**Dental Clinic Address Selection in Kuala Lumpur, Malaysia**

### IBM Applied Data Science Capstone by Coursera (Aug 2021)

SL Tan

## **INTRODUCTION**

## **Background**

Malaysia is a country located in South East Asia, with over 32 million of population [1].

According to the Health Minister Datuk Seri Dr Adham Baba, Malaysia has a total **11,059 dentists** as of June 2020, with the dentist-population ratio at **1:2,963** [2]. There isn't an ideal population-to-dentist ratio recommended by WHO, however, if we compare to the U.S which has ratio of **1:1,638** [3], this suggests that there are plenty of opportunities in providing dentistry service in Malaysia.

## **Business Problem and Interest**

We will use this project to find an optimal location to open a dental clinic. As I reside in Kuala Lumpur, Malaysia, the focus area will be targeting the vicinity of this city. This report would be suited for stakeholders who are interested in starting a **dental clinic in Kuala Lumpur, Malaysia**.

For the selection of an optimal location, we will focus on detecting the area that have **no dental clinics in the vicinity**. It is also preferred if the location are **surrounded by medical centers that offers non-dental medical services**, as this can be an indicator of the potential demand in vicinity. Lastly, the location will have to be **close to Kuala Lumpur**.

We will gather data and use our data science knowledge to generate a list of areas that fits all the criteria above. Advantages of each area will also be clearly stated so the stakeholders can make the best possible location selection.

## **DATA**

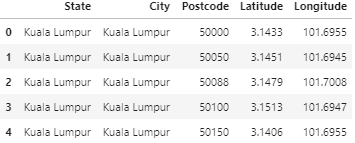
## **Data Sources**

To allow the measurement of the selection criteria listed in the problem statement, we need the followings:

* GPS Location of Kuala Lumpur  
  We will use geo-location function to obtain the GPS coordinate of Kuala Lumpur. This GPS coordination will then serve as a center point in calculating the distance.
* City/Borough/Neighborhood near Kuala Lumpur by Postcode  
  Kuala Lumpur is surrounded by Selangor State, as such the data collection has also extended to Selangor. We have obtained the postcode for cities in both Kuala Lumpur and Selangor from a public website [3][4].
* Medical services in the areas near Kuala Lumpur  
  We will use Foursquare API to populate the medical services nearby the postcode within the area.

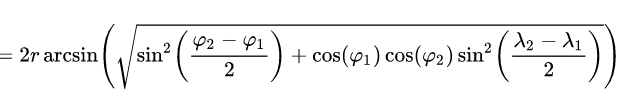
## **Data Preparation**

Data downloaded or scraped from the public website were combined into one table. The table has then been tidied up to a standardized format that provides a list of postcode with details of its latitude and longitude. A total of 559 postcodes is listed in the table.

  
Figure 1. Postcode Table (First 5 rows)

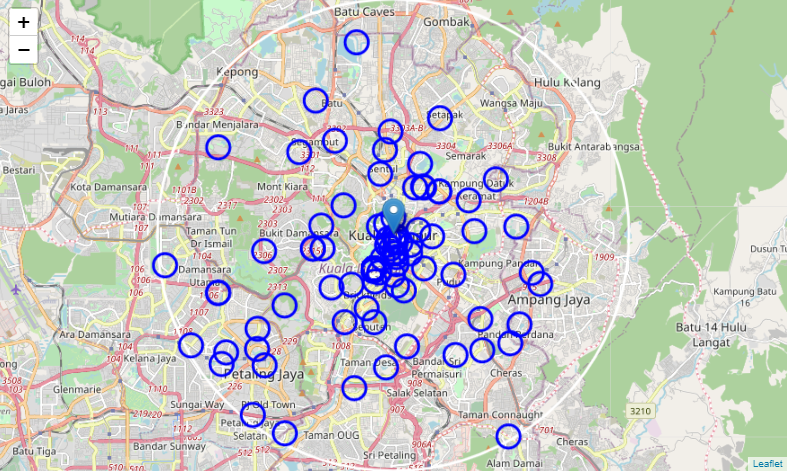
Then, we use geo-location to obtain the latitude and longitude of Kuala Lumpur. The location is (3.1516964, 101.6942371). This location is also representing the city center, and the 10km radius distance from this location will be our focus in this study.

To consider the latitude and longitude is actually on sphere of plant Earth in the distance calculation, Haversine formula [5] is used. Haversine formula:



From the distance calculation obtained, we now able to filter the postcode to consider those are within the 10km distance from city center.

There are 297 postcodes within the 10km radius from Kuala Lumpur city center.

  
Figure 2. Map plot for 297 postcodes with 500m radius

From the map, postcode is not a good neighborhood center point selection, as the radius of 500m was not able to cover all the areas. Further, some of the postcode areas are overlapping with each other.

To solve this issue, we will use Foursquare to locate the medical centers within the 2km of the postcode, then re-assess whether medical centers will be a better neighborhood center point selection.

From the result generated from Foursquare, 1,641 medical centers were found. All the areas are well covered when the radius is 500m. For the purpose of this project, we are focusing on the developed locations, so areas that are green shaded area which represents the vegetation areas are not of our interest.

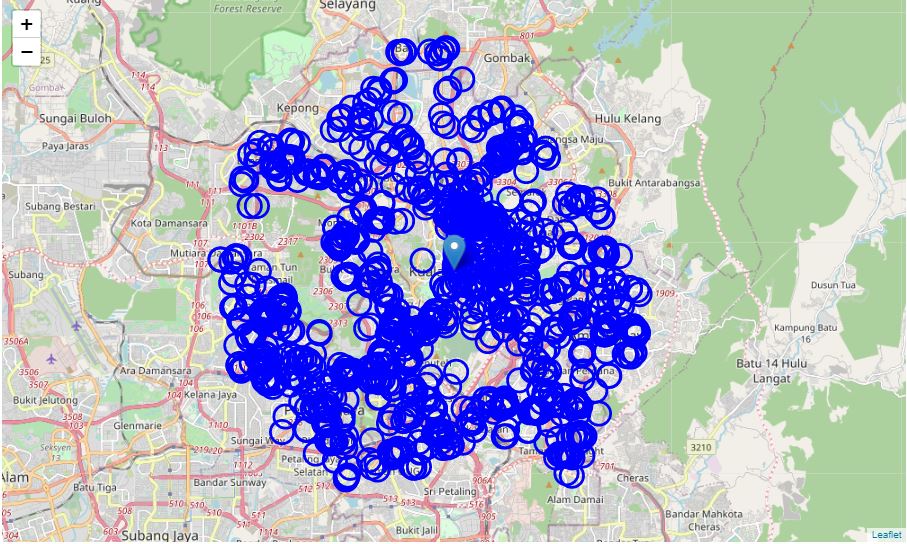
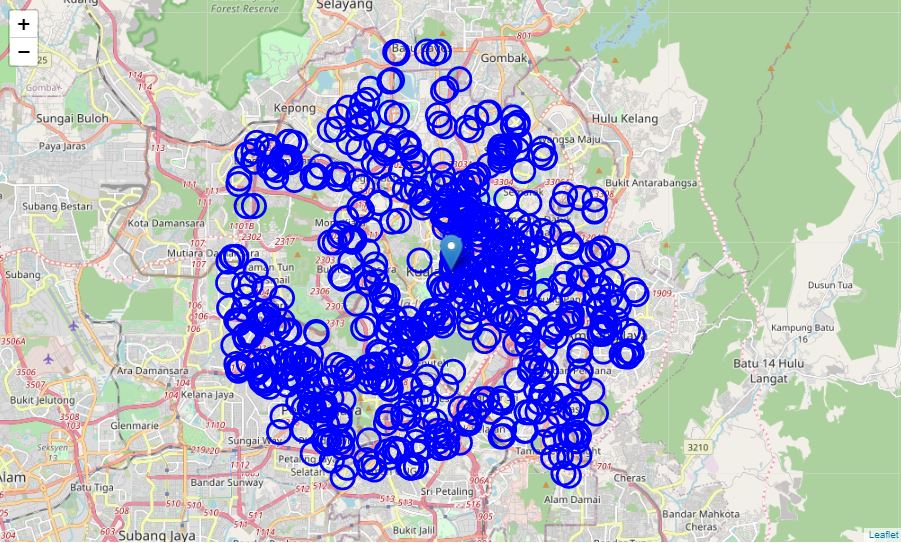


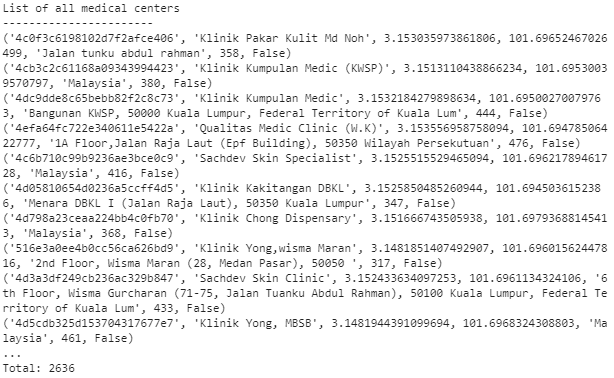
Figure 3. Map plot for 1,641 medical centers with 500m radius

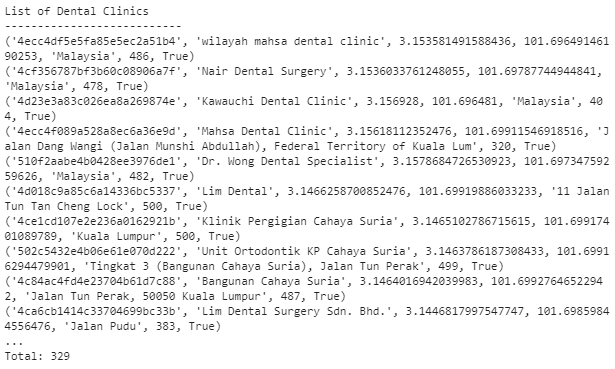
From the map, we have observed that 1,641 medical centers are heavily overlapping with each other. Hence, we will remove the locations that are close to each other. This will effectively reduce the number of query to Foursquare in subsequent steps, while still achieving the same result.

By considering the medical centers that shared the same latitude and longitude up to its 3 decimal points are close to each other, we have then reduced the number of medical centers to 678.

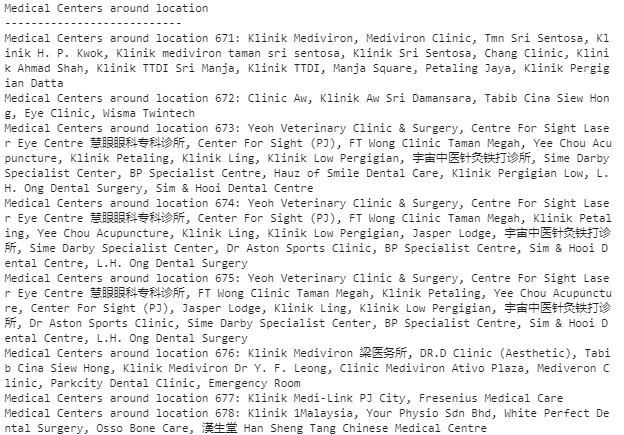
  
Figure 4. Map plot for 678 medical centers with 500m radius

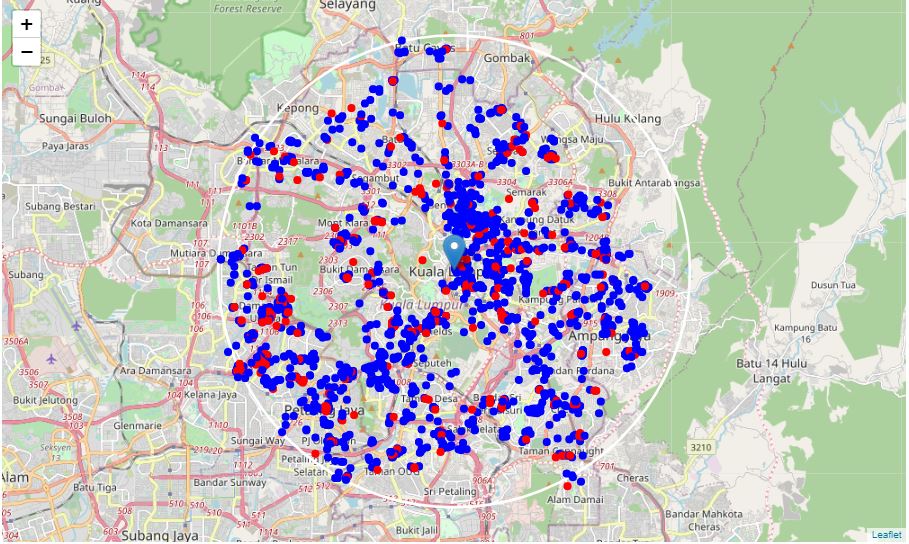
We will use this 678 medical centers as the center point in looking for the nearby medical centers with clear marking on dental clinics. By using Foursquare, we have then obtained a total 2,636 medical centers, of which 329 are dental clinics. We have also obtained a list of medical centers that is within 500m radius distance for each of 678 locations.

1. Part of the list of 2,636 medical centers:
2. Part of the list of 329 dental clinics:



1. Part of the list for the medical centers around each of the 678 locations:



  
Figure 5. Map plot for 2,636 medical centers (blue) and 329 dental clinics (red)

## **METHODOLOGY**

The focus of this project is to identify areas near Kuala Lumpur that do not have dental clinics, preferably surrounded by many medical centers. The location is set to be within 10km from the city center of Kuala Lumpur, the nearer the better.

During the data collection, we have generated a list of **medical centers and dental clinics** that are **within the 10km from Kuala Lumpur** city center. We have also identified the medical centers (include dental clinics) that are within 500m distance.

In the analysis, we will explore areas that fulfill the criteria of **high medical centers density** with **no (or low) dental clinics** within the vicinity. We will the further narrow down the choices to those comparatively **closer to the city center**. The shortlisted areas will then be clustered into zones using k-means, and finally generating a list of recommended locations.

## **ANALYSIS**

## **Data Analysis**

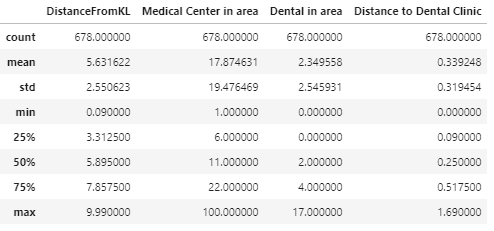
The analysis started by identifying the number of medical centers and number of the dental clinics for each of the 678 locations selected as exploration points.

In addition, the distance to the nearest clinic is also calculated (in km).

  
Figure 6. Table of 678 locations with analysis result.

## **Selection Criteria**

The statistical result of the 678 locations is populated using pandas library build in functions,

  
Figure 7. Table of the statistic for 678 locations result

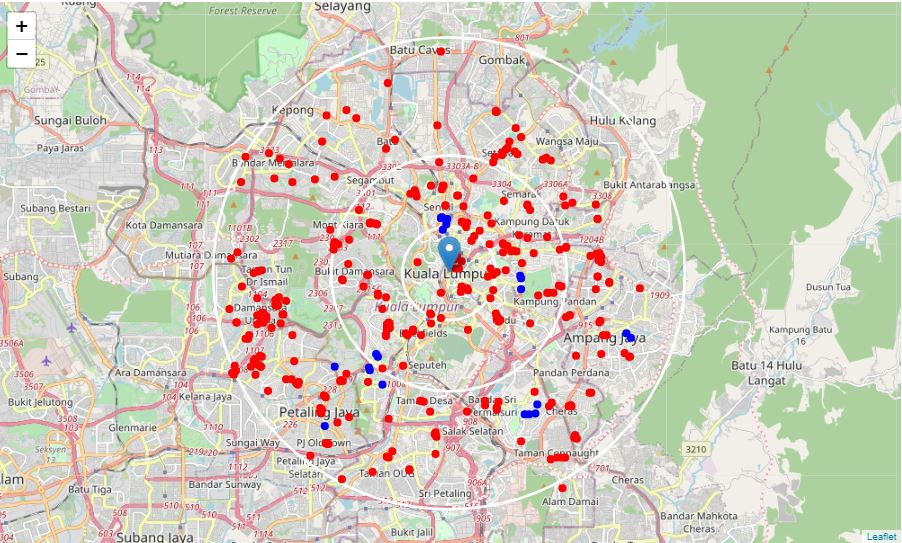
For 678 locations, the number of medical centers is range from 1 to 100, where the average is 17.87 medical centers. 75% of the locations has less than 22 medical centers within 500m. It is then fair to set the selection criteria to focus on the locations that have >=18 medical centers nearby.

In term of the number of dental clinics nearby, it is ranging from between 0 to 17 clinics. It has an average of 2.3 dental clinics, however there are at least 25 percentile has no clinics nearby. To minimize the direct competition, it is suggested to selection locations with no dental clinics nearby.

## **Shortlisted Locations**

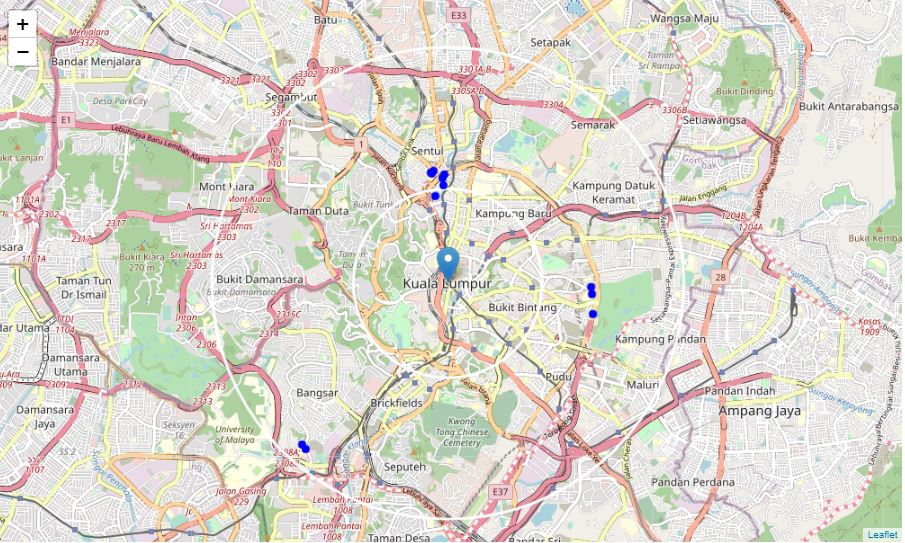
There are 23 locations fulfilled the criteria of having >=18 medical centers within the 500m, while none of them are dental clinics.

The shortlisted location is then plotted on the map and marked with blue dot. Red dots represents dental clinics; while white circle represents 2km, 5km and 10km radius from city center respectively.

  
Figure 8. Map plot with shortlisted locations and dental clinics.

The 23 shortlisted locations are located ~2 to 10km from the Kuala Lumpur city center, where 11 are within 5km from city center.

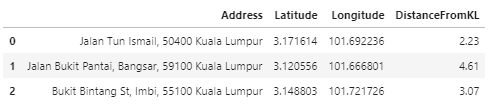
As illustrated in the business problem, locations nearer to the city center are preferred. Hence, the shortlisted locations are narrowed down again to within 5km from city center. These 11 locations is plotted on the map and zoom in to focus within 5km range.

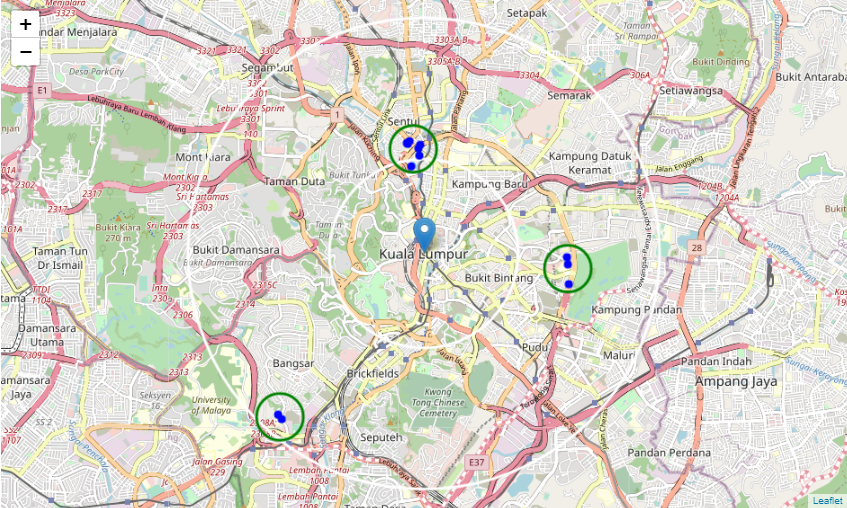
  
Figure 9.Map plot with shortlisted locations within 5km from city center

The shortlisted locations are situated at north, east and south west of the city center.

## **Shortlisted Cluster**

The 11 shortlisted locations are then grouped into 3 clusters using k-means. The center of the clusters (result of k-means) is then translated to the exact address using Google.

  
Figure 10. Table of center of the shortlisted cluster.

  
Figure 11. Map with shortlisted clusters in green.

The 3 clusters have nicely covered the 11 shortlisted locations within its 500m radius. These 3 clusters have formed our recommendation areas to the location selection for a dental clinic. These 3 areas fulfill the requirements of high medical centers density with no dental clinics within 500m. Also, all these locations are reasonably close to city center (within 5km)

## **RESULTS & DISCUSSION**

The 3 recommended areas for locating a dental clinic are:

1. Jalan Tun Ismail, located at north of the city center, is a neighborhood nearby a local residential area.
2. Bangsar, located at south-west of the city center, is an area popular among the expatriates.
3. Bukit Bintang, located at east of the city center, is also located within the golden triangle and a tourist spot.

This 3 areas would be a good starting point in looking for an optimal location for opening up a dental clinics. Each area has its own characteristics and advantages, depending on the business strategy and target customers.

## **CONCLUSION**

In this study, we have explored the location of medical centers and dental clinics that are within the 10km of Kuala Lumpur city center. We noticed the distribution of dental clinics is not evenly spread out, where there are number of pockets that do not have dental clinics.

We have then select the locations that have high medical center density with no existing dental clinics. As there are many choices available, we have then narrow down our selection to look into locations that are within 5km of city center (compared to initial 10km).

This has then arrived our final recommendations of 3 areas – Jalan Tun Ismail, Bangsar and Bukit Bintang. Each area has its own characteristics. To select the most optimal location, it is recommended to also consider the respective business strategy, target customers, social and economic factors during the decision making.

## **APPENDIX**

[1] Current Population Estimates, Malaysia, 2021 <https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=155&bul_id=ZjJOSnpJR21sQWVUcUp6ODRudm5JZz09&menu_id=L0pheU43NWJwRWVSZklWdzQ4TlhUUT09>

[2] 'Msia doctor-population ratio stands at 1: 454' by New Straits Times, 2020 <https://www.nst.com.my/news/nation/2020/08/613844/msia-doctor-population-ratio-stands-1-454>

[3] Kuala Lumpur Postcode list <http://malaysia.postcode.info/kuala-lumpur/kuala-lumpur>

[4] Selangor Postcode list http://malaysia.postcode.info/selangor/

[5] Haversine Formula <https://en.wikipedia.org/wiki/Haversine_formula>