



IBM - Applied Coursera Capstone

Capstone Project – The Battle of Neighbourhood

Topic: Opening a new hotel in Yangon, Myanmar

Author: HAN XIANG CHAN

Date: July 2020

Introduction

Yangon is the largest city and former capital of Myanmar (Burma). It is known for its tourist attractions such as the Shwedagon and Sule pagodas, and British colonial era architecture in the downtown area. Since the country started its political reforms around 2010, so did its economy, and Myanmar has experienced a significant year-on-year increase in tourist arrivals since 2011. Most international tourists enter Myanmar through Yangon, which in turn spurred a hotel construction boom around 2012. [1] However, as supply increased much faster than demand, this resulted in a surplus of hotel rooms available. In 2019, Yangon hotel occupancy rate was only 44% in 2019, which is considerably lower than other cities in the region. Consequently, Yangon hotel prices tend to be more competitive.[3]



Figure 1: The Shwedagon pagoda at dusk

Business problem

An investor has approached me with the interest of opening a hotel in Yangon, and he would like to know which the best area is to start a hotel. While it has been reported that Yangon already has a surplus in the supply of hotel rooms over demand, it seemed that most of the hotels are found only in certain townships of the city popular with tourists, while in other parts of the city accommodation maybe harder to find. My goal would be to identify which areas where there are lower density of hotels found, and yet there is potential for demand, be it tourists or business travelers.

Data

Yangon has a total of 33 townships within the boundaries of the city, each of varying size. The list of townships can be extracted from Wikipedia's page: "https://en.wikipedia.org/wiki/Category:Townships_of_Yangon".

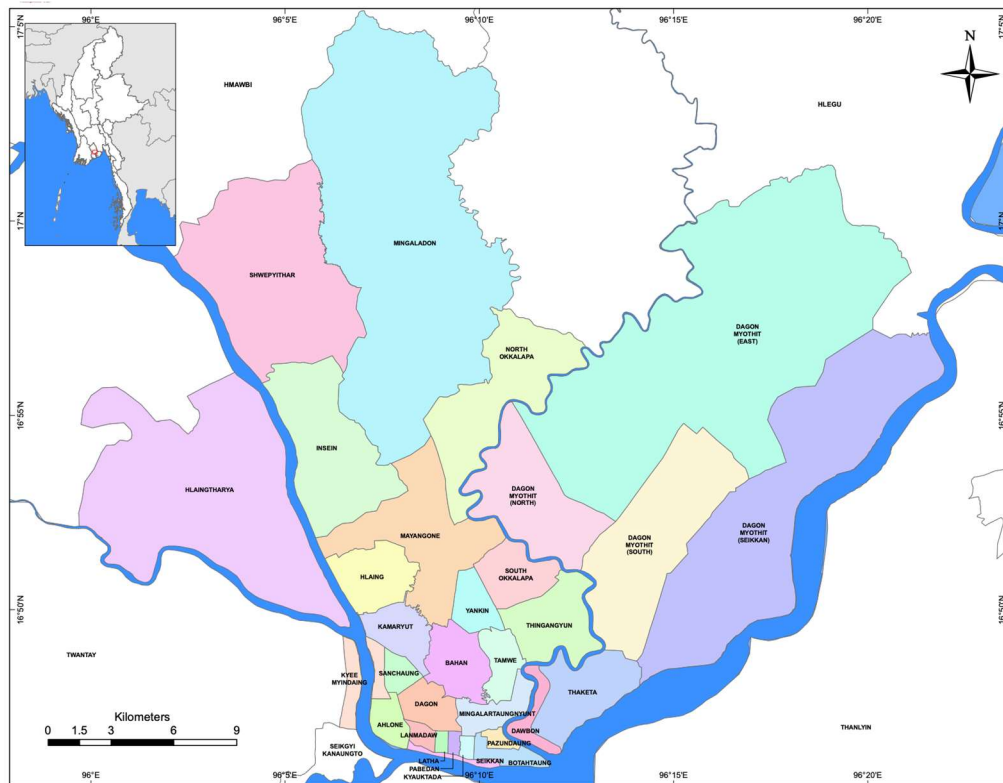


Figure 2: Townships of Yangon

We would also need the latitude and longitude of each township. Using information on the list of township and its corresponding latitude and longitude data, we would use the Foursquare API to retrieve the locality of hotels within Yangon city, which will be used for clustering analysis.

Methodology

1. Import libraries: requests, bs4, pandas, json, geocoder
2. Use BeautifulSoup (from bs4) to parse list of townships
3. Use geocoder package to extract latitude & longitude of each township
4. Store data in Pandas DataFrame, and save it using CSV
5. Import additional libraries: matplotlib, sklearn, folium for mapping & data visualization
6. Using folium, obtain coordinates and plot map of Yangon
7. Connect with Foursquare API, identify the categories of places returned.
8. Perform one-hot encoding on all categorical variables returned by Foursquare API label, which will return a numerical index needed for clustering analysis.
9. For the purpose of this study we select "Hotel". Subsequently, perform clustering analysis for Hotel category.
10. Repeat clustering analysis for "Hostel" category. (See results & discussion for details)

Results

Based on the onehot encoding indices returned for each township, I decided to perform a clustering analysis with $k = 4$ for the hotels. The clustering analysis returned the following hotel clustering indexes of all 33 townships:

Cluster 0 (index value = 0.0, in Green): Thaketa, North Okkapala, Mingalardon, Thingangyun, Ahlon, Insein, Hlaingtharya, Dawbon, Dala, Kyimindaing

Cluster 1 (index value = 0.01 – 0.07, in red): Bahan, Tamwe, Kamayut

Cluster 2 (index value = 0.07 – 0.10, in orange): Kyauktada, Mayangon, Pabedan, Seikkan, Dagon, Dagon Seikkan, Yankin

Cluster 3 (index value = 0.11 - 0.15, in purple): Latha, Botataung, Pazundaung, Mingala Taungnyunt, Lanmadaw

Cluster 4 (index value = 0.20 - 0.30, in turquoise): *South Okkapala* (This is an anomalous result, will explain in discussion).

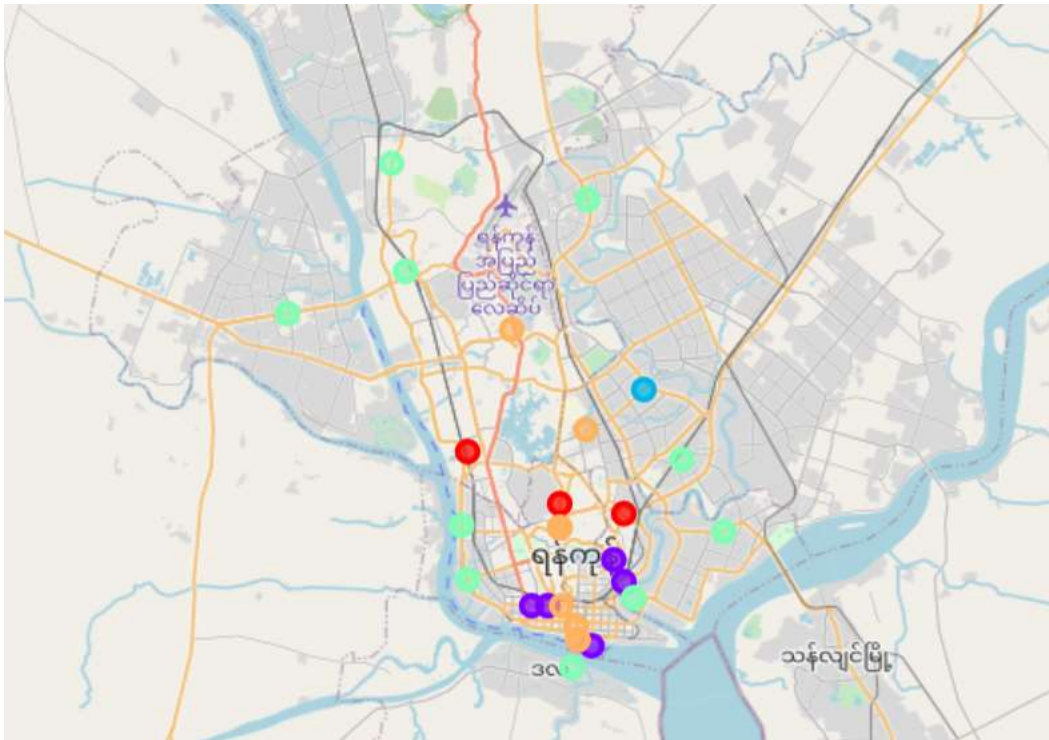


Figure 3: Clustering analysis of Hotels in Yangon

As budget hotels maybe classified as “Hostels” in the Foursquare Database, I decided to carry out clustering analysis for Hostels. Results as follows:

Cluster 0 (index value = 0.0, in red): The majority of townships surrounding Downtown Yangon

Cluster 1 (index value = 0.01 – 0.03, in orange): Botataung

Cluster 2 (index value = 0.03 – 0.05, in purple): Lanmadaw, Pabedan, Seikkan, Dagon Seikkan

Cluster 3 (index value = 0.06 – 0.07, in green): Kyauktada

Cluster 4 (index value = 0.07 – 0.10, in turquoise): Latha

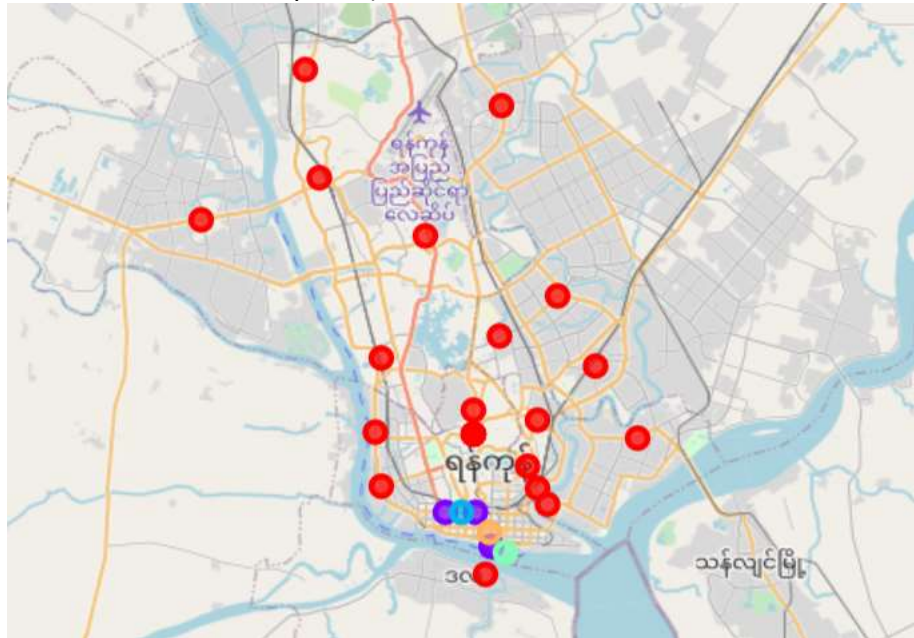


Figure 4: Clustering analysis of Hostels in Yangon

Discussion

The results returned from the clustering analysis of hotels in Yangon showed that most of the hotels are located near or within the city centre. Townships within the city centre, such as Latha, Botataung and Lanmadaw returned the highest hotel indexes as many one, two and three hotels have opened to cater to backpacker, budget and group tours. In wealthier suburbs such as Bahan and Tamwe, hotels are geared more towards the luxury travellers, as most of the inhabitants in these townships come from wealthier backgrounds.

As some of the one or two star hotels maybe identified as hostels by the Foursquare community, I decided to do another clustering analysis using data returned from the “Hostels” category. All the hostels are concentrated in the downtown region of Yangon, which also co-relates to some extent with the higher clustering index of Hotels previously identified.

The clustering analysis for Hotels returned “0.0” for townships further away from the city centre, such as Thaketa, Hlaingtharya and Insein. Clustering analysis of Hostels also returned “0.0” for townships away from Downtown Yangon, including the three aforementioned townships. Thaketa and Hlaingtharya are industrial towns, and the investor may wish to consider developing hotels catering to businessmen, professionals, or budget long term accomodations for industrial workers who do not live in Yangon. As we have mentioned that Yangon already has an oversupply of hotel rooms catering to tourists, it might not be very viable to open another hotel targeting tourists, given the fierce competition among hoteliers in this group.

Limitations

- South Okkapala is mostly a residential area and not known as a tourist destination. Yet, it returned an index of 0.25, which is clearly anomalous. As map inputs in Foursquare are crowdsourced from the community,[5] of which one can easily sign up, false locations can be mapped, possibly leading to the anomalous result.
- While the “Hotel” and “Hostel” clustering maps largely conforms to external observations about locality of hotels in the city, the “Hotel” returned a sample size of 19, whereas “Hostel” returned a sample size of 7. Increasing the radius by 1000 metres only increase the sample size to 27, which is still too low a sample size for the clustering analysis to be deemed as truly credible.
- Myanmar’s internet penetration rate exceeded 50% only in the last three years, which is much later than most other developing and developed countries. We may have expect developer input into Foursquare API to be less extensive, comprehensive and accurate as compared to other localities.

Conclusion

Clustering analysis can be used to distinguish and subsequently aggregate data into different categories based on similar scores or properties. When data comes in the form of a list, and a hypothesis is to be tested and validate business hypothesis, this is a powerful technique which may be applied. However, as clustering analysis returns results from a mathematical/programming perspective, one should also derive from other sources of information such as empirical evidence and literature review of prior studies, and do a cross study with the results of the clustering analysis before making a firm conclusion.

In the case of this study, the mapping results conforms with the general observation that most hotels are located near or within Downtown Yangon. However, the small sample size, as well as the anomalous cluster returned from the API makes it doubtful as to whether this result can be used as a definitive evidence for business making decision. A check with the Agoda website, which caters to hotel bookings, showed that it uses the Google API to return location data. Hence, it is advisable that to gather a more verifiable data and outcome, the analyst might wish to consider perform additional clustering analysis with API from more popular sites such as Google and Bing.

References

- [1] <https://oxfordbusinessgroup.com/overview/be-enchanted-heritage-beauty-and-improved-regulations-attract-new-crop-domestic-and-regional> Myanmar attracts growing number of tourists – Oxford Business Group
- [2] <https://tourism.gov.mm/wp-content/uploads/2020/05/Myanmar-Tourism-Statistics-2019-1.pdf>
Myanmar Tourism Statistics 2019 – Ministry of Hotels and Tourism, Myanmar Government
- [3] <https://www.ttrweekly.com/site/2019/11/yangon-hotel-rates-remain-a-bargain/>
Yangon hotel rates remain a bargain – TTRW
- [4] https://reliefweb.int/sites/reliefweb.int/files/resources/Org%20Map_MSF-CH_MIMU1002v01_Yangon%20City%20Township_05Apr13_A3.pdf
Myanmar Information Management Unit – Yangon City Townships
- [5] <https://rapidapi.com/blog/top-map-apis/> The Top 10 Mapping & Map APIs