Coursera Capstone Project

IBM Data Science Professional certificate

PROBLEM STATEMENT

- Opening up a restaurant is an entrepreneur's dream or passion
- Developing and executing a solid business strategy for the restaurant is extremely important in order to make the business successful
- Objective: To analyse and select the best locations in the city of Bangalore, to open a new Chinese Restaurant in particular area
- So We will use our data analytics powers to generate a few most promising neighborhoods based on this criteria
- In this Project, we will try to find an optimal location for a restaurant. Specifically, this report will be targeted to stakeholders interested in opening a **Chinese restaurant** in **Bangalore**, India.

DATA

- Data required
  - List of neighbourhoods in Bangalore, India.
  - Latitude and longitude coordinates of the neighbourhoods.
  - Venue data, particularly data related to Chinese Restaurants.
- Sources of data
  - Wikipedia page for neighbourhoods (https://commons.wikimedia.org/wiki/Category:Suburbs of Bangalore)
  - Geocoder package for latitude and longitude coordinates
  - Foursquare API for venue data

METHODOLOY

- Web scraping Wikipedia page for neighbourhoods list.
- Get latitude and longitude coordinates using 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 Chinese Restaurants.
- Perform clustering on the data by using k-means clustering.
- Visualize the clusters in a map using Folium.

RESULTS

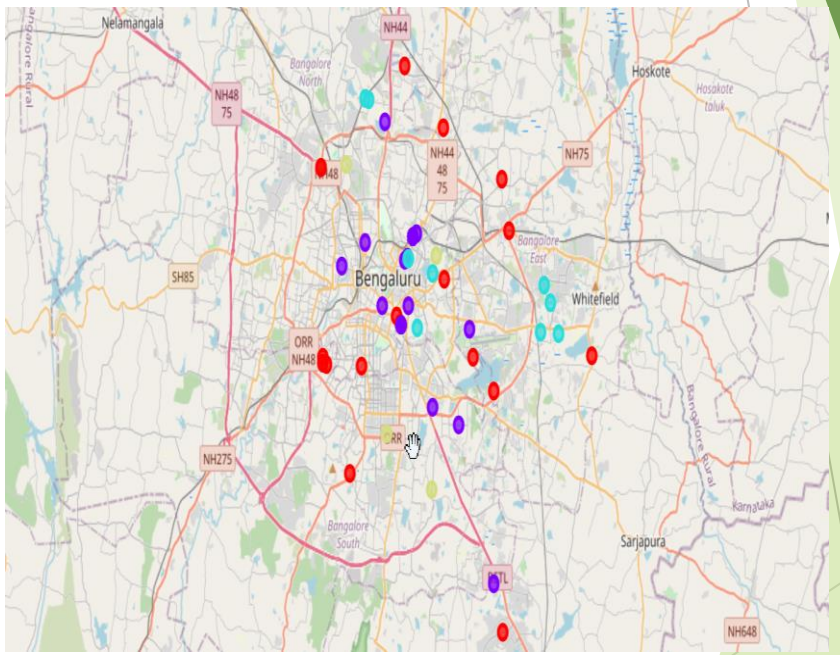
- **Categorized the neighbourhoods into 4 clusters :**

- Cluster 0:Neighbourhoods with low number to no existence of Chinese Restaurants

- Cluster 1:Neighbourhoods with almost equal concentration of Chinese Restaurants

- Cluster 2:Neighbourhoods with moderate concentration Chinese Restaurants

- Cluster 3:Neighbourhoods with high concentration of Chinese Restaurants



DISCUSSION

- Main focusing Chinese Restaurants around the Bangalore city
- Highest number in cluster 3 and moderate number in cluster 2
- Cluster 1 has equal focusing area where the all area is good for openning.
- Cluster 0 has very low number to no shopping mall in the neighbourhoods
- Oversupply of Chinese Restaurants mostly happened in the central area of the city, with the suburb area still have very few shopping malls
- Restaurants Investors are advised to avoid neighborhoods in cluster 3 and 2 which already have high concentration of shopping malls and suffering from intense competition.