EXPLORATORY ANALYSIS

LOCATION RECOMMENDATIONS

Steven L Truong

Friday, 02/04/2021

Four stages of learning Data Science

INTRODUCTION

❖**MOTIVATION**:

- Greenwood LLC wants to expand their business and open the authentic Vietnamese restaurant in NYC.

- They want to leverage Data Science and Analytics skills set to optimize their chance of success.

❖ **OBJECTIVES**:

- Determine the top stations with the most traffic.

- Conclude the best location based on the volume of traffic and surrounding venues.
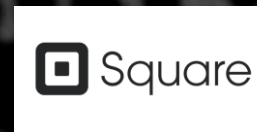
❖ **GOALS**:
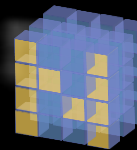
- Logically and analytically present the results.

❖ **Data source**:

  ❖ MTA Turnstile Data

  ❖ Square API

❖ **Tools**:

  ❖ Numpy and Pandas

  ❖ Matplotlib and Seaborn

  ❖ Folium, wordcloud, requests

METHODOLOGY

Exploration

- Pull turnstile data from the source (MTA website)
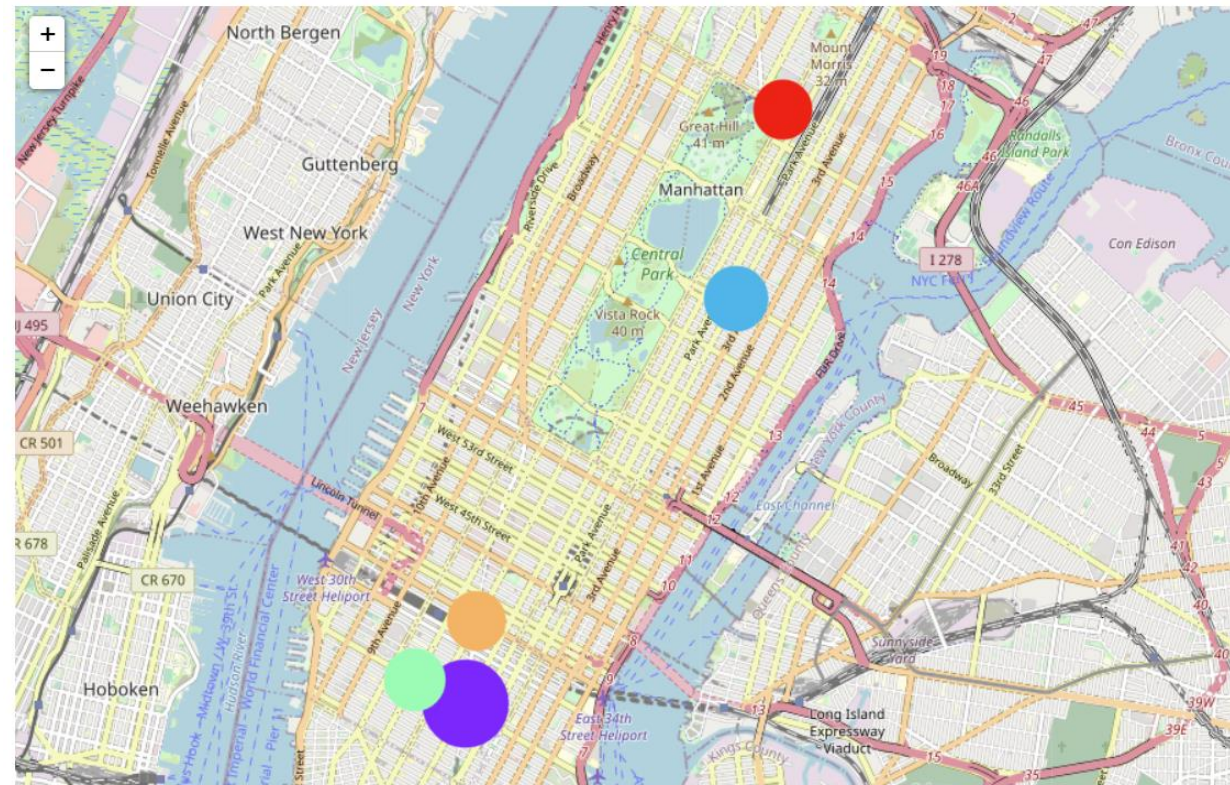- Use Square API to pull data in json format into local machine for further analysis
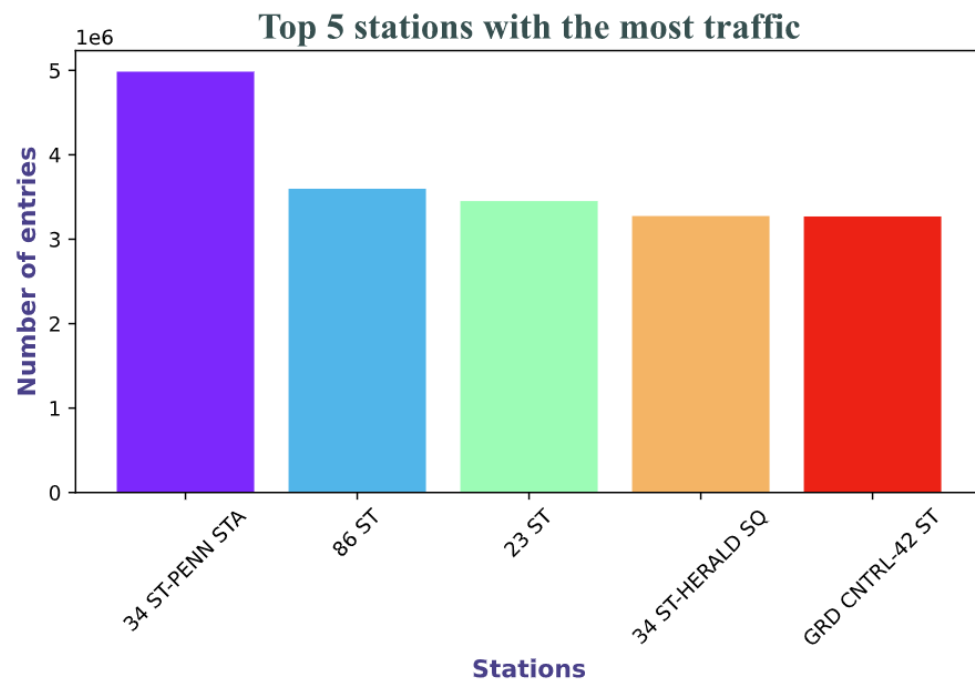
Preparation

- Clean raw data
- Converse data from json to regular pandas dataframe

Visualization

- Create graphs (time series, bar charts)
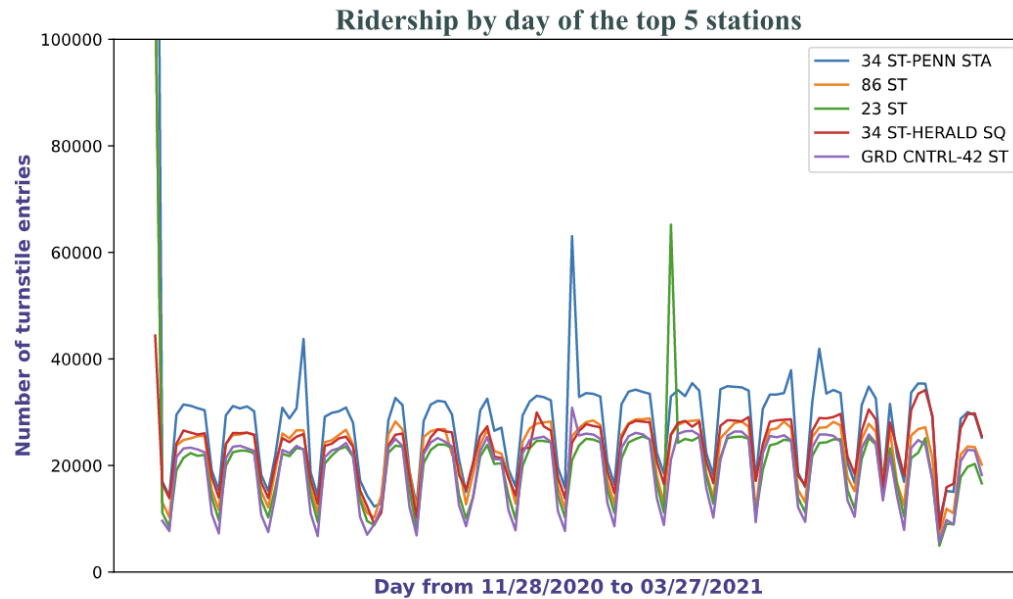- Generate NYC map with traffic density and venues wordclouds

RESULTS

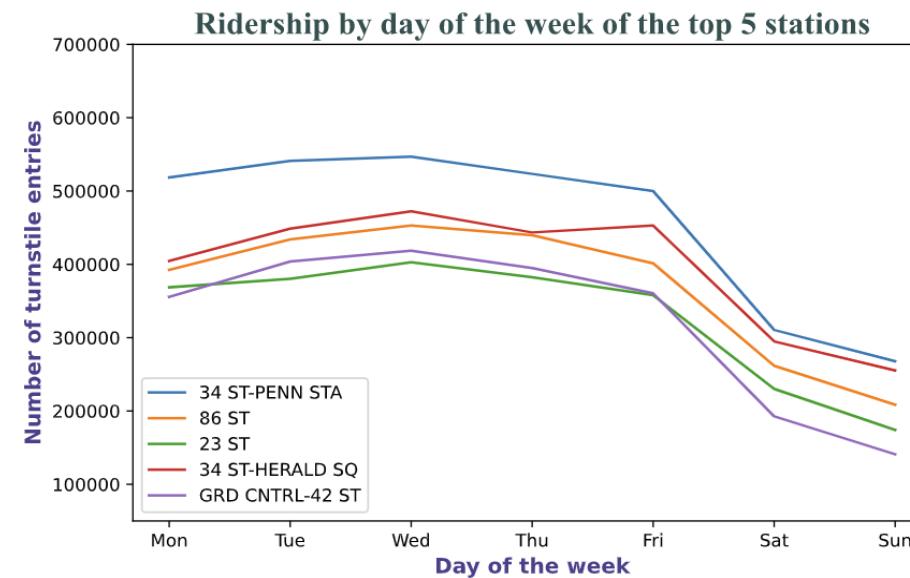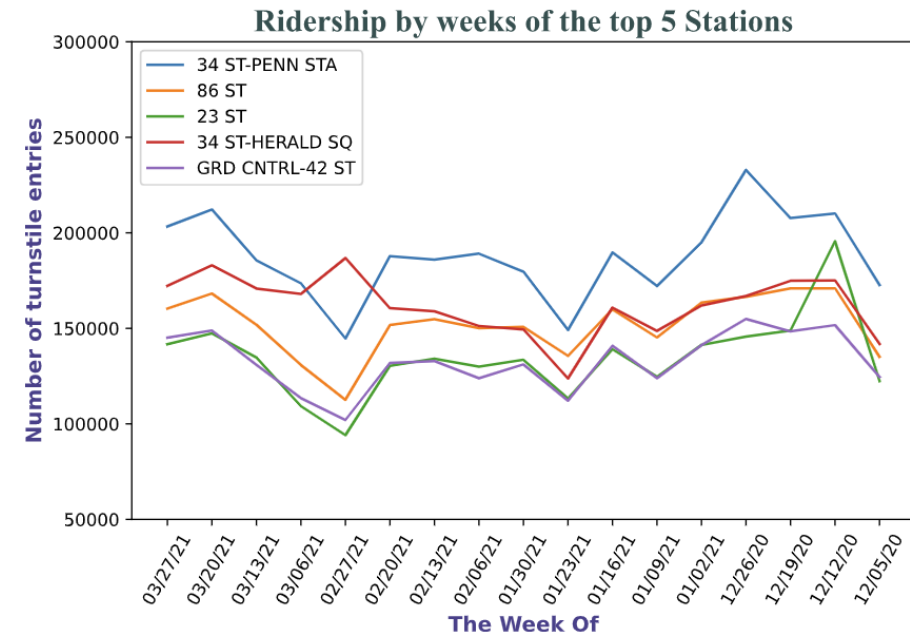
Top 5 stations with the most traffic

❖ **Top 5 stations with the most traffic:**

1. 34 St - Penn Station (near Korean town - Manhattan)

2. 86 St Station (Upper East Side - Yorkville)

3. 23 St Station (near Korean town - Manhattan)

4. 34 St - Herald Square Station (near Korean town - Manhattan)

5. Grand Central - 42 St Station (near Korean town - Manhattan)

RESULTS

Ridership by day of the top 5 stations

Ridership by weeks of the top 5 Stations

Ridership by day of the week of the top 5 stations

What about each neighborhood's nearby venues? Are there already lots of Vietnamese restaurants? Does the demand still need the supply?

DISCUSSION

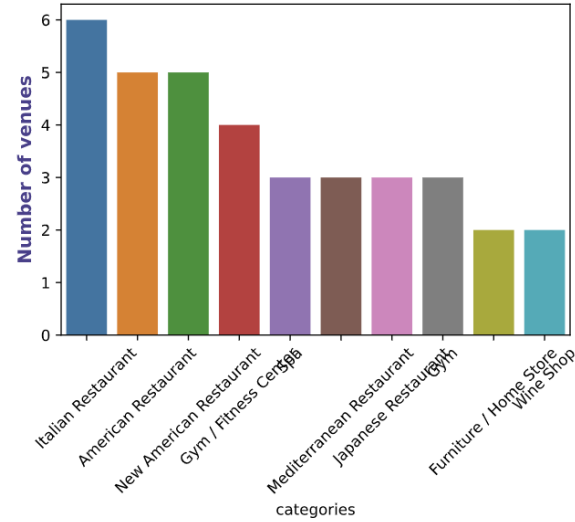Top 10 most common venues around each neighborhood

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	23 St Station, New York, NY	Gym / Fitness Center	Hotel	American Restaurant	Coffee Shop	Japanese Restaurant	Yoga Studio	New American Restaurant	Café	Lounge	Miscellaneous Shop
1	34 St - Penn Station, New York, NY	American Restaurant	New American Restaurant	Gym / Fitness Center	Yoga Studio	Japanese Restaurant	Hotel	Spa	Café	Italian Restaurant	Cycle Studio
2	34 St-Herald Sq, New York, NY	Korean Restaurant	Gym / Fitness Center	Hotel	Coffee Shop	Theater	Spa	American Restaurant	Pizza Place	Boxing Gym	Cycle Studio
3	86th St Station, New York, NY	Italian Restaurant	Coffee Shop	Gym / Fitness Center	Exhibit	Ice Cream Shop	Art Museum	Bakery	Pizza Place	Bookstore	Yoga Studio
4	Grand Central-42 St Station, New York, NY	Mexican Restaurant	Coffee Shop	Park	Cocktail Bar	Bakery	Café	African Restaurant	Restaurant	Fountain	Wine Bar

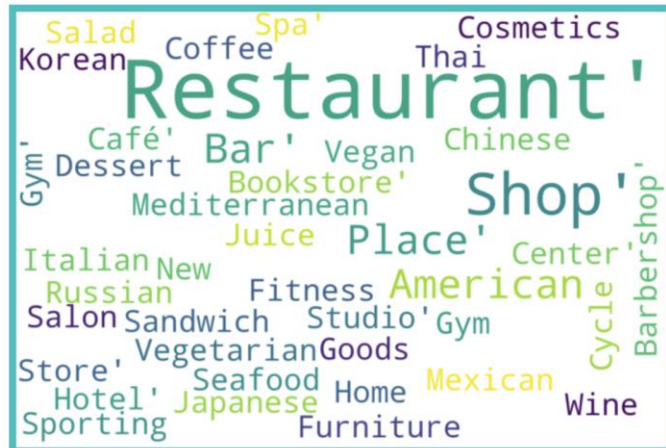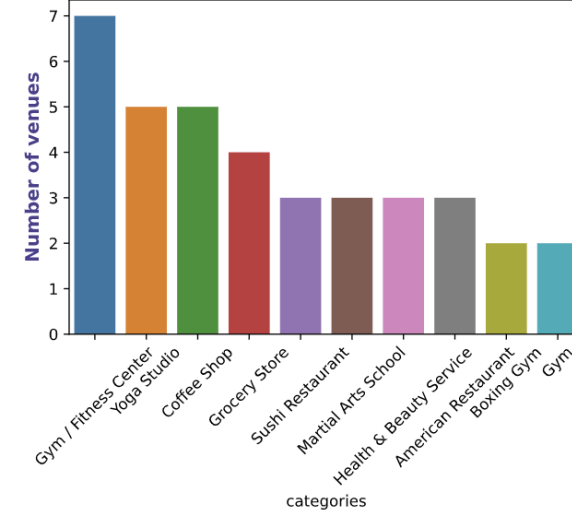Perhaps this is a bit hard to see, let's visualize them!

DISCUSSION

Nearby venues of 34 St – Penn station and 23 St station

Frequencies of venues around 34 St - Penn Station, New York, NY

Frequencies of venues around 23 St Station, New York, NY

CONCLUSION

- The 34 St – Penn Station and the 34 St – Herald Sq neighborhoods have many restaurants around them => not very suitable to open another one.

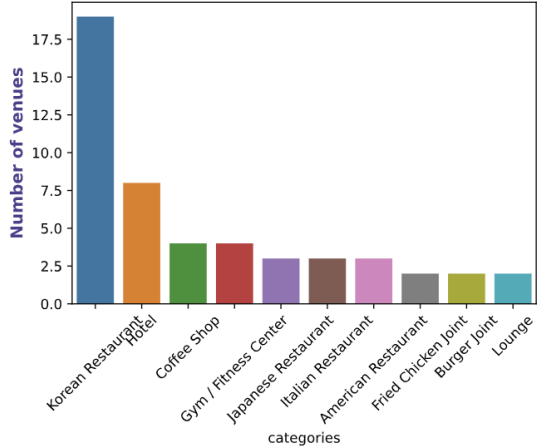- Recommendation: neighborhood around 23 st station.

FUTURE WORK

❖ Focus more on the neighborhood's demographic.

❖ Incorporate other data sources (such as crime rates or median income of certain population) to help shaping the results based on the business' targeted customers.

❖ Investigate customers' behaviors to see whether they are affected by covid.
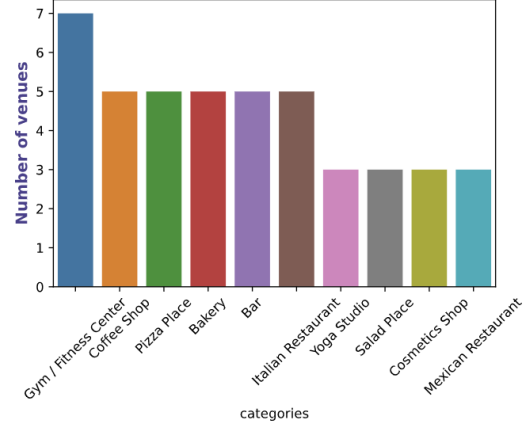
APPENDIX

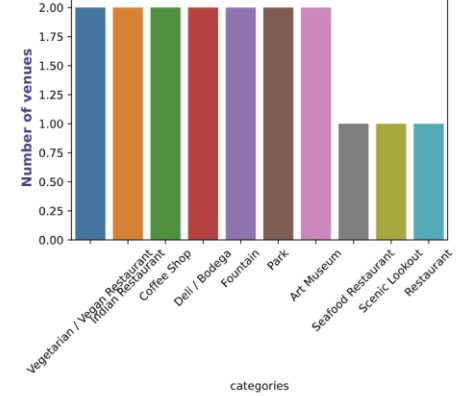❖ Bar plots for the other 3 top stations along with their word clouds

Frequencies of venues around 34 St-Herald Sq, New York, NY

Frequencies of venues around 86th St Station, New York, NY

Frequencies of venues around Grand Central-42 St Station, New York, NY

APPENDIX

❖ First few rows of **one hot encoding** the neighborhood for each station

	Neighborhood	African Restaurant	American Restaurant	Art Gallery	Art Museum	Austrian Restaurant	Bagel Shop	Bakery	Bar	Basketball Stadium	...	Tennis Court	Thai Restaurant	Theater	Toy / Game Store	Vegetarian / Vegan Restaurant	
0	34 St - Penn Station, New York, NY	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0
1	34 St - Penn Station, New York, NY	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0
2	34 St - Penn Station, New York, NY	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0
3	34 St - Penn Station, New York, NY	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0
4	34 St - Penn Station, New York, NY	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0

5 rows × 124 columns

Documentation on how to use Square API to explore nearby venues given the neighborhood's coordinates (latitude and longitude): https://developer.foursquare.com/docs/api-reference/venues/explore/

THANK YOU

STEVEN L TRUONG

https://github.com/luongtruong77

tqluong77@gmail.com

QUESTIONS?