Project Title: United States COVID-19 Cases by State

Which 3 states had the highest and lowest numbers of total cases of COVID-19 in the year 2021?

The 3 states with the highest number of total cases in the year 2021 were California, Illinois and New Jersey. However, when we analyzed the same data relative to state population, the 3 states with the highest percentage of total cases of COVID-19 were Nebraska, New Jersey and Oklahoma.

The 3 states with the lowest number of total cases of COVID-19 in the year 2021 were Wyoming, Maine and Delaware. However, when we analyzed the same data relative to state population, the 3 states with the lowest number of total cases of COVID-19 relative to population were California, New York and Maine.
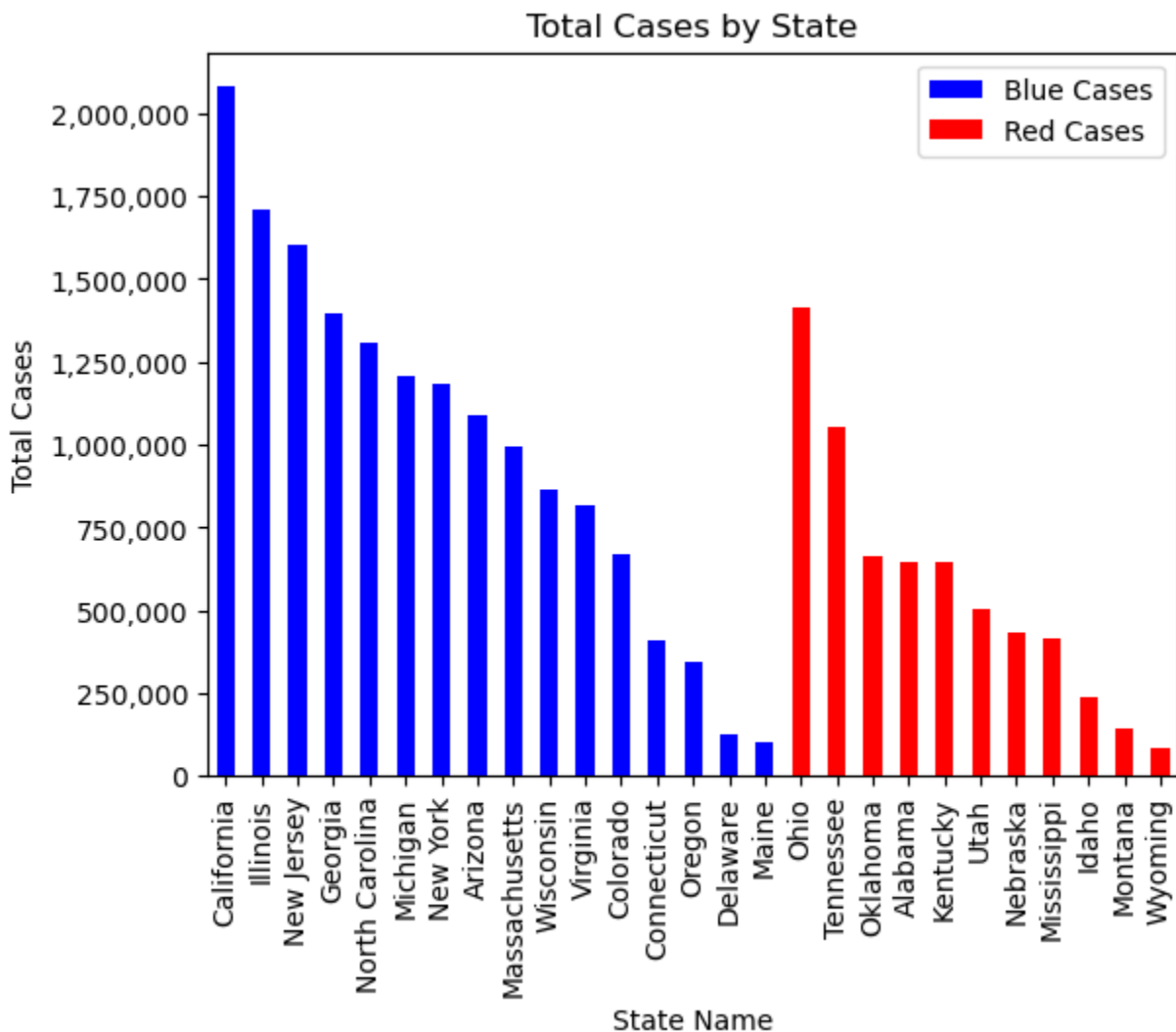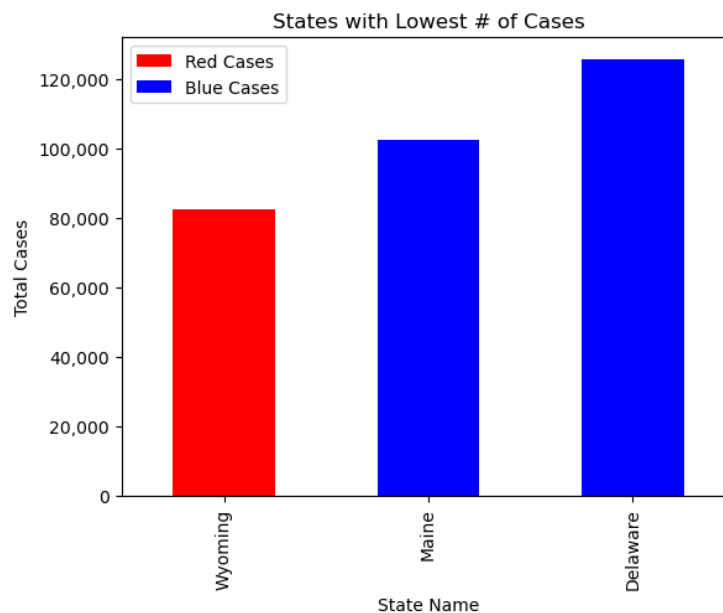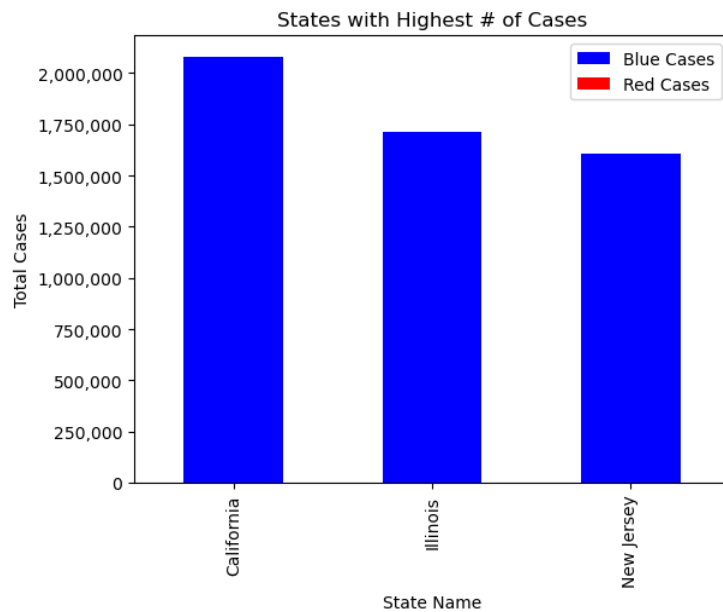
Case Prevalence by Red/Blue States

Total Cases by State
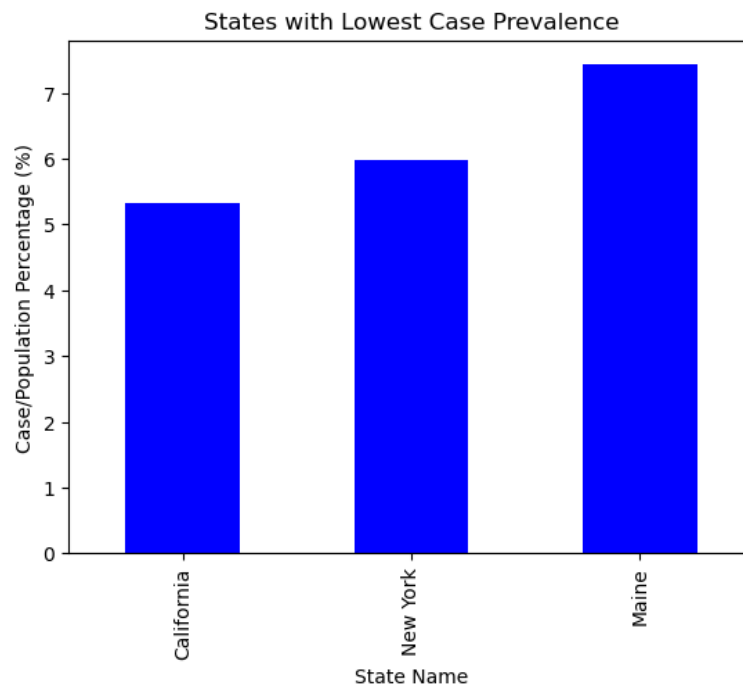

Total Cases by State
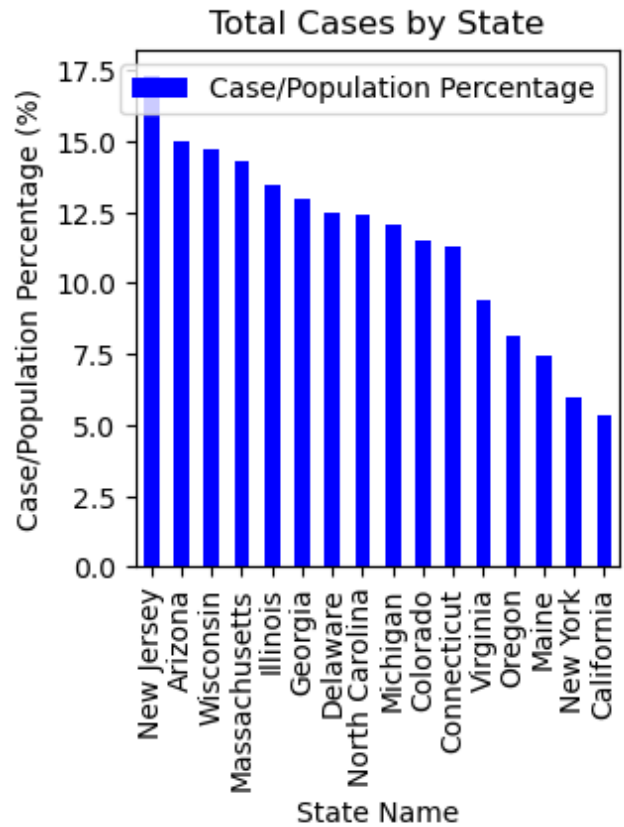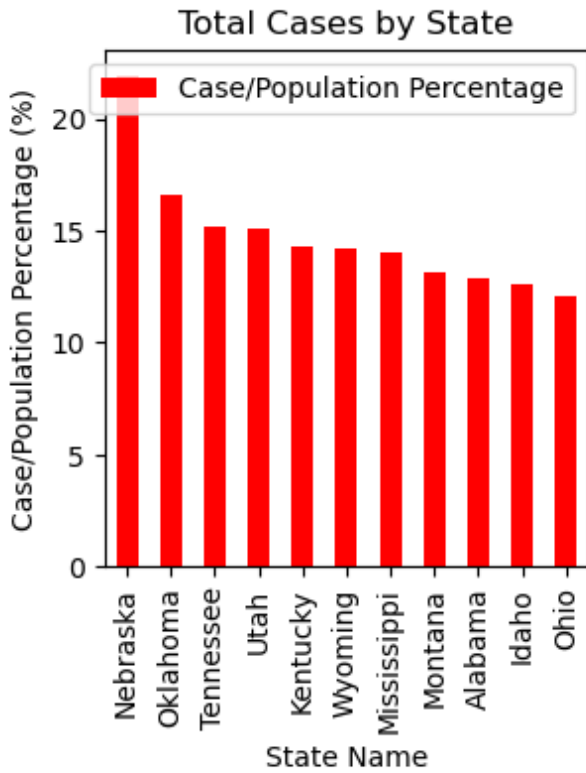
Out of the states mentioned above, which were blue (Democrat) and which were red (Republican)?

The 3 states with the highest number of total cases are all blue states. Whereas, when we take the state population into account the 3 states with the highest percentage of total cases are Nebraska - red and New Jersey - blue and Oklahoma - red.

The 3 states with the lowest number of total cases are Wyoming - red, Maine - blue and Delaware - blue. However, when we take the state population into account the 3 states with the lowest percentage of total cases are all blue.

States with Highest Case Prevalence

Case/Population Percentage (%)

Nebraska
New Jersey
Oklahoma

State Name

States with Lowest Case Prevalence

Case/Population Percentage (%)

California
New York
Maine

State Name
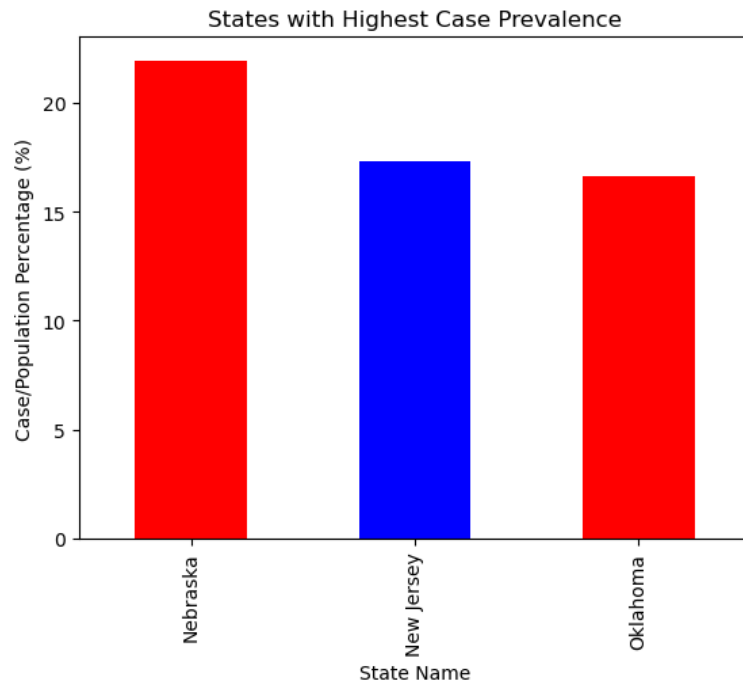
Which states between the blue (Democrat) and the red (Republican) were more affected by COVID-19?

Based on the above analysis, the red (Republican) states were more affected by COVID-19 than the blue (Democrat) states. Even though the blue states had the 3 highest state count of total cases, it is not a fair representation of how the state was affected without taking the state population into consideration. Therefore, after analyzing the same data relative to the state population, we see that the red states had in fact been more affected by COVID-19 with Nebraska and Oklahoma yielding the first and third highest.
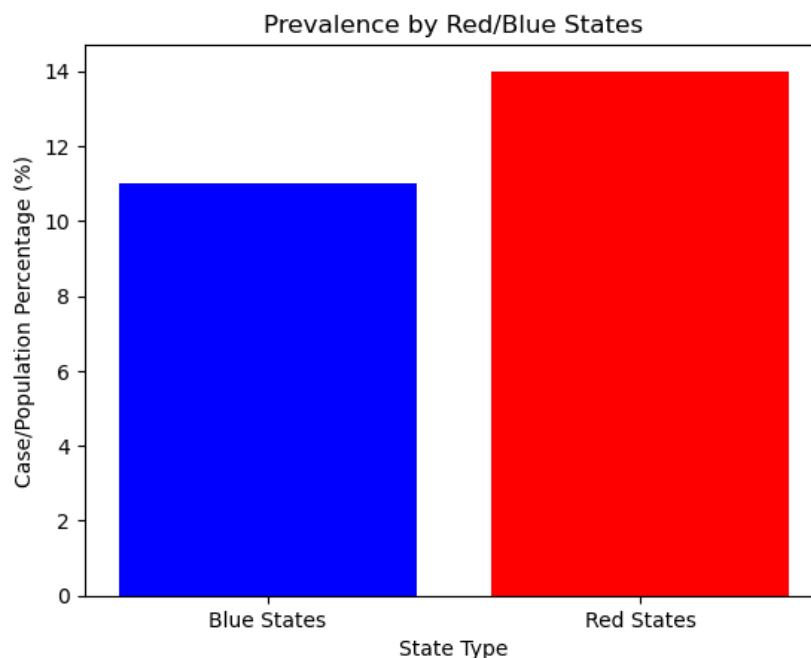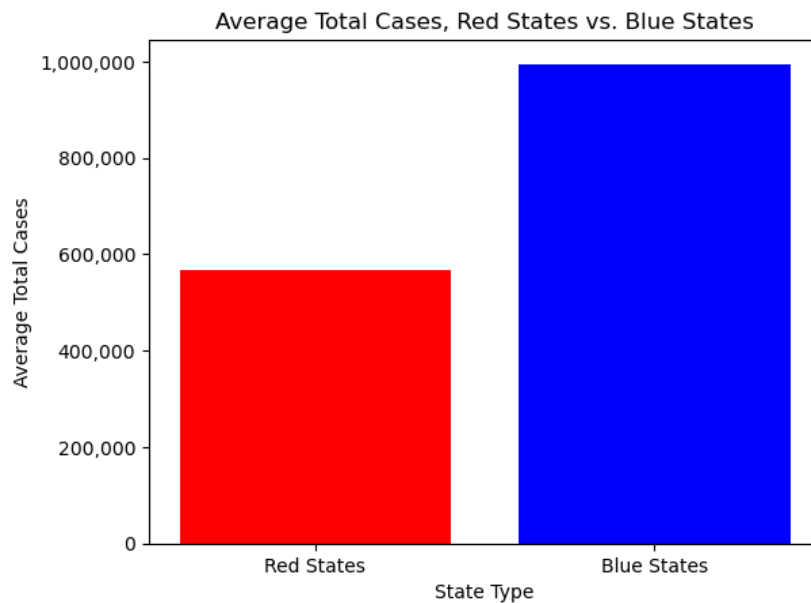
The null and alternative hypotheses

Null Hypothesis: The type of state government (red/Republican or blue/Democrat) has no impact on the number of COVID-19 cases.

Alternative Hypothesis: The type of state government (red/Republican or blue/Democrat) has an impact on the number of COVID-19 cases.

After analyzing the data we found that the t-test was the most appropriate to test the hypotheses. We performed 2 tests to stay consistent with the previous analyses. One was done using the total number of cases per state by color and the other was done using the total number of cases by population percentage per state by color.

The results of the two tests were as follows:

- ○ Total Cases:
    - ■ t-statistic: 33.592271324857755
    - ■ p-value: 8.142141668872524e-242
- ○ Case Prevalence
    - ■ t-statistic: -22.420479534465183
    - ■ p-value: 3.330387755169245e-110

In order to reject the null hypothesis the p-value needs to be less than 0.05 and as we see from the results of both tests that the p-value is significantly lower than 0.05. Based on this we can reject the null hypothesis and therefore we can accept the alternative hypothesis.

Final analysis and conclusion

After conducting all the above analyses and rejecting the null hypothesis we concluded that there were several factors that made an impact on the number of COVID-19 cases.

When we examined the data of the blue states the initial calculation of average total number of cases per state was high. These states generally have larger populations and tend to be clustered around coastlines. They also have some of the country's larger cities and business hubs such as Los Angeles, New York City, Chicago and Atlanta. These cities are also travel hubs with very busy and large airports. These reasons combined with higher populations led to a higher number of COVID-19 cases. However if we look at the percentage of the population that got affected by COVID-19, it was less than the red states. In spite of these circumstances, more highly-regulated masking and vaccination mandates did successfully mitigate spread in these areas.

In comparison the red states have smaller populations and tend to be located in the middle of the country so they are not as easily accessible. The Republican legislation had been more relaxed about vaccination and masking mandates. Even though the actual case count of individuals diagnosed with COVID-19 was low, a greater percentage of the population got sick from COVID-19.