

Bangkok, the capital city of Thailand, have various places to travel and also have very low cost of living. So, this city becomes popular among tourlists. Then, this lures contractors who are trying to start their own business. When people want to start their

1. Introducton

own business, they need to explore the places and try to fetch as much information as possible regarding the city. Thus, It would be great for them if there is information which tell them about venues of each neightborhoods in this city. This project will help find venues and classify them for in each district by using k-means clustering to cluster the venues based on the place category. So, business constructors can gain more insights about types porpotion and number of bussinesses in

Bangkok which can help them to determind which types of businesses they should do. Also, tourlists can determind which district they should travel to for the place they prefer.

2.1)Data

2.Data sets and APIs

Beautiful soup web scraping or read html from pandas can be used to scrap the list of 50 districts of Bangkok from the

following url: https://en.wikipedia.org/wiki/List_of_districts_of_Bangkok Note: I'll try to fill the row with NA in Latitude and Longitude by using Geocorder. if I cannot get lititude and longitude for that row, I will drop that row. 2.2)Foursquare API

This project will use Foursquare API as its main data collecting source. This API gives the ability to perform location search, location sharing and details about a business and then will be used to search the nearby venues. After that, vanues will be

used for k-means clustering analysis to know each types of bussinesses 2.3) Python packages

Pandas:Library used to work with data frame NumPy:Library used to handle data in a vector

- JSON: Library used to handle with JSON files
- Geopy: To retrieve Location Data Requests: Library to used handle http requests
- Matplotlib: used for python plotting Sklearn: Library for machine learning
- Folium: Map rendering Library which will be used to visualizing the result from clustering analysis

District(Khet) Code

50

3.Methodology The work flow of this project begins with web scrapping by using one of the functions from pandas from link above.

Thai

บางบอน

105161

Note: this project will consider 'District' column as nightnborhoods

0 Bang Bon

	1	Bang Kapi	06	บางกะปิ	148465	2	13.765833	100.647778
	2	Bang Khae	40	บางแค	191781	4	13.696111	100.409444
	3	Bang Khen	05	บางเขน	189539	2	13.873889	100.596389
	4	Bang Kho Laem	31	บางคอแหลม	94956	3	13.693333	100.5025
A	fte	e data set have no N r that, rows which ha e after this process	ave 'Lat	itude' or 'Longit		mn) with NaN value will be filled by using	geocorder. T	hen, the NaN

Population No. of Subdistricts (Khwaeng)

Latitude

NaN

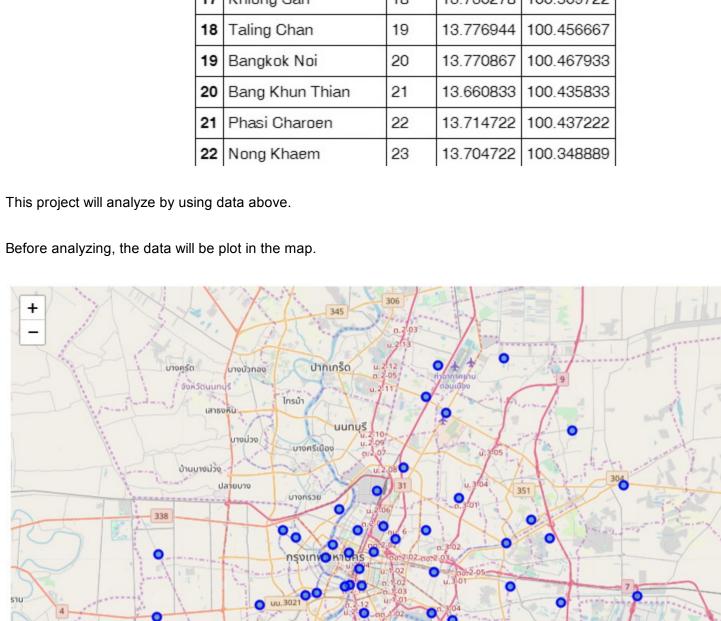
Longitude

NaN

The next process is retriving the column used to analyze which are 'District', 'Code', 'Latitude', 'Longitude'. However, when plotting these location to the map, the row named Bang Bon in district column has wrong location. So, this row will be dropped.

District Code Latitude Longitude

0	Phra Nakhon	01	13.764444	100.499167
1	Dusit	02	13.776944	100.520556
2	Nong Chok	03	13.855556	100.8625
3	Bang Rak	04	13.730833	100.524167
4	Bang Khen	05	13.873889	100.596389
5	Bang Kapi	06	13.765833	100.647778
6	Pathum Wan	07	13.744942	100.5222
7	Pom Prap Sattru Phai	08	13.758056	100.513056
8	Phra Khanong	09	13.702222	100.601667
9	Min Buri	10	13.813889	100.748056
10	Lat Krabang	11	13.722317	100.759669
11	Yan Nawa	12	13.696944	100.543056
12	Samphanthawong	13	13.731389	100.514167
13	Phaya Thai	14	13.78	100.542778
14	Thon Buri	15	13.725	100.485833
15	Bangkok Yai	16	13.722778	100.476389
16	Huai Khwang	17	13.776667	100.579444
17	Khlong San	18	13.730278	100.509722
18	Taling Chan	19	13.776944	100.456667
19	Bangkok Noi	20	13.770867	100.467933
20	Bang Khun Thian	21	13.660833	100.435833
21	Phasi Charoen	22	13.714722	100.437222
22	Nong Khaem	23	13.704722	100.348889



0.28 0.26

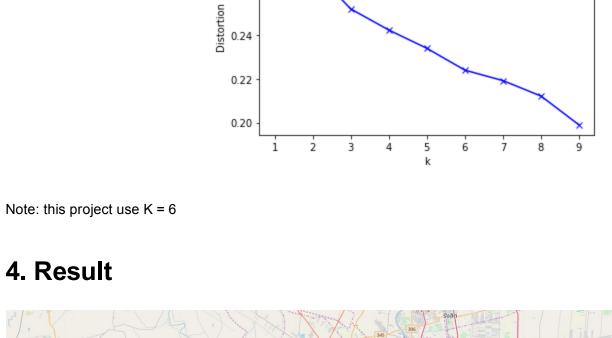
Elbow analysis is used to find an optinum number of K for K-means clustering (the number is not obvious in this case)

Four-square api is used to retrieve top vanues for each neighborhood. K-means clustering is used to cluster each district by their top venues.

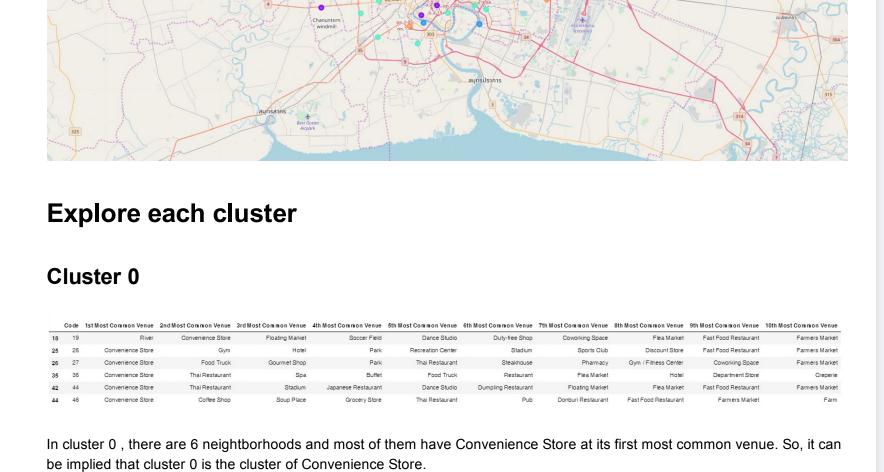
สมุทรปราการ

The Elbow Method showing the optimal k

Leaflet | Data by @ OpenStreetMap, under ODbL.



4. Result



19 20 Noodle House Thai Restaurant Food Truck Caté Park Dessert Shop Coffee Shop Asian Restaurant Auto Garage Chinese Restaurant
22 23 Antique Shop Playground Coffee Shop Market Dessert Shop Steakhouse Noodle House Veterinarian Duty-fee Shop Flea Market
27 28 Noodle House Asian Restaurant Chinese Restaurant Thai Restaurant Dessert Shop Seafood Restaurant Coffee Shop Bakery Convenience Store Supermarket 27 28 Noodle House Asian Restaurant Chinese Restaurant Thai Restaurant Dessert Shop Seafood Restaurant Coffee Shop Bakery Convenience Store Supermarket
30 31 Noodle House Thai Restaurant Chinese Restaurant Vietnamese Restaurant Supermarket Hotpot Restaurant Fast Food Restaurant Convenience Store Comfort Food Restaurant Seafood Restaurant Noodle House Asian Restaurant Coffee Shop Convenience Store Thai Restaurant Chinese Restaurant Food Truck Dessert Shop Pizza Place Clothing Store

In cluster 1, there are 12 neightborhoods and most of them have Noodle House at its first most common venue. So, it can be

Hotel Bar

Convenience Store

Department Store Dim Sum Restaurant

implied that cluster 1 is the cluster of Noodle House. Cluster 2

Cluster 3

Convenience Store

Thai Restaurant Japanese Restaurant

31 32 Convenience Store Comfort Food Restaurant Halal Restaurant

Cluster 1

	Co de	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 Venu
10	11	Thai Restaurant	Flea Market	Restaurant	Asian Restaurant	Café	Steakhouse	Noodle House	Fast Food Restaurant	Farmers Market	Fan
3	24	Thai Restaurant	Paper / Office Supplies Store	Bistro	Coffee Shop	Hotpot Restaurant	Intersection	Chinese Restaurant	Food Truck	Noodle House	Veterinaria
8	29	Noodle House	Coffee Shop	Thai Restaurant	Bar	Train Station	Seafood Restaurant	Hotpot Restaurant	Badminton Court	Flea Market	Fast Food Restaura
5	47	Thai Restaurant	Noodle House	Asian Restaurant	Seafood Restaurant	Satay Restaurant	Restaurant	Café	Coffee Shop	Diner	Dim Sum Restaurar
		•	are 4 neight			hem have T	hai Restaur	ant at its firs	t most com	mon venue.	So, it can t

Code 1st Most Common Venue 2nd 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 01th Most Common Venue 8th Most Common Venue 9th Most Common Venue 10th Most Common Venue 01th Most Common Venue 10th M

5 06 Convenience Store Thail Restaurant Coffee Shop Soccer Field Park Noodle House Ramen Restaurant Museum Multiplex Massage Studio 6 07 Noodle House Convenience Store Thai Restaurant Asian Restaurant Chinese Restaurant Japanese Restaurant Dessert Shop Hotpot Restaurant Som Turn Restaurant Sporting Goods Shop 8 09 Convenience Store Fast Food Restaurant Shopping Mall Italian Restaurant Café Hotel Coffee Shop Pharmacy Pizza Place Residential Building (Apartment / Condo) 9 10 Department Store Thail Restaurant Beach Farm Veterinarian Duty-free Shop Fondue Restaurant Floating Market Flea Market Fast Food Restaurant
11 12 Fast Food Restaurant Chinese Restaurant Japanese Restaurant Electronics Store Bike Rental / Bike Share

Flea Market

Supermarket

Beer Garden Italian Restaurant Dumpling Restaurant

Flea Market Fast Food Restaurant

Food Court

Soccer Stadium

Flea Market Fast Food Restaurant Farmers Market Farm Electronics Store

Hostel Noodle House Thai Restaurant Caté Hotel nience Store Dessert Shop Shopping Mall Jewelry Store Other Repair Shop

Coffee Shop

11 12 Fast Food Restaurant Coffee Shop Brewery Cafe Thai Restaurant Hotoot Restaurant Chinese Restaurant Japanese Restaurant Electronics Store Bike Rental / Bike Share

12 13 Art Gallery Hotel Bar Hotel Restaurant Japanese Restaurant Coffee Shop Creperie Sake Bar Pool

13 14 Café Coffee Shop Thai Restaurant Japanese Restaurant Sushi Restaurant Bar Dessert Shop Som Tum Restaurant Burger Joint Bakery

17 18 Hotel Bar Noodle House Department Store Clothing Store Hotel Chinese Restaurant Coffee Shop Dessert Shop Diner Italian Restaurant

20 21 Thai Restaurant Japanese Restaurant Restaurant Bakery Pizza Place BBQ Joint Noodle House Coffee Shop Shopping Mall Food Court

21 22 Coffee Shop Japanese Restaurant BBQ Joint Hotpot Restaurant Fast Food Restaurant Shabu-Shabu Restaurant Department Store Pizza Place Steakhouse Optical Shop

24 25 Coffee Shop Convenience Store Thai Restaurant Restaurant Restaurant Restaurant Restaurant Restaurant Shabu-Shabu Restaurant Cocktail Bar Café Bus Station Bakery

29 30 Coffee Shop Flea Market Thai Restaurant Cocktail Bar Café Bed & Breakfast Noodle House Badminton Court Convenience Store Fast Food Restaurant 24 25 Corres Ship Onter Ship Onte 39 40 Coffee Shop Café BBQ Joint Thai Restaurant Fast Food Restaurant Noodle House Shopping Mall Convenience Store Dim Sum Restaurant Diner

Shopping Mall

Steakhouse

Grocery Store

Diner Dumpling Restaurant

Cluster 4

In cluster 3, there are 23 neightborhoods and most of them have Cafe or Coffee shop at its first most common venue. So, it can be implied that cluster 3 is the cluster of Cafe and Coffee.

Code 1st Most Common Venue 2nd 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 10th

In cluster 4, there is only 1 neightborhood and it has Seafood Restaurant at its first most common venues. So, it can be implied

Code 1st Most Common Venue 2nd 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 10th

41 42 Bar Food & Drink Shop Grocery Store Thail Restaurant Arts & Entertainment Food Truck Noodle House Veterinarian Flea Market Fast Food Restaurant 43 45 Palace Dessert Shop Cafe History Museum Buddhist Temple Shopping Mall Bakery Soup Place Japanese Restaurant Soiritual Center 47 49 Thai Restaurant Convenience Store Caté Asian Restaurant Vietnamese Restaurant Som Tum Restaurant Japanese Curry Restaurant Food Service Juice Bar Food Court

that cluster 4 is the cluster of Seafood Restaurant. **Cluster 5**

Food Stand Dumpling Restaurant

In cluster 5, there are 2 neightborhoods and most of them have Convenience Store at its first most common venue; however, these neightborhoods also have same venues between 6th-10th most common venues.

5. Discussion

6. Conclusion

Count neightborhood for each cluster

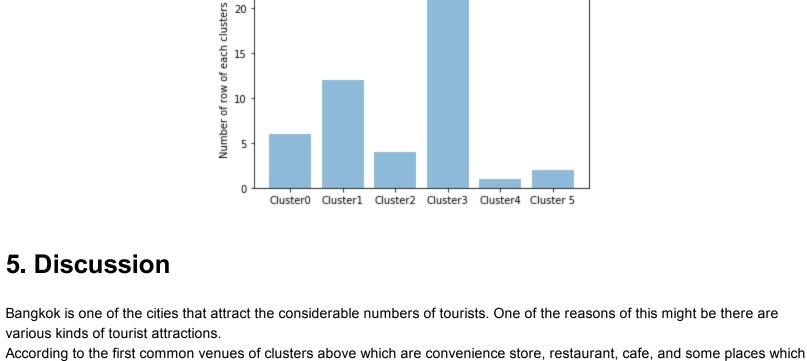
Count each cluster

can travel to. Most of these vanues can attract the tourists.

choosing from the venues around that neightborhood.

With the result above, this project can benefit both tourists and business constructors.

business at by judging from compettion of each types of business in each cluster.



when tourists travel to this city Bangkok, they can easier determine which neighborhood they should live or travel to by

Also, business constructors can which types of businesses they should do of which neighborhood they should set their