

Knicks Home Games and their effect on MTA foot traffic

An exploratory data analysis on the effects of large social gathering in the post-COVID era

By Will Carnevale

OBJECTIVES AND GOALS

Objective: *Analyze the effect of New York Knicks home games on MTA foot traffic in stations proximate to Madison Square Garden (MSG)*

Goals:

1. *Give the city of New York with a better grasp on how large social gatherings might affect public transit congestion in the post-COVID era*
2. *Provide better data to be used as a basis for predictive models of viral outbreaks in New York*

TOOLS AND DATA

Data:

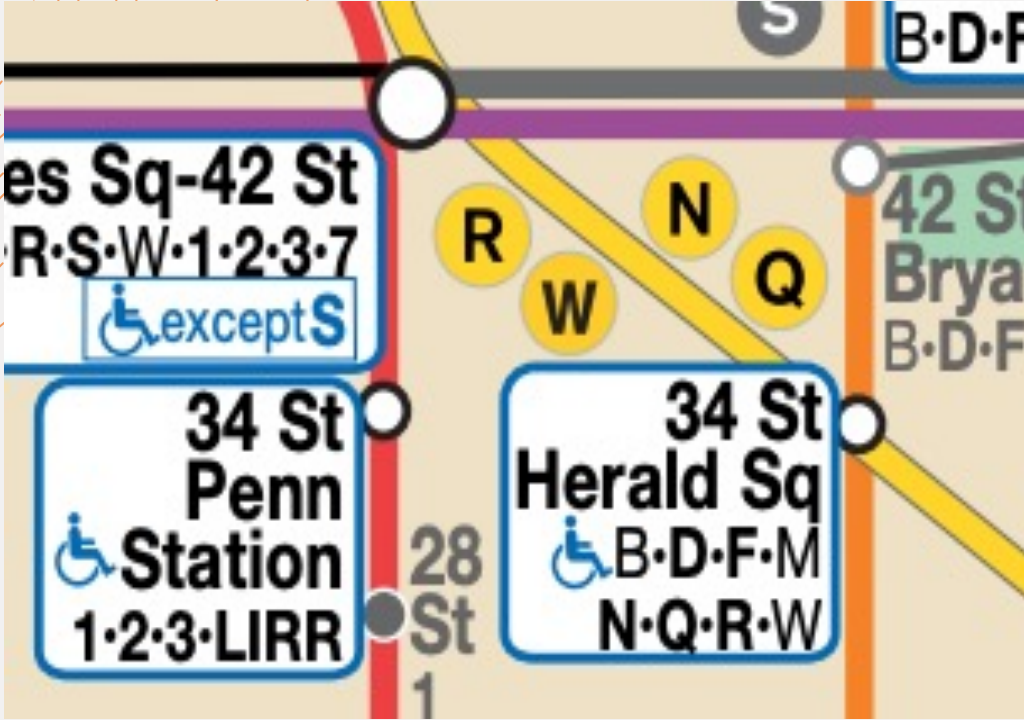
1. MTA Data; publicly accessible and provides records of cumulative entries and exits for every unique turnstile.
2. Basketball Reference Data; provides data on all the games played by the knicks in the 2018-2019 season.

Tools:

1. SQLAlchemy for data queries
2. Pandas for data manipulation and web-scraping
3. Matplotlib for plotting data

METHODS

Three Stations:

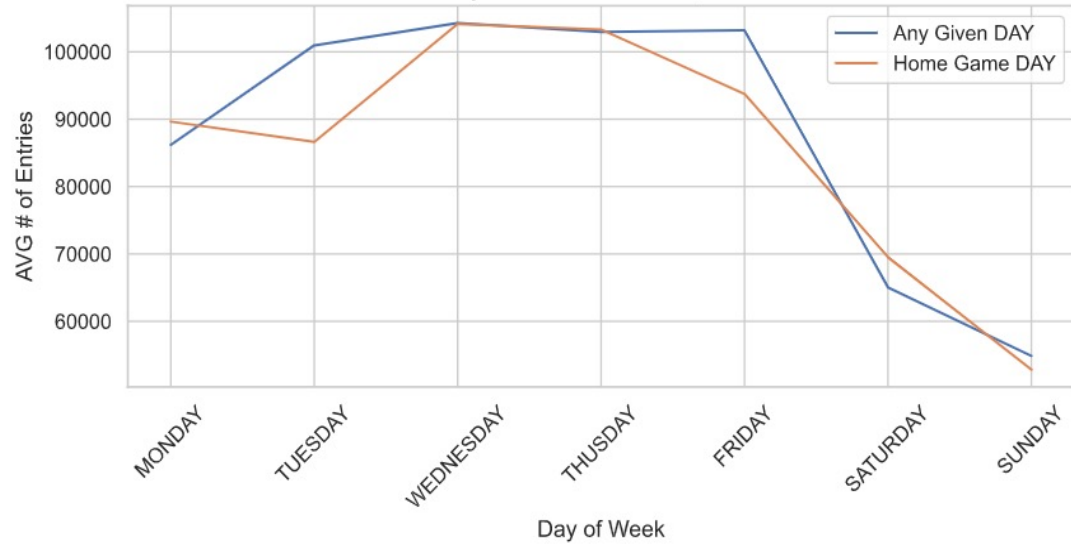


Algorithms:

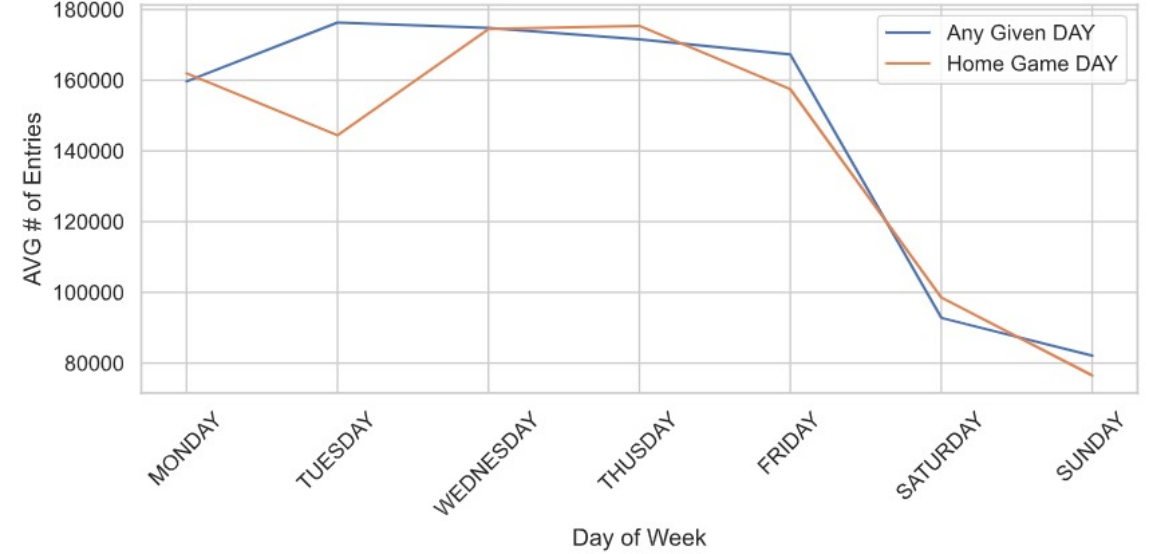
1. Clean data sets and join on a common column (created) in order to match daily entries to days WITH knicks home games
2. Compute daily entries averages grouped by day of the week
3. Use .query and .drop on a left merge/join in order match daily entries with days WITHOUT knicks home games, then repeat step 2

RESULTS

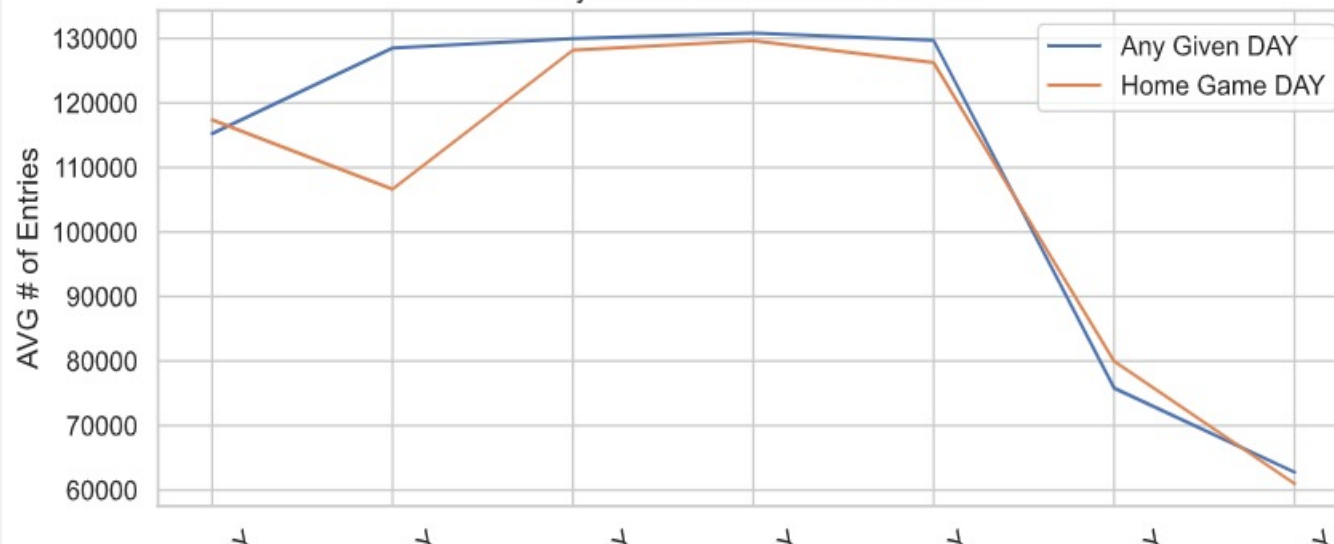
Daily Entries for TIMES SQ-42 ST



Daily Entries for 34 ST-PENN STA

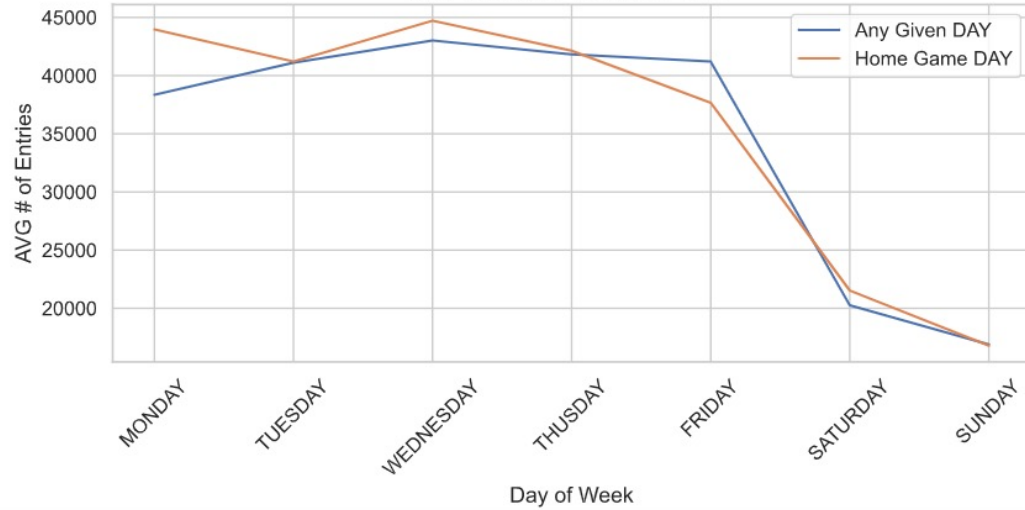


Daily Entries for 34 ST-HERALD SQ

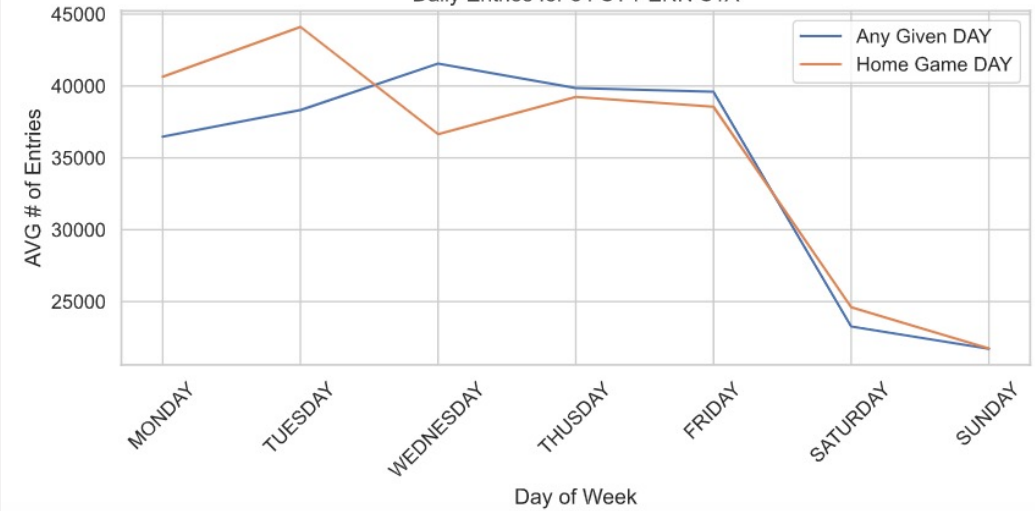


RESULTS PART 2

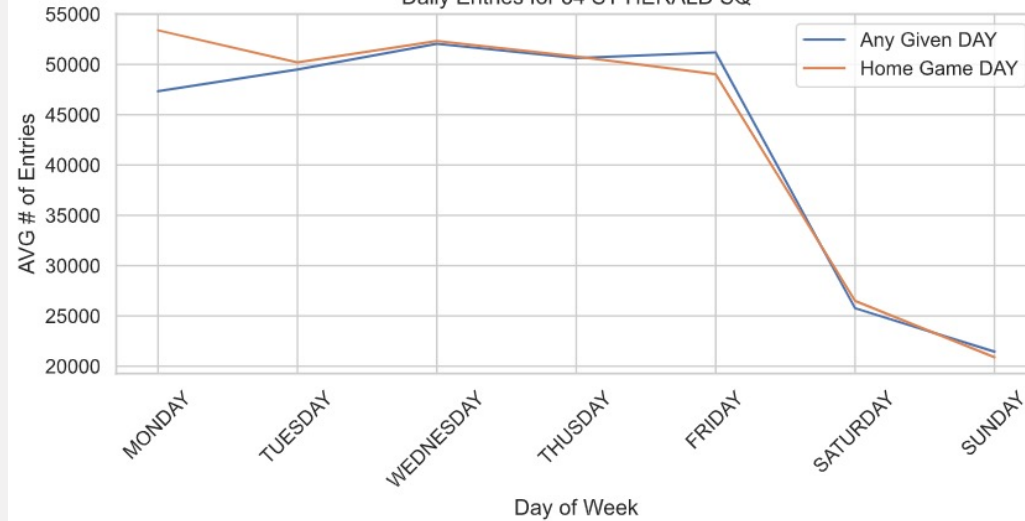
Daily Entries for TIMES SQ-42 ST



Daily Entries for 34 ST-PENN STA



Daily Entries for 34 ST-HERALD SQ



Ex:

```
In [29]: sum_counts_nba3 = merged_data_TIMES1.DAY_OF_WEEK_NUM.value_counts().sort_index()
sum_counts_nba3
```

```
Out[29]: 0    6
         1    4
         2    7
         3    2
         4    7
         5    6
         6    9
         Name: DAY_OF_WEEK_NUM, dtype: int64
```

```
In [34]: sum_counts_not_nba1 = revised1.DAY_OF_WEEK_NUM.value_counts().sort_index()
sum_counts_not_nba1
```

```
Out[34]: 0    19
         1    22
         2    18
         3    24
         4    20
         5    20
         6    16
         Name: DAY_OF_WEEK_NUM, dtype: int64
```

CONCLUSIONS

- + Effect of Knicks home games is **inconclusive**
- + Possible sources of error: **low sample size**
- + **Future work** should address this issue; doing so would require analysis over multiple years with potential adjustments for inflation in MTA usage