Peer-graded Assignment: Capstone Project - The Battle of Neighborhoods (Week 1)

By Prahlad 19 May 2020

Part-I

1. Problem Definition

Objective – Select a suitable location for Indian Multi-Cuisine restaurant in Delhi, the capital of India

1.1 Introduction

This project is sponsored by one of the big restaurant chains from Western India. The sponsors are interested in opening a multi-cuisine Indian restaurant in Delhi. Accordingly, this report has been prepared to evaluate the best location to open a new restaurant in one of the Delhi's localities. The restaurant should be in the area with good footfall and high business potential. . This report will use the Foursquare API in order to analyse the data about the locations of existing restaurants in various venues and find out about trending venues and neighbourhoods that might suggest popular areas to open a new restaurant.

With its diverse culture, come diverse food items. There are many restaurants in Delhi City, each belonging to different categories like - Thai, Intercontinental, Chinese , Italian , French and Indian restaurants – Punjabi, South Indian, Kashmiri, Gujrati , Rajasthani etc. All these have a unique flavour and choice of foods.

1.2 Brief on Delhi City

Delhi is the capital city of India. The population of Delhi is around 19 million. The town is divided into nine districts. It is a historical town. The area around Delhi was probably inhabited before the second millennium BCE and there is evidence of continuous inhabitation since at least the 6th century BCE. The city is believed to be the site of Indraprastha, the legendary capital of the Pandavas in the Indian epic *Mahabharata*. The earliest architectural relics date back to the Mauryan period (c. 300 BCE).

Delhi is a favourite destination for domestic and foreign tourists ((25 million domestic tourists and 5 million international visitors per year) .

Delhi and New Delhi are used interchangeably and referred as **NCT** - National Capital Territory. These are two distinct entities, with New Delhi forming a small part of Delhi. The **NCR**- National Capital Region is a much larger entity comprising the entire NCT along with adjoining districts in neighbouring states of UP, Haryana and Rajasthan.

1.3 Problem Definition

To suggest the best location for a new multi-cuisine Indian restaurant in Delhi by using the descriptive and predictive capability of data science and geospatial data analysis with the help of Foursquare API. Following questions will be answered:

a) Analysis of trending venues across neighbourhoods – the number of existing restaurants

b) Areas have large number of Multi-Cuisine Indian restaurants

c) Areas which have less number of restaurants and where demand for new multi-cuisine Indian restaurants exists

d) Places which have best restaurant in New Delhi

e) Finally finding the best location in Delhi City for Multi-Cuisine Indian restaurant.

1.4 Target Audience

One of the big food chains which owns many multi-cuisine restaurants across India. They have the budget and have entrusted me to submit a report based on the research and analysis of the relevant data.

Part-II

2 Data Description

Based on the preliminary investigation following data will be collected for analysis:

i) Web scrap the Delhi Wikipedia page to get the districts and Neighborhood

https://en.wikipedia.org/wiki/List_of_districts_of_Delhi
Geocode the data by using the library geopy.geocoder (Nominatim)

ii) Data from Kaggle about all existing restaurants working with zomato
https://www.kaggle.com/shaswatd673/delhi-neighborhood-data
https://www.kaggle.com/shrutimehta/zomato-restaurants-data

iii) Foursquare API to get all type of restaurants with 1000 m radius for every neighborhood

www.foursquare.com

2.1 Approach

- Collect the Delhi city data from Wikipedia and Kaggle
- Use Foursquare API get some geospatial information about different neighborhoods, group neighborhoods in clusters and eventually combine the results to reach to our conclusion.
- Onehotencoding – To convert strings as labels for each neighborhood and to digitizing them so that we can use them in our classification algorithm.
- Clustering neighborhoods by using k-means algorithm
- Eliminating over crowded sites i.e. sites having many all type of restaurants as well as Indian restaurants
- Finally select the new site based on popularity of neighborhood