

The Battle of Neighbourhoods-Final Report
(Comparing the similarity of neighbourhoods in the cities of
India)

Karan Thakur
IBM Data Science Professional Certificate

Contents

1. Introduction.....	3
1.1 Background	3
1.2 Problem.....	3
1.3 Interest	3
2. Data collection and cleaning.....	3
2.1 Data Sources	3
2.2 Data Cleaning and feature selection	4
3. Methodology.....	6
3.1 Neighbourhood Visualisation on the map.....	6
3.2 Data pre-processing to apply the machine learning technique	7
3.2 Employing K-Means clustering on the dataset	9
4. Results and Discussion	10
4.1 Visualising clustered neighbourhood on the map	10
4.2 Examining each cluster	11
5. Conclusion.....	15
6. References	15

1. Introduction

1.1 Background

India is the seventh largest country in the world by geographical area and world's second-most populous country with more than 1.38 crore people. According to Census of India 2011, there are 7935 cities. The number of metropolitan cities having million plus population has also increased from 35 to 53 as per 2011 census. The name of such cities can be found [here](#). In the current modern world citizens migrate from one place to another for a better opportunity. As per the [article](#), Bangalore's population comprises of 42% migrants.

1.2 Problem

The world is facing an unprecedented threat from the COVID-19 pandemic. It is difficult for a citizen who wants to explore the whole new city and choose the locality in order to find a place of reside.

1.3 Interest

The aim of the project is to compare the cities of India to find the similarities based on the venues in the neighbourhood. This helps the working class people who may have to move to new city and are looking for the comfortable locality to stay. The locality they would want to choose is dependent on the lifestyle they carry that is in turn dependent majorly on the venues in the nearby locality. For example some people prefer to stay in a neighbourhood which has restaurants, shopping malls in the near vicinity whereas some people prefer staying near to parks, recreational zones etc. In this project, I will be making an attempt to compare the similarities of cities Pune and Bangalore City, and this methodology can be extended to compare any cities of India. Both these cities are culturally diversified, vibrant and full of opportunities for a working professional.

2. Data collection and cleaning

2.1 Data Sources

Since post office is a public facility that provides mail services, including accepting of letters and parcels, providing post office boxes to every nook and corner of the country, I will be using the All India Pincode [Dataset](#) from the Open Government Data(OGD) Platform India to derive the neighbourhoods of the city. The dataset has attributes like Office name, Pincode, Office type, Division, Region, Taluk, District, State, Latitude and Longitude.

```
1 all_india_postal_df = pd.read_csv("all_india_PO_list_without_APS_offices_ver2_lat_long.csv")
2 all_india_postal_df.head()
```

	officename	pincode	officeType	Deliverystatus	divisionname	regionname	circlename	Taluk	Districtname	statename	Telephone	Related Suboffice
0	Achalapur B.O	504273	B.O	Delivery	Adilabad	Hyderabad	Andhra Pradesh	Asifabad	Adilabad	TELANGANA	NaN	Rechin S.C
1	Ada B.O	504293	B.O	Delivery	Adilabad	Hyderabad	Andhra Pradesh	Asifabad	Adilabad	TELANGANA	NaN	Asifabad S.C
2	Adegaon B.O	504307	B.O	Delivery	Adilabad	Hyderabad	Andhra Pradesh	Boath	Adilabad	TELANGANA	NaN	Echod S.C
3	Adilabad Collectorate S.O	504001	S.O	Non-Delivery	Adilabad	Hyderabad	Andhra Pradesh	Adilabad	Adilabad	TELANGANA	08732-226703	NaN
4	Adilabad H.O	504001	H.O	Delivery	Adilabad	Hyderabad	Andhra Pradesh	Adilabad	Adilabad	TELANGANA	08732-226738	NaN

```
1 all_india_postal_df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 154797 entries, 0 to 154796
Data columns (total 15 columns):
officename      154797 non-null object
pincode         154797 non-null int64
officeType      154797 non-null object
Deliverystatus  154797 non-null object
divisionname    154797 non-null object
regionname      154797 non-null object
circlename      154797 non-null object
Taluk           146488 non-null object
Districtname    154761 non-null object
statename       154761 non-null object
Telephone       23418 non-null object
Related Suboffice 125269 non-null object
Related Headoffice 153987 non-null object
longitude       142 non-null float64
latitude        142 non-null float64
dtypes: float64(2), int64(1), object(12)
memory usage: 17.7+ MB
```

Figure 1: Dataset from OGD India platform with data for post offices.

On examining, we find that the dataset is not updated with the latitudes and the longitudes for all the records. To solve this, I will be using geopy, which is a Python client for several popular geocoding web services. The geopy makes it easy to locate the coordinates of addresses, cities, countries, and landmarks across the globe using third-party geocoders and other data sources.

We will also use the FourSquare API to retrieve the venues within the radius of 1.5 km of each neighbourhood. Foursquare is a location data provider with information of all type of venues and events within an area of interest. Such information includes venue names, locations, latitudes, longitudes, venue categories, etc.

2.2 Data Cleaning and feature selection

Since we are interested in comparing the neighbourhood of Pune and Bangalore City, we will filter out the data of other cities. Drop the features like officeType, Deliverystatus, divisionname, regionname, circlename, Telephone, Related Suboffice, Related Headoffice since they are not necessary for our analysis. We will also remove the duplicate data from the dataset. We need latitudes and longitudes, but since the dataset is yet to be updated with the details, we will be dropping these features from the dataset and will use the geopy client to retrieve the latitudes and longitudes.

1	pune_blore_df.head()										
	officename	pincode	officeType	Deliverystatus	divisionname	regionname	circlename	Taluk	Districtname	statename	Telephone
52175	A F Station Yelahanka S.O	560063	S.O	Delivery	Bangalore East	Bangalore HQ	Karnataka	Bangalore North	Bangalore	KARNATAKA	080-28478031
52176	Agram S.O	560007	S.O	Delivery	Bangalore East	Bangalore HQ	Karnataka	Bangalore South	Bangalore	KARNATAKA	080-25300192
52177	Air Force Hospital S.O	560007	S.O	Non-Delivery	Bangalore East	Bangalore HQ	Karnataka	Bangalore North	Bangalore	KARNATAKA	080-25302699
52178	Amruthahalli B.O	560092	B.O	Delivery	Bangalore East	Bangalore HQ	Karnataka	Bangalore North	Bangalore	KARNATAKA	NaN
52179	Anandnagar S.O (Bangalore)	560024	S.O	Non-Delivery	Bangalore East	Bangalore HQ	Karnataka	Bangalore North	Bangalore	KARNATAKA	080-23431761

Figure 2: Filtered data for Pune and Bangalore

To fetch the coordinates of the neighbourhood we need to pass the address, city to the API and hence I will be creating a derived feature Neighbourhood with the concatenation of officename, Districtname and statename. Dataset with the latitudes and longitudes for Bangalore is as below:

1	pune_blore_df.head()								
	officename	pincode	Taluk	Districtname	statename	Neighbourhood	Latitude	Longitude	
0	A F Station Yelahanka	560063	Bangalore North	Bangalore	KARNATAKA	A F Station Yelahanka, Bangalore, KARNATAKA			
1	Agram	560007	Bangalore South	Bangalore	KARNATAKA	Agram, Bangalore, KARNATAKA			
2	Air Force Hospital	560007	Bangalore North	Bangalore	KARNATAKA	Air Force Hospital, Bangalore, KARNATAKA	12.964	77.6275	
3	Amruthahalli	560092	Bangalore North	Bangalore	KARNATAKA	Amruthahalli, Bangalore, KARNATAKA	13.0665	77.5966	
5	Arabic College	560045	Bangalore North	Bangalore	KARNATAKA	Arabic College, Bangalore, KARNATAKA	13.03	77.6209	

Figure 3: Pune and Bangalore dataset with latitudes and longitudes

We have 105 and 147 neighbourhoods along with geographical coordinates for Pune and Bangalore respectively. We will use this data to find the similar neighbourhoods in both the cities.

3. Methodology

3.1 Neighbourhood Visualisation on the map

Visualising the neighbourhoods on the map for Pune and Bangalore using folium package

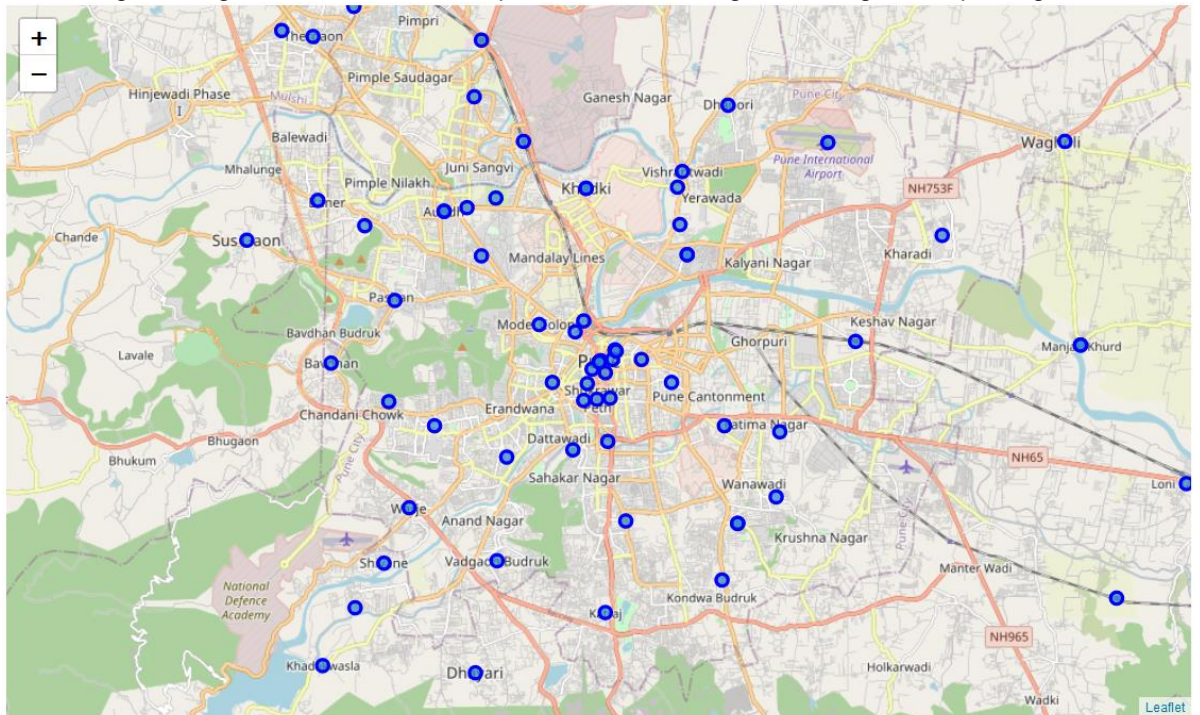


Figure 4: Map of Pune with the neighbourhoods

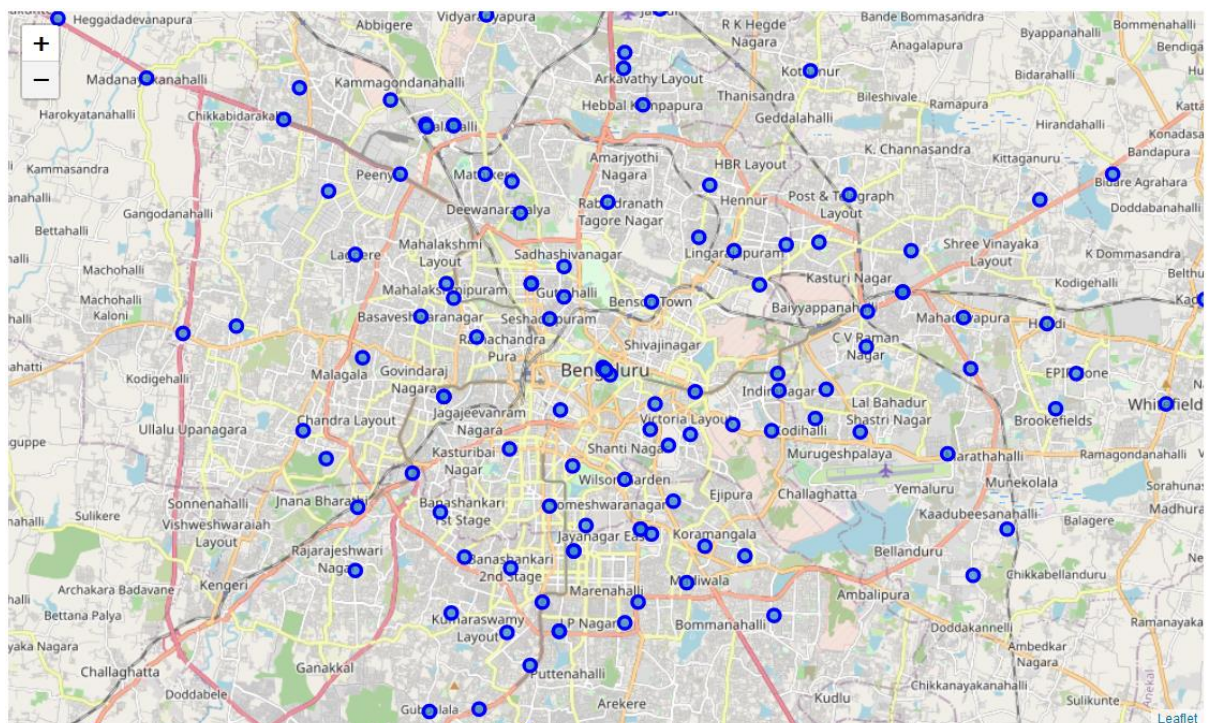


Figure 5: Map of Bangalore with the neighbourhoods

We see that most of the neighbourhood points on the map are closely and uniformly placed. We also see some overlapping, but this should not affect us as we will group them based on

the similarity of the neighbourhoods and overlapping neighbourhoods will come under same group. Hence it would not affect us to achieve the required results.

3.2 Data pre-processing to apply the machine learning technique

To find the similar neighbourhoods in both the cities we need to represent the neighbourhoods in terms of some meaningful attributes and then employ a unsupervised machine learning clustering technique.

One of the ways to represent the neighbourhoods is based on the kind of venues that exist in the neighbourhood. We will use the FourSquare's Explore API to get the list of the venues in the neighbourhood. The Explore API requires the latitude, longitude, radius and a number to limit the venues to be passed. The API returns the list of venues along with the details like: name, address, latitude, longitude, venue type etc. As per the figure 4 and 5, we will choose the radius to be 1500 meters since the neighbourhoods are closely and uniformly placed and a limit of 100 venues.

Using the latitude, longitude for each record, we invoke the Explore REST API provided by FourSquare to fetch the Venue details. From the result returned by the FourSquare API, we will extract the venue name, its latitude, longitude and the venue category. Figure 6 shows the venues in each neighbourhoods of Pune and Bangalore.

```
1 print(punebangalore_venues.shape)
2 punebangalore_venues.head()
```

(9654, 7)

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Air Force Hospital, Bangalore, KARNATAKA	12.964027	77.6275	League of Extraordinary Gamers	12.967099	77.636919	Gaming Cafe
1	Air Force Hospital, Bangalore, KARNATAKA	12.964027	77.6275	Corner House	12.969973	77.637453	Ice Cream Shop
2	Air Force Hospital, Bangalore, KARNATAKA	12.964027	77.6275	Vivanta by Taj	12.973365	77.619951	Hotel
3	Air Force Hospital, Bangalore, KARNATAKA	12.964027	77.6275	Retro Lounge Bar	12.972614	77.626700	Bar
4	Air Force Hospital, Bangalore, KARNATAKA	12.964027	77.6275	Foodhall	12.973486	77.620117	Department Store

Figure 6: Dataframe with the venues in a neighbourhood

Neighbourhood	Venue
AFMC, Pune, MAHARASHTRA	29
Adugodi, Bangalore, KARNATAKA	100
Air Force Hospital, Bangalore, KARNATAKA	87
Akurdi, Pune, MAHARASHTRA	26
Ambegaon BK, Pune, MAHARASHTRA	4
Amruthahalli, Bangalore, KARNATAKA	31
Anjanapura, Bangalore, KARNATAKA	7
Arabic College, Bangalore, KARNATAKA	14
Armament, Pune, MAHARASHTRA	13
Ashoknagar S.O, Bangalore, KARNATAKA	10
Attur, Bangalore, KARNATAKA	14
Aundh T.S., Pune, MAHARASHTRA	49
Aundh, Pune, MAHARASHTRA	49
Austin Town, Bangalore, KARNATAKA	100
Bagalgunte, Bangalore, KARNATAKA	8
Bajirao Road, Pune, MAHARASHTRA	96
Banashankari III Stage, Bangalore, KARNATAKA	25
Banashankari, Bangalore, KARNATAKA	76
Banaswadi, Bangalore, KARNATAKA	42
Baner Gaon, Pune, MAHARASHTRA	49
Baner Road, Pune, MAHARASHTRA	47
Bangalore City, Bangalore, KARNATAKA	87
Bangalore Corporation Building, Bangalore, KARNATAKA	87
Bangalore, Bangalore, KARNATAKA	87
Bannerghatta Road, Bangalore, KARNATAKA	65

Figure 7: Number of venues in each Neighbourhood

In total we find 9654 venues in the 241 neighbourhoods within a radius of 1500 meters, with an average of 40 venues per neighbourhood and 280 unique venue categories. Figure 7 shows the number of venues in each neighbourhood.

The have the data in the expected format to employ the clustering technique to identify the similar neighbourhoods, we will apply one-hot encoding on each of the venues, then we group the neighbourhoods based on the mean frequency of the occurrence of each category.

```
11 punebangalore_onehot.head()
```

	Neighbourhood	ATM	Accessories Store	Afghan Restaurant	African Restaurant	Airport	Airport Service	Airport Terminal	American Restaurant	Andhra Restaurant	...	Warehouse Store	Wedding Hall	Whisky Bar
0	Air Force Hospital, Bangalore, KARNATAKA	0	0	0	0	0	0	0	0	0	...	0	0	0
1	Air Force Hospital, Bangalore, KARNATAKA	0	0	0	0	0	0	0	0	0	...	0	0	0
2	Air Force Hospital, Bangalore, KARNATAKA	0	0	0	0	0	0	0	0	0	...	0	0	0
3	Air Force Hospital, Bangalore, KARNATAKA	0	0	0	0	0	0	0	0	0	...	0	0	0
4	Air Force Hospital, Bangalore, KARNATAKA	0	0	0	0	0	0	0	0	0	...	0	0	0

5 rows × 281 columns

Figure 8: Result of one-hot encoding


```

1 punebangalore_grouped = punebangalore_onehot.groupby('Neighbourhood').mean().reset_index()
2 print(punebangalore_grouped.shape)
3 punebangalore_grouped

```

(241, 281)

	Neighbourhood	ATM	Accessories Store	Afghan Restaurant	African Restaurant	Airport	Airport Service	Airport Terminal	American Restaurant	Andhra Restaurant	...	Warehouse Store	Wedding Hall
0	AFMC, Pune, MAHARASHTRA	0.000000	0.000000	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.000000	...	0.0	0.034485
1	Adugodi, Bangalore, KARNATAKA	0.000000	0.000000	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.000000	...	0.0	0.000000
2	Air Force Hospital, Bangalore, KARNATAKA	0.000000	0.000000	0.000000	0.0	0.000000	0.0	0.0	0.011494	0.000000	...	0.0	0.000000
3	Akurdi, Pune, MAHARASHTRA	0.000000	0.000000	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.000000	...	0.0	0.000000
4	Ambegaon BK, Pune, MAHARASHTRA	0.000000	0.000000	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.000000	...	0.0	0.000000

Figure 9: The neighbourhoods represented in terms of venues by grouping based on mean frequency of the venue category

Figure 8 shows the result of one-hot encoding technique. Figure 9 shows the final dataset with the neighbourhoods represented in terms of the venues grouped based on the mean frequency of the venue category.

Analysing this dataset to get the top 10 most common venues in the dataset, we get the below:

```

20 neighbourhoods_venues_sorted.head()

```

	Neighbourhood	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	AFMC, Pune, MAHARASHTRA	Indian Restaurant	Shopping Mall	Bar	Breakfast Spot	Multiplex	Ice Cream Shop	Pizza Place	Department Store	Burger Joint	Electronics Store
1	Adugodi, Bangalore, KARNATAKA	Indian Restaurant	Café	Dessert Shop	Chinese Restaurant	Ice Cream Shop	Bookstore	Pizza Place	Lounge	Coffee Shop	Mexican Restaurant
2	Air Force Hospital, Bangalore, KARNATAKA	Indian Restaurant	Café	Hotel	Asian Restaurant	Chinese Restaurant	Bar	Ice Cream Shop	Pizza Place	Pub	Restaurant
3	Akurdi, Pune, MAHARASHTRA	Café	Indian Restaurant	Asian Restaurant	Snack Place	Ice Cream Shop	Convenience Store	Diner	Middle Eastern Restaurant	Fast Food Restaurant	Gaming Cafe
4	Ambegaon BK, Pune, MAHARASHTRA	Hotel	Vegetarian / Vegan Restaurant	Asian Restaurant	Indian Restaurant	Electronics Store	English Restaurant	Event Space	Exhibit	Falafel Restaurant	Farm

Figure 10: Dataset with top 10 most common venues for each neighbourhood

3.2 Employing K-Means clustering on the dataset

We first need to choose the number of clusters as 7 and then perform clustering on the punebangalore_grouped dataframe using the k-means clustering. K-means clustering algorithm identifies the k number of centroids, and then allocates every data point to the nearest cluster centroid. K-means clustering is one of the simplest, popular unsupervised machine learning algorithms and is suited to solve the problem of this project.

Figure 12: Neighbourhood Clusters on Bangalore map

4.2 Examining each cluster

- List of neighbourhoods in Cluster 0 (Violet circle markers on the maps):

151	411001	C D A (O), Pune, MAHARASHTRA	18.4624	73.8897	0.0	Coffee Shop	Flea Market	Warehouse Store	Multiplex	Mughlai Restaurant	Department Store	Gym
152	411031	C M E, Pune, MAHARASHTRA	18.4619	73.8895	0.0	Flea Market	Warehouse Store	Multiplex	Mughlai Restaurant	Department Store	Gym	Café
153	411062	Chikhali, Pune, MAHARASHTRA	18.6642	73.8267	0.0	Fast Food Restaurant	Dessert Shop	Accessories Store	Café	Electronics Store	English Restaurant	Event Space
156	411012	Dapodi, Pune, MAHARASHTRA	18.5808	73.8328	0.0	Snack Place	Department Store	Fish Market	Food Truck	Restaurant	Train Station	Hotel
159	411003	East Khadki, Pune, MAHARASHTRA	18.5682	73.8508	0.0	Department Store	Soccer Field	Food Truck	Bakery	Train Station	Chinese Restaurant	Farmers Market
160	411028	Hadapsar, Pune, MAHARASHTRA	18.527	73.9278	0.0	Indian Restaurant	Donut Shop	Coffee Shop	Italian Restaurant	Chinese Restaurant	Multiplex	Sandwich Place
163	411003	Khadki Bazar, Pune, MAHARASHTRA	18.5621	73.838	0.0	Department Store	Bakery	Diner	Indian Restaurant	Soccer Field	Train Station	Food Truck
164	411003	Khadki, Pune, MAHARASHTRA	18.5682	73.8508	0.0	Department Store	Soccer Field	Food Truck	Bakery	Train Station	Chinese Restaurant	Farmers Market
165	411048	Kondhwa BK, Pune, MAHARASHTRA	18.4777	73.8941	0.0	Café	Coffee Shop	Ice Cream Shop	Asian Restaurant	Restaurant	Hookah Bar	Breakfast Spot
166	411048	Kondhwa Lh, Pune, MAHARASHTRA	18.4777	73.8941	0.0	Café	Coffee Shop	Ice Cream Shop	Asian Restaurant	Restaurant	Hookah Bar	Breakfast Spot
168	411048	N I B M, Pune, MAHARASHTRA	18.462	73.8898	0.0	Coffee Shop	Flea Market	Warehouse Store	Multiplex	Mughlai Restaurant	Department Store	Gym
176	411037	Salisbury Park, Pune, MAHARASHTRA	19.0632	72.8262	0.0	Café	Bakery	Indian Restaurant	Seafood Restaurant	Lounge	Asian Restaurant	Ice Cream Shop
183	411040	Wanowarie, Pune, MAHARASHTRA	18.4848	73.9049	0.0	Café	Coffee Shop	Ice Cream Shop	American Restaurant	Bakery	Breakfast Spot	Fast Food Restaurant

Figure 13: Cluster 0 neighbourhoods in Pune

	pincode	Neighbourhood	Latitude	Longitude	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
4	560047	Austin Town, Bangalore, KARNATAKA	12.9613	77.6153	0.0	Hotel	Indian Restaurant	Café	Clothing Store	Bakery	Pub	Chinese Restaurant
13	560036	Devasandra, Bangalore, KARNATAKA	12.7572	77.6473	0.0	Henan Restaurant	Zoo Exhibit	Farm	Eastern European Restaurant	Electronics Store	English Restaurant	Event Space
17	560066	EPIP, Bangalore, KARNATAKA	12.9782	77.7246	0.0	Café	Hotel	Indian Restaurant	Restaurant	Fast Food Restaurant	Brewery	Cafeteria
21	560051	H K P, Road, Bangalore, KARNATAKA	13.0623	77.6613	0.0	Pizza Place	Convenience Store	Vegetarian / Vegan Restaurant	Brewery	Sandwich Place	Stadium	Restaurant
23	560048	Hoodi, Bangalore, KARNATAKA	12.9919	77.7162	0.0	Brewery	Indian Restaurant	Coffee Shop	Hotel	Bakery	Platform	Movie Theater
24	560043	Horamavu, Bangalore, KARNATAKA	13.0273	77.6602	0.0	Department Store	Andhra Restaurant	Badminton Court	Kerala Restaurant	Café	Asian Restaurant	Clothing Store
27	560064	Jakkur, Bangalore, KARNATAKA	13.0785	77.6069	0.0	Café	Hot Spring	Airport	Tennis Court	Lake	Farm	Electronics Store
28	560067	Kadugodi, Bangalore, KARNATAKA	12.9986	77.761	0.0	Hotel	Train Station	South Indian Restaurant	Pizza Place	Fruit & Vegetable Store	Plaza	Indian Restaurant
31	560077	Kothanur, Bangalore, KARNATAKA	13.0614	77.6494	0.0	Mediterranean Restaurant	Stadium	Supermarket	Coffee Shop	Brewery	Café	North Indian Restaurant

Figure 14: Cluster 0 neighbourhoods in Bangalore

We see that the clusters have more restaurants, departmental stores, café, coffee shop, brewery. This neighbourhood will be well liked by people who prefer a dine-out with friends, colleagues and family. Airport, Train are also found in the locality and is also suitable for people who travel to other cities frequently.

- List of neighbourhoods in Cluster 1 (Indigo circle markers on the maps):

	pincode	Neighbourhood	Latitude	Longitude	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
40	560067	Medimallasandra, Bangalore, KARNATAKA	12.9523	77.8301	1.0	Scenic Lookout	Zoo Exhibit	Farm	Electronics Store	English Restaurant	Event Space	Exhibit	Falafel Restaurant	Farm
236	411025	Khanapur Haveli, Pune, MAHARASHTRA	18.4007	73.7318	1.0	Scenic Lookout	Zoo Exhibit	Farm	Electronics Store	English Restaurant	Event Space	Exhibit	Falafel Restaurant	Farm

Figure 15 : Cluster 1 neighbourhoods in Pune and Bangalore.

The neighbourhood has Scenic lookout, zoo, farms, few shops and restaurants etc and is located far away from the city. We see that only 2 neighbourhoods come under this cluster. And is a good place for a day out.

- Cluster 2 (Blue circle markers on the maps):

197	411004	Deccan Gymkhana, Pune, MAHARASHTRA	18.5159	73.8412	2.0	Indian Restaurant	Snack Place	Vegetarian / Vegan Restaurant	Fast Food Restaurant	Seafood Restaurant	Ice Cream Shop	Restaurant	F
199	411004	Film Institute, Pune, MAHARASHTRA	18.5214	73.8545	2.0	Indian Restaurant	Coffee Shop	Restaurant	Fast Food Restaurant	Dessert Shop	Vegetarian / Vegan Restaurant	Café	
200	411007	Ganeshkhind, Pune, MAHARASHTRA	18.5499	73.8208	2.0	Indian Restaurant	Snack Place	Clothing Store	Fast Food Restaurant	Coffee Shop	Restaurant	Café	E
201	411042	Ghorpade Peth, Pune, MAHARASHTRA	18.5246	73.8589	2.0	Indian Restaurant	Restaurant	Ice Cream Shop	Hotel	Seafood Restaurant	Bus Station	Snack Place	
202	411042	Guruwar Peth, Pune, MAHARASHTRA	18.5114	73.8576	2.0	Indian Restaurant	Snack Place	Restaurant	Ice Cream Shop	Vegetarian / Vegan Restaurant	Dessert Shop	Clothing Store	
203	411011	Kasba Peth, Pune, MAHARASHTRA	18.5219	73.8583	2.0	Indian Restaurant	Ice Cream Shop	Fast Food Restaurant	Restaurant	Café	Seafood Restaurant	Snack Place	B
204	411046	Katraj, Pune, MAHARASHTRA	18.4537	73.8563	2.0	Snack Place	Indian Restaurant	Café	Food	Department Store	Convenience Store	Exhibit	B
206	411038	Kothrud, Pune, MAHARASHTRA	18.5039	73.8077	2.0	Indian Restaurant	Ice Cream Shop	Café	Dessert Shop	Pizza Place	Coffee Shop	Fast Food Restaurant	
207	411023	Mandvi BK, Pune, MAHARASHTRA	18.9551	72.8348	2.0	Indian Restaurant	Café	Dessert Shop	Jewelry Store	Snack Place	BBQ Joint	Fast Food Restaurant	
208	411011	Mangalwar Peth S.O, Pune, MAHARASHTRA	18.5243	73.8592	2.0	Indian Restaurant	Restaurant	Ice Cream Shop	Hotel	Fast Food Restaurant	Snack Place	Seafood Restaurant	B
209	411016	Model Colony, Pune, MAHARASHTRA	18.5314	73.8375	2.0	Indian Restaurant	Coffee Shop	Lounge	Café	Sandwich Place	Multiplex	Ice Cream Shop	F
212	411002	Nana Peth, Pune, MAHARASHTRA	18.5184	73.8563	2.0	Indian Restaurant	Fast Food Restaurant	Vegetarian / Vegan Restaurant	Snack Place	Dessert Shop	Restaurant	Seafood Restaurant	
214	411030	Narayan Peth, Pune, MAHARASHTRA	18.5156	73.8511	2.0	Indian Restaurant	Vegetarian / Vegan Restaurant	Restaurant	Dessert Shop	Snack Place	Fast Food Restaurant	Coffee Shop	

Figure 16 : Cluster 2 neighbourhoods in Pune.

	pincode	Neighbourhood	Latitude	Longitude	Cluster Labels	Common Venue	Common Venue	Common Venue	Common Venue	Common Venue	Common Venue	Common Venue
0	560007	Air Force Hospital, Bangalore, KARNATAKA	12.964	77.6275	2.0	Indian Restaurant	Café	Hotel	Asian Restaurant	Chinese Restaurant	Bar	Ice Cream Shop
1	560092	Amruthahalli, Bangalore, KARNATAKA	13.0665	77.5966	2.0	Indian Restaurant	Ice Cream Shop	Department Store	Café	Hotel	Fast Food Restaurant	Brewery
2	560045	Arabic College, Bangalore, KARNATAKA	13.03	77.6209	2.0	Pizza Place	Indian Restaurant	Convenience Store	Clothing Store	BBQ Joint	Jewelry Store	Bus Station
5	560043	Banaswadi, Bangalore, KARNATAKA	13.0142	77.6519	2.0	Indian Restaurant	Bakery	Café	Falafel Restaurant	Coffee Shop	Pizza Place	Lounge
6	560103	Bellandur, Bangalore, KARNATAKA	12.9358	77.6668	2.0	Indian Restaurant	Fast Food Restaurant	Café	Kerala Restaurant	Gym	Lounge	Department Store
7	560046	Benson Town, Bangalore, KARNATAKA	12.9978	77.6042	2.0	Indian Restaurant	Bakery	Middle Eastern Restaurant	Ice Cream Shop	Tea Room	BBQ Joint	Café
8	560049	Bhattarahalli, Bangalore, KARNATAKA	13.0258	77.7143	2.0	Indian Restaurant	Café	Hotel	Korean Restaurant	Pizza Place	Bar	Vegetarian / Vegan Restaurant
9	560092	Byatarayanapura, Bangalore, KARNATAKA	13.0621	77.5964	2.0	Indian Restaurant	Ice Cream Shop	Pizza Place	Café	Fast Food Restaurant	Department Store	Cosmetics Shop
10	560093	C.V.Raman Nagar, Bangalore, KARNATAKA	12.9856	77.665	2.0	Indian Restaurant	Department Store	Coffee Shop	Café	Fast Food Restaurant	Sandwich Place	Restaurant
11	560025	CMP Centre And School, Bangalore, KARNATAKA	12.9583	77.6091	2.0	Indian Restaurant	Hotel	Café	Bakery	Pub	Chinese Restaurant	Coffee Shop
14	560037	Doddanekkundi, Bangalore, KARNATAKA	12.9794	77.6947	2.0	Indian Restaurant	Coffee Shop	Sandwich Place	Motorcycle Shop	Department Store	Fast Food Restaurant	BBQ Joint

Figure 17 : Cluster 2 neighbourhoods in Bangalore.

The neighbourhood has more Indian restaurants, coffee shop, café, fast food, ice cream shops and is suitable for people who like to hang out at evening.

- Cluster 3 (Green circle markers on the maps):

	pincode	Neighbourhood	Latitude	Longitude	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 M Comm Ver	
	63	560108	Anjanapura, Bangalore, KARNATAKA	12.8581	77.5581	3.0	ATM	Asian Restaurant	Pharmacy	Pool	Campground	Photography Lab	Airport Service	Dump Restau
	116	560091	Herohalli, Bangalore, KARNATAKA	12.9912	77.487	3.0	ATM	Market	Gas Station	Fried Chicken Joint	Electronics Store	Performing Arts Venue	Financial or Legal Service	F
	123	560058	Laggere, Bangalore, KARNATAKA	13.0109	77.5207	3.0	ATM	Vegetarian / Vegan Restaurant	Pizza Place	Electronics Store	Gym	Pharmacy	Henan Restaurant	D
	124	560023	Magadi Road, Bangalore, KARNATAKA	12.989	77.4719	3.0	ATM	Auto Workshop	Resort	Gas Station	Eastern European Restaurant	Electronics Store	English Restaurant	Event Sp
	158	411015	Dighi Camp, Pune, MAHARASHTRA	18.6265	73.8661	3.0	ATM	Pharmacy	Coffee Shop	Garden	Hotel	Wine Shop	Falafel Restaurant	Dump Restau
	198	411041	Dhayari, Pune, MAHARASHTRA	18.4374	73.819	3.0	ATM	Cheese Shop	Bus Station	Tea Room	Szechuan Restaurant	English Restaurant	Event Space	Ext
	213	411041	Nanded, Pune, MAHARASHTRA	18.4548	73.7849	3.0	ATM	Pizza Place	Market	Indian Restaurant	Farm	Eastern European Restaurant	Electronics Store	Eng Restau
	224	411023	Shivane, Pune, MAHARASHTRA	18.467	73.7931	3.0	ATM	Pizza Place	Mobile Phone Shop	Market	Diner	Indian Restaurant	Department Store	C
	231	412109	Dehu, Pune, MAHARASHTRA	18.68	73.7343	3.0	ATM	Train Station	Mobile Phone Shop	Juice Bar	Coworking Space	Creperie	Fish Market	Convenie S
	244	412105	Moshi, Pune, MAHARASHTRA	18.6523	73.8442	3.0	ATM	Pizza Place	Indian Restaurant	Restaurant	Farm	Exhibit	Dumpling Restaurant	East Europ Restau
<												>		

Figure 18 : Cluster 3 neighbourhoods in Pune and Bangalore.

The neighbourhood has market places with lots of ATMs, Auto workshops, Restaurants and Pizza shops.

- Cluster 4 (Yellow circle markers on the maps):

	pincode	Neighbourhood	Latitude	Longitude	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
19	560087	Gunjur, Bangalore, KARNATAKA	13.4279	77.5131	4.0	Resort	Trail	Lake	Zoo Exhibit	Exhibit	Eastern European Restaurant	Electronics Store	English Restaurant
46	560064	Rajanakunte, Bangalore, KARNATAKA	13.1692	77.5639	4.0	Resort	Train Station	Fast Food Restaurant	Zoo Exhibit	Falafel Restaurant	Eastern European Restaurant	Electronics Store	English Restaurant
105	560074	Ramohalli, Bangalore, KARNATAKA	12.907	77.4185	4.0	Farm	Zoo Exhibit	Farmers Market	Electronics Store	English Restaurant	Event Space	Exhibit	Falafel Restaurant
139	560090	Tarabanahalli, Bangalore, KARNATAKA	13.1906	77.6279	4.0	Resort	ATM	Diner	Indian Restaurant	Farm	Electronics Store	English Restaurant	Event Space
232	411025	Donje, Pune, MAHARASHTRA	18.4122	73.7666	4.0	Farm	Resort	Hotel	Bed & Breakfast	Indian Restaurant	Electronics Store	English Restaurant	Event Space
248	412115	Sangrun, Pune, MAHARASHTRA	18.3996	73.6768	4.0	Resort	Zoo Exhibit	Donut Shop	Eastern European Restaurant	Electronics Store	English Restaurant	Event Space	Exhibit

Figure 19: Cluster 4 neighbourhoods in Pune and Bangalore

The neighbourhood has resorts, hotels for dining, lake, trails, farms and will be well liked nature lovers. A nearby weekend getaway with lakes and trails.

- Cluster 5 (Orange circle markers on the maps):

	pincode	Neighbourhood	Latitude	Longitude	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
3	560064	Attur, Bangalore, KARNATAKA	13.0995	77.568	5.0	Indian Restaurant	Bakery	BBQ Joint	Soccer Field	Bus Station	Asian Restaurant	Andhra Restaurant
12	560064	CRPF Campus Yelahanka, Bangalore, KARNATAKA	13.1214	77.5709	5.0	Indian Restaurant	Gas Station	Resort	Coffee Shop	Food Truck	Liquor Store	Dhaba
38	560037	Marathahalli Colony, Bangalore, KARNATAKA	12.9559	77.6882	5.0	Indian Restaurant	Fast Food Restaurant	Airport Terminal	Andhra Restaurant	Clothing Store	Hotel	Dessert Shop
57	560047	Viveknagar S.O., Bangalore, KARNATAKA	12.9543	77.6227	5.0	Indian Restaurant	Café	Gym / Fitness Center	General Entertainment	Nightclub	Juice Bar	Golf Course
77	560053	Chickpet, Bangalore, KARNATAKA	12.968	77.5786	5.0	Indian Restaurant	Hotel	Theater	Diner	Bed & Breakfast	Café	Bookstore
82	560100	Electronics City, Bangalore, KARNATAKA	12.8469	77.6769	5.0	Indian Restaurant	Food Court	Café	Vegetarian / Vegan Restaurant	Bakery	Residential Building (Apartment / Condo)	Basketball Court
84	560083	Gottigere, Bangalore, KARNATAKA	12.8565	77.5877	5.0	Indian Restaurant	Department Store	Food Truck	Italian Restaurant	Gym / Fitness Center	Farmers Market	English Restaurant
108	560061	Subramanyapura, Bangalore, KARNATAKA	12.8851	77.5416	5.0	Indian Restaurant	South Indian Restaurant	Bus Station	Garden	Arts & Entertainment	Supermarket	Department Store
110	560109	Thalaghattapura, Bangalore, KARNATAKA	12.8698	77.5357	5.0	Indian Restaurant	Lounge	Gas Station	Supermarket	Zoo Exhibit	Electronics Store	English Restaurant
149	411026	Bhosari I.E., Pune, MAHARASHTRA	18.621	73.8501	5.0	Indian Restaurant	Flea Market	Fast Food Restaurant	Zoo Exhibit	Falafel Restaurant	Eastern European Restaurant	Electronics Store
161	411017	Kalewadi, Pune, MAHARASHTRA	18.4044	74.0214	5.0	Indian Restaurant	Hotel	Farm	Eastern European Restaurant	Electronics Store	English Restaurant	Event Space
162	411034	Kasanwadi, Pune, MAHARASHTRA	18.6081	73.8209	5.0	Indian Restaurant	Bus Station	Food Truck	Cosmetics Shop	Hotel Bar	Platform	Snack Place

Figure 20: Cluster 5 neighbourhoods in Pune and Bangalore

The neighbourhood has Indian restaurants, flea market, Gym, bakery super markets, garden.

5. Conclusion and future directions

The purpose of this project was to demonstrate the methodology to explore and find similarities of the neighborhoods based on the venues in the neighborhoods in any two cities. This idea can be applied to any other cities of India. The venues have been identified using the Foursquare API and have been plotted on the map. Further we could include the cost of living by including the rental data for creating clusters, there are lot of websites which advertise and display the rental price of the properties.

6. References

- <http://mohua.gov.in/cms/number-of-cities--towns-by-city-size-class.php>
- <https://www.thehindu.com/news/cities/bangalore/migrants-constitute-42-of-bengalurus-population/article28734588.ece>