Analysing best location for opening a restaurant in Mumbai, India

Capstone Project
IBM Applied Data Science Certificate Program

Data Description

Kumar Lokesh Singh

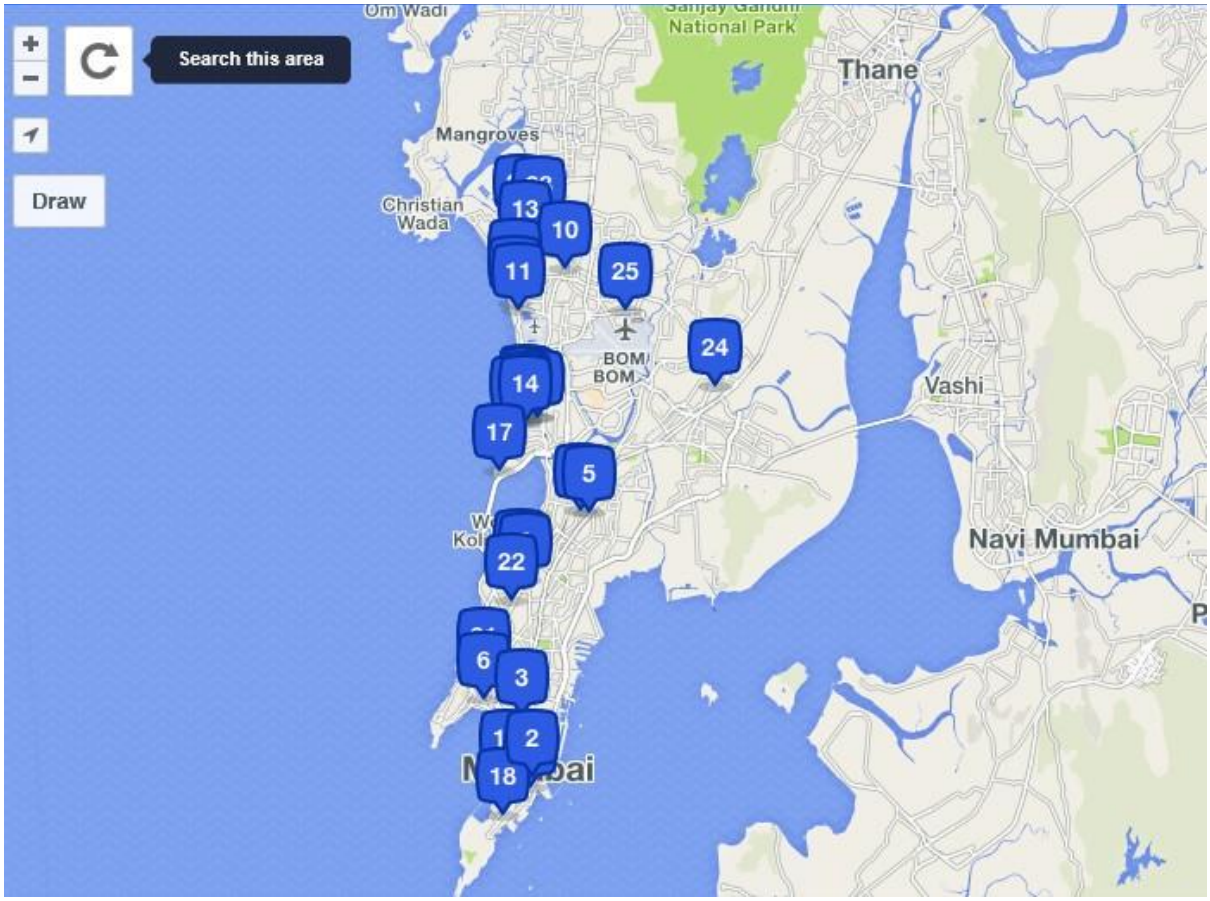
Introduction

The main data used for this project will be from two sources:

- The restaurants data from the Zomato website was obtained from Kaggle website.
- Explore trending venues in a neighbourhood particularly restaurant (Foursquare API).

Other supporting data:

- Coordinates (Geocoder Python)
- https://foursquare.com/explore?mode=url&near=Mumbai%2C%20Mah%C4%81r%C4%81shtra%2C%20India&nearGeoId=72057594039203275&q=Restaurant
- https://www.zomato.com/mumbai



Data Collection Process

The data obtained from Kaggle website for Zomato restaurants was processed in below listed phases.

Phase I

In Phase I involves cleaning of dataset in CSV file and checking each field for different data types. The dataset was optimised to fit as per our project requirement.

Phase II - Foursquare API

For each neighbourhood, Geopy module to convert an address into latitude and longitude values. For each neighbourhood's coordinate, we will call Foursquare API to get the trending venues in that location.

Using data to solve the problem

The basic idea of analysing the Zomato dataset is to get a fair idea about the factors affecting the establishment of different types of restaurant at different places in Mumbai, aggregate rating of each restaurant, Mumbai being one such city has more than 13,790 restaurants with restaurants serving dishes from all over the world.
With each day new restaurants opening the industry is yet to saturate and the demand is increasing regularly. Despite increasing demand, it however has become difficult for new restaurants to compete with established restaurants. Most of them serving similar food and most of the people are dependent on the restaurant food. With such an overwhelming demand of restaurants it has therefore become important to study the demography of a location. What kind of a food is more popular in a locality? Do the entire locality loves vegetarian food etc. This kind of analysis can be done using the data, by studying the factors such as

- Approx. price of food
- Location of the restaurant
- Theme based restaurant or not?
- Which locality of that city serves that cuisines with maximum number of restaurants?
- Customers who are striving to get the best cuisine of the neighbourhood
- Is a neighbourhood famous for its own kind of food?