

Applied Data Science Capstone

By IBM

Opening a new mall in the city of
Bangalore

The Battle of Neighbourhood

REPORT

By Shoban Dinesh

1. Introduction

The objective of this capstone project is to analyse and select the best location in the city of Bangalore to open a new shopping mall. Using data science techniques, the project aims to provide a solution to answer the business question: where would you recommend a property developer to open a new shopping mall?

The project will use data visualization techniques to enable the stakeholder visualize the locations of the neighbourhoods geographically. And also use machine learning to group neighbourhoods in a way to allow a property developer can a better choice in picking a neighbourhood to open a new mall.

2. Data

The following data were required to solve the problem:

- List of all neighbourhood in the city of Bangalore
- The latitude and longitude coordinates of each of the neighbourhood.
- Venue data and data related to the shopping mall in these neighbourhood.

Sources and Methods:

The list of all the neighbourhood was taken from the Wikipedia page(https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Bangalore) were all the neighbourhoods were mentioned.

I used web scraping techniques to extract data which the website, with the help of Python's libraries namely BeautifulSoup and requests.

The latitudes and longitudes of the individual neighbourhoods were obtained by using the Python Geocoder package.

And for the venue data of each neighbourhood, I used Foursquare API. Foursquare API will provide many categories of the venue data, but the category I needed to solve the business problem was 'shopping mall' data particularly.