

Summer 2004

Developing a Library School Course in Government Statistics

John P. Heintz

University Of St. Thomas - Minnesota, jpheintz@stthomas.edu

Follow this and additional works at: http://ir.stthomas.edu/lib_staffpub



Part of the [Higher Education Commons](#), [Information Literacy Commons](#), and the [Social Statistics Commons](#)

Recommended Citation

Heintz, John P., "Developing a Library School Course in Government Statistics," *Journal of Education for Library & Information Science*, vol. 45, no. 3 (Summer 2004), pp. 189-199.

This Article is brought to you for free and open access by the Libraries at UST Research Online. It has been accepted for inclusion in Library Staff Publications & Presentations by an authorized administrator of UST Research Online. For more information, please contact libroadmin@stthomas.edu.

Developing a Library School Course in Government Statistics

John P. Heintz

Statistical information is one of the most important forms of government publication. Librarians receive some training in how to locate and use government statistical information in reference or government documents classes, but the coverage is usually cursory and most librarians could use more education. This paper reviews the literature and available print and electronic statistical resources and outlines a course for librarians in government statistics, suggesting classroom or self-study instructional methods; identifying key information on search strategies, ready reference sources, and detailed topical statistical sources; and suggesting issues for librarians to be aware of in helping their patrons locate and use statistics.

Introduction

Knowledge of the available government publications resources is critical for reference librarians. Many librarians believe that students under use government documents; faculty members contacted in one survey believe that government information is at least "somewhat important" (68 percent) and that their students "rarely or never" (53 percent) use government publications in papers.¹ This mismatch between faculty expectation and student performance suggests a role for reference librarians in helping students identify and use government documents.

Past treatments of the preparation of librarians suggest that most library and information science (LIS) schools only provide one elective course in government publications, that the subject matter is also discussed in a cursory fashion in other reference or collection development courses, and that some writers believe more coursework is necessary.² One survey of government documents librarians suggests that even those who took government publications courses in graduate school often did not feel that their preparation was very good (47 percent responded "satisfactory" or "poor" when asked about the quality of their preparation). In addition, 82 percent felt that library schools should offer more training, which should emphasize federal and state government information and online sources.³

According to Morehead, "the dissemination of statistical data, like legal information, has always been a crucial component of the information policies of governments."⁴ Assuming this is true, one would think that library school government publications courses would cover statistical informa-

About the Author

John P. Heintz is Reference Librarian, Charles J. Keffer Library, University of St. Thomas, Minneapolis, Minnesota. Ms. Received 08/02; accepted 11/02; revised 02/03.

tion in some depth. However, because of the wide range of government publications available, these courses can typically only give cursory treatment to statistical sources. For certain types of librarians (like government documents and academic social science librarians), this will not adequately familiarize them with the search strategies and range of resources available in government statistics. While certainly these librarians will learn on the job, it would be desirable for them to have a formal course available to provide a more in-depth treatment of the subject. Such a course could take the form of a special topics, seminar, or independent study course in a library school, or could be offered as a formal course or online self-study module sponsored by a library association or consortium. This article will outline such a course by: suggesting classroom or self-study instructional methods; identifying key information on search strategies, print and online ready-reference sources, and more detailed topical federal government statistical sources available that such a course should convey to librarians; and suggesting issues for librarians to be aware of while helping their patrons locate and use statistics.

Methods

Traditional reference instruction methods should be utilized to teach this course for librarians if offered in a formal classroom setting.⁵

Lecture

Topics should include overviews of search strategies, issues in finding and using statistics, and information on locating and using the various finding aids and sources. These lectures should be supplemented by guest appearances by reference librarians discussing the types of statistical questions typically posed by patrons and how to address them. Additionally, the instructor could ask professors from various disciplines what types of statistical information they want students to become familiar with and use in their research.

Searches

Library students should be given topical "scavenger hunt" type question and answer assignments requiring them to track down sources in print and online to answer questions posed by the instructor—these would then be discussed in class to reinforce the search strategies and use of the sources, and to share information and experiences on the alternative routes to find-

ing information. These assignments could be done either individually or collaboratively, depending on instructor preferences.

Presentations

Class participants should be required to research particular sources or topics and prepare pathfinders for formal presentation to the class.

Paper

Library students would prepare a formal research paper on an appropriate topic in government statistics.

Tutorials

In a less formal setting, or as a component of a formal course, a web-based curriculum could be developed to cover some of the lecture format and scavenger hunt materials. A number of library-based statistics tutorials exist on the Web. For example, Cruse and DeDecker have developed a web tutorial entitled "How to Effectively Locate Federal Government Information on the World Wide Web."⁶ This useful site (while it could stand some updating of sources) covers topics such as federal government structure, search engines and indexes, legal information, and a section on "statistics, demographics, and other useful data." Each section includes a brief discussion, list of key sources (with hyperlinks), and ends with a series of scavenger hunt questions (with answers provided).

Another useful statistics tutorial site is sponsored by the Milner Library at Illinois State University.⁷ This site provides basic information like: general definitions of statistical terms; descriptions of various types of statistics; how to evaluate authenticity, reliability, and relevance of statistical data; how to locate and retrieve uncompiled data for processing; and search strategies and lists of sources.

Tutorial sites like these should be used as part of the formal instruction for a course, both for their substantive content and as examples of resources that librarians can develop on the web to better serve patrons.

Text

A primary text for the course could be Sears and Moody's *Using Government Information Sources: Electronic and Print*.⁸ This standard reference work contains well over 200 pages specifically targeted to statistics search strategies, finding aids, and explanations of key sources. Readings in the text would be supplemented by appropriate journal articles (some suggested in this article), chapters from other texts, and current newspaper or magazine coverage of statistical issues or current events that can be illuminated by the use of government statistics.

Course Content

The course should be organized to cover five major areas of inquiry:

- The Federal statistics-gathering environment
- Search strategies and finding aids
- Census information
- Topical areas and sources
- Issues for librarians to consider in using statistics

The Federal Statistics-gathering Environment

This topic would provide an institutional overview of which federal agencies collect statistics and why. The collection and dissemination of statistical information in the federal government is a massive and complex undertaking. The president's fiscal year 2002 budget request for major statistical programs in the federal government exceeded \$4.1 billion, employing more than 13,361 federal workers.⁹ There are fifteen "principal" statistical agencies listed on the *Fedstats* website, and a total of one hundred, comprising cabinet departments and bureaus within them, independent agencies, the various national institutes of health, and interagency collaborative efforts.¹⁰ It is important to present an overview of this agency structure, explaining which agency produces which statistics and for what purpose. This information is important both as general context and to inform an "agency" based search strategy, which is one method used to find statistical data. This section should also cover how statistics are used for federal program planning, evaluation, and management; setting grant benefit levels and distribution formulas; legislative oversight; and other issues. Herman's *Locating United States Government Information* has a chapter discussing these and related issues.¹¹ This presentation would also explain the roles played by the Office of Management and Budget (OMB) and the Federal Interagency Council on Statistical Policy. OMB, as part of its regulatory function, issues technical standards defining key statistical terms and collection methods and arbitrates disputes regarding such topics.¹² The Council provides a forum for discussing technical issues among agencies, and produces the *Fedstats* website. Other online databases for statistical information include RDS/Gale's TableBase and the Federal Commerce Department's STAT-USA.

Other key sources for this topic would include: 1) The *Fedstats* website, which provides: a reference shelf of standard statistical compendia; topical, agency, and agency-subject indexes for locating statistics; descriptions of the statistics available; statistics by geography and mapping tools; and information on federal statistical policy. 2) *Statistical Programs of the United States Government*, a congressionally mandated annual report from OMB on funding and policy issues surrounding federal statistics.

Search Strategies and Finding Aids

This section of the course would rely heavily on Sears and Moody's descriptions of statistical search strategies.¹³ They identify a general five-step statistical search strategy: Ready Reference, Statistical Compilations, Census (treated separately below), Indexes/Databases, and Web Search

Engines. The first two of these (Ready Reference and Statistical Compilations) involve checking for needed information in the *Statistical Abstract of the United States* or other standard compilations like the *City and County Databook* and *CIA World Factbook*. These are standard approaches taught in all basic reference courses. However, an additional important source to be covered is the Census Bureau's *Historical Statistics of the United States*, a two-volume set that collects statistics from a myriad of sources and provides annotated notes to additional sources that readers can find in other works.¹⁴ The course should provide a quick review of these ready reference tools, with a reminder to students that they include references to more-detailed original publications concerning a given topic, and as a result can be used as an effective index. Supplementing Sears and Moody, Herman's work on this topic has a separate chapter on ready reference statistical sources that can be consulted for a more detailed treatment.¹⁵

There are a handful of important indexes/databases for finding statistical information with which librarians should become familiar. According to Sears and Moody, "Consulting indexes and databases is an important part of the statistical search strategy and is usually carried out when searches in the known sources . . . have been exhausted."¹⁶ The first of these indexes is Congressional Information Service (CIS) *American Statistical Index*, available in print volumes or online as part of Lexis-Nexis's *Statistical Universe*.¹⁷ While this tool is not available in all libraries, librarians should become familiar with the searching capabilities and techniques of *Statistical Universe* as part of this course. *Statistical Universe* is divided into three major sections: Power Tables, Abstracts, and Links. "Power Tables" allows a searcher to search for specific kinds of statistical tables. A useful search interface allows users to enter keyword terms such as "health insurance coverage" or "population and Minnesota" (or to consult a subject index for terms), and then narrow a search based on geographic, demographic, economic, and frequency criteria. Finally, searchers can select periods of coverage, limit to documents with Excel spreadsheet data, and find forecast or projection data. "Abstracts" features the same search interface, but instead of tables the search returns abstracts and accession information allowing searchers to determine whether or not to obtain a particular publication or view it online. "Links" is simply a rather short list of government statistics agencies that is done better on the *Fedstats* website.

A second index for students to become familiar with is the Gale Group's *Statistics Sources*.¹⁸ This two-volume work contains detailed annotated references to thousands of statistical sources, both governmental and private. The main body of the text is organized as a simple alphabetical subject index. The front matter contains a "Selected Bibliography of Key Statistical Sources," including a summary of Census Bureau and other agency publications and databases. Finally, Sears and Moody itself is an excellent finding aid that should be in all reference collections where patrons often have occasion to seek government documents and statistics.

Sears and Moody assert that "providing access to electronic publications

and information through the World Wide Web is now one of the primary ways that government agencies distribute information. The use of web search engines has become essential in accessing government information."¹⁹ Obviously a course in federal statistics will require a thorough grounding in how to find them on the Web, both as part of general search strategy discussions and in studying specific topical areas. The *FirstGov* portal and search engine is a key web search tool in seeking federal statistical information. Librarians taking the course will also need to become thoroughly familiar with the *Fedstats* site discussed previously, focused search engines like Google Uncle Sam, and some important web directories that contain significant links to statistical resources. Many library websites around the country provide very well done lists of links to statistical sources, which can be a quick way to access the information. Among several, the most important of these is probably the University of Michigan Library Government Document Center's "Statistical Resources on the Web" site.²⁰ This comprehensive site is organized into 24 topical areas, with a much larger subject index of links to almost any conceivable topic. The site also links to the Inter-University Consortium for Political and Social Research (ICPSR), which provides access to large volumes of raw statistical microdata for affiliated member institutions. Another good statistics directory is the University of North Texas' "United States Statistical Agencies and Information," which organizes links to general statistical compilations, statistics by topic, a directory of federal statistical agencies, and a list of statistical indexes.²¹

Finally, any discussion of statistical search strategies must address terminologies to be used in searching. Statistics and demographics have unique terms to describe types of statistics and analyses, and these must be examined. In addition, the complex structure of the agencies involved and the different reasons for which statistics are collected lead inevitably to differences in terminology used to describe data and to confusion about relationships among data series collected by different agencies. Knowledge of the agencies involved, the data collected, the relevant vocabulary and search terms to use, and how to connect information from different sources can enable librarians to help their patrons locate needed information.²² These issues need to be discussed, with examples presented of the various terms used and how researchers and librarians can go about learning the terminology as it exists and as it evolves. There are a number of researchers that examine these topics and study the information-seeking practices of statistics users in an attempt to improve data collection practices, search engines, and website design to improve the accessibility of information. One example of such research is the GovStat Project, whose website (<http://ils.unc.govstat/>) provides summaries and links to some of the research documents, which might be assigned as class readings or research leads.

Census Information

Sears and Moody consider consulting Census publications to be a com-

pletely distinct search strategy, and also include Census documents in their list of sources within individual topical areas. A government statistics course for librarians should have an in-depth review of the Census, including at least two class sessions specifically devoted to it. Possible topics to cover would include:

- the decennial Census of population and housing: constitutional basis, purpose, history
- types of data collected in the Census: 100-percent and sample, review of questionnaires, etc.
- review of data products: print, online publications
- federal government uses of census data: reapportionment, calibration of program formulas and benefit levels, etc.
- other uses by state/local government, businesses, citizens; typical types of reference queries
- access to and uses of microdata
- other population-based surveys: Current Population Survey/American Community Survey, Survey of Program Participation; purpose, frequency, types of data collected
- Current Population Reports series
- Using "American Factfinder," TIGER, FERRET, other data extraction and mapping tools
- Other Census Bureau products: Census of Governments; Economic Censuses; National Survey of Fishing, Hunting, and Wildlife-Associated Recreation; foreign trade statistics²³

In addition to becoming familiar with the sources cited above, class participants would be introduced to Anderson's *Encyclopedia of the U.S. Census* and Lavin's *Understanding the Census: A Guide for Marketers, Planners, Grant Writers, and Other Data Users*.²⁴ Additionally, the instructor should assign current readings from the library literature covering new tools for locating and using the 2000 Census information as these articles and books are published.

Topical Areas and Sources

This section will provide a brief outline of the statistics subject areas the course would cover. The topical coverage was determined by reviewing the subject heading breakdowns in several sources, including Sears and Moody, Morehead, Wasserman's *Statistics Sources*, and the University of Michigan library website. Topic areas to be covered include:

- Population (much coverage in Census discussions, also include immigration)
- Employment and Labor
- Economics (including housing, income, earnings, prices, consumer spending, government finances)

- Business and Industry (including foreign trade, communications)
- Education
- Energy
- Environment and Natural Resources (including weather/climate)
- Health and Vital Statistics
- Crime and Criminal Justice
- Politics
- National Defense and Military
- Transportation

While there is a reasonably good fit between these topics and those highlighted in Sears and Moody, it will be necessary to supplement that text with other materials. For each topical area, the course should review the types of statistics available; which agencies produce them and why; key print, database and Internet sources; explain typical reference queries in that area; and if possible, illustrate how such statistics have been used recently in media treatments of current events and issues.

Issues for Librarians to Consider in Using Statistics

This topic can be divided into two areas: 1) what reference librarians should know about statistical reasoning and usage, and 2) issues surrounding data accuracy, construction, and interpretation. While most of this course would focus on understanding types of statistical queries and how to access sources that can answer them, this section of the course should help provide additional context for helping patrons understand how to use information from such sources appropriately. It would be beyond the scope of this course to give library students a thorough grounding in statistical methods; however, some basic information should be conveyed, if only to give these current or future librarians enough background to warn patrons when statistical information could be invalid, misused, or misunderstood, and to know when to seek additional help or information in interpreting it.

First, for basic information, course participants should be provided with references to key statistics texts that may be in use in their institutions. Often such texts are used in the "methods" courses taught by various disciplines, and it may help librarians better serve students in finding and using statistical information if they are familiar with the texts used and the types of statistical methods being taught in the curriculum. In addition, there are many online statistics texts available on the Internet for quick reference to which librarians should be introduced. A representative sample is the *StatSoft Electronic Statistics Textbook*.²⁵

Second, the course should include a chapter or two and articles on statistical methods and common mistakes. Such material should cover topics like: averages; percentages and percent change calculations; use of charts, graphs, and tables; correlation; confidence intervals and analysis of variance; and basic regression analysis. Rather than trying to give a remedial

statistics course, the emphasis would be on illustrating how such techniques are used in the statistical sources themselves, how researchers use such techniques to manipulate the data they obtain, and common mistakes in the application of the techniques. Excellent sources for this discussion include a chapter on statistical reasoning in Lavin's *Business Information: How to Find It, How to Use It*,²⁶ and two recent articles by Gerhan entitled "When Quantitative Analysis Lies Behind a Reference Question" and "Statistical Significance: How It Signifies in Statistical Reference."²⁷ If time permits, the discussion could be extended to cover how data can be extracted and manipulated in software tools such as spreadsheet and database programs and statistical software packages like SAS or SPSS.

Another thread in this discussion is the issues surrounding data accuracy, construction, and interpretation. Library literature treatment of the authority of reference sources typically views government-supplied information as authoritative, but is it necessarily so? Government statistical information is subject to sampling, measurement, and estimation errors, non-response bias, and other problems of survey construction. Certain types of publications may contain an administration or partisan bias. Government data series are also frequently revised after initial publication. Similar types of data presented by different agencies may be collected using different methods and yield conflicting and confusing results, and the same data series may change methods between collections, resulting in comparability problems over time. The treatment of race and ethnicity in government statistics programs is fraught with controversy and can lead to problems of accuracy, comparability, and proper interpretation. Helping librarians understand and grapple with these issues provides important context to statistical reference. Several sources previously cited touch on these issues; other references are Stratford's "Responding to Reference Queries for Numeric Data and The Problems Inherent in Interpreting Statistical Sources: A Note" and Robbin's "The Problematic Status of U.S. Statistics on Race and Ethnicity: An 'Imperfect Representation of Reality.'"²⁸

Conclusion

Government statistics is an area where most reference librarians could expand their knowledge. Certainly academic, business, and government documents librarians have an expanded need for knowledge of statistical data sources and uses. This paper has attempted to outline an appropriate curriculum for delivering some of that knowledge. While this author makes no claim to complete coverage of the appropriate topics, the material described here is a good representation of the basics useful as preparation for prospective librarians or as continuing education for practicing ones. What remains is for library school professors or practicing librarians to agree with the need to deliver such a curriculum and to do so in formal classes, or for a professional association or library consortium to develop appropriate web-based materials or continuing education classes to address it.

The author would like to thank Professor Helen Humeston and Professor Kate Marek for their suggestions, support, and encouragement in the writing of this article.

References and Notes

1. Nicholas G. Tomaiuolo, Richard Churchill, and Joan Garrett Packer, "Access to Government Publications and Library Instruction: Views of Faculty and Librarians," *College & Undergraduate Libraries* 5, no. 1 (1998): 128-30.
2. Herbert S. White, "Library Education and Government Documents Work," *Government Publications Review* 13 (1986): 135-45.
3. Barbara M. Cross and John Richardson, Jr., "The Educational Preparation of Government Information Specialists," *Journal of Education for Library and Information Science* 30, no. 1 (Summer 1989): 28-38.
4. Joe Morehead, "Reflections on the 'Right To Be Informed,'" *Government Publications Review* 13 (1986): 39-47.
5. There is little current literature addressing how to teach government publications: the White and Cross articles cited previously are representative, and they do not address instructional methods or statistics specifically. The methods described here are largely based on experiences in library school reference classes, and Mary Luebbe, "Teaching Government Publications: A Reference Librarian's Perspective," *Government Information in Canada/Information gouvernementale au Canada* 4, no. 1 (1997). Available at: <http://www.usask.ca/library/gic/v4n1/luebbe/luebbe.html>. Accessed: February 10, 2003.
6. Patricia Cruse and Sherry DeDecker, "How to Effectively Locate Federal Government Information on the World Wide Web." Available at: <http://gort.ucsd.edu/pcruse/universe/intro.html>. Accessed: February 10, 2003.
7. Lisa Janicke Hinchliffe and Lynda Duke, "Finding Statistics—Tutorial and Worksheet," Milner Library, Illinois State University. Available at: <http://www.mlb.ilstu.edu/learn/stat/home.htm>. Accessed: February 10, 2003.
8. Jean L. Sears and Marilyn M. Moody, *Using Government Information Sources: Electronic and Print* (Phoenix: Oryx Press, 2001).
9. Executive Office of the President, Office of Management and Budget, *Statistical Programs of the United States Government, Fiscal Year 2002* (Washington: OMB, 2001). Available at: <http://www.whitehouse.gov/omb/inforeg/statpol.html>. Accessed: February 10, 2003.
10. Fedstats home page. Available at: <http://www.fedstats.gov>. Accessed: February 10, 2003.
11. Edward Herman, "Statistical Sources: Part I—Introduction 'Lies, Damned Lies, and Statistics,'" in *Locating United States Government Information: A Guide to Sources* (Buffalo: William S. Hein, 1997).
12. See the OMB "Statistical Policy" site listed above for a hot-linked list of such bulletins and circulars. Available at: <http://www.whitehouse.gov/omb/inforeg/statpol.html>.
13. Sears and Moody, *Using Government Information Sources*, 8-9.
14. Donna L. Burton, "Historical Statistics of the United States: Colonial Times to 1970," (Review) *Government Information Quarterly* 8, no. 1 (1991): 121-22.
15. Edward Herman, "Ready Reference Statistical Sources," in *Locating United States Government Information: A Guide to Sources* (Buffalo: William S. Hein, 1997).
16. Sears and Moody, *Using Government Information Sources*, 9.
17. Lexis-Nexis *Statistical Universe* (online database) (New York: Reed Elsevier, Inc., 2001).
18. Jacqueline Wasserman O'Brien and Steven R. Wasserman, eds., *Statistics Sources: A Subject Guide to Data in Industrial, Business, Social, Educational, Financial and Other Topics for the U.S. and Internationally* (Detroit: Gale Group, 2001).

19. Sears and Moody, *Using Government Information Sources*, 21.
20. University of Michigan Library, Government Documents Center, *Statistical Resources on the Web* (Ann Arbor: University of Michigan, 2001). Available at: <http://www.lib.umich.edu/govdocs/stats.html>. Accessed: February 21, 2003.
21. University of North Texas Library, Government Documents Department, *United States Statistical Agencies and Information* (Denton, TX: University of North Texas, 2002). Available at: <http://www.library.unt.edu/govinfo/subject/statusa.html>. Accessed: February 21, 2003.
22. Stephanie W. Haas and Carol Hert, "Find What You Need, Understand What You Find: Exploring Integration Issues," (May 15, 2002). Available at: http://www.ils.unc.edu/govstat/papers/scenarios5-151_files/frame.htm. Accessed: February 21, 2003.
23. This list of topics was derived from browsing Sears and Moody (pages 245-60); Joe Morehead, *Introduction to United States Government Information Sources*, 6th Ed., (Englewood, CO: Libraries Unlimited, 1999): 302-03; and the Census Bureau website Available at: <http://www.census.gov/>.
24. Margo J. Anderson, *Encyclopedia of the U.S. Census* (Washington: CQ Press, 2000); and Michael R. Lavin, *Understanding the Census: A Guide for Marketers, Planners, Grant Writers, and Other Data Users* (Phoenix: Oryx Press, 1996).
25. StatSoft, Inc., *Electronic Statistics Textbook* (Tulsa, OK, 2002). Available at: <http://www.statsoft.com/textbook/stathome.html>. Accessed: February 21, 2003.
26. Michael R. Lavin, "Introduction to Statistical Reasoning," in *Business Information: How to Find It, How to Use It* (Phoenix: Oryx Press, 1992).
27. David R. Gerhan, "When Quantitative Analysis Lies Behind a Reference Question" *Reference and User Services Quarterly* 39, no. 2 (Winter 1999): 166-76; and David R. Gerhan, "Statistical Significance: How It Signifies in Statistical Reference," *Reference and User Services Quarterly* 40, no. 4 (Summer 2001): 361-74.
28. Juri Stratford, "Responding to Reference Queries for Numeric Data and The Problems Inherent in Interpreting Statistical Sources: A Note," *Journal of Government Information* 25, no. 5 (1998): 413-17; and Alice Robbin, "The Problematic Status of U.S. Statistics on Race and Ethnicity: An 'Imperfect Representation of Reality,'" *Journal of Government Information* 26, no. 5 (1999): 467-83.