Kochi Neighbourhood Analysis

REASON FOR ANALYSIS:

- KOCHI - A vibrant city in the heat of Kerala , a state in **India**

- Has great potential for growth in food and beverage industry.

- **AIM :** To find a location with prospective customers and a well

    established neighborhood to set up a coffee shop

- Helps reduce manual search and provides efficient solution.

DATA ACQUISITION AND PROCESSING

- Details about the various regions in the city was scrapped from the wikipedia page : https://en.wikipedia.org/wiki/Neighbourhoods_of_Kochi

- Four Square API : prime data gathering source with a large database on places, provides the ability to perform location search, location sharing and details about a business.

- Various Python libraries like Pandas, Folium, Beautiful Soup, Scikit learn, Geocoder for processing and providing result.

ANALYSIS APPROACH

➢ Scraping Wikipedia page to get major locations in Kochi
  ■ 38 Locations have been identified.

➢ Using Geocoder for getting coordinates of the locations.

➢ Foursquare API is used to get the various venues within 500m of the target location.
  ■ 270 venues were located.
  ■ Venue coordinates and category was identified.

- ➤ Created Dataframe was cleansed and normalised
  - ■ Index resseted
  - ■ Frequency of venue categories identified
  - ■ Along with most common venues

- ➤ Clustering of Data Points
  - ■ K Means Clustering Algo was used.
  - ■ Cluster number parameter set to 4

- ➤ Using Folium to display the clusters
  - ■ Displays various locations grouped by colored dots.

FOLIUM MAP WITH CLUSTERS
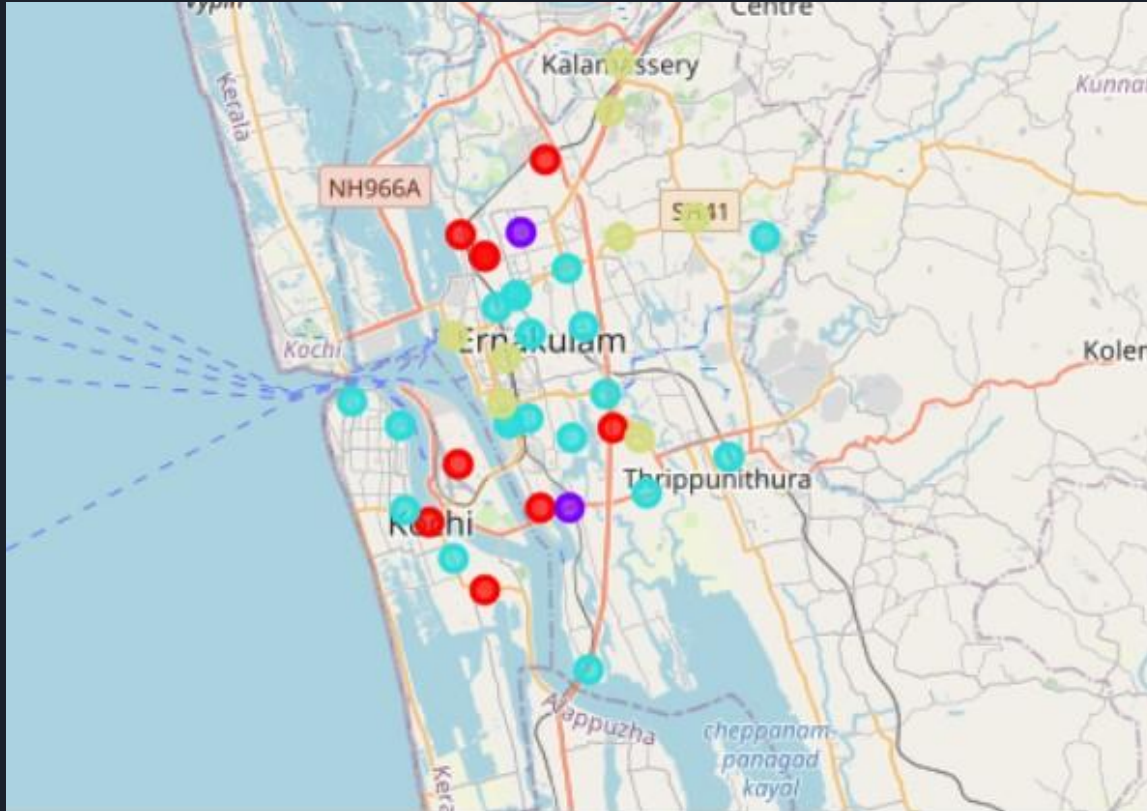
- Cluster 1
- Cluster 2
- Cluster 3
- Cluster 4

- ➤ Analysing various Clusters to find optimum cluster
  - ■ Each cluster is examined based on common venues of cluster elements , number of venues and their relevance

- ➤ Result is devised
  - ■ Optimal location is found from among the optimal clusters

RESULT & CONCLUSION

- The 4 Clusters have been analysed to find the below observation

  - **Cluster 1** - **A region with limited tourist footfall and more family oriented spaces.**
  - **Cluster 2** - **Region with limited access and few attractions and shops.**
  - **Cluster 3** - **Highly populated with extreme competition. A well developed and established location.**
  - **Cluster 4** - **Region with good attractions and on a developing phase. Has great potential.**

- It can be seen that from the various clusters created and analysed, Cluster's 3 and 4 provides good locations for setting up a cafe/coffee shop. Cluster 4 would be the most optimal solution for a new venture given it to be less competitive and developing. This is inferred mainly from the observations regarding the venues present there and also their respective categories. Any of the locations within the cluster is a suitable option.

- Other factors that have not been considered like transportation facilities , population density , land value can also be used to get a more accurate solution.

THANK YOU