June 26, 2024

American Community Survey Resources, Shortcuts, and Tools Workshop

Mary Ana McKay, PhD
Survey Statistician, Outreach and Education Branch
American Community Survey Office

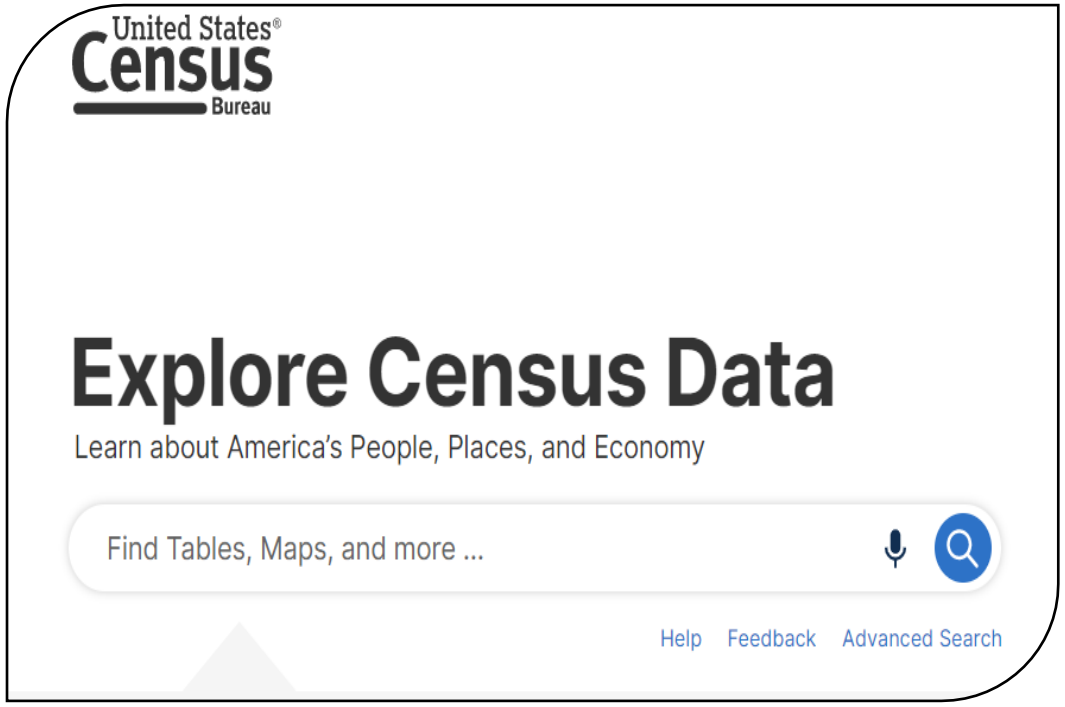
Outline

- Public Use Microdata Sample (PUMS)
  - ACS PUMS basics
  - Run through of Microdata Access Tool
  - ACS PUMS resources

- Application Programming Interface (API)
  - ACS API basics
  - Run through of an API call
  - ACS API resources

American Community Survey Resources, Shortcuts, and Tools Workshop

Introducing data.census.gov



Features:

- Census Bureau's main data dissemination platform
- Filter by geography, surveys, topics, and years
- Download data files and/or create customized maps
- Includes ACS data from 2010 to present

data.census.gov

census.gov/data/what-is-data-census-gov.html

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

ACS Public Use Microdata Sample (PUMS)

Think about…

What are your main goals when accessing ACS data?

How do you access ACS data?

What limitations do you face accessing ACS data?

ACS Public Use Microdata Sample (PUMS)

Outline

- **ACS PUMS basics**
- Run through of Microdata Access Tool
- PUMS Resources

ACS Public Use Microdata Sample (PUMS)

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- PUMS allows data users to create their own estimates which may not be published in data.census.gov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

- Why use the PUMS?
  - *Data needs not supported by standard tables*
    - Example: poverty by educational attainment by children under 18 years ("Households with less than a college degree with young children living under the poverty level")
  - *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
- **Before using PUMS, consider data.census.gov**

ACS Public Use Microdata Sample (PUMS)

PUMS weights

- 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

ACS Public Use Microdata Sample (PUMS)

5-year PUMS Files

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

2018 ACS 1-year PUMS
2019 ACS 1-year PUMS
2020 ACS 1-year PUMS = **2018-2022 ACS 5-year PUMS**
2021 ACS 1-year PUMS
2022 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

ACS Public Use Microdata Sample (PUMS)

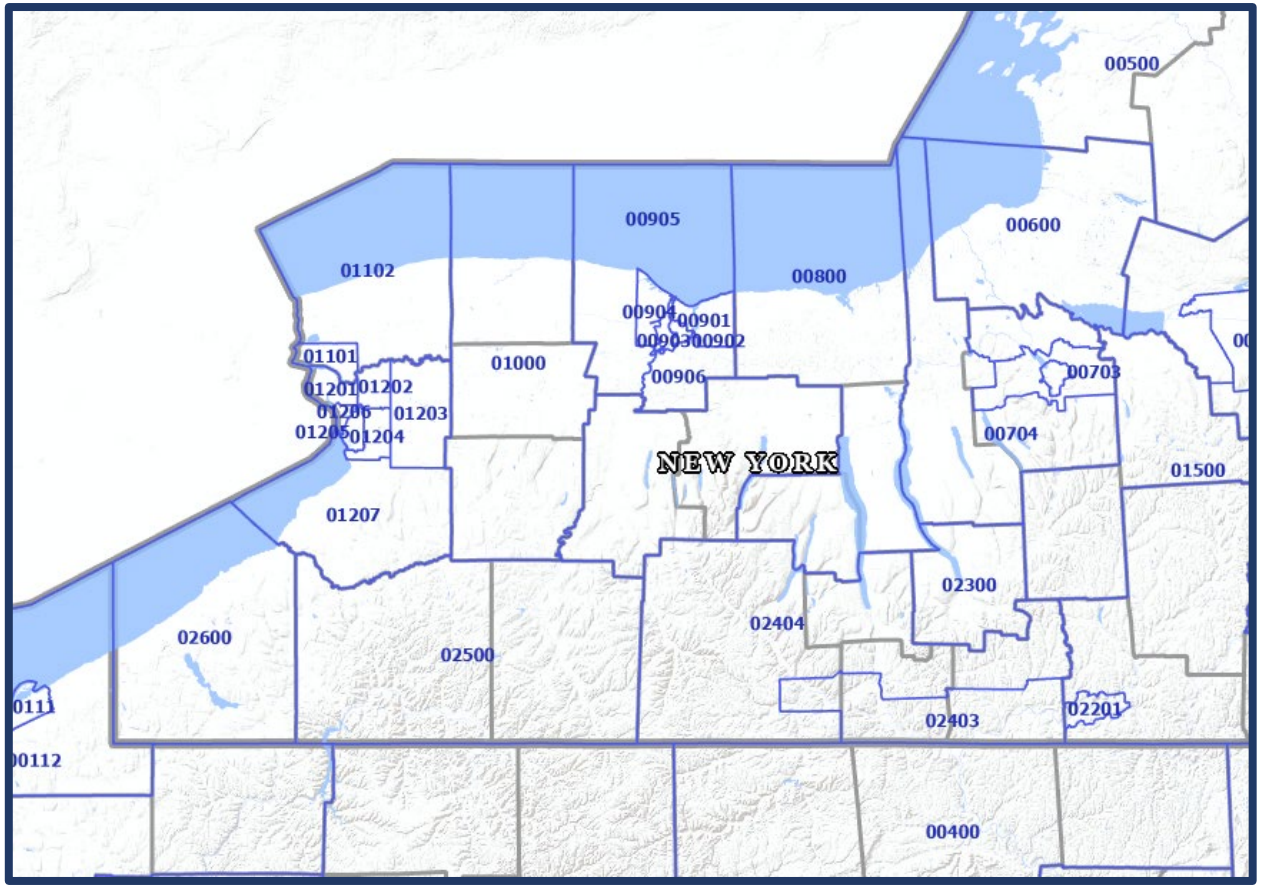
Levels of Geography

- 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**

ACS Public Use Microdata Sample (PUMS)

Public Use Microdata Area (PUMA)

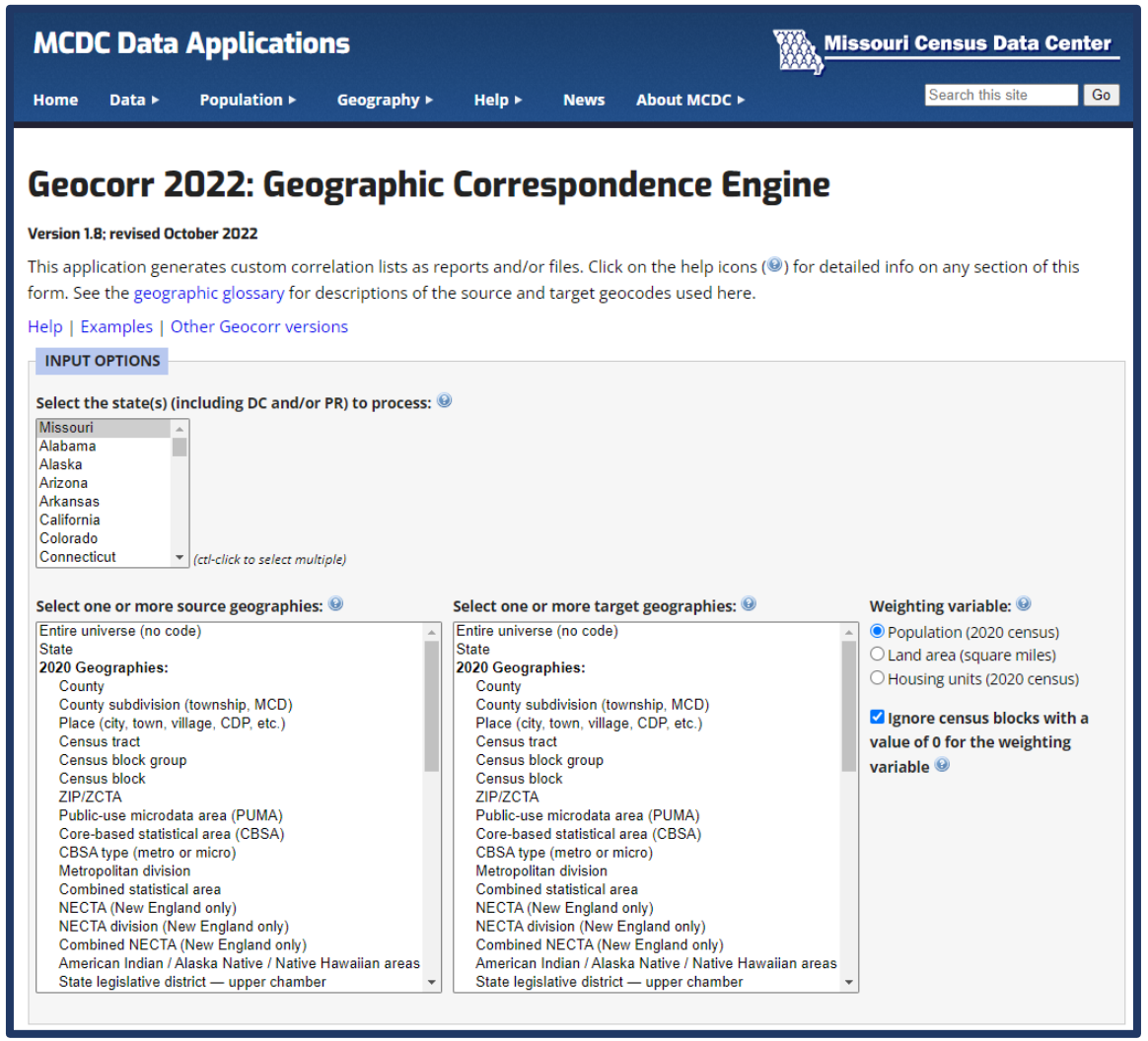
- An area with 100,000+ population
- Identified by five-digit code (unique within each state)
- Nest within states or equivalent entities
- Geographically contiguous
- Defined after each census
  - 2020 Census PUMAs was first used in 2022 ACS
  - Census tracts and counties are the building blocks



census.gov/programs-surveys/geography/guidance/geo-areas/pumas.html

ACS Public Use Microdata Sample (PUMS)

Approximate PUMAs using MCDC's GEOCORR



https://mcdc.missouri.edu/applications/geocorr2022.html

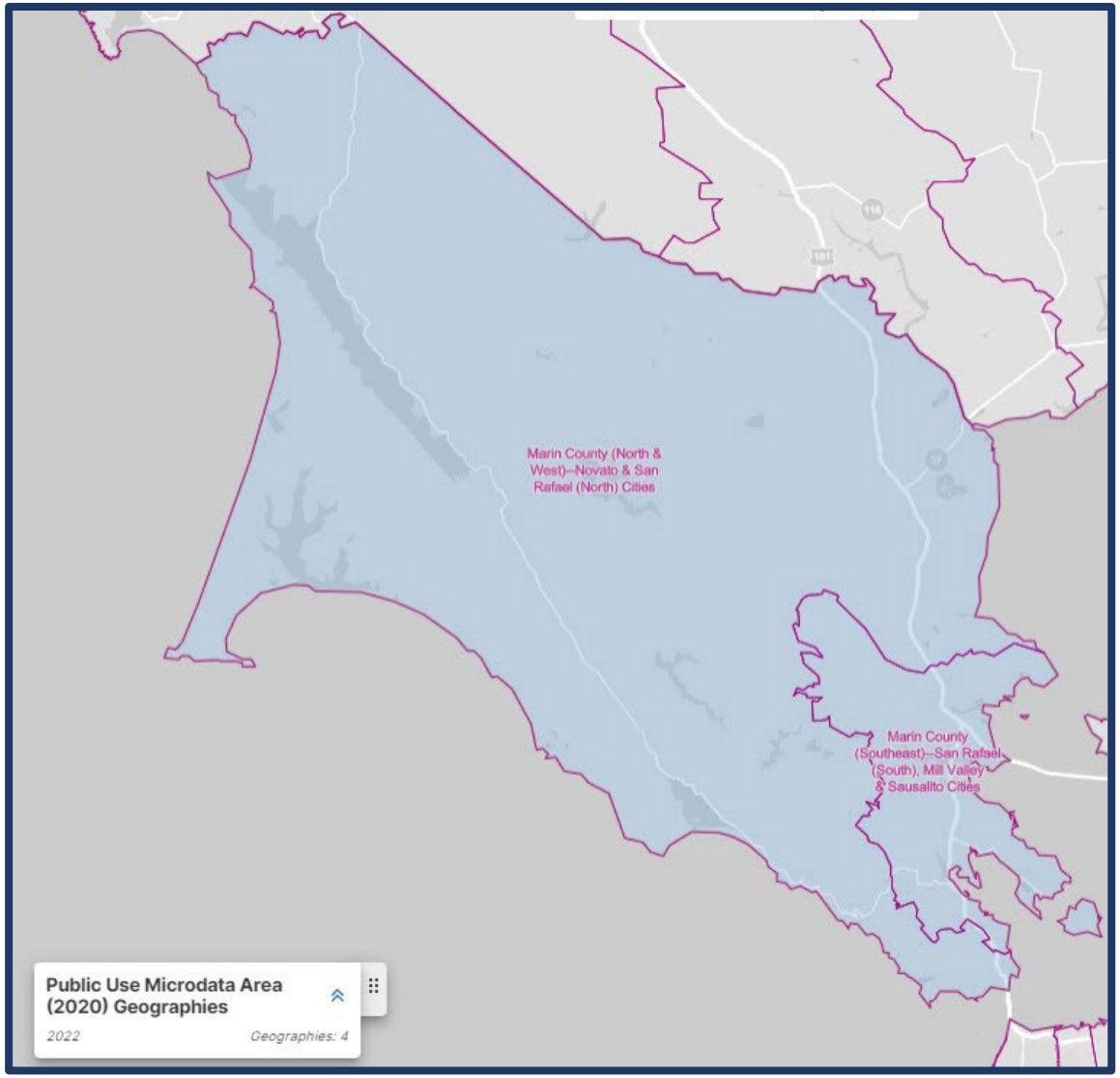
ACS Public Use Microdata Sample (PUMS)

Approximate PUMAs using MCDC's GEOCORR

census.gov/data/academy/data-gems/2021/how-to-use-the-geocorr-to-identify-the-geographies-that-make-up-your-area.html

ACS Public Use Microdata Sample (PUMS)

Visualize PUMA Boundaries for Your Area Using data.census.gov



data.census.gov/map

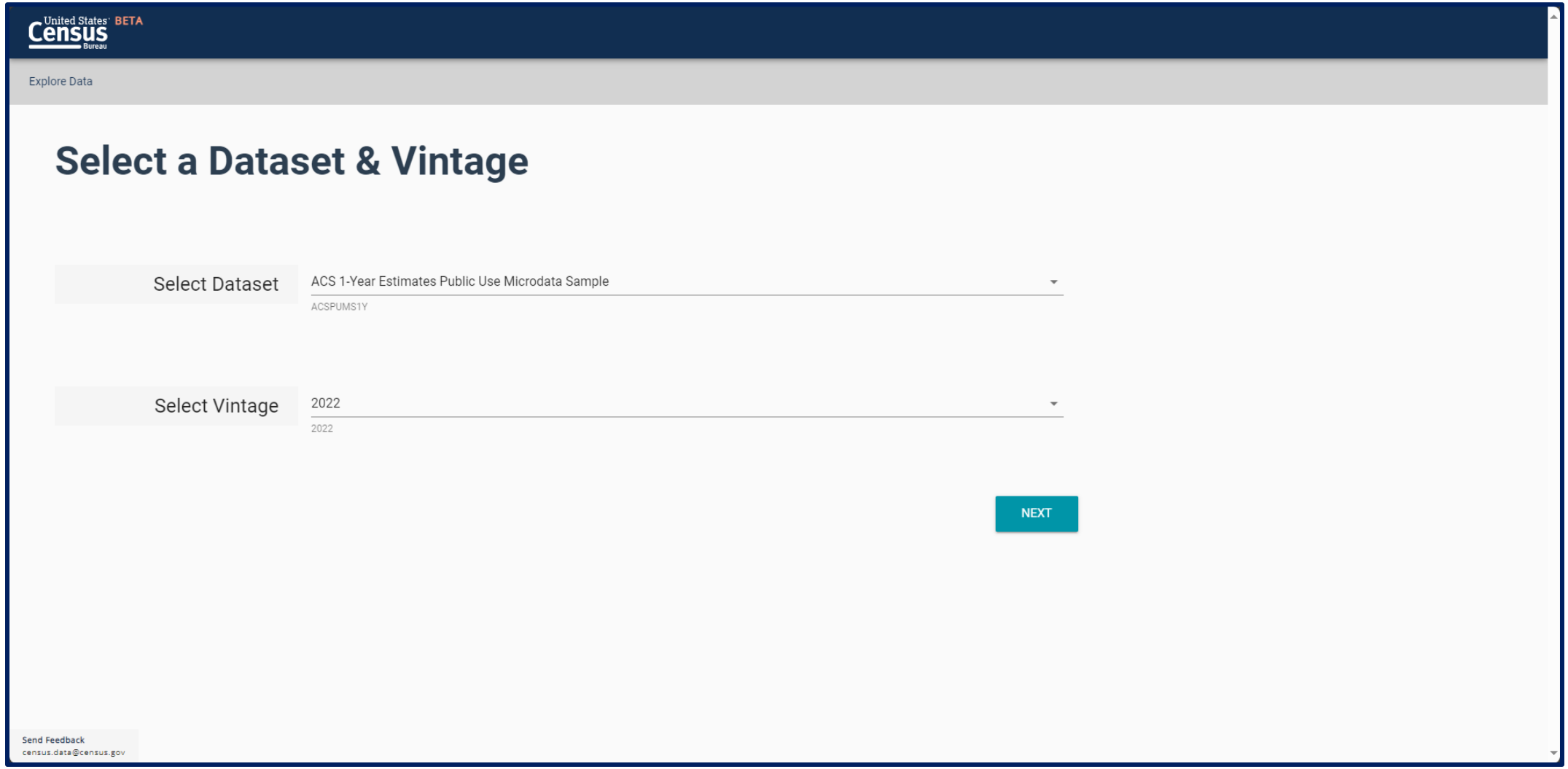
ACS Public Use Microdata Sample (PUMS)

Outline

- ACS PUMS basics
- **Run through of Microdata Access Tool**
- PUMS Resources

ACS Public Use Microdata Sample (PUMS)

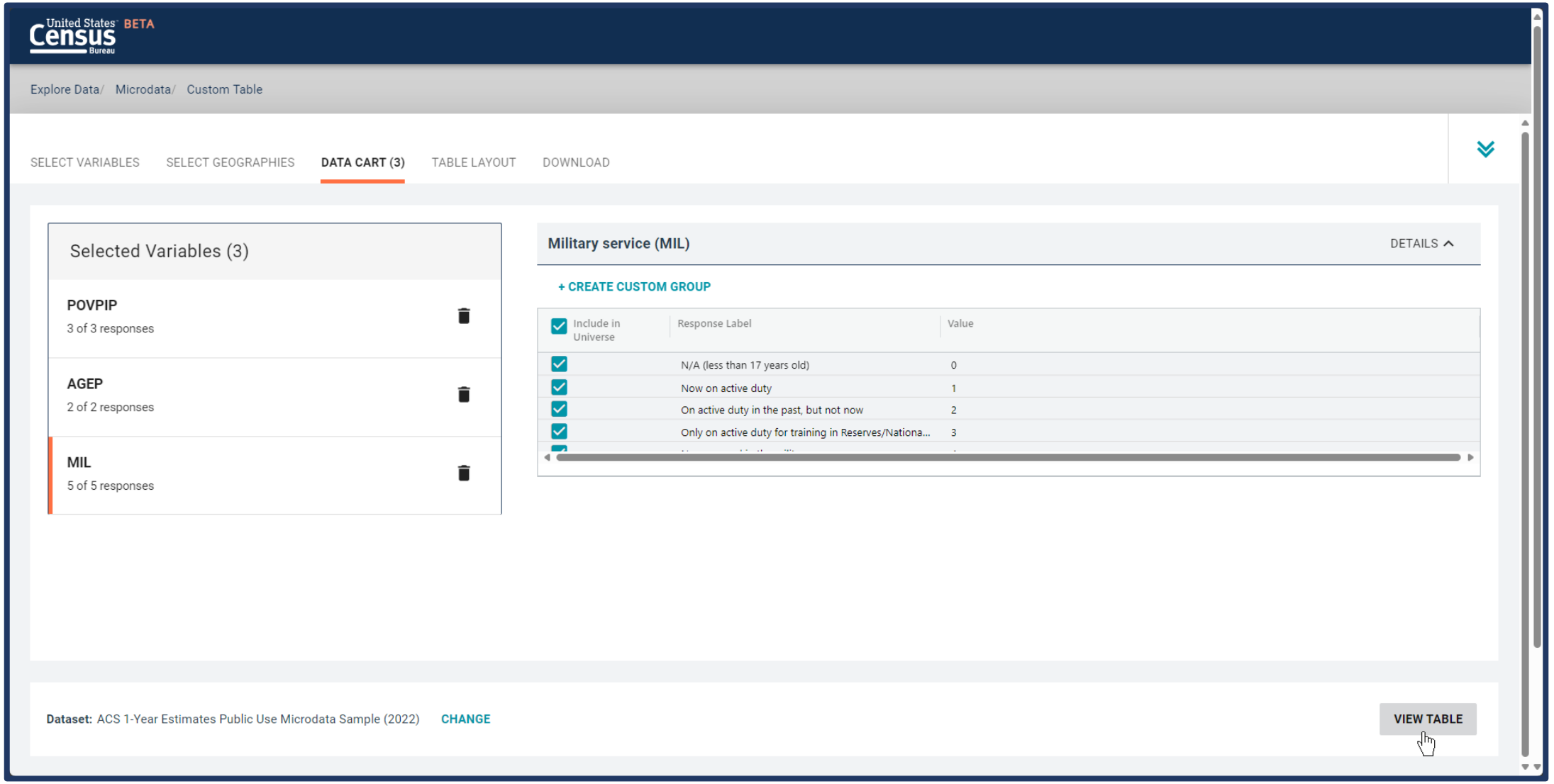
Beginning your PUMS journey



data.census.gov/mdat/#/

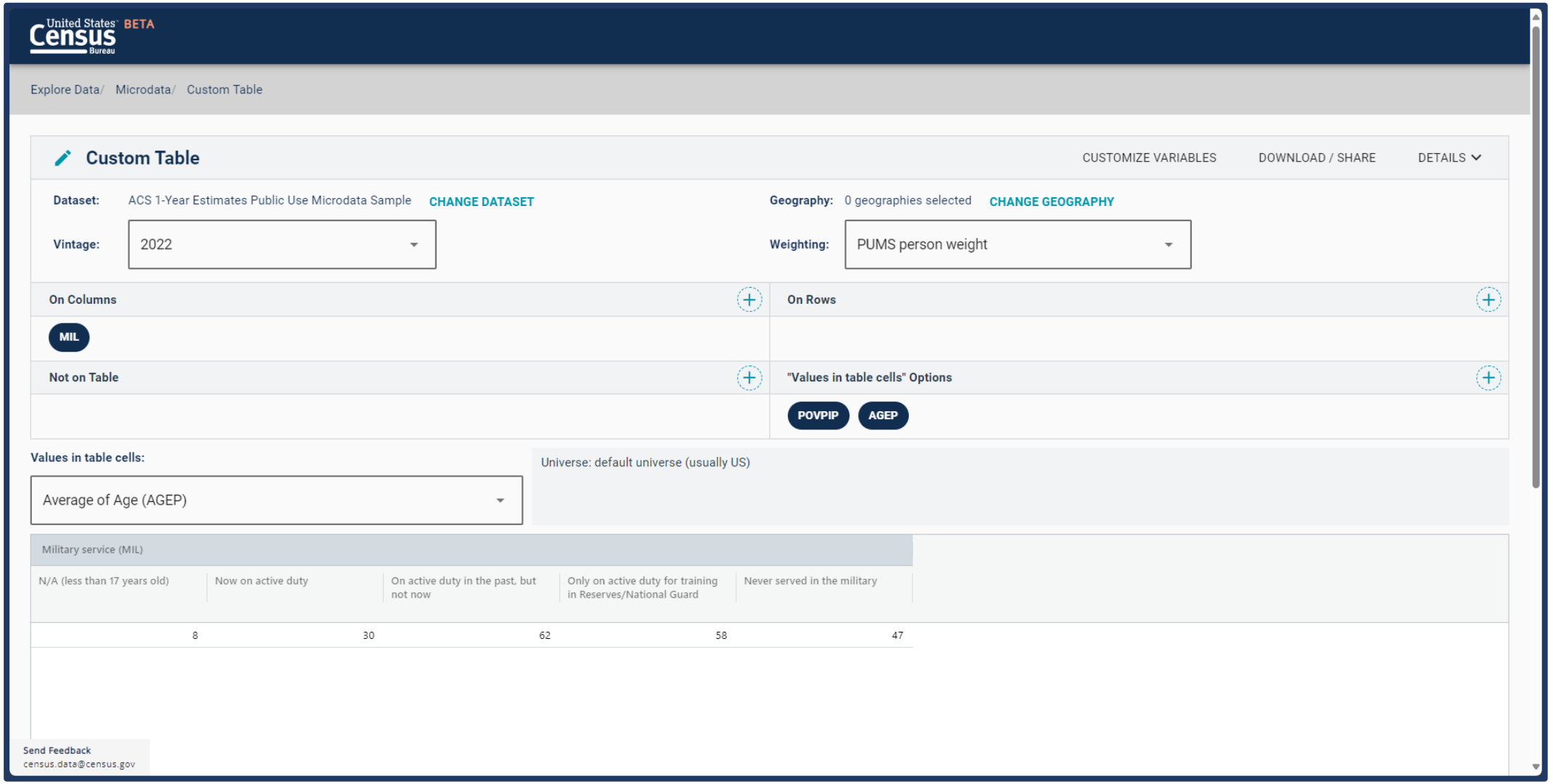
ACS Public Use Microdata Sample (PUMS)

Beginning your PUMS journey



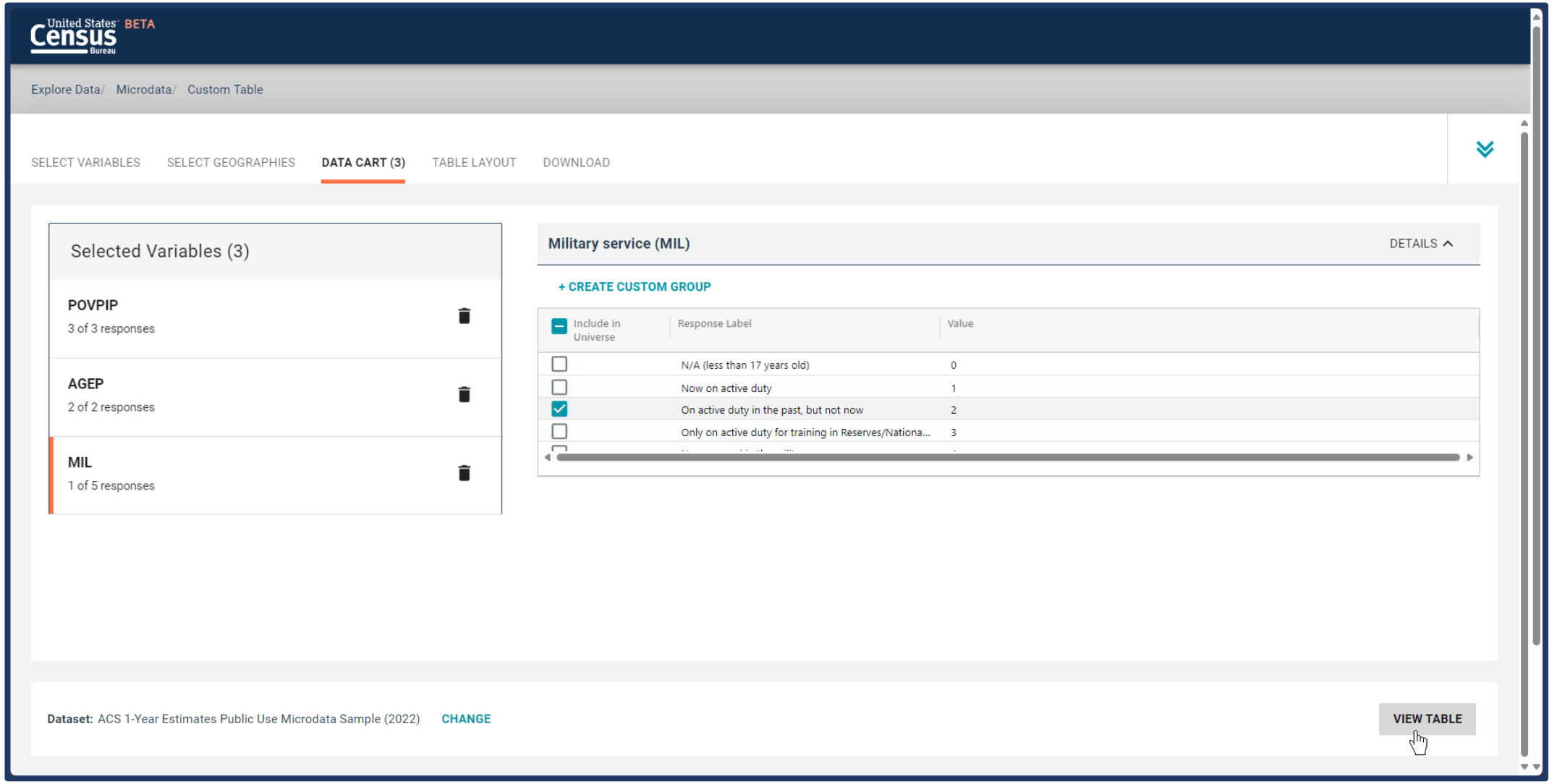
ACS Public Use Microdata Sample (PUMS)

Beginning your PUMS journey



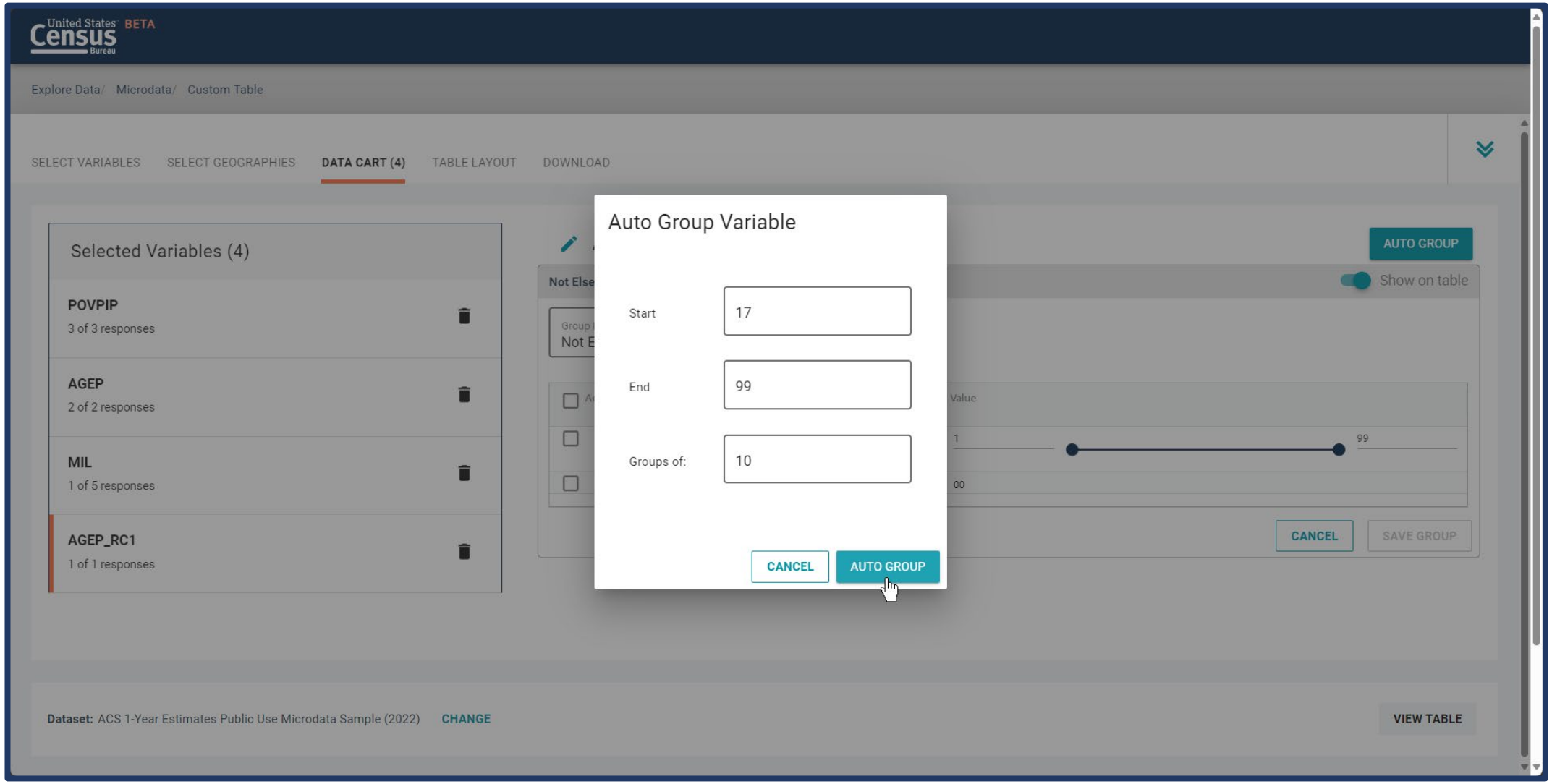
ACS Public Use Microdata Sample (PUMS)

Beginning your PUMS journey



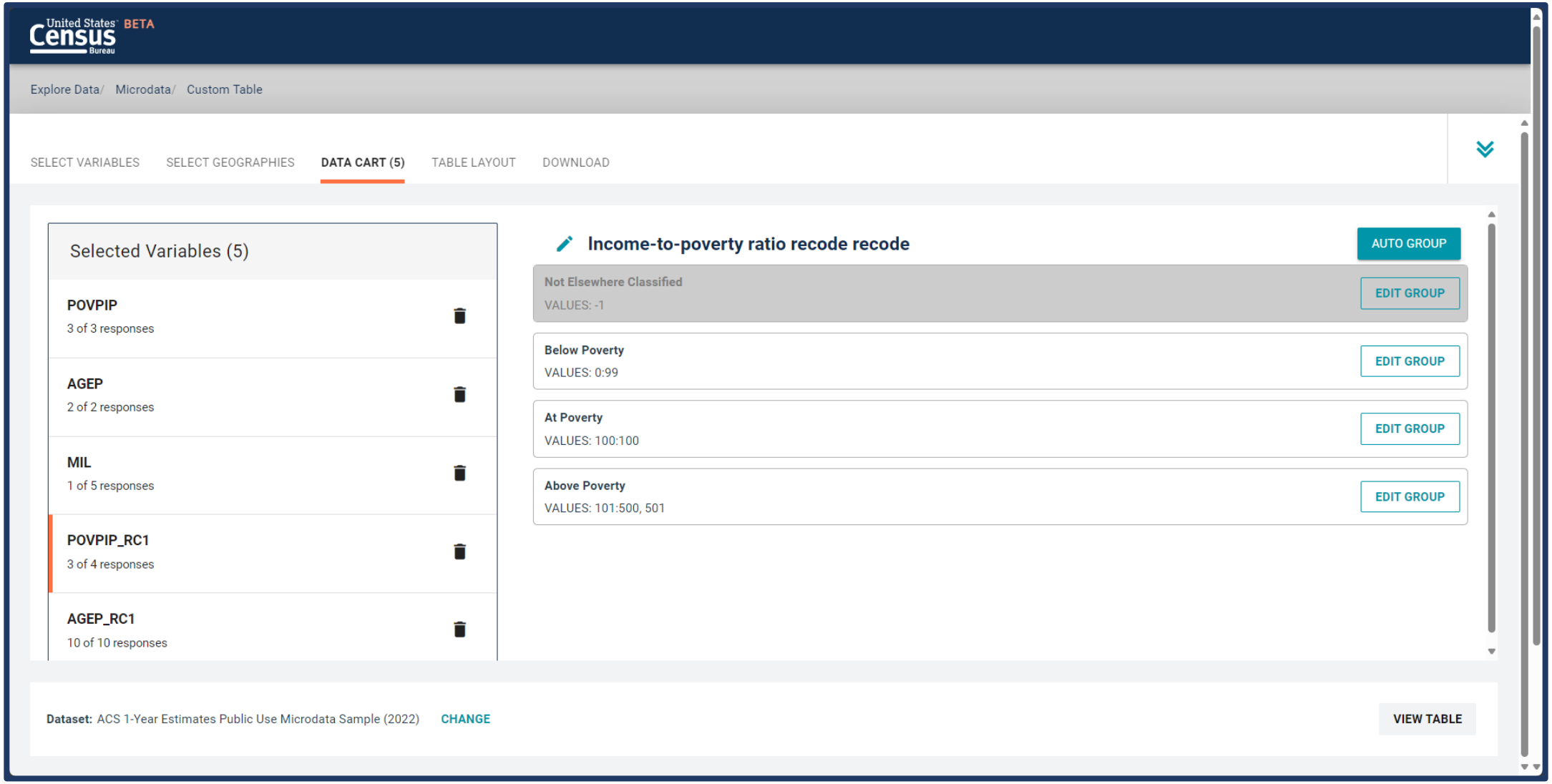
ACS Public Use Microdata Sample (PUMS)

Beginning your PUMS journey



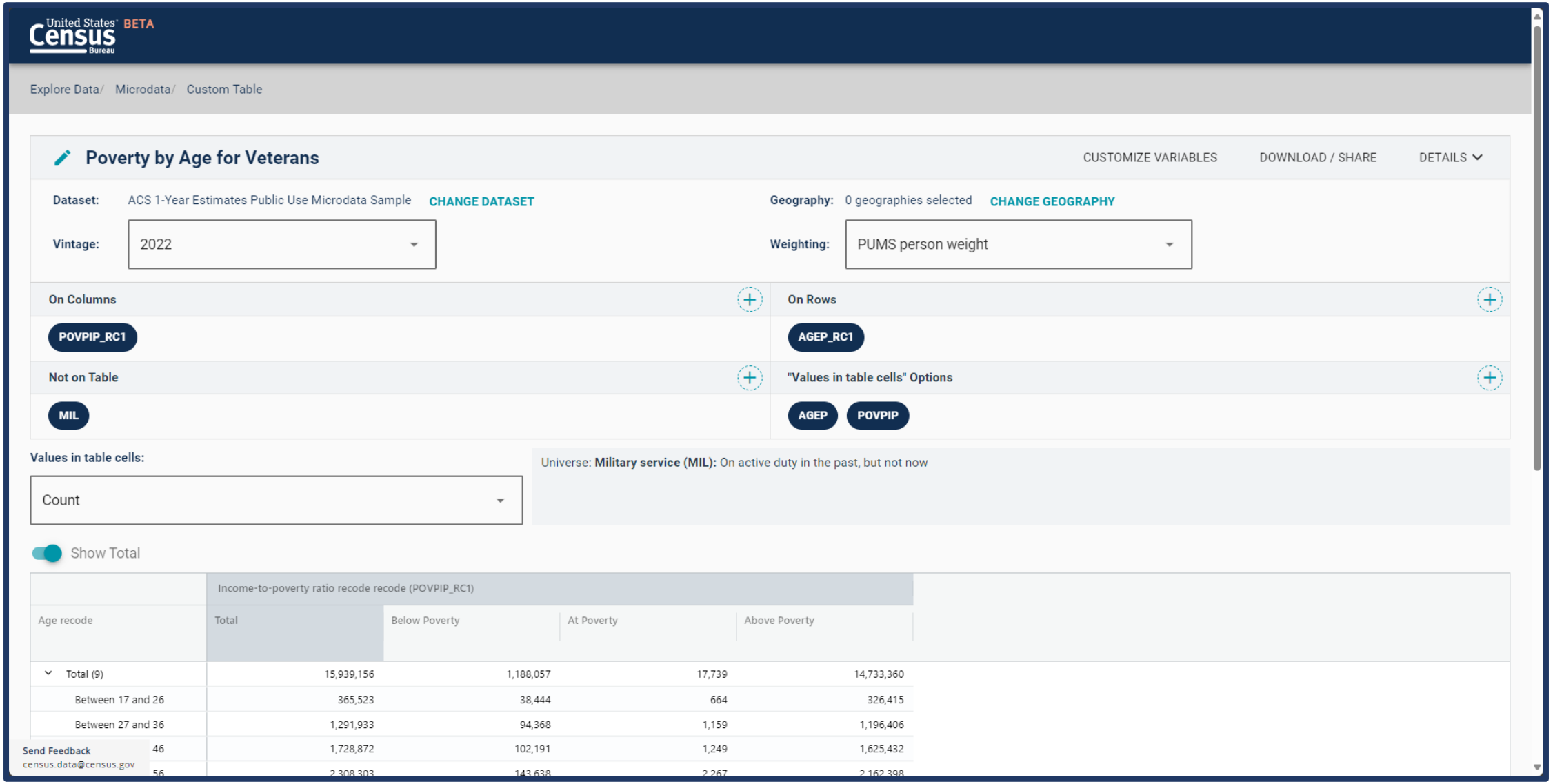
ACS Public Use Microdata Sample (PUMS)

Beginning your PUMS journey



ACS Public Use Microdata Sample (PUMS)

Beginning your PUMS journey



ACS Public Use Microdata Sample (PUMS)

Outline

- ACS PUMS basics
- Run through of Microdata Access Tool
- **PUMS Resources**

ACS Public Use Microdata Sample (PUMS)

Additional resources

Recorded videos/webinars:

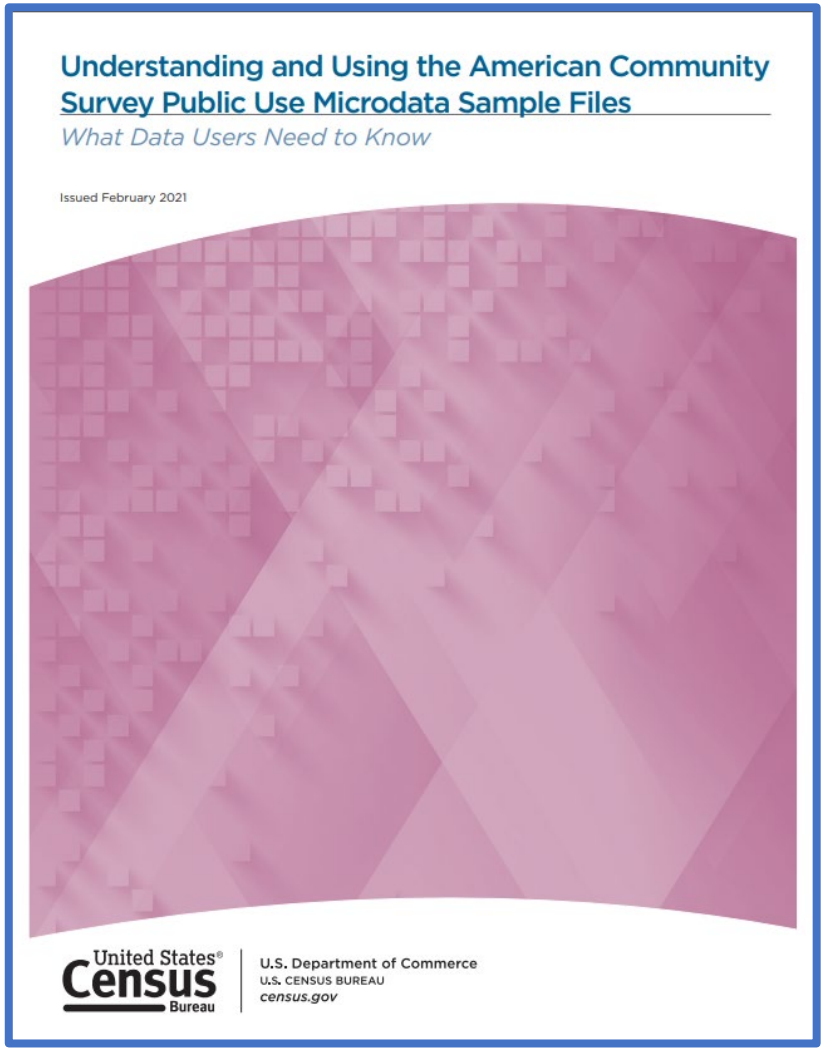
- Webinar: Introduction to ACS PUMS
- Data Gem: Building Custom Table using MDAT
- Data Gem: Learn First Steps to Create your Own Tabulation using Microdata
- Webinar: Using MDAT to Create Custom Tables

PUMS Documentation

- ACS PUMS Documentation webpage
- PUMS Data User Handbook

Additional ways to access PUMS

- PUMS on API



Introduction to the Census Bureau's Application Programming Interface (API)

Application Programming Interface (API)

Think about…

What are your main goals when accessing ACS data?

How do you access ACS data?

What limitations do you face accessing ACS data?

Application Programming Interface (API)

Outline

- **Application Programming Interface (API) basics**
- Run through of an API call
- API Resources

Application Programming Interface (API)

The Census API

- The Census Bureau Application Programming Interface (API)
  - Data Service which enables **Census Bureau programs** to provide data in a standardized way and hosted from a Central location
  - Gives **software and web developers** the ability to create and easily update custom applications
  - Allows **data scientists** to have more direct access to data in order to conduct analyses

Application Programming Interface (API)

Advantages

- Uses of the Census Bureau API include
  - Supporting mobile and web applications
  - Drives interactive data visualizations
  - Connects to statistical analysis software like R

Application Programming Interface (API)

Using the API

- Why use the API?
  - Need specific variables within a table in data.census.gov
  - Uniform geographies within a larger geography?
  - Need data for many geographies
  - Need easy way to update data year after year
  - “Live” connection to data (think: dashboard, online tool)
- **Before using API, consider data.census.gov**

Application Programming Interface (API)
Outline

- Application Programming Interface (API) basics
- **Run through of an API call**
- API Resources

Application Programming Interface (API)

Beginning your API journey

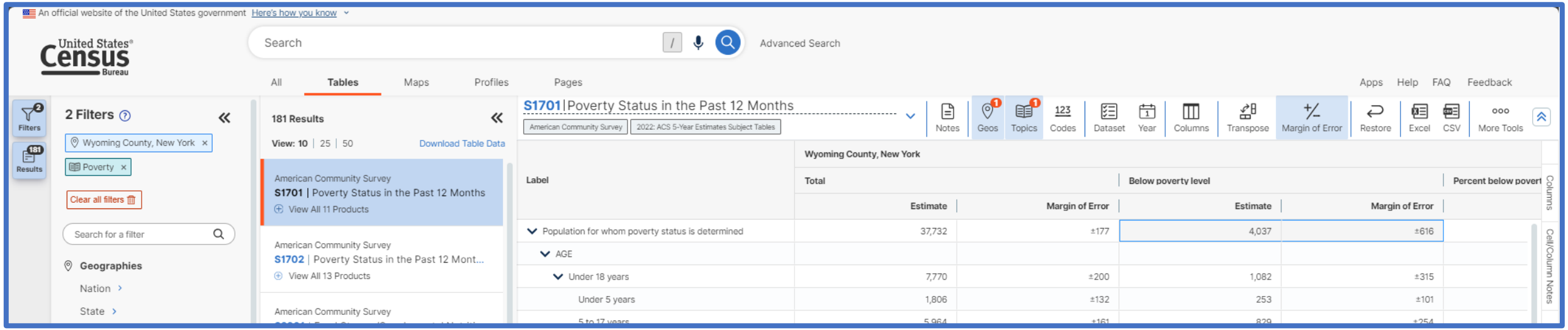
API Dataset	Table ID begins with	API call begins with
Detailed Tables	B or C	api.census.gov/data/2022/acs/acs1
Data Profiles	DP	api.census.gov/data/2022/acs/acs1**/profile**
Comparison Profiles	CP	api.census.gov/data/2022/acs/acs1**/cprofile**
Selected Population Profiles	S0201	api.census.gov/data/2022/acs/acs1**/spp**
Subject Tables	S	api.census.gov/data/2022/acs/acs1**/subject**

API Dataset	Table ID begins with	API call begins with
Detailed Tables	B or C	api.census.gov/data/2022/acs/acs5
Data Profiles	DP	api.census.gov/data/2022/acs/acs5**/profile**
Comparison Profiles	CP	api.census.gov/data/2022/acs/acs5**/cprofile**
Subject Tables	S	api.census.gov/data/2022/acs/acs5**/subject**

https://www.census.gov/programs-surveys/acs/data/data-tables.html

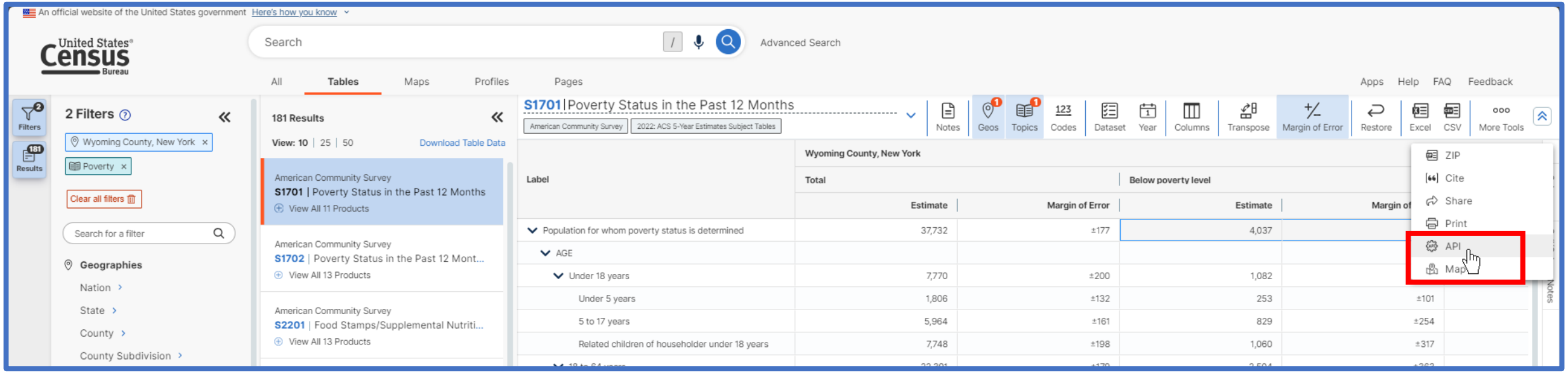
Application Programming Interface (API)

Beginning your API journey



Application Programming Interface (API)

Beginning your API journey



Application Programming Interface (API)

Anatomy of an API Query

https://api.census.gov/data/2022/acs/acs5/subject?get=NAME,S1701_C02_001E,S1701_C02_001M&for=county:121&in=state:36

We want to get the estimate and corresponding margin of error for the number of people living under the poverty level in Wyoming County, NY

Application Programming Interface (API)

Anatomy of an API Query

Base for all Census API queries

https://api.census.gov/data/2022/acs/acs5/subject?get=NAME,S1701_C02_001E,S1701_C02_001M&for=county:121&in=state:36

Application Programming Interface (API)

Anatomy of an API Query

Year and dataset (Subject Table)

https://api.census.gov/data/2022/acs/acs5/subject?get=NAME,S1701_C02_001E,S1701_C02_001M&for=county:121&in=state:36

We are using the 2018-2022 ACS 5-year Subject Tables

Application Programming Interface (API)

Anatomy of an API Query

The **?get=** portion is where you can specify the individual variables or the table that you want.

Using the **NAME** variable means that you will receive the name of the geography written out (e.g., "United States" or "New York city, New York").

Adding the **B27001_001E** variable means that we will get data for the first estimate in table B27001, which is the Total number of people in the civilian noninstitutionalized population.

https://api.census.gov/data/2022/acs/acs5/subject?get=NAME,S1701_C02_001E,S1701_C02_001M&for=county:121&in=state:36

Variable names (Subject Table)

What if you want all variables in the table? Use "group()"

https://api.census.gov/data/2022/acs/acs5/subject?get=group(S1701)&for=county:121&in=state:36

Application Programming Interface (API)

Anatomy of an API Query

The **&for=** portion indicates the geography being used.

https://api.census.gov/data/2022/acs/acs5/subject?get=NAME,S1701_C02_001E,S1701_C02_001M&for=county:121&in=state:36

Geography

What if you want all of one certain geography (ex. All counties in NY?) Use "*"

https://api.census.gov/data/2022/acs/acs5/subject?get=NAME,S1701_C02_001E,S1701_C02_001M&for=county:*&in=state:36

Application Programming Interface (API)

What does it look like?

https://api.census.gov/data/2022/acs/acs5/subject?get=NAME,S1701_C02_001E,S1701_C02_001M&for=county:121&in=state:36

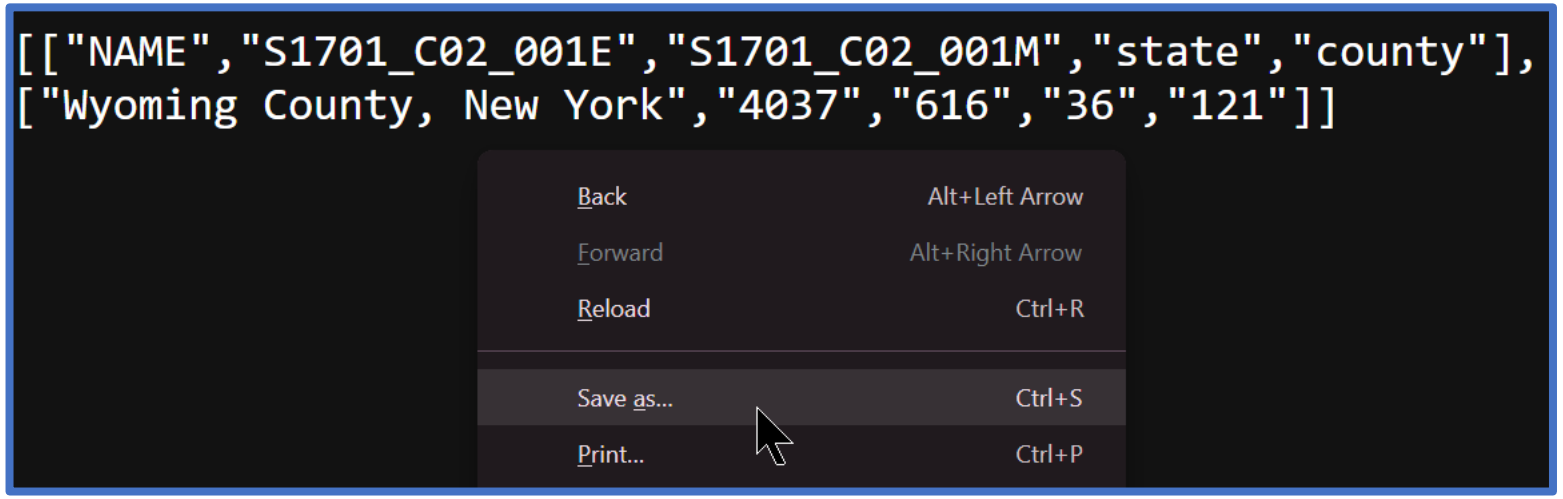

```
[
    [
        "NAME",
        "S1701_C02_001E",
        "S1701_C02_001M",
        "state",
        "county"
    ],
    [
        "Wyoming County, New York",
        "4037",
        "616",
        "36",
        "121"
    ]
]
```

```
[["NAME","S1701_C02_001E","S1701_C02_001M","state","county"],
["Wyoming County, New York","4037","616","36","121"]]
```

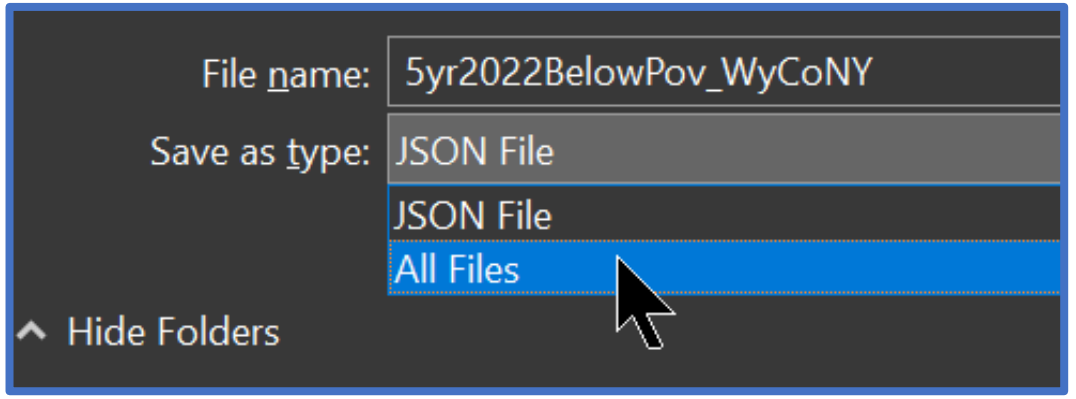
Application Programming Interface (API)

What to do next?

https://api.census.gov/data/2021/acs/acs5/subject?get=NAME,S1701_C02_001E,S1701_C02_001M&for=county:121&in=state:36



1. Right click-”Save as”
2. Name file, add “.csv” to end
3. Save as type “All files”



Application Programming Interface (API)
Outline

- Application Programming Interface (API) basics
- Run through of an API call
- **API Resources**

Application Programming Interface (API)

How can you make your own?

https://www.census.gov/data/developers/data-sets/acs-5year.html

Detailed Tables

- **Example Call:** api.census.gov/data/2022/acs/acs5?get=NAME,group(B01001)&for=us:1&key=YOUR_KEY_GOES_HERE
- 2022 ACS Detailed Tables Variables [html | xml | json]
- ACS Technical Documentation
- Examples
- Supported Geography

Subject Tables

- **Example Call:** api.census.gov/data/2022/acs/acs5/subject?get=NAME,group(S0101)&for=us:1&key=YOUR_KEY_GOES_HERE
- 2022 ACS Subject Tables Variables [html | xml | json]
- ACS Technical Documentation

https://api.census.gov/data/2022/acs/acs5/variables.html

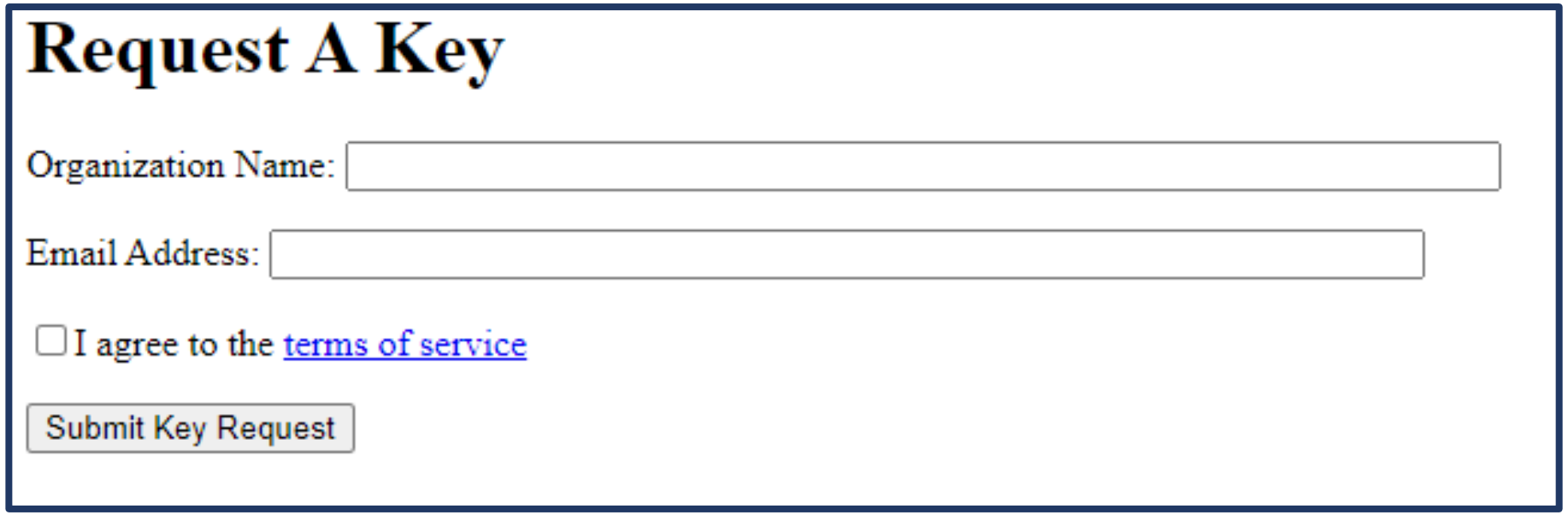
Census Data API: Variables in /data/2022/acs/acs5/variables

Name	Label	Concept	Required	Attributes	Limit	Predicate Type	Group
AIANHH	Geography		not required		0	*(not a predicate)*	N/A
AIHHTL	Geography		not required		0	*(not a predicate)*	N/A
AIRES	Geography		not required		0	*(not a predicate)*	N/A
ANRC	Geography		not required		0	*(not a predicate)*	N/A
B01001_001E	Estimate!!Total:	Sex by Age	not required	B01001_001EA, B01001_001M, B01001_001MA	0	int	B01001
B01001_002E	Estimate!!Total:!!Male:	Sex by Age	not required	B01001_002EA, B01001_002M, B01001_002MA	0	int	B01001
B01001_003E	Estimate!!Total:!!Male:!!Under 5 years	Sex by Age	not required	B01001_003EA, B01001_003M, B01001_003MA	0	int	B01001

Application Programming Interface (API)

What if you need help?

https://api.census.gov/data/key_signup.html



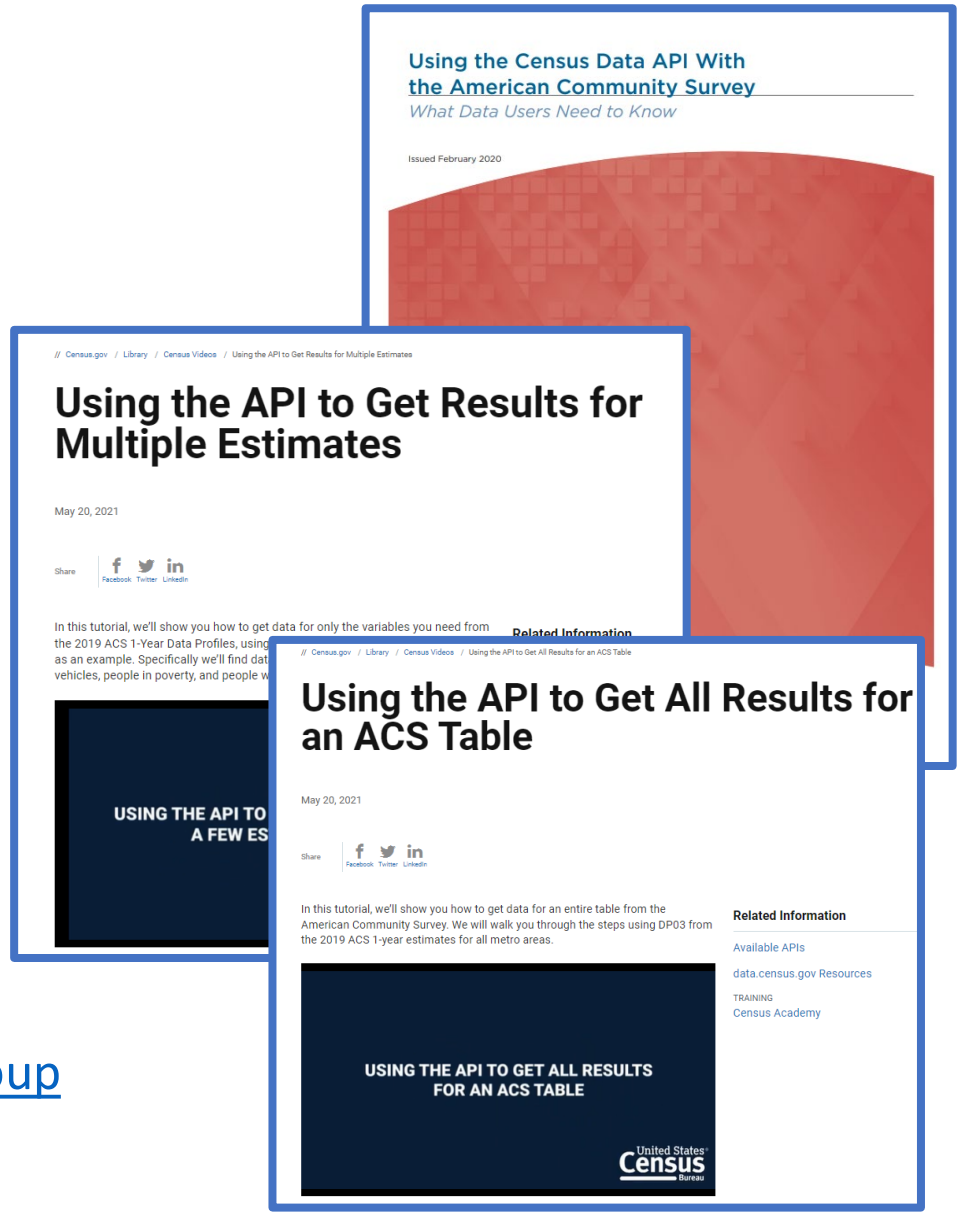
Application Programming Interface (API)
Additional resources

Recorded videos/webinars:

- Webinar: Demystifying the Census API
- Webinar: Using the Census API with the ACS
- Webinar: Getting Started with ACS Data in R and Python
- Webinar: Using API to Get All Results for an ACS Table
- Webinar: Using API to Get Results for Multiple Estimates

API webpages and documentation

- Webpage: ACS Data via API
- Webpage: Using ACS Data with Open-Source Software
- Webpage: How-to Materials for Using the Census API
- Document: ACS API Data Users Handbook
- Document: Census Data API User Guide
- Document: Removing ACS Annotation Columns from API Group Call in Excel



Application Programming Interface (API)

Additional support

U.S. Census Bureau Slack Channel

- Developer's forum to help improve access to public datasets from the U.S. Census Bureau
- Join U.S. Census Bureau Slack Channel

TidyCensus

- U.S. Census Bureau does not maintain or provide support
- Documentation and support: https://walker-data.com/tidycensus/

Additional support

Live Workshops

- Basics of Using MDAT
- Basics of Using the Census API

ACS Data Users Group

- Membership is free and open to all interested ACS data users.
- Discussion forum where you can explore PUMS and API topics or post your own questions and receive response from other data users!
- acsdatacommunity.prb.org

ACS User Support

- acso.users.support@census.gov