

**IBM Data Science****Capstone project****1. Introduction**

This project, I am creating a hypothetical scenario for a concept that there may not be enough Indian Restaurants in Toronto Area. Therefore, it might be a great opportunity for an entrepreneur who is based in Canada. As the Indian food is popular among Asian community, so this entrepreneur might think of opening its business in areas where asian community resides.

With the purpose in mind, finding the location to open such a restaurant is one of the most important decisions for this entrepreneur and I am designing this project to help him find the most suitable location.

**2. Objective**

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. By using data science methods and tools along with machine learning algorithms such as clustering, this project aims to provide solutions to answer the business question: *In Toronto, if an entrepreneur wants to open an Indian Restaurant, where should they consider opening it?*

**3. Data**

I will use explanatory data analysis to explore my data.

To solve this problem, we will need below data:

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

**4. Overall process**

Scrapping of Toronto neighborhoods via Wikipedia

- Getting Latitude and Longitude data of these neighborhoods via Geocoder package
- Using Foursquare API to get venue data related to these neighborhoods