

The Battle of Neighborhoods

To find similarities in neighborhoods of New York and Toronto

Date : 6th Feb 2021

1. Introduction

New York and Toronto are two extremely popular cities in their respective countries. In this project we will highlight the borough and neighborhoods of these respective cities, obtain graphical data of the venues around to have an idea as to the similarity in terms of venues and hence the locality by itself.

1.1. Business focus

We can figure out the neighborhoods in these cities. We will figure out the similarity in terms of venues. We will get to know the frequency of their usage to give us an idea as to what kind of business are really popular in their neck of the woods.

2. Data

Our sources of data are obtained from internet (Wikipedia, websites). We are also using the Foursquare map services to explore and cluster our neighborhoods.

The data is obtained from (Wikipedia for Toronto and a json file for NY from the coursework). This data is further cleaned to exclude any 'not assigned' rows, merge same postal codes with different neighborhoods and so on. Further We will prepare a data frame to analyse by means of bar graph. Also explore the clusters around and prepare maps accordingly.

2.1.Toronto and NY with their boroughs and neighborhoods

	Postal Code	Borough	Neighbourhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government
5	M9A	Etobicoke	Islington Avenue, Humber Valley Village
6	M1B	Scarborough	Malvern, Rouge
7	M3B	North York	Don Mills
8	M4B	East York	Parkview Hill, Woodbine Gardens
9	M5B	Downtown Toronto	Garden District, Ryerson
10	M6B	North York	Glencairn
11	M9B	Etobicoke	West Deane Park, Princess Gardens, Martin Grov...
12	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek
13	M3C	North York	Don Mills
14	M4C	East York	Woodbine Heights
15	M5C	Downtown Toronto	St. James Town

	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
5	Bronx	Kingsbridge	40.881687	-73.902818
6	Manhattan	Marble Hill	40.876551	-73.910660
7	Bronx	Woodlawn	40.898273	-73.867315
8	Bronx	Norwood	40.877224	-73.879391
9	Bronx	Williamsbridge	40.881039	-73.857446
10	Bronx	Baychester	40.866858	-73.835798
11	Bronx	Pelham Parkway	40.857413	-73.854756
12	Bronx	City Island	40.847247	-73.786488
13	Bronx	Bedford Park	40.870185	-73.885512
14	Bronx	University Heights	40.855727	-73.910416
15	Bronx	Morris Heights	40.847898	-73.919672
16	Bronx	Fordham	40.860997	-73.896427
17	Bronx	East Tremont	40.842696	-73.887356
18	Bronx	West Farms	40.839475	-73.877745
19	Bronx	High Bridge	40.836623	-73.926102

2.2. Using Foursquare

Foursquare is a tool/portal with its set of services meant for obtaining geographical data. We use Foursquare pretty much to obtain the coordinates on these cities and their respective localities.

Toronto data

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude
0	Regent Park, Harbourfront	43.65426	-79.360636	Tandem Coffee	43.653559	-79.361809
1	Regent Park, Harbourfront	43.65426	-79.360636	Roselle Desserts	43.653447	-79.362017
2	Regent Park, Harbourfront	43.65426	-79.360636	Cooper Koo Family YMCA	43.653249	-79.358008
3	Regent Park, Harbourfront	43.65426	-79.360636	Body Blitz Spa East	43.654735	-79.359874
4	Regent Park, Harbourfront	43.65426	-79.360636	Impact Kitchen	43.656369	-79.356980
5	Regent Park, Harbourfront	43.65426	-79.360636	Dominion Pub and Kitchen	43.656919	-79.358967
6	Regent Park, Harbourfront	43.65426	-79.360636	Morning Glory Cafe	43.653947	-79.361149
7	Regent Park, Harbourfront	43.65426	-79.360636	Corktown Common	43.655618	-79.356211
8	Regent Park, Harbourfront	43.65426	-79.360636	The Extension Room	43.653313	-79.359725
9	Regent Park, Harbourfront	43.65426	-79.360636	The Distillery Historic District	43.650244	-79.359323
10	Regent Park, Harbourfront	43.65426	-79.360636	Distillery Sunday Market	43.650075	-79.361832

NY 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	Starbucks	40.877531	-73.905582	Coffee Shop
4	Marble Hill	40.876551	-73.91066	Dunkin'	40.877136	-73.906666	Donut Shop

The data was also used to find out the venues around, the frequency of visits on these to gauge what was more popular (eg a Café, Restaurant and so on).

Further besides preparing data frames for 5 venues based on 5 neighborhoods, bar graphs were prepared to know what kind of venues were frequented and hence what was popular.

Toronto data

Berczy Park		venue	freq
0	0	Coffee Shop	0.09
1	1	Cocktail Bar	0.05
2	2	Bakery	0.04
3	3	Cheese Shop	0.04
4	4	Seafood Restaurant	0.04

Central Bay Street		venue	freq
0	0	Coffee Shop	0.18
1	1	Italian Restaurant	0.05
2	2	Sandwich Place	0.05
3	3	Café	0.05
4	4	Salad Place	0.03

Christie		venue	freq
0	0	Grocery Store	0.25
1	1	Cafe	0.19
2	2	Park	0.12
3	3	Nightclub	0.06
4	4	Candy Store	0.06

Church and Wellesley		venue	freq
0	0	Coffee Shop	0.09
1	1	Sushi Restaurant	0.19
2	2	Japanese Restaurant	0.06
3	3	Restaurant	0.06
4	4	Gay Bar	0.06

Commerce Court		venue	freq
0	0	Coffee Shop	0.12
1	1	Restaurant	0.07
2	2	Cafe	0.06
3	3	Hotel	0.06
4	4	Gym	0.04

NY data

Battery Park City		venue	freq
0	0	Coffee Shop	0.09
1	1	Hotel	0.07
2	2	Park	0.06
3	3	Memorial Site	0.04
4	4	Gym	0.04

Carnegie Hill		venue	freq
0	0	Coffee Shop	0.09
1	1	Café	0.05
2	2	Yoga Studio	0.03
3	3	Pizza Place	0.03
4	4	Cosmetics Shop	0.03

Central Harlem		venue	freq
0	0	African Restaurant	0.07
1	1	Bar	0.05
2	2	Chinese Restaurant	0.05
3	3	American Restaurant	0.05
4	4	Seafood Restaurant	0.05

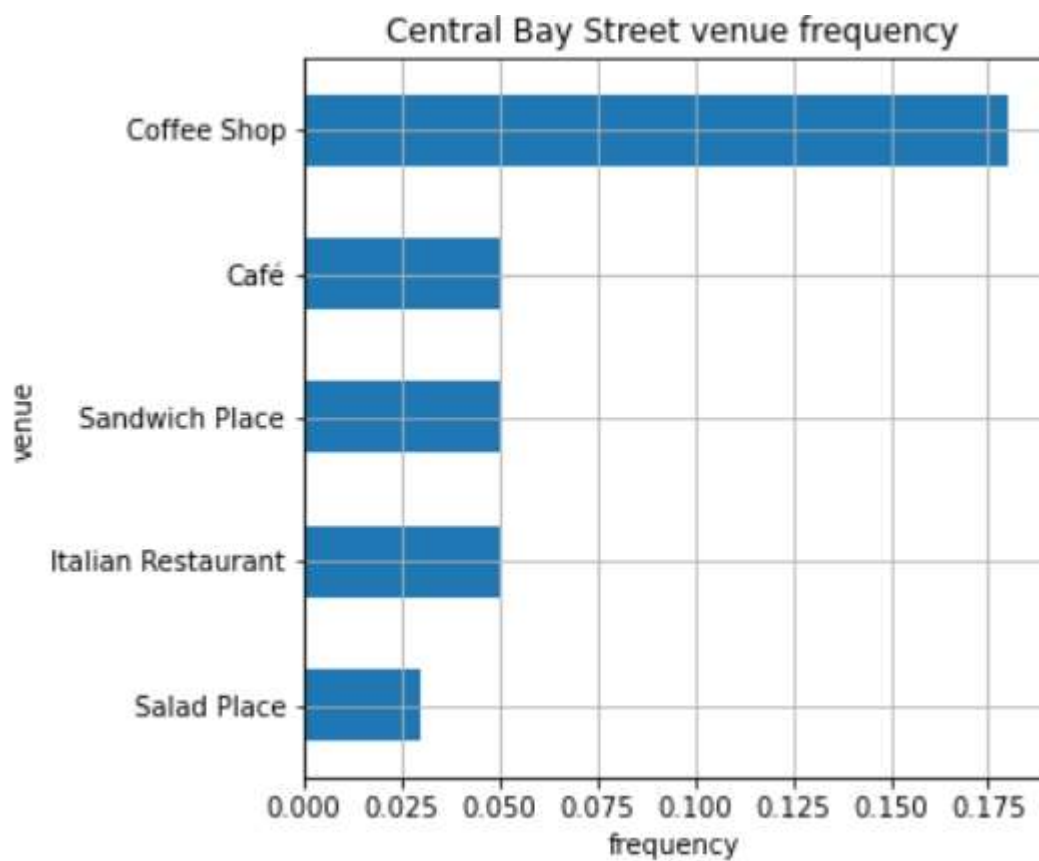
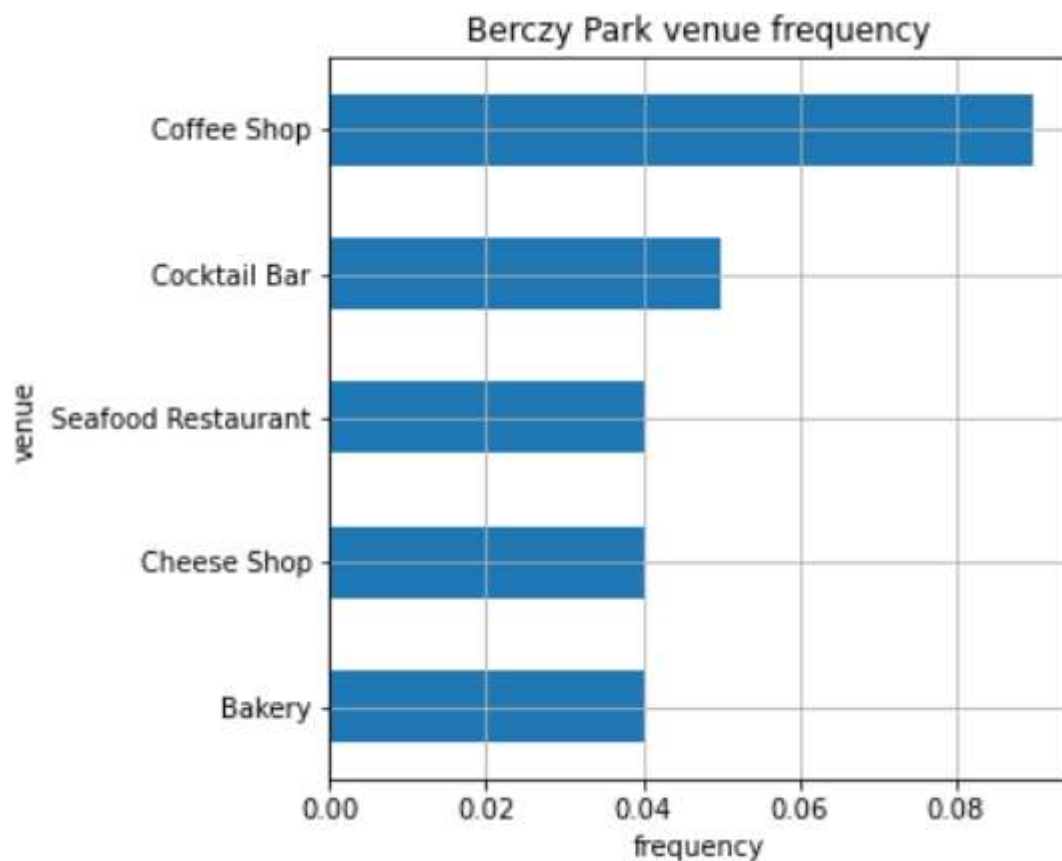
Chelsea		venue	freq
0	0	Art Gallery	0.06
1	1	Coffee Shop	0.06
2	2	Bakery	0.05
3	3	French Restaurant	0.04
4	4	American Restaurant	0.03

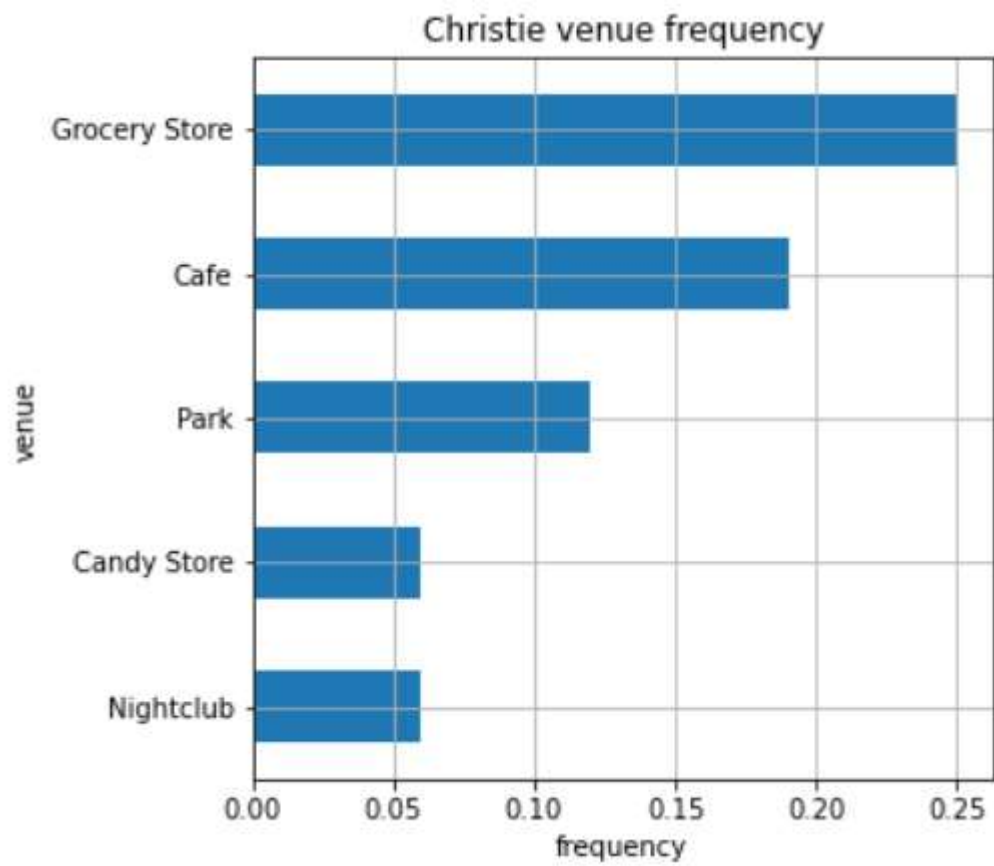
Chinatown		venue	freq
0	0	Chinese Restaurant	0.08
1	1	Bakery	0.07
2	2	Bakery	0.04
3	3	Hotpot Restaurant	0.04
4	4	Cocktail Bar	0.04

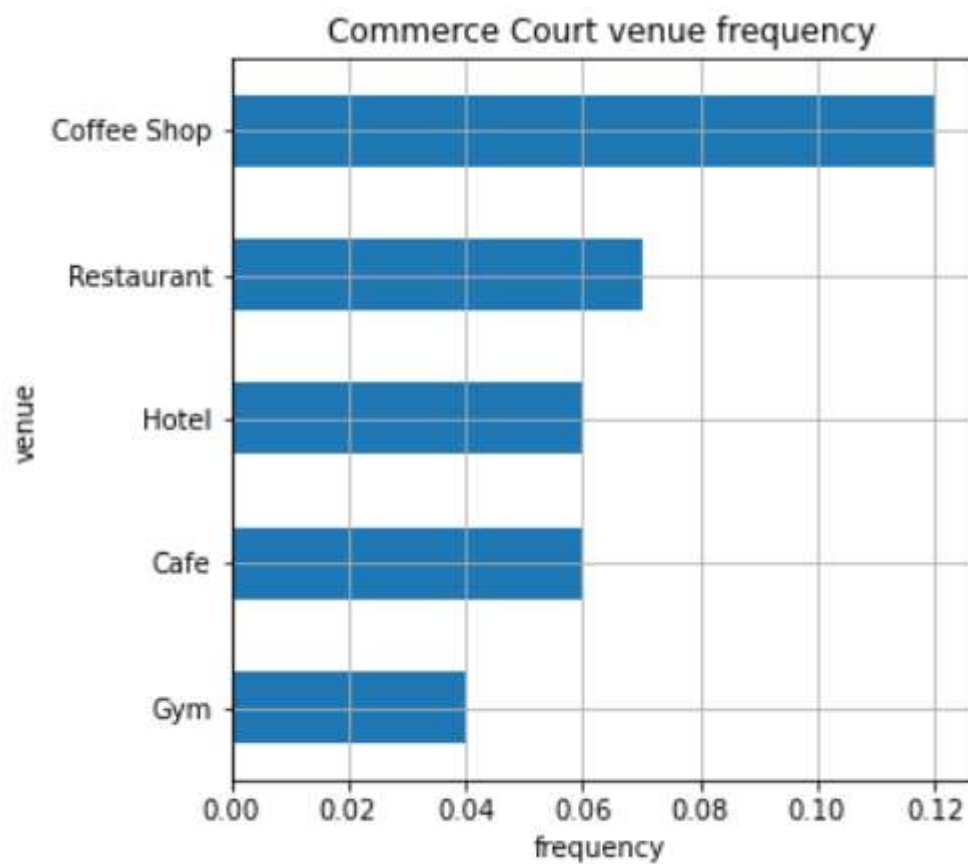
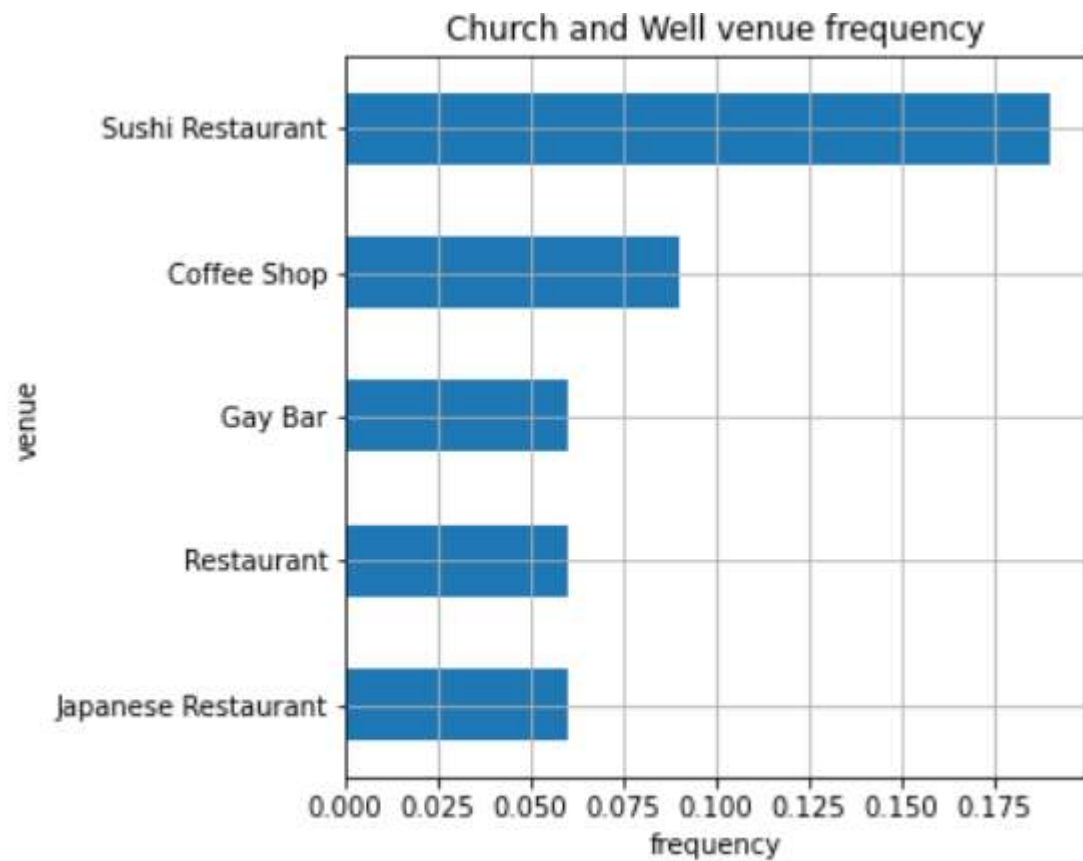
3. Methodology

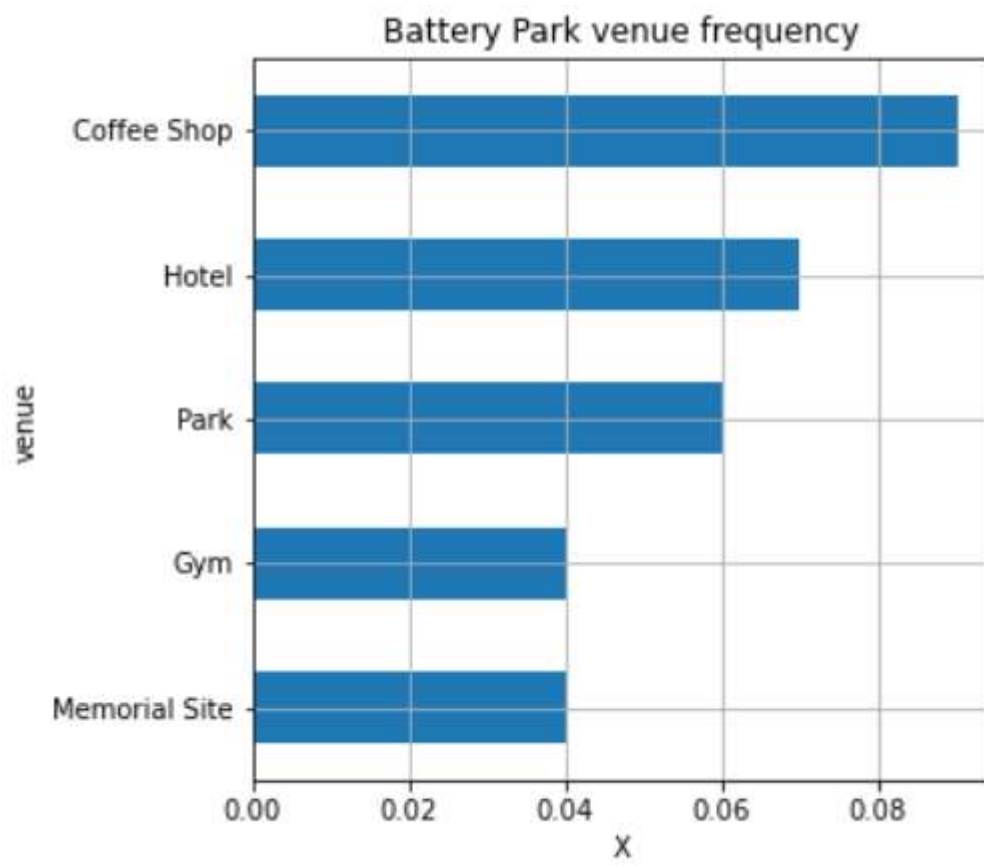
3.1.Exploration

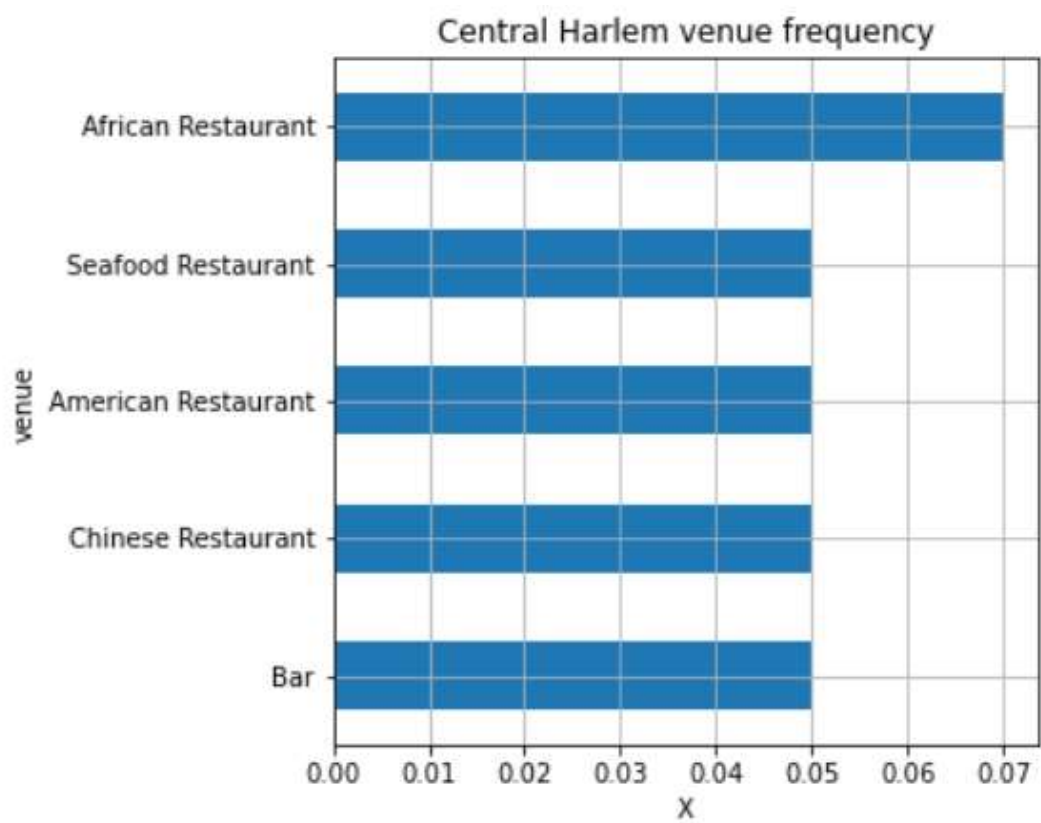
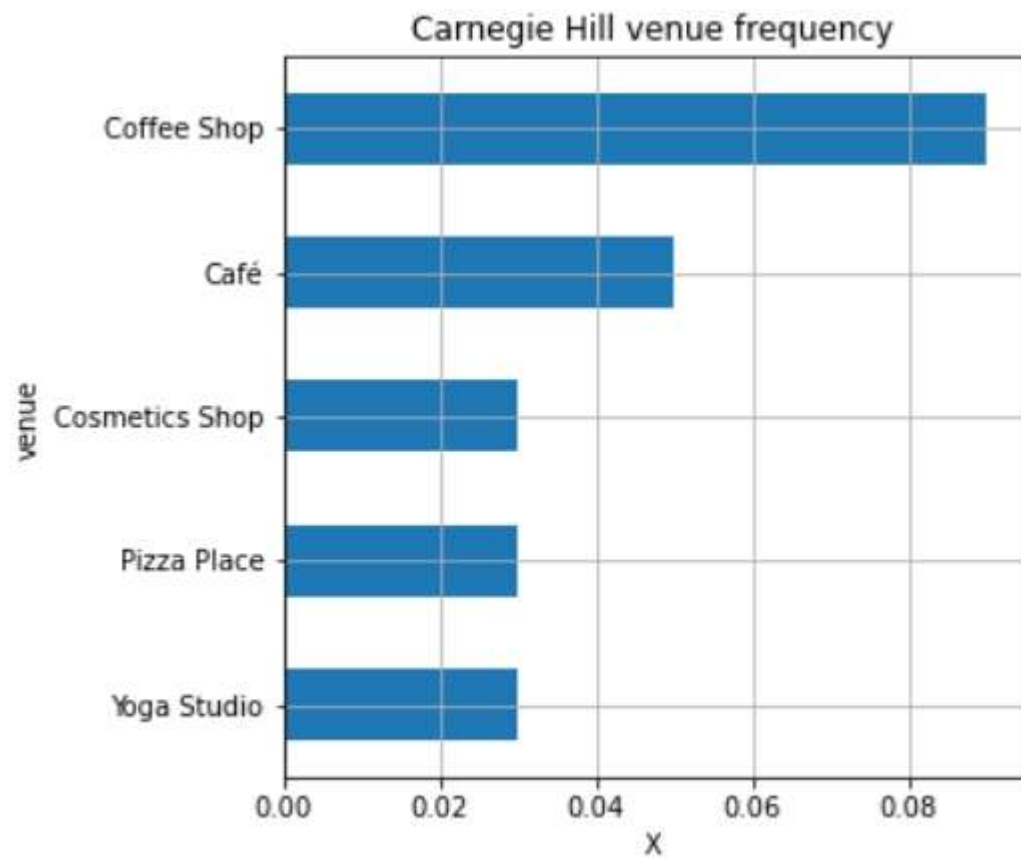
We would use information on 5 venues in 5 neighborhoods, use their frequency to prepare bar charts. The number of venues and neighborhoods can be increased to give a more in depth picture though the data obtained for 5 venues (from 5 neighborhoods) does suffice initially.

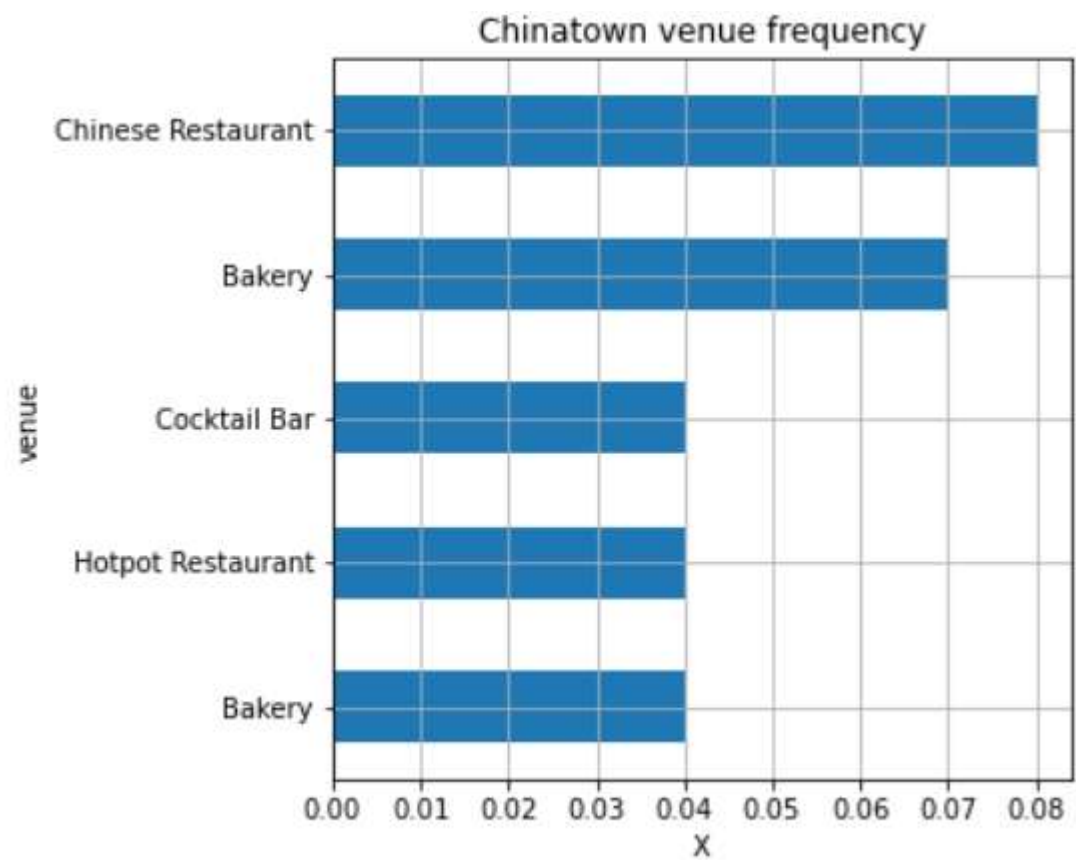
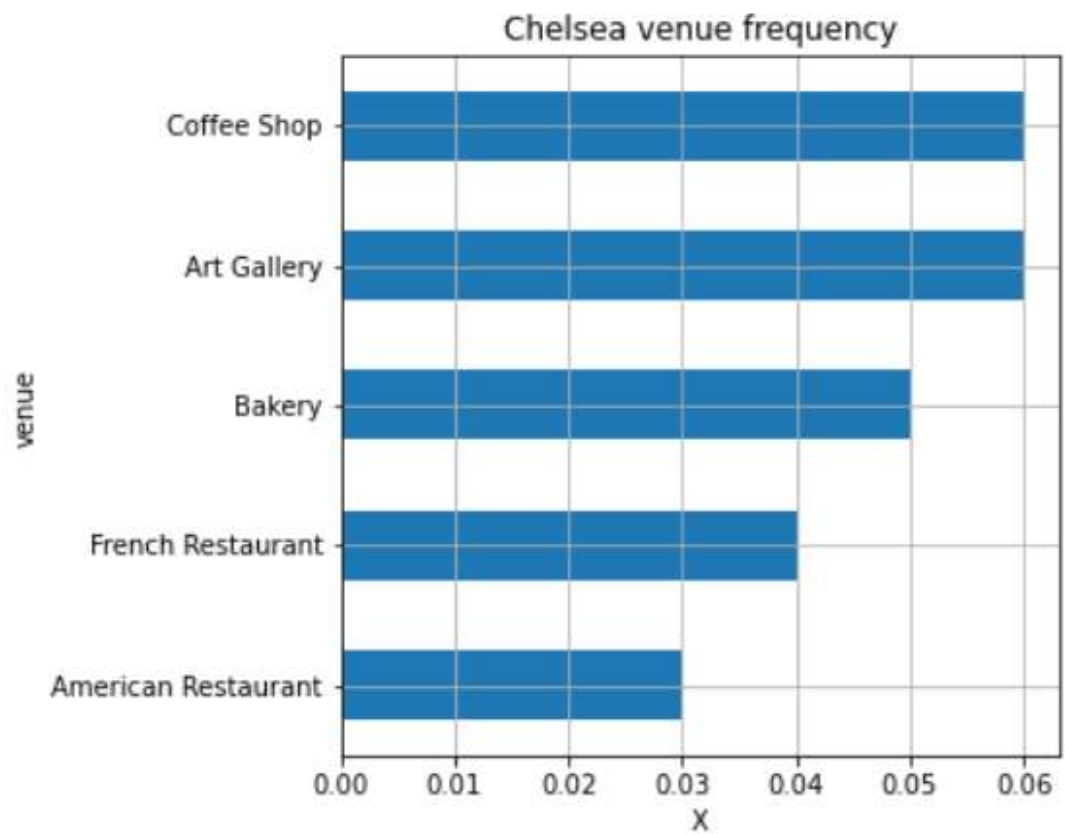
Downtown Toronto neighbourhood bar graphs





Manhattan neighbourhood bar graphs





3.2.Clustering

Using k-means to cluster the neighborhoods into 5 clusters. Preparing data frames for the respective neighborhoods and the top 10 venues.

For clustering, one hot encoding is done to give binary values to venue categories. Data is grouped by borough names to find out how many venues of each category exists in these boroughs. The frequency of the respective or more popular venues in these neighborhoods is seen in accordance.

	Neighbourhood	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
2	Garden District, Ryerson	0	Clothing Store	Coffee Shop	Cosmetics Shop	Bubble Tea Shop	Café	Middle Eastern Restaurant
3	St. James Town	0	Café	Coffee Shop	Cocktail Bar	American Restaurant	Gastropub	Bakery
4	Berczy Park	0	Coffee Shop	Cocktail Bar	Seafood Restaurant	Bakery	Beer Bar	Cheese Shop
7	Richmond, Adelaide, King	0	Coffee Shop	Café	Restaurant	Gym	Deli / Bodega	Thai Restaurant
8	Harbourfront East, Union Station, Toronto Islands	0	Coffee Shop	Aquarium	Hotel	Café	Brewery	Restaurant
9	Toronto Dominion Centre, Design Exchange	0	Coffee Shop	Hotel	Café	Restaurant	Salad Place	Seafood Restaurant
10	Commerce Court, Victoria Hotel	0	Coffee Shop	Restaurant	Hotel	Café	American Restaurant	Italian Restaurant

The clusters are further broken down into 5 across neighborhoods (Toronto and NY). We get an idea of the most common venues in each of these clusters. Like what is popular and accordingly the next one and so on.

Toronto neighbourhood

	Neighbourhood	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
2	Garden District, Ryerson	0	Clothing Store	Coffee Shop	Cosmetics Shop	Bubble Tea Shop	Café	Middle Eastern Restaurant
3	St. James Town	0	Café	Coffee Shop	Cocktail Bar	American Restaurant	Gastropub	Bakery
4	Berczy Park	0	Coffee Shop	Cocktail Bar	Seafood Restaurant	Bakery	Beer Bar	Cheese Shop
7	Richmond, Adelaide, King	0	Coffee Shop	Café	Restaurant	Gym	Deli / Bodega	Thai Restaurant
8	Harbourfront East, Union Station, Toronto Islands	0	Coffee Shop	Aquarium	Hotel	Café	Brewery	Restaurant
9	Toronto Dominion Centre, Design Exchange	0	Coffee Shop	Hotel	Café	Restaurant	Salad Place	Seafood Restaurant
10	Commerce Court, Victoria Hotel	0	Coffee Shop	Restaurant	Hotel	Café	American Restaurant	Italian Restaurant

	Neighbourhood	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
14	Rosedale	1	Park	Trail	Playground	Creperie	Doner Restaurant	Dog Run	Dis

	Neighbourhood	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
6	Christie	2	Grocery Store	Café	Park	Athletics & Sports	Italian Restaurant	Nightclub	Ca Si

	Neighbourhood	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
13	CN Tower, King and Spadina, Railway Lands, Har...	3	Airport Service	Airport Lounge	Airport Terminal	Boutique	Boat or Ferry	Plane	Re Loca

	Neighbourhood	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
0	Regent Park, Harbourfront	4	Coffee Shop	Bakery	Pub	Park	Restaurant	Breakfast Spot	
1	Queen's Park, Ontario Provincial Government	4	Coffee Shop	Sushi Restaurant	Yoga Studio	Diner	Smoothie Shop	Italian Restaurant	J Re
5	Central Bay Street	4	Coffee Shop	Sandwich Place	Italian Restaurant	Café	Bubble Tea Shop	Salad Place	

Manhattan neighbourhood

	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
1	Chinatown	Chinese Restaurant	Bakery	American Restaurant	Hotpot Restaurant	Cocktail Bar	Salon / Barbershop
6	Central Harlem	African Restaurant	Chinese Restaurant	French Restaurant	Gym / Fitness Center	American Restaurant	Bar
12	Upper West Side	Bakery	Café	Bar	Wine Bar	Italian Restaurant	Indian Restaurant
19	East Village	Bar	Pizza Place	Wine Bar	Mexican Restaurant	Korean Restaurant	Vietnamese Restaurant
20	Lower East Side	Chinese Restaurant	Art Gallery	Latin American Restaurant	Bakery	Café	Coffee Shop

	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
33	Midtown South	Korean Restaurant	Hotel	Hotel Bar	Gym / Fitness Center	Burger Joint	Coffee Shop	

	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
0	Marble Hill	Discount Store	Sandwich Place	Coffee Shop	Gym	Tennis Stadium	Steakhouse
8	Upper East Side	Italian Restaurant	Exhibit	Coffee Shop	Bakery	Gym / Fitness Center	American Restaurant
9	Yorkville	Italian Restaurant	Coffee Shop	Bar	Gym	Deli / Bodega	Sushi Restaurant
10	Lenox Hill	Italian Restaurant	Sushi Restaurant	Cocktail Bar	Coffee Shop	Café	Pizza Place
13	Lincoln Square	Plaza	Gym / Fitness Center	Café	Performing Arts Venue	Theater	Italian Restaurant

	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
2	Washington Heights	Café	Bakery	Spanish Restaurant	Grocery Store	Mobile Phone Shop	Deli / Bodega
3	Inwood	Café	Mexican Restaurant	Lounge	Restaurant	Chinese Restaurant	Deli / Bodega
4	Hamilton Heights	Pizza Place	Café	Coffee Shop	Mexican Restaurant	Deli / Bodega	Latin American Restaurant
5	Manhattanville	Coffee Shop	Sushi Restaurant	Seafood Restaurant	Mexican Restaurant	Deli / Bodega	Italian Restaurant

	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
37	Stuyvesant Town	Park	Bar	Boat or Ferry	Coffee Shop	Cocktail Bar	Heliport	Gym / Fitness Center	Farm

4. Results

4.1. Analysis of NY and Toronto data to find out the similarities. As can be seen above we get an idea of the venues frequent in five random neighborhoods. Café, restaurants, parks, gym are pretty much the norm. What can be observed is that 'food' and its avenues are the most sought after eg Coffee shops/Café's/Restaurants shine the most in these neighborhoods. Augmenting to their similarities we can see that parks, gym are the other segments that are pretty much the norm. If we had to pull out a larger sample of neighborhoods, we are mostly likely to come up with representation on similar lines.

5. Conclusion

5.1. From the above data it can be concluded that comparing neighborhoods of two metros (which are also the financial districts of respective countries), one can consider that the most common venues are coffee shops, cafes and restaurants followed by parks, gyms as such. This gives us an idea of the social space around and gives us the prominent venues. A more accurate measure of this data would obviously be to collect more samples in terms of neighborhoods, venues and compare them in case needs to be done.

References:

- 1) NY data "https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN/SkillsNetwork/labs/newyork_data.json"
- 2) https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- 3) Foursquare portal for geographical data.

<https://foursquare.com/>