



EDA | Exploratory Data Analysis

For RSTABUS Imaginary Bus Company

Prepared by
Raghad Rashed Al-Saqer

Supervised by
Dr. Mohamed Baddar.



"RSTABUS Bus Company "

bus company covering several regions in the United States of America. It is concerned with carrying passengers from several specific centres to the destination stations throughout the day. want attract more customers to benefit from its services and also to expand the field of work and cover a larger space for the geographical area at the level of the USA.

A

1

2

Assumptions

1. the busiest stations of demographic or location and do not have bus service. is an aim to work on most of my analysis.
2. I assumed that the date of the event would be on the period form (Jul to Oct 2021), to focus on collecting only the last data from the MTA and to get the latest results.
3. After knowing the locations of the highest stations traffic and not having bus service, RSTABUS will be able to start drawing up a business plan to provide the service to customers.
4. Knowing the traffic throughout the day will be sufficient to determine the number of buses that will be provided and to determine their operating times.
5. Bus stations will be built based on the most traffic stations only.



DATA PREPARATION

Data cleaning

CHECK FULL SUMMARY

A QUICK TRIP TO DECODE THE INFORMATION FROM PROVIDED ON DATASETS.

SOLVING PROBLEM –

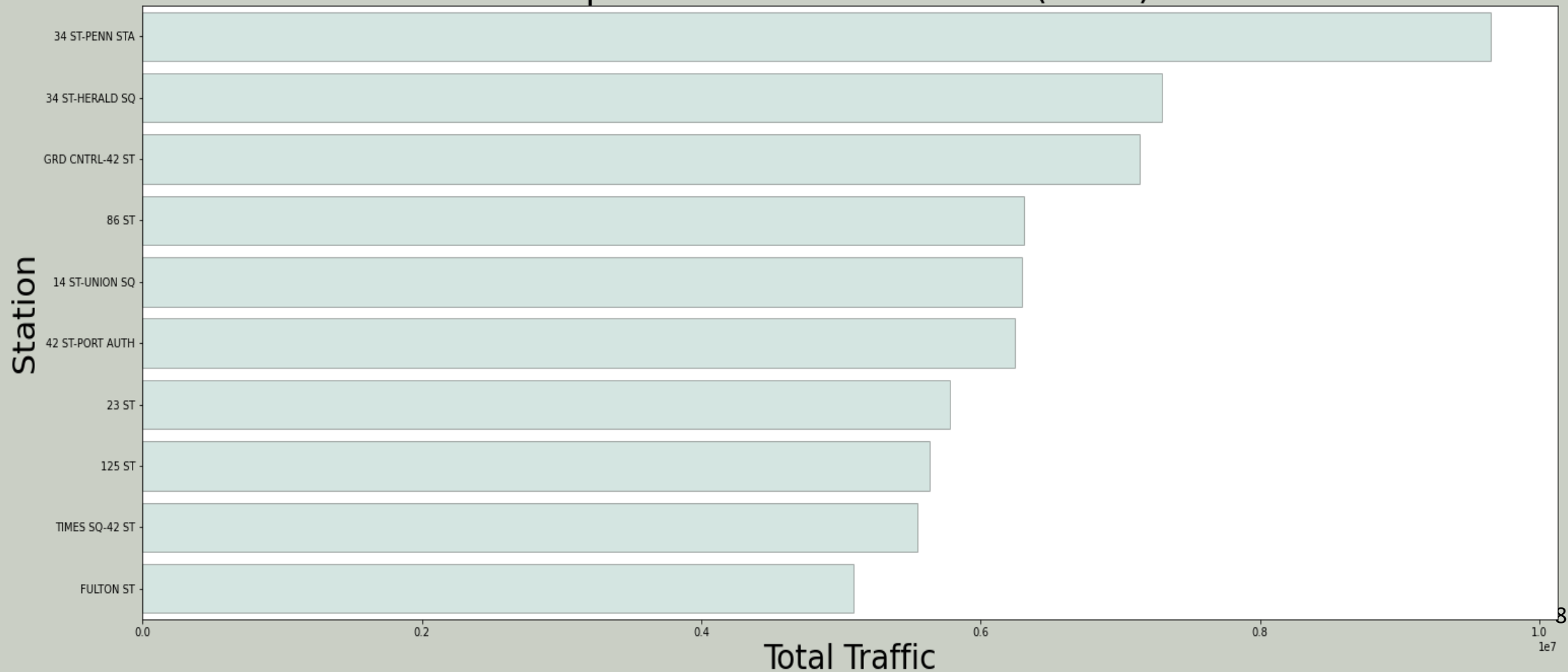
CUMULATIVE ,REST, AND REVERS ISSUE -OUTLIERS

EXTREMELY HIGH COUNTS (OUTLIERS)

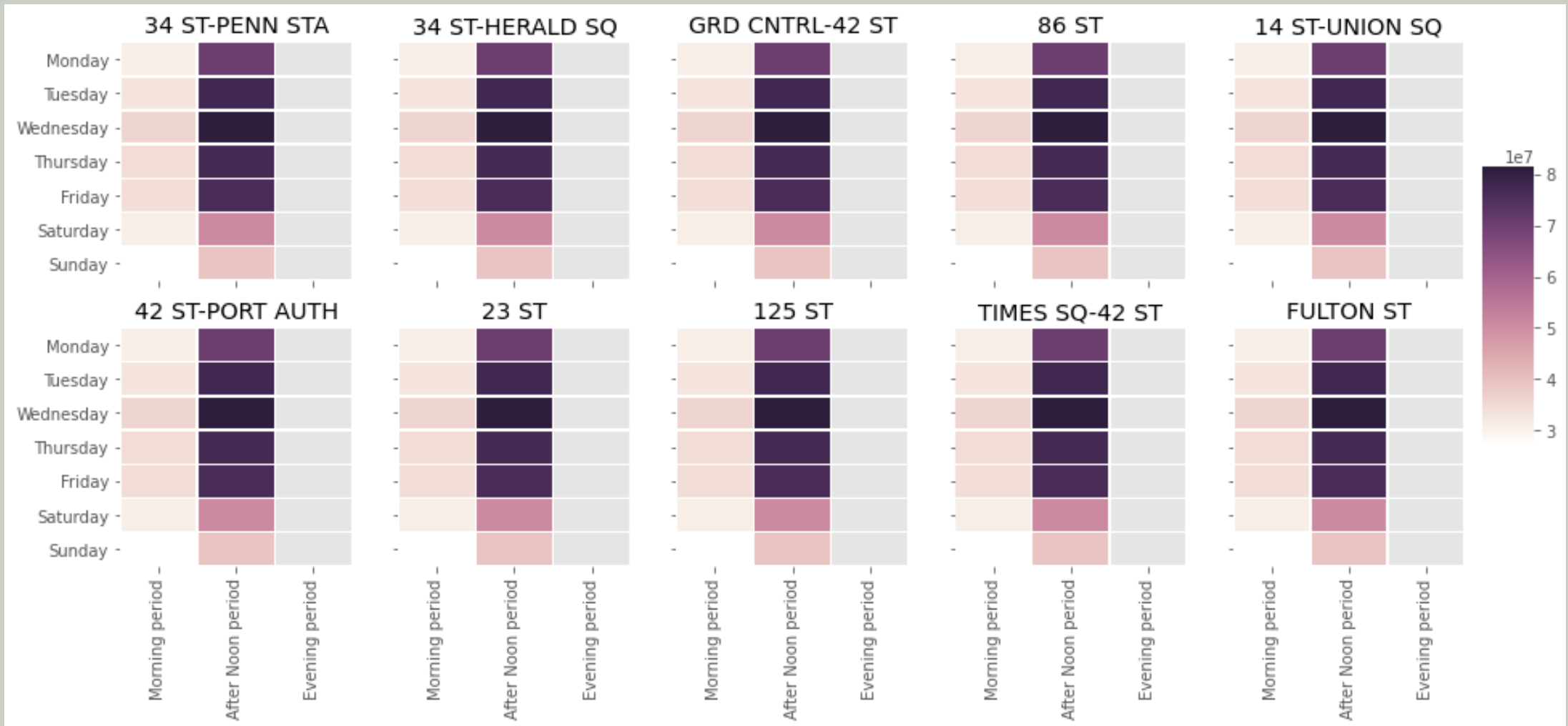


RESULTS

Top Most Trafficed Stations (2021)

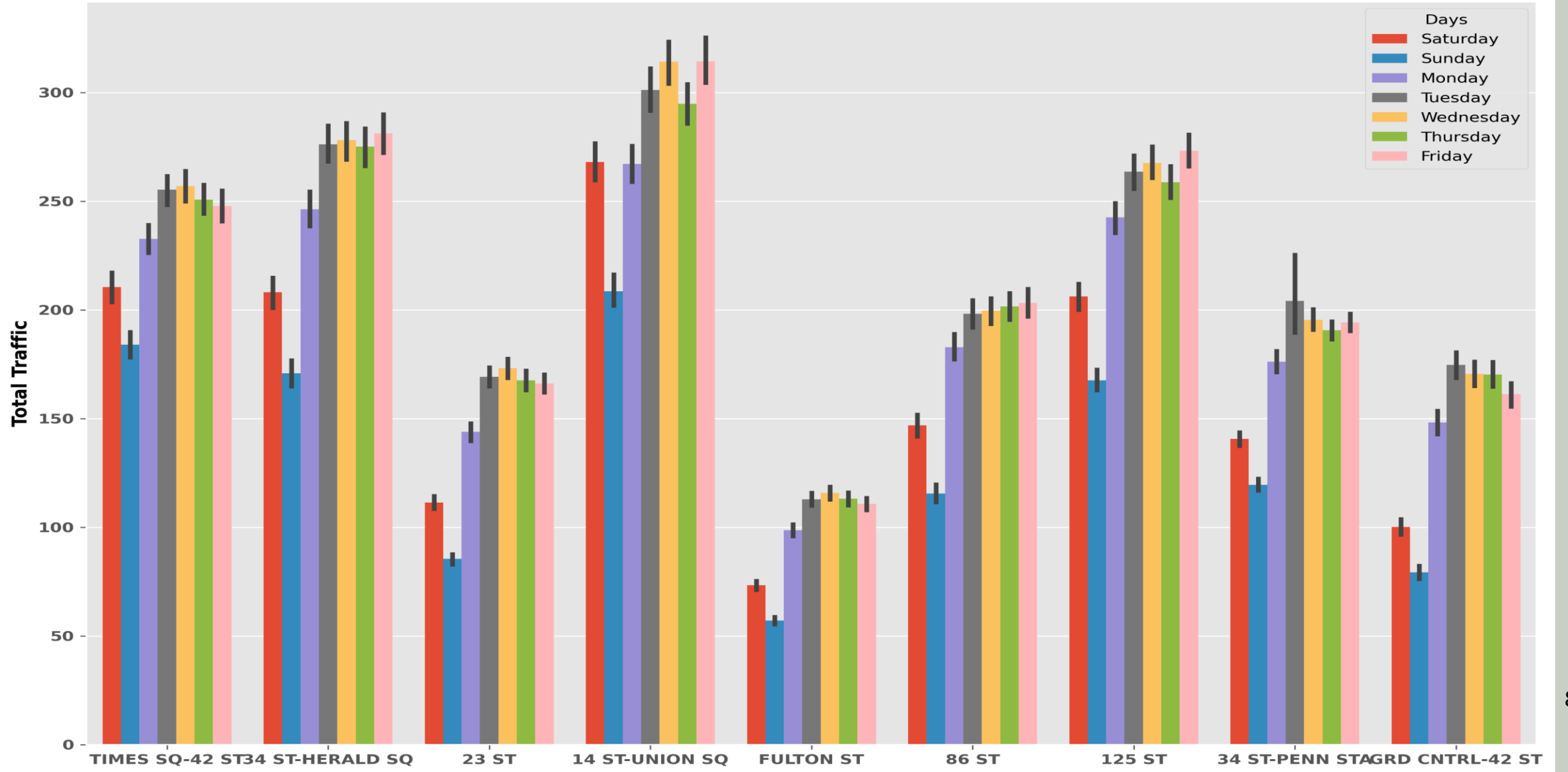


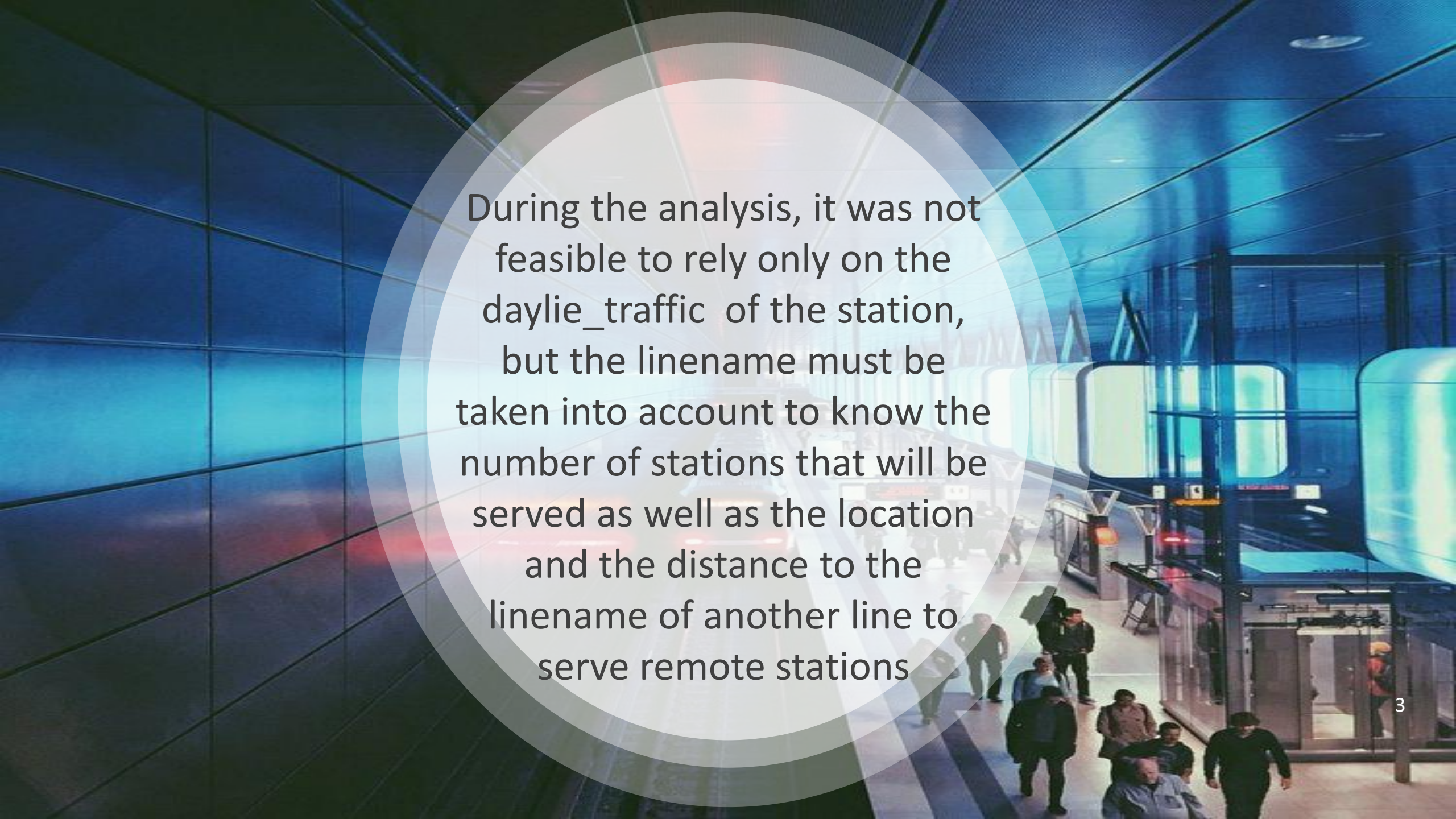
RESULTS



RESULTS

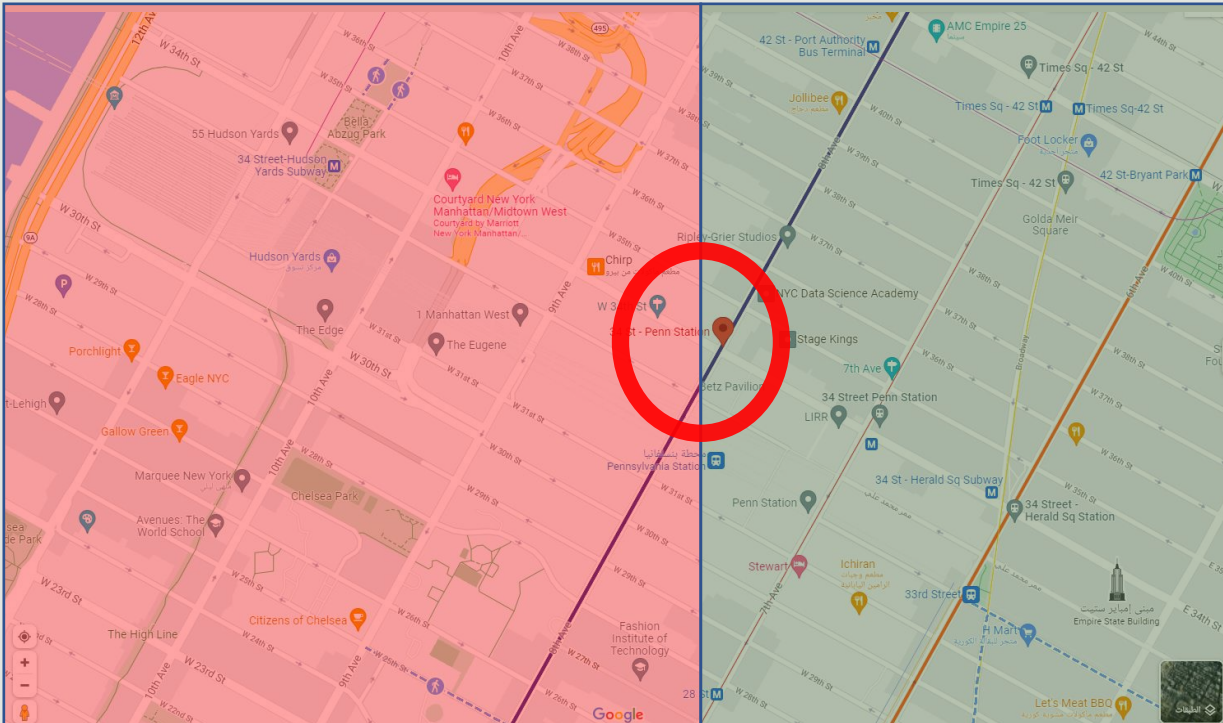
Top 10 Busiest Stations / Day



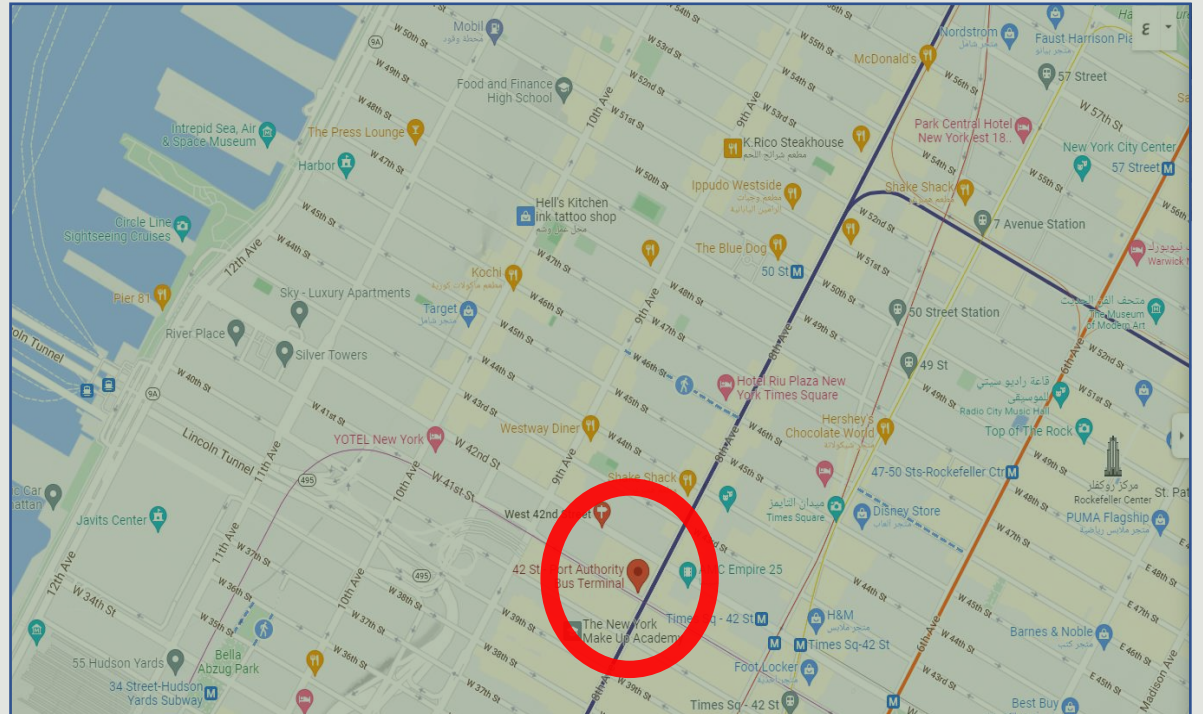


During the analysis, it was not feasible to rely only on the `daylie_traffic` of the station, but the `linename` must be taken into account to know the number of stations that will be served as well as the location and the distance to the `linename` of another line to serve remote stations

34 St - Penn Station



West 42nd Street



Feature work

- GEOMAPPING –

CONNECT THE DATABASE TO THE MAP TO KNOW THE RIGHT LOCATION
ANALYSIS WITH LINE NAME

BUS SERVICE FOR LOW-TRAFFIC AREAS

- OTHER VISUALIZATIONS SUCH AS INTERACTIVE PLOTS
- PLOTLY, BOKEH





Thank you for listening.