Analysis for WTWY Street-team Deployment

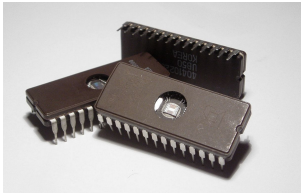
Team Turnstile Hoppers
Gretta Mae Ferguson, Shweta Pai,
Alex Bell, Steven Bierer

Problem Statement

- Optimize placement of **WTWY street teams** to advertise a Gala event in early summer.

- Goal is to **increase attendance** of the event and create **general awareness** about the organization.

Overall Approach



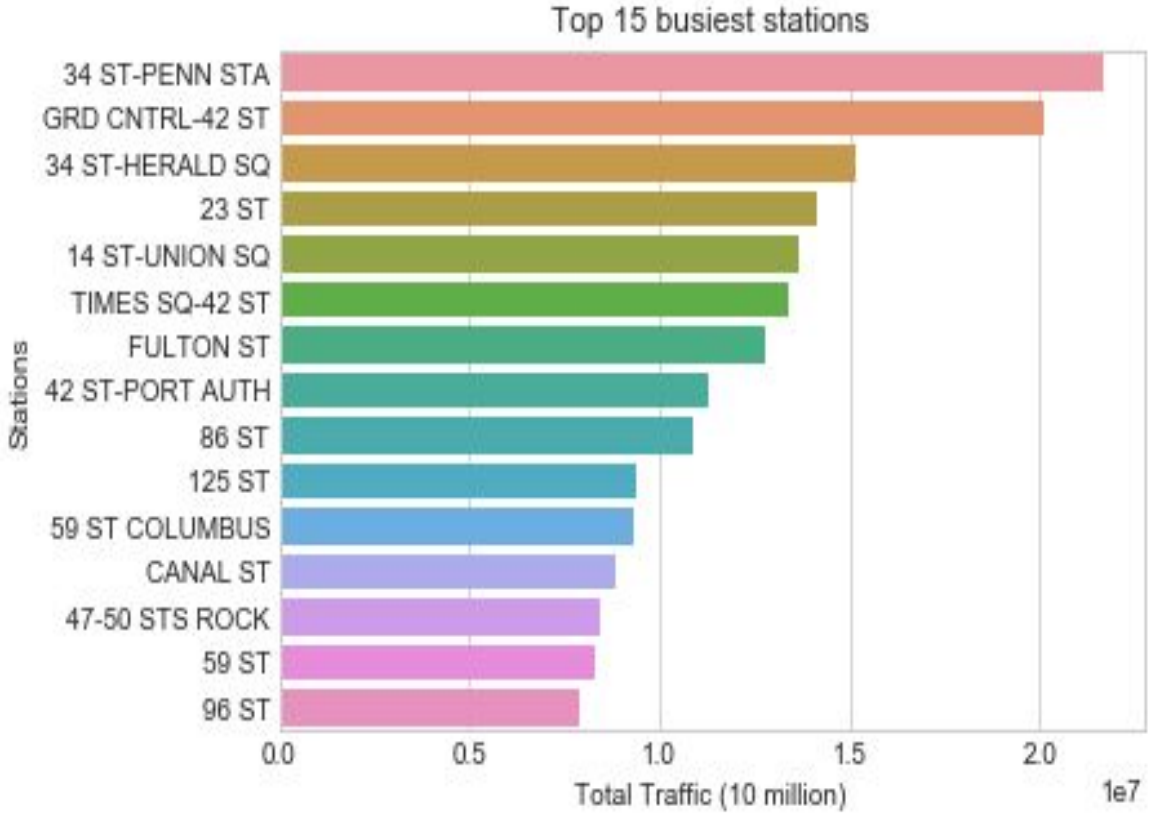
- MTA turnstile data as proxy for pedestrian traffic in NYC
  - focus on stations with high daily usage
- Consider proximity to technology centers, non-subway pedestrian traffic, and demographic data
- Apply data-scraping and geocoding tools

MTA Turnstile Data

- Data was as available as cumulative counts for each turnstile.
- Focused on weeks preceding the event (Spring 2018 data).
- Aggregated **total subway user traffic** for each station.
- Filtered outliers

MTA Turnstile Data

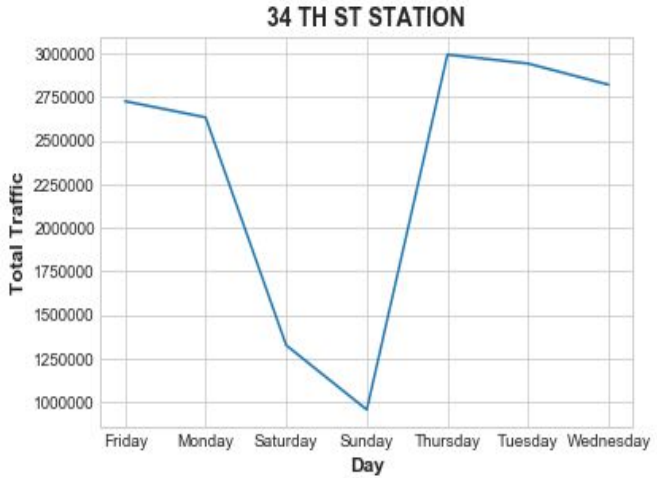
Stations with the highest traffic are mostly located in Manhattan.

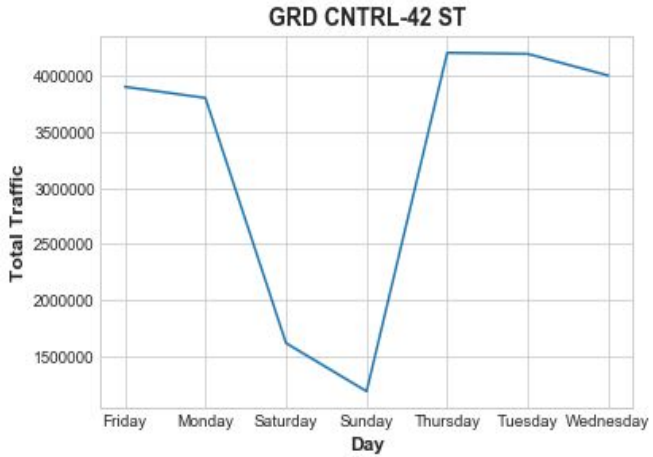


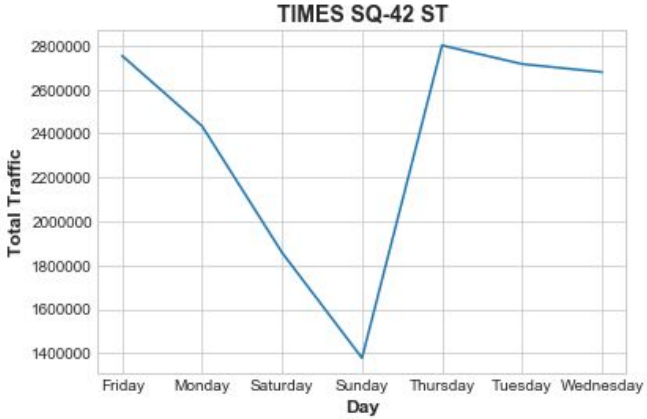
MTA Turnstile Data

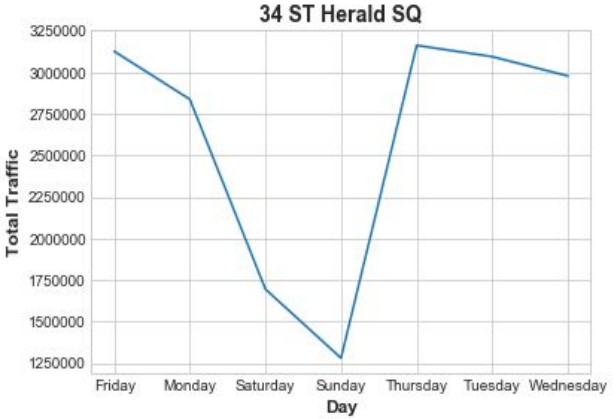
Data is consistent over days of the week.

Investigate the time interval for Thursday and Friday







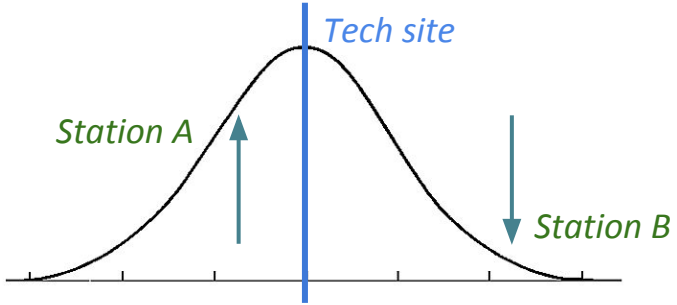


Tech Centers

Top 30 **tech companies**, scaled by # of employees *source: builtinnyc.com*

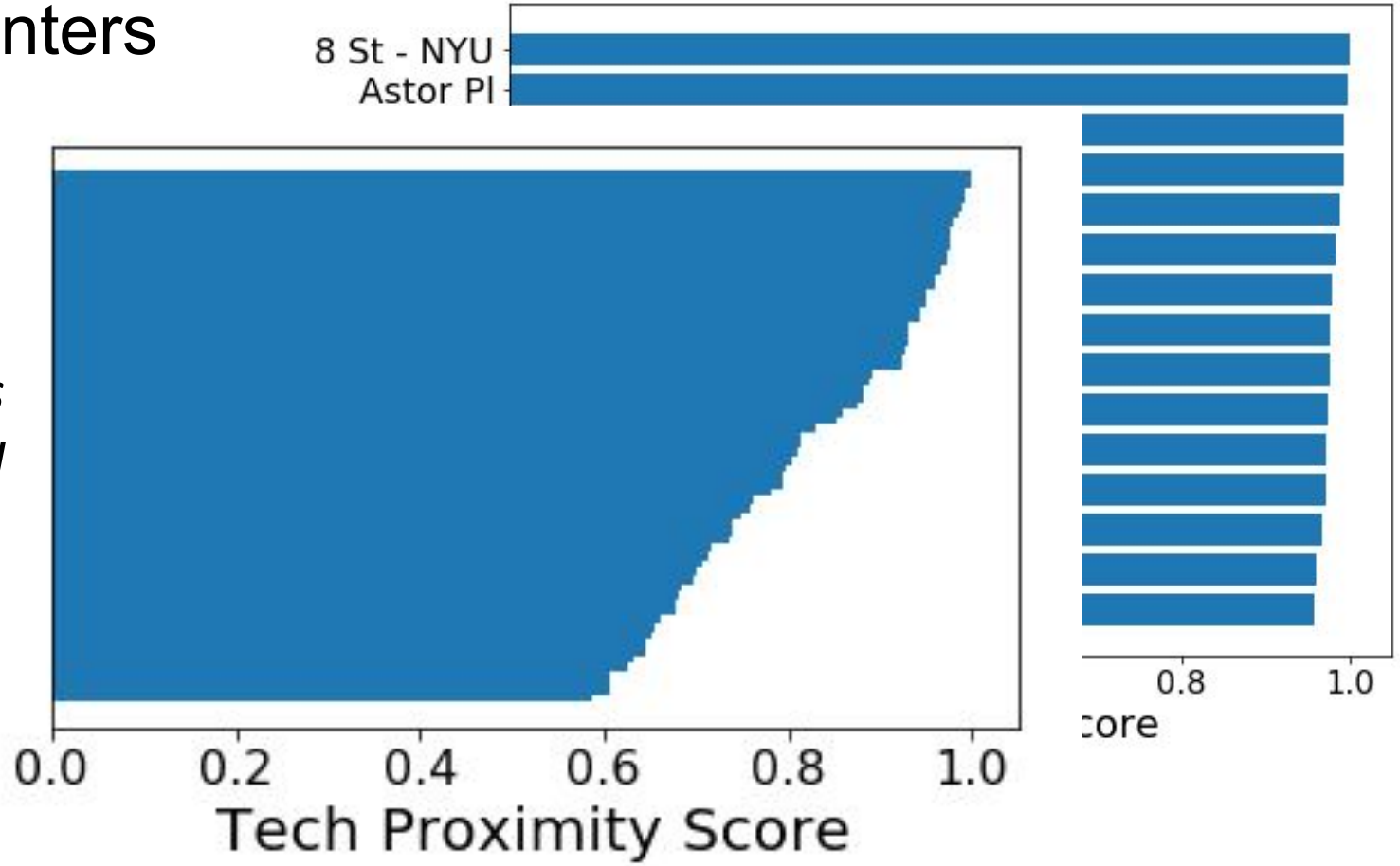
6 **largest universities**, scaled by # of students *source: Google.com*

Transformed site-to-station distances using a gaussian smoothing curve



Tech Centers

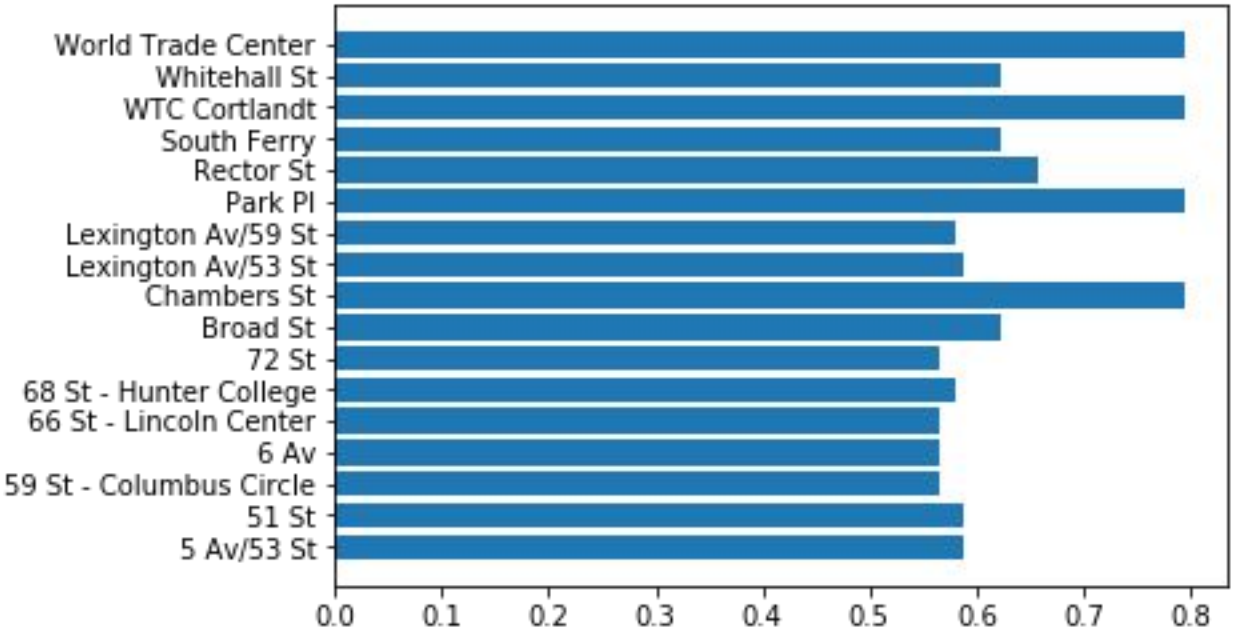
Stations ranked



Large unive
companies
increase fo
subway sta

Census Data

High Income by Station

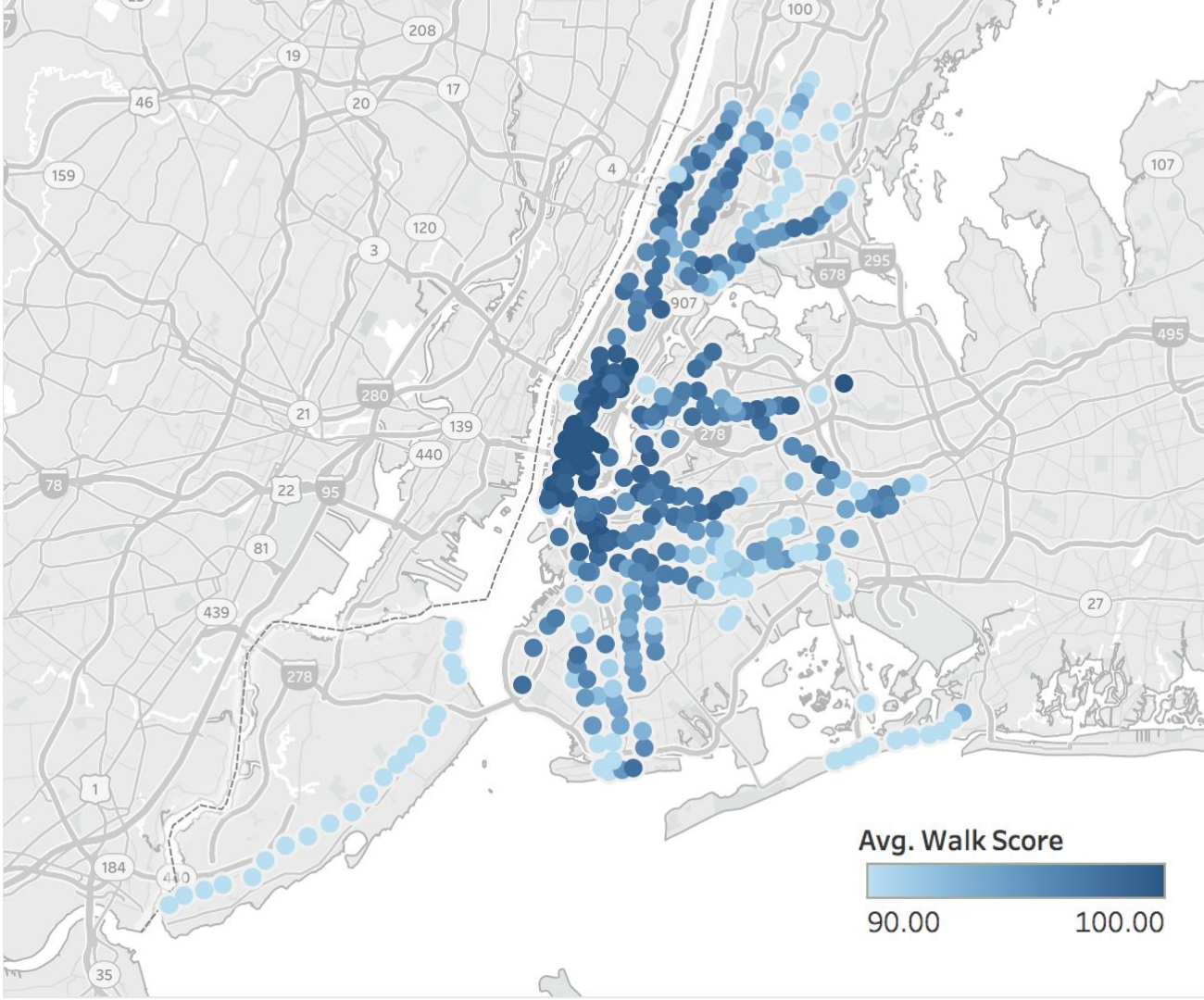


WalkScore

Higher WalkScore

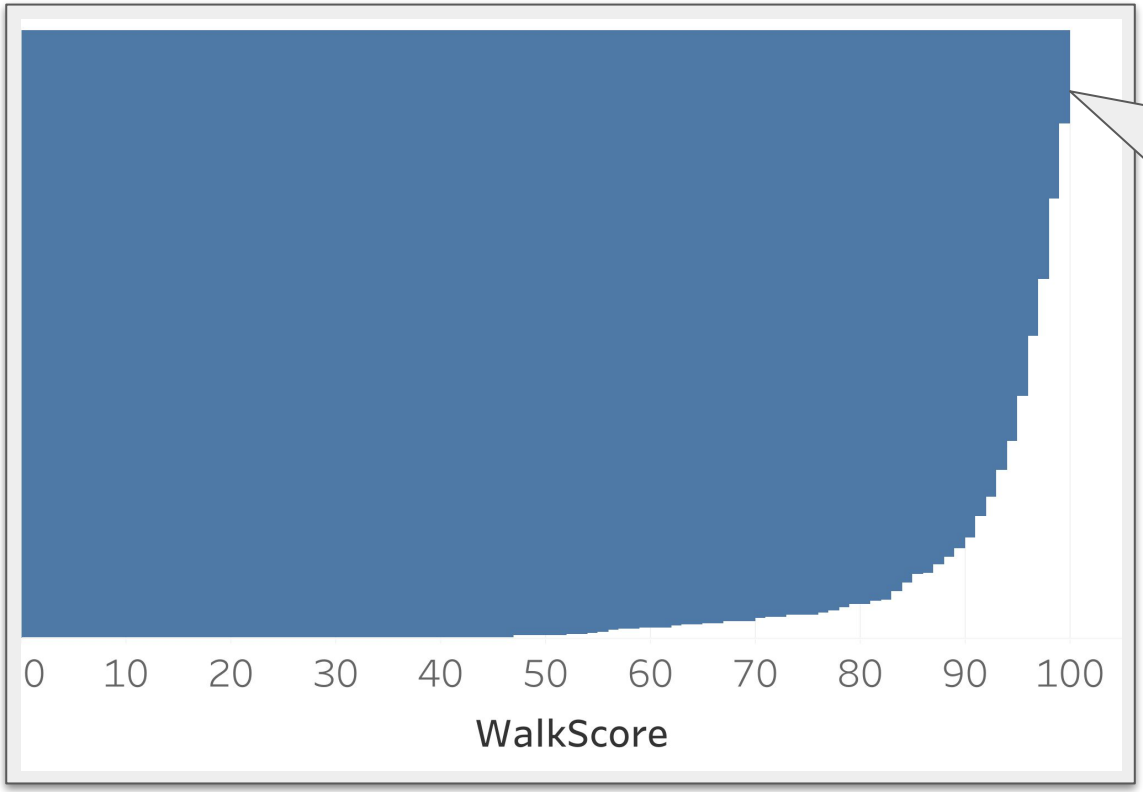
=

More Foot Traffic!



WalkScore

Stations ranked

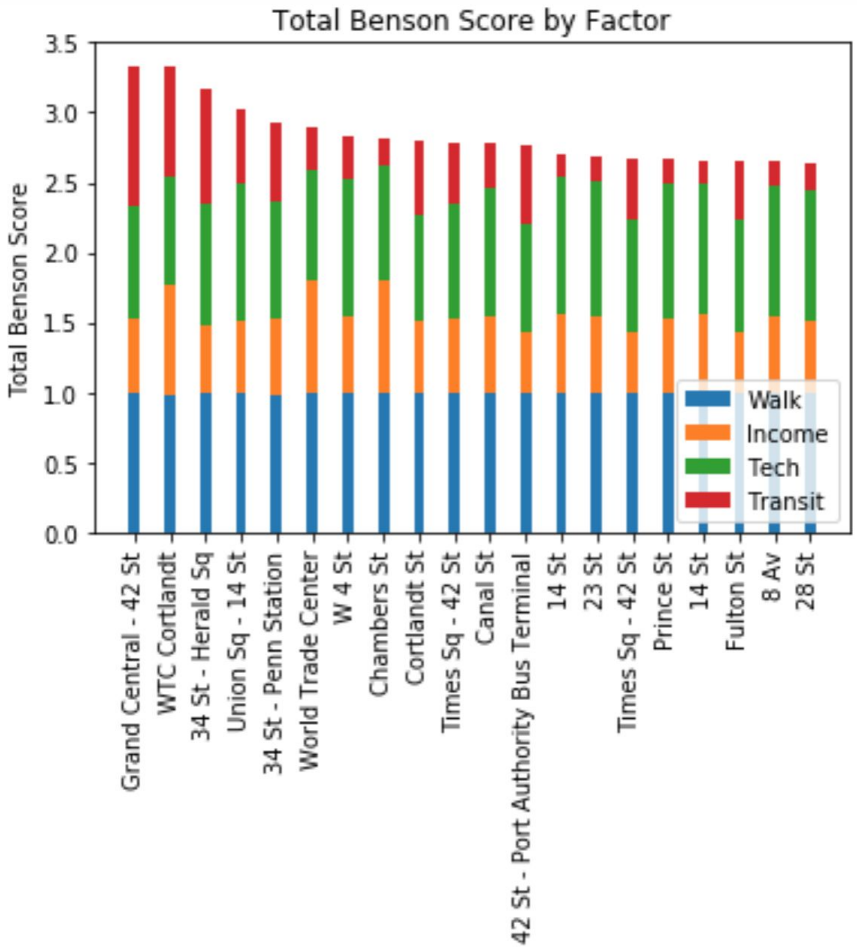


76 stations have a WalkScore of 100, and 60% have a WalkScore > 95

Final Recommendations for Canvassing

Addition of the technology, walking, and affluence factors suggest different stations than the turnstile data alone.

Top five stations are: Grand Central, World Trade Ctr, Herald Sq., Union Sq & 14th, Penn Station



Considerations for Future Analysis

- Analyze MTA-turnstile data to separate different stations with similar names.
- Hone in on the exact location of the top busiest stations.
- Use more tech and educational sites, and more granular demographic data.
- Adjust weights of input factors given client feedback

QUESTIONS?

APPENDIX

Assumptions

- Volunteers would start canvassing in Spring 2018 for a June 2018 gala.
- Target demographic will use the subway stations near their place of work.
- Chosen tech sites are representative of the city tech sector as a whole.