Checklist
Quality of Statistical Output

09

Ir Peter W. M. van Nederpelt EMEA

Publication date Statistics Netherlands' website: 12 February 2009

Statistics Netherlands

The Hague/Heerlen, 2009

Explanation of symbols

.	= data not available
*	= provisional figure
x	= publication prohibited (confidential figure)
–	= nil or less than half of unit concerned
–	= (between two figures) inclusive
0 (0,0)	= less than half of unit concerned
blank	= not applicable
2005-2006	= 2005 to 2006 inclusive
2005/2006	= average of 2005 up to and including 2006
2005/'06	= crop year, financial year, school year etc. beginning in 2005 and ending in 2006
2003/'04–2005/'06	= crop year, financial year, etc. 2003/'04 to 2005/'06 inclusive

Due to rounding, some totals may not correspond with the sum of the separate figures.

Publisher
Statistics Netherlands
Henri Faasdreef 312
2492 JP The Hague

Prepress
Statistics Netherlands - Facility Services

Cover
TelDesign, Rotterdam

Information
Telephone .. +31 88 570 70 70
Telefax .. +31 70 337 59 94
Via contact form: www.cbs.nl/information

Where to order
E-mail: verkoop@cbs.nl
Telefax .. +31 45 570 62 68

Internet
www.cbs.nl

Checklist Quality of Statistical Output

Ir Peter W.M. van Nederpelt EMEA

Abstract
This report describes nineteen characteristics of statistical output. Each characteristic – also called dimension - is elaborated according to a certain structure starting with the definition of the characteristic. For each characteristic possible indicators and measures are formulated and summerized as a checklist in an annex. This report has several purposes. Seven purposes of the report are identified like serve as a knowledge base while making an agreement with customers about quality of statistical output. The report does not contain guidelines for the CBS organization and has no mandatory character. Although it can serve as a starting point for developing guidelines.

Keywords: Quality statstistical output, dimensions statistical output, checklist, indicators, measures.

Content

1 Introduction

1.1 Aim of this document

The aim of this document is to describe what *indicators* can be applied to measure the quality of statistical output, and what *measures* can be taken to control the quality of the statistical output. Together these indicators and measures make up the checklist.

This document can be used for the following purposes:
1. To define the quality of statistical output when (re)designing a statistical process.
2. To come to an agreement about the quality of statistical output (in SLAs and covenants).
3. To design and compose quality reports.
4. To control the quality of the statistical output of existing production processes.
5. To compile a self-assessment of the quality of statistical output: quickscans, deep scans.
6. To set a standard for statistical audits.
7. In some cases: to compile a set of minimum requirements with which CBS statistics should comply.

The report is a body of knowledge and not a mandatory framework. Moreover, the checklist is not meant to be applied as a whole. It aims to map 'all' possible indicators and measures, so that various target groups can make their own choices.

The target groups for the checklist include all managers and employees involved in realising the above-mentioned goals.

1.2 Range

This document describes various dimensions of *statistical output*.
1. relevance
2. accuracy
3. comparability
4. coherence
5. timeliness
6. accuracy
7. accessibility
8. clarity
9. confidentiality

Eurostat distinguished these dimensions in the Code of Practice (COP002, 2005).

The following dimensions are also distinguished:
10. extent of detail
11. completeness
12. numerical consistency
13. plausibility
14. disputability
15. validity
16. reliability
17. verifiability
18. reproducibility
19. availability

These dimensions are used in daily practice at the CBS.

Research reports are available in Dutch about the dimensions accuracy, coherence and comparability. These reports describe these three dimensions in more detail.

For each dimension of statistical output the following elements will be discussed:
1. Definition

2. Relationship with other dimensions of statistical output.
3. Requirements
4. Consequences of problems with the dimension (risks)
5. Importance for the organisation (dependence)
6. Causes of problems with the dimension (threats, vulnerabilities)
7. Indicators
8. Measures
9. Checklist of possible indicators and measures

The *checklist with possible indicators and measures* (item 9) is in the annex to this document. The checklist distinguishes two different kinds of indicators and measures: those that can be used for *individual statistics* (annex 1) and those that can be used for statistical programmes (annex 2).

De *relationships* between dimensions of statistical output (item 2) are described in annex 3 to this document.

The quality of the following object types are *not* discussed in this document: metadata, registers, methodologies, classifications, datasets.

1.3 Definitions, acronyms and abbreviations

Item	Description
CBS	Centraal Bureau voor de Statistiek = Statistics Netherlands
Conceptual metadata	Information about data of a statistical series: population, unit, classifications, reference period, data items.
COP	Code of Practice. Reference (COP002, 2005).
DMK	Division of Methodology and quality
DPK	Department of Quality and process development
Statistics	The table disseminated. This is a restricted meaning of the term statistics.
ESS	European Statistical System
Quality metadata	Information about the quality of the process and data of a statistical series (values of the indicators). NB: This can also refer to process metadata of the measurement process as well as the conceptual metadata of the indicators.
ONS	Office for National Statistics.
Process metadata	Information about the statistical process. This can refer to the statistical methods applied as well as the organisational implementation of the process.
RMSE	Root Mean Square Error

1.4 References

Baker, T.L. (**1988**), Doing Social Research, New York. McGraw-Hill International Editions.

Biemer, Paul P., Groves, Robert M., Lyberg, Lars E., Mathiowetz, Nancy A., Sudman, Seymour (**1991**). Measurement Errors in Surveys, John Wiley & Sons, Inc.

Booleman, Max, Algera, Symon, Eijkhout, Michiel, Garssen, Joop, Mooijman, Paul (**2005**). Transparante statistische output. Version 21 December 2005 (concept).

Daas, Piet J.H. Judit Arends-Tóth, Barry Schouten, Léander Kuijvenhoven (**2008**). Quality framework for the evaluation of administrative data. CBS.

Gelsema, Tjalling (**2007**). ICT Masterplan: CBS Architectuur. Conceptuele Business Architectuur Metadata, version 0.4 dated 12 December 2007. Approved by the Architecture Review Board.

Lessler, Judith T, Kalsbeek, William D.(**1992**), Non sampling errors in surveys, John Wiley & Sons, Inc. ISBN 0-471-86908-2.

Van Nederpelt EMEA, Ir P.W.M. (**2008**). Objectgeörienteerde Kwaliteitszorg (OQM), Een managementmodel voor kwaliteit. Version 1.0 dated 18 August 2008.

Anonymous reports

CBS	CBS Quality System **2008**.
CBS Wet	Act on Statistics Netherlands. 20 November **2003**.
COP002	European Commission. Recommendations of the Commission on he independence, integrity and accountability of the national and Community statistical authorities. Version 25 May **2005**.
COP012b	European Commission. European Statistics Code of Practice. Self Assessment Questionanaire. 21 October **2005**.
DB	Minutes of the Board of Directors 10 April 2007.
DMK	Annual programme of the DMK Division 2007
EU001	Eurostat. Assessment of Quality in Statistics. Standard Quality Report. April **2000**.
EU002	Eurostat, Working Group Assessment of quality in statistics, ITEM 4.2: METHODOLOGICAL DOCUMENTS - DEFINITION OF QUALITY IN STATISTICS. October **2003**
EU003	Eurostat, Working Group Assessment of Quality in Statistics. Methodological Document Standard Report. October **2003**.
EU028	Eurostat. Standard Quality Indicators. May **2005**.
EU034b	Eurostat, ESS Handbook for Quality Reports (EHQR), October **2008**
EU035b	BoP Quality Report **2008.**

EU036 Euro SDMX Metadata Structure, Draft Ocober **2008**.

Gedragscode Code of Conduct Statistics Netherlands. Version June **2003**.

Handboek Statististische Beveiliging
Handboek Statistical Beveiliging. August **2006**.

I008 Startcursus Methodologie (Boom der Statistiek). Module kwaliteit. Enrico Witvliet.

I020a Research report Accuracy of data. Peter van Nederpelt. Version 1 dated October **2008**.

I020b Research report Coherence of statistics. Reinder Banning. Version 1.0 dated October **2008**.

I020c Research report Comparability of data. Reinder Banning. Version 1.0 dated October **2008**.

I020d Annex to the research reports Coherence en Comparability. Reinder Banning. Version 1.0 dated October **2008**.

I021 Elaboration Quality Framework HPE. Projectgroup HPE Quality & Fractions. Version 10 June **2008**. BPA DMV-2008-06-10-BSTN.

Quality declaration
Quality declaration CBS (complete version). Approved by the Board of Directors at 16 June 2008.

Metadata Working Group Metadata. A generic standard for structuring conceptuele metadata. Version 1 dated March **2004**.

Methodenreeks Methods series DMK.

ONS001 Guidelines for measuring statistical quality. ONS. **2006**.

ONS014 Accuracy. 79 Qualiity indicators. Onderdeel van ONS001. **2006**

S-Reeks Supplementaire Methodenreeks DMK

Business- and Information Architecture CBS

C001 Context van verandering. Version 2.0 dated 3 April **2006**.

BI001 Business- en Informatiemodel. R. Huigen. Version 1.0 dated 3 April **2006**.

BI002 Logische Informatie architectuur. Dick Windmeijer. Version 1.0 dated 1 May **2006**.

2 Background and procedure followed

2.1 Statistics Netherlands

DMK's annual programme for 2007 contained the intention to "coordinate the internal reviews of statistical output" (DMK, 2007, page 19). This means that the CBS was to develop a method to review its output in a structural manner similar to the methods used at external institutes like Eurostat and the IMF.

Furthermore, the Board of Directors stated that criteria for considering cost and quality must be part of a standard to be developed. This standard will be applied to innovation of statistical processes.

Criteria costs and quality were discussed in the meeting of the Board of Directors of 10 April 2007. This resulted in to action item 176 "Definite answer about feasibility and time limit review for most import statistics".

The minutes of the meeting of the Board of Directors states:

"As a result of the presentation "cost and quality of statistics" it is concluded that determining a quality standard based on theoretical considerations is not desirable. An assessment based on a model is practically hardly feasible in the short term. Other institutes will be consulted on how they deal with cost/quality considerations".

"It is possible, however, to review the accuracy of current statistical output and the processes underlying statistical output. In the next meeting of the Board of Directors a time limit will be indicated for tan overview of the most important statistics".

The manager of the DPK department proposed to the director of DMK division (email April 22, 2007) that the scope of the project should be the quality of statistics in general. The director agreed with this proposal.

DMK's annual programme for 2008 also includes this project, as it is a biennial project.

2.2 Eurostat

CBS requested a grant in the framework of "Theme 10: Quality Management and Evaluation". The request was granted by Eurostat and on 6 September 6 2007 the Director-General signed the contract with Eurostat.

The contract specifies that the "Accuracy of statistics for the Statistics Netherlands' corporate image" be reviewed for the years 2007-2008. The KIS project ensures that the obligations are met.

2.3 Procedure followed

Before the document at hand was compiled, literature was collected about the quality of statistical output; in the first place from Eurostat, but also from other sources: IMF, OECD and the UN.

National institutes in other countries have also published on the quality of statistical output, and a number of articles in journals were also found on the subject. Lastly, internal documents were also consulted.

All digital literature is stored in a data quality Sharepoint site.

Based on the literature found, three research reports were drafted. These research reports have the same structure as the document at hand. They describe (in Dutch) the following quality dimensions:

1. Data accuracy

2. Comparability of statistics
3. Consistency of statistics

The Checklist has been thouroughly reviewed by a reviewteam: Max Booleman, Piet Daas, Nieno Platteel and Barry Schouten. The Dutch version has been translated by Peter van Nederpelt and Jeanne Roijen. The translation has been reviewed by Lieneke Hoeksma from CBS' Translation Department.

3 Methodology used

This chapter describes the methodology used to compile this report.

A quality management model has been developed and applied at the CBS: Object Oriented Quality Management (OQM). This OQM model (Van Nederpelt, 2008) was used for this report.

Areas of interest
The OQM model starts from the idea that areas of interest for quality assurance are nominated. These areas of interest always consist of a combination of an object type and an object type characteristic. In less formal language: an area of interest is a combination of an object and a dimension of this object.

In our case a statistic of statistical output is the object. This may seem remarkable, because statistics always describe objects. But in this case the statistic itself is the object about which we make statements.

The following dimensions of the statistic are distinguished:
1. relevance
2. accuracy
3. comparability
4. coherence
5. timeliness
6. accuracy
7. accessibility
8. clarity
9. confidentiality
10. extent of detail
11. completeness
12. numerical consistency
13. plausibility
14. disputability
15. validity
16. reliability
17. verifiability
18. reproducibility
19. availability

Steps per area or interest
The OQM model further prescribes that for each area of interest a number of (standard) steps can be taken. The ultimate results of these steps must be that the right measures are chosen to 1) fulfil the requirements, and 2) to keep the risks for the area of interest between acceptable limits.

According to the OQM model, the following steps are taken for each area of interest:

1. **Determine the definition** of the area of interest. The purpose of this step is to find out what it is all about. This is especially important for concepts that are multi interpretable or that are not well-known.

2. **Determine the relationship** with other areas of interest. The purpose of this step is to make explicit the dependencies between areas of interest. It could turn out that areas of interest cannot be examined in an isolated way.

3. **Determine the requirements**. The purpose of this step is to make explicit which requirements must be met (ex ante). These requirements may originate from the outside world or be formulated by the organisation itself.

4. **Determine the consequences of problems with an area of interest (risks)**. The purpose of this step is to perform the part of the risk analysis where the impact of the problem with the area of interest will me mapped.

5. **Determine the importance of the area of interest for the organisation and its environment (dependence).** The purpose of this step is to determine how much attention should be paid to this area of interest compared with other areas of interest. How dependent is the organisation on this area of interest?

6. **Determine the causes of problems with the area of interest (threats, vulnerabilities)**. The purpose of this step is to be able to determine which measures should be taken to control the risk of the area of interest. This step is part of the risk analysis where vulnerabilities of and threats to the area of interest will be mapped.

7. **Determine indicators**. The purpose of this step is to determine what indicators are advisable. These indicators make the state of affairs visible regarding the area of interest by measurement. What norm is laid down for each indicator? Based on the real value of the indicator, it can be determined whether the area of interest meets the standard.

8. **Determine measures**. The purpose of this step is to determine what measures are necessary to take care that risks to the area of interest are limited and that the requirement are met. This is the chief purpose of this method and therefore also of this checklist.

9. **Divide responsibilities**. The purpose of this step is to determine who has which responsibilities regarding the area of interest. There must at least be one owner of the area of interest.

In this report, all the steps are taken except the last one. It is assumed that the process owner is responsible for the quality of the statistical output.

This report comprises a checklist for each area of interest (in this case each dimension). The checklists contain lists of possible indicators and measures. The checklists are incorporated in annexes 1 and 2 to this report.

4 Relationship with existing frameworks

The Checklist report is related to a number of existing frameworks:

Internal
- The mission of the CBS
- The Act on Statistics Netherlands
- The Code of Conduct
- The CBS Quality declaration
- The business and information architecture
- Conceptual business architecture for metadata
- Standard methods
- Quality of registers
- Framework of metadata concepts

External
- The Code of Practice
- ESS Handbook for Quality Report

The relationship with these frameworks is discussed below.

4.1 Mission of the CBS

The mission of CBS gives a concise and overall framework for the quality of statistical output. The mission states: compiling and publishing undisputed, consistent, up-to-date statistical information for practical policy and scientific purposes.

4.2 Act on Statistics Netherlands

In the Act on Statistics Netherlands (2003), section 37 covers data confidentiality. We shall quote this act in the chapter on confidentiality.

4.3 Code of Conduct

The Code of Conduct (2003) contains some principles regarding statistical output: reliability and indisputability of statistics, the relevance of data, and statistical confidentiality.

The Code of Conduct demands for example that an understandable indication is given of the uncertainty of the figures.

4.4 Quality declaration

The Quality declaration of the CBS (**2008**) mentions compliance with the Code of Practice and the Quality Declaration. It also explains how the CBS puts these standards into practice. The measures regarding quality of output and mentioned in the Quality declaration are also found in the report at hand.

4.5 Business and information architecture

The report at hand must fit into the business and information architecture (2006). This architecture contains the following guidelines for the quality of statistical output.

Guidelines from the business and information architecture		
Code	**Direction**	**Comment**
CX3	The solution must assure the **continuity** of the production of statistics.	Continuity = comparability in time.
CX6	While realising the solution, the **quality and the timely delivery** of mandatory statistics may not be compromised.	Timely = punctual
CX7	The solution must be aimed at realising **consistent** information, and the principle of one figure for one phenomenon.	one figure for one phenomenon = numerical consistency

Guidelines from the business and information architecture		
Code	**Direction**	**Comment**
CX16	The solution must realise an effective and a suitable quality control and quality assurance based on registration of audit and quality information.	
CBI02	Data that are processed and the metadata which describe definition, quality and process activities must be distinguished in a strict way.	

This report will adhere to the guidelines of the business and information architecture, but will constitute an addition to this architecture regarding the quality of statistical output. The current architecture is not complete with regard to the quality of statistical output, not even as an outline.

4.6 Conceptual business architecture for metadata

In the business architecture for metadata (Gelsema, 2007) domains of information are distinguished regarding metadata. Relevant domains of information are data source descriptions, data item descriptions, classifications and descriptions of populations. The architecture formulates requirements for these domains of information.

In general, conceptual metadata must posses a certain quality. This is a condition to determine the quality of statistical output.

Therefore population, data items, and classes must, for example, be unambiguously defined to determine whether figures are accurate. It is for example not possible to determine the accuracy of business turnover, if it is not clear what the meaning of this data item is.

Also consistency of statistical output only becomes visible if the population, the classifications and data items are unambiguously defined.

The requirements in the table below are relevant for the quality of statistical output. The requirements regarding the conceptual metadata could be formulated more precisely.

Requirements mentioned in the architecture for metadata		
Code	**Domain of information**	**Requirement**
RQ60	VAR Data item descriptions	Data item descriptions of an understandable definition and explanation must be provided.
RQ62	CLS Classifications	Classes of an understandable explanation must be provided.
RQ63	POP Population descriptions	Population descriptions must be understandable and be available for the general public in a readable format.

4.7 Standard Methods

The purpose of the standard method is to create a framework for the methodologies used at the CBS.

The implementation of a certain methodology always influences – mostly positively – the quality of output. A connection is rarely *explicitly* made between a methodology and the quality of output.

This report tries to bridge the quality of statistical output and methodology. For each dimension of statistical output the Standard Methods will be mentioned that can be applied.

4.8 Framework for the quality of registers

Recently a framework was laid down for the quality of registers (Daas et al., 2008). This framework specifies which quality dimensions could be recognised in a register.

The quality of statistical output is very much dependent on input sources like registers. The framework of the quality of registers is therefore strongly related to this report. For each dimension of statistical output it is indicated what the dependency is on the quality of registers.

It must be understood that the concept “input source” should be broadly interpreted. Auxiliary files that are used to determine weighting factors, microdata from a previous period, administrative data, etc. can also be considered as input sources.

4.9 Glossary for metadata

The glossary for metadata was consulted for statistical concepts (Metadata, 2004).

The purpose of this glossary is to get an overview of the concepts which can be used to describe the conceptual metadata.

However this report deviates from this glossary. Where the glossary mentions **characteristics of an object type** this report applies the expression **dimension or quality dimensions of an object.** This choice has been made to use colloquial language.

4.10 Code of Practice

The Checklist is closely connected to the Code of Practice (COP002, 2005). Principles 11 to 15 of the Code of Practice refer to the quality of statistical output. Principle 5 concerns data confidentiality; this implies confidentiality of statistical output.

Dimensions of statistical output as mentioned in CoP	
Principle Number	**Dimension**
5	Confidentiality of data
11	Relevance
12	Accuracy and reliability
13	Timeliness and punctuality
14	Coherence and comparability
15	Accessibility and clarity

The Code of Practice lays down high level requirement for the quality of statistical output. These requirements are adopted for each dimension.

Furthermore, for each principle indicators are mentioned. These indicators can, however, be characterised as measures. Generally these indicators are not measurable quality data items.

The present report must be seen as an operationalisation of the Code of Practice for CBS.
The added value of this report is that extra dimensions are added to the CoP. These extra dimensions are not described in an integrated manner anywhere. They have been added to the report as we see them in daily practice.

Most documents about the quality of statistical output have a certain focus: definitions, indicators, errors, methodologies, etc. No literature has been found on the consequences of problems with dimensions.

4.11 Handbook for Quality Reports

Eurostat's Handbook for Quality Reports (EU034b, 2008) describes one or more indicators for each dimension of statistical output. These indicators are incorporated in the report.

The Handbook for Quality Reports also describes which process metadata and conceptual metadata should be reported. These elements are not incorporated in this report.

5 Relevance of statistics

This chapter examines the relevance of statistics and gives recommendations for the assurance of the relevance of statistics. We start by defining the relevance of statistics.

5.1 Definition

We define relevance of statistics as follows:

> *The relevance of statistics is the degree to which statistics comply with the needs of current users.*

This definition deviates slightly from the definition of Eurostat (EU034b, 2008). We omitted the needs of potential users, as these are difficult to map.

The dimension relevance can be applied on two levels:
- The statistical programme or a set of statistics
- One statistic

The two levels are explained in the following table:

Level	Further explanation of relevance
Statistical programme	All statistics that are needed are delivered. No statistics are omitted.
One statistic	The concepts used (population, units, data items, classifications and extent of detail) comply with the needs of the users. A statistic is irrelevant if it contains data the user does not need. A statistic must also have sufficient quality. One must think of all dimensions mentioned in this report.

Synonym

A comparable term to relevance is usability. However, a distinction can be made between these two terms:
- A statistic is relevant if there is a need for the concept.
- A statistic is usable if it is also sufficiently accurate, consistent, comparable, up-to-date, punctual and complete.

In this report we use relevance and usability as synonyms.

5.2 Requirements

Code of Conduct

The CBS Code of Conduct mentions that the relevance of CBS data is determined by assessing the need for information of users in all possible manners.

Code of Practice

The Code of Practice (COP002, 2005) recommends in principle 11 that the European statistics must meet user needs. Principle 11 of the CoP is not very specific. Section 5.7 specifies which measures are possible to assure the relevance of statistics.

No further requirements, recommendations or decisions are known regarding the relevance of statistics.

5.3 Consequences of problems with the relevance of statistics

The consequences or risks of irrelevant statistics may vary strongly.

If a statistic does not meet the needs of a user, this user will generally not be satisfied. This may lead to loss of image.

In the most extreme case, if the CBS were to produce only irrelevant statistics, this would endanger its continued existence. Mostly irrelevant statistics will not be accepted by government and society.

5.4 Importance of relevance of statistics for the organisation

The negative consequences of irrelevant statistics proves the importance of the relevance of statistics. The existence of the CBS is strongly dependent on the relevance of its statistics .

As stated in its mission, the CBS regards the relevance of statistics as very important. This mission states that statistical information must be relevant for practical policy and scientific purposes. This means that the statistics of the CBS for Dutch and European society must be significant.

5.5 Causes of problems with the relevance of statistics

We were unable to find any analyses about possible causes of problems with relevance of statistics in the literature about the quality of statistics. Therefore we shall make an attempt to denominate possible causes of such problems.

In our vision problems with the relevance of existing statistics have various causes:

- It is not clear to the CBS why the user needs the statistic.
- There are no agreements with the users of the statistics.
- The agreements are incomplete. They do not state what should be delivered, when and how.
- The agreements are not current because they have not been evaluated or updated for quite some time.

Causes of problems with the relevance of the whole portfolio could be:

- Insufficient contact with potential user of statistics.
- Innovation processes are not in place

5.6 Indicators

The question is: how can the relevance of statistics be measured? We can distinguish statistics that the CBS is obliged to produce, and statistics where CBS has a say in the definition.

If a statistic has to be produced (mostly based on a regulation) then it should be enough to check whether the agreement has been met. Have all data items been delivered? Have the correct classifications been used? Have the results been delivered on time? Has the correct level of detail been used?

If the CBS has a say in determining the content of a statistic, then it is also important whether the user is satisfied with the statistical information. Is the statistic usable in the user’s process?

CBS statistics are always published on the its website. The question is: how can the CBS make agreements with anonymous users? And how can it measure the satisfaction of anonymous users of published data? This can be measured indirectly by counting the hits on StatLine.

One of the tasks of the Central Commission of Statistics (CCS) is to represent Dutch society. It is therefore obvious to come to an agreement with the CCS about statistics on StatLine. This also applies for measuring the satisfaction of the users of StatLine. This can be measured through the members of CCS.

Code of Practice
In the Code of Practice (COP002, 2005) the following indicators are recommended in principle 11:

1. Processes are in place to consult users, monitor the relevance and practical utility of existing indicators in meeting their needs, and advise on their emerging needs and priorities.
2. Priority needs are met and reflected in the work programme.
3. User satisfaction surveys are undertaken periodically.

These indicators are incorporated in the checklist.

5.7 Measures

At the CBS measures are (of course) already taken to control or improve the relevance of statistics.

First of all the CBS is legally required to set a long-term programme. In this programme, less relevant statistics are replaced by more relevant ones. Recent long-term programmes have resulted in a substantial innovation and increased relevance of the CBS portfolio of statistics.

Furthermore the CBS concludes collaboration agreements with external parties, mostly in the form of covenants. The purpose of collaboration may be to develop and supply new statistics. The relevance of the CBS portfolio is thus increased.
In some cases parties collaborate to collect data for existing statistics, in order to increase response and/or reduce the administrative burden.

In recent years users of CBS statistics have been interviewed about what they think needs to be improved or innovated. These interviews were published in a magazine which was sent to all CBS employees, intended to make points of improvement visible.

The CBS has mapped all its important users and designated them “strategic relations”. A contact person has been assigned to each strategic relation, and all communication with that strategic relation goes through this contact person. A streamlined communication promotes the relevance of statistics.

As we will never know who all the users of statistics are, the answer to the relevance question will never be complete. Many hidden users exist.

6 Accuracy of data

This chapter discusses the accuracy of the data and presents recommendations to assure the accuracy of data.

Figures or data are - in more formal language - the same as the values of data items. We shall use these two terms to mean the same in this report

We start by defining accuracy of data.

6.1 Definition

Accuracy

> *The accuracy of data is the degree to which an estimate of these data are close to the exact values of these data.*

The above-mentioned definition is derived from Eurostat (EU002, 2003) which reads as follows:
"Accuracy in the general statistical sense denotes the closeness of computations or estimates to the exact or true values".

In everyday usage statisticians use the term quality and mean accuracy. This restricts the meaning is of the term quality. In this report the term quality is used in a broad sense, including all quality dimensions of statistics.

Total error
Accuracy is also is identified as total error. Total error is calculated on the basis of bias and variance. In other words total error is a function (= dependent on) bias and variance.

The terms bias and variance are defined below (I020a, 2008).

It should be remarked that both bias and variance are connected to an estimation method. Each method that is implemented in a statistical process can cause bias and/or variance.

Bias

> *The bias of an estimate is the average[1] systematic deviation of this estimate compared to the real value of the data item.*

The term bias is also called *systematic error* or *purity*.

If bias appears several times in a process, total bias could increase as well as decrease.

Variance

[1] Avarage over different samples from the same target population. In practice this means samples in different reference periods of the same statistic.

The variance of an estimate is the degree to which the error in the estimate spreads around zero.

The term variance is also called *agility*, *precision* or *random error.*

Sampling causes variance. Non-response increases variance even more.

If variance occurs several times in a process, variance accumulates.

In addition, errors can occur in measuring variance. If, for example, high and low incomes do not respond in an income survey the measured variance of the income will be too low.

6.2 Requirements

Code of Practice

The Code of Practice (COP002, 2005) refers in principle 12 to the dimension accuracy. The CoP states that European statistics must accurately and reliably portray reality. This requirement is formulated on a general level.

No further requirements, recommendations or decisions are known regarding the accuracy of statistics.

6.3 Consequences of problems with the accuracy of the figures

What are the consequences are of problems with data accuracy? The consequences are strongly related to the use of the figures by the customer. Figures are used for very different purposes: to calculate the contribution of the Netherlands to Europe, for economic policy, indexation of contracts, etc. Problems with data accuracy have consequences not only for the CBS, but most of all for society.

Also, inaccuracy has to be discovered. It is also relevant who discovers the inaccuracy and what interest this party has in the outcome of a statistic. In the worst case, a party could make this public in the media, thus damaging the image of the CBS.

Inaccurate statistics still make it impossible to draw strong conclusions and therefore develop adequate policy and research. It the customer does not realise the consequences it is possible that conclusions will be erroneous or that the policy based on these figures will not be effective.

We can conclude that the consequences of problems with accuracy may differ from user to user, and also strongly depend on the statistics concerned.

6.4 Importance of accuracy of data for the organisation

There is no doubt that the interest of the accuracy of statistics for the organisation as well as for society is large.

That is not to say that statisticians should aim for the greatest possible accuracy. An optimum must be found between the accuracy users demand and the costs associated with accuracy of data.

In the CBS mission, the term accuracy is not literally mentioned. However, it is stated that statistical information should be undisputed. The expression disputability of statistics is not defined. We can assume that it is meant that figures should be sufficiently accurate.

6.5 Causes of problems with the accuracy of the figures

This section looks at the causes of problems with data accuracy. We shall call these problems ‘errors’ in the remainder of the text.

In the literature different classifications of possible errors can be found. These classifications sometimes have a hierarchical structure.

In this report we shall give as complete a picture as possible of possible errors. We choose the structure that we found in the Standard Quality Report of Eurostat (EU003, 2003). This structure is reproduced in the module about quality in the internal course “Boom der Statistiek van het CBS” (I008).

Errors are caused by errors in the input of statistical processes and by errors in the statistical process itself. Errors can occur in every step of the statistical process and thus affect the accuracy of data.

The statistical process also includes sub-processes that produce auxiliary information. Errors can be made in these sub-processes too. Weighing factors, for example, could be derived from auxiliary information. Errors could be make in the process of compiling weighing factors, thus affecting the accuracy of the data.

The required data accuracy can hardly be expressed in one figure. However, it is possible to set demands on the degree to which errors of different categories may occur or which measures are required to reduce these errors.

Errors caused by register errors can be distinguished from errors made in the statistical process. These two main causes are elaborated below.

6.5.1 Errors in registers

In this section register errors that affect the accuracy of data are specified.

In the case of registers, the process takes place in whole or in part outside the CBS. In even more general terms: a register is a dataset that is created outside the statistical process that we look at.

Registers can serve directly as input for the statistical process but can also serve as a sampling frame or extrapolation frame. It is characteristic of register is that the quality of its data is a given, and cannot be influenced.

Below we shall formulate the quality dimensions of registers in terms of lack of quality. For a more elaborate explanation we refer to the report Quality dimensions of Registers (Daas et al.,2008).

1. **Insufficient accuracy of the register data**. For accuracy of register data the same definition of accuracy applies as for accuracy of statistical output. It concerns the degree to which figures meet the real values. This applies not only to data items but also to classification variables. In this context, too, the terms validity or correctness are used. However, these terms add no value to the term accuracy.

2. **Insufficient consistency within a unit of a register (internal inconsistency).** Consistency within one unit is the degree to which combinations of figures within one unit (or record) are correct. This kind of error can be seen as an aspect of insufficient accuracy of register data.

3. **Insufficient coverage of the register.** Coverage is the degree to which the size of the register corresponds to the size of the target population.

4. **Insufficient fill rate of the register**. Fill rate is the degree to which the units and/or data items are filled in stead of empty. Insufficient fill rate can be caused by non-response in creating the register.

5. **Insufficient linkage of the register.** Linkage is the degree to which registers can be linked to another dataset. Values of linkage variables may be false, or double values of linkage variables may exist.

6. **Incorrect make-up of the units.** If units chosen are too large or too small this may have consequences for the accuracy of the figures in certain cases.

   If businesses are clustered too much, this may affect the turnover figures of the statistic. Turnover between parts of the cluster will not be reported and will not be seen in the statistical output.

Errors in registers are caused during creation of the register. In the creation process the same errors occur as in the creation of data in our own statistical process. Process errors are discussed below.

6.5.2 Process errors at primary data collection

This section discusses error that may arise in the statistical process of primary data collection.

1. **Sampling errors.** Sampling errors are errors in the statistical output that arise because there is no integrated observation. A sample is drawn from the whole population. An sampling error does not mean that the sample is drawn using an incorrect method.

2. **Response error.** Response errors are errors caused by non-response by part of the sample units.

3. **Measurement errors.** Measurement errors are errors caused by the way data are collected. In the literature different causes of measurement errors are distinguished. These causes of errors are elaborated in a separate section.

4. **Input errors.** Input errors are errors made at data entry of questionnaires in an information system or by importing datasets in an information system.

In the above list, the category coverage is omitted. In the literature coverage is always mentioned in this list. However, coverage errors are always caused by register errors. Coverage is already mentioned under register errors.

6.5.2.1 Measurement errors

In this section the category measurement errors is broken down. Four causes of measurement errors are distinguished. The first three are from Biemer (Biemer et al., 1991).

a. Imperfections in the **questionnaire**. Examples of imperfections in the questionnaires are: no checks on answers to questions, insufficient quality of the questions, illogical flow of the questions, unclear explanations, etc.

   Checks in the questionnaires are for example included in automated questionnaires on the internet or on the interviewers' laptops. In theory it is possible to check the correctness or probability of the answers to the questions.

b. Imperfections in the behaviour of the **interviewers**. For example if they ask questions in other words than on the questionnaire.

c. Imperfections of the **respondents**:
   - Inability to answer the question.
   - Lack of required effort to give the right answer.
   - Psychological factors

d. Imperfections of the **interaction** between interviewers and respondents.

e. Adverse effect of the approach strategy. Questions asked by phone could provide other answers than questions asked in a face to face approach.

6.5.3 Remaining process errors

In this section the category of remaining process errors is elaborated.

1. **Editing errors.** Editing errors are errors made during the correction of errors. Omission of editing can also be seen as an editing error.

   It should be noted that editing is meant to reduce errors. However, in this process errors could occur.

2. **Imputation errors.** Imputation errors are errors made when units are added to a dataset. Incorrect units may be added, e.g. non-existent units. The added units may also contain incorrect values.

   It should be noted that imputation is meant to reduce errors. However, in this process errors could occur.

3. **Classification errors.** Classification errors are errors made during the classification of units. For example, a business classified in the wrong size class.

4. **Outlier errors.** Outlier errors are errors made during the detection of outliers. Outliers may be detected wrongly or overlooked.

   Outliers may be detected at several levels of aggregation, from micro to macro level.

   It should be noted that outlier detection is meant to reduce errors. However, in this process errors could occur.

5. **Linkage errors.** Linkage errors are errors that arise during the matching datasets, e.g. false matches or non-matches.

6. **Transformation errors.** It is possible that the meaning of a population, a unit or data items changes in the course of the process. Errors may occur in this transformation process.

   In the case of secondary data collection, there will be more transformations that in primary data collection. The probability of transformation errors in secondary data collection will therefore be greater.

7. **Weighing errors.** Weighing errors are errors that arise during the calculation of weighing factors.

8. **Extrapolation errors.** Extrapolation errors are errors that arise during the extrapolation of the dataset to the target population. The size of the total population may be estimated wrongly. Extrapolation errors may also be caused by errors in the register that serve as the extrapolation frame.

9. **Computation errors.** Computation errors are errors that arise in the calculation of aggregates like totals, averages, indices, growth rates, ratios, etc.

10. **Publication errors.** Publication errors are errors that occurring in the publication of figures in tables, graphs and publications. These are errors that arise when the figures themselves are no longer edited. Figures may be incorrectly placed in a table, for example. Sometimes datasets have to be manually combined for a publication. Errors may also occur here too.

11. **Remaining processing errors.** Remaining processing errors are all errors not mentioned above. Errors may be made for example by drawing samples. These are not errors caused by sampling, which fall in a separate category of errors.

Model assumption errors. Model assumption errors are errors causes by models used. Models are based on assumptions. This category of errors is not mentioned as a separate category here. Model assumption errors may arise in several categories of errors mentioned above.

6.6 Indicators

In the previous sections 26 categories of errors were inventoried. For each category one or more indicator could be selected or developed.

The English statistical institute ONS published a list of 79 quality indicators regarding data accuracy (ONS014). From this list they selected seven key quality indicators, including, for example, unit response, item response, percentage of edits.

In choosing the indicators we thought it was appropriate to assess which category of errors contributes most to the inaccuracy of the figures. On this basis, indicators of accuracy can then be selected or developed.

Measurement errors can be estimated by:
- Experiments: activities in addition to the regular production.
- Re-approach: approaching respondents a second time
- Comparison of primary data collection with registers. In this case the question is why primary data collection is still in place.

Not all indicators are measurable (quantifiable), but if they are not the error can be described. This is important because it indicates the accuracy of data. The measuring process may also be too costly.

Code of Practice
The Code of Practice (COP002, 2005) mentions the following indicators in principle 12:
1. Source data, intermediate results and statistical outputs are assessed and validated.
2. Sampling errors and non-sampling errors (i.e. all errors) are measured and systematically documented according to the framework of the ESS quality components.
3. Studies and analyses of revisions are carried out routinely and used internally to inform statistical processes.

We incorporated these indicators in our checklist.

Handbook for Quality Reports
In the Handbook for Quality Reports (EU034b, 2008) the following indicators are mentioned:
1. Coefficients of variation
2. Rate of overcoverage

3. Edit failure rate
4. Unit response rate
5. Item response rate
6. Number of mistakes made, by type
7. Average sizes of revisions

6.7 Measures

This section discusses how accuracy of data can be enhanced by finding measures. First of all we shall mention the measures already taken, then the recommended measures.

Within the CBS numerous measures have already been taken to control the accuracy of data. It is the core of the statistical process to transform imperfect input into sufficiently accurate data.

Also, statistical methodology is focused on the dimension accuracy and less on other dimensions of the quality of statistical output.

If a statistical process is well implemented, the choice of sound methodology is a point of focus. Methodology is actually the core business of the CBS. In order to realise this it has a Methodology and Quality division (DMK), a number of “O&O” (R&D) departments. They ensure the application of the methodologies in the design of new processes.

Standard Methods
A more specific measure are the Standards Methods which have been developed by DMK. Methods that affect the accuracy of data are:

Data collection	Processing	Publication
▪ Questionnaire design ▪ Approach strategies ▪ Fieldwork ▪ Sampling theory ▪ Panels	▪ Classification ▪ Editing ▪ Imputation ▪ Weighing as correction for non-response ▪ Outliers ▪ Model/bases estimation ▪ Micro- and meso-integration ▪ Plausibility checks	▪ Macro-integration

Note that the accuracy of data changes for macro-integration. Micro-integration not only aims to enhance data accuracy. Other dimensions are also important, such as coherence and comparability in time and between countries.

CBS Quality system
The DPK department within DMK has developed the CBS Quality System to assure the accuracy of statistical output. The CBS Quality System assumes that the accuracy of statistical output depends on sound methodology but also on other factors, such as the quality of the process, the quality of the information systems, the quality of the employees and the quality of metadata.

7 Coherence of statistics

In this chapter coherence of statistical output is examined and recommendations are given to assure coherence. We start by defining the dimension coherence.

7.1 Definition

Eurostat's definition of coherence (EU034b, 2008) reads as follows: "The coherence of two or more statistics refers to the degree to which the statistical processes by which they were generated used the same concepts – classifications, definitions, and target populations – and harmonised methods".

From this we deduced the following simplified definition.

> *Coherence of statistics is the degree to which statistics use corresponding concepts and can therefore be combined.*

We depart slightly from the Eurostat definition for the following reasons:

- It does not state what coherence is, but to what it refers to.
- It is requires the use of harmonised methods. These come into view only in the matching of the figures. Other methods may yield different figures. We deal with alignment of the figures when discussing numerical consistency.

The concept of a statistic consists of the following components: population, statistical unit, reference period, classification(s) and data items.

An example of combining statistics is: two statistics that refer to the same population, the same reference period and the same geographic areas but another set of data items.

It is also possible that the data items are the same but the reference period, the geographic areas or another domain differs. In this case there is a specific type of coherence, called comparability.

Furthermore there is coherence in the strict sense, i.e. one statistical unit in different statistics. Coherence in a broad sense is that phenomena are connected in theory or in practice, e.g. wages and prices in the wage-price spiral.

The dimension coherence can also be seen as a dimension of the *methods* used: the degree to which statistics use the same methods. This report is about the quality of statistical output, and not the quality of methods. Moreover, coherence of methods is not a necessary condition for coherence in the output.

In the Standard Methods of the CBS a more drastic definition of coherence is given.

> *Coherence of statistics is the degree to which statistical output, such as that resulting from different statistics, comprises the entire field of social phenomena so that they complement each other perfectly, do not overlap and still do not contradict each other.*

This definition shows a higher ambition than our first definition. The latter definition regards statistics as puzzle pieces from the same jigsaw puzzle, with no pieces missing.

There may be differences between conceptual metadata, and these have to be resolved before statistics can be defined.

There may be coherence at the level of microdata. It is, for example, possible to match a dataset with houses and a dataset with persons to determine the number of uninhabited houses. However, this report is about coherence of statistical output and not about coherence of microdata.

7.2 Requirements

Code of Practice

Principle 14 of the Code of Practice (COP002, 2005) refers to coherence and comparability. With respect to coherence the CoP says that is should be possible to combine and make joint use of related data from different sources. This requirement is formulated on a general level.

Business and information architecture

The business and information architecture of the CBS states that the “solution” must aim at producing coherent information.

7.3 Consequences of problems with the coherence of statistics

Possible consequences of problems with coherence of statistics are that users cannot combine statistics, and thus cannot draw any conclusions and cannot use the statistics for their purposes, e.g. policymaking.

7.4 Importance of coherence of statistics for the organisation

The mission of the CBS pays explicit attention to the dimension coherence. It states that the CBS should compile and publish coherent information. It may be concluded that coherence is important for the CBS.

7.5 Causes of problems with coherence of statistics

Various causes of problems with coherence of statistics can be mentioned:

- Ambiguous definitions of populations, statistical units and data items
- Different names of data items with the same definition
- Different definitions of data items with the same names
- Differences in populations
- Differences in statistical units
- Differences in classifications or versions of these classifications
- Differences in the extent of detail
- Differences in data items
- Lack of a standard for permitted statistical units, classifications and data items
- Tables that could be published jointly are still published separately

7.6 Indicators

Code of Practice

In the Code of Practice (COP002, 2005), principle 14 recommends the following indicators with respect to coherence.

1. Statistics are internally coherent and consistent (e.g. arithmetic and accounting principles observed).
2. Statistics are compiled on the basis of common standards with respect to scope, definitions, units and classifications in the different surveys and sources.
3. Statistics from the different surveys and sources are compared and reconciled.

We incorporated these indicators in our checklist.

7.7 Measures

Standard Methods

Within the theme “Introduction to the Standard Methods and the statistical process”, the Standard Methods discuss the sub-theme coherence. It describes how complete coherence can be achieved in six steps. These steps are:

1. Clear and honest definitions
2. Uniform terminology

3. Coordinated concepts and classifications
4. Standard concepts and classifications
5. Consistent data
6. Coherent presentation

We shall discuss step 5 separately in the chapter about numerical consistency of statistics.

Repeated weighing
One of the few methodologies that apply to the quality dimension coherence is "repeated weighing". This method is incorporated in the Standard Methods and used for example in the "virtual census".

8 Comparability of data

In this chapter comparability of data is discussed and recommendations are given to assure comparability of data. We start by defining the dimension comparability.

8.1 Definition

In comparability we distinguish:
1. Comparability in time
2. Comparability between domains

Domains may be geographical domains (regions, countries, provinces, cities, etc.) and non-geographical domains (business activities, products, size class, gender, etc.). For each domain one or more classifications are known.

In the literature comparability between regions and domains are seen as separate categories. However, we see regions as a domain, too. The ONS, for example, calls a region a *spatial* domain.

In our view two figures can be compared if the data items have the same meaning and the figures have pretty much the same level of accuracy.

Comparability always refers to figures of the same data item. Only time or the domain is different.

We could apply another categorisation:
- Comparability of the definition of the data items
- Comparability of accuracy of data

This leads to the following definition of comparability.

> *Comparability of data is the degree to which figures are sufficiently accurate and the data items regarding these figures have the same definition*

Changes in the process may lead to changes in the accuracy of data. Therefore changes in the process may influence the comparability of data.

If definitions of data items differ, then the corresponding data are incomparable. Apples cannot be compared with oranges.

Eurostat definition
The above definition of comparability of data differs slightly from the definition of Eurostat (EU034b, 2008). The latter definition is:

Comparability is a special case of coherence and refers to the example where the statistical outputs refer to the same data items and the aim of combining them is to make comparisons over time, or across regions, or across other domains.

This definition does not completely meet the requirements of a definition. The main drawback is that there is no indication of what comparability is, only what it refers to. Moreover, it refers to the definition of coherence.

Lastly, although it is practical to distinguish between two types of comparability, it is not essential. Statistics always use distribution over time and domains. The basic question is can figure X be compared with figure Y.

Synonym

Instead of comparability in time the word continuity is used.

Examples
In the case of a break in series, the comparability in time between the two parts of the time series is always reduced. Indices are often subject to discontinuity, as base year shifts are a regular occurrence (for example from 1995=100 to 2000=100), and the design of the statistic is often adjusted simultaneously.

If data are comparable then differences and similarities between these data can be determined. For example:
- Unemployment in 2005 is 3 percentage points higher than in 2006 (comparability in time).
- The unemployment rate in the Netherlands in 2007 was the same as in Germany (comparability between geographic domains).

8.2 Requirements

Code of Practice
Principle 14 of the Code of Practice (COP002, 2005) refers to coherence and comparability. With respect to comparability it states that European statistics should be consistent (!?) internally and over time, and comparable between regions and countries. This requirement is formulated at a general level.

Business and information architecture
The business and information architecture states that the “solution” must assure the **continuity** of statistics. The word solution refers to the “Masterplan of the CBS”.

8.3 Consequences of problems with comparability of data

Various consequences can be pinpointed of problems with data comparability.

Indices like the consumer price index, and collectively agreed wage rates are sometimes used in long-term contracts. If these indices are not comparable over time this present problems for contract partners.

In the National Accounts comparability in time is an important requirement. The National Accounts use many source data. If these are not comparable in time this will cause problems for the National Accounts.

One example of a problem with comparability between domains is that data between European countries are not comparable. Consequently, nearly no European policy can be implemented based on these data.

8.4 Importance of comparability of data for the organisation

The importance of comparability of data strongly depends on the specific statistic. Is comparability required? Is it an important statistic? General statements cannot be made about the importance of comparability for the organisation and/or its environment.

8.5 Causes of problems with comparability of data

Various causes of problems with comparability of data can be identified.

A new time series may have been started intentionally, perhaps because a number of principles of the series are outdated after a number of years. Weightings may, for example, may have become obsolete.

There may be many and/or large changes in the statistical process, perhaps to increase the efficiency of the process or reduce the administrative burden

Problems with comparability between domains indicates a lack of coordination between parties. The meaning of data items are still not harmonised.

8.6 Indicators

In (EU028, 2005) four indicators for comparability are defined:

Indicators for comparability		
Comparability in time		
C1	Length of time series with comparable data	Number of years (or fractions) from the last break in the time series. This is an indicator for comparability in time for one statistic.
C2	Contribution of the lengths of time series with comparable data	This indicator is valid for a set of statistics. A contribution is made of the length of different time series with comparable data over different classes e.g.: 1. Less than 5 years 2. Between 5 and 10 years 3. 10 years and more This is an indicator for comparability in time for a set of statistics.
Comparability between geographical domains		
C3	Deviations in concepts and methodologies from the European standard.	This indicator has the status “for further experience”.
C4	Asymmetries between countries which measure the same flows.	This indicator refers to differences in data on incoming and outgoing flows for each pair of countries. The indicator is the sum of the absolute differences of incoming and outgoing flows measured in one country.

NB: These indicators have been replaced by the indicators mentioned in the Handbook for Quality Reports (EU034b, 2008).

Handbook for Quality Reports

In the Handbook for Quality Reports (EU034b, 2008) the following indicators are mentioned:

CC1	Length of time series of key indicators	Number of points in series without series break
CC2	Rate of adherence to ESS regulations	Number of statistical processes in precise accordance with ESS regulations as a proportion of all statistical processes subject to ESS regulations.

Code of Practice

In the Code of Practice (COP002, 2005), principle 14 mentions the following indicators:

1. Statistics are coherent or reconcilable over a reasonable period of time.
2. Cross-national comparability of the data is ensured through periodical exchanges between the European statistical system and other statistical systems.
3. Methodological studies are carried out in close cooperation between the Member States and Eurostat.

These indicators are incorporated in the checklist.

8.7 Measures

To promote comparability in time, a number of possibilities are mentioned below:

- Prevent ageing of series. This will generally mean that data such as weighing factors are kept up to date
- Change the process gradually and spread changes out over time.
- Change retro-actively: correct earlier figures.
- Parallel processes (shadow runs).
- Experimental review. This is an activity alongside the regular production activity.
- Explanation of breaks. An explanation does not correct the lack of comparability but mitigates the pain.
- Estimation of breaks through time series methods. To do this, both old and new series of some length are required. This method has recently become popular.

To promote comparability between domains it is necessary to harmonise concepts between domains. If possible, the statistical process should also be harmonised to obtain data of comparable accuracy.

Standard Methods
One theme in the Standard Methods refers to breaks in time series.

Long time series
At the CBS, the Long Time Series project offers expertise and capacity to compile longer time series.

9 Numerical consistency of data

This chapter looks at numerical consistency of data makes recommendations to assure numerical consistency of data.

We discuss numerical consistency of data separate from coherence of statistics. Statistics can be combined but after combining the statistics may contain inconsistent data.

We start by defining numerical consistency of data.

9.1 Definition

> *Numerical consistency of data is the degree to which data of different statistics that apply to the same data item equal each other.*

In (EU001, 2000), among others, the following types of coherence are distinguished:
1. Coherence between preliminary and definite data.
2. Coherence between annual, quarterly and monthly data.
3. Coherence between statistics and National Accounts

We put these types of coherence under the umbrella of *numerical consistency.*

In addition, there are statistics that compile the same data items. One would expect that the data of these statistics have the same values. We call this the fourth type of numerical consistency.

Numerical consistency matches the *one figure, one phenomenon* principle of the CBS.

Synonyms
The coherence between preliminary and definite data is also called stability (EU035, 2008). Stability is the degree to which the initial figure approaches the definite figure.

9.2 Requirements

Business and information architecture
The business and information architecture states that the solution should aim at realising the one figure one phenomenon principle.

9.3 Consequences of problems with the numerical consistency of data

Possible consequences of problems with the consistency of statistics is that users do not know which data they should use. Also, different users may draw different conclusions because they use different figures for the same data item.

9.4 Importance of numerical consistency of statistics for the organisation

The importance of numerical consistency is great, as numerical inconsistency results in disputability of data.

9.5 Causes of problems with numerical consistency of data

There are various causes of problems with numerical consistency of data:
- The number of data used to compile statistics.
- Differences in statistical methods, for example editing.
- The complexity of the process. The more complex the process, the greater the chance of inconsistencies .
- Adjusting of figures to the National Accounts

9.6 Indicators

No indicators are known.

9.7 Measures

Standard Methods

The theme “Introduction to the Standard Methods and the statistical process” in the Standard Methods discusses the sub-theme Coherence. This sub-theme describes how the state of complete coherence can be achieved six steps. Step 5 refers to consistency of data.

Statistical cube

In the Statistical cube (Booleman et al., 2005) three dimensions are distinguished:

- Reference period: predictive, month, quarter, year
- Status: preliminary and definite
- Degree of integration

Statistics about the same theme can be placed in the Statistical cube. In a certain sense the data of the statistics in this cube should be numerically consistent.

Explanation

Explaining numerical inconsistencies reduces the chance of discussion.

10 Timeliness of publication

In this chapter timeliness of the publication of statistics is discussed and recommendations are presented to assure timeliness of the publication. We start by defining the dimension timeliness of publication.

10.1 Definition

The timeliness of the publication of a statistic is optimal if the statistic is published, according to prior planning, directly after conclusion of the reference period.

> *Timeliness of the publication of a statistic is based on the period between the planned publication of a statistic and the end of the reference period.*

Our definition of timeliness differs from the definition of Eurostat, which reads: "Timeliness of information reflects the length of time between its availability and the event or phenomenon its describes" (EU002, 2003).

The reason for this difference is:

1. The Eurostat definition does not meet the requirements of a definition. It does not define what timeliness is but what it reflects.
2. A phenomenon has no time so the period between the availability and the phenomenon cannot be determined.
3. No difference is made between the planned and realised date of publication. We opt for the planned date of publication so there is no overlap with the dimension punctuality.

10.2 Requirements

Code of Practice

In principle 13, the Code of Practice (COP002, 2005) describes the dimension timeliness and punctuality. The CoP says in reference to timeliness that European statistics should be disseminated in a timely and punctual manner. The requirement is therefore formulated at a general level.

Furthermore one of the indicators of the CoP is: Timeliness meets the highest European and international dissemination standards. This requirements is also formulated at a general level, but it indicates that an ambitious timeline should be chosen.

Business and information architecture

The business and information architecture states that when realising the "solution" the quality and the timely publication of the legally obliged statistics may not be compromised in any way.

In general, it states that under all circumstances, the timeliness of the publication of the statistics should remain intact.

Other

Within the CBS it is an unwritten rule that the 1-to-1 standard be used. This means that the production period is not longer than the reference period. For example a statistic for 2008 should not be published later than the end of 2009.

10.3 Consequences of problems with timely publication

Processes of users who depend on the input of CBS products may be delayed. It is also possible that users will turn to other sources and no longer use CBS data. The relevance of data is reduced if they are published later.

10.4 Importance of publication timeliness for the organisation

The importance of timeliness can be characterised as great. Timeliness is even part of the mission of CBS.

10.5 Causes of problems with publication timeliness

Possible causes of problems with timeliness of the publication of statistical output are:

- A low level of ambition
- Dependence of input in the case of secondary data collection and auxiliary information
- A less efficient process. This may have various reasons: methodological, organisational, information systems.
- No process for production planning and control
- Availability of information systems is not assured
- No fall-back scenario in the case of calamities

10.6 Indicators

Timeliness can be measured directly by determining the period between the end of the reference period and the planned dates of publication (production period).

The production period can be divided by the reference period. For example: an annual statistic is finished 18 months after the end of the year under review: 18/12 = 1,5.

Code of Practice

In the Code of Practice (COP002, 2005), principle 13 recommends the following indicators regarding the dimension timeliness:

1. Timeliness meets the highest European en international dissemination standards.
2. A standard daily time is set for the release of European statistics.
3. Periodicity of European statistics takes into account user requirements as much as possible.
4. Any divergence from the dissemination time schedule is publicised in advance, explained and a new release date set.
5. Preliminary results of acceptable aggregate quality can be disseminated when considered useful.

We incorporated these indicators in our checklist.

Note that measures 2 and 3 are other dimensions than timeliness.

- Measure 2 regards the *predictability* of the time of publication
- Measure 3 regards the *frequency* of publication of a statistic

These measures have the factor time in common with the dimension timeliness of the publication of statistical output.

Handbook for Quality Reports

The Handbook for Quality Reports (EU034b, 2008) mentions as indicator: the number of days from the last day of the reference period to the day of publication.

10.7 Measures

This section discusses what measures may enhance timely publication of statistical output.

There are many possibilities to assure or enhance the timeliness of the publication of statistical output. Fist of all: determining the level of ambition. Furthermore:

In the design stage

- Reduce the dependency on suppliers of data
- Change methodology

In the production stage

- Eliminate organisational bottlenecks
- Production planning and control
- Enhance the adaptability and maintainability of information systems
- Change management of information systems
- Decide to publish preliminary data, provided their accuracy is acceptable
- Fall-back scenario in the case of calamities

Preliminary data
In several cases the CBS publishes preliminary data. This increases the timeliness of the publication of statistical output.

11 Punctuality of the publication

In this chapter punctuality of the publication of statistical output is discussed and recommendations are presented to assure publication of statistical output. We start by defining punctuality of the publication of statistical output.

11.1 Definition

According to (EU002, 2003) punctuality regards the period between the actual publication date and the planned publication date.

We shall define punctuality of publication of statistical output based on this definition as follows:

> *The* punctuality of publication of statistical output *is the time between the actual publication time and the planned publication time.*

11.2 Requirements

The Code of Practice (COP002, 2005) describes the dimensions timeliness and punctuality in principle 13. With respect to punctuality, the CoP states that European statistics must be disseminated in a timely and punctual manner. This requirement is defined at a general level.

No further requirements are known.

11.3 Consequences of problems with punctuality

Negative consequences of problems with punctuality will only appear if these problems are recurrent. However, in some individual cases users of certain statistics are ´waiting´ for the publication of data on the planned publication day.

11.4 Importance of punctuality for the organisation

The importance of punctuality of publication of statistical output is great because it confirms the image of a sound organisation.

11.5 Causes of punctuality problems

Possible causes of problems with punctuality of the publication of statistical output are:

- Input is not delivered on time
- Low response
- Planning not monitored
- Less capacity available than desired. No temporary capacity available in peak periods.
- Information systems are not available in time.

11.6 Indicators

Code of Practice

In the Code of Practice (COP002, 2005), principle 13 contains recommendations regarding indicators for timeliness and punctuality. However, all indicators regard the dimension timeliness.

Handbook for Quality Reports

The Handbook for Quality Reports (EU034b, 2008) mentions as indicator:
the number of days between an previously announced publication date and the actual publication date.

11.7 Measures

Possible measures for punctuality of the publication of statistical output are:

- Monitoring the deliveries of registers.

- Approaching respondent again. In the Supplementary Standard Methods this is discussed in the theme Approach Strategy.
- Monitoring realisation of the planning.
- Extra manpower available when needed
- Take releases of information systems in production in time

CBS Quality system
The DPK department of DMK has also developed the CBS Quality System. This system is also aimed to assure the punctuality of the publication of statistical output.

The CBS Quality System (2008) assumes that the punctuality of the publication of statistical output depends on the quality of the agreements with suppliers, the punctuality of the delivery of input, the soundness of the process, the availability of the right manpower, the availability of information systems and the quality of other resources.

The (standard) vulnerability analysis of the CBS Quality System states what measures could be taken in each area of interest to reduce unwanted risks.

12 Accessibility of statistics

In this chapter the subject accessibility of statistics is discussed and recommendations are given for assuring the accessibility of statistics. We will start by defining the dimension accessibility of a statistic.

12.1 Definition

In (EU002, 2003) it is stated that accessibility refers to the conditions governing the availability of figures to users: where to go, how to order the publication, delivery time, prices policy, are conditions convenient (copyright, etc.), availability of micro and macro data, available formats (paper, files, CD-rom, internet, etc.), etc.

Based on this description we shall define accessibility of a statistic as follows.

> *Accessibility of a statistic is the convenience with which users can get figures and are allowed to use these figures.*

The ease of getting access to the statistics refers to:

1. The place where the figures can be found
2. The procedure to acquire the figures
3. The delivery time of the figures
4. The price of the figures
5. Conditions like copyright
6. The medium on which the figures are available
7. The formats in which the figures are available
8. The availability of microdata

Within the CBS there is an ongoing discussion on whether microdata should be regarded as output. Nevertheless, the CoP leaves this in no doubt.

Accessibility refers to logistic aspects and not to intrinsic aspects of (the supply of) statistical output.

12.2 Requirements

Code of Conduct

The Code of Conduct (*Gedragsregels*) of the CBS states that all parties should have access to the same information at the same time. Thus the CBS takes an impartial position with regard to the various interested parties in society.

Also the results of CBS work commissioned by third parties are never put at disposal of the commissioning client only.

Lastly, in accordance with the Code of Conduct, the CBS announces publicly beforehand when new information is to be released.

Code of Practice

Principle 15 of the Code of Practice (COP002, 2005) deals with the dimensions accessibility and clarity. With regard to accessibility the CoP states that European statistics should be disseminated in a suitable and convenient manner; they should be available and accessible on an impartial basis.

No further requirements, like claims, recommendations, decisions or agreements are known with respect to the accessibility of data.

12.3 Consequences of accessibility problems

Potential consequences of problems concerning the accessibility of statistics are that users will not use the CBS figures and/or they will be dissatisfied with the CBS. If statistics are inaccessible, all the work has been done for nothing.

12.4 Importance of accessibility for the organisation

For a public organisation like the CBS it is of major importance that figures are accessible.

12.5 Causes of accessibility problems

Potential causes of accessibility problems are that no policy has been agreed and/or is pursued in this area.

A product-oriented view instead of a user-oriented view may also lead to diminished accessibility of statistics.

Lastly, the absence of agreements on accessibility may lead to problems. These agreements may be generic for all clients, as well as specific for one client.

12.6 Indicators

Code of Practice

In the Code of Practice (COP002, 2005), principle 15, recommends the following indicators with respect to the dimension accessibility:

1. Dissemination services use modern information and communication technology and, if appropriate, traditional hard copy.
2. Access to microdata may be allowed for research purposes. This access is subject to strict protocols.

These indicators are incorporated in the checklist.

Handbook for Quality Reports

In the Handbook for Quality Reports (EU034b, 2008) the following indicators are mentioned:

1. The number of subscriptions to or purchases of every important publication.
2. The number of hits or downloads of tables.

The first indicator refers to publications and not to tables and will not be included in the checklist. We see the second indicator as a more relevant indicator.

12.7 Measures

The CBS has taken already many measures to guarantee the accessibility of the figures:

- Publication of all figures on the CBS website (StatLine)
- Website accessible free of charge
- Possibility to download and print figures
- “Reproduction of figures is allowed under the condition that CBS is quoted as source”
- Figures are available in three formats: CSV, Excel and SPSS-syntax
- Free download of publications
- Hard copy publications can be ordered by e-mail, telephone or fax
- Users can subscribe to publications
- Availability (under certain conditions) of microdata at the CBS office building and at certain specified locations (remote access).

Other potential measures:

- Agreements with individual customers about making the data accessible
- Publishing the conditions under which figures are published.

13 Clarity of a statistic

This chapter looks at clarity of statistics and gives recommendations to assure clarity. We will start by defining the dimension clarity of a statistic.

13.1 Definition

In (EU002, 2003) it is stated that clarity refers to adequate metadata, illustrations like graphs and maps, availability of information on the quality of data (including restrictions for use) and the extent of additional support provided.

On basis of this description we define the clarity of a statistic as follows:

> *The clarity of a statistic is the extent to which*
>
> 1. *there are adequate metadata*
> 2. *figures are illustrated by graphs and maps*
> 3. *information on the quality of figures is available*
> 4. *restrictions for the use of figures are described, and*
> 5. *there is additional support.*

Metadata is taken to include the titles of tables and the names of rows and columns of tables. Inside StatLine these are called “items”.

We interpret adequate metadata as correct, complete, unambiguous and understandable metadata.

Synonym

Transparency of statistical output can be seen as synonymous with clarity of statistical output.

13.2 Requirements

In the Code of Practice (COP002, 2005), principle 15 deals with the dimensions accessibility and clarity. With respect to the dimension clarity the CoP states that European statistics should be presented in a clear and understandable form; they should be supplied with supporting metadata and guidance. No demands are made with respect to the language in which the metadata are published.

No further requirements like demands, recommendations, decisions or agreements are known with respect to clarity of data.

13.3 Consequences of problems with clarity

Consequences of problems with the clarity of statistics are that in the end users will not use the figures, or will use them wrongly, and will be dissatisfied with CBS.

13.4 Importance of clarity for the organisation

The importance of clarity rises as users’ expectations for the clarity of statistics become greater.

13.5 Causes of problems with clarity

Possible causes of problems with the clarity of a statistic are:

- A product-oriented view instead of a user-oriented view of the production of statistics.
- No metadata or no adequate metadata are available. This applies to conceptual metadata, methodological information on the process, as well as to information on the quality of data.

- The explanation of the StatLine table is absent, inadequate and/or hardly comprehensible.
- Inadequate knowledge, experience or attention to present statistics clearly.
- Lack of facilities to present maps and graphs.
- Lack of a helpdesk for users.

13.6 Indicators

Code of Practice

In the Code of Practice (COP002, 2005), principle 15 recommends the indicators mentioned below for the dimension clarity:

1. Statistics are presented in a form that facilitates proper interpretation and meaningful comparisons.
2. Custom-designed analyses are provided when feasible and are made public.
3. Metadata are documented according to standardised metadata systems.
4. Users are kept informed on the methodology of statistical processes and the quality of statistical outputs with respect to the ESS quality criteria.

These indicators are incorporated in the checklist.

13.7 Measures

Potential measures to enhance the clarity of statistics:

- Develop and implement standards for the description of conceptual metadata, the processing of metadata and the quality of metadata
- Provide training in making statistics presentable (‘table knowledge’)
- Review the metadata of statistics that are made public
- Provide support through a helpdesk

In the meantime, CBS staff have been offered training to write texts for publications.

StatLine

All StatLine tables have an explanation. The fixed lay-out of this explanation is:

1. Explanation
2. Definitions and explanation of symbols
3. Links to relevant tables and articles
4. Description of sources and methods
5. More information

The names of columns and rows *(items)* are always supplied with explanatory notes that can be read via a hyperlink.

The explanations of the table and the items can also be printed together with the table, if desired.

StatLine has facilities to present the figures in the form of a graph and/or a map.

Infoservice

Additional support is offered by the Infoservice. The Infoservice can be contacted via the website (e-mail), telephone 088 570 70 70, and fax. If the Infoservice is not able to answer a question, they pass it on to the appropriate department manager, project manager and/or specialist. The latter option is the third-line support by, among others, the Methodology department .

Centre for Policy Statistics (Centrum voor Beleidsstatistieken)

The Centre for Policy Statistics offers users the possibility to order tailor-made statistics.

14 Extent of detail of a statistic

In this chapter the extent of detail of a statistic is discussed and recommendations are given to assure the extent of detail of a statistic. We start by defining the dimension extent of detail of a statistic.

14.1 Definition

The extent of detail of a statistic is defined as follows:

> *The extent of detail of a statistic is the extent to which subpopulations are distinguished in the statistical output.*

The more subpopulations in a table are aggregated, the less detailed the table is.

Aggregation deals with the classification used and the levels or classes within this classification. The SBI (Standard Business Classification) often speaks of the number of *digits* used in a table.

It is also possible to aggregate data items. This will be treated under the quality dimension completeness.

14.2 Requirements

No requirements are known with respect to the extent of detail of a statistic.

14.3 Consequences of problems with the extent of detail of a statistic

The consequences of a statistic with a low level of detail are that users are not satisfied, as they cannot find what they are looking for.

14.4 Importance of the extent of detail for the organisation

The importance of the extent of detail of a statistic depends on the interest that users have in the statistic.

14.5 Causes of problems with the extent of detail of a statistic

Possible causes of problems with the extent of detail of a statistic are:

- The sample design does not take the extent of detail of the output into account.
- In the observation stage, the necessary categories are not known at the desired level of detail.

No indicators were found in the literature. Neither have any indicators been known to have been implemented within the CBS.

14.6 Measures

Potential measures for securing the extent of detail of a statistic:

- Agree with the users on the extent of detail of the statistical output.
- Consider the desired extent of detail in the sample design.
- Typify the units at the right level of detail.

15 Completeness of a statistic

In this chapter the subject completeness of a statistic is discussed and recommendations are made to assure the completeness of a statistic. We start by defining the dimension completeness of a statistic.

15.1 Definition

The completeness of a statistic is defined as follows:

> *The completeness of a statistic is the extent to which the agreements made with the user on the specifications of the statistic are adhered to.*

Completeness can refer to:
- The supplied data items
- The classifications used
- The extent of detail

The dimension completeness can also be used for the statistical programme. Incompleteness is reported if agreed statistics are not produced and published.

15.2 Requirements

No requirements are known with respect to the completeness of a statistic.

15.3 Consequences of problems with the completeness

One consequence of an incomplete statistic may be that users are not satisfied, as they cannot find what they are looking for.

15.4 Importance of the completeness for the organisation

The importance of completeness of a statistic depends on the interest that users have in the statistic.

15.5 Causes of problems with the completeness of a statistic

Possible causes of problems with the completeness of a statistic are:
- The population is not fully known
- Data items are not available:
  - Observation leads to a too high administrative burden for the respondents.
  - Data items are not included in the questionnaire
  - Data items are not included in a register
- The classification has not been used
- Inadequate specification: too few observations available to calculate reliable figures at the desired extent of detail.

15.6 Indicators

Handbook for Quality Reports

The Handbook for Quality Reports (EU34b, 2008) mentions the indicator: the ratio between the number of data items actually supplied according to the agreement, and the number of data items mentioned in the agreement. In the handbook this indicator is stated under relevance, as completeness is seen as part of relevance in this handbook.

15.7 Measures

Possible measures for securing the completeness of a statistic are:

- Make realistic agreements
- Design statistics based on the agreements with users

- Provide for sufficient observations to be able to publish at the desired level of detail
- Use the agreed classifications in the statistical process
- Make additional observations, if desired, if the original observation is insufficient
- Increase the sample (within certain strata or otherwise)

16 Confidentiality of a statistic

In this chapter the subject confidentiality of a statistic is discussed and recommendations are made to assure the confidentiality of a statistic. We start by defining the dimension confidentiality of a statistic.

16.1 Definition

> *Confidentiality of a statistic is the extent to which data on individual statistical units are no longer identifiable.*

A distinction can be made between i) identification of a unit and ii) publication of further information on the unit.

Confidentiality is not only applicable to statistical output, but to all statistical data.

Synonym
A synonym for confidentiality of data is *statistical secrecy*

16.2 Requirements

Legislation and rules
In the Act on Statistics Netherlands, section 37 includes a number of regulations on data confidentiality. We adopt these regulations integrally below.

1. Data received by the director-general, in the framework of the pursuance of tasks in accordance with this act, will be used for statistical purposes only.
2. Data as meant under 1. will not be supplied to others than those assigned to perform the task of CBS.
3. Data as meant under 1. will be disclosed in such a way that no *identifiable data about an individual person, household, enterprise or institution* can be derived, unless, in the case of data on an enterprise or institution, there is a valid reason to assume that the enterprise or institution concerned has no objection to disclosure.

A comprehensive consideration of the legal framework can be found in chapter 2 of the statistical security handbook (*Handbook Statistische Beveiliging, 2006*). In addition to the Act on Statistics Netherlands, the personal data protection act (*Wet Bescherming Persoonsgegevens*) and community legislation are also applicable.

Code of Practice
Principle 5 of the Code of Practice (COP002, 2005) deals with confidentiality. The CoP states that the privacy of providers of information (households, enterprises, administrations and other respondents), the confidentiality of the information they provide and its use only for statistical purposes, must be absolutely guaranteed.

16.3 Consequences of problems with confidentiality

Problems with the confidentiality of a statistic may have serious consequences for the CBS. The respondents' confidence may be damaged in such a way that they will reduce or stop their response. This could result in a cutback in input for the CBS.

16.4 Importance of confidentiality for the organisation

Confidentiality of statistics is of utmost importance, given the negative consequences of problems with the confidentiality of data in general.

16.5 Causes of problems with the confidentiality of a statistic

Problems with confidentiality of statistics can only be caused by the unsatisfactory application of rules for statistical confidentiality. This goes for statistical output as well as for microdata that are released for research.

16.6 Indicators

Code of Practice

Principle 5 of the Code of Practice (COP002, 2005) recommends the following indicators. We only include indicators concerning confidentiality of the statistical output.

1. Statistical confidentiality is guaranteed by law.
2. Written instructions and guidelines on the protection of statistical confidentiality in production and dissemination processes are provided. These guidelines will be made known to the public.
3. Strict protocols apply to external users accessing statistical microdata for research purposes.

These indicators are incorporated in the checklist.

16.7 Measures

Handbook

Within the CBS, the Board of Directors (DB) has approved the handbook for statistical confidentiality (*Handbook Statistische Beveiliging, 2006*). This handbook is published on the intranet site of the Methodology and Quality Division (DMK).

The handbook distinguishes between protection of:

- Microdata
- Quantitative tables
- Frequency tables
- Results of analysis

17 Remaining quality dimensions

In this chapter other quality dimensions of statistics are discussed.

17.1 Plausibility of data

Plausibility of data can be defined as the extent to which figures are plausible.

The consequences of problems with the plausibility of data are the same as the consequences of problems with the accuracy of data.

Possible causes of plausibility problems:
- Inaccuracy of the figures caused by, among other things, errors in the methodology or its implementation.
- Numerical inconsistency with other statistics.

Possible measures are:
- Using other statistics and sources.
- Announcement of noticeable (changes in the) figures.

Plausibility checks are included as a separate subject In the Standard Methods.

17.2 Disputability of data

In the Van Dale dictionary, disputable is defined as: *insecure, not founded on existing arguments.* It gives as a synonym for indisputable: *incontestable.*

Based on the definitions in Van Dale we come to the following definition. *Disputability of data is the extent to which the accuracy of a figure is opposed on the basis of arguments or perception.*

A possible argument to oppose figures or to question figures is the use of an inferior methodology.

Nevertheless, as it is always possible to debate the methodology used, complete indisputability can never be attained.

A figure may also be opposed if it differs from figures based on other statistics (whether from other sources or not). However, this difference should of course be detected first.

The CBS' Code of Conduct states that "the statistics of CBS should have an indisputable reputation".

The consequences of disputability of data are the same as the problems with the accuracy or plausibility of data.

The importance of indisputable statistical information is reflected by the fact that the supply of *indisputable* statistical information is mentioned in the CBS' mission statement.

The causes of disputability of data are the same as the causes of problems with data accuracy and plausibility.

An indicator for disputability is the number of reactions in the press on the accuracy of a figure.

Possible measures are:
- Stay ahead of reactions and give an explanation for noticeable and unexpected figures.
- Publish metadata so that it is clear how figures have been compiled.

17.3 Validity of a statistic

In general validity is connected with a measuring process. Validity then refers to the question whether *we measure what we intend to measure* (Baker, 1988).

We can see a statistical process also as a measuring process. Validity of a *statistical process* could then be understood as the extent to which the statistical process measures what we intend to measure. The quality of processes, however, is beyond the framework of this report.

The validity of a statistic - as a result of the statistical process - could nevertheless be the extent to which the figures reflect what we intend to describe. With statistics we intend to reflect reality. However, this greatly resembles the dimensions accuracy and plausibility of data. We therefore propose not to use the dimension validity of a statistic.

In the past the validity of a statistic has been applied as a performance indicator (CBS Jaarverslag 2004), i.e. for the deviation of the preliminary from the definite figures. In this report this latter phenomenon is covered by the dimension coherence of statistics.

17.4 Reliability of data

The dimension reliability is often used in combination with the dimension accuracy. The question is whether this dimension adds something more to the dimension accuracy. It could be mentioned as a synonym for the dimension accuracy (accuracy of a figure = reliability of a figure).

The Code of Conduct states that “reliable figures often mean that these are adequately accurate, sufficiently close to the (unknown) real value”. Consequently, there is a strong relationship between reliability and accuracy/plausibility.

However, in the literature reliability of the statistical process is also reported (Lesler et al., 1992). This refers to the extent to which a figure is composed in a reproducible way. It would be more clear, however, to speak of the reproducibility of a figure instead of the reliability of a figure.

In the Euro SDMX Metadata Structure (EU036, 2008) reliability is considered as “closeness of the initial estimated value to the subsequent estimated value”. In this report we give consideration to this phenomenon along with numerical consistency.

Likewise the dimension reliability is used in the context of sample surveys. Reliability or the reliability interval of a sample is then mentioned.

We conclude not to mention the dimension reliability of data further in this report.

17.5 Verifiability of data

Figures are verifiable as long as the output can be fully retraced from the input.

To realise verifiability the following is necessary:

- All datasets that have served as input for the statistical process must be available.
- The version of the software used to produce the output must be known, and what the capabilities of this software are.
- All manual alterations and handlings that have taken place in the statistical process must be known.

17.6 Reproducibility of data

The CBS’ Code of Conduct states that the CBS “models transparency among other things by reproducibility of the figures”. The dimension reproducibility, however, is not defined in the Code of Conduct.

We could define reproducibility as the extent to which figures have been compiled in a reproducible way. The more the process is performed in accordance with certain algorithms, the more reproducible the figures will be.

The more the process is performed according to a certain *fixed* methodology, the more reproducible the figures will be.

Verifiability is a prerequisite for reproducibility. In addition to this, the following is necessary:

- The correct versions of the input files are still available for processing.
- The correct versions of the software are still available and can be executed.
- The manually processed alterations can be executed once again.

Thus, stricter demands apply for reproducibility than for verifiability.

17.7 Availability of data

Availability of data says something about the existence of data. Certain users are, for example, interested in ‘old’ figures. The question then is, are these still available for the users.

We shall not elaborate this dimension further, as there is not a great deal to be said about it.

1 Annex 1: Checklist for individual statistics

This annex lists indicators and measures that can be applied to individual statistics.

M = Measure. The question is whether a certain action has been executed: no/yes.

I = Indicator. Is the result of a measuring process in the form of a measuring process (qualitative) or in the form of a score (quantitative)?

1.1 Relevance

A1	M	**Evidence of agreements**	Have agreements been laid down with the user of the statistic?
A2	M	**Completeness of agreements**	Will the following questions be answered in the agreements: a. What objective does the user have with the statistic? b. What will be supplied by the CBS (population, data items, specifications, reference periods)? c. What quality should the statistical output have in terms of accuracy, coherence, comparability, numerical consistency, timeliness and punctuality? d. How will the statistic be supplied (medium, format)? e. What are the future needs of the user? f. When and how will agreements be evaluated and brought up to date?
A3	M	**Timeliness of agreements**	Have the agreements with the user been evaluated and brought up to date in the last two years?
A4	M	**Compliance with agreements**	Did the statistic meet all the agreements made with the user? This indicator applies if agreements have been made with an individual user. This indicator applies also for the statistics made on the basis of a regulation. Source: Uitwerking kwaliteitskader HPE (I021, 2008; H3.1)
A5	M I	**User satisfaction**	1. Is the user's satisfaction with the statistics measured systematically? Source: Code of Practice (COP002, 2005). 2. What is the user's satisfaction score? Source: Uitwerking kwaliteitskader HPE (I021, 2008; H3.2).
A6	I	**Use of StatLine**	What is the number of hits for the statistic in StatLine? Source: Handbook for Quality Reports (EU034b, 2008; indicator AC2)

1.2 Accuracy in general

B1	M	**Agreements on accuracy of figures**	1. Have agreements (SLAs, covenants, etc.) been made on accuracy of data with the users of the statistic? 2. Will all categories of errors that apply to the figures be addressed? 3. Have standards with respect to possible errors been set up, for example, sample error and non-response?
B2	M	**Evidence of errors**	Have the categories of errors made in the process been described?
B3	M	**Assessment of most important categories of errors**	Has an estimation been made about which categories of errors have the most influence on the accuracy of the figures?
B4	M	**Quantitative indicators**	Have quantitative indicators been applied for the measurement of the most important errors?
B5	M	**Qualitative indicators**	Have the errors been described qualitatively, if they are not or hardly measurable, or if the costs for measuring the indicator are too high?
B6	M	**Plausibility control**	1. Has the output been compared with the output of other internal or external statistics? 2. Has the output been reviewed in relation to known developments in society? 3. Has the output been compared with previous periods?
B7	M	**Quality reports**	Have quality reports been compiled in which the accuracy of the figures is justified? Source: Code of Practice (COP002, 2005)
B8	M	**Revisions**	Have reviews been used to make systematic improvements in the process? Source: Code of Practice (COP002, 2005)
B9	M	**Standardised methods**	Have methods been used that are in the CBS' series of methods and that affect the accuracy of data?
B10	M	**CBS Quality system**	1. Has the CBS' Quality system been used to determine what measures – besides a solid methodology – are needed to achieve sufficient accuracy of the figures? 2. Is the quality document up to date, that is to say not older than one year for critical processes and not older than two years for non-critical processes? The CBS' Quality system is the successor to Procesbeheersing/VIR. A critical process contributes substantially to an image-relevant statistic.
B11	M	**Process metadata**	Has the process been described methodologically? This description includes which knowledge rules, other rules, methods, etc. have been applied. It can be used to analyse where errors may occur.
B12	M	**Checks on data**	Have input, throughput and output been reviewed and validated? Source: Code of Practice (COP002, 2005)

1.2.1 Register errors

B13	M	**Accuracy of figures**	Have audits been performed on the accuracy of important data items in the register and have the outcomes of these audits been described?
B14	I	**Coverage (surplus)**	To what extent is the number of the real population in accordance with the registered population? What is the estimated percentage of surplus coverage? Source: Handbook for Quality Reports (E034b, 2008; indicator A2) and Uitwerking Kwaliteitskader HPE (I021, 2008; H3.2.2).
B15	I	**Coverage (shortage)**	1. What percentage of records is included in the source framework? 2. What is the selectivity measure for shortage coverage? 3. What is the maximum distortion for shortage coverage? 4. What is the maximum RMSE for shortage coverage? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.2.3 Onderdekking). 5. What is the percentage of shortage coverage?
B15	I	**Filling**	1. What percentage of the units is not filled? 2. What percentage of the individual data items is not filled?
B16	I	**Linkability**	1. What percentage of records is linkable? 2. What is the selectivity measure for linkability? 3. What is the maximum distortion for linkability? 4. What is the maximum RMSE for linkability? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.2.2.4 Koppelbaarheid). 5. What percentage of double values of linking variables occurs in the register? 6. What percentage of linking variables does not lead to a link? 7. What percentage of linking variables leads to an incorrect link?

1.2.2 Process errors at first observation

B20	I	**Sampling errors**	1. What is the standard error of the sample estimation? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.2.1) 2. Within what margins are the figures 95% reliable? 3. At what level of detail are the figures still precise enough? 4. Is there a standard for the sampling margin?
B21	M I	**Non-response (unit)**	1. What percentage of records respond to at least one item? 2. What is the selectivity measure for unit response? 3. What is the maximum distortion for unit non-response? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.2.5 Unit-nonrespons) 4. What percentage of units drawn is not included in the figures? Source: Handbook for Quality Reports (E034b, 2008; indicator A4). 5. Is there a reminder procedure? 6. Does the reminder take place in accordance with the (supplementary) series of methods? 7. Is there a standard for non-response? N.B.: Increasing response does not automatically lead to more precise figures, as selectivity cannot increase at the expense of representativeness. Reminder procedures are dealt with under the subject Approach strategy of the S-series.

B22	I	**Non-response (items)**	1. What percentage of records respond to the core variable? 2. What is the selectivity measure for item response? 3. What is the maximum distortion for item non-response? 4. What is the maximum RMSE for item-non response? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.2.6 Item non-response) 5. What percentage of records respond overall to the core variable? 6. What is the selectivity measure for the ultimate response? 7. What is the maximum distortion for the ultimate response? 8. What is the maximum RMSE for the ultimate response? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.2.9 Gevoeligheid) 9. What percentage of every item (variable) is not filled? Source: Handbook for Quality Reports (E034b, 2008; indicator A5) NB: In most cases the measurement of unit non-response is sufficient for the most important data items.
B23	M I	**Questionnaire**	1. Has a *Vragenlab* (question laboratory) test been executed and implemented? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.2.7 Meting) 2. Does the questionnaire contain imperfections and have these imperfections been described?
B24	I	**Interviewers**	Might interviewers make mistakes in the statistic and have these mistakes been described?
B25	I	**Respondents**	Might respondents have difficulties in answering the questions and have these difficulties been described?
B26	I	**Interaction**	Do adverse effects occur for this statistic in the interaction between interviewer and respondent and have these effects described?
B27	I	**Approach strategy**	Do adverse effects occur for this specific statistic with the chosen approaching strategy and have these effects been described?

B28	M	**Estimation of measurement errors**	1. Has parallel testing taken place? 2. What is the relative measurement error? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.2.7 Meting) 3. Have measurement errors been estimated by means of experiments (outside the production process), repeated approach of respondents, or comparison with registers?
B29	I	**Input errors**	What is the percentage of input errors at data entry of the questionnaires?
B30	I	**Covariance**	What are the covariance coefficients? Source: Handbook for Quality Reports (E034b, 2008; indicator A1).

1.2.3 Other process errors

B30	I	**Editing errors**	1. What is the percentage of violated edit rules on the item? 2. What is the percentage of reviewed units on the item? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.2.8 Verwerking) 3. What percentage of units qualifies for editing? Source: Handbook for Quality Reports (E034b, 2008; indicator A3) 4. What errors occurred at editing (automatically and/or manually) and have these errors been described? 5. What percentage of figures is edited per variable? 6. What percentage of units is selectively edited?
B31	I	**Imputation errors**	1. What is the percentage of imputed units on the item? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.2.8 Verwerking) 2. What percentage of units qualifies for imputation? Source: Handbook for Quality Reports (E034b, 2008; indicator A3) 3. What errors occurred in the imputation of units (non-existing units, figures too high or too low) and have these errors been described?
B32	I	**Classification errors**	What errors occurred in the individual classification variables at the assignation of classes to units (typification) and have they been described?

B33	I	**Outlier errors**	1. What errors occurred at the outlier detection and have they been described? 2. What percentage of wrong outliers has been detected? 3. What percentage of outliers has been overlooked? Remarks: If outliers are overlooked, distortion occurs. If outliers are wrongly noticed and not included in the total, variance increases.
B34	I	**Weighing errors**	What errors occurred in the compilation of weighing factors and have these errors been described?
B35	I	**Raising errors**	What errors occurred in the raising of a dataset to the target population and have these errors been described?
B36	I	**Transformation errors**	What errors occurred in the transformation of the meaning of the population, the statistical unit and/or the data items in the process, and have these errors been described?
B37	I	**Computation errors**	What errors occurred in the computation of totals, means, indices, growth percentages, ratios and other derivations? Have these errors been described?
B38	I	**Publication errors**	What risks were incurred with the transformation of an output dataset to a published table, and have these risks been described?

1.3 Coherence

The indicators and measures apply to one statistic in coherence with other statistics in the same domain.

C1	M	**1-to-1 relation between name and definition of components of a statistic**	1. Is there one definition per component of the statistic? 2. Is there one name per component of the statistic? Components of a statistic are: population, statistical unit, reference period, classifications, specifications and data items.
C2	M	**Standard statistical units**	Is there a standard for statistical units? For example: is there a universal definition for business unit or household?
C3	M	**Standard population**	Is there a standard for populations?
C4	M	**Standard classifications**	Is there a standard for (versions of) classifications? This also applies to the classification of the time factor.
C5	M	**Standard for detailing**	Is there a standard for the specification of the figures?
C6	M	**Standard data items**	Is there a standard for data items? Which data items are permitted and which are not?
C7	M	**Relatibility of data items**	Are data items relatable to each other? Is it possible to derive data items from two other data items? For instance, do wage and the number of employees refer to the same concept of employee?
C8	M	**Coherence in presentation**	Are statistics which cohere with each other also presented in one table?

C9	M	**Repeated weighing**	Are statistics attuned to each other, for example by repeated weighing?
C10	M	**Unique definitions**	Are the population definitions, the statistical unit, the (versions of the) classifications, the categories and the data items uniquely defined? Unique definitions are prerequisite to determine coherence. This is not an indicator for coherence itself, but an indicator for whether this requirement is met.
C11	I	**External coherence**	Is there coherence with statistics outside CBS?
C12	I	**Similar population and statistical unit**	Is there coherence with statistics referring to the same population and statistical unit?

1.4 Comparability

Comparability is distinguished in the sub-dimension comparability in time and comparability in domains (geographical or non-geographical).

1.4.1 Comparability in time

D1	I	**Length of time series**	1. What is the length in time of the series to be compared? 2. What is the minimum perceptible difference in time between estimations? 3. What is the exceeding chance in time at a break in methods? 4. Is the break in methods a result of reviews in the statistical process? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.5 Comparability 5. What is the length of the time series of the statistic without break in the series, as long as the statistic has existed? Source: Handbook for Quality Reports (EU034b, 2008; indicator CC1).
D2	M	**Updating time series**	1. Has outdating of a series been prevented? 2. Did this change take place gradually? Explanation: Preventing outdating means e.g. that helpful information (e.g. weighing factors) is kept up to date.
D3	M	**Retroaction**	Have modifications in the figures been executed retroactively? Have older figures from the same time series been adjusted?

1.4.2 Comparability between domains

D4	I	**Asymmetry between countries**	What are the differences in incoming and/or outgoing flows between countries in the statistics of the corresponding countries?

D5	M	**Attuning concepts**	Have concepts been attuned with other statistics, so that concepts are comparable? Attuning may take place within CBS as well internationally.
D6	M	**Attuning methods**	1. Have methods been attuned with other statistical processes, so that comparable accuracy of the figures can be obtained? Attuning may take place within as well as outside CBS. 2. Have methodological studies been performed jointly?

1.5 Numerical consistency

E1	M	**1-figure notion**	Is only one figure published for each statistical phenomenon? In a statistical phenomenon, the metadata are fully identical: population definition, unit, reference period, classification, specification and variable.
E2	I	**Preliminary and definitive figures**	What has the difference been between preliminary and definite figures in the course of time?
E3	I	**Revisions and corrections**	What has the size been of revisions and corrections? Source: Handbook for Quality Reports (EU024b, 2008; indicator A6 p73)
E4	I	**Annual, quarterly and monthly figures**	What have the differences been between annual, quarterly and monthly figures?
E5	I	**National accounts**	1. What have the differences in results been between business statistics and the national accounts? 2. Have the differences been explained?
E6	M	**Statistical cube**	Has the statistical cube (Booleman et al., 2005) been applied as a tool to make numerical consistency visible.
E7	I	**External source**	1. What is the exceeding chance for no deviation from the external source? 2. What is the break with the external source? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.6 Coherence)

1.6 Timeliness

F1	M	**Agreements on period under review**	1. Were agreements in place with the user on the reporting period? 2. Have these agreements been met?
F2	I	**Customers satisfaction**	Was the user satisfied with the timeliness?

F3	I	**Period under review versus reference period**	1. What was the timeliness of the final report? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.3 Timeliness) 2. How long was the period between the end of the reference period and the (planned) date of publication? Source: Handbook for Quality Reports (EU034b, 2008; indicator T1) 3. What was the mean production time divided by the reference period in the last five cycles of the statistic?
F4	M	**Stability in publication moments**	Were planned publication moments delayed by changes in the process (like in a redesign)?
F5	M	**Frequency**	Was the frequency of the statistic attuned with the user and has this agreement been recorded? Source: Code of Practice (COP002, 2005)
F6	I M	**Preliminary results**	1. What is the timeliness of the preliminary results? Source: Uitwerking Kwaliteitskader HPE (I021, 2008; H3.3 Timeliness) 2. Have preliminary results been published? 3. Did these preliminary results have an acceptable accuracy? Source: Code of Practice (COP002, 2005)
F7	M	**Production planning**	Was there a production timetable for the whole chain from which it was evident that the planned publication time is attainable?
F8	M	**International standards**	Were the standards of Eurostat and/or other international organisations met? Source: Code of Practice (COP002, 2005)

1.7 Punctuality

G1	I	**Punctuality**	How many days were there between the planned publication date and the actual publication date? Source: Handbook for Quality Reports (EU034b, 2008; indicator T2)
G2	M	**Agreements with suppliers**	1. Were there agreements on the planned delivery dates with suppliers of register data and have these agreements been recorded? 2. Were these agreements no more that two years old?
G3	I	**Delivery reliability of suppliers of registers**	What was the delivery reliability of the suppliers of the registers?

G4	M	**Response time**	1. Were respondents reminded if they did not answer within the fixed time? 2. Was the reminder timely? 3. When were respondents reminded? And how did this moment relate to the first approach moment?
G5	M	**Production planning and control**	1. Was the production of the whole chain subject to a time schedule? 2. Was the production of the whole chain under control? 3. Were there any adjustments of the process in the case of delay?
G6	M	**Availability of information systems**	1. Were information systems operational in time, for example after changes? 2. Were information systems sufficiently available in day time?
G7	M	**Flexibility of capacity**	Was it possible to use temporary extra capacity if bottlenecks threatened?
G8	M	**Regression scenario**	Was there a regression scenario in the case of calamities?

1.8 Accessibility

H1	M	**Agreements**	Have agreements been made with individual customers on the way data were to be supplied? Think of: ▪ The place where the figures are made available ▪ The procedure to obtain the figures ▪ The delivery time of the figures ▪ The price of the figures ▪ Conditions like copyright ▪ The medium on which the figures are placed ▪ The formats in which figures are available Source: Code of Practice (COP002, 2005)

1.9 Clarity

J1	M	**Review of statistic**	Have the explanation to the table and the items been reviewed for ambiguity and readability?
J2	M	**Completeness of conceptual metadata**	Have all the conceptual metadata of the statistic been explained? Think of: a description of the population (definition), the statistical unit or object type, used classifications (version, deviation from standard), reference periods, definitions of data items. Source: Code of Practice (COP002, 2005)

J3	M	**Completeness of process meta**	Has the methodological process of the statistic been fully explained? Source: Code of Practice (COP002, 2005)
J4	M	**Completeness of quality meta**	Have all the quality dimensions of the output been described in the explanation to the table? Think of: dimensions relevance, accuracy, coherence, comparability, timeliness and punctuality. Relevance also covers the restriction in use of a statistic. Source: Code of Practice (COP002, 2005)
J5	M	**Maps and graphs**	Was the option to present the figures in the form of maps and/or graphs used optimally? Source: Code of Practice (COP002, 2005)
J6	M	**Restrictions**	Have restrictions on use of the figures been described for the user? Source: Code of Practice (COP002, 2005)
J7	M	**Clarity of title**	1. Is the title of the table meaningful? 2. Is any information in the table relevant for choosing the table by the user? 3. Does the title meet the editorial requirements? Source: Beoordelingskader StatLine tabellen (I022, 2008)
J8	M	**Initial presentation of tables**	1. Will a representative idea of the table be given when it is first opened? Have the rows and columns been chosen well? 2. Will the latest information be shown? 3. Is the initial presentation of the table conveniently arranged? Not too extensive? The above goes for the theme pages as well as for the dossiers and StatLine itself. Source: Beoordelingskader StatLine tabellen (I022, 2008)
J9	M	**Table**	1. Can the table and the terms used be understood quickly by the user? 2. Is the table well-designed? Is the number of nestlings not too large? 3. Is the combination of data items in one table logical? 4. Does the table have too many empty cells? 5. Does the table contain timely figures, (excl. historical figures)? Source: Beoordelingskader StatLine tabellen (I022, 2008)
J10	M	**Classifications**	1. Are there dual categories? 2. Is the meaning clear of the category Unknown? Source: Beoordelingskader StatLine tabellen (I022, 2008)

J11	M	**Symbols in cells**	1. Is the indication 'empty cell' well-chosen? The figure should then not occur on logical grounds. Source: Beoordelingskader StatLine tabellen (I022, 2008)
J12	M	**Figures in cells**	1. Are all the figures in the cells significant (no false accuracy)? Source: Beoordelingskader StatLine tabellen (I022, 2008)
J13	M	**Explanation to tables**	1. Have the table items that are explained been chosen correctly? 2. Are the explanation texts in the right section? 3. Are the listings limited in size? 4. Does all the text in the explanation have enough information value? 5. Are the explanation texts comprehensible (short sentences, correct Dutch, no jargon, concrete)? 6. Have the sources as well as the methods in the section "Description of sources and methods" been well-described? 7. Are the texts in the explanation not repeated unnecessarily? 8. Are figures that are not plausible explained? 9. Is the text of the explanation consistent with the contents of the table? 10. Does the explanation not contain obsolete texts? 11. Does the explanation not contain texts that may easily become obsolete? Source: Beoordelingskader StatLine tabellen (I022, 2008)
J14	M	**Corrections**	Does the explanation to corrections have enough information value for the user? Source: Beoordelingskader StatLine tabellen (I022, 2008)

1.10 Extent of detail

K1	M I	**Agreements on level of detail**	1. Have agreements been made on the level of detail of the statistic? 2. Will the agreed specification be applied?
K2	M	**Adequate sampling design**	Has the desired extent of detail of the statistical output been taken into account in the sample design?
K3	M	**Adequate typification**	Is classification done at a low enough level at observation? For example: if classification is done at 2-digit level, later on it is not possible to aggregate at 3-digit level.

1.11 Completeness

L1	I	**Agreements on data items**	1. What is the ratio between supplied data items and agreed data items? Source: Handbook for Quality Reports (EU024b, 2008; indicator R1) 2. Have agreements been made on data items to be delivered? 3. Will the agreed data items be supplied?
L3	M	**Agreements on classifications**	1. Have agreements been made on the classifications to be applied? 2. Will the agreed classifications be applied in the statistic?

1.12 Confidentiality

M1	M	**Data protection policy**	Is the statistic secured in conformity with the Handbook Statistische Beveiliging (handbook on statistical security)?

1.13 Plausibility

N1	M	**Plausibility control**	Is a plausibility check part of the regular process?
N2	I	**Relationship with NA**	Will the unabridged figures be incorporated in the NA?

1.14 Disputability

O1	I	**Comments in the press**	How many comments on data accuracy have appeared in the press in the last three years?

2 Annex 2: Checklist for the statistical programme

This annex contains indicators and measures applicable to the statistical programme or to parts of it.

2.1 Relevance

P1	**Evidence of agreements with the CCS on the statistics in StatLine**	Are there any agreements with the CCS on the contents and quality of <u>all</u> statistics published in StatLine? Source: Code of Practice (COP002, 2005).
P2	**Satisfaction of CCS with the statistics at StatLine**	Is the satisfaction of the CCS with the statistics in StatLine measured systematically? Source: Code of Practice (COP002, 2005).
P3	**Scope of the portfolio**	Has a policy been formulated on the type of statistics CBS wants to produce and does not want to produce?
P4	**Potential demand of users for information**	Is a process in place in which information on potential needs of users of statistics is gathered systematically? Are the results of this discussed periodically by the CBS management? Source: Code of Practice (COP002, 2005).

2.2 Accuracy

P5	**Audit conventions on accuracy of data**	Will all error categories come up for discussion in audits on data accuracy?

2.3 Coherence

P6	**Sets of reconciled statistics**	Is there a description of which statistics can in principle be combined and therefore could be presented in one table?
P7	**Standards**	1. Are there any standards concerning populations, units, classifications, data items, reference periods? 2. Are these standards applied? Source: Code of Practice (COP002, 2005)
P8	**Comparison**	Are statistics compared and attuned to each other systematically? Source: Code of Practice (COP002, 2005)

2.4 Comparability

P9	**Average length of time series**	1. What is the mean or the mode of the length of the time series of statistics of BES, SRS and MSP? Which are the shortest and the longest time series? 2. How have these indicators developed in the last years?

P10	**Compliance with ESS rules**	What percentage of statistics complies with ESS rules? Source: Handbook for Quality Reports (EU034b, 2008; indicator CC2) If a statistic complies with ESS rules, the percentage can be compared with that of other ESS countries.

2.5 Numerical consistency

No indicators and measures concerning the statistical programme have been included.

2.6 Timeliness

P11	**Fixed moment of publication**	Has a fixed time in the day been set for the publication of a statistic? For example: Are figures published at 9:00 a.m.? Source: Code of Practice (COP002, 2005)
P12	**Deviation from schedule**	1. Is every deviation from the publication schedule announced and explained beforehand? 2. Will a new publication date also be set? Source: Code of Practice (COP002, 2005)

2.7 Accuracy

No indicators and measures concerning the statistical programme are included.

2.8 Accessibility

P13	**Conditions**	Have the conditions been published under which data are generally published? These conditions may concern the following items: ▪ The place where the figures are made available ▪ The procedure to obtain the figures ▪ The delivery time of the figures ▪ The price of the figures ▪ Requirements like copyright ▪ The medium on which the figures are published ▪ The formats in which the figures are available These apply to aggregated data as well as microdata. Source: Code of Practice (COP002, 2005)
P14	**Availability of a website**	What is the availability percentage of StatLine on the Internet during 7x24 hours?
P15	**Numbers of visitors and development**	1. How many StatLine visitors are there? 2. How has the number of StatLine visitors developed? It is plausible that the number of visitors also depends on the accessibility of StatLine.
P16	**Search options**	Is it quickly clear to the user what search options he has?

P16	**Search tree**	1. Can tables be found through a search tree? 2. Does the search tree have a logical structure? 3. Do all the branches in the tree have comprehensible terms?
P17	**Keywords**	1. Can tables be found by keywords? 2. Do keywords give good scores? Not too many, not too few hits?
P18	**Identification**	Do tables have a unique reference identification?
P19	**Retracing known tables**	Can known tables easily be retrieved?
P20	**Output**	Can tables be printed and downloaded?
P21	**Views**	Can users compile an own view of each table?
P22	**Copyright**	Is the copyright on the figures regulated clearly?
P23	**Microdata**	Is there a procedure for users to have access to microdata?

2.9 Clarity

P24	**Standards**	1. Are there any instructions for the composition of explanations to the tables and the items? 2. Is training given in the application of the standard?
P25	**Metadata systems**	Are there any standardised metadata systems? Are all metadata documented in them? Source: Code of Practice (COP002, 2005)
P26	**Custom-made**	Can analyses be performed and published, tailored to specific users? Source: Code of Practice (COP002, 2005)

2.10 Extent of detail

No indicators and measures are included that refer to the statistical programme.

2.11 Completeness

P27	**Completeness of statistical programme**	What percentage of statistics is agreed on but not yet published?

2.12 Confidentiality

No indicators and measures have been included that refer to the statistical programme.

2.13 Plausibility

No indicators and measures have been included that refer to the statistical programme.

3 Annex 3: Relationship between characteristics of statistical output

In this annex, the interrelations between characteristics of statistical output are considered. In the overview in this annex the relations are described schematically.

3.1 Relevance and accuracy

A relationship exists between the relevance and the accuracy of data. In one respect statistics may be too inaccurate to be relevant to users, in another respect a less accurate statistic may still be relevant for users.

Although figures become more relevant as they become more accurate, there is a point beyond which more accuracy no longer results in more relevance. The accuracy of the figure is enough, then, to speak of relevant figures.

3.2 Relevance and coherence

There is a relationship between coherence and relevance of statistics. The greater the coherence of statistics, the greater the relevance. After all, more data items and figures can be referred to in combination with each other.

A reverse relationship is also possible. By increasing coherence, specific elements in a statistic can be forfeited, which makes the statistic less relevant. In this respect, coherence goes hand in hand with standardisation of e.g. the data items.

3.3 Relevance and extent of detail

Statistics are irrelevant if data are not sufficiently specified.

3.4 Relevance and comparability, timeliness, accuracy, completeness, clarity and accessibility.

The greater the comparability, timeliness, accuracy and completeness, the more relevance increases, up to a certain limit.

Statistical output is also more relevant as accessibility and clarity increase.

3.5 Accuracy and timeliness

More timeliness leads, in general, to a lower accuracy (‘trade off’).

For methodologists it is, however, a challenge to increase timeliness while preserving accuracy. As a rule this requires adaptation of the applied methodology.

3.6 Accuracy and comparability

Moreover, there is a relationship between comparability of statistics and the accuracy of data. If figures are comparable but the accuracy differs, it is hard to compare these figures with each other, and this could be reported as a break in series.

This phenomenon may occur when the concepts remain the same but the process changes substantially, for instance as a result of the transfer from primary to secondary observation.

3.7 Accuracy and extent of detail

A more detailed statistic may lead to inaccuracy in the figures if it is not taken into account in the design of the statistic.

3.8 Accuracy and plausibility

There is a strong relationship between plausibility and accuracy of data. Inaccurate figures are more likely not to be plausible either. However, even accurate figures have a risk of being judged as non-plausible. This may happen when more

credence is given to figures from another source. The judgement of data plausibility has a strong subjective element.

Plausibility is determined differently than accuracy of data. For accuracy the production process of the figures is considered. For plausibility a comparison is (also) made with other statistics and own perceptions on the phenomenon measured ('professional judgement').

3.9 Accuracy and disputability

From the definition of disputability is turns out that there is a relationship between disputability and accuracy of data. With disputability, accuracy may be disputed.

3.10 Accuracy and confidentiality

To guarantee confidentiality, figures are sometimes made less accurate.

3.11 Accuracy and clarity

A not very accurate statistic is less clear.

3.12 Coherence and comparability

A relationship exists between coherence and comparability of data. If figures are not comparable, there is no coherence either. Comparability can be seen as part of coherence.

3.13 Coherence and numerical consistency

There is a relationship between numerical consistency of data items and coherence of statistics. Numerical consistency of statistics can be reported only when these statistics can be consolidated.

3.14 Coherence and completeness

Incompleteness may result in statistics not being able to be combined meaningfully. Therefore completeness is good for coherence.

3.15 Numerical consistency and plausibility

There is a relationship between plausibility and numerical consistency of statistics. Numerical inconsistency leads to loss of plausibility of figures.

3.16 Numerical consistency and disputability

A relationship exists between numerical consistency and disputability of statistics. Inconsistent figures are disputable.

3.17 Timeliness and punctuality

There is a relationship between timeliness and punctuality. If timeliness is formulated too ambitiously, punctuality comes under pressure. It then becomes more difficult to supply a statistic punctually.

3.18 Accessibility

No relations are found between accessibility and other characteristics of statistical output.

3.19 Clarity and completeness

A more complete statistic is also clearer.

3.20 Completeness and confidentiality

There is an obverse relationship between the confidentiality and completeness of a statistic. It may sometimes be necessary not to publish part of the figures, in favour of the confidentiality of a statistic.

3.21 Extent of detail and confidentiality

A more detailed statistic increases disclosure risk, and will be at the cost of confidentiality.

3.22 Plausibility and disputability

Plausibility opposes disputability. Plausible figures will not quickly be disputed. Indisputable figures will be judged as plausible.

Table: Relationship between characteristics of statistical output																		
	Relevance	Accuracy	Coherence	Comparability	Numerical consistency	Timeliness	Accuracy	Accessibility	Clarity	Completeness	Extent of detail	Confidentiality	Plausibility	Disputability	Validity	Reliability	Verifiability	Reproducibility
Relevance		X	X	X	X	X	X	X	X	X	X							
Accuracy			X		X	X			X		X	X	X	X				
Coherence				X	X					X								
Comparability																		
Numerical consistency													X	X				
Timeliness							X											
Accuracy																		
Accessibility																		
Clarity										X								
Completeness												X						
Extent of detail												X						
Confidentiality																		
Plausibility														X				
Disputability																		
Validity																		
Reliability																		
Verifiability																		
Reproducibility																		