COURSERA CAPSTONE
IBM APPLIED DATA SCIENCE CAPSTONE OPENING A NEW MULTIPLEX IN HYDERABAD, INDIA

BY,

SHARATH KUMAR V

BUSINESS PROBLEM

- Location of the Multiplex is one of the most important decisions to attract the right set of people example young crowd, students etc
- Objective: To analyse and select the best locations in the city of Hyderabad, India to open a new Multiplex
- This project is timely as the city is currently suffering from oversupply and undersupply of Multiplexes

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

DATA

- Data Required
  - ❖ List of neighbourhoods in Hyderabad, India
  - ❖ Latitude and longitude coordinates of the neighbourhoods
  - ❖ Venue data, particularly data related to Multiplexes
- Sources of data
  - ❖ Wikipedia page for neighbourhoods:
    https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Hyderabad,_India
  - ❖ Geocoder package for latitude and longitude coordinates
  - ❖ Foursquare API for venue data

METHODOLOGY
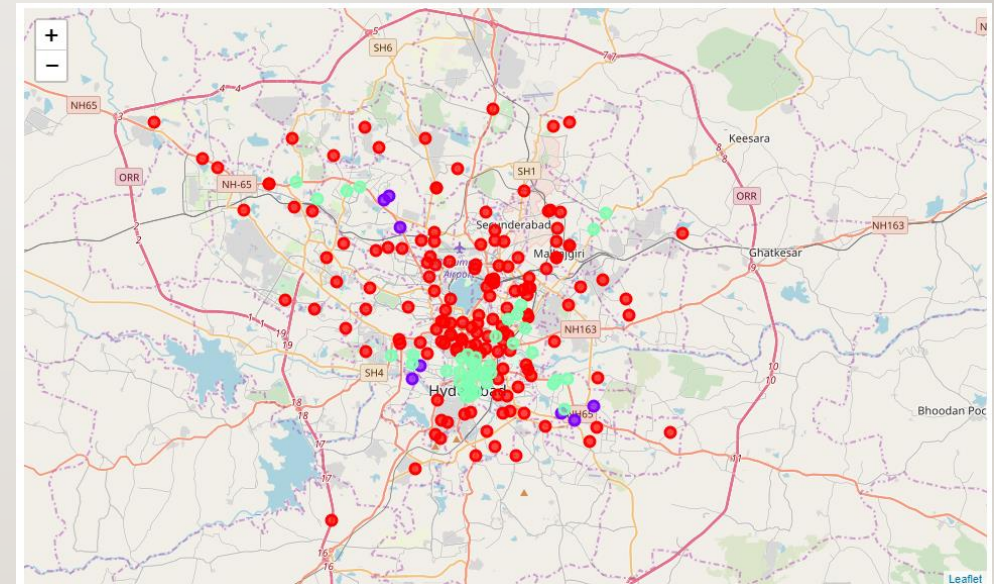
- Web scraping Wikipedia page for neighbourhoods list
- Get latitude and longitude coordinates using Geocoder
- Use Foursquare API to get venue data
- Group data by neighbourhood and taking the mean of the frequency of occurrence of each venue category
- Filter venue category by Multiplex
- Perform clustering 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

  - Cluster 1: Neighbourhoods with low number to no existence of shopping malls

  - Cluster 2: Neighbourhoods with high concentration of shopping malls

CONCLUSION

- Answer to business question: The neighbourhoods in cluster 0 are the most preferred locations to open a new Multiplex
- 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 Multiplex

THANK YOU
Q&A