

1) Introduction

Nowadays, the most of people in Thailand like to go to some place that can work and chill (like cafe, coffee shop, etc.) Thus, I would like to point and ranking the areas that has place like this for someone who want to start this kind of business. I scope my work only in Bangkok which is capital of Thailand.

2) Data acquisition and cleaning

I can use a sub-district data of Thailand which contains latitude and longitude of each sub-district to cluster the area and count number of the place like this using Foursquare API which scope data in some categories. I had to filter it to contains only district in Bangkok and made the assumption that means of latitude and longitude in each district is the center.

3) Methodology

With Foursquare API, I can get venues data in Bangkok for each district.

```
print(bangkok_venues.shape)
bangkok_venues.head()
```

(968, 7)

	District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Bang Bon	13.646	100.3700	ส้มตำ บางบอน5	13.650175	100.370430	Thai Restaurant
1	Bang Kapi	13.771	100.6475	ตลาดนัดแฟลตคลองจั่น	13.770105	100.649124	Flea Market
2	Bang Kapi	13.771	100.6475	ร้านข้าวแกงครัวปักษ์ใต้คลองจั่น	13.770482	100.648638	Thai Restaurant
3	Bang Kapi	13.771	100.6475	RABIKA Coffee	13.770872	100.647133	Coffee Shop
4	Bang Kapi	13.771	100.6475	Ko Boo Rod Sing (โกบูร์สซิ่ง)	13.774971	100.648642	Noodle House

After that, I can count number of each venues and normalize it for ranking.

```
bangkok_grouped = bangkok_onehot.groupby('District').mean().reset_index()
bangkok_grouped
```

	District	Accessories Store	American Restaurant	Art Gallery	Art Museum	Asian Restaurant	Auto Garage	BBQ Joint	Badminton Court	Bakery	...	Theme Park Ride / Attraction	Train Station	Vegetarian / Vegan Restaurant
0	Bang Bon	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
1	Bang Kapi	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
2	Bang Khae	0.000000	0.000000	0.000000	0.000000	0.076923	0.076923	0.076923	0.076923	0.000000	...	0.000000	0.000000	0.000000
3	Bang Khen	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
4	Bang Kho Laem	0.000000	0.000000	0.000000	0.000000	0.333333	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000

Then, I'll find the top 10 most common venues for clustering.

	District	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	Bang Bon	Thai Restaurant	Yoga Studio	Food Court	Food & Drink Shop	Food	Flea Market	Film Studio	Filipino Restaurant	Fast Food Restaurant	Farmers Market
1	Bang Kapi	Noodle House	Thai Restaurant	Bowling Alley	Convenience Store	Flea Market	Coffee Shop	Som Tum Restaurant	Fast Food Restaurant	Food & Drink Shop	Food
2	Bang Khae	Som Tum Restaurant	Café	Bus Stop	Dessert Shop	Asian Restaurant	Auto Garage	BBQ Joint	Badminton Court	Night Market	Chinese Restaurant
3	Bang Khen	Convenience Store	Coffee Shop	Yoga Studio	Event Space	Food & Drink Shop	Food	Flea Market	Film Studio	Filipino Restaurant	Fast Food Restaurant
4	Bang Kho Laem	Asian Restaurant	Convenience Store	Sports Club	Noodle House	Yoga Studio	Farmers Market	Food & Drink Shop	Food	Flea Market	Film Studio

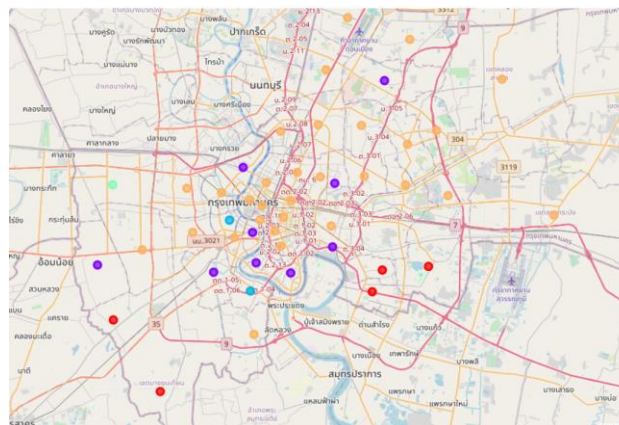
Finally, I used k-means which k=5 which I got cluster labels from clustering.

4) Results

I found each 5 clusters labeled on each district in Bangkok.

	District	District_ID	LAT	LONG	Cluster Labels
0	Bang Bon	1050	13.646000	100.370000	0.0
1	Bang Kapi	1006	13.771000	100.647500	4.0
2	Bang Khae	1040	13.710500	100.397500	4.0
3	Bang Khen	1005	13.867000	100.628000	1.0
4	Bang Kho Laem	1031	13.698667	100.506000	1.0
5	Bang Khun Thian	1021	13.580000	100.414500	0.0
6	Bang Na	1047	13.672000	100.616000	0.0
7	Bang Phlat	1025	13.787000	100.493250	1.0
8	Bang Rak	1004	13.728000	100.523200	4.0
9	Bang Su	1029	13.820000	100.529000	4.0

Then, plot the results with Folium.



There are 5 cluster

First cluster

	District	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	Bang Bon	Thai Restaurant	Yoga Studio	Food Court	Food & Drink Shop	Food	Flea Market	Film Studio	Filipino Restaurant	Fast Food Restaurant	Farmers Market
5	Bang Khun Thian	Thai Restaurant	Soccer Field	Yoga Studio	Food & Drink Shop	Food	Flea Market	Film Studio	Filipino Restaurant	Fast Food Restaurant	Farmers Market
6	Bang Na	Thai Restaurant	Chinese Restaurant	Soccer Field	Beer Garden	Food Truck	Hotel	Dive Shop	Fast Food Restaurant	Dim Sum Restaurant	Food & Drink Shop
32	Phra Khanong	Thai Restaurant	Café	Soup Place	Yoga Studio	Food & Drink Shop	Food	Flea Market	Film Studio	Filipino Restaurant	Fast Food Restaurant
35	Prawet	Thai Restaurant	Noodle House	Halal Restaurant	Café	Coffee Shop	Bakery	Food & Drink Shop	Pool Hall	Duty-free Shop	Dumpling Restaurant

Second cluster

	District	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
3	Bang Khen	Convenience Store	Coffee Shop	Yoga Studio	Event Space	Food & Drink Shop	Food	Flea Market	Film Studio	Filipino Restaurant	Fast Food Restaurant
4	Bang Kho Laem	Asian Restaurant	Convenience Store	Sports Club	Noodle House	Yoga Studio	Farmers Market	Food & Drink Shop	Food	Flea Market	Film Studio
7	Bang Phlat	Convenience Store	Noodle House	Bus Stop	Market	Yoga Studio	Fast Food Restaurant	Food & Drink Shop	Food	Flea Market	Film Studio
14	Chom Thong	Convenience Store	Diner	Noodle House	Café	Massage Studio	Shop & Service	Yoga Studio	Fast Food Restaurant	Food	Flea Market
18	Huai Khwang	Convenience Store	Hotpot Restaurant	Som Tum Restaurant	Hotel	Coffee Shop	Theater	Thai Restaurant	Buffet	Golf Driving Range	Yoga Studio
21	Khlong San	Thai Restaurant	Convenience Store	Noodle House	Chinese Restaurant	Hotpot Restaurant	Asian Restaurant	Badminton Court	Farmers Market	Food	Flea Market
22	Khlong Toei	Spa	Convenience Store	Badminton Court	Farmers Market	Food Court	Food & Drink Shop	Food	Flea Market	Film Studio	Filipino Restaurant
28	Nong Khaem	Lighthouse	Noodle House	Convenience Store	Soup Place	Farmers Market	Food & Drink Shop	Food	Flea Market	Film Studio	Filipino Restaurant
49	Yannawa	Convenience Store	Thai Restaurant	Coffee Shop	Music Store	Yoga Studio	Farmers Market	Food & Drink Shop	Food	Flea Market	Film Studio

Third cluster

	District	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
11	Bangkok Yai	Steakhouse	Thai Restaurant	Convenience Store	Market	Farmers Market	Food & Drink Shop	Food	Flea Market	Film Studio	Filipino Restaurant
36	Rat Burana	Thai Restaurant	Steakhouse	Campground	Food Truck	Food & Drink Shop	Food	Flea Market	Film Studio	Filipino Restaurant	Fast Food Restaurant

Fourth cluster

	District	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
44	Thawi Watthana	Bus Station	Yoga Studio	Department Store	Food Court	Food & Drink Shop	Food	Flea Market	Film Studio	Filipino Restaurant	Fast Food Restaurant

And Final cluster which I can't present it here, I uploaded it on my GitHub repository, you can find it in the link below.

5) Conclusion

I would recommend for anyone who want to start their own business on work and chill place for opening their shops in second cluster which contains a lot of convenience store, sport club (including yoga studio) because it's the place that people use their life and they can stop by your shop everyday.

6) The future directions

The venues data in Bangkok don't have much variety, so there's not much different between each cluster. Thus, for the better clustering, we can expand our work for every province around metropolitan region.