Mobilization of Volunteers

Presentation for Women Tech Women Yes
Seattle, 13th, April
Sirius A Consulting Group

- Travis Trumbly
- Ryan Murray
- Sakura Lin

Agenda

1	Initial Question
2	Assumptions & Cleaning Data
3	EDA Outcomes
4	Further Investigation

Agenda

1. Initial Question
2. Assumptions & Cleaning Data
3. EDA outcome
4. Further investigation

Initial question

- Who is Women Tech Women Yes' target audience?

- Where sd we place volunteers?

- hould we place volunteersWhat days and times?

- How can we maximize contributions?we

Agenda

1	Initial Question
2	Assumptions & Cleaning Data
3	EDA outcome
4	Further investigation

Assumptions

For a gala in July, volunteers should be mobilized in May to allow:

- Collection and processing of emails and signatures
- Time to send email reminders and solicit donations
- Early enough that potential guests' schedules are not already booked

We used data from April - June 2017

- Winter and Spring usage patterns are slightly different
- This is the most current relevant data

Cleaning Data

Two types of abnormal data

1. Missing value

	SCP	STATION	ENTRIES	EXITS	4HR_Entries	4HR_Exits
117329	01-01-00	LACKAWANNA	37436	21210	11.0	3.0
117330	01-01-00	LACKAWANNA	980	504	-36456.0	-20706.0
117331	01-01-00	LACKAWANNA	37642	21218	36662.0	20714.0
117332	01-01-00	LACKAWANNA	37650	21220	8.0	2.0

	SCP	STATION	ENTRIES	EXITS	hr4
117329	01-01-00	LACKAWANNA	37436	21210	11.0
117330	01-01-00	LACKAWANNA	980	504	103.0
117331	01-01-00	LACKAWANNA	37642	21218	103.0
117332	01-01-00	LACKAWANNA	37650	21220	8.0

2. Anomalous jump

1. No clear reason

	SCP	STATION	ENTRIES	EXITS	4HR_Entries	4HR_Exits
174568	00-00-03	5 AVE	165330	25453	18.0	0.0
174569	00-00-03	5 AVE	165345	25469	15.0	16.0
174570	00-00-03	5 AVE	838900529	151035100	838735184.0	151009631.0
174571	00-00-03	5 AVE	838900808	151035128	279.0	28.0
174572	00-00-03	5 AVE	838901512	151035146	704.0	18.0

2. Detection Machine resets

Solution: 1.Take Average 2.Drop data

Result: Only 2% data loss

Agenda

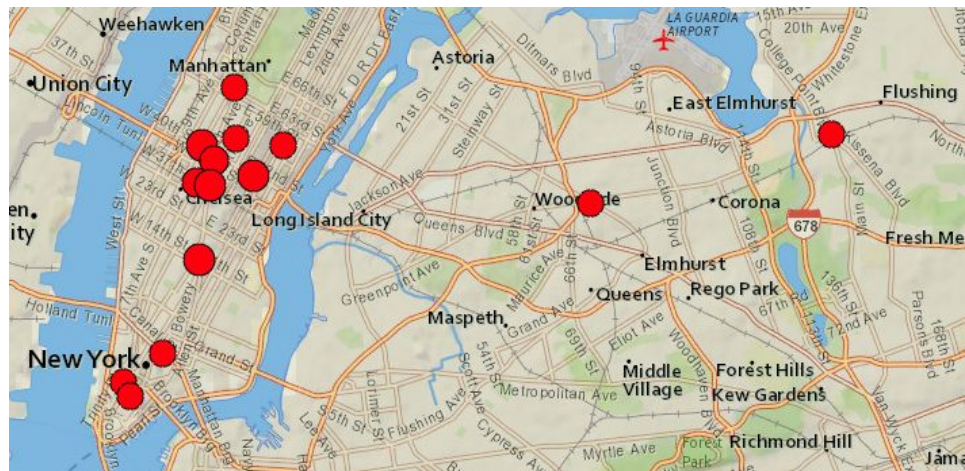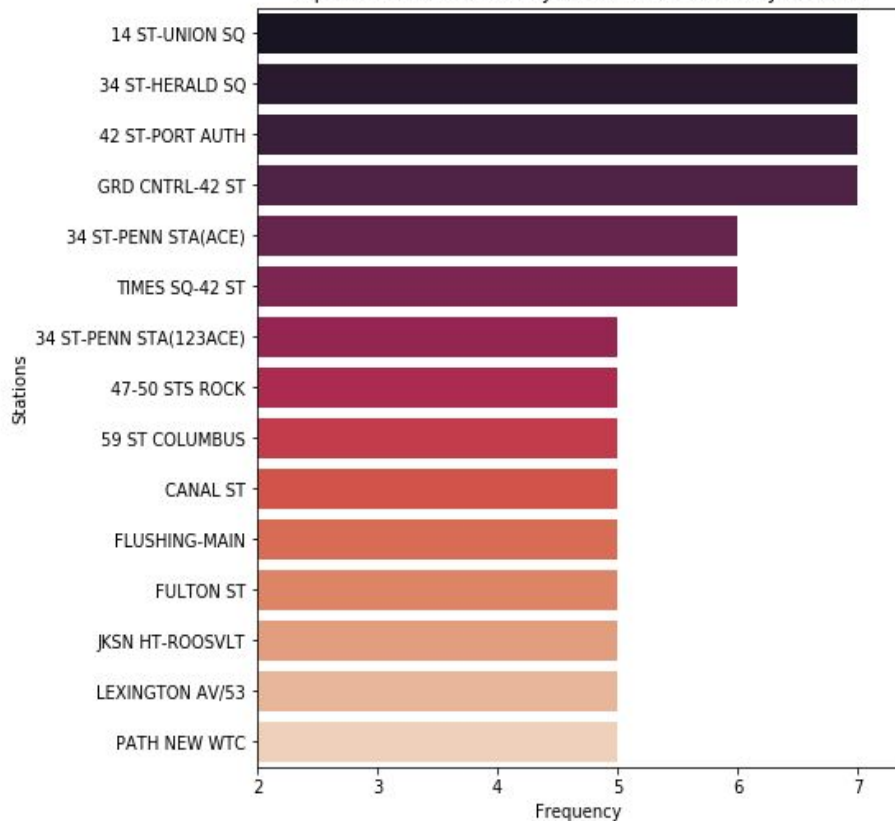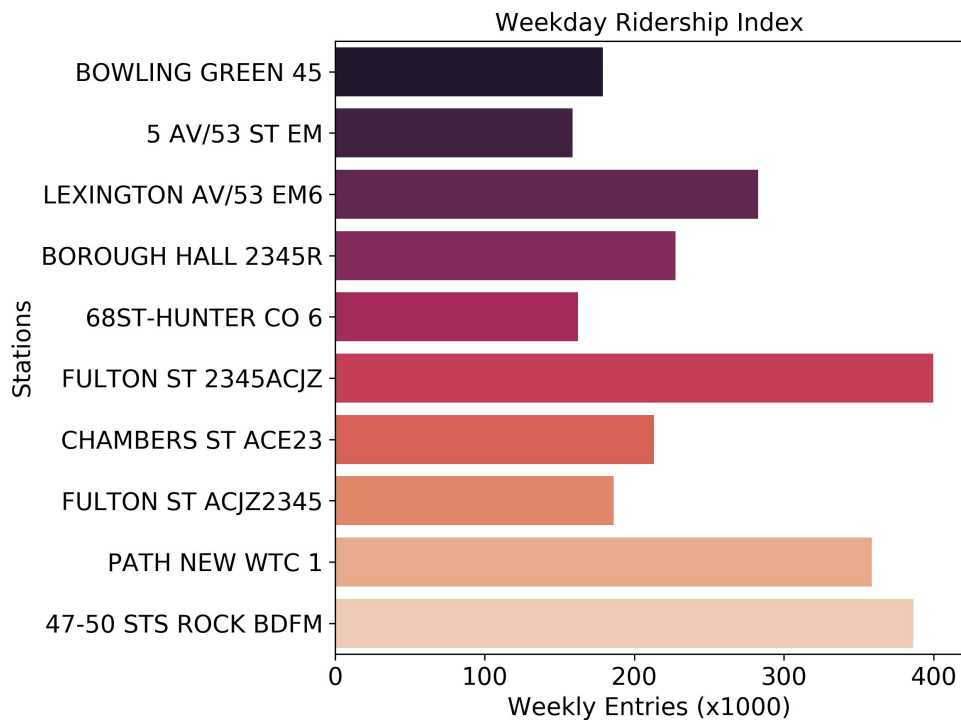|	
1	**Initial Question**
2	**Assumptions & Cleaning Data**
3	**EDA outcome**
4	**Further investigation**

Monthly Traffic
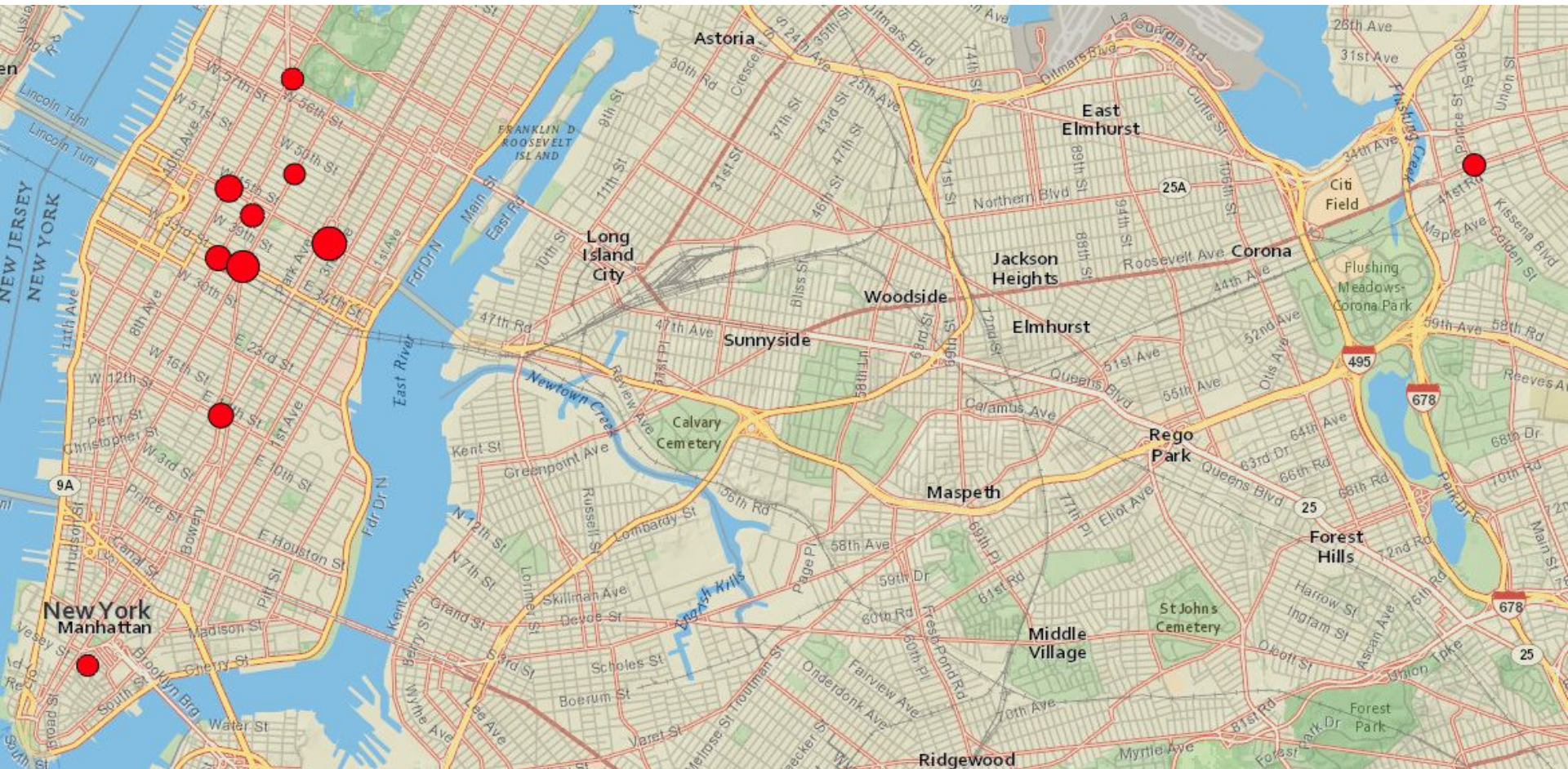
Weekly traffic

Top Stations that are busy at least 5 out of 7 days a week

Weekday Ridership Index
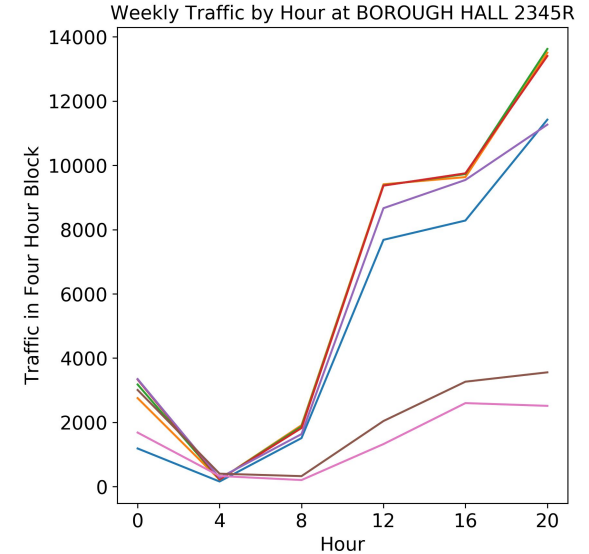
Ratio of *Average* Weekday Riders/Weekend Riders
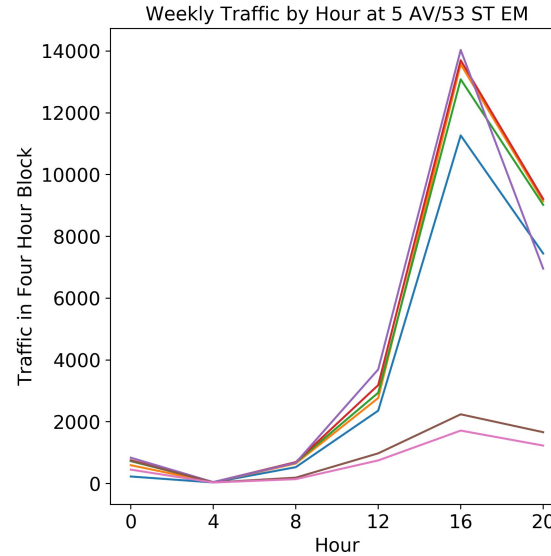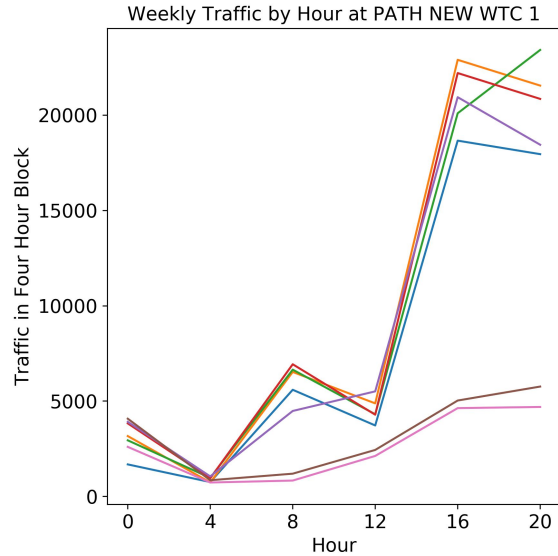

Weekday Ridership Index

5:1

3:1

Daily Traffic

- Stations show slightly different ridership curves throughout the day.
- Traffic peaks in the afternoon or evening.
- Weekdays have significantly more traffic than weekends.

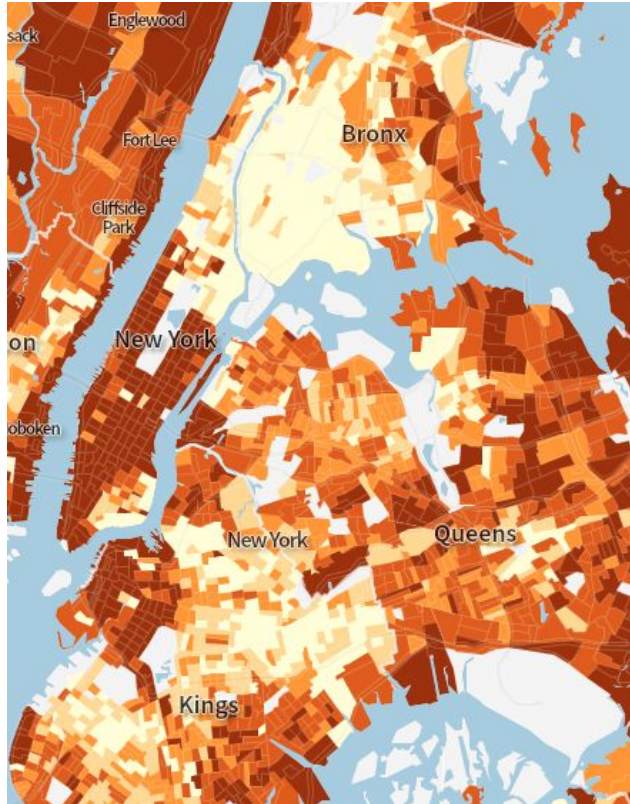# Final List & Selection Standards

Primary considerations:

- Traffic Volume
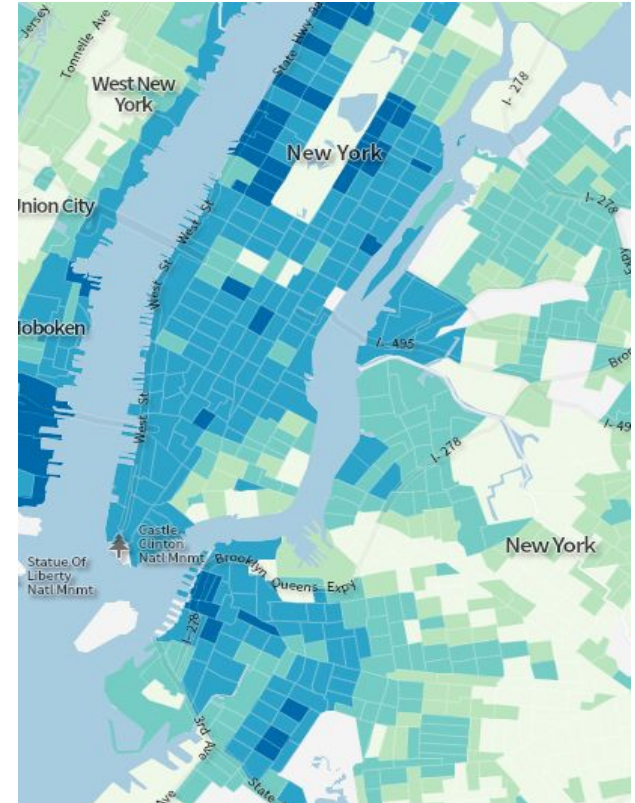- Commuter vs Tourist

Additional considerations:

- Borough representation
- Income demographics
- Education level

Other Selection Criteria

Income

Master's Degrees

Key: Darker colors represent higher percentage of population

Final Recommendations

Fifth Ave - 53rd State Station
47th-50th Street Rockefeller Center
Grand Central - 42nd Street
34th Street - Penn Station
Fulton Street Station
World Trade Center Station
Bowling Green
Borough Hall - Court Street Station
Jackson Heights - Roosevelt Ave
Third Ave - 149th Street Station

Schedule Guide

Visualized Mean Daily Traffic
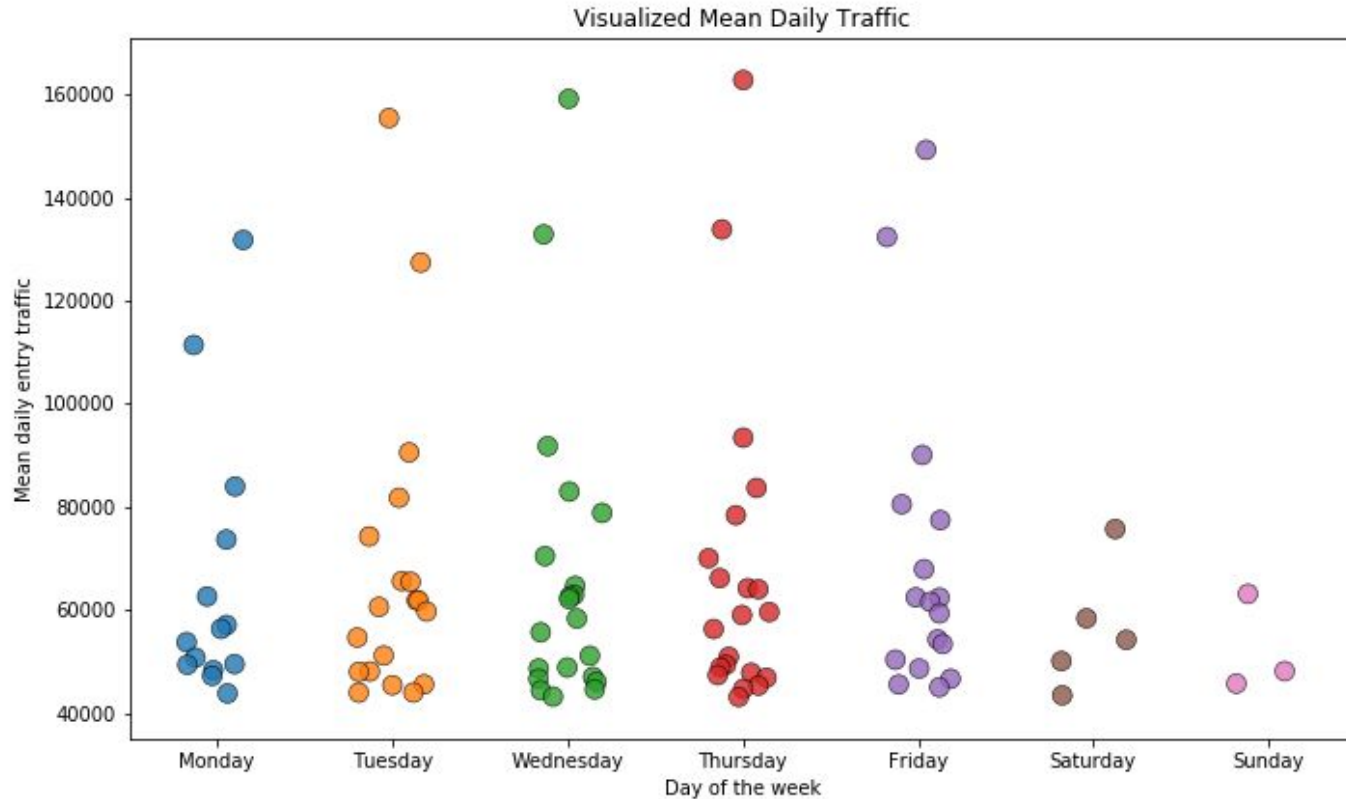
More detailed

time instruction

coming in further

investigation

Agenda

1   Initial Question

2   Assumptions & Cleaning Data

3   EDA outcome

4   Further investigation

Further investigation

Recommended next steps

- Refine selection - integrate census data
- Optimize data collection - integrate technology
- Targeted outreach
  - College / University Departments
  - Technical High Schools
  - Youth programs with tech focus

Where can we use more data?

- Four hour blocks of data
- Only looking at entries
- Method of payment unknown (limiting factor of dataset)

Thank You & Questions