Coursera Capstone

IBM Applied Data Science Capstone

PROJECT NAME:

Opening a New Shopping Mall in Kuala Lumpur, Malaysia



BY :- SANYAM JAIN

Introduction:

Benefits of Shopping Malls

The first malls were built in the 19th century in Paris Arcades. They instantly became popular destination for shoppers. Over time shopping malls have evolved to meet the needs of shoppers. Shopping malls are easily the preferred shopping areas of our time for shoppers across the globe.

The first shops in shopping centers were designed to catered to the elite but this is no longer the case. Nowadays, we have different shops in shopping malls which cater to a variety of budgets.

Shoppers prefer going to shopping malls or shopping centres instead of going to stand-alone shops for various reasons. Bellow we will look at a few benefits of shopping malls for shoppers;

Shopping malls have a wide variety of products available in one space.

Shopping malls have products from competing producers available under one roof. So, making it easier to compare and make purchases.

Malls also have their own parking facility.

Shopping malls have a variety of food courts with a wide variety of cuisine.

Shopping malls tend to have movie theatres.

Mall have gaming zones.

All these benefits of shopping malls make shopping a fun-filled and satisfying experience for shoppers. Now it is important to identify the benefits of shopping malls for business owners as well.

It is very beneficial for a business owner to set up a store in a shopping mall because shopping centers are sought-after shopping destinations, which are usually located in prime and easily accessible locations. Therefore, it is beneficial for retail store owners to rent shop space in a mall.

Renting retail space in a mall allows a business owner to attract clients of competitors who have shops in the same mall.

A business owner does not have to spend time and effort towards the maintaining their shop in a mall. Which means there is more time to focus on running a successful business.

Shopping malls are a popular location for pop-up stores as they are high traffic areas which is ideal for to attract customers to new products.

With all the benefits of shopping malls for shoppers as well as business owners we can conclude that shopping malls & shopping centers will only continue to rise in popularity over time.

Problem Statement

The objective of this capstone project is to analyse and select the best locations in the city of Kuala Lumpur, Malaysia 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 Kuala Lumpur, Malaysia, if a property developer is looking to open a new shopping mall, where would you recommend that they open it?

Data

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

• List of neighbourhoods in Kuala Lumpur. This defines the scope of this project which is confined to the city of Kuala Lumpur, the capital city of the country of Malaysia in South East 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\_in\_Kuala\_Lumpur) contains a list of neighbourhoods in Kuala Lumpur, with a total of 70 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.