IBM DATA SCIENCE CAPSTONE PROJECT

“The Battle Of The Neighborhoods”

Analyzing location for an Indian restaurant in Toronto, Canada

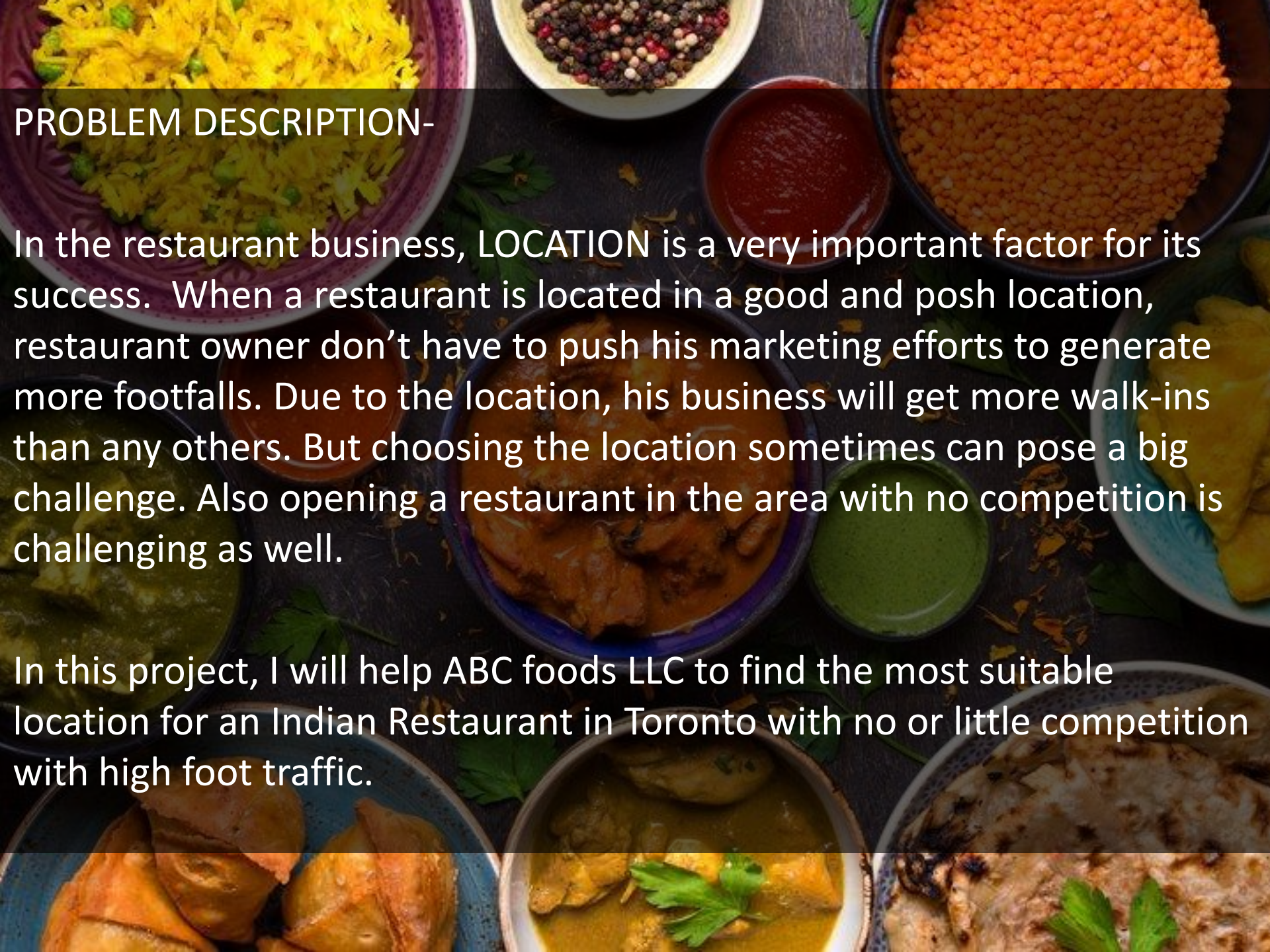
By Ruchir Palkar

INTRODUCTION

-ABC foods LLC wants to open an Indian restaurant in TORONTO city and need best location for same.

BUSINESS PROBLEM

- Choosing best neighborhood for starting new Indian restaurant
- No to Low competition
- High Demand



PROBLEM DESCRIPTION-

In the restaurant business, LOCATION is a very important factor for its success. When a restaurant is located in a good and posh location, restaurant owner don’t have to push his marketing efforts to generate more footfalls. Due to the location, his business will get more walk-ins than any others. But choosing the location sometimes can pose a big challenge. Also opening a restaurant in the area with no competition is challenging as well.

In this project, I will help ABC foods LLC to find the most suitable location for an Indian Restaurant in Toronto with no or little competition with high foot traffic.

Data-

To solve this problem, I used following data:

1. List of neighborhoods in Toronto, Canada.
2. Latitude and Longitude of these neighborhoods.
3. Indian restaurants Venue data.
4. Data of Hotels & Restaurants in each neighborhood.
5. Data of Venues in each neighborhood

Sources of data

1. Wikipedia page for neighborhoods in Toronto city (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
1. Geocoder package for latitude and longitude coordinates
2. Foursquare API for venue data

Methodology-

- Web scraping Wikipedia page for neighborhoods list
- Get latitude and longitude coordinates using Geocoder
- Use Foursquare API to get venue data
- Group data by neighborhood and taking the mean of the frequency of occurrence of each venue category
- Filter venue category by Indian Restaurant
- Perform clustering on the data by using k-means clustering
- Visualize the clusters in a map using Folium

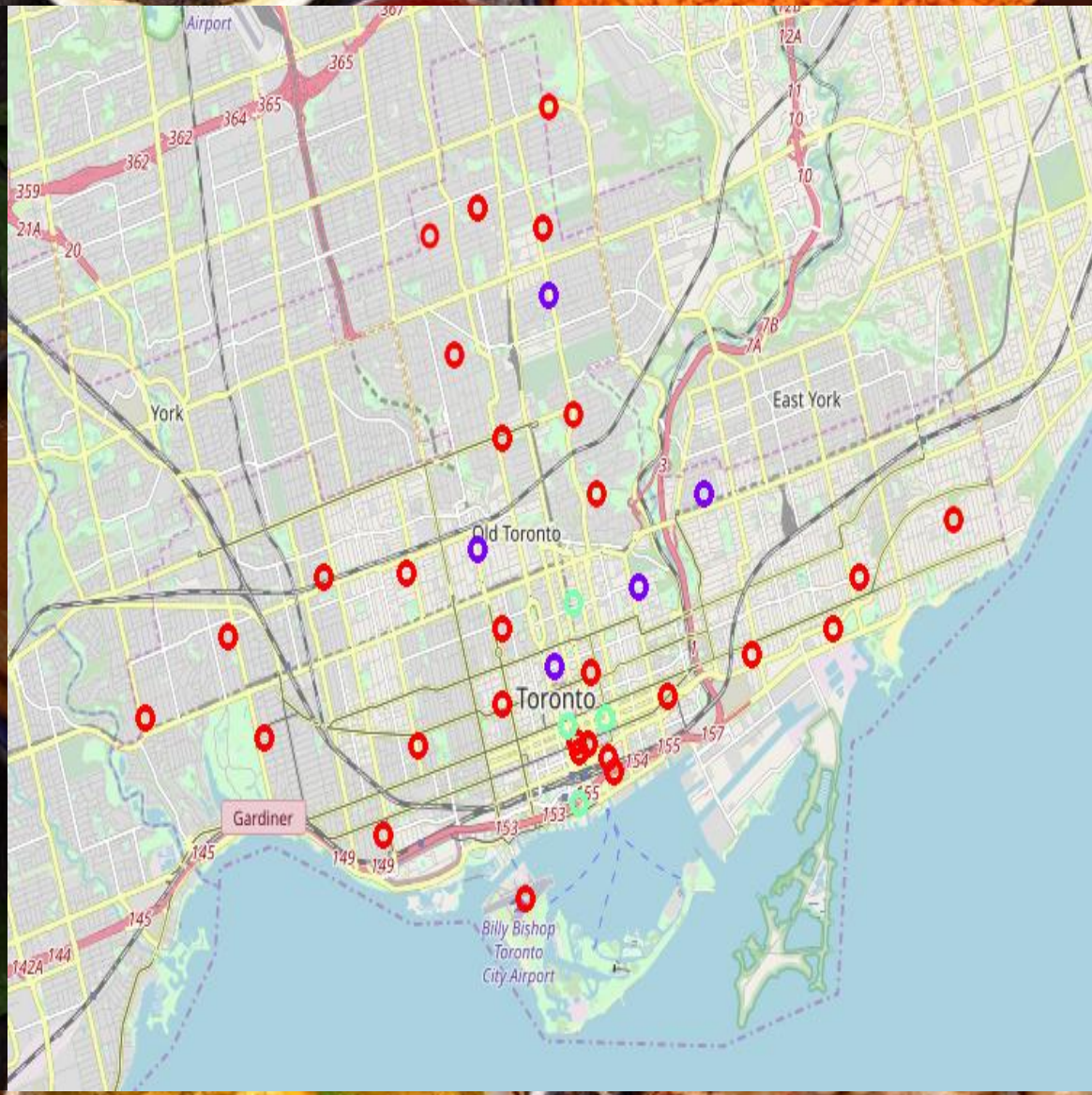
Result-

- **Categorized the neighborhoods into 3 clusters :**

➢Cluster 0: Neighborhoods with no Indian restaurant (Red)

➢Cluster 1: Neighborhoods with low number of Indian restaurants. (Purple)

➢Cluster 2: Neighborhoods with high number of Indian restaurants. (Light Green)



Result – PART 2

Cluster 0 has neighborhoods with no Indian restaurant. Since it has lots of neighborhood, I did further analysis for finding best neighborhoods for starting an Indian restaurant.

First of all, I calculated total number of hotels & restaurants in each neighborhood of Cluster 0 and arranged in descending order.

Venue Type	Hotel	Restaurant	Total H & R
Neighborhood			
Chinatown,Grange Park,Kensington Market	1	34	35
Commerce Court,Victoria Hotel	7	27	34
First Canadian Place,Underground city	5	27	32
Design Exchange,Toronto Dominion Centre	7	22	29
Stn A PO Boxes 25 The Esplanade	3	23	26

This analysis gave neighborhoods which are famous for restaurants & hotels in Cluster 0.

Result- PART 3

More venues in neighborhood mean more foot traffic.

For this part, I calculated total venues in each neighborhood of Cluster 0 and merged with total restaurants & hotels table.

The following are results –

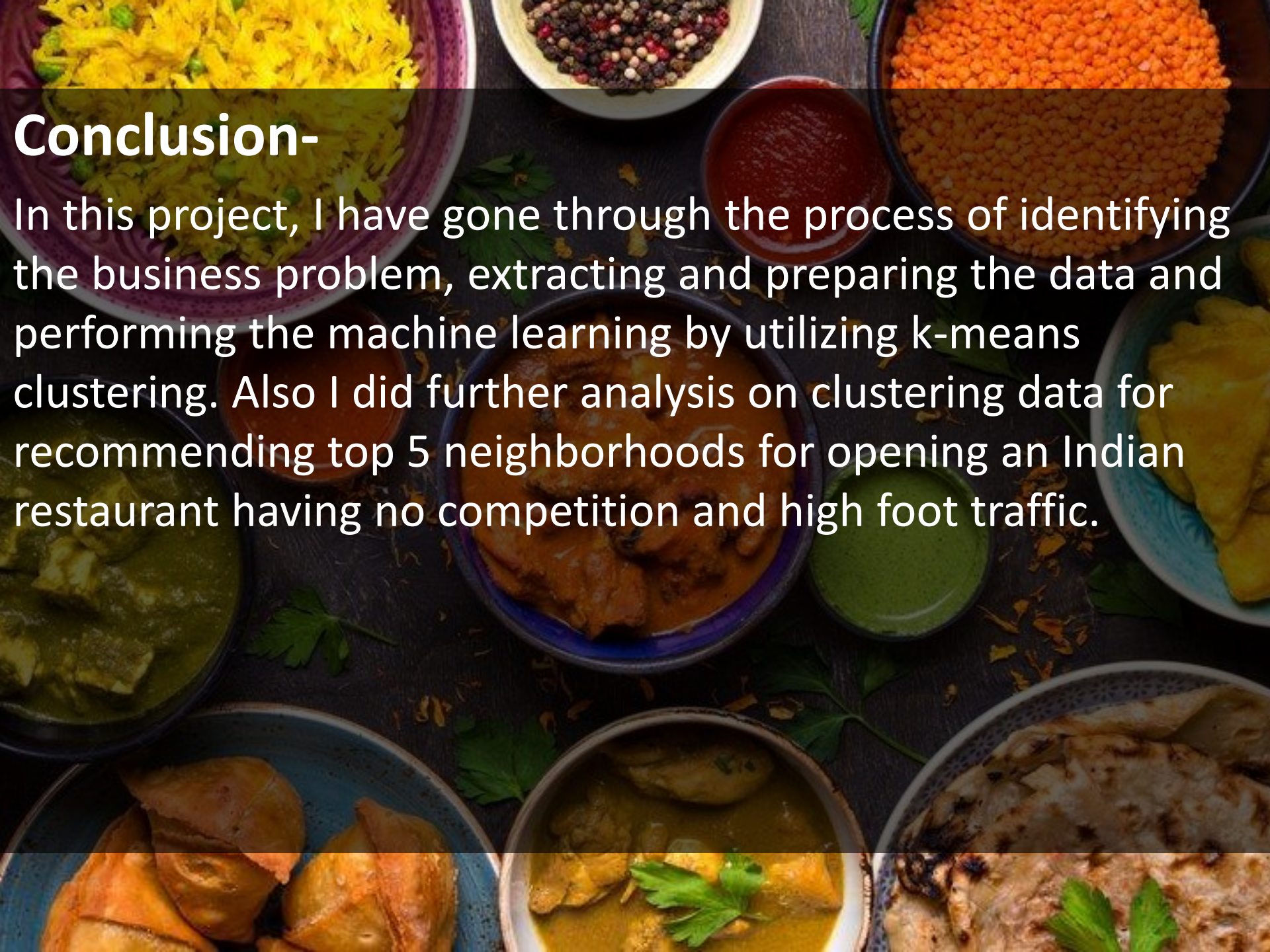
	Hotel	Restaurant	Total H & R	Total Venues
Neighborhood				
Chinatown,Grange Park,Kensington Market	1	34	35	100
Commerce Court,Victoria Hotel	7	27	34	100
First Canadian Place,Underground city	5	27	32	100
Design Exchange,Toronto Dominion Centre	7	22	29	100
Stn A PO Boxes 25 The Esplanade	3	23	26	96

These are top 5 neighborhoods that I can suggest my client for opening an Indian restaurant. Since,I am using free version of Foursquare API,I am getting maximum 100 results only.Paid version will give better results

Discussion-

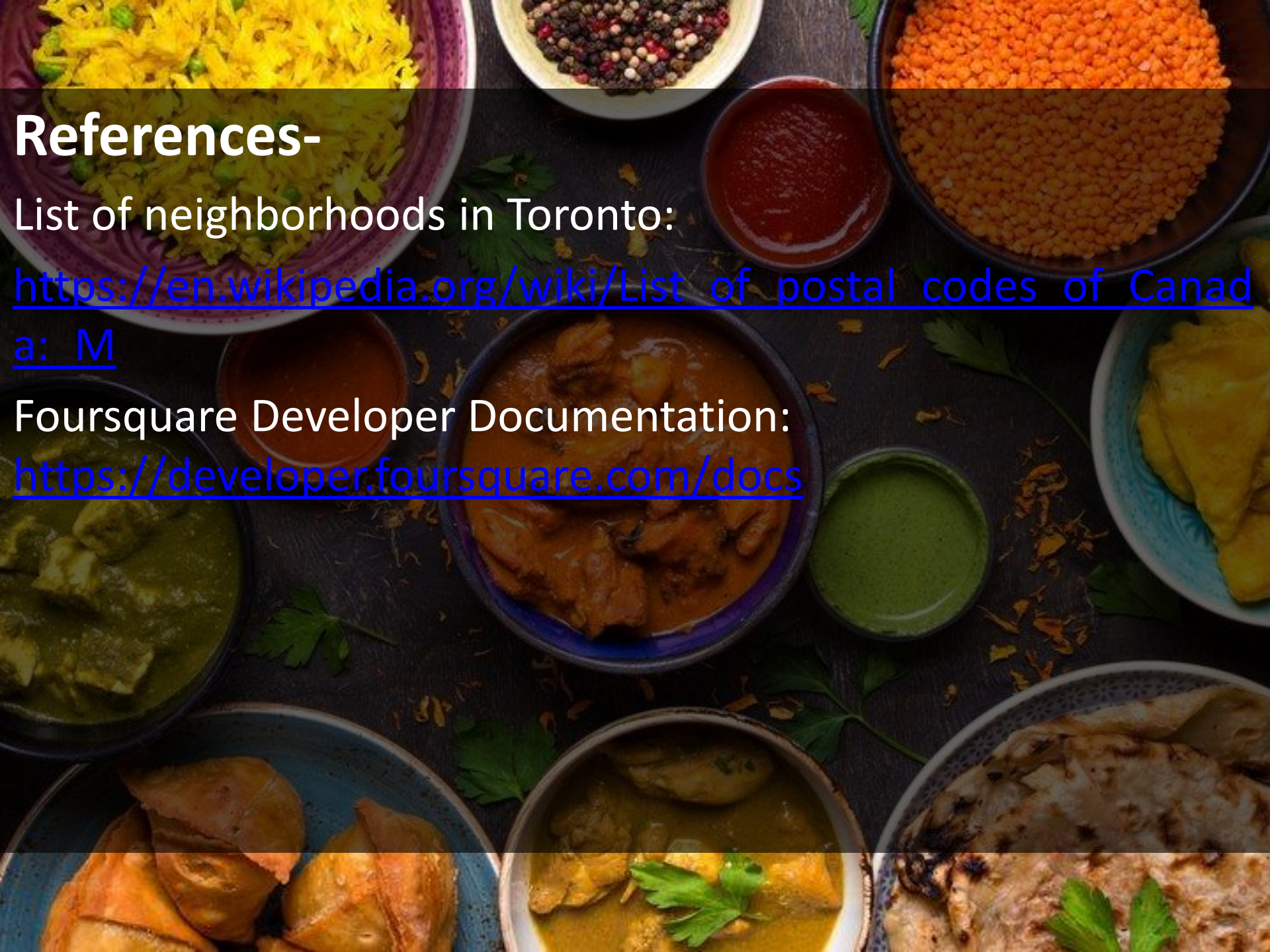
In this project, I took only few considerations for finding ideal location for an Indian restaurant. There are other factors as well that can help to find better location like real estate prices, population density,etc. These factors would have definitely improved my analysis.

However, this project gives a nice start to the process and narrowed down a very long list (from 103 to 5 choices)



Conclusion-

In this project, I have gone through the process of identifying the business problem, extracting and preparing the data and performing the machine learning by utilizing k-means clustering. Also I did further analysis on clustering data for recommending top 5 neighborhoods for opening an Indian restaurant having no competition and high foot traffic.



References-

List of neighborhoods in Toronto:

https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

Foursquare Developer Documentation:

https://developer.foursquare.com/docs