



Ensuring Data Quality

What do we mean when we talk about data quality?

The degree of excellence or accuracy of the factual information being collected in a survey or census needed to make it meet the user's needs for decision making purposes.

“Fitness for use”

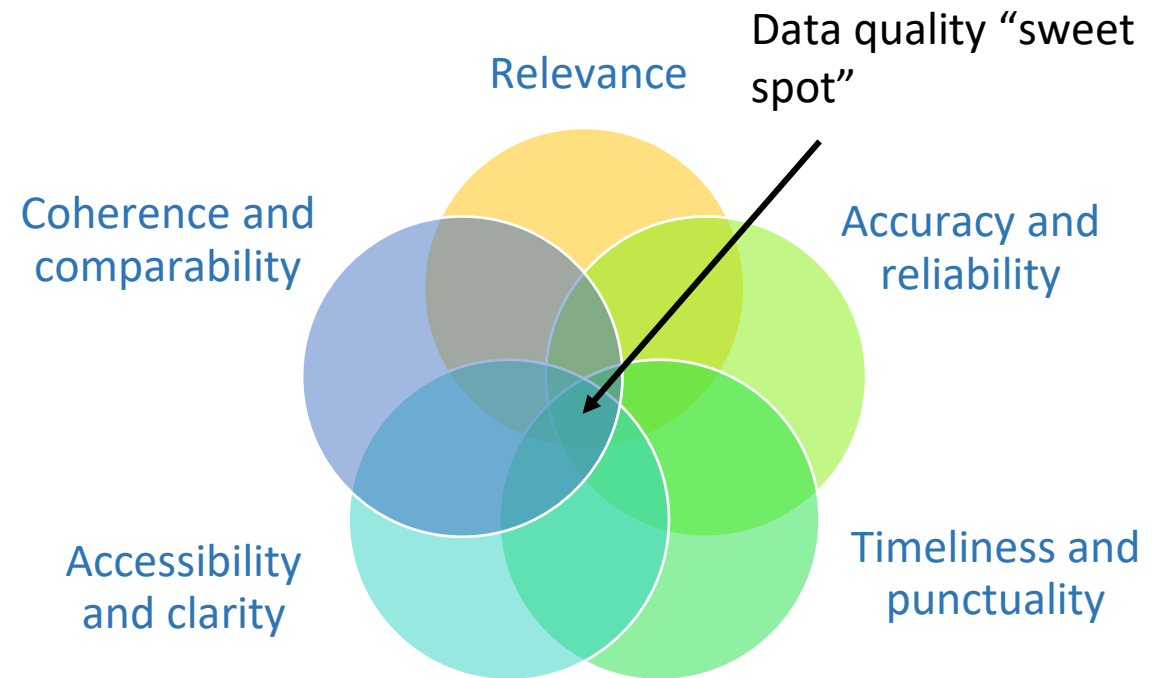


“Fitness for purpose”

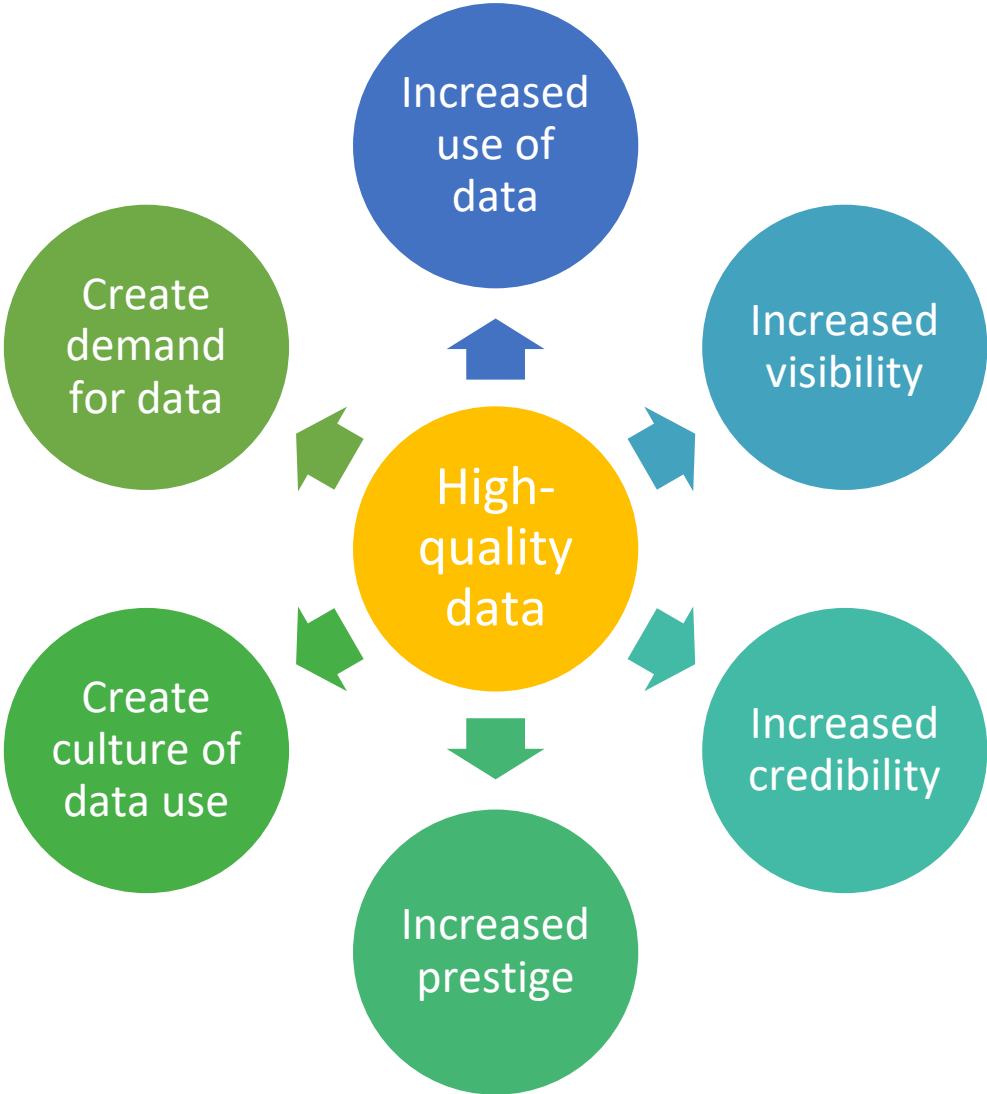
How Data Users See Quality

“Quality” is not determined solely by scientific accuracy.

It is the **data user**, your customer, who decides if your data products have “quality.”



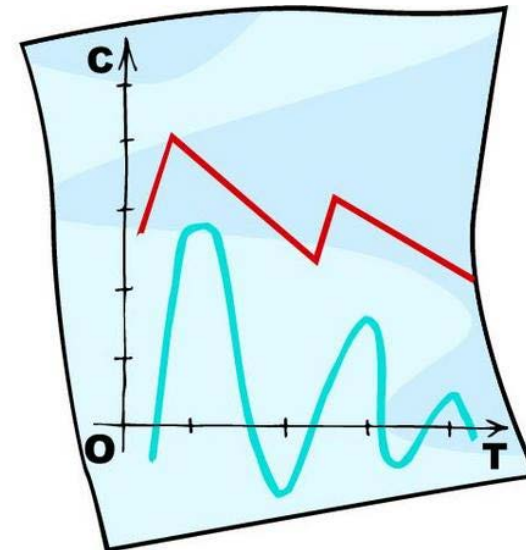
Benefits of High-Quality Data



Why does data quality matter?

The data you produce are used to:

- Make national policy and program decisions
- Track trends in the population
- Make local and national comparisons.

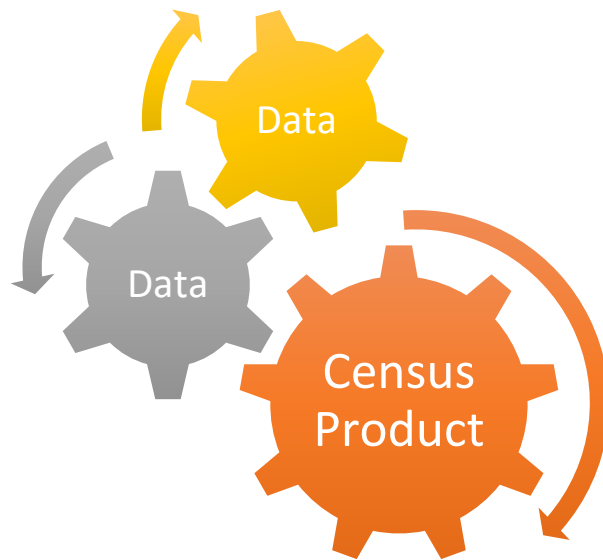


Data Quality Concepts

Quality Assurance

Process oriented

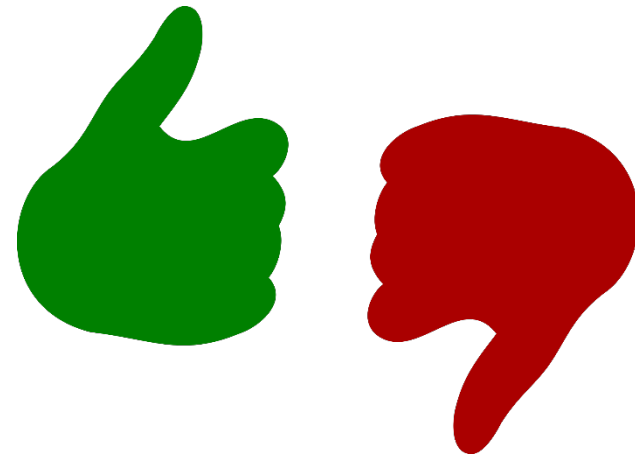
Focused on preventing errors



Quality Control

Product oriented

Focused on identifying errors
and accepting or rejecting
products



Measuring Data Quality

What should we do to measure data quality?

Need a **set of principles and standards** that allows us to effectively identify and implement ways to improve data quality

- At agency level
- At program planning stage
- At program implementation stage
- At post-production stage

Data Quality Assessment Frameworks

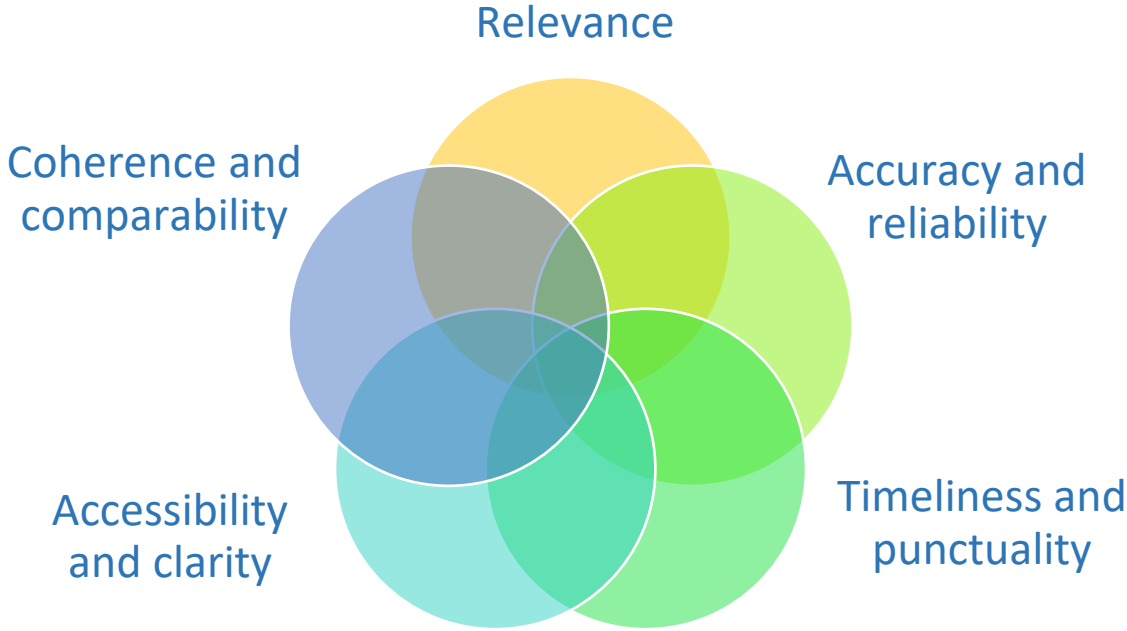
[DQAF and Data Quality Program \(imf.org\)](#)

[IMF -Introduction to the Data Quality Reference Site](#)

[UN National Quality Assurance Frameworks Manual for Official Statistics](#)

[European Statistical System QAF](#)

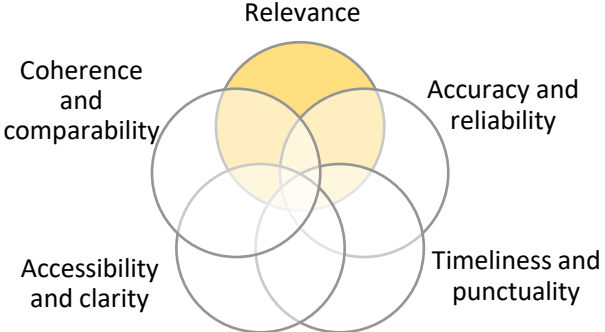
Dimensions of Data Quality





Relevance

Are you is producing information on the right topics and utilizing the appropriate concepts for measurement within these topics?



Questions to ask yourself are:

- Is the information we are producing relevant to topical policy issues or are we still measuring items that are no longer relevant?
- Does my organization utilize a definition of family that is pertinent to today’s society?
- Does my organization’s classification of occupations reflect the current labor market?

Assuring Relevance

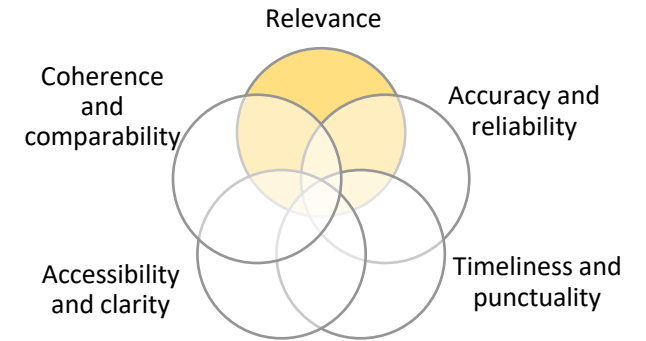
What are some ways to assure relevance in census dissemination?

- Have procedures to consult users about the content of the statistical work program.
- Implement user satisfaction surveys.
- Regularly track of uses of statistical products.
- Create mechanisms to identify users' needs and describe how the data relate to their needs.
- Ensure operational controls for editing and statistical quality consistent with the resultant statistics meeting users' needs.



Set up post-collection evaluations that:

- take stock of outcomes and compare with user needs,
- highlight any issues upon which users should be informed,
- provide users ways to give feedback that can be considered in the planning for future such collections.



Accuracy and Reliability

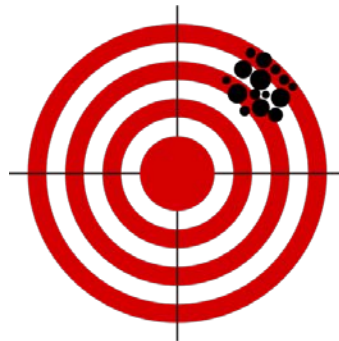
The **objective** of a survey or census is to obtain estimates of the true (unknown) value of a population or economic parameter.



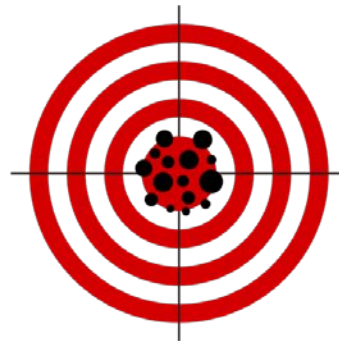
Unreliable & Invalid



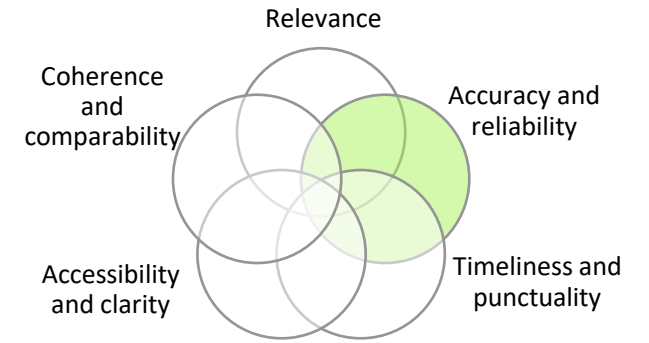
Unreliable, But Valid



Reliable, Not Valid



Both Reliable & Valid

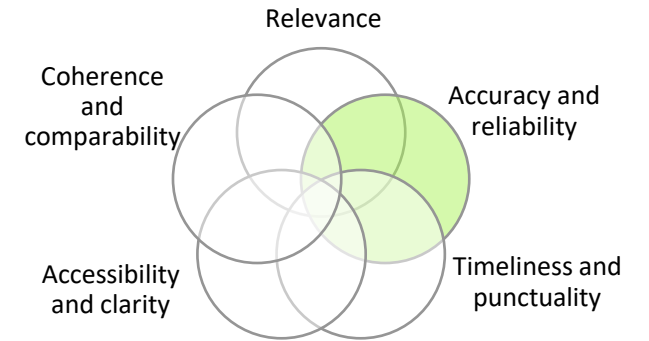


Accuracy (Valid) = hits the **true** value

Reliability = hits the **same** value

Assuring Accuracy and Reliability

What are some ways to assure accuracy and reliability in census dissemination?

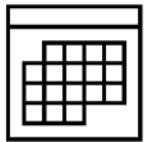


- Develop and manage systems for assessing and validating source data, intermediate results and statistical outputs.
- Measure, evaluate, and systematically document sampling and non-sampling errors.

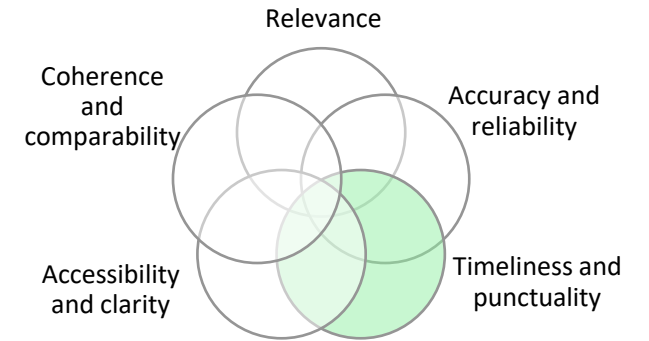
Timeliness and Punctuality



Timeliness refers to the length of time between data availability and the event it describes.



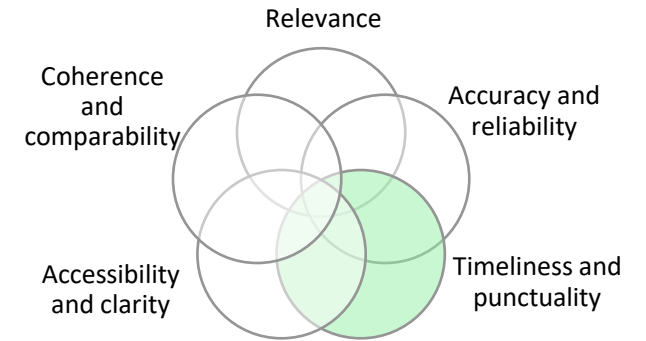
Punctuality refers to whether data are delivered on the dates promised.



Timeliness and punctuality are sometimes a trade off with accuracy.

Assuring Timeliness and Punctuality

What are some ways to assure timeliness and punctuality in census dissemination?

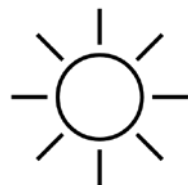


- Publish a release calendar to announce in advance the dates that statistics (at least the major ones) are to be released.
- Define **attainable** schedules for the production processes.

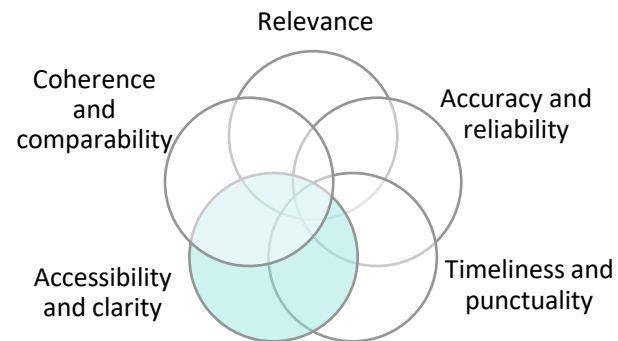
Accessibility and Clarity



Accessibility is the ease with which statistical information can be obtained from statistical offices.



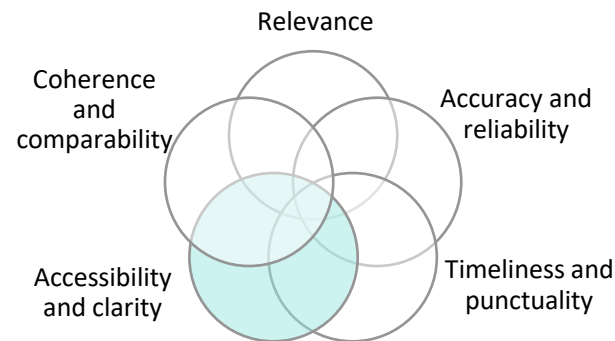
Clarity means that data are presented clearly in such a way that they can be understood.



Clients must be able to determine what is available, how they could obtain it, and what it means.

Assuring Accessibility and Clarity

What are some ways to assure accessibility and clarity in census dissemination?

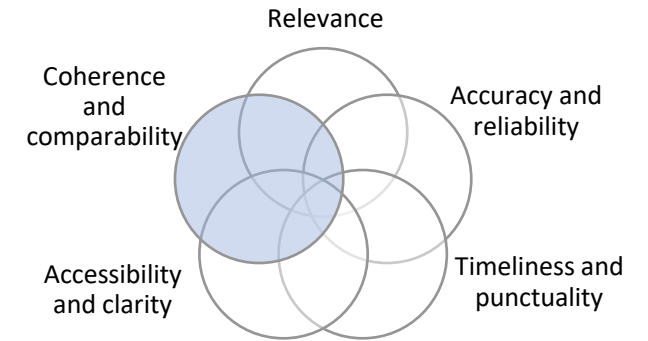


- Create policies and procedures which ensure that the statistical results are released with readily accessible and up-to-date documentation.
- Ensure that statistics and the corresponding metadata are presented and archived, in a form that facilitates proper interpretation and meaningful comparisons.
- Ensure statistics presented in a clear and understandable manner.

Coherence and Comparability



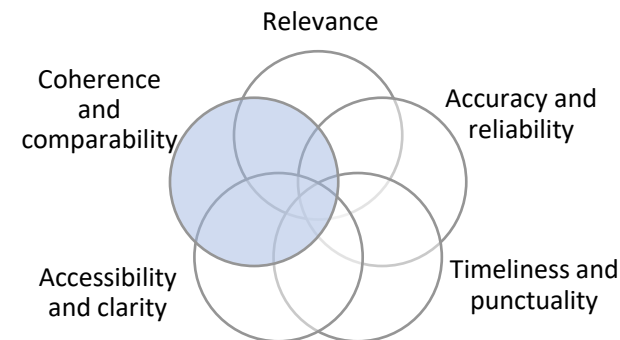
Coherence and comparability refer to the degree to which statistical information is internally **consistent** and **comparable** over time.

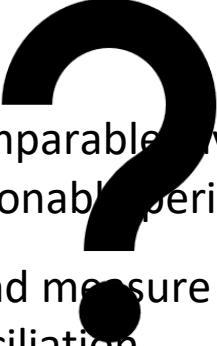


Coherence relates to sets of statistics and considers how well the statistics can be used together – usually through comparability of time and comparability over space.

Assuring Coherence and Comparability

What are some ways to assure coherence and comparability in census dissemination?



- 
- Keep statistics comparable (availability of time series) over a reasonable period
 - Clearly identify and measure changes in methods to facilitate reconciliation