

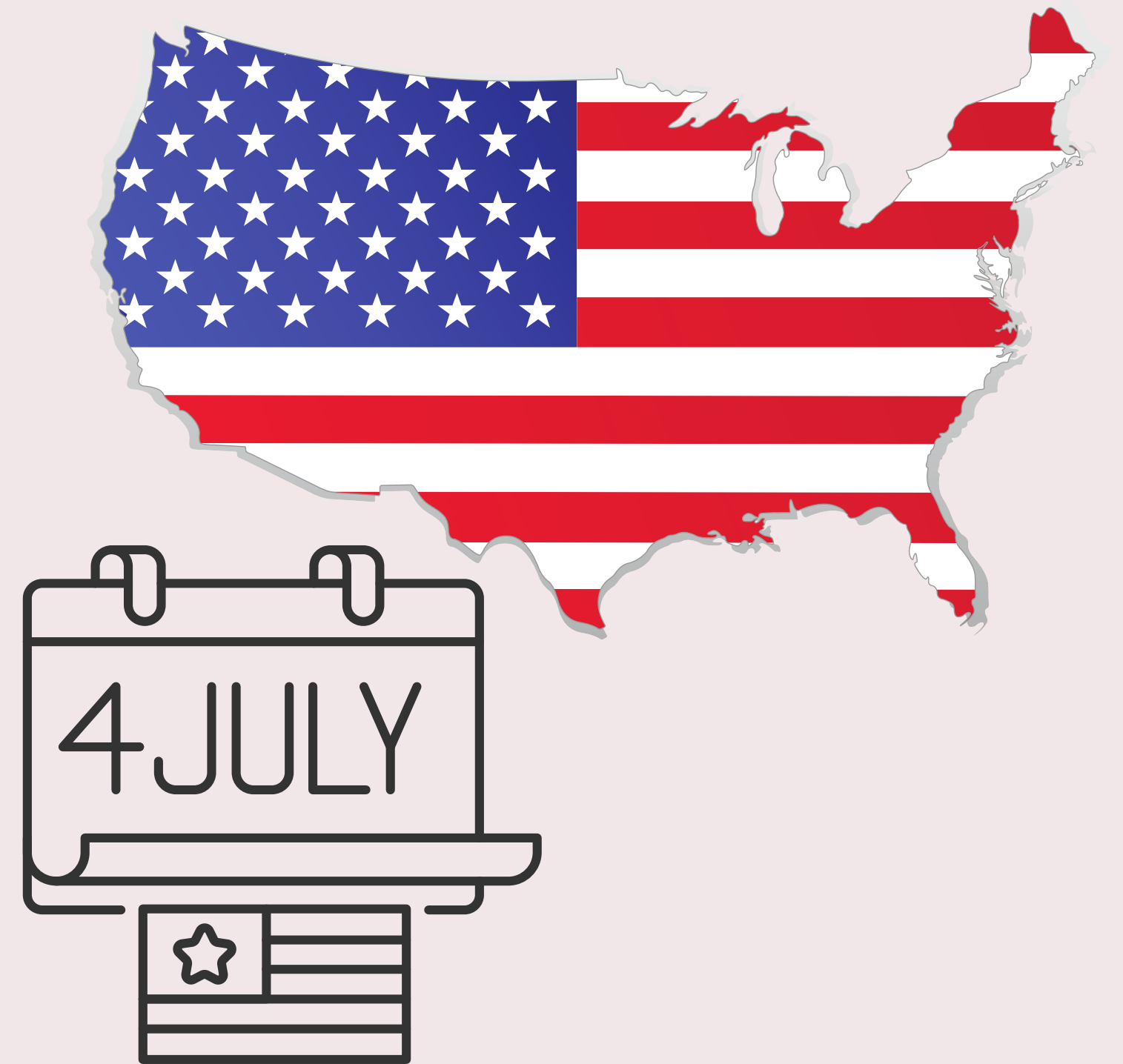
ACTIVITIES BOOTH

BY: AMJAD ALTHINAYYAN



BACKSTORY

Sitting booth to sale
in the MTA before
Independence Day



DATASET

- MTA turnstile data
 - (June - July) from 2016-2019
 - 2383351 row , 11 columns
-
- Stations data to mapping of subway stations to boroughs



TOOLS

- SQLite
- Python
- Jupyter notebook
- SQLAlchemy

- Pandas
- Numpy
- Matplotlib
- Seaborn



DATA CLEANING



DATA CLEANING

- Check nulls

- Modify & add needed columns

- Remove duplicates



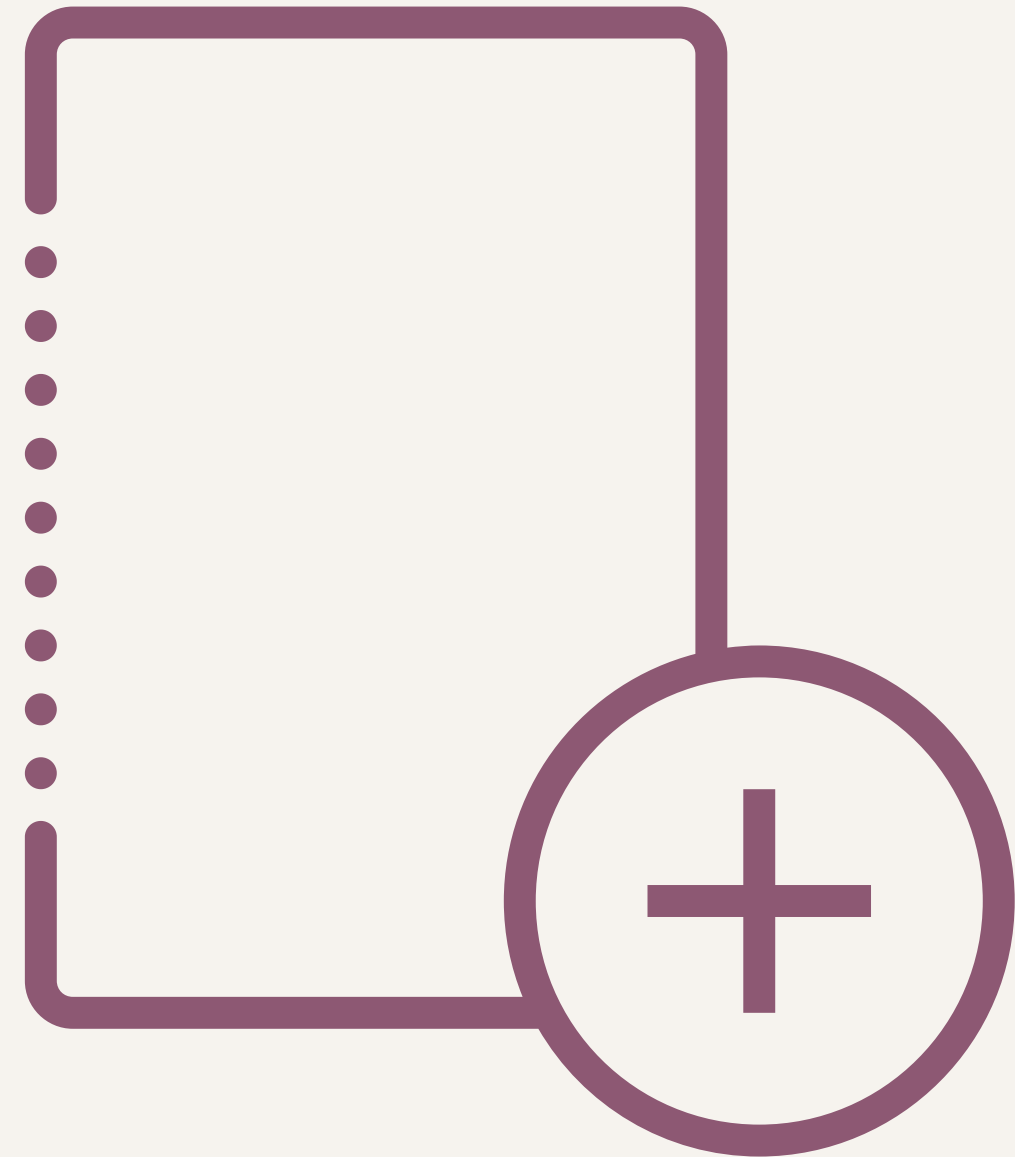
DATA CLEANING

- Check nulls
- Modify & add needed columns
- Remove duplicates



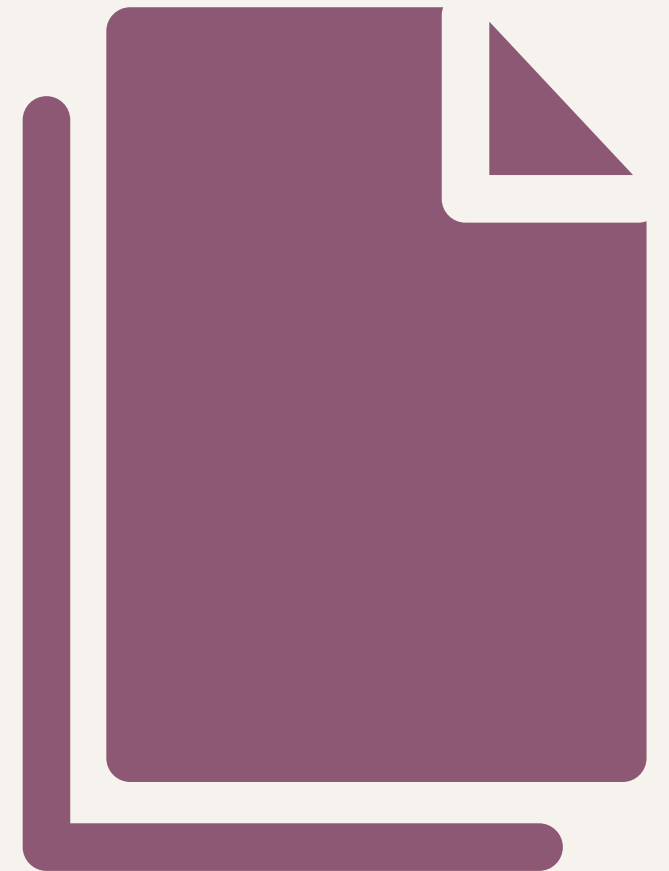
DATA CLEANING

- Check nulls
- Modify & add needed columns
- Remove duplicates



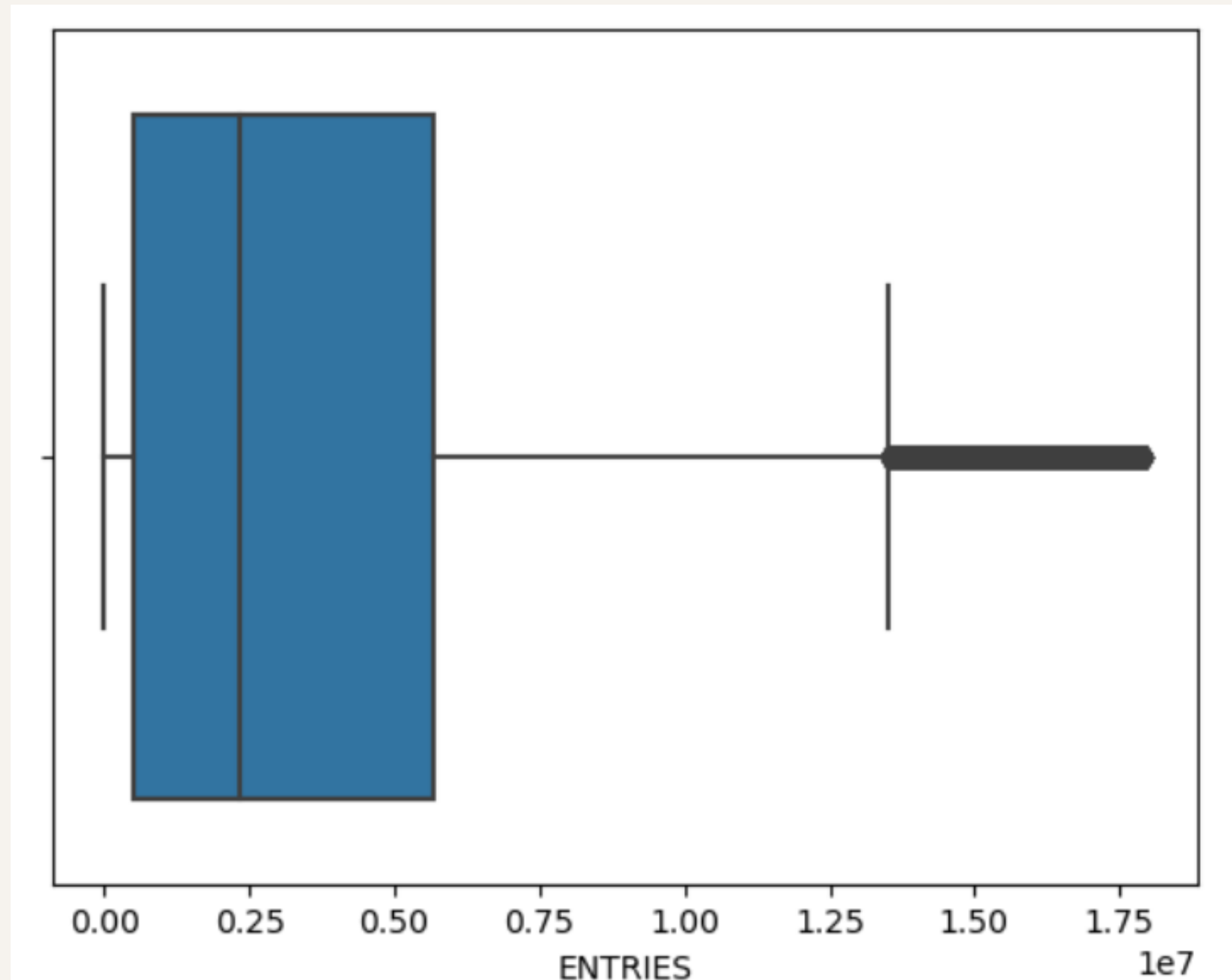
DATA CLEANING

- Check nulls
- Modify & add needed columns
- Remove duplicates



DATA CLEANING

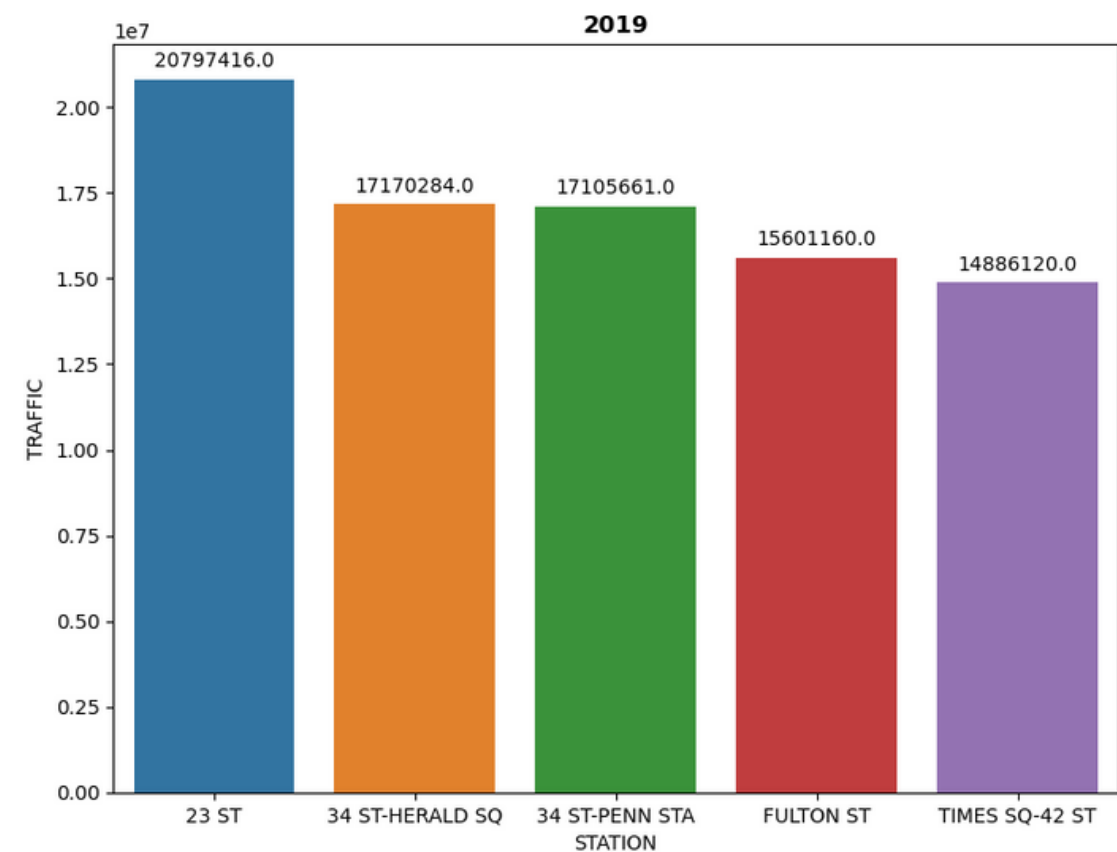
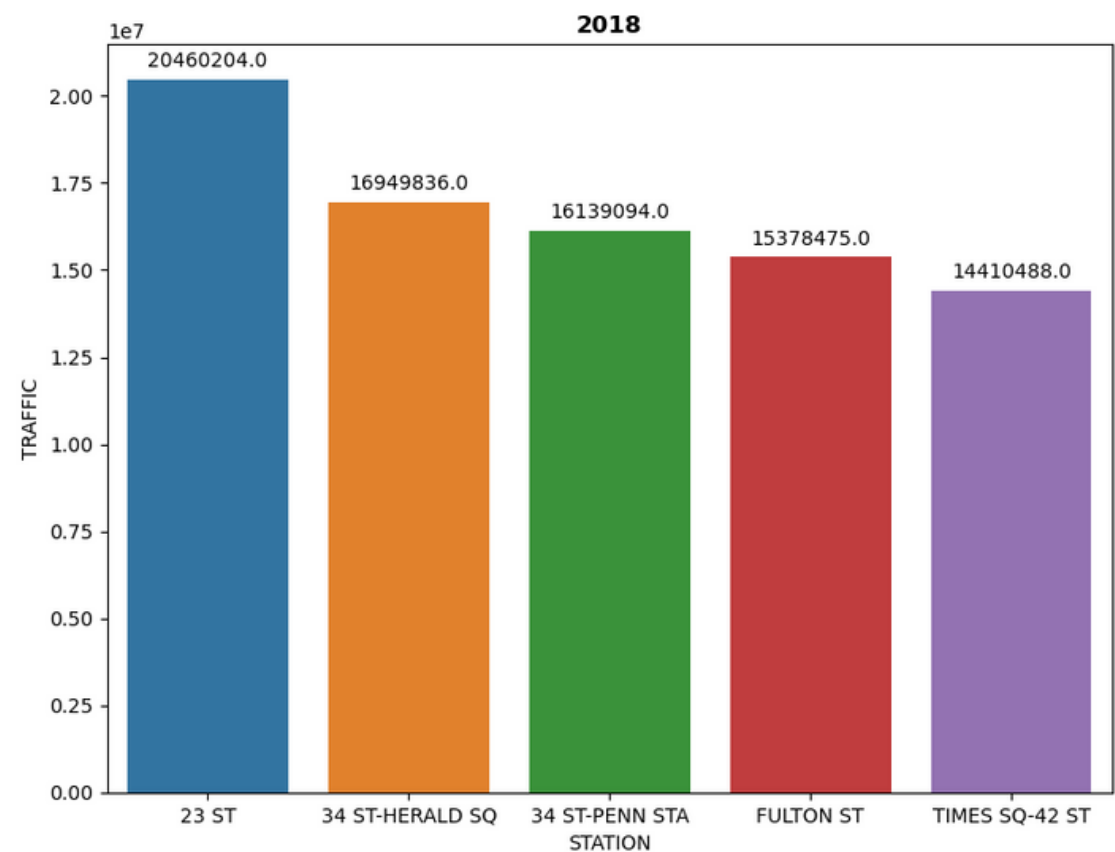
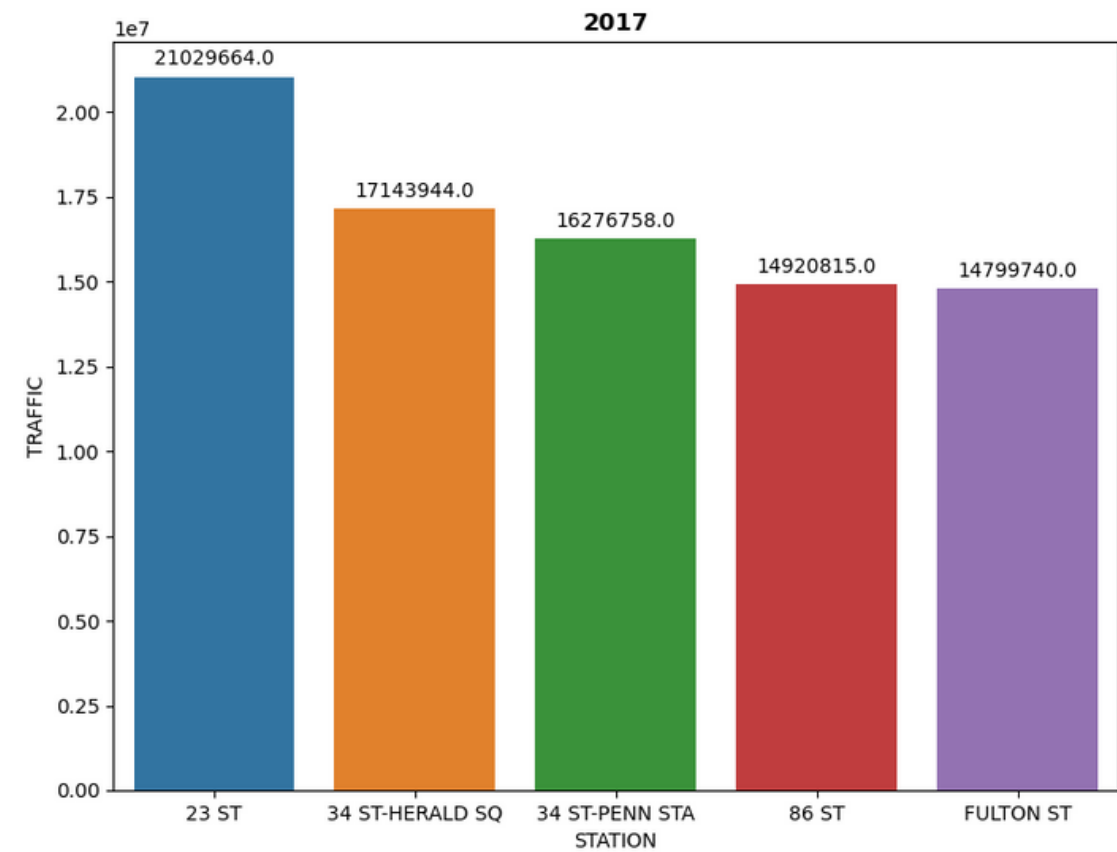
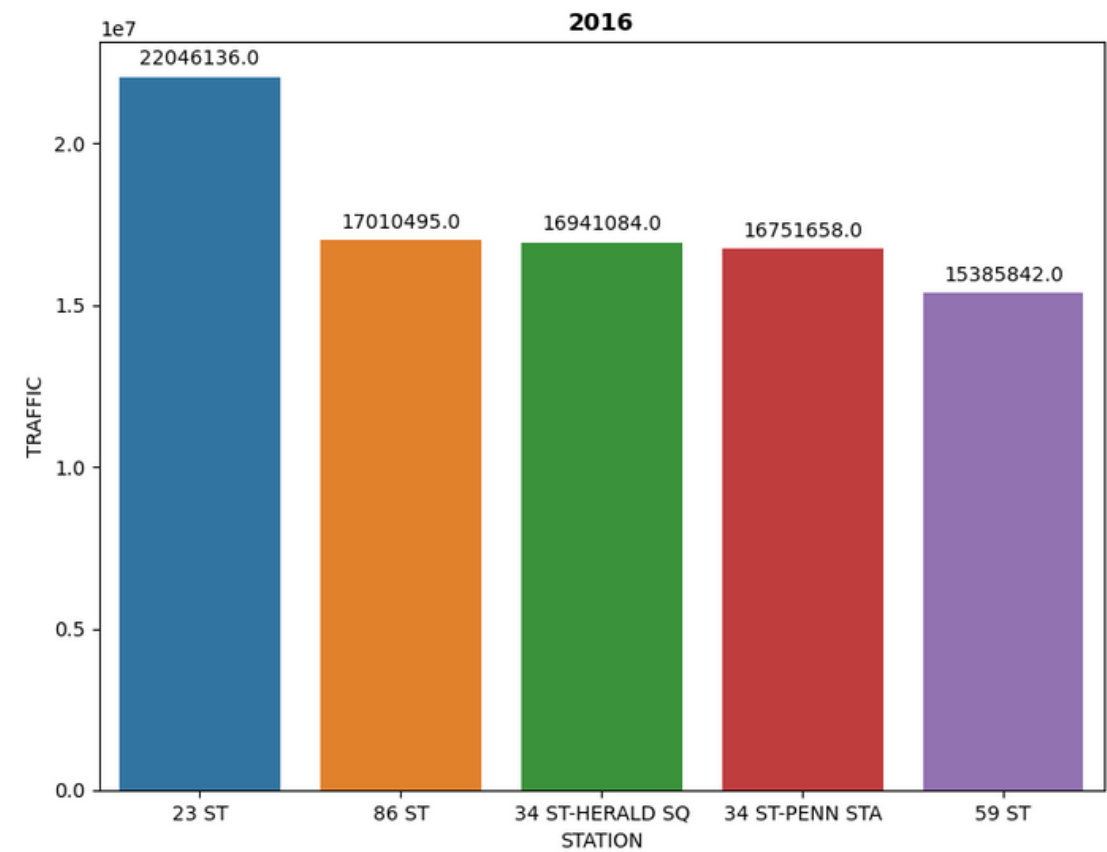
Outliers



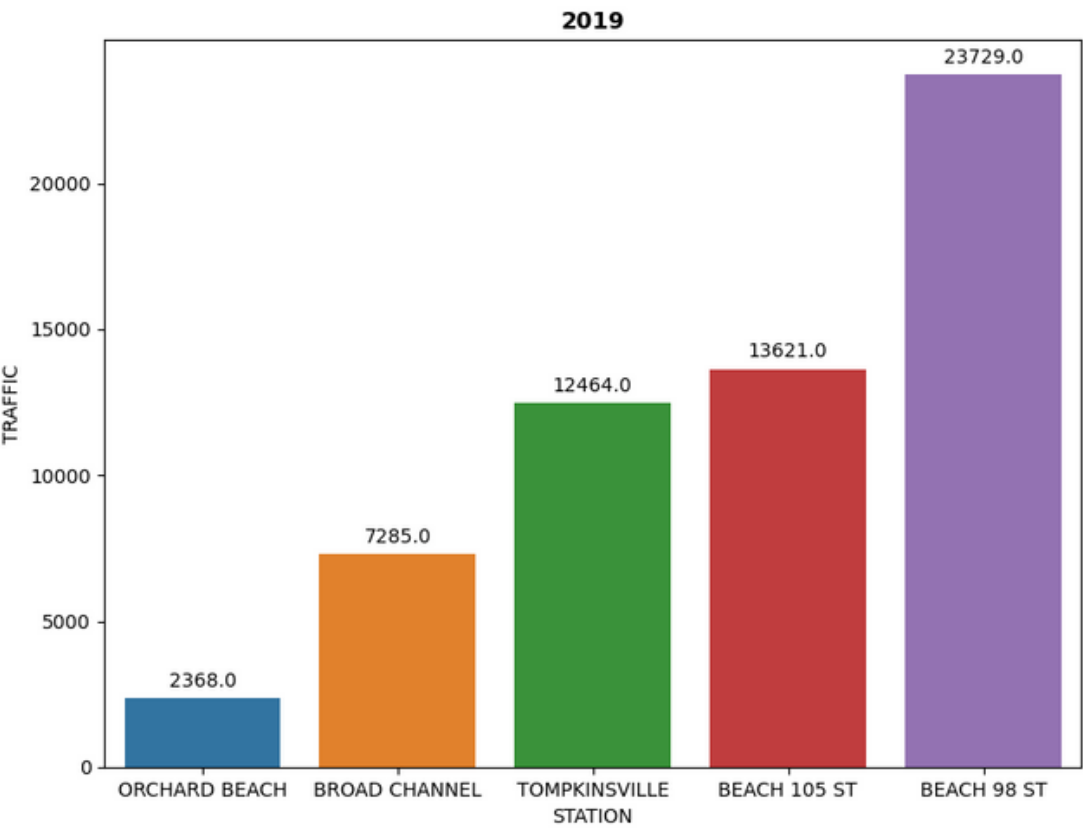
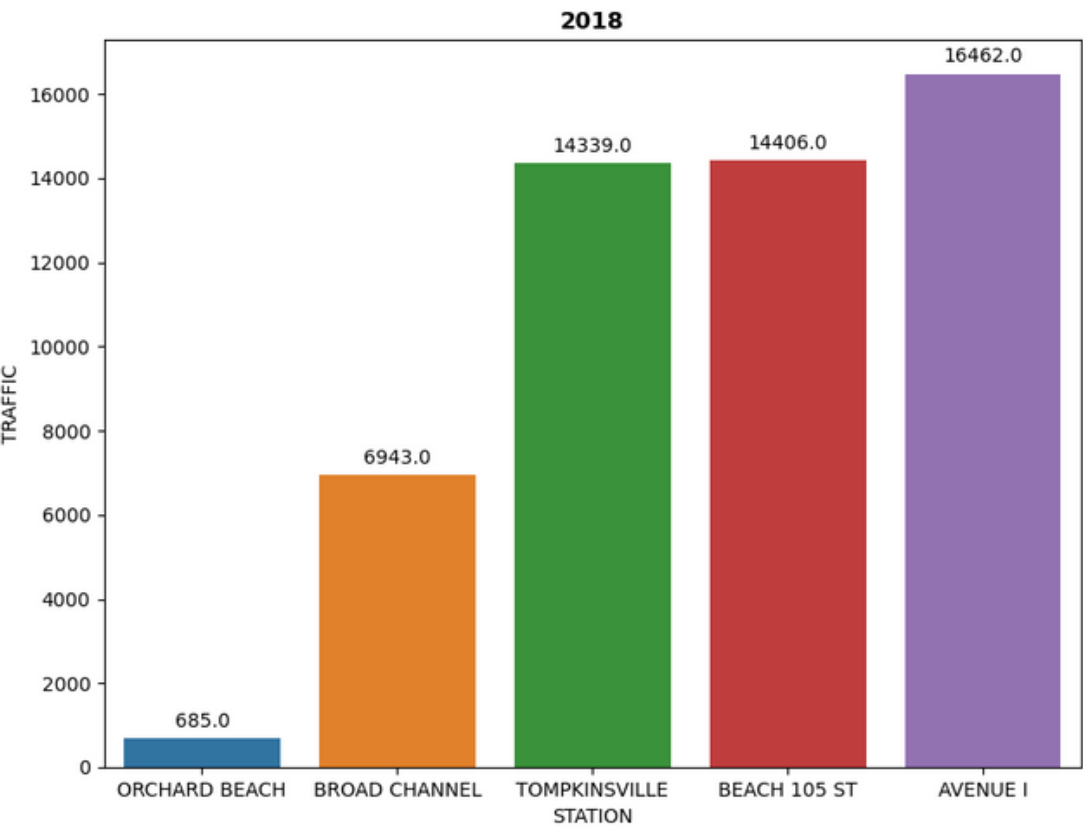
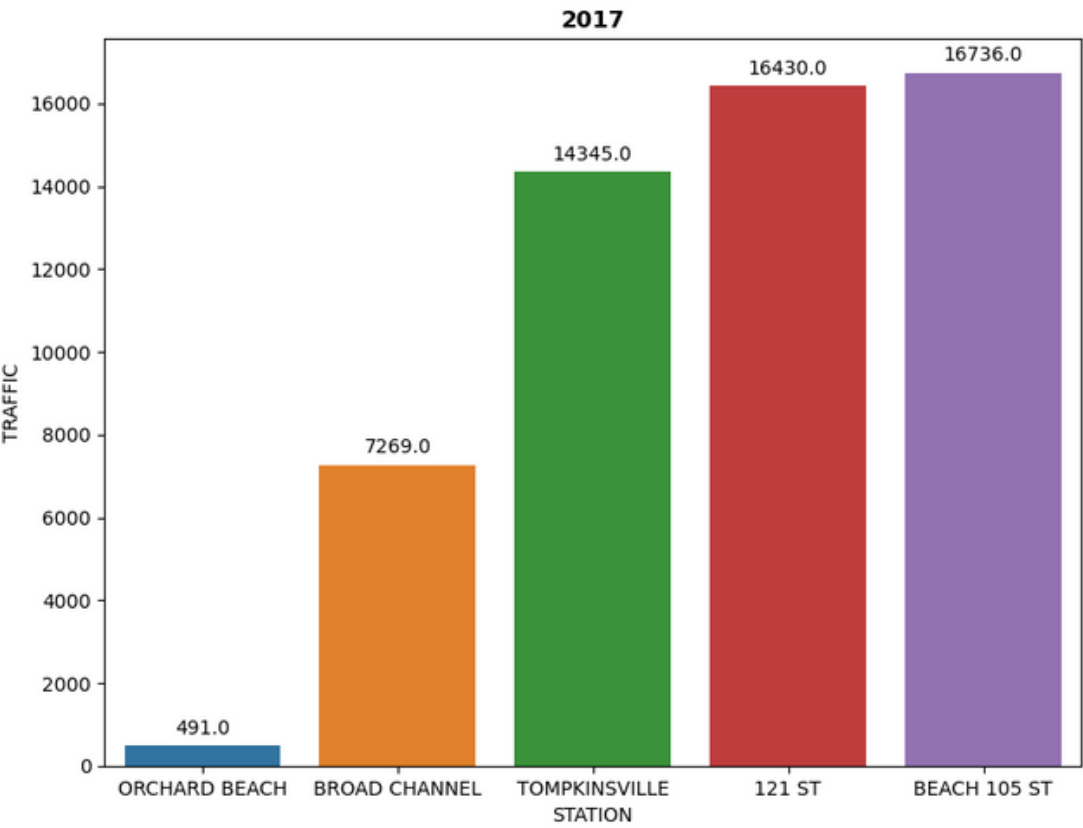
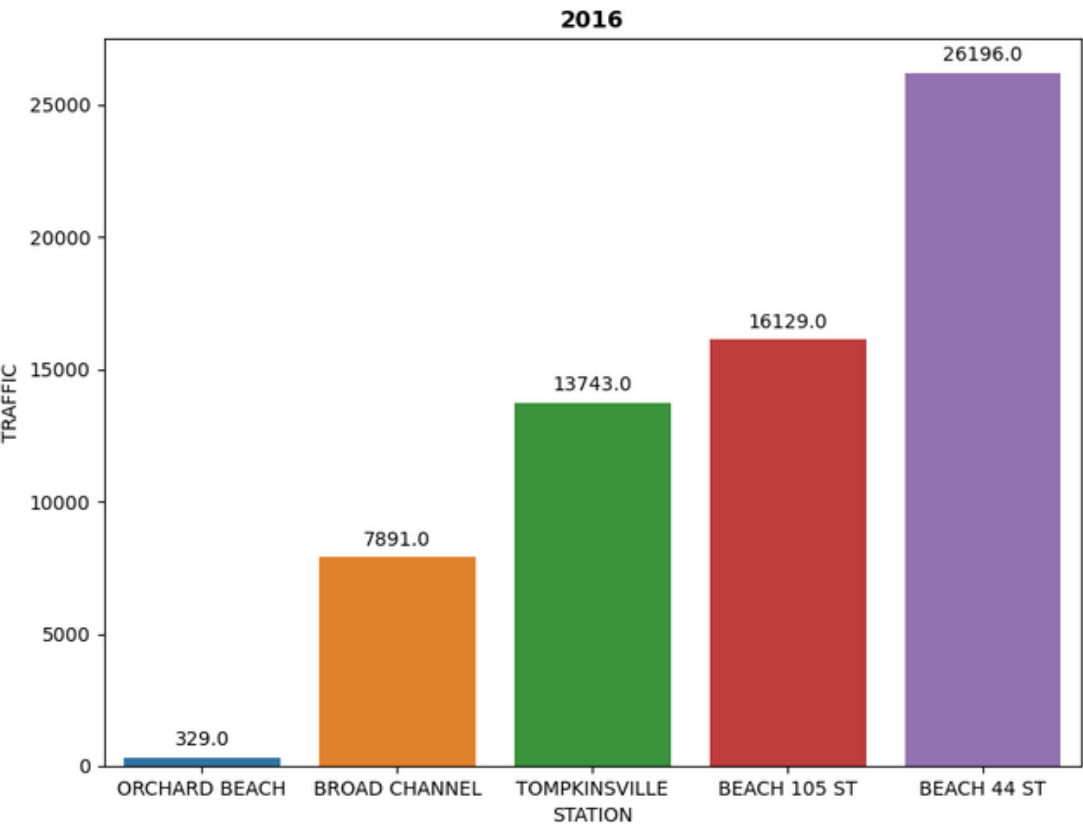
DATA ANALYSIS & VISUALIZATION



TOP 5 BUSSET STATIONS IN (JUNE-JULY)

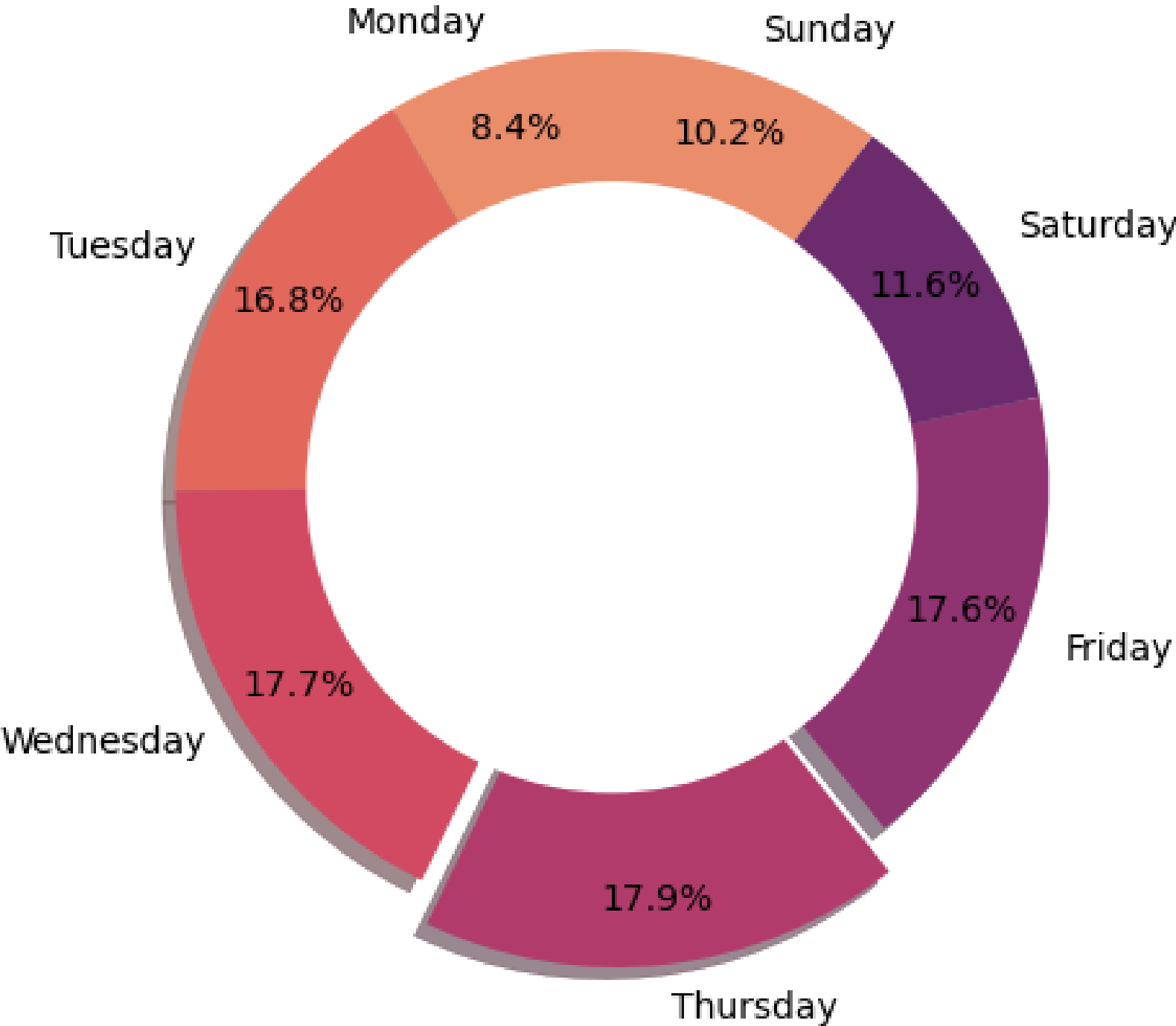


LEAST 5 BUSIEST STATIONS IN (JUNE-JULY)

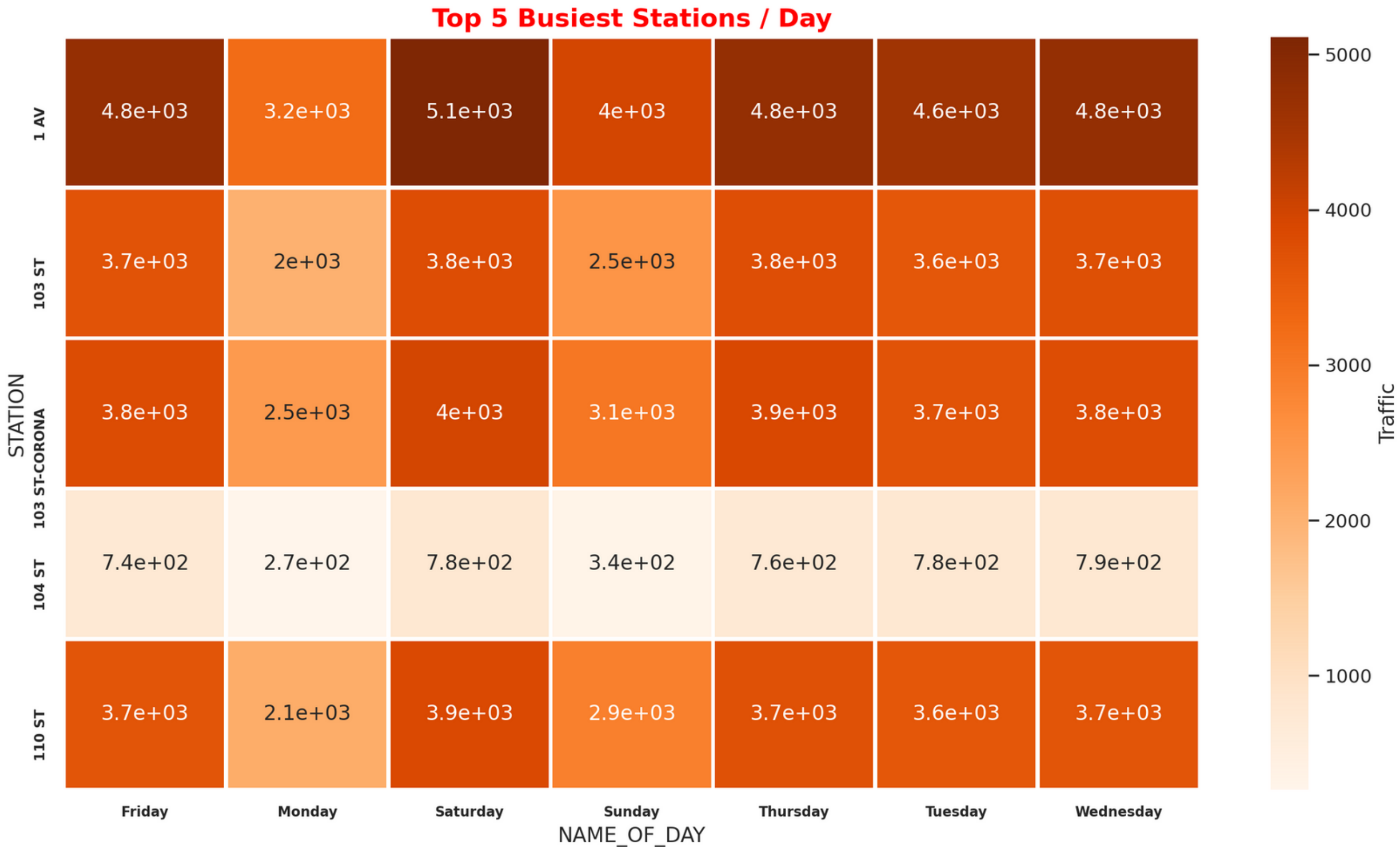


TOTAL TRAFFIC

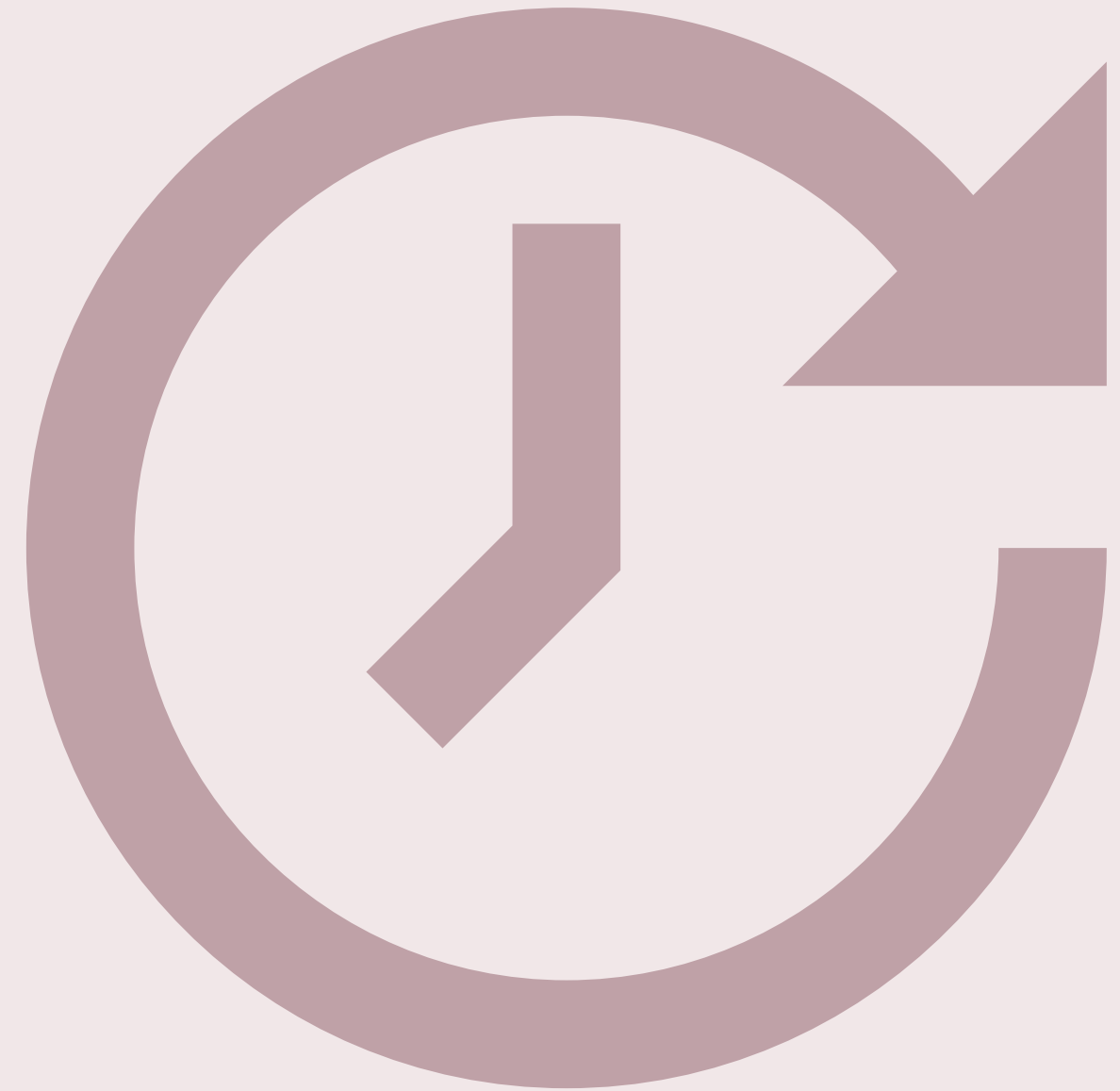
Total Traffic based on day



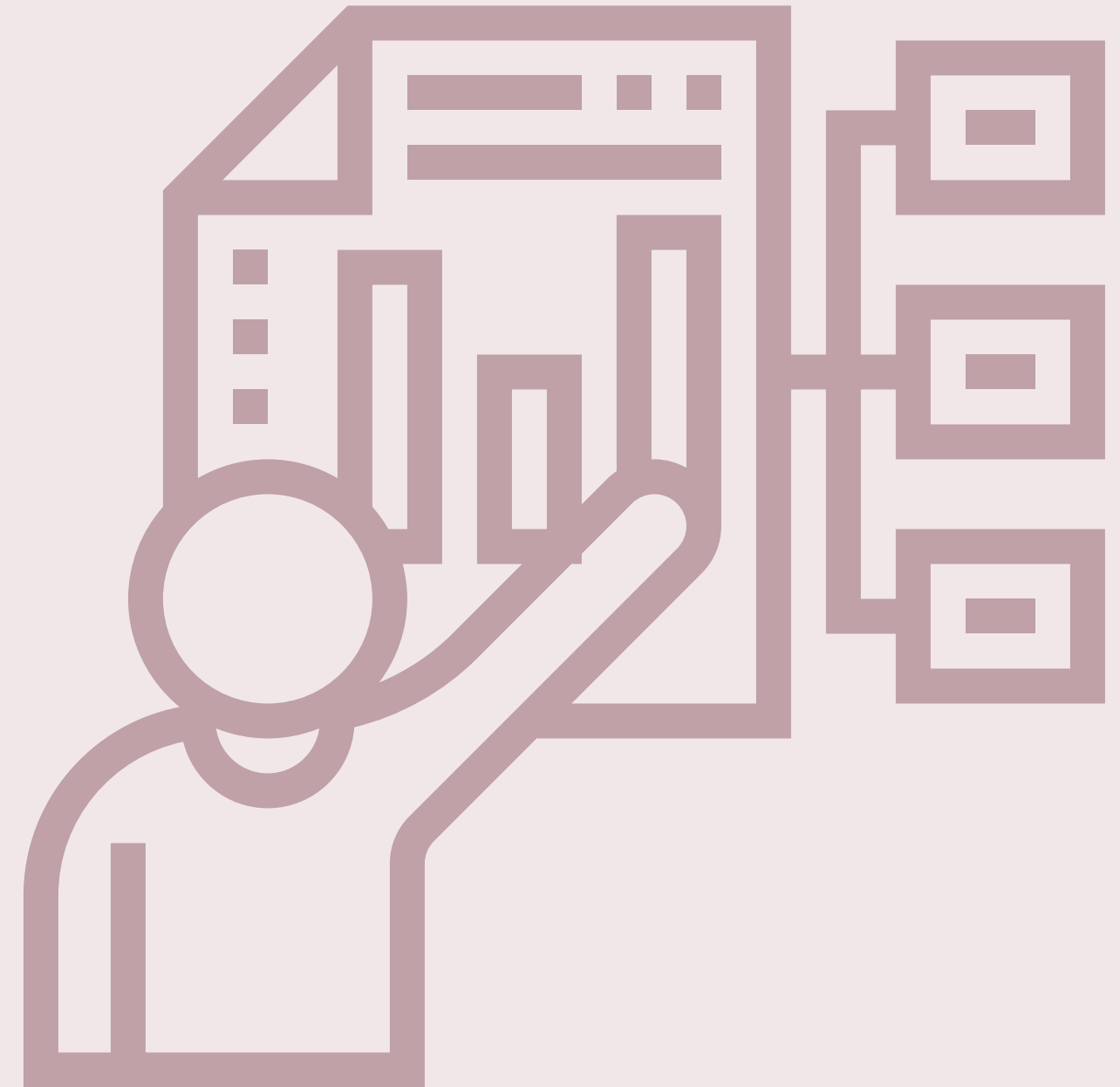
TOP 5 BUSIEST STATIONS / DAY



FUTURE WORK



CONCLUSION



THANK YOU!
