Identifying Ideal Districts in Malaysia for Establishing New Gym Branches Using K-means Clustering

Introduction

Malaysia has the highest rate of obesity and overweight among Asian countries with 64% of the male and 65% of the female population being either obese or overweight. [1] The Malaysian government has also taken many initiatives in the past several years to promote healthier lifestyle and living among Malaysians. As a result, there is increasing awareness about fitness among the younger generation of the country. The health and fitness industry in Malaysia are booming as of now. Currently, there is an increasing demand for more gyms and fitness centres in Malaysia.

Business Problem

Realising the market potential of building a successful business in the fitness sector in Malaysia, a hypothetical fitness-based company called “A Fitness” is interested in establishing a gym chain with branches all over the Peninsular of Malaysia. The company is looking to open their gym branches at locations which would attract a huge customer base and guarantee them a regular profit.

Problem statement

The goal of this project is to help the company identify the most ideal locations to open their gym centres. The problem statement is, “Which are the most suitable districts to open new gym centres in the Peninsular of Malaysia?

Discussion of background

It is important to narrow down the scope of the project and to take certain important criteria into consideration. Important aspects to consider are for example:

i) the location area,
ii) the population of the area
iii) the general age of residents in the particular area,

iv) average income of resident in the area,
v) number of gyms or fitness centres already available in the area
vi) number of shopping malls in the area

a) Location

Image 1 [2]: Map of Peninsular Malaysia

Only the states in the Peninsular of Malaysia are taken into account in this project. The two states in east Malaysia are excluded in this project. There is a total of 11 states and up to 87 districts in Peninsular Malaysia. From the 11 states, the state of Kuala Lumpur (the capital of Malaysia) is excluded also as it already has many gym centres and it would be difficult to make an impact in an already competitive location. It would not make much business sense to open more gyms in an already saturated location. The Klang Valley consists of Kuala Lumpur and areas close to the capital, bordering Selangor.

As of now, the focus of the project would be on the main districts in the states of Perak, Terengganu, Malacca, Johor, Negeri Sembilan, Pahang, Penang, Kedah, Kelantan, and Selangor.

b) Population

The total population of Malaysia in 2020 is estimated to be 32.7 million people [3]. Peninsular Malaysia accounts for the majority (roughly 81.3%) of Malaysia's population [4]. The state of Selangor is the most populous state in Malaysia, while Penang has the

highest population density. In general, urbanisation is more prominent in areas with high population and more shopping malls are available at these locations.

The company prefers to open their gym centres at highly populated locations.

c) Age

The target customer base for the gym centres are people within the ages of 20 to 50 years old. It is assumed that Malaysians in their 30's and 40's are generally more health conscious compared to Malaysian of other age groups. Malaysian above 50's generally prefers more lighter forms of exercise such as brisk walking or jogging and may prefer going to parks instead.

Districts with higher population of people in their 30's and 40's are preferred by the company.

d) Income

Standard gym membership fees of gyms in Klang Valley is within MYR 90 to MYR 300 per month [5]. In 2019, the mean income in Malaysia was RM7,901 while Malaysia's median income was RM5,873 [6]. Median monthly household disposable income was RM5,116 in 2019, and mean monthly household disposable income was RM6,764. Mean disposable income comprises of 85.6 per cent of total mean gross income. [6]. People living in urban districts earn more than those who live in rural districts.

Taking into consideration the general cost of living of a regular Malaysian, those who earn lesser then MYR 3000 per household are unlikely to spend on gym membership fees. The target customers are people who earn MYR 4000 and above. It is also assumed that people with higher income would probably be willing to spend more on exclusive gym membership deals.

e) Gyms and Fitness centre

There are already a few big gym chains in Malaysia as of now and most of them are located in Kuala Lumpur or areas in the Klang Valley. It is preferable to open gyms at districts with little to no gym centres currently available. The competition for market share would be very high in areas with many other gym and fitness centres. It would be very difficult to make a profit in a highly competitive area as customers would prefer to frequent gym which charge the lowest possible gym membership fees.

f) Shopping malls

Most of the gyms in the Klang Valley for example, are located in shopping malls or very close to shopping malls. As many people visit shopping malls on a regular basis, it would make sense to open gyms in such locations. The company “A Fitness” prefers areas with a high number of shopping malls.

Description of Data

The main data required for this project are as follows:

i) the location of area
ii) the population of the area
iii) the main age of residents in the particular area,
iv) average income of resident in the area,
v) number of gyms or fitness centres already available in the area
vi) number of shopping malls in the area

a) Location Data

The list of districts in each state of Peninsular Malaysia were extracted by web scraping the page, https://en.wikipedia.org/wiki/List_of_districts_in_Malaysia [7] using the BeautifulSoup library. These values were stored into a pandas dataframe. A total of 87 district of 11 states were extracted but some districts were dropped from the dataset as no population information were available for those specific districts. The final dataset has 84 districts.

The GeoPy library were used to obtain the latitudes and longitudes of each district in each state. These values were stored in a pandas dataframe. The district names, and their respective latitude and longitudes were merged together and saved in a .csv format file titled “district_2.csv”.

The Folium library was used to create a map of these location points on the map of Peninsular Malaysia (refer Figure 2)

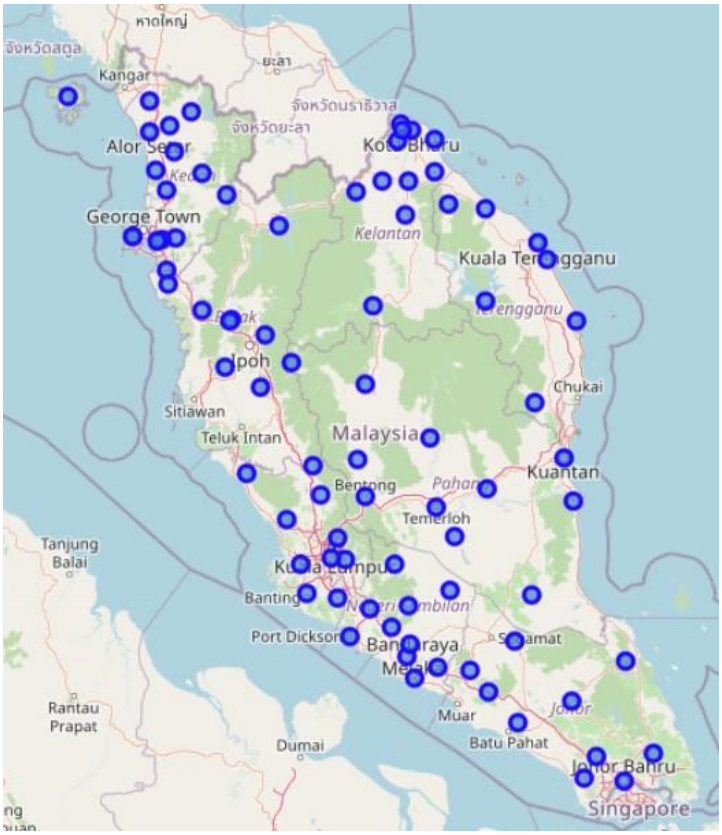
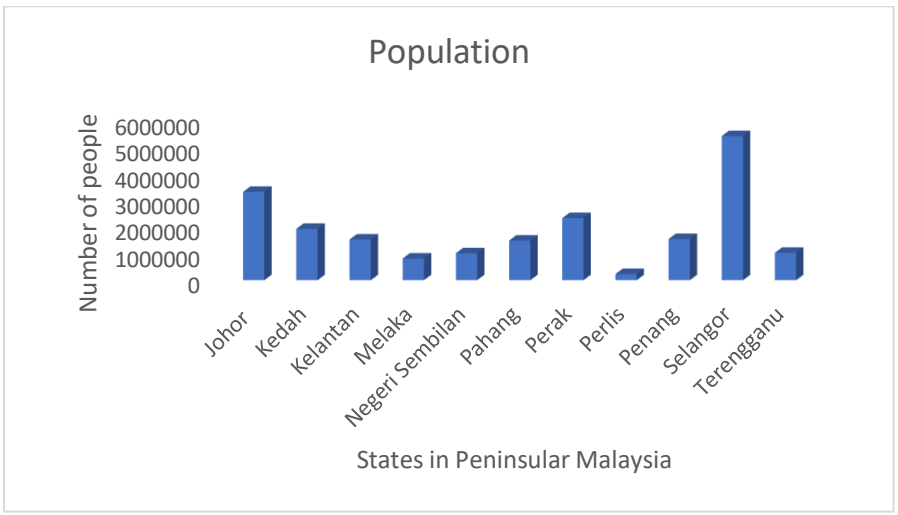


Figure 1 [7]: Map of districts in Peninsular Malaysia

b) Population Data

Population data of each district were obtained from the website https://www.citypopulation.de/en/malaysia/admin/ [8]. The population data were extracted and added to the district_2.csv file. Graph 1 depicts the overall population of each state in Peninsular Malaysia. Selangor has the highest population and Perlis has the lowest population overall.



Graph 1 [8]: Population of states in Peninsular Malaysia

c) Age of Residents

The main age groups of residents of each district were extracted from http://pqi.stats.gov.my/searchBI.php?tahun=2020&kodData=3&kodJadual=1&kodCiri=&kodNegeri= [9] and added to the main dataset in district_2.csv file. The main age groups of residents were divided into several age range categories as shown in Table 2. The majority of the residents in the districts fall within the top 3 age groups. Since there were no specific data for each district, the age data of each state was used instead. The top 3 age groups together account for the more than 50 percent of the district population. Table 3 shows a snippet of the age data used in the final main dataset.

Age Range	Category
0 to 9	1
10 to 19	2
20 to 29	3
30 to 39	4
40 to 49	5

Table 2: Age range and category label

District	1st most age group	2nd most age group	3rd most age group
Batu Pahat	3	4	2
Johor Bahru	3	4	2
Kluang	3	4	2
Kota Tinggi	3	4	2
Kulai	3	4	2
Mersing	3	4	2
Muar	3	4	2
Pontian	3	4	2
Segamat	3	4	2
Tangkak	3	4	2
Baling	3	2	1
Bandar Baharu	3	2	1
Kota Setar	3	2	1

Table 3: Snippet of top 3 age groups of districts dataset

d) Income of residents

The overall mean and median household income of each district was extracted and added to the main dataset. Since there were no specific data for each district, the median and mean household income data of each state was used instead. The mean and median household income of each state was obtained from the Department of Statistics Malaysia Official [10].

e) Number of gyms

The Foursquare API was used to collect information on the locations and number of gyms available in each district. The Foursquare API was used to search venues of category gym with Id= 4bf58dd8d48988d176941735. The number of gyms were extracted into a dataframe and added to the main dataset. The Foursquare API returns results up to a maximum default limit of 50 gym venues for each district. Number of gyms of 50 indicates that there could actually be more the 50 gyms in the specific districts.

f) Number of shopping malls

The Foursquare API was used to collect information on the locations and number of shopping malls available in each district. The Foursquare API was used to search venues of category shopping malls with Id= 4bf58dd8d48988d1fd941735. The number of gyms were extracted into a dataframe and added to the main dataset. The Foursquare API returns results up to a maximum default limit of 50 shopping mall venues for each district. Number of gyms of 50 indicates that there could actually be more the 50 shopping malls in the specific districts.

Methodology

Collecting and Preparing Data

i) The district names were extracted from a Wikipedia page, *List Of Districts In Malaysia* [7] using the BeautifulSoup and request library. The URL of the website was parsed into the request.get() function.

ii) All text with html tag 'a' were extracted using the function call find_all(). The function call returned bs4.element.ResultSet type data. This data was converted into a pandas series data and then converted into a pandas dataframe.

iii) Only rows which contain district names were kept and were converted to string to strip all other text other than the district names string itself. Some rows were also further dropped as some districts did not have population data. The final district names were saved in a .csv file.

iv) The latitude and longitude values of each districts were obtained using the GeoPy library and Nominatim. The latitude and longitude values were obtained by looping thru the district name dataframe extracted earlier.

v) The age, population and income data were extracted from various websites and stored in the district_2.csv file together with the district's names and latitude and longitude values.

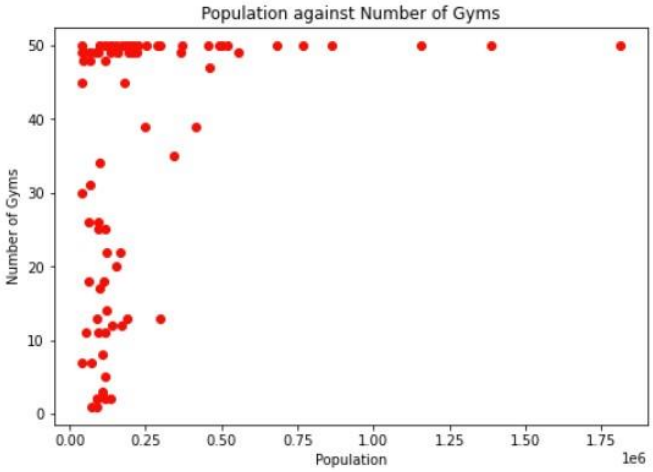
vi) The Foursquare API was used to request a search of id category “gym” within radius of 30000m. The request returned a dictionary type json data with ‘response’ key and value of ‘venues’, for gym. A function call was defined to loop thru the district names in the dataframe and send requests to the Foursquare API for each district. The names of gyms in each district were displayed using print(). And the number of gyms for each district were saved in a separate dataframe and added into the Full.csv main dataset file. The same is repeated for id category “shopping mall’

vii) The final dataset was standardised using StandardScaler()

Exploratory Data Analysis and Inferential Statistic Testing

Scatter plots were plotted to understand the relationship between the population, income and number of shopping malls with the number of gyms available in a district.

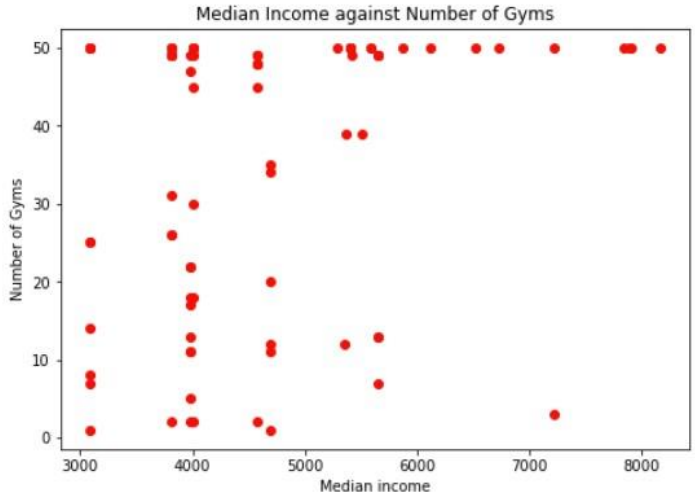
i) Scatter plot 1: Population against Number of Gyms



Observation:

The number of gyms does not depend on the population. Number of gyms is 50 or more for districts with population of 250000 or more in general.

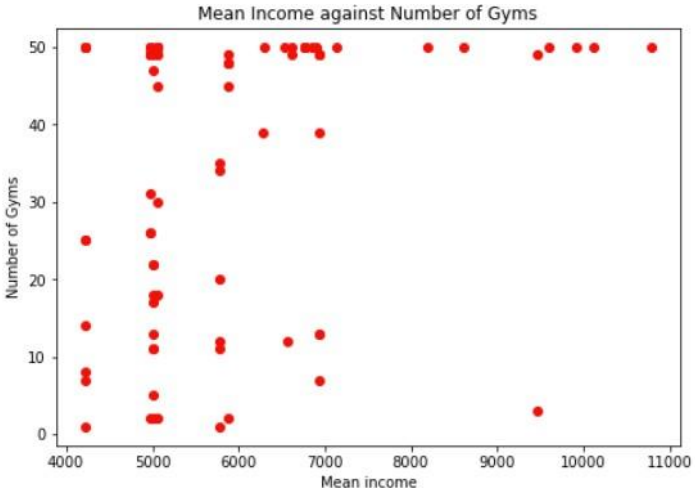
ii) Scatter plot 2: Median Income against Number of Gyms



Observation:

No direct relationship between median income and number of gyms.

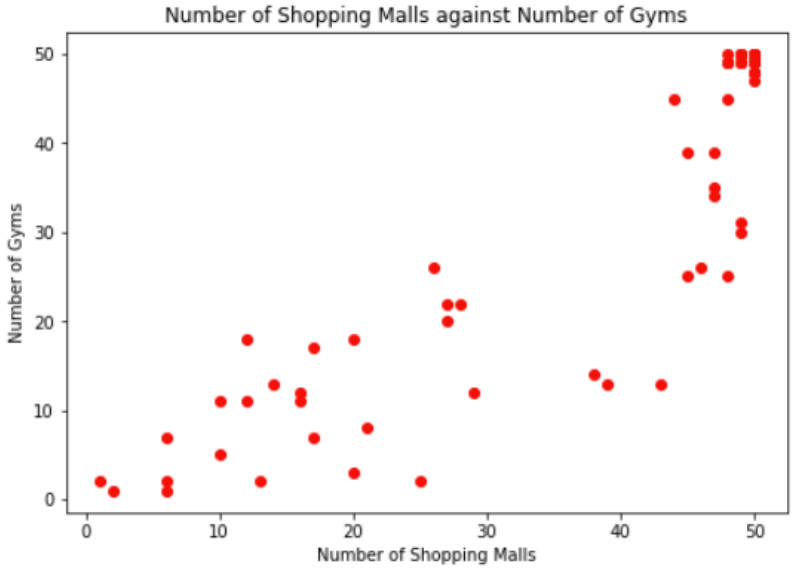
iii) Scatter plot 3: Mean Income against Number of Gyms



Observation:

No direct relationship between mean income and number of gyms.

iv) Scatter plot 4: Number of Shopping Malls against Number of Gyms



Observation:

The number gyms increase as the number of shopping mall increases.

Pearson Correlation Coefficient Calculation was also calculated to determine the correlation between the number of gyms and the other variables. Results are tabulated in the Table 4 below:

Variables	**Correlation Coefficient**
Population	0.3447872721927314
Median Income	0.24786183801352116
Mean Income	0.2514499086542497
Number of Shopping Malls	0.9074854397525329

Table 4: Pearson Correlation Coefficient Calculation

Observation:

There is a weak or no correlation between number of gyms and population, median income and mean income of a district. There is a strong positive correlation between number of shopping malls and number of gyms in a district.

Clustering Using k-Means Clustering

The machine learning method K-Means Clustering were used to cluster the 84 districts in to 7 different cluster group. Each cluster was labelled. The most suitable districts to open new gym branches based on the criteria discussed in the Discussion of background section could be determined based on the characteristics of the districts in the clusters. The clusters are mapped using Folium to visualise the clusters.

Results

The 7 output clusters are labelled as Cluster 0 to Cluster 6 respectively and the resulting clusters are as shown below:

Clusters 0

Cluster Labels	District	Latitude	Longitude	Population	Median Income per month	Mean Income per month	Number of Shopping Malls	Number of Gym/fitness centre	1st most age group	2nd most age group	3rd most age group
0	Baling	5.672928	100.918204	135646	3811	4971	13	2	3	2	1
0	Gua Musang	4.879272	101.965233	90057	3079	4214	2	1	3	1	2
0	Jeli	5.695523	101.844937	40637	3079	4214	17	7	3	1	2
0	Kuala Krai	5.530017	102.199106	109461	3079	4214	21	8	3	1	2
0	Tanah Merah	5.766305	102.034148	121319	3079	4214	38	14	3	1	2
0	Jempol	2.850272	102.508196	116576	4579	5887	25	2	3	1	2
0	Bentong	3.520932	101.910618	119817	3979	5012	28	22	3	2	1
0	Bera	3.236674	102.541772	97882	3979	5012	17	17	3	2	1
0	Jerantut	3.936659	102.371955	91096	3979	5012	14	13	3	2	1
0	Lipis	4.324251	101.908585	89730	3979	5012	6	2	3	2	1
0	Maran	3.582261	102.774742	115424	3979	5012	12	11	3	2	1
0	Pekan	3.490073	103.393699	110633	3979	5012	20	18	3	2	1
0	Raub	3.791261	101.853491	95506	3979	5012	10	11	3	2	1
0	Rompin	2.824577	103.099356	114901	3979	5012	10	5	3	2	1
0	Temerloh	3.448334	102.417571	165451	3979	5012	27	22	3	2	1
0	Hulu Perak	5.447973	101.297149	91218	4006	5065	1	2	3	2	1
0	Muallim	3.738620	101.541056	64254	4006	5065	12	18	3	2	1

Observation:

The districts in Cluster 0 generally have **low population** with close to 100000 residents per district, **low median and mean income** per month, **low total number of shopping malls, low total number of existing gyms,** the majority of residents are within the **age group of 0 to 29 years old.**

Cluster 1

Cluster Labels	District	Latitude	Longitude	Population	Median Income per month	Mean Income per month	Number of Shopping Malls	Number of Gym/fitness centre	1st most age group	2nd most age group	3rd most age group
1	Bandar Baharu	5.131119	100.497279	42341	3811	4971	49	50	3	2	1
1	Kota Setar	6.122603	100.369032	366787	3811	4971	49	49	3	2	1
1	Kuala Muda	5.707315	100.495289	456605	3811	4971	50	50	3	2	1
1	Kubang Pasu	6.335638	100.372322	220740	3811	4971	48	49	3	2	1
1	Kulim	5.366377	100.552571	287694	3811	4971	50	50	3	2	1
1	Langkawi	6.370039	99.792863	94777	3811	4971	26	26	3	2	1
1	Padang Terap	6.256651	100.664580	62896	3811	4971	46	26	3	2	1
1	Pendang	5.972084	100.550185	94962	3811	4971	49	49	3	2	1
1	Pokok Sena	6.167313	100.517432	49506	3811	4971	48	49	3	2	1
1	Sik	5.818854	100.744277	67378	3811	4971	49	31	3	2	1
1	Yan	5.848275	100.415190	68319	3811	4971	49	49	3	2	1
1	Bachok	6.062410	102.399981	133152	3079	4214	48	50	3	1	2
1	Kota Bharu	6.125605	102.239345	491237	3079	4214	49	50	3	1	2
1	Machang	5.764113	102.214217	93087	3079	4214	48	25	3	1	2
1	Pasir Mas	6.049459	102.139944	189292	3079	4214	49	50	3	1	2
1	Pasir Puteh	5.835534	102.405045	117383	3079	4214	45	25	3	1	2
1	Tumpat	6.176396	102.166328	153976	3079	4214	49	50	3	1	2
1	Cameron Highlands	4.478169	101.377600	38471	3979	5012	50	49	3	2	1
1	Batang Padang	6.132683	102.176427	115240	4006	5065	49	50	3	2	1
1	Hilir Perak	4.781497	100.949657	208570	4006	5065	50	49	3	2	1
1	Kampar	4.300660	101.161446	98978	4006	5065	50	50	3	2	1
1	Kerian	5.034575	100.498458	179706	4006	5065	48	45	3	2	1
1	Kinta	4.672468	101.197556	767794	4006	5065	50	50	3	2	1
1	Kuala Kangsar	4.772071	100.940878	159505	4006	5065	50	49	3	2	1
1	Larut	4.853766	100.743605	38399	4006	5065	49	30	3	2	1
1	Perak Tengah	4.449782	100.905392	101128	4006	5065	50	50	3	2	1

Observation:

The districts in Cluster 1 generally have **medium level population** with more than 100000 residents per district, **low median and mean income** per month, **high total number of shopping malls, high total number of existing gyms,** the majority of residents are within the **age group of 0 to 29 years old.**

Cluster 2

Cluster Labels	District	Latitude	Longitude	Population	Median Income per month	Mean Income per month	Number of Shopping Malls	Number of Gym/fitness centre	1st most age group	2nd most age group	3rd most age group
2	Johor Bahru	1.495304	103.755084	1386569	6518	8198	50	50	3	4	2
2	Gombak	3.233044	101.714574	682226	7903	9908	50	50	4	3	1
2	Hulu Langat	3.070825	101.763669	1156585	7851	9593	50	50	4	3	1
2	Hulu Selangor	3.536744	101.594935	198132	5421	9463	50	49	4	3	1
2	Klang	3.043125	101.449611	861189	6724	8606	50	50	4	3	1
2	Kuala Selangor	3.362102	101.345550	209590	7225	6532	50	50	4	3	1
2	Petaling	3.086134	101.664207	1812633	7904	10792	50	50	4	3	1
2	Sabak Bernam	3.687115	101.058059	105777	7225	9463	20	3	4	3	1
2	Sepang	2.800862	101.709401	211361	8174	10121	50	50	4	3	1

Observation:

The districts in Cluster 2 generally have **high population** with more than 100000 residents per district, **high median and mean income** per month, **high total number of shopping malls, high total number of existing gyms,** the majority of residents are within the **age group of 20 to 39 years old.**

Cluster 3

Cluster Labels	District	Latitude	Longitude	Population	Median Income per month	Mean Income per month	Number of Shopping Malls	Number of Gym/fitness centre	1st most age group	2nd most age group	3rd most age group
3	Besut	5.610237	102.506370	140952	4694	5776	29	12	1	3	2
3	Dungun	4.769677	103.416885	154932	4694	5776	27	20	1	3	2
3	Hulu Terengganu	4.910689	102.765070	72052	4694	5776	6	1	1	3	2
3	Kemaman	4.192789	103.119487	171383	5355	6565	16	12	1	3	2
3	Kuala Terengganu	5.329701	103.138313	343284	4694	5776	47	35	1	3	2
3	Marang	5.206546	103.204279	97857	4694	5776	47	34	1	3	2
3	Setiu	5.569059	102.763677	55517	4694	5776	16	11	1	3	2

Observation:

The districts in Cluster 3 generally have **low population** with close to 100000 residents per district, **medium level median and mean income** per month, **medium level total number of shopping malls, medium level total number of existing gyms,** the majority of residents are within the **age group of 0 to 29 years old.**

Cluster 4

Cluster Labels	District	Latitude	Longitude	Population	Median Income per month	Mean Income per month	Number of Shopping Malls	Number of Gym/fitness centre	1st most age group	2nd most age group	3rd most age group
4	Batu Pahat	1.908342	102.995623	417458	5516	6277	47	39	3	4	2
4	Kluang	2.069914	103.381181	298332	5652	6928	43	13	3	4	2
4	Kota Tinggi	1.688710	103.964021	193210	5652	6928	50	49	3	4	2
4	Kulai	1.671807	103.561497	251650	6114	6901	50	50	3	4	2
4	Mersing	2.355646	103.765113	70894	5652	6928	6	7	3	4	2
4	Muar	2.132507	102.784776	247957	5371	6928	45	39	3	4	2
4	Pontian	1.515380	103.474217	155541	5652	6928	50	49	3	4	2
4	Segamat	2.489551	102.972858	189820	5652	6928	39	13	3	4	2
4	Tangkak	2.284004	102.650964	136852	5652	6928	48	49	3	4	2

Observation:

The districts in Cluster 4 generally have **medium level population** with more than 100000 residents per district, **medium level median and mean income** per month, **high total number of shopping malls, medium level total number of existing gyms,** the majority of residents are within the **age group of 10 to 39 years old.**

Cluster 5

Cluster Labels	District	Latitude	Longitude	Population	Median Income per month	Mean Income per month	Number of Shopping Malls	Number of Gym/fitness centre	1st most age group	2nd most age group	3rd most age group
5	Central Seberang Perai	5.353979	100.457346	371975	5409	6771	50	50	3	4	5
5	North Seberang Perai	5.342175	100.427949	295979	5409	6771	50	50	3	4	5
5	Northeast Penang Island	5.378792	100.250172	520242	5409	6771	50	50	3	4	5
5	South Seberang Perai	5.340490	100.427137	171045	5409	6771	50	50	3	4	5
5	Southwest Penang Island	5.378792	100.250172	202142	5409	6771	50	50	3	4	5

Observation:

The districts in Cluster 5 generally have **high level population** with more than 100000 residents per district, **medium level median and mean income** per month, **high total number of shopping malls, high total number of existing gyms,** the majority of residents are within the **age group of 20 to 49 years old.**

Cluster 6

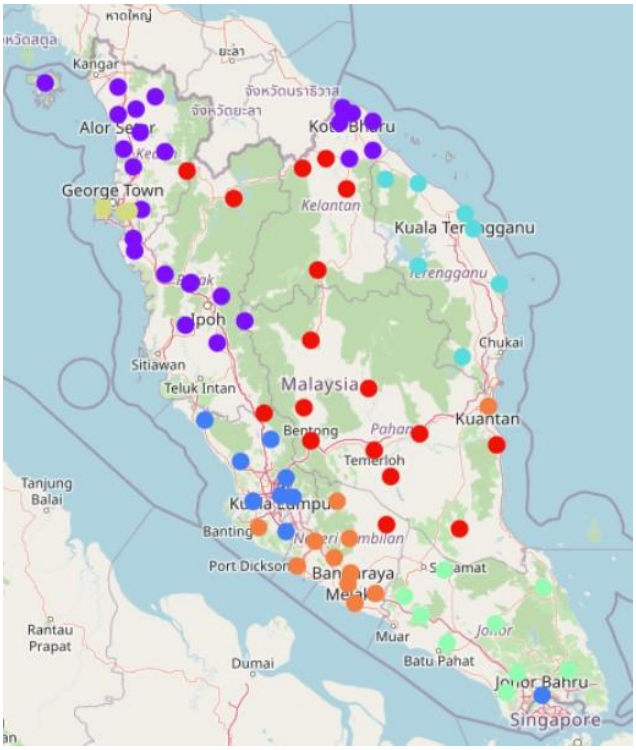
Cluster Labels	District	Latitude	Longitude	Population	Median Income per month	Mean Income per month	Number of Shopping Malls	Number of Gym/fitness centre	1st most age group	2nd most age group	3rd most age group
6	Alor Gajah	2.383668	102.210469	182666	5588	6849	50	50	3	1	2
6	Central Melaka	2.224511	102.261466	503127	5877	7142	50	50	3	1	2
6	Jasin	2.305810	102.422339	135317	5588	6296	50	50	3	1	2
6	Jelebu	3.036626	102.117921	39200	4579	5887	44	45	3	1	2
6	Kuala Pilah	2.743076	102.215488	66092	4579	5887	50	48	3	1	2
6	Port Dickson	2.522894	101.794513	115361	4579	5887	50	48	3	1	2
6	Rembau	2.588869	102.094340	43011	4579	5887	50	48	3	1	2
6	Seremban	2.718788	101.940451	555935	4579	6621	50	49	3	1	2
6	Tampin	2.472154	102.232350	84889	4579	5887	50	49	3	1	2
6	Kuantan	3.798564	103.321990	461906	3979	5012	50	47	3	2	1
6	Kuala Langat	2.836562	101.496182	224648	5293	6615	50	50	4	3	1

Observation:

The districts in Cluster 6 generally have **high level population** with more than 100000 residents per district, **medium level median and mean income** per month, **high total number of shopping malls, high total number of existing gyms,** the majority of residents are within the **age group of 0 to 29 years old.**

Map of Clusters

The clusters were mapped onto the map of Peninsular Malaysia. The output is as shown below:



Discussion

The districts in Cluster 0, Cluster 1, Cluster 2, Cluster 5 and Cluster 6 are unsuitable or less ideal for the company A Fitness to establish new gym branches in. The districts in these clusters already have a high number of gyms available.

The districts in Cluster 3 and Cluster 4 are more suitable and ideal for the company to open new gym branches in when compared to the other clusters. These districts matched the criteria set by A Fitness better then the districts in the other clusters.

Overall districts in Cluster 3 and Cluster 4 have about medium level population with more than 100000 residents per district, medium level median and mean income per month, medium level total number of shopping malls, medium level total number of existing gyms, the majority of residents are within the age group of 10 to 39 years old.

In addition, the results could have been more accurate if the actual mean and median income values for each district were included in the k-means clustering analysis instead of the state mean and median values.

There is a weak correlation between the population, mean and median income per month to the number of gyms available in a district. There is a high correlation between the number of gyms to the number of shopping malls.

Conclusion

The most ideal districts for the company A Fitness to open gym branches in are Besut, Dungun, Hulu Terengganu, Kemaman, Kuala Terengganu, Marang, Setiu, Batu Pahat, Kluang, Kota Tinggi, Kulai, Mersing, Muar, Pontian, Segamat dan Tangkak.

References

[1] Who.int. 2020. *Malaysia And WHO Call For More Investment In Primary Health Care The 21St Century*. [online]

Available at: <https://www.who.int/malaysia/news/detail/08-04-2019-malaysia-and-who-call-for-more-investment-in-primary-health-care-the-21st-century#:~:text=Malaysia%20has%20the%20highest%20rate,being%20either%20obese%20or%20overweight.> [Accessed 6 December 2020].

[2] Maps, M., 2020. *Malaysia Beaches And Islands (Peninsular Malaysia) – Google My Maps*. [online] Google My Maps.

Available at: <https://www.google.com/maps/d/viewer?mid=1Tp3JDwPo_MZ-JH-_diEytGMOKq4&ie=UTF8&source=embed&oe=UTF8&msa=0&ll=4.289239239796553%2C102.854948140625&z=7> [Accessed 6 December 2020].

[3] Dosm.gov.my. 2020. *Department Of Statistics Malaysia Official Portal*. [online]

Available at: <https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=155&bul_id=OVByWjg5YkQ3MWFZRTN5bDJiaEVhZz09&menu_id=L0pheU43NWJwRWVSZklWdzQ4TlhUUT09#:~:text=Malaysia's%20population%20in%202020%20is,to%203.0%20million%20(2020).> [Accessed 6 December 2020].

[4] En.wikipedia.org. 2020. *Peninsular Malaysia*. [online]

Available at: <https://en.wikipedia.org/wiki/Peninsular_Malaysia#:~:text=Peninsular%20Malaysia%20accounts%20for%20the,92%25%20of%20total%20population).> [Accessed 6 December 2020].

[5] Ho, F., 2020. *How Much Do Malaysians Need To Spend To Get Fit In 2016?*. [online] iMoney Malaysia. Available at: <https://www.imoney.my/articles/how-much-do-malaysians-need-to-spend-to-get-fit-in-2016> [Accessed 6 December 2020].

[6] Dosm.gov.my. 2020. *Department Of Statistics Malaysia Official Portal*. [online]

Available at: <https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=120&bul_id=TU00TmRhQ1N5TUxHVWN0T2VjbXJYZz09&menu_id=amVoWU54UTl0a21NWmdhMjFMMWcyZz09> [Accessed 6 December 2020].

[7] En.wikipedia.org. 2020. *List Of Districts In Malaysia*. [online]

Available at: <https://en.wikipedia.org/wiki/List_of_districts_in_Malaysia> [Accessed 6 December 2020].

[8] *Malaysia: Administrative Division (States And Districts) - Population Statistics, Charts And Map*. [online] Citypopulation.de.

Available at: <https://www.citypopulation.de/en/malaysia/admin/> [Accessed 12 December 2020].

[9] Pqi.stats.gov.my. 2020. *Population Quick Info*. [online]

Available at: <http://pqi.stats.gov.my/searchBI.php?tahun=2020&kodData=3&kodJadual=1&kodCiri=&kodNegeri=> [Accessed 13 December 2020].

[10] Dosm.gov.my. 2020. *Department Of Statistics Malaysia Official Portal*. [online]

Available at:
<https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=120&bul_id=RUZ5REwveU1ra1hGL21JWVlPRmU2Zz09&menu_id=amVoWU54UTl0a21NWmdhMjFMMWcyZz09> [Accessed 13 December 2020].