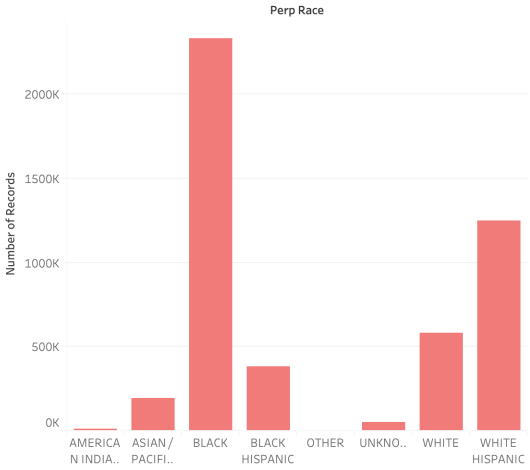
NYPD Arrests Database Exploration

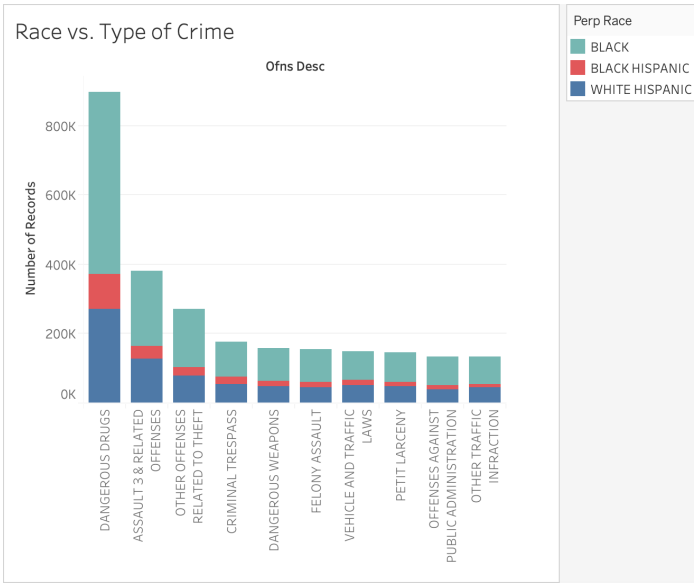
NYPD Arrests Database from Data.gov. Source: https://catalog.data.gov/dataset/nypd-arrests-data-historic. When first looking at my data, I think that this database is a very accurate representation of crime in the New York City area. The database provides different dimensions such as race, gender, age, but also shows the locations of different crimes, the type of crimes, and the different law codes.

My first step in attempting to use a visualization to answer my initial question, I looked at the total number of arrests that occurred from 2006-2018. In total, there were 4,798,339 arrests made, and 3,965,084 of those people arrested were either Black, Black Hispanic, and White Hispanic. I think this is an important statistic to note when analyzing this question, considering that 83% of the arrests made from 2006-2018 were people that were Black or Hispanic.



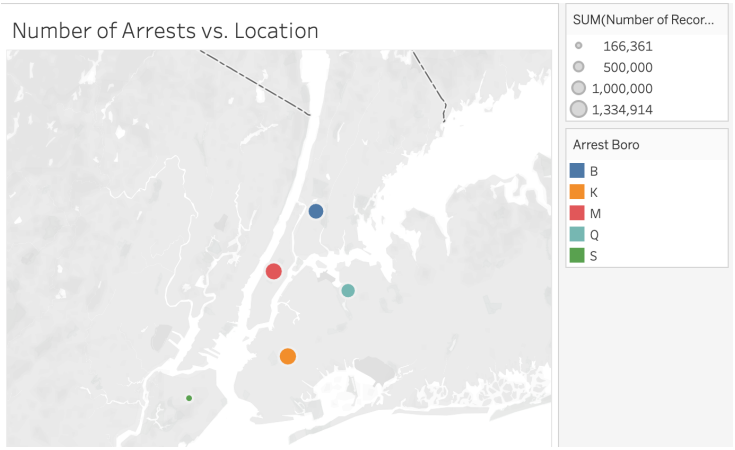
My first visualization as shown above, shows the race breakdown of the different arrests that I mentioned briefly above.

The next question that arose from my analysis of my data, was what type of crimes theses races were committing that was causing them to represent 82% of the database. Due to the immense amount of crime codes and types of crimes in the database, I limited this data down to the top ten types of crimes based on number of arrests.



As depicted in this visualization, the main type of crime for arrests is Dangerous Drugs, which is an obvious issue with New York City and the war on drugs. However, I think that it is important to illustrate how a majority of these crimes are related to very minor offenses. It is crazy to see how many black or hispanic Americans are arrested for minor crimes such as traffic violations. A question that had been lingering in the back of my mind as I was looking at this data was whether there is racial profiling in the NYPD. Lastly, as I was analyzing my data I wanted to finally look at the effect of different boroughs on number of arrests in New York City. This data is important for me to look at when deciding where to potentially live in the years to come.

An important statistic to note before we analyze this data is that the five boroughs of New York City are depicted by B, K, M, Q, and S. The B stands for The Bronx, K stands for Brooklyn, M stands for Manhattan, Q stands for Queens, and the S stands for Staten Island.



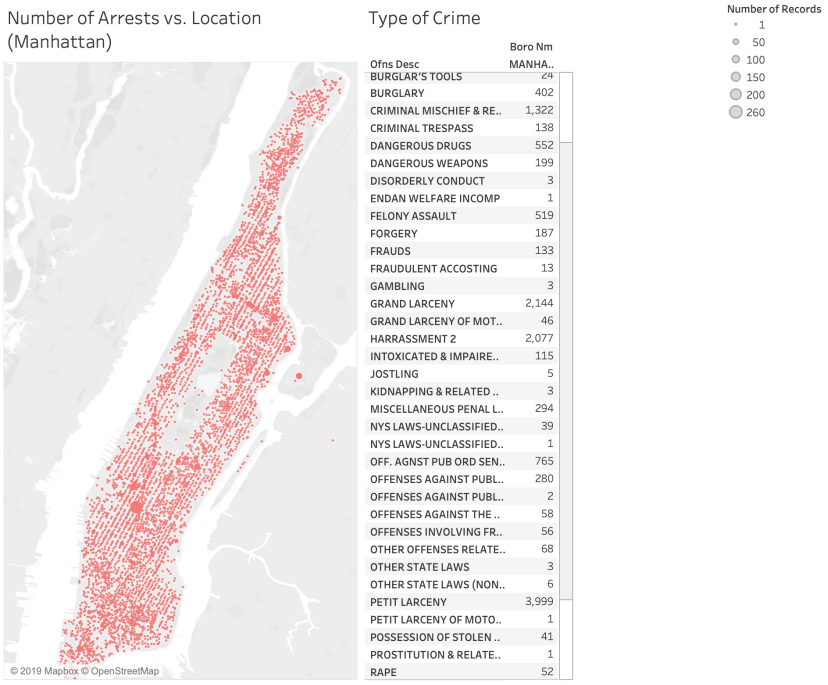
In the visualization I created, you can see the different boroughs depicted by the colors in the legend on the right, and the number of arrests by the size of the circle depicted in the legend in the top right. As you can see, the area with the highest number of arrests is Brooklyn, followed by Manhattan, The Bronx, Queens, and Staten Island. I think that this visualization accurately depicts which areas of New York City are safer however, an important statistic that we need to note is the different population sizes in each location. The populations of each area is:

- Brooklyn: 2,648,771
- Queens: 2,358,582
- Manhattan: 1,664,727
- The Bronx: 1,471,160
- Staten Island: 479,458

These population sizes are an important statistic because the number of arrests per each area must be analyzed based on the number of individuals that live in that borrow. On paper, Staten Island seems like the safest area but with that many arrests per that small of a population, it shows how it isn't as safe as it seems. Based on statistics, the area with the highest population and the lowest percentage of arrests per population size is Queens. Although Manhattan has a lot of arrests per population size, this is the central point of New York City and the top destination for me as I am moving. The next question that crossed my mind as I was doing my analyzation of this data was, Which areas of Manhattan have the lowest numbers of arrests?

In order for me to complete the last part of my data analysis, I had to use a different data set in order to get a more detailed look at the specific latitude and longitude of different crimes in the five boroughs in New York City. The data set that I used for this last portion was Crime Map data from NYC Open Data. Source: https://data.cityofnewyork.us/Public-Safety/Crime-Map-/5jvd-shfj. This data takes a deeper look at the

location of crimes in New York City, and represents all crimes reported to the New York Police Department in the year 2016. Obviously since this data only represents data from the year 2016, it isn't an accurate representation of the historical crime data for this location, however it still gives me good insight on the areas of Manhattan based on crime data.



This data visualization shows the areas of Manhattan where crime has been reported in the year 2016. The legend on the right side of the visualization shows how the size of the circles represent the different number of crimes. The different size circles and the effect of density can be seen in this visualization. On the right side of the visualization, I have also included the different types of crimes and the number of each of those crimes. Attached to my document is a link to my dashboard that I created, which allows you to use this visualization to effectively filter based on both crime and location. I understand that it is hard to see the different crimes per location in Manhattan based off of this screenshot, and it is hard to figure out how to embed an interactive visualization. However, I think this last piece of my analysis of arrests in New York City is the most important part to understanding my questions. Using this visualization for example, you can click on individual circles in each area of Manhattan and it will show you the number and the type of crime. You can also click on any type of crime on the right side, and it will show you where these different crimes have occurred. I used this final step of my analyzation to further look into areas in Manhattan, and decided that the best area for me to live in after graduation would be the Upper East side based on types of crime and number of crimes using my interactive visualization.