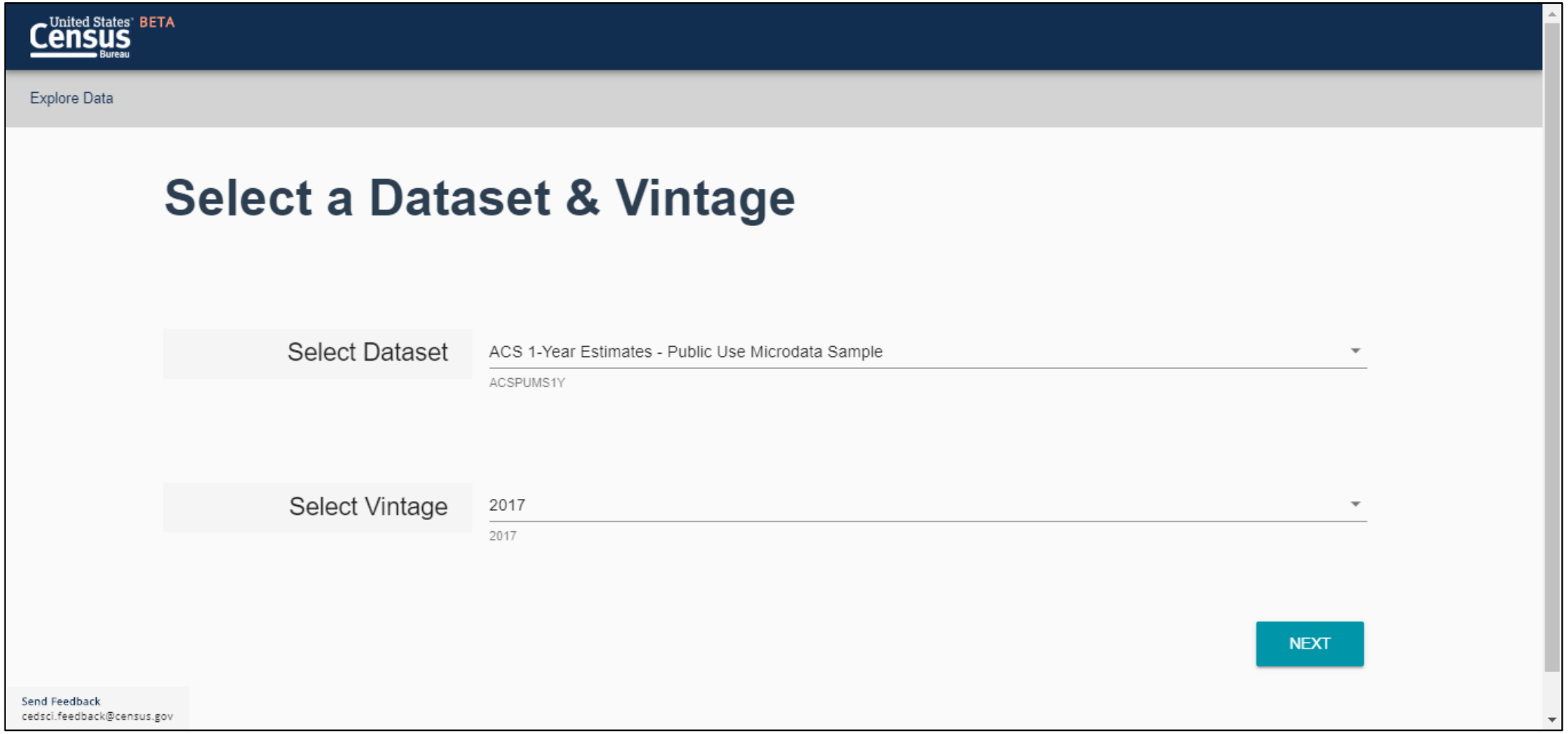
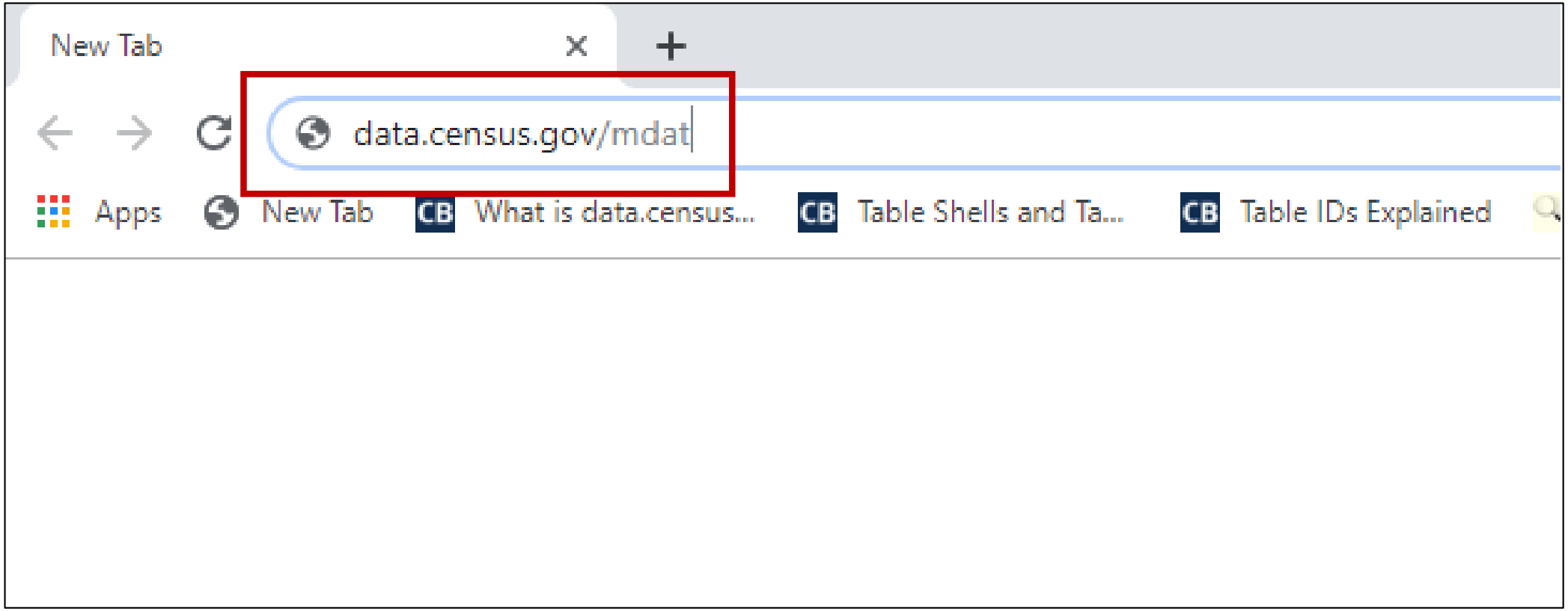
Using Microdata Access

With ACS 1-Year Estimates – Public Use Microdata Sample

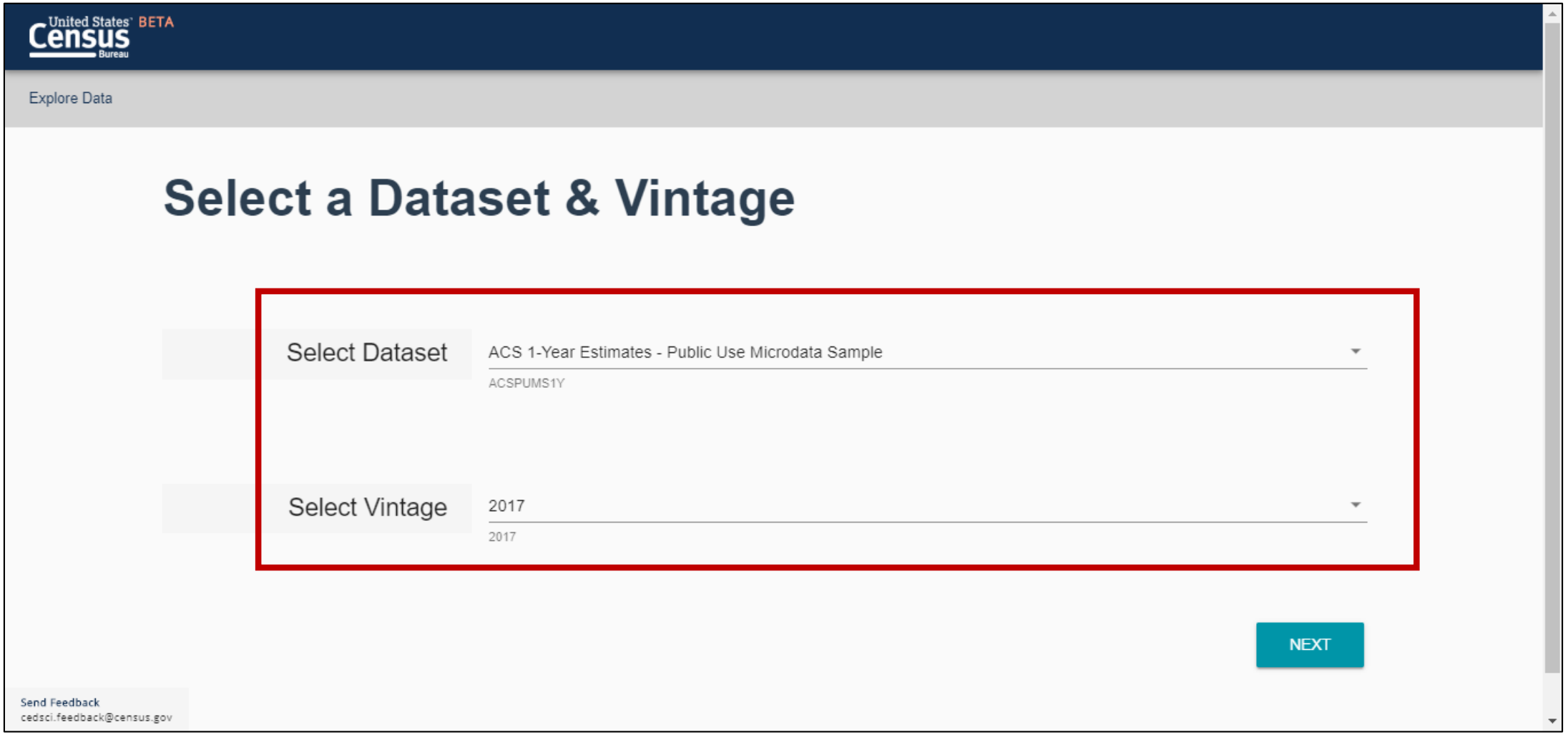
data.census.gov/mdat





To use Microdata Access, go to data.census.gov/mdat.

Microdata Access should work properly using any web browser.



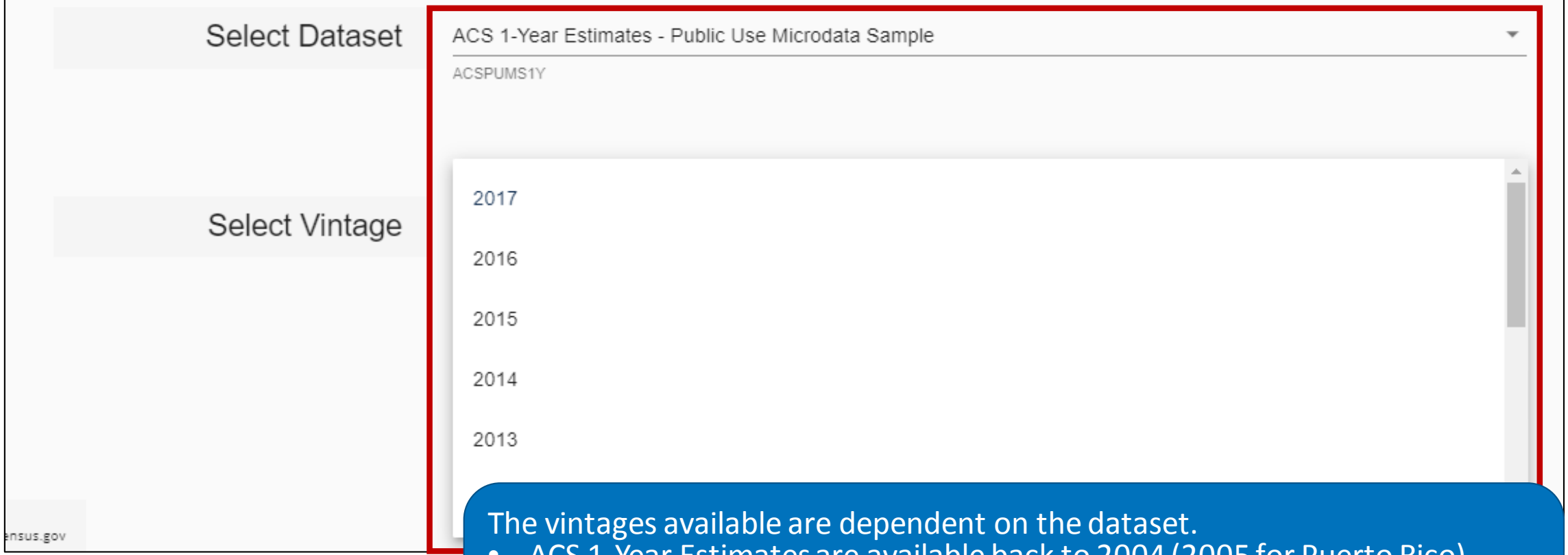
The landing page allows you to select your dataset and vintage.

Select a Dataset & Vintage

Currently, the following datasets can be found in Microdata Access:

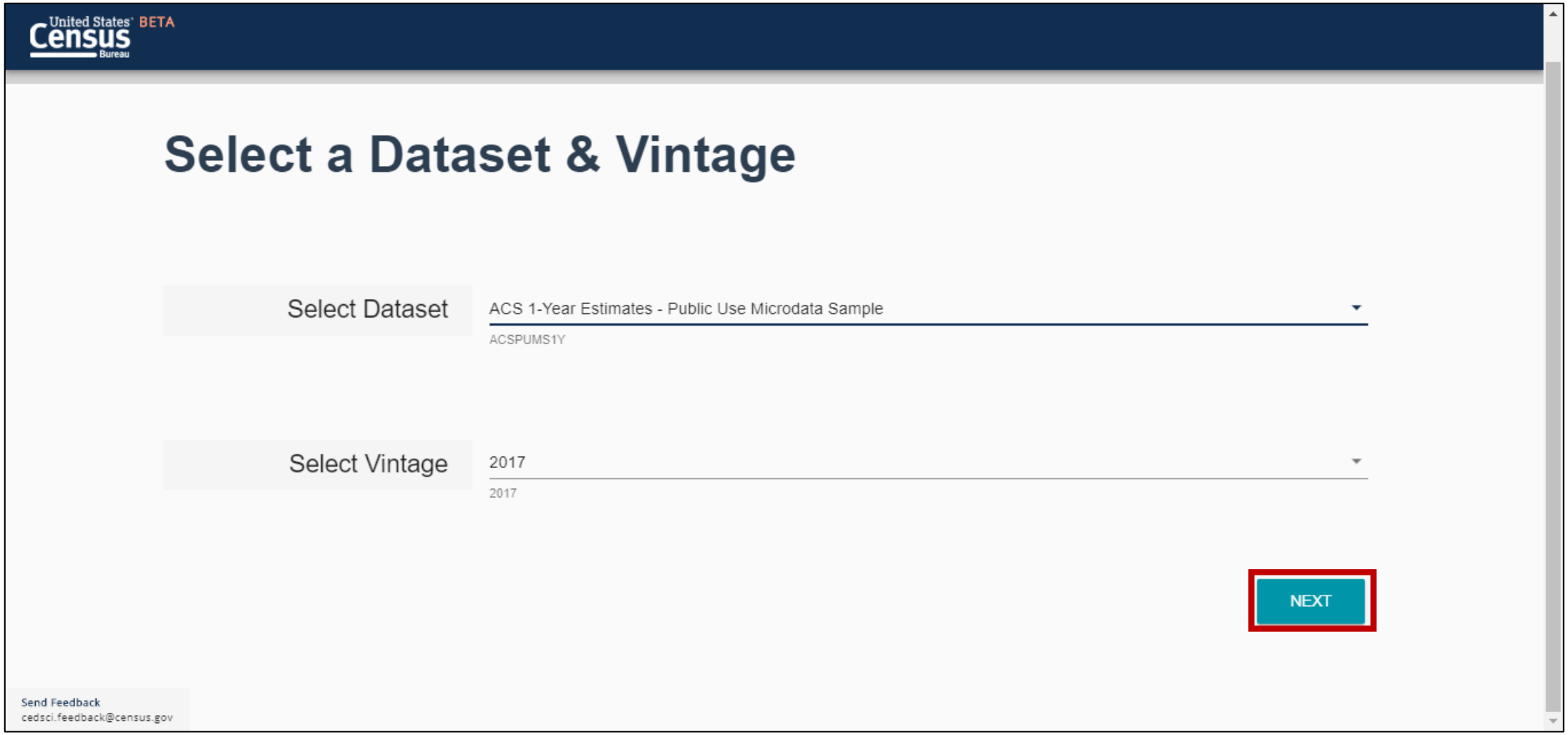
- ACS 1-Year Estimates – Public Use Microdata Sample
- ACS 1-Year Estimates – Puerto Rico Public Use Microdata Sample
- ACS 5-Year Estimates – Public Use Microdata Sample
- ACS 5-Year Estimates – Puerto Rico Public Use Microdata Sample
- CPS Annual Social and Economic (March) Supplement
- CPS Basic Monthly

Select a Dataset & Vintage

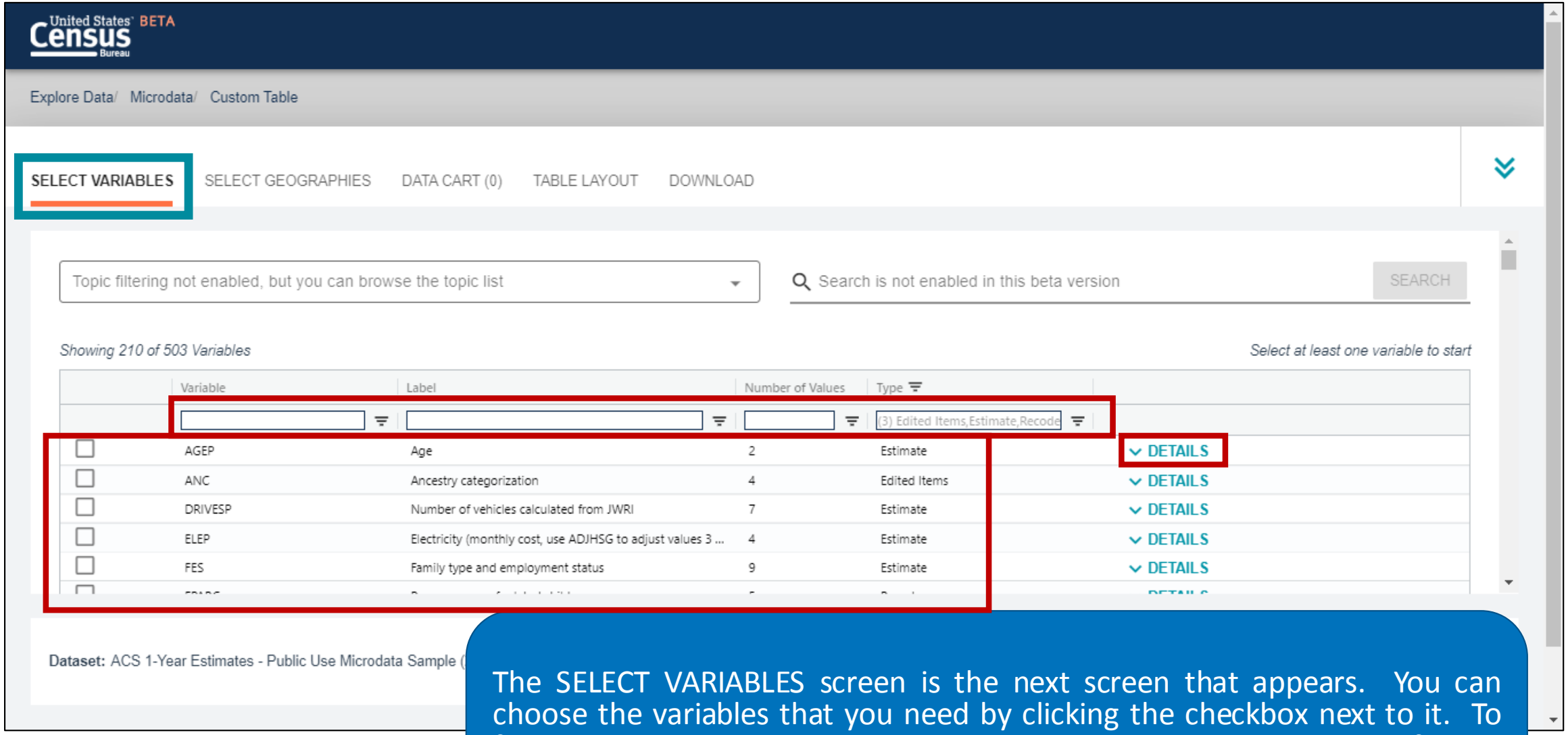


The vintages available are dependent on the dataset.

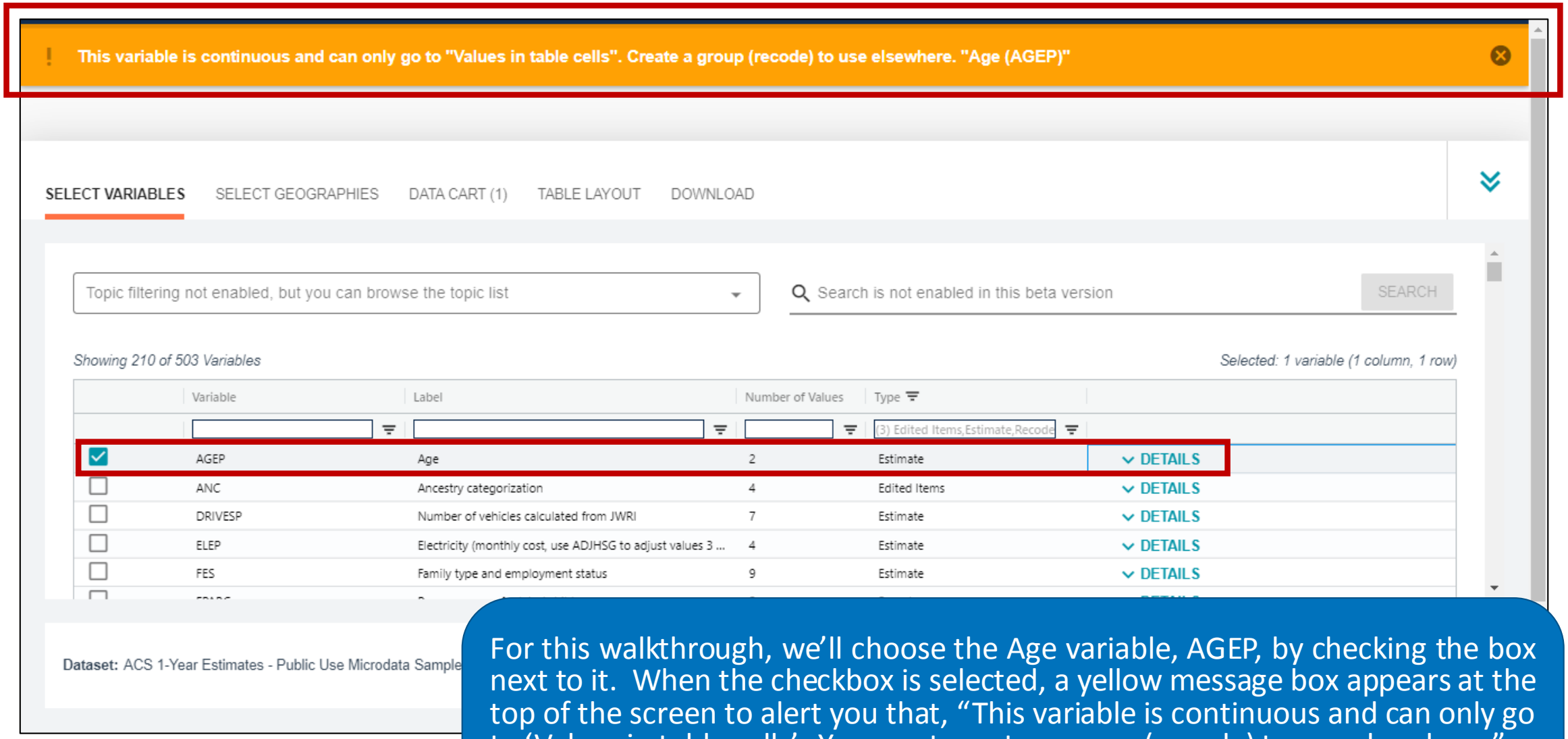
- ACS 1-Year Estimates are available back to 2004 (2005 for Puerto Rico)
- ACS 5-Year Estimates are available back to 2009
- CPS ASEC March Supplement datasets are available back to March 2014
- CPS Basic Monthly datasets are available back to January 1994



For this walkthrough, we'll use the 2017 ACS 1-Year Estimates. Once these are selected, hit the NEXT button found in the lower right of the screen.

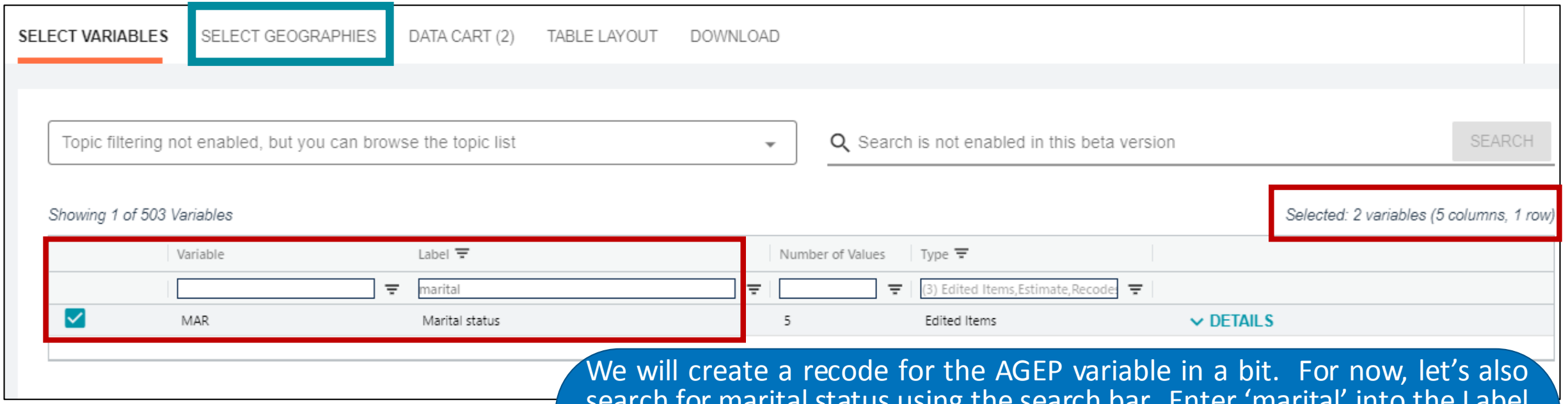


The SELECT VARIABLES screen is the next screen that appears. You can choose the variables that you need by clicking the checkbox next to it. To find the variables more quickly, you can use the search bar to search for the desired variable. You can view specific details about the variable by clicking on the DETAILS dropdown.



For this walkthrough, we'll choose the Age variable, AGEP, by checking the box next to it. When the checkbox is selected, a yellow message box appears at the top of the screen to alert you that, "This variable is continuous and can only go to 'Values in table cells'. You must create a group (recode) to use elsewhere."

This means that a recode must be created in order to use this variable in more than just the 'Values in table cells' option (more on creating recodes later).

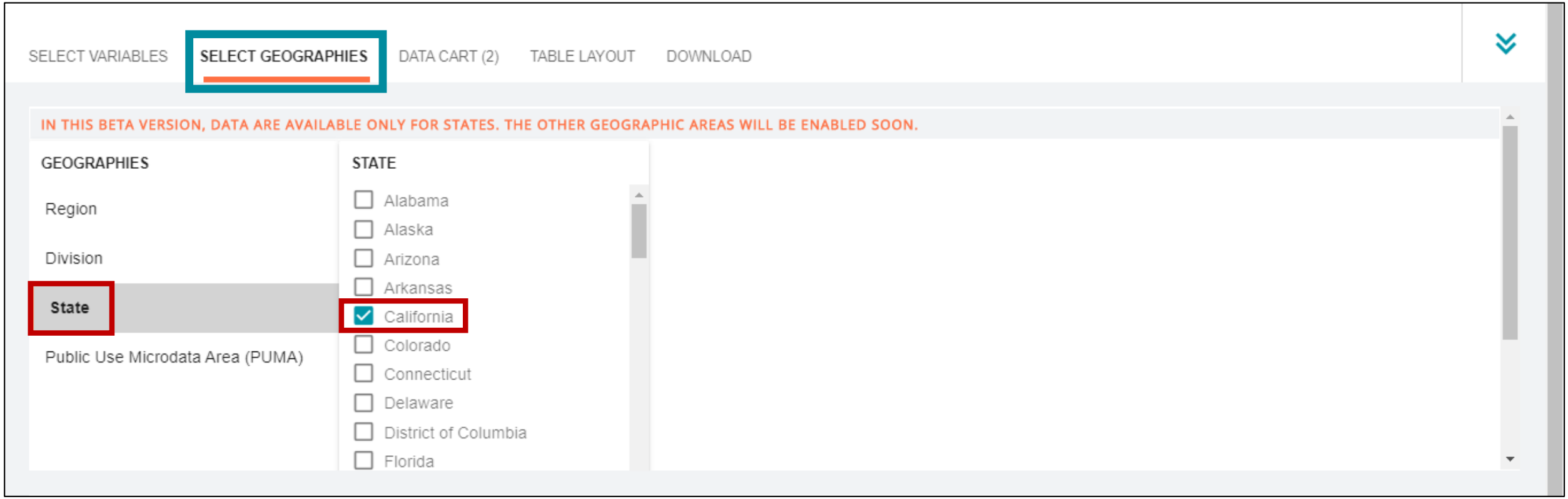


You can explore the full set of variables and values in the ACS PUMS Data Dictionary at https://www.census.gov/programs-surveys/acs/technical-documentation/pums/documentation.html

We will create a recode for the AGEP variable in a bit. For now, let's also search for marital status using the search bar. Enter 'marital' into the Label search bar—it should automatically bring up the variable, MAR, for Marital Status. Check the box next to it.

We can keep track of the number of variables we've selected by looking at the right side of the screen—if we accidentally failed to include one of our desired variables, we'll be able to tell, right away, by checking the number we have selected.

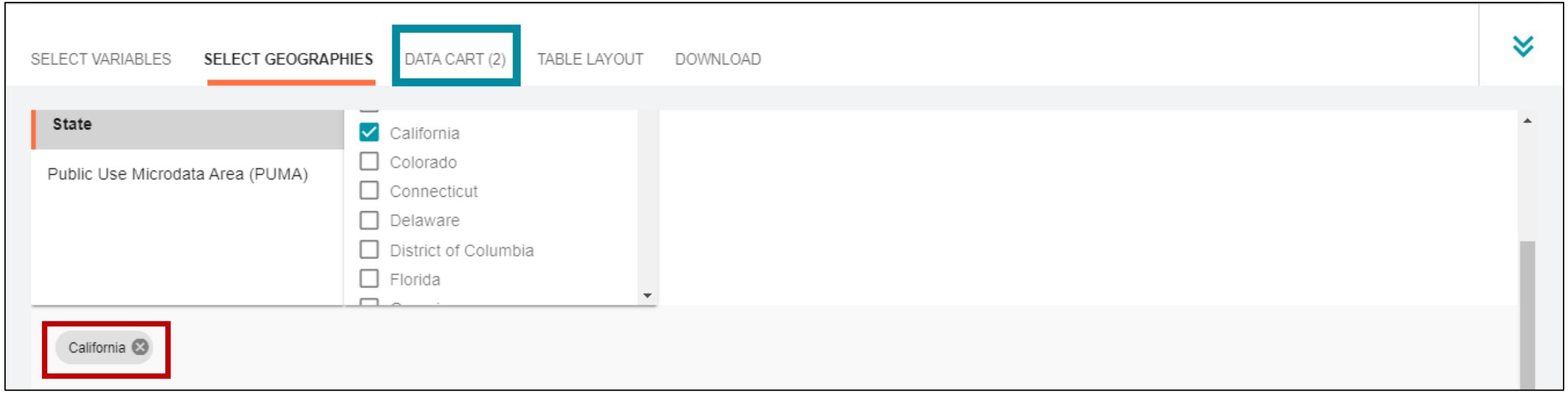
Now that we have our two variables selected, let's click on the SELECT GEOGRAPHIES tab.



Now that we're on the SELECT GEOGRAPHIES tab, let's choose our geography. Nation, State and Public Use Microdata Area (PUMA) geographies are available. If you do not select a geography, Nation will be used as the default geography.

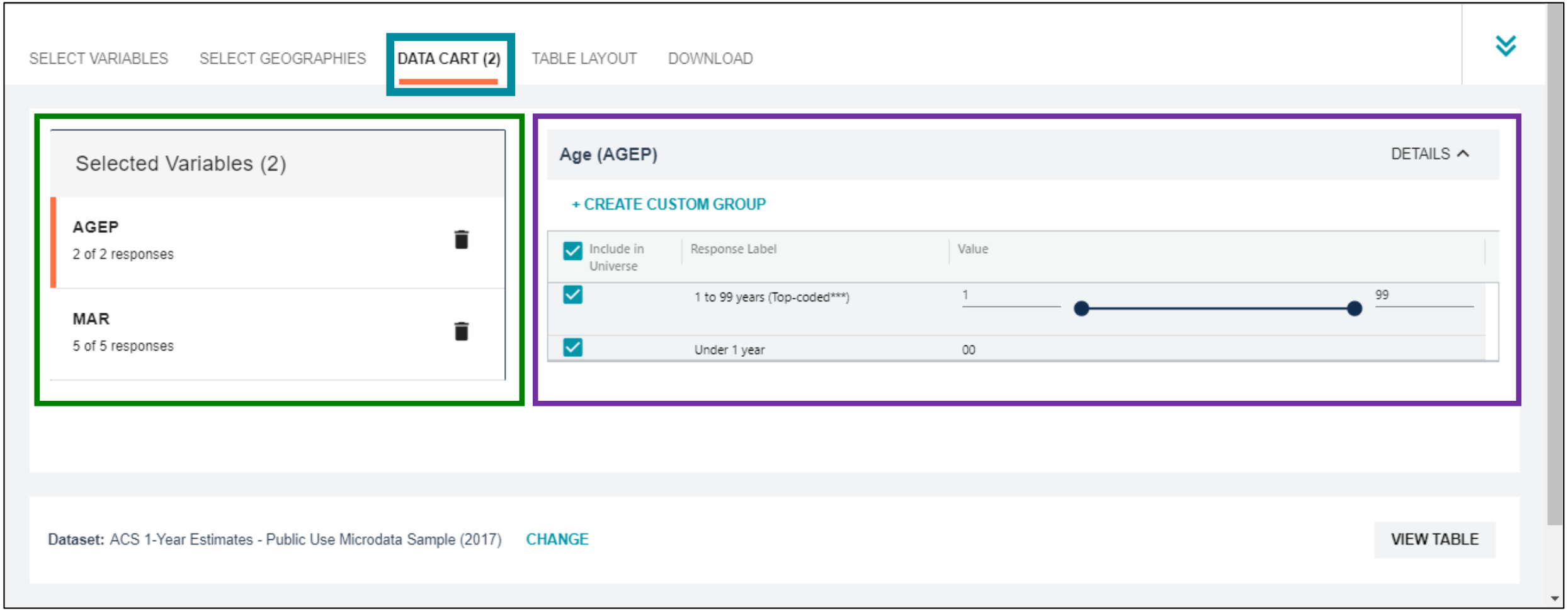
Other geographies will be enabled in the future. Once the other geographies are enabled, you'll see that the list is tailored according to the survey being used.

For this example, let's click on State. Next, we'll check the box next to California.

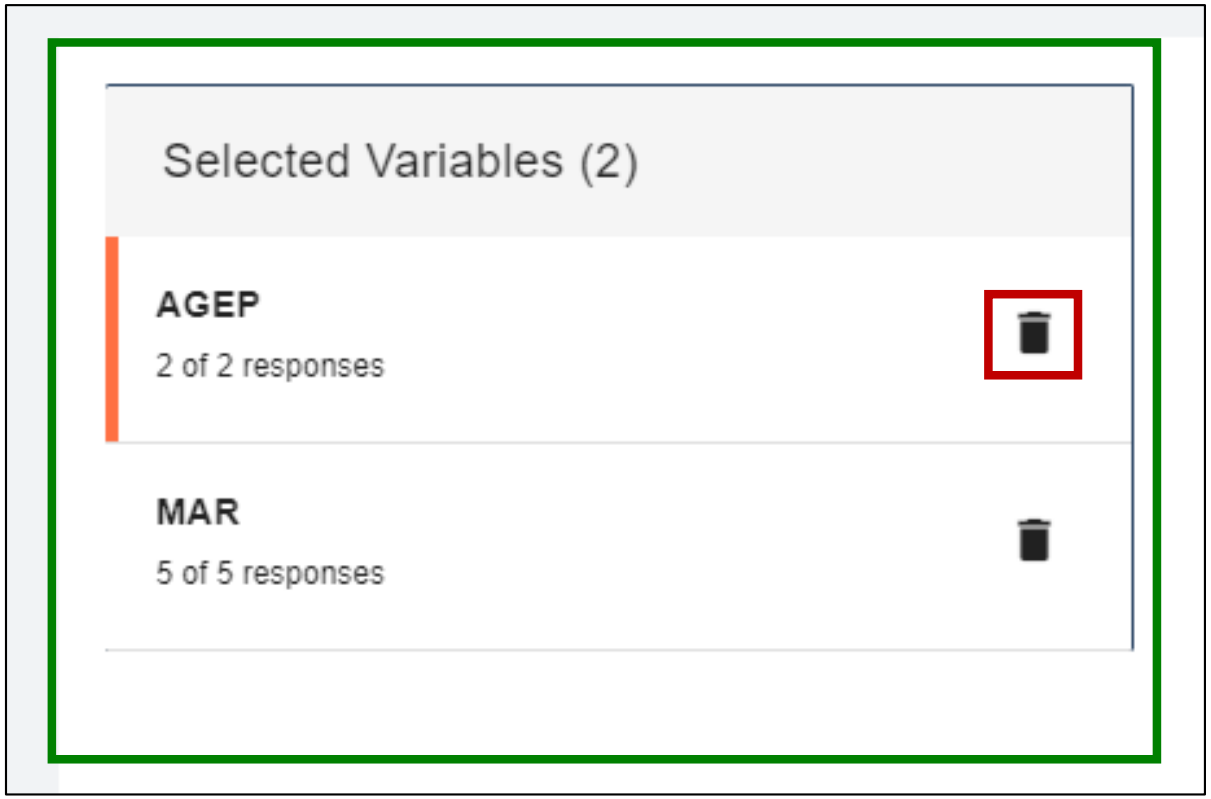


You can keep track of the geographies you've added by scrolling down. As you can see in the example above, we only have California selected. If you would like to remove a geography, you can click on the X located next to the state name.

Now that we have selected our geography, we'll move to the DATA CART tab.

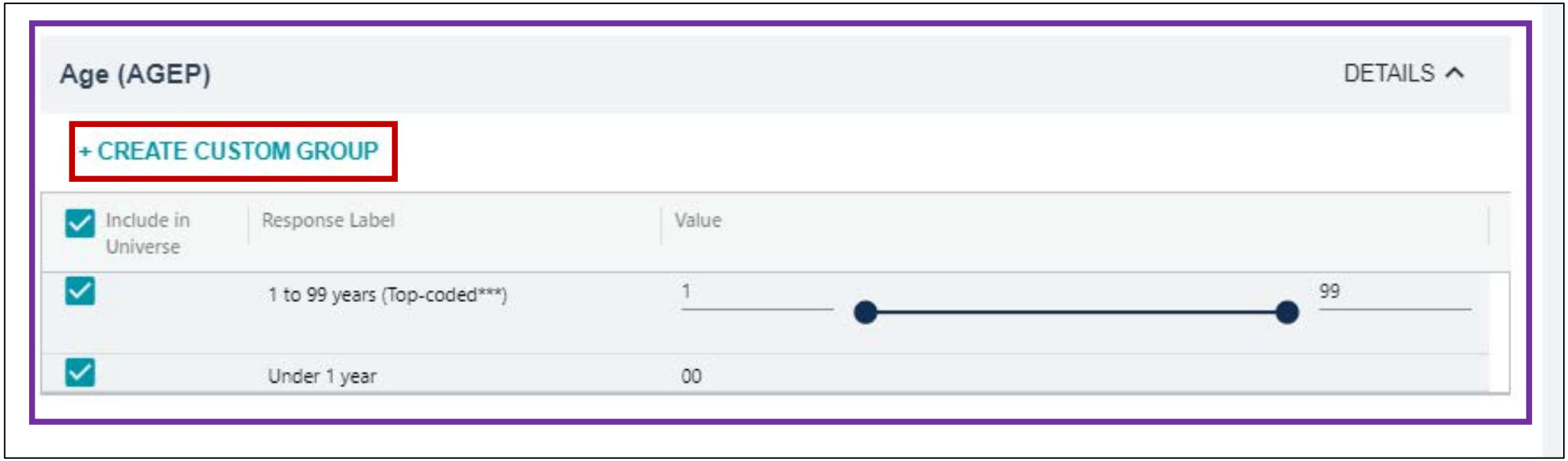


This is how your DATA CART tab should look. Your selected variables should be displayed on the left side of the screen (highlighted in the green box). The information for the variable you have highlighted on the left will be displayed on the right side of the screen (highlighted in the purple box)—this section is used to create the recodes. In our case, right now we have the AGEP variable selected, so we can create a recode for it. If we were to select the MAR variable, we could create a recode for that, as well.



In the Selected Variables box, you can see your selected variables, along with the current number of responses associated with that variable. For example, MAR, the marital status variable, has 5 different response options (Married, Widowed, Divorced, Separated, and Never married or under 15 years old).

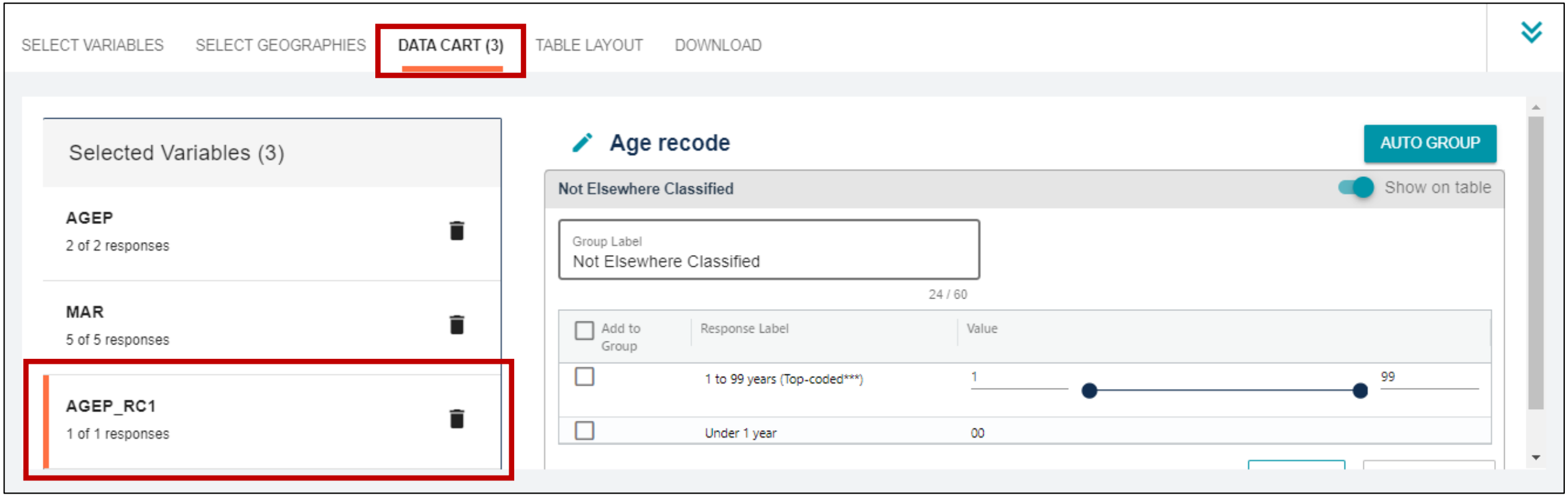
You can also click on the trashcan icon to remove a variable from your list.



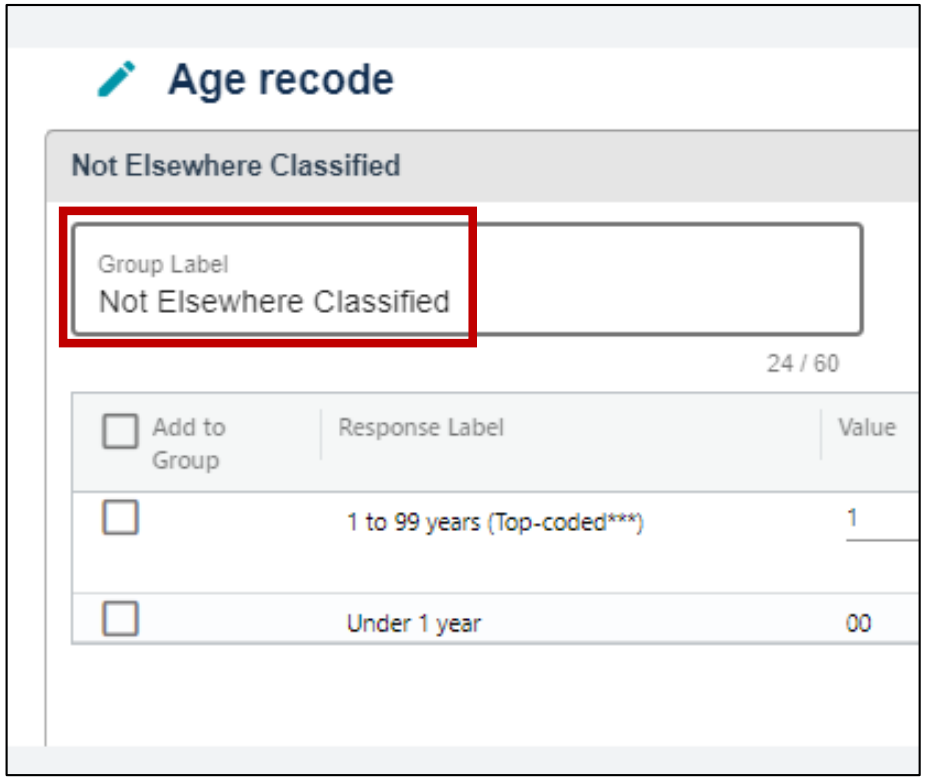
In the Age (AGEP) box, you'll find the current information about the variable, such as the Response Label and the value range.

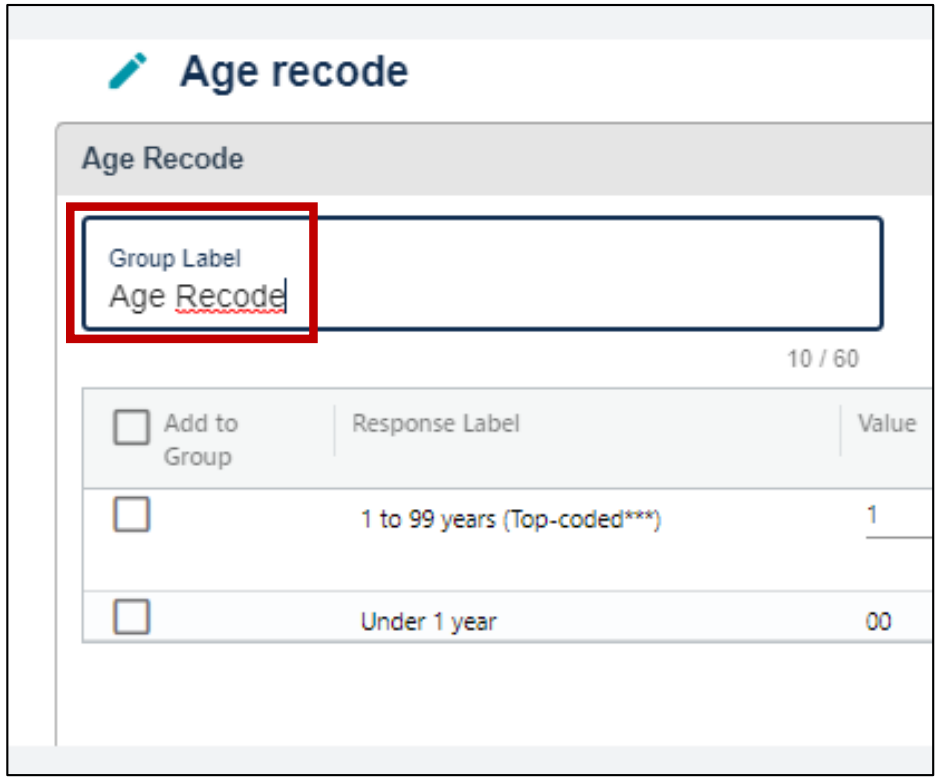
This is where we will create a recode of the Age (AGEP) variable.

To create the recode, select the + CREATE CUSTOM GROUP button.

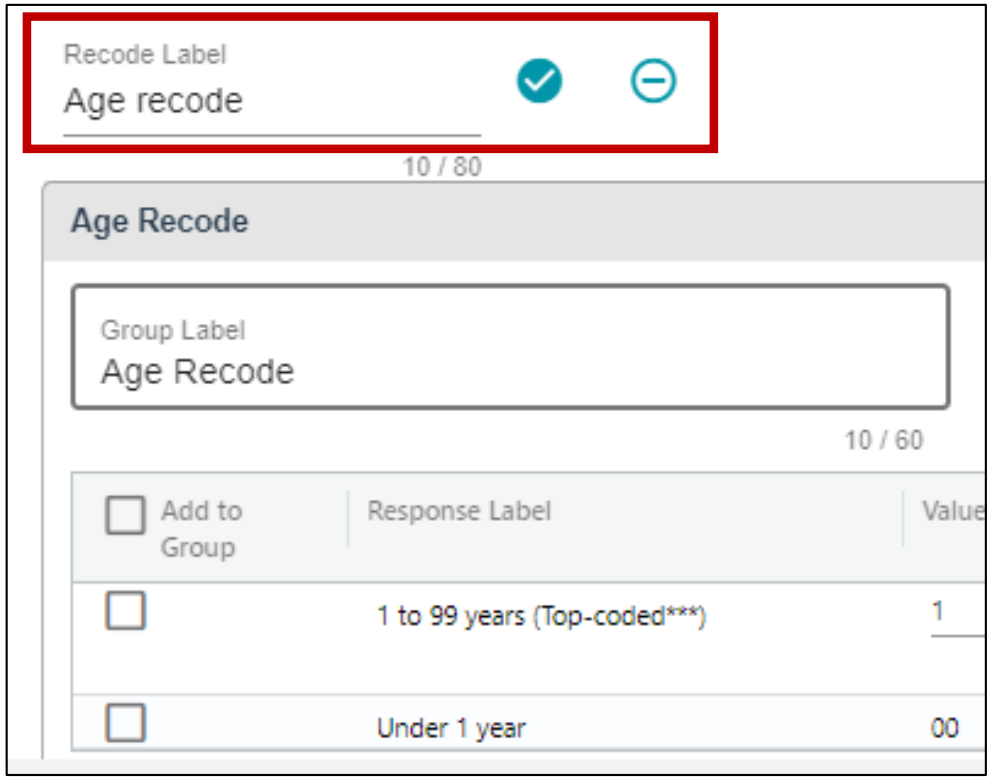


As soon as you click the + CREATE CUSTOM GROUP button, a new variable (the recode you have created) will be added to the Selected Variables list. Our recode is called AGEP_RC1. You'll also notice that we now have three variables in our DATA CART.

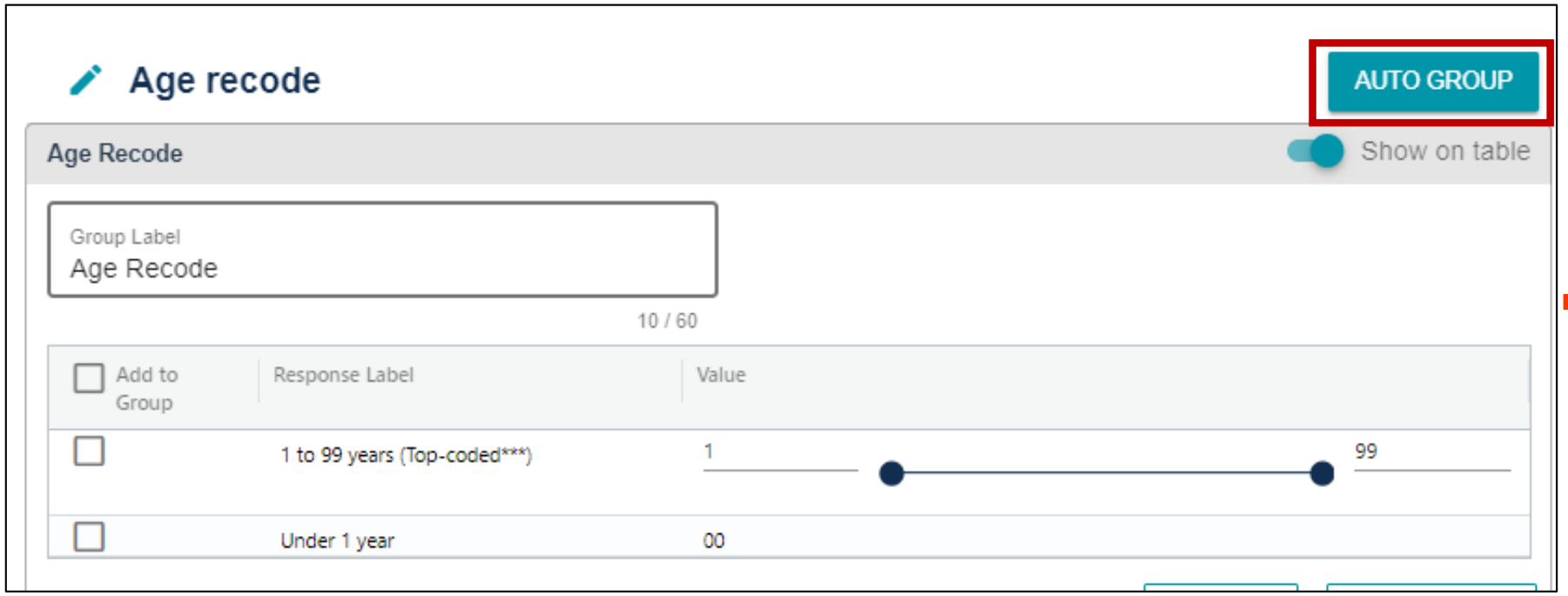


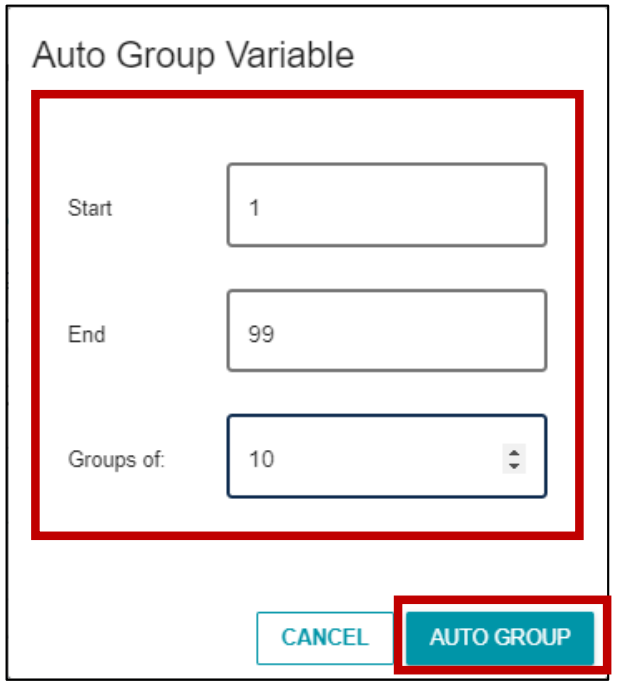


Before we do anything else, let's go over some things about this recode. First, we can rename the recode to something that makes sense. Let's call ours "Age Recode"—click on, or next to, the "Not Elsewhere Classified" text in the Group Label box, delete it, and type "Age Recode" in that spot.



You can also change the Recode Label by selecting the pencil icon next to the Age recode text. Once you are happy with the label, click on the check mark in the green circle.



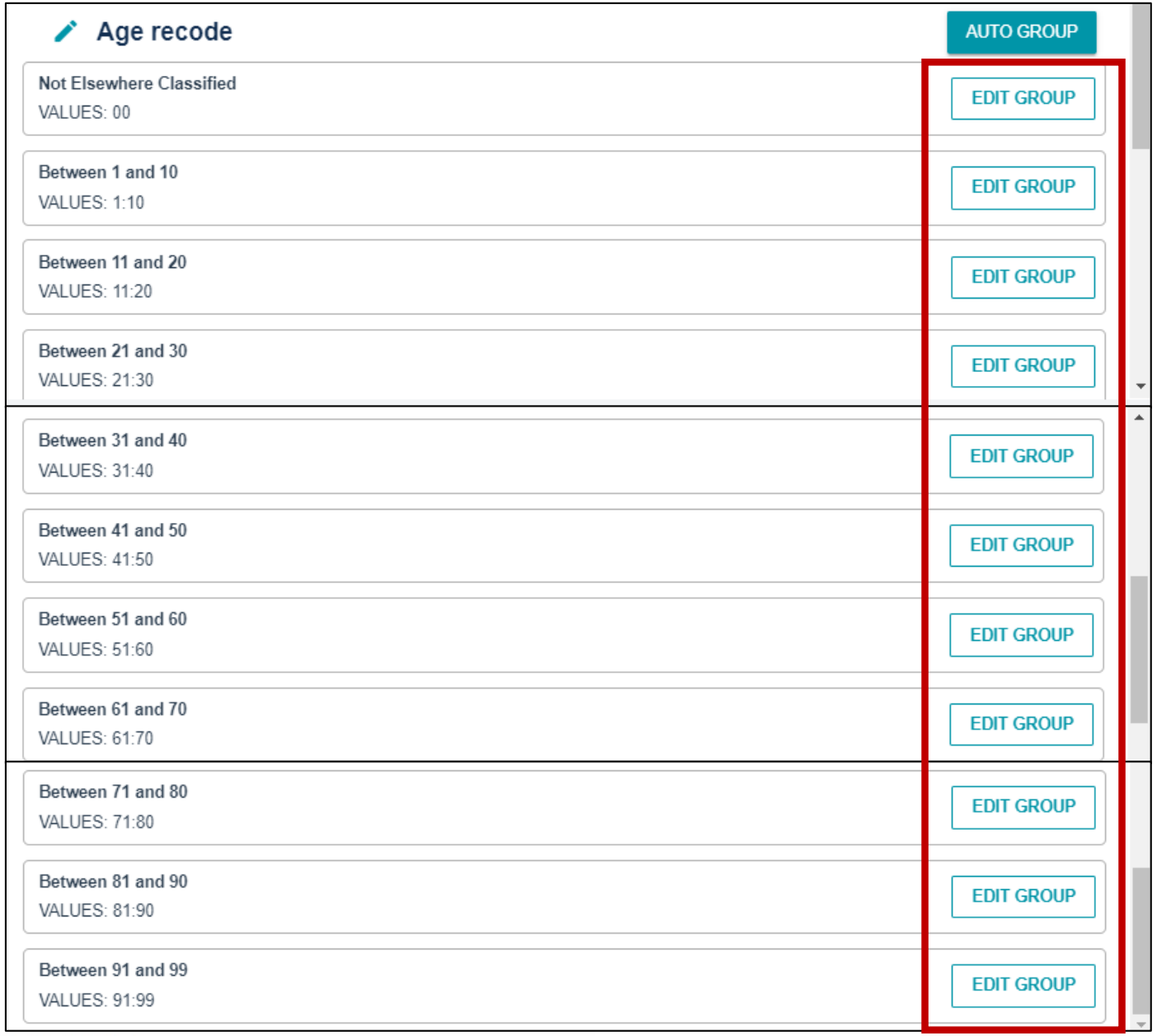


The site offers two ways to create custom groups of continuous variables. You can manually specify each grouping or you can use the Auto Group feature. For this example, we will use the Auto Group feature. Click on the AUTO GROUP button. This will bring up the Auto Group Variable box where you select your starting and ending values, as well as the spacing for the groups you would like to have.

Let's keep the starting value as 1 and the ending value as 99. However, let's change the number of groups from '1' to '10'. This will give us the full range of values split into groups of 10 (i.e., between 1 and 10, between 11 and 20, etc.) Then hit the AUTO GROUP button.

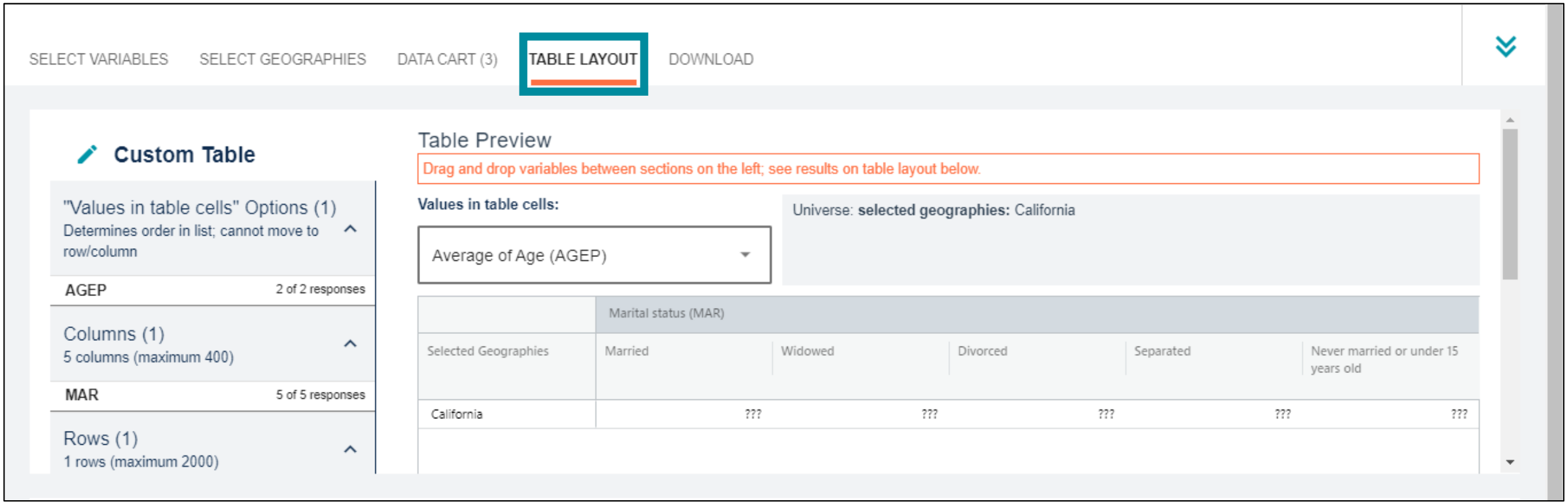
If you'd like to cancel the auto grouping, hit the CANCEL button.

Note: For an example of manually specifying each grouping, go to page 37 of this document, "Manually Specifying a Custom Group."



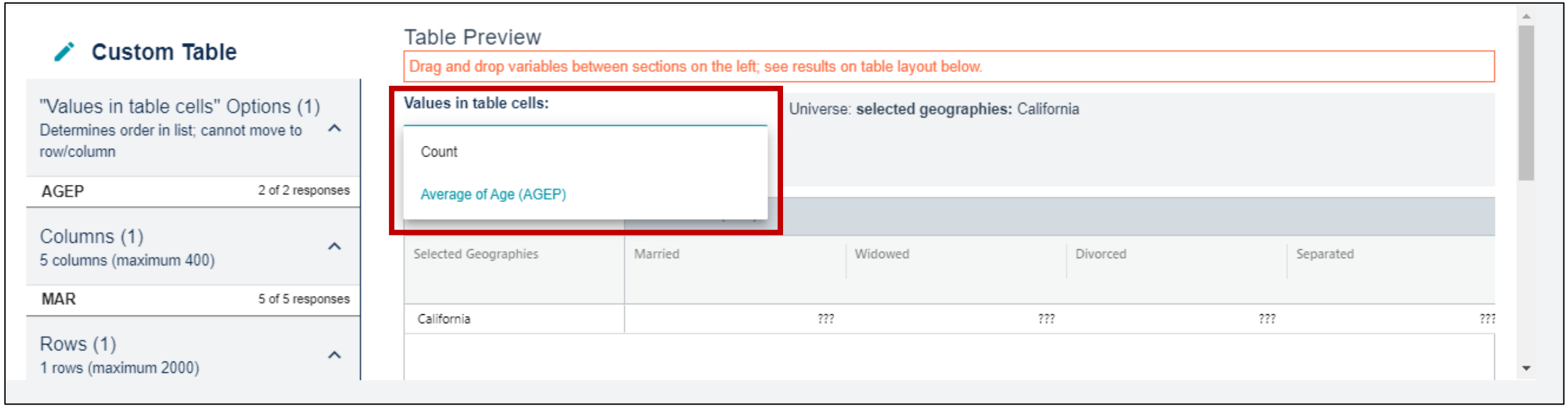
Now we have our groups.

If you wanted to further edit one of the groups, you can select the EDIT GROUP button located to the right of each grouping.

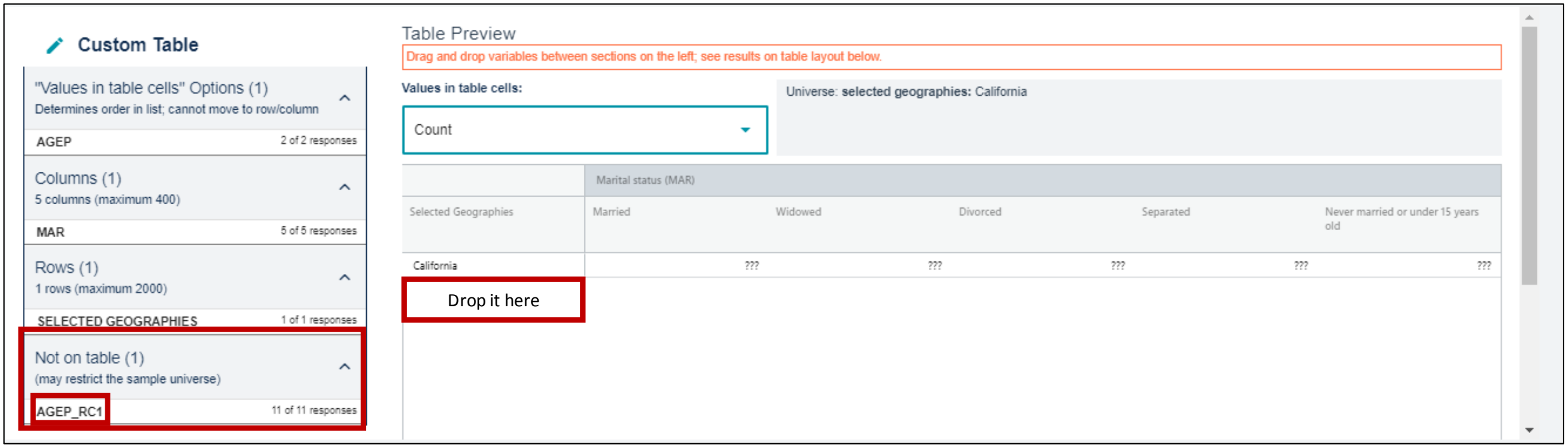


Now that we have our groups, let's move to the TABLE LAYOUT tab. This tab provides a preview of your table. As you can see, we have California placed as our row variable and marital status placed as our column variable. The ??? act as placeholders for the data that will populate the table.

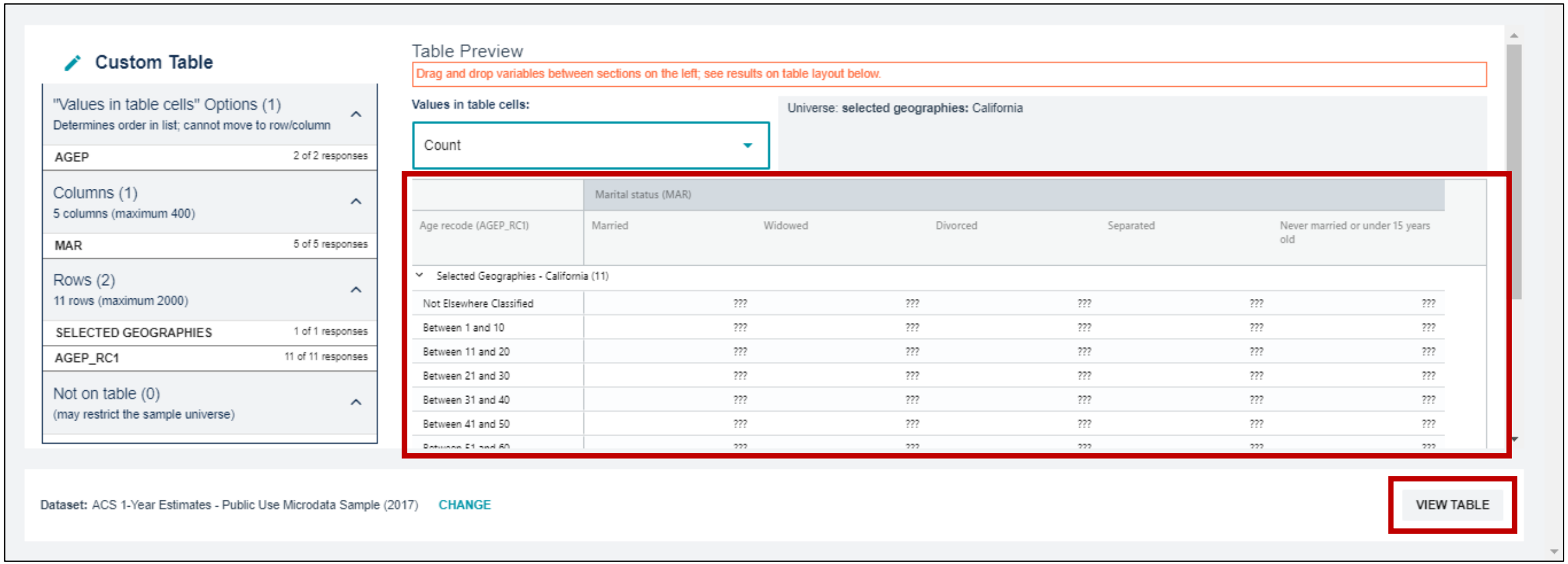
You can make modifications to the table by clicking on a row header or column header, holding the mouse, and dragging it to the spot you would like it to be.



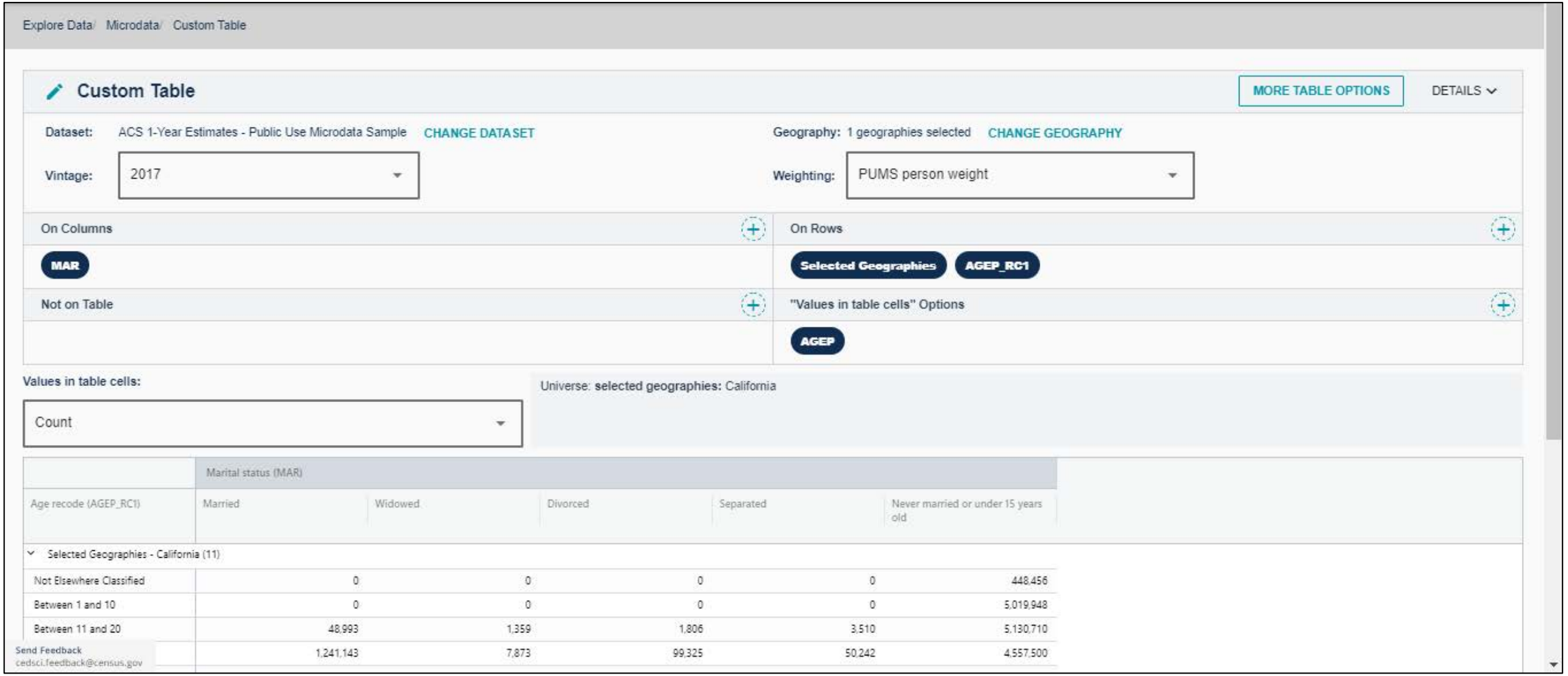
You may notice that the values in the table cells default to “Average of Age (AGEP).” We don’t want the average age for this example, so let’s click on the dropdown menu and select “Count.”



You may notice that our preview table is only using the geography we selected (California) and one of the variables that we selected (MAR, or marital status). To add our age recodes, scroll down until you see the AGEP_RC1 variable. Click on the AGEP_RC1 text. While holding the click on your mouse, drag it over to the table preview. Let's drop it right underneath the cell for California.



As you can see, we now have all the age recode groups included. You can continue to rearrange the table as you see fit. Once you have the table set up the way you want, click on the VIEW TABLE button located near the bottom of the screen.



This screen has a lot of different options on it. Let's take a few moments to go over the different things that are included on it.

Values in table cells:

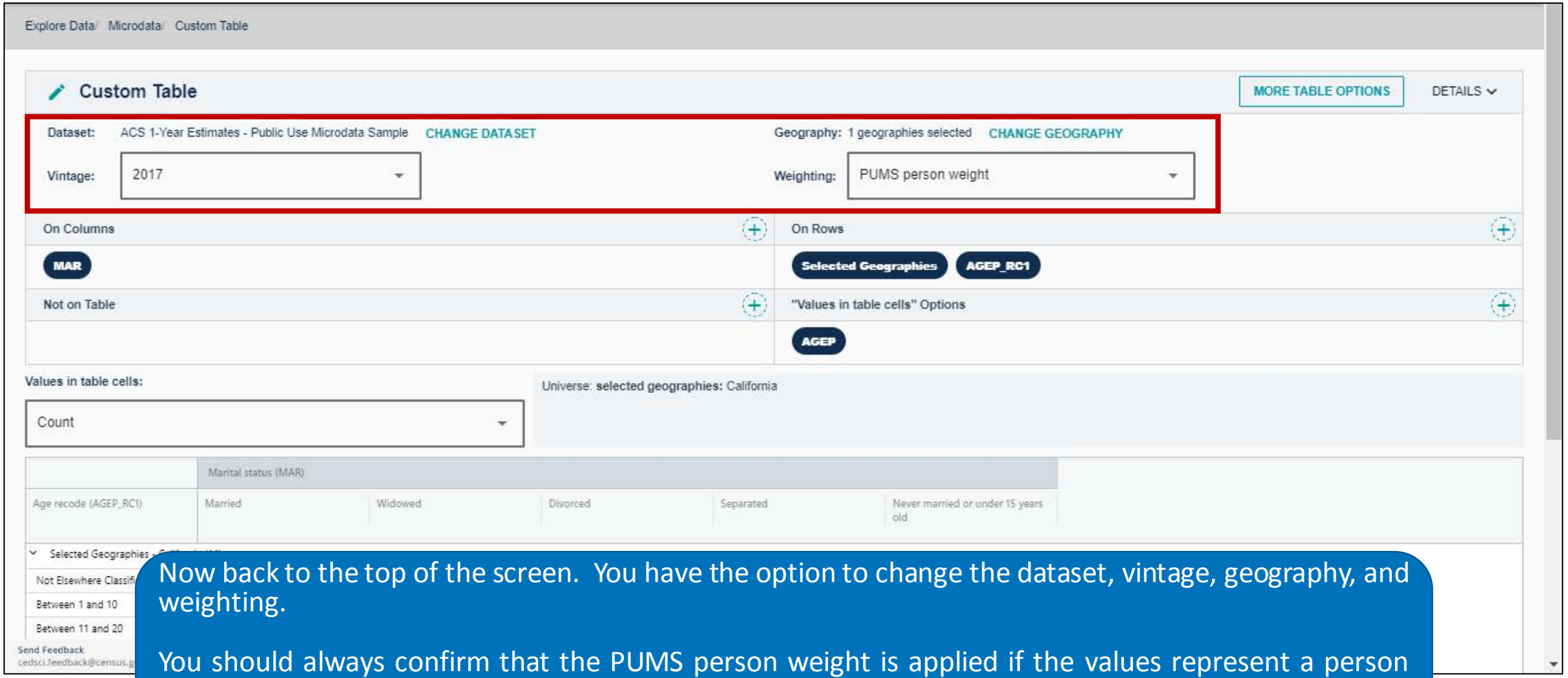
Count

Universe: **selected geographies:** California

Age recode (AGEP_RC1)	Marital status (MAR)				
	Married	Widowed	Divorced	Separated	Never married or under 15 years old
Selected Geographies - California (11)					
Not Elsewhere Classified	0	0	0	0	448,456
Between 1 and 10	0	0	0	0	5,019,948
Between 11 and 20	48,993	1,359	1,806	3,510	5,130,710
Between 21 and 30	1,241,143	7,873	99,325	50,242	4,557,500
Between 31 and 40	3,074,958	19,307	367,927	123,682	1,962,919
Between 41 and 50	3,277,733	57,189	606,603	161,093	994,306
Between 51 and 60	3,179,083	160,065	794,242	149,720	734,590
Between 61 and 70	2,462,445	330,708	674,918	92,765	384,236
Between 71 and 80	1,205,213	419,690	315,039	33,834	131,655
Between 81 and 90	411,965	419,112	103,711	9,999	43,375
Between 91 and 99	45,265	145,881	20,120	1,450	10,990

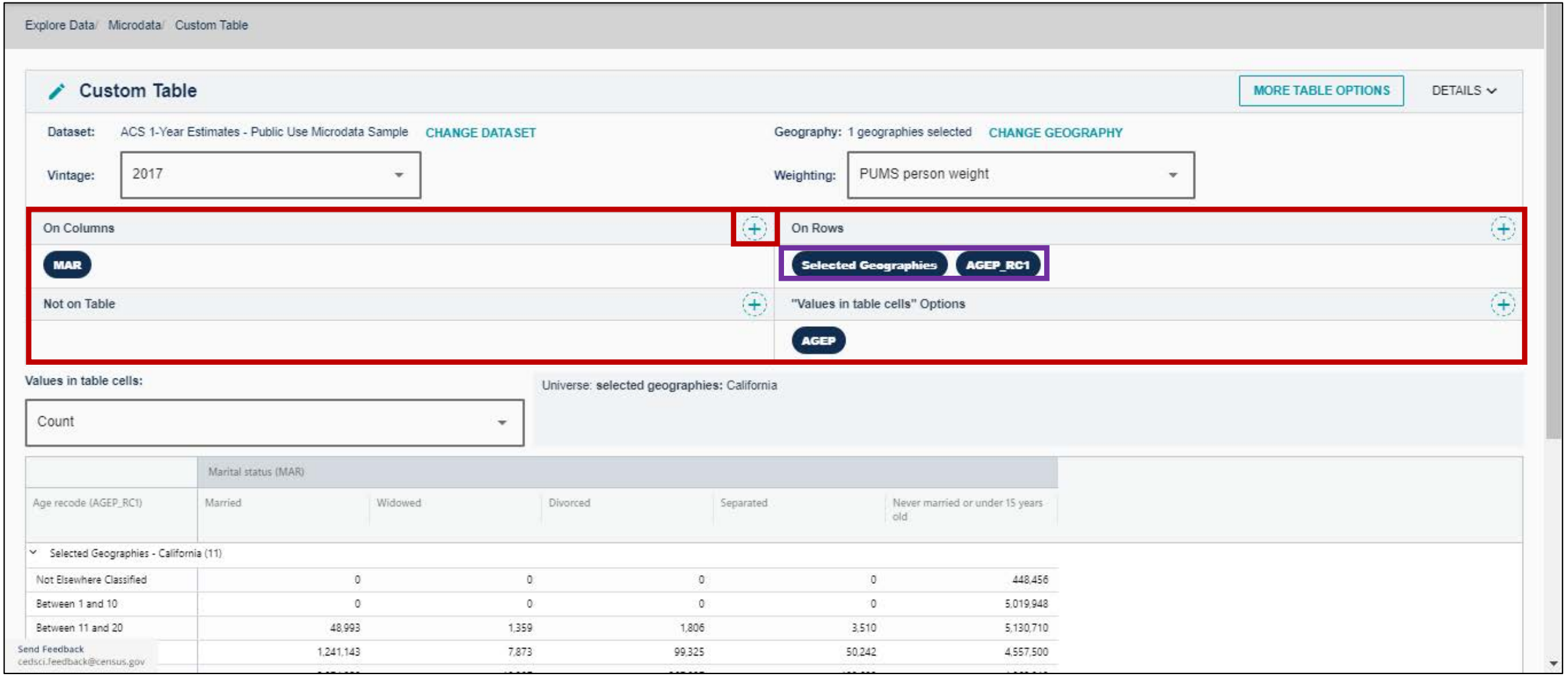
Send Feedback
cedsci.feedback@census.gov

First and foremost, we have our table displayed at the bottom of the page. As you can see, the table is populated with the actual estimates instead of the question marks we saw on the preview. Right above the table, you can see that our Universe is our selected geography, California.

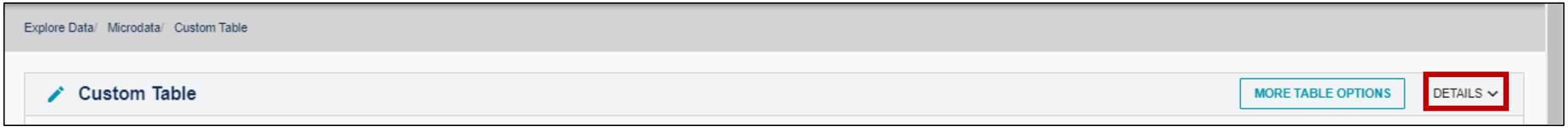


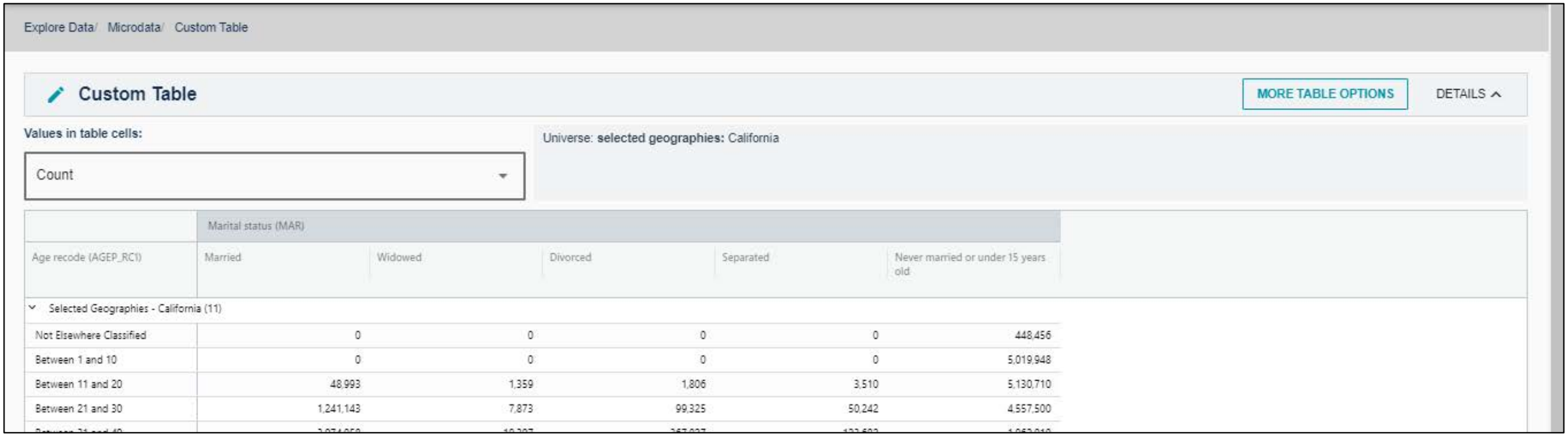
Now back to the top of the screen. You have the option to change the dataset, vintage, geography, and weighting.

You should always confirm that the PUMS person weight is applied if the values represent a person variable and the Housing Weight is applied if the values represent a household variable. For this example, the appropriate weighting is the PUMS person weight, since we are using person-level variables.

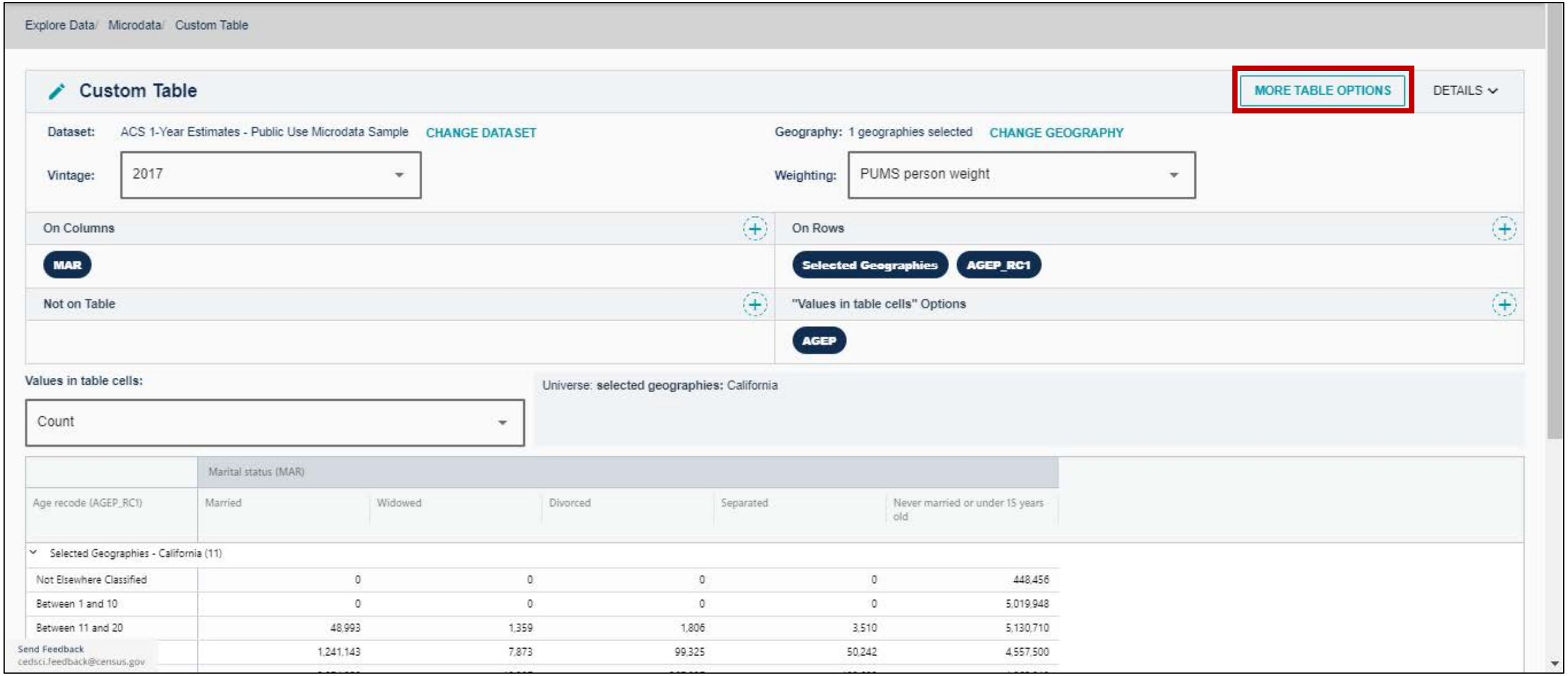


You can also move the variables around from a row to a column, or vice versa, using the black variable pills (examples highlighted in the purple box). If you need to add a variable (or remove one), you can click on the + in the circle—this will take you back to the earlier screen that lists all the variables.

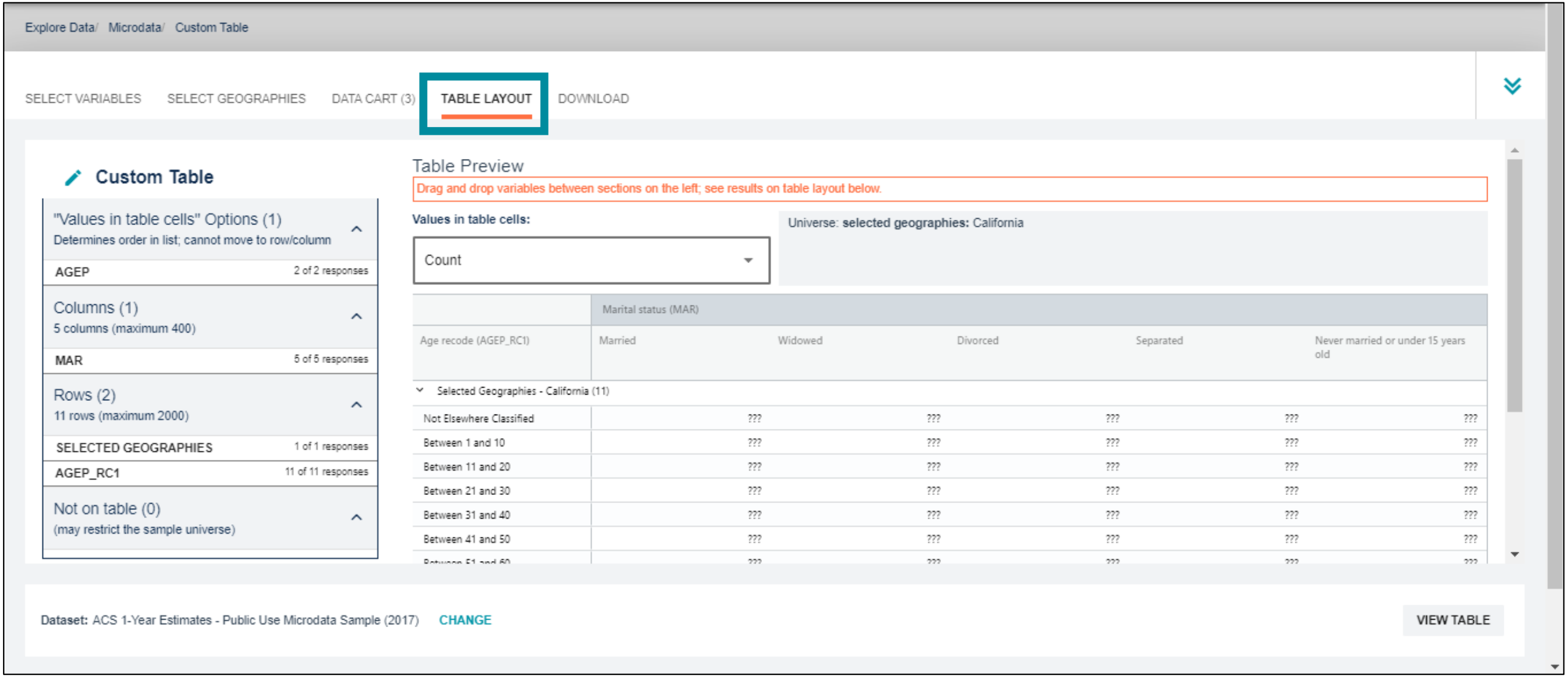




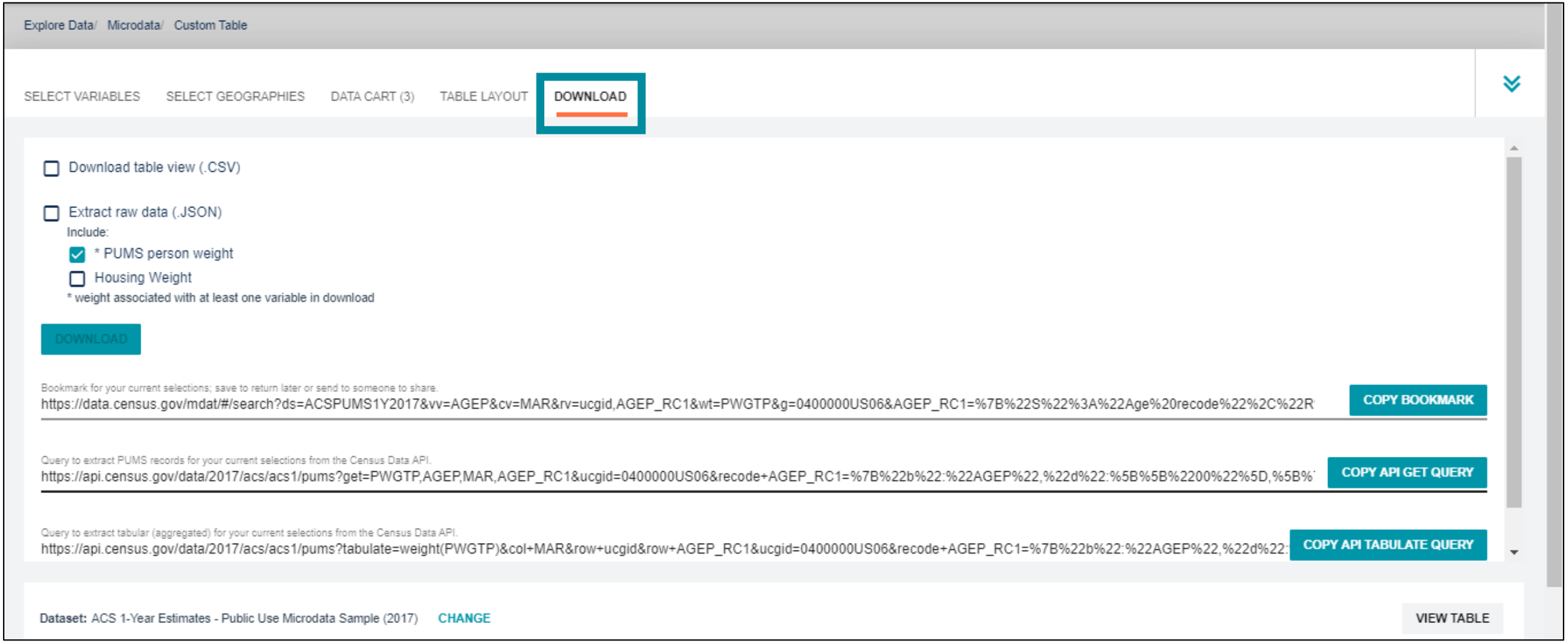
Clicking on the DETAILS button near the top of the screen will hide the upper part of the screen. Click on it again to display the detail pane again.



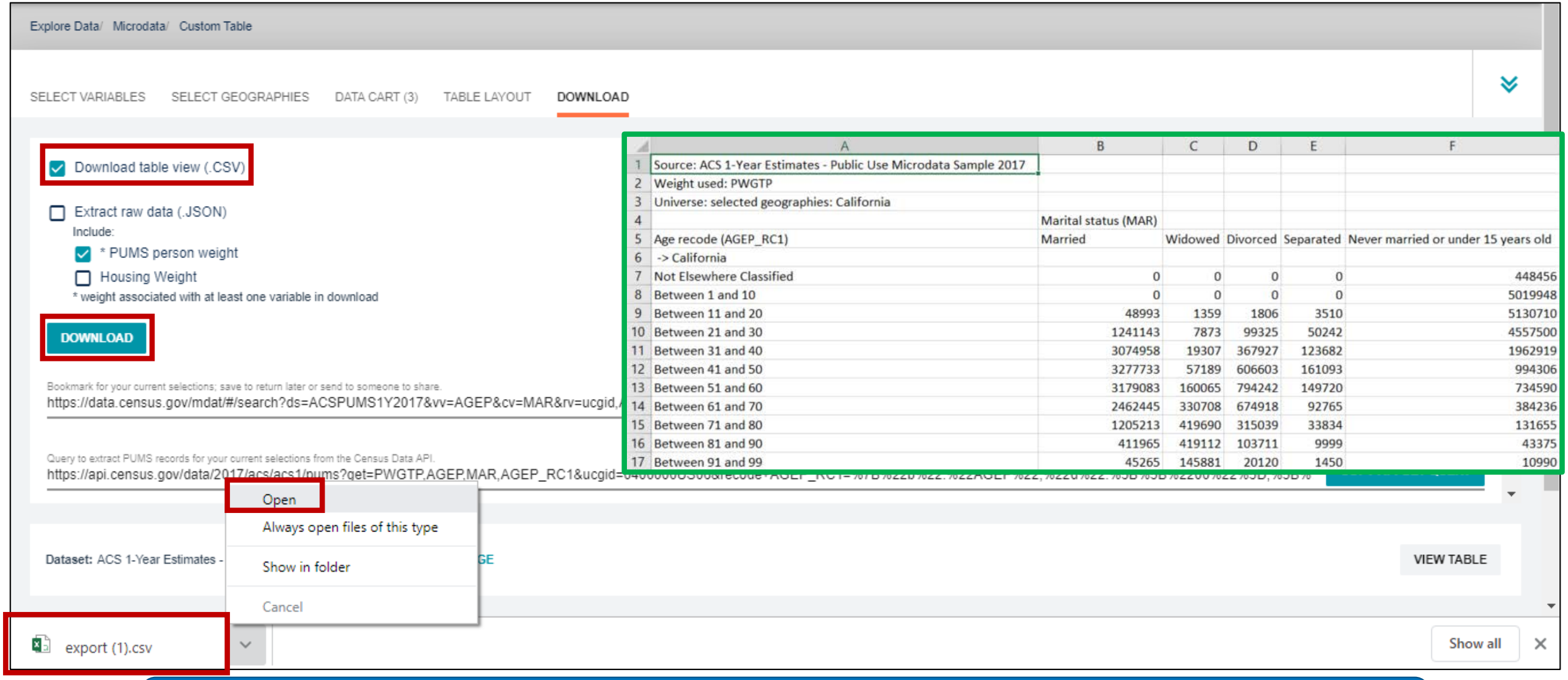
Clicking on the MORE TABLE OPTIONS will give you an opportunity to further customize your table or download/share it.



If you click on Customize Table (from the MORE TABLE OPTIONS button), you will return to the TABLE LAYOUT tab.

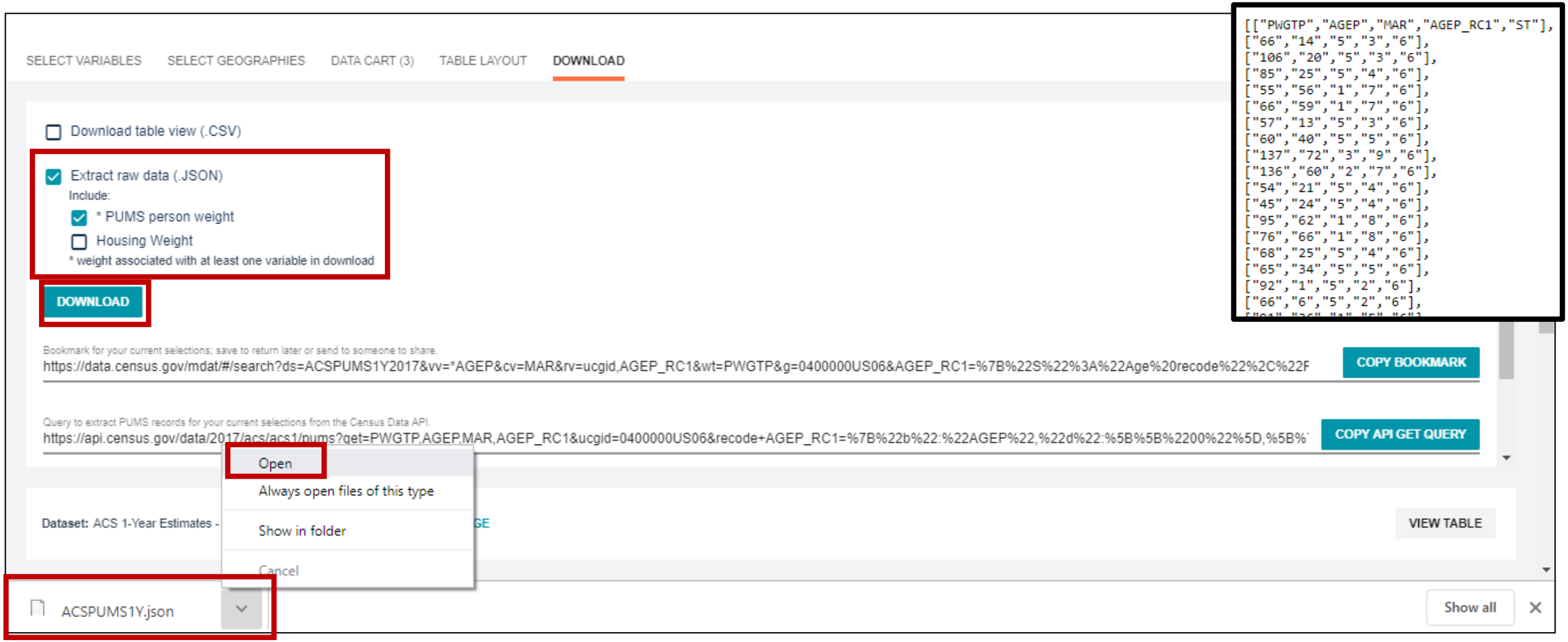


If you click on Download/Share (from the MORE TABLE OPTIONS button), you'll go to the DOWNLOAD tab. Here you can download the table as a CSV file, extract the raw data, copy the bookmark for the table, or copy API queries.

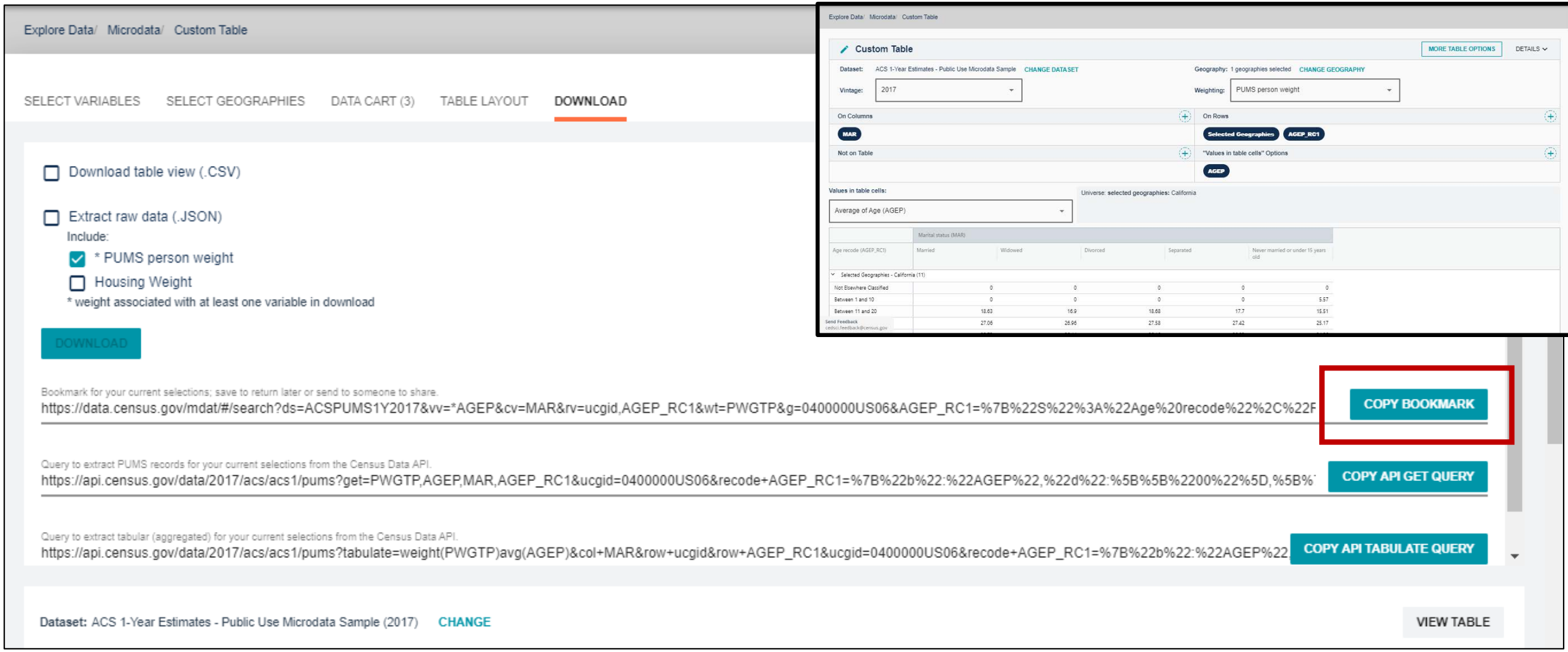


	A	B	C	D	E	F
1	Source: ACS 1-Year Estimates - Public Use Microdata Sample 2017					
2	Weight used: PWGTP					
3	Universe: selected geographies: California					
4		Marital status (MAR)				
5	Age recode (AGEP_RC1)	Married	Widowed	Divorced	Separated	Never married or under 15 years old
6	-> California					
7	Not Elsewhere Classified	0	0	0	0	448456
8	Between 1 and 10	0	0	0	0	5019948
9	Between 11 and 20	48993	1359	1806	3510	5130710
10	Between 21 and 30	1241143	7873	99325	50242	4557500
11	Between 31 and 40	3074958	19307	367927	123682	1962919
12	Between 41 and 50	3277733	57189	606603	161093	994306
13	Between 51 and 60	3179083	160065	794242	149720	734590
14	Between 61 and 70	2462445	330708	674918	92765	384236
15	Between 71 and 80	1205213	419690	315039	33834	131655
16	Between 81 and 90	411965	419112	103711	9999	43375
17	Between 91 and 99	45265	145881	20120	1450	10990

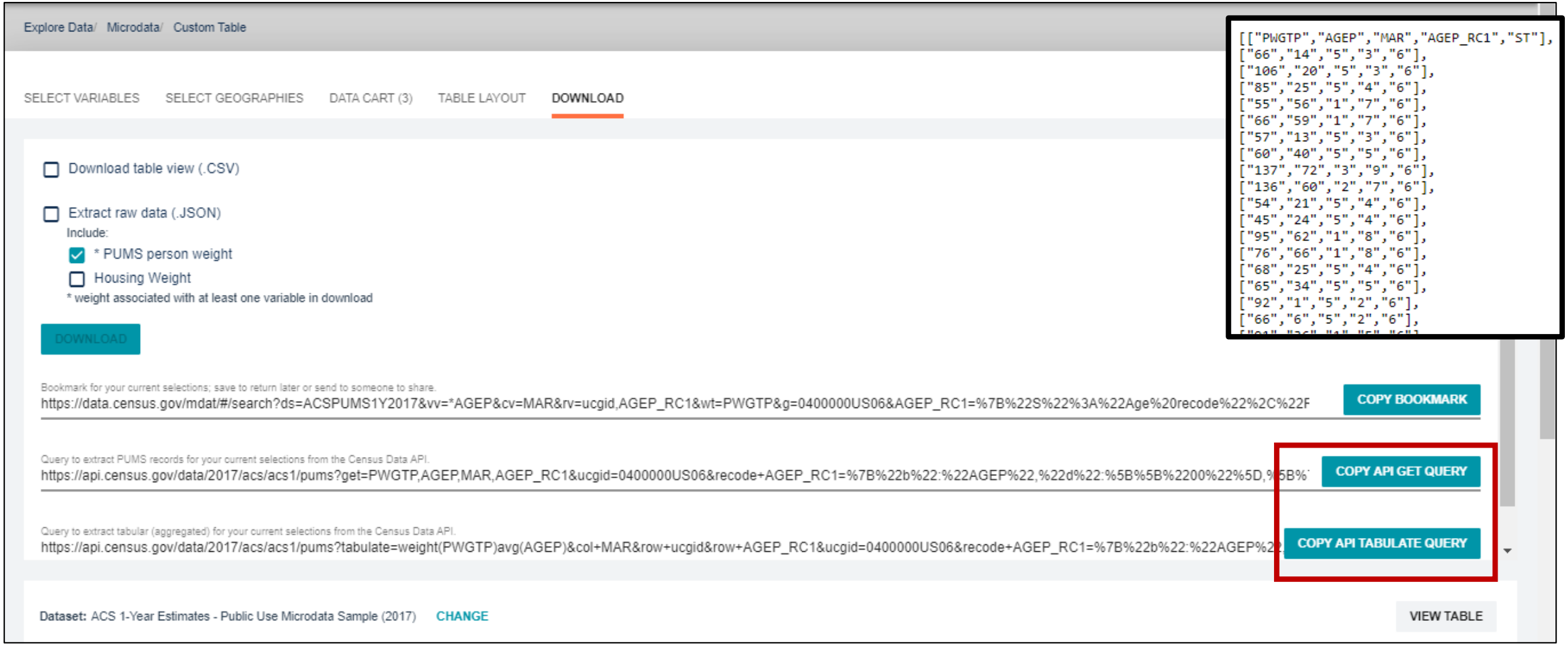
Check the box next to Download table view (.CSV) to download the table as a CSV file and hit the DOWNLOAD button. Open the downloaded file. It will automatically open in Excel.



To extract raw data in a JSON file, check the box next to Extract raw data (.JSON). Be sure that the weight files you would like are also selected. Then hit the DOWNLOAD button. It may take a few moments for the JSON file to be produced. Once it has finished downloading, open it.



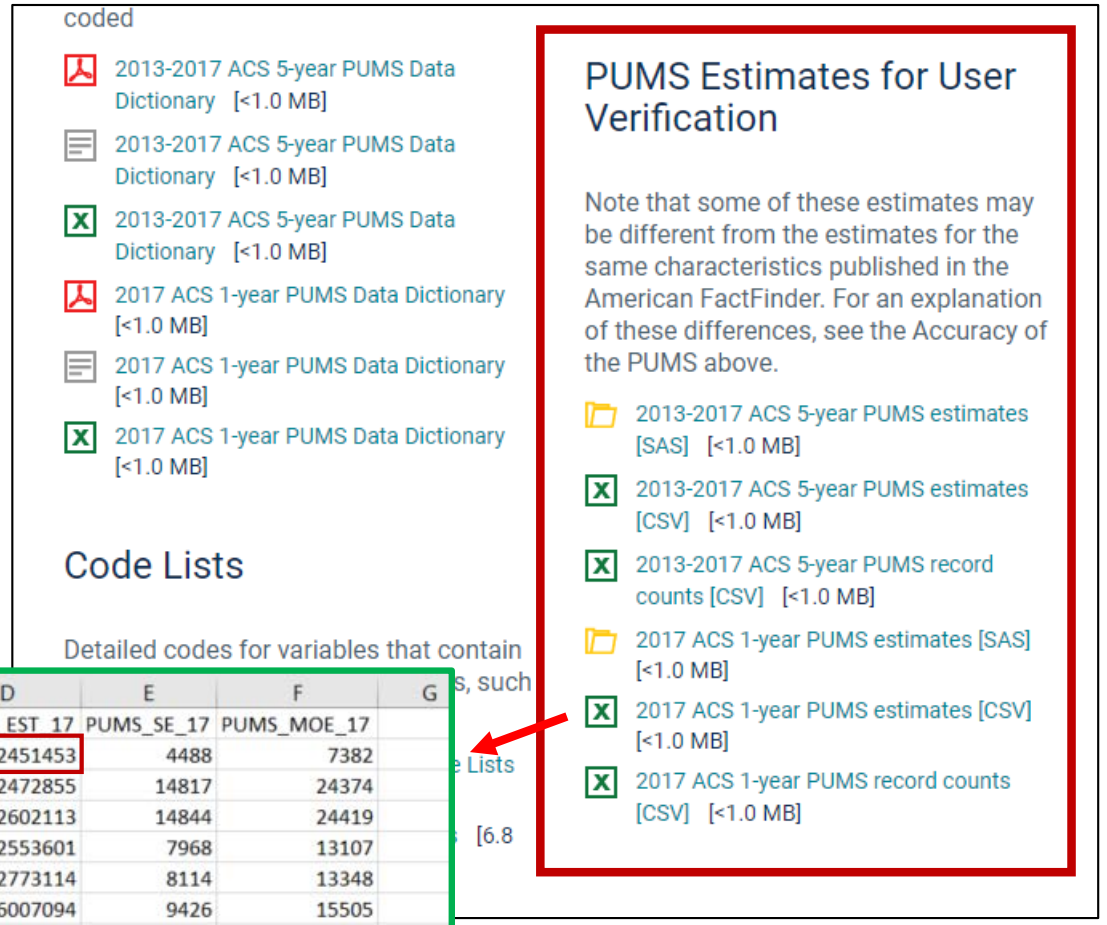
To copy the bookmark, select the COPY BOOKMARK button. Open a new window in your browser and paste the URL into the URL bar. Hit enter. It may initially look as though it has taken you to the landing page for Microdata Access. However, if you wait a few seconds, it will automatically open the Custom Table pane that is shown on page 24.



To copy the API queries, select either the COPY API GET QUERY or the COPY API TABULATE QUERY buttons. Open a new window in your browser and paste the URL into the URL bar. Hit enter. This will make the API call. The data can then be pulled into an Excel document.

HOW CAN I VERIFY THAT I AM CREATING MY TABLE CORRECTLY WITH THE ACS PUMS ON DATA.CENSUS.GOV?

To verify that you are creating your table correctly, walk through a few examples using the PUMS estimates for user verification (https://www.census.gov/programs-surveys/acs/technical-documentation/pums/documentation.html) and confirm that you get the same result as provided in the file. For example, you can see that using the 2017 ACS 1-year PUMS files for California, there were 2,451,453 people age 0-4 and 7,137,291 owner occupied housing units.

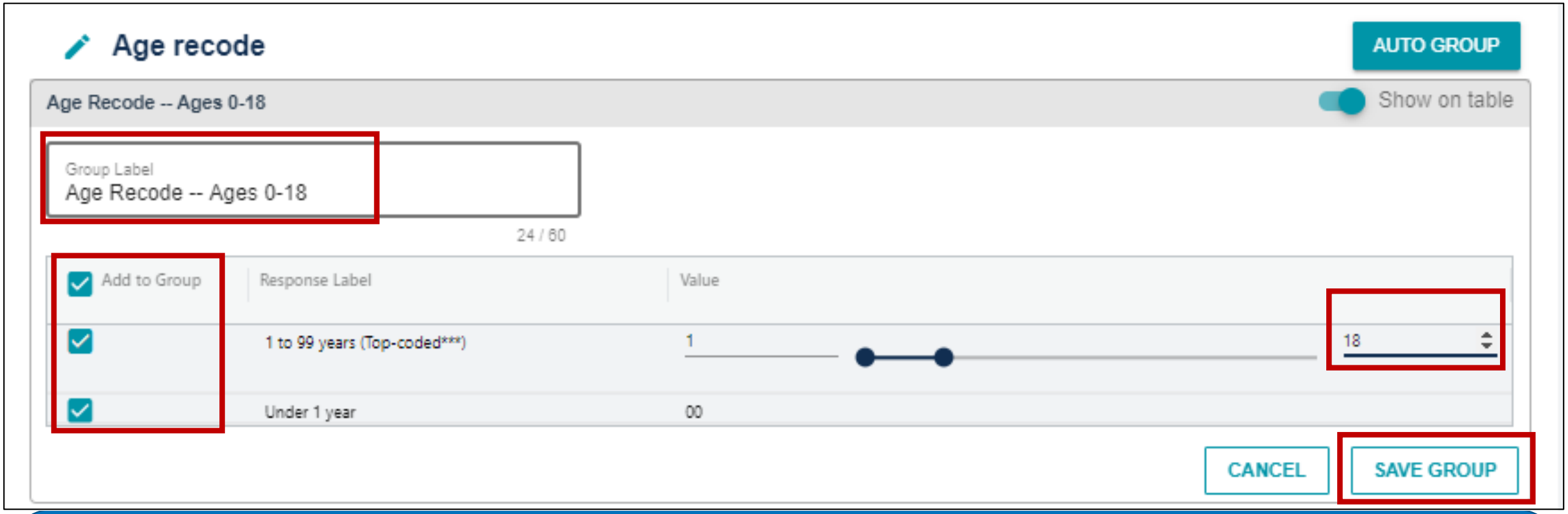


	A	B	C	D	E	F	G
1	ST	STATE	CHARACTERISTIC	PUMS_EST_17	PUMS_SE_17	PUMS_MOE_17	
2	6	California	Age 0-4	2451453	4488	7382	
3	6	California	Age 5-9	2472855	14817	24374	
4	6	California	Age 10-14	2602113	14844	24419	
5	6	California	Age 15-19	2553601	7968	13107	
6	6	California	Age 20-24	2773114	8114	13348	
7	6	California	Age 25-34	6007094	9426	15505	
8	6	California	Age 35-44	5258093	8827	14520	
9	6	California	Age 45-54	5150941	7519	12369	
10	6	California	Age 55-59	2495668	15269	25118	
11	6	California	Age 60-64	2263832	13717	22564	
12	6	California	Age 65-74	3210365	5196	8547	
13	6	California	Age 75-84	1580794	9594	15782	
14	6	California	Age 85 and over	716730	8535	14041	
15	6	California	Total housing units (TYPE=1)	14177270	568	934	
16	6	California	Total occupied units	13005088	10663	17541	
17	6	California	Owner occupied units (TEN in 1,2)	7137291	19666	32350	
18	6	California	Renter occupied units (TEN in 3,4)	5867797	20192	33216	

Manually Specifying a Custom Group

As discussed on a page 18, the site offers two ways to create custom groups of continuous variables. You can manually specify each grouping or you can use the Auto Group feature.

The next few pages provide directions for manually specifying your groupings.



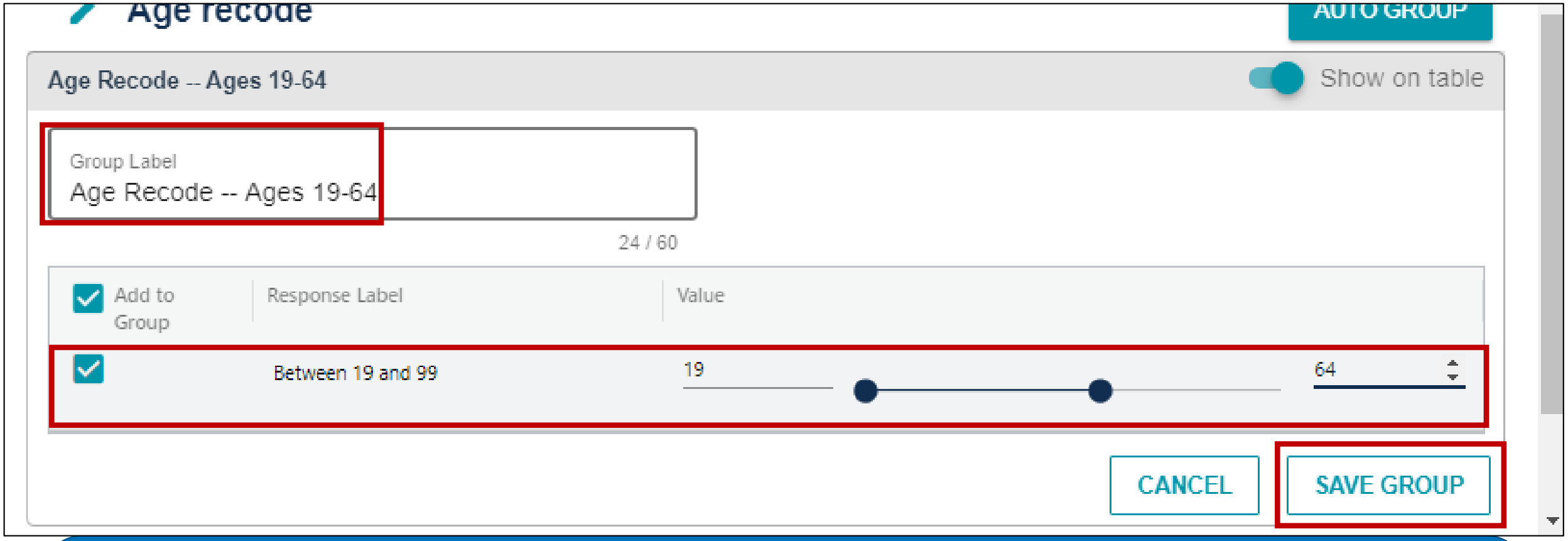
We will pick this example up from page 17. We will be creating our custom groups manually. With this method, each group needs to be created one at a time.

To do this, follow the steps through page 17 of this presentation. Once you have reached the Age recode screen, give your first group a different label. Let's call this first group "Age Recode – Ages 0-18." Enter that in the Group Label box by clicking in the box, deleting the existing text, and replacing it with "Age Recode – Ages 0-18." Next, check the boxes next to both "1 to 99 years (Top-coded***)" and "Under 1 year." This will allow us to add people under 1 year of age to our new recode.

Change the starting/ending values as needed; for this example, we only need to change the end value to 18. Then click on the SAVE GROUP button.

Now you'll notice that you have your new age group of Ages 0-18 and a group with the remaining ages. The values for your new age group are displayed as "1:18, 00." This signifies that people aged 1 to 18 years are included in this groups, as well as people aged 00 (under 1 year).

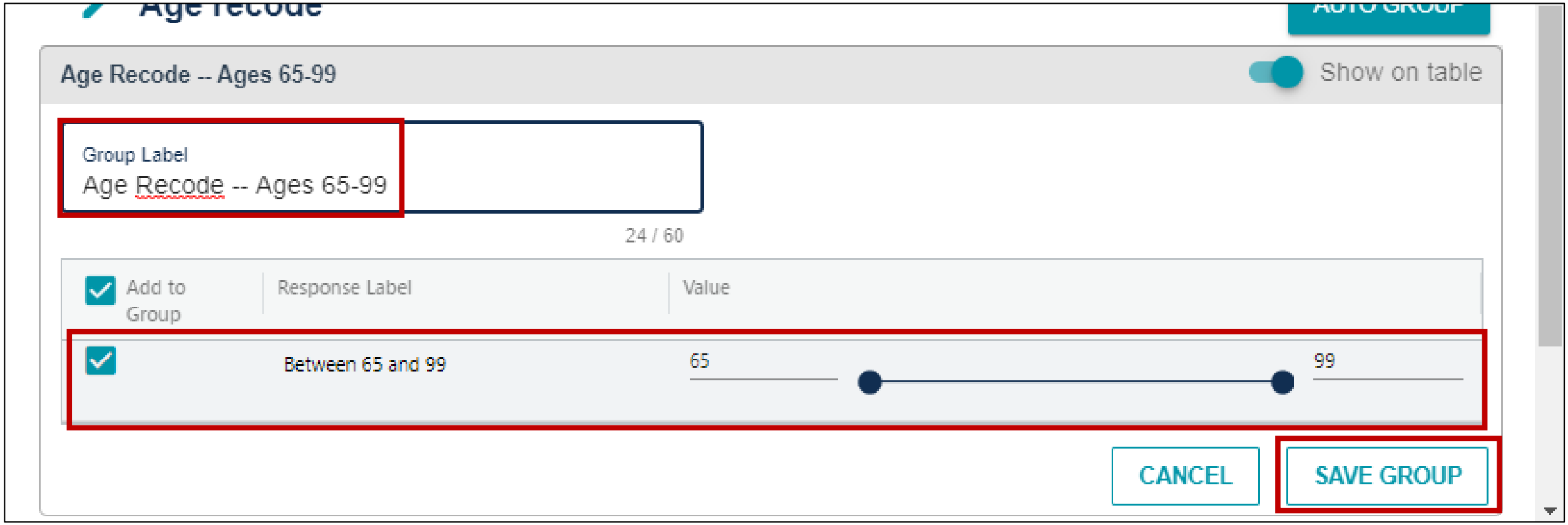
To add an additional age group, select the EDIT GROUP button for the "Not Elsewhere Classified" group.



Like you did before, change the Group Label to something that distinguishes it from the other groups, check the box next to "Between 19 and 99," change the starting/ending range as needed, and click on SAVE GROUP. For this example, we'll create a recode group for people ages 19-64.

NOTE: You can not have any ages that overlap between multiple groups. For example, the tool will not allow you to create a group for people ages 0 to 18 and another group for people ages 17-64. This is not allowed because you are trying to include 17 and 18 year olds in more than one group.

Now you'll see that you have three different groupings listed: one for Ages 0-18, another for Ages 19-64, and the last one for people ages 65-99 (still named "Not Elsewhere Classified"). Let's click on the EDIT GROUP button to change this last label.



To do this, delete the original text and replace it with "Age Recode – Ages 65-99." Check the box next to "Between 65 and 99." There's no need to change the starting or ending values for this recode group. Click SAVE GROUP to save.

Now you have all three custom groups that you manually specified.

From here, you can continue with the steps found on page 20 of this document.

Questions?

cedsci.feedback@census.gov