Anant Venugopal

**Battle of the Neighborhoods Capstone Project**

**Introduction**

1. The information problem I came up with is for prospective restaurant owners in Toronto. The goal of this project is to help a prospective owner decided which neighborhood to open up a sushi restaurant in. Anyone who opens a restaurant obviously wants business to be good and to have a lot of customers, deciding which area to open the restaurant in is crucial to the restaurant’s success. Ideally a good location to open a restaurant would have little competition from similar restaurants and would be close to the busiest areas of the city e.g. The center of the city.

**Data**

1. Using the Foursquare API we can pull relevant information on Toronto neighborhoods and find out the optimal place to open a sushi restaurant. The Foursquare API contains detailed information of venues in different locations of Toronto including restaurants and what kind of restaurant it is. Using this data we can find the frequency of rival sushi restaurants in these neighborhoods to find an optimal spot to open the restaurant in. In addition, I pulled the neighborhood and postal code information for the boroughs of Toronto from Wikipedia. This data also includes useful latitude and longitude information for the various neighborhoods and venues.

**Methodology & Results/Discussion**

1. The first thing to explore was the general layout of the neighborhoods within Toronto. To do this I cleaned the Wikipedia data, put it within a data frame, and combined it with geospatial coordinate data stored on a CSV. Then I plotted these coordinates onto a Folium map so that we can see the general city limits and how neighborhoods are distributed. From this I discovered that most of the neighborhoods are evenly distributed throughout the city, however many neighborhoods are densely consolidated next to the center of the city.
2. Next I pulled data from the Foursquare API to get more information on the kinds of restaurants in the neighborhoods. After consolidating the data I split it into two different data frames. One with venue data on all restaurants in Toronto, and one with only sushi restaurant data. After that I put both data frames into heatmaps. After heatmapping the data I saw that the restaurant data basically followed the trend of the previous map in that most of the restaurants were closely bunched around the center of the city and then relatively evenly distributed around the rest of the city. The sushi restaurants were almost all close to the very center of the city and sparsely scattered throughout the rest of the city. Just northwest of the city was found to be a potential good target for the sushi restaurant given the lack of sushi restaurants in that area while also being quite close to the center of the city.
3. Then I ran onehot encoding on the restaurant data frame and ran the k-means clustering machine learning model. This unsupervised model was run on the restaurant data in order to find out what neighborhoods have similar types of restaurants in them. The model revealed certain types of neighborhoods that would be more or less desirable to open a sushi restaurant in. The ‘0’ cluster involved mainly Asian cuisine, this includes seafood and Japanese restaurants. This makes it an undesirable classification for a sushi restaurant as those cuisines of restaurants directly compete with sushi restaurants. The ‘1’ cluster was the largest cluster, this cluster encompassed all of the restaurants right next to the center and other neighborhoods with very similar and typical restaurant variety. All current sushi restaurants were in neighborhoods that fell under the ‘1’ cluster. Cluster 1 features a wide variety of cuisines, the most popular cuisine varied a lot between neighborhoods. Clusters 3 and 4 were both too small and far away from the city to be considered as potential locations. This leaves mainly clusters 1 and 2 as potential neighborhood cluster classifications for our sushi restaurant.
4. Then I plotted the clusters onto a map to see how that aligns with our previous heatmaps. We can see that cluster 2 has two potentially good locations in the northwest of the city’s center.

**Conclusion**

1. Based on these results we can conclude that the best place to open a sushi restaurant would be somewhere in clusters 1 or 2. Narrowing that down further to neighborhoods close to the center of the city we get many neighborhoods very close to the center of the city. Eliminating any neighborhoods with Japanese restaurants as one of the top 3 most common restaurants we get these following neighborhoods are prime locations to open the restaurant; Kensington Market, Chinatown, Grange Park, Regent Park, Harbourfront, St. James Town, Cabbagetown, The Annex, North Midtown, Yorkville. Compounding this data with the heatmap the prime candidate would be to open the restaurant in the M5R postal code area encompassing The Annex, North Midtown, Yorkville neighborhoods. This area is relatively far from any other competitive restaurant compared to the others on the list and is also still very close to the center of the city. This makes it a very desirable location for a sushi restaurant in my opinion.