

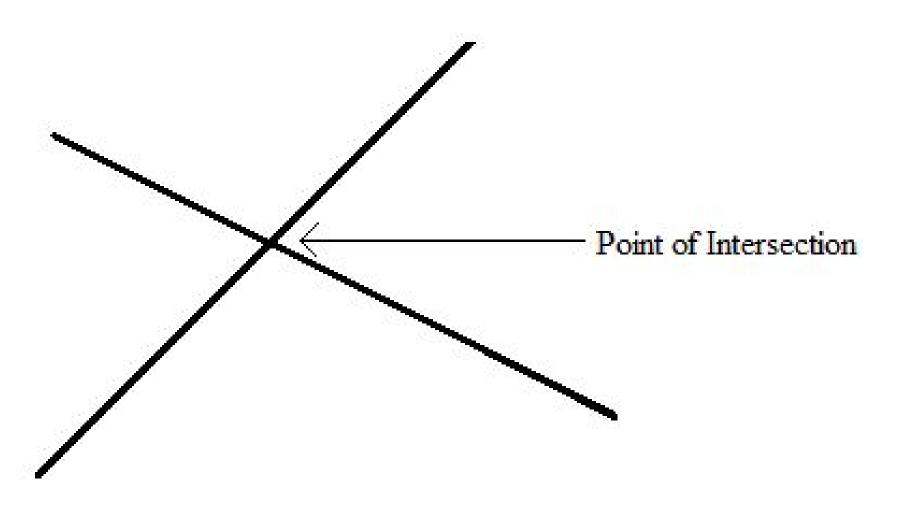


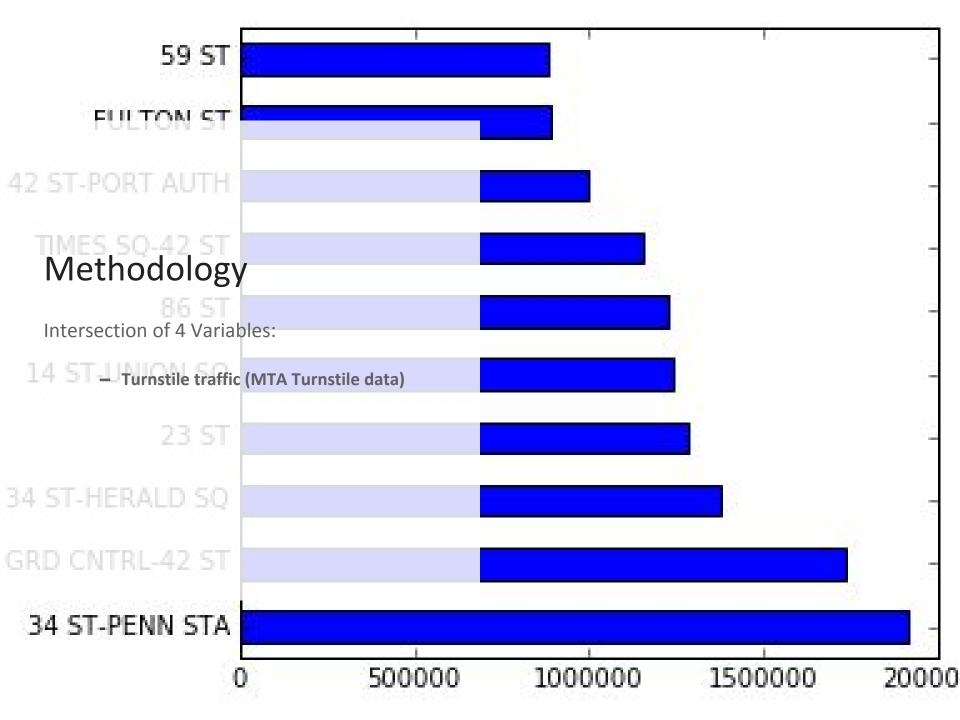
### **Executive Summary**

Determine optimal subway stations to deploy street teams in order to maximize:

- Collection of emails
  - People most likely to attend annual gala
  - People most likely to donate to the WTWY cause
- Build awareness and reach

# Methodology: Intersection of 4 Variables





### Methodology

#### Intersection of 4 Variables:

- Turnstile traffic (MTA Turnstile data)
- Income by Zip Codes (IRS)

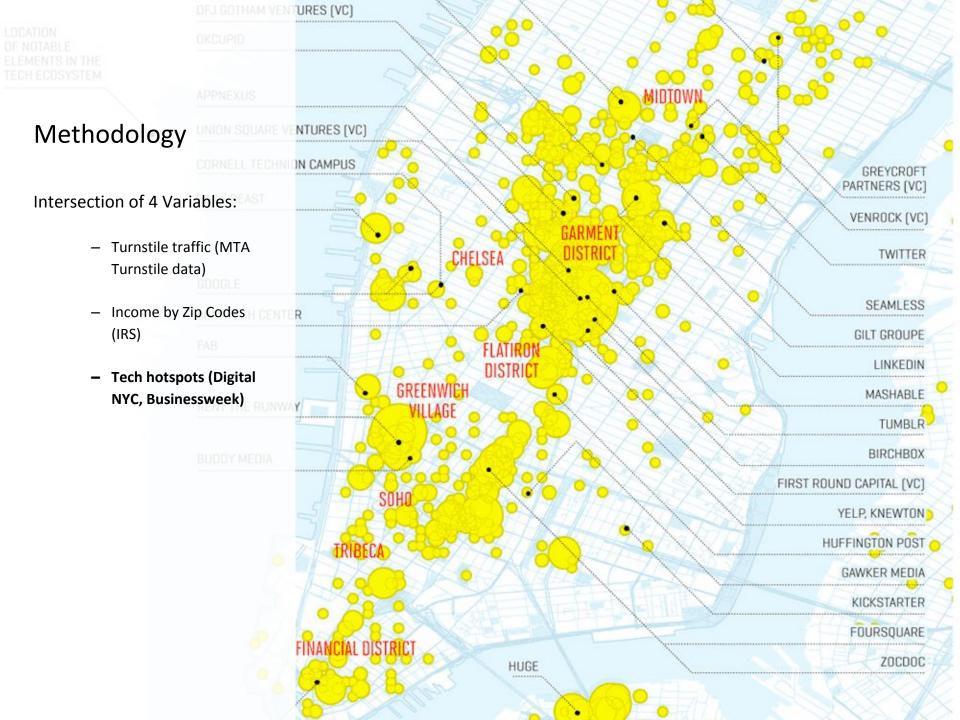
Bay Midtown Murray South Hill Tudor City Hunters Chelsea Point Flatiron Gramercy Stuyvesant Town Median East **Household Income** Greenwich Village Village < \$29,000 \$29,000 - \$50,000 Noho \$50,000 - \$73,000 oho \$73,000 - \$113,000 Little Lower Italy > \$113,000 East Side Chinatown

Turtle

Tribeca

Center

Battery Park City



### Methodology

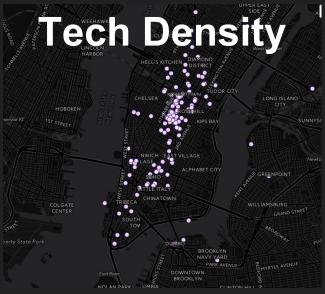
Intersection of 4 Variables:

- Turnstile traffic (MTA Turnstile data)
- Income by Zip Codes (IRS)
- Tech hotspots (Digital NYC, Businessweek)
- Mood Sentiment (Wyst.it)

Theater Island District Midtown East Midtown Garment District Koreatown Murray Hill Flatiron Kips Bay Union Gramercy Sq. Stuy nwich Town age East Village Greenpoint VoHo Alphabet City **UTa** Bowery Lower Williamsburg East Bushwick Chinatown Side East Williamsburg

# Analysis: Combined Data









Sources: CartoDB, OSM, IRS, MTA, Business Week

# Looking Forward: Learn then Do

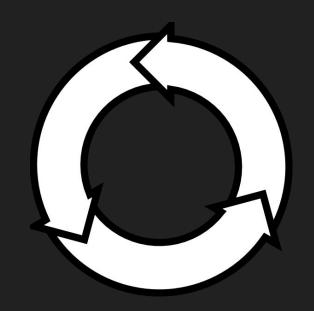
### Learn, iteration 1

#### Look at data

- Weekly traffic counts
- Freq. of top tier income tax payers
- Proximity to tech startups

#### Make a plan

- Identify top 3 clusters
- Deploy teams during high traffic periods



### Do, iteration 1

Deploy field teams to clusters

- Collect signatures / emails at and near subway stations
- Identify best stations from a signatures standpoint

Aggregate and assimilate data

•••



Use results from each Learn-Do iteration to conduct other, deeper, & more targeted analyses to optimize field operations

### Conclusion

- Preliminary hypothesis: Teams will collect more signatures, donations, and build more awareness by deploying teams at or near:
  - High traffic subway stations
  - Affluent neighborhoods
  - Tech density
  - Positive Sentiment
- Based on these criteria, we think Flatiron, Downtown Brooklyn and East Village neighborhoods are the top three areas
- We will refine location evaluation criteria and top deployment areas based on the results of each successive field deployment

# Future Investigations I

- Clean up & improvement of existing data
- More Data -> More Robustness
- Quantitative optimization:
  - Traffic within high income areas
  - Traffic proximal to tech hotspots
  - Formalize Tradeoffs across 4 variables
- Geospatial Analysis

## Future Investigations II

- Time analysis (Days, hours)
- Mood per hour (twitter/ instagram)
- Time Series Analysis -> Forward Projections
- In-Person Polling (de facto) vs Alternate
   Outreaches
  - 'Hijacking' other Meetup groups (in person, email)
  - Contacting the social / outreach director at...
    - Tech companies (Google, Facebook, Metis, etc...)
    - Incubators / co-work spaces (wework ...)
    - Universities (Columbia, NYU, Cornell Tech, etc...)



# **Appendix**



### **Challenges**

- Managing scope
- Merging data sources
- Parsing HTML w/regular expressions
- Exceeding API query limitations

