**Capstone Project - The Battle of the Neighborhoods**

**Table of contents:**

\* Introduction: Business Problem

\* Data

\* Methodology

\* Results

\* Discussion

\* Conclusion

**Introduction: Business Problem:**

Toronto is the provincial capital of Ontario. With a recorded population of 2,731,571 in 2016, it is the most populous city in Canada and the fourth most populous city in North America. The city is the anchor of the Golden Horseshoe, an urban agglomeration of 9,245,438 people (as of 2016) surrounding the western end of Lake Ontario, while the Greater Toronto Area (GTA) proper had a 2016 population of 6,417,516. Toronto is an international center of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world.

A restaurant is a business which prepares and serves food and drink to customers in return for money, either paid before the meal, after the meal or with an open account. Toronto is a famous city. Its food culture includes an array of international cuisines influenced by the city’s immigrant history. So, it is evident that to survive in such competitive market it is very important to strategically plan.

**Data:**

Various factors need to be studied in order to decide on location such as;

Toronto Population and demographics

Who are the competitors in that location?

Cuisine served / Menu of the competitors

Are there any Farmers Markets, Wholesale markets etc. nearby so that the ingredients can be purchased fresh to maintain quality and cost?

Are there any venues like Tourist attractions, Entertainment zones, Parks etc., nearby where floating population is high

Segmentation of the Borough

Untapped markets

Saturated markets etc. And the list goes on… Even though well-funded XYZ Company Ltd. needs to choose the correct location to start its first venture. If this is successful, they can replicate the same in other locations. First move is very important, thereby choice of location is very important.

To build a recommendation model, following datasets and information are considered for analysis;

Scrapped Wikipedia using Pandas, to extract information about 4 toronto boroughs also known as local authority districts. Also, considered local areas or neighborhoods for each borough for detailed analysis.

I used Foursquare API to get information about available restaurants for a given neighborhood and borough in Toronto. The API also provided information about restaurant styles based on cuisine.

**Methodology:**

In this project we will analyze the Boroughs of Toronto and its neighbourhoods. We will use Geopy to find the latitude and longitude of Toronto and will search nearby restaurants in 500km radius area. For this we will use Foursquare API. Then we will cluster it by using k-means clustering which is an unsupervised clustering model. Based on this we will conclude about the project.

**Results:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Cluster** | **1st common restaurant** | **2nd common restaurant** | **Least Common restaurant** |
| **1** | **Restaurant** | **American Restaurant** | **French Restaurant** |
| **2** | **Asian Restaurant** | **Vietnamese Restaurant** | **Eastern European Restaurant** |

**Discussion:**

\* If XYZ company want to open a restaurant in preferred location and irrespective of cuisine, refer to that neighborhood in specific cluster and chose cuisine with the least common restaurant for better profits

\* If XYZ company want to open a restaurant with a preferred cuisine and irrespective of location, refer to the cluster with the least number of restaurants with that specific cuisine and select one among the neighborhoods based on company’s preference.

**Conclusion:**

This analysis is performed on limited data. This may be right or may be wrong. But if good amount of data is available there is scope to come up with better results. If there are lot of restaurants probably there is lot of demand. Toronto has so many restaurants, yet certain neighborhood or borough doesn’t have a specific cuisine restaurant available. As per the neighborhood or restaurant type mentioned like Indian Restaurant analysis can be checked. A venue with lowest risk and competition can be identified.