BATTLE OF NEIGHBOURHOODS

CAPSTONE PROJECT

IBM APPLIED DATA SCIENCE CAPSTONE

Opening an Indian Restaurant in the suburbs near Mumbai, India

Pradnyesh Choudhari

April 2020

Introduction

For many people, these days dining in restaurants has become very common. They can enjoy themselves and the trends have completely changed, the restaurants get crowded not only during holidays or weekends but remain crowded over the complete year. For Indians, it is very common to prefer Indian Cuisine than any other more occasionally. Suburbs near Mumbai include all those places which surround Mumbai which has undergone a lot of development and is becoming center of attraction for many brands and major developers worldwide. These places are a big and a never-ending market for all those who are interested in providing services for the same. For opening a new restaurant considering a location that has less competition and for the developers, they have to consider even the other factors such as rent and all. Particularly, the location of the restaurant is a major reason for the success and failure of the project.

Business Statement

The objective of the capstone project is to analyze the neighborhoods using the techniques of data science and a few machine learning techniques such as clustering, to select the best location for opening an Indian restaurant in the suburbs surrounding to Mumbai, India.

If a restaurant chain wants to open a new restaurant in the suburbs near Mumbai where should you open it?

Data

- List of the Suburbs near Mumbai, India. The list is the most important as it defines the scope of the project and the area to which it is confined.
- Latitude and Longitude coordinates of these cities. These are required to plot the map and using the Foursquare API to get the data for the different surrounding places and the data related to the venue.

The Wikipedia page (https://en.wikipedia.org/wiki/Category:Suburbs_of_Mumbai) contains the data of a complete list of suburbs near Mumbai. We scrap the data from the Wikipedia page where using the **beautifulsoup** library package of python. The **python geocoder** package gives the latitude and longitude data of these places. Further using the **Foursquare API** we get the venue data of the surrounding. **Folium** library is used to plot the map using the location data and finally after data cleaning using the machine learning algorithm **k-means clustering** we segment the data and cluster it into groups. After the clusters are formed, we get a group of places where we can open the restaurant and the problem statement with which we started this project gets fulfilled.