

1 Introduction

This Project is based on a scenario. In this scenario, there is a man called Ismet. He is a fresh MSc Finance Graduate from one of the top universities of Europe. He finally found jobs after seeking about 6 months. Meanwhile, he was working on data analytics skills in order to be successful in this area. These jobs are in Manhattan, NY and Toronto, Ontario. Since these two job offers are almost the same in terms of salary and expectations, he decided to analyze and compare the cities. He has to decide in one week so he needs to be really quick on that task. The choice is too important for him because he will spend at least 4 years there. Nobody wants to live in a place that s/he does not want.

These cities are 2 big financial capitals of their countries and they are too big. He realized that he is capable of doing such an analysis based on data science skills on the area of office which they have to attend. Therefore, he can answer these type of questions;

1-Are these two cities similar?

2-Which city matches with my personal needs-hobbies?

3-Which city has more green area?

4-Which neighborhood should he choose to live after selecting the city?

In order to answer these questions, he needs location data. He uses the Foursquare API for gathering data to make graphical and statistical analysis of venues in the city.

The rest of the report is as follow;

After defining the problem in introduction, dataset and the source is presented. Then, in methodology section, there are explanatory data analysis inferential statistical testing. In addition, there is an explanation about which machine learning model is used and why. In results section, results of the model is discussed and come up with an answer to the questions. Then, I discuss the observations and drawbacks and make some recommendations for future work. Finally, I conclude the report with a conclusion part.

2 Data

There are 3 sources of data in total. They are Foursquare API for venues data, Wikipedia for the Postal Codes, Borough and Neighborhood data of Toronto(gathered by BeautifulSoup) and IBM Cloud which has New York Postal Codes, Borough and Neighborhood data(csv format).

Sources :

- <https://developer.foursquare.com/docs/places-api/>
- https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDriverSkillsNetwork-DS071-labs/newyork_data.json
- https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDriverSkillsNetwork-DS071-labs_v1/Geospatial_Coordinates.csv

The data formats are in csv and json format. Data are transformed into pandas dataframe and combined. The variables in datasets are borough, neighborhood name, postal code, latitude, longitude, venue and venue category.

Note :The first 5 rows of datasets are presented in the next page

Manhattan Data

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	
0	Marble Hill	40.876551	-73.91066	Arturo's	40.874412	-73.910271	Pizza Place
1	Marble Hill	40.876551	-73.91066	Bikram Yoga	40.876844	-73.906204	Yoga Studio
2	Marble Hill	40.876551	-73.91066	Tibbett Diner	40.880404	-73.908937	Diner
3	Marble Hill	40.876551	-73.91066	Dunkin'	40.877136	-73.906666	Donut Shop
4	Marble Hill	40.876551	-73.91066	Starbucks	40.877531	-73.905582	Coffee Shop

Toronto Data

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	
0	Regent Park, Harbourfront	43.65426	-79.360636	Roselle Desserts	43.653447	-79.362017	Bakery
1	Regent Park, Harbourfront	43.65426	-79.360636	Tandem Coffee	43.653559	-79.361809	Coffee Shop
2	Regent Park, Harbourfront	43.65426	-79.360636	Cooper Koo Family YMCA	43.653249	-79.358008	Distribution Center
3	Regent Park, Harbourfront	43.65426	-79.360636	Impact Kitchen	43.656369	-79.356980	Restaurant
4	Regent Park, Harbourfront	43.65426	-79.360636	Body Blitz Spa East	43.654735	-79.359874	Spa

3 Methodology

Before the explanation of k-means clustering, I introduce the maps where all neighborhoods are located and the dataset which I create to show most common venues in neighbourhoods.

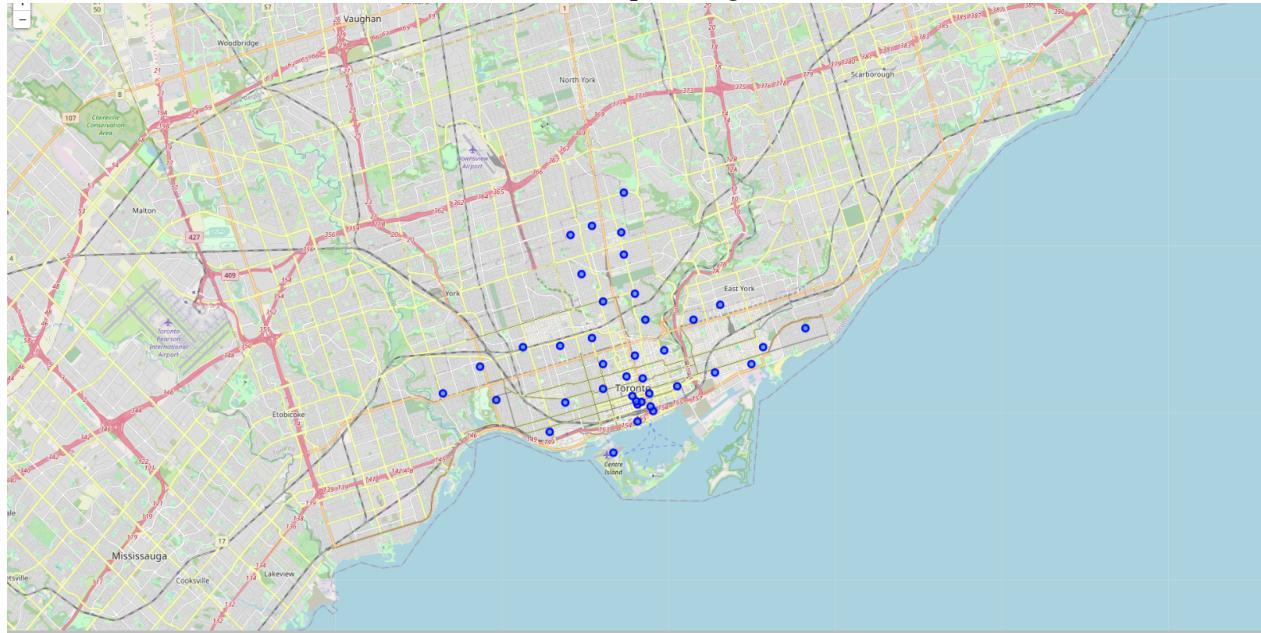
3.1 Maps

Each of the points represents neighbourhoods in the map. Since the jobs are in Manhattan and Toronto City Center, I used filtered data which consist of neighborhoods with names "Manhattan" and "Toronto". You can see 4 maps below which are Main Maps of New York and Toronto (includes every neighborhood in the city) and target maps which are maps of Manhattan and Toronto City Center.

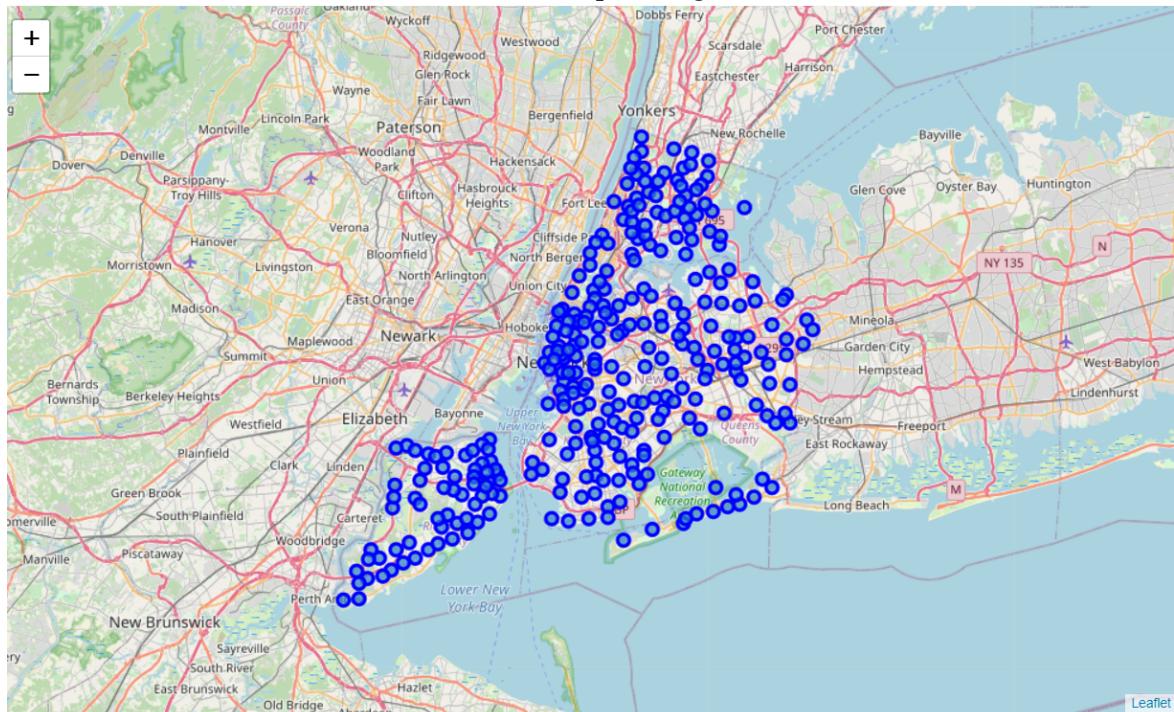
Toronto Map of Neighbourhoods



Toronto Center Map of Neighborhoods

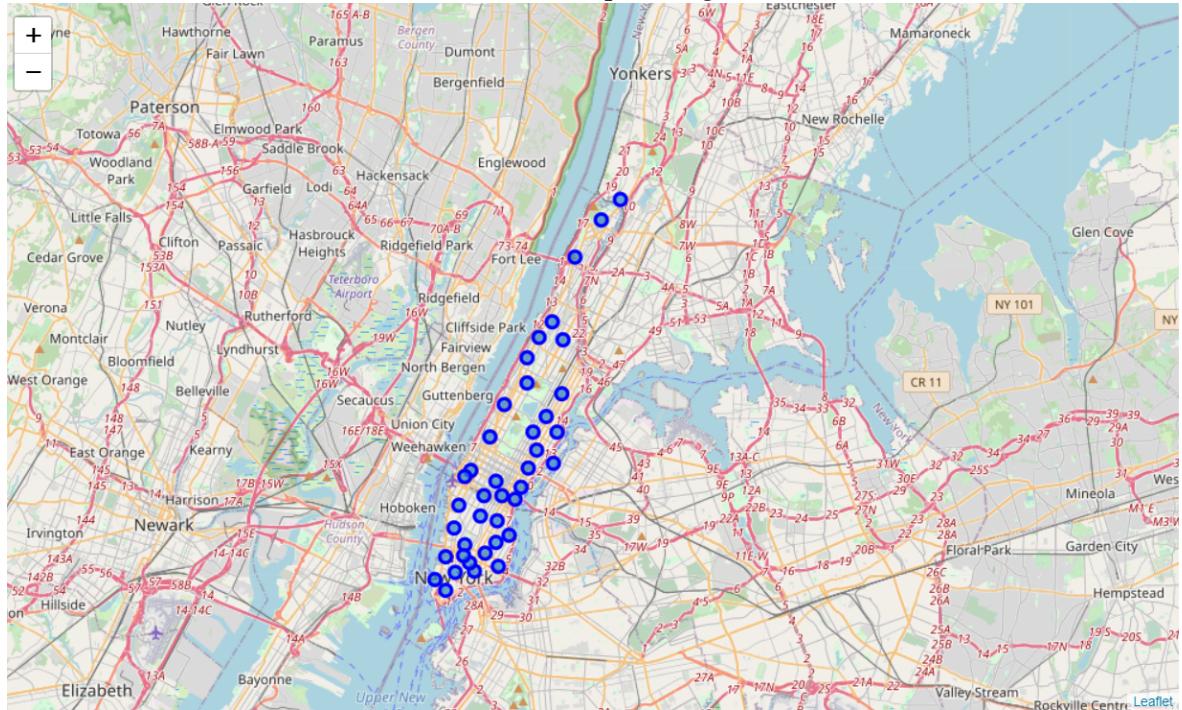


New York Map of Neighborhoods



According to the maps, both cities have unique shapes. Toronto-Ontario is a peninsula whereas Manhattan is an island. This means that there is higher likelihood that Ismet lives along the coast if he chooses Manhattan.

Manhattan Map of Neighborhoods



3.2 Most Common Venues in Neighbourhoods

Tables which are represented below show the most common venues in neighborhoods of Manhattan and Toronto City Center. However, since the dataset is too big to show, I only include 5 neighborhoods but note that all neighborhoods in one cluster are shown in Results section. These 2 data sets are just for to give an insight about the idea, and to show data that is used in clustering.

Most Common Venues in Neighbourhoods of Manhattan, New York

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 Battery Park City	Park	Clothing Store	Hotel	Coffee Shop	Memorial Site	Gym	Playground	Plaza	Burger Joint	Boat or Ferry
1 Carnegie Hill	Coffee Shop	Café	Wine Shop	Yoga Studio	Bar	Bookstore	Cosmetics Shop	French Restaurant	Gym / Fitness Center	Gym
2 Central Harlem	African Restaurant	French Restaurant	Chinese Restaurant	Bar	American Restaurant	Fried Chicken Joint	Public Art	Seafood Restaurant	Gym / Fitness Center	Cycle Studio
3 Chelsea	Coffee Shop	Bakery	American Restaurant	Art Gallery	Hotel	Wine Shop	Ice Cream Shop	French Restaurant	Seafood Restaurant	Bookstore
4 Chinatown	Chinese Restaurant	Bakery	Cocktail Bar	American Restaurant	Dessert Shop	Hotpot Restaurant	Salon / Barbershop	Optical Shop	Asian Restaurant	Bubble Tea Shop

Most Common Venues in Neighbourhoods of Toronto City Center

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
I Berczy Park	Bakery	Coffee Shop	Cocktail Bar	Farmers Market	Pharmacy	Seafood Restaurant	Restaurant	Beer Bar	Cheese Shop	Pub
Brockton, Parkdale Village, Exhibition Place	Café	Coffee Shop	Breakfast Spot	Gym	Intersection	Performing Arts Venue	Pet Store	Grocery Store	Nightclub	Climbing Gym
! CN Tower, King and Spadina, Railway Lands, Har...	Airport Service	Plane	Bar	Coffee Shop	Rental Car Location	Sculpture Garden	Boat or Ferry	Harbor / Marina	Airport Lounge	Airport Food Court
! Central Bay Street	Coffee Shop	Sandwich Place	Café	Italian Restaurant	Restaurant	Spa	Burger Joint	Japanese Restaurant	Bubble Tea Shop	Salad Place
! Christie	Grocery Store	Café	Park	Restaurant	Italian Restaurant	Athletics & Sports	Candy Store	Nightclub	Baby Store	Coffee Shop

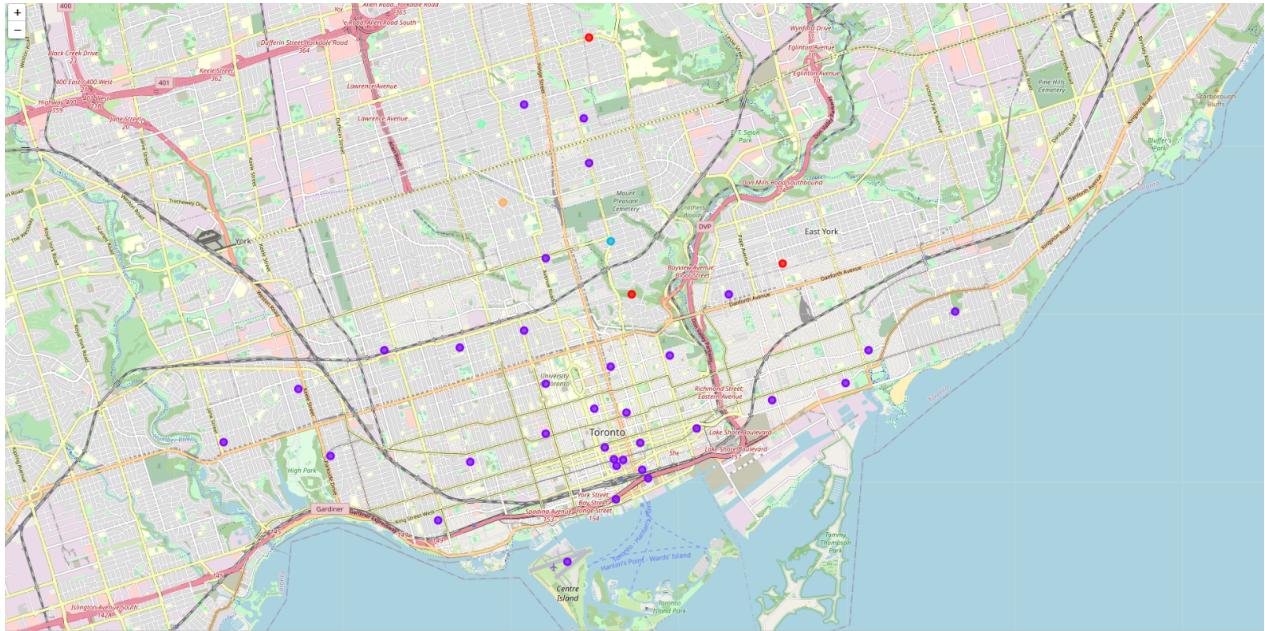
3.3 K-Means Clustering

The reason why I use clustering is based on finding duplicates. A cluster refers to a collection of data points aggregated together because of certain similarities. You will define a target number k, which refers to the number of centroids you need in the dataset. A centroid is the imaginary or real location representing the center of the cluster. Every data point is allocated to each of the clusters through reducing the in-cluster sum of squares. In other words, the K-means algorithm identifies k number of centroids, and then allocates every data point to the nearest cluster, while keeping the centroids as small as possible. The 'means' in the K-means refers to averaging of the data which is finding the centroid. (Garbade,J.M,2018)

4 Results

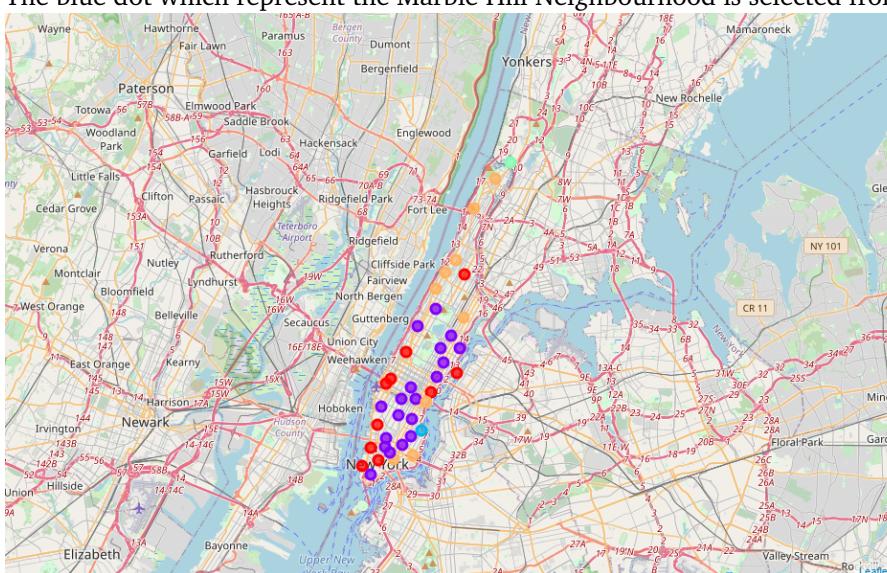
4.1 Cluster Maps

4.1.1 Toronto, Ontario



4.1.2 Manhattan, NY

The blue dot which represent the Marble Hill Neighbourhood is selected from clusters.



4.2 Examining Clusters

Each cluster is examined and one neighborhood is chosen from each cluster. Then, the best neighborhood among chosen ones is selected.(based on the location and venues). Finally, the battle begins between the best neighborhoods of Toronto and Manhattan, and the final decision is made.

4.2.1 Manhattan,NY

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 Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
6	Central Harlem	African Restaurant	French Restaurant	Chinese Restaurant	Bar	American Restaurant	Fried Chicken Joint	Public Art	Seafood Restaurant	Gym / Fitness Center	Cycle Studio
11	Roosevelt Island	Coffee Shop	Playground	Park	Scenic Lookout	Food & Drink Shop	Outdoors & Recreation	Gym / Fitness Center	Gym	Greek Restaurant	Sandwich Place
13	Lincoln Square	Plaza	Performing Arts Venue	Gym / Fitness Center	Theater	Café	Concert Hall	Indie Movie Theater	Wine Shop	French Restaurant	Food Truck
14	Clinton	Italian Restaurant	Theater	Gym / Fitness Center	Gym	American Restaurant	Sandwich Place	Coffee Shop	Hotel	Wine Shop	Cocktail Bar
21	Tribeca	American Restaurant	Italian Restaurant	Park	Café	Wine Bar	Spa	Greek Restaurant	Basketball Court	French Restaurant	Coffee Shop
24	West Village	Italian Restaurant	Cocktail Bar	New American Restaurant	American Restaurant	Park	Coffee Shop	Jazz Club	Ice Cream Shop	Wine Bar	Cosmetics Shop
28	Battery Park City	Park	Clothing Store	Hotel	Coffee Shop	Memorial Site	Gym	Playground	Plaza	Burger Joint	Boat or Ferry
32	Civic Center	Coffee Shop	Spa	Gym / Fitness Center	French Restaurant	Cocktail Bar	American Restaurant	Hotel	Italian Restaurant	Park	Boxing Gym
35	Turtle Bay	Italian Restaurant	Coffee Shop	Japanese Restaurant	Park	Steakhouse	Wine Bar	Deli / Bodega	Sushi Restaurant	Café	Seafood Restaurant
39	Hudson Yards	Gym / Fitness Center	Café	Hotel	Italian Restaurant	Restaurant	American Restaurant	Thai Restaurant	Coffee Shop	Park	Gym

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 Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
8	Upper East Side	Italian Restaurant	Exhibit	Juice Bar	Gym / Fitness Center	Coffee Shop	Bakery	Yoga Studio	Art Gallery	Pizza Place	Cosmetics Shop
9	Yorkville	Italian Restaurant	Coffee Shop	Bar	Gym	Deli / Bodega	Sushi Restaurant	Wine Shop	Ice Cream Shop	Mexican Restaurant	Diner
10	Lenox Hill	Italian Restaurant	Coffee Shop	Sushi Restaurant	Cocktail Bar	Pizza Place	Café	Gym / Fitness Center	Burger Joint	Taco Place	
12	Upper West Side	Italian Restaurant	Café	Bakery	Wine Bar	Coffee Shop	Bar	Bagel Shop	Pizza Place	French Restaurant	Seafood Restaurant
15	Midtown	Hotel	Sporting Goods Shop	Clothing Store	Coffee Shop	Spa	Theater	Bookstore	American Restaurant	Steakhouse	Café
16	Murray Hill	Hotel	American Restaurant	Sandwich Place	Japanese Restaurant	Coffee Shop	Burger Joint	Taco Place	Gym / Fitness Center	Italian Restaurant	Juice Bar
17	Chelsea	Coffee Shop	Bakery	American Restaurant	Art Gallery	Hotel	Wine Shop	Ice Cream Shop	French Restaurant	Seafood Restaurant	Bookstore
18	Greenwich Village	Italian Restaurant	Sushi Restaurant	Clothing Store	Cosmetics Shop	Bakery	Boutique	Dessert Shop	Indian Restaurant	Coffee Shop	Sandwich Place
19	East Village	Bar	Speakeasy	Pizza Place	Mexican Restaurant	Vegetarian / Vegan Restaurant	Wine Bar	Ice Cream Shop	Italian Restaurant	Korean Restaurant	Cocktail Bar
22	Little Italy	Café	Bakery	Chinese Restaurant	Italian Restaurant	Pizza Place	Coffee Shop	Cocktail Bar	Mediterranean Restaurant	Salon / Barbershop	Sandwich Place
23	Soho	Clothing Store	Italian Restaurant	Boutique	Shoe Store	Mediterranean Restaurant	Hotel	Sporting Goods Shop	Bakery	Women's Store	Coffee Shop
25	Manhattan Valley	Bar	Coffee Shop	Mexican Restaurant	Playground	Pizza Place	Vietnamese Restaurant	Thai Restaurant	Falafel Restaurant	Cuban Restaurant	Hawaiian Restaurant
27	Gramercy	Bar	Bagel Shop	Italian Restaurant	Pizza Place	Playground	Coffee Shop	Diner	Cocktail Bar	Mexican Restaurant	Ice Cream Shop
29	Financial District	Coffee Shop	Pizza Place	Cocktail Bar	Bar	Italian Restaurant	Park	Steakhouse	Event Space	American Restaurant	Salad Place
30	Carnegie Hill	Coffee Shop	Café	Wine Shop	Yoga Studio	Bar	Bookstore	Cosmetics Shop	French Restaurant	Gym / Fitness Center	Gym
31	Noho	Pizza Place	Mexican Restaurant	Cocktail Bar	French Restaurant	Hotel	Wine Bar	Sushi Restaurant	Rock Club	Bakery	Italian Restaurant
33	Midtown South	Korean Restaurant	Hotel	Dessert Shop	Hotel Bar	Coffee Shop	Flower Shop	Gym / Fitness Center	Cosmetics Shop	American Restaurant	Japanese Restaurant
34	Sutton Place	Italian Restaurant	Coffee Shop	Gym / Fitness Center	Pizza Place	Park	Furniture / Home Store	Indian Restaurant	Burger Joint	Steakhouse	Gym
38	Flatiron	Japanese Restaurant	Italian Restaurant	New American Restaurant	Spa	American Restaurant	Mexican Restaurant	Sporting Goods Shop	Mediterranean Restaurant	Yoga Studio	Clothing Store

Cluster 3 and 4

Cluster 3

manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 2, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[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 Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	
37 Stuyvesant Town	Bar	Park	Boat or Ferry	Coffee Shop	Pet Service	Fountain	Bistro	Gym / Fitness Center	Harbor / Marina	Farmers Market	

Cluster 4

manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 3, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[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 Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	
0 Marble Hill	Coffee Shop	Sandwich Place	Gym	Yoga Studio	Tennis Stadium	Pizza Place	Diner	Discount Store	Pharmacy	Donut Shop	

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 Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1 Chinatown	Chinese Restaurant	Bakery	Cocktail Bar	American Restaurant	Dessert Shop	Hotpot Restaurant	Salon / Barbershop	Optical Shop	Asian Restaurant	Bubble Tea Shop
2 Washington Heights	Café	Bakery	Mobile Phone Shop	Chinese Restaurant	Grocery Store	Sandwich Place	Supplement Shop	Coffee Shop	Tapas Restaurant	Bank
3 Inwood	Café	Mexican Restaurant	Lounge	Restaurant	Bakery	Caribbean Restaurant	Park	Chinese Restaurant	Pizza Place	Wine Bar
4 Hamilton Heights	Pizza Place	Deli / Bodega	Café	Mexican Restaurant	Coffee Shop	Liquor Store	Park	Bakery	Caribbean Restaurant	Sushi Restaurant
5 Manhattanville	Coffee Shop	Mexican Restaurant	Bar	Italian Restaurant	Seafood Restaurant	Deli / Bodega	Cuban Restaurant	Bus Stop	Bus Station	Ramen Restaurant
7 East Harlem	Mexican Restaurant	Bakery	Deli / Bodega	Thai Restaurant	Spa	Latin American Restaurant	Donut Shop	Sandwich Place	Pharmacy	Gym
20 Lower East Side	Café	Chinese Restaurant	Pizza Place	Bakery	Park	Ramen Restaurant	Caribbean Restaurant	Art Gallery	Yoga Studio	Clothing Store
26 Morningside Heights	Park	Bookstore	American Restaurant	Coffee Shop	Café	Deli / Bodega	Burger Joint	Grocery Store	Greek Restaurant	Outdoor Sculpture
36 Tudor City	Mexican Restaurant	Park	Café	Pizza Place	Diner	Gym	Sushi Restaurant	Coffee Shop	Deli / Bodega	Dog Run

According to the results, for Manhattan, there are 3 big clusters and 2 small clusters. I pick the best ones for each cluster and compare them based on Imset's interests. **Battery Park City** is chosen for cluster 1. The reason is that park is the most common venue around that neighborhood. There are also, gym and clothing stores. Similarly, **Upper East Side** is the best option for cluster 2. In addition, **Tudor City** is chosen for cluster 5 because it has both restaurants and a yoga studio. Sports and nutrition is so important for Ismet. Cluster 3 and 4 has unique neighborhoods so there is no need to choose one.

For choosing the best one, we need to compare 5 neighborhoods. Since Battery Park City and Upper East Side does not have the balance between sport facilities and restaurants, they are eliminated. On the other hand, Tudor city is just one block away the coast. However, Stuyvesant Town and Marble Hill are on the coast. Therefore, as there are a lot of sports facilities, different cuisines, coffee shops etc., Marble Hill is the best choice for Ismet. Although it is on the upper side of the Manhattan, there is not a problem for commute. It takes a short time to go to work from there. (Check the blue dot from map!)

4.2.2 Toronto, Ontario

Cluster 1

Neighborhood	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
35 The Danforth East	0	Pizza Place	Convenience Store	Park	Discount Store	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant
61 Lawrence Park	0	Park	Bus Line	Swim School	Wine Shop	Distribution Center	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store
91 Rosedale	0	Park	Playground	Trail	Wine Shop	Diner	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant

Cluster 2

Neighborhood	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	
2	Ragent Park, Harbourfront	1	Coffee Shop	Bakery	Pub	Park	Café	Breakfast Spot	Theater	Distribution Center	Electronics Store	Mexican Restaurant
9	Garden Street, Ryerson	1	Coffee Shop	Clothing Store	Japanese Restaurant	Bubble Tea Shop	Café	Cosmetic Shop	Fast Food Restaurant	Bookstore	Ramen Restaurant	Hotel
15	St. James Town	1	Coffee Shop	Café	Cosmetics Shop	Hotel	Restaurant	Clothing Store	Cocktail Bar	Park	Seafood Restaurant	Creperie
19	The Beaches	1	Health Food Store	Trail	Pub	Wine Shop	Diner	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant
20	Berczy Park	1	Bakery	Coffee Shop	Cocktail Bar	Farmers Market	Pharmacy	Seafood Restaurant	Restaurant	Beer Bar	Cheese Shop	Pub
24	Central Bay Street	1	Coffee Shop	Sandwich Place	Café	Italian Restaurant	Restaurant	Spa	Burger Joint	Japanese Restaurant	Bubble Tea Shop	Salad Place
25	Christie	1	Grocery Store	Café	Park	Restaurant	Italian Restaurant	Athletics & Sports	Candy Store	Nightclub	Baby Store	Coffee Shop
30	Richmond, Adelaide, King	1	Coffee Shop	Clothing Store	Café	Restaurant	Hotel	Gym	Thai Restaurant	Breakfast Spot	Bakery	Bar
31	Dufferin, Dovercourt Village	1	Bakery	Pharmacy	Wine Shop	Grocery Store	Middle Eastern Restaurant	Music Venue	Park	Coffee Shop	Cafe	Brewery
36	Harbourfront East, Union Station, Toronto Islands	1	Coffee Shop	Aquarium	Café	Hotel	Brewery	Fried Chicken Joint	Scenic Lookout	Restaurant	History Museum	Music Venue
37	Little Portugal, Trinity	1	Bar	Vietnamese Restaurant	Café	Restaurant	Bakery	Men's Store	Asian Restaurant	Yoga Studio	Cuban Restaurant	Brewery
41	The Danforth West, Riverdale	1	Greek Restaurant	Coffee Shop	Italian Restaurant	Ice Cream Shop	Café	Furniture / Home Store	Yoga Studio	Caribbean Restaurant	Indian Restaurant	Spa
42	Toronto Dominion Centre, Design Exchange	1	Coffee Shop	Hotel	Café	Restaurant	Salad Place	Japanese Restaurant	Italian Restaurant	Seafood Restaurant	Bakery	Sushi Restaurant
43	Brockton Parkdale Village, Exhibition Place	1	Café	Coffee Shop	Breakfast Spot	Gym	Intersection	Performing Arts Venue	Pet Store	Grocery Store	Nightclub	Climbing Gym
47	India Bazaar, The Beaches West	1	Sandwich Place	Park	Pub	Movie Theater	Liquor Store	Burrito Place	Fast Food Restaurant	Fish & Chips Shop	Restaurant	Italian Restaurant
48	Commerce Court, Victoria Hotel	1	Coffee Shop	Restaurant	Café	Hotel	Gym	Japanese Restaurant	Seafood Restaurant	Deli / Bodega	Bakery	American Restaurant
54	Studio District	1	Coffee Shop	Café	Bakery	Brewery	American Restaurant	Gastropub	Yoga Studio	Italian Restaurant	Fun Market	Bookstore
67	Davistown North	1	Food & Drink Shop	Dance Studio	Gym / Fitness Center	Playground	Park	Breakfast Spot	Sandwich Place	Department Store	Hotel	Donut Shop
69	High Park, The Junction South	1	Park	Mexican Restaurant	Café	The Restaurant	Flea Market	Bar	Fried Chicken Joint	Italian Restaurant	Speakeasy	Fast Food Restaurant
73	North Toronto West	1	Clothing Store	Coffee Shop	Yoga Studio	Ice Cream Shop	Chinese Restaurant	Diner	Rental Car Location	Restaurant	Café	Salon / Barbershop
74	The Annex, North Midtown, Yorkville	1	Sandwich Place	Café	Coffee Shop	Pizza Place	Pub	Liquor Store	Park	Burger Joint	Donut Shop	Middle Eastern Restaurant
75	Parkdale, Roncesvalles	1	Breakfast Spot	Gift Shop	Restaurant	Dessert Shop	Eastern European Restaurant	Bar	Dog Run	Italian Restaurant	Movie Theater	Bookstore
79	Dundasville	1	Sandwich Place	Dessert Shop	Pizza Place	Gym	Café	Italian Restaurant	Sushi Restaurant	Coffee Shop	Brewery	Seafood Restaurant
80	University of Toronto, Harbord	1	Café	Bookstore	Bar	Japanese Restaurant	Sandwich Place	Bakery	Poutine Place	Pub	Dessert Shop	Sushi Restaurant
81	Ranymade, Swansea	1	Coffee Shop	Café	Sushi Restaurant	Italian Restaurant	Pub	Pizza Place	Dessert Shop	Sandwich Place	Burrito Place	Restaurant
84	Kensington Market, Chinatown, Grange Park	1	Café	Bar	Bakery	Vietnamese Restaurant	Vegetarian / Vegan Restaurant	Coffee Shop	Gaming Cafe	Dumpling Restaurant	Caribbean Restaurant	Burger Joint
86	Summerhill West, Rathnelly, South Hill, Forest...	1	Coffee Shop	Pizza Place	Pub	Liquor Store	Restaurant	Bank	Fried Chicken Joint	Supermarket	Sushi Restaurant	Light Rail Station
87	CN Tower, King & Spadina, Railway Lands, Har...	1	Airport Service	Plane	Bar	Coffee Shop	Rental Car Location	Sculpture Garden	Boat or Ferry	Harbor / Marina	Airport Lounge	Airport Food Court
92	Enclave of MSE	1	Coffee Shop	Seafood Restaurant	Italian Restaurant	Cocktail Bar	Hotel	Pub	Café	Bakery	Japanese Restaurant	Restaurant
96	St. James Town, Cabbagetown	1	Coffee Shop	Bakery	Italian Restaurant	Pizza Place	Restaurant	Chinese Restaurant	Café	Pub	Yoga Studio	Liquor Store
97	First Canadian Place, Underground City	1	Coffee Shop	Café	Hotel	Gym	Japanese Restaurant	Restaurant	Deli / Bodega	Steakhouse	Asian Restaurant	Bakery
99	Church and Wellesley	1	Coffee Shop	Japanese Restaurant	Sushi Restaurant	Gay Bar	Restaurant	Pub	Men's Store	Mediterranean Restaurant	Hotel	Fast Food Restaurant
100	Enclave of MAL	1	Yoga Studio	Spa	Gym / Fitness Center	Garden Center	Garden	Fast food Restaurant	Light Rail Station	Farmers Market	Comic Shop	Park

for Toronto, Ontario, cluster 2 is the biggest one. There are 33 neighbourhoods in this cluster. This means that there are 33 similar neighborhoods in Toronto Center, based on venue statistics. On the other hand, cluster 3, 4 and 5 has unique neighbourhoods and cluster 1 has 3 neighbourhoods. We need to select one neighbourhood from each cluster again. Starting with first one, **Rosedale** is the most suitable neighbourhood to live for Ismet. **The Beaches** is definitely the one to choose as it has a beach and a lot of different restaurants, a trail and a event space. It is so good for both sports and other activities. **Moore Park, Summerhill East** and **Roselawn** are eliminated from the battle because they do not seem very suitable. Contrarily, **Forest Hill North West** from cluster 5 is very suitable for the battle of neighbourhoods of different clusters.

So, we have 3 neighborhoods for the final decision. The winner of the battle is **The Beaches**. The reason for that is the main location. It is at the coast of Toronto, Ontario. It is a little bit away from the center. You can feel the nature there. Also, according to the results, it has a lot of similarity with most of the neighborhoods in Toronto. Therefore, Ismet can find anything he wants there. It has a healthy food store, trail, pub, wine shop, restaurants, electronic store and event space. Great choice for Ismet!

Cluster 3, 4 and 5

Cluster 3

```
[56]: toronto_merged.loc[toronto_merged['Cluster Labels'] == 2, toronto_merged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

Neighborhood	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	
83	Moore Park, Summerhill East	2	Intersection	Restaurant	Wine Shop	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Dumpling Restaurant

Cluster 4

```
[57]: toronto_merged.loc[toronto_merged['Cluster Labels'] == 3, toronto_merged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

Neighborhood	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	
62	Roselawn	3	Fast Food Restaurant	Garden	Wine Shop	Distribution Center	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant

Cluster 5

```
[58]: toronto_merged.loc[toronto_merged['Cluster Labels'] == 4, toronto_merged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

Neighborhood	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	
68	Forest Hill North & West	4	Trail	Mexican Restaurant	Jewelry Store	Sushi Restaurant	Wine Shop	Distribution Center	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant

4.3 Final Decision

Battle of the neighbourhoods has ended, for now. Only 2 neighbourhoods from Toronto and Manhattan left. These are **The Beaches, Toronto** and **Marble Hill, Manhattan**. These are very similar neighborhoods in terms of venues but the location is so important. Since Ismet is coming from an island, he wants to feel the island atmosphere if it is possible. Therefore, **The Beaches** is the winner here. Even though Marble Hill is a very good place to live, it does not have a beach. The beach is the thing to supply the island atmosphere there. Therefore, congratulations to **The Beaches, Toronto** for the win!! It is so clear that Ismet is going to be very happy at this place.

5 Discussion and Conclusion

In this report, I tried to help Ismet to choose the best neighbourhood to live. I want to discuss some points before conclusion. Largeness of observations provide better examination in this context. Better examination means more accurate results. Therefore, I answer the questions that are asked in Introduction. Firstly, these two cities have some similarities based on venues but they are different in general. Based on results, Toronto is much more suitable for Ismet. Although both cities has almost similar venues in chosen neighborhoods, the size of green areas and having open sea have a big effect on this decision.(check the maps!). Cluster Maps showed that there are a lot of similar neighbourhoods in Toronto, Ontario. I make further examination and looked the neighbourhoods in each cluster more deep and analyze them. I found that **The Beaches, Toronto, Ontario** is the best place for Ismet to live. Data science help him to choose the best option. There is no doubt that he is going to be happy.

6 References

Garbade,J.Micheal,13 Sep 2018, retrieved from

<https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1>