IBM Data Science Capstone

Identification of Prospective Location for New Shopping Mall in Bengluru, India

By:

Deepro Sengupta

Table of Content

Contents

Introduction

For many shoppers, visiting shopping malls is a great way to relax and enjoy themselves during weekends and holidays. They can shop for everything ranging from groceries, clothes; electronics as well as Fashion related articles such as accessories, etc. Furthermore, they also provide other entertainment options like fast food, fine dining, movie theatres etic all under one roof. For retailers, the large crowd that it attracts enables them to sell their products and realize their investments quickly as well as have a very less consumer acquisition cost. Additionally, shopping malls also provide a great investment opportunity for property developers as well. They can choose to sell each shops within the Mall or rent it out. Therefore, opening a shopping mall is a good business decision. However, like any business decision, it should be supported with serious consideration. Particularly, location of a mall can be the difference between its success or failure.

Business Problem

The objective of this capstone project is to analyse and select the best locations in the city of Bengaluru, India for new shopping mall. Using data science methodology and machine learning algorithm like clustering, this project aims to provide solutions to answer the business question: In Bengaluru, what should be the location for a new shopping mall?

Target Audience

This project is particularly useful to property developers and investors looking to open or invest in a new shopping mall in the city of Bengaluru in Karnataka, India. This project, however, can be adapted for other purposes as well with minor code modification. This is due to the fact that this project already collects massive amount of data related to the top 100 local venues in Bengaluru before directing our attention towards shopping malls in these neighbourhoods. Therefore, if one wanted to identify a location to open a new food joint, for example, it can be done just by a few minor changes on the code.

Data

In order to work on our problem, we will need the following data:

* Names of neighbourhoods in Bengaluru.

* The geographical co-ordinates of each neighbourhood.

* All shopping malls in Bengaluru

In order to acquire these data, we have used the following sources:

* Foursquare API

* Google Maps Reverse Geocoding

Data Source, extraction and usage

We used the Python Pandas library's pandas.read_html to extract all tables containing a list of all neighbourhoods in Bengaluru, India. Then we used the Python Geoencoder package to get the latitude and longitudinal coordinates of the neighbourhoods.

Following this, we will use Foursquare API to get the data for the top 100 venues in those neighbourhoods. In order to do so, we will use the latitudinal and longitudinal data which we have acquired using the Geocoder package and use the Request library to long with our Foursquare API credentials to get the necessary data as a JSON file. Foursquare has one of the largest database of 105+ million places and is used by over 1,25,000 developers. Although Foursquare API provides many categories of venue data, we will direct our attention to the Shopping Mall category in order to help us solve the aforementioned business problem.