Research Question #1: Is there more crime in Bronx than in Brooklyn?

Barplot of arrest Boroughs
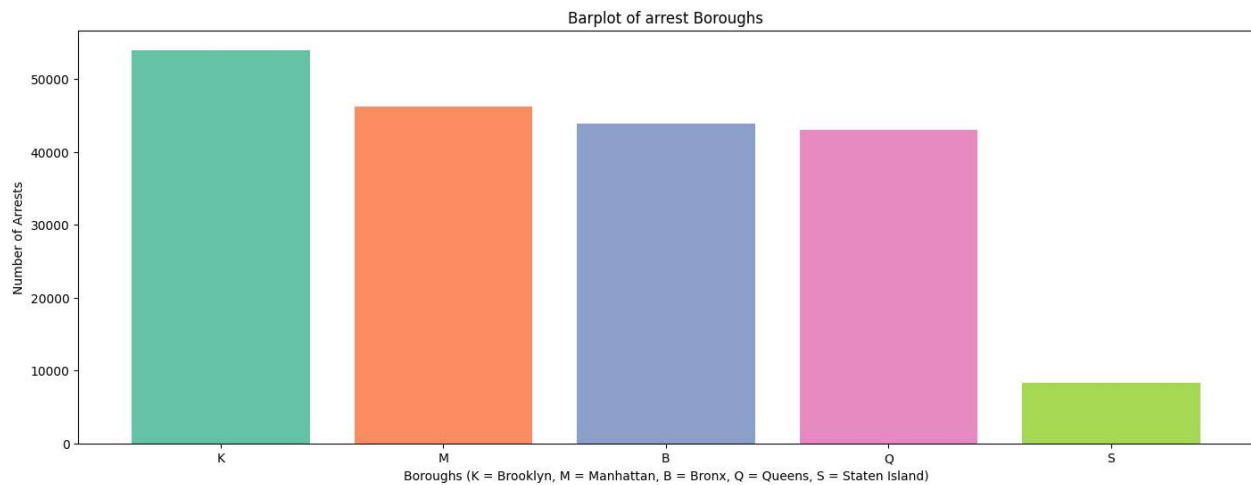
Borough	Number of Arrests
Brooklyn	53994
Manhattan	46235
The Bronx	43899
Queens	42978
Staten Island	8341
Total	195447
Average	39089.4

To visualize arrest data between the boroughs of New York City a bar plot was chosen. Each borough having its distinct arrest count shown can help easily identify the higher crime rate zones. It can be seen by this graphic that Brooklyn has more arrests than the Bronx. The original question asked if there was more crime in the Bronx than Brooklyn, which is disproved.

The bar plot shows the distribution of data across all the different boroughs. The total amount of arrest was 195447. The lowest arrest count is in Staten Island at 8341 arrests, and the highest happens to be in Brooklyn at 53994 arrests. The average amount of arrest per borough is approximately 39089. The first quartile of the mount of arrests was 42978, and the third quartile was 46235. Using that the IQR can be found to be 3257. The lower bound (Q1 – (1.5*IQR)) is 38092.5, and the upper bound (Q1 – (1.5*IQR)) is 41349.5. This shows that the arrests for both Staten Island and Brooklyn would be considered outliers.

One expectation that was set was that the Bronx has a significantly higher poverty rate than Brooklyn, which leads to a higher arrest rate. This expectation led me to believe that the number of arrests would potentially be higher the Bronx than in Brooklyn. However, the graphic has changed my expectation. It seems that the poverty rate does not impact the amounts of arrests happening.

After observing the above bar plot, I would like to investigate further questions. One further question that I have is why are both Staten Island and Brooklyn outliers in this data set?

Research Question #2: Are there more arrests on Federal holidays?

Line Plot of Arrests Over Time
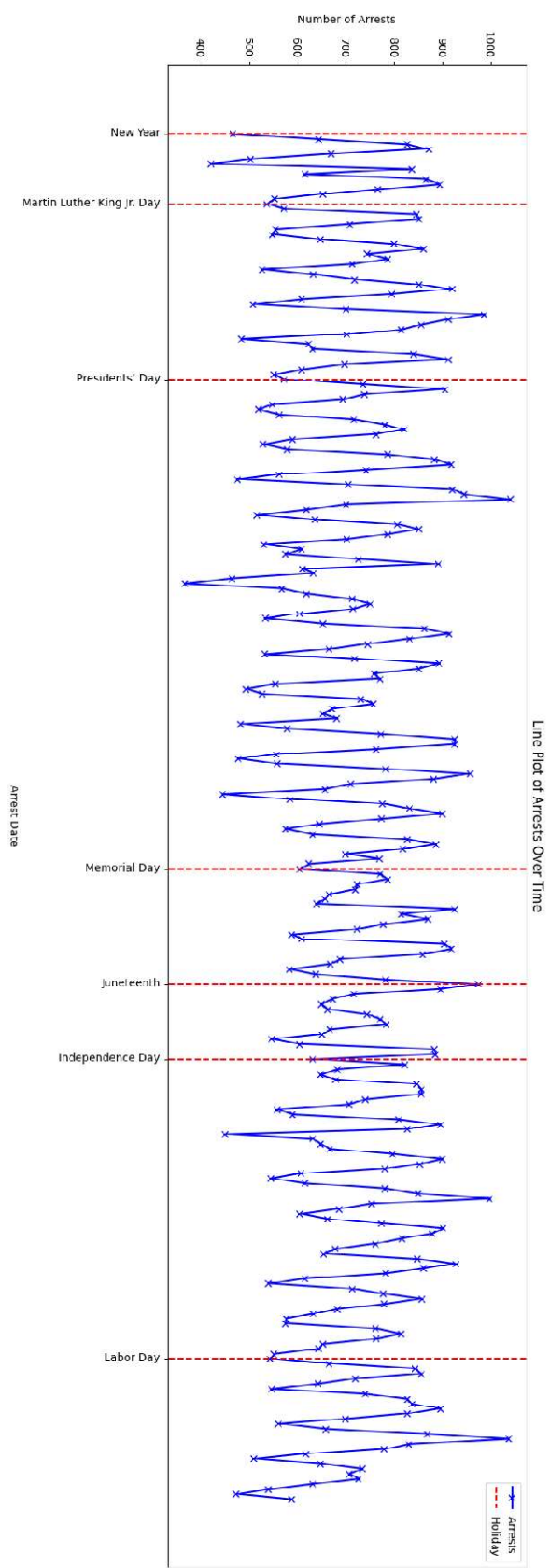
Statistic	Daily arrest amount	Date(s)
Min	366	2024-03-31
Max	1040	2024-03-14
Median	712.5	
Average	713.3102189781022	
Standard deviation	132.53054981820873	
Mode 1	630	2024-05-20, 2024-07-04, 2024-07-20, 2024-09-27
Mode 2	651	2024-01-13, 2024-04-08, 2024-04-26, 2024-08-30
Mode 3	855	2024-02-08, 2024-07-10, 2024-07-11, 2024-09-05
Q1	614	
Q3	824.75	

The graphic above is a line plot with each date's arrest amount for the entire city of New York. In red Federal holidays are marked out and labeled on the x axis. The original question posed (are there more arrests on Federal holidays?) is disproved here except for Juneteenth having a spike in arrests. It is also seen on the graphic that there are relative minimums happening at most of the federal holidays except for presidents' day and Juneteenth.

The highest daily arrest in the city was made on 2024-03-14 at 1040 arrest and the lowest amount was 366 on 2024-03-31. The IQR is 210.75. The upper bound is 1140.875 and the lower bound is 297.875. Using that information there are no outliers in the arrests over time. The standard deviation is 132.53054981820873. There happens to be three modes 630, 651, and 855 daily arrests. The first mode occurs once on the federal holiday of July 4th.

The expectation that was made before creating the graphic was that on federal holidays people do not have work so they would have more time for illicit and illegal activity. This led me to believe that there would generally be spikes on the holidays in arrest amounts. The graphic has led me to change my expectation. The one exception here is the spike on Juneteenth. Juneteenth is the commemoration of the end of slavery in the US. Another expectation that could be an explanation is that African Americans are disproportionately arrested compared to other races. Juneteenth is a celebration of African American freedom so because of festivities perhaps more arrests were made. I would like to investigate further into seeing the races that were arrested at this time.

Presidents' day and Juneteenth happen to not occur at relative minimums in the data set. It is worth nothing though presidents' days daily arrest occurs the day after a relative minimum. Juneteenth happens at a relative maximum. A question that I would like to investigate further is about the celebrations that occur on these holidays that may lead to this phenomenon.

Boxplot of Arrests by Month
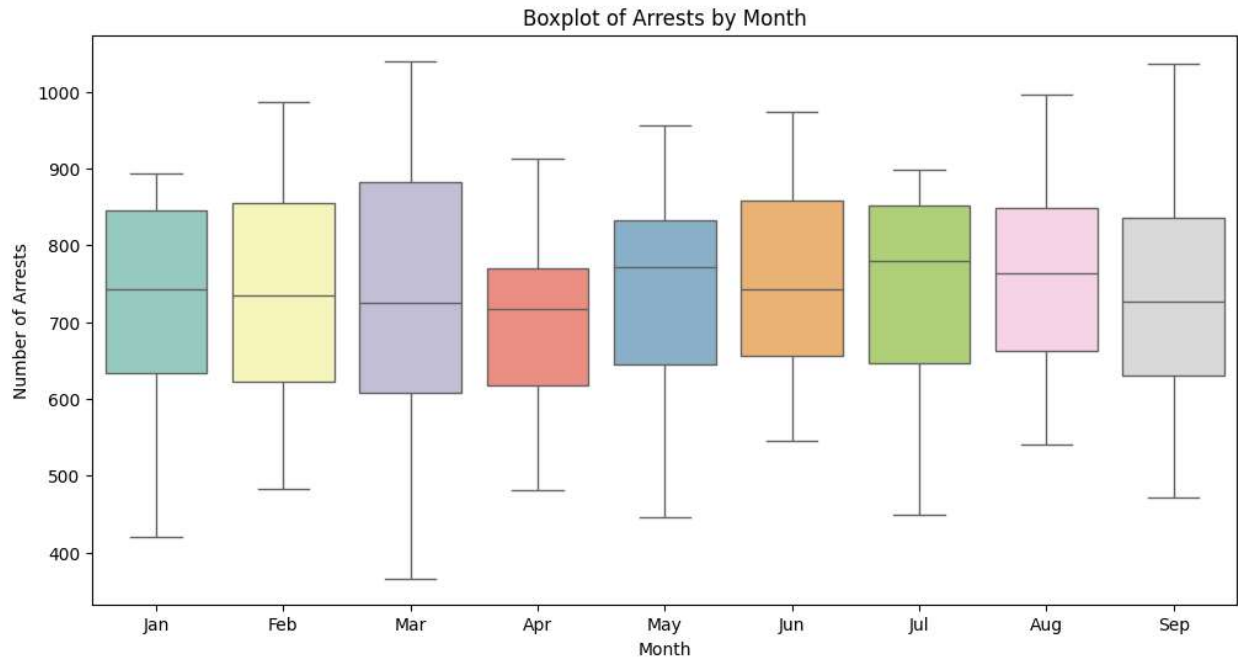
The above graphic is a boxplot of monthly arrest. It helps establish a further understanding of the line plot of daily arrest. It shows the median amount of arrest as well as the upper and lower quantiles. In addition to that the monthly minimums and maximums are shown on the whiskers of the plot.

Some interesting things in this plot are that the median amount of arrest that occurred for the first four months are very close to each other. The 3rd quantiles for January, February, June, July, August, and September are all very close to one another. There appears to be no outliers for any of the months. The month with the lowest minimum appears to be March. March also happened to have the largest maximum number of arrests.

Some expectations I had here were that the months with holidays would have a higher number of arrests, and possibly outliers that occurred. This expectation was proven to be incorrect, in the fact that there are no outliers that occurred in this plot.

I would like to further inquire about other factors that may impact the sample monthly statics here like amount of non-federal holidays in a month impacted results.

Research question #3: Are there more Black people being arrested than Asian people in Queens?

Proportion of Arrests in Queens by Perp Race
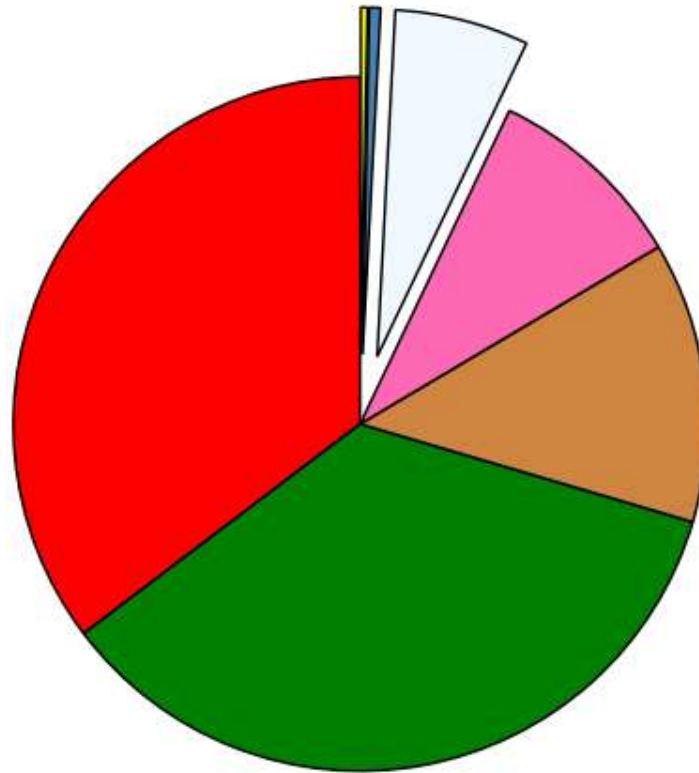
Perp Race

- ■ BLACK: 35.3%
- ■ WHITE HISPANIC: 35.1%
- ■ ASIAN / PACIFIC ISLANDER: 13.1%
- ■ WHITE: 9.4%
- □ BLACK HISPANIC: 6.3%
- ■ AMERICAN INDIAN/ALASKAN NATIVE: 0.5%
- ■ UNKNOWN: 0.3%

	Amount	Race
Max	15181	Black
Min	133	Unknown
Average	6139.714285714285	
Q1	1455.5	
Median	4041.0	
Q3	10356.0	
Upper Bound	23706.75	

The above chart shows the proportions of races arrested in Queens borough. The graphic affirms the posed question (Are there more Black people being arrested than Asian people in Queens?). It is seen on the chart that in Queens in 2024 from January of September 35.3% of arrests made were of Black preparators, whereas only 13.1% peoples of Asian/Pacific Island descent. One note that should be made is that the unknown category is not missing information, it is that the perpetrators race could not be determined.

The max amount of people arrested in Queens by race were Black at 15181 people. The minimum amount was 133 people of unknown race. The average arrests by race were 6139.71. There is no outliers in this set. The upper bound is 23706.75. The lower bound is negative, that cannot occur the as the lowest theoretical arrest that can be made are 0. There are no outliers in this sample.

One expectation I had is that there would be a slightly higher proportion of peoples of Asian descent represented on this chart. Queens is home to Flushing which has a very large community of people from Asian. My expectation was that because there was a large community of Asian people living in Queens, there would have been a higher representation on this graphic. Because of this I would like to investigate further breaking down the arrests precinct to see if this holds true in specific areas within the borough.

Arrests by Race and Sex in Queens
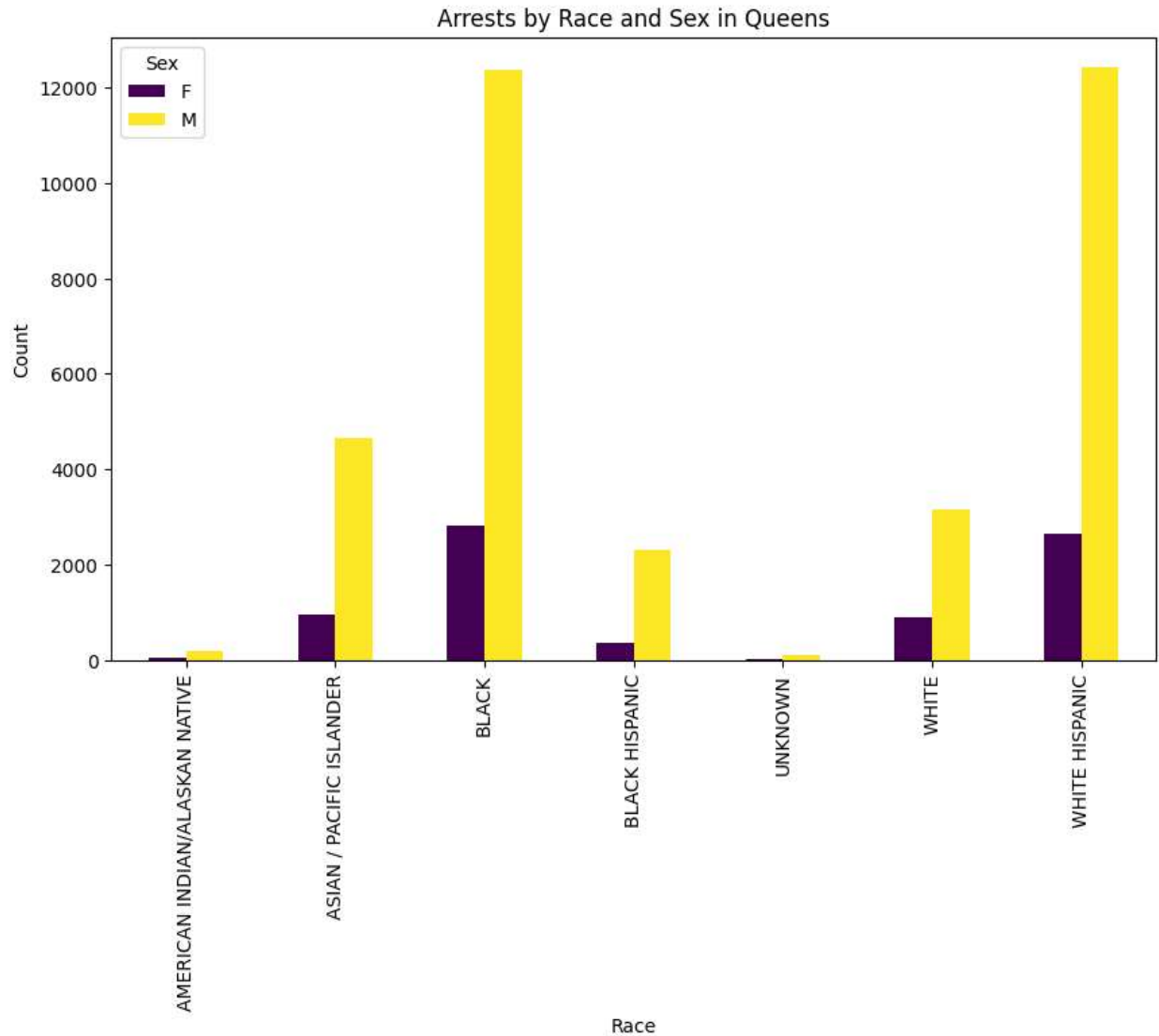
Perp Race	Male count	Female count
AMERICAN INDIAN/ALASKAN NATIVE	183	41
ASIAN / PACIFIC ISLANDER	4653	965
BLACK	12361	2820
BLACK HISPANIC	2326	361
UNKNOWN	117	16
WHITE	3155	886
WHITE HISPANIC	12432	2662

The above graphic further breaks down the race category to further understand the question posed (Are there more Black people being arrested than Asian people in Queens?). It shows a grouped bar graph of each race broken down by sex. This helps breakdown and further understand the question. One thing that would be interesting to investigate here is the total amount of males and females broken down by these races. Something that was unexpected here was how close the amount of white Hispanic women was to the number of black women. From the previous graphic, pie chart, it was seen that the proportion was very similar for two races, it is interesting to note that it seemed to hold true through the sexes as well.

The lowest amount of people arrested were females of unknown races at 16 people. The highest was that of black men. The lowest amount for men was that of people of unknown race at 117 and for females the unknown race as well. I am curious what entails someone being classified as unknown race as there are many races such as middle eastern, or Jewish descent not being included as options.

One expectation I had for this graphic is that there would be a higher arrest amount for men vs. women. This expectation was proven to be true as for each race there are more men that were arrested than there was female. Something that would require further study is the types of crimes being committed broken down by sex as well.

Project Synopsis Tweet:

Black people are more frequently arrested than Asians in queens! Aren't there more Asians in Flushing? Federal holidays decrease the amount of crime occurring EXCEPT on Juneteenth; INCREASE then! Brooklyn has a higher number of arrests than the Bronx! The poverty rate factor?

Seven Keywords:

1. **Crime**
2. **Arrests**
3. **Race**
4. **Borough**
5. **NYPD**
6. **NYC**
7. **Holiday**

Github link:

https://github.com/emirbeg2017/505project