

Gwen
405196467
1/28/2021

Assignment Two: Mapping COVID-19 Data Distributions

Maps are found at the end of this document.

Part One: Covid Mortality Rate Analysis

After completing part one of this assignment, it is very clear to me that the type of data classification method we select greatly affects readers' interpretation of the presented data. This effect was noticeably more present in the county-level maps, so I will primarily talk about them, but affected data at the state level as well. The Equal Interval classification definitely downplays the severity of the global pandemic. Since the data is lopsided, it groups the vast majority of counties into the same category, which is much less informative than other types of classification. Quantile Interval was a much more effective method for representing pandemic data. Each category has an equal amount of data points, which shows a good diversity on the map. However, it still suffers from grouping a large range of numbers into a single category. Natural Breaks is also quite good for representing data - it looks visually similar to Quantile Interval, but groups less into the maximal category. This results in the most affected regions being properly highlighted, while less affected regions are still visible and understandable.

Standard Deviation's diverging color scheme creates a different effect than the other three classification systems: it highlights the states and counties which are outliers in pandemic data. In this case, it specifically highlights states and counties with the highest overall death rates. This can be good depending on what you desire to communicate with these maps, but in my opinion the shift in color downplays the severity of the pandemic in less affected regions. Overall, I think that Natural Breaks (Jenks) is best suited for these datasets, with Quantile Interval at a close second.

Part Two: Covid Case Analysis

State Level Graduated Symbol Map

For my state level graduated symbol map, I decided to use natural breaks (jenks). To be frank, all of the options for classifying graduated symbol data which we talked about in class were inferior to simply using proportional symbols. When representing something as important as a global pandemic, it's important to properly highlight the actual severity in regions. Graduated symbols certainly fail to do this. However, natural breaks (jenks) seemed like the best method of data classification among those graduated options available. Equal Interval is inappropriate for lopsided data, while Quantile suffered from over-normalization of data (hard to tell circles apart, lumped lots together). Natural Breaks' main issue is that it lumps together a large number of disparate data points into a single class, but once again all graduated symbol maps will suffer from this same issue.

County Level Graduated Symbol Map

For my county level graduated symbol map, I decided to use the standard deviation method of data classification. I made this decision because all other methods of data classification which we discussed in lecture content yielded uninformative or inferior results. Quantile distributions resulted in it being very difficult to distinguish any features on the map. As there were lots of circles of similar sizes, it was hard to tell which circle had which size. Equal Interval once again did not make sense, as case data is heavily lopsided. Natural Breaks (Jenks) was the second best, but in this case there were so many unique data points that it did not properly highlight the severity of covid in many counties, grouping many together. Honestly, the Standard Deviation method suffers from the same problem: the largest class ranges from 75,000 cases to over 1,000,000. It's the "best of the worst" options - I believe the superior method in this case is to use proportional symbols, as evidenced by how superior the proportional symbol map I created looks. It's worth noting that although my classes are arranged according to standard deviation, I relabeled them to be indexed by number of cases. This is more intuitive to the reader, who may not have a background in statistics.

Data Sources:

Population Data:

<https://data.census.gov/cedsci/table?t=Populations%20and%20People&g=0100000US.04000.001&tid=ACSST1Y2019.S0101&hidePreview=false>

COVID-19 Data: Provided by professor

County, State shapefiles: TIGER/LINE 2014 dataset and 2018 U.S. Census Data

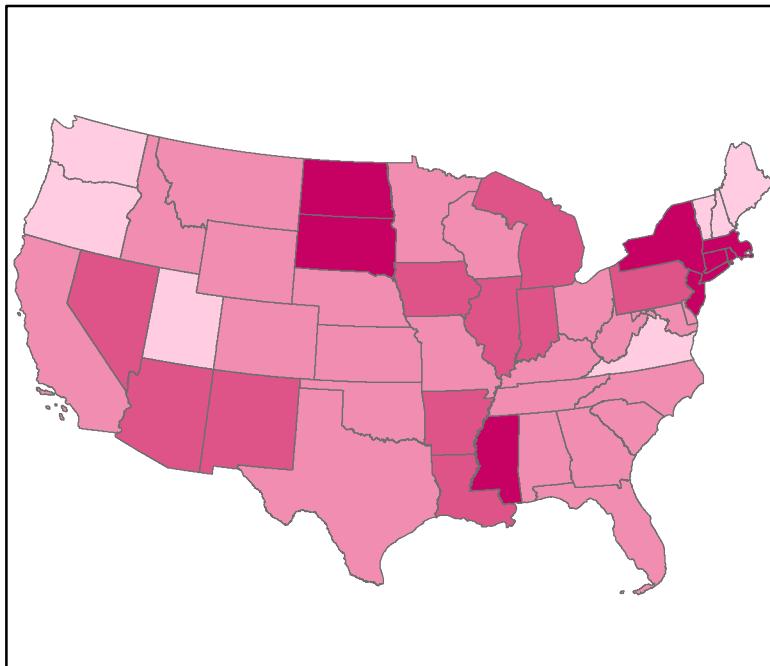
Presidential Election Results:

<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/42MVDX>

Governor Party Alignments: Handmade based off of wikipedia's article:

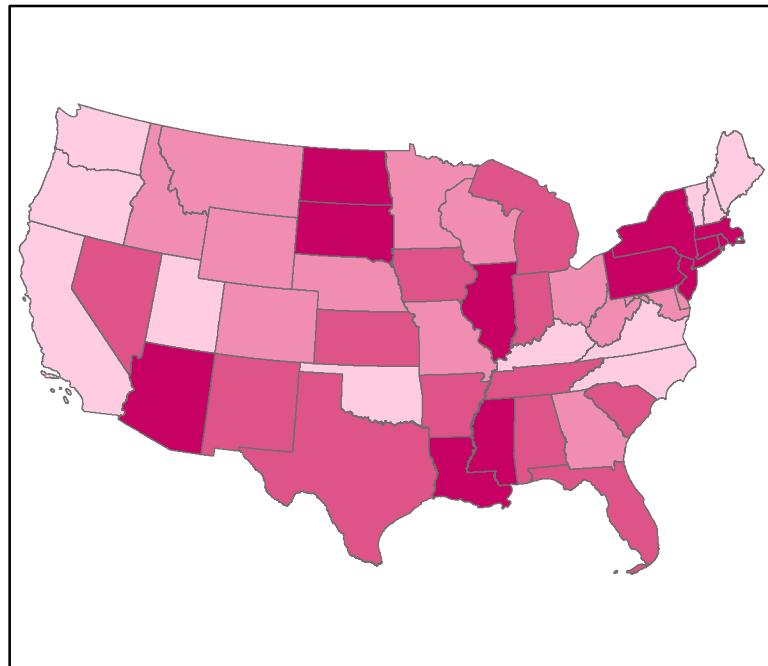
https://en.wikipedia.org/wiki/List_of_current_United_States_governors

Visualizing COVID-19 Death Rates by State



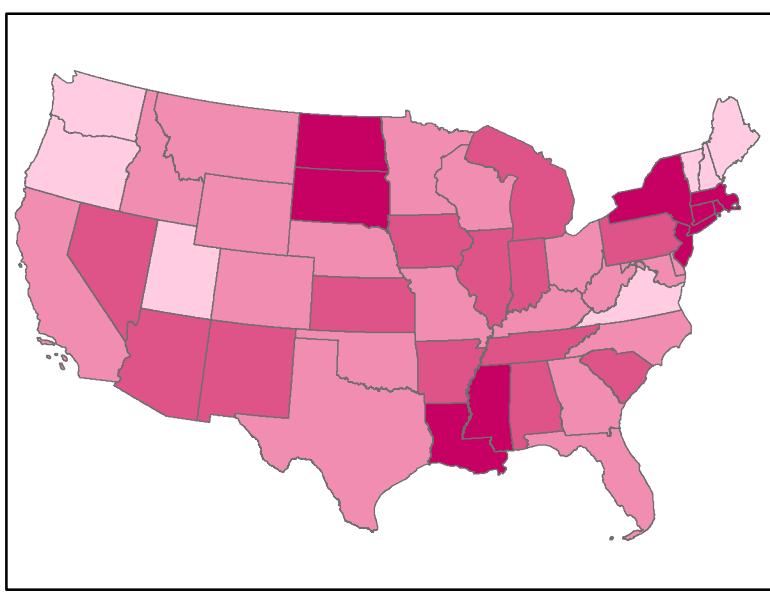
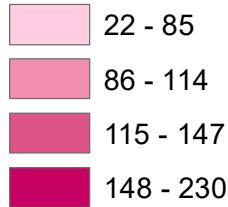
Equal Interval

Deaths per 100,000 People



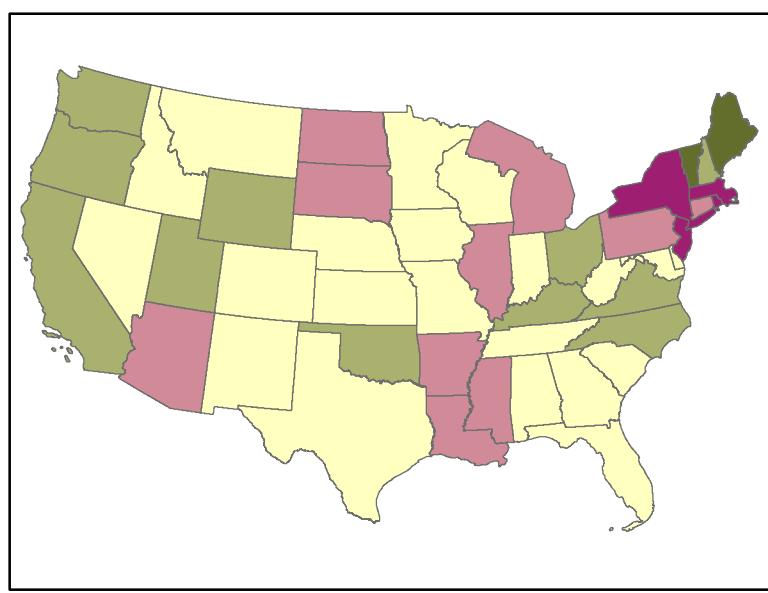
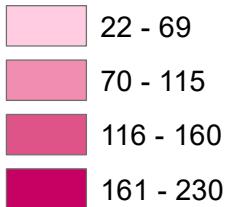
Quantile Interval

Deaths per 100,000 People



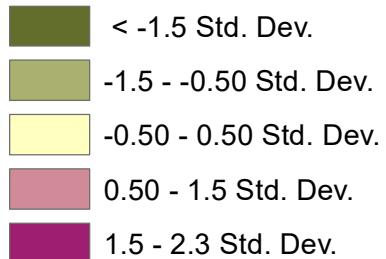
Natural Breaks (Jenks)

Deaths per 100,000 People

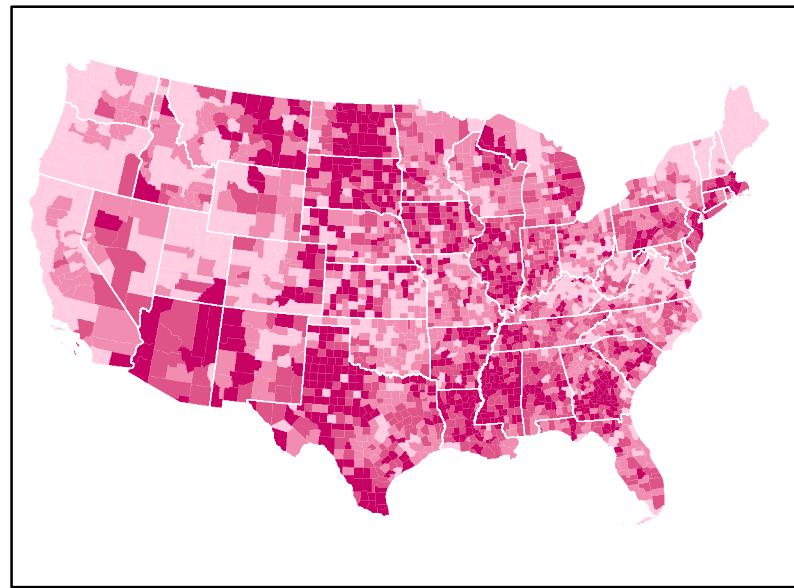
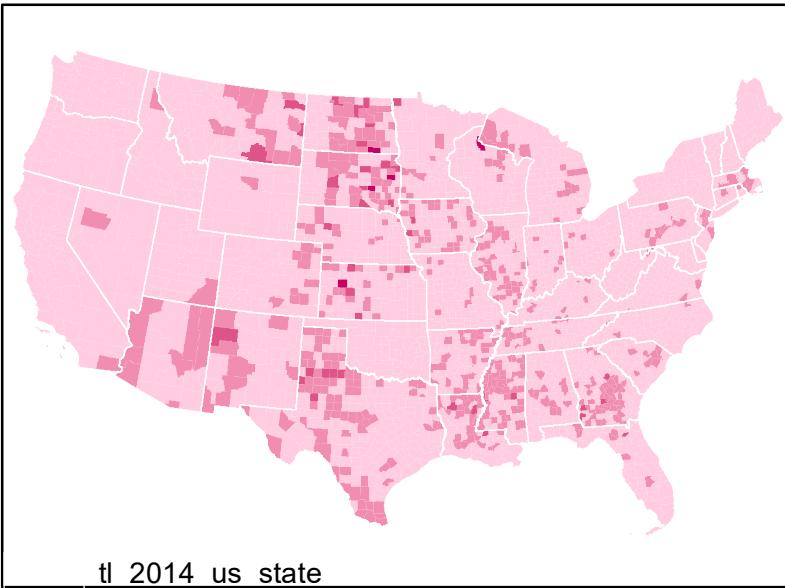


Standard Deviation

Deaths per 100,000 People

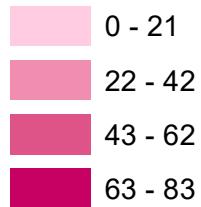


Visualizing COVID-19 Death Rates by County



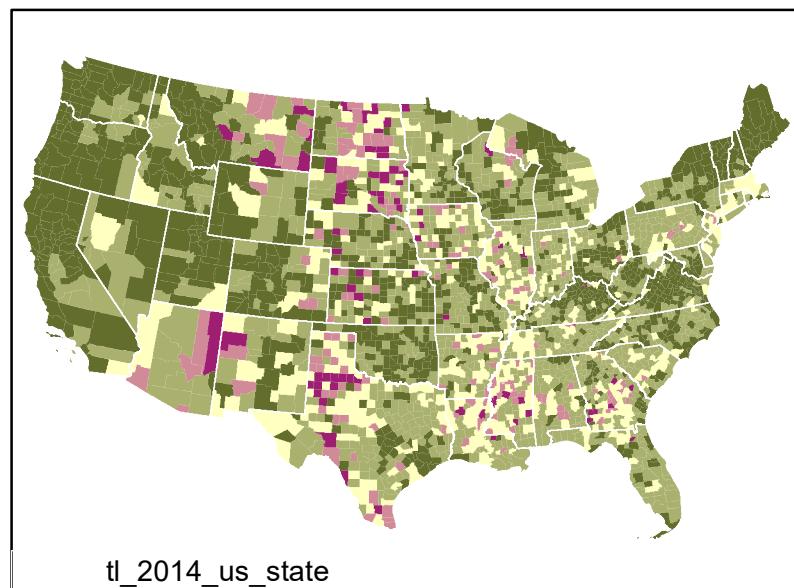
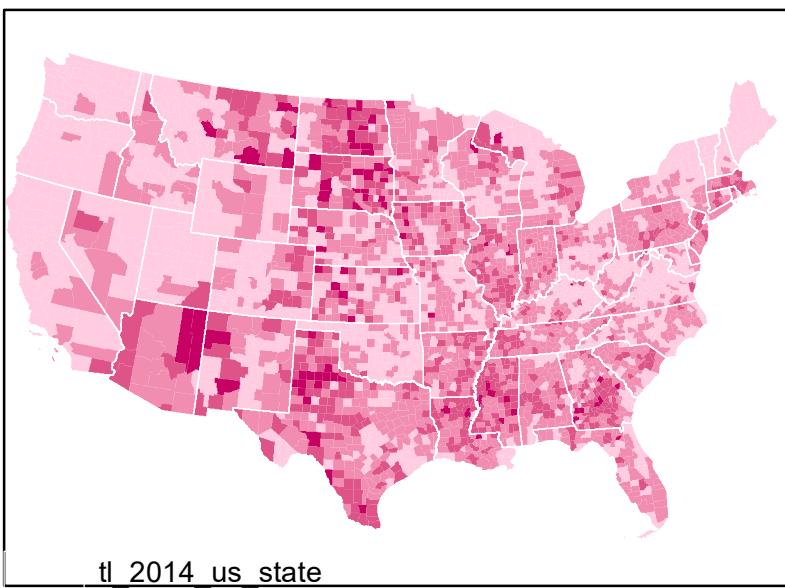
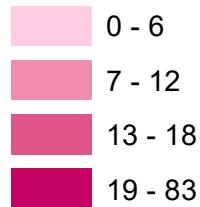
Equal Interval

Deaths per 10,000 People



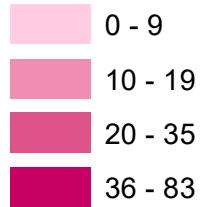
Quantile Interval

Deaths per 10,000 People



Natural Breaks (Jenks)

Deaths per 10,000 People

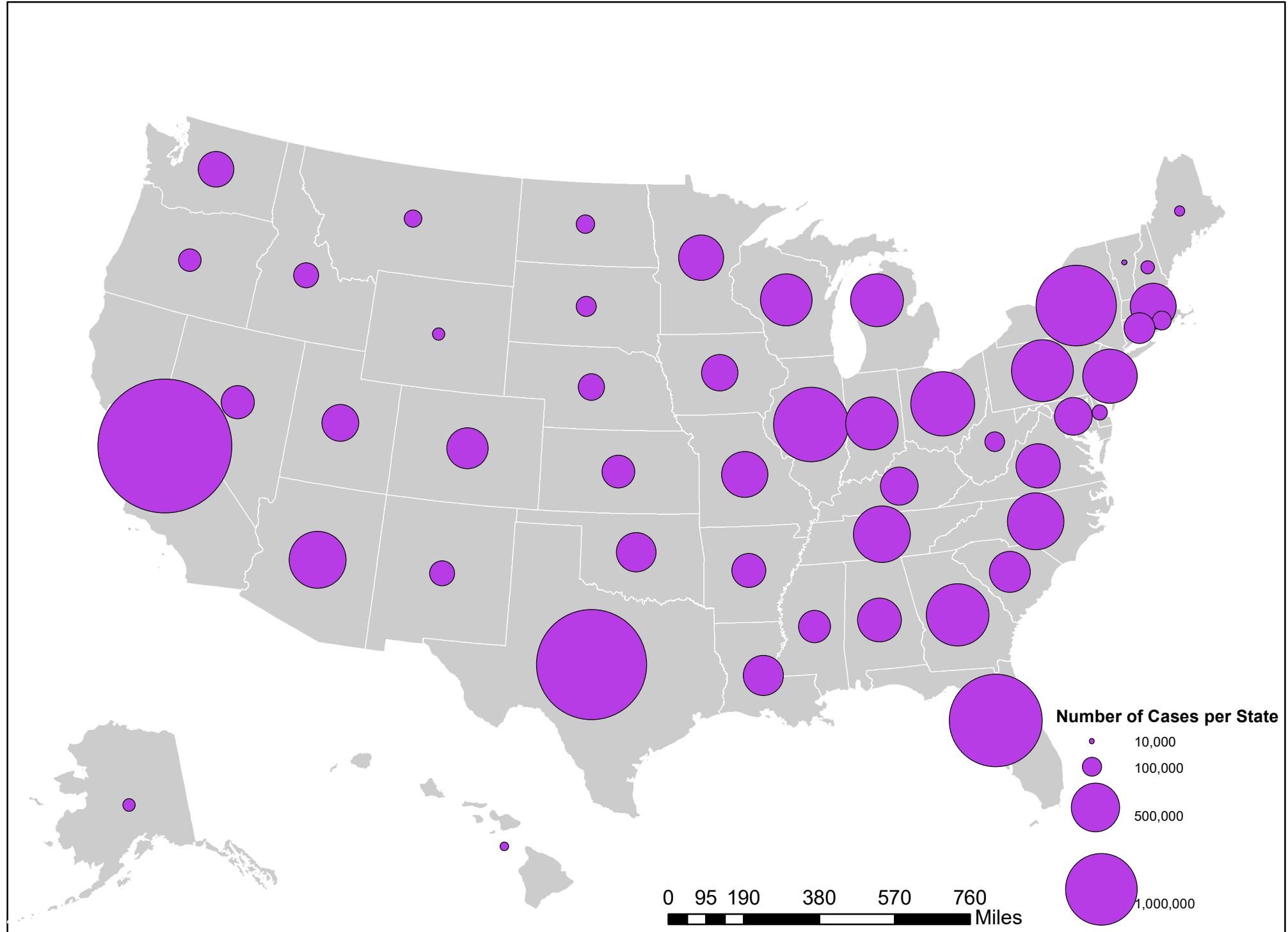


Standard Deviation

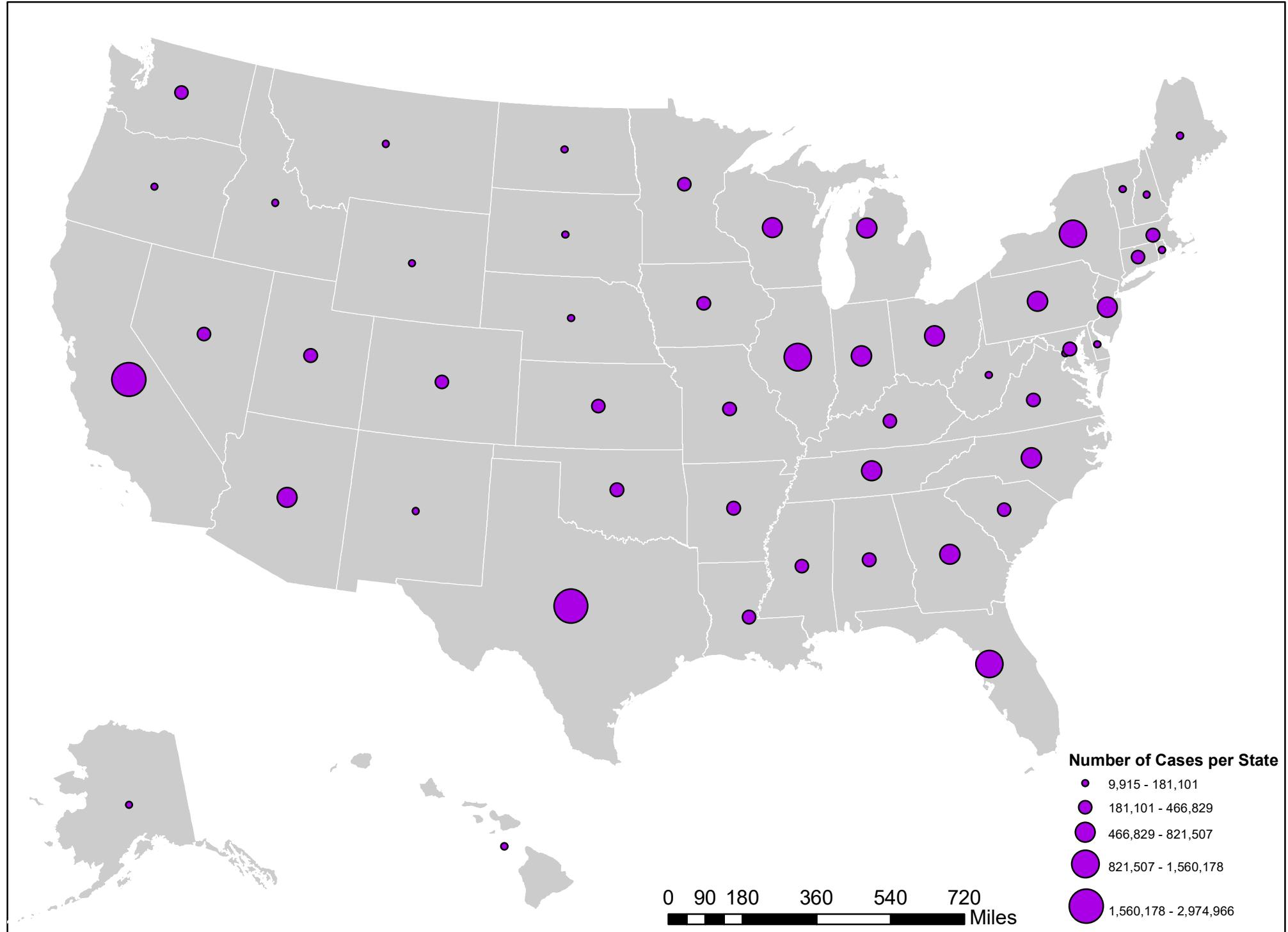
Deaths per 10,000 People



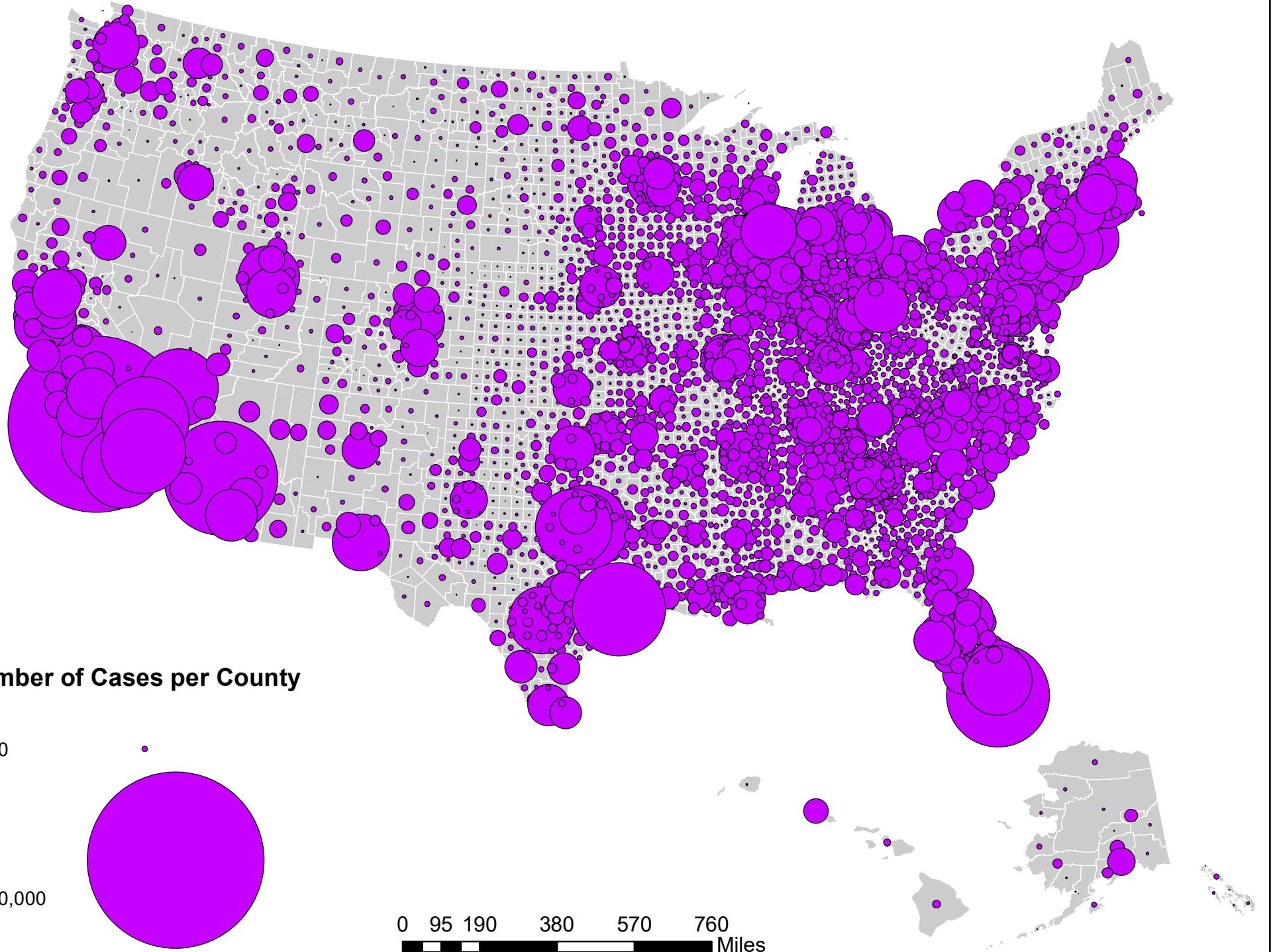
Number of COVID-19 Cases by State (Proportional Symbols)



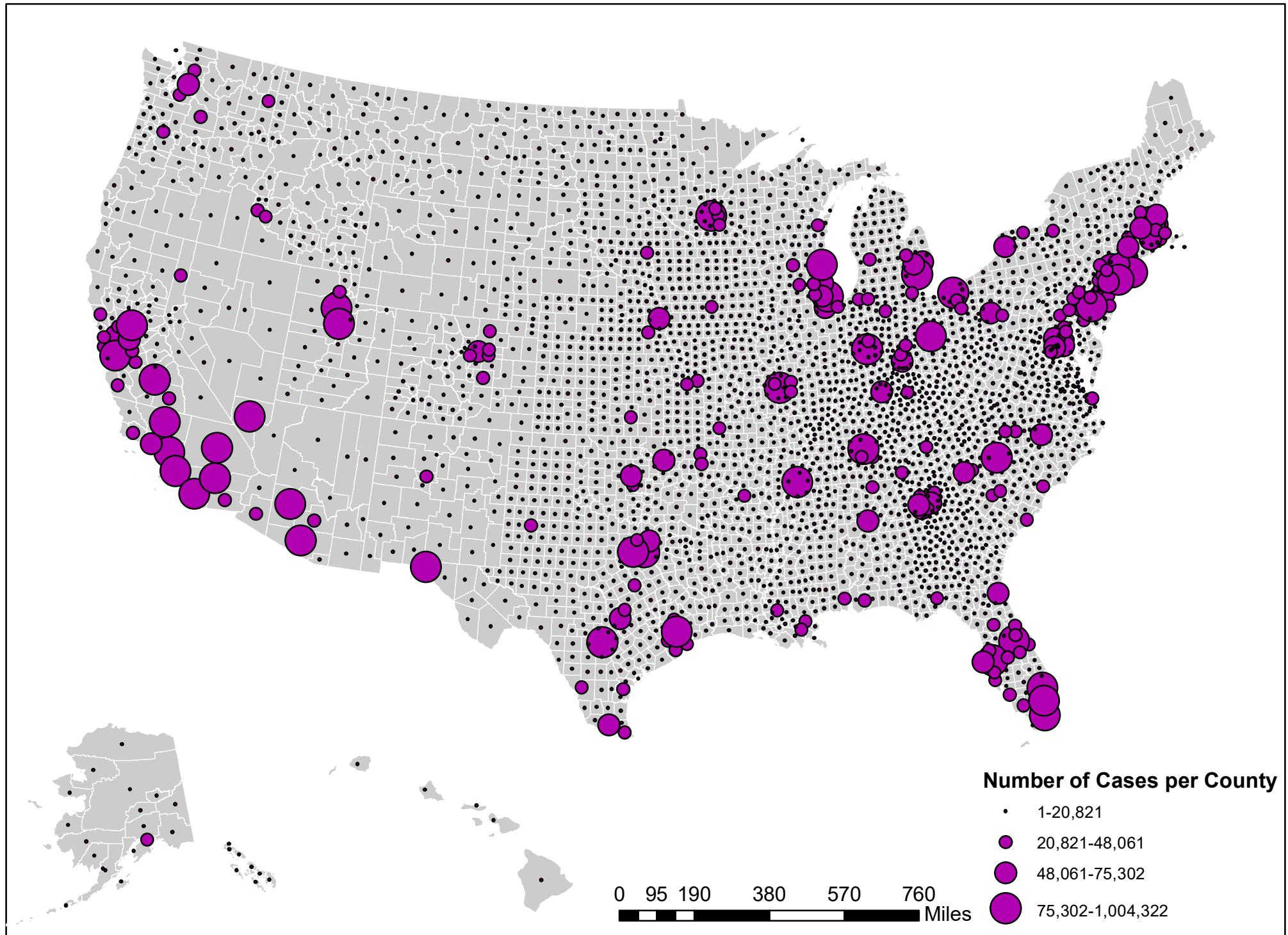
Number of COVID-19 Cases by State (Graduated Symbols)



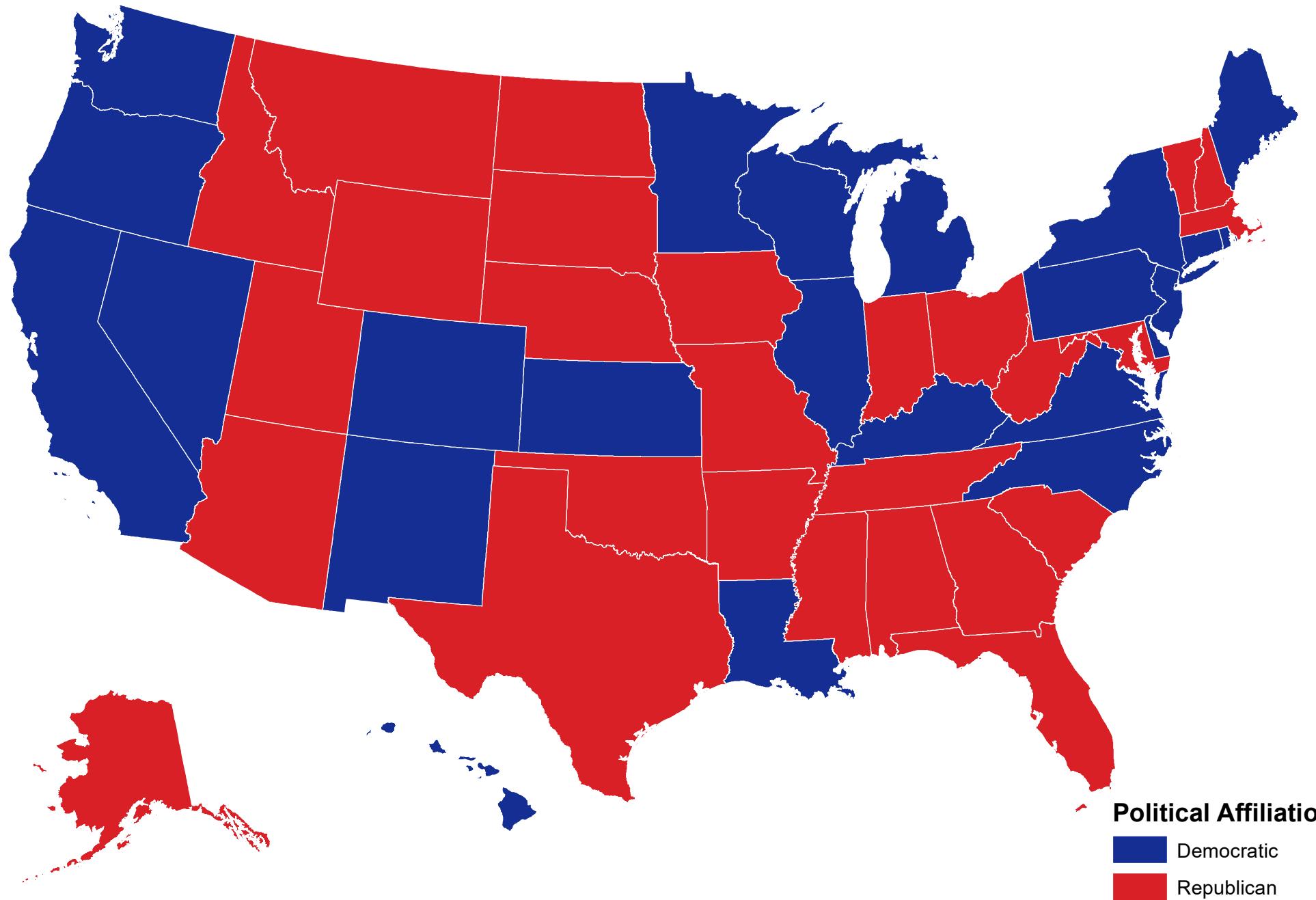
Number of COVID-19 Cases by County (Proportional Symbols)



Number of COVID-19 Cases by County (Graduated Symbols)



Political Affiliation of U.S. Governors by State



2020 Election Candidates' Margin of Victory by State

