Regional Course on
Statistical Business Registers:
Data sources, maintenance and
quality assurance
Perak, Malaysia
21-25 May, 2018

5.1

Quality improvement

UNITED NATIONS
SIAP
Statistical Institute for
Asia and the Pacific

HOW DO WE DEFINE QUALITY?

Quality of official statistics is considered a **multi-dimensional** concept. Quality dimensions include things like timeliness, relevance and accessibility.

HOW DO WE DEFINE QUALITY?

UNSD provides general guidance on quality frameworks for official statistics with its **Generic National Quality Assurance Framework**, but every country should have their own detailed quality assurance policy tailored to its National Statistical System.

SBR QUALITY

With regards to SBR quality, we will discuss the following six dimensions:

- ▸ Relevance
- ▸ Accuracy
- ▸ Timeliness

- ▸ Accessibility
- ▸ Comparability
- ▸ Coherence

SBR QUALITY

Relevance

The SBR contains the units and characteristics necessary to support the production of statistics according to **user needs**.

Accuracy

The information contained in the SBR is correct.

SBR QUALITY

Timeliness

The data provide as **current** a picture of the real world as possible. Users know when to expect frozen frames and SBR statistics.

Accessibility

For internal users, this means access to microdata via direct access. For external users, this means access to aggregate tables.

SBR QUALITY

Comparability

The concepts and methods used to maintain units in the SBR do not change **across regions** or **over time.**

Coherence

Internal coherence refers to coherence of units and and characteristics within the SBR and **coherence with other registers** concerns maintaining links (e.g., with a common ID).

SBR QUALITY

One of the main challenges for SBR quality is that it is highly dependent on the quality of data from administrative sources, which are **outside the control** of the National Statistics Office.

SBR QUALITY

Since the SBR brings together information from a variety of sources, quality considerations will **not be the same** for all units and characteristics.

SBR QUALITY

SBR data should therefore be split into subsets and **quality indicators** should be developed for each. Subsets should be based on:

- ▸ Data source
- ▸ Unit type and/or
- ▸ Group of characteristics

SBR IMPROVEMENT SURVEYS

Two purposes:

- ▸ Error detection and filling in missing values
- ▸ Measuring the accuracy of SBR data

QUALITY ASSESSMENT

In addition to SBR improvement surveys, quality of SBR data can be measured in a number of ways…

QUALITY ASSESSMENT

Comparison with other sources

Comparing SBR data with data available from external sources to see if values are sufficiently close (compliance).

Internal consistency

Checking if values are plausible based on other characteristics of the unit or theoretical constraints.

QUALITY ASSESSMENT

Comparison with past values

Comparing current values to past values to see if their change over time is plausible.

Source and recency of data

Checking how recently values were updated and which sources and methods were used.

QUALITY INDICATORS

Quality indicators can be calculated for SBR:
- ‣ Inputs
- ‣ Processes
- ‣ Outputs

INPUT INDICATORS - EXAMPLES

Time lag

Difference between the date data are available and their reference period.

Missing value rate

Number of missing values for each characteristic of supplied data.

PROCESS INDICATORS - EXAMPLES

Number of new records

Comparison of unit creations over time can help identify problems in supply.

Active status concordance rate

Number of active units in both the source and the SBR out of the number of units in both.

OUTPUT INDICATORS - EXAMPLES

Time lag for survey frames

Time lag between the reference period for a survey and the delivery of the survey frame.

Error rates

For example, percentage of units with wrong contact information or activity code.

UNITED NATIONS
SIAP
Statistical Institute for
Asia and the Pacific

SBR QUALITY POLICY

An SBR quality policy lays out:
- How quality is defined and measured for your SBR
- What information about SBR quality will be communicated to users
- The decision to initiate a quality improvement programme

SBR QUALITY REPORTS

A good practice is to communicate with users about the quality of the SBR data using (ideally sub-annual) quality reports.

SBR QUALITY REPORTS

Quality reports should communicate:
- ▸ An assessment of the quality of the SBR
- ▸ Significant changes that have been made that will affect sampling frames

CASE STUDY: THE NETHERLANDS' DATA SOURCE QUALITY CHECKLIST

The Netherlands uses a checklist, which includes yes or no questions and numeric indicators, to evaluate data sources according to:

▸ General information about the data source
▸ The metadata supplied
▸ The data itself

CASE STUDY: THE NETHERLANDS' DATA SOURCE QUALITY CHECKLIST

General information about the data source includes questions such as:
- Does the data source satisfy information demand?
- Cost of the data source
- How punctually can the data source be delivered?

CASE STUDY: THE NETHERLANDS' DATA SOURCE QUALITY CHECKLIST

Metadata indicators include:
- ▸ Clarity scores for definitions of units and variables
- ▸ Are modified values marked and how?

CASE STUDY: THE NETHERLANDS' DATA SOURCE QUALITY CHECKLIST

Data indicators include:

▸ Percentage of units linked unambiguously

▸ Percentage of cells with missing values

CASE STUDY: THE NETHERLANDS' DATA SOURCE QUALITY CHECKLIST

Only if quality evaluations for each of the three areas are deemed successful can the data source be used.

UNITED NATIONS
SIAP
Statistical Institute for
Asia and the Pacific

CASE STUDY: COLOMBIA'S QUALITY INDICATORS

Colombia has recently proposed using six quality indicators for its SBR data, including:

- ▸ Income precision: ratio of total income by sector in the SBR to the official NSO national accounts data
- ▸ Ratio of the number of units in the SBR for a given year to the number of unique units in the tax register in the same year

UNITED NATIONS
SIAP
Statistical Institute for
Asia and the Pacific

REFERENCES

This presentation was primarily based on UNECE's Guidelines on Statistical Business Registers, in particular Chapter 10. Other material used:

- ▸ The African Development Bank's Guidelines for Building Statistical Business Registers in Africa
- ▸ Eurostat's Business registers: Recommendations manual

Presentation template by SlidesCarnival

UNITED NATIONS
SIAP
Statistical Institute for
Asia and the Pacific