

# **The Battle of Neighborhoods | Business Proposal**

Author: Gawtam

## **1. Introduction**

Toronto is Canada's largest city and a world leader in such areas as business, finance, technology, entertainment and culture. Its large population of immigrants from all over the globe has also made Toronto one of the most multicultural cities in the world.

Its location on the northwestern shore of Lake Ontario, which forms part of the border between Canada and the United States, and its access to Atlantic shipping via the St. Lawrence Seaway and to major U.S. industrial centers via the Great Lakes have enabled Toronto to become an important international trading center. Moreover, the city is positioned on the edge of some of the best farmland in Canada, with a climate favorable to growing a wide range of crops, thereby making Toronto a transportation, distribution, and manufacturing center. Most importantly, its central location, along with a host of political policies favoring international trade, places this city with the greatest economic ties to, and influence from, the United States. Addition to these economic factors, Toronto is also known as popular destination for tourists. All of the factors mentioned allure people from all over the world to move here and settle for better career prospects. Hence the cultural diversity in Toronto.

As we know, more the cultural diversity, more opportunities for food industry. Many successful food franchises from all over world every year look for new places to start their business by investing in new cities. If they are already successful at one place, they invest more in the same city by opening more franchises or branches. So many companies employ a data driven approach to achieve a market insight for opportunity in new cities. This will help them to understand the business environment and allow them to carefully choose strategies which reduces the overall risk and increased ROI.

## **2. Business problem Description/ Problem Which Tried to Solve:**

Our client is a famous Multi-cuisine Cook and a venture capitalist who owns multiple franchise company in South Asia, looking for an opportunity to open a restaurant in the best neighborhood in Toronto, Canada. We as a Data Analysis expert, help him to identify the best places for his investment.

From initial market analysis we found, to be successful in restaurant business location with following criteria matters the most:

1. Availability Raw food material or supplier.
2. Population of location.

3. Market competitors.
4. Type of Locality (residential area /business hub/etc.)
5. Per capita income of people and many more.

Our aim is to develop a report of analysis which allows our client to carefully analyze and make a suitable decision to expand his business in Toronto.

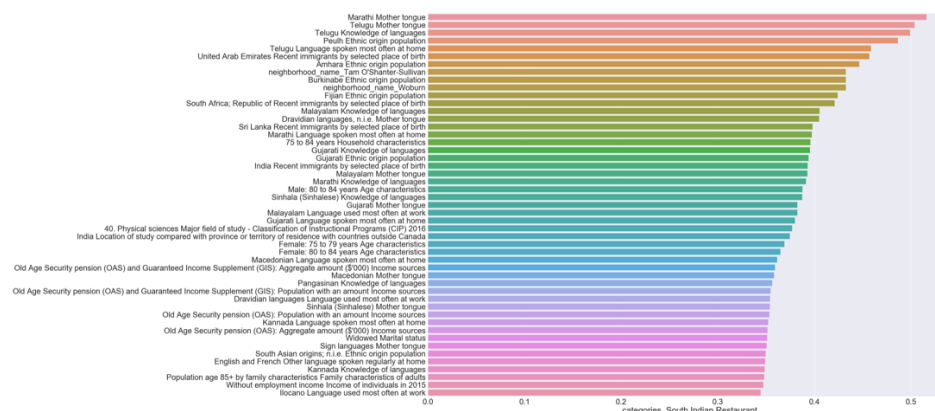
### 3. Data acquisition and cleaning

#### 3.1 Data sources:

Most demographic data for Toronto is available here at url: <https://www.toronto.ca/city-government/data-research-maps/open-data/> . We need to clean and mine the data for our purpose.

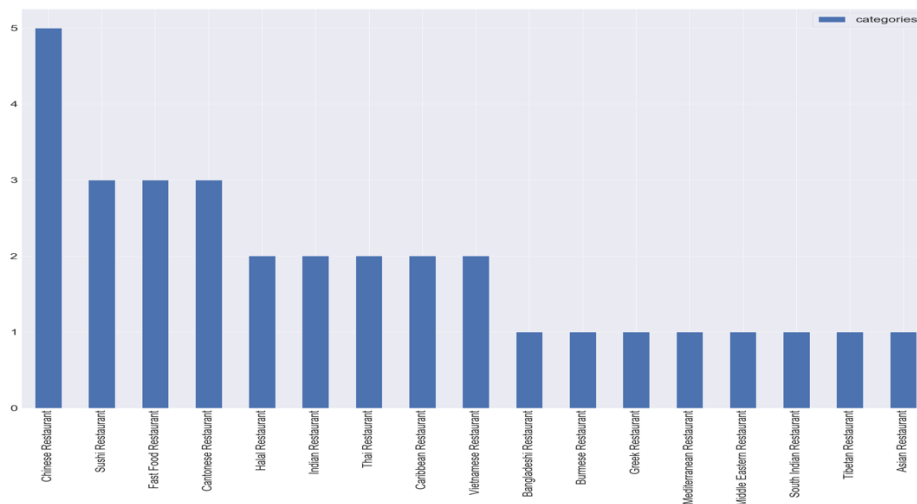
#### 3.2 Data cleaning

Data downloaded or scraped from above mentioned sources were turned into table using pandas in data form. There were a lot of missing values from earlier seasons, because of lack of record keeping. We decided use only relevant data out of all the available demographic data. We later normalized data for our use and fixed many errors like NaN values. After data cleaning, there were 140 samples and 13 features in the data. Upon examining the meaning of each feature, it was clear that there was some redundancy in the features. After discarding redundant features, we inspected the correlation between categories of restaurants and language spoken and found several pairs that were highly correlated. For example, in south Indian restaurants, the data shows majority of people who visit speak Telugu, Tamil and Marathi languages. These languages are widely spoken in southern parts of India, so our analysis shows proper correlation. The pictorial representation is as shown below.



## 4. Location

According to Wikipedia's List of neighborhoods in Toronto, there are several neighborhoods that correspond to different communities such as Little India, Little Italy, Little Portugal, Little Tibet, Greektown, Chinatown, Koreatown and Little Jamaica. We analyzed to see if the majority of restaurants from these cultures were found in their respective areas, so we searched for the perimeter for each neighborhood to compare with the restaurants on my map. The most number of Asian cuisine restaurants are Chinese, followed by Sushi, Halal and then Indian. The pictorial representation is given as below.



## 5. Methodology:

### Jupyter Notebook:

This project utilizes Jupyter Notebook for developing and employing all python codes as a interactive development environment.

### Foursquare API:

This project would use Four-square API as its prime data gathering source as it has a database of millions of places, especially their places API which provides the ability to perform location search, location sharing and details about a business.

Using credentials of Foursquare API features of near-by places of the neighborhoods would be mined. Due to http request limitations the number of places per neighborhood parameter would reasonably be set to 100 and the radius parameter would be set to 500.

### Clustering Approach using KMeans:

To compare the similarities of neighborhood, we decided segment them, and group them into clusters to find similar neighborhoods in the big city. To be able to do that, we need to cluster data which is a form of unsupervised machine learning: k-means clustering algorithm.

### **Python Libraries utilized:**

Pandas: For creating and manipulating data frames.

Folium: Python visualization library would be used to visualize the neighborhoods cluster distribution of using interactive leaflet map.

Scikit Learn: For importing k-means clustering.

JSON: Library to handle JSON files.

XML: To separate data from presentation and XML stores data in plain text format.

Geocoder: To retrieve Location Data.

Beautiful Soup and Requests: To scrap and library to handle http requests.

Matplotlib: Python Plotting Module.

Visualization: Tableau

## **6. Result and Recommendation**

KMeans returned 4 clusters and each cluster has its own characteristics:

- Cluster 1

average Number of Restaurant cluster 1: 32

average score cluster 1: 2.0777936138153117

Cluster 1 returned highest average number of restaurants. This gives us a fair assessment that setting up a restaurant here is little risky as more competition but since it is more populated area, we can consider this place.

- Cluster 2

average Number of Restaurant cluster 2: 1

average score cluster 2: 1.9942055781898618

Cluster 2 shows little concentration of restaurants in neighborhood which is again not in favor of setting up new restaurants would be risky.

- Cluster 4

average Number of Restaurant cluster 4: 7

average score cluster 4: 2.1107313407904496

In clusters made, 4th cluster has 7 average number of restaurants which not favorable for investment.

- Cluster 3

average Number of Restaurant cluster 3: 14

average score cluster 3: 1.9723093321551832

We recommend this cluster of neighborhoods as the number of restaurants is not high nor too low. This shows that the area may be under development and hence it is a recommendation.

Neighborhoods under this cluster are:

1. North St. James Town
2. Clairlea-Birchmount
3. Islington-City Centre West
4. Oakwood Village
5. Rockcliffe-Smythe
6. Weston
7. Mimico (includes Humber Bay Shores)
8. Brookhaven-Amesbury
9. Kingsview Village-The Westway
10. Kennedy Park
11. Niagara
12. Henry Farm
13. Scarborough Village
14. Humbermede

Below is the analysis performed using Tableau where all the Restaurants with category mapped on to Toronto Map. Figure 1 shows the all the different categories of restaurants and their location.

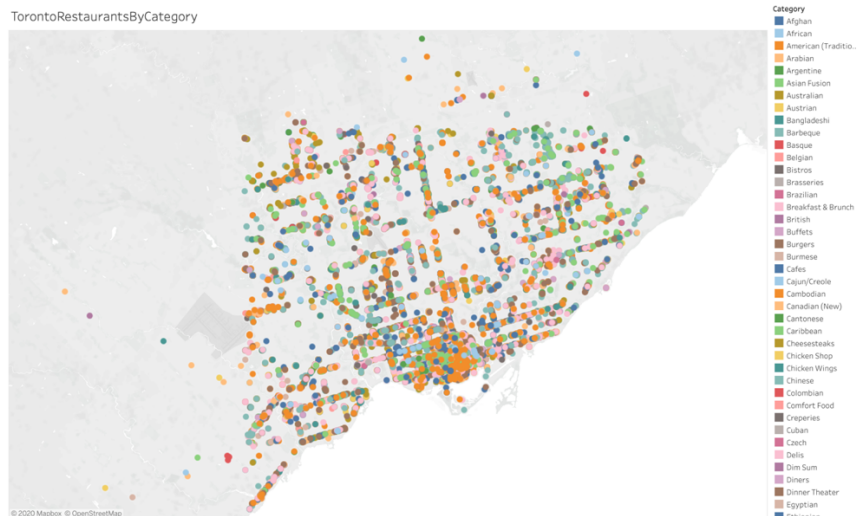


Figure 1

While Indian food is widespread throughout the city, there does appear to be a cluster of restaurants in the area See Toronto describes as Little India - on Gerrard Street East, between Coxwell Avenue and Greenwood Avenue. Figure 2 shows all Indian cuisine restaurants.

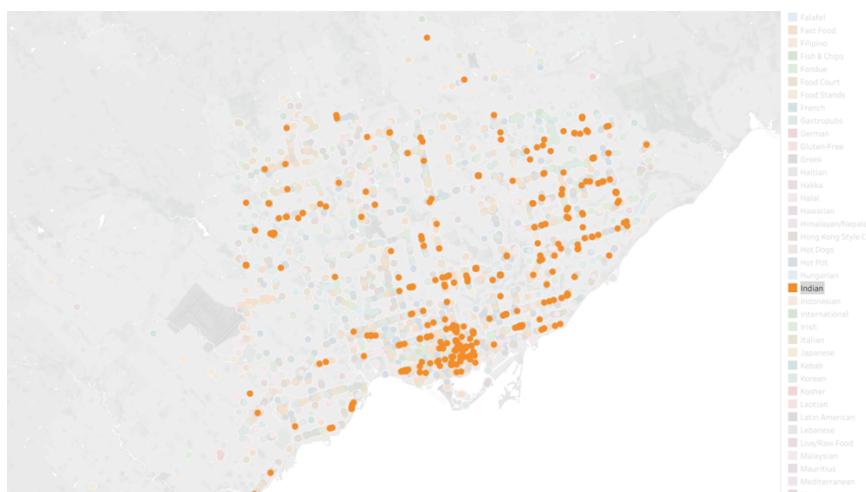


Figure 2

Similar to Indian food, Italian food is also widespread, but it is true that there is a small cluster in the area defined as Little Italy. The majority of Portuguese restaurants are mostly around Toronto's Little Portugal neighborhood and surrounding areas. There are not too many Tibetan restaurants in the city, but there is a cluster of quite a few at the far west of Queen Street West.

## Conclusion

Since the data and the analysis is not precise, our recommendations should be examined further. The limited data provided by foursquare API is a limitation for analysis and hence we believe, a

premium account will give is review of restaurants and facilities around. This will further ease the analysis and help us to recommend better place for investment. There is further scope for this project and analysis can be improved. Thank you.

--- \*\*\*\*\*---