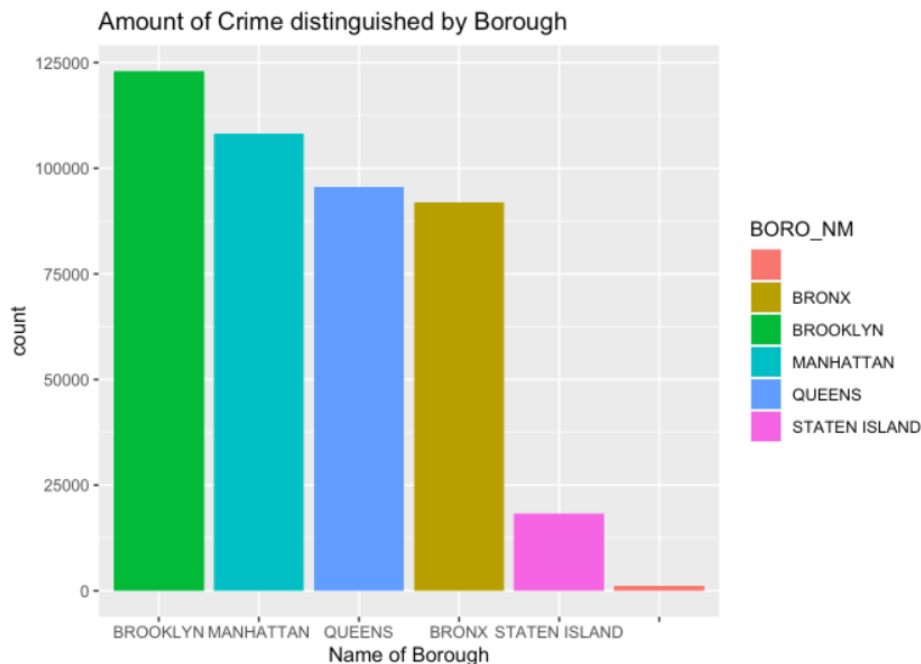


Final Project Process Book

Group B - NYC Public Security

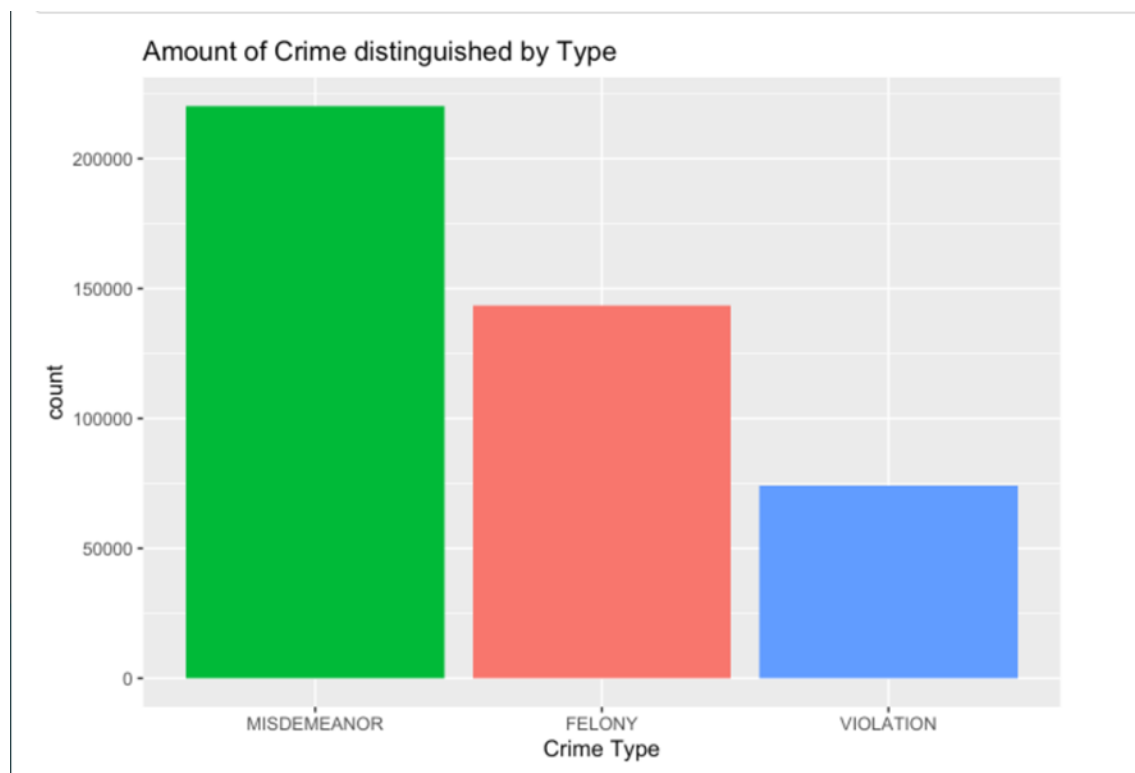
Hanzhi Zhang, Jinghan Ma, Angela Gu, Lin Wang

Our group is trying to figure out the crime distribution and accident pattern in New York city. We used data from the New York Police Department Complaint Dataset including all-valid felony, misdemeanor, and violation crimes reported to the Police. We used mapping and graphs to help us visualize the public security situation in New York.



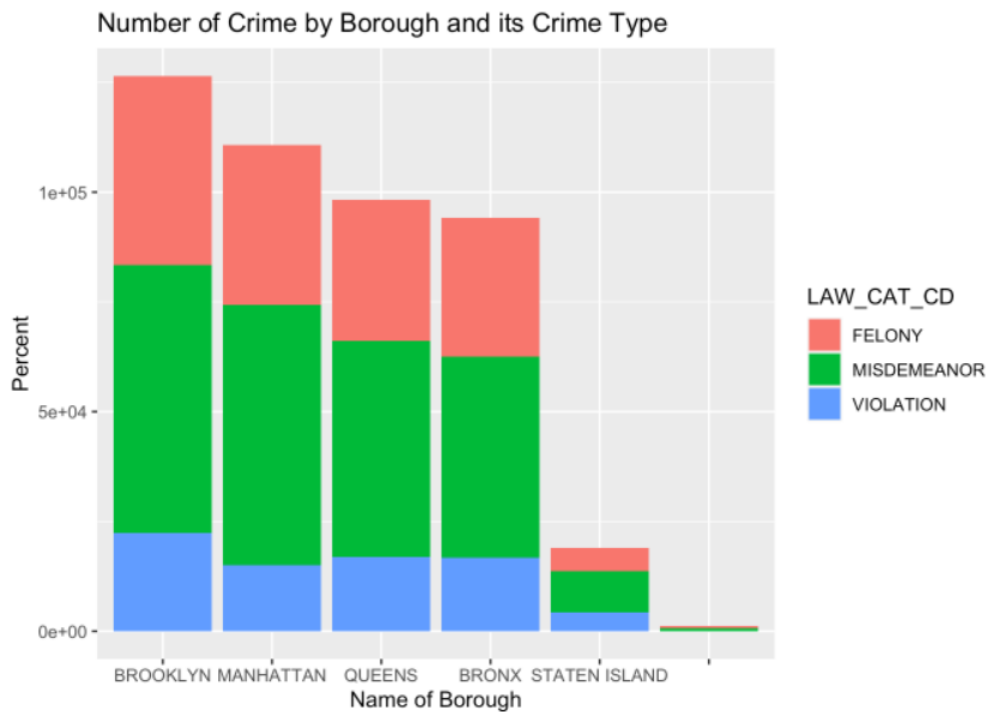
First we want to discover the distribution of crimes happening in each borough, so we used borough names to group and draw the sum of total crimes in the form of a bar chart.

This graph illustrates the number of crimes distinguished by borough among six major living areas in New York. Using the bar chart we can directly make the comparison between different boroughs and easily see that Brooklyn has the highest crime rate with approximately 120000 units. The number of Manhattan and queens ranked second and third place with much lower rates at 105000 and 90000 respectively. The lowest crime report area is another place. Staten Island also contains a low crime rate.

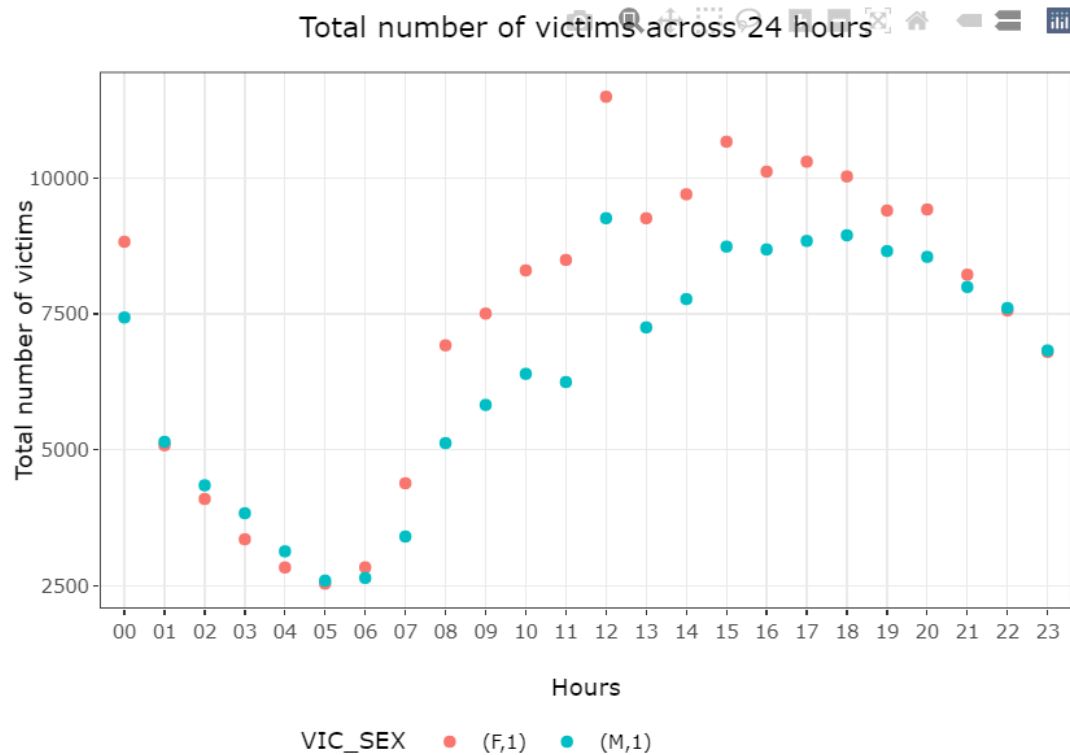


Besides, only analyzing crime in an area is not enough, we also categorized crime into specific types and try to find out what kind of lawbreaking behavior affects most in New York. The second bar chart shows the proportion of different crimes. The result is misdemeanor occupied in top spot with more than 200000 report figures. The second one

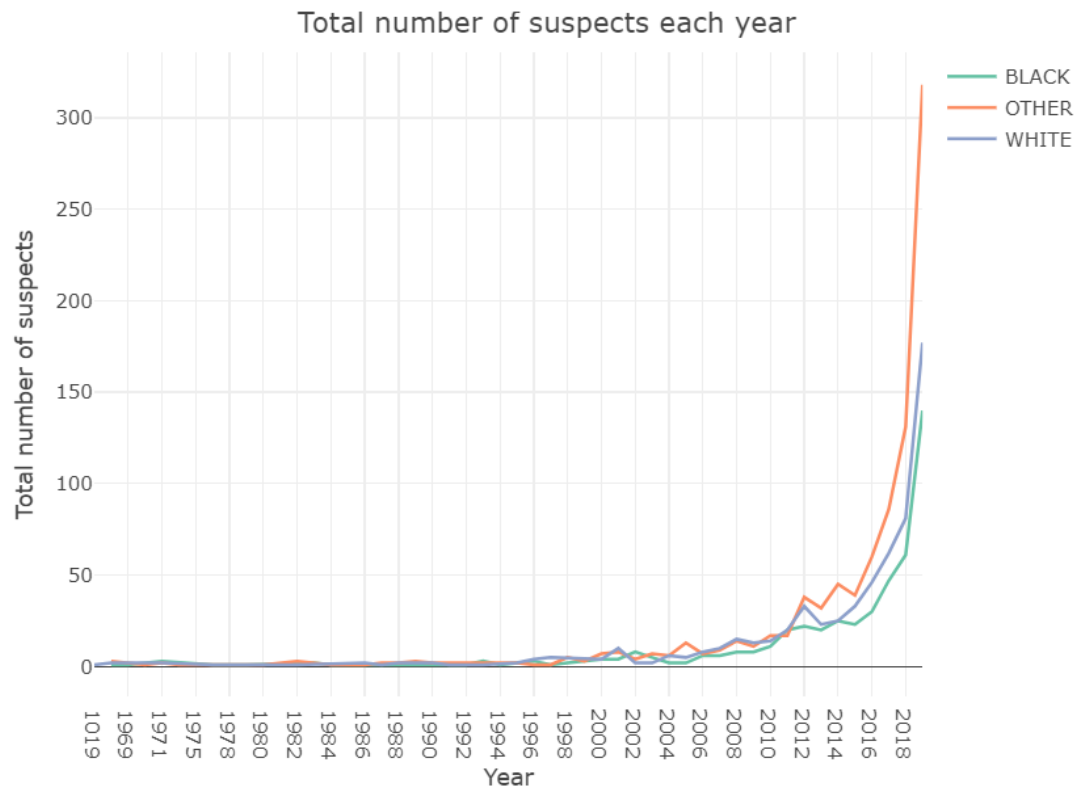
is taken up by Felony with close to 150000. Violation holds a much lower happen rate which is even lower than half of misdemeanor numbers at around 75000.



However, we discarded these two graphs and combined crime numbers and categories into one histogram, so we can reach the same conclusion by using fewer graphs which makes our visualization more clear and concise. This chart demonstrates the total number of crimes by borough with its crime type. Brooklyn still has the largest crime case number. Most crime types are misdemeanors, which are represented by green color.

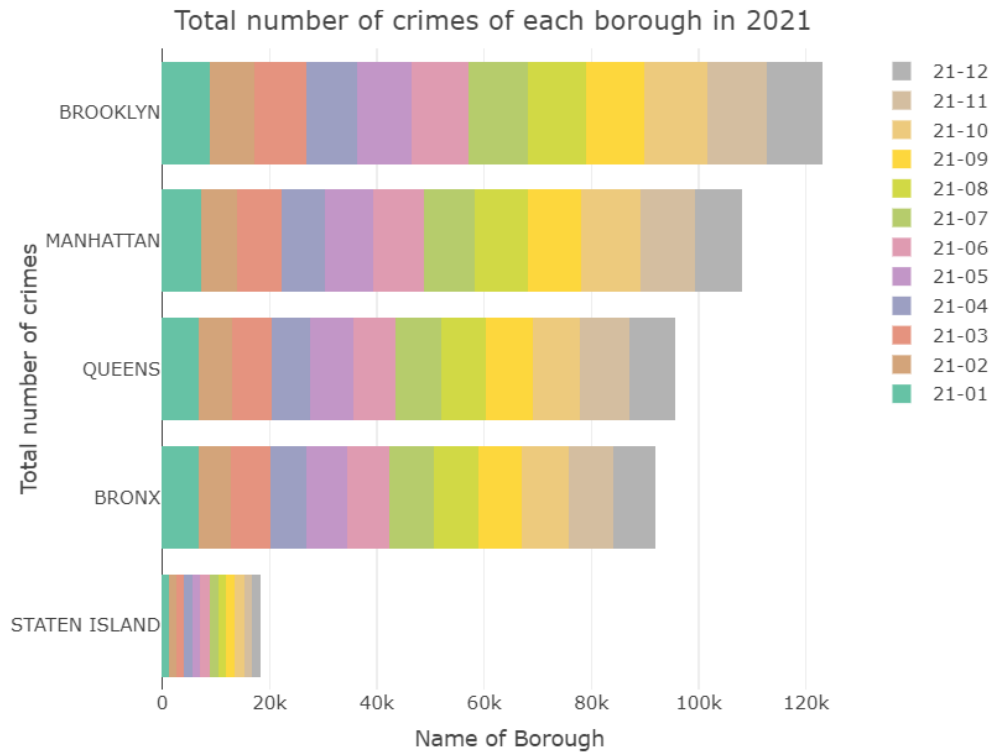


Moreover, we want to analyze the trend of the total victims in 24 hours and find out if there is any difference between genders. We used an interactive point chart to demonstrate the total number of victims over one day (0am - 23pm) and the different color of the dot represents each gender. The curve showed an upward trend from morning to afternoon time. The total number of victims reached peak during 12pm, especially for female victims. Different from what we commonly think, night time from 23 pm to 4pm the number of crimes diminished quickly and reached bottom at 5am. Overall, the number of female crimes is much higher than male.

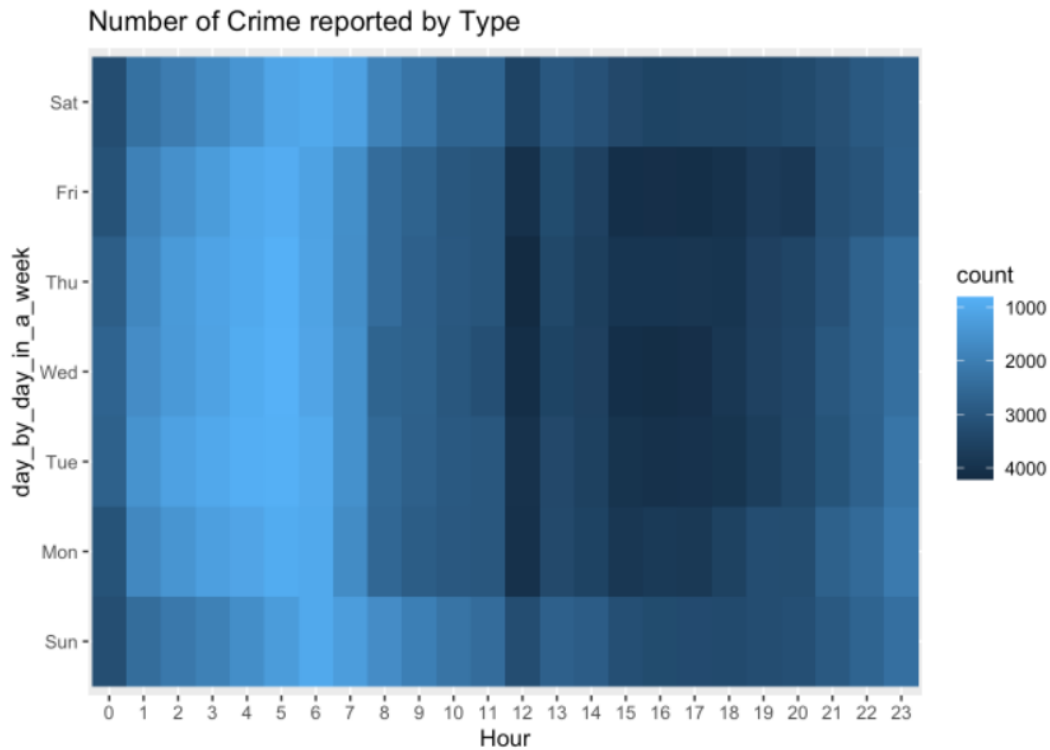


Then, we used ggplot to explore that kind of race with a high crime suspect possibility.

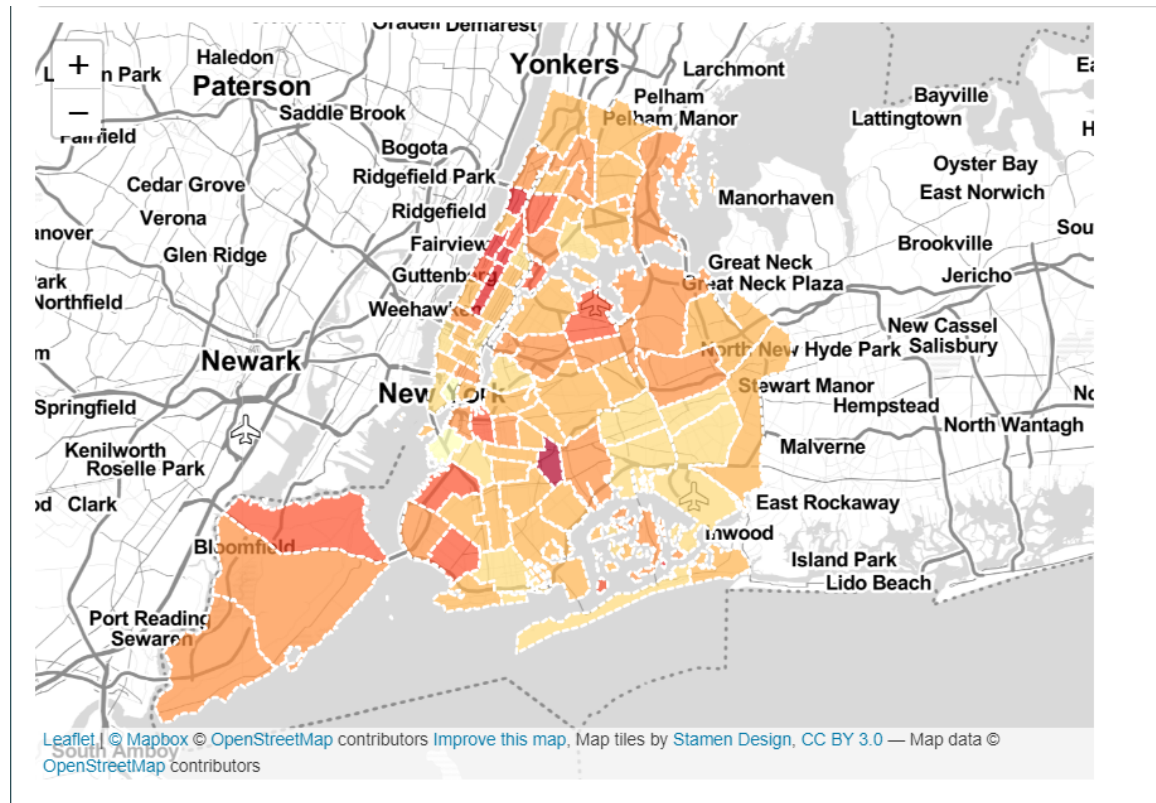
Other race is the most significant race of suspects followed by white and black population. There used to be a similar level from 2019 to 1996. However, the number of other race crimes jumped sharply from less than 50 to over 300 during 2006 to 2018.



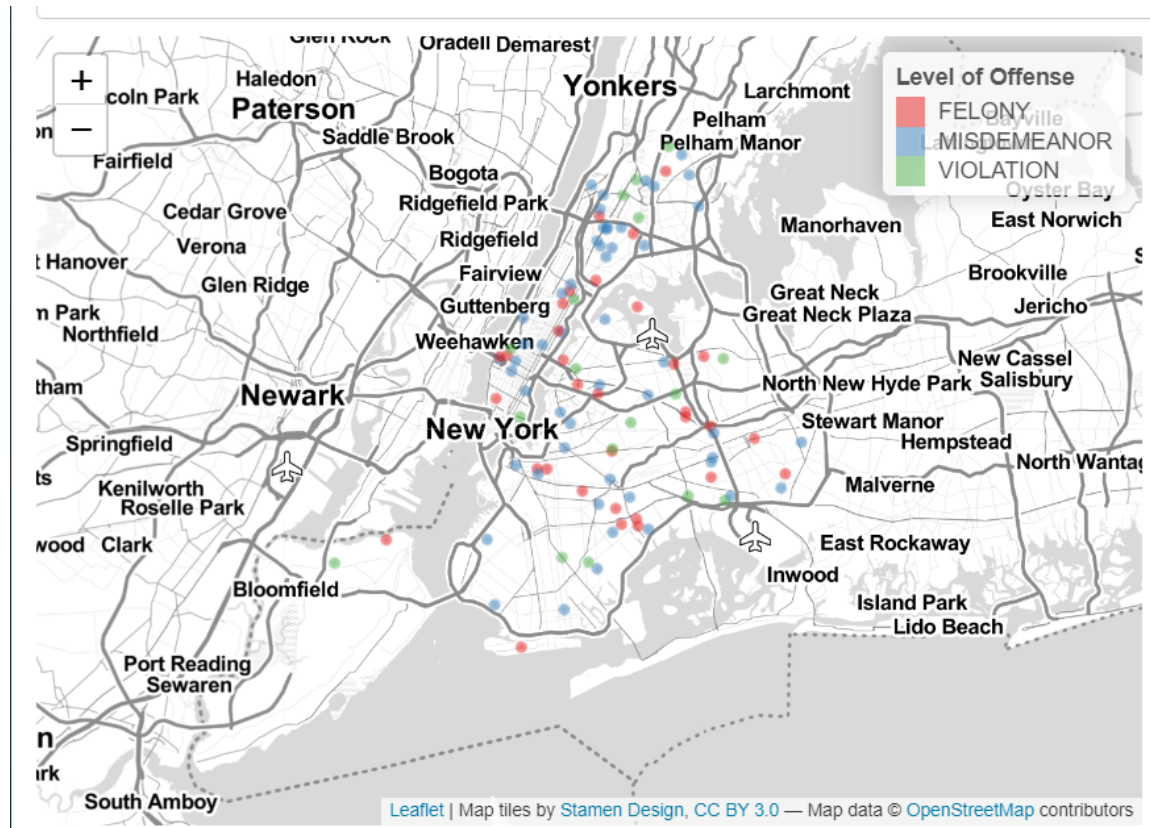
Furthermore, we added more factors in ggplot and selected 2021 as target year to explore the total crime data in New York. Brooklyn still has the highest crime rate at more than 120k and distributes quite evenly every month. Manhattan and queens rank second and third place, with around 110k and 90k respectively. More specifically, from October to December, the crime cases increase quicker than other months in 2021.



Likewise, we used ggplot to create a heatmap in order to more clearly view the crime that happened during one week in New York. The shade of blue represents the number of crimes, the darker the color means the higher the crime. It can be easily found that crime normally happens during noon time, then from 15 - 19 pm during the day. Besides, midnight zero clock also has high incident hour.



Also, we used a leaflet to draw a geo-graph which contains demographic variables in the Census zip code data that may relate to the crime rate in New York City. The graph illustrates that the crime rates are really different in various districts. The darker the color, the higher crime rates. The areas with high proportions are distributed by Brooklyn and Manhattan, especially in the east side of Brooklyn and west upper side of Manhattan.



Lastly, we also use graphs to demonstrate the level of offense distributed in New York. Red means Felony. Misdemeanor is a blue spot. Violation is the green dot. The blue color, which is a misdemeanor, scatters all over the New York City area. It has the highest happen rate which is concentrated in the Brooklyn and Manhattan area. Then there is the felony which is much lower than the first crime type. The violation just maintains a small portion of the whole area.

In Summary, the data visualization helps us better understand the crime destruction and pattern in New York city. Contrary to what we used to believe crime happens during the

time, noon has the highest occurrence possibility. Besides Brooklyn has a high crime report rate, we should pay attention to the safety issue in this area.