# Capstone - Finding suitable locations to New Business – Analysis

# Business Problem

- The gym should be located near a metro station
- Things to be considered:
  - Number of existing gyms in the area
  - Distance from each station to nearest gym

# Data

- Starting point: Syntagma Square

  Obtain its geospatial coordinates

- Using Foursquare API:
  - Find all metro stations in a radius of 15 km from Syntagm Square
  - Find existing gyms in a radius of 750 m from each station
  - Calculate minimum distance from every station to a gym
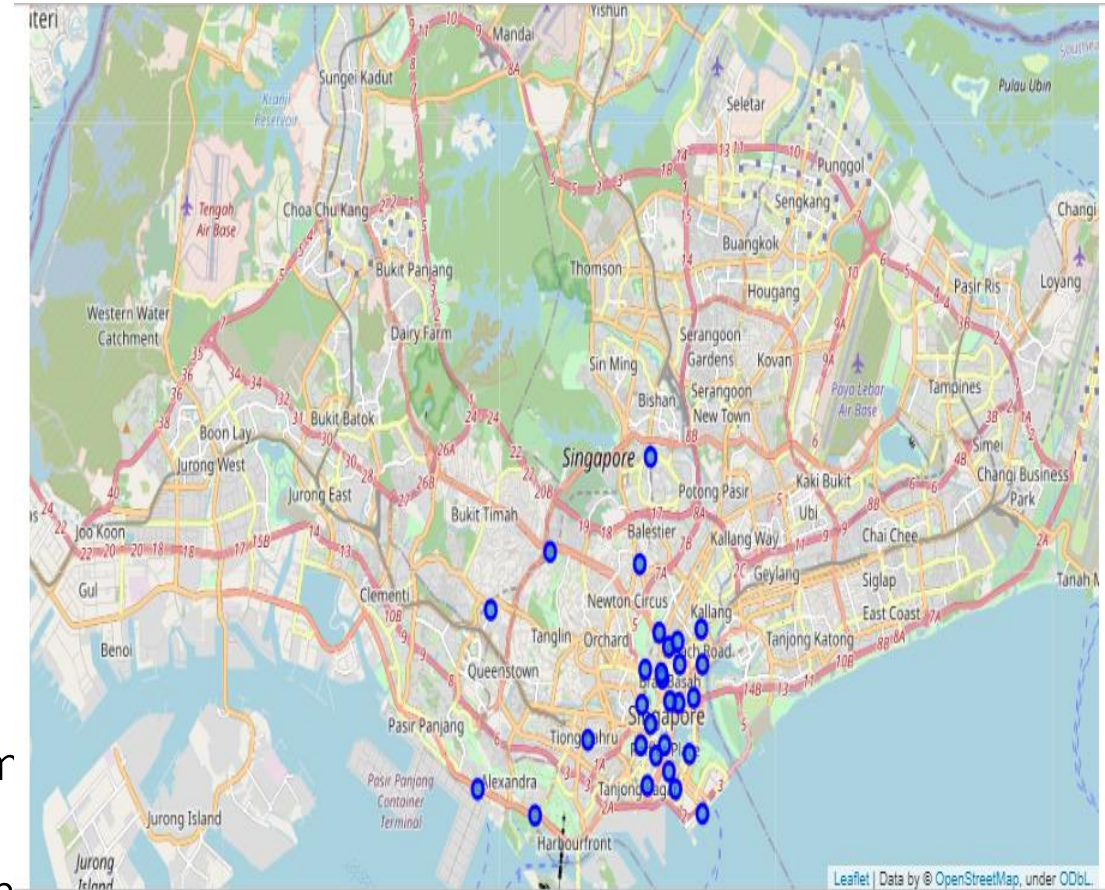
# Data Description, Transformation and Cleaning

- 30 initial metro stations found

  1 stations were train depots so removed from the data set

- Out of the remaining 29 metro stations 28 found with gyms in their vicinity

  No data for 1 stations (either non-existent data or no stations around these stations)

- Only gyms categories with description "Gym / Fitness Center" and "Gym" were considered.

  All other categories (Dance Studio', 'Yoga Studio', 'Martial Arts Dojo' etc.) were excluded

- Data were normalized for the final calculations

# Methodology

- Machine Learning (K-Means clustering algorithm) was used to create clusters of stations
- 3 numbers of clusters were calculated based on the algorithm recommendation (elbow method)


K-Means Elbow curve

# Results
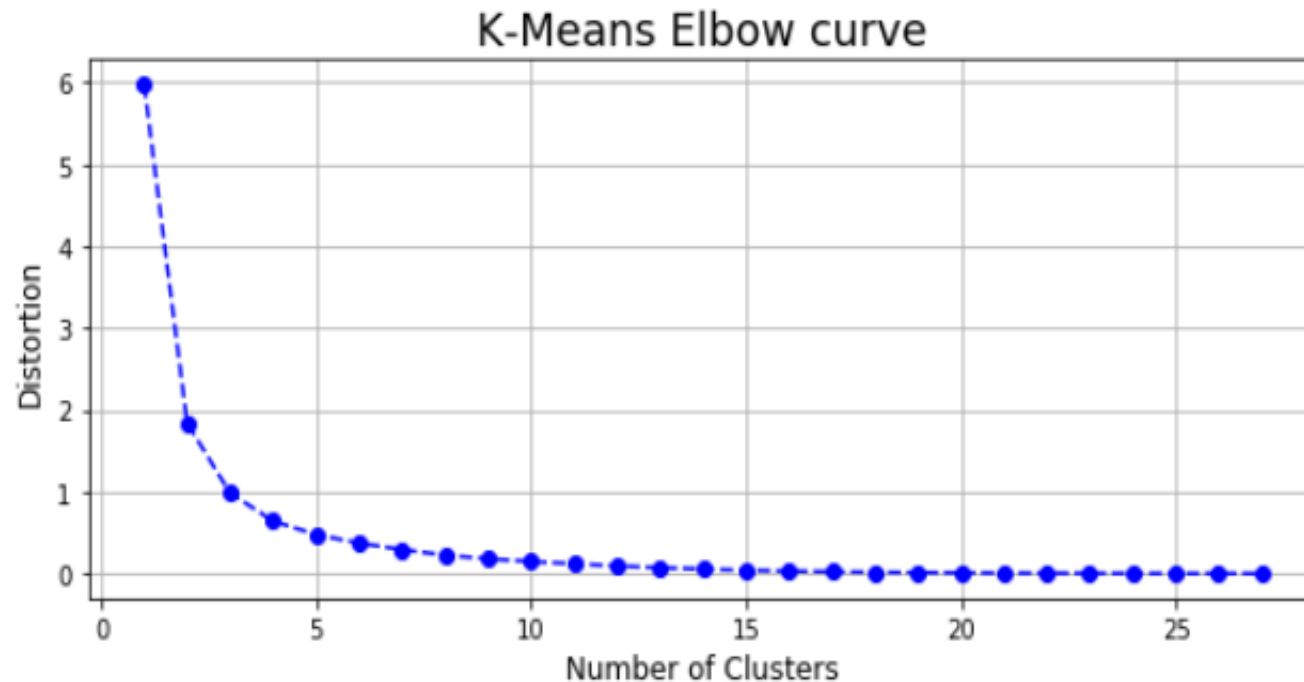


Clusters and Centroids
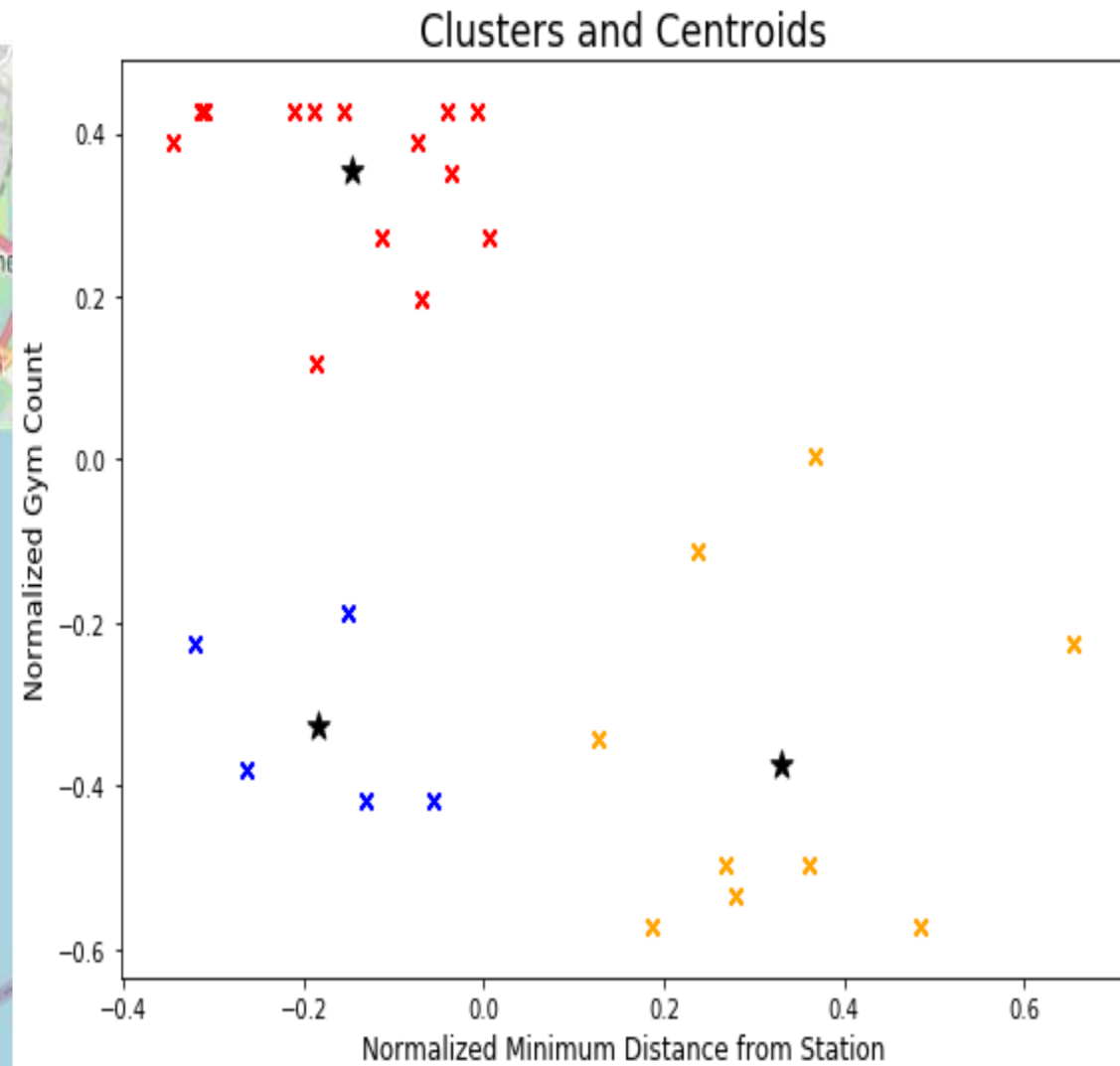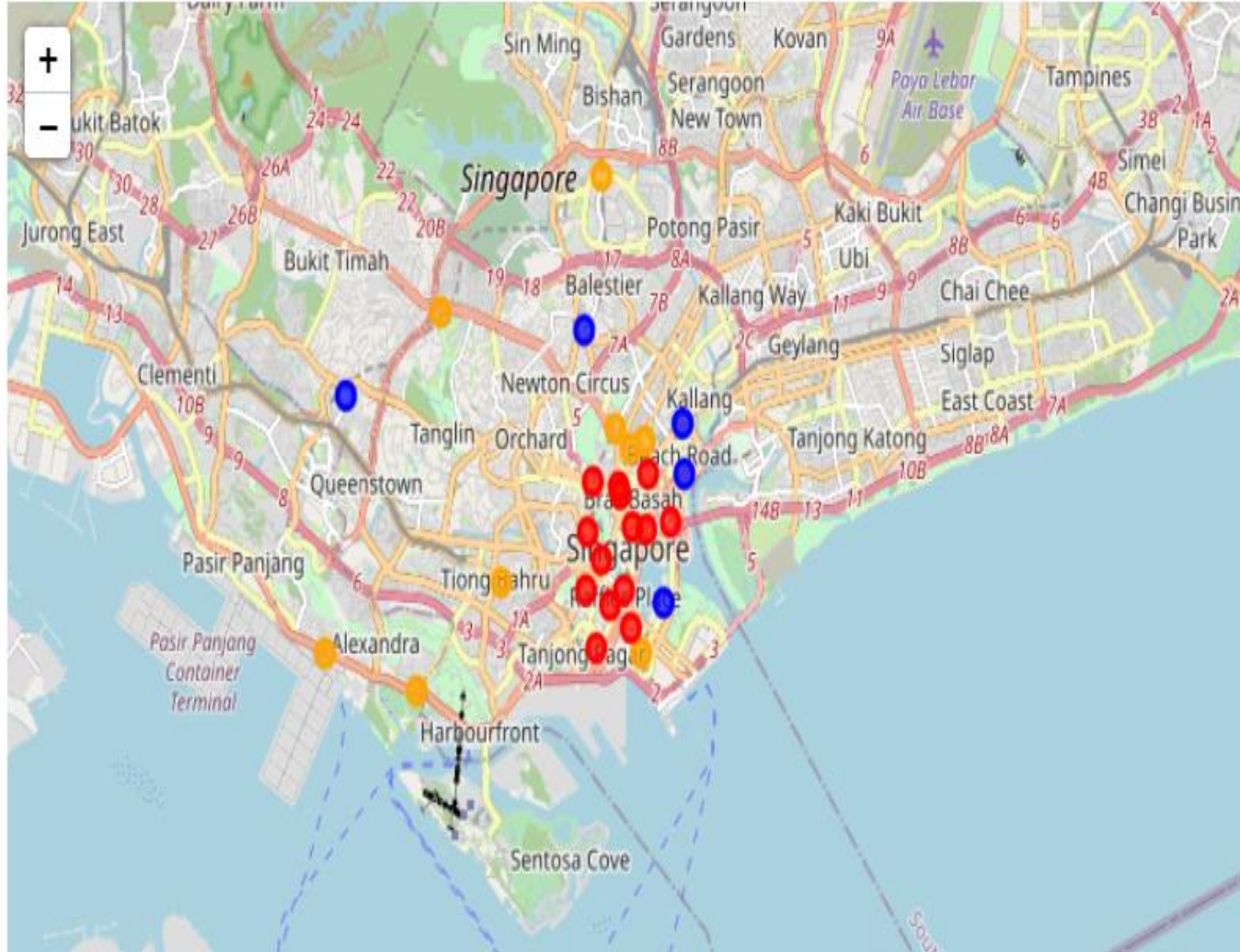
# Results – Cluster # 1

## Cluster # 1 – LOW potential (Red on the map)

- There are already many existing gyms in the area and the nearest gym is in most cases in a relatively short distance from the station. [9 metro stations]

| | station | Cluster Label | lat | lng | Min Distance from Station | Gym Count | Norm Min Distance from Station | Norm Gym Count |
|---|---|---|---|---|---|---|---|---|
| 3 | Telok Blangah MRT Station | 0 | 1.270729 | 103.809998 | 365 | 4 | 0.360202 | -0.495879 |
| 5 | Botanic Gardens MRT Interchange | 0 | 1.322324 | 103.814880 | 278 | 2 | 0.186895 | -0.572802 |
| 12 | Tiong Bahru MRT Station | 0 | 1.285758 | 103.826982 | 325 | 3 | 0.280521 | -0.534341 |
| 15 | Pasir Panjang MRT Station | 0 | 1.276075 | 103.791973 | 427 | 2 | 0.483708 | -0.572802 |
| 18 | Marina Bay MRT Interchange | 0 | 1.276144 | 103.854788 | 303 | 14 | 0.236696 | -0.111264 |
| 19 | Braddell MRT Station | 0 | 1.340742 | 103.847020 | 319 | 4 | 0.268569 | -0.495879 |
| 24 | Rochor MRT Station | 0 | 1.303877 | 103.852635 | 369 | 17 | 0.368170 | 0.004121 |
| 25 | Jalan Besar MRT Station | 0 | 1.305004 | 103.855339 | 513 | 11 | 0.655023 | -0.226648 |
| 27 | Little India MRT Interchange | 0 | 1.306457 | 103.849606 | 248 | 8 | 0.127134 | -0.342033 |

# Results – Cluster # 2

## Cluster # 2 – AVERAGE potential (Orange on the map)

- Although not a prohibitive metro station to open a gym in its vicinity, there is already a fair number of gyms in the area and the nearest one is not far from the metro station. [14 stations]

| | station | Cluster Label | lat | lng | Min Distance from Station | Gym Count | Norm Min Distance from Station | Norm Gym Count |
|---|---|---|---|---|---|---|---|---|
| 0 | Bugis MRT Interchange | 1 | 1.300476 | 103.856094 | 187 | 24 | 0.005620 | 0.273352 |
| 1 | Raffles Place MRT Interchange | 1 | 1.284516 | 103.851446 | 11 | 27 | -0.344977 | 0.388736 |
| 2 | Esplanade MRT Station | 1 | 1.292907 | 103.855946 | 106 | 28 | -0.155734 | 0.427198 |
| 6 | Telok Ayer MRT Station | 1 | 1.282401 | 103.848756 | 26 | 28 | -0.315097 | 0.427198 |
| 7 | Clarke Quay MRT Station | 1 | 1.288804 | 103.846968 | 29 | 28 | -0.309121 | 0.427198 |
| 11 | Promenade MRT Interchange | 1 | 1.293731 | 103.860465 | 91 | 20 | -0.185615 | 0.119505 |
| 13 | Fort Canning MRT Station | 1 | 1.292371 | 103.844272 | 164 | 28 | -0.040196 | 0.427198 |
| 14 | City Hall MRT Interchange | 1 | 1.293146 | 103.853053 | 90 | 28 | -0.187607 | 0.427198 |
| 16 | Dhoby Ghaut MRT Interchange | 1 | 1.299225 | 103.845343 | 149 | 22 | -0.070077 | 0.196429 |
| 17 | Bras Basah MRT Station | 1 | 1.297506 | 103.850506 | 147 | 27 | -0.074061 | 0.388736 |
| 20 | Bencoolen MRT Station | 1 | 1.298700 | 103.850176 | 127 | 24 | -0.113902 | 0.273352 |
| 22 | Chinatown MRT Interchange | 1 | 1.284482 | 103.843842 | 166 | 26 | -0.036212 | 0.350275 |
| 23 | Downtown MRT Station | 1 | 1.279544 | 103.852828 | 180 | 28 | -0.008324 | 0.427198 |
| 26 | Tanjong Pagar MRT Station | 1 | 1.276543 | 103.845943 | 79 | 28 | -0.209519 | 0.427198 |

# Results – Cluster # 3

## Cluster # 3 – LOW potential (Red on the map)

- There are not many already existing gyms in the area and the nearest gym is in most cases relatively not in a short distance to the metro station.[5 stations]

| | station | Cluster Label | lat | lng | Min Distance from Station | Gym Count | Norm Min Distance from Station | Norm Gym Count |
|---|---|---|---|---|---|---|---|---|
| 4 | Bayfront MRT Interchange | 2 | 1.283062 | 103.859167 | 119 | 6 | -0.129838 | -0.418956 |
| 8 | Nicoll Highway MRT Station | 2 | 1.300246 | 103.863451 | 156 | 6 | -0.056133 | -0.418956 |
| 9 | Novena MRT Station | 2 | 1.320086 | 103.843592 | 23 | 11 | -0.321073 | -0.226648 |
| 10 | Lavender MRT Station | 2 | 1.307418 | 103.862860 | 108 | 12 | -0.151750 | -0.188187 |
| 21 | Holland Village MRT Station | 2 | 1.311008 | 103.795959 | 52 | 7 | -0.263304 | -0.380495 |

# Discussion / Conclusions

- Possible areas that were not in the Foursquare database should also be examined so that it can be determined if it is just lack of data about these stations or indeed there are no gyms in the vicinity of the stations.

- A lot more factors can be considered when choosing an appropriate location. Some examples of extra factors can be:
  - Population density in the area
  - Number of businesses operating in the area (people may want to go to a gym close to work)
  - Average age and household income in the area
  - Property prices in the area

The above results can be a good starting point for a prospective businessman that is interested in opening a gym. Similar methodology can be used for other types of businesses probably with customized criteria.