

MTA DATA ANALYSIS STATION NEW YORK

biscuit Company
by Hala Almulhim



Introduction:

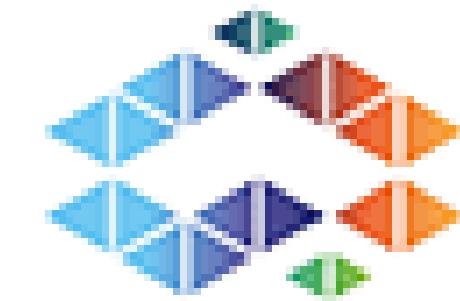
Biscuit company , My goal of this project is promoting a new product for biscuit company that produced new biscuit and must to be real tested from people so we can from MTA dataset know the daily traffic will accrue in which time in the day. and which day of the week, and my strategy to distribute free biscuit when people enter the station so people can know more about my product .my target is the most busiest stations Such as . ,top 5 crowded stations in (NYK),united states





Tools: Data analysis with the following:

- Python
- Pandas
- NumPy
- SQLAlchemy
- Matplotlib
- seaborn



أكاديمية سدايا
SDAIA Academy



Data Sources:

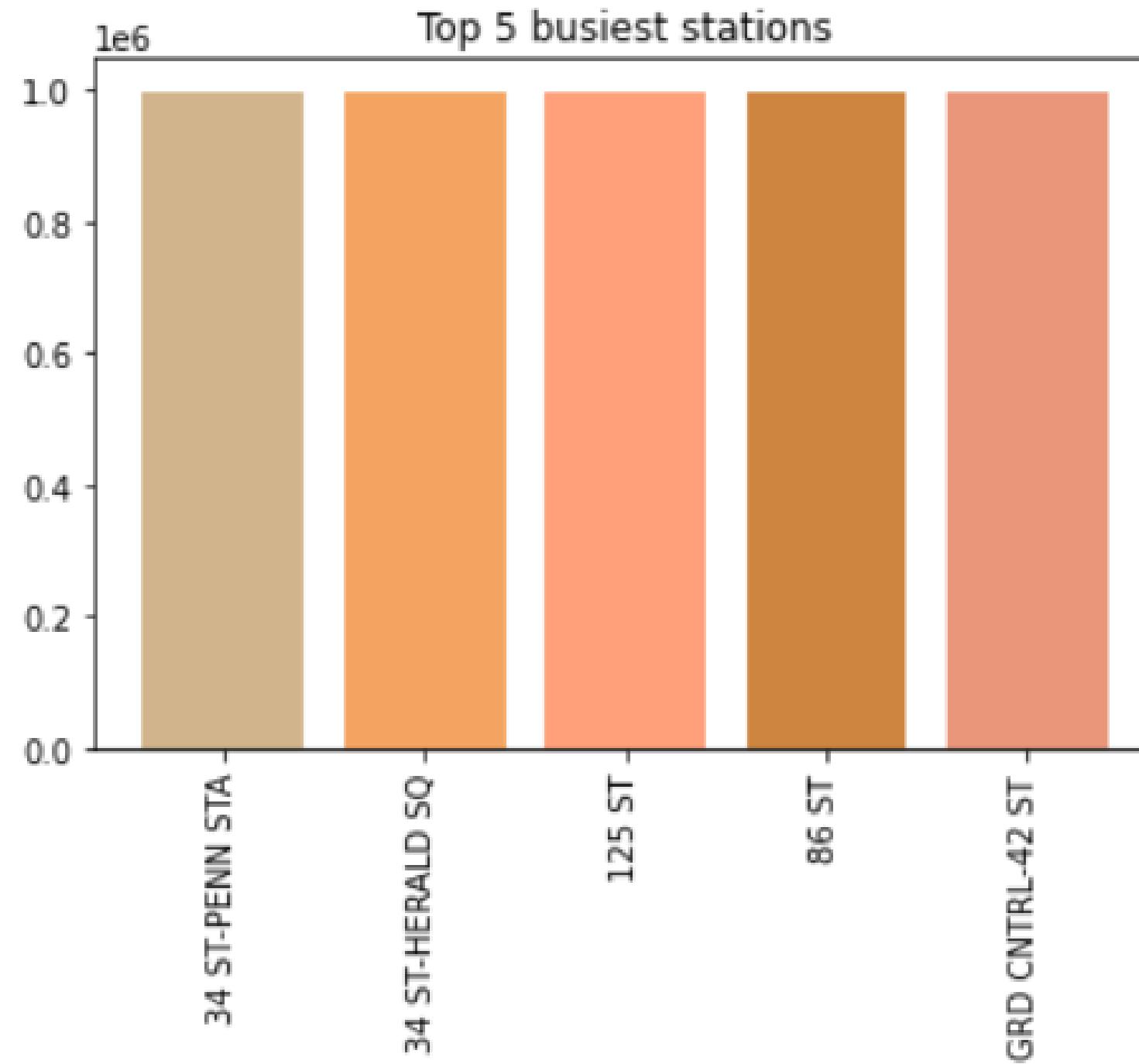
- from the MTA data I took three months (march - April - may in 2021)
- . looking for the most crowded stations in summer
- for each station , I added both their turnstiles daily entries and turnstiles daily exits data to find the busiest stations

MTA DataBase link:

<http://web.mta.info/developers/turnstile.html>:



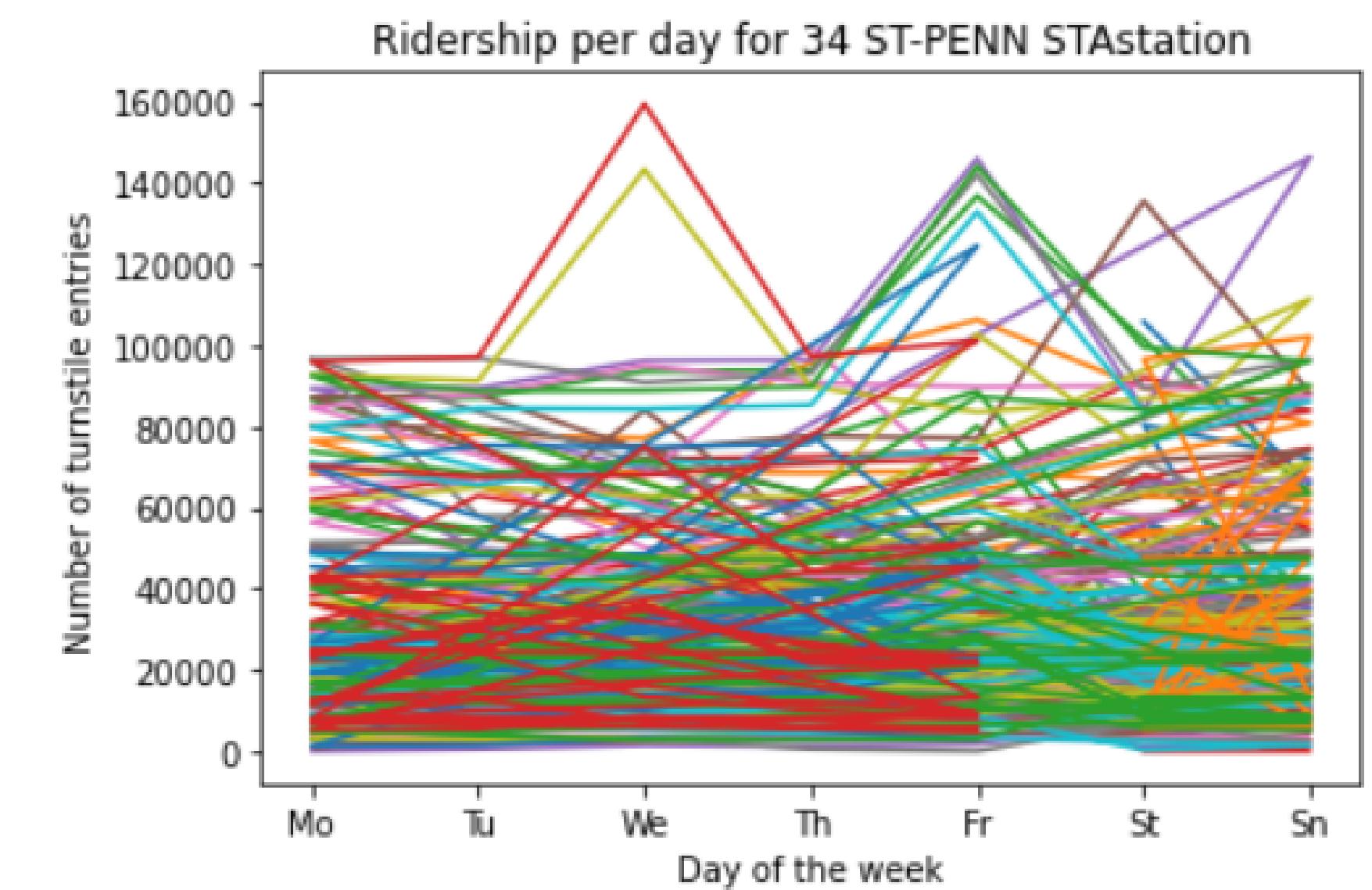
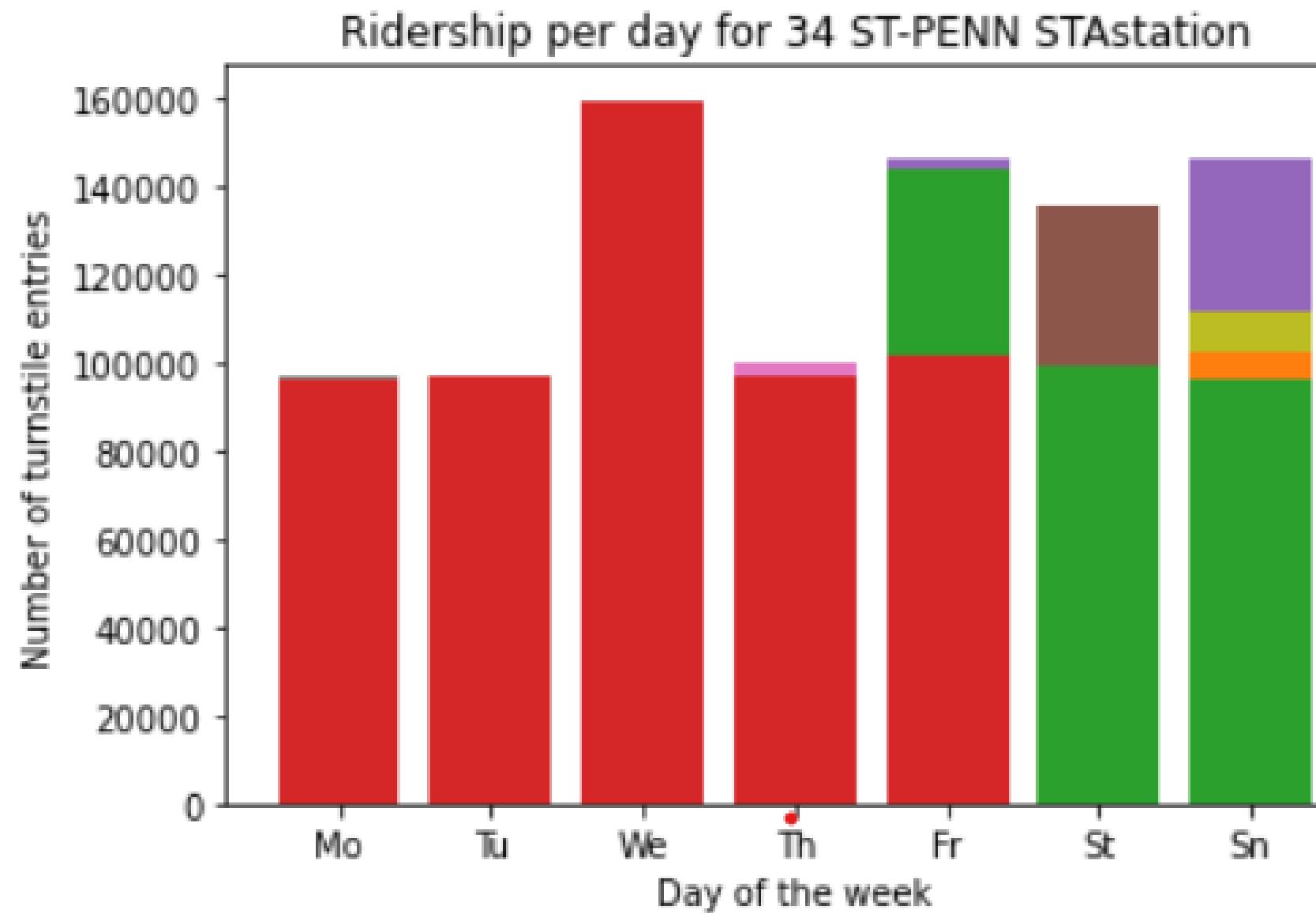
Top 5 busiest stations are:

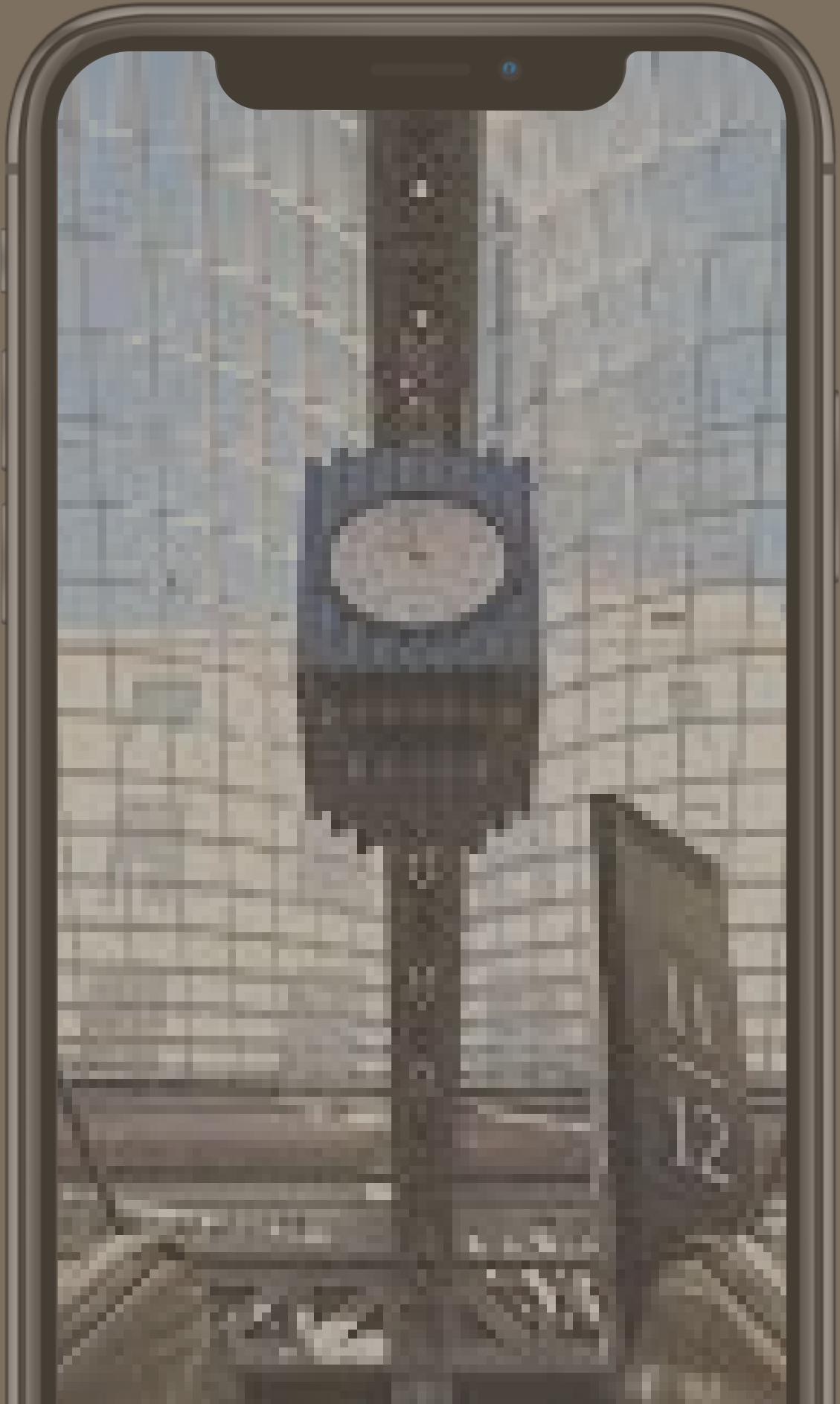


Key Findings:

1. 34 ST-PENN STA.
2. 34 ST-HERALD SQ.
3. 125 ST .
4. 86 ST.
- 5.GRD CNTRL.42 ST

find the first crowded station 34 ST-PENIN STA
station by the number of weekdays:





CONCLUSION:

after analyzing the data that we have I recommend to increase ridership in Wednesday, Friday, Saturday and Sunday by 60%, 40%, 30% and 40% respectively in 34 ST-PENN STA station to have a moderate traffic that will ensure a good level of customer satisfaction and increase in the station revenue due to the continuation of using this station.



GitHub

<https://github.com/HalaAlmulhim>



Thank you