CAPSTONE PROJECT – THE BATTLE OF NEIGHBOURHOODS

Tonny Wong

Contents

Table of Figures

Table of Tables

1.0 PROBLEM & BACKGROUND

New York and Toronto are two of the top-most famous cities located in US and Canada, respectively. Both cities are multicultural and financial hub to their respective countries. Nevertheless, they are also the hot spot for most of the tourist to be visited. However, it is always facing difficulty to opt in between different places for having a wonderful vacation especially in the condition that both high similarity places are in the visitation list. The project here is to present helpful information related to tourism to explore and compare how much they are similar or dissimilar from the aspects as a tourist viewpoint, in term of food, accommodation, point of interests and so on.

Tourism is one of the most important pillars of the economy in most of the countries. Tourist' expectations when they are visiting a particular destination are mostly related to its features such as gastronomy, culture, architecture, infrastructure, landscape, entertainment, and many more. These several features attract tourist visiting to the destination and contribute to the overall tourism experience. Every famous city possessing their own uniqueness in specific features and worthwhile to be visited at least once in the life.

Nevertheless, it is sincerely appreciated for the up-to-date technology that most of the information is widely shared in term of location of places all over the world and obtained easily by a fingertip for further exploration. By wisely utilising the technology, we can explore the specific destinations with our preferably attractive features and comparing between them to come out with the results based on our own interested. This project mainly benefit groups of people who interested to know more about two analysed destinations such as tour group, business founder, fresh graduates who looking for a job, students who going to have a fantastic travelling, even a family who is planning to have a wonderful tourism.

2.0 DATA SECTION

In this project, each famous borough will be selected from the Toronto City and from New York City respectively for analyses and exploration. Downtown Toronto from Toronto City and Manhattan from New York City be selected for the carrying of this data exploration project.

For the New York City Location Dataset,

- The relevant data (Borough, Neighbourhood, Latitude and Longitude) which saved in JSON format file could be obtained from NYU Spatial Data Repository (https://geo.nyu.edu/catalog/nyu_2451_34572). Since the original dataset file is in JSON format, it is needed to download the JSON file and convert it into *pandas* dataframe.
- Foursquare API is one of the location data providers used to retrieve the data from the Foursquare database through the application of API to explore a location even to get trending venues around a point of interest. We use the venues group with explore endpoint, pass the latitude and longitude coordinates of neighbourhoods in Manhattan along with our credentials, then make the call to the database to get a list of the popular spots around the neighbourhoods in return. The explored categories of venues which are in the *items* key could be obtained, clean the JSON file and structured them into *pandas* dataframe.
- With the important data such as Borough, Neighbourhoods with their geographical coordinates, and the venues categories, the analysing, clustering and examining of the data processes could be carried out.

For the Toronto City Location Dataset,

- The relevant data such as Postal Code, Borough and Neighbourhood are scrapped from the Wikipedia webpage (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M) using Beautiful Soup package. The read data then transform into *pandas* dataframe.
- In order to utilize the Foursquare location data, we need to get the geographical coordinates of each neighbourhood. These data in csv file could be accessed through https://cocl.us/Geospatial_data then convert them into *pandas* dataframe.
- With the application of Foursquare API, we use the venues group with explore endpoint, pass the latitude and longitude coordinates of neighbourhoods in Downtown Toronto along with our credentials, then make the call to the database to get a list of the popular spots around the neighbourhoods in return. The explored categories of venues which are in the *items* key could be obtained, clean the JSON file and structured them into *pandas* dataframe.
- With the important data such as Borough, Neighbourhoods with their geographical coordinates, and the venues categories, the analysing, clustering and examining of the data processes could be carried out.

3.0 Methodology

This project is carried out based on the Methodology as follow:

1) Collecting and cleaning the data
2) Analysing the data
3) Machine learning (K-Means Clustering Algorithm)
4) Data visualization

Noticed that first two steps for two cities borough are slightly different due to the data obtained from different sources with different methods. It is always necessary to install or download all the dependencies that are needed for the project into the notebook.

Step 1: Collect and cleaning the location data

As the data for two cities borough are obtained from two different sources, therefore there is slightly different from one another in this step.

For New York City, the location data is downloaded as JSON format file and convert into the *pandas* dataframe. The conversion of location data with specific geographical coordinates data is accomplished by two simple steps, create an empty dataframe and insert the data from JSON format file into the empty dataframe through loop method. All data is cleaned accordingly. For better understanding of the collected data, the Manhattan borough with its neighbourhoods are visualized by creating a map of the borough using Folium library, as shown in Figure 1.

For Toronto City, the location data such as Postal Code, Borough and Neighbourhoods is scraped from Wikipedia webpage using Beautiful Soup package. However, their geographical coordinates are downloaded separately as CSV format, then combined the location data with their specific geographical coordinates using merge method. All data is cleaned accordingly. For better understanding of the collected data, the Downtown Toronto borough with its neighbourhoods are visualized by creating a map of the borough using Folium library, as shown in Figure 2.
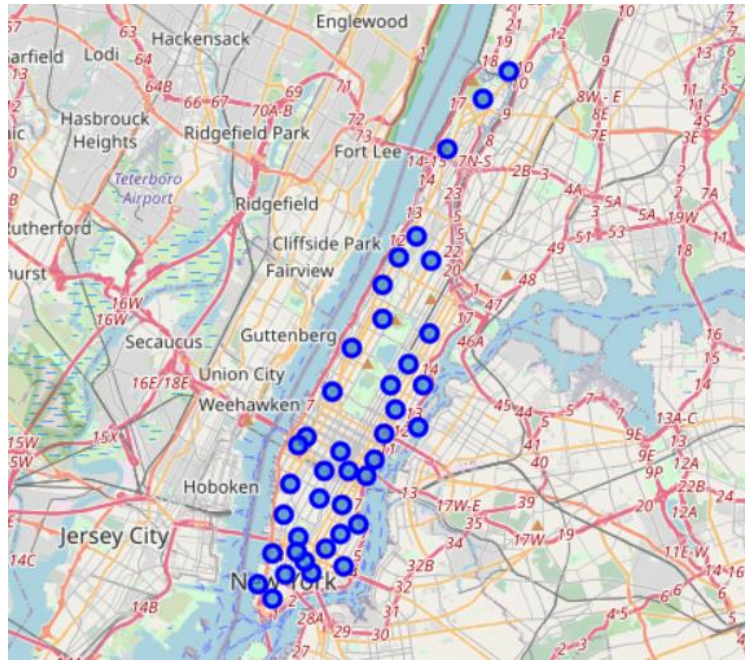
Figure 1: Visualized map for Manhattan, New York City

Figure 2: Visualized map for Downtown Toronto, Toronto City

With the implementation of the Explore function through Foursquare API location data provider, the several venues with their relevant information for each neighbourhood returned by calling of Foursquare API. All extracted information converted into *pandas* dataframe with the help of generated user defined function.

Step 2: Analysing the data

Throughout Step 1, the informative data such as Borough, Neighbourhoods, Geographical Coordinates (latitude and longitude) and several venues with their relevant information for each neighbourhood

are been collected. In this step, the collected venues for each neighbourhood, up to 3004 numbers, are analysed and grouped accordingly. The total numbers of 3004 venues are grouped by neighbourhood and by taking the mean of the frequency of the occurrence of each category. Then, the data of collected venues for neighbourhood are sorted based on their top ten most common venue categories and structured into a new dataframe.

Step 3: Machine learning

K-Means clustering algorithm is implemented on the top ten most common venue categories to identify the clusters which venue categories are high similarity with one another. There is five numbers of cluster initialized in this project, labelled separately from number zero (0) to number four (4) based on their similarity.

Step 4: Data visualization

The clustered data in Step 3 are then visualized by creating a map using Folium library. Unlike the map created in Step 1, this map is created based on the set of five different colour scheme to illustrate five different clusters on the map for better effect of data visualization.

4.0 Result

4.1 For Manhattan, New York City

The map shown below as the clustered map of neighbourhoods in Manhattan, New York City, which visualized through Folium library implementation. Following is a brief discussion of the results observed through the exploration and analysing of the clustered neighbourhoods in Manhattan.
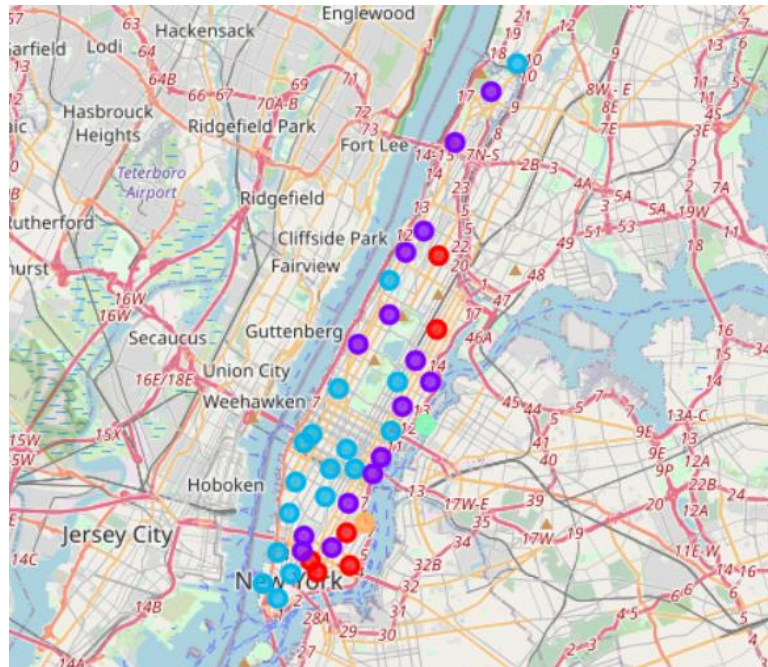

Figure 3: Clustered Map for Manhattan, New York City

- Cluster 0 (Red) Miscellaneous/Night Life

Table 1: Cluster 0 for Neighbourhoods in Manhattan

	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	Cocktail Bar	Bubble Tea Shop	Bakery	Coffee Shop	Salon / Barbershop	Vietnamese Restaurant	Optical Shop	American Restaurant	Spa
6	Central Harlem	African Restaurant	Fried Chicken Joint	American Restaurant	Bar	Cosmetics Shop	French Restaurant	Seafood Restaurant	Chinese Restaurant	Café	Tapas Restaurant
7	East Harlem	Mexican Restaurant	Bakery	Latin American Restaurant	Thai Restaurant	Deli / Bodega	Restaurant	Beer Bar	French Restaurant	Liquor Store	Steakhouse
19	East Village	Cocktail Bar	Pizza Place	Mexican Restaurant	Coffee Shop	Bar	Juice Bar	Wine Bar	Japanese Restaurant	Ramen Restaurant	Speakeasy
20	Lower East Side	Chinese Restaurant	Park	Cocktail Bar	Café	Ramen Restaurant	Art Gallery	Pet Café	Performing Arts Venue	Pharmacy	Ice Cream Shop
22	Little Italy	Bubble Tea Shop	Chinese Restaurant	Bakery	Spa	Mediterranean Restaurant	Italian Restaurant	Thai Restaurant	Hotel	Ice Cream Shop	Coffee Shop

From Cluster 0 which represented as red colour circles, it could be observed that this area mainly covered night life activities such as several kinds of bar, tea-time shop such as bubble tea shop and ice cream shop which could be attended after dinner at several kind of restaurant. Therefore, this cluster can be categorised as night life or miscellaneous activities cluster.

- Cluster 1 (Purple) Gastronomy

Table 2: Cluster 1 for Neighbourhoods in Manhattan

	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
2	Washington Heights	Café	Bakery	Pizza Place	Mobile Phone Shop	Chinese Restaurant	Grocery Store	Latin American Restaurant	Donut Shop	Sandwich Place	Supplement Shop
3	Inwood	Mexican Restaurant	Café	Lounge	Pizza Place	Restaurant	Deli / Bodega	Chinese Restaurant	Spanish Restaurant	Caribbean Restaurant	Park
4	Hamilton Heights	Pizza Place	Coffee Shop	Café	Deli / Bodega	Mexican Restaurant	Cocktail Bar	Indian Restaurant	Sushi Restaurant	Park	Yoga Studio
5	Manhattanville	Coffee Shop	Seafood Restaurant	Italian Restaurant	Deli / Bodega	Park	Sushi Restaurant	Mexican Restaurant	Café	Bike Trail	Boutique
9	Yorkville	Italian Restaurant	Coffee Shop	Gym	Sushi Restaurant	Bar	Deli / Bodega	Mexican Restaurant	Japanese Restaurant	Diner	Wine Shop
10	Lenox Hill	Coffee Shop	Italian Restaurant	Pizza Place	Café	Sushi Restaurant	Cocktail Bar	Gym / Fitness Center	Gym	Burger Joint	Salad Place
12	Upper West Side	Italian Restaurant	Coffee Shop	Dessert Shop	Bar	Middle Eastern Restaurant	Bookstore	Sushi Restaurant	Mediterranean Restaurant	Pizza Place	Bakery
18	Greenwich Village	Italian Restaurant	Coffee Shop	Café	Gym	Indian Restaurant	Ice Cream Shop	Bakery	Comedy Club	Clothing Store	Gourmet Shop
23	Soho	Italian Restaurant	Mediterranean Restaurant	Coffee Shop	Sandwich Place	French Restaurant	Clothing Store	Gym	Bakery	Café	Spa
23	Soho	Italian Restaurant	Mediterranean Restaurant	Coffee Shop	Sandwich Place	French Restaurant	Clothing Store	Gym	Bakery	Café	Spa
25	Manhattan Valley	Coffee Shop	Spa	Pizza Place	Bar	Mexican Restaurant	Grocery Store	Playground	Park	Noodle House	Latin American Restaurant
27	Gramercy	Italian Restaurant	Bar	Coffee Shop	Playground	Pizza Place	Mexican Restaurant	Grocery Store	Bagel Shop	Ice Cream Shop	Diner
30	Carnegie Hill	Coffee Shop	Pizza Place	Yoga Studio	Gym	Wine Shop	Bar	Bookstore	Café	Japanese Restaurant	Grocery Store
31	Noho	Italian Restaurant	Coffee Shop	Pizza Place	Japanese Restaurant	Grocery Store	Sandwich Place	Mexican Restaurant	Wine Shop	Gift Shop	Tea Room
35	Turtle Bay	Coffee Shop	Park	Italian Restaurant	Café	Deli / Bodega	Seafood Restaurant	Wine Bar	Sushi Restaurant	French Restaurant	Karaoke Bar
36	Tudor City	Café	Park	Mexican Restaurant	Deli / Bodega	Diner	Pizza Place	Dog Run	Asian Restaurant	Bar	Garden

From Cluster 1 which represented as purple colour circles, it could be observed that this area can be categorised as gastronomy zone. In this area have many types of restaurant which visitor could be hunted for foods. Café, Italian restaurant, Mexican restaurant, Latin American restaurant, Mediterranean restaurant and many more, are included in this huge zone.

- Cluster 2 (Blue) Entertainment & Accommodation

Table 3: Cluster 2 for Neighbourhoods in Manhattan

	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	Sandwich Place	Coffee Shop	Gym	Ice Cream Shop	Department Store	Pharmacy	Diner	Discount Store	Donut Shop	Bank
8	Upper East Side	Italian Restaurant	Bakery	Gym / Fitness Center	Spa	American Restaurant	Exhibit	Hotel	Juice Bar	Yoga Studio	Wine Shop
13	Lincoln Square	Italian Restaurant	Café	Plaza	Gym / Fitness Center	Theater	Performing Arts Venue	Concert Hall	Wine Shop	French Restaurant	Grocery Store
14	Clinton	Theater	Coffee Shop	Gym / Fitness Center	Hotel	Gym	Italian Restaurant	Wine Shop	Sandwich Place	Spa	American Restaurant
15	Midtown	Coffee Shop	Hotel	Theater	Pizza Place	Cuban Restaurant	Bakery	Japanese Restaurant	Clothing Store	Cosmetics Shop	Spa
16	Murray Hill	Hotel	Sandwich Place	Coffee Shop	Gym / Fitness Center	Steakhouse	Japanese Restaurant	Pizza Place	Chinese Restaurant	Indian Restaurant	Juice Bar
17	Chelsea	Art Gallery	Coffee Shop	Café	Ice Cream Shop	American Restaurant	Market	Seafood Restaurant	Boutique	Cupcake Shop	Cycle Studio
21	Tribeca	Italian Restaurant	Park	Wine Bar	Café	Spa	Hotel	Coffee Shop	Bakery	Steakhouse	Greek Restaurant
24	West Village	Italian Restaurant	American Restaurant	Wine Bar	New American Restaurant	Coffee Shop	Jazz Club	Bakery	Park	Cocktail Bar	Pizza Place
26	Morningside Heights	Park	Coffee Shop	Bookstore	American Restaurant	Deli / Bodega	Burger Joint	Sandwich Place	Tennis Court	New American Restaurant	Supermarket
28	Battery Park City	Park	Coffee Shop	Hotel	Memorial Site	Gym	Boat or Ferry	Gourmet Shop	Food Court	Shopping Mall	Wine Shop
29	Financial District	Coffee Shop	Hotel	Pizza Place	American Restaurant	Sandwich Place	Salad Place	Gym / Fitness Center	Juice Bar	Café	Cocktail Bar
32	Civic Center	French Restaurant	Coffee Shop	American Restaurant	Park	Hotel	Spa	Sushi Restaurant	Gym / Fitness Center	Yoga Studio	Wine Shop
33	Midtown South	Korean Restaurant	Hotel	Japanese Restaurant	Burger Joint	Café	Coffee Shop	Hotel Bar	Gym / Fitness Center	Lounge	Dessert Shop
34	Sutton Place	Italian Restaurant	Coffee Shop	Gym / Fitness Center	Gym	Park	Bagel Shop	Hotel	Bakery	Furniture / Home Store	Grocery Store
38	Flatiron	Gym / Fitness Center	Italian Restaurant	Salon / Barbershop	Wine Shop	Park	Outdoor Sculpture	Cosmetics Shop	Gym	Grocery Store	Mediterranean Restaurant
39	Hudson Yards	Hotel	Gym / Fitness Center	American Restaurant	Café	Italian Restaurant	Restaurant	Park	Gym	Boat or Ferry	Coffee Shop

From Cluster 2 which represented as blue colour circles, it could be observed that this area can be categorised as entertainment & accommodation zone. Hotels accommodation are concentrated at this zone and several types of entertainment such as exhibit, jazz club, theatre, plaza and so on are located at here as well.

- Cluster 3 (Green) Nature

Table 4: Cluster 3 for Neighbourhoods in Manhattan

	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
11	Roosevelt Island	Playground	Park	Gym	Dry Cleaner	Greek Restaurant	Coffee Shop	Outdoors & Recreation	Sandwich Place	Scenic Lookout	School

From Cluster 3 which represented as green colour circles, consists of park, scenic lookout, outdoors & recreation which can be categorised mainly as nature activities or relax zone. People who come to this area mainly for relaxation and enjoy the natural environment.

- Cluster 4 (Yellow) Sport Hub

Table 5: Cluster 4 for Neighbourhoods in Manhattan

	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	Park	Playground	Gas Station	Gym	Gym / Fitness Center	Baseball Field	Cocktail Bar	Harbor / Marina	German Restaurant	Bistro

From Cluster 4 which represented as yellow colour circles, includes several places that could conduct sport activities. These places such as baseball field, gym, fitness centre and others. Therefore, this zone been classified as sport hub.

4.2 For Downtown Toronto, Toronto City

The map shown below as the clustered map of neighbourhoods in Downtown Toronto, Toronto City, which visualized through Folium library implementation. Following is a brief discussion of the results observed through the exploration and analysing of the clustered neighbourhoods in Downtown Toronto.
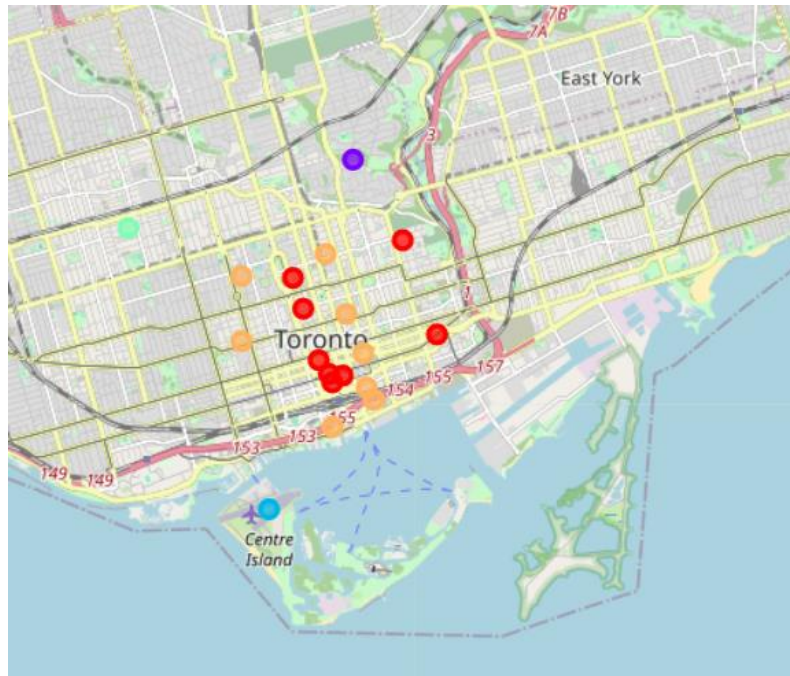

Figure 4: Clustered Map for Downtown Toronto, Toronto City

- Cluster 0 (Red) Gastronomy

Table 6: Cluster 0 for Neighbourhoods in Downtown Toronto

	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
1	St. James Town, Cabbagetown	0	Restaurant	Italian Restaurant	Coffee Shop	Café	Bakery	Liquor Store	Indian Restaurant	Japanese Restaurant	Deli / Bodega	Jewelry Store
3	Regent Park, Harbourfront	0	Coffee Shop	Park	Breakfast Spot	Theater	Bakery	Event Space	Performing Arts Venue	Café	Pub	Restaurant
7	Central Bay Street	0	Coffee Shop	Café	Seafood Restaurant	Sandwich Place	Ramen Restaurant	Poke Place	Bubble Tea Shop	Park	Chinese Restaurant	Modern European Restaurant
8	Richmond, Adelaide, King	0	Coffee Shop	Café	Pizza Place	Sushi Restaurant	Concert Hall	Lounge	Japanese Restaurant	Colombian Restaurant	Seafood Restaurant	Deli / Bodega
10	Toronto Dominion Centre, Design Exchange	0	Coffee Shop	Café	Restaurant	Japanese Restaurant	Hotel	Museum	Gastropub	Pub	Concert Hall	Deli / Bodega
11	Commerce Court, Victoria Hotel	0	Café	Gastropub	Restaurant	Japanese Restaurant	Coffee Shop	Deli / Bodega	Sandwich Place	Seafood Restaurant	Pub	Ice Cream Shop
16	First Canadian Place, Underground city	0	Café	Coffee Shop	Restaurant	Seafood Restaurant	Tea Room	Steakhouse	Gastropub	Pub	Bookstore	Deli / Bodega
18	Queen's Park, Ontario Provincial Government	0	Coffee Shop	Sushi Restaurant	Yoga Studio	Diner	Bar	Smoothie Shop	Beer Bar	Distribution Center	Sandwich Place	Italian Restaurant

From Cluster 0 which represented as red colour circles, it could be observed that many of the eating and drinking places are in this area. Restaurant such as Indian restaurant, Italian restaurant, seafood restaurant, café, and so on, and beverage shop includes coffee shop, smoothie, bar, and many more, are all covered here. Therefore, this cluster can be classified as gastronomy zone.

- Cluster 1 (Purple) Nature

Table 7: Cluster 1 for Neighbourhoods in Downtown Toronto

	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
0	Rosedale	1	Park	Playground	Trail	Comfort Food Restaurant	Deli / Bodega	Dance Studio	Creperie	Cosmetics Shop	Concert Hall	Comic Shop

From Cluster 1 which represented as purple colour circles, could be observed that it is closer to the greenery area in the map as shown in Figure 3. This unique location consists of places such as park, and trail which is ideal place for people to carry out activities that enjoy the natural atmosphere. Therefore, this cluster been categorised as nature activities zone.

- Cluster 2 (Blue) Transportation Hub

Table 8: Cluster 2 for Neighbourhoods in Downtown Toronto

	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
14	CN Tower, King and Spadina, Railway Lands, Har...	2	Airport Lounge	Airport Service	Airport Terminal	Airport	Plane	Boutique	Coffee Shop	Bar	Sculpture Garden	Boat or Ferry

From Cluster 2 which represented as blue colour circles, could be observed that this area located at centre island in the map, and mainly consists of airport, plane and boat or ferry. By this, it could be classified as transportation hub in Downtown Toronto.

- Cluster 3 (Green) Miscellaneous

Table 9: Cluster 3 for Neighbourhoods in Downtown Toronto

	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
17	Christie	3	Grocery Store	Café	Park	Italian Restaurant	Nightclub	Candy Store	Restaurant	Diner	Baby Store	Athletics & Sports

From Cluster 3 which represented as green colour circles, covered several kinds of places that can conduct many types of activities. It includes café, restaurant, grocery store, night club, sports, which able to get all necessities in daily life. Therefore, this cluster been categorised as miscellaneous.

- Cluster 4 (Yellow) Entertainment & Accommodation

Table 10: Cluster 4 for Neighbourhoods in Downtown Toronto

	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	Church and Wellesley	4	Café	Dance Studio	Smoke Shop	Indian Restaurant	Japanese Restaurant	Salon / Barbershop	Bookstore	Restaurant	Juice Bar	Ramen Restaurant
4	Garden District, Ryerson	4	Café	Theater	Coffee Shop	Burger Joint	Electronics Store	Shopping Mall	Music Venue	Plaza	Ramen Restaurant	Hotel
5	St. James Town	4	Gastropub	Café	Cocktail Bar	Coffee Shop	BBQ Joint	New American Restaurant	Park	Middle Eastern Restaurant	Restaurant	Cosmetics Shop
6	Berczy Park	4	Beer Bar	Coffee Shop	Seafood Restaurant	Cocktail Bar	French Restaurant	Breakfast Spot	Creperie	Liquor Store	Bistro	Concert Hall
9	Harbourfront East, Union Station, Toronto Islands	4	Plaza	Park	Hotel	Italian Restaurant	Lounge	Roof Deck	Salad Place	Bubble Tea Shop	Japanese Restaurant	Skating Rink
12	University of Toronto, Harbord	4	Café	Bookstore	Restaurant	Bar	Bakery	Japanese Restaurant	Italian Restaurant	Yoga Studio	Dessert Shop	Beer Bar
13	Kensington Market, Chinatown, Grange Park	4	Café	Mexican Restaurant	Bakery	Vietnamese Restaurant	Comfort Food Restaurant	Record Shop	Pizza Place	Cheese Shop	Organic Grocery	Cocktail Bar
15	Stn A PO Boxes	4	Café	Restaurant	Seafood Restaurant	Beer Bar	Cocktail Bar	Jazz Club	Hotel	Park	Cheese Shop	Museum

From Cluster 4 which represented as yellow colour circles, it could be observed that this area covered dance studio, bookstore, gastropub, concert hall, and many more. It also consists of hotel that is accommodation area. Therefore, this cluster is classified as entertainment and accommodation zone in Downtown Toronto.

5.0 Discussion

From the obtained clustering results, it could be observed the categorization of the top-most common venues in Manhattan and Downtown Toronto is approximately the same. For Manhattan, there are group such as Miscellaneous/Night life, Gastronomy, Entertainment & Accommodation, Nature and Sport Hub, whereas Gastronomy, Nature, Transportation Hub, Miscellaneous and Entertainment & Accommodation are the categories for Downtown Toronto.

In term of gastronomy, both boroughs possessing high numbers of Café and coffee shop. However, there is a different among the types of restaurant between them. For Manhattan majority have Italian restaurant, American restaurant, Japanese restaurant, Mexican restaurant, and so on. For Downtown Toronto majority have Indian restaurant, Seafood restaurant, Japanese restaurant, and restaurant (assume that is local restaurant). Therefore, in the sense of food, Manhattan have more choices than Downtown Toronto, and it is island wide.

In term of entertainment, Manhattan is higher numbers of places comparing to Downtown Toronto. Apart from that, Manhattan possess of certain amount of local landmark for visitation which spread around the area while there is one sculpture garden designed in Downtown Toronto that located at airport terminal area. Besides that, Manhattan consists of wine shop, wine bar and cocktail bar whereas Downton Toronto consists of bar, gastropub, cocktail bar that is suitable to have a night life entertainment.

Nevertheless, there are certain location that enable visitors to conduct nature activities in both analysed boroughs. However, Downtown Toronto has a transportation hub which includes both airplanes and boat or ferry in its area whereas Manhattan merely has boat or ferry in its area. Downtown Toronto is more convenient in the sense of transportation.

6.0 Conclusion

Based on the clustering results obtained through the application Foursquare API with explore endpoint, the top-most common venues for both neighbourhoods in Manhattan and Downtown Toronto were studied. It can be concluded that there is up to 80% of similarity between both boroughs in term of top-most common venues categorization. However, there is discrepancies for the types of selection within each cluster categorised. It is highly depending on the point of interested individually. There is recommended to increase the amounts of neighbourhoods to explore and analyse and to obtain clustered result for examining in order to get the higher accuracy result in future.