Quality Indicators for the Generic Statistical Business Process Model (GSBPM) - For Statistics derived from Surveys and Administrative Data Sources

(Version 2.0, October 2017)

Contents

Abbreviations

ADS – Administrative data sources
BPMN – Business Process Modelling Notation
CATI – Computer-assisted Telephone Interview
CAPI – Computer-assisted Personal Interview
CoP – Code of Practice
CSPA – Common Statistical Production Architecture
CV – Coefficient of Variation
ES – European Statistics
ESMS – Euro SDMX Metadata Structure
ESS – European Statistical System
ESQRS – Standard for Quality Reports Structure
GSBPM – Generic Statistical Business Process Model
ISO – International Organization for Standardization
ICT – Information and Communication Technology
KPIs – Key Performance Indicators
LMI – Logical Information Model
NSO – National Statistical Office
NQAF – National Quality Assurance Framework
PAPI – Paper and Pencil Interview
QPI – Quality and Performance Indicators
SDMX – Statistical Data and Metadata eXchange
SIMS – Single Integrated Metadata Structure
XBRL – Extensible Business Reporting Language

Acknowledgements

Contributors to this version include: Alice Born and Laurie Reedman (Statistics Canada); Sinikka Laurila (Statistics Finland); Giovanna Brancato and Marina Signore (ISTAT, Italy); Felipa Zabala and Anapapa Mulitalo (Stats NZ); Nilgün Dorsan and Deniz Özkan (TurkStat, Turkey); and Christopher Jones and Heini Salonen (UNECE).

I. Introduction

i. Quality concerns organisations, processes and products. This document outlines a set of quality indicators that have been developed for the production of statistics from both survey and administrative data sources (ADS), with reference to the different stages of the Generic Statistical Process Model (GSBPM) Version 5.0.

ii. The main goal of quality management within the statistical business process is to understand and manage the quality of the statistical products. There is general agreement among statistical organisations that quality should be defined according to the ISO 9000-2005 standard: "The degree to which a set of inherent characteristics fulfils requirements".[1] Thus, product quality is a complex and multi-faceted concept, usually defined in terms of several quality dimensions. The dimensions of quality that are considered most important depend on user perspectives, needs and priorities, which vary between processes and across groups of users.[2]

iii. A fundamental role in quality management is played by a set of quality indicators that should be implemented within the sub-processes to prevent and monitor errors. The first version of the quality indicators released in May 2016, focussed on surveys, and complemented the quality management process of the GSBPM. The second version of the quality indicators includes indicators for statistical processes based on both survey data and administrative data. The quality indicators are integrated in each GSBPM sub-process since some indicators apply to both surveys and administrative data sources, while others apply to either surveys or administrative data sources.

iv. Quality indicators are mapped to each sub-process of the GSBPM. The following guiding principles in mapping the quality indicator to the GSBPM were used:

- Indicators cover direct surveys and administrative data sources;
- Develop generic indicators to reflect the nature of the GSBPM as a reference model;
- Be consistent with existing quality assurance frameworks when selecting the quality indicators and determining their related quality dimension;
- No formulas are used to express the indicators, only descriptions or explanations;
- Quantitative indicators were used whenever possible;
- Qualitative indicators in the form of yes/no or large/medium/low were proposed when appropriate;
- Map indicators to the phase they measure even if they might be calculated at a later stage; and
- Allow for a certain degree of redundancy by mentioning the same indicators in different phases or sub-processes.

v. Quality indicators were determined by examining practices within national statistical agencies, United Nations' Statistical Commission National Quality Assurance Framework, European Statistics (ES) Code of Practice, Euro SDMX Metadata Structure (ESMS), national and Eurostat quality assurance frameworks, European Statistical System (ESS) Standard for Quality Reports Structure (ESQRS) and Single Integrated Metadata Structure (SIMS).

vi. While mapping the quality indicators to the GSBPM, the related quality dimensions were determined for each indicator. The global framework of the National Quality Assurance Framework (NQAF) developed by an expert group under the United Nations Statistical Commission was used. Each of the quality indicators is mapped to one of the dimensions of

[1] ISO 9000:2005, Quality management systems – Fundamentals and vocabulary, International Organization for Standardization.
[2] Generic Statistical Business Process Model (GSBPM) Version 5, UNECE, December, 2013.

NQAF. In addition, the quality indicators are mapped to the quality dimensions of the ES Code of Practice Principles.

vii. The quality indicators are presented after each sub-process in a table format – Column 1: Quality dimension; Column 2: Quality indicator; and Column 3: Notes. The NQAF quality dimension is presented in Column 1 in the order of the (nineteen) dimensions of the NQAF. The quality dimension of the ES CoP is noted in column 3 when this differs from the NQAF dimension.

viii. In addition to the quality indicators for each phase and sub-process of the GSBPM, quality indicators are attached to both the quality management and metadata management overarching processes to address the aspect of overall quality management and overall quality of metadata.

ix. Indicators for quality management are related to the availability of a quality policy, quality assurance plan, monitoring procedures and organizational structure for managing quality. Quality indicators for metadata management cover the availability of a policy on metadata documentation, quality of the metadata (i.e., in terms of completeness, accuracy, timeliness, accessibility, etc.), compliance to international metadata standards, availability of a metadata system, and life cycle management of the metadata.

x. The paragraph numbers in the following section correspond to the paragraph numbers in version 5.0 of GSBPM document. Therefore the numbering starts from 36 instead of 1.

II. Quality indicators for the GSBPM phases and sub-processes

36. This section considers each phase in turn, identifying the various sub-processes within that phase, and describing their contents.

Specify Needs Phase

37. This phase is triggered when a need for new statistics is identified, or feedback about current statistics initiates a review. It includes all activities associated with engaging customers to identify their detailed statistical needs, proposing high level solution options and preparing business cases to meet these needs.

38. In this phase, the organisation:

- identifies the need for the statistics;
- confirms, in more detail, the statistical needs of the stakeholders;
- establishes the high level objectives of the statistical outputs;
- identifies the relevant concepts and variables for which data are required;
- checks the extent to which current data sources can meet these needs;
- prepares the business case to get approval to produce the statistics.

39. This phase is broken down into six sub-processes. These are generally sequential, from left to right, but can also occur in parallel, and can be iterative. The sub-processes are:

1.1. Identify Needs

40. This sub-process includes the initial investigation and identification of what statistics are needed and what is needed of the statistics. It may be triggered by a new information request, an environmental change such as a reduced budget. Action plans from evaluations of previous iterations of the process, or from other processes, might provide an input to this sub-process. It also includes consideration of practice amongst other (national and international) statistical organisations producing similar data, and in particular the methods used by those organisations. It may involve consideration of specific needs of different user communities, such as the disabled or different ethnic groups.

Quality Dimension	Indicator	Notes
Relevance	To what extent have stakeholders been identified and included in discussions about statistical needs? To what extent has relevant supporting documentation been gathered?	

1.2. Consult and confirm needs

41. This sub-process focuses on consulting with the stakeholders and confirming in detail the needs for the statistics. A good understanding of user needs is required so that the statistical organisation knows not only what it is expected to deliver, but also when, how, and, perhaps most importantly, why. For second and subsequent iterations of this phase, the main focus will be on determining whether previously identified needs have changed. This detailed understanding of user needs is the critical part of this sub-process.

Quality Dimension	Indicator	Notes
Relevance	To what extent have stakeholders confirmed the detailed statistical needs (what, when, how and why) as documented by the NSO?	Could be a two part indicator; proportion of stakeholders who have confirmed, and proportion of statistical needs confirmed.
Relevance	To what extent does the data source satisfy information demand?	

1.3. Establish output objectives

42. This sub-process identifies the statistical outputs that are required to meet the user needs identified in sub-process 1.2 (Consult and confirm needs). It includes agreeing the suitability of the proposed outputs and their quality measures with users. Legal frameworks (e.g. relating to confidentiality), and available resources are likely to be constraints when establishing output objectives.

Quality Dimension	Indicator	Notes
Statistical Confidentiality and security	To what extent have legal constraints regarding statistical outputs been considered, for example but not limited to ensuring confidentiality of data and preventing the disclosure of sensitive information?	Confidentiality and security includes privacy.
Relevance	To what extent have all statistical needs been addressed by the proposed outputs?	
Accuracy and reliability	To what extent are the proposed outputs and their quality measures suitable to user needs?	
Adequacy of resources	To what extent have resource requirements for the proposed outputs and their quality measures been considered?	Includes extreme value checks, population unit checks, variable checks, combinations of variables checks, etc.

1.4. Identify concepts

43. This sub-process clarifies the required concepts to be measured by the business process from the point of view of the user. At this stage the concepts identified may not align with existing statistical standards. This alignment, and the choice or definition of the statistical concepts and variables to be used, takes place in sub-process 2.2.

Quality Dimension	Indicator	Notes
Relevance	Compliance rate of concepts and definitions of variables with existing standards	
Relevance	Metadata for ADS to determine if relevant variables are available (e.g. presence of useful combinations of variables.)	Metadata about variable definitions is very important when output variables differ from standards. For example, personal income variables on different datasets can include/exclude some income sources – be clear what exactly is included to avoid misinterpretation.
Relevance	Percentage of items that deviate from the target concept or international standards When assessing the usability of the variables for a statistical output, we can weight this indicator for whether or not the variables are key to the statistical output.	

1.5. Check data availability

44. This sub-process checks whether current data sources could meet user requirements, and the conditions under which they would be available, including any restrictions on their

use. An assessment of possible alternatives would normally include research into potential administrative or other non-statistical data sources, to determine whether they would be suitable for use for statistical purposes. When existing sources have been assessed, a strategy for filling any remaining gaps in the data requirement is prepared. This sub-process also includes a more general assessment of the legal framework in which data would be collected and used, and may therefore identify proposals for changes to existing legislation or the introduction of a new legal framework.

Quality Dimension	**Indicator**	**Notes**
Statistical Confidentiality and security	To what extent have legal constraints regarding data collection, acquisition and use been assessed and any necessary changes been proposed?	Confidentiality and security includes privacy.
Relevance	To what extent do current data sources meet user requirements, taking into consideration the conditions under which they would be available and any restrictions on their use? If current data sources do not fully meet user requirements, to what extent has a strategy been proposed to fully meet user requirements?	If an administrative data source is being considered, ensure the following are assessed: continuity of data supply, setting up of responsibilities between data provider and NSO, ICT build resources (data storage, technology required to handle incoming data and data processing). If data are not available, a contingency plan should be in place.
Relevance	Existence of an advance notification plan about the forthcoming changes to the data source Is a contingency plan for changes to the data or data source in place?	
Accuracy and Reliability	Completeness of data source(s), such as:- • Percentage of units not belonging to the target population • Percentage of units missing from the target population • Coverage of the data • Absence of values for key variables • Missing values in the source • Total percentage of empty cells	
Relevance	Availability of a unique key	
Relevance	To what extent does the timeliness of the delivery detract from its relevance?	
Cost effectiveness	For ADS, has the data source been evaluated in terms of its cost effectiveness of data transmission interfaces?	

1.6. Prepare business case

45. This sub-process documents the findings of the other sub-processes in this phase in the form of a business case to get approval to implement the new or modified statistical business process. Such a business case would need to conform to the requirements of the approval body, but would typically include elements such as:

- A description of the "As-Is" business process (if it already exists), with information on how the current statistics are produced, highlighting any inefficiencies and issues to be addressed;
- The proposed "To-Be" solution, detailing how the statistical business process will be developed to produce the new or revised statistics;
- An assessment of costs and benefits, as well as any external constraints.

Quality Dimension	Indicator	Notes
Adequacy of resources	To what extent have resource requirements for the proposed outputs and their quality measures been considered?	Includes extreme value checks, population unit checks, variable checks, combinations of variables checks, etc.
Adequacy of resources	Has the data source been evaluated in terms of its cost effectiveness?	
Relevance	To what extent does the business case conform to the requirements of the approval body?	
Relevance	To what extent does the business case reflect the findings, recommendations and proposals from steps 1.2 to 1.5?	

Design Phase

46. This phase describes the development and design activities, and any associated practical research work needed to define the statistical outputs, concepts, methodologies, collection instruments[3] and operational processes. It includes all the design elements needed to define or refine the statistical products or services identified in the business case. This phase specifies all relevant metadata, ready for use later in the statistical business process, as well as quality assurance procedures. For statistical outputs produced on a regular basis, this phase usually occurs for the first iteration, and whenever improvement actions are identified in the Evaluate phase of a previous iteration.

[3] For GSBPM purposes, collection instruments are defined broadly to include any tool or routine to gather or extract data and metadata, from paper questionnaires to web-scraping tools. In GSIM version 1.1, collection instruments are “exchange channels” used for incoming information.

47. Design activities make substantial use of international and national standards, in order to reduce the length and cost of the design process, and enhance to comparability and usability of outputs. Organisations are also encouraged to reuse or adapt design elements from existing processes. Additionally, outputs of design processes may form the basis for future standards at the organisation, national or international levels.

48. This phase is broken down into six sub-processes, which are generally sequential, from left to right, but can also occur in parallel, and can be iterative. These sub-processes are:

2.1. Design outputs

49. This sub-process contains the detailed design of the statistical outputs, products and services to be produced, including the related development work and preparation of the systems and tools used in the "Disseminate" phase. Disclosure control methods, as well as processes governing access to any confidential outputs are also designed here. Outputs should be designed to follow existing standards wherever possible, so inputs to this process may include metadata from similar or previous collections, international standards, and information about practices in other statistical organisations from sub-process 1.1 (Identify needs).

Quality Dimension	**Indicator**	**Notes**
Statistical Confidentiality and security	Have the confidentiality rules and micro data access procedures been designed?	yes/no indicator
Relevance	Percentage of/Extent to which outputs fulfil users' needs (and/or priority needs)	Link to "identify needs" (sub-process 1.1) and to the "evaluate" phase
Relevance	Percentage of/ Extent to which outputs changed as a result of improvement actions or as a result of user satisfaction surveys/analyses (for outputs produced on a regular basis)	Link to "identify needs" (sub-process 1.1) and to the "evaluate" phase
Relevance	Planned data completeness rate: extent to which the planned outputs will satisfy requirements (e.g. from Regulations or other agreements with users)	Could be calculated as the ratio of the number of data cells planned to the number of data cells required ESS QPI - R1. (Planned) Data completeness rate
Coherence and comparability	Expected length of comparable time series.	Breaks in statistical time series may occur when there is a change in the definition of the parameter to be estimated (e.g. variable or population) or the methodology used for the estimation. Sometimes a break can be prevented, e.g. by linking. The length of comparable time series is applicable: • to all statistical

		processes producing time-series; • to users and producers, with different level of details given. ESS QPI - CC2. Length of comparable time series
Accuracy and reliability	Data revisions are planned (Yes/No)	ESS QPI - A6. Data revision - average size.

2.2. Design variable descriptions

50. This sub-process defines the statistical variables to be collected via the collection instrument, as well as any other variables that will be derived from them in sub-process 5.5 (Derive new variables and units), and any statistical classifications that will be used. It is expected that existing national and international standards will be followed wherever possible. This sub-process may need to run in parallel with sub-process 2.3 (Design collection), as the definition of the variables to be collected, and the choice of collection instrument may be inter-dependent to some degree. Preparation of metadata descriptions of collected and derived variables and classifications is a necessary precondition for subsequent phases.

Quality Dimension	**Indicator**	**Notes**
Cost effectiveness	Percentage of/ Extent to which concepts, definitions and classifications associated to (key) variables and populations, are re-used from other similar surveys and ADS	
Managing metadata	Percentage of/Extent to which concepts, definitions and classifications associated to (key) variables and populations follow international or national standards	See also5.5 for derived variables Corresponds to accessibility and clarity principle in the ES Code of Practice
Managing metadata	Percentage of/Extent to which new concepts, definitions and classifications are introduced (provide motivation for it)	Corresponds to accessibility and clarity principle in the ES Code of Practice
Managing metadata	Percentage of metadata adequately archived (easily retrievable; properly labelled; retention period indicated)	Corresponds to accessibility and clarity principle in the ES Code of Practice
Managing metadata	Percentage of / extent to which collected (survey and ADS) and derived variables and classifications have metadata descriptions.	Corresponds to accessibility and clarity principle in the ES Code of Practice

2.3. Design collection

51. This sub-process determines the most appropriate collection method(s) and instrument(s). The actual activities in this sub-process will vary according to the type of collection instruments required, which can include computer assisted interviewing, paper questionnaires, administrative data interfaces and data integration techniques. This sub-process includes the design of collection instruments, questions and response templates (in conjunction with the variables and statistical classifications designed in sub-process 2.2

(Design variable descriptions)). It also includes the design of any formal agreements relating to data supply, such as memoranda of understanding, and confirmation of the legal basis for the data collection. This sub-process is enabled by tools such as question libraries (to facilitate the reuse of questions and related attributes), questionnaire tools (to enable the quick and easy compilation of questions into formats suitable for cognitive testing) and agreement templates (to help standardise terms and conditions). This sub-process also includes the design of process-specific provider management systems.

Quality Dimension	**Indicator**	**Notes**
Soundness of implementation	Is the process re-using known methods and collection systems, e.g. according to guidelines/recommendations?	yes/no indicator Corresponds to the appropriate statistical procedures principle in the ES Code of Practice
Soundness of implementation	How well does the collection method suit the nature and volume of the information to be gathered?	fully/partly/no indicator Corresponds to the appropriate statistical procedures principle in the ES Code of Practice
Soundness of implementation	When has the data collection technique last been revised/improved?	For outputs produced on a regular basis. Corresponds to the appropriate statistical procedures principle in the ES Code of Practice
Soundness of implementation	Appropriateness of questionnaire to the pre-specified standards.	Corresponds to the appropriate statistical procedures principle in the ES Code of Practice
Managing respondent burden	Percentage of questions used to collect information which will not be published (and motivation).	
Managing respondent burden	Indirect evaluation of response burden: number of questions on the questionnaire	To be evaluated taking into account the complexity of each question, the questionnaire paths and the expected fraction of the sample/population that should fill in each path.
Managing respondent burden	Trend in respondent burden with respect to the previous iteration	For outputs produced on a regular basis.
Managing respondent burden	Is there a communication plan encouraging response by informing potential respondents about the survey and the importance of their contribution?	
Managing respondent burden	The extent to which the respondent can choose among different data collection modes	
Soundness of implementation	Extent to which administrative data collection systems/interfaces are understood and specified.	
Soundness of implementation	Extent to which administrative data integration techniques are understood and specified, both for direct and indirect use of ADS.	

2.4. Design frame and sample

52. This sub-process only applies to processes which involve data collection based on sampling, such as through statistical surveys. It identifies and specifies the population of interest, defines a sampling frame (and, where necessary, the register from which it is derived), and determines the most appropriate sampling criteria and methodology (which could include complete enumeration). Common sources for a sampling frame are administrative and statistical registers, censuses and information from other sample surveys. This sub-process describes how these sources can be combined if needed. Analysis of whether the frame covers the target population should be performed. A sampling plan should be made: The actual sample is created in sub-process 4.1 (Create frame and select sample), using the methodology, specified in this sub-process.

Quality Dimension	**Indicator**	**Notes**
Methodological soundness	Extent to which the survey population matches the target population	See also phase 4 “collect”
Methodological soundness	Timeliness of the frame: how recently was the frame last updated?	See also phase 4 “collect”
Methodological soundness	Impact of coverage errors: assess the likely impact of coverage error on key estimates.	See also phase 4 “collect” ESS QPI - A2. Over-coverage - rate
Methodological soundness	Key indicators for sample design (e.g. estimated size, expected/planned sampling errors for key variables, domains, costs)	See also phase 4 “collect” ESS QPI - A1. Sampling error - indicators
Methodological soundness	Feasibility of estimation (e.g. a complex sample design might force the use of bootstrap variance estimation while a simpler design might not be as efficient but the design based variance might be more desirable)	See also phase 4 “collect”
Methodological soundness	Do unique identification numbers for statistical units exist?	

2.5. Design processing and analysis

53. This sub-process designs the statistical processing methodology to be applied during the "Process" and "Analyse" phases. This can include specification of routines for coding, editing, imputing, estimating, integrating, validating and finalizing data sets.

Quality Dimension	**Indicator**	**Notes**
Cost effectiveness	To what extent is the process planning to re-use systems for coding, E&I, data integration, weighting, estimation	
Soundness of implementation	To what extent is the business process using standard or well-known methods for subsequent phases (e.g. coding, E&I, data integration, weighting, estimation, revision), in a transparent way?	See also phase 5 and 6 yes/partly/no indicator Corresponds to the appropriate statistical procedures principle in the ES Code of Practice
Soundness of	When have the methodologies for	See also phase 5 and 6

implementation	subsequent phases (e.g. coding, E&I, data integration, weighting, estimation, etc.) last been assessed?	for outputs produced on a regular basis Corresponds to the appropriate statistical procedures principle in the ES Code of Practice
Soundness of implementation	Specifications for coding, editing, imputing, estimation, integrating, validating and finalizing datasets take into consideration the type of data being processes, i.e. respondent data or ADS or a combination.	

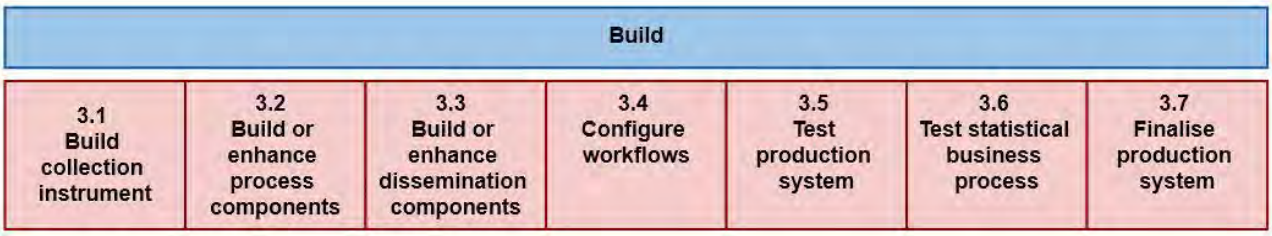
2.6. Design production systems and workflow

54. This sub-process determines the workflow from data collection to dissemination, taking an overview of all the processes required within the whole statistical production process, and ensuring that they fit together efficiently with no gaps or redundancies. Various systems and databases are needed throughout the process. A general principle is to reuse processes and technology across many statistical business processes, so existing production solutions (e.g. services, systems and databases) should be examined first, to determine whether they are fit for purpose for this specific process, then, if any gaps are identified, new solutions should be designed. This sub-process also considers how staff will interact with systems, and who will be responsible for what and when.

Quality Dimension	**Indicator**	**Notes**
Soundness of implementation	Percentage of identified and documented GSBPM processes (with sub-processes) with their flows	Corresponds to the appropriate statistical procedures principle in the ES Code of Practice
Soundness of implementation	Specifications for production systems and workflow take into consideration the type of data being processed (respondent data or ADS or a combination).	
Cost effectiveness	Percentage of/Extent to which corporate solutions (e.g. tools, processes, technologies) are reused in subsequent phases and sub-processes	
Cost effectiveness	Percentage of/Extent to which responsibilities for subsequent phases and sub-processes have been set	
Cost effectiveness	Estimated cost for producing and disseminate designed outputs/Key Performance Indicators (KPIs)	
Accuracy and reliability	Percentage of/ Extent to which quality indicators are planned to be calculated for subsequent sub-processes of GSBPM	
Accuracy and reliability	Amount/percentage of quality indicators used as KPIs	
Timeliness and Punctuality	Planned time frame for subsequent phases and sub-processes	ESS QPI- TP2. Time lag - final results
Accessibility and clarity	The number of social media visitors/followers	

Accessibility and clarity	Metadata - consultations Number of metadata consultations (ESMS) within a statistical domain for a given time period. This indicator is applicable: • to all statistical processes; • to producers	ESS QPI - AC2. Metadata - consultations By "number of consultations" it is meant the number of times a metadata file is viewed. Some information is available through the monthly Monitoring report on Eurostat Electronic Dissemination.
Accessibility and clarity	Number of consultations of data tables within a statistical domain for a given time period	ESS QPI - AC1. Data tables – consultations

Build Phase



55. This phase builds and tests the production solution to the point where it is ready for use in the "live" environment. The outputs of the "Design" phase direct the selection of reusable processes, instruments, information, and services that are assembled and configured in this phase to create the complete operational environment to run the process. New services are built by exception, created in response to gaps in the existing catalogue of services sourced from within the organisation and externally. These new services are constructed to be broadly reusable within the statistical production architecture.

56. For statistical outputs produced on a regular basis, this phase usually occurs for the first iteration, and also following a review or a change in methodology or technology, rather than for every iteration.

57. It is broken down into seven sub-processes, which are generally sequential, from left to right, but can also occur in parallel, and can be iterative. These sub-processes are:

3.1. Build collection instrument

58. This sub-process describes the activities to build the collection instruments to be used during the "Collect" phase. The collection instrument is generated or built based on the design specifications created during the "Design" phase. A collection may use one or more modes to receive the data, e.g. personal or telephone interviews; paper, electronic or web questionnaires; SDMX hubs. Collection instruments may also be data extraction routines used to gather data from existing statistical or administrative data sets. This sub-process also includes preparing and testing the contents and functioning of that instrument (e.g. testing the questions in a questionnaire). It is recommended to consider the direct connection of collection instruments to the statistical metadata system, so that metadata can be more easily captured in the collection phase. Connection of metadata and data at the point of capture can save work in later phases. Capturing the metrics of data collection (paradata) is also an important consideration in this sub-process.

Quality Dimension	Indicator	Notes
Soundness of implementation	Has the questionnaire been tested using appropriate methods (e.g. questionnaire pre-test, pilot in real situation, in depth - interviews, focus groups, interviewer support, etc.)?	Corresponds to the appropriate statistical procedures principle in the ES Code of Practice
Soundness of implementation	Have the test results been taken into account in the process of implementing the final questionnaire, and documented in a report?	Corresponds to the appropriate statistical procedures principle in the ES Code of Practice
Soundness of implementation	Has the data collection tool/instrument (e.g. electronic questionnaire, acquisition web site, SDMX hub) been tested and how?	This indicator refers to the tests of the IT instruments used for data collection (e.g. functionality test, stress test…) Corresponds to the appropriate statistical procedures principle in the ES Code of Practice
Soundness of implementation	To what extent have the test results been taken into account in the process of implementing the final data collection tools	Corresponds to the appropriate statistical procedures principle in the ES Code of Practice
Soundness of implementation	Have administrative data collection systems/interfaces been tested and how?	
Managing respondent burden	Estimated reporting burden (e.g. the time needed to: obtain internal or external expertise; retrieve the required information; handle sensitive information; and answer the questionnaire.)	
Managing respondent burden	Estimated response time (i.e. the interview length)	Can be a proxy indicator of respondent burden
Managing respondent burden	Percentage of questions used to collect information which will not be published (and motivation).	See also 2.3
Managing respondent burden	Trend in respondent burden with respect to the previous iteration (for outputs produced on a regular basis)	See also 2.3
Accuracy and reliability	If mixed or multiple data collection modes are adopted, has the mode effect on data quality been tested?	
Accuracy and reliability	Have the test results been taken into account in the process of implementing the final data collection modes?	

Accuracy and reliability	Extent to which paradata can be captured at the data collection stage?	The collection instrument(s) should allow for capturing paradata to be used for quality assessment
Accessibility and clarity	Extent to which metadata can be captured at the data collection stage and stored in metadata management systems?	The collection instrument(s) should allow for capturing metadata at an early stage
Managing metadata	Do collection instruments capture what is needed to create variables agreed upon in design phase?	See also 2.2 Yes/No indicators; there could be one for each variable and classification. Corresponds to the accessibility and clarity principle in the ES Code of Practice
Managing metadata	Do collection instruments allow for coding to the lowest level of the classifications agreed upon in design phase?	See also 2.2 Yes/No indicators; there could be one for each variable and classification. See also 2.2 Corresponds to the accessibility and clarity principle in the ES Code of Practice.

3.2. Build or enhance process components

59. This sub-process describes the activities to build new and enhance existing components and services needed for the “Process” and “Analyse” phases, as designed in the "Design" phase. Services may include dashboard functions and features, information services, transformation functions, workflow frameworks, provider and metadata management services.

Quality Dimension	**Indicator**	**Notes**
Soundness of implementation	What proportion of functions in the statistical process are built using corporately supported software tools, components or services?	See also 2.5 Yes/No indicator. Corresponds to the appropriate statistical procedures principle in the ES Code of Practice.
Soundness of implementation	Have Enterprise Architecture best practices for software development been followed?	See also 2.5 Yes/No indicator. Corresponds to the appropriate statistical procedures principle in the ES Code of Practice.
Soundness of implementation	Has testing been done throughout the Building process?	See also 2.5 Yes/No indicator. Corresponds to the appropriate statistical procedures principle in the ES Code of Practice.

Soundness of implementation	Have corporate requirements for dashboards and information services been incorporated?	See also 2.5 Yes/No indicator. Corresponds to the appropriate statistical procedures principle in the ES Code of Practice.
Soundness of implementation	Was the testing strategy designed when the process and its components were designed?[4]	Corresponds to the appropriate statistical procedures principle in the ES Code of Practice.
Soundness of implementation	Was additional testing done by someone other than the person(s) who did the programming?	Yes/No indicator on testing if software and IT tools are working properly and not affecting quality/introducing errors. Corresponds to the appropriate statistical procedures principle in the ES Code of Practice.
Soundness of implementation	Were the different types of testing designed, executed, documented and signed-off: Functional testing; volume testing; stress testing; (end-to-end testing and user testing to be done in 3.5 Test production system)	Yes/No indicator on testing if software and IT tools are working properly and not affecting quality/introducing errors. Corresponds to the appropriate statistical procedures principle in the ES Code of Practice.
Soundness of implementation	Was testing done specifically to ensure that the software produces the correct results?	Yes/No indicator on testing if software and IT tools are working properly and not affecting quality/introducing errors. Corresponds to the appropriate statistical procedures principle in the ES Code of Practice. This could be either a parallel run, or in the case where a parallel run is impossible, testing against known results, for example from a prototype or simulations of theoretical results.
Soundness of implementation	Extent to which process components that have complete documentation, support staff, and user training, all available at the same time that the software is put into use.	Yes/No indicator on testing if software and IT tools are working properly and not affecting quality/introducing errors. Corresponds to the appropriate statistical procedures principle in the ES Code of Practice.

[4] Definition from the LEG on Quality (2001): **Efficiency**: produces the desired outcomes cost efficiently; **Effectiveness**: successful in delivering the desired outcomes; **Robustness**: delivers results against challenging demands; **Flexibility**: readily adaptable to changing needs and demands; **Transparency**: open, visible and easily understood and **Integration**: complementary and consistent, both with other processes, and with meeting business needs

Accuracy and reliability	Has the quality of the data after the test of the coding procedure been assessed (e.g. quality indicators such as "recall rate" have been calculated)? The recall rate is calculated as the ratio between the number of values automatically coded and the total number of values submitted to coding.	This is an indicator of the quality of the data obtained by the coding procedure. Indicator of the efficacy of the automated coding procedure
Accuracy and reliability	Have the assessment results been taken into account in the implementation of the final procedure?	
Accuracy and reliability	Has the output of the E&I procedure been assessed? (e.g. by simulation and by calculating indicators, analysing distributions.)	
Accuracy and reliability	Have the assessment results been taken into account in the implementation of the finale procedure?	
Soundness of implementation	Have the process components necessary to manage processing of large data sets been tested and how?	
Accuracy and reliability	Have process components for data linkage been tested and fine-tuned?	

3.3. Build or enhance dissemination components

60. This sub-process describes the activities to build new and enhance existing components and services needed for the dissemination of statistical products as designed in sub-process 2.1 (Design outputs). All types of dissemination components and services are included, from those that are used to produce traditional paper publications to those that provide web services, open data outputs, or access to micro-data.

Quality Dimension	**Indicator**	**Notes**
Managing metadata	Extent to which relevant metadata can be linked to output data.	See also phase 7 Corresponds to the accessibility and clarity principle in the ES Code of Practice
Accessibility and clarity	Extent to which user requirements are fulfilled in terms of e.g. dissemination formats, information systems, graphical supports.	See also phase 7

3.4. Configure workflows

61. This sub-process configures the workflow, systems and transformations used within the statistical business processes, from data collection through to dissemination. It ensures that the workflow specified in sub-process 2.6 (Design production systems and workflow) works in practice.

Quality Dimension	Indicator	Notes
Soundness of implementation	Ratio of the number of sub-processes automated through an IT tool to the total number of sub-processes specified in 2.6	This quality indicator assumes that processes have been specified in Business Process Modelling Notation (BPMN) or using another tool in 2.6 Corresponds to the appropriate statistical procedures principle in the ES Code of Practice
Timeliness and punctuality	Planned timeliness of all subsequent phases and sub-processes	See also 2.6 ESS QPI- TP2. Time lag – final results

3.5. Test production system

62. This sub-process is concerned with the testing of assembled and configured services and related workflows. It includes technical testing and sign-off of new programmes and routines, as well as confirmation that existing routines from other statistical business processes are suitable for use in this case. Whilst part of this activity concerning the testing of individual components and services could logically be linked with sub-process 3.2 (Build or enhance process components), this sub-process also includes testing of interactions between assembled and configured services, and ensuring that the production solution works as a coherent set processes, information and services.

Quality Dimension	Indicator	Notes
Soundness of implementation	Have all programmes, routines and configured services been individually tested and signed off prior to the start of testing the production system?	While 3.2 refers to newly built programmes and routines, 3.5 includes testing new and previously existing programmes and routines.
Soundness of implementation	Has the entire production system been tested and signed off, ensuring that data correctly enters and exits each programme, routine and configured service, and that the functionality of each programme, routine and configured service has been executed according to expectations?	This assumes that there is a business standard in the statistical agency for the system testing. Corresponds to the appropriate statistical procedures principle in the ES Code of Practice. The entire production system, including processing of coding, editing, imputation, integration, validation and finalisation of datasets, takes into consideration the type of data being processed (i.e. respondent data or ADS or a combination thereof)
Soundness of implementation	Has the quality of the linkage procedures been tested and signed off?	

Soundness of implementation	Has the building of statistical units been tested and signed off?	The use of administrative data may require the conversion of administrative units to statistical units. In the same manner, when data integration occurs, the linked units may need to be converted to statistical units. This indicator ensures that building of statistical units in these scenarios is robust during statistical production.

3.6. Test statistical business process

63. This sub-process describes the activities to manage a field test or pilot of the statistical business process. Typically it includes a small-scale data collection, to test collection instruments, followed by processing and analysis of the collected data, to ensure the statistical business process performs as expected. Following the pilot, it may be necessary to go back to a previous step and make adjustments to instruments, systems or components. For a major statistical business process, e.g. a population census, there may be several iterations until the process is working satisfactorily.

Quality Dimension	**Indicator**	**Notes**
Cost effectiveness	Estimated costs for producing and disseminating outputs and divergences from planned costs in design phase	See also 2.6
Accuracy and reliability	Pilot has been carried out and results have been taken into account in final implementation Dimension of the test/field pilot compared to real survey	
Accuracy and reliability	Assessment of major error sources from the Pilot (e.g. coverage, nonresponse, measurement, and process errors)	ESS QPI – A2.Over-coverage rate A4. Unit non-response rate
Timeliness and punctuality	Estimated time frame for subsequent phases and sub-processes and divergences from planned one in design phase	See 2.6 ESS QPI - TP2. Time lag – final results
Soundness of implementation	Testing of formats and timetable for acquiring ADS has been done and taken into account in final implementation.	

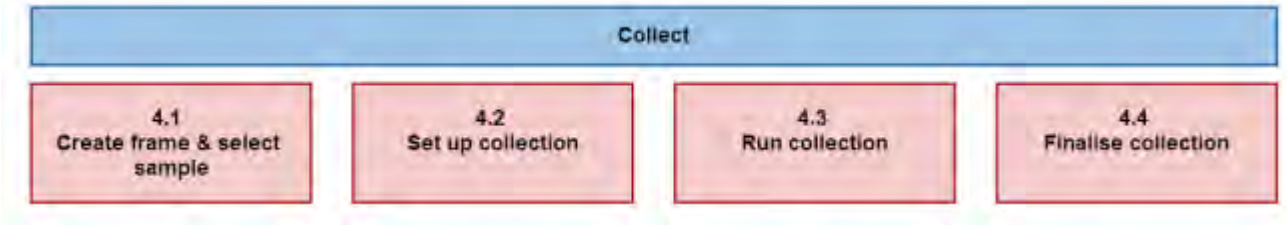
3.7. Finalise production systems

64. This sub-process includes the activities to put the assembled and configured processes and services, including modified and newly-created services into production ready for use by business areas. The activities include:

- producing documentation about the process components, including technical documentation and user manuals
- training the business users on how to operate the process
- moving the process components into the production environment, and ensuring they work as expected in that environment (this activity may also be part of sub-process 3.5 (Test production system)).

Quality Dimension	Indicator	Notes
Accessibility and Clarity	Percentage of materials adequately archived (e.g. easily retrievable; properly labelled; retention period indicated)	

Collect Phase



65. This phase collects or gathers all necessary information (data and metadata), using different collection modes (including extractions from statistical, administrative and other non-statistical registers and databases), and loads them into the appropriate environment for further processing. Whilst it can include validation of data set formats, it does not include any transformations of the data themselves, as these are all done in the "Process" phase. For statistical outputs produced regularly, this phase occurs in each iteration.

66. The "Collect" phase is broken down into four sub-processes, which are generally sequential, from left to right, but can also occur in parallel, and can be iterative. These sub-processes are:

4.1. Create frame and select sample

67. This sub-process establishes the frame and selects the sample for this iteration of the collection, as specified in sub-process 2.4 (Design frame and sample). It also includes the coordination of samples between instances of the same statistical business process (for example to manage overlap or rotation), and between different processes using a common frame or register (for example to manage overlap or to spread response burden). Quality assurance and approval of the frame and the selected sample are also undertaken in this sub-process, though maintenance of underlying registers, from which frames for several statistical business processes are drawn, is treated as a separate business process. The sampling aspect of this sub-process is not usually relevant for processes based entirely on the use of pre-existing sources (e.g. administrative sources) as such processes generally create frames from the available data and then follow a census approach.

Quality Dimension	Indicator	Notes
Accuracy and reliability	The rate of over-coverage: The proportion of units accessible via the frame that do not belong to the target	Need auxiliary data to assess coverage; often cannot assess coverage until after collection

	population (are out-of-scope). The rate of over-coverage is applicable: • to all statistical processes (including use of administrative sources); • to producers. If the survey has more than one unit type, a rate may be calculated for each type. If there is more than one frame or if over-coverage rates vary strongly between sub-populations, rates should be separated.	has happened. ESS QPI - A2. Over-coverage – rate
Accuracy and reliability	Rate of duplicate records identified and corrected during frame creation	
Accuracy and reliability	Rate of missing or suspicious stratification and classification variables; rate of missing contact variables; time elapsed since last successful contact	
Accuracy and reliability	Relative discrepancy between expected and observed sample size; relative discrepancy between expected and observed response, attrition and out of scope rates	Can only be assessed after collection is finished ESS QPI - A4. Unit non-response - rate.
Accuracy and reliability	The sampling error can be expressed: • in relative terms, in which case the relative standard error or, synonymously, the coefficient of variation (CV) is used. • in terms of confidence intervals. Sampling errors indicators are applicable: • to statistical processes based on probability samples or other sampling procedures allowing computation of such information. • to users and producers, with different level of details given.	ESS QPI - A1. Sampling error - indicators
Timeliness and punctuality	Delay between expected and actual creation of frame	It can include the delay due to the acquisition of administrative data sources
Timeliness and punctuality	Delay between expected and actual creation of sample	
Managing respondent burden	Extent to which administrative data supplement direct collection (e.g. % of records from ADS; % of variables from ADS)	
Accuracy and reliability	Extent to which ADS are used to create/ maintain the frame	
Accuracy and reliability	Extent to which ADS are used as auxiliary variables to be used in the construction of the sampling designs	

4.2. Set up collection

68. This sub-process ensures that the people, processes and technology are ready to collect data and metadata, in all modes as designed. It takes place over a period of time, as it includes the strategy, planning and training activities in preparation for the specific instance of the statistical business process. Where the process is repeated regularly, some (or all) of these activities may not be explicitly required for each iteration. For one-off and new processes, these activities can be lengthy. This sub-process includes:

- preparing a collection strategy;
- training collection staff;
- ensuring collection resources are available e.g. laptops;
- agreeing terms with any intermediate collection bodies, e.g. sub-contractors for computer assisted telephone interviewing
- configuring collection systems to request and receive the data;
- ensuring the security of data to be collected;
- preparing collection instruments (e.g. printing questionnaires, pre-filling them with existing data, loading questionnaires and data onto interviewers' computers etc.).

69. For non-survey sources, this sub-process will include ensuring that the necessary processes, systems and confidentiality procedures are in place, to receive or extract the necessary information from the source.

Quality Dimension	**Indicator**	**Notes**
Statistical Confidentiality and security	Risk of a breach while data is being transferred (survey and ADS)	
Adequacy of resources	Rate of HR requirements fulfilled; rate of IT requirements fulfilled	
Adequacy of resources	Success rate for collection staff to perform collection tasks after having been trained	Test collection staff before and after training to assess effectiveness
Soundness of implementation	Success rate for testing collection systems, under expected as well as high volume and extreme situations (survey and ADS)	End to end system testing. Corresponds to the appropriate statistical procedures principle in the ES Code of Practice.
Timeliness and punctuality	Delay between expected and actual sign-off of collection systems (including data transmission, security, collection management systems, and quality control systems) (survey and ADS)	
Timeliness and punctuality	Delay between expected and actual sign-off of collection materials (questionnaire, training materials, etc.)	

4.3. Run collection

70. This sub-process is where the collection is implemented, with the different instruments being used to collect or gather the information, which may include raw micro-data or aggregates produced at the source, as well as any associated metadata. It includes the initial contact with providers and any subsequent follow-up or reminder actions. It may include manual data entry at the point of contact, or fieldwork management, depending on the

source and collection mode. It records when and how providers were contacted, and whether they have responded. This sub-process also includes the management of the providers involved in the current collection, ensuring that the relationship between the statistical organisation and data providers remains positive, and recording and responding to comments, queries and complaints. For administrative and other non-statistical sources, this process is brief: the provider is either contacted to send the information, or sends it as scheduled. When the collection meets its targets, it is closed and a report on the collection is produced. Some basic validation of the structure and integrity of the information received may take place within this sub-process, e.g. checking that files are in the right format and contain the expected fields. All validation of the content takes place in the Process phase.

Quality Dimension	**Indicator**	**Notes**
Managing respondent burden	Are there enough staff responsible for dealing with the respondent's questions?	
Managing respondent burden	Support is provided to respondents (e.g. toll free number).	
Accuracy and reliability	Quality control is used to manage the quality of data collection and data capture processes.	
Accuracy and reliability	Meaningful feedback is provided to interviewers and fieldworkers on a regular basis.	
Accuracy and reliability	Monitoring of fieldwork operations is done during data collection.	
Accuracy and reliability	Interviewer performance is measured for CATI, CAPI, PAPI surveys (e.g. interviewers' productivity).	
Accuracy and reliability	Domain response rates; representativity indicators; achieved CVs of key variables in domains of interest	ESS QPI - A1. Sampling error – indicators A4. Unit non-response - rate
Accuracy and reliability	Unit nonresponse rate; item nonresponse rate; proxy rate	ESS QPI - A4. Unit non-response – rate A5. Item non-response - rate
Accuracy and reliability	Mode effect when more than one collection mode	Can only be assessed after estimation.
Accuracy and reliability	Outgoing error rates; estimate of non-sampling error	Data capture is covered in 4.4
Timeliness and punctuality	Delay between expected and actual start and close of collection	
Accuracy and reliability	Extent to which follow up is based on administrative data	From MIAD (Use ADS to help target follow up); This refers to some of the ideas around 'responsive design' where there may be administrative/auxiliary data about selected sample units that can be used to target follow-up of initial non-respondents to increase the

		likelihood of achieving a 'balanced' (or 'representative') final sample and thereby more effectively utilise what will typically be a limited data collection budget. (There is a corresponding entry in the "design/data collection/construction of sampling designs", too).
Accuracy and reliability	Percentage of data transmitted according to the agreements with administrative data owners (e.g. format, time schedule).	See indicators from (e.g. MIAD, Checklist 2006, Blue-Ets) It includes Quality checks on data transmitted (e.g. completeness).

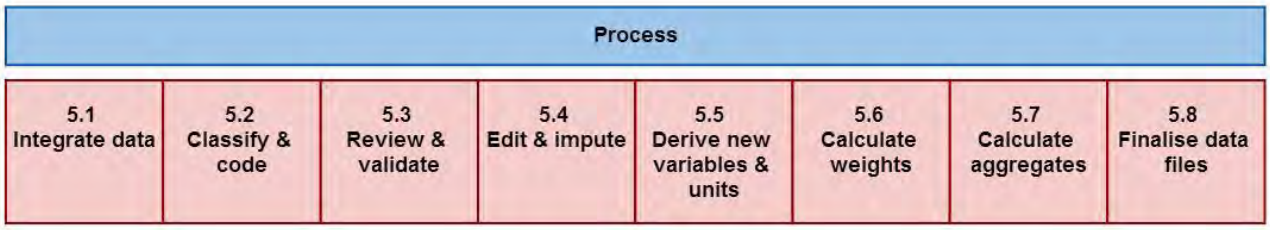
4.4. Finalise collection

71. This sub-process includes loading the collected data and metadata into a suitable electronic environment for further processing. It may include manual or automatic data take-on, for example using clerical staff or optical character recognition tools to extract information from paper questionnaires, or converting the formats of files received from other organisations. It may also include analysis of the process metadata (paradata) associated with collection to ensure the collection activities have met requirements. In cases where there is a physical collection instrument, such as a paper questionnaire, which is not needed for further processing, this sub-process manages the archiving of that material.

Quality Dimension	**Indicator**	**Notes**
Cost-effectiveness	Discrepancy between planned versus actual collection costs Percentage of collection activities that met requirements (assessed through analysis of paradata)	
Accuracy and reliability	Outgoing error rates; estimate of non-sampling error	
Accuracy and reliability	The rate of over-coverage: The proportion of units accessible via the frame that do not belong to the target population (are out-of-scope). The rate of over-coverage is applicable: • to all statistical processes (including use of administrative sources); • to producers. If the survey has more than one unit type, a rate may be calculated for each type. If there is more than one frame or if over-coverage rates vary strongly between sub-	Need auxiliary data to assess coverage; often cannot assess coverage until after collection has happened. ESS QPI - A2. Over-coverage - rate

	populations, rates should be separated.	
Accuracy and reliability	Unit nonresponse rate; item nonresponse rate; proxy rate	ESS QPI - A4. Unit non-response – rate A5. Item non-response - rate
Accessibility and clarity	Percentage of materials adequately archived (easily retrievable; properly labelled; retention period indicated) (survey and ADS)	
Accuracy and reliability	Technical checks on ADS (e.g. 1) controls on the readability of the file (e.g. unknown format, a corrupted file, a file with an unfamiliar character set, or a file that cannot be decoded); 2) File declaration compliance (examples of problems in this area are a file with a missing metadata description and a file with a lay-out that does not comply to the lay-out agreed upon); 3) Convertibility: it focuses on the conversion of the file to the NSO-standard format (examples of problems in this area are file errors while decoding and corrupted data in the file after conversion)	See indicators from (e.g. MIAD, Checklist 2006, Blue-Ets) Pre-processing checks Indicators are drawn from Blue-ETS: technical checks

Process Phase



72. This phase describes the cleaning of data and their preparation for analysis. It is made up of sub-processes that check, clean, and transform input data, so that they can be analysed and disseminated as statistical outputs. It may be repeated several times if necessary. For statistical outputs produced regularly, this phase occurs in each iteration. The sub-processes in this phase can apply to data from both statistical and non-statistical sources.

73. The "Process" and "Analyse" phases can be iterative and parallel. Analysis can reveal a broader understanding of the data, which might make it apparent that additional processing is needed. Activities within the "Process" and "Analyse" phases may commence before the "Collect" phase is completed. This enables the compilation of provisional results where timeliness is an important concern for users, and increases the time available for analysis.

74. This phase is broken down into eight sub-processes, which may be sequential, but can also occur in parallel, and can be iterative. These sub-processes are:

5.1. Integrate data

75. This sub-process integrates data from one or more sources. It is where the results of sub-processes in the "Collect" phase are combined. The input data can be from a mixture of external or internal data sources, and a variety of collection modes, including extracts of administrative data. Administrative data can substitute for all or some of the directly collected survey variables. This sub-process also includes harmonising or creating new figures that agree between sources of data. The result of this sub-process is a set of linked data. Data integration can include:

- combining data from multiple sources, as part of the creation of integrated statistics such as national accounts
- data pooling, with the aim of increasing the effective number of observations of a phenomena
- matching / record linkage routines, with the aim of linking micro or macro data from different sources
- data fusion - integration followed by reduction or replacement
- prioritizing, when two or more sources contain data for the same variable, with potentially different values.

76. Data integration may take place at any point in this phase, before or after any of the other sub-processes. There may also be several instances of data integration in any statistical business process. Following integration, depending on data protection requirements, data may be anonymised, that is, stripped of identifiers such as name and address, to help protect confidentiality.

Quality Dimension	Indicator	Notes
Accuracy and reliability	The proportion of units covered by both the survey and the administrative sources in relation to the total number of units in the survey. The proportion is applicable: • to mixed statistical processes where some variables or data for some units come • from survey data and others from administrative source(s); • to producers.	ESS QPI - A3. Common units - proportion
Accuracy and reliability	Existence of linkage variables (unique identifier) of the register (yes/no question) Linking of microdata to other microdata.	Future development needed in this area, since no agreed upon international/corporate indicators at this point.
Accuracy and reliability	Degree of linkability of the linkage variables (high, medium, low)	Future development needed in this area, since no agreed upon international/corporate indicators at this point.
Accuracy and reliability	Proportion of duplicated records in the linked data	Some datasets may contain erroneous duplicate records that disrupt the linking process and/or the final dataset. Knowing how many units in each dataset are duplicated (or how many are detected as being duplicates) is a useful indicator of the underlying dataset's quality: a very good, well-checked and maintained dataset should have very few duplicates, but one produced for other reasons may not have had the same care taken. If we detect duplicates, they must be resolved in some way. Mistakes in this process can result in errors in the final data, such as when duplicates differ in some variables and we need to choose one set of values.
Accuracy and reliability	Reliability of the linkage results False link and false non-link rates: • False links are record pairs that are deemed to be links but which are actually true non-matches • False non-links are true matches which remain unlinked. In practice they are much harder to identify so it	This indicator will depend on the linkage key, its accuracy in the linked data sets, the validity of the linkage procedure, as well as the comparison between the final resulting linked records and the expected ones.

	may not be very efficient to use the false non-link rate as a regular indicator. Precision and recall: • Precision measures how well the links are made, and is a measure of the goodness of links that are made (it is the complement of the false link rate). • Recall measures how well links are found, and is a measure of how many true links have been captured correctly. It may also be referred as sensitivity of the procedure. These measures are based on false link and false non-link rates.	
Accuracy and reliability	An indicator of the effectiveness of the cut-off weight for determining the threshold of passes in probabilistic matching	
Accuracy and Reliability	Percentage of errors coming from identification and transformation of population, units or data items. • It is possible that the meaning of a population, a statistical unit or data items changes in the course of the process. Errors may occur in this transformation process. • The conversion of one statistical concept into another. For example; measurement units for imported and exported products collected from administrative sources could be different from the measurement units for statistically required data. This type of errors should be measured during the integration of data.	
Accuracy and reliability	Rates of unit change from period to period Birth rate: Ratio of the number of unit births in output period to the total number of units Death rate: Ratio of the number of unit deaths in output period to the total number of units	For many statistical outputs, the target population changes relatively slowly, so any significant changes in the units in input datasets may indicate quality problems with the data, linking, or other aspects of the process. This indicator measures the rate of change in the population.
Accuracy and reliability	Proportion of units that may belong to more than one composite unit	Creating the list of statistical units in a final output dataset may require connecting together

	This indicator records how often a base unit (e.g. a person) doesn't have a single clear composite unit to which it can be assigned without doubt. This could be units that can't be assigned to any composite unit for some reason, or units equally likely to belong to two different composite units.	'base' units (the lowest-level units created during the linking process, or the units on the input datasets). For example, to create a list of household units we might link several input datasets to produce a list of base person units, then link all the person units with the same address to form composite household units.
Methodological soundness	If record linkage is required, report linkage methodology (e.g. exact, probabilistic, etc.) used.	
Accuracy and reliability	Linkage rate - Proportion of units linked from each dataset to a base dataset, or percentage link rates between pairs of datasets	The link rate is a very important measure of the quality of the linked sets (sets of objects from different datasets that were linked together). This measures the proportion of objects in each dataset that can be connected with units in the other datasets. A low link rate may indicate that different datasets cover different parts of the population, or that the linking process is not identifying all the connections that exist between the objects in each dataset.
Accuracy and reliability	Proportion of manually linked units	
Accuracy and reliability	Macro-level comparisons of the distribution of linked objects with reference distributions	By comparing the distribution of units across key variables in the linked datasets with those in a reference dataset, we determine whether the linked set is missing important parts of the target population, or at least whether it represents the target population well.
Accuracy and reliability	Distribution of variables in linked data	Analysis can be done to compare the values of variables in different datasets at various levels of aggregation. This measure isn't easy to give general rules for but we can use statistical data inspection methods to compare measures of variables present in at least two datasets. Histograms and scatter plots are graphical methods to use. Numerical measures such as means, medians, standard deviations, and skewness also help compare distributions of values.

5.2. Classify and code

77. This sub-process classifies and codes the input data. For example automatic (or clerical) coding routines may assign numeric codes to text responses according to a pre-determined classification scheme.

Quality Dimension	Indicator	Notes
Methodological soundness	Compliance rate of classifications of input data to the pre-determined standard international classification and national versions of international classification scheme All international or national classifications and breakdowns which are used for the data set are produced (i.e. although NACE Rev2 is introduced as international classification, using the older version or using a different classification than the proposed classification).	
Methodological soundness	Compliance rate of coding of input data to the pre-determined standard coding scheme	The standard coding scheme in this indicator refers to the compliance with the local codes used in these variables.
Accuracy and reliability	Ratio between the number of values automatically coded and the total number of values submitted to coding.	It measures the efficiency of the automatic coding procedure.
Accuracy and reliability	Extent to which quality control is used to manage the quality of automated and manual coding processes	
Timeliness and punctuality	Delay between expected and actual timing of adaptation of correspondence tables	
Methodological soundness	Proportion of statistical units which cannot clearly be classified or mapped.	This measures the effectiveness of the classification rule.

5.3. Review and validate

78. This sub-process examines data to try to identify potential problems, errors and discrepancies such as outliers, item non-response and miscoding. It can also be referred to as input data validation. It may be run iteratively, validating data against predefined edit rules, usually in a set order. It may flag data for automatic or manual inspection or editing. Reviewing and validating can apply to data from any type of source, before and after integration. Whilst validation is treated as part of the "Process" phase, in practice, some elements of validation may occur alongside collection activities, particularly for modes such as web collection. Whilst this sub-process is concerned with detection of actual or potential errors, any correction activities that actually change the data are done in sub-process 5.4.

Quality Dimension	Indicator	Notes
Accuracy and reliability	Rate of actual errors Identification of incorrect data (actual errors) in the processing stage - missing, invalid or inconsistent entries or that point out data records that are actually in error.	

5.4. Edit and impute

79. Where data are considered incorrect, missing or unreliable, new values may be inserted in this sub-process. The terms editing and imputation cover a variety of methods to do this, often using a rule-based approach. Specific steps typically include:

- the determination of whether to add or change data;
- the selection of the method to be used;
- adding / changing data values;
- writing the new data values back to the data set, and flagging them as changed;
- the production of metadata on the editing and imputation process.

Quality Dimension	Indicator	Notes
Accuracy and reliability	Imputation rate • The indicator is expressed as the ratio of the number of replaced values to the total number of values for a given variable. • The imputation rate is applicable: o to all statistical processes (with micro data (e.g. direct data collection and administrative data); o to producers. Information on the extent to which imputation is used and the reasons for it should be noted. A short description of the methods used and their effects on the estimates. This indicator is influenced by both the item non-response and the editing process. It measures both the relative amount of imputed values and the relative influence on the final estimates from the imputation procedures. The unweighted imputation rate for a variable is the ratio of the number of imputed values to the total number of values expected of the variable.	The unweighted rate shows, for a particular variable, the proportion of units for which a value has been imputed due to the original value being a missing, implausible, or inconsistent value in comparison with the number of units with a value for this variable. The weighted rate shows, for a particular variable, the relative contribution of imputed values to the estimate of this item/variable. Obviously this weighted indicator is meaningful when the objective of a survey is that of estimating the total amount or the average of a variable. When the objective of the estimation is that of estimating complex indices, the weighted indicator is not meaningful. ESS QPI - A7. Imputation rate

	The weighted rate shows the relative contribution of imputed values to a statistic, typically, a total for a quantitative variable. For a qualitative variable, the relative contribution is based on the number of units with an imputed value for the qualitative item.	
Accuracy and reliability	Extent to which administrative data was used for imputation.	
Accuracy and reliability	An indicator of an edit's effectiveness would be the rate of false negative or false positive assessments.	One way to verify this would be to re-interview the respondents of a sample of units to confirm the reported values, and see what proportion of true values were flagged as errors and what proportion of errors were not flagged as errors.
Accuracy and reliability	Edit failure rates can be calculated for key variables and by domains of interest. A sub-class of edits could be those designed to detect outlier observations.	A high/very high edit failure rate for a given variable would suggest possible errors in previous phases (e.g. in the questionnaire or in data collection).
Accuracy and reliability	Rate of robustness of outliers for key variables Robustness of Outliers = Corrected/Discarded Outliers / Total detected outliers This indicator will measure the quality of an outlier detection process	
Accuracy and reliability	Extent to which administrative data was used to determine the consistency of observations from the survey data. (Or vice versa)	Emigration is an example where surveys can be more accurate than ADS, such as population registers.
Accuracy and reliability	Proportion of units with conflicting information Methods for determining if units have conflicting information vary. The simplest measure is to record how many fixes or decisions we made (either manually or automatically) during processing. For example, in linking person-level data from multiple datasets each dataset may have its own address, sex, or date of birth fields. Even a very high probability match may have a disagreement – because of different definitions, reporting periods, or other mistakes.	The focus of this indicator is not on the accuracy of the variables, but the consistency of the units created by linking different datasets. It measures how reliably the units represent the underlying target population units.

5.5. Derive new variables and statistical units

80. This sub-process derives data for variables and units that are not explicitly provided in the collection, but are needed to deliver the required outputs. It derives new variables by applying arithmetic formulae to one or more of the variables that are already present in the dataset, or applying different model assumptions. This activity may need to be iterative, as some derived variables may themselves be based on other derived variables. It is therefore important to ensure that variables are derived in the correct order. New units may be derived by aggregating or splitting data for collection units, or by various other estimation methods. Examples include deriving households where the collection units are persons, or enterprises where the collection units are legal units.

Quality Dimension	Indicator	Notes
Accuracy and reliability	The rate of over-coverage: The proportion of units that, after integration and derivation are classified as belonging to the target population but indeed do not belong to the target population (are out-of-scope). If there is more than one unit type, a rate may be calculated for each type. If the over-coverage rates vary strongly between sub-populations, rates should be separated.	Need auxiliary data to assess coverage. This indicator is especially suggested when the target population is obtained after linking or deriving units.
Coherence and comparability	Rate of comparability for derived variables Definitions, classifications and units of derived variables will be taken as reference for the comparability and coherence checks.	

5.6. Calculate weights

81. This sub process creates weights for unit data records according to the methodology created in sub-process 2.5 (Design processing and analysis). In the case of sample surveys, weights can be used to "gross-up" results to make them representative of the target population, or to adjust for non-response in total enumerations. In other situations, variables may need weighting for normalization purposes.

Quality Dimension	Indicator	Notes
Accuracy and reliability	The weights are adjusted for coverage and non-response error (yes/no indicator)	See also 2.5

5.7. Calculate aggregates

82. This sub-process creates aggregate data and population totals from micro-data or lower-level aggregates. It includes summing data for records sharing certain characteristics, determining measures of average and dispersion, and applying weights from sub-process 5.6

to derive appropriate totals. In the case of sample surveys, sampling errors may also be calculated in this sub-process, and associated to the relevant aggregates.

Quality Dimension	**Indicator**	**Notes**
Accuracy and reliability	The sampling error can be expressed: a) in relative terms, in which case the relative standard error or, synonymously, the coefficient of variation (CV) is used. b) in terms of confidence intervals. Sampling errors are applicable: • to statistical processes based on probability samples or other sampling procedures allowing computation of such information. • - to users and producers, with different level of details given.	This indicator is also included in 4.1. In 4.1 the data has not been so the actual indicator cannot be calculated, but can be calculated in 5.7). ESS QPI - A1. Sampling error - indicators
Accuracy and reliability	The following indicators are proposed to analyze revisions: 1) Mean Absolute Revision (MAR) is the average of absolute revisions over a time period (useful to analyze stability in terms of size). 2) Relative Mean Absolute Revisions (RMAR) is the relative average of absolute revisions over a time period (useful for comparisons and to analyze levels 3) Mean Revision (MR) is the average of revisions over a time period (useful to analyze directions in terms of sign) and its significance (Yes/No). Standard Deviation of Revisions (SDR) is a measure of the variability of the revisions.	The "revision" is defined as the difference between a later and an earlier estimate of the key item. The proposed indicators are the OECD core/basic measures. Indicators MR, MAR and RMAR are Eurostat Quality and Performance Indicators included in the "A6. Data revision - average size" group. A t-test to assess if MR is significantly different from 0 exists and the second indicator proposed in 3) results from it. ESS QPI - A6. Data revision - average size.
Accuracy and reliability	Extent to which administrative data was used to create population benchmarks.	
Accuracy and reliability	Extent to which administrative data provided auxiliary information for estimators.	
Accuracy and reliability	Extent to which administrative data was used for revision.	

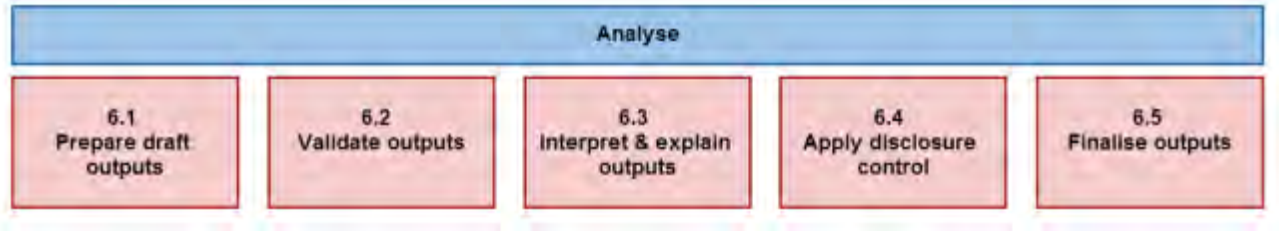
5.8. Finalise data files

83. This sub-process brings together the results of the other sub-processes in this phase

and results in a data file (usually of macro-data), which is used as the input to the "Analyse" phase. Sometimes this may be an intermediate rather than a final file, particularly for business processes where there are strong time pressures, and a requirement to produce both preliminary and final estimates.

Quality Dimension	Indicator	Notes
Timeliness and Punctuality	Delay between expected and actual finalized data file	

Analyse Phase



84. In this phase, statistical outputs are produced, examined in detail and made ready for dissemination. It includes preparing statistical content (including commentary, technical notes, etc.), and ensuring outputs are "fit for purpose" prior to dissemination to customers. This phase also includes the sub-processes and activities that enable statistical analysts to understand the statistics produced. For statistical outputs produced regularly, this phase occurs in every iteration. The "Analyse" phase and sub-processes are generic for all statistical outputs, regardless of how the data were sourced.

85. The "Analyse" phase is broken down into five sub-processes, which are generally sequential, from left to right, but can also occur in parallel, and can be iterative. The sub-processes are:

6.1. Prepare draft outputs

86. This sub-process is where the data are transformed into statistical outputs. It includes the production of additional measurements such as indices, trends or seasonally adjusted series, as well as the recording of quality characteristics.

Quality Dimension	Indicator	Notes
Soundness of implementation	To what extent is the business process using standard or well-known methods (e.g. calculating indices, trends, seasonal adjustment)?	Corresponds to the appropriate statistical procedures principle in the ES Code of Practice.
Accuracy and reliability	Quality Control methods can be applied to ensure that the accuracy of the transformation process itself is sufficient. Indicators could be percentage of outputs reviewed (manually or automated), percentage of errors detected.	
Accuracy and reliability	If the target of estimation is model-based, provide the following: • Model assumptions and associated errors • Non-sampling error being	A short description of the methods used and their effects on the estimates.

	treated or adjusted • For domain specific models, describe the model used and the assessment of validity of the data that had been undertaken. Model assumption errors are errors caused by models used. Models are based on assumptions (see Statistics Netherlands' reports). Model assumption errors occur with 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 (see OECD Glossary). In case of model based seasonal adjustment, indicators include autocorrelation test, seasonal autocorrelation test, skewness, kurtosis and normality test for model residuals provides the opportunity of checking model assumptions satisfied such as Best Linear Unbiased Estimator. Another example of model-based estimation is Small Area Estimation, which is estimation of key variables for small domains. Sample diagnostics include Haussman test and residual-based test depends on the model used.	
Timeliness and punctuality	Delay between the anticipated and actual completion of this step.	

6.2. Validate outputs

87. This sub-process is where statisticians validate the quality of the outputs produced, in accordance with a general quality framework and with expectations. This sub-process also includes activities involved with the gathering of intelligence, with the cumulative effect of building up a body of knowledge about a specific statistical domain. This knowledge is then applied to the current collection, in the current environment, to identify any divergence from expectations and to allow informed analyses. Validation activities can include:

- checking that the population coverage and response rates are as required;
- comparing the statistics with previous cycles (if applicable);
- checking that the associated metadata and paradata (process metadata) are present and in line with expectations

- confronting the statistics against other relevant data (both internal and external);
- investigating inconsistencies in the statistics;
- performing macro editing;
- validating the statistics against expectations and domain intelligence.

Quality Dimension	Indicator	Notes
Accuracy and reliability	Proportion of overall budget dedicated to validation activities; number of validation measures applied	As an example of validation measure the indicator "Asymmetry for mirror flows statistics" can be calculated (ESS QPI- CC1. Asymmetry for mirror flows statistics - coefficient)
Accuracy and reliability	Number or amount of changes made to the data based on validation results	
Coherence and comparability	Availability of backcasting procedures where there is a break in the series	
Coherence and comparability	Degree of coherence with other sources, with provisional data, with quick estimates, and with previous results of the same process	

6.3. Interpret and explain outputs

88. This sub-process is where the in-depth understanding of the outputs is gained by statisticians. They use that understanding to interpret and explain the statistics produced for this cycle by assessing how well the statistics reflect their initial expectations, viewing the statistics from all perspectives using different tools and media, and carrying out in-depth statistical analyses.

Quality Dimension	Indicator	Notes
Accuracy and reliability	Proportion of overall budget dedicated to interpretation and explanation activities; extent to which a report is produced and accepted	

6.4. Apply disclosure control

89. This sub-process ensures that the data (and metadata) to be disseminated do not breach the appropriate rules on confidentiality. This may include checks for primary and secondary disclosure, as well as the application of data suppression or perturbation techniques. The degree and method of disclosure control may vary for different types of outputs, for example the approach used for micro-data sets for research purposes will be different to that for published tables or maps.

Quality Dimension	Indicator	Notes
Statistical Confidentiality	To what extent is the business process using standard or well-known methods	Corresponds to the appropriateness of statistical

and security	identification and protection of sensitive information?	procedures principle in the ES Code of Practice.
Statistical Confidentiality and security	To what extent is the data protected from the risk of disclosure of sensitive information?	Some software provide a diagnostic indicating the level of protection.
Statistical Confidentiality and security	To what extent is the data actually protected? What is the residual risk of disclosure?	
Statistical Confidentiality and security	To what extent has the usability of the data been degraded? What is the loss in precision or level of detail?	

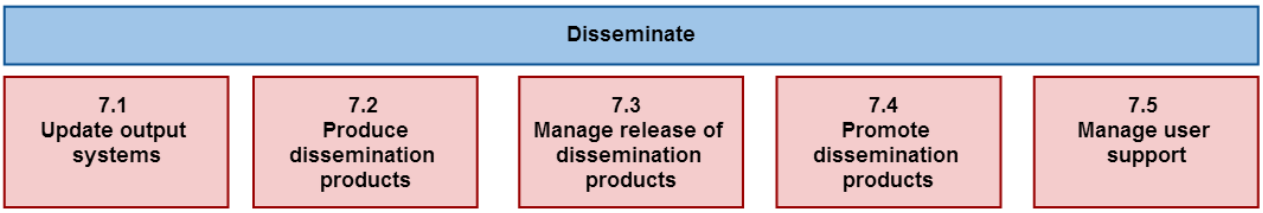
6.5. Finalise outputs

90. This sub-process ensures the statistics and associated information are fit for purpose and reach the required quality level, and are thus ready for use. It includes:

- completing consistency checks;
- determining the level of release, and applying caveats;
- collating supporting information, including interpretation, commentary, technical notes, briefings, measures of uncertainty and any other necessary metadata;
- producing the supporting internal documents;
- pre-release discussion with appropriate internal subject matter experts;
- approving the statistical content for release.

Quality Dimension	**Indicator**	**Notes**
Relevance	Number of planned outputs that were not disseminated	
Accuracy and reliability	Number of errors that were detected and had to be corrected	
Relevance	Data completeness rate: extent to which the outputs satisfy requirements (e.g. from regulations or other agreements with users).	Could be calculated as the ratio of the number of data cells obtained to the number of data cells required. ESS QPI - R1. Data completeness – rate
Accessibility and clarity	Metadata completeness - rate The rate of completeness of metadata is the ratio of the number of metadata elements provided to the total number of metadata elements applicable. The rate of completeness of metadata is applicable: • to all statistical processes; • to producers	ESS QPI - AC3. Metadata completeness – rate

Disseminate Phase



911. This phase manages the release of the statistical products to customers. It includes all activities associated with assembling and releasing a range of static and dynamic products via a range of channels. These activities support customers to access and use the outputs released by the statistical organization.

92. For statistical outputs produced regularly, this phase occurs in each iteration. It is made up of five sub-processes, which are generally sequential, from left to right, but can also occur in parallel, and can be iterative. These sub-processes are:

7.1. Update output systems

93. This sub-process manages the update of systems where data and metadata are stored ready for dissemination purposes, including:

- formatting data and metadata ready to be put into output databases;
- loading data and metadata into output databases;
- ensuring data are linked to the relevant metadata.

94. Formatting, loading and linking of metadata should preferably mostly take place in earlier phases, but this sub-process includes a final check that all of the necessary metadata are in place ready for dissemination.

Quality Dimension	Indicator	Notes
Accessibility and clarity	Date of last update of the content of the metadata. • The date of the latest dissemination of the metadata should be specified. • The date on which the metadata element was inserted or modified in the database should be specified.	
Managing metadata	Extent to which metadata are available and accessible	Corresponds to the accessibility and clarity principle in the ES Code of Practice

7.2. Produce dissemination products

95. This sub-process produces the products, as previously designed (in sub-process 2.1), to meet user needs. They could include printed publications, press releases and web sites. The products can take many forms including interactive graphics, tables, public-use micro-data sets and downloadable files. Typical steps include:

- preparing the product components (explanatory text, tables, charts, quality statements etc.);
- assembling the components into products;
- editing the products and checking that they meet publication standards.

Quality Dimension	Indicator	Notes
Quality commitment	Ratio of statistical products that are disseminated with quality statements/quality reports	
Relevance	The rate of available statistics The indicator is the ratio of the number output data elements provided in accordance to a relevant regulation to those required by the regulation. The extent to which all statistics that are needed are available.	
Relevance	Percentage of/Extent to which "statistical outputs/products" meets users' needs Description of users and their respective needs with respect to the statistical data.	This indicator is also included in 2.1.
Accessibility and clarity	The extent to which relevant metadata is linked to output data	See also 3.3

7.3. Manage release of dissemination products

96. This sub-process ensures that all elements for the release are in place including managing the timing of the release. It includes briefings for specific groups such as the press or ministers, as well as the arrangements for any pre-release embargoes. It also includes the provision of products to subscribers, and managing access to confidential data by authorized user groups, such as researchers. Sometimes an organization may need to retract a product, for example if an error is discovered. This is also included in this sub-process.

Quality Dimension	Indicator	Notes
Impartiality and objectivity	Availability and accessibility of revision policy (Yes/No)	See 2.1 and 5.7
Impartiality and objectivity	Time lag between the release of an output and announcement of the error to the users	
Transparency	Number of press meetings held before and after the release of outputs	Corresponds to the impartiality and objectivity principle in the ES Code of Practice
Accuracy and reliability	Number of errors corrected in disseminated products	Excludes planned revisions. Can include non-data errors.

Timeliness and punctuality	Punctuality of statistical outputs Punctuality is the time lag between the delivery/release date of data and the target date for delivery/release as agreed for delivery or announced in an official release calendar, laid down by Regulations or previously agreed among partners. The punctuality of statistical outputs is applicable: • to all statistical processes with fixed/pre-announced release dates, • to users and producers, with different aspects and calculation formulae.	ESS QPI - TP3. Punctuality - delivery and publication
Timeliness and punctuality	Time lag - first results General definition: The timeliness of statistical outputs is the length of time between the end of the event or phenomenon they describe and their availability. Specific definition: The number of days (or weeks or months) from the last day of the reference period to the day of publication of first results. This indicator is applicable: • to all statistical processes with preliminary data releases; • to producers.	ESS QPI - TP1. Time lag - first results
Timeliness and punctuality	Time lag - final results General definition: The timeliness of statistical outputs is the length of time between the end of the event or phenomenon they describe and their availability. Specific definition: 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. This indicator is applicable:	ESS QPI - TP2. Time lag - final results

	• to all statistical processes; • to users and producers, with different level of details given.	
Timeliness and punctuality	Availability of a dissemination policy defining dissemination practices and its availability on the web site	
Accessibility and clarity	Availability of a release calendar and its availability on the web site	
Accessibility and clarity	Number of analytical and data products accessed	Website metrics for publications and data tables available on the organization's website. Include page views and downloads.
Accessibility and clarity	Percentage of website visitors who found the information that they were looking for	Five-point scale: all, most, half, some, none
Coherence and comparability	Length of comparable time series Number of reference periods in time series from last break. Comment Breaks in statistical time series may occur when there is a change in the definition of the parameter to be estimated (e.g. variable or population) or the methodology used for the estimation. Sometimes a break can be prevented, e.g. by linking. The length of comparable series is applicable: • to all statistical processes producing time-series; • to users and producers, with different level of details given.	ESS QPI- CC2. Length of comparable time series
Statistical Confidentiality and security	• are researchers who have access to micro data legally bound to uphold confidentiality and security protocols of the NSO • are research proposals submitted for approval by NSO analysts (analysts must approve the relevance of the analysis and the appropriateness of the methods) • are there policies in place that ensure outputs are vetted prior to their dissemination • are there confidentiality rules in place, such as a minimum number of units in a cell when doing cross-tabulations, and a	Although disclosure control of individual statistical products is done in 6.4, at 7.3 additional measures should be taken to protect against disclosure that could result from researchers combining several different statistical products.

	maximum number of data requests per day with a maximum number of variables per request (to protect against penetration by an automated data mining process)	

7.4. Promote dissemination products

97. Whilst marketing in general can be considered to be an over-arching process, this sub-process concerns the active promotion of the statistical products produced in a specific statistical business process, to help them reach the widest possible audience. It includes the use of customer relationship management tools, to better target potential users of the products, as well as the use of tools including web sites, wikis and blogs to facilitate the process of communicating statistical information to users.

Quality Dimension	**Indicator**	**Notes**
Relevance	User satisfaction about the metadata availability User satisfaction surveys shall include questions on the opinions of users about metadata availability	
Accessibility and clarity	The number of social media visitors/followers	
Accessibility and clarity	Metadata - consultations Number of metadata consultations (ESMS) within a statistical domain for a given time period. This indicator is applicable: • to all statistical processes; • to producers	ESS QPI - AC2. Metadata - consultations By "number of consultations" it is meant the number of times a metadata file is viewed. Some information is available through the monthly Monitoring report on Eurostat Electronic Dissemination.
Accessibility and clarity	Number of consultations of data tables within a statistical domain for a given time period	ESS QPI - AC1. Data tables – consultations

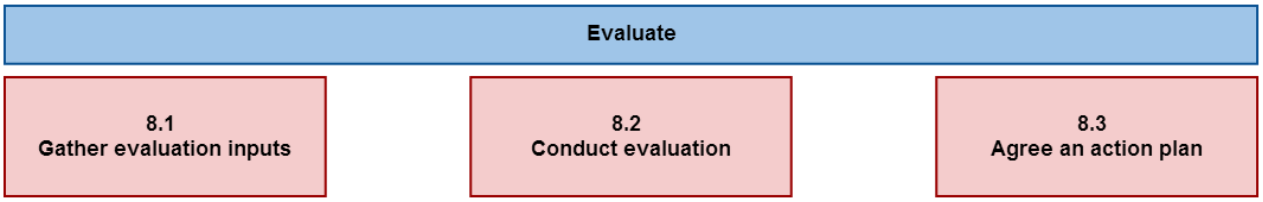
7.5. Manage user support

98. This sub-process ensures that customer queries and requests for services such as micro-data access are recorded, and that responses are provided within agreed deadlines. These queries and requests should be regularly reviewed to provide an input to the over-arching quality management process, as they can indicate new or changing user needs.

Quality Dimension	**Indicator**	**Notes**
Relevance	User satisfaction index Length of time since most recent user satisfaction survey Measures to determine user satisfaction.	

Relevance	The percentage of unmet user needs	
Relevance	Time since last user consultation, in terms of years or months	
Accessibility and clarity	Availability of an information service/unit or a call centre to users to answer enquires about data and metadata issues	

Evaluate Phase



99. This phase manages the evaluation of a specific instance of a statistical business process, as opposed to the more general over-arching process of statistical quality management described in Section VI. It logically takes place at the end of the instance of the process, but relies on inputs gathered throughout the different phases. It includes evaluating the success of a specific instance of the statistical business process, drawing on a range of quantitative and qualitative inputs, and identifying and prioritising potential improvements.

100. For statistical outputs produced regularly, evaluation should, at least in theory occur for each iteration, determining whether future iterations should take place, and if so, whether any improvements should be implemented. However, in some cases, particularly for regular and well established statistical business processes, evaluation may not be formally carried out for each iteration. In such cases, this phase can be seen as providing the decision as to whether the next iteration should start from the Specify Needs phase, or from some later phase (often the Collect phase).

101. This phase is made up of three sub-processes, which are generally sequential, from left to right, but which can overlap to some extent in practice. These sub-processes are:

8.1. Gather evaluation inputs

102. Evaluation material can be produced in any other phase or sub-process. It may take many forms, including feedback from users, process metadata (paradata), system metrics, and staff suggestions. Reports of progress against an action plan agreed during a previous iteration may also form an input to evaluations of subsequent iterations. This sub-process gathers all of these inputs, and makes them available for the person or team producing the evaluation.

Quality Dimension	Indicator	Notes
Output quality	Extent to which quality indicators have been collected for all phases and sub-phases including costs and timeliness of phases and sub-phases.	Indicators and feedbacks should have been collected in previous phases (and some of them probably also analysed) Output Quality gathers all

		dimensions related to the quality of statistics (e.g. relevance, accuracy, timeliness, coherence)
Output quality	Types and relative weight of different measures gathered (e.g. quantitative indicators, feedback from users, paradata or other metrics derived by procedures, staff suggestions, interviewers/supervisors follow ups)	Indicators and feedback should have been collected in previous phases and some of them probably also analysed. Output quality gathers all dimensions related to the quality of statistics (e.g. relevance, accuracy, timeliness, coherence).

8.2. Conduct evaluation

103. This sub-process analyses the evaluation inputs and synthesises them into an evaluation report. The resulting report should note any quality issues specific to this iteration of the statistical business process, and should make recommendations for changes if appropriate. These recommendations can cover changes to any phase or sub-process for future iterations of the process, or can suggest that the process is not repeated.

Quality Dimension	**Indicator**	**Notes**
Soundness of implementation	To what extent process components satisfy process quality requirements such as Efficiency, Effectiveness; Robustness; Flexibility; Transparency and Integration	See also phase 3. Build. For a new process, such an assessment has been carried out in phase 3. Build. For regular processes this stage could represent the opportunity to assess both process components and outputs. Corresponds to the appropriate statistical procedures principle in the ES Code of Practice.
Cost effectiveness	Percentage of GSBPM phases and sub-processes for which there were no gaps between planned and attained costs	
Output quality	Extent to which quality indicators are close to target values (includes all indicators and metadata such as those needed for quality reporting)	Assessment is based on information from 8.1 Output Quality gathers all dimensions related to the quality of statistics (e.g. relevance, accuracy, timeliness, coherence)

Output quality	Trends in quality indicators (e.g. improvements/worsening) for recurring processes.	Output Quality gathers all dimensions related to the quality of statistics (e.g. relevance, accuracy, timeliness, coherence).
Output quality	Percentage of quality dimensions and sub-dimensions (e.g. for accuracy) that was not possible to assess and why.	Output Quality gathers all dimensions related to the quality of statistics (e.g. relevance, accuracy, timeliness, coherence).
Output quality	If an evaluation report has been produced and on which basis (e.g. overall assessment of quality indicators calculated during the process, application of a quality assessment procedure, e.g. self-assessment, audit.)	The indicator can assume values like : 0 (no evaluation report produced) 1 (evaluation report produced on currently available quality indicators) 2 (evaluation report produced on the result of an ad hoc analysis, e.g. a study to estimate the Mean Square Error, MSE) 3 (evaluation report produced on the result of a self-assessment procedure) 4 (evaluation report produced on the result of an audit procedure) Output quality gathers all dimensions related to the quality of statistics (e.g. relevance, accuracy, timeliness, coherence).
Timeliness and punctuality	Percentage of GSBPM phases and sub-processes for which there were no gaps between target and achieved timeliness	ESS QPI - TP2. Time lag final results
Output quality	Have evaluated substantial changes in quality indicators? (Yes/No)	Significant changes in quality indicators and other measures such as edit failure rates, response rates etc. may signal quality issues.

8.3. Agree an action plan

104. This sub-process brings together the necessary decision-making power to form and agree an action plan based on the evaluation report. It should also include consideration of a mechanism for monitoring the impact of those actions, which may, in turn, provide an input to evaluations of future iterations of the process.

Quality Dimension	**Indicator**	**Notes**
Quality	Extent to which the action plan contains	

commitment	mechanisms for monitoring the impact of improvement actions	
Quality commitment	Assuming that an evaluation report was prepared in 8.2 for quality indicators of previous GSBPM phases, and the gaps were identified between the expected and actual quality of the output, cost effectiveness and timeliness; then the decision needs be made to take action for areas where the gaps are identified. The quality indicator is the ratio of: the number of actionable quality issues (quality indicators where problems are identified or targets are not met) / to the total number of quality issues Also a plan can be made to not take an action for all actionable items but for some of them. In that case the quality indicator is: number of quality issues to take action for divided by the number of all actionable quality issues	
Quality commitment	Completion rate of the action plan is: the number of successfully fixed or improved quality issues divided by total number of quality issues planned to be fixed	

III. Over-arching processes

105. The GSBPM also recognises several over-arching processes that apply throughout the production phases, and across statistical business processes. The processes of quality management and metadata management are further elaborated in this Section.

Quality Management

106. Quality concerns organisations, processes and products. In the present framework, quality management over-arching process refers mainly to product and process quality.

107. The main goal of quality management within the statistical business process is to understand and manage the quality of the statistical products. There is general agreement among statistical organisations that quality should be defined according to the ISO 9000-2005 standard: “The degree to which a set of inherent characteristics fulfils requirements"[5]. Thus, product quality is a complex and multi-faceted concept, usually defined in terms of several quality dimensions. The dimensions of quality that are considered most important depend on

[5]ISO 9000:2005, Quality management systems -- Fundamentals and vocabulary. International Organization for Standardization

user perspectives, needs and priorities, which vary between processes and across groups of users.

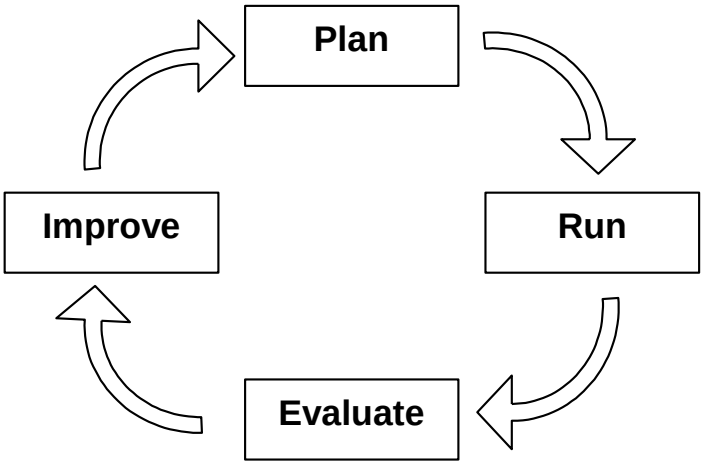
108. In order to improve the product quality, quality management should be present throughout the statistical business process model. It is closely linked to Phase 8 (Evaluate), which has the specific role of post-evaluating individual instances of a statistical business process. However, quality management has both a deeper and broader scope. As well as evaluating iterations of a process, it is also necessary to evaluate separate phases and sub-processes, ideally each time they are applied, but at least according to an agreed schedule. Metadata generated by the different sub-processes themselves are also of interest as an input for process quality management. These evaluations can apply within a specific process, or across several processes that use common components.

109. In addition, a fundamental role in quality management is played by the set of quality control actions that should be implemented within the sub-processes to prevent and monitor errors. The strategy could be reported in a quality assurance plan.

110. Within an organisation, quality management will usually refer to a specific quality framework, and may therefore take different forms and deliver different results within different organisations. The current multiplicity of quality frameworks enhances the importance of the benchmarking and peer review approaches to evaluation, and whilst these approaches are unlikely to be feasible for every iteration of every part of every statistical business process, they should be used in a systematic way according to a pre-determined schedule that allows for the review of all main parts of the process within a specified time period[6].

111. Broadening the field of application of the quality management over-arching process, evaluation of groups of statistical business processes can also be considered, in order to identify potential duplication or gaps.

112. All evaluations result in feedback, which should be used to improve the relevant process, phase or sub-process, creating a quality loop.



113. Examples of quality management activities include:

[6] A suitable global framework is the National Quality Assurance Framework developed by a global expert group under the United Nations Statistical Commission. See: http://unstats.un.org/unsd/dnss/QualityNQAF/nqaf.aspx

- Setting and maintaining of the quality framework;
- Setting of global quality criteria;
- Setting process quality targets and monitoring compliance;
- Seeking and analysing user feedback;
- Reviewing operation and documenting lessons learned;
- Examining process metadata and quality indicators;
- Internal or external auditing on statistical processes.

114. Quality management also involves institutional and organisational factors. Such factors are included in other GSBPM over-arching processes (e.g. Human resources management, Statistical programme management) although they can have an impact on quality.

Quality Dimension	**Indicator**	**Notes**
Quality commitment	Availability of a quality assurance plan, or any other similar scheme, describes the working standards, the formal obligations (such as laws and internal rules) and the set of quality control actions to prevent and monitor errors, to evaluate quality indicators and to control different points at each stage of the statistical process. This indicator is valid for the institutional level.	
Quality commitment	Availability of a quality policy and its availability on the web site A Quality Commitment Statement is made publicly available, laying out principles and commitments related to quality in statistics which are consistent with the goals set out in the mission and vision statements. This indicator is valid for the institutional level.	
Quality commitment	Availability of procedures to plan and monitor the quality of the statistical production process.	
Quality commitment	Availability of a clear organizational structure for managing quality within the statistical authority. Examples of such a structure are: • Quality Committee; • Quality Manager; • Centralized Quality unit; • Other structures (e.g. a selected group of staff trained as “quality pilots” to act as project/processes	

	coach/advisers).	
Quality commitment	For what proportion of GSBPM sub-processes are standardised corporate solutions used?	
Quality commitment	Is a process of risk identification and management in place? (Yes/No) Time since risk management plans were last reviewed? (Years and Months)	
Quality commitment	Extent of HR requirements fulfilled (e.g. training, staffing)	
Quality commitment	Extent to which quality indicators, metadata and paradata are compliant to standards	
Managing respondent burden	Is there a communication strategy encouraging response by informing potential respondents about the survey?	
Managing respondent burden	Percentage of statistics produced from administrative data and other data sources instead of survey	Covers all statistical domains

Metadata Management

115. Good metadata management is essential for the efficient operation of statistical business processes. Metadata are present in every phase, either created or carried forward from a previous phase. In the context of this model, the emphasis of the over-arching process of metadata management is on the creation, use and archiving of statistical metadata, though metadata on the different sub-processes themselves are also of interest, including as an input for quality management. The key challenge is to ensure that these metadata are captured as early as possible, and stored and transferred from phase to phase alongside the data they refer to. Metadata management strategy and systems are therefore vital to the operation of this model, and these can be facilitated by the GSIM.

116. The GSIM is a reference framework of information objects, which enables generic descriptions of the definition, management and use of data and metadata throughout the statistical production process. The GSIM supports a consistent approach to metadata, facilitating the primary role for metadata envisaged in Part A of the Common Metadata Framework[7] "Statistical Metadata in a Corporate Context", that is, that metadata should uniquely and formally define the content and links between objects and processes in the statistical information system.

117. Part A of the Common Metadata Framework also identifies the following sixteen core principles for metadata management, all of which are intended to be covered in the over-arching Metadata Management process, and taken into the consideration when preparing the statistical metadata system (SMS) vision and global architecture, and when implementing the SMS. The principles can be presented in the following groups:

[7] See: http://www.unece.org/stats/cmf/PartA.html

Metadata handling	**i.**	***Statistical Business Process Model***: Manage metadata with a focus on the overall statistical business process model.
	ii.	***Active not passive***: Make metadata active to the greatest extent possible. Active metadata are metadata that drive other processes and actions. Treating metadata this way will ensure they are accurate and up-to-date.
	iii.	***Reuse***: Reuse metadata where possible for statistical integration as well as efficiency reasons
	iv.	***Versions***: Preserve history (old versions) of metadata.
Metadata Authority	**i.**	***Registration***: Ensure the registration process (workflow) associated with each metadata element is well documented so there is clear identification of ownership, approval status, date of operation, etc.
	ii.	***Single source***: Ensure that a single, authoritative source ('registration authority') for each metadata element exists.
	iii.	***One entry/update***: Minimize errors by entering once and updating in one place.
	iv.	***Standards variations***: Ensure that variations from standards are tightly managed/approved, documented and visible.
Relationship to Statistical Cycle / Processes	**i.**	***Integrity***: Make metadata-related work an integral part of business processes across the organization.
	ii.	***Matching metadata***: Ensure that metadata presented to the end-users match the metadata that drove the business process or were created during the process.
	iii.	***Describe flow***: Describe metadata flow with the statistical and business processes (alongside the data flow and business logic).
	iv.	***Capture at source***: Capture metadata at their source, preferably automatically as a bi-product of other processes.
	v.	***Exchange and use***: Exchange metadata and use them for informing both computer based processes and human interpretation. The infrastructure for exchange of data and associated metadata should be based on loosely coupled components, with a choice of standard exchange languages, such as XML.
Users	**i.**	***Identify users***: Ensure that users are clearly identified for all metadata processes, and that all metadata capturing will create value for them.
	ii.	***Different formats***: The diversity of metadata is recognized and there are different views corresponding to the different uses of the data. Different users require different levels of detail. Metadata appear in different formats depending on the processes and goals for which they are produced and used.
	iii.	***Availability***: Ensure that metadata are readily available and useable in the context of the users' information needs (whether an internal or external user).

Quality Dimension	Indicator	Notes
Managing metadata	Availability of a policy on metadata documentation and standards on updating metadata. The policy is communicated to internal users and accessible on the web site. Performance indicators for the accuracy, completeness, timeliness and accessibility of disseminated metadata at the institutional level, assessed annually against predetermined targets.	Quality of the metadata can be assessed for each statistical programme and then rolled up to higher levels of institutional units
Managing metadata	Extent to which metadata and metadata terminology are compliant to existing metadata standards.	Metadata standards include GSIM, GSBPM, CSPA and LIM
Managing metadata	Extent to which the life cycle of the metadata is managed across the GSBPM. Use of a metadata system (data or process metadata) in the production process. (Yes / No) Extent to which metadata are adequately stored and archived using a metadata model (easily retrievable; properly labelled; retention period indicated) Extent to which metadata are accurately and completely registered in a corporate metadata repository/registry.	The importance of a metadata model and data metadata stored in a metadata system is crucial in processing and delivering data. The maintenance of the production process is easier when it uses information from general metadata systems and information is not coded in production programmes.
Managing metadata	Extent to which metadata are available in different formats and available to internal and external users Are metadata available in machine-readable, searchable and accessible formats? Are metadata available in open data portals? Are metadata and data accessible in standard exchange formats such as SDMX, DDI or XBRL?	

References

Aabo, T., Fraser, J., & Simkins, B. J. (2005). The rise and evolution of the chief risk officer: Enterprise risk management at hydro one. *Journal of Applied Corporate Finance*, 17(3), 62–75.]

Biemer, P.P. and L.E. Lyberg. 2003. *Introduction to Survey Quality.* New York: John Wiley

Biemer, P., Trewin, D., Bergdahl, H., and Japec, L. (2014). A system for managing the quality of official statistics. *Journal of Official Statistics*, 30, 3, 381–415, http://dx.doi.org/10.2478/JOS-2014-0022.

Daas, P., Ossen, S., Vis-Visschers, R., and Arends-Tóth, J. (2009). Checklist for the Quality evaluation of Administrative Data Sources, Statistics Netherlands, The Hague Heerlen, 2009

Daas, P., Ossen, S. (2011). Report on methods preferred for the quality indicators of administrative data sources, Blue – ETS Project, Deliverable 4.2.

ESSnet AdminData (2013). Final list of quality indicators and associated guidance. Deliverable 2011/6.5 of ESSnet on Admin Data https://ec.europa.eu/eurostat/cros/content/use-administrative-and-accounts-data-business-statistics_en

Eurostat, (2011). The European Statistics Code of Practice. http://ec.europa.eu/eurostat/web/quality/european-statistics-code-of-practice

Eurostat, (2014). ESS Guidelines for the Implementation of the ESS Quality and Performance Indicators (QPI). http://ec.europa.eu/eurostat/documents/64157/4373903/02-ESS-Quality-and-performance-Indicators-2014.pdf/5c996003-b770-4a7c-9c2f-bf733e6b1f31

Eurostat (2015). Quality Assurance Framework of the European Statistical System. Version v1.2. http://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V1-2final.pdf/bbf5970c-1adf-46c8-afc3-58ce177a0646

MIAD - Methodologies for an Integrated Use of Administrative Data (2015) Deliverables A.1., A.2., B.1., B.2., B.3. Available at: https://ec.europa.eu/eurostat/cros/content/miad-methodologies-integrated-use-administrative-data-statistical-process_en

Reedman, L., Dorsan, N., Signore, M., and Özkan, D.. Challenges and Discoveries in Developing Quality Indicators for the GSBPM, European Conference on Quality in Official Statistics (Q2016) Madrid, 31 May-3 June 2016

Reid, G., Zabala, F., and Holmberg A. (2017). Extending TSE to Administrative Data: A Quality Framework and Case Studies from Stats NZ. *Journal of Official Statistics*, Vol. 33, No. 2, 2017, pp. 477–511, http://dx.doi.org/10.1515/JOS-2017-0023.

United Nations Statistics Division (2015). National Quality Assurance Frameworks https://unstats.un.org/unsd/dnss/QualityNQAF/nqaf.aspx

Wallgren, A. & Wallgren, B. (2014). Register-based Statistics: Administrative Data for Statistical Purposes. Second Edition. John Wiley & Sons, Chichester, UK.

Zhang L.C. (2012). Topics of statistical theory for register-based statistics and data integration. *Statistica Neerlandica*, Vol 66, nr.1, pp. 41-63.

Annex 1

In the following table, the European Statistical System Quality and Performance Indicators (ESS QPI) are mapped to the corresponding indicators in the GSBPM, and to the phases and sub-processes.

ESS QPI	Name in GSBPM	Phase and sub-process
R1. Data completeness – rate	Planned data completeness rate: extent to which the planned outputs will satisfy requirements (e.g. from Regulations or other agreements with users)	Design phase. 2.1. Design outputs
	Data completeness rate: extent to which the outputs satisfy requirements (e.g. from regulations or other agreements with users)	Analyse Phase 6.5. Finalise outputs
A1. Sampling error - indicators	Key indicators for sample design (e.g. estimated size, expected/planned sampling errors for key variables, domains, costs)	Design phase. 2.4. Design frame and sample
	The sampling error can be expressed: a) in relative terms, in which case the relative standard error or, synonymously, the coefficient of variation (CV) is used. b) in terms of confidence intervals	Collect phase. 4.1. Create frame and select sample
	Domain response rates; representativity indicators; achieved CVs of key variables in domains of interest	Collect phase. 4.3. Run collection
	The sampling error can be expressed: a) in relative terms, in which case the relative standard error or, synonymously, the coefficient of variation (CV) is used. b) in terms of confidence intervals.	Process phase. 5.7. Calculate aggregates
A2. Over-coverage - rate	Impact of coverage errors: assess the likely impact of coverage error on key estimates.	Design phase. 2.4. Design frame and sample
	Assessment of major error sources from the pilot (e.g. coverage, nonresponse, measurement, and process errors)	Build phase. 3.6. Test statistical business process
	The rate of over-coverage: The proportion of units accessible via the frame that do not belong to the target population (are out-of-scope).	Collect phase. 4.1. Create frame and select sample

	The rate of over-coverage: The proportion of units accessible via the frame that do not belong to the target population (are out-of-scope).	Collect phase. 4.4. finalise collection
A3. Common units - proportion	The proportion of units covered by both the survey and the administrative sources in relation to the total number of units in the survey	Process phase. 5.1. integrate data
A4. Unit non-response - rate	Assessment of major error sources from the pilot(e.g. coverage, nonresponse, measurement, and process errors)	Build phase. 3.6. Test statistical business process
	Relative discrepancy between expected and observed sample size; relative discrepancy between expected and observed response, attrition and out of scope rates	Collect phase. 4.1. Create frame and select sample
	Domain response rates; representativity indicators; achieved CVs of key variables in domains of interest	Collect phase. 4.3. Run collection
	Unit nonresponse rate; item nonresponse rate; proxy rate	Collect phase. 4.3. Run collection
	Unit nonresponse rate; item nonresponse rate; proxy rate	Collect phase. 4.4. finalise collection
A5. Item non-response - rate	Unit nonresponse rate; item nonresponse rate; proxy rate	Collect phase 4.3. Run collection
	Unit nonresponse rate; item nonresponse rate; proxy rate	Collect phase. 4.4. finalise collection
A6. Data revision - average size	Data revisions are planned (Yes/No)	Design phase. 2.1 design outputs
	The following indicators are proposed to analyze revisions: Mean Absolute Revision (MAR), i.e. the average of absolute revisions over a time period (useful to analyze stability in terms of size). Relative Mean Absolute Revisions (RMAR), i.e. the relative average of absolute revisions over a time period (useful for comparisons and to analyze levels) 3) Mean Revision (MR), i.e. the average of revisions over a time period (useful to analyze	Process phase. 5.7. Calculate aggregates

	directions in terms of sign) and its significance (Yes/Not) Standard Deviation of Revisions (SDR), i.e. a measure of the variability of the revisions	
A7. Imputation - rate	Imputation rate - The indicator is expressed as the ratio of the number of replaced values to the total number of values for a given variable.	Process phase. 5.4. Edit and impute
TP1. Time lag - first results	Time lag - first results	Disseminate phase. 7.3. Manage release of dissemination products
TP2. Time lag - final results	Planned time frame for subsequent phases and sub-processes	Design phase. 2.6. Design production systems and workflow Build phase. 3.4. Configure workflows
	Estimated time frame for subsequent phases and sub-processes and divergences from planned one in design phase	Build phase. 3.6 Test statistical business process
	Time lag - final results	Disseminate phase. 7.3. Manage release of dissemination products
	Percentage of GSBPM phases and sub-processes for which there were no gaps between target and achieved timeliness	Evaluate phase. 8.2. Conduct evaluation
TP3. Punctuality - delivery and publication	Punctuality of statistical outputs Punctuality is the time lag between the delivery/release date of data and the target date for delivery/release as agreed for delivery or announced in an official release calendar, laid down by Regulations or previously agreed among partners.	Disseminate phase. 7.3. Manage release of dissemination products
CC1. Asymmetry for mirror flows statistics - coefficient	Proportion of overall budget dedicated to validation activities; number of validation measures applied;	Analyse phase. 6.2 Validate outputs
CC2. Length of comparable time series	Expected length of comparable time series.	Design phase. 2.1 design outputs

	Length of comparable time series	Disseminate phase. 7.3. Manage release of dissemination products
AC1. Data tables – consultations	Number of consultations of data tables within a statistical domain for a given time period	Design phase 2.6. Design production systems and workflow
AC1. Data tables – consultations	Number of consultations of data tables within a statistical domain for a given time period	Disseminate phase. 7.4. Promote dissemination products
AC2. Metadata - consultations	Metadata – consultations Number of metadata consultations (ESMS) within a statistical domain for a given time period.	Design phase 2.6. Design production systems and workflow
AC2. Metadata - consultations	Metadata - consultations Number of metadata consultations (ESMS) within a statistical domain for a given time period.	Disseminate phase. 7.4. Promote dissemination products
AC3. Metadata completeness – rate	Metadata completeness - rate	Analyse Phase. 6.5. Finalise outputs