Introduction

- Motivation: to increase overall womens particapation in tech fields
- Objective: raise awareness and involvement for the oganization WTWY
- Goald: find how WTWY can be most effective in gaining support around NYC

Methodology

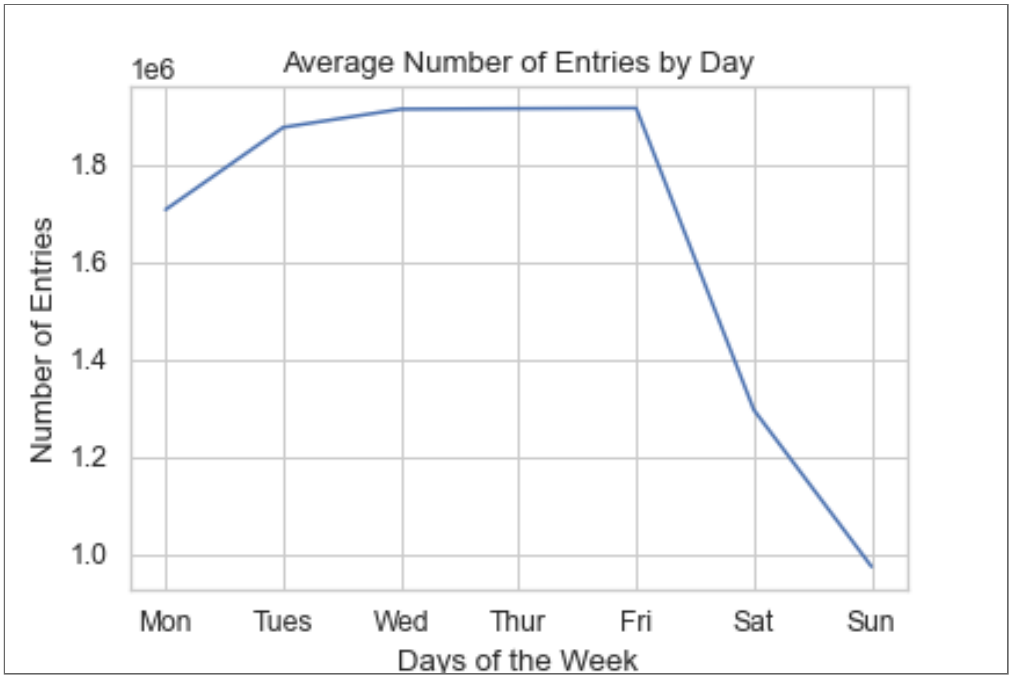
Primary Data source:

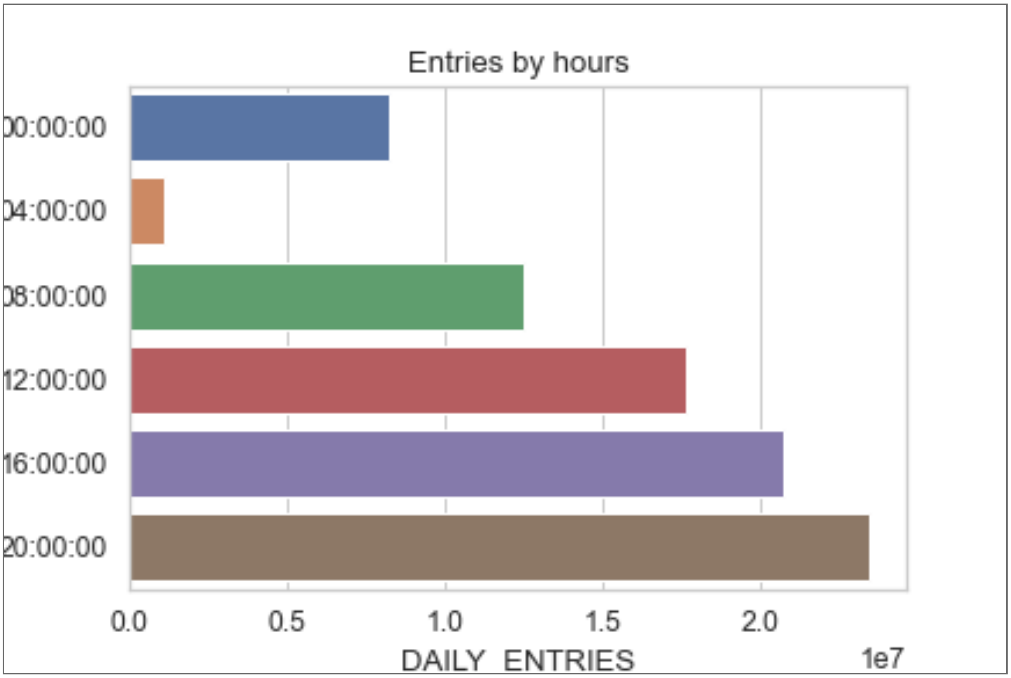
Most of the primary data came from the MTA NYC turnstile data, with some tertiary data coming from CollegeSimply for enrollment numbers, and google maps to check locations.

Methodology Cont

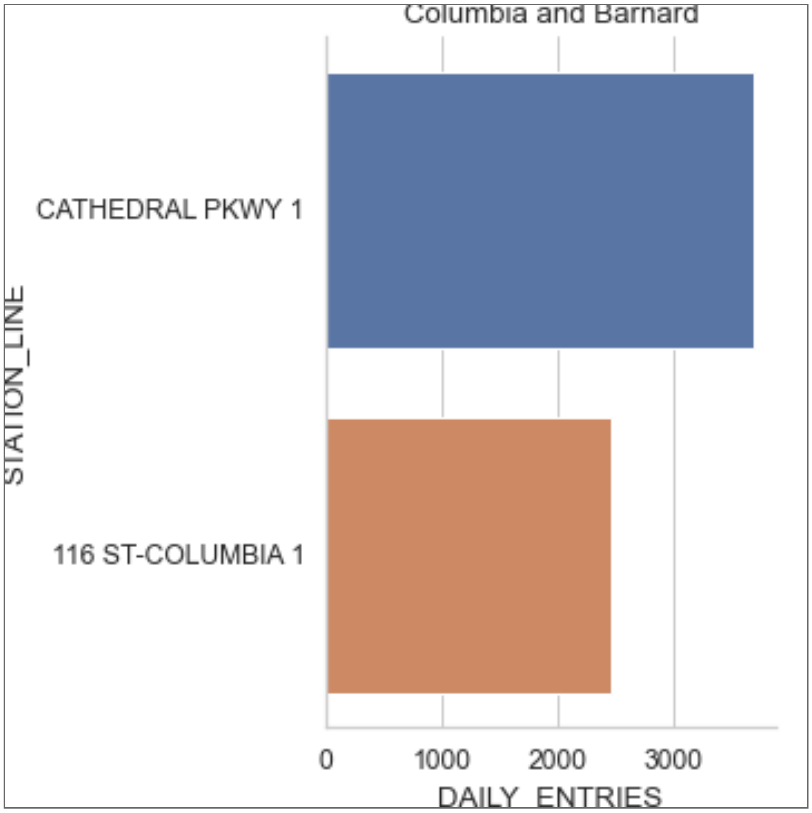
- Using Python and SQL we are able to clean up our data sources and present it graphically making it easier to read and more accessible.
- Key metrics we are looking at include busiest subways, peak times, including days and hours, and proximity to universities/colleges.

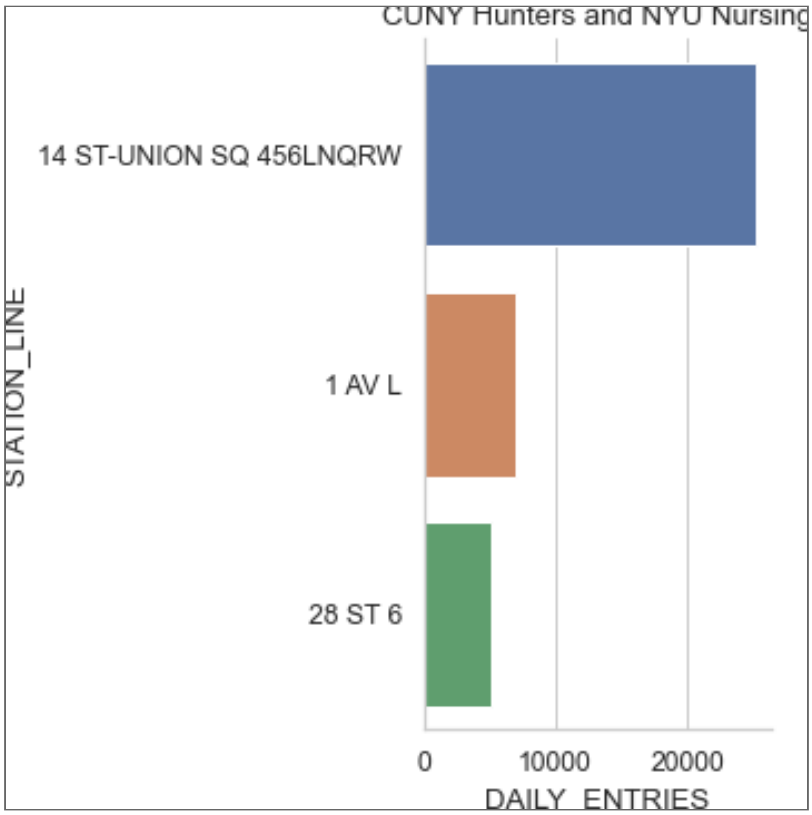
Results


Average Number of Entries by Day
1e6
1.8
1.6
1.4
1.2
1.0
Number of Entries
Mon
Tues
Wed
Thur
Fri
Sat
Sun
Days of the Week

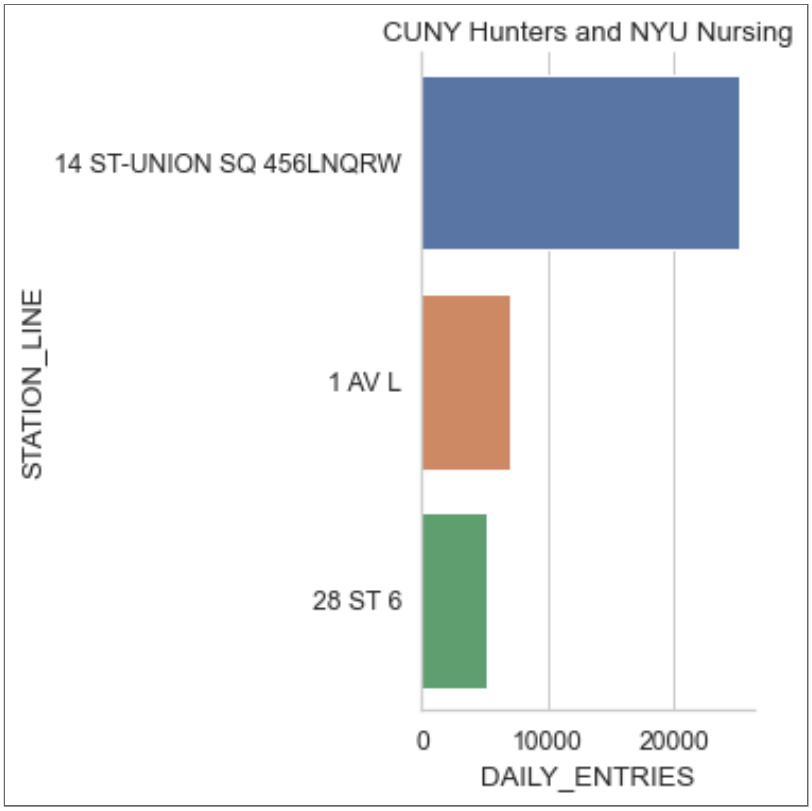

Entries by hours
00:00:00
04:00:00
08:00:00
12:00:00
16:00:00
20:00:00
0.0
0.5
1.0
1.5
2.0
DAILY_ENTRIES
1e7

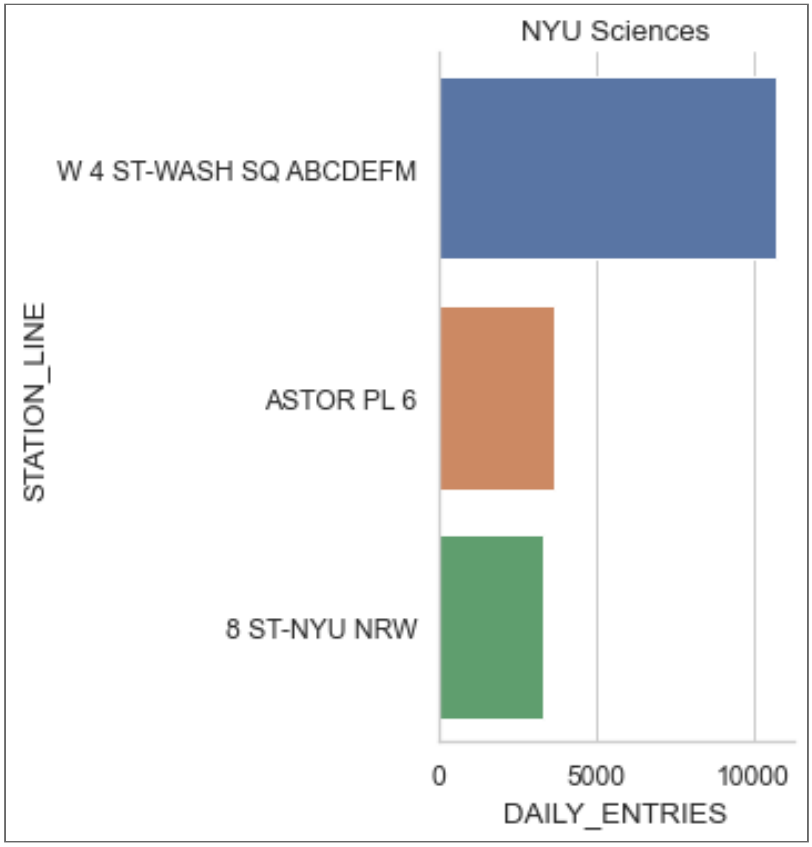
Results Cont



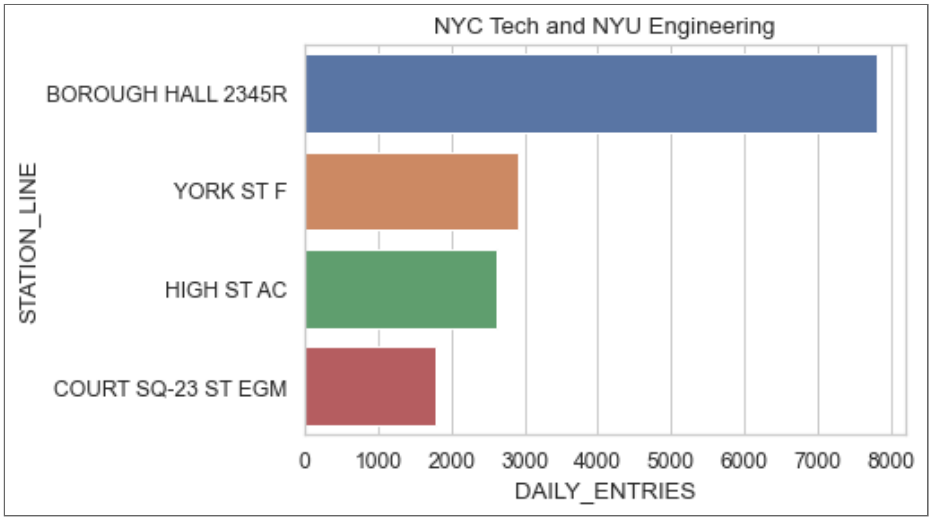

CUNY Hunters and NYU Nursing
14 ST-UNION SQ 456LNQRW
1 AV L
28 ST 6
STATION_LINE
0
10000
20000
DAILY_ENTRIES

Results Cont




NYU Sciences
W 4 ST-WASH SQ ABCDEFM
ASTOR PL 6
8 ST-NYU NRW
STATION_LINE
0
5000
10000
DAILY_ENTRIES

Results Cont



Conclusion

- Would focus efforts during the middle of the week (Tuesday-Friday)
- Heaviest traffic for subway stops is centered around the early/late evening
- For the most part there is a clear heavier use subway stop around the various campus. Would try at those stops first.

Future Work

- With a little more given time we could further isolate the most heavily trafficked subway stations in close proximity to university campuses.
- Could incorporate more geo location data to make sure the WTWY representatives are most efficient in collecting support.

In [ ]:

```
get_station_traffic()
```