COURSERA CAPSTONE PROJECT

IBM Applied Data Science Capstone

Deciding Best Place to Open New Shopping Centre in Mumbai , India

By : Mayur Rindhe

August , 2019

Introduction

Shopping centre are buildings form a complex of shops with interconnecting walkways, usually indoors. In the 21th century, with the rise of the suburb and automobile culture in the India, a new style of shopping centre was created away from downtown.

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 Mumbai and many more are being built. Being the economic capital of India , Mumbai got its own uniqueness .

Opening shopping Centre allows property developers to earn consistent rental income. Of course, as with any business decision, opening a new shopping Centre requires serious consideration and is a lot more complicated than it seems. Particularly, the location of the shopping centre is one of the most important decisions that will determine whether the Centre will be a success or a failure.As India is a place of opportunity and if somebody pick the right thing at right time at right location then there is nothing will be there to stop him.

Business Problem :

The objective of this capstone project is to analyse and select the best locations in the city of Mumbai , Maharashtra (India) to open a new shopping Centre. Using data science methodology and machine learning techniques like clustering also using web scrapping , this project aims to provide solutions to answer the business question: In the city of Mumbai , India if a property developer , builders , foreign investors are looking to open a new shopping Centre, where would you recommend that they open it?

Target Audience of this project :

This project is particularly useful to property developers and investors looking to open or invest in new Property like Shopping Cenre in the economic capital city of India i.e. Mumbai. This project is timely as the city is currently suffering from oversupply of shopping malls.

This project may also help the government agencies to decide and take control over development in city .

Data :

To solve the problem, we will need the following data:

• List of neighbourhoods in Mumbai city. This defines the scope of this project which is confined to the city of Mumbai, the Economic capital city of the country of India.

• Latitude and longitude coordinates of those neighbourhoods. This is required in order to plot the map and also to get the venue data.

• Venue data, particularly data related to shopping Centres. We will use this data to perform clustering on the neighbourhoods.

Sources of data and methods to extract them :

This Wikipedia page mentioned in the following bracket (https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Mumbai) contains a list of neighbourhoods in Mumbai, with a total of 50 neighbourhoods. We will use web scraping techniques to extract the data from the Wikipedia page, with the help of Python requests and beautifulsoup packages.
Then we will get the geographical coordinates of the neighbourhoods using Python Geocoder package which will give us the latitude and longitude coordinates of the neighbourhoods.
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. Foursquare API will provide many categories of the venue data, we are particularly interested in the Shopping Centre 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.