Implementation Glossary

Version 1.0

March 2006

GLOSSARY OF TERMS

Eurostat already has a glossary of statistical terms called CODED (http://forum.europa.eu.int/irc/dsis/coded/info/data/coded/en.htm). Most terms in the glossary below, as well as their sources, have been extracted from it. Only those terms considered of relevance for the NACE implementation project have been included.

Note: This glossary includes definitions and acronyms. The numbers in brackets are referred to the information sources, which are listed in the annexe.

A

- ACTR (14)

    Automatic Coding by Text Recognition

- Administrative data (3)

    Administrative data is the set of units and data derived from an administrative source.

- Administrative source (3)

    An administrative source is the organisational unit responsible for implementing an administrative regulation (or group of regulations) corresponding register of units and the transactions are viewed as a source of statistical for which the data.

B

- Back casting (14)

    A break in time series occurs in a particular point in time, for instance due to a change of classification. Back-casting is the process that allows to re-construct the time series back in time from this point in terms of the new classification.

- Base period (18)

    Base period is the period of time for which data used as the base of an index number, or other ratio, have been collected. This period is frequently one year but it may be as short as one day or as long as the average of a group of years.

- Base weight (18)

    The weights of a weighting system for an index number computed according to the information relating to the base period instead, for example, of the current period.

- Base year (9)

    The base year is the year used to determinate the pattern of prices on which the constant prices figures are calculated.

- Business register (*Statistical business register*) (10)

  Statistical business registers are lists of enterprises and other units, as required by the registers Regulation or recorded on voluntary basis, whose activities contribute to the Gross Domestic Product of the Member State. All Member States of the European Union maintain business registers for statistical purposes. According to Article 1 of the registers Regulation "Member States shall set up for statistical purposes one or more harmonised registers with the definitions and scope specified in the following Articles".

  C

- CATI (12)

  Computer Aided Telephonic Interviewing

- Central Product Classification (CPC) (4)

  The Central Product Classification (CPC) of the United Nations is a complete product classification covering goods and services.

- Changes in classifications and structure [Statistical methodologies] (17)

  A new version of a classification differs in essential ways from the previous version. Essential changes are changes that alter the borders between categories, i.e. a statistical unit may belong to different categories in the new and the older version. Border changes may be caused by creating or deleting categories, or moving a part of a category to another category.

- Classification (16)

  A classification is a set of discrete, exhaustive and mutually exclusive items, which can be assigned to one or more units to be measured in the collation and/or presentation of data. The terms "classification" and "nomenclature" are often used interchangeably, despite the definition of a "nomenclature" being narrower than that of a "classification".

- Classification of activities (5)

  The main purpose of a classification of activities is to classify productive economic activities. The main aim is to provide a set of activity categories that can be utilised when presenting statistics according to such activities.

- Classification of Products by Activity (CPA) (13)

  The CPA is the official product classification by activity of the European Union.

- Coding (14)

  Coding is a technical procedure for converting verbal information into numbers or other symbols which can be more easily counted and tabulated.

- Coding error (14)

  The assignment of an incorrect code to a survey response.

- Confidential data (1)

Confidential data is defined in Article 13 of Council Regulation No 322/97, as:

1. Data used by the national authorities and the Community authority for the production of Community statistics shall be considered confidential when they allow statistical units to be identified, either directly or indirectly, thereby disclosing individual information.

To determine whether a statistical unit is identifiable, account shall be taken of all the means that might reasonably be used by a third party to identify the said statistical unit.

The two main reasons for declaring data to be primary confidential are:
- Too few units in a cell
- Dominance of one or two units in a cell.

The limits of what constitutes "too few" or "dominance" vary between statistical domains.

- Conversion matrix (12

It is a theoretical correspondence that presents probabilities of conversion from an A code to different B codes. Some ancillary information is used for calculating these probabilities, for instance turnover, number of employees, etc.

- Correspondence (theoretical) (12)

A theoretical correspondence between two classifications (A and B), or two versions of the same classification, is made of pairs of codes, one from each classification. When a pair is included, this means that the content (or part of it) of the A code is included in the content of the B code.

D

- Data confrontation (19)

Data confrontation is the process of comparing data that has generally been derived from different surveys or other sources, especially those of different frequencies, in order assess their coherency, and the reasons for any differences identified. Data confrontation may also take place between statistics produced in different countries. Such processes may or may not attempt to quantify the impact of any differences identified.

Such data may not be coherent for a number of reasons including the use of different data item definitions, classifications, scope, reference period, etc.

- Data dissemination (12)

Dissemination is the release to users of information obtained through a statistical activity.

Data dissemination consists of distributing or transmitting statistical data to users.

- Dual coding (12)

Data coding in two different versions of the classification simultaneously.

- Double reporting (12)

  Transmission of aggregates referred to two different classifications (or versions of the same classification) simultaneously.

  E

- ECB (13)

  European Central Bank

- Economic Activity (NACE REV 1.1)

  An activity is said to take place when resources such as equipment, labour, manufacturing techniques, information networks or products are combined, leading to the creation of specific goods or services. An activity is characterised by an input of products (goods or services), a production process and an output of products.

  In practice the majority of units carry on activities of a mixed character. One can distinguish between three types of economic activity:

  - Principal activity: The principal activity is identified by the top down method as the activity which contributes most to the total value added of the entity under consideration. The principal activity so identified does not necessarily account for 50% or more of the entity's total value added.

  - Secondary activity: A secondary activity is any other activity of the entity that produces goods or services for third parties.

  - Ancillary activity: Principal and secondary activities are generally carried out with the support of a number of "ancillary activities", such as accounting, transportation, storage, purchasing, sales promotion, repair and maintenance etc. Thus ancillary activities are those that exist solely to support the main productive activities of an entity by providing non-durable goods or services for the use of that entity.

- Elaboration period (12)

  Period in which the data collection, analysis and elaboration of the information take place. It may or not coincide with the reference or the transmission periods.

  F

- Forward-casting (12)

  A break in time series occurs in a particular point in time, for instance due to a change of classification. Forward-casting is the process that allows to re-construct the time series forward in time from this point in terms of the old classification.

G

- GBR (12)

General Business Register

- Glossary (20)

An alphabetised list of terms with definitions often created by an organization to reflect its needs. Normally lacks hierarchical arrangement or cross references. Also known as a term list.

I

- ICT(12)

Information and Communication Technology

- IMF (13)

International Monetary Fund

- Implementation measures(12)

Set of measures ensuring the smooth transition from NACE Rev. 1.1 to NACE Rev.2, for instance related to breaks in time series, including double reporting and back-casting of time series.

- Implementation plan (12)

Schedule for the implementation of the new classification (NACE Rev.2) setting up the first data reference period in which the new classification must be applied, as well as the first transmission period in which data must be sent using the new classification.

- Index for classification purposes (12)

List of descriptions (economic activities, products…) and the corresponding codes of a specific classification.

- ISIC Rev. 4 (11)

 International Standard Industrial Classification of all Economic Activities, Revision 4. (by the NU). The ISIC is intended to be a standard classification of productive economic activities. Its main purpose is to provide a set of activity categories that can be utilised when dissecting statistics according to such activities.

K

- KAU

Kind of Activity Unit

L

- LCI (12)

  Labour Cost Index

- Level of detail (12)

  Level of the classification (section, division, group, class) to which the macro-data are provided.

M

- Macrodata, statistical (6)

  Observation data gained by a purposeful aggregation of statistical microdata conforming to statistical methodology.

- Microdata, statistical (6)

  An observation data collected on an individual object - statistical unit.

- MIG (13)

  Main Industrial Groupings

- Misclassification (18)

  When a subject is falsely classified into a category in which the subject does not belong. It may result from misreporting by study subjects, from the use of less than optimal measurement devices, or from random error.

N

- NACE

  Nomenclature statistique des Activités économiques dans la Communauté Européenne

- NACE Rev.1.1

  NACE Rev. 1 ("Nomenclature statistique des Activités économiques dans la Communauté Européenne" - Statistical classification of economic activities in the European Community) was adopted in order to establish a common statistical classification of economic activities within the European Community in order to ensure comparability between national and community classifications and hence national and community statistics.

  NACE Rev. 2
  21 Sections - letters A to Q
  88 Divisions - 2-digit codes
  272 Groups - 3-digit codes
  618 Classes - 4-digit codes

- NAICS (4)

  North American Industry Classification System

  O

- Observation, pre-break (12)

  The pre-break observation is the observation, at a time series break period, that was calculated using the "old" classification, with the "observation" being calculated following the "new" classification.

  P

- PRODCOM List (13)

  The PRODCOM List is a list of products, PRODCOM Headings. The PRODCOM List is based on the Statistical classification of products by activity in the European Economic Community (the CPA) and linked with the Combined Nomenclature (the CN).

  Q

- Quality Regulations (12)

  Regulations that set a coefficient of variation value, over which the data are not acceptable. A change in the classification structure will affect the coefficient of variation value, and therefore, the quality.

  R

- RAMON (13)

  Eurostat's Classifications Server

  http://europa.eu.int/comm/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC

- Reference period (12)

  Period which the data are referred to. It may or not coincide with the data collection period.

  S

- Sample (6)

  A sample is a subset of a frame where elements are selected based on a randomised process with a known probability of selection.

- Sample design (15)

  The sample design provides information on the target and final sample sizes, strata definitions and the sample selection methodology.

- Sample size (18)

  The number of sampling units which are to be included in the sample. In the case of a multi-stage sample this number refers to the number of units at the final stage in the sampling.

- SIC (12)

  Standard Industrial Classification

- SBS (12)

  Structural Business Statistics

- Source (Data source or provider) (19)

  The source of data (often referred to as "data provider") is the organisation or individual from where data are obtained. The organisation or individual may have been the primary source or secondary source of the data.

- Stock series/data (19)

  Statistical data presented as stock series/data are measured at the end of the reference period, for example, money supply data which can refer to an observation on the last working day of the reference period.

- Stratification (8)

  Stratification consists of dividing the population into subsets (called strata) before the selection of a sample within each of these subsets. Regarding economic statistics, the stratification is usually based on the classification of economic activities.

- STS (4)

  Short-Term Statistics

- SNA (4)

  System of National Accounts

  T

- TFI (12)

  Task Force on implementation

- Time series (18)

  A set of ordered observations on a quantitative characteristic of an individual or

  collective phenomenon taken at different points of time.

- Time series breaks (19)

  Breaks in statistical time series occur when there is a change in the standards for defining and observing a variable over time. Such changes may be the result of a single change or the combination of multiple changes at any one point in time of observation of the variable.
  The specific causes of breaks in a statistical time series include changes in: classifications used, definitions of the variable, coverage; etc.

  Finally, the impact of a time series break is often a matter of judgement on the part of the user and depends on the use(s) to which the data are put.

- Transmission period (12)

  Period in which data are sent to the Institution or Organisation that requests them.

    W

- Weight (18)

  The importance of an object in relation to a set of objects to which it belongs; a numerical coefficient attached to an observation, frequently by multiplication, in order that it shall assume a desired degree of importance in a function of all the observations of the set.

- Weight (index) (12)

  Weight used in the construction of an index (for instance, LCI).

- Weight (stratification) (12)

  Weight used in the definition of the stratification during a sample design.

INFORMATION SOURCES

Regulations

1. European Union, Council Regulation (EC) No 322/97 of 17 February 1997 on Community Statistics (Official Journal of the European Communities No L 052, 22/02/1997, p. 1)

2. NACE Rev. 1.1, Statistical Classification of Economic Activities in the European Community, Introduction, Eurostat, May 1996, P. 14-15, European Union, Council Regulation (EEC) No 696/93 of 15 March 1993 on the statistical units for the observation and analysis of the production system in the Community (Official Journal of the European Communities No L 076, 30/03/1993, p. 1), Section IV B1 and B4 of 15.03.1993 on the statistical units for the observation and analysis of the production system in the Community and Eurostat

Organisations

1. OECD, IMF, ILO, Interstate Statistical Committee of the Commonwealth of Independent States, "Measuring the Non-Observed Economy: A Handbook", Annex 2, Glossary, Paris, 2002.

2. United Nations Statistics Division, "Central Product Classification (CPC), Version 1.0", Statistical Papers, Series M No. 77, Ver. 1.0, United Nations, New York, 1998, P. 118

3. Statistical Office of the United Nations, "International Standard Industrial Classification of all Economic Activities, Third Revision", Statistical Papers Series M No. 4, Rev. 3, United Nations, New York, 1990, par. 16; 2) Eurostat, "NACE Rev. 1 - Statistical Classification of Economic Activities in the European Community - Introduction", Office for Official Publications of the European Communities, Luxembourg, 1996

4. Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000.

5. Organisation for Economic Co-operation and Development (OECD), "Main Economic Indicators", monthly

6. Statistics Canada, "Statistics Canada Quality Guidelines", 3rd edition, October 1998, page 20

Recommendation manuals

1. Handbook on quarterly national accounts, 1999 Edition, Eurostat, p.82
2. Business Register Recommendation Manual, paragraph 1.11
3. Eurostat, "Methodological manual on the design and implementation of surveys on inbound tourism", Office for Official Publications of the European Communities, Luxembourg, 2000, Appendix A, Classification and Glossary
4. NACE Rev.2 Implementation handbook. Implementation TF, NACE/CPA WG, Eurostat.

Working Groups and Task Forces in Eurostat

1. Eurostat
2. Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003
3. Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, May 2002

Expert Groups

1. "United Nations Glossary of Classification Terms" prepared by the Expert Group on International Economic and Social Classifications; unpublished on paper.

2. Neuchâtel Group, "Neuchâtel Terminology: Classification database object types and their attributes - Version 2", September 2002

Other projects

1. The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

2. Statistical Data and Metadata Exchange (SDMX) http://www.sdmx.org/

   The BIS, ECB, EUROSTAT, IMF, OECD, UN, and the World Bank have joined together to focus on business practices in the field of statistical information that would allow more efficient processes for exchange and sharing of data and metadata within the current scope of our collective activities.

   The goal is to explore common e-standards and ongoing standardization activities that could allow us to gain efficiency and avoid duplication of effort in our own work and possibly for the work of others in the field of statistical information.

3. The Dublin Core Metadata Initiative is an open forum engaged in the development of interoperable online metadata standards that support a broad range of purposes and business models. DCMI's activities include consensus-driven working groups, global conferences and workshops, standards liaison, and educational efforts to promote widespread acceptance of metadata standards and practices   (www.dublincore.org)