



Exploratory Analysis of MTA Data

Alex Stake



Introduction

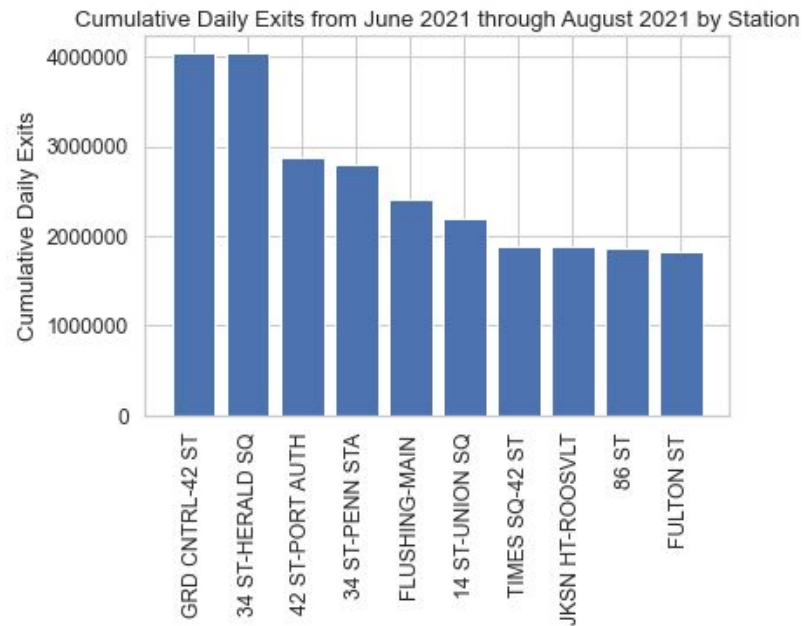
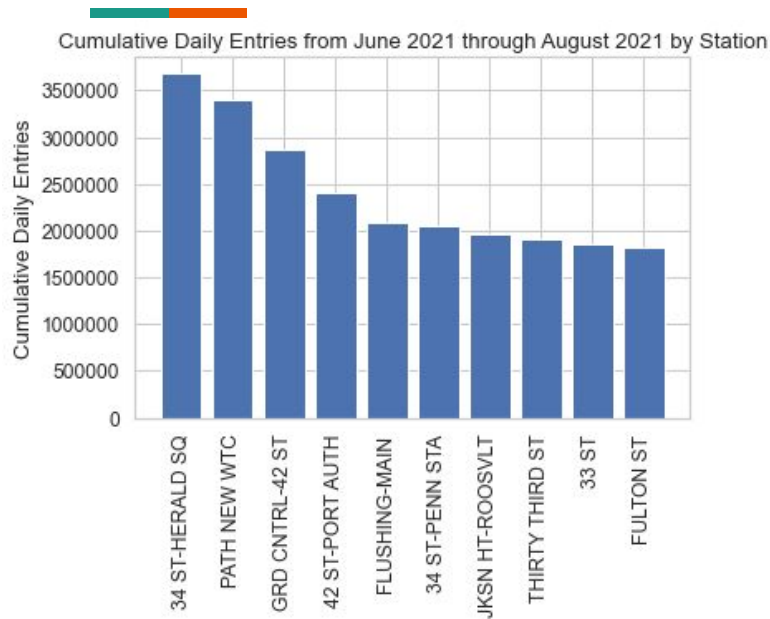
- Women Tech Women Yes (WTWY) want to increase gala attendance and membership
- Achievable through increased contact points at subway stations
- Higher traffic stations -> increased contact points -> increased attendance



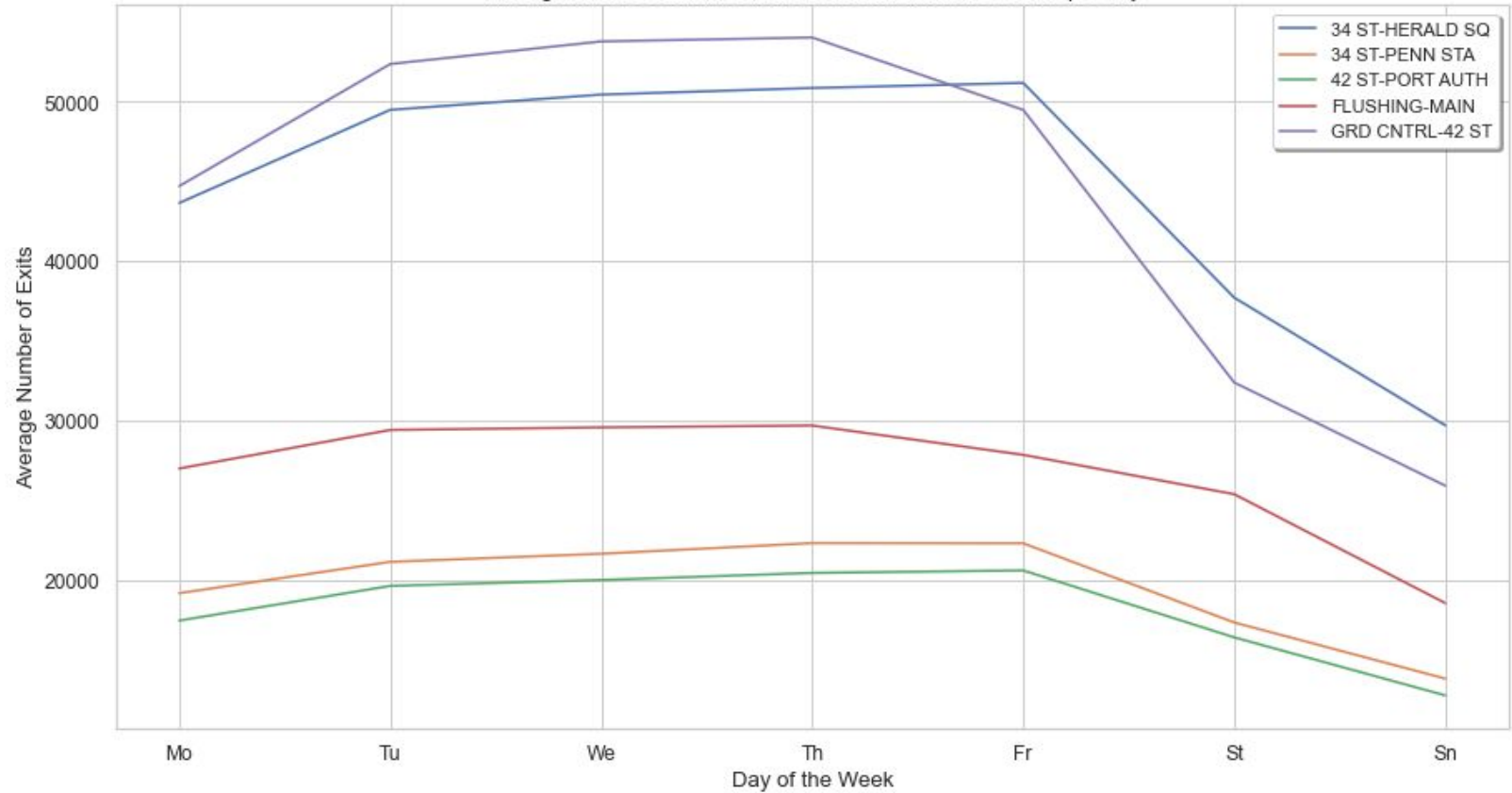
Methodology

- Data used:
 - MTA Turnstile data June 2021-August 2021
 - Google Maps
- Focus on exits
- Tools used:
 - SQLite, DB Browser, SQLAlchemy
 - Numpy and Pandas
 - Matplotlib and Seaborn

Initial Results



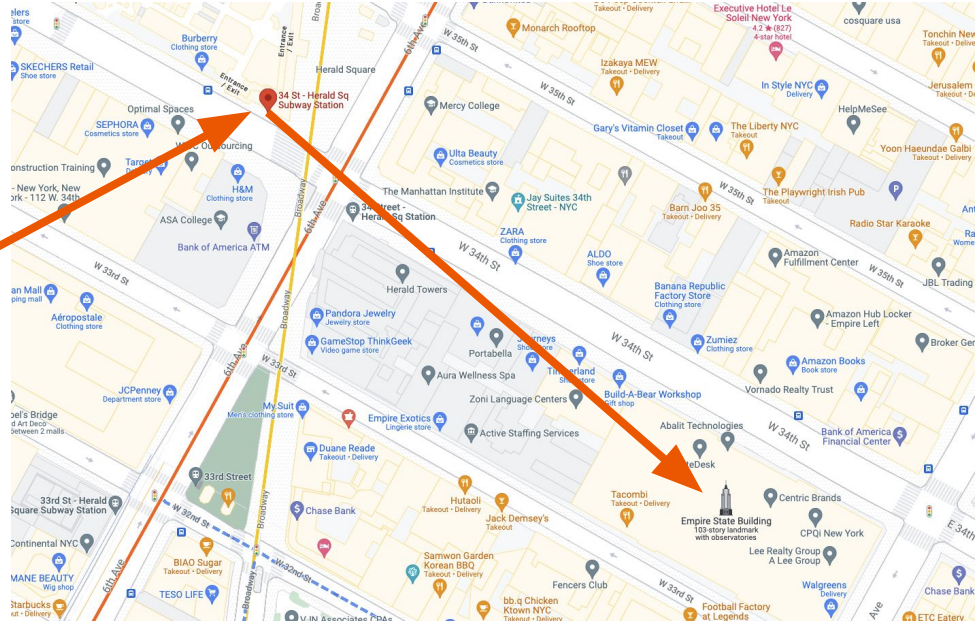
Average Number of Exits for The 5 Most Trafficked Stations per day



Conclusion

- Day of Most Traffic: Thursday
- Top 5 Stations:
 - 34th St and Herald Square
 - Grand Central Station at 42nd St
 - Flushing and Main
 - 34th St and Penn Station
 - 42nd St and Port Authority

Consider: Tourism Impact





Future Work

- More in-depth analysis of station by turnstile control area
- Use of a tourism database
- Incorporation of a vaccination database