How will Differential Privacy affect data in the 2020 Census?

The Census Bureau has changed its privacy protection method to Differential Privacy. Instead of other methods formerly used, such as swapping, this method uses an algorithm called the Top-Down Algorithm (TDA) to randomly inject error into data, making it harder for computers to identify personal data. Almost all data in the redistricting product is subject to the TDA. Much of it can still be considered reliable, while other data points should be used with caution. Below is a guide produced by the Maine State Data Center for using data from the 2020 Redistricting Data product.

Green Light 🟢

This data is highly reliable. It has low levels of error and can be used as usual.

Green light variables:

- State-level total population, all other demographic & housing variables
- County-level total population, race & ethnicity except for Native Hawaiian and Pacific Islander, all other housing & Group Quarters variables
- Total population in medium and large cities & towns (>500 population)
- Total Population in Census Tracts
- Number and type of occupied group quarters units at the block level
- Total housing units at the block level
- Any data point at the county or county subdivision/place level with a count of at least 250
- Any data point at the Tract level with a count of at least 500

Yellow Light 🟡

This data is slightly less reliable, but can still be used with caution. Understand that these variables may have moderate error.

Yellow light variables:

- Race and ethnicity in large towns & cities (>5,000 population)
- Total population in small towns & cities
- Any cell at the county or county subdivision level that has a count between 60-250
- Any data point at the Tract level with a count between 250-500

Red Light 🔴

This data is subject to high levels of error. Use these variables with extreme caution or consider aggregating them together to mitigate error.

Red light variables:

- Block-level data
- Race and ethnicity data in small towns
- Total population in county subdivisions with fewer than 60 population
- Any statistic that is divided across tables; for example, persons per household
- Any data point at the county subdivision level that has a count of less than 60
- Any data point at the tract level with a count of less than 250