Coursera Capstone

IBM Applied Data Science Capstone

Opening a New Shopping Mall in Mumbai, India

By: Aditya

March 2021

Introduction

A shopping mall is a large single building or multiple interconnected buildings housing a great number and variety of shopping units that are all commonly accessible. The concept of gigantic malls as we know them today are a twentieth century phenomenon, predominantly conceptualised in the west in the second half of the last century. Over a period of time, these gained huge popularity in cities as they combined retail diversity, social experience and a place to get together. The concept caught on and proliferated across the globe. Asia readily accepted the idea and some of the largest malls that got built are now located in Asia. "China's New South China Mall in Dongguan stands at the top of the heap with 2.9 million square meters of space."

The Great Indian Mall Boom began innocuously in the early 2000s, with just three malls in existence in the entire country. Nineteen years later, the number has mushroomed to a whopping 650 malls, enabled by the fact that the Indian audience discovered clean, vibrant, climate-controlled and technology-enabled malls, replete with a plethora of brands. This organised retail space is dedicated to shopping, entertainment, dining and cinema. The shopping mall culture has well and truly gripped the nation. In fact, as per an ANAROCK report, more than 30 new shopping malls covering nearly 14 million sq. ft. of area are expected to come up across top eight cities by 2020.

"Rapid urbanisation, digitisation, increasing disposable incomes and lifestyle changes in the middle-class society are leading to a major revolution in the Indian retail sector, which is pegged to grow by 60 percent to reach US$ 1.1 trillion by 2020. Cities that have seen maximum malls include Gurgaon, Noida, Greater Noida and Delhi in NCR, Mumbai, Chennai, Bengaluru and Pune. Over the next 5 years, nearly 85 malls are expected to come up in India," explains MD & CEO of ANAROCK Retail.

According to a report released by Boston Consulting Group in (BCG) 2017, though 70 percent of nearly 90 million online shoppers in India are influenced by the information they glean from internet, only 16 percent of them actually end up buying online. This means that purely offline pathway remains dominant, accounting for 78 percent of purchases and 58 percent of value. Loosely translated, this means that while most Indians research and compare products online, they eventually need to touch and feel a product, so wind up going to a store to buy what they need. And this is what makes the mall industry click in the country.

Opening shopping malls generates continuous source of rental income for the developer/investor. As with any business decision, opening a mall requires serious consideration in terms of analysis and research of the location and business opportunities available. The location of the mall is paramount to the success or failure of the project.

Business Problem

The objective of this capstone project is to analyse and select the best locations in the city of Mumbai, India to open a new shopping mall. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business question: In the city of Mumbai, India, if a property developer is looking to open a new shopping mall, 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 shopping malls in the financial capital of India i.e. Mumbai. This project is timely as the city is currently suffering from oversupply of shopping malls. Data from the National Property Research Org released last year showed that an additional 25 per cent will be added to existing mall space, and the agency predicted that total occupancy may dip below 75 per cent. The local newspaper The Business line also reported in February that the true occupancy rates in malls may be as low as 30 per cent in some areas, quoting a Financial Times (FT) article cataloguing the country's continued obsession with building more shopping space despite chronic oversupply further aggravated by the effects of the Covid19 pandemic.

Data

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

- List of neighbourhoods in Mumbai. This defines the scope of this project which is confined to the city of Mumbai, the financial capital of the country of India in South Asia.
- 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 malls. We will use this data to perform clustering on the neighbourhoods.

Sources of data and methods to extract them

This Wikipedia page (https://en.wikipedia.org/wiki/Category:Suburbs_of_Mumbai) contains a list of neighbourhoods in Mumbai, with a total of 41 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 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.