Finding suitable locations for a Restaurant in Downtown Toronto, Canada

Ashwini Anand

August 22, 2021

1. Introduction:

1.1. Business Problem:

This project aims at identifying a suitable location to open a new restaurant in the city of Toronto, Canada, specifically Downtown Toronto. Through this project, I aim to ease the process of shortlisting the most suitable locations for such a venture using map data from Foursquare.

Although a lot of factors must be considered in such cases, our primary goal is to find areas with a low-to-moderate density of restaurants, which are also close to the city centre. This will be helpful to ensure good footfall in our restaurant, while also maintaining a healthy level of competition with other restaurants for maximum growth of our business.

1.2. Interest:

This project will be of great use to stakeholders interested in opening a new restaurant in the Downtown Toronto area, and especially to those looking to bounce back after the coronavirus pandemic has pushed many out of business, and to those looking to cash in on the surge of customers getting out of their homes after the pandemic scare and repeated lockdowns.

2. Data and Sources:

2.1. Data Acquisition:

First, the list of boroughs and neighbourhoods in Toronto city can be found on this Wikipedia page. I scraped the page and created a data frame of all neighbourhoods by postal codes and boroughs. However, this data frame lacks coordinates. The location data of these postal codes can be found here. The location data, in csv format, is already sorted by postal codes with coordinates to the centre of each area. Then, I merged the two data frames to create a data frame having the postal codes, borough, neighbourhoods, and the latitude and longitude data of each postal code in Toronto.

The remaining data about restaurants in each neighbourhood was collected using Foursquare API, which will be covered in detail in the methodology section.

2.2. Data Cleaning:

Some borough names were erroneous, which were replaced with the corrected names. In the end, a complete data frame was created with 103 rows. Here is a snapshot of the data frame:

	Postal Code	Borough	Neighbourhood	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
...	...	...	...	...	...
98	M8X	Etobicoke	The Kingsway, Montgomery Road, Old Mill North	43.653654	-79.506944
99	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160
100	M7Y	East Toronto Business	Enclave of M4L	43.662744	-79.321558
101	M8Y	Etobicoke	Old Mill South, King's Mill Park, Sunnylea, Hu...	43.636258	-79.498509
102	M8Z	Etobicoke	Mimico NW, The Queensway West, South of Bloor,...	43.628841	-79.520999

103 rows × 5 columns

Now, since our business problem's focus is Downtown Toronto, I created a new data frame out of the present data frame, containing postal codes and neighbourhoods from Downtown Toronto only. Here is a snapshot of the final data frame obtained:

	Postal Code	Borough	Neighbourhood	Latitude	Longitude
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
1	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
2	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
3	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306
4	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383
5	M6G	Downtown Toronto	Christie	43.669542	-79.422564
6	M5H	Downtown Toronto	Richmond, Adelaide, King	43.650571	-79.384568
7	M5J	Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	43.640816	-79.381752
8	M5K	Downtown Toronto	Toronto Dominion Centre, Design Exchange	43.647177	-79.381576
9	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817
10	M5S	Downtown Toronto	University of Toronto, Harbord	43.662696	-79.400049
11	M5T	Downtown Toronto	Kensington Market, Chinatown, Grange Park	43.653206	-79.400049
12	M5V	Downtown Toronto	CN Tower, King and Spadina, Railway Lands, Har...	43.628947	-79.394420
13	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529
14	M4X	Downtown Toronto	St. James Town, Cabbagetown	43.667967	-79.367675
15	M5X	Downtown Toronto	First Canadian Place, Underground city	43.648429	-79.382280
16	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160

3. Methodology, Analysis and Observations:

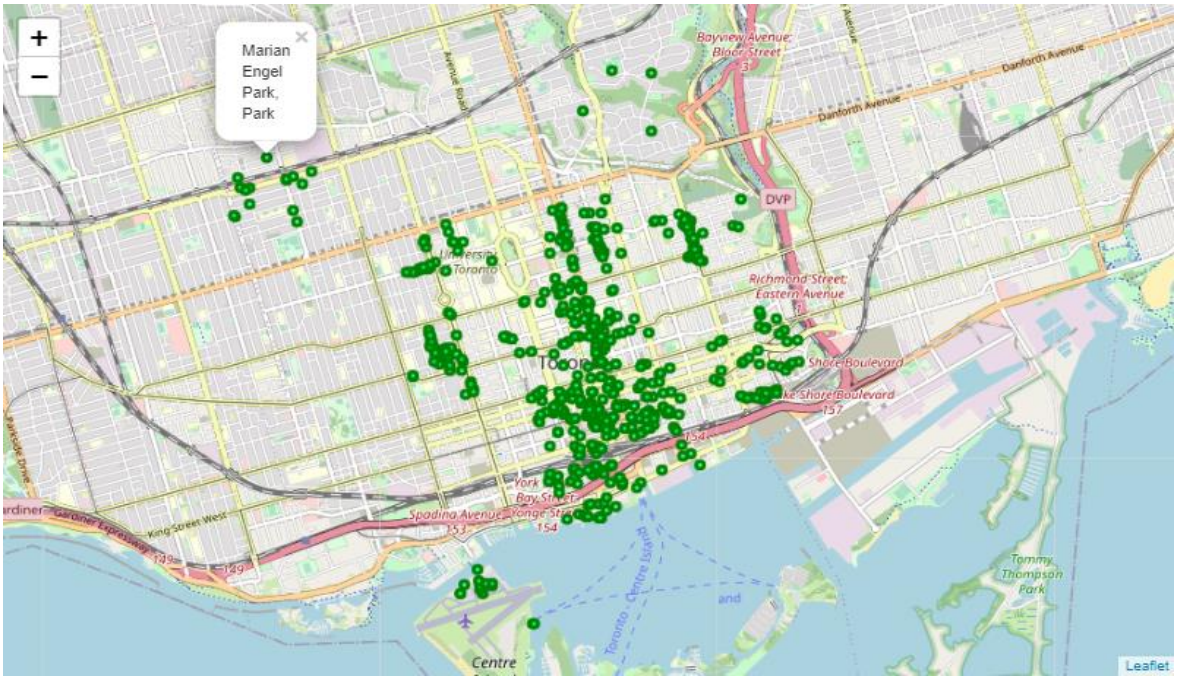
3.1. Exploring Neighbourhoods using Foursquare's Places API:

Foursquare's Places API enables users to access its massive location database. I used it to get a list of 50 venues within a radius of 500 meters of the coordinates in our data frame, by making API calls for every set of coordinates in the Downtown Toronto data frame (16 calls). The returned data was used to create another data frame containing the name of the venue, its type, and its coordinates:

	Postal Code	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	M5A	43.65426	-79.360636	Roselle Desserts	43.653447	-79.362017	Bakery
1	M5A	43.65426	-79.360636	Tandem Coffee	43.653559	-79.361809	Coffee Shop
2	M5A	43.65426	-79.360636	Cooper Koo Family YMCA	43.653249	-79.358008	Distribution Center
3	M5A	43.65426	-79.360636	Impact Kitchen	43.656369	-79.356980	Restaurant
4	M5A	43.65426	-79.360636	Body Blitz Spa East	43.654735	-79.359874	Spa
...	...	...	...	...	...	...	...
711	M4Y	43.66586	-79.383160	The Yoga Sanctuary	43.661499	-79.383636	Yoga Studio
712	M4Y	43.66586	-79.383160	Rooster Coffee House	43.669654	-79.379871	Coffee Shop
713	M4Y	43.66586	-79.383160	Wow! Sushi	43.668514	-79.386686	Sushi Restaurant
714	M4Y	43.66586	-79.383160	Coffee Island	43.664271	-79.386972	Coffee Shop
715	M4Y	43.66586	-79.383160	Cawthra Square Dog Park	43.666583	-79.380040	Dog Run

716 rows × 7 columns

Let us display these venues on a map of Downtown Toronto:



3.2. Filtering out the Restaurants:

In this step, I simply ran a loop through the data frame and searched for the rows containing the term 'restaurant' in the venue category column and created a new data frame containing the data satisfying the condition. This gives the following new data frame:

	Latitude	Longitude	Postal Code	Restaurant Latitude	Restaurant Longitude	Restaurant Name	Restaurant Type
0	43.65426	-79.360636	M5A	43.656369	-79.356980	Impact Kitchen	Restaurant
1	43.65426	-79.360636	M5A	43.650565	-79.357843	Cluny Bistro & Boulangerie	French Restaurant
2	43.65426	-79.360636	M5A	43.650601	-79.358920	El Catrin	Mexican Restaurant
3	43.65426	-79.360636	M5A	43.649970	-79.360153	Izumi	Asian Restaurant
4	43.65426	-79.360636	M5A	43.653475	-79.355458	Copper Branch	Vegetarian / Vegan Restaurant
...	...	...	...	...	...	...	...
160	43.66586	-79.383160	M4Y	43.667872	-79.385659	Kothur Indian Cuisine	Indian Restaurant
161	43.66586	-79.383160	M4Y	43.664665	-79.380641	Loaded Pierogi	Polish Restaurant
162	43.66586	-79.383160	M4Y	43.668759	-79.385694	Wish	Restaurant
163	43.66586	-79.383160	M4Y	43.663894	-79.380210	Kawa Sushi	Japanese Restaurant
164	43.66586	-79.383160	M4Y	43.668514	-79.386686	Wow! Sushi	Sushi Restaurant

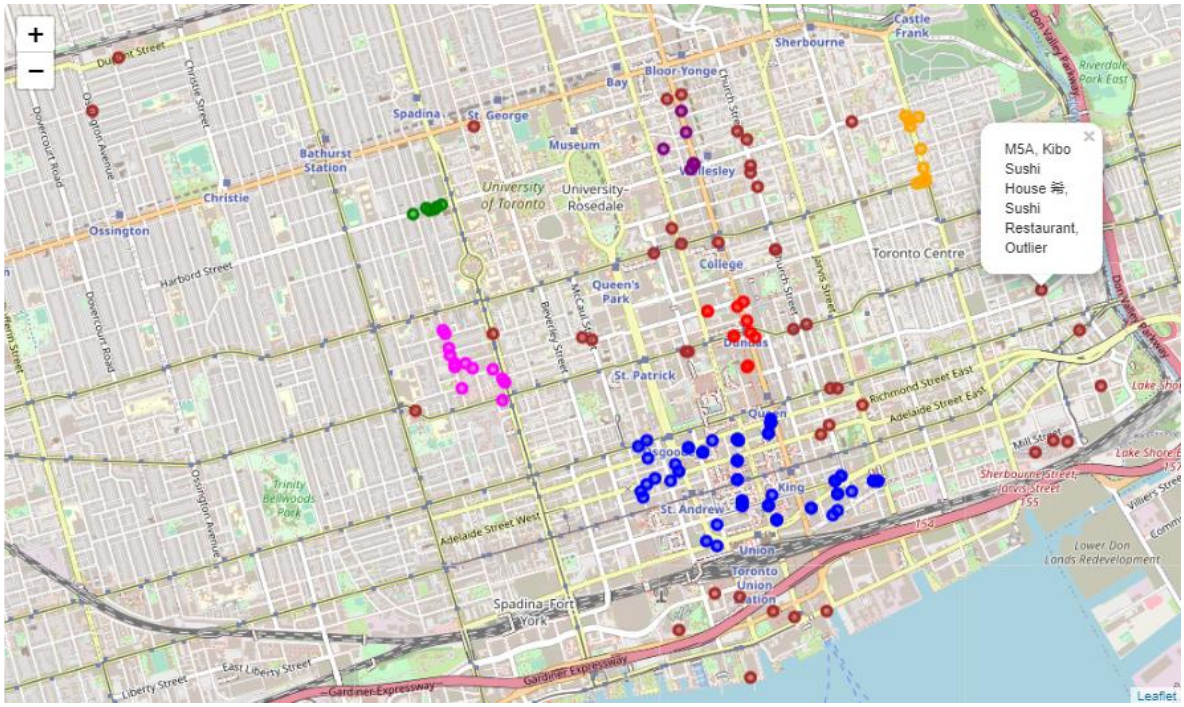
165 rows × 7 columns

I have included all restaurants in the data frame, irrespective of their cuisine or the type of food served, as they are all direct competitors to our restaurant which has no specified type. The following image shows these restaurants on a map of Downtown Toronto:

3.3. Clustering the Restaurants:

To determine the optimal locations for a new restaurant, I used a density-based clustering approach rather than k-means or any other. The reason for this is that our business problem does not specify any specific type (or cuisine) for our restaurant. So, it is of no use to cluster the restaurants by type. Rather, our goal is to simply find locations which are not too densely packed with restaurants. Clusters with a minimum of 5 restaurants within a radius of 250 meters from each other can be considered a sweet spot. Denser areas would mean cutthroat competition, whereas sparse areas could mean there may not be as many customers. Both situations are bad for business, and the full potential of our restaurant may not be realised.

Using Scikit-Learn's DBSCAN, I clustered the restaurants according to the parameters mentioned above, and the following results were obtained, visualised on a map:



4. Results and Discussions:

I have, using Foursquare data, obtained the locations of several restaurants in Downtown Toronto, and clustered them by density, i.e., by the number of restaurants per unit area in Downtown Toronto.

To define the clusters, I have chosen 250 metres, and at least 5 restaurants, meaning a restaurant belongs to a cluster if there are a minimum of 5 other restaurants within a radius of 250 meters from it.

This resulted in the formation 6 clusters (as of now - results may differ as Foursquare data gets updated), with several restaurants remaining outliers, i.e., which could not be put into any of the clusters. For our restaurant, I chose a region that is not too dense with restaurants, as it would be too competitive. When faced with too many choices, consumers tend to default to their one regular choice and it may not be the new restaurant we intend to open, which would be bad for our business.

I also did not choose an outlier area, as it could mean the area might not be competitive enough, i.e., there may not be enough customers, or that it may not be a well-developed commercial area, which is again bad for business.

A restaurant in these areas can work, but we'd rather stay on the safe side. Hence, I chose locations from either low-density clusters, or from outliers that are decently close enough to each other, as suitable places for our restaurant.

Hence, as per our selection criteria the following locations make good candidates for our final list:

- Richmond.
- St. James Town.
- Bay Street.
- Church and Wellesley.
- Harbourfront and Union Station.

All the above locations are very close to the centre of Downtown Toronto, and in areas that are not too dense, but just dense enough with restaurants.

5. Conclusion:

The stated goal of this project was to identify the most suitable locations for opening a new restaurant in Downtown Toronto, Canada. With the help of Foursquare, I identified the top restaurants in the neighbourhoods of Downtown Toronto. Then these restaurants were clustered together based on density, and then locations from low-density clusters and high-density outliers were finalised.

However, this is by no means a comprehensive study, and follow-up studies must be carried out to evaluate other factors such as emergency services, accessibility, look and feel of the area, prices etc. This project also does not consider fast food chains, pubs, bars, coffee shops and other eateries and diners, which can also be addressed in the follow-up studies.

But as a preliminary study, this project does its job of shortlisting suitable locations for a new restaurant in Downtown Toronto, Canada.