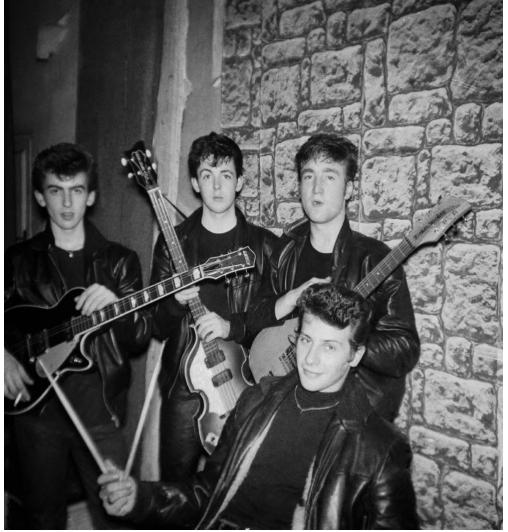
1960



1961

1962

So You Want to Be a Rockstar

How to Polish Your Brand & Increase Exposure in NYC Using MTA Data (beyond landing gigs)

10,000 Hours: The Story of Success

Motivation

- Find ‘stages’ for musicians’ exposure
- Live, extended performance experience
- Larger, more diverse audience
- Develop band personality
- Create cohesive musicianship

Objectives

- Choose 8 hour time periods (12pm-12am)
- Analyze which stations have the most traffic
- Identify the top 10 stations for each 8 hour time period

Goals

- Model benefits independent artists and small recording labels
  - expand audience
  - polish chops
  - develop people skills

Methodology

Data

Afternoon times

Spring/Summer/Fall (April-October)

2019 data

- Model is applicable to any MTA turnstile data

Metrics

Clean - eliminate corrupt data, duplicates, turnstile resets

Explore - min/max, negative values, kept data w/in +-2.7 std

Aggregate - turnstile ent/ex per 8 hr intervals, add turnstiles/station, sort by F/S/Su, avg F/S/Su

Visualize - plot to check anomalies, plot and heatmap to show results

Tools

SQL database

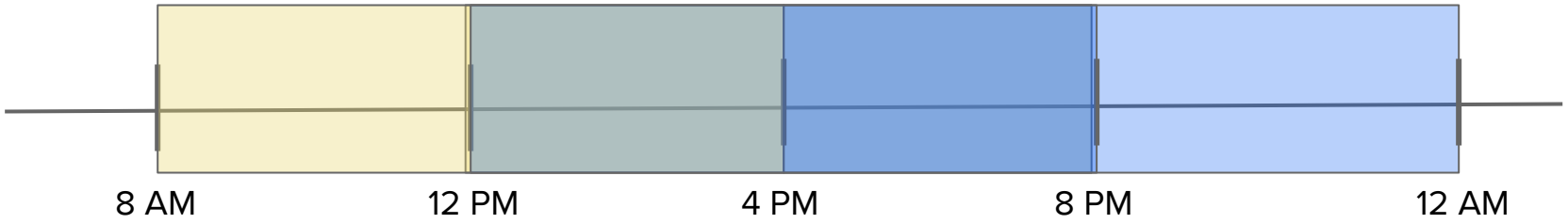
SQLAlchemy query from Python (WHERE, AND, NOT IN, ORDER BY)

Exploratory Data Analysis in pandas and numpy

Matplotlib/seaborn

Data Aggregation

It’s all relative - to the turnstile.

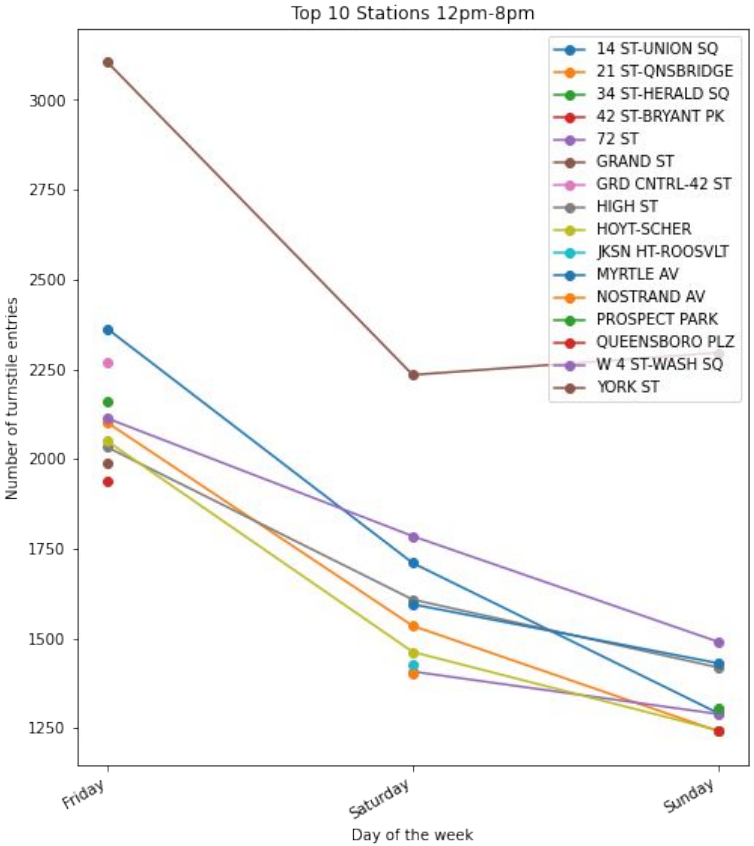


Early 8 hour period of entries/exits to the station

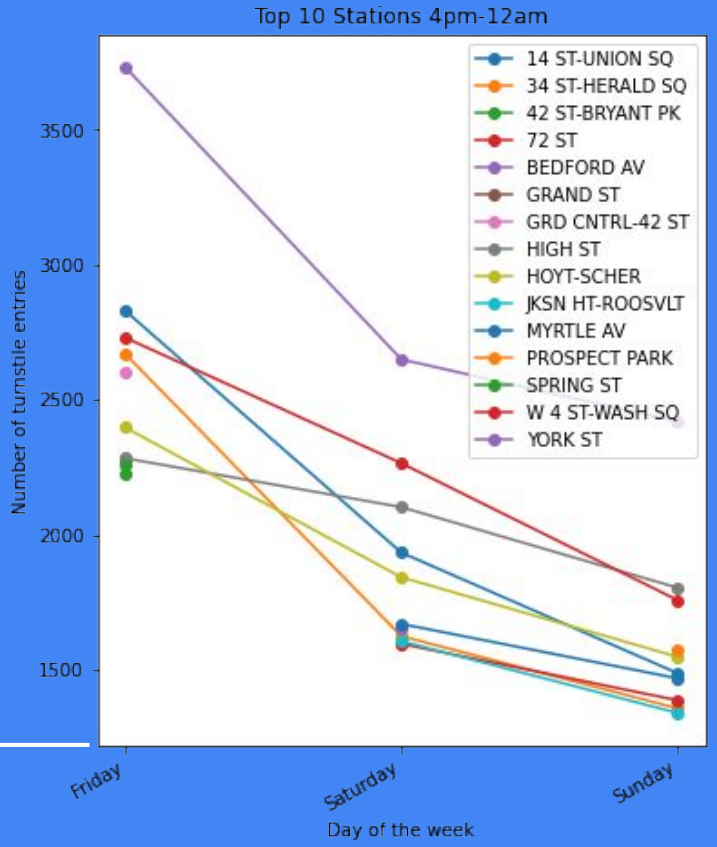
Late 8 hour period of entries/exits to the station

Results & Conclusions

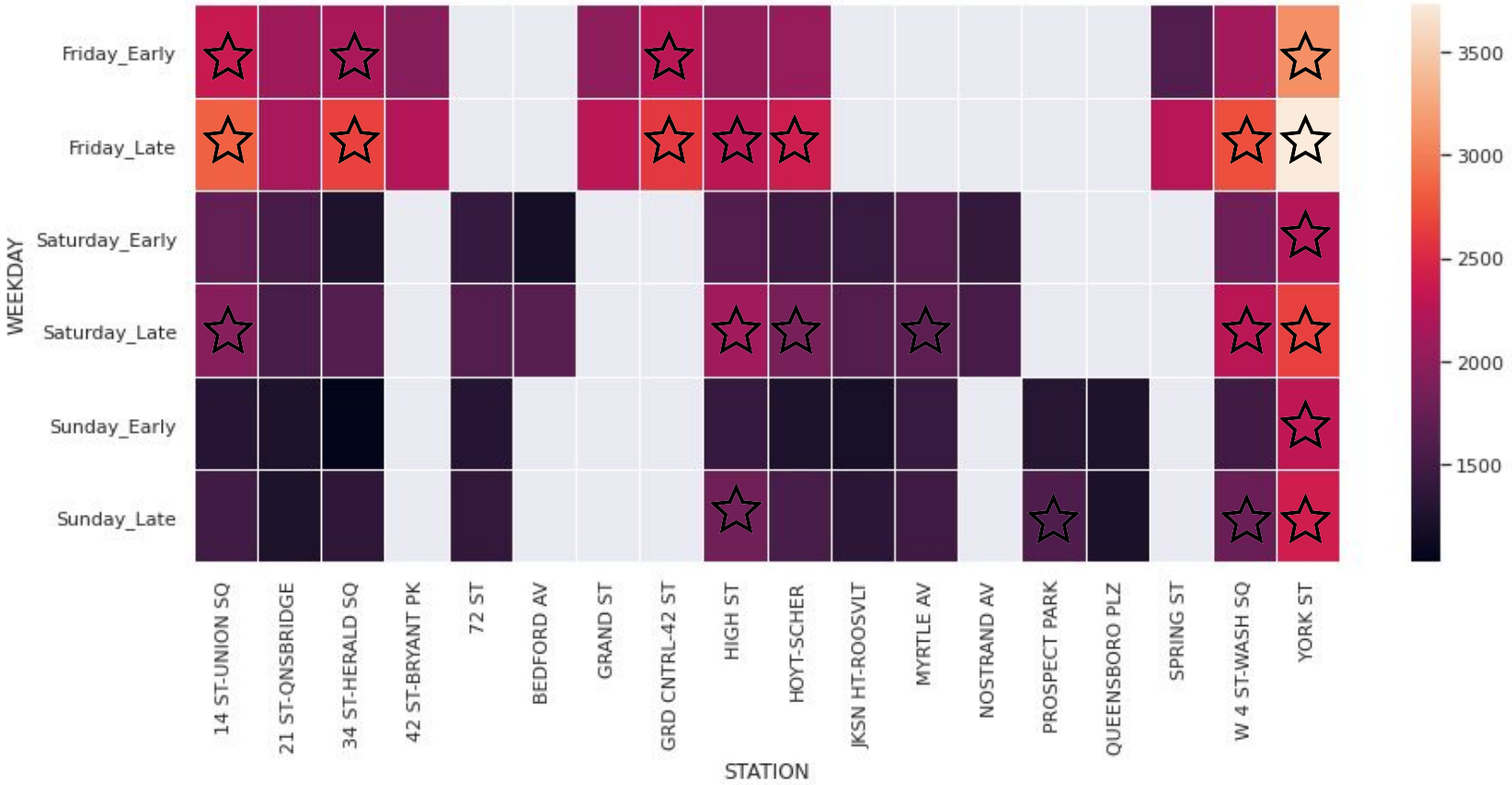
Early Period



Late Period



Heatmap of Top 10 Stations



Future Work

- Overlay top stations on a map of NYC
- Find popular stations near nightlife locations
- Match top stations to type of music based on various demographics
- Apply this model to other cities' public transit systems