Interpreting The Effect Of COVID On Motor Vehicle Accidents in the New York Metropolitan Area

Jian Ruan, Zach Yu, Carrie Hashimoto, Sophia Lyu (Undergraduate), Patrick Yee (Graduate)

March 8, 2023

1 Introduction.

COVID-19 quarantine limited people to remote work and in-home activities. It's likely that the COVID-19 restriction of transportation decreased the number of car crashes in NYC. To test out our hypothesis, we first compare and contrast the car crash distribution through geographical mapping, and later through time-series analysis and free-text analysis based on New York Times articles.

2 Geographical Distribution Of Car Crashes Before And After COVID

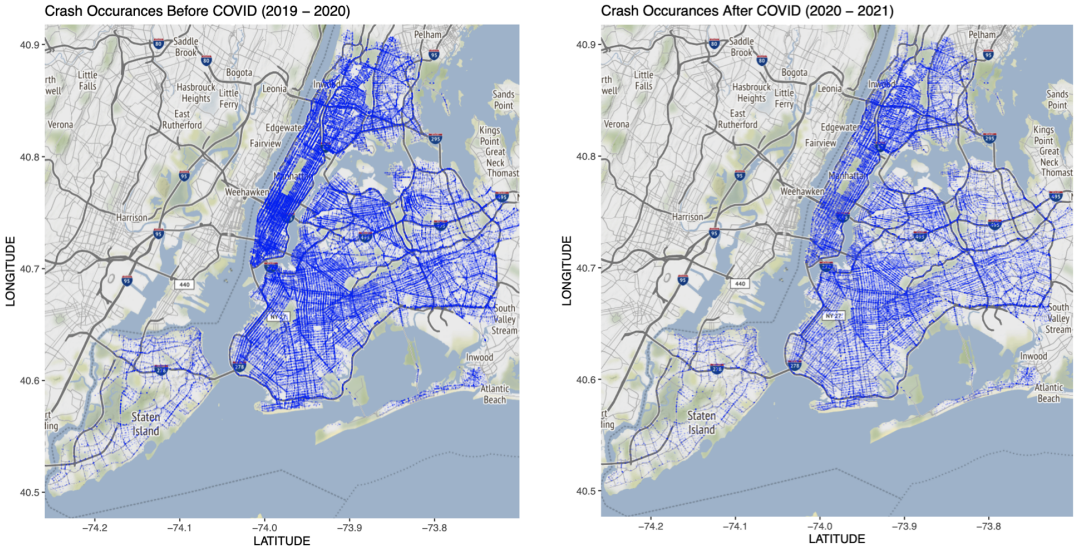

Figure 1: Accidents Distribution Before vs After COVID.

Each incident is mapped as a blue dot (x = lattitude,y = longitude) based on the OpenStreetMap API for New York City. The bluer the area shows, the more accidents happened.

As shown in Figure 1, there is a sudden decrease in the number of accidents coinciding with the start of the COVID-19 pandemic, especially around the two most populous areas - Manhattan and Brooklyn. Since COVID transmitted quickly among large group of population, Manhattan and Brooklyn were the top two areas with people tested positive at the start of 2020. As a result, the quarantine policy for Manhattan and Brooklyn was more strict than other places, and people adapted to online activities to avoid infection.

3 Time Series Ribbon Plot

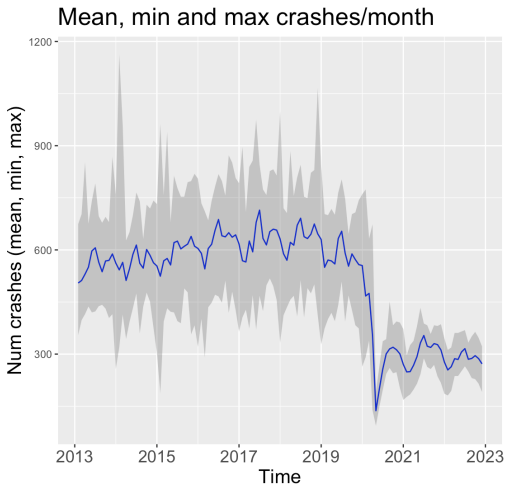

Figure 2: Ribbon Plot

As base and ggplots were being made in the previous analysis, it came to our attention that the number of crashes changed significantly around the onset of the COVID-19 pandemic.The ribbon plot created above shows a clear drop in the number of crashes for the beginning of 2020, which was roughly when the pandemic started, followed by a partial rebound later in the year. The blue line represents the mean value of car crashes per month, bounded by the maximum and minimum values in the grey area.

4 Free-text Analysis Based On New York Times Articles

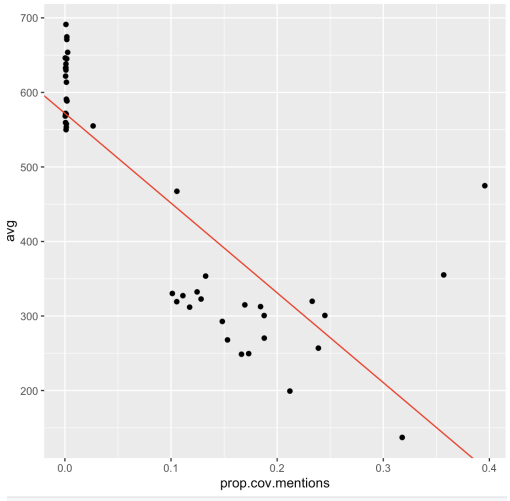

Figure 3: Linear Regression - Average Car Crashes vs Mention Of COVID In New York Times

The trend from the Ribbon Plot led us to ask the question of how the COVID-19 pandemic affected the transportation in New York City. To further see if the transportation within New York City was really affected by the pandemic, a dataset with headlines of every article from The New York Times from 2018 was collected. A plot of average number of crashes per day in a month vs the number of mentions of something

virus/pandemic related in the headline is plotted.

Through this linear regression model, we see there's a negative correlation between the average number of crashes per day with the mention of COVID in news. It is reasonable since people were more aware of the COVID outbreaks and chose to stay at home instead of going out, leading to a lower probability of car crashes.

5 Contributing Factors Of Accidents Across Time

Figure 4: Contributing Factors Piechart

Since a notable decrease in the number of crashes per day was observed at the onset of the COVID-19 pandemic, we wanted to explore whether any change in the composition of the causes of crashes was observed when the pandemic began. To this end, pie plots showing the relative frequency of each of the top ten contributing factors in the data set by year were created. However, the distribution does not show a significant difference among different years.