

# **Facial Treatment Shop Analysis in Singapore**

MF Chan

2 December 2019

## **1. Introduction**

The purpose of this case study is to determine suitable location for business setup that provide facial treatment service in Singapore. Nowadays the market is very competitive with high rental fees. Starting a business can be challenging and if they are not careful enough, it might cost them a lot of time and money. Many businesses had failed because unable to attract the customers to their business location. This is often more critical for service businesses that required customers to reach their business location. Therefore, proper site selection is important to bring success for a business especially when they are still new to the local communities.

Sometimes it is quite hard for the business owner to decide the location for them to start their business as there are many things to consider. By doing location analysis before starting any business, it will greatly reduce the risk that the business owner might face as they have a better understanding for the environment of the location they choose. They will be able to discover the best location to setup their business and explore where to target an ad campaign to potential customers as these are keys for a successful business.

## **2. Data**

### **2.1 Data Collection**

By using Foursquare API, we are able to collect existing facial treatment shop location in Singapore. As for our case study, we only covered 5km within the Singapore downtown area. Besides that, we also obtained demographic data from Department of Statistic Singapore website that contain some basic information such as age group, gender, population etc. based on different planning areas in Singapore.

### **2.2 Data Processing**

As some of the data collected from the API are quite redundant in our case, we exclude the unrelated features from the dataset and keep the ones that are only useful for our data analysis. The main information that we really need is the coordinate of the shop location so we are able to plot into the map.

As for the demographic data obtained from the website, we have data separated by age group in the range of 5 from 0 to 100. To make the data analysis easier, we do some feature engineering to make it smaller group with labels shown below according to different age groups. There are also some redundancies in this dataset such as subzone of the areas and the time period which are not include in our study case. After discarding them, the total population based on different features are calculated.

The following is the age group for different label:

- i. *child* (age from 0 to 9)
- ii. *teenager* (age from 10 to 19)
- iii. *young adult* (age from 20 to 34)
- iv. *adult* (age from 35 to 54)
- v. *seniors* (age from 55 and above)

### 3. Methodology

By using foursquare API, the locations of facial treatment shops which are located within 5000 meters from downtown Singapore are obtained. The coordinates of the shops are then used to create a distribution map of the facial treatment shops as shown below.

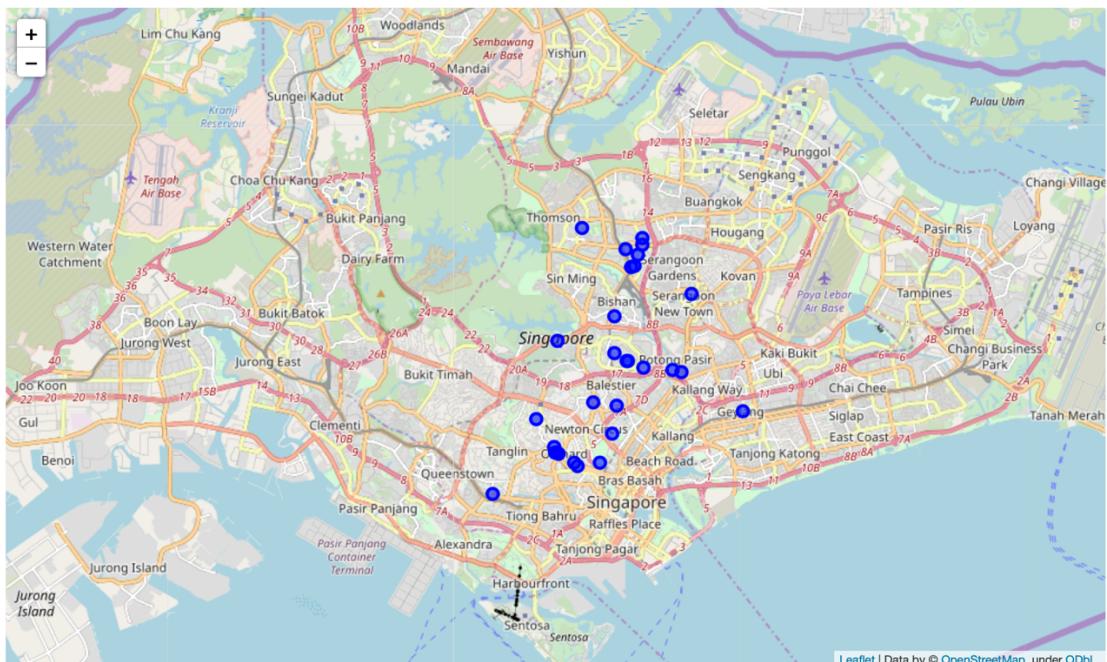


Figure I

From the map above, we can pick out which areas have more shops than others by counting the number of dots at the local areas. Another method is to group all the

shops that located in the same area and choose the area which have higher number of shops. Below is the map display after applying clustering technique.

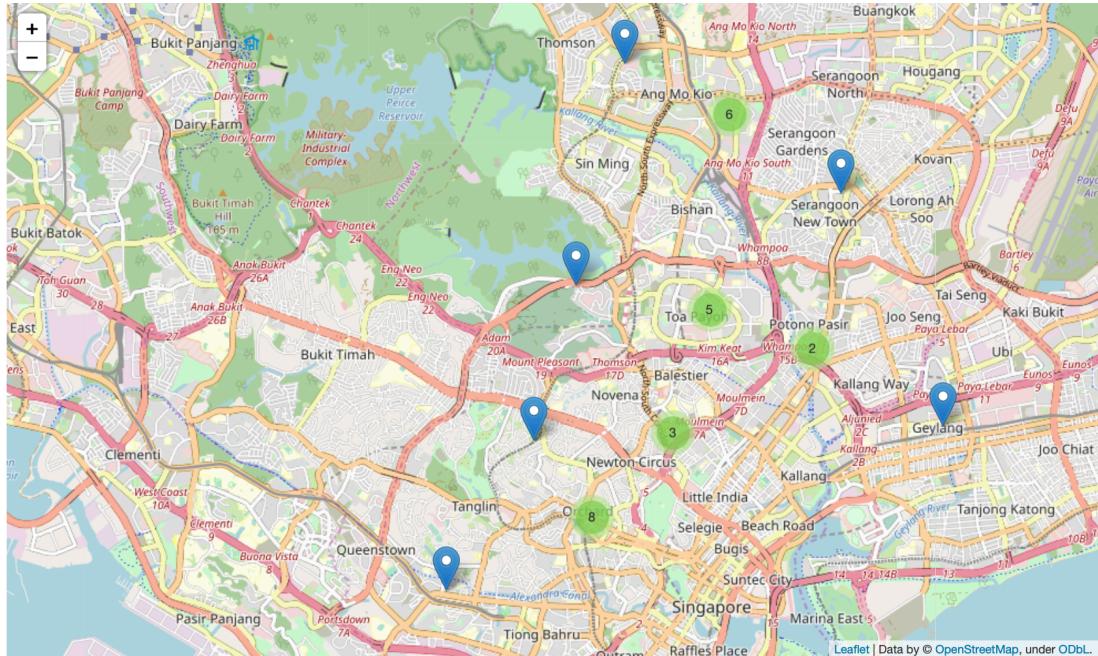


Figure 2

From the map result shown, we can identify three major locations that facial treatment business owners pick which have more than 5 number of shops in the area. As for my case study, I will avoid selecting the Orchard area although it has the highest shop number because this area are well known for shopping paradise.

Before I am able to start, I need to obtain existing demographic data which is available at Department of Statistic Singapore webpage. Then I do some analysis on the demographic data for the chosen areas to find the features that might cause higher demand. Below is the part result extract from the original dataset after excluding some unwanted information.

PA	SZ	AG	Sex	TOD	Pop	Time	AC
785536	Ang Mo Kio	Ang Mo Kio Town Centre	0_to_4	Males	HDB 1- and 2-Room Flats	0	2019 Child
785537	Ang Mo Kio	Ang Mo Kio Town Centre	0_to_4	Males	HDB 3-Room Flats	10	2019 Child
785538	Ang Mo Kio	Ang Mo Kio Town Centre	0_to_4	Males	HDB 4-Room Flats	10	2019 Child
785539	Ang Mo Kio	Ang Mo Kio Town Centre	0_to_4	Males	HDB 5-Room and Executive Flats	20	2019 Child
785540	Ang Mo Kio	Ang Mo Kio Town Centre	0_to_4	Males	HUDC Flats (excluding those privatised)	0	2019 Child

Table 1

Besides that, I also tried to collect more information about the areas that we choose for study. I use foursquare API to find out what are the shops nearby for that particular area. These will help us to have better understanding and determine which location are good for new business setup. Below is the part result obtained from foursquare API showing shop names and types for one of my study areas.

	<b>name</b>	<b>categories</b>	<b>location.lat</b>	<b>location.lng</b>
<b>0</b>	Passenger Pick-Up / Alighting Point	General Travel	1.369700	103.849420
<b>1</b>	Ang Mo Kio MRT Station (NS16)	Train Station	1.369557	103.849685
<b>2</b>	FairPrice Xtra	Supermarket	1.369279	103.848886
<b>3</b>	Cheers	Convenience Store	1.369263	103.849777
<b>4</b>	Bus Stop 54261 (Ang Mo Kio Station)	Bus Stop	1.369306	103.850736

Table 2

I also tried to find areas which are outside from my scope of study but within Singapore that might have similar demographic from the dataset obtained earlier. I decide to use k-means clustering technique to group them. Before that, we need to find the k-value so that we can know the best number of clusters for the dataset. By using elbow method, we are able to determine the optimal value for k-means clustering. From the chart below, we can see shape like an elbow.

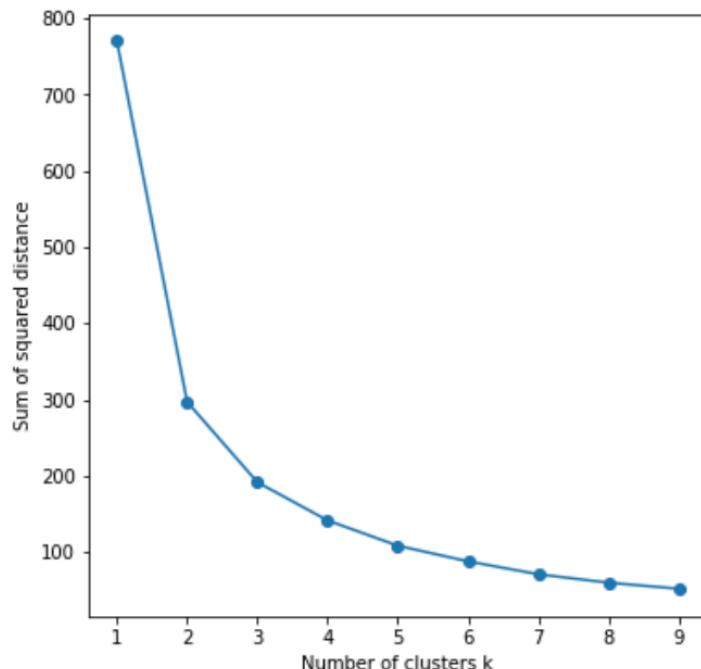


Figure 3

## 4. Results

As the dataset obtained from the webpage had already been processed, we can directly use the data without needing to spend too much time for data cleaning process. I just need to do some feature engineering to the dataset so that we can extract some insights from it. I had tried to get the total population based on gender and different age groups. Below is the results for my study area.

Age category	AMK		TP	
	Male	Female	Male	Female
Child	5880	5770	4780	4520
Teenager	7750	7560	5580	5340
Young Adult	14810	15240	10760	11370
Adult	16490	18380	12150	13670
Seniors	27720	32260	19680	23710
<b>Total</b>	<b>164430</b>		<b>121060</b>	

Table 3

The original dataset also contains type of dwelling for populations in Singapore. To get a better visual information on this, I tried to plot it with bar chart. Below is the chart that shows side by side comparison that indicates number of populations with different type of dwelling for my study areas.

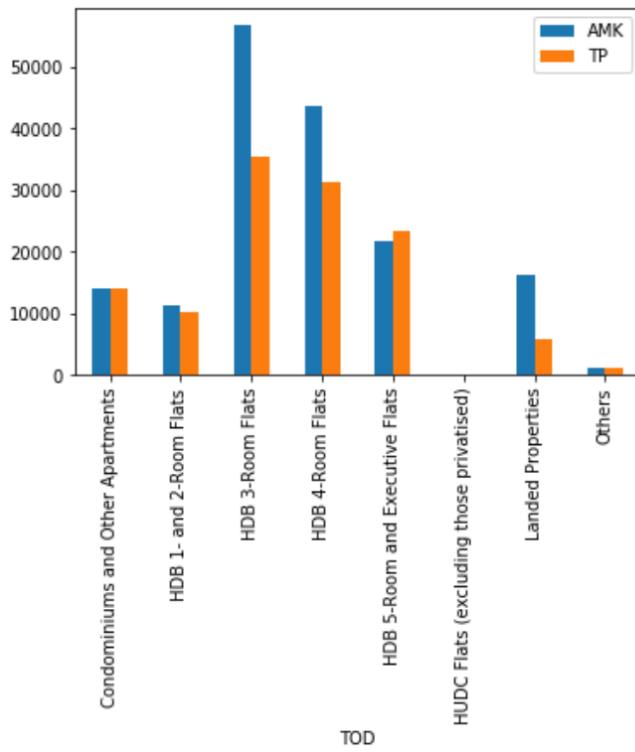


Figure 4

In my case, it is quite hard for me to figure out what is the best number of clusters to be used when I want to group the local areas for having similar demographic. The curve does not have obvious point show sign of flattening. Therefore I have pick the midpoint of the curve which is 5 number of clusters. The areas that fall within the same group with my study areas might considered to have similar demographic.

<b>Group</b>	<b>Area Name</b>
1	Bedok, Hougang
2	Bishan, Bukit Batok, Bukit Panjang, Bukit Timah, Clementi, Jurong East, Marine Parade, Novena, Pasir Ris, Sembawang, Serangoon
3	Chao Chu Kang, Jurong West, Punggol, Sengkang, Tampines, Woodlands, Yishun
4	Boon Lay, Central Water Catchment, Changi, Changi Bay, Downtown Core, Lim Chu Kang, Mandai, Marina East, Marina South, Museum, Newton, North-Eastern Islands, Orchard, Outram, Paya Lebar, Pioneer, River Valley, Rochor, Seletar, Simpang, Singapore River, Southern Islands, Straits View, Sungei Kadut, Tanglin, Tengah, Tuas, Western Islands, Western Water Catchment
5	Ang Mo Kio, Bukit Merah, Geylang, Kallang, Queenstown, Tao Payoh

*Table 4*

## 5. Discussion

From table 3 above, we can see that the largest group of population in both areas are senior citizen. The second largest group would be the adult group with age between 35 and 54 years old. We are also able to observed that there are more females for both of my study locations. As the facial treatment customers are more likely to be female adults, this result can help to explain why there are more shops here.

From figure 4 above, we notice that the numbers of 3-Room and 4 Room Flats for Ang Mo Kio are much higher compare to Tao Payoh. There are also more landed properties at Ang Mo Kio compare to Toa Payoh. As the price for landed properties in Singapore are much higher than the HDB flats, the average income for people who stay at Ang Mo Kio would be much higher than Tao Payoh. And for the number of shops for both areas are about the same, we might suggest that Ang Mo Kio still have growing space compare to Toa Payoh.

Besides that, Ang Mo Kio also have higher population compare to Tao Payoh. That could mean the former location is easier to build customer base for the business. I also tried to use foursquare API to get the names and types of shops nearby for my study areas. From there, we can discover what are the most common shop type for that particular area.

For those who wish to choose other places beside our study areas or having plan to expand their business to other areas, we can pick the areas having similar demographics. Based on table 4 above, both of our study areas fall under same category which is group number 5.

## **6. Conclusion**

As a result from my study, for those who are new in this industry but wish to start their business here in Singapore, they could consider to setup their business at Ang Mo Kio. But if selecting this area happens to be hard for them or any other reasons that preventing them to do so, they might consider to pick areas that fall under group 5 in table 4 above.