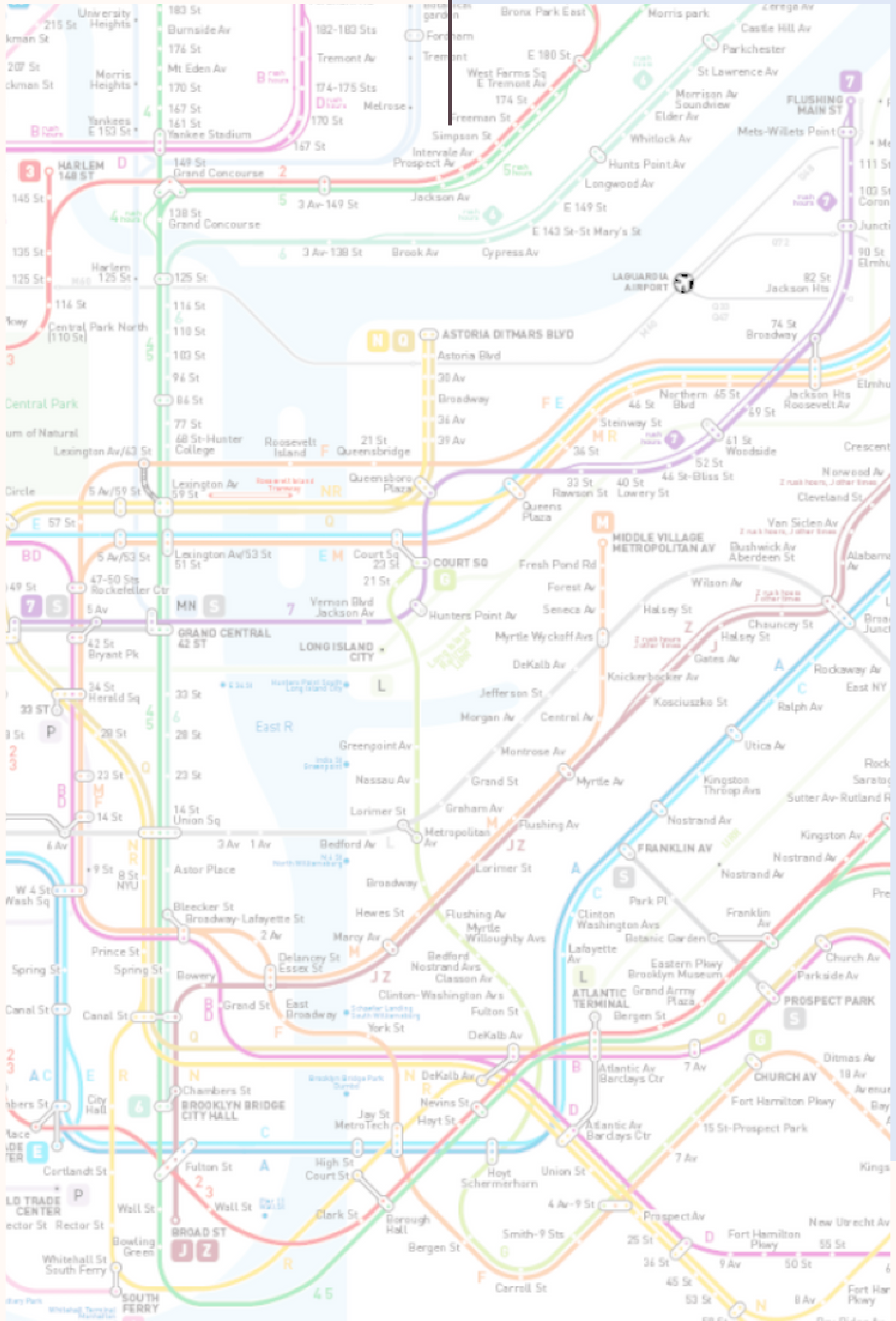


EDA Project

Background

New York city consists of 5 boroughs: Manhattan, Brooklyn, the Bronx, Queens, and Staten Island. It has the highest population density of any major city in the United States. Recently, the crime figures are threatening to undermine the sense of safety in the city according to police statistics.

This project helps the police department to find the most dangerous stations in NYC so they can increase the level of surveillance over these stations and the areas nearby so that the crime rate will be lower.

Summary of Steps

Determine the crime rate in each borough in NYC.

Determine the hours during which the crimes rate is high in the most dangerous borough.

Determine which stations are the busiest in the most dangerous borough during those hours.

Database and Datasets

A database instance has been created using amazon RDS.

Datasets:

- MTA Turnstile data July-September 2021.
- Crimes Dataset from New York City Police Department.
  Information about crimes that happened in NYC from 2012 until 2019.
- NYC Turnstile Stations List.
  Information about the stations in NYC.

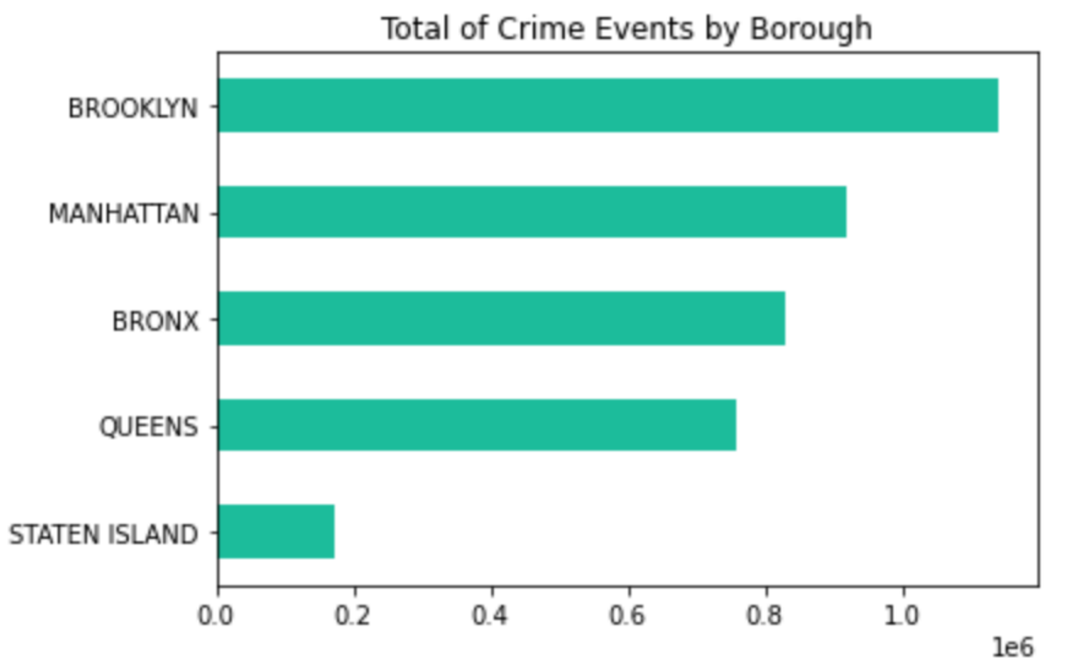
Data Cleaning

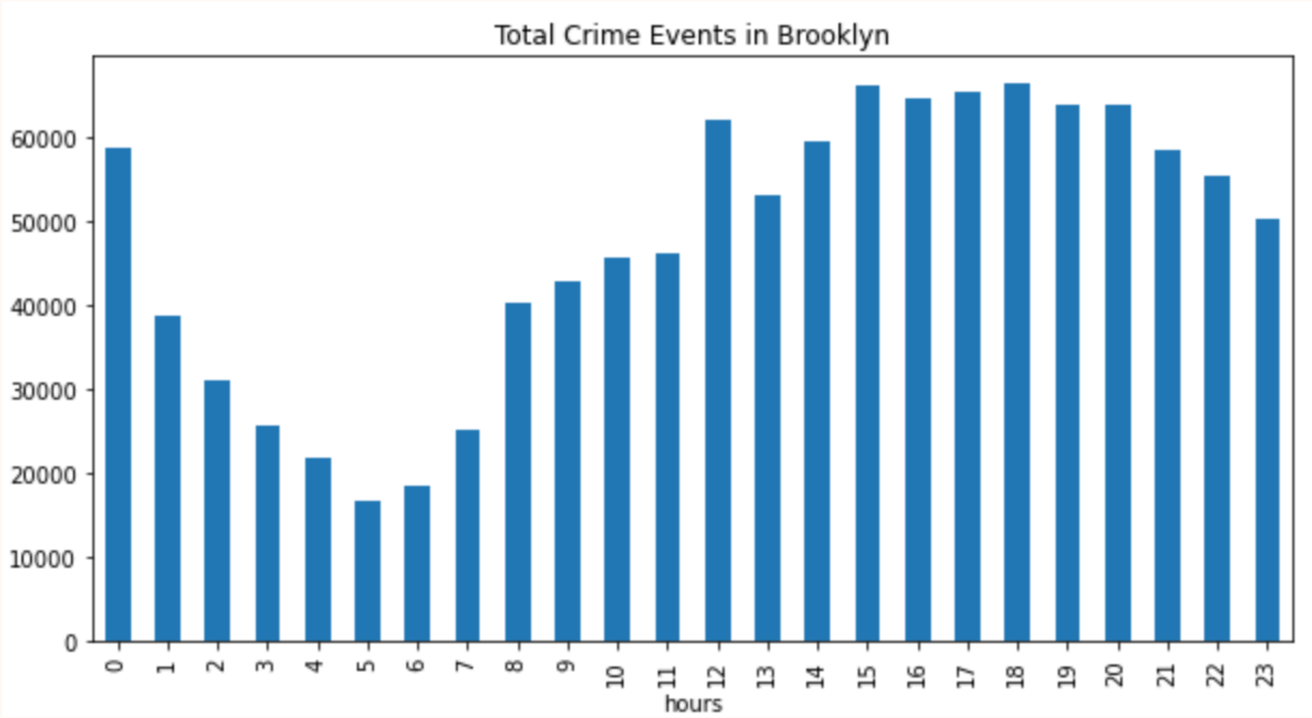
- For the MTA turnstile:
  Consistency in column names.
  Dropping duplicate rows
  Adding columns for the entries and exits per for hours for each turnstile.
  Fixing the outliers in the new values for entries and exits.

- For the crimes data:
  Dropping columns containing a huge number of missing values and are not important for the analysis.
  Fixing inconsistencies in the values of one of the columns.
  Dropping the rows containing missing values.

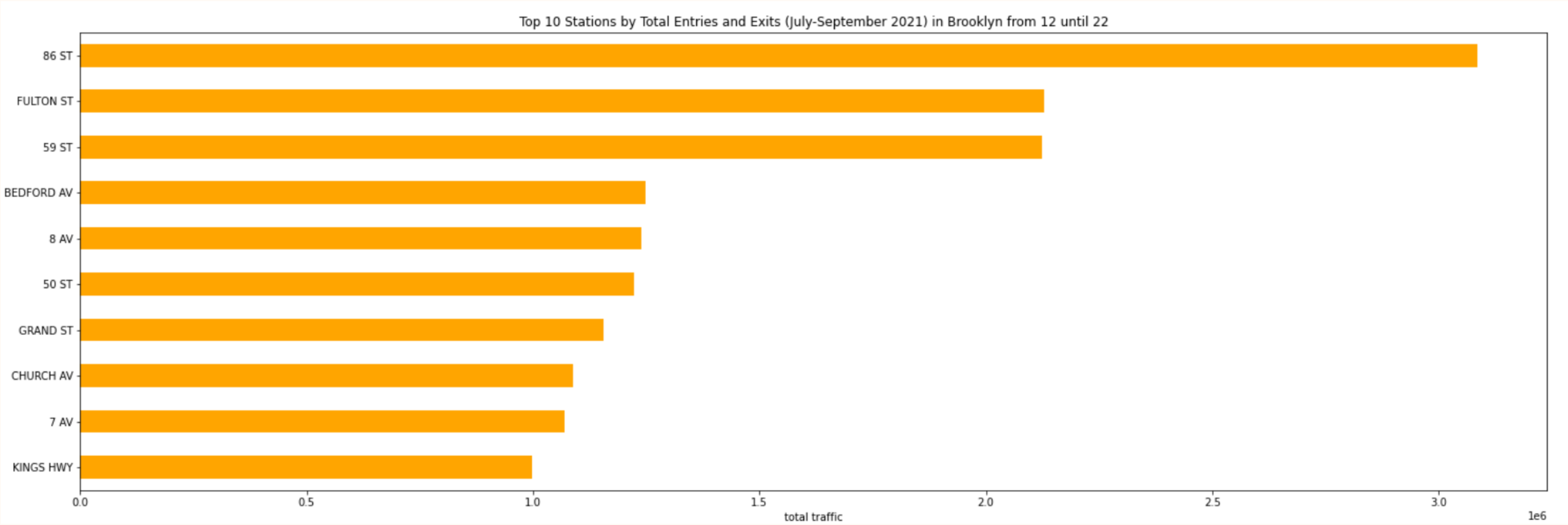
- For the stations list dataset:
  Dropping columns containing missing values and are not important for the analysis

Visualizations





Visualizations



Conclusion

Based on the findings, I highly recommend the police department to increase the number of security guards in the stations below (ordered from the most dangerous):

1. 86 ST.
2. FULTON ST.
3. 59 ST.
4. BEDFORD AV
5. 8 AV.

And during the hours from 12 to 22.