# Report

## Introduction

With a population of more than 10 million, the city of Bangalore in India is one of the most populous cities in India with a significant migrant population. With the blossoming number of apartments across the city, the choice of apartment to rent is determined not just by the proximity to the workplace but the quality of the neighbourhood. This project will cluster together similar neighborhoods based on the popular venue categories in each neighbourhood and map the various apartments in these neighborhoods.

This will help people who are looking to move into the city or to see if there are similar neighborhoods to their current neighborhoods which they can consider relocating.

## Data

The data used for this project will be obtained from the following sources:

a) For the list of neighborhood in Bangalore, although the dataset available in Kaggle (https://www.kaggle.com/rmenon1998/bangalore-neighborhoods), was initially considered, later, the wikipedia page on neighborhoods in Bangalore which is available here: https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Bangalore was used.

b) For latitude and longitude details for the above neighborhoods, Google API has been used.

c) Finally, Foursquare API will be used to collate neigborhood venue information.

## Methodology

Exploratory Data Analysis: Kaggle has a dataset of Bangalore neighborhoods along with their latitude and longitude details available. Initially, the idea was to go with this available data set. An extract of the data set downloaded from Kaggle is as follows:

| | Unnamed: 0 | Neighborhood | Latitude | Longitude |
|---|---|---|---|---|
| 0 | 0 | Agram | 45.813177 | 15.977048 |
| 1 | 1 | Amruthahalli | 13.066513 | 77.596624 |
| 2 | 2 | Attur | 11.663711 | 78.533551 |
| 3 | 3 | Banaswadi | 13.014162 | 77.651854 |
| 4 | 4 | Bellandur | 58.235358 | 26.683116 |
| 5 | 5 | Bhattarahalli | 13.025800 | 77.714279 |
| 6 | 6 | Bidrahalli | 14.577426 | 74.928560 |
| 7 | 7 | Byatarayanapura | 13.062074 | 77.596392 |
| 8 | 8 | Devanagundi | 12.973613 | 77.839402 |
| 9 | 9 | Devasandra | 12.757227 | 77.647280 |
| 10 | 10 | Doddagubbi | 14.452076 | 75.510799 |
| 11 | 11 | Doddanekkundi | 12.975720 | 77.694042 |

Sample of the data set available on Kaggle

There were two main issues with the above data:

- one- The neighborhoods listed here are not commonly understood as neighborhoods by residents of Bangalore. On further analysis, it was realised that this data set equates post offices as neighborhoods. While this list may be useful in some use cases, for the purposes of this Project, neighborhood is to be given a more layman interpretation- i.e., we want a list which identifies neighborhoods as commonly understood by the residents of Bangalore.
- two- the latitude and longitude data in this is very evidently incorrect. For instance, for neigborhood 'Agram', latitude is shown as 45.813 and for 'Amruthahalli', latitude is shown as 13.06. No two neighborhoods within the same city can have such vide variation in latitudes.

Accordingly, it was decided not to use this dataset, but to create a new data set of Bangalore neigbhbourhoods by scraping from the net. Wikipedia has a page on Bangalore neighborhoods which lists out the correct neighborhoods, as below.

Central [ edit ]

| Name | Image | Summary |
|---|---|---|
| Cantonment area | | The Cantonment area in Bangalore was used as a military cantonment during the British Raj in the 19th century. After the Indian independence, the area merged with rest of the city but still retains names for localities and streets used by the British as well as the colonial architecture in many buildings. Some localities which come under this area are Richmond Town, Austin Town, Murphy Town, Fraser Town, and Cox Town. |
| Domlur |  | Formerly part of the Cantonment area, Domlur has become a major hub for IT outsourcing firms and military establishments. The Embassy GolfLinks Business Park and HAL Airport are located in/near Domlur. |
| Indiranagar |  | Indiranagar is a sought-after residential and commercial neighbourhood which is located between the central and eastern parts of the city. The neighbourhood is known for its cosmopolitanism, nightlife, pubs and retail stores. The 100 Feet Road which runs through Indiranagar is one of the major commercial hubs in the city. |
| Jeevanbheemanagar | | |
| Malleswaram |  | |

snapshot of the wiki page on Bangalore neighborhoods

However, as evident above, this list does not have the latitude and longitude data. Further, the names of neighborhoods are provided in different tables in the wiki page along with images and summary which are not useful for our context.

Accordingly, I scraped only the neighborhood names from the Wikipedia page and created a new dataframe:

| | Name |
|---|---|
| 0 | Cantonment area |
| 1 | Domlur |
| 2 | Indiranagar |
| 3 | Jeevanbheemanagar |
| 4 | Malleswaram |

But as you can see, this dataframe does not have the latitude and longitude details of these neighborhoods. We would need the locational data to easily plot these neighborhoods on map.

Using Google Maps' API, latitude and longitude details were pulled for each of these neighborhoods and the latitude and longitude columns were added to the dataframe:

| | Name | long | lat |
|---|---|---|---|
| 0 | Cantonment area | -87.3446 | 30.6054 |
| 1 | Domlur | 77.6387 | 12.961 |
| 2 | Indiranagar | 77.6408 | 12.9784 |
| 3 | Jeevanbheemanagar | 77.6581 | 12.9642 |
| 4 | Malleswaram | 77.5692 | 13.0055 |

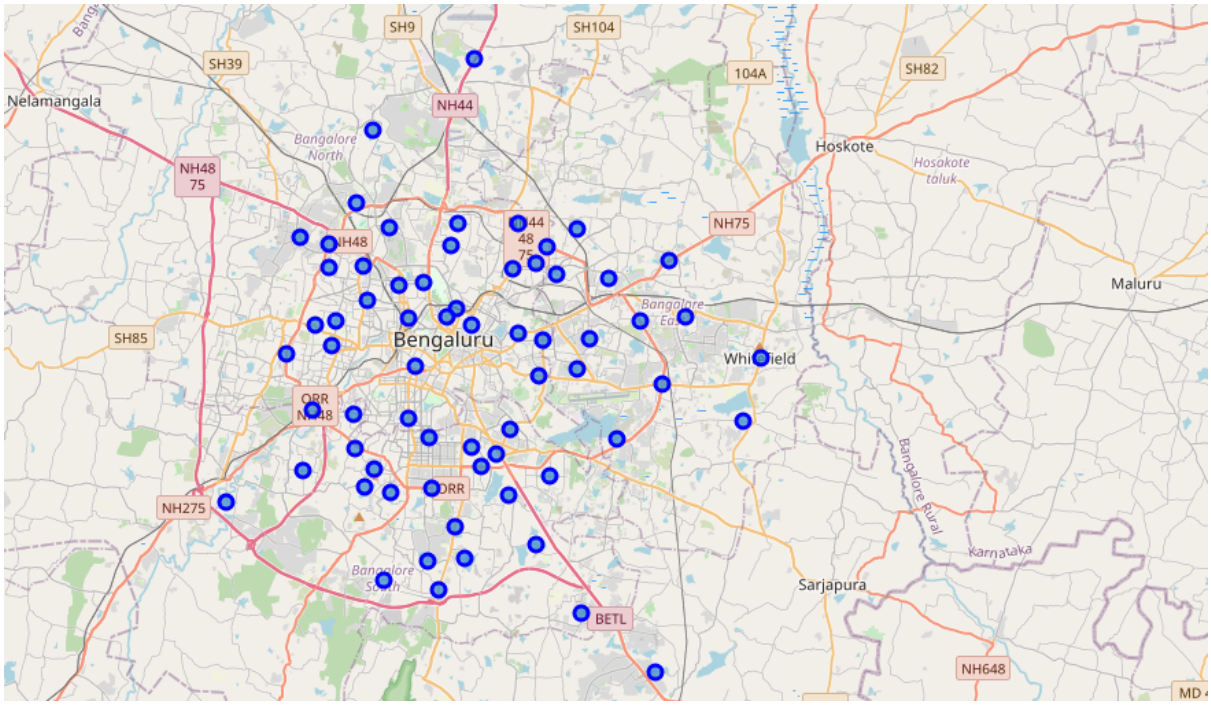
Dataframe with latitude and longitude details added

On an analysis, it turned out that even in this, a couple of longitude and latitude details were wrong. For these, further address details were added and the correct latitude and longitude details were included. For future references purposes, this dataframe was also saved as a csv file.

| | Name | long | lat |
|---|---|---|---|
| 0 | Cantonment area | 77.597800 | 12.993800 |
| 1 | Domlur | 77.638732 | 12.960986 |
| 2 | Indiranagar | 77.640836 | 12.978369 |
| 3 | Jeevanbheemanagar | 77.658076 | 12.964163 |
| 4 | Malleswaram | 77.569236 | 13.005511 |

Dataframe after correcting the latitude and longitude details

Using Folium, these neighborhoods were plotted on the map of Bangalore:




Bangalore Neighborhoods

The next step was to use Foursquare API to pull the 100 top venues within 500 metre radius of each of these neighborhoods.

```
(880, 7)
There are 147 unique categories.
```

| | Neighborhood | Neighborhood Latitude | Neighborhood Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category |
|---|---|---|---|---|---|---|---|
| 0 | Cantonment area | 12.9938 | 77.5978 | Ujwal Bar & Restaurant | 12.9923 | 77.5945 | Indian Restaurant |
| 1 | Cantonment area | 12.9938 | 77.5978 | Alliance Française | 12.9912 | 77.5967 | Concert Hall |
| 2 | Cantonment area | 12.9938 | 77.5978 | Millers 46 | 12.9917 | 77.5942 | Steakhouse |
| 3 | Cantonment area | 12.9938 | 77.5978 | Jayamahal Palace Hotel | 12.9968 | 77.5972 | Indian Restaurant |
| 4 | Cantonment area | 12.9938 | 77.5978 | Watson's | 12.9942 | 77.5946 | Pub |
| 5 | Cantonment area | 12.9938 | 77.5978 | Jaymahal Palace Grounds | 12.9953 | 77.5969 | Music Venue |

Top venues within 500m radius of each of the neighborhoods

Next step was to generate one-hot vectors for the venues in each of these neighborhoods and generate the most common venues and get the top 10 venues.

| | Neighborhood | 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 | Anjanapura | ATM | Women's Store | Electronics Store | Food | Flower Shop | Flea Market | Fast Food Restaurant | Farmers Market | Falafel Restaurant | Event Space |
| 1 | Arekere | Sporting Goods Shop | Indian Restaurant | Pizza Place | Athletics & Sports | Breakfast Spot | Liquor Store | Department Store | Business Service | Fast Food Restaurant | Bar |
| 2 | BTM Layout | Indian Restaurant | Ice Cream Shop | Bakery | Chinese Restaurant | Snack Place | Pizza Place | Fast Food Restaurant | Sandwich Place | Coffee Shop | Vegetarian / Vegan Restaurant |
| 3 | Banashankari | Clothing Store | Café | Breakfast Spot | North Indian Restaurant | Fast Food Restaurant | Indian Restaurant | Fried Chicken Joint | Pizza Place | Men's Store | Sporting Goods Shop |
| 4 | Banaswadi | Indian Restaurant | Vegetarian / Vegan Restaurant | Pharmacy | Kerala Restaurant | Women's Store | Electronics Store | Flea Market | Fast Food Restaurant | Farmers Market | Falafel Restaurant |
| 5 | Basavanagudi | Indian Restaurant | Fast Food Restaurant | Park | Sandwich Place | Food | Hookah Bar | Pharmacy | Café | Plaza | Metro Station |
| 6 | Basaveshwaranagar | Ice Cream Shop | Fast Food Restaurant | Indian Restaurant | Juice Bar | Burger Joint | Miscellaneous Shop | Gastropub | Sporting Goods Shop | Pizza Place | Gym |
| 7 | Begur | Supermarket | Mobile Phone Shop | Women's Store | Flower Shop | Flea Market | Fast Food Restaurant | Farmers Market | Falafel Restaurant | Event Space | Electronics Store |
| 8 | Bellandur | Hotel Bar | Soccer Field | Fast Food Restaurant | Shopping Mall | Indian Restaurant | Hookah Bar | Eastern European Restaurant | Hotel | Farmers Market | Falafel Restaurant |
| 9 | Bidadi | Indian Restaurant | Breakfast Spot | Women's Store | Electronics Store | Flower Shop | Flea Market | Fast Food Restaurant | Farmers Market | Falafel Restaurant | Event Space |
| 10 | Bommanahalli | Indian Restaurant | Department Store | Women's Store | Electronics Store | Flower Shop | Flea Market | Fast Food Restaurant | Farmers Market | Falafel Restaurant | Event Space |

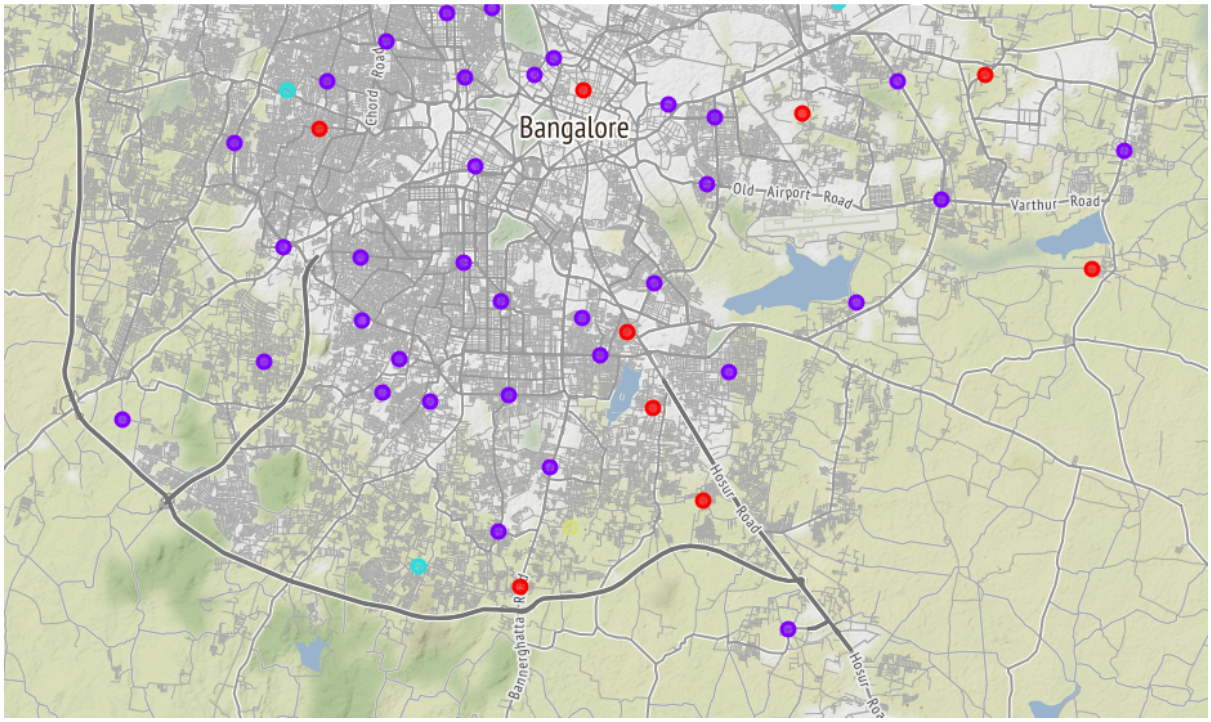
Top ten venues in each neighborhood

Next step is to segment these neighborhoods based on similarities in their common venues. For this, I have used K-means clustering. I have set the clusters as four.

| | Neighborhood | long | lat | Cluster Labels | 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Cantonment area | 77.597800 | 12.993800 | 1 | Indian Restaurant | Bakery | Music Venue | Gym / Fitness Center | Concert Hall | Tea Room | Pub |
| 1 | Domlur | 77.638732 | 12.960986 | 1 | Indian Restaurant | Café | BBQ Joint | Rajasthani Restaurant | Chinese Restaurant | Asian Restaurant | Sandwich Place |
| 2 | Indiranagar | 77.640836 | 12.978369 | 1 | Indian Restaurant | Café | Pub | Bakery | Vegetarian / Vegan Restaurant | Fast Food Restaurant | Chinese Restaurant |
| 3 | Malleswaram | 77.569236 | 13.005511 | 1 | Ice Cream Shop | Coffee Shop | Chinese Restaurant | Fast Food Restaurant | South Indian Restaurant | Donut Shop | Café |
| 4 | Pete area | 77.577000 | 12.965600 | 1 | Indian Restaurant | Plaza | Market | Flower Shop | Historic Site | Food Truck | Diner |
| 5 | Sadashivanagar | 77.581285 | 13.006818 | 1 | Coffee Shop | Department Store | Ice Cream Shop | Café | Indian Restaurant | Seafood Restaurant | Cafeteria |
| 6 | Seshadripuram | 77.574044 | 12.988905 | 1 | Ice Cream Shop | Indian Restaurant | Clothing Store | Fast Food Restaurant | Bakery | Electronics Store | Donut Shop |
| 7 | Shivajinagar | 77.605693 | 12.985650 | 0 | Indian Restaurant | Clothing Store | Donut Shop | Fast Food Restaurant | Tea Room | Market | South Indian Restaurant |
| 8 | Ulsoor | 77.628415 | 12.981700 | 1 | Café | Burger Joint | Bakery | Event Space | Flower Shop | Flea Market | Fast Food Restaurant |
| 9 | Vasanth Nagar | 77.592795 | 12.989619 | 1 | Indian Restaurant | Coffee Shop | Chinese Restaurant | Italian Restaurant | Dessert Shop | Pizza Place | Sandwich Place |
| 10 | Bellandur | 77.678404 | 12.930428 | 1 | Hotel Bar | Soccer Field | Fast Food Restaurant | Shopping Mall | Indian Restaurant | Hookah Bar | Eastern European Restaurant |
| 11 | CV Raman Nagar | 77.664184 | 12.979325 | 0 | Indian Restaurant | Coffee Shop | Food Court | Café | Department Store | Dessert Shop | Dim Sum Restaurant |
| 12 | Hoodi | 77.712657 | 12.989555 | 0 | Indian Restaurant | Bus Station | Women's Store | Electronics Store | Flower Shop | Flea Market | Fast Food Restaurant |

Each neighborhoods with Cluster labels

Finally, these clusters are plotted on the map to get a spatial information about the relative locations of these similar neighborhoods.




Neighborhoods clustered

# Results

*First Cluster*

| | Neighborhood | long | lat | Cluster Labels | 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 Venu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7 | Shivajinagar | 77.605693 | 12.985650 | 0 | Indian Restaurant | Clothing Store | Donut Shop | Fast Food Restaurant | Tea Room | Market | South Indian Restaurant | Event Space | Flowe Sho |
| 11 | CV Raman Nagar | 77.664184 | 12.979325 | 0 | Indian Restaurant | Coffee Shop | Food Court | Café | Department Store | Dessert Shop | Dim Sum Restaurant | Diner | Donu Sho |
| 12 | Hoodi | 77.712657 | 12.989555 | 0 | Indian Restaurant | Bus Station | Women's Store | Electronics Store | Flower Shop | Flea Market | Fast Food Restaurant | Farmers Market | Falafe Restaurar |
| 13 | Krishnarajapuram | 77.704433 | 13.016999 | 0 | Indian Restaurant | Bakery | Tibetan Restaurant | Women's Store | Electronics Store | Flower Shop | Flea Market | Fast Food Restaurant | Farmer Marke |
| 18 | Banaswadi | 77.648194 | 13.010376 | 0 | Indian Restaurant | Vegetarian / Vegan Restaurant | Pharmacy | Kerala Restaurant | Women's Store | Electronics Store | Flea Market | Fast Food Restaurant | Farmer Marke |
| 27 | Mathikere | 77.563976 | 13.033419 | 0 | Indian Restaurant | Fast Food Restaurant | Ice Cream Shop | Department Store | Pizza Place | Bus Station | Convenience Store | Halal Restaurant | Histori Sit |
| 32 | Bommanahalli | 77.624194 | 12.902980 | 0 | Indian Restaurant | Department Store | Women's Store | Electronics Store | Flower Shop | Flea Market | Fast Food Restaurant | Farmers Market | Falafe Restaurar |
| 33 | Bommasandra | 77.697437 | 12.816730 | 0 | Indian Restaurant | Women's Store | Electronics Store | Flower Shop | Flea Market | Fast Food Restaurant | Farmers Market | Falafel Restaurant | Ever Spac |

Extract from the first cluster

*Second Cluster*

| | Neighborhood | long | lat | Cluster Labels | 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 | 8 C |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Cantonment area | 77.597800 | 12.993800 | 1 | Indian Restaurant | Bakery | Music Venue | Gym / Fitness Center | Concert Hall | Tea Room | Pub | F |
| 1 | Domlur | 77.638732 | 12.960986 | 1 | Indian Restaurant | Café | BBQ Joint | Rajasthani Restaurant | Chinese Restaurant | Asian Restaurant | Sandwich Place | Pizz |
| 2 | Indiranagar | 77.640836 | 12.978369 | 1 | Indian Restaurant | Café | Pub | Bakery | Vegetarian / Vegan Restaurant | Fast Food Restaurant | Chinese Restaurant | Dess |
| 3 | Malleswaram | 77.569236 | 13.005511 | 1 | Ice Cream Shop | Coffee Shop | Chinese Restaurant | Fast Food Restaurant | South Indian Restaurant | Donut Shop | Café | Came |
| 4 | Pete area | 77.577000 | 12.965600 | 1 | Indian | Plaza | Market | Flower | Historic | Food Truck | Diner | |

Extract from the Second cluster

*Third Cluster*

| Neighborhood | long | lat | Cluster Labels | 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ramamurthy Nagar | 77.673703 | 13.008489 | 2 | Shoe Store | ATM | Electronics Store | Food | Flower Shop | Flea Market | Fast Food Restaurant | Farmers Market | Falafel Restaurant | Event Space |
| Anjanapura | 77.561594 | 12.861652 | 2 | ATM | Women's Store | Electronics Store | Food | Flower Shop | Flea Market | Fast Food Restaurant | Farmers Market | Falafel Restaurant | Event Space |
| Kamakshipalya | 77.526883 | 12.985524 | 2 | ATM | Warehouse Store | Electronics Store | Food | Flower Shop | Flea Market | Fast Food Restaurant | Farmers Market | Falafel Restaurant | Event Space |
| Doddaballapura / li\nHoskote\n\nMagad... | 77.536391 | 13.295714 | 2 | ATM | Women's Store | Electronics Store | Food | Flower Shop | Flea Market | Fast Food Restaurant | Farmers Market | Falafel Restaurant | Event Space |

Extract from the Third Cluster

# Discussion

Looking at the clustered data, the most common venue across the first and second cluster is Indian restaurants. Therefore, while this may not be helpful in distinguishing between the first and the second neighbourhood, if you are looking to stay in a neigborhood with enough number of eating out places, then neighborhoods in the third cluster may not be your ideal choice.

Between the first cluster and the second cluster, if you are looking a place with more number of neighborhood cafes, pubs and fitness centres, any of the neighborhoods in the second cluster will be your ideal choice.

# Conclusion

This project looked at the qualitative factors in choosing a neighborhood. Rather than looking at price and proximity to work place, this project looks at the common venues in a neighborhood and using unsupervised k-cluster machine learning algorithm, segmented the similar neighborhoods into clusters.

If you are new to the city and if you prefer a neighborhood with a lot of cafes, pubs and fitness centres, neighborhoods in the second cluster will be your ideal choice.