

Capstone Project

Battle of neighborhoods

Amit Sarkar | IBM Data Science Professional Certificate | 29 Mar 2019

## Introduction

I live in a city called Kolkata. It's one of metropolitan city in India. French author Dominique Lapierre called this City of Joy, It can be considered as food capital of India. So recently one of my friend Reshmi Mazumdar wished to open a Chinese restaurant in Kolkata. But she is not sure about the location where she should open her restaurant. So she discussed with me about her dream project as she was aware about my learning Data science in Coursera.

So I thought below business problem for her:

* She should open her restaurant in a location where there are enough customers and minimum competitor restaurant.
* We should explore all the qualified venues in Kolkata to decide best location for Chinese restaurant.
* Need to know what category of customers are available whether they are target customers by their age.
* Based on all the information we need to maximize score and decide the best location.

## DATA

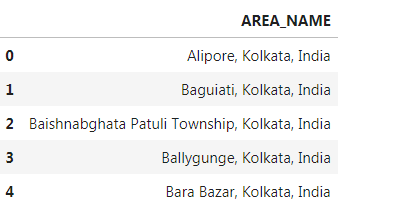
To help Reshmi Mazumdar for her dream project we need below data:

* All area names of Kolkata which we have researched and make a csv file.
* Based on area name we have accessed longitude and latitude data by using geolocator.
* Based on area name, longitude and latitude data we have accessed nearby venues by foursquare API.
* For each location we got N numbers of venues with different categories i.e. Chinese restaurant, cinema, station etc.

## Implementation

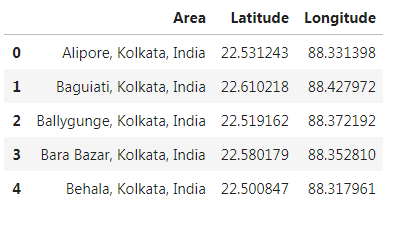
We have collected below data from different source and make a csv file which has been kept in Github repository:

<https://github.com/amit58/Coursera_Capstone/blob/master/Kolkata_Neighborhood_Data.csv>

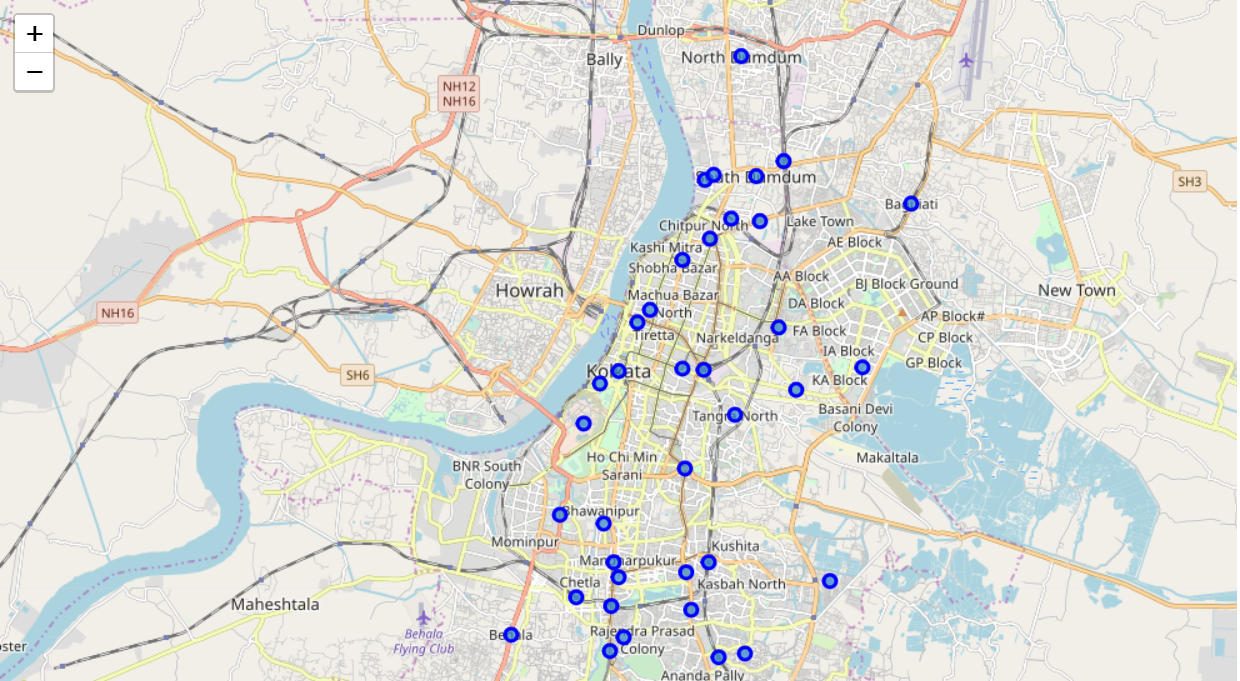


Next we have used GeoPy library to access longitude and latitude data for each location in the csv file and saved in another csv file.

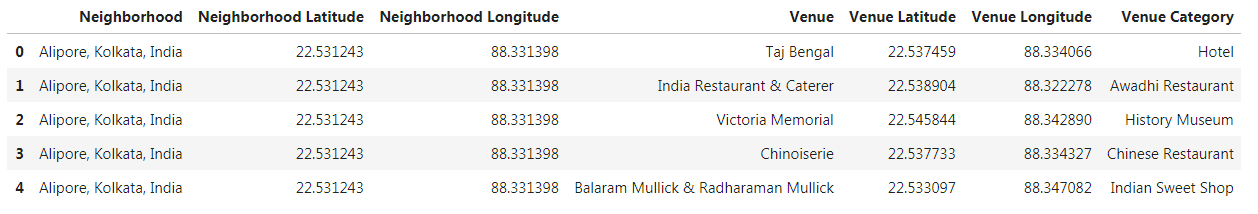
<https://github.com/amit58/Coursera_Capstone/blob/master/Kolkata_Location_Data.csv>



Using above location data we have created below neighborhood map of Kolkata:



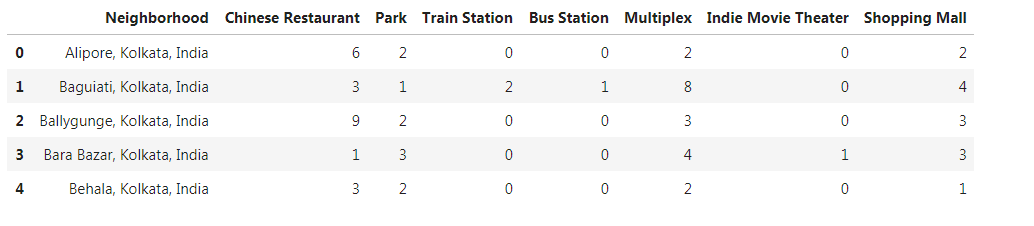
We have accessed venues data by using FourSquare API



Next we have analyzed each venue based on below categories as these categories influence the location of the restaurant mostly.

1. Chinese restaurant
2. Railway station
3. Bus station
4. Shopping mall
5. Movie theater
6. Theme park

We have given negative weightage if Chinese restaurant is available around the venue as this this the direct competitor of the business. So we have grouped data as below.

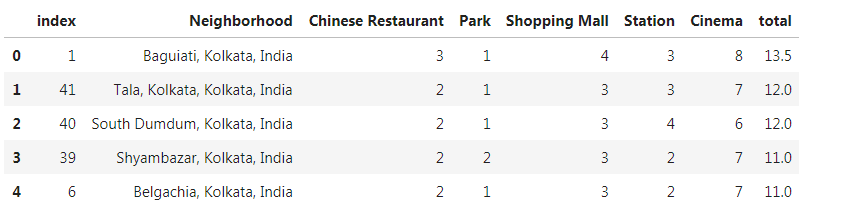


Below is the final data frame to calculate total score



## RESULT

Finally we have ordered data descending by Total score to get best location at the top of the data frame.



## CONCLUSION

Based on the maximum score we have found we can conclude the she should open her Chinese restaurant in below Location:

Baguiati, Kolkata, India