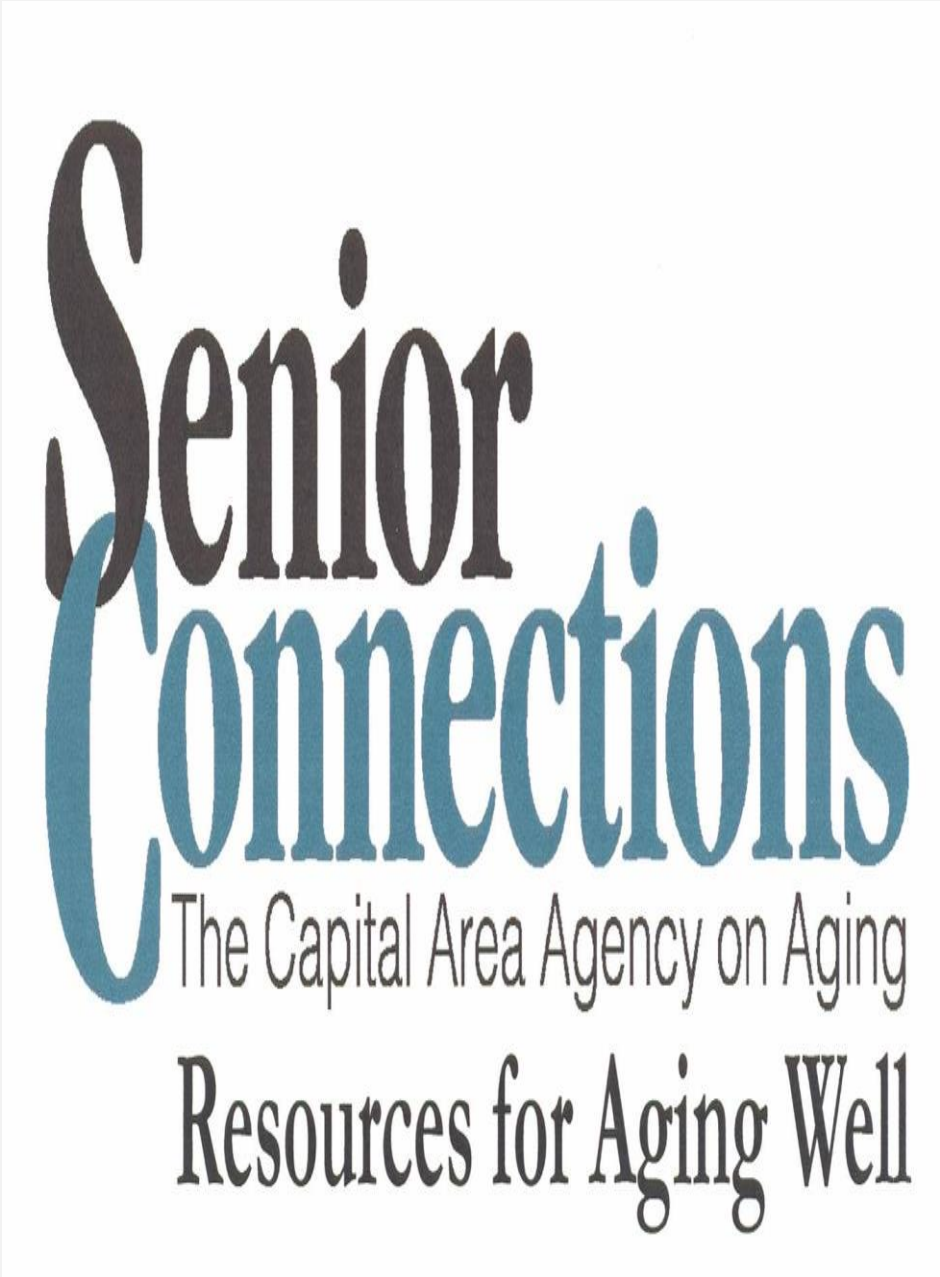
MTA DATA ANALYSIS OF NY SUBWAY

Senior connections- fundraising analyze

By Khulud Mansour Amer

BACKSTORY

- Senior connection decides to launch a new location in NY city.
- The organization is planning to have a fundraising in New York city which they hope to fill their event space with people passionate about increasing the participation of old people in society by improving quality of life for seniors and find usefulness in their daily tasks.
- To support this gala Senior connections is doubling their effort by sending members to the most busiest location "sub way NY stations" to collect emails from travelers and send free tickets to attend the fundraising .
- My duty is targeting the perfect(most crowded) location at NY subway stations(entries) to place our team by analyzing MTA database during June, July, August 2021 to ease their mobility and collect signatures from the top 5 most crowded entries .



TOOLS

- ANALYSING DATA :
- Python
- Panda
- NumPy
- SQLalchemy
- Matplotlib
- Seaborn

DATA SOURCES

constrict three months- June, July, August from the MTA database.

Perform a thorough Exploratory Data Analysis of the MTA turnstile data (http://web.mta.info/developers/turnstile.html) ;

ANALYTICAL STRUCTURING

clean, explore, aggregate, and visualize the data for ENTRIES as appropriate to address the top 5 crowded stations.

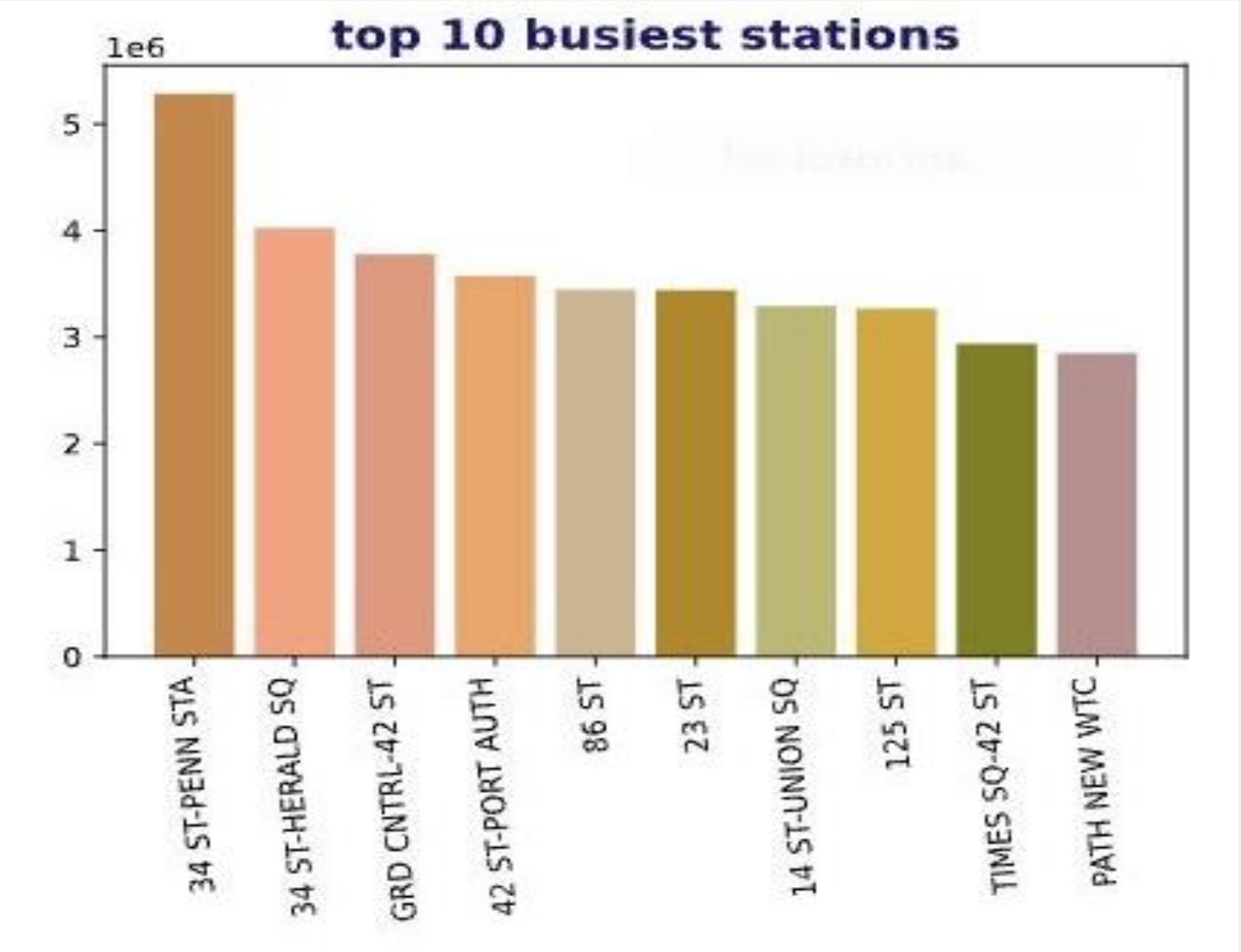
From the top 5 stations plot the daily time series for a station.

Make the same list for another week, and another week ,ets using pandas datetime.

Use Visualization

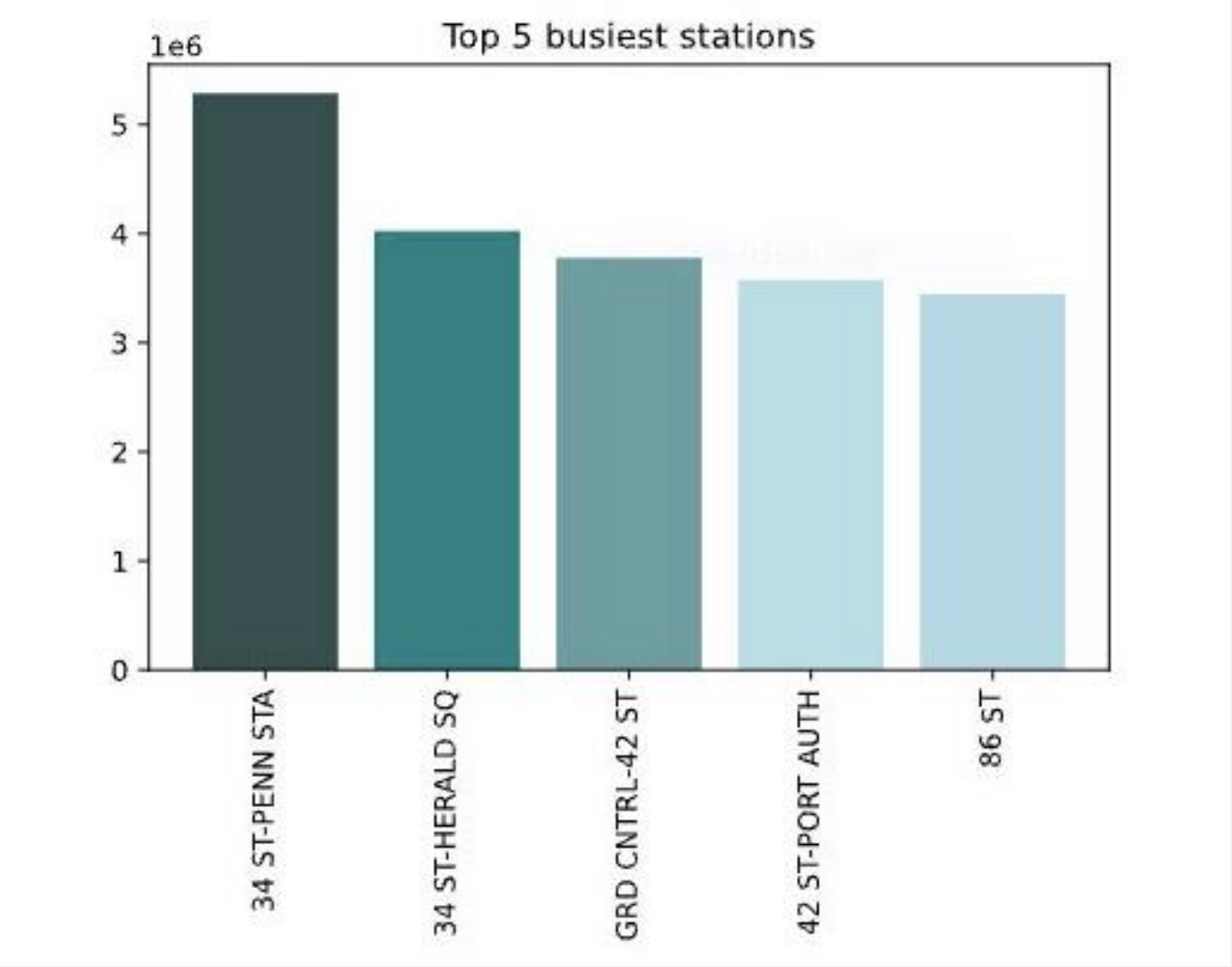
VISUALIZATION

- Bar plot of the top 10 NY subway busiest stations at entries.

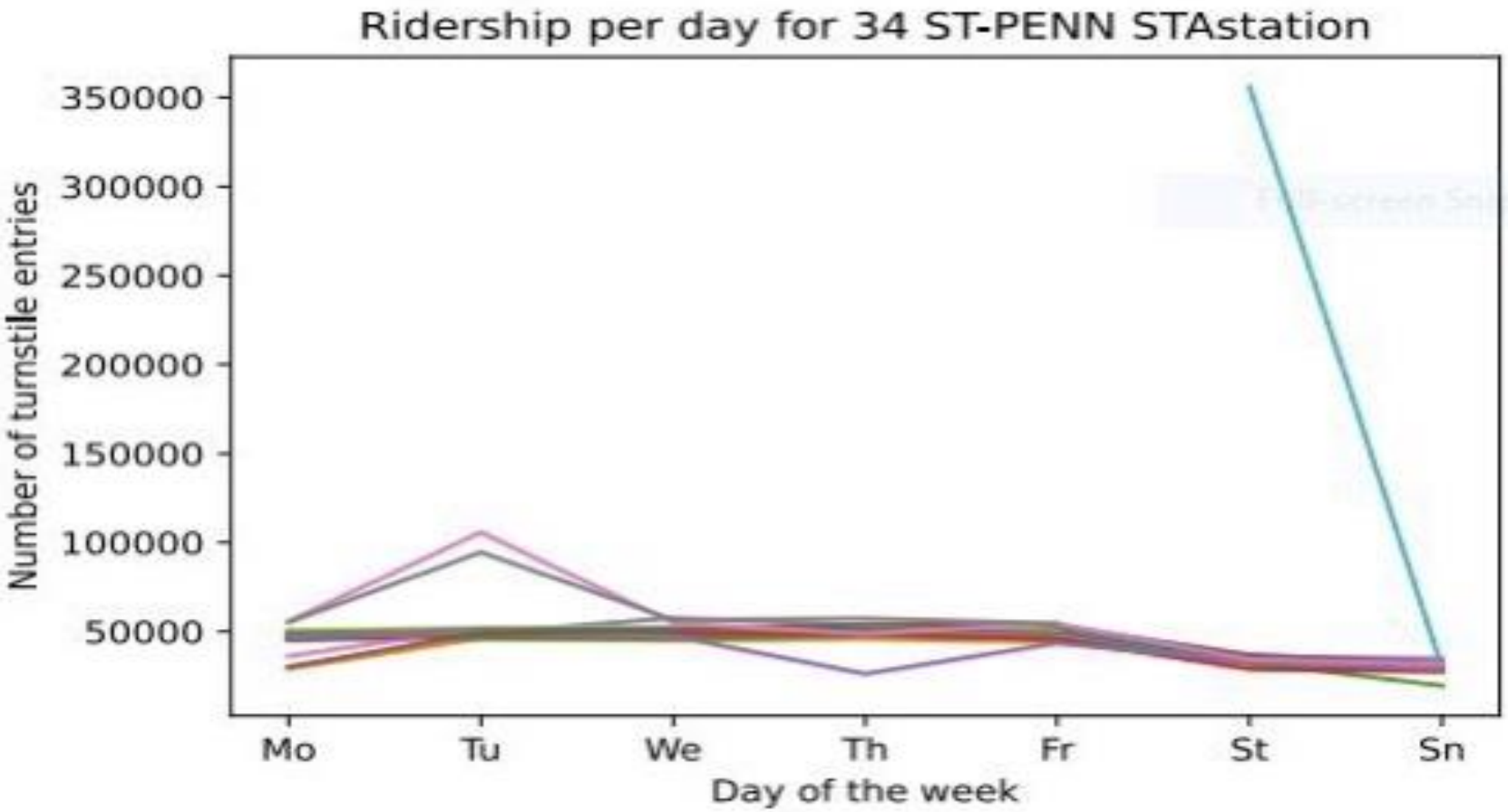


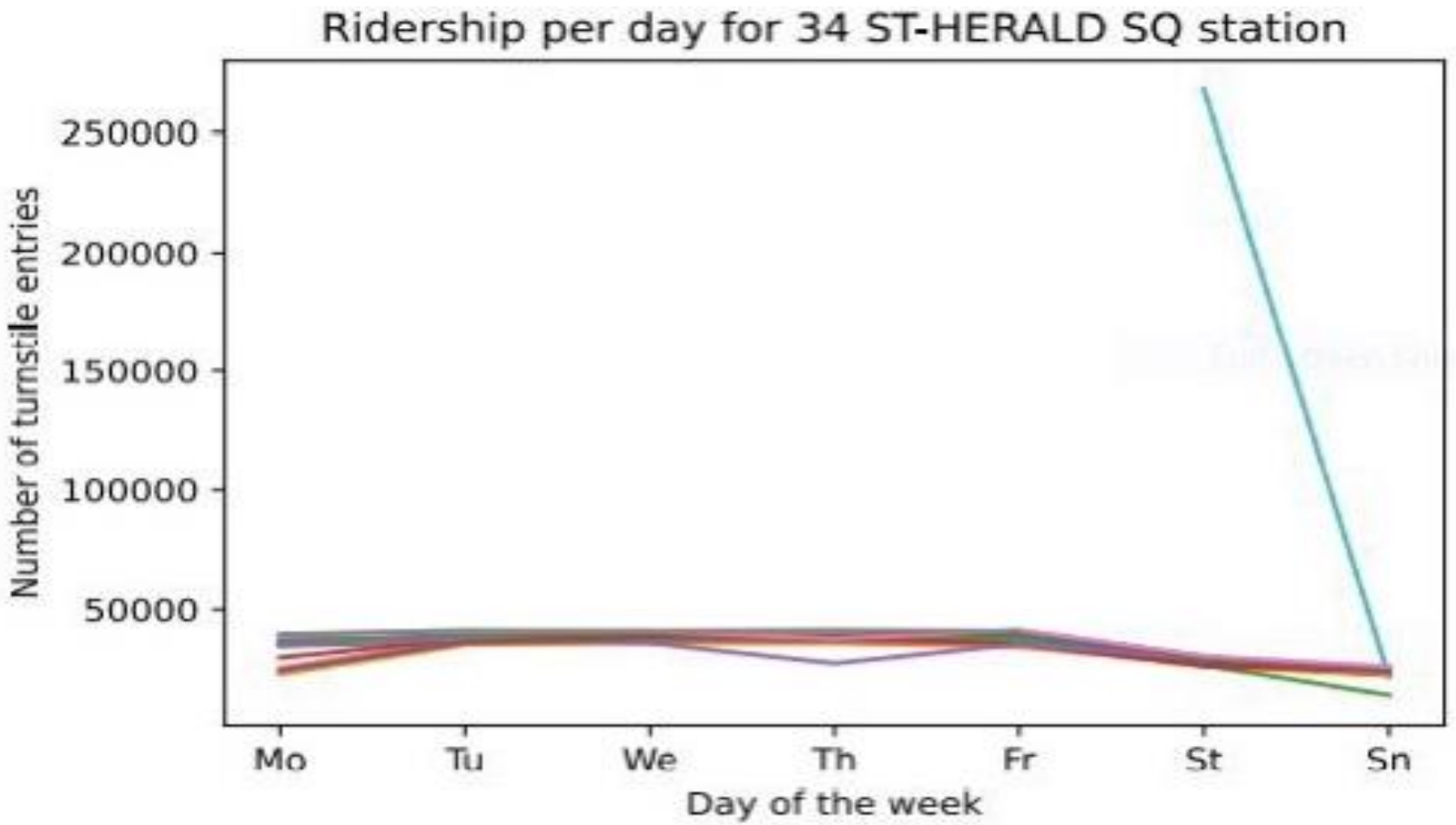
VISUALIZATION

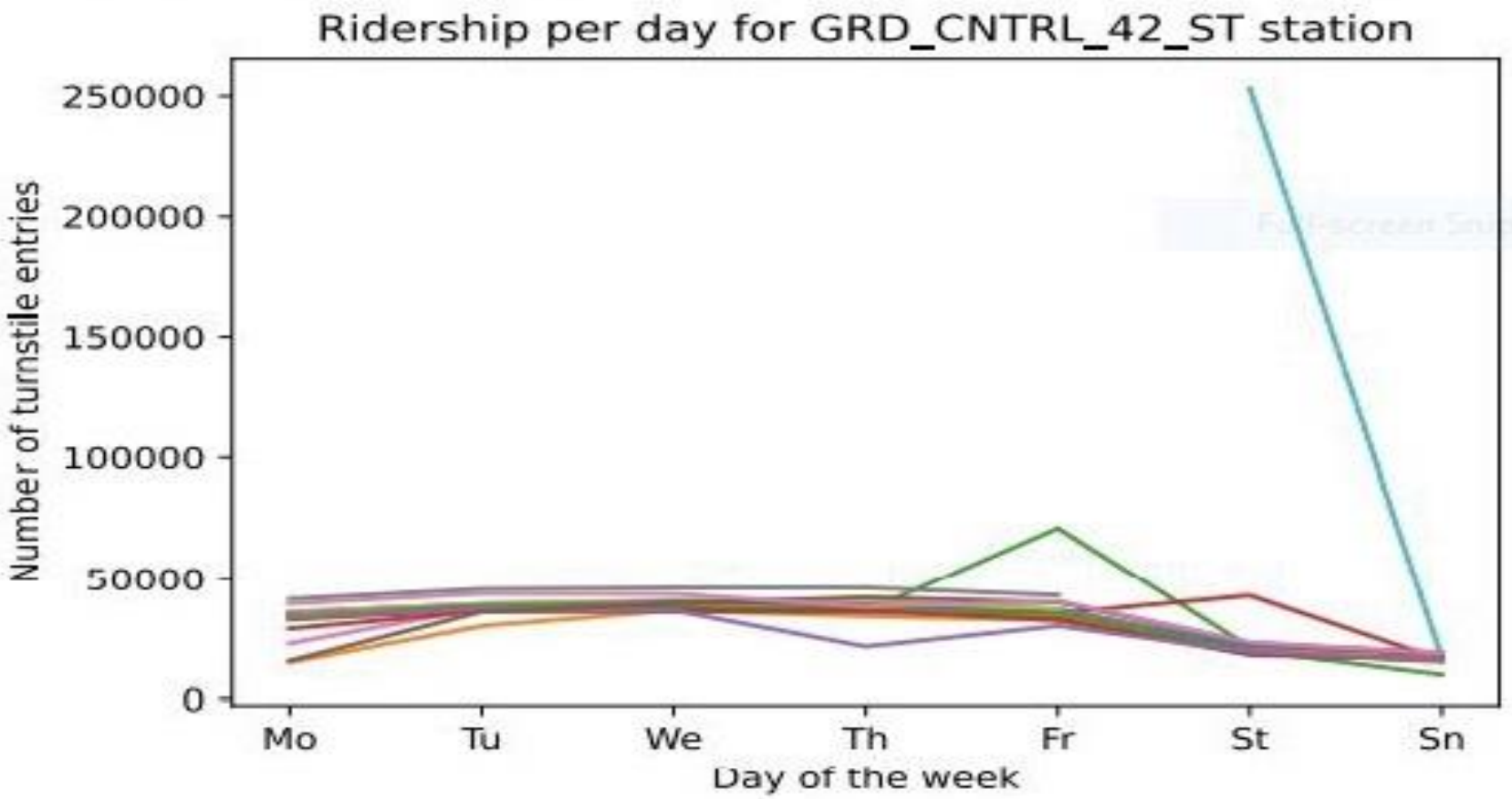
Bar plot of the top 5 NY sub way busiest stations at entries.



VISUALIZATION OF THE TOP 3 BUSIEST STATIONS AT ENTERIES


Ridership per day for 34 ST-PENN STAstation
Number of turnstile entries
350000
300000
250000
200000
150000
100000
50000
Mo
Tu
We
Th
Fr
St
Sn
Day of the week


Ridership per day for 34 ST-HERALD SQ station
Number of turnstile entries
250000
200000
150000
100000
50000
Mo
Tu
We
Th
Fr
St
Sn
Day of the week


Ridership per day for GRD_CNTRL_42_ST station
Number of turnstile entries
250000
200000
150000
100000
50000
0
Mo
Tu
We
Th
Fr
St
Sn
Day of the week

QUESTIONS!