



This glossary is part of the

# Memobust Handbook

on Methodology of Modern Business Statistics

26 March 2014

# Memobust Glossary

Edited by Rob van de Laar, Statistics Netherlands

The glossary contains a list of words and concepts with a description of their meaning. It also contains several terms outside the SDMX and GSBPM standards. The work on the glossary started at the beginning of the Memobust project (2011), and as the project progressed, new words and concepts were added. In the final stage of the project (2014) a lot of work has been done on the glossary to integrate and harmonise statistical terms and definitions originating from different modules and different topics in the Handbook. Definitions that are not part of a standard glossary and were formulated by the authors of the Memobust modules are the so-called *Memobust definitions*. For these the “ISO/IEC 11179-4 Part 4: Formulation of data definitions” standard was applied. These concepts were used by the authors in the indicated modules. Intentionally different definitions for the same term from different standards have been kept as separate definitions in the glossary, so that in future work differences in standard definitions can be removed or definitions can be combined to arrive at even better definitions. In some cases where two or more definitions for the same term exist with one definition from a standard source and another definition a Memobust definition, the author was not able to use the standard definition and provided a definition for the purpose of his module. Accidental differences in Memobust definitions have been removed by the editor of the Memobust glossary. This work of harmonising and integrating terms and definitions could be continued after the end of the Memobust project if resources exist for this task. At the moment we integrated and harmonised the Memobust glossary for 764 definitions and 695 terms. Some of these definitions are intended to be different, so homonyms have not been prevented completely. They are indicated with a light background colour. Web addresses of sources for standard definitions are provided at the end of the document.

This Memobust glossary was used during the writing of the Handbook in order to facilitate the use of harmonised vocabulary right from the start. From the beginning this glossary was based on the SDMX glossary, and contains all concepts relevant to the Memobust handbook. For internal reviews this glossary was used as it helped reviewers to check the specific vocabulary of a module. It is intended for readers of the modules in the Memobust handbook as an easy reference, but it can also be used to find quickly modulus of the Handbook with relevant information from key terms. For each term references are provided to the relevant modules. Definitions are not repeated as part of the modules, so maintenance of the glossary is limited to this ‘global’ Memobust glossary.

Term	Definition	Source of definition	Synonyms (optional)	Module
(n,k) rule	A cell is regarded as confidential, if the n largest units contribute more than k % to the cell total, e.g. n=2 and k=85 means that a cell is defined as risky if the two largest units contribute more than 85 % to the cell total. The n and k are given by the statistical authority. In some NSIs the values of n and k are confidential.	Glossary on Statistical Disclosure Control (2014)	Dominance rule	Theme: Statistical disclosure control methods for quantitative tables
(p,q) rule	It is assumed that prior to publication of tabular data the contribution of one individual to a cell total can be estimated to within q per cent (a priori relative error in estimating the individual contribution). If after publication of the statistic the value can be estimated to within p per cent (a posteriori relative error in estimating the individual contribution), the cell is declared as confidential. The parameters p and q are determined by the statistical authority. In some NSIs the values of p and q are confidential.	Glossary on Statistical Disclosure Control (2014)	Ambiguity rule; prior posterior rule	Theme: Statistical disclosure control methods for quantitative tables
$\mu$ -ARGUS	Software that creates safe micro-data files.	Argus (2013)		Theme: Logging
Acceptance region	A component of an edit rule that defines, for a given edit group, for which values of the test variable the edit is satisfied.	Norberg (2011)		Method: Manual Editing
Accepted burden	An allowable level of response burden created e.g. by increasing nonresponse rates, which has a positive effect on response burden. To avoid such undesirable "rewards" and, consequently, a less alert attitude towards declining response rates, survey managers should be confronted with burden figures which include hypothetical non response burden as well	Willeboordse <i>et al.</i> (1997)		Theme: Response Burden
Accessibility	The ease and conditions under which statistical information can be obtained.	Eurostat's Concepts and Definitions Database (2013)		(1) Theme: Quality of Statistics; (2) Theme: Overall Design
Accessibility of a log	The ease and conditions under which logs can be obtained.	Memobust definition (2014)		Theme: Logging
Accuracy	The closeness of estimates to the unknown true values.	ESS Handbook for Quality Reports (2009) (2009)		(1) Theme: Overall Design; (2) Theme: Repeated Surveys
Accuracy	Closeness between the estimated value and the true value measured by the statistic (usually unknown)	OECD (2006)		Theme: Revisions of Economic Official Statistics
Accuracy	Closeness of computations or estimates to the exact or true values that the statistics were intended to measure.	SDMX (2009)		(1) Method: Denton for Benchmarking; (2) Method: RAS; (3) Method: Stone; (4) Theme: Macro Integration.

Accuracy (of estimates)	The closeness of estimates to the true values.	ESS Handbook for Quality Reports (2009)		Theme: Quality of Statistics
Accuracy (of estimates)	Closeness of computations or estimates to the exact or true values that the statistics were intended to measure. Context: The accuracy of statistical information is the degree to which the information correctly describes the phenomena. It is usually characterized in terms of error in statistical estimates and is often decomposed into bias (systematic error) and variance (random error) components. Accuracy is associated with the "reliability" of the data, which is defined as the closeness of the initial estimated value to the subsequent estimated value.	SDMX (2009)		(1) Theme: Methods and Quality; (2) Theme: Quality and Risk Management Models
Active enterprise	Within the Business Demography context, activity is defined as any turnover and/or employment in the period from 1st January to 31st December in a given year.	Eurostat-OECD Manual on Business Demography Statistics (chapter 6)		Theme: Business Demography
Activity	An activity can be 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 and services), a production process and an output of products. Activities can be determined by reference to a specific level of NACE Rev. 2.	CODED		Theme: Derivation of Statistical Units
Activity	The combination of actions that result in a certain set of products. An activity can be said to take place when resources such as equipment, labour, manufacturing techniques or products are combined, leading to specific goods or services. Thus, an activity is characterised by an input of resources, a production process and an output of products. Context: 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.	RAMON, Eurostat's metadata server		(1) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (2) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys

Activity	An activity can be 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 and services), a production process and an output of products. Activities can be determined by reference to a specific level of NACE Rev. 2. If a unit carries out more than one activity, all the activities, which are not ancillary activities are ranked according to the gross value added. On the basis of the preponderant gross value added generated, a distinction can then be made between principal activity and secondary activities. Ancillary activities are not isolated to form distinct entities or separated from the principal or secondary activities of entities they serve.	RAMON, Eurostat's metadata server		Theme: Statistical Registers and Frames – The statistical units and the business register
Actual burden	The burden based on a realistic level of difference between signals on non-response and response. More precisely, it is a reasonably allowable level of non-response.	Hedlin <i>et al.</i> (2005)		Theme: Response Burden
Adjacency matrix	0-1 matrix that indicates which nodes in a graph (or a digraph) are connected by an edge (or an arrow).	Hacking & Willenborg (2012)		Method: Automatic coding based on semantic networks
Administrative data	The data derived from an administrative source, before any processing or validation by the NSIs.	Essnet Admin Data Glossary 1.1		(1) Theme: Collection and Use of Secondary Data; (2) Theme: Editing Administrative Data; (3) Theme: Estimation with administrative data
Administrative data holder	The organisational unit holding an administrative source	Essnet Admin Data Glossary 1.1		Theme: Collection and Use of Secondary Data
Administrative data provider	The administrative data holder who is bound to provide their data to the NSI, by law or by virtue of a specific agreement	Essnet Admin Data Glossary 1.1		Theme: Collection and Use of Secondary Data
Administrative population	The set of units that an administrative source is meant to cover, as defined by the relevant administrative regulation. This population may or may not correspond exactly to a given target	Essnet Admin Data Glossary 1.1		Theme: Collection and Use of Secondary Data
Administrative register	Administrative registers come from administrative sources and become statistical registers after passing through statistical processing in order to make it fit for statistical purposes (production of register based statistics, frame creation, etc.).	UN/ECE Glossary of Terms on Statistical Data Editing (2007)		Theme: Collection and Use of Secondary Data
Administrative regulation	A set of detailed directions having force of law, developed to put a policy into practice (such as decrees, ordinances, and other similar provisions). It is normally addressed to a designated population of natural and/or juridical persons, which are bound to observe it.	Essnet Admin Data Glossary 1.1		Theme: Collection and Use of Secondary Data
Administrative source	A data holding containing information collected and maintained for the purpose of implementing one or more administrative regulations.	Essnet Admin Data Glossary 1.1 (first part) & SDMX, 2009		Theme: Collection and Use of Secondary Data

Administrative source	A data holding containing information collected and maintained for the purpose of implementing one or more administrative regulations. In a wider sense, any data source containing information that is not primarily collected for statistical purposes.	Essnet Admin Data Glossary 1.1		Theme: Editing Administrative Data
Administrative source	A data holding containing information collected and maintained for the purpose of implementing one or more administrative regulations. Context: A wider definition of administrative sources, is used in the Eurostat Business Registers Recommendations Manual: a data holding containing information which is not primarily collected for statistical purposes. The organisational unit responsible for maintaining one or more administrative sources is known as an administrative organisation.	SDMX (2009)		(1) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (2) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (3) Theme: Statistical Registers and Frames – Statistical register and survey frame design; (4) Theme: Statistical Registers and Frames – The statistical units and the business register
Administrative units	With reference to the use of administrative data for statistical purposes, the units for which administrative data are recorded. These units may or may not be the same as those required for the statistical output (which are referred to as statistical units).	Essnet Admin Data Glossary 1.1		Theme: Editing Administrative Data
Aggregation	Aggregation in a system of time series is commonly referred in a literature as benchmarking to contemporaneous constraints.	Stuckey et.al. (2004)		(1) Theme: Issues on Seasonal Adjustment; (2) Theme: Seasonal adjustment – introduction and general description.
AIC	Measure of the relative goodness of fit of a statistical model $AIC = 2k - 2\log(lik)$ , where k is the number of parameters in the model and <i>lik</i> is maximum value assumed by the likelihood function.	Memobust definition (2014)		(1) Method: EBLUP Unit level for Small Area Estimation; (2) Method: EBLUP Area Level for Small Area Estimation (Fay-Herriot)
Allocating (sample elements to interviewers)	The allocation consists of associating each telephone number (belonging to a sample element) with an interviewer. So the allocation of interviewers to sample elements is via their telephone numbers.	Memobust definition (2014)		Theme: CATI Allocation
Ambiguity rule	See: (p,q) rule.	Glossary on Statistical Disclosure Control (2014)	(p,q) rule; prior posterior rule	Theme: Statistical disclosure control methods for quantitative tables
Annual Alignment	The constraint that annual data has to be consistent with sub annual data. Annual and sub annual are used in a broad sense here. It can be any combination of two periods with a difference frequency, such that one annual period covers a whole number of sub annual periods.	Memobust definition (2014)		Method: Denton for Benchmarking
Anticipated value	Anticipated values are used in score functions and are predictions for the values which are expected in the actual survey.	EDIMBUS Manual	Predicted values	Theme: Selective Editing

ARGUS	Two software packages for Statistical Disclosure Control are called Argus. $\mu$ -Argus is a specialized software tool for the protection of microdata. The two main techniques used for this are global recoding and local suppression. In the case of global recoding several categories of a variable are collapsed into a single one. The effect of local suppression is that one or more values in an unsafe combination are suppressed, i.e. replaced by a missing value. Both global recoding and local suppression lead to a loss of information, because either less detailed information is provided or some information is not given at all. $\tau$ -Argus is a specialized software tool for the protection of tabular data. $\tau$ -Argus is used to produce safe tables. $\tau$ -Argus uses the same two main techniques as $\mu$ -Argus: global recoding and local suppression. For $\tau$ -Argus the latter consists of suppression of cells in a table.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
ARIMA models	These are a versatile family of models for modelling and forecasting time series data. Seasonal ARIMA models have a special form for efficiently modelling many kinds of seasonal time series and are heavily used in seasonal adjustment. ARIMA is an acronym for AutoRegressive Integrated Moving Average	US Census Bureau		Method: Seasonal adjustment of economic time series
Assisted coding	Coding of textual variable performed during the interview	Memobust definition (2014)		Theme: Data Collection: Techniques and Tools
Attribute	A quality of feature, especially one that is considered to be good or useful. Examples: availability, accuracy, integrity, confidentiality, effectiveness.	Longman (2010)		(1) Theme: Methods and Quality; (2) Theme: Quality and Risk Management Models
Attribute disclosure	Attribute disclosure is attribution independent of identification. This form of disclosure is of primary concern to NSIs involved in tabular data release and arises from the presence of empty cells either in a released table or linkable set of tables after any subtraction has taken place. Minimally, the presence of an empty cell within a table means that an intruder may infer from mere knowledge that a population unit is represented in the table and that the intruder does not possess the combination of attributes within the empty cell.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables

Attribute of register unit	Attribute of a register unit is a regularly updated characteristic of a register unit. Remark: Attributes of statistical register units can be arranged in groups. Accordingly, attributes referring to identification, contact, classification, demographic characteristics, relation to other register units, attributes supporting register maintenance and statistical processes (for example organization of data collection, sampling, etc.) can be defined. In respect of maintainability and changes of attributes over time, administrative and statistical attributes are distinguished	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (2) Theme: Statistical Registers and Frames – Survey frames for business surveys
Automatic coding	Coding (in batch) using a program. The program takes all of the decisions.	Hacking & Willenborg (2012)		(1) Method: Automatic coding based on pre-coded datasets; (2) Method: Automatic coding based on semantic networks; (3) Theme: Coding
Automatic coding	The computer assigns codes to the verbal responses working in “batch” processing	Macchia S. and Murgia M (2002)	AUC	Theme: Different Coding Strategies
Automatic coding precision	The percentage of correctly coded descriptions ((Number of correctly coded description/Number coded descriptions)	Memobust definition (2014)		Method: Automatic coding based on pre-coded datasets
Automatic coding rate	The percentage of coded descriptions(Number of coded description/Number descriptions to be coded)	Memobust definition (2014)		Method: Automatic coding based on pre-coded datasets
Automatic editing	An umbrella term for editing methods in which the data are checked and adjusted by a computer.	Memobust definition (2014)		(1) Method: Automatic Editing; (2) Method: Deductive Editing; (3) Theme: Editing Administrative Data; (4) Theme: Statistical Data Editing
Autoregressive model	A representation of a type of random process; as such, it describes certain time-varying processes. The autoregressive model specifies that the output variable depends linearly on its own previous values.	Memobust definition (2014)		Method: Chow-Lin Method for Temporal Disaggregation
Autoregressive model	An econometric model-based upon the autoregressive process but also containing lagged versions of some or all of the endogenous variables considered in the model specification.	Memobust definition (2014)		Method: Preliminary estimates with model-based methods
Auxiliary variable	A variable that correlates with the target variable and is observed for all units.	CBS Methods Series Glossary		(1) Theme: Donor Imputation; (2) Theme: Imputation; (3) Theme: Imputation for Longitudinal Data; (4) Theme: Model-based Imputation; (5) Theme: Sample selection; (6) Theme: Design of Estimation – Some Practical Issues; (7) Method: Assigning random numbers when co-ordination of surveys based on different unit types is considered

Bag-of-words assumption	The assumption that, for a description, only the separate words that occur play a role, and not the order and the combinations of these words in the description.	Hacking & Willenborg (2012)		(1) Method: Automatic coding based on pre-coded datasets; (2) Theme: Coding
Barnardisation	A method of disclosure control for tables of counts that involves randomly adding or subtracting 1 from some cells in the table.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Base register	Registers kept as a basic resource for public administration. The function of base registers is typically to keep stock of the population at any given time. In addition, they have to maintain identification information to be used by other sources.	UN/ECE Glossary of Terms on Statistical Data Editing (2007)		Theme: Collection and Use of Secondary Data
Benchmarking	Achieving consistency between data that are published at different frequencies (for instance quarterly data that has to comply with annual data).	Memobust definition (2014)		Method: Denton for Benchmarking
Benchmarking	Achieving consistency between data that are published at different frequencies (for instance quarterly data that has to comply with annual data).	Memobust definition (2014)		Theme: Macro Integration
Benchmarking	Achieving consistency between data that are published at different level of aggregation.	SDMX (2009)		(1) Method: EBLUP Area Level for Small Area Estimation (Fay-Herriot); (2) Method: EBLUP Unit level for Small Area Estimation
Benchmarking	Benchmarking (to temporal constraints) involves enforcing consistency across time with respect to another time series.	Stuckey et.al. (2004)		(1) Theme: Issues on Seasonal Adjustment; (2) Theme: Seasonal adjustment – introduction and general description.
Bias	An effect which deprives a statistical result of representativeness by systematically distorting it, as distinct from a random error which may distort on any one occasion but balances out on the average.	Eurostat's Concepts and Definitions Database (2013)	Systematic error.	Theme: Quality of Statistics
Bias	The bias of an estimator is the difference between its mathematical expectation and the true value of the parameter. In case it is zero, the estimator is said to be unbiased. Expectation is usually calculated on the set of all possible samples (Randomization approach). Otherwise is calculated with respect to the assumed model (model-based approach).	Memobust definition (2014)		(1) Theme: Weighting and Estimation; (2) Theme: Estimation with administrative data; (3) Method: EBLUP Unit level for Small Area Estimation
Bias	The bias of an estimator is the difference between its mathematical expectation and the true value it estimates. If this difference is zero, the estimator is said to be unbiased. Expectation is usually calculated on the set of all possible samples.	SDMX (2009)		(1) Method: EBLUP Area Level for Small Area Estimation (Fay-Herriot); (2) Method: Preliminary estimates with design-based methods; (3) Method: Deductive Editing; (4) Theme: Statistical Data Editing

Bias	The bias of an estimator is the difference between its mathematical expectation and target parameter. In the case it is zero, the estimator is said to be unbiased. Expectation is usually calculated on the set of all possible samples.	Statistical Data and Metadata Exchange (SDMX)		Method: Generalised regression estimator
Bias (of an estimator)	An effect which deprives a statistical result of representativeness by systematically distorting it, as distinct from a random error which may distort on any one occasion but balances out on the average.	SDMX (2009)		(1) Theme: Sample co-ordination; (2) Method: Denton for Benchmarking; (3) Theme: Macro Integration
BIC	This is a criterion for model selection among a finite set of models. It is based, in part, on the likelihood function.	Memobust definition (2014)		Method: EBLUP Unit level for Small Area Estimation
BIC	Measure of the relative goodness of fit of a statistical model $BIC=k \log(n)-2\log(lik)$ where k is the number of parameters in the model, n is the number of observation and lik is the maximum value of the likelihood function.	Memobust definition (2014)		Method: EBLUP Area Level for Small Area Estimation (Fay-Herriot)
Binding constraint	See hard constraint	Memobust definition (2014)		Method: Denton for Benchmarking
Birth rate	The birth rate of a given reference period is the number of births as a percentage of the population of active enterprises.	Memobust definition (2014)		Theme: Business Demography
Blocking variable	A variable that is used to partition matching data sets, that is, divide in a number of subfiles, with the intention of reducing the search space.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching; (5) Method: Fellegi-Sunter and Jaro Approach to Record Linkage
BLUE (Best Linear Unbiased Estimator)	Estimator minimizing the square loss in the class of linear unbiased estimators (unbiasedness is referred to the model distribution).	Memobust definition (2014)		Method: Small area estimation methods for time series data
BLUP (Best Linear Unbiased Predictor)	Predictor which minimizes the square loss in the class of linear unbiased predictors (unbiasedness is referred to the model distribution).	Memobust definition (2014)		Method: Small area estimation methods for time series data
Bounds	The range of possible values of a cell in a table of frequency counts where the cell value has been perturbed or suppressed. Where only margins of tables are released it is possible to infer bounds for the unreleased joint distribution. One method for inferring the bounds across a table is known as the Shuttle algorithm.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Break of time series	Break occurring when there is a change in the standards for defining and observing a variable over time.	SDMX (2009)	Time series break	Theme: Repeated Surveys

Break-up	This event involves a splitting of the production factors of an enterprise into two or more new enterprises, in such a way that the previous enterprise is no longer recognisable. There is no continuity or survival, but the closure of the previous enterprise is not considered to be a death. Similarly the new enterprise are not considered to be births.	Eurostat-OECD Manual on Business Demography Statistics (chapter 4).		Theme: Business Demography
BSDG	Bussiness Statistics Directors Group	Eurostat website/CROS portal		Theme: The European Statistical System
Business register for statistical purposes	Regulation (EC) No 177/2008 of the European Parliament and of the Council establishes a common framework for business registers for statistical purposes in the Community. Member States shall set up one or more harmonised registers for statistical purposes, as a tool for the preparation and coordination of surveys, as a source of information for the statistical analysis of the business population and its demography, for the use of administrative data, and for the identification and construction of statistical units. The registers shall be compiled of: All enterprises carrying on economic activities contributing to the gross domestic product (GDP), and their local units; The legal units of which those enterprises consist; Truncated enterprise groups and multinational enterprise groups; and All-resident enterprise groups.	Business Register Regulation (EC) No 177/2008, Articles 1 and 3 (1)	Statistical business register	Theme: Statistical Registers and Frames – Main module; Theme: Statistical Registers and Frames – Quality of statistical registers and frames
CAI	Computer Assisted Interviewing. The use of computer during interviewing.	Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva (2000)		(1) Theme: Electronic Questionnaire Design; (2) Theme: Editing During Data Collection; (3) Theme: Testing the Questionnaire; (4) Theme: Questionnaire Design; (5) Theme: Data Collection: Techniques and Tools; (6) Theme: CATI Allocation
cAIC	As model selection measure, cAIC is well -suited for small area estimation. It is relevant to inferences regarding the clusters, or areas, in the context of linear mixed models. inferences regarding the clusters, or areas, in the context of linear mixed models. The criterion is based on the conditional likelihood for fixed and random effects vectors evaluated at their estimated values, and $y$ is the data. The effective number of degrees of freedom is essentially given by the trace of the hat matrix $H$	Memobust definition (2014)		Method: EBLUP Unit level for Small Area Estimation

cAIC	As model selection measure, cAIC is well-suited for small area estimation. It is relevant for inferences regarding clusters, or domains, in the context of linear mixed models. The criterion is $cAIC = 2peff - 2\log(lik)$ , where $lik$ is maximum values assumed by the conditional likelihood, that is the likelihood function when fixed and random effects vectors evaluated at their estimated values. The effective number of degrees of freedom $peff$ is essentially given by the trace of the hat matrix $H$ .	Memobust definition (2014)		Method: EBLUP Area Level for Small Area Estimation (Fay-Herriot)
CAII	Computer Assisted Internet Interview	Willeboordse <i>et al.</i> (1997)		Theme: Response Burden
Calculated interval	The interval containing possible values for a suppressed cell in a table, given the table structure and the values published.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Calendar adjustment	Calendar adjustment refers to the correction for calendar variations. Such calendar adjustments include working day adjustments or the incidence of moving holidays (such as Easter and Chinese New Year)	OECD (2006)		Method: Seasonal adjustment of economic time series
Calendar effects	Influences deriving from differences in the number of working days or the dates of particular days which can be statistically proven and quantified	Eurostat (2009)		(1) Method: Seasonal adjustment of economic time series; (2) Theme: Seasonal adjustment – introduction and general description
Calibration	One of the most important methods of weighting commonly used by many statistical agencies in survey sampling, whose main aim is to compute weights to be used in estimation, given an input of auxiliary information.	Memobust definition (2014)		Method: Calibration
Calibration equation	In the calibration procedure for totals, equations in which calibration weights applied to all auxiliary variables in the sample exactly reproduce the known population totals of the auxiliary variables.	Memobust definition (2014)		Method: Calibration
Calibration estimator	An estimator which is a weighted sum of sample observation, whose weights are obtained in order to minimize a distance with the design weights subject to the constraint that the weighted sum of an auxiliary variables reproduce the known amount. See Module XIX 2.c for further details	Memobust definition (2014)		Method: Generalised regression estimator
Calibration estimator	Estimator which takes into account calibration weights which satisfy calibration equations.	Memobust definition (2014)		Method: Calibration
Calibration weights	Weights which replace the original initial design weights and satisfy calibration equations.	Memobust definition (2014)		Method: Calibration
Call scheduler	Software that runs in the scheduling system according to the values of the scheduling parameters set by survey responsible	Memobust definition (2014)		Theme: Data Collection: Techniques and Tools

Capacity of call room	The maximum number of interviewers that can work simultaneously in the call room for CATI survey work.	Memobust definition (2014)		Theme: CATI Allocation
CAPI	Computer Assisted Personal Interviewing. A method of data collection in which an interviewer uses a computer to display questions and accept responses during a face-to-face interview.	United States Bureau of Census, Glossary of Selected Abbreviations and Acronyms.		(1) Theme: Data Collection; (2) Theme: CATI Allocation; (3) Theme: Data Collection: Techniques and Tools; (4) Theme: Response Burden; (5) Method: Computer-assisted coding; (6) Theme: Questionnaire Design; (7) Theme: Mixed Mode Data Collection – design issues; (8) Theme: Electronic Questionnaire Design; (9) Theme: Editing During Data Collection; (10) Theme: Coding; (11) Theme: Quality of Statistics
CASI	Computer Assisted Self-Interviewing. The technique whereby respondents independently complete electronic questionnaires, assisted only by specially-designed computer programs.	Glossary, Adapting new technologies to census operations (2001)		(1) Theme: Electronic Questionnaire Design; (2) Theme: Editing During Data Collection; (3) Theme: Testing the Questionnaire; (4) Theme: Questionnaire Design
CASI	Computer Assisted Self Interviewing is a method of data collection in which the respondent operates the computer: questions are read from the computer screen and responses are entered directly in the computer. A well-known form of CASI is the web survey.	Memobust definition (2014)		Theme: Mixed Mode Data Collection – design issues
CATI	Computer Assisted Telephone Interviewing. A method of data collection by telephone with questions displayed on a computer and responses entered directly into a computer.	United States Bureau of Census, Glossary of Selected Abbreviations and Acronyms.		(1) Theme: Questionnaire Design; (2) Theme: Mixed Mode Data Collection – design issues; (3) Theme: Data Collection; (4) Theme: CATI Allocation; (5) Theme: Data Collection: Techniques and Tools; (6) Theme: Response Burden; (7) Method: Computer-assisted coding; (8) Theme: Electronic Questionnaire Design; (9) Theme: Editing During Data Collection; (10) Theme: Testing the Questionnaire; (11) Theme: Coding; (12) Theme: Quality of Statistics
CATI Interviewer	A person who on behalf of a statistical office carries out interviews by telephone. In this module we assume that these people work from a call room.	Memobust definition (2014)		Theme: CATI Allocation

CAWI	Computer Assisted Web Interviewing. A method of data collection based on web questionnaire. The respondent accesses the questionnaire via a web connection and fills it in.	Memobust definition (2014)	Web Survey	(1) Theme: Coding; (2) Theme: Quality of Statistics; (3) Theme: Data Collection: Techniques and Tools; (4) Method: Computer-assisted coding; (5) Theme: Mixed Mode Data Collection – design issues
CBA	Cost Benefit Analysis – a model enabling identification of which cost and benefits to include to evaluate effects of participating in the survey, discounting future benefits and costs over time to obtain a present day value and identification of relevant constraints.	Haraldsen <i>et al.</i> (2013), Pres and Turvey (1965)		Theme: Response Burden
Cell suppression	In tabular data the cell suppression SDC method consists of primary and complementary (secondary) suppression. Primary suppression can be characterised as withholding the values of all risky cells from publication, which means that their value is not shown in the table but replaced by a symbol such as 'x' to indicate the suppression. According to the definition of risky cells, in frequency count tables all cells containing small counts and in tables of magnitudes all cells containing small counts or presenting a case of dominance have to be primary suppressed. To reach the desired protection for risky cells, it is necessary to suppress additional non- risky cells, which is called complementary (secondary) suppression. The pattern of complementary suppressed cells has to be carefully chosen to provide the desired level of ambiguity for the risky cells with the least amount of suppressed information.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Chain-linking	Joining together two indices that overlap in one period by rescaling one of them to make its value equal to that of the other in the same period, thus combining them into single time series. More complex methods may be used to link together indices that overlap by more than period	OECD (2006)		Theme: Issues on Seasonal Adjustment
Changes in inventories	Changes in inventories are measured by the value of the entries into inventories less the value of withdrawals and the value of any recurrent losses of goods held in inventories.	ESA (2010)		Theme: Manual Integration
Characteristic	See: Attribute.	Memobust definition (2014)		Theme: Methods and Quality
Checking rule	see Edit	Memobust definition (2014)	Edit	Theme: Statistical Data Editing
Clarity	The ease with which users can understand the statistics.	ESS Handbook for Quality Reports (2009)		Theme: Overall Design

Clarity	The extent to which easily comprehensible metadata are available (for the user), where these metadata are necessary to give a full understanding of statistical data.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Clarity of log information	The degree to which the log information can be read, understood and interpreted.	Memobust definition (2014)	Readability, interpretability	Theme: Logging
Classification scheme	A hierarchical arrangement of kinds of things (classes) or groups of kinds of thing	Wikipedia, English edition		(1) Method: Manual coding; (2) Theme: Coding
Cluster sampling	A sampling technique used when 'natural' groupings are evident in a statistical population	Wikipedia Cluster Sampling		Theme: Sample selection
Coder	A specialist trained to interpret and classify descriptions (in a certain area) in the light of a classification used for that purpose.	Hacking & Willenborg (2012)		(1) Method: Manual coding; (2) Method: Automatic coding based on pre-coded datasets; (3) Method: Computer-assisted coding; (3) Theme: Coding
Coding	The activity in the statistical process in which it is determined whether a code from a classification can be assigned to a description, and, if so, which code this could be.	Hacking & Willenborg (2012)		(1) Method: Manual coding; (2) Method: Automatic coding based on pre-coded datasets; (3) Method: Computer-assisted coding; (4) Method: Automatic coding based on semantic networks; (5) Theme: Coding
Coding	The process of converting verbal or textual information into codes representing classes within a classification scheme, to facilitate data processing, storage or dissemination	Memobust definition (2014)		Theme: Data Collection: Techniques and Tools
Coding	The process of converting verbal or textual information into codes representing classes within a classification scheme, to facilitate data processing, storage or dissemination.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Coding error	The assignment of an incorrect code to a data item.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Coding precision	The percentage of <u>correctly</u> coded descriptions ((Number of correctly coded description/Number coded descriptions)	Memobust definition (2014)	Automatic coding precision	Theme: Coding
Coding rate	Percentage of coded texts on the total of texts to be coded	D'Orazio M. and Macchia S (ROS) (2002)	Efficacy, Automatic coding rate	(1) Theme: Coding; (2) Theme: Measuring Coding Quality
Coefficient of variation	The ratio of the square root of the variance of the estimator to its expected value.	ESS Handbook on Precision Requirements and Variance Estimation for Household Surveys		(1) Method: Generalised regression estimator; (2) Theme: Quality of Statistics
Coherence	The degree to which the statistical processes by which statistics were generated used the same concepts – classifications, definitions and target populations – and harmonised methods.	ESS Handbook for Quality Reports (2009)		Theme: Quality of Statistics

Coherence	Adequacy of statistics to be reliably combined in different ways and for various uses.	ESS Handbook for Quality Reports (2009) (2009)		(1) Theme: Overall Design; (2) Theme: Repeated Surveys
Coherence	Adequacy of statistics to be combined in different ways and for various uses.	SDMX (2009)		Theme: Weighting and Estimation
Coherence	Adequacy of statistics to be combined in different ways and for various uses.	SDMX (2009)		(1) Method: Denton for Benchmarking; (2) Method: RAS; (3) Method: Stone; (4) Theme: Macro Integration.
Cold deck imputation	A form of donor imputation in which the donor record comes from a different data set than the recipient record.	Memobust definition (2014)		Theme: Donor Imputation
Collection unit	Collection unit is the unit from which data are obtained and by which questionnaire survey forms are completed. Data supplier and data provider are collection units.	United Nations, DEPARTMENT of Economic and Social Affairs, Statistics Division [2007]: Statistical Units. United Nations, New York		(1) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (2) Theme: Statistical Registers and Frames – Survey frames for business surveys; (3) Theme: Statistical Registers and Frames – Statistical register and survey frame design.
Commodity	Goods and services produced and used in an economy.	Memobust definition (2014)		Theme: Manual Integration
Communication mode	A channel used in a survey to contact businesses, seek survey cooperation, communicate information, instructions, procedures and non-response follow-up, and to support businesses.	Snijkers & Jones (2013)		Theme: Mixed Mode Data Collection – design issues
Communication strategy	How businesses are contacted and followed-up in case of non-response, aimed at receiving timely, accurate and complete responses.	Memobust definition (2014)		Theme: Mixed Mode Data Collection – design issues
Comparability	The degree to which the same data items can be compared but for different reference periods or different sub populations (regions or domains).	ESS Handbook for Quality Reports (2009)		Theme: Quality of Statistics
Comparability	Adequacy of statistics to be reliably compared; measurement of the impact of differences in applied statistical concepts, measurement tools and procedures where statistics are compared.	ESS Handbook for Quality Reports, 2009; modified and expanded.		(1) Theme: Overall Design; (2) Theme: Repeated Surveys.
Comparison functions	Functions that compute the distance between records compared on the chosen matching variables.	Memobust definition (2014)		Theme: Probabilistic Record Linkage
Complementary suppression	See: Secondary suppression	Glossary on Statistical Disclosure Control (2014)	Secondary suppression	Theme: Statistical disclosure control methods for quantitative tables
Completeness of log information	The degree to which log information meets all current and potential needs of the user of the log information.	Memobust definition (2014)		Theme: Logging

Composite estimator	A weighted sum of two component estimators defined to reduce the mean-squared-error (MSE) of the resulting estimator.	Memobust definition (2014)		(1) Method: Preliminary estimates with design-based methods; (2) Method: Composite Estimators for Small Area Estimation
Composite unit	A unit that is composed of units from a lower order. A household is an example of a composite unit; 'persons' are the simple units from which 'households' are composed.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Computer assisted coding (CAC)	The operator assigns codes working interactively with the computer, that gives him a support in "navigating" inside the dictionary to search for codes to be assigned to the input descriptions.	Macchia S. and Murgia M (2002)	Interactive coding	Theme: Different Coding Strategies
Computer assisted survey information collection	Computer assisted survey information collection (CASIC) encompasses computer assisted data collection and data capture. CASIC may be more broadly defined to include the use of computer assisted, automated, or advanced computing methods for data editing and imputation, data analysis and tabulation, data dissemination, or other steps in the survey or census process.	UN Statistical Commission, UNECE, 2000. Glossary of Terms on Statistical Data Editing.		Theme: Testing the Questionnaire
Computer supported coding	See Computer-assisted coding	Memobust definition (2014)		Method: Computer-assisted coding
Computer-assisted coding	A form of coding in which a coder makes all the coding decisions, possibly while using an electronic file or index.	Hacking & Willenborg (2012)	Interactive coding	(1) Method: Computer-assisted coding; (2) Theme: Coding
Concentration rule	Rule to assess whether a cell is a risky cell, based on comparing the size of the individual contributions to the cell. Examples are the dominance rule and the p% rule.	Hundepool et al. (2012)	Dominance rule; P% rule	Theme: Statistical disclosure control methods for quantitative tables
Conditional mean matching	Model based imputation method: imputes the missing value with its expectation given the observed variables	Memobust definition (2014)		Method: Statistical Matching Methods
Confidentiality of log information	The degree to which log information cannot be made available to users of log information.	Memobust definition (2014)		Theme: Logging
Connected component	A maximal connected subgraph of a graph.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Consistency	Sum of sub-annual values of a time series is equal to the annual values; in case of aggregation, the total values are equal to the aggregated values	Dagum and Cholette (2006)		Theme: Issues on Seasonal Adjustment

Consistency	Logical and numerical coherence.	SDMX (2009)		(1) Method: Chow-Lin Method for Temporal Disaggregation; (2) Method: Denton for Benchmarking; (3) Method: RAS; (4) Method: Stone; (5) Theme: Macro Integration; (6) Theme: Data fusion at micro level; (7) Theme: Quality of Statistics; (8) Theme: Weighting and Estimation
Consistency	Data values are said to be consistent if they conform to specified edit rules.	SDMX (2009)		(1) Method: Generalised Ratio Adjustments; (2) Method: Minimum Adjustment Methods; (3) Method: Prorating; (4) Method: Reconciling Conflicting Micro-Data
Consistency	An estimator is called consistent if it converges in probability to its estimand as sample increases	The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press (2003).		Theme: Small area estimation
Constrained distance hot deck	The donor can be chosen just once and the subset of the donors is selected in order to minimize the overall matching distance.	Memobust definition (2014)		Method: Statistical Matching Methods
Constraint	Specification of what may be contained in a data or metadata set in terms of content or, for data only, in terms of the set of key combinations to which specific attributes (defined by the data structure) may be attached.	SDMX (2009)		(1) Method: Chow-Lin Method for Temporal Disaggregation; (2) Method: Denton for Benchmarking; (3) Method: RAS; (4) Method: Stone; (5) Theme: Macro Integration
Consumption of government	Final consumption expenditure consists of expenditure incurred by resident institutional units on goods or services that are used for the direct satisfaction of the collective needs of members of the community.	ESA (2010)		Theme: Manual Integration
Consumption of households	Final consumption expenditure consists of expenditure incurred by resident institutional units on goods or services that are used for the direct satisfaction of individual needs or wants.	ESA (2010)		Theme: Manual Integration
Contact strategy	when and how respondents are contacted, and what material (questionnaire, cover letter, instructions et cetera) is used in each contact.	Memobust definition (2014)		(1) Theme: Data Collection; (2) Theme: Design of data collection (part 2) – Contact strategies
Contempeous constraints	Constraints within one period, between different time-series	Memobust definition (2014)		Method: Denton for Benchmarking

Control	Control is the ability to determine general corporate policy by choosing appropriate directors. Control is when owning more than half of the voting shares or otherwise controlling half of the shareholders' voting power (e.g. by controlling the shareholder or by a contract of control). This type of control can be registered as it has a legal basis. Control can be direct but can also be indirect .	European System of Accounts (ESA 1995), paragraph 2.26		Theme: Derivation of Statistical Units
Controlled rounding	Controlled rounding: To solve the additivity problem, a procedure called controlled rounding was developed. It is a form of random rounding, but it is constrained to have the sum of the published entries in each row and column equal to the appropriate published marginal totals. Linear programming methods are used to identify a controlled rounding pattern for a table.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Controlled Tabular Adjustment	A method to protect tabular data based on the selective adjustment of cell values. Sensitive cell values are replaced by either of their closest safe values and small adjustments are made to other cells to restore the table additivity. Controlled tabular adjustment has been developed as an alternative to cell suppression.	Glossary on Statistical Disclosure Control (2014)	CTA	Theme: Statistical disclosure control methods for quantitative tables
Conventional Rounding	A disclosure control method for tables of counts. When using conventional rounding, each count is rounded to the nearest multiple of a fixed base. For example, using a base of 5, counts ending in 1 or 2 are rounded down and replaced by counts ending in 0 and counts ending in 3 or 4 are rounded up and replaced by counts ending in 5. Counts ending between 6 and 9 are treated similarly. Counts with a last digit of 0 or 5 are kept unchanged. When rounding to base 10, a count ending in 5 may always be rounded up, or it may be rounded up or down based on a rounding convention.	Glossary on Statistical Disclosure Control (2014)	Deterministic rounding	Theme: Statistical disclosure control methods for quantitative tables
Co-ordination of samples	Increasing the sample overlap for some surveys rather than drawing the samples independently is known as positive co-ordination. Reducing the overlap between samples for different surveys is known as negative co-ordination.	SDMX (2009)		(1) Method: Assigning random numbers when co-ordination of surveys based on different unit types is considered; (2) Method: Sample co-ordination using simple random sampling with permanent random numbers; (3) Theme: Sample co-ordination; (4) Theme: Design of Estimation – Some Practical Issues; (5) Theme: Sample selection

CoP	The European Code of Practice provides 15 principles covering the institutional environment, the statistical production processes and the output of statistics. A set of indicators of good practice for each of the principles provides a reference for reviewing the implementation of the Code.	European Code of Practice (2011)		(1) Theme: Specification of User Needs for Business Statistics; (2) Theme: Dissemination of Business Statistics; (3) Theme: Methods and Quality
Corpus	Coded set of descriptions.	Hacking & Willenborg (2012)		(1) Method: Automatic coding based on pre-coded datasets; (2) Method: Computer-assisted coding; (3) Theme: Coding
Correction rule	An if-then rule that is used to treat a particular error in a deterministic manner.	loosely based on UN/ECE Glossary of Terms on Statistical Data Editing		Method: Deductive Editing
Correctness of log information	The degree to which log information reflects reality.	Memobust definition (2014)		Theme: Logging
Covariance matrix	A mathematic measure of reliability.	Memobust definition (2014)		Method: Stone
Coverage	The definition of the population that statistics aim to cover.	SDMX (2009)		Theme: Sample selection
Coverage error	Error caused by a failure to cover adequately all components of the population being studied, which results in differences between the target population and the sampling frame.	Eurostat's Concepts and Definitions Database, SDMX Metadata Common Vocabulary ( <a href="http://sdmx.org">http://sdmx.org</a> ), 2009		(1) Method: Denton for Benchmarking; (2) Theme: Quality of Statistics
Creative editing	A process whereby manual editors invent editing procedures to avoid reviewing another error message from subsequent machine editing.	UN/ECE Glossary of Terms on Statistical Data Editing (2007)		Method: Manual Editing
Cross validation	CV methods allow to test the robustness of the models, quantifying their predictive power by leaving out one or more observations when fitting the models, and subsequently assessing the model predictions for the left-out observation(s). It can be quantified in alternative ways, for instance averaging the prediction errors	Memobust definition (2014)		(1) Method: EBLUP Unit level for Small Area Estimation; (2) Method: EBLUP Area Level for Small Area Estimation (Fay-Herriot)
Cross-section	This involves some observations of all population, or a representative subset at one specific point in time.	Memobust definition (2014)		Method: Little and Su Method
CTA	See: Controlled Tabular Adjustment	Memobust definition (2014)	Controlled Tabular Adjustment	Theme: Statistical disclosure control methods for quantitative tables
Cut-off sampling	A sampling procedure in which a predetermined threshold is established with all units in the universe at or above the threshold being included in the sample and all units below the threshold being excluded. The threshold is usually specified in terms of the size of some known relevant variable. In the case of establishments, size is usually defined in terms of employment or output.	Memobust definition (2014)		Method: Subsampling for Preliminary Estimates

Cut-off survey	A survey in which all the entities falling above or below a threshold determined according to one or more characteristics of those entities are either included or excluded	SDMX (2009)		Theme: Sample selection
Cut-off threshold	A threshold used, mainly for cost or burden reasons, to exclude from the target population (hence from the frame) units contributing very little to the requested statistics, small businesses for instance.	SDMX (2009)		Theme: Sample selection
Cut-off value	A value to limit the matching weights (upwards or downwards).	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Damerau-Levenshtein distance	A metric defined to measure the distance between strings. It measures the minimum number of elementary steps to transform one string into another.	Memobust definition (2014)	Levenshtein distance	(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Data	Characteristics or information, usually numerical, that are collected through observation.	SDMX (2009)		(1) Method: Chow-Lin Method for Temporal Disaggregation; (2) Method: Denton for Benchmarking; (3) Method: RAS; (4) Method: Stone; (5) Theme: Macro Integration
Data cleaning	see Editing	Memobust definition (2014)	Editing	Theme: Statistical Data Editing
Data collection mode	The technical set-up for presenting and answering survey questions to respondents, and the collection of the survey data to the central administration.	Memobust definition (2014)		Theme: Mixed Mode Data Collection – design issues
Data Integration	The process of combining <u>data</u> from two or more sources to produce statistical outputs.	SDMX (2009)		(1) Method: Denton for Benchmarking; (2) Method: RAS; (3) Method: Stone; (4) Theme: Macro Integration.
Data provider	The unit that actually reports the data about the reporting unit in the name of the data supplier. This could be a representative, e.g. an accounting firm.	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys. (2) Theme: Statistical Registers and Frames – Survey frames for business surveys; (3) Theme: Collection and Use of Secondary Data

Data reconciliation	The process of adjusting <u>data</u> derived from two different sources to remove, or at least reduce, the impact of differences identified.	SDMX (2009)		(1) Method: Denton for Benchmarking; (2) Method: RAS; (3) Method: Stone; (4) Theme: Macro Integration; (5) Method: Reconciling Conflicting Micro-Data; (6) Method: Generalised Ratio Adjustments; (7) Method: Minimum Adjustment Methods; (8) Method: Prorating; (9) Theme: Data fusion at micro level
Data set	Any organised collection of <u>data</u>	SDMX (2009)		(1) Method: Denton for Benchmarking; (2) RAS; (3) Stone; (4) Macro Integration; (5) Method: Chow-Lin Method for Temporal Disaggregation
Data supplier	The unit which is formally responsible to provide data about its reporting unit(s). The survey organization has a legal relationship with the data supplier.	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (2) Theme: Statistical Registers and Frames – Survey frames for business surveys; (3) Theme: Statistical Registers and Frames – Statistical register and survey frame design
Data validation	see Editing	Memobust definition (2014)	Editing	Theme: Statistical Data Editing
DBMS	Database Management System.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Death rate	The death rate of a given reference period is the number of deaths as a percentage of the population of active enterprises	Memobust definition (2014)		Theme: Business Demography
Deductive editing	An umbrella term for editing methods that use logical reasoning to derive adjustments from the unedited data.	Memobust definition (2014)		(1) Method: Automatic Editing; (2) Method: Deductive Editing; (3) Theme: Statistical Data Editing
Deductive imputation	An umbrella term for imputation methods that use logical reasoning to derive imputed values in a deterministic manner.	CBS Methods Series Glossary	Logical imputation	(1) Method: Deductive Imputation; (2) Theme: Imputation; (3) Theme: Imputation for Longitudinal Data; (4) Theme: Imputation under Edit Constraints
Deduplication	Taking the duplicate records out of a file, one by one, that occur multiple times, and that all relate to the same unit (in a certain period).	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching

Definition	Step 1 in the OQRM model, where the object and the focus area is defined.	Van Nederpelt (2012)		Theme: Quality and Risk Management Models
Definitive interruption rate	The proportion of observation units for which the reporting unit has been successfully contacted, but has interrupted in cooperation before the very end of the questionnaire	Memobust definition (2014)		Theme: Data Collection: Techniques and Tools
Degree	The degree of a point in a graph is the number of edges in the graph connected to this point.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Delphi method	A research tool in which opinions are solicited from many experts about a topic for which there is no consensus. The answers of other experts are fed back anonymously in several rounds until consensus is reached. The method is named after the Oracle of Delphi.	Daas and Arends-Toth (2012)		Theme: Collection and Use of Secondary Data
Dependencies	Step 10 in the OQRM model, where dependencies of a focus area with other focus areas are determined. Example: The soundness of methodology contributes to the accuracy of estimates.	Van Nederpelt (2012)		Theme: Quality and Risk Management Models
Design burden	The burden which includes all aspects of the survey environment that are not directly associated with the respondent e.g. method of data collection, mode of collection and the contents of the survey, errors in sampling frame, incorrect sampling, etc.	Hedlin <i>et al.</i> (2005)		Theme: Response Burden
Design weight	For a sampling unit, it is the inverse probability of selection.	ESS Handbook on Precision Requirements and Variance Estimation for Household Surveys, EUROSTAT 2013		(1) Theme: Weighting and Estimation; (2) Method: Generalised regression estimator; (3) Method: Preliminary estimates with design-based methods
Design weight	Weight which is the inverse of the inclusion probability.	Memobust definition (2014)		Method: Calibration
Design-consistency	Convergence in probability as the sample size increased.	Memobust definition (2014)		Theme: Weighting and Estimation
Deterministic imputation	A deterministic imputation method determines one unique value for the imputation of a missing or inconsistent data item. This means that when the imputation process is repeated, the same values will be imputed.	EDIMBUS Manual		(1) Theme: Donor Imputation; (2) Theme: Imputation; (3) Theme: Model-based Imputation
Deterministic record linkage	Linkage method that detects links if and only if there is a full agreement of unique identifiers or a set of common identifiers, the matching variables.	Memobust definition (2014)	Object identifiers	Theme: Probabilistic Record Linkage
Deterministic rounding	See: conventional rounding	Glossary on Statistical Disclosure Control (2014)	Conventional rounding	Theme: Statistical disclosure control methods for quantitative tables
DIME	Directors of Methodology	Eurostat website/CROS portal		Theme: The European Statistical System

DIMESA	Direcors' Meetings of Environmental Statistics and Accounts	Eurostat website/CROS portal		Theme: The European Statistical System
Direct estimator	An estimator of the target parameter for a given sub-population (domain) is said to be a direct estimator when it is based only on sample information from the sub-population itself. The more common direct estimators used in large scale business surveys are Calibration estimators.	Memobust definition (2014)		Theme: Weighting and Estimation
Direct estimator	An estimator which takes into account only domain-specific data. In many cases this estimator gives unacceptable results due to the fact that small areas are not represented in the sample by many units.	Memobust definition (2014)		(1) Method: Composite Estimators for Small Area Estimation; (2) Method: Synthetic Estimators for Small Area Estimation
Disaggregation	The breakdown of observations, usually within a common branch of a hierarchy, to a more detailed level to that at which detailed observations are taken.	SDMX (2009)		(1) Method: Denton for Benchmarking; (2) Theme: Macro Integration
Disclosive cells	See: risky cells.	Glossary on Statistical Disclosure Control (2014)	Risky cells	Theme: Statistical disclosure control methods for quantitative tables
Disclosure	Disclosure relates to the inappropriate attribution of information to a data subject, whether an individual or an organisation. Disclosure has two components: identification and attribution.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical Disclosure Control
Disclosure risk	A disclosure risk occurs if an unacceptably narrow estimation of a respondent's confidential information is possible or if exact disclosure is possible with a high level of confidence.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical Disclosure Control
Dissemination	Supply of data in any form whatever: publications, access to databases, microfiches, telephone communications, etc.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical Disclosure Control
Dissemination	Dissemination is the release to users of information obtained through a statistical activity.	OECD Glossary of Statistical Terms		Theme: Dissemination of Business Statistics
Dissimilarity measure	A measure to express the differences between two objects or entities. Somewhat similar to a metric.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Distance function	See Metric	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Distance function	In the calibration procedure, a function which measures the distance between initial design weights and calibration weights.	Memobust definition (2014)		Method: Calibration
DMES	Directors of Macro-Economic Statistics	Eurostat website/CROS portal		Theme: The European Statistical System

Dominance rule	See: (n,k) rule.	Memobust definition (2014)	(n,k) rule; concentration rule	Theme: Statistical disclosure control methods for quantitative tables
Donor file	File where one variable (say Z) has been observed and that will be used for imputation purposes on a file where Z is missing (recipient file)	Memobust definition (2014)		Theme: Statistical Matching
Donor imputation	An imputation method for which the imputed value is copied from a donor record that closely matches the recipient record on many features.	CBS Methods Series Glossary		(1) Theme: Donor Imputation; (2) Theme: Imputation; (3) Theme: Imputation for Longitudinal Data; (4) Theme: Imputation under Edit Constraints
Doubt category	A category that can be used if a description cannot be classified with sufficient certainty. The same or other coders can review the descriptions designated as such at a later stage in the process.	Hacking & Willenborg (2012)		(1) Method: Automatic coding based on pre-coded datasets; (2) Method: Automatic coding based on semantic networks; (3) Theme: Coding
DPoD	Day - Part of Day combination. The basic time unit for allocating CATI interviewers. For the sake of concreteness we have assumed three DPoD's in this module: morning, afternoon, evening. Other choices are possible and allowed, however.	Memobust definition (2014)		Theme: CATI Allocation
DSS	Directors of Social Statistics	Eurostat website/CROS portal		Theme: The European Statistical System
EBLUP	Empirical Best Linear Unbiased Predictor – estimator obtained by plugging in the estimation of variance components in a BLUP estimator, that is the estimator that in the class of all linear unbiased estimator minimize square loss. Unbiasedness is referred to model distribution.	Memobust definition (2014)		(1) Method: EBLUP Unit level for Small Area Estimation; (2) Method: EBLUP Area Level for Small Area Estimation (Fay-Herriot)
EBLUP (Empirical Best Linear Unbiased Predictor)	Predictor obtained by plugging in the estimates of the variance components in the BLUP.	Memobust definition (2014)		Method: Small area estimation methods for time series data
ECSC	European Coal and Steel Community	ECSC		Theme: The European Statistical System
EDI	Electronic Data Interchange	Willeboordse <i>et al.</i> (1997)		Theme: Response Burden
Edit	A logical condition or a restriction to the value of a data item or a data group which must be met if the data is to be considered correct.	EDIMBUS Manual		(1) Method: Generalised Ratio Adjustments; (2) Method: Minimum Adjustment Methods; (3) Method: Prorating; (4) Method: Reconciling Conflicting Micro-Data; (5) Theme: Data fusion at micro level; (6) Theme: Imputation for Longitudinal Data; (7) Theme: Imputation under Edit Constraints; (8) Theme: Editing for Longitudinal Data

Edit	A check (logical condition or a restriction to the value of a data item or a group of data items) that identifies missing, invalid or inconsistent values or that points to data records that are potentially in error.	EDIMBUS Manual	Edit rule, Checking rule	(1) Method: Automatic Editing; (2) Theme: Statistical Data Editing
Edit	A logical condition or a restriction to the value of a data item or a data group which must be met.	EDIMBUS Manual		Theme: Selective Editing
Edit constraints	see Edit	Memobust definition (2014)		
Edit distance	Distance that returns the minimum cost in terms of insertion, deletions and substitutions needed to transform a string of one record into the corresponding string of the compared record	Memobust definition (2014)		Theme: Probabilistic Record Linkage
Edit group	A component of an edit rule that identifies a (homogeneous) subset of the units for which the acceptance region is applicable for the test variable.	Norberg (2011)		Method: Manual Editing
Edit rule	see Edit	Memobust definition (2014)	Edit	
Editing	The application of checks that identify missing, invalid or inconsistent entries or that point to data records that are potentially in error.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Editing	An activity that aims to detect, understand, and correct missing values and erroneous values in data.	Memobust definition (2014)	Data cleaning, Data validation	Theme: Statistical Data Editing
Editor	A person who performs interactive or manual editing.	Memobust definition (2014)		Method: Manual Editing
EEA	European Economic Area	EEA		Theme: The European Statistical System
EFQM	European Foundation for Quality Management	Memobust definition (2014)		Theme: Quality and Risk Management Models
EFTA	European Free Trade Association.	EFTA		Theme: The European Statistical System
Employer enterprise birth	Birth of an enterprise with at least one employee. This population consists of enterprise births that have at least one employee in the birth year and of enterprises that existed before the year in consideration, but were below the threshold of one employee	Eurostat-OECD Manual on Business Demography Statistics (chapter 5).		Theme: Business Demography
Employer enterprise death	An Employee Enterprise death occurs either as an enterprise death with at least one employee in the year of death or as an exit by decline, moving below the threshold of one employee.	Eurostat-OECD Manual on Business Demography Statistics (chapter 7).		Theme: Business Demography

Enterprise	The enterprise is the smallest combination of legal units that is an organizational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources. An enterprise carries out one or more activities at one or more locations. An enterprise may be a sole legal unit. Note: The definition does not limit enterprise to one country. However, by convention this is generally done in the European statistical context. Enterprise may thus be used elsewhere in the meaning of enterprise group, in America also in the meaning of truncated enterprise group	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, Annex Section III A		(1) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (2) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (3) Theme: Statistical Registers and Frames – Survey frames for business surveys; (4) Theme: Statistical Registers and Frames – Statistical register and survey frame design; (5) Theme: Statistical Registers and Frames – The statistical units and the business register.
Enterprise Birth	A birth amounts to the creation of a combination of production factors with the restriction that no other enterprises are involved in the event. Births do not include entries into the population due to mergers, break-ups, split-off or restructuring of a set of enterprises. It does not include entries into a sub-population resulting only from a change of activity.	Definition of SBS Regulation variables, Eurostat-OECD Manual on Business Demography Statistics (chapter 5).	Real Birth Enterprise	Theme: Business Demography
Enterprise Death	A death amounts to the dissolution of a combination of production factors with the restriction that no other enterprises are involved in the event. Deaths do not include exits from the population due to mergers, take-overs, break-ups or restructuring of a set of enterprises. It does not include exits from a sub-population resulting only from a change of activity.	Definition of SBS Regulation variables, Eurostat-OECD Manual on Business Demography Statistics (chapter 7).	Real Death Enterprise	Theme: Business Demography
Enterprise group	An enterprise group is an association of enterprises bound together by legal and/or financial links. A group of enterprises can have more than one decision-making centre, especially for policy on production, sales and profit. It may centralise certain aspects of financial management and taxation. It constitutes an economic entity, which is empowered to make choices, particularly concerning the units that it comprises.	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, Annex Section III C		(1) Theme: Derivation of Statistical Units; (2) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (3) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (4) Theme: Statistical Registers and Frames – The statistical units and the business register.
Error	In general, a mistake or error in the colloquial sense.	Eurostat's Concepts and Definitions Database (2013)	Mistake	Theme: Quality of Statistics
Error message	For electronic questionnaire: a window containing a text that described what sort of inconsistency happened and the list of variables involved in it	Memobust definition (2014)		Theme: Data Collection: Techniques and Tools

ESAC	European Statistical Advisory Committee	Regulation N° 99/2013		Theme: The European Statistical System
ESCB	European System of Central Banks	Regulation N° 223/2009		Theme: The European Statistical System
ESGAB	European Statistical Governance Advisory Board	Decision N° 235 (2008)		Theme: The European Statistical System
ESS	European Statistical System. The ESS is the partnership comprising Eurostat, National Statistical Institutes (NSIs) and other national statistical bodies responsible in each Member State (MS) for producing and disseminating European statistics.	ESS Regulation No 223 (2009)		(1) Theme: Different types of surveys; (2) Theme: The European Statistical System
ESSC	European Statistical System Committee	Regulation N° 223/2009		Theme: The European Statistical System
ESS-VIP	ESS Vision Implementation Project	Eurostat website/CROS portal		Theme: The European Statistical System
Establishment	An establishment is defined by the System of National Accounts (SNA) as an enterprise, or part of an enterprise, that is situated in a single location and in which only a single (non-ancillary) productive activity is carried out or in which the principal productive activity accounts for most of the value added. According to the Regulation on statistical units the local kind-of-activity unit (local KAU) corresponds to the operational definition of the establishment. According to the European System of Accounts (ESA) the local KAU is called the establishment in the SNA and ISIC Rev. 3.	System of National Accounts (SNA) 1993, (5.21), P. 116, European System of Accounts (ESA) 1995, [2.106] footnote 15 and 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, Annex Section III G (2)		(1) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (2) Theme: Statistical Registers and Frames – The statistical units and the business register; (3) Theme: Derivation of Statistical Units; (4) Theme: Derivation of Statistical Units
Estimate	The particular value yielded by an estimator in a given set of circumstances.	SDMX (2009)	Estimated value.	(1) Method: Preliminary estimates with design-based methods; (2) Theme: Design of Estimation – Some Practical Issues; (3) Theme: Quality of Statistics; (4) Method: Balanced Sampling for Multi-Way Stratification; (5) Method: Subsampling for Preliminary Estimates; (6) Theme: Methods and Quality; (7) Theme: Quality and Risk Management Models
Estimator	A rule or method of estimating a parameter of a population.	SDMX (2009)		(1) Method: Balanced Sampling for Multi-Way Stratification; (2) Method: Little and Su Method; (3) Method: Subsampling for Preliminary Estimates; (4) Method: Preliminary estimates with design-based methods; (5) Theme: Design of Estimation – Some Practical Issues; (6) Theme: Quality of Statistics

Estimator effect	Ratio between variance of the estimator and variance of the HT estimator for the same sampling design.	Memobust definition (2014)		Method: Generalised regression estimator
ETL	Extract Transform Load. A set of operations to make an external data set suitable for further processing, e.g. at a statistical office.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Evaluation	The systematic and objective assessment of an on-going statistical production process, its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability.	GSBPM (2009)		Theme: Evaluation of Business Statistics
Experiment embedded	The sample of a survey is randomly divided into several groups, which are differently treated and then compared with regard to a hypothesis about treatment effects.	Memobust definition (2014)		Theme: Repeated Surveys
Exports of goods and services	Exports of goods and services consist of transactions in goods and services (sales, barter, and gifts) from residents to non-residents.	ESA (2010)		Theme: Manual Integration
Failure rate	The proportion of records in the unedited data that fail a given edit.	Memobust definition (2014)		Method: Manual Editing
False match rate	number of incorrectly linked record pairs divided by the total number of linked record pairs	Memobust definition (2014)	False positive rate	Method: Fellegi-Sunter and Jaro Approach to Record Linkage
False matches	Matched records which do not represent the same entity	Memobust definition (2014)	mismatch, false positive match, Type I error	Theme: Probabilistic Record Linkage
False negative match	See Missed Match	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
False non-match rate	number of incorrectly unlinked record pairs divided by the total number of true match record pairs	Memobust definition (2014)	False negative rate, Missed match rate	Method: Fellegi-Sunter and Jaro Approach to Record Linkage
False nonmatches	Unmatched records not correctly classified, that imply truly matched entities were not linked.	Memobust definition (2014)	mismatch, false negative match, Type II error	Theme: Probabilistic Record Linkage
False positive match	See Mismatch	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Fatal edit rule	see Hard edit rule	Memobust definition (2014)		

Feasible matching graph	A subgraph of an MC graph that satisfies the criteria that are established for the matching graph. These criteria relate at least to the maximum degree of the points or a part thereof (degree restrictions). The word 'feasible' is used in the sense of 'feasible solution'.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Fellegi-Sunter method	Matching method described in Fellegi and Sunter (1969).	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
First order autoregressive process AR(1)	Model belonging to the class of autoregressive (AR), in which the current level is modelled on the basis of the previous levels.	Memobust definition (2014)		Method: Small area estimation methods for time series data
Flow variable	A flow variable is measured over an interval of time. (see also stock variable)	Memobust definition (2014)		Theme: Macro Integration
Focus area	Combination of an object and one accompanying attribute. Examples: accuracy of estimates, soundness of methodology, clarity of a description.	Van Nederpelt (2012)		Theme: Quality and Risk Management Models
Follow-up	The work performed by an editor during manual editing to handle an edit failure.	Memobust definition (2014)		(1) Method: Manual Editing; (2) Theme: Macro-Editing
Follow-up	A further attempt to obtain information from an individual or a reporting unit in a survey or field experiment because the initial attempt has failed or later information is available.	SDMX (2009)		(1) Method: Preliminary estimates with design-based methods; (2) Method: Subsampling for Preliminary Estimates; (3) Theme: Data Collection; (4) Theme: Data Collection: Techniques and Tools
Foreign key	A key value that occurs in a record but is not suitable to identify the record itself. A foreign key is therefore located outside the key of a data set. The purpose of a foreign key is to make a match with a record in another data set which, for example, includes additional data based on that key.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Frame	A list, map or other specification of the units, which define a population to be completely enumerated or sampled	SDMX (2009), CODED		(1) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (2) Theme: Sample selection; (3) Theme: Asymmetry in Statistics – European Register for Multinationals (EGR)
Frame error	Error caused by imperfections in the frame (business register, population register, area sample, etc.) from which units are selected for inclusion in surveys or censuses.	NQAF (2012)		Theme: Quality and Risk Management Models

Frame population	Frame population is the set of population units described in the survey frame. Remark: Because of the coverage error of the frame population (reference scope) the frame population and the target population is not overlap each other. The part of the frame population belonging to the target population is the survey population	Memobust definition (2014)	Reference scope	(1) Theme: Statistical Registers and Frames – Survey frames for business surveys; (2) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (3) Theme: Asymmetry in Statistics – European Register for Multinationals (EGR); (4) Theme: Statistical Registers and Frames – Main module; Theme: Statistical Registers and Frames – Quality of statistical registers and frames
Frequency	The time interval at which observations occur over a given time period.	SDMX (2009)		(1) Method: Chow-Lin Method for Temporal Disaggregation; (2) Method: Denton for Benchmarking; (3) Theme: Macro Integration
Frequency of register maintenance	Frequency of register maintenance is the time interval of the register content alterations. Remark: Registers can be maintained from different sources with different frequencies. In such cases, the most frequently used source determines the frequency of the register maintenance	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (2) Theme: Statistical Registers and Frames – Statistical register and survey frame design
Frequency tables	See: Tables of frequency (count) data	Memobust definition (2014)		Theme: Statistical disclosure control methods for quantitative tables
Fuzzy string matching	The comparison of two texts, for which the outcome (usually) is a scalar that indicates the extent to which the texts are similar.	Hacking & Willenborg (2012)		(1) Method: Automatic coding based on pre-coded datasets; (2) Method: Automatic coding based on semantic networks; (3) Method: Computer-assisted coding
Gazelle	A gazelle is a high-growth enterprise that is up to 5 years old.	Eurostat-OECD Manual on Business Demography Statistics.		Theme: Business Demography
Generalised regression estimator	An estimator that can be written as the sum of the Horvitz Thompson estimator (HT) and a weighted difference between known totals and their HT estimator.	Memobust definition (2014)	GREG	Method: Generalised regression estimator
Global score function	A global score function is the combination of all defined local score functions, i.e., score functions defined for individual variables.	EDIMBUS Manual		Theme: Selective Editing
Gross burden	All additional costs to businesses arising from their inclusion in a survey if all sampled businesses respond.	Willeboordse <i>et al.</i> (1997), DETI (2009)		Theme: Response Burden

Gross domestic product	Gross Domestic Product (GDP) is one of the key aggregates in the ESA. GDP is a measure of the total economic activity taking place on an economic territory which leads to output meeting the final demands of the economy. There are three ways of measuring GDP at market prices: (1) The production approach, as the sum of the values added by all activities which produce goods and services, plus taxes less subsidies on products; (2) The expenditure approach, as the total of all final expenditures made in either consuming the final output of the economy, or in adding to wealth, plus exports less imports of goods and services; (3) The income approach, as the total of all incomes earned in the process of producing goods and services plus taxes less subsidies on products.	Memobust definition (2014)	GDP	Theme: Manual Integration
Gross measurement errors	are observations that are not true values	Memobust definition (2014)		Method: Outlier Treatment
GSBPM	The Generic Statistical Business Process Model provides a framework to describe the statistical production process in terms of standard components (phases and sub-processes).	GSBPM (2009)		(1) Theme: Specification of User Needs for Business Statistics; (2) Theme: Dissemination of Business Statistics; (3) Theme: Evaluation of Business Statistics
Hamming distance	Distance between two records on a matching key, measured by counting the number of variables with different scores.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching; (5) Theme: Probabilistic Record Linkage
Hard Constraint	A constraint that should hold exactly	Memobust definition (2014)		Method: Denton for Benchmarking
Hard edit rule	An edit rule that identifies data errors with certainty.	EDIMBUS Manual	Fatal edit rule, Logical edit rule	(1) Method: Automatic Editing; (2) Method: Manual Editing; (3) Theme: Statistical Data Editing
Heteroscedasticity	A collection of random variables is heteroscedastic if there are sub-populations that have different variabilities than others.	Memobust definition (2014)		Method: Generalised regression estimator
High-growth enterprise	A high-growth enterprise is an enterprise with average annualised growth greater than 20% per annum, over a three year period. Growth can be measured by the number of employees or by turnover.	Eurostat-OECD Manual on Business Demography Statistics.		Theme: Business Demography
History	Step 9 in the OQRM model, where the history of the focus area is formulated.	Van Nederpelt (2012)		Theme: Quality and Risk Management Models
Hit rate	(1) The proportion of error flags that an edit generates which point to true errors. Or (2) the proportion of error flags generated by an edit that are associated with adjustments made to the data. [Note: this is a practical approximation to the formal definition given under (1).]	EDIMBUS Manual		Method: Manual Editing

Horizontal aggregation	Horizontal aggregation: aggregation, e.g. by country	European Communities (2001)		Theme: Seasonal adjustment – introduction and general description
Horvitz-Thompson estimator	Weighted sum with weights given by the inverse of inclusion probabilities.	Module XIX.0	HT	Method: Generalised regression estimator
Hot-deck imputation	A donor record is found from the same survey as the record with the missing item(s). This donor record is used to supply values for the missing or inconsistent data item(s).	EDIMBUS Manual	Donor imputation	(1) Method: Minimum Adjustment Methods; (2) Method: Reconciling Conflicting Micro-Data; (3) Method: Statistical Matching Methods; (4) Theme: Data fusion at micro level; (5) Theme: Statistical Matching; (6) Theme: Donor Imputation
Hypercube method	A heuristic method for protecting tables through cell suppression.	Memobust definition (2014)		Theme: Statistical disclosure control methods for quantitative tables
Hypernym	A generalisation of a term or a more general term. Opposite of 'hyponym'.	Hacking & Willenborg (2012)		(1) Method: Automatic coding based on semantic networks; (2) Method: Computer-assisted coding
Hyponym	A specialisation of a term or a more specific term. Opposite of 'hypernym'.	Hacking & Willenborg (2012)		(1) Method: Automatic coding based on semantic networks; (2) Method: Computer-assisted coding
Importance	Step 7 in the OQRM model, where the importance of the focus area will be determined, related to the output quality or other objectives.	Van Nederpelt (2012)		Theme: Quality and Risk Management Models
Imports of goods and services	Imports of goods and services consist of transactions in goods and services (purchases, barter, and gifts) from non-residents to residents.	ESA (2010)		Theme: Manual Integration
Imputation	(1) A procedure for entering a value for a specific data item where the response is missing or unusable. Or (2) a value that is filled in during the process described under (1).	UN/ECE Glossary of Terms on Statistical Data Editing (2007), CBS Methods Series Glossary.	Imputing, Imputed value	(1) Theme: Donor Imputation; (2) Theme: Imputation; (3) Theme: Imputation for Longitudinal Data; (4) Theme: Imputation under Edit Constraints; (5) Theme: Model-based Imputation; (6) Method: Minimum Adjustment Methods; (7) Method: Reconciling Conflicting Micro-Data; (8) Method: Statistical Matching Methods; (9) Theme: Data fusion at micro level; (10) Theme: Design of Estimation – Some Practical Issues; (11) Theme: Quality of Statistics; (12) Theme: Statistical Data Editing
Imputation class	A subpopulation for which imputation is carried out, without using any information from the rest of the population. Different imputation methods can be used for different imputation classes.	CBS Methods Series Glossary	Imputation group	(1) Theme: Donor Imputation; (2) Theme: Imputation; (3) Theme: Imputation for Longitudinal Data; (4) Theme: Model-based Imputation

Imputed value	see Imputation (2)	Memobust definition (2014)	Imputation	
Imputing	see Imputation (1)	Memobust definition (2014)	Imputation	
In control	Step 6 in the OQRM model, where is determined if the requirements for a focus area are met and/or if the residual risk is acceptable.	Van Nederpelt (2012)		Theme: Quality and Risk Management Models
Inclusion probability	For a sampling design without replacement, the probability that a particular unit from the population is drawn. This probability may vary between units, depending on the sampling design.	CBS Methods Series Glossary		(1) Theme: Imputation; (2) Method: Balanced Sampling for Multi-Way Stratification
Inclusion probability	The probability that a member of a population will appear in a given sample.	Memobust definition (2014)		(1) Method: Calibration; (2) Method: Synthetic Estimators for Small Area Estimation
Indicator	A data element that represents statistical data for a specified time, place, and other characteristics, and is corrected for at least one dimension (usually size) to allow for meaningful comparisons.	SDMX (2009)		(1) Method: Chow-Lin Method for Temporal Disaggregation; (2) Method: Denton for Benchmarking; (3) Theme: Macro Integration; (4) Theme: Data Collection: Techniques and Tools
Indirect estimator	An estimator that “borrows strength” by taking into account values of the variable under study from outside the domain or time period. These values are brought into the estimation process through a properly chosen model and may come from different sources, for instance censuses or administrative registers.	Memobust definition (2014)		Method: Synthetic Estimators for Small Area Estimation
Industry	A group of producing units, having similar output and production processes; the classification of industries is based on NACE	Memobust definition (2014)		Theme: Manual Integration
Influential error	An error that has a significant influence on figures to be published.	CBS Methods Series Glossary		(1) Theme: Editing Administrative Data; (2) Theme: Editing for Longitudinal Data; (3) Theme: Macro-Editing; (4) Theme: Selective Editing; (5) Theme: Statistical Data Editing
Input editing	Editing that is performed as data is input, e.g., during an interview.	EDIMBUS Manual		Theme: Selective Editing

Institutional unit	The institutional unit is an elementary economic decision-making centre characterized by uniformity of behavior and decision-making autonomy in the exercise of its principal function. A unit is regarded as constituting an institutional unit if it has decision-making autonomy in respect of its principal function and keeps a complete set of accounts. In order to be said to have autonomy of decision in respect of its principal function, a unit must be responsible and accountable for the decisions and actions it takes. In order to be said to keep a complete set of accounts, a unit must keep accounting records covering all its economic and financial transactions carried out during the accounting period, as well as a balance sheet of assets and liabilities. Remark: According to the Regulation on statistical units an institutional unit corresponds to an enterprise in the corporate enterprises sector.	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, Annex, Section III B.		(1) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (2) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (3) Theme: Statistical Registers and Frames – The statistical units and the business register
Interaction burden	A product of the relationship between respondent burden and design burden, e.g. requirement concerning memory and effort to be made, familiarity of the respondent with IT methods and tools, etc.	Hedlin <i>et al.</i> (2005)		Theme: Response Burden
Interactive coding	Coding using an interactive program, which presents the necessary background or other information to a coder, who makes all the coding decisions. The program also processes the answers (and the possible reason for the choices as indicated by the coder).	Hacking & Willenborg (2012)		(1) Method: Computer-assisted coding; (2) Theme: Coding
Interactive editing	An editing method for which a computer program checks the data and a human editor adjusts the data.	CBS Methods Series Glossary	Manual editing	(1) Method: Automatic Editing; (2) Method: Manual Editing; (3) Theme: Editing Administrative Data; (4) Theme: Editing for Longitudinal Data; (5) Theme: Macro-Editing; (6) Theme: Selective Editing; (7) Theme: Statistical Data Editing
Intermediate consumption	Intermediate consumption consists of goods and services consumed as inputs by a process of production, excluding fixed assets whose consumption is recorded as consumption of fixed capital. The goods and services are either transformed or used up by the production process.	ESA (2010)		Theme: Manual Integration
Internet survey	See Web Survey.	Memobust definition (2014)		Theme: Mixed Mode Data Collection – design issues
Interviewer effect	Effects on respondents' answers stemming from the different ways that interviewers administer the same survey.	Memobust definition (2014)		(1) Theme: Data Collection; (2) Theme: Data Collection: Techniques and Tools
Interviewer error	Effects on respondents' answers stemming from the different ways those interviewers administer the same survey. .	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics

Interviewer-administered mode	An interviewer administers and guides the respondent when answering the survey questions.	Memobust definition (2014)		(1) Theme: Data Collection; (2) Theme: Design of data collection (part 1) – Choosing the appropriate data collection method
Intruder	A data user who attempts to link a respondent to a microdata record or make attributions about particular population units from aggregate data. Intruders may be motivated by a wish to discredit or otherwise harm the NSI, the survey or the government in general, to gain notoriety or publicity, or to gain profitable knowledge about particular respondents.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical Disclosure Control
Investment	Gross fixed capital formation consists of resident producers' acquisitions, less disposals, of fixed assets during a given period plus certain additions to the value of non-produced assets realised by the productive activity of producer or institutional units. Fixed assets are produced assets used in production for more than one year.	ESA (2010)	Gross fixed capital formation	Theme: Manual Integration
Inward FATS	'Inward statistics on foreign affiliates' shall mean statistics describing the activity of foreign affiliates resident in the compiling economy.	Foreign Affiliates Statistics (FATS) recommendation manual, version 2012		Theme: Asymmetry in Statistics – European Register for Multinationals (EGR)
Irregular component	This is the residual time series that results from the removal of estimated seasonal and other systematic calendar-related components of an observed time series, along with the removal of an estimated trend-cycle component	US Census Bureau		Method: Seasonal adjustment of economic time series
ITDG	Information Technology Directors Group	Eurostat website/CROS portal		Theme: The European Statistical System
Item non-response	Item non-response occurs when a respondent provides some, but not all, of the requested information, or if some of the reported information is not usable.	EDIMBUS Manual	Partial non-response	(1) Theme: Questionnaire Design; (2) Theme: Editing During Data Collection; (3) Theme: Testing the Questionnaire; (4) Theme: Response Process; (5) Theme: Imputation; (6) Theme: Imputation for Longitudinal Data; (7) Method: Little and Su Method; (8) Theme: Quality of Statistics
Item response rate	The ratio of the number of units which have provided data for a given data item to the total number of units from which data are to be collected or to the number of units that have provided information at least for some data items. It can indirectly measure the level of response burden.	Eurostat (2009).		Theme: Response Burden
Iterative proportional fitting	See multiplicative weighting.	Memobust definition (2014)		(1) Method: RAS; (2) Method: Stone

Jaro distance	Distance counts the number of common characters and the number of transpositions of characters (same character with a different position in the string) between two strings;	Memobust definition (2014)		Theme: Probabilistic Record Linkage
Joining	A form of matching used for databases and in which matching is based on matching keys being identical.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Kalman filter	An iterative technique of dynamic linear modelling, used mainly for estimating the parameters of autoregressive moving-average time series models with Gaussian residuals.	Memobust definition (2014)		Method: Preliminary estimates with model-based methods
Key	See: Object identifier	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Key word	Word in a description that is usable for coding, in contrast to a stop word.	Hacking & Willenborg (2012)		Theme: Coding
Kind-of-activity unit	The kind of activity unit (KAU) groups all the parts of an enterprise contributing to the performance of an activity at class level (4-digits) of NACE Rev. 2 and corresponds to one or more operational subdivisions of the enterprise. The enterprise's information system must be capable of indicating or calculating for each KAU at least the production value, intermediate consumption, manpower costs, the operating surplus and employment and gross fixed capital formation.	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, Annex Section III D		(1) Theme: Derivation of Statistical Units; (2) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (3) Theme: Statistical Registers and Frames – The populations, frames and units of busine; (4) Theme: Statistical Registers and Frames – Statistical register and survey frame design; (5) Theme: Statistical Registers and Frames – The statistical units and the business register.
K-Nearest neighbor imputation	the imputed value is an average of the closest k donors chosen in such a way that some measure of distance between the donors and the recipient is minimized.	EDIMBUS Manual	Distance hot deck	Theme: Statistical Matching
Lagrange multiplier technique	In mathematical optimization, the technique of Lagrange multipliers (named after Joseph Louis Lagrange) provides a strategy for finding the maxima and minima of a function subject to constraints.	Memobust definition (2014)		Method: Denton for Benchmarking
Large outlier	the Y values are extremely larger than the other Y values of the "normal" units	Memobust definition (2014)		Method: Outlier Treatment
Least median of squares	statistical method that attempts to minimise the median of all sample squared residuals	Memobust definition (2014)		Method: Outlier Treatment

Least squares method	One of the most popular methods of finding estimates based on fitting a mathematical model to data, aiming at minimizing the sum of squares of deviations between observed and fitted values.	Memobust definition (2014)		Method: Synthetic Estimators for Small Area Estimation
Legal form	The legal form is defined according to national legislation. It is useful for eliminating ambiguity in identification searches and as the possible criterion for selection or stratification for surveys. It is also used for defining the institutional sector. Statistics according to legal form are produced in business demography. The character of legal or natural person is decisive in fiscal terms, because the tax regime applicable to the unit depends on this. It means that any statistical register fed with fiscal records will have that information. Experience has shown that legal form will often be useful to make adjustments to information collection processes and questionnaires on the legal unit operating an enterprise. A code representing the legal form should therefore be recorded in accordance with the classification of legal forms or categories. The following legal forms can be found in most Member States: Sole proprietorship, Partnership, Limited liability companies, Co-operative societies, Non-profit making bodies, Enterprises with other forms of legal constitution.	Business Register Recommendations Manual (edition 2010), chapter 5, characteristic 1.6	Legal status	Theme: Statistical Registers and Frames – Survey frames for business surveys
Legal local unit	A legal local unit is a part of a legal unit that is located at a certain address. A legal local unit can operate in several different industries. In practice, a legal local unit is the same as a local unit.	Memobust definition (2014)		Theme: Derivation of Statistical Units
Legal unit	Legal units include: - Legal persons whose existence is recognized by law independently of the individuals or institutions which may own them or are members of them. - Natural persons who are engaged in an economic activity in their own right. The legal unit always forms, either by itself or sometimes in combination with other legal units, the legal basis for the statistical unit known as the 'enterprise'.	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, Annex Section II A 3 - 4.		(1) Theme: Derivation of Statistical Units; (2) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (3) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys
Levenshtein distance	Distance measure between two strings, defined as the minimum number of mutations needed to transform one string into the other. A mutation is one of three operations: insertion, deletion or substitution of a character/l into a stringl.	Hacking & Willenborg (2012)	Damerau-Levenshtein distance	(1) Method: Automatic coding based on pre-coded datasets; (2) Method: Automatic coding based on semantic networks; (3) Method: Computer-assisted coding; (4) Theme: Object matching; (5) Method: Object Identifier Matching; (6) Method: Unweighted Matching; (7) Method: Weighted Matching
Leverage	Outlier in the x-direction	Memobust definition (2014)		Method: Outlier Treatment

Linear mixed model (LMM)	Linear model containing both fixed and random effects.	Memobust definition (2014)		Method: Small area estimation methods for time series data
Linked tables	A set of tables with common cells.	Memobust definition (2014)		Theme: Statistical disclosure control methods for quantitative tables
Local kind-of-activity unit	The local kind-of-activity unit (local KAU) is the part of a KAU which corresponds to a local unit. The local KAU corresponds to the operational definition of the establishment. According to the European System of Accounts (ESA) the local KAU is called the establishment in the SNA and ISIC Rev. 3.	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, Annex Section III G, and European System of Accounts (ESA) 1995, [2.106], footnote 15		(1) Theme: Derivation of Statistical Units; (2) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (3) Theme: Statistical Registers and Frames – The populations, frames and units of business; (4) Theme: Statistical Registers and Frames – Survey frames for business surveys; (5) Theme: Statistical Registers and Frames – Statistical register and survey frame design; (6) Theme: Statistical Registers and Frames – The statistical units and the business register
Local unit	The local unit is an enterprise or part thereof (e.g. a workshop, factory, warehouse, office, mine or depot) situated in a geographically identified place. At or from this place economic activity is carried out for which - save for certain exceptions - one or more persons work (even if only part-time) for one and the same enterprise.	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, Annex Section III F.		(1) Theme: Derivation of Statistical Units; (2) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (3) Theme: Statistical Registers and Frames – The statistical units and the business register
Local unit of homogeneous production	The local unit of homogeneous production (local UHP) is the part of a unit of homogeneous production which corresponds to a local unit.	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, Annex Section III H		(1) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (2) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (3) Theme: Statistical Registers and Frames – The statistical units and the business register
LOCF	Last Observation Carried Forward. One method of handling missing data based on existing data.	Memobust definition (2014)		Theme: Imputation for Longitudinal Data
Log	A file that contains log information.	Memobust definition (2014)		Theme: Logging
Log information	Metadata produced during a specific run of a process.	Memobust definition (2014)	A set of logging indicators	Theme: Logging
Logging	Activity of producing log information in a log	Memobust definition (2014)	Tracing	Theme: Logging

Logging indicator	A variable that is logged.	Memobust definition (2014)	Log item	Theme: Logging
Logical edit rule	see Hard edit rule	Memobust definition (2014)	Hard edit rule	
Logical imputation	see Deductive imputation	Memobust definition (2014)	Deductive imputation	(1) Method: Deductive Imputation; (2) Theme: Imputation; (3) Theme: Imputation under Edit Constraints
Longitudinal data	Longitudinal data occurs when the same variables of the same units are measured several times at different moments.	Memobust definition (2014)		(1) Theme: Design of Estimation – Some Practical Issues; (2) Theme: Repeated Surveys; (3) Method: Little and Su Method; (4) Method: Subsampling for Preliminary Estimates; (5) Theme: Different types of surveys; (6) Method: Preliminary estimates with design-based methods
Longitudinal imputation	An umbrella term for imputation methods that make use of observed values for the same variable at other times, either for the same object or for different objects.	CBS Methods Series Glossary		(1) Theme: Imputation; (2) Theme: Imputation for Longitudinal Data
Longitudinal sampling design	Sampling design over time of a given unit of the population.	Memobust definition (2014)		Method: Sample co-ordination using Poisson sampling with permanent random numbers
Lower bound	The lowest possible value of a cell in a table of frequency counts where the cell value has been perturbed or suppressed.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Macro integration	Integrating data from different sources on an aggregate level, to enable a coherent analysis of the data, and to increase the accuracy of estimates.	Memobust definition (2014)	Balancing	(1) Method: Denton for Benchmarking; (2) Method: RAS; (3) Method: Stone; (4) Theme: Macro Integration; (5) Theme: Manual Integration
Macrodata	The result of a statistical transformation process in the form of aggregated information.	SDMX (2009)	Tabular data	(1) Method: Denton for Benchmarking; (2) Method: RAS; (3) Method: Stone; (4) Theme: Macro Integration; (5) Theme: Statistical disclosure control methods for quantitative tables
Macro-editing	An umbrella term for editing methods that (initially) check the data on an aggregate level.	CBS Methods Series Glossary	Output editing	Theme: Editing Administrative Data
Macro-editing	A procedure for tracking suspicious data by checking aggregates or applying statistical methods on all records or on a subset of them.	SDMX (2009)	Output editing	(1) Theme: Macro-Editing; (2) Theme: Statistical Data Editing
Manual coding	Coding performed by a coder, without substantial support from a program.	Hacking & Willenborg (2012)		(1) Method: Computer-assisted coding; (2) Theme: Coding
Manual editing	see Interactive editing	Memobust definition (2014)	Interactive editing	

Marginal table	Table derived from a bigger table by aggregation.	Memobust definition (2014)	Sub table	Theme: Statistical disclosure control methods for quantitative tables
Master frame	Master frame is a snapshot of a register (union of registers) to assign the survey frames based on the given register (registers). Remark: An example of the master frame is the snapshot of the business register to define the survey frames of different economic statistical data collections. Another example can be the snapshot of the address register to make a common frame for population surveys. The common master frame, the common reference period helps the integration and linking of statistical data coming from different surveys.	Handbook on the design and implementation of business surveys		(1) Theme: Statistical Registers and Frames – Main module; Theme: Statistical Registers and Frames – Quality of statistical registers and frames; (2) Theme: Statistical Registers and Frames – Survey frames for business surveys; (3) Theme: Asymmetry in Statistics – European Register for Multinationals (EGR)
Matching	The process of bringing together data (represented in records) relating to units and spread over two data sets, based on common or very similar characteristics in the form of primary or object characteristic values.	Memobust definition (2014)	Record linkage, object linkage	(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Matching candidate graph	A bipartite graph that represents the possible matches between records from two data sets. A bipartite graph is one where the set of points is the union of two disjoint sets, such that each edge has its endpoints in each of these sets.	Memobust definition (2014)	MC graph	(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Matching key	One or multiple key variables that are used in two or more data sets to be matched.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Matching noise	Discrepancy between the data generation mechanism and the imputation generation mechanism. The larger the matching noise, the more distant the usual inferences on the matched data set will be from the inferences that could have been done if the sample was completely observed	Memobust definition (2014)		Theme: Statistical Matching
Matching variables	Common identifiers, either quantitative or qualitative, chosen in order to compare records among files	Memobust definition (2014)	Matching keys	(1) Theme: Probabilistic Record Linkage; (2) Method: Fellegi-Sunter and Jaro Approach to Record Linkage
Matching weight	A nonnegative function defined on the edges of a graph, which associates a non-negative value $G$ with each edge of the $G$ . When matching, this weight expresses how well/poorly records match.	Memobust definition (2014)	Weight	(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching

Maximum Likelihood	Method to estimate a parameter of a probability distribution. More specifically it is the value that maximizes the likelihood function.	Memobust definition (2014)	ML	(1) Method: EBLUP Area Level for Small Area Estimation (Fay-Herriot); (2) Method: EBLUP Unit level for Small Area Estimation; (3) Method: Small area estimation methods for time series data
MC graph	See: Matching candidate graph	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Mean Square Error	Expected value of the square of the difference between the estimator and the parameter. It is the sum of variance and squared bias.	Eurostat's Concepts and Definitions Database (2013)	MSE	(1) Theme: Weighting and Estimation; (2) Method: EBLUP Unit level for Small Area Estimation; (3) Theme: Estimation with administrative data; (4) Method: EBLUP Area Level for Small Area Estimation (Fay-Herriot); (5) Theme: Quality of Statistics
Measure	Step 5 in the OQRM model where action are determined to manage the focus area in order to be in control of the focus area. Context: Preventive, corrective and signalling measures can be distinguished.	Van Nederpelt (2012)		Theme: Quality and Risk Management Models
Measurement error	Error in reading, calculating or recording numerical value.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Measurement error	Measurement is characterized as the difference between the observed value of a variable and the true, but unobserved, value of that variable.	Measuring and Reporting Sources of Error in Surveys, FCSM 2001		Theme: Data fusion at micro level
Measurement error	Error in reading, calculating or recording numerical value. Context: Measurement errors occur when the response provided differs from the real value. Such errors may be attributable to the respondent, the interviewer, the questionnaire, the collection method or the respondent's record-keeping system. Errors may be random or they may result in a systematic bias if they are not random.	SDMX (2009)		Theme: Quality and Risk Management Models
Measurement error	The difference between the observed value of a variable and the true, but unobserved, value of that variable.	SDMX (2009)		(1) Method: Denton for Benchmarking; (2) Theme: Macro Integration
Measurement errors	Measurement errors occur when the response provided differs from the real value; such errors may be attributable to the respondent, the interviewer, the questionnaire, the collection method or the respondent's record-keeping system. Such errors may be random or they may result in a systematic bias if they are not random.	Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 59.		(1) Theme: Questionnaire Design; (2) Theme: Editing During Data Collection; (3) Theme: Testing the Questionnaire; (4) Theme: Response Process

Merger	This event can be seen as the opposite of a break-up. It involves a consolidation of the production factors of two or more enterprises into one new enterprise, in such a way that the previous enterprises are no longer recognisable. There is no continuity or survival, but the closures of the previous enterprises are not considered to be deaths. Similarly the new enterprise is not considered to be a birth.	Eurostat-OECD Manual on Business Demography Statistics (chapter 4).		Theme: Business Demography
Metadata	Information that is needed to be able to use and interpret statistics.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Methodology	A structured approach to solve a problem.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Metric	A metric $d$ for a set $X$ is defined as nonnegative function that measures how far two points in $X$ are apart.	Memobust definition (2014)	Distance; Distance Function	(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Micro data	Non-aggregated observations or measurements of characteristics of individual units.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Micro editing	An exhaustive check to find errors by inspecting each individual observation.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Micro integration	A method that matches data on individual statistical units from different sources, to obtain a combined data file with better information. The quality of the data is measured in terms of validity, reliability and consistency.	Memobust definition (2014)		(1) Method: Denton for Benchmarking; (2) Method: Stone
Microdata	Non-aggregated observations or measurements of characteristics of individual units.	SDMX (2009)		(1) Method: Denton for Benchmarking; (2) Method: RAS; (3) Method: Stone; (4) Theme: Macro Integration; (5) Theme: Methods and Quality; (6) Theme: Statistical Disclosure Control
Micro-selection	see Selective editing	Memobust definition (2014)	Selective editing	Theme: Statistical Data Editing
Misclassification	Erroneous classification of a subject into a category in which the subject does not belong. For instance, a business is classified in Trade instead of Industry.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Mismatch	A match that has been made erroneously.	Memobust definition (2014)	False positive match; Type I error	(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Missed error rate	The proportion of errors in the unedited data that were not flagged by any edits in a given set.	Memobust definition (2014)		Method: Manual Editing

Missed match	A match that should have been made but was not.	Memobust definition (2014)	False negative match; Type II error	(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Missing data	Observations which were planned and are missing	SDMX (2009)		(1) Method: Chow-Lin Method for Temporal Disaggregation; (2) Method: Denton for Benchmarking; (3) Method: RAS; (4) Method: Stone; (5) Theme: Macro Integration
Mixed-mode survey	A survey where multiple modes are used to collect data from the sampled units in the data collection period of one survey.	Memobust definition (2014)		Theme: Mixed Mode Data Collection – design issues
Mode effect	The effect that using a specific mode has on the responses that are obtained in that mode. Mode effects may be interpreted as a form of measurement bias.	De Leeuw, Hox & Dillman (2008)		Theme: Mixed Mode Data Collection – design issues
Mode effect	A pure mode effect is essentially a measurement bias that is specifically attributable to the mode. In some surveys the mode effects are small because the same questionnaire can be used across all modes. Most problems occur when mail is combined with an interviewer-administered mode.	Memobust definition (2014)		(1) Theme: Data Collection; (2) Theme: Design of data collection (part 1) – Choosing the appropriate data collection method
Mode of data collection	Mode refers to what medium is used when contacting the sample members to get their responses.	Memobust definition (2014)		(1) Theme: Data Collection; (2) Theme: Design of data collection (part 1) – Choosing the appropriate data collection method
Model assumption error	Error that occurs due to the use of methods, such as calibration, generalized regression estimator, 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.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Model assumption error	Model assumption errors occur with the use of methods, such as calibration, generalised regression estimator, 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.	SDMX (2009)		(1) Theme: Methods and Quality; (2) Theme: Quality and Risk Management Models
Model based imputation	Imputation based on an explicitly described statistical model. E.g. use of averages, medians, regression equations, etc. to impute a variable.	EDIMBUS Manual		(1) Method: Statistical Matching Methods; (2) Theme: Imputation; (3) Theme: Imputation for Longitudinal Data; (4) Theme: Imputation under Edit Constraints; (5) Theme: Model-based Imputation
Movement preservation principle	The property that <i>all</i> changes of the sub annual data are kept as much as possible at their initial values.	Memobust definition (2014)		Method: Denton for Benchmarking

Moving holiday effects	These are systematic changes in the values of a time series that are associated with the timing of moving holidays, i.e. holidays whose dates vary from year to year, such as Easter, Passover, Ramadan, Chinese New Year and U.S. Labor Day. Estimates of one or a combination of such effects define the moving holiday component of time series	US Census Bureau		Method: Seasonal adjustment of economic time series
Moving seasonality	Moving seasonality is a form of seasonality that accounts for the variability in the seasonal component of a time series from year to year	ABS (2008)		(1) Method: Seasonal adjustment of economic time series; (2) Theme: Seasonal adjustment – introduction and general description
Multiple activity business	A business operating in several economic activities	Memobust definition (2014)		Method: Assigning random numbers when co-ordination of surveys based on different unit types is considered
Multiple imputation	An observation with failing and/or missing values is imputed several times stochastically. Multiple imputation allows under certain conditions the correct estimation of the variance due to imputation. This estimation is based on a combination of the within and the between variance of the multiply imputed data.	EDIMBUS Manual		Theme: Imputation
Multiple location business	A business operating in several geographical locations	Memobust definition (2014)		Method: Assigning random numbers when co-ordination of surveys based on different unit types is considered
Multiplicative weighting	A form of weighting for which the weights are obtained by multiplying relevant weight factors, determined in an iterative process. Multiplicative weighting is also referred to as raking or iterative proportional fitting.	Memobust definition (2014)		(1) Method: RAS; (2) Method: Stone
Multistage sampling	A complex form of cluster sampling	Wikipedia Multistage Sampling		Theme: Sample selection
Multivariate imputation	Imputing several missing values in a record.	CBS Methods Series Glossary		Theme: Imputation
NACE	Classification of economic activities in the European Community (referred to as 'NACE Rev. 1' or 'NACE Rev. 1.1').	Council Regulation (EEC) No 3037/90.		Theme: Small area estimation
NACE	NACE (Statistical classification of economic activities) is the European standard classification of productive economic activities. NACE presents the universe of economic activities partitioned in such a way that a NACE code can be associated with a statistical unit carrying them out. NACE provides the framework for collecting and presenting a large range of statistical data according to economic activity in the fields of economic statistics.	NACE Rev.2		(1) Theme: Different types of surveys; (2) Theme: Estimation with administrative data

NACE	Nomenclatures statistique des activités économiques dans la Communauté Européenne.	NACE Rev. 2 (ISSN 1977-0375)		Theme: The European Statistical System
NACE	General Industrial Classification of Economic Activities within the European Communities (1970 version); Statistical classification of economic activities in the European Community (after 1970)	RAMON, Eurostat's metadata server -classification	NACE Rev. 2	(1) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (2) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (3) Theme: Statistical Registers and Frames – Survey frames for business surveys; (4) Theme: Statistical Registers and Frames – The statistical units and the business register
Nearest neighbor imputation	The donor is chosen in such a way that some measure of distance between the donor and the recipient is minimized.	EDIMBUS Manual	Distance hot deck	Method: Statistical Matching Methods
Negative co-ordination	Minimize the overlap between samples	Memobust definition (2014)		Theme: Sample co-ordination
Net burden	The opposite of gross burden – the total costs actually incurred by responding businesses; this type of burden accounts for “benefits” enjoyed by respondents for their contribution whereas gross burden ignores them.	Willeboordse <i>et al.</i> (1997)		Theme: Response Burden
no terms	no terms	Memobust definition (2014)		Theme: GSBPM: Generic Statistical Business Process Model
No-answer	In telephone surveys: the line sounds but nobody answer the telephone	Memobust definition (2014)		Theme: Data Collection: Techniques and Tools
Non probability sample	A sample in which the selection of units is based on factors other than random chance, e.g. convenience, prior experience or the judgement of a researcher.	SDMX (2009)		Theme: Weighting and Estimation
Non response	A form of non observation present in most surveys. Non response means failure to obtain a measurement on one or more study variables for one or more elements k selected for the survey. The term encompasses a wide variety of reasons for non observation: "impossible to contact", "not at home", "unable to answer", "incapacity", "hard core refusal", "inaccessible", "unreturned questionnaire", and others. In the first two cases contact with the selected element is never established.	SDMX (2009)		Method: Subsampling for Preliminary Estimates
Non response error	Error that occurs when the survey fails to get a response to one, or possibly all, of the questions.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics

Non response rate	In sample surveys, the failure to obtain information from a designated individual for any reason (death, absence or refusal to reply) is often called a non-response and the proportion of such individuals of the sample aimed at is called the non-response rate.	SDMX (2009)		Method: Subsampling for Preliminary Estimates
Nonbinding benchmarking	A benchmarking problem with at least one nonbinding annual alignment constraint.	Memobust definition (2014)		Method: Denton for Benchmarking
Nonbinding constraint	See soft constraint	Memobust definition (2014)		Method: Denton for Benchmarking
Non-probability sample	A sample in which the selection of units is based on factors other than random chance, e.g. convenience, prior experience or the judgement of a researcher.	SDMX (2009)		(1) Method: Balanced Sampling for Multi-Way Stratification; (2) Method: Subsampling for Preliminary Estimates; (3) Theme: Sample selection
non-representative outlier	are unique in population (in the sense that there is no other unit like them)	Memobust definition (2014)		Method: Outlier Treatment
Non-response	A form of non observation present in most surveys. Nonresponse means failure to obtain a measurement on one or more study variables for one or more elements k selected for the survey. The term encompasses a wide variety of reasons for non observation: "impossible to contact", "not at the address", "unable to answer", "incapacity", "hard core refusal", "inaccessible", "unreturned questionnaire", and others. In the first two cases contact with the selected element is never established.	SDMX (2009)		(1) Theme: Questionnaire Design; (2) Theme: Editing During Data Collection; (3) Theme: Testing the Questionnaire; (4) Theme: Response Process; (5) Method: Preliminary estimates with design-based methods
Non-response error	Context: Non-sampling error may arise from many different sources such as defects in the sampling frame, faulty demarcation of sample units, defects in the selection of sample units, mistakes in the collection of data due to personal variations, misunderstanding, bias, negligence or dishonesty on the part of the investigator or of the interviewer, mistakes at the stage of the processing of the data, etc.	Memobust definition (2014)		Theme: Quality and Risk Management Models
Non-response error	Non- sampling errors may be categorised as: § Coverage errors (or frame errors) due to divergences between the target population and the frame population ; § Measurement errors occurring during data collection. § Non-response errors caused by no data collected for a population unit or for some survey variables. § Processing errors due to errors introduced during data entry, data editing, sometimes coding and imputation. § Model assumption errors.	Memobust definition (2014)		Theme: Quality and Risk Management Models
Non-response error	Error that occurs when the survey fails to get a response to one, or possibly all, of the questions.	NQAF (2012)		Theme: Quality and Risk Management Models
Non-response error	Error in sample estimates which cannot be attributed to sampling fluctuations.	SDMX (2009)		Theme: Quality and Risk Management Models

Non-response rate	In sample surveys, the failure to obtain information from a designated unit for any reason is often called a nonresponse and the proportion of such units of the sample aimed at is called the nonresponse rate.	SDMX (2009)		Method: Preliminary estimates with design-based methods
Non-sampling error	Error in sample estimates which cannot be attributed to sampling fluctuations.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Non-sampling error	An error in sample estimates which cannot be attributed to sampling fluctuations.	The International Statistical Institute, The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003.		Theme: Editing During Data Collection
Normal distribution	One of the most widely known and used of all distributions, sometimes referred to as Gaussian distribution. The continuous probability distribution with two parameters: the expected value and the variance.	Memobust definition (2014)		Method: Synthetic Estimators for Small Area Estimation
Not at Random sample	A sample in which the selection of units is based on factors other than random chance, e.g. convenience, prior experience or the judgement of a researcher	SDMX (2009)		Method: Subsampling for Preliminary Estimates
Nowcast	A forecast relating to the current time (or, rather, to the recent past) and produces an estimate for the period just behind us, but for which no direct statistical observation has been made.	Daas and Arends-Toth (2012)		(1) Theme: Collection and Use of Secondary Data; (2) Theme: Estimation with administrative data
NSA	National Statistical Authority: a non-NSI also responsible for official statistics.	Memobust definition (2014)		Theme: The European Statistical System
NSI	A National Statistical Institute is the leading statistical agency within a national statistical system.	OECD Glossary of Statistical Terms		(1) Theme: Specification of User Needs for Business Statistics; (2) Theme: Dissemination of Business Statistics; (3) Theme: Evaluation of Business Statistics; (4) Theme: Estimation with administrative data; (5) Theme: Different types of surveys; (6) Theme: Response Burden; (7) Theme: The European Statistical System; (8) Theme: Statistical Disclosure Control
NSO	National statistical office - NSI or other office producing official statistics	ESS Handbook for Quality Reports. Eurostat Methodologies and Working Papers. Luxembourg: Office for Official Publications of the European Communities.		Theme: Different types of surveys
NSTR	Nomenclature uniformes des marchandises pour les Statistiques de Transport, Révisé	Eurostat website/CROS portal		Theme: Coding

NUTS	Common regional classification, (called 'Nomenclature of territorial units for statistics' or NUTS).	Council Regulation (EEC) No. 1059/2003.		Theme: Small area estimation
NUTS	The NUTS classification (Nomenclature of territorial units for statistics) is a hierarchical system for dividing up the economic territory of the EU.	NUTS classification		Theme: Different types of surveys
Object	Everything that can be perceived or conceived. Examples: output, process, input, staff, software, methodology, document. Context: For an organization, a specific set of objects are relevant like customers, products, processes, input, data, software, staff, etc.	ISO 1179 (2004)	Component or entity	(1) Theme: Methods and Quality; (2) Theme: Quality and Risk Management Models; (3) Theme: Quality of Statistics
Object characteristic	A combination of variables that can be used in the identification of units, but which are not used as object identifier. Often, this concerns variables (or a combination thereof) such as name, address, place of residence, date of birth, profession, education, gender, etc. None of these variables can identify the record by themselves, but the combination can be used as a proxy for a object identifier, if this is missing.	Memobust definition (2014)	Secondary key	(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Object identifier	In database technology, the object identifier is the name for a variable or a combination of variables that satisfy the following requirements: - the value of the variable (or the combination of variables) is unique in the table (or data set) and therefore unambiguously defines the record in which it occurs. - the variable (or the combination of variables) is filled in everywhere and therefore cannot be empty. - The combination of variables is minimal: by eliminating one of the variables, the record is no longer unambiguously defined. If related tables refer to the table in which the variable (or combination) of variables occur, this is used to establish a relationship between tables.	Memobust definition (2014)	Primary key; Key	(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Objective burden	Burden referring directly to the actual cost of completing questionnaires by respondents; subjective burden reflects their perception	Willeboordse <i>et al.</i> (1997)		Theme: Response Burden
Observation unit	An observation unit represents an identifiable entity about which data can be obtained and for which data is recorded. It should be noted that this may, or may not be, the same as the reporting unit. Remark: It may not be known in advance (e.g. commodities).	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (2) Theme: Data Collection: Techniques and Tools
OECD	Organisation for Economic Co-operation and Development	OECD		Theme: The European Statistical System

On-site facility	A facility that has been established on the premises of several NSIs. It is a place where external researchers can be permitted access to potentially disclosive data under contractual agreements which cover the maintenance of confidentiality, and which place strict controls on the uses to which the data can be put. The on-site facility can be seen as a ' safe setting' in which confidential data can be analysed. The on-site facility itself would consist of a secure hermetic working and data storage environment in which the confidentiality of the data for research can be ensured. Both the physical and the IT aspects of security would be considered here. The on-site facility also includes administrative and support facilities to external users, and ensures that the agreed conditions for access to the data were complied with.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical Disclosure Control
Open-ended question	Question that let respondent answer using their own words	Memobust definition (2014)		(1) Theme: Data Collection: Techniques and Tools; (2) Theme: Design of data collection (part 1) – Choosing the appropriate data collection method
Opportunities	Step 8 in the OQRM model, where opportunities are analysed if a focus area meets the requirements and more.	Van Nederpelt (2012)		Theme: Quality and Risk Management Models
OQRM	Object-oriented Quality and Risk Management	Van Nederpelt (2012)		(1) Theme: Methods and Quality; (2) Theme: Quality and Risk Management Models
Ordinary rounding	See: Conventional rounding.	Glossary on Statistical Disclosure Control (2014)	Conventional rounding	Theme: Statistical disclosure control methods for quantitative tables
OS, PS, TS	Observed Sample - respondent units for the final estimates; Preliminary Sample - quick respondent units; Theoretical Sample-planned sample.	Memobust definition (2014)		Method: Preliminary estimates with model-based methods
Outlier	An outlier is an observation which is not fitted well by a model for the majority of the data. For instance, an outlier may lie in the tail of the statistical distribution or "far away from the centre" of the data.	EDIMBUS Manual		(1) Theme: Statistical Data Editing; (2) Theme: Design of Estimation – Some Practical Issues; (3) Theme: Editing Administrative Data; (2) Theme: Macro-Editing; (4) Method: Outlier Treatment
Outlier in the x-direction	See x-outliers	Memobust definition (2014)		
Outlier in the y-direction	See y-outliers	Memobust definition (2014)		
Outliers	An outlier is a data value that lies in the tail of the statistical distribution of a set of data values.	OECD (2006)		(1) Method: Seasonal adjustment of economic time series; (2) Theme: Issues on Seasonal Adjustment; (3) Theme: Seasonal adjustment – introduction and general description

Output editing	A procedure for tracking suspicious data by checking aggregates or applying statistical methods on all records or on a subset of them.	SDMX (2009)	Macro editing	(1) Theme: Editing for Longitudinal Data; (2) Theme: Selective Editing; (3) Theme: Statistical Data Editing; (4) Theme: Macro-Editing
Outward FATS	'Outward statistics on foreign affiliates' shall mean statistics describing the activity of foreign affiliates abroad controlled by the compiling economy.	Foreign Affiliates Statistics (FATS) recommendation manual, version 2012		Theme: Asymmetry in Statistics – European Register for Multinationals (EGR)
Over-coverage	Over-coverage arises from the presence in the frame of units not belonging to the target population and of units belonging to the target population that appear in the frame more than once.	Eurostat, "Assessment of Quality in Statistics: Glossary",		(1) Theme: Weighting and Estimation; (2) Theme: Quality of Statistics; (3) Theme: Sample selection; (4) Theme: Design of Estimation – Some Practical Issues
Overediting	Editing of data beyond a certain point after which as many errors are introduced as are corrected.	UN/ECE Glossary of Terms on Statistical Data Editing (2007)		(1) Method: Automatic Editing; (2) Method: Manual Editing
Panel	A set of units, which is included several times in a repeated survey according to a specified pattern	Memobust definition (2014)		(1) Method: Little and Su Method; (2) Theme: Imputation for Longitudinal Data; (3) Theme: Design of Estimation – Some Practical Issues
Panel survey	A survey where elements are followed over time.	Memobust definition (2014)	Longitudinal survey	Theme: Weighting and Estimation
Paper and Pencil Interviewing	"Paper" is a method of data collection without the assistance of an interviewer. A questionnaire is sent to respondents, they write in their responses and send it back to the data collection organization.	Memobust definition (2014)	PPI	Theme: Mixed Mode Data Collection – design issues
PAPI	Pencil And Paper Interviewing.	Hacking & Willenborg (2012)		(1) Theme: CATI Allocation; (2) Theme: Data Collection; (3) Theme: Coding
Paradata	Paradata, also termed process data contain information about the primary data collection process (e.g. survey duration, interim status of a case, navigational errors in a survey questionnaire). They can provide a means of additional control over or understanding of the quality of the primary data (the responses to the survey questions).	Memobust definition (2014)		Theme: Data Collection
Parallel mixed mode	Using two or more modes at the same time, e.g. letting the respondent choose his preferred mode.	Memobust definition (2014)		(1) Theme: Data Collection; (2) Theme: Design of data collection (part 2) – Contact strategies
Partial non-response	Also known as item non-response, defines the case of unit may that may respond to the questionnaire incompletely	Memobust definition (2014)	Item non-response	Theme: Data Collection: Techniques and Tools
Pencil And Paper Interviewing.	See PAPI	Memobust definition (2014)		(1) Theme: CATI Allocation; (2) Theme: Data Collection; (3) Theme: Coding

Perceived burden	Burden felt subjectively by the respondent, e.g. connected with the length of the questionnaire, difficulty of the questions, effort required to answer these questions, time spent, etc. or disadvantageous perception of the survey by some respondents, i.e. weak willingness to respond, insufficient awareness of the usefulness of participation, etc.	Willeboordse <i>et al.</i> (1997), Hedlin <i>et al.</i> (2005)	Subjective burden	Theme: Response Burden
Permanent Random Number	A unique random number permanently associated with a unit in a register	Memobust definition (2014)		(1) Method: Assigning random numbers when co-ordination of surveys based on different unit types is considered; (2) Method: Sample co-ordination using simple random sampling with permanent random numbers; (3) Theme: Sample co-ordination
Pilot survey	A survey, usually on a small scale, carried out prior to the main survey, primarily to gain information to improve the efficiency of the main survey. For example, it may be used to test a questionnaire, to ascertain the time taken by field procedure or to determine the most effective size of sampling unit.	Dictionary of Statistical Terms, 5th edition, prepared for the International Statistical Institute by F.H.C. Marriott, Longman Scientific and Technical	Exploratory survey	Theme: Testing the Questionnaire
Planning period	Period in which CATI interviewers are scheduled. This can be a period, of say, 4 weeks starting from the current date, or a calendar month, depending on the allocation variant applied.	Memobust definition (2014)		Theme: CATI Allocation
Poisson sampling design	Sampling design whereby the selection of any unit of the population into the sample is decided independently from the selection of other units.	Memobust definition (2014)		Method: Sample co-ordination using Poisson sampling with permanent random numbers
Population	Population is the total membership or population or "universe" of a defined class of people, objects or events. There are two types of population, viz, target population and survey population. A target population is the population outlined in the survey objects about which information is to be sought and a survey population is the population from which information can be obtained in the survey. The target population is also known as the scope of the survey and the survey population is also known as the coverage of the survey. For administrative records the corresponding populations are: the "target" population as defined by the relevant legislation and regulations, and the actual "client population".	RAMON, Eurostat's metadata server		(1) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (2) Theme: Statistical Registers and Frames – Survey frames for business surveys; (3) Theme: Statistical Registers and Frames – Statistical register and survey frame design
Positive co-ordination	Maximize the overlap between samples	Memobust definition (2014)		Theme: Sample co-ordination
Positive predictive value	number of correctly linked record pairs divided by the total number of linked record pairs (one minus the false match rate)	Memobust definition (2014)	Precision	Method: Fellegi-Sunter and Jaro Approach to Record Linkage

p-percent rule	A (p,q) rule where q is 100 %, meaning that from general knowledge any respondent can estimate the contribution of another respondent to within 100 % (i.e., knows the value to be nonnegative and less than a certain value which can be up to twice the actual value).	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
PPOS	Planned Preliminary Observed Sample: The respondents of the PTS.	Memobust definition (2014)		Method: Subsampling for Preliminary Estimates
pps	Probability proportional to size	Memobust definition (2014)		Theme: Sample selection
PRB	Perceived Response Burden Study. A survey aimed at recognition and assessment of perceived response burden conducted using some common methodological framework, e.g. PRB Core Questions are a basis to construct relevant target-adjusted questionnaires on burden perceived by respondents in relation to a given statistical survey.	Dale and Haraldsen (2007)		Theme: Response Burden
Precision	number of correctly linked record pairs divided by the total number of linked record pairs	Memobust definition (2014)	Positive predicted value	Method: Fellegi-Sunter and Jaro Approach to Record Linkage
Precision rate	Percentage correct coded texts on the total of coded texts	D'Orazio M. and Macchia S (ROS) (2002)	Accuracy	Theme: Measuring Coding Quality
Predicted values	See anticipated values.	Memobust definition (2014)		Theme: Selective Editing
Preliminary estimates	Estimates based on a preliminary sample	Memobust definition (2014)		(1) Theme: Weighting and Estimation; (2) Theme: Estimation with administrative data
Preliminary sample	Partial sample based on early respondents.	Memobust definition (2014)		Theme: Weighting and Estimation
Preventive measure	Measure to avoid a quality problem.	Memobust definition (2014)		Theme: Quality and Risk Management Models
Price index	The result of a formula in which price changes of various goods and services are weighed together in order get an index for the aggregate.	Memobust definition (2014)		Theme: Manual Integration
Primacy effect	A given response alternative is more likely to be chosen when presented at the beginning rather than at the end of a list of response alternatives.	Memobust definition (2014)	Primacy	Theme: Design of data collection (part 1) – Choosing the appropriate data collection method
Primary confidentiality	It concerns tabular cell data, whose dissemination would permit attribute disclosure. 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.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Primary data	Data collected on behalf of an NSI and for which the NSI has defined the conceptual and process metadata	Daas and Arends-Toth (2012)		Theme: Collection and Use of Secondary Data
Primary data collection	The gathering of primary data by an NSI	Daas and Arends-Toth (2012)		Theme: Collection and Use of Secondary Data

Primary key	See Object identifier	Memobust definition (2014)		
Primary protection	Protection using disclosure control methods for all cells containing small counts or cases of dominance.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Primary research	Research that uses primary data	Golden (1976)		Theme: Collection and Use of Secondary Data
Primary source	A source containing primary data	Golden (1976)		Theme: Collection and Use of Secondary Data
Primary suppression	This technique can be characterized as withholding all disclosive cells from publication, which means that their value is not shown in the table, but replaced by a symbol such as 'x' to indicate the suppression. According to the definition of disclosive cells, in frequency count tables all cells containing small counts and in tables of magnitudes all cells containing small counts or representing cases of dominance have to be primary suppressed.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Principal activity	The principal (or main) activity is identified as the activity which contributes most to the total value added of a unit under consideration. The principal activity so identified does not necessarily account for 50 % or more of the unit's total value added. The classification of principal activity is determined by reference to NACE Rev. 2, first at the highest level of classification and then at more detailed levels ("top-down" method).	Business Register Recommendations Manual (edition 2010), chapter 5, characteristic 2.6, 3.6, 4.7		(1) Theme: Derivation of Statistical Units; (2) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (3) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (4) Theme: Statistical Registers and Frames – Survey frames for business surveys; (5) Theme: Statistical Registers and Frames – The statistical units and the business register
Prior-posterior rule	See: (p,q) rule.	Glossary on Statistical Disclosure Control (2014)	(p,q) rule; ambiguity rule	Theme: Statistical disclosure control methods for quantitative tables
Probabilistic record linkage	Linkage method that makes an explicit use of probabilities for deciding when a given pair of records is actually a match or not	Memobust definition (2014)	Weighted matching	Theme: Probabilistic Record Linkage
Probability sample	A sample selected by a method based on the theory of probability (random process), that is, by a method involving knowledge of the likelihood of any unit being selected.	SDMX (2009)		(1) Method: Balanced Sampling for Multi-Way Stratification; (2) Theme: Sample co-ordination; (3) Theme: Sample selection; (4) Theme: Weighting and Estimation
Probing	Follow-up questions that interviewers can ask in addition to those written on the questionnaire to get more adequate information from respondents.	Memobust definition (2014)		(1) Theme: Data Collection: Techniques and Tools; (2) Theme: Design of data collection (part 1) – Choosing the appropriate data collection method

Processing error	Error in final survey results arising from the faulty implementation of correctly planned implementation methods. Context: In survey data, for example, processing errors may include transcription errors, coding errors, data entry errors and errors of arithmetic in tabulation.	NQAF (2012)		(1) Theme: Quality and Risk Management Models; (2) Theme: Quality of Statistics
PRODCOM	A classification of industrial products	Eurostat website/CROS portal		Theme: Coding
Production	Production is an activity carried out under the control, responsibility and management of an institutional unit that uses inputs of labour, capital and goods and services to produce outputs of goods and services.	ESA (2010)	Output	Theme: Manual Integration
Profiling	Profiling is a method to analyse the legal, operational and accounting structure of an enterprise group in order to establish the statistical units within that group and their links and the most efficient structures for the collection of statistical data	Business Register Recommendations Manual (edition 2010), chapter 19B		Theme: Derivation of Statistical Units
Pro-rata method	A straightforward, widely known benchmarking method that achieves consistency between annual and sub annual time series by multiplying all sub annual periods by correction factors defined by the ratio between an annual value and the sum of all underlying sub annual values. These correction factors are called <i>proportional annual discrepancies</i> .	Memobust definition (2014)		Method: Denton for Benchmarking
Provider load	The effort, in terms of time and cost, required for respondents to provide satisfactory answers to a survey.	Australian Bureau of Statistics, Service Industries Statistics, "Glossary of Terms"; unpublished on paper	Respondent burden	(1) Theme: Testing the Questionnaire; (2) Theme: Response Process
PTS	Preliminary Theoretical Sample. The planned sample draws to obtain the provisional estimates	Memobust definition (2014)		Method: Subsampling for Preliminary Estimates
Punctuality	Time lag between the release date of data and the target date on which they were scheduled for release as announced in an official release calendar.	ESS Handbook for Quality Reports (2009)		(1) Theme: Quality of Statistics; (2) Theme: Overall Design
Punctuality of a log	The period between the delivery time of the log and the planned delivery and time.	Memobust definition (2014)		Theme: Logging
PUPOS	Partially Unplanned Preliminary Observed Sample. A subset of the final sample with a specific follow-up plan. Usually the large units of the final sample	Memobust definition (2014)		Method: Subsampling for Preliminary Estimates
Purposive sample	See non-random sample	Memobust definition (2014)		Method: Subsampling for Preliminary Estimates

Qualitative data	Data describing the attributes or properties that an object possesses.	Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva (2000)		Theme: Testing the Questionnaire
Quality	Quality is the degree to which a set of (inherent) characteristics fulfils requirements.	Eurostat's Concepts and Definitions Database (2013), ISO 9000 (2005)		(1) Theme: Quality of Statistics; (2) Theme: Quality and Risk Management Models; (3) Theme: Quality and Risk Management Models
Quality assurance	Part of quality management focused on providing confidence that quality requirements will be fulfilled.	ISO 9000 (2005)		Theme: Overall Design
Quality Circles	Structured employee involvement groups operating in designated work areas that meet regularly to identify work related problems and to suggest solutions or improvements to management.	OECD Glossary of Statistical Terms		Theme: Evaluation of Business Statistics
Quality control	Part of quality management focused on fulfilling quality requirements.	ISO 9000 (2005)		Theme: Overall Design
Quality dimension	Characteristic	Memobust definition (2014)	Criterion, Quality component, Quality aspect, Attribute	(1) Theme: Quality of Statistics; (2) Theme: Methods and Quality
Quality indicator	Variable that represents the quality of data or process.	Memobust definition (2014)		Theme: Quality and Risk Management Models
Quantitative tables	See: Tables of magnitude data	Memobust definition (2014)	Tables of magnitude data	(1) Theme: Statistical disclosure control methods for quantitative tables; (2) Theme: Statistical Disclosure Control
Query edit rule	see Soft edit rule	Memobust definition (2014)	Soft edit rule	
Question format	The way the question is structured. Possible formats: single-choice questions, multi-choice question, table, matrix, partial open-ended question (single or multi choice with other specify), open-ended question.	Memobust definition (2014)		(1) Theme: Data Collection; (2) Theme: Data Collection: Techniques and Tools
$R^2$	Coefficient of determination. It provides a measure of how well observed outcomes are replicated by the model, as the proportion of total variation of outcomes explained by the model.	Memobust definition (2014)		Method: Chow-Lin Method for Temporal Disaggregation
Raking	See multiplicative weighting.	Memobust definition (2014)		(1) Method: RAS; (2) Method: Stone
Random error	Antonym of Systematic error	Memobust definition (2014)		(1) Method: Automatic Editing; (2) Method: Deductive Editing; (3) Theme: Statistical Data Editing

Random error	The degree to which the error in the estimate spreads around zero.	Van Nederpelt (2009)	Variance, Precision.	Theme: Quality of Statistics
Random hot deck	A donor record is randomly selected for each recipient record (record with missing information). Usually selection is carried out after grouping units according to some characteristics (e.g. same gender, region, etc.)	Memobust definition (2014)		Method: Statistical Matching Methods
Random rounding	In order to reduce the amount of data loss that occurs with suppression, alternative methods have been investigated to protect sensitive cells in tables of frequencies. Perturbation methods such as random rounding and controlled rounding are examples of such alternatives. In random rounding cell values are rounded, but instead of using standard rounding conventions a random decision is made as to whether they will be rounded up or down. The rounding mechanism can be set up to produce unbiased rounded results.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Random sample	A sample selected by a method based on the theory of probability (random process), that is, by a method involving knowledge of the likelihood of any unit being selected.	SDMX (2009)	Probability sample	(1) Method: Subsampling for Preliminary Estimates; (2) Method: Balanced Sampling for Multi-Way Stratification
random walk model	Model formalization of a random path consisting of a succession of random steps.	Memobust definition (2014)		Method: Small area estimation methods for time series data
Rank hot deck	The donor is chosen in such a way that some measure of distance among percentage points of the empirical distribution is minimized.	Memobust definition (2014)		Method: Statistical Matching Methods
Raw description	Description recorded in an interview or specified by a respondent and that has not been (thoroughly) checked. This may contain various errors, along with insufficient or unnecessary (stop words) information. This is why descriptions are first subjected to several grammatical treatments. This creates a clean or cleansed string, which is used for automatic coding. This string is not intended to be readable, but is utilised as input for the coding program used.	Hacking & Willenborg (2012)		Method: Manual coding
Real-time dataset	dataset showing how estimates change over time providing further information about the dissemination policy, the timing of revisions, the explanation of revision sources, the status of the published data	OECD (2006)	Revision triangle	Theme: Revisions of Economic Official Statistics
Recall	number of correctly linked record pairs divided by the total number of true match record pairs	Memobust definition (2014)	Sensitivity	Method: Fellegi-Sunter and Jaro Approach to Record Linkage
Recency effect	A given response alternative is more likely to be chosen when presented at the end rather than at the beginning of a list of response alternatives.	Memobust definition (2014)	Recency	Theme: Design of data collection (part 1) – Choosing the appropriate data collection method
Recipient file	File where one variable (say Z) is completely missing, and that will be imputed making use of the observed Z in the donor file	Memobust definition (2014)		Theme: Statistical Matching

Reconciliation	The series of a system must be reconciled in order to satisfy cross-sectional (contemporaneous) aggregation constraints (see aggregation above)	Dagum and Cholette (2006)	Data reconciliation	(1) Theme: Issues on Seasonal Adjustment; (2) Theme: Seasonal adjustment – introduction and general description
Record linkage	See: matching.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Reference period	The period of time or point in time to which the measured observation is intended to refer.	RAMON, Eurostat's metadata server - Statistical concept		(1) Theme: Statistical Registers and Frames – Main module; Theme: Statistical Registers and Frames – Quality of statistical registers and frames; (2) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (3) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (4) Theme: Statistical Registers and Frames – Survey frames for business surveys; (5) Theme: Statistical Registers and Frames – Statistical register and survey frame design
References	Step 11 in the OQRM model where Information is collected related to the focus area.	Van Nederpelt (2012)		Theme: Quality and Risk Management Models
Referential integrity	In a relational database, this is the basic principle that is required for internal consistency of the different tables in that database. This means that a table always has a key if it is referenced by another table in a key field, possibly a foreign key field. Database systems guarantee consistency and ensure that a transaction that violates the consistency cannot be performed.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Refusal rate	The proportion of observation units for which the reporting unit has been successfully contacted, but has refused to give the information sought.	Memobust definition (2014)		(1) Theme: Data Collection; (2) Theme: Data Collection: Techniques and Tools
Reg-ARIMA	In the seasonal adjustment context, a hybrid model in which some features of the time series, such as moving holiday, trading day and outlier effects, are modeled with linear regression variables while the remaining features (those of the regression residuals, including trend, cycle and seasonal components) are modelled with a seasonal ARIMA model	US Census Bureau		Method: Seasonal adjustment of economic time series

Register	A written and complete record containing regular entries of items and details on particular set of objects. Administrative registers come from administrative sources and become statistical registers after passing through statistical processing in order to make them fit for statistical purposes (production of register based statistics, frame creation, etc.).	Business Register Recommendations Manual (edition 2010), Glossary		(1) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (2) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (3) Theme: Statistical Registers and Frames – Survey frames for business surveys; (4) Theme: Statistical Registers and Frames – Statistical register and survey frame design; (5) Theme: Statistical Registers and Frames – The statistical units and the business register
Register	A set of files (paper, electronic, or a combination) containing the assigned data elements and the associated information.	SDMX (2009)		Theme: Sample selection
Register	A systematic collection of unit-level data organized in such a way that updating is possible. Updating is the processing of identifiable information with the purpose of establishing, bringing up to date, correcting or extending the register, i.e. keeping track of any changes in the data describing the units and their attributes. As a rule, a register will contain information on a complete group of units, a target population (e.g. persons, buildings, firms). These units are defined by a precise set of rules (for instance resident population in a country), and the attributes are updated in line with changes undergone by the units.	UN/ECE Glossary of Terms on Statistical Data Editing (2007)		Theme: Collection and Use of Secondary Data
Register unit	Register unit is the unit, entity of the register population with related descriptive information on identification, accessibility and other attributes. Remark: Register unit type – that is the collection of a given type of individual units – and register unit instance – that is a concrete, individual register unit – are distinguished. In the surveying process, data processing and dissemination phases, register units might function as data supplier, data provider or statistical (reporting, observation, analytical, dissemination) units.	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (2) Theme: Statistical Registers and Frames – Survey frames for business surveys; (3) Theme: Statistical Registers and Frames – Statistical register and survey frame design
Regression	A statistical technique for estimating the relationships among variables. In the univariate case only one explanatory variable is used. For the multivariate case, the number of explanatory variables equals two or more.	Memobust definition (2014)		Method: Chow-Lin Method for Temporal Disaggregation
Rejection region	antonym of Acceptance region	Memobust definition (2014)		Method: Manual Editing

Relevance	The degree to which statistical outputs meet current and potential user needs.	ESS Handbook for Quality Reports (2009)	Usability	(1) Theme: Quality of Statistics; (2) Theme: Overall Design
Relevance of log information	The degree to which log information is useful.	Memobust definition (2014)	Usability	Theme: Logging
Reliability	Closeness of the initial estimate to subsequent (revised) estimates	OECD (2006)		(1) Theme: Quality of Statistics; (2) Theme: Revisions of Economic Official Statistics; (3) Theme: Overall Design
Remote access	On-line access to protected microdata.	Memobust definition (2014)		Theme: Statistical Disclosure Control
Remote execution	Submitting scripts on-line for execution on disclosive microdata stored within an institute's protected network. If the results are regarded as safe data, they are sent to the submitter of the script. Otherwise, the submitter is informed that the request cannot be acquiesced. Remote execution may either work through submitting scripts for a particular statistical package such as SAS, SPSS or STATA which runs on the remote server or via a tailor made client system which sits on the user's desk top.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical Disclosure Control
Repeated survey	A survey which is carried out more than once, often regularly and often designed with some overlap over time between sampled units, taking both accuracy and response burden into account.	Memobust definition (2014)		Theme: Design of Estimation – Some Practical Issues
Repeated survey	A survey which is carried out more than once, often regularly and often designed with some overlap over time between sampled units, taking both accuracy and response burden into account.	Memobust definition (2014)		Theme: Repeated Surveys
Reporting unit	A unit that supplies the data for a given survey instance. The reporting unit is the unit about which data are reported. When, for a specific survey, the book keeping office completes questionnaires for each of the locations of a business, these locations are the reporting units.	Memobust definition (2014)		Theme: Data Collection
Reporting unit	The unit to which the questionnaire is tied and for which the questionnaire is filled in. It may be the observation unit, or it may be a means to reach the observation units.	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (2) Theme: Statistical Registers and Frames – Survey frames for business surveys; (3) Theme: Statistical Registers and Frames – Statistical register and survey frame design; (4) Theme: Data Collection: Techniques and Tools
representative outlier	represent other population units similar in value to the observed outliers	Memobust definition (2014)		Method: Outlier Treatment

Requirement	Step 3 in the OQRM model where the requirements for the focus area are formulated. Related: norm, standard, prescription, rule, principle, and indicator.	Van Nederpelt (2012)		Theme: Quality and Risk Management Models
Respondent	Respondents are businesses, authorities, individual persons, etc, from whom data and associated information are collected for use in compiling statistics.	Memobust definition (2014)		Theme: Data Collection
Respondent	The physical person at the data provider who answers the questionnaire.	Memobust definition (2014)		Theme: Data Collection: Techniques and Tools
Respondent	The physical person who answers the questionnaire. This is a person at the data provider.	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (2) Theme: Statistical Registers and Frames – Survey frames for business surveys
Respondent burden	Burden concerning behavioural and attitudinal attributes of respondents that affect the survey and cannot be changed by the supervisor or organiser of the survey. This concept also includes attitudes towards the survey itself such as the belief in the usefulness of surveys in general.	Hedlin et al. (2005)		Theme: Response Burden
Respondent burden	The effort, in terms of time and cost, required for respondents to provide satisfactory answers to a survey.	SDMX (2009)	Respondent/provider load	(1) Theme: Data Collection; (2) Theme: Data Collection: Techniques and Tools; (3) Theme: Sample selection; (4) Theme: Design of Estimation – Some Practical Issues
Response	The reaction of an individual unit to some form of stimulus. It may be to a drug, as in bioassay, or the reaction to a request for information, as in sample surveys of human beings.	A Dictionary of Statistical Terms, 5th edition, prepared for the International Statistical Institute by F.H.C. Marriott. Published for the International Statistical Institute by Longman Scientific and Technical.		Theme: Response Process
Response	In classical statistics we talk of response when each subject, or experimental units, gives rise to a single (case univariate) or vector (case multivariate) measurement on some relevant variables.	Memobust definition (2014)		Method: Little and Su Method

Response burden	The effort, in terms of time and cost, required for respondents to provide satisfactory answers to a survey.	SDMX (2009)	Statistical burden, Respondent burden	(1) Method: Sample co-ordination using simple random sampling with permanent random numbers; (2) Theme: Sample co-ordination; (3) Method: Assigning random numbers when co-ordination of surveys based on different unit types is considered; (4) Theme: Design of Estimation – Some Practical Issues; (5) Theme: Data Collection: Techniques and Tools; (6) Method: Balanced Sampling for Multi-Way Stratification; (7) Theme: Response Burden
Response process	The result of the interaction between a respondent and a questionnaire	Edwards W.S. & Cantor D. <i>Towards a Response Model in Establishment Surveys</i> In P. P. Biemer, et al., eds., <i>Measurement Error in Surveys</i> , New York: John Wiley & Sons, pp. 211-233		(1) Theme: Testing the Questionnaire; (2) Theme: Response Process
Response rate	The number of observation units for which data have been received, as a proportion of the number of observation units for which data was sought.	Memobust definition (2014)		(1) Theme: Data Collection; (2) Theme: Data Collection: Techniques and Tools
Response variable	A variable that is used to define the values in a table. The other kind of variable used to define a table is a spanning variable.	Memobust definition (2014)		Theme: Statistical disclosure control methods for quantitative tables
Responsibilities	Step 2 in the OQRM model, where the distribution of responsibilities of a focus area are determined. Context: There must be at least an owner of each focus area.	Van Nederpelt (2012)		Theme: Quality and Risk Management Models
Responsive design	There is more than one phase of the data collection and, according to the design, changes between phases are made, based on observed process data, typically indicators of quality and costs.	Memobust definition (2014)		(1) Theme: Data Collection; (2) Theme: Design of data collection (part 2) – Contact strategies; (3) Theme: Overall Design
Restricted (or Residual) Maximum Likelihood (REML)	Particular form of maximum likelihood estimation. It is based on maximizing a likelihood of transformed data not depending on nuisance parameters. In the case of estimation of variance components, the nuisance parameters are the regression coefficients.	Memobust definition (2014)		(1) Method: EBLUP Area Level for Small Area Estimation (Fay-Herriot); (2) Method: Small area estimation methods for time series data

Revision	This is a regular procedure in case of unadjusted (raw) data and seasonally adjusted data. Raw data may be revised due to improved information set (in terms of coverage and/or reliability). Revisions of seasonally adjusted data can also take place because of a better estimate of the seasonal pattern due to new information provided by new components. A revision shows the degree of closeness of an initial estimate to a subsequent or final estimate.	ESS Guidelines (2009), ESS Handbook on Quality Reports (2009)		(1) Theme: Overall Design; (2) Theme: Repeated Surveys; (3) Theme: Design of Estimation – Some Practical Issues; (4) Theme: Issues on Seasonal Adjustment
Revision	Difference between revised and preliminary estimate (Lt - Pt)	OECD (2006)		Theme: Revisions of Economic Official Statistics
Revision error	Difference between final and preliminary estimate	Memobust definition (2014)		(1) Theme: Weighting and Estimation; (2) Theme: Estimation with administrative data
Risk analysis	Step 4 in the OQRM model, where possible causes and possible effects with problems with a focus area are analysed. Example: Software errors cause problems with the accuracy of estimates.	Van Nederpelt (2012)		Theme: Quality and Risk Management Models
Risky cells	The cells of a table which are non-publishable due to the risk of statistical disclosure are referred to as risky cells. By definition there are three types of risky cells: small counts, dominance and complementary suppression cells.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Rotating panel	Limiting the length of time in which units stay in the survey panel by dropping a proportion of them after a certain period of time and replacing them with new ones. It is generally done only with the smaller respondents, for whom it is felt that responding to surveys imposes a significant burden. Rotation is designed to keep the sample up to date. It also helps to alleviate the problems caused by sample depletion.	Memobust definition (2014)		Method: Subsampling for Preliminary Estimates
Rotating panel survey	A panel survey where a portion (e.g. 25% of elements are replaced regularly.	Memobust definition (2014)		Theme: Weighting and Estimation
Rounding	Rounding belongs to the group of disclosure control methods based on output-perturbation. It is used to protect small counts in tabular data against disclosure. The basic idea behind this disclosure control method is to round each count up or down either deterministically or probabilistically to the nearest integer multiple of a rounding base. The additive nature of the table is generally destroyed by this process. Rounding can also serve as a recoding method for microdata.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Safe data	Microdata or macrodata that have been protected by suitable Statistical Disclosure Control methods.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical Disclosure Control
Safe setting	An environment such as a microdata lab whereby access to a disclosive dataset can be controlled.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical Disclosure Control

Safety interval	The minimal calculated interval that is required for the value of a cell that does not satisfy the primary suppression rule.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Sample	A subset of a frame where elements are selected based on a randomised process with a known probability of selection.	SDMX (2009)		Theme: Sample selection
Sample co-ordination	<i>From topic Sample selection, but long there now</i>	Memobust definition (2014)		Theme: Repeated Surveys
Sample size	The number of observation units which are to be included in the sample.	SDMX (2009)		(1) Method: Balanced Sampling for Multi-Way Stratification; (2) Method: Subsampling for Preliminary Estimates
Sample size dependent estimator	A sample size dependent estimator is a composite estimator with a subjectively chosen weight for the direct component which depends on true and estimated domain population sizes.	Memobust definition (2014)		Method: Composite Estimators for Small Area Estimation
Sample splitting	statistical method that splits the data into two halves, a regression model is performed on each statistically independent sub-sample	Memobust definition (2014)		Method: Outlier Treatment
Sampling	The process of selecting a number of cases from all the cases in a particular group or universe.	SDMX (2009)		Theme: Sample selection
Sampling design	Design that provides information on the target and final sample sizes, strata definitions and the sample selection methodology.	SDMX (2009)		(1) Method: Balanced Sampling for Multi-Way Stratification; (2) Method: Subsampling for Preliminary Estimates
Sampling error	An error caused by the fact that only a sample of values is observed and therefore there is a difference between a population value and an estimate.	Memobust definition (2014)		Method: Synthetic Estimators for Small Area Estimation
Sampling error	That part of the difference between a population value and an estimate thereof, derived from a random sample, which is due to the fact that only a sample of values is observed; as distinct from errors due to imperfect selection, bias in response or estimation, errors of observation and recording, etc. The totality of sampling errors in all possible samples of the same size generates the sampling distribution of the statistic which is being used to estimate the parent value.	The International Statistical Institute, The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003.		(1) Theme: Quality and Risk Management Models; (2) Theme: Quality of Statistics; (3) Theme: Editing During Data Collection
Sampling fraction	The ratio of the sample size to the population size.	SDMX (2009)		Method: Balanced Sampling for Multi-Way Stratification

Sampling frame	A list, map or other specification of the units which define a population to be completely enumerated or sampled.	SDMX (2009)		(1) Theme: Statistical Registers and Frames – Survey frames for business surveys; (2) Theme: Statistical Registers and Frames – Statistical register and survey frame design; (3) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (4) Method: Balanced Sampling for Multi-Way Stratification
Sampling strategy	Sampling design and estimation methodology	Memobust definition (2014)		Theme: Weighting and Estimation
Satellite register	Satellite register records a given subpopulation of the business register and fulfill the following conditions: - They are not an integral part of the statistical business register as referred to in the business registers Regulation, but are capable of being linked to it. - They are more limited in scope than the statistical business register, e.g. in terms of NACE, but within that scope they may have more extensive coverage of units and/or variables. - They contain one or more variables that are not found in the statistical business register. Such variables are generally capable of being used for stratification purposes.	Business Register Recommendations Manual (edition 2010), paragraph 20.40 - modified	Associated register	Theme: Statistical Registers and Frames – The statistical units and the business register
SBS	Structural Business Statistics. SBS are statistical surveys covering industry, construction, trade and services. They are conducted in each Member State of the European Union (UE) in order to describe the structure, conduct and performance of businesses across the EU.	Council Regulation (EC, Euratom) No 58/97 of 20 December 1996 concerning structural business statistics amended by Council Regulation (EC, Euratom) No 410/98 of 16 February 1998		Theme: Different types of surveys
Scheduling of interviewers	Production of a planning which indicates which interviewers work on what days and part of day (DPoD) combinations in the planning period.	Memobust definition (2014)		(1) Theme: CATI Allocation; (2) Theme: Data Collection
Scheduling system	IN CATI surveys is the IT module to manage telephone contacts.	Memobust definition (2014)		Theme: Data Collection: Techniques and Tools
SCM	Standard Cost Model – an international method model aimed at reducing administrative burdens in the business environment by adopting a policy based on costs of regulations	ISCM (2003)		Theme: Response Burden

Scope of data suppliers	Scope of the data supplier is the set of entities of the frame population assigned for data reporting from which data can be retrieved for the investigated population (statistical and observation units). Remark: In full scope data collection, the scope of data suppliers corresponds to the frame population. In representative or combined data collection, the scope of data suppliers is only a part of the frame population. It doesn't contain the statistical units of the frame population not selected into the sample.	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (2) Theme: Statistical Registers and Frames – Survey frames for business surveys
SDC	See: Statistical Disclosure Control	Memobust definition (2014)	Statistical Disclosure Control	(1) Theme: Statistical Disclosure Control; (2) Theme: Statistical disclosure control methods for quantitative tables
Seasonal adjustment	Seasonal adjustment is a statistical technique to remove the effects of seasonal calendar influence operating on series	OECD (2006)	SA	(1) Method: Seasonal adjustment of economic time series; (2) Theme: Issues on Seasonal Adjustment; (3) Theme: Seasonal adjustment – introduction and general description
Seasonal adjustment software	There is a wide range of software and interfaces available to perform seasonal adjustment. For official statistics, the two most commonly used seasonal adjustment methods are X-12-ARIMA (US Census Bureau) and TRAMO-SEATS (Bank of Spain). Recently, Eurostat has released a new software (in which both X-12-ARIMA and TRAMO-SEATS are available), called DEMETRA+	Memobust definition (2014)		Method: Seasonal adjustment of economic time series
Seasonal component	A time series whose values quantify (usually in percents or in the units of data measurement, e.g. dollars) variations in the level of the observed series that recur with the same direction and a similar magnitude at time intervals of length one year. (Length is measured in the calendar units of the observed series--usually quarters or months, sometimes semesters, weeks, or other units.)	US Census Bureau	Seasonality	(1) Method: Seasonal adjustment of economic time series; (2) Theme: Seasonal adjustment – introduction and general description
Secondary data	Data that is collected by others (i.e. not the NSI), used by an NSI for producing statistics and where the NSI has not defined the conceptual or process metadata	Daas and Arends-Toth (2012)		Theme: Collection and Use of Secondary Data
Secondary data collection	The acquisition of secondary data by an NSI	Daas and Arends-Toth (2012)		Theme: Collection and Use of Secondary Data
Secondary Key	See Object characteristic	Memobust definition (2014)		
Secondary research	Research that uses secondary sources	Golden (1976)		Theme: Collection and Use of Secondary Data
Secondary source	A source containing secondary data	Golden (1976)		Theme: Collection and Use of Secondary Data

Secondary suppression	To reach the desired protection for risky cells, it is necessary to suppress additional non- risky cells, which is called secondary suppression or complementary suppression. The pattern of complementary suppressed cells has to be carefully chosen to provide the desired level of ambiguity for the disclosive cells at the highest level of information contained in the released statistics.	Glossary on Statistical Disclosure Control (2014)	Complementary suppression	Theme: Statistical disclosure control methods for quantitative tables
Segmentation effect	It is a characteristics typical of electronic questionnaire and consists in the display of one question per screen thus restricting the view of the questionnaire	Memobust definition (2014)		Theme: Data Collection: Techniques and Tools
Selective editing	An umbrella term for methods that select records which are likely to contain influential errors for interactive editing, on a record-by-record basis.	CBS Methods Series Glossary	Micro-selection, significance editing	(1) Theme: Editing for Longitudinal Data; (2) Theme: Selective Editing
Selective editing	A procedure which targets only some of the micro data items or records for review by prioritizing the manual work and establishing appropriate and efficient process and edit boundaries.	UN/ECE Glossary of Terms on Statistical Data Editing (2007)	Micro-selection	(1) Method: Automatic Editing; (2) Theme: Editing Administrative Data; (3) Theme: Macro-Editing; (4) Theme: Statistical Data Editing
Self-administered mode	The questions in the survey are administered and answered by the respondent without any assistance or help from an interviewer.	Memobust definition (2014)		(1) Theme: Data Collection; (2) Theme: Design of data collection (part 1) – Choosing the appropriate data collection method
Self-Administered Questionnaire	A questionnaire used in Paper and Pencil interviewing.	Memobust definition (2014)		Theme: Mixed Mode Data Collection – design issues
Semantic network	A network(or grph) consisting of words and concepts and semantic relationships between them. Examples of such relationships are synonyms, hypernyms and hyponyms.	Hacking & Willenborg (2012)		Method: Computer-assisted coding
Semi-automatic coding	Synonymous with computer-supported coding and computer-assisted coding.	Hacking & Willenborg (2012)		(1) Method: Automatic coding based on pre-coded datasets; (2) Method: Computer-assisted coding; (3) Theme: Coding
Sensitivity	number of correctly linked record pairs divided by the total number of true match record pairs. Sensitivity measures the percentage of correctly classified record matches,	Memobust definition (2014)	Recall	Method: Fellegi-Sunter and Jaro Approach to Record Linkage
Sequential mixed mode	Using different modes one after another, maximizing the use of one mode before switching to another.	Memobust definition (2014)		(1) Theme: Data Collection; (2) Theme: Design of data collection (part 2) – Contact strategies
Shrinkage factor	The parameter used in composite estimator formulas to decide about the contribution of the direct and synthetic estimators.	Memobust definition (2014)		Method: Composite Estimators for Small Area Estimation
SIC	Standard Industrial Classification, a classification of industries by type of economic activity created and maintained by the Department of Labour, United States of America and used also e.g. in UK.	US Department of Labour		Theme: Response Burden

Signalling measure	Measure to detect a quality problem.	Memobust definition (2014)		Theme: Quality and Risk Management Models
Significance editing	Synonym of Selective editing.	Memobust definition (2014)	Selective editing	Theme: Selective Editing
Similarity measure	A measure that indicates the extent to which two units are similar. This type of measure (or its complement: the dissimilarity measure) is also used in the multivariate analysis.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Simple random sampling	A sampling design in which the inclusion probability of each unit of the population is given by the sampling fraction.	SDMX (2009)		(1) Method: Balanced Sampling for Multi-Way Stratification; (2) Method: Subsampling for Preliminary Estimates
Single activity business	A business operating in only one economic activity	Memobust definition (2014)		Method: Assigning random numbers when co-ordination of surveys based on different unit types is considered
Single location business	A business operating from only one geographical location	Memobust definition (2014)		Method: Assigning random numbers when co-ordination of surveys based on different unit types is considered
Skewness	Measure of the asymmetry of a distribution.	Memobust definition (2014)		(1) Method: EBLUP Area Level for Small Area Estimation (Fay-Herriot); (2) Theme: Weighting and Estimation
Small outlier	the Y values are extremely smaller than the other Y values of the “normal” units	Memobust definition (2014)		Method: Outlier Treatment
Smith-Waterman Distance	Distance that uses dynamic programming to find the minimum cost to convert one string into the corresponding string of the compared record; the parameters of this algorithm are the insertions cost, deletions cost and transposition cost	Memobust definition (2014)		Theme: Probabilistic Record Linkage
Snapshot of register	Snapshot of a register is its frozen state on a given date. Remark: Instead of a register, snapshots are used for statistical processing because, unlike register units (that can be updated frequently), population units and their attributes must be unchanged during the data collection and statistical processing.	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – Main module; Theme: Statistical Registers and Frames – Quality of statistical registers and frames; (2) Theme: Statistical Registers and Frames – Survey frames for business surveys; (3) Theme: Statistical Registers and Frames – Statistical register and survey frame design
Social desirability bias	Systematic underreporting of something to “fit in” in what the respondent thinks is “normal” or accepted in society. For instance, alcohol consumption is often underreported to avoid embarrassment.	Memobust definition (2014)	Social desirable answers	Theme: Design of data collection (part 1) – Choosing the appropriate data collection method
Soft constraint	A constraint that does not have to hold exactly, but approximately.	Memobust definition (2014)		Method: Denton for Benchmarking

Soft edit rule	An edit rule whose failure indicates an error with probability less than 1.	EDIMBUS Manual	Query edit rule	(1) Method: Automatic Editing; (2) Method: Manual Editing; (3) Theme: Statistical Data Editing
Soundex	Indexing technique based on the sound (or pronunciation) of words (and not how they are written), originally only for English, but later developed for Dutch as well.	Hacking & Willenborg (2012)		(1) Method: Automatic coding based on pre-coded datasets; (2) Method: Computer-assisted coding
Soundex algorithm	Originally a phonetic algorithm to index names based on sound (in English). Later, a similar algorithm was developed for words in the Dutch language. Improvements of the Soundex algorithm for English include Metaphone and Double Metaphone.	Memobust definition (2014)		(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Soundness of methodology	The extent to which the methodology used to compile statistics complies with the relevant international standards, including the professional standards enshrined in the Fundamental Principles for Official Statistics.	SDMX (2009)		(1) Theme: Methods and Quality; (2) Theme: Quality of Statistics
Source	A specific data set, metadata set, database or metadata repository from where data or metadata are available.	SDMX (2009)		Theme: Collection and Use of Secondary Data
Source Data	Characteristics and components of the raw statistical data used for compiling statistical aggregates.	SDMX (2009)		(1) Method: RAS; (2) Method: Stone
Spanning variable	A variable that is used to define the rows, columns etc. of a table. The other kind of variable used to define a table is a response variable.	Memobust definition (2014)		Theme: Statistical disclosure control methods for quantitative tables
Specificity	number of correctly unlinked record pairs divided by the total number of true non-match record pairs. Specificity measures the percentage of correctly classified non-matches.	Memobust definition (2014)		Method: Fellegi-Sunter and Jaro Approach to Record Linkage
Split-off	This event is similar to a break-up, but in this case the original enterprise does survive in a recognisable form, and therefore there is both continuity and survival. There is no death, but one or more new enterprises are created.	Eurostat-OECD Manual on Business Demography Statistics (chapter 4).		Theme: Business Demography
Spreading activation	Method to search in a semantic network.	Hacking & Willenborg (2012)		Method: Computer-assisted coding
State space model	A time-series model that predicts the future state of a system from its previous states probabilistically, via a process model. The state space models describes mathematically how observations of the state of the system are generated via an observation model.	Memobust definition (2014)		(1) Method: Preliminary estimates with model-based methods; (2) Method: Small area estimation methods for time series data
Statistical burden	The burden of the sampled unit to respond to the survey questionnaire.	Memobust definition (2014)		Method: Balanced Sampling for Multi-Way Stratification
Statistical data	Data that are collected and/or generated by statistics in process of statistical observations or statistical data processing.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics

Statistical data collection	Statistical data collection is the operation of statistical data processing aimed at gathering of statistical data and producing the input object data of a statistical survey.	Terminology on Statistical Metadata, Conference of European Statisticians Statistical Standards and Studies, No. 53, UNECE, Geneva 2000,		Theme: Testing the Questionnaire
Statistical data editing	The process of editing a data file for statistical purposes.	Memobust definition (2014)		Theme: Statistical Data Editing
Statistical disclosure control	Statistical Disclosure Control techniques can be defined as the set of methods to reduce the risk of disclosing information on individuals, businesses or other organisations. Such methods are only related to the dissemination step and are usually based on restricting the amount of or modifying the data released.	Glossary on Statistical Disclosure Control (2014)	Statistical disclosure limitation; SDC;SDL	(1) Theme: Statistical Disclosure Control; (2) Theme: Statistical disclosure control methods for quantitative tables
Statistical edit	A statistical edit is a set of checks based on statistical analysis of respondent data, e.g., the ratio of two fields lies between limits determined by a statistical analysis of that ratio for presumed valid reporters. A statistical edit may incorporate cross-record checks, e.g., the comparison of the value of an item in one record against a frequency distribution for that item for all records. A statistical edit may also use historical data on a firm-by-firm basis in a time series modeling procedure.	Glossary of Terms Used in Statistical Data Editing Located on K-Base, the knowledge base on statistical data editing, UNECE Data Editing Group		Theme: Editing During Data Collection
Statistical matching	Matching records with information from units which do not necessarily have to be the same, but are similar. In terms of intention, this method deals with an entirely different problem than is discussed in this report. This is actually an imputation method. This method is not further discussed in this report for this reason.	Memobust definition (2014)	Synthetic matching	(1) Theme: Object matching; (2) Method: Object Identifier Matching; (3) Method: Unweighted Matching; (4) Method: Weighted Matching
Statistical measure	A summary (means, mode, total, index, etc.) of the individual quantitative variable values for the statistical units in a specific group (study domains).	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Statistical output	Results from a statistical process to be accessed by the final users. Context: Can take the form of aggregate statistics, analysis, and microdata releases and can include different forms of media.	NQAF (2012)		Theme: Methods and Quality
Statistical output	Results from a statistical process to be accessed by the final users..	NQAF (2012)		Theme: Quality of Statistics
Statistical register	Statistical register is a continuously updated set of objects for a given population containing information on identification, accessibility of population units and other attributes, supporting the surveying process of the population. The register contains the current and historical statuses of the population and the causes, effects and sources of alterations in the population. Register data of population units are stored in a structured database	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – Main module; Theme: Statistical Registers and Frames – Quality of statistical registers and frames; (2) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys

Statistical register	A regularly updated list of units and their characteristics to be used for statistical purposes.	SDMX (2009)		Theme: Collection and Use of Secondary Data
Statistical source	A source containing information collected and maintained for statistical purposes. It contains statistical units and statistical variables	Parallel definition to administrative source		Theme: Collection and Use of Secondary Data
Statistical unit	Statistical units are defined on the basis of three criteria: Legal, accounting or organizational criteria; Geographical criteria; Activity criteria. The relationship between different types of statistical units can be summarized in the following way: Units with one or more activities and one or more locations; Enterprise; Institutional unit; Units with one or more activities and a single location; Local unit; Units with one single activity and one or more locations; KAU; UHP; Units with one single activity and one single location; Local KAU; Local UHP. The Council Regulation (EEC), No 696/93 of 15 March 1993 on statistical units for the observation and analysis of the production system in the Community lays down a list of eight (types of) statistical units: The enterprise; The institutional unit; The enterprise group; The kind-of-activity unit (KAU); The unit of homogeneous production (UHP); The local unit; The local kind-of-activity unit (local KAU); The local unit of homogeneous production (local UHP).	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, Annex Section.		(1) Theme: Statistical Registers and Frames – The statistical units and the business register; (2) Theme: Derivation of Statistical Units
Step problem	The phenomenon of a large gap between the last sub annual period of one annual period and the first sub annual period of the next annual period. (for instance: a large gap between December and Januar). Annual and sub annual are used in a broad sense here. It can be any combination of two periods with a difference frequency, such that one annual period covers a whole number of sub annual periods.	Memobust definition (2014)		Method: Denton for Benchmarking
Stochastic imputation	In stochastic imputation the imputed value contains a random component. Repetition of the imputation leads to a different result.	EDIMBUS Manual		(1) Theme: Donor Imputation; (2) Theme: Imputation; (3) Theme: Imputation for Longitudinal Data; (4) Theme: Model-based Imputation
Stochastic regression imputation	Model based imputation method: : imputes the missing value with a value obtained as the sum of the predicted value by the regression model being considered and a random error term	Memobust definition (2014)		Method: Statistical Matching Methods
Stock Variable	A stock variable is measured at one specific time, and represents a quantity existing at that point in time. See also flow variable	Memobust definition (2014)		Theme: Macro Integration
Stop word	Word in a description that does not contain any information or contains too little information, because it occurs too frequently. A stop word can therefore be deleted by an automatic coding system.	Hacking & Willenborg (2012)		Theme: Coding

Stratification	A sampling procedure in which the population is divided into homogeneous subgroups or strata and the selection of samples is done independently in each stratum.	SDMX (2009)		(1) Method: Sample co-ordination using simple random sampling with permanent random numbers; (2) Theme: Sample selection
Stratified simple random sampling	A sampling design in which the population is divided into homogeneous subgroups or strata and the selection of samples is done independently in each stratum.	SDMX (2009)		(1) Method: Balanced Sampling for Multi-Way Stratification; (2) Method: Subsampling for Preliminary Estimates
Structural zero (cell)	A zero in a table cell corresponding to a situation where there can be no population elements, because this is impossible, on logical or as a matter of fact or principle. (For instance: a pregnant man.)	Memobust definition (2014)		Theme: Statistical disclosure control methods for quantitative tables
STS	Short-Term Statistics. STS are statistical surveys conducted in each Member State of the UE with a monthly or quarterly frequency. Output data (indicators) deliver information on supply, demand, factors of production and prices in four main domains: industry, construction, retail trade and other services.	Regulation EC No 1165/98, amended by Regulation EC No 1158/2005 and the regulations implementing and amending these two instruments		(1) Theme: Different types of surveys; (2) Theme: The European Statistical System; (3) Theme: Estimation with administrative data
Study domains	A segment of the population for which separate statistics are needed.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Subadditivity	One of the properties of the (n,k) rule or (p,q) rule that assists in the search for complementary cells. The property means that the sensitivity of a union of disjoint cells cannot be greater than the sum of the cells' individual sensitivities (triangle inequality). Subadditivity is an important property because it means that aggregates of cells that are not sensitive are not sensitive either and do not need to be tested.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Subpopulation	Subpopulation is a subset of a population. Remark: Subpopulation refers to populations that require different handling in the statistical working process. Subpopulations are usually specified to understand the distinguishing characteristics of these populations	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (2) Theme: Statistical Registers and Frames – Survey frames for business surveys; (3) Theme: Statistical Registers and Frames – Statistical register and survey frame design
Supplier manager	Person responsible to acquire and manage products and or resources to needed to run a business	Memobust definition (2014)		Theme: Collection and Use of Secondary Data
Supply use tables	An accounting framework in which supply and use of goods and services and the generation of value added is described, detailed to commodities and industries. It is the ideal framework for making estimates of gross domestic product (GDP)	Memobust definition (2014)		Theme: Manual Integration

Survey	Survey is an investigation on the characteristics of a given population by means of collecting data and estimating their characteristics through the systematic use of statistical methodology. Remark: Included are: - censuses, which attempt to collect data from all members of a population; - sample surveys, in which data are collected from a (usually random) sample of population members. Surveys can be unique in time or repeated with regular or irregular periodicity. A single wave of a repeated survey is called survey instance. A wider definition under which the term survey covers any activity that collects or acquires statistical data (including censuses, sample surveys, the collection of data from administrative records and derived statistical activities) has also been proposed. (see Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 7, available at <a href="http://www.statcan.ca:8096/bsolc/english/bsolc?catno=12-539-X&amp;CHROPG=1">http://www.statcan.ca:8096/bsolc/english/bsolc?catno=12-539-X&amp;CHROPG=1</a> ).	RAMON, Eurostat's metadata server – Statistical concepts		Theme: Statistical Registers and Frames – The populations, frames and units of business surveys
Survey	A investigation about the characteristics of a given population by means of collecting data from a sample of that population and estimating their characteristics through the systematic use of statistical methodology.	SDMX (2009)		Theme: Sample selection
Survey (1)	Survey is an investigation on the characteristics of a given population by means of collecting data and estimating their characteristics through the systematic use of statistical methodology. Remark: Included are: censuses, which attempt to collect data from all members of a population; - sample surveys, in which data are collected from a (usually random) sample of population members. Surveys can be unique in time or repeated with regular or irregular periodicity. A single wave of a repeated survey is called survey instance.	RAMON, Eurostat's metadata server – Statistical concepts		Theme: Statistical Registers and Frames – Main module; Theme: Statistical Registers and Frames – Quality of statistical registers and frames
Survey (2)	A wider definition under which the term survey covers any activity that collects or acquires statistical data (including censuses, sample surveys, the collection of data from administrative records and derived statistical activities) has also been proposed. (see Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 7, available at <a href="http://www.statcan.ca:8096/bsolc/english/bsolc?catno=12-539-X&amp;CHROPG=1">http://www.statcan.ca:8096/bsolc/english/bsolc?catno=12-539-X&amp;CHROPG=1</a> )	Memobust definition (2014)		Theme: Statistical Registers and Frames – Main module; Theme: Statistical Registers and Frames – Quality of statistical registers and frames

Survey feedback	Information obtained from a survey used to update the Register	Memobust definition (2014)		(1) Theme: Sample co-ordination; (2) Theme: Repeated Surveys
Survey frame	Survey frame is the set of survey population units together with their attributes referring to a given reference period. The frame contains the identification, contact, classification attributes of the frame units for a given reference period.	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (2) Theme: Statistical Registers and Frames – Survey frames for business surveys; (3) Theme: Statistical Registers and Frames – Statistical register and survey frame design; (4) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (5) Theme: Statistical Registers and Frames – Main module; Theme: Statistical Registers and Frames – Quality of statistical registers and frames
Survey instance	Survey instance is a particular survey and reference period in which data are collected from respondents	RAMON, Eurostat's metadata server - UN metadata terminology		(1) Theme: Statistical Registers and Frames – Main module; Theme: Statistical Registers and Frames – Quality of statistical registers and frames; (2) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (3) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (4) Theme: Statistical Registers and Frames – Survey frames for business surveys
Survey population	Survey population is the population for which information during the survey process can be obtained. Remark: Concurrence or difference of survey and target populations is measured by coverage	Handbook on the design and implementation of business surveys		(1) Theme: Asymmetry in Statistics – European Register for Multinationals (EGR); (2) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys
Surveying department	The surveying department is the unit of the statistical office that responsible for the data collection phase of the survey	Memobust definition (2014)		(1) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (2) Theme: Statistical Registers and Frames – Survey frames for business surveys

Survival rate	The survival rate of newly born enterprises in a given reference period is the number of enterprises that were born in year t-i (i=1,...,n) and survived to year t as a percentage of all enterprises born in year t-i.	Memobust definition (2014)		Theme: Business Demography
Surviving enterprise	In the Business Demography context, survival occurs if an enterprise is active in terms of employment and/or turnover in the year of birth and the following year(s). Two types of survival can be distinguished: 1) An enterprise born in year t-1 is considered to have survived in year t if it is active in terms of turnover and/or employment in any part of year t (= survival without changes). 2) An enterprise is also considered to have survived if the linked legal unit(s) have ceased to be active, but their activity has been taken over by a new legal unit set up specifically to take over the factors of production of that enterprise (= survival by take-over).	Eurostat-OECD Manual on Business Demography Statistics		Theme: Business Demography
Swapping (or switching)	Swapping (or switching) involves selecting a sample of the records, finding a match in the data base on a set of predetermined variables and swapping all or some of the other variables between the matched records. Swapping (or switching) was illustrated as part of the confidentiality edit for tables of frequency data.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Synonym	Word or concept with the same meaning as another word, possibly in a special context.	Hacking & Willenborg (2012)		(1) Method: Computer-assisted coding; (2) Theme: Coding
Synthetic estimator	An indirect estimator based on the assumption that small areas have the same characteristics as a large area and a reliable direct estimator for the large area is used in the estimation process for small areas.	Memobust definition (2014)		Method: Synthetic Estimators for Small Area Estimation
Synthetic matching	See: Statistical matching	Memobust definition (2014)		
Systematic error	(1) An error reported consistently over time and/or between responding units. Or (2) a type of error for which the error mechanism and the imputation procedure are known.	UN/ECE Glossary of Terms on Statistical Data Editing, EDIMBUS Manual.		(1) Method: Automatic Editing; (2) Method: Deductive Editing; (3) Theme: Editing Administrative Data; (4) Theme: Statistical Data Editing
Systematic error	The systematic deviation of the estimate from the true value.	Van Nederpelt (2009)	Bias, purity	Theme: Quality of Statistics
Table	A special form of aggregate data, where the information is divided into cells, each corresponding to a group of individual entities	Memobust definition (2014)		Theme: Statistical Disclosure Control
Table redesign	See: table restructuring	Memobust definition (2014)	Table restructuring	Theme: Statistical disclosure control methods for quantitative tables
Table restructuring	A technique to produce safe tables by combining rows or columns	Memobust definition (2014)	Table redesign	Theme: Statistical disclosure control methods for quantitative tables

Tables of frequency (count) data	These tables present the number of units of analysis in a cell. When data are from a sample, the cells may contain weighted counts, where weights are used to bring sample results to the population levels. Frequencies may also be represented as percentages.	Glossary on Statistical Disclosure Control (2014)	Frequency tables	(1) Theme: Statistical Disclosure Control; (2) Theme: Statistical disclosure control methods for quantitative tables
Tables of magnitude data	Tables of magnitude data present the aggregate of a "quantity of interest" over all units of analysis in the cell. When data are from a sample, the cells may contain weighted aggregates, where quantities are multiplied by units' weights to bring sample results up to population levels. The data may be presented as averages by dividing the aggregates by the number of units in their cells.	Glossary on Statistical Disclosure Control (2014)	Quantitative tables	(1) Theme: Statistical Disclosure Control; (2) Theme: Statistical disclosure control methods for quantitative tables
Tabular data	Aggregate information on entities presented in tables.	Glossary on Statistical Disclosure Control (2014)	Macrodata	Theme: Statistical disclosure control methods for quantitative tables
Take-over	This event can be seen as the opposite of a split-off. Enterprises taken over are not considered to be deaths. In this case, one of the original enterprise does survive in a recognisable form, and therefore there is both continuity and survival. The remaining original enterprises are closed.	Eurostat-OECD Manual on Business Demography Statistics (chapter 4).		Theme: Business Demography
Target population	Target population is the set of units about which information is wanted and estimates are required. Remark: We differentiate the ideal and the intended target population. The ideal target population is the user demand, the intended target population is the realisable population of the survey.	CODED – Statistical concept - modified		(1) Theme: Statistical Registers and Frames – Main module; Theme: Statistical Registers and Frames – Quality of statistical registers and frames; (2) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (3) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (4) Theme: Statistical Registers and Frames – Survey frames for business surveys; (5) Theme: Statistical Registers and Frames – Statistical register and survey frame design; (6) Theme: Asymmetry in Statistics – European Register for Multinationals (EGR)
Target variable	A variable that is observed or derived and that measures an aspect of a phenomenon of interest during a survey; a goal of the survey will be to estimate population parameters for such a variable.	CBS Methods Series Glossary		(1) Theme: Donor Imputation; (2) Theme: Imputation; (3) Theme: Imputation for Longitudinal Data; (4) Theme: Model-based Imputation

t-ARGUS	t-Argus is a specialized software tool for the protection of tabular data. t-Argus is used to produce safe tables. t-Argus uses the same two main techniques as μ-Argus: global recoding and local suppression. For t-Argus the latter consists of suppression of cells in a table.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical Disclosure Control
TDE	Telephone/Touchtone Data Entry is a data entry mode in which a telephone is used by the respondent to communicate his/her answers. It is a form of self-administered telephone survey that does not require interviewer assistance.	Memobust definition (2014)		Theme: Mixed Mode Data Collection – design issues
Temporal constraint	Constraints in the same time-series for different periods	Memobust definition (2014)		Method: Denton for Benchmarking
Temporal Disaggregation	Deriving sub annual data (for instance quarterly data) from annual data, by using indicators of the sub annual data (i.e. related time series), see disaggregation. Annual and sub annual are used in a broad sense here. It can be any combination of two periods with a difference frequency, such that one annual period covers a whole number of sub annual periods.	SDMX (2009)		(1) Method: Chow-Lin Method for Temporal Disaggregation; (2) Method: Denton for Benchmarking; (3) Theme: Macro Integration
Test variable	A component of an edit rule that defines, for a given edit group, the expression (in terms of one or more observed variables) that is to be evaluated with respect to the acceptance regions for edit groups.	Norberg (2011)		(1) Method: Manual Editing; (2) Theme: Editing for Longitudinal Data
TF-IDF Distance	Distance that is used to match strings in a document. It assigns high weights to frequent tokens in the document and low weights to tokens that are also frequent in other documents	Memobust definition (2014)		Theme: Probabilistic Record Linkage
Threshold rule	Usually, with the threshold rule, a cell in a table of frequencies is defined to be sensitive if the number of respondents is less than some specified number. Some agencies require at least five respondents in a cell, others require three. When thresholds are not respected, an agency may restructure tables and combine categories or use cell suppression, rounding or the confidentiality edit, or provide other additional protection in order to satisfy the rule.	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Time series	A set of ordered observations on a quantitative characteristic of an individual or collective phenomenon taken at different points of time.	SDMX (2009)		(1) Method: Chow-Lin Method for Temporal Disaggregation; (2) Method: Denton for Benchmarking; (3) Theme: Macro Integration
Time series	A sequence of measurements of an economic (or other) variable made at approximately equally spaced times. It is important that the definition of the variable and the method used to measure it be consistent over time	US Census Bureau		Method: Seasonal adjustment of economic time series

Timeliness	The length of time between the event or phenomenon the statistical outputs describe and their availability.	ESS Handbook for Quality Reports (2009)		(1) Theme: Quality of Statistics; (2) Theme: Overall Design
Timeliness	The lapse of time between the end of a reference period and availability of the data.	SDMX (2009)		(1) Method: Subsampling for Preliminary Estimates; (2) Method: Preliminary estimates with design-based methods
Top-of-the-head responses	The respondent is feeling stressed and pressured to give an quick answer and therefore picks the first response category presented to them.	Memobust definition (2014)		Theme: Design of data collection (part 1) – Choosing the appropriate data collection method
Total survey error	The accumulation of all errors that may arise in the design, collection, processing, and analysis of survey data.	Biemer (2010)		Theme: Quality of Statistics
Training set	A corpus where the codes linked to the descriptions are verified. The codes originate from a classification. A training set is used in the coding methods that are based on supervised classification.	Hacking & Willenborg (2012)		(1) Method: Automatic coding based on pre-coded datasets; (2) Method: Computer-assisted coding
Transversal sampling design	Sampling design of one of the surveys, at one sampling occasion.	Memobust definition (2014)		Method: Sample co-ordination using Poisson sampling with permanent random numbers
Trend-cycle	The trend is the underlying long-term movement lasting many years. The cycle, also called business-cycle, is a quasi-periodic oscillation lasting for more than a year around the long-term trend. It is characterized by alternating periods of expansion and contraction. The trend and the cycle are difficult to estimate separately and thus are considered and analysed as a whole as the trend-cycle	Statistics Canada (2009)	TC	(1) Method: Seasonal adjustment of economic time series; (2) Theme: Seasonal adjustment – introduction and general description
Trigram	String consisting of three consecutive characters. They are used in fuzzy string matching. The more trigrams two strings have in common, compared to the trigrams they have not in common, the more similar they are.	Hacking & Willenborg (2012)		(1) Method: Automatic coding based on pre-coded datasets; (2) Method: Automatic coding based on semantic networks; (3) Method: Computer-assisted coding
Trimmed least absolute value	robust statistical method that attempts to minimise the sum of absolute deviation (residuals) over a subset of $k$ points which yields the lowest sum of absolute residuals ( $k < n$ )	Memobust definition (2014)		Method: Outlier Treatment
Trimmed least square	robust statistical method that attempts to minimise the sum of squared residuals over a subset, $k$ points which yields the lowest sum of squared residuals ( $k < n$ )	Memobust definition (2014)		Method: Outlier Treatment
True value	The actual population value that would be obtained with perfect measuring instruments and without committing any error of any type, both in collecting the primary data and in carrying out mathematical operations.	Eurostat's Concepts and Definitions Database (2013)		Theme: Quality of Statistics
Type I error	See: Mismatch	Memobust definition (2014)		
Type II error	See: Missed match	Memobust definition (2014)		

Unbiased	Estimator whose bias is zero.	Memobust definition (2014)		Method: Generalised regression estimator
Unbiasedness	An estimator is said to be unbiased if the bias (difference between its mathematical expectation and the true value it estimates) is zero.	The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press (2003).		Theme: Small area estimation
Under-coverage	There are target population units that are not accessible via the frame.	ESS Handbook on Quality Reports (2009)		Theme: Design of Estimation – Some Practical Issues
Under-coverage	Under-coverage results from the omission from the frame of units belonging to the target population.	OECD Glossary		Theme: Weighting and Estimation
Under-coverage	Failure to include required units in the frame, which results in the absence of information for those units.	SDMX (2009)		(1) Theme: Quality of Statistics; (2) Theme: Sample selection
Unequal probability sampling	A sampling design in which the inclusion probability may be different for each unit of the population.	SDMX (2009)		(1) Method: Balanced Sampling for Multi-Way Stratification; (2) Method: Subsampling for Preliminary Estimates
Unit	Units refer to entities, respondents to a survey or things used for the purpose of calculation or measurement. Their statistics are collected, tabulated and published. They include, among others, businesses, government institutions, individual organisations, institutions, persons, groups, geographical areas and events. They form the population from which data can be collected or upon which observations can be made. Remark: In this handbook chapter the unit can belong to the population, frame, register. The type of unit can be statistical unit, collection unit, reporting unit, observation unit, analytical unit, legal unit.	RAMON, Eurostat's metadata server		Theme: Statistical Registers and Frames – The populations, frames and units of business surveys
Unit non-response	The event that no data are obtained from a unit that was supposed to be observed.	CBS Methods Series Glossary		(1) Theme: Imputation; (2) Theme: Imputation for Longitudinal Data; (3) Theme: Quality of Statistics; (4) Theme: Data Collection: Techniques and Tools

Unit of homogeneous production	The unit of homogeneous production (UHP) is characterised by a single activity which is identified by its homogeneous inputs, production process and outputs. The products which constitute the inputs and outputs are themselves distinguished by their physical characteristics and the extent to which they have been processed as well (as) by the production technique used, by reference to a product classification. The unit of homogeneous production may correspond to an institutional unit or a part thereof; on the other hand, it can never belong to two different institutional units	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, Annex Section III E; SBS Regulation No 58/97, variable (12 11 0)	(1) Theme: Statistical Registers and Frames – Building and maintaining statistical registers to support business surveys; (2) Theme: Statistical Registers and Frames – The populations, frames and units of business surveys; (3) Theme: Statistical Registers and Frames – The statistical units and the business register
Unit of measurement error	An error that occurs when respondents report values that are consistently too high or too low by a constant factor.	Memobust definition (2014)	(1) Method: Deductive Editing; (2) Theme: Statistical Data Editing
Unit response rate	The ratio of the number of units for which data for some variables have been collected to the total number of units from which data are to be collected. It can indirectly measure response burden.	Eurostat (2009)	Theme: Response Burden
Unit types	A Business Register generally consists of several unit types, for example the enterprise unit, the kind of activity unit	Memobust definition (2014)	Method: Assigning random numbers when co-ordination of surveys based on different unit types is considered
Unity measure error	An error that occurs when respondents report the value of a variable in a wrong unity measure.	EDIMBUS Manual	Theme: Editing for Longitudinal Data
UPOS	Unplanned Preliminary Observed Sample. Early respondents used for provisional estimates. No specific follow-up has been planned	Memobust definition (2014)	Method: Subsampling for Preliminary Estimates
Upper bound	The highest possible value of a cell in a table of frequency counts where the cell value has been perturbed or suppressed.	Glossary on Statistical Disclosure Control (2014)	Theme: Statistical disclosure control methods for quantitative tables
User	A person or an organization that employs or applies statistical information, within or outside an NSI. Institutional users can also be stakeholders in the specification of which statistical information is produced.	Memobust definition (2014)	Theme: Specification of User Needs for Business Statistics
User needs	User needs refer to the data and metadata requirements of persons or organizations to meet a particular use or set of uses. Such needs may be specified in terms of the quality dimensions promulgated by international organizations or national agencies.	OECD Glossary of Statistical Terms	Theme: Specification of User Needs for Business Statistics
Value index	The ratio of transaction in current prices of the present and previous period	Memobust definition (2014)	Theme: Manual Integration
Variance	Expectation of the square difference between the estimates and its means value.	ESS Handbook on Precision Requirements and Variance Estimation for Household Surveys	Method: Generalised regression estimator

Variance	The variance is the mean square deviation of the variable around the average value. It reflects the dispersion of the empirical values around its mean.	Eurostat's Concepts and Definitions Database (2013)	Precision, random error.	Theme: Quality of Statistics
Variance	Expectation of the square difference between the estimates and its means value over the possible values.	See also Glossary of the Handbook on precision requirement and variance estimation for household surveys.		Theme: Weighting and Estimation
Vertical aggregation	Vertical aggregation: aggregation by sector or branch	European Communities (2001)		Theme: Seasonal adjustment – introduction and general description
Volume index	The result of a formula in which volume changes of various goods and services are weighed together in order get an index for the aggregate.	Memobust definition (2014)		Theme: Manual Integration
VVK	The Dutch Association of Chambers of Commerce. VVK means Vereniging Van Kamers van Koophandel	Hacking & Willenborg (2012)		Method: Automatic coding based on semantic networks
Waiver approach	Instead of suppressing tabular data, some agencies ask respondents for permission to publish cells even though doing so may cause these respondents' sensitive information to be estimated accurately. This is referred to as the waiver approach. Waivers are signed records of the respondents' granting permission to publish such cells. This method is most useful with small surveys or sets of tables involving only a few cases of dominance, where only a few waivers are needed. Of course, respondents must believe that their data are not particularly sensitive before they will sign waivers	Glossary on Statistical Disclosure Control (2014)		Theme: Statistical disclosure control methods for quantitative tables
Web forms	A form on a website that enables visitors to communicate with the host by filling in the fields and submitting the information. Information received via a form can be received by email and processed by other specific software.	OECD, 2004, Promise and Problems of E-Democracy: Challenges of Online Citizen Engagement, OECD, Paris, Annex 1: Commonly used E-Engagement Terms.		(1) Theme: Questionnaire Design; (2) Theme: Editing During Data Collection; (3) Theme: Testing the Questionnaire
Web Survey	A form of CASI in which a computer administers a questionnaire on a web site. In on-line surveys the questions are viewed and answered using a standard web browser on a PC, laptop or tablet. In an off-line survey the electronic questionnaire is downloaded and completed off-line. The responses are transferred through the internet to the server.	Memobust definition (2014)	CAWI	Theme: Mixed Mode Data Collection – design issues
Weight	The importance of an object in relation to a set of objects to which it belongs.	SDMX (2009)	Matching weight	(1) Method: Denton for Benchmarking; (2) Theme: Weighting and Estimation
Weight trimming	reduction of weights larger than some value	Memobust definition (2014)		Method: Outlier Treatment

Weighted Least Square	The parameter are obtained as those value the minimize a weighted square of distance between predicted and observed.	Memobust definition (2014)		Method: EBLUP Unit level for Small Area Estimation
Weighted Least Square (WLS)	Parameter estimates are obtained as the values maximizing the weighted square of distance between predicted and observed values.	Memobust definition (2014)		Method: EBLUP Area Level for Small Area Estimation (Fay-Herriot)
Weighting	The act of assigning weights to survey respondents, which are then used to obtain estimates of population parameters by calculating weighted sums of observed values.	CBS Methods Series Glossary		(1) Theme: Imputation; (2) Theme: Imputation for Longitudinal Data
Winsorization	modifying values in the sample so that the estimator becomes robust and isn't affected by large residuals	Memobust definition (2014)		Method: Outlier Treatment
Working/trading day effects	These are systematic effects in monthly times series related to changes in the day-of-week composition of each month and, in some cases, also to changes in the length of February. For flow series (monthly accumulations of daily activity e.g. monthly sales), the increases or decreases from average day-of-week activity associated with the days that occur five times in the month in a given year are important. For flow series, the length of February can have an impact. For stock series, such as end-of-month inventories, the extent to which inventories tend to rise or fall on the day of measurement (e.g. the last day of the month) can have an impact that is different from year to year. Attempts to measure analogous effects in quarterly series are seldom successful. A series of estimated trading day effects defines a trading day component for the time series	US Census Bureau		Method: Seasonal adjustment of economic time series
X-outliers	the X values of a few sample units that are very distant from the X-values of the other sample units.	Memobust definition (2014)	Outlier in the x-direction	Method: Outlier Treatment
Y-outliers	the Y values of a few sample units that are very distant from the Y-values of the other sample units.	Memobust definition (2014)	Outlier in the y-direction	Method: Outlier Treatment
T-ARGUS	Software program designed to protect statistical tables.	Argus (2013)		Theme: Logging

Source	Website
Argus (2013)	<a href="http://neon.vb.cbs.nl/casc/glossary.htm">http://neon.vb.cbs.nl/casc/glossary.htm</a> . Retrieved 25 October 2013
ESA (2010)	Eurostat (2010), European system of accounts (forthcoming).
ESS Regulation No 223 (2009)	<a href="http://www.ons.gov.uk/ons/about-ons/what-we-do/relationships-abroad/european-statistical-system--ess-/index.html">http://www.ons.gov.uk/ons/about-ons/what-we-do/relationships-abroad/european-statistical-system--ess-/index.html</a> , Regulation No 223/2009
Eurostat's Concepts and Definitions Database (2013)	<a href="http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL_GL_OSSARY&amp;StrNom=CODED2&amp;StrLanguageCode=EN">http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL_GL_OSSARY&amp;StrNom=CODED2&amp;StrLanguageCode=EN</a> . Retrieved 25 October 2013
Glossary on Statistical Disclosure Control (2014)	<a href="http://neon.vb.cbs.nl/casc/index.htm">http://neon.vb.cbs.nl/casc/index.htm</a>
Glossary, Adapting new technologies to census operations (2001)	Adapting new technologies to census operations, Arij Dekker, Symposium on Global Review of 2000 Round of Population and Housing Censuses: Mid-Decade Assessment and Future Prospects, Statistics Division, Department of Economic and Social Affairs, United Nations Secretariat New York, 7-10 August 2001, Glossary.
Golden (1976)	Golden, M.P. (1976), <i>The research experience</i> . F.E. Peacock Publishers Inc., Itasca, Illinois, USA
Hacking & Willenborg (2012)	Hacking, W. and L. Willenborg (2012), <i>Coding – interpreting short descriptions using a classification</i> . Translation of a contribution to the CBS Methods Series, Report, Statistics Netherlands, The Hague and Heerlen.
Hacking, W. and L. Willenborg (2012) Memobust definition (2014)	Hacking, W. and L. Willenborg (2012), <i>Coding – interpreting short descriptions using a classification</i> . Translation of a contribution to the CBS Methods Series, Report, Statistics Netherlands, The Hague and Heerlen. <a href="http://www.cros-portal.eu/content/memobust">http://www.cros-portal.eu/content/memobust</a>
NACE Rev.2	<a href="http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-07-015/EN/KS-RA-07-015-EN.PDF">http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-07-015/EN/KS-RA-07-015-EN.PDF</a>
NUTS classification	<a href="http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction">http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction</a>
OECD Glossary	<a href="http://stats.oecd.org/glossary/detail.asp?ID=5069">http://stats.oecd.org/glossary/detail.asp?ID=5069</a>
RAMON, Eurostat's metadata server	
SDMX (2009)	<a href="http://sdmx.org">http://sdmx.org</a>

SDMX (2009) [http://sdmx.org/wp-content/uploads/2009/01/04\\_sdmx\\_cog\\_annex\\_4\\_mcv\\_2009.pdf](http://sdmx.org/wp-content/uploads/2009/01/04_sdmx_cog_annex_4_mcv_2009.pdf)

US Census Bureau <http://www.census.gov/srd/www/x13as/glossary.html>

US Department of Labour [https://www.osha.gov/pls/imis/sic\\_manual.html](https://www.osha.gov/pls/imis/sic_manual.html)

Wikipedia Cluster Sampling [http://en.wikipedia.org/wiki/Cluster\\_sampling](http://en.wikipedia.org/wiki/Cluster_sampling)

Wikipedia Multistage Sampling [http://en.wikipedia.org/wiki/Multistage\\_sampling](http://en.wikipedia.org/wiki/Multistage_sampling)