Battle of Neighborhoods

Bangkok, Thailand 🇹🇭

1

Introduction

"It would be great for business constructors and tourists if there is information which tell them about venues of each neighborhood in this city."

This project will help find venues and classify them for in each neighborhood.

2

Data sets and APIs

Dataset

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

	District(Khet)	Code	Thai	Population	No. ofSubdistricts(Khwaeng)	Latitude	Longitude
0	Bang Bon	50	บางบอน	105161	4	NaN	NaN
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

Foursquare API

- Foursquare API gives the ability to perform location search, location sharing and details about a business.

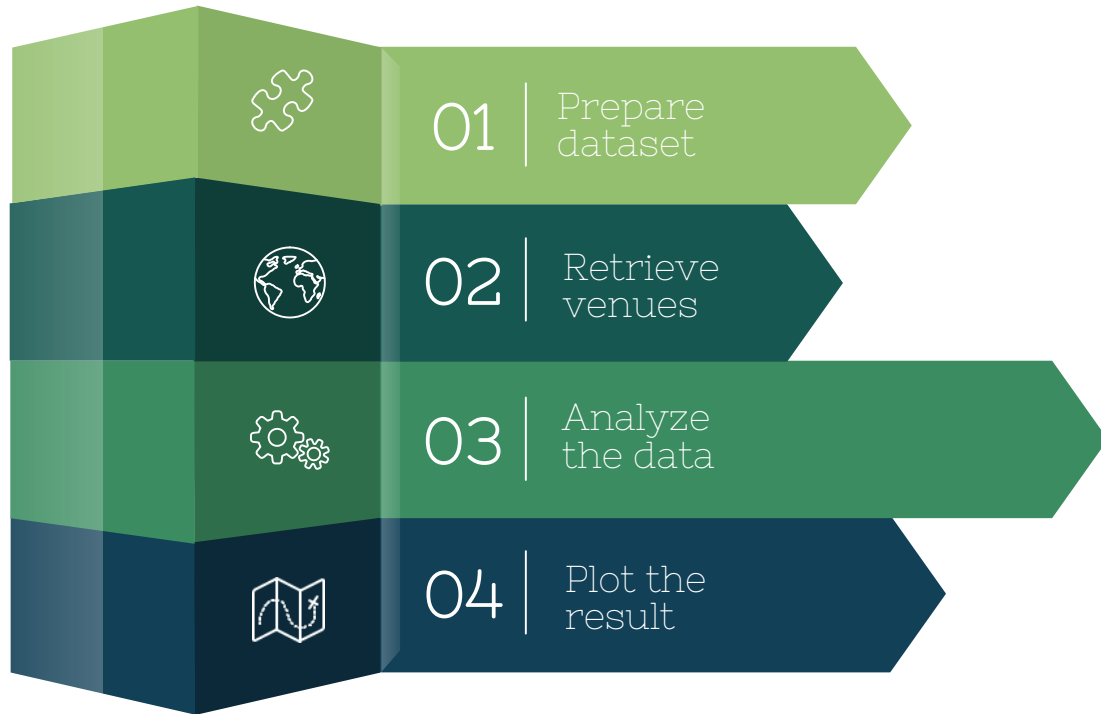- It will be used to search the nearby venues.

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

3

Methodology

Work flow:

01 | Prepare dataset

02 | Retrieve venues

03 | Analyze the data

04 | Plot the result

01 Prepare the data

Use Geocorder package to fill latitude and longitude for the data

Drop the rows which have no values in latitude and longitude

Plot locations of each row into the map and drop when it has wrong coordinate

01 Prepare the data

- This is a result data frame which will be used from now.

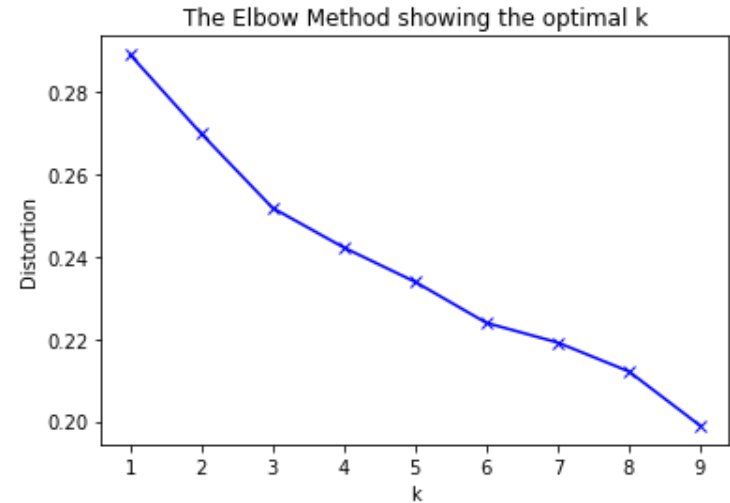|	District	Code	Latitude	Longitude
0	Phra Nakhon	01	13.764444	100.499167
1	Dusit	02	13.776944	100.520556
2	Nong Chok	03	13.855556	100.8625
3	Bang Rak	04	13.730833	100.524167
4	Bang Khen	05	13.873889	100.596389
5	Bang Kapi	06	13.765833	100.647778
6	Pathum Wan	07	13.744942	100.5222
7	Pom Prap Sattru Phai	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

02 Retrieve venues

- Foursquare API retrieve top vanues for each neighbourhood.
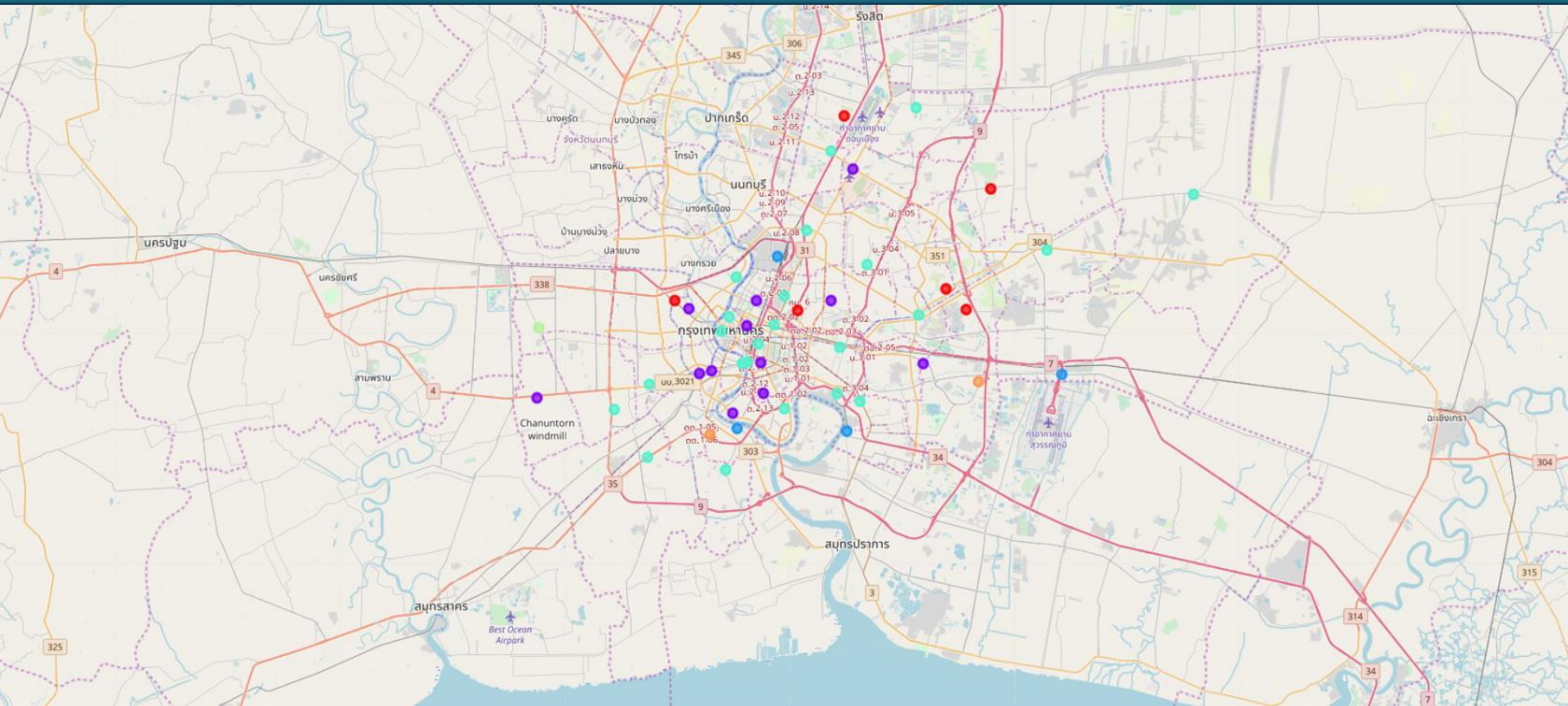
- K-means clustering is used to cluster each district by their top venues.

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

The Elbow Method showing the optimal k

This project use K = 6
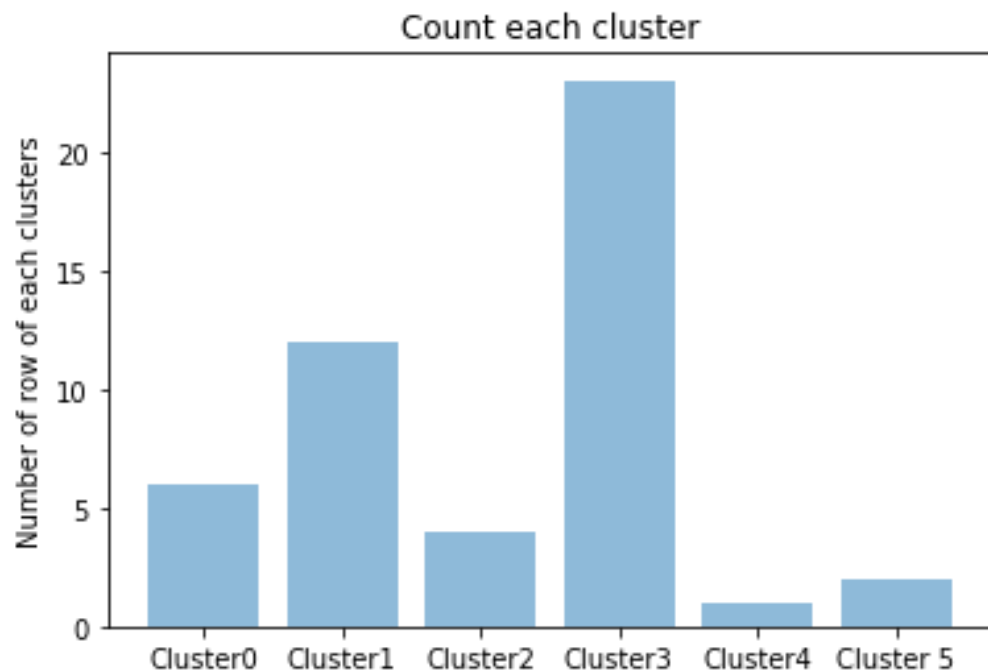
Plot the result

4

Result

Explore each cluster

- Cluster 0 is the cluster of Convenience Store.
- Cluster 1 is the cluster of Noodle House.
- Cluster 2 is the cluster of Thai Restaurant.
- Cluster 3 is the cluster of Cafe and Coffee.
- cluster 4 is the cluster of Seafood restaurant.
- Cluster 5 is similar to cluster 0 but have the same venues in 6th–10th most common venues. So, it is divided into another cluster.

Count neighborhoods for each clusters

5

Discussion

Discussion

- Bangkok is one of the cities that attract the considerable numbers of tourists. One of the reasons of this might be there are various kinds of tourist attractions.

- According to the first common venues of clusters above which are convenience store, restaurant, cafe, and some places which can travel to. Most of these venues can attract the tourists.

6

Conclusion

Conclusion

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

- when tourists travel to this city Bangkok, they can easier determine which neighborhood they should live or travel to by choosing from the venues around that neighborhood.

- Also, business constructors can which types of businesses they should do of which neightborhood they should set their business at by judging from competition of each type of business in each cluster.