



APPLIED DATA SCIENCE
WITH PYTHON

**THE BATTLE OF
NEIGHBORHOODS:
TORONTO VS. NEW YORK**

FARZIN MANIEI



I. INTRODUCTION

- **Background:** One of the important questions for business ventures willing to open a new location is “*What are the good potential site locations for a new store to maximize their sales?*”
- To answer this question, the first step is to layout a good demographic of current business ventures in the neighborhood of the interest. This would provide information about top current venues, their locations, the competitors business, population preferences, etc.
- **Problem:** As described in the background, the problem is to find potential site locations for a new venue in a neighborhood. This project investigates the potential site locations and business ventures in a neighborhood of Toronto, Canada. Then, the same exercise will be performed for the city of New York, U.S.A.
- **Interest:** This is a very interesting topic for entrepreneurs and business owners who are interested to start or expand their business in a neighborhood. The result of this project will be beneficial to layout a business development plan.

2. DATA ACQUISITION AND CLEANING

- **Data Sources:** The postal code of Canada are extracted from a Wikipedia page https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M and stored in a csv file as a table. The data for New York City is obtained from a dataset provided in https://cocl.us/new_york_dataset.
- **Data Cleaning:** After creating the datasets, the datasets need to be cleaned and wrangled to exclude the invalid or incomplete datapoints.

3. EXPLORATORY DATA ANALYSIS

- **Data Explanation:** The dataset for the city of Toronto includes Postcodes, Borough, and Neighborhood. The Postcode in the dataset is all the postcodes in the city of the Toronto that starts with M. The following figure shows a part time of the dataset as a table:
 - **Data Preparation:** In this step, dataset will be prepared for data analysis.

Postcode	Borough	Neighbourhood
0 M9A	Etobicoke	Islington Avenue
1 M4H	East York	Thorncliffe Park
2 M1B	Scarborough	Rouge, Malvern
3 M9L	North York	Humber Summit
4 M4Y	Downtown Toronto	Church and Wellesley
5 M9N	York	Weston
6 M3J	North York	Northwood Park, York University
7 M2H	North York	Hillcrest Village
8 M2J	North York	Henry Farm
9 M5S	Downtown Toronto	University of Toronto
10 M1T	Scarborough	Tam O'Shanter
11 M6L	North York	Downsview
12 M8Y	Etobicoke	Humber Bay, Mimico NE, Old Mill South, The Queen's Quay
13 M3H	North York	Bathurst Manor, Wilson Heights

Figure 1. City of Toronto's Neighborhood
(Postcode starting with 'M')

	id	name	categories	referrallid	hasPerk	location.address	location.lat	location.lng	location.labeledLatLngs	location.distance	location.postalCode	location.state
0	4fa862b3e4b0ebff2f749f06	Harry's Pizza	4fb5fdd8d48988dfca847393... "name": "P..."	1571111237	V-	225 Murray St	40.715218	-70.041739	[{"label": "display", "lat": 40.715217, "lng": -70.041739, "place_id": "ChIJyvJ..."}]	58	10282	NY
1	4f3232e21983e91c7bfe94	Conca Cucina Italian Restaurant	4d4b7f05d754a06374a8f1939... "name": "C..."	1571111237	V-	63 W Broadway	40.714466	-74.010086	[{"label": "display", "lat": 40.71446, "lng": -74.010086, "place_id": "ChIJG..."}]	446	10007	NY
2	3fd6200f98e4ab204fe4ee3	Ecco	4b58dd8d48988df01044713... "name": "E..."	1571111237	V-	124 Chambers St	40.715337	-74.008848	[{"label": "display", "lat": 40.715337, "lng": -74.008848, "place_id": "ChIJF..."}]	549	10007	NY

Figure 2. City of New York's Neighborhood

Postcode	Borough	Neighbourhood
0 M1A	Not assigned	Not assigned
1 M2A	Not assigned	Not assigned
2 M3A	North York	Parkwoods
3 M4A	North York	Victoria Village
4 M5A	Downtown Toronto	Harbourfront

Figure 3. City of Toronto Dataset (Sorted).

3. EXPLORATORY DATA ANALYSIS (CONT.)

Postcode	Borough	Neighbourhood
0	M1A	Not assigned
1	M1B	Scarborough
2	M1C	Scarborough
3	M1E	Scarborough
4	M1G	Scarborough
5	M1H	Scarborough
6	M1J	Scarborough
7	M1K	Scarborough
8	M1L	Scarborough
9	M1M	Scarborough
10	M1N	Scarborough
11	M1P	Scarborough
12	M1R	Scarborough
13	M1S	Scarborough
14	M1T	Scarborough
15	M1V	Scarborough
16	M1W	Scarborough
17	M1X	Scarborough
18	M1Y	Not assigned
19	M1Z	Not assigned
20	M2A	Not assigned
		Queen's Park
		Rouge, Malvern
		Highland Creek, Rouge Hill, Port Union
		Guildwood, Morningside, West Hill
		Woburn
		Cedarbrae
		Scarborough Village
		East Birchmount Park, Ionview, Kennedy Park
		Clairlea, Golden Mile, Oakridge
		Cliffcrest, Cliffside, Scarborough Village West
		Birch Cliff, Cliffside West
		Dorset Park, Scarborough Town Centre, Wexford ...
		Maryvale, Wexford
		Agincourt
		Clarks Corners, Sullivan, Tam O'Shanter
		Agincourt North, L'Amoreaux East, Milliken, St...
		L'Amoreaux West
		Upper Rouge
		Queen's Park
		Queen's Park
		Queen's Park

Figure 4. City of Toronto Dataset
 (After Replacing unknown Neighborhood cells).

	Postcode	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

Figure 5. City of Toronto – Geospatial Data for the Postcodes.

Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M1B	Scarborough	Rouge, Malvern	43.806686 -79.194353
1	M1T	Scarborough	Tam O'Shanter	43.781638 -79.304302
2	M1X	Scarborough	Upper Rouge	43.836125 -79.205636
3	M1M	Scarborough	Cliffcrest, Cliffside	43.716316 -79.239476
4	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535 -79.160497
5	M1R	Scarborough	Maryvale, Wexford	43.750072 -79.295845
6	M1J	Scarborough	Scarborough Village	43.744734 -79.239476

Figure 6. City of Toronto – Dataset combined with the Geospatial Data.

Figure 7. New York City Dataset.

4. EXPLORING THE DATA

- **Data Visualization:** Data Visualization is to represent data graphically. One of the interesting data visualization methods is map visualization which is used to graphically display data in the form of map. Dealing with the geospatial dataset, it is recommended to produce map visualization for better understanding the data. This helps to discover the potential stories that are hidden prior to map visualization (or data visualization). Using the ‘folium’ library in Python, the datasets for Toronto’s and New York City’s Neighborhoods can be represented on a map using northing and easting which provide the map visualization for the datasets.
- Toronto’s and New York City’s datasets are represented on a map that shown in Figures 8 and 9, respectively.

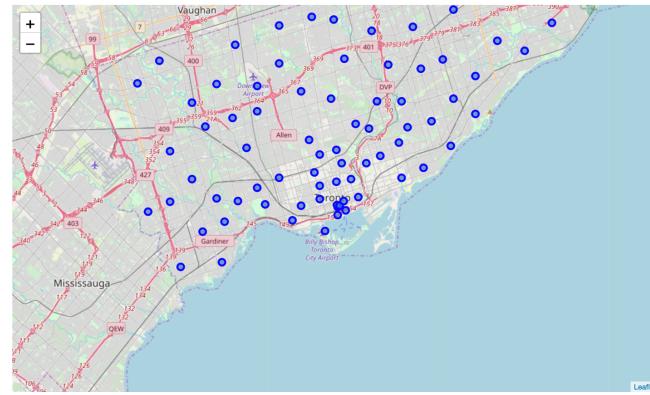


Figure 8. Toronto’s Neighborhood Dataset – Map Visualization.

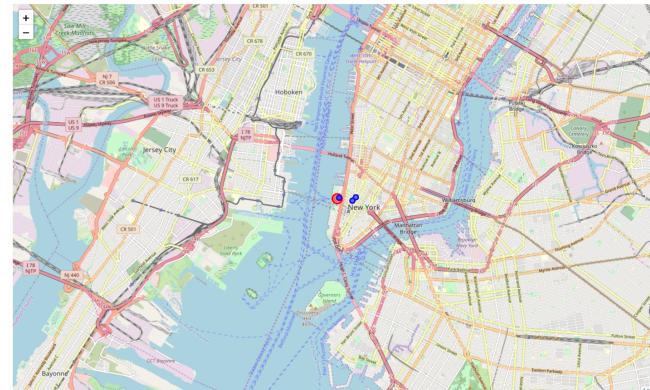


Figure 9. New York City’s Neighborhood Dataset – Map Visualization.

4. EXPLORING THE DATA (CONT.)

- **Identifying the Target in the Dataset:** The dataset for Toronto's neighborhood includes several neighborhoods. For this project, '**Scarborough**' neighborhood is selected as a target neighborhood from the dataset of Toronto's Neighborhood. Then, the dataset is filtered by '**Borough**' variable with value of '**Scarborough**'. Figure 10 shows this filtered dataset.

Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M1B	Scarborough	Rouge, Malvern	43.806686 -79.194353
1	M1T	Scarborough	Tam O'Shanter	43.781638 -79.304302
2	M1X	Scarborough	Upper Rouge	43.836125 -79.205636
3	M1M	Scarborough	Cliffcrest, Cliffside	43.716316 -79.239476
4	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535 -79.160497
5	M1R	Scarborough	Maryvale, Wexford	43.750072 -79.295849
6	M1J	Scarborough	Scarborough Village	43.744734 -79.239476

Figure 10. Toronto's Neighborhood Dataset – Filtered Dataset.

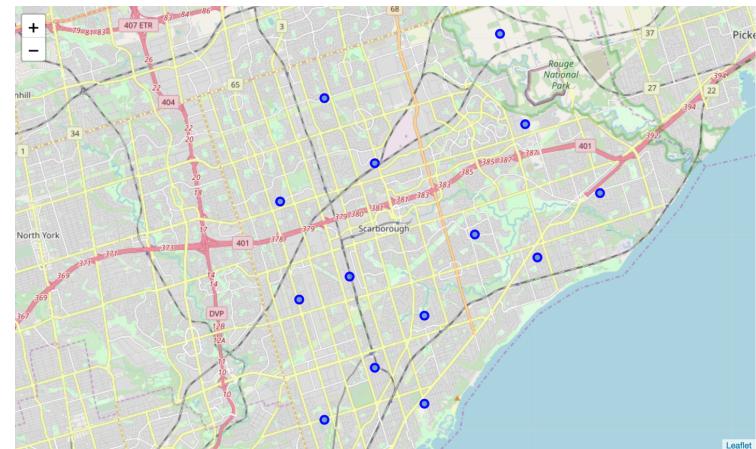


Figure 11. Toronto's Neighborhood Sub-Dataset – Map Visualization.

4. EXPLORING THE DATA (CONT.)

- **Problem Definition:** As discussed in the background, this project is focused on exploring current venues in the neighborhoods of Toronto in Canada and New York City in U.S.A in order to find potential site locations for a profitable venue in the neighborhoods.
- **Identifying Current Venues in the Neighborhoods**
 - **Neighborhoods of Toronto:** To understand the market, the sub-dataset for '**Scarborough**' neighborhood of Toronto will be ordered by current venues. The following figure show the top nine venues in '**Scarborough**' neighborhood:
 - The statistics of the current venues in Toronto's neighborhood is provided in Figure 13.

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0 Rouge, Malvern	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant
1 Tam O'Shanter	43.781638	-79.304302	Remezzo Italian Bistro	43.778649	-79.308264	Italian Restaurant
2 Tam O'Shanter	43.781638	-79.304302	The Royal Chinese Restaurant 遇風樓小炒	43.780505	-79.298844	Chinese Restaurant
3 Tam O'Shanter	43.781638	-79.304302	Kub Khao	43.780438	-79.299837	Thai Restaurant
4 Tam O'Shanter	43.781638	-79.304302	Eight Noodles	43.778234	-79.308299	Noodle House
5 Tam O'Shanter	43.781638	-79.304302	TD Canada Trust	43.779169	-79.303617	Bank
6 Tam O'Shanter	43.781638	-79.304302	KFC	43.779440	-79.303371	Fast Food Restaurant
7 Tam O'Shanter	43.781638	-79.304302	Popeyes Louisiana Kitchen	43.780476	-79.298460	Fried Chicken Joint
8 Tam O'Shanter	43.781638	-79.304302	Little Caesars	43.780563	-79.298624	Pizza Place
9 Tam O'Shanter	43.781638	-79.304302	A&W Canada	43.780327	-79.299060	Breakfast Spot

Figure 12. Toronto's Neighborhood Sub-Dataset – 'Scarborough' Neighborhood.
(Sub-dataset is ordered by venue, top nine venues)

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Agincourt	4	4	4	4	4	4
Agincourt North, Milliken	3	3	3	3	3	3
Birch Cliff	4	4	4	4	4	4
Clairlea, Golden Mile, Oakridge	8	8	8	8	8	8
Cliffcrest, Cliffside	2	2	2	2	2	2
Dorset Park, Scarborough Town Centre, Wexford Heights	6	6	6	6	6	6
Highland Creek, Rouge Hill, Port Union	1	1	1	1	1	1
Ionview, Kennedy Park	5	5	5	5	5	5
Maryvale, Wexford	7	7	7	7	7	7
Morningside, West Hill	7	7	7	7	7	7
Rouge, Malvern	1	1	1	1	1	1
Scarborough Village	2	2	2	2	2	2
Tam O'Shanter	11	11	11	11	11	11
Woburn	3	3	3	3	3	3

Figure 13. Toronto's Neighborhood – Statistics of the Current Venues.

4. EXPLORING THE DATA (CONT.)

- Now, it is important to find the top 10 most common venues for Toronto's neighborhoods. Figure 14 provides the top 10 most common venues for Toronto's neighborhoods. All the venues are clustered based on the type of venue. Figure 15 shows all the venues in Toronto's neighborhood by type of venue. These clusters are represented as map visualization in Figure 16. By taking these steps, the dataset will be classified such that it reveals the top 10 most common venues in the target sub-dataset **Scarborough** neighborhood. This method is called '**Data Classification**'.

	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	Agincourt	Skating Rink	Lounge	Breakfast Spot	Clothing Store	Vietnamese Restaurant	Coffee Shop	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store
1	Agincourt North, Milliken	Park	Playground	Vietnamese Restaurant	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store	Convenience Store	College Stadium
2	Birch Cliff	General Entertainment	Skating Rink	Café	College Stadium	Vietnamese Restaurant	Coffee Shop	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store
3	Clairea, Golden Mile, Oakridge	Bakery	Bus Line	Soccer Field	Fast Food Restaurant	Bus Station	Metro Station	Vietnamese Restaurant	College Stadium	Fried Chicken Joint	Electronics Store
4	Cliffcrest, Cliffside	American Restaurant	Motel	Coffee Shop	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store	Convenience Store
5	Dorset Park, Scarborough Town Centre, Westford ...	Indian Restaurant	Vietnamese Restaurant	Chinese Restaurant	Latin American Restaurant	Pet Store	Bar	Breakfast Spot	Fried Chicken Joint	Auto Garage	Fast Food Restaurant
6	Highland Creek, Rouge Hill, Port Union	Bar	Vietnamese Restaurant	Coffee Shop	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store	Convenience Store
7	Jonview, Kennedy Park	Coffee Shop	Discount Store	Department Store	Convenience Store	Chinese Restaurant	Vietnamese Restaurant	Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store
8	Maryvale, Wexford	Middle Eastern Restaurant	Auto Garage	Smoke Shop	Shopping Mall	Sandwich Place	Breakfast Spot	Vietnamese Restaurant	Fast Food Restaurant	Electronics Store	Discount Store
9	Morningside, West Hill	Intersection	Electronics Store	Breakfast Spot	Rental Car Location	Pizza Place	Medical Center	Mexican Restaurant	Vietnamese Restaurant	Coffee Shop	Fast Food Restaurant
10	Rouge, Malvern	Fast Food Restaurant	Vietnamese Restaurant	Intersection	General Entertainment	Fried Chicken Joint	Electronics Store	Discount Store	Department Store	Convenience Store	College Stadium
11	Scarborough Village	Playground	Convenience Store	Vietnamese Restaurant	Clothing Store	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store
12	Tam O'Shanter	Pizza Place	Italian Restaurant	Thai Restaurant	Fast Food Restaurant	Noodle House	Pharmacy	Chinese Restaurant	Fried Chicken Joint	Breakfast Spot	Bank
13	Woburn	Coffee Shop	Korean Restaurant	Vietnamese Restaurant	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store	Convenience Store

Figure 14. Toronto's Neighborhood – Top 10 Most Common Venues.

4. EXPLORING THE DATA (CONT.)

Postcode	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	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	M1B	Scarborough	Rouge, Malvern	43.806666	-79.194353	1	Fast Food Restaurant	Vietnamese Restaurant	Intersection	General Entertainment	Fried Chicken Joint	Electronics Store	Discount Store	Department Store	Convenience Store	College Stadium
1	M1T	Scarborough	Tan O' Shanter	43.778138	-79.304302	1	Pizza Place	Italian Restaurant	Thai Restaurant	Fast Food Restaurant	Noodle House	Pharmacy	Chinese Restaurant	Fried Chicken Joint	Breakfast Spot	Bank
2	M1X	Scarborough	Upper Rouge, Cliffcrest, Cawthron	43.836125	-79.205636	1	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN
3	M1M	Scarborough	Hightower Creek, Highgate, Highgate Creek, Highgate Hill, Port Union	43.716316	-79.239476	1	American restaurant	Moda	Coffee Shop	Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store	Convenience Store
4	M1C	Scarborough	Rouge Hill, Port Union	43.784535	-79.160497	4	Bar	Vietnamese Restaurant	Coffee Shop	Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store	Convenience Store
5	M1R	Scarborough	Merryvale, Wexford	43.750072	-79.295849	1	Middle Eastern Restaurant	Auto Garage	Smoke Shop	Shopping Mall	Sandwich Place	Breakfast Spot	Vietnamese Restaurant	Fast Food Restaurant	Electronics Store	Discount Store
6	M1J	Scarborough	Scarborough Village	43.744734	-79.239476	2	Playground	Convenience Store	Vietnamese Restaurant	Clothing Store	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store
7	M1V	Scarborough	Agincourt North, Milliken	43.815252	-79.284577	1	Park	Playground	Vietnamese Restaurant	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store	Convenience Store	College Stadium
8	M1L	Scarborough	Clairlea, Goldring, Mts, Oakridge	43.711112	-79.284577	1	Bakery	Bus Line	Soccer Field	Fast Food Restaurant	Bus Station	Metro Vietnamese Restaurant	College Stadium	Fried Chicken Joint	Electronics Store	
9	M1Q	Scarborough	Woburn	43.770992	-79.216917	1	Coffee Shop	Korean Restaurant	Vietnamese Restaurant	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store	Convenience Store
10	M1K	Scarborough	Ionview, Kempen Park	43.727919	-79.302029	3	Coffee Shop	Discount Store	Department Store	Convenience Store	Chinese Restaurant	Vietnamese Restaurant	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store
11	M1N	Scarborough	Birch Cliff	43.692657	-79.264848	1	General Entertainment	Skating Rink	Café	College Station	Vietnamese Restaurant	Coffee Shop	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store
12	M1E	Scarborough	Morningside, West Hill	43.763573	-79.188711	1	Intersection	Electronics Store	Breakfast Spot	Rental Car Location	Pizza Place	Medical Center	Mexican Restaurant	Vietnamese Restaurant	Coffee Shop	Fast Food Restaurant
14	M1P	Scarborough	Dorset Park, Scarborough Town Centre, Westford	43.757410	-79.273304	0	Indian Restaurant	Vietnamese Restaurant	Chinese Restaurant	Latin American Restaurant	Pet Store	Bar	Breakfast Spot	Fried Chicken Joint	Auto Garage	Fast Food Restaurant

Figure 15. Toronto's Neighborhood – Top 10 Most Common Venues (Clustered).

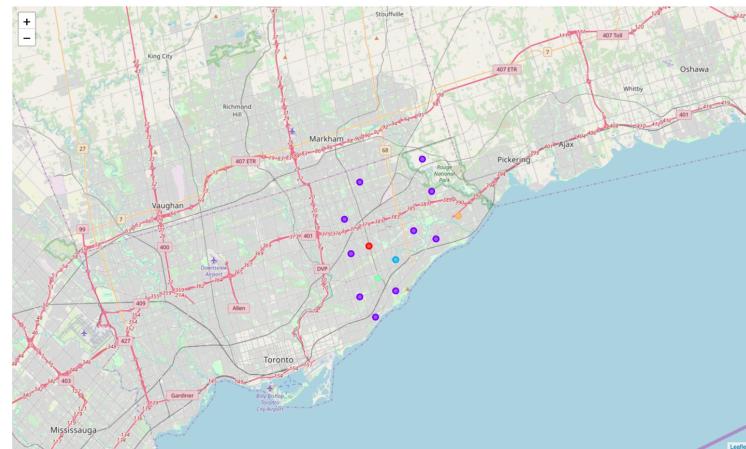


Figure 16. Toronto's Neighborhood – Top 10 Most Common Venues (Clustered).

4. EXPLORING THE DATA (CONT.)

Borough	Cluster Labels	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
14 Scarborough	0	Indian Restaurant	Vietnamese Restaurant	Chinese Restaurant	Latin American Restaurant	Pet Store	Bar	Breakfast Spot	Fried Chicken Joint	Auto Garage	Fast Food Restaurant

Figure 17. 'Scarborough' Neighborhood – Top 10 Most Common Venues in (Clustered).

Borough	Cluster Labels	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 Scarborough	1	Fast Food Restaurant	Vietnamese Restaurant	Intersection	General Entertainment	Fried Chicken Joint	Electronics Store	Discount Store	Department Store	Convenience Store	College Stadium
1 Scarborough	1	Pizza Place	Italian Restaurant	Thai Restaurant	Fast Food Restaurant	Noodle House	Pharmacy	Chinese Restaurant	Fried Chicken Joint	Breakfast Spot	Bank
2 Scarborough	1	NaN	NaN	NaN	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store	Convenience Store
3 Scarborough	1	American Restaurant	Motel	Coffee Shop	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store	Convenience Store
5 Scarborough	1	Middle Eastern Restaurant	Auto Garage	Smoke Shop	Shopping Mall	Sandwich Place	Breakfast Spot	Vietnamese Restaurant	Fast Food Restaurant	Electronics Store	Discount Store
7 Scarborough	1	Park	Playground	Vietnamese Restaurant	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store	Convenience Store	College Stadium
8 Scarborough	1	Bakery	Bus Line	Soccer Field	Fast Food Restaurant	Bus Station	Metro Station	Vietnamese Restaurant	College Stadium	Fried Chicken Joint	Electronics Store
9 Scarborough	1	Coffee Shop	Korean Restaurant	Vietnamese Restaurant	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store	Convenience Store
11 Scarborough	1	General Entertainment	Skating Rink	Café	College Stadium	Vietnamese Restaurant	Coffee Shop	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store
12 Scarborough	1	Intersection	Electronics Store	Breakfast Spot	Rental Car Location	Pizza Place	Medical Center	Mexican Restaurant	Vietnamese Restaurant	Coffee Shop	Fast Food Restaurant

Figure 18. 'Scarborough' Neighborhood – Top 10 Most Common Venues in (Clustered). (Different Cluster Labels)

4. EXPLORING THE DATA (CONT.)

Neighborhoods of New York City:
The same process has been followed for New York City's neighborhoods and the following figure is obtained:

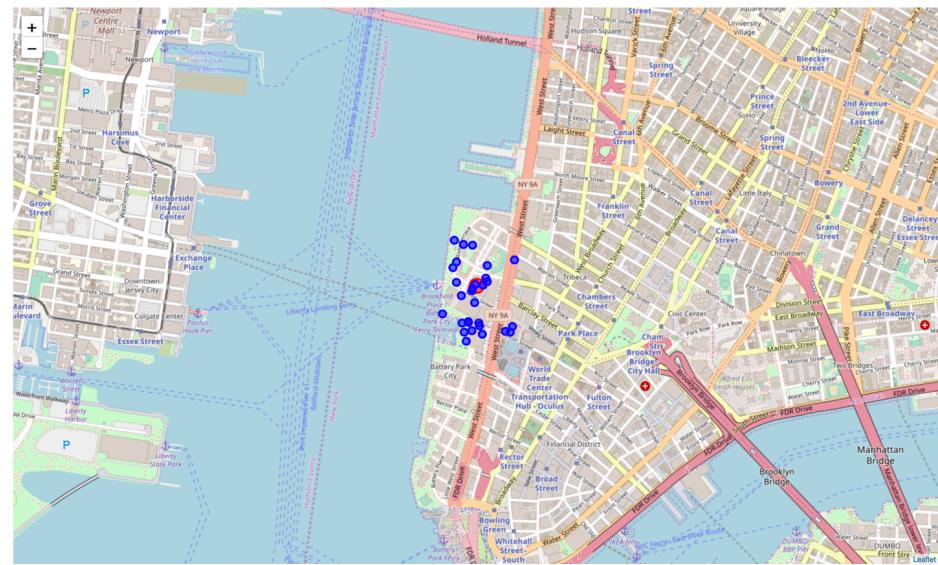


Figure 19. New York City Neighborhood – Most Common Venues in (Clustered).
(Map Visualization).

5. CONCLUSION

- According to the data exploration, if an investor is willing to open a new business in **Scarborough** neighborhood of Toronto, Canada then Vietnamese restaurants is the best most common restaurant in the neighborhood while the best most common restaurant in New York City neighborhoods is Italian restaurants.

6. FUTURE DEVELOPMENT

- This project can be extended by acquiring the census data for the target neighborhoods in Toronto and New York City. By including the census data, the current market can be explored for reaching better decisions with more details.