Coursera Capstone Project

Battle of Neighbourhoods Week 1

Description & Discussion of the Background

Bangalore is one of the fast-growing cities in India where over 12 million people live and it has a population density of 4381 people per square kilo meter. As a resident of this city, I decided to use Bangalore in my project. People from different regions of India are thickly populated in Bangalore, especially from northern and southern regions of India. Due to is diversity in culture, the city needs to deal with diverse food items. There are many restaurants in Bangalore, each belonging to different categories like Northern, Southern, Chinese, Andhra, Kerala etc. As part of this project, I am analysing the data from Zomato Data set, where we can find the details of different restaurants in Bangalore. Here I am trying to identify the best neighbourhoods for getting the food of our choice. In this project, I am analysing the data based on the availability of Kerala foods, and the best restaurants and areas in Bangalore for getting it. Please note that we can select any category of food here including North, South, or the cuisine of your choice.

Business Problem:

• What is best location in Bangalore for Kerala Cuisine

• Which areas have large number of Kerala Restaurant Market

• Which all areas have less number of Restaurant

• Which is the best place to stay if I prefer Kerala Cuisine

• What places are have best restaurant in Bangalore

Interested Audience:

The question of best areas for food choices of interest is applicable to the all the people, especially for the ones who stayed in a different geographical region. With the diversity in culture, Bangalore become a major area where we need to find a solution for this question. Since the it is developed based on the Zomato dataset for Bngalore, the primary targeted audience are the people staying in Bangalore from different regions with different interest in food choices. Though this example is made based on Kerala cuisines, this can be used for the ones who choose different food choice.

Data Section

For this project we need the following data:

1. Zomato data set that contains the details for restaurants in Bangalore and the ratings.
2. Data Source: https://www.kaggle.com/himanshupoddar/zomato-bangalore-restaurants
3. Description: This data set contains the required information. And we will use this data set to explore various neighborhoods of Bangalore.
4. Kerala restaurants in Bangalore neighborhood.
5. Data Source: Foursquare API
6. Description: By using this API we will get all the venues in bangalore. We can filter these venues to get only Kerala restaurants.

Collect the Bangalore Zomato data from https://www.kaggle.com/himanshupoddar/zomato-bangalore-restaurants

* Using Foursquare API we will get all venues for each neighborhood.

* Filter out all venues which are Kerala Restaurants.

* Data Visualization and some statistical analysis.

* Analysing using Clustering (Specially K-Means):

1. Find the best value of K

2. Visualize the neighborhood with number of Kerala Restaurants.

List the various cuisines available in Zomato dataset for Bangalore

Just to start with Data Preprocessing, getting the details of various cuisines available for Bangalore Zomato Data Set

```
blr_data.cuisines.value_counts()
#blr_data.rest_type.value_counts()
```

North Indian	2284
North Indian, Chinese	2036
South Indian	1320
Cafe	653
Bakery, Desserts	644
Biryani	632
South Indian, North Indian, Chinese	601
Fast Food	580
Desserts	572
Chinese	449
Bakery	432
Ice Cream, Desserts	390
Chinese, North Indian	345
Mithai, Street Food	323
North Indian, Chinese, Biryani	280
Desserts, Ice Cream	272
Finger Food	261
Desserts, Beverages	258
South Indian, North Indian	257
North Indian, South Indian	254
Chinese, Momos	238
North Indian, South Indian, Chinese	234
Biryani, North Indian	223
Beverages, Fast Food	221
Street Food	213
Cafe, Fast Food	204
North Indian, Biryani	202

Beverages	201
South Indian, Chinese	185
North Indian, Mughlai	184

Methodology & Approach

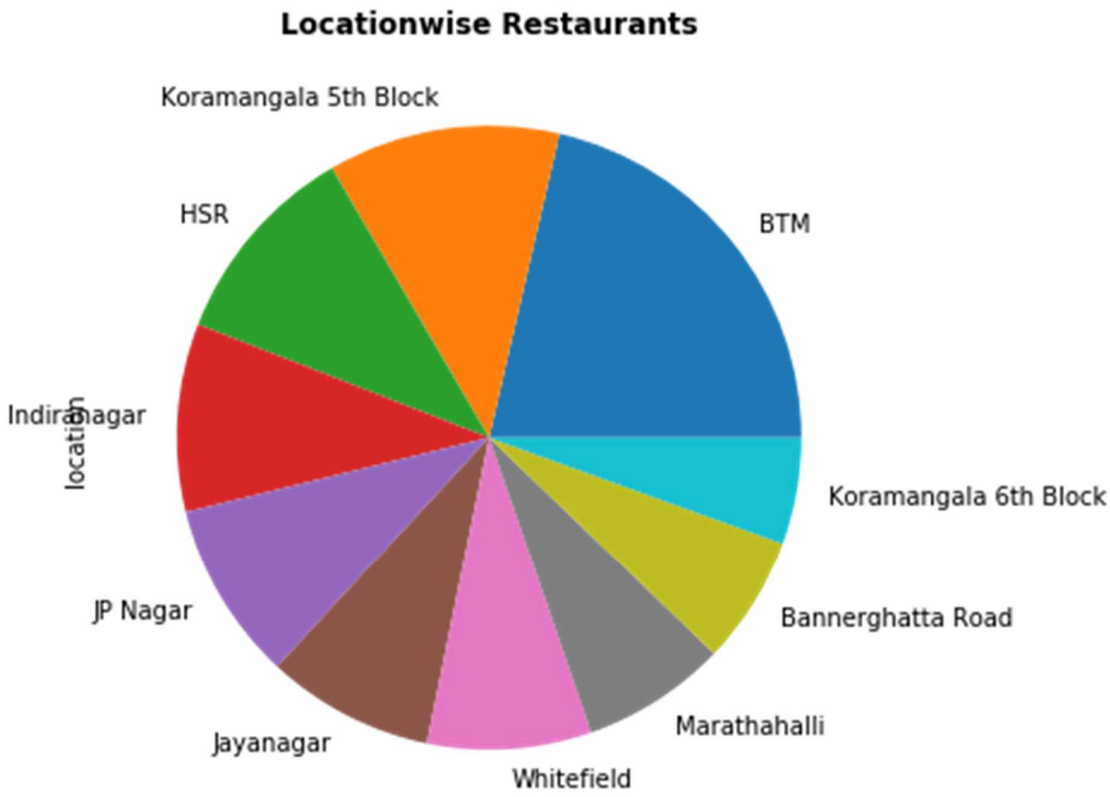
Approach

We use Zomato kaggel dataset Set for collecting the Bangalore restaurant data. We downloaded the dataset and use read_csv method to access the dataset and create dataframe from it. We use Geocoder for getting the latitude and Longitudes of all the locations received from the data set. We added the Latitude and longitude information for all the locations to the original data from by connecting it with location field. Using FourSquare API we find all venues for each neighbourhood and Filter out all venues that are nearby by locations. We Plot the data in bars using mathplotlib, for displaying the top ten highest rated Kerala restaurants area in Bangalore.

We use aggregative rating for each restaurant to find the best places for Kerala restaurants. Visualise the Ranking of neighbourhoods in Maps using folium library.

Modelling

We use KMeans Clustering Model to group the neighbourhood restaurants into five clusters. We use GeoCoder to get Geocode for all the locations listed in the Dataset. We mapped it to the original data frame using location identifier.



Results and Conclusion

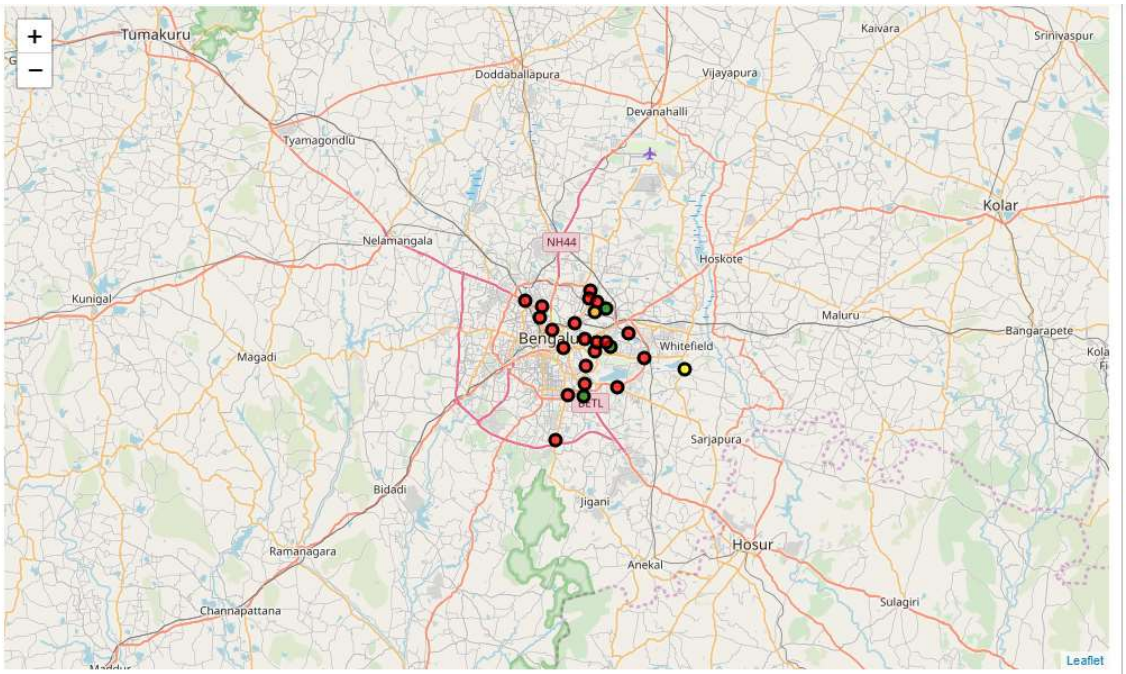
With the data now ready, we run k-means to cluster the neighborhoods into five (5) clusters.

MG Road, Seshadripuram and Chirch Street are some of the best neighborhoods for Kerala cuisine.

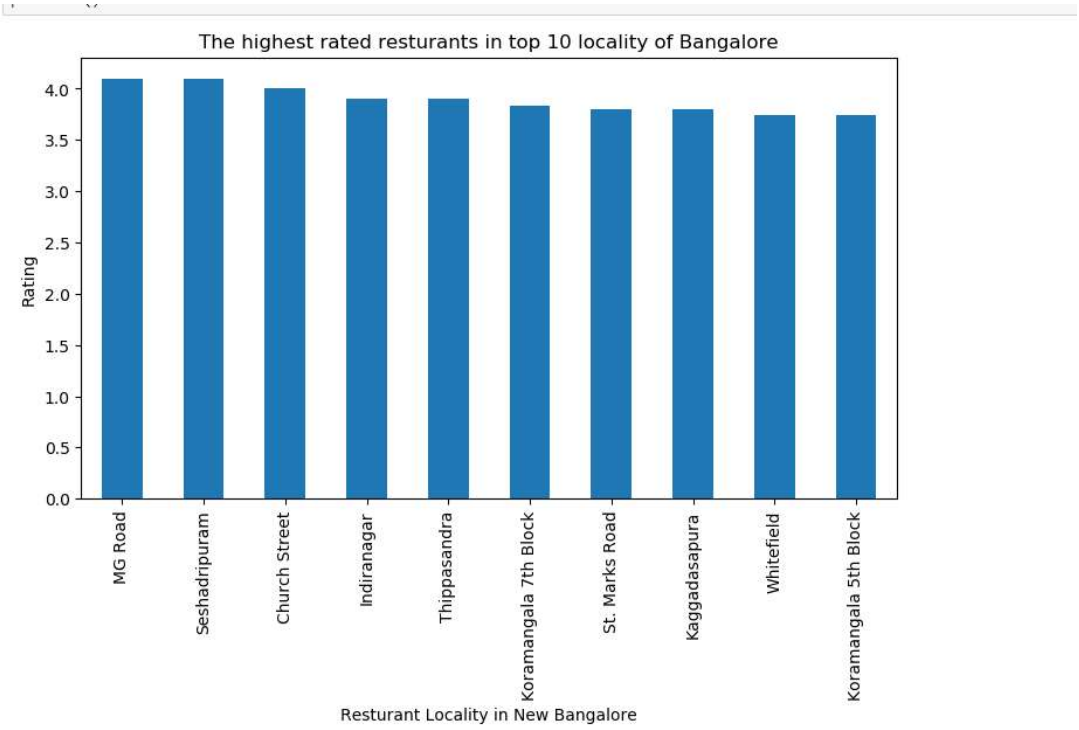
Cunningham Road,Commercial Street and Sivajinagar are not great places for Kerala cuisine.

BTM Layout and Koramangala 1st block are the best localities for Keala cuisine.

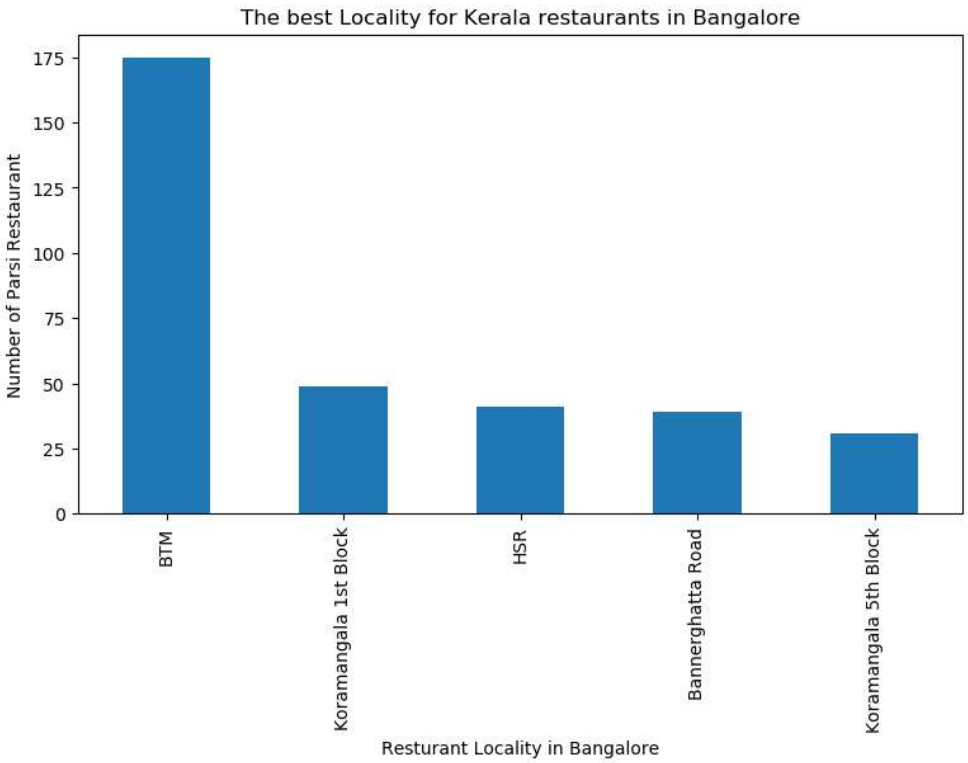
KMeans Clusters Visualization using Folium Maps



Best Rated Restaurants in Bangalore



Best Locality for Kerala Restaurants



Discussions

Please note that the result set can be updated for other cuisines just by putting the cusine name instead Kerala in this project. I used the Kerala to reduce the data processing time, since it has moderate amount of records for restaurants