Coursera Capstone: Opening a Specialty Bakery in the city of Toronto, Canada

Introduction

For the Capstone project, I am reviewing the scenario in which an enterpreneur wants to explore the possibility of opening a new organic bakery in the city of Toronto, Canada. The bakery will specialize in gluten, wheat and dairy free products. In the last years, there has been a noticeable improvement in the diagnosis of food allergies. Previously, what was an unknown problem which affected many people has become in a manageable ailment which is easily treated through modification of diets.

Business Problem

The new bakery aims to offer products targeted to people who suffer from food allergies. The range of products offered will be spelt breads, muffins and pastries. Spelt will be used since there are less people who are allergic to this grain. A range of yeast-free breads, dairy-free muffins and pastries, as well as egg-free products will be offered.

Target Audience

The new specialty bakery will focus on two customer segments. One is the Allergy segment and the second is the Dieter segment, consisting of people who are making modifications in their diet based on a desire to lose weight.

Data necessary for the project

The data necessary for the project will be collected from the following sources:

Neighborhood data from Toronto¶
This will be acquired from a list of neighborhoods in Wikipedia page: "List of postal codes of Canada: M" in the webpage https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
Description: List of neighborhoods and postal codes of Toronto

Geographical coordinates¶
The data set http://cocl.us/Geospatial_data will be used to acquire the coordinates (latitude and longitude) corresponding to the different postal codes.
Description: The data base provides the latitude and longitude of a given location based on the postal code.

Foursquare Venues data¶
API call to Foursquare and acquisition of .json data related to the venues located at a given radius around each neighborhood.
Description: It provides data about venues around a given neighborhood. Venues are catergorized and information is provided about ratings, reviews and tips given by users of Foursquare.

Additional data from specialty bakeries in Toronto
Foursquare provides information related to location of bakeries in the area of Toronto and detailed information regarding rating, likes, tips from users. Additional information regarding specialty bakeries in Toronto was found in http://www.blogTo.com/bakeries

Methodology

A list of the neighborhoods in Toronto was obtained from the Wikipedia page https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:M. The list contains the names of the different Boroughs, neighborhoods and postal codes in Toronto. Web-scraping was used to consolidate data frame information pulling tabular data from the webpage into a data frame.

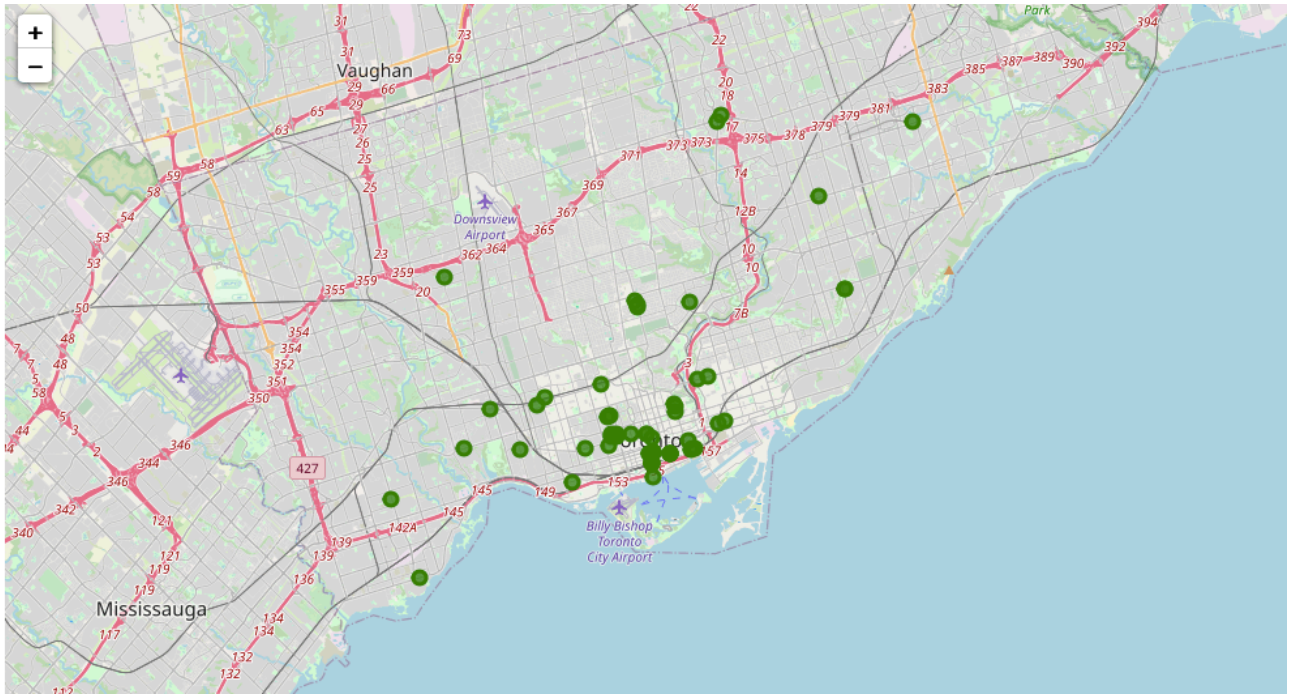
Geographical coordinates for each neighbourhood were obtained from a .csv file provided by the IBM team for this course (dataset: http://cocl.us/Geospatial.data. The coordinates and neighborhood locations were visualised in a map of Toronto created using the Folium package.

The Foursquare API was used to pull a list of 200 venues within 500 m radius from each neighborhood, based on the latitude and longitude of each neighborhood. The information also included categories for each venue some other detailed data as popularity or each venue based on ratings, likes and count of tips given by users of Foursquare. From the list of venues, a data frame was created containing only the existing bakeries in the Toronto area.

Based on the users ratings, likes and tips, the neighborhoods of Toronto were clustered, using the k-means algorithm. The optimal number of clusters for the data set was determined using the Elbow method, which resulted in four clusters. k-means allocates every data point to the nearest cluster. Recommendations and conclusions were made based on the results of the clustering.

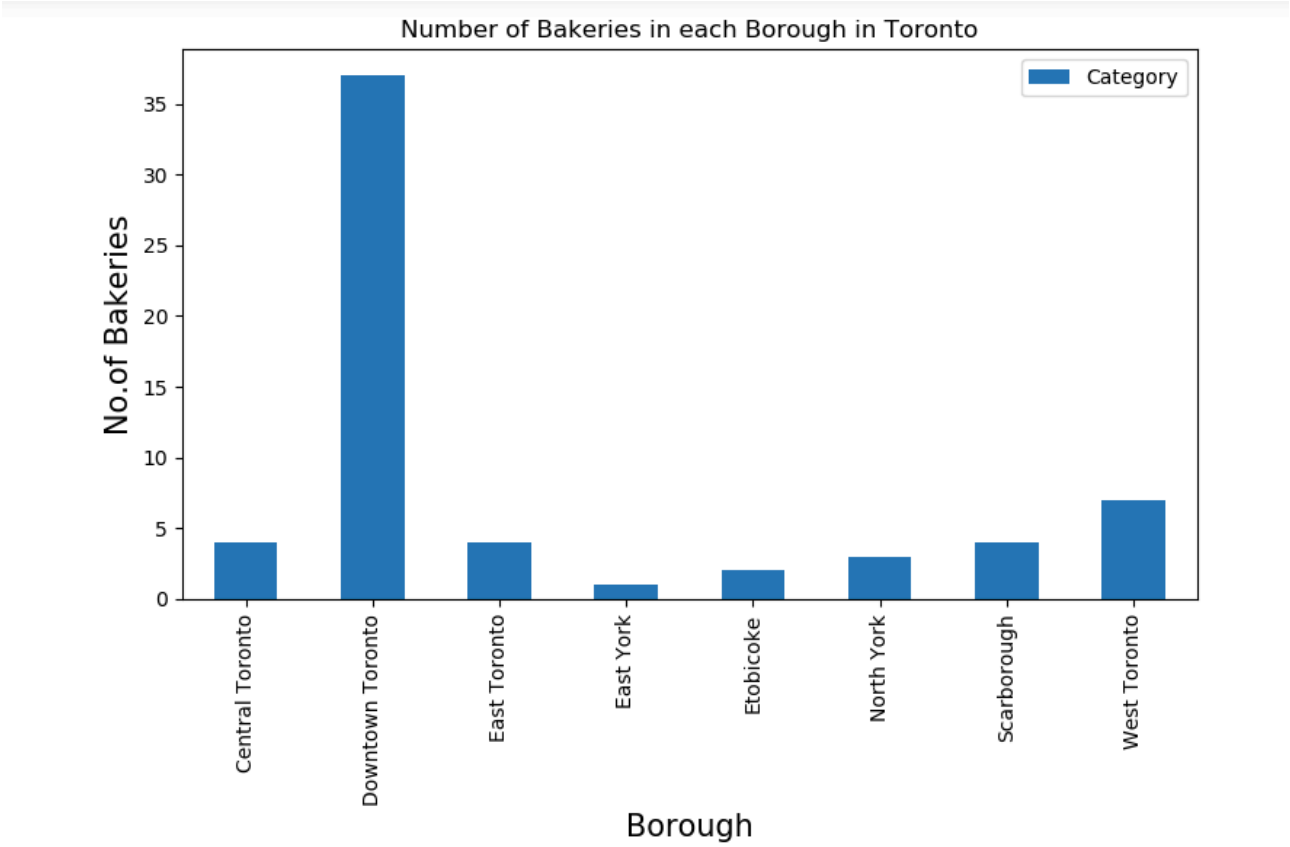
Discussion of the Results

Using the information retrieved using the Foursquare API related to the 200 venues within a radius of 500 m from each neighborhood in Toronto, the location of the bakeries in the area of Toronto was determined as showing the map below.

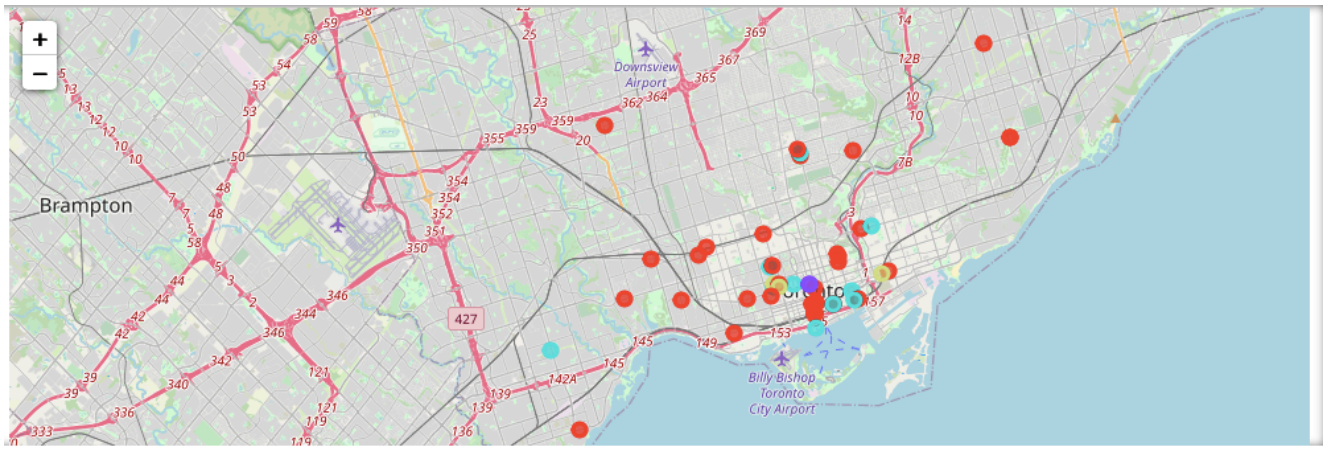


Location of bakeries in Toronto area (source: Foursquare)

Careful examination of the map above and from the data retrieved from Foursquare, shows that the majority of bakeries in Toronto are located in Downtown Toronto, followed by bakeries in West Toronto and Central Toronto. The rest are distributed in smaller number around other Boroughs, as shown in the figure below.



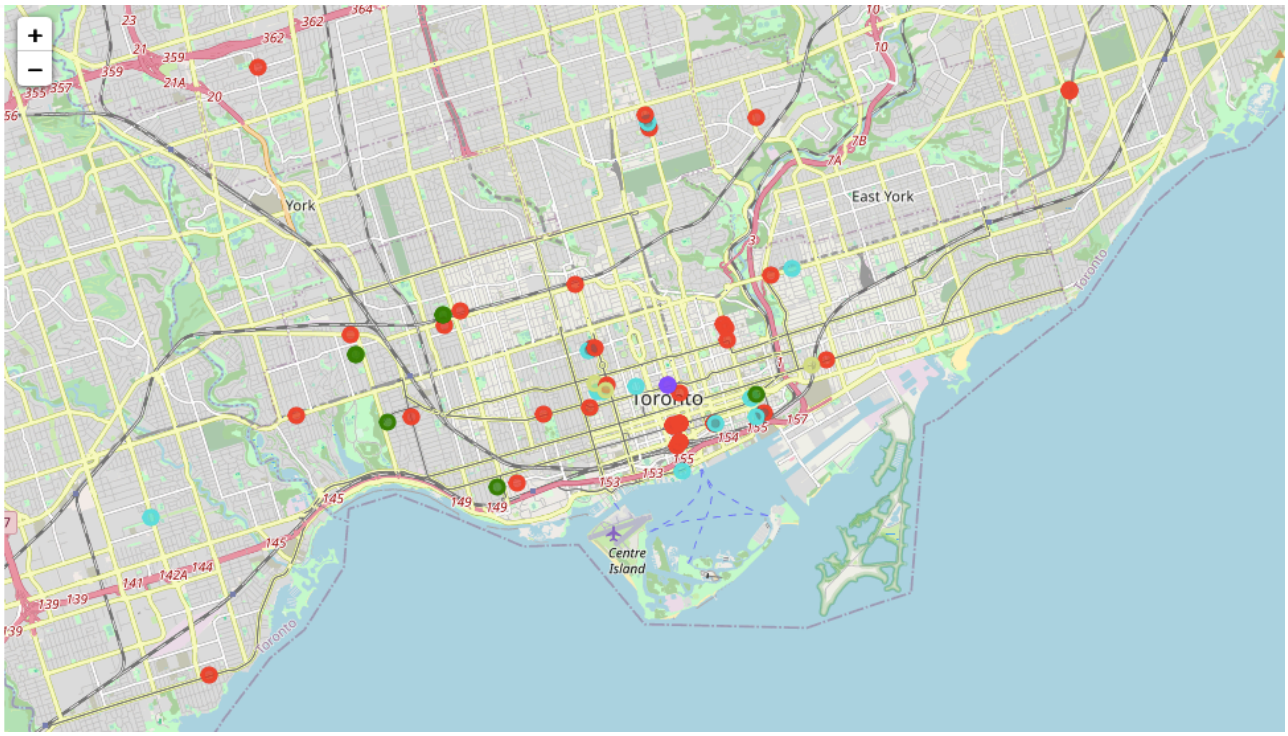
Clustering the neighborhoods in Toronto using k-means algorithm, showed that the neighborhoods can be categorised into four clusters according to the ratings, likes and tips given by users of Foursquare. The map below shows the distribution of clusters.



Examining the results from the clusters, it can be seen that:
Cluster 1 (red) shows the biggest number of bakeries spread in Downtown Toronto. However for this group 70% of the bakeries received an average to lower than average rating
Cluster 2 (purple) located in Downtown Toronto, in the area of Little Tokyo shows the bakery with highest number of likes from the data set
Cluster 3 (blue) shows bakeries located mainly in Downtown Toronto, in the vicinity of the University of Toronto, Chinatown and Regent Park with 46% of them having a higher rating (8-9)

Cluster 4 (green) shows bakeries located around Chinatown, which have a number of likes higher than average compared to the other clusters.

In addition to the information acquired from Foursquare, we used information offered at http://www.blogTo/bakeries, in which we could find the location of specialty bakeries located in the area of Toronto. These specialty bakeries offer the same type of products that our intended specialty bakery would offer, so they would be direct competitors for our business. These bakeries are located predominantly at the West of Downtown Toronto and one shop in the area of Regent Park and they are shown in dark green in the map below, which also contains the clusters.



Map showing existing specialty bakeries (dark green) and clusters from k-means analysis

Conclusions

Based on the results of our analyses we can conclude that the number of specialty bakeries in the area of Toronto is very limited, so the prospect of opening a specialty bakery is good. The results from the clustering of the neighborhoods based on the ratings, likes and tips counts for the different bakeries show that even though Downtown Toronto has a high concentration of bakeries, the bakeries located in the vicinity of the University of Toronto, China Town and Regent Park have the higher ratings (8-9) and the rest have an average rating. The clusters also showed that bakeries in the area of Little Tokyo received the highest number of likes compared to bakeries in other locations.

Areas away from Downtown or West Toronto offer very little competition since there is a very low concentration of bakeries for those locations. Even if the enterpreneur decided to open a specialty bakery in Downtown Toronto, the higher competition is from traditional

type of bakeries, the number of specialty bakeries offering the same type of intended products is very low to non-existing, so this would represent a good opportunity.

Further analyses to help in narrowing the choice of adequate areas to open the specialty bakery would take into consideration population density, income of the residents and costs of rents.