Exploring venues (eateries) in Singapore using Foursquare API

Sanat Madkar

April 20, 2020

1. Introduction

    1.1. Background

    With the rapid urbanization of the 'Lion City' came the additional demand for gastronomic options around the city-state. Over the 80s and the 90s, it was a prime example of a successfully built independent post-colonial state and attracted many migrants from all over South, East and Southeast Asia. This meant that a lot of the food options available in the city are heavily influenced by these countries. Over the turn of the century, the city has transformed itself, making up a GDP of over USD 300 billion, the 8th largest by capita (USD 65,000) out of any country in the world, thus increasing the amount of prospects for potential migrants and tourists likewise. This has resulted in a multitude of choices of eateries one can dine out at. In this analysis, we will be taking a look at some of the neighborhoods located in the Eastern Region of the city-state i.e. neighborhoods located the closest to Changi Airport.

1.2. Target Audience

This analysis is potentially useful for people who are looking to move in the Eastern Region of Singapore and tourists who perhaps may be in Singapore for a few hours on transfer and may prefer to grab a quick bite.

2. Data

The information about the neighborhoods (scope) is taken from Wikipedia and scraped to create the data frame

   https://en.wikipedia.org/wiki/Planning_Areas_of_Singapore

To obtain the data in a CSV (comma separated values) format, the Beautiful Soup library is used to scrape the details out of the above Wikipedia link

To obtain the individual latitudes and longitudes of the neighborhoods, the Geocoder API is used for plotting these on a map

The venue data and its categories (eateries) is obtained by the Foursquare API and this data is used for clustering

3. Methodology and Exploratory Data Analysis

Firstly, the data is scraped from a Wikipedia article – Planning Areas of Singapore. This is then converted to a csv file which forms the main data frame df_singapore. The individual neighborhoods are passed into the Geocoder API to retrieve their respective coordinates and are added to df_singapore. Next, with the help of Folium, these neighborhoods are plotted on an OpenStreetMap template shown below

NOTE: Not all neighborhoods have been displayed accurately as the API does not exactly recognize the exact location of these neighborhoods. Also many of the outlying islands do not consist of civilian settlements or may be exclusively used by the military, hence these areas are exempted. An example is the Pulau Tekong Camp located northeast in the given map.

The Foursquare API collects the neighborhood names and passes them in the API call, retrieving details of all the eateries located in a 2km radius in those neighborhoods.

3.1. Categories

The eateries are sorted according to categories - what type of food they serve: Chinese, Japanese, cafes, bakeries etc.

From the following figure, it is to be noted that there are 83 different categories of eateries in the city-state, with the most common being Chinese restaurants.

Count of eateries of each category
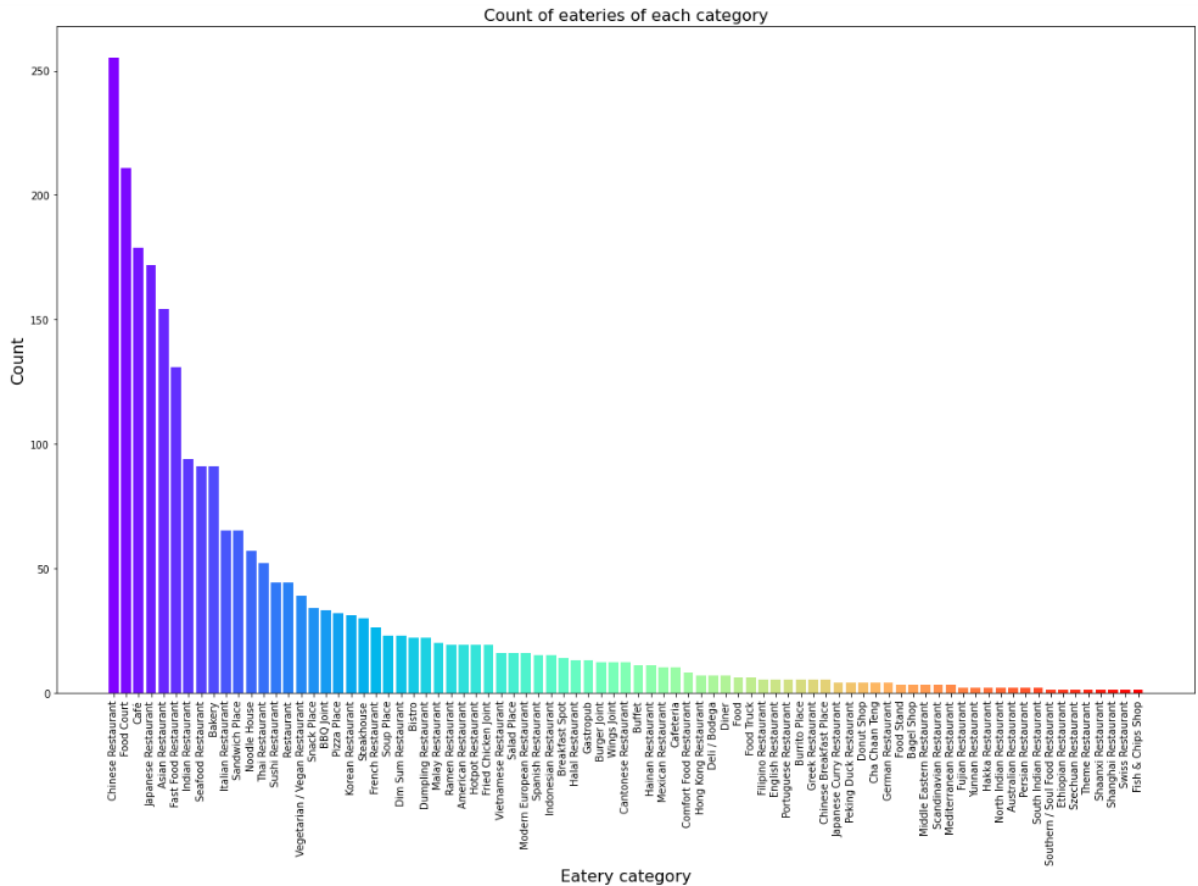
There are 83 different categories of eateries in the city

Because this analysis will be done exclusively in the Eastern region of Singapore, we will only take specific neighborhoods (Bedok, Changi, Changi Bay, Geylang, Marina East, Marine Parade, Pasir Ris, Paya Lebar, Siglap and Tampines). A new data frame is created with eateries belonging to the aforementioned neighborhoods. Below are some of the eateries located in Bedok.

	Neighborhood	Latitude	Longitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
50	Bedok	1.32425	103.95297	58 Minced Meat Mee (Minced Pork Noodle) 58肉脞面 ...	1.324106	103.941222	Noodle House
51	Bedok	1.32425	103.95297	New Mahamoodiya Restaurant	1.331267	103.947671	Indian Restaurant
52	Bedok	1.32425	103.95297	Badoque Cafe	1.331426	103.948504	Café
53	Bedok	1.32425	103.95297	Huggs JPM @ Changi City (huggs coffee)	1.334441	103.964183	Café
54	Bedok	1.32425	103.95297	Hua Yu Wee Restaurant	1.314572	103.941589	Chinese Restaurant
55	Bedok	1.32425	103.95297	The Burning Oak	1.331380	103.948312	BBQ Joint
56	Bedok	1.32425	103.95297	Turo Turo	1.335085	103.964109	Filipino Restaurant
57	Bedok	1.32425	103.95297	Fatty Fong Seafood Resturant	1.325251	103.940759	Seafood Restaurant
58	Bedok	1.32425	103.95297	Changi Lor 108 Fei Lao Seafood	1.332402	103.939209	Asian Restaurant
59	Bedok	1.32425	103.95297	CHAN B.B.Q	1.331829	103.938760	Seafood Restaurant
60	Bedok	1.32425	103.95297	SONG FA Bak Kut Teh	1.335363	103.964323	Chinese Restaurant
61	Bedok	1.32425	103.95297	Bedok Food Centre (Bedok Corner)	1.320303	103.955491	Food Court
62	Bedok	1.32425	103.95297	The Marketplace @ 58	1.324038	103.941145	Food Court
63	Bedok	1.32425	103.95297	Century Village Eating House	1.335003	103.950822	Seafood Restaurant
64	Bedok	1.32425	103.95297	Greenwitch	1.333785	103.964999	Sandwich Place
65	Bedok	1.32425	103.95297	SaladStop!	1.335053	103.964224	Salad Place
66	Bedok	1.32425	103.95297	The Bedok Market Place	1.331325	103.948341	Food Court
67	Bedok	1.32425	103.95297	New Changi Eating House	1.323117	103.945954	Asian Restaurant
68	Bedok	1.32425	103.95297	85 Fengshan Centre	1.331799	103.938691	Food Court
69	Bedok	1.32425	103.95297	YummySoy	1.332614	103.937650	Café
70	Bedok	1.32425	103.95297	Tang Tea House Hong Kong Cafe	1.331714	103.947322	Asian Restaurant
71	Bedok	1.32425	103.95297	Spize	1.331338	103.948198	Malay Restaurant
72	Bedok	1.32425	103.95297	Jalan Tua Kong Lau Lim Mee Pok Kway Teow Mee 老...	1.331205	103.948501	Noodle House
73	Bedok	1.32425	103.95297	Chin Lee Restaurant 深利美食馆	1.331970	103.937156	Chinese Restaurant

All eateries located in Eastern Singapore are then plotted on a Folium map as shown below
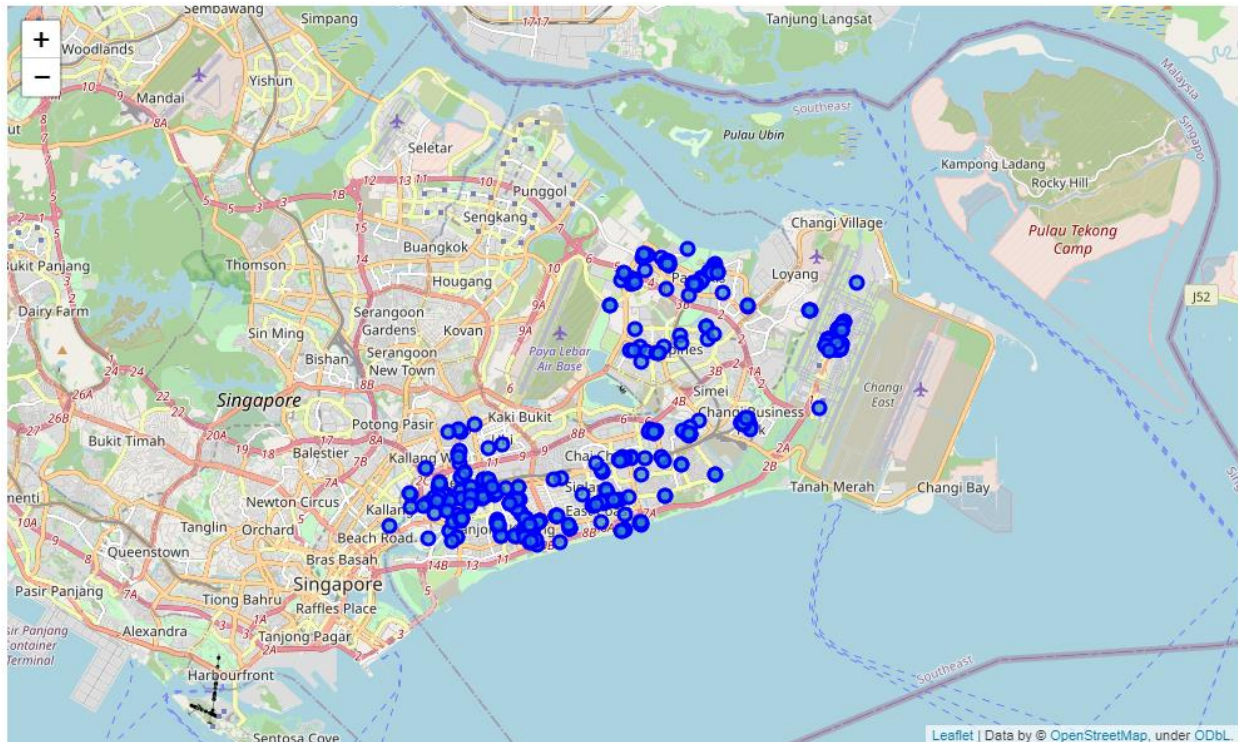
3.2. Clustering

We have made use of one hot encoding to convert given categorical variables into binary vectors as follows

	Neighborhood	American Restaurant	Asian Restaurant	BBQ Joint	Bakery	Bistro	Breakfast Spot	Buffet	Burger Joint	Burrito Place	Cafeteria	Café	Cantonese Restaurant	Chinese Restaurant	Comfort Food Restaurant
0	Bedok	0.00	0.060000	0.04	0.020000	0.00	0.02	0.00	0.00	0.00	0.000000	0.120000	0.00	0.140000	0.00
1	Changi	0.00	0.060000	0.04	0.080000	0.00	0.00	0.00	0.02	0.00	0.000000	0.060000	0.02	0.080000	0.00
2	Changi Bay	0.00	0.080000	0.02	0.080000	0.00	0.00	0.00	0.02	0.00	0.000000	0.060000	0.02	0.060000	0.00
3	Geylang	0.00	0.060000	0.08	0.000000	0.00	0.00	0.00	0.00	0.02	0.000000	0.040000	0.04	0.140000	0.00
4	Marina East	0.00	0.020000	0.02	0.060000	0.04	0.00	0.00	0.00	0.00	0.000000	0.060000	0.02	0.100000	0.00
5	Marine Parade	0.00	0.040000	0.00	0.060000	0.04	0.00	0.00	0.00	0.00	0.000000	0.020000	0.00	0.100000	0.00
6	Pasir Ris	0.00	0.080000	0.00	0.020000	0.00	0.00	0.02	0.02	0.00	0.000000	0.040000	0.00	0.060000	0.02
7	Paya Lebar	0.00	0.100000	0.08	0.000000	0.00	0.00	0.00	0.00	0.02	0.000000	0.060000	0.02	0.160000	0.00
8	Siglap	0.04	0.060000	0.02	0.040000	0.00	0.02	0.00	0.00	0.02	0.000000	0.000000	0.00	0.240000	0.02
9	Tampines	0.00	0.128205	0.00	0.051282	0.00	0.00	0.00	0.00	0.00	0.025641	0.051282	0.00	0.153846	0.00

The following function retrieves top 10 categories of eateries by neighborhood based on the one hot encoded binary vector.

```python
def return_most_common_eat(row, num_top_eat):
    row_categories = row.iloc[1:]
    row_categories_sorted = row_categories.sort_values(ascending=False)

    return row_categories_sorted.index.values[0:num_top_eat]
```

```python
num_top_eat = 10

indicators = ['st', 'nd', 'rd']

columns = ['Neighborhood']
for ind in np.arange(num_top_eat):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

# create a new dataframe
neighborhoods_venues_sorted = pd.DataFrame(columns=columns)
neighborhoods_venues_sorted['Neighborhood'] = sg_east_grouped['Neighborhood']

for ind in np.arange(sg_east_grouped.shape[0]):
    neighborhoods_venues_sorted.iloc[ind, 1:] = return_most_common_eat(sg_east_grouped.iloc[ind, :], num_top_eat)

neighborhoods_venues_sorted
```

	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	Bedok	Chinese Restaurant	Café	Food Court	Noodle House	Japanese Restaurant	Malay Restaurant	Asian Restaurant	Seafood Restaurant	BBQ Joint	Salad Place
1	Changi	Sandwich Place	Snack Place	Bakery	Chinese Restaurant	Japanese Restaurant	Café	Fast Food Restaurant	Asian Restaurant	Donut Shop	Restaurant
2	Changi Bay	Sandwich Place	Snack Place	Asian Restaurant	Japanese Restaurant	Bakery	Chinese Restaurant	Café	Fast Food Restaurant	Food Court	Donut Shop
3	Geylang	Chinese Restaurant	Food Court	Seafood Restaurant	BBQ Joint	Dim Sum Restaurant	Asian Restaurant	Noodle House	Cantonese Restaurant	Soup Place	Vegetarian / Vegan Restaurant
4	Marina East	Indian Restaurant	Chinese Restaurant	Japanese Restaurant	Bakery	Café	Italian Restaurant	Seafood Restaurant	Vietnamese Restaurant	Dim Sum Restaurant	Noodle House
5	Marine Parade	Indian Restaurant	Chinese Restaurant	Japanese Restaurant	Noodle House	Bakery	Italian Restaurant	Bistro	Food Court	Dim Sum Restaurant	Asian Restaurant
6	Pasir Ris	Food Court	Fast Food Restaurant	Asian Restaurant	Chinese Restaurant	Sandwich Place	Seafood Restaurant	Italian Restaurant	Café	Restaurant	Thai Restaurant
7	Paya Lebar	Chinese Restaurant	Food Court	Asian Restaurant	BBQ Joint	Seafood Restaurant	Café	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	Cantonese Restaurant	Noodle House
8	Siglap	Chinese Restaurant	Seafood Restaurant	Asian Restaurant	Food Court	American Restaurant	Wings Joint	Indian Restaurant	Italian Restaurant	Pizza Place	Noodle House
9	Tampines	Chinese Restaurant	Asian Restaurant	Pizza Place	Food Court	Fast Food Restaurant	Bakery	Café	Snack Place	Seafood Restaurant	Indian Restaurant

The given data is now clustered by using k-means clustering where the number of clusters are 3

```python
kclusters = 3

sg_east_grouped_clustering = sg_east_grouped.drop('Neighborhood', 1)

kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(sg_east_grouped_clustering)

kmeans.labels_
```
```
array([0, 2, 2, 0, 1, 1, 2, 0, 0, 2])
```

# 4. Results and Findings

The neighborhoods we divided into clusters as follows

| | Name | 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 | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Bedok | 1.32425 | 103.95297 | 0.0 | Chinese Restaurant | Café | Food Court | Noodle House | Japanese Restaurant | Malay Restaurant | Asian Restaurant | Seafood Restaurant | BBQ Joint | Salad Place |
| 9 | Changi | 1.35514 | 103.99006 | 2.0 | Sandwich Place | Snack Place | Bakery | Chinese Restaurant | Japanese Restaurant | Café | Fast Food Restaurant | Asian Restaurant | Donut Shop | Restaurant |
| 10 | Changi Bay | 1.36996 | 103.99311 | 2.0 | Sandwich Place | Snack Place | Asian Restaurant | Japanese Restaurant | Bakery | Chinese Restaurant | Café | Fast Food Restaurant | Food Court | Donut Shop |
| 13 | Geylang | 1.31147 | 103.88218 | 0.0 | Chinese Restaurant | Food Court | Seafood Restaurant | BBQ Joint | Dim Sum Restaurant | Asian Restaurant | Noodle House | Cantonese Restaurant | Soup Place | Vegetarian / Vegan Restaurant |
| 20 | Marina East | 1.29579 | 103.89544 | 1.0 | Indian Restaurant | Chinese Restaurant | Japanese Restaurant | Bakery | Café | Italian Restaurant | Seafood Restaurant | Vietnamese Restaurant | Dim Sum Restaurant | Noodle House |
| 22 | Marine Parade | 1.30306 | 103.90778 | 1.0 | Indian Restaurant | Chinese Restaurant | Japanese Restaurant | Noodle House | Bakery | Italian Restaurant | Bistro | Food Court | Dim Sum Restaurant | Asian Restaurant |
| 28 | Pasir Ris | 1.37194 | 103.94994 | 2.0 | Food Court | Fast Food Restaurant | Asian Restaurant | Chinese Restaurant | Sandwich Place | Seafood Restaurant | Italian Restaurant | Café | Restaurant | Thai Restaurant |
| 29 | Paya Lebar | 1.32503 | 103.89049 | 0.0 | Chinese Restaurant | Food Court | Asian Restaurant | BBQ Joint | Seafood Restaurant | Café | Vegetarian / Vegan Restaurant | Vietnamese Restaurant | Cantonese Restaurant | Noodle House |
| 39 | Siglap | 1.31059 | 103.92540 | 0.0 | Chinese Restaurant | Seafood Restaurant | Asian Restaurant | Food Court | American Restaurant | Wings Joint | Indian Restaurant | Italian Restaurant | Pizza Place | Noodle House |
| 45 | Tampines | 1.36819 | 103.92948 | 2.0 | Chinese Restaurant | Asian Restaurant | Pizza Place | Food Court | Fast Food Restaurant | Bakery | Café | Snack Place | Seafood Restaurant | Indian Restaurant |

There are 3 clusters: 0, 1 and 2. These are marked on a Folium map as seen below.
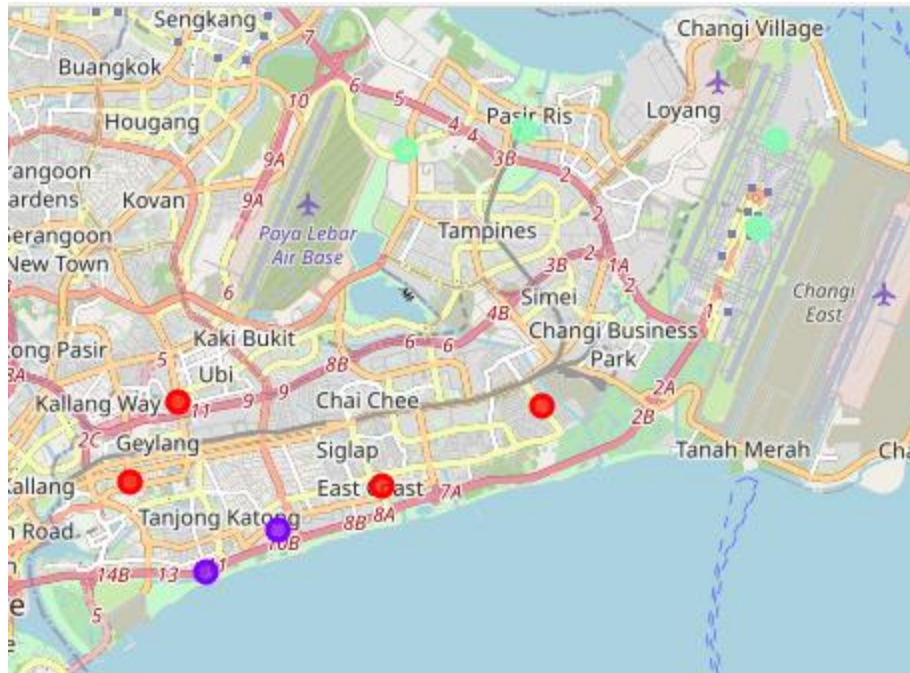
There is a significant presence of cluster 2 (light blue) in and around the Airport area. Cluster 0 (red) is located quite close to the rail (MRT) line. Cluster 1 (violet) is mainly concentrated in one region

| | Name | 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 | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Bedok | 1.32425 | 103.95297 | 0.0 | Chinese Restaurant | Café | Food Court | Noodle House | Japanese Restaurant | Malay Restaurant | Asian Restaurant | Seafood Restaurant | BBQ Joint | Salad Place |
| 13 | Geylang | 1.31147 | 103.88218 | 0.0 | Chinese Restaurant | Food Court | Seafood Restaurant | BBQ Joint | Dim Sum Restaurant | Asian Restaurant | Noodle House | Cantonese Restaurant | Soup Place | Vegetarian / Vegan Restaurant |
| 29 | Paya Lebar | 1.32503 | 103.89049 | 0.0 | Chinese Restaurant | Food Court | Asian Restaurant | BBQ Joint | Seafood Restaurant | Café | Vegetarian / Vegan Restaurant | Vietnamese Restaurant | Cantonese Restaurant | Noodle House |
| 39 | Siglap | 1.31059 | 103.92540 | 0.0 | Chinese Restaurant | Seafood Restaurant | Asian Restaurant | Food Court | American Restaurant | Wings Joint | Indian Restaurant | Italian Restaurant | Pizza Place | Noodle House |

Cluster 0 (above) consists of the suburbs of Bedok, Geylang, Siglap and Paya Lebar, which are mainly dominated by public housing and residences, explaining the high presence of food courts (hawker centres)

and Chinese restaurants in the area

| | Name | 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 | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | Marina East | 1.29579 | 103.89544 | 1.0 | Indian Restaurant | Chinese Restaurant | Japanese Restaurant | Bakery | Café | Italian Restaurant | Seafood Restaurant | Vietnamese Restaurant | Dim Sum Restaurant | Noodle House |
| 22 | Marine Parade | 1.30306 | 103.90778 | 1.0 | Indian Restaurant | Chinese Restaurant | Japanese Restaurant | Noodle House | Bakery | Italian Restaurant | Bistro | Food Court | Dim Sum Restaurant | Asian Restaurant |

Cluster 1 (above) consists of Marine Parade and Marina East areas, primarily consisting of shopping malls and other retail zones; various categories of eateries are prevalent here

| | Name | 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 | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | Changi | 1.35514 | 103.99006 | 2.0 | Sandwich Place | Snack Place | Bakery | Chinese Restaurant | Japanese Restaurant | Café | Fast Food Restaurant | Asian Restaurant | Donut Shop | Restaurant |
| 10 | Changi Bay | 1.36996 | 103.99311 | 2.0 | Sandwich Place | Snack Place | Asian Restaurant | Japanese Restaurant | Bakery | Chinese Restaurant | Café | Fast Food Restaurant | Food Court | Donut Shop |
| 28 | Pasir Ris | 1.37194 | 103.94994 | 2.0 | Food Court | Fast Food Restaurant | Asian Restaurant | Chinese Restaurant | Sandwich Place | Seafood Restaurant | Italian Restaurant | Café | Restaurant | Thai Restaurant |
| 45 | Tampines | 1.36819 | 103.92948 | 2.0 | Chinese Restaurant | Asian Restaurant | Pizza Place | Food Court | Fast Food Restaurant | Bakery | Café | Snack Place | Seafood Restaurant | Indian Restaurant |

Cluster 2 (above) consists of the suburbs located the closest to the Airport; fast-food is prevalent here more than the other suburbs

## 5.  Conclusion

By the analysis, it can be concluded that there is a demographic distribution in the number of eatery types that dominate each neighborhood. The neighborhoods located the closest to the airport are dominated by fast-food joints or short eats, the neighborhoods located close to public housing are dominated by economical food courts and Chinese restaurants and the shopping district has varied options of eateries located in its precinct. This data can be used by new residents and tourists likewise to settle or plan their stay in these areas respectively.