Where to open a new restaurant in Toronto?

Back ground

Toronto is a big city judging from the city size. By looking through the food and industry busines in the city, the density does not reach its limit. Therefore, there are rooms for a new restaurant. Then we have to answer the most important question, where?

Business Problem

Opening a restaurant is a very challenging task. As a start, it is very important to decide the location wisely. The current situation is that an entrepreneur usually carelessly relies on common sense and domain knowledge to choose a restaurant type and a good spot. Needless to say that too often an inconsiderate decision leads to a poor income and inevitable bankruptcy. Where to open it?

I do not know the business rules in Toronto. In the Netherlands, it is not allowed that you open a restaurant where there are no restaurants nearby. In other words, the government tries to group all restaurants together in a city center or a shopping area. You can not buy a house near a school and start a fast food restaurant there.

As a starter, an easy way is to open a restaurant where it is surrounded by other restaurants. Because you do not need to worry about customers. You can simply share them with other business. Then the restaurant can build up its reputation and grow in to a mature business. Once it is getting famous, we may move it because customers will follow.

To sum up as a starter, opening a new restaurant should be where most of restaurants are.

Let us do the restaurants data analysis in Toronto and find the concentrated restaurants area for a new restaurant’s starter.

Target audience

Investors, entrepreneurs, and chefs interested in opening a restaurant in Toronto, who may need a piece of objective advice on where exactly it should be opened.

Data and methods

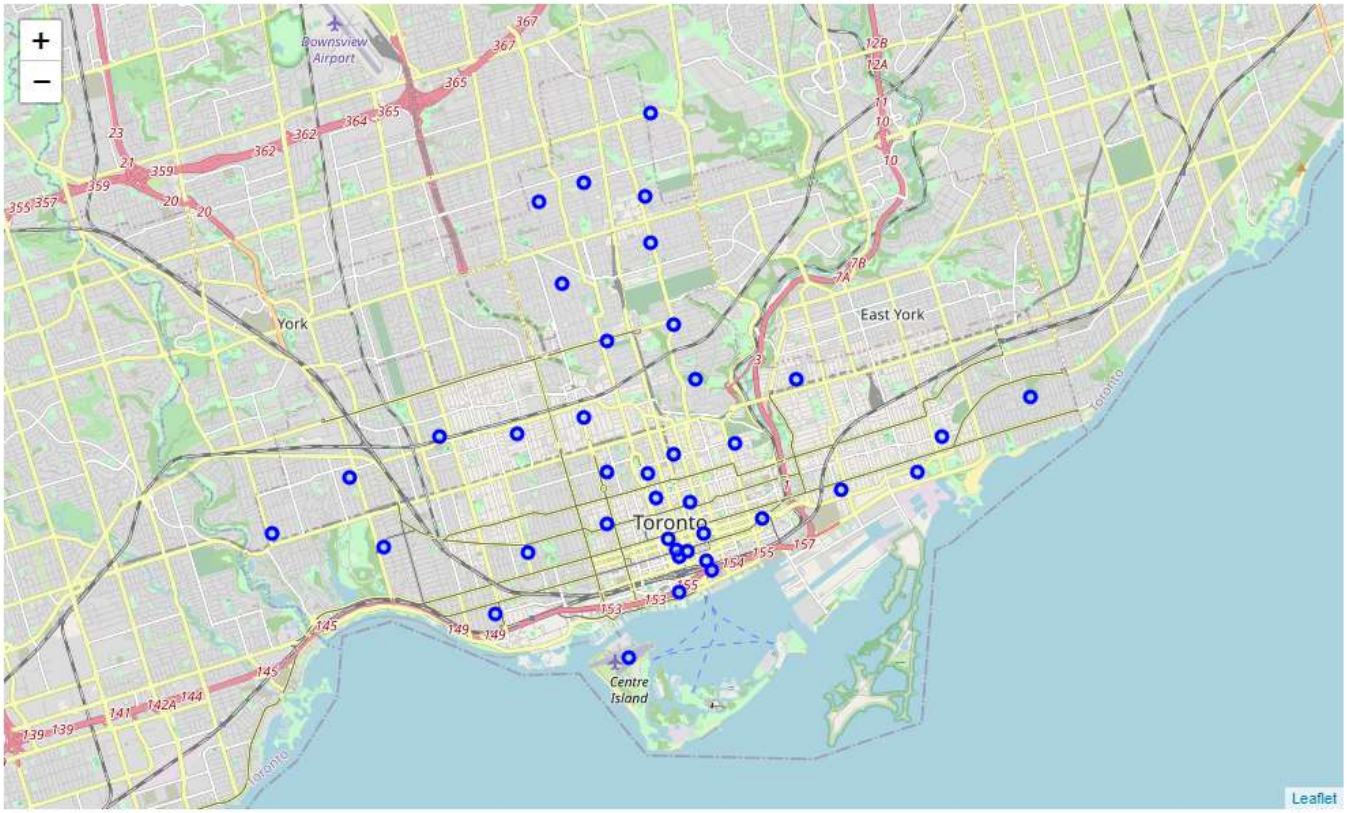
1. Using a table on https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M) collect information about Toronto boroughs and locations and Postal Code.
2. Use the Geopy and Folium library to get the coordinates of every locations and map geospatial data on a Toronto map.
3. Use Foursquare API, collect the all restaurants in Toronto and their location via an exploring query.
4. Group collected restaurants by their locations using K-mean algorithm and find out the cluster which has the most nodes. This cluster represents the busy restaurants area.
5. Calculate the distance (Euclidean distance) from each borough to the biggest cluster center and select a few boroughs which have smaller distances than others. These boroughs will be my recommendations.
6. Visualize on the map and see if my recommendations are convincing.

Methodology

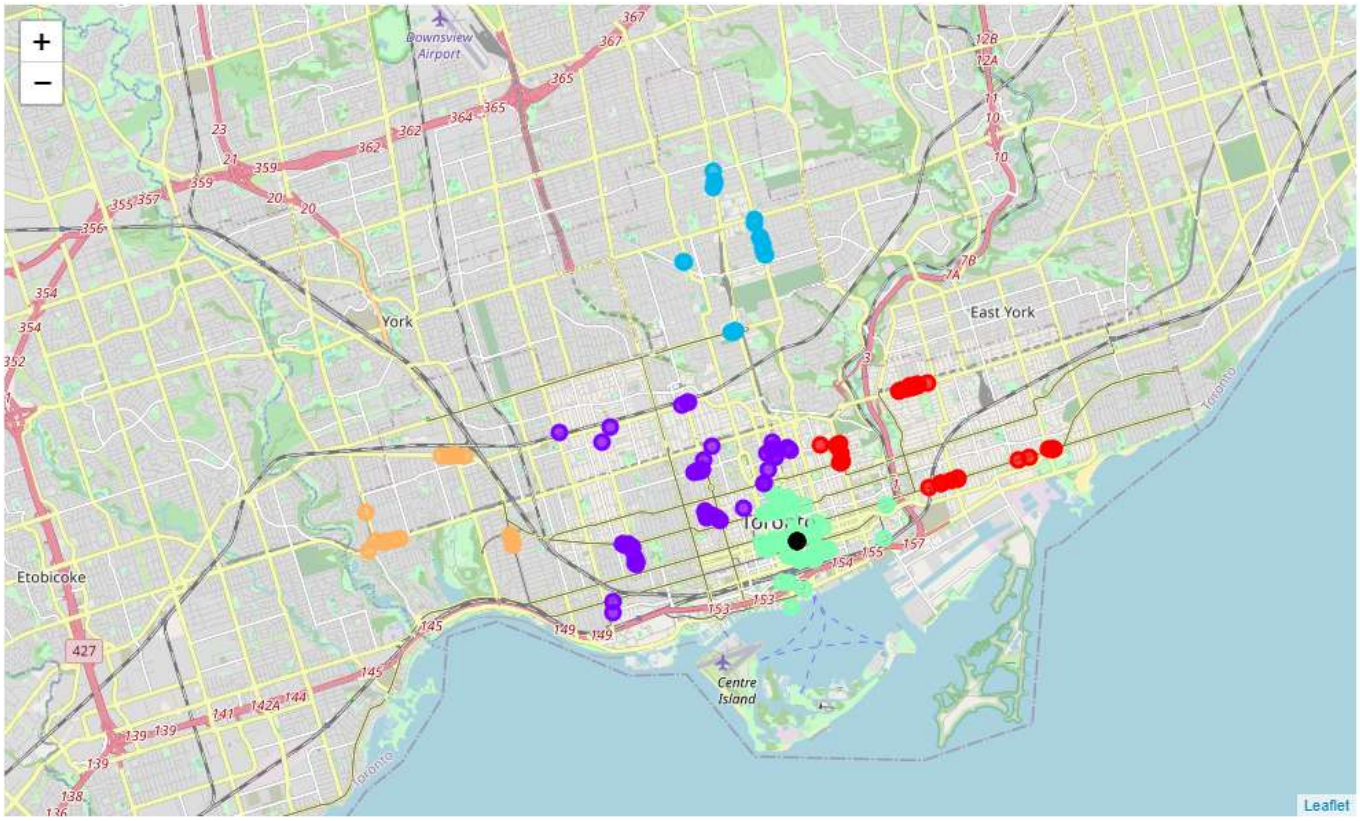
By taking a table on https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M), and filling in the missing information, we can make a table of Toronto boroughs and neoghbourhoods. Here is a snapshot of the first 5 rows of this table.

	PostalCode	Borough	Neighborhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government

As you can see, the locations are missing. We read locations from geospatial data, 'http://cocl.us/Geospatial_data (http://cocl.us/Geospatial_data)'. Together with Toronto locations that we retrived, we can demonstrate all neiborhoods on the map centered at Toronto. However since there are quite some, we select neighborhood from Boroughs, which name has 'Toronto' in. I use these neighborhhod data as an example to demonstrate my methods. The same method can be applied to other boroughs as well. The slected neighborhoods are: 'Downtown Toronto', 'East Toronto', 'West Toronto' and 'Central Toronto'. The following image shows such a map.



Using Foursquare API, I collect all restaurants in Toronto and their location via an exploring query. The number of resuarants in the above neighborhood are in total 196.
After grouping collected restaurants by their locations using K-mean algorithm, we can see that the biggest cluser is close to the seashore. This is shown in the following map.
The center location of the biggest cluster is [43.649928, -79.379402] and it is marked in balck on the map.



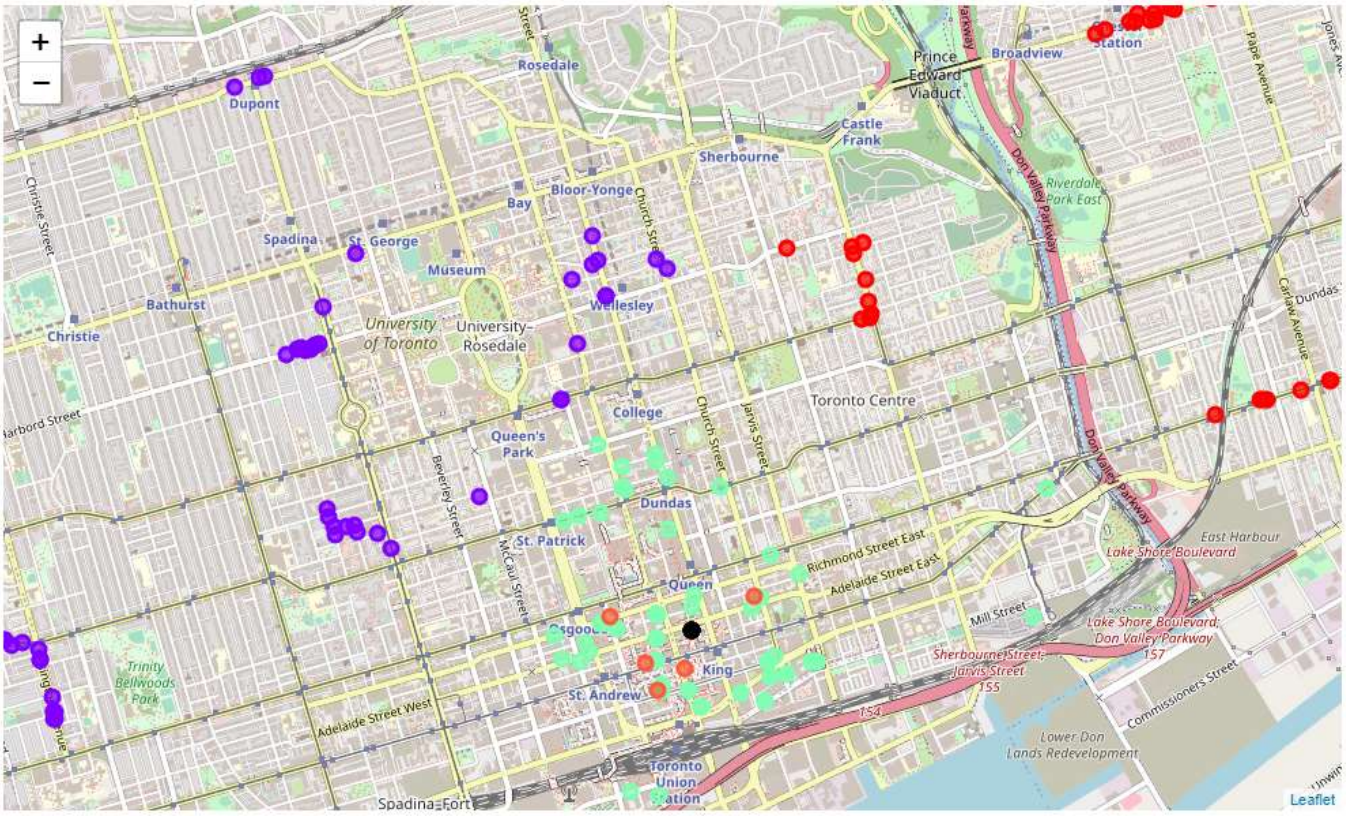
results

Now we have prepared locations of neighborhoods, and the clusters of restaurant in these neighborhoods. Using this informations, we can decide which neighborhoods are closest to the biggest cluster center, i.e., the area center where most restaurants are located. The center location is calculated by averging of all the nodes locations in this cluster.

By calculating Euclidean distance from each neighborhood to the cluster center, we pick up the first 5 boroughs which are closest (least distances). Their names are listed in the following table with order of asceding. The top one on the list,'Commerce Court, Victoria Hotel' has the shortest distance.

	Latitude	Longitude	Neighborhood
16	43.648198	-79.379817	Commerce Court, Victoria Hotel
36	43.648429	-79.382280	First Canadian Place, Underground city
13	43.647177	-79.381576	Toronto Dominion Centre, Design Exchange
3	43.651494	-79.375418	St. James Town
8	43.650571	-79.384568	Richmond, Adelaide, King

The suggested neighborhoods locations no the map are marked in orange. They are: 'Commerce Court, Victoria Hotel','First Canadian Place, Underground city', 'Toronto Dominion Centre, Design Exchange', 'St. James Town' and 'Richmond, Adelaide, King'.



Discussions

As we have recommended 5 neighborhoods that suitable for a starup restaurant business. Our methodology is that the more restaurants the neighborhood is surround by, the more customers the restaurants will have. Thus it is a good location of starup restaurant business.
The futher question we could investigate is what type of restaurant to open. However, this question requires more domain knowledge on restaurant business. If there are all Italian restaurants in the neighborhood, it could mean that people living there are fond of Italian food. It could also mean that there is no more room for another Italian restaurant. Then we have to study the menu and specialties of each Italian restaurant. We need someone who understands different pizza's.

Conclusion

In this report we work out a methodology to determine where are the most locations if you want to open a new restaurants. We collected information about Toronto boroughs from Wikipedia, and using geospatial libraries mapped them. Using Foursquare API, we collect the all restaurants in these neighborhood. Then we clustered restaurants by the k-means algorithm and find out where the biggest cluster is located. By analising distances from neighborhoods to the center of the biggest cluster, we select top 5 neighborhoods which have the least distance. Eventually we visualized clusters on the map, as well as the best locations/neighborhoods for opening the chosen type of restaurant.
This type of analysis can be applied to any city of and to any type of venue (shopping, clubs, etc.) that is available in Foursquare database.

In [ ]: