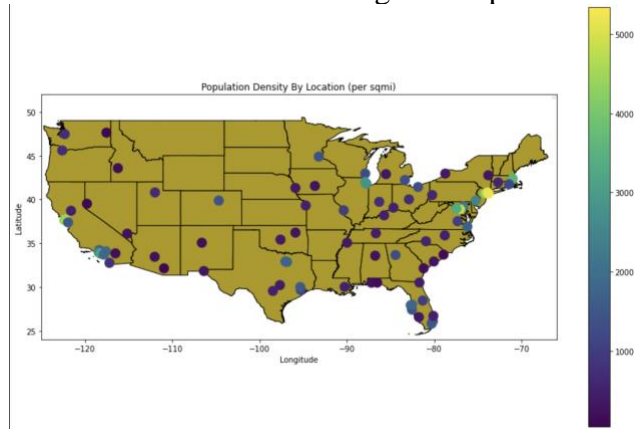
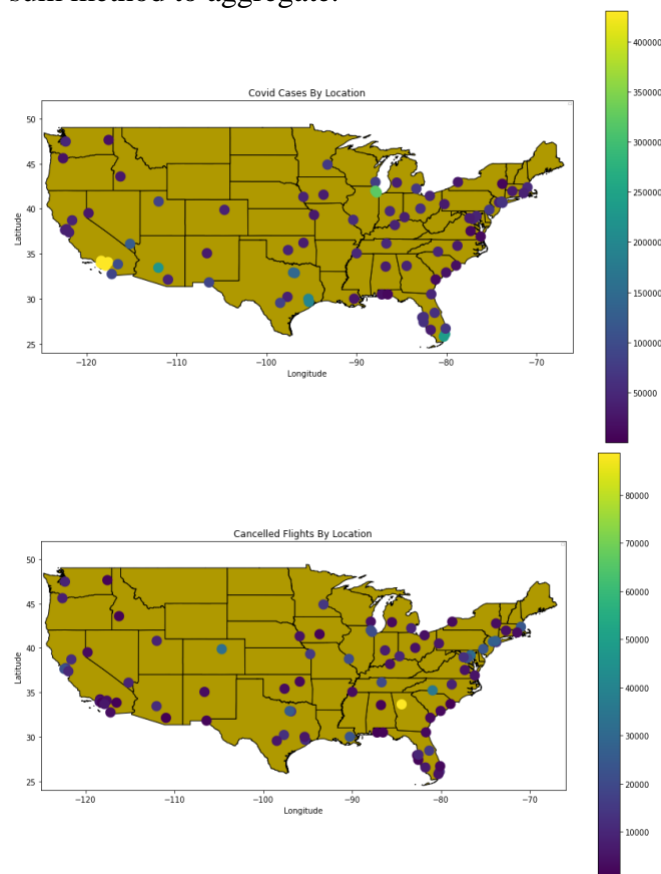


I have a dataset that contains the covid data and the count of flights for each county by week with 16302 rows \times 45 columns. This dataset contains the geometry location for each county and airport. Firstly, I create a geo map showing the population density of each county where the main airports in the US are located using geopandas, excluding some small private airports. And I create a US state border using the shape file from the US census.



Then I display the total cases in each county in the same way. I grouped the data frame by each county and aggregated the cases using the max of cases in the last week, which is the total number of cases finally. And I did the same to the total number of canceled flights but using the sum method to aggregate.



However, I could only see the relation between the population and cases, but covid data generally have nothing to do with the number of canceled flights. This result does not make

sense to me. So finally, I selected Atlanta, which had the greatest number of canceled flights from this dataframe as an example using its 'ORIGIN_AIRPORT_ID' to find if there is any relation. I plotted a line graph that shows the weekly changes in the numbers of airlines and canceled airlines in Atlanta. And specifically, I chose a date when the stay-at-home policy was announced in Atlanta as a flag, coloring the weeks before and after this date with different colors. And there was a clear increase in canceled flights and a huge drop in airlines after the policy was exactly effective.

