

# Minding the store: An internal audit program for demographic programs at the U.S. Census Bureau<sup>1</sup>

Richard Levy\* and Jimmie Scott  
*U.S. Census Bureau, Washington, DC, USA*

**Abstract.** The U.S. Census Bureau collects and disseminates information products on people, businesses and governments. To ensure the information products are at an acceptable level of quality before release, the Census Bureau has developed and codified a set of statistical quality standards. And to ensure that Census Bureau programs are following the standards, the agency has developed several internal audit programs. This paper examines the audit program which covers demographic data, reviewing five years (2011–2015) of administrative data from the audit program. Analysis revealed four most common areas of noncompliance with the quality standards: Measuring and evaluating measures of nonsampling error; conducting statistical and policy reviews; developing timetables for reviewing files; and following policies for archiving information products. Analysis also reveals that full-scale surveys have both a greater level and share of noncompliance than other types of programs. Reviews of surveys, given their scope, also tend to take more time. While statistical programs vary across countries in organization and scope, the results of this research suggest that data programs may benefit from more organizational support on the back end of data production.

Keywords: Audits, quality standards

## 1. Background

The U.S. Census Bureau ([www.census.gov](http://www.census.gov)) is a leader in collecting data and disseminating information products on the American people, businesses and governments. These information products range from micro data sets and table packages to methodological papers and analytical reports. Information products are prepared for sponsoring agencies and for the general public. The bureau collects data for these products through surveys and censuses of households, businesses, and governments.

To ensure its information products meet an acceptable quality standard before dissemination, the Census Bureau has developed a set of statistical quality standards. These standards not only address the U.S. Office of Management and Budget's (OMB's) Stan-

dards and Guidelines for Statistical Surveys, but also go beyond the OMB requirements to “provide additional guidance that focuses on the Census Bureau's statistical programs and activities and that addresses the Census Bureau's unique methodological and operational issues” [1].

The standards are organized sequentially, covering each of the stages of the survey life cycle. Each standard contains multiple requirements towards meeting the standard's intent.

To ensure the various data programs are meeting the quality standards, the bureau has developed two internal audit programs. Other countries, such as Statistics Sweden, conduct quality audits on their statistical programs too [2].

This paper focuses on the results of the quality audits which cover demographic programs at the Census Bureau; that is, surveys, estimation programs, value-added programs, and inputs to these programs (process/participant programs). A *survey* refers to a data collection operation in which, among other tasks, a questionnaire is developed, a frame is drawn, a sam-

<sup>1</sup>This paper was originally presented at the European Conference on Quality in Official Statistics in Madrid, Spain, on June 2, 2016.

\*Corresponding author: Richard Levy, U.S. Census Bureau, Washington, DC, USA. E-mail: [richard.a.levy@census.gov](mailto:richard.a.levy@census.gov).

ple is selected, data are collected, and estimates and weights are produced, to gather information about subjects of interest. An *estimation program* takes data from surveys, censuses, and administrative records to produce estimates. A *value-added* product takes survey or other data and produces tables and reports based on these data. A *process/participant* program either provides some background research that affects various data operations, or refers to situations where the program area plays a supporting role to a data product being developed in another area.

Audits of demographic programs are conducted in accordance with Generally Accepted Government Auditing Standards ([www.gao.gov/yellowbook](http://www.gao.gov/yellowbook)) on a predetermined schedule coordinated among an audit steering committee, audit staff, and the various programs to be audited.

### 1.1. Audit stages

Audits are conducted in five stages:

- In the first stage (preparation), data program and audit staff meet to discuss logistics, budget and audit scope. That is, which of the quality standards apply to the program to be audited? The first stage also includes training for auditors and auditees.
- In the second stage (self-assessment), the data program staff completes an audit checklist. The checklist includes instructions, administrative data, and 263 ratable items which correspond to each of the various requirements. The data program staff rates its compliance (compliant, not applicable, or not compliant) with each of the requirements in the standards in scope for the audit, and provides evidence to justify their self-rating.
- In the third stage (independent review), a team of three or four auditors reviews the evidence in the checklist provided by the data program staff, and makes independent determinations whether each applicable item on the checklist complies with the corresponding requirement. Auditors not only determine compliance and applicability, but also make recommendations and identify best practices. Auditors present findings to the data program staff and senior management at a formal finding meeting.
- In the fourth stage (program action plan), the data program staff drafts an action plan to address any noncompliant items identified during the audit.

- In the fifth stage (follow-up), the audit staff work with the data program staff over the course of a year to ensure completion of the program action plan.
- The first audit of demographic data programs was conducted in 2011. As of December 2015, thirty-two quality audits have been conducted through the independent review stage. The audit program tracks the results of each audit (compliant, not applicable, and noncompliant findings) and administrative data around completing each audit (number of staff, hours spent by the program and the auditors in completing each audit). The data are maintained and made available for customers and stakeholders as well as used for the analysis below.

## 2. Review of audit findings

Overall, a review of audit results reveals that demographic data programs demonstrate high rates of compliance with the quality standards. To date, data programs have averaged a 94% compliance rate. (The compliance rate is the ratio of the compliant items over both compliant and noncompliant items.) The first year of the audit program, the compliance rate was 90%. Since then, the annual average compliance rate has hovered between 93 and 96%. This higher level of compliance in more recent audits most likely reflects greater familiarity with the quality standards and with the audit process.

### 2.1. Standards with complete compliance

Analysis of the audits indicates that data programs are most likely to be compliant with the following standards, i.e., no audits were found to be noncompliant with any of the requirements in these standards:

- Developing and implementing a sample design
- Establishing and implementing data collection methods
- Acquiring and using administrative records
- Capturing data
- Linking data records
- Producing direct estimates from models
- Addressing information quality guideline complaints

The complete compliance, in some cases, reflects that not many programs carry out procedures, in the course of their data operations, to which the standards would apply. For example, only four programs linked data records from different sources as part of their data collection and dissemination. If a particular standard is determined out of scope for an audit, an area would be implicitly compliant with the requirements in that standard.

Other standards, such as establishing and implementing data collection methods or developing and implementing a sample design, were in scope for quite a few audits. As a whole, these standards tend to cover activities carried out in the early and middle phases of data program operations.

## 2.2. Noncompliant findings

Programs averaged five noncompliant items per audit, ranging from no noncompliant items to twenty-five (the first audit conducted, during the pilot phase). The average varied by type of program. Surveys averaged 13 noncompliant items per audit, the most of any type of program. Surveys averaged not only the greatest number of noncompliant items but also had the highest share of noncompliant items, as indicated in the table below. At the same time, surveys averaged more standards in scope, as more activities are associated with conducting a survey than with other data programs. (See Table 1).

Except in the standards noted in Section 2.1 above, programs were not compliant with specific requirements across various standards. Yet noncompliant findings emerged primarily in four standards: Managing Data and Documents; Producing Measures and Indicators of Nonsampling Error; Reviewing Information Products; and Releasing Information Products. Within each standard, certain requirements were identified repeatedly as not being met. Table 2 lists these requirements.

The requirements for releasing information products and producing measures of nonsampling error are related, as several of the restrictions for releasing information products are based on nonsampling error measures. In several audits, auditees had not calculated the required measures of nonsampling error, and as such could not determine whether there were quality issues severe enough to place restrictions on data release.

## 2.3. Staff time to conduct audits

Data programs and audit staff participate in the audits in addition to their usual responsibilities. Auditors are identified and selected across programs to participate based on their experience, independence, and availability. The audit program attempts to identify staff early in the process to accommodate current responsibilities and the audit requirements. To help program managers plan and manage responsibilities including the audits, the audit staff provide information such as estimated time required based on previous audits.

Below are some of the resource findings:

- On average, program staff required 186 hours to complete their self-assessments and answer auditor questions during the independent review. One early audit required nearly fourteen hundred hours to complete their checklist, exceeding the next highest figure by over nine hundred hours. Removing this outlier, the average drops to 147 hours per audit.
- Over time, the program staff generally required fewer hours to complete their checklists. This most current year, data program staff averaged only 118 hours to complete their audit work.
- Surveys required on average the most hours (210), followed by process/participant programs (157), value-added programs (90) and estimation programs (87, excluding the 1379 hour outlier).
- On average, auditors took 110 hours of staff time to complete their independent reviews. Surveys averaged the most hours (153), followed by estimation programs (99), process/participant programs (87), and value-added programs (75). The hours the auditors required to complete the audit was not strongly associated with the number of findings ( $R = 0.57$ ). Some audits take longer than others because they may present issues which take more time to research, whether such research results in noncompliant findings. Also, some auditors are either more efficient or more thorough in their independent review.

## 3. Discussion and recommendations

One key point from the data analysis is that noncompliant findings in the demographic programs cluster around a few standards, and that these unmet requirements concern activities that take place near the end of

Table 1  
Type of data program by number and share of noncompliant items

Type of program	Average number of noncompliant items	Average rate of noncompliance	Average number of standards originally determined in scope
Surveys (N = 11)	13	10%	14
Estimation Program (N = 8)	1	2%	11
Process/Participant (N = 7)	1	2%	9
Value-Added (N = 6)	2	4%	9

Source: Analysis of 2011–2015 data from Demographic Quality Audit Program, U.S. Census Bureau.

Table 2  
Most commonly not compliant requirements, by standard

Standard	Requirements Out of Compliance
Managing Data and Documents	Knowing, documenting, and adhering to data and document retention schedules Creating and adhering to timetables to review files for usefulness and readability
Producing Measures and Indicators of Nonsampling Error	Developing a plan for measures and indicators of nonsampling error to be produced Developing a plan to evaluate measures and indicators of nonsampling error Using standard formulae for nonsampling error measures Defining sufficient partial interviews Verifying and testing the calculation of measures of nonsampling error Evaluating measures of nonsampling error
Reviewing Information Products	Supervisory review Policy and sensitivity review Documentation of review
Releasing Information Products	Releasing information products without restrictions only when of sufficient quality Releasing information products with restrictions when data quality issues are known

Source: Analysis of 2011–2015 data from Demographic Quality Audit Program, U.S. Census Bureau.

a data program life cycle: Measuring nonsampling error, reviewing products, releasing products, and managing data and documents. This is not surprising, as it is easier to start projects than to complete them. (One is reminded of the chapter in *The Dilbert Principle* [3] on projects: Under the section for “Completing the Project,” Scott Adams deliberately leaves the section blank.) As such, this research suggests that data programs should anticipate and dedicate more resources to completing the final stages of their projects.

In following programs with noncompliant findings after their audits, it became apparent that program areas were not complying not due to lack of expertise, but for lack of knowledge of the standards. Even with the most common audit finding – not knowing or adhering to records schedules for pertinent findings – twelve out of fourteen programs that failed this item during their audits have since identified or established new records schedules.

Such experience suggests that data programs require more education around the quality standards. The audit process itself serves as implicit education for the auditees and for the auditors about the quality standards. The demographic audit staff has also developed, and will continue to develop, training for Census Bureau staff on the quality standards.

In addition, it is incumbent on the data programs that produce data for customers – most often, other

federal agencies – to educate their customers on the quality standards, and how following these standards will improve the quality and utility of the data products they are receiving. This will help ensure that the sponsor and Census Bureau requirements are better aligned. Between the quality audits and training on the standards, the Census Bureau will help ensure all its data products meet an acceptable quality standard.

### Acknowledgments

The authors would like to thank Cynthia Rothhaas, audit staff, for her statistical review and Jim Treat, Division Chief, Demographic Statistical Methods Division for his support and guidance in writing this report. The authors would also like to thank Jill Harbison, former Lead Auditor of the Demographic Quality Audit Program, for her work in starting the audit program and the compilation of much of the administrative data used in writing this report.

### Conflict of interest

Any views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

## References

- [1] U.S. Census Bureau. Statistical Quality Standards. 2013. p. ii. Available from: <http://www.census.gov/about/policies/quality/standards.html>.
- [2] G. Eiderbrandt-Nilsson, Quality Audits at Statistics Sweden. Presented at the Joint Statistical Meetings of the American Statistical Association; 2005. Available from: <http://www.amstat.org/sections/srms/proceedings/y2005/Files/JSM2005-000474.pdf>.
- [3] S. Adams, *The Dilbert Principle: A Cubicle's-Eye View of Bosses, Meetings, Management Fads & Other Workplace Afflictions*. New York: Harper Collins Publishers; 1996. p. 236.