August 2020

Hi. I am Alexandra Barker. There are many Census Bureau data products that I find useful. But microdata is my go-to product when I need an estimate that can’t be found within any of our pre-tabulated data products. Microdata are un-tabulated records about individual people in housing units, and you can use them to create custom tabulations. So I invited my colleague, Jessica Barnett to introduce you to microdata and help you understand when you should consider using it. Let’s check it out.

Hi, I’m Jessica. Have you ever needed a data estimate and found that it’s not published? Or do you need more detailed estimates than the pre-tabulated tables provide? You may want to consider using Microdata Access, which allows you to pick and choose your own variables and create custom estimates. In this data gem, I’ll give you more information on Microdata Access and when you should use it.

Microdata Access uses the Public Use Microdata, or PUMS, which are the individual responses from a survey or program, to create estimates. Let’s start with defining exactly what that means. Public use microdata takes all of the responses collected, removes personally identifiable information, and puts additional edits in place specifically to protect the confidentiality of respondents. PUMS data are accessible to the general public in a variety of formats for free through the API, FTP sites, and data.census.gov/mdat. In order to aggregate the individual responses and make a table from the microdata, you have to tabulate the data and apply weights.

You don’t have to use microdata access to create tables from PUMS data, but it’s an easy-to-use tool that allows you to create these tabulations from the microdata without having your own statistical software or doing your own programming. The tool helps you find the right variables and even automatically pulls the appropriate weights for you. This tool really helps everyday people create custom tabulations.

So when would you want to use Microdata Access instead of using the pre-tabulated tables? We recommend that you consider using Microdata Access whenever your data needs are not supported by the pre-tabulated tables.
If you’ve found that the pre-tabulated tables do not provide the detailed estimates that you need, you may want to give Microdata Access a shot. Since it allows you to create custom tables, you can get a lot more detail by topic.

Conversely, if the estimate that you need can be found in the pre-tabulated tables, we recommend that you use those estimates instead of creating your own estimates using Microdata Access. The main reason for this is that the estimates provided in the pre-tabulated tables are more accurate than those found using Microdata Access.

You should also take into consideration the geographic level that you need the data for. The pre-tabulated tables found in data.census.gov can be found for a wide range of geographies. But the available geographies in public use microdata are very limited; typically limited to nation, state and, for some surveys, PUMAs, which are areas of around 100,000 people.

Another thing to consider is the complexity in creating the estimates. With the pre-tabulated tables, the correct variables have already been determined and everything is defined for you. With Microdata Access, it is necessary for you to know the variables that you need, which can make the overall process more challenging. Variable definitions are available in the tool to help you navigate the most appropriate variable for your needs, but a good tip is to consult the survey or program pages for data dictionaries.

In summary, you may find Microdata Access to be a great alternative to using the pre-tabulated tables if you need an estimate that is not found in those and you only need data for a high-level geography.

I hope this is helpful! Thank you.

Thank you, Jessica, for introducing us to microdata. I hope you found this data gem useful. For more videos like this please visit Census.gov/Academy, and subscribe. Thank you.