Generic Statistical Business Process Model (GSBPM)

Overview

The Generic Statistical Business Processes Model (Version 5) lists the process followed by the Office for National Statistics (ONS) to produce statistical outputs. In this document we consider each phase in turn, identifying the various sub-processes within that phase, and describing their contents.

Background

In autumn 2012, the ONS Design Authority made the decision that ONS will adopt GSBPM terminology for its business processes which brought us in line with other National Statistical Institutes (NSIs).

GSPBM (previously known as the Statistical Value Chain (SVC)) was produced by UN/OECD/Eurostat. It ensures the same terminology is used by all when referring to statistical business processes. A GSBPM revision Task Force was initiated in 2013 with representatives from a range of NSIs, including ONS. There are limited changes to the model itself, but many more improvements, additions and clarifications to the supporting documentation.

The main changes introduced in Version 5

- the former phase 8 (out of 9 phases previously) entitled 'Archive' has been removed, and incorporated into the over-arching process of data and metadata management, to reflect the view that archiving can happen at any stage in the statistical production process. There are now just eight, not nine, phases to the GSBPM. These are shown below. There are also ONS examples available for each of the phase descriptions in GSBPM.

- a new sub-process: 'Build or enhance dissemination components' has been added within the 'Build' phase to reflect the growing importance of having a range of dissemination options,

- several sub-processes have been re-named to improve clarity,

- the descriptions of sub-processes have been updated and expanded where necessary. The terminology used has been changed to be less survey based, in recognition of the growing use of non-survey sources (administrative data, big data etc).

Please email ONS Quality Centre if you require training, or have any questions to ensure quality is maintained.

1. GSBPM

Each sub-process is also expressed as a linear version.

2. Over-arching processes: quality management and metadata management

This page forms part of the overall Generic Statistical Business Process Model (GSPBM) listing the processes followed by ONS to produce statistical outputs.

Contents

- Over-arching process

- Quality Management

- Metadata Management

- Help and Questions

Over-arching process

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. The processes of quality management and metadata management are further elaborated in this section. Other over-arching processes include data management; customer management; project management. For more details on these overarching processes, please see paragraphs 13 and 14 of [this document](#).

Quality Management

This process is present throughout the model. It is closely linked to Phase 8 (Evaluate), which has the specific role of evaluating individual instances of a statistical business process. The over-arching quality management process, however, 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.

Quality management also involves the evaluation of groups of statistical business processes, and can therefore identify potential duplication or gaps. All evaluations should result in feedback, which should be used to improve the relevant process, phase or sub-process, creating a quality loop.

For more information on Quality within ONS, please see the ONS Quality Management Strategy.

Metadata Management

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 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 effective operation of this process.

Relevant parts of this metadata should also be reported to users if the information is useful for their understanding of the statistical output.

For more information on this reporting within ONS contact the ONS Metadata Project.

Linear model

While this is shown below for illustrative purposes, it is designed to be a matrix through which there are many possible paths, including iterative loops with and between phases.
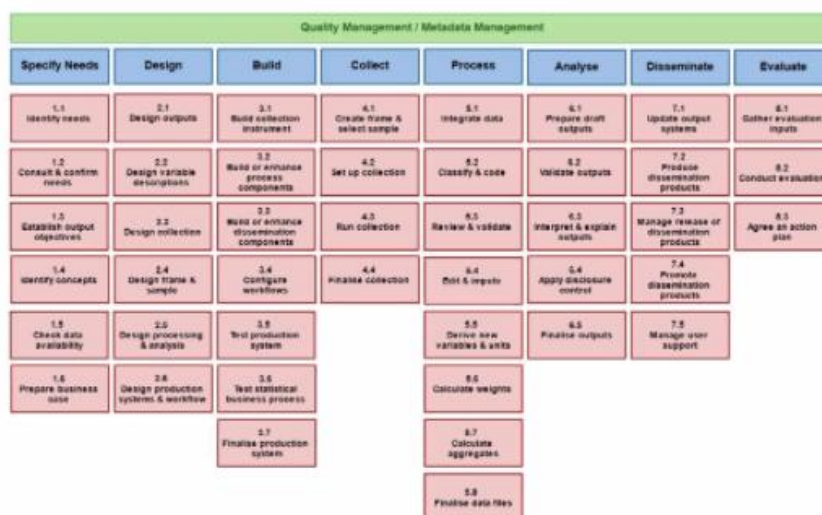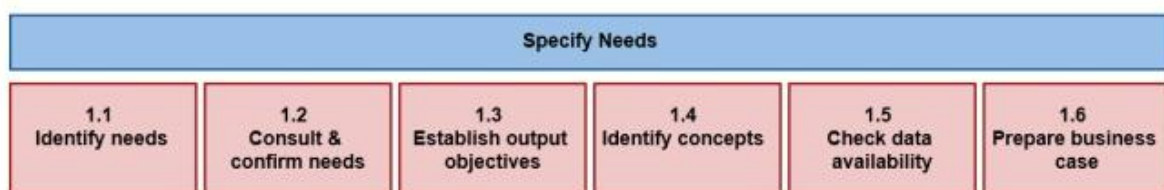

3. Phase 1 – Specify Needs

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.

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 demands, and
- prepares the business case to get approval to produce the statistics.

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.

Specify Needs					
1.1 Identify needs	1.2 Consult & confirm needs	1.3 Establish output objectives	1.4 Identify concepts	1.5 Check data availability	1.6 Prepare business case

The sub-processes are:

1.1. Title – 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.

1.2. Consult and confirm needs – 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 organization 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.

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

1.4. Identify concepts – 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.

1.5. Check data availability – 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.

1.6. Prepare business case – this sub-process documents the findings of the other sub-processes in this phase in the form 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;
- the proposed 'To-Be' solution, detailing how the statistical business process model will be developed to produce the new or revised statistics, and
- an assessment of costs and benefits, as well as any external constraints.

4. Phase 2 – Design

This phase describes the development and design activities, and any associated practical research work needed to define the statistical outputs, concepts, methodologies, collection instruments and operational processes. It includes all the design elements needed to define or refine the statistical products or services identified in a 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.

Design activities make substantial use of international and national standards, in order to reduce the length and cost of the design process, and enhance 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. 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 – 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 organizations from sub-process 1.1 (Identify needs).

2.2 Design variable descriptions – 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 data 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.

2.3 Design collection – this sub-process determines the most appropriate data 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.

2.4. Design frame and sample – 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.

2.5. Design processing and analysis – 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 finalising data sets.
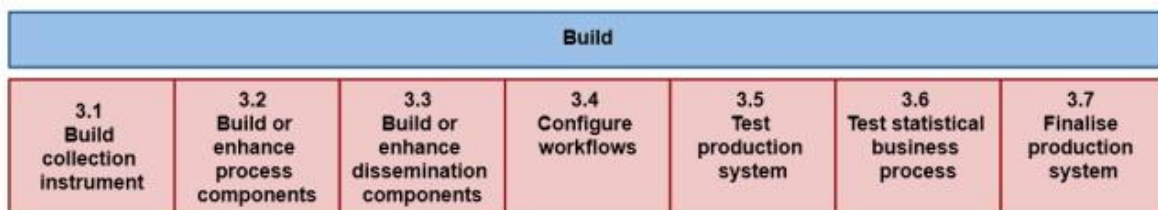
2.6. Design production systems and workflow – this sub-process determines the workflow from data collection to dissemination, taking an overview of all the processes required within the whole

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 (for example 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.

5. Phase 3 – Build

This phase builds and tests the production solution to the point where they are 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 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.

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

Build						
3.1 Build collection instrument	3.2 Build or enhance process components	3.3 Build or enhance dissemination components	3.4 Configure workflows	3.5 Test production system	3.6 Test statistical business process	3.7 Finalise production system

These sub-processes are:

3.1. Build collection instrument – 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, for example personal or telephone interviews; paper, electronic or web questionnaires and/or 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 (for example, 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 Collect 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.

3.2. Build or enhance process components – this sub-process describes the activities to build new and enhance existing software components 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 tools.

3.3. Build or enhance dissemination components – 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.

3.4. Configure workflows – 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.

3.5. Test production system – 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.

3.6. Test statistical business process – 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, for example a population census, there may be several iterations until the process is working satisfactorily.

3.7 Finalise production systems – 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, and

- 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).

6. Phase 4 – Collect

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.

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.
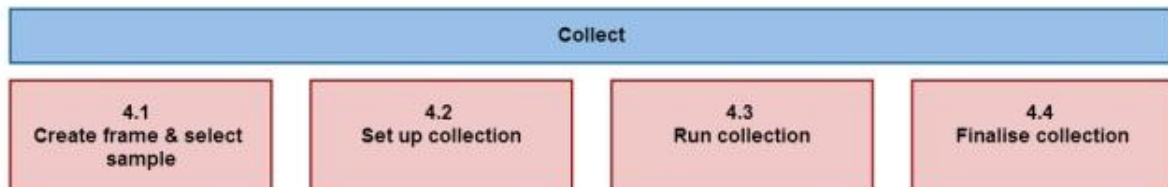
These sub-processes are:

4.1. Create frame and select sample – 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 data sources (for example, administrative data) as such processes generally create frames from the available data and then follow a census approach.

4.2. Set up collection – 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 for example, laptops,

- agreeing terms with any intermediate collection bodies, for example sub-contractors for computer assisted telephone interviewing configuring collection systems to request and receive data,

- ensuring the security of data to be collected and

- preparing collection instruments (for example, printing questionnaires, pre-filling them with existing data, loading questionnaires and data onto interviewers' computers etc).

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.

4.3. Run collection – 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, for example checking that files are in the right format and contain the expected fields. All validation of the content takes place in the Process phase.

4.4. Finalize collection – 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 the material.

7. Phase 5 – Process

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 (with the possible exception of sub-process 5.6 (Calculate weights)), which is usually specific to survey data).

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. This phase is broken down into eight sub-processes, which may be sequential, from left to right, but can also occur in parallel, and can be iterative.

Process							
5.1 Integrate data	5.2 Classify & code	5.3 Review & validate	5.4 Edit & impute	5.5 Derive new variables & units	5.6 Calculate weights	5.7 Calculate aggregates	5.8 Finalise data files

These sub-processes are:

5.1. Integrate data – 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. The result 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,

- matching /record linkage routines, with the aim of linking micro or macro data from different sources, and

- prioritising, when two or more sources contain data for the same variable with potentially different values.

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 to protect confidentiality.

5.2. Classify and code – 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.

5.3. Review and validate – 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 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. Edit and impute – where data are considered as incorrect, missing or unreliable, estimates, 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, and

- the production of metadata on the editing and imputation process.

5.5. Derive new variables and units – 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.

5.6. Calculate weights – 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, variations may need weighting of normalisation purposes.

5.7. Calculate aggregates – 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.

5.8. Finalize data files – 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.

8. Phase 6 – Analyse

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 for all statistical outputs, regardless of how the data were sourced.

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.
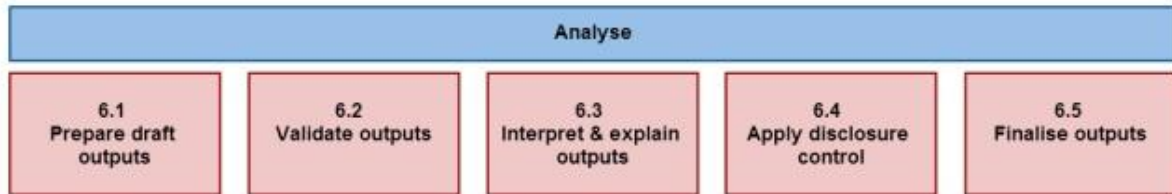
The sub-processes are:

6.1. Prepare draft outputs – 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.

6.2. Validate outputs – 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 the associated metadata and paradata (process data) 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, and

- validating the statistics against expectations and domain intelligence.

6.3. Interpret and explain outputs – 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.

6.4. Apply disclosure control – 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.

6.5. Finalize outputs – 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, and

- approving the statistical content for release.

9. Phase 7 – Disseminate

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 organisation. 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 – this sub-process manages the update of systems where data and metadata are stored for dissemination purposes, including:

- formatting data and metadata ready to be put into output databases,

- loading data and metadata into output databases, and

- ensuring data are linked to the relevant metadata.

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

7.2. Produce dissemination products – 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, and

- editing the products and checking that they meet publication standards.

7.3. Manage release of dissemination products – 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 authorised user groups, such as researchers. Sometimes an organisation may need to retract a product, for example if an error is discovered. This is also included in the sub-process.

7.4. Promote dissemination products – 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.

7.5. Manage user support – 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 should be regularly reviewed to provide an input to the over-arching quality management process, as they can indicate new or changing user needs.

10. Phase 8 – Evaluate

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.

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).

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 – 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.

8.2. Conduct evaluation – this sub-process analyzes the evaluation inputs and synthesizes 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.

8.3. Agree an action plan – 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.