

...

WTWY MTA TRAFFIC ANALYSIS

ÖZLEN SÖNMEZ
KÜBRA ERENŞOY
HÜSEYİN AYDOĞAN
ATALAY BARUT
MUHAMMED KARAÇOBANOĞLU



As First Project of
Istanbul Data Science Academy

Introduction

WomenTechWomenYes is holding an annual gala, and hopes to invite as many interested individuals as possible.



Our Goal

To find out the which stations and days in a week are the busiest in the resource.

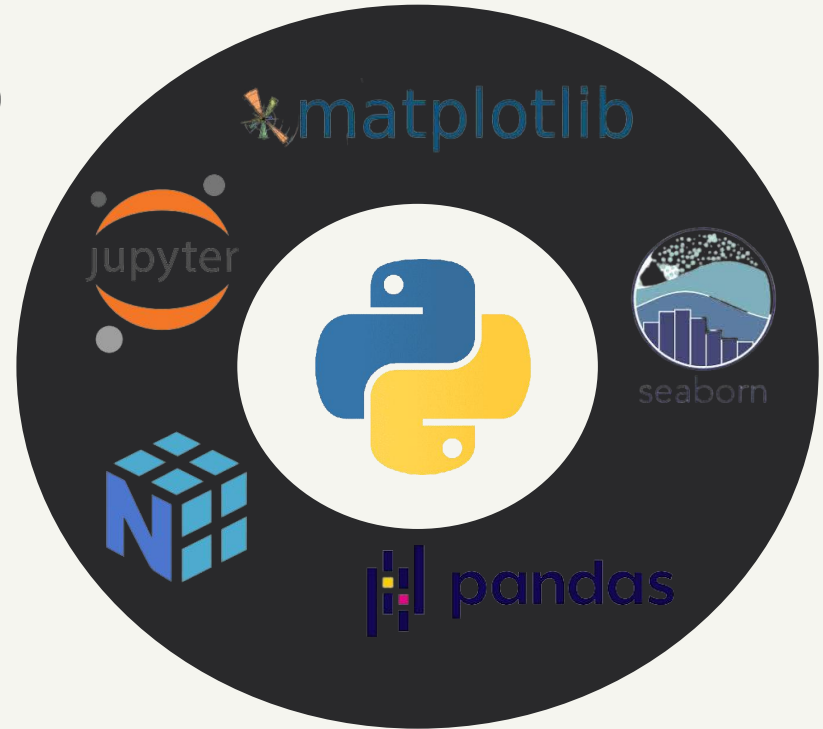


Methodology

- MTA Turnstile Data (January 2018 – 2021)

Tools

- Python
- Matplotlib
- Pandas
- Numpy
- Seaborn
- Jupyter Notebook





DATA EXPLORATION

Looked For Duplicated And Null Values In The Data Set.

Dropped Some Columns From The Data And Changed The Names Of Them.

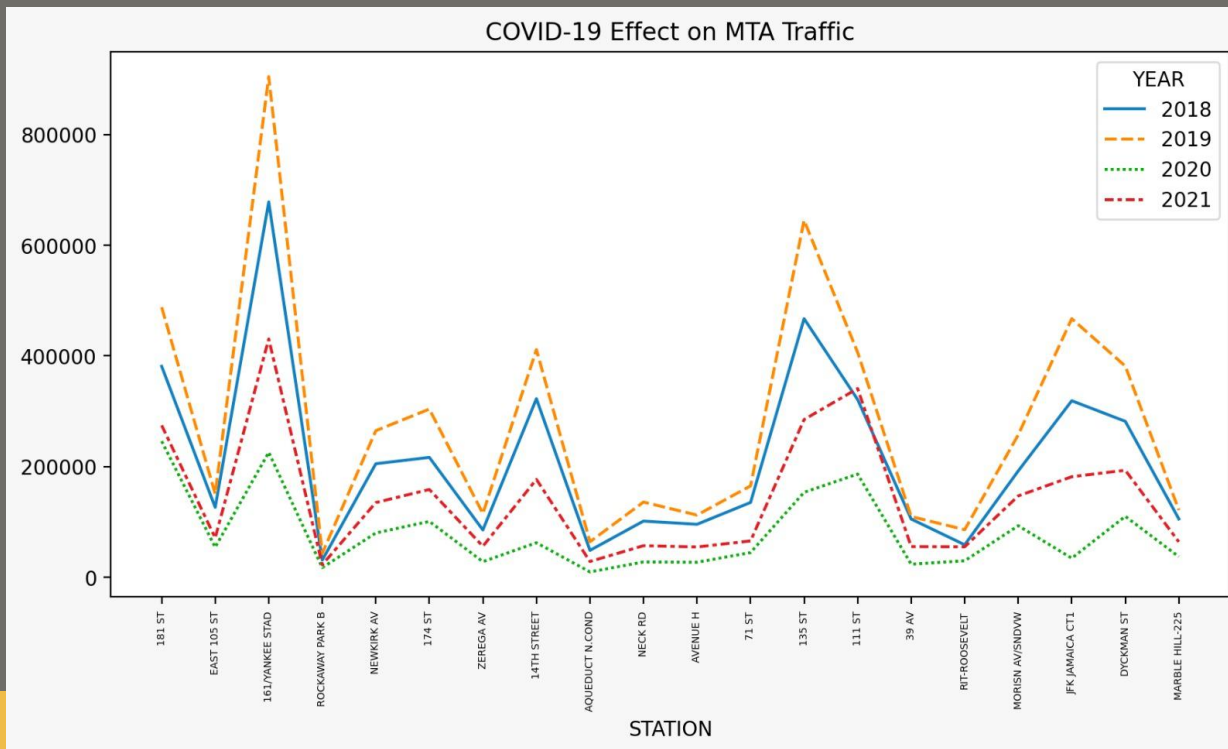
Added The Day Column To Data Set.

```
mta.head(5)
```

	ca	unit	scp	station	date	time	desc	entries	exits	datetime	yeardate	day	turnstile	day_date
0	A002	R051	02-00-00	59 ST	06/19/2021	00:00:00	REGULAR	7589159	2594185	2021-06-19 00:00:00	2021	Saturday	59 ST-A002-R051-02-00-00	2021-06-19_Saturday
1	A002	R051	02-00-00	59 ST	06/19/2021	04:00:00	REGULAR	7589165	2594190	2021-06-19 04:00:00	2021	Saturday	59 ST-A002-R051-02-00-00	2021-06-19_Saturday
2	A002	R051	02-00-00	59 ST	06/19/2021	08:00:00	REGULAR	7589174	2594207	2021-06-19 08:00:00	2021	Saturday	59 ST-A002-R051-02-00-00	2021-06-19_Saturday
3	A002	R051	02-00-00	59 ST	06/19/2021	12:00:00	REGULAR	7589214	2594253	2021-06-19 12:00:00	2021	Saturday	59 ST-A002-R051-02-00-00	2021-06-19_Saturday
4	A002	R051	02-00-00	59 ST	06/19/2021	16:00:00	REGULAR	7589300	2594278	2021-06-19 16:00:00	2021	Saturday	59 ST-A002-R051-02-00-00	2021-06-19_Saturday



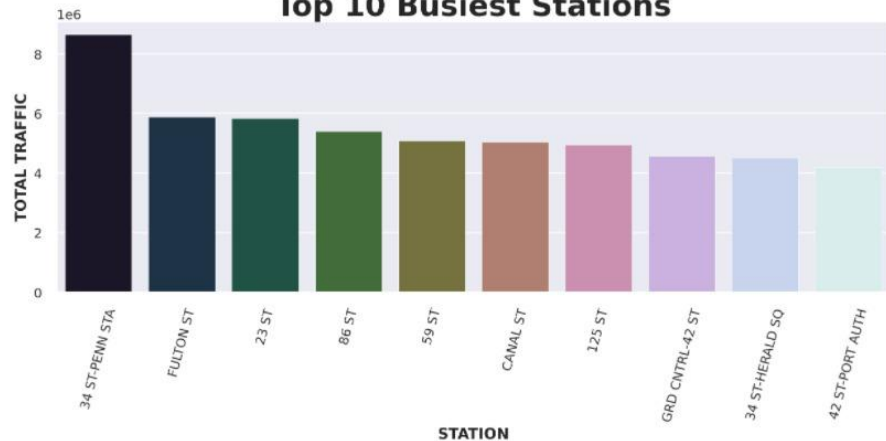
PANDEMIC EFFECT



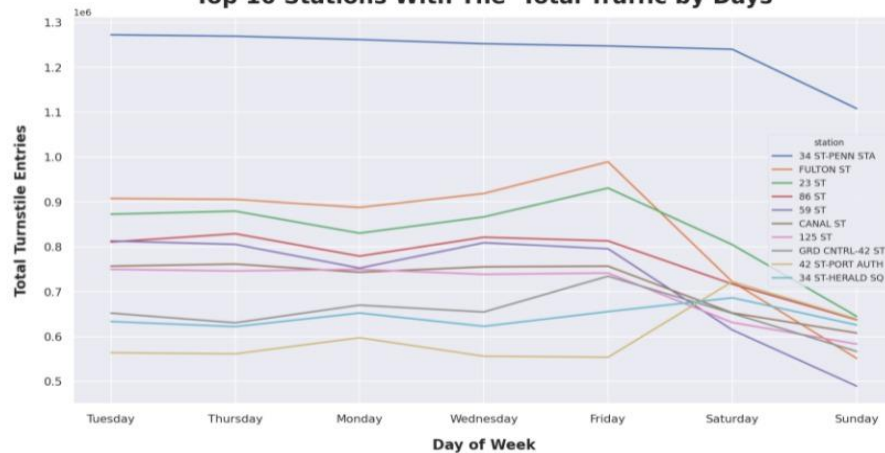


RESULTS

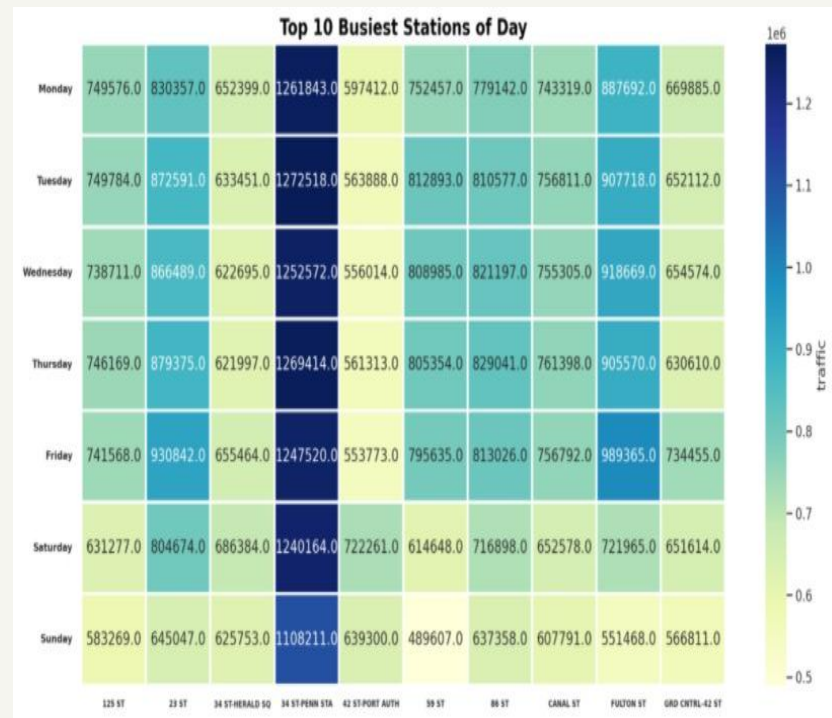
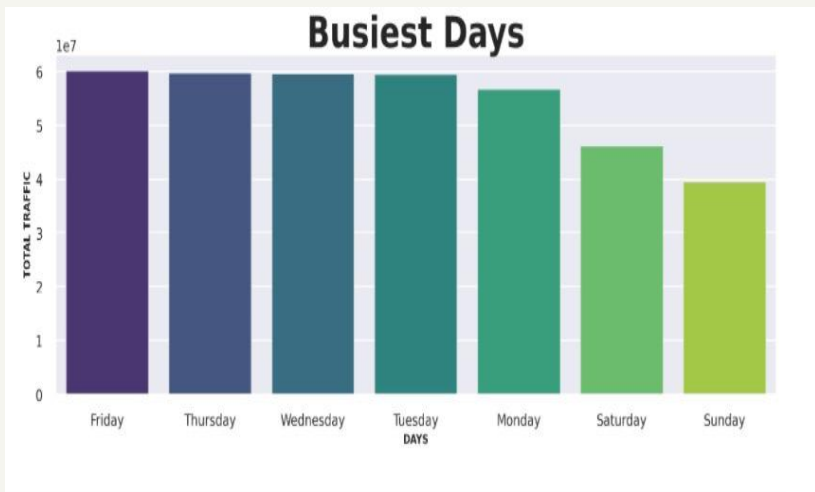
Top 10 Busiest Stations



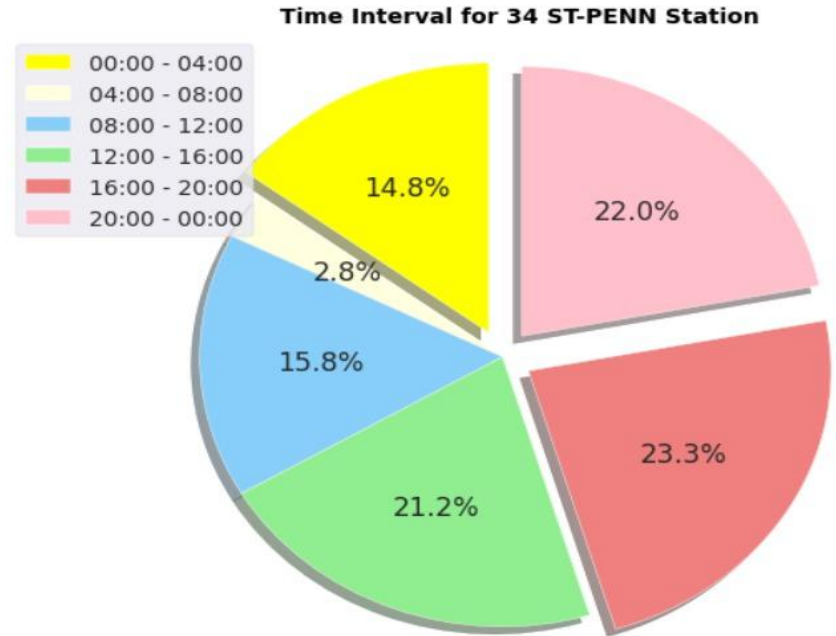
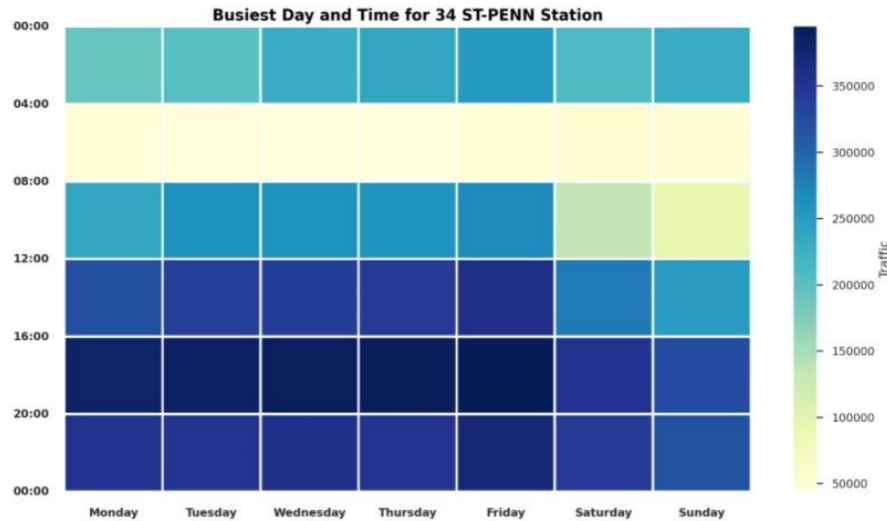
Top 10 Stations With The Total Traffic by Days



RESULTS



RESULTS



CONCLUSION

Weekdays are busier than Weekends.
The busiest days are Tuesday,
Thursday and Wednesday.

WTWY should focus their street
efforts in the top 10 stations,
preferably in the busiest days to avoid
wasting their resources.

34 ST-PENN STATION should be the
first option If time and resources are
limited. The best time interval is 16:00
to 20:00 due to the highest traffic
during the day.



FUTURE WORK

- Improve our understanding of available resources.
- Expanding the analysis to include previous months as well as new incoming data.
- Using different dataset.

