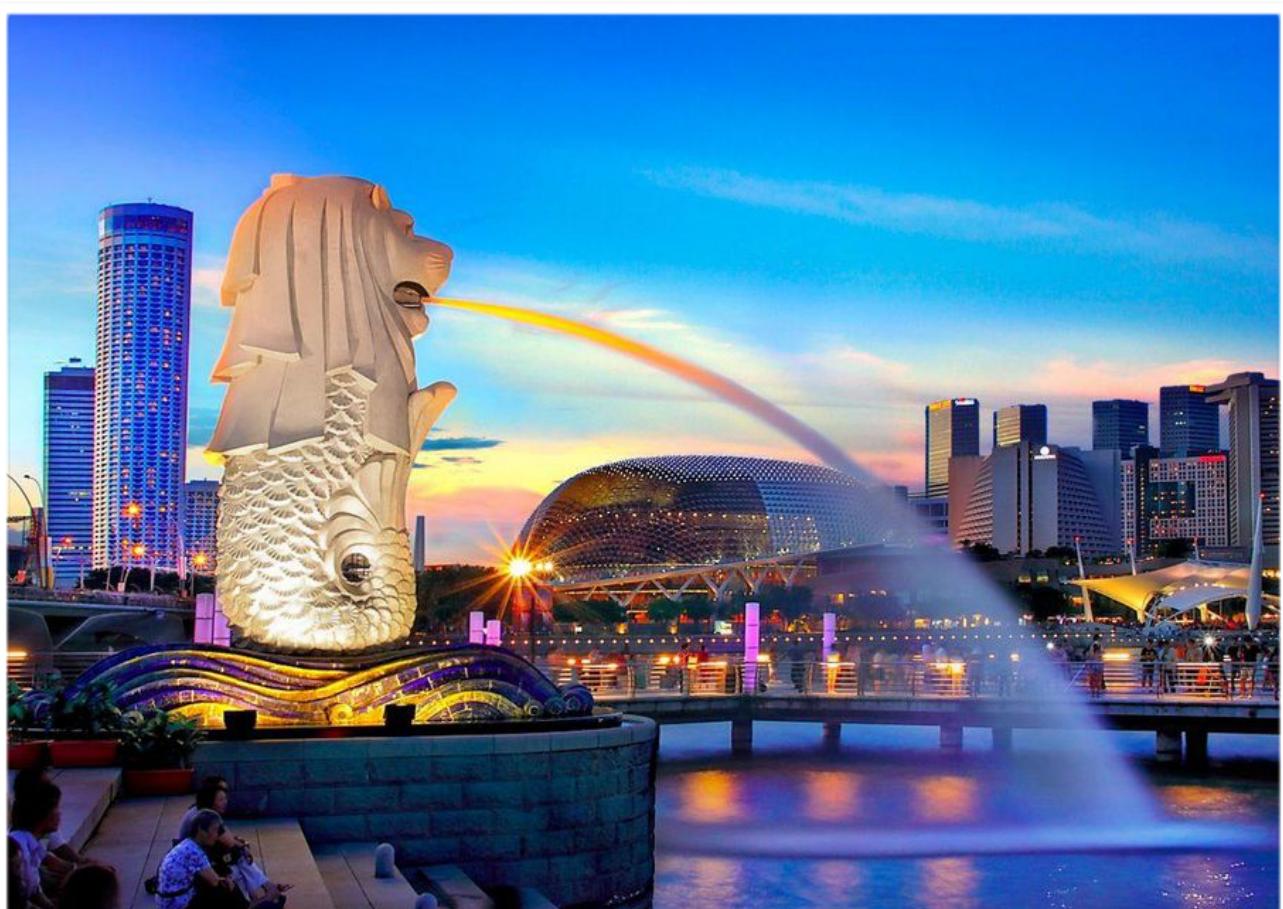


Capstone Project

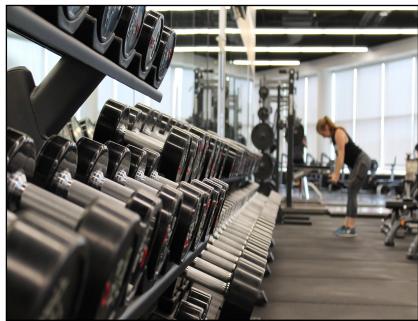
The Battle of Neighbourhoods

24 January 2020



Introduction

Being a Certified Personal Trainer, I dreamt of owning my own Gym in future. One major success of running a Gym is selecting its location. You can have a world-class Gym but if it is in an inaccessible area, the Gym will not be visited and it will hardly take off. Thus, I would like to take this opportunity to find the most suitable location to set up a Gym. So this report is for anyone who is interested in setting up a fitness facility around Singapore.



Business Problem

To help prospective Gym owners select a location which is accessible to the public and lack fitness facilities in the present day.

Data

Since Singapore is well-connected by the Mass Rapid Transit (MRT) System, one key data is a pre-consolidated locations (Longitude & Latitude) of all MRT Stations in Singapore. This data was consolidated and made available on GitHub, link provided below. An enquiry to Foursquare will locate the nearby fitness facilities and identify which MRT Station lack fitness facilities and will be a suitable Gym location. As the Venue Category in Foursquare might refer to fitness facilities as "Climbing Gym", "Yoga Studio", etc., it will be necessary to consider them as well before analyse and concluding on the location.

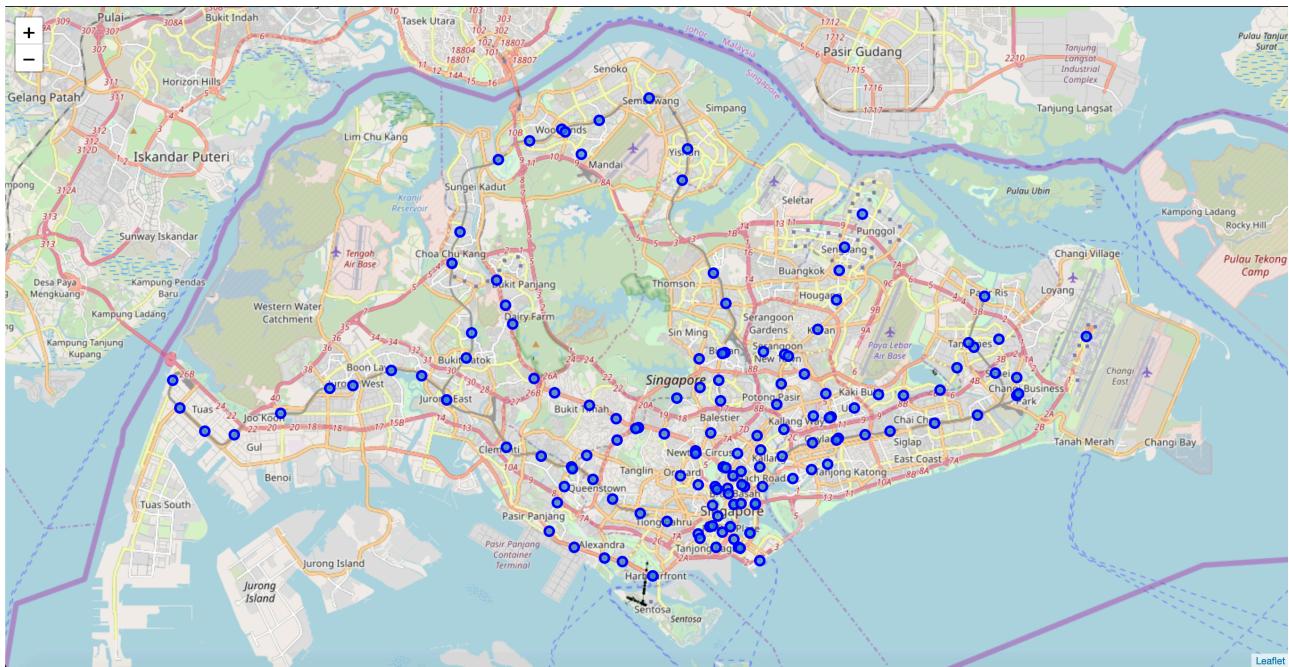
Methodology

Step 1: Areas that are accessible by MRT are considered accessible, thus, the locations of all MRT Stations are imported and their headers are renamed. As the consolidated list also

contained the Light Rail Transit (LRT) stations, a new dataframe (sg_mrt) was defined to only values with “MRT”.

ID	Neighborhood	Number	X	Y	Latitude	Longitude	Line	
0	12	ADMIRALTY MRT STATION	NS10	24402.1063	46918.1131	1.440585	103.800998	RED
1	16	ALJUNIED MRT STATION	EW9	33518.6049	33190.0020	1.316433	103.882893	GREEN
2	33	ANG MO KIO MRT STATION	NS16	29807.2655	39105.7720	1.369933	103.849553	RED
3	153	BARTLEY MRT STATION	CC12	33168.3039	36108.7003	1.342828	103.879746	YELLOW
4	115	BAYFRONT MRT STATION	DT16	30867.0093	29368.6250	1.281874	103.859073	BLUE

Step 2: The Map is plotted based on the Latitude and Longitude above to ensure all MRT stations are represented in the data.



Step 3: Using the Latitude and Longitude of the MRT Stations, data of the facilities nearby are retrieved from Foursquare.

Neighborhood	Latitude	Longitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
ADMIRALTY MRT STATION	1.440585	103.800998	Kampung Admiralty Hawker Centre	1.439939	103.800774	Food Court
ADMIRALTY MRT STATION	1.440585	103.800998	Superpets Warehouse / HQ	1.436848	103.806059	Pet Store
ADMIRALTY MRT STATION	1.440585	103.800998	McDonald's	1.445931	103.798101	Fast Food Restaurant
ADMIRALTY MRT STATION	1.440585	103.800998	Asia Ghani	1.437431	103.795388	Malay Restaurant
ADMIRALTY MRT STATION	1.440585	103.800998	Starbucks	1.439761	103.800659	Coffee Shop

Step 4: There are a few unique categories that are related to fitness, such as "Climbing Gym", "Yoga Studio", etc.. These are renamed to "Gym" to consider them as part of these analysis.

```
In [36]: venues_df.replace('Gym / Fitness Center', 'Gym', inplace = True)
venues_df.replace('Climbing Gym', 'Gym', inplace = True)
venues_df.replace('Rock Climbing Spot', 'Gym', inplace = True)
venues_df.replace('Yoga Studio', 'Gym', inplace = True)
venues_df.head()
```

Out[36]:

	Neighborhood	Latitude	Longitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	ADMIRALTY MRT STATION	1.440585	103.800998	Kampung Admiralty Hawker Centre	1.439939	103.800774	Food Court
1	ADMIRALTY MRT STATION	1.440585	103.800998	Superpets Warehouse / HQ	1.436848	103.806059	Pet Store
2	ADMIRALTY MRT STATION	1.440585	103.800998	McDonald's	1.445931	103.798101	Fast Food Restaurant
3	ADMIRALTY MRT STATION	1.440585	103.800998	Asia Ghani	1.437431	103.795388	Malay Restaurant
4	ADMIRALTY MRT STATION	1.440585	103.800998	Starbucks	1.439761	103.800659	Coffee Shop

Step 5: Based on the data, 103 fitness facilities are identified. Now we look at where these facilities are by breaking them into 4 clusters.

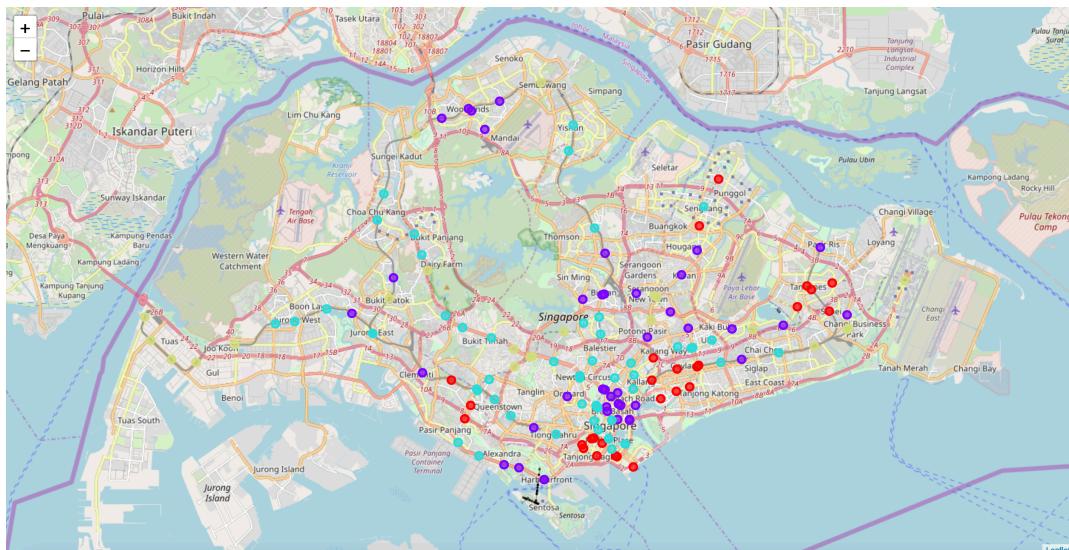
```
In [36]: venues_df.replace('Gym / Fitness Center', 'Gym', inplace = True)
venues_df.replace('Climbing Gym', 'Gym', inplace = True)
venues_df.replace('Rock Climbing Spot', 'Gym', inplace = True)
venues_df.replace('Yoga Studio', 'Gym', inplace = True)
venues_df.head()
```

Out[36]:

	Neighborhood	Latitude	Longitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	ADMIRALTY MRT STATION	1.440585	103.800998	Kampung Admiralty Hawker Centre	1.439939	103.800774	Food Court
1	ADMIRALTY MRT STATION	1.440585	103.800998	Superpets Warehouse / HQ	1.436848	103.806059	Pet Store
2	ADMIRALTY MRT STATION	1.440585	103.800998	McDonald's	1.445931	103.798101	Fast Food Restaurant
3	ADMIRALTY MRT STATION	1.440585	103.800998	Asia Ghani	1.437431	103.795388	Malay Restaurant
4	ADMIRALTY MRT STATION	1.440585	103.800998	Starbucks	1.439761	103.800659	Coffee Shop

Results

The map below shows which of the MRT stations in 4 clusters. It shows which MRT stations are crowded with fitness facilities and which stations lack such facilities.



Discussion

Based on the analysis, Cluster 0 (Red) is the most crowded with fitness facilities while Cluster 3 (Green) lacks fitness facilities. Cluster 2 (Blue) is sufficiently equipped, while Cluster 1 (Purple) is equipped. This shows that Central Singapore and East Singapore are crowded with fitness facilities already and I should avoid those areas. While the western part of Singapore might have opportunity for me to set up a Gym.

Conclusion

With the help of Foursquare, I am able to get a quick sense of where the opportunities for a Gym are. It seems like Western Singapore, such as Jurong East, Chinese Gardens MRT Stations are possible areas for a new Gym.