Coursera Capstone

IBM Data Science Professional Certificate

Exploration of Eateries in Kolkata

By: Riddhiman Basu

April 2020

Business Problem

- For anyone who visits Kolkata as a tourist or for work assignments:

- How does he/she choose a cheap place which offers good food?
- How can the person find out the highly rated restaurants in Kolkata?
- What good options for lunch/dinner are present near the locality where the person is staying?

Approach

- This project explores various eateries in Kolkata and attributes the data based on user ratings and average price.
- A map of the venues with specific color attributes has been plotted to highlight their position, and information about these venues.
- This enables any visitor to take a quick glance and decide what place to visit.

Target Audience

- Tourists from abroad and other cities, as well as people who visit Kolkata on work assignments. This analysis could help them pick out restaurants which fit their budget and also near their current locality.

- In addition, travel portals like TripAdvisor may also use such data to recommend restaurants for visitors.

Data Requirements

- Data will be acquired from two main sources:

- Foursquare API

  Name, Category, Latitude, Longitude

- Zomato API
- Name, Address, Rating, Price Range, Price for two, Latitude, Longitude

Data Processing

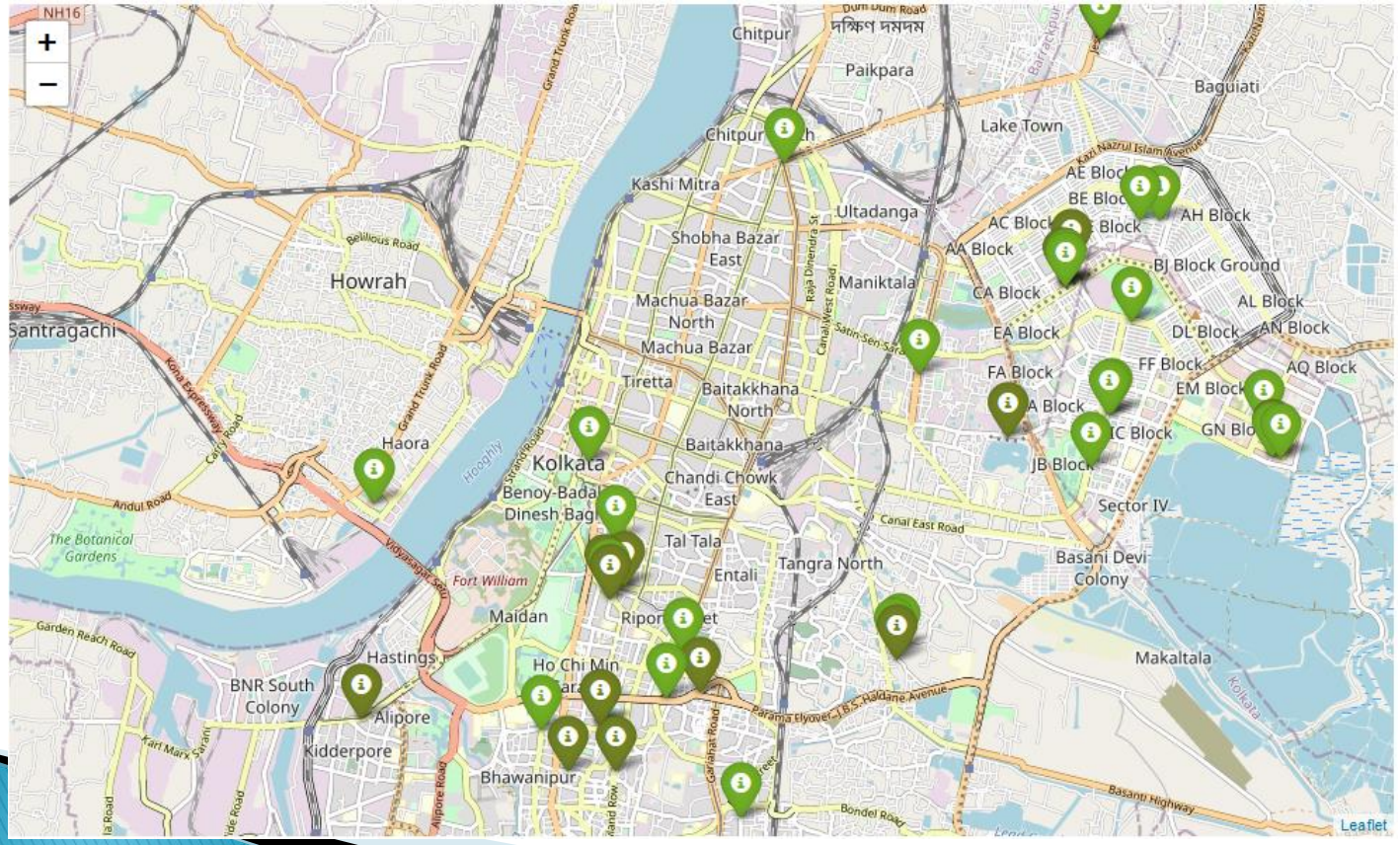
- Erroneous data like venues showing as 0 were removed
- Foursquare and Zomato Data were combined by calculating the difference between latitude and longitude. Venues having difference less than 0.0004 were considered as one.
- Average price for all the eateries were calculated from zomato data

Methodology

- Identified places where many venues are located so that any visitor can go to one place and make a choice amongst various venue options.
- Explored and plotted areas that are high rated and low rated.
- Also plotted the locations of high and low priced venues.
- Clustered the available data on venues to identify common patterns.

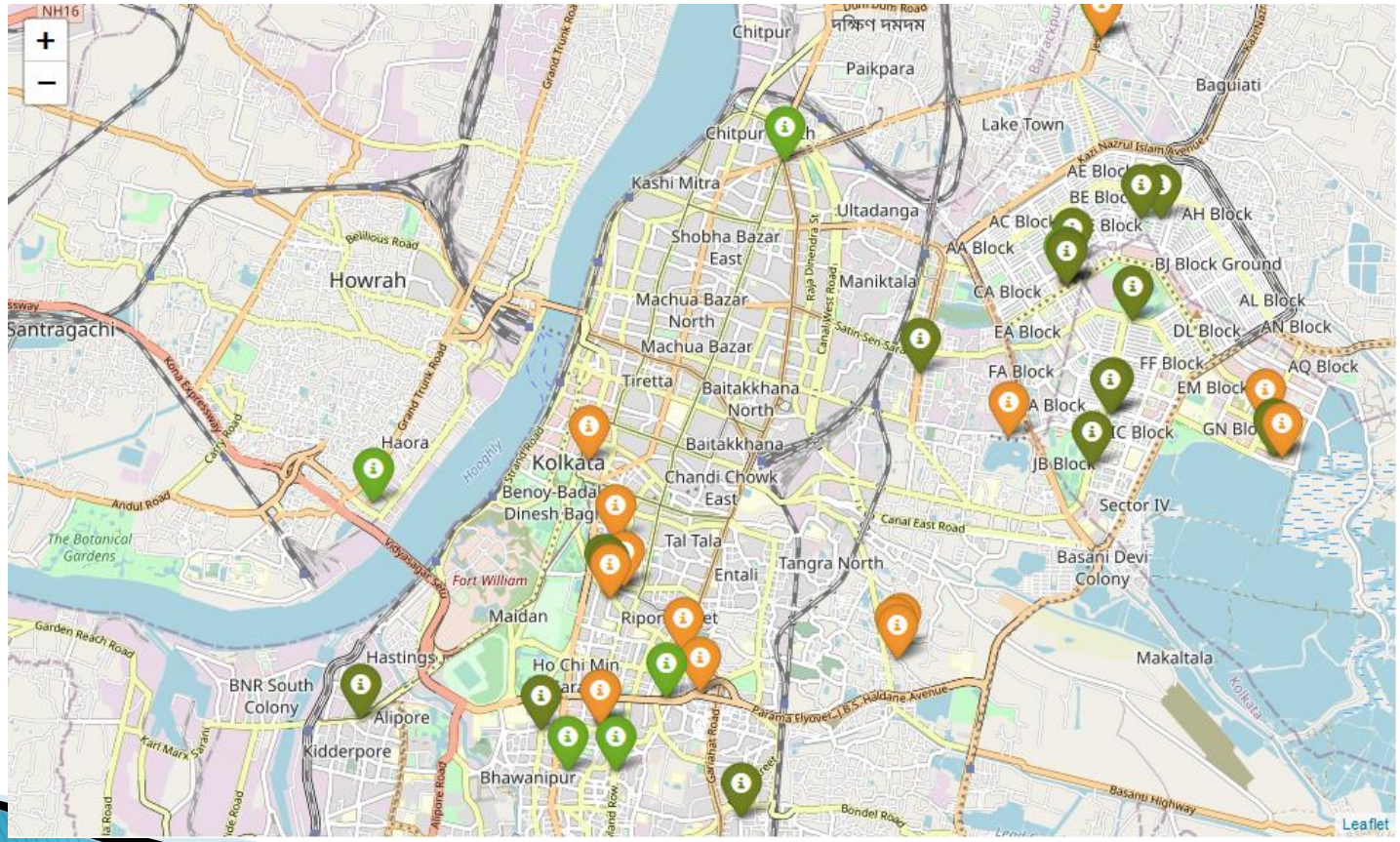
Analysis

- Plot of venues based on ratings – Red for Lowest to Dark Green for Highest



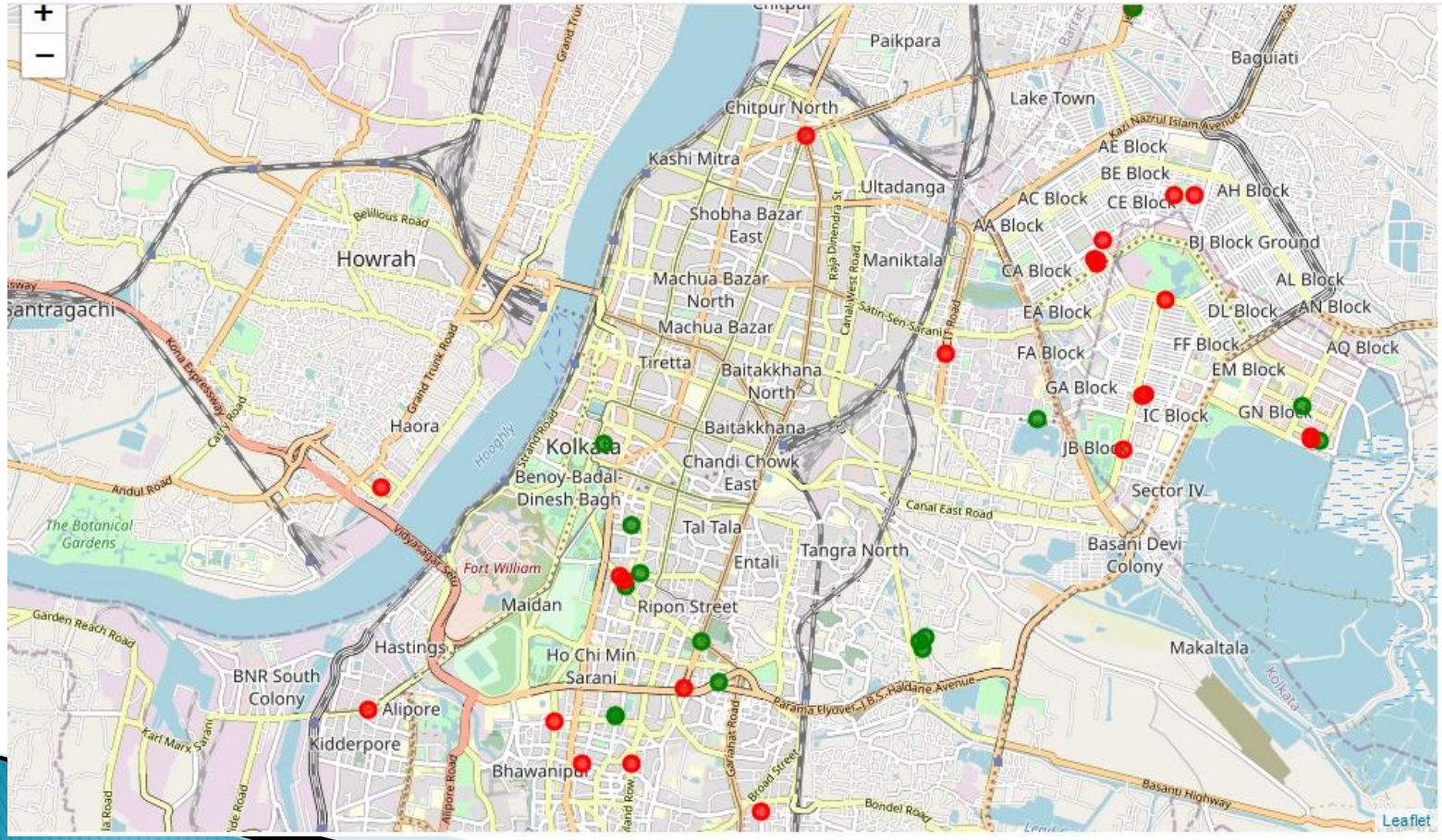
Analysis – Contd

- Plot of venues based on price range – Green for Lowest to Red for Highest



Analysis – Contd

- The data was clustered into two clusters – Red and Green



Analysis – Contd

- Cluster 0 – Green had mean price range of 3.06 and rating spread around 4.06.

- Cluster 1 – Red had mean price range of 1.62 and rating spread around 3.95

Results and Discussions

- There is a concentration of venues around Central Kolkata and Salt Lake City
- Majority of the venues have ratings between 3.4 and 4.2
- Majority of the venues have a price range of Rs 200 to Rs 400 for one person
- The first cluster has few venues spread throughout the city as well as concentrated around Central Kolkata which are relatively high priced and have an average rating of 4.06
- The second cluster has many venues concentrated around Salt Lake City which are moderately priced and have average rating of 3.95.

Conclusion

- If one is looking for cheap places with moderately high ratings (close to 4.0), one should check out **Salt Lake City**.
- If one is looking for the best places, with the highest rating but might also carry a relatively higher price tag, one has to visit **Central Kolkata, as well as may explore parts of North and South Kolkata**.
- A company can use this information to build up an online website/mobile application, to provide users with up to date information about various venues in the city based on the search criteria (name, rating and price).