

Applied Data Science Capstone

Capstone Project - The Battle of Neighbourhoods (Week 2)

1. Purpose

This document provides the details of my final peer reviewed assignment for the IBM Data Science Professional Certificate program – Coursera Capstone.

2. Problem statement

Singapore is City state located in busiest shipping lanes in the world; it's a small island nation with a thriving tourism industry. The internet provides host of information to visit and stay in and around Singapore. However, most of the recommendations are simply based on usual tourist attractions or residential areas that are mostly expensive or already known for travellers based on certain keywords like "Hotel", or "Backpackers" etc. The intention on this project is to collect and provide a data driven recommendation that can supplement the recommendation with statistical data. This will also be utilizing data retrieved from Singapore open data sources and FourSquare API venue recommendations.

3. Data Sources

This demonstration will make use of the following data sources:

- Singapore Towns and median residential rental prices
Data will retrieved from Singapore open dataset from [median rent by town and flattype](#) from <https://data.gov.sg> website.
- Singapore Towns location data retrieved using Google maps API
- Singapore Top Venue Recommendations from FourSquare API

4. Methodology

Singapore Towns List with median residential rental prices.

The idea is to run analysis by using the average rental prices of all available flat type.

Data Cleansing and preprocessing

The retrieved table contains some un-wanted entries and needs some cleanup. The following tasks will be performed:

- Drop/ignore cells with missing data.
- Use most current data record.
- Fix data types.

5. Segmenting and Clustering Towns in Singapore

Using the Foursquare API, the **explore** API function was used to get the most common venue categories in each neighborhood, and then used this feature to group the neighborhoods into clusters. The *k*-means clustering algorithm was used for the analysis. Finally, the Folium library is used to visualize the recommended neighborhoods and their emerging clusters.

In the ipynb notebook, the function **getNearbyVenues** extracts the following Information for the dataframe it generates:

- Venue ID
- Venue Name
- Coordinates : Latitude and Longitude
- Category Name

6. Conclusion

On this notebook, Analysis of best town venue recommendations based on Food venue category has been showcased. Recommendations based on other user searches like available outdoor and recreation areas are also available. As Singapore is a small country with a whole host of interesting venues scattered around the town, the information extracted in this notebook present on the town areas, will be a good supplement to web based recommendations for visitors to find out nearby venues of interest and be a useful aid in deciding a place to stay or where to go during their visits.

Result shows that most popular food venue among Singaporeans, residents and visitors are **Food Courts, Coffee Shops and Fast Food Restaurants**. The major cluster being one's with most restaurants and nightlife, see Table 1 and Figure 2. Other details refer Annex A and B.

Top 10 Recommended Venue

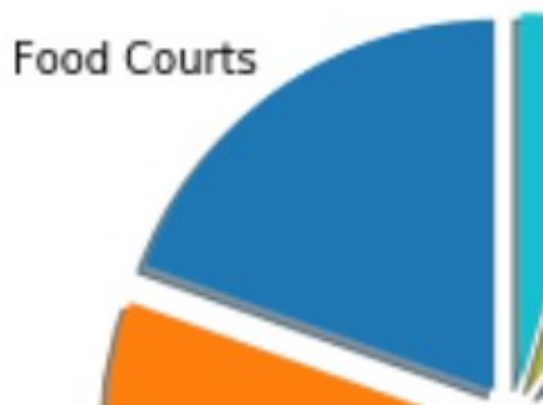


Figure 1: Top 10 Food Category type

	median_rent	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
Town							
KALLANG/ WHAMPOA	2300	1.32975	103.8638	0	Food Courts	Coffee Shops	Chinese Restaurants
GEYLANG	2200	1.318186	103.8871	0	Coffee Shops	Asian Restaurants	Chinese Restaurants
QUEENSTOWN	2200	1.294623	103.806	0	Coffee Shops	Chinese Restaurants	Food Courts
TOA PAYOH	2166.667	1.335391	103.8497	0	Food Courts	Chinese Restaurants	Coffee Shops
BUKIT MERAH	2162.5	1.270439	103.8283	0	Food Courts	Cafés	Coffee Shops
BEDOK	2120	1.323976	103.9302	0	Coffee Shops	Chinese Restaurants	Food Courts
HOUGAN G	2025	1.370682	103.8925	0	Coffee Shops	Chinese Restaurants	Food Courts

Table 1: Major cluster for food venues

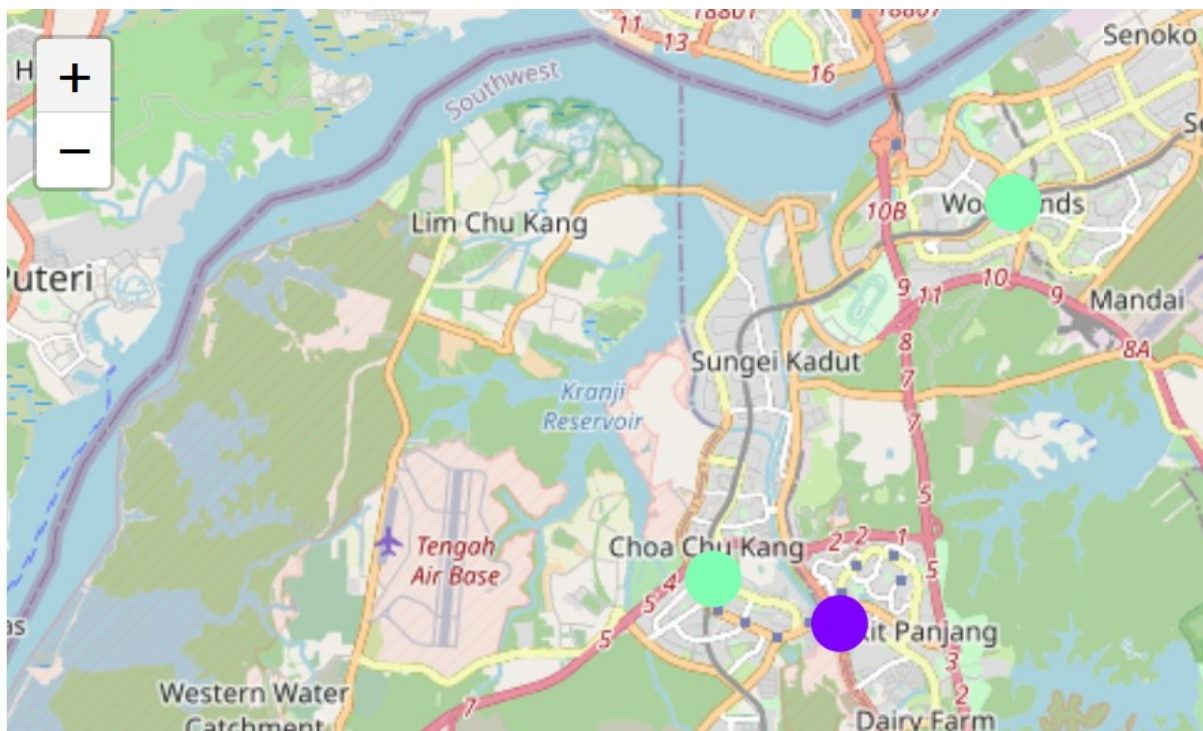


Figure 2: Clusters generated based on the dataset

Annex A - Analyze Singapore Town most visited venues

Town=< ANG MO KIO >

	venue	freq
0	Food Courts	0.16
1	Bubble Tea Shops	0.10
2	Fast Food Restaurants	0.08
3	Coffee Shops	0.08
4	Chinese Restaurants	0.08
5	Malay Restaurants	0.04
6	Japanese Restaurants	0.04
7	Noodle Houses	0.04
8	Dessert Shops	0.04
9	Fried Chicken Joints	0.02

Town=< BEDOK >

	venue	freq
0	Coffee Shops	0.16
1	Food Courts	0.12
2	Chinese Restaurants	0.12
3	Cafés	0.08
4	Bakeries	0.06
5	Fast Food Restaurants	0.04
6	Indian Restaurants	0.04
7	Burger Joints	0.04
8	Noodle Houses	0.04
9	Sandwich Places	0.04

Town=< BISHAN >

	venue	freq
0	Food Courts	0.12
1	Coffee Shops	0.12
2	Chinese Restaurants	0.12
3	Fast Food Restaurants	0.06
4	Bakeries	0.06
5	Cafés	0.04
6	Japanese Restaurants	0.04
7	Bubble Tea Shops	0.04
8	Snack Places	0.04
9	Noodle Houses	0.02

Town=< BUKIT BATOK >

	venue	freq
0	Food Courts	0.22
1	Chinese Restaurants	0.12
2	Coffee Shops	0.08
3	Asian Restaurants	0.06
4	Fast Food Restaurants	0.04
5	Bakeries	0.04
6	Food Trucks	0.02
7	Event Spaces	0.02
8	Fish & Chips Shops	0.02
9	Sandwich Places	0.02

Town=< BUKIT MERAH >

	venue	freq
0	Cafés	0.10
1	Food Courts	0.10
2	Coffee Shops	0.08
3	Chinese Restaurants	0.06
4	Thai Restaurants	0.06
5	Seafood Restaurants	0.06
6	Asian Restaurants	0.06
7	Veg / Vegan Restaurants	0.04
8	Bakeries	0.04
9	Snack Places	0.04

Town=< BUKIT PANJANG >

	venue	freq
0	Asian Restaurants	0.12
1	Fast Food Restaurants	0.08
2	Bakeries	0.08
3	Coffee Shops	0.08
4	Chinese Restaurants	0.06
5	Japanese Restaurants	0.06
6	Italian Restaurants	0.04
7	Food Courts	0.04
8	Sandwich Places	0.04
9	Snack Places	0.04

Town=< CENTRAL >

	venue	freq
0	Restaurants	0.25
1	Tea Rooms	0.25
2	Coffee Shops	0.25
3	Cafés	0.25
4	Middle Eastern Restaurants	0.00
5	Poke Places	0.00
6	Pizza Places	0.00
7	Pharmacies	0.00
8	Noodle Houses	0.00
9	Night Markets	0.00

Town=< CHOA CHU KANG >

	venue	freq
0	Coffee Shops	0.14
1	Bakeries	0.10
2	Noodle Houses	0.10
3	Dessert Shops	0.06
4	Chinese Restaurants	0.06
5	Food Courts	0.06
6	Cafés	0.06
7	Snack Places	0.04
8	Fast Food Restaurants	0.04
9	Restaurants	0.04

Town=< CLEMENTI >

	venue	freq
0	Food Courts	0.20
1	Chinese Restaurants	0.08
2	Bakeries	0.06
3	Fast Food Restaurants	0.06
4	Coffee Shops	0.06
5	Asian Restaurants	0.04
6	Snack Places	0.04
7	Dessert Shops	0.04
8	Soup Places	0.04
9	Sushi Restaurants	0.04

Town=< JURONG WEST >

	venue	freq
0	Fast Food Restaurants	0.14
1	Bubble Tea Shops	0.10
2	Asian Restaurants	0.10
3	Food Courts	0.08
4	Snack Places	0.08
5	Bakeries	0.06
6	Noodle Houses	0.04
7	Thai Restaurants	0.04
8	Cafés	0.04
9	Sushi Restaurants	0.04

Town=< GEYLANG >

	venue	freq
0	Coffee Shops	0.18
1	Asian Restaurants	0.10
2	Chinese Restaurants	0.10
3	Noodle Houses	0.08
4	Fast Food Restaurants	0.06
5	Veg/ Vegan Restaurants	0.06
6	Seafood Restaurants	0.04
7	Dessert Shops	0.04
8	Food Courts	0.04
9	Dim Sum Restaurants	0.02

Town=< KALLANG/WHAMPOA >

	venue	freq
0	Coffee Shops	0.12
1	Food Courts	0.12
2	Chinese Restaurants	0.10
3	Noodle Houses	0.06
4	Cafés	0.04
5	BBQ Joints	0.04
6	Snack Places	0.04
7	Fast Food Restaurants	0.04
8	Asian Restaurants	0.04
9	Fried Chicken Joints	0.04

Town=< HOUGANG >

	venue	freq
0	Coffee Shops	0.14
1	Chinese Restaurants	0.14
2	Food Courts	0.12
3	Cafés	0.08
4	Fast Food Restaurants	0.06
5	BBQ Joints	0.06
6	Bakeries	0.06
7	Snack Places	0.06
8	Dessert Shops	0.06
9	Soup Places	0.02

Town=< MARINE PARADE >

	venue	freq
0	Coffee Shops	0.10
1	Food Courts	0.08
2	Bakeries	0.08
3	Dessert Shops	0.08
4	Chinese Restaurants	0.08
5	Indian Restaurants	0.06
6	Fast Food Restaurants	0.04
7	Japanese Restaurants	0.04
8	Thai Restaurants	0.04
9	Seafood Restaurants	0.04

Town=< JURONG EAST >

	venue	freq
0	Food Courts	0.12
1	Cafés	0.10
2	Coffee Shops	0.08
3	Fast Food Restaurants	0.08
4	Juice Bars	0.04
5	Sushi Restaurants	0.04
6	Chinese Restaurants	0.04
7	Japanese Restaurants	0.04
8	Steakhouses	0.04
9	Thai Restaurants	0.04

Town=< PASIR RIS >

	venue	freq
0	Chinese Restaurants	0.10
1	Bakeries	0.10
2	Food Courts	0.08
3	Fast Food Restaurants	0.06
4	Dessert Shops	0.06
5	Bubble Tea Shops	0.04
6	Japanese Restaurants	0.04
7	Cafés	0.04
8	Asian Restaurants	0.04
9	Snack Places	0.04

#	Town=< PUNGGOL >	venue	freq
0		Food Courts	0.12
1		Cafés	0.08
2		Japanese Restaurants	0.08
3		Chinese Restaurants	0.08
4		Fast Food Restaurants	0.06
5		Juice Bars	0.04
6		Fried Chicken Joints	0.04
7		Bakeries	0.04
8		Coffee Shops	0.04
9		Sushi Restaurants	0.04

#	Town=< QUEENSTOWN >	venue	freq
0		Chinese Restaurants	0.17
1		Coffee Shops	0.17
2		Food Courts	0.10
3		Cafés	0.08
4		Asian Restaurants	0.06
5		Indian Restaurants	0.06
6		Noodle Houses	0.04
7		Cafeterias	0.04
8		American Restaurants	0.02
9		Seafood Restaurants	0.02

#	Town=< SEMBAWANG >	venue	freq
0		Food Courts	0.16
1		Asian Restaurants	0.08
2		Coffee Shops	0.06
3		Chinese Restaurants	0.06
4		Food	0.04
5		BBQ Joints	0.04
6		Food Trucks	0.04
7		Dessert Shops	0.04
8		Seafood Restaurants	0.04
9		Pizza Places	0.04

#	Town=< SENGKANG >	venue	freq
0		Bakeries	0.12
1		Coffee Shops	0.10
2		Asian Restaurants	0.08
3		Food Courts	0.08
4		Fast Food Restaurants	0.08
5		Cafés	0.06
6		Chinese Restaurants	0.06
7		Noodle Houses	0.04
8		Bubble Tea Shops	0.04
9		Thai Restaurants	0.04

#	Town=< SERANGOON >	venue	freq
0		Coffee Shops	0.14
1		Fast Food Restaurants	0.10
2		Food Courts	0.06
3		Chinese Restaurants	0.06
4		Cafés	0.06

5	Japanese Restaurants	0.06
6	Dessert Shops	0.04
7	Bakeries	0.04
8	Soup Places	0.04
9	Sushi Restaurants	0.04

#	Town=< TAMPINES >	venue	freq
0		Food Courts	0.12
1		Bakeries	0.10
2		Fast Food Restaurants	0.10
3		Bubble Tea Shops	0.08
4		Coffee Shops	0.08
5		Chinese Restaurants	0.06
6		Indian Restaurants	0.06
7		Sushi Restaurants	0.04
8		Donut Shops	0.04
9		Snack Places	0.04

#	Town=< TOA PAYOH >	venue	freq
0		Food Courts	0.20
1		Chinese Restaurants	0.16
2		Noodle Houses	0.10
3		Coffee Shops	0.10
4		Fast Food Restaurants	0.08
5		Thai Restaurants	0.06
6		Cafés	0.04
7		Asian Restaurants	0.04
8		Steakhouses	0.04
9		Japanese Restaurants	0.02

#	Town=< WOODLANDS >	venue	freq
0		Fast Food Restaurants	0.12
1		Food Courts	0.10
2		Chinese Restaurants	0.08
3		Bakeries	0.06
4		Snack Places	0.06
5		Coffee Shops	0.06
6		Cafés	0.04
7		Pizza Places	0.04
8		Sushi Restaurants	0.04
9		Juice Bars	0.02

#	Town=< YISHUN >	venue	freq
0		Food Courts	0.14
1		Fried Chicken Joints	0.08
2		Chinese Restaurants	0.08
3		Noodle Houses	0.06
4		Cafés	0.06
5		Fast Food Restaurants	0.04
6		Bakeries	0.04
7		Sushi Restaurants	0.04
8		Bubble Tea Shops	0.04
9		Coffee Shops	0.04

Annex B – Top venues sorted by town

Town	median_rent	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
KALLANG /WHAMPOA	2300	1.32975	103.8638	0	Food Courts	Coffee Shops	Chinese Restaurants
GEYLANG	2200	1.318186	103.8871	0	Coffee Shops	Asian Restaurants	Chinese Restaurants
QUEENSTOWN	2200	1.294623	103.806	0	Coffee Shops	Chinese Restaurants	Food Courts
TOA PAYOH	2166.667	1.335391	103.8497	0	Food Courts	Chinese Restaurants	Coffee Shops
BUKIT MERAH	2162.5	1.270439	103.8283	0	Food Courts	Cafés	Coffee Shops
BEDOK	2120	1.323976	103.9302	0	Coffee Shops	Chinese Restaurants	Food Courts
HOUGANG	2025	1.370682	103.8925	0	Coffee Shops	Chinese Restaurants	Food Courts
JURONG WEST	2012.5	1.339636	103.7073	1	Fast Food Restaurants	Asian Restaurants	Bubble Tea Shops
SENGKANG	1983.333	1.391654	103.8954	1	Bakeries	Coffee Shops	Asian Restaurants
BUKIT PANJANG	1750	1.378629	103.7621	1	Asian Restaurants	Coffee Shops	Bakeries
CLEMENTI	2250	1.3151	103.7652	2	Food Courts	Chinese Restaurants	Fast Food Restaurants
ANG MO KIO	2066.667	1.370073	103.8495	2	Food Courts	Bubble Tea Shops	Fast Food Restaurants
BUKIT BATOK	1950	1.349057	103.7496	2	Food Courts	Chinese Restaurants	Coffee Shops
SEMBAWANG	1933.333	1.449093	103.8201	2	Food Courts	Asian Restaurants	Chinese Restaurants

Town	media n_rent	Latitude	Longitud e	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
TAMPINES	2125	1.354653	103.9436	3	Food Courts	Fast Food Restaurants	Bakeries
PASIR RIS	2116.667	1.373031	103.9493	3	Chinese Restaurants	Bakeries	Food Courts
SERANGOON	2100	1.349862	103.8737	3	Coffee Shops	Fast Food Restaurants	Cafés
JURONG EAST	2000	1.333115	103.7423	3	Food Courts	Cafés	Fast Food Restaurants
CHOA CHU KANG	1933.333	1.384749	103.7445	3	Coffee Shops	Bakeries	Noodle Houses
MARINE PARADE	1925	1.302689	103.9074	3	Coffee Shops	Dessert Shops	Bakeries
PUNGGOL	1900	1.405258	103.9023	3	Food Courts	Chinese Restaurants	Cafés
WOODLANDS	1737.5	1.436897	103.7862	3	Fast Food Restaurants	Food Courts	Chinese Restaurants
CENTRAL	2400	1.340863	103.8304	4	Cafés	Tea Rooms	Restaurants