
EDA Project

Gideon Metrikin

June 2022

Introduction

Background

- > Food truck business that wants to maximize sales by parking their trucks in the busiest locations
- > Specialize in offering breakfast food and are a popular choice for commuters looking for something to eat before heading into the office for the day or during a morning break
- > Owns 10 food trucks that open at 8am and close at 2pm

Objective

- > Use publicly available MTA turnstile data to determine optimal locations for parking food trucks

Methodology

Data

- > MTA turnstile data from February 2022 to April 2022

Metrics

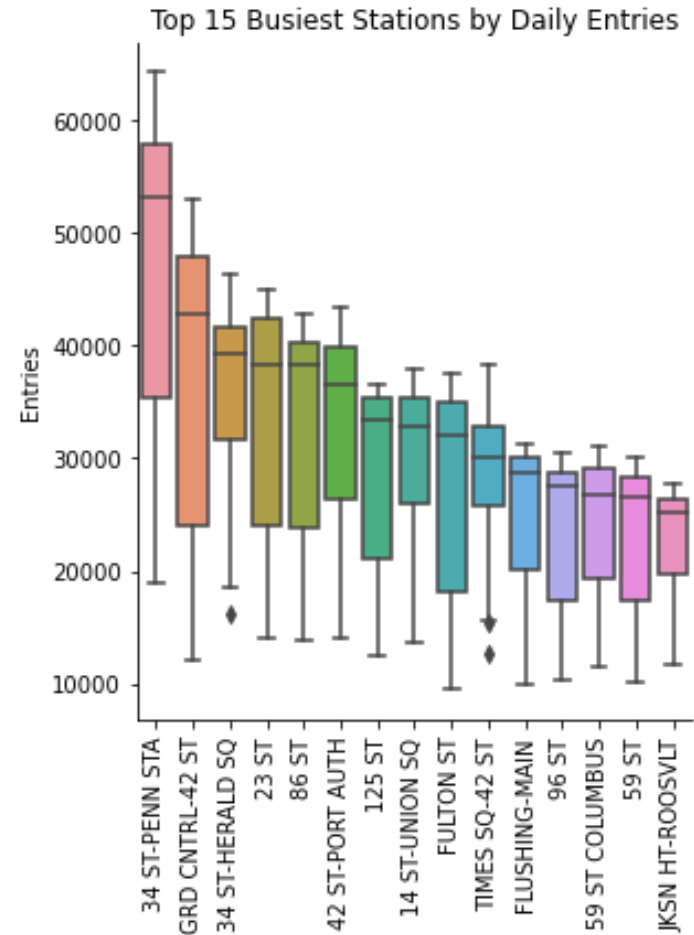
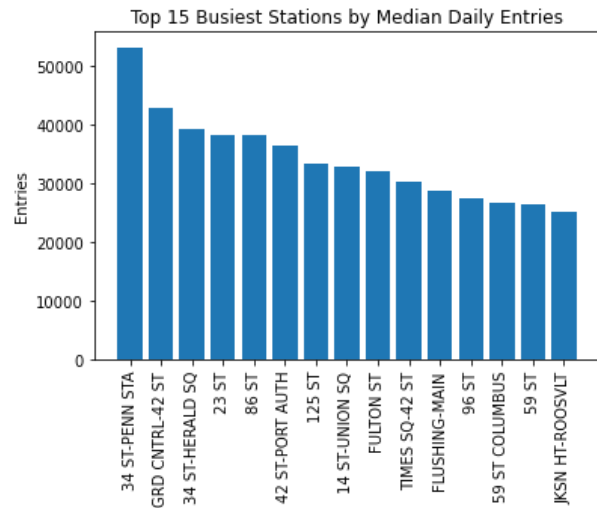
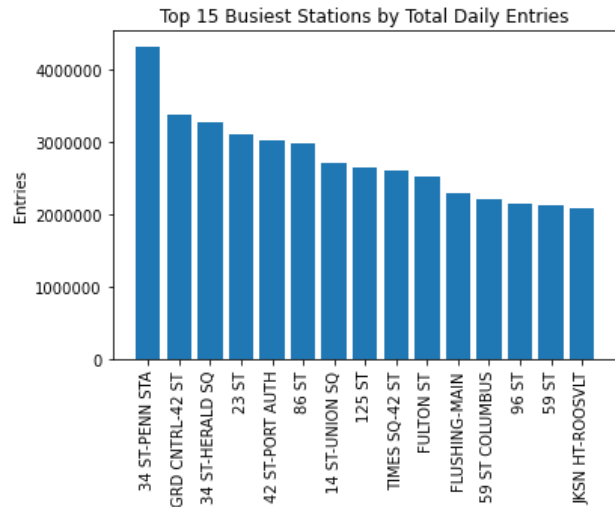
- > For each station:
 - Total daily entries/exits
 - Total morning (9am – 1pm) entries/exits
 - Median daily entries/exits
 - Median morning (9am – 1pm) entries/exits

Tools

- > SQLAlchemy for querying data from database into Python
- > NumPy and Pandas for data manipulation
- > Matplotlib and Seaborn for plotting

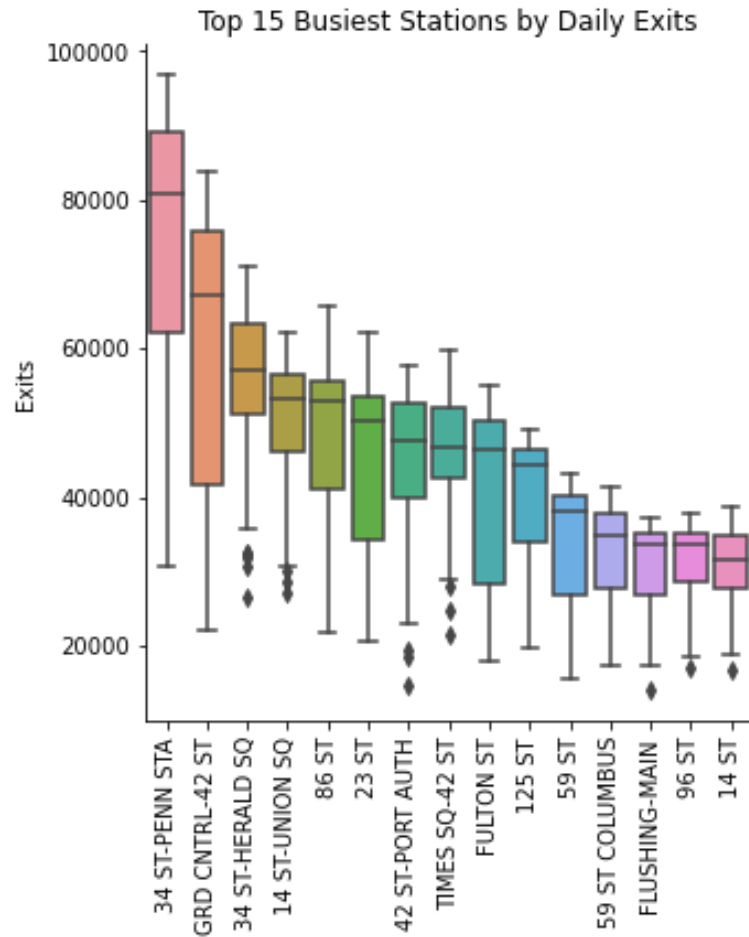
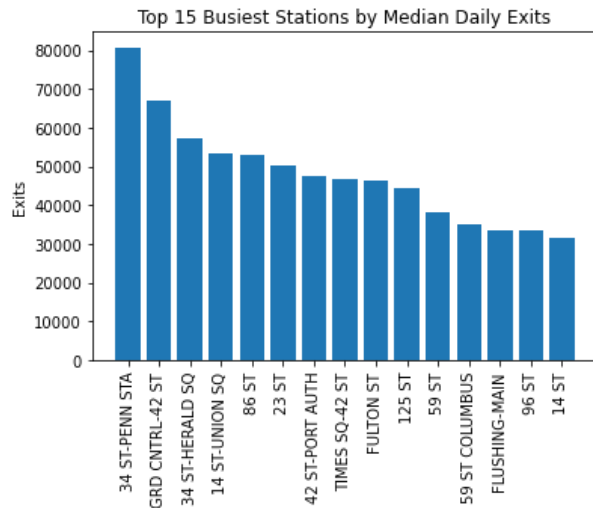
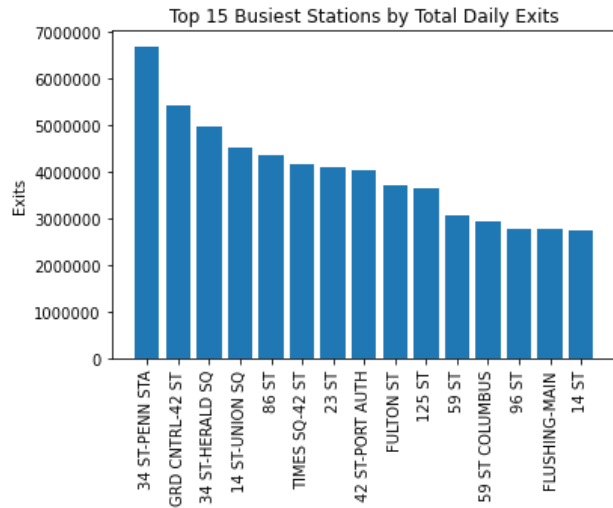
Results

Daily Entries (Feb 2022 – Apr 2022)



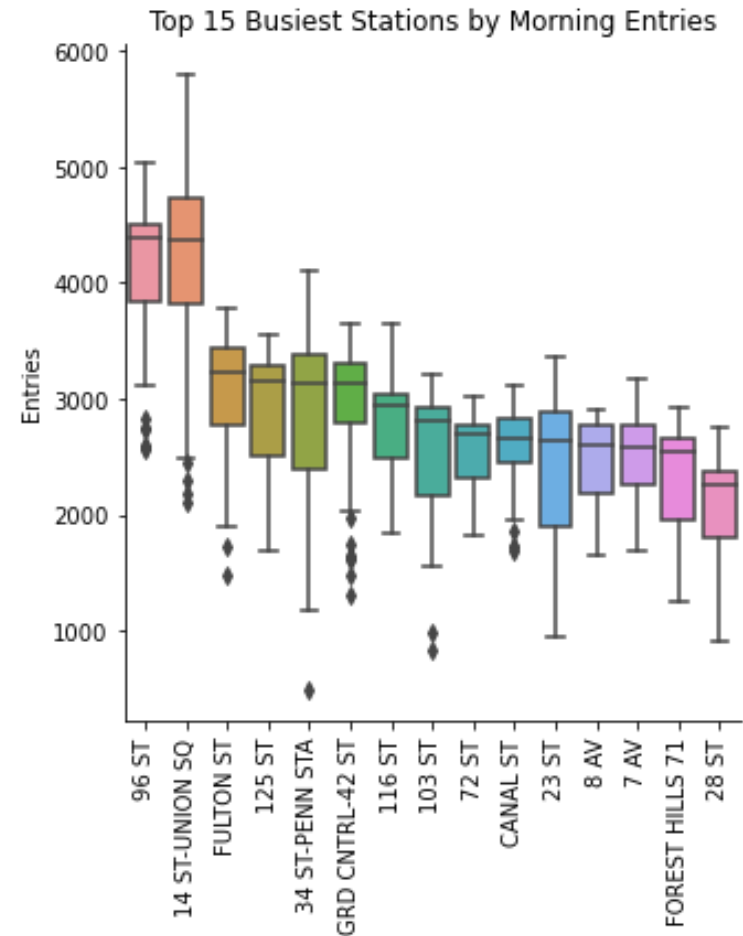
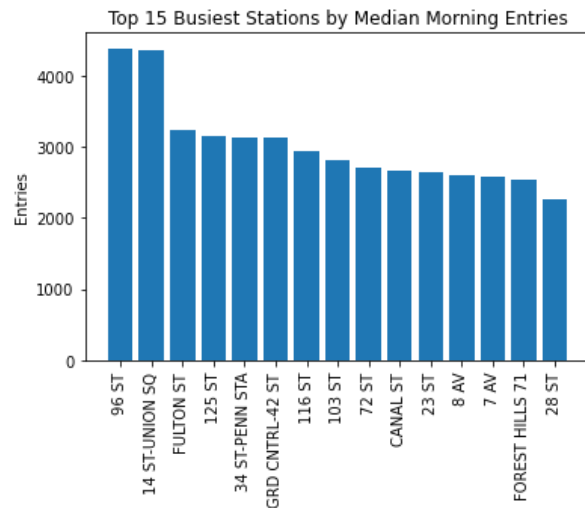
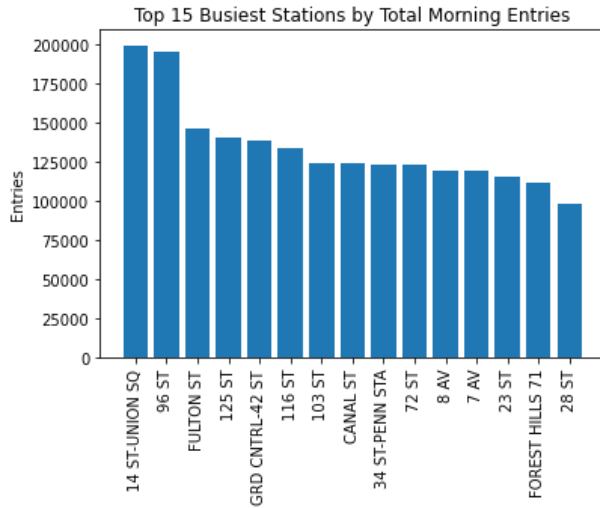
Results

Daily Exits (Feb 2022 – Apr 2022)



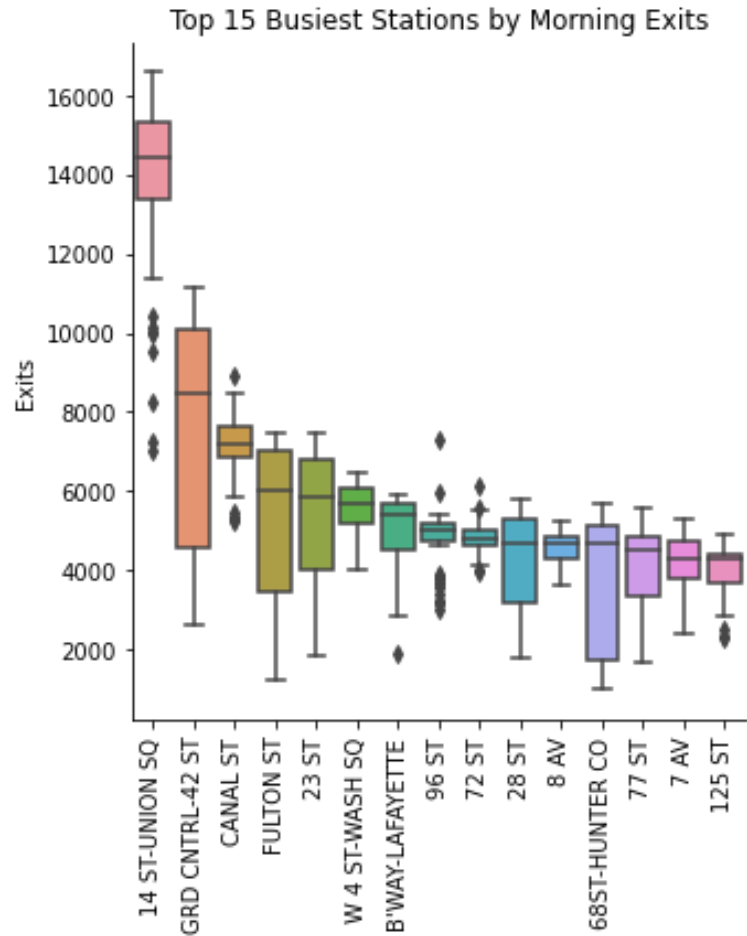
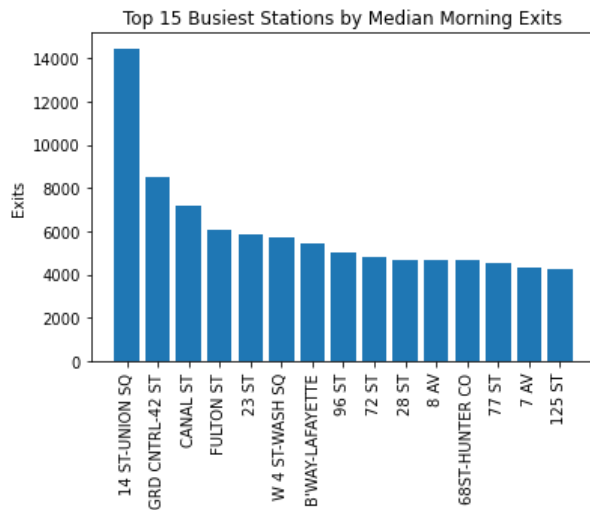
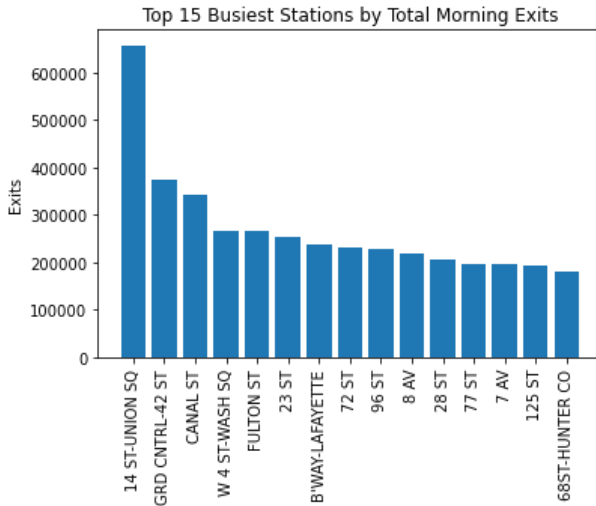
Results

Entries from 9am to 1pm (Feb 2022 – Apr 2022)



Results

Exits from 9am to 1pm (Feb 2022 – Apr 2022)



Conclusions and Further Analysis

Conclusions

- > Better to use top locations by exits instead of by entries
 - Higher daily and morning numbers
 - People more likely to buy breakfast after exiting subway

Further Analysis

- > Incorporate price of food truck parking permits into analysis to determine which locations have the best bang-for-buck
- > Look into busiest locations by days of the week to see if there are any differences