Fine Great Food @ BKK
Bangkok Food & Restaurant Data-assisted Suggestion

Phumrapee Pisutsin

Project Title

Fine Great Food @ BKK: Bangkok Food & Restaurant Data-assisted Suggestion

Introduction

Thailand is one of the world's most visited travelling destination, due to its fame on food, climate, and natural scenery. In 2019, Thailand attracts around 40 million of visitors, placing it amongst the top 5 countries with the most annual visitors. Bangkok, Thailand's capital city, has a large number of tourists around 20 million each year, including 10 million of local resident living together in this city. The number of people living temporarily and permanently in Bangkok is extremely high; therefore, there is a large consumption, social activities, and usage on food services and restaurants.

Business Problem

There are currently mainly two main problems associated with food and restaurant services industries that this project would like to address. First, due to the ongoing global COVID-19 pandemic (as of May 2019), people all around the world are practicing safe distancing measure and be more aware of eating at restaurant, gathering with families and friends, and travelling around their own cities or other countries. With the hope that the economy would recover back to normal, there is a huge demand from both public and private sectors to encourage people not to reduce their consumption from food and beverage services, once it is safe to do so. Second, people, when they gather, often have a tough time to decide the location for eating food together. If there is a service to suggest the right restaurant at the right location that is close to the rail public transportation station, it could effectively resolve this issue and give convenience to people that gather. This will not only encourage the residents and tourist to be more convenient in using food and restaurant services, but also help the business, especially in the food and restaurant industry, to get more income and recover from the recession. Therefore, the economy can recover faster with this solution benefiting both the people and the nation. This project will work on a data-assisted restaurant suggestion system. The project would suggest the restaurant at the right location for people who would like to gather.

Data

There are two types of data used in this project. The types of data are as follows:

1. Location (latitude and longitude) of stations of rail transportation service in Bangkok
2. Information about restaurants and food services, such as, rating, location (latitude, longitude)

Currently there are three common rail public transportation services used by residents of Bangkok. Name of these services are BTS, MRT, and Airport Link. I have acquired the data (latitude and longitude) about stations of each rail public transportation service in Bangkok using Google Maps. The data is stored in the csv format. The detail about each column is as follow

- Service: type of service of the station
- Station: station name in English
- Lines: the color of the rail line that the station is located on (grey color if the station is a hub connecting more than one rail line together at the same station.
- Latitude: latitude of the station
- Longitude: longitude of the station
- Is Operating: Boolean variable about whether the station is still in use

Data Source: Google Maps

Service	Station	Lines	Latitude	Longitude	Is Operating
BTS	Siam	grey	13.75	100.53	TRUE
BTS	National Stadiun	blue	13.75	100.53	TRUE
BTS	Ratchadamri	blue	13.74	100.54	TRUE
BTS	Sala Daeng	blue	13.73	100.53	TRUE
BTS	Chong Nonsi	blue	13.72	100.53	TRUE
BTS	Surasak	blue	13.72	100.52	TRUE
BTS	Saphan Taksin	blue	13.72	100.51	TRUE
BTS	Krung Thon Buri	blue	13.72	100.5	TRUE
BTS	Wongwian Yai	blue	13.72	100.5	TRUE
BTS	Pho Nimit	blue	13.72	100.49	TRUE
BTS	Talat Phlu	blue	13.71	100.48	TRUE
BTS	Wutthakat	blue	13.71	100.47	TRUE
BTS	Bang Wa	blue	13.72	100.46	TRUE
BTS	Kasetsart Univer	green	13.84	100.58	TRUE
BTS	Sena Nikhom	green	13.84	100.57	TRUE
BTS	Ratchayothin	green	13.83	100.57	TRUE
BTS	Phahonyothin 24	green	13.82	100.57	TRUE
BTS	Ha Yaek Lat Phra	green	13.82	100.56	TRUE

Figure 1: Data frame storing detail about rail public transportation in Bangkok

The data collection from Google Maps in this part is done by the author is the data is not available only from any sources. The station data from the csv file is firstly read in and then stored in transportation_df, which will be used throughout this project.

Latitude and longitude data of each station would be very helpful when there is a need to locate the station on the map accurately. isOperating column would be beneficial for further update when there is some station not operating temporarily.

The location and other information, such as rating, about restaurant and food services can be obtained from Foursquare through Foursquare API.

Methodology

In the very first step, the locations of all rail service transportation stations are visualized on the map of Bangkok generated by the folium library to show a broad picture of how each station is located around and spread across Bangkok.
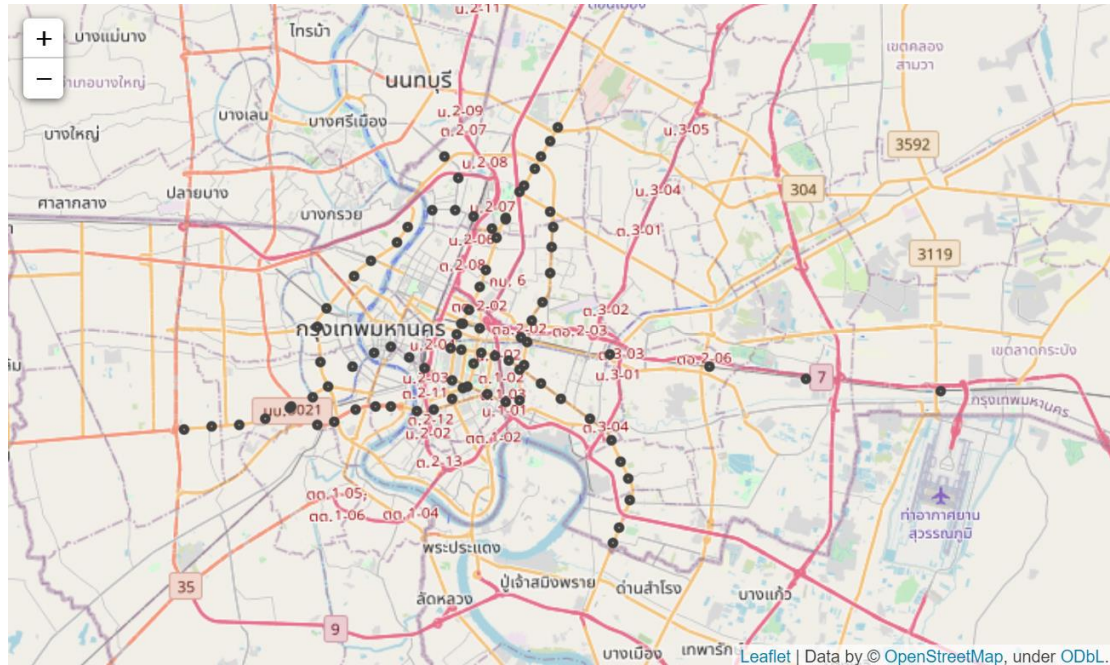

Figure 2: Map of Bangkok with all rail public transportation stations plotted in dark grey

Since there are many services and rail lines of rail public transportation in Bangkok, it is better to plot these points in colors that are according to the rail lines or services of the stations. For some stations, they are the transportation hub connecting more than one rail line or service together. These stations will be plotted specially in grey to make some contrast from other normal stations.

For the next step, the color of plotting changes to represent the rail line or services that the rail public transportation station is located on.
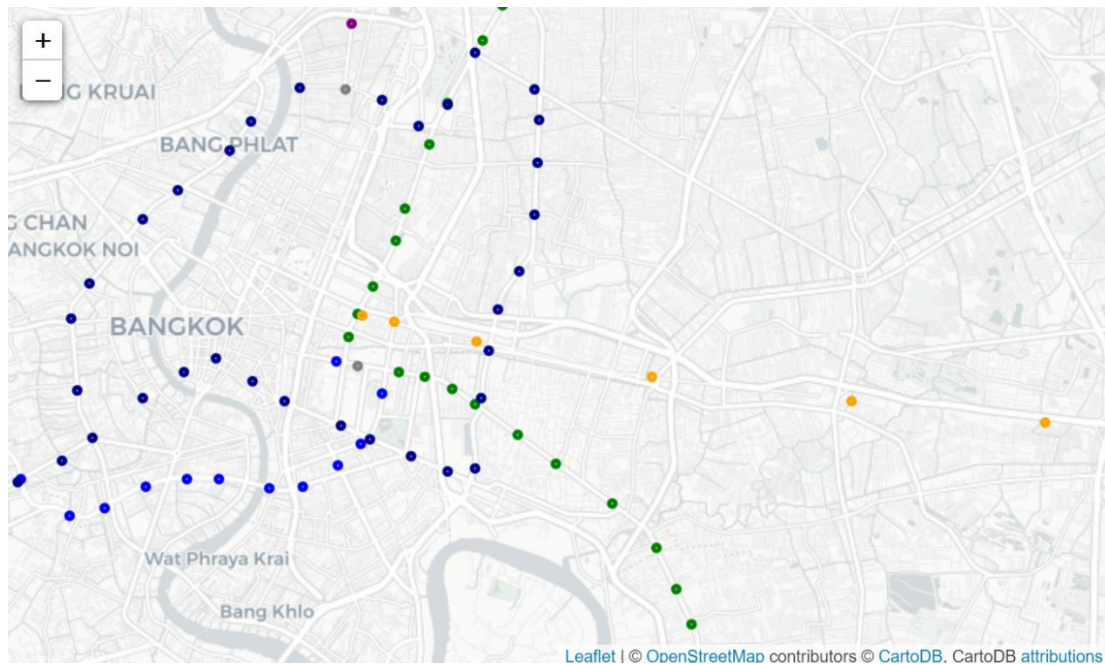

Figure 3: Map of Bangkok with all rail public transportation stations plotted in the colors according the rail line or service color that the station is located on.

For the next part, the project would ask the group of people who would like to gather together to enter their location information (latitude and longitude) so that the project can suggest the location of the restaurant and the restaurant accordingly.

Throughout the report and the notebook, the example used will be the example of two people wishing to gather together at some food and restaurant services. They have the following location (latitude and longitude) as follows:

1. Person No. 1: Latitude is 13.751725, Longitude is 100.531142
2. Person No. 2: Latitude is 13.740558, Longitude is 100.525142

Unit: Degree

The location and number of people can be changed to anywhere subjected to the conditions, needed to be satisfied, that these all people must be in Bangkok region and the number of people must be at least two.

For the next step, in order to visualize the position of all people who would like to gather together. The location of every person is plotted in red on the same map to show their location together with public transportation stations.
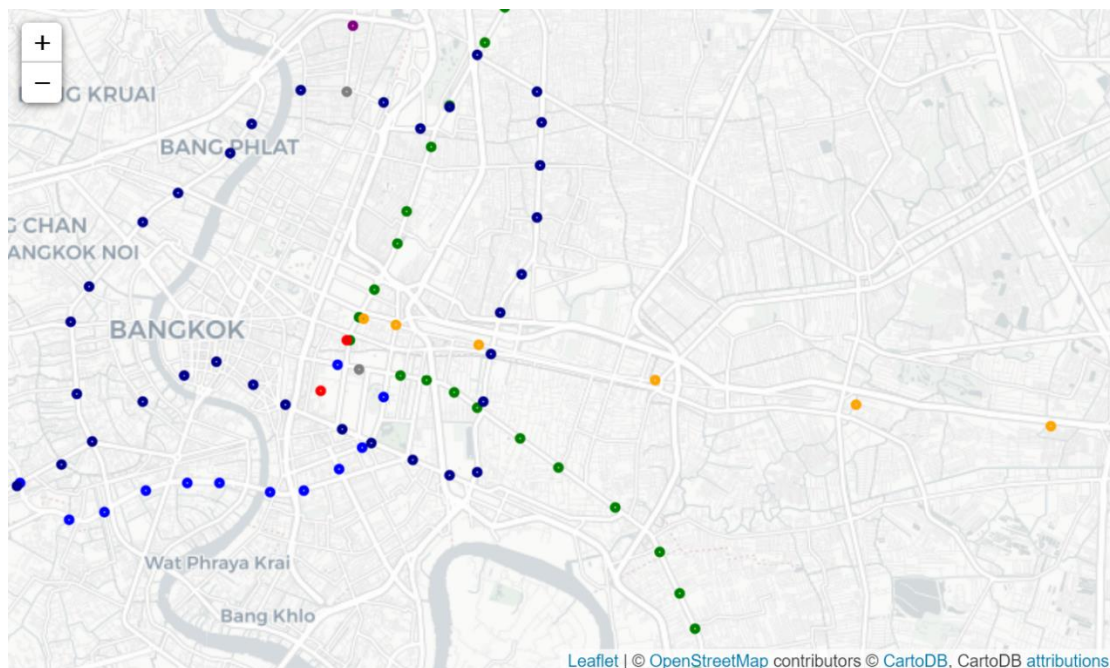

Figure 4: Map of Bangkok with all rail public transportation stations and location of gathering people plotted in the colors according the rail line or service color that the station is located on and in red, respectively.

The best point to meet up should be, for simplicity, the point that is equally far from everyone in term of distance. Therefore, the meet up point, which will be called midpoint throughout the report, will be the point that has its latitude equals average latitude of all people involved and its longitude equals average longitude of all people involved.

The project would calculate the midpoint of people's location data and then plot on the map to visualize the best meet up point (in term of distance) for all people involved on the map in pink
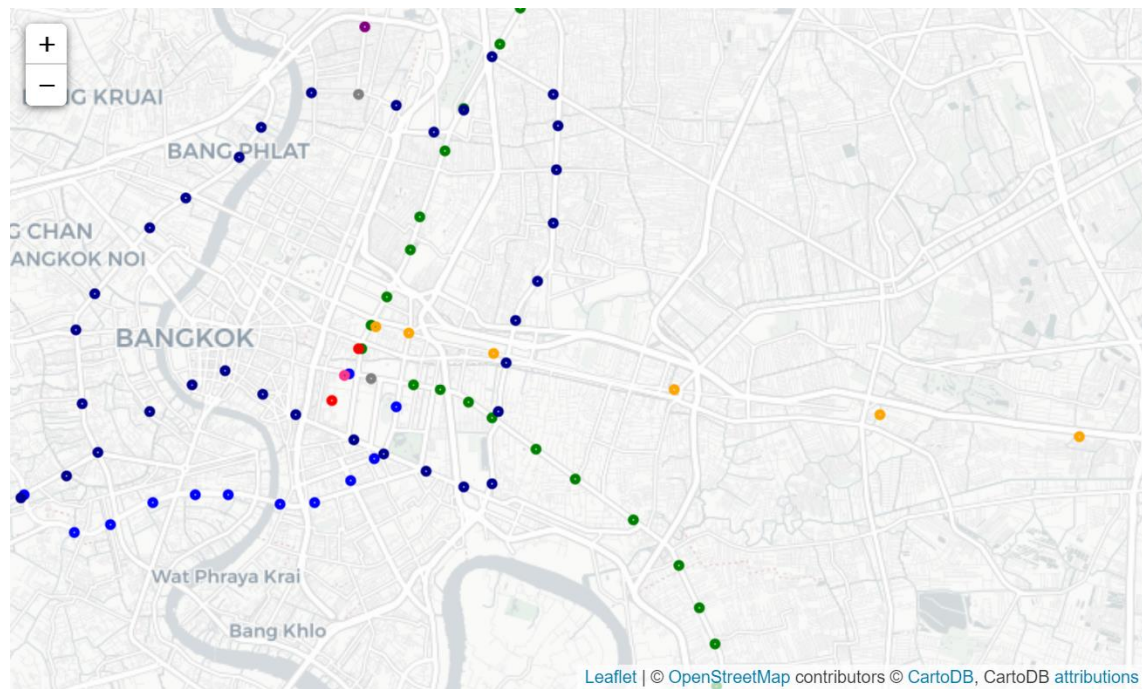

Figure 5: Map of Bangkok with all rail public transportation stations, location of gathering people, and location of the midpoint plotted in the colors according the rail line or service color that the station is located on, in red, and in pink, respectively.

However, there is a real-world constraint in term of transportation. Since many people nowadays tend to utilize more of rail public transportation service, it may not be really feasible for gathering people to meet others exactly at the midpoint as it might not be convenient for public transportation user.

Instead, the project would suggest the restaurants that are close to the public transportation station that is nearest to the midpoint, not the midpoint itself.

The project then calculates distance from the midpoint to all rail public transportation station and then store this information in the new 'Distance' column of transportation data frame.

	Service	Station	Lines	Latitude	Longitude	isOperating	Distance
0	BTS	Siam	grey	13.745511	100.533880	True	0.005773
1	BTS	National Stadium	blue	13.746345	100.528863	True	0.000749
2	BTS	Ratchadamri	blue	13.739500	100.539267	True	0.012957
3	BTS	Sala Daeng	blue	13.728426	100.534436	True	0.018800
4	BTS	Chong Nonsi	blue	13.723851	100.529450	True	0.022329

Figure 6: Updated transportation data frame with information about distance from each station to the midpoint in 'Distance' column of the data frame.

The project then selects the station that is closest to the midpoint, which is the station that has the minimum value in the 'Distance' column. The program firstly selects the index of the row and then call the name of the station of that row.

```
indexMin = findMin(distances)[0]
print('The closest station is \'{}\''.format(transportation_df['Station'][indexMin]))
The closest station is 'National Stadium'
```

In this example, the project tells that closest station to the midpoint is 'National Stadium'. Hence, people gathering together should consider the restaurants near 'National Stadium' station.

Generally, approximately 10 minutes duration of walking from a rail transportation station is reasonable, which is not too far from people to walk from the station. This is equivalent to 800 meters of walking distances from the station. Therefore, the project should suggest the restaurants that are less than 800 meters away from the selected public transportation station.

The project firstly creates query to get the information of all restaurants within 800 meters from the station and use that query to get the data from Foursquare API.

Please note that CLIENT_ID and CLIENT_SECRET are intentionally hidden for a privacy purpose.

```
CLIENT_ID = 'Unavailable'
CLIENT_SECRET = 'Unavailable'
VERSION = '20180604'
LIMIT = 20
search_query = 'Restaurant'
radius = 800
url = 'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius=
{}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, lat_centre, long_centre, VERSION, search_query, radius, LIMIT)
```

```
results = requests.get(url).json()
venues = results['response']['venues']
restaurant_df = json_normalize(venues)

restaurant_df.head()
```

The following data frame is obtained as a result of the Foursquare API call.

	id	name	categories	referralId	hasPerk	location.address	location.crossStreet	locati
0	4bdd057bb0f5c92817924be3	PH1 -Party House One Bar & Restaurant	[{'id': '4bf58dd8d48988d1d5941735', 'name': 'H...	v-1590343182	False	Siam@Siam Design Hotel & Spa	G Fl. & M Fl.	13.
1	4f3b48bde4b06a3b8f065498	City Center Restaurant (ห้องอาหารกลางเมือง)	[{'id': '4bf58dd8d48988d1c4941735', 'name': 'R...	v-1590343182	False	กฤษณ์ไทย แมนชั่นส์	NaN	13.
2	4d1f3b88dd6a236a4b132e38	Yana Restaurant	[{'id': '52e81612bcbc57f1066b79ff', 'name': 'H...	v-1590343182	False	MBK Center	5th Fl.	13.
3	509fadd0e4b00adc5ee9170f	Grandmother Restaurant (ร้านคุณยาย)	[{'id': '4bf58dd8d48988d149941735', 'name': 'T...	v-1590343182	False	Phaya Thai Rd	NaN	13.
4	4f0514cb9a523e111eeba1d6	Jim Thompson Bar & Restaurant	[{'id': '4bf58dd8d48988d149941735', 'name': 'T...	v-1590343182	False	6 Soi Kasem San 2	NaN	13.

Figure 7: Data frame storing information of all restaurants within 800 meters from the station.

The previous data frame is then cleaned to obtain the data frame below.

	id	name	location.lat	location.lng	location.distance
0	4bdd057bb0f5c92817924be3	PH1 -Party House One Bar & Restaurant	13.747002	100.526976	216
1	4f3b48bde4b06a3b8f065498	City Center Restaurant (ห้องอาหารกลางเมือง)	13.746248	100.528488	41
2	4d1f3b88dd6a236a4b132e38	Yana Restaurant	13.745851	100.530206	155
3	509fadd0e4b00adc5ee9170f	Grandmother Restaurant (ร้านคุณยาย)	13.750530	100.531317	536
4	4f0514cb9a523e111eeba1d6	Jim Thompson Bar & Restaurant	13.749316	100.528363	335

Figure 8: Cleaned data frame storing information of all restaurants within 800 meters from the station.

In order to return the best recommendation to users, all restaurants are to be ranked by their rating. The restaurant with highest rating should be recommended first. The restaurant with no rating, in this project, are ignored from the recommendation since there is no other ways using entirely Foursquare API to obtain the important information.

The query is created and then used to get rating information of all restaurants. If the restaurant is rated, the name, the location, and the rating are stored. If not, the restaurant is ignored.

```python
ratedList= []

for id in list(restaurant_df['id']):
    url = 'https://api.foursquare.com/v2/venues/{}?client_id={}&client_secret={}&v={}'.format(id, CLIENT_ID,
CLIENT_SECRET, VERSION)
    result = requests.get(url).json()
    try:
        ratedList.append([result['response']['venue']['rating'], result['response']['venue']['name'], result
['response']['venue']['location']['lat'], result['response']['venue']['location']['lng'] ])
    except:
        pass
```

Figure 8: Query for getting rating information of restaurants that are less than 800 meters away from the station through Foursquare API

The ratedList, which store the rating, the location, and the name of the restaurants, are obtained. Then the ratedList are sorted so that the project can print and suggest the top rated restaurant easily.

```
ratedList

[[7.1,
  'PH1 -Party House One Bar & Restaurant',
  13.747002390429184,
  100.52697572012217],
 [7.5, 'Yana Restaurant', 13.745851059524055, 100.53020551662877],
 [7.8,
  'Jim Thompson Bar & Restaurant',
  13.749316116244351,
  100.52836349409179],
 [7.2,
  'Scala Restaurant (ภัตตาคารสกาล่า)',
  13.745749901265727,
  100.53126870602846],
 [7.3, 'The Great Wall Restaurant', 13.751497465166223, 100.53034327118345],
 [5.2, 'The Eight Restaurant', 13.746647620627424, 100.5289768821219]]
```

The program then prints and suggests the restaurant with top three highest rating from the all restaurant data.

```python
print('Suggested Restaurant')
try:
    print("1.", sortedRatingList[-1][1])
except:
    pass
try:
    print("2.", sortedRatingList[-2][1])
except:
    pass
try:
    print("3.", sortedRatingList[-3][1])
except:
    pass

Suggested Restaurant
1. Jim Thompson Bar & Restaurant
2. Yana Restaurant
3. The Great Wall Restaurant
```

Finally, the three suggested restaurants are plotted in yellow on the same map to provide users with visualization of the location of the suggested restaurant.
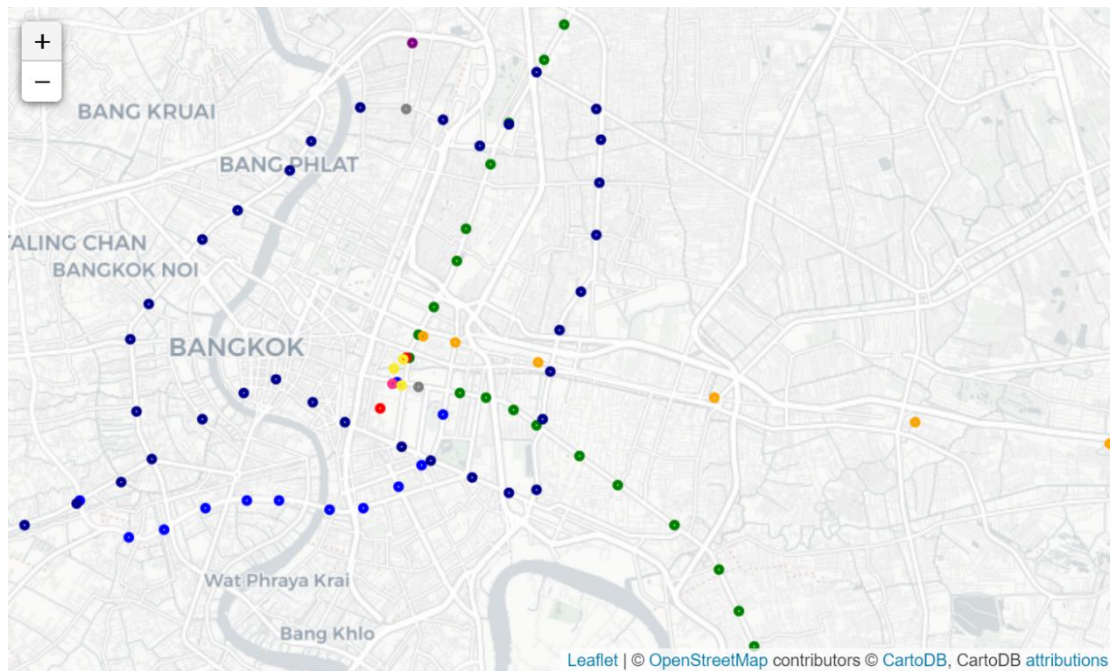

Figure 9: Map of Bangkok with all rail public transportation stations, location of gathering people, location of the midpoint plotted, and location of suggested restaurant in the colors according the rail line or service color that the station is located on, in red, in pink, and in yellow, respectively.

Result

With the input information of the rail transportation stations in Bangkok, Thailand together with the location data of gathering people and data from Foursquare API, the project is able to suggest the restaurants with top three highest rating near the rail transportation station (within 800 meters) that is closest to the midpoint of everyone who gathers.

The project prints three suggested restaurants.

```
Suggested Restaurant
1. Jim Thompson Bar & Restaurant
2. Yana Restaurant
3. The Great Wall Restaurant
```

The project also plots the top three restaurant points in yellow and label them the name and the rating when the curser is hover on the plot.
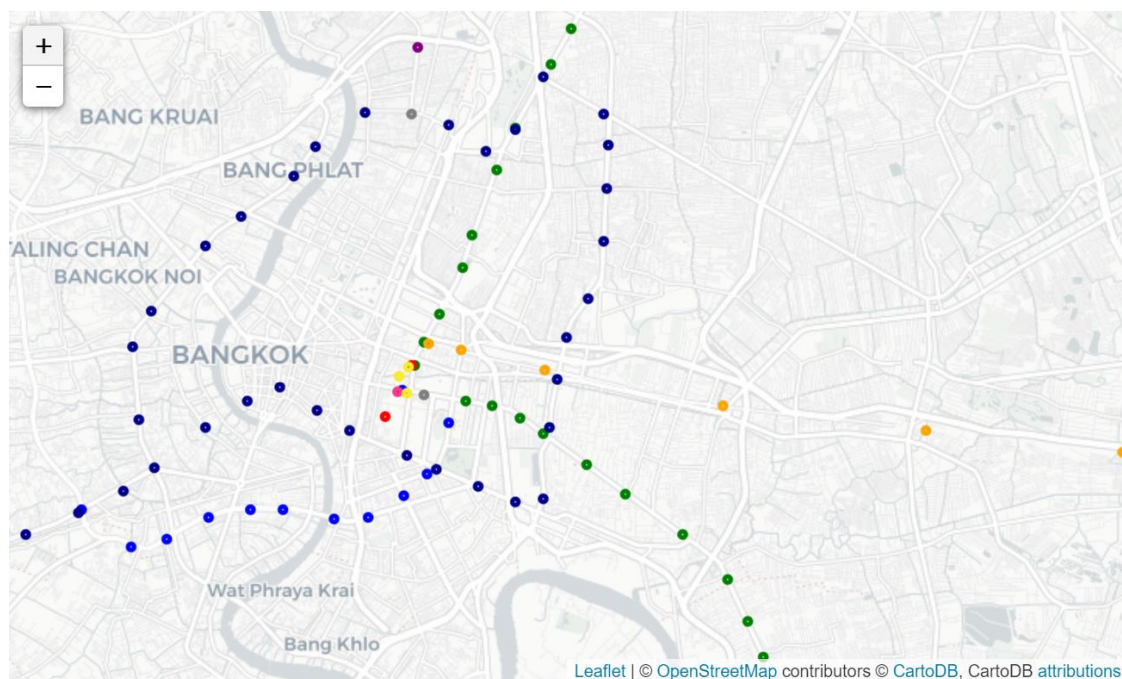
Discussion

The project can be used to suggest restaurant and food services with top three highest rating at the location that is approximately equally far from everyone, based on public transportation station. However, as of the current version, the restaurants must be rated first in order to be suggested to the users since there is no way to rank the unrated restaurant as of now. It may cause the problem in case that all restaurants in some regions are unrated. The current scale of the project is now limited to Bangkok region only, where there are networks of public transportation services. The future development may include the way to rank restaurants in absence of ratings from Foursquare API. The extra information used may be number of positive ratings and comments.

Conclusion

The Fine Great Food @ BKK: Bangkok Food & Restaurant Data-assisted Suggestion is a project aiming to address the current big economic problem caused by COVID-19 pandemic. To address the problem, this project is designed to suggest the restaurant at the appropriate location. By suggesting the restaurants, people will get more information and restaurants are promoted, increasing their revenue. With the location suggestion, people can have convenience in finding restaurant location that is equally far from everyone and can get there by public transportation. The projects take in the location data of gathering people, location of all rail public transportation as inputs. With data science techniques and the help of Foursquare API, the project is able to suggest the top three restaurants ranked by rating which is within 10 minutes walking distance from the rail public transportation that is located approximately equally far from everyone.

Thank you for your time reviewing my work. I hope you have a nice day, stay healthy, and keep safe.