PROBLEM: USING MACHINE LEARNING TO SUGGEST FOOD, COFFEE, NIGHTLIFE, FUN AND SHOPPING PLACES TO TOURISTS

Abhishek Paul
National Institute of Technology, Durgapur
Github Profile: https://github.com/abhishekpawl

1 Introduction

Business Problem:

The objective of this project is to suggest tourists visiting Kolkata, India places for food, coffee, nightlife, fun and shopping based on the location of the hotel in which they will be staying to save their time making plans. I have used an *Unsupervised Machine Learning* method, namely, **Clustering** for recommending the tourists places. This project aims to answer the question: What are the places for food, coffee, nightlife, fun and shopping that tourists can try out?

Background:

For this capstone project, I am concentrating on the scenario of tourists visiting Kolkata, India. Kolkata is called ‘the city of joy’. The city is known for its grand colonial architecture, art galleries, cultural festivals, sweets, food culture and many more. Kolkata is a common destination for vacations. This city attracts tourists from all over the world. So, this project is concerned about suggesting tourists places where they can hangout such as places for food, coffee, nightlife, fun and shopping and thereby saving their valuable vacation-time. This project takes into account the geocoordinates (latitude, longitude) of the hotel and based on the same, makes the suggestions.

Target Audience:

The target audience for this project are the tourists visiting spending their vacation in Kolkata, IN.

Clustering:

Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group and dissimilar to the data points in other groups. The particular method used in the project is k-means clustering. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed apriori. The main idea is to define k centers, one for each cluster. The next step is to take each point belonging to a given data set and associate it to the nearest center. When no point is pending, the first step is completed. At this point we need to re-calculate k new centroids as barycentre of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new center. A loop has been generated. As a result of this loop we may notice that the k centers change their location step by step until no more changes are done or in other words centers do not move any more.

2 Data

All the data used in the project are fetched using the **FOURSQUARE API**.

Data for hotels:

'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, Latitude, Longitude, VERSION, ‘Hotel’, RADIUS, LIMIT)

Data for food places near the hotel:

'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, HOTEL_Latitude, HOTEL_Longitude, VERSION, ‘Food’, RADIUS, LIMIT)

Data for coffee places near the hotel:

'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, HOTEL_Latitude, HOTEL_Longitude, VERSION, ‘Coffee’, RADIUS, LIMIT)

Data for nightlife places near the hotel:

'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, HOTEL_Latitude, HOTEL_Longitude, VERSION, ‘Nightlife’, RADIUS, LIMIT)

Data for fun places near the hotel:

'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, HOTEL_Latitude, HOTEL_Longitude, VERSION, ‘Fun’, RADIUS, LIMIT)

Data for shopping places near the hotel:

'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, HOTEL_Latitude, HOTEL_Longitude, VERSION, ‘Shopping’, RADIUS, LIMIT)

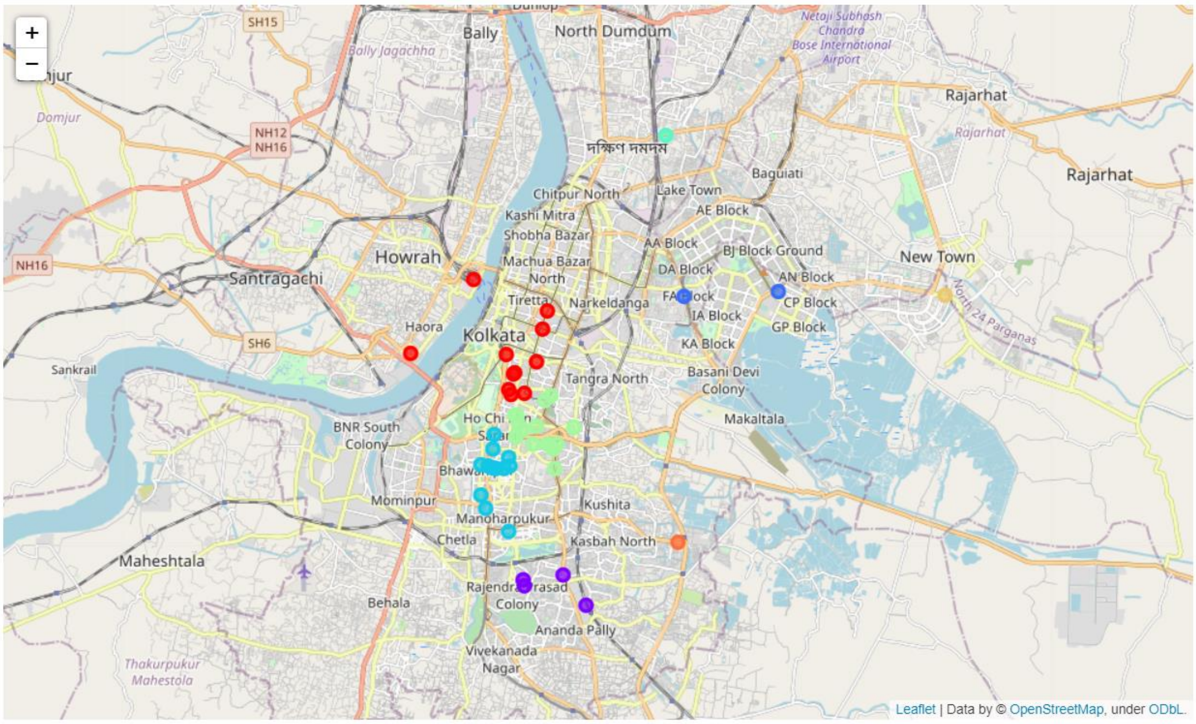
3 Methodology

The methodology used in this project is Unsupervised Machine Learning. I have used Clustering method for clustering the places for food, coffee, nightlife, fun and shopping. The particular method used here is K-Means Clustering. The main idea is to define k centers, one for each cluster. The next step is to take each point belonging to a given data set and associate it to the nearest center. When no point is pending, the first step is completed. At this point we need to re-calculate k new centroids as barycentre of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new center. A loop has been generated. As a result of this loop we may notice that the k centers change their location step by step until no more changes are done or in other words centers do not move any more. I have clustered the places based on their geo-coordinates. Finally, I calculated the Euclidean distances of each cluster from the hotel and return the places belonging to the cluster that is the nearest to the hotel.

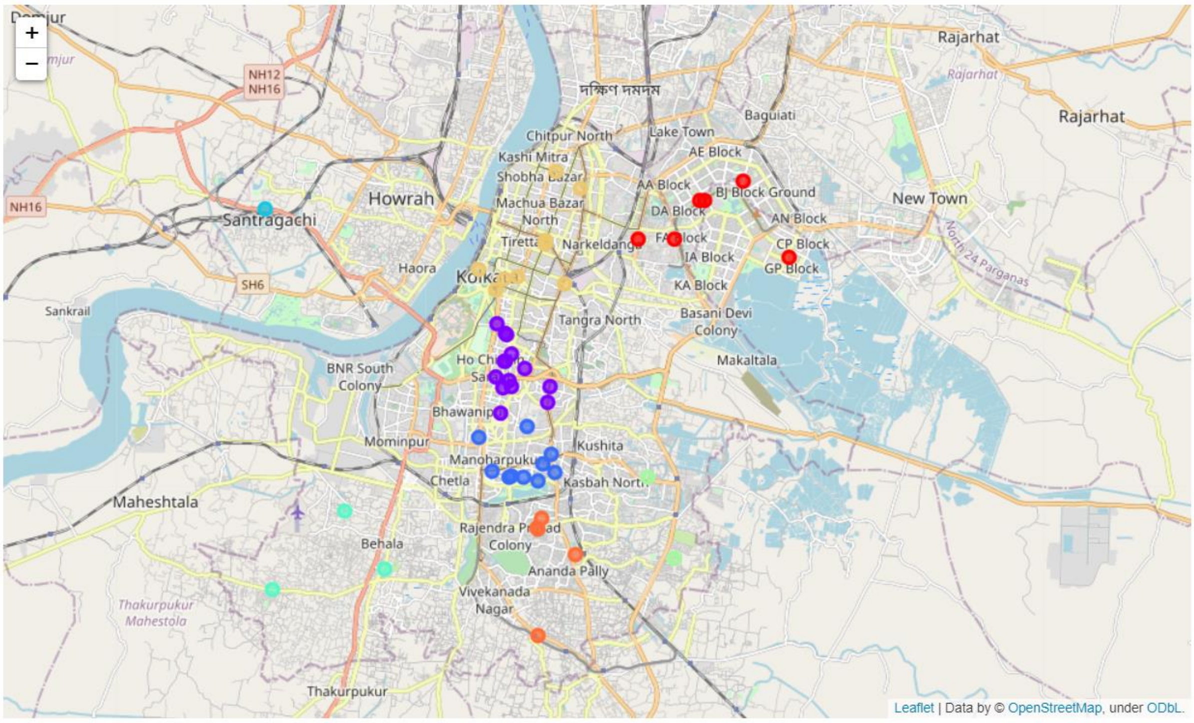
4 Results

The clustered places visualised on the folium map are as follows:

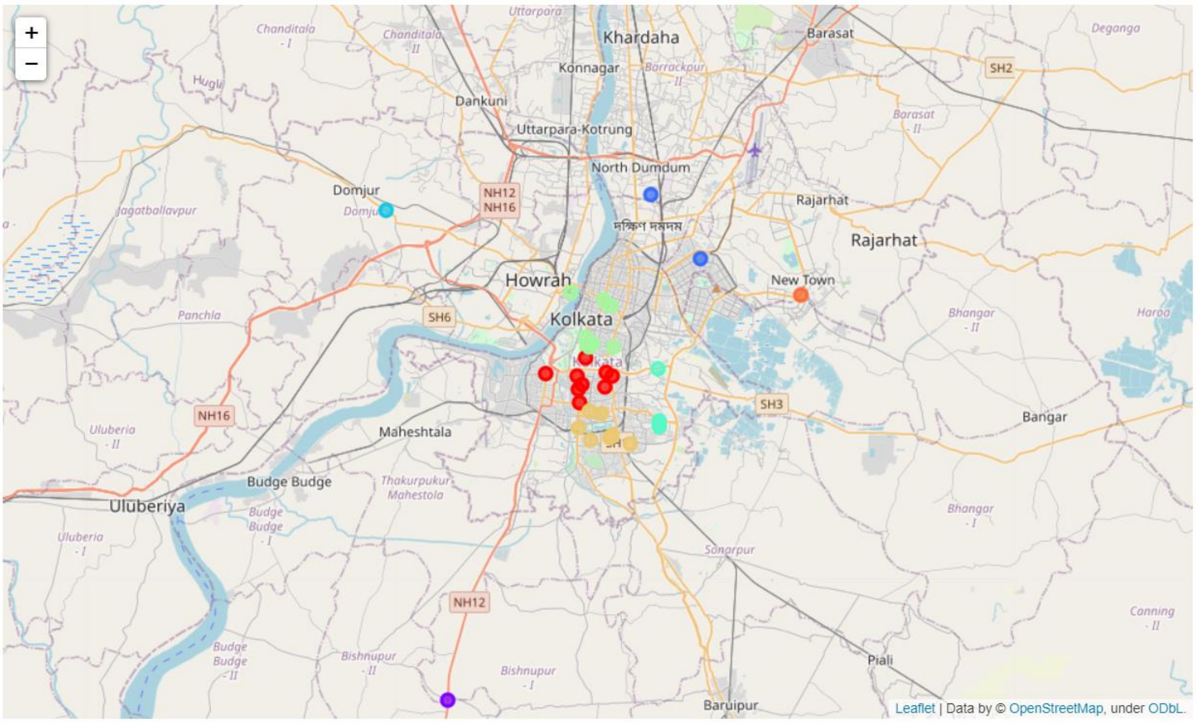
Places for food: (8 clusters)



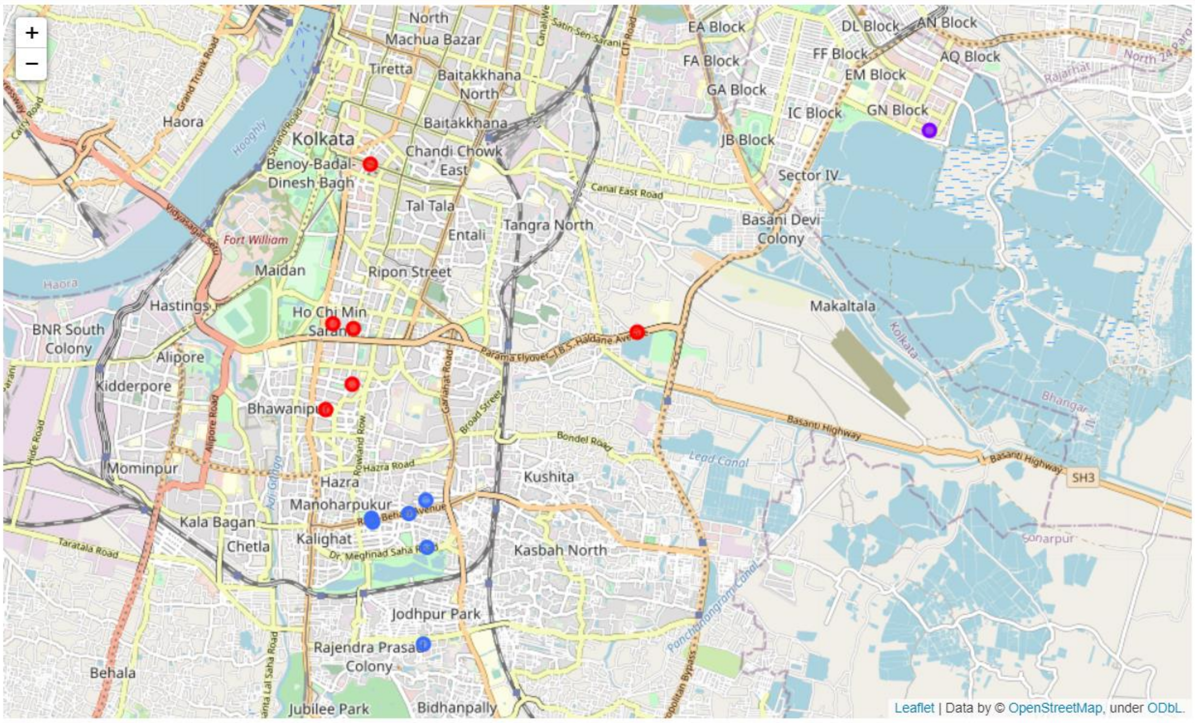
Places for coffee: (8 clusters)



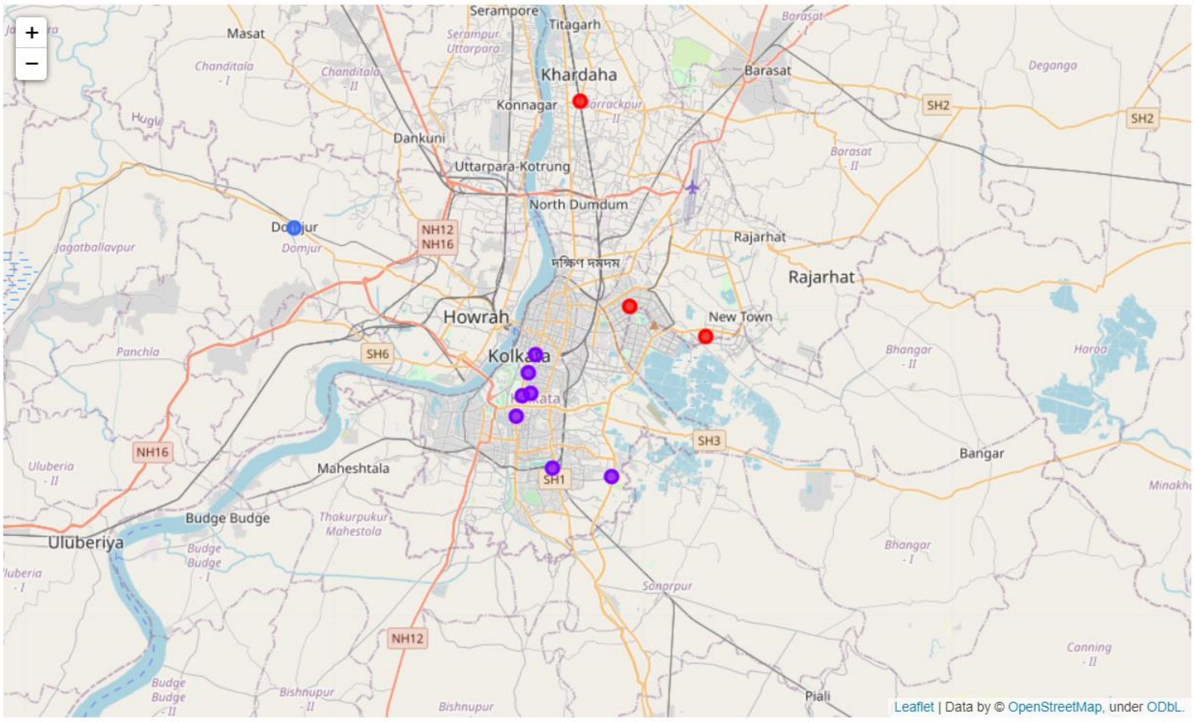
Places for nightlife: (8 clusters)



Places for fun: (3 clusters)



Places for shopping: (3 clusters)



The suggested places are as follows:

Suggested places for food:

	Name	ID	Latitude	Longitude	Address	Label
0	Food Bazaar	4fc631bbe4b005fcba5f8871	22.546905	88.353392	['Pantaloons', 'Kolkata 700016', 'West Bengal', 'India']	5
2	R C S Fast Food	5b0ab58ee1f228002c5797d7	22.541059	88.353649	['229, Acharya Jagadish Chandra Bose Rd', 'Kolkata 700020', 'West Bengal', 'India']	5
3	Food Court	598eb4d98d0a5310231f7437	22.539206	88.365499	['Quest Mall', 'Kolkata 700017', 'West Bengal', 'India']	5
9	Aqua Java Fast Food Pvt. Ltd.	4e71c57ce4cd7aa6d6f40a53	22.546865	88.354229	['10, Wood St', 'Kolkata 700017', 'West Bengal', 'India']	5
13	Aqua Java Fast Food Pvt. Ltd.	4bbe3eda79eaef3b07e4aec3	22.542949	88.355867	['8/2, Loudon St', 'Kolkata 700016', 'West Bengal', 'India']	5
14	Hunger Food Court	5ca04aaf18d43b002c8c4b59	22.544605	88.354576	['700017', 'West Bengal', 'India']	5
16	Arsh Fast Food	51dac3e2498e0188907115ec	22.550649	88.362094	['India']	5
17	City Fast Food Center	5916c41f6bd36b0dc9458694	22.551789	88.363664	['Acharya Jagadish Chandra Bose Rd (Beniapukur Road)', 'Kolkata 700017', 'West Bengal', 'India']	5
19	Al falah fast food	5088fd48e4b0b15cd5f580ad	22.543978	88.359619	['Kolkata', 'West Bengal', 'India']	5
21	Shimnan Fast Food Centre	50c3145ce4b0e121076eabb6	22.541098	88.353767	['Acharya Jagadish Chandra Bose Rd', 'Kolkata 700020', 'West Bengal', 'India']	5
28	Whats App Fast Food	5781d311498ec67efba5a726	22.539295	88.358703	['66, Ballygunge Circular Rd', 'Kolkata 700019', 'West Bengal', 'India']	5
32	Alisha Biryani & Fast Food	55291ae5498eddb919a4d113	22.539354	88.362753	['101/1A, Karaya Rd', 'Kolkata 700017', 'West Bengal', 'India']	5
34	Skyline Health Food	52dd2397498e516317c0c075	22.539265	88.362989	['101/1A, Karaya Rd', 'Kolkata 700017', 'West Bengal', 'India']	5
35	Haldiram Food City	4c13a3f777cea5935c1acf60	22.532822	88.364788	['24, Ballygunge Pk', 'Kolkata 700019', 'West Bengal', 'India']	5
39	Eat Good Food	58babbf2dfa6ff23d8c04cd9	22.537885	88.364477	['19, Tarak Dutta Rd', 'Kolkata 700019', 'West Bengal', 'India']	5
40	Shudh Veg Food Court	58babc20c30f797a72740ee7	22.537800	88.364500	['24, Tarak Dutta Rd', 'Kolkata 700019', 'West Bengal', 'India']	5
41	Anis Fast Food	4fb4c2ffe4b06fedadbb804e	22.543513	88.370381	['India']	5

Suggested places for coffee:

	Name	ID	Latitude	Longitude	Address	Label
0	Cafe Coffee Day	4c6e9eae4ee776b0b4b0581b	22.545478	88.352501	['4/1, Camac St', 'Kolkata 700016', 'West Bengal', 'India']	1
1	Cafe Coffee Day	4c6ea5834ee776b055e1581b	22.552775	88.352717	['Park St', 'Kolkata 700017', 'West Bengal', 'India']	1
2	Cafe Coffee Day	5167a710e4b08b494a5adeed	22.555407	88.350313	['Park St', 'Kolkata 700016', 'West Bengal', 'India']	1
4	The Coffee Bean & Tea Leaf	52c16ed711d2e75296b8c7f6	22.538809	88.365524	['Quest Mall', 'Kolkata 700017', 'West Bengal', 'India']	1
6	Cafe Coffee Day	4c6eaf425d089c7487bf07c3	22.547381	88.354433	['Wood St', 'Kolkata 700016', 'West Bengal', 'India']	1
7	Cafe Coffee Day Kolkata Office	50347ce4e4b0bf8f2d057132	22.545424	88.352783	['Camac St', 'Kolkata 700017', 'West Bengal', 'India']	1
12	Coffee World	4d0e13a8801054816a042296	22.534602	88.364883	['1, Ballygunge Park Rd', 'Kolkata 700019', 'West Bengal', 'India']	1
13	Cafe Coffee Day	4e5b9da345dd705592b256ab	22.543596	88.358161	['42/A, Rawdon St', 'Kolkata 700017', 'West Bengal', 'India']	1
17	Cafe Coffee Day	5cf94dea8f90d9002347d428	22.541287	88.349830	['235/2A, Acharya Jagadish Chandra Bose Rd', 'Kolkata 700020', 'West Bengal', 'India']	1
19	Cafe Coffee Day	4c6ea0814ee776b0f4be581b	22.538542	88.351954	['Woodburn Pk', 'Kolkata 700020', 'West Bengal', 'India']	1
23	Starbucks	5ab0e6b16adbf502ef89feb3	22.552444	88.353199	['57A, Park St', 'Kolkata 700016', 'West Bengal', 'India']	1
27	Artsy Coffee And Culture	5aa4fb48916bc12e2eda19e1	22.540373	88.353806	['230/B, Acharya Jagadish Chandra Bose Rd', 'Kolkata 700020', 'West Bengal', 'India']	1
28	Dosa Coffee	5d9dcd8589f8b00008503f5f	22.538953	88.354282	['11/1, Sarat Bose Rd', 'Kolkata 700020', 'West Bengal', 'India']	1
41	Cafe Coffee Break	5cade58c89e490002ce36ab0	22.531843	88.351520	['52A, Paddapukur Rd', 'Kolkata 700020', 'West Bengal', 'India']	1

Suggested places for nightlife:

	Name	ID	Latitude	Longitude	Address	Label
5	London Pub	583d8075ea29b85601cdf158	22.547619	88.349700	['Golden Parkk, Chowringhee', 'India']	0
7	Rishirich Soda Pub	52dd1ccc498e9f6c1d09befe	22.540038	88.361610	['Beck Bagan Row', 'Kolkata 700017', 'West Bengal', 'India']	0
8	Soda Pub	4fdb69ede4b08187e9dabda9	22.538066	88.345485	['Sambhu Nath Pandit St', 'Kolkata 700020', 'West Bengal', 'India']	0
9	Rishirich Soda Pub	529e921511d255ec089af691	22.533329	88.347609	['23A/1/B, Justice Dwarka Nath Rd', 'Kolkata 700020', 'West Bengal', 'India']	0

Suggested places for fun:

	Name	ID	Latitude	Longitude	Address	Label
0	Kolkata Fun Club	5446e0e6498ee9a7ef4d7957	22.543556	88.348875	['25, Jawaharlal Nehru Rd', 'Kolkata 700016', 'West Bengal', 'India']	0
2	COLORZ - The Funked-Up Store	4ffab423e4b0ae40ee877f91	22.564408	88.354348	['1/A, Grant St', 'Kolkata 700013', 'West Bengal', 'India']	0
4	Funda & Yener Karakaş	549fc3a0498ef7f11ab5872d	22.532181	88.347939	['Bbsnsbbaba', 'India']	0
5	Personalized Funda	51c8fb85498e8a70de1e97c4	22.535519	88.351729	['9/1, Ram Mohan Dutta Rd', 'Kolkata 700020', 'West Bengal', 'India']	0
8	Fung Fa	52e7cbc2498e9bdd9a1eb232	22.542376	88.392159	['119/A/1, Matheshwartala Rd', 'Kolkata 700046', 'West Bengal', 'India']	0

Suggested places for shopping:

	Name	ID	Latitude	Longitude	Address	Label
0	The Globe Shopping Mall	56d54c51498e3b3c9ff86414	22.558964	88.352437	['21, Humayun Pl', 'Kolkata 700087', 'West Bengal', 'India']	1
1	Shopping	5711e3dd498e7d4892cd6082	22.568306	88.356557	['31, Ganesh Chandra Ave', 'Kolkata 700013', 'West Bengal', 'India']	1
2	Vardaan Shopping Market	4d4faa1a8a592c0f752c8f9f	22.548193	88.353810	['25A, Camac St', 'Kolkata 700016', 'West Bengal', 'India']	1
3	Homeland Shopping Mall	4fead6a4011c5c131a1e8fa8	22.536282	88.346075	['18B, Ashutosh Mukherjee Rd', 'Kolkata 700020', 'West Bengal', 'India']	1
4	Metro Plaza	4db2a089f7b12ee8facdb145	22.547103	88.349063	['1, Ho Chi Minh Sarani', 'Kolkata 700071', 'West Bengal', 'India']	1
5	Dakshinapan Shopping Complex	4bcf06c8cc8cd13a0e33c5cf	22.508749	88.366601	['2, Gariahat Rd (Near Dhakuria Bridge)', 'Kolkata 700068', 'West Bengal', 'India']	1
6	Avishar Shopping Mall	5102aca2e4b077344fd5b9c4	22.504112	88.400030	['369/4, Eastern Metropolitan Bypass', 'Kolkata 700099', 'West Bengal', 'India']	1

5 Discussion

The project did not take into account the details of the places as there were problems regarding Foursquare API calls, that could have contributed more to the legitimacy of the project.

6 Conclusion

The project took into consideration, that the tourist had put up in The Park Hotel, and then returned the suggest places for food, coffee, nightlife, fun and shopping that the tourist can try out.