**Introduction**

Our project studied the impact of COVID-19 on all US states throughout 2020. We aimed to discover if the number of available hospital beds was correlated to COVID-19 death rates, identify any specific US states or regions where COVID-19 deaths were more prevalent, and find out if state population size was correlated to COVID-19 death rates. Our analysis involved looking at the total cases and deaths by state and evaluating how these metrics were influenced by hospital availability and state population.

To accomplish this, we collected data from The Covid Tracking Project API for cases and death rate, kff.org for hospital beds per 1000 capita and the US census API for population data. We pulled the death and cases data for each state from the Covid Tracking API and loaded it into a Pandas dataframe. This dataframe was then merged with the US census API data per state and hospital beds per 1000 capita. The hospital bed per capita data was acquired through a CSV file. We removed extraneous rows and columns, and performed a merge based on states with the existing dataframe. The output of our data cleaning process was a dataframe with the following key information:

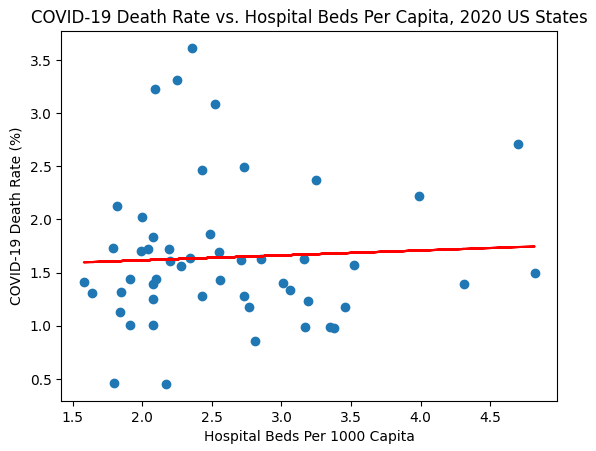
* + States
  + Hospital beds per 1000 capita
  + COVID-19 death rate
  + Population size

**Hypothesis 1**

Greater hospital bed availability will reduce the death rate by COVID due to increased probability of care.

**Results and Conclusions**

A scatter plot and linear regression model was used to evaluate the correlation between hospital beds per capita and COVID-19 death rate. The number of hospital beds available is not correlated with the COVID death rate. The analysis showed a 0.05 correlation coefficient (weak positive correlation).



**Possible Explanations**

High hospital bed availability may not be the best measure of care. Hospitals may not be accessible to everyone in the state equally.

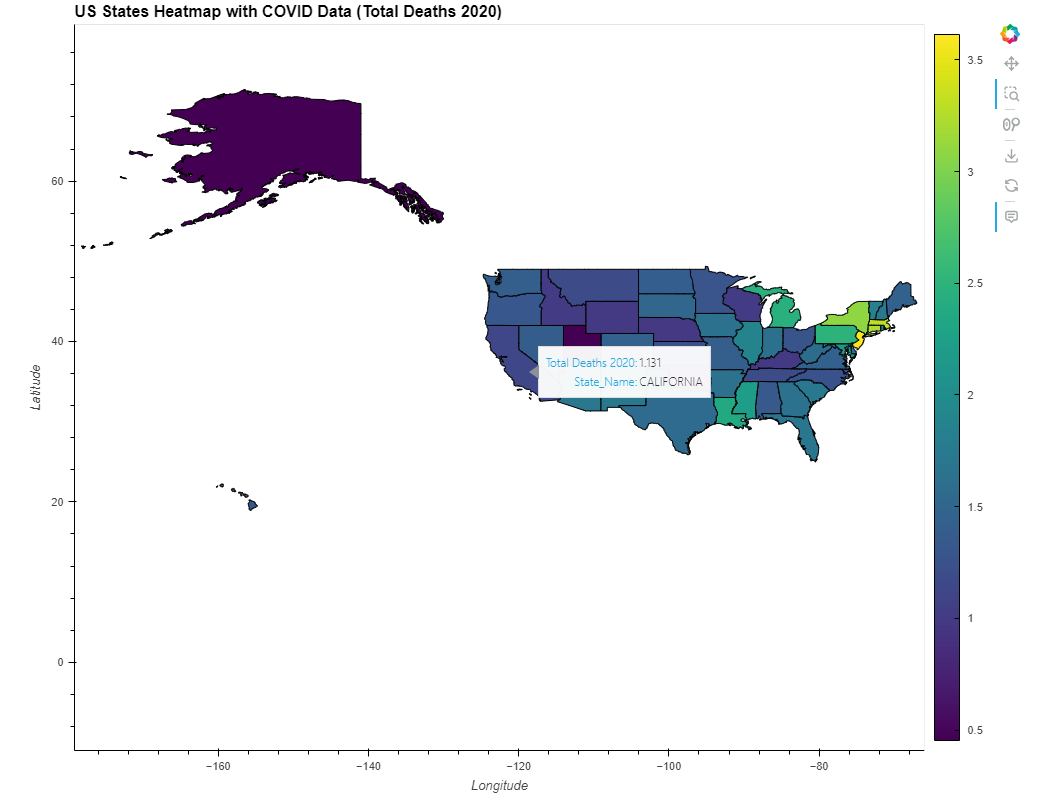
**Hypothesis 2**

COVID deaths will be more prevalent in states/regions with higher population density, such as the Northeast.

**Results and Conclusions**

A heatmap was created to visualize the prevalence of COVID-19 deaths across states.

Covid deaths were most prevalent in densely populated states such New Jersey, Massachusetts, Connecticut, New York, and Pennsylvania. The lowest death rates were in states with low population density: Alaska, Utah, Oklahoma, Nebraska, and Wyoming. With the exception of Alaska, the lowest death rates occurred in the Midwest/Great Plains region..



**Possible Explanations**

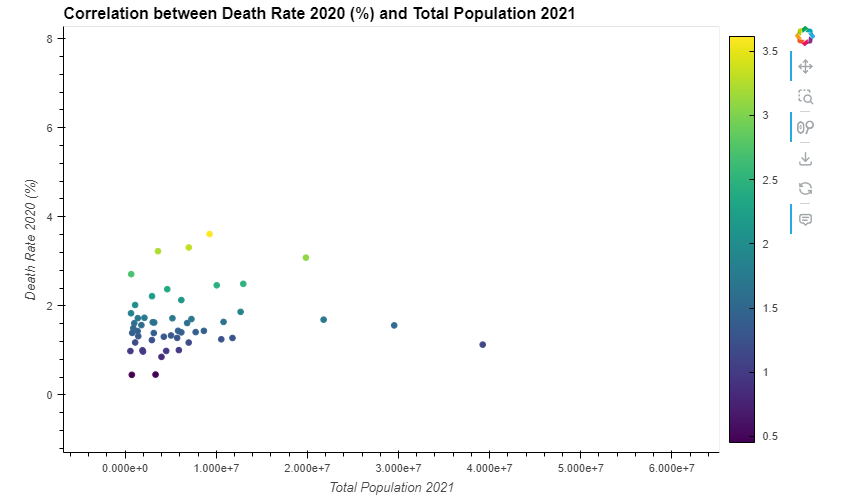
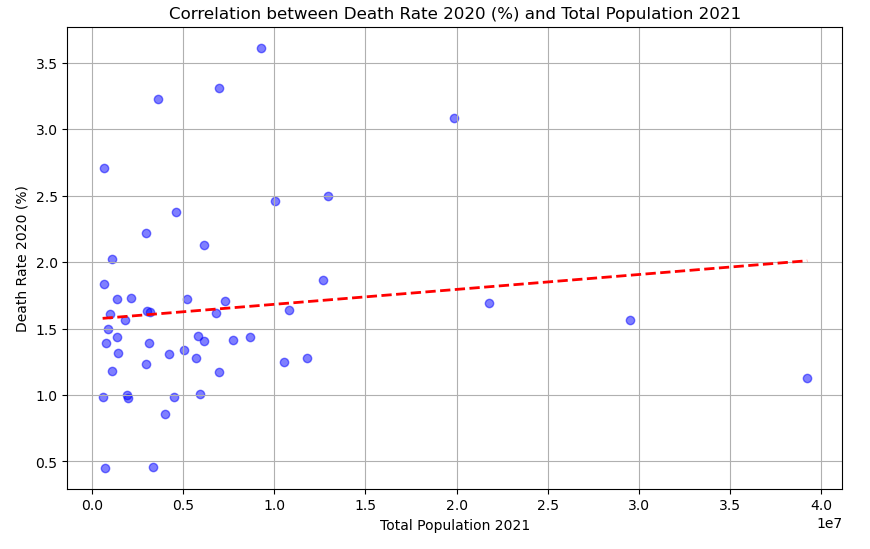
High population density increases risk of spread.

**Hypothesis 3**

COVID-19 death rates will be positively correlated with state population size.

**Results and Conclusions**

An additional scatter plot and linear regression model was used to evaluate the correlation between hospital beds per capita and COVID-19 death rate. COVID-19 death rate and state population size are not correlated. The analysis showed a 0.12 correlation coefficient.



**Possible Explanations**

Population size is likely not a large contributing factor to COVID-19 death rate compared to population density, which indicates proximity of individuals within certain areas. This explains why states with large population sizes but low population density (CA, TX, FL) have lower death rates than states with large population sizes and high population density (NY).