

MTA Data Analysis to find the best location for a coffee shop in NYC

Praveena Suresh

Aug 18, 2021

Business Need and Objective

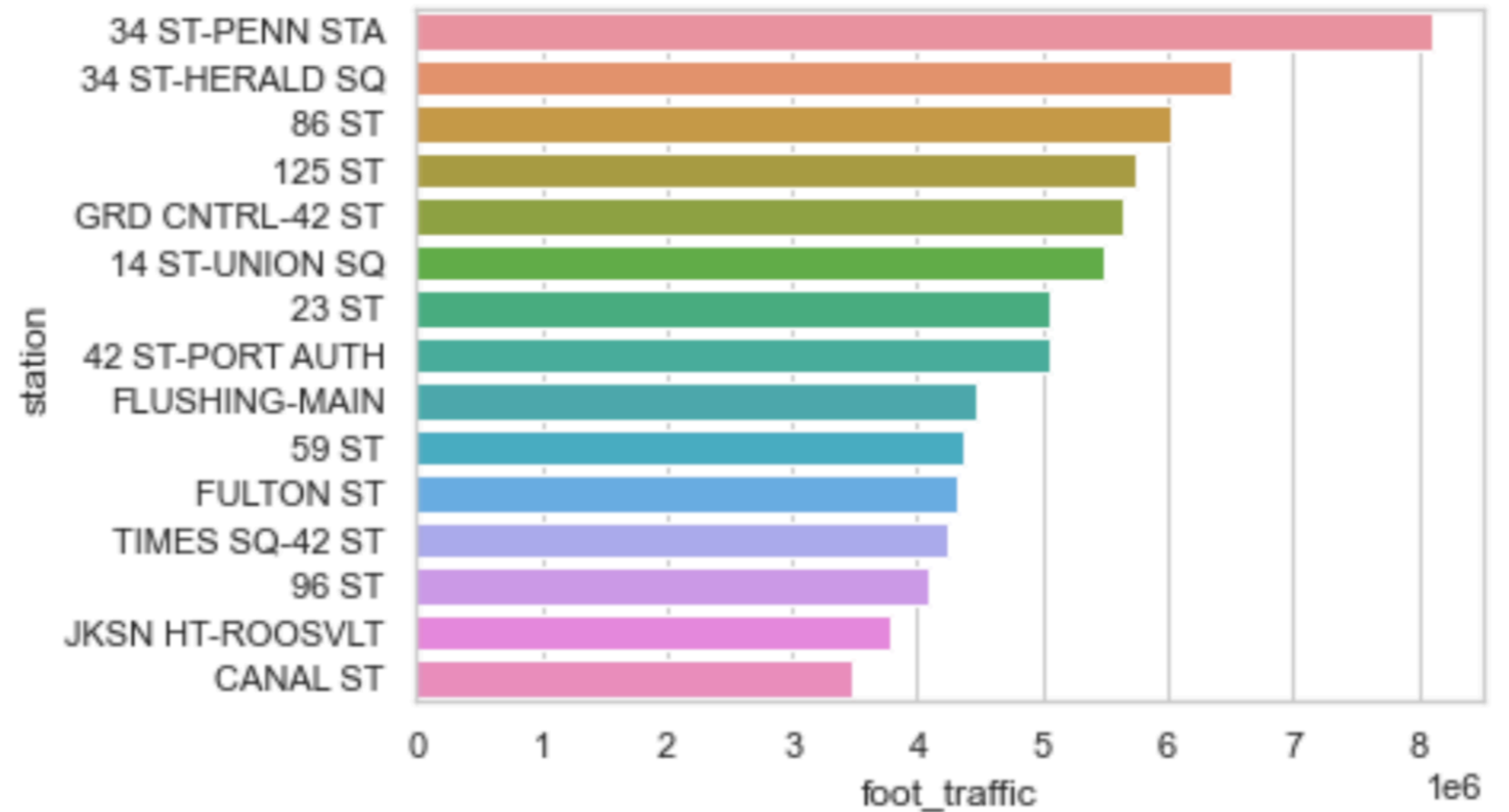
- Find an optimal location for a coffee shop in NYC which has easy access to subway stations
- Commuters are more likely to have coffee on their way to work or when they return home
- Identify the busiest stations and also the time when there is maximum foot traffic
- Pay attention to change in patterns based on different seasons

Data Exploration and Cleaning

- I imported the data for the MTA turnstiles to a local database using SQLite and used SQL to query from that database into Python via SQLAlchemy.
- The turnstile data is a series of data files containing cumulative number of entries and exits by station, turnstile, date and time. Data is typically collected every 4 hours.
- In this analysis I use 4 months worth of data which includes the months of Sep 2020, Dec 2020, Mar 2021 and June 2021.
- I calculated the actual entries and exits for each time interval and used their sum to come up with the “foot traffic” at each subway station.

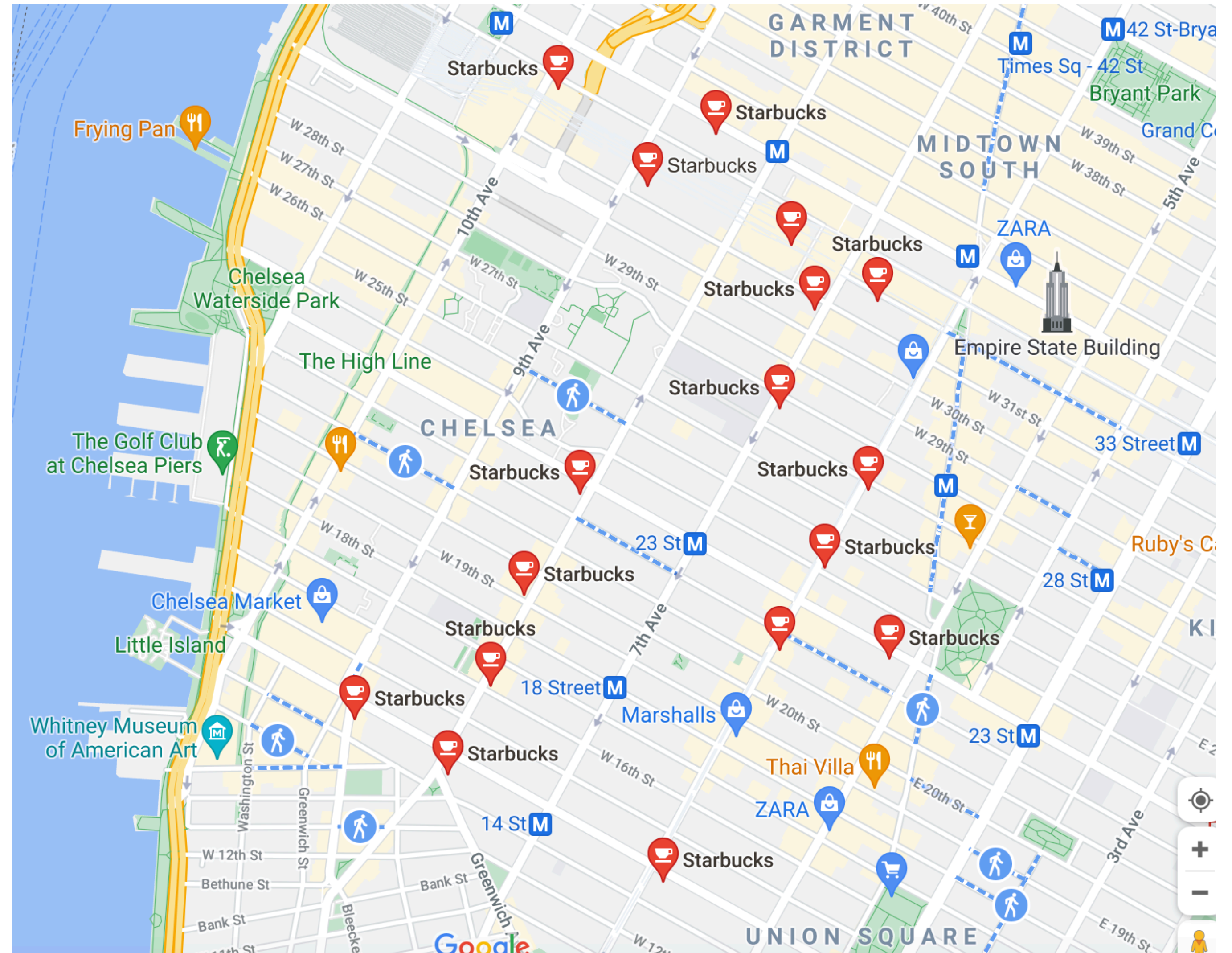
Top 15 busiest stations in NYC

- Identify the top 15 locations NY



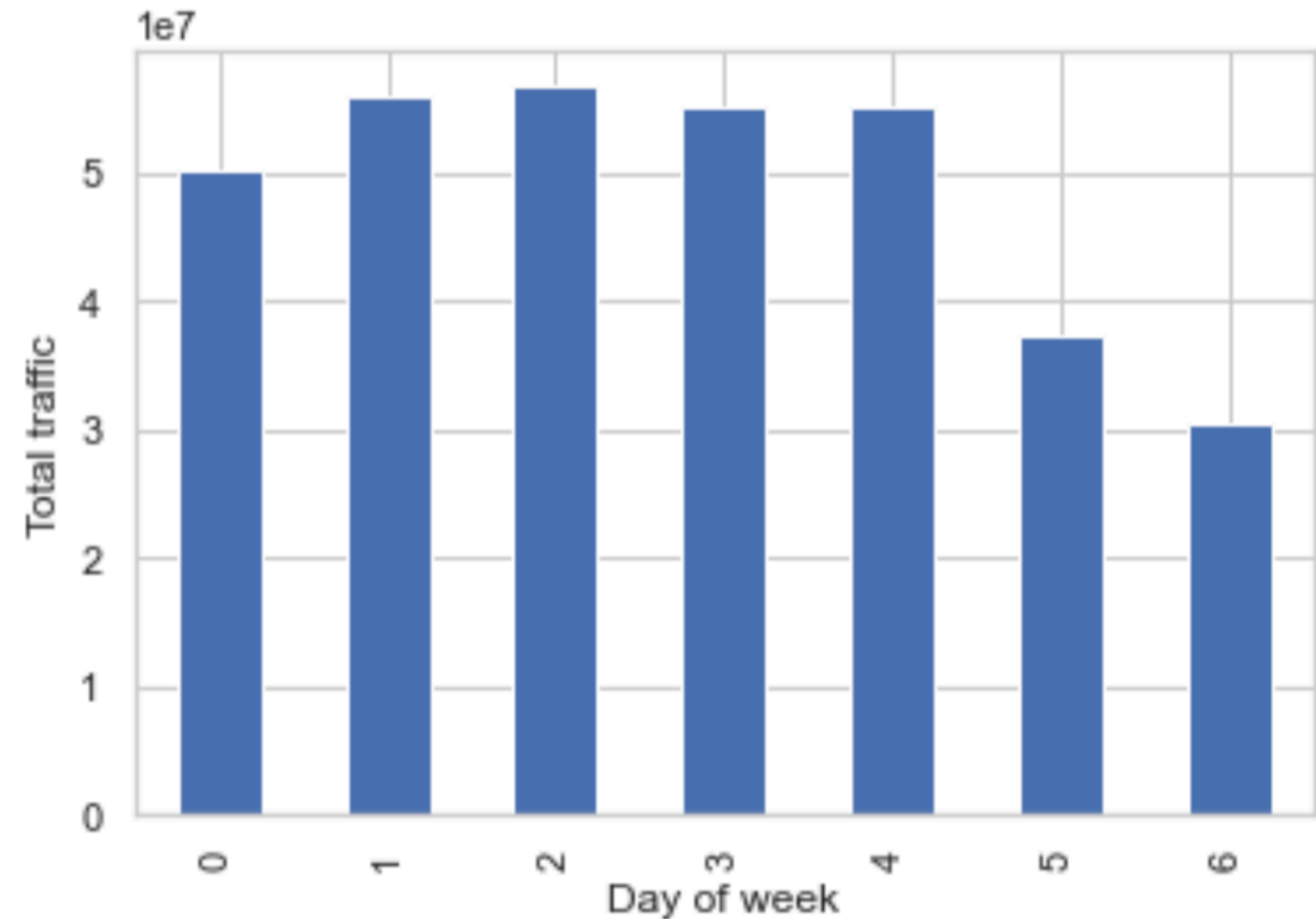
Competitors

- Identify other coffee shops that are located near MTA stations and compare their foot traffic, they are not necessarily near the top busiest station, take the case of 23 ST.
- Selecting a location which has the top most traffic might imply that it is too busy for people to stop and buy coffee
- Instead picking a location which is in the bottom 5 of the 10 locations would be more beneficial, hence we are going with 42 ST- Port Authority which has foot traffic in the range of 23 ST.



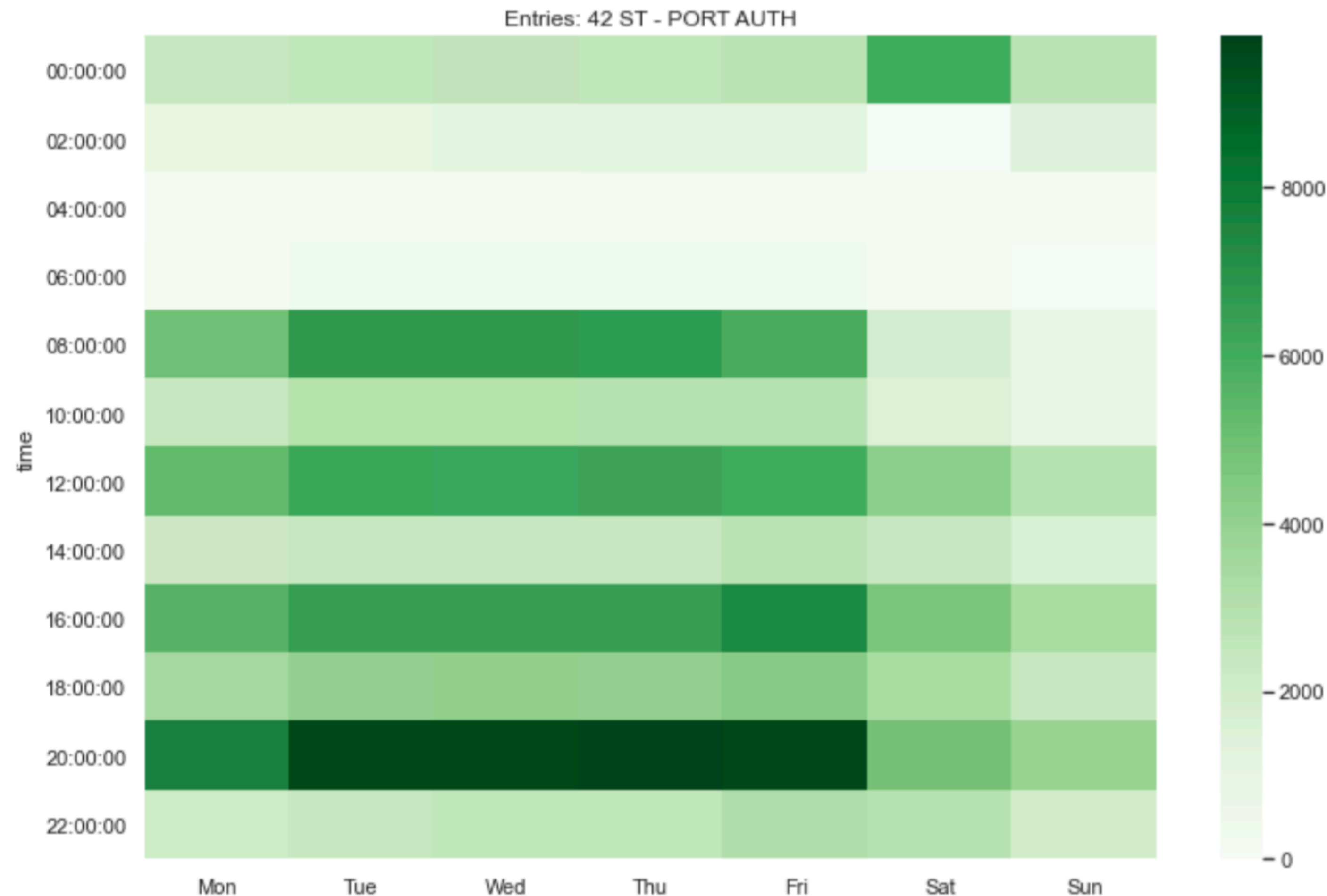
Traffic patterns on Weekday vs Weekend

- Looking at the traffic patterns on Weekdays vs Weekends, it clearly shows the Weekdays are more busy than weekends, especially Tuesday and Wednesdays seem to be more popular.

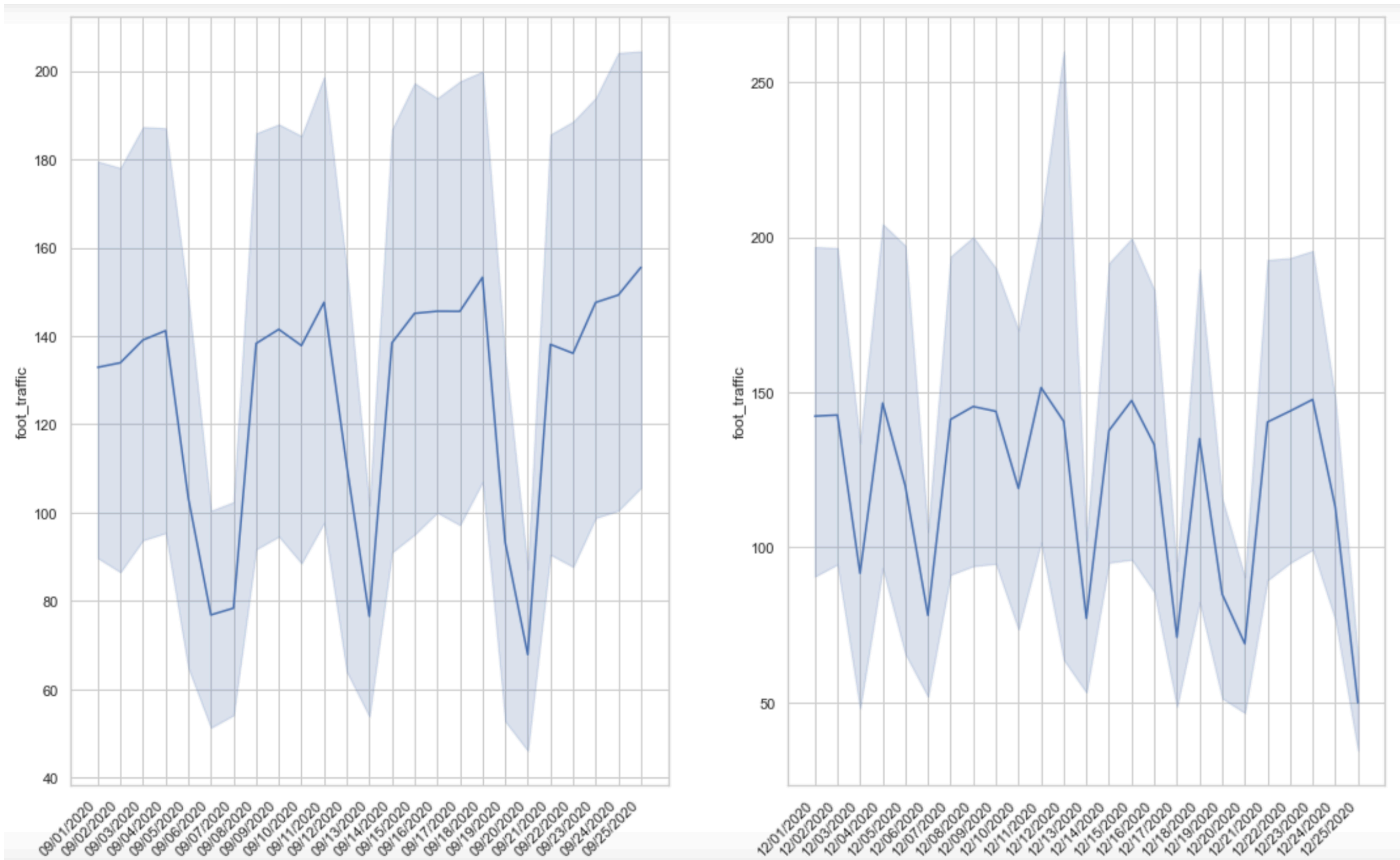


Busiest times of the day

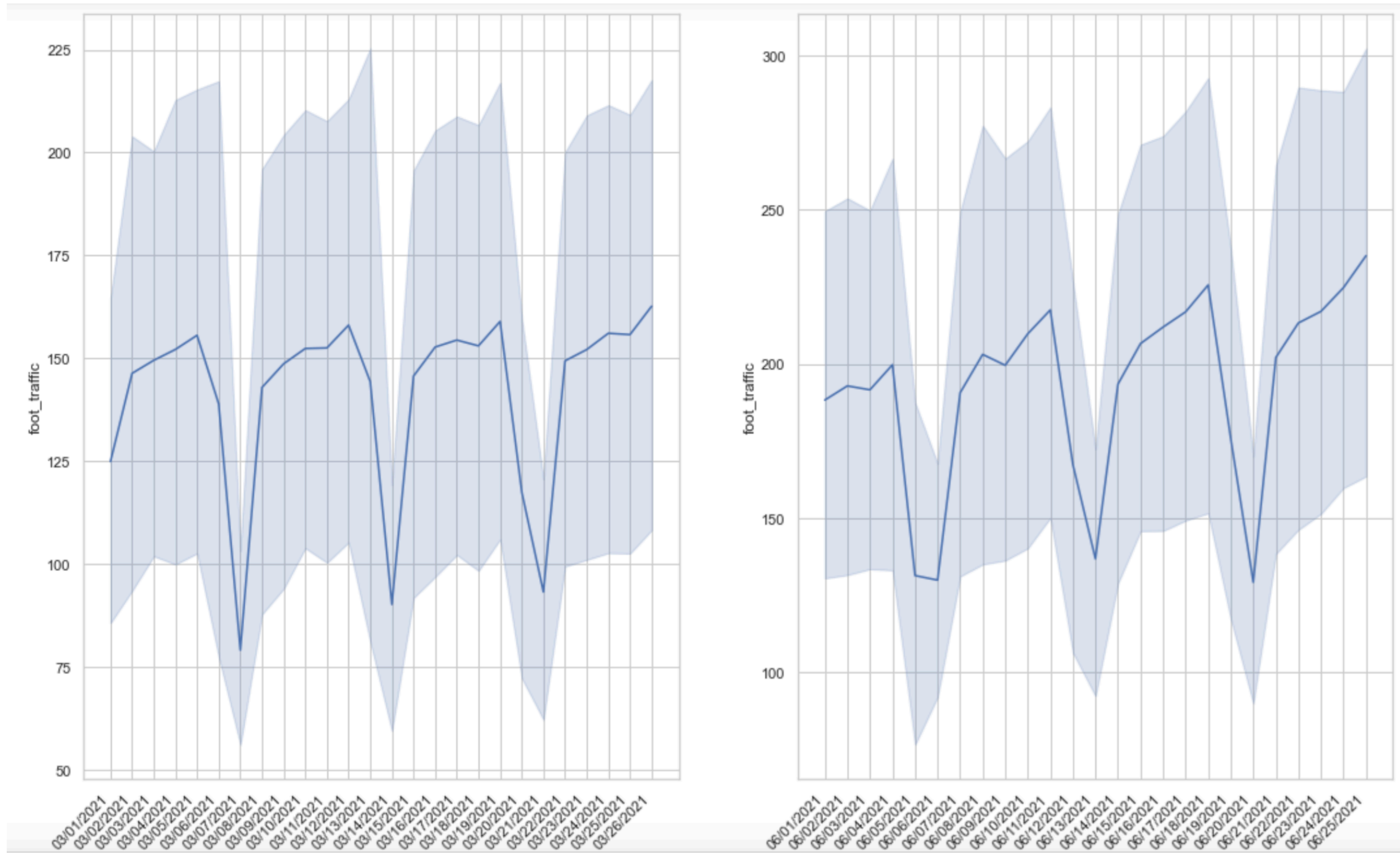
- When looking at the foot traffic in 42 ST - Port Authority, it seems the traffic is highest at 8pm. This is not surprising given the fact many companies in NYC let their employees work from home due to the pandemic which has resulted in less crowd during the morning rush hour. Other popular times are:
 - 8am - Morning is by far the time when the majority of people have coffee
 - Noon - This is a good time to expect people who like to take in caffeine when they pick up lunch
 - 4pm - Plan to accommodate people who might grab coffee on their way home



Traffic patterns in Fall and Winter



Traffic patterns in Spring and Summer



Conclusion

- From my analysis of the MTA data, I can conclude that weekdays are more busier than weekends.
- Also you can expect more crowd at a coffee shop around 8am, noon and 4pm.
- Summer time seems to have the maximum traffic while Fall and Spring have moderate traffic, not surprisingly Winter seems to have the least.
- The above insights can be used to pick an optimal location and also plan to staff the shop based on demand during different times.

Next steps

- Compare the performance of other coffee shops in the area and try to see what are other factors that can impact the business...