eurostat

Doc. Eurostat/A4/Quality/00/General/Standard report
Orig. : EN
Available in **EN**

# ASSESSMENT OF THE QUALITY IN STATISTICS

**Luxembourg - 4-5/04/2000 at 9h 30**

Room Quetelet, BECH-building

**Item 4 of the agenda**

## Standard quality report

# CONTENTS

# STANDARD QUALITY REPORT

## INTRODUCTION

This paper sets out the proposals for a quality report on statistics. It describes the subjects that should be covered in a report on quality, and lists particular information that Eurostat expects Member States to provide. Such a report makes Member State deviations from Eurostat norms explicit. Some background to the proposal below is included in the documents "Eurostat/A4/Quality/00/General/Definition" and "Eurostat/A4/Quality/00/General/Glossary".
The standard quality report aims at covering general topics for the assessment of the quality. **The implementation of such reports has to be done within domains regarding weaknesses, relevance, priorities and costs.**

## 1. RELEVANCE OF STATISTICAL CONCEPTS

To make sure that the produced statistics remain relevant for the users, the Member States should conduct regularly a survey of the opinion of their main users. This information may come from regular liaison committees / steering groups which consider needs expressed by users.

The main statistical concepts, which are of interest to users are detailed as follows:
The *Statistical measure* is a summary (means, mode, total, index, etc.) of the individual *variable* values for the statistical units in a specific sub-population.
The *target population* is the population of interest. Normally the statistics are presented for different sub-populations, so called *domains*. These domains can be geographical as well as non-geographical. Often these domains are according to classifications.
Population, statistical units and variables relate to specific times, which may be limited to a *reference time point* or a *reference period*.

| To be reported on a multi-yearly basis: |
| --- |
| • A summary including users' description, origin and satisfaction of user's needs, and relevance of statistical concepts for the users. |

## 2. ACCURACY

The overall accuracy is one of the main needs of users. This information should be given by the publication of confidence intervals together with statistics. The confidence intervals should take into account all the effects of remaining errors.
The remaining errors are broken down into sub-components as follows:

1. Sampling errors;
2. Non sampling errors
    1. Coverage errors;
    2. Measurement errors;
    3. Processing errors;
    4. Non response errors;
    5. Model assumption errors.

Since **the overall accuracy is of primary interest**, this breakdown according to error sources is a way of collecting information. Member States should provide information on the different types of errors.

By definition, the sampling variance is associated with an estimator. Corrections are often done in order to reduce bias using statistical techniques. They may be included in the estimator. Other errors occurring at the time of data collection or data entry can not be corrected neither by editing process nor by estimators, but they should be taken into account in a confidence interval.

In order to obtain comparable confidence intervals between Member States, the same components should be included in the variance estimation process. Statistical methods are not enough developed in order to deal with the variance estimation of all components as a whole or separately. Therefore, the implementation should follow the improvement of statistical techniques.

The assessment of the accuracy of individual data from administrative sources is similar to the one for data collected for statistical purpose, even if may be more difficult to apply.

System of account are compiled using many sources and under balancing constraints. Methodological researches are in progress in order to assess the accuracy. Therefore, the accuracy of system of account is not yet considered in this document.

Selection of references:

[1] Biemer, P., Groves, R.M., Lyberg, L.E., Mathiowetz, N.A. and Sudman, S. (eds) (1991) *Measurement Errors in Surveys*, New York: Wiley
[2] Cox B. G. Binder D. A., Chinnappa B. N., Christianson A., Colledge M.J. and Kott P. S. (eds) (1995) *Business Survey Methods*, New York: John Wiley.
[3] Groves, R.M. (1989) *Survey Errors and Survey Costs*, New York: John Wiley
[4] Lyberg L., Biemer P., Collins M., De Leeuw E., Dippo C., Schwarz N., Trevin D. (eds) (1997). *Survey Measurement and Process Quality*. Wiley.
[5] Särndal C-E, Swensson B, & Wretman (1992). *Model Assisted Survey Sampling*. Wiley.

## 2.1 Sampling errors

### *2.1.1 Probability sampling*

If we assume we are estimating from a random probability sample, and if we ignore all other sources of error (coverage error, measurement error, etc.), then a measure of the accuracy of an estimate is the square root of the mean square error. The mean square error of an estimator is equal to the variance of the estimator plus the square of the bias. Obviously, whenever the methods of sampling and measurement produce an unbiased estimator, then the mean square error is equal to the variance of the estimator.

Probability samplings are used where registers are available and accuracy can be estimated for variables related to statistical units. Sampling in several stages may use non-probability sampling for drawing units at one stage of a multi-stage design.

#### 2.1.1.1 Estimator bias

Estimators of totals are often unbiased. For example, estimates produced using the Horvitz-Thompson estimator are unbiased. Frequently, however, estimates are calculated using the ratio or regression estimator, and these are biased but consistent. Occasionally more robust, but biased, estimators are used, for instance outlier robust estimators.

| To be reported on a multi-yearly basis: |
|---|
| • Yearly biases due to the estimation method, when measurable.<br>• The yearly median and range of bias for quarterly and monthly statistics. |

#### 2.1.1.2 Sampling variance

For most of the standard sample designs (for example simple random or stratified random sampling) used with standard estimation methods (for example Horvitz-Thompson or ratio estimator), there are standard formulae to estimate the variance and standard error associated with an estimator.

For complex sample - multistage or multi-phase sampling - it may occur that a stage of the selection procedure is not a probability sample. The sampling variance can be estimated for the stages where probability sampling is used. However, this leads to an under-estimation of the variance. The variance of the non probability selection stage may be roughly approximated under the assumption of probability sampling. Another possibility is to assess the impact of this selection stage under next item: non-probability sampling.

| To be reported:<br><br>Standard error of: | Frequency of statistics | | |
|---|---|---|---|
| | Monthly (resp. Quarterly) | Yearly | Multi-yearly |
| Total (resp. Level for an Index) | Y | Y | Y |
| Growth rates Consecutive periods | Y | Y | |
| Same period for consecutive years | Y | | |
| Frequency of Transmission to Eurostat: | Monthly (resp. Quarterly) | Yearly | Multi-yearly |

*2.1.2 Non probability sampling*

In some circumstances, non-random sampling is used. National Statistical Institutes may use expert samples (also called judicious) based on a high coverage of relevant characteristics (for instance production, employment, turnover...).
In these circumstances it is impossible to obtain an objective assessment of the accuracy of the estimates. Nevertheless, some rough accuracy indicators of quality can be designed using sensitivity analyses. When units are selected with certainty following a structural auxiliary variable, such as yearly value added, a more sophisticated one could be built using an econometric model in order to estimate the effect of units not selected.

| To be reported | Transmission frequency |
|---|---|
| • Estimates of the effects | Same as statistics |
| or failing that: | |
| • The yearly coverage rate | Yearly |
| • An accuracy indicator | Same as statistics |
| • The methods used to obtain these indicators | Multi-yearly |

## 2.2 Non sampling errors

*2.2.1 Coverage errors*

Coverage errors are due to the disagreement between the frame and the target population and sub-populations.
The frame is the listing or listings of units that delimit, identify, and allow access to the elements or sets of elements of the target population.
When statistics are directly based on a statistical register whatever the sources used (Censuses, administrative data), the statistical register is considered as the frame.
By extension to statistics of flows, when they are supposed to be exhaustively registered, the frame is defined as the list of recorded flows.

Coverage errors include overcoverage, undercoverage, according to the target population, and misclassification according to the target sub-populations.
They may be caused by the inherent limitations of input data, or by delays and errors in data acquisition and processing.

***Overcoverage*** relates either to wrongly classified units that are in fact out of scope or to units that do not exist in practice.

***Undercoverage*** refers to units not included in the frame.
Undercoverage due to cut-off sampling or cut-off threshold for data collection is included in model assumption component (3.2.5).

*Misclassification* refers to wrong classifications that belong to the target population. The following distinction is important:
- When the frame information is used for the classification of units, there is an error as long as the erroneous classification is used.
- When the frame information is used for the sampling design but not for the statistics (based on corrected classification mainly resulting from the sample), misclassifications lead to a higher sampling variance.

These misclassifications may reflect economic, social phenomena or special cases, and should be corrected or not. What are the rules for taking account of misclassifications? How do these corrections affect statistics? Do they occur when statistics revised?
For stratified sample, the misclassified units lead to estimates for a domain over several strata. The corresponding sampling variance should take into account the variance of all units of the considered domain over the several initial strata.

| To be reported: | Transmission frequency |
|---|---|
| • The effects of undercoverage , over coverage, and misclassification on the main characteristics | Yearly |
| • The yearly rate of undercoverage , over coverage, and misclassification broken-down according to the main characteristics of the population | Yearly |
| • The methods used to obtain these figures. | Multi-Yearly / when changes occurs |
| • Methodological notes on the processing of misclassifications. | Multi-yearly / when changes occurs |

### *2.2.2 Measurement errors*

Measurement errors are errors that occur at the time of data collection. There are a number of sources of measurement error, including:
- the *survey instrument* (the form or questionnaire),
- the *respondent*,
- the *information system* (respondent's report-keeping system),

- the *mode of data collection* (face to face interviewing, telephone interviewing, self-administered mail survey, diary surveys, administrative records, direct observation, and electronic observation),
- and the *interviewer*.

This breakdown is use to collect information on measurement errors, but there is no need to report for each sub-component when they are all taken into account in an overall assessment.

In practice, the errors due to the respondent are difficult to separate from questionnaire and interviewer effects. It requires to ask again the same questions to the respondents (for a sub-sample) in order to measure the response variability. Bias could be assessed by more expensive methods using appropriate staff and mode of data collection in order to obtain "true" values for a sub-sample.

*Errors due to the survey instrument*
The quality of a questionnaire is difficult to assess in an objective way.
A description of the pilot survey and the conclusion of the analysis of the questionnaire by a cognitive laboratory are necessary in order to assess the questionnaire.
The primary objective of cognitive laboratory research methods is to identify through careful analysis questioning strategies that will yield more accurate answers.
For instance, a well-known type of measurement error occurs in a survey where individuals are asked to report their age. There is a tendency to report age data that are rounded off. It is preferable to ask the date of birth.

*Errors due to the reporting unit*
The respondent may unconsciously give incorrect response. For example it may concern the effects of recalling past events (*memory error*), the tendency to over-report characteristics perceived to be valued by others (*social desirability effects* or *conditioning*), the lack of attention of the respondents, the effects of age and education etc.
Statisticians try to minimise these errors by improving questionnaire design and interviewer training.
When individual administrative data are used, the errors made when filling the administrative form have to be considered.

*Errors due to the information system and mode of data collection*
The information system of the respondent and the use of administrative record may not allow to obtain an accurate measurement of a concept. For instance:
- The reference period for which statistics are requested may differ from the period data are based on. (Calendar year and accounting year)
- The measure of pure price evolution needs to take account of quality changes. (for instance using hedonic methods)
- The correspondence between the administrative concept and the statistical concept. (for example the delineation of unit)
The assessment of the inaccuracy due to the mode of data collection can be made by comparing different data collection methods on the same surveyed units. Information on the respondent information system is necessary for assessing how the concepts used fit the statistical concepts.
The mode of data collection has to be especially considered if electronic data interchange (EDI)), computer-assisted telephone or personal interviewing (CATI CAPI) are not used.

*Errors due to the interviewer*

The interviewer may influence the respondent in such way that measurement errors results. Groves [3] distinguishes four interviewer effects. First, interviewer demographic and socioeconomic characteristics can affect the behaviour of the respondent. Second, interviewers can administer the questionnaire in different ways, especially where computer-assisted telephone or personal interviewing (CATI CAPI) are not used. Third, interviewers can emphasize different words. Fourth, in reaction to respondent difficulties with the questionnaire, interviewers assist the respondent in different ways.

Systematics errors, (ie: errors common to the interviewers) are rare when interviewers are trained. However, the use of CAPI-CATI increases training costs and reduces the number of interviewers, that may lead to systematic errors.

The random error due to the interviewers can be cheap to measure by using interpenetrating sub samples (See [5] p627).

Special techniques may be used to assess measurement errors (re-interview, estimation of an “interviewer effect”, etc.). But the costs of this sort of exercise (re-interview) are often prohibitively high, both for data providers and data collectors. The estimation of the bias due to measurement errors need to obtain "true values" for a sub sample. Re-interview may be followed by a reconciliation process where the respondent indicate the true value from both interviews. See Forsman & Schreiner (in [1])

Rao & Sitter (in [4]) compare different variance estimator for the measurement bias based on subsample (re-interviews) in order to obtain trues value.

Very often, NSIs limit their investigation to consistency edits between characteristics of the questionnaire (or between characteristics of the questionnaire and any external information of good quality).

| To be reported on a multi-yearly basis: |
|---|
| • The variance and bias due to the reporting unit, questionnaire design, interviewer, model used to correct measurement errors. |
| or failing that: |
| • Specific studies done by Member States on these errors; |
| • Methods used to reduce this kind of error; |
| • Error rates drawn from the editing process by sources of measurement errors |

### *2.2.3 Processing errors*

Processing errors are errors in post-data-collection processes such as coding, keying, editing, weighting, and tabulating.
For coding, some measures of the error rates can be achieved through standard techniques of quality control, for instance by verifying the quality of a sub-sample of the processed questionnaires (to check the level of errors either during the keying phase or during the processing of edits by NSI staff).
For editing, weighting, estimation and tabulating processes, the errors are difficult to measure.
The description of the edits should consider the breakdown into macro-edits/micro-edits, the latter being divided into (missing values/errors/anomalies). Missing values and errors are fatals edits and anomalies are query edits according to the distinction made by Granquist (in [2]): Fatal edits identify data that are clearly erroneous, whereas query edits point to data that have a high probability of being erroneous.

| To be reported on a multi-yearly basis: |
|---|
| • The variance and bias due to processing errors.<br>Or failing that:<br>• The rates of remaining processing errors broken down according to the different steps.<br>• Methodological notes on their estimation<br>Or failing that:<br>  • The description of the edits system<br>  • The rates of failed edits breakdown into missing values, errors and anomalies. |

### *2.2.4 Non-response errors*

**Item non-response** and **unit non-response** may bias the estimates in an uncontrolled way. A traditional and simple indicator of non-response is the response rate, either by unit or by item. It can be either weighted (by the sampling weights or other variables such as numbers of employees, turnover etc.) or non-weighted. The distinction between actual non-respondents and out-of-scope units is very important. It would be advisable to keep as a separate group the category of units for which the NSI does not know if the non-measurement is a non-response or a coverage error.
When administrative data are used, one considers the non availability of an administrative data as a unit non-response and missing data in a record as an item non-response.
An important question is how respondent units differ from non-respondent ones. One kind of non-responses is due to data coming late. The effect of such non-responses can be measured by comparison a posteriori. The second kind of non-response can be investigated through a supplementary survey of a sub-sample of non-respondents, generally using more efficient data-collection techniques (for instance, interviews instead of mail inquiries).

Non-responses (item- or unit-) have to be processed during the production process, either through imputation (partial or total) or through reweighting of the responding units. This process introduces randomness, which should be integrated in the estimation of biases and variances.

| To be reported | Transmission frequency |
|---|---|
| • The variance due to imputation and reweighting | Same as statistics |
| • Unit and item response rates | |
| • A description of the methods used for imputation and/or reweighting for non-response | Multi-yearly |
| • Unit and item non-response bias | |
| • Methodological notes on the above estimations | |

### *2.2.5 Model assumption errors*

Model assumption errors occur with the choice of methods, such as the use of auxiliary variables for the estimators, calculation based on full scope or constant scope, benchmarking, seasonal adjustment and other models not included in the preceding accuracy components, in order to calculate statistics or indexes.
The quality of such models should be assessed in term of variance and bias. This evaluation should be integrated in the confidence intervals related to statistics, but there is no standard method for this integration.

Sampling estimators using auxiliary variables such as calibration may be used to deal with several kinds of errors. For these estimators, the classical sampling variance is estimated regardless of the accuracy of the auxiliary variables. Errors in the auxiliary variables may introduce errors in statistics, which should be assessed together with sampling errors.

Cut-off thresholds and unavailability of data for sub-populations
Cut-off threshold are often used to avoid burden on reporters who contribute very few to the requested statistics. It may be a deliberate choice of NSIs or stated in regulation.
When no data collection for a sub-population is available, when a part of needed data are not yet available (for instance the weights for index calculation), one may chose an estimator that compensate for the lack of data.
Therefore, estimated statistics are considered as proxies for the measure of the characteristics on the target population.

Seasonal adjustment model. A very important point related to the quality of seasonal adjustment is the detection and correction of outliers prior the seasonal adjustment. Outliers of aggregated data are due to irregular phenomena (such as a strike) and have to be corrected for estimating the seasonal components.

Benchmarking:
Benchmarking procedures are generally used to produce revised infra-annual series consistent with yearly statistics. This method should be described when it is applied. The comparison between initial values and final value should be reported so as to assess the impact of the method, and consequences for the accuracy of the benchmarked series should be evaluated (diminution of bias, increase of variance).

Other models may be used for the estimation of part of the statistics for instance when aggregated administrative data is used.

The selection procedure of these models (i.e. why a given model has been chosen in preference to alternative ones), together, perhaps, with the associated estimation error of the corresponding estimates, should be reported to Eurostat.

| To be reported on a multi-yearly basis for each model used to adjust the series |
|---|
| • Biases and variances introduced by the models.<br>Or failing that, comments on:<br>• The verification of the assumptions underlying the model<br>• The test of the predictive power of the model by using historical data to 'predict' known quantities<br>• The comparison of the results generated by the model with other related sources of data<br>• The use of screening and cross-validation studies<br>• The tests of sensitivity of the model to parameters' estimation<br>• The validation of the data inputs to the model |

## 3. TIMELINESS AND PUNCTUALITY

In order to analyse the impact of each step in the production process, and with a view to facilitating the discussion of delays with the Member States or developing advance estimates at the European level, the following information would be useful:

- the use of electronic data transmission for the statistical data collection in the Member State
- the legal deadline imposed on respondents in the Member State
- key data-collection dates, specifying when the questionnaires and recalls and follow-ups were sent out and when the field work took place;
- a graph of response in terms of number of questionnaires received by week or by working day
- starting and finishing dates for the editing phase;
- dates for the imputation phase;
- details and date of the event which indicates that the data collection phase can be considered to have ended;
- date on which the advance results were calculated and disseminated;
- date of the quality check (congruency of results) and non-disclosure measures;
- date of availability of the camera-ready publication
- date of publication

Punctuality of data transmission to Eurostat will be evaluated according to delays stated in regulations, or in other agreements.

## 4. ACCESSIBILITY AND CLARITY

Eurostat is interested in the following information on the dissemination process:
- a copy of the publication(s)
- information on what, if any, results are sent to reporting units included in the sample
- information on the dissemination scheme for the results (e.g. to whom the results are sent, who the users are)
- a copy of any methodological documents relating to the statistics provided.

This information is useful not only in order to design a consistent dissemination scheme at Eurostat level and to better inform users calling from Member States, but also as a way of obtaining any national comments included in these publications which may assist in interpretation of the statistics within Eurostat. The information will also be of help in reporting on the benefits of the availability of statistics.

## 5. COMPARABILITY

The comparability component aims at measuring the impact of differences in applied statistical concepts and definitions when statistics are compared between geographical areas, non-geographical domains, or reference periods.
Systematic deviations in definitions and methods used for the building of statistics that have an impact on these comparisons should be assessed and provided to the users. Random errors, which are measured under accuracy component have to be kept in mind during the assessment process.

### 5.1 Geographical comparability

Where there is a difference between National concepts and European concepts (definition of statistical units, reference population, classifications, definition of observed characteristic), details of the differences in terms of concept and estimation of consequences for resulting statistics should be mentioned[1].

Similarly, Member States should report if they follow the existing European recommendations for measurement, and possible deviations from these norms and the corresponding assessment of consequences on the estimates.

Combination of both pieces of information will provide an assessment of the level of comparability of the statistics.
Some complementary (voluntary?) survey on a small sub-sample should assess deviations from the European Standards.

[1] When national statistics are derived not only from a survey sampling estimate but also from a model to adjust for some discrepancy between the statistic produced by the survey and the objective aimed at by the Member State, the adjustment due to choice of model should be included in paragraph 2.2.5 "model assumption errors". Only the differences between the European norms and the norms aimed at after the possible national model adjustment have to be reported in this paragraph. To get more detailed information, see document Eurostat/A4/Quality/98/General Comparability.

| To be reported on a multi-yearly basis or when significant changes occur: |
|---|
| • A comparison between Member state practices and Eurostat recommendation.<br>• An assessment of the effects of these differences.<br>• Comments on results |

### 5.2 Comparability between domains

Users frequently compare statistics based on domains, which are often defined according to classifications. For instance these classifications may be related to economic activities, size classes, products, modes of transport, sexes, etc.
The difference in concepts used for the estimation of the statistics should be reported. It concerns mainly the definition of the characteristic, the reference period, the definition of the statistical unit, and the statistical measure.
These consequences of these differences are of main importance for the comparability of statistics between domains.

| To be reported on a multi-yearly basis for the main comparisons between domains. |
|---|
| • The differences in concepts.<br>• An Assessment of the effects of these differences.<br>• Comments on results |

### 5.3 Comparability over time

Some users, especially those working in the macroeconomic field, use time series rather than point estimates for a date. The stability of the concepts and methods of measurement is very important to them. Details of changes in definitions, coverage or methods should be provided. When the comparison of statistics from two consecutive periods may be more affected by legal events than by the socio-economic trends, it is important to report on this aspect. Any change in legislation at the national level having consequences on continuity should be reported. The consequences of non-negligible changes should also be reported.
For seasonal adjusted series, comparison should take into account changes in adjustment procedures, but the seasonal adjustment model should be reported under paragraph 2.2.5 "Model assumption errors".

| To be reported on a multi-yearly basis or when significant changes occur: |
|---|
| • The difference in concepts and methods of measurements between last and previous reference period.<br>• An Assessment of the effects of these differences.<br>• Comments on results |

## 6. COHERENCE

Like the comparability component, this one aims at measuring the impact of differences in applied statistical concepts and definitions, but for a wide range of uses of statistics. It focuses on the joint use of statistics that are produced for different primary purposes.
Systematic deviations in definitions and methods used for the building of statistics may lead to meaningless joint uses of statistics. The impact of such deviation should be assessed, keeping in mind random errors, which are measured under accuracy component.

### 6.1 Coherence between provisional and final statistics

Provisional and final statistics are in most cases based on the same concepts and data collection method. However, for provisional statistics the available information is lesser and the processing must be quicker. Therefore, the lack of coherence between provisional and final statistics is mainly due to difference in accuracy.

| To be reported on a multi-yearly basis: |
|---|
| • Yearly differences for the main characteristics broken down by sampling errors, coverage errors, measurement errors, processing errors, non-response errors, and model assumption errors.<br>• Comments on results |

### 6.2 Coherence of annual and short-term statistics

For several characteristics, statistics have to be produced with both infra-annual and annual frequencies. It is important for the image to the users to aim at a certain coherence of the disseminated information. That means lack of consistency between annual and short term estimates have to be avoid as much as possible. A simple validation method consists in comparing estimates of the same concept but derived from these statistics of different frequencies. It seems most natural to compare average annual levels or totals when both frequencies provide estimators in level, and growth rates when at least one of these statistics is an index. These comparisons should be made before possible benchmarking by the NSI, and the Member State should indicate separately if it benchmarks its infra-annual statistics on the considered annual statistic.

| To be reported on a multi-yearly basis: |
|---|
| • A comparison on an annual basis of statistics and growth rates if relevant taking into account the overall accuracy which should be estimated for both kind of statistics.<br>• If differences are not fully explained by the accuracy components, differences in national concepts should be investigated and assessed. |

### 6.3 Coherence of statistics in the same socio-economic domain

Frequently, a group of statistics, possibly of different type (in monetary value, in volume or constant price, price indicators) brings a representation of the same phenomenon under different angles. It is very important to check that these representations are compatible in order to anticipate users' questions and to prepare corrective actions.

For instance, business short-term statistics like turnover, value-added, variations in stocks should be compared on a yearly basis with Prodcom statistics after deflation by production prices. Statistics on population, unemployment and employment should be compared. Similarly, opinion surveys and quantitative indicators should be compared. Employment as measured by social statistics (censuses for instance) and by business statistics should ideally have limited discrepancies. Surveys on individuals' holidays should have some consistency with quantitative indicators of tourism statistics, or short-term indicators in the service sector (Horeca). Etc.

| To be reported on a multi-yearly basis: |
|---|
| • Yearly differences for the common characteristics according to accuracy component and differences in national concepts.<br>• Comments on results |

For statistics of flows (e.g. trade, transport, balance of payment, tourism) the use of mirror statistics give an idea of the accuracy through discrepancies or asymmetries between statistics based on two sources. However, there are often differences in the concepts used. For instance date for registration due to the time lag, cif/fob registration for external trade, definition of population in the countries etc.

| To be reported on a yearly basis: |
|---|
| • Summary of the mirror statistics<br>• Estimation of the asymmetries due to the differences in concepts and due to accuracy.<br>• Comments on results |

### 6.4 Comparison of statistics with national accounts

In order to be in a position to advise users on the source of information best suited to their needs, it may also be interesting to compare statistics from the surveys to national accounts. So an indication of the overall consistency of these surveys with other sources of information could be provided to users. The methodology used in the national account statistics would need to be described for the considered statistics, including the primary data source and the adjustments that are made. Discrepancies in the concepts should also be taken into account.

| To be reported on a multi-yearly basis: |
|---|
| • A summary of the comparison. |

## 7. COMPLETENESS

The completeness component assesses the suitability between domains of available statistics and needs and priorities expressed by the users of the European Statistical System. It is assessed at the general level and should be based partly on relevance quality component, comparability and coherence. However, for each domain the rate of available statistics in comparison with the requested statistics by regulation or other means is needed for the assessment. Eurostat should calculate and distribute this information. In the framework of standard quality report of statistical products, Member States are not asked to assess this component at national level.

| To be reported: |
| --- |
| • In case of non-available statistics, the reasons why. |

## 8. RELATED INFORMATION: BURDEN, COST AND BUDGET

### 8.1 Burden to respondents

Member States should provide Eurostat annually with the burden to respondents of providing the information requested. It should be expressed in hours, days (24 hours), weeks (7 days), months (30 days) or years (12 months).
For enterprises, the total cost to the population should be provided using the average cost per enterprise by the sample size.

If relevant, indicators for measuring the dispersion of the total statistical burden between individual units should be provided.

### 8.2 Costs and budget for the statistical office

Member States should provide Eurostat with the total costs associated with the survey. On a multi-yearly basis, information on the number of employees (broken down by statisticians, methodologists, programmers and analysts, interviewers, clerical staff) should be provided. Computer resources (hardware and software) and other processing equipment (Optical Character Reading equipment, etc.) used should also be detailed.

## 9. TO REPORT: SUMMARY

| |
|---|
| **1. Relevance of statistical concepts**<br>A summary including users' description, origin and satisfaction of user's needs, and relevance of statistical concepts for the users |
| **2. Accuracy**<br>*2.1 Sampling errors*<br>*2.1.1 Probability sampling*<br>Standard error of total and growth rates |
| *2.1.2 Non probability sampling*<br>Estimates of the effects<br>or failing that:<br>The yearly coverage rate,<br>An accuracy indicator,<br>The methods used to obtain these indicators |
| *2.2 Non sampling errors*<br>*2.2.1 Coverage errors*<br>The effects of undercoverage , over coverage, and misclassification on the main characteristics.<br>The yearly rate of undercoverage , over coverage, and misclassification broken-down according to the main characteristics of the population<br>The methods used to obtain these figures.<br>Methodological notes on the processing of misclassifications. |
| *2.2.2 Measurement errors*<br>The variance and bias due to the reporting unit, questionnaire design, interviewer, model used to correct measurement errors.<br>or failing that:<br>Specific studies done by Member States on these errors;<br>Methods used to reduce this kind of error;<br>Error rates drawn from the editing process by sources of measurement errors |
| *2.2.3 Processing errors*<br>The variance and bias due to processing errors.<br>Or failing that:<br>The rates of remaining processing errors broken down according to the different steps. Methodological notes on their estimation<br>Or failing that:<br>The description of the edits system. The rates of failed edits breakdown into missing values, errors and anomalies |

*2.2.4 Non-response errors*

The variance due to imputation and reweighting. Unit and item response rates

A description of the methods used for imputation and/or reweighting for non-response. Unit and item non-response bias

Methodological notes on the above estimations

*2.2.5. Model assumption errors.*

Biases and variances introduced by the models.

Or failing that, comments on:

The verification of the assumptions underlying the model

The test of the predictive power of the model by using historical data to 'predict' known quantities

The comparison of the results generated by the model with other related sources of data

The use of screening and cross-validation studies

The tests of sensitivity of the model to parameters' estimation

The validation of the data inputs to the model

**3. Timeliness and punctuality**

the use of electronic data transmission for the statistical data collection in the Member State

The legal deadline imposed on respondents in the Member State

Key data-collection dates, specifying when the questionnaires and recalls and follow-ups were sent out and when the field work took place;

A graph of response in terms of number of questionnaires received by week or by working day

Starting and finishing dates for the editing phase;

Dates for the imputation phase;

Details and date of the event which indicates that the data collection phase can be considered to have ended;

Date on which the advance results were calculated and disseminated;

Date of the quality check (congruency of results) and non-disclosure measures;

Date of availability of the camera-ready publication

Date of publication

**4. Accessibility and clarity**

A copy of the publication(s)

Information on what, if any, results are sent to reporting units included in the sample

Information on the dissemination scheme for the results (e.g. To whom the results are sent, who the users are)

A copy of any methodological documents relating to the statistics provided.

**5. Comparability**

*5.1 Geographical comparability*

A comparison between Member state practices and Eurostat recommendation.

An assessment of the effects of these differences.

Comments on results

*5.2 Comparability between domains*

The differences in concepts. An Assessment of the effects of these differences. Comments on results

*5.3 Comparability over time*

The difference in concepts and methods of measurements between last and previous reference period.. An Assessment of the effects of these differences. Comments on results

**6. Coherence**

*6.1 Coherence between provisional and final statistics*

Yearly differences for the main characteristics broken down by sampling errors, coverage errors, measurement errors, processing errors, non-response errors, and model assumption errors. Comments on results

*6.2 Coherence of annual and short-term statistics*

A comparison on an annual basis of statistics and growth rates if relevant taking into account the overall accuracy which should be estimated for both kind of statistics.

If differences are not fully explained by the accuracy components, differences in national concepts should be investigated and assessed.

*6.3 Coherence of statistics in the same socio-economic domain*

Yearly differences for the common characteristics according to accuracy component and differences in national concepts. Comments on results.

For statistics of flows

Summary of the mirror statistics. Estimation of the asymmetries due to the differences in concepts and due to accuracy. Comments on results

*6.4 Comparison of statistics with national accounts*

A summary of the comparison

**7. Completeness**

In case of non-available statistics, the reasons why.

## 10. INDEX OF TERMS DEFINED IN THE GLOSSARY[2]



[2] See Doc Eurostat/A4/Quality/00/General/Glossary