**Coursera Capstone Project Of Neighbourhoods**

**IBM Applied Data Science Capstone**

****

**Introduction:**

A **shopping mall** is a specially built covered area containing **shops**

and restaurants which people can walk between, and where cars are

not allowed. In any city Shooping malls are the best way for people

to relax and for shopping in Weekends or Holidays. People can chill

and relax in Shopping malls .Malls are different from bazaars where

the shops are not tiny booths. In Malls each shop has its own space.

Shopping malls consists of shops i.e fashion stores, Theatres and

What not. In Shopping malls people can enjoy relax and they can do

Shopping. So, the shopping mall should be located in the centre of

city or the place with good neighborhoods. Opening shopping malls

should make developers to earn their consistent rental income.

Literally, the location of Shopping mall plays a crucial role in deter-

Ming whether it is success or not. So, after the great scrutiny of

Neighborhoods the location of Shopping mall should be decided.

**Business Problem:**

The Ultimate aim of this **Capstone** is to select the best Locations in

The city of Banglore, India to open a new Shopping mall. Using the

Data science methodology and Machine learning techniques like

**Clustering K-means .** The main aim of this capstone is to answer the

Business question is:

If the some property Developer wants to build a shopping mall in the

City of **Banglore, India,** where would you suggest to open it?

**Target audience of this Project:**

This project is mainly useful for the property developers and the

Investors who are looking to open new shopping malls in Banglore.

**Data Required:**

->List of Neighborhoods in Banglore,India. This defines the scope of

Project which is confined to the city of Banglore, this city is the third

Most populated city in India.

->Latitude and Longitude from geopy which is used to get the venues

and plot in the Map.

->Venues or venue data which is used in to performclustering of the

neighborhoods.

**Source of Data and Methods for Extracting Them**

This Wikipedia(“<https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Bangalore>”)

page consists of nearly 140 neighborhoods in Banglore. By using web

scraping and Beautiful soup and also Python requests we will extract

the list of Neighborhood. Then we will get the geographical co-ordina

tes of the neighborhoods such as Latitude and Longitudes with the

help of python geocoder package.

After that, we will use Foursquare API to get the venue data for

those neighbourhoods. Foursquare has one of the largest database

of 105+ million places and is used by over 125,000 developers. Four

square API will provide many categories of the venue data, we are

particularly intrested 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.