Relocation Analysis:

From Birmingham to Bangkok

Relocation is not easy

- Moving to a foreign country to live in a new neighborhood with different culture, lifestyles, environment, etc., while exciting, can be challenging and stressful for everyone.
- I am currently in such process to relocate from Birmingham, UK to Bangkok, Thailand and would value any insights that could help me select a neighborhood in Bangkok that share some similarity to my current neighborhood – Edgbaston – and in turn lessen the challenges of living in a foreign land.
- In addition to looking for similar neighborhood, I also want to maintain my current lifestyle and want the following venues within a specific radius:
  - Gym: within 2000 meters
  - Coffee Shop: within 1000 meters
  - Grocery Store/Supermarket: within 2000 meters
- While the problem may seem to focus on my life event, the analysis is adaptable to suit anyone's needs and should be useful for those who are moving either domestically or internationally as well.

Data Acquisition and Cleaning

- Bangkok's districts and coordinates were scraped from Wikipedia via the following URL: https://en.wikipedia.org/wiki/List_of_districts_of_Bangkok
  Only the following features from the data were kept:
  - District
  - Post Code
  - Latitude
  - Longitude
- Data for my current neighborhood – Edgbaston – is provided manually but the coordinate used for the analysis was obtained using geopy package.
- The cleaned data for Bangkok contains 4 features and 50 rows, while data for my neighborhood has 4 features and one row.

Data Acquisition – Foursquare API

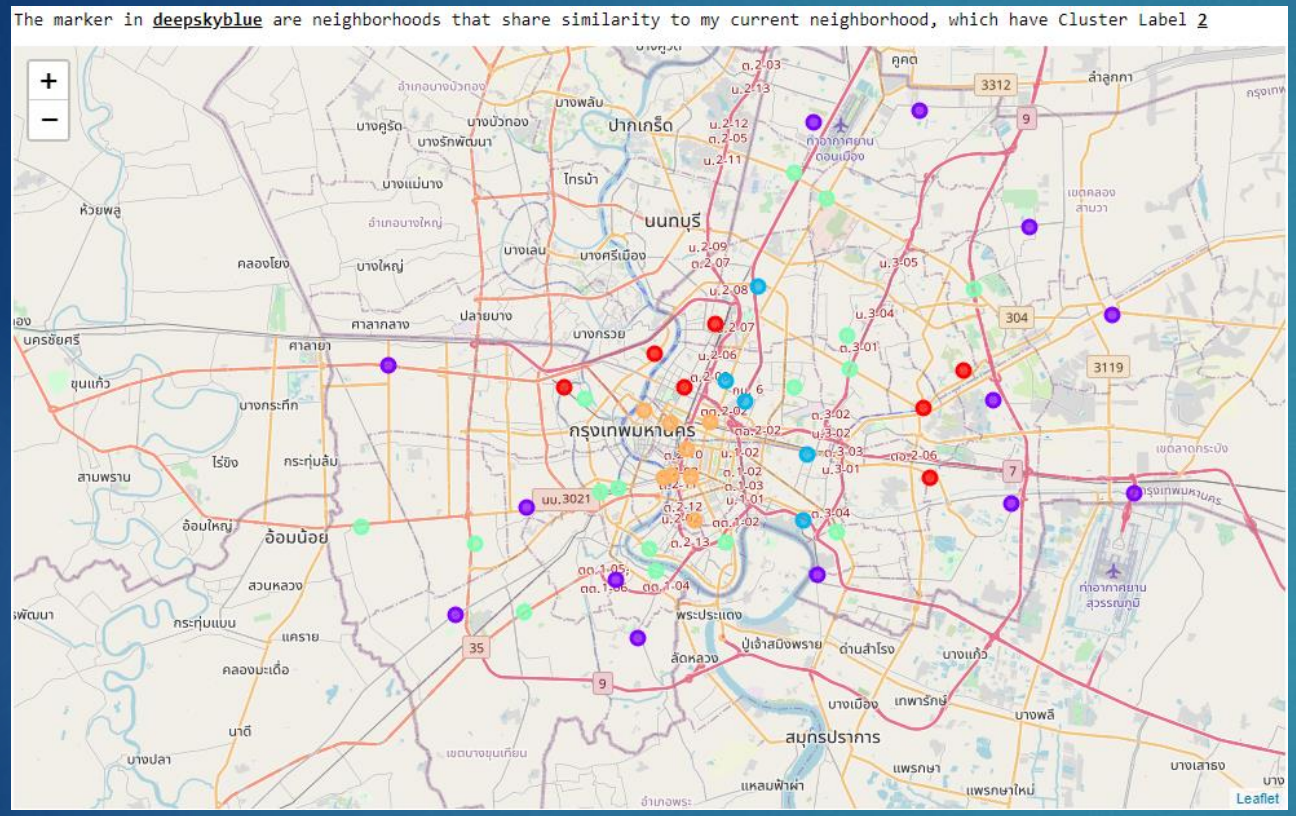
	District	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Bang Bon	Convenience Store	Coffee Shop	Noodle House	Fast Food Restaurant	Asian Restaurant	Thai Restaurant	Flea Market	Badminton Court	Hotpot Restaurant	Market
1	Bang Kapi	Noodle House	Coffee Shop	Thai Restaurant	Japanese Restaurant	Dessert Shop	Som Tum Restaurant	Fast Food Restaurant	Steakhouse	Clothing Store	Soccer Field
2	Bang Khae	Noodle House	Coffee Shop	Fast Food Restaurant	Convenience Store	Thai Restaurant	Dessert Shop	Asian Restaurant	Café	Shopping Mall	Department Store
3	Bang Khen	Coffee Shop	Noodle House	Convenience Store	Hotpot Restaurant	Asian Restaurant	Thai Restaurant	Som Tum Restaurant	Golf Course	Bookstore	Vietnamese Restaurant
4	Bang Kho Laem	Noodle House	Thai Restaurant	Coffee Shop	Hotel	Asian Restaurant	Pub	Chinese Restaurant	BBQ Joint	Bar	Massage Studio

- Coordinates from Bangkok and my neighborhood were used to get nearby venues or point of interests from Foursquare API.
- A limit is set for 100 venues, and a radius of 3,000meters from the coordinates. A wide radius was used due to my neighborhood located in a large residential area with few venues nearby, and smaller radius also seems to return very few venues for many districts in Bangkok as well.
- The obtained data was cleaned and transformed to display 10 most common venues for each district.

Data Segregation

- K-Means clustering was used as a segregation method to label district that shares similarity.
- Number of clusters was set to 5. This returns a reasonable number of districts per cluster. Larger or smaller number of clusters produced results that have few clusters with most districts and most clusters with few districts or none.

Data Visualization



- A mapping package Folium was used to visualize the clustered data.
- A map of Bangkok was overlaid with colored markers representing position of the district and color for each cluster label.
- A printout also indicate color and label of district similar to my neighborhood.

Applying Venues Requirements

	District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Distance(km)
0	Bang Rak	13.730833	100.524167	Rod Dee Det (รสดีเด็ด)	13.734807	100.524340	Noodle House	0.440069
1	Bang Rak	13.730833	100.524167	Amara Bangkok	13.728673	100.527564	Hotel	0.438295
2	Bang Rak	13.730833	100.524167	iSanook Residence	13.732560	100.521723	Hotel	0.326184
3	Bang Rak	13.730833	100.524167	สุกี้สามย่าน	13.734000	100.527565	Thai Restaurant	0.507782
4	Bang Rak	13.730833	100.524167	Jok Sam Yan (โจ๊กสามย่าน)	13.734714	100.526080	Breakfast Spot	0.476646

- Recall the requirements of venues within a distance I set at the beginning:
  - Gym: within 2000 meters
  - Coffee Shop: within 1000 meters
  - Grocery Store/Supermarket: within 2000 meters
- Data with 10 most common venues were filtered to display only data from similar district to my neighborhood.
- Distance between venues and the district was calculated using geopy.distance package and appended to the data as column ‘Distance’.

Applying Venues Requirements

District	Venue Category	Venue
Chatuchak	Coffee Shop	4
	Gym / Fitness Center	3
	Supermarket	2
Din Daeng	Supermarket	2
Khlong Toei	Coffee Shop	1
	Convenience Store	2
	Gym / Fitness Center	3
Phaya Thai	Coffee Shop	7
	Gym / Fitness Center	1
	Supermarket	1
Watthana	Coffee Shop	2
	Supermarket	2

- Using .loc to filter the previous table with the venues and distance requirements and summarize the result table by grouping the district by count of venues.
- From the table, the following districts have at least a gym within 2km, supermarket within 2km, and coffee shop within 1km:
  - Chatuchak
  - Khlong Toei
  - Phaya Thai
- The analysis has narrowed down 50 districts in Bangkok to three that are similar to my neighborhood and has all the required venues within a specific distance. Thus saving me a significant amount of time to do a research and allowing me to focus for house hunting in smaller number of areas!

Conclusion

- Built an analysis model to search for neighborhoods in Bangkok that share similarity to my current neighborhood (Edgbaston, Birmingham, UK) based on venues within a set radius. A set of required venues within a specific distance was also created to fulfill my required lifestyle.
- The model resulted in three districts that matched my neighborhood out of 50.
- The model is crude as it only looks to match neighborhood based on venues in the nearby area only. One would need more insights in the neighborhood than just types of venues. Additional data such as house prices, crime rate, proximity to public transport, etc. could be added as part of the analysis to search for a suitable neighborhood.