**Capstone Project: Choice of Neighborhood in Toronto for Punjabi Restaurant**

1. Introduction to the Idea:

One of the largest South-East Asian ethnic groups to live in Canada are Punjabis, hailing from the Northwestern parts of India and Eastern Pakistan. Punjabis first arrived in Canada during the late 19th century to work in the forestry industry. Primarily concentrated in the western province of British Columbia, the Punjabi population initially peaked in 1908 before an ensuing period of population decline and stagnation followed. In the mid-20th-century Canadian immigration laws were relaxed, fostering rapid population growth into the present day. Today, the largest Punjabi communities in Canada are situated in the province of British Columbia, concentrated in Vancouver, and the province of Ontario, particularly in Toronto. The dominant religion of Punjabis, Sikhism is the 7th largest in Canada, accounting for about 1.4% of the total population, roughly 500,000 people as of 2011. But numbers and statistics aside, it is the cultural influence of the Punjabis that has contributed much to the social entropy of the country. Although the largest settlement of Punjabis still remains in British Columbia, By the 1980s, the traditional Punjabi immigration patterns began to shift. Ontario soon became an important center of immigration to Canada. Large Punjabi populations began to appear across the Greater Toronto Area. Later in the 1990s, Alberta became another important immigration destination for Punjabis. Most concentrated in Edmonton and Calgary. Today, the Punjabi population of Canada is 668,240 with the largest community still located in British Columbia. In addition, Punjabi is the third most spoken language of the Parliament of Canada. So it is quite imperative that there is a burgeoning demand for South-East Asian cuisines across the different neighborhoods of Toronto, and it is this prospect that draws many Punjabi youths from both India and Pakistan to immigrate to Canada, which is also reflected by the data pertaining to worldwide immigration to Canada over the last few decades. So for an aspiring Punjabi youth, qualified in culinary arts or hotel management to open a restaurant in Canada, Toronto is likely to come up as one of the foremost choices. This is where our program comes in, where we can analyze them and visualize the distribution of such region-specific restaurants across different neighborhoods in Toronto so that the person can have a tentative idea of the possible demand that will help him to better decide the final location of his enterprise.

1. Description of the Dataset used:

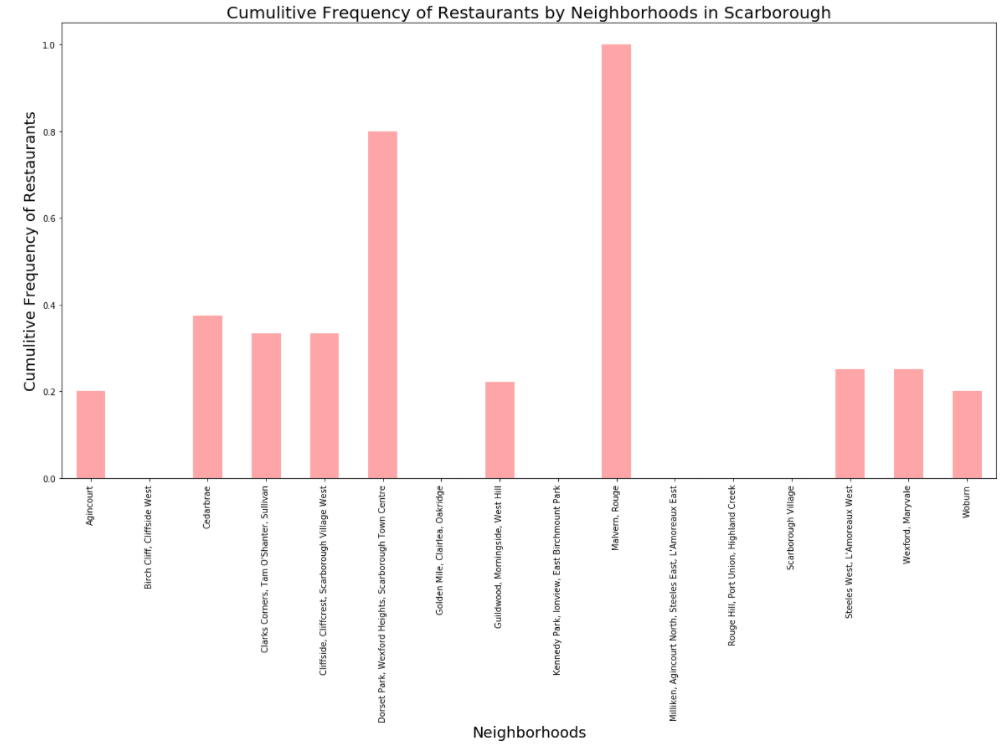
The data to be used were obtained from Wikipedia, converted to an excel file named List 1.csv which contains a list of different neighborhoods in Toronto pertaining to the listed postal codes and their corresponding boroughs. The geospatial data containing the latitudes and longitudes for the corresponding postal codes have been obtained from an excel file called Geospatial\_Coordinates. The combined dataset with both the geospatial and neighborhood datasets was combined using pandas, and the specific neighborhoods of interest were obtained from Foursquare location data using the Foursquare API, to obtain the data regarding various places of interests like restaurants, coffee shops, parks, toy stores, etc. given a certain radius of interest. The areas of interest to be investigated as prospective locations for the restaurant were obtained from the Wikipedia page dedicated to Punjabi Canadians.

1. Methodology:

The analysis was organized on the basis of the information on the boroughs in Toronto that are commonly known to have a larger presence of the Punjabi community as obtained from Wikipedia, and the localities on which data were available in the above stated files for investigation. Thus we narrowed down our investigation on the Scarborough and Greater Toronto neighborhoods. First a combined dataframe comprising of both the neighborhood and the corresponding geospatial data were created, wherefrom the data pertaining to each neighborhoods were queried into separate dataframes. Using the Foursquare API venues specific to each neighborhood were obtained, and from there, the data points under restaurant category were separated and were subjected to clustering, a form of unsupervised machine learning, in order to better organize the geospatial data, a visualization of which was then created using the Folium package.

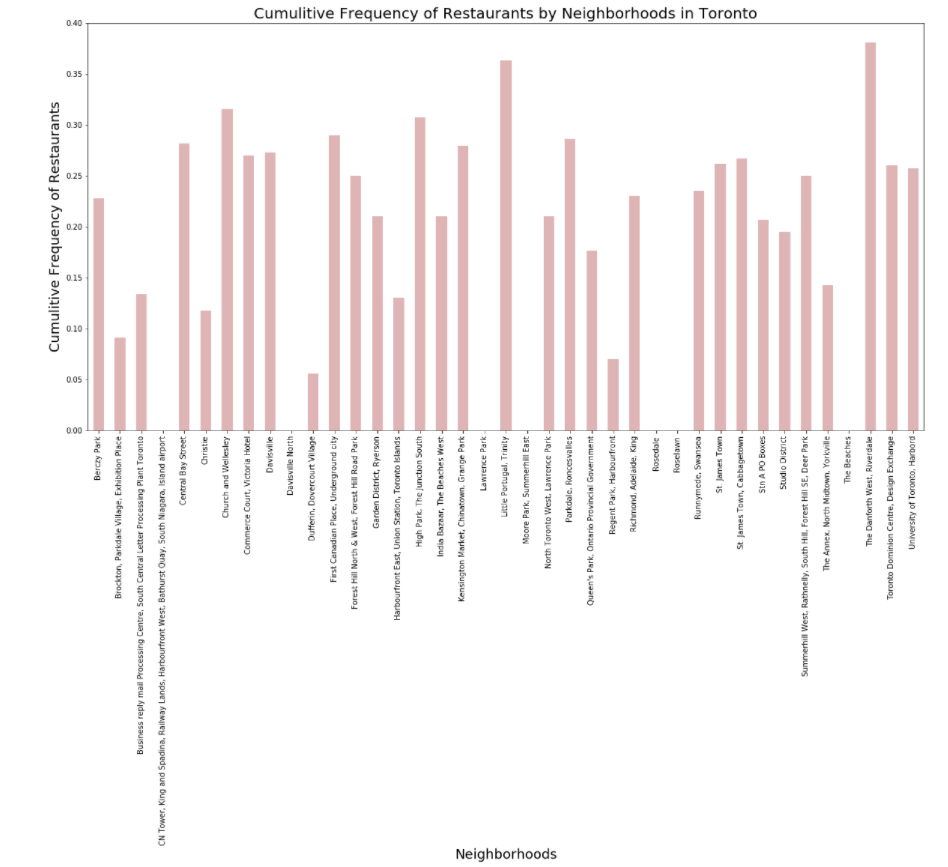
1. Results:

On the basis of restaurant frequency, as shown by the following graphs, and the clusters we selected the suitable neighborhoods in the 2 separate cases and the combined the results to further narrow down the spots for setting up the Punjabi restaurant.



From the bar charts showing the frequency of restaurants in the different localities, the ideal locations in the 2 areas are summarized in the following tables. The selections have been made on the basis of lower restaurant frequency.

|  |  |
| --- | --- |
| Scarborough | Greater Toronto |
| Golden Mile, Clairlea, Oakridge | Davisville North |
| Kennedy Park, Ironview, Birchmount Park | Lawernce Park |
| Scarboroug Village | Rosedale |
| Scarborough | **Greater Toronto** |
| Milliken, Agincourt North, Steeles East, l’Amoreux East | CN Tower, King and Spadina, Railway Lands, Island Airport, |
| Rouge Hill, Port Union, Highland Creek | Roselawn |
| Birch Cliff, Cliffside West | The Beaches |



* 1. Conclusion about Scarborough Neighborhood:

Assuming that none of the restaurants in the list that do not have their genre specified are Indian restaurants, it is advisable not to choose the neighborhoods of Dorset Park, Wexford Heights, Scarborough Town for setting up a Punjabi restaurant, since it is likely to be designated as an Indian restaurant by the local populace who do not belong to the Punjabi community. Along with that, the neighborhoods of Wexford and Maryvale should be preferably avoided since Middle-East and North-West India shares a taste palette. The most suitable location for the restaurant are chosen to be the Wodburn, West Hill, Guildwood, Morningside and Agincourt since all the restaurants located over there are of completely different genre.

* 1. Conclusion about Greater Toronto Neighborhood:

To begin with, this section encompasses a larger area than that covered in the previous section. Hence we are going to analyse the results clusterwise, starting with cluster 7, covering the neighborhoods of High Park and The Junction South. Among the top 5 venues, there are virtually no restaurants that are anything like the taste palette of Punjabi food, the closest being Mexican food which as we all know tend to be very spicy. Same goes for cluster 6 and cluster 2 and 3. Going on to cluster 5, the Davisville neighbourhood will probably not be a good choice given the 2nd most common venue. Same goes for the 4th most common venue of Runnymeade and Swansea neighborhoods where the vegan restaurant is likely to be preferred by the Sikh community who make up the largest section of the Punjabi diaspora. In the 4th cluster all the neighborhoods but St. James Town and Cabbagetown may be considered as suitable restaurant spots. Cluster 1 is also comprised of mostly suitable neighborhoods barring Commerce Court, Victoria Hotel, Church and Wellesly. As for Cluster 0, given the large number of spots that it encompasses, we had to resort to a code in order to have a list of suitable spots for a Punjabi restaurant, that returned the following list:

["Queen's Park, Ontario Provincial Government", 'Garden District, Ryerson', 'Berczy Park', 'Christie', 'Richmond, Adelaide, King', 'Little Portugal, Trinity', 'Toronto Dominion Centre, Design Exchange', 'Brockton, Parkdale Village, Exhibition Place', 'Commerce Court, Victoria Hotel', 'Studio District', 'North Toronto West, Lawrence Park', 'University of Toronto, Harbord', 'Summerhill West, Rathnelly, South Hill, Forest Hill SE, Deer Park', 'Stn A PO Boxes', 'First Canadian Place, Underground city', 'Business reply mail Processing Centre, South Central Letter Processing Plant Toronto']

1. Conclusion:

Keeping the analysis results in mind, the spots chosen to be the most suitable for setting up a Punjabi restaurant are as follows:

For Scarborough: **Golden Mile, Clairlea, Oakridge, Woodburn and Agincourt**

**For Greater Toronto Area: India Bazaar, University of Toronto, Lawrenece Park**

**Note: The above report is based on a particular instance of running the code and noting the outputs all at a time. Outputs are be subjected to minute changes in each run due to the clusters created.**