Segmenting and Clustering Neighbourhoods in Singapore to Provide Recommendations for Visitors

Linus Fones

7 May, 2020

1. Introduction

1.1 Background

Singapore is a cosmopolitan metropolis just like New York and Toronto. Singapore is divided into 28 separate districts (or boroughs) and each district is made up of one or more neighbourhoods which offer distinctively different living experience. For example, a fascinating blend of old and new greets visitors in Singapore's bustling Chinatown (in District 01), a haven for foodies, shoppers and history buffs. In contrast, Orchard Road (in District 09) is the ultimate browsing and buying destination in Singapore. Known for high-end boutiques and gleaming malls that beckon shoppers, this is Asia's most famous shopping street.

Chinatown (in District 01)

Orchard Road (in District 09)

1.2 Problem Statement

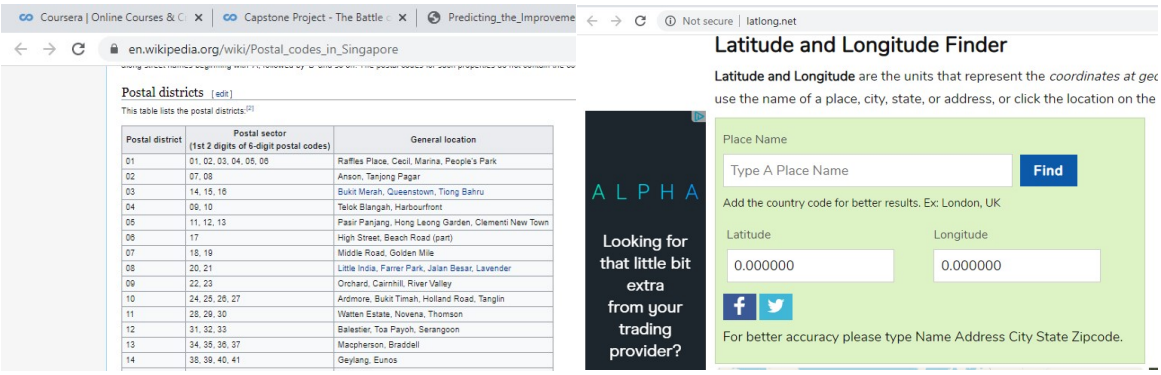
If you are an expat or tourist, you might need recommendations on where to eat, where to stay, and where to explore in Singapore. In this capstone project, I will attempt to answer these questions with the help of location data provide by Foursquare API.

The findings of this study can also be used by potential business owners when they are trying to start their businesses in Singapore. For example, if someone is looking to open a restaurant in Singapore, he might need recommendations on where to locate his restaurant.

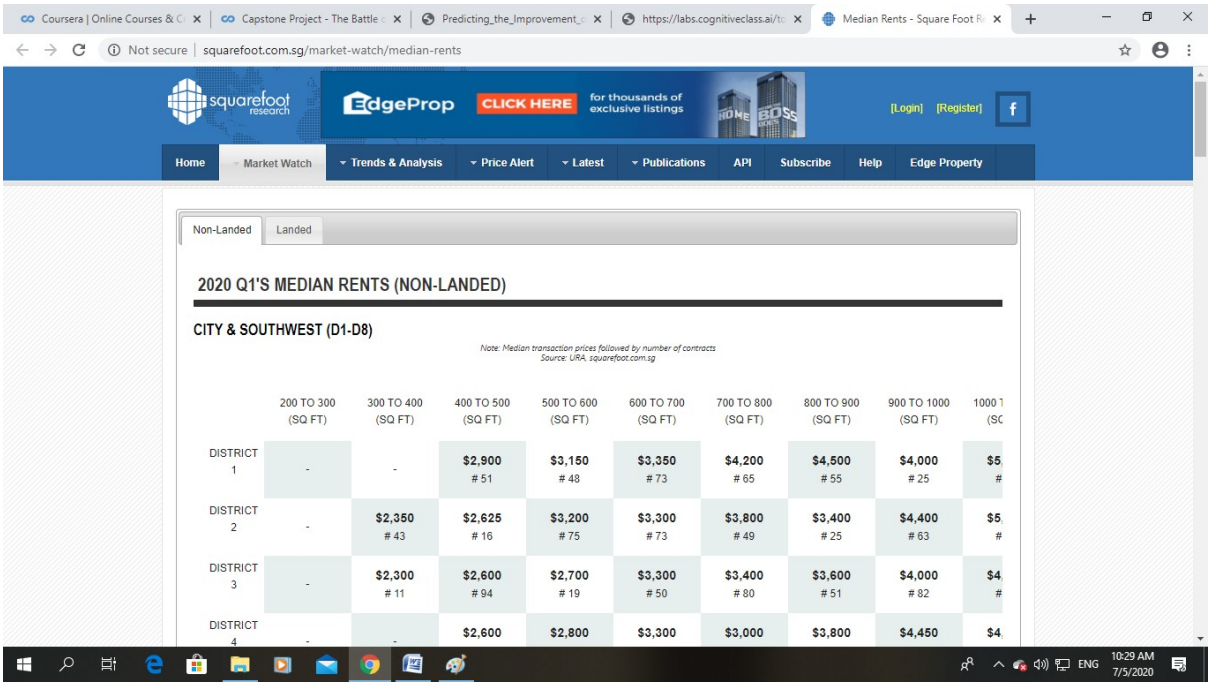
2. Data Acquisition and Cleaning

2.1 Data Sources

A Wikipedia page exists that has all the postal districts and neighbourhoods needed to explore and cluster the districts (or boroughs) in Singapore. There is also a website (https://www.latlong.net/) that enables the user to find the latitude and longitude of each postal district.



For the cost-conscious, there is a website that publishes the median rents for each district on a quarterly basis (https://www.squarefoot.com.sg/market-watch/median-rents).



2.2 Data Preparation

Firstly, the Wikipedia page was scraped using the Beautiful Soup Python package [1], and the data wrangled and cleansed before it was read into a pandas dataframe (see picture below).

Launcher | Segmenting and Clustering

Code

[3]:

	Postal District	Postal Sector	General Location
1	01	01, 02, 03, 04, 05, 06	Raffles Place, Cecil, Marina, People's Park
2	02	07, 08	Anson, Tanjong Pagar
3	03	14, 15, 16	Bukit Merah,Queenstown,Tiong Bahru
4	04	09, 10	Telok Blangah, Harbourfront
5	05	11, 12, 13	Pasir Panjang, Hong Leong Garden, Clementi New...
6	06	17	High Street, Beach Road (part)
7	07	18, 19	Middle Road, Golden Mile
8	08	20, 21	Little India,Farrer Park,Jalan Besar,Lavender
9	09	22, 23	Orchard, Cairnhill, River Valley
10	10	24, 25, 26, 27	Ardmore, Bukit Timah, Holland Road, Tanglin
11	11	28, 29, 30	Watten Estate, Novena, Thomson

Next, the latitude and longitude of each district was found using latlong.net and added to the dataframe. Rental prices were also inserted to provide a cost dimension for analysis. Foursquare API [2] was then called to obtain a list of venues for each district (see pictures below).

Launcher | Segmenting and Clustering

Code

[5]:

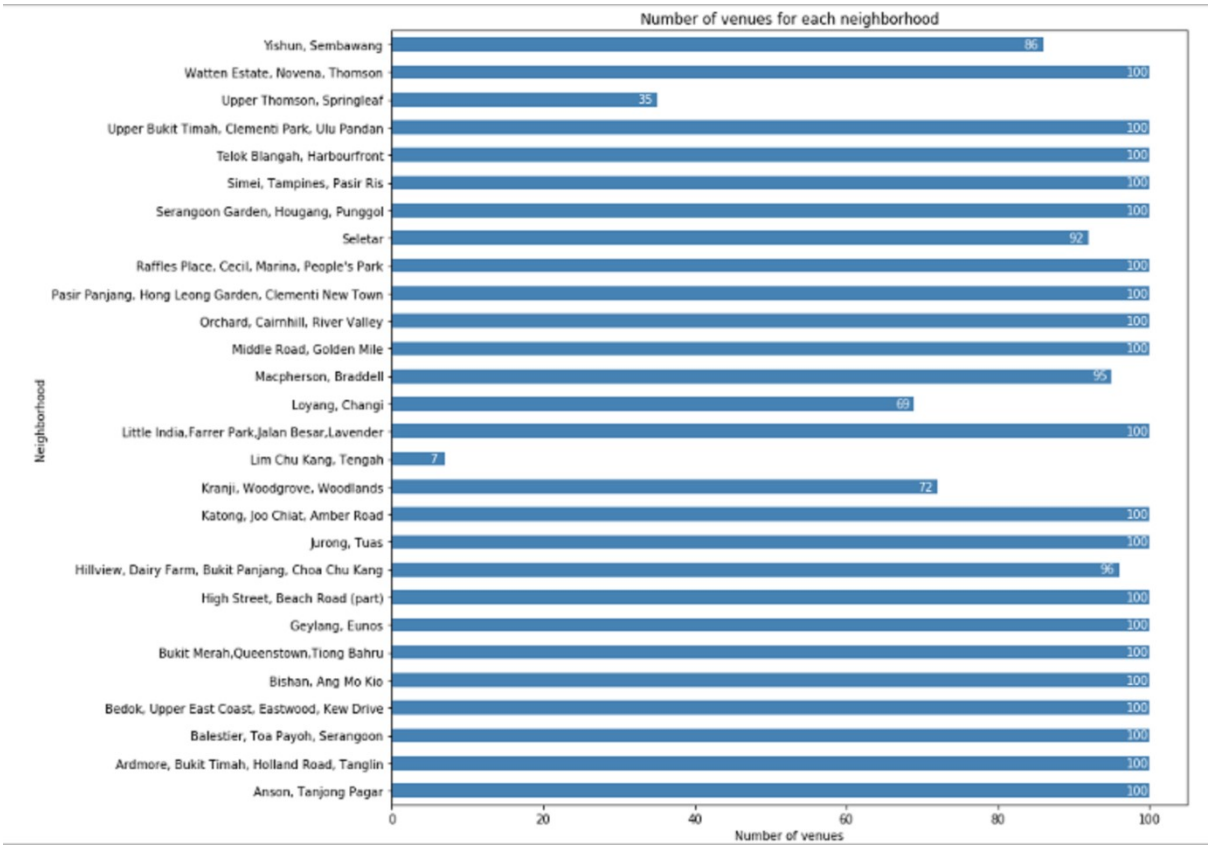
	Postal District	Postal Sector	Neighborhood	Latitude	Longitude	Rental Rates
0	1	01, 02, 03, 04, 05, 06	Raffles Place, Cecil, Marina, People's Park	1.283000	103.851300	5200
1	2	07, 08	Anson, Tanjong Pagar	1.278000	103.840400	5050
2	3	14, 15, 16	Bukit Merah,Queenstown,Tiong Bahru	1.289600	103.816800	4100
3	4	09, 10	Telok Blangah, Harbourfront	1.270700	103.809900	4000
4	5	11, 12, 13	Pasir Panjang, Hong Leong Garden, Clementi New...	1.296600	103.776400	3000
5	6	17	High Street, Beach Road (part)	1.290950	103.848750	5738
6	7	18, 19	Middle Road, Golden Mile	1.299540	103.851940	5000
7	8	20, 21	Little India,Farrer Park,Jalan Besar,Lavender	1.305500	103.855400	3000
8	9	22, 23	Orchard, Cairnhill, River Valley	1.304833	103.831833	4600
9	10	24, 25, 26, 27	Ardmore, Bukit Timah, Holland Road, Tanglin	1.311040	103.804380	3575
10	11	28, 29, 30	Watten Estate, Novena, Thomson	1.325740	103.843030	3350
11	12	31, 32, 33	Balestier, Toa Payoh, Serangoon	1.316360	103.858110	3050
12	13	34, 35, 36, 37	Macpherson, Braddell	1.331760	103.879100	3000

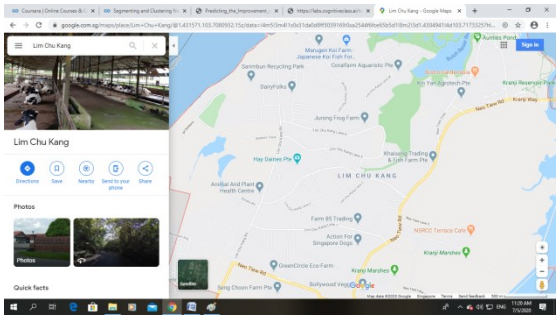
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Raffles Place, Cecil, Marina, People's Park	1.283	103.8513	The Fullerton Bay Hotel	1.283878	103.853314	Hotel
1	Raffles Place, Cecil, Marina, People's Park	1.283	103.8513	CITY Hot Pot Shabu shabu	1.284173	103.851585	Hotpot Restaurant
2	Raffles Place, Cecil, Marina, People's Park	1.283	103.8513	Virgin Active	1.284608	103.850815	Gym / Fitness Center
3	Raffles Place, Cecil, Marina, People's Park	1.283	103.8513	Amoy Hotel	1.283118	103.848539	Hotel
4	Raffles Place, Cecil, Marina, People's Park	1.283	103.8513	Fat Saigon Boy	1.282977	103.849068	Vietnamese Restaurant

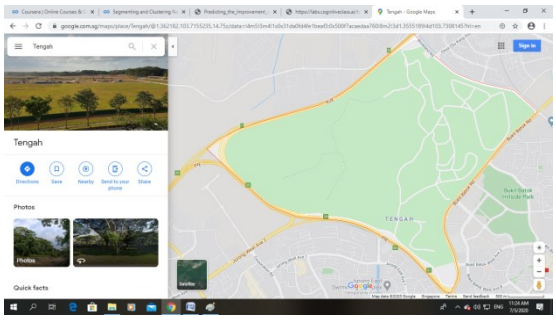
3. Exploratory Data Analysis

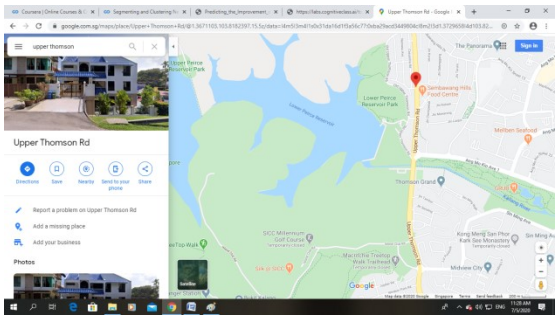
3.1 Descriptive Statistics

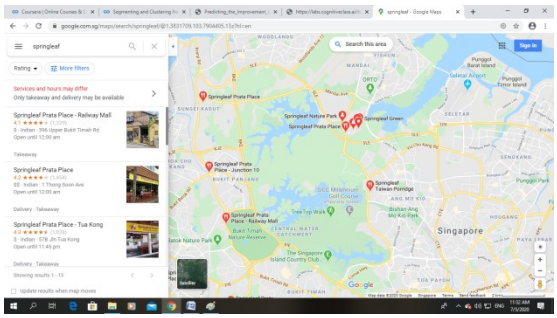
First of all, a bar-chart was generated to display the number of venues that the Foursquare API returned for each neighbourhood. The chart showed that only seven venues were returned for the Lim Chu Kang, Tengah neighbourhood and thirty five for Upper Thomson, Springleaf neighbourhood. Google map was used to verify the results, and it turned out that Lim Chu Kang is filled with farmland and Tengah is just a vacant piece of land. On the other hand, the Upper Thomson, Springleaf neighbourhood is surrounded by reservoirs and golf courses (see pictures below).











Next, the neighbourhoods were segmented by printing out the top five most common venues to see how they differ. A dataframe was then created to display the top ten venues of each neighbourhood.

```
Launcher                    Segmenting and Clustering
Code
----Anson, Tanjong Pagar----
                venue  freq
0               Hotel  0.08
1                Café  0.05
2           Food Court  0.05
3  Japanese Restaurant  0.05
4   Korean Restaurant  0.04


----Ardmore, Bukit Timah, Holland Road, Tanglin----
               venue  freq
0  Chinese Restaurant  0.06
1      Ice Cream Shop  0.05
2              Bakery  0.04
3              Garden  0.04
4   French Restaurant  0.03


----Balestier, Toa Payoh, Serangoon----
               venue  freq
0   Indian Restaurant  0.12
1  Chinese Restaurant  0.07
2              Bakery  0.04
3         Coffee Shop  0.04
4                Café  0.04


----Bedok, Upper East Coast, Eastwood, Kew Drive----
                venue  freq
0          Coffee Shop  0.10
1           Food Court  0.08
2  Japanese Restaurant  0.05
3     Asian Restaurant  0.05
4                Café  0.05


----Bishan, Ang Mo Kio----
                venue  freq
0   Chinese Restaurant  0.13
1           Food Court  0.08
2          Coffee Shop  0.07
3                Café  0.05
```

Launcher | Segmenting and Clustering

Code | Python

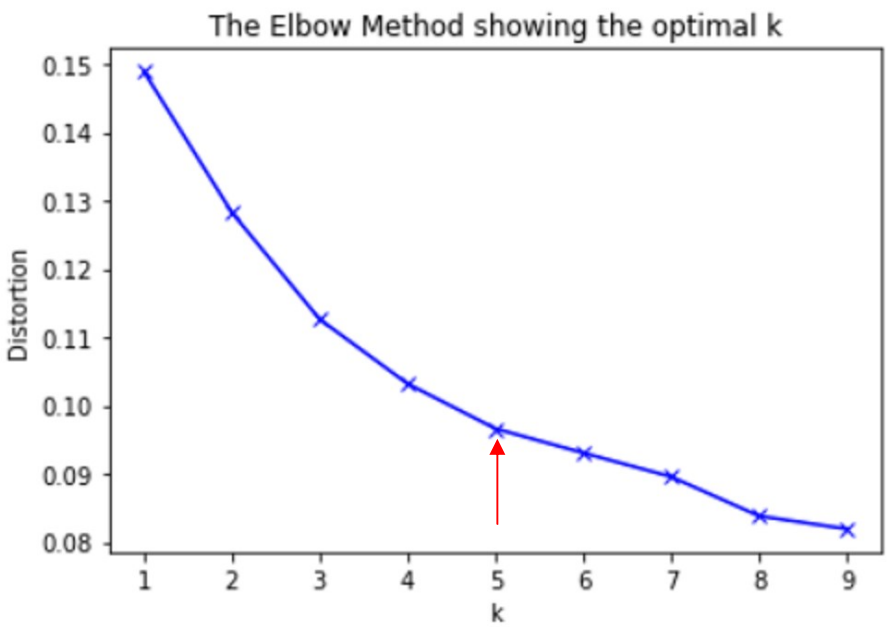
[18]:

	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	Anson, Tanjong Pagar	Hotel	Japanese Restaurant	Food Court	Café	Coffee Shop	Korean Restaurant	Tapas Restaurant	Cocktail Bar	Spanish Restaurant	Mexican Restaurant
1	Ardmore, Bukit Timah, Holland Road, Tanglin	Chinese Restaurant	Ice Cream Shop	Garden	Bakery	Dessert Shop	Asian Restaurant	French Restaurant	Indian Restaurant	Pizza Place	Performing Arts Venue
2	Balestier, Toa Payoh, Serangoon	Indian Restaurant	Chinese Restaurant	Hotel	Coffee Shop	Bakery	Café	Food Court	Hostel	Thai Restaurant	Dessert Shop
3	Bedok, Upper East Coast, Eastwood, Kew Drive	Coffee Shop	Food Court	Café	Japanese Restaurant	Asian Restaurant	Noodle House	Chinese Restaurant	Fast Food Restaurant	Seafood Restaurant	Dessert Shop
4	Bishan, Ang Mo Kio	Chinese Restaurant	Food Court	Coffee Shop	Café	Japanese Restaurant	Park	Thai Restaurant	Ice Cream Shop	Spa	Supermarket
5	Bukit Merah,Queenstown,Tiong Bahru	Café	Park	Food Court	Japanese Restaurant	Coffee Shop	Chinese Restaurant	Hotel	Indian Restaurant	Seafood Restaurant	Bakery
6	Geylang, Eunos	Chinese Restaurant	BBQ Joint	Café	Thai Restaurant	Supermarket	Hostel	Coffee Shop	Food Court	Soup Place	Restaurant
7	High Street, Beach Road (part)	Hotel	Japanese Restaurant	Waterfront	Wine Bar	Cocktail Bar	French Restaurant	Italian Restaurant	Lounge	Shopping Mall	Spanish Restaurant
8	Hillview, Dairy Farm, Bukit Panjang, Choa Chu ...	Trail	Coffee Shop	Nature Preserve	Fast Food Restaurant	Café	Shopping Mall	Italian Restaurant	Sushi Restaurant	Food Court	Park
9	Jurong, Tuas	Fast Food Restaurant	Food Court	Asian Restaurant	Coffee Shop	Japanese Restaurant	Dessert Shop	Chinese Restaurant	Supermarket	Cafe	Seafood Restaurant
10	Katong, Joo Chiat, Amber Road	Chinese Restaurant	Hotel	Bar	Coffee Shop	Multiplex	BBQ Joint	Bakery	Café	Dessert Shop	Dim Sum Restaurant
11	Kranji, Woodgrove, Woodlands	Food Court	Coffee Shop	Asian Restaurant	Japanese Restaurant	Park	Café	Clothing Store	Indian Restaurant	Racetrack	Fast Food Restaurant
12	Lim Chu Kang, Tengah	Farm	Bus Station	Café	Outdoors & Recreation	Wings Joint	Fish & Chips Shop	Flea Market	Flower Shop	Food	Food & Drink Shop
13	Little India,Farrer Park,Jalan Besar,Lavender	Indian Restaurant	Hotel	Chinese Restaurant	Café	Coffee Shop	Vegetarian / Vegan Restaurant	Thai Restaurant	Cocktail Bar	Italian Restaurant	Movie Theater
14	Loyang, Changi	Coffee Shop	Fast Food Restaurant	Gym	Supermarket	Resort	Thai Restaurant	Food Court	Café	Chinese Restaurant	Karaoke Bar
15	Macpherson, Braddell	Food Court	Chinese Restaurant	Coffee Shop	Asian Restaurant	Noodle House	Café	Convenience Store	Seafood Restaurant	Hotel	Indian Restaurant
16	Middle Road, Golden Mile	Hotel	Indian Restaurant	Cocktail Bar	Café	Japanese Restaurant	Shopping Mall	Performing Arts Venue	French Restaurant	Italian Restaurant	Buffet
17	Orchard, Cairnhill, River Valley	Hotel	Japanese Restaurant	Shopping Mall	Department Store	Coffee Shop	Clothing Store	Boutique	Bakery	Chinese Restaurant	Cosmetics Shop
18	Pasir Panjang, Hong Leong Garden, Clementi New...	Coffee Shop	Sandwich Place	Café	Japanese Restaurant	Supermarket	Italian Restaurant	Chinese Restaurant	Indian Restaurant	Food Court	Breakfast Spot
19	Raffles Place, Cecil, Marina, People's Park	Hotel	Waterfront	Coffee Shop	Italian Restaurant	Spanish Restaurant	Salad Place	Plaza	Performing Arts Venue	Shopping Mall	Lounge
20	Seletar	Food Court	Coffee Shop	Café	Supermarket	Fast Food Restaurant	Snack Place	Noodle House	Asian Restaurant	Shopping Mall	Bus Station
21	Serangoon Garden, Hougang, Punggol	Food Court	Coffee Shop	Fast Food Restaurant	Café	Chinese Restaurant	Sandwich Place	Supermarket	Asian Restaurant	Shopping Mall	Trail

3.2 Machine Learning

As it was not possible to identify patterns or trends from the 28 x 10 dataframe (see above) using visual observation, machine learning was used to cluster the neighbourhoods based on the top ten venues. The k-means clustering algorithm was chosen for its simplicity and ability to quickly discover insights from unlabeled data.

Firstly, the optimal value of k was determined using an elbow plot. The value of k at the "elbow" was chosen.



Next, the scikit-learn KMeans algorithm was used to cluster the neighbourhoods with k set to five.

Run k-means to cluster the neighborhood into 5 clusters

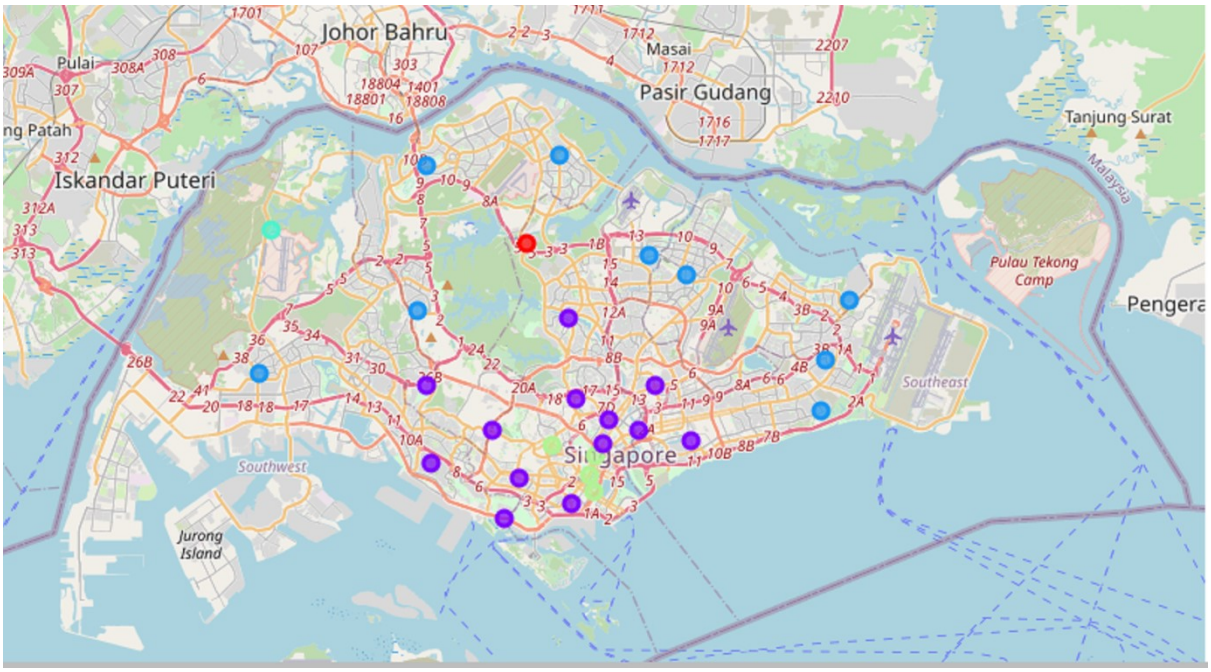
```
# set number of clusters
kclusters = 5

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(singapore_grouped_clustering)

# check cluster labels generated for each row in the dataframe
print(kmeans.labels_)
```

```
[1 1 1 2 1 1 1 4 2 2 1 2 3 1 2 1 4 4 1 4 2 2 2 1 1 0 1 2]
```

4. Results

4.1 Visualization

The resulting clusters were displayed on a Leaflet map of Singapore using the Folium Python library [3]. Five distinct clusters were shown in different colours (see picture below).



4.2 Examine the Clusters

Details of the neighbourhoods belonging to each cluster were printed out for observation. Each cluster was given a descriptive label based on observations made on the data collected (see pictures below).

Cluster 1: A neighborhood with a golf course and lake that is reasonably priced

```
cluster1 = cluster_df.loc[cluster_df['Cluster Labels'] == 0]
cluster1
```

	Postal District	Neighborhood	Rental Rates	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
25	26	Upper Thomson, Springleaf	2125	0	Golf Course	Lake	Chinese Restaurant	Café	Gun Range	Bus Stop	Indian Restaurant	German Restaurant	Smoke Shop	Military Base

Cluster 2: Upmarket neighborhoods with multiple eateries serving a wide range of cuisines

```
cluster2 =cluster_df.loc[singapore_merged['Cluster Labels'] == 1]
cluster2
```

	Postal District	Neighborhood	Rental Rates	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	2	Anson, Tanjong Pagar	5050	1	Hotel	Café	Coffee Shop	Food Court	Japanese Restaurant	Chinese Restaurant	Tapas Restaurant	Spanish Restaurant	Bookstore	Beer Garden
2	3	Bukit Merah,Queenstown,Tiong Bahru	4100	1	Café	Park	Food Court	Japanese Restaurant	Chinese Restaurant	Coffee Shop	Hotel	Indian Restaurant	Seafood Restaurant	Bakery
3	4	Telok Blangah, Harbourfront	4000	1	Chinese Restaurant	Café	Scenic Lookout	Beach	Trail	Bakery	Food Court	Park	Sporting Goods Shop	Coffee Shop
4	5	Pasir Panjang, Hong Leong Garden, Clementi New...	3000	1	Coffee Shop	Sandwich Place	Italian Restaurant	Japanese Restaurant	Café	Chinese Restaurant	Supermarket	Indian Restaurant	Performing Arts Venue	Shopping Mall
7	8	Little India,Farrer Park,Jalan Besar,Lavender	3000	1	Hotel	Indian Restaurant	Café	Chinese Restaurant	Thai Restaurant	Coffee Shop	Bakery	Vegetarian / Vegan Restaurant	Italian Restaurant	Shopping Mall
9	10	Ardmore, Bukit Timah, Holland Road, Tanglin	3575	1	Chinese Restaurant	Japanese Restaurant	Bakery	Ice Cream Shop	Garden	Indian Restaurant	Asian Restaurant	Pizza Place	Dessert Shop	French Restaurant
10	11	Watten Estate, Novena, Thomson	3350	1	Chinese Restaurant	Hotel	Bakery	Café	Dessert Shop	Food Court	Coffee Shop	Italian Restaurant	Supermarket	Flower Shop
11	12	Balestier, Toa Payoh, Serangoon	3050	1	Indian Restaurant	Chinese Restaurant	Hotel	Coffee Shop	Bakery	Thai Restaurant	BBQ Joint	Food Court	Hostel	Noodle House
12	13	Macpherson, Braddell	3000	1	Food Court	Chinese Restaurant	Coffee Shop	Asian Restaurant	Noodle House	Café	Seafood Restaurant	Convenience Store	Fast Food Restaurant	Furniture / Home Store
13	14	Geylang, Eunos	2700	1	Chinese Restaurant	BBQ Joint	Café	Thai Restaurant	Coffee Shop	Supermarket	Hostel	Food Court	Restaurant	Hotel
14	15	Katong, Joo Chiat, Amber Road	3100	1	Chinese Restaurant	Coffee Shop	Hotel	Bar	Yoga Studio	Asian Restaurant	Dessert Shop	Dim Sum Restaurant	Multiplex	Food Court
19	20	Bishan, Ang Mo Kio	3300	1	Chinese Restaurant	Food Court	Coffee Shop	Café	Thai Restaurant	Japanese Restaurant	Park	Ice Cream Shop	Supermarket	Spa
20	21	Upper Bukit Timah, Clementi Park, Ulu Pandan	2350	1	Café	Chinese Restaurant	Korean Restaurant	Bakery	Food Court	Indian Restaurant	Coffee Shop	Thai Restaurant	Seafood Restaurant	Asian Restaurant

Cluster 3: Suburban neighborhoods with many coffee shops and food courts that are affordably priced

```
cluster3 = cluster_df.loc[cluster_df['Cluster Labels'] == 2]
cluster3
```

	Postal District	Neighborhood	Rental Rates	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
15	16	Bedok, Upper East Coast, Eastwood, Kew Drive	2600	2	Coffee Shop	Food Court	Japanese Restaurant	Asian Restaurant	Chinese Restaurant	Café	Noodle House	Malay Restaurant	Beach	Dessert Shop
16	17	Loyang, Changi	2200	2	Coffee Shop	Fast Food Restaurant	Supermarket	Café	Thai Restaurant	Food Court	Resort	Bus Stop	Park	Asian Restaurant
17	18	Simei, Tampines, Pasir Ris	2600	2	Coffee Shop	Japanese Restaurant	Food Court	Café	Sandwich Place	Bakery	Fast Food Restaurant	Gym	Dessert Shop	Noodle House
18	19	Serangoon Garden, Hougang, Punggol	2700	2	Food Court	Fast Food Restaurant	Coffee Shop	Café	Sandwich Place	Supermarket	Trail	Chinese Restaurant	Shopping Mall	Gym
21	22	Jurong, Tuas	2750	2	Fast Food Restaurant	Food Court	Japanese Restaurant	Asian Restaurant	Coffee Shop	Chinese Restaurant	Supermarket	Indian Restaurant	Seafood Restaurant	Shopping Mall
22	23	Hillview, Dairy Farm, Bukit Panjang, Choa Chu ...	2400	2	Trail	Coffee Shop	Fast Food Restaurant	Café	Shopping Mall	Park	Nature Preserve	Italian Restaurant	Supermarket	Food Court
24	25	Kranji, Woodgrove, Woodlands	2100	2	Food Court	Coffee Shop	Asian Restaurant	Japanese Restaurant	Park	Café	Fast Food Restaurant	Sandwich Place	Malay Restaurant	Chinese Restaurant
26	27	Yishun, Sembawang	2400	2	Coffee Shop	Chinese Restaurant	Indian Restaurant	Asian Restaurant	Fried Chicken Joint	Food Court	Thai Restaurant	Sushi Restaurant	Park	Vegetarian / Vegan Restaurant
27	28	Seletar	2700	2	Food Court	Coffee Shop	Café	Supermarket	Bus Station	Noodle House	Asian Restaurant	Fast Food Restaurant	Snack Place	Shopping Mall

Cluster 4: A neighborhood surrounded by farmland

```
cluster4 = cluster_df.loc[cluster_df['Cluster Labels'] == 3]
cluster4
```

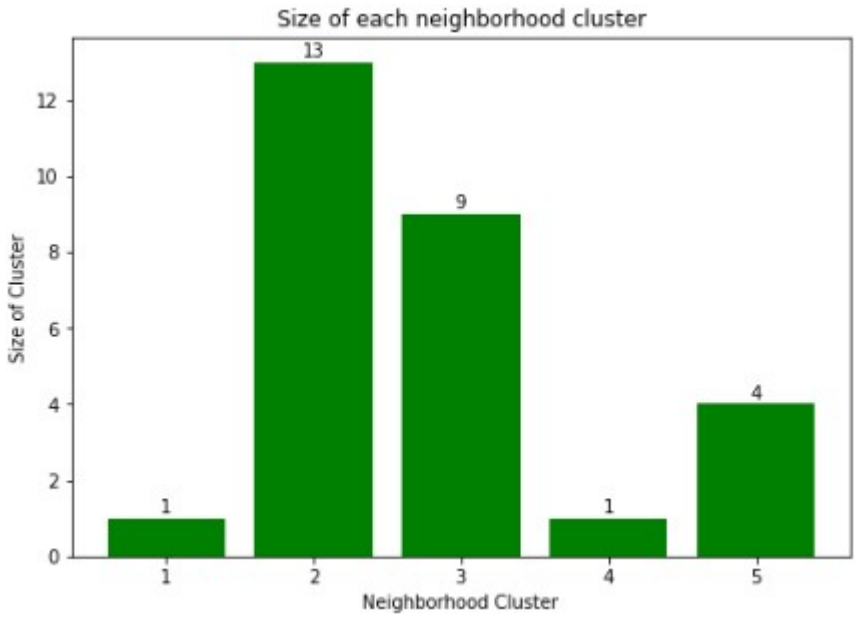
	Postal District	Neighborhood	Rental Rates	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
23	24	Lim Chu Kang, Tengah	2250	3	Farm	Cafeteria	Bus Station	Military Base	Department Store	Dessert Shop	French Restaurant	Football Stadium	Food Truck	Food Stand

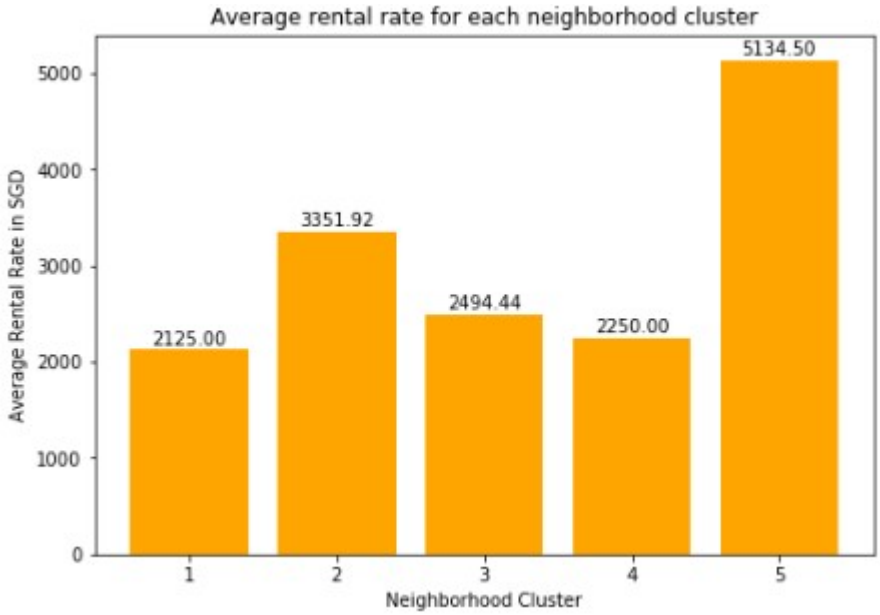
Cluster 5: Neighborhoods near the city centre where high-end hotels and shopping malls are located

```
cluster5 = cluster_df.loc[singapore_merged['Cluster Labels'] == 4]
cluster5
```

	Postal District	Neighborhood	Rental Rates	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	1	Raffles Place, Cecil, Marina, People's Park	5200	4	Hotel	Waterfront	Coffee Shop	Spanish Restaurant	Lounge	Italian Restaurant	Beer Garden	Event Space	Café	Performing Arts Venue
5	6	High Street, Beach Road (part)	5738	4	Hotel	Japanese Restaurant	Waterfront	Cocktail Bar	Wine Bar	Performing Arts Venue	Buffet	Shopping Mall	Plaza	Park
6	7	Middle Road, Golden Mile	5000	4	Hotel	Japanese Restaurant	Café	Cocktail Bar	Shopping Mall	Buffet	Indian Restaurant	Italian Restaurant	French Restaurant	Bakery
8	9	Orchard, Cairnhill, River Valley	4600	4	Hotel	Japanese Restaurant	Shopping Mall	Department Store	Clothing Store	Café	Chinese Restaurant	Bakery	Coffee Shop	Steakhouse

The size and average rental rates of each cluster were also plotted out for analysis.





5. Observations and Recommendations

5.1 Observations

Some observations can be made from the unique characteristics of the different neighbourhood clusters. For instance, cluster 2 is the biggest cluster that is made up of 13 matured estates with multiple eateries serving a wide range of cuisines. Most of the neighbourhoods in this cluster are located at the fringe of the city centre where the average rental rate for a 1000-1100 sq ft apartment is around SGD 3351.92.

The second biggest cluster (3) is made up of 9 suburban or less-matured neighbourhoods that have many coffee shops and food courts selling a wide variety of local food at affordable prices. As all of these neighbourhoods are further from the city centre (>10km), the average rental rate for a 1000-1100 sq ft apartment is around SGD 2494.44.

The third biggest cluster (5) consists of 4 neighbourhoods that make up the city centre where most high-end hotels and shopping malls are located. The average rental rate for a 1000-1100 sq ft apartment is around SGD 5134.54.

It was observed in section 3.1 that the number of venues for the Lim Chu Kang, Tengah neighbourhood (7) and Upper Thomson, Springleaf neighbourhood (35) was small as compared to the rest of the neighbourhoods. It was also observed in section 4.3 that the top two venues of the Lim Chu Kang, Tengah neighbourhood and Upper Thomson, Springleaf neighbourhood were quite different from the rest of the neighbourhoods. The top two venues for the Lim Chu Kang, Tengah neighbourhood were farm and cafeteria, and golf course and lake for the Upper Thomson, Springleaf neighbourhood. This explains why they ended up as single-neighbourhood clusters.

5.2 Recommendations

For visitors who are used to eating at restaurants and can fork out more than SGD3000 per month for a 1000-1100 sq ft apartment should explore neighbourhood cluster 2. For those who are more cost-conscious and are used to eating local food, they can consider neighbourhood cluster 3. When cost isn't a concern and proximity to the city centre is a priority then cluster 5 will be the perfect choice. Visitors who are keen to visit a farm or have a farm stay should explore the Lim Chu Kang, Tengah neighbourhood (cluster 4) where most of Singapore's farms are located. Golfers who want to stay near the golf course and have a scenic lake-view should consider the Upper Thomson, Springleaf neighbourhood (cluster 0) where rental rates are quite affordable.

6. Conclusion

In this study, I have segmented the 28 postal districts of Singapore based on the venue information provided by Foursquare API. I have performed statistical analysis to determine the mean of the frequency of occurrence of each venue category by neighbourhood, and have determined the top 10 venues of each neighbourhood. I have then clustered the neighbourhoods using the k-means machine learning algorithm, and visualised the results using the Folium Python library. Some observations and recommendations were made from the unique characteristics of the different neighbourhood clusters. For example, visitors who are used to eating at restaurants and can fork out more than SGD3000 per month for a 1000-1100 sq ft apartment should explore neighbourhood cluster 2.

7. Future Directions

The observations made in this study can also be used by potential business owners when they are trying to start their businesses in Singapore. For instance, if someone is looking to open a restaurant in Singapore, he might want to consider locating it in neighbourhood cluster 2 where residents are more likely to eat at restaurants regularly. This is inferred from restaurants being the first and second most common venues in cluster 2 and has to be

confirmed by a separate study or survey. A contractor who is trying to start a business to serve the hospitality and retail sectors should consider locating its head office in neighbourhood cluster 5 where majority of high-end hotels and shopping malls are located. However, that will also mean higher operating costs as the rental rates are higher.

References:

[1] https://www.crummy.com/software/BeautifulSoup/bs4/doc/

[2] https://developer.foursquare.com/

[3] https://python-visualization.github.io/folium/