

Coursera IBM Data Science Capstone Project

Analyzing location for an Indian restaurant in Toronto, Canada

**Prepared by:
Ruchir Palkar**

1.Introduction & Business Problem :

Problem Background:

Toronto is the densely populated city in Canada. It provides a lot of business opportunities and a business-friendly environment. Toronto is an international center of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world. Its economy is highly diversified with strengths in technology, design, financial services, life sciences, education, arts, fashion, business services, environmental innovation, food services, and tourism.

Toronto has attracted big business players from all around the world. It means the Toronto market is highly competitive. Also, the cost of doing business in such a market is high. Hence, any business venture or expansion must be analyzed very carefully. The insights derived from the analysis will give a good understanding of the business environment which help in strategically targeting the market. It also reduces risk and helps in getting a better Return on Investment.

Problem Description

A restaurant is a business establishment which prepares and serves food and drink to customers in return for money. Toronto is famous for its excellent cuisine. Its food culture includes an array of international cuisines influenced by the city's immigrant history.

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.

Hence, 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.

Target Audience

Anyone who wants to open a restaurant in Toronto, Canada.

2.Data Section

To solve this problem, I will be using the following data:

- List of neighborhoods in Toronto, Canada.
- Latitude and Longitude of these neighborhoods.
- Indian restaurants Venue data.
- Data of Hotels & Restaurants in each neighborhood.
- Data of Venues in each neighborhood

Sources of data and methods to extract them

This Wikipedia page ("https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M") contains a list of neighborhoods in Toronto, with a total of 130 neighborhoods. I will be using web scraping techniques to extract the data from the Wikipedia page, with the help of Python requests and BeautifulSoup packages. Then I will get the geographical coordinates of the neighborhoods using Python Geocoder package which will give us the latitude and longitude coordinates of the neighborhoods.

After that, we will use the Foursquare API to get the venue data for those neighborhoods. Foursquare has one of the largest database of 105+ million places and is used by over 125,000 developers. Foursquare API will provide many categories of the venue data, we are particularly interested in the Restaurant category to solve the business problem. In this project, I will use many data science skills like web scraping (Wikipedia), working with API (Foursquare), data cleaning, data wrangling, machine learning (K-means clustering) and map visualization (Folium).