Predicting the Optimal Location for a Syrian Restaurant in Downtown Toronto

Purvasha Patnaik

1. Introduction

1.1 Background

Due to the crisis in Syria, there has been a great influx of Syrian refugees all of over Canada, especially in Toronto. For competition and marketing purposes, it is important to keep the location in a hotspot of restaurants, yet away from other Syrian restaurants. It would also be essential for the location to be as close to downtown Toronto as much as possible, as that is the area with the most local and tourist popularity.

1.2 Problem

The business problem I would like to address is the lack of Syrian restaurants in Toronto. This report will demonstrate to stakeholders where the best location would be to open a Syrian restaurant to increase the market for Syrian food as well as make the refugees feel more at home.

2. Data Acquisition and Cleaning

2.1 Data Sources

There are three factors that will influence our decision of optimal location for the Syrian restaurant:

1. Number of existing restaurants in the vicinity
2. Number of existing Syrian restaurants in the vicinity
3. Distance of optimal location to the downtown Toronto area

Therefore, we will be using the following data to address these factors:

1. Toronto postal code data scraped from Wikipedia page to make a data frame of boroughs, neighbourhoods, latitude, and longitude
2. Number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API
3. Coordinates of downtown Toronto will be obtained using Google latitude and longitude of the CN tower, which is a well-known tourist attraction located in downtown Toronto

2.2 Data Cleaning and Feature Selection

Firstly, the data of postal codes, boroughs, and neighbourhoods were scraped from a Wikipedia page (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada) with the information. From this, a table was made where all of the rows that were not assigned to boroughs were dropped. All of the neighbourhoods were grouped by their assigned boroughs and separated by commas. Secondly, a csv file containing the geographical coordinates of each postal code was opened. This data was combined with the previous table to make a complete data frame of the post codes, neighbourhoods, boroughs, and their latitudes and longitudes.

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M1B	Scarborough	Rouge, Malvern	43.806686	-79.194353
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476

Table 1. Preview of pandas data frame showing the postal code, borough, neighbourhood, latitude, and longitudes in Toronto.

3. Exploratory Analysis

3.1 Relationship between venues and hotspots

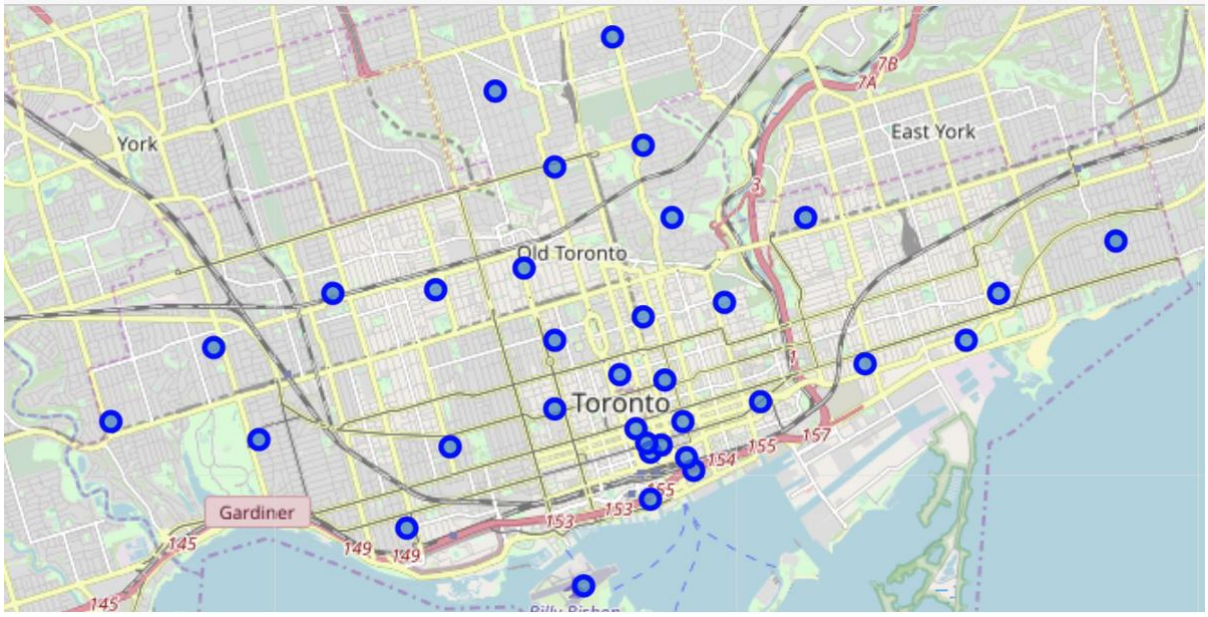

Figure 1. Map created by Folium of neighbourhoods in Toronto.

Using latitude and longitude of Toronto from Google, I was able to create a map on Folium to visualize the neighbourhoods in Toronto. Using a function that required the Foursquare API URL and GET request, a pandas data frame was made containing all the venue names and categories with their latitudes and longitudes.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	The Beaches	43.676357	-79.293031	Glen Manor Ravine	43.676821	-79.293942	Trail
1	The Beaches	43.676357	-79.293031	The Big Carrot Natural Food Market	43.678879	-79.297734	Health Food Store
2	The Beaches	43.676357	-79.293031	Grover Pub and Grub	43.679181	-79.297215	Pub
3	The Beaches	43.676357	-79.293031	Glen Stewart Ravine	43.676300	-79.294784	Other Great Outdoors
4	The Beaches	43.676357	-79.293031	Upper Beaches	43.680563	-79.292869	Neighborhood

Table 2. Preview of pandas data frame with neighbourhood and venue information.

Each neighbourhood was then analyzed and organized so that each neighbourhood was listed with its top 5 venues.

```
----Adelaide, King, Richmond----
                venue  freq
0         Coffee Shop  0.08
1                Café  0.05
2                 Bar  0.04
3  Thai Restaurant  0.04
4          Steakhouse  0.04


----Berczy Park----
                venue  freq
0         Coffee Shop  0.09
1        Cocktail Bar  0.05
2  Farmers Market  0.04
3           Beer Bar  0.04
4          Steakhouse  0.04
```

Figure 2. Preview of neighbourhoods with their top 5 venues.

The neighbourhoods were clustered and analyzed, after which a pandas data frame was made to contain each neighbourhood and their top 10 venues. This will allow us to sort for the hotspots depending on which neighbourhoods have 10 or more venues that are restaurants.

	Postcode	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	9th Most Common Venue	10th Most Common Venue
0	M4E	East Toronto	The Beaches	43.676357	-79.293031	0	Health Food Store	Other Great Outdoors	Trail	Pub	Dive Bar	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant
1	M4K	East Toronto	The Danforth West, Riverdale	43.679557	-79.352188	0	Greek Restaurant	Coffee Shop	Ice Cream Shop	Italian Restaurant	Furniture / Home Store	Pizza Place	Bookstore	Brewery	Bubble Tea Shop	Caribbean Restaurant
2	M4L	East Toronto	The Beaches West, India Bazaar	43.668999	-79.315572	0	Gym	Fish & Chips Shop	Sushi Restaurant	Sandwich Place	Brewery	Steakhouse	Ice Cream Shop	Pub	Movie Theater	Italian Restaurant

Table 3. Preview of neighbourhoods with top 10 venues.

The following data cleaning allowed for the isolation of the restaurant hotspots. All other venue categories that did not include food items in their venue were excluded and removed from the data frame. For example, any neighbourhoods that included any of the following: yoga studio, museums, parks, gardens, spas, bookstores, airport, shops and pharmacies, and stadiums were removed from the data set. What remained were 6 restaurant hotspot locations that all had venues all in the category of food as their top 10 venues.

	Postcode	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	9th Most Common Venue	10th Most Common Venue
7	M4S	Central Toronto	Davisville	43.704324	-79.388790	0	Pizza Place	Sandwich Place	Dessert Shop	Café	Italian Restaurant	Coffee Shop	Sushi Restaurant	Gourmet Shop	Deli / Bodega	Seafood Restaurant
11	M4X	Downtown Toronto	Cabbagetown, St. James Town	43.667967	-79.367675	0	Coffee Shop	Restaurant	Café	Flower Shop	Italian Restaurant	Bakery	Pub	Pizza Place	Pet Store	Breakfast Spot
16	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306	0	Coffee Shop	Cocktail Bar	Cheese Shop	Farmers Market	Bakery	Café	Beer Bar	Seafood Restaurant	Steakhouse	Irish Pub
26	M5T	Downtown Toronto	Chinatown, Grange Park, Kensington Market	43.653206	-79.400049	0	Café	Vegetarian / Vegan Restaurant	Chinese Restaurant	Bar	Mexican Restaurant	Coffee Shop	Dumpling Restaurant	Bakery	Vietnamese Restaurant	Dessert Shop
32	M6J	West Toronto	Little Portugal, Trinity	43.647927	-79.419750	0	Bar	Coffee Shop	Asian Restaurant	Vietnamese Restaurant	Café	Restaurant	French Restaurant	Cocktail Bar	Bakery	New American Restaurant
36	M6S	West Toronto	Runnymede, Swansea	43.651571	-79.484450	0	Coffee Shop	Pizza Place	Café	Sushi Restaurant	Italian Restaurant	Burrito Place	Food	Fish Market	Fish & Chips Shop	Smoothie Shop

Table 4. All 6 hotspot restaurant locations as sorted for their top 10 venues to be in the food venue category.

Using Folium, another map was generated to show how the restaurant hotspot locations are situated from each other. We can now use these hotspots as potential options of where to place our new Syrian restaurant.

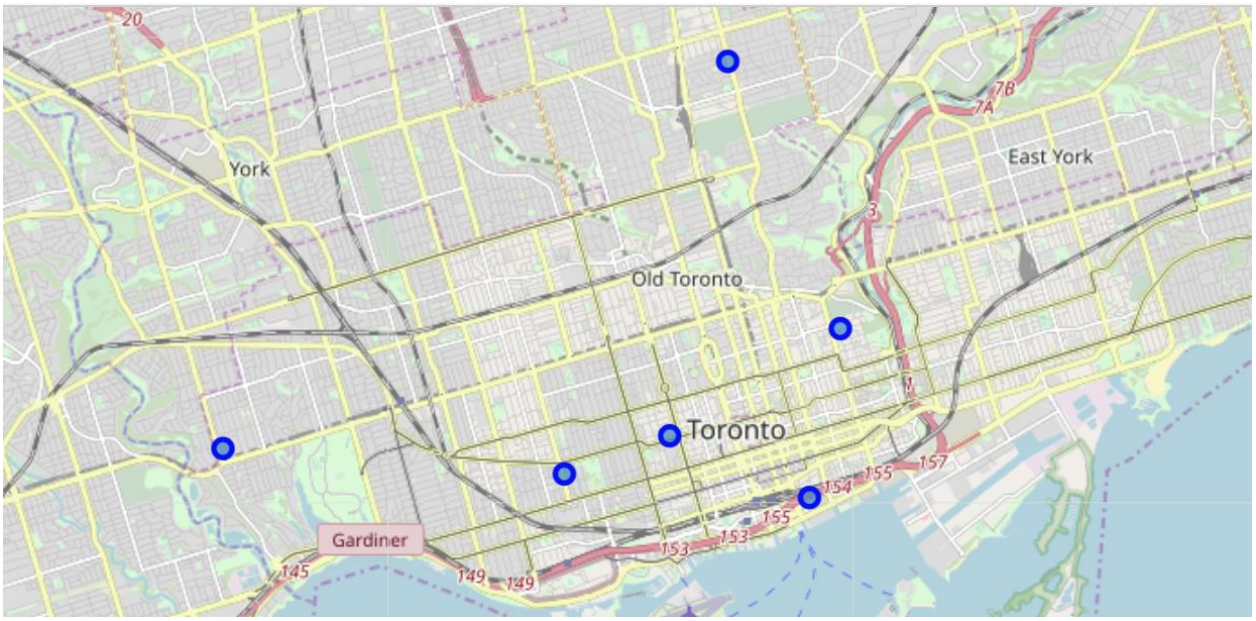

Figure 3. Map of all 6 restaurant hotspot locations.

3.2 Relationship between hotspots and Syrian restaurants

After generating 6 restaurant hotspots and mapping them on Folium, the next step is to use Foursquare API locations data using the explore URL to search for any Syrian restaurants in the vicinity. After applying the explore URL with the Syrian restaurant venue category ID to all of the restaurant hotspots, the results showed that there are not many results near any of the hotspots. The parameters included in the explore URL were to

limit the results to a 500m radius and allow for 100 results. The results concluded ‘There are not many results near you. Try something more general, reset your filters, or expand the search area.’ From this result, we can conclude that all 6 restaurant hotspots don’t have many Syrian restaurants within their vicinity of a 500m radius, and all of the hotspots can be potential locations to open our Syrian restaurant.

3.3 Relationship between hotspots and CN tower

The CN tower, or Canadian National Tower built by a railway company of the same name, is the main tourist attraction of Downtown Toronto. It is situated in the core of Downtown Toronto and is surrounded by several other tourist attractions, such as Ripley’s Aquarium and the Roger’s Center Stadium (home to many BlueJays baseball games). Therefore, the CN tower was chosen as a landmark to be in close proximity to the neighbourhood that we will situate our Syrian restaurant. This will ensure marketing success and will be essential in the growth of the restaurant as it will be in an area of high foot traffic. Using geopy, the distances were mapped out and the output gave us which coordinates are closest to the coordinates of the CN tower (which was found on Google). The code selected ‘Berczy Park’ as the neighbourhood closest to the CN tower. With all elements in place, ‘Berczy Park’ satisfies our conditions of being in a restaurant hotspot location that has 10 or more food category venues provided by Foursquare API, it does not have many Syrian restaurants in a 500m radius vicinity, and it is the closest to our chosen tourist attraction, the CN tower.

4. Conclusions

The data was gathered by assembling the neighbourhood, boroughs and postal codes in Toronto with their latitude and longitude coordinates. A pandas data frame was created by this data, and using Toronto's coordinates, a map was generated using Folium. The neighbourhoods were clustered and segmented and shown on the map. Using this data and Foursquare API, venues were explored in these neighbourhoods. Particularly, restaurants and food areas were highlighted in the pandas data frame and sorted to isolate the restaurant hotspots in Toronto. This was done by assuming that if a neighbourhood had a total of 10 restaurants in its area, then it would be considered a hotspot. This brought the data down to 6 hotspots, which brings us to the next challenge: Syrian restaurants.

Foursquare API exploration resulted in a lack of Syrian restaurants in any of the isolated restaurant hotspots. This is excellent news! We can choose any one of these areas for the Syrian restaurant. However, to optimize the location and ensure marketing success, we found the hotspot that is the closest to the CN tower. The CN tower is the main tourist attraction of Toronto: it is a 553.3 m high communications and observations tower located in the heart of Downtown Toronto. After mapping the distances, we found that Berczy Park is within the closest proximity to the CN tower, and it is the restaurant hotspot that would be the most optimal location for a Syrian restaurant.

There is the lack of Syrian restaurants in Toronto compared to others ex. Indian, Chinese, Italian etc. It is concluded at the end of this report that Berczy Park will demonstrate is the best location to open a Syrian restaurant to increase the market for Syrian food as well as make the refugees feel more at home. The location is in a hotspot of restaurants, as demonstrated by Foursquare API, yet away from other Syrian restaurants. This report also evaluated the location as within close proximity to CN tower in Downtown Toronto, attracting an abundance of local and tourist attention.

5. Future Directions

There are numerous factors that need to be considered when selecting an optimal location to start any type of business. For example, we chose areas that had at least 10 venues per neighbourhood to select that area as a restaurant hotspot. However, it is highly likely that there are many more areas that have 10+ restaurant venues in their neighbourhood in the core of downtown Toronto. One possible project would be to find which neighbourhood has the most amount of restaurant venues rather than limiting it to 10.

Another prospective research question would be to factor in the various concentrated populations of Syrian communities around Toronto. It would definitely to be beneficial if the restaurant were closer to the Syrian population of Toronto as well as be a hotspot venue and close to downtown Toronto.