# IBM DATA SCIENCE

WEEK-5 FINAL PRESENTATION

# INTRODUCTION

- ► This project is based on a hypothetical scenario where there may not be enough Indian Restaurants in Toronto Area.
- As Indian food is popular among the Asian community, we are thinking of opening this business in areas where the Asian community resides.
- Finding a location to open a restaurant is one of the most important factors.

#### BUSINESS PROBLEM

- The objective of this capstone project is to find the most suitable location for the entrepreneur to open a new Indian Restaurant in Toronto, Canada.
- We are using data science methods and tools along with machine learning algorithms such as clustering.
- Business Question: In Toronto, if an entrepreneur wants to open an Indian Restaurant, where should they consider opening it?

# DATA

- List of neighbourhoods in Toronto, Canada.
- ► Latitude and Longitude of these neighbourhoods.
- Venue data related to Indian restaurants. This will help us find neighbourhoods that are more suitable to open an Indian Restaurant.

#### EXTRACTING THE DATA

- ▶ The scrapping of Toronto neighbourhoods via Wikipedia.
- Getting Latitude and Longitude data of these neighbourhoods via Geocoder package.
- Using Foursquare API to get venue data related to these neighbourhoods.

#### METHODOLOGY

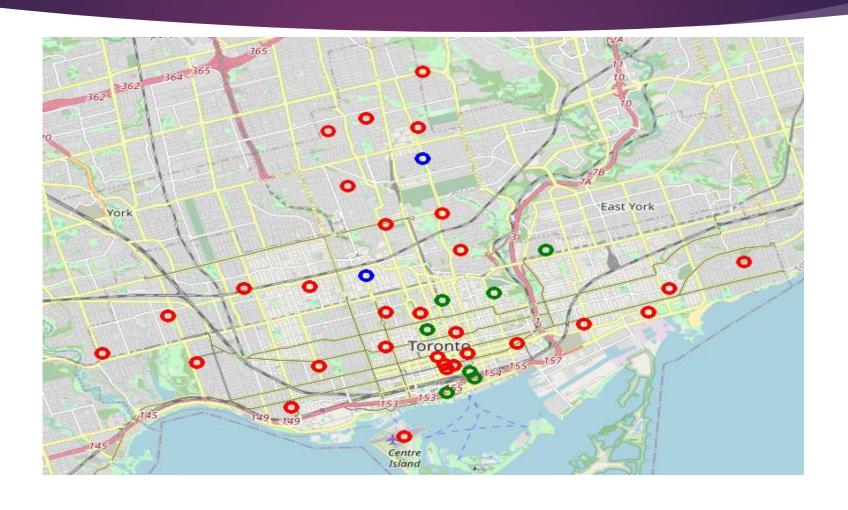
- ▶ Get the list of neighbourhoods in Toronto, Canada.
- ▶ Web scraping by utilizing pandas HTML table scraping method.
- Utilize Foursquare to pull the list of venues near these neighbourhoods.
- ▶ Use Foursquare API to pull the list of top 100 venues within 500 meters radius.

#### METHODOLOGY

Analyse each neighbourhood by grouping the rows by neighbourhood and taking the mean on the frequency of occurrence of each venue category.

- Making a justification to specifically look for "Indian restaurants".
- Performed the clustering method by using k-means clustering.

# RESULT



# RESULT

- ► Cluster 0: Neighbourhoods with no Indian restaurants.
- Cluster 1: Neighbourhoods with the smaller number of Indian restaurants.
- Cluster 2: Neighbourhoods with the highest number of Indian restaurants.
- ▶ The results are visualized in the above map with Cluster 0 in red, Cluster 1 in green, Cluster 2 in blue.

#### RECOMMENDATION

- ► The least number of restaurants lie in cluster 0 areas which are North Toronto West, Parkade areas, St James Town, Cabbagetown etc.
- Looking at nearby venues it seems cluster 0 might be a good location as there are not a lot of Indian restaurants in these areas.
- ► Therefore, this project recommends the entrepreneur to open an authentic Indian restaurant in these locations.