Finding the optimal spot for our authentic Italian Restaurant

The Battle of Neighbourhoods
Sebastiaan Vrij

Table of Contents

Introduction

In this is a project for IBM Data Science Professional Certificate. In this project, I am creating a hypothetical scenario for a concept that there may not be enough Italian Restaurants in Toronto Area. With the purpose in mind, finding the location to open such a restaurant is one of the most important decisions for this entrepreneur and I am designing this project to help him find the most suitable location.

Business Problem

In this project we will try to find an optimal location for a restaurant. Specifically, this report will be targeted to stakeholders interested in opening an Italian restaurant in Toronto, Canada.
Since there are lots of restaurants in Toronto we will try to detect **locations that are not already crowded with restaurants**. We are also particularly interested in **areas with no or almost Italian restaurants in vicinity**.
We will use our data science powers to generate a few most promising neighbourhoods based on these criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

Data

- We want to map the area with information from Wikipedia + Geospatial_Coordinates.csv
- We want to map the amount of restaurants in Toronto, this data we will get from Foursquare API

Methodology

First, I need to get the list of neighbourhoods in Toronto, Canada. This is possible by extracting the list of neighbourhoods from Wikipedia: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M I did the web scraping by utilizing pandas HTML table scraping method as it is easier and more convenient to pull tabular data directly from a webpage into the data frame.

However, it is only a list of neighbourhoods names and postal codes. I need to get their coordinates to utilize Foursquare to pull the list of venues near these neighbourhoods. To get the coordinates, I used Geocoder Package(Geospatial_Coordinates.csv).

Match the coordinates of Toronto neighbourhoods. After gathering these coordinates, I visualize the map of Toronto using Folium package to verify whether these are correct coordinates. Next, I use Foursquare API to pull the list of top 100 venues within 500 meters radius. I am pulled the names, categories, latitude, and longitude of the venues. With this data, I can also check how many unique categories that I can get from these venues.

Then, I analyse each neighbourhoods by grouping the rows by neighbourhoods and taking the mean on the frequency of occurrence of each venue category. This is to prepare clustering to be done later.

Here, I made a justification to specifically look for "Italian restaurants". Lastly, I performed the clustering method by using k-means clustering. K-means clustering algorithm identifies k number of centroids, and then allocates every data point to the nearest cluster while keeping the centroids as small as possible. It is one of the simplest and popular unsupervised machine learning algorithms and it is highly suited for this project as well. I have clustered the neighbourhoods in Toronto into 3 clusters based on their frequency of occurrence for "Italian food". Based on the results (the concentration of clusters), I will be able to recommend the ideal location to open the restaurant.

Results

Clusters
With the K-mean we deterimed 3 clusters to see how many restaurants are in each cluster.
Below you can find the results of the amount of restaurants in each cluster.

Segment 0

`Out[220]:`

	Neighborhood	Italian Restaurant	Cluster Labels	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
13	Church and Wellesley	0.013699	0	43.665860	-79.38316	Dal Moro's Fresh Pasta To Go	43.666641	-79.38540	Italian Restaurant
34	First Canadian Place, Underground city	0.010000	0	43.648429	-79.38228	Mercatto	43.650243	-79.38082	Italian Restaurant

Segment 1

`Out[221]:`

	Neighborhood	Italian Restaurant	Cluster Labels	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
76	St. James Town	0.024390	1	43.651494	-79.375418	Terroni	43.650927	-79.375602	Italian Restaurant
77	St. James Town, Cabbagetown	0.043478	1	43.667967	-79.367675	Hey Lucy	43.664075	-79.368655	Italian Restaurant
76	St. James Town	0.024390	1	43.651494	-79.375418	Mercatto	43.650243	-79.380820	Italian Restaurant
11	Central Bay Street	0.048387	1	43.657952	-79.387383	Tosto	43.661198	-79.386414	Italian Restaurant
63	Ontario Provincial Government	0.033333	1	43.662301	-79.389494	Mercatto	43.660391	-79.387664	Italian Restaurant
77	St. James Town, Cabbagetown	0.043478	1	43.667967	-79.367675	F'Amelia	43.667536	-79.368613	Italian Restaurant
79	Studio District	0.027778	1	43.659526	-79.340923	Baldini	43.661300	-79.339027	Italian Restaurant
31	Enclave of M5E	0.030612	1	43.646435	-79.374846	Carisma	43.649617	-79.375434	Italian Restaurant
22	Don Mills South	0.047619	1	43.725900	-79.340923	Sorento Restaurant	43.726575	-79.341989	Italian Restaurant
31	Enclave of M5E	0.030612	1	43.646435	-79.374846	Mirto	43.650702	-79.376307	Italian Restaurant
31	Enclave of M5E	0.030612	1	43.646435	-79.374846	Metropolitan Resto Bar	43.650062	-79.377181	Italian Restaurant
17	Commerce Court, Victoria Hotel	0.040000	1	43.648198	-79.379817	Mercatto	43.650243	-79.380820	Italian Restaurant
17	Commerce Court, Victoria Hotel	0.040000	1	43.648198	-79.379817	Terroni	43.650927	-79.375602	Italian Restaurant
17	Commerce Court, Victoria Hotel	0.040000	1	43.648198	-79.379817	Metropolitan Resto Bar	43.650062	-79.377181	Italian Restaurant
17	Commerce Court, Victoria Hotel	0.040000	1	43.648198	-79.379817	Carisma	43.649617	-79.375434	Italian Restaurant
11	Central Bay Street	0.048387	1	43.657952	-79.387383	Scaddabush Italian Kitchen & Bar	43.658920	-79.382891	Italian Restaurant
7	Brockton, Parkdale Village, Exhibition Place	0.040000	1	43.636847	-79.428191	Caffino	43.639021	-79.425289	Italian Restaurant
53	Little Portugal, Trinity	0.022727	1	43.647927	-79.419750	Ufficio	43.649439	-79.423014	Italian Restaurant
46	India Bazaar, The Beaches West	0.043478	1	43.668999	-79.315572	Casa di Giorgio	43.666645	-79.315204	Italian Restaurant

4

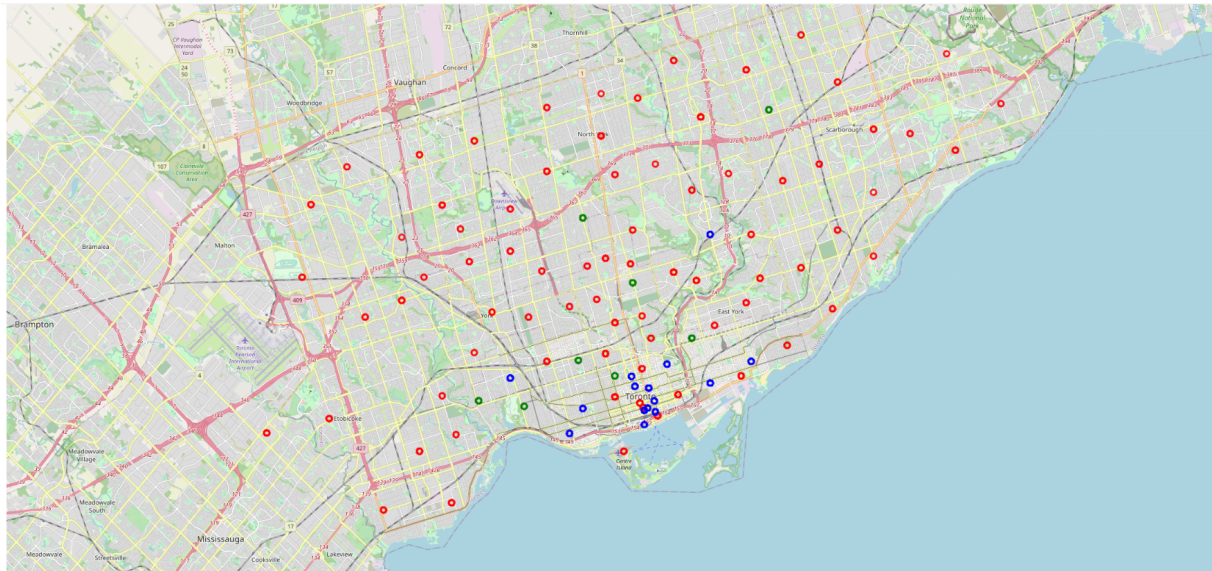|	Neighborhood	Italian Restaurant	Cluster Labels	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
40	Harbourfront East, Union Station, Toronto Islands	0.030000	1	43.640816	-79.381752	Taverna Mercatto	43.642625	-79.383257	Italian Restaurant
40	Harbourfront East, Union Station, Toronto Islands	0.030000	1	43.640816	-79.381752	Piazza Manna	43.641460	-79.377719	Italian Restaurant
40	Harbourfront East, Union Station, Toronto Islands	0.030000	1	43.640816	-79.381752	Amano Pasta	43.644919	-79.382358	Italian Restaurant
87	Toronto Dominion Centre, Design Exchange	0.030000	1	43.647177	-79.381576	Mercatto	43.650243	-79.380820	Italian Restaurant
87	Toronto Dominion Centre, Design Exchange	0.030000	1	43.647177	-79.381576	Metropolitan Resto Bar	43.650062	-79.377181	Italian Restaurant
87	Toronto Dominion Centre, Design Exchange	0.030000	1	43.647177	-79.381576	TOCA	43.645431	-79.387059	Italian Restaurant
11	Central Bay Street	0.048387	1	43.657952	-79.387383	Mercatto	43.660391	-79.387664	Italian Restaurant
41	High Park, The Junction South	0.040000	1	43.661608	-79.464763	nodo	43.665303	-79.465621	Italian Restaurant
36	Garden District, Ryerson	0.020000	1	43.657162	-79.378937	Scaddabush Italian Kitchen & Bar	43.658920	-79.382891	Italian Restaurant
36	Garden District, Ryerson	0.020000	1	43.657162	-79.378937	Trattoria Mercatto	43.654453	-79.380974	Italian Restaurant

Segment 2

Out[222]:

	Neighborhood	Italian Restaurant	Cluster Labels	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
84	The Danforth West, Riverdale	0.069767	2	43.679557	-79.352188	Cafe Fiorentina	43.677743	-79.350115	Italian Restaurant
84	The Danforth West, Riverdale	0.069767	2	43.679557	-79.352188	7 Numbers	43.677062	-79.353934	Italian Restaurant
18	Davisville	0.054054	2	43.704324	-79.388790	Positano	43.704558	-79.388639	Italian Restaurant
72	Runnymede, Swansea	0.054054	2	43.651571	-79.484450	Goodfellas Wood Oven Pizza	43.648224	-79.486356	Italian Restaurant
18	Davisville	0.054054	2	43.704324	-79.388790	Florentia Ristorante	43.703594	-79.387985	Italian Restaurant
84	The Danforth West, Riverdale	0.069767	2	43.679557	-79.352188	IL FORNELLO on Danforth	43.678604	-79.346904	Italian Restaurant
72	Runnymede, Swansea	0.054054	2	43.651571	-79.484450	Campo	43.655191	-79.487067	Italian Restaurant
88	University of Toronto, Harbord	0.062500	2	43.662696	-79.400049	Piano Piano	43.662949	-79.402898	Italian Restaurant
12	Christie	0.062500	2	43.669542	-79.422564	Vinny's Panini	43.670679	-79.426148	Italian Restaurant
64	Parkdale, Roncesvalles	0.066667	2	43.648960	-79.456325	Domani Restaurant & Wine Bar	43.649235	-79.450229	Italian Restaurant
4	Bedford Park, Lawrence Manor East	0.086957	2	43.733283	-79.419750	Francobollo	43.734557	-79.419549	Italian Restaurant
4	Bedford Park, Lawrence Manor East	0.086957	2	43.733283	-79.419750	Il Fornaro	43.734073	-79.419870	Italian Restaurant
88	University of Toronto, Harbord	0.062500	2	43.662696	-79.400049	Daddyo's	43.664622	-79.402685	Italian Restaurant
15	Clarks Corners, Tam O'Shanter, Sullivan	0.071429	2	43.781638	-79.304302	Remezzo Italian Bistro	43.778649	-79.308264	Italian Restaurant

Below you can find a overview summary of the amount of Italian restaurants in each area.

Segment 0	Segment 1	Segment 2
2	29	14

We can see that segment 0 has the least amount of restaurants. Therefore we want to be in this area.

- Cluster 0: Neighbourhoods with a few Italian restaurants.
- Cluster 1: Neighbourhoods with the most Italian restaurants.
- Cluster 2: Neighbourhoods with some Italian restaurants

In the picture above you can see on overview of the segments visually displayed. We want to be in the green area.

Recommendations

Most of the Italian restaurants are in cluster 1 lowest in Cluster 0.
Looking at nearby venues it seems cluster 1 might be a good location as there are not a lot of Italian restaurants in these areas. We therefore recommend the entrepreneur to open an authentic Italian restaurant in these locations.