

MTA EDA

By: Yi Chiang
Dan Allison
Cyrus Rustomji

Rank

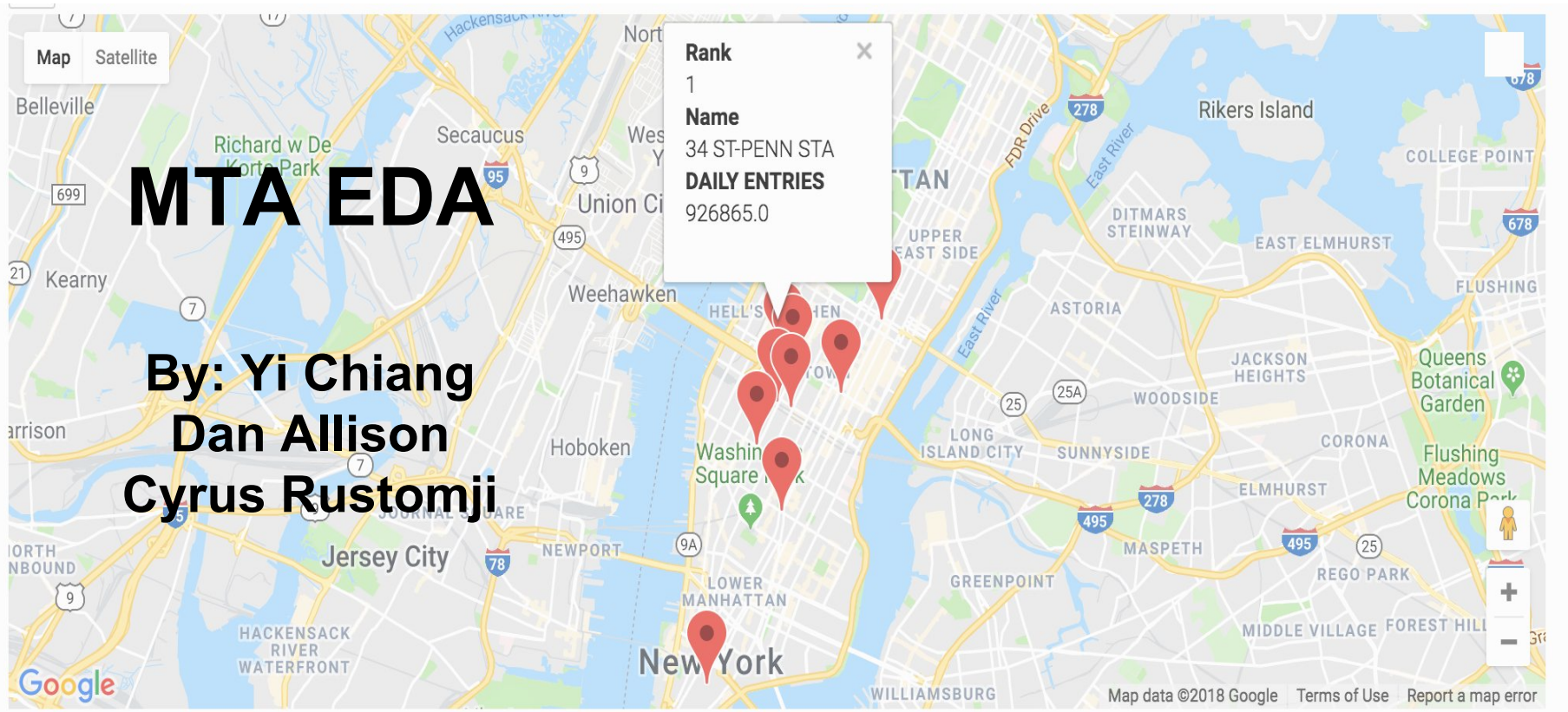
1

Name

34 ST-PENN STA

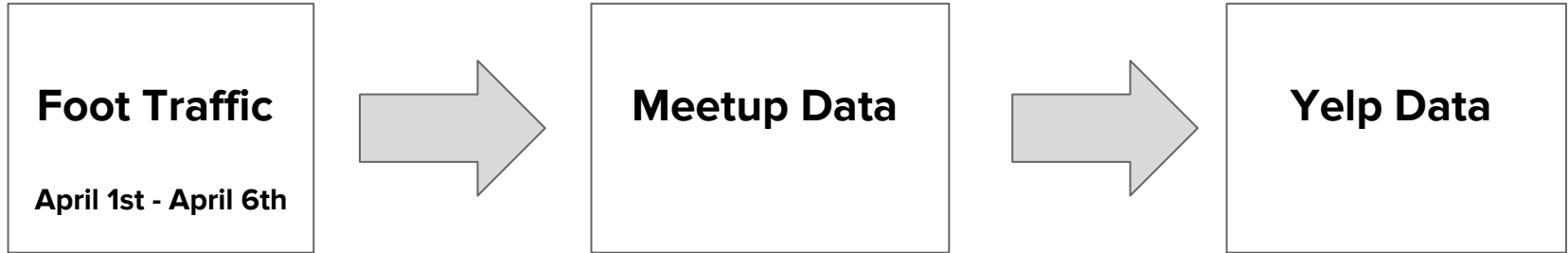
DAILY ENTRIES

926865.0



Abstract

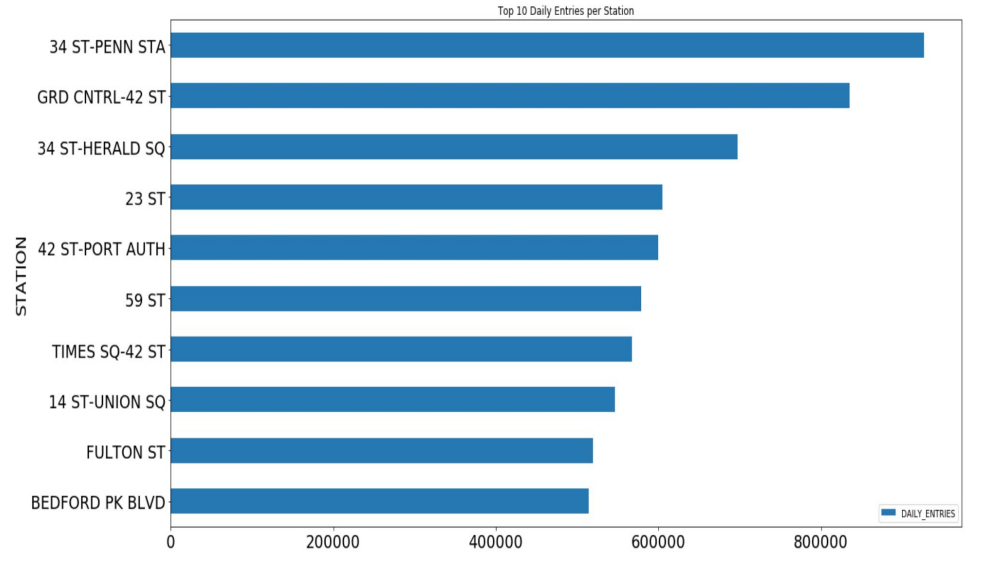
Our goal is to identify the optimal transit stations to place the WTWY street team.
The factors we considered in our investigation are:



Foot Traffic

Which stations have the most traffic?

Rank	Station	Daily Entries
1	34 St-Penn Sta	926,865
2	Grd Cntrl-42 St	835,361
3	34 St-Herald Sq	697,252
4	23 St	604,723
5	42 St-Port Auth	600,254
6	59 St	579,229
7	Times Sq-42 St	567,137
8	14 St-Union Sq	546,984
9	Fulton St	519,745
10	Bedford Pk Blvd	514,473

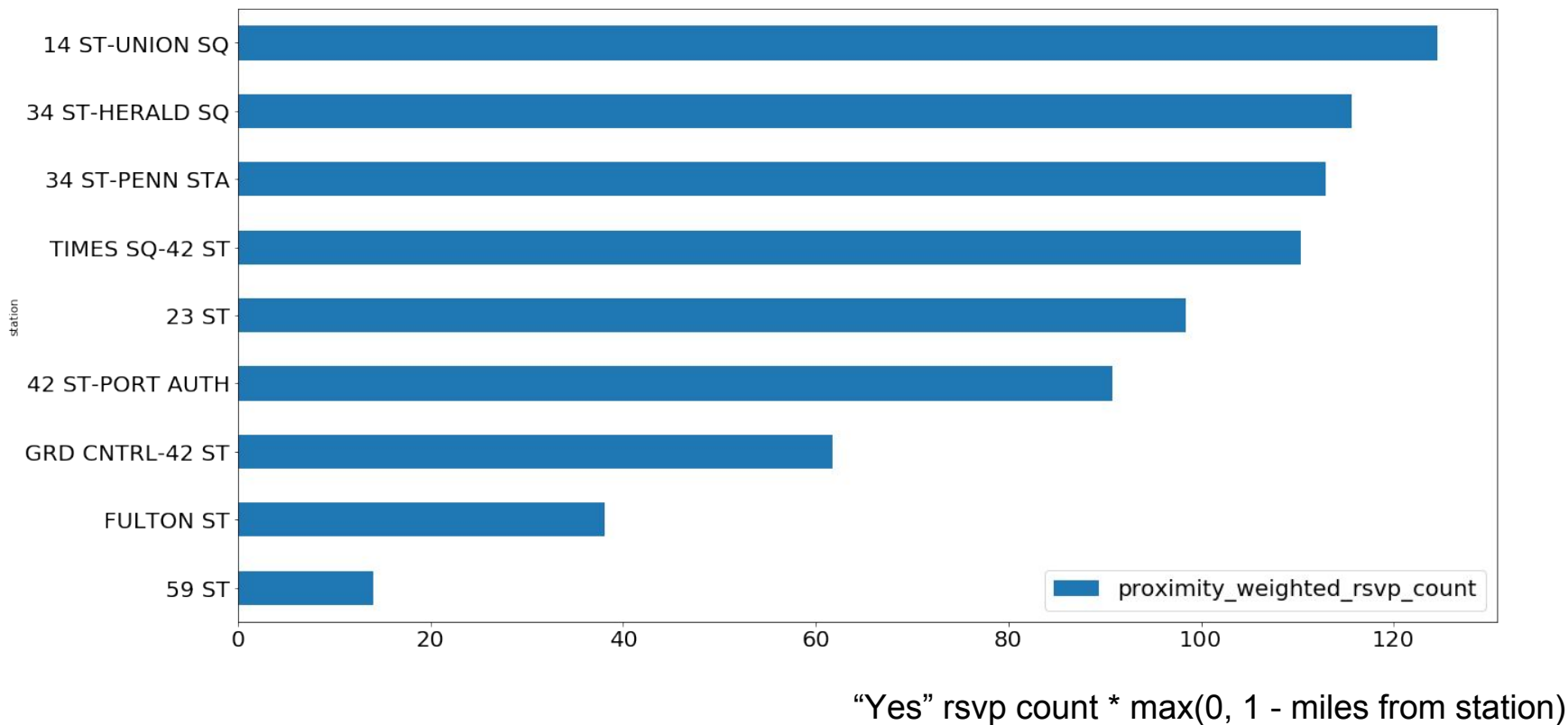


meetup

Women in Tech Meetup Groups

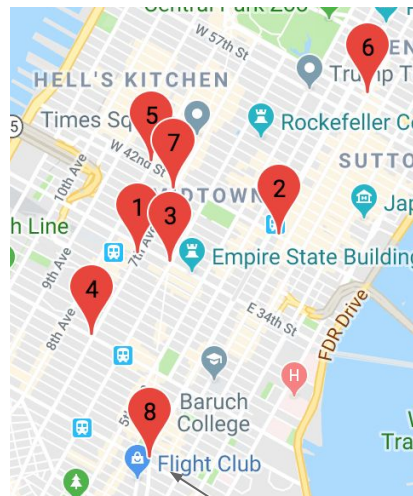
- NYC Women in Machine Learning & Data Science
- Women Who Code NYC
- NY Women in Tech
- Girl Develop It NYC
- League of Women Coders
- Women Investing in Women Entrepreneurs
- Women in Hyper Growth Companies
- Women in Agile New York City
- New York Women in Crypto & Blockchain Investment Club Meetup
- ...

WiT Meetup RSVPs weighted by proximity

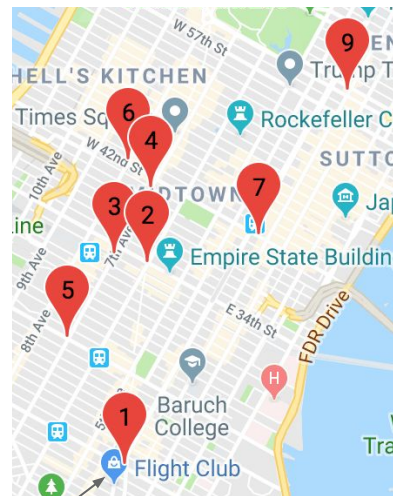


Traffic vs. Meetup Rank

Rank by Traffic



Rank by Meetup RSVPs



14 St Union Sq



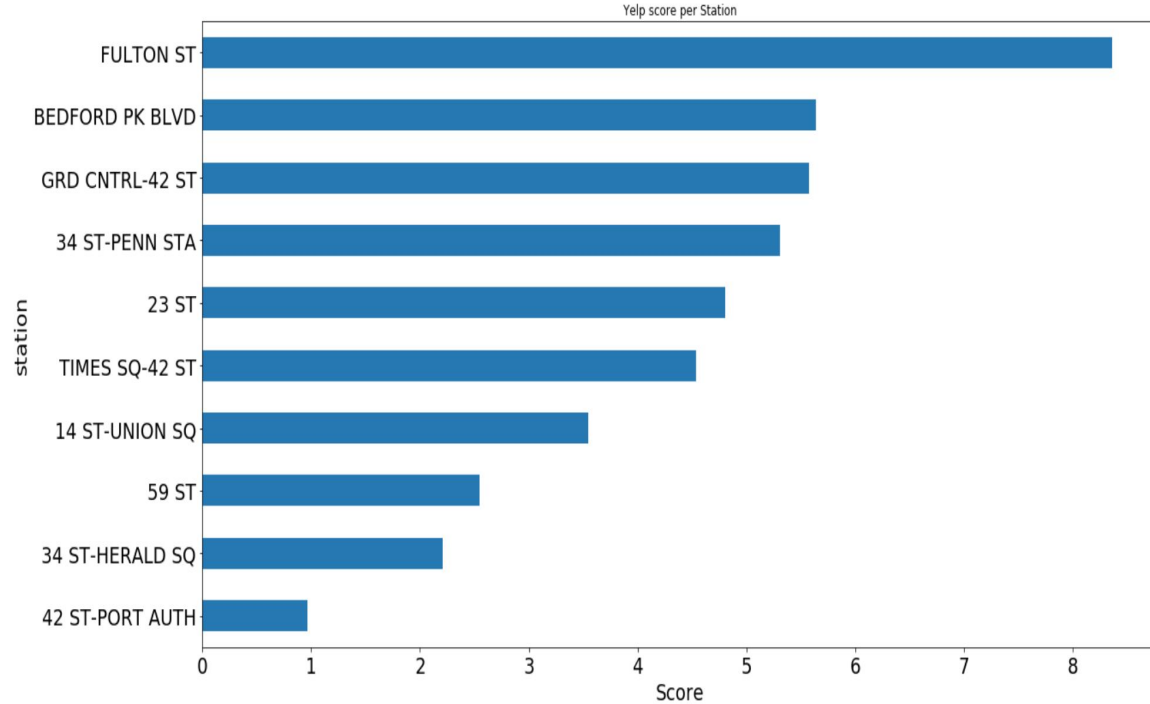
Yelp Score

Score =

(Mean rating of restaurants * 2)

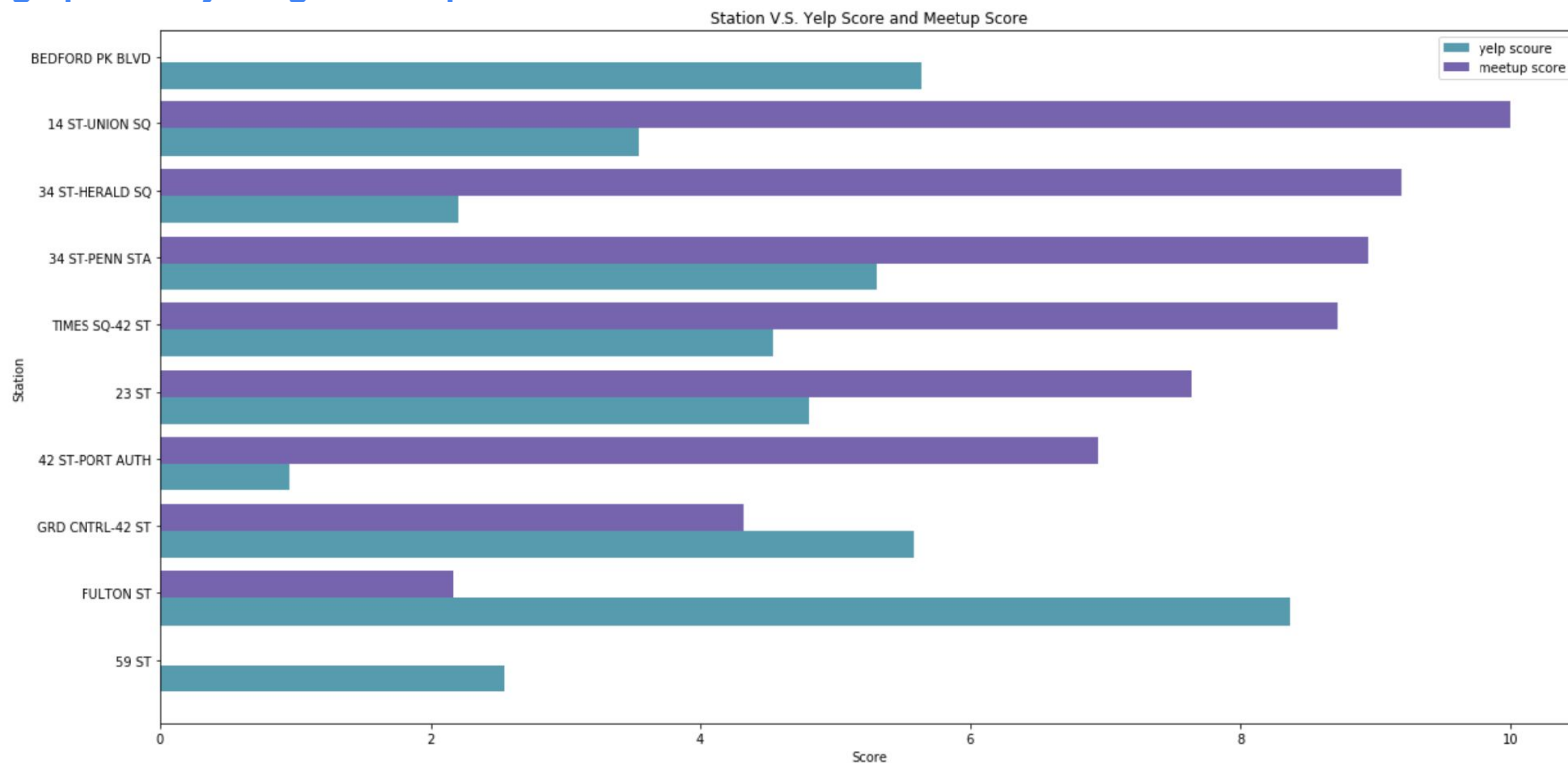
+ (# of restaurants * 4.5)

+ (Distance between the station * 3.5)



Yelp score and Meetup score

Average proximity weighted rsvp count

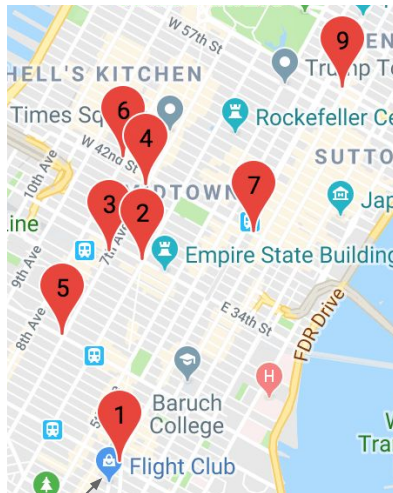


Traffic vs. Meetup vs. Yelp Rank

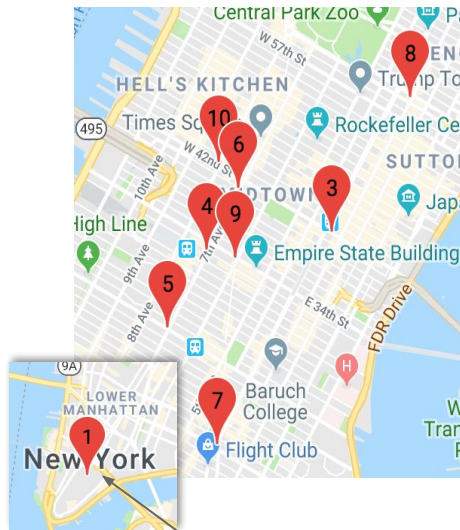
Rank by Traffic



Rank by Meetup RSVPs



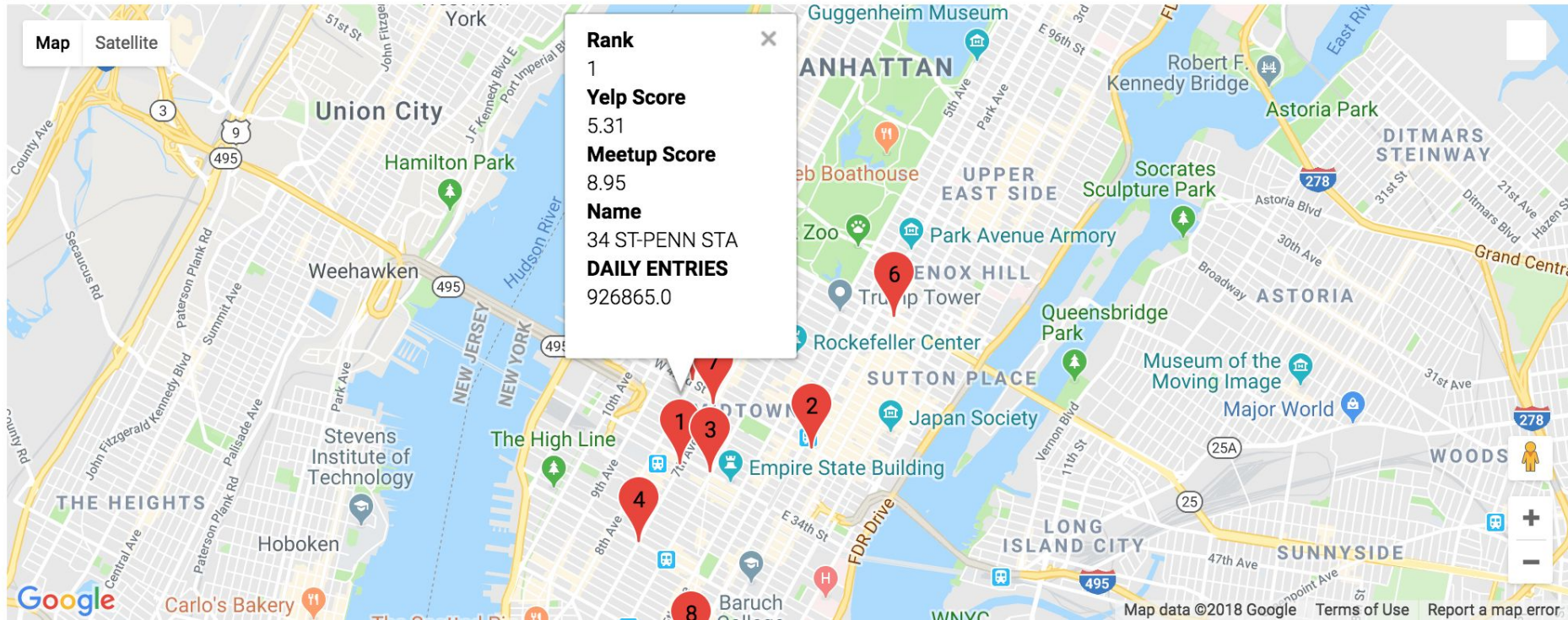
Rank by Yelp Score



14 St Union Sq

Fulton St

Map



Overall

Final Rank	Station	Daily Entries	Meetup Score	Yelp Score	Total Score
1	34 St-Penn Sta	926,865	8.95	5.31	18.735
2	14 St-Union Sq	546,984	10	3.55	18.55
3	Times Sq-42 St	567,137	8.72	4.53	17.61
4	34 St-Herald Sq	697,252	9.2	2.21	16.01
5	Grd Cntrl-42 St	835,361	4.32	5.58	12.06
6	Fulton St	519,745	2.17	8.36	11.615
7	42 St-Port Auth	600,254	6.95	0.96	11.385
8	Bedford Pk Blvd	514,473	0	5.64	5.64
9	59 St	579,229	0	2.55	2.55
10	23 St	604,723	7.64	4.81	1

Total Score = (Meetup Score * 1.5) + (Yelp Score * 1)

Data Source

- **MTA Stops:** [New York City Transit Subway](#)
- **MTA Traffic Data:** [http://web.mta.info/developers/data/nyct/turnstile/turnstile_180407.txt](#)
- **Meetup:** [https://www.meetup.com/meetup_api/docs/find/upcoming_events/](#)
- **Yelp:** [https://www.yelp.com/developers/documentation/v3/business_search](#)

Thank you

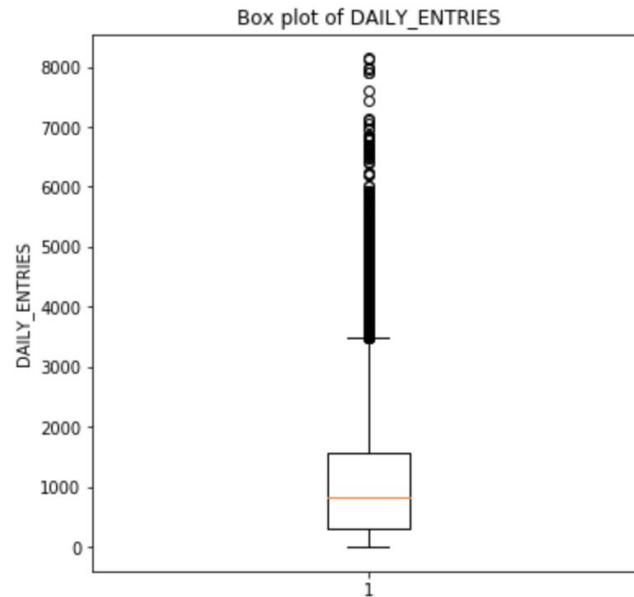
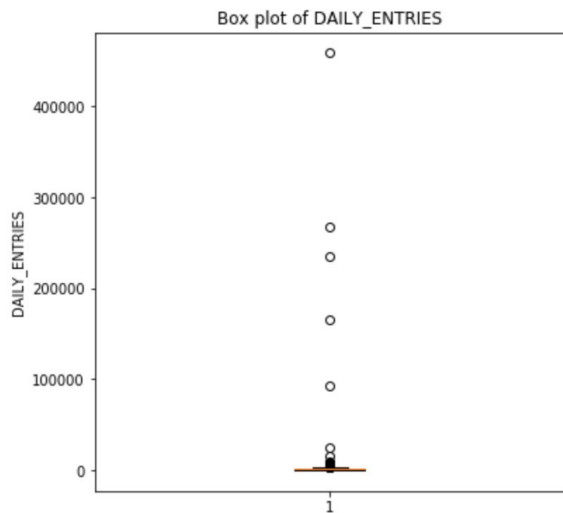
Backup Slides

Future work

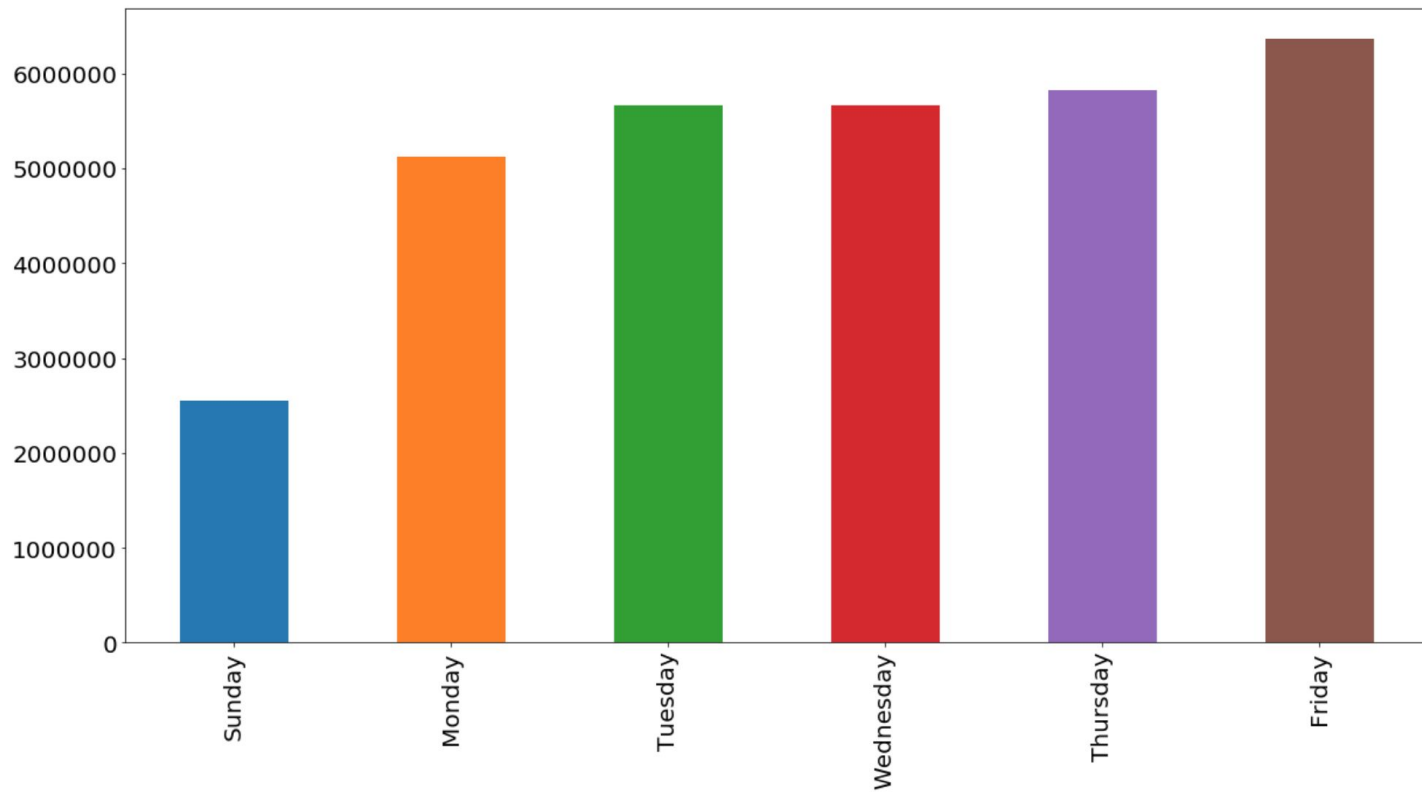
- Time
- Length
- Did not target certain events (Ex: Women's March)
- “Gala is in the summer”
- Used only one week

Other EDA

- Outliers
 - Outliers = 2 standard deviation away from mean
 - Replace outliers with mean value



Total Traffic by date



Time of day

When should the street teams arrive?

- Utilize outreach teams from 8 am - 7 pm on the weekdays (Monday - Thursday)
- All teams will be working from from 7 am - 8 am, 11 am - 1 pm, and 4 pm - 6 pm
- All teams will be working 11 pm - 1 am on the weekends (Friday - Sunday)

Scatter Matrix Plot

