IBM Data Science Capstone Project

Danh Thai Hoang
danhthaihoang@gmail.com

April 2020

Opening Restaurants Business in Kular Lumpur

1. Introduction

Starting a Restaurant Business in Malaysia

We can see that more and more restaurants are popping up in the city. The phenomenon proves one thing: the restaurant business is one of the common business ventures in today's business world. That said, building a restaurant from scratch is not easy – all of those advantages come at a price.

Snapshot of Advantages of a Restaurant Business
Though starting a restaurant business is a hard and expensive process, there are advantages you can get when running a restaurant business. As restaurants are in huge demand, a good restaurant will always be filled with customers- that means it will generate revenues. The probability of earning big profits (and encountering less

loss) is high in opening a restaurant business. In other words, the right restaurant at the right location can be a lucrative business.

Things to Consider Before Starting a Restaurant Business in Malaysia. Having a business plan is crucial to success. A carefully planned business plan will steer the direction of your restaurant business thus ensuring lucrative return. Let us dive in to have a look at what you should know before starting a restaurant business in Malaysia.

Concept:

Know and determine your restaurant's concept before you launch it. If you are unsure, you can always run some research; consult the professionals to streamline your idea. Generally, the concept is like the road map that includes your restaurant's menu, investment's capital, and source of financial aids, expected profit, and marketing plan.

Location:

The location of your restaurant business will affect its success nearly as much as the menu. As such, you may want to research an ideal location with high traffic, easy access, and good visibility. Other factors that could

determine your restaurant's location is the target market of your business as well as the competitors nearby.

Menu:
The menu of a restaurant is another big consideration when you are planning a restaurant business. When you are deciding what to include in the menu, the first thing you should know is the target market, the current food trend and what you plan to serve. A good restaurant menu is the key to any restaurant's marketing plan.

Marketing:
Marketing plan is a plan of how to get the word out about your new restaurant. Nowadays the marketing plan sees a lot of variety when you can take advantage of the many social media platforms. Thinking it through with brilliant ideas will actually help your restaurant to succeed and generate huge profits.

Licensing and permits:
Before you can open the doors to your restaurant, you need to make sure you apply for the proper licenses. In Malaysia, a Malaysia business license is issued on various terms depending on whether the investor is local or foreign. For restaurant business, a Signboard License, Alcohol License (if you are planning to sell alcohol in your

restaurant), Halal License (if applicable) MACP and PPM (for restaurants where they will play music in the premises). If the restaurant is a foreign-owned restaurant, an approved Wholesale, Retail Trade License is needed. Please note that such license is only granted for restaurants with unique concepts (like Arabian restaurants, French cuisine) and a minimum space of 1,500 sq. ft.

The problem statement

As the final assignment of IBM Data Science professional certificate on the Coursera, the project is intended to apply possible knowledge and skills in Data Science gained from different courses along the certificate to resolve the problem in the real world.

Because the location of the Restaurant – This is the key to the future of your business the project tries to answer the questions: **"Should an entrepreneur start a restaurant business in Kular Lumpar City, Malyasia" and "Which location is suitable and recommended to open it"**. With applying various data science methodology and machine learning techniques from linear regression to clustering, classification the project work aims to answer these questions.

2. Data

Data Requirements

Because the location is the most important factor for further analysis, visualizing, modelling and evaluation as well so it requires the Kular Lumpur's **neighborhoods data**. 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 SouthEast Asia.

Latitude and longitude coordinates of those neighbourhoods are also required to visualize and plot the map and also to get the venue data.

Venues data, particularly data related to restaurants. This is the crucial data for clustering on the neighbourhoods.

Data collection

- The list of Kular Lumpur's **neighborhoods data** is collected from the wiki page: https://en.wikipedia.org/wiki/Category:Suburbs_in_Kuala_Lumpur with more than 70 **neighborhoods** in the list. We will crawl the data from the page by using

BeautifulSoup library, Python requests and Pandas.
The collection process looks like this:

Data Collection

Wiki page to get the list of Kular Lumpur

```
In [10]: url = 'https://en.wikipedia.org/wiki/Category:Suburbs_in_Kuala_Lumpur'
```

Call the bs4 lib to start crawling the data

```
In [14]: nData = requests.get(url).text
```

Parse the data collected from wiki page

```
In [15]: nSoup = BeautifulSoup(nData, 'html.parser')
```

Create the list of neighborhoods

```
In [20]: neighborhoodList = []
         for row in nSoup.find_all("div", class_="mw-category")[0].findAll("li"):
             neighborhoodList.append(row.text)

Out[20]: ['Alam Damai',
          'Ampang, Kuala Lumpur',
          'Bandar Menjalara',
          'Bandar Sri Permaisuri',
          'Bandar Tasik Selatan',
          'Bandar Tun Razak',
          'Bangsar',
```

- **The latitude and longitude coordinates** of the neighbourhoods are retrieved using Google Maps Geocoding API. The geometric location values are then stored into the initial dataframe.

Get the geographical coordinates

Define a function to get lat and long

```
In [47]: def get_latlng(neighborhood):
             # initialize your variable to None",
             lat_long_coords = None
             #loop until you get the coordinates",
             while(lat_long_coords is None):
                 g = geocoder.arcgis('{}, Kuala Lumpur, Malaysia'.format(neighborhood))
                 lat_long_coords = g.latlng
             return lat_long_coords
```

Call the function to get lat and long list and store in the new list (include the neigborhoods and their coordinates

```
In [48]: coords = [get_latlng(neighborhood) for neighborhood in kuLu_df["Neighborhood"].tolist()]
```

Merge the neigborhoods data into coordinates data into one dataframe

```
In [49]: neigh_coors_df = pd.DataFrame(coords, columns=['Latitude', 'Longitude'])
         neigh_coors_df.head()
```

Out[49]:

	Latitude	Longitude
0	3.057690	101.743880
1	3.148494	101.696729
2	3.190350	101.625450
3	3.103910	101.712260
4	3.072750	101.714610

- From the location data obtained after Web Scraping and Geocoding, the venue data is found out by passing in the required parameters to the **FourSquare API (https://developer.foursquare.com/)**, and creating another DataFrame to contain all the venue details along with the respective neighbourhoods.

3. **Methodology**
4. **Results**
5. **Discussion**
6. **Conclusion**

7.   References