

data.census.gov: A Brief Demonstration

Frank Donnelly, Geospatial Data Librarian, Baruch College CUNY

August 30, 2019

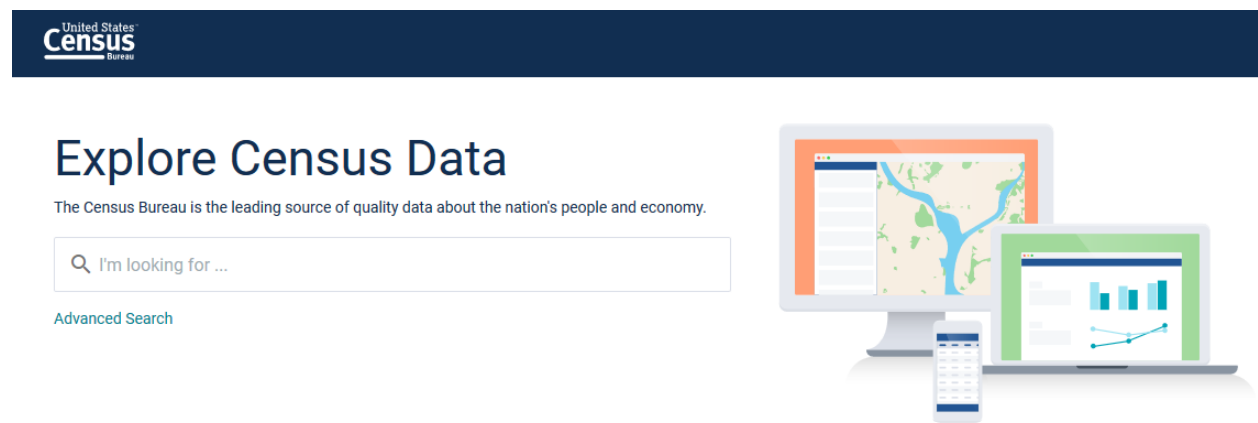
Abstract

This tutorial demonstrates how to use data.census.gov to access and download US census data, highlighting a specific approach using the advanced search feature. Given the size and complexity of census datasets, we'll use the filtering options to limit our choices based on dataset and geography with two examples: getting a profile for a specific census tract and downloading rental data for all counties in a state from the American Community Survey (ACS). We'll conclude with summaries of the census datasets and geographies.

1 Methods for Searching data.census.gov

In mid-2019 data.census.gov (at <https://data.census.gov/>) was launched as the new portal for accessing the Census Bureau's datasets, replacing the American Factfinder. There are two principle options for obtaining data:

1. To get basic data for a state, county, city, town, or ZIP Code, type the name in the search box and hit go. You will be able to preview recent data for these places, and navigate to individual tables. You can even type the name of a topic, like educational attainment, to find all tables related to this subject. This type of search allows you to casually browse and access data.
2. If you want to chart a straight path to access the data you need, and you want to see the full range of datasets, topics, and geographies, use the Advanced Search.



The following two sections focus on using the Advanced Search, which has a steeper learning curve but brings you more directly to the data you're looking for. The final section discusses the different datasets and geographies, which are important to understand for doing advanced searches.

2 Downloading a Profile from the ACS

In this section we'll download a demographic profile for census tract 68 in New York City (where Baruch College is located) that will give us a broad overview of the population for that area. We'll look at 5-year estimate data from the 2013-2017 American Community Survey (ACS). Census tracts are only included in the 5-year ACS and the decennial census due to their small size, and the ACS is the dataset that contains detailed socio-economic characteristics of the population.

1. Go to data.census.gov at <https://data.census.gov/cedsci/>.
2. Under the search box select the Advanced Search option.
3. In the filters menu on the left, select Surveys. In the list of Surveys check ACS 5-Year Estimates Data Profiles.

Advanced Search

Narrow search with filters

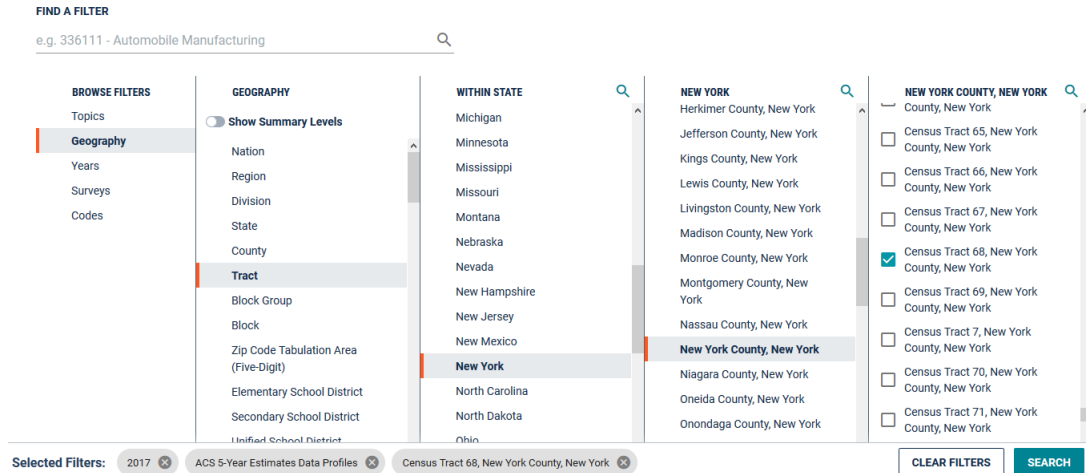
FIND A FILTER

e.g. 336111 - Automobile Manufacturing



BROWSE FILTERS	SURVEYS
Topics	<input type="checkbox"/> ACS 1-Year Estimates Comparison Profiles
Geography	<input type="checkbox"/> ACS 1-Year Estimates Data Profiles
Years	<input type="checkbox"/> ACS 1-Year Estimates Detailed Tables
Surveys	<input type="checkbox"/> ACS 1-Year Estimates Selected Population Profiles
Codes	<input type="checkbox"/> ACS 1-Year Estimates Subject Tables
	<input type="checkbox"/> ACS 5-Year Estimates Comparison Profiles
	<input checked="" type="checkbox"/> ACS 5-Year Estimates Data Profiles

4. In the menu on the left select Years, and check 2017.
5. In the menu select Geography. Choose Tract. When the next menu opens, choose New York under Within State, then New York County within the state, then within New York County scroll down and check Census Tract 68.



6. Hit the Search button in the lower right-hand corner.
7. This takes you to a page that has results for tables, maps, and pages. At the top of the screen where it says ALL, click on the option for TABLES.
8. This brings you to the Tables result screen. The four tables that are returned, with ID numbers DP02 to DP05, are data profile tables. They contain a large swath of data divided into broad topics: demographic, social, economic, and housing. Clicking on a table listed on the left will provide a preview on the right. Click on the Selected Social Characteristics table to preview it. You'll see estimates with margins of error and percent totals for each value. In this example, we would say with 90% confidence that census tract 68 has 3,547 households, plus or minus 306.

Census Tract 68, New York County, New York				
	Estimate	Margin of Error	Percent	Percent Margin of Error
HOUSEHOLDS BY TYPE				
Total households	3,547	+/-306	3,547	(X)
Family households (fa...)	780	+/-223	22.0%	+/-6.4
With own children of...	245	+/-122	6.9%	+/-3.5
Married-couple family	585	+/-216	16.5%	+/-6.2
With own children of...	184	+/-121	5.2%	+/-3.4
Male householder, no ...	33	+/-52	0.9%	+/-1.5
With own children of...	33	+/-52	0.9%	+/-1.5
Female householder, ...	162	+/-120	4.6%	+/-3.4
With own children of...	28	+/-47	0.8%	+/-1.3
Nonfamily households	2,767	+/-365	78.0%	+/-6.4
Householder living a...	2,044	+/-392	57.6%	+/-8.2
65 years and over	493	+/-170	13.9%	+/-4.3

9. Click on the Customize Table button in the upper right-hand corner of the screen. This screen allows you to sort, hide, or filter data in the table.

// Search / Tables / DP02

SELECTED SOCIAL CHARACTERISTICS IN THE UNITED STATES

Survey/Program: American Community Survey TableID: DP02 Product: 2017: ACS 5-Year Estimates Data Profiles

Data Notes | 1 Geography | 1 Year | Topic | 1 Survey | 123 Code | Hide | Filter | Sort | Transpose Table | Margin of Error | Restore Layout | Download | Print | Share

Census Tract 68, New York County, New York			
	Estimate	Margin of Error	Percent
▼ HOUSEHOLDS BY TYPE			
▼ Total households	3,547	+/-306	3,547
▼ Family households (families)	780	+/-223	22.0%
With own children of the hous...	245	+/-122	6.9%

- Click the Download button on the toolbar above the table. Keep the default option for CSV. Hit the Download button on this screen, and Download Now once the file is ready. When prompted save the file on your computer.

Download / Print / Share

DOWNLOAD EMBED SHARE API PRINT MORE DATA

Select Table Vintages

	All	2017
DP02 5-Year	<input type="checkbox"/>	<input checked="" type="checkbox"/>

File Type

CSV

PDF

What You're Getting

- 1 .csv files (metadata)
- 1 .csv files (data)
- 1 .txt files (table title)

Uncompressed Estimated Size: 69.2 kB

DOWNLOAD

- All data that you download will be compressed in a zip file. Navigate to where the file was downloaded, right click, and choose the option to unzip the file (this step will vary based on your web browser and operating system). The csv file that contains the word 'data' is the actual data file, while the csv with the word 'metadata' contains a list of all the columns in the file. CSV files can be opened with any spreadsheet package, like Excel or Calc.

3 Downloading a Detailed Table from the ACS

The data profile tables are useful for obtaining a cross-section of different variables. If you want something more focused, either fewer variables or a crosstab of data not published in the profiles, you'll want to search for detailed tables instead. You may also want to compare data for multiple geographies. In this section we'll download ACS data on rent for all counties in New York State.

- Go to <https://data.census.gov/cedsci/> and select the Advanced Search option.
- If you just followed the previous example, select the Clear Filters button at the bottom of the Advanced Search screen to remove the filters from the previous search.
- In the filter menu, choose Surveys, then select ACS 5-Year Estimates Detailed Tables.

FIND A FILTER

e.g. 336111 - Automobile Manufacturing

BROWSE FILTERS	SURVEYS
Topics	<input type="checkbox"/> ACS 1-Year Estimates Subject Tables
Geography	<input type="checkbox"/> ACS 5-Year Estimates Comparison Profiles
Years	<input type="checkbox"/> ACS 5-Year Estimates Data Profiles
Surveys	<input checked="" type="checkbox"/> ACS 5-Year Estimates Detailed Tables
Codes	

- In the menu on the left select Years, and check 2017.
- In the menu select Geography. Choose County. In the County menu choose New York as the state. In the New York menu check the first box for All counties in New York.

BROWSE FILTERS	GEOGRAPHY	COUNTY	NEW YORK
Topics	<input type="checkbox"/> Show Summary Levels	Missouri	Within Other Geographies
Geography	Nation	Montana	<input checked="" type="checkbox"/> All counties in New York
Years	Region	Nebraska	<input type="checkbox"/> Albany County, New York
Surveys	Division	Nevada	<input type="checkbox"/> Bronx County, New York
Codes	State	New Hampshire	<input type="checkbox"/> Broome County, New York
	County	New Jersey	<input type="checkbox"/> Cattaraugus County, New York
	Tract	New Mexico	<input type="checkbox"/> Cayuga County, New York
	Block Group	New York	
		North Carolina	

- In the filters menu select Topics. Choose Housing, then Financial Characteristics, then Renter Costs.

BROWSE FILTERS	TOPICS	HOUSING	FINANCIAL CHARACTERISTICS
Topics	Business and Economy	<input type="checkbox"/> Housing	<input type="checkbox"/> Financial Characteristics
Geography	Education	<input type="checkbox"/> Absorption Rate	<input type="checkbox"/> Delinquent Payments, Foreclosures, and Evictions
Years	Employment	Financial Characteristics	<input type="checkbox"/> Government Subsidies and Tax Credit
Surveys	Families and Living Arrangements	Health and Safety Characteristics	<input type="checkbox"/> Home Improvement and Repairs
Codes	Government	<input type="checkbox"/> Homeownership Rate	<input type="checkbox"/> Housing Value and Purchase Price
	Health	<input type="checkbox"/> Housing Units	<input type="checkbox"/> Insurance, Utilities, and Other Fees
	Housing	New and Existing Units	<input type="checkbox"/> Mortgage Costs
	Income and Poverty	Occupancy Characteristics	<input type="checkbox"/> Real Estate Taxes
	Populations and People	Owner/Renter (Householder) Characteristics	<input type="checkbox"/> Rental Property Ownership
	Race and Ethnicity	Physical Characteristics	<input checked="" type="checkbox"/> Renter Costs
		Vacancy	

- Hit the Search button in the lower right-hand corner.
- This takes you to a page that has results for tables, maps, and pages. At the top of the screen where it says ALL, click on the option for TABLES.

- The Table results window contains over thirty tables that are associated with renter costs. Select a table in the list on the left to view it on the right. Each of the detailed tables begins with a letter 'B' for base table, which means it's a detailed table focused on a narrow characteristic. Tables that begin with a 'C' are consolidated tables with fewer categories. Data is summarized in different ways; the Contract Rent table estimates the number of households in different rent brackets, while Median Contract Rent provides the median or middle dollar value.

The screenshot shows the 'MEDIAN CONTRACT RENT (DOLLARS)' table for Schoharie County, New York. The table has three columns: 'Median contract rent', 'Estimate', and 'Margin of Error'. The data row shows a median contract rent of 643, an estimate of 643, and a margin of error of +/-30.

MEDIAN CONTRACT RENT (DOLLARS)		
Schoharie County, New York		
	Estimate	Margin of Error
Median contract rent	643	+/-30

- Select the Median Contract Rent table in the results. Hit the Customize Table button. On the Customize screen hit the Download button. Keep the file as type CSV and Download the table.
- Unzip the download file, then open the csv file that contains the word 'data' in Excel or Calc. You'll notice that each row represents a piece of geography and each column is an attribute, which is the opposite of the way it was displayed on the website.

	A	B	C	D
1	GEO_ID	NAME	B25058_001E	B25058_001M
2	id	Geographic Area Name	Estimate!!Median contract rent	Margin of Error!!Median contract rent
3	0500000US36095	Schoharie County, New York	643	30
4	0500000US36005	Bronx County, New York	1024	5
5	0500000US36067	Onondaga County, New York	710	7
6	0500000US36035	Fulton County, New York	574	16
7	0500000US36019	Clinton County, New York	657	18
8	0500000US36099	Seneca County, New York	581	27
9	0500000US36083	Rensselaer County, New York	769	15
10	0500000US36079	Putnam County, New York	1191	41
11	0500000US36033	Franklin County, New York	549	20
12	0500000US36043	Herkimer County, New York	529	12

4 Other Tips

- The data.census.gov platform is relatively new, and doesn't support many filtering options for capturing all geographies within another geography. The options should improve over time.
- In the ACS, Subject tables are an alternative that sit in-between the profile and detailed tables. They contain a wider array of variables in one table relative to the detailed tables, but they are more focused than the profile tables. Access them via the Surveys menu, their ID numbers begin with the letter 'S'.
- 1-year ACS estimates are only available for geographies that have 65,000 people or more. If you need data for small areas (tracts, ZCTAs) or need complete coverage for a geography that contains a mix of places above and below 65k (counties, places), use the 5-year ACS.

4. The decennial census (2020, 2010) contains fewer variables but is an actual count of the population. Under Surveys the primary file is DEC Summary File 1. The table designation system differs from the ACS; the letter P is for population counts, H is for housing counts, and QT are quick tables similar to ACS Subject tables. There is one data profile table, DP-01.

5 Dataset Summary

Decennial Census

The decennial census is a 100% count of the population taken every ten years. The data represents a snapshot of the nation at a fixed point in time. The number of variables is limited to basic demographic characteristics of the population: race, sex, age, households, family relations, housing units, housing occupancy, and tenure (owner versus renter occupied). Decennial census data is available for practically all geographic areas. The most commonly used series in the dataset is Summary File 1 (SF1).

The 2020 census population count for states will be published in December 2020. The first significant release of 2020 census data will occur in spring 2021: the redistricting file, PL 94-171, contains few variables but will be published for most geographies down to the smallest ones. Census 2020 Summary File 1 will be released on a rolling basis beginning in the summer of 2021.

American Community Survey (ACS)

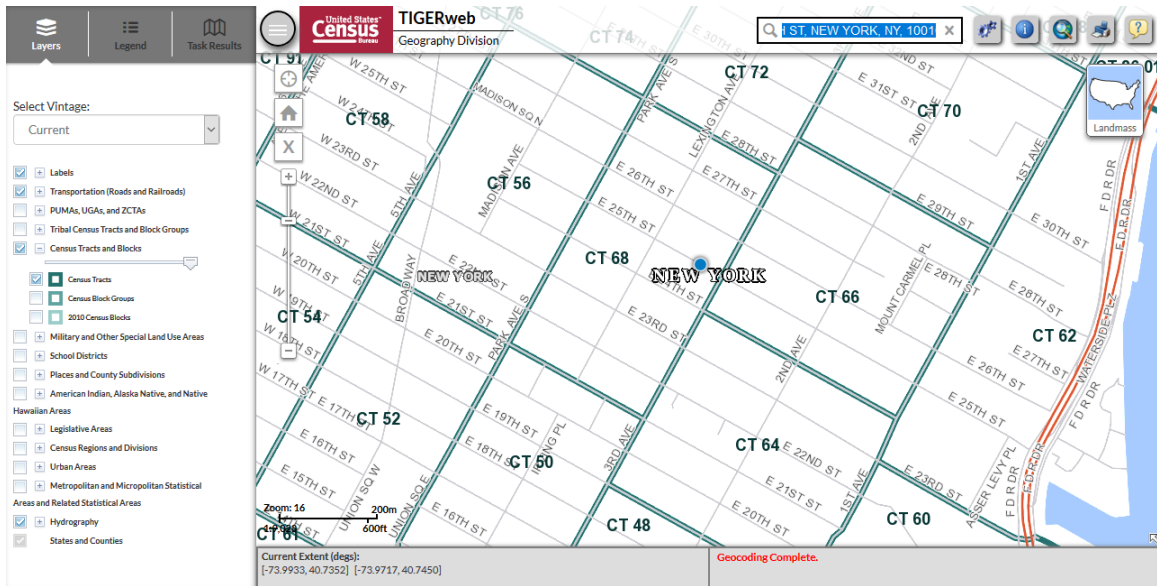
The ACS is a sample survey of the population that's compiled every year. Since it's a survey (based on rolling monthly samples) and not a count, the values represent estimates with a 90% confidence interval and margins of error for 1-year and 5-year periods. The number of variables is more extensive than the decennial census; in addition to basic demographic characteristics the ACS also includes socio-economic data like citizenship, educational attainment, income, occupation, home value, and much more. Geographically the data is more limited; geographic areas with 65k residents or more are published annually, and areas with less than 65k residents down to census tracts are published as 5-year averages. New data is published towards the end of each calendar year.

Population Estimates

Population estimates are a third census dataset for population and housing. Annual estimates are calculated using the last decennial census as a baseline and factoring in births, deaths, and migration at the county level from year to year. Variables are limited to age, sex, race, and housing units and are only published for states, counties, metropolitan areas, and incorporated places. These estimates are treated like counts, since they are generated from calculations and not from a sample survey.

Geography in Brief

To view census geography for an area, visit TIGERweb at <https://tigerweb.geo.census.gov/tigerweb/>. Choose the type of geography you want to see, check some reference layers to turn them on, and do a search for an address. An address search for Baruch College's Newman Library displays its location within New York County census tract 68:



Census Tract : a statistical area designed to have an optimal size of 4,000 residents, with a general range of 1,200 to 8,000. Tracts never cross county boundaries. Available in the decennial census and 5-year ACS.

County : the legal and administrative subdivisions of states; there are over 3,000 counties in the US. Available in the decennial census and the ACS (partial 1-year, fully 5-year).

Metropolitan Area : a concentrated, urban population center surrounded by adjacent areas that have a high degree of social and economic interaction with that center, created by aggregating counties. Available in the decennial census and the ACS.

Place : a concentrated settlement of population. Places include legally incorporated municipalities and statistical areas defined by the Census Bureau. Available in the decennial census and the ACS (partial 1-year, fully 5-year).

PUMA : Public Use Microdata Area - a statistical area designed to have an optimal size of 100k residents, created by aggregating census tracts. PUMAs never cross state boundaries. Available in the ACS.

ZCTA : ZIP Code Tabulation Area - the Census Bureau's approximation of residential USPS ZIP Codes, created by aggregating census blocks. ZCTAs may cross any boundary, including state boundaries, and do not precisely represent USPS delivery areas. Available in the decennial census and 5-year ACS.

For More Info

Visit these research guides:

- US Census Data: http://guides.newman.baruch.cuny.edu/us_census
- New York City Data: http://guides.newman.baruch.cuny.edu/nyc_data