

Version 1.3

Definition of Data Quality

Issued: 14 Jun 06

Census Bureau Principle

Authored for:

Census Bureau Methodology & Standards Council
Alan R. Tupek, Chair

USCENSUSBUREAU

Helping You Make Informed Decisions



Document Management & Control ¹

Version	Issue Date	Approval	Description
1.0	22 Nov 05	M&S Council	Initial Release
1.1	13 Dec 05	Configuration Mgr.	Re-Edit
1.2	05 Jan 06	Configuration Mgr.	Reformat
1.3	14 Jun 06	Associate Directors	Redefined transparency and clarified concepts. Initial concurrence of the Associate Directors.

¹ **The most current version of this document is maintained on the Census Bureau Intranet and may be accessed from the Quality Management Repository.**

Census Bureau Principle: Definition of Data Quality

The Census Bureau defines quality as “fitness for use.” We are guided by the needs of our customers to help determine that our data products are fit for their use. Our customers include all branches and levels of the federal government, state and local governments, and the public. When planning each data product we identify our primary customers and their uses for the data. We provide this definition of data quality to help our customers and employees understand the foundation of our data quality principles and standards.

To be fit for use, data products must possess all three attributes of quality defined by the Office of Management and Budget (OMB):

Utility - refers to the usefulness of the information for its intended users.

Objectivity - refers to whether information is accurate, reliable, and unbiased, and is presented in an accurate, clear, and unbiased manner.

Integrity - refers to the security or protection of information from unauthorized access or revision.

To help apply limited resources to best achieve fitness for use, the Census Bureau further defines utility, objectivity, and integrity in terms of six dimensions of data quality: relevance, accuracy, timeliness, accessibility, interpretability, and transparency. Thinking of data quality in these dimensions helps the Census Bureau to analyze the trade-offs between the sometimes conflicting dimensions, to meet our customers’ needs, and ultimately, to satisfy our customers.

Relevance refers to the degree to which our data products provide information that meets our customers’ needs.

Accuracy refers to the difference between an estimate of a parameter and its true value. We characterize the difference in terms of systematic (bias) and random (variance) errors.

Timeliness refers to the length of time between the reference period of the information and when we deliver the data product to our customers.

Accessibility refers to the ease with which customers can identify, obtain, and use the information in our data products.

Interpretability refers to the availability of documentation to aid customers in understanding and using our data products. This documentation typically includes: the underlying concepts; definitions; the methods used to collect, process, and analyze the data; and the limitations imposed by the methods used.

Transparency refers to providing documentation about the assumptions, methods, and limitations of a data product to allow qualified third parties to reproduce the information, unless prevented by confidentiality or other legal constraints.

In planning, developing, and delivering a data product, these dimensions of data quality may come into conflict. When this occurs, the Census Bureau analyzes all the dimensions to achieve a suitable balance among them. For example, timeliness and accuracy may present conflicting priorities. To achieve the best accuracy might require a longer data collection period to follow up nonrespondents and obtain interviews. But if obtaining near-perfect response would delay the data product until it was too late to inform the customer's decision, then a trade-off might be needed between accuracy and timeliness to achieve fitness for use. Similarly, unexpected circumstances in collecting or processing the data might require the Census Bureau to re-evaluate the dimensions of quality and redirect resources aimed at satisfying one dimension to focus on another, in the interest of achieving fitness for use.

Approved by the Census Bureau Methodology and Standards Council:

<u><i>Alan R. Tupek</i></u> Alan R. Tupek Chair, Methodology and Standards Council	<u><i>5/25/06</i></u> Date
<u><i>David C. Whitford</i></u> David C. Whitford Acting Chief, Decennial Statistical Studies Division	<u><i>5/25/06</i></u> Date
<u><i>Rita Petroni</i></u> Rita Petroni Chief, Office of Statistical Methods and Research for Economic Programs	<u><i>5/25/06</i></u> Date
<u><i>Ruth A. Killion</i></u> Ruth Ann Killion Senior Methodologist for Mathematical Statistician and Methodology Staffing and Development	<u><i>5/25/06</i></u> Date
<u><i>Tommy Wright</i></u> Tommy Wright Chief, Statistical Research Division	<u><i>5/25/06</i></u> Date

Concurrence:

Ted A. Johnson

Ted A. Johnson
Associate Director Administration and
Chief Financial Officer

6/14/06

Date

Nancy M. Gordon

Nancy M. Gordon
Associate Director for Strategic Planning
and Innovation

6/14/06

Date

Thomas L. Mesenborg

Thomas L. Mesenborg
Associate Director for Economic Programs

6/14/06

Date

Preston Jay Waite

Preston Jay Waite
Associate Director for Decennial Census

6/14/06

Date

Howard Hogan

Howard Hogan
Associate Director for Demographic Programs

6/14/06

Date