

Coursera Capstone Project

The Battle of Neighborhoods (Week 2)

Contents

1. [Introduction](#)
2. [Data](#)
3. [Methodology](#)
4. [Results](#)
5. [Discussion](#)
6. [Conclusion](#)

Introduction

Yangon is the capital of Myanmar where local migrant workers are residing to work in various industries, and foreign investors are having their eye on new opportunities.

The Map of Yangon



Recently, a local investor is looking for a place to run a burmese restaurant in a certain neighborhood of Yangon. The area of interest will be in the following criteria:

1. near to offices, hotels, and public attractions such as supermarket, shopping malls and cinemas
2. fewer competitors in the vicinity
3. ample space for parking

Data

To analyse the each neighborhoods, the coordinates of each neighborhoods in Yangon city will be required and can be found at the website, www.themimu.info.

	SR_Pcode	SR_Name_Eng	District/SAZ_Pcode	District/SAZ_Name_Eng	Tsp_Pcode	Township
0	MMR013	Yangon	MMR013D004	Yangon (West)	MMR013037	
1	MMR013	Yangon	MMR013D004	Yangon (West)	MMR013044	
2	MMR013	Yangon	MMR013D002	Yangon (East)	MMR013017	
3	MMR013	Yangon	MMR013D003	Yangon (South)	MMR013032	
4	MMR013	Yangon	MMR013D004	Yangon (West)	MMR013043	

With this data, we will be working to find the best place to run a burmese restaurant considering the criteria.

Methodology

We will find the neighborhoods of Yangon where:

- 1) low restaurants exists,
- 2) populated with public areas, offices, hotels, supermarkets, cinemas, etc.,
- 3) and population density is higher than average;

which will be:

- a) good for customer acquisition with less competition,
- b) and lower rental cost for spacious areas in return.

Firstly, for the requirements: population, neighborhood areas, we will need to **scrape from Wikipedia** and **explore venues using Foursquare API** based on coordinates of neighborhoods.

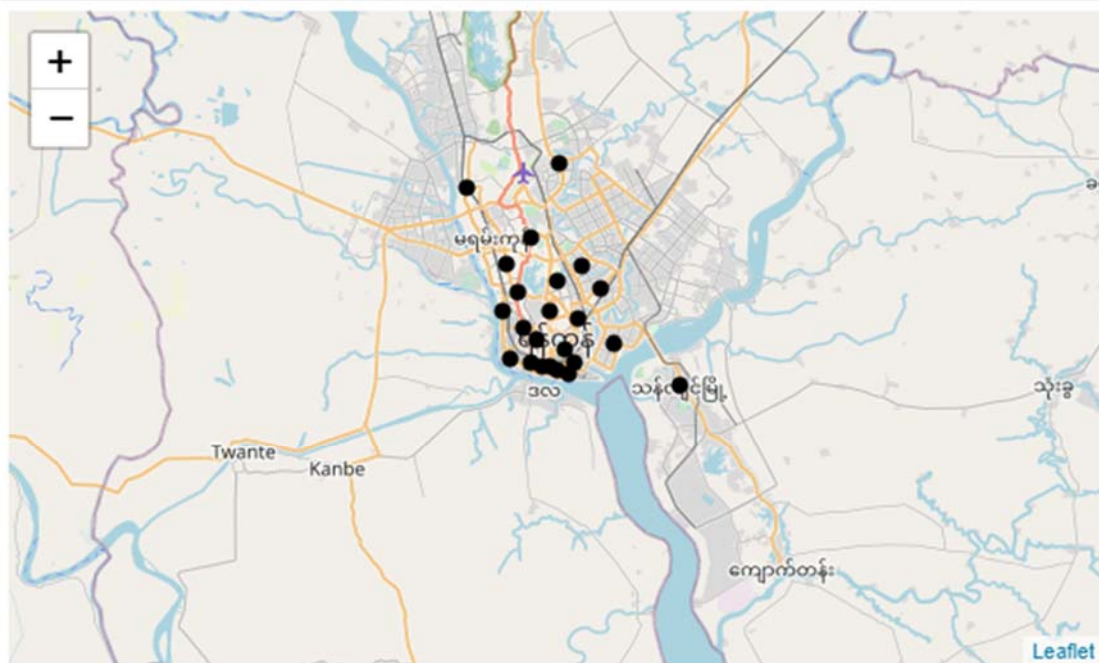
Secondly, as a exploratory analysis, we will find out **restaurant density, population density and existence of public and office areas** at every neighborhoods.

Finally, we will focus on areas with highest potential and highlight the selective areas for final decision on the map along with clusters of all neighborhoods using K-Means algorithm for further investigation at ground level.

1. Analysis

According to the data, there are 45 neighborhoods in the region.

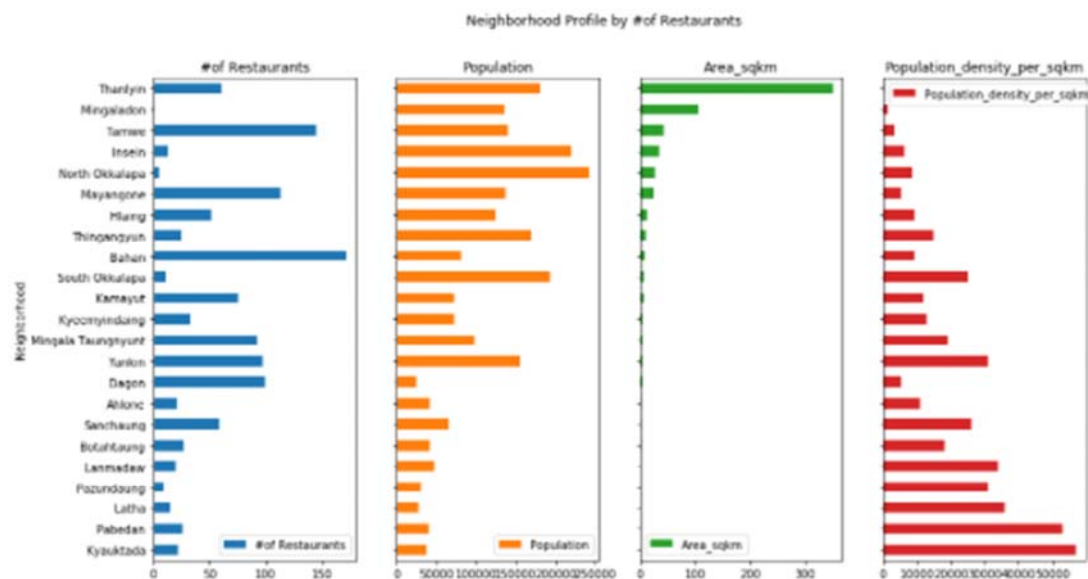
First, we need to select a few areas that are popular for business location because of infrastructure availability and population with higher income.



Based on the filtered areas, required data such population and areas of neighborhoods will be scraped from www.wikipedia.org.

	Neighborhood	Area	Population	Population Density
0	Ahlon	4 km2 (1.4 sq mi)	41,200	11,000/km2 (29,000/sq mi)
1	Bahan	8.84 km2 (3.413 sq mi)	81,000	9,200/km2 (24,000/sq mi)
2	Botahtaung	2.4 km2 (0.92 sq mi)	42,000	18,000/km2 (46,000/sq mi)
3	Dagon	4.7 km2 (1.8 sq mi)	25,082	5,300/km2 (14,000/sq mi)
4	Hlaing	13.7 km2 (5.29 sq mi)	125,000	9,100/km2 (24,000/sq mi)

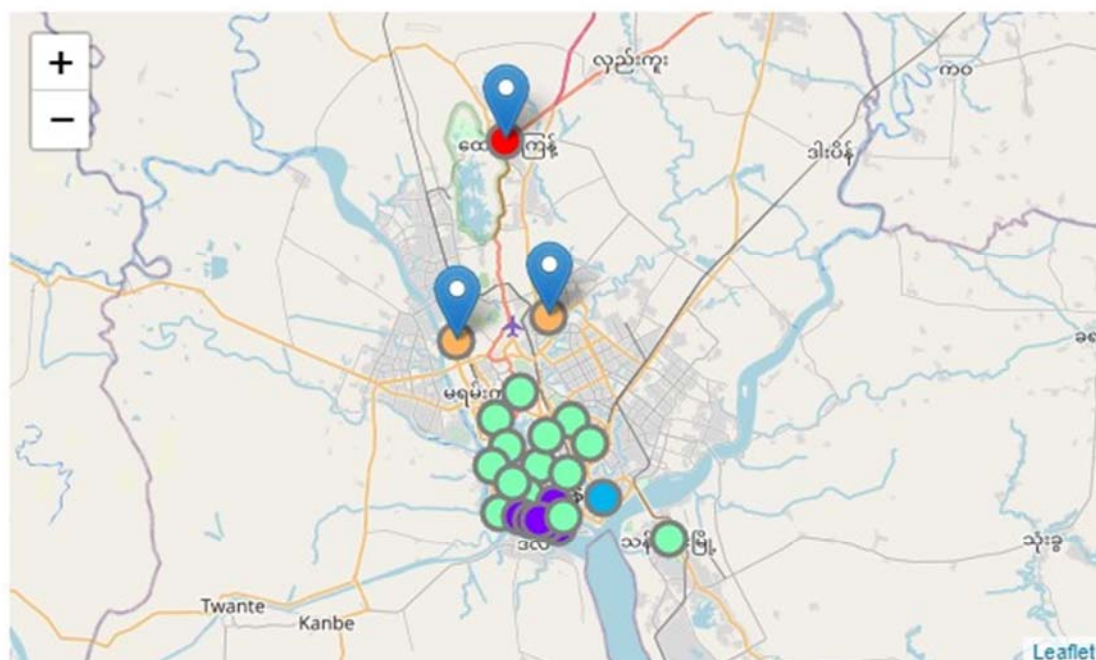
The data contains the required data but in formatted text, so, we will clean texts and characters from the data first.



According to above comparison, top 3 neighborhoods populated with restaurants have correlation to big areas with low population density while the existence of offices and income of the population are considered as top reasons.

With that details, we can estimate, the second neighborhood **Mingaladon** is the best option following **Insein** and **North Okkalapa**.

Let's focus on those 3 areas and plot them on to clusters of neighborhoods using K-Means algorithm below.



Obviously, the 3 areas are located away from the downtown area but have potentials too.

Let's look at the info of those below.

Results

As we can see, both of the clusters where our top 3 prioritized areas have top popular venues in **Transportation**, **Public Space** and **Market**.

The existence of the competitors are also low and the areas have lesser population density comparing to other parts of the region which is the opportunity for acquiring new customers and the location with ample space at lower cost.

Besides, the areas are located near to the exist ways of Yangon; Pyay road and international airport of Yangon, to other parts of Myanmar and foreign countries. That will be an advangeous for attracting travellers in the holiday seasons as well.

So, we recommend to keep our priority in finding the location in those areas.

Discussion

We have explored the venues and examined the features to oversee the potential areas. But there are still limitations that:

- 1) exploring nearby venues from the center of neighborhood can not get venues within the area correctly since neighborhood boundaries are not in circular shape
- 2) venues obtained from Foursquare is limited and may not include new venues that are not registered on the web

But, with this solid information, we can start exploring at the ground level for futher project development works.

Conclusion

Eventhough the recommendation is for kicking off the project, we can improve our report on giving more time on:

- 1) using **SVM** that enables explored venues to be clustered into neighborhoods with better accuracy,
- 2) improving K-Means by finding best cluster size with **elbow method**

As a result, the recommendation will be more reliable