Microdata Access on data.census.gov

New York Data Users Meeting
June 5, 2024

Maria Valdisera
Center for Enterprise Dissemination (CED)
U.S. Census Bureau

United States® Census Bureau

Microdata = PUMS Files

Public **U**se **M**icrodata

Anonymized

- No personally identifiable information
- Edits to protect confidentiality

Individual Responses

- Must be tabulated and weighted by user

Accessible

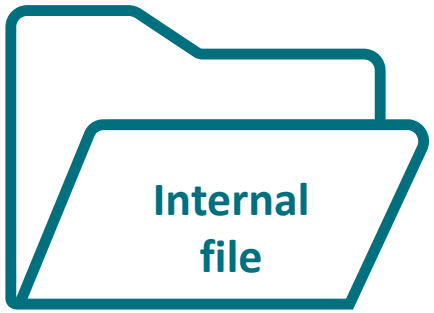
- data.census.gov/mdat
- Application Programming Interface (API)
- Download through FTP sites

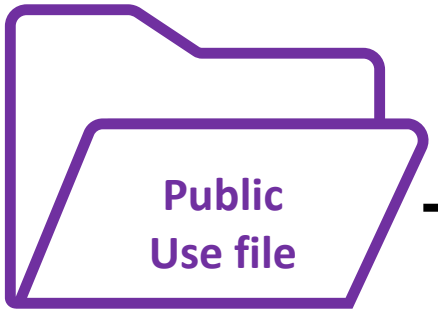
United States® Census Bureau

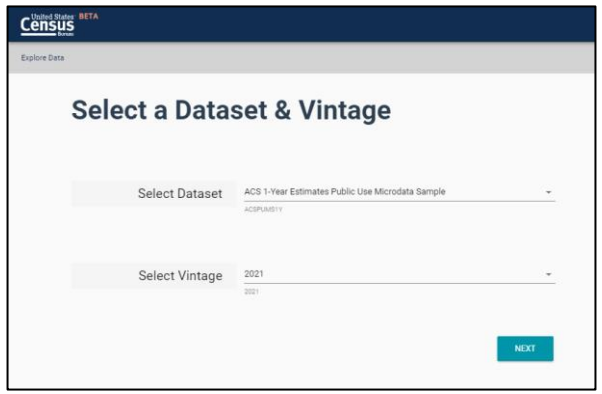
What's the difference between data.census.gov and Microdata Access?

Survey Responses (individual records)

Edits to protect confidentiality applied







data.census.gov

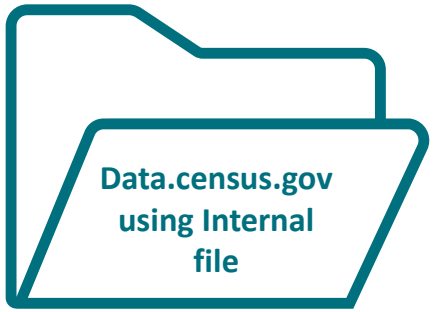
- Estimates are created using individual records that are only available to Census program area staff

Microdata Access (internally known as MDAT)

- Estimates are created using a sample of individual records that have been processed for use by the public

United States® Census Bureau

What are the pros and cons of using data.census.gov and Microdata Access?



Pros and Cons of Using Tables found in data.census.gov	
Pros: • Provides more precise estimates • Wider range of datasets • Fewer limitations to available geographies • No in-depth knowledge of variables required	Cons: • Limited to crosstabulations and tables that are predetermined by data providers • Limited ability to customize tables

Pros and Cons of Creating Tables in Microdata Access	
Pros: • Provides custom estimates when a pre-tabulated Census table is not available • More historical data available • Includes datasets not available in data.census.gov	Cons: • Limited geographies • Provides less precise estimates • Requires in-depth knowledge of variables • No margins of error provided

United States® Census Bureau

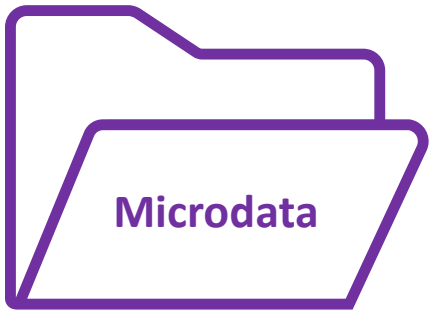
What's the difference between tabulated data and microdata?

	Maryland	
Label	Estimate	Margin of Error
Total:	3,098,870	±17,785
Male:	1,565,561	±11,667
Management, business, science, and arts occupations:	682,858	±11,323
Management, business, and financial occupations:	286,831	±7,906
Management occupations	195,401	±6,483
Business and financial operations occupations	91,430	±5,335
Computer, engineering, and science occupations:	212,203	±6,790
Computer and mathematical occupations	121,160	±5,830
Architecture and engineering occupations	54,967	±3,693

data.census.gov

Aggregated tables for a geography:

"In 2019 in Maryland, approximately 121,160 males worked in computer and mathematical occupations."



SERIALNO	SPORDER	ST	SEX	OCCP
2019HU0045422	4	24	1	4710
2019HU0045422	5	24	2	9
2019HU0045422	6	24	2	9
2019HU0045644	1	24	1	2100
2019HU0045764	1	24	2	5740
2019HU0045764	2	24	1	1031
2019HU0046210	1	24	1	150

Microdata Access (MDAT)

Microdata (a set of edited survey responses):

"This male in Maryland is a web developer."

United States® Census Bureau

Data Dictionaries

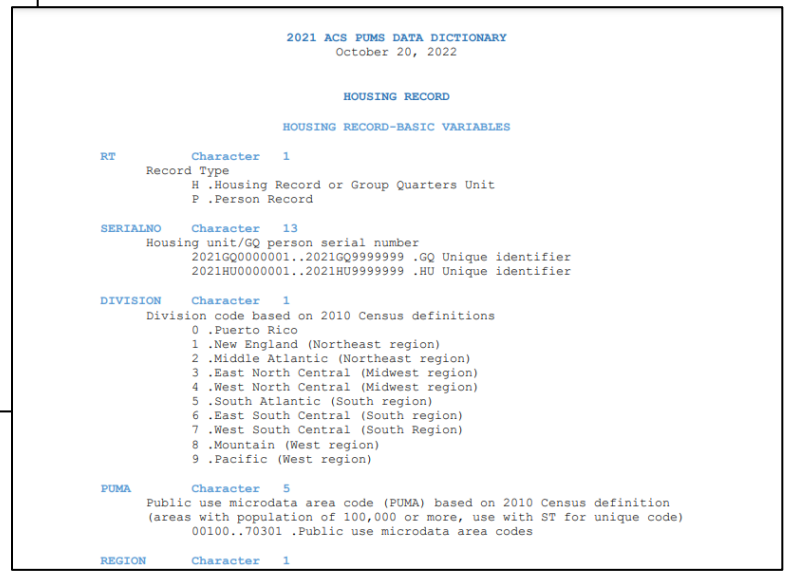
American Community Survey

https://www.census.gov/programs-surveys/acs/microdata/documentation.html

Data Dictionaries for PUMS Files

Includes variables available for each release of PUMS files on the Census Bureau FTP site, and how each variable is coded.

- 2021 ACS 1-year PUMS Data Dictionary [PDF] [<1.0 MB]
- 2021 ACS 1-year PUMS Data Dictionary [TXT] [<1.0 MB]
- 2021 ACS 1-year PUMS Data Dictionary [CSV] [<1.0 MB]
- 2017-2021 ACS 5-year PUMS Data Dictionary [PDF] [<1.0 MB]
- 2017-2021 ACS 5-year PUMS Data Dictionary [TXT] [<1.0 MB]
- 2017-2021 ACS 5-year PUMS Data Dictionary [CSV] [<1.0 MB]



2021 ACS PUMS DATA DICTIONARY
October 20, 2022

HOUSING RECORD

HOUSING RECORD-BASIC VARIABLES

RT Character 1
Record Type
H .Housing Record or Group Quarters Unit
P .Person Record

SERIALNO Character 13
Housing unit/GQ person serial number
2021GQ0000001..2021GQ9999999 .GQ Unique identifier
2021HU0000001..2021HU9999999 .HU Unique identifier

DIVISION Character 1
Division code based on 2010 Census definitions
0 .Puerto Rico
1 .New England (Northeast region)
2 .Middle Atlantic (Northeast region)
3 .East North Central (Midwest region)
4 .West North Central (Midwest region)
5 .South Atlantic (South region)
6 .East South Central (South region)
7 .West South Central (South Region)
8 .Mountain (West region)
9 .Pacific (West region)

PUMA Character 5
Public use microdata area code (PUMA) based on 2010 Census definition
(areas with population of 100,000 or more, use with ST for unique code)
00100..70301 .Public use microdata area codes

REGION Character 1

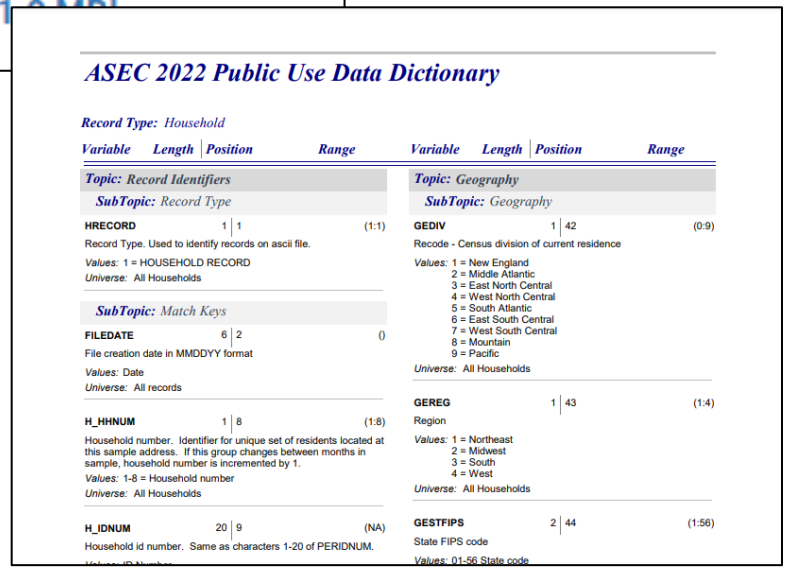
United States® Census Bureau

Current Population Survey Annual Social and Economic Supplement (CPS ASEC)

https://www.census.gov/data/datasets/2022/demo/cps/cps-asec-2022.html

Data and Documents

Data Dictionary [1



ASEC 2022 Public Use Data Dictionary

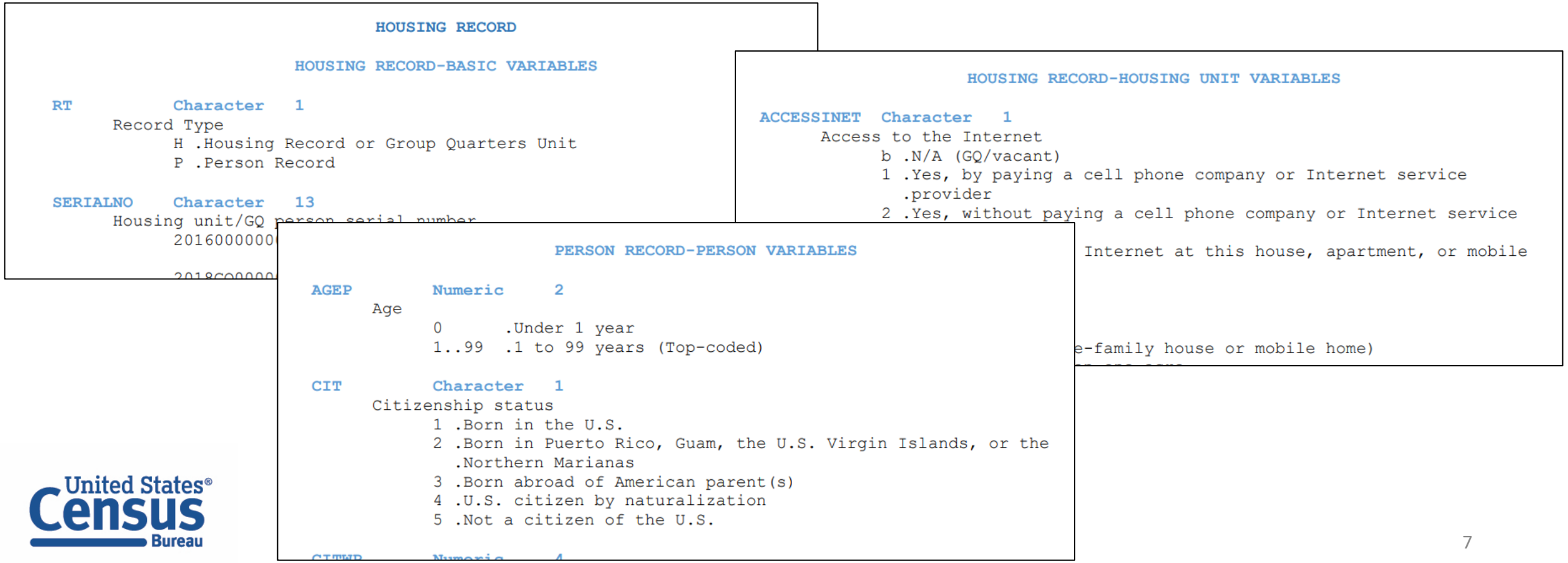
Record Type: Household

Variable	Length	Position	Range	Variable	Length	Position	Range
Topic: Record Identifiers				*Topic:* Geography			
SubTopic: Record Type				*SubTopic:* Geography			
HRECORD	1	1	(1:1)	GEDIV	1	42	(0:9)
Record Type. Used to identify records on ascii file.				Recode - Census division of current residence			
Values: 1 = HOUSEHOLD RECORD				*Values:* 1 = New England 2 = Middle Atlantic 3 = East North Central 4 = West North Central 5 = South Atlantic 6 = East South Central 7 = West South Central 8 = Mountain 9 = Pacific			
Universe: All Households				*Universe:* All Households			
SubTopic: Match Keys							
FILEDATE	6	2	()	GEREG	1	43	(1:4)
File creation date in MMDDYY format				Region			
Values: Date				*Values:* 1 = Northeast 2 = Midwest 3 = South 4 = West			
Universe: All records				*Universe:* All Households			
H_HHNUM	1	8	(1:8)				
Household number. Identifier for unique set of residents located at this sample address. If this group changes between months in sample, household number is incremented by 1.							
Values: 1-8 = Household number							
Universe: All Households							
H_IDNUM	20	9	(NA)	GESTFIPS	2	44	(1:56)
Household id number. Same as characters 1-20 of PERIDNUM.				State FIPS code			
				Values: 01-56 State code			

Data Dictionaries

American Community Survey
https://www.census.gov/programs-surveys/acs/microdata/documentation.html

The ACS PUMS data dictionary is broken out into different sections of variables, including basic variables, housing unit variables, and person variables.



HOUSING RECORD

HOUSING RECORD-BASIC VARIABLES

RT Character 1
Record Type
H .Housing Record or Group Quarters Unit
P .Person Record

SERIALNO Character 13
Housing unit/GQ person serial number
2016000000(
2018GQ0000(

HOUSING RECORD-HOUSING UNIT VARIABLES

ACCESSINET Character 1
Access to the Internet
b .N/A (GQ/vacant)
1 .Yes, by paying a cell phone company or Internet service
.provider
2 .Yes, without paying a cell phone company or Internet service

Internet at this house, apartment, or mobile

e-family house or mobile home)

PERSON RECORD-PERSON VARIABLES

AGEP Numeric 2
Age
0 .Under 1 year
1..99 .1 to 99 years (Top-coded)

CIT Character 1
Citizenship status
1 .Born in the U.S.
2 .Born in Puerto Rico, Guam, the U.S. Virgin Islands, or the
.Northern Marianas
3 .Born abroad of American parent(s)
4 .U.S. citizen by naturalization
5 .Not a citizen of the U.S.

CITWP Numeric 4

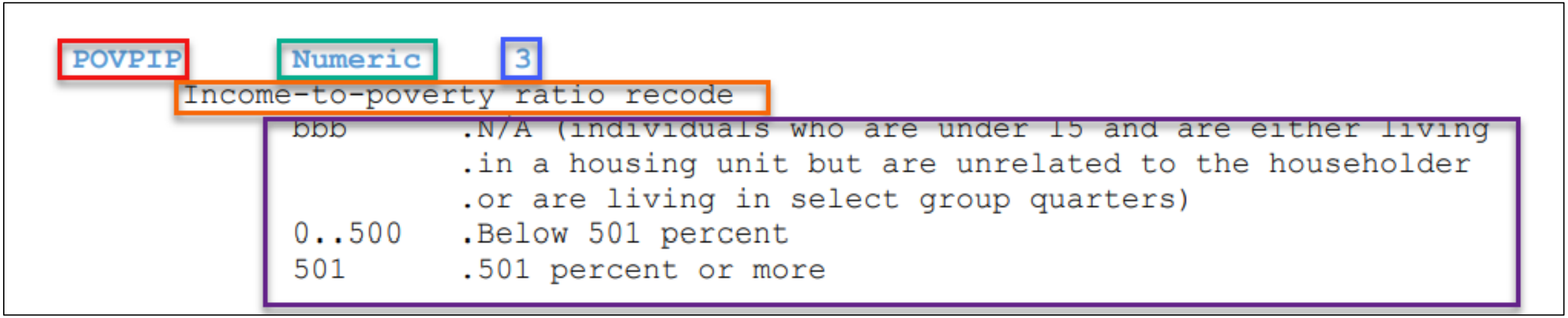
United States® Census Bureau

Data Dictionaries

American Community Survey
https://www.census.gov/programs-surveys/acs/microdata/documentation.html

Find all the variables that are available in the PUMS dataset for any given year.

The dictionary will give you the name of the variable, whether it's a character or numeric variable, the length of the variable, a brief description of the variable, and the possible response options or recoded values.



United States® Census Bureau

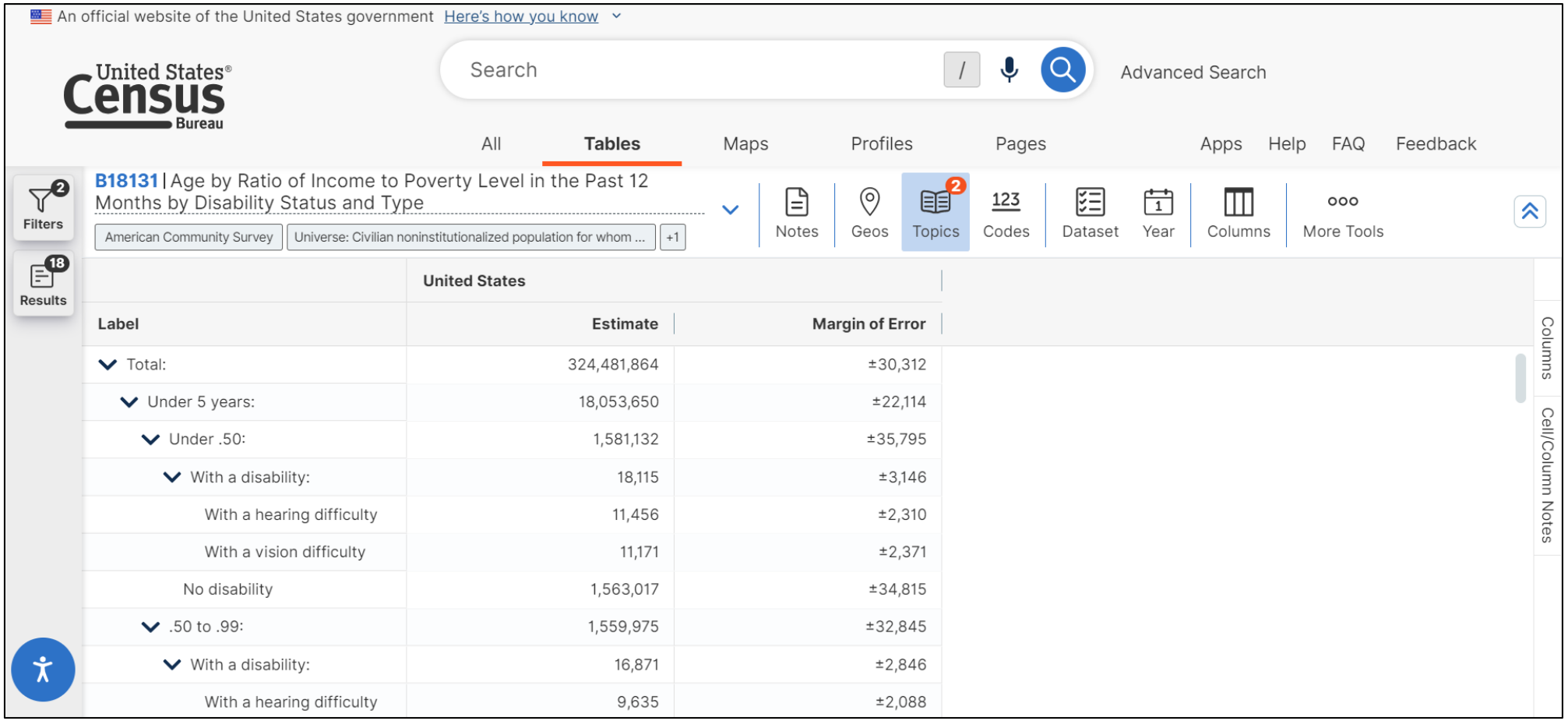
Demo

Example 1:

Poverty status by disability in New York state

United States® Census Bureau

Table B18131 – Age by Ratio of Income to Poverty Level in the Past 12 Months by Disability Status and Type



An official website of the United States government Here's how you know

United States Census Bureau

Search / Advanced Search

All | Tables | Maps | Profiles | Pages | Apps | Help | FAQ | Feedback

Filters 2 | Results 18

B18131 | Age by Ratio of Income to Poverty Level in the Past 12 Months by Disability Status and Type

American Community Survey | Universe: Civilian noninstitutionalized population for whom ... | +1

Notes | Geos | Topics 2 | Codes | Dataset | Year | Columns | More Tools

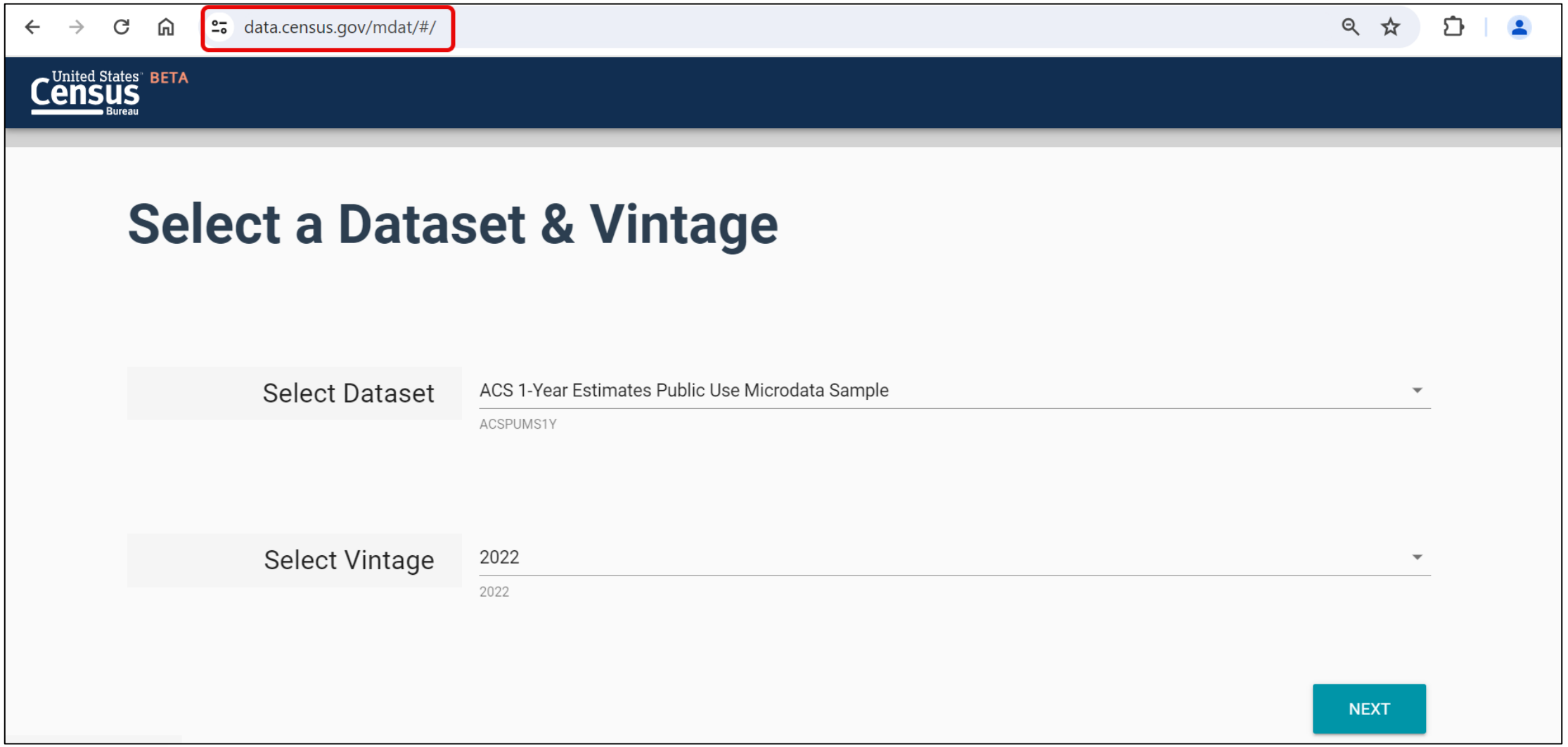
Label	United States Estimate	Margin of Error
Total:	324,481,864	±30,312
Under 5 years:	18,053,650	±22,114
Under .50:	1,581,132	±35,795
With a disability:	18,115	±3,146
With a hearing difficulty	11,456	±2,310
With a vision difficulty	11,171	±2,371
No disability	1,563,017	±34,815
.50 to .99:	1,559,975	±32,845
With a disability:	16,871	±2,846
With a hearing difficulty	9,635	±2,088

Columns | Cell/Column Notes

Prefabricated ACS tables poverty ratios by disability, but what if we need a different poverty break?

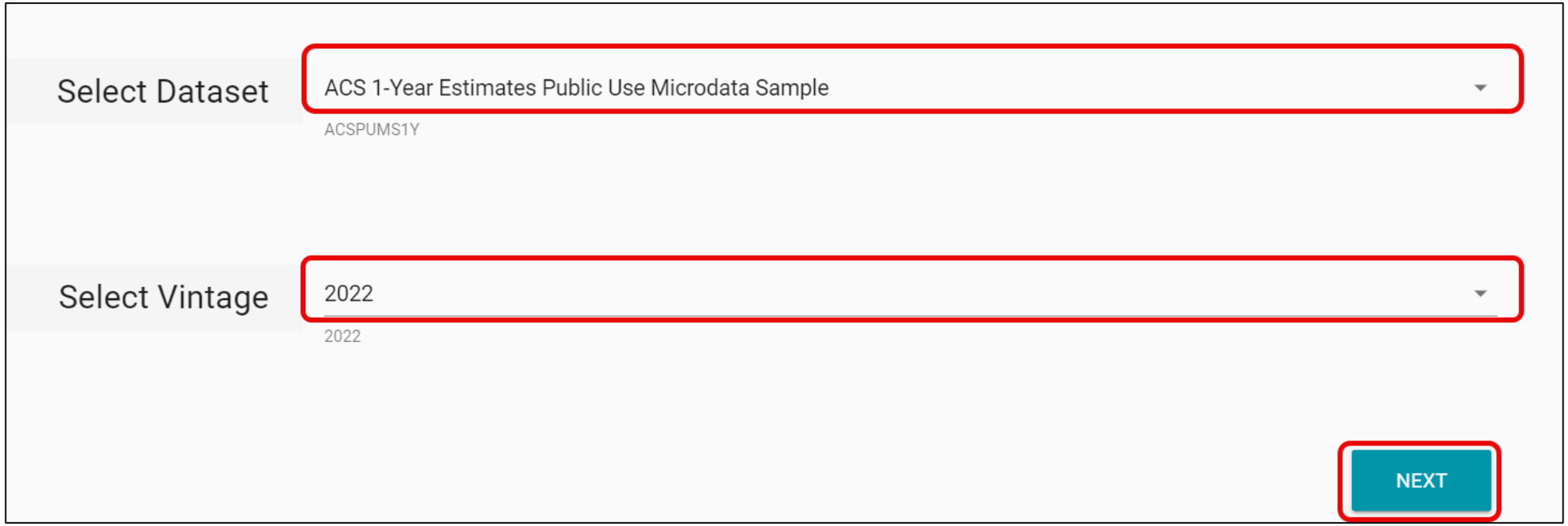
United States Census Bureau

- Visit Microdata Access at data.census.gov/mdat



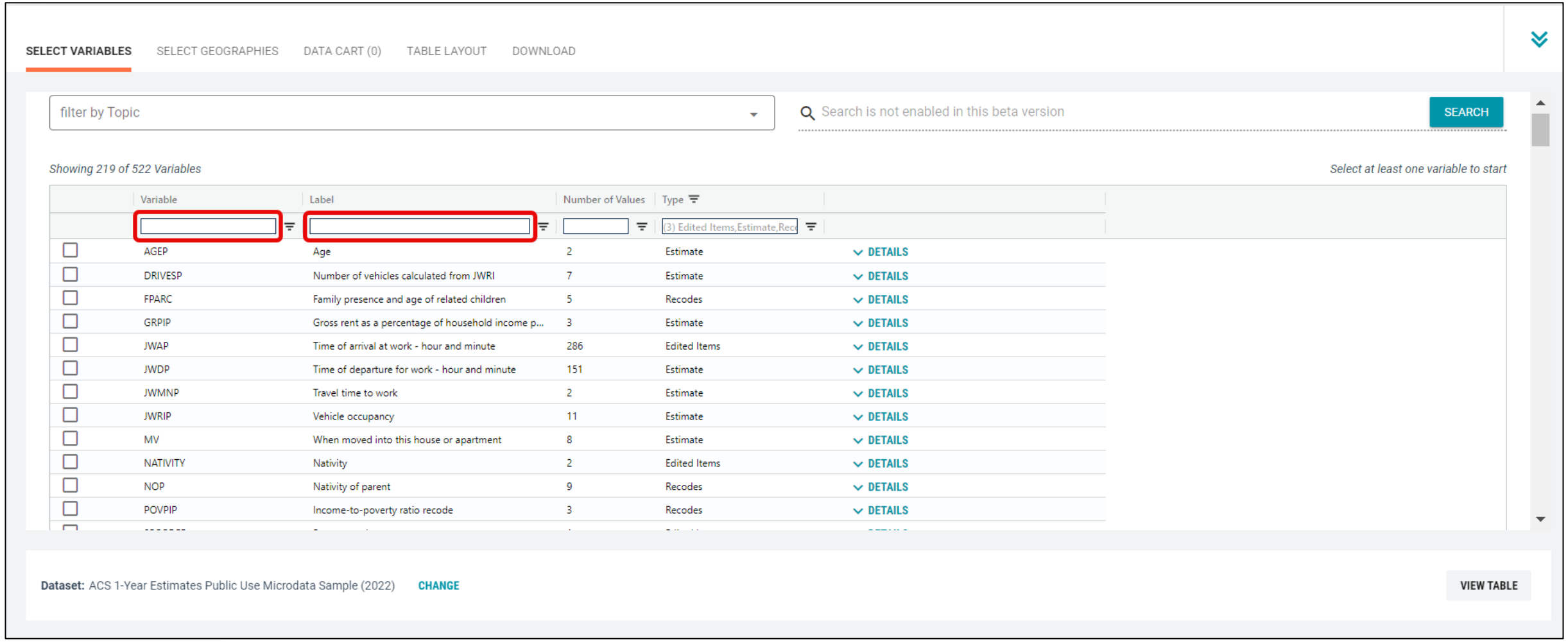
data.census.gov/mdat

- Choose Dataset and Vintage:
  - Dataset – **ACS 1-Year Estimates – Public Use Microdata Sample**
  - Vintage – **2022**
  - Click **Next** in the lower right

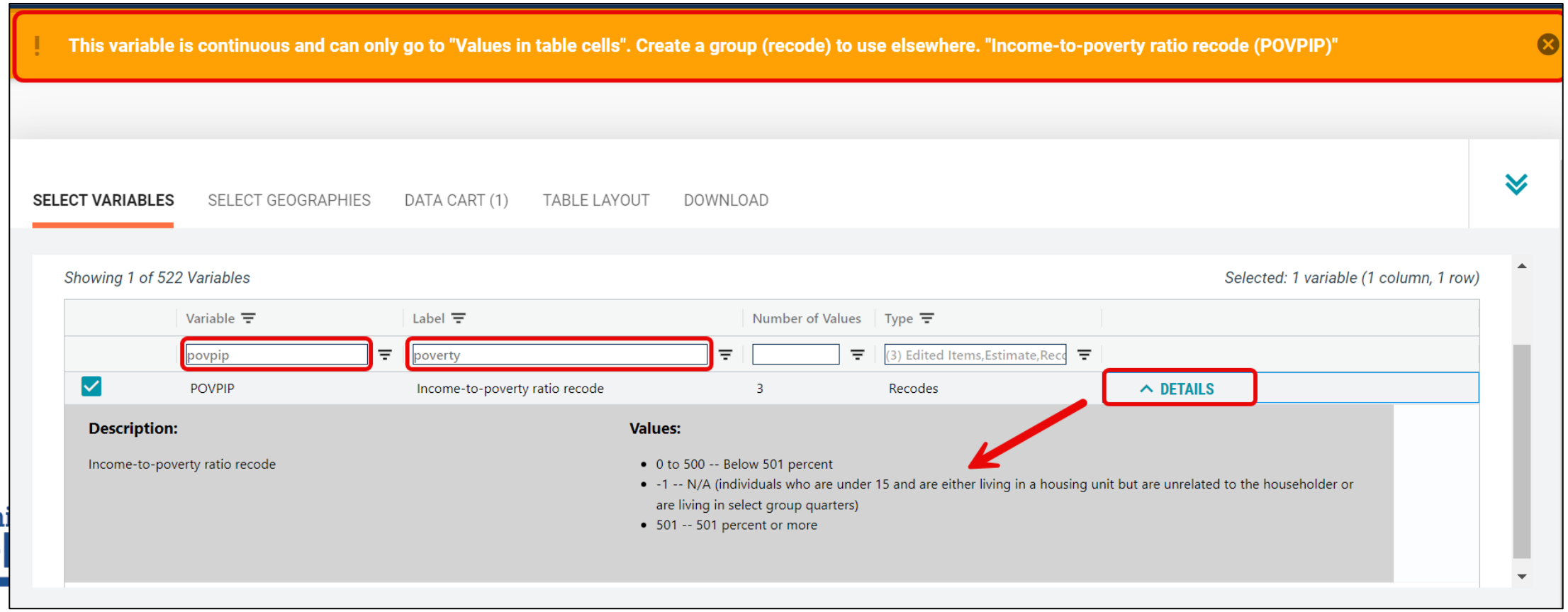


United States® Census Bureau

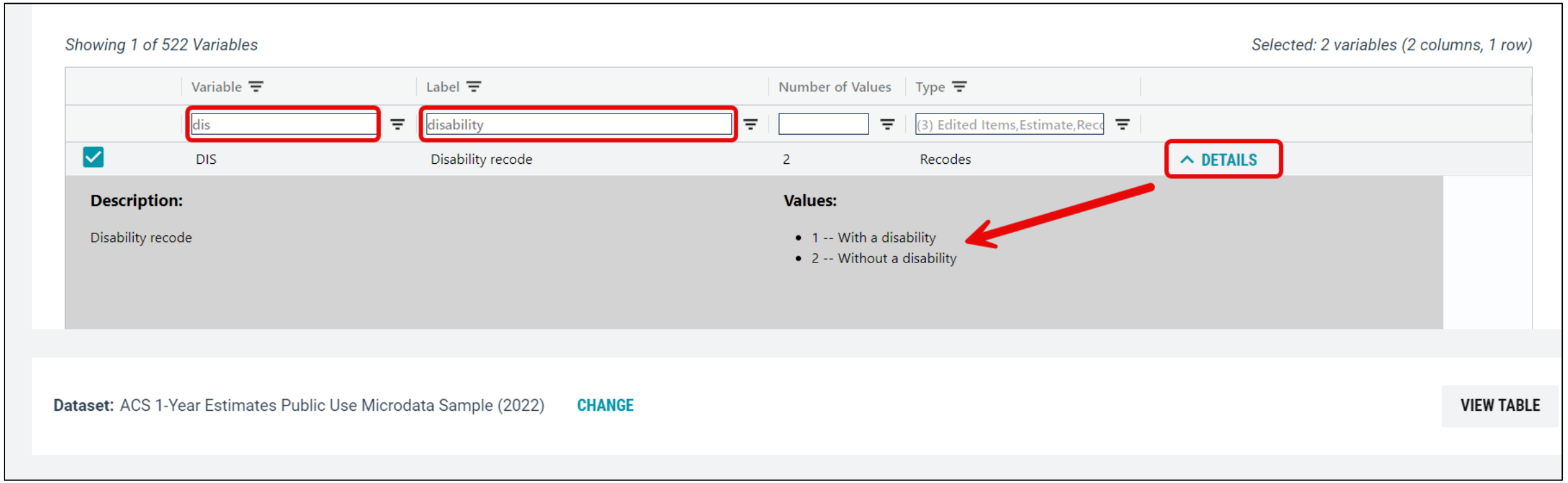
- **Search for Variables**: Use the search box below “Variable” or “Label” to find your variables of interest



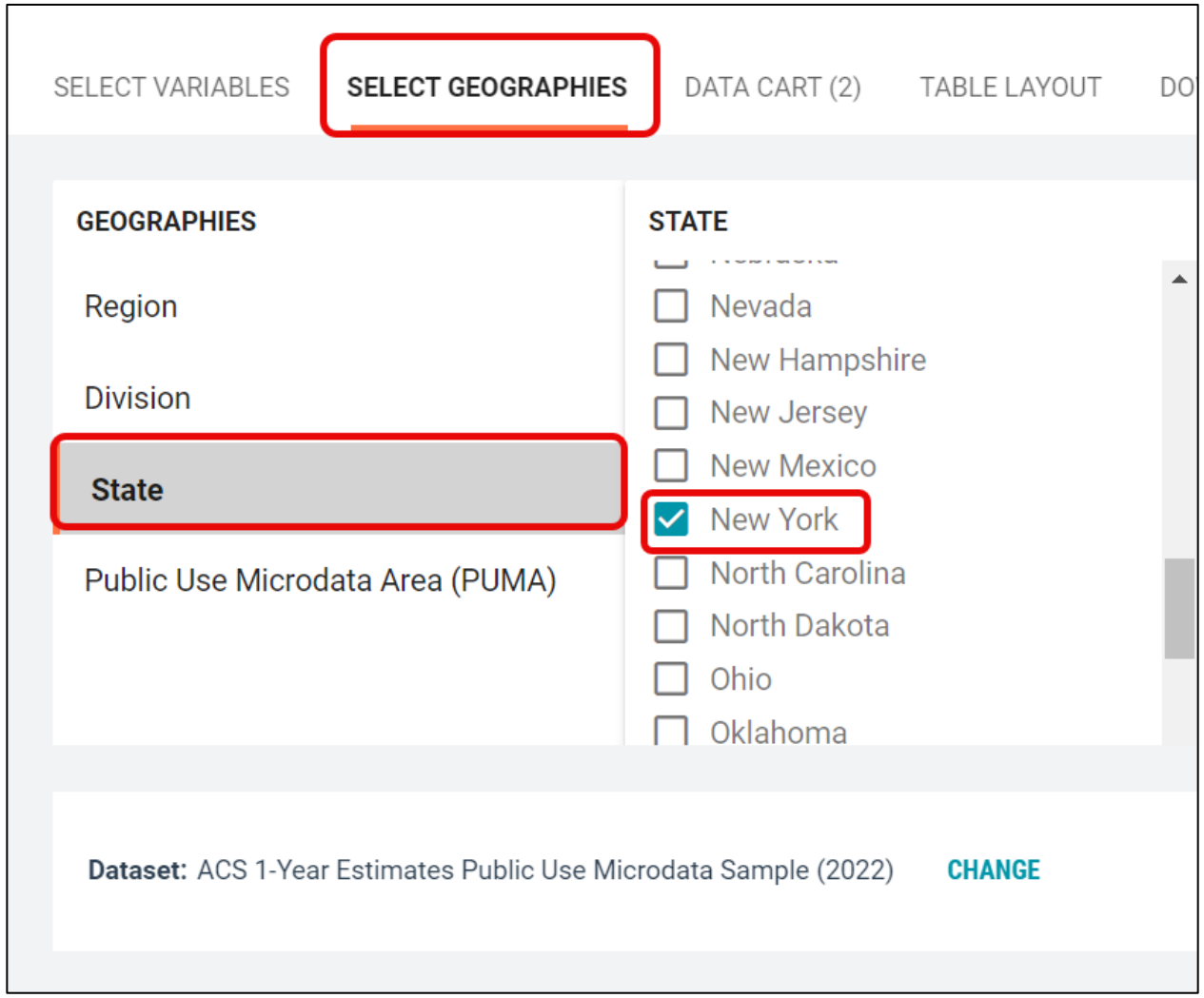
- **Select variable for Income to Poverty Ratio**:
  - Type “POVPIP” in the Variable search box or type “poverty” in the label search box
  - Check the box to the left of POVPIP to add the variable to your data cart
  - Notice the message at the top of the screen saying you will need to create your own categories (or recodes) for this variable if you want it shown in the table. (You will do this action in the Data Cart)



- **Select variable for Disability**:
  - Type “DIS” in the Variable search box or type “Disability” in the label search box
  - Check the box to the left of DIS to add the variable to data cart

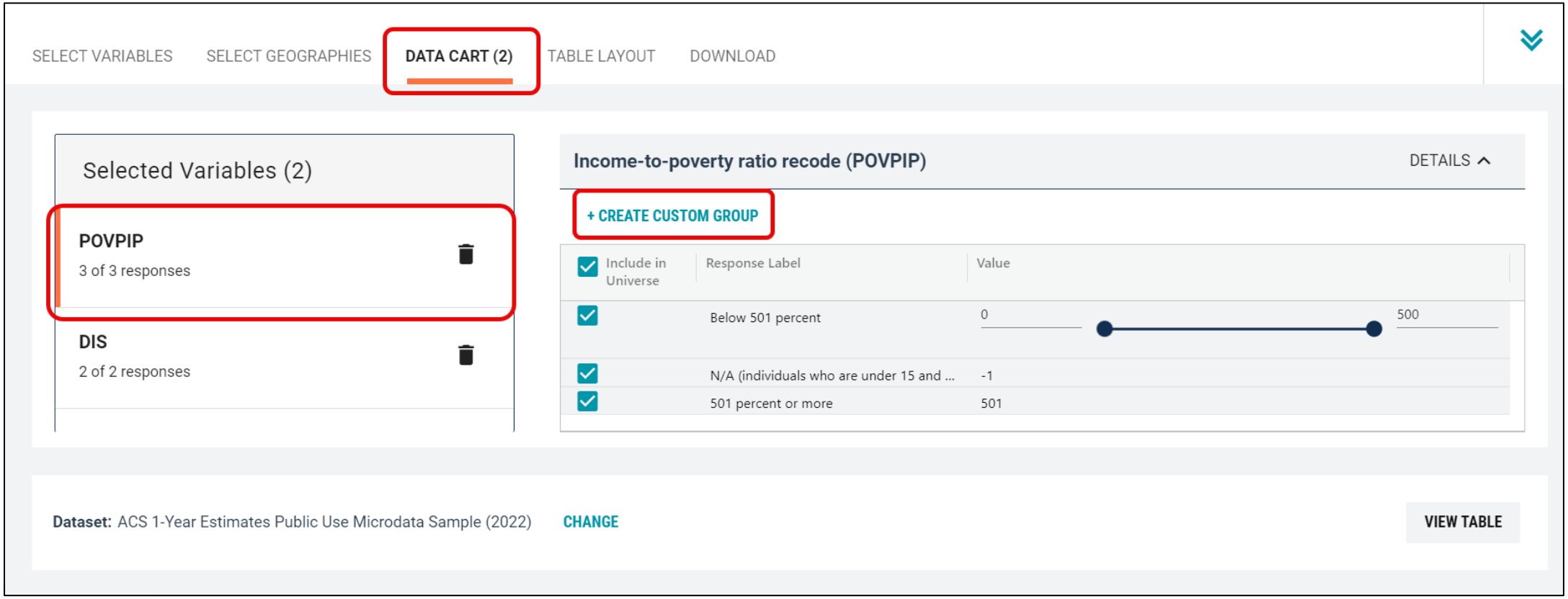


- Select geography:
  - Move to the **Select Geographies** tab
  - Click **State** and click on **New York**

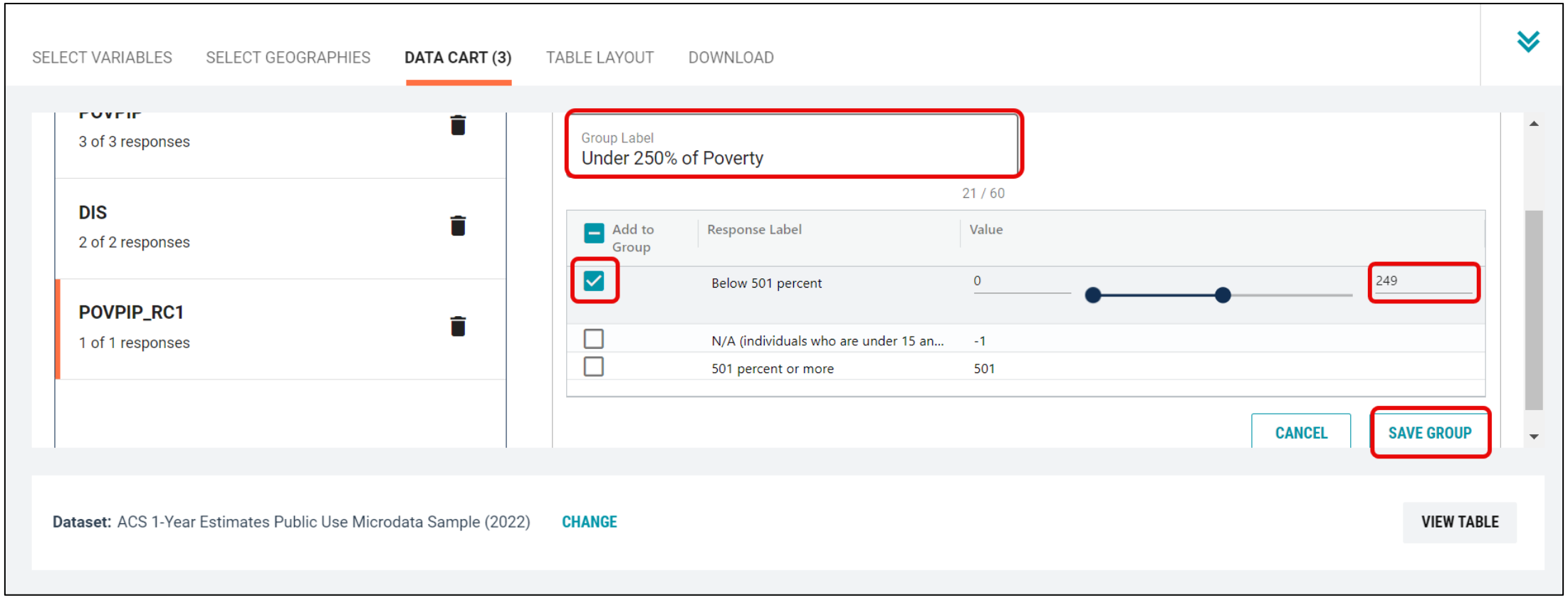


United States® Census Bureau

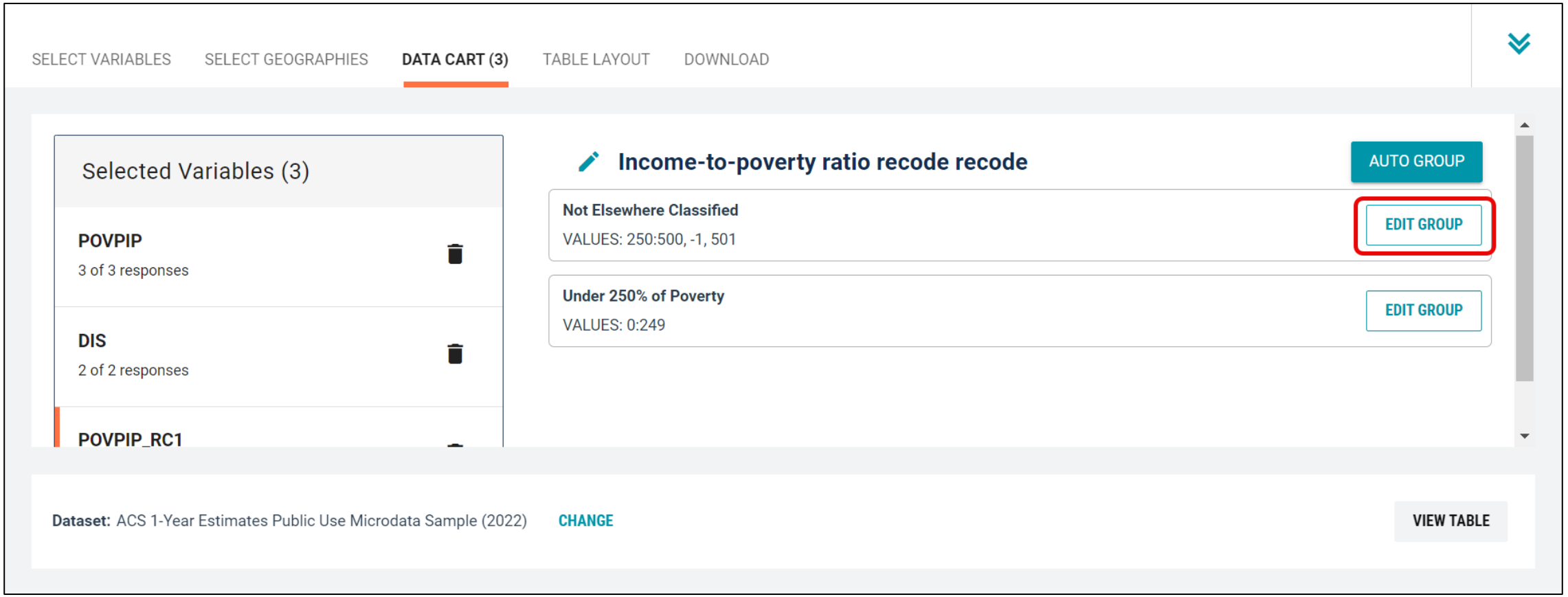
- Create recode for POVPIP variable:
  - Move to the **Data Cart** tab
  - Click the **POVPIP** variable on the left
  - Select '**Create Custom Group**' to create recode for custom income to poverty ratio categories



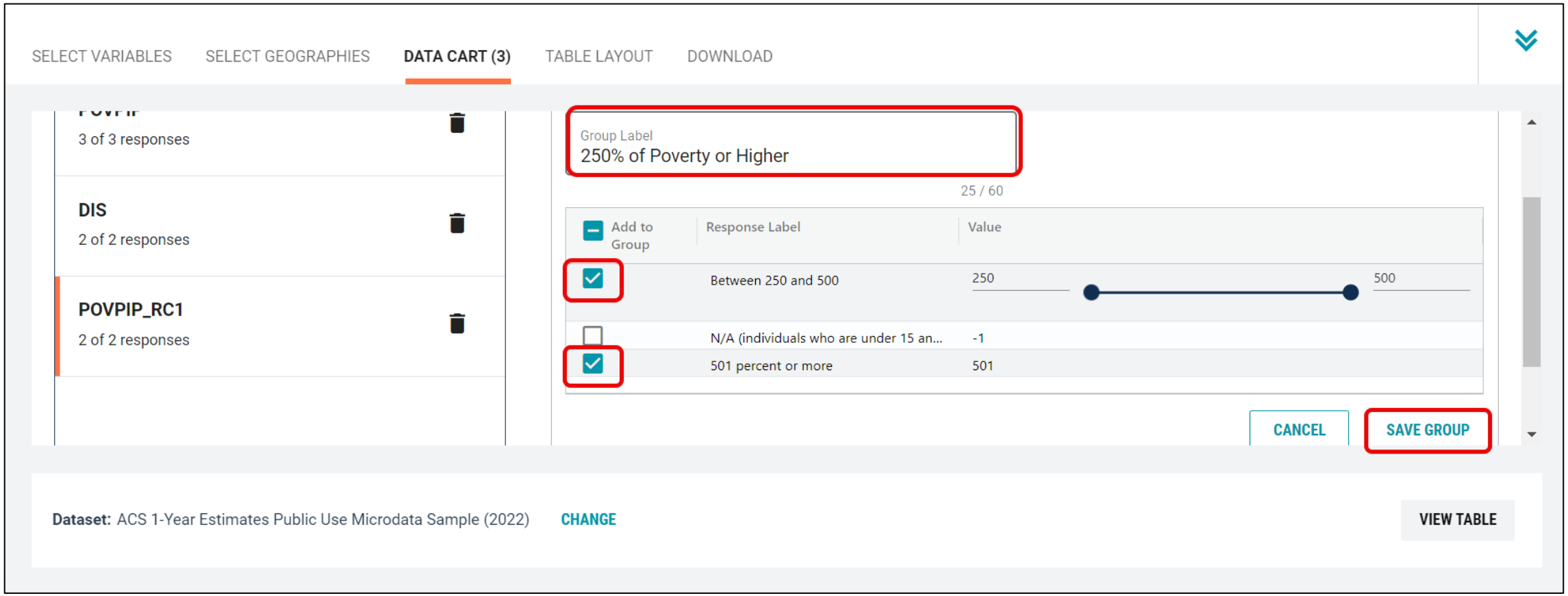
- Create recode for POVPIP variable:
  - Change Group Label to '**Under 250% of Poverty**'.
  - Click on the checkbox next to '**Below 501 percent**' and edit end range from 500 to **249**.
  - Click on the '**Save Group**' button.



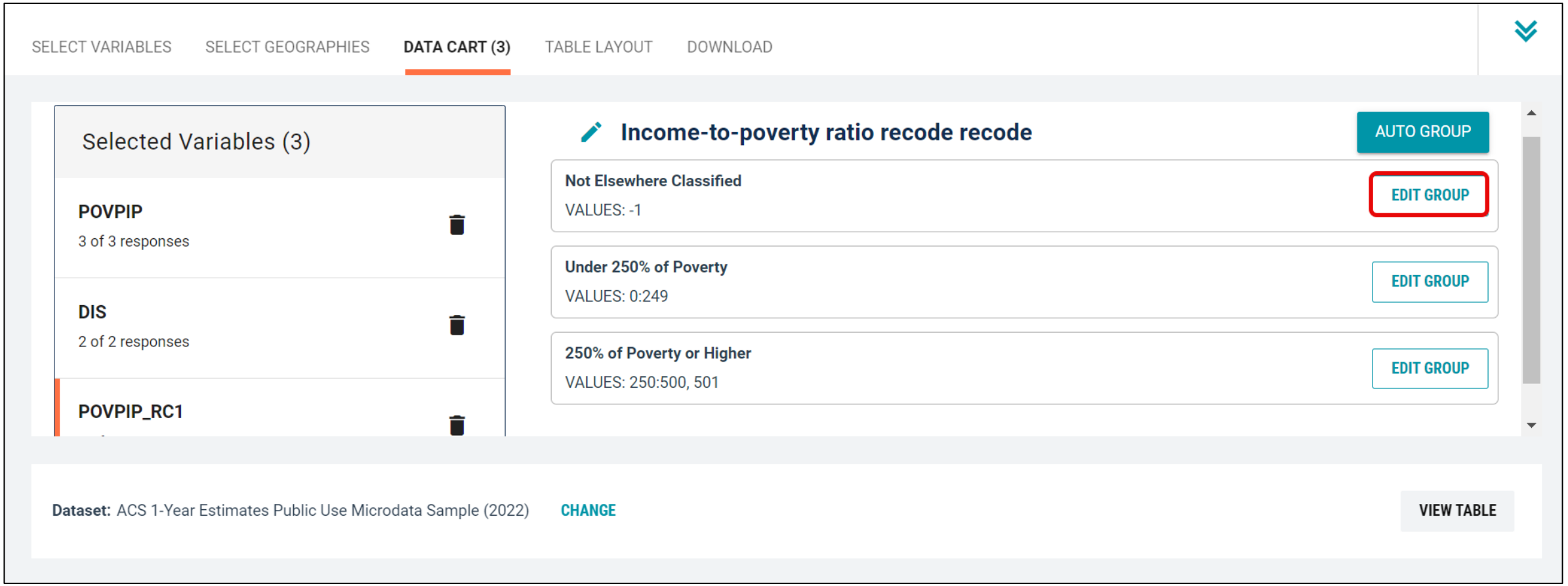
- Create recode for POVPIP variable:
  - Your first category, Under 250% of Poverty, appears just below “Not Elsewhere Classified”
  - Click **Edit Group** for “Not Elsewhere Classified” to verify and rename the category



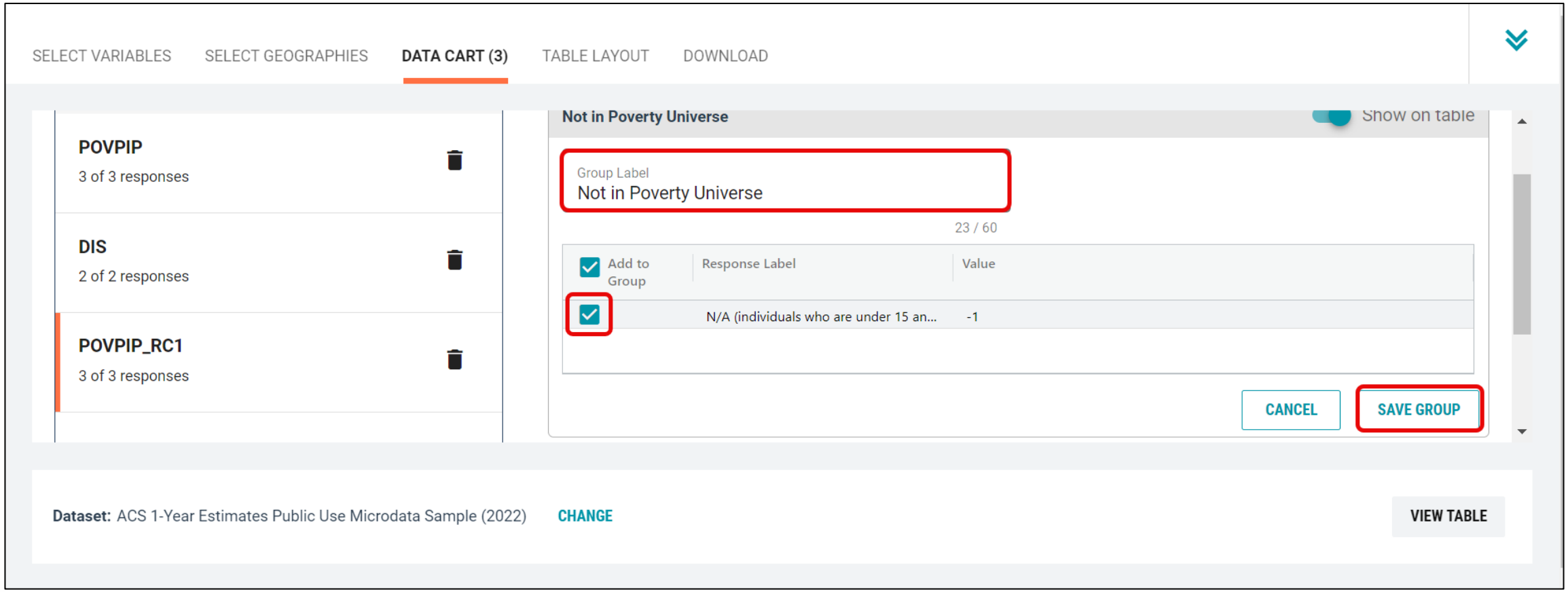
- Create recode for POVPIP variable:
  - Change Group Label to '**250% of Poverty or Higher**'.
  - Click on the checkboxes next to '**Between 250 and 500**' and '**501 percent or more**'.
  - Click on the '**Save Group**' button.



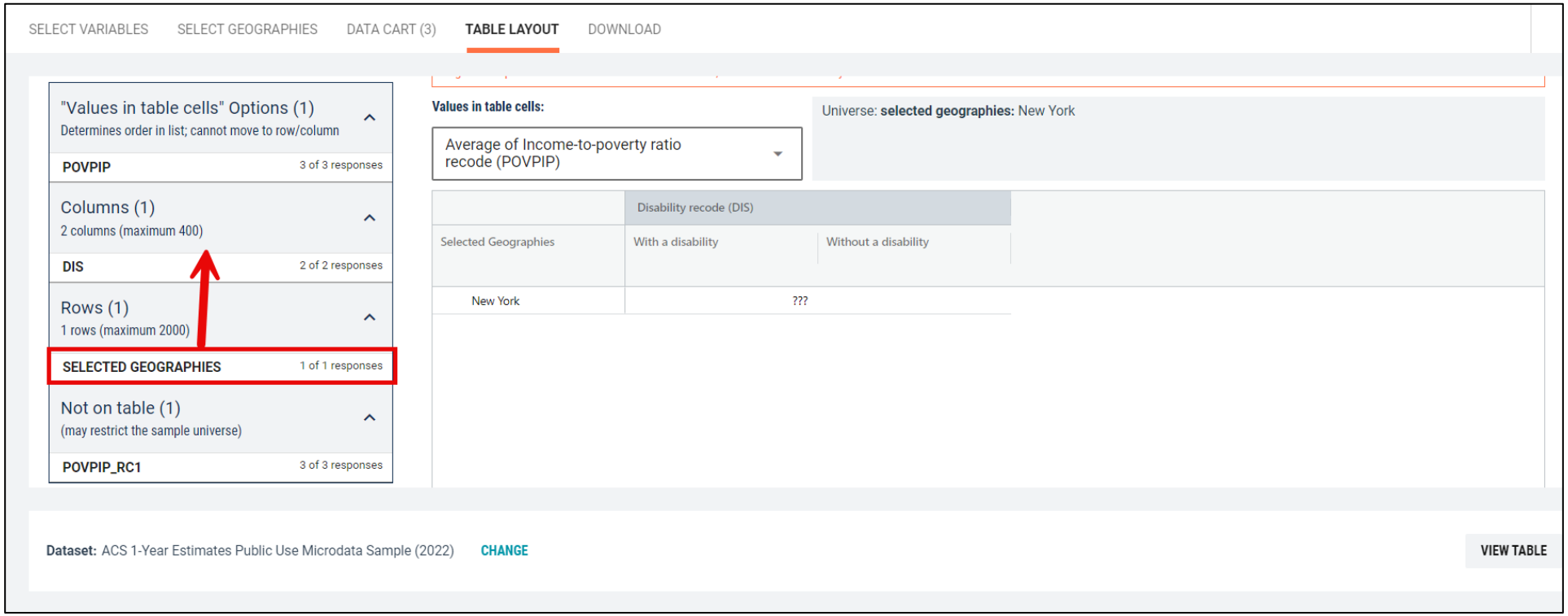
- Create recode for POVPIP variable:
  - Click **Edit Group** for “Not Elsewhere Classified” to verify and rename the final category



- Create recode for POVPIP variable:
  - Change Group Label to '**Not in Poverty Universe'.**
  - Click on the checkbox next to the '**N/A**' value.
  - Click on the '**Save Group**' button.

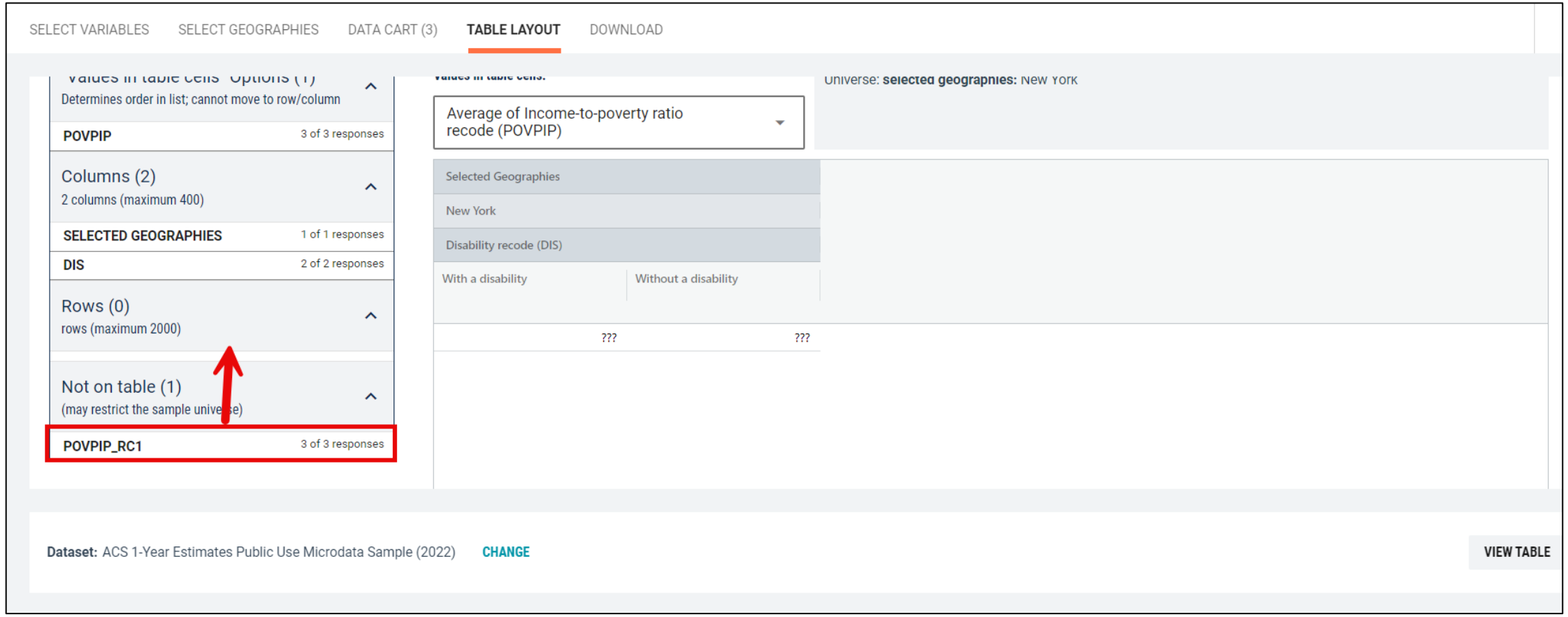


- View variable placement in the default table layout:
  - Move to the **Table Layout** tab
  - **Columns/Rows – Variables will be shown in the table.** By default, the table is providing the DIS variable in the Columns and the Selected Geographies in the Rows
  - Drag **Selected Geographies** up to Columns to display it above the DIS variable

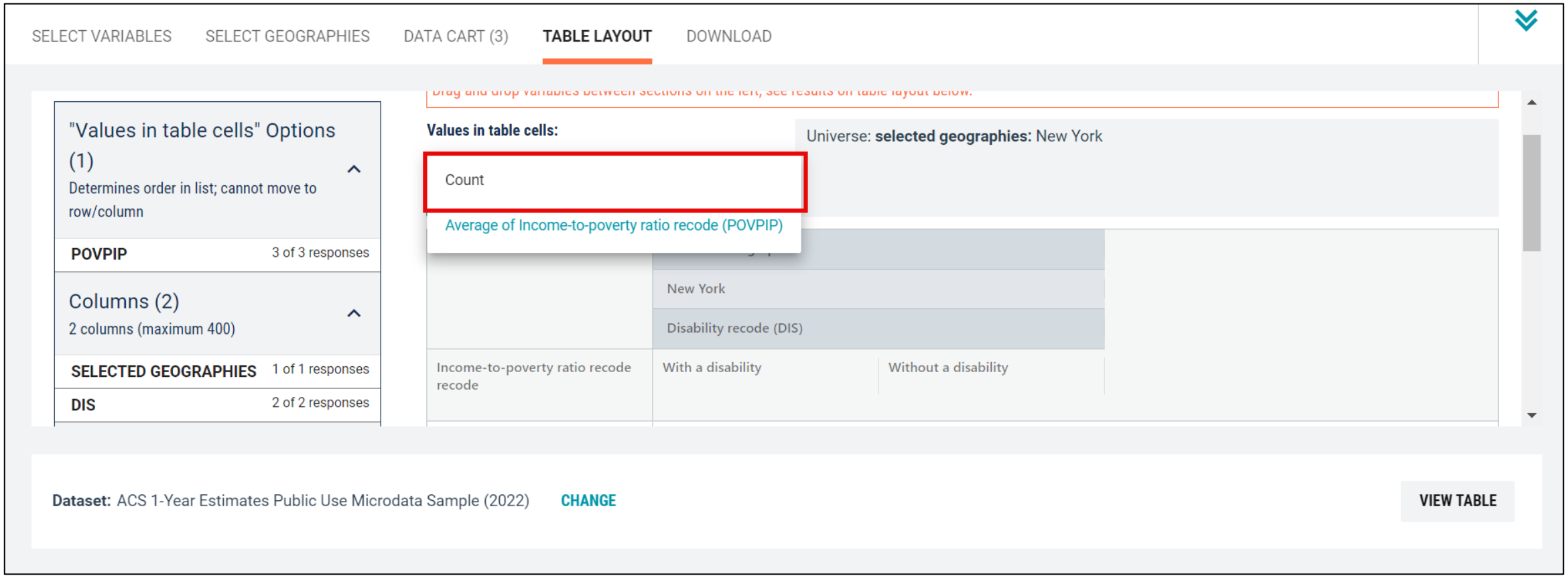


United States® Census Bureau

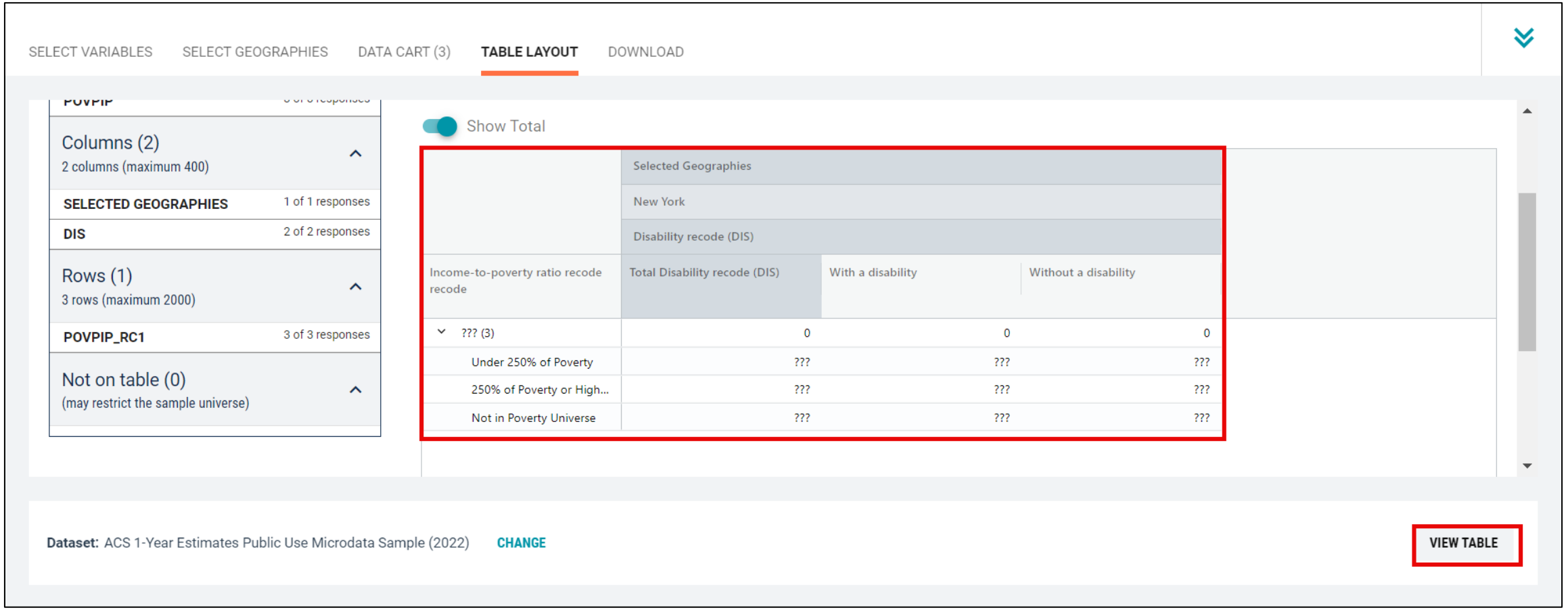
- Edit Table Layout:
  - **Move Poverty variable to rows:**
    - Click, hold and drag **POVPIP_RC1** up to the Rows heading.



- Choose type of values in table cells
  - Change the “Value in table cells” option from Average Income-to-poverty ratio recode to **Count.** This will give you data for the total number of people within the requested categories.

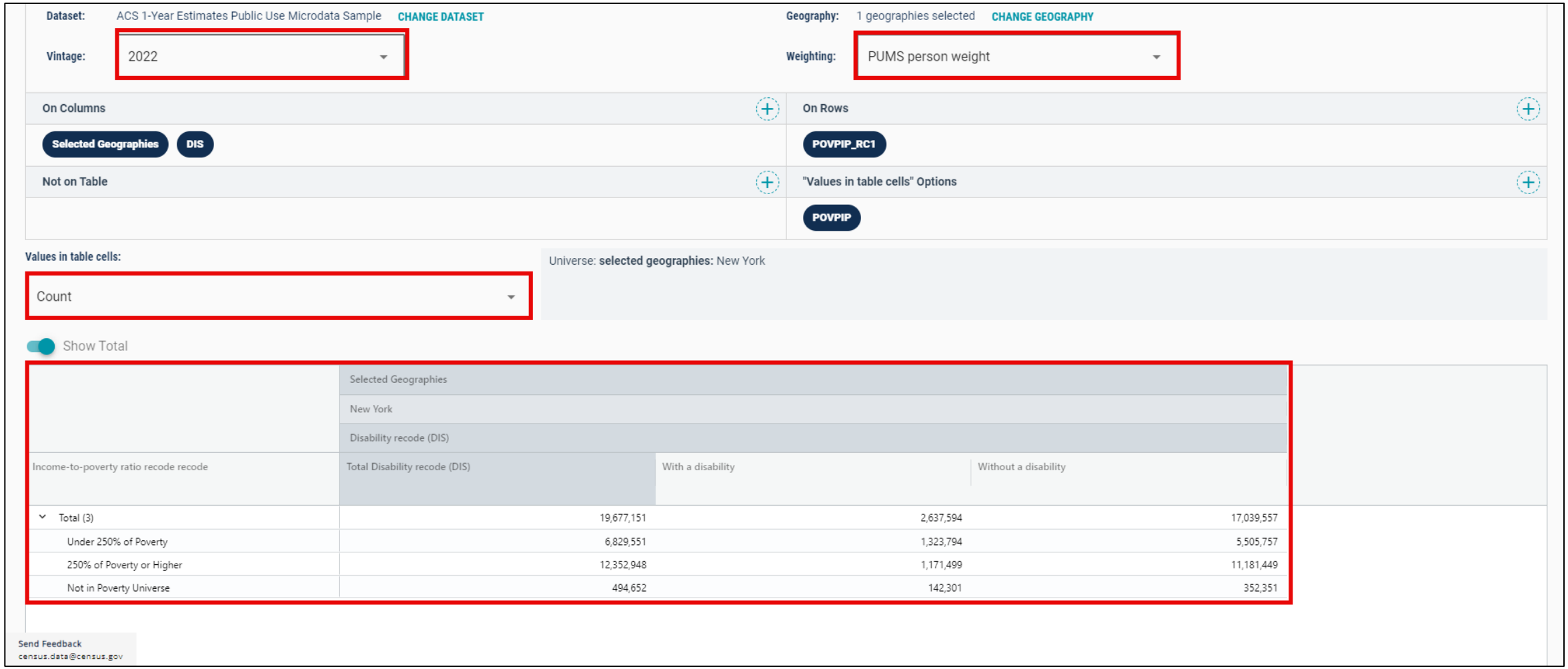


- Confirm Table Layout:
  - Confirm table layout and click **View Table** in the lower right



View Table

Note that the site automatically chooses a weight for you. You do have the option to change the weight if you want.



Dual Vintages for 2022 5-Year Public Use Microdata Areas (PUMAs)

United States® Census Bureau

What are Public Use Microdata Areas (PUMAS)?

- Non-overlapping
- **Statistical** geographic areas
- Partition each state or equivalent entity into geographic areas containing **no fewer than 100,000 people each**
- Cover the **entirety** of the United States, Puerto Rico and Guam
- Created for the tabulation and dissemination of
  - Decennial Census and American Community Survey (ACS) Public Use Microdata Sample (PUMS)
  - Puerto Rico Community Survey (PRCS)

United States® Census Bureau

What are Public Use Microdata Areas (PUMAS)?

- Delineation of new PUMAs occurs **after** the completion of the decennial census as part of a program involving the State Data Centers (SDCs)
- Created by using the latest decennial census population counts/census tracts

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

United States® Census Bureau

Dual Vintages

- 2022 5-year estimates are the first to use both 2010 and 2020 PUMA boundaries
  - 2022 5-year estimates comprised of 2018, 2019, 2020, 2021, and 2022
    - 2018 – uses 2010 PUMA boundaries
    - 2019 – uses 2010 PUMA boundaries
    - 2020 – uses 2010 PUMA boundaries
    - 2021 – uses 2010 PUMA boundaries
    - 2022 – uses 2020 PUMA boundaries
  - 2023 5-year estimates comprised of 2019, 2020, 2021, 2022, and 2023
    - 2019 – uses 2010 PUMA boundaries
    - 2020 – uses 2010 PUMA boundaries
    - 2021 – uses 2010 PUMA boundaries
    - 2022 – uses 2020 PUMA boundaries
    - 2023 – uses 2020 PUMA boundaries

Dual Vintages Continue Until Release of 2026 5-Year PUMS

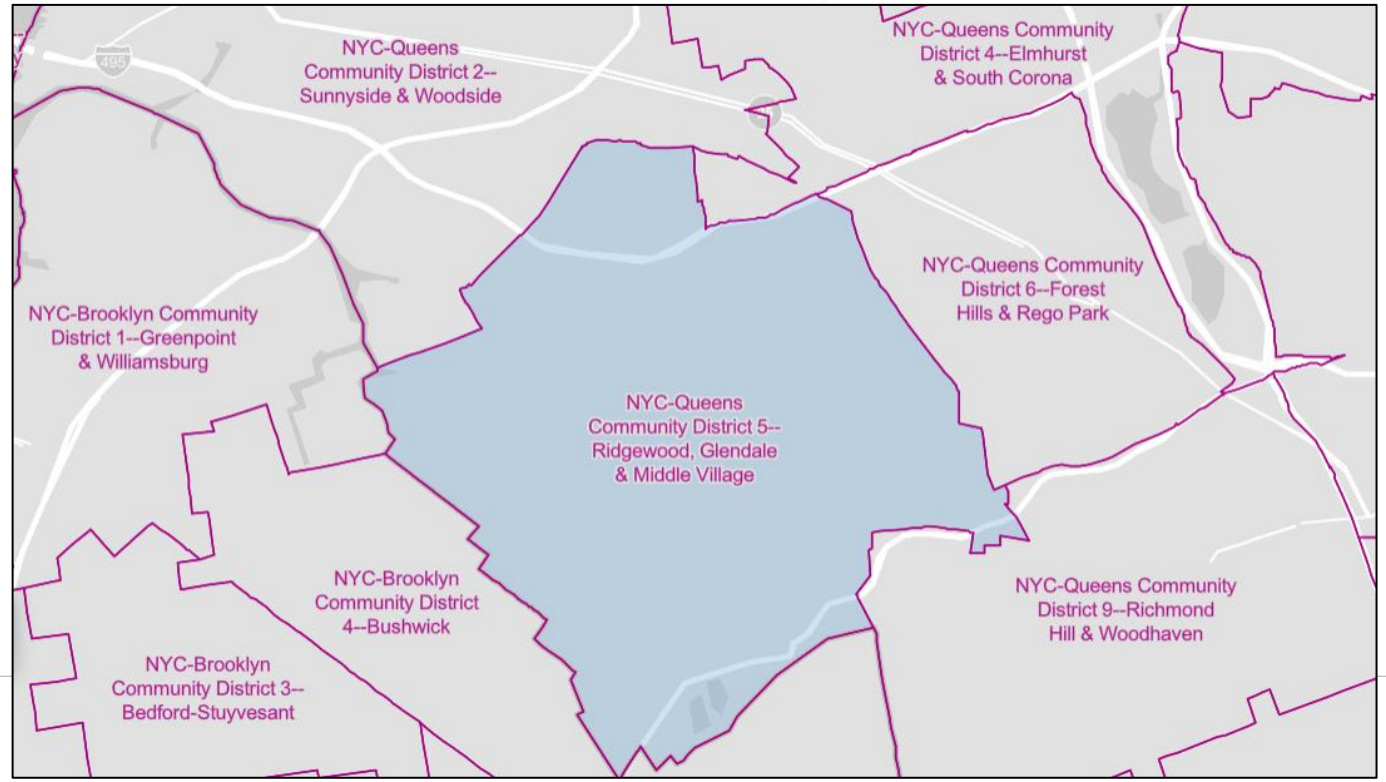
- **2022 5-year estimates** comprised of 2018, 2019, 2020, 2021, and 2022
  - 2018 – uses 2010 PUMA boundaries
  - 2019 – uses 2010 PUMA boundaries
  - 2020 – uses 2010 PUMA boundaries
  - 2021 – uses 2010 PUMA boundaries
  - 2022 – uses 2020 PUMA boundaries
- **2023 5-year estimates** comprised of 2019, 2020, 2021, 2022, and 2023
  - 2019 – uses 2010 PUMA boundaries
  - 2020 – uses 2010 PUMA boundaries
  - 2021 – uses 2010 PUMA boundaries
  - 2022 – uses 2020 PUMA boundaries
  - 2023 – uses 2020 PUMA boundaries
- **2024 5-year estimates** comprised of 2020, 2021, 2022, 2023, and 2024
  - 2020 – uses 2010 PUMA boundaries
  - 2021 – uses 2010 PUMA boundaries
  - 2022 – uses 2020 PUMA boundaries
  - 2023 – uses 2020 PUMA boundaries
  - 2024 – uses 2020 PUMA boundaries
- **2025 5-year estimates** comprised of 2021, 2022, 2023, 2024, and 2025
  - 2021 – uses 2010 PUMA boundaries
  - 2022 – uses 2020 PUMA boundaries
  - 2023 – uses 2020 PUMA boundaries
  - 2024 – uses 2020 PUMA boundaries
  - 2025 – uses 2020 PUMA boundaries

No more dual vintages until 2030 Census

- **2026 5-year estimates** comprised of 2022, 2023, 2024, 2025, and 2026
  - 2022 – uses 2020 PUMA boundaries
  - 2023 – uses 2020 PUMA boundaries
  - 2024 – uses 2020 PUMA boundaries
  - 2025 – uses 2020 PUMA boundaries
  - 2024 – uses 2020 PUMA boundaries

2010 PUMA 04110 in New York
NYC-Queens Community District 5—Ridgewood, Glendale, and Middle Village

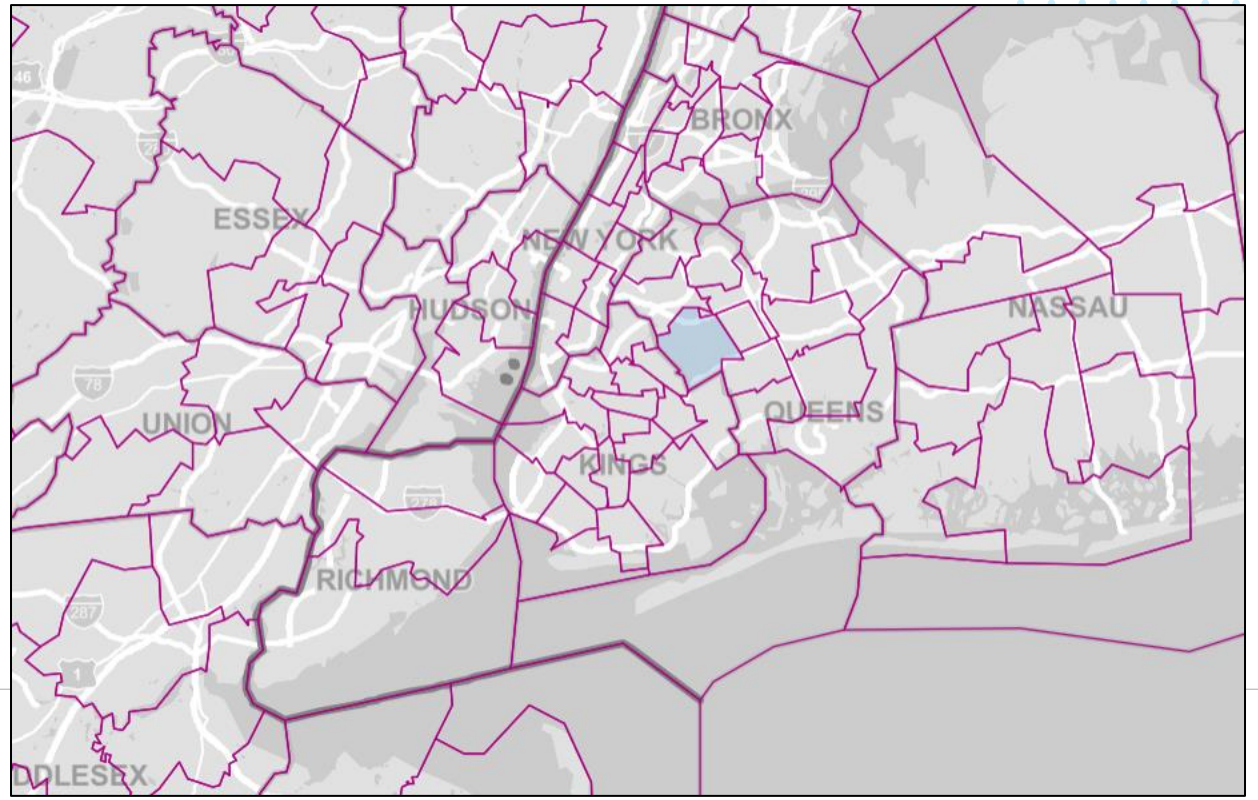
2010 PUMA Names File		
36	04108	NYC-Queens Community District 6--Forest Hills & Rego Park
36	04109	NYC-Queens Community District 2--Sunnyside & Woodside
36	04110	NYC-Queens Community District 5--Ridgewood, Glendale & Middle Village
36	04111	NYC-Queens Community District 9--Richmond Hill & Woodhaven
36	04112	NYC-Queens Community District 12--Jamaica, Hollis & St. Albans



United States® Census Bureau

2010 PUMA 04110 in New York
NYC-Queens Community District 5—Ridgewood, Glendale, and Middle Village

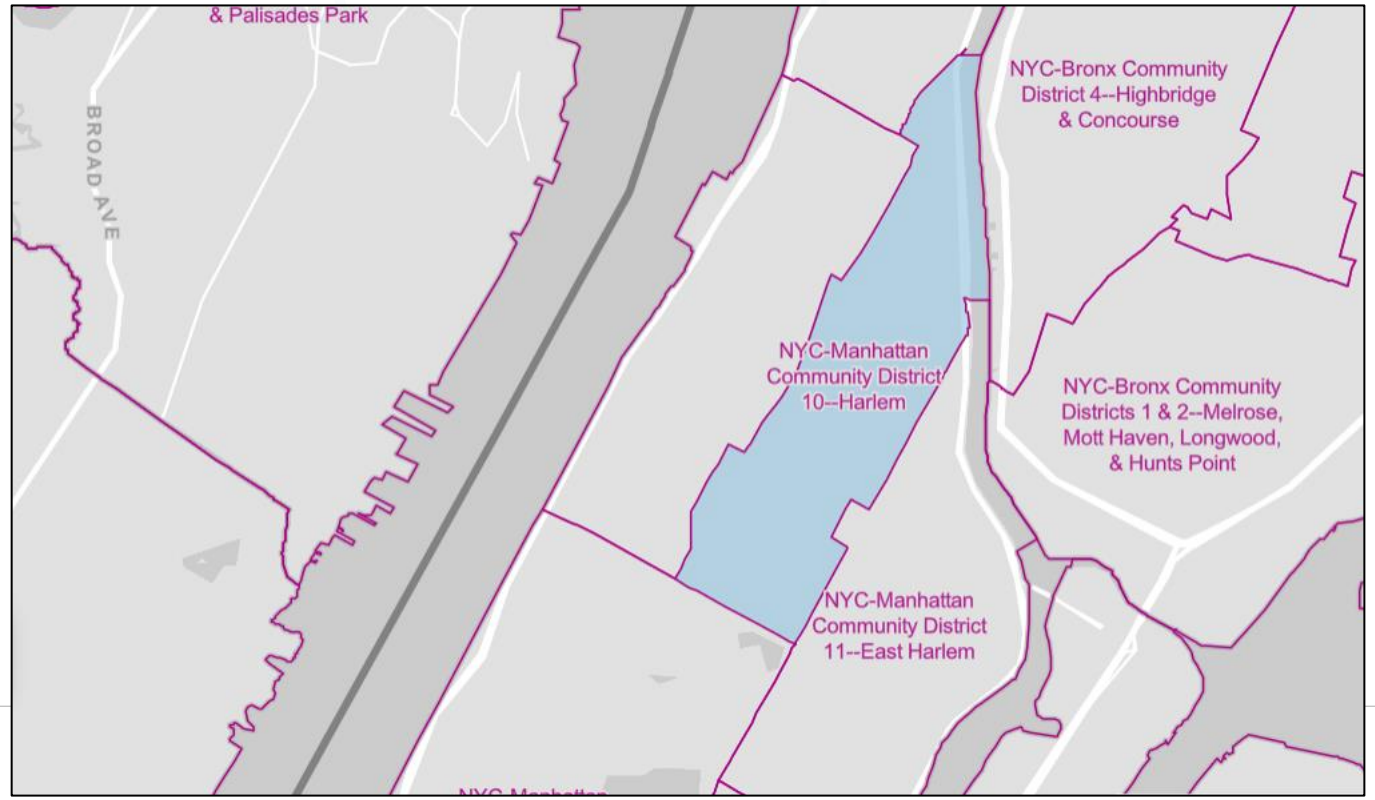
2010 PUMA Names File		
36	04108	NYC-Queens Community District 6--Forest Hills & Rego Park
36	04109	NYC-Queens Community District 2--Sunnyside & Woodside
36	04110	NYC-Queens Community District 5--Ridgewood, Glendale & Middle Village
36	04111	NYC-Queens Community District 9--Richmond Hill & Woodhaven
36	04112	NYC-Queens Community District 12--Jamaica, Hollis & St. Albans



2020 PUMA 04110 in New York
NYC-Manhattan Community District 10—Harlem

2020 PUMA Names File

36	04108	NYC-Manhattan Community District 8--Upper East Side & Roosevelt Island
36	04109	NYC-Manhattan Community District 9--Morningside Heights & Hamilton Heights
36	04110	NYC-Manhattan Community District 10--Harlem
36	04111	NYC-Manhattan Community District 11--East Harlem
36	04112	NYC-Manhattan Community District 12--Washington Heights & Inwood
36	04121	NYC-Manhattan Community Districts 1 & 2--Financial District & Greenwich Vil

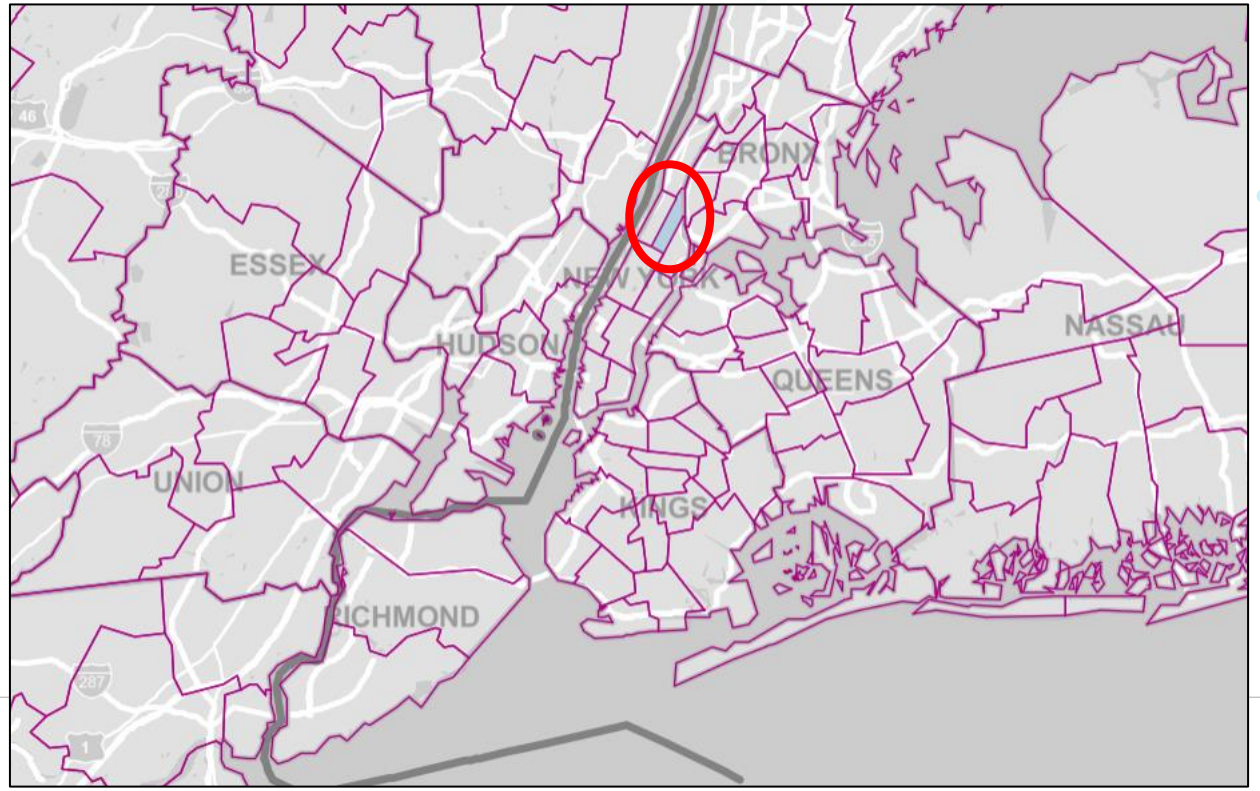


United States® Census Bureau

2020 PUMA 04110 in New York
NYC-Manhattan Community District 10—Harlem

2020 PUMA Names File

36	04108	NYC-Manhattan Community District 8--Upper East Side & Roosevelt Island
36	04109	NYC-Manhattan Community District 9--Morningside Heights & Hamilton Heights
36	04110	NYC-Manhattan Community District 10--Harlem
36	04111	NYC-Manhattan Community District 11--East Harlem
36	04112	NYC-Manhattan Community District 12--Washington Heights & Inwood
36	04121	NYC-Manhattan Community Districts 1 & 2--Financial District & Greenwich Vil

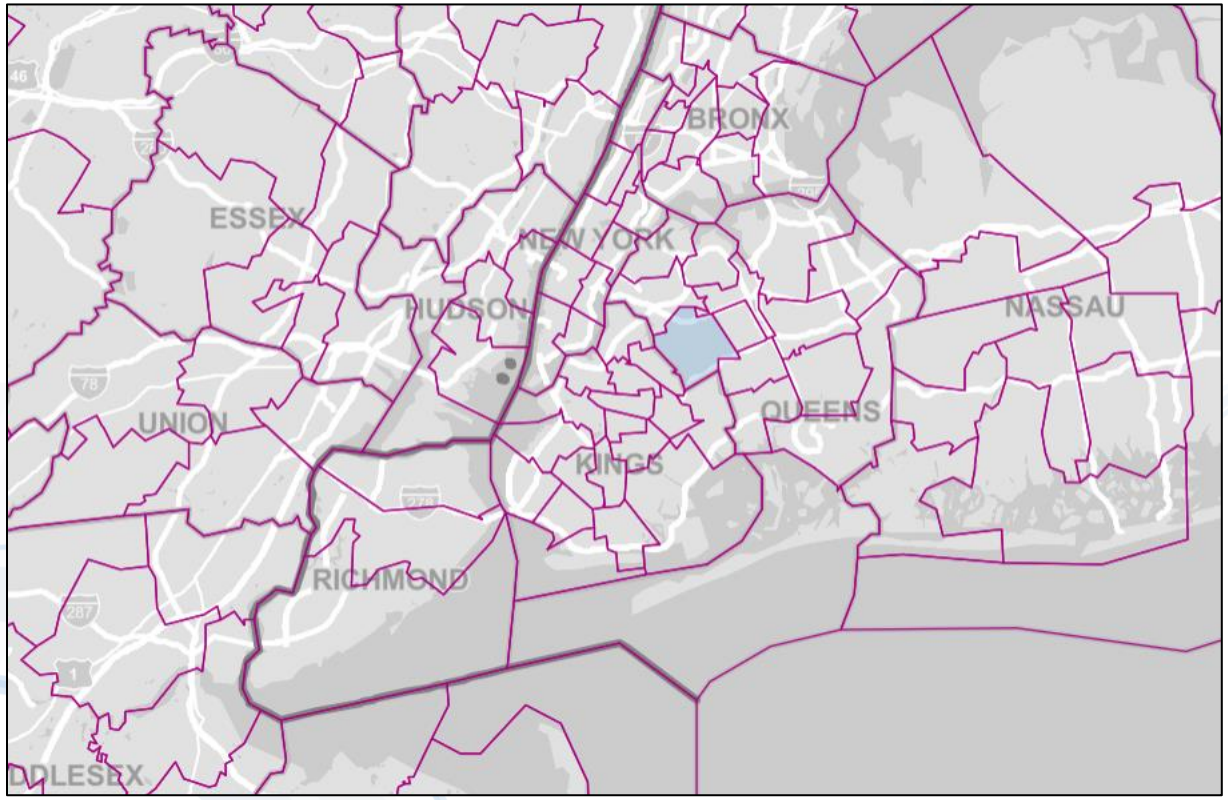


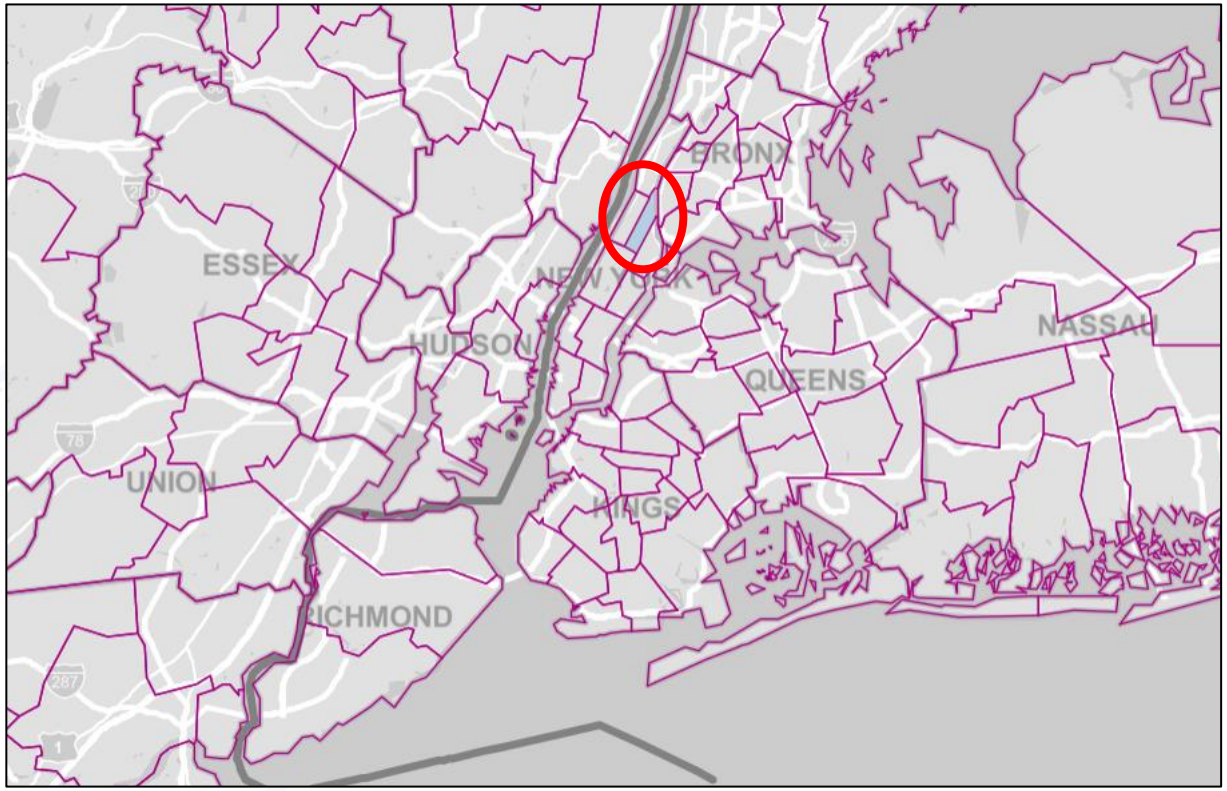
United States® Census Bureau

2010 PUMA 04110 in New York NYC-Queens Community District 5—Ridgewood, Glendale, and Middle Village

vs.

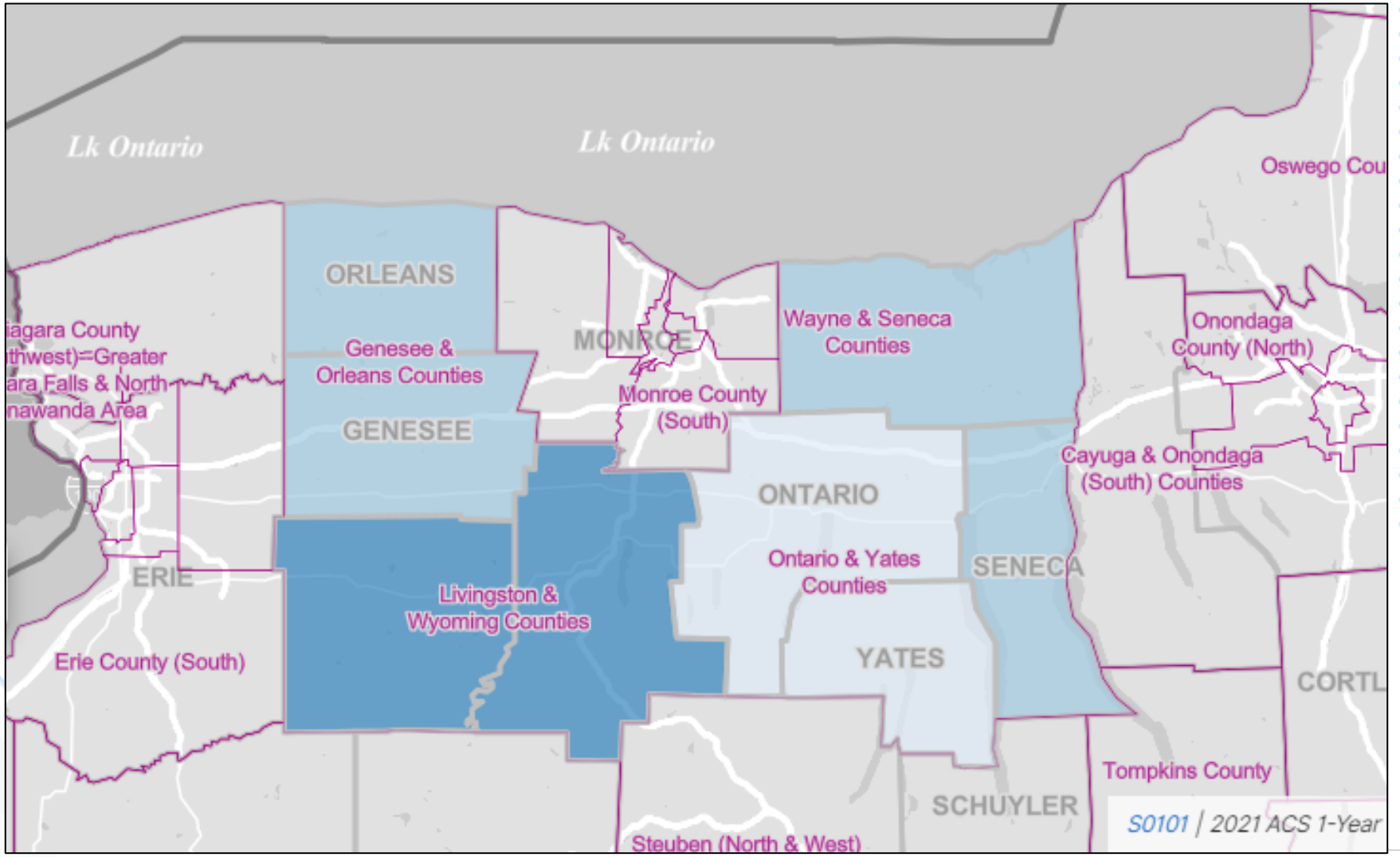
2020 PUMA 04110 in New York NYC-Manhattan Community District 10—Harlem



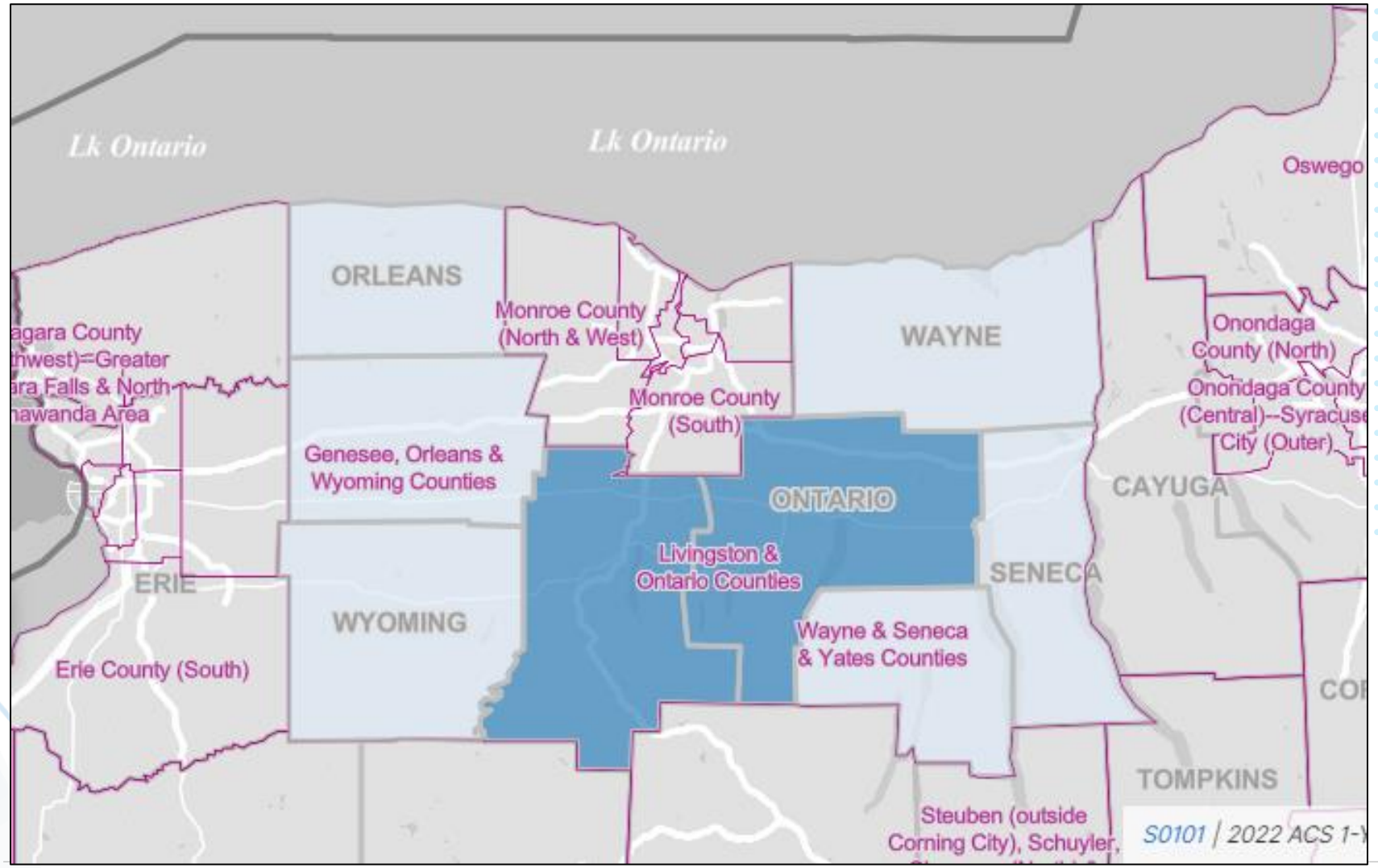


United States® Census Bureau

2021 Map of PUMAs near Lake Ontario (Using 2010 PUMA Boundaries)



2022 Map of PUMAs near Lake Ontario (Using 2020 PUMA Boundaries)



United States® Census Bureau

2010 IDs and Names vs. 2020 IDs and Names for PUMAs near Lake Ontario

2010 ID	2010 Name	2020 ID	2020 Name
01000	Genessee & Orleans Counties	01000	Genessee, Orleans, & Wyoming Counties
01300	Livingston & Wyoming Counties	01300	Livingston & Ontario Counties
01400	Ontario & Yates Counties		
00800	Wayne & Seneca Counties	00800	Wayne & Seneca & Yates Counties

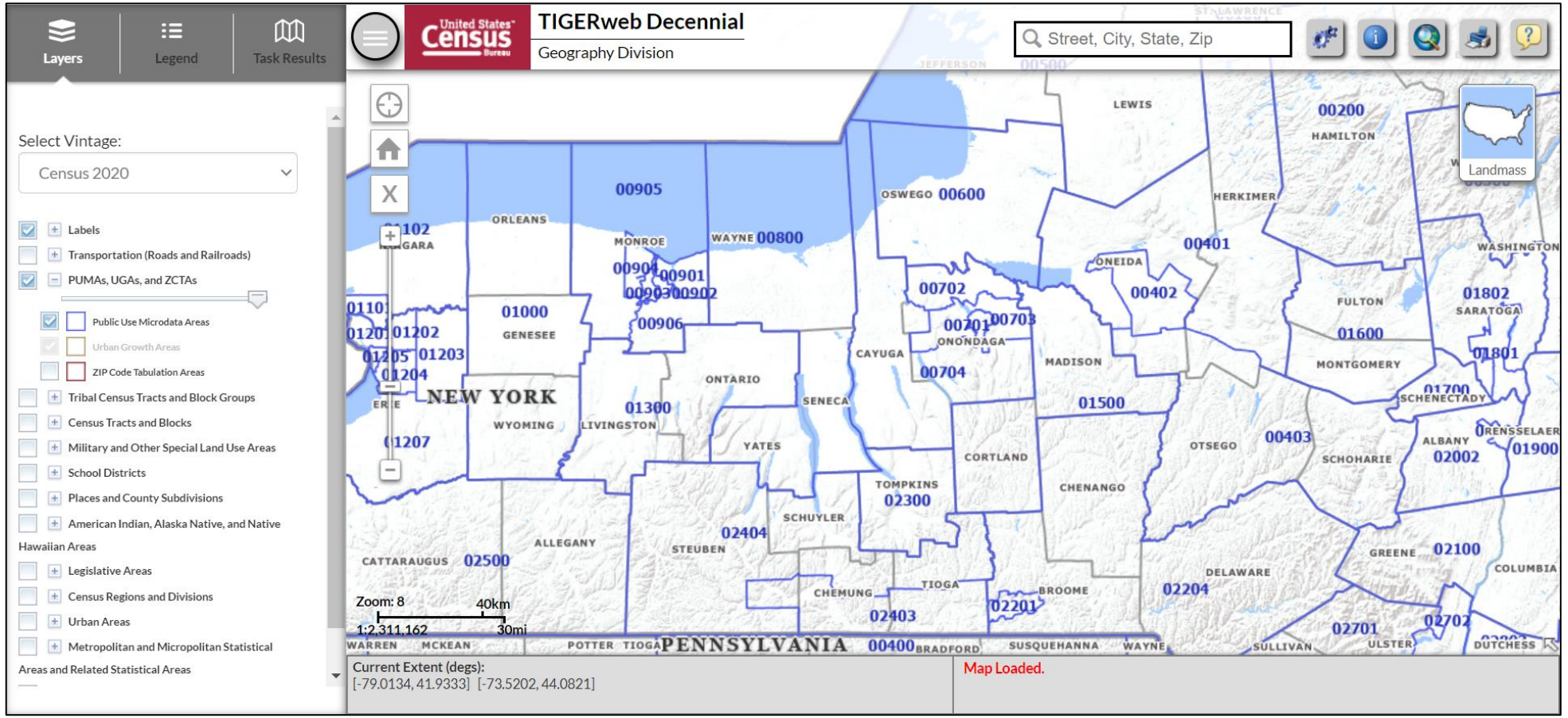
2010 PUMA Names vs. 2020 PUMA Names

TIGERweb as a Resource to find different GEOIDs

- **Along with data.census.gov, [TIGERweb Decennial] can act as a resource to find information about PUMAs that have GEOIDs that may have changed.**

- **The Census 2020 Vintage will have the 2020 PUMA boundaries.**

- **Switch to the Census 2010 Vintage to use the 2010 PUMA boundaries.**



For questions on TiGERweb, please reach out to [geo.tigerweb@census.gov]

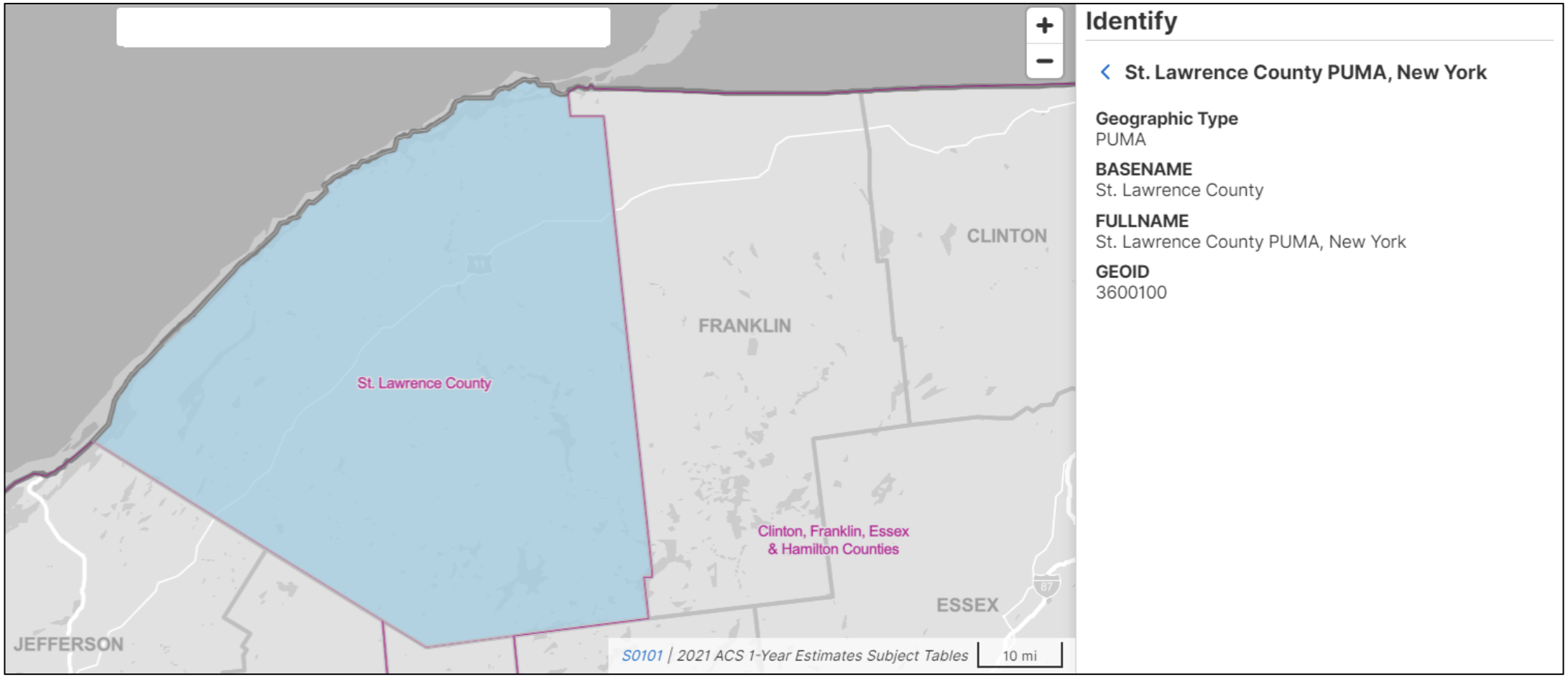
Demo

Example 2:

PUMA with no change in name or GEOID

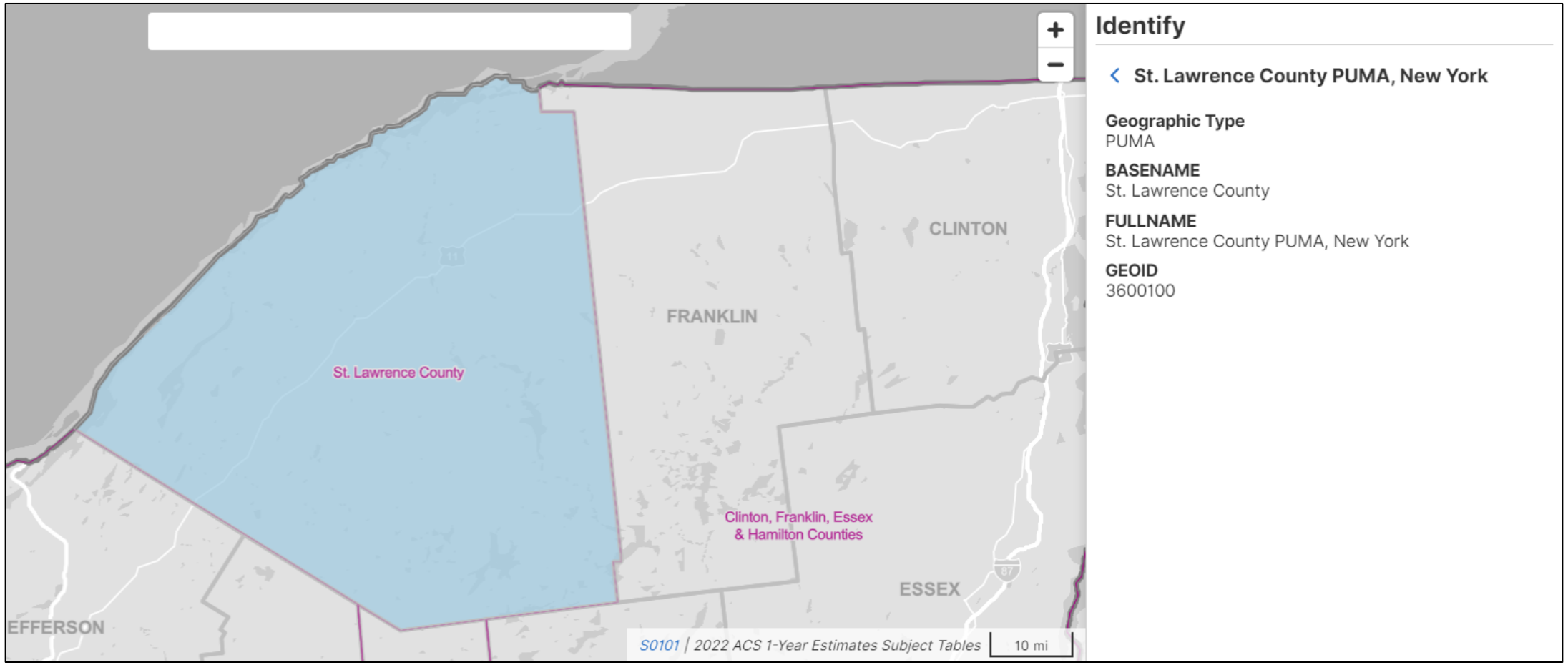
United States® Census Bureau

2021 Map for St. Lawrence County PUMA, New York (Using 2010 PUMA Boundaries)



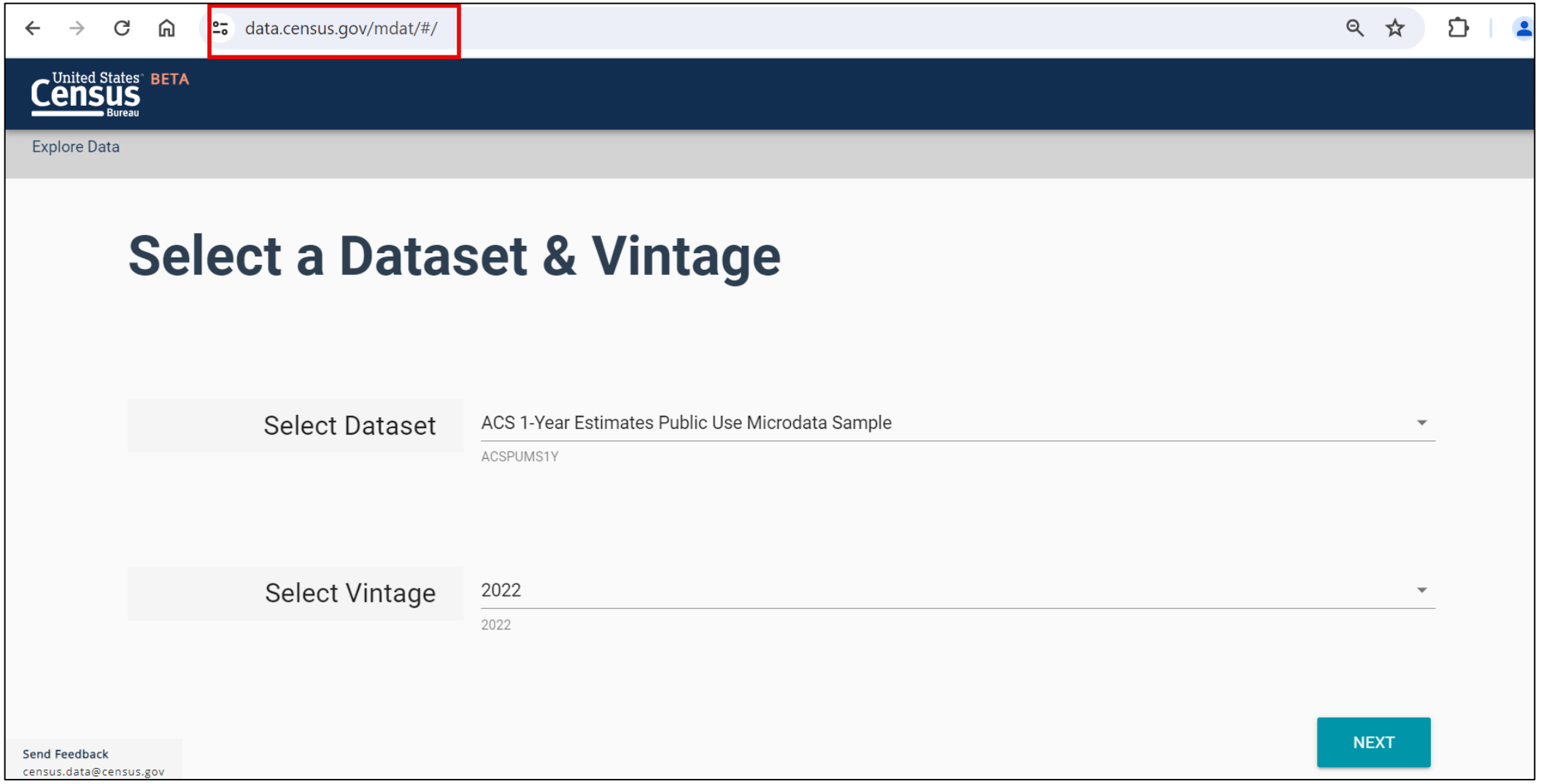
United States® Census Bureau

2022 Map for St. Lawrence County PUMA, New York (Using 2020 PUMA Boundaries)



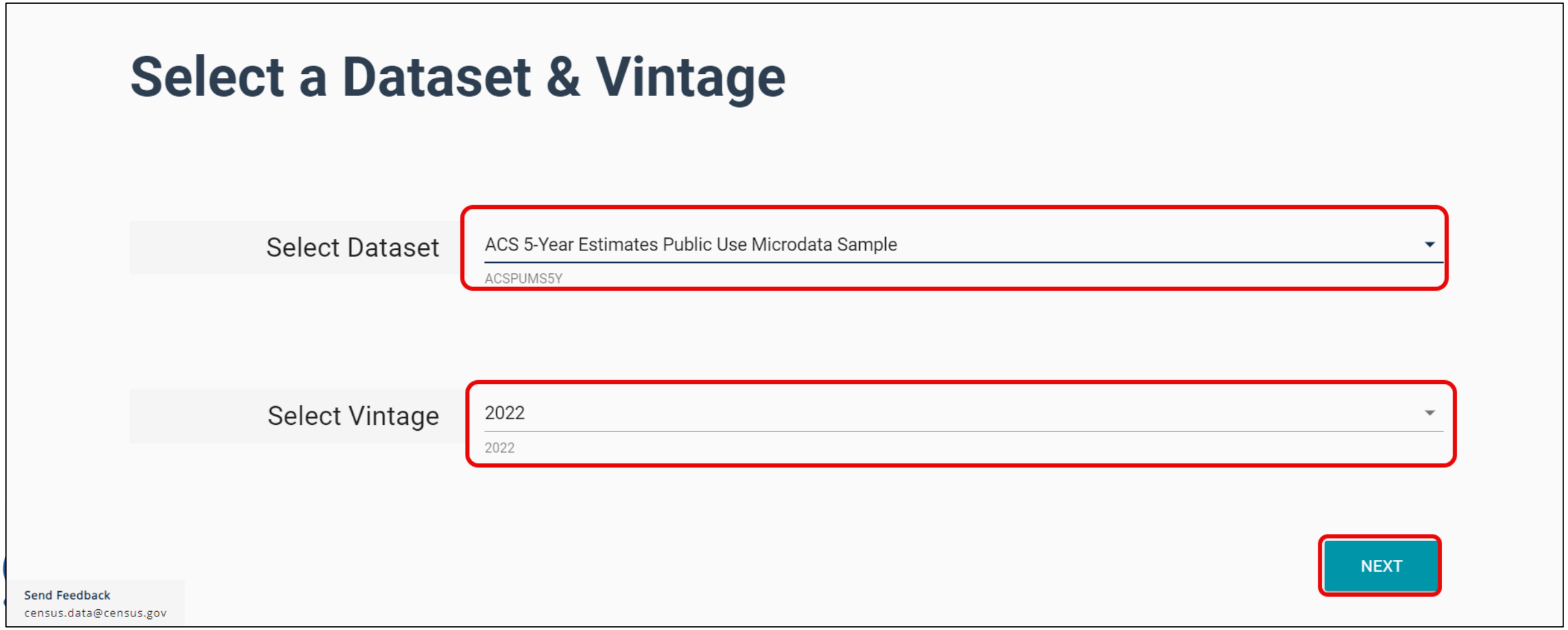
United States® Census Bureau

- Visit Microdata Access at data.census.gov/mdat

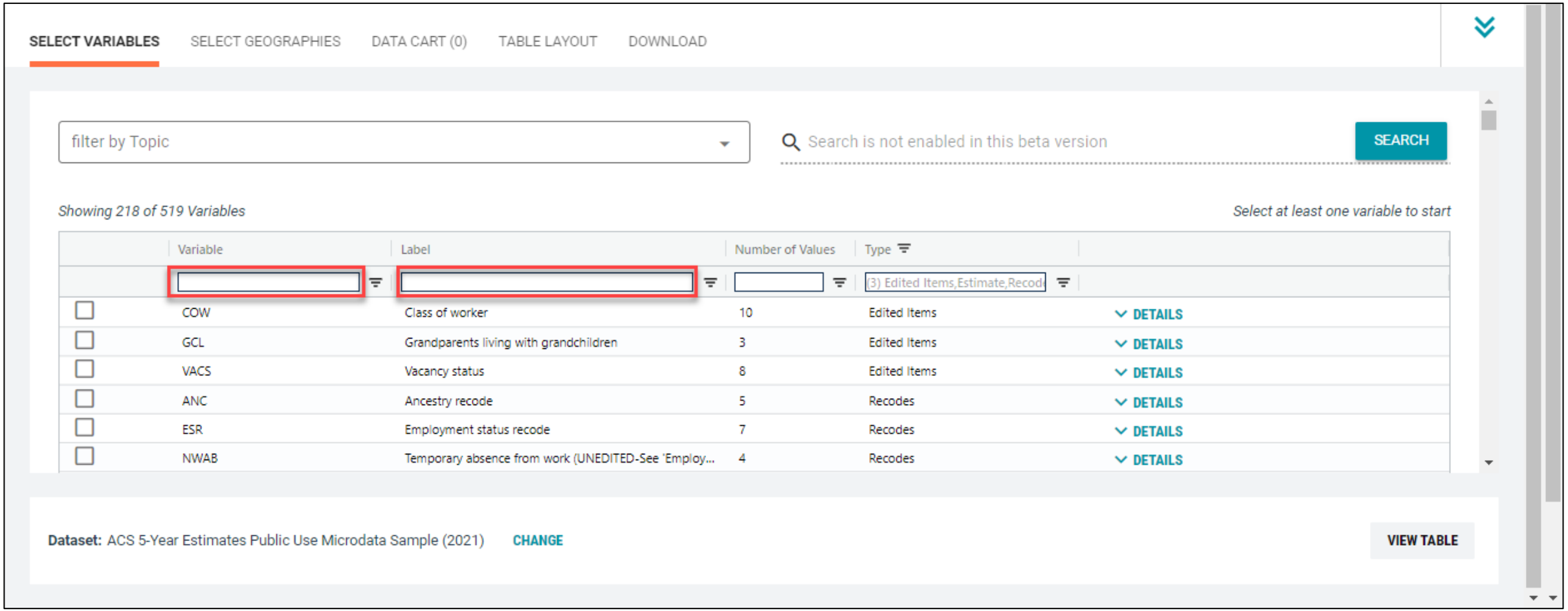


data.census.gov/mdat

- Choose Dataset and Vintage:
  - Dataset – **ACS 5-Year Estimates – Public Use Microdata Sample**
  - Vintage – **2022**
  - Click **Next** in the lower right

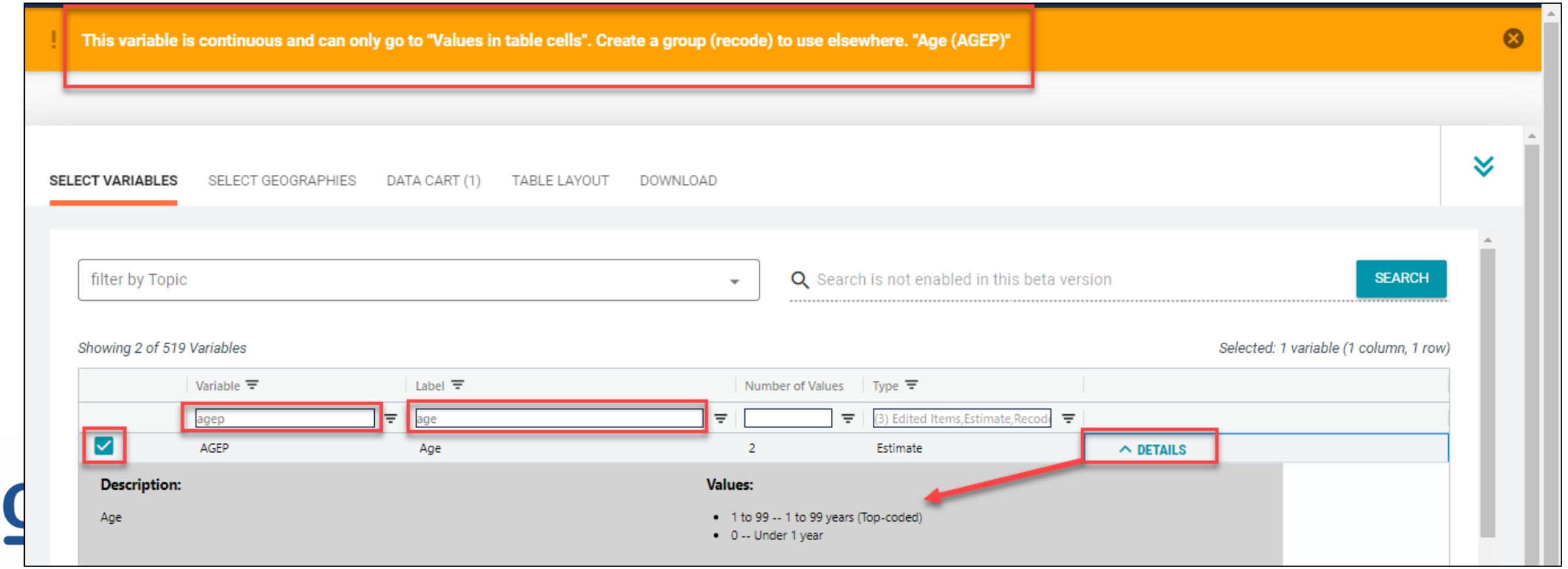


- **Search for Variables**: Use the search box below “Variable” or “Label” to find your variables of interest

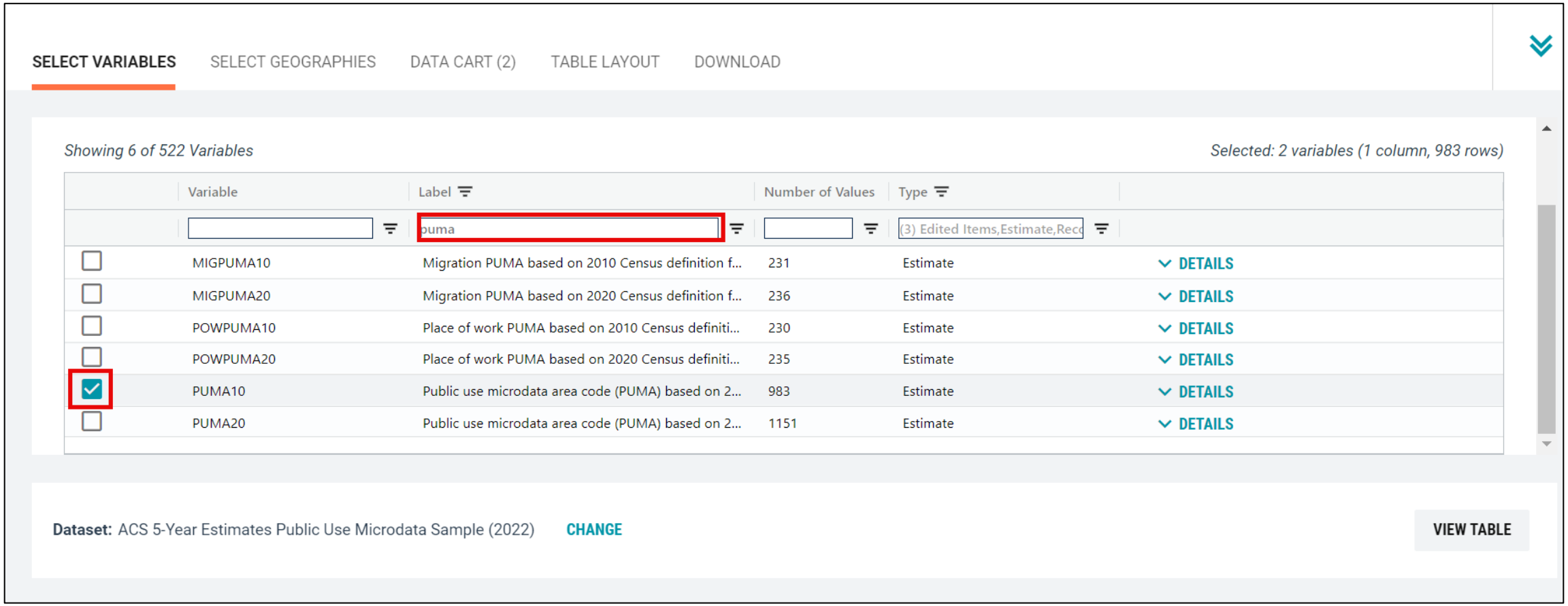


United States® Census Bureau

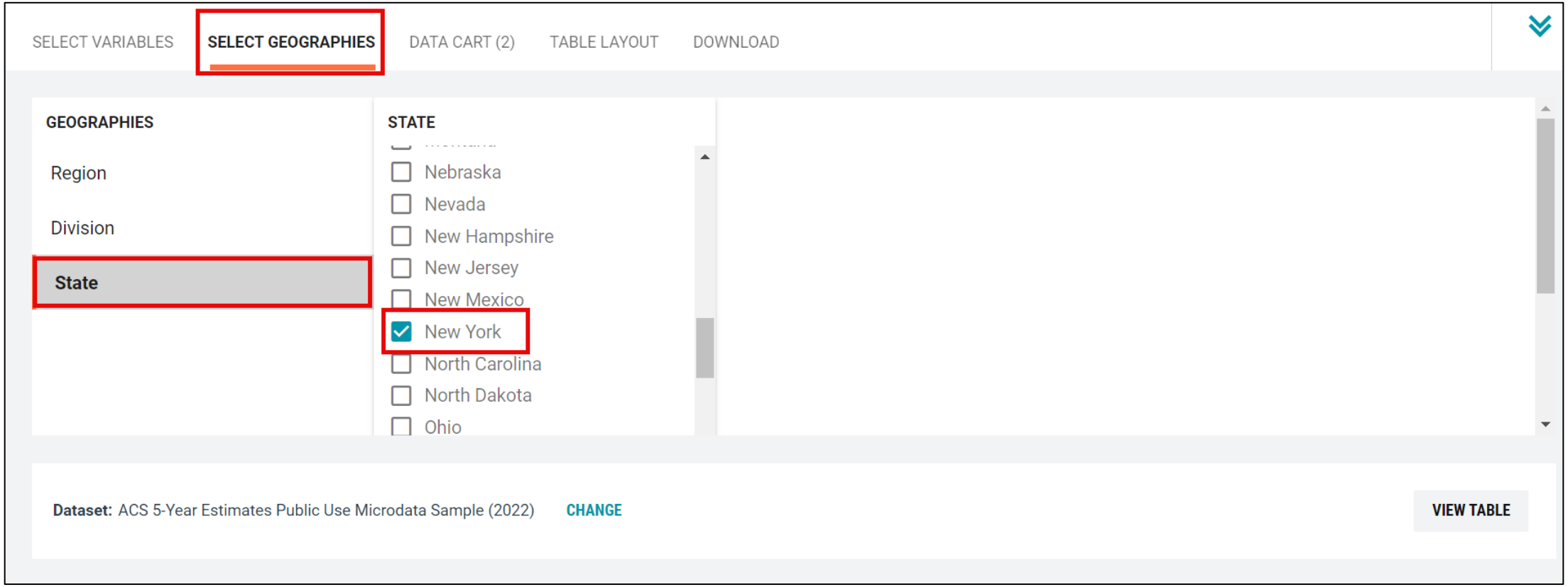
- Select variable for Age:
  - Type “AGEP” in the Variable search box or type “Age” in the label search box
  - Check the box to the left of AGEP to add the variable to your data cart
  - Notice the message at the top of the screen saying you will need to create your own categories (or recodes) for this variable if you want it shown in the table. (You will do this action in the Data Cart)



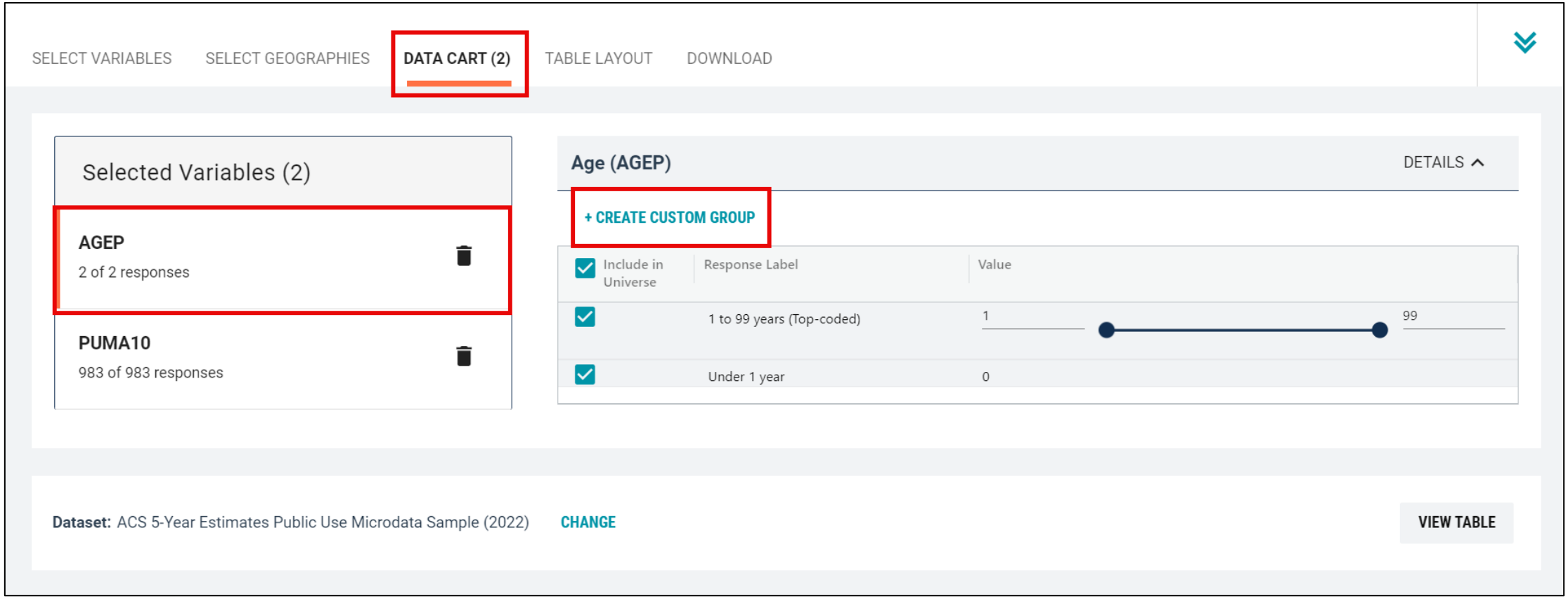
- Select variable for 2010 PUMAs:
  - Type ‘PUMA’ in the label search box
  - Check the box to the left of ‘PUMA10’ to add the variable to data cart



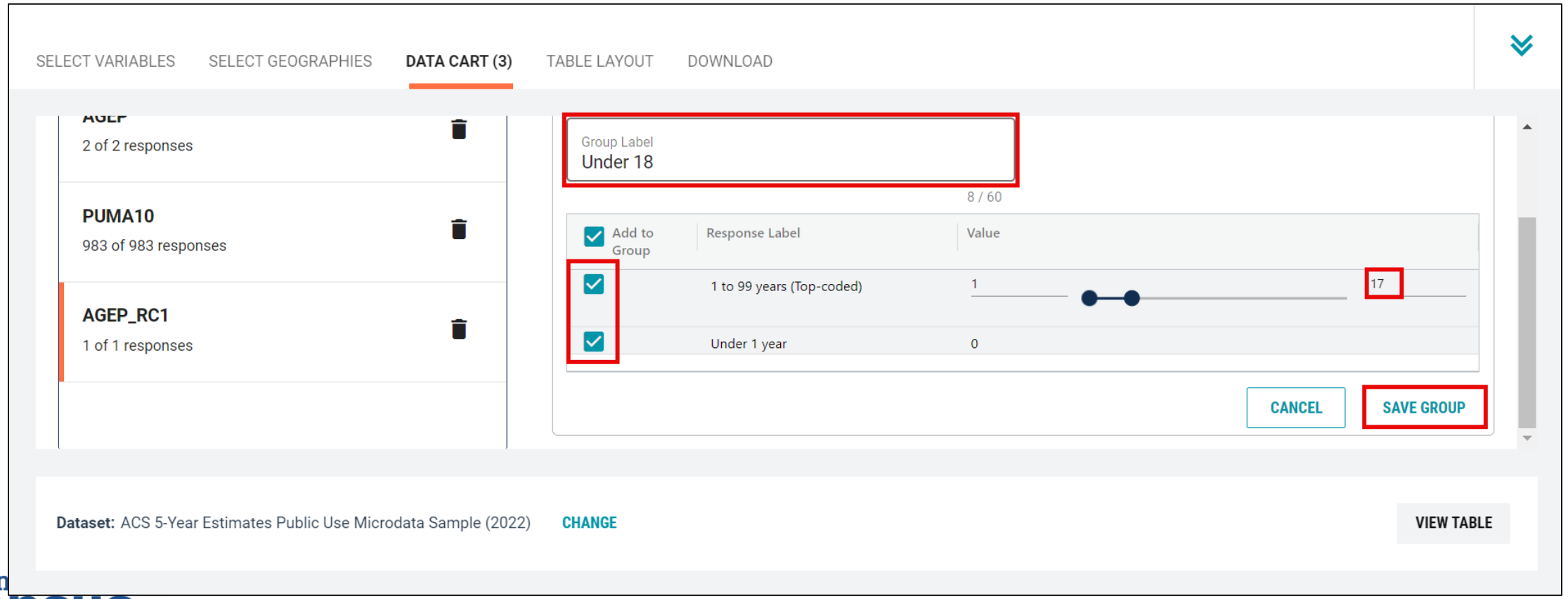
- **Select New York state geography.**
  - Click on the Select Geographies tab
  - Check the box to the left ‘New York’ to only pull up data for PUMAs from New York state



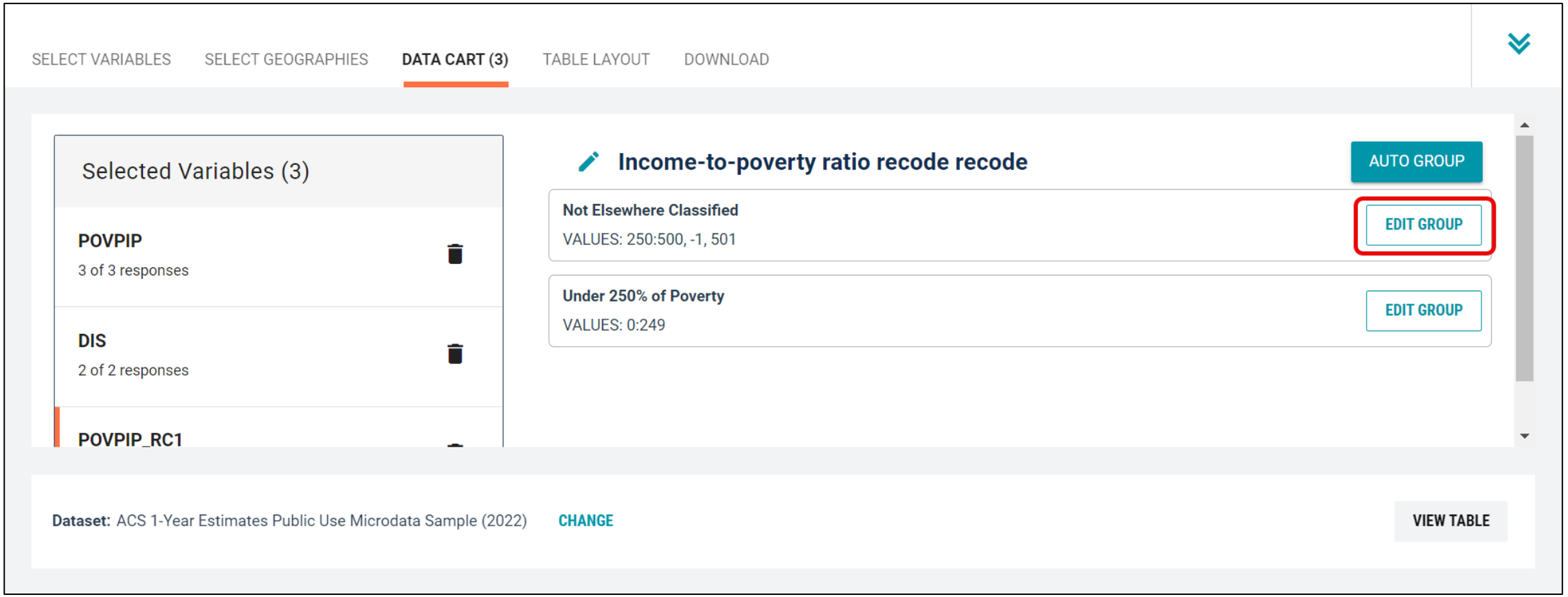
- Categorize (recode) your variable:
  - Move to the **Data Cart** tab
  - Click the **AGEP** variable on the left
  - Click **Create Custom Group** to begin specifying your age groups (e.g. under 18 years)



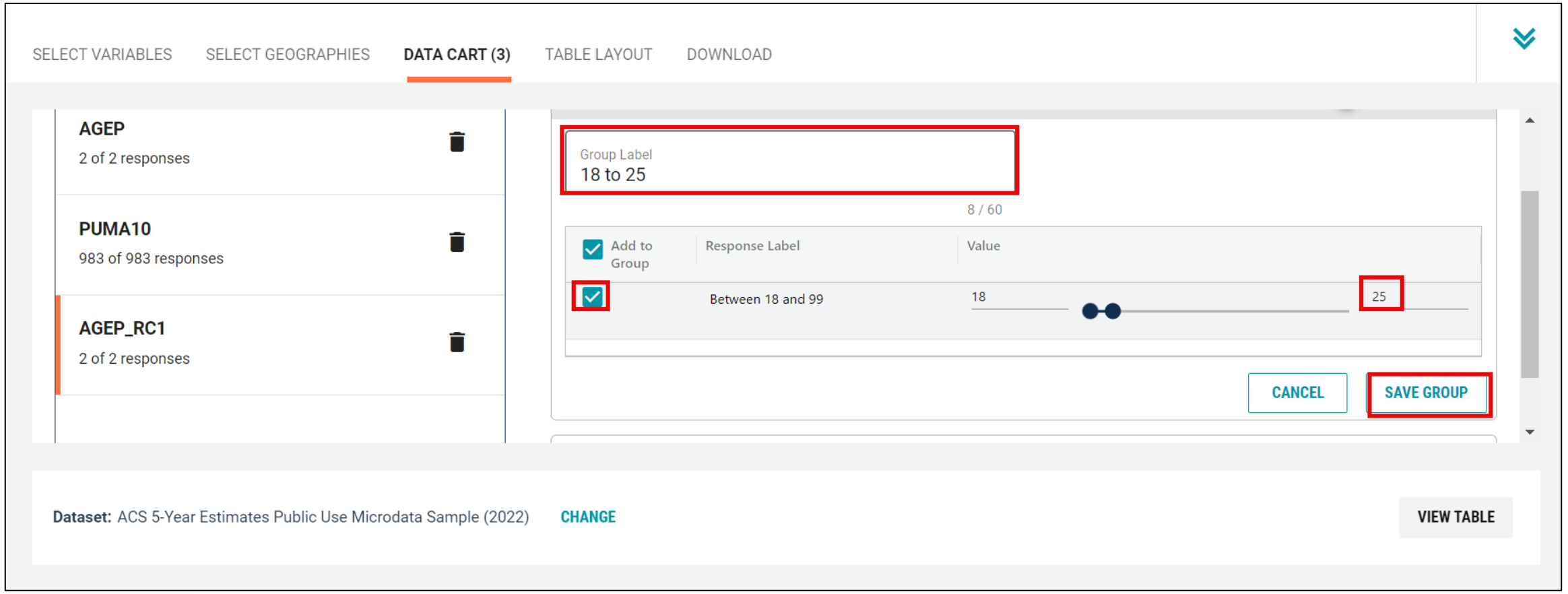
- Create recode for AGEP variable:
  - Change Group Label to **‘Under 18’**
  - Click on the checkbox next to **1 to 99** and edit the end range to **17** and click the checkbox next to **Under 1 Year**
  - Click on the ‘**Save Group**’ button.



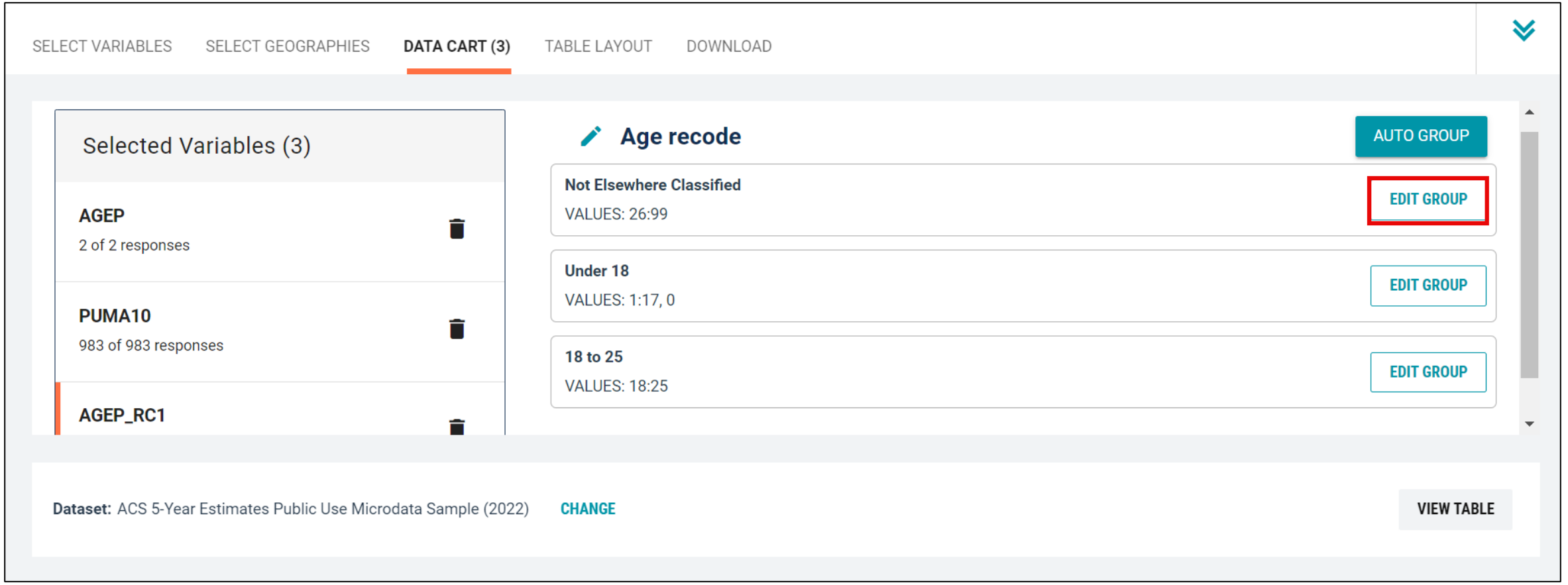
- Create recode for AGEP variable:
  - Your first category, Under 18, appears just below “Not Elsewhere Classified”
  - Click **Edit Group** for “Not Elsewhere Classified” to verify and rename the category



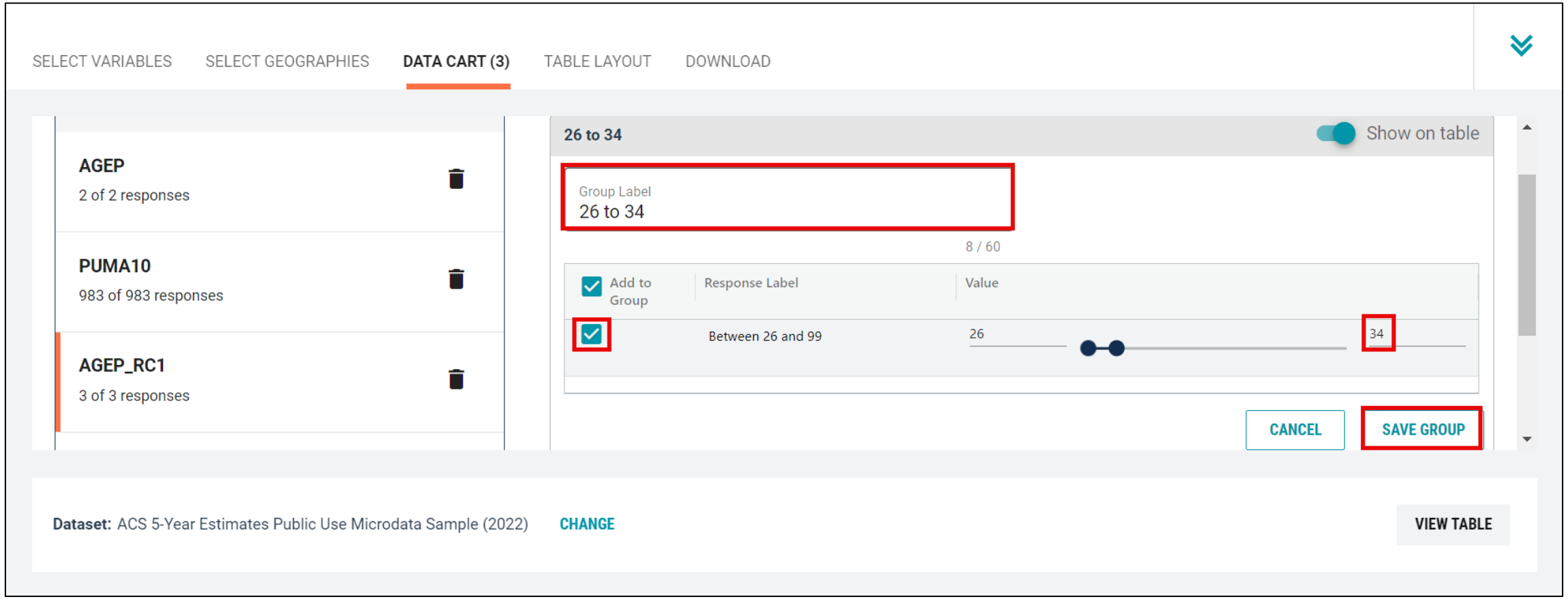
- Create recode for AGEP variable:
  - Change Group Label to **’18 to 25’**
  - Click on the checkbox next to **Between 18 to 99** and edit the end range to **25**
  - Click on the ‘**Save Group**’ button.



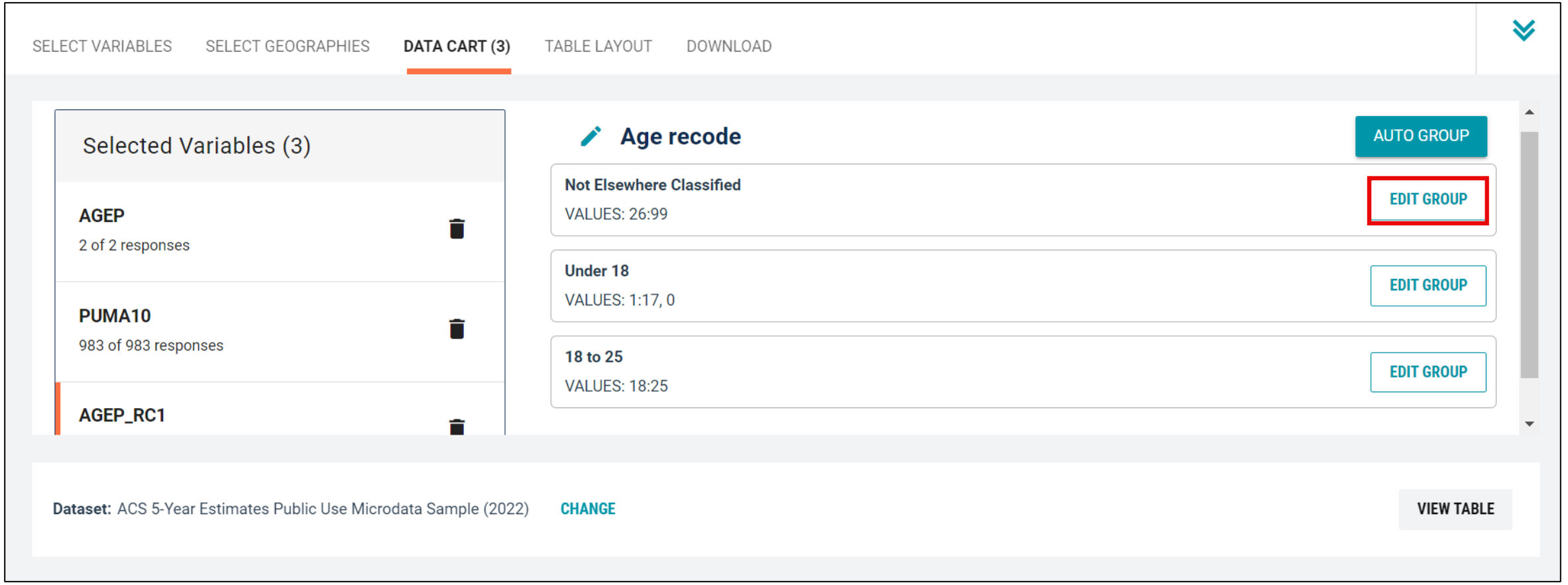
- Create recode for AGEP variable:
  - Click **Edit Group** for “Not Elsewhere Classified”



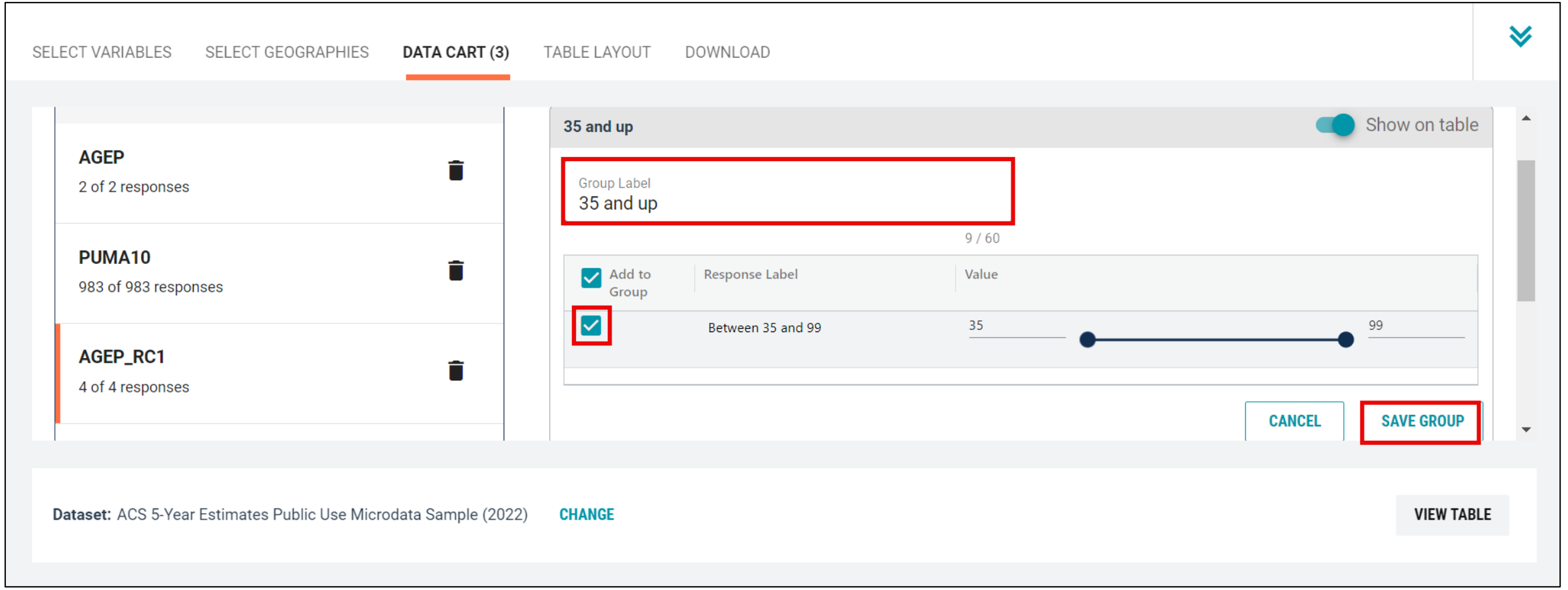
- Create recode for AGEP variable:
  - Change Group Label to **’26 to 34’**
  - Click on the checkbox next to **Between 26 to 99** and edit the end range to **34**
  - Click on the ‘**Save Group**’ button.



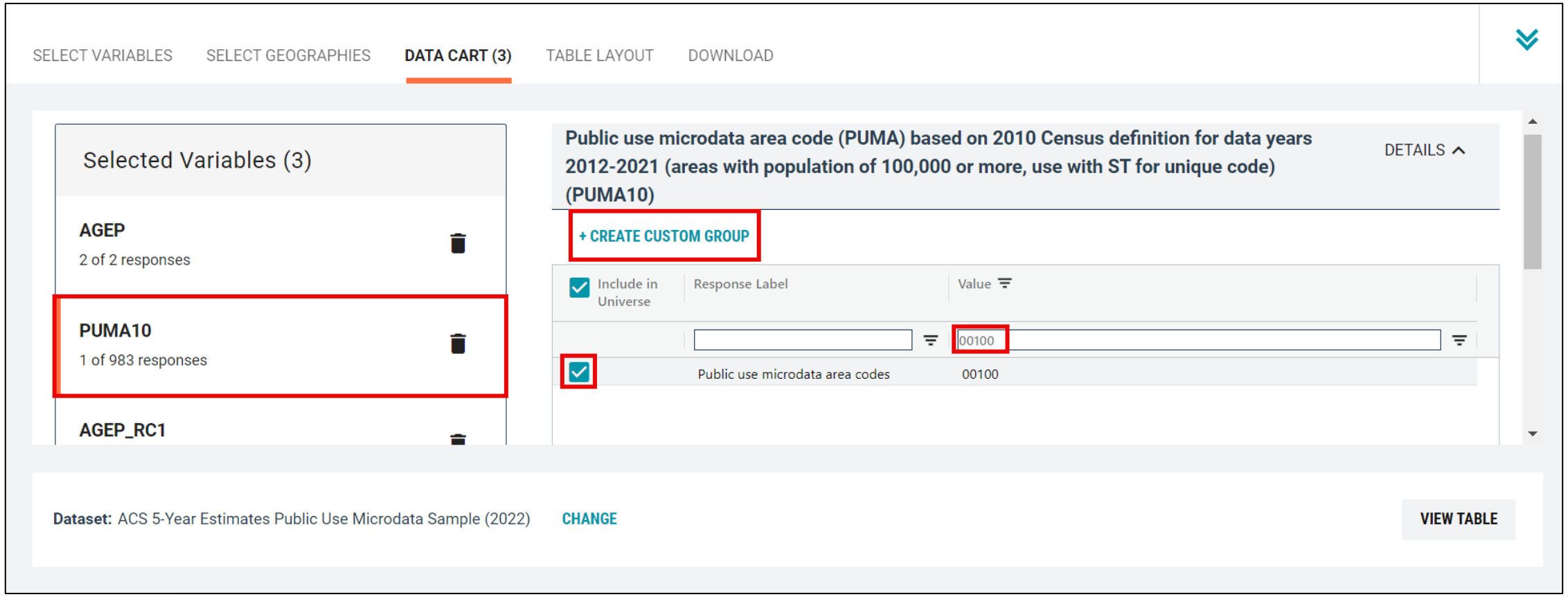
- Create recode for AGEP variable:
  - Click **Edit Group** for “Not Elsewhere Classified”



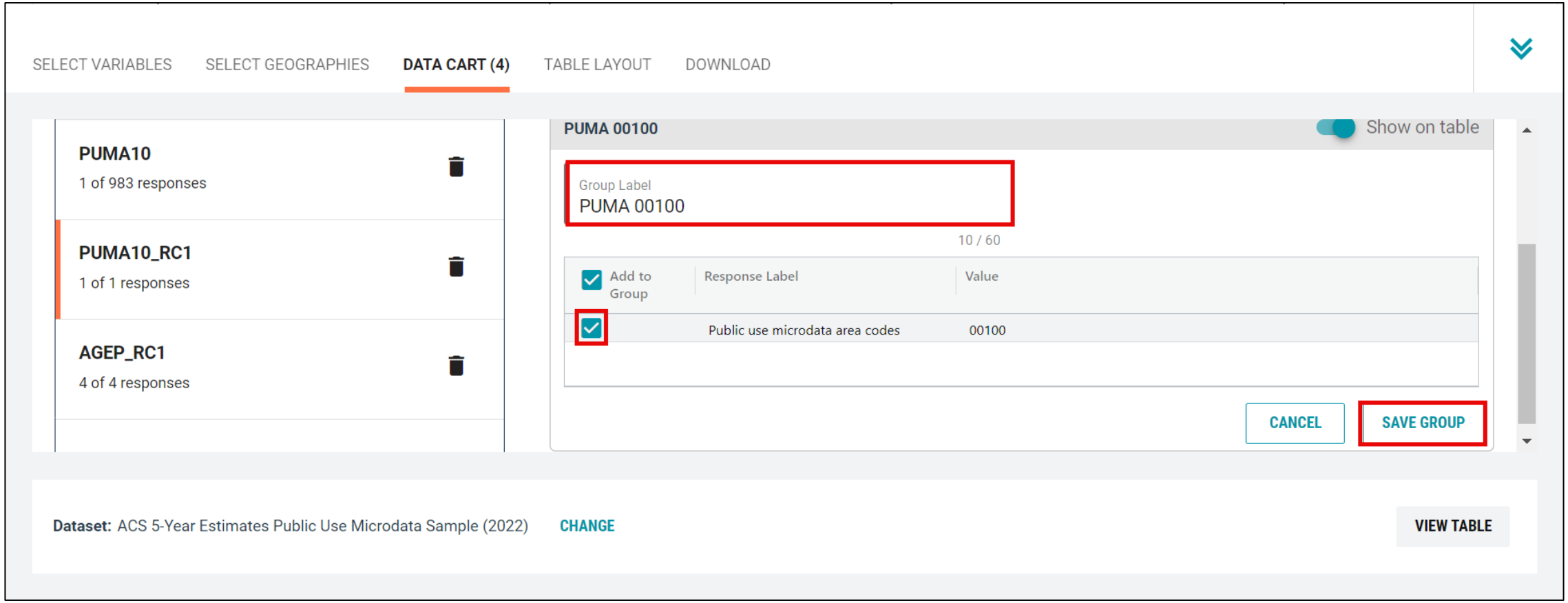
- Create recode for AGEP variable:
  - Change Group Label to **35 and up**
  - Click on the checkbox next to **Between 35 to 99** click on the '**Save Group**' button.



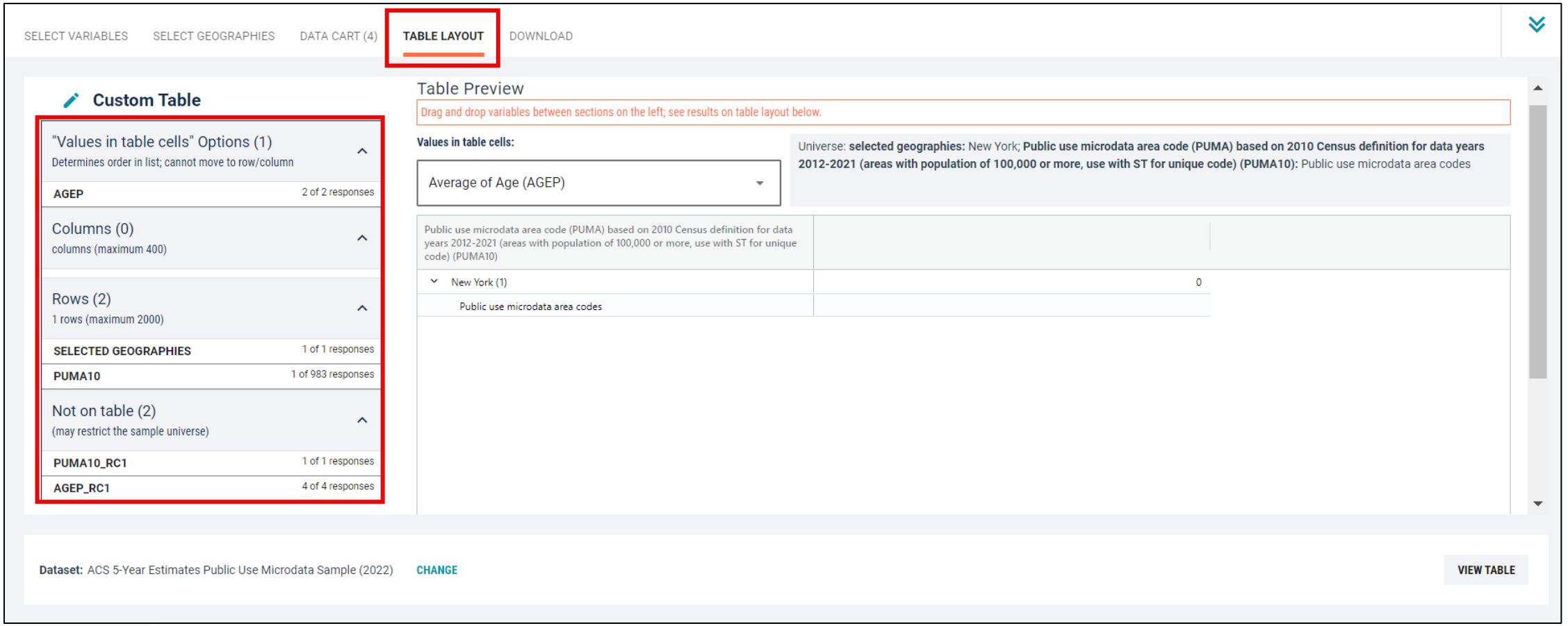
- Create recode to name PUMA10 variable:
  - **Select PUMA10** and click on the '**Include in Universe**' checkbox to uncheck all selected PUMAs
  - Reselect **00100** or use the Value search box to search for your desired PUMAs
  - Click on the **Create Custom Group** button to name your PUMA



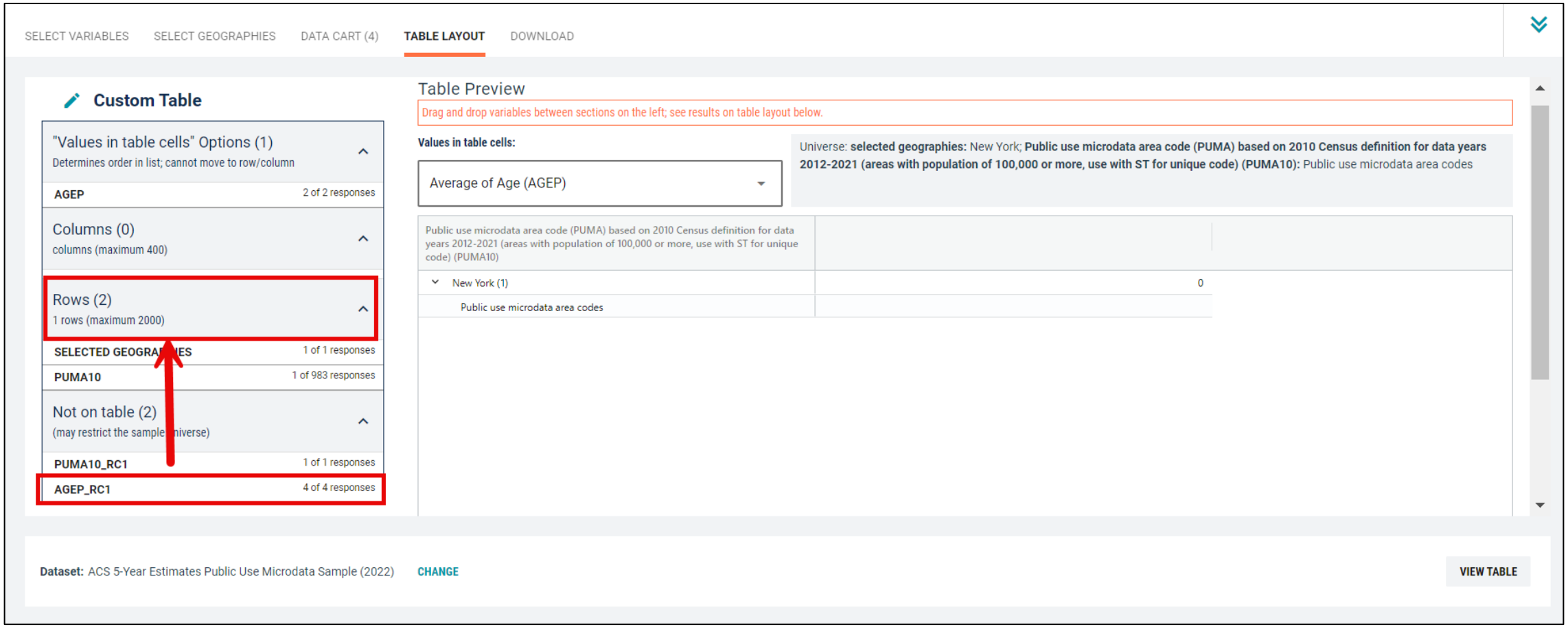
- Create recode to name PUMA10 variable:
  - Use the **Group Label box to type in 'PUMA 00100'** and **select your PUMA** by clicking on the checkbox
  - Click the **Save Group** button



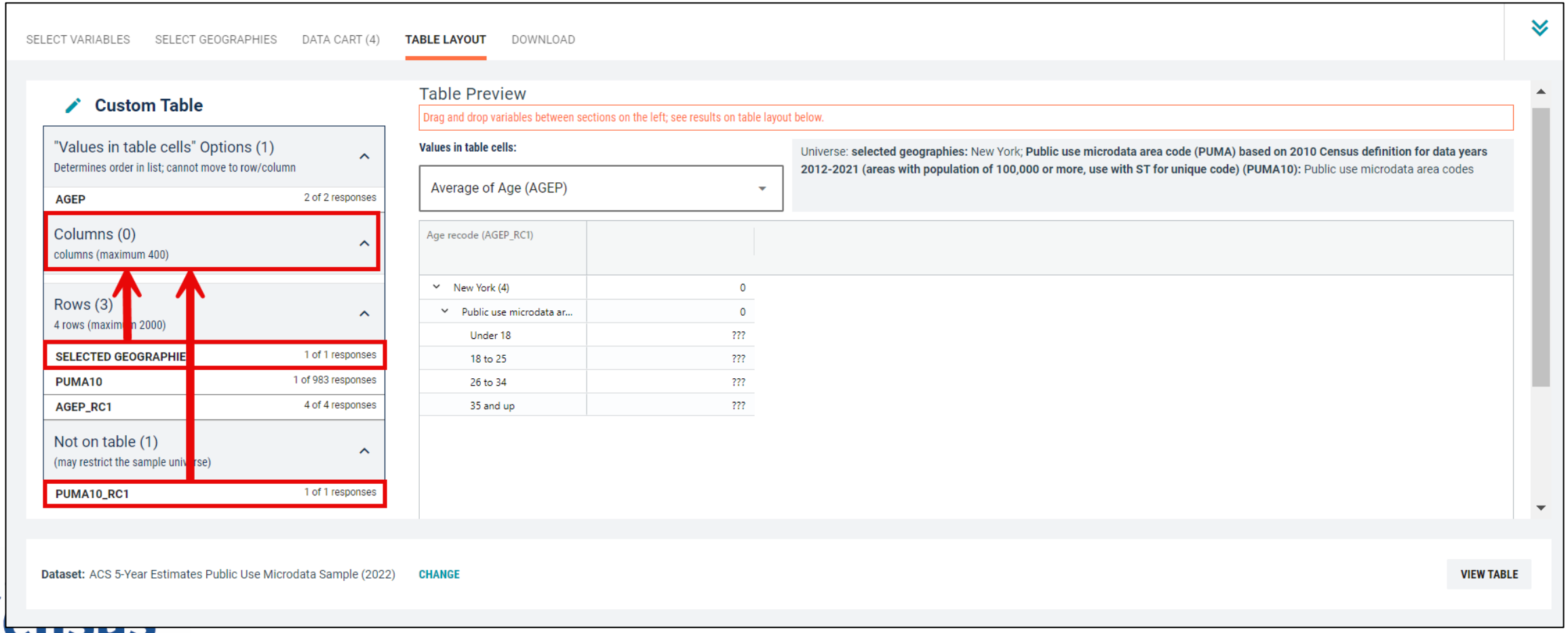
- View variable placement in the default table layout:
  - Move to the **Table Layout** tab
  - **Columns/Rows – Variables will be shown in the table.** By default, the table is providing the average age with the original PUMA variable and Selected Geographies in the Rows.



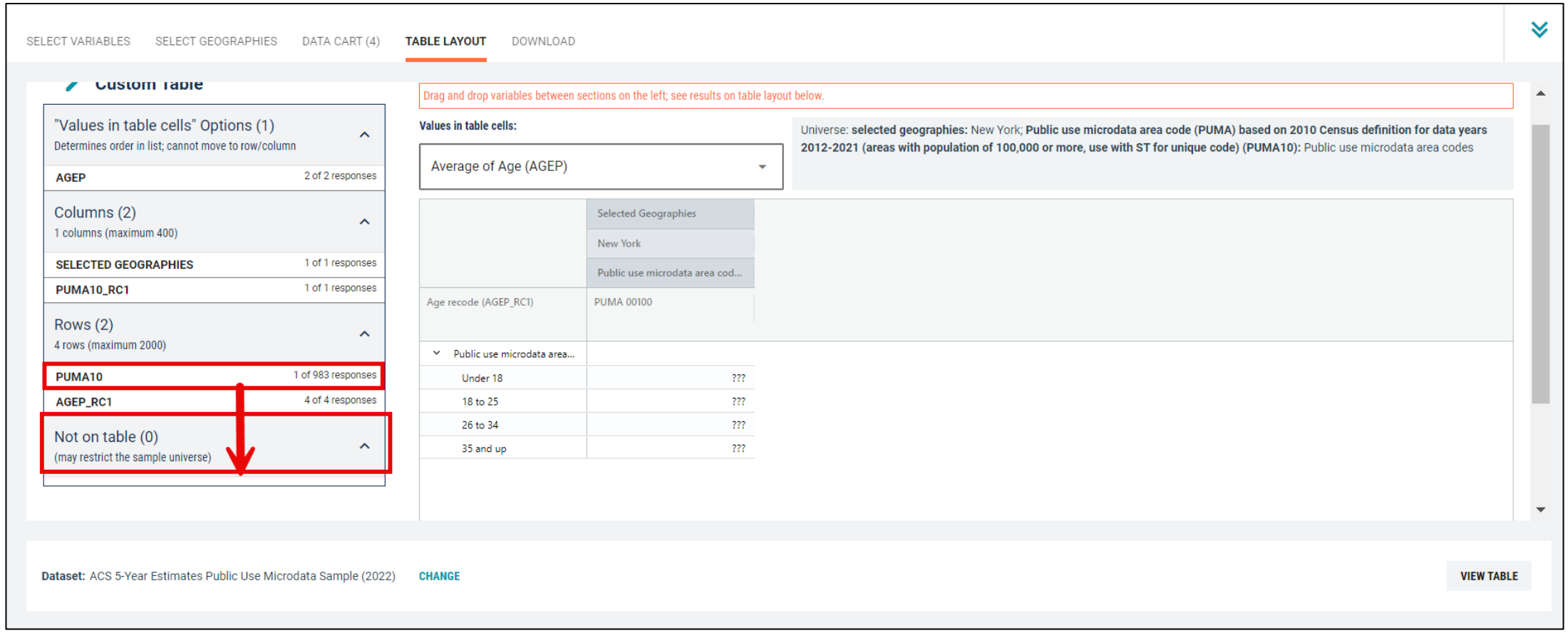
- Edit Table Layout:
  - Move Age Recode to Rows:
    - **Click, hold and drag AGEP_RC1 on the left side of the page up to the rows heading.** This will give you a table layout that includes the age categories that were created as the rows.



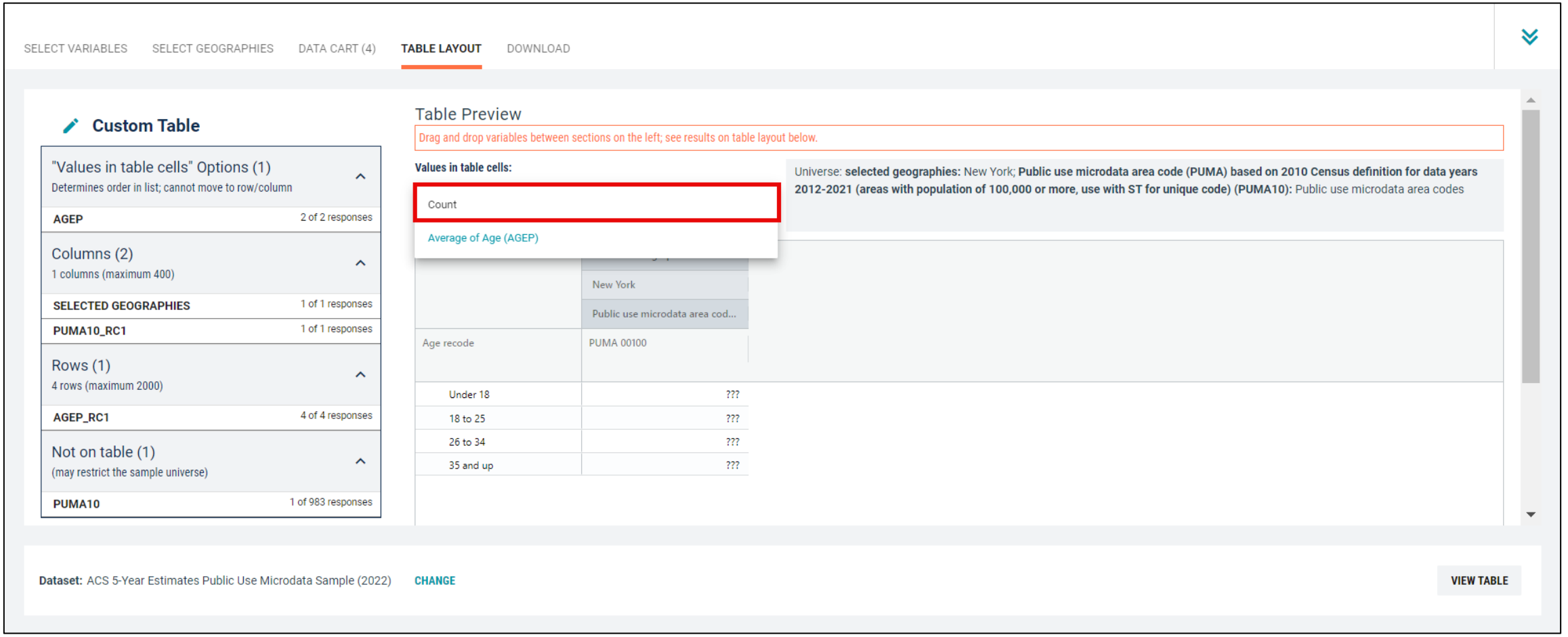
- Edit Table Layout:
  - Move Selected Geographies and PUMA recode to Columns:
    - Click, hold and drag Selected Geographies and PUMA10_RC1 on the left side of the page up to the **columns heading.** This will give you a table layout that includes the selected PUMA 00100 from New York as the columns.



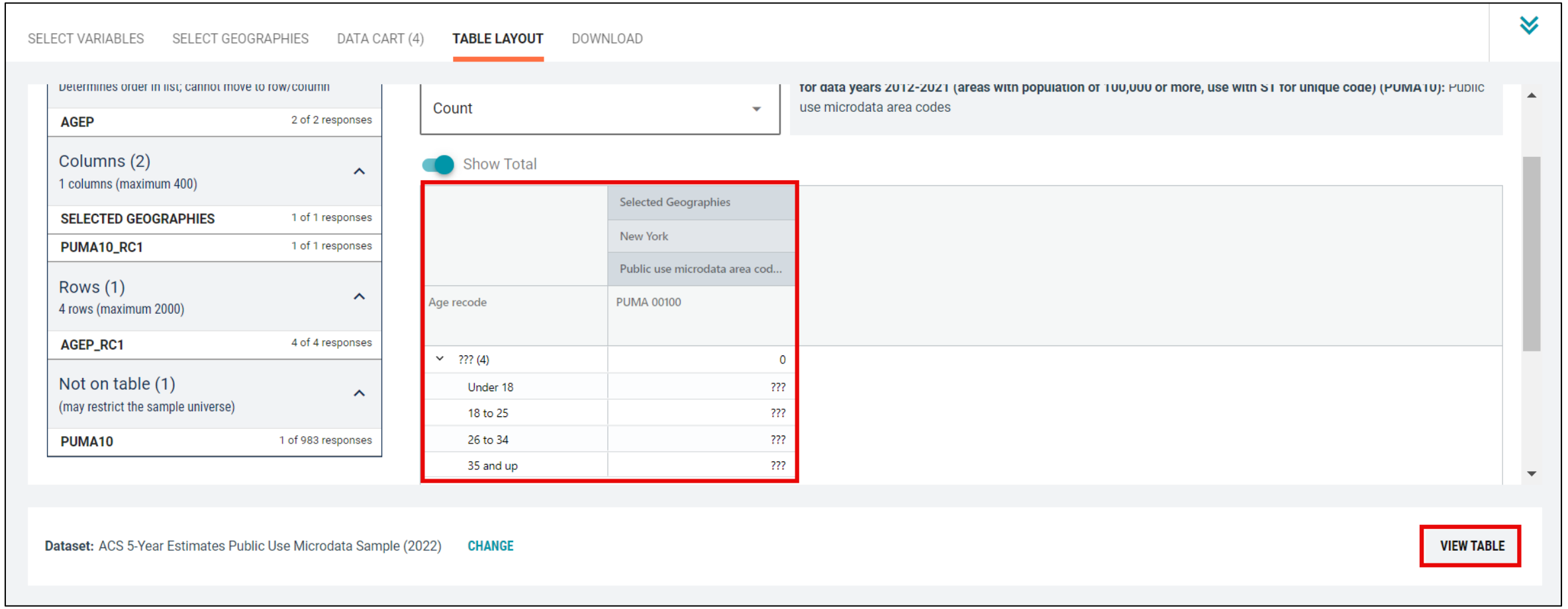
- Edit Table Layout:
  - Move original PUMA10 variable to Not on table section:
    - **Click, hold and drag PUMA10 to the Not on table section.** This will give you a table layout that includes the selected PUMA from New York as the columns.



- Choose type of values in table cells
  - Change the “Value in table cells” option from Average of Age (AGEP) to **Count.** This will give you data for the total number of people within the requested categories.

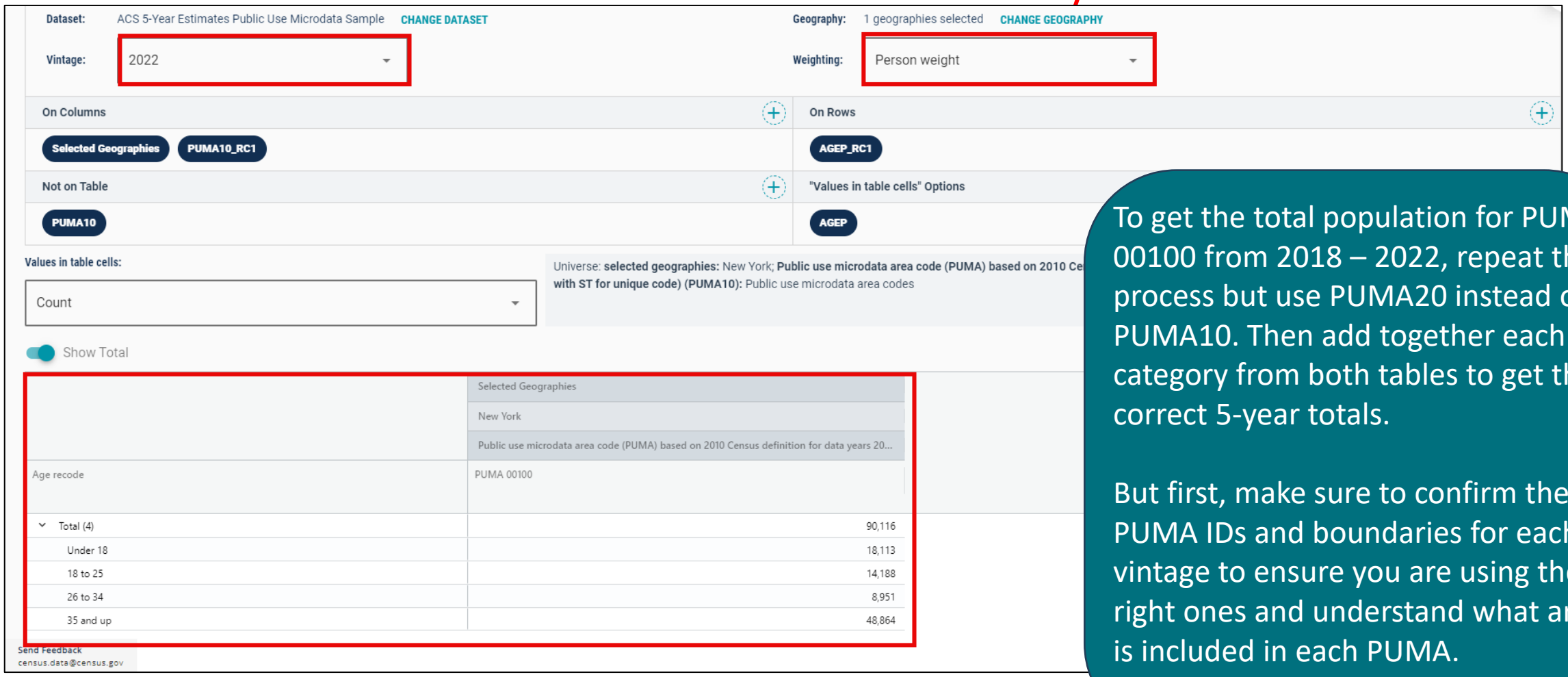


- Confirm Table Layout:
  - Confirm table layout and click **View Table** in the lower right



View Table

Note that the site automatically chooses a weight for you. You do have the option to change the weight if you want.



To get the total population for PUMA 00100 from 2018 – 2022, repeat this process but use PUMA20 instead of PUMA10. Then add together each age category from both tables to get the correct 5-year totals.

But first, make sure to confirm the PUMA IDs and boundaries for each vintage to ensure you are using the right ones and understand what area is included in each PUMA.

United States® Census Bureau

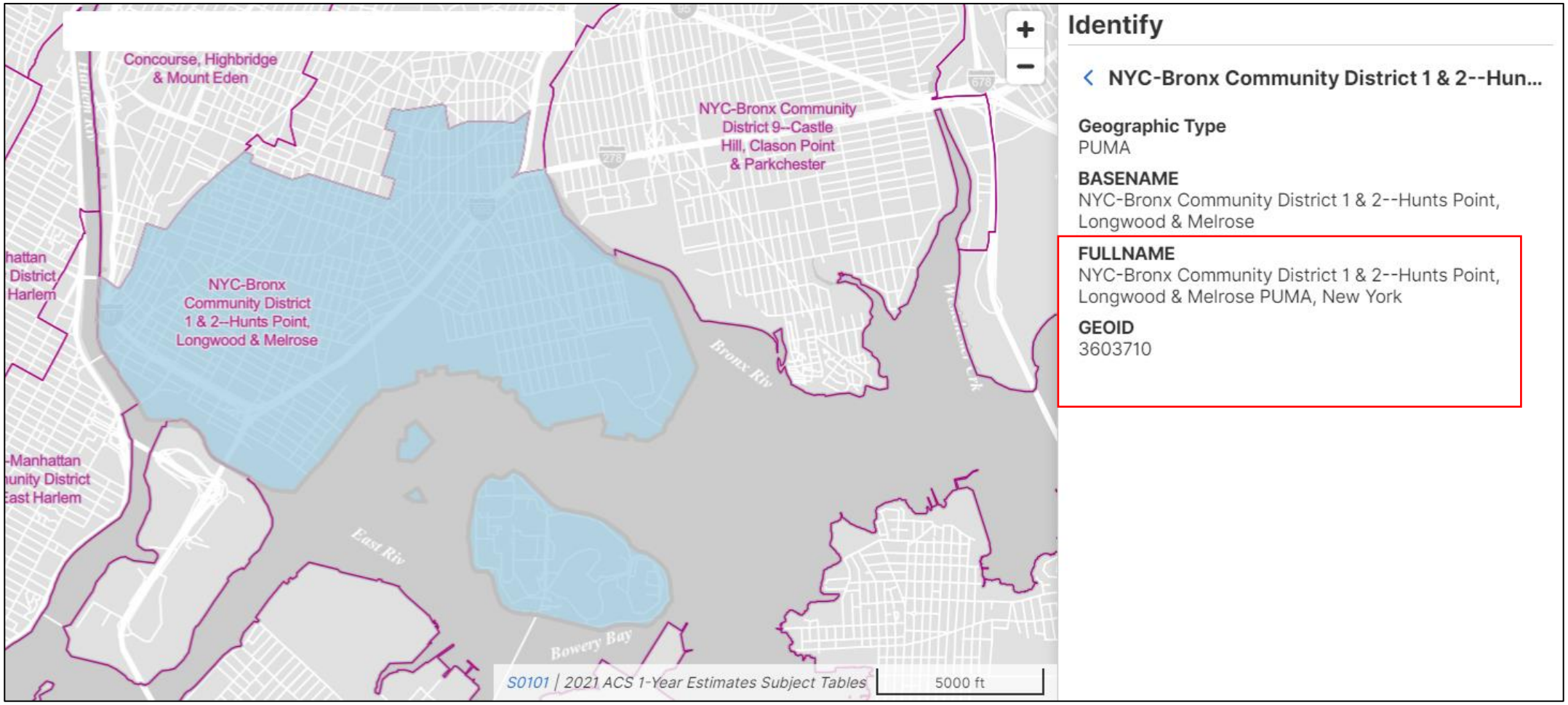
Demo

Example 3:

PUMA with slight name change, and different GEOIDs

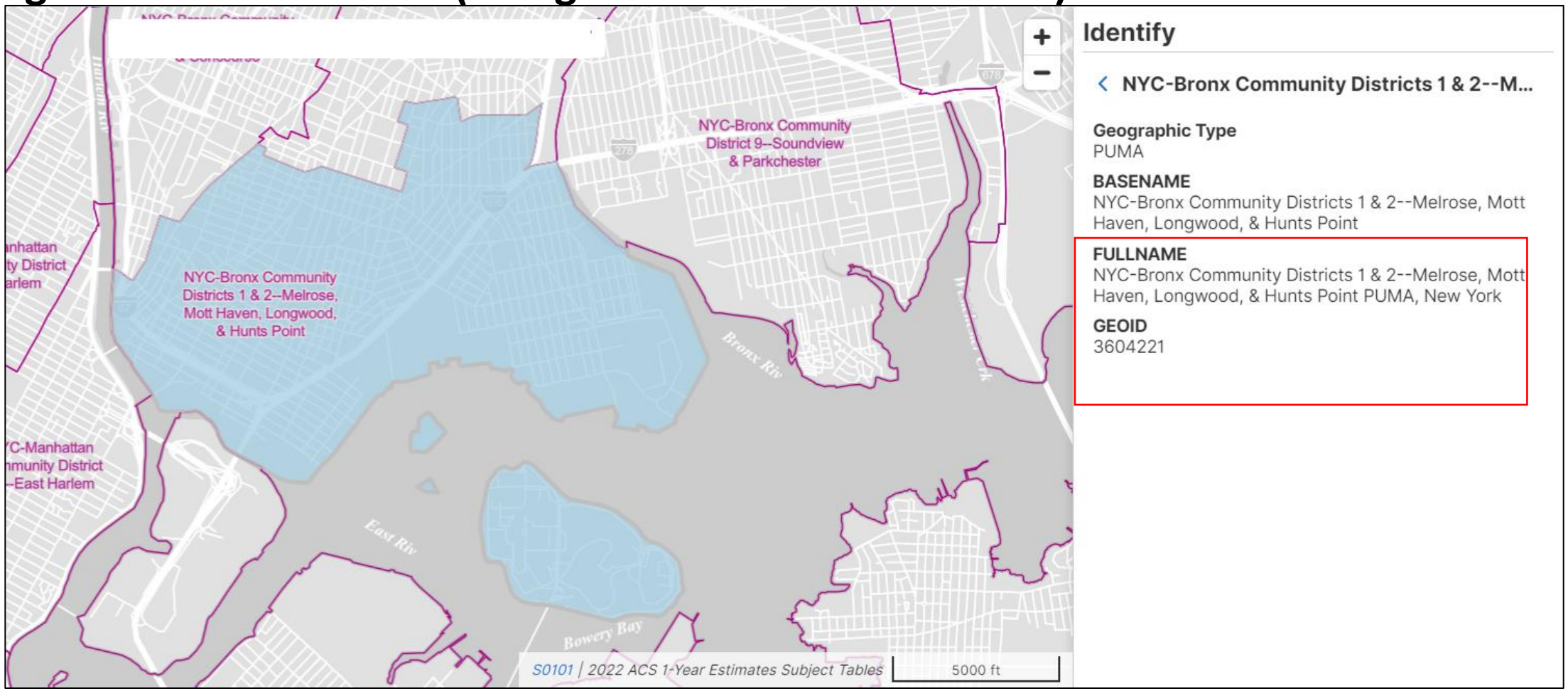
United States® Census Bureau

2021 Map for NYC Bronx Community District 1 & 2 – Hunts Point, Longwood & Melrose (Using 2010 PUMA Boundaries)



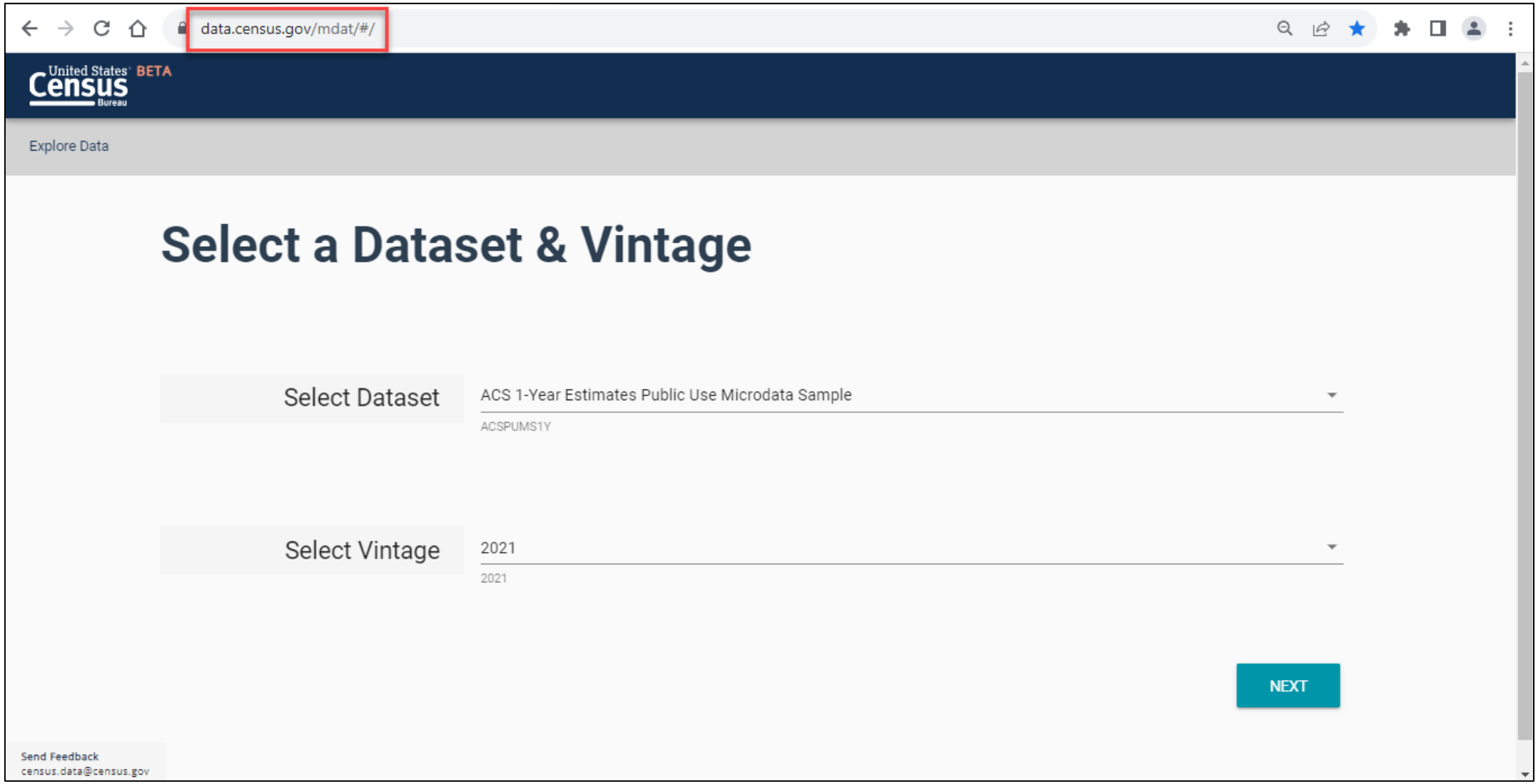
United States® Census Bureau

2022 Map for NYC Bronx Community District 1 & 2 –Melrose, Mott Haven, Longwood & Hunts Point (Using 2020 PUMA Boundaries)



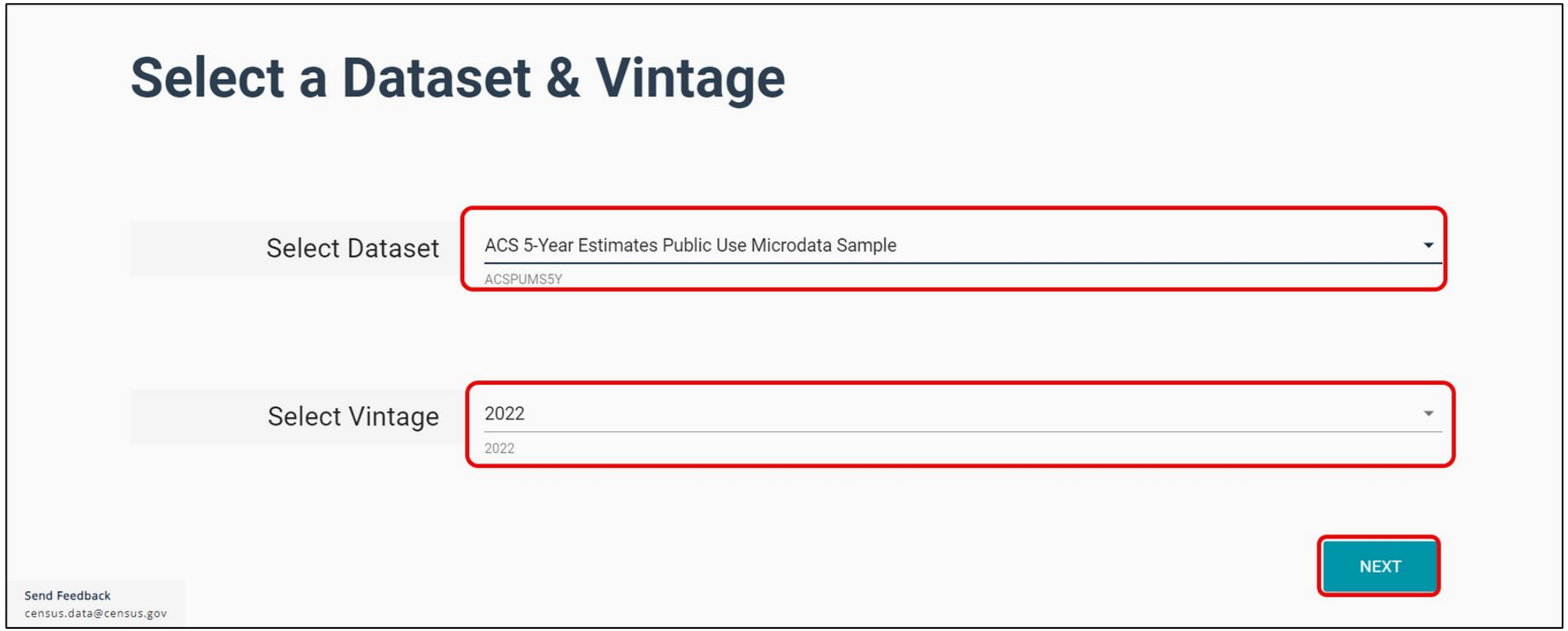
United States® Census Bureau

- Visit Microdata Access at data.census.gov/mdat

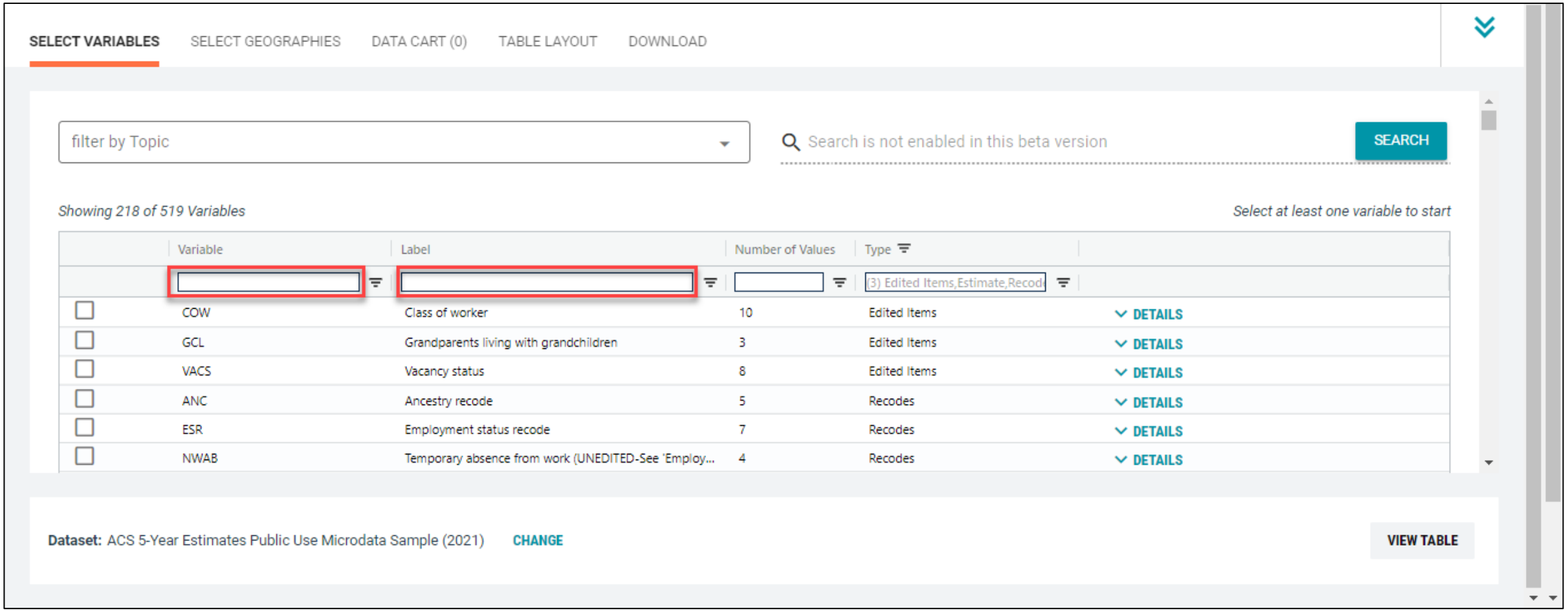


data.census.gov/mdat

- Choose Dataset and Vintage:
  - Dataset – **ACS 5-Year Estimates – Public Use Microdata Sample**
  - Vintage – **2022**
  - Click **Next** in the lower right

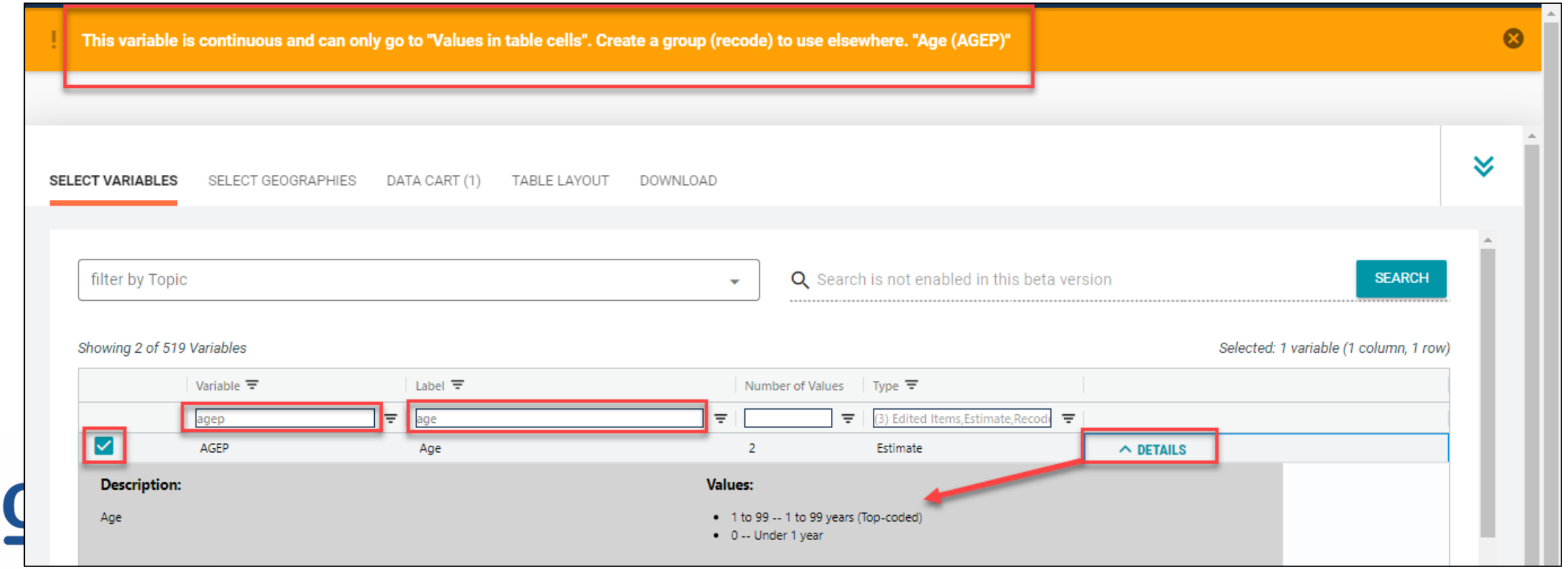


- **Search for Variables**: Use the search box below “Variable” or “Label” to find your variables of interest

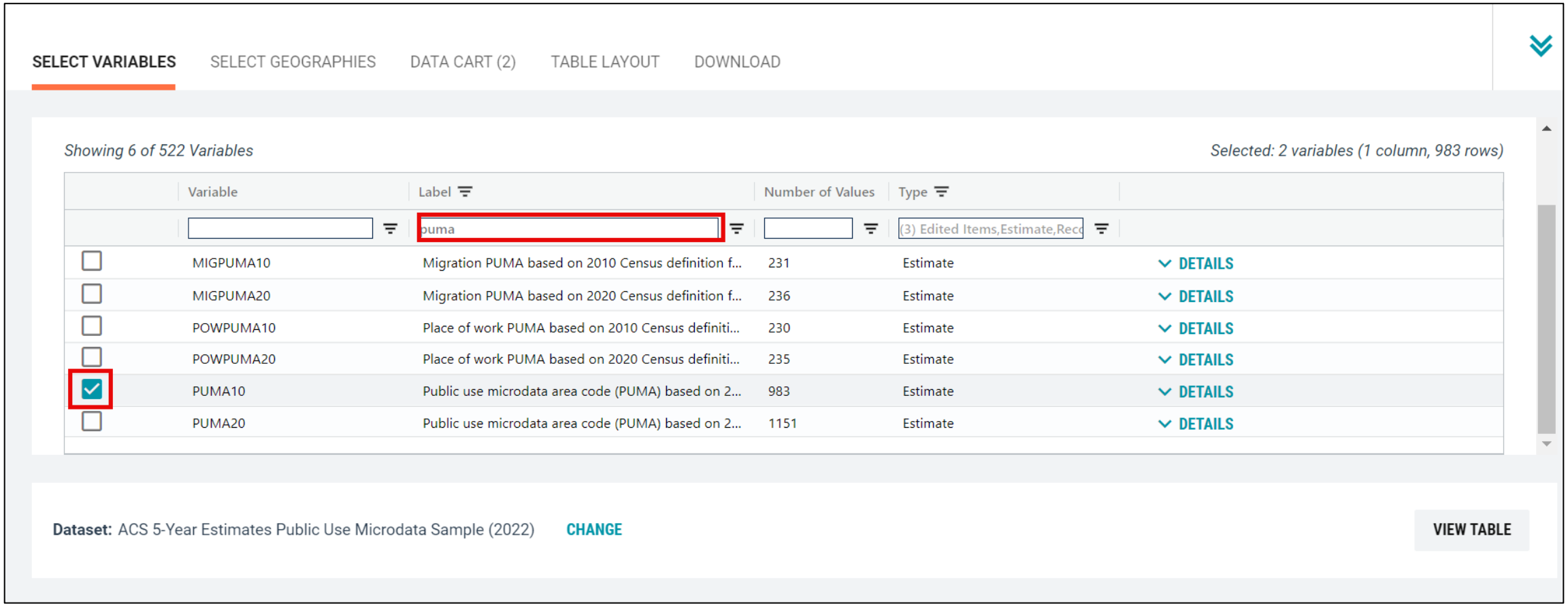


United States® Census Bureau

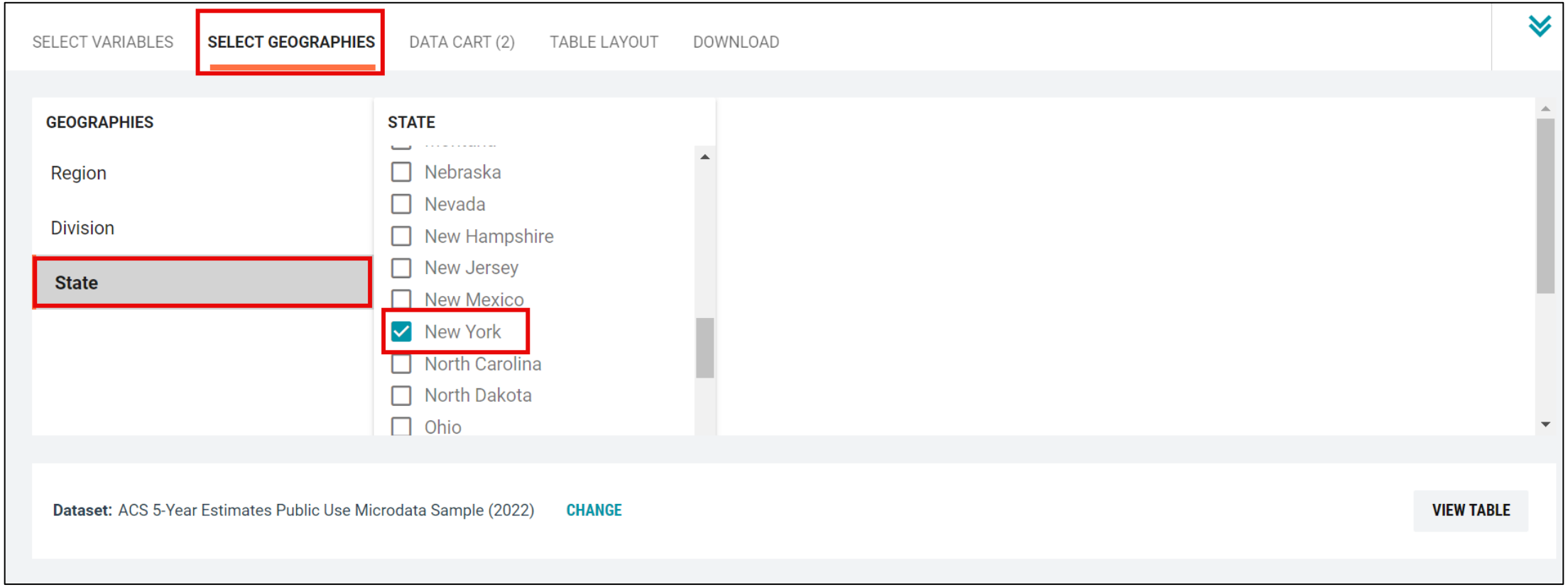
- Select variable for Age:
  - Type “AGEP” in the Variable search box or type “Age” in the label search box
  - Check the box to the left of AGEP to add the variable to your data cart
  - Notice the message at the top of the screen saying you will need to create your own categories (or recodes) for this variable if you want it shown in the table. (You will do this action in the Data Cart)



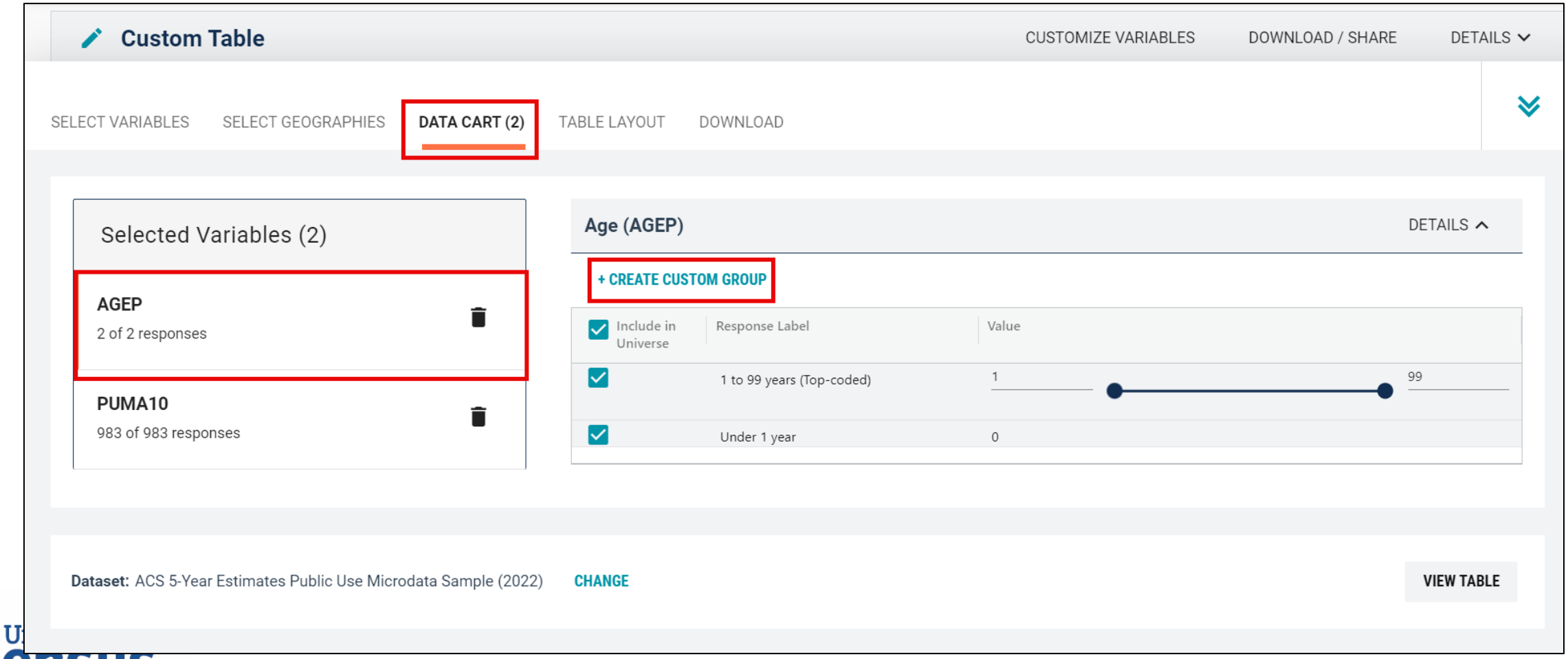
- Select variable for 2010 PUMAs:
  - Type ‘PUMA’ in the label search box
  - Check the box to the left of ‘PUMA10’ to add the variable to data cart



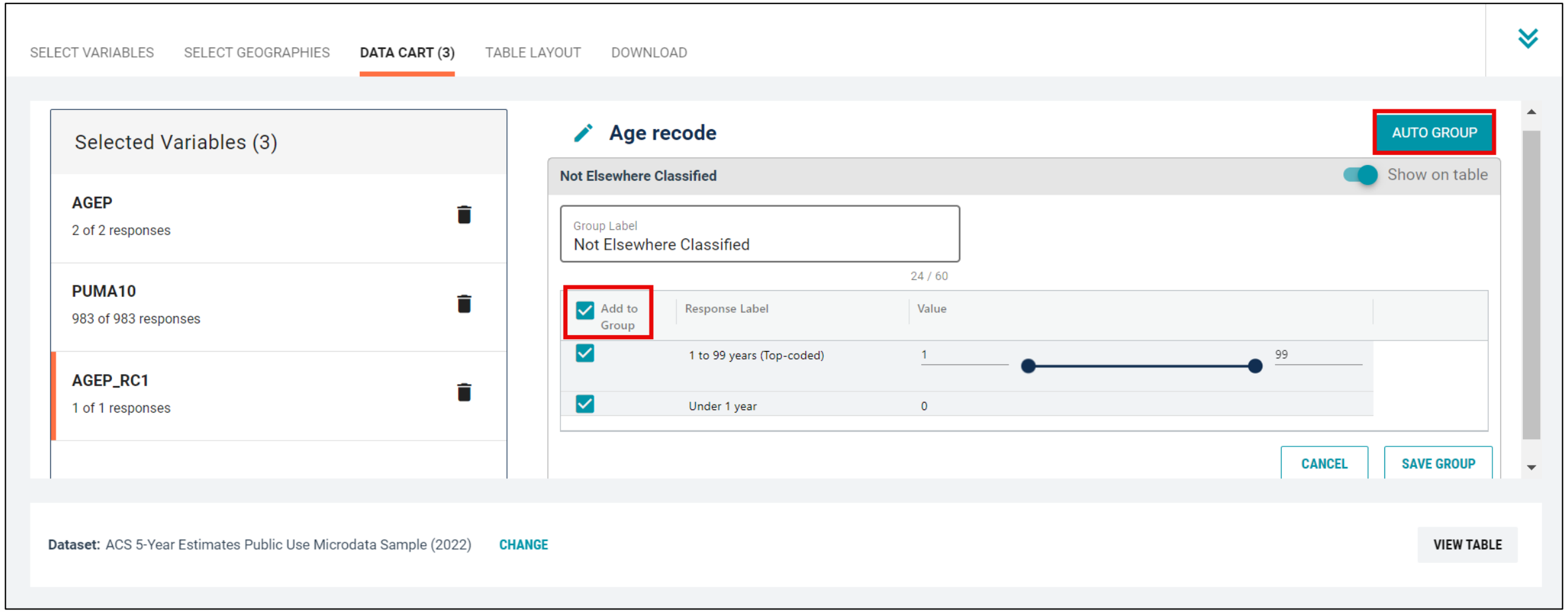
- **Select New York state geography.**
  - Click on the Select Geographies tab
  - Check the box to the left ‘New York’ to only pull up data for PUMAs from New York state



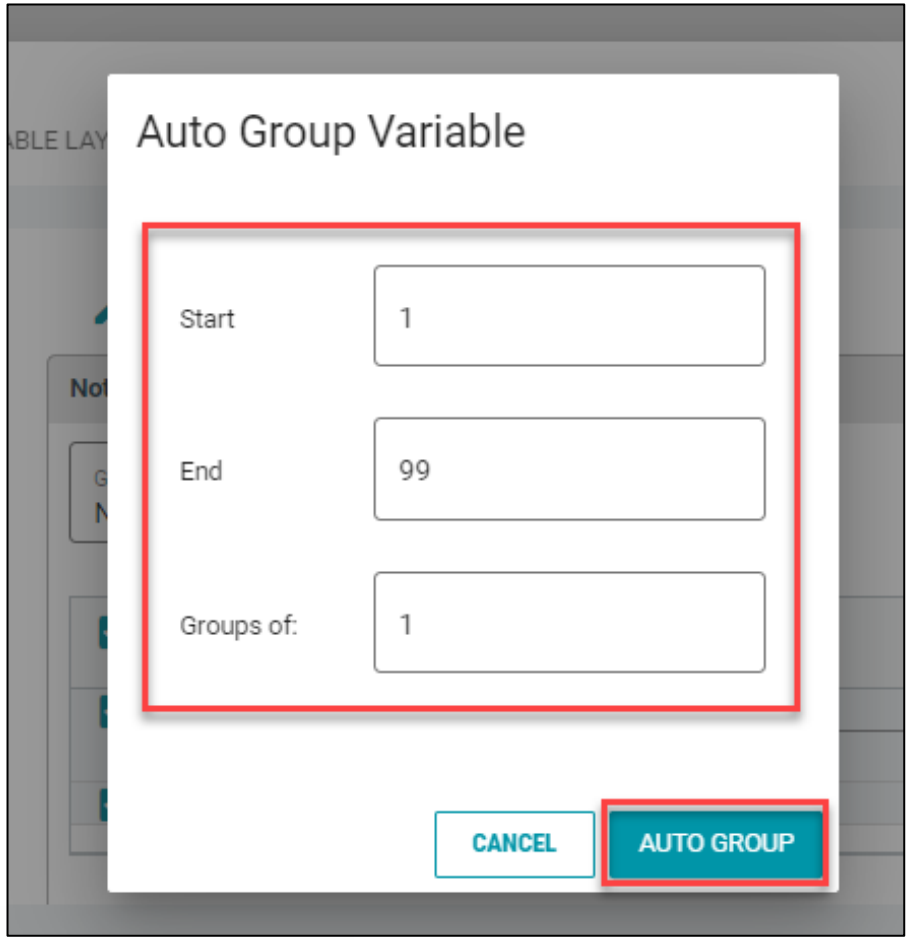
- Categorize (recode) your variable:
  - Move to the **Data Cart** tab
  - Click the **AGEP** variable on the left
  - Click **Create Custom Group** to begin specifying your age groups (e.g. single years of age)

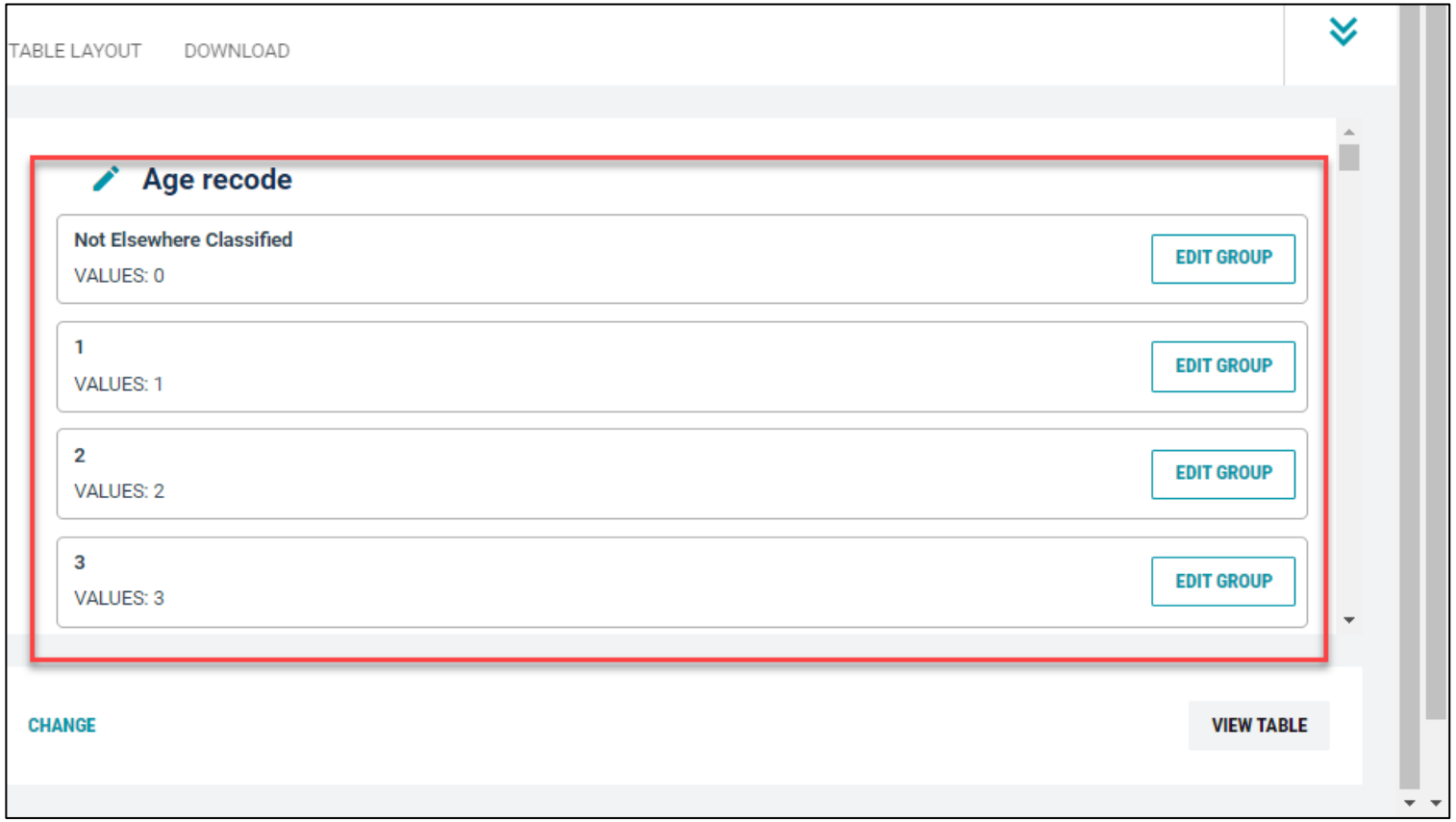


- Categorize (recode) your variable:
  - Check the box next to Add to Group to add both categories to the recode
  - Click on **Auto Group**

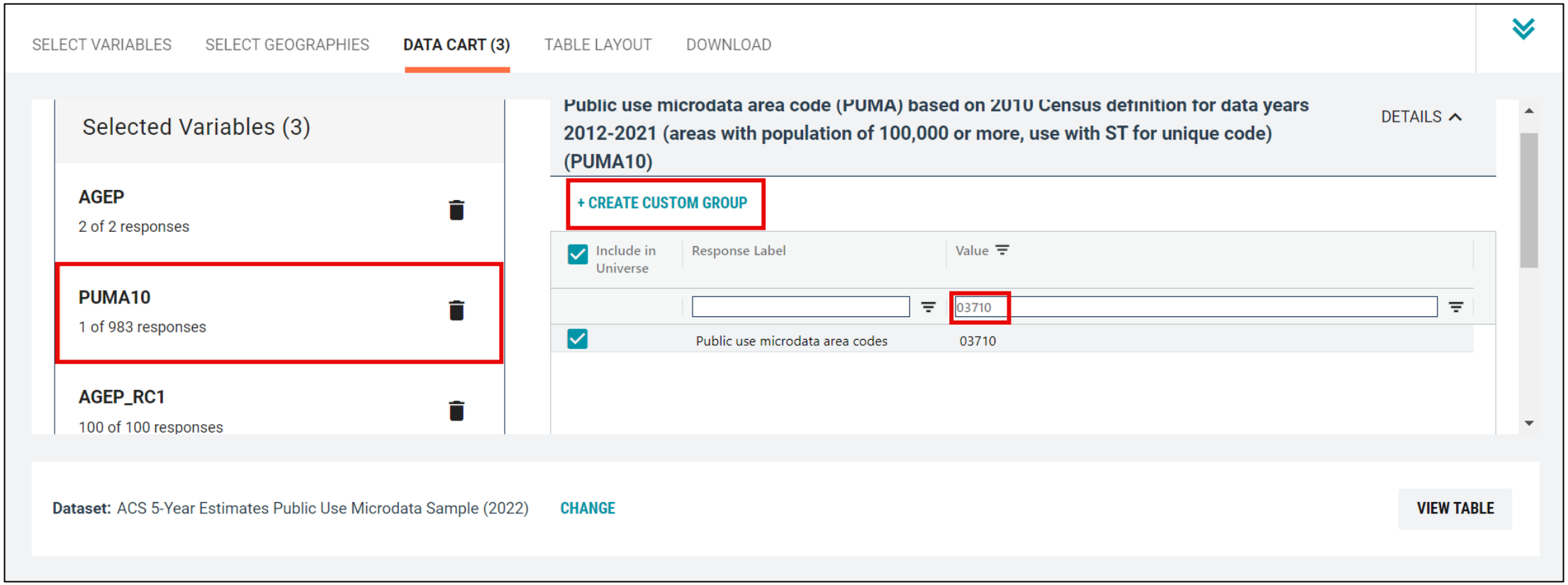


- **Categorize (recode) your variable**:
  - Confirm that the Start value is ‘1’, the End value is ‘99’, and the Groups of value is ‘1’
  - Click **Auto Group**. This will automatically create each year of age as its own group.

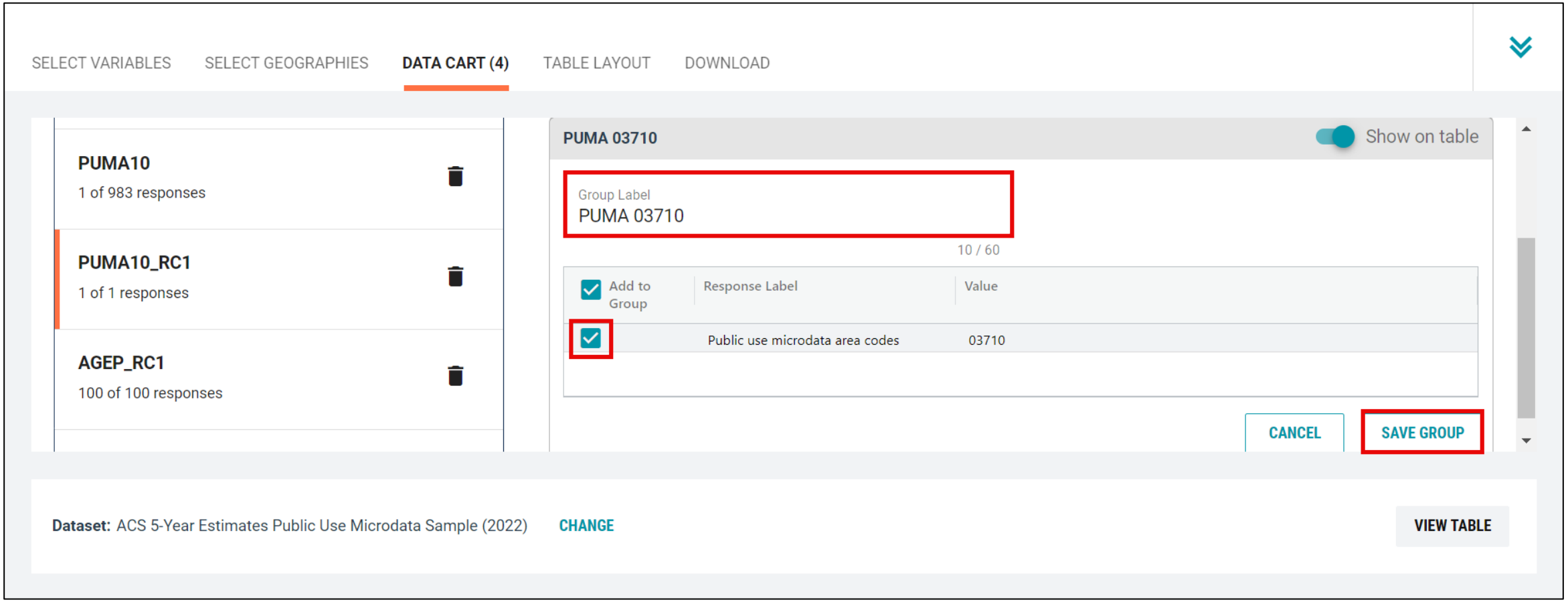




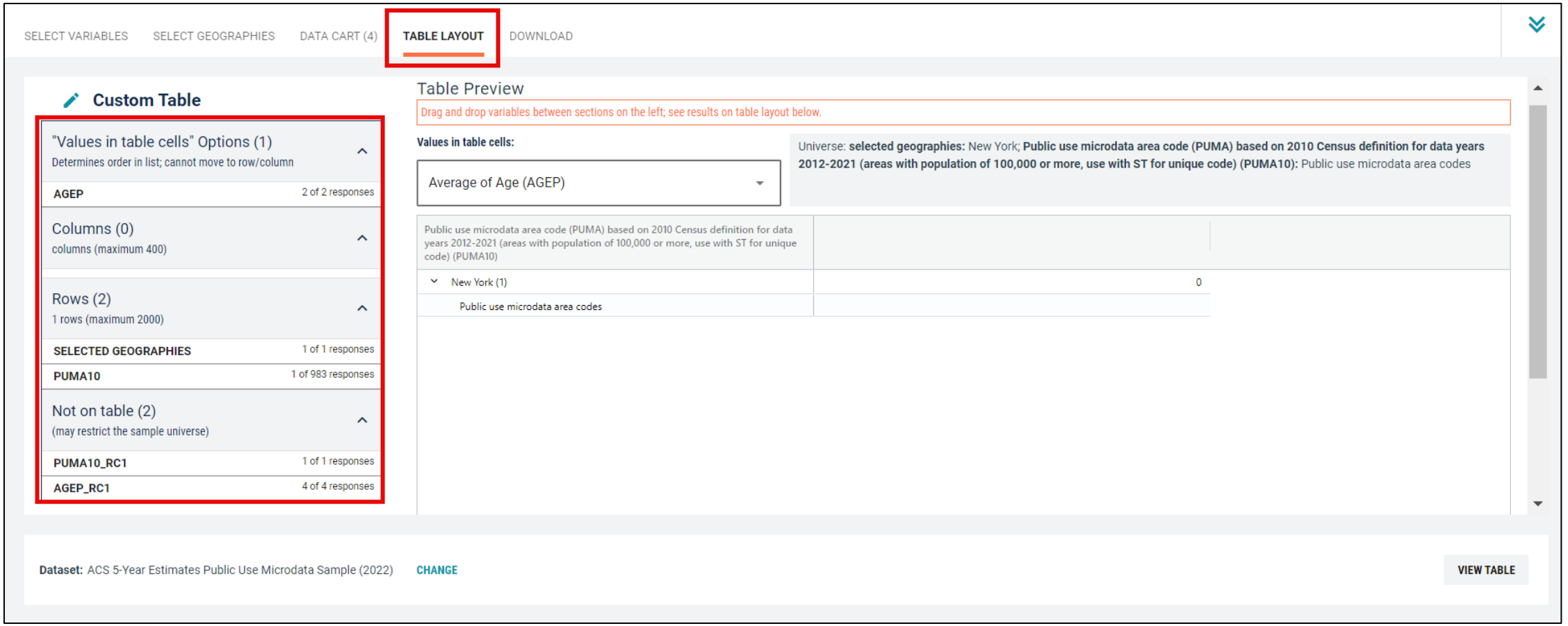
- Create recode to name PUMA10 variable:
  - **Select PUMA10** and click on the '**Include in Universe**' checkbox to uncheck all selected PUMAs
  - Reselect **03710** or use the Value search box to search for your desired PUMAs
  - Click on the **Create Custom Group** button to name your PUMA



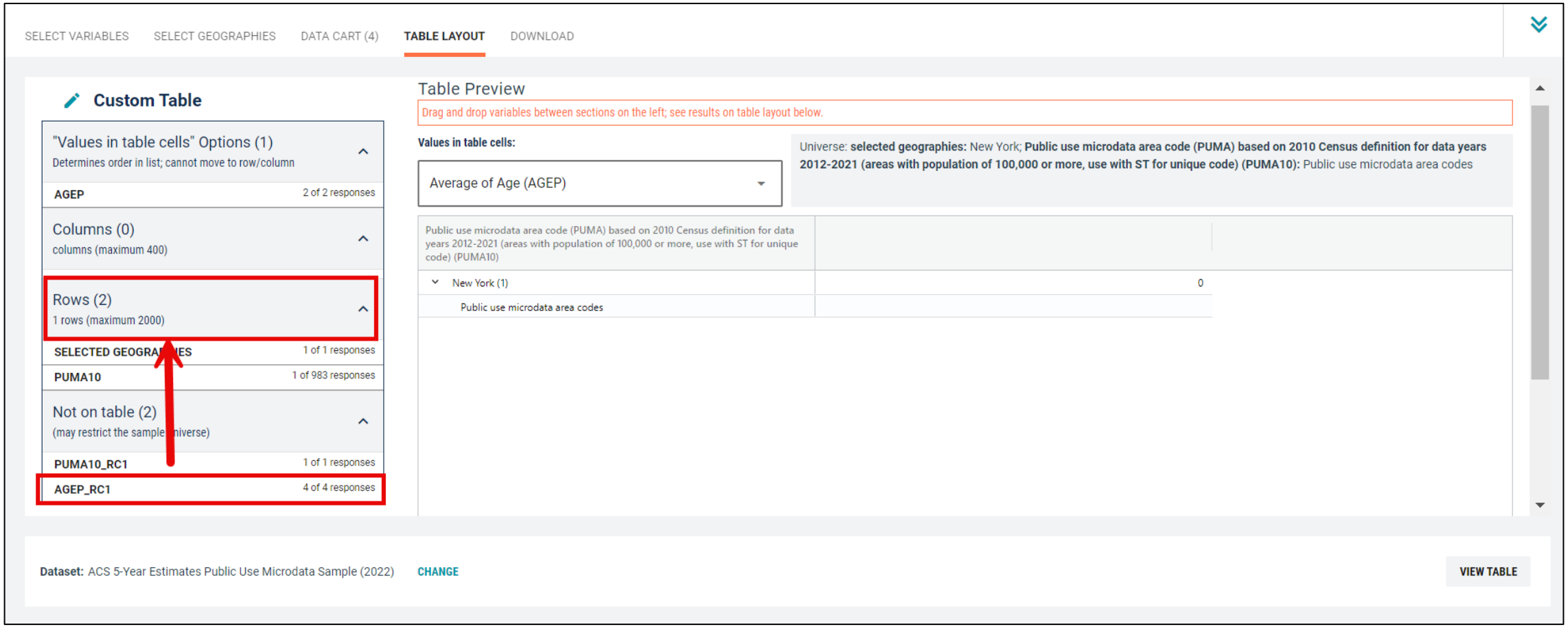
- Create recode to name PUMA10 variable:
  - Use the **Group Label box to type in 'PUMA 03710'** and **select your PUMA** by clicking on the checkbox
  - Click the **Save Group** button



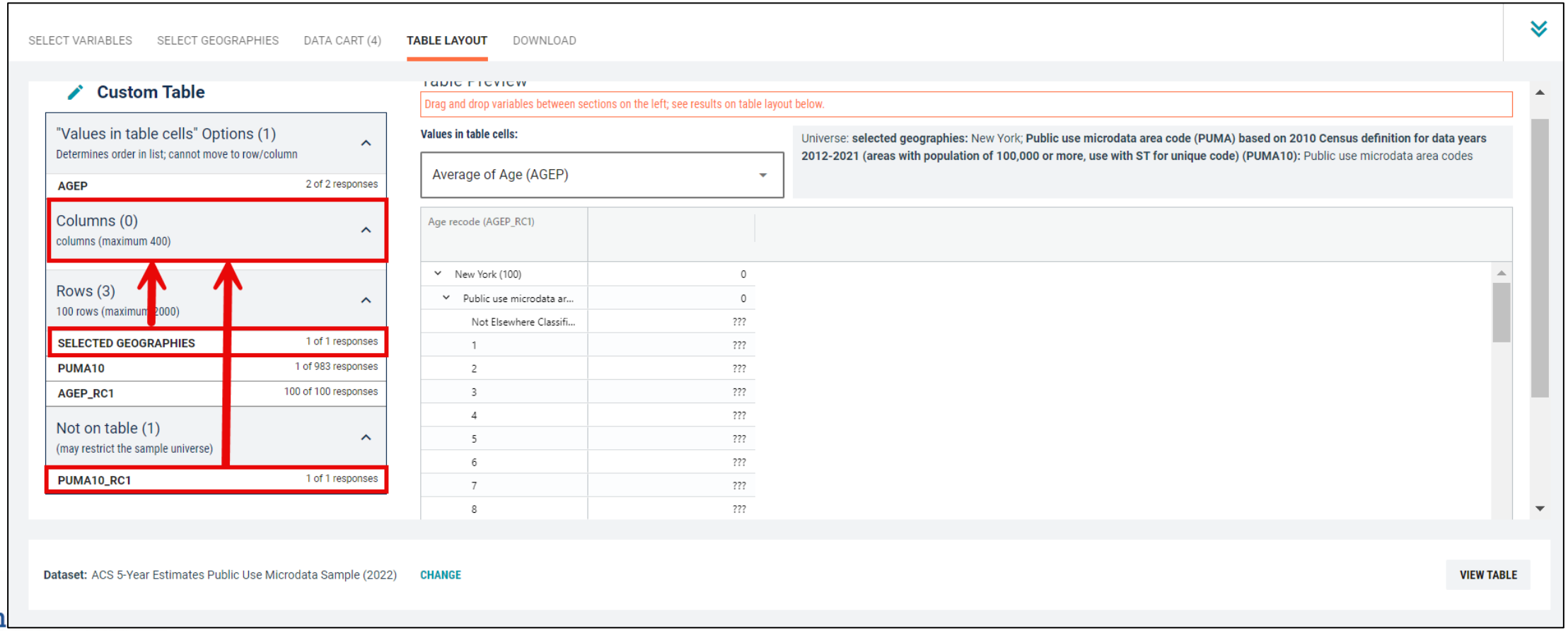
- View variable placement in the default table layout:
  - Move to the **Table Layout** tab
  - **Columns/Rows – Variables will be shown in the table.** By default, the table is providing the average age with the original PUMA variable and Selected Geographies in the Rows.



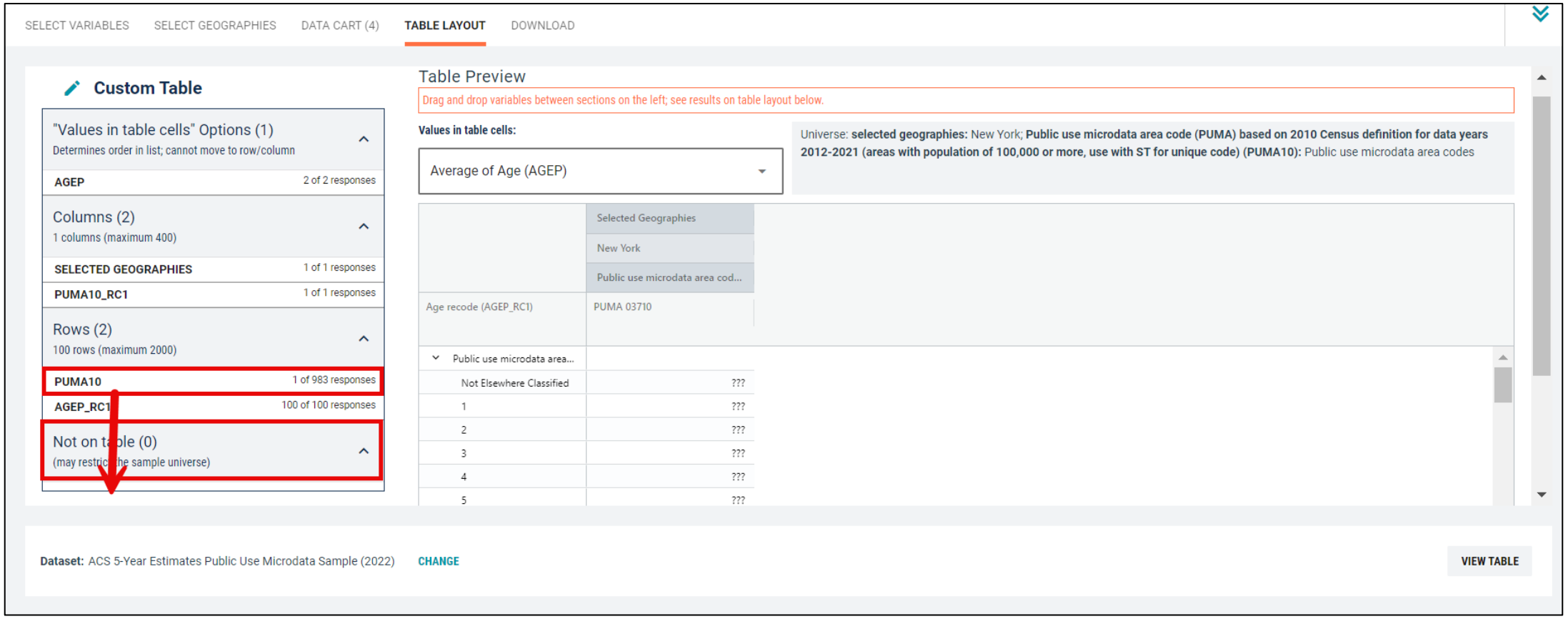
- Edit Table Layout:
  - Move Age Recode to Rows:
    - **Click, hold and drag AGEP_RC1 on the left side of the page up to the rows heading.** This will give you a table layout that includes the age categories that were created as the rows.



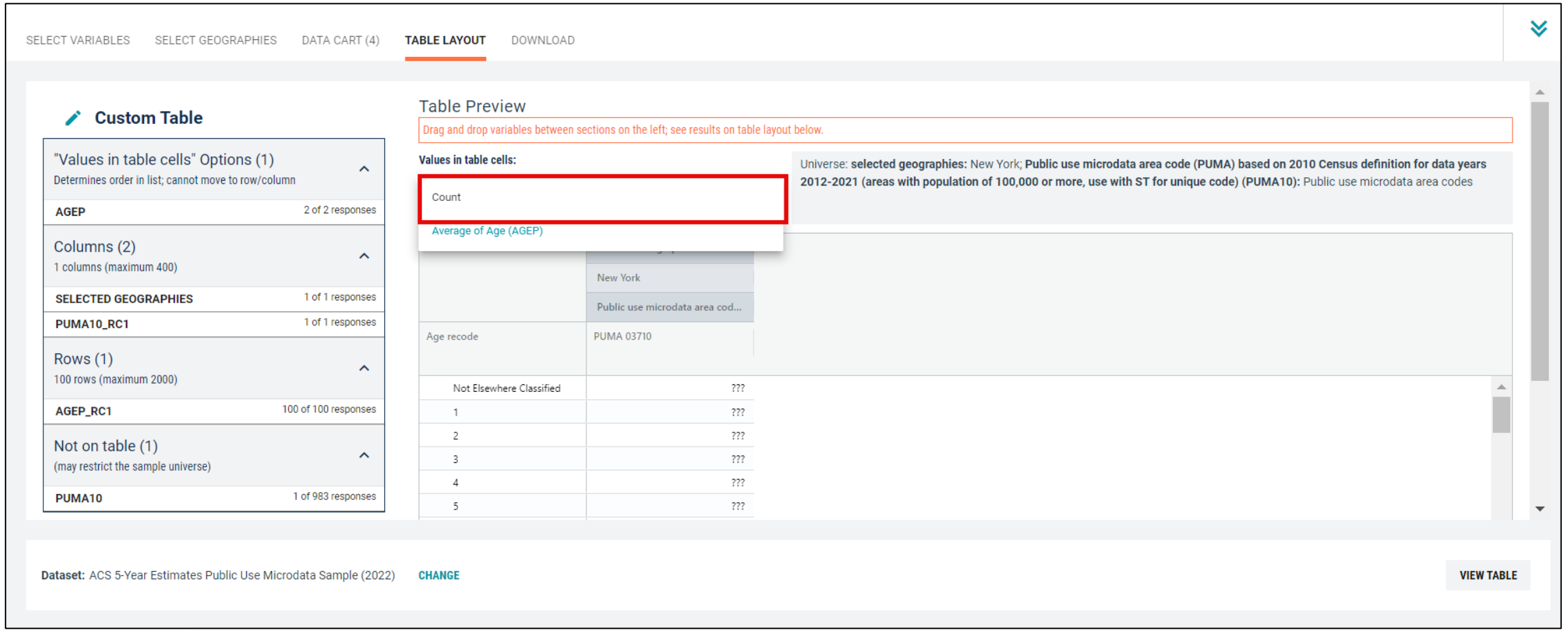
- Edit Table Layout:
  - Move Selected Geographies and PUMA recode to Columns:
    - Click, hold and drag Selected Geographies and PUMA10_RC1 on the left side of the page up to the **columns heading.** This will give you a table layout that includes the selected PUMA 03710 from New York as the columns.



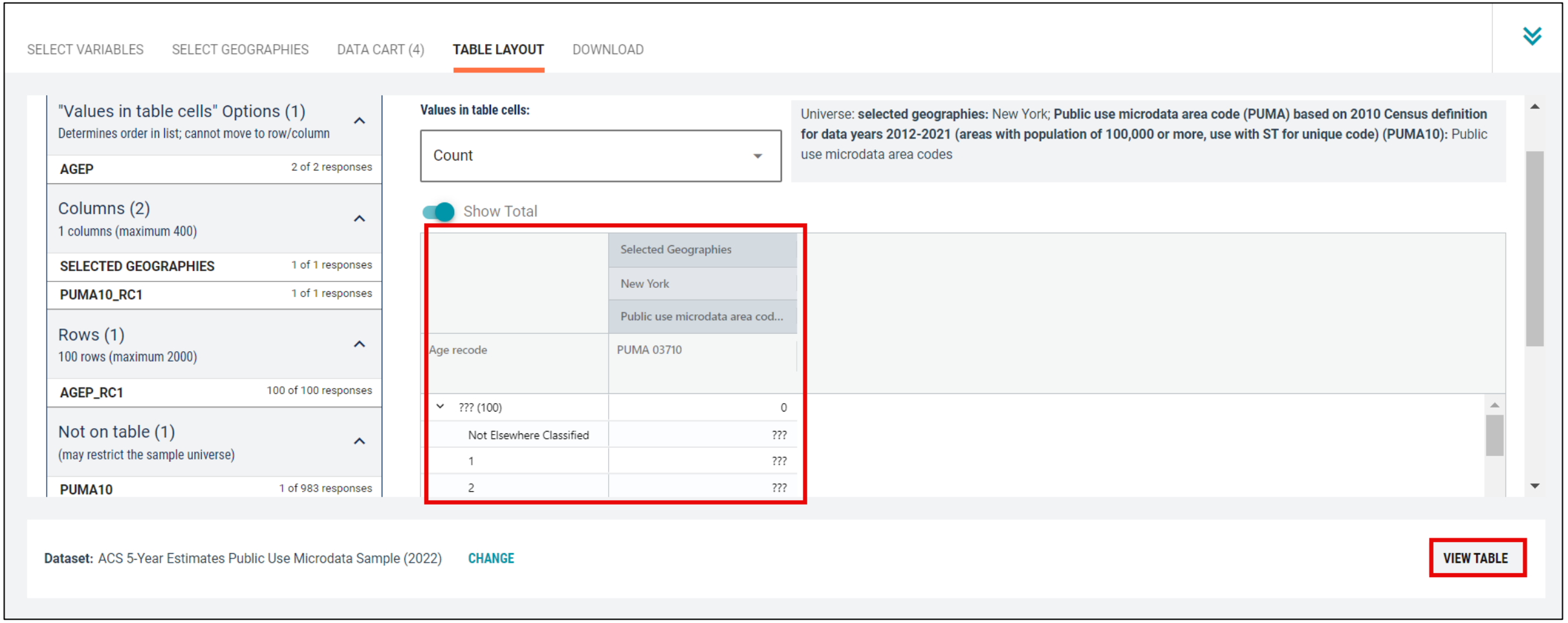
- Edit Table Layout:
  - **Move original PUMA10 variable to Not on table section:**
    - **Click, hold and drag PUMA10 to the Not on table section.** This will give you a table layout that includes the selected PUMA from New York as the columns.



- Choose type of values in table cells
  - Change the “Value in table cells” option from “Average of Age” to **Count.** This will give you data for the total number of people within the requested categories.

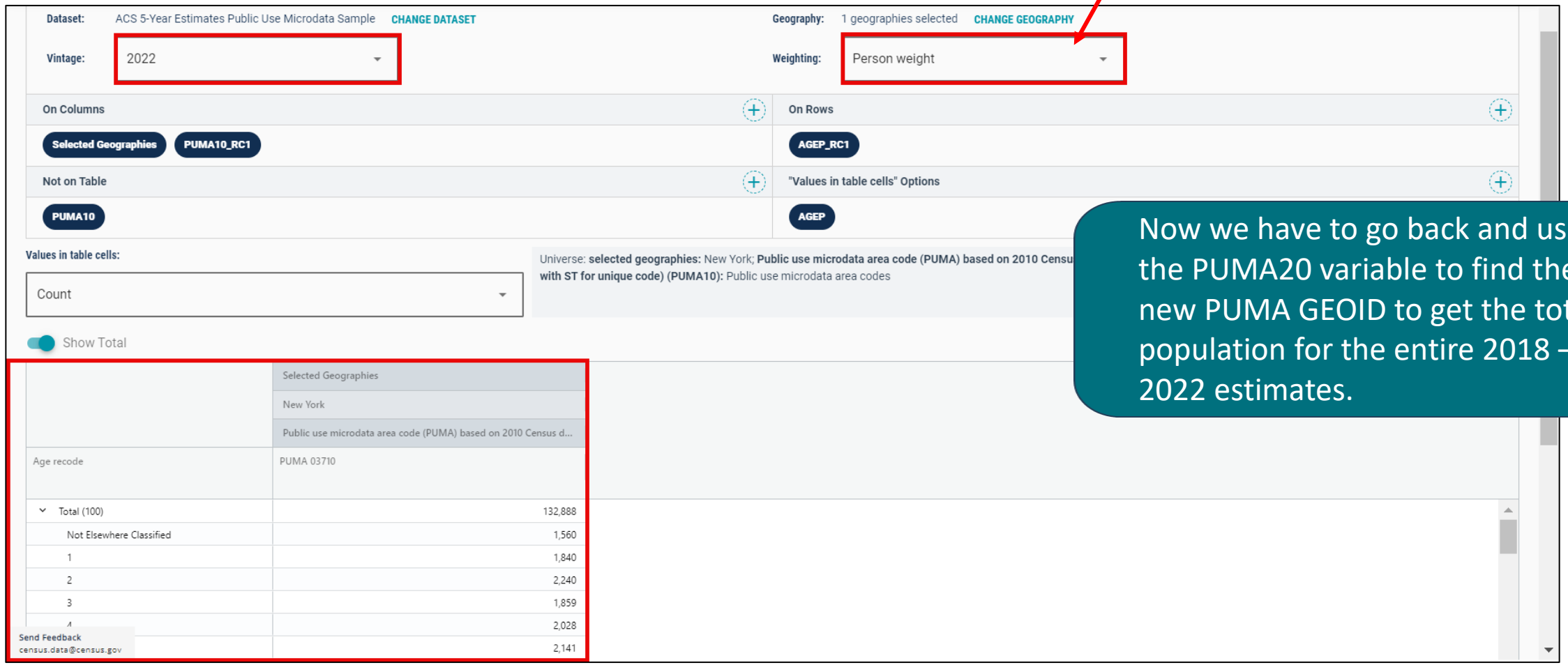


- Confirm Table Layout:
  - Confirm table layout and click **View Table** in the lower right



View Table

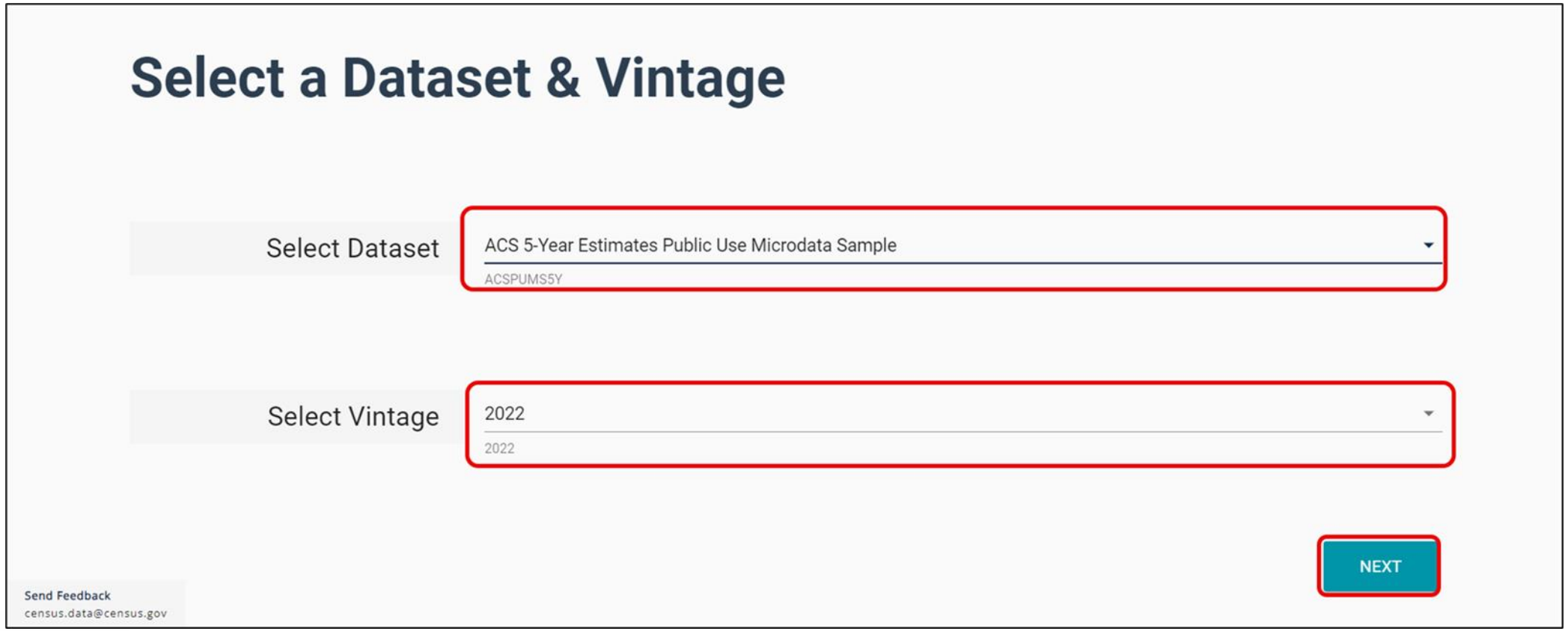
Note that the site automatically chooses a weight for you. You do have the option to change the weight if you want.



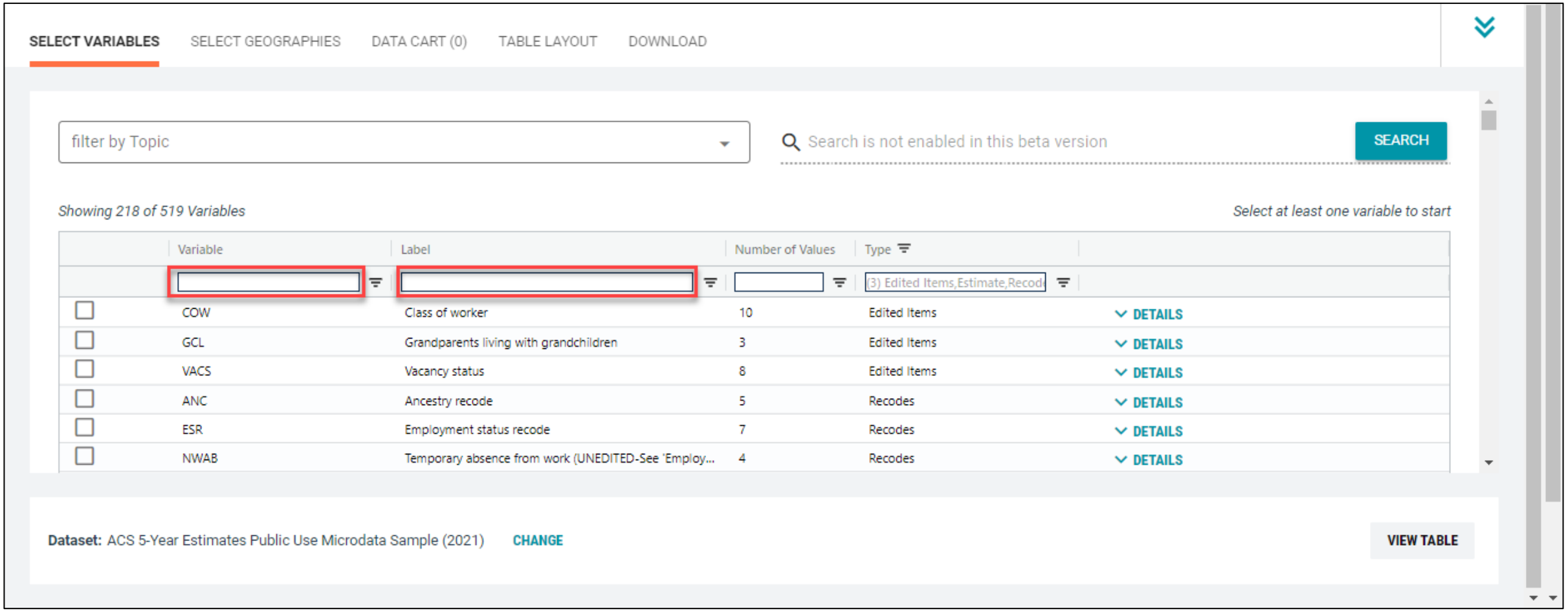
Now we have to go back and use the PUMA20 variable to find the new PUMA GEOID to get the total population for the entire 2018 – 2022 estimates.

United States® Census Bureau

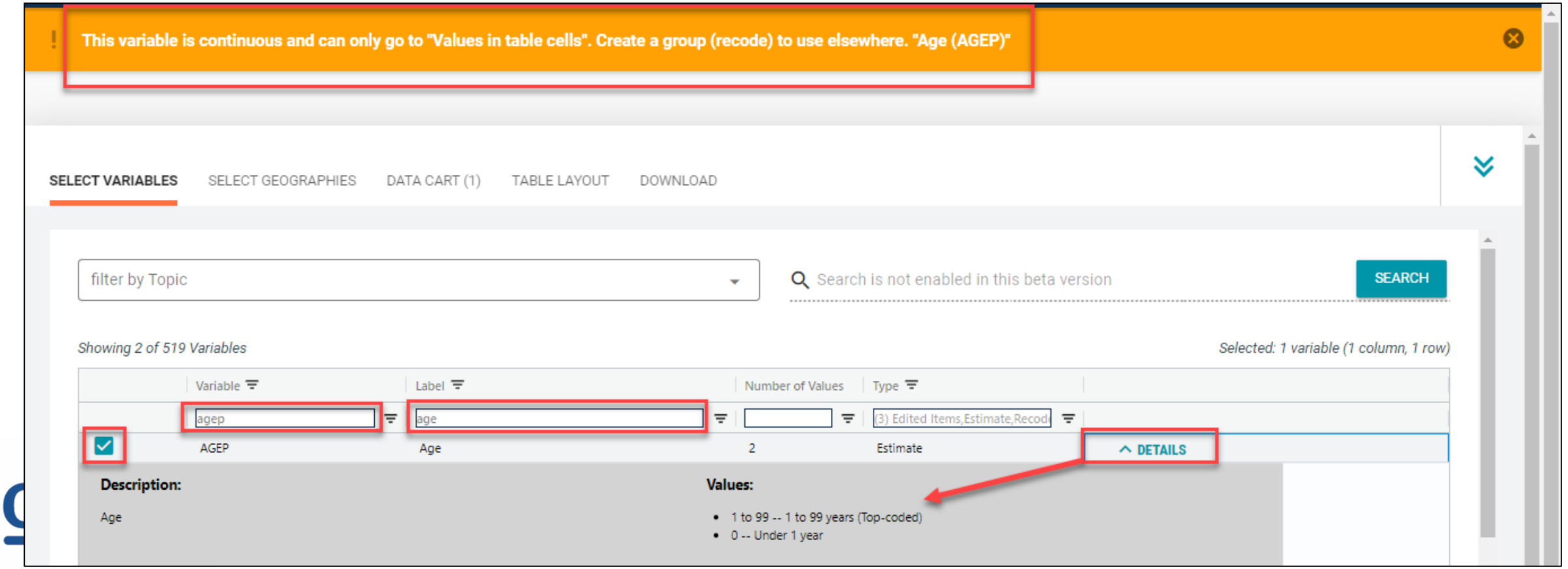
- Choose Dataset and Vintage:
  - Dataset – **ACS 5-Year Estimates – Public Use Microdata Sample**
  - Vintage – **2022**
  - Click **Next** in the lower right



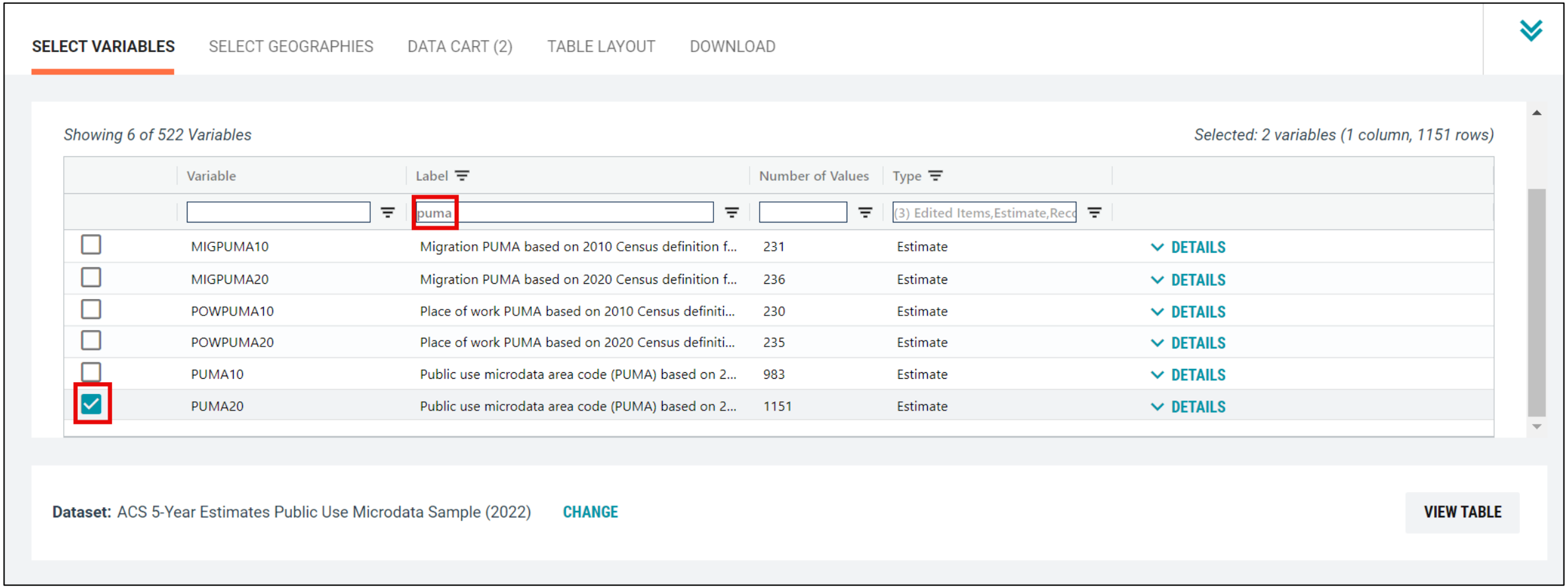
- **Search for Variables**: Use the search box below “Variable” or “Label” to find your variables of interest



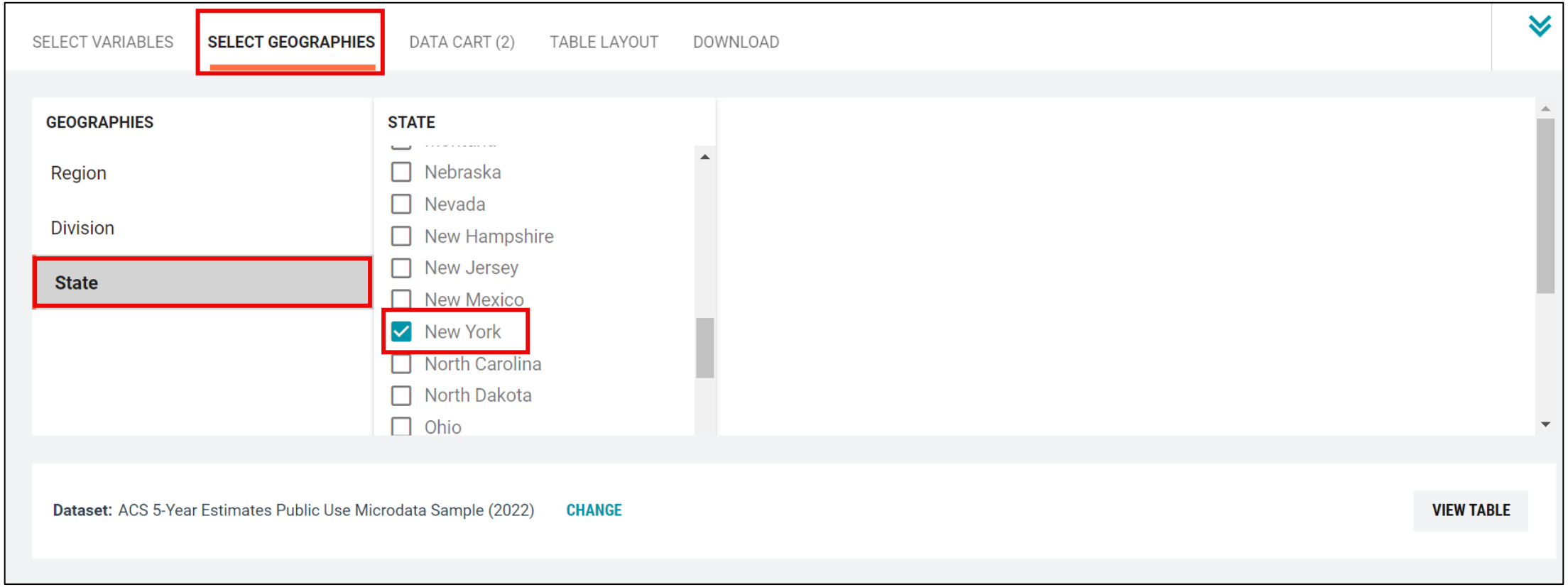
- Select variable for Age:
  - Type “AGEP” in the Variable search box or type “Age” in the label search box
  - Check the box to the left of AGEP to add the variable to your data cart
  - Notice the message at the top of the screen saying you will need to create your own categories (or recodes) for this variable if you want it shown in the table. (You will do this action in the Data Cart)



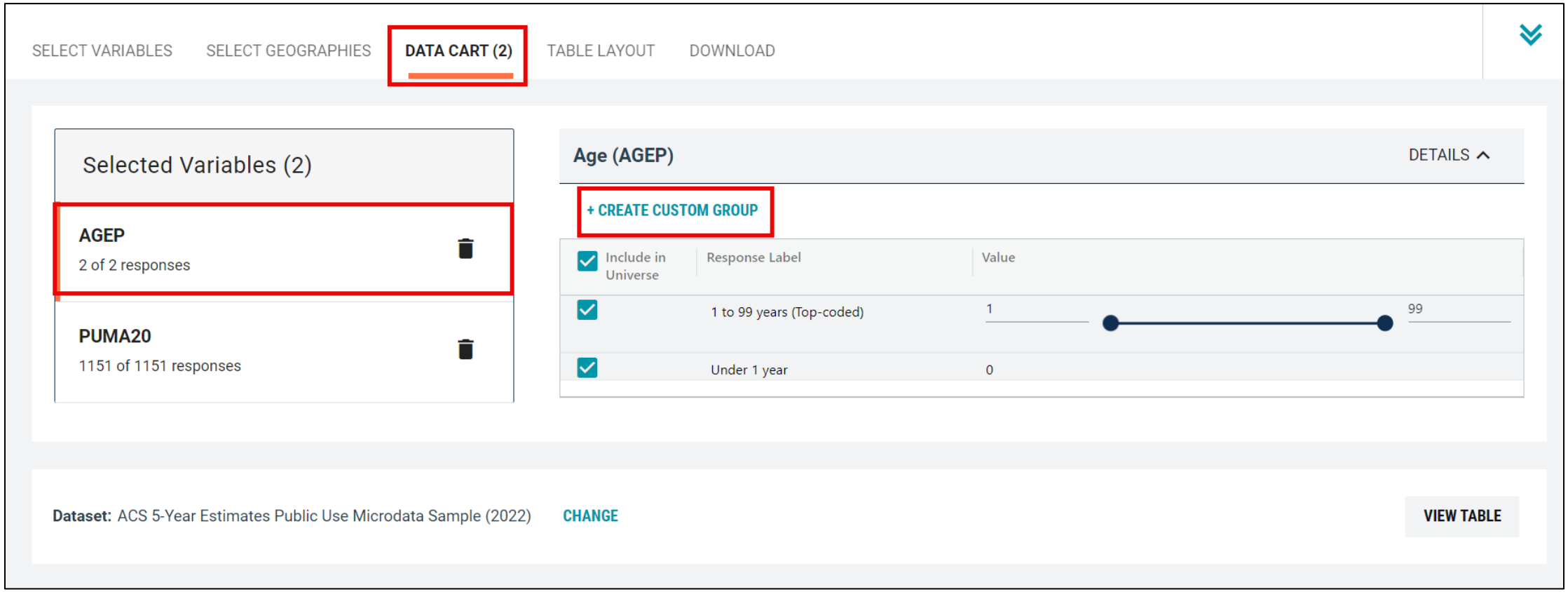
- Select variable for 2020 PUMAs:
  - Type ‘PUMA’ in the label search box
  - Check the box to the left of ‘PUMA20’ to add the variable to data cart



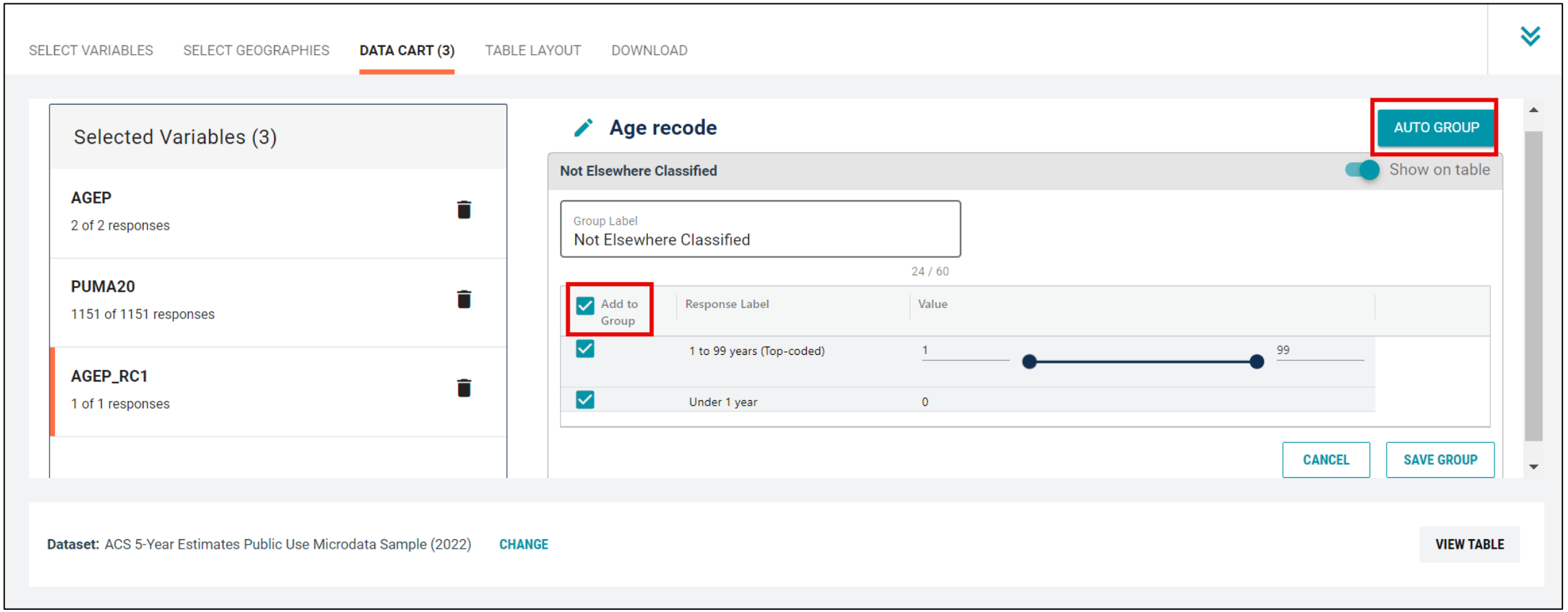
- **Select New York state geography.**
  - Click on the Select Geographies tab
  - Check the box to the left ‘New York’ to only pull up data for PUMAs from New York state



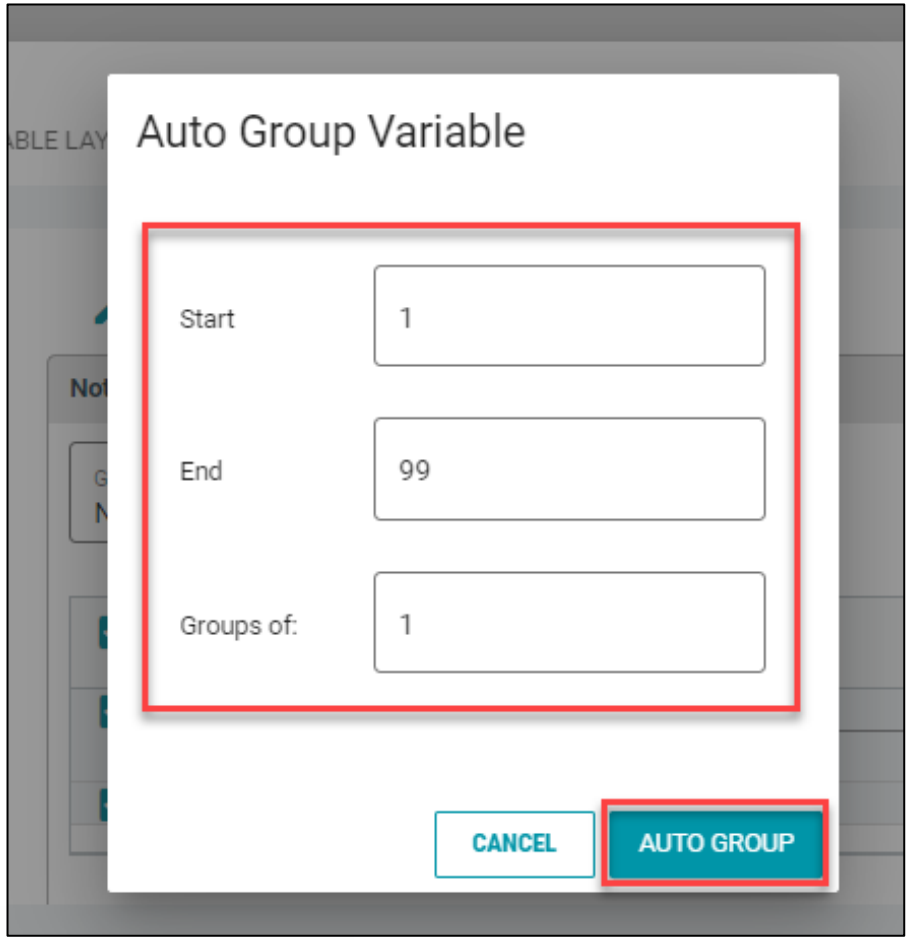
- Categorize (recode) your variable:
  - Move to the **Data Cart** tab
  - Click the **AGEP** variable on the left
  - Click **Create Custom Group** to begin specifying your age groups (e.g. single years of age)

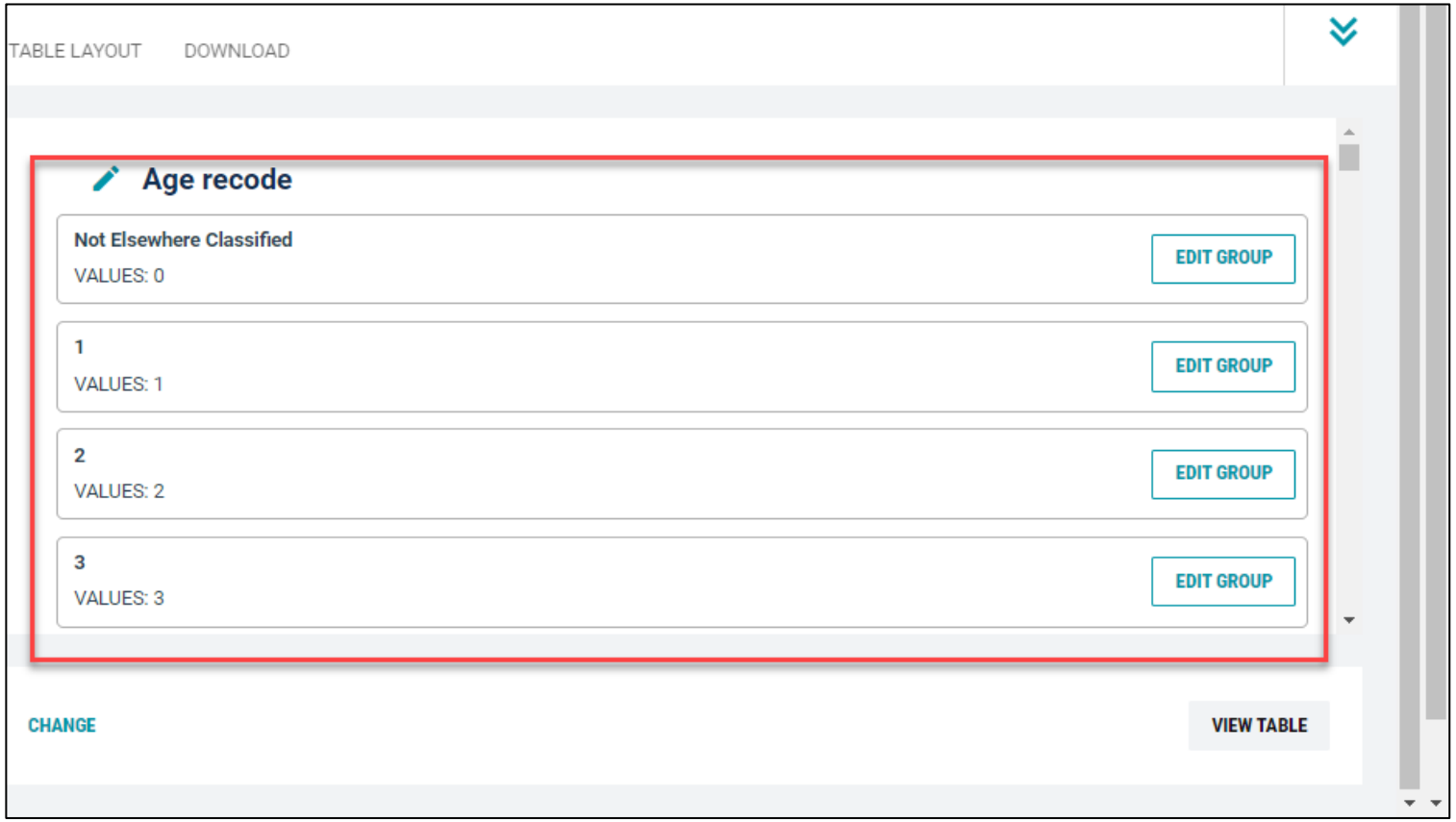


- Categorize (recode) your variable:
  - Check the box next to Add to Group to add both categories to the recode
  - Click on **Auto Group**

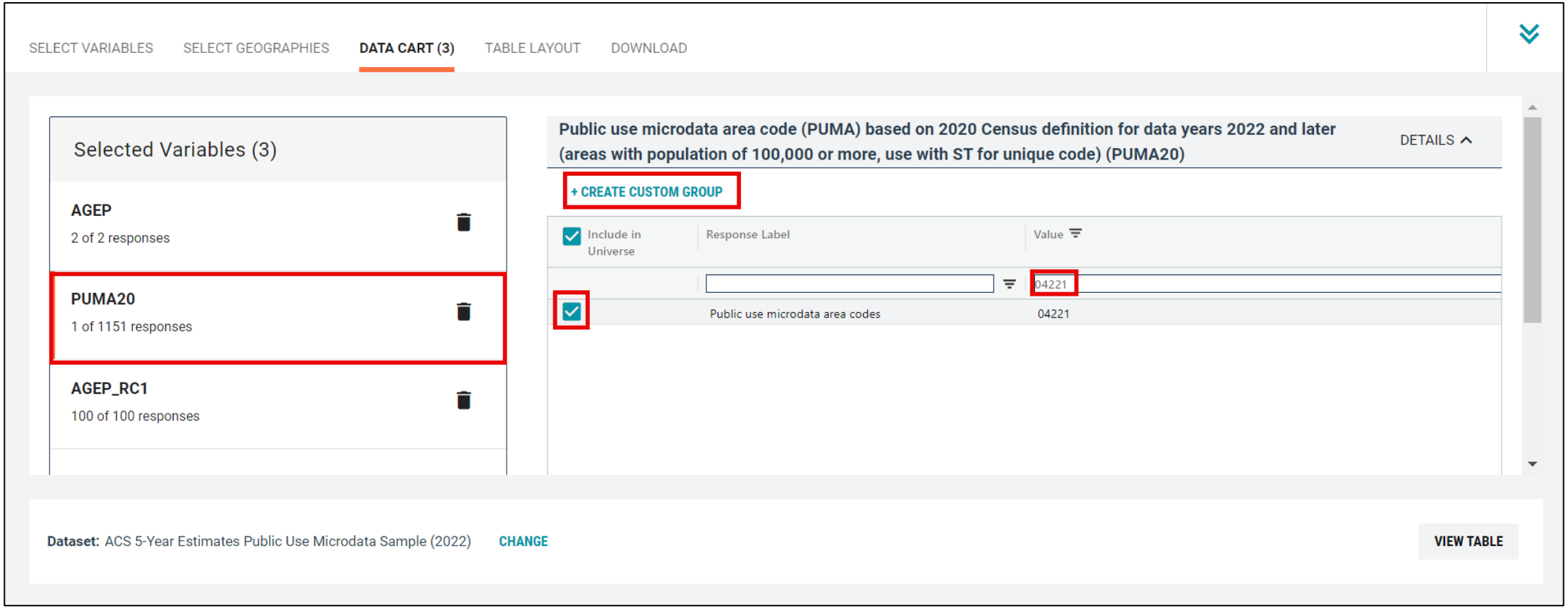


- **Categorize (recode) your variable**:
  - Confirm that the Start value is ‘1’, the End value is ‘99’, and the Groups of value is ‘1’
  - Click **Auto Group**. This will automatically create each year of age as its own group.

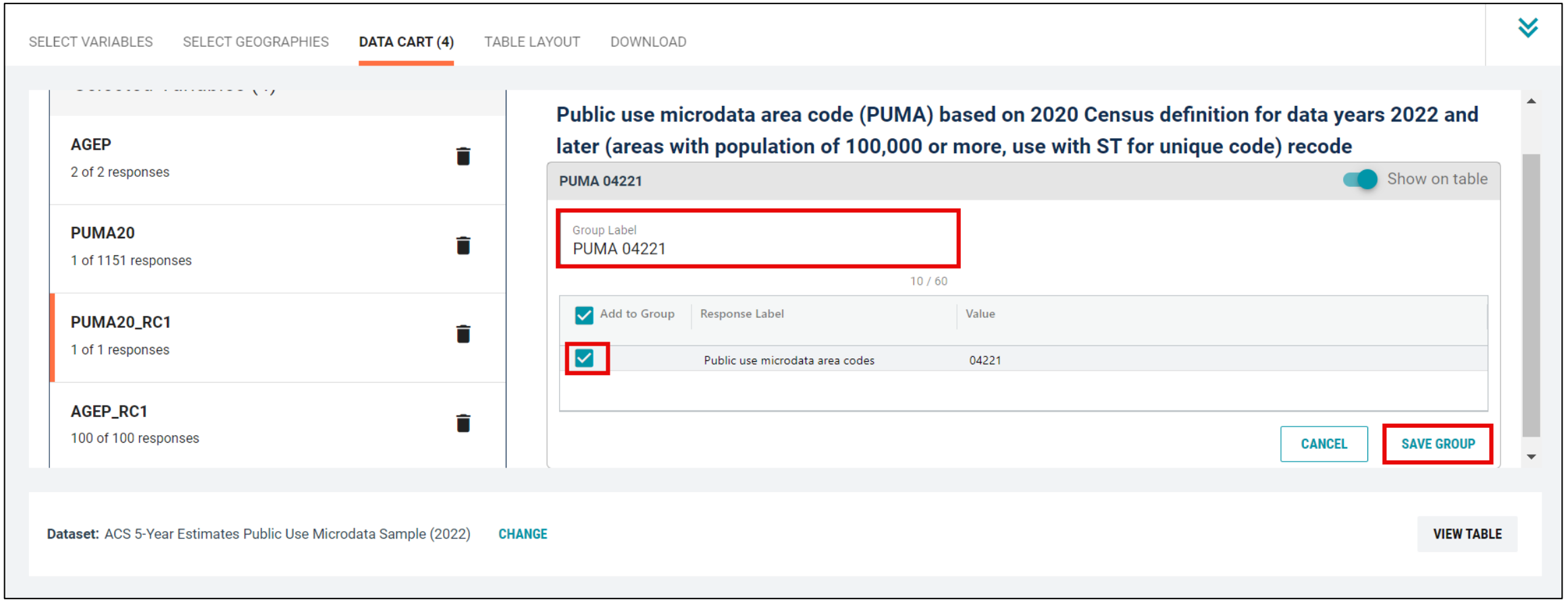




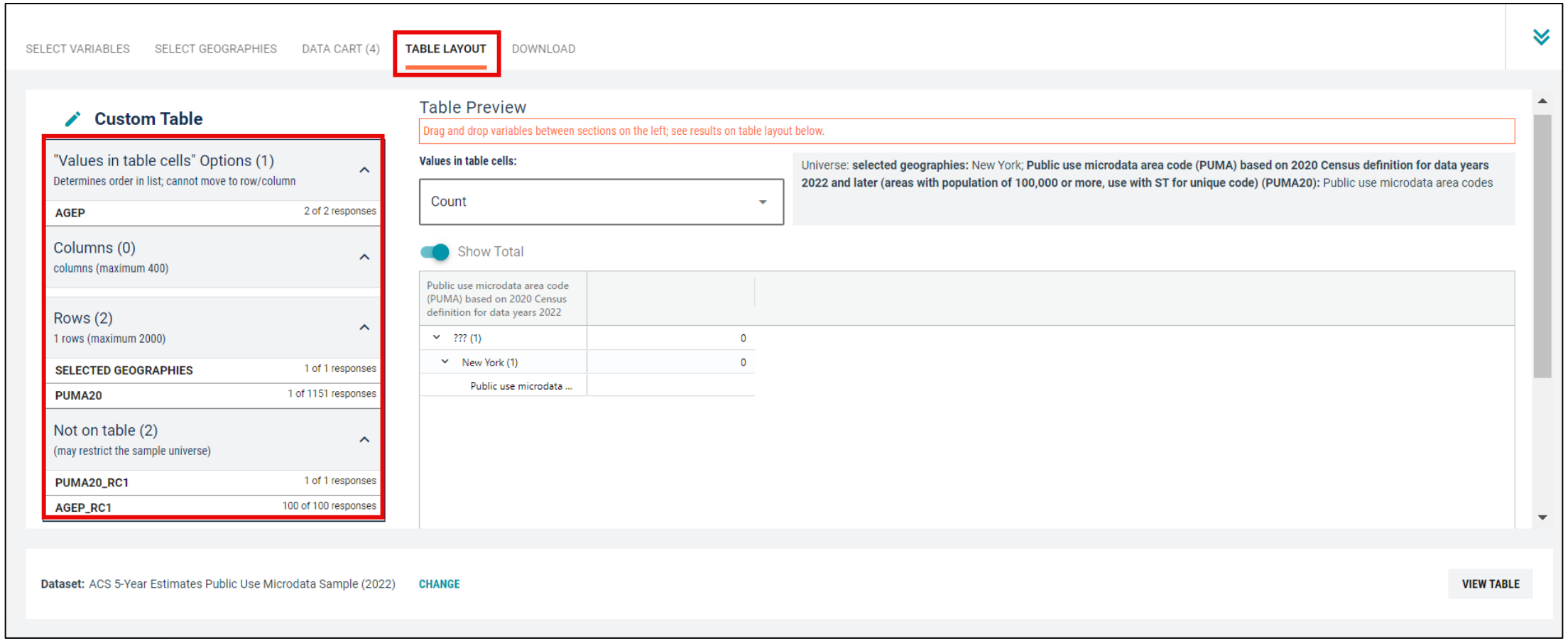
- Create recode to name PUMA20 variable:
  - **Select PUMA20** and click on the '**Include in Universe**' checkbox to uncheck all selected PUMAs
  - Reselect **04221** or use the Value search box to search for your desired PUMAs
  - Click on the **Create Custom Group** button to name your PUMA



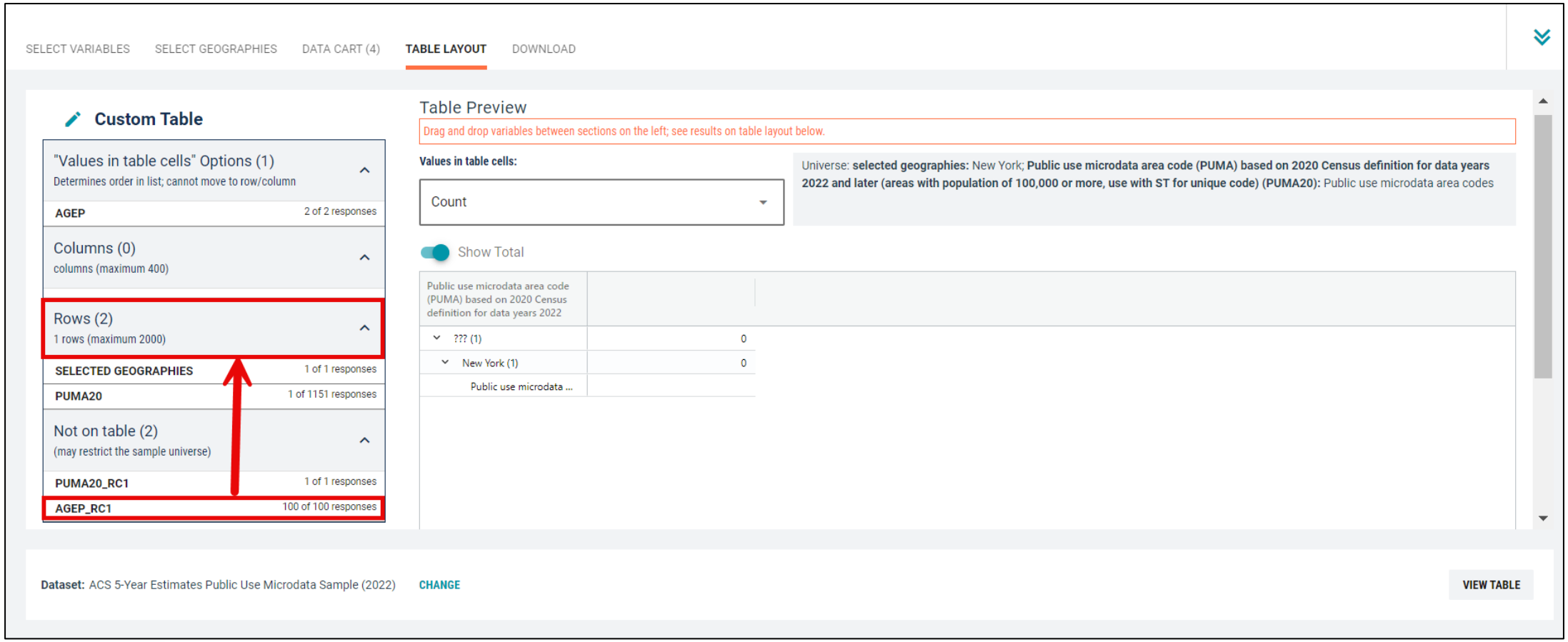
- Create recode to name PUMA20 variable:
  - Use the **Group Label box to type in 'PUMA 04221'** and **select your PUMA** by clicking on the checkbox
  - Click the **Save Group** button



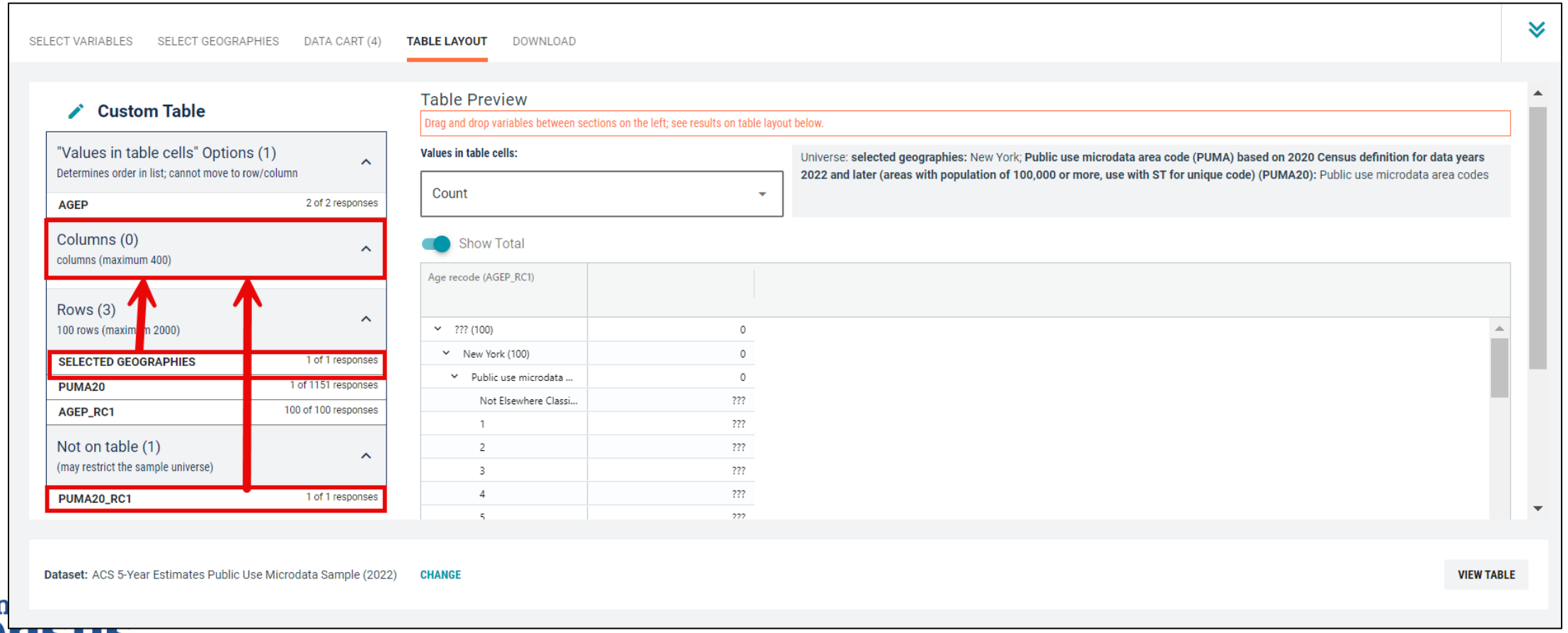
- View variable placement in the default table layout:
  - Move to the **Table Layout** tab
  - **Columns/Rows – Variables will be shown in the table.** By default, the table puts the 2020 PUMAs in the rows



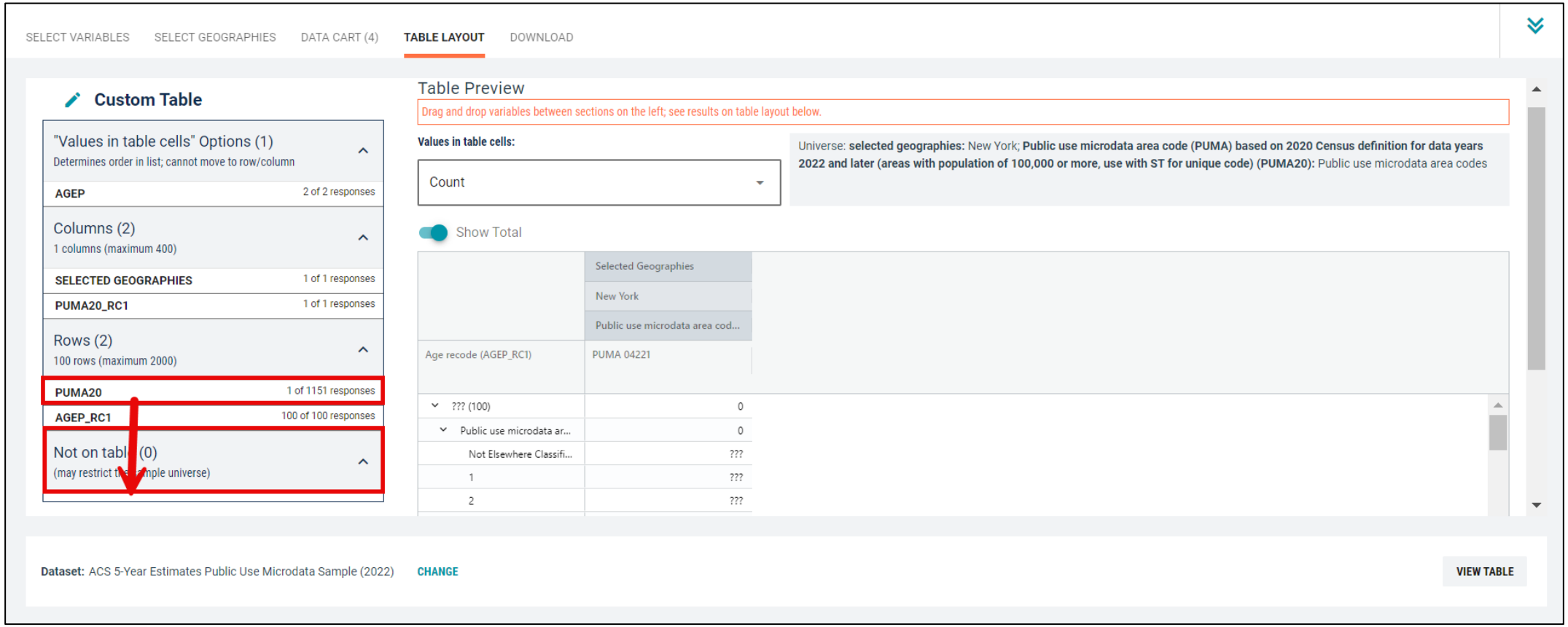
- Edit Table Layout:
  - Move Age Recode to Rows:
    - **Click, hold and drag AGEP_RC1 on the left side of the page up to the rows heading.** This will give you a table layout that includes the age categories that were created as the rows.



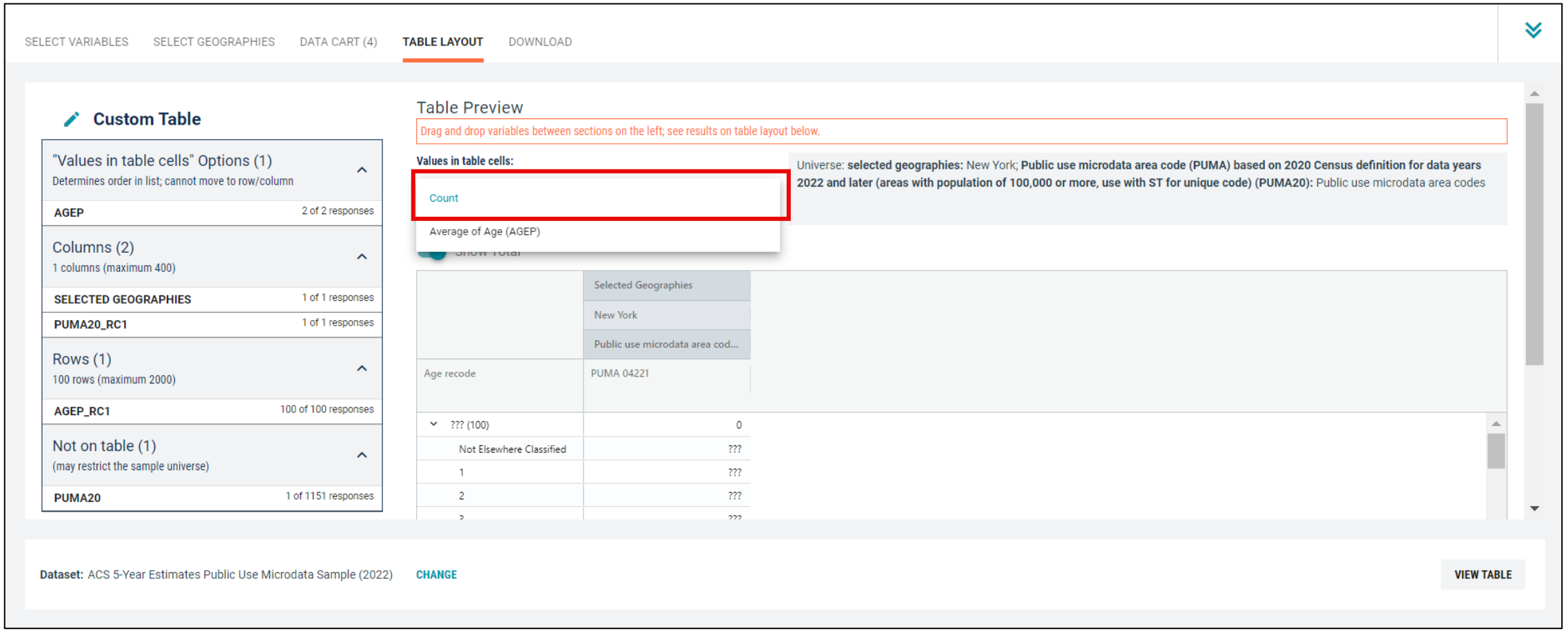
- Edit Table Layout:
  - Move Selected Geographies and PUMA recode to Columns:
    - Click, hold and drag Selected Geographies and PUMA20_RC1 on the left side of the page up to the **columns heading.** This will give you a table layout that includes the selected PUMA 04221 from New York as the columns.



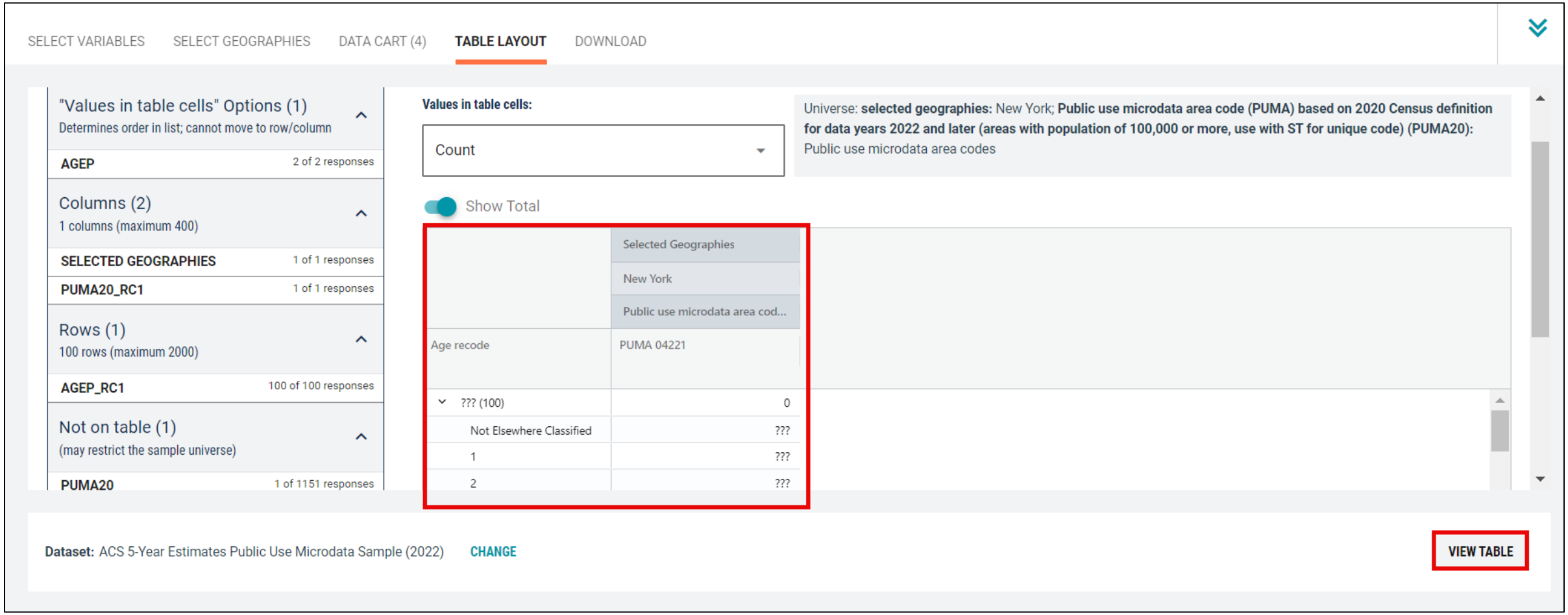
- Edit Table Layout:
  - Move original PUMA20 variable to Not on table section:
    - **Click, hold and drag PUMA20 to the Not on table section.** This will give you a table layout that includes the selected PUMA from New York as the columns.



- Choose type of values in table cells
  - Change the “Value in table cells” option from “Average of Age” to **Count.** This will give you data for the total number of people within the requested categories.

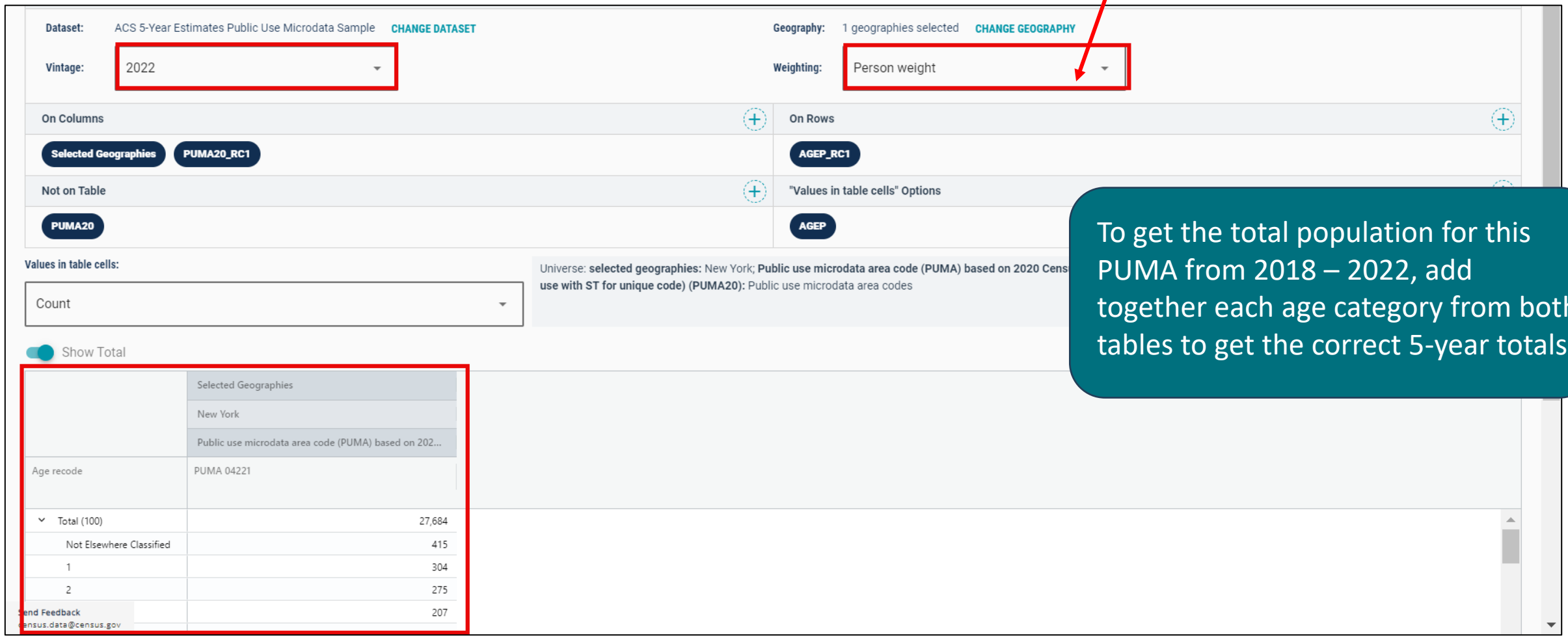


- Confirm Table Layout:
  - Confirm table layout and click **View Table** in the lower right



View Table

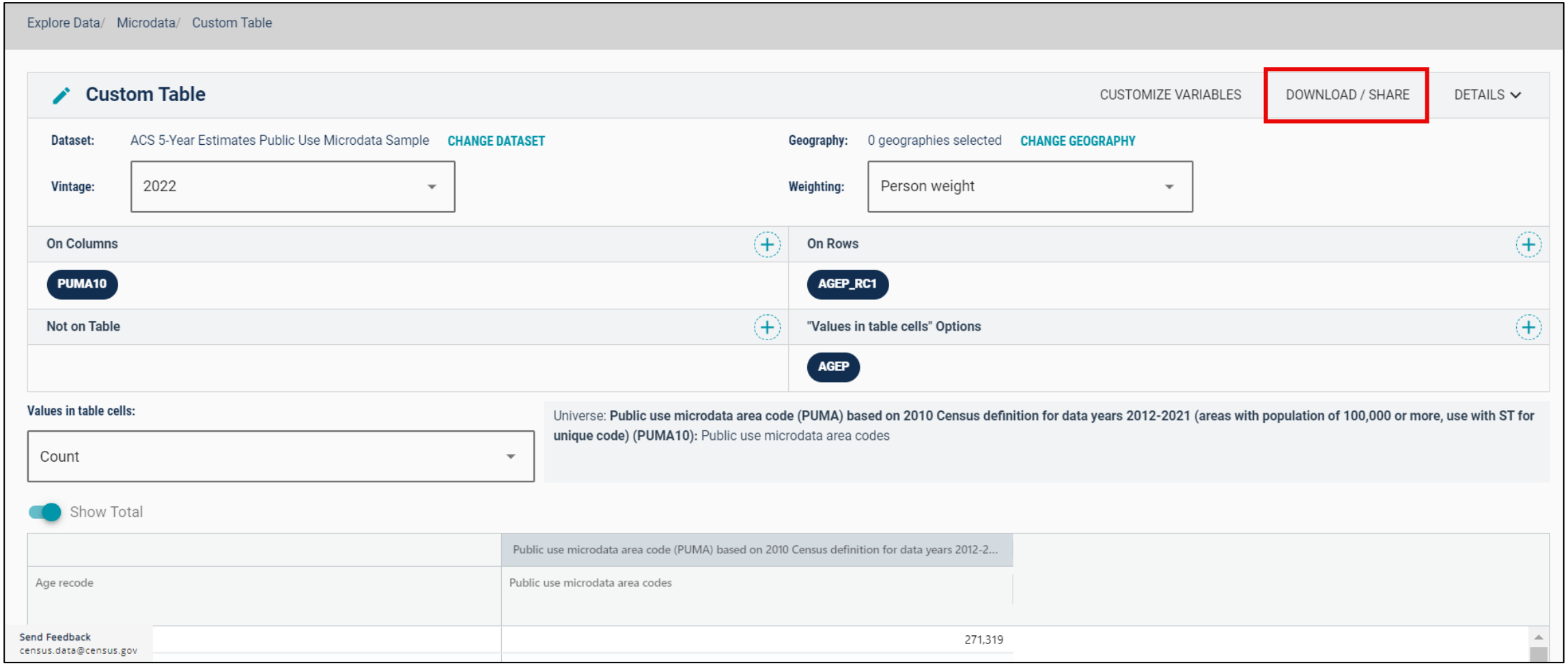
Note that the site automatically chooses a weight for you. You do have the option to change the weight if you want.



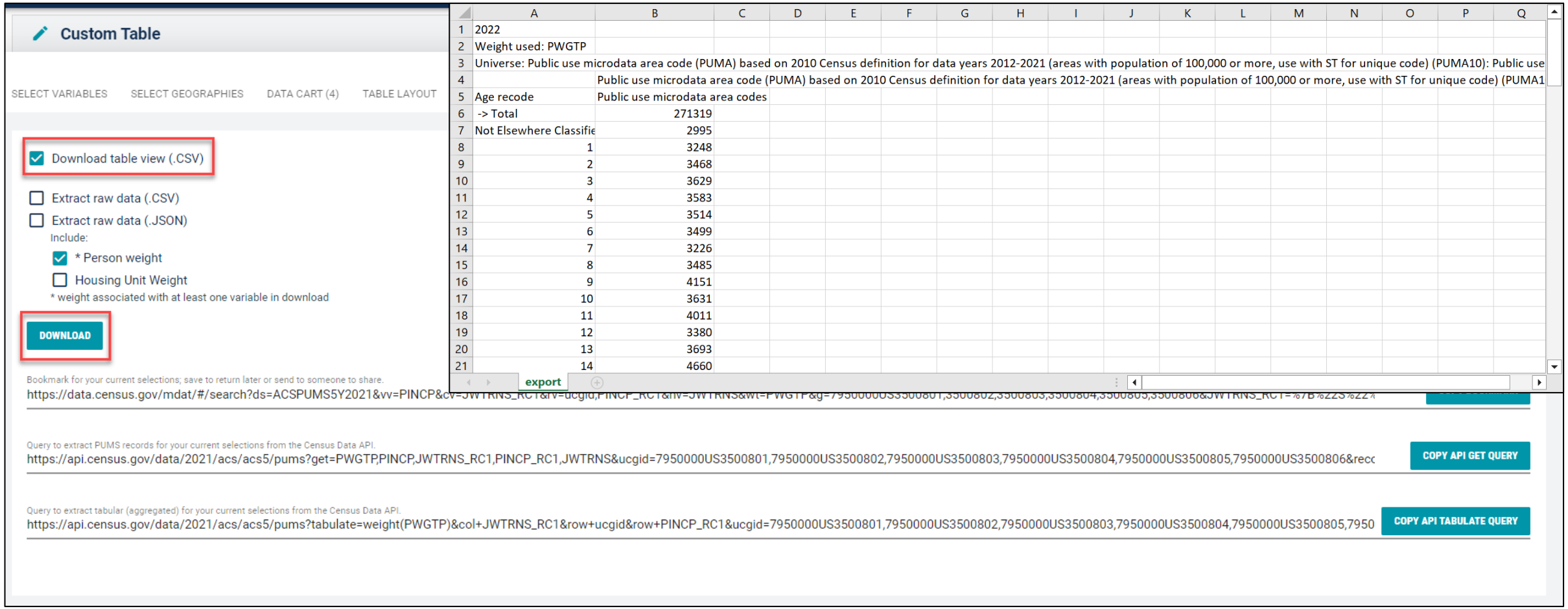
To get the total population for this PUMA from 2018 – 2022, add together each age category from both tables to get the correct 5-year totals.

United States® Census Bureau

- Download:
  - Click **Download/Share** at the top of the table

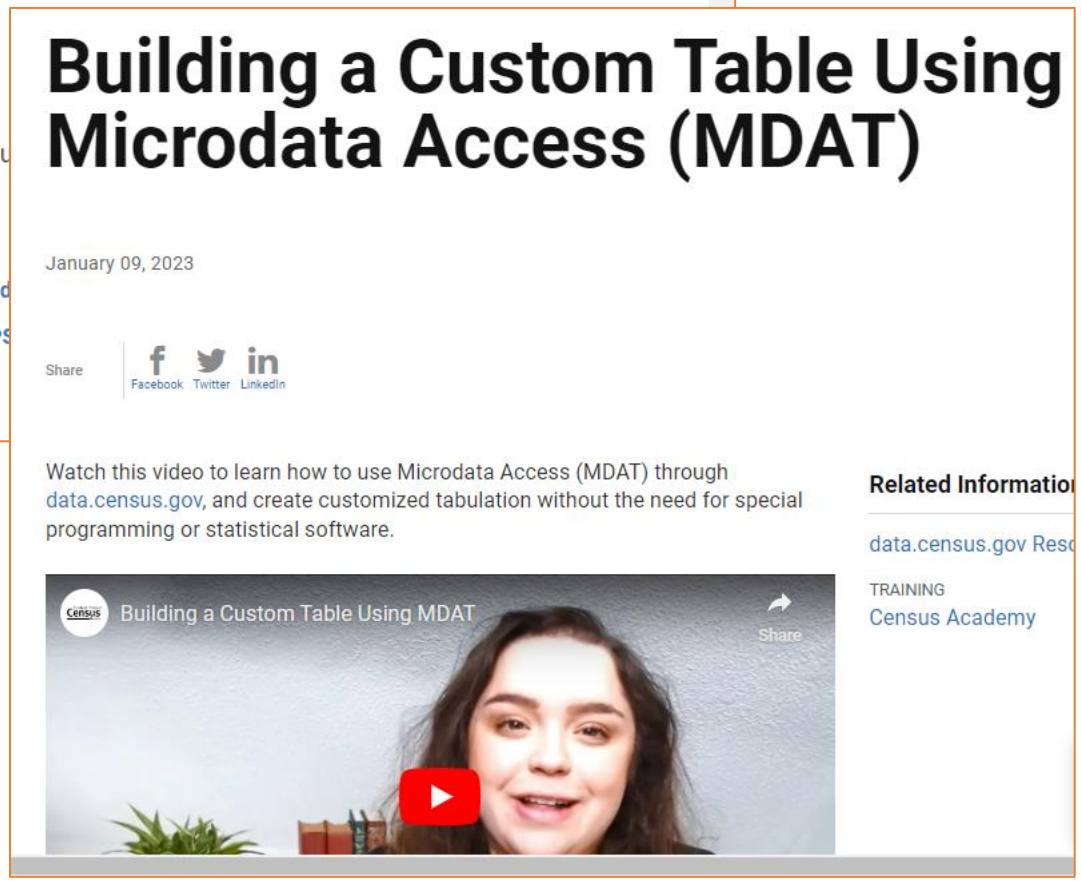


- Download:
  - Select **Download table view (.CSV),** then click **DOWNLOAD**
  - Click on **export.csv** to view your downloaded table



	A	B
1	2022	
2	Weight used: PWGTP	
3	Universe: Public use microdata area code (PUMA) based on 2010 Census definition for data years 2012-2021 (areas with population of 100,000 or more, use with ST for unique code) (PUMA10): Public use	
4		Public use microdata area code (PUMA) based on 2010 Census definition for data years 2012-2021 (areas with population of 100,000 or more, use with ST for unique code) (PUMA1
5	Age recode	Public use microdata area codes
6	-> Total	271319
7	Not Elsewhere Classifie	2995
8	1	3248
9	2	3468
10	3	3629
11	4	3583
12	5	3514
13	6	3499
14	7	3226
15	8	3485
16	9	4151
17	10	3631
18	11	4011
19	12	3380
20	13	3693
21	14	4660

Guidance for Data Users



MDAT Resources Page:

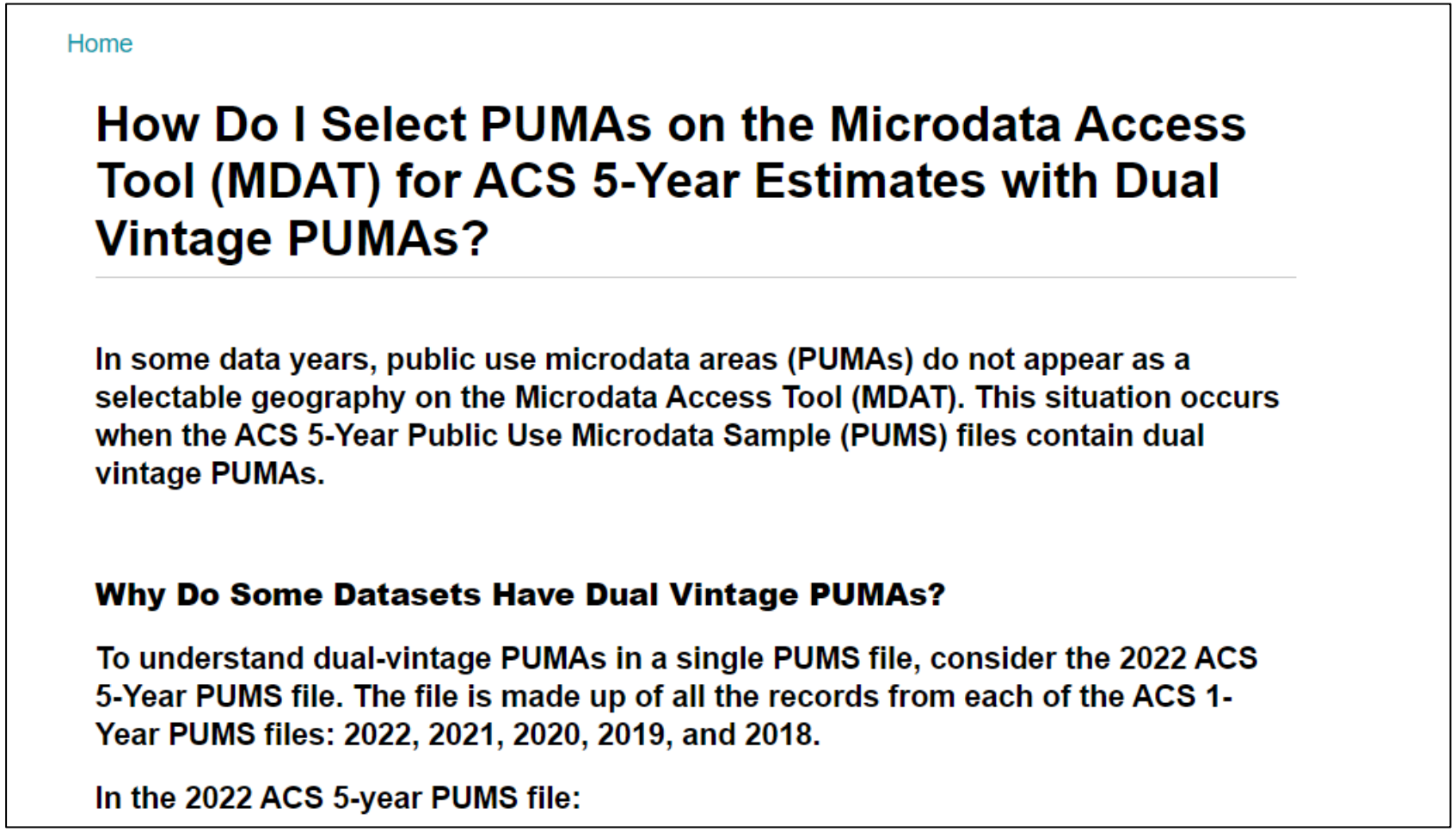
https://www.census.gov/data/what-is-data-census-gov/guidance-for-data-users/how-to-materials-for-using-the-microdata-access.html

United States® Census Bureau

PUMA FAQ Page

PUMA FAQ Page:

https://ask.census.gov/prweb/PRServletCustom?pyActivity=pyMobileSnapStart&ArticleID=KCP-18715



Home

How Do I Select PUMAs on the Microdata Access Tool (MDAT) for ACS 5-Year Estimates with Dual Vintage PUMAs?

In some data years, public use microdata areas (PUMAs) do not appear as a selectable geography on the Microdata Access Tool (MDAT). This situation occurs when the ACS 5-Year Public Use Microdata Sample (PUMS) files contain dual vintage PUMAs.

Why Do Some Datasets Have Dual Vintage PUMAs?

To understand dual-vintage PUMAs in a single PUMS file, consider the 2022 ACS 5-Year PUMS file. The file is made up of all the records from each of the ACS 1-Year PUMS files: 2022, 2021, 2020, 2019, and 2018.

In the 2022 ACS 5-year PUMS file:

United States® Census Bureau

2020 Public Use Microdata Areas (PUMA) Guidance Page

Includes:

- **Summary Guide**
- **Coding and Naming Guidelines**
- **2020 PUMA FAQs**

https://www.census.gov/programs-surveys/geography/guidance/geo-areas/pumas/2020pumas.html

// Census.gov / Guidance for Geography Users / About Geographic Areas / Public Use Microdata Areas (PUMAs) / 2020 Public Use Microdata Areas (PUMA) Program

2020 Public Use Microdata Areas (PUMA) Program

Share Facebook X (Twitter) LinkedIn

The Census Bureau partnered with State Data Centers (SDCs) from each state, the District of Columbia, and the Commonwealth of Puerto Rico to delineate PUMAs during the 2020 PUMA program. While the SDCs were the official program participants, the Census Bureau encouraged SDCs to involve other interested data users to ensure newly delineated PUMAs met the needs of a variety of data users.

Three training webinars, two in October and one in December, were conducted by the Census Bureau for participants in late 2021. Delineation occurred over a 90-day review period from November 2021 through January 2022, after the publication of population counts and census tracts from the 2020 Census. The final PUMAs and associated data will be available online for public use beginning in the summer of 2022.

The Census Bureau is no longer accepting submissions for the 2020 PUMA. The

Top

For questions regarding the 2020 PUMA program, contact the Census Bureau at geo.puma@census.gov.

United States® Census Bureau

Email Updates

Get data.census.gov updates delivered to your inbox!

Sign up for email updates: https://public.govdelivery.com/accounts/USCENSUS/signup/15450

United States® Census Bureau

Measuring America's People, Places, and Economy

Sign up to stay up to date on the latest Census Bureau data releases, new data visualizations, alerts for developers, and new tools for data users.

Email *

census.data@census.gov

Select One or More: *

- [ ] COVID-19 Data Hub
- [ ] Weekly Pulse Newsletter
- [ ] Experimental Data Products
- [ ] Data Viz Newsletter
- [ ] Census Business Builder
- [x] data.census.gov Updates
- [ ] Census Data API
- [ ] Developer Newsletter

Select your state: *

Maryland

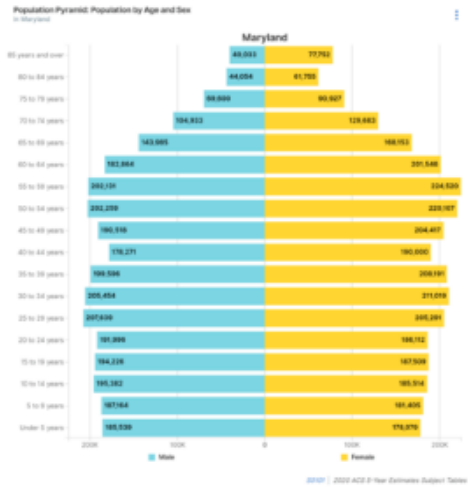
By checking this box, you consent to our data privacy policy. *

Sign Up

Data.census.gov Newsletter – January 2023

United States® Census Bureau

Data.census.gov Newsletter – January 2023



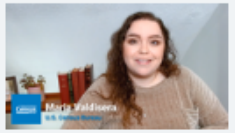
Learn about the latest system updates, data releases, and educational opportunities for data.census.gov.

Latest System Updates

In December, we released new updates to improve your experience on data.census.gov

Featured Video Tutorial:

Explanation of Mapping Buttons on data.census.gov



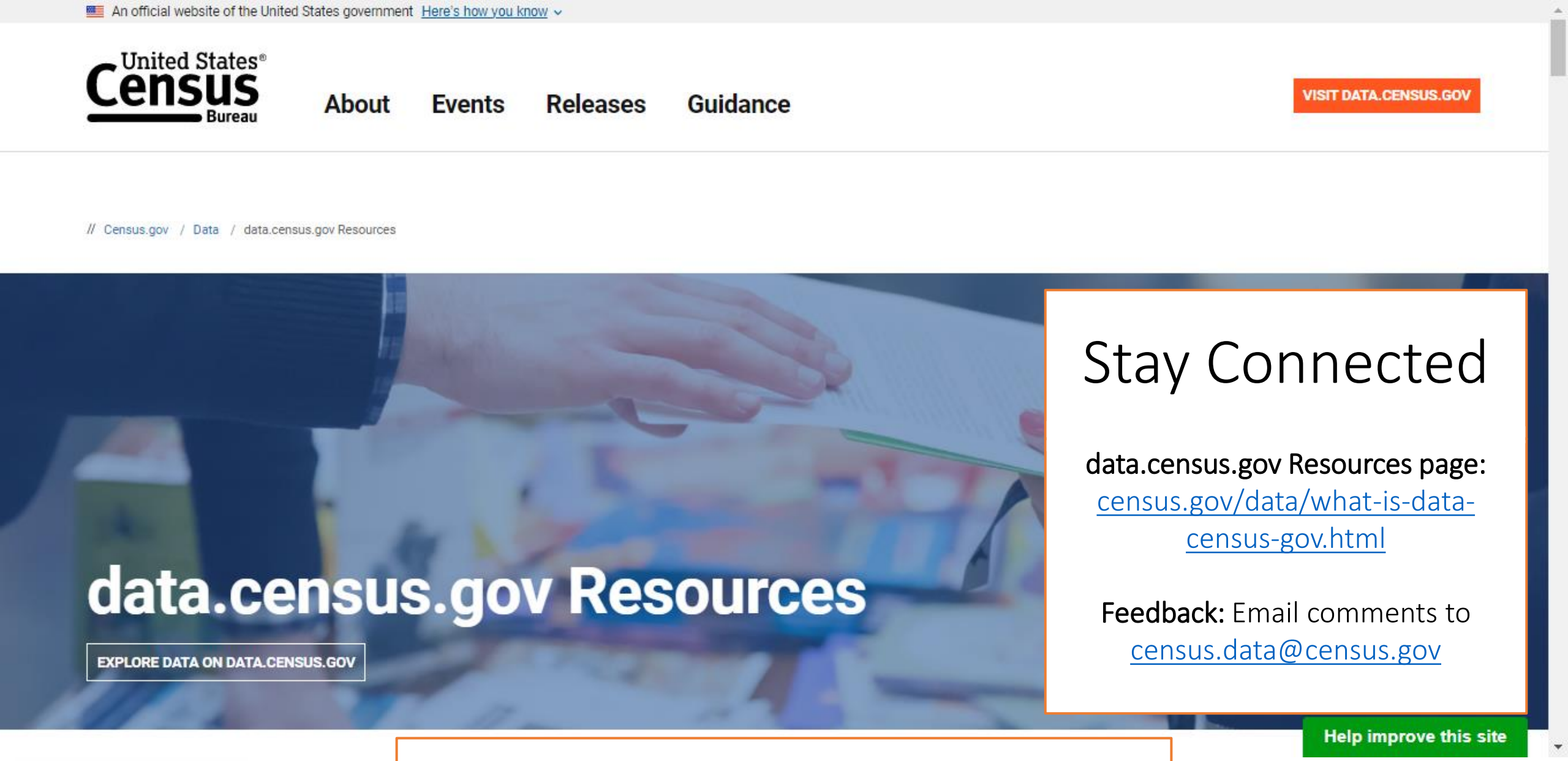
More Videos

Upcoming Webinar:

data.census.gov News and Updates: January 2023

Join us for this webinar where we will cover updates on data.census.gov for the months of October, November, and December. Updates include simplified URLs, new accessibility

United States® Census Bureau



United States Census Bureau