APPLIED DATA SCIENCE CAPSTONE

THE BATTLE OF THE NEIGHBORHOODS

BY

ALTAF JAMALIA

Contents

1. Introduction: Business Problem

Mr. John currently stays in **Downtown Toronto** with his family and now has a good job opportunity in the **United States of America**. He needs to move to **New York City** with his family.

Mr. John has never been to New York City and is looking to us to help him find a **good location within New York for him and his family to settle down**. Mr. John and his family are accustomed to the busy downtown life with good options and is looking for a similar location in New York.

We will use Data Science to generate a few most promising neighbourhoods based on this criteria so that Mr. John can take an informed decision and move to New York, USA.

2. Data

Mr. John has made it clear that his family enjoys the Downtown life, so based on this criteria, we are going look at the locations in New York which has:

- Good Downtown neighbourhood that has good Options in Restaurants (any type of restaurant), Shopping, Schooling

Following data sources will be needed to extract/generate the required information:

- New York Neighbourhood data available from ** https://cocl.us/new_york_dataset **

- Restaurants and their type, Shopping, Schools and other venues in the neighbourhoods of New York will be obtained using **Foursquare API**

- Coordinate of New York and the Boroughs in New York will be obtained using **GeoPy Library**

3. Methodology

The data shows us that there are 5 Boroughs in New York - **Manhattan Brooklyn Bronx Staten Island Queens**

Mr. John has not given any preference for any of the Boroughs, we are left with no other option but to do an analysis of all the Boroughs and its neighbourhoods. Our approach will be as follows:

- Do an analysis of one Borough and all its neighbourhood at a time

- Use Foursquare API to explore the neighbourhood

- Fetch the top 100 venues in a radius of 500 meters for each of these neighbourhoods

- Do an analysis and get the Top 10 most common venues in each of the neighbourhoods

- Use K-Means clustering to cluster the neighbourhoods

- The end result will be the clustered neighbourhoods in that borough with the top 10 common venues in it

This will be repeated for all the Boroughs and its neighbourhoods.

Mr. John's criteria was for a good neighbourhood that has **good options in restaurants (any type), Shopping, Schools etc.**

Once our analysis is complete, it will present a detailed view of all the neighbourhoods in each Borough along with its most common venue. Based on this Mr. John will be able to make an informed decision of which Borough and neighbourhood he would like to relocate to.

Let's start off with Manhattan and then go onto the other Boroughs

4. Analysis

We do an exploratory data analysis on each of the Boroughs – Manhattan, Staten Island, Brooklyn, Bronx and Queens and each of its neighbourhoods. Our goal is to analyse each neighbourhood and give enough insights to Mr. John about each neighbourhood so that he can take an informed decision about which neighbourhood that he would like to relocate to.

- We begin by copying the Neighbourhoods of each Borough to a separate dataframe

- We then use the Geopy Library and get the co-ordinates for each Borough

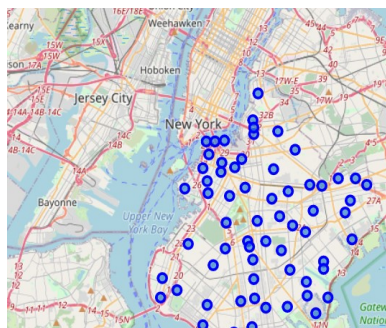- Here we create a folium map of each Borough and super impose only the neighbourhood data of that Borough
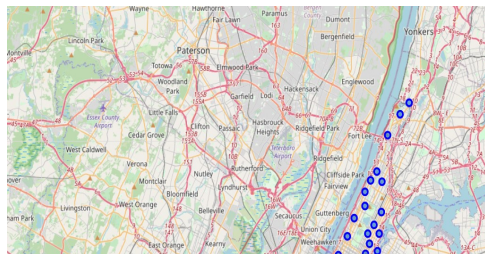
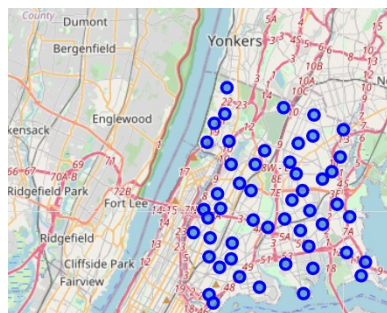Manhattan   Staten Island
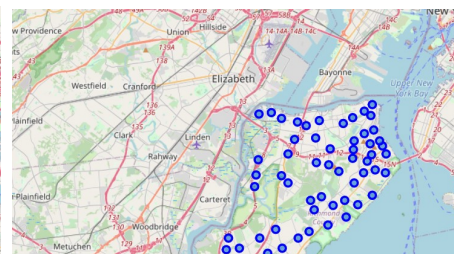
Brooklyn   Bronx

Queens
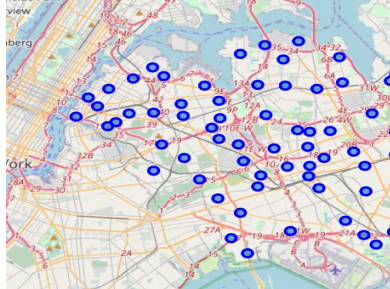
- We will then use Foursquare API to extract the venues in the neighbourhoods of Manhattan, Staten Island, Brooklyn, Bronx and Queens. We will extract the top 100 venues within a radius of 500m in each neighbourhood of these Boroughs and store it in separate dataframes.

- We then rank the top 10 venues in each neighbourhood of these Boroughs. Ranking is done by the frequency of the venue type in these neighbourhoods.

- Here we get to use one of the popular machine learning techniques - K-Means clustering. We cluster the neighbourhoods and its top 10 venues in 5 separate clusters. We do this separately for all the Boroughs.

- Let's view the clusters of each of these Boroughs.

Manhattan
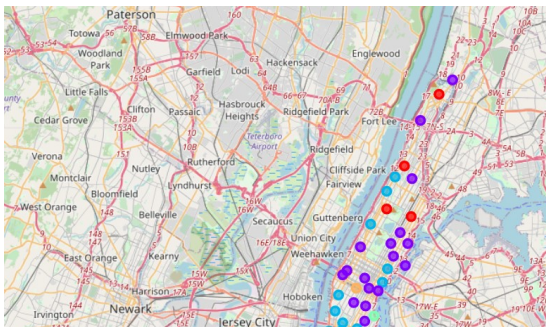
Brooklyn
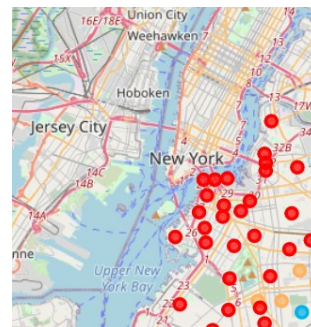
Bronx

Queens

- We can also examine these clusters in each of these Boroughs in detail:

Manhattan

Cluster 1

Cluster 1

```
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 0, manhattan_merged.columns[[1] + list(range(5, manhatt
3]]:
```

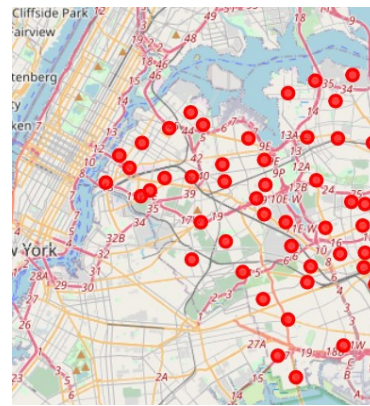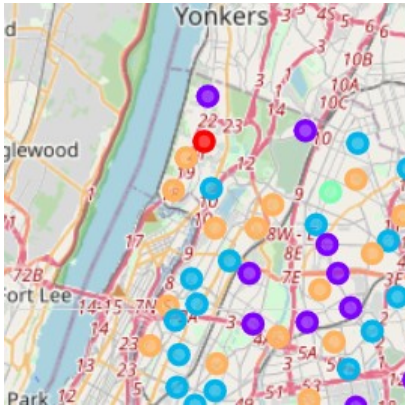|	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th M
3	Inwood	Mexican Restaurant	Café	Bakery	Pizza Place	Lounge	Restaurant	
4	Hamilton Heights	Pizza Place	Café	Coffee Shop	Mexican Restaurant	Deli / Bodega	Cocktail Bar	

Cluster 2

	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 M
0	Marble Hill	Sandwich Place	Gym	Coffee Shop	Yoga Studio	Deli / Bodega	Supplement Shop	
1	Chinatown	Chinese Restaurant	Bakery	Cocktail Bar	Optical Shop	American Restaurant	Spa	
2	Washington Heights	Café	Bakery	Grocery Store	Chinese Restaurant	Pizza Place	Mobile Phone Shop	
6	Central Harlem	Bar	Chinese Restaurant	Fried Chicken Joint	Gym / Fitness Center	African Restaurant	American Restaurant	Fren
8	Upper East Side	Italian Restaurant	Bakery	Exhibit	Gym / Fitness Center	Spa	American Restaurant	
9	Yorkville	Coffee Shop	Italian Restaurant	Gym	Bar	Sushi Restaurant	Deli / Bodega	
10	Lenox Hill	Italian Restaurant	Coffee Shop	Pizza Place	Cocktail Bar	Sushi Restaurant	Gym / Fitness Center	
11	Roosevelt Island	Park	Baseball Field	Gym / Fitness Center	Kosher Restaurant	Greek Restaurant	Restaurant	
13	Lincoln Square	Plaza	Café	Italian Restaurant	Gym / Fitness Center	Theater	Performing Arts Venue	
14	Clinton	Theater	Gym / Fitness Center	Coffee Shop	Wine Shop	Gym	Spa	Itali

Cluster 3

	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 M
5	Manhattanville	Coffee Shop	Seafood Restaurant	Italian Restaurant	Park	Deli / Bodega	Mexican Restaurant	
12	Upper West Side	Italian Restaurant	Coffee Shop	Wine Bar	Bakery	Middle Eastern Restaurant	Mediterranean Restaurant	
17	Chelsea	Art Gallery	Coffee Shop	Italian Restaurant	Seafood Restaurant	Ice Cream Shop	Juice Bar	
18	Greenwich Village	Italian Restaurant	Coffee Shop	Café	Sushi Restaurant	Wine Bar	Dessert Shop	Chine
21	Tribeca	Italian Restaurant	Park	Wine Bar	Café	Spa	American Restaurant	Gre
23	Soho	Italian Restaurant	Coffee Shop	Mediterranean Restaurant	Dessert Shop	Gym	French Restaurant	
24	West Village	Italian Restaurant	Wine Bar	New American Restaurant	Park	Coffee Shop	American Restaurant	Seaf
26	Morningside	Park	Bookstore	American Restaurant	Coffee Shop	Burger Joint	Deli / Bodega	

Cluster 4

	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 Mo
37	Stuyvesant	Boat or Ferry	Park	Gym / Fitness Center	Baseball Field	Harbor / Marina	Cocktail Bar	

Cluster 5

	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 M
33	Midtown South	Korean Restaurant	Hotel	Japanese Restaurant	Hotel Bar	Gym / Fitness Center	Burger Joint	

Staten Island

Cluster 1

	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 C
0	St. George	Clothing Store	Sporting Goods Shop	Bar	Italian Restaurant	Deli / Bodega	Pizza Place	
2	Stapleton	Discount Store	Bank	Sandwich Place	Pizza Place	Mexican Restaurant	Park	Don
3	Rosebank	Italian Restaurant	Pharmacy	Grocery Store	Museum	Storage Facility	Breakfast Spot	Pizz
4	West Brighton	Bar	Coffee Shop	Music Store	Pharmacy	Breakfast Spot	Bank	Italian Res
8	Port Richmond	Rental Car Location	Mexican Restaurant	Deli / Bodega	Construction & Landscaping	Donut Shop	Food Truck	Food & Drin
11	Castleton Corners	Pizza Place	Bank	Ice Cream Shop	Go Kart Track	Mini Golf	Grocery Store	Tatto
12	New Springville	Pizza Place	Health & Beauty Service	Chinese Restaurant	Coffee Shop	Mobile Phone Shop	Accessories Store	Shopp
13	Travis	Hotel	Deli / Bodega	Baseball Field	Comedy Club	Spanish Restaurant	Bowling Alley	
14	New Dorp	Italian Restaurant	Pizza Place	Vietnamese Restaurant	Gym / Fitness Center	Dessert Shop	Deli / Bodega	Coffe

Cluster 2

	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 Mo
1	New Brighton	Bus Stop	Park	Deli / Bodega	Bowling Alley	Discount Store	Playground	
7	South Beach	Pier	Beach	Deli / Bodega	Bus Stop	Athletics & Sports	Yoga Studio	
15	Oakwood	Bus Stop	Lawyer	Playground	Bar	Yoga Studio	Falafel Restaurant	
24	Park Hill	Bus Stop	Coffee Shop	Hotel	Athletics & Sports	Gym / Fitness Center	Falafel Restaurant	
27	Arlington	Bus Stop	Deli / Bodega	Intersection	Grocery Store	Coffee Shop	Farmers Market	
29	Grasmere	Bus Stop	Bank	Italian Restaurant	Japanese Restaurant	Juice Bar	Grocery Store	

Cluster 3

	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 Mo
6	Todt Hill	Park	Yoga Studio	Intersection	French Restaurant	Food Truck	Food & Drink Shop	

Cluster 4

	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 Mo
5	Grymes Hill	Bus Stop	Dog Run	Yoga Studio	Farmers Market	French Restaurant	Food Truck	Food &
56	Willowbrook	Bus Stop	Intersection	Deli / Bodega	Pizza Place	Falafel Restaurant	Food Truck	Food &

Cluster 5

	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 M
9	Mariner's Harbor	Italian Restaurant	Deli / Bodega	Bus Stop	Other Repair Shop	Furniture / Home Store	Farmers Market	
20	Tottenville	Italian Restaurant	Thrift / Vintage Store	Lawyer	Mexican Restaurant	Cosmetics Shop	Bus Stop	
34	New Dorp Beach	Italian Restaurant	Deli / Bodega	Diner	Restaurant	Food	Sports Bar	S
47	Richmond Town	Italian Restaurant	Café	Spa	Bagel Shop	History Museum	Farmers Market	
48	Shore Acres	Deli / Bodega	Italian Restaurant	Bus Stop	Bar	Intersection	Supermarket	

Brooklyn

Cluster 1

	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 M
0	Bay Ridge	Spa	Pizza Place	Italian Restaurant	Chinese Restaurant	American Restaurant	Greek Restaurant	C
1	Bensonhurst	Chinese Restaurant	Italian Restaurant	Sushi Restaurant	Ice Cream Shop	Donut Shop	Grocery Store	
2	Sunset Park	Bakery	Pizza Place	Deli / Bodega	Bank	Mexican Restaurant	Latin American Restaurant	Fried
3	Greenpoint	Pizza Place	Coffee Shop	Cocktail Bar	Bar	Yoga Studio	Deli / Bodega	Sus

Cluster 2

	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 M

Cluster 3

	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 Mc
10	East Flatbush	Chinese Restaurant	Food	Supermarket	Park	Fast Food Restaurant	Moving Target	
14	Brownsville	Park	Restaurant	Chinese Restaurant	Airport Terminal	Playground	Pharmacy	

Cluster 4

	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 Mc
28	Sea Gate	Bus Station	Beach	Spa	Home Service	Food & Drink Shop	Farmers Market	

Cluster 5

	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 Mo
26	East New York	Salon / Barbershop	Pizza Place	Plaza	Event Service	Convenience Store	Food Truck	
28	Canarsie	Caribbean Restaurant	Asian Restaurant	Grocery Store	Gym	Chinese Restaurant	Shipping Store	Tha
29	Flatlands	Pharmacy	Caribbean Restaurant	Fast Food Restaurant	Fried Chicken Joint	Arts & Crafts Store	Electronics Store	
32	Coney Island	Baseball Stadium	Caribbean Restaurant	Brewery	Clothing Store	Monument / Landmark	Pizza Place	
47	Prospect Park South	Caribbean Restaurant	Pizza Place	Fast Food Restaurant	Mobile Phone Shop	Grocery Store	Fried Chicken Joint	Mexica
55	Wingate	Fried Chicken Joint	Fast Food Restaurant	Park	Donut Shop	BBQ Joint	Nightclub	Gym / Fi
56	Rugby	Caribbean Restaurant	Bank	Grocery Store	Pizza Place	Diner	Fast Food Restaurant	Fried (
		Caribbean	Fast Food					

Bronx

Cluster 1

	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
3	Fieldston	Plaza	Bus Station	River	Donut Shop	Flower Shop	Fish Market	Fish & (

Cluster 2

	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
6	Woodlawn	Pizza Place	Deli / Bodega	Pub	Playground	Food & Drink Shop	Convenience Store	L
10	Pelham Parkway	Pizza Place	Italian Restaurant	Frozen Yogurt Shop	Deli / Bodega	Bakery	Sandwich Place	Me
16	East Tremont	Pizza Place	Breakfast Spot	Restaurant	Mobile Phone Shop	Deli / Bodega	Spanish Restaurant	
27	Throgs Neck	Deli / Bodega	Pizza Place	Juice Bar	Sports Bar	Bar	Asian Restaurant	C
31	Van Nest	Deli / Bodega	Pizza Place	Bakery	Playground	Coffee Shop	Donut Shop	
32	Morris Park	Pizza Place	Bakery	Deli / Bodega	Burger Joint	Buffet	Supermarket	C
33	Belmont	Italian Restaurant	Pizza Place	Deli / Bodega	Bakery	Grocery Store	Dessert Shop	D
35	North Riverdale	Pizza Place	Italian Restaurant	Bank	Sushi Restaurant	Donut Shop	Bus Station	Mexican

Cluster 3

	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 Mo
0	Wakefield	Pharmacy	Ice Cream Shop	Laundromat	Gas Station	Caribbean Restaurant	Donut Shop	Sar
2	Eastchester	Caribbean Restaurant	Deli / Bodega	Diner	Bus Stop	Bus Station	Metro Station	
5	Kingsbridge	Pizza Place	Bar	Mexican Restaurant	Supermarket	Latin American Restaurant	Sandwich Place	
9	Baychester	Donut Shop	Other Great Outdoors	Sandwich Place	Fried Chicken Joint	Electronics Store	Bank	Ma
13	University Heights	Fried Chicken Joint	Pizza Place	Chinese Restaurant	Pharmacy	Bakery	Shoe Store	
15	Fordham	Mobile Phone Shop	Shoe Store	Bank	Gym / Fitness Center	Spanish Restaurant	Donut Shop	
19	Melrose	Pizza Place	Pharmacy	Gym / Fitness Center	Grocery Store	Discount Store	Sandwich Place	
20	Mott Haven	Pizza Place	Spanish Restaurant	Gym	Donut Shop	Latin American Restaurant	Pharmacy	B
21	Port Morris	Furniture / Home Store	Latin American Restaurant	Spanish Restaurant	Restaurant	Food Truck	Storage Facility	B

Cluster 4

	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 Mo
8	Williamsbridge	Caribbean Restaurant	Convenience Store	Soup Place	Nightclub	Bar	Women's Store	

Cluster 5

	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 M
1	Co-op City	Bus Station	Restaurant	Grocery Store	Pharmacy	Fast Food Restaurant	Fried Chicken Joint	
4	Riverdale	Park	Medical Supply Store	Plaza	Gym	Bus Station	Food Truck	
7	Norwood	Pizza Place	Park	Deli / Bodega	Bank	Pharmacy	Coffee Shop	Chine
11	City Island	Harbor / Marina	Park	Seafood Restaurant	Thrift / Vintage Store	History Museum	Bar	
12	Bedford Park	Chinese Restaurant	Diner	Pizza Place	Mexican Restaurant	Sandwich Place	Deli / Bodega	Span
14	Morris Heights	Spanish Restaurant	Pharmacy	Pizza Place	Recreation Center	Supermarket	Grocery Store	
17	West Farms	Bus Station	Park	Playground	Basketball Court	Lounge	Donut Shop	
18	High Bridge	Pharmacy	Pizza Place	Bus Station	Chinese Restaurant	Supermarket	Deli / Bodega	S
23	Hunts Point	Shipping Store	Juice Bar	Gourmet Shop	Bank	Grocery Store	BBQ Joint	
25	Soundview	Chinese Restaurant	Grocery Store	Breakfast Spot	Latin American Restaurant	Playground	Discount Store	
26	Clason Point	Park	South American Restaurant	Pool	Home Service	Boat or Ferry	Bus Stop	

Queens

Cluster 1

	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
0	Astoria	Bar	Middle Eastern Restaurant	Mediterranean Restaurant	Greek Restaurant	Hookah Bar	Pizza Place	Seafc
1	Woodside	Grocery Store	Thai Restaurant	Latin American Restaurant	Bakery	Filipino Restaurant	Donut Shop	
2	Jackson Heights	Latin American Restaurant	Peruvian Restaurant	South American Restaurant	Bakery	Thai Restaurant	Mobile Phone Shop	Mexi
3	Elmhurst	Thai Restaurant	Mexican Restaurant	South American Restaurant	Vietnamese Restaurant	Chinese Restaurant	Pizza Place	Gym /
4	Howard Beach	Italian Restaurant	Clothing Store	Pharmacy	Sandwich Place	Bagel Shop	Fast Food	

Cluster 2

	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 Mo
63	Somerville	Park	Women's Store	Farm	Egyptian Restaurant	Electronics Store	Empanada Restaurant	

Cluster 3

	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 M
64	Brookville	Deli / Bodega	Women's Store	Farmers Market	Egyptian Restaurant	Electronics Store	Empanada	

Cluster 4

	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 M
54	Jamaica	Indian Restaurant	Intersection	Women's Store	Farmers Market	Egyptian Restaurant	Electronics Store	

Cluster 5

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 Mo
50	Neponsit	Beach	Beach Bar	Women's Store	Farmers Market	Empanada	Event Space	Falafe

- The above clusters with the information on each neighbourhood in each Borough would give enough insight to Mr. John on which Neighbourhood in which Borough that he would like to relocate to.

5. Results and Discussion

We have completed our analysis of all the 5 Boroughs and its neighbourhoods. Mr. John has not given any preference for any of the Boroughs and we did analysis of all the Boroughs and its neighbourhoods.

We now have a 5 clustered set of neighbourhoods for each of the Boroughs and **the cluster is formed basis the 10 most common venues in each of the neighbourhoods**. Remember Mr. John's criteria was for a good neighbourhood which has good options for restaurants, shopping, schools etc.

We put this criteria to the fore in our analysis and formed this cluster. By looking at each of these clusters in each of these neighbourhoods, Mr. John will now be able to make an informed decision on which neighbourhood and Borough he would like to move to.

6. Conclusion

Our purpose was to provide John with as much as information about the Neighbourhoods and Boroughs in New York with respect to the options they provide so that Mr. John can do a comparison of his existing neighbourhood in Toronto to the one that he would like to move to.

I believe we have done just that and provided the options that are available in each of the neighbourhood and Borough in New York and did a clustering basis that.

It is now up to Mr. John to make a decision of which neighbourhood and Borough he would like to move into. The other factors that Mr. John will also need to consider but not within the purview of this project is the affordability and also the proximity to his workplace.

Thank You and Hope you enjoyed this report.