Capstone Project (The Battle of Neighborhoods)

IBM Data Science

By Rupali Panda

Photo source: https://en.wikipedia.org/wiki/Bangalore

Introduction

In metro cities, visiting a shopping malls is a great way of relaxation for this generation. They can do shopping (including grocery, clothes and home essentials), dine at restaurants, go to movies and even kids can have their own play area.

Bangalore is one of the biggest cities of India, also known as the IT city of India. It has a population of over ten million and it is either fourth- or fifth-most productive metro area. Thus opening a shopping mall is not only a lucrative business for the property developers but also will be very useful for the customers. However, opening a shopping mall requires lot of strategies and planning. Particularly, the location of the shopping mall is the most important decision as it will impact the success of the mall.

Business Problem

The property developers needs to review and then strategically target so that the return on the investment will be profitable. The objective of this project to analyze and select the best locations in the Bangalore neighbourhood to open a new shopping mall. For this analysis, Data science methodology and machine learning techniques such as clustering were used.

Target Audience

This project will particularly be very useful for the property developers who are looking for opening a new shopping mall in Bangalore metropolitan area. Moreover, it will also be very helpful for the investors/business who can open the stores inside the mall.

Data

The following data are required for this project:

- List of all the neighbourhoods in Bangalore. This is needed to determine the location of the shopping mall.
- Latitude and Longitude of the neighbourhoods. It is required to plot the map and to get the venue data.
- Venue data. It will be used to perform the clustering of the neighbourhoods.

Data Sources and Data Science Methodology:

- The data will be collected by using the web scarping techniques from the Wikipedia page: https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Bangalore.
- Then the Python Geocoder package will be used to get the geographical co-ordinations of the Bangalore neighbourhoods.
- We will use the Foursquare API to get the venue data for the neighbourhoods.
- Finally to analyze each neighbourhood and clustering the k-means clustering is used.