

Supporting Document A
Metadata That Must Accompany All Census Bureau Data
Releases

Version 1.1

Issued: 29 May 07

*Census Bureau Standard
Quality Requirements for Releasing Data Products*

Authored by:

Larry Cahoon (DSSD)
Gia Donnalley (DSMD)
Edison Gore (DMD)
Donna Kostanich (DSSD)
Ruth Runyan (CSD)
Ruth Detlefsen (SSSD)
Sharon Stern (HHES)

USCENSUSBUREAU

Helping You Make Informed Decisions



Document Management & Control*

| Version | Issue Date | Approval | Description |
|---------|------------|-------------------------------------|--|
| 1.0 | 22 Mar 07 | M&S Council and Associate Directors | Initial Release |
| 1.1 | 29 May 07 | Quality Program Manager | Correct requirement identification labels. |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

* The most current version of this document is maintained on the Census Bureau Intranet and may be accessed from the website for the Methodology & Standards Council.

Metadata That Must Accompany All Census Bureau Data Releases

Most data released by the Census Bureau are very high quality. Measures or indicators of data quality and metadata describing estimates should accompany all data releases. This supporting document describes the types of metadata that must accompany or be made available for all Census Bureau data releases.

- A. Descriptions of the procedures used to collect and process the data, and to produce estimates, must be made available in sufficient detail to allow competent data users to understand the procedures that generated the products and to assess their quality.

If providing this information would breach confidentiality protections, then it should not be provided, but the metadata explanation should state that the information is not being provided in order to protect confidentiality.

- B. Public use data files must include records representing all units selected for the sample (even those out of scope, if needed for variance calculations) and indicate all missing items within the data.
- C. Summary data on coverage loss, unit non-response, and item non-response must be made available.
- D. When imputation for item nonresponse is not carried out, the reasons for not imputing the data must be described in the metadata documentation.
- E. For tabulated data, an adjustment for nonresponse (both unit and item) must be included when appropriate, or the reason for not making such an adjustment must be described. When the adjustment is not made, all unadjusted tables should include cells that show the number of units not responding.
- F. A program may introduce changes in procedures or methodology that result in estimates that are not comparable with previous estimates. When data from such a program are used to produce time-series or change-estimate comparisons, metadata with the documentation describing the changes and the known impact on the data must be made available.

Two additional requirements apply to custom tabulations:

- G. Custom tabulations must include imputed and missing data to be released as category 1.
- H. Metadata, including coefficients of variation and sample sizes as available for the data sets used to create the tabulations, must accompany the custom tabulations. A reference or URL citing the metadata is not sufficient.

Occasionally the Census Bureau may identify a severe problem in data quality after having released a product as “Category 1” (meeting the Public Data Release Criteria).

- I. Within four weeks after identifying such a problem, the Census Bureau will make available documentation that:

identifies the problem;

describes what is known about the effect on the data;

discusses what is known at that time about its cause; and

describes plans for addressing the problem.

This documentation should be issued even if the problem is not yet fully understood, with more complete information provided when available.