

# **Coursera Capstone Project**

## **IBM Applied Data Science Capstone**

**Identify new place for a shopping mall in Bangalore, India**

***By : Ravi Teja Jalagam***

***April 2020***



## **Introduction :**

For many shoppers, visiting shopping malls is a great way to relax and enjoy themselves during weekends and holidays. They can do grocery shopping, dine at restaurants, shop at the various fashion outlets, watch movies and perform many more activities. Shopping malls are like a one-stop destination for all types of shoppers. For retailers, the central location and the large crowd at the shopping malls provides a great distribution channel to market their products and services. Property developers are also taking advantage of this trend to build more shopping malls to cater to the demand. As a result, there are many shopping malls in the city of Bangalore. Opening shopping malls allows property developers to earn consistent rental income. Ofcourse, as with any business decision, opening a new shopping mall requires serious consideration and is a lot more complicated than it seems. Particularly, the location of the shopping malls is one of the most important decisions that will determine whether the mall will be a success or a failure.

## **Business Problem :**

The objective of this capstone project is to analyse and select the best locations in the city of Bangalore, India to open a new shopping mall. Using Data Science methodology and Machine Learning techniques like clustering, this project aims to provide solutions to answer the business question: In the city of Bangalore, India if a property developers is looking to open a new shopping mall, where would you recommend that they open it?

## **Target Audience of the project :**

This project is particularly useful to property developers and investors looking to open or invest in new shopping malls in the capital city of Karnataka i.e Bangalore. This project timely helps as the city is currently suffering from oversupply of shopping malls.

## **Data :**

### **To solve the problem, we need the following data**

1. List of Neighbourhoods in Bangalore city.
2. Latitude and Longitude of those Neighbourhoods. This is required to plot the map and also to get the venue data.
3. Venue data, particularly data related to shopping malls. We will use this data to perform clustering on the neighbourhoods.

### **Sources of data and methods to extract them :**

This Wikipedia page [https://en.wikipedia.org/wiki/Category:Neighbourhoods\\_in\\_Bangalore](https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Bangalore) contains a list of neighbourhoods in Bangalore, with a total of 67 neighbourhoods. We have already gathered data of Neighbourhoods and their geographical coordinates in a csv file. We will be creating a dataframe from a csv file.

After that, we will use FourSquare API to get venue data for those neighbourhoods. Foursquare has one of the largest databases 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 put forward. This is a project that will make use of many data science skills, from web scraping(Wikipedia), working with API(Foursquare), data cleaning, data wrangling to Machine learning(K-means Clustering) and map Visualization(Folium).In the next section, we will present the methodology section where we will discuss the steps taken in this project, the data analysis that we did and the machine learning technique that was used.