Data Analysis for Alsayegh Charitable Foundation

Nisreen Alsayegh

Back story:

Al-Sayegh Charitable Foundation wants to establish a charity market during the summer season Alsayegh charitable foundation wants to take advantage of crowded stations to establish the market near to one of the train stations.

Dataset

MTA: data files containing numbers of cumulative entries and exits and stations, turnstile, dates and time

Dataset: from June 22, 2019 to September to 28, 2019. Because it is the duration that corresponds with the summer season .

- Rows 3100324
- columns 11

Tools :

Python

Jupyter notebook

Matplotlib
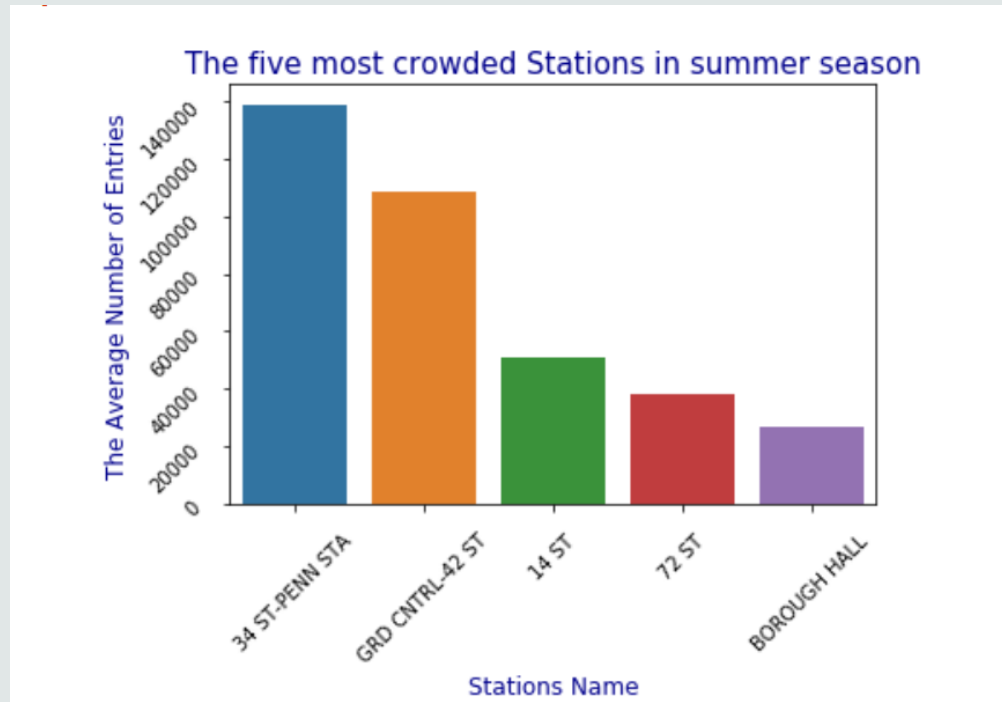
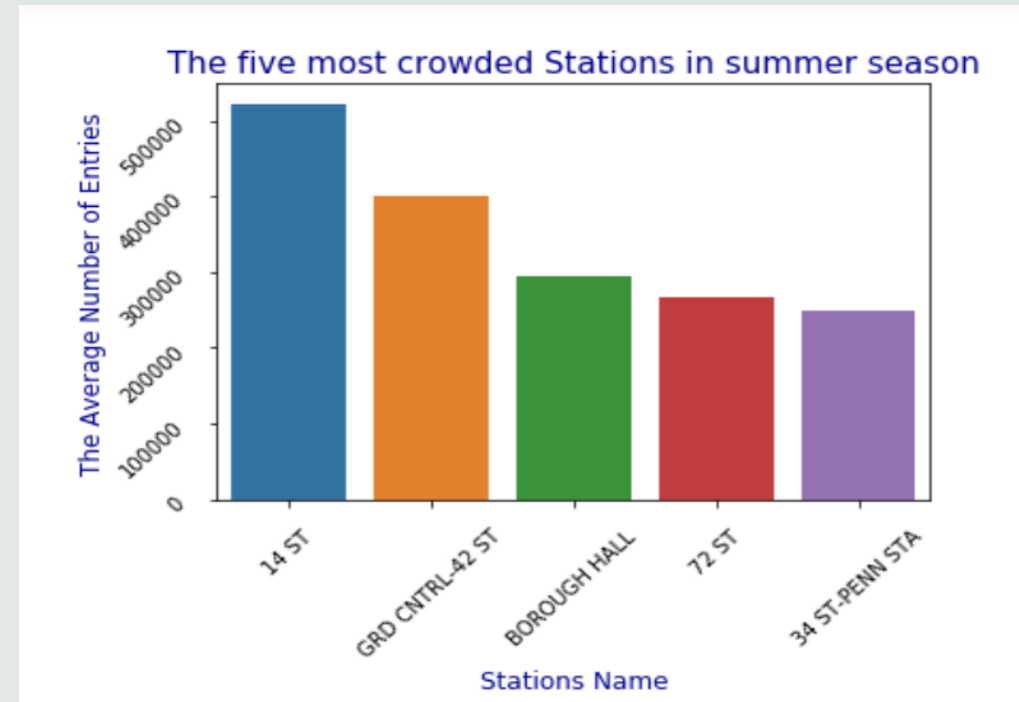NumPy

SQLAlchemy

Seaborn

Pandas

Cleaning data :

- Check nulls: no nulls

- Fix Columns : Strip column names

- Add new columns : DATE and Daily Entries After

- Check Duplicates: drop the duplicate
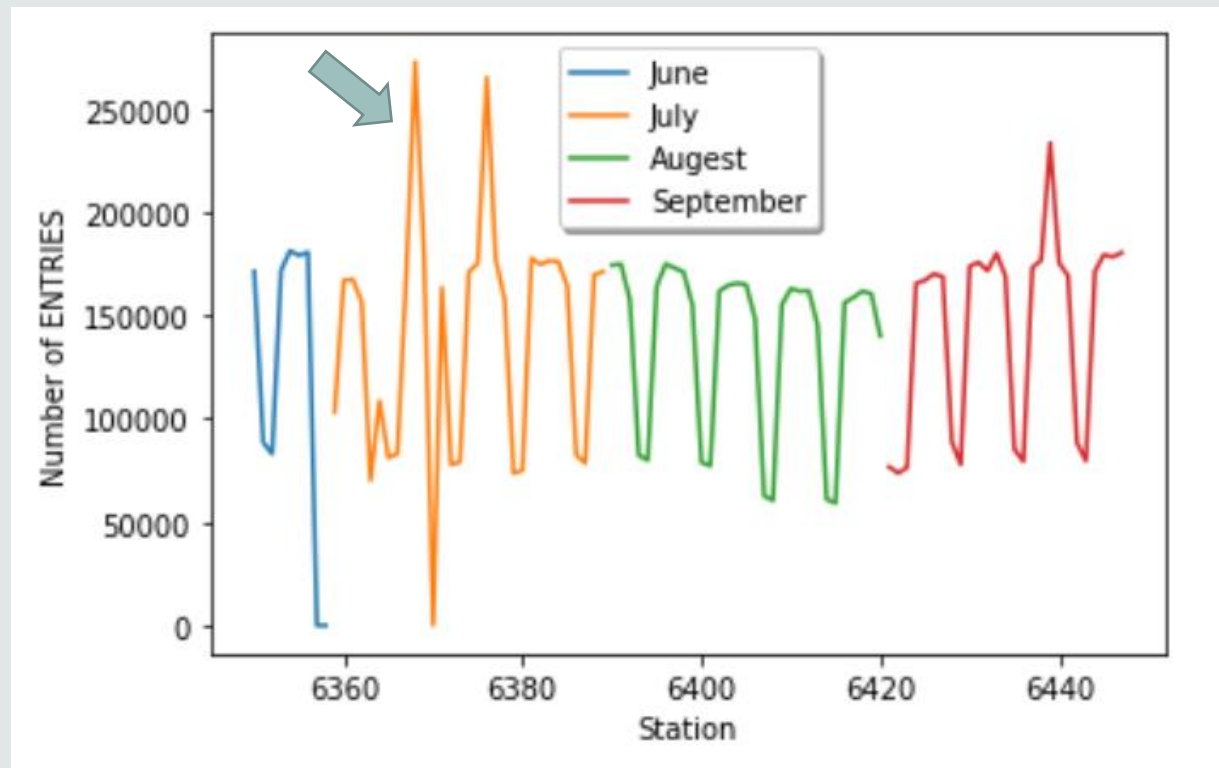
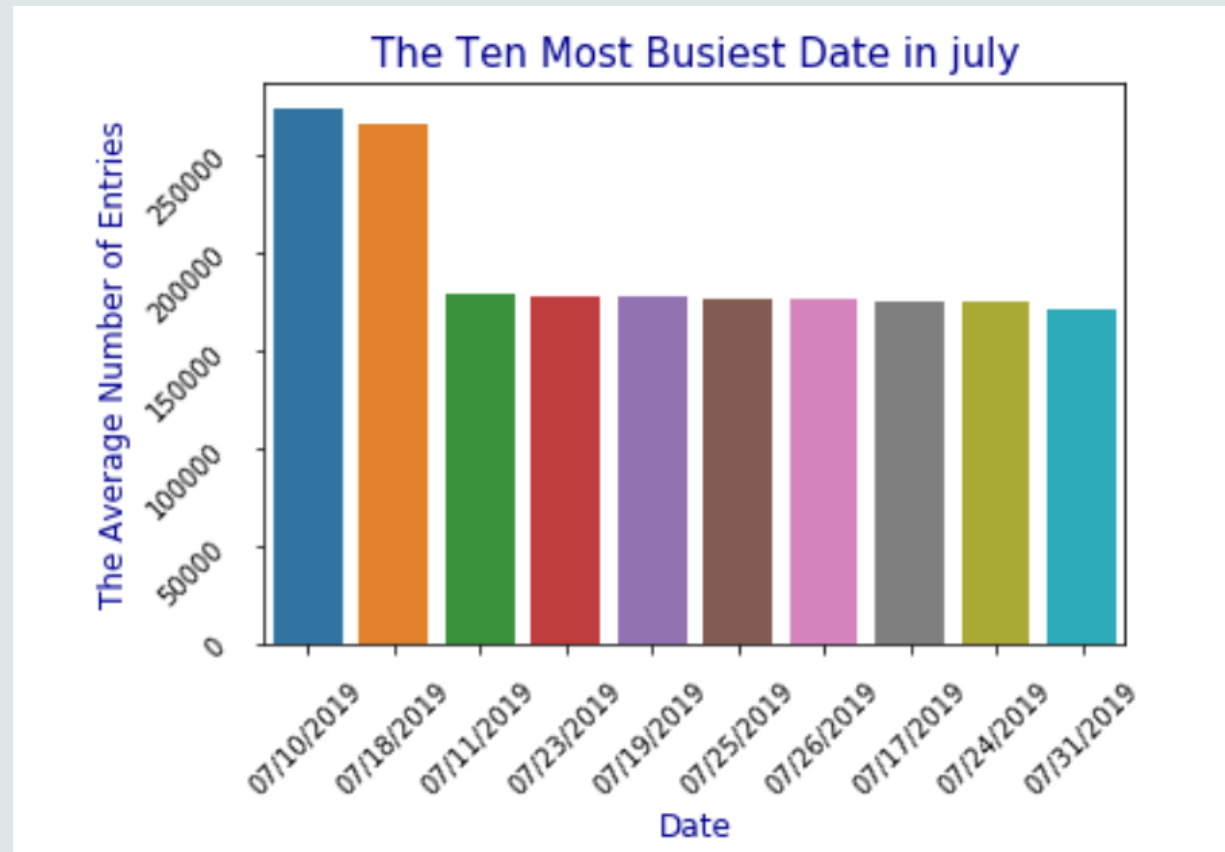# Bar chart before and after cleaning the data

After

Before

The most crowded month

The most crowded date

The Ten Most Busiest Date in july

Conclusion:

- We suggest to alsayegh charitable foundation to establish the market in 10th of July near to 34 ST-PENN STA that located at the intersection of 34th Street and Seventh Avenue.

Conclusion:

- We suggest to alsayegh charitable foundation to establish the market in 10th of July near to 34 ST-PENN STA that located at the intersection of 34th Street and Seventh Avenue.

Future work :

- Determine the best time to establish the market
- Get additional data to study the income for the residents of the area

Thank you