



---

# MOVING TO YOUR DREAM AREA

---

IBM Data Science Professional Certificate



OCTOBER 5, 2018

A DATA SCIENCE PROJECT (COURSERA)

By Amir Fahmi bin Jalilluddin

## Contents

<b>Background .....</b>	<b>2</b>
<b>Problem .....</b>	<b>3</b>
<b>Data .....</b>	<b>4</b>
<b>Methodology .....</b>	<b>5</b>
<b>Result .....</b>	<b>6</b>
<b>Discussion .....</b>	<b>11</b>
<b>Conclusion .....</b>	<b>11</b>

## Background

The booming of location services has created new capabilities to explore new area without even physically be near the location. One such early invention that use these capabilities is Google Maps. This application manages not only give an accurate map of an area but also can estimate the time taken to travel from point 'A' and to point 'B'. This not only give user the ability to planned ahead, but also give peace of mind when travel to unfamiliar places. Location services also been used by developer to make location-based game like Pokémon GO by Niantic to provide a much challenging gameplay. Some developer also uses these capabilities to identify consumer patterns, tastes, and trends in certain area like TripAdvisor and Foursquare

In this project, we will be using the power of data science to create a database in which help an individual or families to move to their dream area or neighbourhood by comparing venues that available in certain area in ahead of time before moving in.

## **Problem**

Moving can be hassle if doing it unplanned especially when try to find and decide the best area for your family.

Some families may consist only 2 persons, husband and wife, and such want to be near work places and entertainment places. Some families may start to have children of their own and wants to be near to schools and some community centres to enjoy some activities with family members. Some families may consist of old folks that wants to be have easy access to hospitals and somewhere quiet.

By creating a database of the venues of these areas, we then may provide a service where people can refer to before moving in into the area to ensure the area in question suites the needs and wants of the family.

## Data

This project will be projected upon the venues in area of Selangor, Malaysia especially in Ampang, Shah Alam, Bangi and Bandar Baru bangi.

This project will use data collected from these 2 sources;

- Foursquare using Foursquare API
- World Postal Code website ([worldpostalcode.com/malaysia/selangor](http://worldpostalcode.com/malaysia/selangor))

We will use this data collection to compare the density of venues according to 10 types;

- Entertainment, include but not limited to, Stadium, Movie Theatre, Museum and Music Club
- University, include but not limited to College and Sport facilities
- Event, that currently be held
- Food, consist any place that sell food
- Nightlife, entertainment places that usually open after dark
- Recreational facilities, include but not limited to Sport facilities and Indoor play area
- Residence building, like apartments and condos
- Shop and Services, include but not limited to ATM, Child Daycare and luxury item shop
- Transport, any terminal points
- Others, included but not limited to Factory, Office building and Medical Center

We also can assume 3 groups of people, in which we will use as case study for this project;

- a single man in mid 20s that like to party and entertainment (Case Study A)
- a family with 3 children whom age ranges from 3 years old to 18 years old (Case Study B)
- an old couple that recently retired and what to move away from busy city (Case Study C)

## Methodology

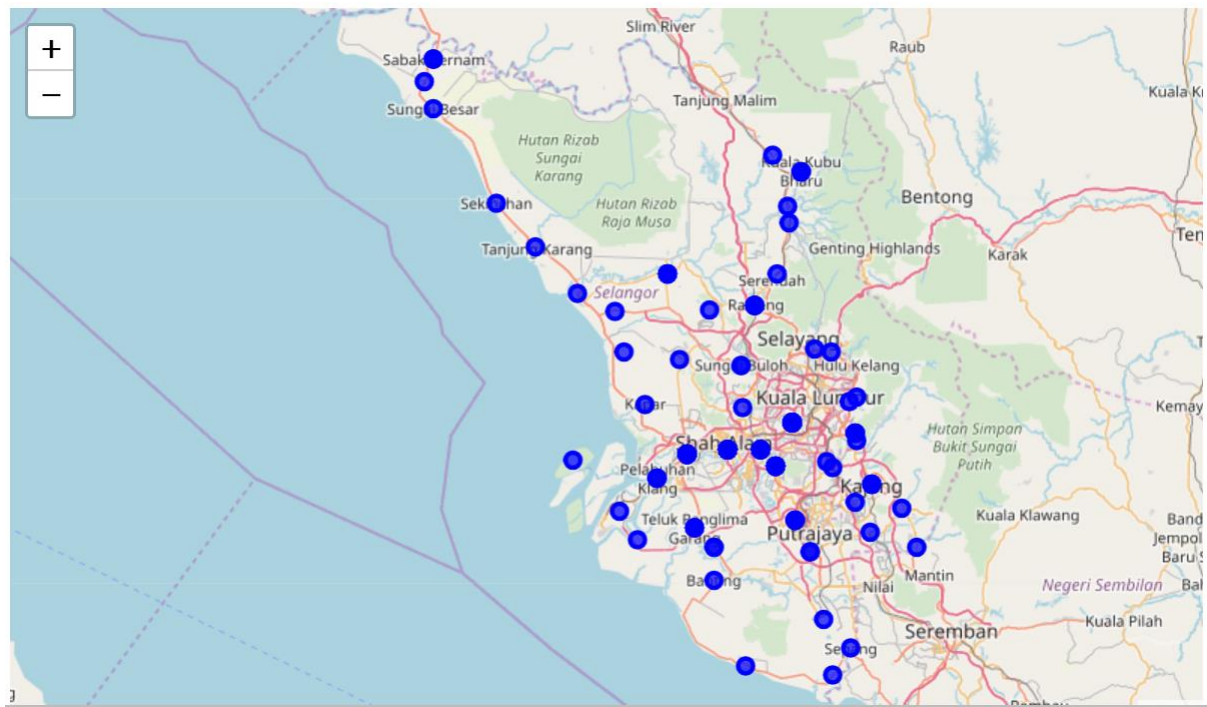
The data was collected through World Postal Code and Foursquare API. The data then been clean for this project purpose.

The data gathered using World Postal Code then use Nominatim geocoder to get the latitude and longitude of each area. This area will be visualize using Folium package to ensure the latitude and longitude collected is the same on world maps.

By using the Foursquare API, we will collect enough information on each venue that included in surrounding 20km (10km radius) of the latitude and longitude provided by Nominatim geocoder. The data collection will then be used to compare the venues in Ampang, Shah Alam, Bangi and Bandar Baru Bangi.

## Result

Below is the map of Selangor, Malaysia that mapped with all the area in it;



We then compare the 4 area (Ampang, Shah Alam, Bangi, Bandar Baru Bangi);

	Venue_Category	Ampang_Count	SAlam_Count	Bangi_Count	BBBangi_Count
0	Arts_Entertainment	1	1	2	1
1	College_University	1	9	7	0
2	Event	0	0	0	0
3	Food	8	11	16	21
4	Nightlife	2	0	0	0
5	Outdoors_Recreation	3	7	1	8
6	Professional_Other	18	12	6	4
7	Residence	1	3	0	1
8	Shop_Service	10	0	9	12
9	Travel_Transport	5	5	3	2

We can also get each area's venues more detailed data;

- Ampang

	Venue_Type	Venue_Category	Venue_Count				
0	Outdoors_Recreation	Athletics & Sports	1	15	Professional_Other	Dentist's Office	1
1	Professional_Other	Auditorium	1	16	Food	Dessert Shop	1
2	Shop_Service	Automotive Shop	1	17	Professional_Other	Doctor's Office	1
3	Shop_Service	Bank	2	18	Professional_Other	Event Space	1
4	Shop_Service	Bookstore	1	19	Shop_Service	Furniture / Home Store	1
5	Shop_Service	Boutique	1	20	Shop_Service	Gas Station	2
6	Food	Breakfast Spot	1	21	Shop_Service	Health Food Store	1
7	Professional_Other	Building	3	22	Nightlife	Hookah Bar	1
8	Travel_Transport	Bus Line	2	23	Professional_Other	Hospital	1
9	Food	Café	2	24	Residence	Housing Development	1
10	Arts_Entertainment	Casino	1	25	Food	Italian Restaurant	1
11	Food	Chinese Restaurant	1	26	College_University	Law School	1
12	Nightlife	Cocktail Bar	1	27	Travel_Transport	Light Rail Station	1
13	Shop_Service	Convenience Store	1	28	Food	Malay Restaurant	1
14	Professional_Other	Coworking Space	1	29	Professional_Other	Medical Center	1
30	Professional_Other	Military Base	1				
31	Professional_Other	Non-Profit	1				
32	Professional_Other	Office	1				
33	Outdoors_Recreation	Park	1				
34	Professional_Other	Parking	2				
35	Travel_Transport	Pier	1				
36	Professional_Other	Post Office	2				
37	Food	Restaurant	1				
38	Professional_Other	Temple	1				
39	Outdoors_Recreation	Trail	1				
40	Travel_Transport	Travel Lounge	1				



- Shah Alam

Venue_Type		Venue_Category	Venue_Count				
0	Food	Asian Restaurant	1	15	Outdoors_Recreation	Farm	1
1	Outdoors_Recreation	Bathing Area	1	16	Professional_Other	Fire Station	1
2	Food	Bubble Tea Shop	1	17	Food	Food Court	1
3	Food	Burger Joint	1	18	Food	Food Truck	1
4	Travel_Transport	Bus Station	1	19	Food	French Restaurant	1
5	Travel_Transport	Bus Stop	1	20	College_University	General College & University	1
6	Food	Café	1	21	Outdoors_Recreation	Gym	1
7	College_University	College Administrative Building	1	22	Outdoors_Recreation	Gym / Fitness Center	1
8	College_University	College Classroom	5	23	Professional_Other	Hospital	1
9	College_University	College Lab	1	24	Food	Malay Restaurant	2
10	Professional_Other	Conference Room	1	25	Professional_Other	Meeting Room	1
11	Professional_Other	Convention Center	1	26	Professional_Other	Mosque	1
12	Food	Diner	1	27	Professional_Other	Office	1
13	Professional_Other	Embassy / Consulate	1	28	Professional_Other	Parking	1
14	Professional_Other	Factory	1	29	Residence	Residential Building (Apartment / Condo)	3

30	Food	Sandwich Place	1
31	Outdoors_Recreation	Soccer Field	2
32	College_University	Student Center	1
33	Travel_Transport	Train Station	3
34	Outdoors_Recreation	Vineyard	1
35	Professional_Other	Voting Booth	1
36	Professional_Other	Warehouse	1
37	Arts_Entertainment	Zoo Exhibit	1

- Bangi

	Venue_Type	Venue_Category	Venue_Count				
0	Shop_Service	Adult Boutique	1	15	Shop_Service	Electronics Store	1
1	Arts_Entertainment	Art Gallery	1	16	Arts_Entertainment	Exhibit	1
2	Food	Asian Restaurant	3	17	Professional_Other	Factory	1
3	Food	Breakfast Spot	1	18	Shop_Service	Flower Shop	1
4	Food	Bubble Tea Shop	1	19	Food	Food Court	1
5	Travel_Transport	Bus Line	1	20	Food	Food Truck	1
6	Food	Café	2	21	Outdoors_Recreation	Garden	1
7	Professional_Other	Capitol Building	1	22	Shop_Service	Health Food Store	1
8	College_University	College Academic Building	1	23	Travel_Transport	Hotel	1
9	College_University	College Classroom	1	24	Food	Indian Restaurant	1
10	College_University	College Lab	2	25	Food	Malay Restaurant	3
11	College_University	College Tennis Court	1	26	Professional_Other	Medical Center	1
12	Shop_Service	Convenience Store	1	27	Professional_Other	Mosque	1
13	Food	Creperie	1	28	Shop_Service	Motorcycle Shop	1
14	Food	Diner	1	29	Shop_Service	Pharmacy	1
		30	Professional_Other	Prayer Room			1
		31	Shop_Service	Salon / Barbershop			1
		32	Professional_Other	School			1
		33	Food	Seafood Restaurant			1
		34	Shop_Service	Shopping Mall			1
		35	College_University	Student Center			2
		36	Travel_Transport	Train Station			1

- Bandar Baru Bangi

	Venue_Type	Venue_Category	Venue_Count				
0	Shop_Service	Arts & Crafts Store	1	15	Shop_Service	Drugstore	1
1	Residence	Assisted Living	1	16	Shop_Service	Fabric Shop	1
2	Shop_Service	Bank	1	17	Food	Fast Food Restaurant	1
3	Outdoors_Recreation	Baseball Field	1	18	Shop_Service	Flea Market	1
4	Arts_Entertainment	Baseball Stadium	1	19	Food	Food Court	1
5	Shop_Service	Bookstore	1	20	Food	Food Truck	2
6	Professional_Other	Building	1	21	Food	Halal Restaurant	2
7	Food	Burger Joint	1	22	Travel_Transport	Hotel	1
8	Food	Café	5	23	Travel_Transport	Intersection	1
9	Outdoors_Recreation	Campground	1	24	Food	Korean Restaurant	1
10	Shop_Service	Convenience Store	1	25	Outdoors_Recreation	Lake	1
11	Shop_Service	Department Store	1	26	Food	Malay Restaurant	1
12	Food	Dessert Shop	1	27	Food	Noodle House	1
13	Food	Diner	2	28	Professional_Other	Office	2
14	Professional_Other	Doctor's Office	1	29	Outdoors_Recreation	Park	2
		30	Outdoors_Recreation	Playground			2
		31	Food	Restaurant			2
		32	Shop_Service	Salon / Barbershop			1
		33	Shop_Service	Shopping Mall			3
		34	Outdoors_Recreation	Skate Park			1
		35	Food	Thai Restaurant			1

## Discussion

From the df\_compare, we can conclude the following

- Ampang is suitable for the need of Case Study A as he like party (assuming night time party) and entertainment. Ampang also provide many work places and near with many travels terminal point for easy way to travel around.
- Shah Alam and Bangi is suitable for Case Study B as there are many schools for the young children and work places for the the adults. The eldest also can attend the nearest college or university and travel back home everyday
- Sadly, for Case Study C, none of these 4 areas is suitable for their needs. However, choose between these 4 areas, Shah Alam can provide a better place since it has hospital and based on the maps, is actually quite large area compare to another areas.

## Conclusion

Data science have proven that it can be use as tools for everyday uses.