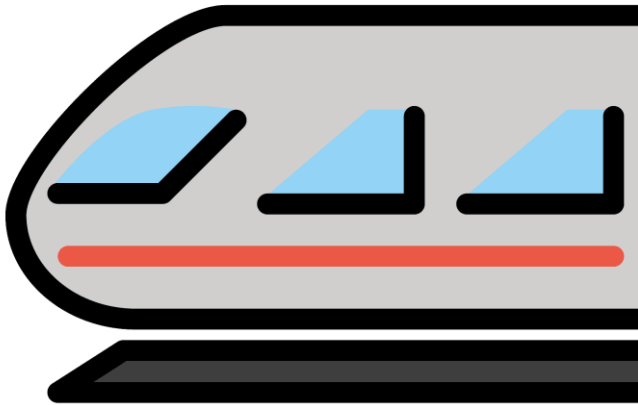
أكاديمية سدايا
SDAIA Academy



EDA ON NEWYORK MTA DATA

09-Septermper-2021

Prepared by: Sara Al Abdulsalam

TABLE OF CONTENTS

01 **INTRO AND PROPLEM**
What is the problem? How the MTA data can solve it?

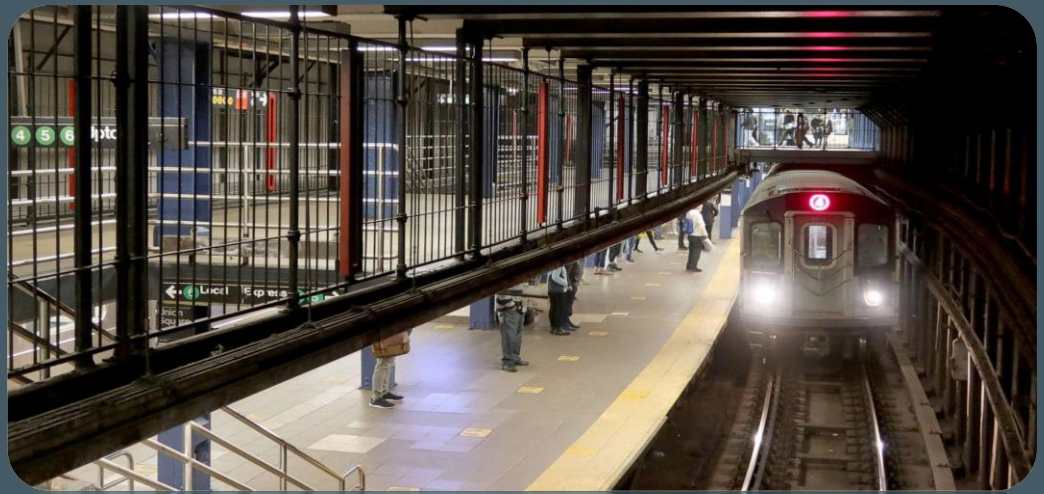
02 **DATA AND PRE-PROCCING**
Info, Features and Pre-proccing steps

03 **CHARTS**
Exploring data and visluize it

04 **CONCLUSION**
How the information extracted can be valuable to solve the problem?

CROWDED STREETS

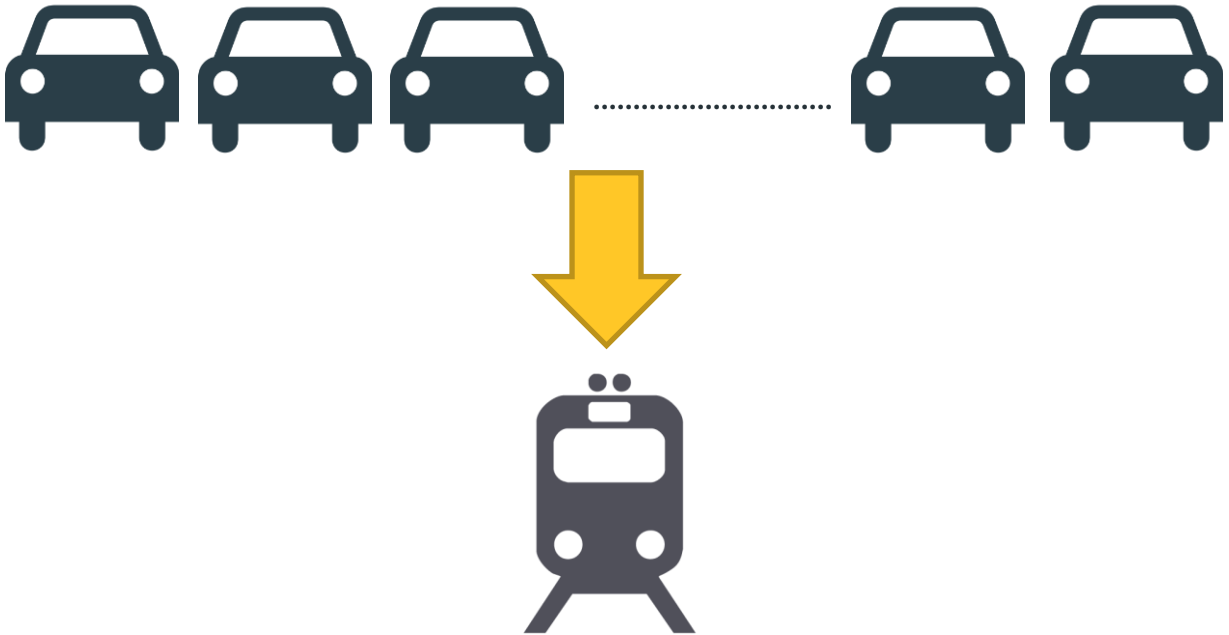
SOLUTION



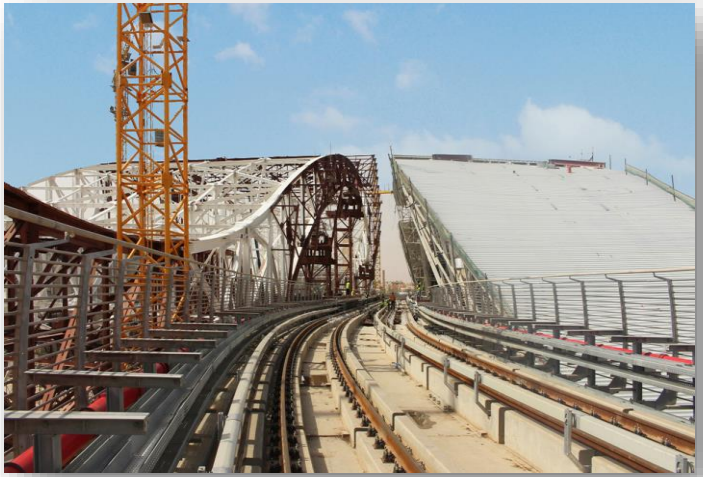
Public Transportation

SOLUTION

600 Cars



RECONSTRUCTION



Propose ideas and pinpoints the crowded stations and give idea of riders' behavior pattern to be reflect it on the MTA reconstruction design

MTA DATA

First Five Months of 2021

MTA DATA

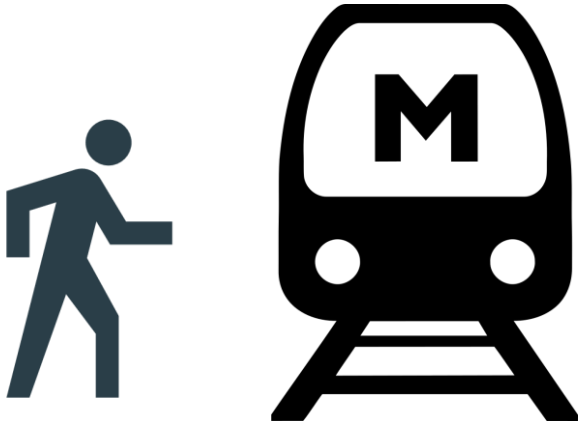
- Turnstile

- C/A : Control Area
- UNIT: Remote Unit for a station
- SCP: Subunit Channel Position
- STATION: station name the device is located at

MTA DATA

- DATE: the date (MM-DD-YY)
- TIME: the time (hh:mm:ss) for a scheduled audit event

MTA DATA



- ENTRIES: cumulative entry register value
- EXIST: cumulative exit register value
- Entries_num: exact number of entries
- Exits_num: exact number of exits
- Weekday: the weekday name
- Weekday_type: weekend or weekday
- Congestion: number of entries and exists added up

PRE-PROCESSING STEPS

01

FEATURE ENGINEERING

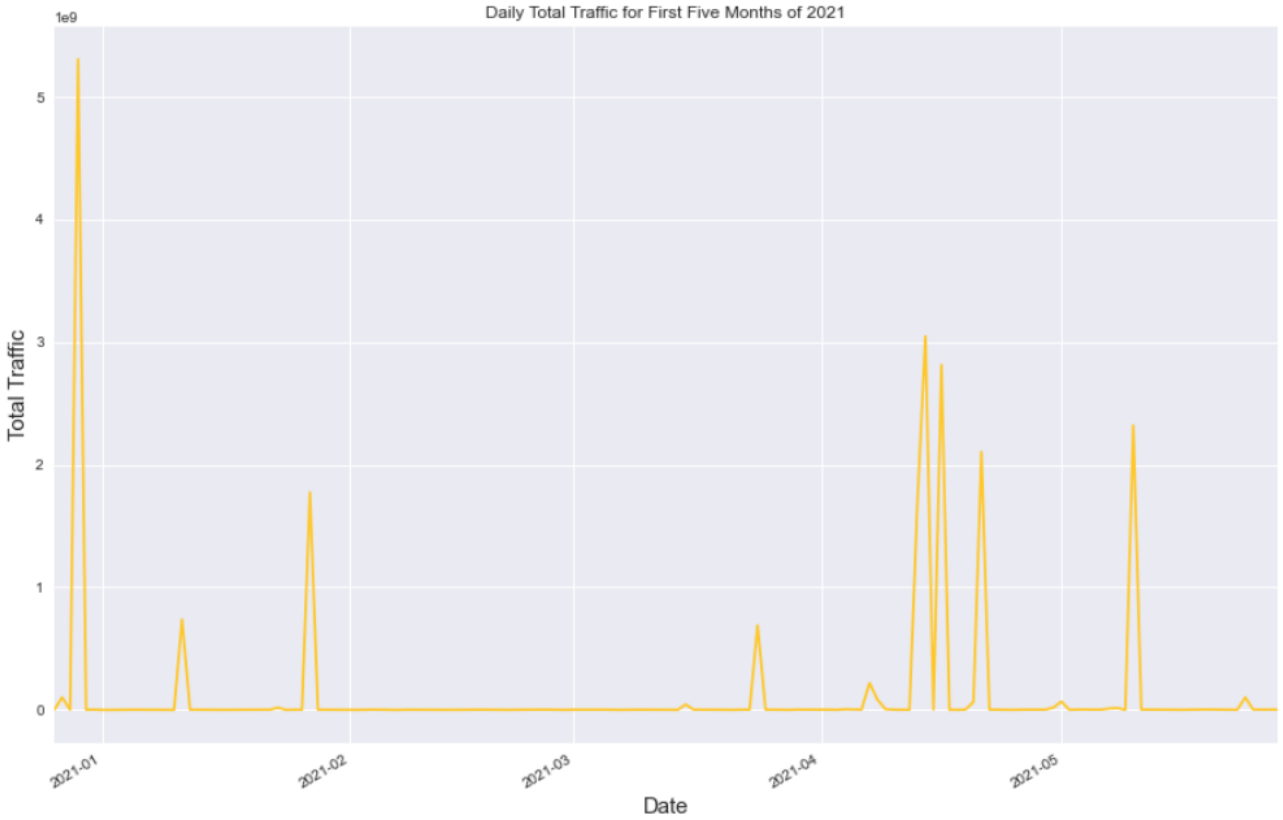
02

REMOVE DUPLICTIES

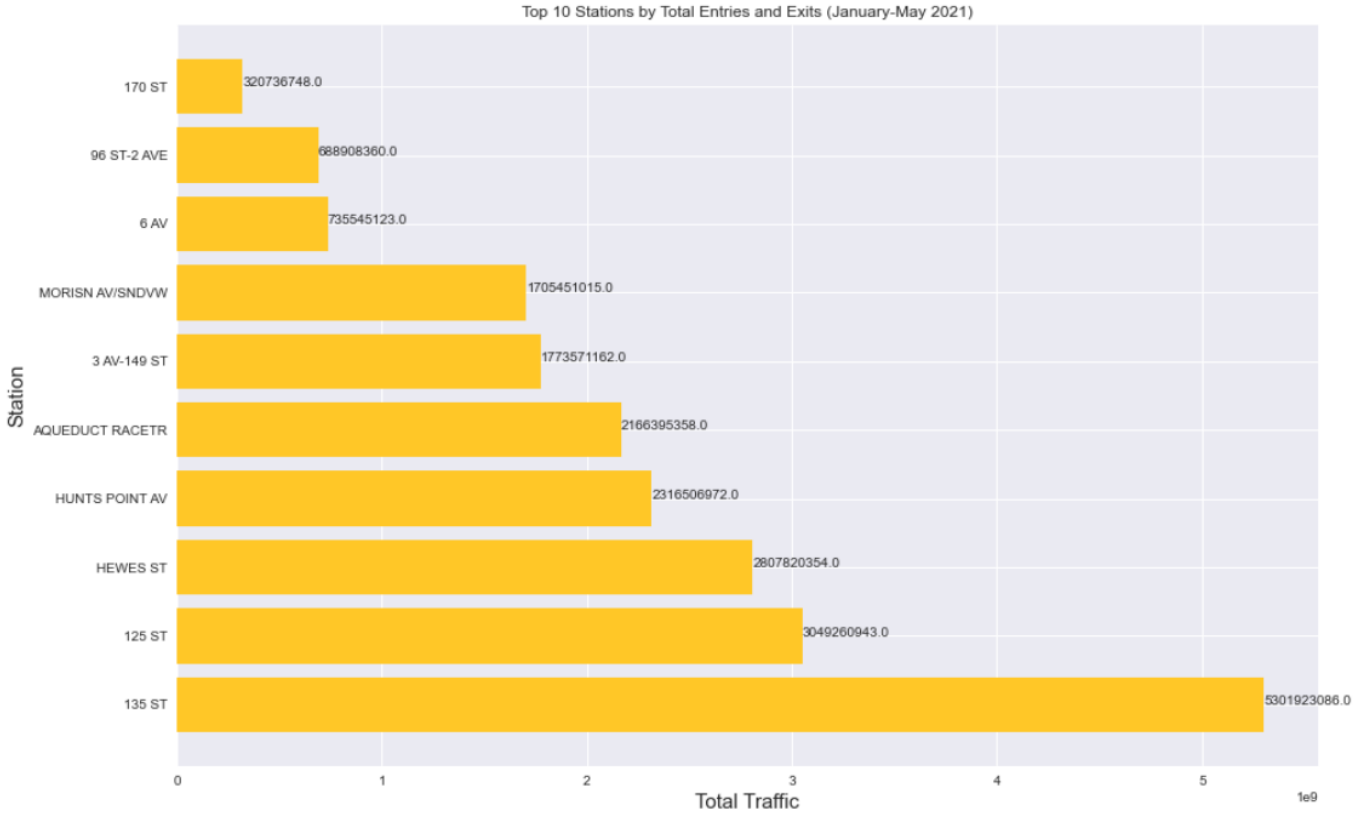
03

REMOVE NULLS

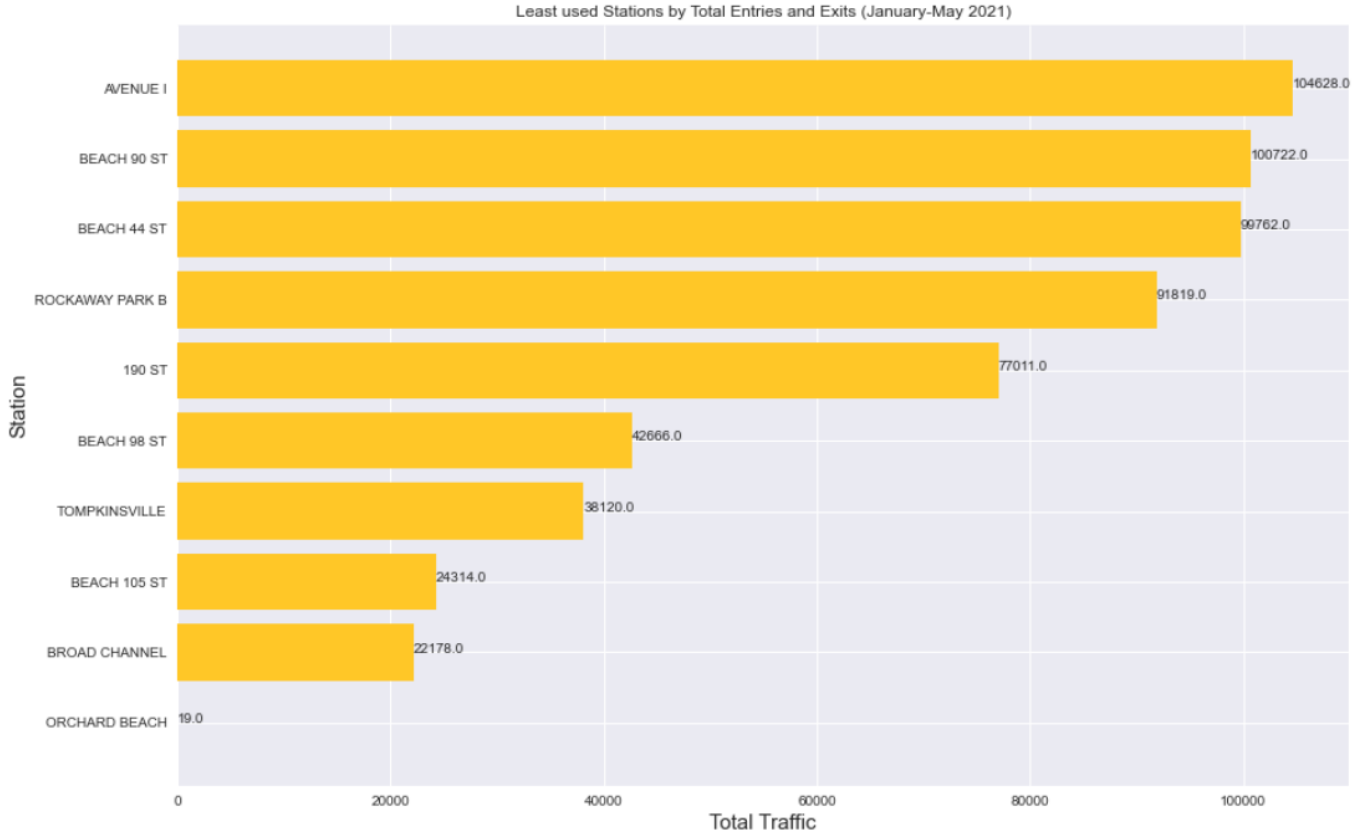
DAILY TOTAL TRAFFIC FOR THE FIRST FIVE MONTHS OF 2021?



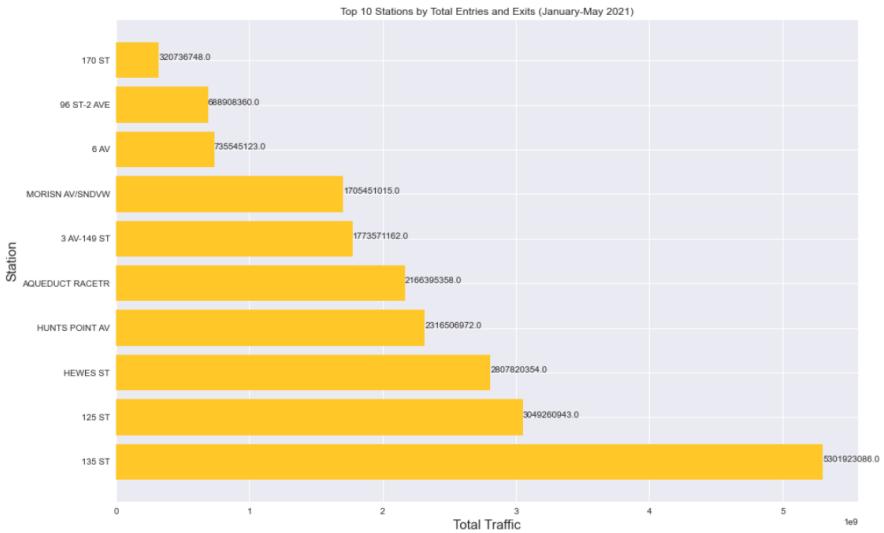
WHAT ARE THE TOP 10 STATIONS IN TERMS OF TOTAL TRAFFIC ?

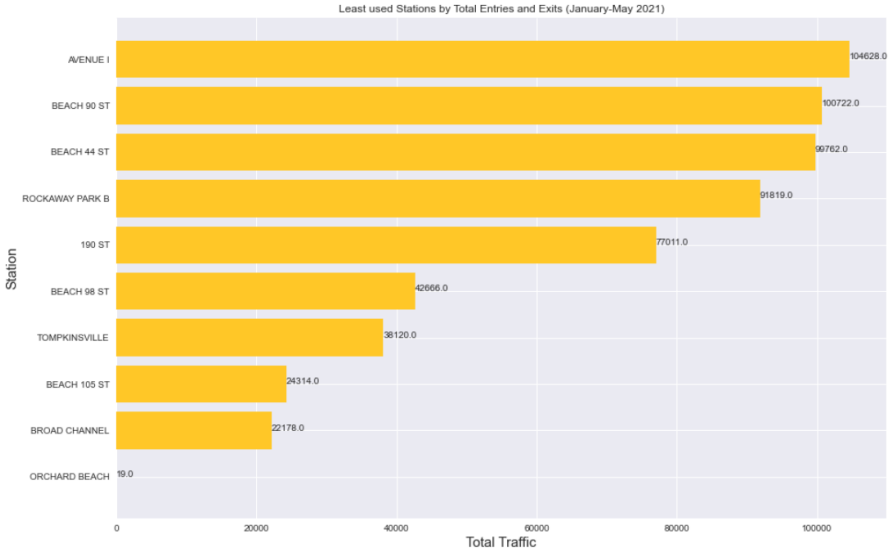


WHAT ARE THE LEAST USED 10 STATIONS IN TERMS OF TOTAL TRAFFIC ?



TOP/BOTTOM 10 STATIONS IN TERMS OF TOTAL TRAFFIC ?



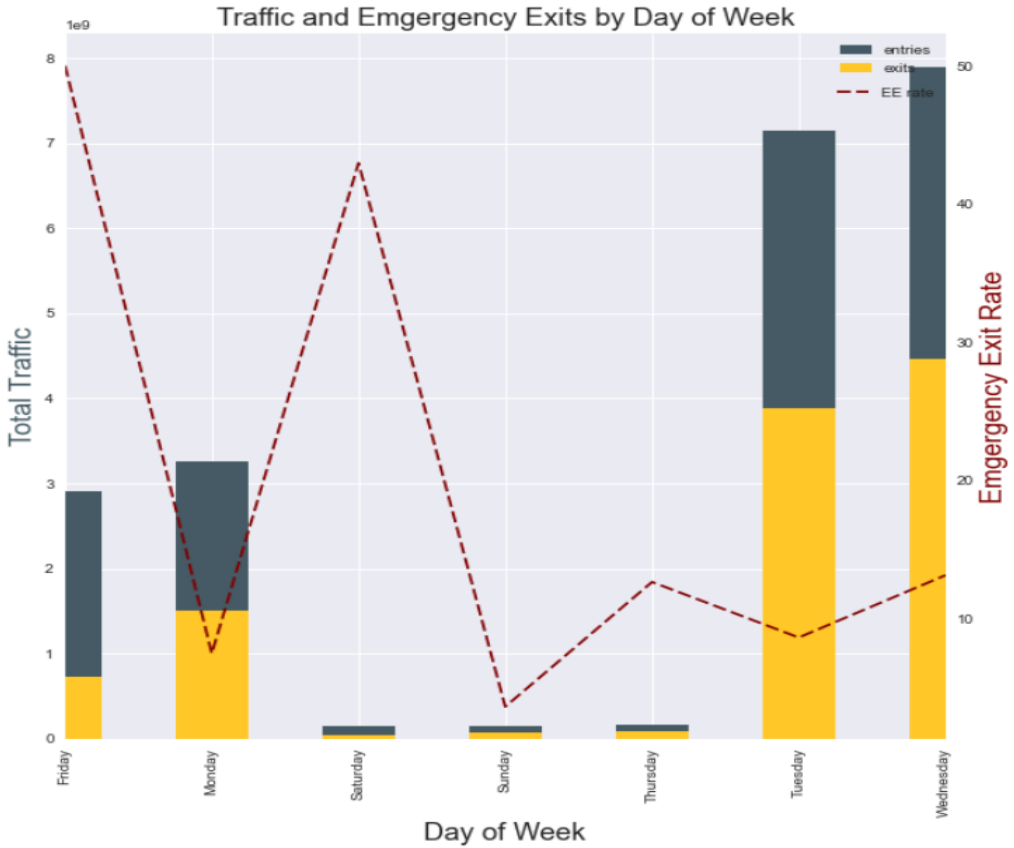


HAVE YOU NOTICED THE DIFFERENCE BETWEEN THE ENTRIES AND EXISTS ?

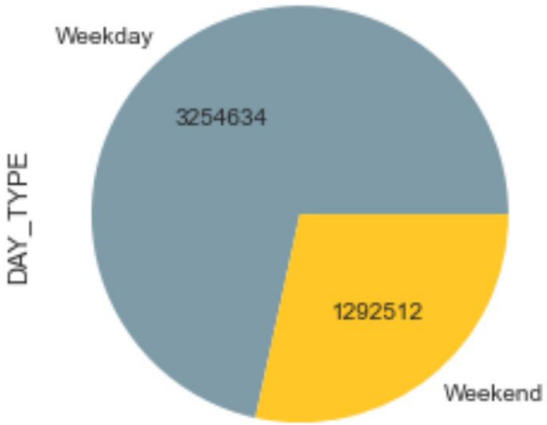
HAVE YOU NOTICED THE DIFFERENCE BETWEEN THE ENTRIES AND EXISTS ?

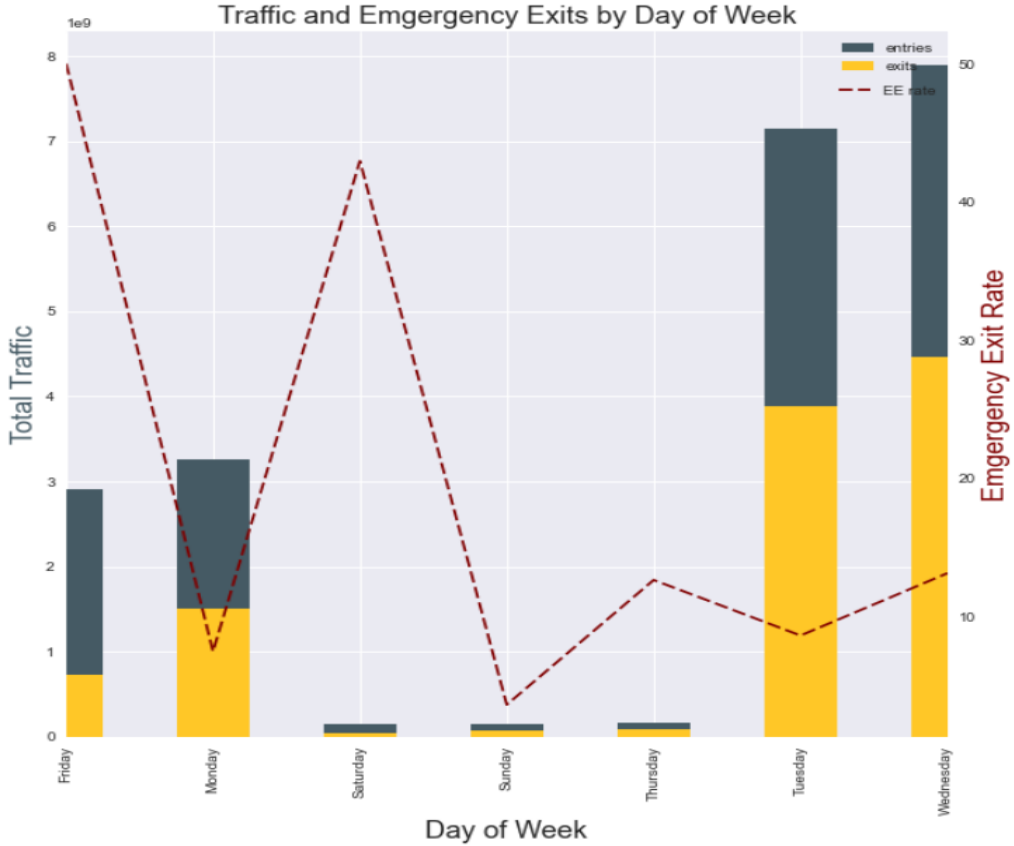
Emergency Exits Rate =(Exact Entries - Exact Exits) / Total Traffic

HAVE YOU NOTICED THE DIFFERENCE BETWEEN THE ENTRIES AND EXISTS ?



HAVE YOU NOTICED THE DIFFERENCE BETWEEN THE ENTRIES AND EXISTS ?





CONCLUSION

- Least trafficked stations can be closed and enlarge the most trafficked stations.
- Riders use the metro stations more during weekdays therefor more management records can be applied
- Insights can be tricky due to the fact we have skewed data –More weekdays records-
- Emergency exists are used for non-emergence events

THANK YOU FOR LISTNING