DATA SCINCE CAPSTONE PROJECT

BATTLE OF NEIGHBORHOODS

Amit Sarkar

IBM DATA SCIENCE PROFESSIONAL CERTIFICATE

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.

INTRODUCTION Contd.

- 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.

Libraries used for the project

- Pandas – A data manipulation and analysis library
- Numpy - Scientific computing with Python
- Geopy - To locate the coordinates of addresses, cities, countries, and landmarks across the globe
- Folium – Map creation library

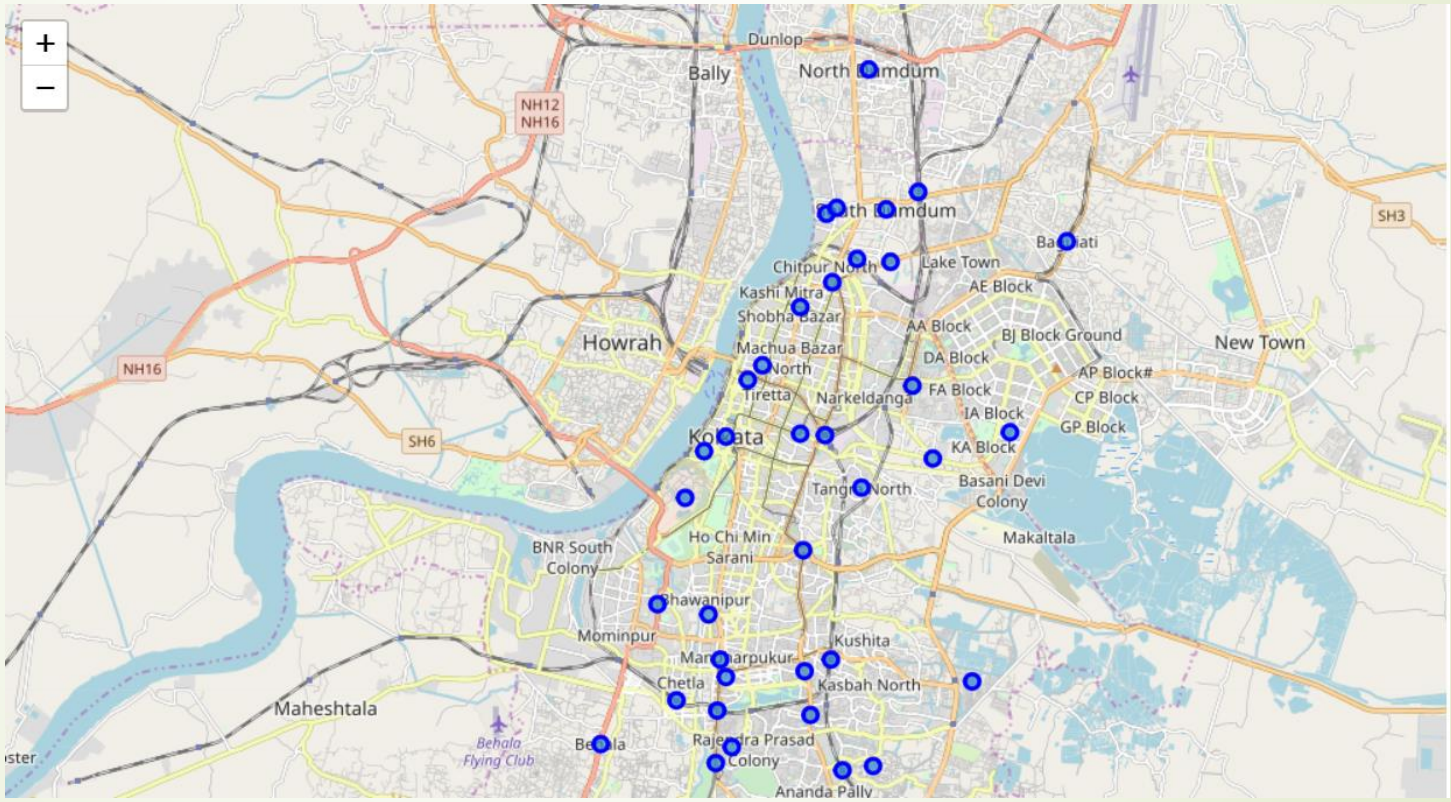
Chinese restaurant

- Below criteria have been chosen to decide best location to open a Chinese restaurant in Kolkata.
  - There should be minimum direct competitor in that location i.e. there shouldn't be so many Chinese restaurants.
  - It's better to have some shopping places around the chosen location.
  - It's a plus to have some movie theater around the location.
  - The location to be chosen where commute station is present i.e. bus station, railway station etc. as bus and railway are major public transportation mode in Kolkata.

Implementation

- We have collected area data from different source and make a csv file which has been kept in GitHub repository.
- we have used GeoPy library to access longitude and latitude data for each location in the csv file and saved in another csv file.
- Using above location data we have created neighborhood map of Kolkata.
- We have accessed venues data by using FourSquare API
- We have given negative weightage if Chinese restaurant is available around the venue as this this the direct competitor of the business.

Kolkata neighborhood map



Score calculation

- Total score = -0.5 * Chinese restaurant + Station + Cinema + Park + Shopping mall
- Based on total score formula we have chosen location with maximum total score.

Result

	index	Neighborhood	Chinese Restaurant	Park	Shopping Mall	Station	Cinema	total
0	1	Baguiati, Kolkata, India	3	1	4	3	8	13.5
1	41	Tala, Kolkata, Kolkata, India	2	1	3	3	7	12.0
2	40	South Dumdum, Kolkata, India	2	1	3	4	6	12.0
3	39	Shyambazar, Kolkata, India	2	2	3	2	7	11.0
4	6	Belgachia, Kolkata, India	2	1	3	2	7	11.0

Conclusion

- Below is best location to open Chinese restaurant in Kolkata

Baguiati, Kolkata, India