

WTWY Outreach Strategy

by Ahmed Abid & Amy Zhou

Background

Women at Tech Companies



Total positions filled by women



Leadership positions filled by women



Technical positions filled by women

Source: CNET

- Women Tech Women Yes (WTWY) - NYC non-profit
- Promoting annual summer gala at subway stations
- Promotion dates May 1-May 14
- 6 street teams, 2 promoters per team

(source: [cnet](#))

How might we identify which stations to post WTWY's street teams to, in order to maximize outreach for the summer gala?





Approach

→ DATA

MTA turnstile data

Range April 23rd - May 21st, 2015 to 2017

→ CLEANING

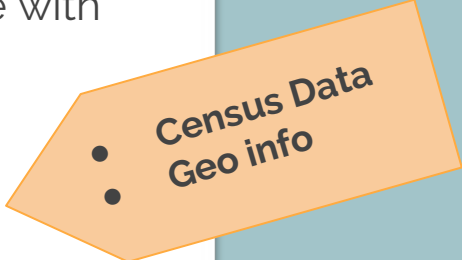
Remove outliers, out of scope data

→ ANALYSIS

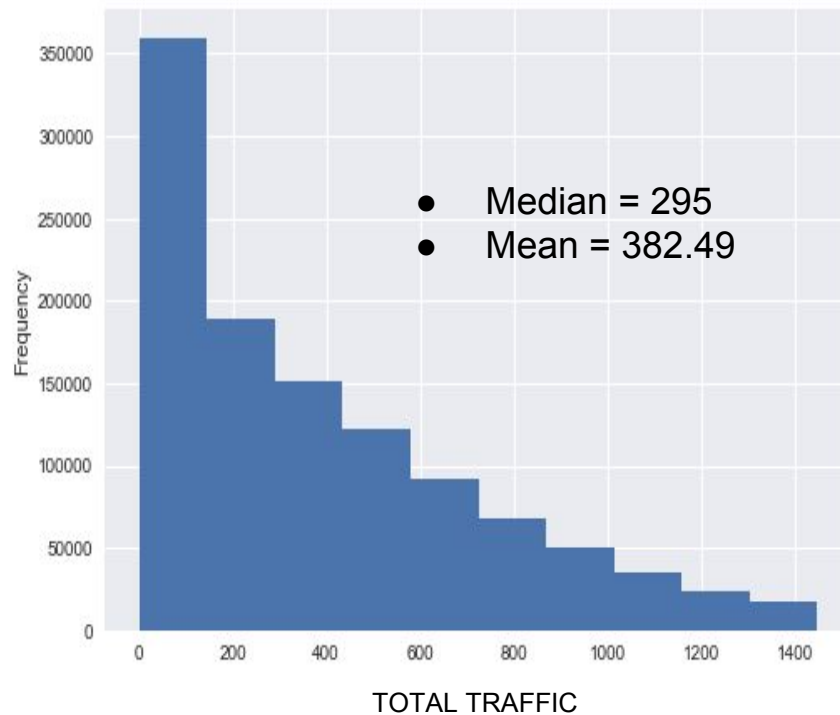
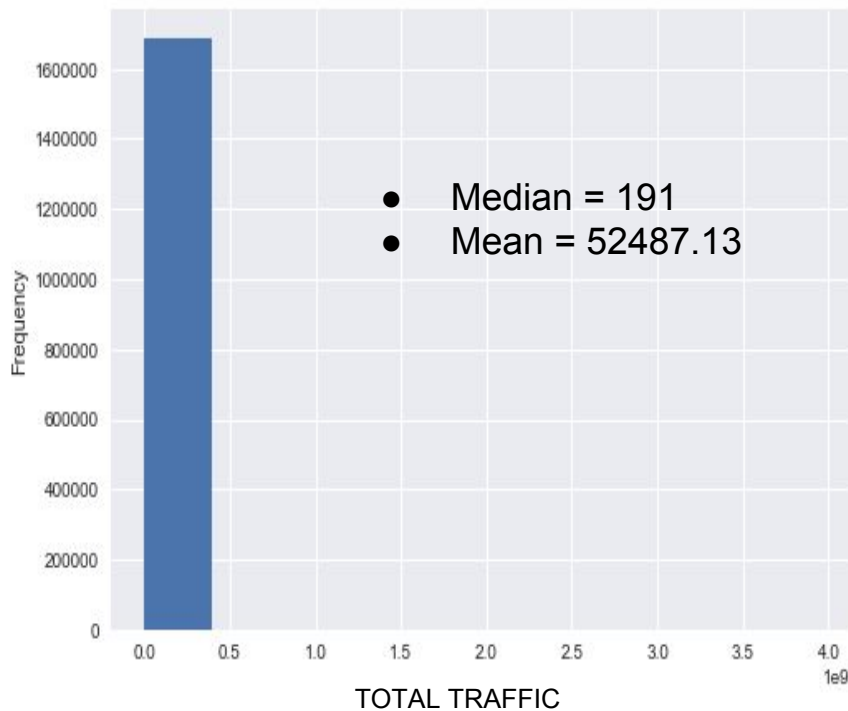
Find stations with highest traffic, compare with key factors

→ CONSIDERATIONS

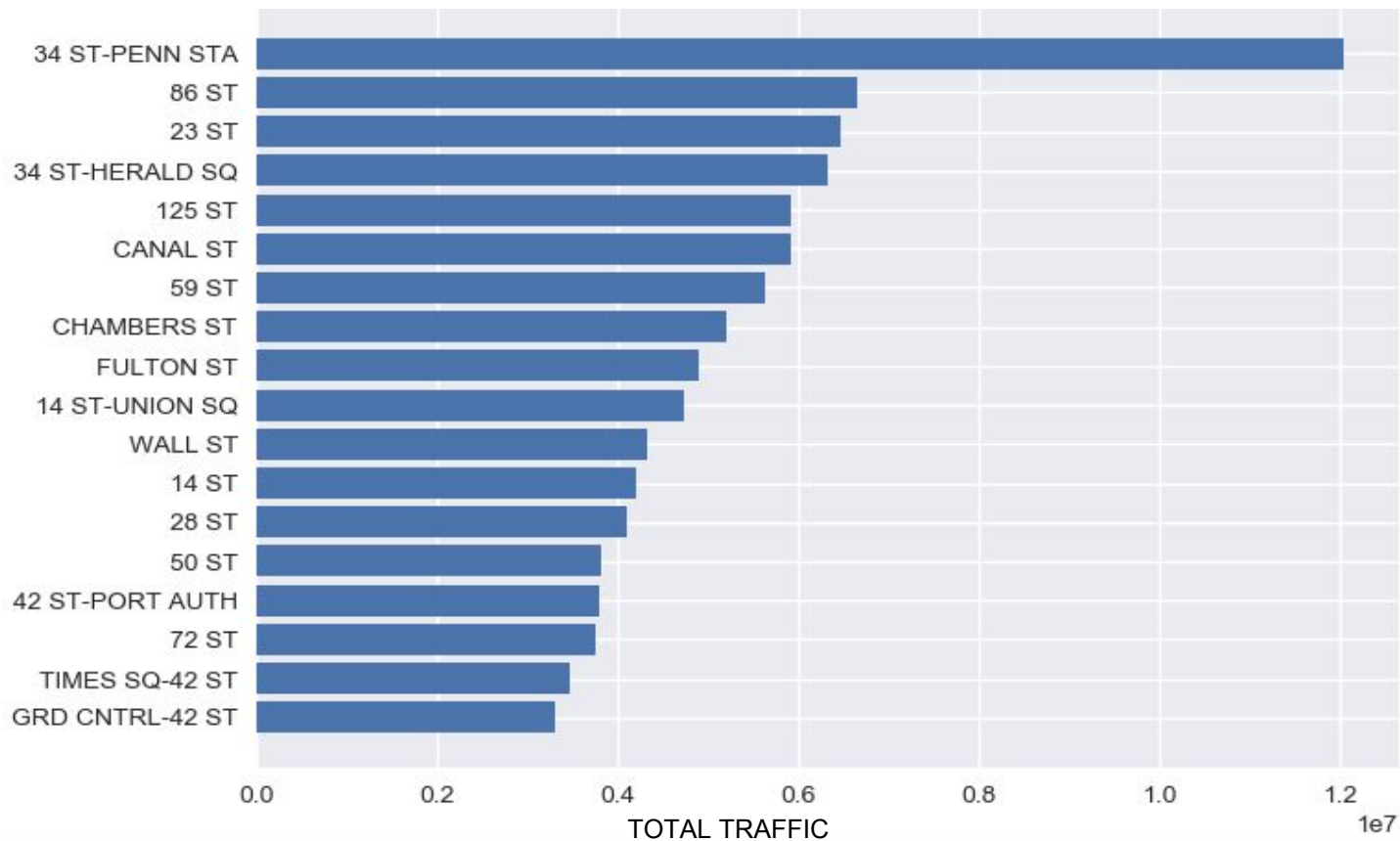
Distance, Target Population, Tech Hubs

- 
- Census Data
 - Geo info

Before & After Cleaning



Top Stations



34 Penn St

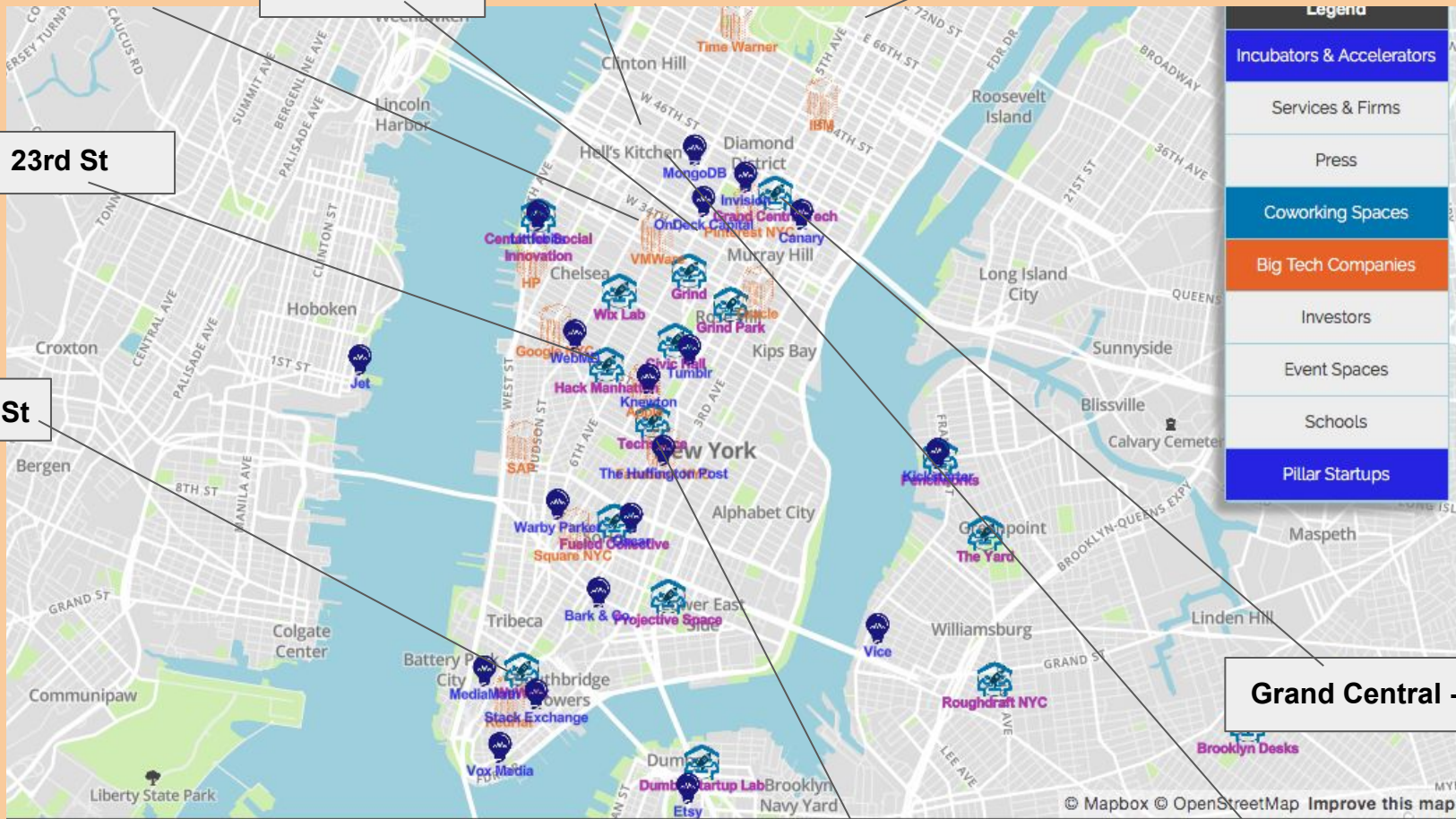
34 Herald St

42 St Port Auth

86th St

23rd St

Fulton St

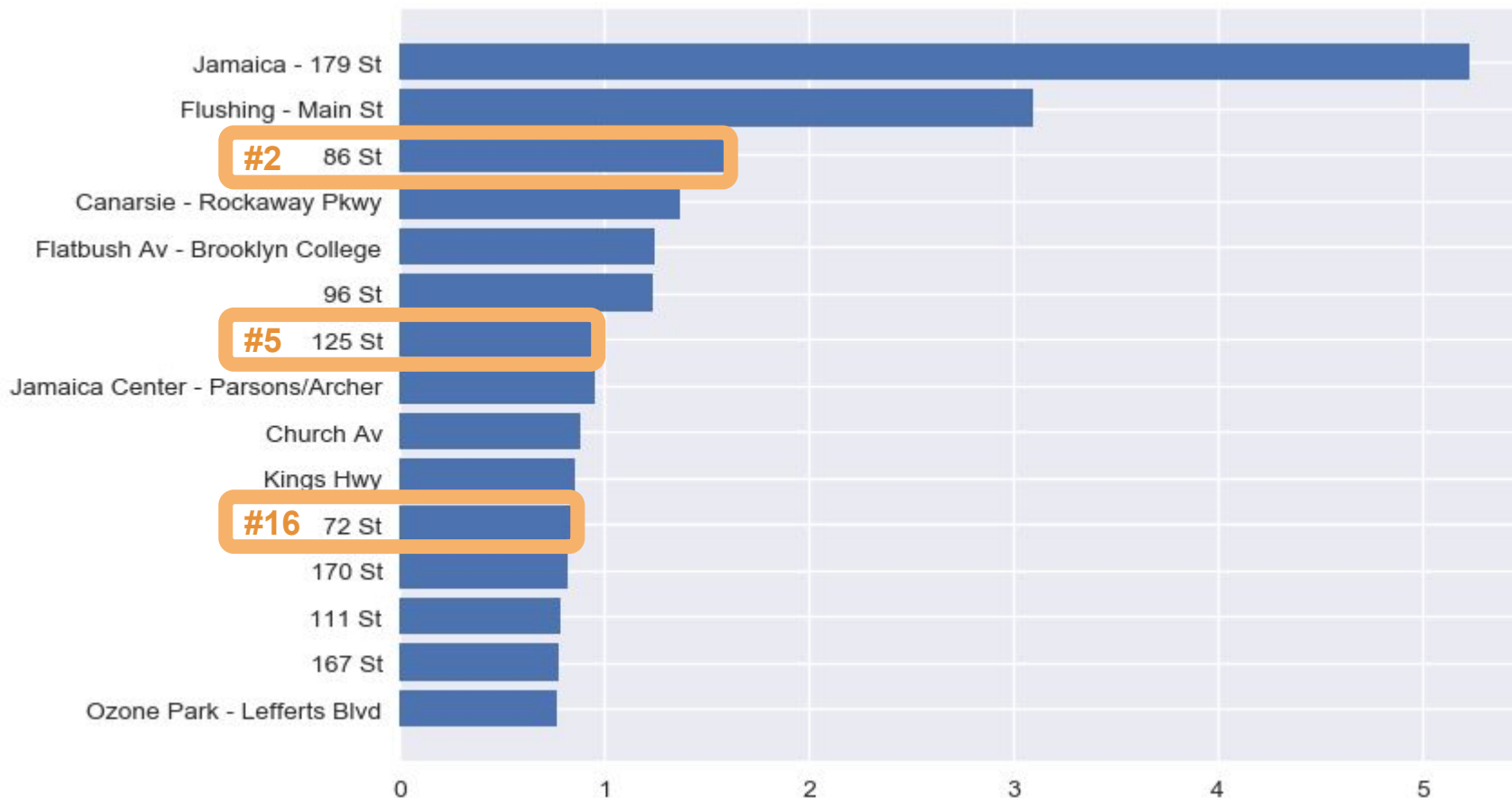


Grand Central -42 St

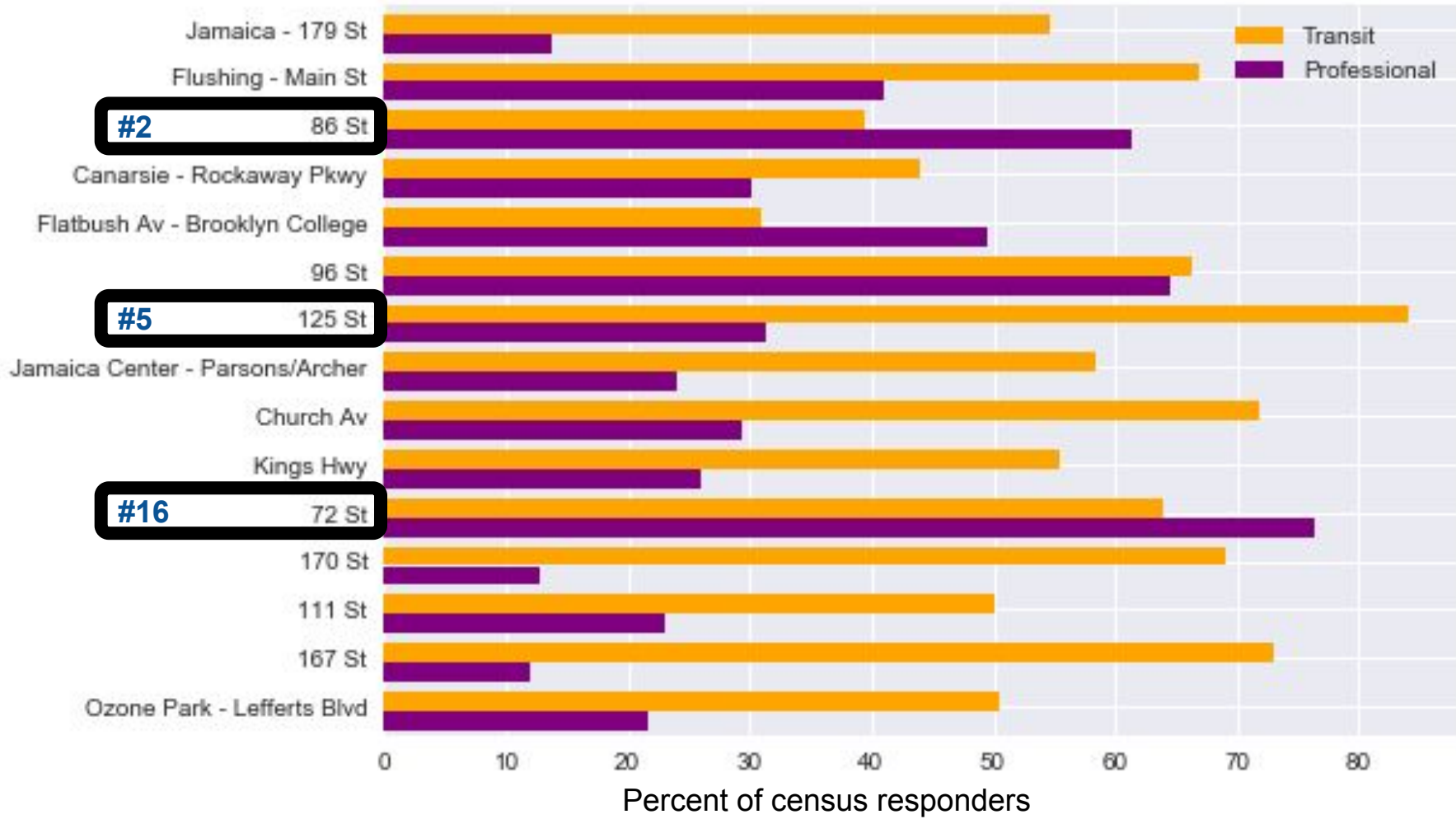
Map from startupguide.nyc

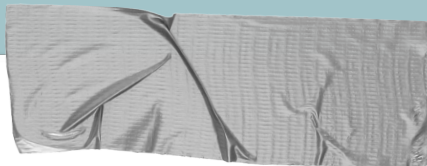
14 St-Union Sq

Time Sq-42 St



Percent of Women responding to census per station compared to all stations





Recommendations

→ **During commuter hours**

2 teams at Penn Station

1 team per 86 ST, 23 ST, 34th St Herald SQ, 125 ST

→ **After commuter hours**

Partial reallocation to 86 ST, 125 ST, 72 ST

→ **What's next?**

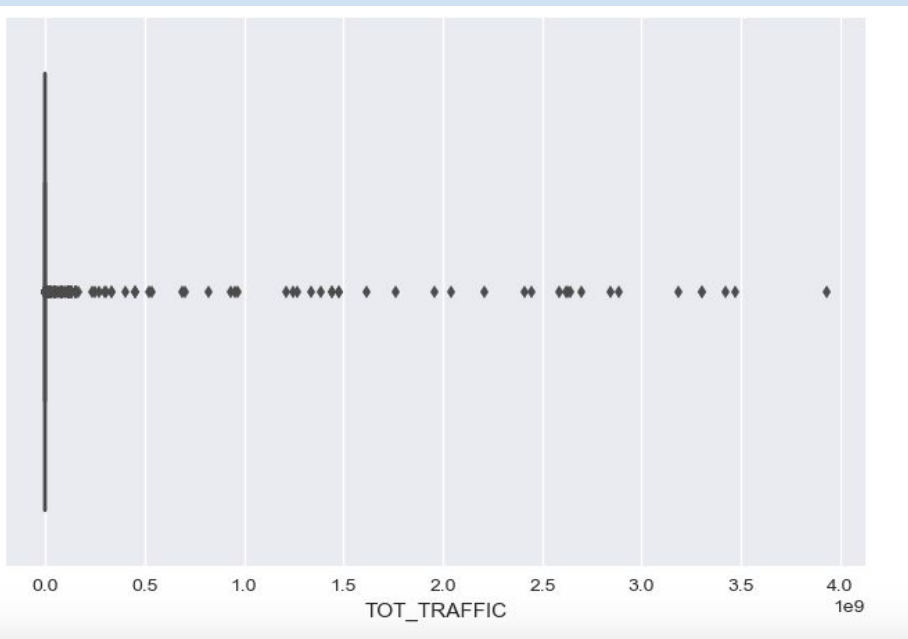
Target attendees of similar events

More specific schedule (time, day of week)

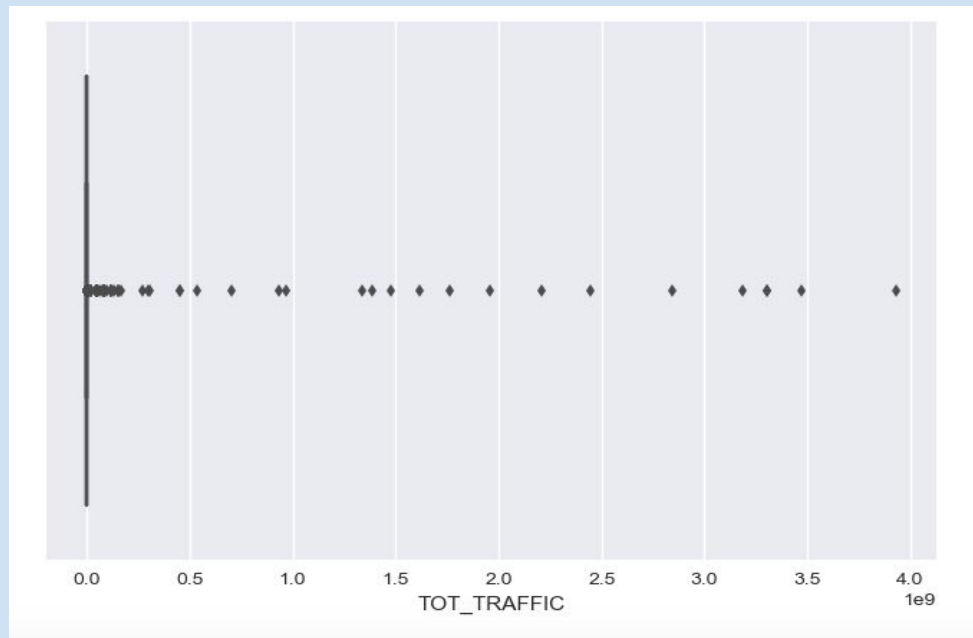
Optimize travel time for teams between stations

Explore expanding promotion period

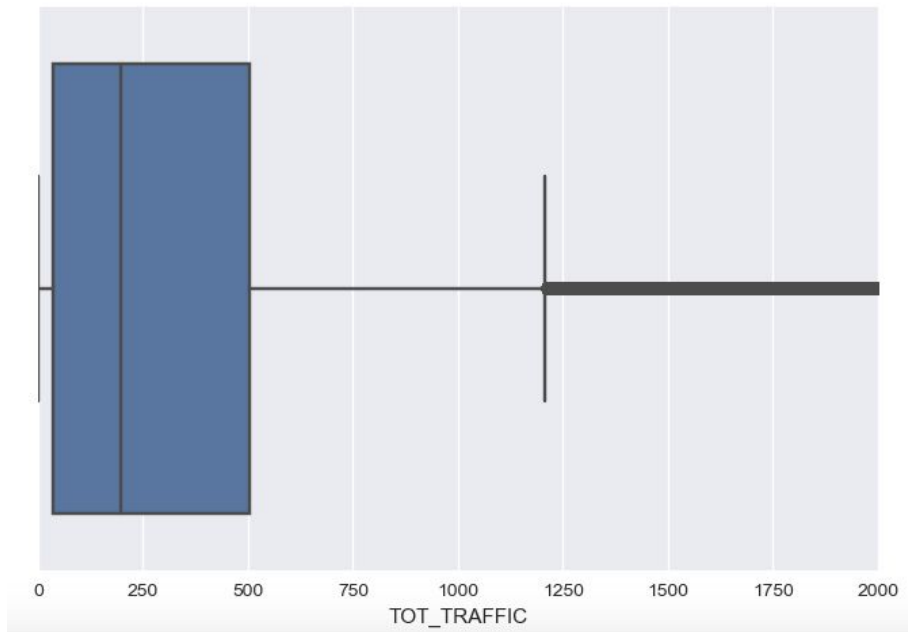
APPENDIX



Box Plot before
removing outliers

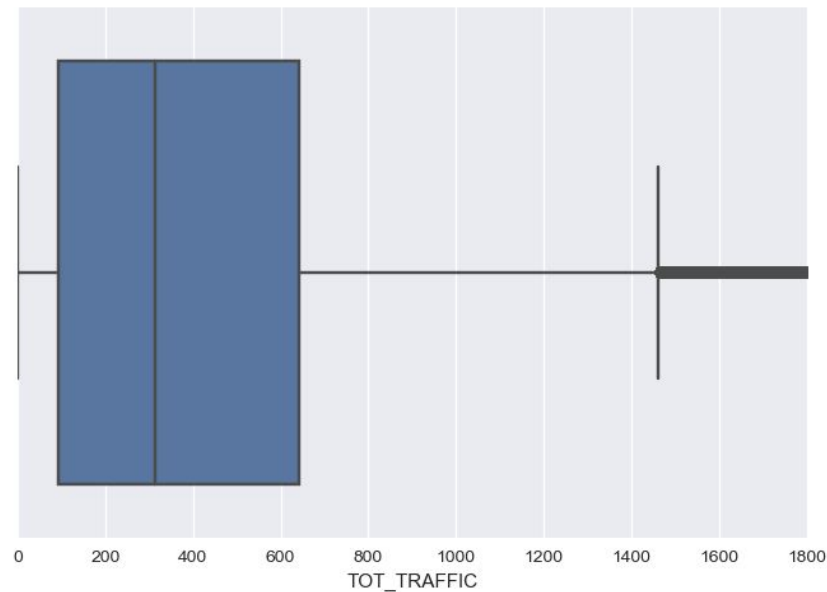


Box Plot after filtering time
period 12am - 6am



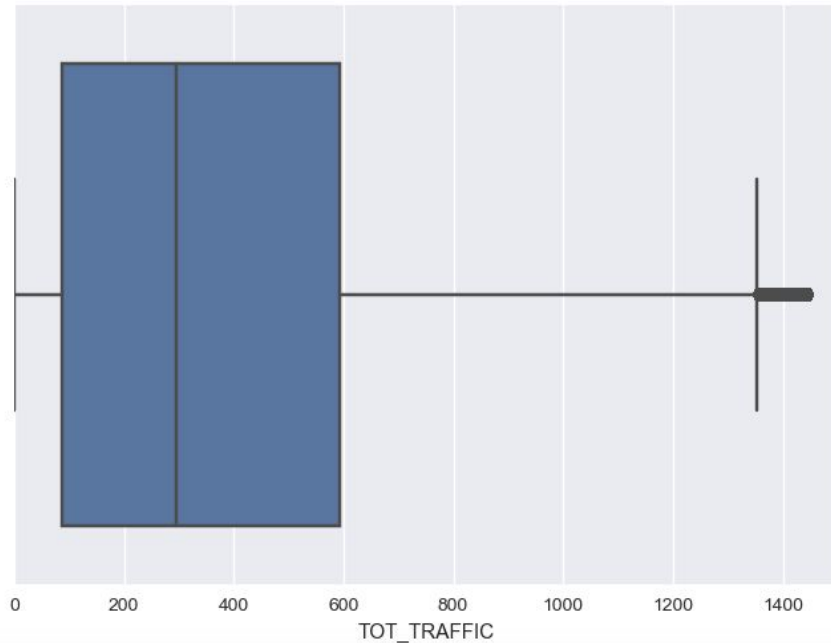
Before filtering out time period
between 12am - 6am

Median = 191
Mean = 52487.13



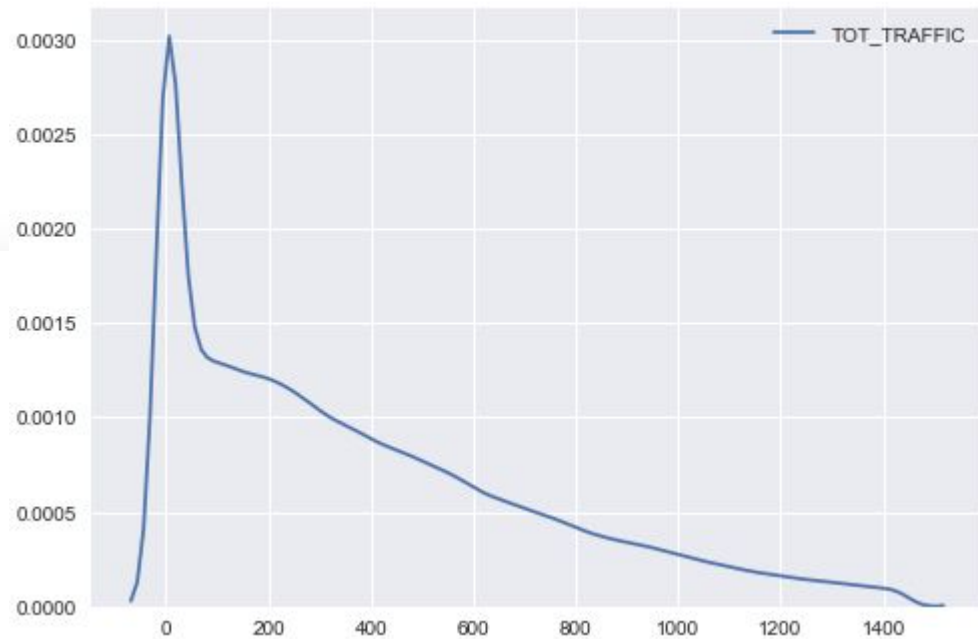
After filtering out time period
between 12am - 6am

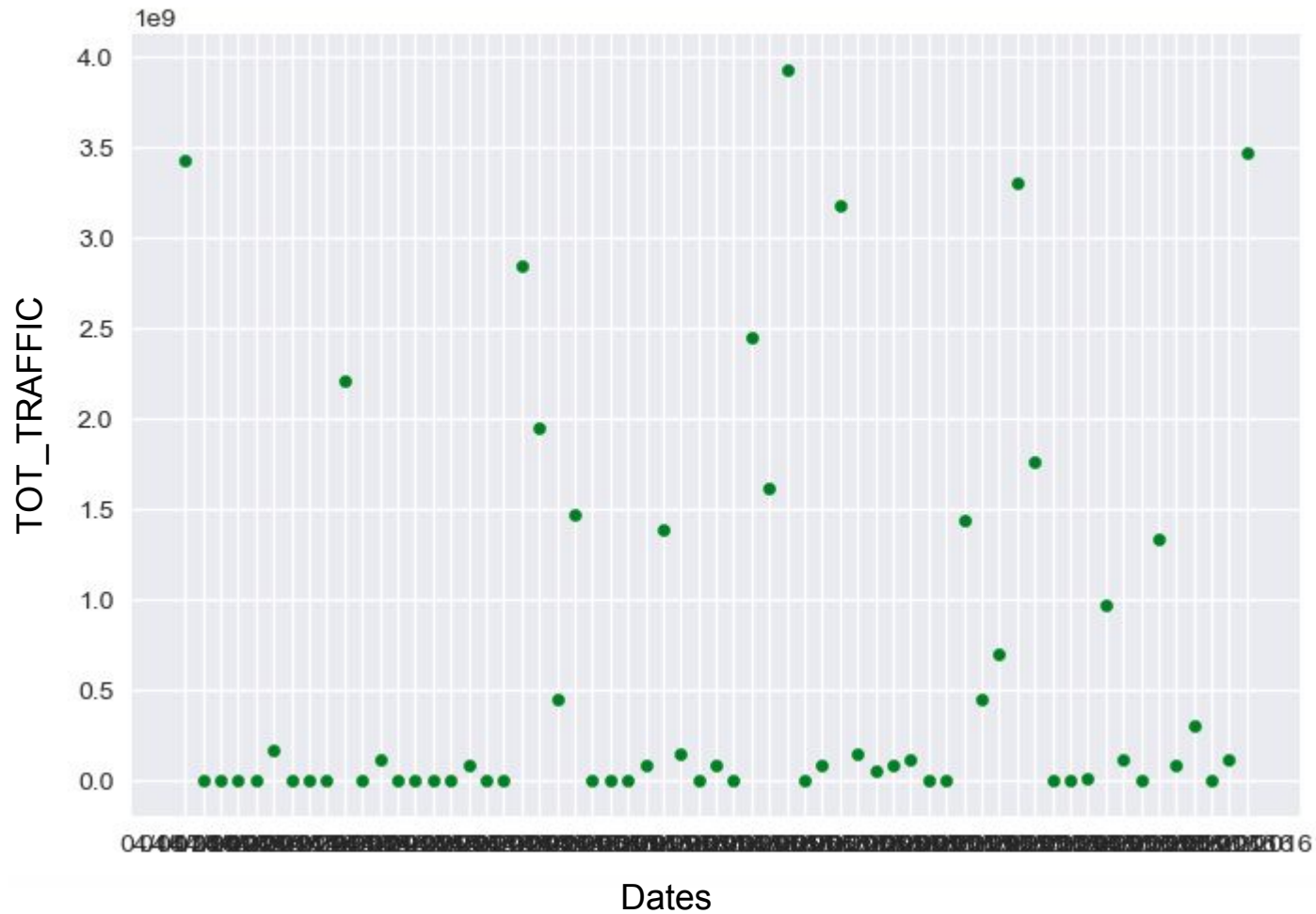
Median = 314
Mean = 38643.54



Kept 96.2% of data after cleaning

- Median = 295
- Mean = 382.49





Plot shows total traffic by day, showing that there are too many points by day to remove all days with Traffic over a certain threshold