IBM
Applied Data Science Capstone – Week 5
Report - Final

The Battle of Neighborhoods

Introduction:

The city of Bangkok, a capital of Thailand is the most populous city of the country which occupies 1,568.7 square kilometers (605.7 sq. mi) has a population of over eight million. This city also presents diversities in both culture and food choices.

Bangkok is one of the best places for people to eat tasty and authentic Thai food. So many restaurants and street vendors selling Thai food are welcoming locals and foreigners every single day.

I have been in Bangkok for years and have tried a lot of Thai restaurants ranging from fancy restaurants to food streets. Individual area has its own unique style and taste. Famous areas such as Ekkamai and Khaosan Road attract foreigners and even locals with tasty Thai food.

Here we will give you our favorite places to eat authentic Thai food. These are places that not many tourists visit, mostly Thais and expats living in Bangkok.

For this capstone project, we will look into the restaurant categories with various food menus and classify into 3 clusters. Then we will make a recommendation on what types of restaurant are popular for clusters, where we recommend that entrepreneur open and where contractors would setup their office.

Business Problem

When people decide where to live or work, the very first reason they can think of is choices of food and restaurants. Of course, we can't deny that food is a key factor for us to decide where to stay or even travel. Places with best food items will attract more business units, offices, workers or even a lot of jobs to the area.

Thus, we need to search for the perfect and right place to solve our business problems as the following:

- To search for the best locations with popular Thai restaurants

- To list the locations of all areas and streets in Bangkok which have great Thai restaurants.
- To find the best place to live close to fancy food choice
- To recommend entrepreneurs or vendors about locations with potential food business and companies that want to find new location based on choice of foods

Target audience

Target for this project can range from vendors or entrepreneurs who want to open new food business or restaurants. This also includes business units that decide to open a new branch in Bangkok as well as people who make decision where to live based on food choices.

Data Section:

The city of Bangkok, a capital of Thailand is the most populous city of the country which occupies 1,568.7 square kilometers (605.7 sq. mi) has a population of over eight million. This city also presents diversities in culture and food choices.

Picture1: Bangkok Montage
Source: https://upload.wikimedia.org/wikipedia/commons/b/b4/Bangkok_montage_3.jpg

With numerous food choices, there are many food markets, restaurants ranging from local eateries to world-class gourmet ones. This city is indeed a hub for food lovers.

To explore more choice, the following data will be used for analysis:

List of Bangkok areas and neighborhoods

- **Data source**: https://en.wikipedia.org/wiki/Category:Neighbourhoods_of_Bangkok

Description: This set of data contains the required information and will be used to explore all neighborhoods of this city

- Thai restaurants and food streets in Bangkok neighborhoods.

Description: To use API we will be able to get all the streets in each neighborhood. Then we will be able to segment Thai restaurants.

- **Data source**: Foursquare API

GeoSpace data

Description: This geo space data will get us all areas, boundaries and streets which help us visualize the map.

- Getting Latitude and Longitude data via Geocoder package
- Using Foursquare API to get venue data related to these neighborhoods

Methodology

1. Collect data of Bangkok areas and neighborhoods from https://en.wikipedia.org/wiki/Category:Neighbourhoods_of_Bangkok
2. Use Foursquare API to get neighborhood location, latitude and longitude data using geocoder library as the following:

[22]:

	Neighborhood	Latitude	Longitude
0	Ban Bat	13.749176	100.507046
1	Ban Chang Lo	13.755101	100.478444
2	Ban Khamin	13.751690	100.483570
3	Ban Khrua	13.822480	100.514350
4	Ban Mo, Bangkok	13.696820	100.498750
5	Ban Noen, Bangkok	13.684080	100.498130
6	Bang Krabue, Bangkok	13.789170	100.514860
7	Bang Lamphu	13.790270	100.515210
8	Bobae Market	13.752930	100.517060
9	Captain Bush Lane	13.728446	100.514330
10	Chan Road	13.708545	100.539839
11	Chinatown, Bangkok	13.737710	100.512930
12	Clinton Plaza	13.760390	100.624200
13	Dao Khanong	13.697619	100.482470
14	Ekkamai Road	13.730497	100.586341
15	Fai Chai Junction	13.754940	100.469540

3. Categorize vendor's names and categories for each neighborhood area

```
array(['Asian Restaurant', 'Noodle House', 'Flea Market', 'Dessert Shop',
       'Park', 'Fast Food Restaurant', 'Shopping Mall', 'Thai Restaurant',
       'Seafood Restaurant', 'Hostel', 'Chinese Restaurant',
       'Dim Sum Restaurant', 'Market', 'Convenience Store',
       'History Museum', 'Café', 'Hobby Shop', 'Night Market', 'Bus Stop',
       'Camera Store', 'Bakery', 'Farmers Market', 'Juice Bar',
       'Arts & Crafts Store', 'Coffee Shop', 'Hotpot Restaurant',
       'Supermarket', 'Food Court', 'Sushi Restaurant',
       'Japanese Restaurant', 'Shabu-Shabu Restaurant', 'Bar', 'Pharmacy',
       'Som Tum Restaurant', 'Hotel Bar', 'Ice Cream Shop', 'Hotel',
       'Steakhouse', 'Jewelry Store', 'Diner', 'Art Gallery', 'Pub',
       'BBQ Joint', 'Department Store', 'Restaurant', 'Big Box Store',
       'Museum', 'Multiplex', 'Ramen Restaurant', 'Tea Room'],
      dtype=object)
```

	Neighborhood	Latitude	Longitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	Ban Bat	13.749176	100.507046	Nakhon Sanook (ข้าวหมูแดง นครสนุก)	13.747888	100.507420	Asian Restaurant
1	Ban Bat	13.749176	100.507046	ยินดี นู้ดเดิ้ล	13.747940	100.508970	Noodle House
2	Ban Bat	13.749176	100.507046	Worachak Chicken Noodle (ก๋วยเตี๋ยวคั่วไก่วรจักร)	13.749234	100.506615	Noodle House
3	Ban Bat	13.749176	100.507046	ก๋วยเตี๋ยวคั่วไก่ เจ๊เค็ง & เจ๊จิ๋ม	13.749389	100.510210	Noodle House
4	Ban Bat	13.749176	100.507046	โต ก๋วยจั๊บ	13.751322	100.504376	Noodle House

	Neighborhood	Latitude	Longitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	Ban Bat	13.749176	100.507046	Nakhon Sanook (ข้าวหมูแดง นครสนุก)	13.747888	100.507420	Asian Restaurant
1	Ban Bat	13.749176	100.507046	ยินดี นู้ดเดิ้ล	13.747940	100.508970	Noodle House
2	Ban Bat	13.749176	100.507046	Worachak Chicken Noodle (ก๋วยเตี๋ยวคั่วไก่วรจักร)	13.749234	100.506615	Noodle House
3	Ban Bat	13.749176	100.507046	ก๋วยเตี๋ยวคั่วไก่ เจ๊เค็ง & เจ๊จิ๋ม	13.749389	100.510210	Noodle House
4	Ban Bat	13.749176	100.507046	โต ก๋วยจั๊บ	13.751322	100.504376	Noodle House

	Neighborhoods	Accessories Store	American Restaurant	Arcade	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant	Auto Garage	Automotive Shop
0	Ban Bat	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.166667	0.000000	0.0000
1	Ban Chang Lo	0.00	0.000000	0.000000	0.000000	0.000000	0.035714	0.000000	0.000000	0.0000
2	Ban Khamin	0.00	0.000000	0.000000	0.027027	0.000000	0.000000	0.027027	0.000000	0.0000
3	Ban Khrua	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0000
4	Ban Mo, Bangkok	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.076923	0.000000	0.0000
5	Ban Noen, Bangkok	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.076923	0.000000	0.0000
6	Bang Krabue, Bangkok	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0000
7	Bang Lamphu	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0000
8	Bobae Market	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0000
9	Captain Bush Lane	0.00	0.000000	0.000000	0.063830	0.000000	0.000000	0.000000	0.000000	0.0000
10	Chan Road	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.018519	0.000000	0.0000

4. Apply K-Means algorithm to cluster the restaurants:

```
kclusters = 3

mum_clustering = mum_rest.drop(["Neighborhoods"], 1)

kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(mum_clustering)

kmeans.labels_[0:10]
```

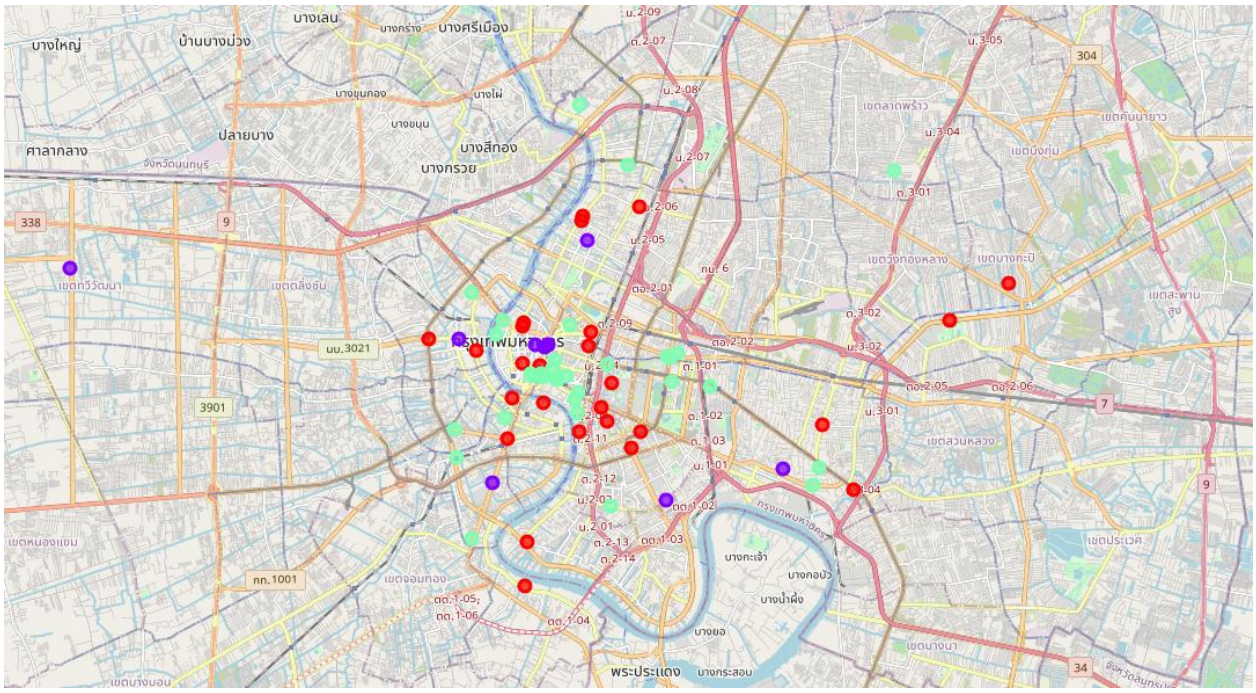
```
array([2, 1, 0, 2, 0, 0, 0, 0, 0, 0], dtype=int32)
```

```
mum_merged = mum_rest.copy()
mum_merged["Cluster Labels"] = kmeans.labels_
```

```
mum_merged.rename(columns={"Neighborhoods": "Neighborhood"}, inplace=True)
mum_merged.head()
```

	Neighborhood	Thai Restaurant	Cluster Labels
0	Ban Bat	0.023810	2
1	Ban Chang Lo	0.142857	1
2	Ban Khamin	0.054054	0
3	Ban Khrua	0.000000	2
4	Ban Mo, Bangkok	0.076923	0

5. Plot the map using Folium to visualize clusters



Result

There are 3 clusters for Thai restaurants as:

Purple: Low number of Thai restaurants

Red: Moderate number of Thai restaurants

Green: High number of Thai restaurants