Using the U.S. Census Bureau's Statistical Testing Tool

One of the most important uses of American Community Survey (ACS) data is to make comparisons between estimates—across different geographic areas, different time periods, or different population subgroups. In addition to determining whether an estimate is comparatively higher or lower, data users also need to account for the sampling error associated with each estimate by conducting statistical testing.

All estimates produced from sample surveys have uncertainty associated with them as a result of being based on a sample of the population rather than the full population. Sampling error is the difference between an estimate based on a sample and the corresponding value that would be obtained if the estimate were based on the entire population. Statistical testing shows whether the observed difference between estimates likely represents a true difference that exists within the full population and that the difference is unlikely to have occurred by chance. In such cases where the observed differences between estimates are statistically significant, making statements regarding their differences is valid. In other cases where the observed differences between estimates are not statistically significant, making comparison statements regarding their differences outside of similarity is invalid as it misrepresents the data.

The U.S. Census Bureau developed a [Statistical Testing Tool](#) to allow all data users to carry out statistical testing with ACS data. The tool is an Excel spreadsheet with tabs that automatically calculate statistical significance when comparing estimates. While the tool was designed for ACS data, it may be used to conduct statistical testing with data from other Census Bureau surveys.[1] When using data from other surveys, data users should consult the relevant technical documentation for how to handle special cases. Otherwise, this tool may return a result of "not applicable" when a statistical test could be carried out.

This guide provides step-by-step instructions for using the Statistical Testing Tool to determine whether differences between ACS estimates are statistically different.

1. Key Benefits to Using the Statistical Testing Tool

To use the Statistical Testing Tool, you first need to obtain estimates and their associated margins of error (MOE) through one of the Census Bureau's data tools. These estimates and MOEs will then be copied into the Statistical Testing Tool and the results of the statistical test will be displayed. The results will show "Yes"—the estimates are statistically different from each other, "No"—the estimates are not statistically different, or "N/A"—statistical testing is not applicable.

The Statistical Testing Tool uses Z-scores to determine statistical significance at the 90 percent confidence level. In the case of a statistically significant result, this means that there is less than a 10 percent chance that the observed difference between the two estimates occurred randomly. The tool does not include an option for multiple comparison adjustments (e.g., Bonferroni test).

The Statistical Testing Tool is flexible and designed for data users to paste data directly into a spreadsheet with very little editing or adjustments needed. You can use the Two-Estimates spreadsheet to compare over 3,200 pairs of estimates, and you can use the Multiple Estimates spreadsheet to compare up to 150 estimates with each other. You can also easily change the parameters to modify the significance level (e.g., from 90 percent to 95 percent) or to conduct a test based on Standard Errors instead of MOEs.

---

[1] If you are looking for information on how to use ACS Variance Replicate Estimates or the Public Use Microdata Sample (PUMS) to create margins of error, please see our Variance Replicate Estimate Tables <www.census.gov/programs-surveys/acs/data/variance-tables.html> and PUMS Documentation <www.census.gov/programs-surveys/acs/microdata/documentation.html> pages.

2. Accessing the Statistical Testing Tool

The Statistical Testing Tool can be found on the Census Bureau's Website.

From the ACS homepage, click on "Guidance for Data Users" (see Figure 2.1).

Figure 2.1
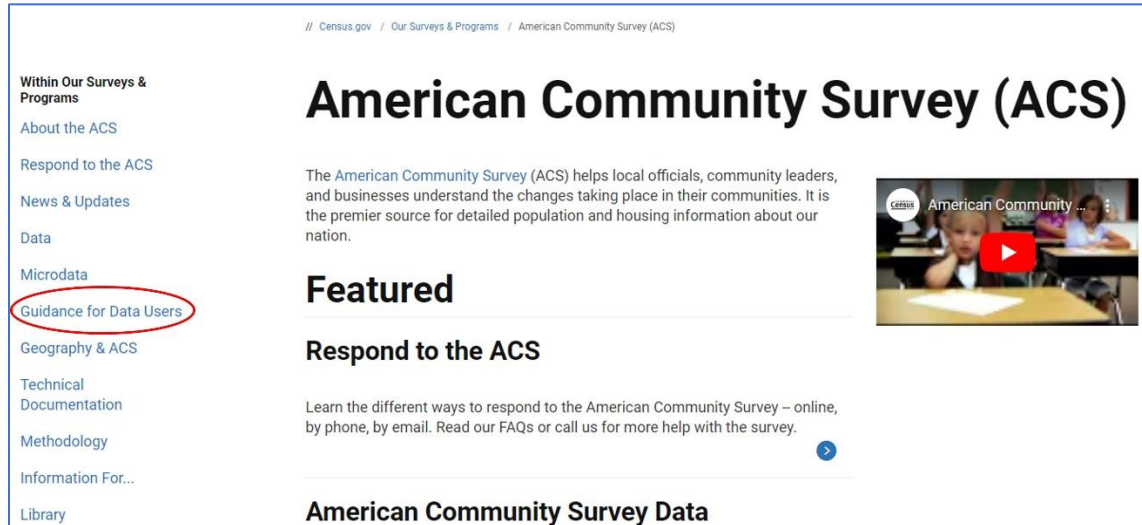
Source: U.S. Census Bureau, American Community Survey (ACS), <www.census.gov/programs-surveys/acs>.
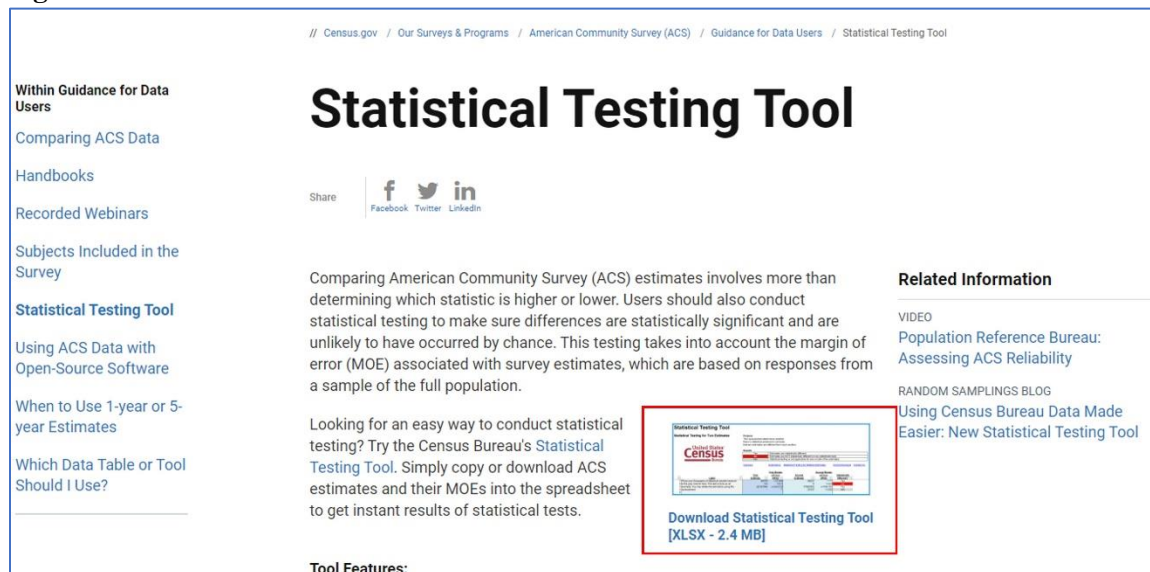
Then click on "Statistical Testing Tool" (see Figure 2.2).

Figure 2.2

Source: U.S. Census Bureau, American Community Survey (ACS), <www.census.gov/programs-surveys/acs/guidance.html>.

The Statistical Testing Tool Web page provides some background information and a link to download the tool (see Figure 2.3).

Figure 2.3

Source: U.S. Census Bureau, American Community Survey (ACS), <www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html>.

3. Data Requirements to Use the Statistical Testing Tool

The data requirements for using the Statistical Testing Tool are minimal. All you need to get started are two or more comparable estimates and their associated margins of error (or standard errors). You can use the tool to conduct statistical tests for counts (such as the number of people or households), means, medians, percentages, and other types of estimates.

While the Statistical Testing Tool is designed for data obtained through data.census.gov, you can also use Census Bureau data from other sources, such as the Application Programming Interface (API) and My Congressional District.

The Statistical Testing Tool can be used to compare two estimates to each other—for example, comparing change across time for the same geography, comparing two geographies to each other, or comparing two population groups.

The tool can also be used to compare many estimates to each other, such as comparing estimates for all counties in a state to each other. This comparison can be useful when trying to rank or order data across many geographies or groups.

Data users may wish to compare more than 150 estimates to each other. For example, they may wish to find the county with the highest or lowest estimate for a characteristic. One way to handle this is to use the two statistical testing tabs together. For example, to find the highest county estimate, first sort the estimates. Then compare the top 150 counties using the "Stat_Test_for_Multiple_Est" tab to find the one with the highest estimate. To ensure that this is indeed the highest estimate, copy all county estimates into the first estimate column in the "Stat_Test_for_Two_Est" tab. In the second estimate column, put the

highest county estimate for all of the rows. This will ensure that the statistical results hold for all counties and not just the top 150. In addition, this method may be repeated for the 2nd highest estimate, 3rd highest estimate, etc.

4. Data Comparability

When using the Statistical Testing Tool, it is important to only compare estimates that are measured in equivalent ways. Be sure to review survey documentation to ensure that the estimates are comparable.

Variables may change over time, for example, and new question wording or response categories are occasionally introduced.

Geographic boundaries can also change. For example, congressional districts, metro areas, counties and cities occasionally redraw boundaries based on changes in governing structures. Many statistical areas (like census tracts and block groups) are updated once per decade to reflect the most recent decennial census.

The Census Bureau's Comparing ACS Data Web page provides annually updated information about the year-to-year comparability of ACS data across different subjects.

With each ACS data release, the Census Bureau also provides a list of Table and Geography changes that may affect data users' ability to compare ACS data over time.

5. Worksheets in the Statistical Testing Tool

The Excel-based Statistical Testing Tool is comprised of multiple sheets, or tabs:
- **Overview:** Provides information on the importance of statistical testing and basic features of the Statistical Testing Tool.

- **Instructions**: Provides brief step-by-step instructions on how to copy data into the tool and how to interpret results.

- **Statistical Testing for Two Estimates (Stat_Test_for_Two_Est)**: Use this tab to insert estimates and margins of error when comparing two estimates, such as comparing data across two time periods or two population subgroups.

- **Statistical Testing for Multiple Estimates (Stat_Test_for_Mult_Est)**: Use this tab to insert estimates and margins of error when comparing two or more estimates to each other, such as estimates for all counties in a state.

- **Worked Example:** Provides step-by-step instructions with screenshots on how to use the tool for conducting statistical testing for multiple estimates.

- **Contact information (Contact_US):** Provides contact information for help/questions at the Census Bureau as well as links to more information about the ACS.

## 6. Using the Statistical Testing Tool to Compare Two Estimates

- **Example:** A city planner wants to understand how commuting patterns have changed over time.

- **Data:** 2019 ACS 1-year estimates, 2015 ACS 1-year estimates

- **Table:** S0801: "Commuting Characteristics by Sex"

- **Variables:** Number of workers, Means of transportation to work, mean travel time

- **Method of obtaining data:** Copy-paste from data.census.gov

A city planner in Henderson, NV wants to understand how commuting patterns have changed since 2015. The planner needs data on the number of workers, how they travel to work, and the average travel time to work in 2015 and 2019.

They start their query by going to the Census Bureau's data.census.gov Website and clicking on "Advanced Search" (see Figure 6.1).

Figure 6.1
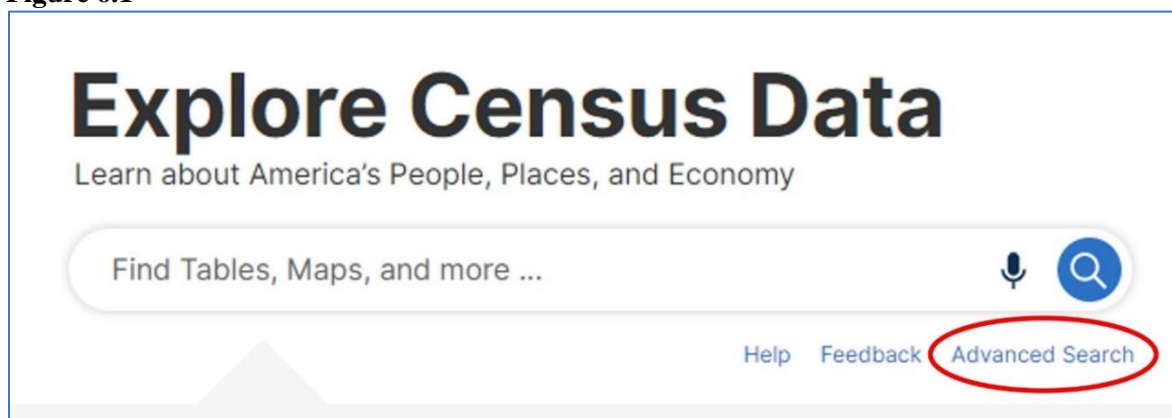
Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

They select "Geography," then "Place" (see Figure 6.2).

Figure 6.2
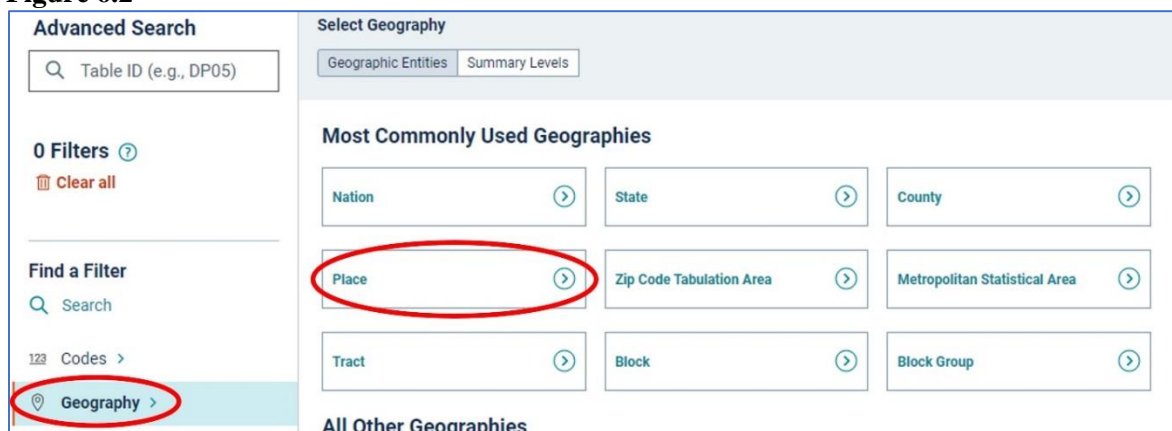
Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

Within the Place menu, they select "Nevada" and then scroll down to click on "Henderson city, Nevada" (see Figure 6.3).

Figure 6.3
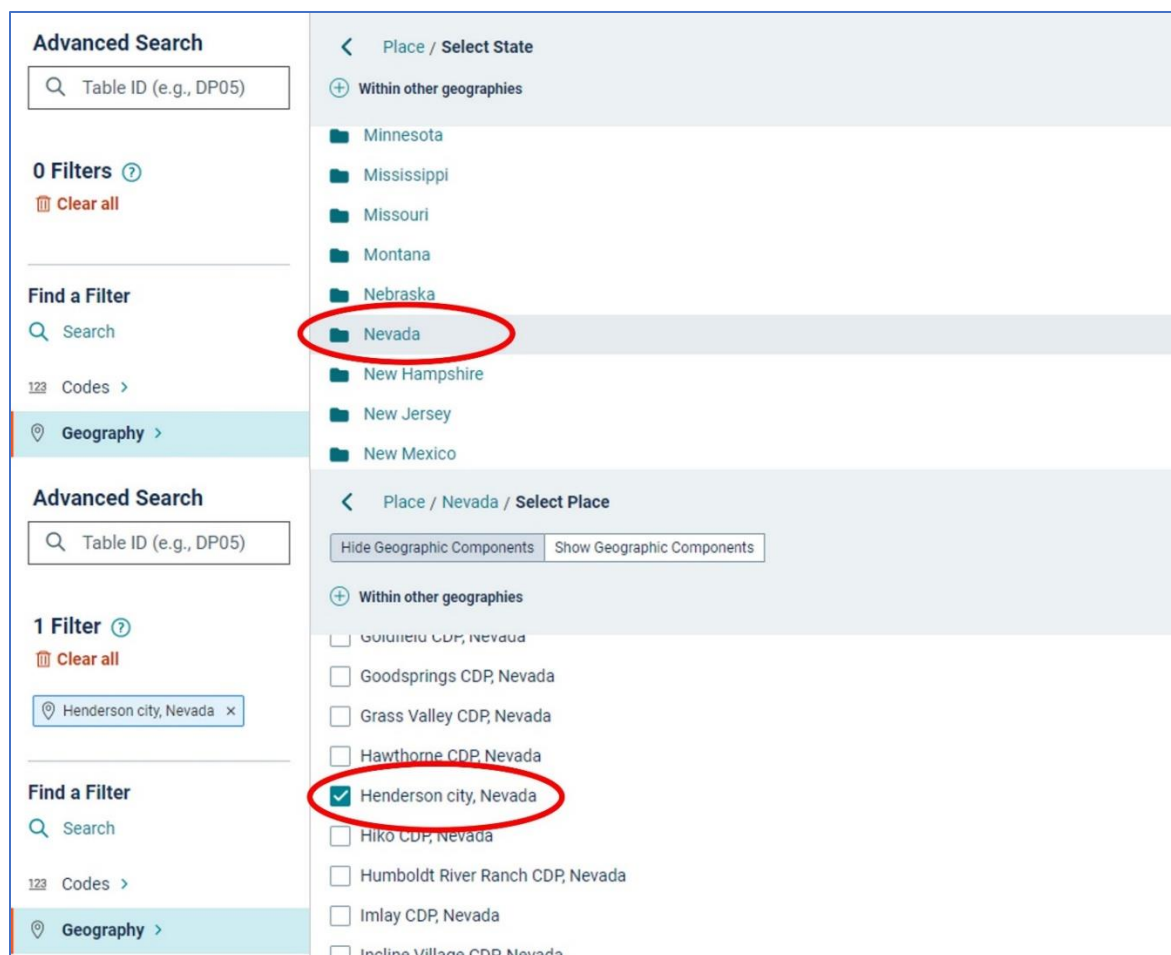
Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

They select "Topics" from the filter menu. From here, they select "Employment" and click on "Commuting." Then they click on "Search" in the lower right corner of the screen (see Figure 6.4).

Figure 6.4

Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

They Select Table S0801: "Commuting Characteristics by Sex" (see Figure 6.5).

Figure 6.5
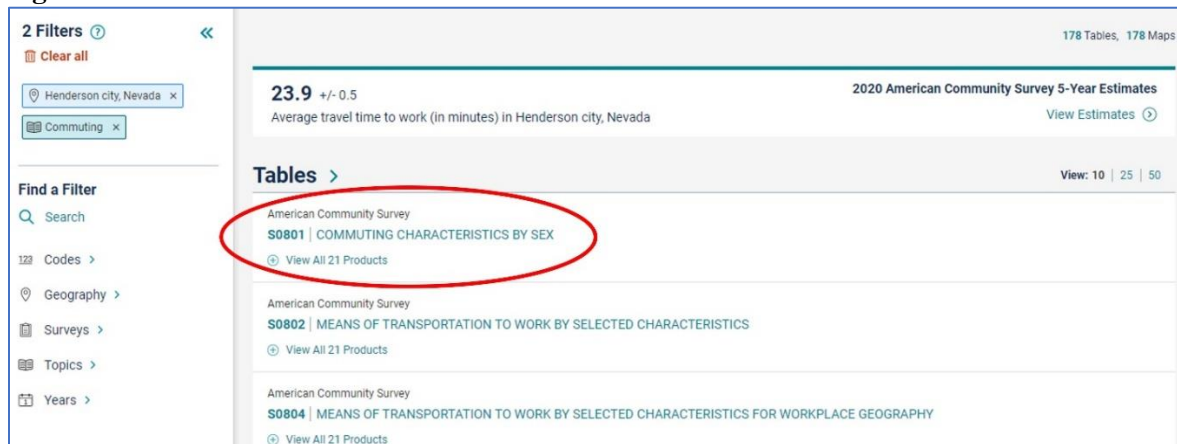
Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

Next, they select the desired year and data product for the comparison. By default, the table displays the most recent ACS data available, but for this example, the city planner first selects the 2015 ACS 1-Year Estimates Subject Tables from the drop-down menu (see Figure 6.6).
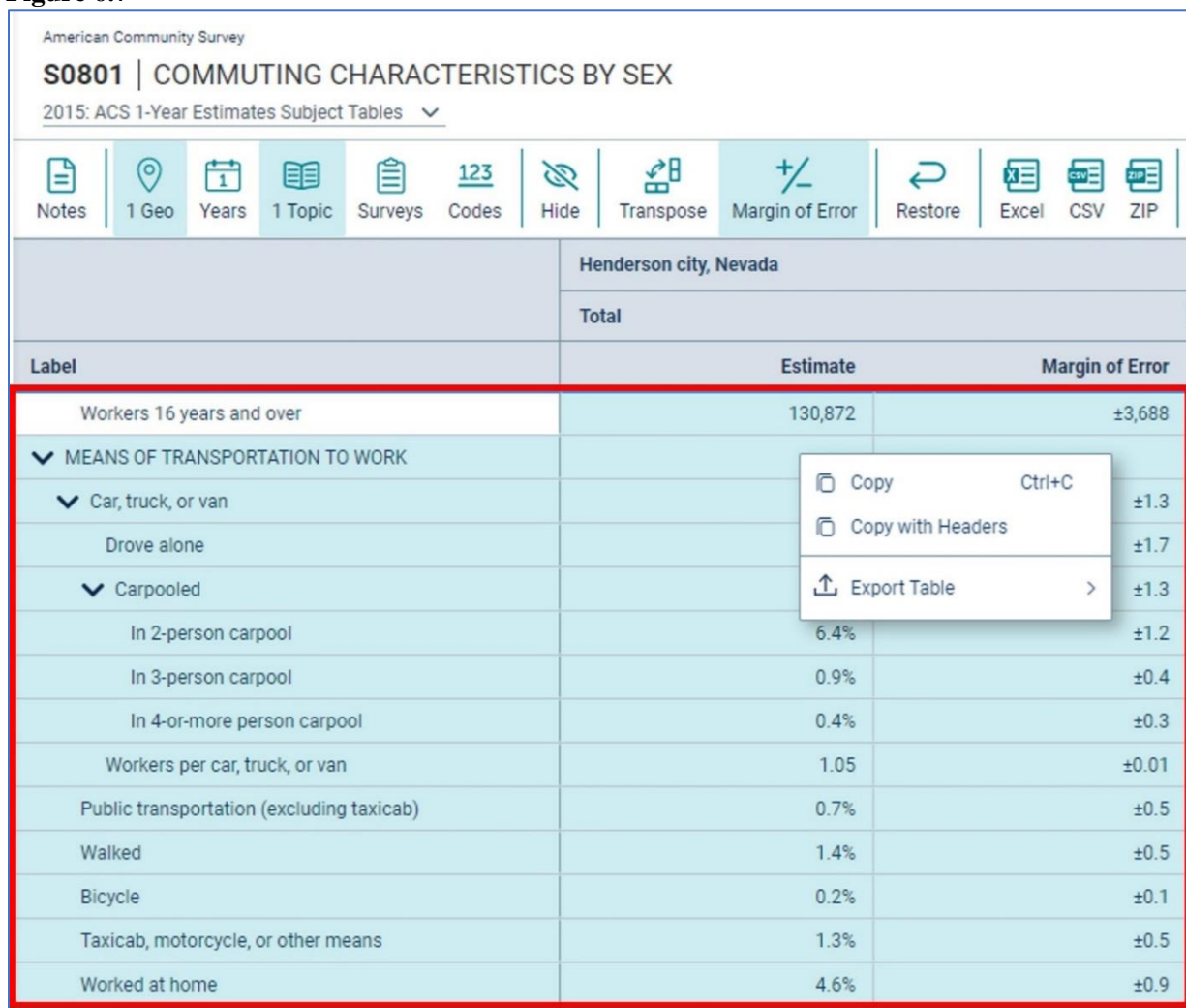
Figure 6.6
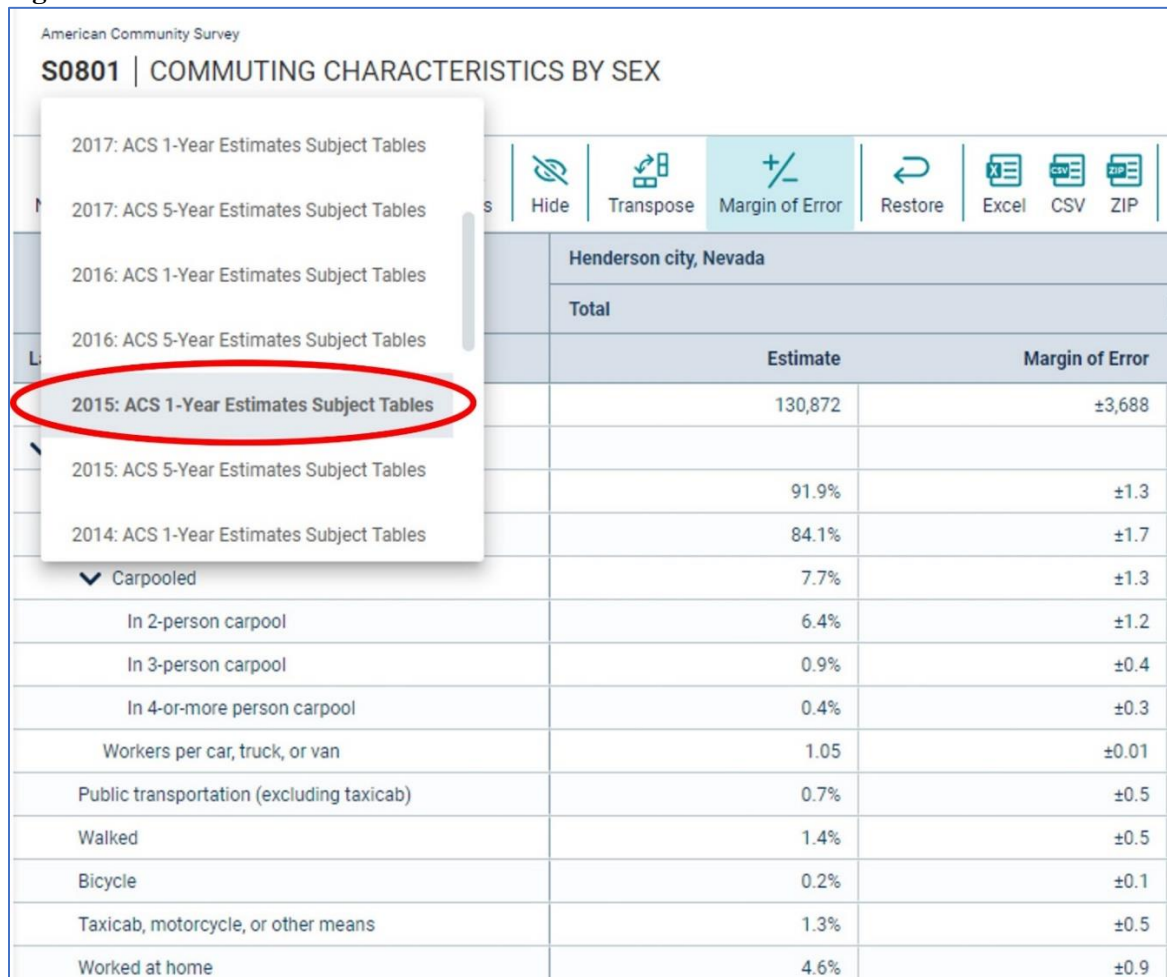
Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

Once the data are displayed in the table, they need to copy and paste the desired row labels, estimates, and margins of error into the Statistical Testing for Two Estimates tool. For this example, the city planner selects the labels and data for the total number of workers 16 years and over, means of transportation to work, and mean travel time to work. To copy cells, they hold a left-click on their mouse to highlight the desired cells. Then they right-click on the highlighted cells and have the option to "Copy" or "Copy with Headers" (see Figure 6.7). Either option will work with the Statistical Testing Tool.

Figure 6.7

Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

Then, they copy the desired labels, estimates, and margins of error into the Statistical Testing Tool (see Figure 6.8).

Figure 6.8

	Label	First Estimate	First Margin of Error (MOE)	Second Estimate	Second Margin of Error (MOE)	Statistically Different?
1	Workers 16 years and over	130872	3688			
2	MEANS OF TRANSPORTATION TO WORK					
3	Car, truck, or van	91.9	1.3			
4	Drove alone	84.1	1.7			
5	Carpooled	7.7	1.3			
6	In 2-person carpool	6.4	1.2			
7	In 3-person carpool	0.9	0.4			
8	In 4-or-more person carpool	0.4	0.3			
9	Workers per car, truck, or van	1.05	0.01			
10	Public transportation (excluding taxicab)	0.7	0.5			
11	Walked	1.4	0.5			
12	Bicycle	0.2	0.1			
13	Taxicab, motorcycle, or other means	1.3	0.5			
14	Worked at home	4.6	0.9			
15	Mean travel time to work (minutes)	23.5	1			

Source: U.S. Census Bureau, Statistical Testing Tool, <www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html>.

They repeat these steps to copy and paste the 2019 ACS estimates and margins of error into the next two columns in the spreadsheet (see Figure 6.9).

Figure 6.9

	Label	First Estimate	First Margin of Error (MOE)	Second Estimate	Second Margin of Error (MOE)	Statistically Different?
1	Workers 16 years and over	130872	3688	148205	4962	Yes
2	MEANS OF TRANSPORTATION TO WORK					
3	Car, truck, or van	91.9	1.3	88	1.4	Yes
4	Drove alone	84.1	1.7	78.9	2.1	Yes
5	Carpooled	7.7	1.3	9.1	1.6	No
6	In 2-person carpool	6.4	1.2	7.6	1.6	No
7	In 3-person carpool	0.9	0.4	0.7	0.3	No
8	In 4-or-more person carpool	0.4	0.3	0.8	0.4	No
9	Workers per car, truck, or van	1.05	0.01	1.06	0.01	No
10	Public transportation (excluding taxicab)	0.7	0.5	1.5	0.6	Yes
11	Walked	1.4	0.5	1.2	0.6	No
12	Bicycle	0.2	0.1	0.2	0.1	No
13	Taxicab, motorcycle, or other means	1.3	0.5	1.8	0.6	No
14	Worked at home	4.6	0.9	7.4	1.3	Yes
15	Mean travel time to work (minutes)	23.5	1	26.5	1.6	Yes
16						

Source: U.S. Census Bureau, Statistical Testing Tool, <www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html>.

The "Statistically Different?" Column displays the results from the statistical test, comparing the 2015 and 2019 commuting estimates for Henderson, NV. The results show that the number of workers ages 16 and over increased. The share of workers who traveled by car, truck, or van decreased, as well the share who drove alone. While the percent of workers who traveled by carpool appears to have increased from 7.7% in 2015 to 9.1% in 2019, this difference is not statistically significant. The share who traveled by public transportation increased, as did the share who worked from home. The mean travel time to work also increased. There were no statistically significant differences between the remaining estimates.

## 7. Using the Statistical Testing Tool to Compare Multiple Estimates

- **Example**: A state policymaker in Pennsylvania is comparing child poverty rates across counties and wants to understand whether the rates are statistically different.

- **Data**: 2016-2020 ACS 5-year estimates

- **Table**: S1701: "Poverty Status in the Past 12 Months."

- **Variable**: Percent of children under age 18 below poverty
- **Method of obtaining data**: download zip files

The Statistical Testing for Multiple Estimates tool can be used to compare many groups or geographies to each other. This tool is also useful for comparing ACS estimates that have been ranked from highest to lowest or lowest to highest. For example, a state policymaker in Pennsylvania is interested in understanding how child poverty rates differ across counties in Pennsylvania. These data can be obtained from data.census.gov and then copied into the Statistical Testing for Multiple Estimates tool to conduct statistical testing across all the counties in Pennsylvania.

The policymaker begins by going to data.census.gov and typing "poverty" in the search bar. They select Table S1701: "Poverty Status in the Past 12 Months" (see Figure 7.1). This table provides the number and percent of people in poverty by selected social and demographic characteristics.

Figure 7.1
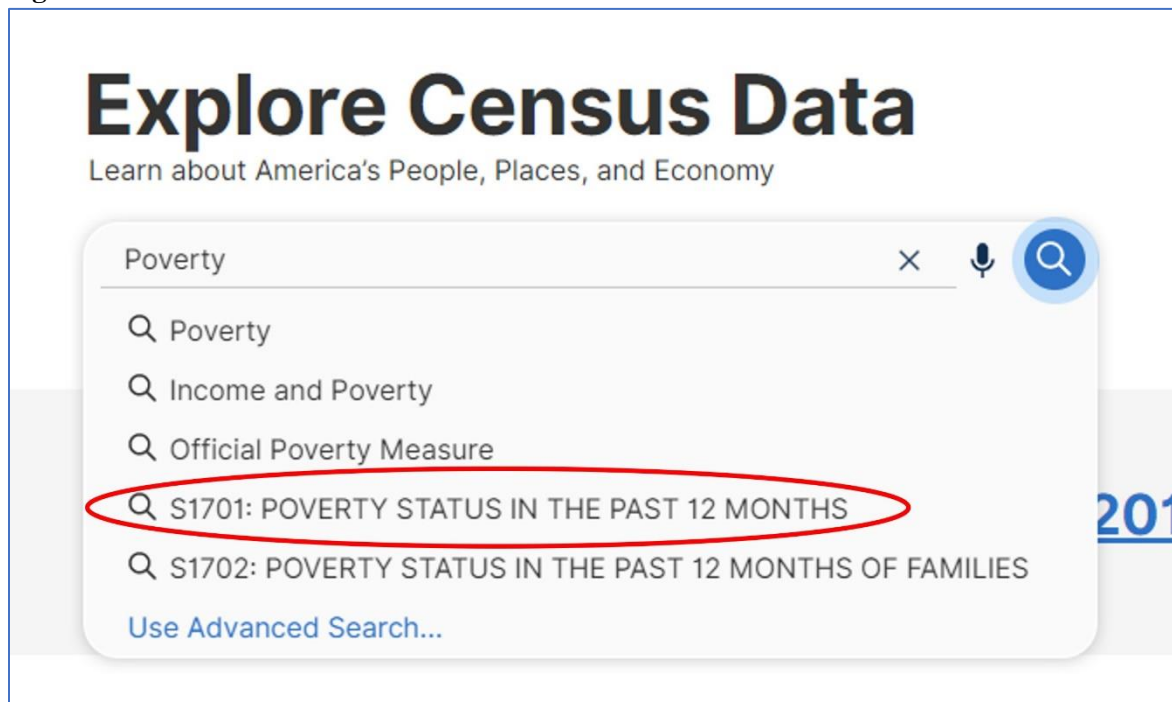
Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

Next, they click on the "Geos" icon (see Figure 7.2).

Figure 7.2

American Community Survey

S1701 | POVERTY STATUS IN THE PAST 12 MONTHS

2020: ACS 5-Year Estimates Subject Tables ⌄

Notes	Geos	Years	Topics	Surveys	Codes	Hide	Transpose	Margin of Error	Restore	Excel	CSV	ZIP	Pri
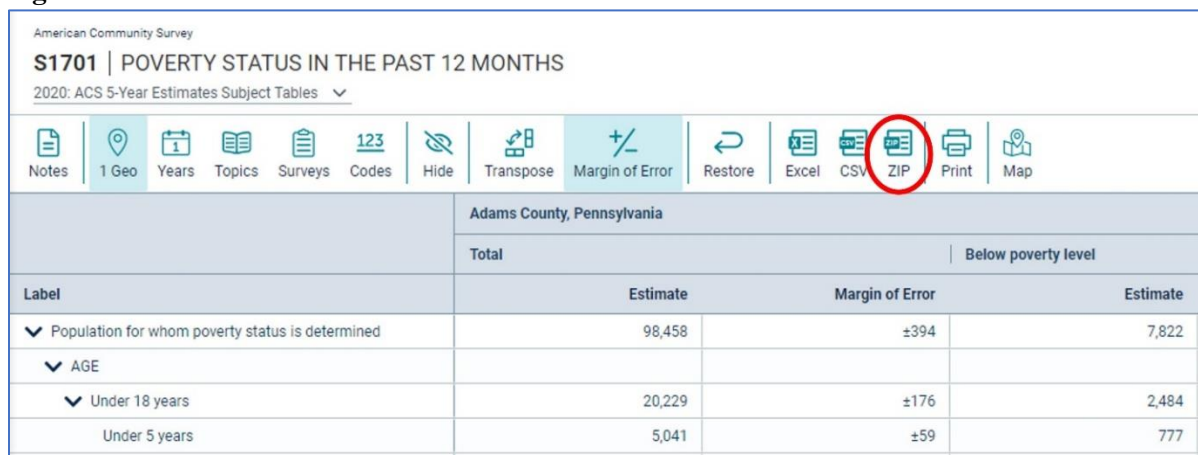
	United States	
	Total	
Label	Estimate	Margin of Error
⌄ Population for whom poverty status is determined	318,564,128	±12,307
⌄ AGE		
⌄ Under 18 years	72,065,774	±11,957
Under 5 years	19,294,872	±7,365
5 to 17 years	52,770,902	±8,534
Related children of householder under 18 years	71,749,414	±13,609

Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

They click on "County," select the state of Pennsylvania, and then select "All counties within Pennsylvania." Then, they click on the "X" in the upper right corner to return to the results page (see Figures 7.3 and 7.4).

Figure 7.3

Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

Figure 7.4

Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

The estimates displayed in the table are from the 2016-2020 ACS 5-year data—the most recent data available at the time. To obtain these estimates in an easy-to-use format for the Statistical Testing for Multiple Estimates tab, the policymaker downloads the data using the "Zip" option (see Figure 7.5).
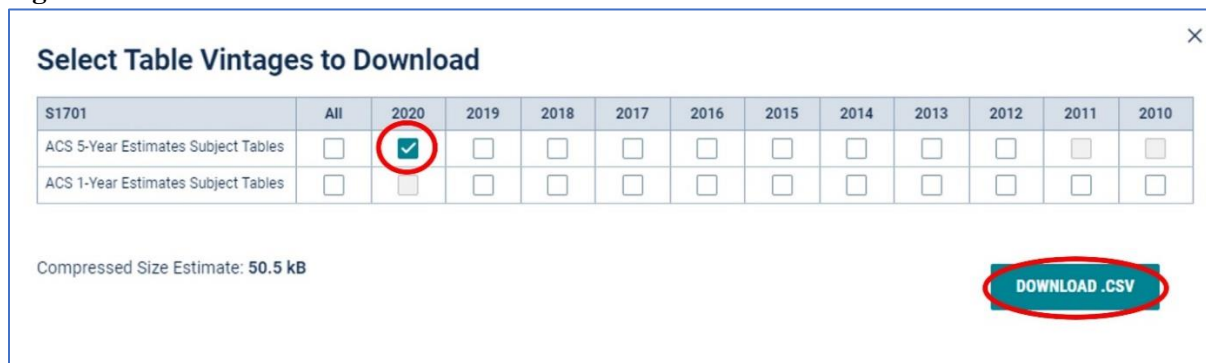
Figure 7.5

Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

They mark the box labeled "2020 ACS 5-Year Estimates Subject Tables" and then click on "Download .CSV" (see Figure 7.6).

Figure 7.6

Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

They open the folder on their computer with the zipped data files. The data are stored in the file with "ACSST5Y2020.S1701_data_with_overlays" in the title (see Figure 7.7). They open this file.

Figure 7.7
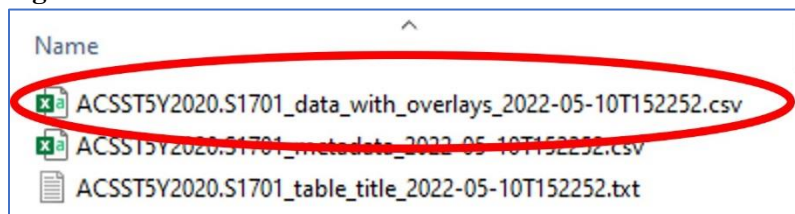
Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

Next, they find the desired estimates from the downloaded data. They select the columns for the percent estimate and the associated margins of error for children living below the federal poverty line and hide or delete the other columns.

To make the results easier to read, they sort the data from lowest poverty rate to highest poverty rate (see Figures 7.8 and 7.9).

Figure 7.8

Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

Figure 7.9

	A	B	C	D	
1	GEO_ID	NAME	S1701_C03_002E	S1701_C03_002M	S17
2	id	Geographic Area Name	Estimate!!Percent below poverty level!!Population for whom poverty status is determined!!AGE!!Under 18 years	Margin of Error!!Percent below poverty level!!Population for whom poverty status is determined!!AGE!!Under 18 years	Est
3	0500000US42001	Bucks County, Pennsylvania	6.7	0.8	
4	0500000US42003	Montgomery County, Pennsylvania	6.9	0.8	
5	0500000US42005	Chester County, Pennsylvania	7.4	1	
6	0500000US42007	Butler County, Pennsylvania	7.8	1	
7	0500000US42009	Cumberland County, Pennsylvania	8.9	1.3	
8	0500000US42011	Centre County, Pennsylvania	9.7	1.8	
9	0500000US42013	Washington County, Pennsylvania	11.1	1.4	
10	0500000US42015	Perry County, Pennsylvania	11.7	2	
11	0500000US42017	Union County, Pennsylvania	11.7	3.2	
12	0500000US42019	York County, Pennsylvania	11.9	1.1	
13	0500000US42021	Adams County, Pennsylvania	12.3	1.9	
14	0500000US42023	Lancaster County, Pennsylvania	12.3	1.5	
15	0500000US42025	Westmoreland County, Pennsylvania	12.6	1.4	
16	0500000US42027	Northampton County, Pennsylvania	13.3	1.7	
17	0500000US42029	Delaware County, Pennsylvania	13.5	1.1	
18	0500000US42031	Bedford County, Pennsylvania	13.7	2.3	
19	0500000US42033	Wayne County, Pennsylvania	13.7	3.2	
20	0500000US42035	Franklin County, Pennsylvania	13.8	2.8	
21	0500000US42037	Fulton County, Pennsylvania	14	2.5	

Note: This figure displays only a portion of the columns and rows in the spreadsheet.
Source: U.S. Census Bureau, data.census.gov, <https://data.census.gov/cedsci/>.

They copy and paste the Geography Area Names into the column named "Label" in the Statistical Testing for Multiple Estimates tool (see Figure 7.10). They adjust the height and width of the rows and columns as needed to make the labels easier to read.

Figure 7.10

Note: This figure displays only a portion of the columns and rows in the spreadsheet.
Source: U.S. Census Bureau, Statistical Testing Tool, <www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html>.

Next, they select the desired estimates from the downloaded data. They select the columns for the percent estimate and the associated margins of error for children living below the federal poverty line and paste the data into the Statistical Testing for Multiple Estimates tool in the "Estimates" and "Margin of Error (MOE)" columns (see Figure 7.11).

Figure 7.11

Note: This figure displays only a portion of the columns and rows in the spreadsheet.
Source: U.S. Census Bureau, Statistical Testing Tool, <www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html>.

The statistical testing results appear in the columns on the right side of the table. Results from the tests can be found by reading across a row or down a column for a particular county. For example, the first cell

for Bucks County is shaded gray and marked with an X; this is the marker for comparing the county to itself/not a comparison. The next three cells to the right are shaded red with "No." This means that Butler County, Chester County, and Montgomery County all have child poverty rates that are not statistically different from Bucks County's rate. Continuing across the row for Bucks County, the remaining cells are "Yes," which means that the child poverty rate in Bucks County is statistically different than the child poverty rate in the remaining Pennsylvania counties at the 90% confidence level.

8. Additional Options

You can change the confidence level by scrolling to the bottom of the spreadsheet and changing the value in the "Parameters" column to the desired level (e.g., 95 or 99) (see Figure 8.1).

Figure 8.1

Description	Parameters	Notes
Constant to convert MOE to SE (default = 1.645)	1.645	For 95% conf. level MOE, change to 1.96
Confidence level (default = 90)	90	For 95% conf. level, change this to 95.
Cutoff value for statistical testing	1.645	Automatically changes based on conf. level.

Source: U.S. Census Bureau, Statistical Testing Tool, <www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html>.

Changing the confidence level to 99 shows that at this higher level of confidence, the child poverty rate in Bucks County is no longer statistically different than the rates in Centre, Cumberland, Forest, Sullivan, or Union Counties.

Figure 8.2

Estimate	Margin of Error (MOE)	Label
Estimate!!Percent below poverty level!!Population for whom poverty status is determined!!AGE!!Under 18 years	Margin of Error!!Percent below poverty level!!Population for whom poverty status is determined!!AGE!!Under 18 years	Geographic Area Name
6.7	0.8	Bucks County, Pennsylvania
6.9	0.8	Montgomery County, Pennsylvania
7.4	1	Chester County, Pennsylvania
7.8	1	Butler County, Pennsylvania
8.9	1.3	Cumberland County, Pennsylvania
9.7	1.8	Centre County, Pennsylvania
11.1	1.4	Washington County, Pennsylvania
11.7	2	Perry County, Pennsylvania
11.7	3.2	Union County, Pennsylvania
11.9	1.1	York County, Pennsylvania
12.3	1.9	Adams County, Pennsylvania
12.3	1.5	Lancaster County, Pennsylvania
12.6	1.4	Westmoreland County, Pennsylvania
13.3	1.7	Northampton County, Pennsylvania
13.5	1.1	Delaware County, Pennsylvania
13.7	2.3	Bedford County, Pennsylvania
13.7	3.2	Wayne County, Pennsylvania

Column headers (1–13): Geographic Area Name; Bucks County, Pennsylvania; Montgomery County, Pennsylvania; Chester County, Pennsylvania; Butler County, Pennsylvania; Cumberland County, Pennsylvania; Centre County, Pennsylvania; Washington County, Pennsylvania; Perry County, Pennsylvania; Union County, Pennsylvania; York County, Pennsylvania; Adams County, Pennsylvania; Lancaster County, Pennsylvania.

Note: This figure displays only a portion of the columns and rows in the spreadsheet.
Source: U.S. Census Bureau, Statistical Testing Tool <www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html>.

9. For More Information

Understanding and Using American Community Survey Data: What All Data Users Need to Know
www.census.gov/programs-surveys/acs/library/handbooks/general.html
This handbook provides an overview of the ACS to help data users understand the basics of the survey, how the data can be used, how to judge the accuracy of ACS estimates, and how to access ACS data.

Comparing ACS Data
www.census.gov/programs-surveys/acs/guidance/comparing-acs-data.html
This Web page provides guidance on making valid comparisons between the latest ACS data and ACS data from the previous year, the 2000 Census, and the 2010 Census.

Sample Size and Data Quality
www.census.gov/acs/www/methodology/sample-size-and-data-quality
This Web page describes the steps the Census Bureau takes to ensure that ACS data are accurate and reliable. It also includes several measures of ACS data quality for the nation and states.

Code Lists, Definitions, and Accuracy
www.census.gov/programs-surveys/acs/technical-documentation/code-lists.html
View the detailed codes and definitions for variables, statistical testing, and an explanation of sample design, methodology, and accuracy for the ACS.