

Introduction

Kuala Lumpur is a fast-growing area in Asia. As a result, there are many shopping malls in the city of Kuala Lumpur being built. The shopping malls allows REIT to earn consistent and strong income. Opening a mall however is relatively complicated and requires a bunch of research. The best place for a shopping mall will allow for more success.

Business Problem

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?

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 capital city of Malaysia i.e. Kuala Lumpur. This project is timely as the city is currently suffering from oversupply of shopping malls. Data from the National Property Information Centre (NAPIC) released last year showed that an additional 15 per cent will be added to existing mall space, and the agency predicted that total occupancy may dip below 86 per cent. The local newspaper The Malay Mail also reported in March last year that the true occupancy rates in malls may be as low as 40 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. 3

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

I will use (https://en.wikipedia.org/wiki/Category:Suburbs_in_Kuala_Lumpur) which has a list of all the neighbourhoods in Kuala Lumpur. I will use Python requests and beautifulsoup packages. Using Python Geocoder package will give 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.

Conclusion

We hope to 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.