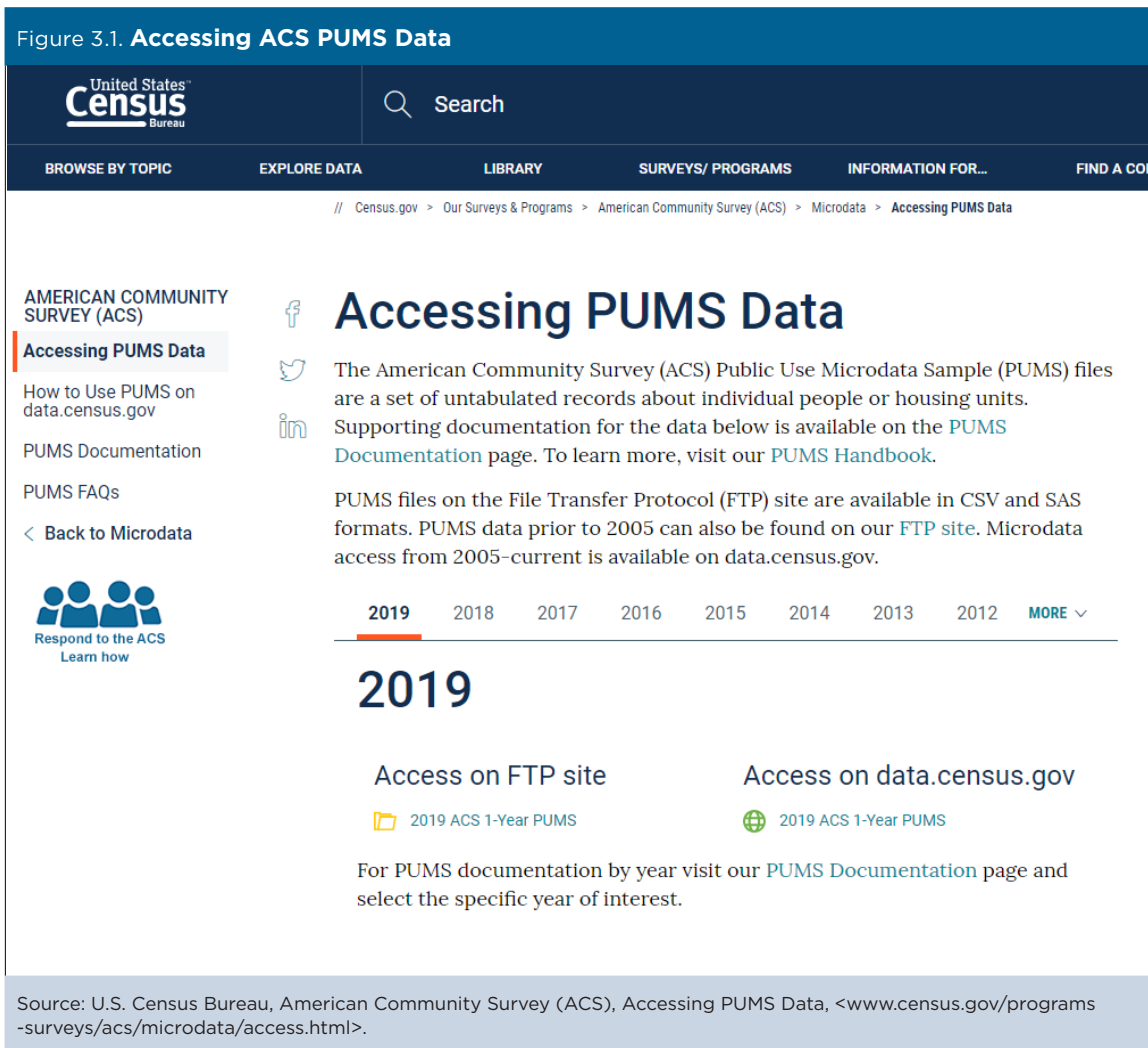


3. ACCESSING ACS PUMS DATA

Data users can access the American Community Survey (ACS) Public Use Microdata Sample (PUMS) files through the U.S. Census Bureau’s “Accessing PUMS Data” Web page (see Figure 3.1).¹⁶ First, select a year of interest in the year tabs. From there, you can click

on the link under “Access on FTP site” to navigate to the PUMS files on the Census Bureau’s FTP site. You can also click on a link under “Access on data.census.gov” to access the PUMS data files using the microdata access tool on data.census.gov, the Census Bureau’s dissemination tool.

¹⁶ U.S. Census Bureau, American Community Survey (ACS), Accessing PUMS Data, <www.census.gov/programs-surveys/acs/microdata/access.html>.



From the FTP site, you have already selected your desired year and data set of interest (see Figure 3.1).¹⁷ ACS 1-year PUMS data sets for 2005 and later years contain data on approximately 1 percent of the U.S. population, while ACS 5-year PUMS files for 2005–2009 and later years contain data on approximately 5 percent of the population.

The next step is to download the data set(s) of interest (see Figure 3.2). PUMS files on the Census Bureau's

¹⁷ The Census Bureau previously released 3-year estimates based on 36 months of data collection. In 2015, the 3-year products were discontinued. The 2011–2013 ACS 3-year estimates, released in 2014, are the last release of this product.

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.

- File formats are comma separated value files (CSV) and SAS data sets for UNIX.
- Record types are housing files (h) or person files (p).
- 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.”

Figure 3.2. Downloading ACS 5-Year PUMS Files From the Census Bureau's FTP Server

Name	Last modified	Size	Description
Parent Directory		-	
PUMS_file_naming_convention.pdf	31-Jan-2019 12:21	36K	
csv_hak.zip	31-Jan-2019 12:54	2.4M	
csv_hal.zip	31-Jan-2019 12:54	17M	
csv_har.zip	31-Jan-2019 12:54	10M	
csv_haz.zip	31-Jan-2019 12:54	22M	
csv_hca.zip	31-Jan-2019 12:54	105M	
csv_hco.zip	31-Jan-2019 12:54	17M	
csv_hct.zip	31-Jan-2019 12:54	11M	
csv_hdc.zip	31-Jan-2019 12:54	2.4M	
csv_hde.zip	31-Jan-2019 12:54	3.2M	
csv_hfl.zip	31-Jan-2019 12:54	67M	
csv_hga.zip	31-Jan-2019 12:54	32M	
csv_hhi.zip	31-Jan-2019 12:54	4.1M	

Source: U.S. Census Bureau, FTP server, <<https://www2.census.gov/programs-surveys/acs/data/pums/2017/5-Year/>>.

PUMS data sets can be downloaded for the entire nation, individual states, the District of Columbia, or for Puerto Rico. Data for Puerto Rico are not included in the national files. To create a complete national data set, you need to combine two or more files together. The 1-year data are divided into two files—an “a” and a “b” file—while the 5-year PUMS is divided into “a,” “b,” “c,” and “d” files.

TIP: If you only need data for a single state or a few states, you can save hard disk space and computer-processing time by downloading data only for those areas rather than the entire nation. For example, the 2016 ACS 1-year PUMS SAS data set for the nation as a whole—including all person and housing records—is 4.0 GB in size, while the 2012–2016 ACS 5-year data set is 18.8 GB in size.

Housing and Population Records

The ACS PUMS data sets are hierarchical, with separate records for housing units and population. The housing records include information about the housing unit—such as the number of rooms, whether the unit is vacant, access to the Internet, and mortgage payments—and household characteristics such as household income and the presence of children. The housing data set contains one record for each sampled household.

The basic unit in the ACS PUMS is an individual housing unit. The PUMS also includes a sample of people living in group quarters (GQ). The population sample is defined as all persons living in households selected

in the housing unit sample, plus the persons selected from the GQ sample. Each GQ record is assigned a unique SERIALNO, and data users cannot determine whether two individuals reside in the same GQ.

The population records contain information about individuals such as age, marital status, and educational attainment. The population data set includes one record for each individual in each sampled household or group quarters. Because PUMS data sets are large, you should only download the records that you need. If you need both housing and population records, then you need to download both files and merge them together. (See the section on “Preparing ACS PUMS Data Files for Analysis” for more information.)

Click on the selected data set to download the PUMS data. The data set is provided as a compressed ZIP file along with a README document that provides information about how to start using PUMS data.

Microdata Access on Data.census.gov

The Census Bureau recently launched a new microdata access tool on data.census.gov that may be used to generate ACS estimates online without the use of statistical software. This tool is still under development and in beta form on the Census Bureau’s Web site.¹⁸ 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.¹⁹

¹⁸ U.S. Census Bureau, Microdata Access Tool, <<https://data.census.gov/mdat/>>.

¹⁹ U.S. Census Bureau, Using Microdata Access, <<https://www2.census.gov/data/api-documentation/using-microdata-access/microdata-access-1-year-acs-pums.pdf>>.