COURSERA CAPSTONE
IBM APPLIED DATA SCIENCE CAPSTONE
SIMILARITY OF NEIGHBOURHOODS BETWEEN CHENNAI AND MUMBAI

By:Malavika Rajesh Vikraman

August 13th,2019

PROBLEM

- People from Mumbai migrating to Chennai and People from Chennai find it difficult to to find a neighbourhood similar to the one they were living in.Similarly , when you shift from one neighbourhood to another neighbourhood in Chennai or Mumbai itself you find it difficult to find a similar neighbourhood with somewhat similar amenities and facilities.

DATA

Required Data:

- List of Neighbourhoods in Chennai

- List of Neighbourhoods in Mumbai

- Latitude and Longitudes coordinates of these neighbourhoods. This is required in order to plot the map and also to get venue details

- Venue data, especially regarding all type of venues found in a particular neighbourhood, so that it is easy for clustering

Sources of Data:

- https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai

- This link gives us the major neighbourhoods in Mumbai city.

- https://en.wikipedia.org/wiki/List_of_neighbourhoods_of_Chennai

- This link gives us the major Neighbourhoods in Chennai city.

- Geocoder package for latitude and longitude and Foursquare API for venue data

METHODOLOGY

- Web scraping Wikipedia page for neighbourhood 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

- Perform clustering by k-means clustering

- Visualise the clusters in a map using Folium

RESULTS:

- **Clustering Of Mumbai**
- Cluster labels
- 0
- Basically consists of neighbourhoods having more Indian Restaurants,Resrtaurants, Ice Cream Parlors,Fast Food Restaurants
- 1
- Basically the neighbourhoods having more Juice Bars and Electronics Stores
- 2
- Basically the neighbourhoods having more Bakery and Department stores.
- 3
- Neighbourhoods having more Café's ,Trials and Coworking Spaces.
- 4
- Has Hot Dog Joints, Food trucks and pub more in these neighbourhoods
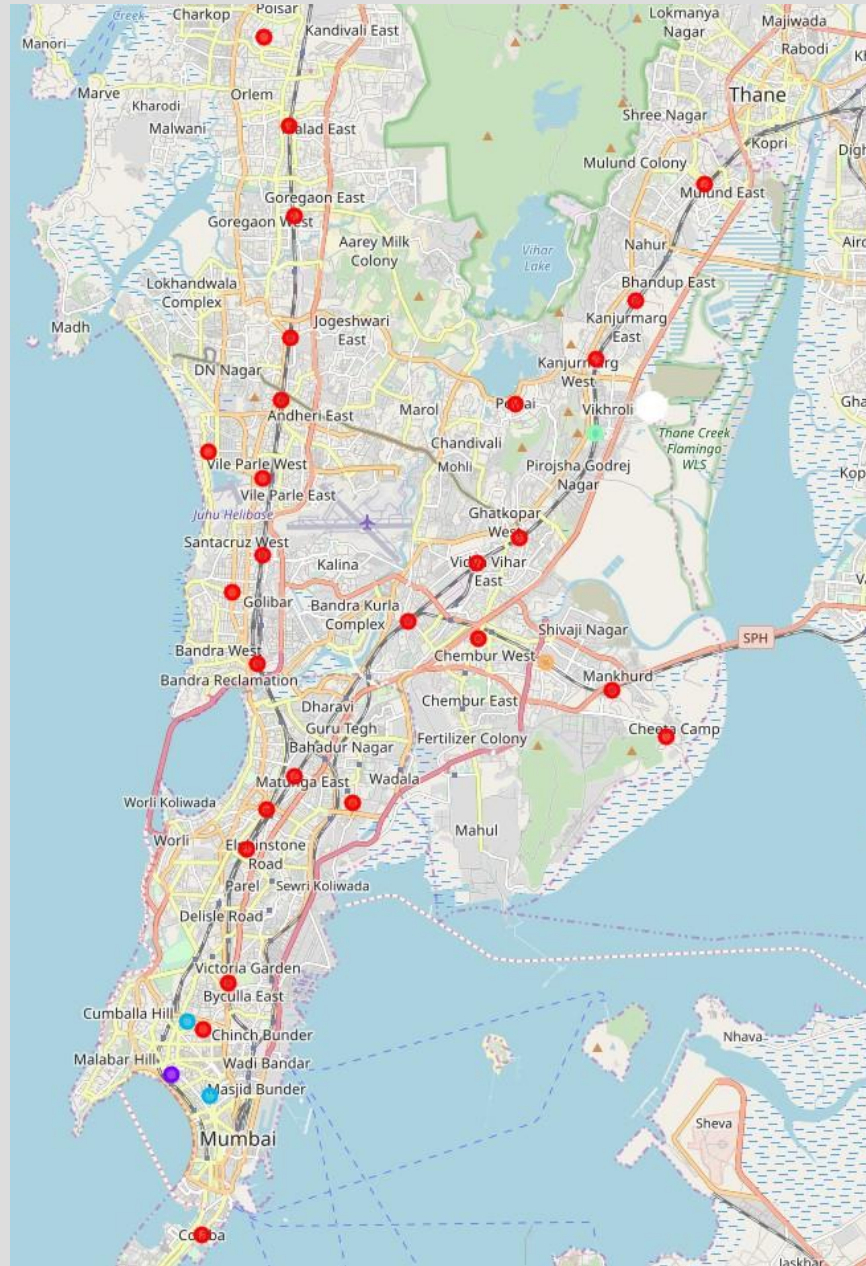
- **<u>Clustering of Chennai</u>**

- 0

The neighborhoods that have more resorts and Golf Courses are grouped in 0 cluster.

- 1

The neighbourhoods that have more Indian restaurants , Convenience Stores, Cosmetics Stores, Daycares with Bus Station/Train Station .
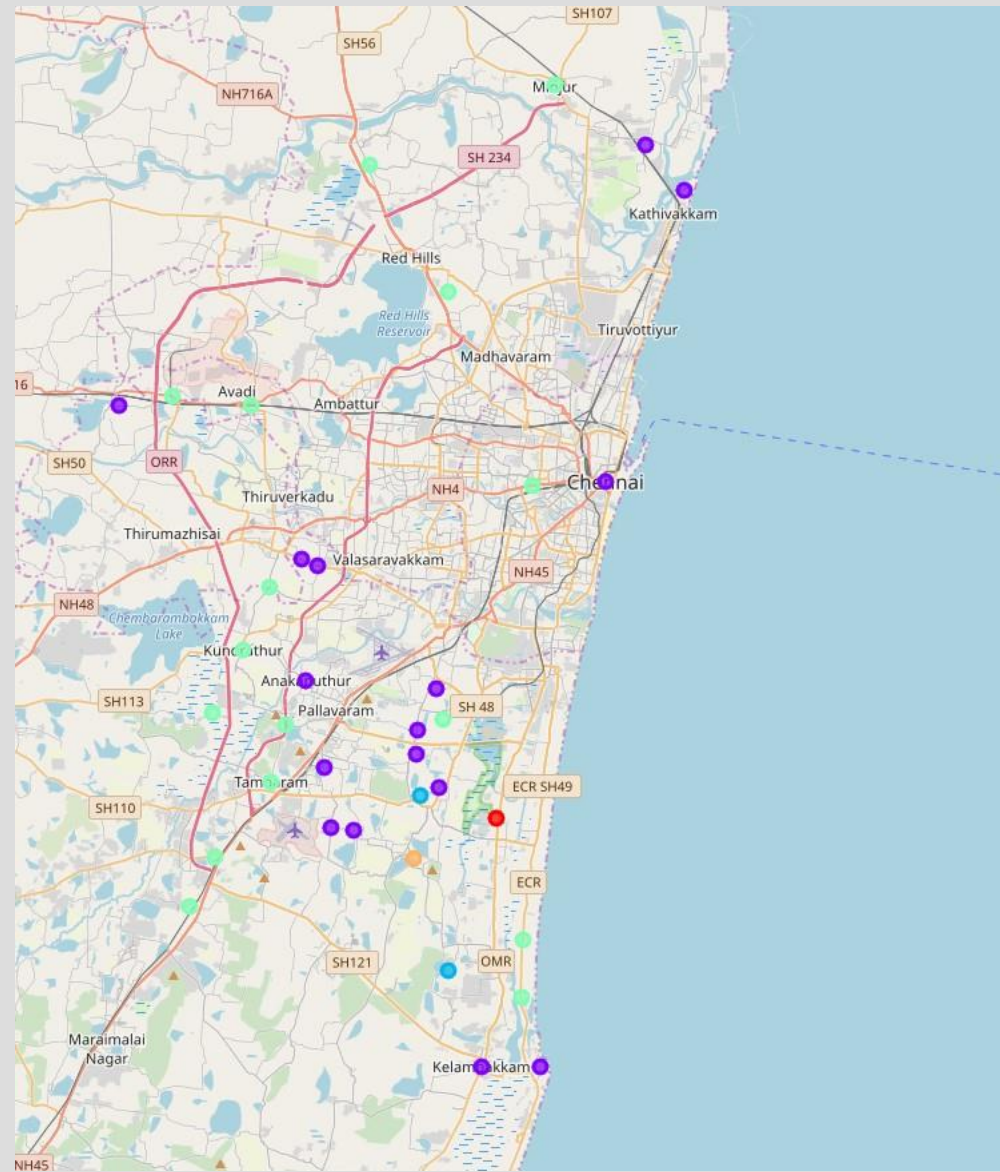
- 2

Neighborhoods having more Tea Stations and Indian Restaurants

- 3

Neighborhoods having Cosmetics stores, Convenience Stores , Daycare , Department Stores and Electronics Store more are grouped in 3.

- 4

Neighbourhoods having Bakery and Men's Store more are grouped together in this cluster.

- **<u>Clustering of Mumbai and Chennai</u>**
- 0

Neighbourhoods that contain Pizza Places, Food and Fast Facilities More

- 1

Neighbourhoods that contain Restaurants (different kinds) More.

- 2

Neighbourhoods that conatin Tea Rooms more

- 3

Neighbourhoods that contain Cafes and Multiplexes More.

- 4

Neighbourhoods that contain Tea Rooma and Men' Stores More.

- 5

Neighbourhoods that contain Resort and Golf Course more.

- 6

Neighbourhoods that contain Convenience stores , Cosmetic shops , Day care ,Department stores and Electronics Stores more.

DISCUSSIONS

- So since we have clusters formed now its easy for us to identify , where to shift towards.

- Like a person who's living in neighbourhood which in cluster 0 can shift to a similar neightbourhood by going for another neighbouhood of cluster 0.Sometimes we can see the clusters having only one neighbourhood which means that there will be no other neighbourhood that you can shift to with most facilities same.

CONCLUSION

In this project,we have gone through the process of identifying a problem , specifying the data required , extracting and preparing the data, performing machine learning by clustering the data into 5 to 7 clusters based on similarities and lastly providing information to people regarding how to shift from one neighbourhood to another without feeling much of a change.

THANK YOU !