

This module is part of the

Memobust Handbook

on Methodology of Modern Business Statistics

26 March 2014

Theme: GSBPM: Generic Statistical Business Process Model

Contents

General section	3
1. Summary	3
2. General description.....	3
2.1 Versions of the GSBPM	3
2.2 The process structure of the GSBPM	3
2.3 Different uses of the GSBPM.....	5
2.4 The contents of the handbook.....	6
2.5 Remarks	8
3. Design issues	8
4. Available software tools	8
5. Decision tree of methods	8
6. Glossary.....	8
7. References	8
Interconnections with other modules.....	10
Administrative section.....	11

General section

1. Summary

The Generic Statistical Business Process Model (GSBPM) is a means to describe statistics production in a general and process-oriented way. It is used both within and between statistical offices as a common basis for work with statistics production in different ways, such as quality, efficiency, standardisation, and process-orientation. It is used for all types of surveys, and “business” is not related to “business statistics” but refers to the statistical office, simply expressed. The GSBPM has been used as a basis for this handbook, in line with the instruction from Eurostat on using standards. The model is described below in brief.

The handbook is structured with some twenty topics. These topics are related to the GSBPM, but they do not correspond precisely with the phases and sub-processes. There is an inherent difference in that this handbook is restricted to methodology. A mapping between the GSBPM sub-processes and the handbook topics is provided, and some further comments are made.

2. General description

2.1 *Versions of the GSBPM*

Version 4 of the Generic Statistical Business Process Model (GSBPM), which was released in 2009, has been used by the Memobust project. In December 2013 version 5.0 was released, too late to use in the current handbook. The main differences are the following, as stated by UNECE (2013b):

- Phase 8 (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.
- 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 the sub-processes have been updated and expanded where necessary. The terminology used has been changed to be less survey-centric, in recognition of the growing use of non-survey sources (administrative data, big data etc.).

Hence, there are no principal differences, which would cause negative effects. Please note that in this module the term survey is not restricted to direct data collection as above; it is used in a broad sense.

2.2 *The process structure of the GSBPM*

The Generic Statistical Business Process Model (GSBPM) is used by statistical offices. Statistics New Zealand was probably first and presented such a model around 2006. Much work has been done in joint UNECE/Eurostat/OECD Work Sessions on Statistical Metadata (METIS), see for instance UNECE (2009) and documentation from METIS. The GSBPM currently comprises four levels:

- Level 0, the statistical business process;
- Level 1, the nine phases of the statistical business process;
- Level 2, the sub-processes within each phase;
- Level 3, a description of those sub-processes.

The nine phases on the first level are the following:

1. Specify Needs;
2. Design;
3. Build;
4. Collect;
5. Process;
6. Analyse;
7. Disseminate;
8. Archive;
9. Evaluate.

The phases 1-3 can be regarded as preparatory, phases 4–7 correspond to the “obvious” production, phase 8 is a saving for the future (of essential data and metadata), and phase 9 summarises and formulates an action plan. Figure 1 below is the same as that presented by UNECE (2009), i.e., sub-processes on a two-digit level are included. Statistics Sweden, for instance, has found it useful to have further levels of sub-processes in some cases to elaborate more and give more detail and support. There is then a hierarchy with numbering on successive, more detailed, levels with 3 digits etc. This is the case for example for phase 4 *Collect* where the sub-processes are specified in more detail considering different collection modes, reminders etc. (There is unfortunately no reference in English.)

Quality Management / Metadata Management								
1 Specify Needs	2 Design	3 Build	4 Collect	5 Process	6 Analyse	7 Disseminate	8 Archive	9 Evaluate
1.1 Determine needs for information	2.1 Design outputs	3.1 Build data collection instrument	4.1 Select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Define archive rules	9.1 Gather evaluation inputs
1.2 Consult & confirm needs	2.2 Design variable descriptions	3.2 Build or enhance process components	4.2 Set up collection	5.2 Classify & code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Manage archive repository	9.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design data collection methodology	3.3 Configure workflows	4.3 Run collection	5.3 Review, Validate & edit	6.3 Scrutinize & explain	7.3 Manage release of dissemination products	8.3 Preserve data and associated metadata	9.3 Agree action plan
1.4 Identify concepts	2.4 Design frame & sample methodology	3.4 Test production system	4.4 Finalize collection	5.4 Impute	6.4 Apply disclosure control	7.4 Promote dissemination products	8.4 Dispose of data & associated metadata	
1.5 Check data availability	2.5 Design statistical processing methodology	3.5 Test statistical business process		5.5 Derive new variables & statistical units	6.5 Finalize outputs	7.5 Manage user support		
1.6 Prepare business case	2.6 Design production systems & workflow	3.6 Finalize production system		5.6 Calculate weights				
				5.7 Calculate aggregates				
				5.8 Finalize data files				

Figure 1. The GSBPM according to UNECE (2009)

All models are simplifications. Statistics production is not a simple process with successive sub-processes. There are, for instance, feedback loops, which are not explicit in the model. Many sub-processes need much more detail than shown here.

UNECE (2009) states that according to general process modelling theory, each sub-process should have a number of clearly identified attributes. It is clearly important to be aware of input(s), output(s), and purpose or value added. Furthermore, it is stated that other characteristics to take into account are the process owner with responsibility for the process, guides (for example manuals and documentation), and enablers: both people and systems. These attributes should be stated explicitly in order to make the statistics production flow smoothly. It is also stated that the attributes are likely to differ, at least to some extent, between statistical business processes, and between organisations.

2.3 *Different uses of the GSBPM*

The GSBPM model is useful as a common framework for all types of statistical surveys or processes, for instance both with direct data collection and when administrative data are used. Different types of statistical processes are described, for example by Eurostat (2009) in the European context and also in the theme module “General Observations – Different Types of Surveys”.

The six European types of statistical process described by Eurostat (2009) include “Statistical compilation”, which assembles a variety of primary sources to obtain an aggregate, with a special conceptual significance. Many of these are economic aggregates such as the National Accounts and the Balance of Payments. The GSBPM is then useful at least on a high level with data collection referring to the primary sources.

The GSBPM has been used also to describe production of registers, see UNECE (2011). There is some dissimilarity between maintenance of statistical business registers and production of statistics. The former is more or less constantly updated from a set of administrative and statistical sources. The latter is more from-start-to-end for each production round of a survey. The outputs from statistical registers are (i) registers and frames, and (ii) statistics based on the register. See the topic “Statistical Registers and Frames” (several modules) and the module “Dynamics of the Business Population – Business Demography”, for such outputs. The module mentioned gives examples of statistics based on the business register, for instance different types of birth depending on enterprise characteristics.

The GSBPM is useful for a new survey, when a survey is re-designed, and for continuous improvements of a repeated survey. The balance between the phases varies, for instance between situations such as those just mentioned. The three first phases may require considerable efforts for a new design and a re-design. They may be brief in repeated surveys, but they must not be skipped, when there is information to use, typically from the evaluation phase. The preparatory phases may then include, for instance, renewed contacts with users about priorities, modifications of some variables and possibly a data source, a renewed sample allocation, changes in the allocation of survey resources, and improvements of the IT-system.

Each sub-process involves methods, tools, and routines. The statistical office may choose to standardise or limit the ways in which a process can be run, at least for the majority of its statistical surveys. There could, for instance, be standard tools for sampling, for electronic data collection, for coding, and for imputation – tools that are broad enough to be useful for many surveys.

Similarly, the GSBPM is used also on the international level.

2.4 *The contents of the handbook*

It is necessary to note that this handbook is focused on methodology. This is especially obvious in method modules but also in theme modules. Hence, some sub-processes of the GSBPM get attention in several modules, whereas others are hardly mentioned. It has to be remembered also that this is an early version of the handbook. Some topics will expand later on, when there has been more development and agreement on recommended practices, and also more time for writing. Figure 2 below is a modified illustration of the GSBPM that is tied to the current Memobust project and handbook. The figure has two main goals.

Firstly, this version indicates the degree of methodological content in each sub-process. A scale with three categories is used: green for high methodological content, light green for an intermediate methodological content, and light yellow for little or no methodological content. The judgements are rough, and they have been made in the project based on UNECE (2009).

Secondly, there is a mapping between this GSBPM-version and the topics of the handbook. Again, the goal is to give an overview without details. Most topics are sorted into one or sometimes two sub-processes. This shows the main content of the topic. Hence, Figure 2 gives an overview in both directions between handbook topics and GSBPM sub-processes.

There are many handbook modules for phase 5 Process and phase 6 Analyse. Several of these include design aspects, in the module itself or in a separate module (an overall theme or a specific design module). Sub-process 2.5 Design statistical processing methodology is treated in several handbook modules. In Figure 2 they are shown in phases 5 and 6 but not in sub-process 2.5.

Some topics are quite broad and do not fit well into just one or a few sub-processes. Three such topics are:

- “General observations”
- “Overall Design”, which takes most of phase 2 *Design* into account with choices, allocations, and also coordination and optimisation aspects. This topic could, to some extent, be put in sub-process 2.6 *Design production systems and workflow* to emphasise its focus on overall issues.
- “Repeated Surveys”, which emphasises both user aspects and producer possibilities; these are typical when a survey is made regularly.

A few further topics are:

- “Response”, with modules about the response process and response burden, provides information to several sub-processes, for example about variables, data collection, and sampling with regard to response burden.
- “Quality Aspects”; this topic is shown in sub-process 6.3 *Scrutinize and explain*. It is related also to sub-process 6.5 *Finalize outputs* and it has more information, for instance about variance estimation and quality components in general.

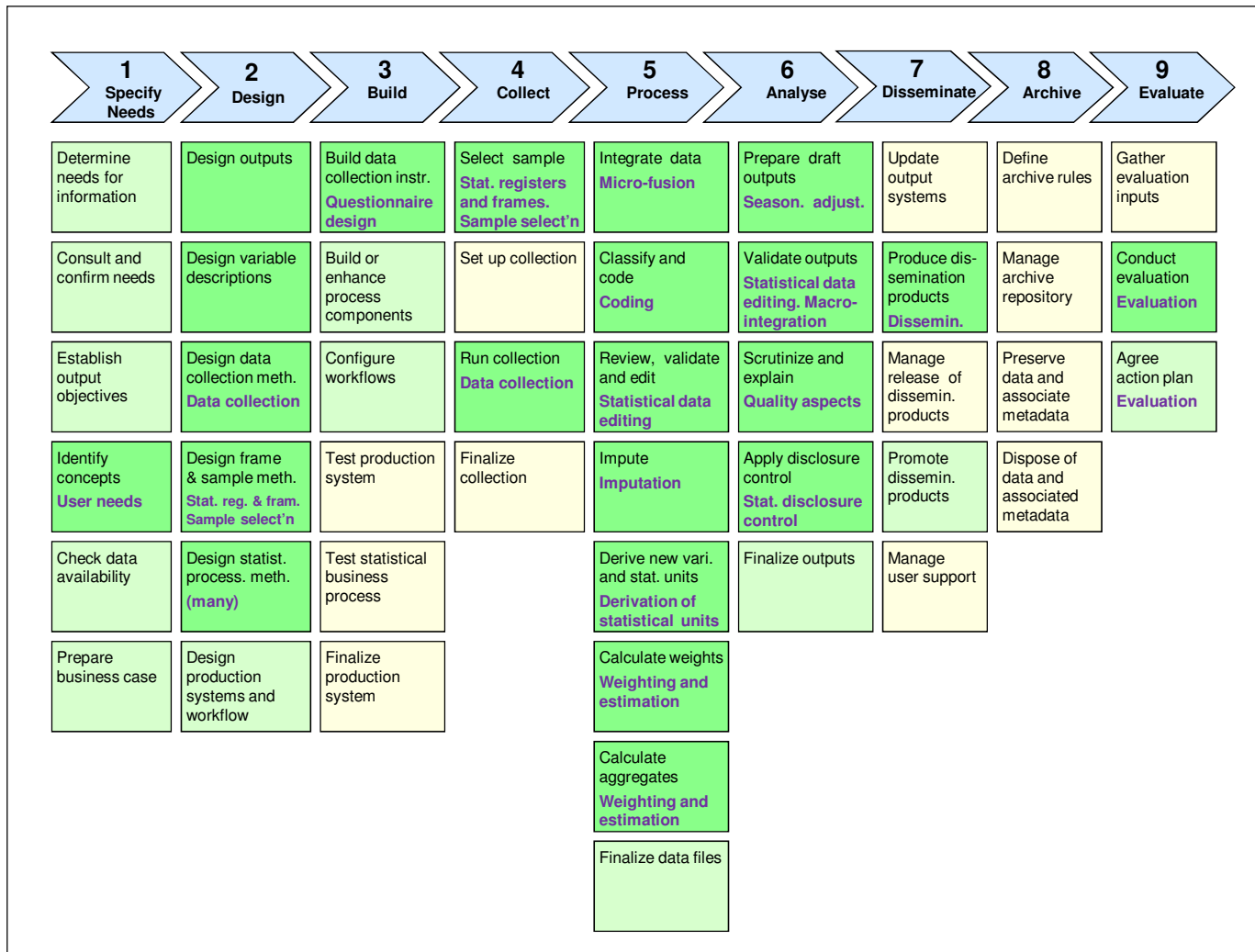


Figure 2. The GSBPM modified to describe methodological content and the handbook topics.

The topic “Statistical Registers and Frames” can be regarded in two different ways when it comes to the GSBPM. One of them was described above in section 2.1, i.e., the on-going building and maintaining of the statistical register and frames; this is a statistical process in its own right. The second way to consider this topic is in the context of a production round of a survey: then the topic provides information for sub-processes 2.4 *Design frame and sample methodology* and the sub-process 4.1 *Select sample*, which includes the establishment of the frame. The topic “Dynamics of the Business Population has contents that are related to the topic “Statistical Registers and Frames”.

Comments on some sub-processes that have no handbook module in Figure 2 follow. Sub-process 2.1 *Design outputs* is related to phase 1 *Specify Needs* (especially sub-process 1.4 *Identify concepts*) and also phases 6 *Analyse* and 7 *Disseminate*. Sub-process 2.2 *Design variable descriptions* have similar relationships and even more to sub-process 2.3 *Design data collection methodology* and sub-process 3.1 *Build data collection instrument*.

2.5 Remarks

The GSBPM is a practical reference when developing a standard set of methods, tools, and routines for the statistical office. Statistics Netherlands and Statistics Sweden, which both developed their versions before the joint international one, have long experience of the value of such a model as a fundamental part of statistics production, improvements, and developments. Statistics Sweden has an internal support system based on its GSBPM.

Communication is further facilitated and standardised through the more recent GSIM, Generic Statistical Information Model. This model is too recent to have been used in this handbook. GSIM is intended as a complement to the GSBPM. Information is easily found on the Internet, for example in the description by UNECE (2013a) about different activities for modernisation of statistical production and services; there is also a link to GSIM.

3. Design issues

4. Available software tools

5. Decision tree of methods

6. Glossary

For definitions of terms used in this module, please refer to the separate “Glossary” provided as part of the handbook.

7. References

Eurostat (2009), *ESS Handbook for Quality Reports (EHQR)*. This handbook (planned to be revised soon) is accessible on the webpage of Eurostat, currently:

http://epp.eurostat.ec.europa.eu/portal/page/portal/ver-1/quality/documents/EHQR_FINAL.pdf

UNECE (2009), Generic Statistical Business Process Model. Version 4.0 – April 2009 (prepared by the UNECE Secretariat). Joint UNECE/Eurostat/OECD Work Session on Statistical Metadata (METIS).

UNECE (2011), Applying the Generic Statistical Business Process Model to business register maintenance. UNECE Conference of European Statisticians, Group of experts on Business Registers, Twelfth session, Paris, 14-15 September 2011.

UNECE (2013a), What's New from the High-Level Group? Working paper from the UNECE (prepared by S. Vale) to the Meeting on the Management of Statistical Information Systems (MSIS 2013).

UNECE (2013b), Generic Statistical Business Process Model. Version 5.0 – December 2013. The United Nations Economic Commission for Europe (UNECE). See:
<http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0>

Interconnections with other modules

8. Related themes described in other modules

(Restricted to modules on a high level)

1. General Observations – Different Types of Surveys
2. Overall Design – Overall Design
3. Repeated Surveys – Repeated Surveys
4. Statistical Registers and Frames – Main Module

9. Methods explicitly referred to in this module

- 1.

10. Mathematical techniques explicitly referred to in this module

- 1.

11. GSBPM phases explicitly referred to in this module

1. All nine GSBPM phases

12. Tools explicitly referred to in this module

- 1.

13. Process steps explicitly referred to in this module

1. The process steps on level two are shown in Figure 1 (and Figure 2 without number).

Administrative section

14. Module code

General Observations-T-GSBPM

15. Version history

Version	Date	Description of changes	Author	Institute
0.1	28-02-2013	first version	Eva Elvers	Statistics Sweden
0.2	17-04-2013	extended, mapping	Eva Elvers	Statistics Sweden
0.2.5	16-11-2013	updates, figure	Eva Elvers	Statistics Sweden
0.2.6	26-11-2013	most of the EB comments	Eva Elvers	Statistics Sweden
0.2.7	26-11-2013	preliminary release		
0.2.8	20-01-2014	note: new GSBPM version	Eva Elvers	Statistics Sweden
0.3	10-02-2014	updates	Eva Elvers	Statistics Sweden
1.0	26-03-2014	final version within the Memobust project		

16. Template version and print date

Template version used	1.0 p 4 d.d. 22-11-2012
Print date	21-3-2014 17:23