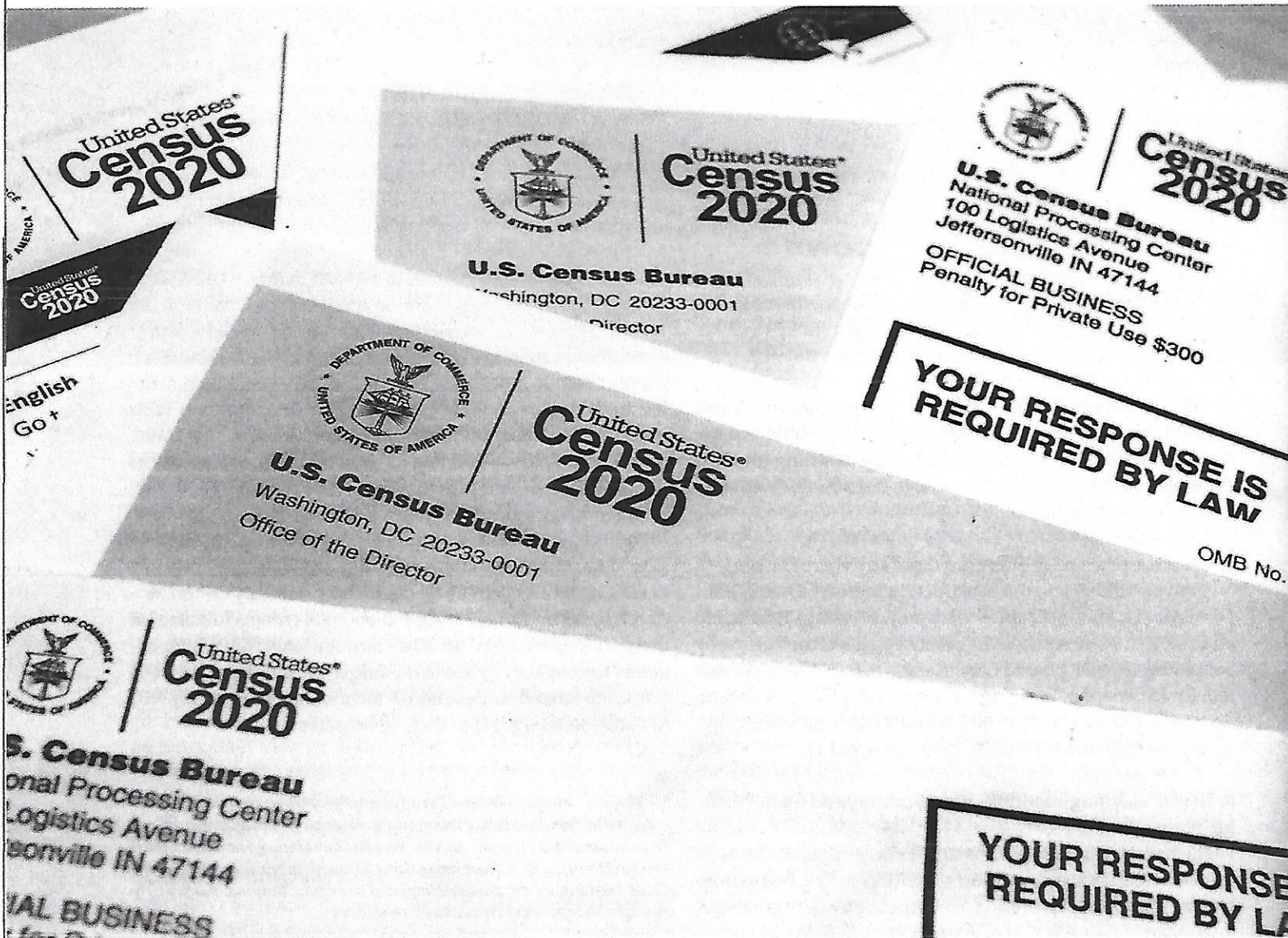


# Finding and Evaluating Census Data

By Roger Magnus

## A 360-Degree Perspective





People who regularly use data from the U.S. Census Bureau's website ([census.gov](https://www.census.gov)) often consult only a particular data program or a small group of programs. Frequently, these are the better-known ones, including the Decennial Census or the Economic Census. Yet the Census Bureau's data reach is actually much broader. It covers an astonishing 130-plus programs in subject areas as diverse as education, income, health insurance, poverty, and time use.

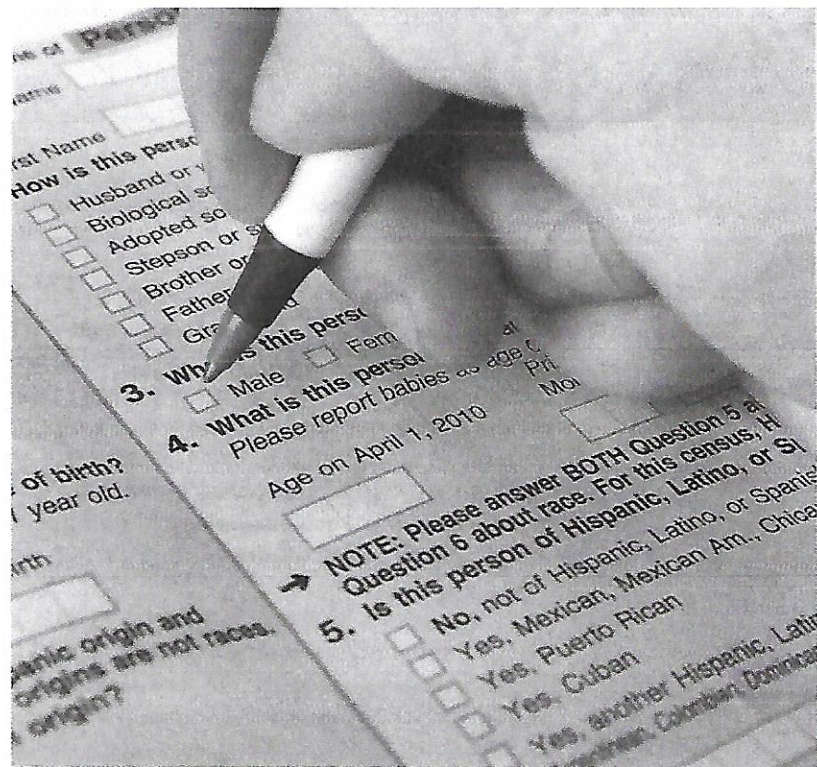
**G**iven this volume and diversity of data all in one place, are there ways to locate and compare multiple census programs to optimize the final data provided? This article will demonstrate how to approach the Census Bureau website and data programs in a broader way than is typically done in order to use this amazing resource to its fullest potential.

#### WHERE TO BEGIN

There are at least five major searching interfaces or listings on the Census Bureau's site that allow for searching data programs and datasets. (Note that datasets are usually data programs broken down by a particular year or topic.)

- [data.census.gov](https://data.census.gov)
- Index A-Z
- List of All Surveys and Programs
- Datasets
- Tables

Census data is also searchable in more granular and integrated ways. For example, there is a Browse by Topic link from the homepage that allows searching by 19 various subject areas in addition to Index A-Z. Clicking on Age and Sex, for instance, shows links to Data (tables) in reverse-chronological order and a targeted linked list of Surveys and Programs that contains this type of data. Another example is the link to Surveys and Programs. From the homepage, it connects directly to more than 15 of these data programs. Also, Datasets and Tables contain filters by topic, data program, and year.

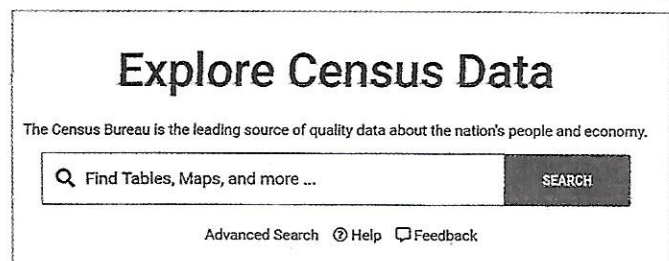


#### DATA.CENSUS.GOV

Since June 2019, [data.census.gov](https://data.census.gov) has been the successor data portal to the American FactFinder (which was phased out in March 2020) and is still having data added to it. Currently, it covers more than 10 data programs back to 2000. Sometime soon, [data.census.gov](https://data.census.gov) also plans to add Small Area Income and Poverty Estimates (SAIPE) as well as Equal Employment Opportunity (EEO) tabulations.

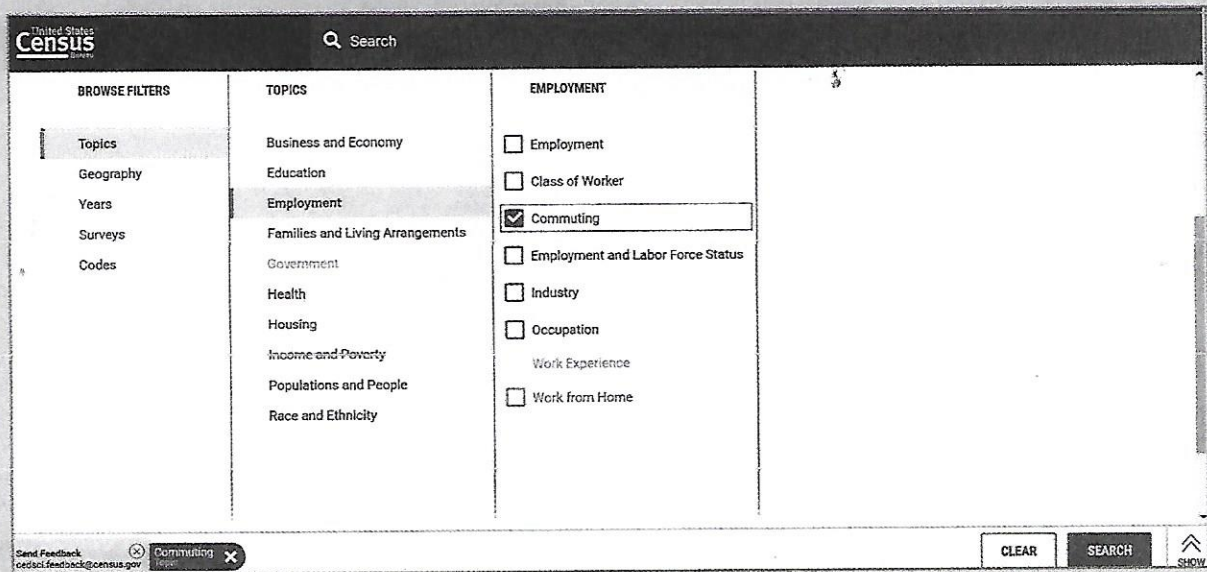
There are two ways to search [data.census.gov](https://data.census.gov): Basic (keyword) and Advanced (browse filters and drill down into sub-filters or search for a filter or Table ID). The main list of filters includes five main types: Topics, Geography, Years, Surveys, and Codes. You can select filters in more than one type. The main menu has 10 topics, ranging from Business and Economics to Race and Ethnicity. In Geography, there are 56 Summary Levels, from Nation to Block. Years covered are 2010–2020 and 2000. Surveys show the data programs covered, which is a small subset of the Decennial Census, American Community Survey, the Economic Census, and a few others. Codes are primarily for industries, products, and commodities.

When browsing, either check a filter or sub-filter (if that option is available), or click on the filter name to find additional sub-filters that can be checked. Clicking on the Search



*The Basic Search box for [data.census.gov](https://data.census.gov)*





The Advanced Search on data.census.gov showing filters for topics and subtopics

button in the bottom right-hand corner pops up a customized page with the three results formats, Tables, Maps, and Pages, as well as a combination in lists of up to 10. There is also a View All button below each section to access a complete list of links. To the right, there is a link to Public Use Microdata (currently in beta). Search within results with the search box at the top of the page.

Because data.census.gov is still being developed, some filters are grayed out and not yet available. There is also a Resources page (census.gov/data/what-is-data-census-gov.html) and FAQ page (census.gov/data/what-is-data-census-gov/guide-for-data-users/frequently-asked-questions.html).

A Basic Search on **Commute**, for instance, retrieves 74 tables, 74 maps, and 286 pages. The initial list of tables and maps come from the American Community Survey and the Economic Census Bureau. The pages here link to general website articles on the topic and can sometimes link to additional data programs not found in data.census.gov

## INDEX A-Z

Index A-Z (census.gov/about/index.html) can be accessed from the Census Bureau homepage top banner under Browse by Topic and then A-Z. When accessing it this way, you can select one of 19 specific topics for precision searching. The Index A-Z page is one continuous page: You can browse by a particular letter, browse the whole page, or search on a specific term. The index covers a wide variety of census subject areas, programs, concepts, and more. Clicking on a link frequently brings up this item in more detail, including further explanations, data programs, and actual data.

## LIST OF ALL SURVEYS AND PROGRAMS

Similar to Index A-Z, List of All Surveys and Programs (census.gov/programs-surveys/surveys-programs.html) can be accessed from the Census Bureau homepage top banner under

Surveys/Programs and then All Surveys and Programs, bypassing the 19 specific data program links.

Unlike Index A-Z, the page accessed is one of seven. Along the left banner is a link to the only three true census programs the Census Bureau conducts: Decennial, Economic, and the Census Bureau of Governments. Every other data program is a sample survey—so maybe the U.S. Census Bureau is a misnomer. The survey and program links are arranged alphabetically by name and can be browsed or searched page by page. Unfortunately, a list of all programs is not found on one page, which would make for more efficient searching. Clicking on a link brings you directly to the data program page that explains its data, methodology, and purpose.

## DATASETS

Datasets (census.gov/data/datasets.html) are portions of data programs that can be accessed from the Census Bureau homepage top banner under Explore Data and Data Tools and Apps. From the Data Tools page, look at the left banner and click on Datasets. The listing, as of mid-December 2020, had 545 datasets covering 28 pages. To make this more manageable, along the left banner there are specific filters by Topics, Surveys and Programs, and Years. One advantage over data.census.gov is that datasets can be filtered back all the way to 1970. They can also be sorted chronologically, alphabetically, or the reverse of these.

## TABLES

Tables (census.gov/data/tables.html) are also found under Data Tools and Apps and have the same setup for filters (although these have some different subject categories and subcategories) and sorting mechanisms as datasets. The Years filter goes all the way back to 1890. There are currently 2,173 tables covering 109 pages.



**Data Goal: Find Educational Attainment Data for the State of Massachusetts.**

<b>Considerations</b>	<b>ACS</b>	<b>CPS</b>
<b>Survey?</b>	Yes. (Note: Margins of Error listed and affect geographic and time comparisons.)	Yes.
<b>How current?</b>	2019	2020
<b>Updated how frequently?</b>	Yearly and 5-year average data (survey is monthly)	Annual Social and Economic Supplements
<b>How far back?</b>	2005 (in some cases, 1990s)	1940
<b>Limited or extensive variable coverage?</b>	Extensive. Questions about Educational Attainment and Field of Study. Also see <a href="http://census.gov/programs-surveys/acs/technical-documentation/table-shells.html">census.gov/programs-surveys/acs/technical-documentation/table-shells.html</a> .	Extensive. Educational Attainment cross-tabbed for Personal Income, Age, Race, Poverty, etc.
<b>Geographic coverage?</b>	Down to Block Group. For a list of geographies available, go to <a href="http://census.gov/programs-surveys/acs/geography-acs/areas-published.html">census.gov/programs-surveys/acs/geography-acs/areas-published.html</a> .	Nation, State
<b>Overlapping/complementary datasets?</b>	Decennial Census, Annual Population Estimates	NA
<b>Usage complexities?</b>	Methodological: Margin of Error, 1-year and 5-year average data, variable definitions changes, geographic changes, etc.  Technological: <a href="http://Data.census.gov">Data.census.gov</a> , Microdata, API, FTP, Summary File Variance Replication Tables, etc.	Methodological: Weighting of data  Technological: CPS Data Creator (unique interface) and Microdata
<b>How are data used?</b>	Population Characteristics (See <a href="http://census.gov/acs/www/about/why-we-ask-each-question/education">census.gov/acs/www/about/why-we-ask-each-question/education</a> .)	Labor Force Data
<b>Other</b>	Sent to 3.5 million households annually.  Page: Guidance for Data Users  Response Rates: See <a href="http://census.gov/acs/www/methodology/sample-size-and-data-quality/response-rates/index.php">census.gov/acs/www/methodology/sample-size-and-data-quality/response-rates/index.php</a> .	60,000 households interviewed monthly over 8 months, covering those 16 years of age and older.



## 6. Geographic Coverage

This criterion can sometimes make or break finding a solution to a data request and often requires consulting multiple census programs (or even other government datasets) if a smaller geographic level is needed. For example, the Economic Census only goes down to the county level, but ZIP Code Business Patterns (a similar survey) covers ZIP code data.

Even within the same program, certain variable tables may not go down to the same level if data need to be suppressed to preserve respondent privacy. The Census Bureau additionally has created several custom geographies, notably Census Tract (1,200–8,000 people), Block Group (600–3,000 people), Block (<3,000 people or the equivalent of a city block in urban areas), and a few others for the Decennial Census and ACS.

## 7. Industry Coverage

For the Census Bureau's business and economic data programs, the industry coverage can range from specific to broad. Specific programs include the Annual Survey of Manufacturers (there are several other programs for manufacturers) and Annual Survey of Retail Trade. These surveys go into greater variable detail than some of the broader programs but often only at the national level. For example, the Annual Survey of Manufacturers covers just NAICS codes 31 and 32 at the U.S. level but includes data for more than 25 different variables. A broader industry survey such as the Economic Census (others are County Business Patterns, ZIP Code Business Patterns, and Nonemployer Statistics) covers all NAICS codes to the county level but only includes a handful of variables.

## 8. Overlapping/Complementary Datasets

Data programs can contain the same variables, but their numbers differ due to different data definitions and/or methodology. A simple example is that the PEP, which produces population data between each Decennial Census and provides data for July 1 of every year, while the Decennial is April 1. That means these two data programs will have different numbers in the same year when the Decennial Census is also collected. The Economic Census (as well as County Business Patterns and ZIP Code Business Patterns) and Nonemployer Statistics programs contain complementary data because the former covers businesses with employees and the latter does not. This is rare, and there is actually a special tabulation combining County Business Patterns and Nonemployer Statistics programs.

Note that sometimes census datasets overlap with other free sources of government data (they each differ somewhat in methodology and data availability) such as income data in the ACS that is also available from the Bureau of Economic Analysis, Bureau of Labor Statistics, and the IRS.

## 9. Usage Complexities

Some of the Census Bureau's data programs and ways of presenting data come with technological hurdles. One example is having to learn how to search and customize tables in data.census.gov for the data programs available there.

There are also several notable methodological obstacles in using the American Community Survey that can apply to other data programs as well. They include the following:

- Understanding margins of error and standard error—how they differ and how you can convert one to the other
- Awareness of how margins of error affect data comparisons across time and geography
- Determining if data definitions and geographic boundaries have changed across time, which may affect historical comparisons
- Comprehending what multi-year average data are and how they can only be compared to data covering the same time span in discrete periods

## 10. Uses of Data

The Census Bureau now provides section links in many data program homepages (terms such as About/About This Program, Uses section, or Using Census Bureau Stats) to show real-world examples of how the data are used. While not matching your data need, the uses may show if the dataset approximates it or provides additional possibilities previously not considered.

## REVIEWING A DATA PROGRAM

If you find a potentially relevant data program, look closely at its homepage to become acquainted with its inner workings to determine its usefulness.

For example, take a look at the American Housing Survey (AHS; [census.gov/programs-surveys/ahs.html](http://census.gov/programs-surveys/ahs.html)) for housing data. Note that most census data programs have a *menu* (and sometimes submenus) of links along the left banner. (There may also be relevant links on the homepage.) First, always read the About This Survey link. It provides a general overview of what the survey covers. Toward the bottom of the About page, there are also links to AHS Introduction and History, Methodology, and Respondent Information, all of which should be reviewed. On the left banner, look at *Technical Documentation*, which has a subpage with links to Definitions, Questionnaires, Help Guides, and User Notes. Finally, along the left panel, there are also links to Data and Visualizations.

There are several options to locate data across the entire Census Bureau website. While taking a holistic perspective is more challenging than looking at a particular data program, as it requires rigorously finding, evaluating, and comparing multiple data programs, the potential rewards of finding more data and additional and richer ways of seeing data for the same topic cannot be overestimated.

---

*Roger Magnus* ([roger@rogermagnusresearch.com](mailto:roger@rogermagnusresearch.com)) is the owner of Roger Magnus Research, based in Amherst, Mass.  
Comments? Email the editor-in-chief ([marydee@xmission.com](mailto:marydee@xmission.com)).