Yoga Today

EDA Metis June/July 2022
Linda Ngo

Introduction

- Background Story
    - Yoga Today was an established yoga studio that had closed in 2020 due to the pandemic
    - Now that the Pandemic is over, the team is believes that now is a good time to reopen their studio
- Request asked
    - Requesting our expertise to find ideal areas for a new Yoga Studio

Methodology

- Tools that were used
  - SQL
    - DB Browser
    - SQL Alchemy/SQLite
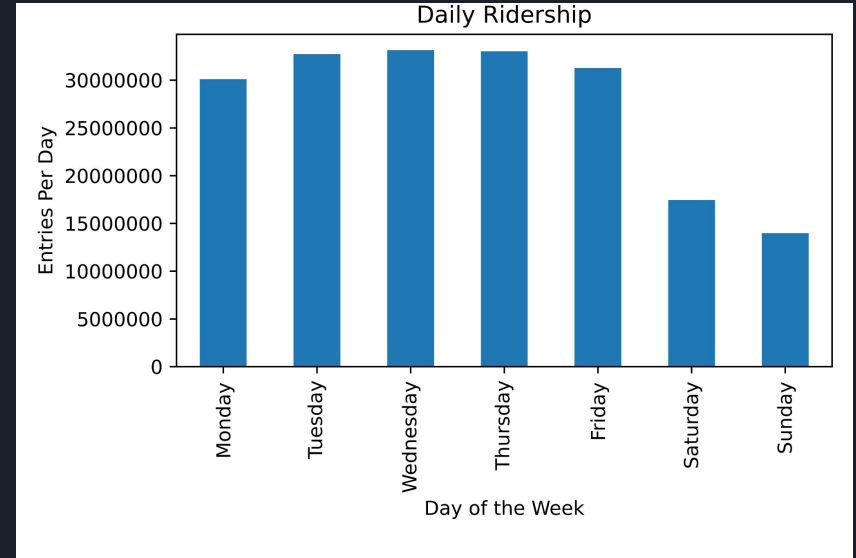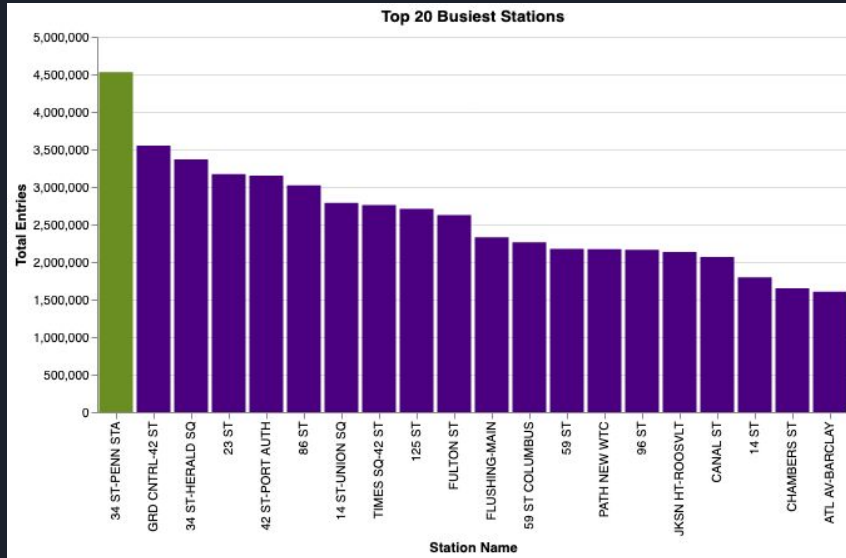 - Pandas
  - Numpy
  - Matplotib
  - Altair
  - Folium

```python
%matplotlib inline
import matplotlib.pyplot as plt
import numpy as np
import datetime
import pandas as pd
import altair as alt
import folium
from sqlalchemy import create_engine
from vega_datasets import data
from sqlite3 import connect
# con = connect('mta_data.db')
engine = create_engine('sqlite:////Users/lindango/(
con = engine.connect()
```

Results

- Top 20 Busiest Stations
- Total Riders by the days of the week

```
indexes = ['Monday','Tuesday', 'Wednesday','Thursday','Friday','Saturday','Sunday']
dotw_entries.index = indexes
dotw_entries
```
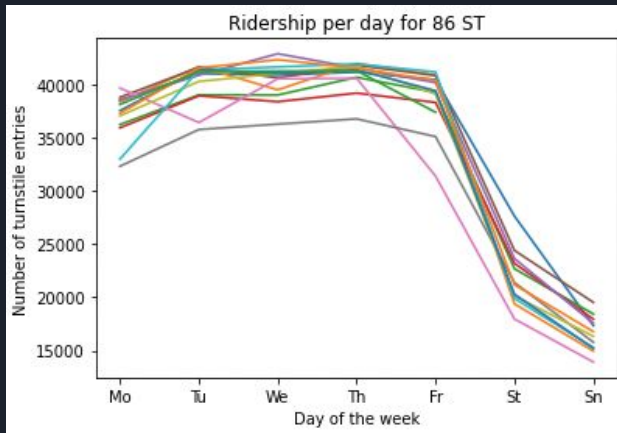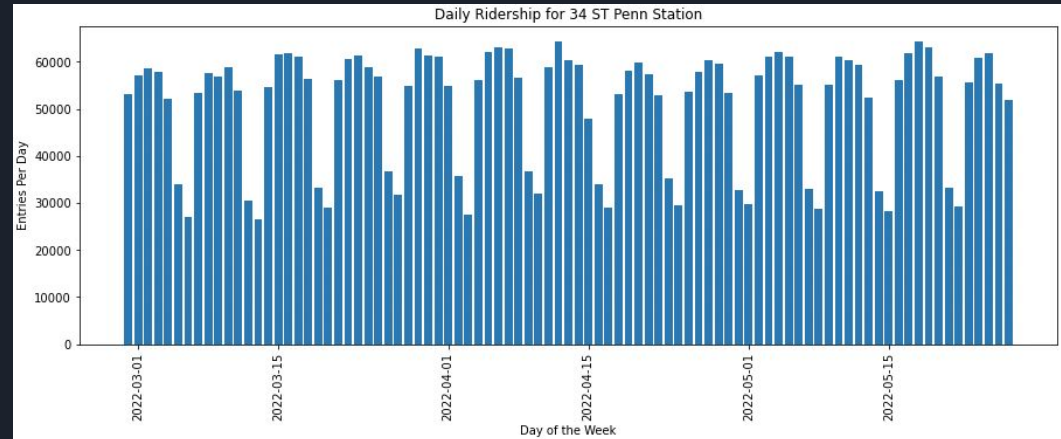
```
Monday        30105079.0
Tuesday       32720029.0
Wednesday     33149756.0
Thursday      33024618.0
Friday        31277725.0
Saturday      17453268.0
Sunday        13971928.0
Name: DAILY_ENTRIES, dtype: float64
```
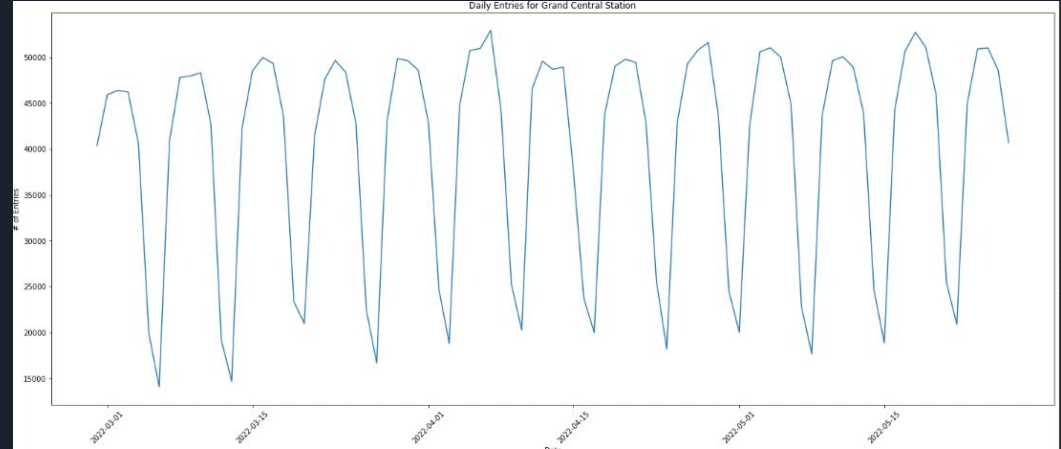
Results

- Anomalies:
  - counter sometimes counted the wrong way
  - Data for counters was not gathered at the same time
  - Some stations had traffic that was close to the weekend but but most stations had a drop in activity for the weekends - tourism?

Daily Entries for Grand Central Station

Ridership per day for 86 ST

Daily Ridership for 34 ST Penn Station

Conclusion

- Recommendations
  - Focus on the Top 10 Stations
  - Focus on Weekday Traffic
  - NYPD YTD Arrest Data
    - High crime in certain areas

Future Work

- Things to further explore
  - Using the 311 data to see the crime in the area
  - Use the Census data to see the demographics for gender and age group
  - Use Real Estate data to gauge the cost of setting up a new studio in the neighborhoods
  - Determine peak hours to better understand the flow of traffic
  - Take income data into consideration

Thank you!

Questions?