eurostat

Doc. Eurostat/A4/Quality/03/General/Definition
Available in EN

Working Group

"Assessment of quality in statistics"

Sixth meeting

Luxembourg, 2-3 October 2003 at 9 h 30

Room Ampere, Bech building

ITEM 4.2:
METHODOLOGICAL DOCUMENTS - DEFINITION OF QUALITY IN STATISTICS

Definition of quality in statistics

Introduction

This paper defines the Eurostat quality concept[1]. The major topics that should be addressed in a report on quality are described according to the six quality dimensions. The items to be included in the report are described in the Eurostat document 'Standard quality report' of October 2003.

A guide "How to make a Quality Report" is also available for providing the detailed guidelines for the completion of the Quality Report.

DEFINITION OF QUALITY

Quality is defined in the ISO 8402 - 1986 as: "the totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs".

EUROSTAT'S DEFINITION OF QUALITY OF STATISTICS

Quality of statistics is defined by Eurostat with reference to the following six criteria:

- relevance;
- accuracy;
- timeliness and punctuality;
- accessibility and clarity;
- comparability and
- coherence

.

Although not a measure of quality, the costs involved in the production of statistics as well as the burden on respondents act as a constraint on quality. When assessing the ability of a Member State to comply with quality guidelines, it is necessary to take into account the cost and burden of statistics.

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.) reflects user needs.

Accuracy

Accuracy in the general statistical sense denotes the closeness of computations or estimates to the exact or true values

Timeliness and punctuality

[1] It replaces "Eurostat/A4/Quality/02/General/Definition".

Punctuality refers to the time lag between the release 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.

Timeliness of information reflects the length of time between its availability and the event or phenomenon it describes

Accessibility and clarity

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.

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 also available (including limitation in use…) and the extend to which additional assistance is provided by the NSI.

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. We can say 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*.

- Comparability over time refers to comparison of results, derived normally from the same statistical operation, at different times.

- The geographical component of comparability emphasises the comparison of statistics between countries and/or regions in order to ascertain, for instance, the meaning of aggregated statistics at European level.

  Geographic comparability is not of course limited to the comparability within EU. The EU statistics can be compared with other international statistics, for example with Japan and USA. In the EU context, it can exist a *European reference*, to which each Member State statistic could be benchmarked

- Comparability between domains refers to non-geographical domains, for instance between industrial sectors, between different types of households, etc.

Coherence

Coherence of statistics is therefore their adequacy to be reliably combined in different ways and for various uses. It is, however, generally easier to show cases of incoherence than to prove coherence.

When originating from a single source, statistics are normally coherent in the sense that elementary results derived from the concerned survey can be reliably combined in numerous ways to produce more complex results.

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. Conveying neighbouring results, they may also convey not completely coherent messages, the possible effects of which, users should be clearly informed of.

REMARKS

1. There is a trade-off between the different components of quality, especially: timeliness/accuracy, accuracy/geographic comparability, relevance/comparability over time, relevance/accuracy, coherence for large domains/relevance for sub-domains, etc.

2. The above breakdown of quality into components is not unique neither invariant over time. Other organisations use slightly different sets of quality dimensions. For instance,

- Statistics Canada uses six dimensions: relevance, accuracy, timeliness, accessibility, interpretability, and coherence (from Statistics Canada, 2002, "Statistics Canada's Quality Assurance Framework", Catalogue nr. 12-586-XIE, http://www.statcan.ca/english/freepub/12-586-XIE/12-586-XIE02001.pdf); and

- Statistics Sweden uses five: content, accuracy, timeliness, comparability/ coherence, and availability/clarity (from Rosén, B., and Elvers, E., 1999, "Quality Concept for Official Statistics" pp. 621-629 in S. Kotz, C.B. Read, and D.L. Banks (eds.), Encyclopedia of Statistical Science, Update Vol. 3, Wiley, New York).

Other frameworks for assessing the quality are also in place.

- The International Monetary Fund (IMF) has developed its own framework (Carson, Carol S., February 2001, "Toward a Framework for Assessing Data Quality", IMF Working Paper, WP/01/25). IMF uses six components: Integrity, Methodological Soundness, Accuracy and Reliability, Serviceability, Accessibility and Prerequisites of Quality

- The OECD has developed a quality framework with eight components: Relevance, Accuracy, Credibility, Timeliness, Punctuality, Accessibility, Interpretability and Coherence (OECD (2002) Quality framework for OECD statistics. OECD. Paris. Available at www.oecd.org/doc/m00029000/m00029990.doc.).