IBM Applied Data Science

Coursera Capstone Project

Opening a new Shopping mall in Bengaluru

By – utkal bharti

Introduction:

The concept of shopping mall came into existence when people started demanding all products at one place. A shopping mall is a modern, chiefly North American, term for a form of shopping precinct or shopping center in which one or more buildings form a complex of shops with interconnecting walkways, usually indoors. It is a place they could go and hangout with their friends and be relatively safe. Shopping malls are like a one-stop destination for all types of shoppers. For many shoppers visiting shopping mall is a great way to relax and enjoy themselves during weekends and holidays. Property dealers are also taking advantage of this trend to build more shopping malls to cater the demand. There are many shopping malls in the city Bengaluru and many more are being built. Now a days shopping mall is a great source of income as well as for generating employement. Opening a new shopping mall requires serious consideration. Particularly the location of shopping mall is one of the most important decision that will determine whether the mall will be successful or a failure

Business Problem:

The objective of this capstone project is to analyse and select the best locations in the city Bengaluru, India to open a new shopping mall. If a property dealer wants to open a new shopping mall in Bengaluru , where would we recommend them to build, this answer will be given by this project. We will be using Data science methodology and machine learning techniques like clustering in this project.

Data Set

We need following data set to solve the problem:

- List of neighborhoods in Bengaluru, the capital of state Karnataka , India which is situated at south east of Asia.
- Latitude and longitude coordinates of neighborhoods of the city Bengaluru.
- Coordinates are required to plot the map of the city Bengaluru and also to get the venues.
- Venue data, particularly data related to shopping malls. We will use this data set to perform clustering on the neighborhoods.

Source of data and methods to extracts them:

This Wikipedia page (”https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Bangalore”) contains the list of neighborhoods in Bengaluru,with a total of 129 neighborhood .We will use web scrapping techniques to extracts the data from the Wikipedia page with the help of python request and beautiful soup packages. Then we will use Python geocoder package and with the help of this package we get the geographical coordinates of the neighborhoods of Bengaluru and also with the help of this package we will get the latitude and longitude of neighborhoods.

After that we will use Foursquare API to get the venue data for those neighborhoods. Foursquare has one of the largest database of 105+ million places and is used by over 125,000 developers. Foursquare API will provide many categories of the venue data , we are particularly interested in the shopping mall category in order to help us to solve the business problem . This is a project that will make use of many data science skills , from web scrapping(Wikipedia) and map visualization(folium).