

Current cluster by environment of the branch of After You Desert Café

Puttavee Charoenwanthanang

16 July 2019

Background

After You is very famous Desert Cafe in Thailand. You can follow via <https://www.afteryoudessertcafe.com/en/>. Now, there are 74 branches in Thailand which is fantastic amount of branch. All branches are going well from first start until now and then After You are listed in stock exchange in Thailand. They try to open new branch with now is not easy to find out the place since they are saturation of new branch.

Problem

In addition, there is some lesson learn in 2017 that they tried to open new branch in the airport however it is not work and have to be closed in 2018 to stop loss of investment.

Interest

In this case, it would be better if data science can be differentiated type of current branch into 4-5 cluster and provide this data to After You Company especially for CEO to understand the point of view in data science data support. With additional analysis with unofficial information e.g. same store sales growth or sales per month in each branch/cluster. It will help to predict the new appropriate location of new branch.

Method - Data Acquisition and Cleaning

Data Source

From the specific question to make cluster of current branch of After You Company

There are available data of all current branch (74 branches) that you can find via <https://www.afteryoudessertcafe.com/en/branches/>.

Data Cleaning & Analysis

Therefore, it is needed to scrape data from this page to get 'Branch Name', Location ('Latitude' and 'Longitude') of each branch. The data can be scraping by using beautifulsoup. However, it is not all in format of table so it is needed to use regular expression to extract out some unused word.

The data is collected in form of table by using Panda. After getting all branches name and location, they will be put into Foursquare to find nearby place (assuming that nearby place = less than 200 meter far away from each branch) and track the type of location that nearby each branch.

Then, summarizing and analysis the type of location in each branch and quantify by percentage of each type in each branch. Then, cluster each branch depend on location nearby each branch.

After that, showing on the map by folium to provide the specific location of each branch with different color to show each cluster.

Results

We got data of all 74 Branches of After You by using BeautifulSoup. The data contains branch name, Latitude, Longitude. This data is ready for the next step to get pin in the folium.

Please note that some Branch Name is show in local language (Thai) which are shown in Unicode. However, there is no method to change to local language since it would be unable to read by reader who is not Thai.

	Branch Name	Latitude	Longitude
0	Central Festival Hatyai	6.9915948	100.4829401
1	\u0e40\u0e0b\u0e47\u0e19\u0e17\u0e23\u0e31\u0e...	6.9915948	100.4829401
2	\u0e2d\u0e35\u0e40\u0e25\u0e1f\u0e40\u0e27\u0e...	13.74488	100.557056
3	ELEVEN BANGKOK	13.74488	100.557056
4	THE MARKET BANGKOK	13.74763	100.541319
5	\u0e40\u0e14\u0e2d\u0e30 \u0e21\u0e32\u0e23\u0e...	13.74763	100.541319
6	\u0e40\u0e0b\u0e47\u0e19\u0e17\u0e23\u0e31\u0e...	7.892787	98.297998
7	CENTRAL PATONG	7.892787	98.297998
8	101 The Third Place	13.68614	100.611678
9	\u0e27\u0e31\u0e19-\u0e42\u0e2d-\u0e27\u0e31\u0e...	13.68614	100.611678
10	ICONSIAM	13.726917	100.511189
11	\u0e44\u0e2d\u0e04\u0e2d\u0e19\u0e2a\u0e22\u0e...	13.726917	100.511189
12	\u0e40\u0e17\u0e2d\u0e21\u0e34\u0e19\u0e2d\u0e...	12.950206	100.888678
13	TERMINAL 21 PATTAYA	12.950206	100.888678
14	\u0e40\u0e0b\u0e47\u0e19\u0e17\u0e23\u0e31\u0e...	13.663684	100.439341
15	CENTRALPLAZA RAMA 2	13.663684	100.439341
16	\u0e40\u0e0b\u0e47\u0e19\u0e17\u0e23\u0e31\u0e...	17.406483	102.800075
17	CENTRALPLAZA UDONTHANI	17.406483	102.800075
18	\u0e40\u0e0b\u0e47\u0e19\u0e17\u0e23\u0e31\u0e...	18.807268	99.018122
19	CENTRALFEFSTIVAL CHIANGMAI	18.807268	99.018122
20	\u0e40\u0e0b\u0e47\u0e19\u0e17\u0e23\u0e31\u0e...	13.747398937284	100.54025077354
21	THE OFFICES AT CENTRALWORLD	13.747398937284	100.54025077354
22	\u0e40\u0e0b\u0e47\u0e19\u0e17\u0e23\u0e31\u0e...	13.876644	100.411027
23	CENTRALPLAZA WESTGATE	13.876644	100.411027
24	\u0e40\u0e14\u0e2d\u0e30\u0e21\u0e2d\u0e25\u0e...	14.974594	102.077241
25	THE MALL KORAT	14.974594	102.077241
26	THE PROMENARD	13.826958	100.676697

27	\u0e40\u0e14\u0e2d\u0e30 \u0e1e\u0e23\u0e2d\u0...	13.826958	100.676697
28	\u0e40\u0e2d\u0e47\u0e21\u0e1a\u0e35\u0e40\u0e...	13.745042	100.529946
29	\u0e40\u0e0b\u0e47\u0e19\u0e17\u0e23\u0e31\u0e...	13.669344	100.634311
30	\u0e40\u0e2d\u0e2a\u0e1e\u0e25\u0e32\u0e19\u0e...	13.766406	100.569521
31	BANGKOK PLAZA @\u0e42\u0e23\u0e07\u0e1e\u0e22\...	13.74879	100.582551
32	\u0e40\u0e14\u0e2d\u0e30 \u0e1e\u0e2d\u0e23\u0...	13.911431	100.550129
33	\u0e40\u0e0b\u0e47\u0e19\u0e17\u0e23\u0e31\u0e...	13.777812	100.475898
34	\u0e0b\u0e35\u0e04\u0e2d\u0e19\u0e2a\u0e41\u0e...	13.694498	100.648361
35	\u0e40\u0e17\u0e2d\u0e21\u0e34\u0e19\u0e2d\u0e...	13.737894	100.560384
36	\u0e2a\u0e40\u0e1b\u0e25\u0e25\u0e4c \u0e1f\u0...	13.989542	100.617846
37	\u0e40\u0e14\u0e2d\u0e30\u0e21\u0e2d\u0e25\u0e...	13.714309	100.407695
38	\u0e17\u0e48\u0e32\u0e21\u0e2b\u0e32\u0e23\u0e...	13.755026	100.4887
39	\u0e40\u0e21\u0e01\u0e32\u0e1a\u0e32\u0e07\u0e...	13.647195443986	100.68056459005
40	\u0e40\u0e14\u0e2d\u0e30 \u0e04\u0e23\u0e34\u0...	13.809624501965	100.44853678847
41	\u0e2d\u0e34\u0e19\u0e17\u0e4c-\u0e2d\u0e34\u0...	13.67455683634	100.54500885308
42	\u0e40\u0e08 \u0e2d\u0e40\u0e27\u0e19\u0e34\u0...	13.733679	100.58136
43	\u0e40\u0e14\u0e2d\u0e30 \u0e04\u0e23\u0e34\u0...	13.811267222969	100.61893198037
44	\u0e2a\u0e22\u0e32\u0e21\u0e1e\u0e32\u0e23\u0e...	13.747092238823	100.53514458836
45	\u0e40\u0e0b\u0e47\u0e19\u0e17\u0e23\u0e31\u0e...	13.817211910538	100.56148756085
46	\u0e25\u0e32 \u0e27\u0e34\u0e25\u0e25\u0e48\u0...	13.779965111953	100.54499564851
47	\u0e40\u0e14\u0e2d\u0e30\u0e21\u0e2d\u0e25\u0e...	13.765148	100.643113
48	\u0e40\u0e0b\u0e47\u0e19\u0e17\u0e23\u0e31\u0e...	13.747398937284	100.54025077354
49	\u0e2a\u0e22\u0e32\u0e21\u0e2a\u0e41\u0e04\u0e...	13.745414	100.534012
50	\u0e2a\u0e35\u0e25\u0e21\u0e04\u0e2d\u0e21\u0e...	13.72852	100.535186
51	MBK CENTER	13.745042	100.529946
52	CENTRALPLAZA BANGNA	13.669344	100.634311
53	ESPLANADE RATCHADAPISEK	13.766406	100.569521
54	BANGKOK PLAZA - Bangkok Hospital	13.74879	100.582551
55	THE PORTAL, Muang Thong Thani	13.911431	100.550129
56	CENTRAL PLAZA PINKLAO	13.777812	100.475898
57	SEACON SQUARE	13.694498	100.648361
58	TERMINAL 21 ASOK	13.737894	100.560384
59	ZPELL - FUTURE PARK RANGSIT	13.989542	100.617846
60	THE MALL BANGKAE	13.714309	100.407695

61	THA MAHARAJ	13.755026	100.4887
62	MEGABANGNA	13.647195443986	100.68056459005
63	THE CRYSTAL RATCHAPRUEK	13.809624501965	100.44853678847
64	INT INTERSECT, Rama 3	13.67455683634	100.54500885308
65	J AVENUE THONGLOR 13	13.733679	100.58136
66	THE CRYSTAL PARK	13.811267222969	100.61893198037
67	SIAM PARAGON	13.747092238823	100.53514458836
68	CENTRALPLAZA LARDPRAO	13.817211910538	100.56148756085
69	LA VILLA AREE (PHAHOLYOTHIN)	13.779965111953	100.54499564851
70	THE MALL BANGKAPI	13.765148	100.643113
71	CENTRALWORLD	13.747398937284	100.54025077354
72	SIAM SQUARE ONE	13.745414	100.534012
73	SILOM COMPLEX	13.72852	100.535186

From the data above the map is created by folium

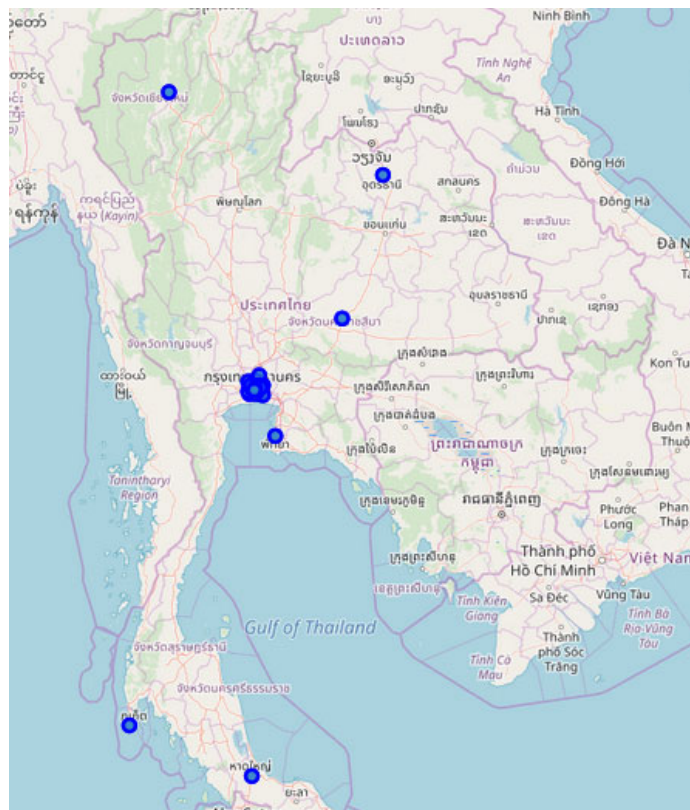


Figure: Thailand Map with After You Branch in blue dot

Data is shown in Thailand Maps and it is found that most of the branches are in Bangkok. In this case, the maps of Bangkok are shown below.

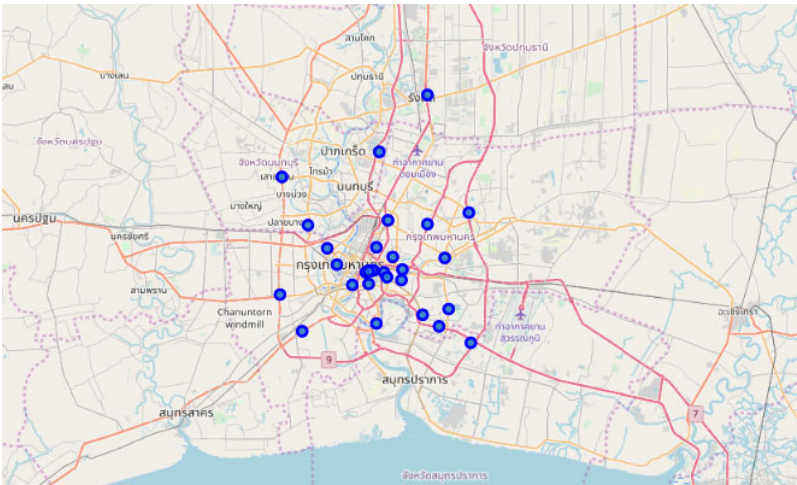


Figure: Bangkok Map with After You Branch in Blue dot

Then, the data of foursquare are loaded and mapping with nearby place which is defined as the place is not longer than 200 meter. Some of the branches with the most common place are shown below.

	BranchName	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	101 The Third Place	Thai Restaurant	Fast Food Restaurant	Book store	Korean Restaurant	Steak house	Super market	Vietnamese Restaurant	Bakery	Bar	Shopping Mall
1	BANGKOK PLAZA - Bangkok Hospital	Coffee Shop	Convenience Store	Kids Store	Café	Fast Food Restaurant	Dessert Shop	Sandwich Place	Flea Market	Food Court	Shopping Mall
2	BANGKOK PLAZA @\u0e42\u0e23\u0e07\u0e1e\u0e22\...	Coffee Shop	Convenience Store	Kids Store	Café	Fast Food Restaurant	Dessert Shop	Sandwich Place	Flea Market	Food Court	Shopping Mall
3	CENTRAL PATONG	Bar	Hotel	Thai Restaurant	Massage Studio	Night club	Pizza Place	Ice Cream Shop	Coffee Shop	Restaurant	Dive Spot

					o						
4	CENTRAL PLAZA PINKLAO	Japanese Restaurant	Café	Chinese Restaurant	Ice Cream Shop	Shabu-Shabu Restaurant	Coffee Shop	Clothing Store	Noodle House	Steakhouse	Dessert Shop

Then, K-means method are run by using 4 clusters with random state = 0. The example of branches with cluster information is shown below.

	BranchName	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Central Festival Hatyai	0	Japanese Restaurant	Clothing Store	Fast Food Restaurant	Coffee Shop	Multiple x
1	\u0040\u00b\u0047\u0019\u0017\u0023\u0031\u00e...	0	Japanese Restaurant	Clothing Store	Fast Food Restaurant	Coffee Shop	Multiple x
2	\u002d\u0035\u0040\u00e2\u001f\u0040\u0027\u00e...	2	Hotel	Bar	Burger Joint	Nightclub	Coffee Shop
3	ELEVEN BANGKOK	2	Hotel	Bar	Burger Joint	Nightclub	Coffee Shop
4	THE MARKET BANGKOK	0	Japanese Curry Restaurant	Japanese Restaurant	Ice Cream Shop	Spa	Chinese Restaurant
5	\u0040\u0014\u002d\u0030\u0021\u0032\u0023\u00...	0	Japanese Curry Restaurant	Japanese Restaurant	Ice Cream Shop	Spa	Chinese Restaurant
6	\u0040\u00b\u0047\u0019\u0017\u0023\u0031\u00e...	2	Bar	Hotel	Thai Restaurant	Massage Studio	Nightclub
7	CENTRAL PATONG	2	Bar	Hotel	Thai Restaurant	Massage Studio	Nightclub
8	101 The Third Place	0	Thai Restaurant	Fast Food Restaurant	Bookstore	Korean Restaurant	Steakhouse
9	\u0027\u0031\u0019-	0	Thai	Fast Food	Bookstore	Korean	Steakho

	BranchName	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
	\u0e42\u0e2d-\u0e27\u0e31\u0e...		Restaurant	Restaurant		Restaurant	use
10	ICONSIAM	0	Thai Restaurant	Dessert Shop	Japanese Restaurant	Clothing Store	Café

With assessed to all place, it can be defined for four clusters as

- Cluster0: Nearby Restaurant
- Cluster1: Nearby Cafe
- Cluster2: Nearby Hotel/Bar/Spa
- Cluster3: Nearby Shopping center

Then the Folium plot is run to provide map data with different cluster.

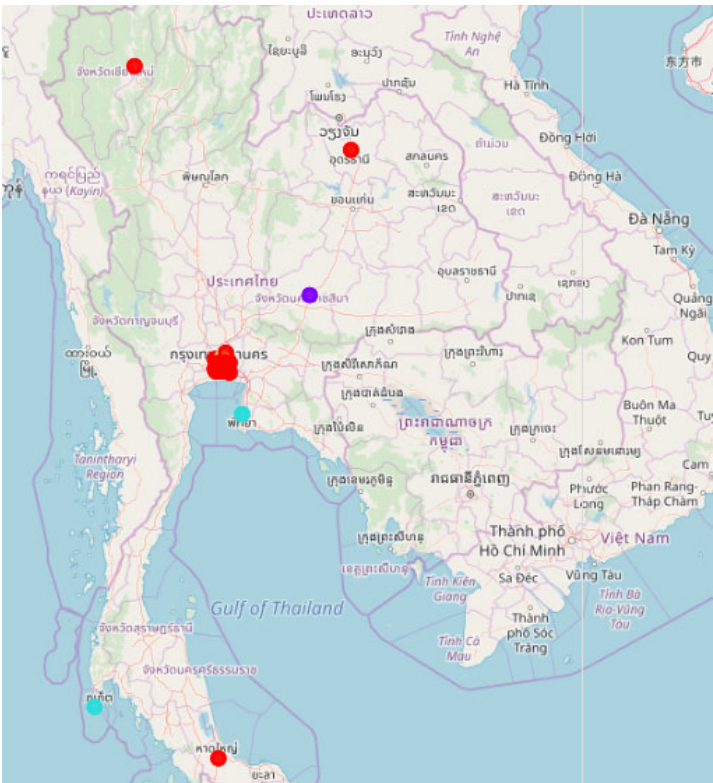


Figure: Thailand Map with After You Branch in different color (Red = Nearby Restaurant; Purple = Nearby Café; Blue = Nearby Hotel/Bar/Spa; Yellow: Nearby Shopping center

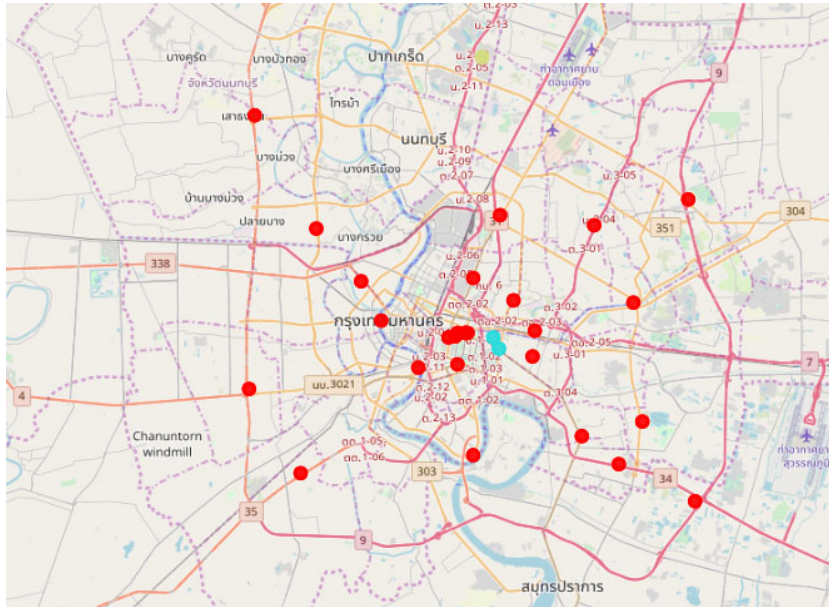


Figure: Bangkok Map with After You Branch in different color (Red = Nearby Restaurant; Purple = Nearby Café; Blue = Nearby Hotel/Bar/Spa; Yellow: Nearby Shopping center

Discussion

From the different between branch in Bangkok and other provinces are cluster. This may be because of Bangkok are main economy city with containing numbers of restaurant. It should be focus on performance of different cluster (Phuket and Nakhon Ratchasima) if it perform well or not?

If it perform well, it is possible to open new branch which have nearby area like these province It should be focus on performance of different cluster in Bangkok (The Portal Mueng Thong, Eleven bangkok, Terminal 21 - Nearby Hotel and shopping center).

It should be compared with other branches (nearby restaurant). Which cluster perform better?

Conclusion

Most of the branches of After You will open nearby restaurant. For non-Bangkok, it also mainly nearby restaurant but some is nearby restaurant and café It should be explored more information compare between the branches nearby restaurant and other types in term of sales. However, the data is limited. It should be proposed to After You Company.