EDA Project

Background

New York consists of 5 boroughs: Manhattan, Brooklyn, the Bronx, Queens, and Staten Island. It is known that NYC is one of the most dangerous cities in the US.

This project will help the police department to find out what is the borough with the highest crime rate, which hours have the highest crime rate, and what stations are the busiest in this borough during these hours in order to increase the surveillance level in these stations.

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.
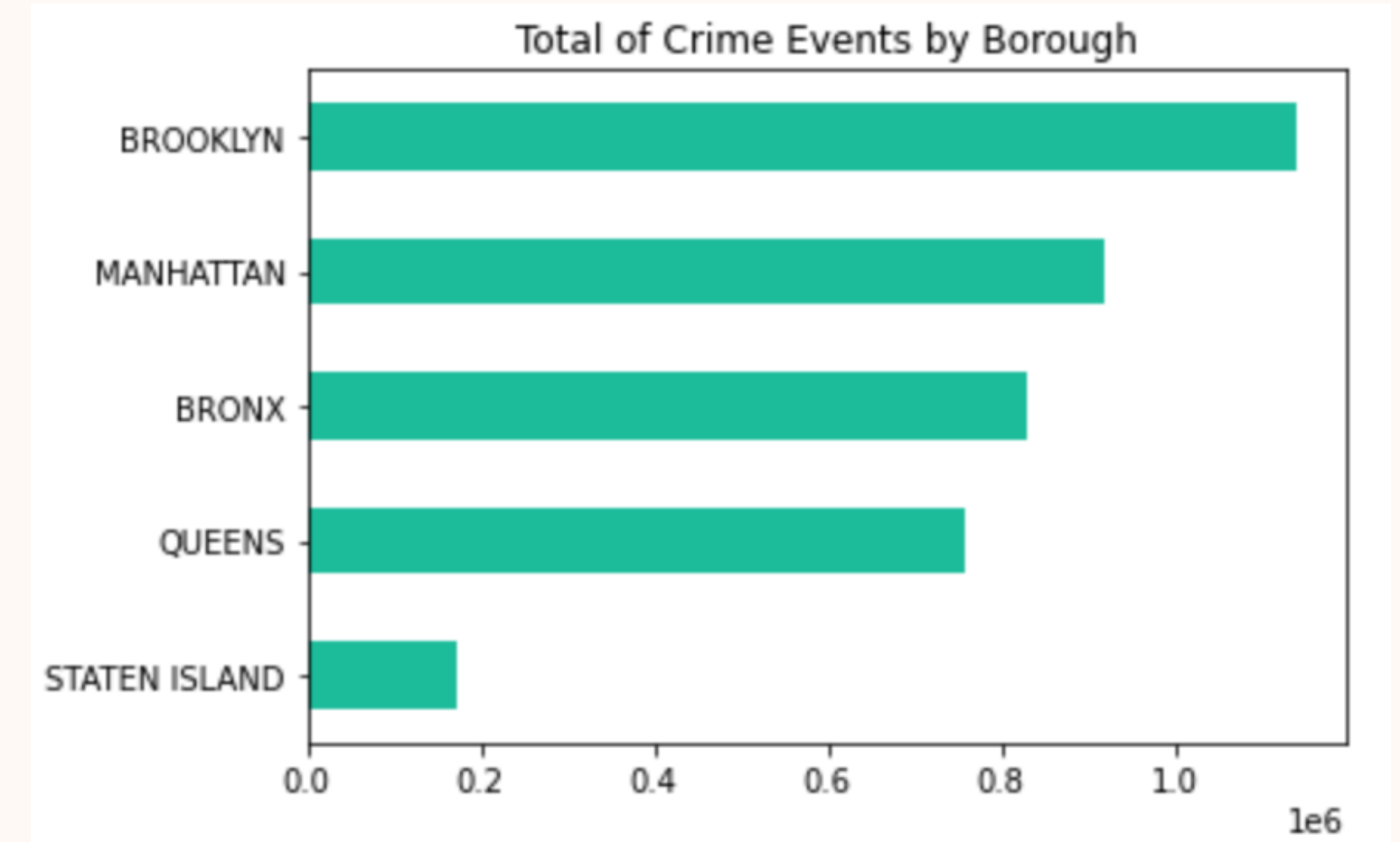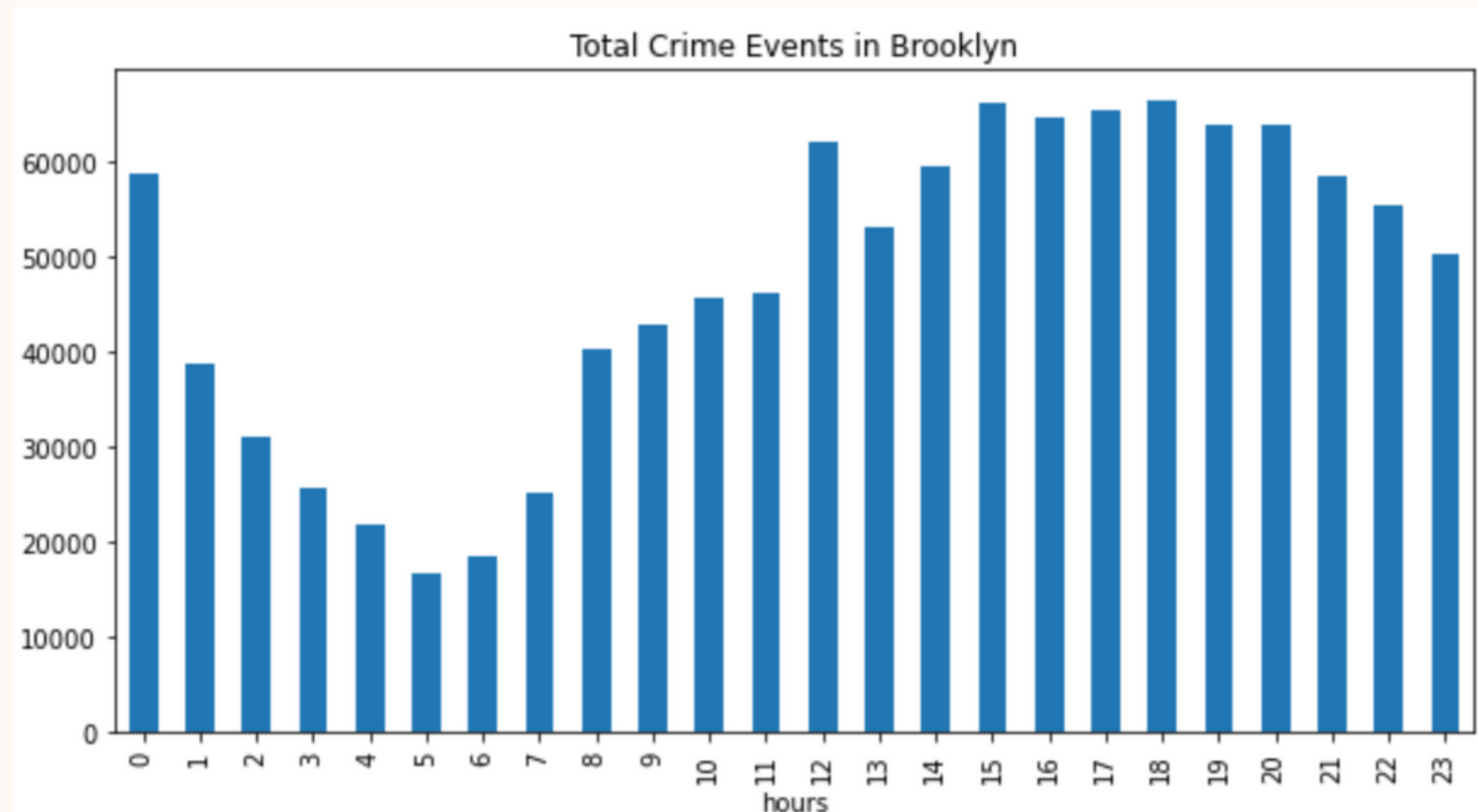 Information about the stations in NYC.
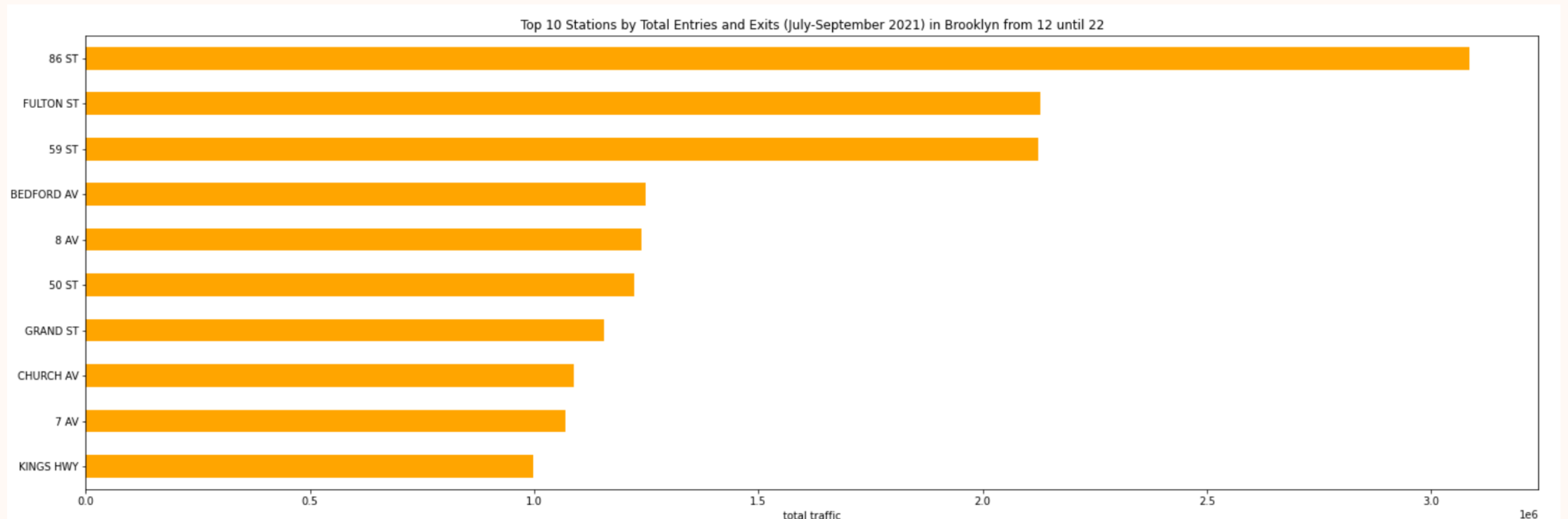
Data Cleaning

- 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

Total Crime Events in Brooklyn

Total of Crime Events by Borough

Visualizations

Top 10 Stations by Total Entries and Exits (July-September 2021) in Brooklyn from 12 until 22

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 likely that crimes will occur in):

1. 86 ST.
2. 59 ST.
3. FULTON ST.
4. 8 AV.
5. 50 ST.

And during the hours from 12 to 22.