Finding The Hotspot For Mobile Phone Expo - 2020
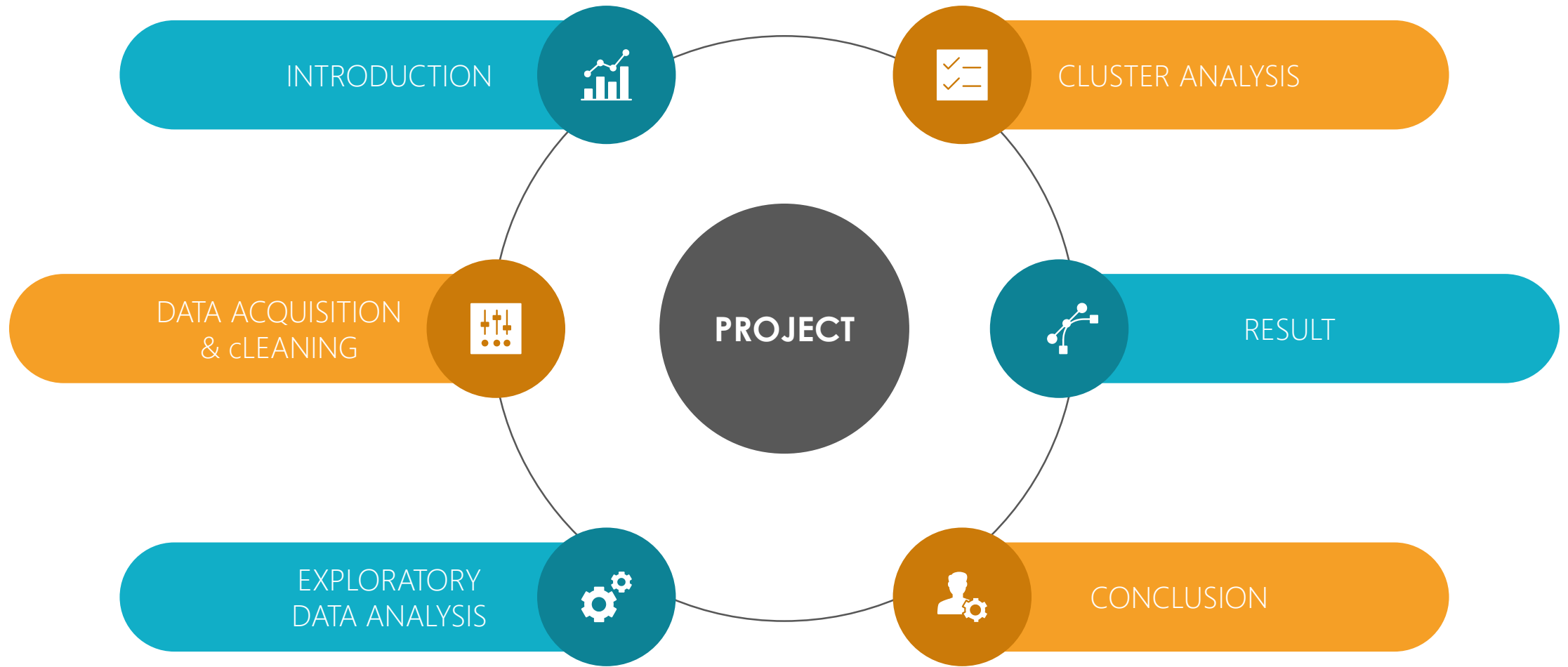
Iyyanar M
Jan 02, 2020

Advanced Data Science Capstone -Coursera

CONTENTS

INTRODUCTION

CLUSTER ANALYSIS

DATA ACQUISITION & cLEANING

PROJECT

RESULT

EXPLORATORY DATA ANALYSIS

CONCLUSION

Introduction

- A Mobile Phone Expo in India for the year 2020 has been planned and the location for the expo is still undecided.

- This Expo will act as a marketing campaign for major mobile phone manufacturers and local mobile stores around the location.

- Customers can purchase a new mobile phone and also they can get to know the local stores, service points nearby from the expo.

- The problem is to find the central hub-point, which should be located in middle of many mobile shops in a lesser distance. So, that expected amount of sponsors and customers around that area will be guaranteed to participate in the expo. Therefore, the cost of installing the expo can be reduced by more number of sponsors and profit of the expo will be increased.

Data Acquisition & Cleaning

- I have used the Foursquare API for collecting the mobile shop location data's which includes latitude and longitude values for 5 major cities in India.

- The Cities are :

  - Bangalore,

  - Chennai,

  - Delhi,

  - Hyderabad and

  - Mumbai.

- The reason behind these 5 Cities are that these cities are well known for their daily market trade and populations.

Data Acquisition & Cleaning

- The Data retrieved from the FourSquare API is in JSON format containing many information, which are not needed.

- The information only needed for this project are shop name, latitude, longitude and city name. By using python script, only the needed data columns are selected from the JSON results.

Fig. 1 – Screenshot of the Dataset.

	Name	Address	Lat	Lng	City
0	Ritche Street	Ritchie St	13.069477	80.271374	Chennai
1	Croma	2nd Ave,	13.084631	80.216891	Chennai
2	Reliance Digital Express	Anna Nagar	13.084756	80.213343	Chennai
3	Vodafone Store	Anna nagar	13.085272	80.203079	Chennai
4	Aircel Store	Aircel Towers, No.301,Poonamalle High Road,Kil...	13.075956	80.236309	Chennai

Exploratory Data Analysis

DATA VISUALIZATION :

- The Dataset which we collected needs to be analyzed. Since, it consists of latitude and longitude values. We can plot it in a graph for analysis.

- Data visualization is done using folium libraries.

Fig. 2 – Screenshot of the Chennai and Delhi Dataset.
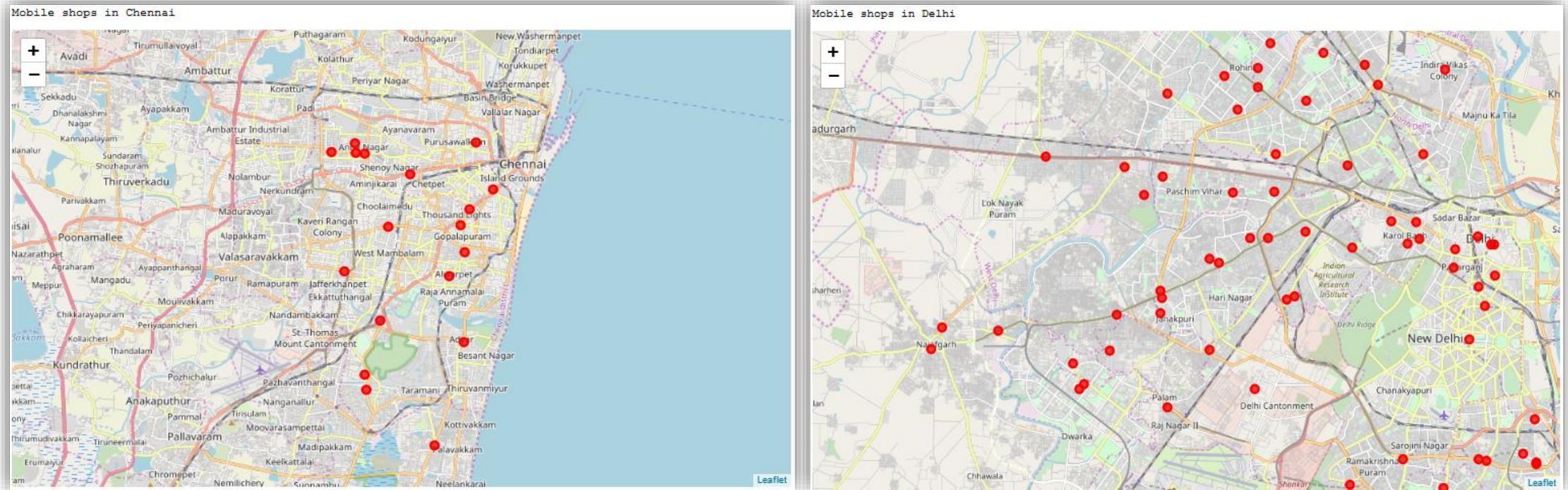
Exploratory Data Analysis

DISTANCE FRO MEAN POINTS IN EACH CITY :

• The calculation of the distance from a central point of those mobile shops in each city would be helpful in identifying the best city with higher number of mobile shops and also helps to find the average distance to be travelled for the Expo-2020.

• These centre points are calculate by the mean average of the latitude and longitude values of each cities.

Fig. 3 – Distance from Center point in each Cities.

```
Chennai
Distance from the centre point : 0.05105306360595714
Hyderabad
Distance from the centre point : 0.043075877532875756
Bangalore
Distance from the centre point : 0.035131045589305084
Mumbai
Distance from the centre point : 0.08167181371420372
Delhi
Distance from the centre point : 0.09574720569597839
```

Cluster Analysis

K- Means Clustering Algorithm :

- K-Means algorithm is an unsupervised learning algorithm.

- It clusters, partitions or segments the available data points by means of similar characteristics based on labels.

- The available data points are divided into non-overlapping subsets called clusters without any internal cluster structure.

- Algorithm :

  1. Randomly placing k Centroids, one for each cluster

  2. Calculate the distance of each point from each centroid

  3. Assign each data point to its closest centroid, creating a cluster

  4. Re-calculate the position of the k centroids

  5. Repeat the steps 2-4, until the centroids no longer move.

Cluster Analysis

- After applying the k-means algorithm to the data points for delhi city with a *k* cluster value. We got the cluster labels for each row in the delhi dataset.

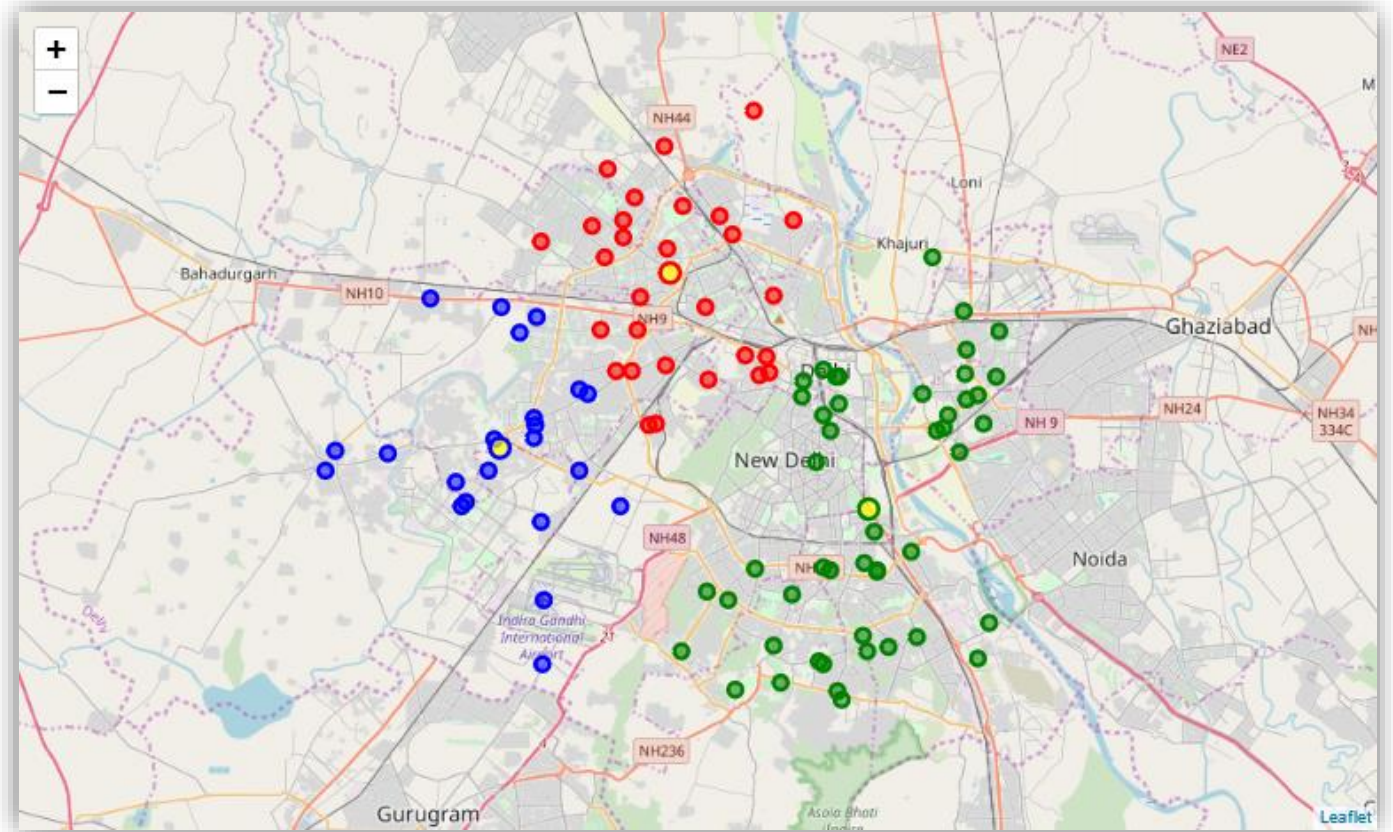- By using this cluster label, we can find the center point for each cluster to start expo in that place.

Fig. 4 – Delhi data with their Cluster labels after applying K-means algorithm

	Cluster Labels	Lat	Lng	Name
0	0	28.683666	77.196670	My Idea Store
1	1	28.535816	77.196995	My Idea Store
2	0	28.713319	77.109163	My Idea Store
3	1	28.649738	77.304540	My Idea Store
4	1	28.641635	77.295433	Trendy cover for HTC

Result

- The Result is the center hub point for each segmented cluster in Delhi.
- There are 3 clusters of mobile shops located in this city and the center mid-point for each cluster can be used for starting the Expo-2020.

- *The three clusters are plotted in the graph having a differentiation of markings in colors as red, green, and blue.*
- *. Each cluster has its own center points highlighted with yellow color.*

Fig. 5 – Cluster of Mobile Shops in Delhi with Central Hub point (Hotspot)

Conclusion

The project work for finding the hotspot location to start a Mobile Expo-2020 in India has been found as Delhi. Therefore, the distance from the mobile shops to the expo will be lesser and will result in participation of higher number of mobile stores, sponsors, and customers. Which will increase the popularity and sales profit of all the mobile stores around Delhi.

Thank You