

Workshop on the Implementation of a National Quality Assurance Framework for Official Statistics in countries of the Africa Region

Addis Ababa, Ethiopia, 6-9 December 2022

Session 4.3: Review of additional examples and practices

Matthias Reister, reister@un.org
Chief, Development Data Section
Statistics Division, Development Data and Outreach Branch
United Nations Department of Economic and Social Affairs



Example 1: Use of novel data sources for consumer prices

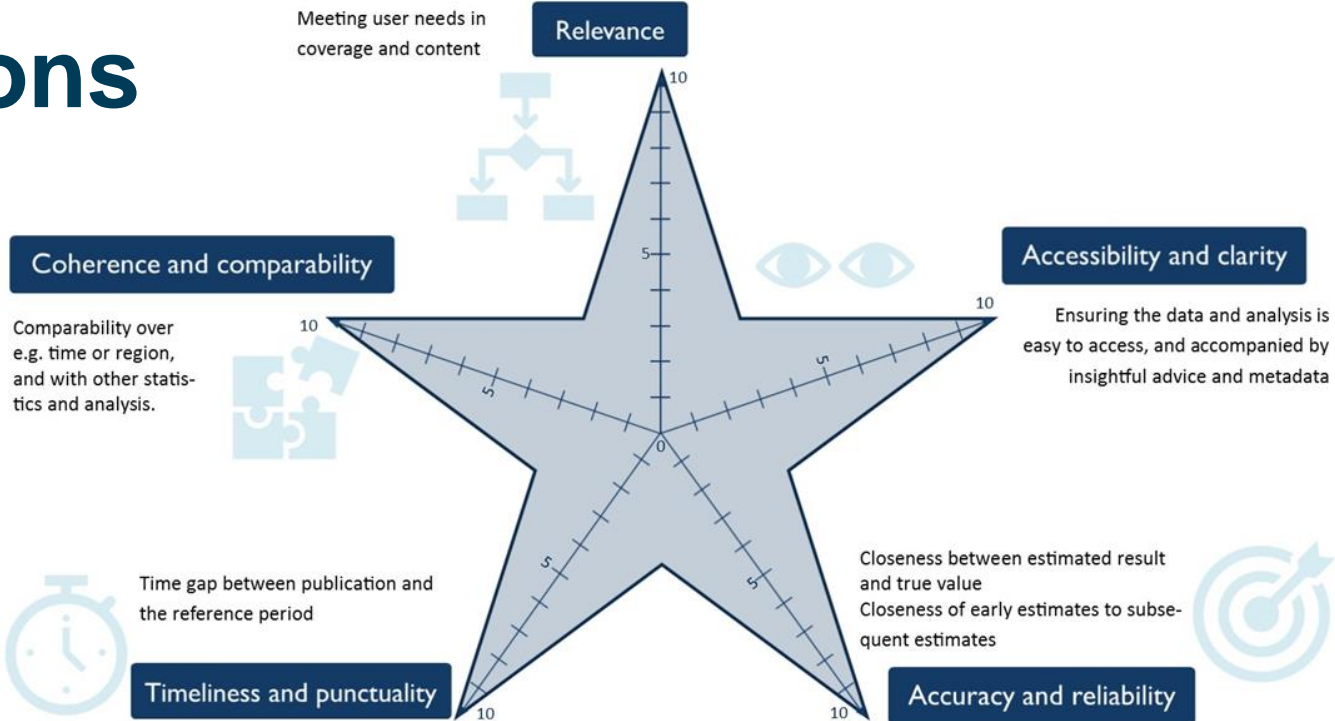
Catherine McGirr

Office for National Statistics, UK

Quality assurance and new data sources, 9th Feb 2021

Available at: <https://unstats.un.org/unsd/statcom/52nd-session/side-events/20210209-1M-Quality-assurance-and-new-data-sources/>

Quality dimensions



What are alternative data sources?

- Scanner data
- Web scraped data
- Used in conjunction with data as currently collected

Why use alternative data sources?

- Improved coverage – for example regional
- High frequency of collection
- Reduced burden
- More granular information (eg expenditure)

Comparison of different data sources

<i>Data dimension</i>	Traditional	Scanner data	Web scraping
<i>Data acquisition</i>	Manual	Automated	Automated
<i>Completeness/scope</i>	Sample items from sample of retailers	All transactions (census) from medium to large retailers	Bulk or sample from online retailers
<i>Metadata</i>	Item description + some attributes	Item description + (some) attributes Backseries may be available	Item description + attributes
<i>Quantity data</i>	None	Quantities sold	None
<i>Timing</i>	Single collection day	Daily	Daily/weekly

Ensuring quality of the data

- Initial data checks
- Standardising the data
- Monitoring the regular feeds
- Data sharing agreements
- Quality Assurance of Admin Data (QAAD)

Ensuring quality of the outputs

- Robust statistical systems
- Research into new methods
- Impact analysis
- Incorporate QA into regular monthly production
- Advisory panels – stakeholder and technical
- International best practice

Example2: Quality reports for registers and administrative data

4 MARCH 2020, NY

LASSE.SANDBERG@SSB.NO

See <https://unstats.un.org/unsd/statcom/51st-session/side-events/20200304-2L-quality-assurance/>



Statistisk sentralbyrå
Statistics Norway

Summary

- Starting around 2011/2012, Statistics Norway introduced a new approach to professionalise cooperation with data owners, which included MoUs and Quality reports. Experience with first register-based population and census showed need to correct data at the source.
- The statistics act gives Statistics Norway the right to **access** registers, but just as important, it also gives the right to **influence** registers. (standards, variable definitions, ..)
- “Agreements on cooperation” are supplemented by an **annual quality report for each register**, which included feedback on the micro-level for the first time.

Summary

- ESS Code of Practice was the underlying framework when developing the new system for quality reporting for registers.
- The indicators selected for the report can be grouped in the following five dimensions of quality: **Technical checks, Accuracy, Completeness, Integrability, and Time.**

Oman – E-Census

- Sources



Note: Reference provided as separate document

Oman – E-Census

In order for the administrative records to be used in the electronic census, it must pass through a large set of stages, which ensures accurate, comprehensive and timely data

- Refining and processing data that is missing some variables
- Linking and matching between variables by ID number
- Determining the time reference period
- Application of quality assurance standard
- Contacting record providers to ensure the accuracy of the data
- Knowing the reasons for non-responding
- Knowing the reasons for non-conformity
- Evaluate the quality of the variables
- Report inconsistencies in records
- Report missing metadata
- Update incoming records periodically

Oman – E-Census

Among the most important conditions that must be met to ensure data quality are the following:

- Available data must meet the needs of users.
- Ease of obtaining data in a timely manner.
- The data should be clear, high-accuracy and credible.
- The available data should enable international comparisons to be made.
- The possibility of the availability and use of metadata, with clarification of the used terms and definitions.
- Taking into account the provision of data at the lowest possible cost.

Other examples (random selection)

- Tools, metrics and quality reporting of administrative data sources (Poland) (Note: Reference provided as separate document)
 - Data profiling is the basic data quality check procedure.
 - Cohesion analysis of address data
 - The Variable Quality System (VQS) to monitor quality of source data
 - provides quality reports for all administrative data sources
- Quality assurance of administrative data (QAAD) report, e.g. for Value Added Tax turnover data
(available at [Quality assurance of administrative data \(QAAD\) report for Value Added Tax turnover data - Office for National Statistics \(ons.gov.uk\)](#))

THANK YOU

