Best location for a Portuguese Restaurant in Toronto

Ricardo Pereira

24 of August, 2021

1. Introduction: Business Problem

1.1. Background

With the globalization of food, people have tended to want to try new cuisines. Portuguese cuisine, despite influencing other cuisines, due to colonization, is little known.

As such, my client wants to open a typical Portuguese food restaurant in Toronto, ON.

The objective of this project is to find the best localization to open a Portuguese restaurant in Toronto, ON.

1.2. Problem

In order for the business to be successful, i need to find the best localization. For that purpose i will gather data about other restaurant localization, competitors (Portuguese restaurant) and top venues.

1.3. Interest
This study is interesting to anyone interested in open a business, because most of the times the location of business can dictate the success or unsuccess of the business.

2. Data

Based on the definition of our problem, factors that may impact our decision are:

. Demographic information, e.g. population, density, education, age, income. Number of existing restaurants in the neighborhood and nearby. Number of Portuguese restaurants in the neighborhood and nearby.

2.1. Data Sources

Data used in the analysis are listed below:

· Neighborhoods in Toronto -- 'https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'.

. Population in Toronto -- 'https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/pd-pl/Tables/CompFile.cfm?Lang=Eng&T=1201&OFT=FULLCSV'

· Using Geopy to get geological location by address name.

. Using geospatial data of Borough of Toronto --
'https://raw.githubusercontent.com/BlizzWiz/IBM_Capstone_Project/master/toronto_m.geojson'

· Using Foursquare API to get the most common venues of given Borough of Toronto.

· Using Foursquare API to get the venues' record of given venues of Toronto.
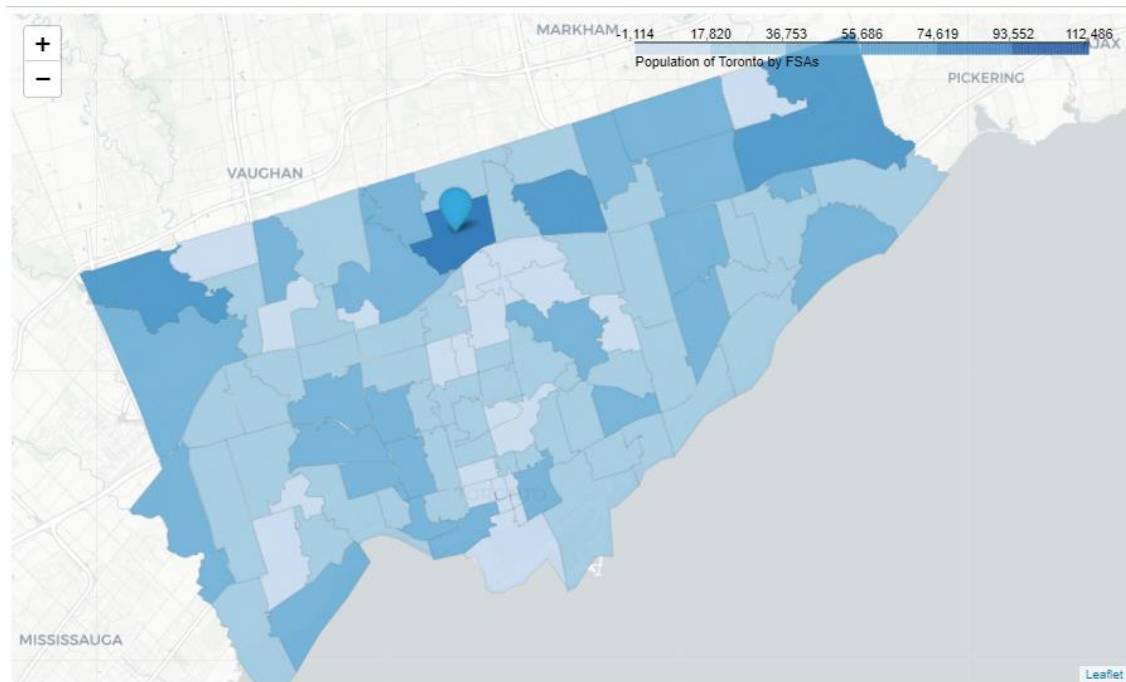
2.2. Data Cleaning

First we used Beautifulsoap 4 as a web scraping package to get the Postal Codes, boroughs and neighborhoods from Wikipedia and organize a table. Then i used the GeoSpatial Dataset with the coordinates of each postal code provided in week 3 and merge a table with postal codes, boroughs neighborhoods, latitude and longitude from Toronto.

	Postal Code	Borough	Neighborhood	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	Queen's Park	Ontario Provincial Government	43.662301	-79.389494

Then we gathered the population data of Toronto, latest as of 2016, by Forward Sortation Area (FSA) from the Statistics Canada and created a single table. For that purpose we extracted the useful columns, such as geographic code, province and population data and then, merge the two tables.

	Postal Code	Borough	Neighborhood	Latitude	Longitude	Geographic code	Geographic name	Province or territory	Population, 2016
0	M3A	North York	Parkwoods	43.753259	-79.329656	M3A	M3A	Ontario	34615.0
1	M4A	North York	Victoria Village	43.725882	-79.315572	M4A	M4A	Ontario	14443.0
2	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763	M6A	M6A	Ontario	21048.0
3	M3B	North York	Don Mills North	43.745906	-79.352188	M3B	M3B	Ontario	13324.0
4	M6B	North York	Glencairn	43.709577	-79.445073	M6B	M6B	Ontario	28522.0

To help with the data visualizion we created a cloroplethmap of Toronto population density.

For last we used the foursquare API to get all the top venues, restaurants and Portuguese Restaurants.

3. Methodology

3.1. Explanatory data analysis

After we realized North York was the densest Borough, we decide to explore (using the Foursquare API) the top venues, restaurants and Portuguese restaurants near de centre area and confirm if it would be a good localization.
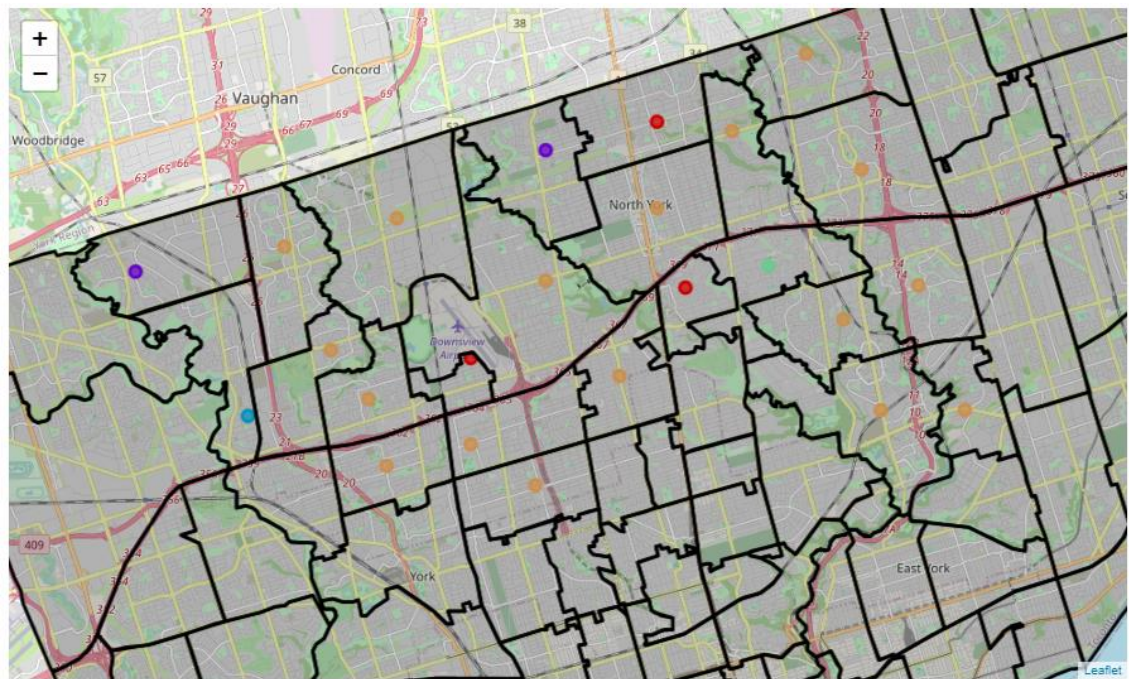
3.2. Clustering Approach

In order to confirm if the centre of North York (Willowdale) was a good localization, we decide to explore the venues, restaurants and Portuguese restaurants, and then segment and group similar neighborhoods. To be able to do that, we used K-Means Clustering approach (unsupervised machine learning).

Cluster Neighborhood

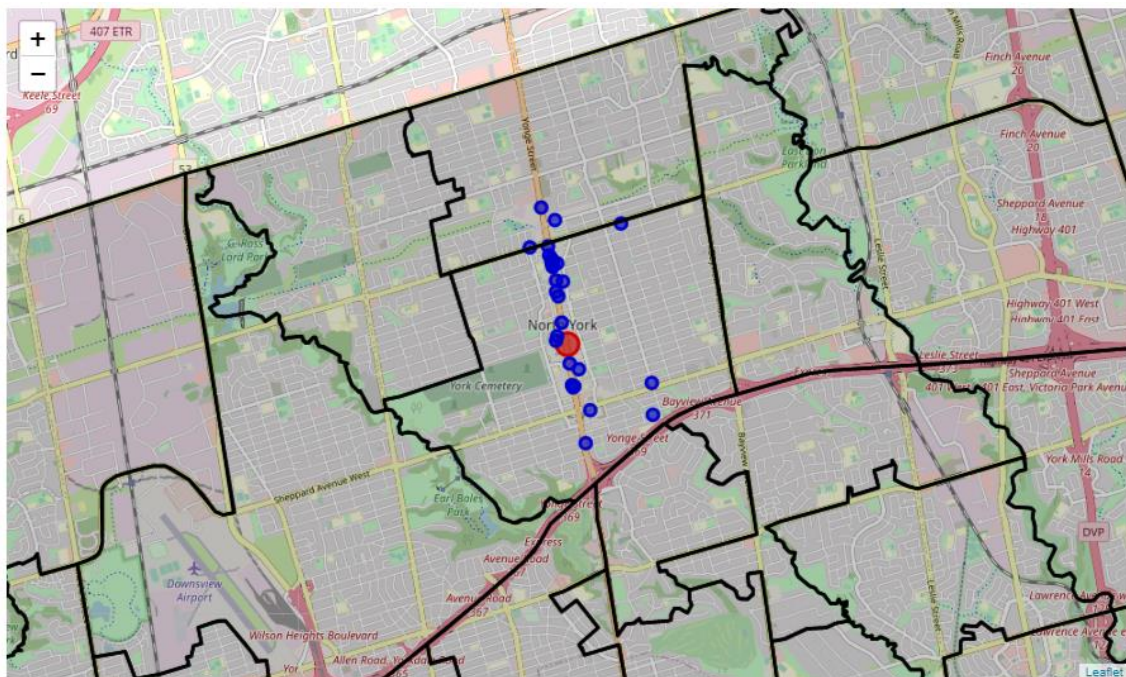|	index	Postal Code	Borough	Neighborhood	Latitude	Longitude	Geographic code	Geographic name	Province or territory	Population, 2016	Cluster Labels	1st Most Common Venue	2nd Most Common Venue
0	20	M2N	North York	Willowdale South	43.770120	-79.408493	M2N	M2N	Ontario	75897.0	4	Ramen Restaurant	Pizza Place
1	8	M2J	North York	Fairview, Henry Farm, Oriole	43.778517	-79.346556	M2J	M2J	Ontario	58293.0	4	Clothing Store	Coffee Shop
2	21	M3N	North York	Downsview Northwest	43.761631	-79.520999	M3N	M3N	Ontario	41958.0	4	Liquor Store	Grocery Store
3	23	M2R	North York	Willowdale West	43.782736	-79.442259	M2R	M2R	Ontario	40792.0	1	Coffee Shop	Supermarket
4	5	M3C	North York	Don Mills South	43.725900	-79.340923	M3C	M3C	Ontario	39153.0	4	Coffee Shop	Restaurant

Map of Clusters neighborhood

We used the Foursquare API again to explore Restaurants in radius of 1500 meters around North York centre.

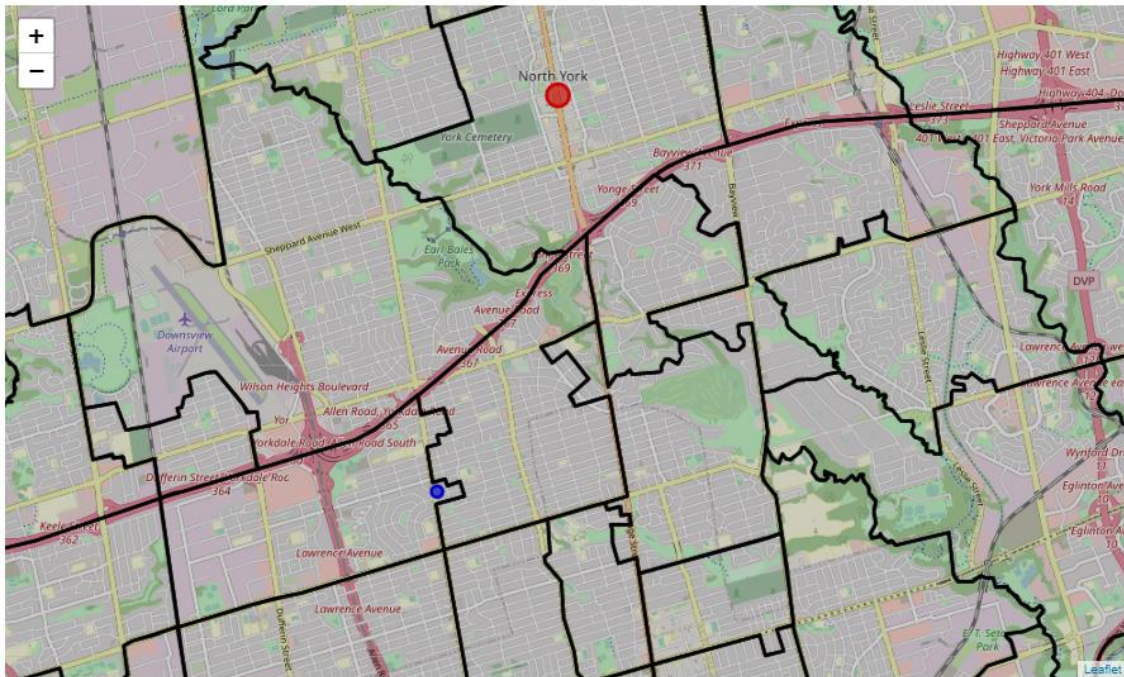|	name	categories	address	crossStreet	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	countr
0	Symposium Cafe Restaurant & Lounge	Restaurant	5221 Yonge St	Parkview	43.771075	-79.413396	[{'label': 'display', 'lat': 43.77107544003332...	273	M2N 5P8	CA	North York	ON	Canad
1	Cottage Restaurant	Pizza Place	15 Northtown Way	NaN	43.775572	-79.413234	[{'label': 'display', 'lat': 43.77557162610398...	768	NaN	CA	North York	ON	Canad
2	Lee Town Restaurant	Chinese Restaurant	4907 Yonge St.	at Spring Garden Ave.	43.764195	-79.411626	[{'label': 'display', 'lat': 43.76419503107473...	505	M2N 5N4	CA	Toronto	ON	Canad
3	Shanghai Chinese Restaurant 상해반점	Chinese Restaurant	5451 Yonge St.	NaN	43.777283	-79.414769	[{'label': 'display', 'lat': 43.7772828010856,...	973	M2N 5S1	CA	North York	ON	Canad
4	Restaurant At Novotel North York Toronto	Asian Restaurant	NaN	NaN	43.769609	-79.414167	[{'label': 'display', 'lat': 43.769609, 'lng':...	161	NaN	CA	Toronto	ON	Canad
5	Trio Restaurant	Restaurant	3 Park Home Avenue	btw Finch Ave E and Sheppard Ave W	43.769058	-79.414377	[{'label': 'display', 'lat': 43.76905769173501...	146	M2N 6L3	CA	North York	ON	Canad
6	Koryo's Korean Restaurant & Bar	Korean Restaurant	NaN	NaN	43.764317	-79.411910	[{'label': 'display', 'lat': 43.764317, 'lng':...	488	M2N 5N5	CA	Toronto	ON	Canad
7	Leetown Chinese Restaurant	Chinese Restaurant	4907 Yonge Street	Spring Garden	43.764054	-79.411733	[{'label': 'display', 'lat': 43.76405354842137...	519	NaN	CA	Toronto	ON	Canad
8	Chattanooga Restaurant	Diner	NaN	NaN	43.773981	-79.413992	[{'label': 'display', 'lat': 43.77398070199035...	600	NaN	CA	NaN	NaN	Canad
9	Bei Jing Restaurant	Chinese Restaurant	NaN	NaN	43.775650	-79.414250	[{'label': 'display', 'lat': 43.77565, 'lng': ...	786	M2N 5R6	CA	Toronto	ON	Canad

Map of Restaurants:

In radius of 1500 meters the only cuisine more similar to Portuguese cuisine is a Pizza place (Italian).

In order to be sure the direct competition was far from our localization we used the Foursquare API to explore Portuguese Restaurants in radius of 5000 meters around North York centre.

	name	categories	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	
0	Portuguese Chicken Place	Portuguese Restaurant	43.72553	-79.43089	[{'label': 'display', 'lat': 43.72553, 'lng': ...	5022	M6A 2B7	CA	Toronto	ON	Canada	[Toronto ON M6A 2B7, Canada]	5d0d6e60672:

Map of Portuguese Restaurants:

4. Results

The objective was to help our client to identify the best Borough and Neighborhood in Toronto to open a Portuguese restaurant. To do that, we have in consideration the population (https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/pd-pl/Tables/CompFile.cfm?Lang=Eng&T=1201&OFT=FULLCSV) and visualized the information with a choropleth map. Through the analysis we realize that the most populated area in Toronto is North York, therefore, the best candidate localization. After selecting the borough/Area, it was imperative to choose the right neighborhood, where Portuguese restaurants and similar cuisine (as Italian and Mediterranean) were not among venues in a close proximity, the best neighborhood was Willowdale South,which also surrounds the North York Centre.

In radius of 15000 meters around North York centre we realize theres is a lot of Asian offers and only on Pizza place (Italian cuisine) wath for our purpose is good.

The nearest Portuguese Resaurant is about 5000 meters from centre.

We achieved this by grouping the neighborhoods into clusters assist the client by providing them with relevant data about venues, restaurants and Portuguese restaurants of a given neighborhood (Willowdale).

5. Conclusion

We find that there is a lot of Asian food near the North York Center (Willowdale) and nearest Portuguese restaurant is about 5000 meters making a good candidate location, based on population density and Portuguese/ Mediterranian Cuisine competition.

However this is a preliminary and small study. And in order for a more accuracy decision, we should have to take in account more extra data, as rental prices, parking and public transport.