

Finding suitable district for substitute inner district of Bangkok

Pokpong Klongchaipakdee

11 July 2019

1. Introduction/Business Problem

1.1 Background

In Bangkok, capital of Thailand, the land price of this province has been considered one of the highest land prices in Thailand. Especially in the area of inner district of Bangkok which has been in a top 5 highest land price in the country such as Pathum Wan, Bang Rak, Sathon and Watthana District. These areas contain with places that essentially important to many Thai people for their living and their lifestyle such as major workplace and major department store. These areas have been considered desirable by many of Thai People. However, due to the high land price of these areas, most of the people are unable to afford to have a residential place in these areas. While there is possibility that the similar neighborhoods and places to these areas in other parts of Bangkok, The crucial of these areas still high for many of the people. The viable solution that will be effective to most of the people is to locate the area that like the inner district of Bangkok while still be able to travel into inner district of Bangkok conveniently. The areas might be able to locate in other districts of Bangkok.

1.2 Problem

Our purpose is how to find the suitable district for substituting the inner district of Bangkok. By determine the similarity of the high value areas in inner district of Bangkok to other districts in Bangkok throughout the commonality of category in places that locate within the area and the travel distant from those areas to inner district of Bangkok to determine the easiness of travel which will contribute into determine the suitability of that area as well

1.3 Interest

The information from this project will be valuable for many parties, such as real estate related entrepreneur that interest in developing the residential places that can attract customer that in demand for good neighborhood but cannot afford the place in inner district of Bangkok. The developing cost will be lower due to the lower land cost or Real Estate Broker that which to sell their residential places in the district outside of inner district of Bangkok that can be identify as easy to travel to inner district of Bangkok and similar to its.

2. Data acquisition

2.1 Data Source

The data for geocoordinate (Ex. Latitude, Longitude) of districts in Bangkok, route and travel information will come from "here.com". The data of surrounding places of area in each district of Bangkok can be obtain from "Foursquare". The Name of each district in Bangkok can be found in Wikipedia.

2.2 Data Cleaning, Data Processing

Bangkok district name has been parsed from Wikipedia has contain geocoordinate of each district in Bangkok. However, the geocoordinate that has been provide is missing in some district and some are not accurate. For this reason, the name of the district will be the only data from Wikipidia to be used.

For data of places in each district that obtain from “Foursquare” has contain names, geocoordinate and category of that place. The search limit to 100 places in the radius of 600 meters of coordinate of each district. Because the exact geocoordinate data of each place in district is irrelevant in this finding, the data has been dropped and only the data of category of place is use. The number of places in each district has been grouped and count by their category. Then the number of places in each category has been divided by total number of places that has been found in that district in order to compare that category to other district.

Data of distance from each district to others has been obtaining base on shortest distance to travel by car to each other district.

3. Data Analysis

3.1 Determine the parameters for DBSCAN

Since there is no known number of groups of districts to determine the number of cluster, the Density-based spatial clustering of applications with noise (**DBSCAN**) method has been chosen to determine the similarity of each district. The important parameters for this method

1. **Minpoint**: The least amount of datapoint that can be reach within certain radius in order to be consider as core point.
2. **Epsilon**: The value of radius from datapoint

First, the Minpoint has been chosen which **Minpoint** = $\ln(n)$ where “n” is total number of datapoint. In this finding “n” is equal to 50 because Bangkok has a 50 district which make **Minpoint** equal to approximately 4.

The Epsilon will be determined by finding the change in distance between K-nearest datapoint and amount of datapoint or “knee-point”. For this finding, the number of nearest neighbor or “K” is equal to Minpoint. The graph of distance between K-nearest datapoint and amount of datapoint has shown as below

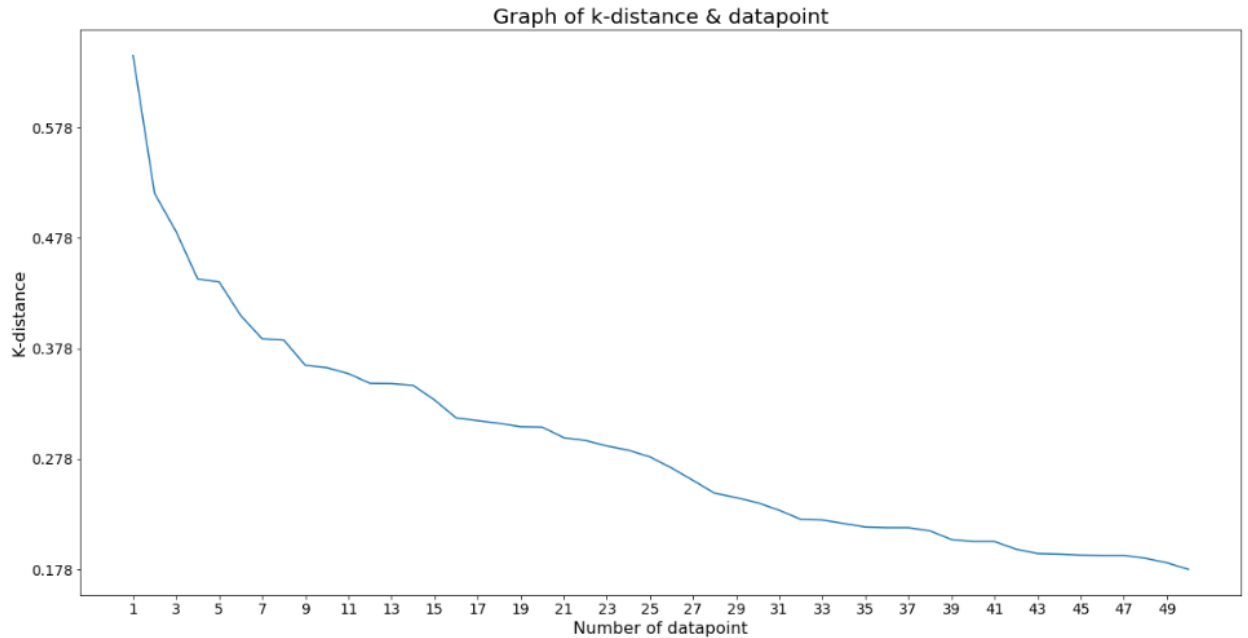


Fig.1 The graph of distance between K-nearest datapoint and amount of data point

From the graph, the knee-point at data point equal to 15 has been chosen and the k-distance at this point is 0.33129689995814543

3.2 Determine the cluster of districts

The data of places in district group by the category has been used to determine the cluster in order to find similarity of each district. The DBSCAN method has been used to determine the cluster based on the Epsilon and Minpoint parameters that has been obtain from previous step. The result shown that there are 1 cluster which has been designate as “cluster0”. The number of districts belong to cluster0 is 40 district and the other 10 districts are outlier which mean it does not belong to any cluster. The list of members of cluster0 and Outliner are shown as below.

Index	District
0	Bang Khae
1	Bang Khun Thian
2	Chom Thong
3	Don Mueang
4	Dusit
5	Khlong Sam Wa
6	Prawet
7	Rat Burana
8	Saphan Sung
9	Thawi Watthana

Fig.2 The table of Outliner district

Index	District
0	Bang Bon
1	Bang Kapi
2	Bang Khen
3	Bang Kho Laem
4	Bang Na
5	Bang Phlat
6	Bang Rak
7	Bang Sue
8	Bangkok Noi
9	Bangkok Yai
10	Bueng Kum
11	Chatuchak
12	Din Daeng
13	Huai Khwang
14	Khan Na Yao
15	Khlong San
16	Khlong Toei
17	Lak Si
18	Lat Krabang
19	Lat Phrao
20	Min Buri
21	Nong Chok
22	Nong Khaem
23	Pathum Wan
24	Phasi Charoen
25	Phaya Thai
26	Phra Khanong
27	Phra Nakhon
28	Pom Prap Sattru Phai
29	Ratchathewi
30	Sai Mai
31	Samphanthawong
32	Sathon
33	Suan Luang
34	Taling Chan
35	Thon Buri
36	Thung Khru
37	Wang Thonglang
38	Watthana
39	Yan Nawa

Fig.3 The table of Cluster0 district

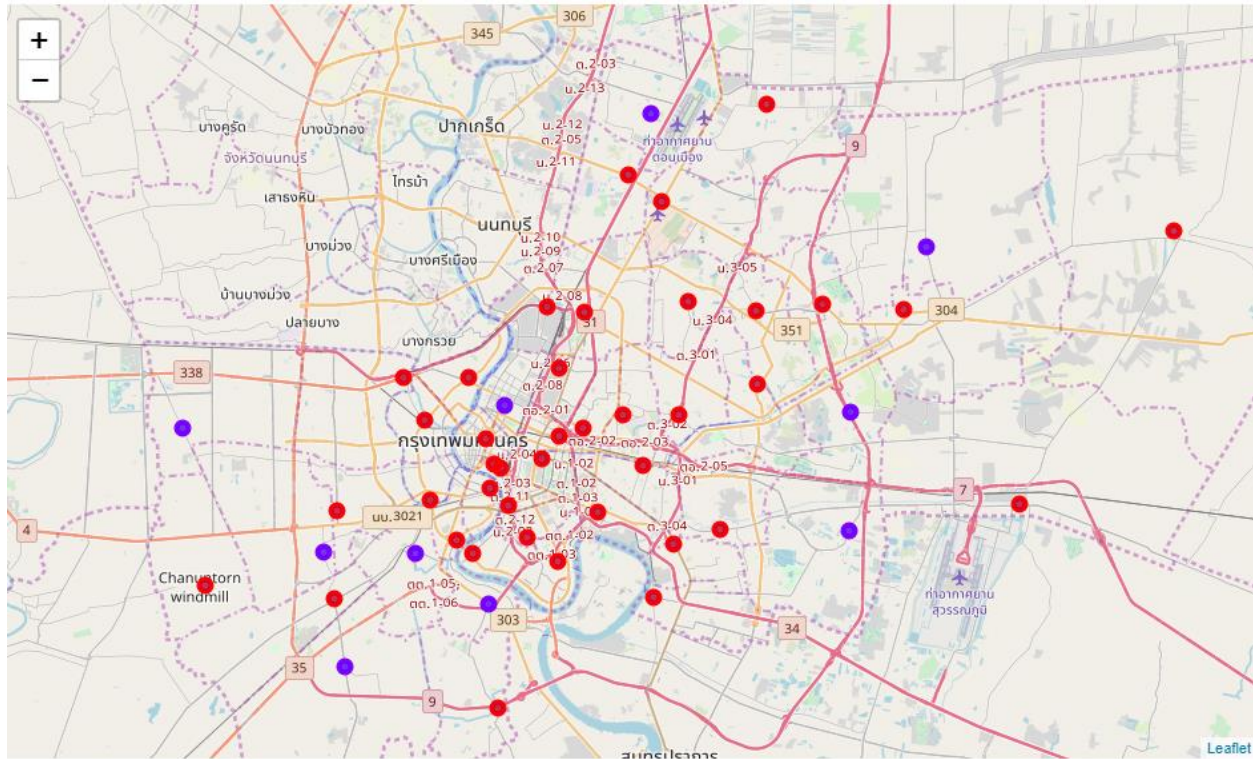


Fig.4 Maps of coordinate of district where red=Cluster 0 , purple = Outliner

3.3 Obtain mean distance to each district for each cluster

In analyzing the cluster0, the inner district of Bangkok within this cluster is Pathum Wan, Bang Rak, Sathon, Wattana District. Since the most suitability district in cluster0 will depend on the distance to these inner district, the district that have shortest average distance from to other 4 inner districts The distance from each district to these 4 districts with exclusion of inner district sorted by mean distance are shown as below.

District Destination	<u>Bang Rak</u>	<u>Pathum Wan</u>	<u>Sathon</u>	<u>Watthana</u>	<u>Mean Distance</u>
District Origin					
Khlong Toei	5,380.00	5,264.00	5,120.00	5,559.00	5,330.75
Pom Prap Sattru Phai	2,815.00	3,332.00	5,043.00	10,748.00	5,484.50
Ratchathewi	6,442.00	3,067.00	8,075.00	7,512.00	6,274.00
Yan Nawa	5,857.00	7,599.00	2,548.00	9,580.00	6,396.00
Samphanthawong	3,969.00	3,744.00	6,197.00	11,902.00	6,453.00
Phra Nakhon	5,077.00	3,645.00	7,305.00	11,619.00	6,911.50
Bang Kho Laem	3,349.00	7,580.00	3,646.00	13,147.00	6,930.50
Din Daeng	8,064.00	4,482.00	9,490.00	6,843.00	7,219.75
Khlong San	3,576.00	7,445.00	5,889.00	12,703.00	7,403.25
Thon Buri	4,784.00	9,198.00	5,081.00	14,456.00	8,379.75
Phra Khanong	10,631.00	9,189.00	10,371.00	5,957.00	9,037.00
Phaya Thai	8,834.00	6,144.00	11,185.00	11,766.00	9,482.25
Huai Khwang	11,600.00	8,245.00	12,520.00	7,575.00	9,985.00
Bangkok Yai	7,086.00	9,066.00	8,689.00	16,213.00	10,263.50
Wang Thonglang	13,130.00	9,775.00	14,050.00	5,556.00	10,627.75
Bang Na	12,189.00	11,578.00	11,807.00	8,346.00	10,980.00
Suan Luang	12,842.00	11,400.00	12,582.00	8,168.00	11,248.00
Bangkok Noi	9,611.00	8,595.00	11,968.00	16,427.00	11,650.25
Bang Sue	12,307.00	10,349.00	14,658.00	16,112.00	13,356.50
Bang Phlat	12,386.00	10,012.00	14,737.00	16,822.00	13,489.25
Chatuchak	12,984.00	11,146.00	15,335.00	16,357.00	13,955.50
Taling Chan	12,278.00	11,262.00	14,635.00	19,094.00	14,317.25
Phasi Charoen	12,622.00	14,583.00	14,139.00	21,749.00	15,773.25
Thung Khru	14,423.00	17,906.00	12,586.00	20,650.00	16,391.25
Bang Bon	13,566.00	17,013.00	13,535.00	22,693.00	16,701.75
Lat Phrao	19,346.00	15,235.00	19,886.00	14,172.00	17,159.75
Bang Kapi	20,250.00	16,895.00	21,170.00	12,346.00	17,665.25
Bang Khen	19,955.00	16,295.00	21,303.00	20,045.00	19,399.50
Buang Kum	22,266.00	18,911.00	23,186.00	14,692.00	19,763.75
Lak Si	20,885.00	19,047.00	23,236.00	23,116.00	21,571.00
Khan Na Yao	26,244.00	22,889.00	27,164.00	18,670.00	23,741.75
Nong Khaem	22,161.00	24,122.00	23,678.00	31,288.00	25,312.25
Min Buri	28,607.00	25,252.00	29,527.00	20,703.00	26,022.25
Sai Mai	28,838.00	25,178.00	30,186.00	26,728.00	27,732.50
Lat Krabang	30,653.00	29,618.00	30,393.00	24,895.00	28,889.75
Nong Chok	43,895.00	40,540.00	44,815.00	35,991.00	41,310.25

Fig.4 The table of distance form each district to inner district of Bangkok

From the table, the top 5 District in cluster0 which has the shortest distance to other 4 inner districts are

1. Khlong Toei
2. Pom Prap Sattru Phai
3. Ratchathewi
4. Yan Nawa
5. Samphanthawong

4. Result and Discussion

From our analysis, based on a similarity of places in each district, there is only 1 cluster and from 50 districts in Bangkok, there are 40 districts that belong to the same cluster which has been designate as "cluster 0". While the other 10 districts have been classified as outlier which does not belong to any cluster. From 4 of the inner districts of Bangkok that we are interested to compare, all of them, Phatum Wan, Bang Rak, Sathon and Wattana district, are in the same cluster, within the cluster0, based on the mean of distance to inner districts which is the member of the same cluster, the suitable districts have been determine by the shortest of mean of the distance. The top 5 suitable districts for this cluster are as follow:

1. Khlong Toei
2. Pom Prap Sattru Phai
3. Ratchathewi
4. Yan Nawa
5. Samphanthawong

However, this finding has based on the similarity of places in each district which has been used to create the cluster of districts. The similarity of places in this finding has been based on the number of places in each category that can be obtain from each district along with the assumption that each category has the same important value. However, the important of each place can be different depend which category of places. For example, the supermarket, gas station or department store can be considered important to daily life than spa. The similarity between districts might be determine by a few places in category that important than high number of places in non-important category. The important of each category of places can be include into calculating the cluster to determine the similarity by weight in on category of place that important to Thai people than other category of places. The result can be different.

5. Conclusion

In this finding, we intend to find the suitable district for substituting the inner district of Bangkok, Based on similarities and distance of each district to inner district of Bangkok. By using DBSCAN method to find the similarity of each district and distance travel by car to another district, we found that 40 out of 50 Districts of Bangkok have similarity to each other. This include the inner district of Bangkok. Using the mean distance to inner district we found that Khlong Toei District is the most suitable for substituting inner district.