



THE BATTLE OF NEIGHBORHOODS

Applied Data Science Capstone
Project

Yee Tung Lee

Background

- Increasing number of people in Hong Kong planning to relocate to other countries
- Factors affecting to decide the country to move to include economic environment, political stability, education system, cost of living, living environment, etc
- Living environment is utmost important as similar neighborhoods make it easier for people to adapt to changes during relocation process
- Singapore is one of the popular option for relocation due to similarities between Hong Kong and Singapore

Project Objectives

- Currently there are limited available resources in the comparison of neighborhoods between Singapore and Hong Kong
- Being new to the city, people who would like to relocate to Singapore find it difficult to decide the suitable neighborhood to stay in
- This project aim to provide recommendation in neighborhoods that is suitable for our target customer to help them in exploring and evaluating suitable neighborhoods. A list of points of interest and neighborhood map will also be prepared for them to explore the recommended neighborhoods
- Recommendation was done through analyzing the similarities between different neighborhoods in Singapore and Hong Kong, with the requirements and preferences from our target customer

Target audience

- Below is the profile of the target customers:
 - A young couple without any children
 - Work in banking industry in CBD area
 - Do not drive
 - Currently staying in Sha Tin in Hong Kong
 - Enjoy hanging out in cafes and shopping during rest days

Data Acquisition and Cleaning

- Neighborhood Data
 - List of neighborhoods scraped from below Wikipedia pages, columns not required removed
https://en.wikipedia.org/wiki/Planning_Areas_of_Singapore
https://en.wikipedia.org/wiki/Districts_of_Hong_Kong
 - List of Singapore MRT Railway stations from below government website, columns not required removed
<https://data.gov.sg/dataset/train-station-chinese-names>
 - Neighborhoods coordinates obtained using Nominatim geocoding API
 - Distance of neighborhoods from city centres calculated using haversine formula
 - Total 55 neighborhoods in Singapore and 19 neighborhoods in Hong Kong

	City	Neighborhood	Latitude	Longitude	Distance from city centre
0	Singapore	Ang Mo Kio	1.370080	103.849523	8.713848
1	Singapore	Bedok	1.323976	103.930216	9.607766
2	Singapore	Bishan	1.350986	103.848255	6.594669
3	Singapore	Boon Lay	1.338550	103.705812	16.844590
4	Singapore	Bukit Batok	1.349057	103.749591	12.849412
5	Singapore	Bukit Merah	1.270439	103.828318	3.371441
6	Singapore	Bukit Panjang	1.379149	103.761413	13.833367
7	Singapore	Bukit Timah	1.354690	103.776372	10.768272
8	Singapore	Central Water Catchment	1.375708	103.801743	10.768461
9	Singapore	Changi	1.351080	103.990064	16.903210

Snapshot of Singapore neighborhoods dataset

	City	Neighborhood	Latitude	Longitude	Distance from city centre
0	Hong Kong	Central and Western	22.284918	114.158304	0.331478
1	Hong Kong	Eastern Hong Kong	22.284918	114.158304	0.331478
2	Hong Kong	Southern Hong Kong	22.219263	114.225230	9.810707
3	Hong Kong	Wan Chai	22.279015	114.172483	1.518831
4	Hong Kong	Sham Shui Po	22.328190	114.160854	5.147140
5	Hong Kong	Kowloon City	22.330160	114.189937	6.280976
6	Hong Kong	Kwun Tong	22.312937	114.225610	7.752648
7	Hong Kong	Wong Tai Sin	22.341654	114.193859	7.587756
8	Hong Kong	Yau Tsim Mong	22.307404	114.165526	2.931598
9	Hong Kong	Islands Hong Kong	22.244541	114.205376	6.399822

Snapshot of Hong Kong neighborhoods dataset

Data Acquisition and Cleaning

- Venue Location Data
 - List of venues in each neighborhood retrieved using Foursquare API
 - Extracted relevant data including venue name, coordinates, categories and addresses

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Venue Address
0	Ang Mo Kio	1.37008	103.849523	Old Chang Kee	1.369094	103.848389	Snack Place	[AMK Hub (53 Ang Mo Kio Avenue 3), 569933, Sin...
1	Ang Mo Kio	1.37008	103.849523	FairPrice Xtra	1.369279	103.848886	Supermarket	[#B2-26 AMK Hub (53 Ang Mo Kio Ave 3), 569933,...]
2	Ang Mo Kio	1.37008	103.849523	MOS Burger	1.369170	103.847831	Burger Joint	[#01-33, AMK Hub (53 Ang Mo Kio Ave 3), 569933...]
3	Ang Mo Kio	1.37008	103.849523	Face Ban Mian 非板面 (Ang Mo Kio)	1.372031	103.847504	Noodle House	[722 Ang Mo Kio Ave 8 (Ang Mo Kio Central), 56...
4	Ang Mo Kio	1.37008	103.849523	PLAYe	1.369109	103.848225	Hobby Shop	[#02-30/31, AMK HUB (53 Ang Mo Kio Ave 3), 569...

Snapshot of Singapore neighborhood venue dataset

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Central and Western	22.284918	114.158304	Four Seasons Hotel Hong Kong (香港四季酒店)	22.286554	114.156929	Hotel
1	Central and Western	22.284918	114.158304	Apple ifc mall	22.284612	114.159070	Electronics Store
2	Central and Western	22.284918	114.158304	The Spa at Four Seasons	22.286279	114.157623	Spa
3	Central and Western	22.284918	114.158304	ifc mall (國際金融中心商場)	22.284613	114.158156	Shopping Mall
4	Central and Western	22.284918	114.158304	city'super	22.286319	114.157806	Supermarket

Snapshot of Hong Kong neighborhood venue dataset

Data Acquisition and Cleaning

- Venue Location Data
 - Combine venues by category using one hot encoding to obtain frequency of different categories and top 10 common venue categories for each neighborhood

	Neighborhood	Zoo Exhibit	ATM	Accessories Store	Airport	Airport Service	American Restaurant	Arcade	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant	Astrologer	Athletics & Sports
0	Ang Mo Kio	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.024390	0.0000	0.000000
1	Bedok	0.0	0.000000	0.000000	0.000000	0.000000	0.016949	0.000000	0.000000	0.000000	0.000000	0.050847	0.0000	0.000000
2	Bishan	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.046512	0.0000	0.000000
3	Boon Lay	0.0	0.000000	0.000000	0.000000	0.000000	0.014085	0.000000	0.000000	0.000000	0.000000	0.084507	0.0000	0.000000
4	Bukit Batok	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0000	0.000000
5	Bukit Merah	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0000	0.000000
6	Bukit Panjang	0.0	0.000000	0.000000	0.000000	0.000000	0.027027	0.000000	0.000000	0.000000	0.000000	0.054054	0.0000	0.000000
7	Bukit Timah	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0000	0.000000
8	Central and Western	0.0	0.000000	0.000000	0.000000	0.013514	0.000000	0.000000	0.013514	0.000000	0.000000	0.013514	0.0000	0.000000
9	Changi	0.0	0.000000	0.000000	0.428571	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0000	0.000000

Snapshot of dataframe with frequency of venue categories for each neighborhood

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0 Ang Mo Kio	Coffee Shop	Dessert Shop	Bubble Tea Shop	Japanese Restaurant	Snack Place	Supermarket	Sandwich Place	Food Court	Multiplex	Hobby Shop
1 Bedok	Japanese Restaurant	Coffee Shop	Asian Restaurant	Sandwich Place	Food Court	Chinese Restaurant	Sushi Restaurant	Noodle House	Fast Food Restaurant	Supermarket
2 Bishan	Food Court	Coffee Shop	Chinese Restaurant	Bubble Tea Shop	Café	Cosmetics Shop	Ice Cream Shop	Supermarket	Japanese Restaurant	Asian Restaurant
3 Boon Lay	Japanese Restaurant	Asian Restaurant	Fast Food Restaurant	Chinese Restaurant	Dessert Shop	Coffee Shop	Park	Indian Restaurant	Gym / Fitness Center	Café
4 Bukit Batok	Coffee Shop	Food Court	Chinese Restaurant	Fast Food Restaurant	Department Store	Frozen Yogurt Shop	Mobile Phone Shop	Malay Restaurant	Sandwich Place	Café
5 Bukit Merah	Resort	Beach	Wine Shop	Clothing Store	Cafeteria	Farm	Flea Market	Fast Food Restaurant	Filipino Restaurant	Fish & Chips Shop
6 Bukit Panjang	Coffee Shop	Sushi Restaurant	Fast Food Restaurant	Shopping Mall	Light Rail Station	Asian Restaurant	Bubble Tea Shop	Bookstore	Supermarket	Fried Chicken Joint

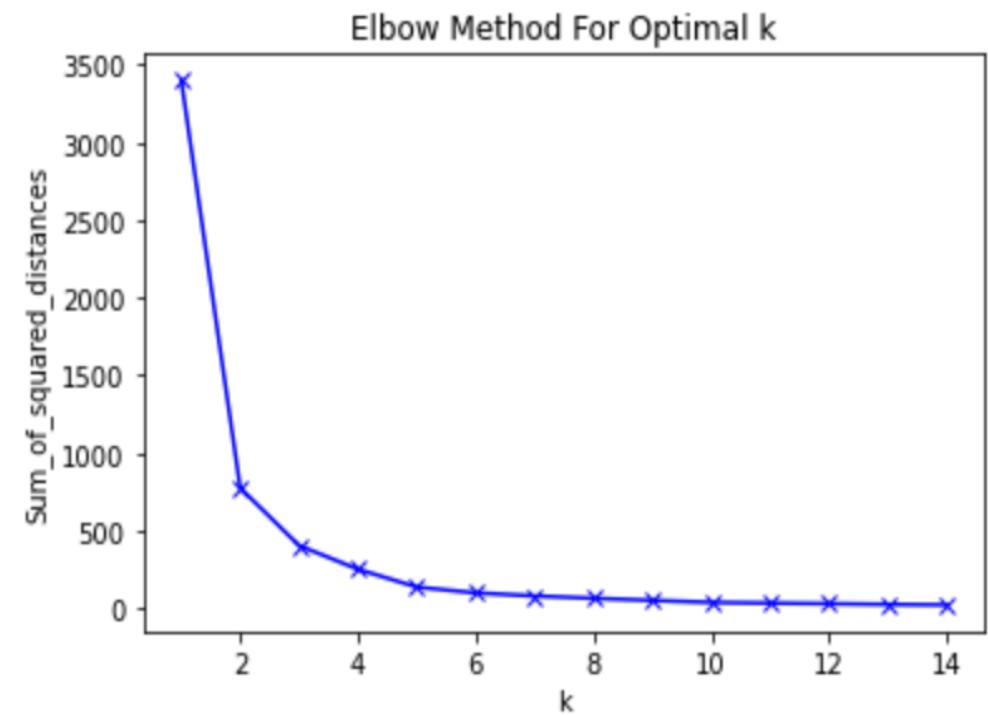
Top 10 common venue categories in each neighborhood

Methodology

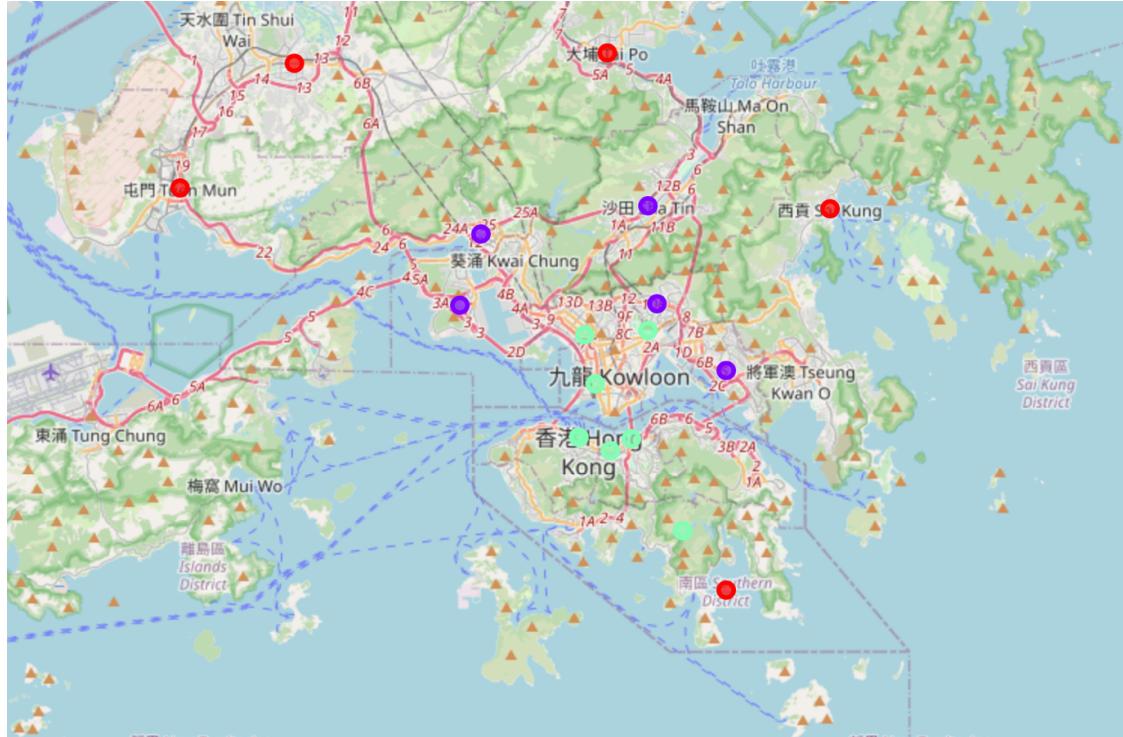
- Analyze similarities and dissimilarities of neighborhoods by patterns of venues available in each neighborhood and their distance from city centres
- K-means clustering algorithm to group neighborhoods into clusters
- Filter clustering results according to target customers' requirements and preferences to identify suitable neighborhoods as recommendations

Neighborhood Clustering using K-means

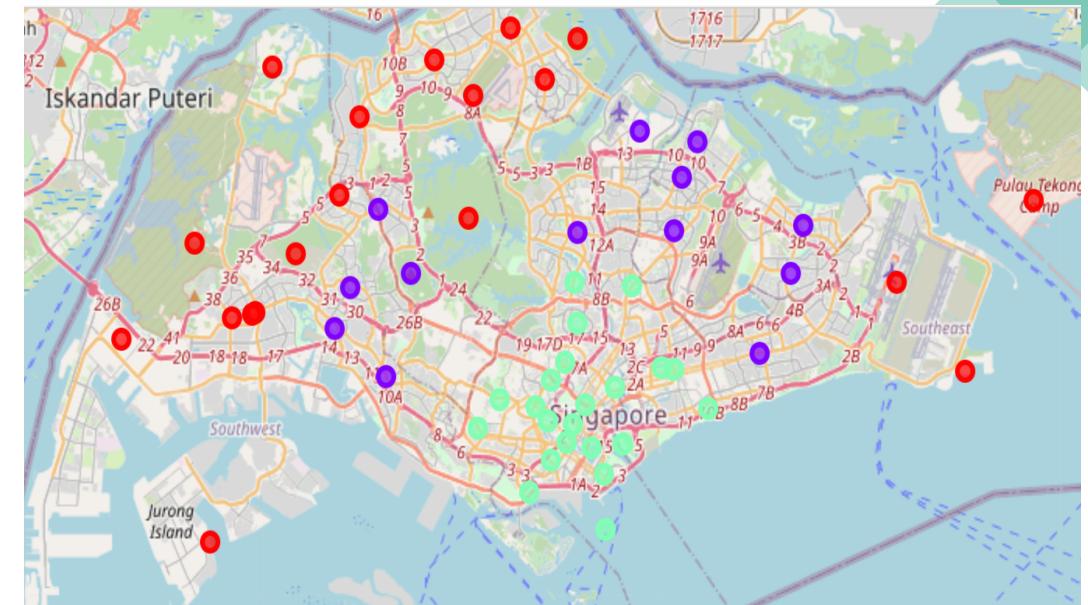
- According to results from Elbow Method, the optimal K is 3
- K-means used to segment the neighborhoods into 3 clusters:
 - Cluster 0 - Countryside area**, close to nature (e.g. mountains, parks, seaside), further away from city centre
 - Cluster 1 - Suburban area**, with plenty of eateries, shops and markets
 - Cluster 2 - Urban area**, with tourist attractions and shopping malls



Clustering Results



Map of Hong Kong with clusters



Map of Singapore with clusters

Recommendation Selection Criteria

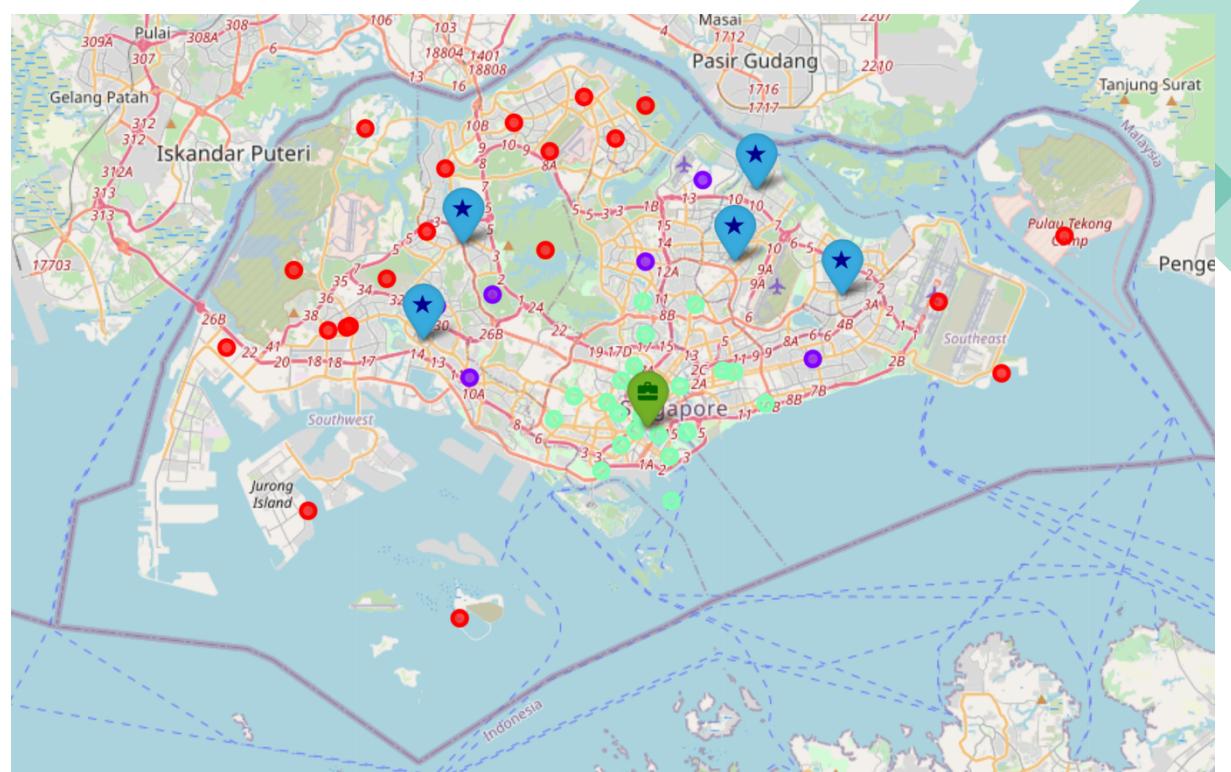
To recommend suitable Singapore neighborhoods for our target customers, below are the criteria:

- Same Cluster
 - Similar neighborhood as their current neighborhood in Hong Kong so that they can adapt to changes during relocation easily
 - Current neighborhood (Shatin) is under Suburban Cluster
- Accessibility to MRT Railway network
 - As our target customers do not drive, MRT network provides easy access around Singapore and allow them to commute to work or get around easily
- Lifestyle preferences (availability of cafes and shopping malls)

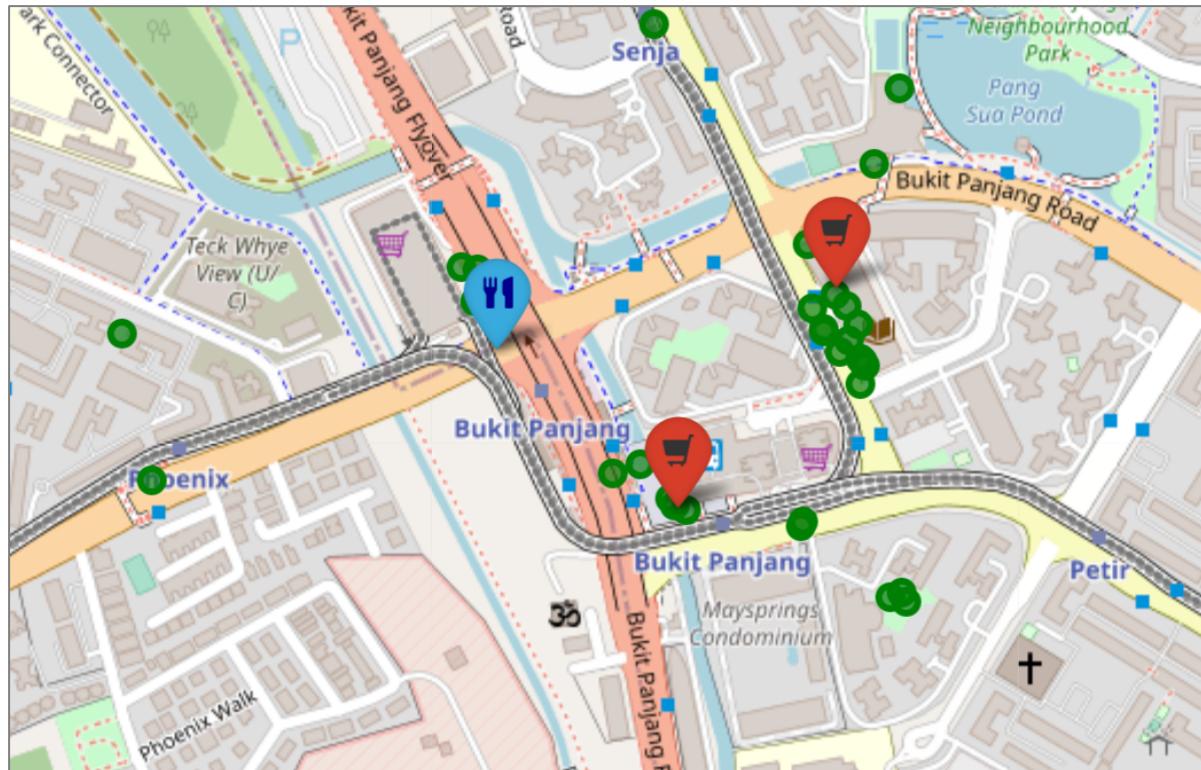
Recommendations

There are 5 recommended neighborhoods:

- Bukit Panjang, Hougang, Jurong East, Punggol, Tampines
- Located in suburban area, away from the busy urban environment while easily accessible via public transport (e.g. MRT railway)
- There are shopping malls and cafes available in the neighborhood, other than many more other dining options



Recommendation 1 – Bukit Panjang



Red marker: Shopping Mall

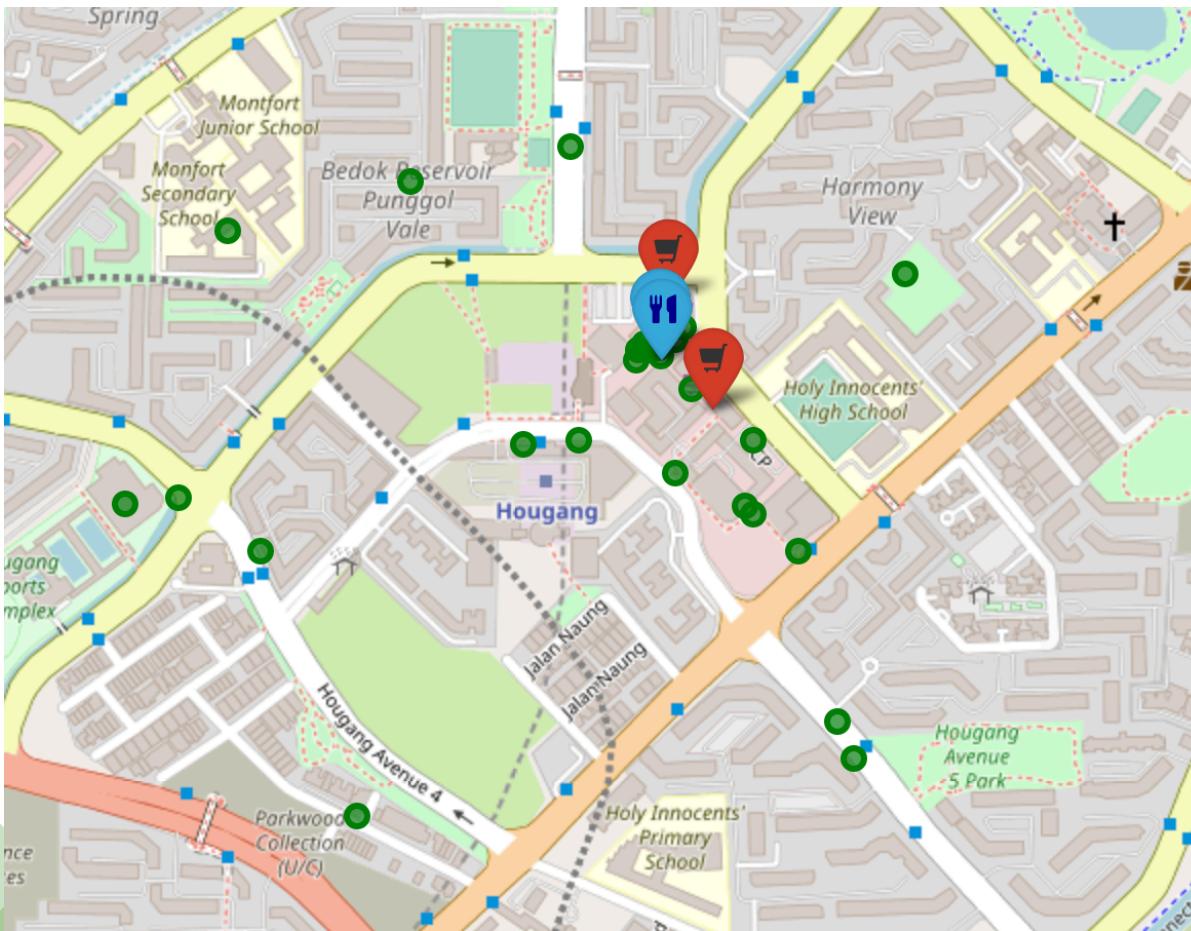
Blue marker: Cafe

Complete list of venues and map, refer to Annex in The Battle of Neighborhoods Report

Point of interest:

Venue	Venue Category
Starbucks	Coffee Shop
Starbucks	Coffee Shop
Churros Factory	Café
Wang Cafe	Café
Nice Cafe	Asian Restaurant
Maki-san	Sushi Restaurant
Osaka Ohsho	Restaurant
Hillion Mall	Shopping Mall
Bukit Panjang Plaza	Shopping Mall
The Gym Nation	Gym
Gym@Maysprings	Gym

Recommendation 2 – Hougang



Red marker: Shopping Mall

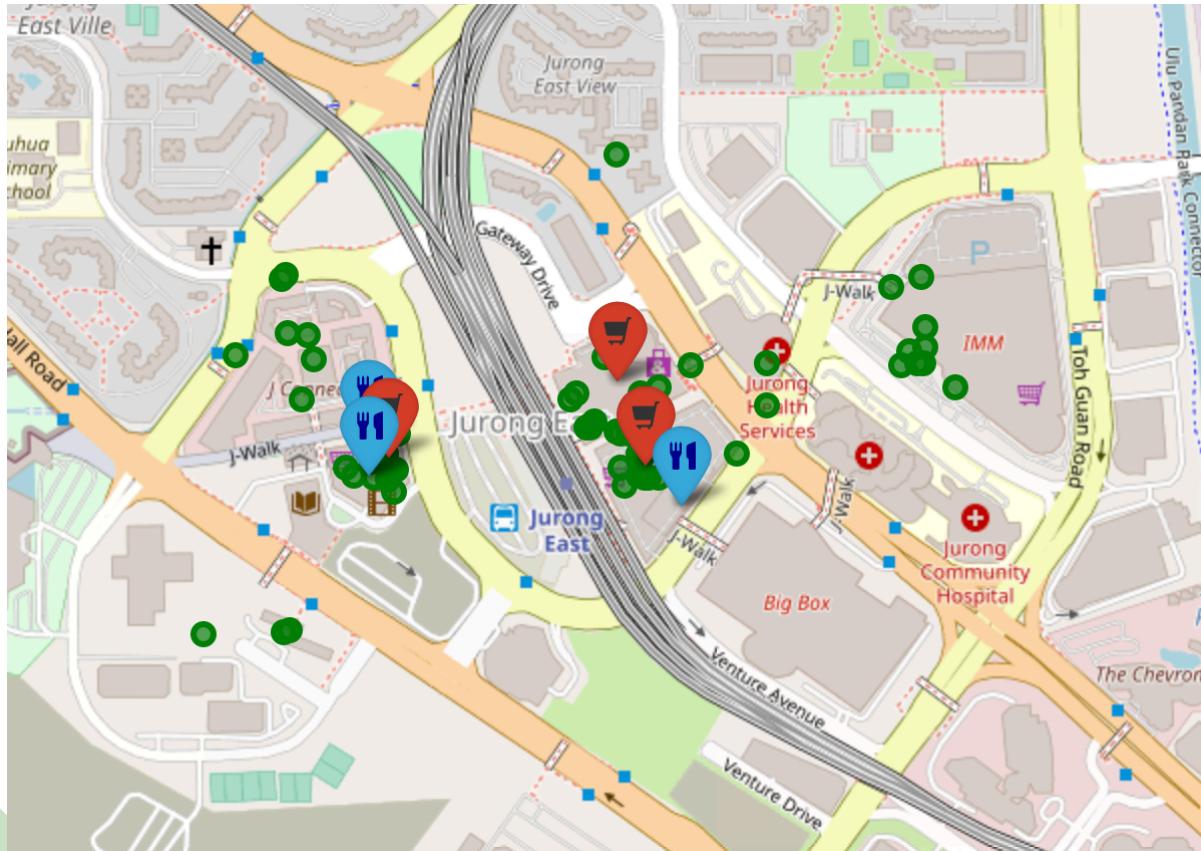
Blue marker: Cafe

Complete list of venues and map, refer to Annex in The Battle of Neighborhoods Report

Point of interest:

Venue	Venue Category
5 Star Corner Western Food	Coffee Shop
Toast Box	Café
Starbucks	Coffee Shop
Kawa Cafe	Café
Dessert Story	Dessert Shop
Yolé	Frozen Yogurt Shop
Subway	Sandwich Place
Hougang Sports Hall	Athletics & Sports
Watsons	Pharmacy
NTUC FairPrice	Supermarket
Hougang Mall	Shopping Mall
Kang Kar Mall	Shopping Mall
Hougang ActiveSG Gym	Gym

Recommendation 3 – Jurong East



Red marker: Shopping Mall

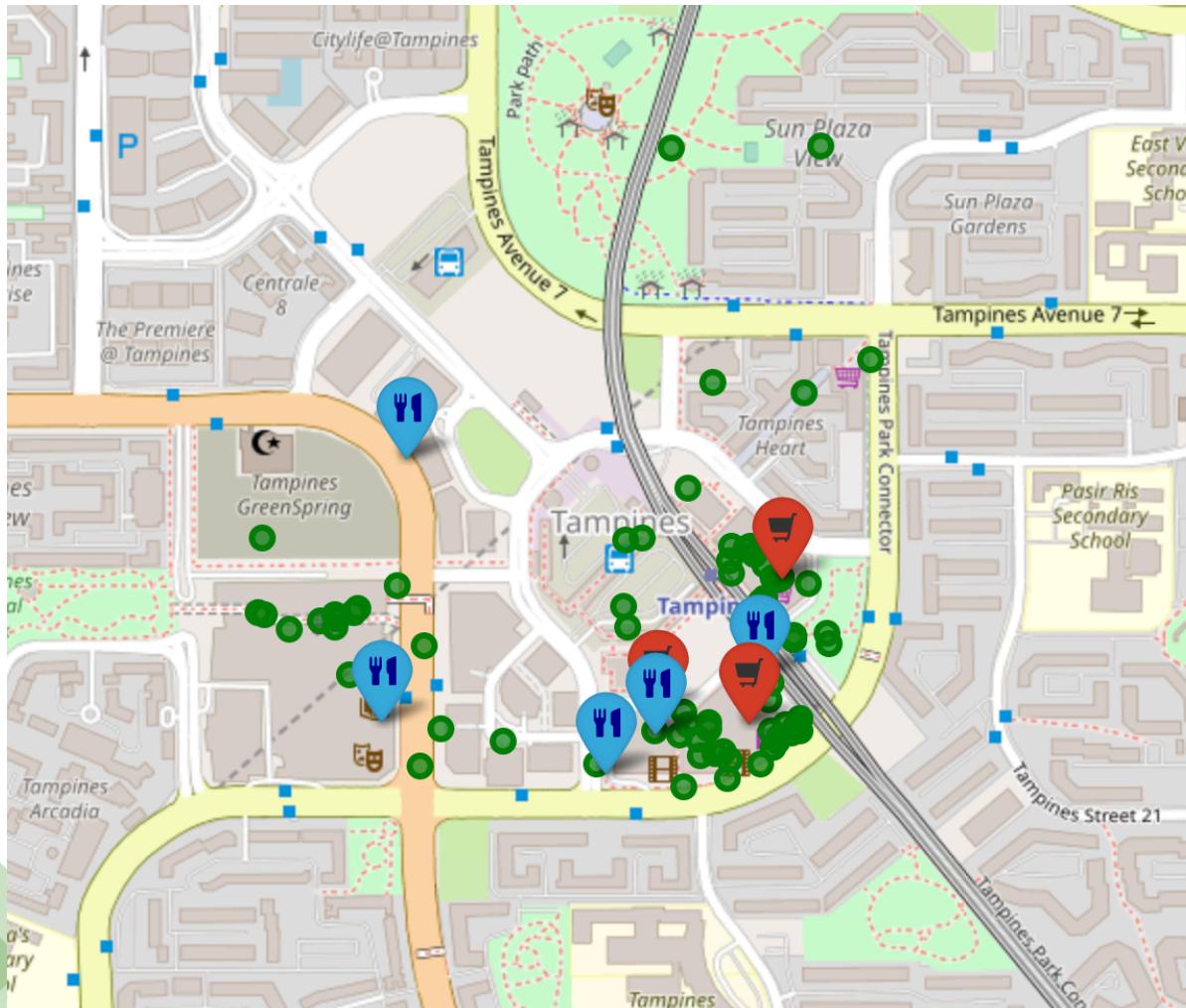
Blue marker: Cafe

Complete list of venues and map, refer to Annex in The Battle of Neighborhoods Report

Point of interest:

Venue	Venue Category
The Coffee Bean & Tea Leaf	Coffee Shop
Starbucks	Coffee Shop
Paris Baguette	Café
OldTown White Coffee	Café
JCube	Shopping Mall
Westgate	Shopping Mall
Jem	Shopping Mall
Robinsons	Department Store
FairPrice Xtra	Supermarket
Platinum Movie Suites	Movie Theater
Cathay Cineplexes	Multiplex
Fitness First Platinum	Gym / Fitness Center

Recommendation 4 – Punggol



Red marker: Shopping Mall

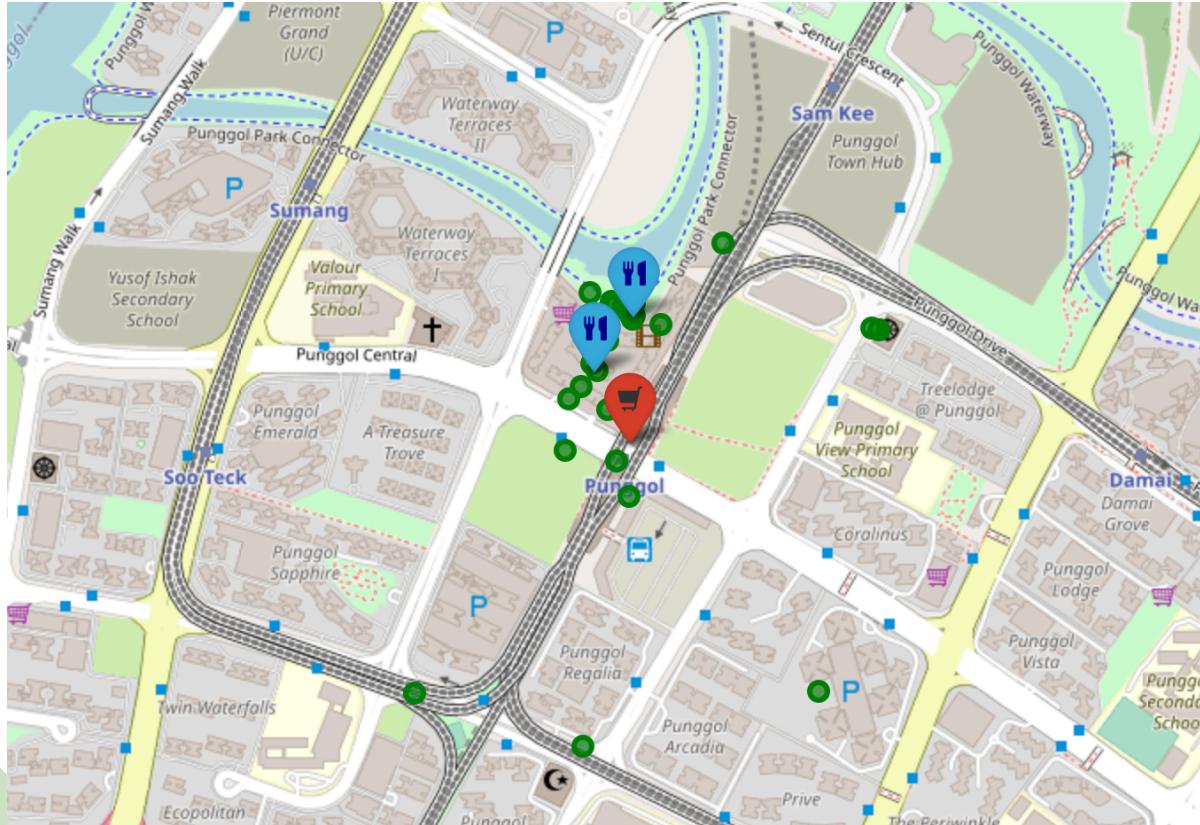
Blue marker: Cafe

Complete list of venues and map, refer to Annex in The Battle of Neighborhoods Report

Point of interest:

Venue	Venue Category
Cedele	Café
Starbucks	Coffee Shop
The Coffee Bean & Tea Leaf	Café
Water Drop Teahouse	Tea Room
Waterway Point	Shopping Mall
Sheng Siong Supermarket	Supermarket
Fairprice Finest	Supermarket
Doraemon Theme Park	Theme Park
Punggol Discovery Cube	Plaza
Shaw Theatres	Multiplex
Gymnasium @ PRIVÉ	Gym

Recommendation 5 – Tampines



Red marker: Shopping Mall

Blue marker: Cafe

Complete list of venues and map, refer to Annex in The Battle of Neighborhoods Report

Point of interest:

Venue	Venue Category
Starbucks Reserve	Coffee Shop
Commons	Café
Rocky Master	Café
The Coffee Bean & Tea Leaf	Coffee Shop
McCafe	Café
Tampines 1	Shopping Mall
Tampines Mall	Shopping Mall
Century Square	Shopping Mall
Isetan	Department Store
NTUC Fairprice	Supermarket
Amore Fitness	Gym
Fitness First	Gym
Healing Touch	Spa
Golden Village	Multiplex
SunPlaza Park	Park

Conclusions

- There are many factors people will need to consider when choosing neighborhoods for relocation, including housing availability, neighborhood environment, etc
- Analysis was based on the current preferences and requirements of our target customers
- Recommendations aimed to be an overview of neighborhoods that they would be interested in to further explore and decide on the suitable neighborhood for their family
- Other considerations or requirements (e.g. housing prices, availability of schools) can be added to future analysis to improve on recommendations or provide recommendations to different customers