

# **Battle of Singaporean Neighborhoods**

## **The Best Location for New Thai Restaurant in Singapore**

Topp Nattawat Theeralerttham

3 January 2021

### **1. Introduction**

Lord Anantara is an ambitious food entrepreneur from Thailand. He is planning to open his first restaurant overseas in Singapore. He is looking for a location with high traffic of customers and low competition of restaurants. The inconvenient truth is locations with high demand are often already crowded by existing restaurants with cutthroat competition. Location with low competition that still offers high demand is a gem and is a key success factor for Lord Anantara's restaurant. The challenge is finding it. Lord Anantara decided to hire Topp, a charming data scientist, to conduct a location analysis and recommend him the best location. Topp is tasked to use his data science prowess to generate the most promising neighborhood option(s) that can satisfy Lord Anantara's criteria, along with a clear explanation of their advantages and disadvantages.

In this report, we will detail the process Topp carried out to find the best food store location for Lord Anantara.

### **2. Data**

Based on definition of our problem, factors that will influence our decision are:

- total number of restaurants in the neighborhood, which is a proxy for food demand.
- most common types of restaurants in the neighborhood, as their numbers reflect the level of supply and general competition in food business.
- number of existing Thai restaurants in the neighborhood, which reflects the specific competition for Thai cuisine.

Following data sources will be needed to extract and/or generate the required information:

- list of candidate neighborhoods will be scraped from Wikipedia
- coordinates of each neighborhood will be generated using arcGIS API
- information on restaurants in each neighborhood will be obtained using Foursquare API

Using these datasets, our charming data scientist Topp will explore, segment, and cluster the different neighborhoods in Singapore. He will scrape a list of candidate

neighborhoods from Wikipedia, extract their coordinates from arcGIS, and use Foursquare API to explore the different neighborhoods in the city of Singapore. Specifically, he will analyze the total number of restaurants, the most common types of restaurants and the number of existing Thai restaurants for each neighborhood. Then, he will use this analysis to group the neighborhoods into clusters with K-means clustering algorithm. Also, he will use Folium library to visualize the neighborhoods and their clusters. Finally, he will make a recommendation as to where would be the best store location for Lord Anantara in Singapore.

### 3. Methodology

In this project, we were tasked to identify neighborhoods in Singapore that have high demand and low supply (competition). Our role was to make a thorough analysis of the entire country of Singapore and find the 'hidden gem' for Lord Anantara.

In the first step we collected data list and location values of neighborhoods. We also obtained data on different types of restaurants within 500 meters from each neighborhood center (according to Foursquare categorization).

*Figure 1: List of neighborhoods in Singapore on Wikipedia*

#### Postal districts [\[ edit \]](#)

This table lists the postal districts:<sup>[2]</sup>

Postal district	Postal sector (1st 2 digits of 6-digit postal codes)	General location
01	01, 02, 03, 04, 05, 06	Raffles Place, Cecil, Marina, People's Park
02	07, 08	Anson, Tanjong Pagar
03	14, 15, 16	<a href="#">Bukit Merah</a> , <a href="#">Queenstown</a> , <a href="#">Tiong Bahru</a>
04	09, 10	Telok Blangah, Harbourfront
05	11, 12, 13	Pasir Panjang, Hong Leong Garden, Clementi New Town
06	17	High Street, Beach Road (part)
07	18, 19	Middle Road, Golden Mile
08	20, 21	<a href="#">Little India</a> , <a href="#">Farrer Park</a> , <a href="#">Jalan Besar</a> , <a href="#">Lavender</a>
09	22, 23	Orchard, Cairnhill, River Valley
10	24, 25, 26, 27	Ardmore, <a href="#">Bukit Timah</a> , Holland Road, Tanglin
11	28, 29, 30	Watten Estate, <a href="#">Novena</a> , Thomson
12	31, 32, 33	<a href="#">Balestier</a> , <a href="#">Toa Payoh</a> , Serangoon
13	34, 35, 36, 37	Macpherson, Braddell
14	38, 39, 40, 41	<a href="#">Geylang</a> , Eunos, Aljunied
15	42, 43, 44, 45	Katong, Joo Chiat, Amber Road

Figure 2: Map visualization of neighborhoods with coordinates from arcGIS API

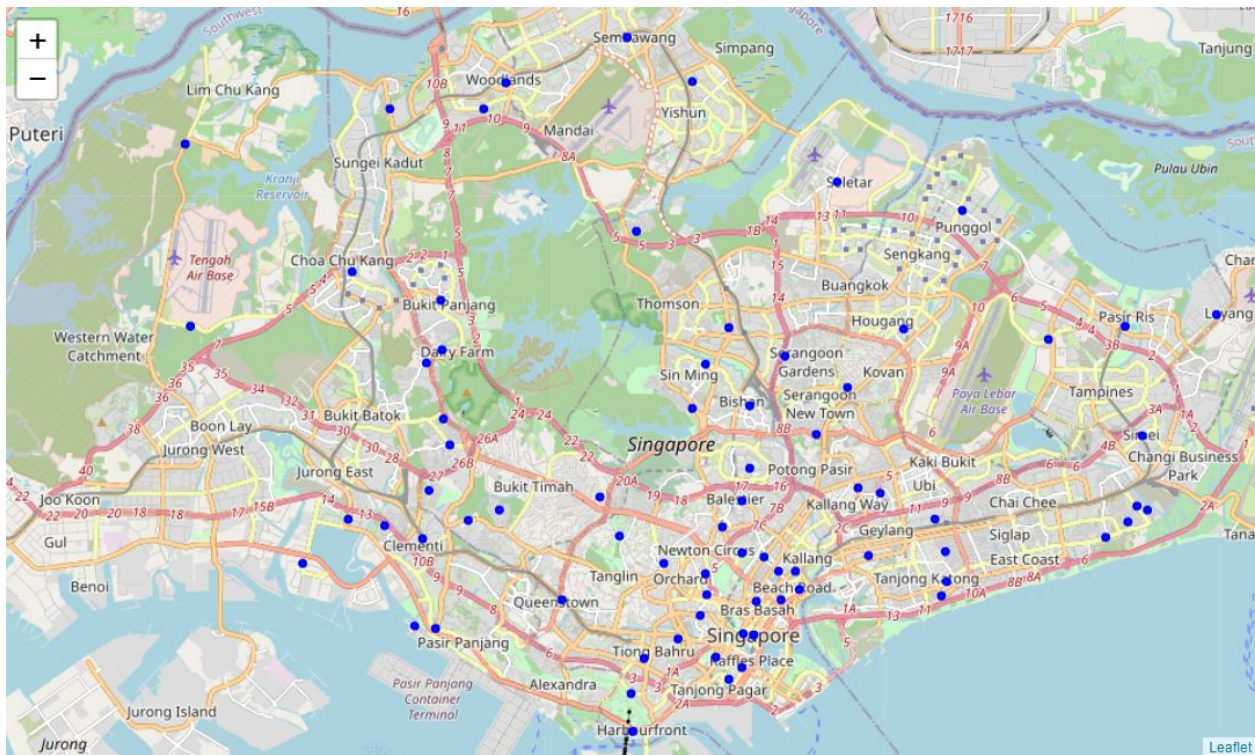


Figure 3: List of restaurants and their types in each neighborhood from Foursquare API

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Raffles Place	1.28189	103.84912	Luke's Oyster Bar & Chop House	1.282459	103.847240	Seafood Restaurant
1	Raffles Place	1.28189	103.84912	Grain Traders	1.281468	103.850039	Restaurant
2	Raffles Place	1.28189	103.84912	Wang Dae Bak Korean BBQ	1.282901	103.848116	Korean Restaurant
3	Raffles Place	1.28189	103.84912	Magal BBQ 마포갈매기	1.281299	103.847932	Korean Restaurant
4	Raffles Place	1.28189	103.84912	왕대박 Wang Dae Bak Korean BBQ Restaurant	1.281345	103.847551	Korean Restaurant
...	...	...	...	...	...	...	...
891	Sembawang	1.44794	103.81891	White Restaurant	1.448468	103.819869	Chinese Restaurant
892	Sembawang	1.44794	103.81891	McDonald's	1.448552	103.819927	Fast Food Restaurant
893	Sembawang	1.44794	103.81891	KFC	1.447950	103.818971	Fast Food Restaurant
894	Sembawang	1.44794	103.81891	Sushi Express	1.448118	103.819348	Japanese Restaurant
895	Sembawang	1.44794	103.81891	Ananas Cafe	1.448861	103.819900	Asian Restaurant

896 rows × 7 columns

Second step in to analyze the \*\*numbers of different types of restaurants for each neighborhood\*\* with a particular attention to Thai restaurants, and also identify the \*\*most common types of restaurants\*\* in all areas of Singapore.

*Figure 4: Numbers of different types of restaurants in each neighborhood*

	Neighborhood	American Restaurant	Asian Restaurant	Australian Restaurant	Cantonese Restaurant	Chinese Restaurant	Comfort Food Restaurant	Dim Sum Restaurant	Dumpling Restaurant	Eastern European Restaurant
0	Aljunied	0	1	0	0	1	0	0	0	0
1	Amber Road	0	1	1	0	4	0	1	1	0
2	Ang Mo Kio	1	2	0	0	1	0	0	0	0
3	Beach Road (part)	0	1	0	0	2	0	0	0	0
4	Braddell	0	2	0	0	1	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...
59	Telok Blangah	0	1	0	0	3	0	0	0	0
60	Upper Bukit Timah	0	0	1	0	0	0	0	0	0
61	Upper Thomson	0	1	0	0	6	0	0	0	0
62	Watten Estate	0	3	0	0	0	0	0	0	0
63	Yishun	0	1	0	0	0	0	0	0	0

64 rows × 56 columns

*Figure 5: Most common types of restaurants in each neighborhood*

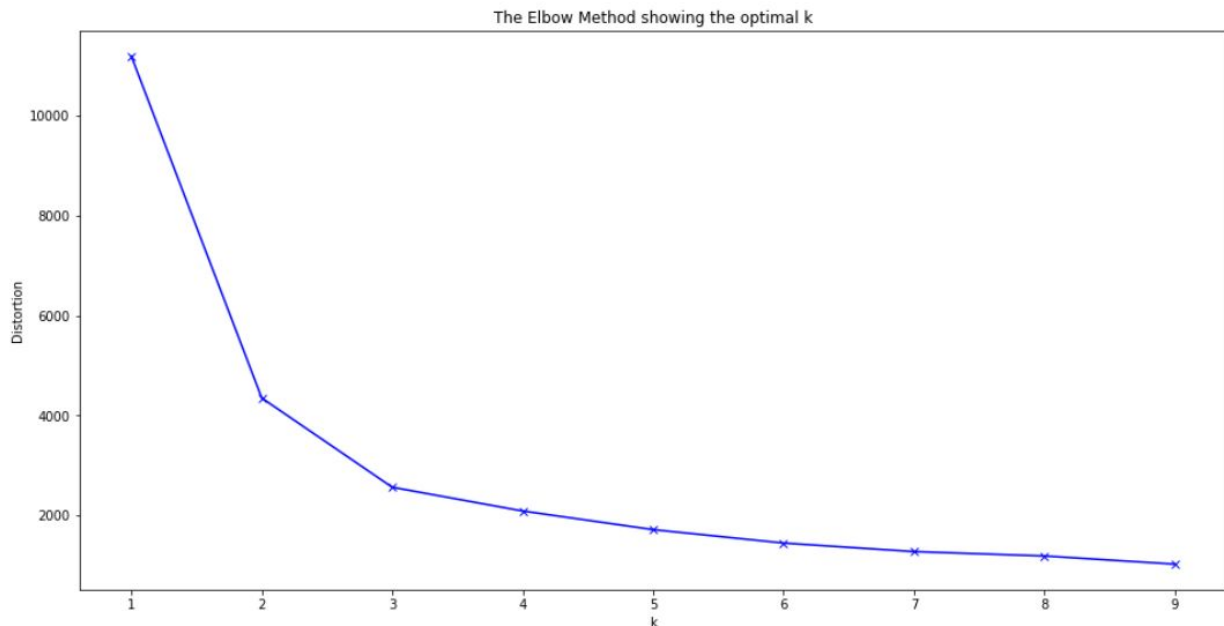
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	Number of Thai Restaurant	Total Number of Restaurants
0	Aljunied	Asian Restaurant	Chinese Restaurant	Indian Restaurant	0	8
1	Amber Road	Chinese Restaurant	Indian Restaurant	Japanese Restaurant	1	38
2	Ang Mo Kio	Fast Food Restaurant	Asian Restaurant	Japanese Restaurant	0	36
3	Beach Road (part)	Japanese Restaurant	Restaurant	Italian Restaurant	3	70
4	Braddell	Asian Restaurant	Seafood Restaurant	Chinese Restaurant	0	10
...	...	...	...	...	...	...
59	Telok Blangah	Chinese Restaurant	Asian Restaurant	Hong Kong Restaurant	0	12
60	Upper Bukit Timah	Australian Restaurant	Yunnan Restaurant	Fujian Restaurant	0	2
61	Upper Thomson	Chinese Restaurant	Thai Restaurant	Indian Restaurant	3	34
62	Watten Estate	Asian Restaurant	Thai Restaurant	Malay Restaurant	3	30
63	Yishun	Fast Food Restaurant	Vegetarian / Vegan Restaurant	Asian Restaurant	1	12

64 rows × 6 columns



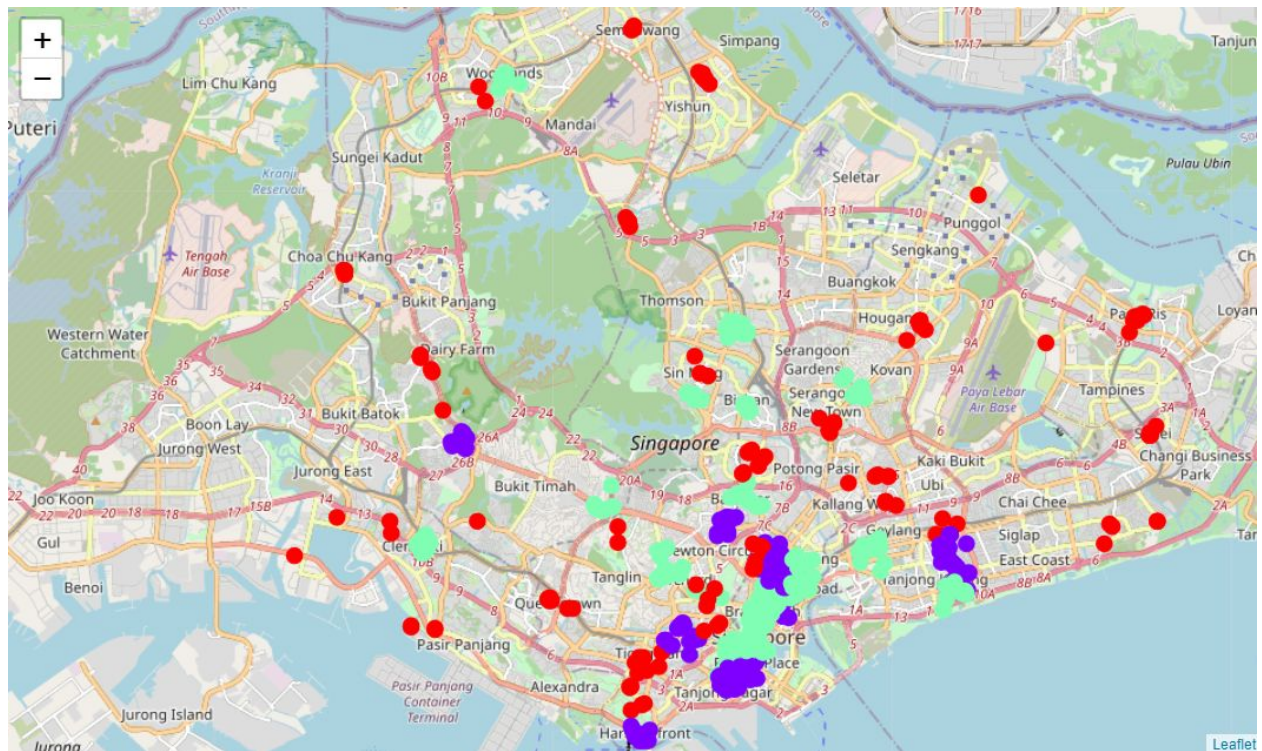
In the third, we used the above analysis to group all areas into clusters with **K-means clustering algorithm**. The key component of this algorithm is the value of  $k$ . We can perform elbow method to find the best value of  $k$ . This method calculates the sum of squared distances from each point to its assigned center (distortions). When the distortions are plotted and the plot looks like an arm then the 'elbow' (the point of inflection on the curve) is the best value of  $k$ .

*Figure 6: Elbow method showing optimal  $k$  for K-Means algorithm*



We can observe that the 'elbow' is the number 3 which is optimal for this case. Now we can run a K-Means using as `n_clusters` the number 3. Then, we used **Folium map visualization** library to visualize the neighborhoods and their clusters. We used the resulting clusters to identify the most promising neighborhoods.

Figure 7: Map visualization showing 3 clusters of neighborhoods



Finally, we discussed the resulting options of neighborhoods against Lord Anantara's criteria and made a recommendation as to where would be the best store location in Singapore.

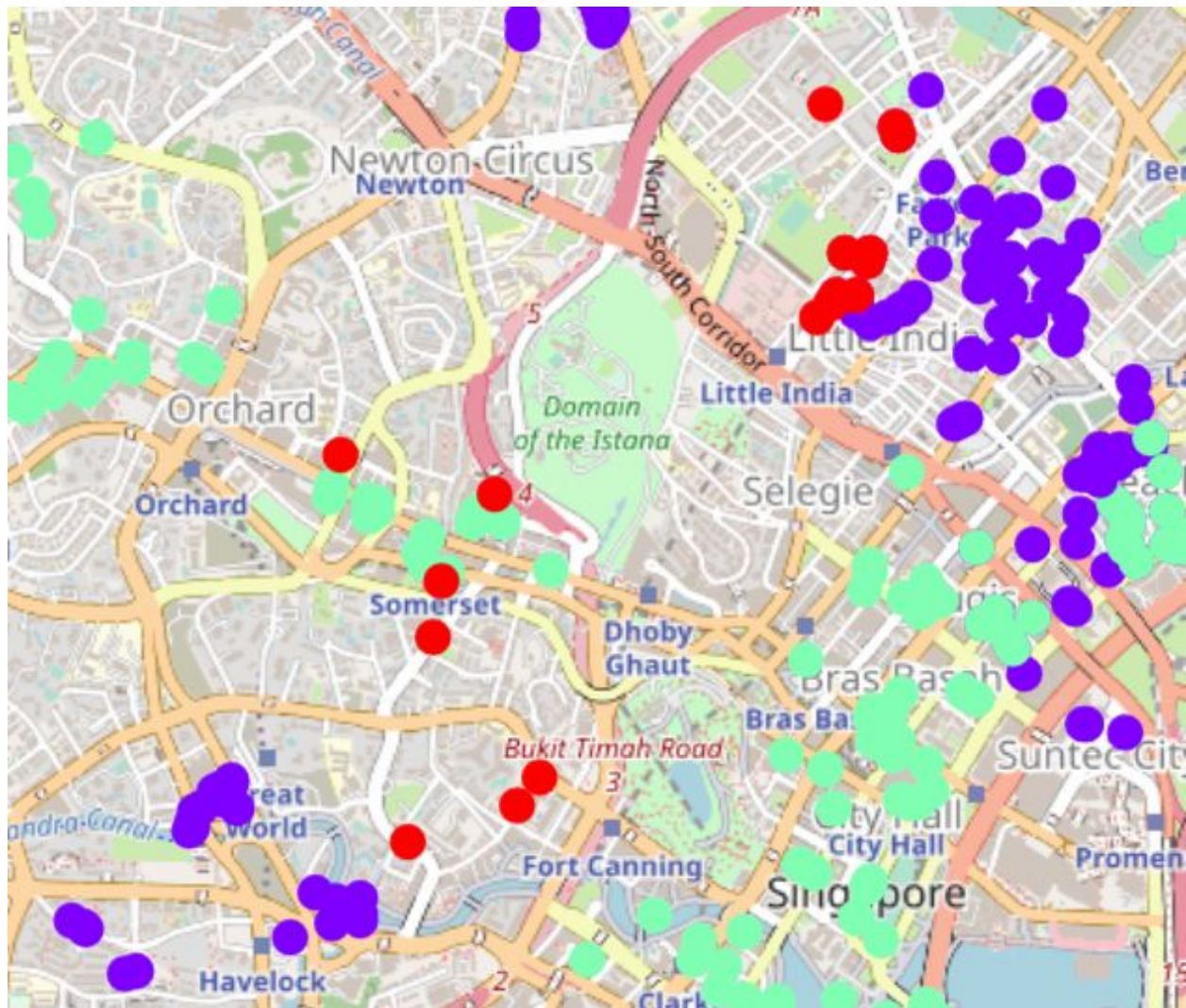
#### 4. Result & Discussion

Based on the analysis, we have obtained 3 clusters of neighborhoods as followed.

- Cluster 1: Low-to-moderate demand & low supply and competition
- Cluster 2: Low-to-moderate demand & moderate-to-high supply and competition
- Cluster 3: Moderate-to-high demand & low-to-moderate supply and competition

Let's examine each of the 3 clusters in detail with a particular attention to the discriminating features that distinguish them.

Cluster 1 (Red): Low-to-moderate demand & low supply and competition



The first cluster is characterized by a low number of Thai restaurants. Only 5 out of the 34 neighborhoods in this cluster have a Thai restaurant and the number of restaurants in a neighborhood is only 1–2.

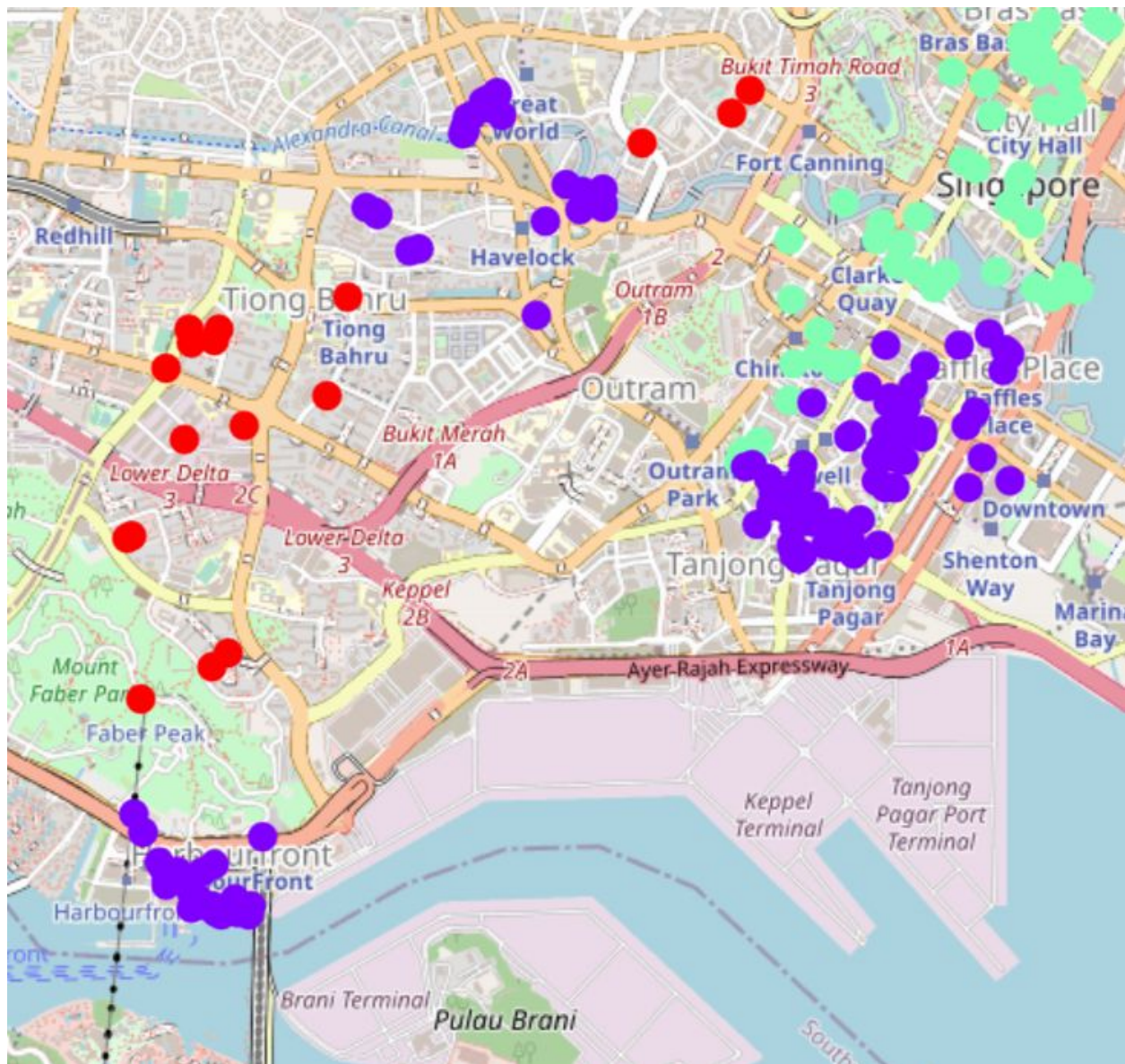
This is partly explained by the relatively low number of restaurants in these neighborhoods in general. The average number of restaurants in a neighborhood is only 5. Most of these neighborhoods are far from the city centers and leisure destinations such as Marina Bay, Clarke Quay, and Orchard. So the demand and traffic of people are expected to be generally low.

There is nonetheless an exception. There are a few neighborhoods in this cluster that are very close to, if not in the middle of, city centers and leisure destinations such as



Cairnhill neighborhood (Orchard) and River Valley neighborhood (Clarke Quay). Furthermore, Farrer Park (Little India) turns out to have a moderately high number of restaurants, 11 to be exact, but no existing Thai restaurant is found.

Cluster 2 (Purple): Low-to-moderate demand & moderate-to-high supply and competition



The second cluster is characterized by a low number of Thai restaurants with a high number of restaurants in general. Out of 11 neighborhoods, 9 already have at least 1 Thai restaurant. In addition, half of them have 2 or more Thai restaurants in the neighborhood. In addition, these neighborhoods distinctively have a high number of

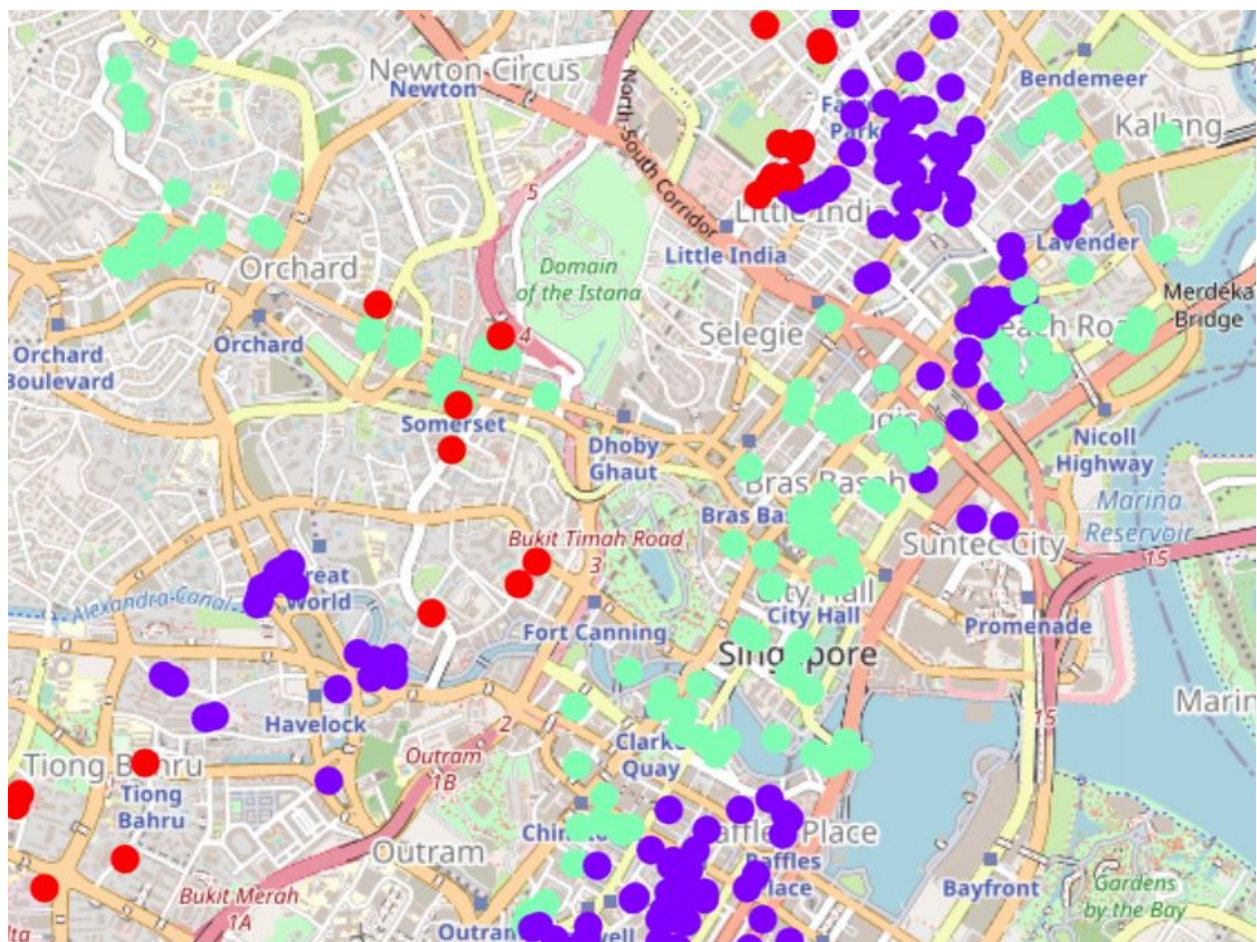


restaurants in general with an average of 35 and it can go up to 47. This concentration is the highest among the three clusters.

Looking at the map visualization, we can see that the neighborhoods in this cluster are closest to the most buzzing part of Singapore. For instance, Tanjong Pagar neighborhood (Chinatown), Raffle Place neighborhood (financial district), Harbourfront (Sentosa Island), and Beach Road neighborhood (Marina Bay). Fierce competition among restaurants is expected, along with possibly the highest rental rate.

Virtually all of these neighborhoods are in close proximity to Marina Bay, a must-see tourist attraction, as well as the financial districts of Singapore. Likewise, Harbourfront neighborhood is the gateway to Sentosa Island, one of Asia's most famous casino destinations and family holiday spots.

Cluster 3 (Green): Moderate-to-high demand & low-to-moderate supply and competition



The third cluster is characterized by a moderate competition for restaurants in general with moderate-to-low competition for Thai restaurants. More than half have only 1 Thai restaurant or none. For a total of 19 neighborhoods, 7 do not have any Thai restaurant yet. Supply and competition of Thai restaurants in this cluster is moderate, if not low. An apparent outlier is Golden Miles neighborhood which is also known as 'Thai Town'.

Demand for Thai food is expected to be high, extrapolating from the high number of restaurants in general as well as the concentration of Asian restaurants in many of these neighborhoods. This is explained by the fact that these neighborhoods are close to People's Park neighborhood and High Street neighborhood which are the spinal cord of Chinatown. A number of neighborhoods are also close to City Hall and Orchard which are the government and shopping district of Singapore respectively.

One neighborhood worthy of special attention is Clementi New Town. This neighborhood is in the vicinity of a large university town of the National University of Singapore as well as a high concentration of world's acclaimed higher education institutions like INSEAD, Duke (Asia campus), and Yale (Asia campus).

## **5. Conclusion**

The purpose of this report is to identify the best location for Lord Anantara's first Thai restaurant in Singapore.

The most strategic location for Lord Anantara is cluster 3 which displays moderate-to-high demand & low-to-moderate supply and competition. Neighborhoods. One of the most promising locations in this cluster is Clementi New Town which is in close proximity to many large and famous university campuses. Other attraction options are Anson (City Hall) and Orchard (Orchard) which boast the government and shopping districts of Singapore. Supply and competition of Thai restaurants in this cluster is moderate-to-low. This is a golden opportunity for Lord Anantara's!

Another strategic location is cluster 2. While neighborhoods in this cluster generally have low traffic and demand for Thai food. A few neighborhoods are an exception. Neighborhoods such as Cairnhill (Orchard), River Valley (Clarke Quay), and Farrer Park (Little India) are very close to the city centers and leisure destinations and promise high traffic of customers. One indicator is the moderately high number of restaurants, 11 to be exact. Interestingly, no existing Thai restaurant is found. Lord Anantara's restaurant will be the first!