Creating Custom Tables Using the American Community Survey Public Use Microdata Sample

Amanda Klimek, Survey Statistician, American Community Survey Office

February 23, 2022



Have a question?

Send your questions through the Q&A panel located to the right of your screen.

You may need to activate the Q&A by clicking on the Q&A Icon (lower right portion of your Or click screen). on 3 dots



≥ Participants





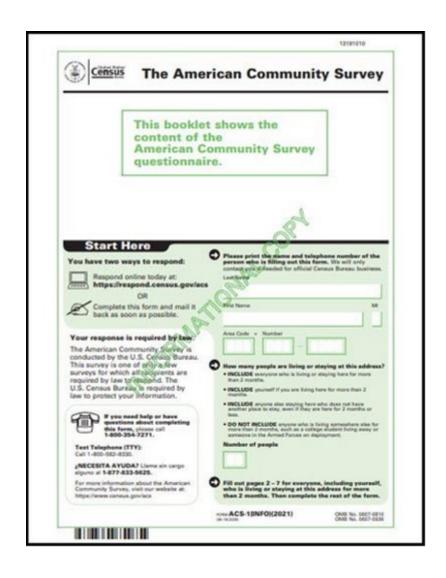
- American Community Survey (ACS) and PUMS basics
- PUMS Geography
- Accessing PUMS Data
- Creating Custom PUMS Tables in data.census.gov
- Introducing the Microdata API
- Resources for Learning More



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)





The American Community Survey Content Overview

POPUI	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, 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 in 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 FUL Survey/Program: American Community Survey									ar Estimates	Detailed T	ables 🗸		
Data Notes Selections Geographies	Years Topic	Survey Coo		Tilter	Jr Sort	₽ Transpose Table	+∕_ Margin of Error	Restore Layout	Ū Download		⇔ Share M	More Data	₩ Map
	United States												
							Es	timate		М	largin of I	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	268097	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			



The American Community Survey 5-year PUMS Files

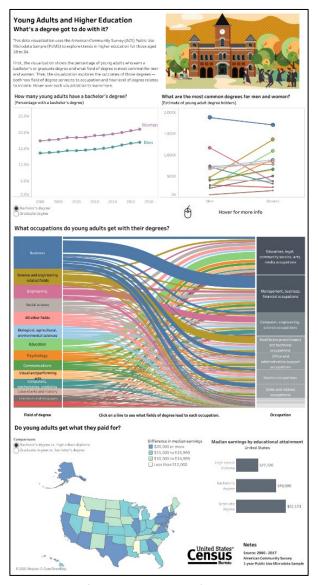
 5-year PUMS files contain the same cases as their component 1-year files

```
2016 ACS 1-year PUMS
2017 ACS 1-year PUMS
2018 ACS 1-year PUMS = 2016-2020 ACS 5-year PUMS
2019 ACS 1-year PUMS
2020 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





- American Community Survey (ACS) and PUMS basics
- PUMS Geography
- Accessing PUMS Data
- Creating Custom PUMS Tables in data.census.gov
- Introducing the Microdata API
- Resources for Learning More



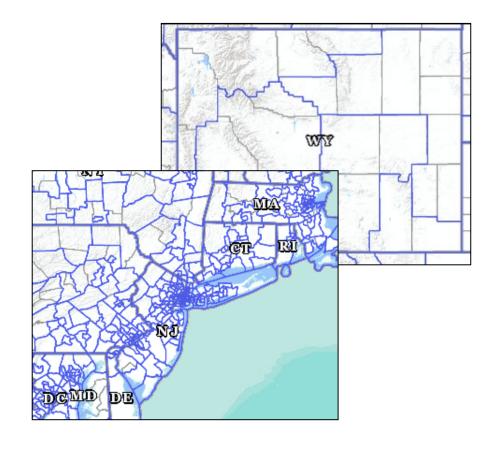
The American Community Survey Levels of Geography

- 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



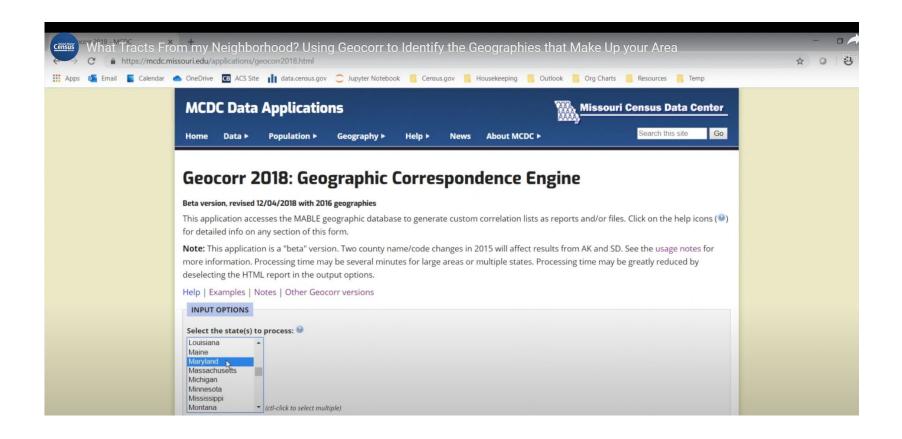


Approximate PUMAs using MCDC's GEOCORR



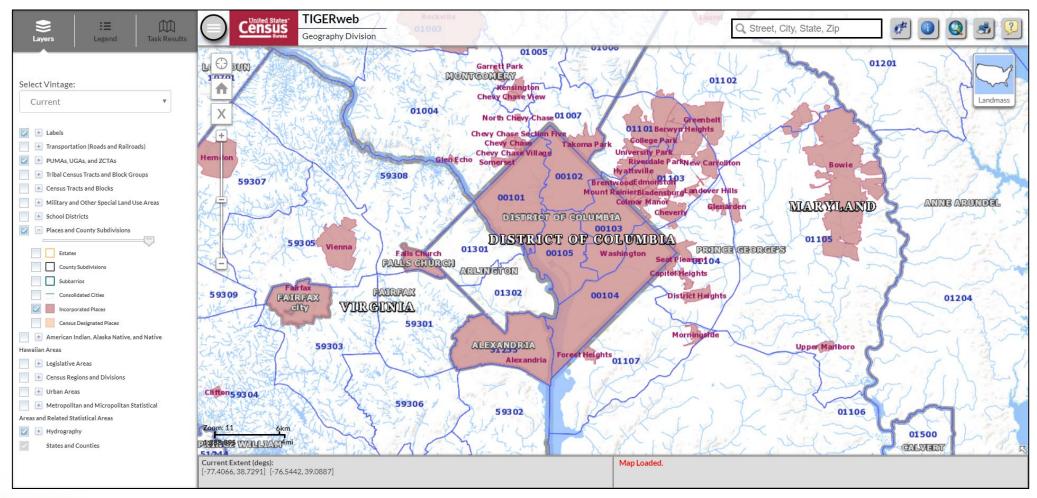


Approximate PUMAs using MCDC's GEOCORR





Visualize PUMA Boundaries for Your Area Using TIGERweb

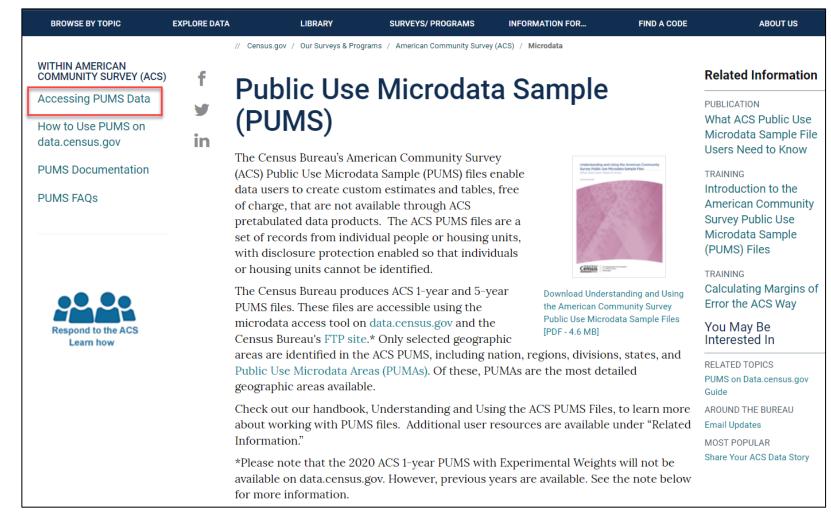




- American Community Survey (ACS) and PUMS basics
- PUMS Geography
- Accessing PUMS Data
- Creating Custom PUMS Tables in data.census.gov
- Introducing the Microdata API
- Resources for Learning More



Accessing PUMS Data on census.gov/acs





Accessing PUMS Data on the File Transfer Protocol (FTP) Site

- Go to census.gov/acs
- Navigate to Microdata > Accessing 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)





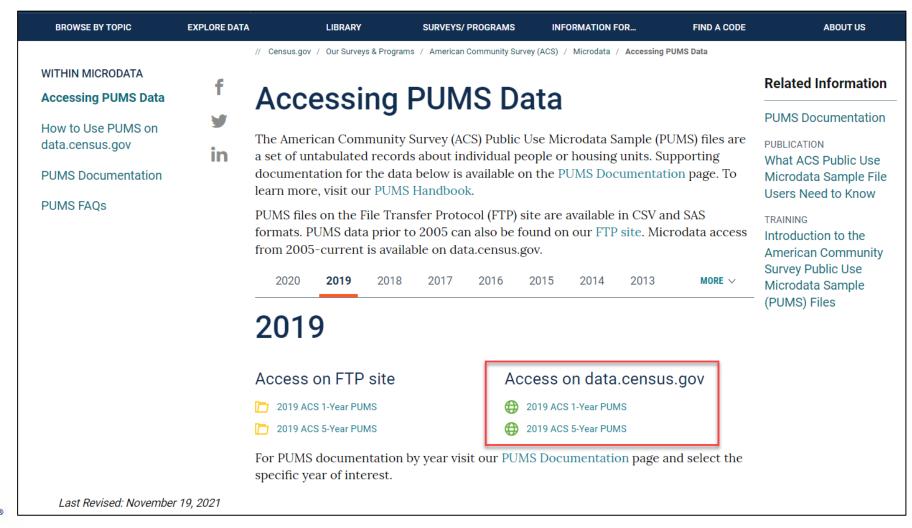
The American Community Survey Accessing PUMS Data on census.gov/acs

- 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)





Accessing PUMS Data on data.census.gov





- American Community Survey (ACS) and PUMS basics
- PUMS Geography
- Accessing PUMS Data
- Creating Custom PUMS Tables in data.census.gov
- Introducing the Microdata API
- Resources for Learning More



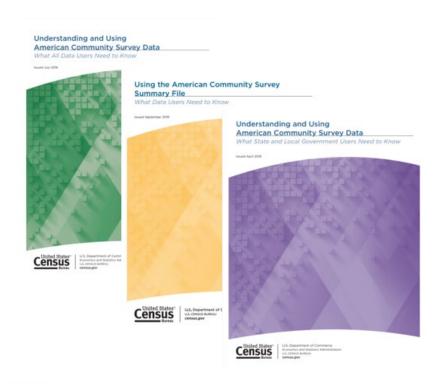
- American Community Survey (ACS) and PUMS basics
- PUMS Geography
- Accessing PUMS Data
- Creating Custom PUMS Tables in data.census.gov
- Introducing the Microdata API
- Resources for Learning More



- American Community Survey (ACS) and PUMS basics
- PUMS Geography
- Accessing PUMS Data
- Creating Custom PUMS Tables on data.census.gov
- Introducing the Microdata API
- Resources for Learning More



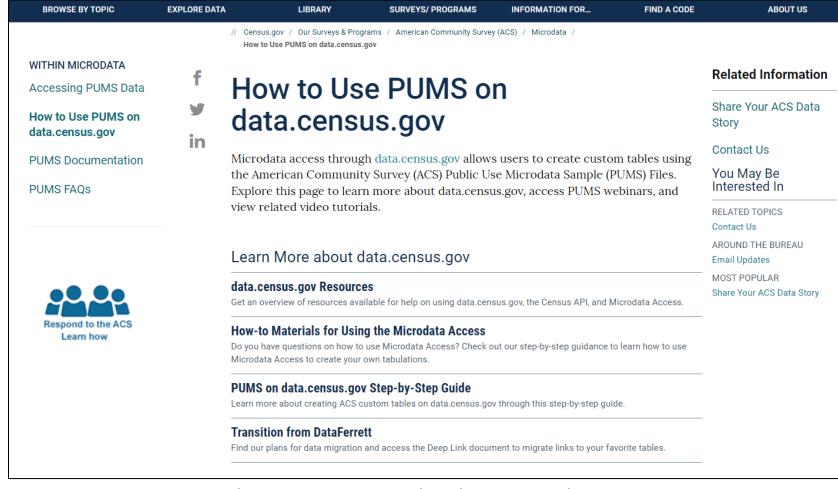
Understanding and Using American Community Survey Data



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



How to Use PUMS in data.census.gov





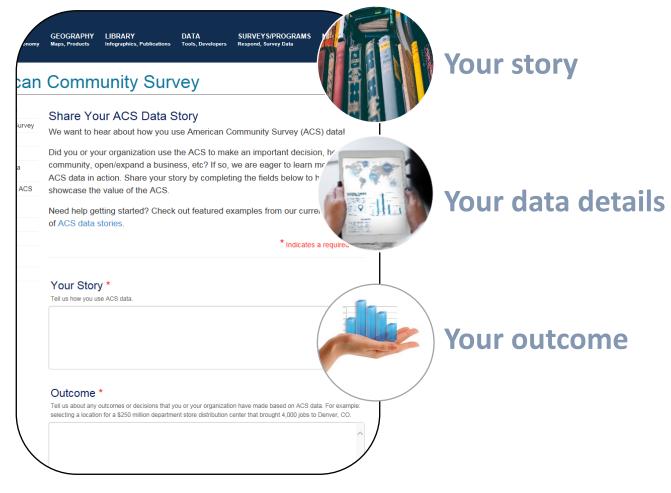
PUMS Documentation Page

PUMS ReadMe		PUMS Top Coded and Bottom Coded Values					
Important information about 2019 geography and variable changes, as well as guidance for novice ACS PUMS files users. 2015-2019 ACS 5-year PUMS ReadMe [<1.0 MB] 2019 ACS 1-year PUMS ReadMe [<1.0 MB]			List of variables with responses exceeding a state- specific value that are replaced with a predetermined value. 2019 ACS PUMS Top & Bottom Coded [<1.0 MB]				
Subjects in the PUMS				[<1.0 MB]			
A list of topics included in each of the h population record files	ousing and	Ac	curacy of the PUMS				
2015-2019 ACS 5-year PUMS Subjects [<1.0 MB] 2019 ACS 1-year PUMS Subjects [<1.0 MB]			A basic explanation of the sample design, estimation methodology, and accuracy of the data.				
			2015-2019 ACS 5-year PUMS Accuracy [PDF]	[<1.0 MB]			
PUMS Data Dictionary		X	2015-2019 ACS 5-year PUMS Design Factors [CSV]	[<1.0 MB]			
	Includes variables available for each PUMS release			[<1.0 MB]			
and how each variable is coded 2015-2019 ACS 5-year PUMS Data Dictionary	[<1.0 MB]	X	2019 ACS 1-year PUMS Design Factors [CSV]	[<1.0 MB]			
2015-2019 ACS 5-year PUMS Data Dictionary	[<1.0 MB]	Pι	JMS Estimates for User				
2015-2019 ACS 5-year PUMS Data Dictionary	[<1.0 MB]		rification				
Z019 ACS 1-year PUMS Data Dictionary	[<1.0 MB]		e that some of these estimates may				
2019 ACS 1-year PUMS Data Dictionary	[<1.0 MB]	different from the estimates for the same characteristics published in data.census.gov. For					
X 2019 ACS 1-year PUMS Data Dictionary	[<1.0 MB]	an explanation of these differences, see the Accuracy of the PUMS above.					



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

The American Community Survey Data Tell Stories. Tell us Yours!





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



• 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 Continue the Conversation



Sign up for and manage alerts at https://public.govdelivery.com/accounts/USCENSUS/signup/12426



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 [1/5]-year [estimates/statistics/data release]

