# Methodology

To generate estimates of the number of deaths attributed to COVID-19, researchers used the available county-level figures (collected and reported by the Center for Systems Science and Engineering at John Hopkins University).

In cases where a county is divided across multiple congressional districts, the deaths in that county were proportionally assigned to the congressional districts it falls into based on the share of the county's population that falls into each district (using the Geocorr tool from the Missouri Census Data Center).

The John Hopkins University CSSE groups deaths from the five counties making up New York City (Kings, Queens, Bronx, Richmond and New York counties). The City has been tracking cases by borough, and their data was used to fill in the missing data from the Johns Hopkins data.

Similarly, the Johns Hopkins data does not break out Rhode Island counties separately. Data collected by The New York Times was used to fill in the missing John Hopkins data in Rhode Island.

Data was pulled from GitHub repositories of each organization on May 21, 2020, and reflects reported coronavirus deaths through May 20.

The profile information for congressional districts was drawn mostly from the 2014-2018 American Community Survey estimates published by the Census Bureau (accessed through the tidycensus package in R on May 21). The estimates of urban and dense suburban population by congressional district come from CityLab's estimates (accessed on May 21).