# EDA MTA Project

Stanley Guo

# Introduction

We have been tasked to optimize engagement and awareness for WomenTechWomenYes (WTWY) and their upcoming gala by using MTA subway data to help them strategically place street teams around the city.

The goal is to conclude which stations are the busiest and on which days to optimize street team placement for maximum campaign outreach.

# Assumptions
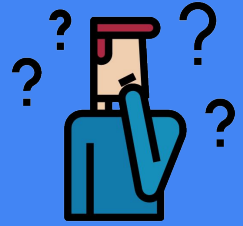
1. Focus on total exits per station

- Allows for data collection on travellers from surrounding areas - greater outreach than entries

2. Busier stations would ideally provide more engagement

- More people = more emails collected, more diverse demographic

# Process
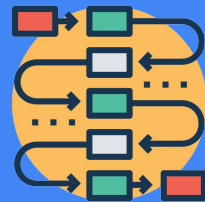


1. Create shell of the database using SQLite 3

2. Scrape data from MTA website using a Python script

3. Export database into Jupyter notebook and explore using SQLAlchemy

4. Clean and model data using Pandas

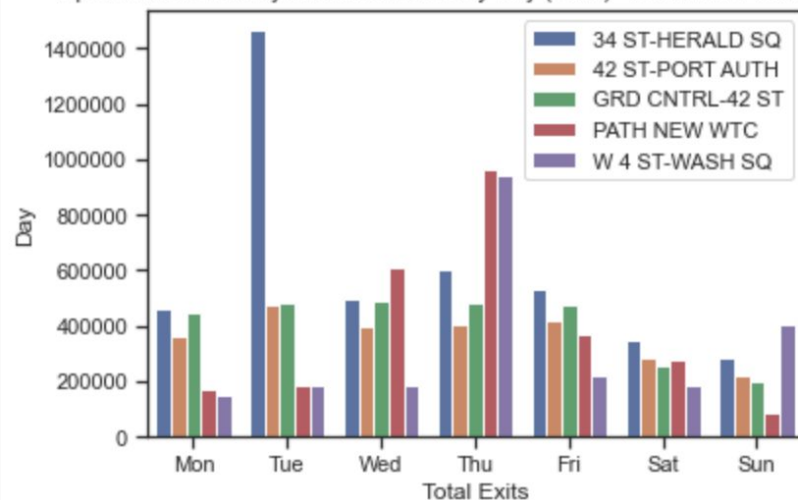5. Use Seaborn and Matplotlib for visualization

# Issues With The Data

Issue 1: Duplicate data, some rows contained a 'RECOVR AUD' value in its 'DESC' column
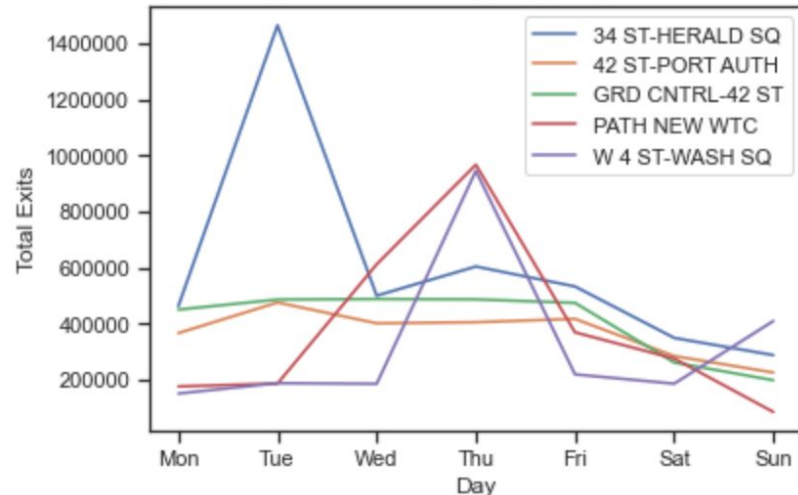
Issue 2: Some rows had negative counts - previous entries/exits were greater than current entries/exits
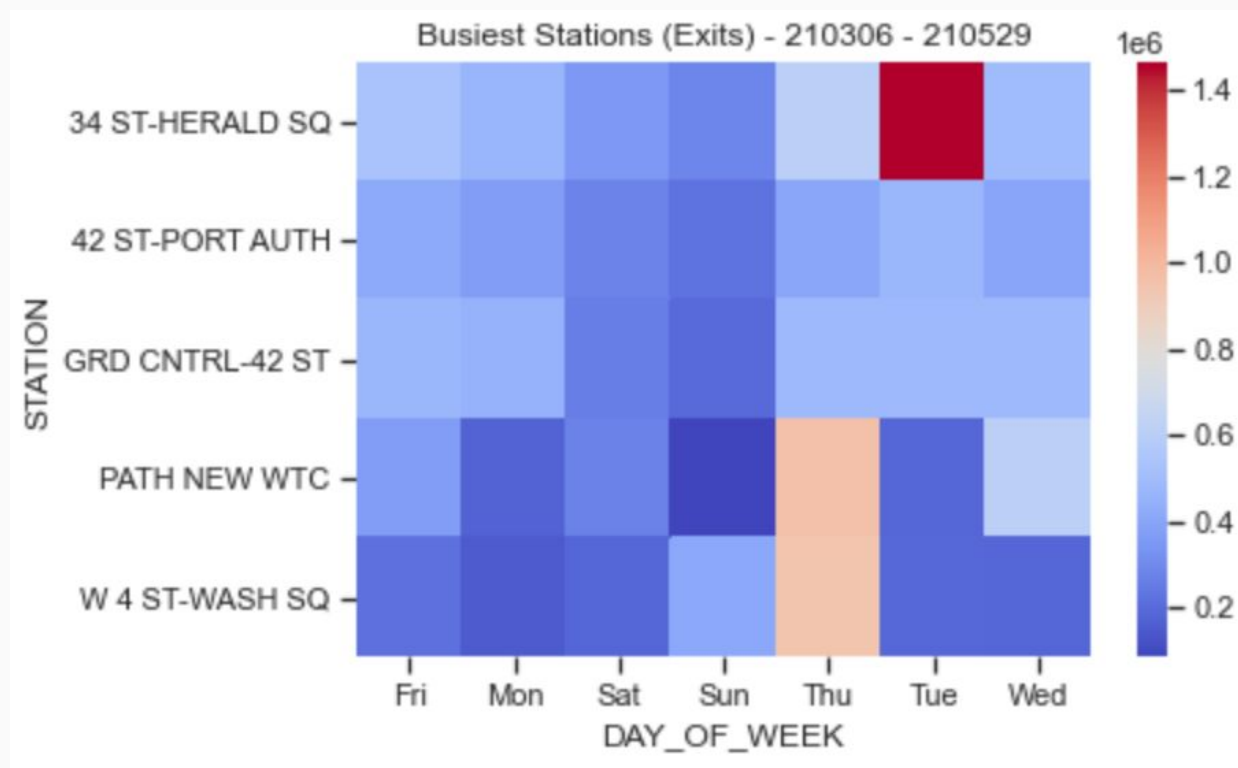
# Results



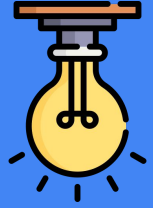Top 5 Busiest Subway Stations in NYC By Day (Exits) - 03/06/2021 to 05/29/2021

Top 5 Busiest Subway Stations in NYC By Day (Exits) - 03/06/2021 to 05/29/2021

# Results



Busiest Stations (Exits) - 210306 - 210529

# Conclusion

For the most effective street team efforts, WTWY should focus on..

34 ST-HERALD SQ and GRD CNTRL-42 ST on Mondays, Tuesdays and Fridays

GRD CNTRL-42 ST and PATH NEW WTC on Wednesdays,

W 4 ST-WASH SQ and PATH NEW WTC on Thursdays,

34 ST-HERALD SQ and PATH NEW WTC on Saturdays,

W 4 ST-WASH SQ and 34 ST-HERALD SQ on Sundays

Since 34 ST-HERALD SQ is extremely busy on Tuesdays, it might make sense to concentrate the majority or all efforts there on Tuesdays.
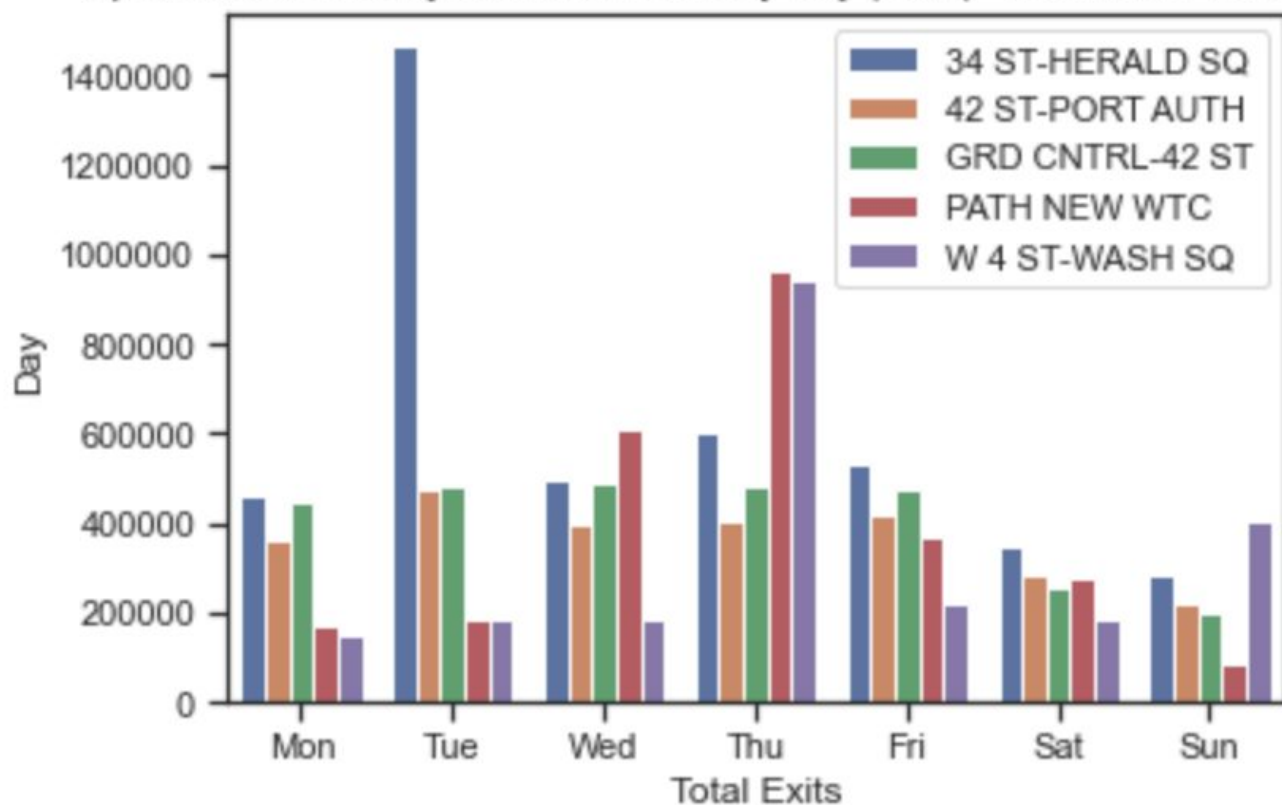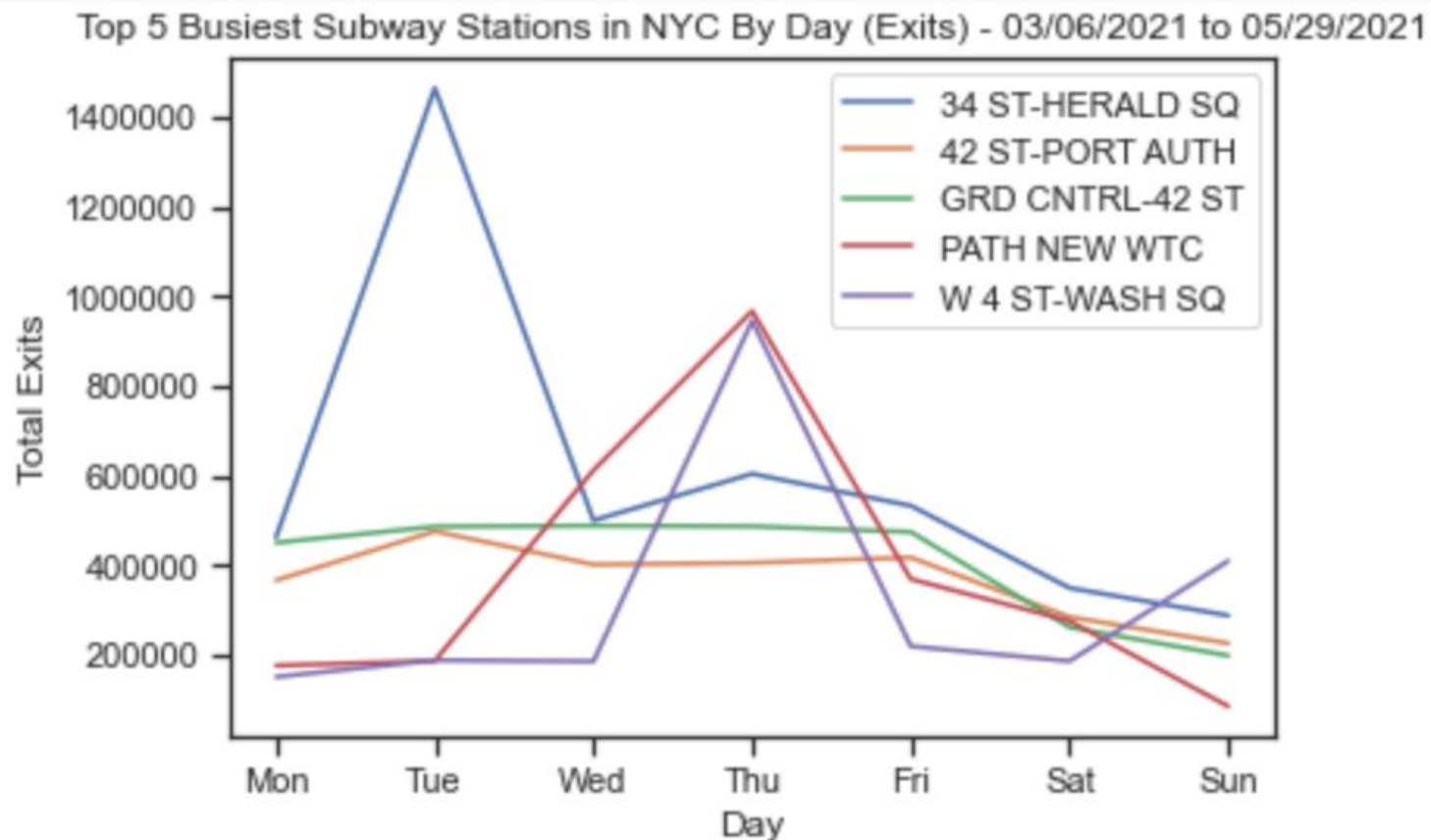
# Future Work

If there was more time, I would dissect the data to an even more granular level to not only split busy stations by day, but also by time of day. That will allow street teams to hit up multiple stations on the same day.

Top 5 Busiest Subway Stations in NYC By Day (Exits) - 03/06/2021 to 05/29/2021

Top 5 Busiest Subway Stations in NYC By Day (Exits) - 03/06/2021 to 05/29/2021

Busiest Stations (Exits) - 210306 - 210529