



OMAR ALHADI

EDA (EXPLORATORY DATA ANALYSIS)

Of MTA TURNSTILE DATA NEWYORK CITY



Agenda

Part 1: Introduction

Part 2: Methodology

Part 3: Results

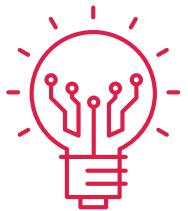
Part 4: Conclusion

Introduction:



Business Need

Locate the best Ad places in MTA stations to meet the company advertising requirements



Solutuation

Using MTA Turnstile dataset, and analyze it from January 2021 to March 2021 to find best stations for advertisement



Objectives

- Finding most busiest subway stations, most crowded day and time slot
- decide which places are the best for advertisement



Methodology :



Data Sources:

MTA Turnstile Data:

- Between January 2021 to March 2021
- 3 Months ,13*.txt file



Tools

- Data : Pandas, Numpy
- Data Visualization: Seaborn, Matplotlib
- Program: Python , SQL

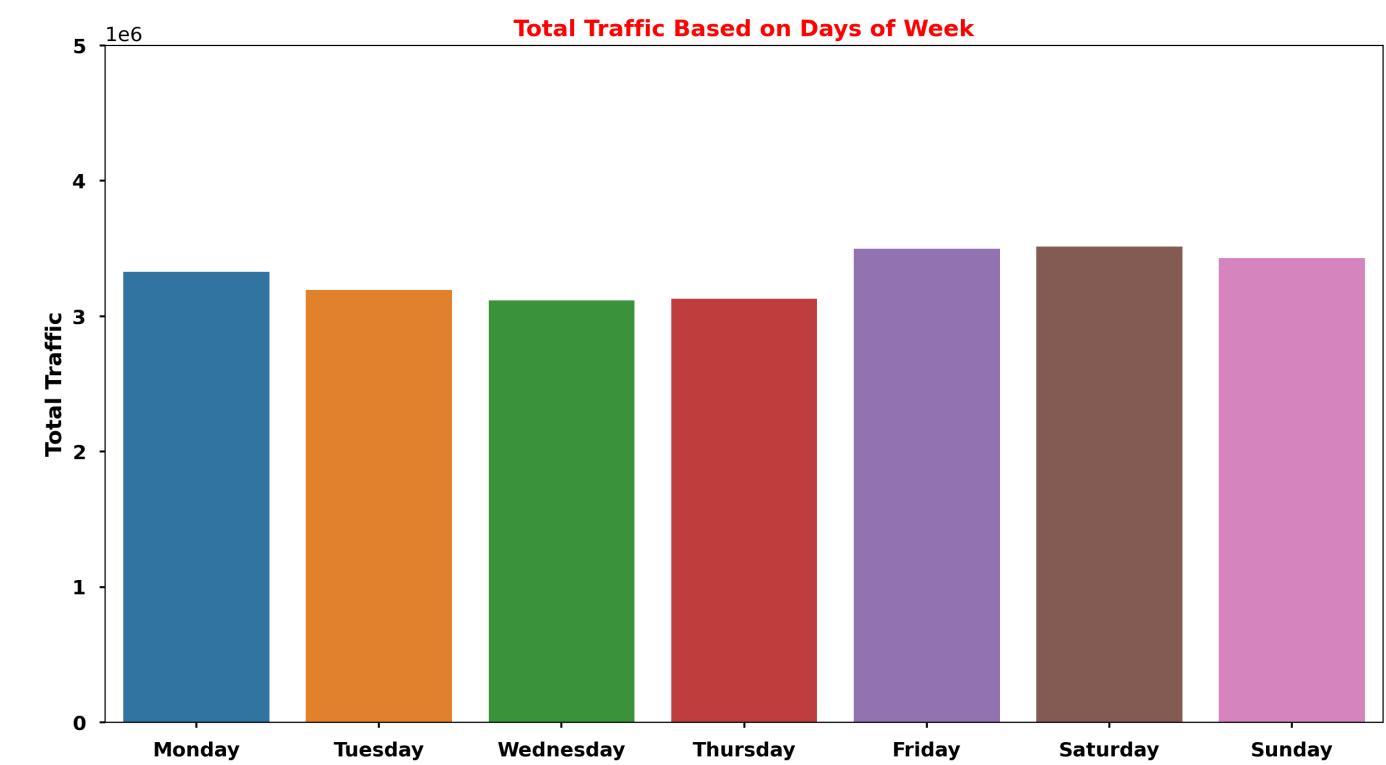
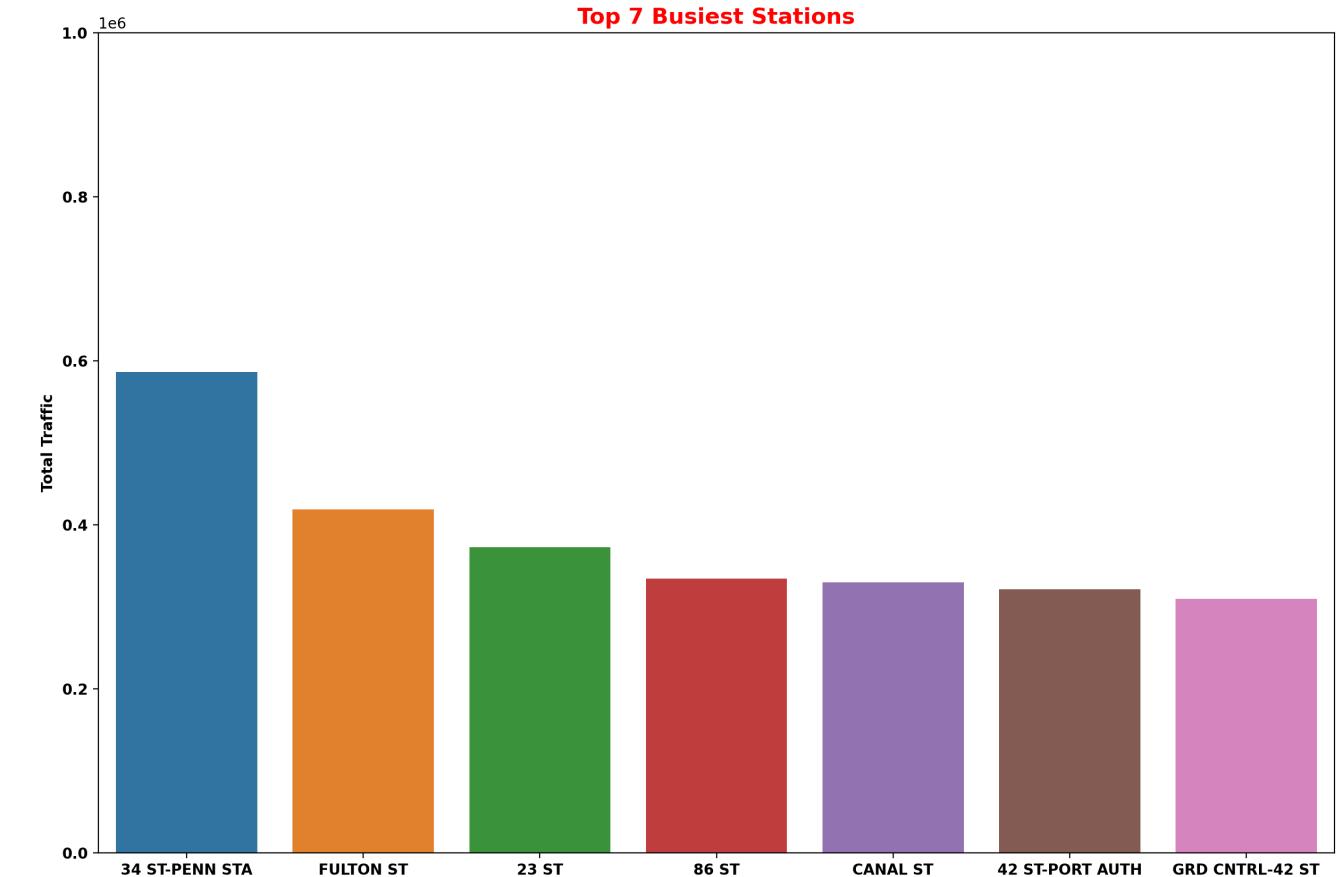
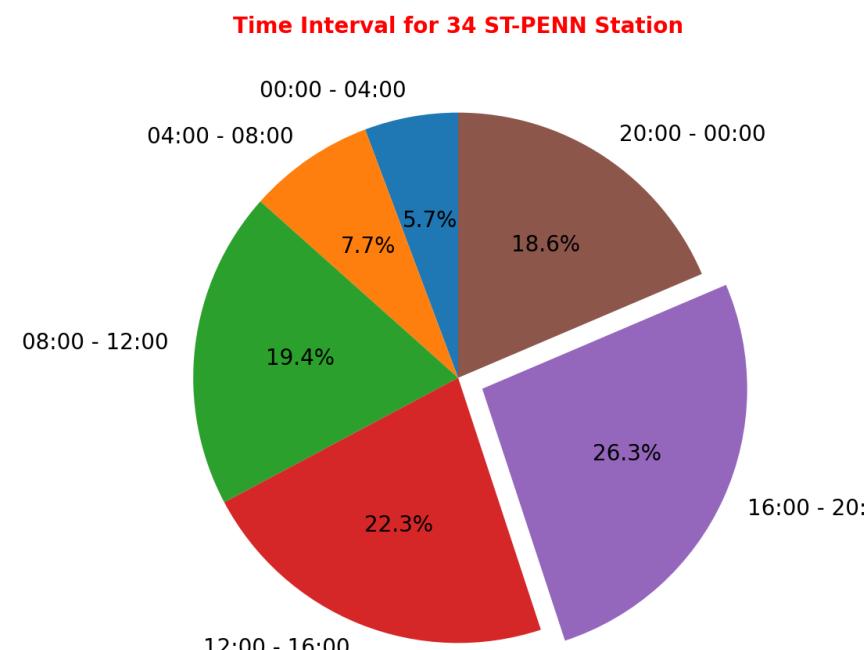


Results :



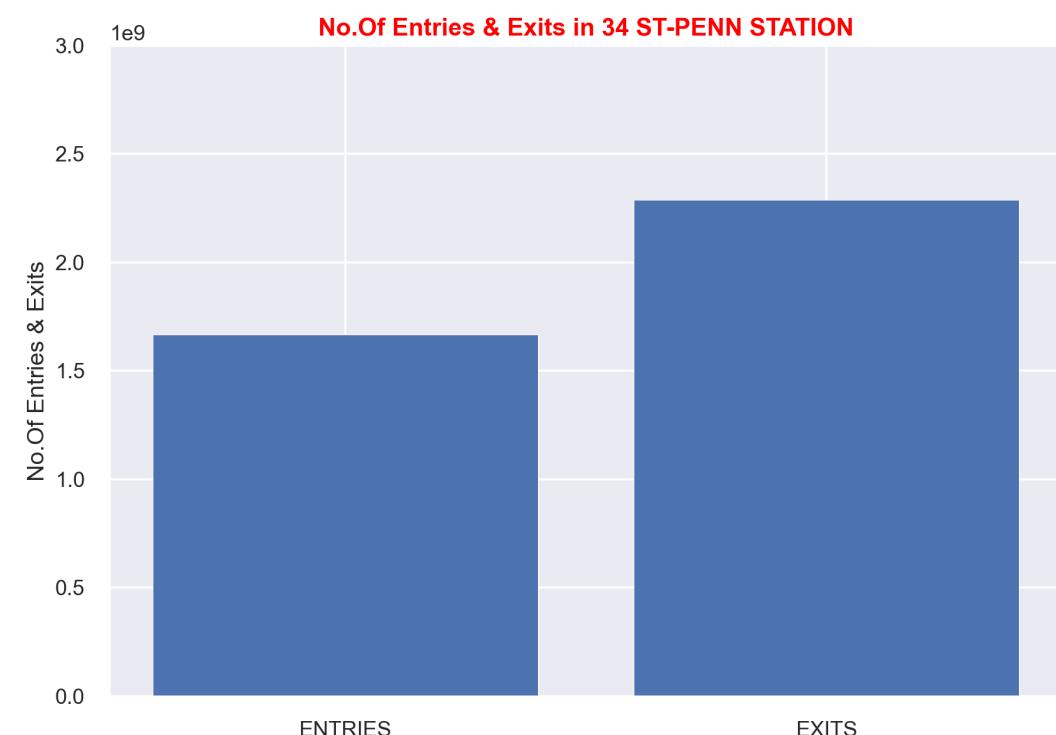
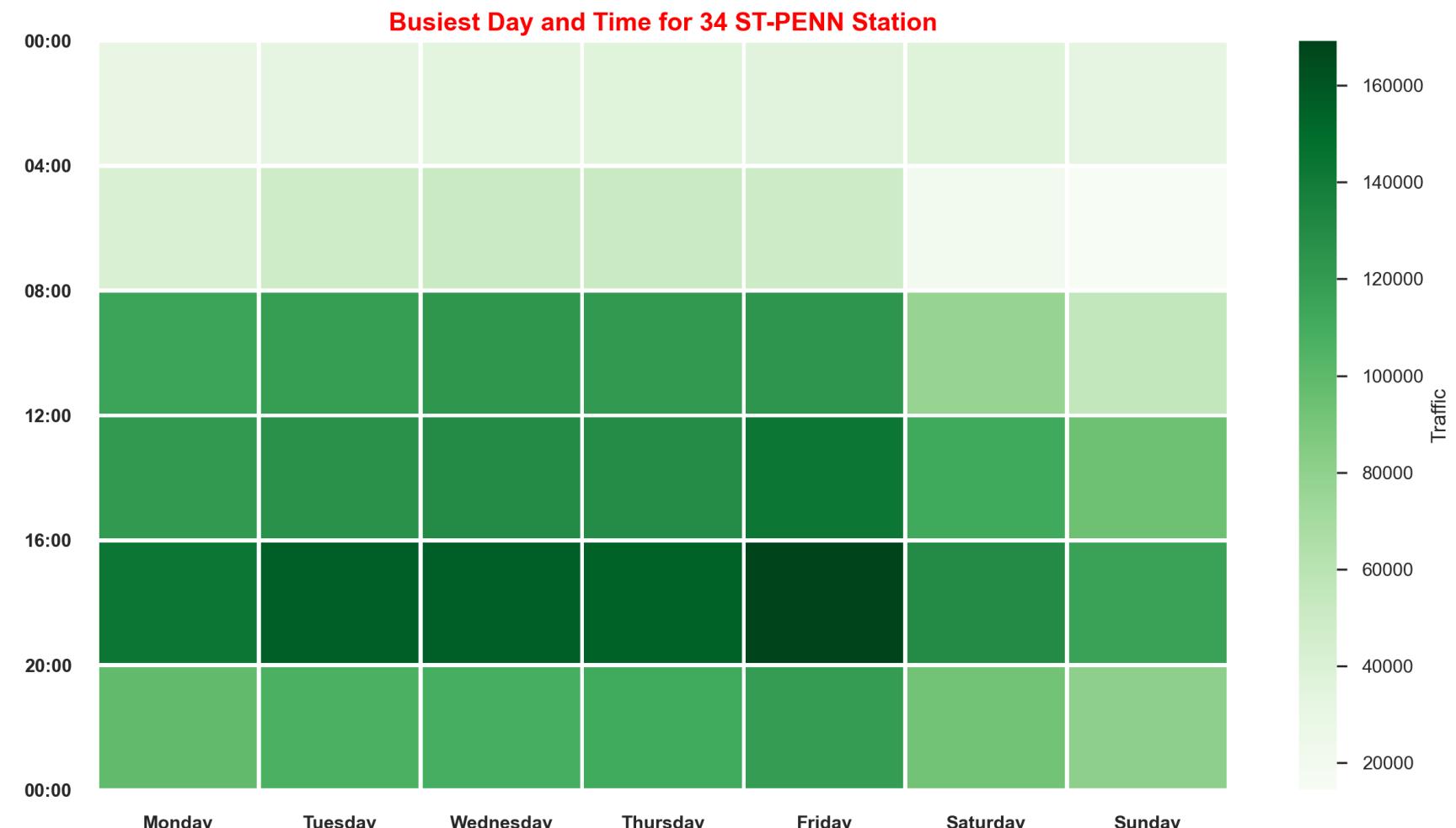
Brief Insight From Data

- 378 Stations
- 5037 Turnstile
- Most Busiest Station: 34 ST-PENN STATION
- Most Busiest Day: Friday
- Most Busiest Time Slot: 16:00 - 20:00



Conclusion:

- Weekdays are more busy
- Tuesday, Wednesday, Thursday and Friday are the best days
- 16:00 - 20:00 Time Slot is the best Time
- The number of EXITS is more than the number of ENTRIES which mean it is a commercial area
- The best option to reach max people is **Friday, 16:00 - 20:00**





Thank you!

