



Applied Data Science

Capstone Project

Business Problem

- Recommendation for office location for a distributor of sea food to Chinese restaurants in New York (Manhattan)
- Target Audience for this solution are Distributors of sea food

Constraints :

- Distributor is selling perishable sea food
- Wants to be close to his customers to deliver on an as needed basis

Data used

- Used data from the internet which is readily available
 - data from geo.nyu.edu
 - Four square for restaurant data

Exploring the Data set

- Geo data from geo.nyu.edu
 - Latitude and longitude data for 40 neighborhoods in Manhattan
 - This data is used to get the top food venue data from Foursquare
- Foursquare Data
 - Explore API used with “section = food” filter to get data relevant for our business problem

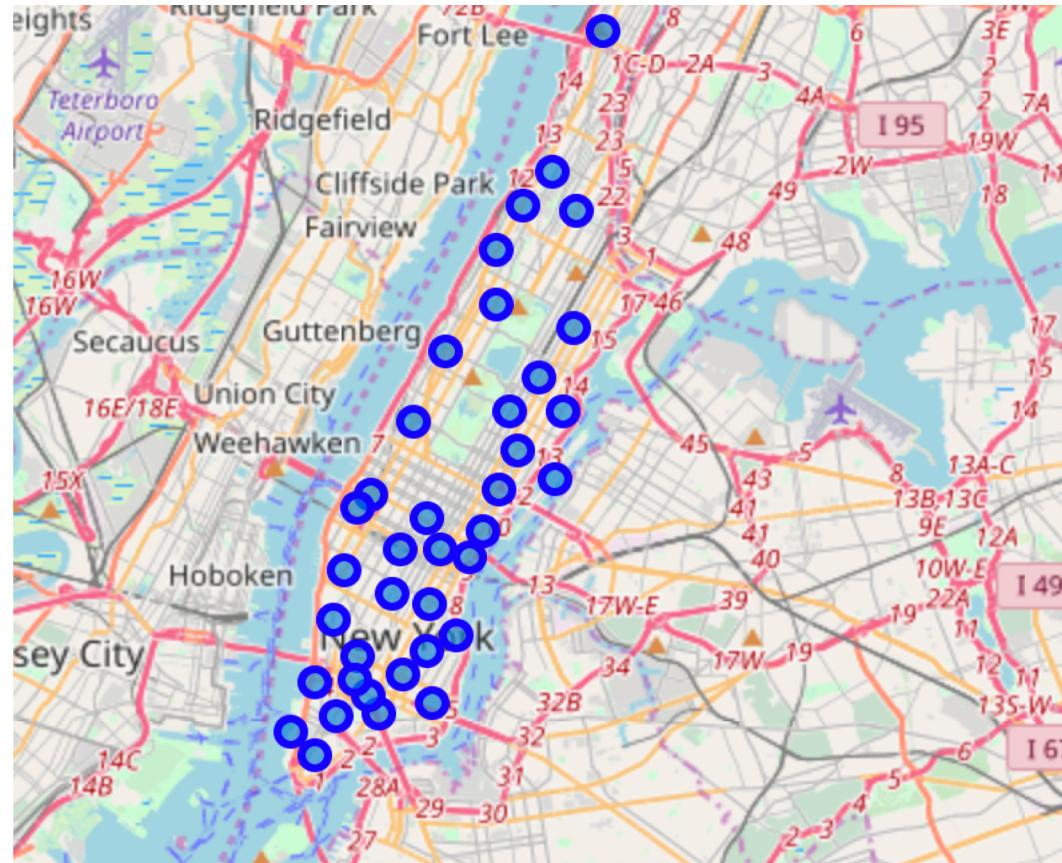
Analyzing the data

- Group and classify the data into top common venues for every neighborhood
- Use one hot encoding to encode the data so it can be clustered and analyzed

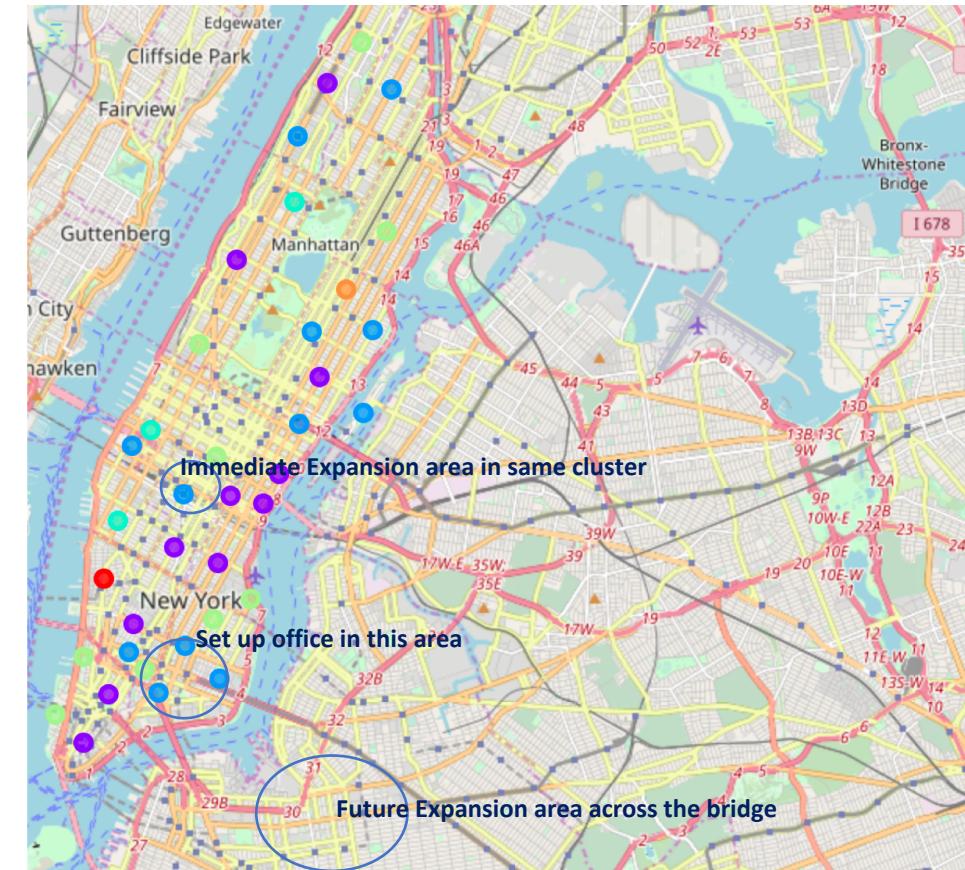
Clustering Neighborhoods

- Run K-means to cluster the neighborhoods into 6 clusters
- Various cluster numbers are used to do multiple runs to get to the optimal cluster number

Unanalyzed data



Clustered and Analyzed data



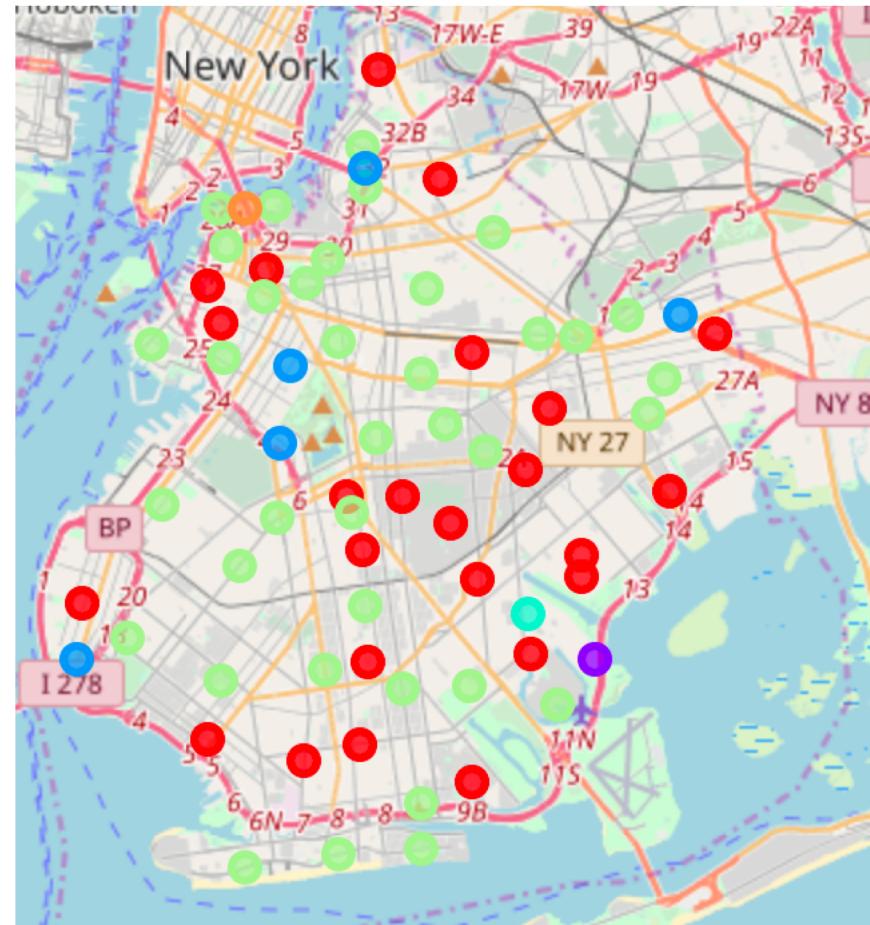
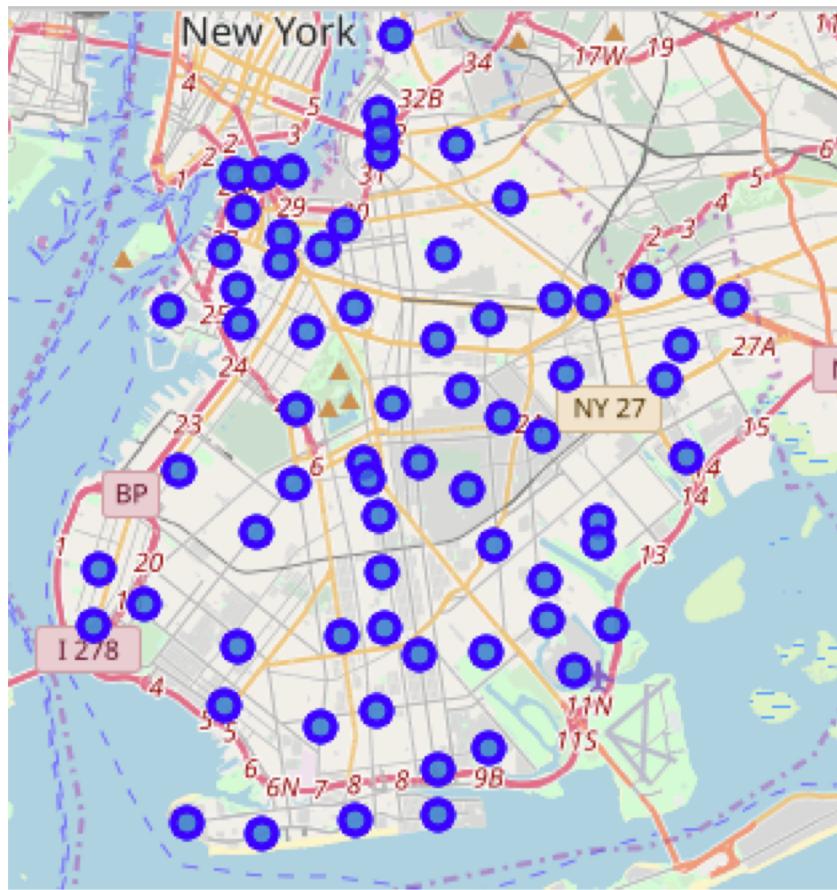
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
1	Chinatown	Chinese Restaurant	Bakery	Vietnamese Restaurant	Dim Sum Restaurant
20	Lower East Side	Pizza Place	Chinese Restaurant	Deli / Bodega	Bakery

Conclusion

- Based on the results it is clear that the distributor should setup office close to China town and Lower East side
- This is based on popularity of Chinese restaurants in that cluster
- As a immediate expansion they can look at midtown south (in the same cluster)
- They can do future expansion to Brooklyn which is across the bridge on the other side

Expansion plan for the business –
Additional Analysis beyond the original Business
problem

Future Expansion plans - Brooklyn



- Run K-means to cluster the neighborhoods into 6 clusters
- Cluster 1 has Chinese restaurants in common venues in multiple neighborhoods. Perfect for the expansion plan in the cluster

Future Expansion plans - Brooklyn Cluster 1 data

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	Bay Ridge	Pizza Place	Italian Restaurant	Chinese Restaurant Bagel Shop
3	Greenpoint	Pizza Place	Deli / Bodega	Chinese Restaurant Restaurant
4	Gravesend	Chinese Restaurant	Italian Restaurant	Pizza Place Deli / Bodega
7	Manhattan Terrace	Pizza Place	Bakery	Donut Shop Chinese Restaurant
8	Flatbush	Deli / Bodega	Caribbean Restaurant	Mexican Restaurant Bakery
10	East Flatbush	Deli / Bodega	Chinese Restaurant	Fast Food Restaurant Caribbean Restaurant
14	Brownsville	Restaurant	Chinese Restaurant	Deli / Bodega Fried Chicken Joint
19	Cobble Hill	Deli / Bodega	Pizza Place	Italian Restaurant Bakery
20	Carroll Gardens	Italian Restaurant	Deli / Bodega	Pizza Place Bakery
27	Starrett City	American Restaurant	Caribbean Restaurant	Pizza Place Chinese Restaurant
28	Canarsie	Deli / Bodega	Food	Asian Restaurant Thai Restaurant
29	Flatlands	Fast Food Restaurant	Caribbean Restaurant	Restaurant Asian Restaurant
33	Bath Beach	Chinese Restaurant	Pizza Place	Donut Shop Fast Food Restaurant
36	Gerritsen Beach	Pizza Place	American Restaurant	Food Court Deli / Bodega
40	Downtown	Pizza Place	Food Truck	Sandwich Place Burger Joint
44	City Line	Food Truck	Donut Shop	Chinese Restaurant Bakery
47	Prospect Park South	Caribbean Restaurant	Pizza Place	Deli / Bodega Fast Food Restaurant
49	East Williamsburg	Deli / Bodega	Mexican Restaurant	Café Bakery
57	Remsen Village	Caribbean Restaurant	Sandwich Place	Fast Food Restaurant Deli / Bodega
59	Paerdegat Basin	Food	Deli / Bodega	Asian Restaurant Chinese Restaurant
60	Mill Basin	Chinese Restaurant	Pizza Place	Bagel Shop Japanese Restaurant
63	Weeksville	Deli / Bodega	Chinese Restaurant	Café American Restaurant
66	Homecrest	Donut Shop	Chinese Restaurant	Café Mexican Restaurant
69	Erasmus	Caribbean Restaurant	Pizza Place	Fast Food Restaurant Donut Shop

Expansion Plan

- Based on the results it is clear that the distributor should setup office close to China town and Lower East side
- This is based on popularity of Chinese restaurants in that cluster
- As a immediate expansion they can look at midtown south (in the same cluster)
- They can do future expansion to Brooklyn which is across the bridge on the other side

Assumptions / Limitations

- Four square data has limitation on the number of free API calls
 - Analyzed Brooklyn and other data from Four square over multiple days because of the above limitation.
 - Use of paid account should be explored to get all the data in the same run.