MTA STATION FOOT TRAFFIC

Motivation:

Have there been significant changes in Station usage pre-COVID versus the current year that would affect retailers dependent on foot traffic?

Objectives:

Analyze which stations had the greatest change in usage.

Methodology

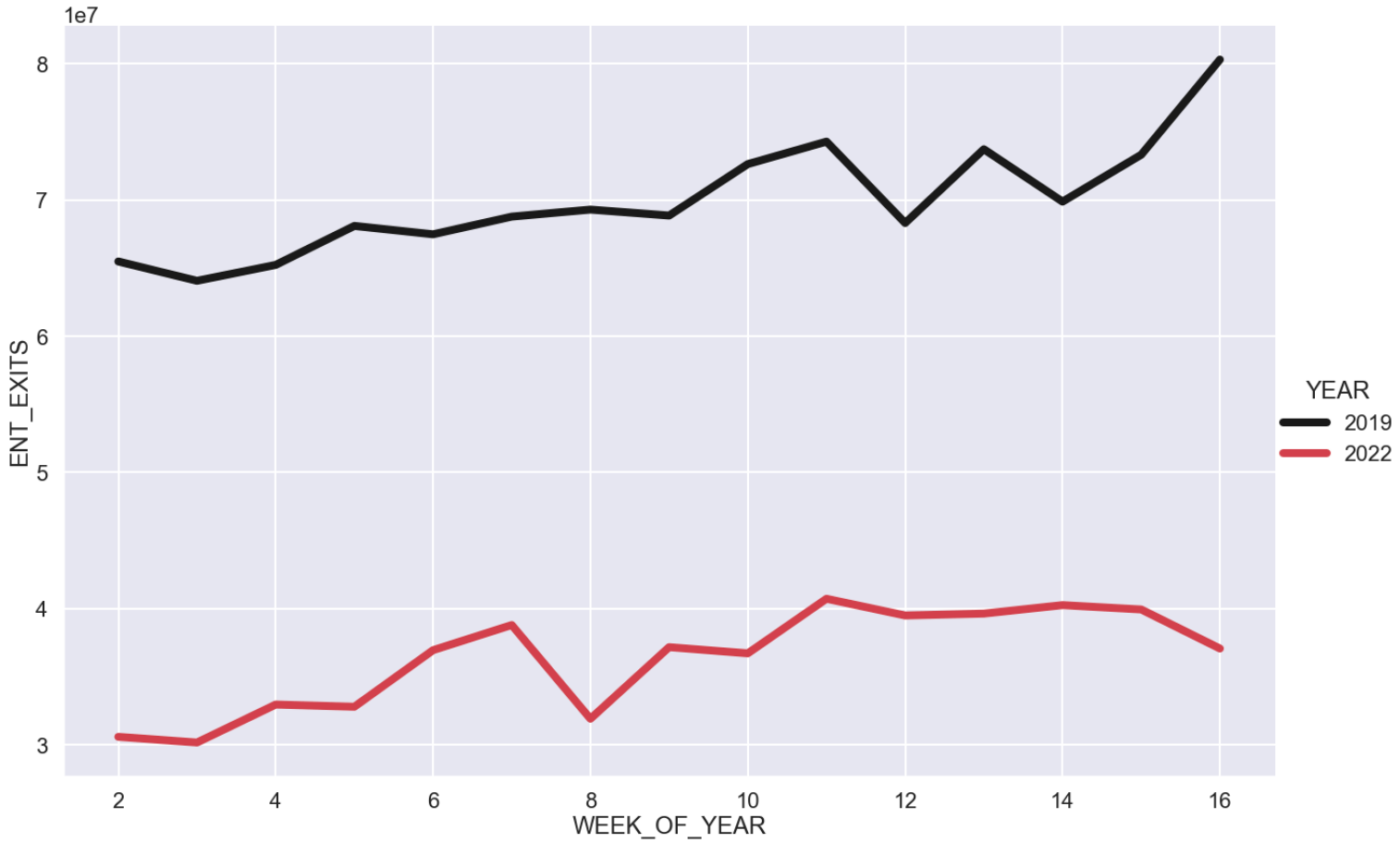
- Utilize readily available MTA turnstile from the first 4 months of 2019 and the first 4 month of 2022.
- SQLite and Python to analyze data
- Graph data with Matplot lib and Seaborn
- Focus on Change in turnstile usage from 2019 to 2022
- Main Metrics are total usage change by station and usage percentage change by station

2019 vs 2022 Entries and Exits by Week

Significant Decrease in 2022 vs 2019

Both years show some Seasonality

Diverge on Weeks 15 and 16

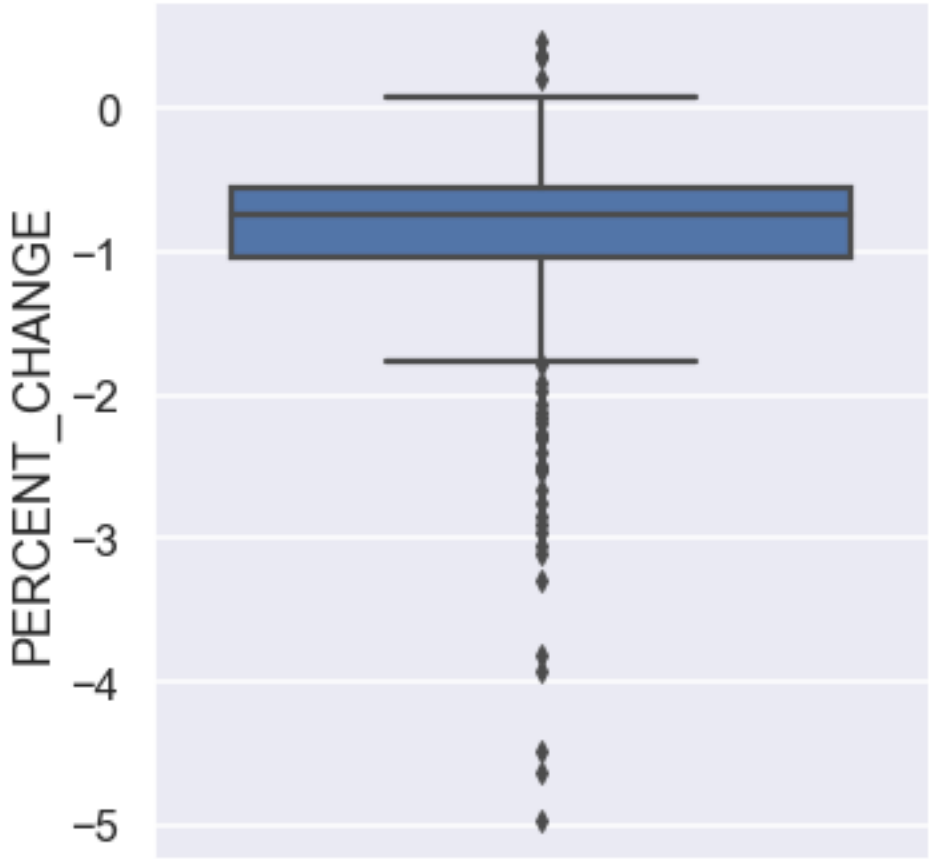


Percent Change from 2019 Exits and Entries

Majority of the Stations are in the -50% to -100% range

Quite a few outliers on the negative side

There are a few Stations with positive increase in usage.

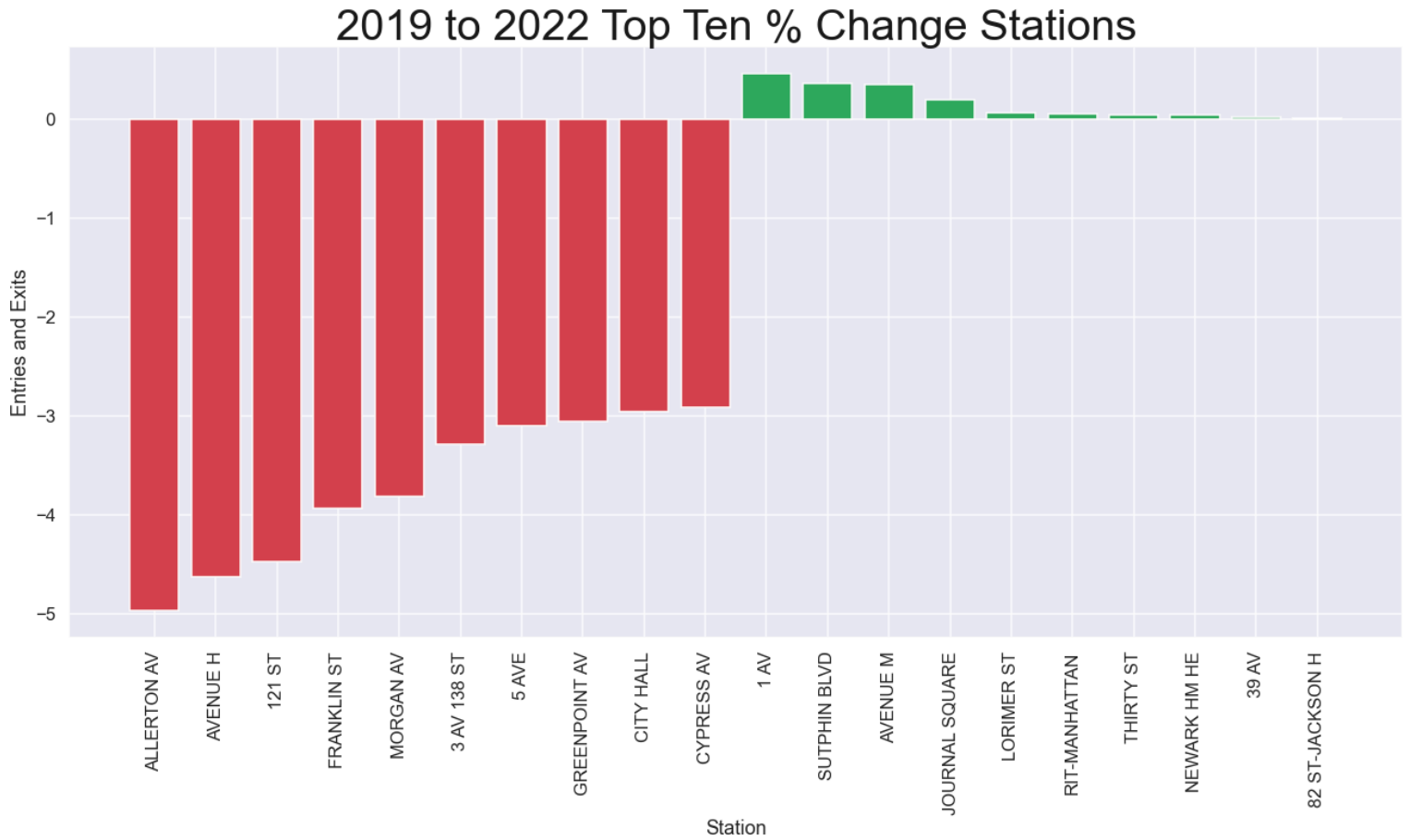


Entries and Exits Percent Change by Station

Focus on outliers as most Impactful changes

There are a few stations with significant increases

Need to view all the stations



Conclusions

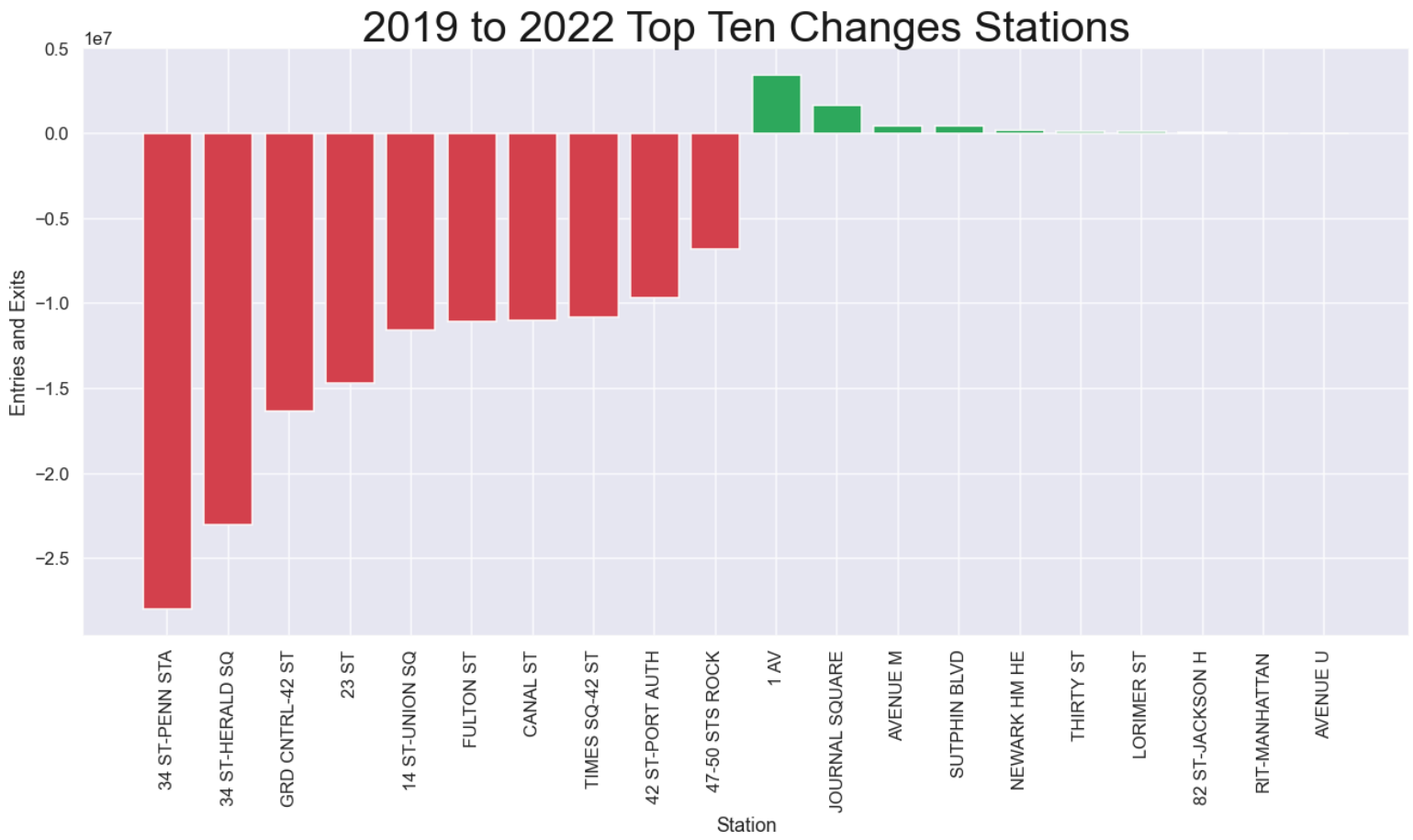
- Significant decrease in turnstile usage at most stations could have a strong negative impact on retailers relying on foot traffic

- There are some stations with increases or almost flat changes that could be focused on

Next Steps

- Use Geo-location data to better understand where the changes are
- Create a Tableau dashboard to better view and filter on all or a subset of the stations
- Compare additional years of data

Entries and Exits Total Change by Station



Box plot with removed outlier

Removed outlier with -250000%