PROJECT BENSON

Collin, Morgan, Aleksandra, Chris

OBJECTIVE

Identify optimal days, locations, and conditions to deploy the Women Tech Women Yes subway canvassing team in order to maximize ticket purchases for the gala.

ASSUMPTIONS

Gala on Sep 26, 2017

Attendees include tech workers, wealthy individuals and philanthropists.

Subway ridership includes above.

Geographical distribution of women is even.

WTWY has a team of 50 volunteers, working twice per week

APPROACH

1) Select 20 Busiest Stations by Ridership
2) Identify busiest days of the week by Ridership
3) Cross Reference busiest stations w/income and donation data
4) Investigate effect of weather on priority stations

DATA SOURCES

NYC MTA: Subway station name and date time stamped counters for entry and exits through the turnstiles, Jun 2016 - Sep 2016.

NOAA: Daily Precipitation in New York, Jun 2016 - Sep 2016

U.S. Census: Median household income in New York, 2010

IRS: Total tax deduction in New York, 2014

SURPRISE

It wasn't as straightforward as we'd hoped

EDA IS CHALLENGING

- Turnstiles
  - Takes 4 columns to uniquely specify
  - Audits reset at different times
  - Crazy turnstiles

- Station names
  - Inconsistent
  - Not unique
  - Mapping to location

FIRST QUESTION

WHICH STATIONS HAVE
THE MOST RIDERS?

Grand Central-42nd St

34 St - Herald Sq

Times Square-42nd St

14 St - Union Sq

Fulton St

34 St - Penn Station

42 St - Port Authority Bus Terminal

59 St - Columbus Circle

59 St - Lex Ave

86 St

47-50 Sts - Rockefeller Ctr

Flushing - Main St

34 St - Penn Station

Jackson Hts - Roosevelt Av

42 St - Bryant Pk

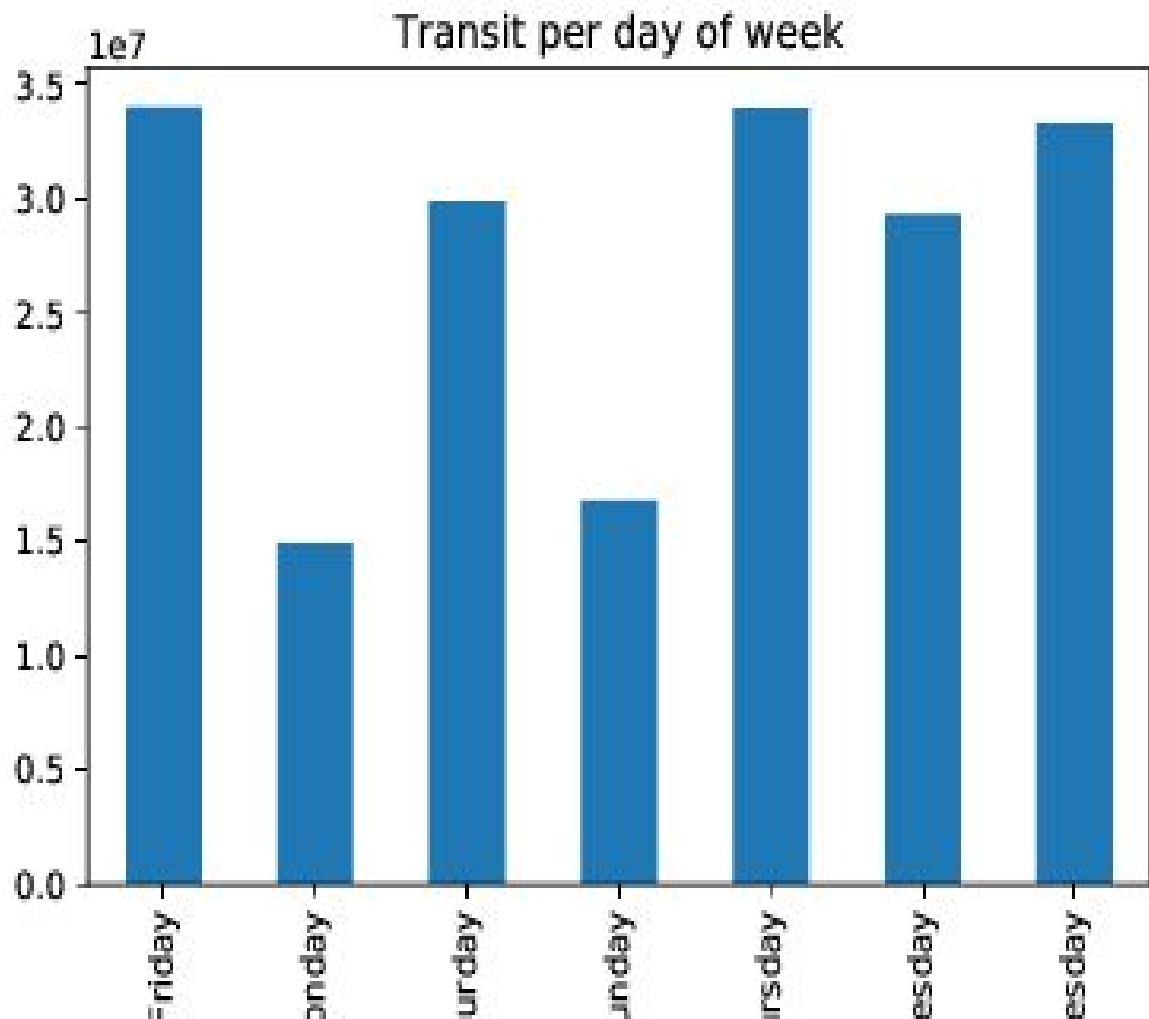Atlantic Av - Barclays Ctr

Canal St

Lexington Av/53 St

96 St

14 St - 7th Ave
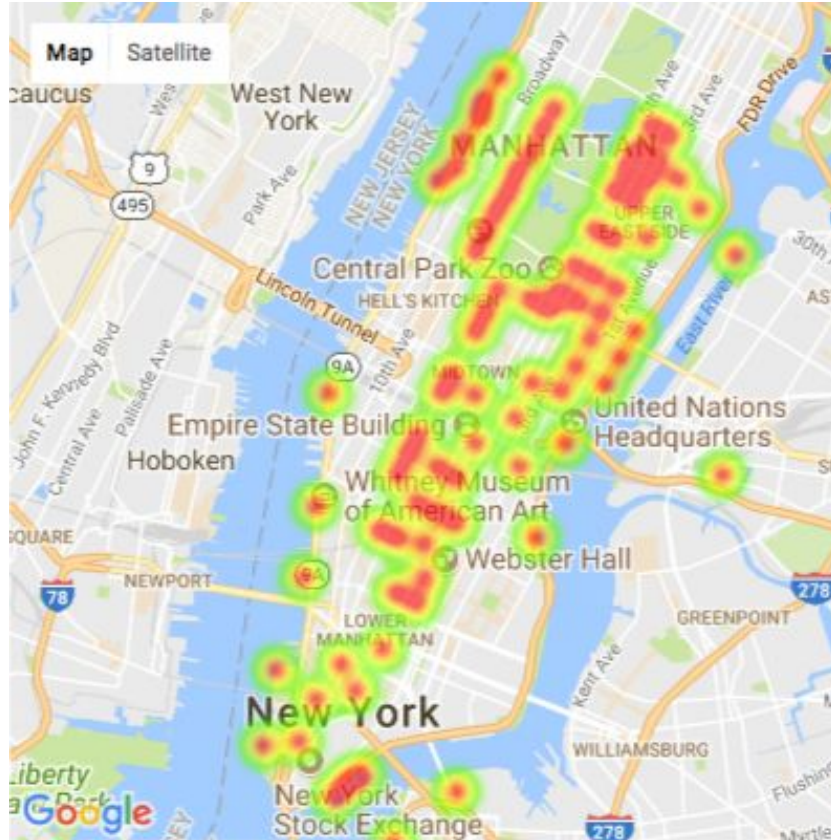
SECOND QUESTION

WHICH DAYS OF THE WEEK ARE BUSIEST?

Recommendation: avoid Sunday and Monday
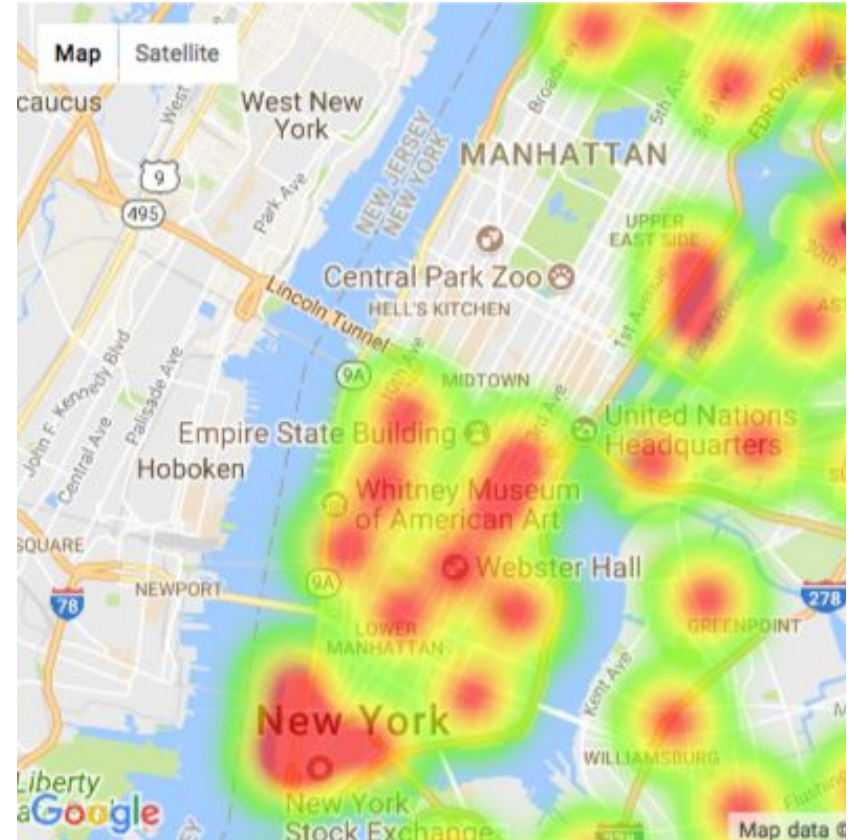
THIRD QUESTION

WHICH STATIONS ARE LOCATED
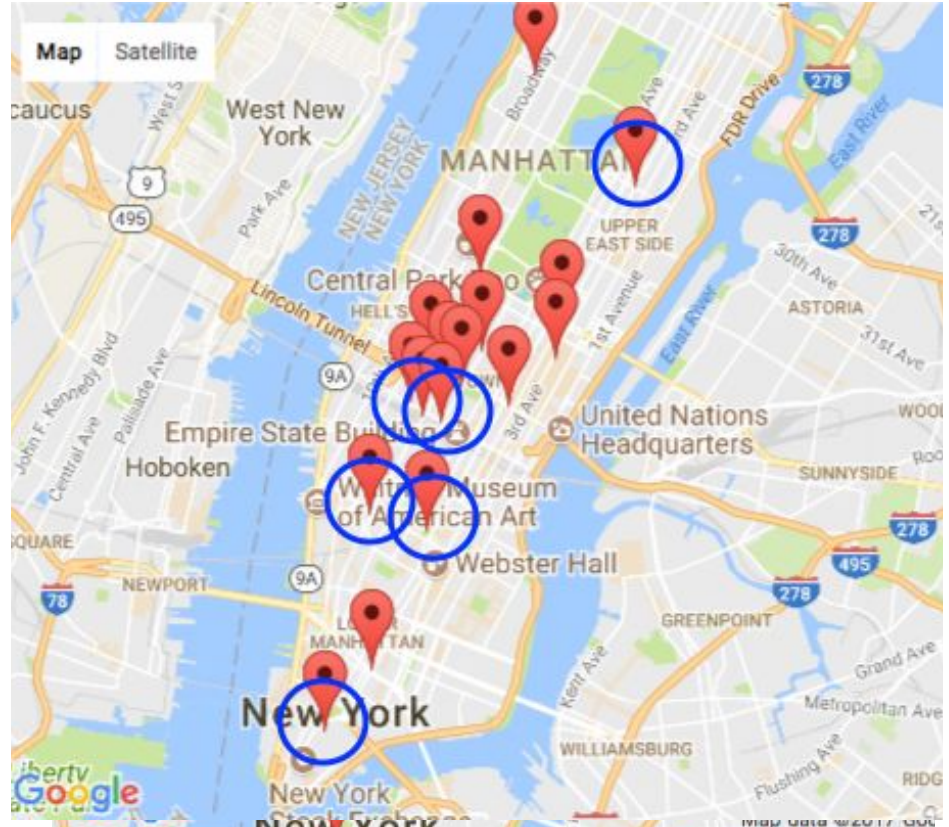IN WEALTHY/PHILANTHROPIC AREAS?

Household Income

Charitable Gifts (Est.)

RECOMMENDATIONS

- Fulton
- Union Square
- 14th Street and 7th Ave
- Herald Square
- Penn Station
- 86th Street
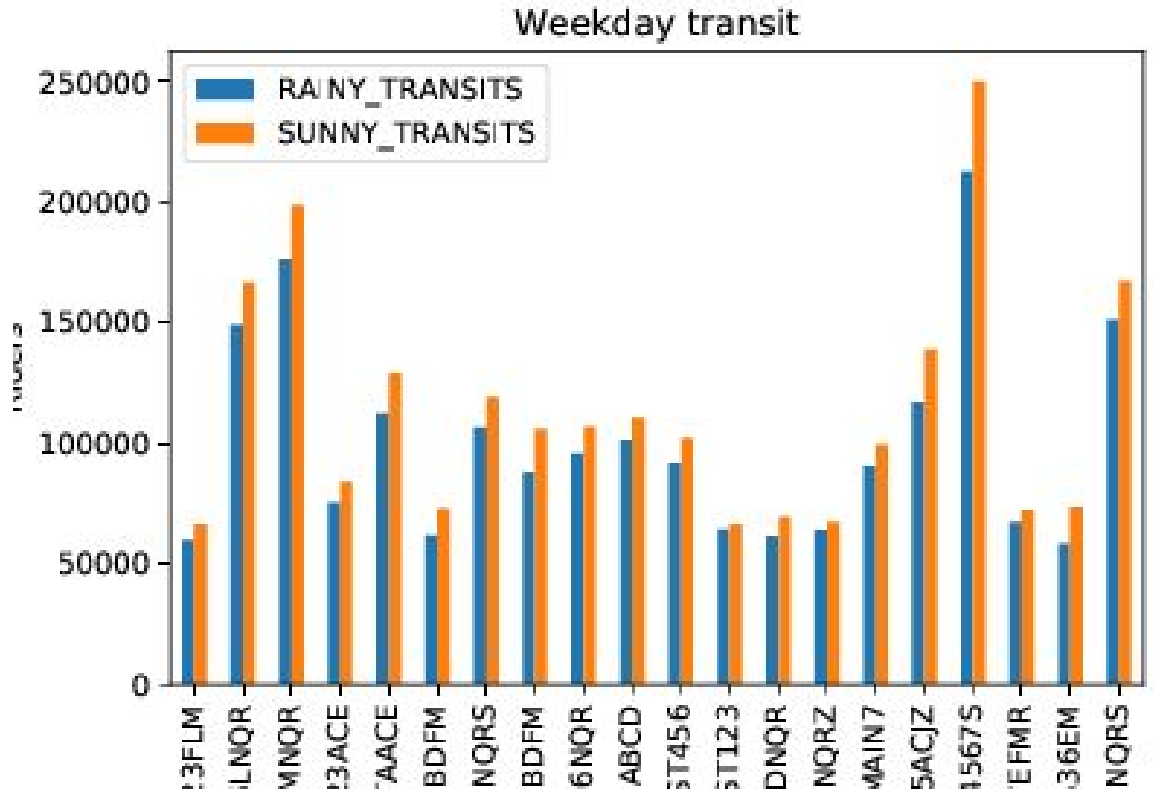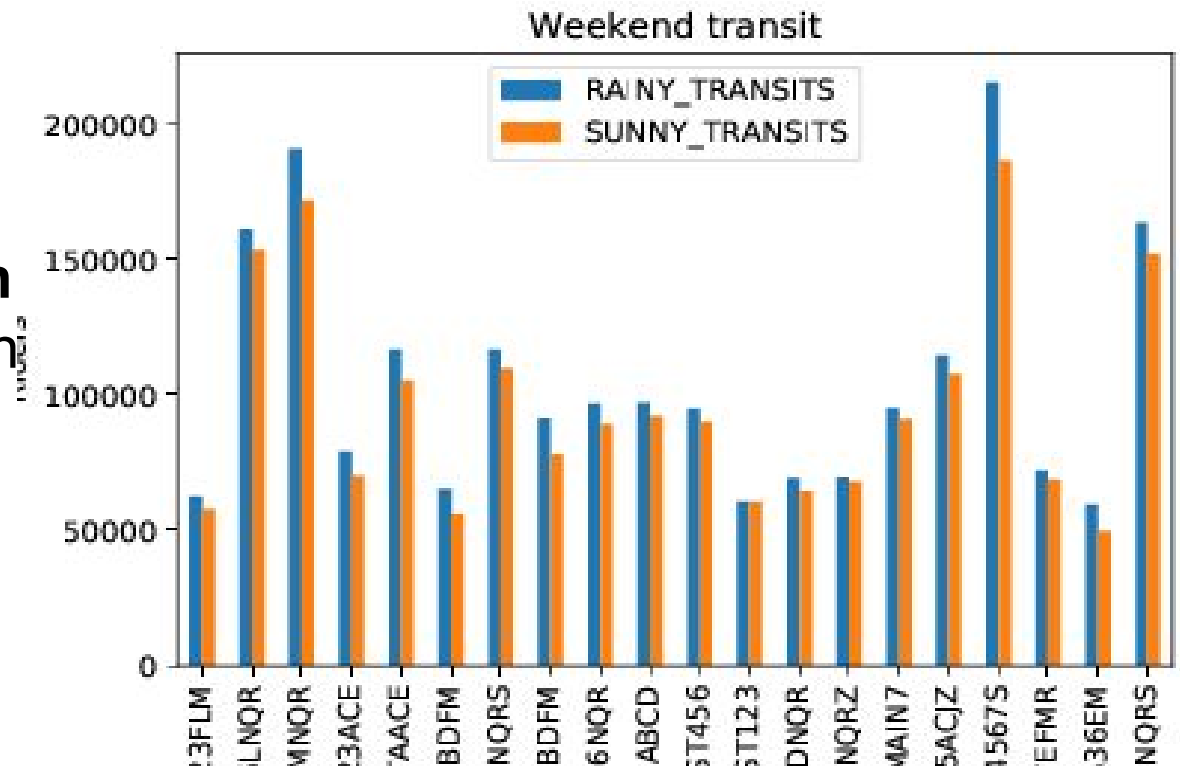- Barclay's Center
- Lexington Ave - 53 St

FOURTH QUESTION

DOES WEATHER HAVE AN IMPACT ON SUBWAY RIDERSHIP

Ridership is **higher on Sunny Weekdays** than Rainy Weekdays

Weekday transit

Legend:
- RAINY_TRANSITS
- SUNNY_TRANSITS

Ridership is **higher on Rainy Weekends** than Sunny Weekends

Weekend transit

LESSONS LEARNED

Spend time understanding what the customer really wants to do with the data science recommendations.

Expect to spend 80% of the time cleaning

Be comfortable with the tools. Still getting up to speed.

Teamwork! Strengths complement each other. Confer, divide and conquer.

Simple is Elegant

NEXT STEPS

- Improve EDA
  - -Link Location Data to Turnstile Data
  - -Target Subway Station w Bottlenecks
  - -A/B Testing for Weather
- External Data
  - -Demographics (Occupation)
  - -Donors dataset
- ML:
  - -Scoring for matching characters
  - -Classification vs. regression

Source:

Data:

[MTA: Turnstile](#)

[MTA: Subway Stations](#)

[US Census: Median Household Income](#)

[Internal Revenue Service: Tax Stats](#)