3. CASE STUDIES USING ACS DATA

Case Study #1: San Diego Region: City Data Profile

Skill Level: Introductory/Intermediate
Subject: County- and city-level demographic and socioeconomic data
Type of Analysis: Analysis and visualization of American Community Survey (ACS) data trends across the nation, state, county, and subcounty areas
Tools Used: Data.census.gov, spreadsheet, data visualization tools
Authors: Marcela Alvarez, Research Coordinator, San Diego Regional Economic Development Center (EDC); and Kirby Brady, Research Director, San Diego Regional EDC

San Diego Regional Economic Development Corporation (EDC) uses a multitude of publicly available data sets coupled with visualization software to create powerful economic development tools that provide clarity and insight into the region's demographic and economic landscape. As part of a greater initiative to uncover and highlight subregional trends within the county, EDC produced a demographic and socioeconomic profile for the City of San Marcos, California. This profile and accompanying online interactive dashboard were created to better inform business leaders, other decision makers, and the public about the city's social, economic, and demographic characteristics. City-level data can be compared with data for the county, state, and nation to provide additional context.

We used data from the U.S. Census Bureau's American Community Survey (ACS) and Longitudinal Employer-Household Dynamics (LEHD) program to inform parts of the profile (see Table 3.1). The Census Bureau releases single-year ACS data for geographic areas with 65,000 people or more. Since the City of San Marcos—the smallest geographic area under consideration—meets this threshold, we were able to use 1-year estimates throughout the analysis. EDC recognizes that all estimates have an associated level of uncertainty. However, after careful consideration, margins of error were not included in the profile.

Other data sources used in the profile include the San Diego Association of Governments, Emsi (a labor market analytics firm), the State of California Employment Development Department, and ReferenceUSA.

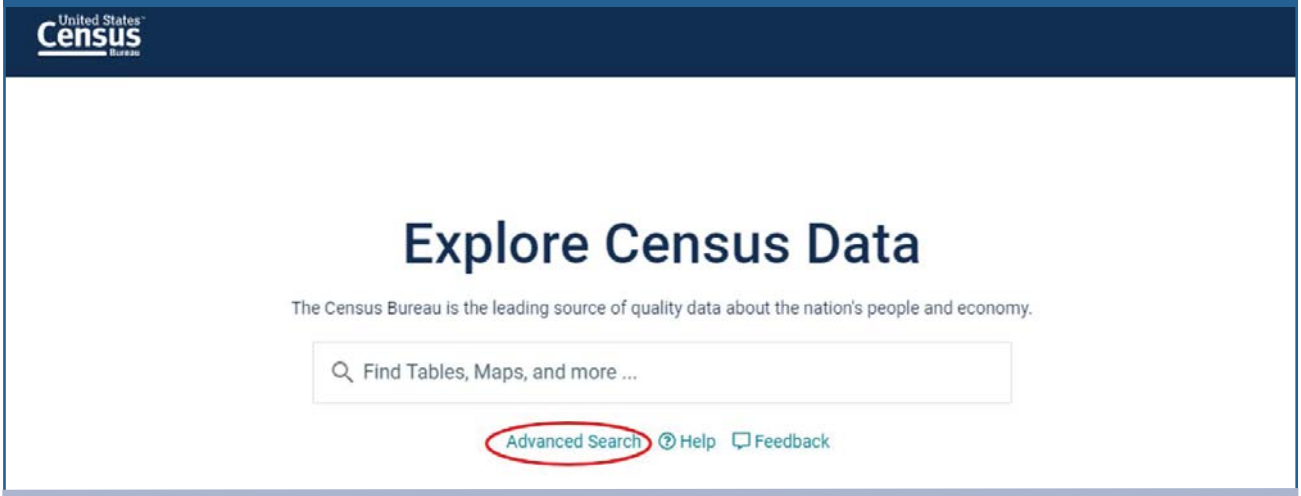
Table 3.1. **List of Key Variables and Data Sources**

Variable	Data Set
Median household income	ACS 1-Year Estimates
Population below poverty level	ACS 1-Year Estimates
Educational attainment	ACS 1-Year Estimates
Field of bachelor's degree	ACS 1-Year Estimates
Median income by industry	ACS 1-Year Estimates
Work destination	LEHD OnTheMap
Distance to work	LEHD OnTheMap

To obtain the desired data, use the data.census.gov "Advanced Search" option, as follows:

- Go to the data.census.gov Web site at <https://data.census.gov>.
- Select "Advanced Search" below the search bar (see Figure 3.1).

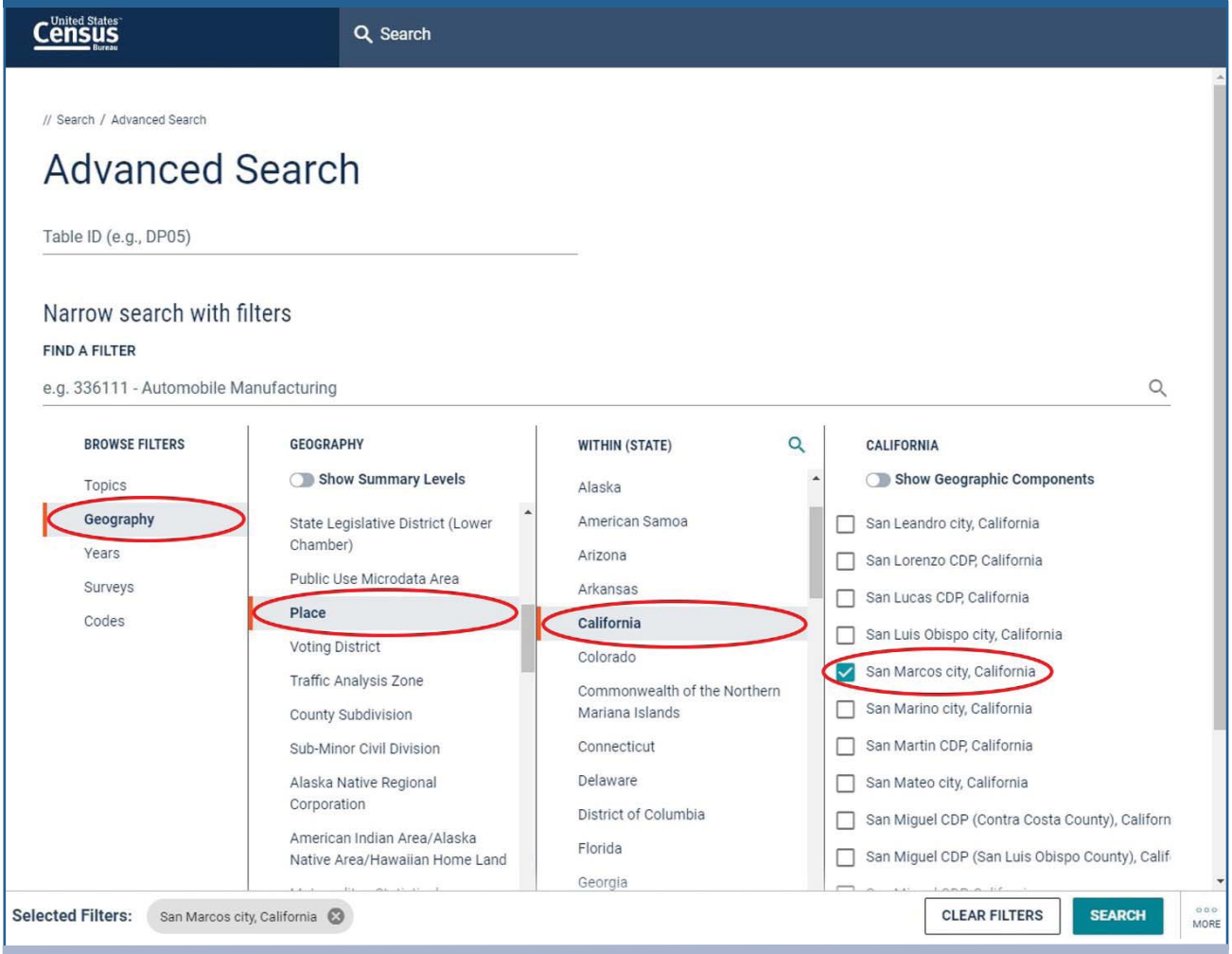
Figure 3.1. **Selecting Advanced Search in Data.census.gov**



Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov>.

- Select “Geography” in the navigation pane on the left side of the screen to display a list of available geographies (see Figure 3.2).
- Select “Place,” then “California.” From here, you can scroll to select “San Marcos city, California” or you can use the magnifying glass icon to search for places within California. This geography will appear in the “Selected Filters” at the bottom of the page.

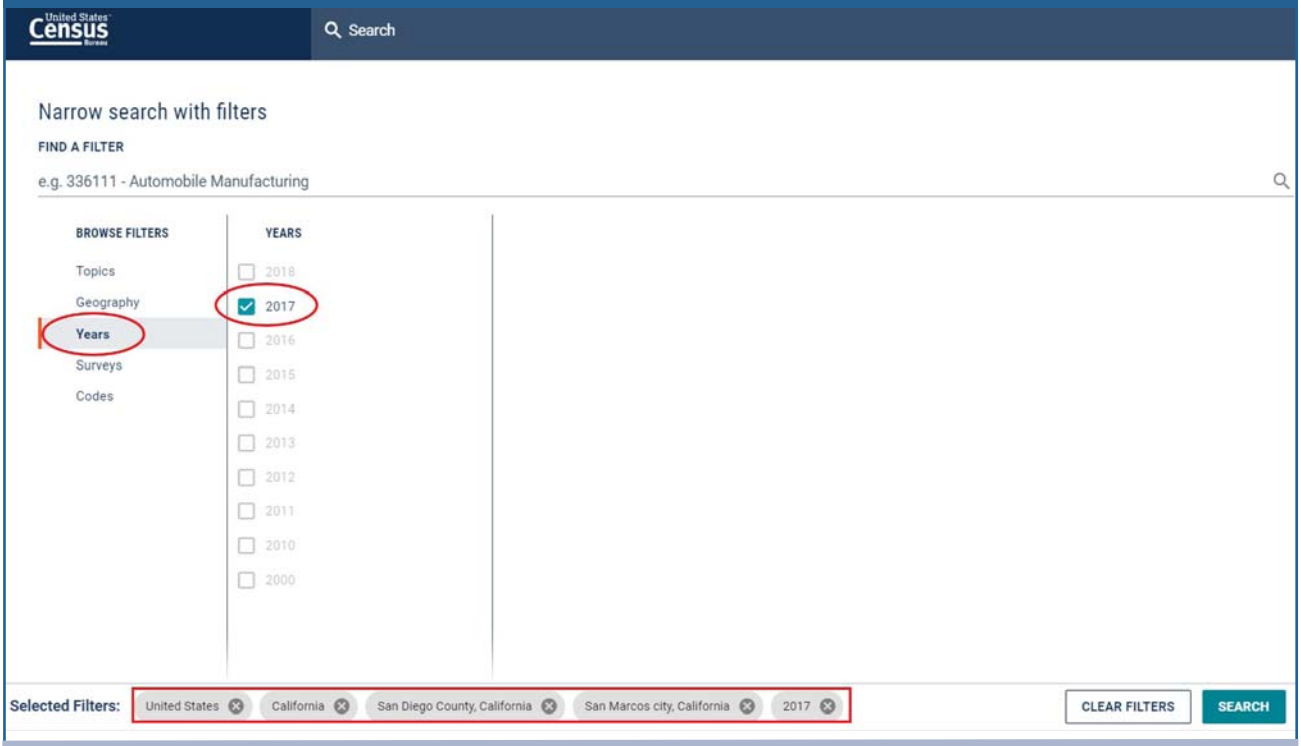
Figure 3.2. **Selecting Geographic Areas in Data.census.gov**



Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov>.

- Use the Geography filter to add the United States, California, and San Diego County to your selections. At the end of this step, you should have four geographic areas specified in your "Selected Filters" (see Figure 3.3).
- Then click on the "Years" filter and select "2017." The year "2017" should appear next to the previously selected geographic areas in "Selected Filters." For the purposes of this case study, we used 2017 because they were the most recent data available at the time.)

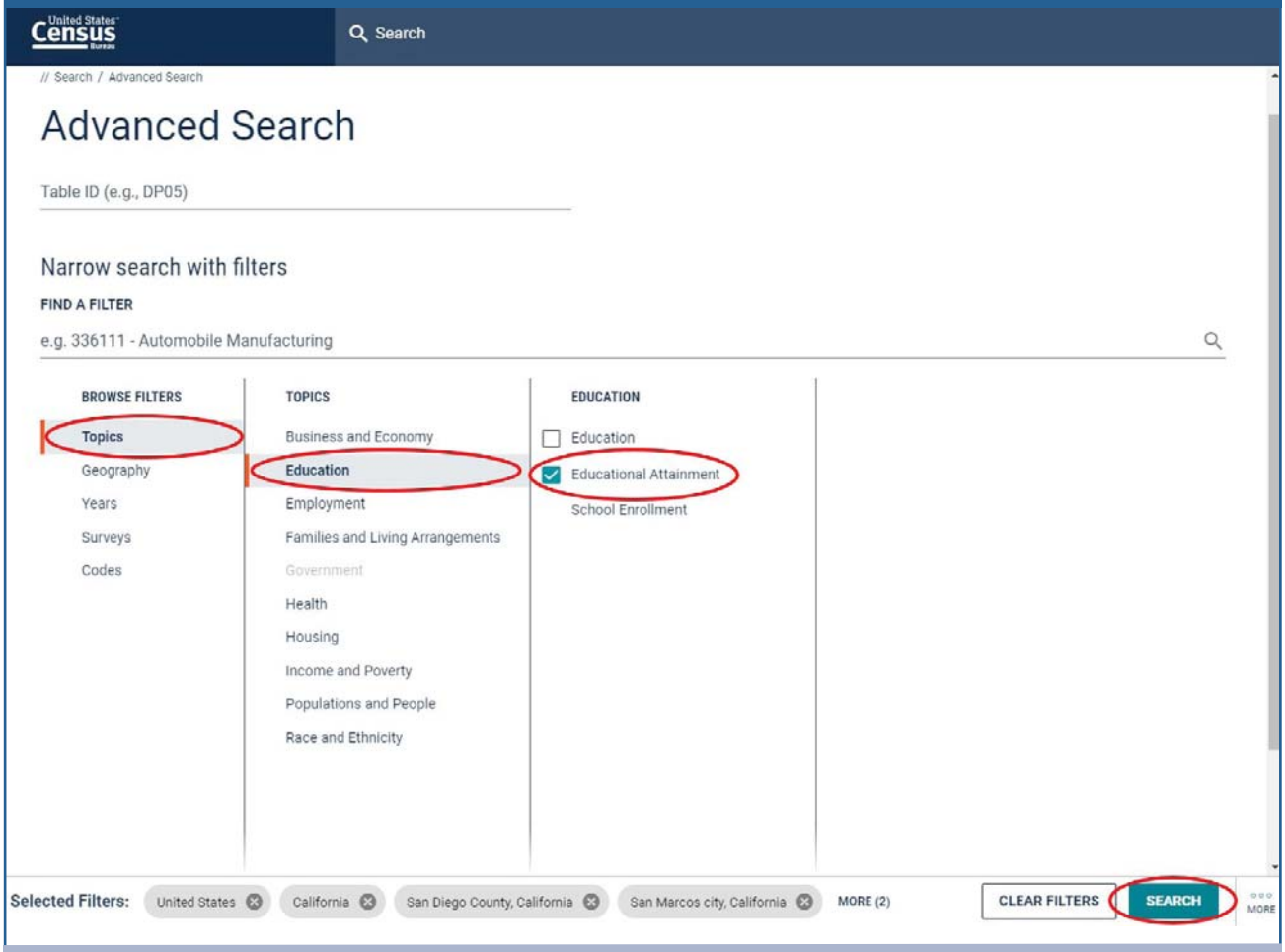
Figure 3.3. **Selecting Years and Verifying Selected Filters in Data.census.gov**



Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov>.

- Since we are interested in the city's socioeconomic characteristics, we begin by searching for information about educational attainment. Select "Topics," "Education," and "Educational Attainment." Then, click "Search" in the lower right corner (see Figure 3.4).

Figure 3.4. **Selecting the Educational Attainment Topic in Data.census.gov**

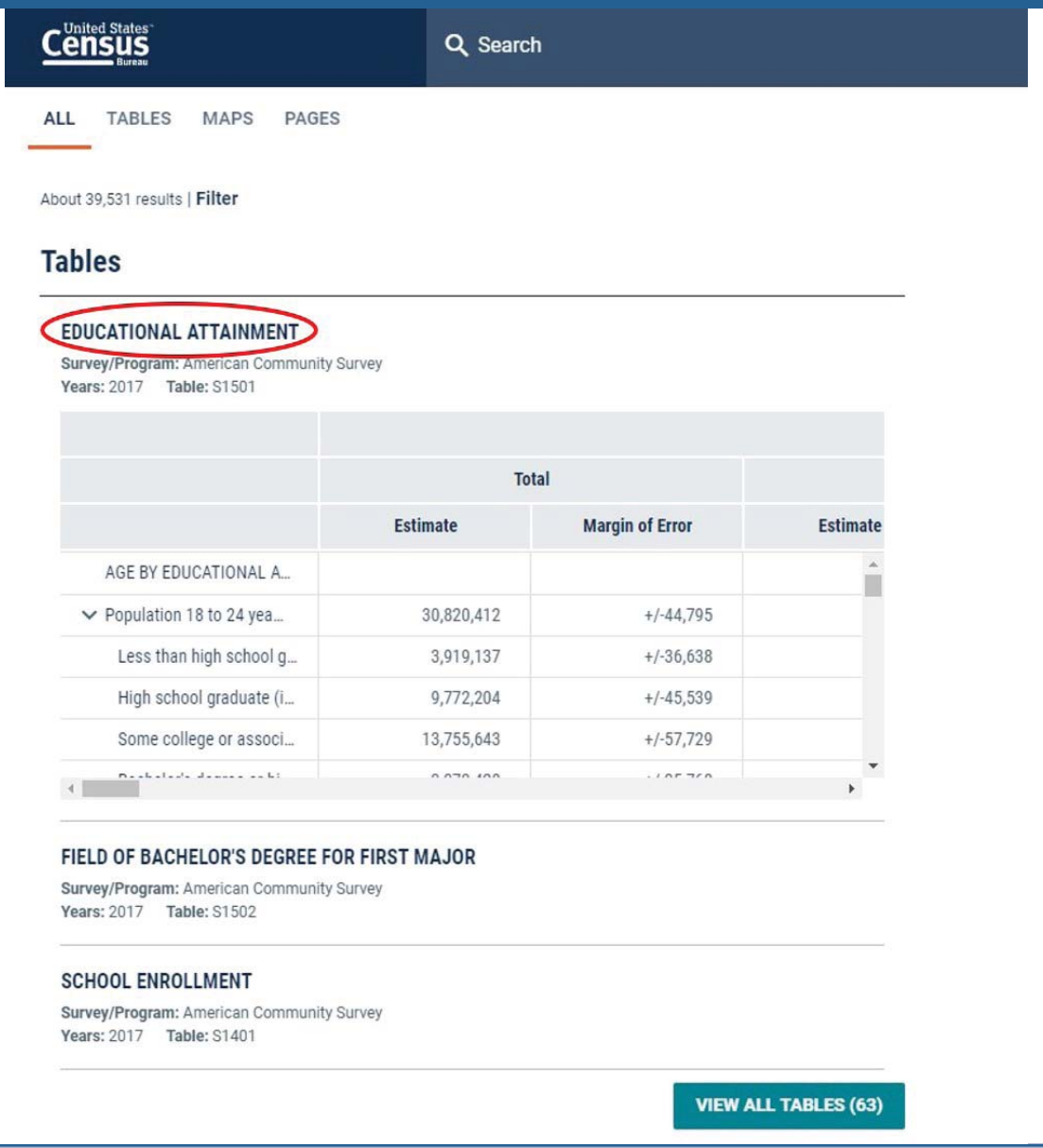


Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov>.

- Use the descriptive table titles to make your selection. We know we are searching for a summary table, so we select the first result, S1501: "Educational Attainment" (see Figure 3.5).[20]

[20] Tables beginning with "S" are Subject Tables that focus on particular ACS subjects and include both estimates and percentages. More information about Table IDs is available on the Census Bureau's Web site, <www.census.gov/programs-surveys/acs/guidance/which-data-tool/table-ids-explained.html>.

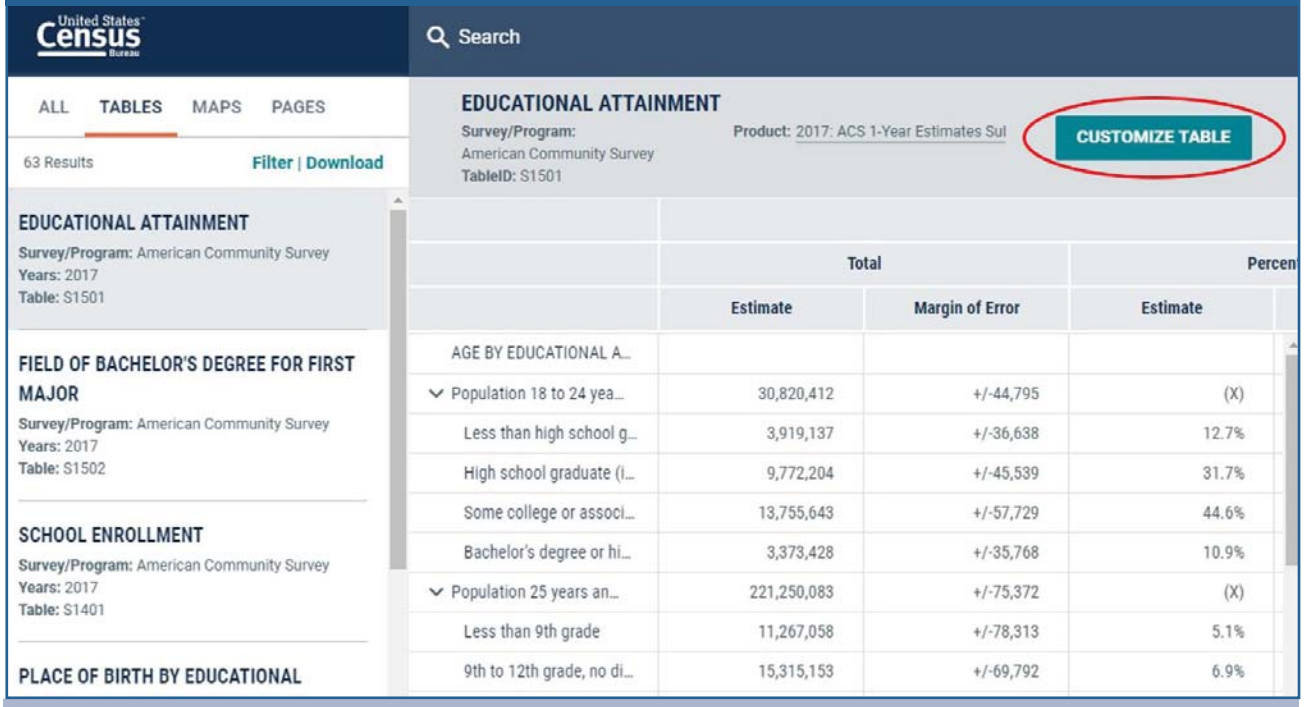
Figure 3.5. **Selecting an Educational Attainment Table in Data.census.gov**



Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov>.

- Next, select “Customize Table” (see Figure 3.6).

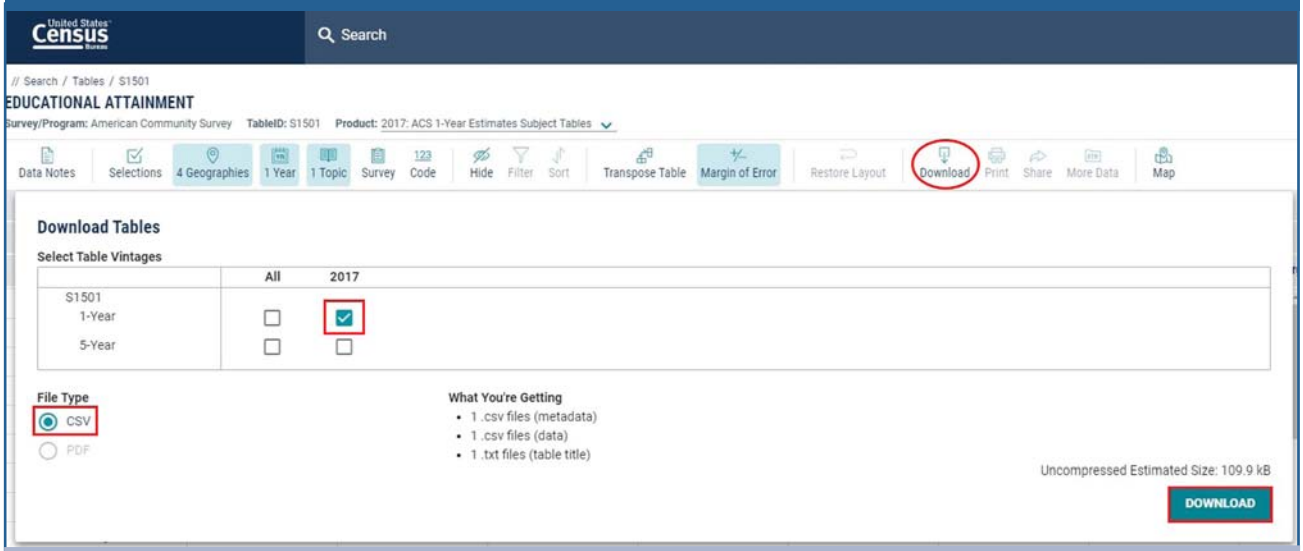
Figure 3.6. **Customizing a Table in Data.census.gov**



Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov>.

- Select “Download” at the top of the window. Then, use the Download Tables window to check the box for the 2017 ACS 1-year data. Select “CSV” as the file type and click on “Download” (see Figure 3.7).

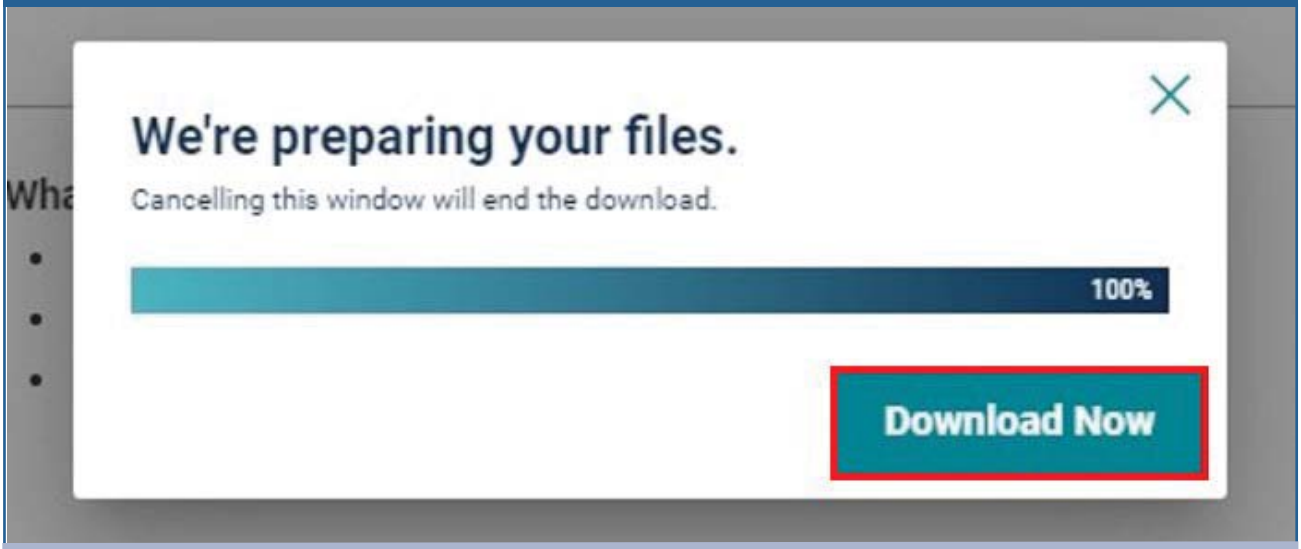
Figure 3.7. **Selecting the Survey Year(s) and File Type in Data.census.gov**



Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov>.

- After the files are prepared, click the “Download Now” button (see Figure 3.8).

Figure 3.8. **Downloading Tables in Data.census.gov**



Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov>.

- This produces a ZIP file. Open the ZIP file to extract your table. The data are located in a file with prefix “ACSST1Y2017.S1501_data_with_overlays” (see Figure 3.9).
- We repeat this process for all data.census.gov topics in the profile.

Figure 3.9. **Portion of Data Table S1501 Downloaded From Data.census.gov**

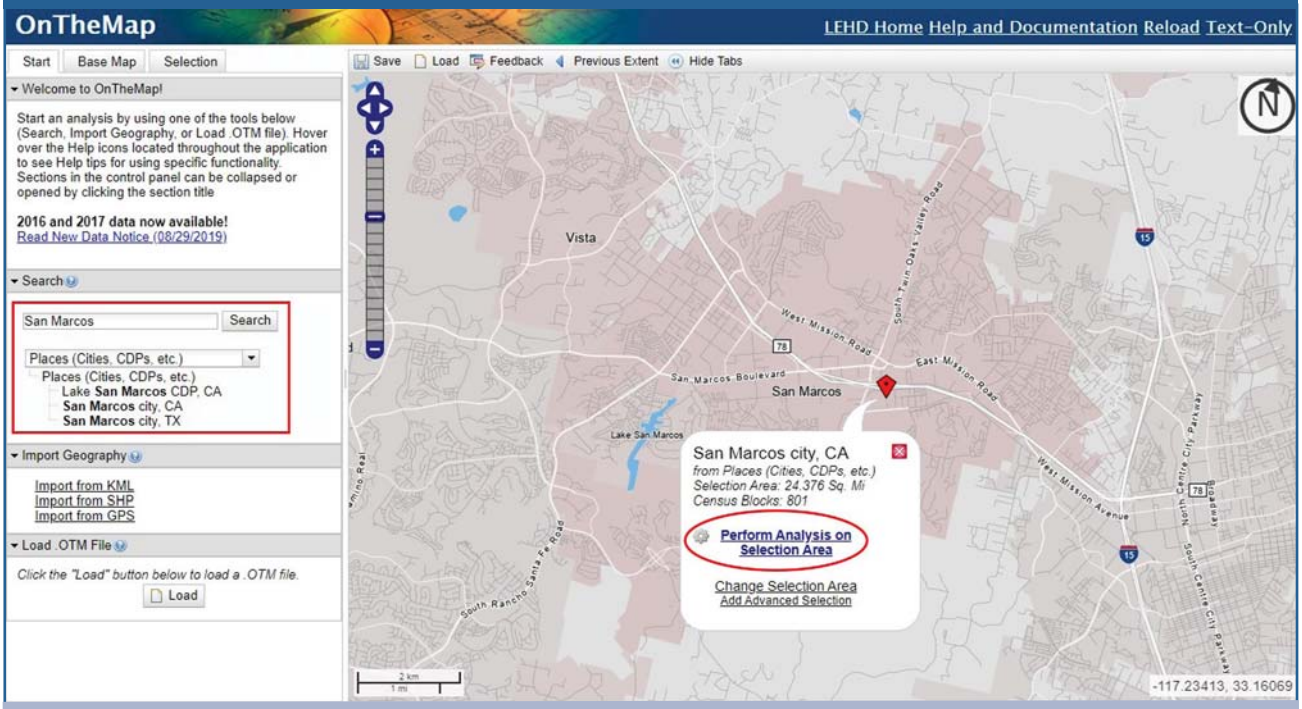
	A	B	C	D
1	GEO_ID	NAME	S1501_C01_001E	S1501_C01_001M
2	id	Geographic Area Name	Total!!Estimate!!Population 18 to 24 years	Total!!Margin of Error!!Population 18 to 24 years
3	0100000US	United States	30820412	44795
4	0400000US06	California	3799994	5436
5	0500000US06073	San Diego County, California	346417	273
6	1600000US0668196	San Marcos city, California	8528	1958
7				
8				

Source: U.S. Census Bureau, data.census.gov, Table S1501: “Educational Attainment,” <https://data.census.gov>.

For data on commuting patterns, we access LEHD data, as follows:

- Go to the LEHD Web site at <https://lehd.ces.census.gov/research/>.
- Click "OnTheMap" on the left sidebar. This action will open a new window.
- Under the "Search" box, type "San Marcos" in the search bar, using the dropdown menu to select "Places" as the geography type. Then click "Search." Be sure to select San Marcos, California, rather than San Marcos, Texas. Our selection refreshes the map view and zooms in to San Marcos (see Figure 3.10).
- Use the information bubble next to the location pin to select "Perform Analysis on Selection Area."

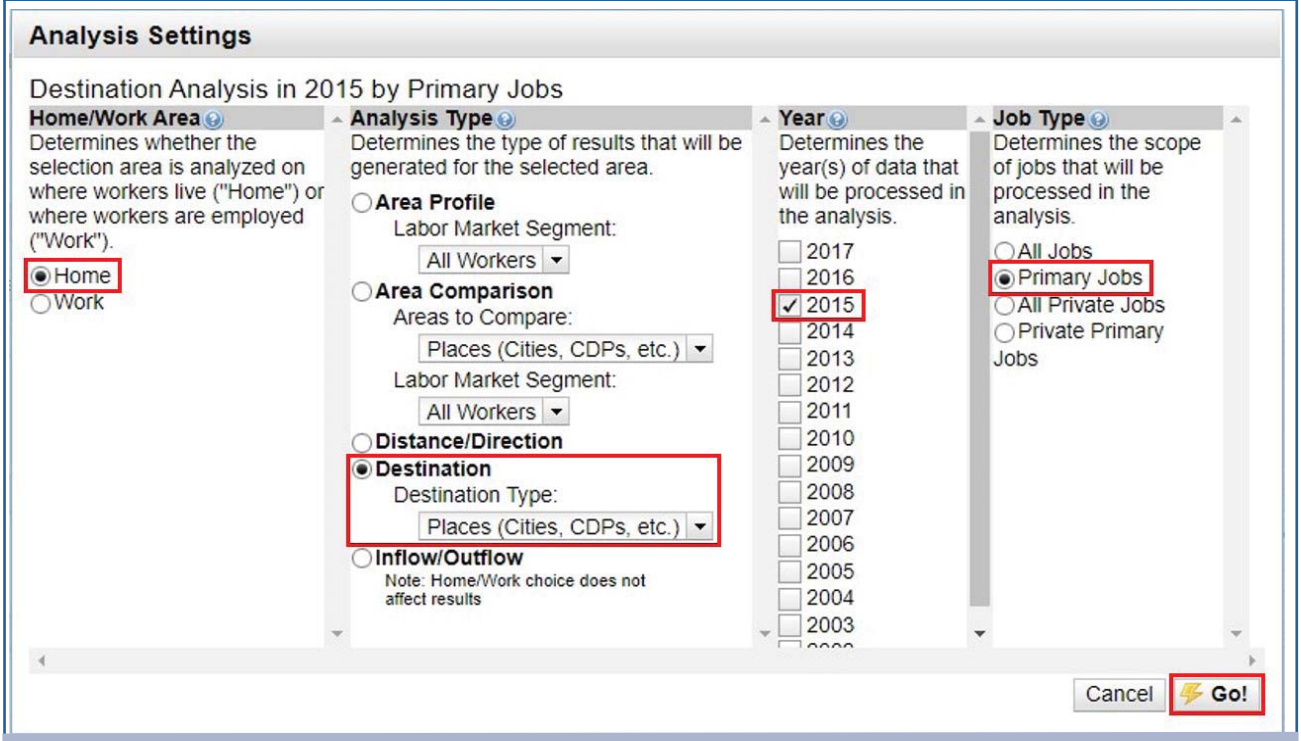
Figure 3.10. **Searching for a Geographic Area in OnTheMap**



Source: U.S. Census Bureau, OnTheMap, <https://onthemap.ces.census.gov/>.

- The “Analysis Settings” window opens, allowing us to refine our search.
    - Home/Work Area: Since we are interested in understanding commuting patterns, we select “Home.” This yields data based on San Marcos’ residents, rather than workers.
    - Analysis Type: We are interested in understanding where residents work. We select “Destination” and “Places” as our destination type.
    - Year: We select the most recently available data.
    - Job Type: Select “Primary Jobs.”
- Once all selections have been made, we click on “Go!” (see Figure 3.11).

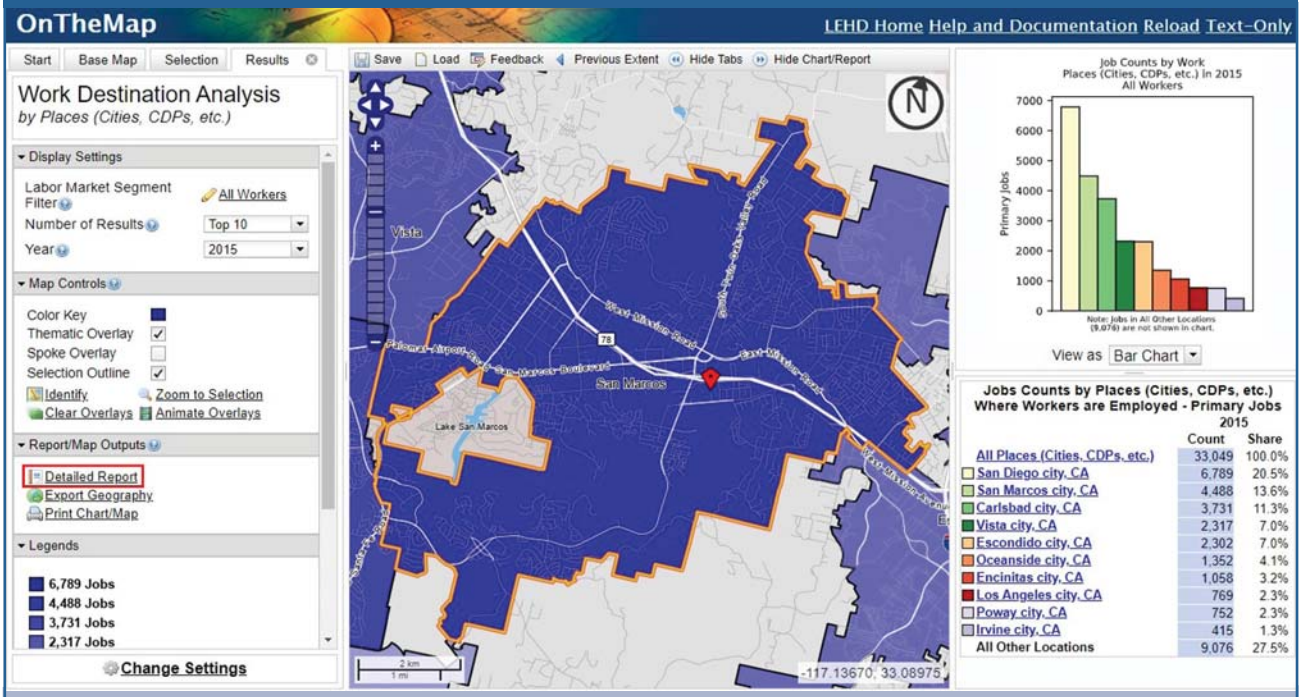
Figure 3.11. **Selecting Variables for Analysis in OnTheMap**



Source: U.S. Census Bureau, OnTheMap, <https://onthemap.ces.census.gov/>.

- The page refreshes and populates our search results. We export the results by selecting "Detailed Report" under the left sidebar's "Report/Map Outputs" bin and export the data in spreadsheet format (see Figure 3.12).

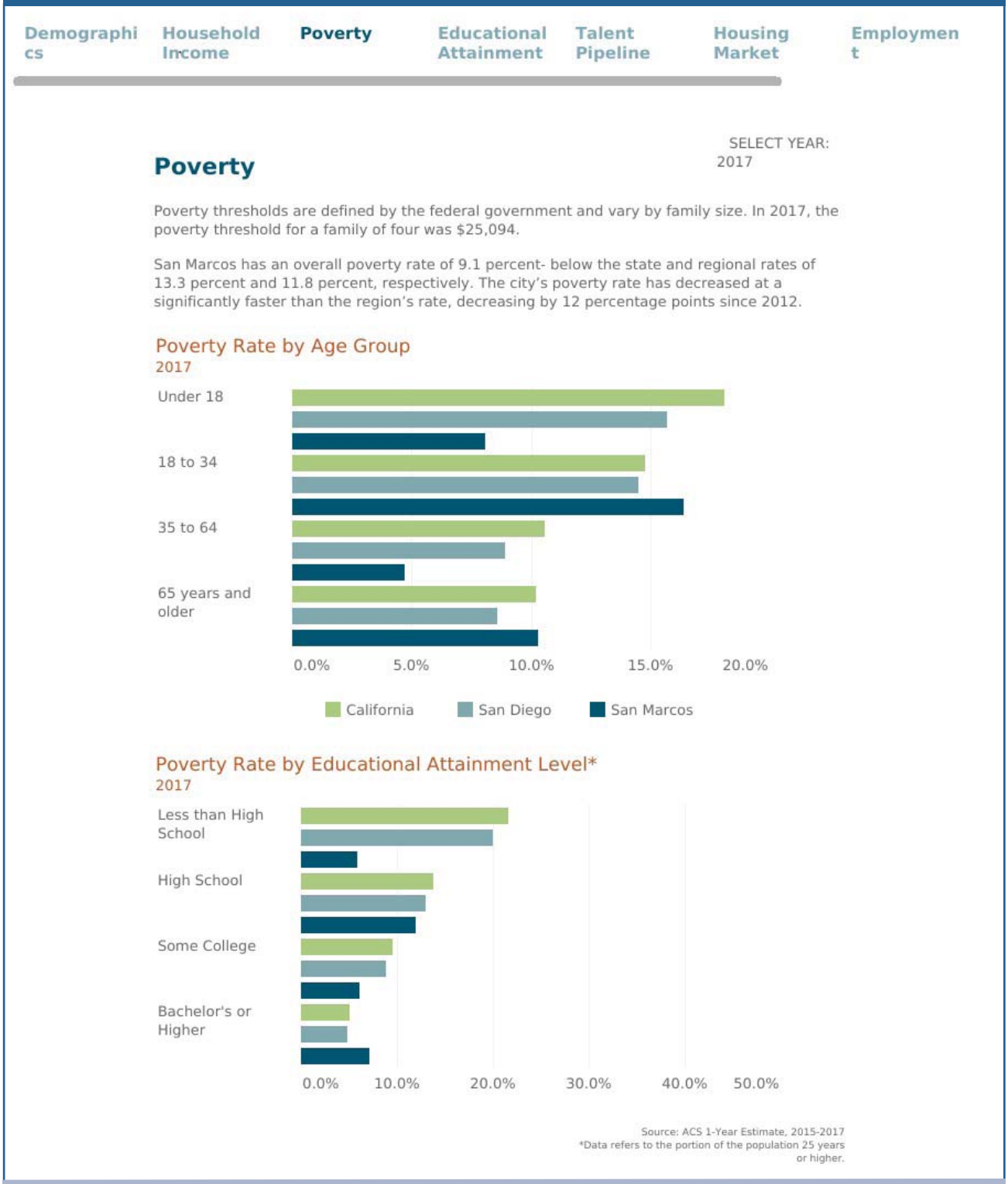
Figure 3.12. **Creating a Detailed Report in OnTheMap**



Source: U.S. Census Bureau, OnTheMap, <https://onthemap.ces.census.gov/>.

Initial analysis was conducted in a spreadsheet, by sorting results from highest to lowest and comparing the city of San Marcos to the county, state, and nation. The resulting charts were created using a mixture of spreadsheets and visualization software. Figure 3.13 illustrates how poverty data were visualized in the dashboard.

Figure 3.13. **City of San Marcos Community Profile: Poverty**



Source: San Marcos City Hall, San Marcos Regional Profile, <www.san-marcos.net/work/economic-development/community-profile>.

Figure 3.14 shows an interactive flow map that was created to visualize the city's commuting patterns.

The City of San Marcos Community Profile and interactive dashboard were launched in the fall of 2017. Since then, the dashboard has been actively used by the city and its partners to better inform the region's business community. In partnership with the city of San Marcos, the EDC plans on updating the profile and dashboard on an annual basis as data become available. Both products can be found on the city of San Marcos Web site.[21]

[21] San Marcos City Hall, San Marcos Community Profile, <www.san-marcos.net/work/economic-development/community-profile>.

Figure 3.14. **City of San Marcos Regional Profile: Commuting Patterns**

Source: San Marcos City Hall, San Marcos Community Profile, <www.san-marcos.net/work/economic-development/community-profile>.

Case Study #2: Housing Affordability Gap

Skill Level: Intermediate
Subject: Median household income, housing affordability
Type of Analysis: Analysis of national housing trends, combining American Community Survey (ACS) data with other sources
Tools Used: ACS Public Use Microdata Sample, data.census.gov, and spreadsheet
Authors: Svenja Gudell, Chief Economist, Zillow; and Aaron Terrazas, Senior Economist, Zillow

In the years since the 2007–2009 Financial Crisis, there has been a boom in housing market data available to the public. Each month, Zillow—an online resource for home buyers and renters—produces and publishes over 100 housing market data series—including median home values and median rents—aggregated at the national, state, metro, county, ZIP code, and neighborhood levels.[22]

Zillow data provide a comprehensive view into the state of the U.S. housing stock, but Zillow's economic research and data science teams also rely on critical data inputs from the U.S. Census Bureau to analyze the full implications of changing home values and rents for American households. One prominent example is how Zillow combines proprietary data on home values with household income data from the Census Bureau's American Community Survey (ACS) to explore the share of income that the typical household spends on a monthly mortgage payment. (Zillow also produces a similar statistic for the share of income spent on rent.)

By the middle of 2016—the most recent data available at the time of this writing—the typical American household earning the country's national median income and buying the median-valued U.S. home could expect to pay 15.5 percent of their income on a mortgage. In the nation's priciest markets—such as San Jose, Los Angeles, and San Francisco—these shares exceed 40 percent.

However, calculating housing affordability estimates based on median income and median home values by themselves does not provide a complete picture of affordability. Home values and incomes have consistently grown more at the top of the housing market and at the top of the labor market than at the bottom. Increasingly, affordability varies dramatically depending on a given buyer's income level and the type of home they are trying to buy.

To illustrate these diverging trends between more affluent and less affluent Americans, Zillow also computes mortgage affordability—the share of income spent on a mortgage—for the bottom, middle, and top one-third of households by income. This more granular (and, some would argue, more powerful) tier analysis is conducted by combining Zillow's property-level data with ACS microdata.

Calculating Mortgage Affordability

To calculate mortgage affordability, Zillow first estimates the mortgage payment for the median-valued home in a metropolitan statistical area (based on the Zillow Home Value Index for a given quarter) and the 30-year, fixed mortgage interest rate during that quarter (from the Freddie Mac Primary Mortgage Market Survey), assuming a 20 percent down payment. Zillow uses their internal property-level data to estimate mortgage payments, but data users can also access housing data—including home value and mortgage payment estimates—from the ACS.

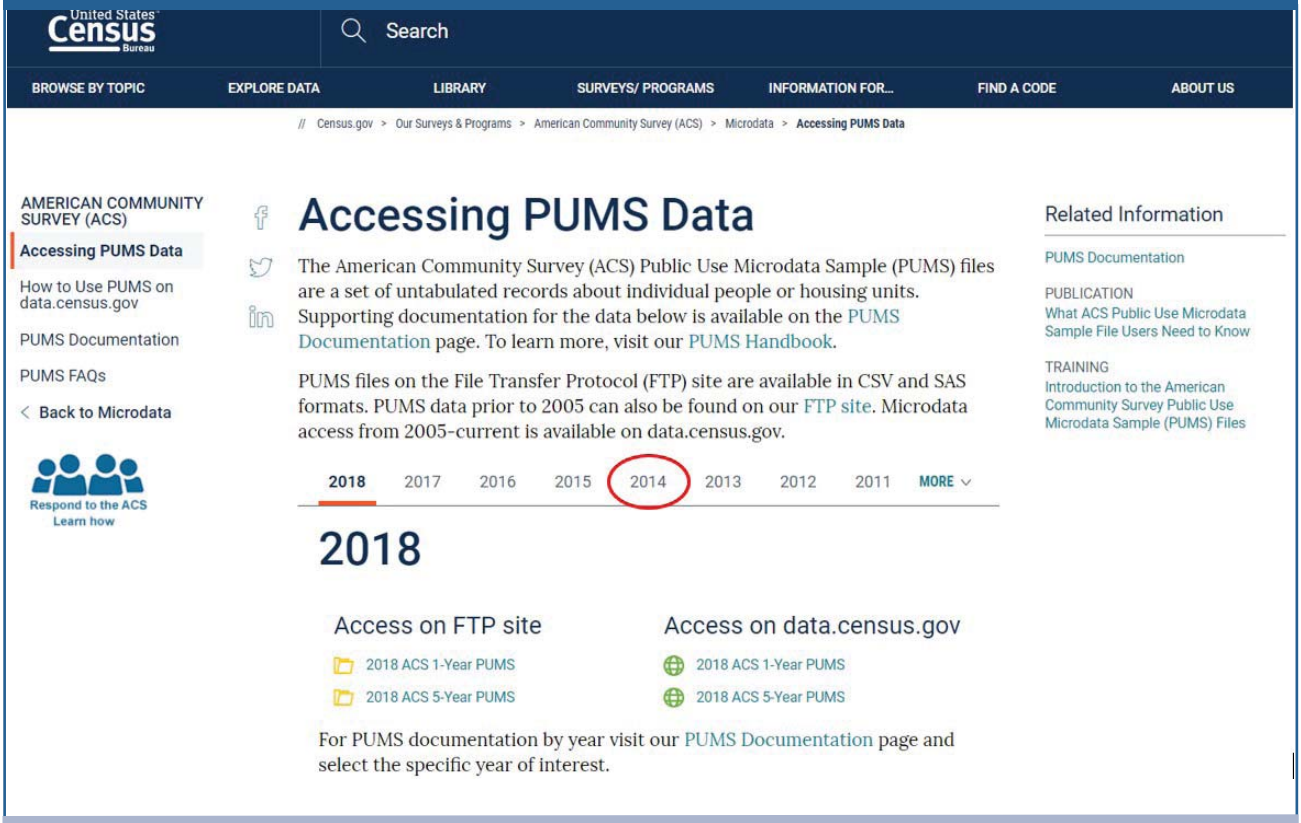
Data on median household income are from the Census Bureau's ACS Public Use Microdata Sample (PUMS). The PUMS files allow data users to conduct a custom analysis of the ACS data using a sample of actual responses to the survey. They are much more flexible than the aggregate data available on data.census.gov, though the PUMS files also require familiarity with statistical analysis software.

With access to appropriate software—the most common are SAS, SPSS, STATA, R, and Python—using the ACS PUMS data is straightforward. Analysts without statistical software can create detailed cross tabulations using the microdata access tool on data.census.gov.[23] The Census Bureau created a step-by-step guide on how to use this tool to produce custom estimates from the ACS 1-year PUMS file.[24]

[22] Data are published at <www.zillow.com/data>.
[23] This tool is still under development and is available in beta form on the Census Bureau's Web site: <https://data.census.gov/mdat/>.
[24] U.S. Census Bureau, Using Microdata Access, <https://www2.census.gov/data/api-documentation/using-microdata-access/microdata-access-1-year-acs-pums.pdf>.

To download PUMS data, go to the Census Bureau's ACS PUMS Data "Accessing PUMS Data" Web page and select the desired timeframe for the PUMS data (see Figure 3.15). Data are available from the year 2005 to the present. You can select single-year or multiyear data. The example below uses data from the 2014 ACS 1-year file.

Figure 3.15. **ACS PUMS Data Repository**



Source: U.S. Census Bureau, American Community Survey (ACS), Accessing PUMS Data, <www.census.gov/programs-surveys/acs/microdata/access.html>.

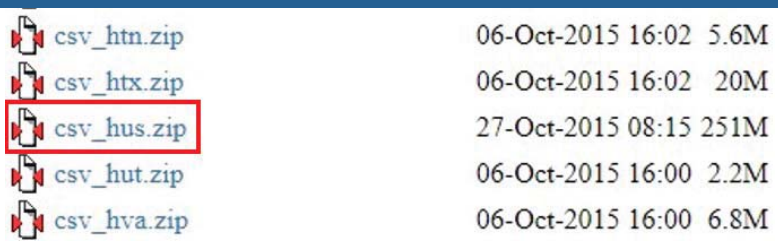
On the FTP site, click on "1-Year" to access the 2014 ACS 1-year PUMS files (see Figure 3.16).

Figure 3.16. **Selecting ACS PUMS File on the Census Bureau's FTP Server**

Source: U.S. Census Bureau, American Community Survey (ACS), Accessing PUMS Data, <www.census.gov/programs-surveys/acs/microdata/access.2014.html>.

The next step is to download the data set by selecting "csv_hus.zip" (see Figure 3.17). PUMS files on the Census Bureau's FTP site are stored as ZIP files. The naming convention for PUMS files on the FTP server is based on three file features: the file format, the record type, and the state abbreviation.

Figure 3.17. **ACS PUMS Downloads by Geography (Nation and States) and File (Population and Housing)**



Source: U.S. Census Bureau, FTP server, <https://www2.census.gov/programs-surveys/acs/data/pums/2014/1-Year/>.

- File formats are comma-separated value (CSV) files and SAS data sets for UNIX.
- Record types are housing files (h) or person files (p). Zillow uses the housing file.
- State (or state equivalent) abbreviations are two letter labels such as "tx" for Texas and "dc" for District of Columbia. The abbreviation for the file containing all records in the United States is "us." Zillow uses the United States file.

After downloading the file, Zillow loads the ACS data into R (software for statistical computing) to produce estimates of monthly owner costs as a percentage of household income for households at different income levels. The PUMS Data Dictionary can help you find the variable needed for this analysis: HINCP (Household income in the past 12 months).[25]

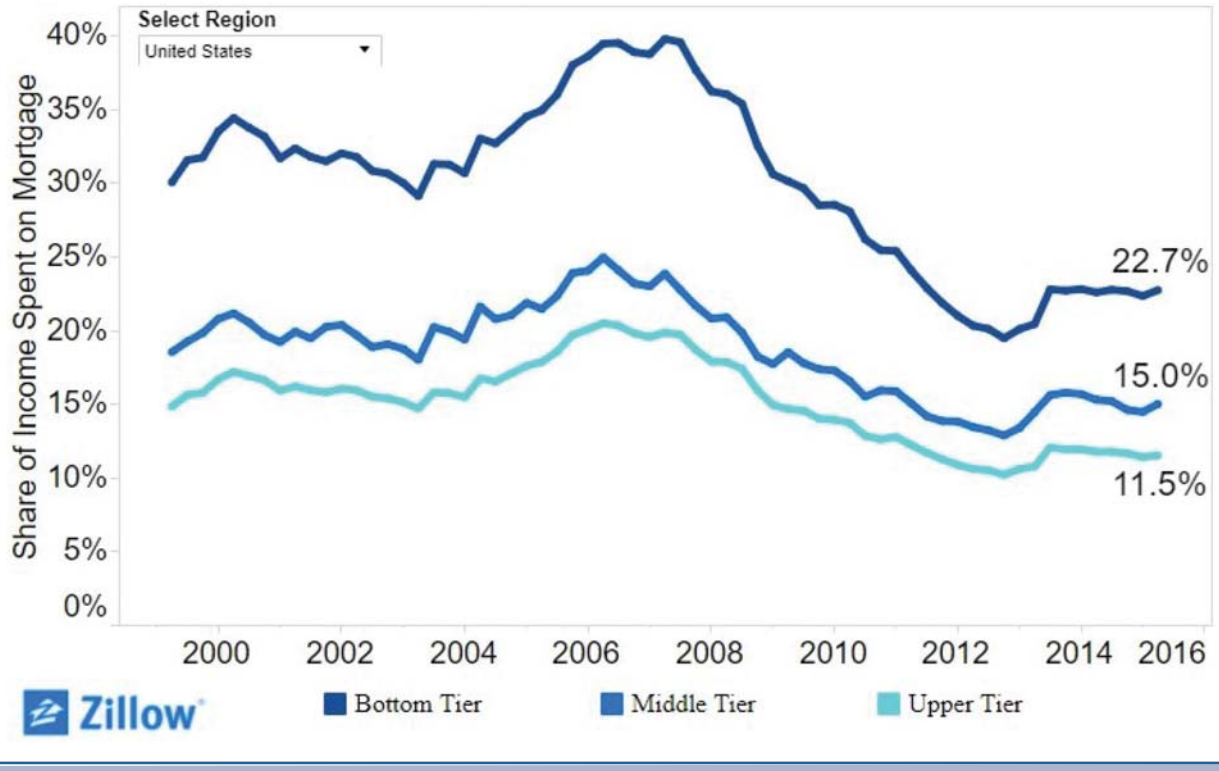
Using the HINCP variable to estimate median household income for households in the bottom one-third of the income distribution, Zillow then assumes that these households purchase a median bottom-tier home (i.e., 16.7th percentile of homes in a metro area) and then calculates the percentage of monthly income that household would have to spend to pay the mortgage.[26]

[25] U.S. Census Bureau, PUMS Data Dictionary, <www.census.gov/programs-surveys/acs/microdata/documentation.html>.

[26] For median income used in the overall affordability analysis, Zillow chains the ACS income data forward using the Employment Cost Index (ECI), which is updated quarterly with a one-quarter lag. Income tier data are not directly published in the ECI, so to calculate tiers, Zillow relies on the Consumer Expenditure Survey (CES), which is published with a 1-year lag. For this reason, tier data are only available through Q2 2016, but overall affordability data are available through Q2 2017.

The key results from Zillow's analysis are shown in Figure 3.18. Depending on a given buyer's household income level and the kind of home they are trying to buy, affordability can vary dramatically. Nationwide, a buyer at the median household income in the bottom one-third of all incomes and who wants to buy a home valued in the bottom one-third of all homes would need to spend 23 percent of their household income on a mortgage as of Q2 2016, the latest quarter for which data were available at the time of this writing. A buyer in the top one-third of household incomes and who wants to buy a more expensive home in the top one-third of homes by value would only spend 11.5 percent of their income on a mortgage.

Figure 3.18. **Mortgage Affordability by Income Tier**



Source: Svenja Gudell, "Housing Highs & Lows: How the Home Affordability Gap Between the Rich and Poor is Widening," <www.zillow.com/research/affordability-2016q1-12763/>.

A decade after the 2007–2009 Financial Crisis and associated housing market bust, there has been a proliferation of housing market data available to consumers and researchers. Census Bureau data, particularly from the ACS, are a critical complement to Zillow's housing market data.

In an era when the gap in net worth between higher- and lower-wealth households is increasing, medians or averages cannot tell the full story of the American housing market.[27] Exploring within the distribution of home values and household incomes—analysis that is possible with Zillow's property-level home values and the Census Bureau's ACS PUMS data—allows for a much richer and more complete perspective on what is, for most Americans, their largest single asset.[28]

[27] U.S. Census Bureau, Gap Between Higher- and Lower-Wealth Households Widens, Census Bureau Reports, <www.census.gov/newsroom/press-releases/2014/cb14-156.html>.

[28] Zillow updates this analysis periodically. For the most recent data and analysis, see <www.zillow.com/research>.

Case Study #3: USAA Potential Market Size Estimation

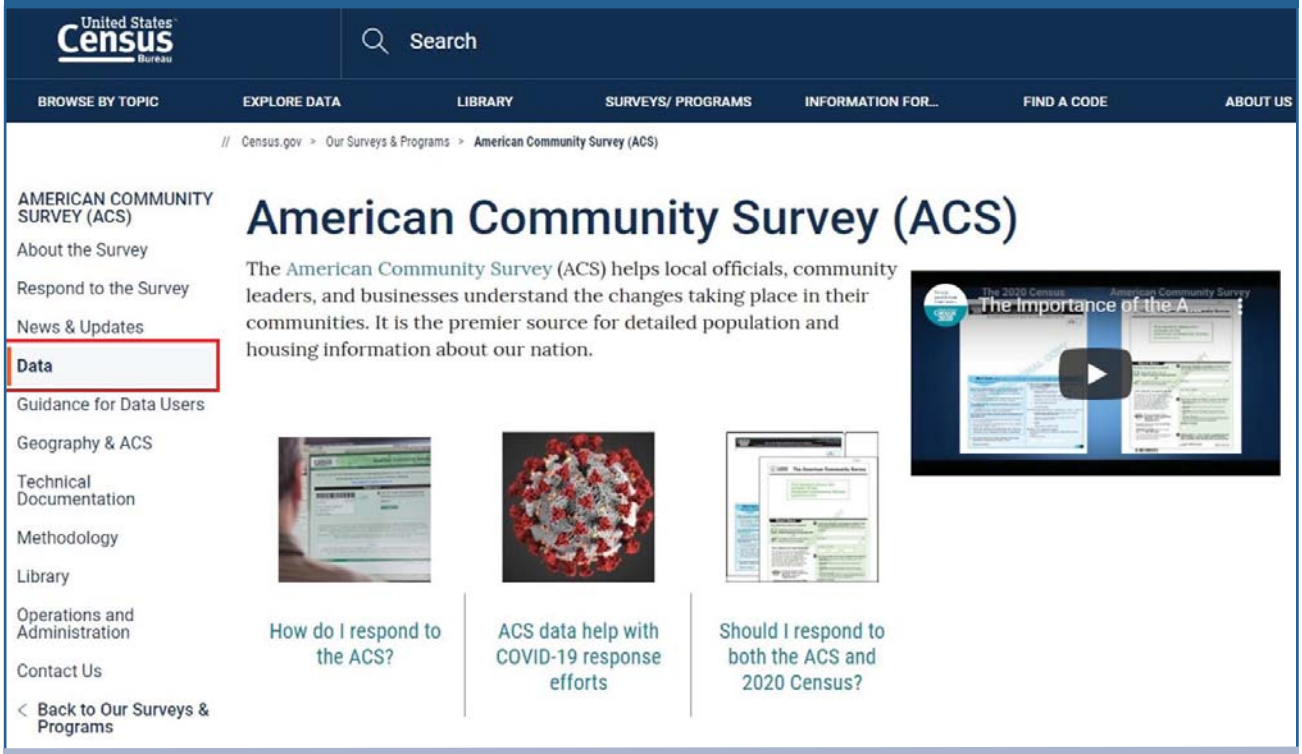
Skill Level: Introductory/Intermediate
Subject: Housing
Type of Analysis: Comparisons of American Community Survey (ACS) data over time, across demographic groups, and across states; identifying business market(s)
Tools Used: Data.census.gov and spreadsheet
Author: Rob Galbraith, Director of Underwriting Research, Property & Casualty Insurance Group, USAA

Rob Galbraith is a director of underwriting research with the Property & Casualty Insurance Company at USAA, a large financial services member-owned association that focuses on serving active duty military, veterans, and their families. The underwriting area is challenged with providing insurance to as many eligible members as possible while not taking on excessive exposure to natural hazards that could put the company at risk of bankruptcy. American Community Survey (ACS) data are used at USAA in conjunction with advanced analytics and predictive modeling to identify segments to target for marketing three of its major lines of business: homeowners insurance, sold to those with owner-occupied homes; renters insurance, sold to those who are renting a dwelling from a landlord who owns the property; and rental property insurance, sold to those who are landlords renting their property to one or more tenants. ACS data are also used to track trends in owner- and renter-occupied housing to target marketing expenditures to achieve profitable growth.

Rob scans the documentation available for the ACS and learns that while single-year data are available for states and many large geographic areas, data for smaller areas with fewer than 65,000 people are only available in 5-year tabulations (e.g., 2012–2016). He also learns that there is uncertainty—margin of error—associated with ACS estimates, particularly for smaller geographic areas and small population groups. For this example, Rob is interested in accessing ACS housing data at the state level.

- Rob navigates to the U.S. Census Bureau's ACS home page at <www.census.gov/programs-surveys/acs> and clicks on the "Data" link on the left side of the page (see Figure 3.19).

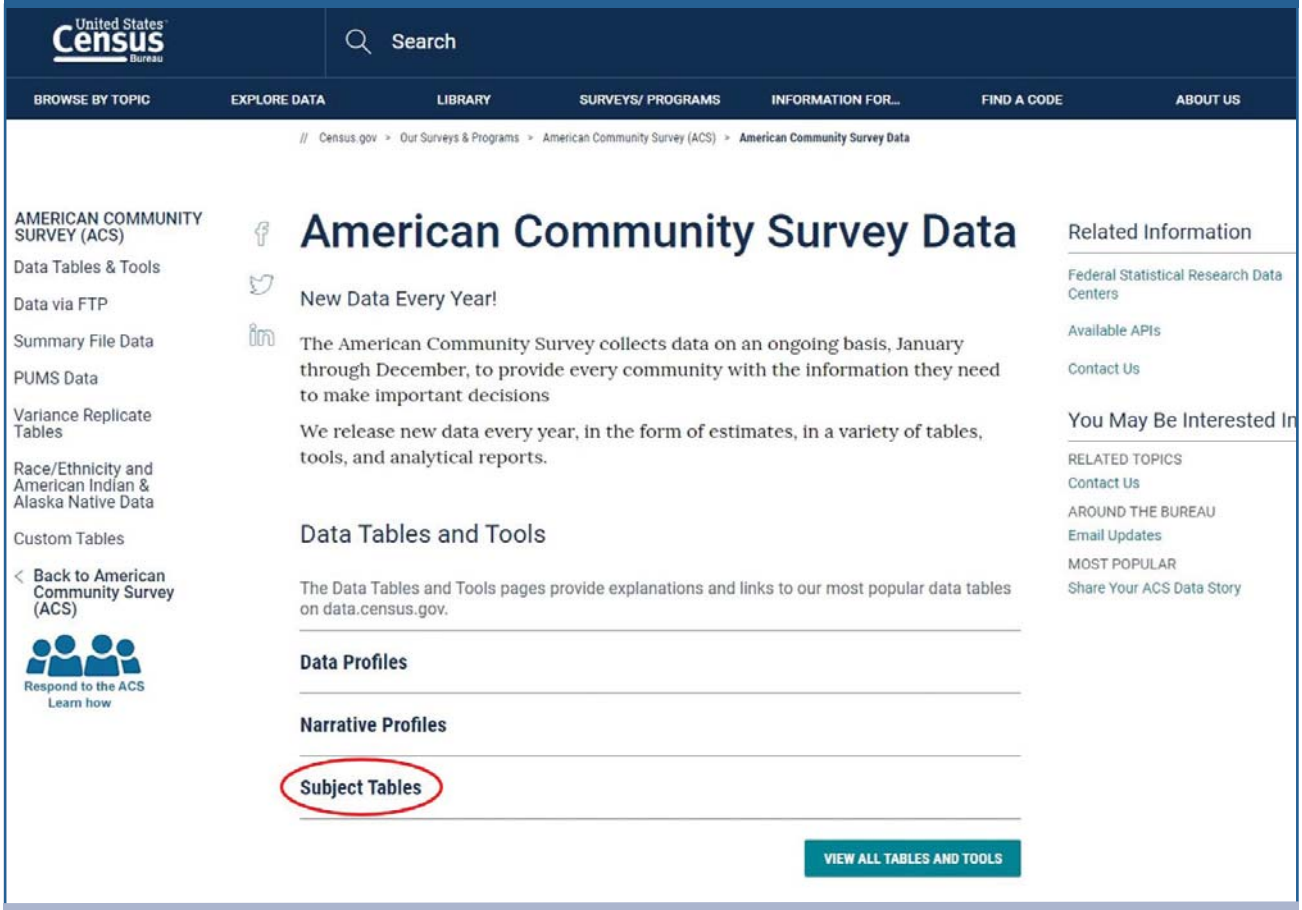
Figure 3.19 **Selecting Data From the ACS Home Page**



Source: U.S. Census Bureau, American Community Survey (ACS), <www.census.gov/programs-surveys/acs/>.

- Under the subheading “Data Tables and Tools,” he clicks on “Subject Tables” (see Figure 3.20).

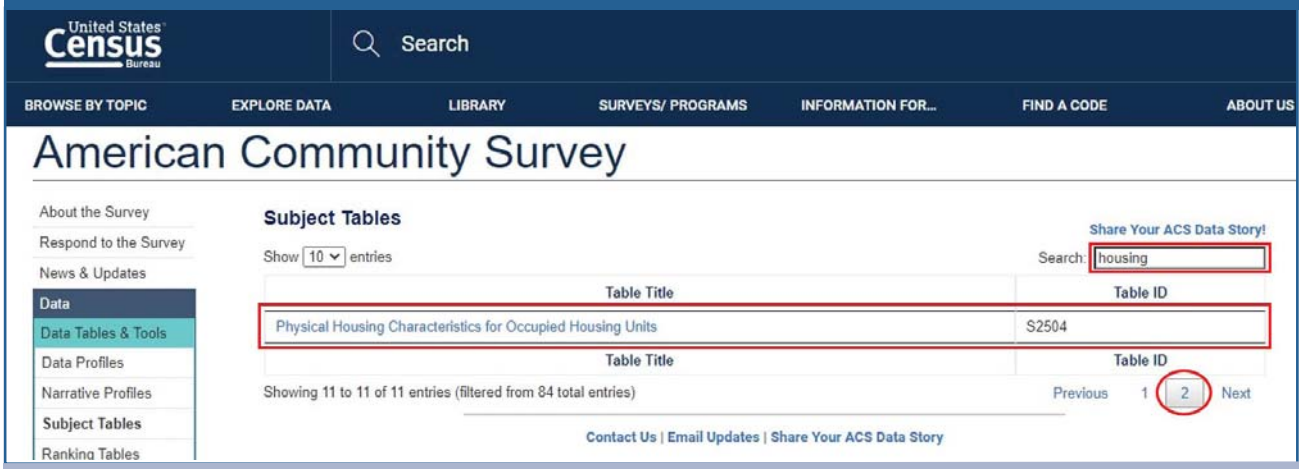
Figure 3.20. **Selecting ACS Subject Tables**



Source: U.S. Census Bureau, American Community Survey (ACS), <www.census.gov/acs/www/data/data-tables-and-tools/subject-tables/>.

- In the search box near the top right corner of the page, he types in "housing," navigates to the second page of results, and then selects Table S2504: "Physical Housing Characteristics for Occupied Housing Units" (see Figure 3.21). When he clicks on this link, he is redirected to data.census.gov, the Census Bureau's primary tool for accessing data from the ACS and many other Census Bureau data sets.

Figure 3.21. **Searching for ACS Housing Tables**



Source: U.S. Census Bureau, American Community Survey (ACS), <www.census.gov/acs/www/data/data-tables-and-tools/subject-tables/>.

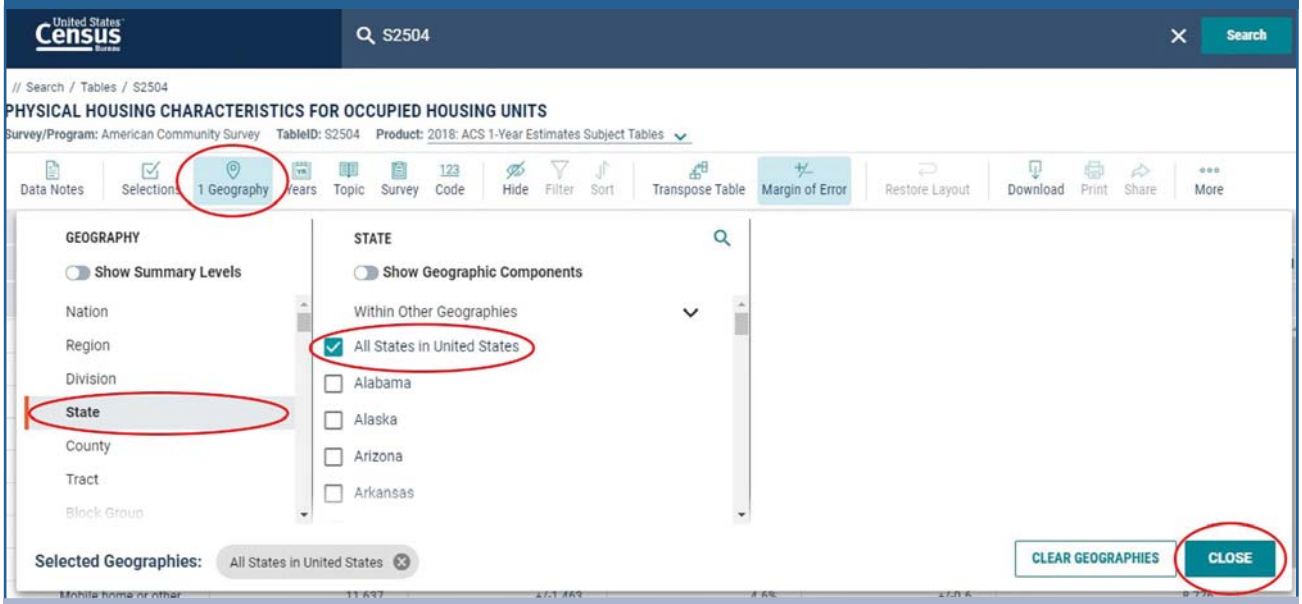
- Next, he clicks on "Customize Table" (see Figure 3.22).
- The resulting table shows the United States as the default geography, so he selects "Geographies" to change the geography filter.

Figure 3.22. **Customizing a Table in Data.census.gov**

Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov>.

- Next, he selects "State," "All States in United States," and then closes the window to view the results (see Figure 3.23).

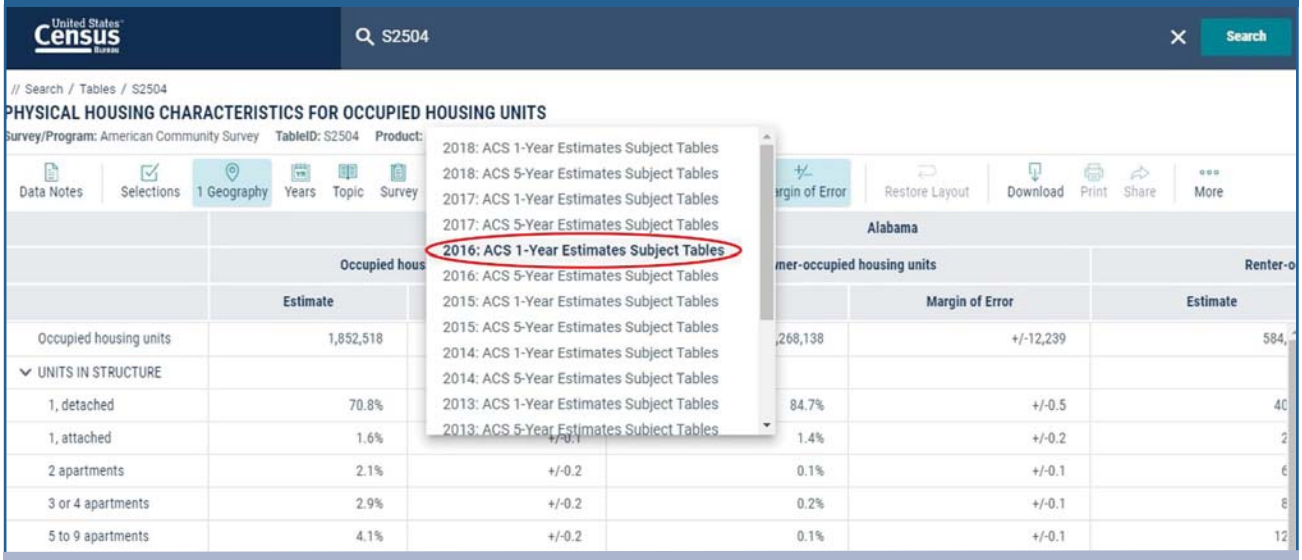
Figure 3.23. **Selecting Geographies in Data.census.gov**



Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov>.

- He selects the desired survey year by clicking on the current "Product" selection. For the purposes of this case study, he is using 2016 ACS 1-year estimates. The header should read "2016 ACS 1-Year Estimates Subject Tables" (see Figure 3.24).

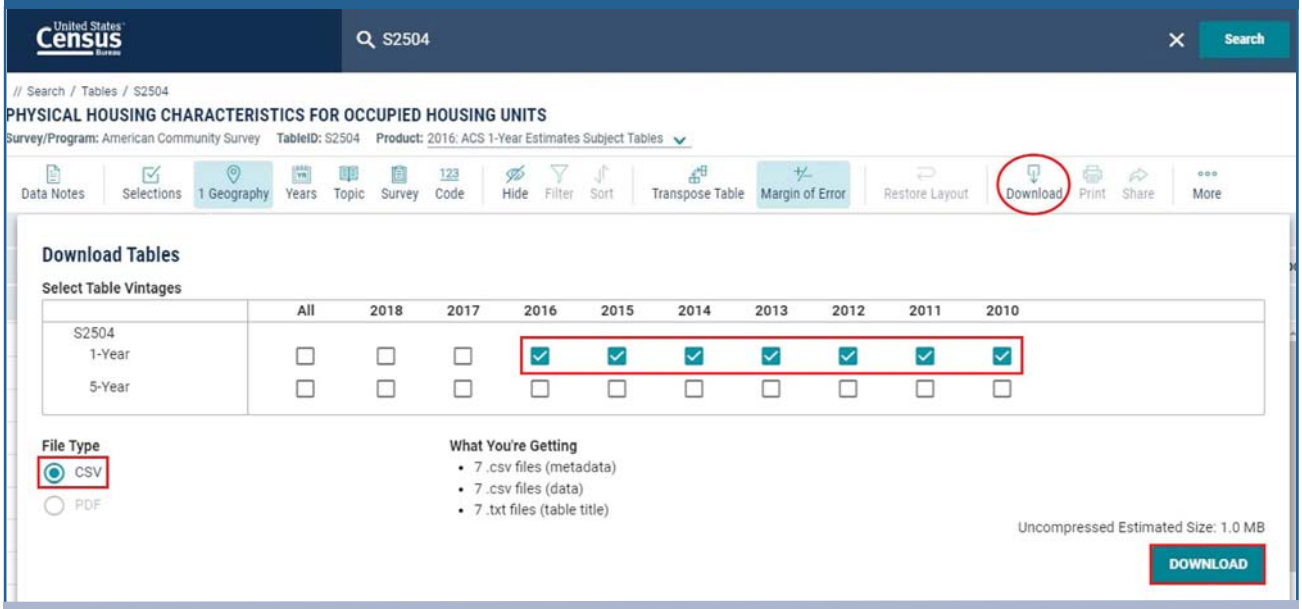
Figure 3.24. **Changing the Survey Year in Data.census.gov**



Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov>.

- He selects “Download” at the top of the window.
- Then, he uses the Download Tables window to check the boxes for the 2016 through 2010 1-year data. Rob downloads data for multiple years to assist with the selection of target markets to recommend.
- He selects “CSV” as the file type and clicks on “Download” (see Figure 3.25).

Figure 3.25. **Selecting ACS Data for Previous Years**



Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov>.

Rob opens the CSV file and copies and pastes the owner-occupied and renter-occupied data for each state into a new table showing the data for the last 5 to 10 years (see Table 3.2). He can then use this table, compare it with internal USAA data showing the number of policies in force for each of the three lines of business, and compare the two tables to determine potential market size by state. Since the purpose of this analysis is to identify target market opportunities, the margin of error is not considered as a critical factor, as the goal is merely to gain reasonable estimates of potential market size.

If desired, this process can be replicated at a county level for further refinement of market segmentation.

Table 3.2. **Owner-Occupied Housing Units by State and Year**

State	2012	2013	2014	2015
Alabama	1,268,565	1,240,017	1,246,080	1,253,595
Alaska	159,427	156,107	156,006	159,922
Arizona	1,496,650	1,490,031	1,484,857	1,524,828
Arkansas	757,722	739,987	744,318	746,000
California	6,781,817	6,804,639	6,855,688	6,910,823
Colorado	1,278,158	1,291,945	1,302,533	1,322,618
Connecticut	908,452	888,478	900,039	889,035
Delaware	241,050	243,047	245,808	249,681
District of Columbia	110,681	110,513	112,492	112,555
Florida	4,724,428	4,672,482	4,693,821	4,760,071

Source: U.S. Census Bureau, data.census.gov, Table S2504: “Physical Housing Characteristics for Occupied Housing Units.”