Opening a Shopping Mall in suburbs of Mumbai, India

Capstone Project - Applied Data Sciences by IBM

By: Shulin Saraswat

Content

- Introduction
- Business Problem
- Data
- Methodology
- Results
- Discussion
- Conclusion

Introduction

- Visiting shopping malls is a good way to enjoy, have fun and relax during holidays or weekends.
- Grocery shopping, dining, watching movies, etc. some of the activities that a shopping mall provides to its customers.
- Building new malls allows the developer to earn consistent rental income as well as retailers that buy shops in the mall earn from providing from channelling their products and services to the market.
- Location of the mall place a critical role to determine the success or failure of the mall.

Business Problem

- The aim of this capstone project is to analysis and suggest best locations in Mumbai for the developers to build a new mall.
- Using different data science techniques like clustering, this project aims to answer the business problem.
- In the city of Mumbai, if a property developer is looking to open a new shopping mall, where would he/she be recommended to open it ?

Data

- Data Required:
  - List of suburbs neighbourhood in Mumbai
  - Latitude and Longitude coordinates of the neighbourhood.
  - Venue data related to shopping malls.

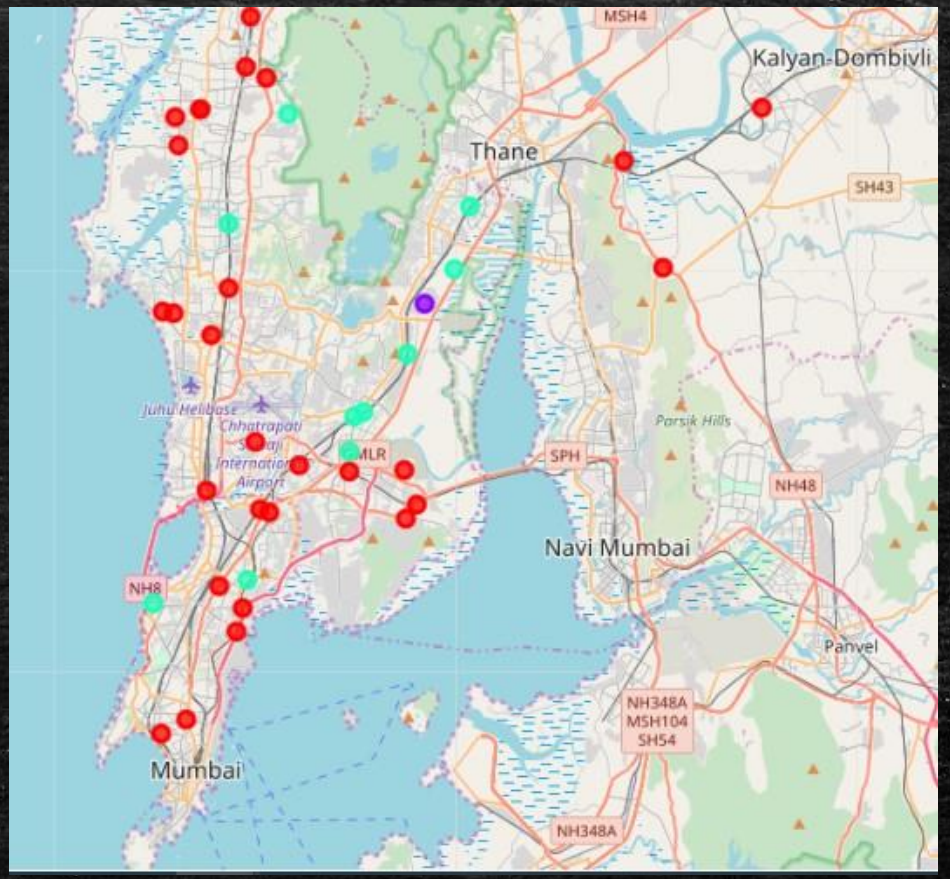
- Source of data
  - Wikipedia page https://en.wikipedia.org/wiki/Category:Suburbs_of_Mumbai
  - Geocoder package for longitude and latitude.
  - Foursquare API for venue data

Methodology

- Web scarping the Wikipedia page.
- Get the coordinates using geocoder.
- Using Foursquare API to get venue data.
- Grouping the neighborhood on the basis of the number of occurenece of the Shopping mall.
- Clustering the data using K-means clustering.
- Visualizing the cluster on the map of Mumbai.

Results

- The results from k-mean clustering show that we can group the neighbourhood into 3 clusters on the basis of the number of occurrence for Shopping mall
- Cluster 0: It has the cluster 0 has a low number of shopping malls.
- Cluster 1: It has moderate number of shopping malls.
- Cluster 2: It has the highest number of shopping malls.



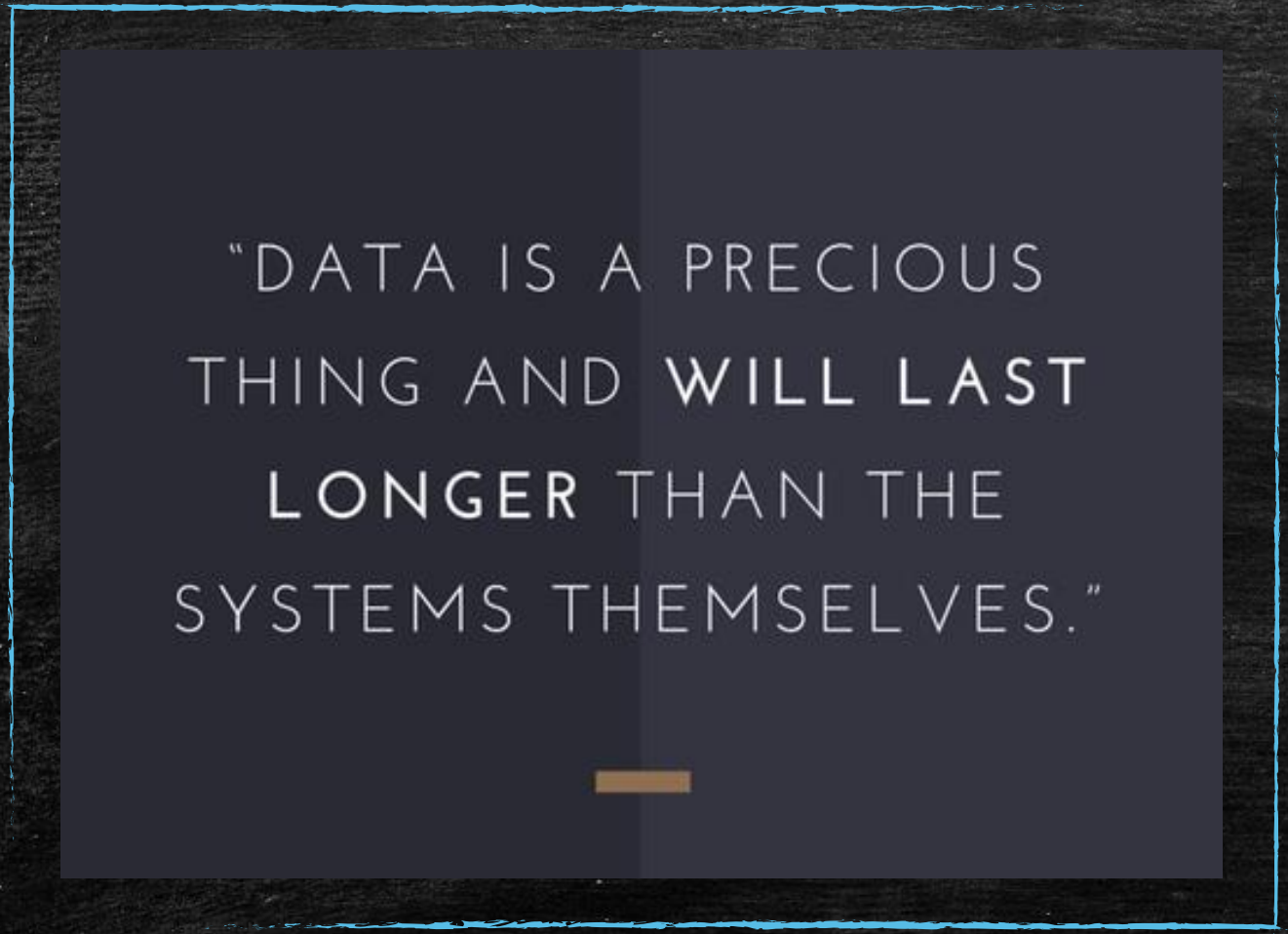
Discussion

- From the observation from the developed results we can say that most shopping malls in Mumbai are in cluster 2. Cluster 0 has very low number of malls in neighborhood.
- This represents a great opportunity and high potential areas to open new shopping malls as there is very little to no competition from existing malls.
- Meanwhile, shopping malls in cluster 2 are likely suffering from intense competition due to oversupply and high concentration of shopping malls.
- Property developers are advised to avoid neighborhoods in cluster 2 which already have high concentration of shopping malls and suffering from intense competition

Conclusion

- In this project, we were able to identify a business problem and provide answers to the problem by using data science methodology and techniques such as clustering.
- The answer to the problem raised would be a great opportunity and high potential areas to open new shopping malls as there is very little to no competition from existing malls in cluster 0.

Thank You !



Capstone project by:

Shulin Saraswat