COURSERA CAPSTONE PROJECT REPORT

INTRODUCTION:

One of a renowned MNC want to open an outlet of their Restaurant Business in Kolkata. They want to open the outlet near Park Street area. Before investing into the business they want to study the present market.

The management of the MNC wants to know about the Restaurants presently running in that area. They mainly want to know about the Cost-Per-Two person and Ratings provide by customers.

By analysing these datas they want to finalise the location for the Restaurant.

DATA:

Four Square API is used to fetch Restaurants data within the 5 Kms distance from Park Street, Kolkata. Restaurants name, address, latitude and longitude are taken from Four Square API.

Zomato API is used to fetch mainly ratings and price for two people data.

Finally these two data frames are combined to get the final data for the project.

]:

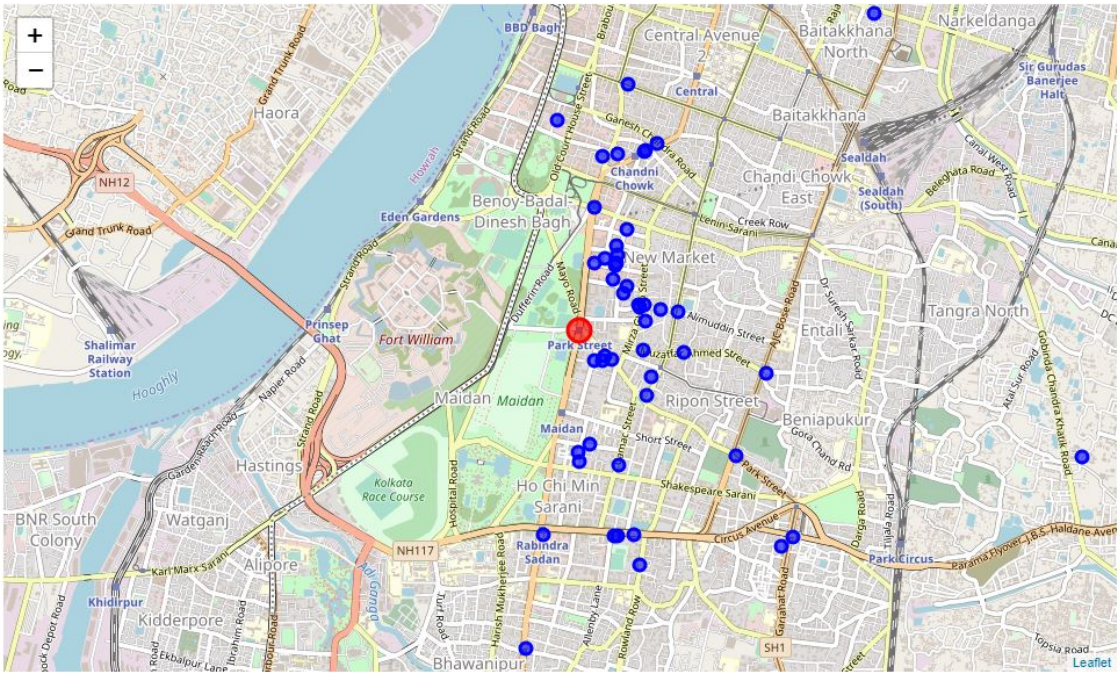
	name	categories	address	city	country	postalCode	state	latitude	longitude	price_for_two	price_range	rating
0	New Cathay Bar and Restaurant	Chinese Restaurant	17, Jawaharlal Nehru Rd	Kolkata	India	700087	West Bengal	22.5601	88.3517	250	1	3.2
1	Radhuni Restaurant	Indian Restaurant	17-G, Park St	Kolkata	India	700016	West Bengal	22.5570	88.3547	600	2	3.4
2	Shiraz Golden Restaurant	Mughlai Restaurant	135, Park St	Kolkata	India	700016	West Bengal	22.5465	88.3617	500	2	4.1
3	Kamal Restaurant	Vegetarian / Vegan Restaurant	33, Bentinck St	Kolkata	India	700069	West Bengal	22.5670	88.3521	600	2	3.5
4	Khandani Rajdhani Restaurant	Vegetarian / Vegan Restaurant	21, Park St	Kolkata	India	700016	West Bengal	22.5534	88.3520	100	1	3.7

Methodology:

In the very first section I have fetched the longitude and latitude of Park Street, Kolkata from Four Square API and the plot a map of Kolkata, cantered at Park Street with a zoom level of 15. Then I defined my Foursquare API credentials i.e. Foursquare client id, Foursquare secrete code and Foursquare version.

After defining the Foursquare credentials, I search the restaurants within 5 Km radius of Park Street from Foursquare API and put the received data into pandas dataframe. I put the data into the pandas dataframe so that I can easily clean the data to my desired level. I delete the unwanted columns from the dataframe and keep only the useful columns.

When the data cleaning is done then I plot a map of Park Street and put a red circle at the exact location of Park Street. Blue circles are added which are showing the positions of existing restaurants within the 5km area of Park Street.

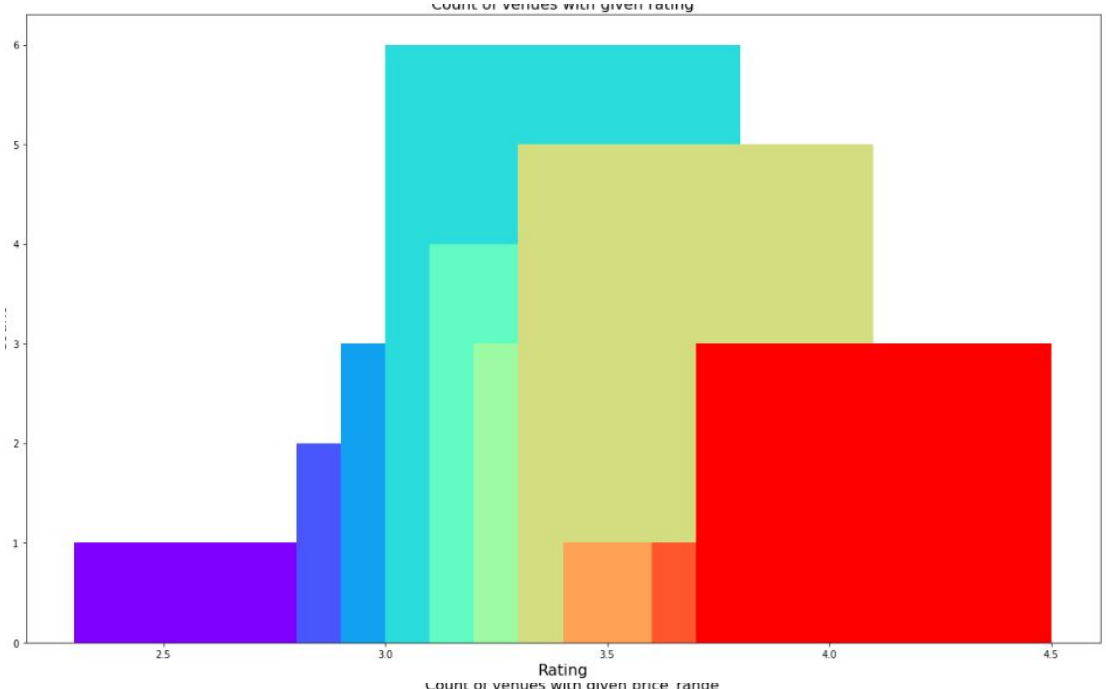


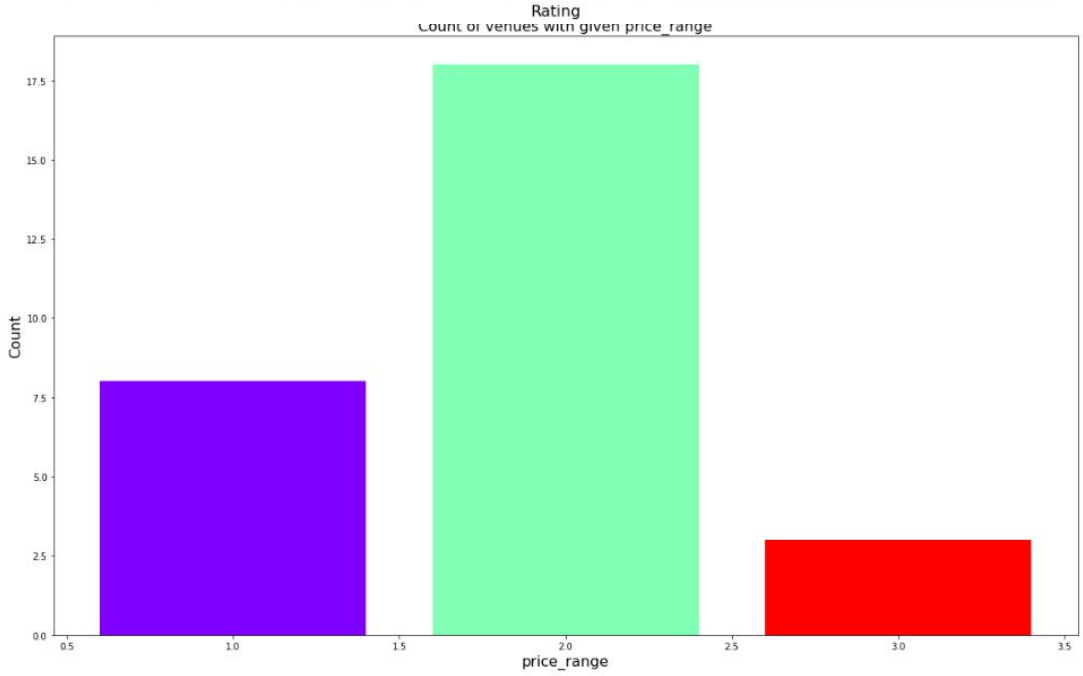
Now my aim is to get the user ratings and price range of the restaurants shown in the map. From user ratings we can predict the quality of that restaurants and from price range it can be predict whether it is costly or not. To get the user rating I use the Zomato API. I can get this

from foursquare API also but there is a limit in foursquare API. So I am not able to fetch the same.

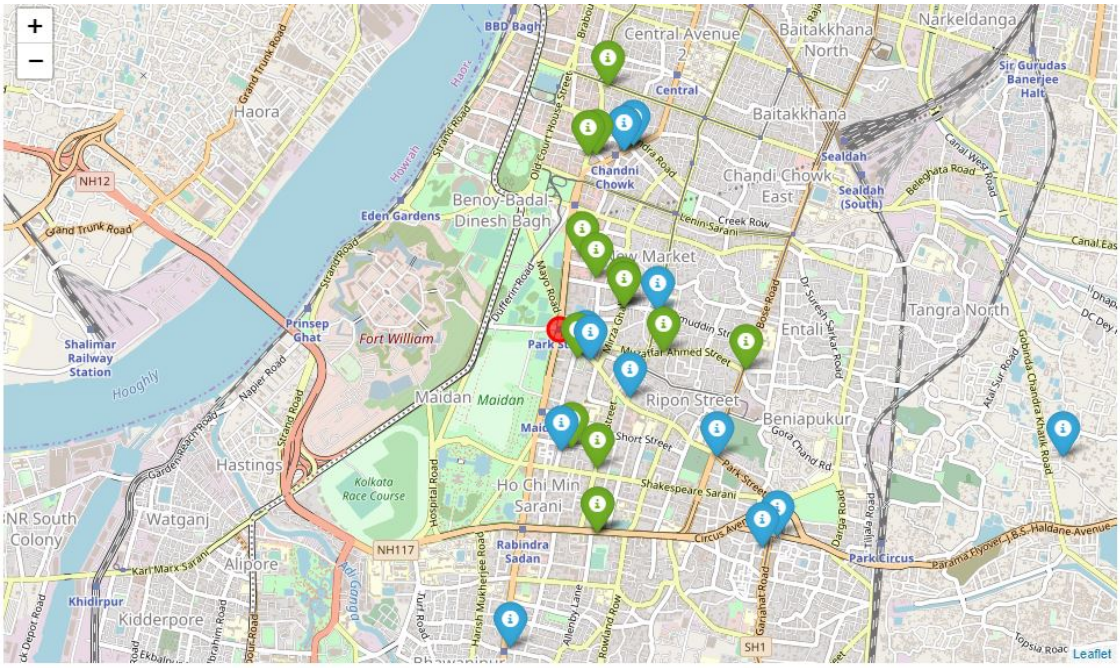
After fetching the Zomato API data, I put the data into a pandas dataframe and make it clean. I drop all columns except latitude, longitude and rating. Then I combine the two dataframe into a single dataframe and drop all unnecessary columns.

I drop the venues with 0 rating as it seems that the venues are not rated yet. After that I make a bar plot for showing the number of venues with given user ratings. From the plot it is shown that maximum venues have a rating in between 3 and 3.5. Then I plot the price range of the venues. Maximum venues have a price range of 2.



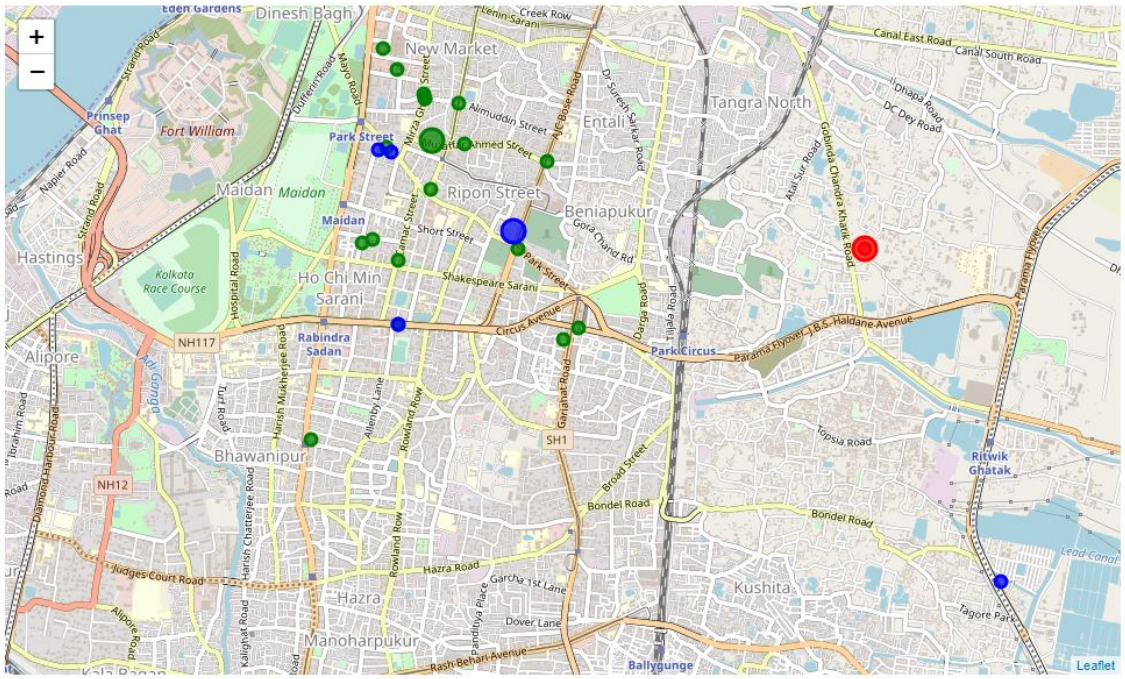


After that I classify the venues in four categories ‘Low’, ‘Okay’, ‘Good’ and ‘Very Good’ by their user ratings. Now again I plot the venues with the marker as 'Low': 'red', 'Okay': 'yellow', 'Good': 'green', 'Very good': 'blue'.



Then I cluster the venues in 3 cluster. Cluster 0 having mean price range of 1.65 and rating spread around 3.57. Cluster 1 having mean price range of 3 and rating spread around 4. Cluster 2 having mean price range of 2.4 and rating spread around 3.32.

At last I plot the cluster into a folium map.



RESULT:

From the final map it is clear that around the Park Street, low price rage venues have medium user rating, medium price range venues have low user rating and high price range venues have high user ratings. It is also observed that number of high price range venues are very less but users are like that venues.

DISCUSSION:

From the data I have observed that most of the venues are rated by users and according to their comments they can afford price for a good quality food and pleasant atmosphere of the venue.

CONCLUSION:

From the above discussion I came to this conclusion that the people of Kolkata who visited the Park Street area for taking their lunch or dinner can afford high price for best quality. But number of such venues are very less. So if anybody want to open a restaurant around the Park Street area he can open a venue with best quality of food and surroundings.