SHOPPING MALL IN BANGKOK, THAILAND

INTRODUCTION

Bangkok in Thailand is a city where majority of the income comes from tourism. Thailand is also a place where different forms of arts and literatures are respected. Investment in market such as café, pubs, restaurant, shop, etc. can lead to a good amount of income. As per the latest trend an individual search all the above in once place and thus the concept of *malls* is growing. Property developers are taking advantage of this trend to build a Shopping Centre which offers all the leisure and fun at one place. Opening a shopping mall in KL sounds like a good idea but requires proper planning.

BUSINESS PROBLEM

As with any business decision, opening a mall also requires lot of aspects to be considered like;

- Selecting a large area of plot.
- Income of the nearby localities.
- Easily accessible.

These questions can help to decide on an location for the shopping mall which can decide its success or failure.

DATA

1. For these exercises, we will be using neighborhoods data of Bangkok from Wikipedia page.

   https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Bangkok

2. ArcGIS geocoder to get coordinates for the suburbs data extracted from Wikipedia page.

3. Foursquare API, venue data to get list of existing shopping malls

METHODOLOGY

We will use below steps;

1. Scrapping Wikipedia Page using *Beautiful Soup*

2. Getting latitudes and longitudes using *ArcGIS Geocoder*

3. Using Foursquare API to extract *venue data for shopping malls*

4. *Grouping data by neighborhood* and taking the *mean of the frequency of occurrence* of each venue category.

5. Clustering the data using *k-means*

6. Visualizing the clustered data on map using *folium*

First, we will scrap the neighborhood data within Bangkok using Beautiful Soup. The data is present on Wikipedia page. These will give us list of neighborhood names which can be converted into a pandas dataframe. Geo locations can be added by arcgis geocoder package. Foursquare API will be used to extract venue data and filtering them only for shopping malls. Then grouping, clustering with k-means data can be clustered. Lastly using folium data can be visualized. These will help to perform a sanity check on the returned data in Bangkok.

With Foursquare API, we limit the data to 100 records and search within 2000 meters of radius from the neighborhood location.

RESULTS

The results from the k-means clustering show that we can categorize the data into 2 clusters based on the frequency of occurrence for Shopping Mall, Cluster 2 is only at 1 location:

- Cluster 0 – Neighborhood with moderate number of shopping malls (Red marker)
- Cluster 1 – Neighborhood with low number of shopping malls (Blue marker)

The results of clustering are visualized in the map below;
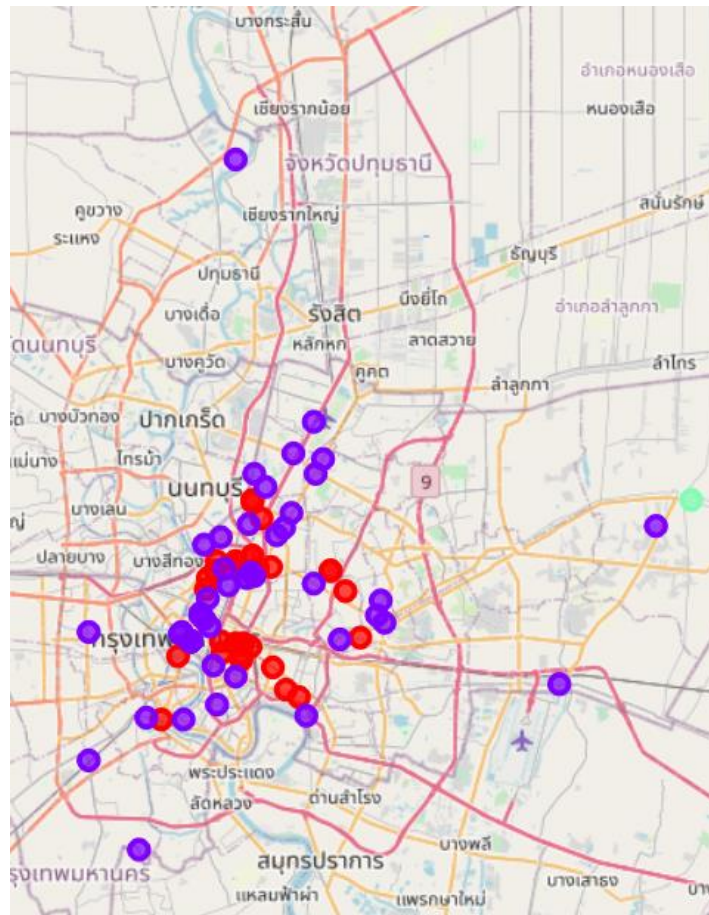
DISCUSSION

The observation says that cluster 0 focuses in the core city part of Bangkok with maximum number of malls while cluster 1 is with lower count of mall somewhat extended to the main city part of Bangkok. That means for our requirement cluster 1 area with low competition can be considered.

CONCLUSION

With all the analysis done using different machine learning algorithms we can fulfill our requirement mentioned in the introduction part of the report. With all the analysis results and discussions, it seems that cluster 1 area can be considered to open a new shopping mall in Bangkok. Although it is not the core city part, but the area can be suggested by developers which consume lower investment and higher returns. People now also take into account crowd gathered in an single location.