COURSERA CAPSTONE PROJECT PRESENTATION

INTRODUCTION :

- The basic idea behind this project is to cluster the neighborhoods in Toronto to get the best places to open a restaurant, a coffee shop, or a hotel in Toronto.
- Since Toronto is the financial capital of Canada it is safe to assume that there will be an interest to open new hotels, restaurants, location or coffee shops in there
- In this project I will try to explore the best location to do so.

DATA

- The data used in this project was originally obtained from here
- The raw data didn’t give much information since a lot of its borough and neighborhoods were labeled as “Not assigned”, so I cleaned the data and added the latitude and longitude to each neighborhood. (Images shown in the following slide)
- A map of Toronto with different coloring for each borough is in the next slide

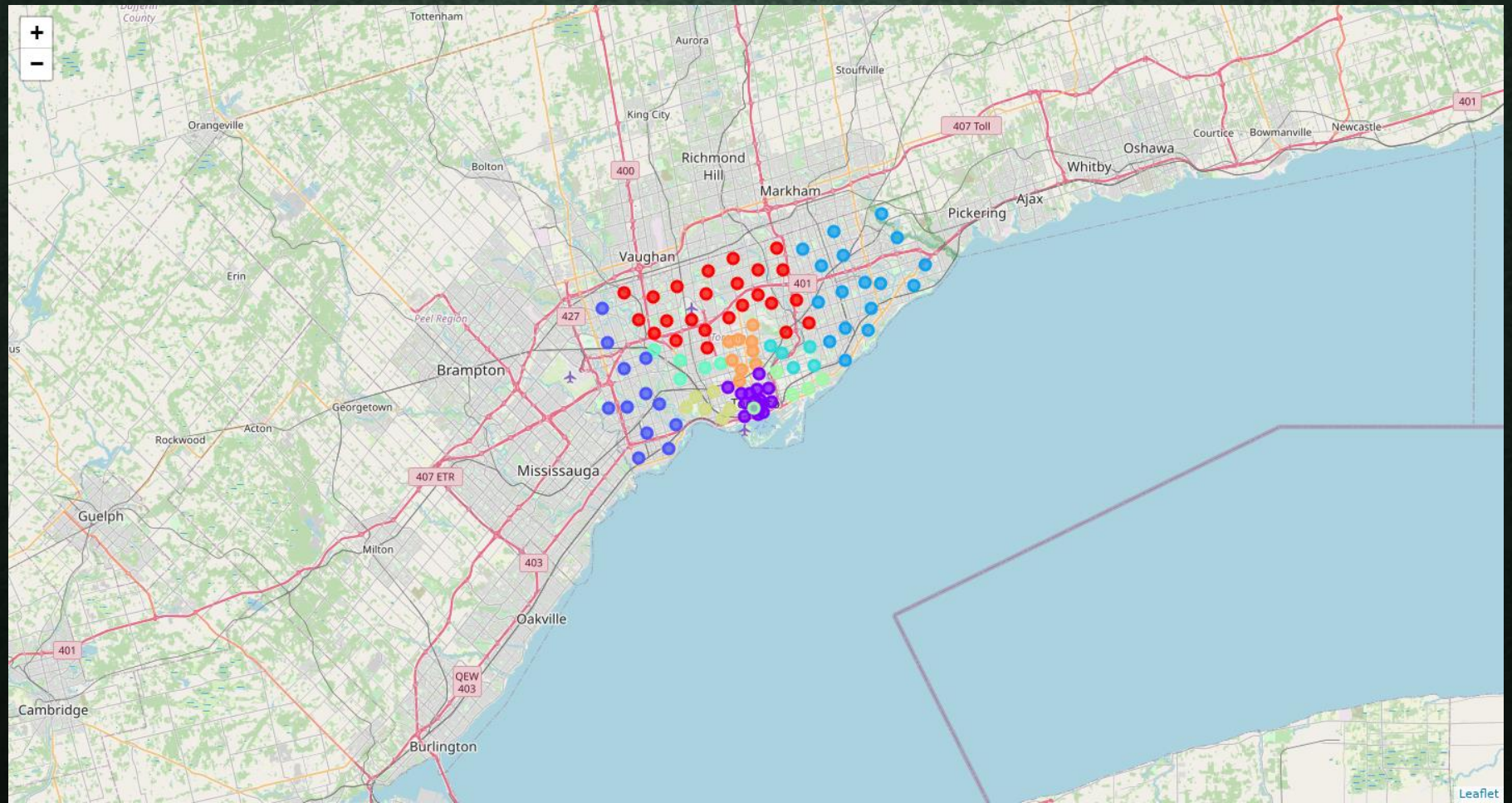
How it was

Postal Code	Borough	Neighbourhood
M1A	Not assigned	Not assigned
M2A	Not assigned	Not assigned
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Regent Park, Harbourfront
M6A	North York	Lawrence Manor, Lawrence Heights
M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government
M8A	Not assigned	Not assigned
M9A	Etobicoke	Islington Avenue, Humber Valley Village
M1B	Scarborough	Malvern, Rouge
M2B	Not assigned	Not assigned

How it became

	Postal Code	Borough	Neighborhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.75245	-79.32991
1	M4A	North York	Victoria Village	43.73057	-79.31306
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.65512	-79.36264
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.72327	-79.45042
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.66253	-79.39188
5	M9A	Etobicoke	Islington Avenue, Humber Valley Village	43.66263	-79.52831
6	M1B	Scarborough	Malvern, Rouge	43.81139	-79.19662
7	M3B	North York	Don Mills	43.74923	-79.36186
8	M4B	East York	Parkview Hill, Woodbine Gardens	43.70718	-79.31192
9	M5B	Downtown Toronto	Garden District, Ryerson	43.65739	-79.37804
10	M6B	North York	Glencairn	43.70687	-79.44812
11	M9B	Etobicoke	West Deane Park, Princess Gardens, Martin Grov...	43.65034	-79.55362
12	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek	43.78574	-79.15875

THE MAP


Tottenham
Aurora
Stouffville
King City
Orangeville
Bolton
Richmond Hill
Markham
Pickering
Ajax
Whitby
Oshawa
Courtice
Bowmanville
Newcastle
Vaughan
Erin
Peel Region
Brampton
Georgetown
Acton
Rockwood
Mississauga
Guelph
Milton
Oakville
Cambridge
Burlington
400
401
427
407 Toll
407 ETR
403
QEW
Leaflet

- The last thing to do was use the foursquare API to get the needed data however I had to use the API to search for each venue alone then I combined the data and it was like this .

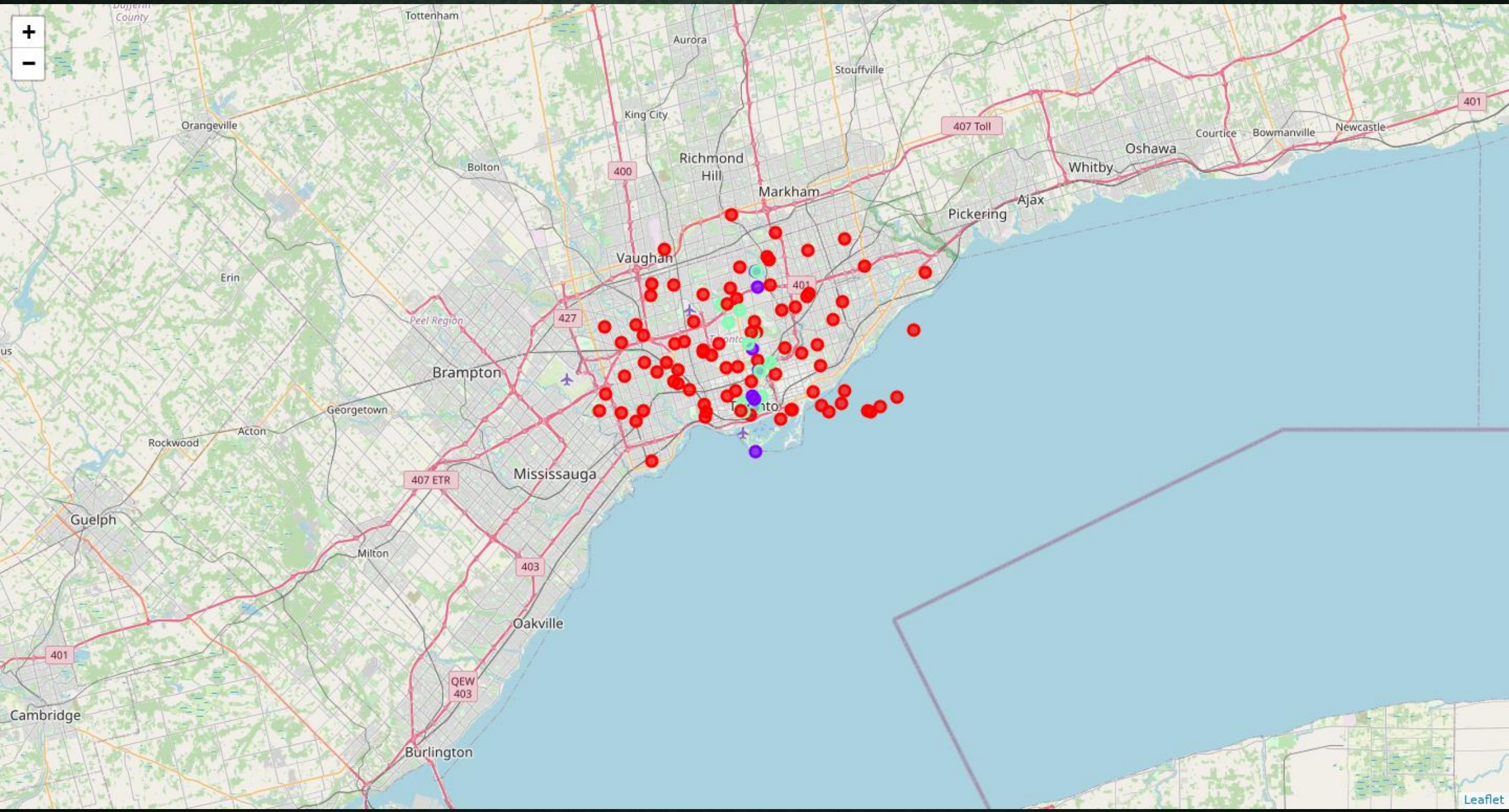
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue Category	Venue	Venue Latitude	Venue Longitude
0	Parkwoods	43.75245	-79.32991	Hotel	Towns On The Ravine	43.754754	-79.332552
1	Garden District, Ryerson	43.65739	-79.37804	Hotel	The Grand Hotel & Suites Toronto	43.656449	-79.374110
2	Garden District, Ryerson	43.65739	-79.37804	Hotel	Marriott Downtown at CF Toronto Eaton Centre	43.654728	-79.382422
3	Garden District, Ryerson	43.65739	-79.37804	Hotel	The Saint James Hotel	43.659398	-79.380932
4	Garden District, Ryerson	43.65739	-79.37804	Hotel	Chelsea Hotel	43.658498	-79.383097

- In order to be able to use the machine learning algorithm I had to change the categorical data to values the final data frame before using the algorithm was like this.

	Neighborhood	Coffee Shop	Hotel	Restaurant	Postal Code	Borough	Latitude	Longitude
0	Agincourt	4	0	6	M1S	Scarborough	43.79452	-79.26708
1	Alderwood, Long Branch	3	0	3	M8W	Etobicoke	43.60124	-79.53879
2	Bathurst Manor, Wilson Heights, Downsview North	2	0	0	M3H	North York	43.75788	-79.44847
3	Bayview Village	1	0	0	M2K	North York	43.78112	-79.38060
4	Bedford Park, Lawrence Manor East	3	0	12	M5M	North York	43.73545	-79.41916
...	...	...	...	...	...	...	...	...
96	Willowdale, Willowdale East	14	1	6	M2N	North York	43.76774	-79.40728
97	Willowdale, Willowdale West	1	0	0	M2R	North York	43.77989	-79.44678
98	Woburn	2	0	3	M1G	Scarborough	43.76812	-79.21761
99	Woodbine Heights	9	0	2	M4C	East York	43.68970	-79.30682
100	York Mills West	4	0	1	M2P	North York	43.74778	-79.40033

METHODOLOGY

- K-means algorithm was used, with k set to be 3, in order to cluster the neighborhood in three different categories.
- A map of the corresponding clusters is in the next slide.


Tottenham
Aurora
Stouffville
King City
Orangeville
Bolton
Richmond Hill
Markham
Pickering
Ajax
Whitby
Oshawa
Courtice
Bowmanville
Newcastle
407 Toll
401
400
Vaughan
Erin
Peel Region
427
Brampton
Georgetown
Acton
Rockwood
Mississauga
407 ETR
Guelph
Milton
403
Oakville
QEW
Cambridge
Burlington
Leaflet

THE FOLLOWING SLIDES CONTAIN THE RESULTING DATA FRAMES OF EACH CLUSTER.

FIRST CLUSTER

	Coffee Shop Count	Hotel Count	Restaurant Count	Latitude	Longitude	Cluster Labels	Neighborhood	Borough
0	4	0	6	43.79452	-79.26708	0	Agincourt	Scarborough
1	3	0	3	43.60124	-79.53879	0	Alderwood, Long Branch	Etobicoke
2	2	0	0	43.75788	-79.44847	0	Bathurst Manor, Wilson Heights, Downsview North	North York
3	1	0	0	43.78112	-79.38060	0	Bayview Village	North York
4	3	0	12	43.73545	-79.41916	0	Bedford Park, Lawrence Manor East	North York
...	...	...	...	...	...	...	...	...
96	14	1	6	43.76774	-79.40728	0	Willowdale, Willowdale East	North York
97	1	0	0	43.77989	-79.44678	0	Willowdale, Willowdale West	North York
98	2	0	3	43.76812	-79.21761	0	Woburn	Scarborough
99	9	0	2	43.68970	-79.30682	0	Woodbine Heights	East York
100	4	0	1	43.74778	-79.40033	0	York Mills West	North York

SECOND CLUSTER

	Coffee Shop Count	Hotel Count	Restaurant Count	Latitude	Longitude	Cluster Labels	Neighborhood	Borough
8	46	17	56	43.64869	-79.38544	1	Business reply mail Processing Centre, South C...	East Toronto
11	46	17	56	43.64869	-79.38544	1	Canada Post Gateway Processing Centre	Mississauga
18	43	19	51	43.64840	-79.37914	1	Commerce Court, Victoria Hotel	Downtown Toronto
33	43	21	47	43.64828	-79.38146	1	First Canadian Place, Underground city	Downtown Toronto
35	53	12	50	43.65739	-79.37804	1	Garden District, Ryerson	Downtown Toronto
69	45	21	49	43.64970	-79.38258	1	Richmond, Adelaide, King	Downtown Toronto
80	46	17	56	43.64869	-79.38544	1	Stn A PO Boxes	Downtown Toronto
88	41	23	46	43.64710	-79.38153	1	Toronto Dominion Centre, Design Exchange	Downtown Toronto

THIRD CLUSTER

	Coffee Shop Count	Hotel Count	Restaurant Count	Latitude	Longitude	Cluster Labels	Neighborhood	Borough
5	39	4	31	43.64536	-79.37306	2	Berczy Park	Downtown Toronto
7	29	1	24	43.63941	-79.42676	2	Brockton, Parkdale Village, Exhibition Place	West Toronto
9	41	2	30	43.64082	-79.39818	2	CN Tower, King and Spadina, Railway Lands, Har...	Downtown Toronto
13	56	7	26	43.65609	-79.38493	2	Central Bay Street	Downtown Toronto
15	54	6	35	43.66659	-79.38133	2	Church and Wellesley	Downtown Toronto
39	40	10	21	43.64285	-79.38076	2	Harbourfront East, Union Station, Toronto Islands	Downtown Toronto
48	74	2	24	43.65351	-79.39722	2	Kensington Market, Chinatown, Grange Park	Downtown Toronto
53	47	0	26	43.64848	-79.41774	2	Little Portugal, Trinity	West Toronto
67	68	0	9	43.66253	-79.39188	2	Queen's Park, Ontario Provincial Government	Downtown Toronto
68	32	0	7	43.65512	-79.36264	2	Regent Park, Harbourfront	Downtown Toronto
77	44	14	34	43.65215	-79.37587	2	St. James Town	Downtown Toronto
89	55	1	17	43.66311	-79.40180	2	University of Toronto, Harbord	Downtown Toronto

RESULTS

- Referring to the above images it can be found that :
- The first cluster has relatively low coffee shops and restaurants , and almost no hotels
- The second cluster ,however, has the highest count of any cluster
- The third cluster is in the middle

CONCLUSION

- Having checked the data we can presume that the best places to open the venues are in cluster 2 for hotels since it has the most count ,the third for restaurants and coffees can be opened anywhere