

ESS Standard for Quality Reports Structure (ESQRS) 2.0

The ESS Standard for Quality Reports Structure (ESQRS) contains the description and representation of statistical metadata concepts to be used for providing detailed information for assessing data quality. The broad concepts used are compatible with the SDMX cross-domain concepts and with the common terminology as published within the SDMX Glossary (2016). The detailed quality concepts are based on the ESS Standard for Quality Reports (ESQR) from 2009.

The ESQRS is addressed to the European Statistical System. It is implemented at Eurostat and at national level: the application of the concepts and sub concepts at European level and at national level are provided in the ESS Handbook for Quality Reports (EHQR) from 2014 and the ESS Guidelines for the implementation of the ESS Quality and Performance Indicators from 2014.

	Concept Name	Concept Code	Descriptions	
1	Contact	CONTACT	Individual or organisational contact points for the data or metadata, including information on how to reach the contact points.	
1.1	Contact organisation	CONTACT_ORGANISATION	The name of the organisation of the contact points for the data or metadata.	The full name of your organisation.
1.2	Contact organisation unit	ORGANISATION_UNIT	An addressable subdivision of an organisation	The name of the unit responsible for the metadata file (it can also include a unit number).
1.3	Contact name	CONTACT_NAME	The name of the contact points for the data or metadata.	The name of the person responsible for the statistical domain (first name and family name).
1.4	Contact person function	CONTACT_FUNCT	The area of technical responsibility of the contact, such as "methodology", "database management" or "dissemination".	The title of the person responsible for the statistical domain (this title can contain the precise area of responsibility such as methodologist or data base manager)
1.5	Contact mail address	CONTACT_MAIL	The postal address of the contact points for the data or metadata.	The postal address of the person responsible for the statistical domain.
1.6	Contact email address	CONTACT_EMAIL		The email address of the person responsible for the statistical domain (this can be an individual mail address or a functional mailbox).
1.7	Contact phone number	CONTACT_PHONE	The telephone number of the contact points for the data or metadata.	The phone number of the person responsible for the statistical domain.
1.8	Contact fax number	CONTACT_FAX	Fax number of the contact points for the data or metadata.	The fax number of the person responsible for the statistical domain.
2	Statistical presentation	STAT_PRES		
2.1	Data description	DATA_DESCR	Main characteristics of the data set described in an easily understandable manner, referring to the data and indicators disseminated.	Describe the main characteristics of the data set in an easily understandable manner, referring to the main data and indicators disseminated. This short description should be understood immediately and easily by the users.
2.2	Classification system	CLASS_SYSTEM	Arrangement or division of objects into groups based on characteristics which the objects have in common.	
2.3	Sector coverage	COVERAGE_SECTOR	Main economic or other sectors covered by the statistics.	

ESS Standard for Quality Reports Structure (ESQRS) 2.0

2.4	Statistical concepts and definitions	STAT_CONC_DEF	Statistical characteristics of statistical observations.	
2.5	Statistical unit	STAT_UNIT	Entity for which information is sought and for which statistics are ultimately compiled.	
2.6	Statistical population	STAT_POP	The total membership or population or "universe" of a defined class of people, objects or events.	
2.7	Reference area	REF_AREA	The country or geographic area to which the measured statistical phenomenon relates.	
2.8	Time coverage	COVERAGE_TIME	The length of time for which data are available.	
2.9	Base period	BASE_PER	The period of time used as the base of an index number, or to which a constant series refers.	
3	Statistical processing	STAT_PROCESS	Statistical processing	
3.1	Source data	SOURCE_TYPE	Characteristics and components of the raw statistical data used for compiling statistical aggregates.	Indicate if the data set is based on a survey or on administrative data sources. If sample surveys are used, some sample characteristics should also be given (e.g. gross and net sample size, type of sampling design, reporting domain etc.). If administrative registers are used, the description of registers should be given (source, year, primary purpose, potential deficiencies etc.)
3.2	Frequency of data collection	FREQ_COLL	Frequency with which the source data are collected.	Indicate the frequency of data collection (e.g. monthly, quarterly, annually, continuous). The frequency can also be expressed in using the codes released in the harmonised code list available for the European Statistical System.
3.3	Data collection	COLL_METHOD	Systematic process of gathering data for official statistics.	Describe the method used to gather data from respondents (e.g. postal survey, CAPI, on-line survey, etc.). Some additional information on questionnaire design and testing, interviewer training, methods used to monitor non-response etc. should be provided here. Reference: The Handbook of Recommended Practices for Questionnaire Development and Testing Methods in the ESS (Eurostat, 2006).
3.4	Data validation	DATA_VALIDATION	Process of monitoring the results of data compilation and ensuring the quality of the statistical results.	Describe the procedures for checking and validating the source data and how the results of these validations are monitored and used.

ESS Standard for Quality Reports Structure (ESQRS) 2.0

3.5	Data compilation	DATA_COMP	Operations performed on data to derive new information according to a given set of rules.	Describe the data compilation process (e.g. data editing, imputation, weighting, adjustment for non-response, calibration, model used etc.). Main reference: Survey sampling reference guidelines – Introduction to sample design and estimation techniques (Eurostat, 2008).
3.6	Adjustment	ADJUSTMENT	The set of procedures employed to modify statistical data to enable it to conform to national or international standards or to address data quality differences when compiling specific data sets.	Describe the statistical procedures used for adjusting the data (such as seasonal adjustment methods, time series decomposition, or other similar methods). Main reference: ESS Guidelines on seasonal adjustment (2008).
4	Quality management	QUALITY_MGMNT	Systems and frameworks in place within an organisation to manage the quality of statistical products and processes.	Describe briefly the quality management system used in the organisation (EFQM, ISO- series etc.) and as well hyperlink to the European Statistics Code of Practice.
4.1	Quality assurance	QUALITY_ASSURE	All systematic activities implemented that can be demonstrated to provide confidence that the processes will fulfil the requirements for the statistical output.	Hyperlink to the general quality assurance framework (or similar) and brief description how it is implemented for the domain-specific quality assurance activities (the use of best practices, quality reviews, self-assessments, compliance monitoring etc.).
4.2	Quality assessment	QUALITY_ASSMNT	Overall assessment of data quality, based on standard quality criteria.	The standard quality criteria are provided in concepts 12 -15.A qualitative assessment of the overall quality of the statistical outputs should be provided by summarising the main strengths and possible quality deficiencies. Any trade-offs between quality aspects can be mentioned as well as planned quality improvements. Main reference: "ESS Handbook for Quality Reports" (2014).
5	Relevance	RELEVANCE	The degree to which statistical information meets the real or perceived needs of clients.	
5.1	User Needs	USER_NEEDS	Description of users and their respective needs with respect to the statistical data.	Provide a classification of users with some indication of their importance , an indication of the uses for which they want the statistical outputs and as well users and uses given special considerations. Unmet user needs and the reasons for not meeting them should be included as well.
5.2	User Satisfaction	USER_SAT	Measures to determine user satisfaction.	Describe how the views and opinions of the users are collected. In addition the main results regarding the user satisfaction should be shown (in the form of a user satisfaction index if available) and the date of most recent user satisfaction survey.

ESS Standard for Quality Reports Structure (ESQRS) 2.0

5.3	Completeness	COMPLETENESS	The extent to which all statistics that are needed are available.	Provide information on completeness compared with relevant regulations/ guidelines. Applicable for Eurostat: if any Member States are not producing the statistics in question.
5.3.1	Data completeness - rate	COMPLETENESS_RATE	The ratio of the number of data cells provided to the number of data cells required.	
6	Accuracy and reliability	ACCURACY_RELIABILITY	Accuracy: closeness of computations or estimates to the exact or true values that the statistics were intended to measure Reliability: closeness of the initial estimated value to the subsequent value.	The accuracy of statistical outputs in the general statistical sense is the degree of closeness of computations or estimates to the exact or true values that the statistics were intended to measure. Reliability refers to the closeness of the initial estimated value to the subsequent estimated value.
6.1	Accuracy - overall	ACCURACY_OVERALL	Assessment of accuracy, linked to a certain data set or domain, which is summarising the various components into one single measure.	Provide a summary of the main sources of error and an assessment of the potential for bias (sign and order of magnitude) for each key indicator in quantitative or qualitative terms.
6.2	Sampling error	SAMPLING_ERR	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 subset of the population is enumerated.	If probability sampling is used, it should be provided estimates of the accuracy, normally in the form of cv's, standard errors or confidence intervals. It should be stated if adjustments for non-response, misclassifications and other uncertainty sources such as outlier treatment are included. - If non-probability sampling is used, the responsible for the statistical domain should provide estimates of the accuracy, a motivation for the invoked model for this estimation, and brief discussion of sampling bias.
6.2.1	Sampling error - indicators	SAMPLING_ERR_IND	Precision measures for estimating the random variation of an estimator due to sampling.	
6.3	Non-sampling error	NONSAMPLING_ERR	Error in sample estimates which cannot be attributed to sampling fluctuations.	Provide an assessment, preferable quantitative, on the non-sampling errors and the bias risks associated with: - Overcoverage, undercoverage and multiple listings. - Survey instrument, respondent and interviewer where relevant. - Unit (non)response including causes for nonresponse and measures to reduce nonresponse. - Item nonresponse for key variables. - Data editing, coding and imputation where relevant. - Specific models used in estimation (. Actions undertaken to reduce the different types of errors could be provided as well.
6.3.1	Coverage error	COVERAGRE_ERR	Divergence between the frame population and the target population.	
6.3.1.1	Over-coverage - rate	OVERCOVERAGE_RATE	The proportion of units accessible via the frame that do not belong to the target population.	

ESS Standard for Quality Reports Structure (ESQRS) 2.0

6.3.1.2	Common units - proportion	COMMON_UNIT_SHARE	The proportion of common units covered by both the survey and the administrative sources in relation to the total number of units in the survey.	
6.3.2	Measurement error	MEASUREMENT_ERROR	Error in reading, calculating or recording numerical value.	
6.3.3	Non response error	NONRESPONSE_ERROR	The difference between the statistics computed from the collected data and those that would be computed if there were no missing values.	
6.3.3.1	Unit non-response - rate	UNIT_NONRESPONSE_RATE	The ratio of the number of units with no information or not usable information to the total number of in-scope (eligible) units.	
6.3.3.2	Item non-response - rate	ITEM_NONRESPONSE_RATE	The ratio of the in-scope (eligible) units which have not responded to a particular item and the in-scope units that are required to respond to that particular item	
6.3.4	Processing error	PROCESSING_ERROR	The error in final data collection process results arising from the faulty implementation of correctly planned information methods.	
6.3.4.1	Imputation - rate	IMPUTATION_RATE	The ratio of the number of replaced values to the total number of values for a given variable.	
6.3.5	Model assumption error	MODEL_ASSUMPTION_ERROR	Error due to domain specific models needed to define the target of estimation.	
6.4	Seasonal adjustment	SEASONAL_ADJ	The statistical technique used to remove the effects of seasonal calendar influences operating on a series.	
6.5	Data revision - policy	REV_POLICY	Policy aimed at ensuring the transparency of disseminated data, whereby preliminary data are compiled that are later revised.	
6.6	Data revision - practice	REV_PRACTICE	Information on the data revision practice.	
6.6.1	Data revision - average size	DATA_REV_AVGSIZE	The average over a time period of the revisions of a key item. The 'revision' is defined as the difference between a later and an earlier estimate of the key item.	
7	Timeliness and punctuality	TIMELINESS_PUNCT	Timeliness and punctuality	
7.1	Timeliness	TIMELINESS	Length of time between data availability and the event or phenomenon they describe	Provide, for annual or more frequent releases, the average production time for each release of data. Applicable for Eurostat: - National data deliveries: the agreed time frame for deliveries should be included as well as the achieved dates for deliveries during a past period.
7.1.1	Time lag - first result	TIMELAG_FIRST	The number of days (or weeks or months) from the last day of the reference period to the day of publication of first results.	

ESS Standard for Quality Reports Structure (ESQRS) 2.0

7.1.2	Time lag - final result	TIMELAG_FINAL	The number of days (or weeks or months) from the last day of the reference period to the day of publication of complete and final results.	
7.2	Punctuality	PUNCTUALITY	Time lag between the actual delivery of the data and the target date when it should have been delivered.	Provide, for annual or more frequent releases: - The percentage of releases delivered on time, based on scheduled release dates. - The reasons for non-punctual releases explained.
7.2.1	Punctuality - delivery and publication	PUNCTUALITY_RELEASE	The number of days between the delivery/release date of data and the target date on which they were scheduled for delivery/release.	
8	Coherence and comparability	COHER_COMPAR	Adequacy of statistics to be reliably combined in different ways and for various uses and the extent to which differences between statistics can be attributed to differences between the true values of the statistical characteristics.	Coherence measures the adequacy of the statistics to be combined in different ways and for various uses. Comparability is a measurement of the impact of differences in applied statistical concepts, measurement tools and procedures where statistics are compared between geographical areas or over time.
8.1	Comparability - geographical	COMPAR_GEO	Extent to which statistics are comparable between geographical areas.	Describe any problems of comparability between countries or regions. The reasons for the problems should be described and as well the order of magnitude of the effects of the main sources of errors. Information on discrepancies from the ESS/ international concepts and definitions should be included. Also asymmetries for statistical mirror flows should be described.
8.1.1	Asymmetry for mirror flow statistics - coefficient	ASYMMETRY_COEFF	The difference or the absolute difference of inbound and outbound flows between a pair of countries divided by the average of these two values.	
8.2	Comparability - over time	COMPAR_TIME	Extent to which statistics are comparable or reconcilable over time.	Provide information on the length of comparable time series, reference periods at which series breaks occur, the reasons for the breaks and treatments of them.
8.2.1	Length of comparable time series	COMPAR_LENGTH	The number of reference periods in time series from last break.	
8.3	Coherence - cross domain	COHER_X_DOM	Extent to which statistics are reconcilable with those obtained through other data sources or statistical domains.	Describe the differences of the statistical outputs in question to other related statistical outputs. The order of magnitude of the effects of the differences should be assessed as well.
8.4	Coherence - sub annual and annual statistics	COHER_FREQSTAT	The extent to which statistics of different frequencies are reconcilable	
8.5	Coherence - National Accounts	COHER_NATACCOUNTS	The extent to which statistics are reconcilable with National Accounts.	
8.6	Coherence - internal	COHER_INTERNAL	Extent to which statistics are consistent within a given data set.	Describe if statistical outputs within the data set in question are not consistent and the reasons for publishing such

ESS Standard for Quality Reports Structure (ESQRS) 2.0

				results.
9	Accessibility and clarity	ACCESSIBILITY_CLARITY	The conditions and modalities by which users can obtain, use and interpret data.	The various means of dissemination used for making the data set available to users should be described (including the various dissemination formats available as well as their accessibility).
9.1	News release	NEWS_REL	Regular or ad-hoc press releases linked to the data.	Regular or ad-hoc press releases linked to the data set in question should be described.
9.2	Publications	PUBLICATIONS	Regular or ad-hoc publications in which the data are made available to the public.	The publications using the data set in question should be described in short.
9.3	Online database	ONLINE_DB	Information about on-line databases in which the disseminated data can be accessed.	The on-line database available for the data set in question should be described. This includes the domain names as released on the website.
9.3.1	Data tables - consultations	DATATABLE_CONSULT	Number of consultations of data tables within a statistical domain for a given time period displayed in a graph.	
9.4	Microdata access	MICRO_DAT_ACC	Information on whether micro-data are also disseminated.	Describe if and how the data set is accessible as micro-data (e.g. for researchers). Also the micro-data anonymization rules should be described in short.
9.5	Other	DISS_OTHER	References to the most important other data dissemination done.	The most important other data dissemination means should be described (e.g. within other publications, policy papers, etc.).
9.6	Documentation on methodology	DOC_METHOD	Descriptive text and references to methodological documents available.	Describe the availability of national reference metadata files, important methodological papers, summary documents or other important handbooks.
9.7	Quality documentation	QUALITY_DOC	Documentation on procedures applied for quality management and quality assessment.	Describe the availability of quality reports and studies. For Eurostat: The responsible of the statistical domain should also describe the availability of national quality reports.
9.7.1	Metadata completeness - rate	METADATA_COMPLETE	The ratio of the number of metadata elements provided to the total number of metadata elements applicable.	
9.7.2	Metadata - consultations	METADATA_CONSULT	Number of consultations within a statistical domain for a given time period.	
10	Cost and Burden	COST_BURDEN	Cost associated with the collection and production of a statistical product and burden on respondents.	Provide a summary of costs for production of statistical data and of the burden on respondents (in general measured in time used). Objectives/actions concerning burden reduction could be additionally provided. Main references: Handbook for Monitoring and Evaluating Business Survey Response Burdens (Eurostat, 2007) International Standard Cost Model Manual (SCM network)

ESS Standard for Quality Reports Structure (ESQRS) 2.0

11	Confidentiality	CONF	A property of data indicating the extent to which their unauthorized disclosure could be prejudicial or harmful to the interest of the source or other relevant parties.	The legislation (or any other formal provision) related to statistical confidentiality applied for the data set in question as well as the actual confidentiality data treatment done (also with regard to the aggregated data disseminated) should be described.
11.1	Confidentiality - policy	CONF_POLICY	Legislative measures or other formal procedures which prevent unauthorized disclosure of data that identify a person or economic entity either directly or indirectly.	The European and national legislations related to statistical confidentiality should be described.
11.2	Confidentiality - data treatment	CONF_DATA_TR	Rules applied for treating the data set to ensure statistical confidentiality and prevent unauthorized disclosure.	The rules applied for treating the data set with regard to statistical confidentiality should be described (e.g. controlled rounding, cell suppression, aggregation of disclosive information, aggregation rules on aggregated confidential data, primary confidentiality with regard to single data values, etc.). Main reference: Handbook on Statistical Disclosure Control (2007).
12	Comment	COMMENT_DSET	Supplementary descriptive text which can be attached to data or metadata	