



MTA EXPLORATORY DATA ANALYSIS

Yingqing Qiu
Metis Data Science Bootcamp
September 2021

Introduction

- **Motivation**

NYC DOT and MTA want to add/construct several new stations to satisfy the growing post-pandemic daily public transit needs.

- **Objectives and goals**

We plan to send out several groups of surveyors to collect information from the passengers entering the subway stations. So that it is necessary to come out a reasonable and efficient survey plan to save time and labor. We want to find out the busiest station and time blocks for the preliminary survey plan.

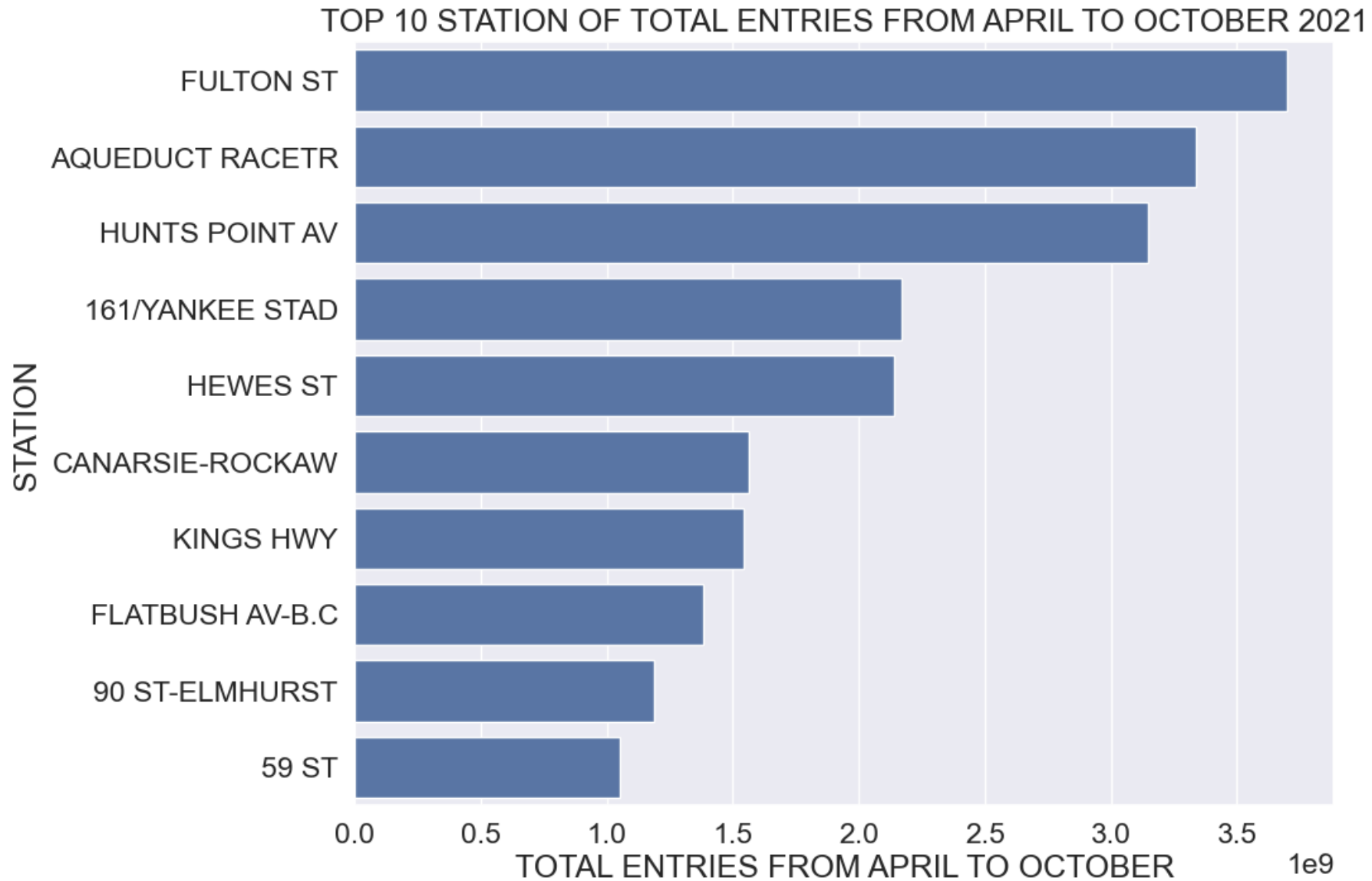
- **Final results**

After the data analysis, we selected several stations for the survey plan. Additional metrics that can be discussed in the future are considered and presented.

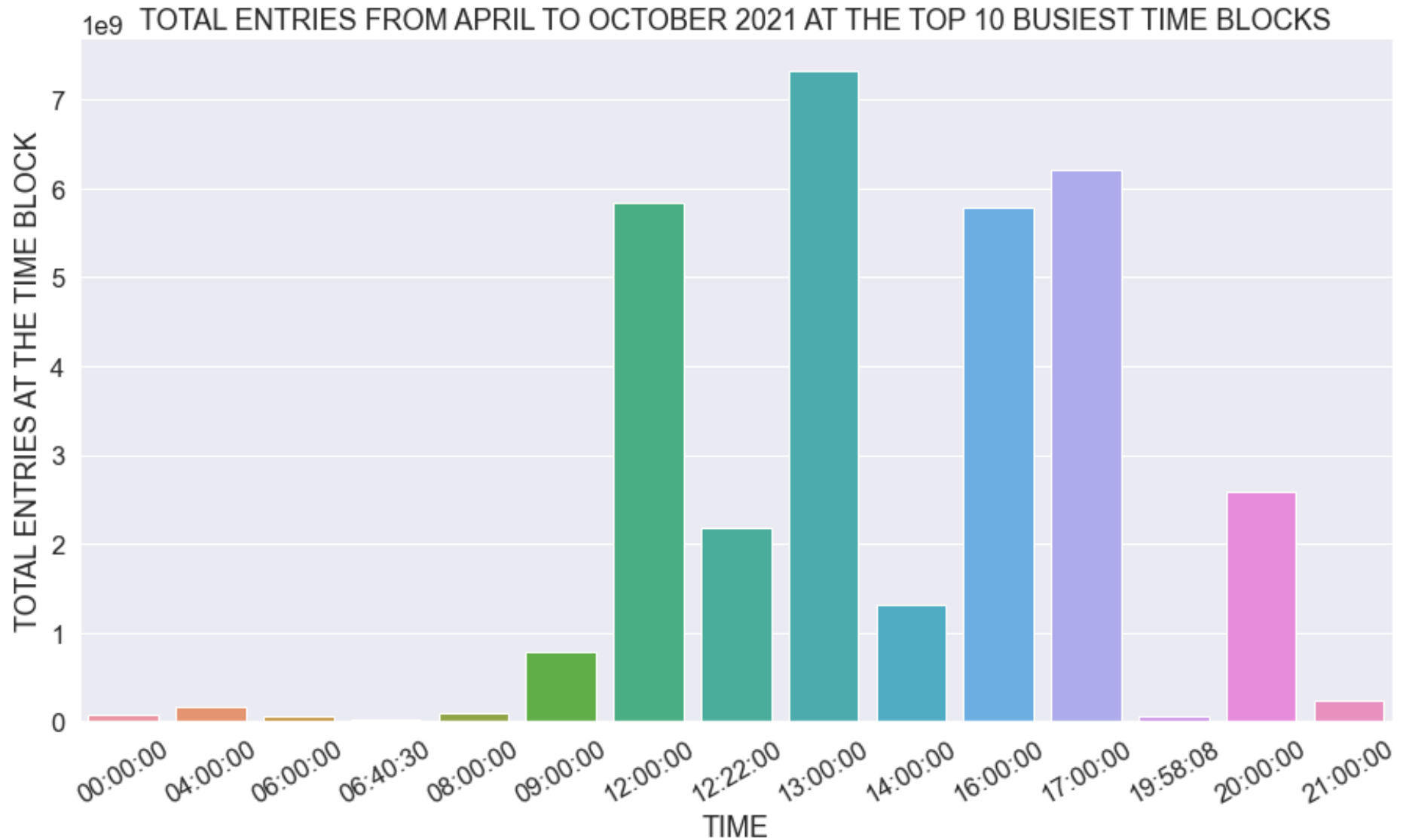
Methodology

- **Pandas**
- **Numpy**
- **Matplotlib**
- **Seaborn**

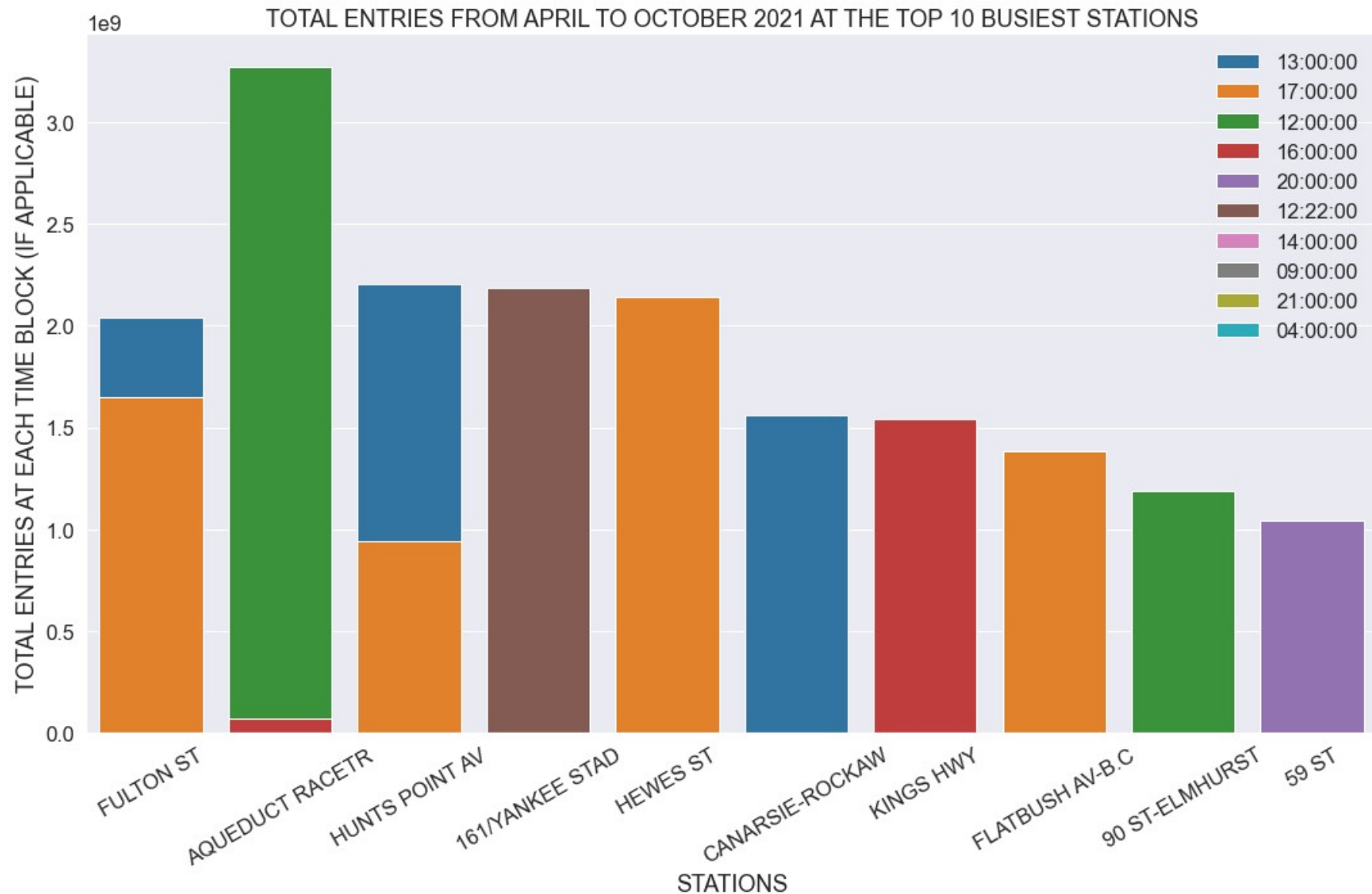
Results



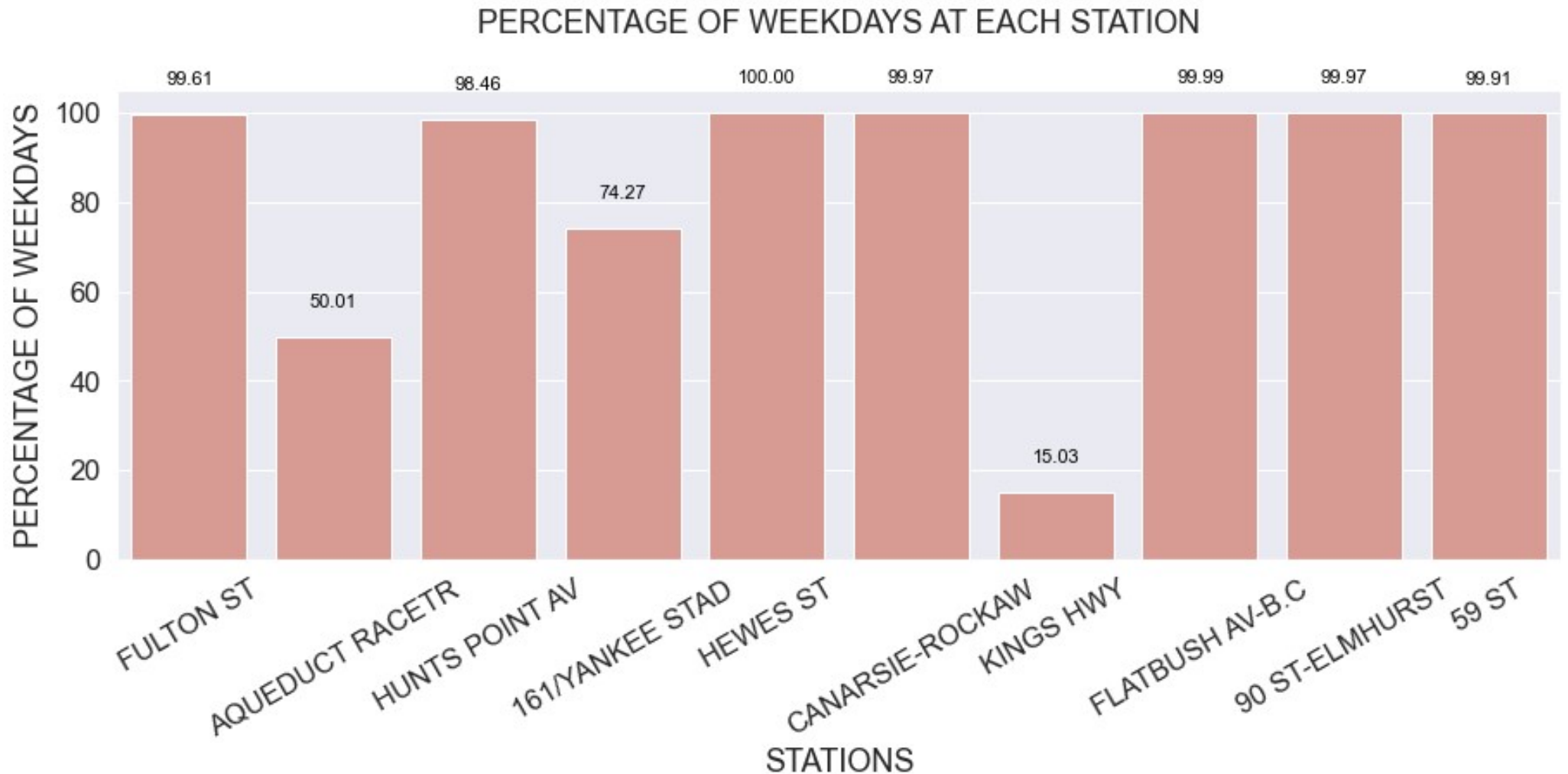
Results



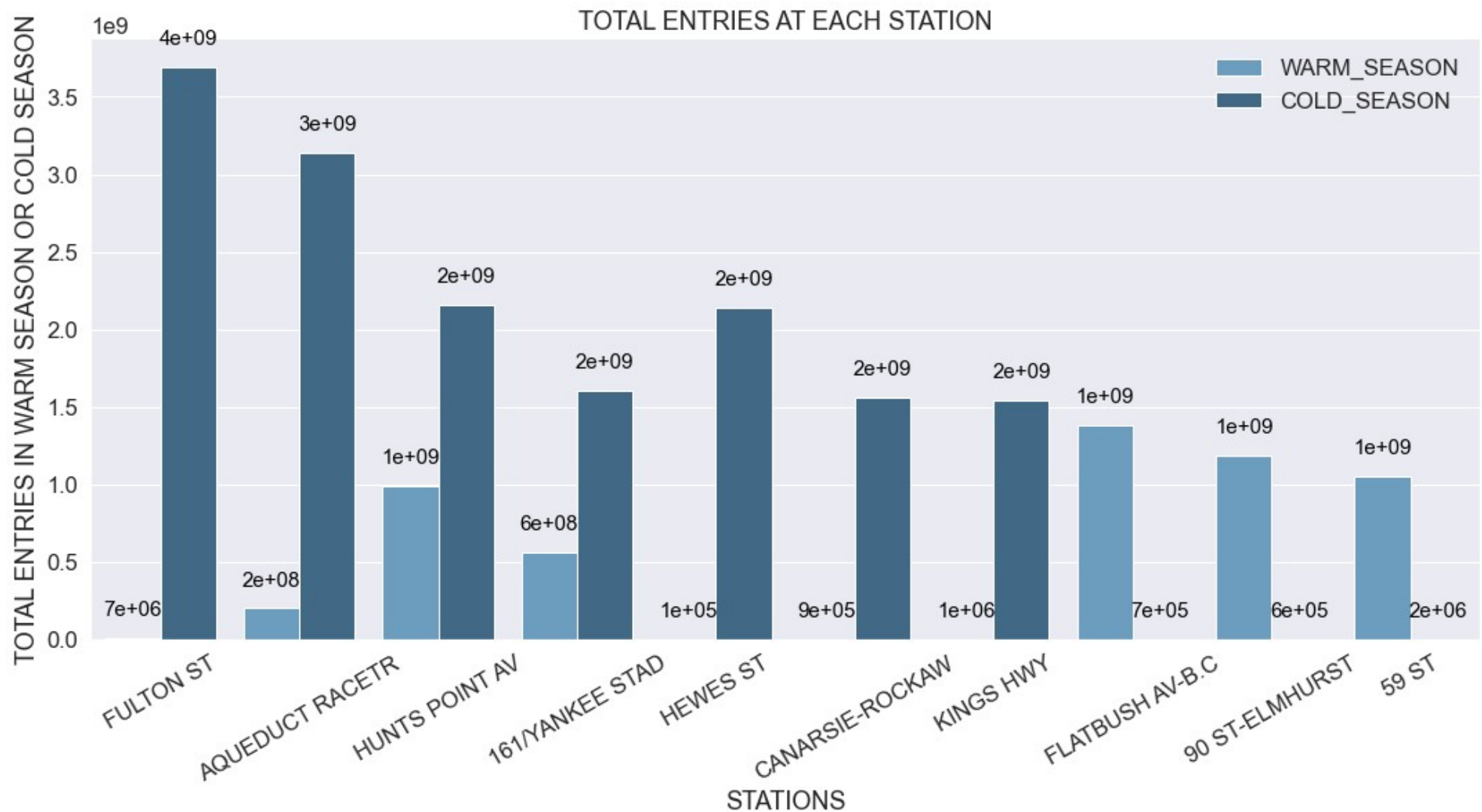
Results



Results



Results



Conclusions

PRELIMINARY SURVEY PLAN – TIME AND LOCATION

Station	Time	Weekday/ Weekends	Season
FULTON ST	17:00:00	Weekday	Cold
AQUEDUCT RACETR	12:00:00	Both Weekday/W eekends	Cold
HUNTS POIN AV	13:00:00	Weekday	Cold/Warm
... ..			

Future Work

- **Distance between stations**
- **Correlate the census characteristics near each station with the daily entries**
- **... ..**



Thank you!