Analyzing Neighborhoods of Two Cities

Capstone Project – The Battle of the Neighborhoods

Applied Data Science Capstone – IBM/Coursera

By

Rachel Olufunmilayo Olokungbemi

June 19, 2021

Table of Contents

1. Introduction: Business Problem

1.1 Background

In this Capstone Project – The Battle of the Neighborhoods, I chose to analyze neighborhoods of two cities by comparing the neighborhoods of City of Toronto and New York City to determine how similar or dissimilar they are. The City of Toronto and New York City are both culturally and economically diverse, they are the financial capitals of their respective countries. The standard of living in each city is expensive compared to the rest of the country and in terms of population, Toronto is the largest city in Canada while New York City is the largest city in the United States. On the contrary, New York City is a Coastal City while Toronto is a Great Lakes City. NYC is built on series of islands while Toronto is completely built on the mainland. The way of life in NYC is more fast-paced, intense and feels like it has more “energy”. NYC is a lot more crowded, especially in the city center, while Toronto is actually a governmental capital, the public transportation is of low degree compared to NYC’s and really isn’t that extensive. If you don’t live in Toronto’s core, you need a car. Not so in NYC.

1.2 Interested Stakeholders

- Investors who want to know more about each of these cities in order to make informed decisions about the location that is best for their investment would be interested in this analysis.
- Individuals or group of people who want to make decisions about which city meets their preferences of work-lifestyle-balance would definitely benefit.
- Moreover, a contractor who wants to start a business would be helped with the analysis of these two cities in terms of getting some preliminary understanding of the location data that shows the venues in these cities in order to determine venue categories that will best serve their business interest.

2. Data Description

2.1 Data Sources

The Toronto neighborhood data is not readily available on the internet like the New York data, therefore I acquired the data by scraping this Wikipedia page; https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M for all the information needed to explore the neighborhoods.

Moreover, I acquired the New York City dataset provided on the server for this course from https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs/newyork_data.json

I imported the csv file containing the latitude and longitudes of various postal codes in Canada from https://cocl.us/Geospatial_data

I used the Geopy library - Geocoder Python package to get the latitude and longitude values of the City of Toronto and New York City.

Also, I used Foursquare API to explore neighborhoods in Toronto and New York City.

2.2 Data Wrangling

2.2.1 Toronto Dataset

For the city of Toronto, I scraped the Wikipedia page using BeautifulSoup Library of Python, retrieved the URL and created a BeautifulSoup object. With this, I was able to display the raw Wikipedia page contents consisting of postal codes, boroughs and neighborhoods information.

The next process was to display the table contents, I first created an empty list and found the table and the table data, I created a dictionary having three keys; PostalCode, Borough and Neighborhood, I extracted the postal codes containing up to three characters. Then I used split, strip and replace functions to get Borough and Neighborhood information and appended to the list. In the next task, I created a dataframe with the list of table contents. I used the DataFrame function to transform the data into a pandas dataframe with the resulting dataframe shape of 103rows x 3columns.

2.2.2 New York City Dataset

The New York City dataset contains the 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood. This dataset was provided on the server for this course and running a wget command enabled me to access the data. I loaded the data, extracted all the relevant data in the features key and transformed the data of nested Python dictionaries into a pandas dataframe by looping through the data and filling the dataframe one row at a time. The resulting dataframe has all the 5 boroughs and 306 neighborhoods.

2.3 Data Exploration

2.3.1.1 Exploring the Neighborhoods of City of Toronto

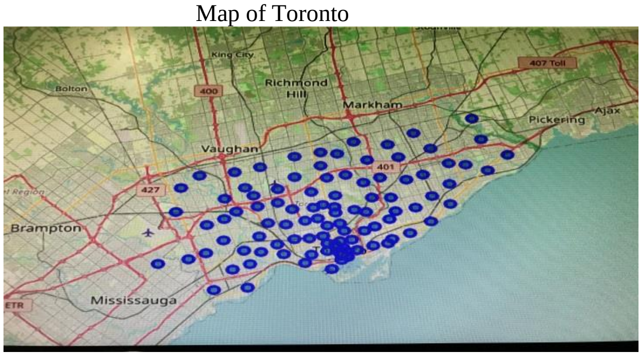
I created a new dataframe that includes the geographical coordinates of each postal code. First I got the geographical coordinates of the neighborhoods by importing the csv file containing the latitudes and longitudes of each postal code in Canada and then merged the geospatial table with the existing dataframe. I displayed the number of boroughs and neighborhoods of the dataframe resulting in 15 boroughs and 103 neighborhoods.

Dataframe of City of Toronto

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Queen's Park	Ontario Provincial Government	43.662301	-79.389494
5	M9A	Etobicoke	Islington Avenue	43.667856	-79.532242
6	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353
7	M3B	North York	Don Mills North	43.745906	-79.352188
8	M4B	East York	Parkview Hill, Woodbine Gardens	43.706397	-79.309937
9	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
10	M6B	North York	Glencairn	43.709577	-79.445073
11	M9B	Etobicoke	West Deane Park, Princess Gardens, Martin Grov...	43.650943	-79.554724
12	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek	43.784535	-79.160497
13	M3C	North York	Don Mills South	43.725900	-79.340923
14	M4C	East York	Woodbine Heights	43.695344	-79.318389
15	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
16	M6C	York	Humewood-Cedarvale	43.693781	-79.428191
17	M9C	Etobicoke	Eringate, Bloordale Gardens, Old Burnhamthorpe...	43.643515	-79.577201
18	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711
19	M4E	East Toronto	The Beaches	43.676357	-79.293031

I used the geopy library to get the latitude and longitude values of the City of Toronto, the geographical coordinates are 43.6534817, -79.3839347. I then use these coordinates to create the map of Toronto with neighborhoods superimposed on top.

Map of Toronto



2.3.1.2 Exploring the Neighborhoods in Downtown Toronto

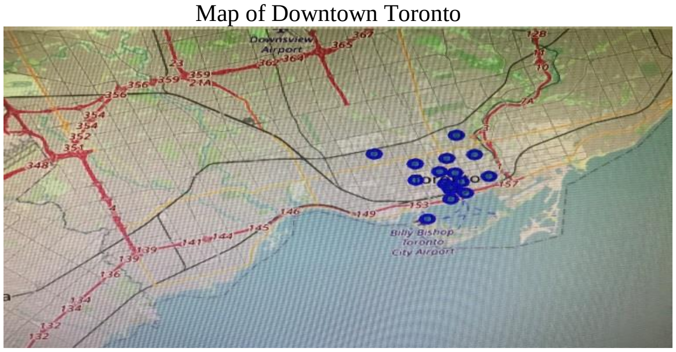
To simplify the above map, I decided to explore only the boroughs of Downtown Toronto. In that instance, I sliced the original dataframe and created a new dataframe of the Downtown data.

Dataframe of Downtown Toronto

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
1	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
2	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
3	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306
4	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383
5	M6G	Downtown Toronto	Christie	43.669542	-79.422564
6	M5H	Downtown Toronto	Richmond, Adelaide, King	43.650571	-79.384568
7	M5J	Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	43.640816	-79.381752
8	M5K	Downtown Toronto	Toronto Dominion Centre, Design Exchange	43.647177	-79.381576
9	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817
10	M5S	Downtown Toronto	University of Toronto, Harbord	43.662696	-79.400049
11	M5T	Downtown Toronto	Kensington Market, Chinatown, Grange Park	43.653206	-79.400049
12	M5V	Downtown Toronto	CN Tower, King and Spadina, Railway Lands, Har...	43.628947	-79.394420
13	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529
14	M4X	Downtown Toronto	St. James Town, Cabbagetown	43.667967	-79.367675
15	M5X	Downtown Toronto	First Canadian Place, Underground city	43.648429	-79.382280

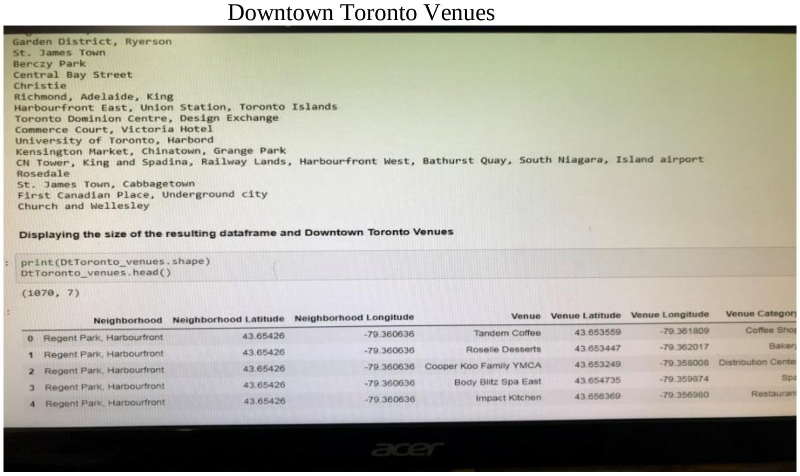
Using the geopy library, I got the geographical coordinates of Downtown Toronto, 43.6541737, -79.38081162653639 and then I use these coordinates to create the map to visualize the neighborhoods of Downtown Toronto.

Map of Downtown Toronto



I utilized the Foursquare API to explore the neighborhoods of Downtown Toronto and segment them. I used a function to get the nearby venues of all the neighborhoods and then run the function on each neighborhood, and created a new dataframe called DtToronto_venues. I displayed the size of the resulting dataframe of 1070rows x 7columns and Downtown Toronto venues.

Downtown Toronto Venues



```
Garden District, Ryerson
St. James Town
Berczy Park
Central Bay Street
Christie
Richmond, Adelaide, King
Harbourfront East, Union Station, Toronto Islands
Toronto Dominion Centre, Design Exchange
Commerce Court, Victoria Hotel
University of Toronto, Harbord
Kensington Market, Chinatown, Grange Park
CN Tower, King and Spadina, Railway Lands, Harbourfront West, Bathurst Quay, South Niagara, Island airport
Rosedale
St. James Town, Cabbagetown
First Canadian Place, Underground city
Church and Wellesley
```

Displaying the size of the resulting dataframe and Downtown Toronto Venues

```
print(DtToronto_venues.shape)
DtToronto_venues.head()
```

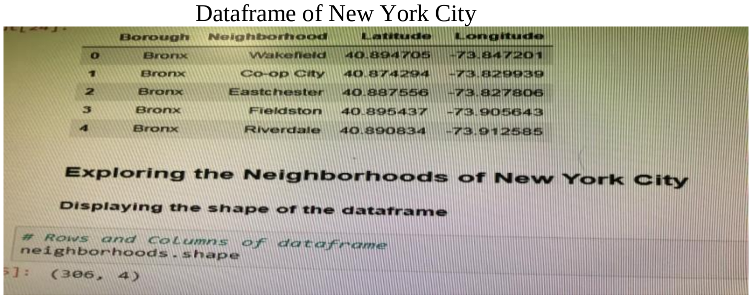
```
(1070, 7)
```

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

2.3.2.1 Exploring the Neighborhoods of New York City

First, I displayed the shape of the dataframe and wrote a script to confirm that the dataframe has 5 boroughs and 306 neighborhoods.

Dataframe of New York City



	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

Exploring the Neighborhoods of New York City

Displaying the shape of the dataframe

```
# Rows and Columns of dataframe
neighborhoods.shape
```

(306, 4)

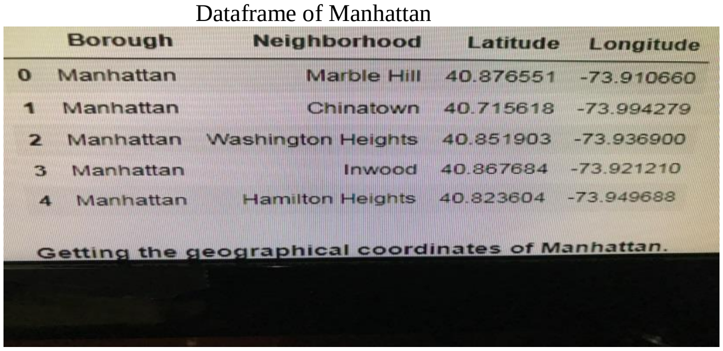
I used the geopy library to get the latitude and longitude values of the New York City, with the geographical coordinates of 40.7127281, -74.0060152. I then use these coordinates to create the map of New York City with neighborhoods superimposed on top.

Map of New York City

2.3.2.2 Exploring the Neighborhoods in Manhattan

To simplify the above map, I chose to explore only the borough of Manhattan in order to compare it to Downtown Toronto since part of Manhattan is often referred to as Downtown as well. In that instance, I sliced the original dataframe and created a new dataframe of the Manhattan data.

Dataframe of Manhattan

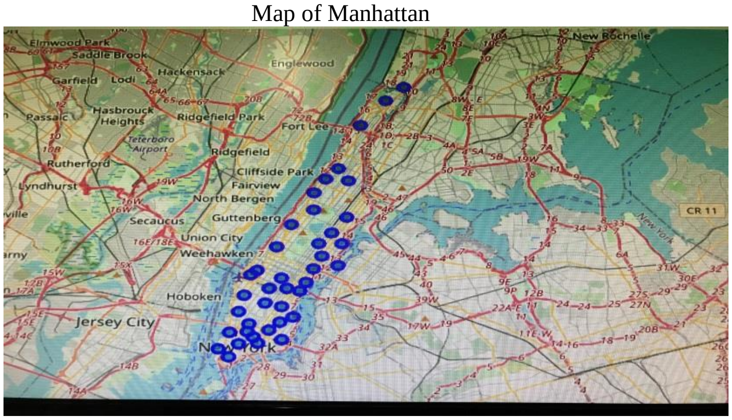


	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

Getting the geographical coordinates of Manhattan.

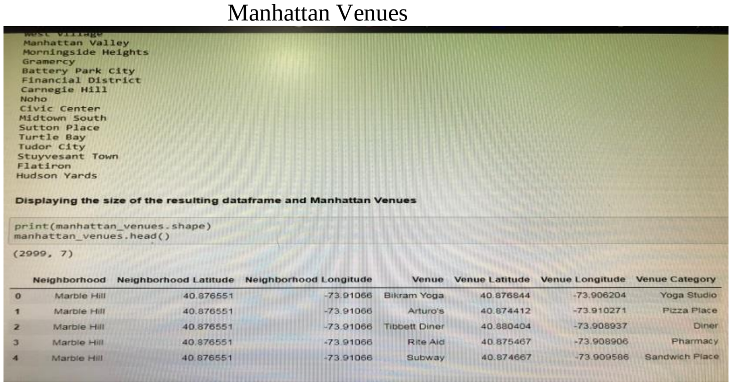
Using the geopy library, I got the geographical coordinates of Manhattan, 40.7896239, -73.9598939 and then I use these coordinates to create the map to visualize the neighborhoods of Manhattan.

Map of Manhattan



I utilized the Foursquare API to explore the neighborhoods of Manhattan and segment them. I used the same previous function to get the nearby venues of all the neighborhoods and then run the function on each neighborhood, and created a new dataframe called Manhattan_venues. I displayed the size of the resulting dataframe of 2999rows x 7columns and Manhattan venues.

Manhattan Venues



```
Manhattan Valley
Morningside Heights
Gramercy
Battery Park City
Financial District
Carnegie Hill
Noho
Civic Center
Midtown South
Sutton Place
Turtle Bay
Tudor City
Stuyvesant Town
Flatiron
Hudson Yards
```

Displaying the size of the resulting dataframe and Manhattan Venues

```
print(manhattan_venues.shape)
manhattan_venues.head()
```

(2999, 7)

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	Bikram Yoga	40.876844	-73.906204	Yoga Studio
1	Marble Hill	40.876551	-73.91066	Arturo's	40.874412	-73.910271	Pizza Place
2	Marble Hill	40.876551	-73.91066	Tibbett Diner	40.880404	-73.908937	Diner
3	Marble Hill	40.876551	-73.91066	Rite Aid	40.875467	-73.908906	Pharmacy
4	Marble Hill	40.876551	-73.91066	Subway	40.874667	-73.909586	Sandwich Place

3. Methodology

In this project, I am analyzing neighborhoods of two cities by comparing the neighborhoods of City of Toronto and New York City to determine how similar or dissimilar they are.

In the first step, I collected the required data, explored the City of Toronto and New York City and displayed the map of each city with the neighborhoods superimposed on top. I further simplified the maps by exploring only the neighborhoods in Downtown Toronto and Manhattan New York and displayed the maps for these neighborhoods.

I utilized the Foursquare API to explore the neighborhoods and segment them. According to Foursquare categorization, I was able to identify the nearby venues of all the neighborhoods of Downtown Toronto and Manhattan New York.

The second step is analyzing each neighborhood of Downtown Toronto and Manhattan New York. I will base this on calculation and explanatory data analysis to derive some additional information from our raw data.

In the third and final step, I will create clusters of neighborhoods in Downtown Toronto and clusters of neighborhoods in Manhattan New York that will provide useful insights to stakeholders in making informed decisions for their personal or business interest. My focus will be to identify clusters of neighborhoods that are conducive to business ventures, investment interests and residential life based on stakeholders' criteria. I will be using K-Means Clustering to cluster the neighborhoods and Folium to display the map of the clustered neighborhoods.

4. Analysis

4.1 Downtown Toronto

4.1.1 Analyzing Each Neighborhood of Downtown Toronto

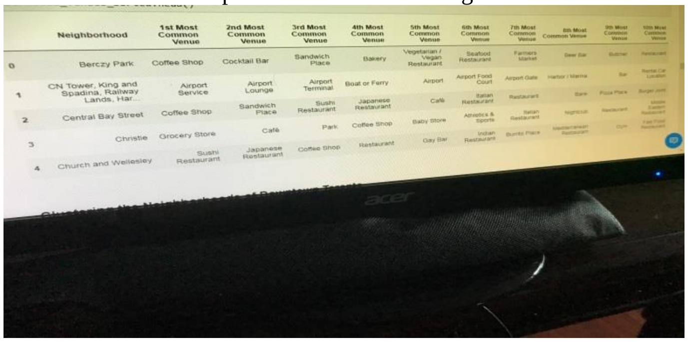
Performing some explanatory data analysis and deriving some additional information from our raw data, I first displayed the number of venues returned for each neighborhood and the number of unique categories that can be curated from all the returned venues, the result is 187 unique categories.

Unique Categories

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Berczy Park	46	46	46	46	46	46
'est, Bathurst Quay, ıgara, Island airport	16	16	16	16	16	16
Central Bay Street	62	62	62	62	62	62
Christie	15	15	15	15	15	15
urch and Wellesley	69	69	69	69	69	69
ourt, Victoria Hotel	100	100	100	100	100	100
, Underground city	100	100	100	100	100	100
n District, Ryerson	100	100	100	100	100	100
n, Toronto Islands	100	100	100	100	100	100
town, Grange Park	59	59	59	59	59	59
Park, Harbourfront	45	45	45	45	45	45
nd, Adelaide, King	99	99	99	99	99	99
Rosedale	4	4	4	4	4	4
St. James Town	82	82	82	82	82	82

I performed one hot encoding to convert the categorical variables to binary variables and appended them to the dataframe. Examining the dataframe size resulted in 1071rows x 187colums. Grouping rows by neighborhood and taking the mean of the frequency of occurrence of each category resulted in a new dataframe size of 17rows x 187columns.
I printed each neighborhood along with the top 5 most common venues, applied function to sort the venues in descending order. I then created a new dataframe and displayed the top 10 venues for each neighborhood

Top 10 Venues for each neighborhood

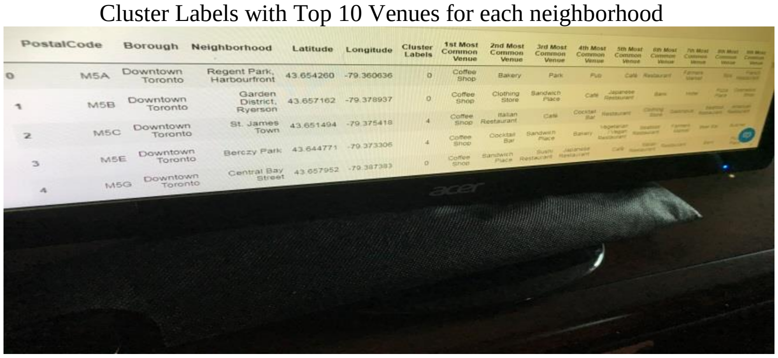


	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Berczy Park	Coffee Shop	Cocktail Bar	Sandwich Place	Bakery	Vegetarian / Vegan Restaurant	Seafood Restaurant	Farmers Market	Beer Bar	Butcher	Restaurant
1	CN Tower, King and Spadina, Railway Lands, Har...	Airport Service	Airport Lounge	Airport Terminal	Boat or Ferry	Airport	Airport Food Court	Airport Gate	Harbor / Marina	Bar	Rental Car Location
2	Central Bay Street	Coffee Shop	Sandwich Place	Sushi Restaurant	Japanese Restaurant	Café	Italian Restaurant	Restaurant	Bar	Pizza Place	Burger Joint
3	Christie	Grocery Store	Café	Park	Coffee Shop	Baby Store	Athletics & Sports	Italian Restaurant	Nightclub	Restaurant	Middle Eastern Restaurant
4	Church and Wellesley	Sushi Restaurant	Japanese Restaurant	Coffee Shop	Restaurant	Gay Bar	Indian Restaurant	Burrito Place	Mediterranean Restaurant	Gym	Fast Food Restaurant

4.1.2 Clustering the Neighborhoods of Downtown Toronto

I used K-Means to cluster the neighborhoods into 5 clusters and created a new dataframe that included the cluster labels as well as the top 10 venues for each neighborhood.

Cluster Labels with Top 10 Venues for each neighborhood



	PostalCode	Borough	Neighborhood	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
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	0	Coffee Shop	Bakery	Park	Pub	Café	Restaurant	Farmers Market	[illegible]	[illegible]
1	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	0	Coffee Shop	Clothing Store	Sandwich Place	Café	Japanese Restaurant	Bar	Hotel	Pizza Place	[illegible]
2	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	4	Coffee Shop	Italian Restaurant	Café	Cocktail Bar	Restaurant	Clothing Store	Gastropub	Seafood Restaurant	American Restaurant
3	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306	4	Coffee Shop	Cocktail Bar	Sandwich Place	Bakery	Vegetarian / Vegan Restaurant	Seafood Restaurant	Farmers Market	Beer Bar	Butcher
4	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383	0	Coffee Shop	Sandwich Place	Sushi Restaurant	Japanese Restaurant	Café	Italian Restaurant	Restaurant	Bar	Pizza Place

Finally, I created map of Downtown Toronto to visualize K-Means clusters of the neighborhoods.

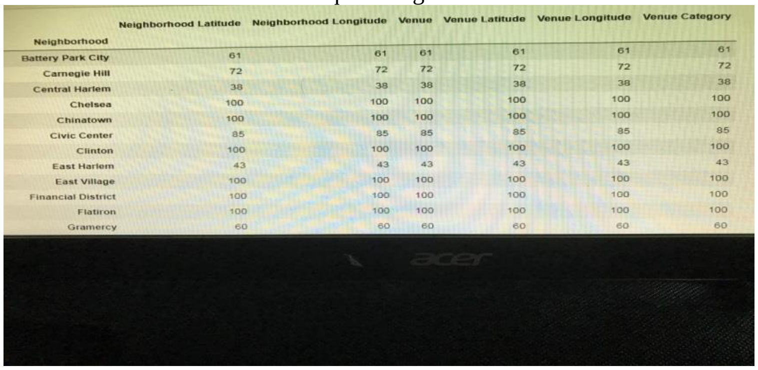
Clusters of Downtown Toronto Neighborhood

4.2 Manhattan

4.2.1 Analyzing Each Neighborhood of Manhattan

Likewise, performing some explanatory data analysis and deriving some additional information from our raw data, I first displayed the number of venues returned for each neighborhood and the number of unique categories that can be curated from all the returned venues, the result is 326 unique categories.

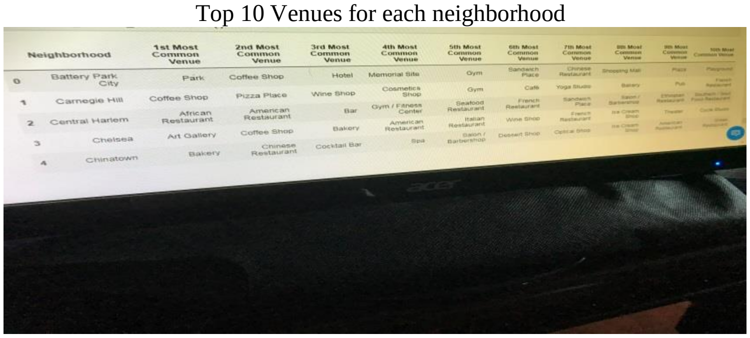
Unique Categories



Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Battery Park City	61	61	61	61	61	61
Carnegie Hill	72	72	72	72	72	72
Central Harlem	38	38	38	38	38	38
Chelsea	100	100	100	100	100	100
Chinatown	100	100	100	100	100	100
Civic Center	85	85	85	85	85	85
Clinton	100	100	100	100	100	100
East Harlem	43	43	43	43	43	43
East Village	100	100	100	100	100	100
Financial District	100	100	100	100	100	100
Flatiron	100	100	100	100	100	100
Gramercy	60	60	60	60	60	60

I performed one hot encoding to convert the categorical variables to binary variables and appended them to the dataframe. Examining the dataframe size resulted in 2995rows x 327colums. Grouping rows by neighborhood and taking the mean of the frequency of occurrence of each category resulted in a new dataframe size of 40rows x 327columns.
I printed each neighborhood along with the top 5 most common venues, applied function to sort the venues in descending order. I then created a new dataframe and displayed the top 10 venues for each neighborhood

Top 10 Venues for each neighborhood

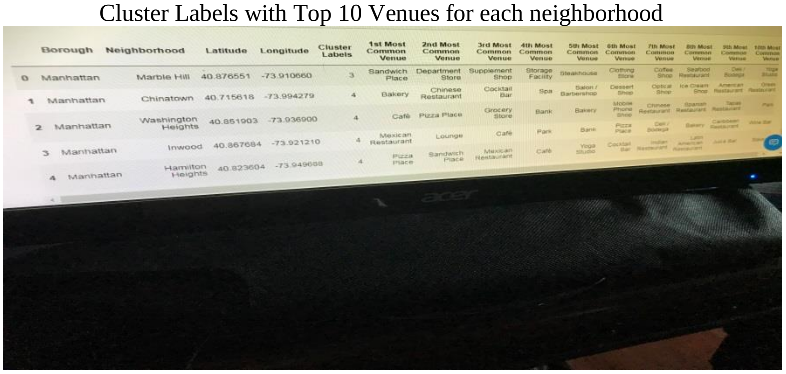


	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Battery Park City	Park	Coffee Shop	Hotel	Memorial Site	Gym	Sandwich Place	Chinese Restaurant	Shopping Mall	Plaza	Playground
1	Carnegie Hill	Coffee Shop	Pizza Place	Wine Shop	Cosmetics Shop	Gym	Café	Yoga Studio	Bakery	Pub	French Restaurant
2	Central Harlem	African Restaurant	American Restaurant	Bar	Gym / Fitness Center	Seafood Restaurant	French Restaurant	Sandwich Place	Salon / Barbershop	Ethiopian Restaurant	Southern / Soul Food Restaurant
3	Chelsea	Art Gallery	Coffee Shop	Bakery	American Restaurant	Italian Restaurant	Wine Shop	French Restaurant	Ice Cream Shop	Theater	Cycle Studio
4	Chinatown	Bakery	Chinese Restaurant	Cocktail Bar	Spa	Salon / Barbershop	Dessert Shop	Optical Shop	Ice Cream Shop	American Restaurant	Greek Restaurant

4.2.2 Clustering the Neighborhoods of Manhattan

I used K-Means to cluster the neighborhoods into 5 clusters and created a new dataframe that included the cluster labels as well as the top 10 venues for each neighborhood.

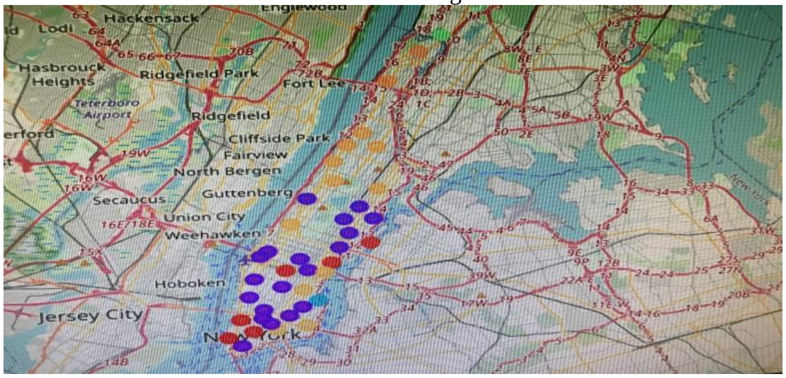
Cluster Labels with Top 10 Venues for each neighborhood



	Borough	Neighborhood	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	Manhattan	Marble Hill	40.876551	-73.910660	3	Sandwich Place	Department Store	Supplement Shop	Storage Facility	Steakhouse	Clothing Store	Coffee Shop	Seafood Restaurant	Deli / Bodega	Yoga Studio
1	Manhattan	Chinatown	40.715618	-73.994279	4	Bakery	Chinese Restaurant	Cocktail Bar	Spa	Salon / Barbershop	Dessert Shop	Optical Shop	Ice Cream Shop	American Restaurant	Greek Restaurant
2	Manhattan	Washington Heights	40.851903	-73.936900	4	Café	Pizza Place	Grocery Store	Bank	Bakery	Mobile Phone Shop	Chinese Restaurant	Spanish Restaurant	Tapas Restaurant	Park
3	Manhattan	Inwood	40.867684	-73.921210	4	Mexican Restaurant	Lounge	Café	Park	Bank	Pizza Place	Deli / Bodega	Bakery	Caribbean Restaurant	Wine Bar
4	Manhattan	Hamilton Heights	40.823604	-73.949688	4	Pizza Place	Sandwich Place	Mexican Restaurant	Café	Yoga Studio	Cocktail Bar	Indian Restaurant	Latin American Restaurant	Juice Bar	[illegible]

Finally, I created map of Manhattan to visualize K-Means clusters of the neighborhoods.

Clusters of Manhattan Neighborhood



5. Results and Discussion

In my analysis of each neighborhood of Downtown Toronto and Manhattan New York, I found that there are many diverse common venues in all the neighborhoods. It is shown that New York City cover larger geographical locations, therefore the location data are widely spread out more than that of Toronto. I examined the Downtown Toronto clusters as well as Manhattan clusters and determined the discriminating venue categories that distinguish each cluster, I assigned the following names to each cluster based on the defining common categories.

Downtown Toronto Clusters:

- DtToronto_Cluster 1 – COFFE SHOP, RESTAURANT, DIVERSE STORES, BANK
  These neighborhoods are represented on the map in Red with Cluster Label 0. They are close to the city center.

- DtToronto_Cluster 2 – NORTHERN RESIDENTIAL
  There is only 1 neighborhood (Rosedale) in this cluster, represented on the map in Purple with Cluster Label 1. It’s close to the northern part of the city center. Common venues in this neighborhood has residential feel to it like Park, Playground and so on.

- DtToronto_Cluster 3 – AIRPORT CITY, SEA PORT
  These neighborhoods are represented on the map in Blue with Cluster Label 2. It’s close to the southern part of the city center.

- DtToronto_Cluster 4 – WESTERN RESIDENTIAL
  There is only 1 neighborhood (Christie) in this cluster, represented on the map in Green with Cluster Label 3. It’s close to the western part of the city center. Common venues in this neighborhood has residential feel to it like Grocery Store, Café, Park, and so on.

- DtToronto_Cluster 5 – COFFE SHOP, RESTAURANT, EATRY

These neighborhoods are represented on the map in Orange with Cluster Label 4. They are close to the city center.

Manhattan Clusters:

- Manhattan_Cluster 1 – PARK, RESTAURANT, COFFE SHOP, DIVERSE STORES, These neighborhoods are represented on the map in Red with Cluster Label 0. They spread out on the center-south to the city center.

- Manhattan_Cluster 2 – COFFE SHOP, RESTAURANT, EATRY
  These neighborhoods are represented on the map in Purple with Cluster Label 1. They spread out from the south to the city center.

- Manhattan_Cluster 3 – PARK, PORT, STORES
  There is only 1 neighborhood (Stuyvesant Town) in this cluster, represented on the map in Blue with Cluster Label 2. It's close to the southern part of the city center.

- Manhattan_Cluster 4 – SANDWICH PLACE, STORES, EATRY
  There is only 1 neighborhood (Marble Hill) in this cluster, represented on the map in Green with Cluster Label 3. It's situated on the northern part of the city center.

- Manhattan_Cluster 5 – RESTAURANT, EATRY, DIVERSE STORES
  These neighborhoods are represented on the map in Orange with Cluster Label 4. They spread out from north to south of the city center.

6. Conclusion

In my analysis of each neighborhood of Downtown Toronto and Manhattan New York, it is shown that there are similarities and dissimilarities between the neighborhoods of these two cities, there are many diverse common venues in all the neighborhoods. New York City cover larger geographical locations, with location data widely spread out more than in the City of Toronto. I observed that residential areas are easily identified in Toronto more than in New York City where different venue categories are diversely spread out among residential areas. Stakeholders who want to make decisions about work-lifestyle-balance would definitely benefit from this location data as well as those who are making decisions about business venture and investment. There are many factors that could be considered which are based on different views, ideas and opinions of individual and to what interest are being considered, these can be for further studies.