IBM APPLIED DATA SCIENCE CAPSTONE BY COURSERA

Opening a New Shopping Mall in Kolkata, India

By: Barnendra Mohan Chaudhuri

January 2020

BUSINESS PROBLEM

- Location of the shopping mall is one of the most important decisions that will determine whether the mall will be a success or a failure

- Objective: To analyse and select the best locations in the city of Kolkata, India to open a new shopping mall

- This project is timely as the city is currently suffering from a disproportionate distribution of shopping malls which leads over-congestion is some parts of the city, traffic management issues and adequate public transportation issues

- Business question
  - In the city of Kolkata, India, if a property developer is looking to open a new shopping mall, where would you recommend that they open it?

DATA

- Data required
  - List of neighbourhoods in Kolkata
  - Latitude and longitude coordinates of the neighbourhoods
  - Venue data, particularly data related to shopping malls

- Sources of data
  - Wikipedia page for Kolkata neighbourhoods (https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Kolkata)
  - Open Cage Geocoder package for latitude and longitude coordinates
  - Foursquare API for venue data

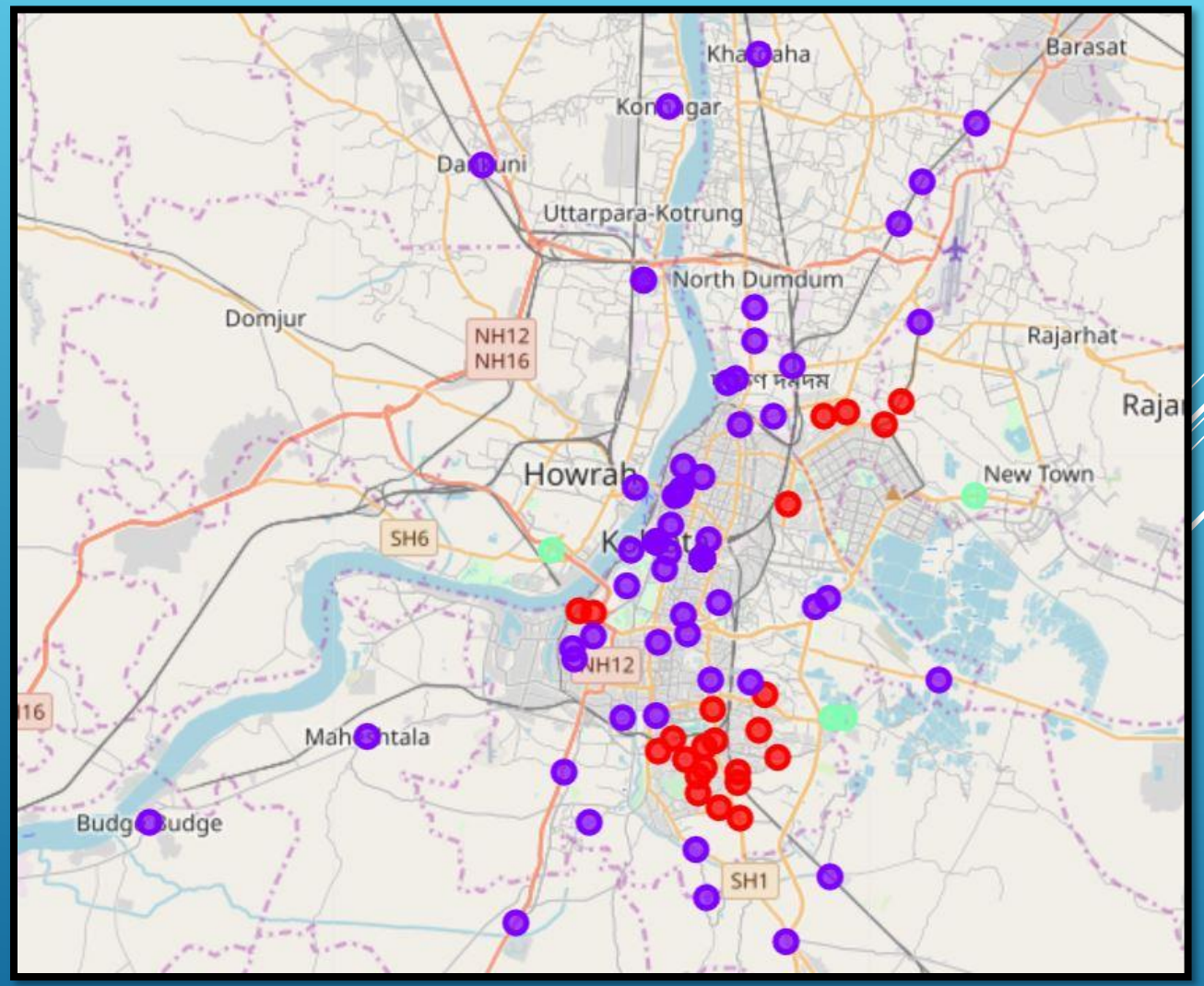
METHODOLOGY

- Web scraping Wikipedia page for neighbourhoods list
- Get latitude and longitude coordinates using OpenCage Geocoder
- Use Foursquare API to get venue data
- Group data by neighbourhood and taking the mean of the frequency of occurrence of each venue category
- Filter venue category by Shopping Mall
- Perform clustering on the data by using k-means clustering
- Visualize the clusters in a map using Folium

RESULTS

Categorized the neighbourhoods into 3 clusters :

- Cluster 0: Neighbourhoods with moderate number of shopping malls (red color dots)
- Cluster 1: Neighbourhoods with low number to no existence of shopping malls (purple color dots)
- Cluster 2: Neighbourhoods with high concentration of shopping malls (mint green color dots)



DISCUSSION

- Shopping malls are highly concentrated in 4 neighborhoods which form cluster 2. They are Anandapur, Bidhannagar, Jetia, Kasba

- Moderate concentration of shopping malls in cluster 0 comprising of 23 neighborhoods

- Cluster 1 has very low number to no shopping mall in the neighbourhoods. This cluster comprise about 171 neighborhoods

RECOMMENDATIONS

- Open new shopping malls in neighbourhoods in cluster 1 with little to no competition
- Can also open in neighbourhoods in cluster 0 with moderate competition if have unique selling propositions to stand out from the competition
- Avoid neighbourhoods in cluster 2, already high concentration of shopping malls and intense competition

CONCLUSION

- Answer to business question: The neighbourhoods in cluster 1 are the most preferred locations to open a new shopping mall
- Findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a new shopping mall
- This will also help alleviate the traffic management and public transportation issues currently being faced by the city authorities

THANK YOU!