

# **COURSERA CAPSTONE PROJECT**

## **WEEK 4 – THE BATTLE OF NEIGHBOURHOODS**

### **Kuala Lumpur, Malaysia vs Tokyo, Japan**

Prepared by: Muhammad Najmy Bin Hashim

Github Id: TrufflegarLaw



# Introduction

In this project, I would like to answer two questions:

- 1) Is there any similarity between two different cities?
- 2) If developer would like to open a new shopping mall, where should they open it?

For question 1, I will use Kuala Lumpur, Malaysia [3.139003,101.686852] and Tokyo, Japan [35.689487,139.691711]. As we all know, Japan is a highly developed country and most of the countries in Asia including Malaysia idolize Japan and set it as their benchmark. On the other hand, Malaysia is a developing country, despite undergoing rapid economic development over the past five decades. In Kuala Lumpur, we can see that it has a lot of amenities provided as offered by developed country. Therefore, I would like to know is there any similarity between Kuala Lumpur and Tokyo in terms of amenity.

For question 2, I will only focus on Kuala Lumpur, Malaysia. As of now, the number of shopping mall in Malaysia is very high and this is not a good sign. This is because there are many old shopping malls start to close due to cannot compete with the brand-new shopping malls and this will result many vacant buildings in very limited space.

## Hypothesis

To answer the first question, how should we define similarity or dissimilarity between two different cities. From my perspective, I always find different city has their own differences as compared to another city. Therefore, to tackle this question, I need to find the nearby venues within the areas/districts of Kuala Lumpur and Tokyo respectively. I will do the clustering model and cluster all of them and see is there any districts in Kuala Lumpur that are clustered together with districts in Tokyo. If all districts are clustered together, this implies that there is no difference between Kuala Lumpur and Tokyo (which is completely untrue and I can prove it to you). My hypothesis for this question is districts of Kuala Lumpur will be separately clustered from districts of Tokyo, Japan. This is because the two countries have two different people, culture and therefore the demand for the area is different.

For second question, my hypothesis is that there will be one cluster that has high number of shopping malls and there will be another cluster that has few or zero number of shopping malls. It is better to open it in an area that has few or zero shopping mall and the reasons are 1) to give the area new attraction and 2) to avoid the oversupply of shopping mall in one area. In my point of view, shopping mall is very essential as people can buy anything at just one place. Therefore, it always be a demand from the people to have the shopping centre near to their area.

# Data

In order to proceed with the analysis, the following data are required:

- 1) List of districts in Kuala Lumpur, Malaysia. This data can be found in the link here ([https://en.wikipedia.org/wiki/Kuala\\_Lumpur](https://en.wikipedia.org/wiki/Kuala_Lumpur))

## Districts [\[ edit \]](#)



Kuala Lumpur's eleven districts, with estimated population and percentage of the total, serve as administrative subdivisions under the authority of the [Kuala Lumpur City Hall](#) authority.<sup>[89]</sup>

1. [Bukit Bintang](#) (103,820 - 5.8%)
2. [Titiwangsa](#) (198,690 - 11.1%)
3. [Setiawangsa](#) (179,000 - 10.0%)
4. [Wangsa Maju](#) (227,330 - 12.7%)
5. [Batu](#) (91,290 - 5.1%)
6. [Kepong](#) (10,740 - 0.6%)
7. [Segambut](#) (125,300 - 7%)
8. [Lembah Pantai](#) (189,740 - 10.6%)
9. [Seputeh](#) (230,910 - 12.9%)
10. [Bandar Tun Razak](#) (273,870 - 15.3%)
11. [Cheras](#) (159,310 - 8.9%)

- 2) List of districts in Tokyo, Japan. This data can be found in the link here (<https://en.wikipedia.org/wiki/Tokyo>)

Special Wards of Tokyo			
	Place Name		Color
	Rōmaji	Kanji	
1	Adachi	足立区	Red
2	Arakawa	荒川区	Green
3	Bunkyo	文京区	Yellow
4	Chiyoda	千代田区	Orange
5	Chūō	中央区	Green
6	Edogawa	江戸川区	Green
7	Itabashi	板橋区	Yellow
8	Katsushika	葛飾区	Yellow
9	Kita	北区	Orange
10	Kōtō	江東区	Yellow
11	Meguro	目黒区	Orange
12	Minato	港区	Yellow
13	Nakano	中野区	Yellow
14	Nerima	練馬区	Green
15	Ōta	大田区	Yellow
16	Setagaya	世田谷区	Green
17	Shibuya	渋谷区	Red
18	Shinagawa	品川区	Green
19	Shinjuku	新宿区	Green
20	Suginami	杉並区	Orange
21	Sumida	墨田区	Orange
22	Taitō	台東区	Red
23	Toshima	豊島区	Red



A map of the Special Wards of Tokyo, showing the 23 wards and their corresponding colors. The map includes a scale bar for 10km. The wards are labeled with their names in English: Adachi, Arakawa, Bunkyo, Chiyoda, Chūō, Edogawa, Itabashi, Katsushika, Kita, Kōtō, Meguro, Minato, Nakano, Nerima, Ōta, Setagaya, Shibuya, Shinagawa, Shinjuku, Suginami, Sumida, Taitō, and Toshima. The map shows the geographical distribution of these wards, with some overlapping areas and a clear distinction between the different colors used to represent each ward.

- 3) Get the latitude and longitude by using the geocoder library or extract one by one from <https://www.latlong.net>. In this project, I'd used the latter because the former gave inaccurate latitude and longitude thus it the markers were pointed wrongly on the map that I've created using folium function.
- 4) Nearby places for each Kuala Lumpur and Tokyo districts. I can get this data by using Foursquare API.

# Methodology

## A. Question 1: Find the similarity/dissimilarity of two different cities of your choice.

### 1.1 Data preparation

I collect the required data from Wikipedia websites as specified above. Then, we can acquire their latitudes and longitudes by using the data retrieved from <https://www.latlong.net>. The other option is by using the geocoder library. However, I have tried this method and the location given is not so accurate. I realized it once I created a map of Kuala Lumpur, Malaysia. Next, Once the process of data collection is finished, I compile everything and merge them into two data frames called lat\_Ing\_kl for Kuala Lumpur and lat\_Ing\_tokyo for Tokyo.

#### Kuala Lumpur (first 5 rows)

	District	Latitude	Longitude
0	Bukit Bintang	3.14679	101.711
1	Titiwangsa	3.18067	101.703
2	Setiawangsa	3.18253	101.747
3	Wangsa Maju	3.20299	101.732
4	Batu	3.20467	101.672
5	Kepong	3.19987	101.639

#### Tokyo (first 5 rows)

	District	Latitude	Longitude
0	Adachi	35.7748	139.805
1	Arakawa	35.7361	139.783
2	Bunkyo	35.7076	139.752
3	Chiyoda	35.6939	139.754
4	Chūō	35.6706	139.772
5	Edogawa	35.707	139.868

### 1.2 Data analysis

Next, I will use the Foursquare API to retrieve the nearby venues within each district of Kuala Lumpur and Tokyo. In this process, I set the radius to be 5km. After we have found the nearby venues for each district, we can proceed with finding the top 10 most common nearby venues based on category. After some analysis is made, I can get the idea on the differences between Kuala Lumpur and Tokyo. Below are the figures to show first 5 rows of Kuala Lumpur and Tokyo districts plus their top 10 most common venues.

#### a) Kuala Lumpur (first 5 rows)

	District	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
0	Bandar Tun Razak	Chinese Restaurant	Café	Coffee Shop	Asian Restaurant	Ice Cream Shop	BBQ Joint	Burger Joint	Spa	Market	Malay Restaurant
1	Batu	Chinese Restaurant	Café	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Restaurant	Dessert Shop	Gym / Fitness Center	Motorcycle Shop	Malay Restaurant
2	Bukit Bintang	Hotel	Café	Shopping Mall	Spa	Chinese Restaurant	Cosmetics Shop	Bar	Bookstore	Hotel Bar	Clothing Store
3	Cheras	Chinese Restaurant	Café	Asian Restaurant	Hotel	Noodle House	Restaurant	Shopping Mall	Nightclub	Korean Restaurant	Trail
4	Kepong	Chinese Restaurant	Café	Dessert Shop	Coffee Shop	Park	Asian Restaurant	Thai Restaurant	Restaurant	Spa	Ice Cream Shop
5	Lembah Pantai	Hotel	Ice Cream Shop	Indian Restaurant	Shopping Mall	Chinese Restaurant	Café	Supermarket	Bakery	Convenience Store	Malay Restaurant

b) Tokyo (first 5 rows)

	District	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
0	Adachi	Ramen Restaurant	Café	Sake Bar	Dessert Shop	Japanese Restaurant	BBQ Joint	Coffee Shop	Convenience Store	Supermarket	Steakhouse
1	Arakawa	BBQ Joint	Sake Bar	Ramen Restaurant	Café	Art Museum	Japanese Restaurant	History Museum	Sukiyaki Restaurant	Soba Restaurant	Tonkatsu Restaurant
2	Bunkyo	Ramen Restaurant	Café	Hotel	BBQ Joint	Coffee Shop	Wagashi Place	History Museum	Tonkatsu Restaurant	Sake Bar	Japanese Curry Restaurant
3	Chiyoda	Hotel	Wagashi Place	Ramen Restaurant	BBQ Joint	Japanese Restaurant	Yoshoku Restaurant	Coffee Shop	Japanese Curry Restaurant	Sake Bar	Chinese Restaurant
4	Chūō	Hotel	Coffee Shop	Sushi Restaurant	Yoshoku Restaurant	Japanese Restaurant	Art Museum	Soba Restaurant	Sake Bar	Garden	Hotel Bar
5	Edogawa	Ramen Restaurant	Sake Bar	BBQ Joint	Dumpling Restaurant	Sushi Restaurant	Italian Restaurant	Park	Café	Indian Restaurant	Coffee Shop

### 1.3 Clustering and segmentation

Before I start clustering the data, I need to combine the two data frames into one. Once it is done, I proceed with the clustering process by using KMeans model from sklearn.cluster package. I will use the 'Elbow Method' to find which no. of cluster (n) gives a significant different in sum of square for error (SSE). I believe that  $n = 3$  gives the best result. After we have clustered the data, I proceed with analysing each cluster.

## B. Question 2: Where should developer open the new shopping mall in Kuala Lumpur?

### 1.1 Data preparation

I collect the required data from Wikipedia website to get the districts of Kuala Lumpur as specified above. Then, we can acquire their latitudes and longitudes by using the data retrieved from <https://www.latlong.net>. The other option is by using the geocoder library. However, I have tried this method and the location given is not so accurate. I realized it once I created a map of Kuala Lumpur, Malaysia. Next, Once the process of data collection is finished, I compile everything and merge them into one data frames called lat\_lng\_kl.

### 1.2 Data analysis

Next, I will use the Foursquare API to retrieve the nearby venues within each district of Kuala Lumpur. In this process, I set the radius to be 5km. After we have found the nearby venues for each district, we can proceed with finding the top 10 most common nearby venues based on category.

The category that I would like to focus here is the 'Shopping Mall'. Therefore, I drop all the categories except the 'Shopping Mall' as shown in the figure below and then I proceed to find the average of the 'Shopping Mall' category for district. The result as follows:

	District	Shopping Mall
0	Bukit Bintang	1
1	Bukit Bintang	0
2	Bukit Bintang	0
3	Bukit Bintang	0
4	Bukit Bintang	1
...	...	...
1095	Cheras	0
1096	Cheras	0
1097	Cheras	0
1098	Cheras	0
1099	Cheras	0

→

	District	Shopping Mall
0	Bandar Tun Razak	0.00
1	Batu	0.01
2	Bukit Bintang	0.03
3	Cheras	0.03
4	Kepong	0.01
5	Lembah Pantai	0.03
6	Segambut	0.02
7	Seputeh	0.04
8	Setiawangsa	0.02
9	Titiwangsa	0.03
10	Wangsa Maju	0.01

From the figure on the right, we can see that there Bandar Tun Razak has zero shopping mall value meanwhile Seputeh has the highest value. From the same figure, we also can see that almost all districts have their own shopping malls (note that the radius for this data is set at 5km) and this indicates that Kuala Lumpur has a lot of shopping malls.

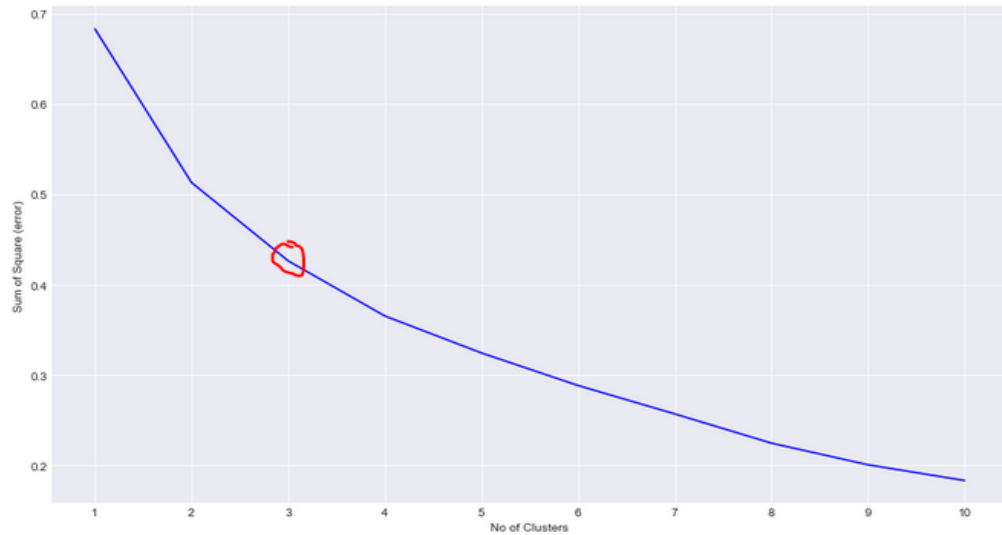
### 1.3 Clustering and segmentation

Next, I proceed with the clustering process by using KMeans model from sklearn.cluster package. I use the 'Elbow Method' to find which no. of cluster (n) gives a significant different in sum of square for error (SSE). I believe that  $n = 2$  gives the best result. After I have clustered the data, I proceed with analysing each cluster and then do data visualization.

# Results

## A. Question 1: Find the similarity/dissimilarity of two different cities of your choice.

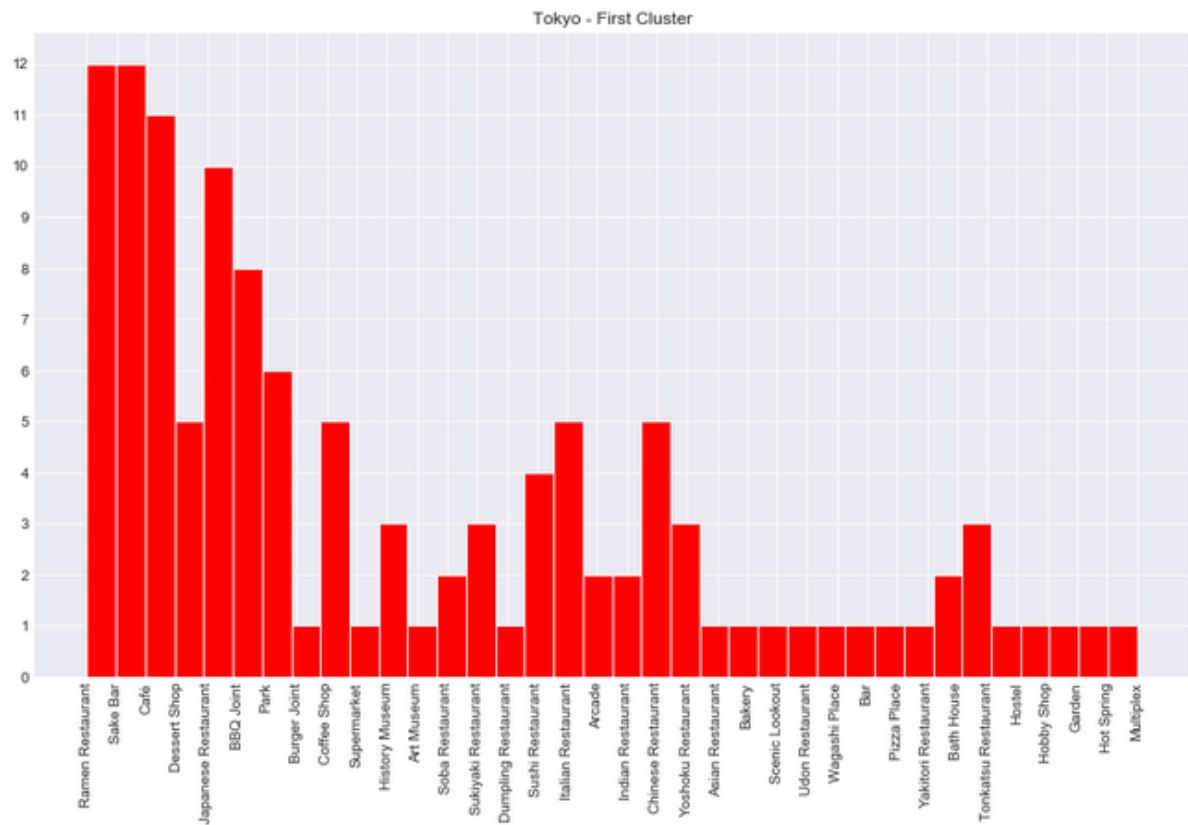
Based on Elbow Method that I have done, I found out that  $n=3$  that gives optimal result to the model.



Therefore, I have created 3 clusters which are Cluster0, Cluster1 and Cluster2. The details of each cluster as follows:

- Cluster0 (red marker)

District	Latitude	Longitude	Cluster	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Adachi	35.7748	139.805	0	Ramen Restaurant	Sake Bar	Café	Dessert Shop	BBQ Joint	Japanese Restaurant	Burger Joint	Dumpling Restaurant	Supermarket	Coffee Shop
Arakawa	35.7361	139.783	0	BBQ Joint	Sake Bar	Ramen Restaurant	Café	Japanese Restaurant	Art Museum	History Museum	Sukiyaki Restaurant	Soba Restaurant	Wagashi Place
Edogawa	35.707	139.868	0	Ramen Restaurant	Sake Bar	BBQ Joint	Dumpling Restaurant	Sushi Restaurant	Italian Restaurant	Park	Arcade	Coffee Shop	Café
Itabashi	35.7511	139.709	0	Sake Bar	Café	Sushi Restaurant	Japanese Restaurant	Ramen Restaurant	Chinese Restaurant	Yoshoku Restaurant	Asian Restaurant	Coffee Shop	Steakhouse
Katsushika	35.7435	139.847	0	Ramen Restaurant	Sake Bar	Dessert Shop	Park	Café	Arcade	Japanese Restaurant	Scenic Lookout	Italian Restaurant	BBQ Joint
Kita	35.7528	139.734	0	Sake Bar	Ramen Restaurant	Japanese Restaurant	Café	Yoshoku Restaurant	Chinese Restaurant	Sushi Restaurant	Dessert Shop	Udon Restaurant	Burger Joint
Nerima	35.7357	139.652	0	Ramen Restaurant	Sake Bar	BBQ Joint	Park	Café	Italian Restaurant	Coffee Shop	Bar	Shabu-Shabu Restaurant	Thai Restaurant
Suginami	35.6997	139.636	0	Ramen Restaurant	Park	Sake Bar	Café	Chinese Restaurant	Yakitori Restaurant	BBQ Joint	Italian Restaurant	Bath House	Indian Restaurant
Sumida	35.7107	139.802	0	Sake Bar	Ramen Restaurant	BBQ Joint	History Museum	Bath House	Japanese Restaurant	Tonkatsu Restaurant	Hostel	Sukiyaki Restaurant	Wagashi Place
Taitō	35.7126	139.78	0	Café	Ramen Restaurant	BBQ Joint	Sake Bar	Tonkatsu Restaurant	Electronics Store	History Museum	Hobby Shop	Sukiyaki Restaurant	Wagashi Place
Toshima	35.7362	139.714	0	Sake Bar	Café	Ramen Restaurant	Garden	Sushi Restaurant	Japanese Restaurant	Dessert Shop	Yoshoku Restaurant	Chinese Restaurant	Indian Restaurant
Ōta	35.5614	139.716	0	Ramen Restaurant	Chinese Restaurant	Sake Bar	Café	Tonkatsu Restaurant	Japanese Restaurant	Park	BBQ Joint	Hot Spring	Sushi Restaurant

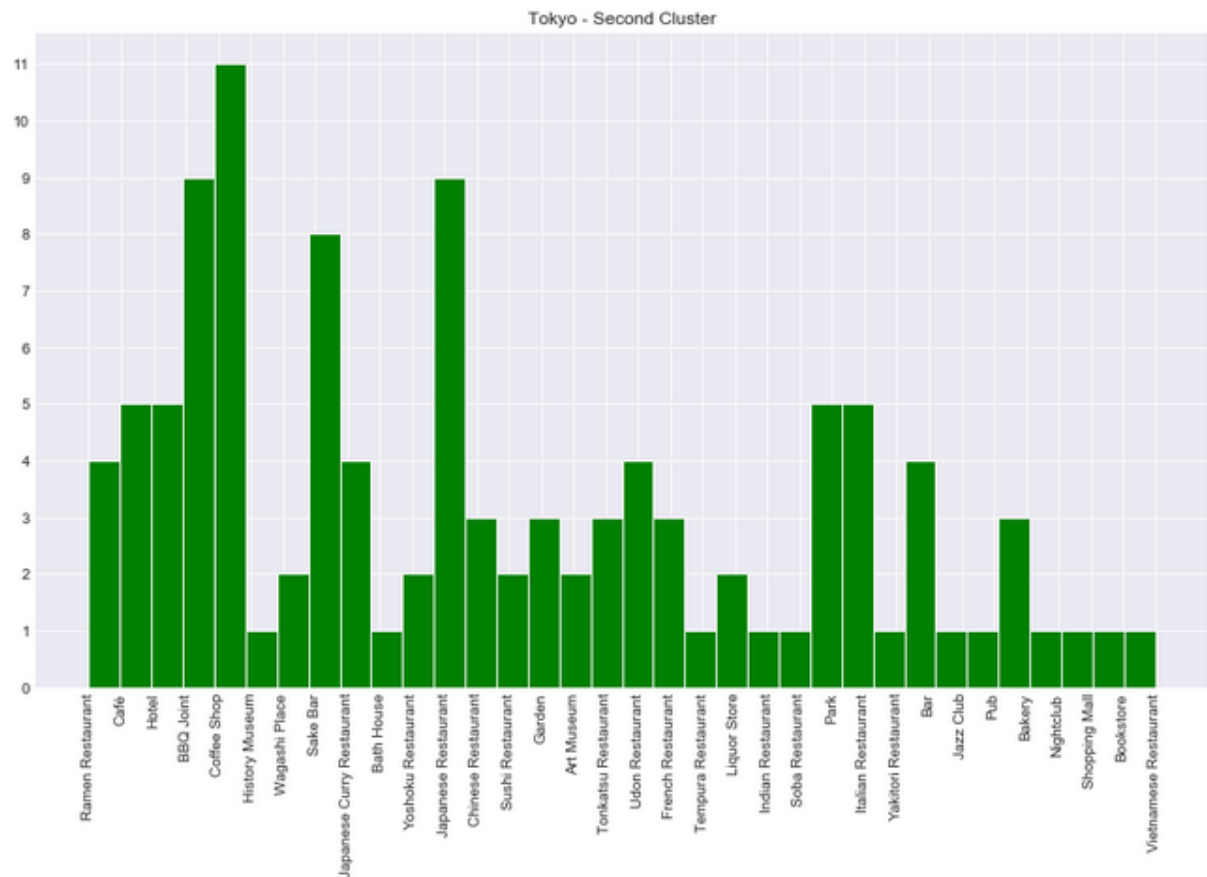


In this cluster, the top 5 most common venues are Ramen Restaurant, Sake Bar, Café, BBQ Joint and Japanese Restaurant.

- Cluster1 (green marker)

	District	Latitude	Longitude	Cluster	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
0	Bunkyo	35.7076	139.752	1	Ramen Restaurant	Hotel	BBQ Joint	Café	Coffee Shop	History Museum	Wagashi Place	Sake Bar	Tonkatsu Restaurant	Japanese Curry Restaurant
1	Chiyoda	35.6939	139.754	1	Hotel	Wagashi Place	Ramen Restaurant	BBQ Joint	Yoshoku Restaurant	Japanese Restaurant	Japanese Curry Restaurant	Sake Bar	Chinese Restaurant	Coffee Shop
2	Chūō	35.6706	139.772	1	Hotel	Yoshoku Restaurant	Sushi Restaurant	Japanese Restaurant	Coffee Shop	Soba Restaurant	Garden	Art Museum	Sake Bar	Concert Hall
3	Kōtō	35.6728	139.817	1	Japanese Restaurant	Sake Bar	Sushi Restaurant	BBQ Joint	Coffee Shop	Soba Restaurant	French Restaurant	Indian Restaurant	Tempura Restaurant	Seafood Restaurant
4	Meguro	35.6415	139.698	1	Coffee Shop	Japanese Restaurant	BBQ Joint	Park	Italian Restaurant	Café	Yakitori Restaurant	Bar	Bakery	Udon Restaurant
5	Minato	35.658	139.752	1	Hotel	Garden	BBQ Joint	Jazz Club	Tonkatsu Restaurant	Chinese Restaurant	Art Museum	Sushi Restaurant	French Restaurant	Italian Restaurant
6	Nakano	35.7073	139.664	1	BBQ Joint	Japanese Restaurant	Sake Bar	Italian Restaurant	Park	Udon Restaurant	Ramen Restaurant	Bar	Coffee Shop	Cocktail Bar
7	Setagaya	35.6465	139.653	1	Coffee Shop	Café	Park	Bakery	Sake Bar	Japanese Restaurant	BBQ Joint	Ramen Restaurant	Chinese Restaurant	Indian Restaurant
8	Shibuya	35.6618	139.704	1	Coffee Shop	Japanese Restaurant	BBQ Joint	Nightclub	Café	Bar	Dessert Shop	Italian Restaurant	Japanese Curry Restaurant	Sushi Restaurant
9	Shinagawa	35.6092	139.73	1	BBQ Joint	Sake Bar	Udon Restaurant	Coffee Shop	Park	Japanese Restaurant	Italian Restaurant	Tonkatsu Restaurant	Café	Hotel
0	Shinjuku	35.6938	139.703	1	Coffee Shop	Japanese Restaurant	Garden	Sake Bar	Bar	Hotel	Italian Restaurant	Vietnamese Restaurant	Bakery	Ramen Restaurant

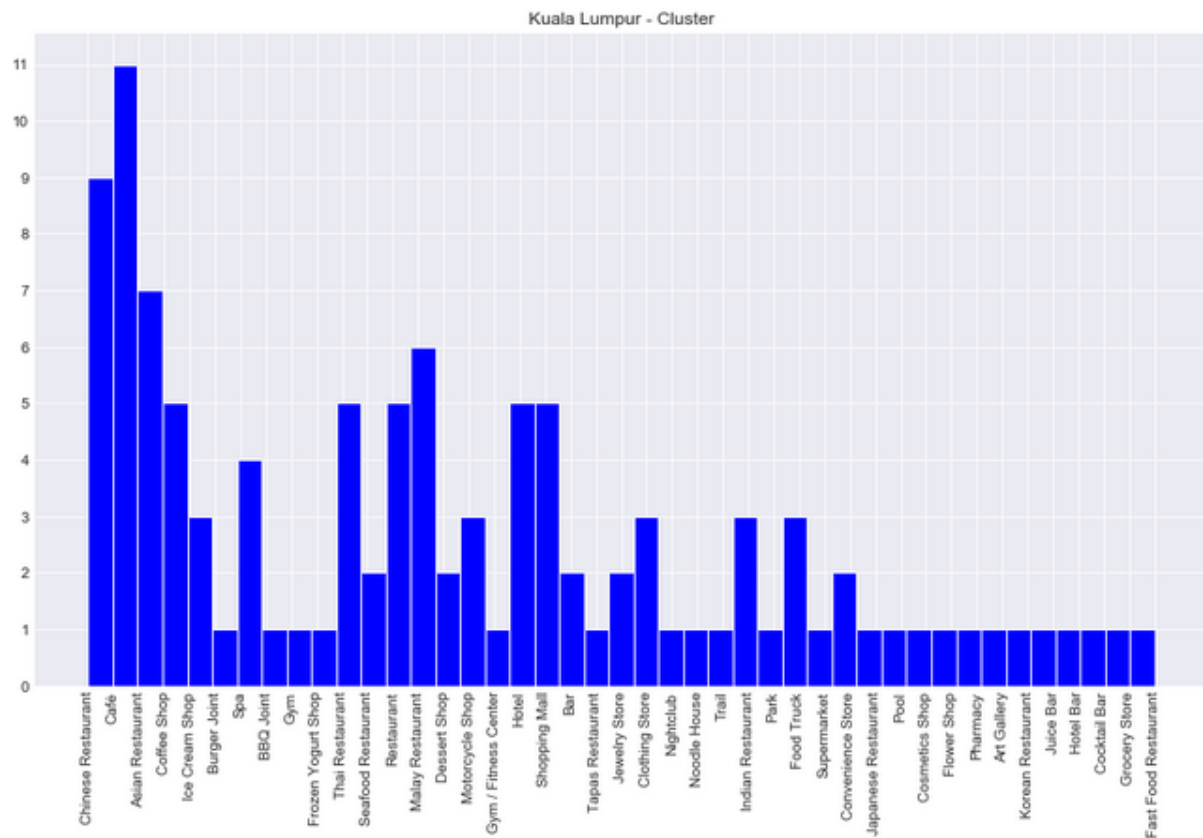




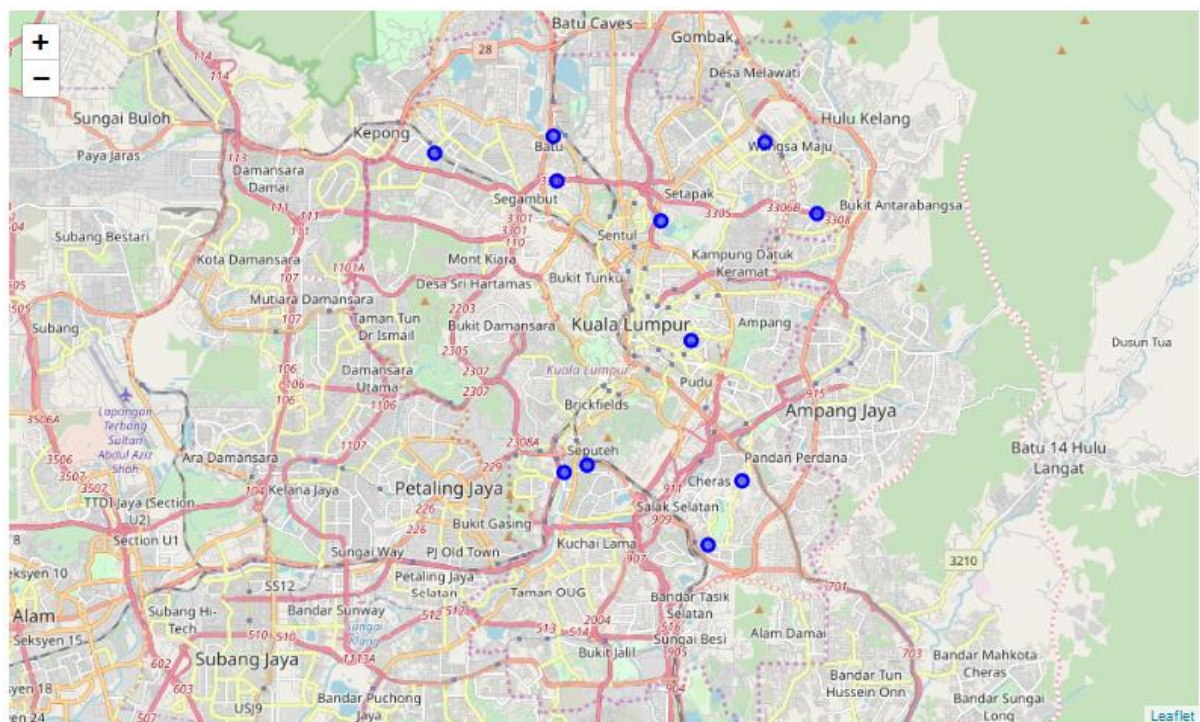
In this cluster, the top 5 most common venues are Coffee Shop, BBQ joint, Japanese Restaurant, Sake Bar and Hotel/Italian Restaurant.

- Cluster2 (blue marker)

District	Latitude	Longitude	Cluster	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Bandar Tun Razak	3.08863	101.716	2	Chinese Restaurant	Café	Ice Cream Shop	Coffee Shop	Asian Restaurant	Burger Joint	Spa	BBQ Joint	Japanese Restaurant	Seafood Restaurant
Batu	3.20467	101.672	2	Chinese Restaurant	Asian Restaurant	Café	Thai Restaurant	Seafood Restaurant	Restaurant	Malay Restaurant	Motorcycle Shop	Gym / Fitness Center	Dessert Shop
Bukit Bintang	3.14679	101.711	2	Hotel	Café	Shopping Mall	Spa	Cosmetics Shop	Chinese Restaurant	Bar	Lounge	Shoe Store	Hostel
Cheras	3.10684	101.726	2	Chinese Restaurant	Café	Asian Restaurant	Hotel	Nightclub	Restaurant	Shopping Mall	Noodle House	Furniture / Home Store	Malay Restaurant
Kepong	3.19987	101.639	2	Chinese Restaurant	Café	Dessert Shop	Coffee Shop	Park	Asian Restaurant	Thai Restaurant	Spa	Restaurant	Grocery Store
Lembah Pantai	3.1094	101.675	2	Hotel	Indian Restaurant	Ice Cream Shop	Café	Clothing Store	Shopping Mall	Malay Restaurant	Cosmetics Shop	Convenience Store	Bakery
Segambut	3.19209	101.673	2	Café	Chinese Restaurant	Japanese Restaurant	Asian Restaurant	Thai Restaurant	Spa	Seafood Restaurant	Motorcycle Shop	Pool	Shopping Mall
Seputeh	3.11138	101.682	2	Hotel	Indian Restaurant	Shopping Mall	Ice Cream Shop	Clothing Store	Cosmetics Shop	Chinese Restaurant	Malay Restaurant	Optical Shop	Noodle House
Setiawangsa	3.18253	101.747	2	Malay Restaurant	Coffee Shop	Café	Food Truck	Restaurant	Art Gallery	Korean Restaurant	Juice Bar	Asian Restaurant	Thai Restaurant
Titivangsa	3.18067	101.703	2	Hotel	Malay Restaurant	Shopping Mall	Hotel Bar	Dance Studio	Chinese Restaurant	Bar	Coffee Shop	Fountain	Dessert Shop
Wangsa Maju	3.20299	101.732	2	Malay Restaurant	Chinese Restaurant	Coffee Shop	Café	Asian Restaurant	Food Truck	Motorcycle Shop	Convenience Store	Thai Restaurant	Boutique



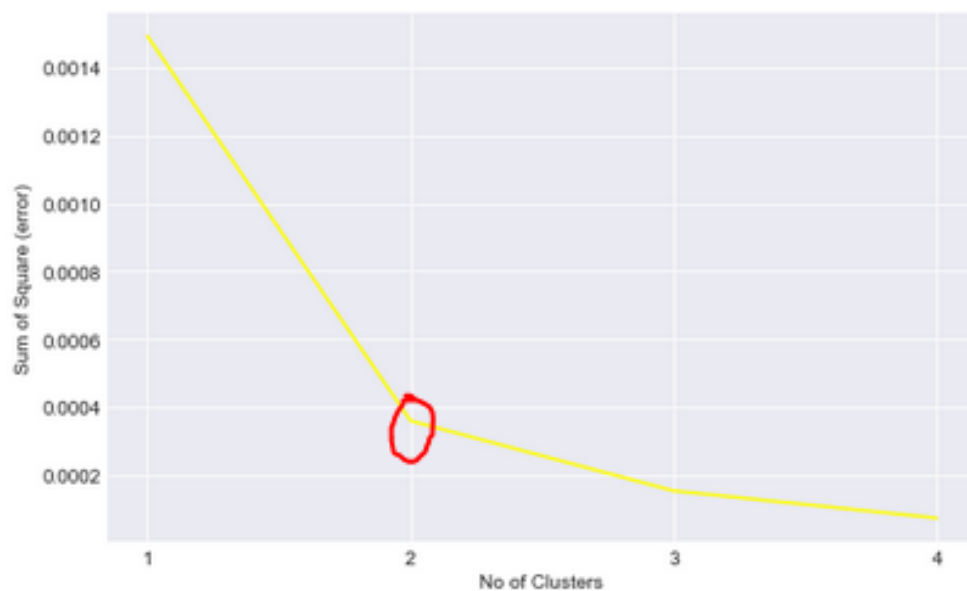
In this cluster, the top 5 most common venues are Café, Chinese Restaurant, Asian Restaurant, Malay Restaurant and Hotel/Shopping Mall.





## B. Question 2: Where should developer open the new shopping mall in Kuala Lumpur?

Based on Elbow Method that I have done, I found out that  $n=2$  that gives optimal result to the model.





Then, I clustered the data into two namely cluster0 and cluster1.

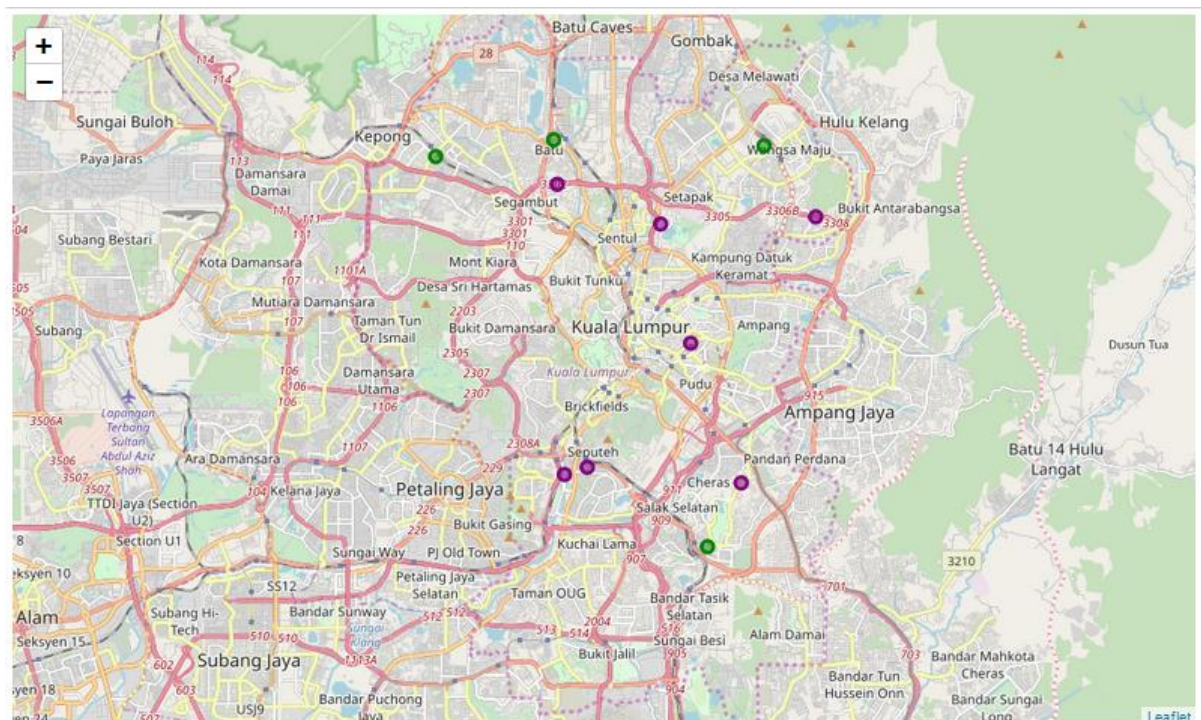
- Cluster0 (green marker)

	District	Shopping Mall	Latitude	Longitude	Cluster
0	Bandar Tun Razak	0.00	3.08863	101.716	0
1	Batu	0.01	3.20467	101.672	0
2	Kepong	0.01	3.19987	101.639	0
3	Wangsa Maju	0.01	3.20299	101.732	0

- Cluster1 (purple marker)

	District	Shopping Mall	Latitude	Longitude	Cluster
0	Bukit Bintang	0.03	3.14679	101.711	1
1	Cheras	0.03	3.10684	101.726	1
2	Lembah Pantai	0.03	3.1094	101.675	1
3	Segambut	0.02	3.19209	101.673	1
4	Seputeh	0.04	3.11138	101.682	1
5	Setiawangsa	0.02	3.18253	101.747	1
6	Titiwangsa	0.03	3.18067	101.703	1

As we can see, the Cluster0 has 'Shopping Mall' value of less and equal to 0.01. Meanwhile, Cluster1 has the value or more than 0.01.



# Discussion

## A. Question 1: Find the similarity/dissimilarity of two different cities of your choice.

From the result that I have obtained, we can see that cluster0 and cluster1 consist of Tokyo districts only and on the other hand, cluster2 only consist of Kuala Lumpur districts. Cluster0 and cluster1 have Japanese Restaurant and Sake Bar in their top 5 most common venues near the area. For Kuala Lumpur city, the most demand category is Chinese Restaurant and then followed by Café. This shows that the demand from the people who live within the area plays an important role to shape the city. As for Japanese, they love to eat Japanese/local food and enjoy drinking sake after the work time is over. Meanwhile, Malaysian loves to eat rice and usually can be found at Chinese Restaurant. They also love to spend their time having a discussion/meeting in a café.

## C. Question 2: Where should developer open the new shopping mall in Kuala Lumpur?

From the result that I have obtained, I found that almost all districts of Kuala Lumpur have their own shopping malls. In this project, I use the radius of 5 km and to have so many shopping malls within this small radius indicates that there is oversupply in shopping mall. To build more shopping mall has its own pros and cons. The pros are can attract more people and therefore can encourage them to spend more and this will contribute to positive growth in GDP of a country. The cons are will be many 'dead' shopping malls or abandoned buildings that tear down the landscape of the area. However, there is a good location to open a new shopping mall which is at Bandar Tun Razak as the number of existing shopping mall is zero at the moment. This will give a new attraction to the place. In my point of view, shopping mall is very essential to the people as this make their life easier as they can buy everything under one roof. Therefore, it is always a demand for shopping mall and it is good to open it at a location that has less or zero number of shopping malls nearby.

# Conclusion

My earlier hypotheses are true for question 1 and 2. Kuala Lumpur and Tokyo are completely two different cities and we cannot find the similarity based on amenity. This is because Malaysia and Japan have different people and culture that shape the city to be different from one another. In Kuala Lumpur, we can cluster the districts into two which are 'high' and 'low' in terms of number of shopping malls. The best location to open a new shopping mall is in cluster0 or to be precise at Bandar Tun Razak as it has zero shopping mall at the moment.