

Open a Kitchen Supply Store in Penang

Capstone Project - Battle of Neighborhoods

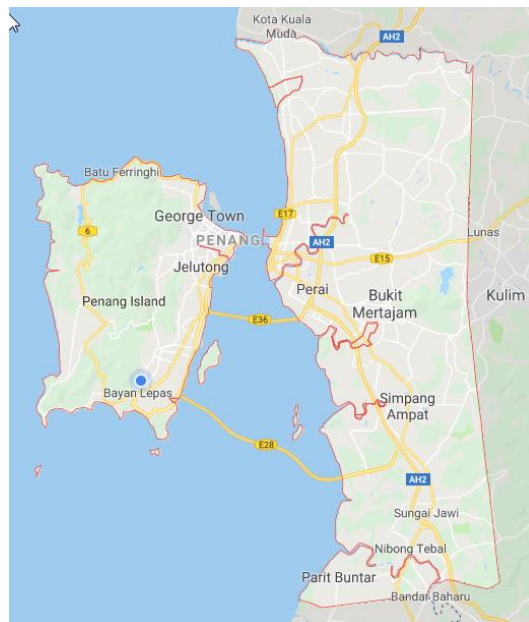
Wei Chong Tin

18 September 2019

1. Introduction

Background/Business Problem

Penang is one of the Malaysia states located on the northwest coast of Peninsular Malaysia. It has two parts: Penang Island and Seberang Perai (formerly Province Wellesley) on the Malay Peninsula.



Penang is famous for its

- UNESCO World Heritage Site and street foods at George Town
- Silicon Valley of the East for its industries in Bayan Lepas
- Beaches on the north of Penang Island - Batu Ferringhi
- Durian (King of the Fruit) at the west of Penang Island - Balik Pulau

There are many food venues (restaurants, food courts, coffee shops, food trucks) in Penang. To open a Food & Beverage (F&B) shop in Penang means that you have to face fierce competition from these food and drinks providers. Instead of compete with them, you can open a shop to supply following items to them:

- Kitchen utensils (knife, cooking pots, cooking pans, etc.)
- Dining supplies (plates, bowls, glasses, spoons, etc.)
- Raw materials (beef, pork, chicken meat, vegetables, etc.)

The location of a new kitchen supply store will decide if the store will success by getting enough customers with less or no competitor.

Business Questions

If a business owner is looking to open a kitchen supply store in Penang, where would you recommend them to open it?

Objective

To analyze and select the best location in the Penang to open a new kitchen supply store.

Target Audience

Business owners who are looking to open a new kitchen supply store in Penang.

2. Data Acquisition

Data Required

- A list of neighborhoods in Penang and their latitude and longitude coordinates.
- A list of food venues with the following data:
 - Venue name
 - Venue category
 - Latitude and longitude coordinates
 - Neighborhood
- A list of competitors - shops that provide kitchen supplies (kitchen supply stores, supermarkets) with the following data:
 - Venue name
 - Venue category
 - Latitude and longitude coordinates
 - Neighborhood

This study exclude the shops/venues that sell raw material of foods, large kitchen appliances/equipment and shops that build kitchen cabinets.

Source of Data

Penang neighborhood list can be extracted from Wikipedia page ([https://en.wikipedia.org/wiki/Category:Districts of Penang](https://en.wikipedia.org/wiki/Category:Districts_of_Penang)). The neighborhood latitude and longitude coordinates can be found using Python Nominatim geocoder.

Foursquare API is used to find out all the food venues (venue name, category, latitude and longitude coordinates) that are within 1km radius of each neighborhood using the Category ID. It is also used to find out the shops that provide kitchen supplies and its latitude and longitude coordinates.

Additional kitchen supply stores and supermarkets and their latitude and longitude coordinates are searched using Google Search and Google Map and added to the study as competitors.

3. Methodology

Data Cleaning - Neighborhoods

A total of 23 neighborhoods in Penang are extracted from Wikipedia page ([https://en.wikipedia.org/wiki/Category:Districts of Penang](https://en.wikipedia.org/wiki/Category:Districts_of_Penang)). George Town and Bayan Baru are added to the list of neighborhoods because these two areas have many food venues and might not covered by the 23 neighborhoods. So, a total of 25 neighborhoods are studied. Python Nominatim geocoder is used to find out the longitude and latitude coordinates of each neighborhood.

No	Neighborhood	Latitude	Longitude	No	Neighborhood	Latitude	Longitude
1	Air Itam	5.40187	100.27783	14	Nibong Tebal	5.1701	100.47852
2	Balik Pulau	5.34821	100.22741	15	Paya Terubong	5.3718	100.27616
3	Batu Ferringhi	5.46826	100.25033	16	Perai	5.37777	100.38196
4	Batu Kawan	5.2626	100.43046	17	Permatang Pauh	5.4105	100.41365
5	Batu Lanchang	5.38333	100.3	18	Pulau Tikus	5.43182	100.31177
6	Bayan Baru	5.32284	100.28904	19	Simpang Ampat	5.28171	100.4766
7	Bayan Lepas	5.29539	100.28576	20	Sungai Jawi	5.25111	100.45205
8	Bukit Mertajam	5.36634	100.4477	21	Tanjung Bungah	5.45747	100.287
9	Butterworth	5.39352	100.36627	22	Tanjung Tokong	5.44614	100.30525
10	Gelugor	5.36751	100.31213	23	Tasek Gelugor	5.48323	100.49048
11	George Town	5.41457	100.3298	24	Teluk Bahang	5.4578	100.21477
12	Jelutong	5.38691	100.31209	25	Teluk Kumbar	5.28976	100.23151
13	Kepala Batas	5.51468	100.42846				

Data Cleaning - Food Venues (Customers)

Foursquare API is used to find out all the food venues (venue name, category, latitude and longitude coordinates) that are within 1km radius of each neighborhood using the "Food" Category ID 4d4b7105d754a06374d81259. The Category ID for other groups can be found in Foursquare website (<https://developer.foursquare.com/docs/resources/categories>).

Note that there is a maximum query limit of 50 in Foursquare API for free subscription if you query the venues by Category ID.

Below are the sample of food venues that are returned by the Foursquare API.

	Neighborhood	Latitude	Longitude	VenueID	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	Air Itam	5.401865	100.277833	4d19d3966c8b5481c953fbcc	Air Itam Asam Laksa	5.401591	100.278277	Food Stands
1	Air Itam	5.401865	100.277833	4d4814f611618cfaba90d023	Air Itam Market Annex Food Court	5.401456	100.278058	Food Courts
2	Air Itam	5.401865	100.277833	4ce9f988948f224b6442f95d	Sin Ka Oon Bakery 新加恩	5.401650	100.278303	Bakeries
3	Air Itam	5.401865	100.277833	4dc49b4a7d8b14fb4615eba3	Kopitiam Bawah Pokok	5.401579	100.277494	Coffee Shops
4	Air Itam	5.401865	100.277833	4c950b8772dd224b8ee29f91	Air Itam Duck Rice (亞依淡鴨飯)	5.401482	100.277989	Food Stands

Foursquare API returns food venues with some categories that are suspicious as shown below:

- Flea Markets
- Food Services
- Parks
- Farmers Markets
- Hotels
- Delis / Bodegas
- Pet Services
- Government Buildings
- Fish Markets
- Wings Joints
- Markets

The venue names in the above categories are then filtered. The following venues are deleted because they are not applicable in this study:

- Sunway Hotel
- Pejabat FAMA Negeri Pulau Pinang - Government Agriculture Agency
- Pasar Awam Teluk Bahang (Market) - Wet market
- Tanjung Bungah Market- Wet market
- Pulau Tikus Market- Wet market
- Jelutong Morning Market- Wet market

Fast food restaurants are also deleted because they have their own supply chain and will not buy things from the kitchen supply shop.

A total of 32 duplicates food venues are deleted. The duplicate of venue name is due to the shop located within 1km from two or more neighborhoods.

A total of 1101 food venues are used for further analysis. The data is saved as "foodVenues_df.csv" and the file is uploaded to [Github](#).

Data Cleaning & Regrouping- Kitchen Supply Stores (Competitors)

When Foursquare API is used to find "Kitchen Supply Store" (Category ID: 58daa1558bbb0b01f18ec1b4) in all the neighborhoods in Penang, only the last row of data are applicable. The first two venues are college that provide cooking class and are deleted.

	Neighborhood	Latitude	Longitude	VenueID	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	George Town	5.414568	100.329803	59b5dfad66fc6523c3ea5c87	Ptpl Kitchen 1	5.414789	100.319187	Kitchen Supply Stores
1	George Town	5.414568	100.329803	59b5e0021499466c7a9628f8	Ptpl Kitchen 2	5.414759	100.319197	Kitchen Supply Stores
2	Permatang Pauh	5.410498	100.413651	53fabbb3a498e7a693209438d	Win Wins Kitchen Equipments Sdn. Bhd.	5.409110	100.407841	Hardware Stores

When Foursquare API is used to find "Supermarket" (Category ID: 52f2ab2ebcb57f1066b8b46) in all the neighborhoods in Penang, it returns 13 rows of data. Ten rows of data are applicable (see below). Not applicable data are deleted.

	Neighborhood	Latitude	Longitude	VenueID	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	Balik Pulau	5.348206	100.227405	53e5ad14498e7c5b4fa1ca85	Ceria 99 - Balik Pulau	5.351872	100.235224	Department Stores
1	Bayan Baru	5.322842	100.289039	4b554650f964a52087e027e3	Giant Hypermarket	5.321193	100.284257	Supermarkets
2	George Town	5.414568	100.329803	5bfa305a33e118002c942d72	Pacific Market Place	5.415766	100.330555	Supermarkets
3	George Town	5.414568	100.329803	4c307539213c2d7f427e325d	Giant Supermarket	5.421229	100.325000	Supermarkets
5	Jelutong	5.386907	100.312090	575680e0498ebb75e97d0f3e	Sunshine Jelutong	5.394208	100.314752	Supermarkets
7	Kepala Batas	5.514678	100.428464	57777711498ecac8cec464cc	sweet world enterprise	5.517612	100.429965	Supermarkets
8	Nibong Tebal	5.170096	100.478519	4de4a8a952b1d38d29a21955	Yawata	5.166896	100.475567	Supermarkets
9	Perai	5.377766	100.381963	4cfe58f5ee9cb60cc84a9fad	Mydin Wahab Enterprise	5.386533	100.384618	Supermarkets
10	Pulau Tikus	5.431822	100.311768	4d108514ba378cfa8a267893	Supermarket Bandar Baru (新城)	5.431112	100.311470	Supermarkets
12	Tanjung Tokong	5.446139	100.305254	4c24a0fdb7b8a593ebb63ce8	Cold Storage	5.450468	100.306444	Grocery Stores

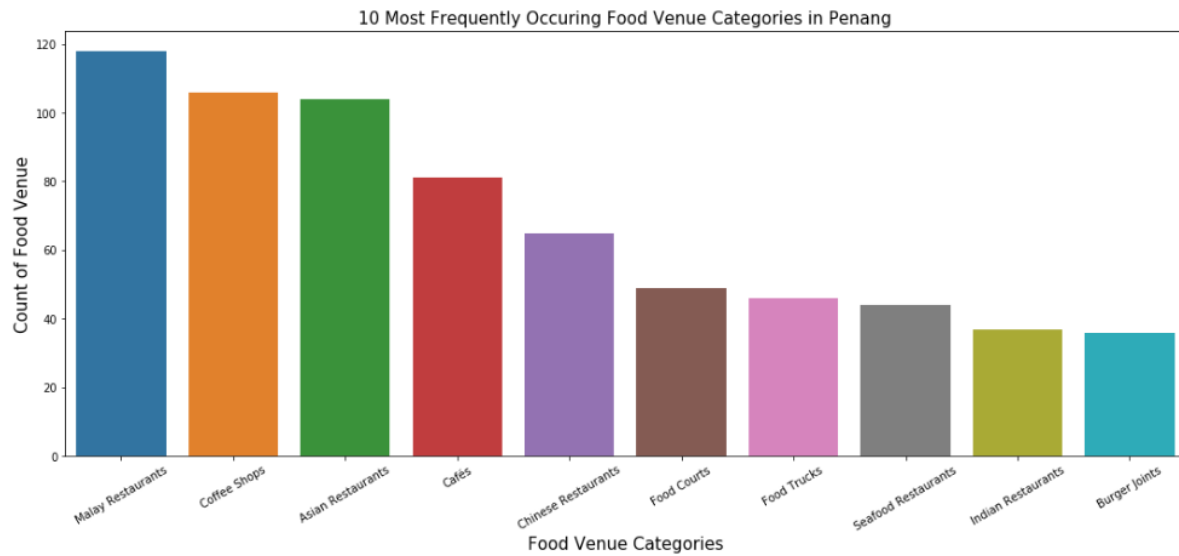
Since Foursquare API does not provide many competitor data, Google Search and Google Map are used to search kitchen supply stores/supermarkets and their latitude and longitude coordinates. The data ("competitors_not_foursquare.csv") is then imported into a dataframe. Each shop is then assigned to the nearest neighborhood using a Python `geopy.distance.geodesic` function.

The data from Google Map and Foursquare API are then merged. Duplicate venues again are deleted. The "VenueID" column is also dropped.

The data is saved as "competitor_df_final.csv" and the file is uploaded to [Github](#).

4. Results

What are the Top 10 Most Common Food Venue Categories?

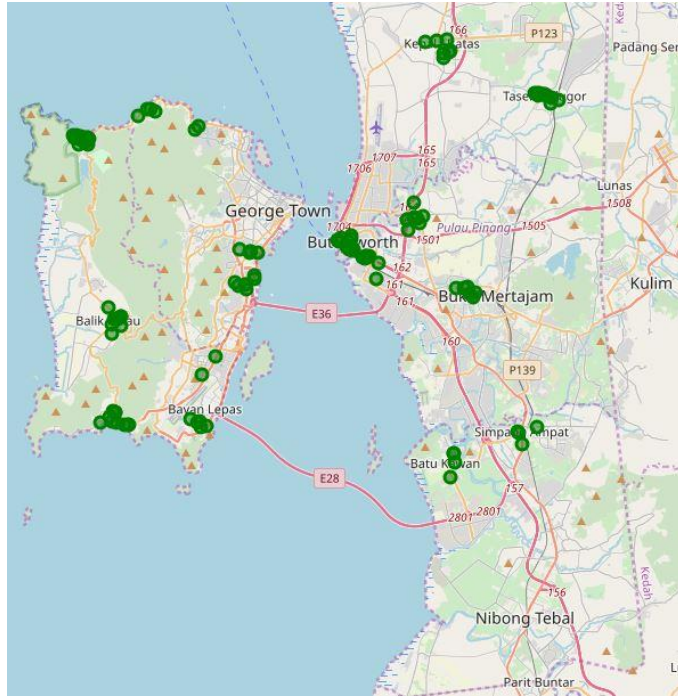


	FoodVenueCategory	Frequency
0	Malay Restaurants	118
1	Coffee Shops	106
2	Asian Restaurants	104
3	Cafés	81
4	Chinese Restaurants	65
5	Food Courts	49
6	Food Trucks	46
7	Seafood Restaurants	44
8	Indian Restaurants	37
9	Burger Joints	36

Malay restaurants are the most common food venues in Penang, followed by coffee shops and Asian restaurants.

Where are the Top 10 Locations that have the Most Malay Restaurants?

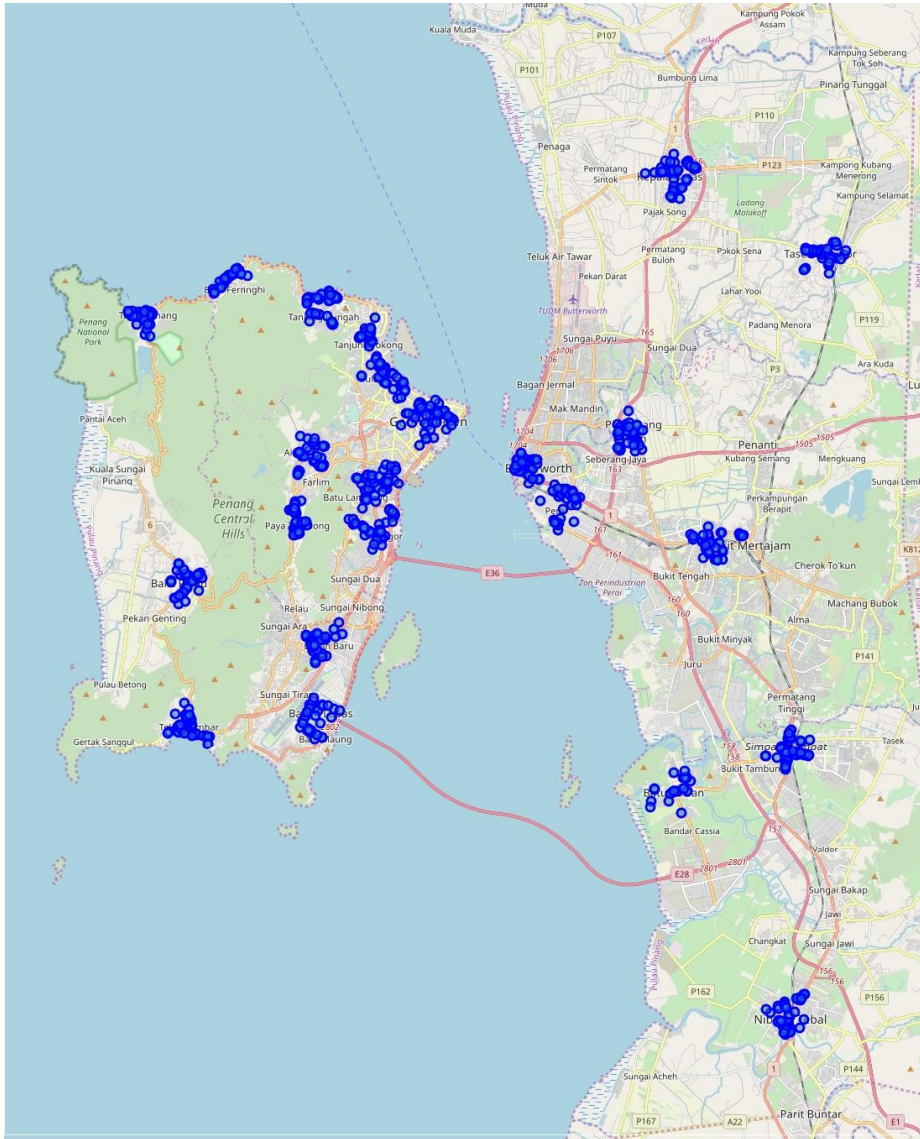
	Count
Teluk Bahang	16
Tasek Gelugor	11
Butterworth	10
Teluk Kumbar	10
Balik Pulau	9
Perai	8
Permatang Pauh	8
Gelugor	7
Kepala Batas	7
Bukit Mertajam	6



Refer to "map_pg_malay_restaurants.html" in [Github](#).

Teluk Bahang has the most Malay restaurants followed by Tasek Gelugor, Butterworth and Teluk Kumbar. If you plan to supply raw materials to Malay restaurants, these neighborhoods are the target.

Where are all the Food Venues located?



Refer to "map_pg_foodvenues.html" in [Github](#).

By plotting the location of all the food venues on Penang map, we will know where our customers are located. These locations have the potential to open a kitchen supply store.

However, we also need to find out how many shops that provide kitchen supplies in each neighborhood. These shops are our competitors. This information can be shown in the map below (red dots represent competitors, blue dots represent food venues):

	FoodVenueCount	CompetitorCount	AvgFoodVenue
Neighborhood			
Bayan Lepas	50	NaN	NaN
Batu Lanchang	49	NaN	NaN
Tanjung Bungah	49	NaN	NaN
Teluk Bahang	49	NaN	NaN
Batu Ferringhi	48	NaN	NaN
Teluk Kumbar	48	NaN	NaN
Tasek Gelugor	46	NaN	NaN
Batu Kawan	29	NaN	NaN
Air Itam	50	1.0	50.00
Gelugor	49	1.0	49.00
Nibong Tebal	49	1.0	49.00
Balik Pulau	48	1.0	48.00
Perai	48	1.0	48.00
Paya Terubong	47	1.0	47.00
Kepala Batas	48	2.0	24.00
Simpang Ampat	48	3.0	16.00
Bukit Mertajam	49	5.0	9.80
Butterworth	47	5.0	9.40
Tanjung Tokong	31	4.0	7.75
Jelutong	30	4.0	7.50
Pulau Tikus	47	7.0	6.71
Permatang Pauh	49	8.0	6.12
Bayan Baru	46	9.0	5.11
George Town	47	12.0	3.92

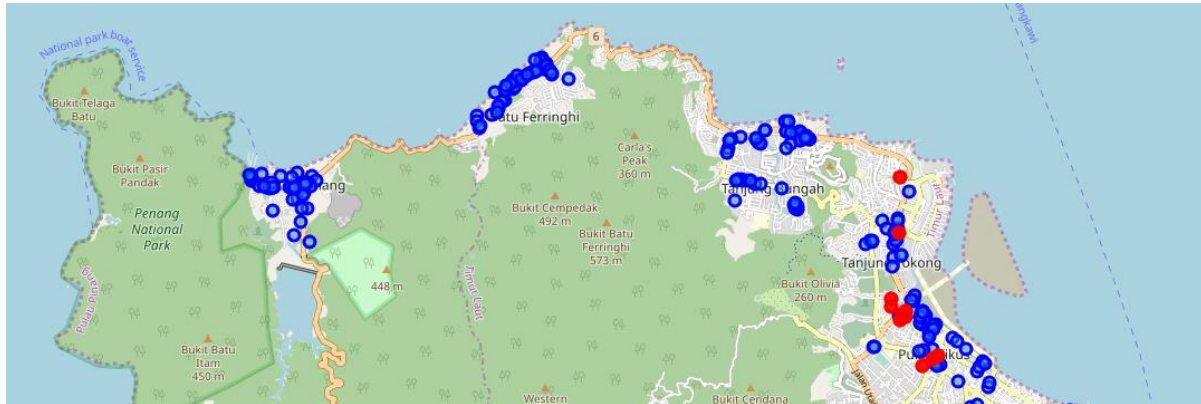
NaN means no competitor.

5. Discussions

Based on the average food venues per competitor table, there is no competition in Bayan Lepas, Batu Lanchang, Tanjung Bungah, Teluk Bahang, Batu Ferringhi, Teluk Kumbar, Tasek Gelugor, Batu Kawan.

The detail analysis of each neighborhood will be discussed – started from north of Penang Island to south of island, followed by neighborhoods from north to south Penang in main land.

North of Penang Island

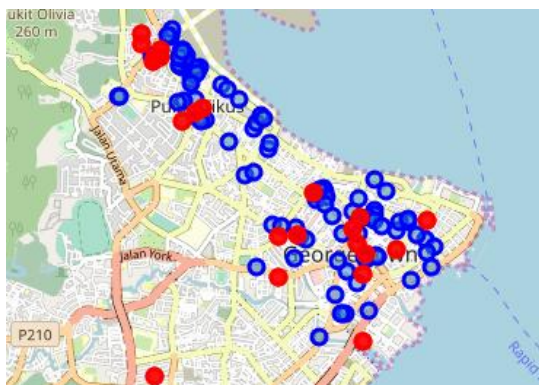


- Tanjung Bungah - 49 food venues with no competitor.
- Batu Ferringhi - 48 food venues with no competitor.
- Teluk Bahang - 49 food venues with no competitor.
- Tanjung Tokong - 31 food venues with 4 competitors.

There are total 177 food venues served by 4 competitors, average food venue served by one competitor is 44.25. Teluk Bahang is a less developed area compared to Tanjung Bunga and Batu Ferringhi, the shops that provide foods and drinks may not have many business.

Recommendation: Tanjung Bungah, Batu Ferringhi

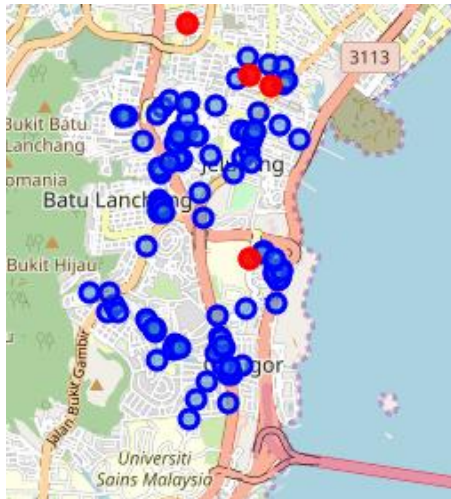
George Town and Pulau Tikus (East of Penang Island)



These 2 neighborhoods are not good places to open a kitchen supply store because of high competition, compared to other neighborhoods.

Recommendation: None

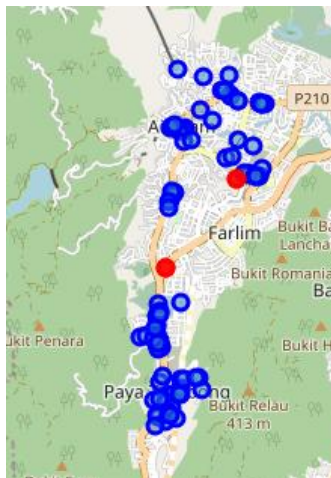
Jelutong, Batu Lanchang and Gelugor (East of Penang Island)



- Jelutong - 30 food venues with 4 competitors, less attractive because less food venues with more competitors, compared to Gelugor and Batu Lanchang.
- Gelugor - 49 food venues with 1 competitor.
- Batu Lanchang - 49 food venues with no competitor.

Recommendation: Gelugor and Batu Lanchang

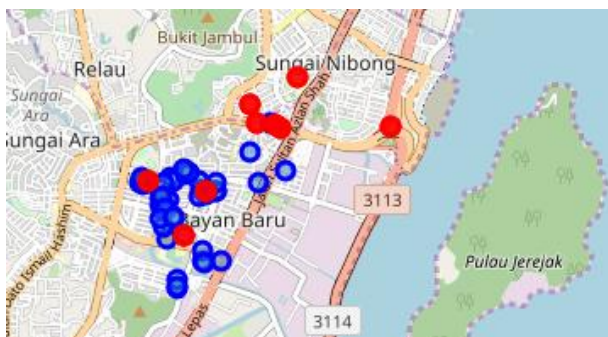
Air Itam and Paya Terubong (East of Penang Island)



- Air Itam - 50 food venues with 1 competitor.
- Paya Terubong - 47 food venues with 1 competitor.

Recommendation: Air Itam & Paya Terubong

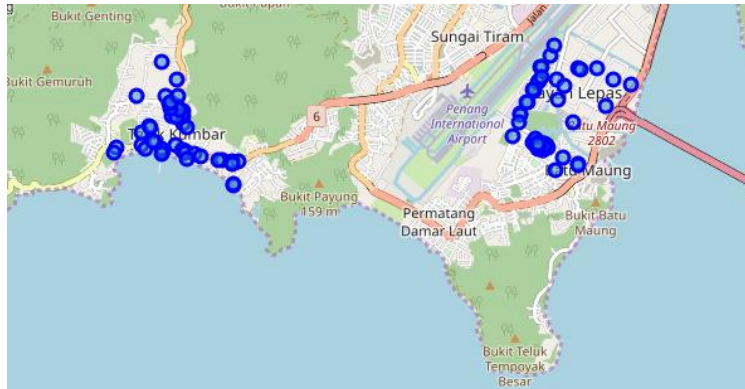
Bayan Baru (East-south of Penang Island)



This neighborhood is a not good place to open a kitchen supply store because of high competition, compared to other neighborhoods.

Recommendation: None

Teluk Kumbar, Bayan Lepas (South of Penang Island)



- Bayan Lepas - 50 food venues with no competitor.
- Teluk Kumbar - 48 food venues with no competitor. It is a less developed area compared to Bayan Lepas.

Recommendation: Bayan Lepas

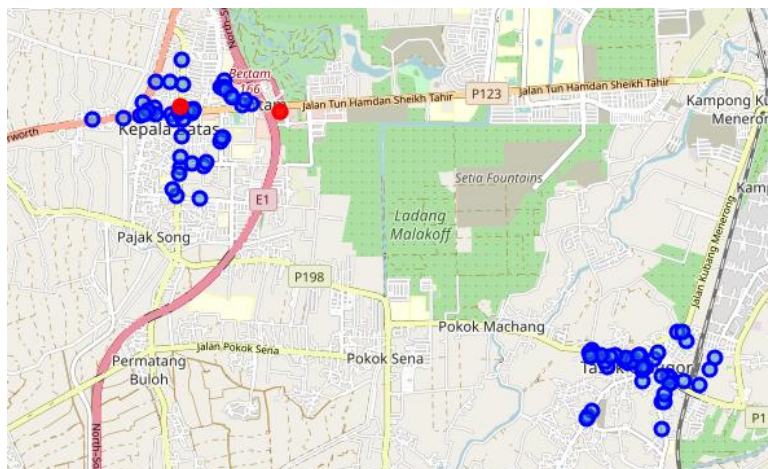
Balik Pulau (West of Penang Island)



Balik Pulau - 48 food venues with 1 competitor.

Recommendation: Balik Pulau

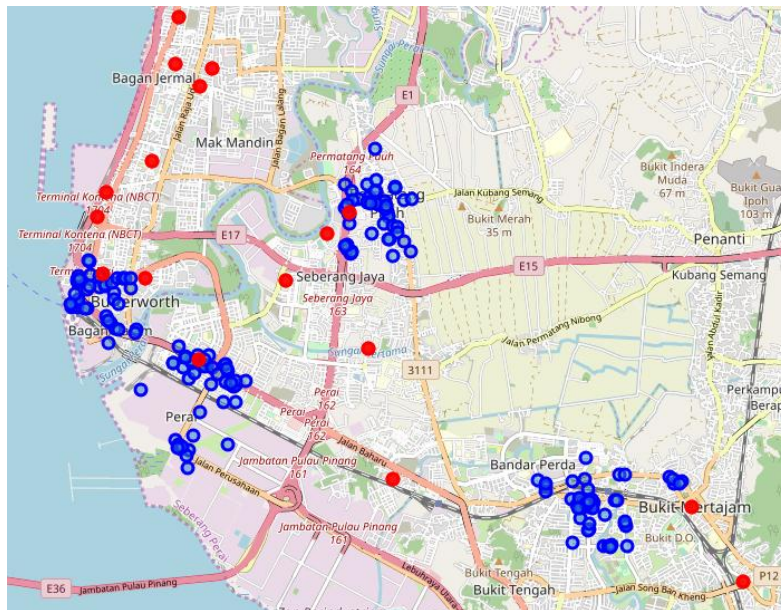
Kepala Batas and Tasek Gelugor (North of Penang mainland)



- Tasek Gelugor - 46 food venues with no competitor.
- Kepala Batas - 48 food venues with 2 competitors.

Recommendation: Tasek Gelugor

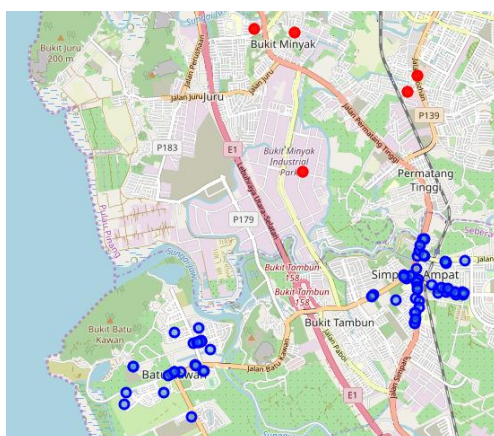
Perai, Butterworth, Bukit Mertajam, Permatang Pauh (Center of Penang mainland)



- Perai - 48 food venues with 1 competitor (0 competitor is shown in table).
- Bukit Mertajam - 49 food venues with 1 competitor (5 competitors are shown in table). Additional four competitors shown in the table are assigned to Bukit Mertajam because it is the nearest neighborhood from this competitors. If you look at the map, the competitors are located far away from Bukit Mertajam.
- Butterworth, Permatang Pauh neighborhoods have many kitchen supply stores based on the map, the competition are high in these areas and not suitable to open a kitchen supply store.

Recommendation: Perai, Bukit Mertajam

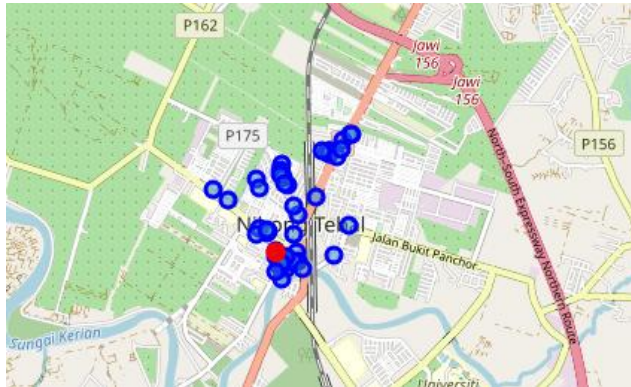
Simpang Ampat and Batu Kawan (South of Penang mainland)



- Batu Kawan - 29 food venues with no competitor. Even though there is no competitor in this neighborhood, the number of food venues is not high compared to other neighborhoods.
- Simpang Ampat - 48 food venues with 3 competitors. However, the competitors are located far away from the town.

Recommendation: Simpang Ampat

Nibong Tebal (South of Penang mainland)



Nibong Tebal - 49 food venues with 1 competitor.

Recommendation: Nibong Tebal

Finally, it is recommended to open a kitchen supply store in the following neighborhoods:

- Tanjung Bungah and Batu Ferringhi
- Gelugor and Batu Lanchang
- Air Itam and Paya Terubong
- Bayan Lepas
- Balik Pulau
- Tasek Gelugor
- Perai and Bukit Mertajam
- Simpang Ampat
- Nibong Tebal

You need to visit all the neighborhoods to see how good is the food venue business and whether there are other competitors that are not listed in Foursquare database. The shop rental price is also a factor to be considered.

Challenges & Limitations

There are many challenges and limitations during the data extraction:

- The longitude and latitude coordinates of each neighborhood are extracted from Nominatim geocoder. The coordinates returned may or may not represent the location where majority of the people live. It is best if the geocoder can provide the border of each neighborhood.
- Foursquare API is used to find out the food venues and kitchen supply store venues within 1km radius (defined by us) from the neighborhood. If a shop is located out of 1km away from the neighborhood, it is excluded from the study. If we extended the distance to 2km or more and later remove the duplicate venues using the Pandas

dataframe function, it sometime removes a wrong one - a food venue in Penang Island is assigned to the neighborhood in the main land of Penang. Thus, we have to manually verify the correct neighborhood is assigned and this take up a lot of time.

- Foursquare API is depending on the users to supply venue data and assign category. Many kitchen supply stores in Penang are not included in Foursquare database. Many shops are assigned to the wrong category, like "Government Buildings", "Pet Services, Parks" appear when "Food" Category ID is used to search the venue. To improve the accuracy of the analysis, we may need to go through the food venue list (1101 rows) one by one to remove the rows that are not applicable. The venue category "Asian Restaurants" contains many restaurants, food courts, coffee shops that also belong to other category like "Chinese Restaurants", "Malay Restaurants", "Coffee Shops" and "Food Courts". To improve the data accuracy, this "Asian Restaurants" may need to regroup into "Chinese Restaurants", "Malay Restaurants", "Coffee Shops" and "Food Courts".
- The longitude and latitude coordinates returned by Foursquare API or Nominatim geocoder sometime are not accurate. If you zoom in the map, you will find out that the location is not correct.
- In this study, we make assumption that one food venue is one customer. However, many coffee shops and food courts have more than one food stall which contribute more than one customer.
- The venue data returned by Foursquare API is different each time the data is queried because its database is frequently updated. Thus, the analysis has to be redo. To avoid this problem, the data is saved to a CSV file once it is extracted and cleaned up.
- Foursquare API and Nominatim geocoder sometime give timeout error with no data returned. To avoid this problem, the data is saved to a CSV file once it is extracted and cleaned up.

6. Conclusions

In this study, Python Nominatim geocoder is used to find out the latitude and longitude coordinates of each neighborhood in Penang. Using Foursquare API, food venues and kitchen supply shops and their latitude and longitude coordinates are gathered for analysis. Maps are plotted to show the locations of food venues and kitchen supply stores.

A bar chart is also plotted to show the top 10 most common food venue categories in Penang. Malay restaurants, coffee shops and Asian restaurants are the most common food venues in Penang. Many neighborhoods have many food venues with 1 or no competitor. Although these neighborhoods are good places to open a kitchen supply store, you may need to visit the places to see if there are any competitors not given by the Fourquare API, how big are their shops and what kinds of kitchen utensils sold by them. The shop rental and how good are the business of the customers also contribute to the success of a kitchen supply shop.