Inside Investigators

Devesh Shah

Libby Gough

Sierra Stubs

William Creamer

Introduction

<u>Research Question:</u>

What is the relationship between borough and perpetrator sex for violent crimes committed throughout all 5 boroughs?

<u>Motivation:</u>

Our project aims to analyze any potential trends between boroughs, types of violent crimes committed, and perpetrator demographics in the New York City Metro Area. We believe modeling these relationships could help understand crime characteristics within the boroughs and potentially help prevent future crime.

<u>Sourcing:</u>

Our data comes from NYC OpenData, a website operated by the NYC Office of Technology and Innovation. The team works with city agencies to compile data and make it available to the public.

Data

Unmodified Data:

- 19 unique variables & 227,000 observations of NYC arrests in 2023
- 64 different types of crimes were reported → focused on 7 types of **violent crime**
  - Highly reported & discussed
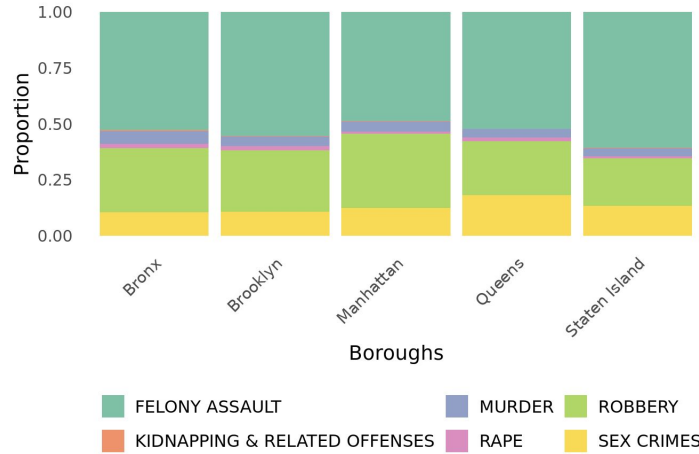 - ~40,000 observations

Data Cleaning:

- 7 types of violent crime
  - Combined "sex crime" and "felony sex crime"
- Expanded acronyms of borough names
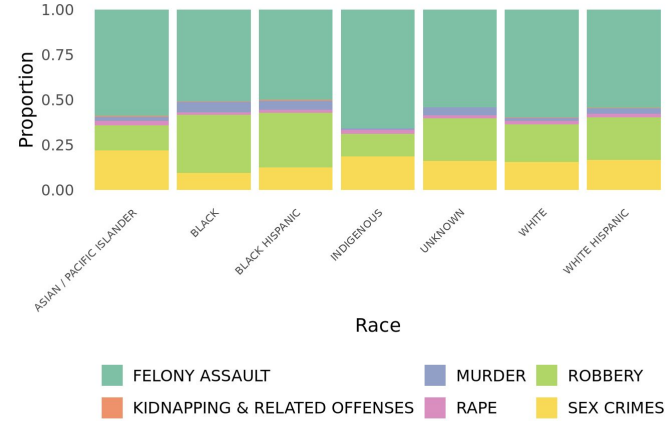- Removed observations with an unidentified perpetrator sex

Variables:

- Explored the type of violent crime, perpetrator sex, age, and race, and borough of arrest
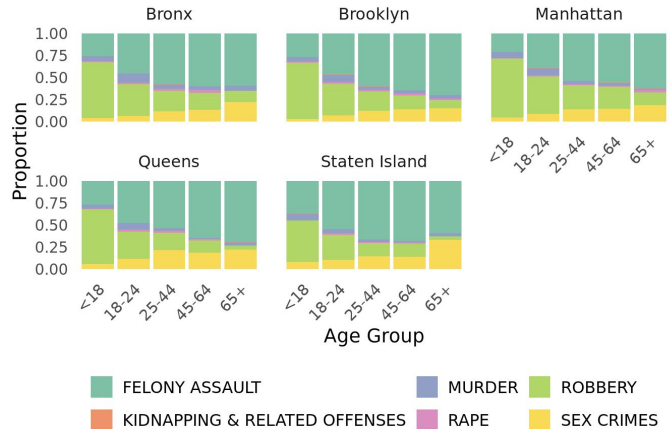
Violent Crimes in NYC by Borough
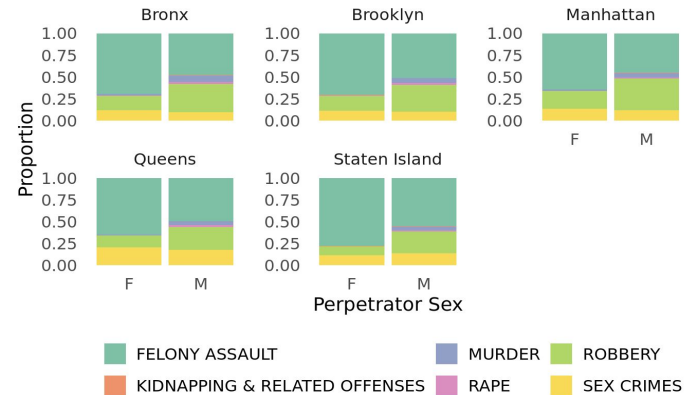
Violent Crimes by Race

Violent Crime Proportions by Age Group by Borough

Violent Crime Proportions by Perpetrator Sex by Borough

Legend: FELONY ASSAULT, KIDNAPPING & RELATED OFFENSES, MURDER, RAPE, ROBBERY, SEX CRIMES

Methodology

Motivation:

We built a logistic regression model predicting the gender of the perpetrator based upon the arrest borough. We were interested in seeing if there were any gender disparities in violent crimes across the New York City metropolitan area. We wanted to challenge any potential bias we have may have for violent crimes.
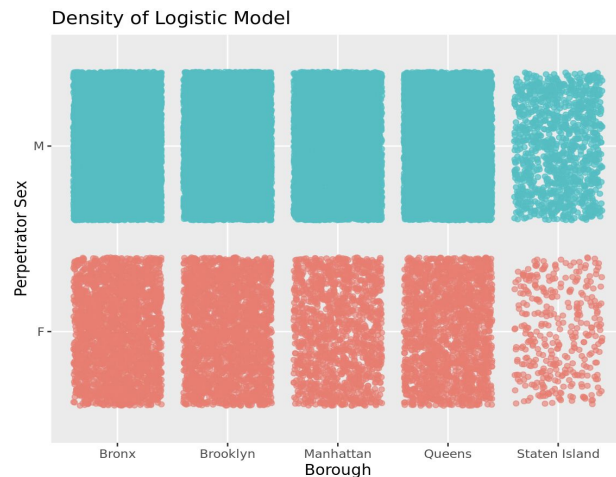
Null Hypothesis:

There is no significant difference in the proportion of felonies committed by female and male perpetrators across the different NYC boroughs.

Alternative Hypothesis:

There is a significant difference in the proportion of felonies committed by female and male perpetrators across the different NYC boroughs.

term	estimate	std.error	statistic	p.value
(Intercept)	1.204	0.023	52.006	0.000
arrest_boroBrooklyn	0.037	0.033	1.121	0.262
arrest_boroManhattan	0.254	0.037	6.843	0.000
arrest_boroQueens	0.183	0.035	5.259	0.000
arrest_boroStaten Island	-0.057	0.065	-0.867	0.386

Density of Logistic Model

Conclusions and Future Work

<u>Results:</u>
- We received a p-value that was reported to 0, which could indicate the actual p-value is extremely small, thus we can reject the null hypothesis. The data provides sufficient evidence that there is a difference between the proportions of perpetrator sex committing felonies, which is also supported by prior visualizations.
- Overall, more violent crimes were committed by male perpetrators across nearly all categories except for Sex Crimes within the different boroughs.

<u>Considerations - Reliability & Validity:</u>
- When handling demographic data, there are biases that can arise in the collection and production of the data.
- Minority communities and lower income neighborhoods tend to have heavier policing which might be reflected in the data had that been a parameter explored.
- Amongst the different perpetrator demographics, prejudice can lead to different arrest rates, such as through racial profiling
- Building a model off of such data would inherently possess biases, and combating such requires systemic changes to be made to law enforcement practices.

<u>Future Work:</u>
- Avenues for future work include considering multinomial logistic regression models, which could be used to predict types of violent crime committed based on any demographic characteristic, borough location, or other independent variable in the dataset.