

**Data Science Bootcamp
SADAIA**

**Project Report
securing MTA stations**

Norah Alqahtani

Summery:

Metropolitan Transportation Authority (MTA) has been receiving a lot of complaining messages from commuters that there are many stealing and harm events in stations, and they realize that their stations need more securing controls. But unfortunately, they don't have enough budget to do so. So, they decided to hire Security First Agency (SFA) and ask them to make MTA's stations more secure with less cost as possible.

As I am part of Security First Agency (SFA) team that they responsible for this project, we take advantage of publicly available Metropolitan Transportation Authority (MTA) Turnstile Dataset. We select 12 weeks from July to September and analyze it to come out with proper recommendations to make Metropolitan Transportation Authority (MTA) more secure.

Assumptions:

- We assume that we should focus on the top 7 most crowded stations since the budget is limited
- The security controls that will be provided are security cameras, security men, and Warning panels. In additions more advanced turnstiles in the most crowded units in the top 7 stations. Also Alarm devices link the stations to the concerned authorities will be provided.
- We assume that when the traffic increases the need of security controls will increase too

Question/need:

- How to make MTA's stations more secure as possible with less cost as possible?
- What are the top 7 most crowded stations to focus on?
- What are the days of the weak that had more traffic in each station?
- Security First Agency (SFA) that hired by Metropolitan Transportation Authority (MTA) for securing their stations and commuters.

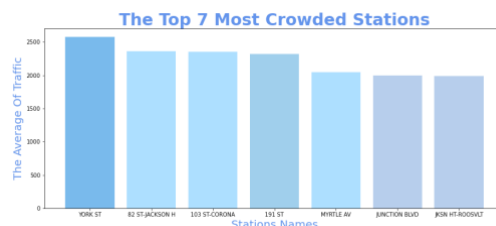
Data Description:

- We planning to benefit from the publicly available Metropolitan Transportation Authority (MTA) Turnstile Data from New York City.
<http://web.mta.info/developers/turnstile.html>

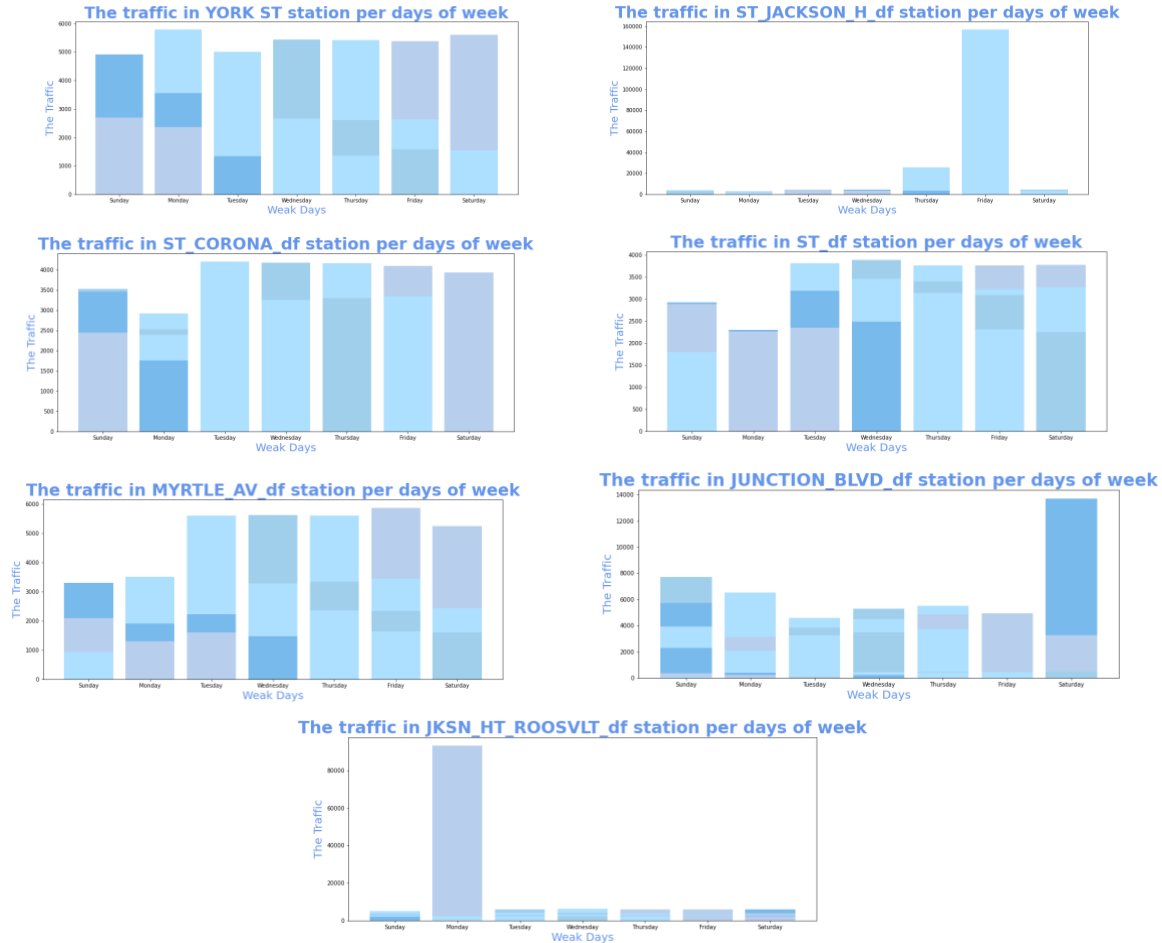
Tools:

- SQL database, SQLAlchemy and SQLite
- Exploratory data analysis in pandas
- Python visualization libraries (such as matplotlib and seaborn)

Communications:



Here in figure 1 we have selected the top 7 stations



From figure 2 to Figure 8 the most active days in each station have been determined

Recommendations:

- The focus of our project needs to be on the top 7 stations which are YORK ST, 82 ST-JACKSON H, 103 ST-CORONA, 191 ST, MYRTLE AV, JUNCTION BLVD, and JKSN HT-ROOSVLT
- The YORK ST station will be provided by most security controls as possible in all 7 days of week specially security men and Alarm devices link the stations to the concerned authorities
- In ST_CORONA, 119 ST and MYRTLE_AV the days that will be focused on is Thursday, Wednesday, Thursday, Friday, and Saturday
- Unlike T_JACKSON_H station which will be having more focus only on Friday. And JKSN_HT_ROOSVLT station on Monday only
- And JUNCTION_BLVD station was more crowded in weekend (Saturday and Sunday)

Future work:

- We are planning to propose a schedule of security men shifts that depend in the traffic by time instead of depending only on days
- Also, we were planning to select the most crowded Control Area in each station to get it more attention when applying security controls