' quality preferences.

Most producers also do have such knowledge, at least to the effect that they know if users will regard a particular production change as a step in a positive or negative "quality direction". As regards direction, essentially all users agree, but due to different use purposes they often disagree as regards the weight they assign to a specific quality change. Moreover, a production change may have a positive effect on some quality components and a negative effect on others. Hence, conflicting interests often prevail between users, and also within the same user.

The following elaborates on the contents of the quality components, their descriptive side as well as indications of conflicting interests.

CONTENTS

Concerns the statistical target characteristics

Users' requirements for ***statistical information***, i.e. information on values of statistical characteristics, emanate from their ***subject matter problems***. These may concern issues of economics, demography, environment, and many more. The preferable choices of units, population, variables, etc. in the target characteristics depend on the subject matter problem. Hence, ***relevance*** is not an intrinsic property of statistics, but relates to the subject matter problem. Within the same subject matter field a specific set of target characteristics can make the statistics highly relevant for some users, but less relevant for others. Conflicting interests often turn up, and compromises have to be made. Moreover, even if there is consensus about the most suitable target characteristics, considerations concerning cost, timeliness, measurement difficulties, etc. may lead to "second best" choices.

Specification of target characteristics

Units, population, variables and statistical measures

In the descriptive context, this point concerns precise definitions of population units, delimitation of the population, definitions of variables and statistical measures.

Study domains

The descriptive aspect of this quality component concerns answers to the following questions;

- Which types of classifications are used to form study domains?
- How far-reaching subdivisions into study domains are made?

Users commonly present very extensive requirements in both of the above respects, in particular as regards statistics for "small domains". There are restraining factors, however. One is that data on a requested classification variable may not be available (e.g. for cost reasons), another that statistics for many domains require overly extensive publication space.

For sample surveys the following additional restriction is at hand. When sample survey statistics is "broken down" to smaller and smaller domains, its accuracy deteriorates. Ultimately it becomes so low that the statistics no longer are meaningful. Hence, the breaking down process has to be terminated at an appropriate level.

Reference time

Units and variable values relate to specific times, which may be narrowly delimited and referred to as ***reference time points*** (e.g. a specific day), or be a time interval, a ***reference period*** (e.g. a calendar year). Usually reference times agree for all variables and units in a target characteristic, but they may differ. (Example: A survey target could concern salaries in 1985 and in 1995 for students who graduated from a particular education in 1975.)

On the descriptive side this quality aspect concerns precise specification of reference times for units and variables in the target characteristics.

Comprehensiveness

This aspect refers to a system of statistics for a specific subject matter field. (Example: The totality of economic statistics from the "national statistics system".) Many users want the statistics system to provide information on "all vital respects". The better this request is met, the more comprehensive is the statistics system. However, in practice no national statistics system can make all users satisfied relative to their opinion of "all vital respects", but it can be made better or worse.

ACCURACY

Concerns the agreement between statistics and target characteristics

For a sample survey it is evident that the resulting statistics do not provide exact values of the target characteristics. Moreover, as has already been emphasized, total enumeration surveys are usually subject to so many disturbances that the resulting statistics should be regarded as estimates rather than exact values. Normally there is a ***discrepancy*** between the values of a statistic and its target characteristic, also referred to as an ***error***. The (relatively) smaller the discrepancy is, the more ***accurate*** is the statistics. User requirements are, of course, that discrepancies should be small, preferably negligible.

Often, however, discrepancies are not negligible, in particular not for sample survey statistics. Then, at least statistically knowledgeable users want numerical bounds for the discrepancies, called ***accuracy measures*** or ***uncertainty measures***. Exhibition of accuracy measures is a somewhat intricate matter, since the discrepancies are defined in terms of target values that are unknown. (If they were known, it would be unnecessary to estimate them.) Statements concerning accuracy are therefore inevitably statements about states of uncertainty, a conceptually difficult topic. The usual structure for information about accuracy is as follows (in layman formulation).

It is most likely that a specified interval of the type

accuracy (or ***uncertainty) interval*** =

= value for the statistic ± ***margin of uncertainty*** (***error***)

comprises the true value of the target characteristic.

Sometimes such an interval can be interpreted as a ***confidence interval*** with a specified confidence level. In official statistics the confidence level is often chosen to be 95%. There are other accuracy measures, which in essence are equivalent to a confidence interval ; the estimator's ***standard deviation***, ***relative margin of error*** and ***coefficient of variation***.

Overall accuracy

The user's interest is focused on overall reliability of a statistic, in other words on the magnitude of the ***total error***.

In some cases the producer can provide precise overall accuracy intervals, but this is rather the exception than the rule. However, lacking precise bounds for total errors the producer should do his/her best to provide information on, at least judgments of, how certain source(s) of inaccuracy have affected the statistics. This is considered under the next quality component.

Sources of inaccuracy

As already stated, many different ***error sources*** affect the accuracy of statistics from a survey. The main ones are listed and discussed below.

Classifications of error source usually employ the duality ***sampling errors*** versus ***general survey errors*** (often called ***non - sampling errors***). The former relate to sample surveys, and emanate from the fact that only a sample of population units, not all, are observed. The latter relate to error sources which all types of surveys are subject to, total enumeration surveys as well as sample surveys.

Another common classification duality is ***systematic errors***, which lead to bias in the statistics, versus ***random errors***. The former relate to errors which (for the majority of observations) go in the same direction, while the latter relate to errors which spread randomly around 0. In this context the accuracy is commonly divided into the components ***bias*** (= size of the systematic error) and ***precision*** (= bound for the random error).

The total error (i.e. the discrepancy between a statistic and its target value) is often viewed as a sum of partial errors, emanating from different error sources,

Total error = sampling error + coverage error + measurement error + non-response error + ...

Even if it is difficult to give quantitative bounds for the total error, it is often possible to provide accuracy information for at least some of the partial errors.

In quality declarations the producer should, in addition to potential numerical error bounds, provide a verbal account of the data collection, including encountered obstacles.

We now turn to the main sources of inaccuracy.

SAMPLING

The fact that only a sample of population units are observed in a sample survey contributes to the inaccuracy of the resulting statistics.

A distinction in this context is that of ***probability samples*** (yielding control of sample inclusion probabilities) versus ***non - probability samples*** ("expert samples", "subjective samples" are existing synonyms).

Probability sampling is a safeguard against bias. Moreover, under probability sampling bounds for the sampling error can usually be given in terms of confidence intervals.

GENERAL SURVEY ERROR SOURCES

Coverage

Disagreement between survey frame and target population contributes to statistics inaccuracy.

Measurement

A measurement error occurs if a respondent's answer differs from the true variable value. Measurement errors may be systematic (e.g. underreporting of income) as well as random.

Systematic measurement errors lead to biased statistics. The contribution to inaccuracy from random measurement errors is mostly covered by the sampling error confidence interval.

Non - response

Non - response occurs when observation variable(s) value(s) for a designated observation unit have not been collected at the time when the estimation process starts.

Non - responses may lead to bias if there is correlation between "not respond" and the value of the survey variable. Various procedures exist for adjustment, in the best possible manner, for non - responses.

Non - response rates are commonly reported. They indicate the quality of the data collection process, but do not give information about the crucial quantity, the order of magnitude of the non - response error.

Data processing

On the way to statistics, collected data are processed in different stages such as data entry, coding, editing, and estimation/aggregation. At each step mistakes/mishaps may occur, contributing to inaccuracy.

Model assumptions

Some statistics rely on assumptions (e.g. stability of a consumption pattern), also referred to as ***models***. A model assumption that is not perfectly fulfilled contributes to inaccuracy.

Adjustment procedures (for non-response, coverage deficiencies, seasonal variations, etc.) also rely on assumptions/models. In such cases, the inaccuracy due to using models should be accounted for under the specific quality aspect. This component comprises model issues that are not covered by other quality components.

Presentation of accuracy measures

Statistics with accuracy deficiencies may lead to fallacious conclusions if used uncritically. Knowledgeable users can avoid fallacies if appropriate accuracy measures are presented. Statistics with accompanying accuracy measures are more informative than "bare" statistics.

This quality component refers to whether or not disseminated statistics are accompanied by uncertainty measures.

TIMELINESS

Concerns the appearance of the statistics over time

Many users want statistics from repeated surveys in order to monitor some specific development, prepared to take appropriate action if "alarming" levels are reached. In such situations, as in many others, a main requirement is that available statistics should be up to date. A vital aspect here is the time lag between "now" and the reference time for last available statistics. This lag depends on how frequently the survey is repeated and its processing time. A user's quality judgment in this respect does not, however, solely depend on the maximal time lag, his/her opinion of the pace of change for the development under consideration is also crucial. User opinions on timeliness quality depend jointly on the factors frequency, production time and punctuality.

Frequency

Statistics from repeated surveys are usually produced according to a regular scheme (monthly, quarterly, annually, etc.). In such situations it is natural to talk of ***frequency*** (or ***periodicity***). This notion has (at least) the following sub-aspects.

Data collection frequency concerns the periodicity in the producer's data collection.
Reference time frequency concerns the periodicity of reference times for published statistics.
Dissemination frequency concerns the periodicity with which statistics are made public.

Normally the three frequencies agree, but they may differ. (Example: Swedish crime statistics is published quarterly, comprising statistics for each month in the quarter.) Users normally care most about reference time and dissemination frequencies.

Production time

Production time is the lag between the reference time point (or end of the reference period) and the time for publication of a statistic.

Normally users' quality assessment direction for this component is that, ceteris paribus, the shorter the production time the better. However, if the statistics carries "unpleasant" messages, some users/actors may wish delayed publication. The common policy for handling the dilemma is that official statistics which do not have, or are late relative to, a promised publication date should be published immediately when ready.

The quality aspects accuracy and production time may come in conflict with each other. Shortening of a production time often leads to increased non-response as well as more hasty editing, which in turn affects accuracy adversely.

Punctuality

Punctuality refers to the agreement between promised and factual dissemination time.

The interest in punctuality varies considerably among users. An extreme example: For economic statistics that affect stock market prizes, punctuality may be a question of parts of a second.

COHERENCE especially COMPARABILITY

Concerns how well different statistics can be used together

Coherence relates to ***sets of statistics***, and takes into account how well the statistics can be used together. Two sub-aspects are of special importance. When the statistics set is a time series, one speaks of ***comparability over time***. When it comprises statistics for different domains with similar target characteristics, one speaks of ***comparability over space***.

In comparison contexts, one ideally wants to compare true values of the same characteristic. This ideal situation is not always at hand, and as a second best alternative one wants to compare statistics with as similar target statistics and as good accuracy as possible.

When judging similarity of target statistics, their ***definitions*** (regarding units, population and domain delineation, variables, etc.), of course, play a central role. The more stable a definition has been over time, the better comparability is over time. Analogously, for good comparisons over space, similarity in definitions of target characteristics is crucial. ***Statistical standard classifications*** (e.g. NACE for the classification of industries) are vital instruments to achieve agreement, or at least good similarity, between target statistics.

The acuteness of comparisons also depends on the ***accuracy*** of the statistics used for the comparison, their bias and precision. If the statistics compared are severely inaccurate, observed differences may reflect "noise play" rather than true differences. Of course biases disturb comparisons, but the harm caused by bias is often, but certainly not always, mitigated if the statistics compared have similar bias structures. An important mean for achieving good comparability is to require that statistics should be produced with the ***same methodology***, notably as regards questionnaire, data collection and estimation. The common methodology will hopefully be a good one, leading to negligible bias in all statistics to be compared. In any case it will lead to similar bias structures. The importance of a common methodology is enhanced by the fact that the precise content/definition of a variable often depends on the measurement and data collection procedures.

Comparability over time

Surveys that are repeated over time yield ***statistical time series***, which enable users to follow developments over time. Basic aspects of this quality component concern the extent to which the statistics in a time series in fact estimate the "same thing" in the "same way". Stability over time of target statistic definition and survey methodology work in the direction of good comparability over time, for reasons that are discussed above.

As regards the stability of definitions there are often conflicting user interests. A user with a particular interest in some special issue (or with "historical" interests), puts great emphasis on the stability of definitions and methods. Users whose main interest is the present and future state of affairs want reality changes (e.g. changes in industry structure) to be met by appropriate changes in the statistics. Modifications of target characteristics to meet reality changes usually have adverse effects on comparability over time.

A further aspect of comparability over time is that certain users, notably users of economic statistics indicating short term changes in economic activity, are anxious to be able to separate changes "in substance" from effects due to fairly regular seasonal variations. Technical means for this purpose are ***seasonal adjustments*** and ***calendar adjustments***. The mentioned type of users require adjusted series as complements to the basic time series.

Comparability over space

A common usage of statistics is for the comparison of conditions in different geographical regions (e.g. of average wage in different countries). The "space dimension" may also be of a non - geographical nature. (Example: Comparison of average disposable incomes for families with 1,2,3,... children.) Also when judging comparability over space, similarity of definitions of target characteristics and of survey methodology are crucial aspects.

When the statistics (for different domains) emanate from the same survey (by the same producer), problems regarding comparability over space are usually reduced to questions about the precision of the statistics. However, the farther apart the producers are (different surveys at the same agency, different agencies in the same country, offices in different countries, etc.) the greater are the comparability problems met. When statistics come from different surveys an additional aspect enters the picture, namely similarity in reference time.

Coherence in general

As stated, coherence relates to the general feasibility of making joint use of statistics from different sources, not only for comparison purposes. (Example: In order to judge the consequences of a potential change in taxation and benefits rules it might be of interest to combine statistics from the "Income survey", the "Expenditure survey" and the "Rent survey". Then it is of course important that the statistics are coherent, for instance that the same definition of "household" is used in the different surveys.)

The most important prerequisites for good coherence have already been mentioned. There should be agreement in definitions of basic target characteristic quantities (units, population, domains, variables and reference times). Similarity in survey methodology also works in the right direction.

AVAILABILITY and CLARITY

Concerns physical availability and intellectual clarity of statistics

Forms of dissemination

This aspect refers to what ***dissemination media*** (print on paper, diskette, CD/ROM, etc.) and what ***distribution channels*** are used.

Presentation

This aspect refers to how statistics are presented in printed publications, databases, etc. Specifically it concerns: presence, layout and clarity of texts, tables, charts and other figures; referencing; etc. It also covers how well particularly interesting features of the statistics are emphasized.

Documentation

This aspect refers to users' possibility to acquire documentation relating to published statistics.

Most users want an easily readable quality declaration. More advanced users are often also interested in precise documentation of the production process. This last requirement is particularly important when the user gets access to micro data for own use.

Access to micro data

Users may be interested in statistics that are not provided by the producer, but could be derived from already collected data. There are two main options in this context.

- The producer makes ***special derivations*** from available data, in accordance with requests formulated by the user.
- The ***user gets access to micro data*** for his/her own "statistics production".

Users with well specified problems mostly prefer the first alternative. Important points are then how fast the derivations can be carried out, and at what cost.

Researchers are commonly interested in getting hold of micro data for their own processing. Thereby they can make analyses more flexibly, faster and cheaper than via special derivations by the producer. Release of micro data is, however, associated with problems of secrecy and special precautions have to be taken by the producer. Removal of the means of identification is a minimum requirement.

Information services

A main aspect of information services is what assistance a user can get to find his/her way in the "statistics storage". Another important aspect is the possibility to get answers to questions about published statistics: their interpretation; specifics of definitions; details about data collection; etc.

4 Selected references on quality work at some statistical agencies

Official statistics has a long tradition. It has developed considerably during this century due to new demands (e.g. as regards subject matter areas), new methodology (e.g. survey sampling), new technology (e.g. for data collection and processing) etc. The numbers of uses and users have increased greatly. To give a comprehensive review of the development of the notion of quality of official statistics over time and space is too big a task to be covered here. We restrict ourselves to some recent "milestones" and a brief review of current views and activities.

Milestones

First, much survey development work has its origin in statistical agencies of the US Federal Government, notably the Bureau of the Census. The US role in this development is described by Bailar [1] and by Fienberg and Tanur [7].

Second, works on quality issues by Statistics Canada are often cited by other agencies. An important example is Quality Guidelines [11], which is a manual "providing advice for the production, maintenance and promotion of quality for statistical processes". Related works [12] and [13] focus on how to inform users.

Third, but not least, instrumental work has been carried out by international statistical organizations. The task of informing users was discussed in the 1980s. See for instance the reference United Nations [16], which was influenced by work by Statistics Canada, Statistics Sweden and US Federal statistical agencies. The latter work is presented in Gonzales et al. [8]. The UN guidelines emphasize two main types of quality presentations: (i) extensive presentations with technical orientation, with professional statisticians (producers as well as users) as target group, and (ii) presentations for statistics users in general, to assist them in interpretation of the statistics and in the decision of whether, and how, to use them.

Some current views and activities

Only a few papers discuss the quality concept in such structural detail as in Section 2. As already indicated, Statistics Sweden and Eurostat are two exceptions. However, quality concepts emerge implicitly from papers on quality endeavors. We try to emphasize this aspects in the review below.

Statistics Sweden [15] presents a quality definition and recommendations for quality declarations. This document updates 1979 guidelines for the presentation of quality of statistics and a 1983 policy for a user-oriented presentation.

Eurostat has an internal quality policy document, drafted in 1996. Moreover, there are documents on quality of business statistics tied to regulations on business statistics, the first being [5].

Harmonization and coordination of statistical systems are important activities in international statistics work aiming at good comparability and coherence of statistics. These quality components are emphasized in the UN guidelines and in the Eurostat quality concept for business statistics. The SNA system for national accounts is an important example of a world-wide harmonized system, which also influences other branches of economic statistics. Beekman and Struijs [2] discuss economic concepts and the quality of the statistical output.

Statistisches Bundesamt [14] provides a collection of discussions of statistics quality from a user's point of view, for political decision makers, scientists, in econometric uses, etc. The structure and emphasis of the quality concept vary between the contributions. Quality components that recur in several discussions are timeliness, accuracy, and comparability.

Dippo [4] considers survey measurement and process improvement. The paper links early work on non - sampling errors and different components of the overall error with recent work on process improvement. It includes the quality measurement model – which has the user in the center – of the US Bureau of Labor Statistics.

McLennan [10] describes the history of UK official statistics and developments in the 1990s, and lists some operational principles for the Central Statistical Office (CSO) under three headings: "definitions and methodology", "integrity and validity of CSO output", and "timing and coverage of publications".

Linacre [9], when describing methodology work in a statistical agency, refers to the objectives of the Australian Bureau of Statistics as: "Informed and satisfied clients through an objective, relevant, and responsive statistical system". A statistical product should comprise "reliable, timely, and coherent statistics".

Characteristics of an effective statistical system are discussed by Fellegi [6] who states that the "objective of national statistical systems is to provide relevant, comprehensive, accurate and objective (politically untainted) statistical information".

Colledge and March [3] report on a study, comprising 16 national statistical agencies around the world, on the existence of "quality practices" (classified as policies, standards, guidelines, and recommended practices) as well as the degree of compliance with prescribed practices.

References

[1] Bailar, B. A. (1990). Contributions to Statistical Methodology from the U.S. Federal Government. *Survey Methodology* **16**, 51-61.

[2] Beekman, M. M. and Struijs P. (1993). The Quality of Economic Concepts and Definitions. *Statistical Journal of the United Nations ECE* **10**, 1-15.

[3] Colledge, M. and March M. (1997). Quality Policies, Standards, Guidelines, and Recommended Practices at National Statistical Agencies. In *Survey Measurement and Process Quality* (editors: Lyberg, Biemer, Collins, de Leeuw, Dippo, Schwartz and Trewin). Wiley, New York.

[4] Dippo C. S. (1997). Survey Measurement and Process Improvement: Concepts and Integration. In *Survey Measurement and Process Quality* (editors: Lyberg, Biemer, Collins, de Leeuw, Dippo, Schwartz and Trewin). Wiley, New York.

[5] Eurostat (1996). *Quality in Business Statistics.* Eurostat/D3/Quality/96/02-final for structural business statistics.

[6] Fellegi, I. P. (1996). Characteristics of an Effective Statistical System. *International Statistical Review* **64**, 165 - 197.

[7] Fienberg S. E. and Tanur J.M. (1990). A Historical Perspective on the Institutional Bases for Survey Research in the United States. *Survey Methodology* **16**, 31 - 50.

[8] Gonzales, M. E., Ogus J. L., Shapiro G. and Tepping B. J. (1975). Standards for Discussion and Presentation of Errors in Survey and Census Data. *Journal of the American Statistical Association* **70**, 5 - 23.

[9] Linacre, S. (1995). Planning the Methodology Work Program in a Statistical Agency. *Journal of Official Statistics* **11**, 41 - 53.

[10] McLennan, B. (1995). You Can Count on Us – with Confidence. *J. R. Statist. Soc. A* **158**, 467-489.

[11] Statistics Canada (1987a). *Quality Guidelines*. Second Edition. Statistics Canada, Ottawa.

[12] Statistics Canada (1987b). Statistics Canada's Policy on Informing Users of Data Quality and Methodology. *Journal of Official Statistics* **3**, 83-91.

[13] Statistics Canada (1992). *Policy on Informing Users of Data Quality and Methodology. Policy manual.* Statistics Canada, Ottawa.

[14] Statistisches Bundesamt (1993). *Qualität statistischer Daten.* (Quality of Statistics.) Beiträge zum wissenschaftlichen Kolloquium am 12./13. November 1992 in Wiesbaden. Band **25** der Schriftenreihe Forum der Bundesstatistik herausgegeben vom Statistischen Bundesamt. (In German)

[15] Statistics Sweden (1994). *Kvalitetsbegrepp och riktlinjer för kvalitetsdeklaration av officiell statistik.* (Quality Definition and Recommendations for Quality Declarations of Official Statistics). MIS 1994:3. Statistics Sweden, Stockholm. (In Swedish)

[16] United Nations (1983). *Guidelines for Quality Presentations that are Prepared for Users of Statistics.* Statistical Commission and Economic Commission for Europe, Conference of European Statisticians, Meeting on Statistical Methodology, 21-24 November 1983.

Förteckning över utkomna R&D Reports

R&D Reports är en för U/ADB och U/STM gemensam publikationsserie, som 1988-01-01 ersatte de tidigare "gula" och "gröna" serierna. I serien ingick fram till årsskiftet 1992-93 även **Abstracts** (sammanfattning av metodrapporter från SCB).

Reports published during 1996–

1996:1 (grön)	On Sampling with Probability Proportional to Size *(Bengt Rosén)*
1996:2 (grön)	Bortfallsbarometern nr 11 *(Antti Ahtiainen, Stefan Berg, Margareta Eriksson, Åsa Greijer, Dan Hedlin, Monica Rennermalm, Anita Ullberg)*
1996:3 (grön)	Regression Estimators in Theory and in Practice *(Tomas Garås)*
1996:4 (gul)	Quality Aspects of a Modern Database Service *(Pat Dean and Bo Sundgren)*
1996:5 (gul)	Metadata: A Quality Element in Official Statistics - the Swedish Approach *(Bo Sundgren and Pat Dean)*
1996:6 (gul)	Our Legacy to Future Generation - Using Databases for Better Availability and Documentation *(Gösta Guteland and Erik Malmborg)*
1997:1 (grön)	Bortfallsbarometern nr 12 *(Anti Ahtiainen, Stefan Berg, Mats Bergdahl, Fredrik Granström, Dan Hedlin,Lena Otterskog, Monica Rennermalm)*
1997:2 (grön)	Quality Concept for Official Statistics - Entry in the forthcoming update of the Encyclopedia of Statistical Sciences, Wiley & Sons *(Eva Elvers, Bengt Rosén)*

090693-002101- 1

CATARINA ELFFORS

BR/BY-S

ISSN 0283-8680

Tidigare utgivna *R & D Reports* kan beställas genom Ingvar Andersson, SCB, U/SIB, Box 24300, 115 81 STOCKHOLM (telefon 08-783 41 47, fax 08-783 45 99, e-post ingvar.andersson@scb.se).

R & D Reports listed above as well as issues from 1988-1994 can - in case they are still in stock - be ordered from Statistics Sweden, att. Ingvar Andersson U/SIB, Box 24300, S-115 81 STOCKHOLM (telephone +46 8 783 41 47, fax +46 8 783 45 99, e-mail ingvar.andersson@scb.se).

SCB-Tryck, Örebro 1997.10