Capstone Project –
Battle of the neighborhoods

SINGAPORE
NEIGHBORHOOD
RECOMMENDATION

Finding a suitable place to rent in Singapore

- Singapore is a highly developed market economy. Many multinational companies has their office located here

- Singapore is also one of the world's most expensive city to live in. Finding a suitable place to live is a big challenge.

- Expats have a difficult time in selecting the place to live with recommendations flying from different sources.

- Having places of interest of individual liking at nearby location is an important factor

- Students from other countries and Tourists may also find the output of project very useful

Data Acquisition and Cleaning

1. **Singapore Towns and median residential rental prices -** Data is retrieved from Singapore open dataset from median rent by town from https://data.gov.sg website.

    - Most recent recorded rental prices from this data source (2019-Q4) are being perused for the study
    - Unwanted columns are dropped
    - Simplified the analysis by using the average rental prices of all available flat type
    - The coordinates of towns are attached with the Singapore Town and Rent table using geopy library

2. **Singapore Top Venue Recommendations from FourSquare API**
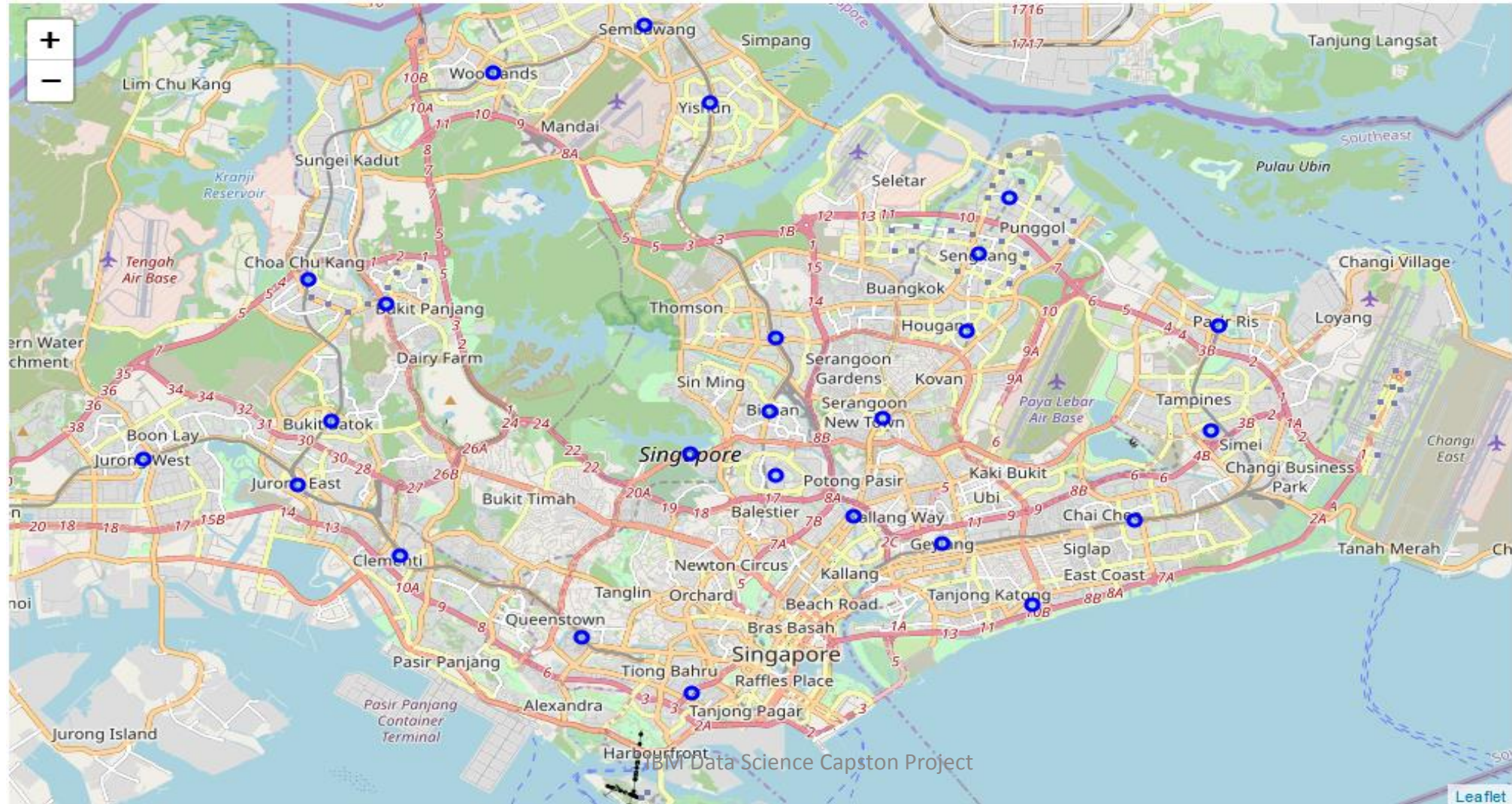
Data Acquisition and Cleaning

Singapore Towns with Rent and Coordinates

Town	median_rent	Latitude	Longitude
ANG MO KIO	2033.333333	1.370073	103.849516
BEDOK	2112.500000	1.323976	103.930216
BISHAN	2200.000000	1.351452	103.848250
BUKIT BATOK	1950.000000	1.349057	103.749591
BUKIT MERAH	2175.000000	1.280628	103.830591
BUKIT PANJANG	1757.500000	1.378629	103.762136
CENTRAL	2375.000000	1.340863	103.830392
CHOA CHU KANG	1883.333333	1.384749	103.744534
CLEMENTI	2250.000000	1.315100	103.765231
GEYLANG	2183.333333	1.318186	103.887056
HOUGANG	1982.500000	1.371904	103.892725
JURONG EAST	2100.000000	1.333115	103.742297
JURONG WEST	1982.500000	1.339636	103.707339
KALLANG/WHAMPOA	2293.333333	1.325110	103.867168
MARINE PARADE	1950.000000	1.302689	103.907395
PASIR RIS	2116.666667	1.373031	103.949255
PUNGGOL	1900.000000	1.405258	103.902330
QUEENSTOWN	2170.000000	1.294623	103.806045
SEMBAWANG	1900.000000	1.449093	103.820055
SENGKANG	1887.500000	1.391330	103.895294
SERANGOON	2162.500000	1.349763	103.873721
TAMPINES	2100.000000	1.346893	103.947341
TOA PAYOH	2166.666667	1.335391	103.849741
WOODLANDS	1762.500000	1.436897	103.786216
YISHUN	1887.500000	1.429384	103.835028

Singapore and its towns in a maps

The geograpical coordinate of Singapore 1.357107, 103.8194992.

Methodology

- After data acquisition and cleaning we will explore the neighborhoods and analyze each Singapore Town nearby recommended venues from the data downloaded using Foursquare.

- Then we will rank the top ten most common nearby place of interest.

- We will also run k-means to cluster the Towns into 5 clusters based on Top ten most common nearby venue and present our findings in a map.

Explore Neighborhoods

A single function **getNearbyVenues** is defined. It single handedly performs following tasks

- Uses the Foursquare credentials and version to get the data of nearby most common venues

- Get the top 100 venues that are within a radius of 500 meters of each of the towns

- The venues are categorized based on their basic functions. For example Fairprice shop is categorized as Supermarket, MacDonald is categorized burger Joint.

- Clean the json and structure it into a *pandas* dataframe

Explore Neighborhoods

A single function **getNearbyVenues** returns the beside table

Town	Town Latitude	Town Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
ANG MO KIO	42	42	42	42	42	42
BEDOK	58	58	58	58	58	58
BISHAN	44	44	44	44	44	44
BUKIT BATOK	22	22	22	22	22	22
BUKIT MERAH	39	39	39	39	39	39
BUKIT PANJANG	43	43	43	43	43	43
CENTRAL	4	4	4	4	4	4
CHOA CHU KANG	19	19	19	19	19	19
CLEMENTI	64	64	64	64	64	64
GEYLANG	35	35	35	35	35	35
HOUGANG	35	35	35	35	35	35
JURONG EAST	75	75	75	75	75	75
JURONG WEST	64	64	64	64	64	64
KALLANG/WHAMPOA	16	16	16	16	16	16
MARINE PARADE	37	37	37	37	37	37
PASIR RIS	31	31	31	31	31	31
PUNGGOL	47	47	47	47	47	47
QUEENSTOWN	16	16	16	16	16	16
SEMBAWANG	14	14	14	14	14	14
SENGKANG	31	31	31	31	31	31
SERANGOON	40	40	40	40	40	40
TAMPINES	15	15	15	15	15	15
TOA PAYOH	45	45	45	45	45	45
WOODLANDS	55	55	55	55	55	55
YISHUN	45	45	45	45	45	45

Analyze Each Neighborhood

One hot encoding technique is used analyze categories of most common venue in the vicinity of the town

Grouping rows by neighborhood and by taking the mean of the frequency of occurrence of each category gives beside table

	Town	Accessories Store	American Restaurant	Arcade	Arts & Crafts Store	Asian Restaurant	BBQ Joint	Bakery	Bank	Basketball Court	Bay	Bike Rental / Bike Share
0	ANG MO KIO	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.023810	0.02381	0.000000	0.00	0.000000
1	BEDOK	0.000000	0.017241	0.000000	0.000000	0.051724	0.000000	0.017241	0.00000	0.000000	0.00	0.000000
2	BISHAN	0.000000	0.000000	0.000000	0.000000	0.022727	0.000000	0.000000	0.00000	0.000000	0.00	0.000000
3	BUKIT BATOK	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00000	0.000000	0.00	0.000000
4	BUKIT MERAH	0.000000	0.025641	0.000000	0.000000	0.102564	0.000000	0.025641	0.00000	0.000000	0.00	0.000000
5	BUKIT PANJANG	0.000000	0.023256	0.000000	0.000000	0.069767	0.023256	0.023256	0.00000	0.000000	0.00	0.000000
6	CENTRAL	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00000	0.000000	0.25	0.000000
7	CHOA CHU KANG	0.000000	0.000000	0.000000	0.000000	0.052632	0.000000	0.052632	0.00000	0.000000	0.00	0.000000
8	CLEMENTI	0.000000	0.015625	0.000000	0.015625	0.046875	0.015625	0.015625	0.00000	0.000000	0.00	0.000000
9	GEYLANG	0.000000	0.000000	0.000000	0.000000	0.085714	0.000000	0.028571	0.00000	0.028571	0.00	0.000000
10	HOUGANG	0.000000	0.000000	0.000000	0.000000	0.028571	0.000000	0.000000	0.00000	0.000000	0.00	0.000000
11	JURONG EAST	0.013333	0.000000	0.000000	0.000000	0.013333	0.000000	0.013333	0.00000	0.000000	0.00	0.000000
12	JURONG WEST	0.000000	0.015625	0.000000	0.000000	0.109375	0.015625	0.015625	0.00000	0.000000	0.00	0.000000
13	KALLANG/WHAMPOA	0.000000	0.000000	0.000000	0.000000	0.062500	0.000000	0.000000	0.00000	0.000000	0.00	0.000000
14	MARINE PARADE	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.027027	0.00000	0.000000	0.00	0.000000
15	PASIR RIS	0.000000	0.000000	0.000000	0.000000	0.032258	0.000000	0.064516	0.00000	0.000000	0.00	0.032258
16	PUNGGOL	0.000000	0.000000	0.000000	0.000000	0.042553	0.000000	0.042553	0.00000	0.000000	0.00	0.000000
17	QUEENSTOWN	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00000	0.000000	0.00	0.000000
18	SEMBAWANG	0.000000	0.000000	0.000000	0.000000	0.000000	0.071429	0.000000	0.00000	0.000000	0.00	0.000000

Analyze Each Neighborhood

Function **return_most_common_v enues** to sort the venues in descending order.

Then create the new dataframe and display the top 10 venues for each neighborhood. The output is indicated:

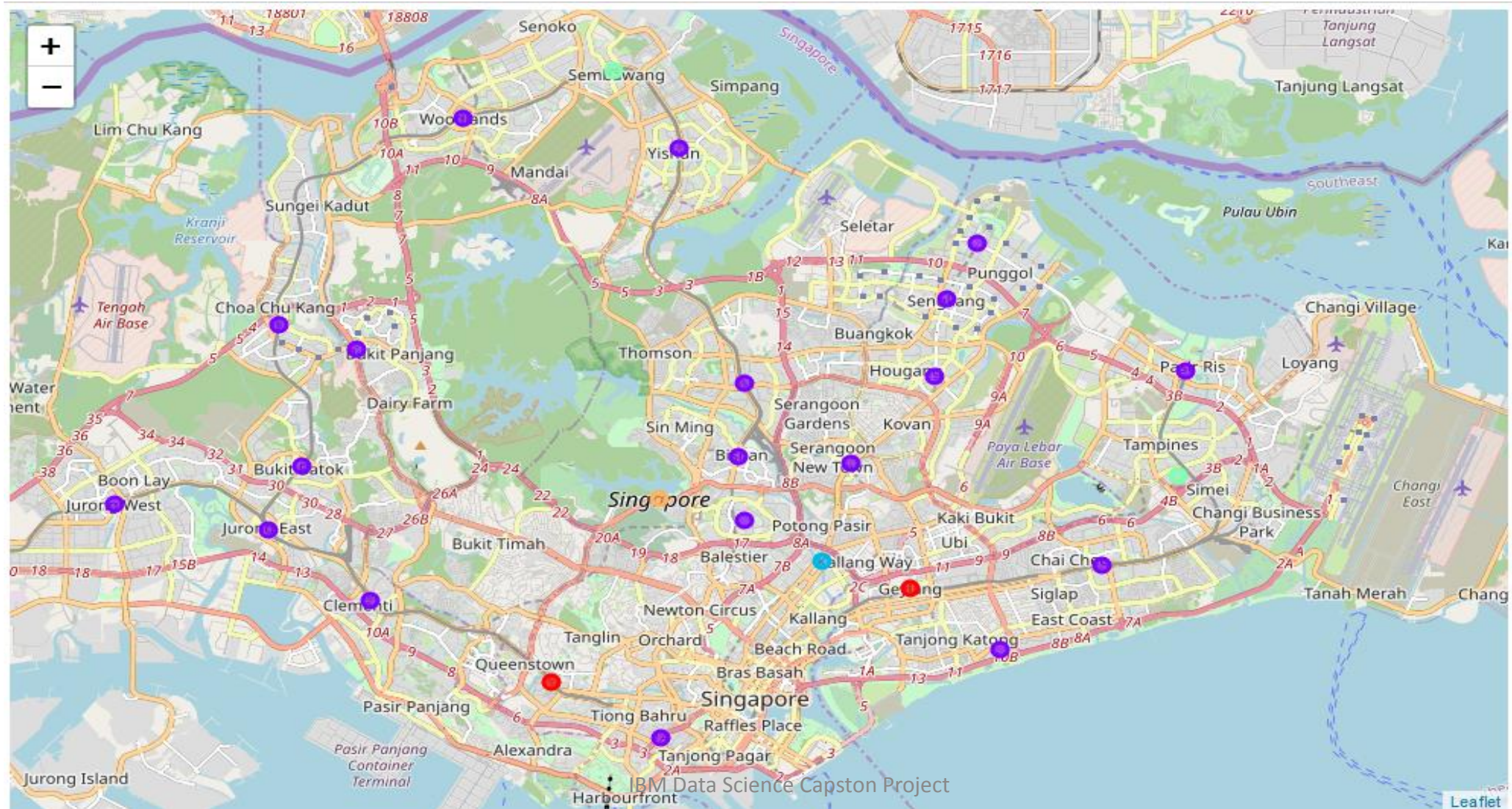|	Town	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	ANG MO KIO	Coffee Shop	Food Court	Supermarket	Snack Place	Japanese Restaurant	Bubble Tea Shop	Dessert Shop	Modern European Restaurant	Chinese Restaurant	Ramen Restaurant
1	BEDOK	Chinese Restaurant	Coffee Shop	Asian Restaurant	Japanese Restaurant	Sandwich Place	Food Court	Café	Supermarket	Fast Food Restaurant	Malay Restaurant
2	BISHAN	Food Court	Bubble Tea Shop	Café	Coffee Shop	Chinese Restaurant	Cosmetics Shop	Supermarket	Ice Cream Shop	Japanese Restaurant	Shopping Mall
3	BUKIT BATOK	Coffee Shop	Fast Food Restaurant	Chinese Restaurant	Food Court	Department Store	Grocery Store	Sandwich Place	Multiplex	Café	Shopping Mall
4	BUKIT MERAH	Café	Asian Restaurant	Chinese Restaurant	Food Court	Coffee Shop	Bookstore	Noodle House	Restaurant	Yoga Studio	Bus Station

Cluster Neighborhoods

- Run *k*-means to cluster the neighborhood into 5 clusters
- create a new dataframe that includes the cluster as well as the top 10 venues for each neighborhood.

	Town	median_rent	Latitude	Longitude	Cluster Labels	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
0	ANG MO KIO	2033.333333	1.370073	103.849516	1	Coffee Shop	Food Court	Supermarket	Snack Place	Japanese Restaurant	Bubble Tea Shop	Dessert Shop
1	BEDOK	2112.500000	1.323976	103.930216	1	Chinese Restaurant	Coffee Shop	Asian Restaurant	Japanese Restaurant	Sandwich Place	Food Court	Café
2	BISHAN	2200.000000	1.351452	103.848250	1	Food Court	Bubble Tea Shop	Café	Coffee Shop	Chinese Restaurant	Cosmetics Shop	Supermarket
3	BUKIT BATOK	1950.000000	1.349057	103.749591	1	Coffee Shop	Fast Food Restaurant	Chinese Restaurant	Food Court	Department Store	Grocery Store	Sandwich Place
4	BUKIT MERAH	2175.000000	1.280628	103.830591	1	Café	Asian Restaurant	Chinese Restaurant	Food Court	Coffee Shop	Bookstore	Noodle House

Cluster Neighborhoods

Examine Clusters

Cluster 1

More Chinese food centric with Chinese food, noodle house and food court appearing frequently in topmost common venue.

	Town	median_rent	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
9	GEYLANG	2183.333333	Chinese Restaurant	Noodle House	Asian Restaurant	Food Court	Vegetarian / Vegan Restaurant	Seafood Restaurant	Grocery Store	Dim Sum Restaurant	Pool
17	QUEENSTOWN	2170.000000	Food Court	Chinese Restaurant	Noodle House	Steakhouse	Pool	Train Station	Italian Restaurant	Spa	Seafood Restaurant

Examine Clusters

Cluster 2

More diverse than cluster 1 in terms of food choices. coffee shop, food court, fast food restaurants are among the topmost common venue

	Town	median_rent	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
0	ANG MO KIO	2033.333333	Coffee Shop	Food Court	Supermarket	Snack Place	Japanese Restaurant	Bubble Tea Shop	Dessert Shop	Modern European Restaurant	Chinese Restaurant
1	BEDOK	2112.500000	Chinese Restaurant	Coffee Shop	Asian Restaurant	Japanese Restaurant	Sandwich Place	Food Court	Café	Supermarket	Fast Food Restaurant
2	BISHAN	2200.000000	Food Court	Bubble Tea Shop	Café	Coffee Shop	Chinese Restaurant	Cosmetics Shop	Supermarket	Ice Cream Shop	Japanese Restaurant
3	BUKIT BATOK	1950.000000	Coffee Shop	Fast Food Restaurant	Chinese Restaurant	Food Court	Department Store	Grocery Store	Sandwich Place	Multiplex	Café
4	BUKIT MERAH	2175.000000	Café	Asian Restaurant	Chinese Restaurant	Food Court	Coffee Shop	Bookstore	Noodle House	Restaurant	Yoga Studio
5	BUKIT PANJANG	1757.500000	Coffee Shop	Fast Food Restaurant	Café	Sushi Restaurant	Asian Restaurant	Ramen Restaurant	Noodle House	Pharmacy	Shopping Mall
7	CHOA CHU KANG	1883.333333	Coffee Shop	Fast Food Restaurant	Supermarket	Café	Food Court	Smoke Shop	Furniture / Home Store	Bubble Tea Shop	Sandwich Place

Examine Clusters

Cluster 3

The Chinese restaurant and coffee shop are the most common venue in this cluster, with very high frequency than any other cluster

	Town	median_rent	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
13	KALLANG/WHAMPOA	2293.333333	Chinese Restaurant	Coffee Shop	Pet Store	Food Court	Park	Convenience Store	Restaurant	Bus Station	Asian Restaurant

Examine Clusters

Cluster 4

Coffee shop and fast food restaurant are the most common venue in this cluster

	Town	median_rent	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
18	SEMBAWANG	1900.0	Coffee Shop	Fast Food Restaurant	Chinese Restaurant	Snack Place	Bus Station	Bus Stop	Bistro	Supermarket	Shopping Mall
21	TAMPINES	2100.0	Coffee Shop	Supermarket	Chinese Restaurant	Fast Food Restaurant	Seafood Restaurant	Indian Restaurant	Bus Station	Food Court	Noodle House

Examine Clusters

Cluster 5

Outgoing and adventure-oriented Cluster

	Town	median_rent	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
6	CENTRAL	2375.0	Bay	Reservoir	Trail	Park	Yoga Studio	Food Truck	Food Court	Flower Shop	Fish & Chips Shop	Filipino Restaurant

Conclusion

- The findings of the Project will be a good supplement to web based recommendations for expats, visitors, students to find out nearby venues of interest

- Sourcing from the venue recommendations from Foursquare has its limitation, the list of venues is not exhaustive and need update time to time

- The towns in cluster 2 has a wide range of rent and offer more diverse set of choices for food. Therefore for expats choosing a town in cluster 2 is preferred