Introduction to the American Community Survey Public Use Microdata Sample (PUMS) Files

March 11, 2020

Toll free audio access: 888-946-3820 (corrected)

Passcode: 5500300#

Amanda Klimek Survey Statistician, American Community Survey Office



The American Community Survey About the U.S. Census Bureau

- The U.S. Census Bureau is the federal government's largest statistical agency.
- We conduct more than 100 censuses and surveys each year, including
 - The Decennial Census the once-a-decade population and housing count of all 50 states and U.S. territories
 - The American Community Survey the ongoing, annual survey of the social, economic, housing, and demographic characteristics of the nation's population
 - The Economic Census the official five-year measure of American businesses
- Our mission is to serve as the nation's leading provider of quality data about its people and economy.





The American Community Survey Outline

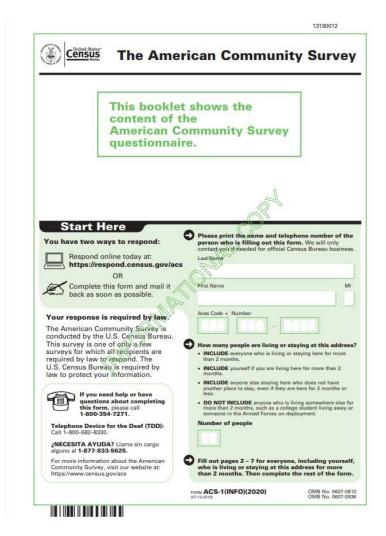
- American Community Survey (ACS) and PUMS basics
- PUMS Geography
- Accessing PUMS Data through the File Transfer Protocol (FTP)
- Creating custom PUMS Tables on data.census.gov
- Resources for learning more
- Questions



The American Community Survey The Foundation

The American Community Survey is on the leading edge of survey design, continuous improvement, and data quality

- The nation's most current, reliable, and accessible data source for local statistics on critical planning topics such as age, children, veterans, commuting, education, income, and employment
- Surveys 3.5 million addresses and informs over \$675 billion of Federal government spending each year
- Covers 40+ topics, supports over 300 evidence-based Federal government uses, and produces 11 billion estimates each year
- Three key annual data releases:
 - 1-year Estimates (for large populations)
 - 1-year Supplemental Estimates (for small populations)
 - 5-year Estimates (for very small populations)





Content Overview

POPULATION		HOUSING	
SOCIAL Ancestry Citizenship Citizen Voting Age Population Disability Educational Attainment Fertility Grandparents Language Marital Status Migration School Enrollment Veterans	DEMOGRAPHIC Age Hispanic Origin Race Relationship Sex ECONOMIC Class of Worker Commuting Employment Status Food Stamps (SNAP) Health Insurance Hours/Week, Weeks/Year Income Industry & Occupation	Computer & Internet Use Costs (Mortgage, Rent, Taxes, Insurance) Heating Fuel Home Value Occupancy Plumbing/Kitchen Facilities Structure Tenure (Own/Rent) Utilities Vehicles Year Built/ Year Moved In	



The American Community Survey Overview of PUMS

- PUMS stands for the Public Use Microdata Sample
 - Microdata file is a subsample of the full ACS records
 - 1-year (approximately 1% of U.S. households)
 - 5-year (approximately 5% of U.S. households)
 - Additional data disclosure measures are applied to PUMS to protect confidentiality
- PUMS allows data users to create their own estimates which may not be published on data.census.gov
- Statistical programs, such as SAS, R, Python, or STATA are recommended to calculate PUMS estimates and MOEs



The American Community Survey Why Use PUMS?

DETAILED OCCUPATION FOR THE FULL-TIME, YEAR Survey/Program: American Community Survey Universe: Full-time,			ar Estimates Detailed Tables 🔻
Data Notes Selections Geographies Years Topic	123 Survey Code Hide Filter Sort Transpose	Table Margin of Error Restore Layout	Download Print Share More Data Map
		Uni	ited States
		Estimate	Margin of Error
✓ Total:		47,989,329	+/-85,819
Chief executives		268,474	+/-10,150
General and operations managers		283,257	+/-9,175
Legislators		4,890	+/-1,111
Advertising and promotions managers		24,483	+/-2,628
Marketing managers		263,399	+/-9,099

- Data needs not supported by standard tables
 - Example: occupation by sex by marital status ("married female actuaries")
- Can create new measures with unique combinations of person and/or household variables
 - Example: spouse's occupation
- Users want to conduct sophisticated statistical analysis to understand relationship between variables
 - Example: correlation analysis



Summary Data vs. Microdata

	Louisiana		
	Estimate	Margin of Error	
✓ Total:	2,015,672	+/-17,879	
✓ Male:	1,029,182	+/-10,608	
Management, business, science, and arts occupations:	302,411	+/-8,186	
Management, business, and financial occupations:	138,208	+/-6,478	
Computer, engineering, and science occupations:	64,855	+/-4,695	
Computer and mathematical occupations:	22,827	+/-2,629	
→ Computer occupations:	22,439	+/-2,626	
Computer and information research scientists and analysts	2,287	+/-781	
Software developers and programmers	6,874	+/-1,772	

Aggregated tables for a geography: "In 2018 in Louisiana, approximately 22,827 males work in computer and mathematical occupations."

Individual responses:

"This male in Louisiana is a web developer."

RT	SERIALNO	SPORDER	ST	SEX	OCCP
P	267855	2	22	1	6600
Р	267870	1	22	2	1020
Р	267870	2	22	1	1030
Р	267913	1	22	2	430
P	267913	2	22	1	9620
P	268097	1	22	2	4110
D	268007	າ	22	1	6260



The American Community Survey Weighting in PUMS

- Data users must aggregate PUMS records to create weighted estimates
- PUMS weights and replicate weights are provided
 - WGTP: PUMS household weights
 - used to produce housing unit estimates
 - PWGTP: PUMS person weights
 - used to produce population estimates
 - WGTP1 WGTP80, PWGTP1 PWGTP80: PUMS Replicate Weights
 - used to calculate standard errors
- Use the variance formula to calculate the MOE



The American Community Survey What is the Margin of Error (MOE)?

Definition: An MOE is a measure of the possible variation of the estimate around the population value

- At a given confidence level, the estimate and actual population value will differ by no more than the MOE
- 90% confidence level is the Census standard and ACS MOEs are provided in the same unit as the estimate

Example: How many males under age 5 live in WY?

Lower bound = 19,174 - 346 = 18,828

Upper bound = 19,174 + 346 = 19,520

We are 90% confident the true number of males under age 5 in Wyoming falls **between 18,828 and 19,520**

	Wyoming		
	Estimate	Margin of Error	
✓ Total:	583,200	****	
✓ Male:	298,301	+/-673	
Under 5 years	19,174	+/-346	
5 to 9 years	20,796	+/-771	
10 to 14 years	20,227	+/-740	
15 to 17 years	11,677	+/-367	
18 and 19 years	7,605	+/-364	



Webinar: Calculation Margins of Error the ACS Way

Calculating Margins of Error the ACS Way Using Replicate Methodology to Calculate Uncertainty

February 12, 2020

For audio call: 1-888-950-9568 (passcode: 8679472)

Sirius Fuller Decennial Statistical Studies Division

Charles Gamble American Community Survey Office





The American Community Survey 5-year PUMS Files

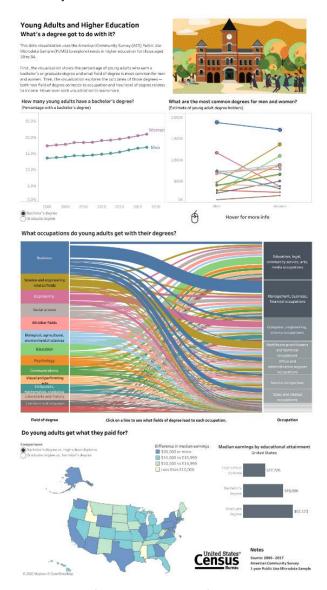
 5-year PUMS files contain the same cases as their component 1-year files

```
2014 ACS 1-year PUMS
2015 ACS 1-year PUMS
2016 ACS 1-year PUMS = 2014-2018 ACS 5-year PUMS
2017 ACS 1-year PUMS
2018 ACS 1-year PUMS
```

- Why do we release multiyear PUMS?
 - New weights are produced using latest population estimate "vintages"
 - Coding schemes and dollar amounts are standardized to latest year in the multi-year file



Example of PUMS in Action





The American Community Survey Outline

- American Community Survey (ACS) and PUMS basics
- PUMS Geography
- Accessing PUMS Data through the File Transfer Protocol (FTP)
- Creating custom PUMS Tables on data.census.gov
- Resources for learning more
- Questions



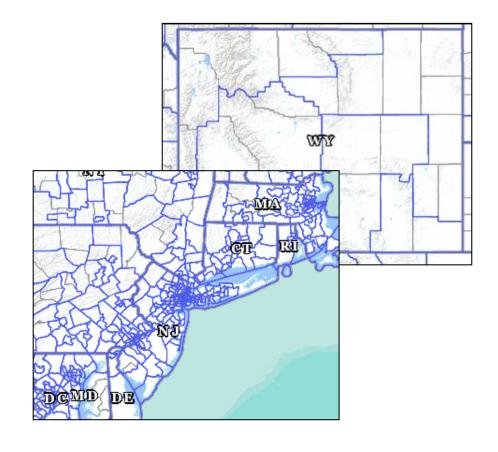
The American Community Survey Levels of Geography

- Region, division, state, and Public Use Microdata Area (PUMA)
- PUMAs can identify most cities of 100,000+ and many metropolitan areas, but not all
- PUMS is not designed for statistical analysis of small geographic areas



The American Community Survey Public Use Microdata Area (PUMA)

- An area with 100,000+ population
- Identified by five-digit code (unique within each state)
- Nest within states or equivalent entities
- Geographically contiguous
- Defined after each census
 - 2010 Census PUMAs first used in the 2012 ACS
 - Census tracts and counties are the building blocks





Approximate PUMAs using MCDC's Geocorr 2014



Geocorr 2018: Geographic Correspondence Engine

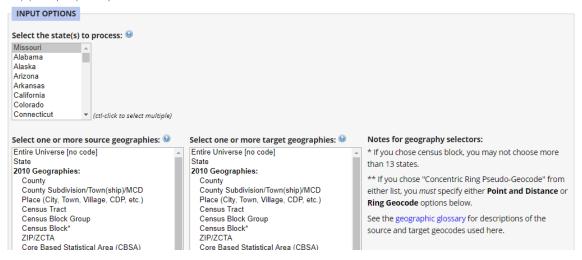
Beta version, revised 12/04/2018 with 2016 geographies

This application accesses the MABLE geographic database to generate custom correlation lists as reports and/or files. Click on the help icons (

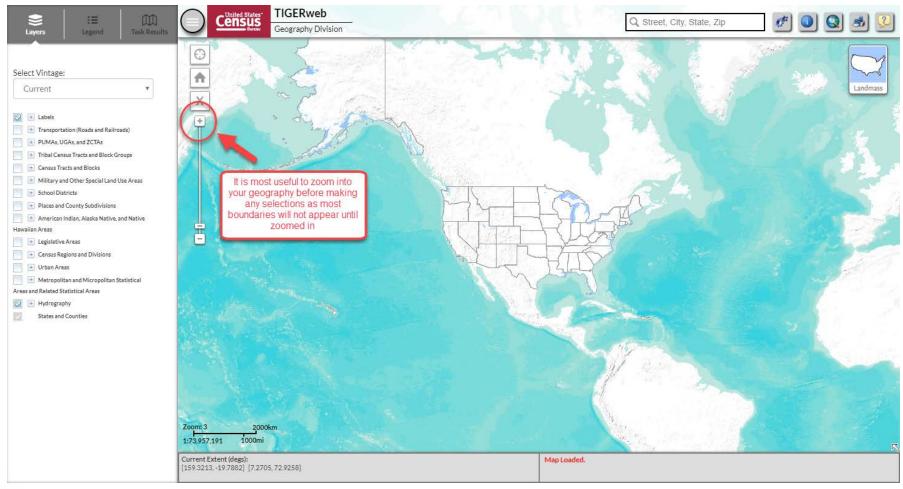
for detailed info on any section of this form.

Note: This application is a "beta" version. Two county name/code changes in 2015 will affect results from AK and SD. See the usage notes for more information. Processing time may be several minutes for large areas or multiple states. Processing time may be greatly reduced by deselecting the HTML report in the output options.

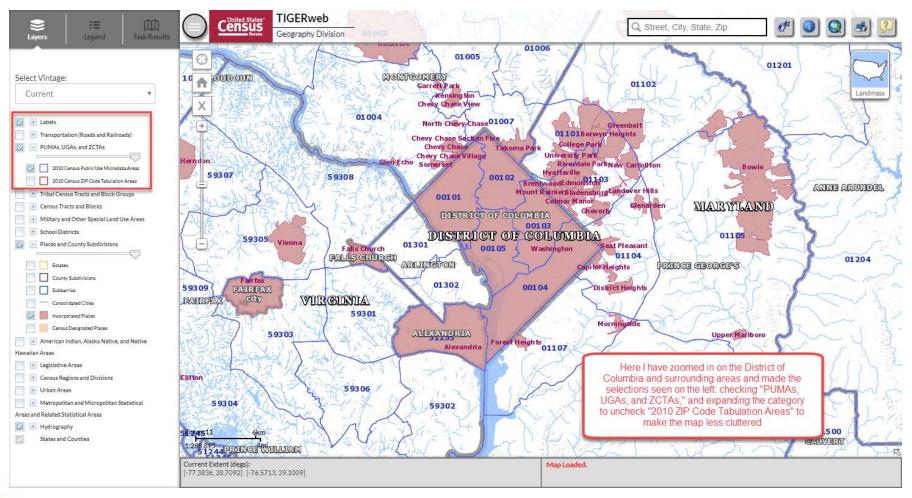
Help | Examples | Notes | Other Geocorr versions



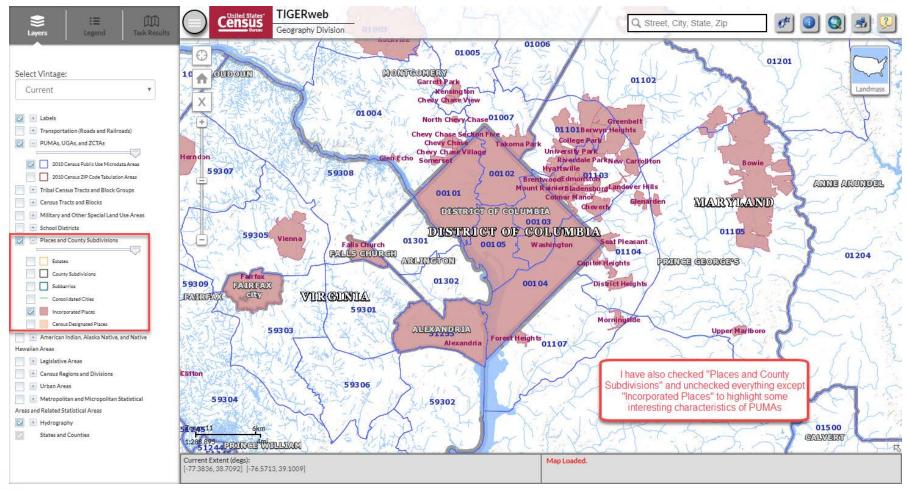




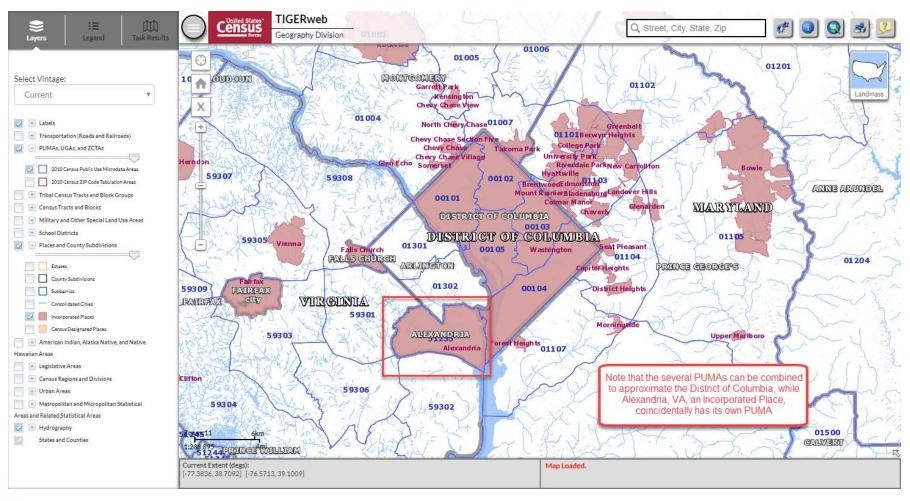




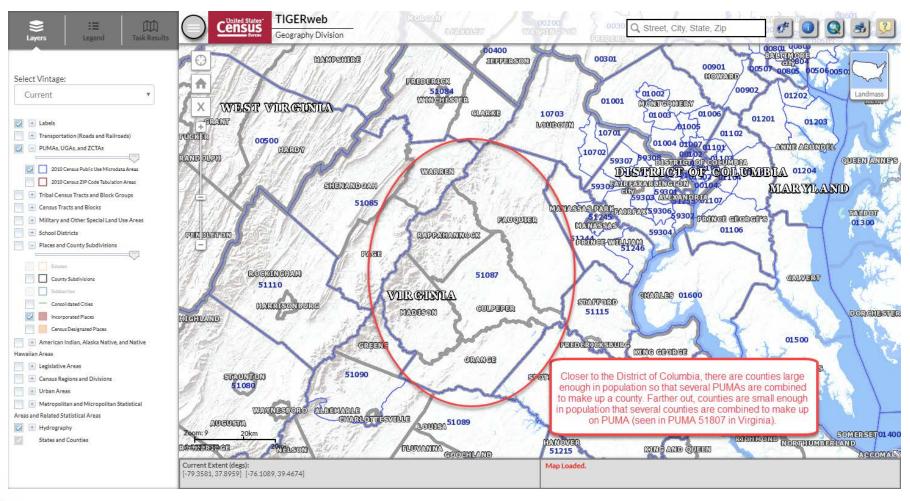














The American Community Survey Outline

- American Community Survey (ACS) and PUMS basics
- PUMS Geography
- Accessing PUMS Data through the File Transfer Protocol (FTP)
- Creating custom PUMS Tables on data.census.gov
- Resources for learning more
- Questions



The American Community Survey Accessing PUMS Data on census.gov/acs

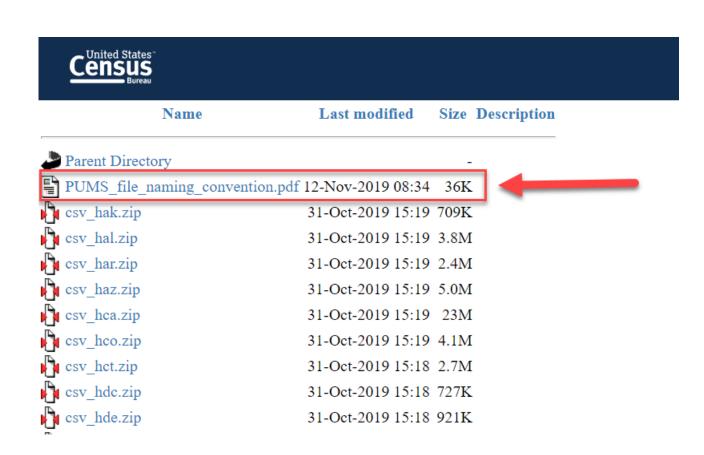
- Go to the census.gov/acs
- Navigate to Data > PUMS Data
- Select the desired year and dataset (1-year vs. 5-year)
- Select format (CSV or SAS)
- Select geography and type
 - (Housing or Population)
- Unzip the files
 - (National will have two to four files)





The American Community Survey Accessing PUMS Data on census.gov/acs

- Go to the census.gov/acs
- Navigate to Data > PUMS Data
- Select the desired year and dataset (1-year vs. 5-year)
- Select format (CSV or SAS)
- Select geography and type
 - (Housing or Population)
- Unzip the files
 - (National will have two to four files)

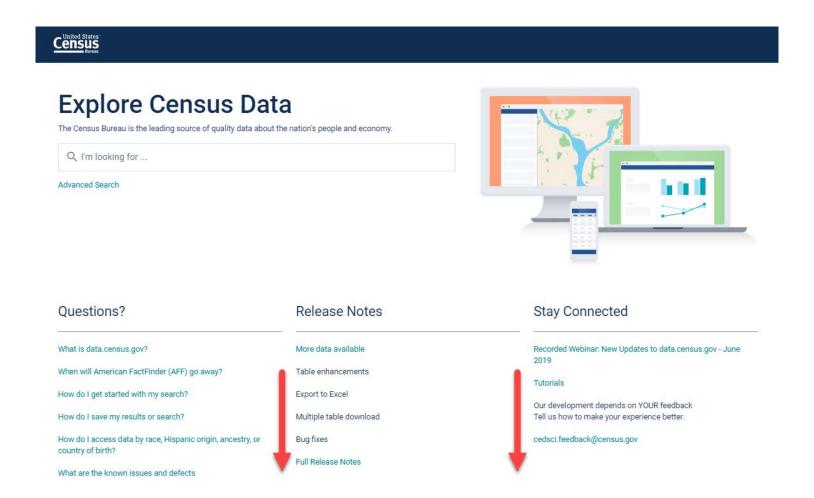




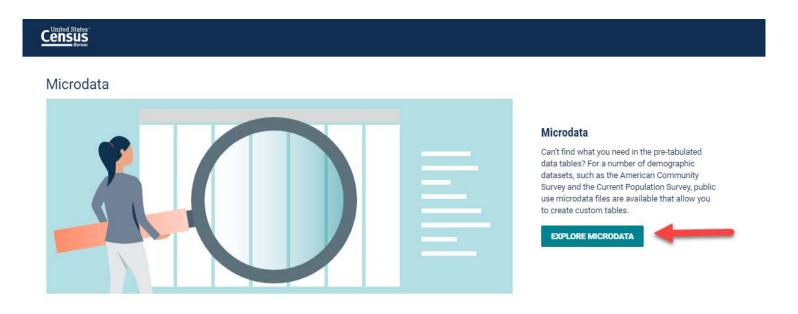
The American Community Survey Outline

- American Community Survey (ACS) and PUMS basics
- PUMS Geography
- Accessing PUMS Data through the File Transfer Protocol (FTP)
- Creating custom PUMS Tables on data.census.gov
- Resources for learning more
- Questions

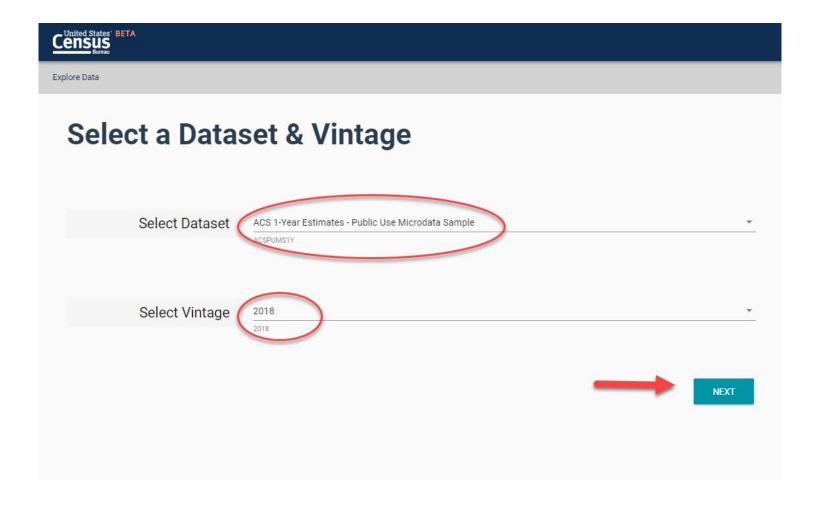




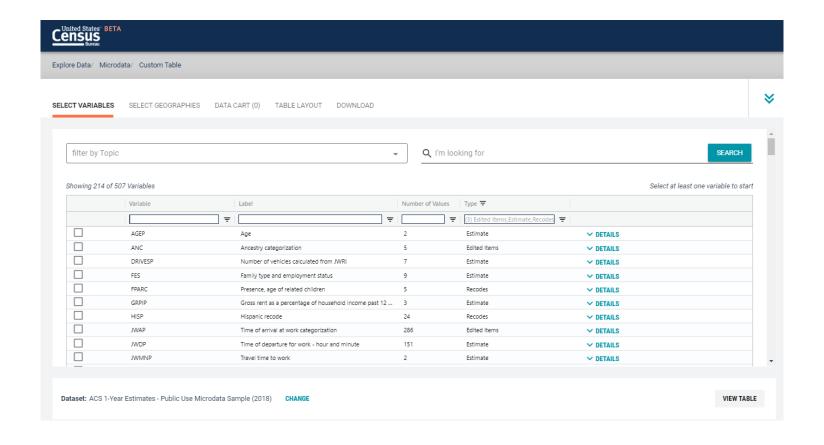




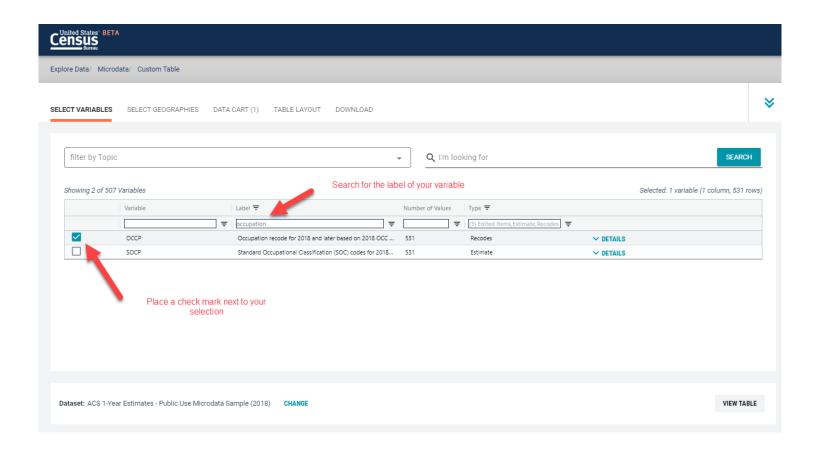




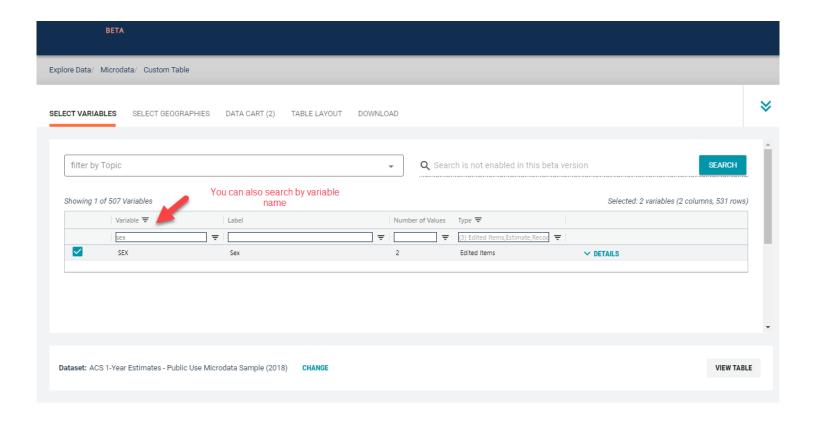




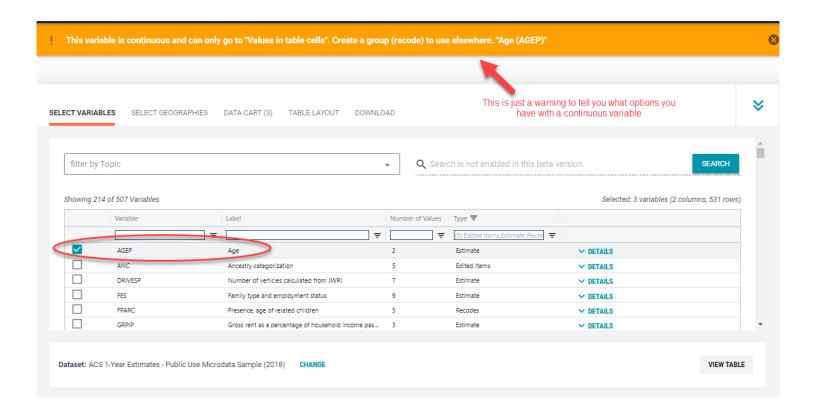




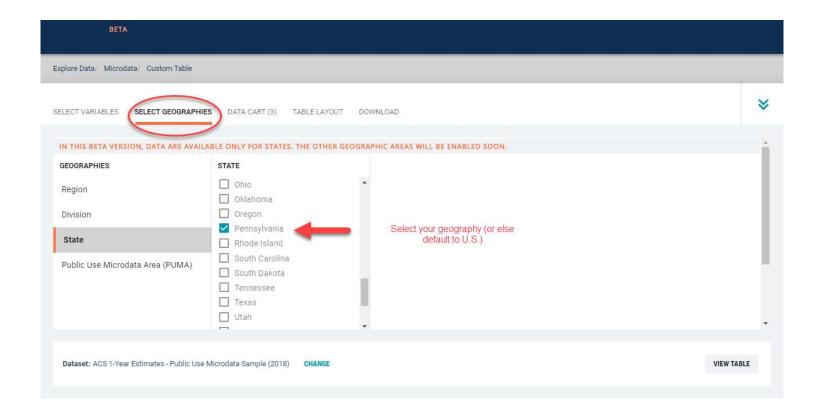




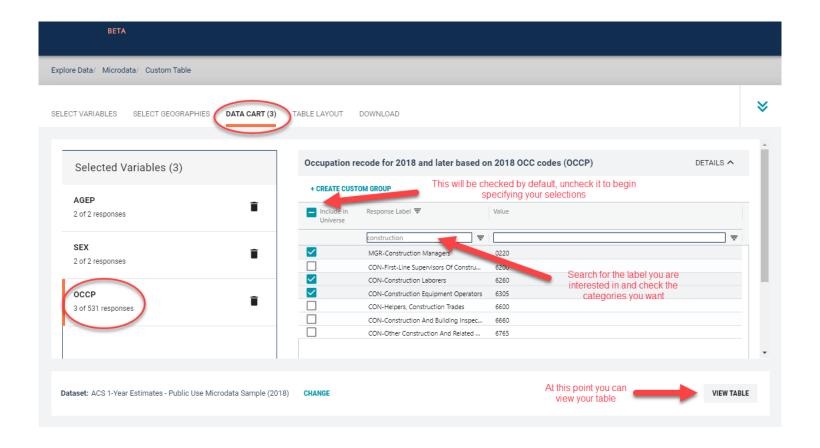




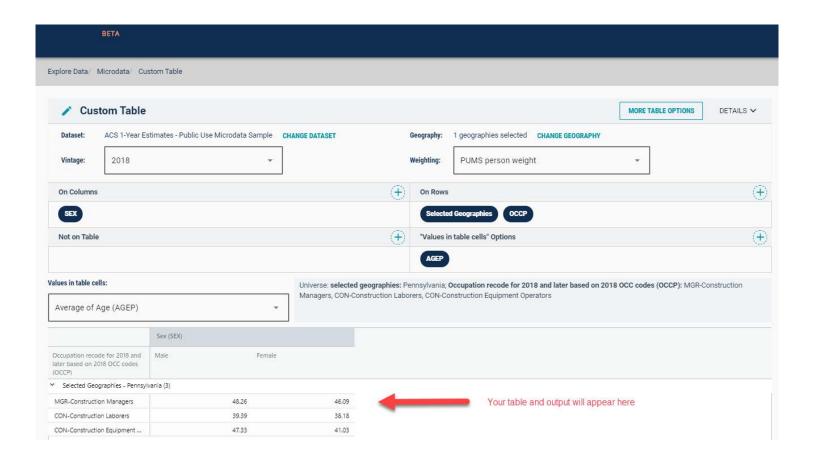




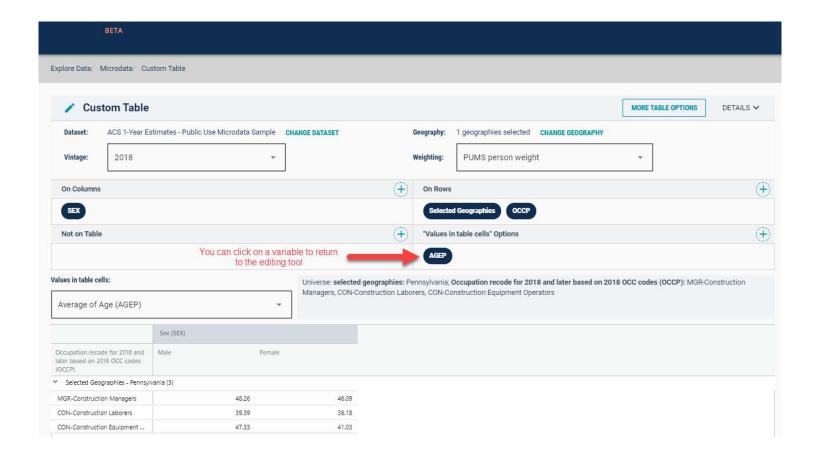




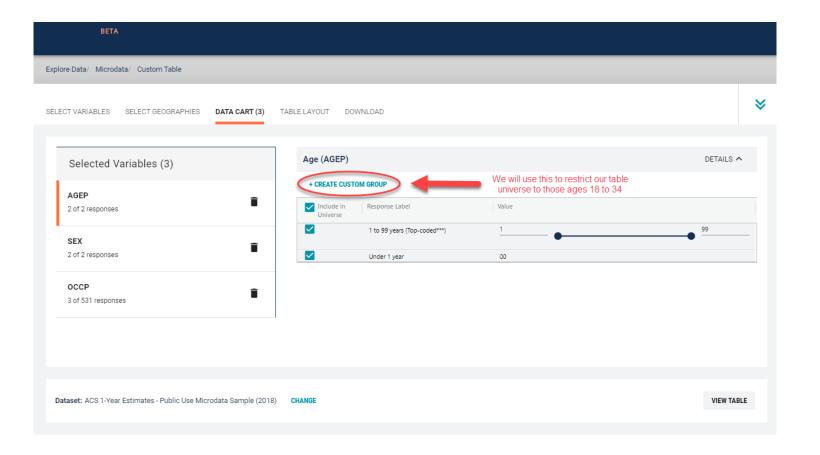




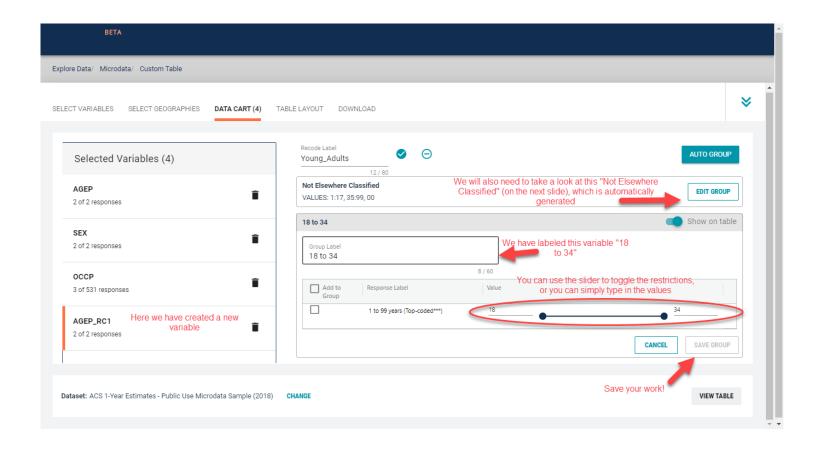




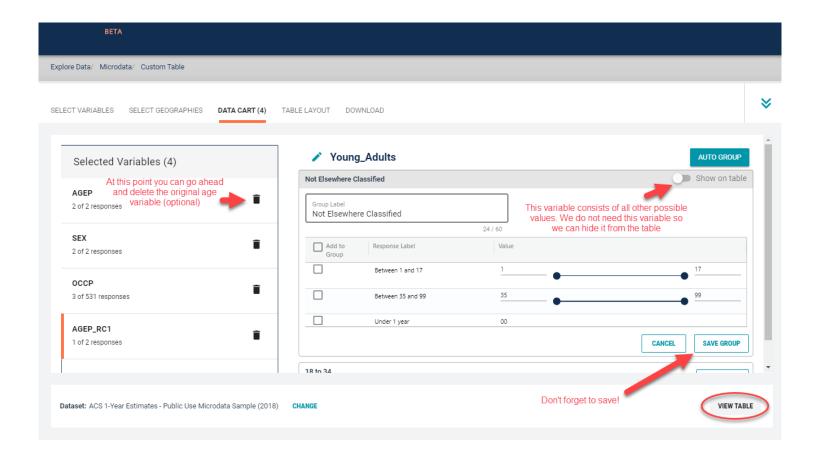




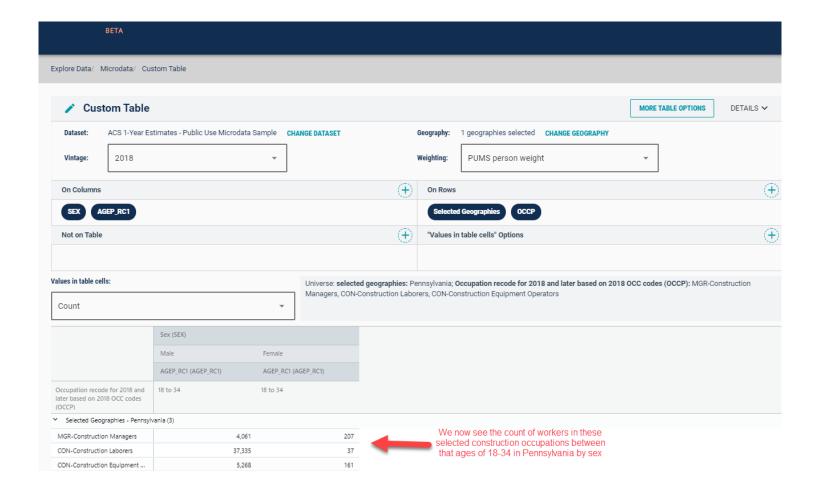




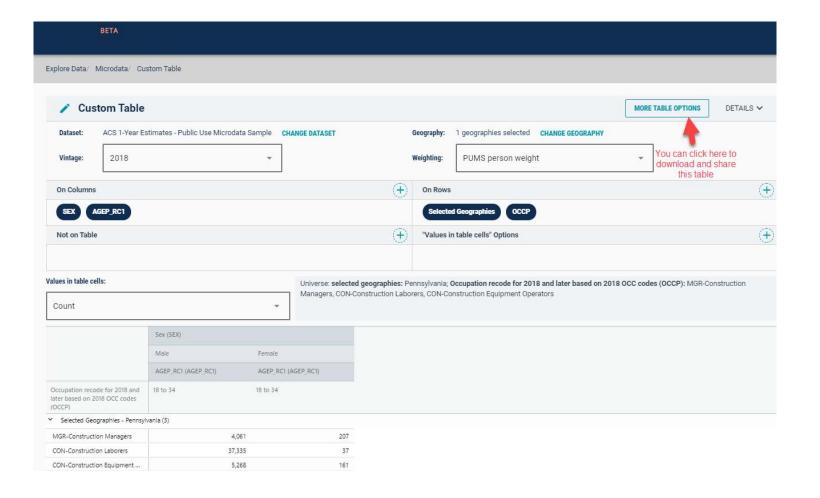




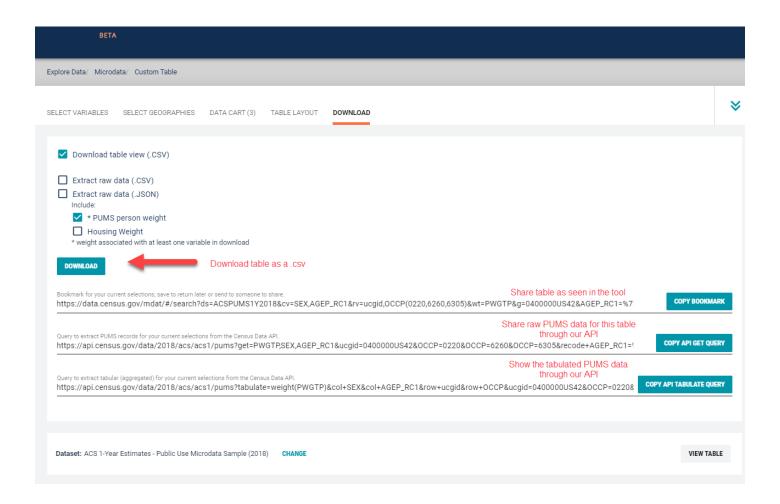














The American Community Survey Outline

- American Community Survey (ACS) and PUMS basics
- PUMS Geography
- Accessing PUMS Data through the File Transfer Protocol (FTP)
- Creating custom PUMS Tables on data.census.gov
- Resources for learning more
- Questions



Understanding and Using American Community Survey Data



Coming soon - Understanding and Using American Community Survey Public Use Microdata Sample Files: What Data Users Need to Know



Microdata Access Tool Step-by-Step Guide

Using Microdata Access

With ACS 1-Year Estimates - Public Use Microdata Sample

data.census.gov/mdat





www2.census.gov/data/api-documentation/using-microdata-access/microdata-access-1-year-acs-pums.pdf

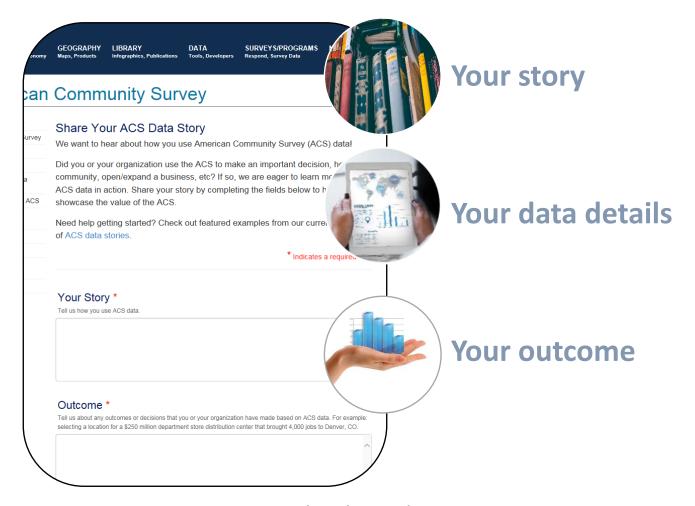
The American Community Survey PUMS Technical Documentation Page

- PUMS ReadMe: Important information about 2016 geography and variable changes as well as guidance for novice ACS PUMS files users
- Subjects in the PUMS: A list of topics included in each of the housing and population record files
- PUMS Data Dictionary: Includes variables available for each PUMS release and how each variable is coded
- **PUMS Code Lists**: Detailed codes for variables that contain a large number of coded responses, such as ancestry and occupation.
- **PUMS Top Coded and Bottom Coded Values**: List of variables with responses exceeding a state specific value that are replaced with a predetermined value
- Accuracy of the PUMS: A basic explanation of the sample design, estimation methodology, and accuracy of the data
- PUMS Estimates for User Verification: A set of Census-created PUMS estimates for data users to verify their methods



census.gov/programs-surveys/acs/technical-documentation.html

Data Tell Stories. Tell us Yours!





census.gov/acs/www/share-your-story

The American Community Survey Continue the Conversation



Sign up for and manage alerts at https://public.govdelivery.com/accounts/USCENSUS/subscriber/new



Social media: @uscensusbureau #ACSdata



More information online: census.gov/acs



acso.users.support@census.gov



Census Customer Service Center 800-923-8282



Source Us:

U.S. Census Bureau's [YYYY – YYYY] American Community Survey [1/3/5]-year [estimates/statistics/data release]



The American Community Survey Outline

- American Community Survey (ACS) and PUMS basics
- PUMS Geography
- Accessing PUMS Data through the File Transfer Protocol (FTP)
- Creating custom PUMS Tables on data.census.gov
- Resources for learning more
- Questions





• Purpose:

- Improve understanding of the value and utility of ACS data.
- Promote information sharing among data users about key ACS data issues and applications.
- Includes users group website and online community
- ACS Data Users Conference
- Membership is free and open to all interested ACS data users



The American Community Survey Need Local Stats?

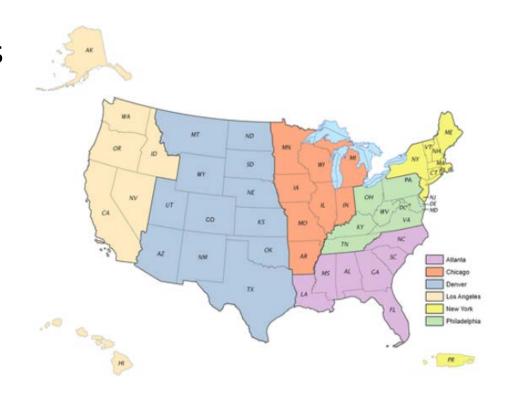
FREE Assistance Near You!
 Our regional data staff can help you access local statistics from the ACS or offer training to help build your skills.

Contact us at:

1-844-ASK-DATA

(1-844-275-2382)

census.askdata@census.gov





The American Community Survey Upcoming/Recent Webinars

Upcoming Webinars:

- March 31: Exploring Census Data Webinar Series: Transportation
- April 22: American Community Survey Administrative Record Research
- May 27: Exploring Census Data Webinar Series: Education
- June 16: Exploring Census Data Webinar Series: Economic Geography

Recent Webinars:

- Calculating Margins of Error the ACS Way Using Replicate Methodology to Calculate Uncertainty
- Census Bureau Webinar on Release of 2014-2018 ACS 5-Year Estimates

