

Segmenting and Clustering Neighborhoods of Mumbai City to find Optimal Location for Opening an Indian Restaurant

Maitri A. Prajapati Final project

1. Introduction

1.1 Background Information:

Mumbai, formerly known as Bombay, is the capital city of Maharashtra. The Greater Mumbai area occupies a long, narrow peninsula in the Arabian Sea on the west coast of India. Mumbai is the most populous city in India and in the top five in the world. While the 2011 census estimated population at 12.4 million. Mumbai is the financial center, economic powerhouse, and industrial hub of India. In 2018, a report by New World Wealth ranked the city as the 12th wealthiest global city with a total private wealth of US Dollar 950 billion.

The Mumbai city is one of the world's top centers of commerce in terms of financial flow. It is also home to important financial institutions, such as the Reserve Bank of India, the Bombay Stock Exchange, the National Stock Exchange of India, and corporate headquarters of many Indian companies and multinational corporations.

1.2 Discription of problem:

Mumbai is rich in its Indian cuisine. It has huge variety of Indian foods from different parts of India. Mumbai seems to be a better place for setting up an Indian restaurant. Since there are lots of restaurants in Mumbai we will try to detect locations that are not already crowded with restaurants. We are also particularly interested in areas with no Indian restaurants in vicinity.

1.3 Target Audience:

Companies or Individuals looking into opening a restaurant would be interested in prediction of optimal location in Mumbai City. This project will provide an analysis whether the venture is feasible or not.

2. Data acquisition and cleaning

2.1- Data Sources

First, I search the information of Mumbai neighbourhoods from the various web information sources. Finally, I successfully search the information from following link of Wikipedia.

https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai

As per the information given on this web page, Mumbai has a total of 31 boroughs and 92 neighbourhoods. This web page contains one table, in which information of Mumbai neighbourhoods, location, latitude and longitude are given. I consider location as a borough of particular neighbourhood. I

converted the html table from the website to a pandas dataframe using webscrapping methods of BeautifulSoup package.

Based on definition of our problem, factors that will influence our decision is the number of existing indian restaurants in the neighbourhood. Number of restaurants and their type and location in every neighbourhood will be obtained using Foursquare API .

2.2- Data Cleaning

Data downloaded from wikipedia website contains Borough, their Neighbourhoods and latitude and longitude coordinates of each neighbourhood. The details of data cleaning methods are given below.

1. The names of some neighborhoods/borough were found wrong and therefore, such names were corrected.
2. I found some wrong coordinates and therefore I verified this information using geopy package of python. Most of the coordinates received using above python package differed from the coordinates given in the website. Therefore, I replaced the coordinates of the dataframe.

Finally, I got the dataframe shown below:

```
In [87]: Mumbai= Mumbai.assign(Latitude=latitude, Longitude=longitude)
print(Mumbai.shape)
Mumbai.head()

(93, 4)
```

Out[87]:

	Neighbourhood	Borough	Latitude	Longitude
0	Amboli	Andheri,Western Suburbs	19.07599	72.877393
1	Chakala	Andheri,Western Suburbs	19.07599	72.877393
2	D.N. Nagar	Andheri,Western Suburbs	19.07599	72.877393
3	Four Bungalows	Andheri,Western Suburbs	19.07599	72.877393
4	Lokhandwala	Andheri,Western Suburbs	19.07599	72.877393

Next week work plan

We get the Initial Data-Frame with Names of Boroughs, and corresponding neighbourhoods and the coordinates of those neighbourhoods. Before comparing all the neighbourhoods, since we want to concentrate only on Indian restaurants, we need to get the idea about the best areas for Indian Restaurants in Mumbai. So as the next step we will use Foursquare data and obtain information on restaurants. With these, we can start with our battle of neighborhoods for finding optimal location to start Indian restaurant in Mumbai.