

---

# One Hundred and Eighty Years of Great Statistical Collections from the US Census Bureau

## *A Retrospective*

**Mark Anderson, Guest Columnist**

*Mark Anderson is Reference Librarian/Associate Professor of University Libraries, James A. Michener Library, University of Northern Colorado*

*Correspondence concerning this column should be addressed to **Marianne Ryan**, Associate University Librarian for Public Services, Northwestern University, 1970 Campus Drive, Evanston, IL 60208; e-mail: marianne-ryan@northwestern.edu.*

Generations of information seekers have relied on the statistical compendia—most notably, the *Statistical Abstract of the United States*—produced by the Census Bureau and its predecessor agencies. As the federally-funded versions of these information resources become a thing of the past, it is fitting to remember their unique benefit to researchers of all stripes and libraries of every type. In this retrospective, Mark Anderson recalls the launch of these publications, reviews their coming of age, and reminds us why it will be difficult to do without them.—*Editor*

In early 2011, librarians around the country were dismayed to learn that the Census Bureau budget for FY 2012 made no provision for the Statistical Compendia Branch, the agency which produced not only the renowned and much beloved *Statistical Abstract of the United States (SA)*, but also its major supplements the *County and City Data Book (CCDB)* and the *State and Metropolitan Area Data Book (SMADB)*. As justification, the Bureau pointed out that as more and more government data is accessible in user-manipulated electronic formats, print is an obsolete medium for most research purposes. No doubt there is some truth to that, but information seekers had come to rely on these unparalleled resources for good reason.

Single-volume collections of condensed social, economic, and demographic data have a long and esteemed history. The role of the census enumerators, who were originally charged by the Constitution with counting the free and slave populations of the various states for the purpose of congressional apportionment, expanded to collect a large amount of demographic and economic information about American families. After only a few census cycles, early demographers, social scientists, public planners, and businessmen saw the need for a summary of the most important data sets from the most recent censuses collected and available in one place.

The first census volume to bear the word “Abstract” in the title was published as a volume of the 1830 decennial census and was titled *Abstract of the Fifth Census*. This *Abstract* bore little resemblance to the present *SA*, and was merely a state-by-state and county-by-county summary of the populations of each, divided into the free and slave segments, and included a calculation of the number of House of Representatives seats each state would receive, based on those counts. According to Article 1, Section 2 of the U. S. Constitution, slaves counted as three fifths of a person.

The real grandparent of today’s *Statistical Abstract* was the brainchild of James Dunwoody Brownson De Bow, director of the 1850 Census. Mr. De Bow lobbied Congress to fund

the printing of a single volume of basic demographic, economic, and social statistics that would compare the current year with previous years. He wrote in the introduction to the *Statistical View of the United States*, “statistics . . . constitute the ledger of a nation, in which, like the merchant in his books, the citizen can read, at one view, all of the results of a year or of a period of years, as compared with other periods, and deduce the profit or the loss which has been made in morals, education, wealth and power.” Congress agreed, and a House resolution of January 12, 1854 authorized the publication of a hundred thousand copies of a compendium of the Seventh Census, so long as it was printed in royal octavo form and did not exceed 400 pages.

The censuses of 1860, '70, '80, and '90 each contained a compendium which summarized time-series data collected from many diverse agencies. However, the census bureaus that were responsible for the first 12 censuses (1790–1900) were ad hoc operations, each created by an act of Congress, placed under the jurisdiction of the State or the Interior Department, charged with collecting data and publishing the current census, and then dissolved when that work was done. Since census bureaus were not permanent, an annual SA was not possible.

The first volume carrying the title *Statistical Abstract of the United States* and published as an annual appeared in 1879. It was a product of the Statistics Division of the Treasury Department, so its contents were very different from the SA as we know it today. *Statistical Abstract 1878* heavily emphasized the economic and the commercial, such as federal revenues and expenditures, bank assets and debits, miles of railroad, agriculture and mineral production, and imports and exports. One demographic item included was counts of immigrants arriving at the major ports. This data was readily available to the Treasury Department because, in the days before the personal income tax, the major source of federal revenue was the tariffs levied on imported goods. The Internal Revenue Service had an office in each US port authority, and its agents inspected all foreign cargo bound for sale in the United States to calculate the appropriate tax. In addition to recording quantities of teakwood, coffee, tea, sugar, and “fancy articles, perfumery, etc.,” IRS agents counted passengers and divided them into the categories “alien passengers” and “immigrants,” thus recording for posterity not only numbers of immigrants and their countries of origin, but also data on the US as a world tourist destination.

Throughout the 1870s, '80s, and '90s, the SA continued to be published by the Treasury Department and was gradually expanded to include more detailed demographic and education data, though it continued to emphasize US economic and commercial expansion. By an Act of Congress in March of 1902, the Census Bureau was established as a permanent independent agency that answered directly to the President. In February of 1903, it was placed under the jurisdiction of the newly-created Department of Commerce and Labor. The *Statistical Abstract of 1903* was the first published with the Commerce Department's imprint. The next 110 volumes

appeared in roughly the current format, including those for multiple years, with annually updated tables of data collected from diverse government agencies and some non-government entities. This format has made the SA among the most frequently consulted reference tools in library collections.

Early in the twentieth century, demographers recognized a need for more locally-focused census geography concepts. Economic changes spurred by technological advances had resulted in major population shifts. Farm populations declined, and the national workforce was concentrated in the industrialized urban areas. Urban industrial economies were more complex, more integrated, and more interdependent than their rural agricultural counterparts. The Census Bureau, acknowledging these new economic realities, made its first attempt to quantify the different lifestyles of the urban and rural populations in the 1910 Census of Population. An anonymous source noted in the introductory essay that “cities represent, in comparison with the remainder of the county, a distinct type of economic and industrial life” and therefore, “for convenience of comparison, . . . statistics regarding the number, composition and characteristics of the population have been presented separately for urban communities as a group and for rural communities as a group.”<sup>1</sup> Incorporated places with more than 2500 inhabitants were considered urban; the remainder of the country was rural. The urban/rural distinction was employed in the censuses of 1920 and 1930, but for the 1940 census the concept was further expanded and “standard metropolitan areas” were created and defined. Demographers recognized that in very large urban areas, with populations over 200,000 and which might include more than one major city, a correspondingly larger degree of economic and social integration existed. There might be several commercial centers whose economic and social interdependence extended into the surrounding suburbs and rural areas. Metropolitan areas were defined to include all the counties that were deemed to be included in one of these “metropolitan areas,” and some of them, such as Cincinnati and Kansas City, crossed state lines.

The changing census geographies fueled a demand for new compendia that summarized statistics with more local foci. The first companion volume to the *Statistical Abstract* was the *Cities Supplement, Statistical Abstract of the United States: Selected Data for Cities Having 25,000 or More Inhabitants in 1940*, which appeared in 1944. The *Cities Supplement* contained 79 data items on population, labor force, housing, industry, retail and service establishments, city government finances and employment, vital statistics, and employment for about 400 cities of 2500 or more inhabitants.

The *County Data Book: A Supplement to the Statistical Abstract of the United States* followed in 1947. The *County* volume included 91 variables, 75 of them taken from the 1940 Census. The author of the introductory essay acknowledged additional contributions from the American Telephone and Telegraph Company, the US Departments of Agriculture and Treasury, the War Production Board, and various other offices in the Department of Commerce.

In addition to updated census geographies, some traditional concepts from earlier censuses that had been updated and modernized for the 1940 Census were included. For example, census takers in 1930 and before, attempting to measure the general state of education in America, had counted the “illiterate” and used a variety of tests and measures to ascertain which individuals qualified to be labeled as such. With the 1940 Census, “illiteracy” evolved into the more objective and measurable concept of “educational attainment.” Thus, the *Census of 1940* and *Counties* data book contained statistics on the percentage of county residents over the age of 25, who had completed “five or more grades” (the national average in 1940 was 85.1 percent) and “high school” (the national average was 24.1 percent).<sup>2</sup>

In 1949, the *Cities* supplement and the *Counties* collection were combined to form the first *County and City Data Book*. The introductory essay stated that its primary purpose was to “assist public officials, businessmen, research workers, teachers, librarians, and students who have a recurring need for a compact single source of statistics relating to counties, cities, and metropolitan areas.”<sup>3</sup> With 92 data items and 51 maps, *CCDB 1949* was also remarkable as an early attempt to organize data using Federal Information Processing Standards (FIPS) codes for geographies and Standard Industrial Classification (SIC) codes for businesses. The *CCDB* continued to be published at irregular intervals throughout the remainder of the twentieth century and into the twenty-first, combining data from the most recent census with data that was more frequently updated. Issues of the *CCDB* were produced for the years 1952, 1956, 1962, 1967, 1972, 1977, 1983, 1988, 1994, 2000, and 2007.

A spinoff publication, the *State and Metropolitan Area Data Book*, appeared in 1977. Like the *CCDB*, it tabulated statistics for “the entire area in and around a city, in which the activities from an integral economic and social system, needs to be considered as a unit.”<sup>4</sup> *SMADB 1977* contained 121 data items for states, regions, and the 277 designated Metropolitan Statistical Areas (MSAs). Later, *State and Metropolitan Area Data Books* were published for the years 1979, 1982, 1986, 1991, 1997–98, 2006, and 2010.

Acknowledging its original purpose as outlined in the US Constitution, the Census Bureau also produced four *Congressional District Data Books*, in 1961 for the 87th Congress, 1963 for the 88th, 1973 for the 93rd, and 1984 for the 98th. In addition to the familiar economic and demographic data that defined the constituencies of each congressional representative, these volumes included the total votes cast for the two major parties, for President, and for representative in the several most recent elections, along with a brief history of re-apportionment in the United States and the role of the Census Bureau.

Now the Census Bureau statistical compendia program is being phased out. Data users will have to search for the original sources of the non-census material. Each cabinet-level department has an agency specifically charged with collecting and publishing statistics. Thus, the National Center for

Education Statistics produces the *Digest of Education Statistics* and the *Condition of Education*. The National Center for Health Statistics publishes *Health United States* and the *Vital Statistics of the United States*. The Bureau of Justice Statistics publishes *Crime in the United States (Uniform Crime Reports)* and *Sourcebook of Criminal Justice Statistics*, and so on. These collections of national and state-level statistics are universally available in pdf format from agency websites. Supplementing these, most statistics-gathering agencies also make available collections of raw data attached to user-friendly search tools. County- and city-level statistics can usually be accessed, multiple years can be searched in one session, and data can be saved in a variety of electronic formats. With little effort and without relying on someone else’s selection criteria, an enterprising data seeker can create a personal collection of cross-tabulations and ask questions of the data that have not been asked before. With current technology, a tenacious data user, student, teacher, or librarian can assemble a personal statistical compendium containing many statistics at every geographic level.

Bernan has announced that it will continue to publish the annual *Statistical Abstract of the United States*, and Proquest will make it available digitally. Perhaps the other publications will eventually be picked up by other vendors. But will the original purposes of these valuable information resources survive in a commercial environment? Mr. De Bow and the other architects of the original *Statistical Abstract* believed they were proving with empirical evidence the doctrine of Manifest Destiny. Increasing populations, industrialization, agricultural productivity, property values, and bank assets were directly related to an increase of the number of educated citizens, rising wages, and declining poverty. The *Statistical View* of 1850 quantified no less than the advance of civilization itself, and showed how its benefits were universally manifest in the comparatively more comfortable lifestyles of ordinary Americans.

Subsequent generations of researchers and government planners were confident that they had access to the objective results of national studies. If the private sector takes over publishing these valuable reference tools, it is possible that only what is commercially viable will be produced and that their quality may be compromised or content biased. Librarians will no doubt be very interested in these developments and remain vigilant as they unfold.

## References

1. US Department of Commerce, Bureau of the Census, *Thirteenth Census of the United States Taken in the Year 1910*, vol. 1, *Population 1910* (Washington, DC: GPO, 1913).
2. US Department of Commerce, Bureau of the Census, *Cities Supplement, Statistical Abstract of the United States: Selected Data for Cities Having 25,000 or More Inhabitants in 1940* (Washington, DC: GPO, 1944).
3. US Department of Commerce, Bureau of the Census, *County and City Data Book 1949: A Statistical Abstract Supplement* (Washington, DC: GPO, 1952).
4. US Department of Commerce, Bureau of the Census, *Statistics for States and Metropolitan Areas: A Preprint from County and City Data Book 1977* (Washington, DC: GPO, 1977).