

# Statistical Metadata Glossary

Metadata Glossary is a list of definitions and explanatory texts for main terms used in ModernStats Models: GAMSO, GSBPM, GSIM and CSPA. For GSIM which already has its own glossary, Metadata Glossary team reviewed terms and made proposals for changes (column Note).

Click [here](#) to download excel version

Term	Standard	Definition	Relationships	Explanatory text	Note
Agree An Action Plan	GSBPM 5.1	Sub-process of deciding what should be improved in future iterations of the statistical business process	Evaluate	This Sub-process includes a consideration of time, cost, available resources, impact, and priorities of the organisation. These decisions may be applied to other statistical business processes.	
Analyse	GSBPM 5.1	Phase of the statistical process in which data are carefully examined before dissemination to users	Phase; Prepare Draft Outputs; Validate Outputs; Interpret And Explain Outputs; Apply Disclosure Control; Finalise Outputs	The Analyse Phase ensures that data to be disseminated are fit for purpose. It includes the validation, interpretation and adjustment of the data and the preparation of documents such as technical or methodological notes explaining the statistical content.	
Apply Disclosure Control	GSBPM 5.1	Sub-process of protecting data in order to ensure that no person or organisation is identifiable in the released data	Analyse	This Sub-process is where the likelihood of identification of a person or an organisation in the data and metadata is assessed, and protection techniques are considered and implemented.	
Build	GSBPM 5.1	Phase of the statistical process concerned with the construction and deployment of the ready-to-use production system, based on results from the Design phase	Phase; Reuse Or Build collection instrument; Reuse Or Build Processing And Analysis Components; Reuse Or Build Dissemination Components; Configure Workflows; Test Production System; Test Statistical Business Process; Finalise Production Systems	This Phase builds, assembles, and tests the production system 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. The Sub-processes include building modules, services, and sub-systems for collection, processing, and dissemination, which constitute a production system. Each of the constituent parts of the resulting production system may be automated or not.	
Business Process	GSBPM 5.1	process employed to implement the necessary activities of an organisation	Process GSBPM	Necessary activities are those needed to help fulfilling the mandate of an organisation. Examples from a statistical office are those needed to conduct a survey or manage personnel. Unnecessary activities might include holiday celebrations in an office.	
Calculate Aggregates	GSBPM 5.1	Sub-process of transforming data to generate the required macro data or to derive statistical indicators	Process; Calculate Weights; Design Processing And Analysis	This Sub-process can include the summing of data for records sharing certain characteristics and determining measures of average and dispersion. In the case of sampling surveys, weights determined in 5.6 (Calculate weights) are used to obtain the aggregated values as well as measures of sampling errors.	
Calculate Weights	GSBPM 5.1	Sub-process of calculating coefficients with the aim of ensuring estimates approximate the population totals	Process; Design Processing And Analysis	The Sub-process generates the weights associated to unit data records according to sound methodologies and as designed in sub-process 2.5 (Design processing and analysis).	
Check Data Availability	GSBPM 5.1	Sub-process of determining if there are existing data sources that could be transformed to meet the user information requirements	Specify Needs	This Sub-process assesses existing data sources, such as administrative, statistical or non-statistical, to determine if they are suitable. When existing sources have been assessed, a strategy for filling any remaining gaps in the data requirement is prepared. An evaluation of the need for changes in existing legal frameworks to use the data may be considered.	
Classify And Code	GSBPM 5.1	Sub-process of transforming data into codes using statistical classifications and code lists	Process	This Sub-process often uses international or national standards. Classification and coding may be performed by automatic routines or manually (clerically).	
Collect	GSBPM 5.1	Phase of the statistical process concerned with the selection of units from which to obtain data and the collection or extraction of that data and related metadata	Phase; Create Frame And Select Sample; Set Up Collection; Run Collection; Finalise Collection	This Phase is about collecting or gathering all necessary information (data and metadata), using (possibly) different methods (e.g. traditional usage of question instruments as well as extractions from statistical, administrative, or other non-statistical registers and databases), managing the logistics, and loading the data into the appropriate environment for further processing. The Sub-processes include all aspects of the data collection or gathering process including selecting from where, whom, how, and what data will be gathered, and to where and how the data will be available for subsequent phases.	
Conduct Evaluation	GSBPM 5.1	Sub-process of analysing the evaluation inputs in order to provide feedback on the statistical business process	Evaluate	This Sub-process produces feedback (e.g. reports and recommendations) based on the analysis of evaluation inputs collected and compiled. The feedback can cover changes to any Phase or Sub-process for future iterations of the process, or can suggest that the process is not repeated.	

Configure Workflows	GSBPM 5.1	Sub-process of assembling the systems and transformations used within the statistical business process, checking the interactions among them, and adjusting as necessary	Build; Design Production Systems And Workflow	This Sub-process ensures that the workflow as specified in Sub-process Design Production Systems And Workflow works in practice.	
Consult And Confirm Needs	GSBPM 5.1	Sub-process of evaluating the information requirements with the users, and agreeing upon the deliverables	Specify Needs	This Sub-process acquires the detailed requirements of the users including stakeholder consultation, expected deliverables, and the purpose of the statistical products or services.	
Create Frame And Select Sample	GSBPM 5.1	Sub-process of constructing a frame and producing a sample from the frame	Collect; Frame	This Sub-process establishes the frame and selects the sample for each iteration of the collection, as specified in Sub-process Design Frame And Sample.	
Derive New Variables And Units	GSBPM 5.1	Sub-process of creating new variables or units from existing ones	Process	The Sub-process generates variables that are not available, as needed for statistical purposes. New units may be created by aggregation or by fragmentation of observed ones.	
Design	GSBPM 5.1	Phase of the statistical process which describes the development and design of the statistical outputs, methodology, and later phases of the GSBPM for the statistical activity	Phase; Design Outputs; Design Variable Descriptions; Design Collection; Design Frame And Sample; Design Processing And Analysis; Design Production Systems And Workflow	This Phase includes all the design elements needed to define the statistical products and processes identified in the business case from the Specify Needs phase. It specifies all the relevant metadata and quality assurance procedures needed to produce the data. The Sub-processes include the design of outputs, collection and processing methods, and production systems.	
Design Collection	GSBPM 5.1	Sub-process of developing the appropriate collection exchange channels for acquiring data	Design	This Sub-process includes the design of questions and response templates when the exchange channel is a questionnaire; agreements between the data supplier and consumer; and design of systems to exchange data.	
Design Frame And Sample	GSBPM 5.1	Sub-process of identifying the target population, defining the sampling frame, and selecting the sampling criteria and methodology in order to select a sample	Create Frame and Select Sample; Design	This Sub-process currently applies to survey-based processes using administrative and statistical registers.	
Design Outputs	GSBPM 5.1	Sub-process of designing statistical products or services to be released in the Disseminate phase	Design; Disseminate	This Sub-process develops the specifications of statistical outputs and the related systems used in the Disseminate Phase. Rules for accessing confidential outputs and disclosure control methods are developed here. Existing statistical standards, metadata, and practices from other statistical organisations should be used.	
Design Processing And Analysis	GSBPM 5.1	Sub-process of designing the methodology to be used in processing and analysing the data	Analyse; Design; Process	This Sub-process can include the specification routines for coding, editing, imputing, estimating, integrating, validating, and finalising data outputs.	
Design Production Systems And Workflow	GSBPM 5.1	Sub-process of specifying the systems and the sequence of activities from acquiring to disseminating data to ensure that there are no gaps or redundancies	Design	Existing processes and systems across the GSBPM should be evaluated first before designing new solutions. This Sub-process also determines the responsibilities of staff for these downstream processes and systems.	
Design Variable Descriptions	GSBPM 5.1	Sub-process of defining collected and derived variables and their related statistical classifications	Design; Design Collection; Reuse Or Build Collection Instruments	This Sub-process is where variables, either collected or derived, and classifications used in subsequent phases are either selected from existing standards or are created. The Sub-process may be conducted in parallel with the design of the collection method and instrument.	
Disseminate	GSBPM 5.1	Phase of the statistical process concerned with making statistical products available to customers by assembling and releasing a range of static and dynamic products through various channels	Phase; Update Output Systems; Produce Dissemination Products; Manage Release Of Dissemination Products; Promote Dissemination Products; Manage User Support	In this Phase, statistical organisations provide information services and products oriented to different types of users who have different needs. In some cases, the dissemination must be accompanied by technical support or special tasks to customise the information products and services to solve special requests. The Phase covers releasing and promoting the information and services, as well as providing support to the users.	
Edit And Impute	GSBPM 5.1	Sub-process of changing incorrect, inconsistent, and/or missing values by modifying or inserting new values according to pre-determined rules	Design Processing And Analysis; Process	This Sub-process is run according to pre-determined routines for editing and imputation as designed in Sub-process Design Processing And Analysis. It also includes the analysis of the edited and imputed data and the production of metadata on the editing and imputation process itself.	

Establish Output Objectives	GSBPM 5.1	Sub-process of determining the outputs in terms of quality, confidentiality and availability of resources	Specify Needs	This Sub-process includes consideration of legal frameworks, costs, and quality dimensions, such as timeliness and accuracy of the outputs.	
Evaluate	GSBPM 5.1	Phase of the statistical process concerned with the review and assessment of the experiences gained from the specific instance of a statistical business process	Phase; Gather Evaluation Inputs; Conduct Evaluation; Agree An Action Plan	This Phase is related to the review of the lessons learned and the management of the knowledge acquired each time the statistical process has taken place. This Phase reviews the experiences from each one of the other Phases. It includes evaluating the quality indicators from previous Phases of the statistical business process, drawing on a range of quantitative and qualitative inputs, and identifying and prioritising potential improvements. It logically takes place at the end of each execution of the statistical business process, but relies on inputs gathered throughout all the previous phases and is used to improve the practices of the statistical office.	
Finalise Collection	GSBPM 5.1	Sub-process of loading acquired data into a suitable electronic environment for further action	Collect	This Sub-process includes extraction and loading of data from exchange channels. It may also include analyses of paradata (process metadata) to ensure acquisition activities meet requirements.	
Finalise Data Files	GSBPM 5.1	Sub-process of creating a data set that brings together the results of the other sub-processes of the Process phase for further processing or analysis purposes	Analyse; Process	This Sub-process provides inputs into the Analyse Phase.	
Finalise Outputs	GSBPM 5.1	Sub-process of undertaking final methodological and quality checks to ensure the statistical outputs are fit for purpose and ready for release	Analyse	This Sub-process ensures that the statistics and associated information are fit for purpose and reach the required quality level, and are thus ready for use. It includes various activities such as consistency checks, drafting of all supporting documentation, clearance meetings with senior management, getting formal approval for dissemination.	
Finalise Production Systems	GSBPM 5.1	Sub-process of moving the statistical business process system into a production-ready state	Build	This Sub-process includes: <ul style="list-style-type: none"> <li>• producing documentation about the process components, including technical documentation and user manuals;</li> <li>• training the business users on how to operate the process; and</li> <li>• moving the process components into the production environment, and ensuring they work as expected in that environment.</li> </ul>	
Gather Evaluation Inputs	GSBPM 5.1	Sub-process of collecting material produced in other phases or sub-processes in order to assess the statistical business process as conducted	Evaluate	This Sub-process gathers evaluation material produced in any other phase or sub-process, including feedback from users, process metadata (paradata), quality indicators, and system metrics. The collection of some of these evaluation materials can be automated and take place in a continuous way throughout the whole process. For the evaluation of certain Sub-processes it can be necessary to perform specific activities like small surveys, e.g. coverage assessment or effectiveness of dissemination.	
Identify Concepts	GSBPM 5.1	Sub-process of identifying the concepts to be measured in the subsequent phases of the business process	Design Variable Descriptions; Specify Needs	The selection of concepts may not align with international or national standards or practices in this Sub-process. The alignment to existing standards or creation of new standards occurs in the Design Variable Descriptions Sub-process.	
Identify Needs	GSBPM 5.1	Sub-process of identifying statistical products or services needed in response to a new requirement	Specify Needs	This Sub-process is the initial step in responding to a new information request or organisational changes such as budget and resources. The Sub-process may review inputs from previous iterations or the practices of other statistical organisations.	
Integrate Data	GSBPM 5.1	Sub-process of combining data from two or more sources to obtain new and enhanced data	Process	This Sub-process includes input sources that may be of different nature (e.g. survey-based or administrative-based). Record linkage procedures or statistical matching methodologies may be used to integrate different data sources.	
Interpret And Explain Outputs	GSBPM 5.1	Sub-process of increasing understanding of the raw statistical outputs through in-depth analysis	Analyse	This Sub-process is where statisticians examine the outputs generated during the previous Sub-processes in order to interpret and explain the data and produce statistical analyses and/or reports.	
Manage Release Of Dissemination Products	GSBPM 5.1	Sub-process of ensuring that the delivery of statistical products to users is made in a timely manner and according to the organisation's release policy	Disseminate	This Sub-process manages the activities needed to be sure that the release of statistical products will be timely and satisfactory. It includes briefings for specific groups; arrangements for any pre-release embargoes; release of subscription based products; access to confidential data by authorised user groups; retracting products if errors or other wrong conditions are discovered.	
Manage User Support	GSBPM 5.1	Sub-process of ensuring that customer queries and requests for services are correctly managed and satisfied	Disseminate	This Sub-process ensures that customer queries and requests are tracked, satisfied in a timely manner, and according to organisational service levels.	

Phase	GSBPM 5.1	distinct stage of the statistical business life cycle based on interrelated activities addressing a given purpose			
Prepare And Submit Business Case	GSBPM 5.1	Sub-process of documenting the outcomes of the other sub-processes in the Specify Needs phase in order to get approval	Specify Needs	In this Sub-process, the business case is prepared to conform to the requirements of the approval body, either formal or informal. The business case typically includes the current business process, if it exists; options for new solutions; and a cost-benefit analysis.	
Prepare Draft Outputs	GSBPM 5.1	Sub-process of generating statistical outputs in preparation for dissemination	Analyse	This Sub-process includes the production of tables, charts, microdata files, and additional measurements (e.g. trends, indices, or seasonally adjusted data). It can also incorporate the recording of quality characteristics.	
Process	GSBPM 5.1	Phase of the statistical process in which input data are transformed to produce the target outputs	Phase; Integrate Data; Classify And Code; Review And Validate; Edit And Impute; Derive New Variables And Units; Calculate Weights; Calculate Aggregates; Finalise Data Files	In this Phase, the data are checked and cleaned from errors (such as inconsistencies and missing values), classified according to classification schemes. To obtain output data, input data can be integrated with data obtained from other sources and new variables can be derived. This Phase includes weights calculation and data aggregation.	
Produce Dissemination Products	GSBPM 5.1	Sub-process of generating outputs for release that satisfy user needs	Design Outputs; Disseminate	This Sub-process creates products to meet user needs according to the design specified in Sub-process 2.1 Design Outputs. These products could include interactive graphics, tables, public-use micro-data sets, data services and downloadable files. Some of the activities undertaken are preparing the product components (explanatory texts, tables, charts, quality statements etc.); assembling the components into products; editing the products; and checking that they meet publication standards.	
Promote Dissemination Products	GSBPM 5.1	Sub-process of informing and communicating about the availability of products to potential users	Disseminate	This Sub-process concerns the active promotion of the statistical products. 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.	
Reuse Or Build Collection Instruments	GSBPM 5.1	Sub-process of producing the system through which data are obtained	Build	This Sub-process describes the activities to build the collection instruments or extraction routines to be used during the Collect phase to obtain data. The collection instrument or the extraction routines are generated or built based on the design specifications created during the Design phase.	
Reuse Or Build Dissemination Components	GSBPM 5.1	Sub-process of producing or updating the systems and services needed for the dissemination of statistical products	Build	This Sub-process describes the activities for building all types of dissemination components and services, from those that are used to produce traditional paper publications to those that provide web services, open data outputs, or access to micro-data.	
Reuse Or Build Processing And Analysis Components	GSBPM 5.1	Sub-process of producing or updating systems for handling and transforming data	Build	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.	
Review And Validate	GSBPM 5.1	Sub-process of analysing and evaluating data in order to identify potential or actual errors, missing values, or outliers	Process	The Sub-process only focuses on data verification disregarding the actual correction of data. When computer assisted modes are used, part of this validation process can happen in the collection phase.	
Run Collection	GSBPM 5.1	Sub-process of actually acquiring data through collection exchange channels	Collect	This Sub-process may use different exchange channels 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 all contacts with data providers and suppliers, capturing paradata, and ensuring a continued good relationship with each data source. Simple validation steps are taken during this Sub-process as well.	
Set up Collection	GSBPM 5.1	Sub-process of identifying and establishing the people, processes, and technology to conduct data collection	Collect	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. 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.	
Specify Needs	GSBPM 5.1	Phase of the statistical process identifying what the statistical activity is going to produce and the data required to produce the expected output to satisfy the users' requirements	Phase; Identify Needs; Consult And Confirm Needs; Establish Output Objectives; Identify Concepts; Check Data Availability; Prepare And Submit Business Case	This phase is triggered when a need for new information is detected or feedback initiates a review of current statistics. The sub-processes include the identification or the review of the statistics needed, the concepts to be measured and the available data sources.	
Sub-process	GSBPM 5.1	distinct stage of the statistical business process that consists of interrelated activities	Process		

Test Production Systems	GSBPM 5.1	Sub-process concerned with the testing of assembled and configured services and related workflows	Build; Reuse Or Build Processing And Analysis Components	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 (Reuse or build processing and analysis 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 of processes, information and services.	
Test Statistical Business Process	GSBPM 5.1	Sub-process of managing a field test or pilot of the statistical business process	Build	In this Sub-process, the entire system from data acquisition to dissemination is tested and its effectiveness is evaluated. Adjustments are made based on the evaluation to ensure the system implements the process correctly.	
Update Output Systems	GSBPM 5.1	Sub-process of assuring that the systems where data and metadata are stored are ready for dissemination purposes	Disseminate	This Sub-process comprises activities to make a final check and assure that data and metadata are ready to be disseminated, 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.	
Validate Outputs	GSBPM 5.1	Sub-process of assessing the quality of outputs in accordance with a general quality framework and expectations	Analyse	This Sub-process includes building a body of knowledge about a specific statistical subject domain. This knowledge is applied to the current situation to identify divergence from quality targets and to allow informed analyses. This process is usually undertaken by a subject matter expert.	