

Compare Vancouver and Toronto based on top common venues

Francis Tao

Introduction:

As the top 2 populated and famous regions in Canada, Greater Vancouver Area and Toronto are two symbol cities of east and west coast. While the differences among geographical locations and development patterns lead these two places to the separate ways. So in this capstone project, I'll try to find the similarities and dissimilarities between them based on the top most common venues, which can be used to organize the neighborhoods into several clusters and then easily for comparison. Moreover, I believe the similarities and dissimilarities will display what is abundant and what is lacking in these two regions because the proportion of venues' types can be calculated from the data, and that is exactly what the current and future investors and developers want to learn about. Hence, it somehow tells the investors what is the situation and developing stage of the cities now as well as indicating the next steps or goals they could attempt to make.

Data Source:

1. Wikipedia(Toronto): [List of postal codes of Canada: M](#)
2. Wikipedia(Vancouver): [List of postal codes of Canada: V](#)
3. [Latitudes and longitudes\(Vancouver\)](#)
4. [Latitudes and longitudes\(Toronto\)](#)

Data Description:

Since the analysis is carried out based on the clusters of neighborhoods, I'll grab the data from Wikipedia, which will use the postal codes as a way to categorize the neighborhoods in both regions. The tables from the webpages already contain the borough and neighborhoods sections. However, the detailed location information, that is longitude and latitude are not provided on these pages, so I'll get Toronto's from the [previous lab](#), meanwhile, Vancouver's will be extracted from my github repository, where I downloaded it [here](#).

Data Wrangling:

For Vancouver data, I need to extract the postal codes(starts with V as initial), borough and neighborhoods respectively because the loaded data sets' formats are not what we want or properly displayed. After the success extraction, I'll assign each of them as a column, which could be denoted as 'Postcode', 'Borough' and 'Neighborhoods', and create a new pandas dataframe with these 3 columns.

A brief preview is provided below.

	Postcode	Borough	Neighborhood
0	V5A	Burnaby	Government Road / Lake City / SFU / Burnaby Mo...
1	V5B	Burnaby	Parkcrest-Aubrey / Ardingley-Sprott
2	V5C	Burnaby	Burnaby Heights / Willingdon Heights / West Ce...
3	V5E	Burnaby	Lakeview-Mayfield / Richmond Park / Kingsway-B...
4	V5G	Burnaby	Cascade-Schou / Douglas-Gilpin
5	V5H	Burnaby	Maywood / Marlborough / Oakalla / Windsor
6	V5J	Burnaby	Suncrest / Sussex-Nelson / Clinton-Glenwood / ...
7	V5K	Vancouver	North Hastings-Sunrise
8	V5L	Vancouver	North Grandview-Woodland

For Toronto data, it is pretty similar to what I did in Vancouver's but much easier. The original data from the webpage is somehow well organized and I just need to drop the not assigned values and eliminate the duplicate rows. For the postcodes which have more than one neighborhoods, I'll combine them according to the unique postal code so that the dataframe will be clear and straightforward to view.

A brief preview below:

	Postcode	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	Queen's Park	Not assigned
5	M9A	Etobicoke	Islington Avenue
6	M1B	Scarborough	Malvern, Rouge
7	M3B	North York	Don Mills North
8	M4B	East York	Parkview Hill, Woodbine Gardens

Data combination:

For both Vancouver and Toronto tables, I'll merge the latitudes and longitudes to them, joining on the postal codes. Preview below.

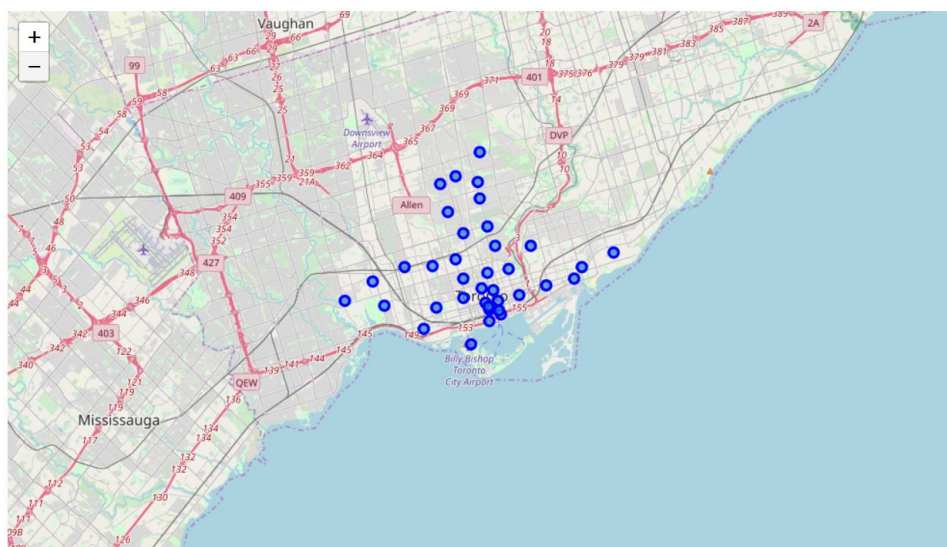
	Postcode	Borough	Neighborhood	Latitude	Longitude
0	V5A	Burnaby	Government Road / Lake City / SFU / Burnaby Mo...	49.264485	-122.941439
1	V5B	Burnaby	Parkcrest-Aubrey / Ardingley-Sprott	49.273582	-122.977384
2	V5C	Burnaby	Burnaby Heights / Willingdon Heights / West Ce...	49.276181	-123.007376
3	V5E	Burnaby	Lakeview-Mayfield / Richmond Park / Kingsway-B...	49.226319	-122.957311
4	V5G	Burnaby	Cascade-Schou / Douglas-Gilpin	49.247746	-123.000456
5	V5H	Burnaby	Maywood / Marlborough / Oakalla / Windsor	49.230178	-122.998919
6	V5J	Burnaby	Suncrest / Sussex-Nelson / Clinton-Glenwood / ...	49.213681	-122.995491
7	V5K	Vancouver	North Hastings-Sunrise	49.279416	-123.041032
8	V5L	Vancouver	North Grandview-Woodland	49.278348	-123.065858

	Postcode	Borough	Neighbourhood	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	M4E	East Toronto	The Beaches	43.676357	-79.293031
4	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306
5	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383
6	M6G	Downtown Toronto	Christie	43.669542	-79.422564
7	M5H	Downtown Toronto	Richmond, King, Adelaide	43.650571	-79.384568
8	M6H	West Toronto	Dufferin, Dovercourt Village	43.669005	-79.442259

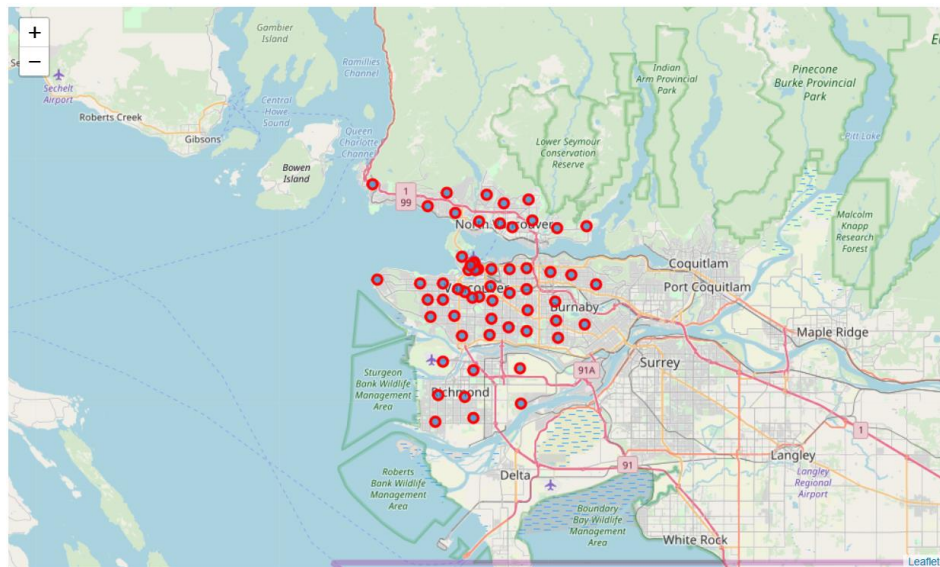
Methodology:

Once the data wrangling is done, I'll get the geographical coordinates of Toronto and Vancouver using the function 'Nominatim' imported from the package 'geopy.geocoders'. After the centers are found, I'll produce two maps of these 2 regions, which demonstrate the neighborhoods by circles.

Toronto:



Vancouver:



Next, I'll use API FOURSQUARE to get the detailed information of each city/region - that is venues and their categories, locations, etc. For each neighborhood, I'll set the radius to 500m and limit the number of venues to 100, which is a representative and great value so that we could make effective statistical analysis without having too much useless information.

Then, I'll display the top common venues in each neighborhood based on its frequency. Take top 5 as an example, Toronto's are shown below:

```
----Berczy Park----
      venue  freq
0    Coffee Shop 0.07
1    Cocktail Bar 0.05
2         Pub     0.03
3 Seafood Restaurant 0.03
4    Cheese Shop  0.03

----Business Reply Mail Processing Centre 969 Eastern--
      venue  freq
0 Light Rail Station 0.12
1    Smoke Shop     0.06
2 Recording Studio  0.06
3    Brewery        0.06
4    Auto Workshop  0.06
```

```
----Central Bay Street----
```

Vancouver's are here:

```
----Cascade-Schou / Douglas-Gilpin---
      venue  freq
0  Snack Place 0.22
1  Coffee Shop 0.11
2  Bus Station 0.11
3   Bookstore  0.11
4   Bus Stop   0.11
```

```
----Central----
      venue  freq
0 Convenience Store 0.29
1   Sushi Restaurant 0.14
2      Pharmacy     0.14
3 Clothing Store    0.14
4 Chinese Restaurant 0.14
```

```
----Central Kitsilano / Greektown----
```

If we put them into a table, see below:

Top 10 common venues (Toronto):

	Neighbourhood	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	Restaurant	Farmers Market	Cheese Shop	Café	Pub	Bakery	Steakhouse	Italian Restaurant
1	Business Reply Mail Processing Centre 969 Eastern	Yoga Studio	Auto Workshop	Pizza Place	Park	Restaurant	Butcher	Burrito Place	Brewery	Skate Park	Smoke Shop
2	Central Bay Street	Coffee Shop	Café	Italian Restaurant	Burger Joint	Bar	Chinese Restaurant	Thai Restaurant	Salad Place	Bubble Tea Shop	Spa
3	Christie	Café	Grocery Store	Park	Convenience Store	Italian Restaurant	Baby Store	Diner	Athletics & Sports	Nightclub	Restaurant
4	Church and Wellesley	Japanese Restaurant	Coffee Shop	Sushi Restaurant	Gay Bar	Restaurant	Burger Joint	Pub	Mediterranean Restaurant	Gastropub	Fast Food Restaurant

Top 10 common venues (Vancouver):

	Neighbourhood	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	Bentall Centre	Hotel	Restaurant	Italian Restaurant	Seafood Restaurant	Bakery	Breakfast Spot	Coffee Shop	Concert Hall	Yoga Studio	Burger Joint
1	Burnaby Heights / Willingdon Heights / West Ce...	Shopping Plaza	Yoga Studio	Comfort Food Restaurant	Filipino Restaurant	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Fair
2	Cascade-Schou / Douglas-Gilpin	Snack Place	Bookstore	Auto Garage	Bus Station	Bus Stop	Sandwich Place	Park	Coffee Shop	Department Store	Dessert Shop
3	Central	Convenience Store	Pharmacy	Clothing Store	Sushi Restaurant	Chinese Restaurant	Sandwich Place	Donut Shop	Fast Food Restaurant	Farmers Market	Farm
4	Central Kitsilano / Greektown	Coffee Shop	Grocery Store	Restaurant	Japanese Restaurant	Pharmacy	Pizza Place	Fast Food Restaurant	Supplement Shop	Food Truck	Breakfast Spot

Machine learning method used:

Now we'll discuss which machine learning method to use for this project. Since each neighborhood is not a number value, this is an unsupervised learning which is not continuous. Moreover, since each of them is not labeled, we have to use clustering as our exploratory data analysis. While the criterion is based on the frequency of venues within each neighborhood, we'll use K-means to solve the problem because it is one of the greatest and most convenient ways to evaluate the similarities.

Result:

After the k-means clustering is applied, I get the following two tables which indicate the labels of the neighborhoods:

For Vancouver:

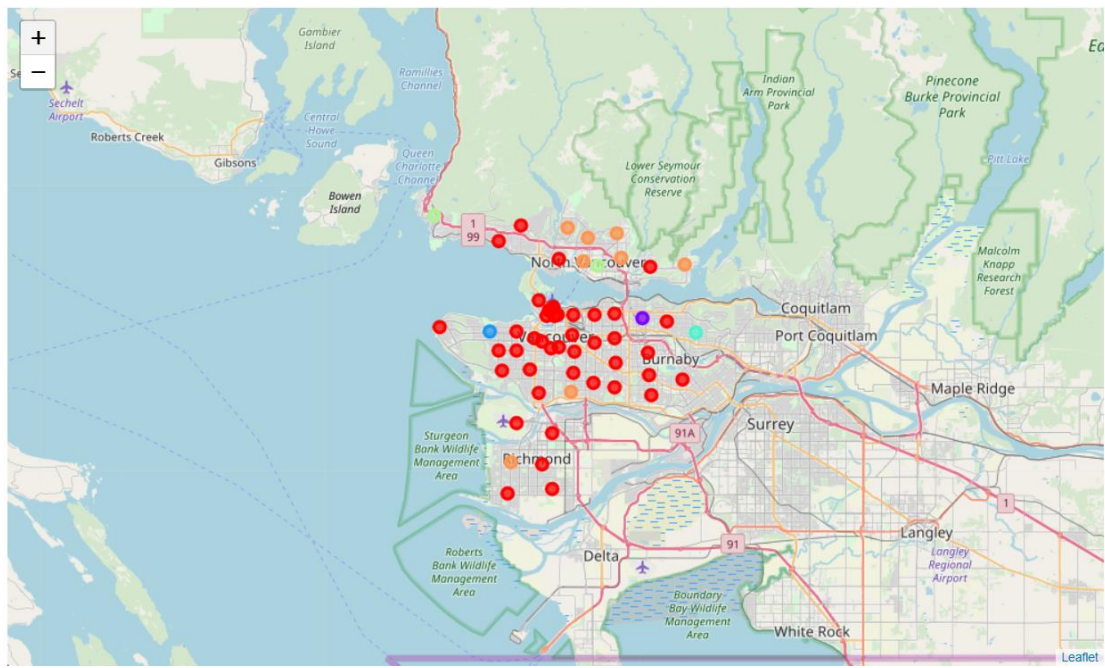
Postcode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st most Common Venue	2nd most Common Venue	3rd most Common Venue	4th most Common Venue	5th most Common Venue	6th most Common Venue	7th most Common Venue	8th most Common Venue
0	V5A	Burnaby	Government Road / Lake City / SFU / Burnaby Mo...	49.264485	-122.941439	3.0	Burger Joint	Golf Course	Golf Driving Range	Yoga Studio	Fish & Chips Shop	Field	Fast Food Restaurant
1	V5B	Burnaby	Parkcrest-Aubrey / Ardingley-Sprott	49.273582	-122.977384	0.0	Stadium	Bus Stop	Playground	Martial Arts Dojo	Golf Course	Filipino Restaurant	Field
2	V5C	Burnaby	Burnaby Heights / Willingdon Heights / West Ce...	49.276181	-123.007376	1.0	Shopping Plaza	Yoga Studio	Comfort Food Restaurant	Filipino Restaurant	Field	Fast Food Restaurant	Farmers Market
3	V5E	Burnaby	Lakeview-Mayfield / Richmond Park / Kingsway-B...	49.226319	-122.957311	0.0	Mobile Phone Shop	Bus Stop	Yoga Studio	Concert Hall	Filipino Restaurant	Field	Fast Food Restaurant
4	V5G	Burnaby	Cascade-Schou / Douglas-Gilpin	49.247746	-123.000456	0.0	Snack Place	Bookstore	Auto Garage	Bus Station	Bus Stop	Sandwich Place	Park

For Toronto:

	Postcode	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	1st most Common Venue	2nd most Common Venue	3rd most Common Venue	4th most Common Venue	5th most Common Venue	6th most Common Venue	7th most Common Venue
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	0	Coffee Shop	Pub	Bakery	Park	Café	Breakfast Spot	Theater
1	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	0	Coffee Shop	Clothing Store	Cosmetics Shop	Café	Diner	Thai Restaurant	Ramen Restaurant
2	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	0	Coffee Shop	Restaurant	Café	Hotel	Breakfast Spot	Park	Gastropub
3	M4E	East Toronto	The Beaches	43.676357	-79.293031	3	Coffee Shop	Health Food Store	Pub	Other Great Outdoors	Neighborhood	Event Space	Falafel Restaurant
4	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306	0	Coffee Shop	Cocktail Bar	Bakery	Cheese Shop	Pub	Café	Restaurant

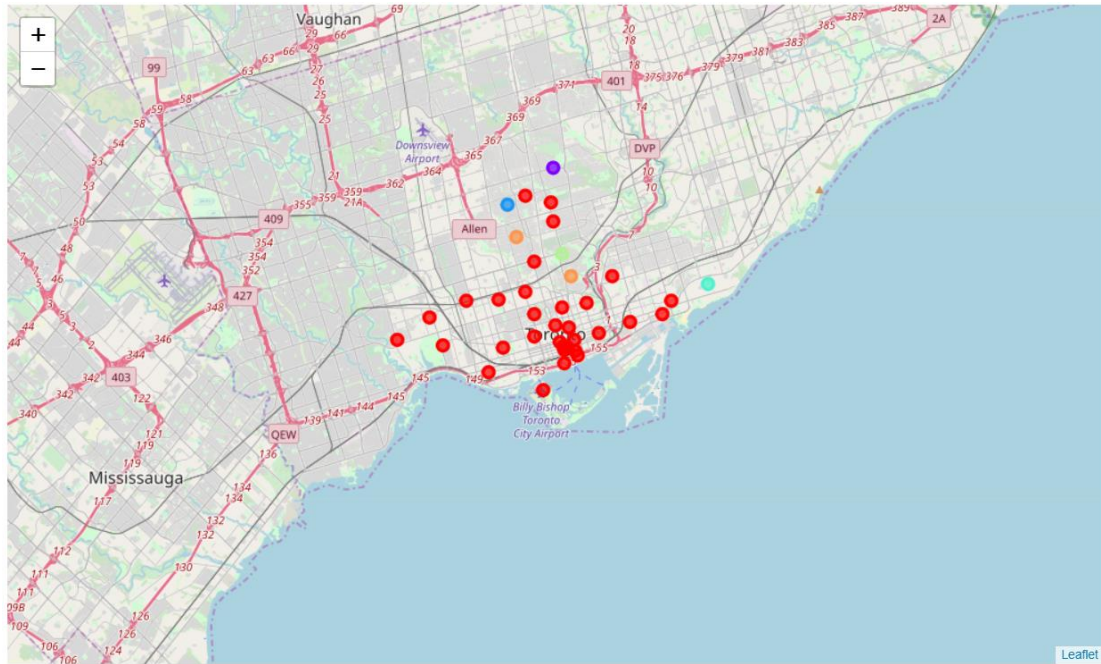
Let's create two maps and get a more straightforward and clear visualization.

Vancouver:



```
0.0    43
5.0     8
4.0     2
2.0     1
1.0     1
3.0     1
Name: Cluster Labels, dtype: int64
```

Toronto:



```
0    32
5     2
4     1
3     1
2     1
1     1
```

Name: Cluster Labels, dtype: int64

Cluster 1 of Toronto(first 10):

	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Downtown Toronto	0	Coffee Shop	Pub	Bakery	Park	Café	Breakfast Spot	Theater	Restaurant	Mexican Restaurant	Gym / Fitness Center
1	Downtown Toronto	0	Coffee Shop	Clothing Store	Cosmetics Shop	Café	Diner	Thai Restaurant	Ramen Restaurant	Bubble Tea Shop	Plaza	Japanese Restaurant
2	Downtown Toronto	0	Coffee Shop	Restaurant	Café	Hotel	Breakfast Spot	Park	Gastropub	Italian Restaurant	Cosmetics Shop	Clothing Store
4	Downtown Toronto	0	Coffee Shop	Cocktail Bar	Bakery	Cheese Shop	Pub	Café	Restaurant	Farmers Market	Seafood Restaurant	Steakhouse
5	Downtown Toronto	0	Coffee Shop	Café	Italian Restaurant	Bubble Tea Shop	Bar	Burger Joint	Sandwich Place	Chinese Restaurant	Thai Restaurant	Spa
6	Downtown Toronto	0	Café	Grocery Store	Park	Convenience Store	Nightclub	Diner	Baby Store	Italian Restaurant	Restaurant	Athletics & Sports
7	Downtown Toronto	0	Coffee Shop	Café	Bar	Steakhouse	Thai Restaurant	Hotel	Restaurant	Burger Joint	Bakery	American Restaurant
8	West Toronto	0	Discount Store	Pharmacy	Supermarket	Bakery	Music Venue	Liquor Store	Brewery	Middle Eastern Restaurant	Café	Bar
9	Downtown Toronto	0	Coffee Shop	Hotel	Aquarium	Café	Pizza Place	Italian Restaurant	Brewery	Restaurant	Scenic Lookout	History Museum

Cluster 1 of Vancouver(7 of all):

6	Burnaby	0.0	Liquor Store	Sandwich Place	Grocery Store	Yoga Studio	Donut Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
7	Vancouver	0.0	Theme Park Ride / Attraction	Vietnamese Restaurant	Event Space	Theme Park	Office	Bridal Shop	Bus Line	Bus Station	Sandwich Place	Fair
8	Vancouver	0.0	Pizza Place	Coffee Shop	Asian Restaurant	Grocery Store	Theater	Fried Chicken Joint	Italian Restaurant	Bakery	Japanese Restaurant	Sushi Restaurant
9	Vancouver	0.0	Coffee Shop	Pizza Place	Arts & Crafts Store	Bus Stop	Sandwich Place	Restaurant	Liquor Store	Chinese Restaurant	Grocery Store	Office
10	Vancouver	0.0	Café	Bookstore	Church	Farmers Market	Park	Bike Shop	Sushi Restaurant	Beach	Lake	Dog Run
11	Vancouver	0.0	Japanese Restaurant	Pizza Place	Convenience Store	Asian Restaurant	Fried Chicken Joint	Bus Stop	Fish Market	Bus Station	Thrift / Vintage Store	Noodle House
12	Vancouver	0.0	Hotel	Asian Restaurant	Tennis Court	Other Great Outdoors	Dessert Shop	Hotpot Restaurant	Fast Food Restaurant	Farmers Market	Farm	Eastern European Restaurant

Rest of Toronto:

	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
17	Central Toronto	1	Park	Swim School	Bus Line	Yoga Studio	Discount Store	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store
18	Central Toronto	2	Garden	Pool	Yoga Studio	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant
3	East Toronto	3	Coffee Shop	Health Food Store	Pub	Other Great Outdoors	Neighborhood	Event Space	Falafel Restaurant	Farmers Market	Ethiopian Restaurant	Diner
28	Central Toronto	4	Playground	Diner	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Dumpling Restaurant
20	Central Toronto	5	Park	Trail	Sushi Restaurant	Jewelry Store	Yoga Studio	Discount Store	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant
32	Downtown Toronto	5	Park	Playground	Trail	Diner	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store

Rest of Vancouver:

	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
2	Burnaby	1.0	Shopping Plaza	Yoga Studio	Comfort Food Restaurant	Filipino Restaurant	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Fair
32	Vancouver	2.0	Gym / Fitness Center	Yoga Studio	Donut Shop	Filipino Restaurant	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Fair
0	Burnaby	3.0	Burger Joint	Golf Course	Golf Driving Range	Yoga Studio	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
48	North Vancouver	4.0	Playground	Café	Convenience Store	Cosmetics Shop	Filipino Restaurant	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
56	West Vancouver	4.0	Playground	Tapas Restaurant	Harbor / Marina	Dog Run	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Fair
17	Vancouver	5.0	Indian Restaurant	Park	Yoga Studio	Donut Shop	Filipino Restaurant	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
42	Richmond	5.0	Eastern European Restaurant	Park	Sushi Restaurant	Farm	Basketball Court	Trail	Yoga Studio	Filipino Restaurant	Field	Fast Food Restaurant
44	North Vancouver	5.0	Dog Run	Pub	Park	Soccer Field	Yoga Studio	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Fair
46	North Vancouver	5.0	Distribution Center	Park	Trail	Dog Run	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Fair
47	North Vancouver	5.0	Convenience Store	Dance Studio	Park	Eastern European Restaurant	Filipino Restaurant	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
49	North Vancouver	5.0	Park	Mediterranean Restaurant	Insurance Office	Athletics & Sports	Eastern European Restaurant	Fish & Chips Shop	Filipino Restaurant	Field	Fast Food Restaurant	Farmers Market

Discussion and Conclusion:

After the mapping and cluster checking of these two regions, we can find that the dominating clusters of Vancouver and Toronto are mainly composed of drinking and eating places, such as coffee shop, café and pizza places. Furthermore, most of them gather of locate in the center of these two regions. On the other hand, the rest clusters can be

combined to see what kind of venues are listed. For Vancouver, there is only a few exercising places like fitness center, studios, park and other entertainment such as pub. For Toronto, we can see that swimming pools, park, garden, playground emerged.

For both cities, we can approach to the point that they are quite similar in the distribution or proportion of the venues diversity. Eating and drinking places are common, while the other places which represent the entertainment and social activities only occupy a small portion.

Although in our potential mind, Toronto and Vancouver are two totally different regions, however, the analysis shows that the dissimilarities only or mostly exist in the diversity catering industry, which can be ignored at the current stage for the reason that both of them have achieved a mature and developed level.

Yet, I suggest the investors to build or carry out the project and programs such as establishing more cinemas, fitness centers, amusement parks to foster the development of sporting goods, recreation and entertainment industry in both regions.