

Steady States – An Overview

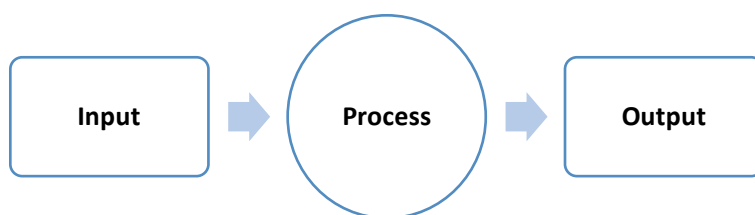
Purpose

This document provides an overview of value chains and steady states, the benefits of and examples of using steady states in the statistical process, explains how steady states relate to other production models at Statistics New Zealand and gives examples of how steady states can enable data sharing and improve data management.

Value chains

1. What is a value chain?

A value chain is a business management concept where production is viewed as a sequence of inputs, transformation processes and outputs. Each transformation process adds value, which accumulates to the final product.



The value chain is a concept often used in manufacturing and a common example used to explain this concept is the diamond industry. The stages in the diamond value chain include mining, sorting, cutting, polishing, and jewellery creation. Each stage adds significant monetary value to the final product.

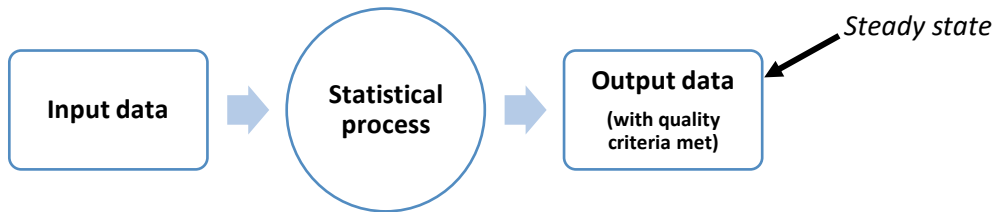
2. What is the statistical value chain?

The flow of information through the statistical process can be viewed as a statistical value chain. We collect raw data and transform it into data with higher value. Data goes through many transformation processes and each adds value to the final product.

Steady states

1. What are steady states?

Steady states are specified stages in the value chain where the output of a process meets a set of pre-agreed quality indicators. This concept is easily applied to the statistical value chain where data enters, a process occurs and data that meets certain criteria exits. This data is a steady state.



Data goes through many processes and not all are equally important. The use of steady states identifies the milestones in the process which provide assurance that an acceptable level of value has been added to the input data.

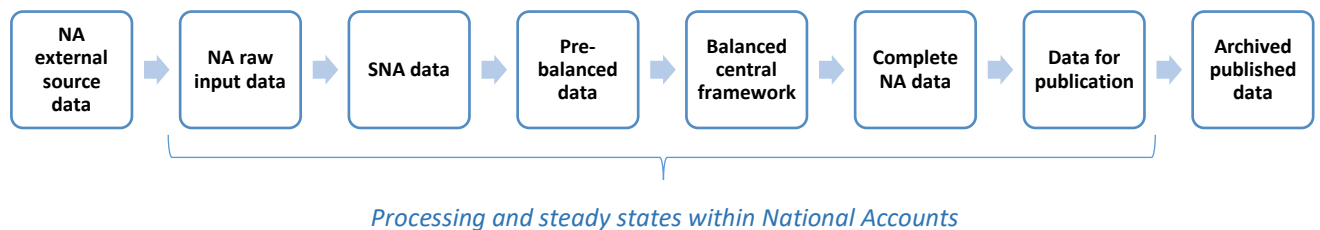
2. An example of steady states: Statistics Netherlands

The business architecture of Statistics Netherlands recognises six steady states, shown in the simplified graphical representation below. In practise the states are not necessarily linear. An output dataset that is a steady state might be consumed by a different statistical process or a step may be repeated.



1. *Pre-input base*: Raw input source data.
2. *Inputbase*: Raw source data as it enters processing stage.
3. *Microbase*: Statistical microdata that has completed processing (editing, imputation, integration).
4. *Statbase*: Aggregated statistical data.
5. *Outputbase*: Secured micro datasets and aggregated datasets for (external) publication purposes that form the starting point for creating products for external users (see below).
6. *Post-output base*: Information that has been disseminated, including published data, press releases, analyses and other contextual information.

3. An example of steady states: Statistics New Zealand National Accounts high-level design for the MATAI project



The MATAI project recognises eight steady states for National Accounts, six that are part of the core National Accounts process. Again, this is a simplified graphical representation and the states are not necessarily linear in practice.

1. *National Accounts external source data*: data supply external to the core annual National Accounts, either from other areas within Statistics New Zealand or from external sources.
2. *National Accounts raw input data*: source data as it enters the National Accounts production process.
3. *System of National Accounts data*: the raw input data has been transformed to SNA variables, classifications and is at the right working level.
4. *Pre-balanced data*: key variables and indicators have been derived. There are no large measurement errors in the data. Time series are fit-for-purpose.
5. *Balanced central framework*: data where all identities in the central integration framework are balanced.
6. *Complete National Accounts*: data is complete and complies with all (NZ) SNA definitions and quality requirements. No adjustments after this point.
7. *Data for publication and story-telling*: data is ready for meeting the customer and data quality requirements and steady state guarantees confidential data isn't disclosed.
8. *Archived published data*: published data that is archived and accessible in the long-term.

4. Attributes of data in a steady state

Since steady states provide points in the data value chain where an acceptable level of quality is achieved, the quality criteria along with the measures and tolerances for each of those criteria must be agreed as a start.

Establishing these criteria prior to processing means that steady states can serve as strategic monitoring points, helping to identify and isolate potential issues before they become increasingly significant problems further along the value chain.

The criteria or attributes of steady states are determined based on the needs of the organisation but can include the following:

Well-defined

Steady state data must have quality metadata. The concepts, variables, reference period and other relevant metadata must be defined, as well as being standardised across multiple datasets.

Quality standards met

To be a steady state the data must satisfy quality standards. These standards must be clearly defined in advance. Data can be reprocessed until standards are met.

User

Each steady state must have at least one defined user to be relevant and valuable.

Custodianship

Each steady state must have an 'owner' responsible for the data design and realisation of quality. At Statistics New Zealand this role would currently be referred to as the 'data custodian'. The data custodian is responsible for authorising use.

Available for common reuse

A steady state should be able to be picked up and become an input into another relevant business process. Structural, descriptive and quality metadata allows for this reuse. Centralised storage of steady states (i.e. datasets) can help facilitate this in practice.

5. Benefits of using steady states

There are numerous benefits for identifying and using steady states in statistical production:

- Steady states allow for the adoption of a modern approach to statistical production where data is used for several purposes/outputs. A steady state can become an input into another process at any point.
- Steady states can subsequently be used for production design and even for designing organisational structures.
- Steady states allows the business to clearly identify the information assets it holds, allowing for improved information risk, security and capacity management.
- Having quality defined upfront is more efficient than trying to continuously improve data quality throughout the production process.
- Steady states fundamentally have good metadata management, which supports an environment where data is more open and reusable allowing more value to be extracted.
- Steady states allow for fall-back scenarios. If a step in the production process fails, there is a predefined and steady based to return to.
- The different pieces in the statistical value chain are clearly identified, making the work planning and prioritisation process clearer.

Steady states and other models

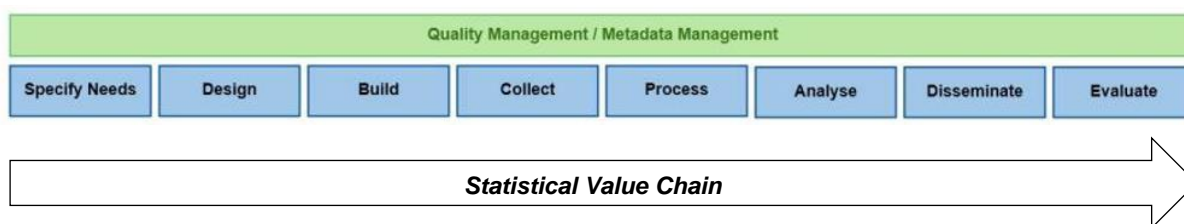
1. Steady states and quality gates

Steady states are similar to the concept of quality gates. Quality gates are a project management concept and in simple terms they refer to points in a project's life cycle when a steering committee or change board can decide whether to continue with the project or not.

In wider process management, a quality gate is a point in a process where the quality of a product is assessed to see if it is meeting the need for which it was designed. Based on a set of agreed quality criteria and associated measures defined by acceptable tolerances, quality gates can be used to approve output datasets to move on to the next point in the value chain or to initiate corrections or re-processing to mitigate unacceptable levels of errors uncovered in the output datasets.

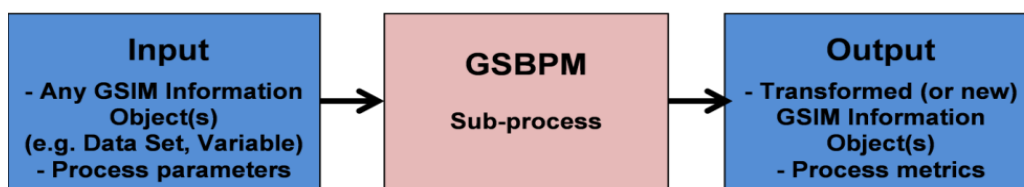
2. Steady states and the Generic Statistical Business Process model (GSBPM)

The statistical process detailed in the GSBPM is also known as a statistical value chain. The GSBPM describes the processes or transformations that data undergoes along this value chain. Steady states can therefore also be mapped to the processes and sub-processes of the GSBPM.



3. Steady states and the Generic Statistical Information Model (GSIM)

The GSIM is a reference framework for statistical information. It provides a set of standardised, consistently described information objects, which are the inputs and outputs along the statistical value chain.



The key use of GSIM is to provide a common language for statistical information and it can be applied in this way when referring to the inputs and outputs of statistical processes that result in steady states.

4. Steady states and Statistics New Zealand’s statistical disposal authorities

Statistics New Zealand has five recognised states along the statistical value chain where data and its associated metadata have agreed disposal outcomes. Disposal refers to what happens to the data once it has fulfilled its statistical or business use. These outcomes are agreed with the Chief Archivist in documents called disposal authorities (DAs).

These points in the value chain are similar to steady states in that they recognise specific instances of output data throughout the statistical process. However, the value added to obtain each state is not clearly enough articulated for these to formally be called steady states.

These Disposal Authority states are:



1. *Received admin data and statistical schedules (DA271 & DA272)*: Administrative data and statistical schedules as received by Statistics New Zealand. These can be destroyed once data has been extracted/input into the next stage of the production process.
2. *Operational datasets (DA379 1.1.1)*: The first versions of datasets created from collected/received data. These datasets can be destroyed once data has been extracted/input into the next stage of the production process.
3. *Processing datasets (DA379 1.1.2)*: These are working versions of data created during processing and are considered transitory and can be destroyed.
4. *Final definitive datasets (DA379 1.2)*: Datasets that have experienced all quality checks and editing procedures and are considered the final definitive microdata for are particular series/output. These datasets must be retained for the long-term.
5. *Archived or subject to legal deposit*: Data and associated information that has been disseminated, is subject to legal deposit with the National Library of New Zealand. Under legislation, all publishers in New Zealand must deposit their publications with the National Librarian enabling the National Library to collect, preserve, and make available the documentary heritage of New Zealand.

Statistics New Zealand also operates a data archive for microdata. Copies of *final definitive datasets (DA379 1.2)* are archived, meeting disposal authority criteria for data in this state.

Enabling data sharing and improving data management with steady states

Because steady states reflect pre-defined quality standards, they can be implemented to help enable data sharing and improve data governance and data management across the System.

Example 1

A steady state can be identified in the business process of another agency where their data is described and classified to a pre-agreed specification that is most conducive to integration with other data. This steady state then becomes the preferred point in the value chain where data is shared, before the supplying agency moves the data further through their normal business process.

Example 2

Data from a Statistics New Zealand survey collection has been received and cleaned to an agreed standard and is now a steady state. This data can then provide a consistent input to multiple statistical activities.

Example 3

In the statistical data value model where data no longer necessarily flows in a linear manner from collection to dissemination, identified steady states across the organisation can form the basis of a data governance model. Roles and responsibilities, such as data custodianship, data policies and data quality standards can all be mapped onto the organisation's statistical process model, using the steady states as anchor points.

Sources/Further reading

Braaksma, B. and Zeelenberg, K. (2014). Information Management as tool for standardization in statistics. <https://www.cbs.nl/nl-nl/achtergrond/2014/07/information-management-as-tool-for-standardization-in-statistics>

Dunnet, G. et al (2015). Managing National Accounts Redevelopments in New Zealand. [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=STD/CSSP/WPNA\(2015\)11&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=STD/CSSP/WPNA(2015)11&docLanguage=En)

Herrman, I. et al. (2013). The Statistical Value Chain – a Benchmarking Checklist for Decision Makers to Evaluate Decision Support Seen from a Statistical Point-Of-View. http://orbit.dtu.dk/files/101179378/THE_STATISTICAL_VALUE_CHAIN.pdf

Kole, J. (2015). MATAI High-level Design. <Notes://DRDOMMAIL01/CC2574AB000FA882/735CE27ED55E107CCC256E8B001C686A/7EEB9FE71978D8AECC257E9D0026BC7D>

Renssen, R and Goossens, H. (2010). Underlying Concepts of the Data Service Centre at Statistics Netherlands. <Notes://DRDOMMAIL01/CC25755700713AE8/735CE27ED55E107CCC256E8B001C686A/032A1D5974F908DECC257B9C007DE8BE>

Van Delden, A. et all. (2009). Chain management in statistics: Best practices. <Notes://DRDOMMAIL01/CC25755700790F5B/F6E01B6BE8D67A1ACC256E8B001C6868/5F115B6D1E680F3FCC257A5C00753535>
https://www.researchgate.net/profile/Arnout_Delden/publication/318453616_Chain_management_in_statistics_Best_practices/links/596b44e8458515e9afb1c3ba/Chain-management-in-statistics-Best-practices.pdf

Van Delden, A.& Renssen, R. (2009). Standardisation of design and production of statistics. <Notes://DRDOMMAIL01/CC25755700790F5B/F6E01B6BE8D67A1ACC256E8B001C6868/5F115B6D1E680F3FCC257A5C00753535>
<https://www.cbs.nl/-/media/imported/documents/2009/34/2009-34-x10-pub.pdf?la=nl-nl>
https://www.researchgate.net/publication/318453336_Standardisation_of_design_and_production_of_statistics_a_service_oriented_approach_at_Statistics_Netherlands