**Location for opening a Restaurant in Mumbai**

Tanmay Sharma

17th October,2019

1. **INTRODUCTION**
   1. **Background**

Mumbai is the central capital of India. Every person in Mumbai strives hard each day to achieve their goals in life. People from other parts of the country also visit Mumbai in order to understand the culture of India. India is one of the most multicultural country in the world. Moreover, food is a must try for each individual who visits India.

* 1. **Problem Statement**

While opening a restaurant at a particular place, what are the important traits and the best possible location which are to be kept in mind will be evaluated from this project.

1. **DATA ACQUISITION AND CLEANING**
   1. **Data sources**

The data which will be used for this project is an excel file created from listings of various restaurants in Mumbai on Zomato. The following link provides the final excel file to be used while creating this project - [Link](https://www.kaggle.com/srivpuneet16/zomato-mumbai-restaurant-analysis/version/1%23zomato_res_final.csv)

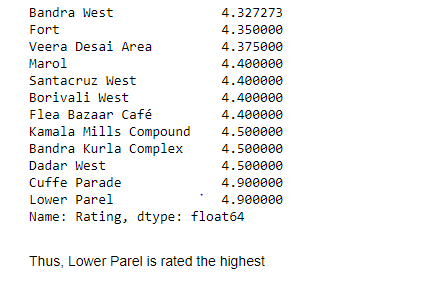
* 1. **Data Cleaning**

All the restaurants with duplicate names were dropped. The biggest problem which we faced while using such a huge dataset was that the number of calls via Foursquare were limited in nature. Thus, the original dataset comprised of 6181 restaurants, ie 6181 tuples in the excel sheet. Thus, importing such a huge dataset would not be possible by using the general ‘Sandbox’ account in Foursquare. Since I did not want to spend any money on increasing my calls on Foursquare, I manually selected the first 100 rows and dropped the rest of the rows from the dataset. However, all the commands will work incase we have some other kind of account in Foursquare.

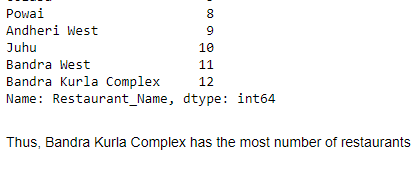


Upon analyzing the data, we found that restaurants in Lower Parel were ranked the highest. We also analyzed that the number of restaurants in Bandra Kurla Complex were the most – 12. These analysis were done by using basic python commands and there was no need of Foursquare.





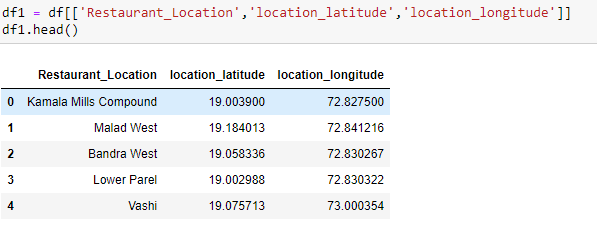




* 1. **Feature Selection**

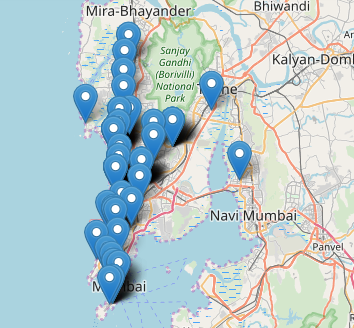
Once the data was cleaned and we were ready for analysis, we found the latitudes and longitudes of all the restaurants locations. Some of the locations for which location could not be retrieved were manually fit into the data. Once the latitudes and longitudes were determined, map of Mumbai along with all the restaurant locations were plotted using Folium.

Once the latitudes, longitudes and Restaurants locations were selected, Foursquare API was used to determine the neighborhood of all the places in Mumbai



1. **USING FOURSQUARE API**
   1. **Process**

Once the Client ID and Client Secret were provided, Foursquare API was used to determine top 100 venues in a 2000 meters radius. This call provided various aspects like Venue location, venue category, etc. Venue category was the main aspect as it specified whether a venue was considered as a restaurant or not.

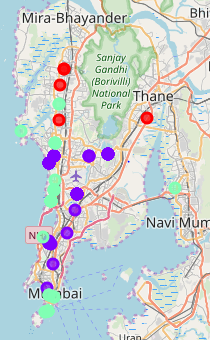


* 1. **Venue Category**

Once the venue category was analyzed, we could observe that there were a wide categories of restaurants like American Restaurant, Japanese Restaurant, Indian Restaurant, etc. Thus, taking all types of restaurants as a factor, we collectively grouped the data in terms of Restaurant location and its category. This helped us in fetching the data for all kinds of restaurants/cafes/breakfast spots/ etc into a single data frame and perform analysis on them.

1. **CLUSTERING**

Once the data was analyzed in terms of their type, all the data were stored into a cluster. We made 3 clusters, namely, cluster 0, cluster 1 and cluster 2. All of the restaurants locations were stored in one of the clusters. We made use of K means clustering to carry out the segregation of data. We also plotted all of the data onto a map to have a visualization of the entire cluster set using Folium.



1. **OBSERVATION**

We could observe from the clusters that cluster 0 has the least number of restaurants in Mumbai. Cluster 1 and cluster 2 have a large number of restaurants in the city, thus, a lot of competition will be faced if the restaurant is opened in either location of cluster 1 or cluster 2. Opening a restaurant in Borivali West, Kandivali West, Goregaon West, or Mahakali will be a safe bet as competition is very low in these areas. Moreover, these areas are not in the core locations of Mumbai like BKC, Kamala Mills Compound, Juhu, Powai, etc. Thus, opening a restaurant will be easier and cheaper for any individual. We can also observe that most of the restaurants are of Cafe or Indian restaurant type. Thus, opening a restaurant like American Restaurant or Japanese Restaurant will be more beneficial as it will provide diversity to the area. Thus, opening a restaurant in some other type than Cafe or Indian Restaurant will be more beneficial for any individual