

CAPSTONE PROJECT

By Manish Bhatia

Unification of Restaurants based on Cuisine type in Downtown TORONTO, Canada and suggesting optimum cluster to Open an Indian Restaurant

Introduction: -

Toronto city in Canada is one of the most multicultural & lively cities in the world, this is the reason Toronto comes in Top 10 best cities to live in and attract a lot of tourists from all over the Globe. The city is famous of beautiful lakes, sceneries and outdoor life. The one major characteristic of this city is that it is having more than 50% population which are migrated from outside of Canada.

This has resulted in diverse food choices of people who are living here or the tourists who usually visit in seasonal time. Due to this there are around 7500 restaurants of different cuisine styles are opened like German, Italian, Japanese, Indian, Asian and so on.

Description of Problem Statement: -

As it becomes hard to conclude the choices of which area/restaurant to visit in a long list of restaurants or which are the **best restaurant venues nearby** or a suggestion to open an **Indian restaurant** in which neighborhood where it could attract more people and could possibly generate more revenue.

In this project we will analyze most famous area '**Downtown Toronto**' of Toronto city using its demographics characteristics and will visualize which specific cuisine restaurants are most common in certain neighborhoods using **clustering** methodology so that tourists could plan the trip based on their food choices and also we will try to identify the neighborhoods where an Indian restaurant could be opened based on existing restaurant number and food style priorities.

Description of Data: -

Based on problem statement, below sources were utilized to extract the desired data

1. Wikipedia: - 'https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'

This data will provide information about the 'Postal codes', 'Boroughs' and 'Neighborhoods' of Toronto city which can help in filtering 'Downtown Toronto' from the Boroughs and its associated neighborhoods.

	Postal Code	Borough	Neighbourhood
0	M1A	Not assigned	Not assigned
1	M2A	Not assigned	Not assigned
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Regent Park, Harbourfront

2. Geospatial Data: - 'http://cocl.us/Geospatial_data'

With the help of geospatial data, we can extract information about 'Latitude' and 'Longitude' of each and every postal code of Toronto city and with the help of merging technique, the location information can be imposed on Neighborhoods on the table extracted from Data source 1.

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

3. Foursquare API:-

https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, VERSION, neighborhood_latitude, neighborhood_longitude, radius, LIMIT)

This API will help in retrieving information about the most famous venues in 'Downtown Toronto' and the category of those venues like Restaurants, Bars, park etc.

The main advantage of this API is that the data can be extracted based on geographical parameters and which can be modified to extract desired information in a certain radius and desired number of venues are also changeable.

4. Demographic information from Wikipedia: - https://en.wikipedia.org/wiki/Demographics_of_Toronto_neighbourhoods

This data from wiki webpage is to get demographic information of specific neighborhoods, this table can help in studying living population in certain neighborhoods and in which areas the most spoken language was either 'Hindi' or 'Punjabi' so that the proposal of Indian restaurants could be decided and chances of higher revenue could get increased.

How Data will be used to solve the problem: -

The Toronto neighborhood data will be extracted from wiki page 1 and this will provide an association of all the neighborhoods in 'Downtown Toronto' area of Toronto city. After extracting this information, the table will be mapped with Geospatial data which will provide latitude and longitude of each neighborhood in Downtown Toronto. Once it is done then by using Foursquare API extracted table will be utilized to extract venues and their categories in each neighborhood. Thus, using this API Categories of different restaurants will be filtered out for each neighborhood and we can extract information of most famous restaurants in the neighborhoods and also can categorize them by rank.

After categorizing, the data can help in proper clustering of restaurants in each neighborhood, this could be a valuable information for tourists to filter out the neighborhoods where specific cuisine restaurant is most famous.

Using Folium map, we will try to plot each neighborhood on Toronto map and will visualize the distribution of clusters of restaurants of different cuisine. This data will also support in extracting information about the clusters where Indian restaurants are most famous or in specific neighborhoods Indian restaurants are not in top 15 venues.

Thus by imposing demographic information like 'population' and the spoken language in any neighborhood we can finalize a cluster where other restaurants are most famous but the locality is of people who belongs to Indian or Asian community and proposal of an Indian restaurant could be favorable.


Analysis Methodology and Results outcome: -

The extracted data was in raw form which required initial data cleaning analogy to convert it into usable form. after filtering out the tables from wiki webpage certain columns like postal code, Borough and Neighborhood were selected.

As many of the entries were having invalid values like 'Not assigned' in Borough column so processing & cleaning was done by removing such non informational values and Neighborhood columns were mapped with Borough values if there were 'not assigned' values.

Final conversion and outcome from Dataset1 are as shown below-

	Postal Code	Borough	Neighbourhood
0	M1A	Not assigned	Not assigned
1	M2A	Not assigned	Not assigned
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Regent Park, Harbourfront



	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


Now second Dataset was extracted from Geospatial data link to extract latitude and longitude of each Postal code, this helped in mapping each Borough/neighborhood with latitude and longitude coordinates.

The result of this step is as shown below: -

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

Once the coordinates of each postal code were extracted, table 2 was merged with table 1 and we were having a combined table representing Boroughs/neighborhoods with individual location coordinates.

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

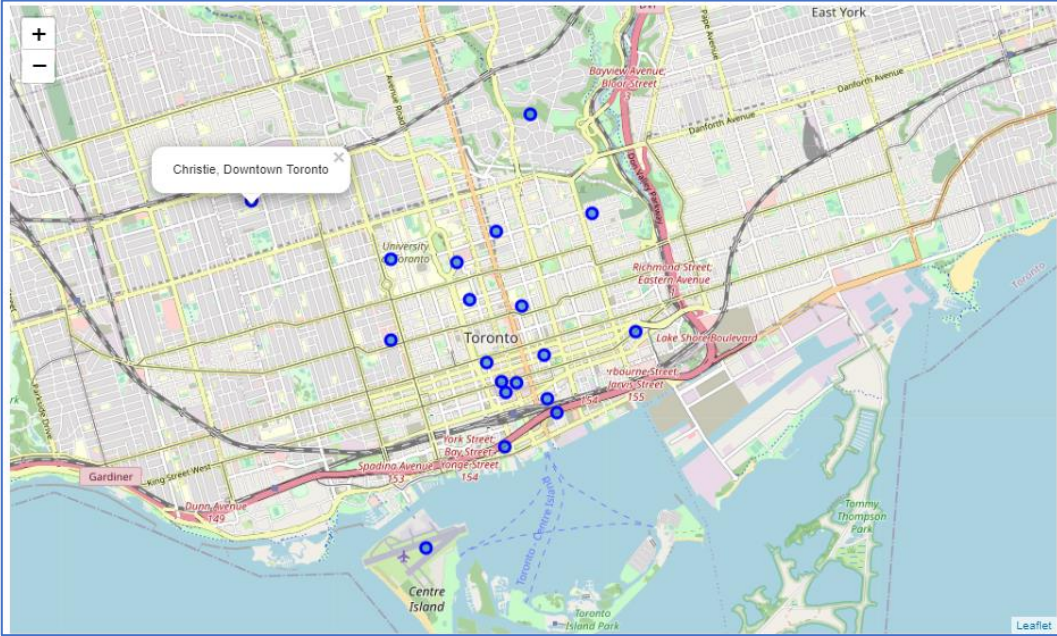


	Postal Code	Borough	Neighbourhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
5	M9A	Etobicoke	Islington Avenue, Humber Valley Village	43.667856	-79.532242
6	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353
7	M3B	North York	Don Mills	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	Brims Hill, Port Union, Highland Creek	43.784535	-79.160497

Now as our problem statement was to analyze area of 'Down town Toronto' so filtering methodology was performed on 'Borough' column to select 'Downtown Toronto' and the final outcome is :

	Postal Code	Borough	Neighbourhood	Latitude	Longitude
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
1	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
2	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
3	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
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, Adelaide, King	43.650571	-79.384568
8	M5J	Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	43.640816	-79.381752
9	M5K	Downtown Toronto	Toronto Dominion Centre, Design Exchange	43.647177	-79.381576
10	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817

Now once the Database was finalized, it was plotted on map using Folium library to better understand the neighborhood distribution in Downtown Toronto area. The circles shown in below map shows each neighborhood in the area.

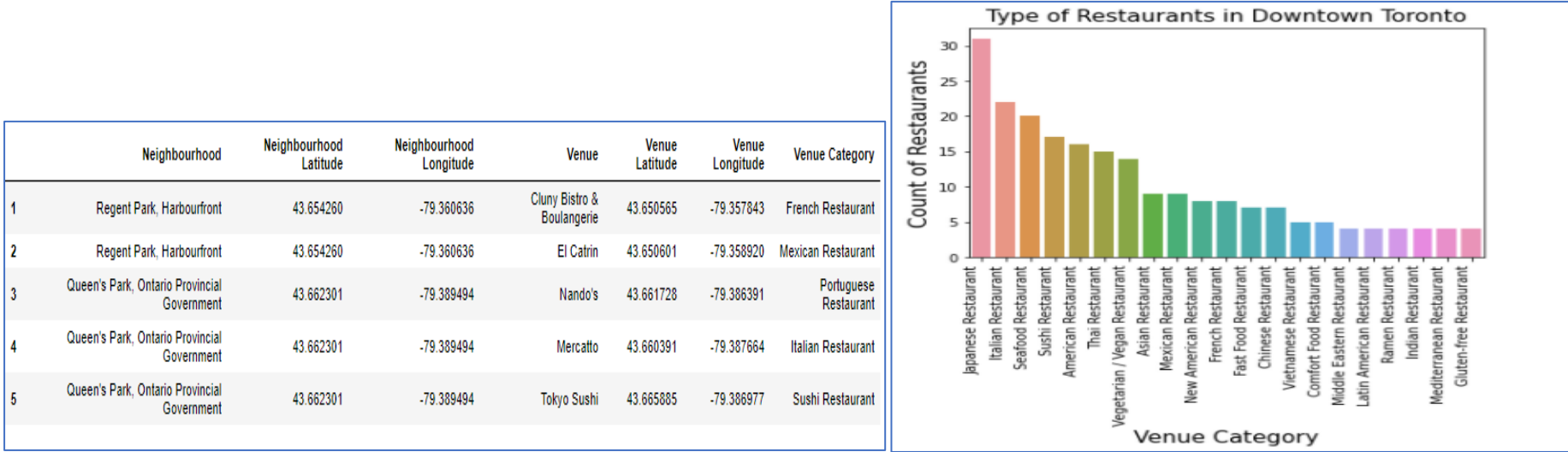


Now the next step was to explore nearby venues in Downtown Toronto area , This was performed using “FOURSQUARE API” In which we provided our finalized table as input and the output was generated initially in form of .json files , From which the data was preprocessed and only desired information was selected.

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

As it is visible that the API categorized all the venues located in a specific neighborhood and provided additional information like name of the venue, category of the venues e.g. Bakery, Coffee shop, Restaurants, spas etc.

Now in consideration of our problem statement, we extracted all Restaurants available in all the neighborhoods which will be used for clustering and tried to convert it into graphical form to get an idea of restaurants which are in Highest number in downtown Toronto. The extraction can support the tourists to identify diverse cuisine restaurants in a specific locality.



Also in consideration of the business statement, Data was further filtered out to check which neighborhoods were having Indian restaurants , the information helped in identifying the neighborhoods where there were many other restaurants and the Indian languages being spoken were in high numbers but Indian restaurant was missing.

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Central Bay Street	43.657952	-79.387383	Colaba Junction	43.660940	-79.385635	Indian Restaurant
1	Harbourfront East, Union Station, Toronto Islands	43.640816	-79.381752	Indian Roti House	43.639060	-79.385422	Indian Restaurant
2	St. James Town, Cabbagetown	43.667967	-79.367675	Butter Chicken Factory	43.667072	-79.369184	Indian Restaurant
3	Church and Wellesley	43.665860	-79.383160	Kothur Indian Cuisine	43.667872	-79.385659	Indian Restaurant

Another information which could be of a great support for tourists is to identify the neighborhood with Highest number of restaurants. Database was refined and it was concluded that First Canadian Place, Underground city is having highest number of restaurants and is visible as follows:

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
	First Canadian Place, Underground city	25	25	25	25	25	25
	Commerce Court, Victoria Hotel	23	23	23	23	23	23
	Kensington Market, Chinatown, Grange Park	21	21	21	21	21	21
	Toronto Dominion Centre, Design Exchange	21	21	21	21	21	21
	Church and Wellesley	21	21	21	21	21	21
	Garden District, Ryerson	20	20	20	20	20	20
	St. James Town	20	20	20	20	20	20
	Richmond, Adelaide, King	19	19	19	19	19	19
	Stn A PO Boxes	19	19	19	19	19	19
	Central Bay Street	19	19	19	19	19	19
	Berczy Park	11	11	11	11	11	11
	St. James Town, Cabbagetown	10	10	10	10	10	10
	Harbourfront East, Union Station, Toronto Islands	10	10	10	10	10	10
	University of Toronto, Harbord	6	6	6	6	6	6
	Queen's Park, Ontario Provincial Government	5	5	5	5	5	5
	Regent Park, Harbourfront	2	2	2	2	2	2
	Christie	1	1	1	1	1	1

Now as the data is cleaned, pre-processed and grouped, all the neighborhoods were analyzed and one-hot encoding technique was applied on the refined dataset in order to rank and extract top 10 restaurants in each neighborhood. The restaurants were grouped based on closest distance and connected to a single neighborhood. Below is the example of information extracted for 'Berczy park' where Seafood restaurants are having highest frequency which states that these are situated at very closest distance and in large number.

----Berczy Park----			
	venue	freq	
0	Seafood Restaurant	0.18	
1	Eastern European Restaurant	0.09	
2	Comfort Food Restaurant	0.09	
3	Italian Restaurant	0.09	
4	Greek Restaurant	0.09	
5	French Restaurant	0.09	
6	Japanese Restaurant	0.09	
7	Vegetarian / Vegan Restaurant	0.09	
8	Thai Restaurant	0.09	
9	Sushi Restaurant	0.09	

Neighborhoods were also grouped to see average(mean) occurrence of each type of restaurant in the specific neighborhoods.

	Neighbourhood	Afghan Restaurant	American Restaurant	Asian Restaurant	Belgian Restaurant	Brazilian Restaurant	Caribbean Restaurant	Chinese Restaurant	Colombian Restaurant	Comfort Food Restaurant	Dim Sum Restaurant	Doner Restaurant	Dumplings Restaurant
0	Berczy Park	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.090909	0.000000	0.000000	0.000
1	Central Bay Street	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
2	Christie	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
3	Church and Wellesley	0.047619	0.047619	0.000000	0.000000	0.000000	0.047619	0.047619	0.000000	0.000000	0.000000	0.000000	0.000
4	Commerce Court, Victoria Hotel	0.000000	0.130435	0.086957	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
5	First Canadian Place, Underground city	0.000000	0.120000	0.120000	0.000000	0.040000	0.000000	0.000000	0.040000	0.000000	0.000000	0.000000	0.000
6	Garden District,	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.050000	0.000000	0.000000	0.000000	0.000000	0.000

Also, information related to most famous restaurant in any specific neighborhood was extracted by grouping neighborhoods and API information. Below are the results:

	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	11th Most Common Venue
0	Berczy Park	Seafood Restaurant	Comfort Food Restaurant	Thai Restaurant	Greek Restaurant	Sushi Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	Japanese Restaurant	French Restaurant	Eastern European Restaurant	Dumpling Restaurant
1	Central Bay Street	Italian Restaurant	Thai Restaurant	Japanese Restaurant	Indian Restaurant	New American Restaurant	French Restaurant	Vegetarian / Vegan Restaurant	Korean Restaurant	Falafel Restaurant	Modern European Restaurant	Middle Eastern Restaurant
2	Christie	Italian Restaurant	Vietnamese Restaurant	Greek Restaurant	German Restaurant	French Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	Eastern European Restaurant	Dumpling Restaurant
3	Church and Wellesley	Sushi Restaurant	Japanese Restaurant	Mediterranean Restaurant	Vietnamese Restaurant	Mexican Restaurant	American Restaurant	Caribbean Restaurant	Chinese Restaurant	Ethiopian Restaurant	Fast Food Restaurant	Indian Restaurant
4	Commerce Court, Victoria Hotel	American Restaurant	Seafood Restaurant	Japanese Restaurant	Thai Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	French Restaurant	Latin American Restaurant	Sushi Restaurant	Gluten-free Restaurant
5	First Canadian Place Underground city	Japanese Restaurant	American Restaurant	Asian Restaurant	Seafood Restaurant	Thai Restaurant	Sushi Restaurant	Gluten-free Restaurant	Greek Restaurant	Mediterranean Restaurant	Vegetarian / Vegan Restaurant	Colombian Restaurant
6	Garden District, Ryerson	Japanese Restaurant	Italian Restaurant	Middle Eastern Restaurant	Fast Food Restaurant	Ramen Restaurant	Vietnamese Restaurant	New American Restaurant	Ethiopian Restaurant	Mexican Restaurant	Chinese Restaurant	Modern European Restaurant
7	Harbourfront East, Union Station, Toronto Islands	Italian Restaurant	Indian Restaurant	Mexican Restaurant	Vegetarian / Vegan Restaurant	Sushi Restaurant	Seafood Restaurant	Japanese Restaurant	Chinese Restaurant	New American Restaurant	Fast Food Restaurant	Falafel Restaurant
8	Kensington Market, Chinatown, Grange Park	Vegetarian / Vegan Restaurant	Mexican Restaurant	Vietnamese Restaurant	Dumpling Restaurant	Comfort Food Restaurant	Filipino Restaurant	Japanese Restaurant	Doner Restaurant	Dim Sum Restaurant	Caribbean Restaurant	Belgian Restaurant
9	Queen's Park, Ontario Provincial Government	Italian Restaurant	Sushi Restaurant	Chinese Restaurant	Portuguese Restaurant	Mexican Restaurant	Vietnamese Restaurant	Doner Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant
10	Regent Park, Harbourfront	French Restaurant	Mexican Restaurant	Vietnamese Restaurant	Greek Restaurant	German Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	Eastern European Restaurant	Dumpling Restaurant
11	Richmond, Adelaide, King	Thai Restaurant	American Restaurant	Sushi Restaurant	Modern European Restaurant	New American Restaurant	Mediterranean Restaurant	Vegetarian / Vegan Restaurant	Fast Food Restaurant	Colombian Restaurant	Gluten-free Restaurant	Japanese Restaurant
12	St. James Town	American Restaurant	Japanese Restaurant	Seafood Restaurant	Moroccan Restaurant	Comfort Food Restaurant	New American Restaurant	Italian Restaurant	Middle Eastern Restaurant	Vegetarian / Vegan Restaurant	Latin American Restaurant	German Restaurant

After analyzing each scenario of extracting highest number of restaurants, most famous restaurants, neighborhoods with Indian restaurants, **Clustering machine algorithm** was applied on the database to group the neighborhoods in relation with most common restaurants in each specific area.

By analyzing the neighborhood information, **Clusters** were divided into 5 categories, this was set using value of **Kmeans**.

Clusters formed are [0,1,2,3,4]

	Postal Code	Borough	Neighbourhood	Latitude	Cluster_Labels	Longitude	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	4	-79.360636	French Restaurant	Mexican Restaurant	Vietnamese Restaurant	Greek Restaurant	German Restaurant	Filipino Restaurant	Fast Food Restaurant
1	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	1	-79.389494	Italian Restaurant	Sushi Restaurant	Chinese Restaurant	Portuguese Restaurant	Mexican Restaurant	Vietnamese Restaurant	Doner Restaurant
2	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	1	-79.378937	Japanese Restaurant	Italian Restaurant	Middle Eastern Restaurant	Fast Food Restaurant	Ramen Restaurant	Vietnamese Restaurant	New American Restaurant
3	M5C	Downtown Toronto	St. James Town	43.651494	3	-79.375418	American Restaurant	Japanese Restaurant	Seafood Restaurant	Moroccan Restaurant	Comfort Food Restaurant	New American Restaurant	Italian Restaurant
4	M5E	Downtown Toronto	Berczy Park	43.644771	3	-79.373306	Seafood Restaurant	Comfort Food Restaurant	Thai Restaurant	Greek Restaurant	Sushi Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant

The above list was a combined one, so each cluster group was extracted separately to analyze the parameters.

CLUSTER1													
	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	8th Most Common Venue
11	Downtown Toronto	University of Toronto, Harbord	43.662696	-79.400049	0	Japanese Restaurant	French Restaurant	Italian Restaurant	Sushi Restaurant	Comfort Food Restaurant	Vietnamese Restaurant	Dumpling Restaurant	Filipino Restaurant

CLUSTER2													
	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	8th Most Common Venue
1	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	1	Italian Restaurant	Sushi Restaurant	Chinese Restaurant	Portuguese Restaurant	Mexican Restaurant	Vietnamese Restaurant	Doner Restaurant	Res
2	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	1	Japanese Restaurant	Italian Restaurant	Middle Eastern Restaurant	Fast Food Restaurant	Ramen Restaurant	Vietnamese Restaurant	New American Restaurant	Et Res
5	Downtown Toronto	Central Bay Street	43.657952	-79.387383	1	Italian Restaurant	Thai Restaurant	Japanese Restaurant	Indian Restaurant	New American Restaurant	French Restaurant	Vegetarian / Vegan Restaurant	Res
8	Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	43.640816	-79.381752	1	Italian Restaurant	Indian Restaurant	Mexican Restaurant	Vegetarian / Vegan Restaurant	Sushi Restaurant	Seafood Restaurant	Japanese Restaurant	C Res
12	Downtown Toronto	Kensington Market, Chinatown, Grange Park	43.653206	-79.400049	1	Vegetarian / Vegan Restaurant	Mexican Restaurant	Vietnamese Restaurant	Dumpling Restaurant	Comfort Food Restaurant	Filipino Restaurant	Japanese Restaurant	Res
16	Downtown Toronto	St. James Town, Cabbagetown	43.667967	-79.367675	1	Italian Restaurant	Chinese Restaurant	Indian Restaurant	Thai Restaurant	Taiwanese Restaurant	Sri Lankan Restaurant	Japanese Restaurant	Ca Res
18	Downtown Toronto	Church and Wellesley	43.665860	-79.383160	1	Sushi Restaurant	Japanese Restaurant	Mediterranean Restaurant	Vietnamese Restaurant	Mexican Restaurant	American Restaurant	Caribbean Restaurant	C Res

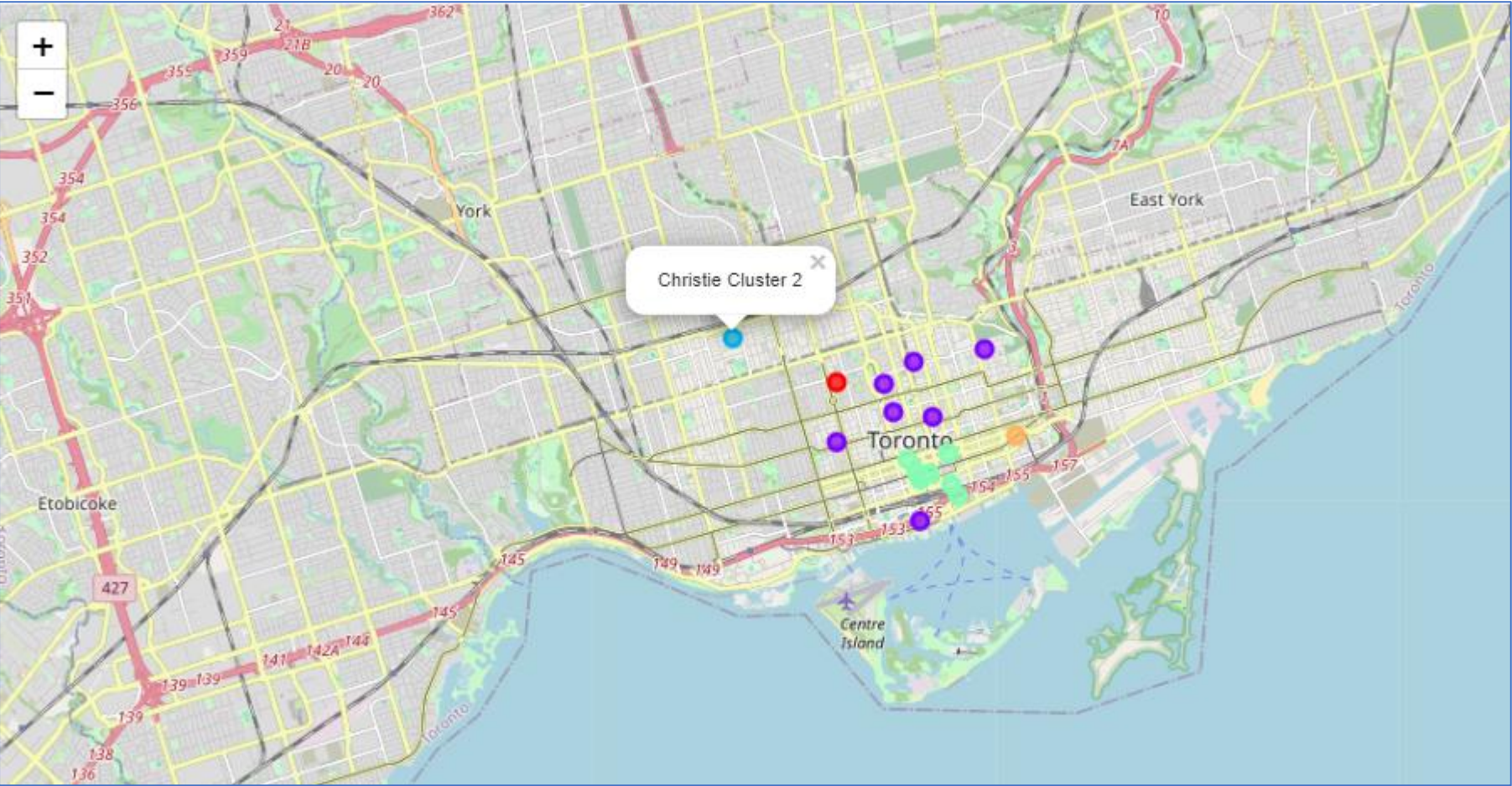
CLUSTER3													
	Borough	Neighbourhood	Latitude	Longitude	Cluster_Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
6	Downtown Toronto	Christie	43.669542	-79.422564	2	Italian Restaurant	Vietnamese Restaurant	Greek Restaurant	German Restaurant	French Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant

CLUSTER4													
	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	8th Most Common Venue
3	Downtown Toronto	St. James Town	43.651494	-79.375418	3	American Restaurant	Japanese Restaurant	Seafood Restaurant	Moroccan Restaurant	Comfort Food Restaurant	New American Restaurant	Italian Restaurant	Fast Food Restaurant
4	Downtown Toronto	Berczy Park	43.644771	-79.373306	3	Seafood Restaurant	Comfort Food Restaurant	Thai Restaurant	Greek Restaurant	Sushi Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	Fast Food Restaurant
7	Downtown Toronto	Richmond, Adelaide, King	43.650571	-79.384568	3	Thai Restaurant	American Restaurant	Sushi Restaurant	Modern European Restaurant	New American Restaurant	Mediterranean Restaurant	Vegetarian / Vegan Restaurant	Fast Food Restaurant
9	Downtown Toronto	Toronto Dominion Centre, Design Exchange	43.647177	-79.381576	3	American Restaurant	Seafood Restaurant	Japanese Restaurant	Asian Restaurant	Italian Restaurant	French Restaurant	Fast Food Restaurant	Fast Food Restaurant
10	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817	3	American Restaurant	Seafood Restaurant	Japanese Restaurant	Thai Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	Fast Food Restaurant
15	Downtown Toronto	Stn A PO Boxes	43.646435	-79.374846	3	Italian Restaurant	Seafood Restaurant	Japanese Restaurant	Molecular Gastronomy Restaurant	American Restaurant	Thai Restaurant	Fast Food Restaurant	Fast Food Restaurant
17	Downtown Toronto	First Canadian Place, Underground city	43.648429	-79.382280	3	Japanese Restaurant	American Restaurant	Asian Restaurant	Seafood Restaurant	Thai Restaurant	Sushi Restaurant	Gluten-free Restaurant	Fast Food Restaurant

CLUSTER5													
	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	8th Most Common Venue
0	Downtown Toronto	Regent Park, Harbourfront	43.65426	-79.360636	4	French Restaurant	Mexican Restaurant	Vietnamese Restaurant	Greek Restaurant	German Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant

The grouped clusters created using machine learning were also presented on 'Downtown Toronto' map to check geographical information of the clusters. This was performed using 'Folium library'.

By visualization of map, once cursor is moved on the circles it will represent the grouped cluster number and the neighborhood associated with it. Important thing to note here is that cluster number starts with 0 on map (which is considered as cluster1).

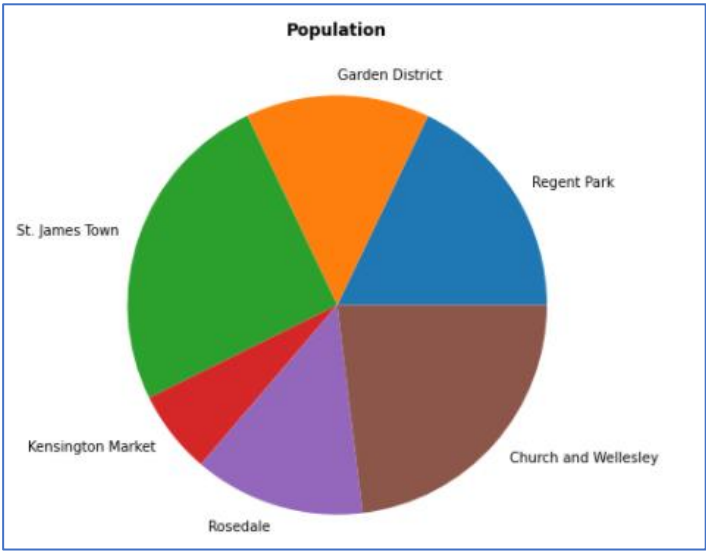


After identifying most famous restaurants and clustering them based on neighborhoods, Next step was to extract demographic information from Wikipedia page for Toronto city. As the data was not clean so cleaning process was performed by modifying the neighborhoods and converting the names in a format which was matching with Toronto neighborhood data.

	Second most common language (after English) by name	Population	Neighbourhood
0	NaN	5113149.0	Toronto CMA Average
1	Cantonese (19.3%)	44577.0	Agincourt
2	Cantonese (17.9%)	4355.0	Alexandra Park
3	Russian (1.4%)	2513.0	Allenby
4	Spanish (6.1%)	17318.0	Amesbury

After cleaning and processing the tables, data was merged with Toronto neighborhood data to extract information of Population, most common language being spoken in the specific neighborhoods. Now the output of the table as shown below was mapped with grouped clusters which were created using machine learning algorithm to identify which specific cluster and neighborhood are having many restaurants but any 'Indian Restaurant' is not in even top 15 restaurants, The filtering results were mapped with the Neighborhoods where second most common language being spoken was of Indian community and the population was in mid or high category.

	Borough	Neighbourhood	Second most common language (after English) by name	Population
0	Downtown Toronto	Regent Park	Bengali (10.5%)	10387.0
1	Downtown Toronto	Garden District	Unspecified Chinese (3.0%)	8240.0
2	Downtown Toronto	St. James Town	Filipino (8.1%)	14666.0
4	Downtown Toronto	Kensington Market	Cantonese (14.3%)	3740.0
5	Downtown Toronto	Rosedale	Unspecified Chinese (1.0%)	7672.0
6	Downtown Toronto	Church and Wellesley	Spanish (1.8%)	13397.0



Here in cluster 5 There is no Indian restaurant in Regent park area, but the most spoken language is ‘Bengali’

CLUSTER5														
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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	11th Most Common Venue
Regent Park, Harbourfront	43.65426	-79.360636	4	French Restaurant	Mexican Restaurant	Vietnamese Restaurant	Greek Restaurant	German Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	Indian Restaurant	R

Discussion: -

After performing Exploratory data analysis and visualization techniques on extracted datasets, it was observed that in Downtown Toronto there are variety of Cuisine choices available , Also as there is more than 45% of population migrated from all around the globe, Toronto city is one of the best choices where Tourists from every community can visit and live here. From the data analyzed it was noticed that there are 2 Clusters (Cluster 2 and 4) which should be in the first to visit list of tourists as there are variety of Restaurants in almost all the neighborhoods.

Cluster 2 should be the best choice for tourists or residence which belong to European countries or someone who would like to prefer Japanese food.

The reason was that in most of the neighborhoods Italian and Japanese restaurants are among the topmost venues. Also, Cluster 2 is having ‘Indian Restaurants’ and vegan restaurants in many neighborhoods so it would be a best choice for Indian / vegetarian community, But this cluster **cannot** be the best choice for opening an Indian restaurant which is a part of our Business statement as there are already many Indian restaurants existing in many areas.

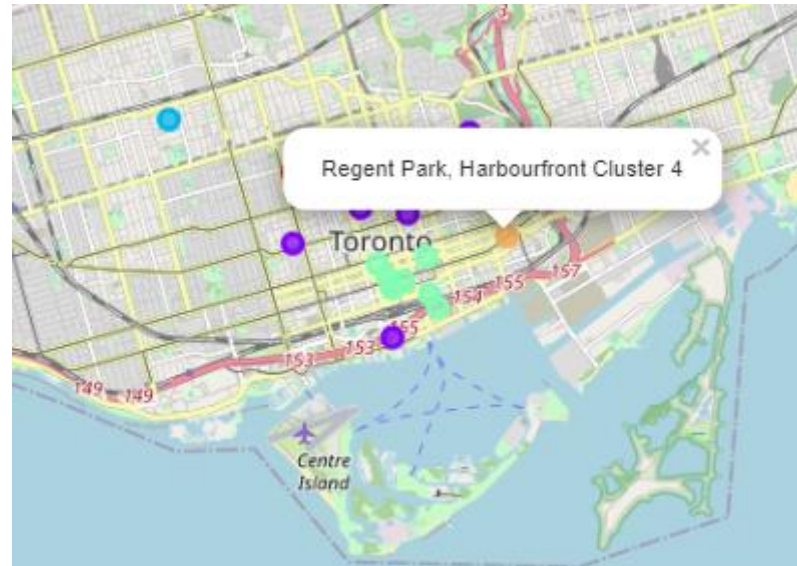
CLUSTER2									
	Borough	Neighbourhood	Latitude	Longitude	Cluster_Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	
1	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	1	Italian Restaurant	Sushi Restaurant	Chinese Restaurant	
2	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	1	Japanese Restaurant	Italian Restaurant	Middle Eastern Restaurant	
5	Downtown Toronto	Central Bay Street	43.657952	-79.387383	1	Italian Restaurant	Thai Restaurant	Japanese Restaurant	
8	Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	43.640816	-79.381752	1	Italian Restaurant	Indian Restaurant	Mexican Restaurant	
12	Downtown Toronto	Kensington Market, Chinatown, Grange Park	43.653206	-79.400049	1	Vegetarian / Vegan Restaurant	Mexican Restaurant	Vietnamese Restaurant	
16	Downtown Toronto	St. James Town, Cabbagetown	43.667967	-79.367675	1	Italian Restaurant	Chinese Restaurant	Indian Restaurant	
18	Downtown Toronto	Church and Wellesley	43.665860	-79.383160	1	Sushi Restaurant	Japanese Restaurant	Mediterranean Restaurant	

Cluster 4 should be the best choice for American community as most of the neighborhoods are having American restaurants in Top 3 most common venues. Another thing related to our business problem statement is that as this cluster is having many restaurants in all the neighborhoods but an Indian restaurant is not in Top 15 most common venues so Cluster 4 can be considered as one of the choices to open an Indian restaurant.

CLUSTER4									
	Borough	Neighbourhood	Latitude	Longitude	Cluster_Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	
3	Downtown Toronto	St. James Town	43.651494	-79.375418	3	American Restaurant	Japanese Restaurant	Seafood Restaurant	
4	Downtown Toronto	Berczy Park	43.644771	-79.373306	3	Seafood Restaurant	Comfort Food Restaurant	Thai Restaurant	
7	Downtown Toronto	Richmond, Adelaide, King	43.650571	-79.384568	3	Thai Restaurant	American Restaurant	Sushi Restaurant	
9	Downtown Toronto	Toronto Dominion Centre, Design Exchange	43.647177	-79.381576	3	American Restaurant	Seafood Restaurant	Japanese Restaurant	
10	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817	3	American Restaurant	Seafood Restaurant	Japanese Restaurant	
15	Downtown Toronto	Stn A PO Boxes	43.646435	-79.374046	3	Italian Restaurant	Seafood Restaurant	Japanese Restaurant	
17	Downtown Toronto	First Canadian Place, Underground Parking	43.648429	-79.382280	3	Japanese Restaurant	American Restaurant	Asian Restaurant	

After performing demographic analysis and predicting the community living in the area Cluster 5 should be the best choice for opening an Indian restaurant because of following reasons

1. Cluster 5 is having many restaurants in 'Regent park' which is the 4th largest community in terms of population and the 2nd most common language being spoken is 'Bengali'
2. In Cluster 5, No Indian restaurant is in top 15 most common venues thus opening an Indian restaurant may attract more population from the neighborhood and hence might yield more revenue and increase in sales.



Cluster 4 is a naming convention generated by Clustering algorithm. Which we considered as cluster 5.

Conclusion: -

With the help all the datasets analyzed, Foursquare API was of a great support in extracting valuable insights related to most common and famous venues like bars, restaurants etc. These insights were a great support in analyzing our business problem statement. By combining all information and performing visualization we were able to Unify variety of restaurants and were mapped with specific Neighborhoods. This could be of a great support for tourists or people who prefer to have Cuisine choices of their countries that they can easily follow clusters of restaurants in their nearby Neighborhoods.

By further analyzing clusters created using Kmeans clustering algorithm and the demographic data, there were certain clusters/neighborhoods identified where there is a business opportunity of opening an Indian restaurant and the recommendation of opening in certain cluster or in certain neighborhood could be followed to yield more revenue and increase in sales.