Recommending the ideal location to open a restaurant

DANIEL BOEY

JUNE 7, 2021

Introduction

- Penang is well known domestically and in Singapore for being the "food paradise" of Malaysia. In this project we will try to find an optimal location for a restaurant.

- This report will be targeted to stakeholders interested in opening a **Japanese restaurant** in **Penang**, Malaysia

- We will try to detect **locations that are not already crowded with restaurants**. We are also particularly interested in **areas with no Japanese restaurants in vicinity**. We would also prefer locations **as close to city center as possible**, assuming that first two conditions are met.

- We will use data science to generate a few most promising neighborhoods based on this criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.
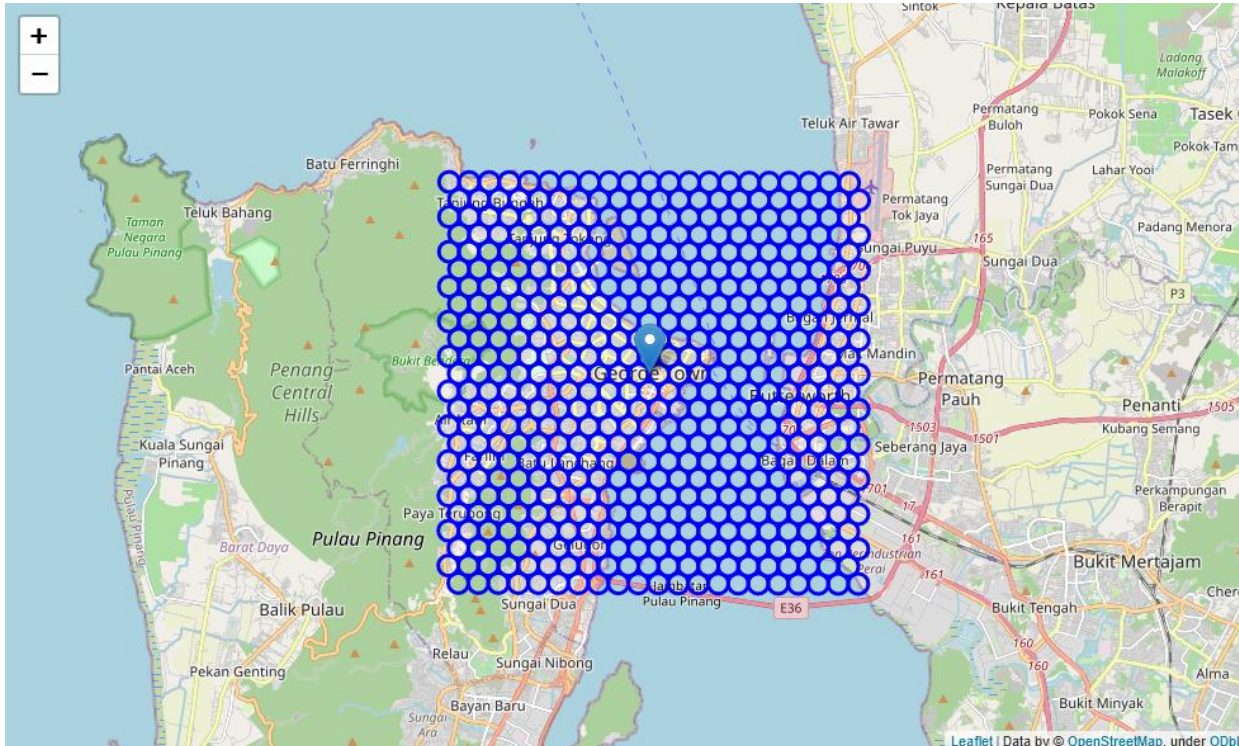
Data

- Based on definition of our problem, factors that will influence our decision are:
  - o number of existing restaurants in the neighborhood (any type of restaurant)
  - o number of and distance to Japanese restaurants in the neighborhood, if any
  - o distance of neighborhood from city center

- Following data sources will be needed to extract/generate the required information:
  - o centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using Google Maps API reverse geocoding
  - o number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API
  - o coordinate of Penang center will be obtained using Google Maps API geocoding of Penang's main city, George Town

Exploratory Data Analysis

A visualization of the city center location and candidate neighborhood centers, the 12x12 kilometers of hexagonal grid centered around George Town

Exploratory Data Analysis

	Address	Latitude	Longitude	X	Y	Distance from center
0	Mukim 13 Paya Terubong, Penang	5.357847	100.277194	641514.479933	592367.674215	8448.076704
1	130, Halaman Bukit Gambir 4, 11700 Gelugor, Pu...	5.357836	100.282608	642114.479933	592367.674215	8055.432949
2	52, Halaman Bukit Gambir 9, Sunway Bukit Gambi...	5.357824	100.288022	642714.479933	592367.674215	7689.603371
3	1, Halaman Bukit Gambir, Sunway Bukit Gambier,...	5.357813	100.293436	643314.479933	592367.674215	7354.590403
4	Doping Control Centre, 11700 Gelugor, Pulau Pi...	5.357801	100.298850	643914.479933	592367.674215	7054.785610
5	10, Lengkok Sastera, 11800 Gelugor, Pulau Pinang	5.357790	100.304265	644514.479933	592367.674215	6794.850992
6	6, 11900 Gelugor, Pulau Pinang	5.357778	100.309679	645114.479933	592367.674215	6579.513660
7	Unnamed Road, 11700, 11700 Gelugor, Penang	5.357767	100.315093	645714.479933	592367.674215	6413.267498
8	Jambatan Pulau Pinang, Gelugor, Pulau Pinang	5.357755	100.320507	646314.479933	592367.674215	6300.000000
9	Jambatan Pulau Pinang, Gelugor, Pulau Pinang	5.357743	100.325921	646914.479933	592367.674215	6242.595614

We'll then use Google Maps API to approximate addresses of those locations. We'll then place them all into a Pandas dataframe and clean up any invalid addresses

Exploratory Data Analysis
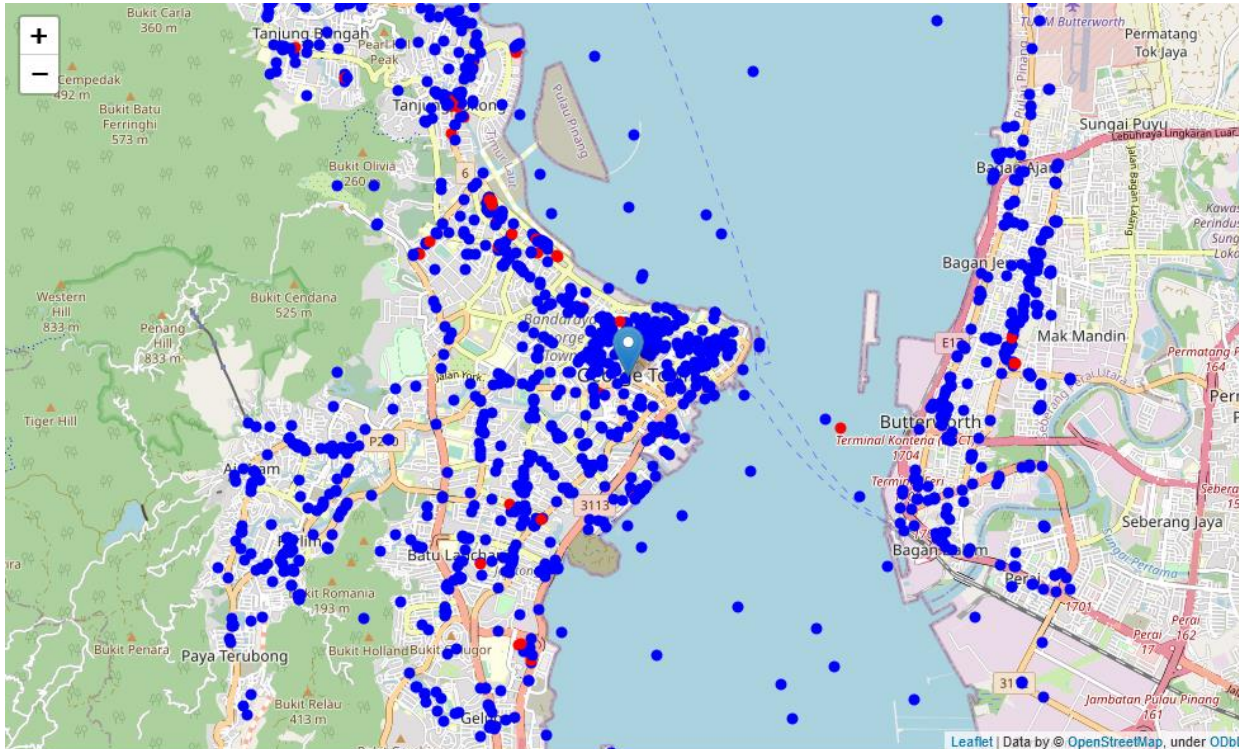
Foursquare API has given us:
- 1299 total number of restaurants,
- 51 Japanese restaurants (3.93%),
- an average of 2.25 restaurants / neighborhood

Plotted the restaurants in our area of interest on the map, and coloring Japanese restaurants in red.

Methodology

- We will now direct our efforts on detecting areas of Penang that have low restaurant density, particularly those with low number of Japanese restaurants.

- We will focus on the most promising areas and within those create clusters of locations that meet some basic requirements established in discussion with stakeholders

  o We will take into consideration locations with no more than two restaurants in a radius of 250 meters, and we want locations without Japanese restaurants in a radius of 400 meters

  o We will present a map of all such locations but also create clusters (using k-means clustering) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final "street level" exploration and search for optimal venue location by stakeholders
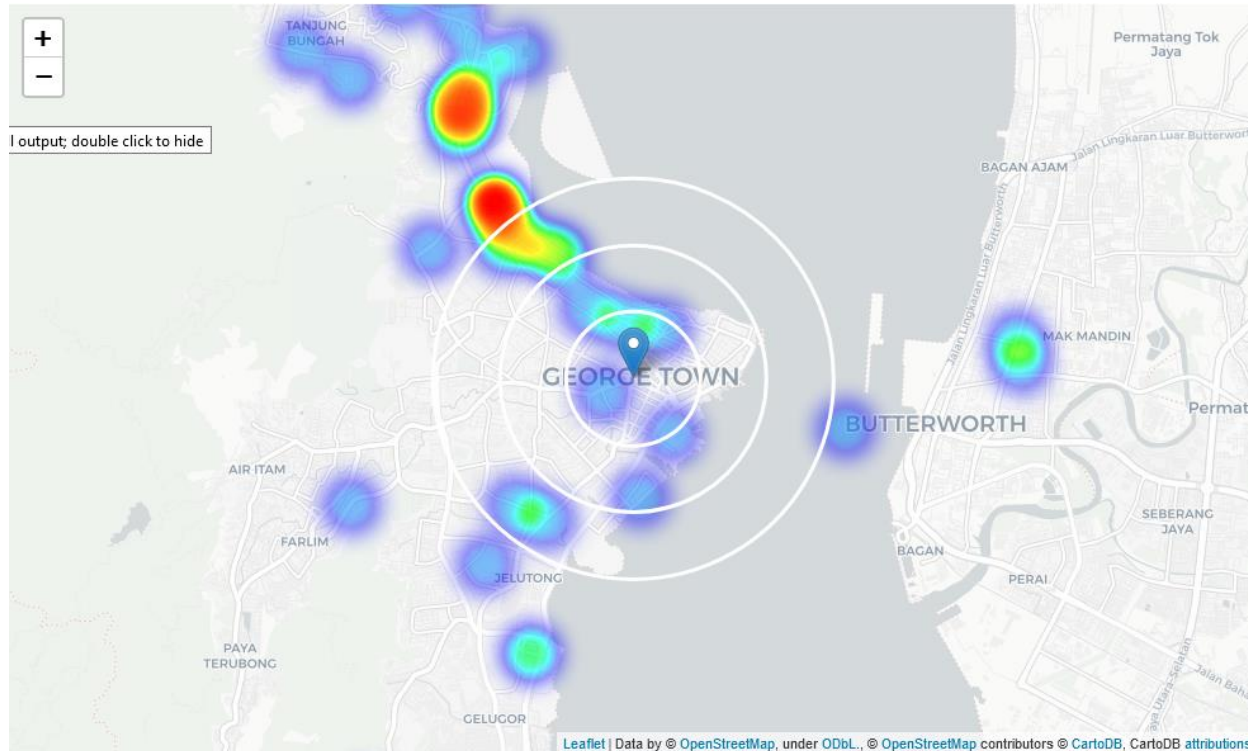
Analysis & Clustering
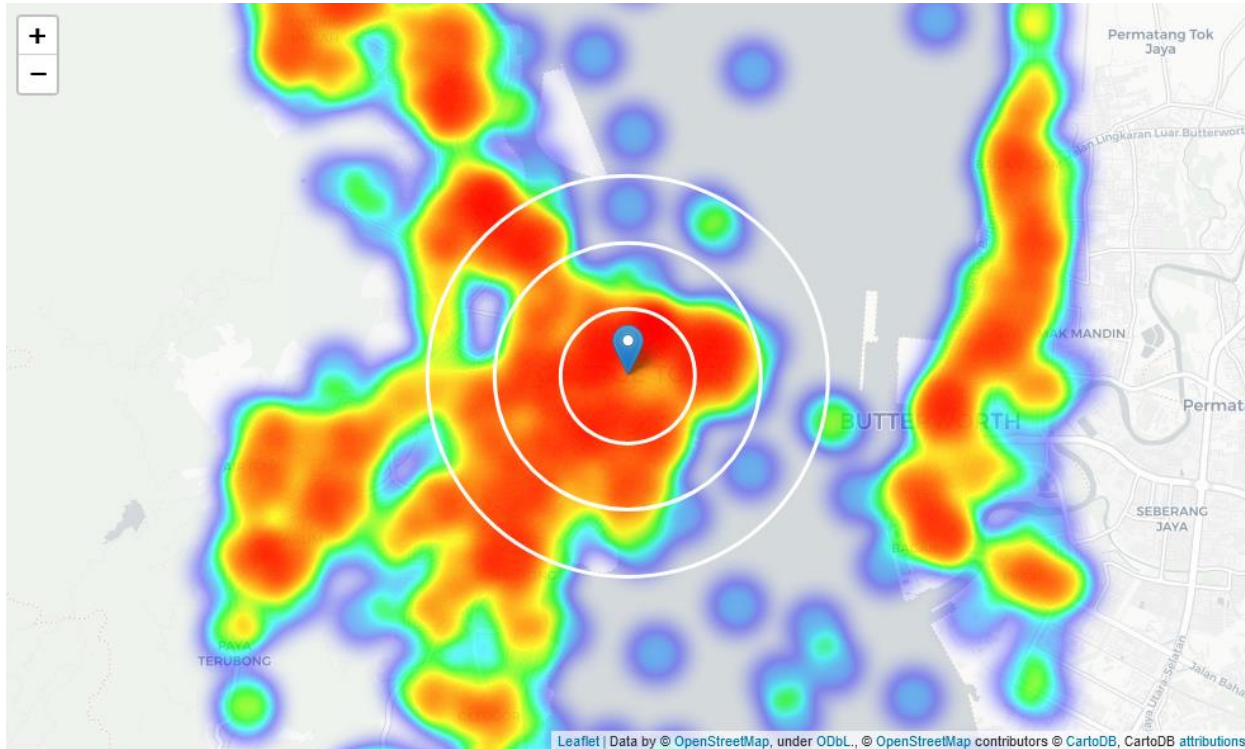
By creating a heatmap showing the density of restaurants, we can observe pockets of low restaurant density closest to the city center to the west of George Town
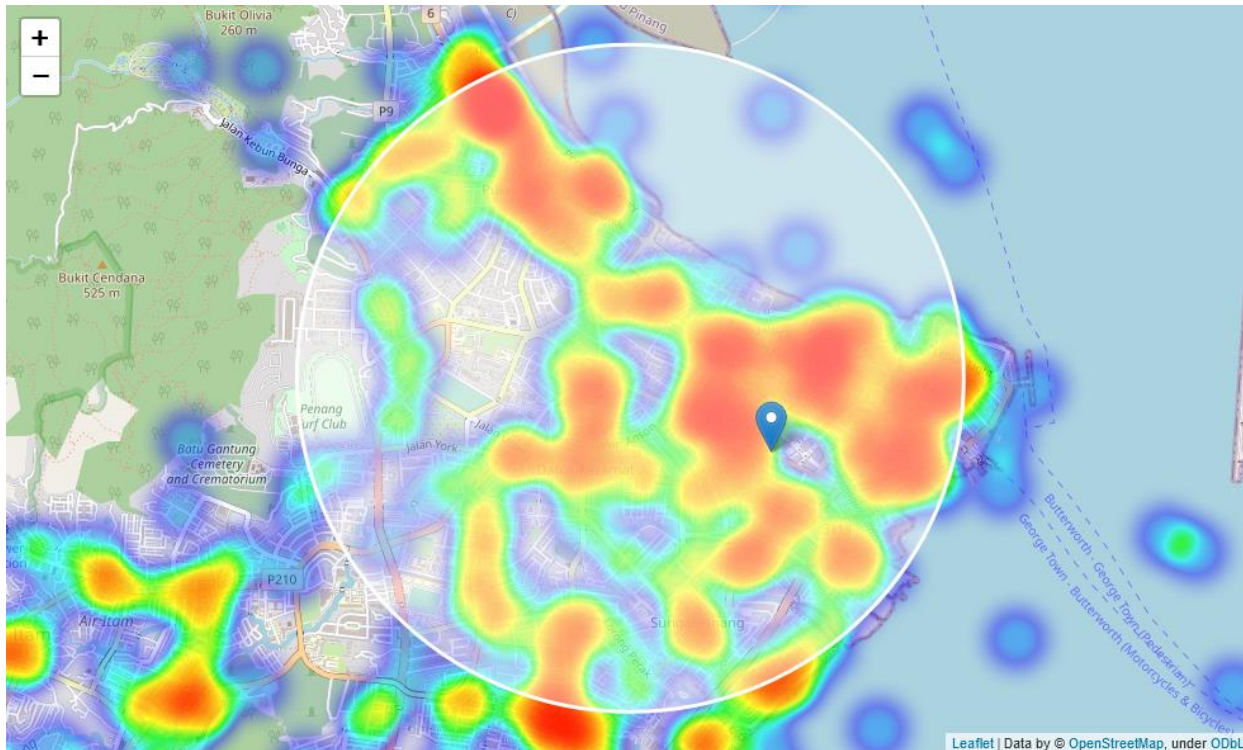
Analysis & Clustering

A reverse, doing the same but a heatmap showing the density of Japanese restaurants only.

It indicates higher density of existing Japanese restaurants directly northwest from George Town, with closest pockets of **low Japanese restaurant density positioned west and southwest from city center**.

Vicinities of Jalan Logan

Based on this information we will now focus our analysis on areas **west from George Town center** - we will move the center of our area of interest and reduce its size to have a radius of **2.5km**.

So, we define a new, narrower region of interest, which will include low-restaurant-count parts of Jalan Logan closest to George Town.

Vicinities of Jalan Logan

|	Latitude	Longitude	X	Y	Restaurants nearby	Distance to Japanese restaurant
0	5.412588	100.310247	645164.479933	598428.057122	6	1615.229149
1	5.412586	100.311150	645264.479933	598428.057122	9	1515.289746
2	5.412584	100.312052	645364.479933	598428.057122	8	1415.358902
3	5.412582	100.312955	645464.479933	598428.057122	7	1315.438570
4	5.412580	100.313857	645564.479933	598428.057122	7	1215.531341
5	5.412578	100.314759	645664.479933	598428.057122	7	1115.640735
6	5.412577	100.315662	645764.479933	598428.057122	8	1015.771656
7	5.412575	100.316564	645864.479933	598428.057122	8	915.931147
8	5.412573	100.317467	645964.479933	598428.057122	9	816.129693
9	5.412571	100.318369	646064.479933	598428.057122	2	716.383616

We calculate the number of restaurants in the vicinity (we used a radius of 250 meters) and distance to closest Japanese restaurant.

Filtering only locations with no more than two restaurants in a radius of 250 meters, we'll end up with:
- 156 locations with no more than 2 restaurants nearby
- 328 locations with no Japanese restaurants within 400m
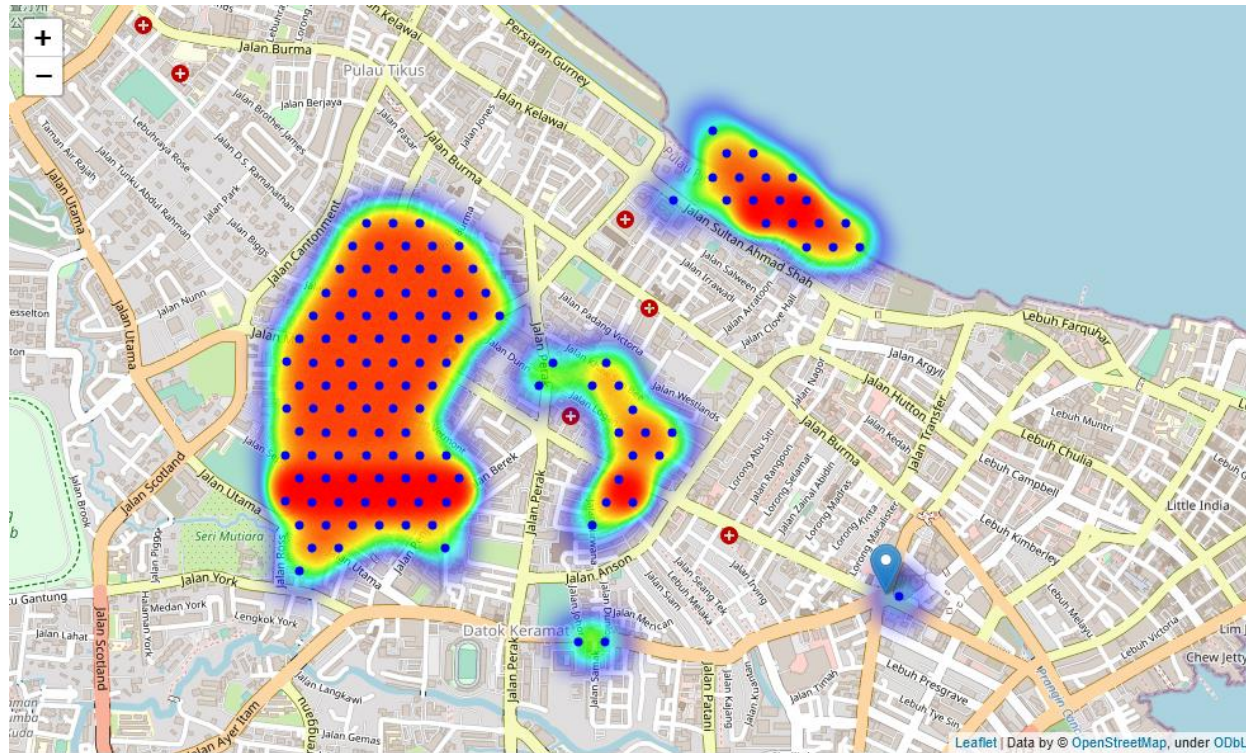- 128 locations that meets both conditions

Vicinities of Jalan Logan
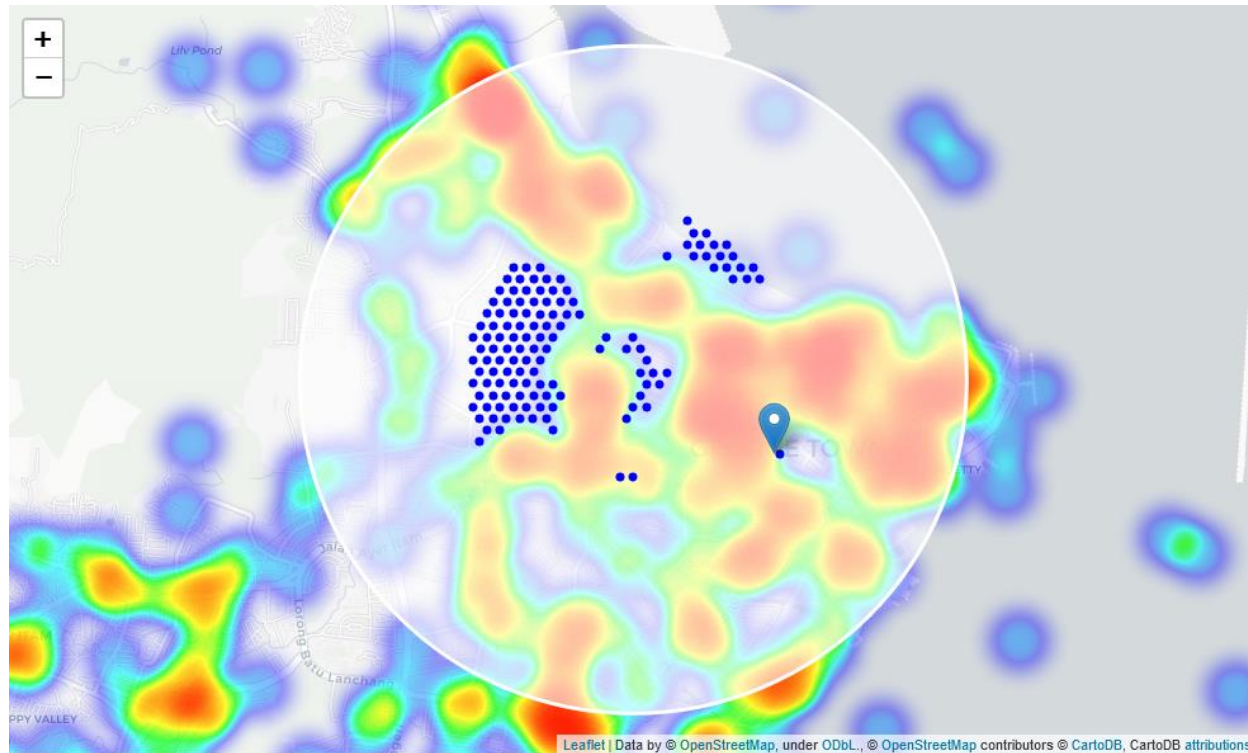
Overlaying these potential locations onto the heatmap.
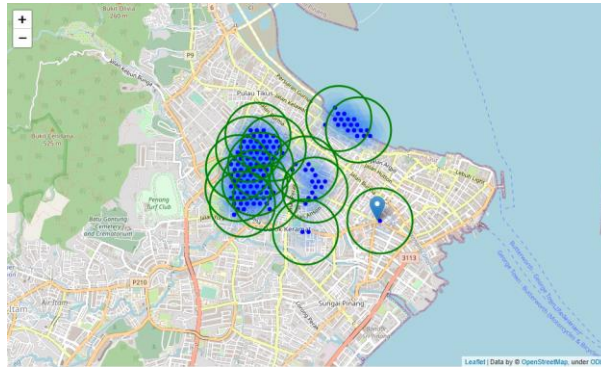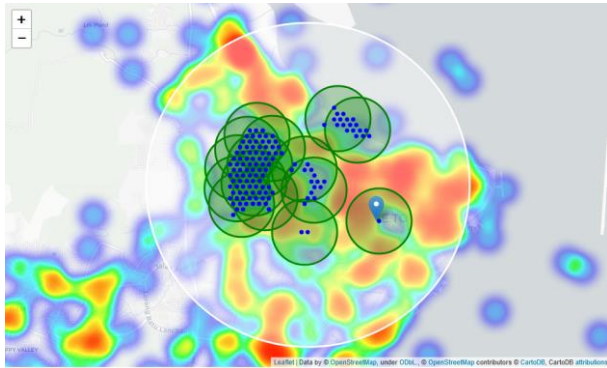
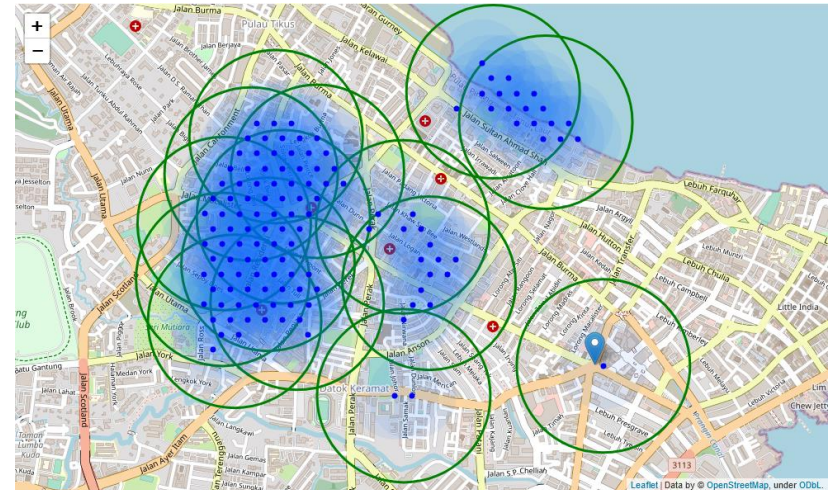# Vicinities of Jalan Logan

How it looks reversing the heatmap to the potential locations only.

Clustering with kMeans

Clustering with machine learning algorithm to create centers of zones containing good locations

Zoomed in:

Clustering with kMeans

48,B,C,D, Jln Sultan Ahmad Shah, 10050 George Town, Pulau Pinang => 1.4km from George Town
1, Jalan Tull, George Town, 10450 George Town, Pulau Pinang => 2.0km from George Town
1, Jalan Pierce, George Town, 10350 George Town, Pulau Pinang => 1.9km from George Town
233, Jalan Macalister, George Town, 11400 George Town, Pulau Pinang => 1.1km from George Town
18-1-1, Casabella APT, Pulau Tikus, 10350 George Town, Penang => 2.3km from George Town
65, Jalan Utama, George Town, 10450 George Town, Pulau Pinang => 2.1km from George Town
219, Jalan Magazine, George Town, 10300 George Town, Pulau Pinang => 0.1km from George Town
4, Jalan Turf Club, 10450 George Town, Pulau Pinang => 2.2km from George Town
111, Jalan Macalister, George Town, 10450 George Town, Pulau Pinang => 2.3km from George Town
395, Jalan Dato Keramat, George Town, 10150 George Town, Pulau Pinang => 1.1km from George Town
12, Jalan Khaw Sim Bee, George Town, 10400 George Town, Pulau Pinang => 1.4km from George Town
15, Lebuhraya Codrington, Pulau Tikus, 10350 George Town, Pulau Pinang => 2.3km from George Town
58, Jln Sultan Ahmad Shah, 10050 George Town, Pulau Pinang => 1.7km from George Town
219, Jalan Macalister, George Town, 10400 George Town, Pulau Pinang => 2.0km from George Town
34, Jalan Berek, George Town, 10450 George Town, Pulau Pinang => 1.8km from George Town

By geocode-reversing these candidate area centers and getting their addresses, these are what we can recommend to our stakeholders for further analysis

Results & Conclusions

- Result of all this is 15 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general and Japanese restaurants particularly

- Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met

- Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.