Leverage Analytics to select the Next Best Location for Italian Restaurant

Introduction

For this capstone, we will solve a business problem which deals with opening of a new restaurant in the city of Toronto by leveraging insights from Data Analytics and Machine Learning. The objective is find the optimal solution based on data understanding, inference and find the one that is feasible.

Problem Definition

In this scenario, Person A wants to open an Italian Restaurant in Toronto. The objective is to solve the problem by using location data from Foursquare and by doing a study of restaurants on all neighborhoods on interest through a combination of location profiling and machine learning.

We are going to explore the neighborhoods of Toronto and their venues in order to gain a better understanding on existing places and to narrow down our neighborhood choices to make the best decision.

Data Understanding

We are going to explore the neighborhoods of Toronto and their venues in order to gain a better understanding on existing places and to narrow down our neighborhood choices. For this purpose, we were able to explore a total of 38 unique neighborhoods with a total of 674 venues spanning across 187 different venue categories. We further preprocess this data to extract only restaurant categories and we then drill down only to Italian restaurants in specific to solve the problem.

Methodology

For the purpose of this project, we leveraged four square data using their API to explore all the neighborhoods of Toronto. Leveraging the API, we can study the top venues in each neighborhoods and can hence fine tune the optimal location for the new Italian Restaurant in Toronto. We then use Machine Learning algorithms like Clustering to find hidden behaviors in the data and then use the results of the clustering to drive the Segmentation Analysis which can help us to narrow down the location. The following steps shows us various steps in our methodology.

a) Visualize the Data

As like every other project, we perform data visualization once we extract the data from Foursquare API. The following figure shows a map of the neighborhoods based on geo coordinates leveraged through Four Square API. The blue circle represents all the different neighborhoods.

Figure 1: Neighborhoods of Toronto

The following figure shows a snapshot of the neighborhoods and the corresponding geo coordinates and the venue categories leveraged through Four Square API.

b) Explore the Data

The second step is to do Data Exploration where we deep dive into specific attributes of interest to validate our gut feeling into inferences from data. Here we explore the distribution of different venues for each neighborhood and then deep dive into Restaurant Categories.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	The Beaches	43.676357	-79.293031	The Big Carrot Natural Food Market	43.678879	-79.297734	Health Food Store
1	The Beaches	43.676357	-79.293031	Grover Pub and Grub	43.679181	-79.297215	Pub
2	The Beaches	43.676357	-79.293031	Starbucks	43.678798	-79.298045	Coffee Shop
3	The Beaches	43.676357	-79.293031	Upper Beaches	43.680563	-79.292869	Neighborhood
4	Riverdale,The Danforth West	43.679557	-79.352188	Pantheon	43.677621	-79.351434	Greek Restaurant
5	Riverdale,The Danforth West	43.679557	-79.352188	Dolce Gelato	43.677773	-79.351187	Ice Cream Shop
6	Riverdale,The Danforth West	43.679557	-79.352188	MenEssentials	43.677820	-79.351265	Cosmetics Shop
7	Riverdale,The Danforth West	43.679557	-79.352188	Messini Authentic Gyros	43.677827	-79.350569	Greek Restaurant
8	Riverdale,The Danforth West	43.679557	-79.352188	Mezes	43.677962	-79.350196	Greek Restaurant
9	Riverdale,The Danforth West	43.679557	-79.352188	Cafe Fiorentina	43.677743	-79.350115	Italian Restaurant

Figure 2: Exploring venues

The following figure shows a breakdown of number of venues that was extracted from each neighborhood. 30 is the maximum number of venues that we can extract per neighborhood. It was set as cap for the purpose of this project. You can see that some neighborhood have very few venues.

```
Neighborhood
Adelaide,Richmond,King                                                                30
Berczy Park                                                                           30
Business Reply Mail Processing Centre 969 Eastern                                     19
Central Bay Street                                                                    30
Christie                                                                              16
Church and Wellesley                                                                  30
Commerce Court,Victoria Hotel                                                         30
Davisville                                                                            30
Davisville North                                                                       9
Design Exchange,Toronto Dominion Centre                                               30
Dovercourt Village,Dufferin                                                           20
Exhibition Place,Parkdale Village,Brockton                                            19
Forest Hill West,Forest Hill North                                                     4
Harbourfront East,Toronto Islands,Union Station                                       30
Harbourfront,Regent Park                                                              30
High Park,The Junction South                                                          23
Kensington Market,Grange Park,Chinatown                                               30
Lawrence Park                                                                          4
Little Portugal,Trinity                                                               30
North Toronto West                                                                    21
Parkdale,Roncesvalles                                                                 16
Riverdale,The Danforth West                                                           30
Rosedale                                                                               4
Roselawn                                                                               3
Ryerson,Garden District                                                               30
South Niagara,Bathurst Quay,King and Spadina,Railway Lands,CN Tower,Harbourfront West,Island airport    13
St. James Town                                                                        30
St. James Town,Cabbagetown                                                            30
Stn A PO Boxes 25 The Esplanade                                                       30
Studio District                                                                       30
Summerhill East,Moore Park                                                             3
Summerhill West,Forest Hill SE,Deer Park,South Hill,Rathnelly                         15
Swansea,Runnymede                                                                     30
The Beaches                                                                            4
The Beaches West,India Bazaar                                                         19
Underground city,First Canadian Place                                                 30
University of Toronto,Harbord                                                         30
Yorkville,The Annex,North Midtown                                                     24
Name: Venue Longitude, dtype: int64
```

Figure 3: Distribution of venues

The following figure a breakdown of all venues for a particular neighborhood called St. James Town as an example. We can see that venue categories range from Hostels, Church, Gym, etc. So the ideal next step is to filter the data by Restaurants as it is the Venue Category of Interest.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
289	St. James Town	43.651494	-79.375418	Terroni	43.650927	-79.375602	Italian Restaurant
290	St. James Town	43.651494	-79.375418	Gyu-Kaku Japanese BBQ	43.651422	-79.375047	Japanese Restaurant
291	St. James Town	43.651494	-79.375418	GEORGE Restaurant	43.653346	-79.374445	Restaurant
292	St. James Town	43.651494	-79.375418	Crepe TO	43.650063	-79.374587	Creperie
293	St. James Town	43.651494	-79.375418	Fahrenheit Coffee	43.652384	-79.372719	Coffee Shop
294	St. James Town	43.651494	-79.375418	Triple A Bar (AAA)	43.651658	-79.372720	BBQ Joint
295	St. James Town	43.651494	-79.375418	Pearl Diver	43.651481	-79.373600	Gastropub
296	St. James Town	43.651494	-79.375418	Hogtown Smoke	43.649287	-79.374689	Food Truck
297	St. James Town	43.651494	-79.375418	Mystic Muffin	43.652484	-79.372655	Middle Eastern Restaurant
298	St. James Town	43.651494	-79.375418	St James Anglican Cathedral	43.650110	-79.374292	Church
299	St. James Town	43.651494	-79.375418	Aveda Institute Toronto	43.650096	-79.373630	Cosmetics Shop
300	St. James Town	43.651494	-79.375418	Hawthorne Food and Drink	43.652270	-79.376318	Gastropub
301	St. James Town	43.651494	-79.375418	Club 120	43.652100	-79.375522	Performing Arts Venue
302	St. James Town	43.651494	-79.375418	Beerbistro	43.649419	-79.377237	Gastropub
303	St. James Town	43.651494	-79.375418	The Carbon Bar	43.653367	-79.374965	Restaurant
304	St. James Town	43.651494	-79.375418	Cambridge Suites Toronto	43.651836	-79.378107	Hotel
305	St. James Town	43.651494	-79.375418	GoodLife Fitness	43.651136	-79.378797	Gym
306	St. James Town	43.651494	-79.375418	NAMI	43.650853	-79.375887	Japanese Restaurant
307	St. James Town	43.651494	-79.375418	Versus Coffee	43.651213	-79.375236	Coffee Shop
308	St. James Town	43.651494	-79.375418	Dineen Coffee	43.650497	-79.378765	Café
309	St. James Town	43.651494	-79.375418	The Omni King Edward Hotel	43.649191	-79.376006	Hotel
310	St. James Town	43.651494	-79.375418	Richmond Station	43.651569	-79.379266	American Restaurant
311	St. James Town	43.651494	-79.375418	Elgin And Winter Garden Theatres	43.653394	-79.378507	Theater
312	St. James Town	43.651494	-79.375418	Carisma	43.649617	-79.375434	Italian Restaurant
313	St. James Town	43.651494	-79.375418	The Poké Box	43.650469	-79.376317	Poke Place
314	St. James Town	43.651494	-79.375418	Sukhothai	43.648487	-79.374547	Thai Restaurant
315	St. James Town	43.651494	-79.375418	The Chase	43.650952	-79.379422	New American Restaurant
316	St. James Town	43.651494	-79.375418	The George Street Diner	43.652803	-79.371563	Diner
317	St. James Town	43.651494	-79.375418	La Bettola Di Terroni	43.651993	-79.378056	Italian Restaurant
318	St. James Town	43.651494	-79.375418	HI-Toronto Hostel	43.650623	-79.375190	Hostel

Figure 4: Example of St. James Town

Results

Once we completed data visualization and some preliminary data discovery, we then proceed to exploration of the data to solve the business case in hand.

a) Filter only be Restaurants

The following figure shows a breakdown of venues only by Restaurant Venue Category. We found 4 neighborhoods do not have any restaurants and hence are not considered for the analysis.

	Neighborhood	Neighborhood_Latitude	Neighborhood_Longitude	Venue	Venue_Latitude	Venue_Longitude	Venue_Category
4	Riverdale,The Danforth West	43.679557	-79.352188	Pantheon	43.677621	-79.351434	Greek Restaurant
7	Riverdale,The Danforth West	43.679557	-79.352188	Messini Authentic Gyros	43.677827	-79.350569	Greek Restaurant
8	Riverdale,The Danforth West	43.679557	-79.352188	Mezes	43.677962	-79.350196	Greek Restaurant
9	Riverdale,The Danforth West	43.679557	-79.352188	Cafe Fiorentina	43.677743	-79.350115	Italian Restaurant
10	Riverdale,The Danforth West	43.679557	-79.352188	Christina's On The Danforth	43.678240	-79.349185	Greek Restaurant
14	Riverdale,The Danforth West	43.679557	-79.352188	7 Numbers	43.677062	-79.353934	Italian Restaurant
18	Riverdale,The Danforth West	43.679557	-79.352188	Rikkochez	43.677267	-79.353274	Restaurant
19	Riverdale,The Danforth West	43.679557	-79.352188	Pan on the Danforth	43.678263	-79.348648	Greek Restaurant
20	Riverdale,The Danforth West	43.679557	-79.352188	Astoria Shish Kebob House	43.677689	-79.351892	Greek Restaurant
24	Riverdale,The Danforth West	43.679557	-79.352188	Ouzeri	43.678193	-79.348908	Greek Restaurant

Figure 5: Filter only by Restaurant Categories

b) Use Frequency Distribution to Find the Popular Restaurant Category

We then use one hot encoding to list all restaurants for each of the 34 neighborhoods. The following figure shows a snapshot of the data. For each of the neighborhoods, we create one hot encoding for 36 restaurant categories. 1 means the restaurant exits and 0 means it does not.

	Neighborhood	American Restaurant	Asian Restaurant	Belgian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Comfort Food Restaurant	Cuban Restaurant	Dim Sum Restaurant	...	Portuguese Restaurant	Ramen Restaurant	Restaurant	Seafood Restaurant	Sushi Restaurant	Taiwanese Restaurant	Thai Restaurant	Theme Restaurant	Vegetarian / Vegan Restaurant	Vietnamese Restaurant
4	Riverdale,The Danforth West	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
7	Riverdale,The Danforth West	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
8	Riverdale,The Danforth West	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
9	Riverdale,The Danforth West	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
10	Riverdale,The Danforth West	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0

5 rows × 37 columns

Figure 6: One Hot Encoding of Restaurant Venue Categories

We then counts all the 1s and get a percentage frequency value for each neighborhood for each category.

	Neighborhood	American Restaurant	Asian Restaurant	Belgian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Comfort Food Restaurant	Cuban Restaurant	Dim Sum Restaurant	...	Portuguese Restaurant	Ramen Restaurant	Restaurant	Seafood Restaurant	Sushi Restaurant	Taiwanese Restaurant	Thai Restaurant	Theme Restaurant	Vegetarian / Vegan Restaurant	Vietnamese Restaurant
0	Adelaide,Richmond,King	0.142857	0.285714	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	...	0.0	0.000000	0.000000	0.142857	0.142857	0.0	0.000000	0.000	0.142857	0.000000
1	Berczy Park	0.000000	0.000000	0.125000	0.000000	0.000000	0.000000	0.125000	0.000000	0.0	...	0.0	0.000000	0.125000	0.250000	0.000000	0.0	0.125000	0.000	0.000000	0.000000
2	Business Reply Mail Processing Centre 969 Eastern	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	...	0.0	0.000000	0.500000	0.000000	0.000000	0.0	0.000000	0.000	0.000000	0.000000
3	Central Bay Street	0.000000	0.000000	0.000000	0.000000	0.000000	0.100000	0.000000	0.000000	0.0	...	0.1	0.100000	0.000000	0.100000	0.100000	0.0	0.100000	0.000	0.100000	0.000000
4	Christie	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	...	0.0	0.000000	0.500000	0.000000	0.000000	0.0	0.000000	0.000	0.000000	0.000000

Figure 7: Percent Frequencies for each Neighborhood

We then use the frequency distribution for each neighborhood to select the top 7 restaurant categories as shown in the figure below. The figure shows the top 5 distributions for each neighborhood as an example.

```
----Adelaide,Richmond,King----
                venue  freq
0     Asian Restaurant  0.29
1     Sushi Restaurant  0.14
2     Greek Restaurant  0.14
3   Seafood Restaurant  0.14
4  American Restaurant  0.14


----Berczy Park----
                venue  freq
0  Seafood Restaurant  0.25
1          Restaurant  0.12
2  Belgian Restaurant  0.12
3  Italian Restaurant  0.12
4     Thai Restaurant  0.12


----Business Reply Mail Processing Centre 969 Eastern----
                      venue  freq
0       Fast Food Restaurant   0.5
1                 Restaurant   0.5
2        American Restaurant   0.0
3      Portuguese Restaurant   0.0
4  Latin American Restaurant   0.0


----Central Bay Street----
                        venue  freq
0             Italian Restaurant   0.2
1    Modern European Restaurant   0.1
2  Vegetarian / Vegan Restaurant   0.1
3              Thai Restaurant   0.1
4            Chinese Restaurant   0.1


----Christie----
                      venue  freq
0         Italian Restaurant   0.5
1                 Restaurant   0.5
2      Portuguese Restaurant   0.0
3          Korean Restaurant   0.0
4  Latin American Restaurant   0.0
```

Figure 8: Frequency Distribution of Neighborhoods based on Restaurant Categories

Finally, we create a data frame that has top 7 common restaurant categories for each of the 34 different neighborhood as shown in the figure below.

	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
0	Adelaide,Richmond,King	Asian Restaurant	American Restaurant	Sushi Restaurant	Greek Restaurant	Seafood Restaurant	Vegetarian / Vegan Restaurant	Cajun / Creole Restaurant
1	Berczy Park	Seafood Restaurant	Italian Restaurant	Comfort Food Restaurant	Restaurant	French Restaurant	Belgian Restaurant	Thai Restaurant
2	Business Reply Mail Processing Centre 969 Eastern	Fast Food Restaurant	Restaurant	Vietnamese Restaurant	Dim Sum Restaurant	French Restaurant	Falafel Restaurant	Ethiopian Restaurant
3	Central Bay Street	Italian Restaurant	Thai Restaurant	Vegetarian / Vegan Restaurant	Sushi Restaurant	Seafood Restaurant	Ramen Restaurant	Portuguese Restaurant
4	Christie	Italian Restaurant	Restaurant	Dim Sum Restaurant	French Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant
5	Church and Wellesley	Vietnamese Restaurant	Theme Restaurant	Indian Restaurant	Japanese Restaurant	Ethiopian Restaurant	Restaurant	Ramen Restaurant

Figure 9: Top 7 Restaurant Venue Categories

c) Cluster the Neighbors based on Frequency

Once we have the frequency values, we use k-means algorithm to cluster the neighborhoods based on the frequency distribution values. Clustering helps to segment data by grouping similar neighborhoods based on the frequency values. So within a cluster we have homogenous data points. Also, the clusters themselves are heterogeneous to one another as each cluster. By studying each cluster separately, we can tap hidden information in the data. In the picture below, we can see four different clusters.

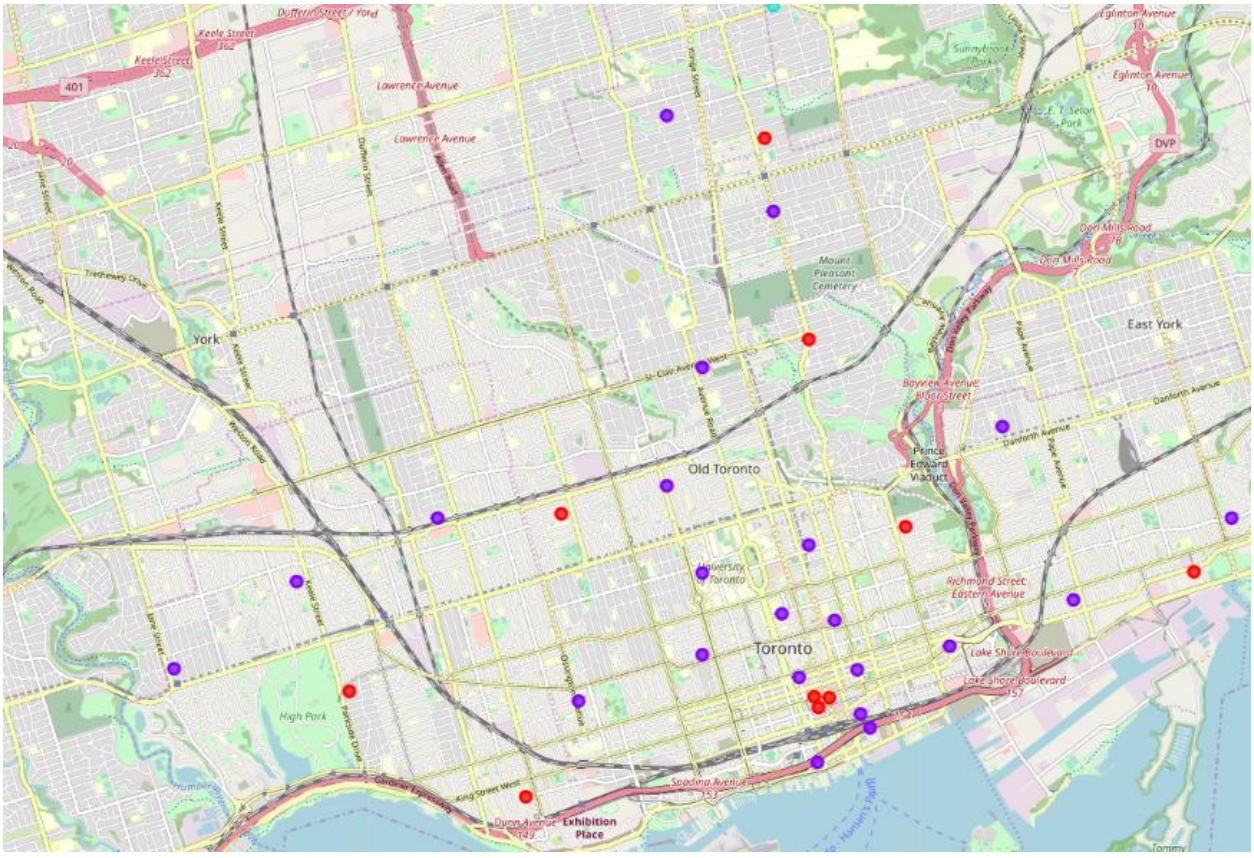

Figure 10: Neighborhoods after Clustering

d) Cluster Inference

The four figures in this sections shows the groups of neighborhoods that are separated into 4 different clusters. We can see that Cluster 3 and 4 have one neighborhood in them whereas 1 and 2 are more highly concentrated. Each of these clusters have completely different behaviors based on popularity as frequency distribution was the primary factor which was used to drive the segmentation

	Cluster Labels	Neighborhood	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	Postcode	Borough	Latitude	Longitude
2	0	Business Reply Mail Processing Centre 969 Eastern	0	Fast Food Restaurant	Restaurant	Vietnamese Restaurant	Dim Sum Restaurant	French Restaurant	Falafel Restaurant	Ethiopian Restaurant	M7Y	East Toronto	43.662744	-79.321558
4	0	Christie	0	Italian Restaurant	Restaurant	Dim Sum Restaurant	French Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	M6G	Downtown Toronto	43.669542	-79.422564
6	0	Commerce Court,Victoria Hotel	0	Restaurant	American Restaurant	Seafood Restaurant	Japanese Restaurant	New American Restaurant	Caribbean Restaurant	Cajun / Creole Restaurant	M5L	Downtown Toronto	43.648198	-79.379817
8	0	Davisville North	0	Asian Restaurant	Restaurant	Vietnamese Restaurant	Dim Sum Restaurant	French Restaurant	Fast Food Restaurant	Falafel Restaurant	M4P	Central Toronto	43.712751	-79.390197
9	0	Design Exchange,Toronto Dominion Centre	0	Restaurant	American Restaurant	Japanese Restaurant	Dim Sum Restaurant	French Restaurant	Fast Food Restaurant	Falafel Restaurant	M5K	Downtown Toronto	43.647177	-79.381576
11	0	Exhibition Place,Parkdale Village,Brockton	0	Italian Restaurant	Caribbean Restaurant	Restaurant	Dim Sum Restaurant	French Restaurant	Fast Food Restaurant	Falafel Restaurant	M6K	West Toronto	43.636847	-79.428191
20	0	Parkdale,Roncesvalles	0	Italian Restaurant	Restaurant	Eastern European Restaurant	Cuban Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	M6R	West Toronto	43.648960	-79.456325
24	0	St. James Town,Cabbagetown	0	Restaurant	Italian Restaurant	Japanese Restaurant	Caribbean Restaurant	Indian Restaurant	Taiwanese Restaurant	Thai Restaurant	M4X	Downtown Toronto	43.667967	-79.367675
27	0	Summerhill East,Moore Park	0	Restaurant	Vietnamese Restaurant	Dim Sum Restaurant	French Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	M4T	Central Toronto	43.689574	-79.383160
31	0	Underground city,First Canadian Place	0	Restaurant	American Restaurant	Japanese Restaurant	Seafood Restaurant	Greek Restaurant	Dim Sum Restaurant	Fast Food Restaurant	M5X	Downtown Toronto	43.648429	-79.382280

Figure 11: Cluster 1

	Cluster Labels	Neighborhood	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	Postcode	Borough	Latitude	Longitude
0	1	Adelaide,Richmond,King	1	Asian Restaurant	American Restaurant	Sushi Restaurant	Greek Restaurant	Seafood Restaurant	Vegetarian / Vegan Restaurant	Cajun / Creole Restaurant	M5H	Downtown Toronto	43.650571	-79.384568
1	1	Berczy Park	1	Seafood Restaurant	Italian Restaurant	Comfort Food Restaurant	Restaurant	French Restaurant	Belgian Restaurant	Thai Restaurant	M5E	Downtown Toronto	43.644771	-79.373306
3	1	Central Bay Street	1	Italian Restaurant	Thai Restaurant	Vegetarian / Vegan Restaurant	Sushi Restaurant	Seafood Restaurant	Ramen Restaurant	Portuguese Restaurant	M5G	Downtown Toronto	43.657952	-79.387383
5	1	Church and Wellesley	1	Vietnamese Restaurant	Theme Restaurant	Indian Restaurant	Japanese Restaurant	Ethiopian Restaurant	Restaurant	Ramen Restaurant	M4Y	Downtown Toronto	43.665860	-79.383160
7	1	Davisville	1	Italian Restaurant	Sushi Restaurant	Indian Restaurant	Restaurant	Seafood Restaurant	Greek Restaurant	Thai Restaurant	M4S	Central Toronto	43.704324	-79.388790
10	1	Dovercourt Village,Dufferin	1	Middle Eastern Restaurant	Vietnamese Restaurant	Indian Restaurant	French Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	M6H	West Toronto	43.669005	-79.442259
13	1	Harbourfront East,Toronto Islands,Union Station	1	Italian Restaurant	New American Restaurant	Japanese Restaurant	Chinese Restaurant	French Restaurant	Fast Food Restaurant	Falafel Restaurant	M5J	Downtown Toronto	43.640816	-79.381752
14	1	Harbourfront,Regent Park	1	Mexican Restaurant	Italian Restaurant	French Restaurant	Restaurant	Dim Sum Restaurant	Fast Food Restaurant	Falafel Restaurant	M5A	Downtown Toronto	43.654260	-79.360636
15	1	High Park,The Junction South	1	Mexican Restaurant	Italian Restaurant	Thai Restaurant	Fast Food Restaurant	Cajun / Creole Restaurant	Dumpling Restaurant	French Restaurant	M6P	West Toronto	43.661608	-79.464763
16	1	Kensington Market,Grange Park,Chinatown	1	Vietnamese Restaurant	Caribbean Restaurant	Mexican Restaurant	Dumpling Restaurant	Comfort Food Restaurant	Vegetarian / Vegan Restaurant	Belgian Restaurant	M5T	Downtown Toronto	43.653206	-79.400049
18	1	Little Portugal,Trinity	1	Vietnamese Restaurant	Asian Restaurant	New American Restaurant	Cuban Restaurant	French Restaurant	Greek Restaurant	Korean Restaurant	M6J	West Toronto	43.647927	-79.419750
19	1	North Toronto West	1	Fast Food Restaurant	Mexican Restaurant	Chinese Restaurant	Vietnamese Restaurant	Dumpling Restaurant	French Restaurant	Falafel Restaurant	M4R	Central Toronto	43.715383	-79.405678
21	1	Riverdale,The Danforth West	1	Greek Restaurant	Italian Restaurant	Restaurant	Indian Restaurant	Asian Restaurant	Belgian Restaurant	Cajun / Creole Restaurant	M4K	East Toronto	43.679557	-79.352188
22	1	Ryerson,Garden District	1	American Restaurant	Vegetarian / Vegan Restaurant	Thai Restaurant	Japanese Restaurant	Ramen Restaurant	Mexican Restaurant	Dim Sum Restaurant	M5B	Downtown Toronto	43.657162	-79.378937
23	1	St. James Town	1	Italian Restaurant	Restaurant	Japanese Restaurant	Middle Eastern Restaurant	New American Restaurant	American Restaurant	Thai Restaurant	M5C	Downtown Toronto	43.651494	-79.375418
25	1	Stn A PO Boxes 25 The Esplanade	1	Seafood Restaurant	Restaurant	Italian Restaurant	Comfort Food Restaurant	French Restaurant	Belgian Restaurant	Thai Restaurant	M5W	Downtown Toronto	43.646435	-79.374846
26	1	Studio District	1	Italian Restaurant	American Restaurant	Comfort Food Restaurant	Seafood Restaurant	Chinese Restaurant	New American Restaurant	Middle Eastern Restaurant	M4M	East Toronto	43.659526	-79.340923
28	1	Summerhill West,Forest Hill SE,Deer Park,South...	1	Vietnamese Restaurant	Sushi Restaurant	American Restaurant	Thai Restaurant	Greek Restaurant	Asian Restaurant	Belgian Restaurant	M4V	Central Toronto	43.686412	-79.400049
29	1	Swansea,Runnymede	1	Italian Restaurant	Sushi Restaurant	French Restaurant	Falafel Restaurant	Latin American Restaurant	Vegetarian / Vegan Restaurant	Caribbean Restaurant	M6S	West Toronto	43.651571	-79.484450
30	1	The Beaches West,India Bazaar	1	Italian Restaurant	Fast Food Restaurant	Sushi Restaurant	Dim Sum Restaurant	French Restaurant	Falafel Restaurant	Ethiopian Restaurant	M4L	East Toronto	43.668999	-79.315572
32	1	University of Toronto,Harbord	1	Japanese Restaurant	Restaurant	Italian Restaurant	Sushi Restaurant	Chinese Restaurant	Comfort Food Restaurant	French Restaurant	M5S	Downtown Toronto	43.662696	-79.400049
33	1	Yorkville,The Annex,North Midtown	1	American Restaurant	Indian Restaurant	Jewish Restaurant	Vegetarian / Vegan Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	M5R	Central Toronto	43.672710	-79.405678

Figure 12: Cluster 2

	Cluster Labels	Neighborhood	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	Postcode	Borough	Latitude	Longitude
17	2	Lawrence Park	2	Dim Sum Restaurant	Vegetarian / Vegan Restaurant	Greek Restaurant	French Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	M4N	Central Toronto	43.72802	-79.38879

Figure 13: Cluster 3

	Cluster Labels	Neighborhood	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	Postcode	Borough	Latitude	Longitude
12	3	Forest Hill West,Forest Hill North	3	Sushi Restaurant	Vietnamese Restaurant	Dim Sum Restaurant	French Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	M5P	Central Toronto	43.696948	-79.411307

Figure 14: Cluster 4

Conclusion

Based on the drilldown of each of the clusters, we can see that optimal choice for opening an Italian Restaurant should be a neighborhood from Cluster 1 or 2. This is because cluster 3 and 4 have only one neighborhood and you can see that both of them have no popularity for Italian cuisine as it is not in top 7 common venues. So, it is high risk bet but also has high rewards because there is no competition for Italian in Clusters 3 and 4. Now let us profile Cluster 1. Cluster 1 has around 4 Italian Restaurants in the

top 3 venues. So, the Italian Favorability Factor (IFF) is 4/10 = 40%. On the other hand, Cluster 2 has 13 Italian Restaurants in the top 3 venues. So, the Italian Favorability Factor is 13/22 = 59%. Hence, Cluster 2 is the obvious choice as it has strong propensity for Italian Cuisine. Now we have to just narrow down to 22-13 = 9 locations as the optimal choice for Italian Restaurants.

To conclude, if we have pick a candidate from this 9 locations we have to deep dive on each of the individual neighborhoods in detail. We have bring in external demographics data to understand the potential customer behavior, population details and also how many have Italian ethnicity. So, for the purpose of this use case, anyone of the 9 neighborhoods is an ideal candidate for the next best location. By leveraging data and machine learning, we were able to bring down our choices from 38 possible choices to just 9 neighborhoods which showcases the power of analytics.