Background/Business Problem

As different health ailments are on the rise and people look to be healthier with their bodies and minds, the trend to convert to a vegetarian/vegan diet is nothing new. In a study conducted by EPIC-Oxford, it was found that vegetarians were, on average, 25% less likely to die of heart disease. Further to this, if you stop eating red meat (whether or not you become a vegetarian), you'll eliminate a risk factor for colon cancer.

People choose to become vegetarian for several reasons, including: religious convictions, concerns about animal welfare or the use of antibiotics and hormones in livestock, or a desire to eat in a way that avoids excessive use of environmental resources.

With the rise in individuals around the globe turning to a vegetarian diet and businesses taking advantage of this trend, it is clear that this type of analysis would be great for a restaurant owner who is looking to open up a restaurant in the neighborhoods of Toronto, Canada.

Data

The data used in this project are:

- List of postal codes of Canada Wiki: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M for access to neighborhood data of Toronto region.
- Geographical coordinates of the neighborhoods: http://cocl.us/Geospatial_data for getting the longitude and latitude data for the neighborhoods.
- Foursquare database: https://Foursquare.com to be used in order to explore the desired neighborhood data for various venue details and access the JSON files. This data shall be utilized to map the venue in neighborhoods.

Research Methodology

In this project, I used multiple sources such as wiki data, geocodes, Foursquare API, beautifulsoup to conduct web scrapping and preprocessing of data. Below I have outlined all this information

As you can see in the code, there were many Boroughs not assigned or NaN so they were removed and the Postal Code variable was changed for indexing done later. Once the data was cleaned up, the latitude and longitudes were added for each location and columns were rearranged so the data frame was easier to read. Furthermore, new tables were created with latitudes and longitudes corresponding to different Postal Codes. Then, the foursquare API was used to get venues in the radius of 500 meters.

Neighborhoods were clustered and k-means was used to determine how many neighborhoods per cluster exists. It was then determined which cluster has the most amount of vegan/vegetarian restaurants and which neighborhood is best for opening a new restaurant.

Discussion

Most of the Vegan/Vegetarian Restaurants are in cluster 1 represented by the red clusters. The Neighborhoods located in the Toronto area that have the highest average of Vegan Restaurants is North York. Even though there is a huge amount of Neighborhoods in cluster 2, there are little to no Vegan Restaurant. We see that in the Downtown Toronto area (cluster 3) has the second last

average of Vegan Restaurants. Looking at the nearby venues, the optimum place to put a new Vegan Restaurant is in Cluster 3 as there are many Neighborhoods in the area but little to no Vegan Restaurants therefore, eliminating any competition. The second best Neighborhoods that have a great opportunity would be in areas in Cluster 2. Having 70 neighborhoods in the area with little to no Vegan Restaurants gives a good opportunity for opening up a new restaurant. This concludes the optimal findings for this project and recommends the entrepreneur to open an authentic Vegan restaurant in these locations with little to no competition.