

VINIL M. THOMBREY

IBM DATA SCIENCE SPECIALIZATION - CAPSTONE

---

**INDIAN RESTAURANT IN  
TORONTO, CANADA**

PART 1

---

# INTRODUCTION

## INTRODUCTION

---

A successful Businessman with more than 50 restaurants in India is planning to expand his business in Toronto, Canada considering there are lot of Asians (Indians) staying there. There are already lot of Indian restaurants in Canada, especially in Toronto and he is not sure where exactly to open. He needs an analysis of the neighborhoods in Toronto and want to understand the type of population in the neighborhood. Thus, as a Data Scientist, I am helping him in analyzing all the neighborhoods in Canada and identify the best location to open an Indian restaurant.

PART 2

---

# BUSINESS PROBLEM

## BUSINESS PROBLEM

---

The objective of this project is to identify the best location to open an Indian restaurant in Toronto, Canada. Thus, with the help of data science methods and machine learning techniques, the project aims in finding the most suitable location to open an Indian restaurant by considering the type of already existing restaurants, type of locality and demographics of that area.

PART 3

---

# DATA EXTRACTION

## DATA SOURCES

---

- Get the neighborhood data from the Wikipedia page for Toronto for Latitude and Longitudes of each neighborhood in Json format and use Python to clean and convert it into the pandas dataframe.
- Combine this data with the venues related data for Indian restaurants using the Foursquare API which was introduced in this course.
- Use **Beautiful Soup** Library for reading HTML file, **Foursquare** API for getting lat/long info about the nearby venues and **Folium** for Geo spatial visualization

PART 4

---

# METHODOLOGY



- ▶ Data Extraction and Processing
- ▶ Extracting Venues data using Four Square API
- ▶ Analyzing the Data using Cluster Analysis
- ▶ Using Folium library to Visualize the Data
- ▶ Infer conclusions from the final map

PART 5

---

RESULT

# OUTPUT OF CLUSTER ANALYSIS USING FOLIUM

The geographical coordinates of Toronto are 43.667262184210514, -79.38988323421053



## OBSERVATIONS

---

There are lot of Indian restaurants in Toronto but after the cluster analysis, we see that there are no Indian restaurants in cluster 0 while there are many of them in cluster 1 and cluster 2. Cluster 0 consists of areas like Dominion Centre, St. James street, North Toronto, Davisville and thus these areas lack Indian restaurants. There are many Asian restaurants in the area which serves India/Chinese/Thai and Malaysian food as well, but lacks authentic Indian restaurant. Thus, there is a scope to have a restaurant in these areas.

PART 8

---

# FUTURE WORK

## FUTURE WORK

---

We can have further analysis about the demographics of that region which will help us in making this decision quicker. Understanding the average age in the area, income and the ethnicity can add value to our analysis but the data for demographics is not publicly available.