

The Battle of Neighborhoods – Kannur and Kozhikode

Introduction

Kannur and Kozhikode are two major cities in Kerala, India. Both cities become a center of attention for residential, job employment, tourism, education, and shopping and sports activity. Both cities are well known in Kerala, India.

Brief information about both cities:

Kannur is a city and a Municipal Corporation in Kannur district, state of Kerala, India. Kannur is the largest city of North Malabar region.[1] As of 2011 census population of Kannur was 232,486.[2] Kannur is one of the million-plus urban agglomerations in India with a population of 1,642,892 in 2011. (Source: <https://en.wikipedia.org/wiki/Kannur>)

Kozhikode, also known as Calicut, is an Indian city, second-largest urban agglomeration in the State of Kerala and 20th largest in the country with a population of 2 million as of 2011 (Source: <https://en.wikipedia.org/wiki/Kozhikode>)

Objective

In this project, we will study in details the area classification using Foursquare data and machine learning segmentation and clustering. The aim of this project is to segment areas of Kannur and Kozhikode based on the most common places captured from Foursquare.

Using segmentation and clustering, we hope we can determine:

1. The similarity or dissimilarity of both cities
2. Classification of area located inside the city whether it is residential, tourism places, or others.

Data

The data acquired from wikipedia pages are converted to csv file for easier manipulation and reading. Both files uploaded to my github for references. Link to the files are:

<https://github.com/sharikaavi/CapstoneProject/blob/master/Kannur.csv>

<https://github.com/sharikaavi/CapstoneProject/blob/master/Kozhikode.csv>

This data includes the pincode, Taluk and Areas in Kannur and Kozhikode

In this project, we will study in details the area classification using foursquare data and machine learning segmentation and clustering. The aim of this project is to segment areas of Kannur and Kozhikode based on the most common places captured from Foursquare.

Methodology

At first, latitude and longitude of every area of Kannur and Kozhikode are found.

[5]:

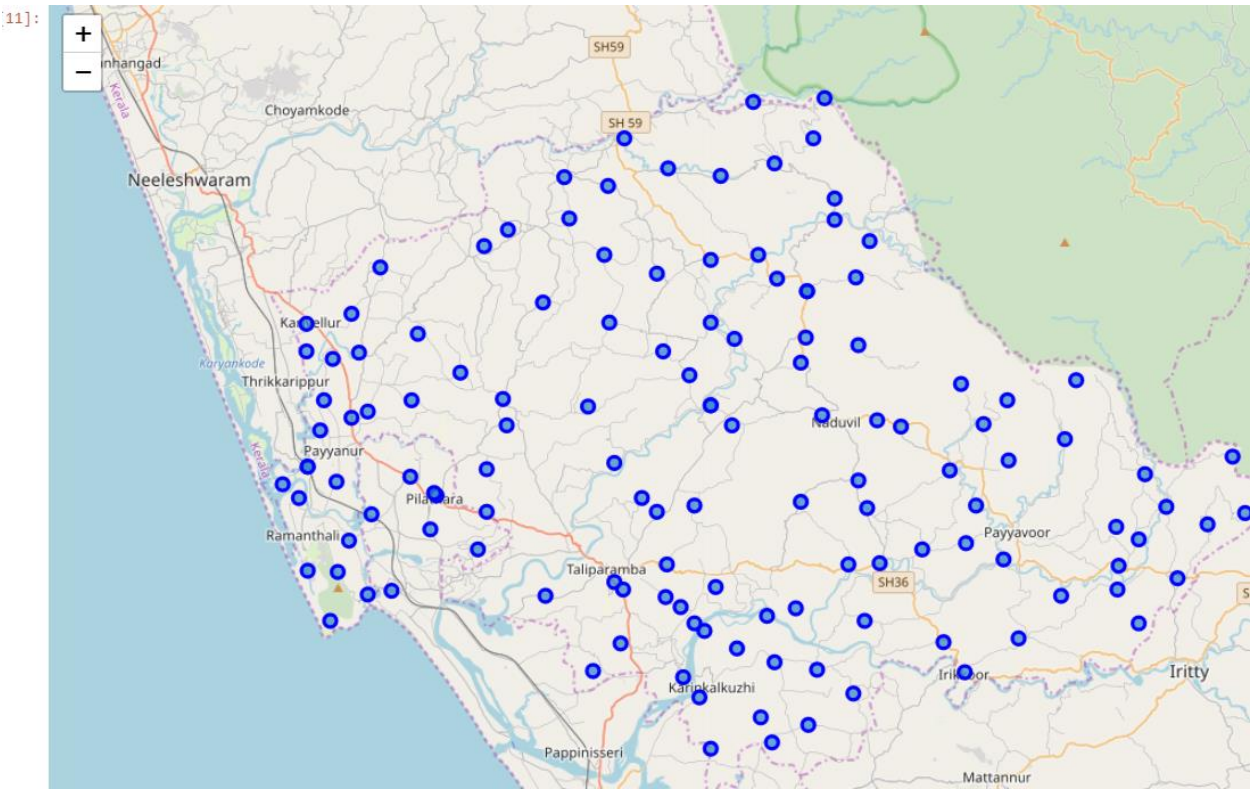
	Pincode	Taluk	Area	Latitude	Longitude
0	670571	Taliparamba	Alakode	12.190978	75.467331
1	670571	Taliparamba	Alakode kuttaramba	12.190978	75.467331
2	670008	Kannur	Alavil	11.901894	75.346538
3	670331	Kannur	Anchampeedika	11.532411	75.730335
4	670612	Thalassery	Anjarakandy	11.885834	75.485089
5	670307	Taliparamba	Annur	12.132956	75.202438
6	670582	Taliparamba	Arang	8.556909	76.983869
7	670582	Taliparamba	Areekamala	12.132892	75.577157
8	670143	Taliparamba	Aril	10.850516	76.271083
9	670571	Taliparamba	Arivilanjapoyil	11.350407	75.914712

[6]:

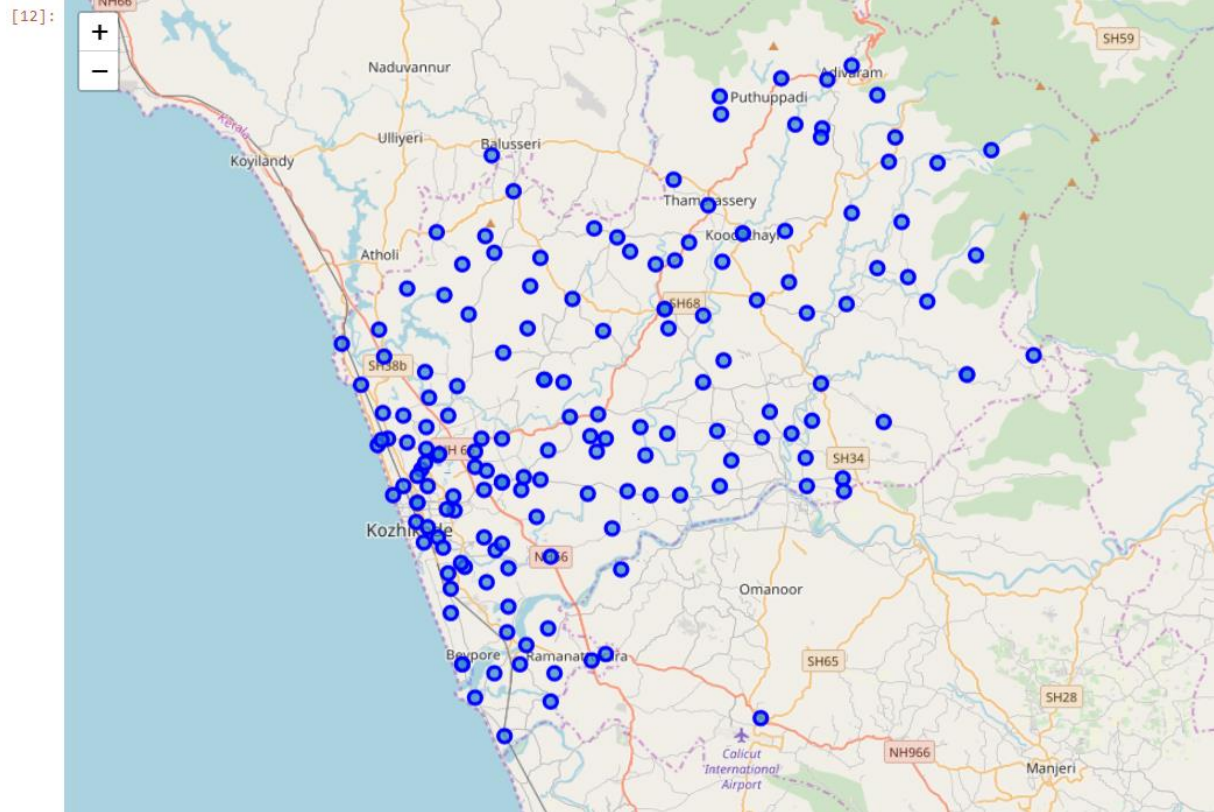
	Pincode	Taluk	Area	Latitude	Longitude
0	673586	Kozhikode	Adivaram Pudupadi	11.488039	76.013078
1	673602	Kozhikode	Alli	9.967117	76.288968
2	673603	Kozhikode	Anakampoyil	11.436896	76.058772
3	673028	Kozhikode	Arakinar	11.200921	75.798190
4	673572	Kozhikode	Avilora	11.383467	75.908301
5	673015	Kozhikode	Bey pore North	11.173585	75.804002
6	673015	Kozhikode	Bey pore	11.173585	75.804002
7	673018	Kozhikode	Calicut Arts & Science College	11.213722	75.798346
8	673032	Kozhikode	Calicut Beach	11.262623	75.767309
9	673020	Kozhikode	Calicut Civil Station	11.283644	75.791106

Foursquare API is used to explore neighborhoods in both cities, Kannur and Kozhikode.

Taliparamba Area in Kannur is found have the highest number of areas.



Kozhikode Area in Kozhikode is found to have the highest number of areas.



Most common venue categories in each neighborhood are found and then used this feature to group the neighborhoods into clusters.

Results

K-mean clustering algorithm is used to complete this task. And the **Folium library** to visualize the neighborhoods in Kannur and Kozhikode and their emerging clusters.

Analyze Kannur

Kannur –Cluster 1

Area	Bus Station	Café	Clothing Store	Coffee Shop	Cosmetics Shop	Currency Exchange	Fabric Shop	Fast Food Restaurant	Gym / Fitness Center	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	Latitude	Longitude
Arang	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.00	0.0	0.000000	Video Store	Cluster Labels	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	11.901694	75.346538
Chengalayi	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.00	0.0	0.000000	Plaza	Cluster Labels	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	11.885834	75.485089
Cherupuzha	0.333333	0.000000	0.0	0.0	0.0	0.000000	0.00	0.0	0.000000	Bus Station	Music Store	Indian Restaurant	Cluster Labels	Jewelry Store	Bakery	Basketball Court	Burger Joint	8.556909	76.983869
Thundakunnu	0.000000	0.000000	0.0	0.0	0.0	1.000000	0.00	0.0	0.000000	Currency Exchange	Cluster Labels	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	12.132892	75.577157
Irakkur	0.500000	0.000000	0.0	0.0	0.0	0.000000	0.00	0.0	0.000000	Bus Station	Playground	Cluster Labels	Jewelry Store	Bakery	Basketball Court	Burger Joint	Café	10.850516	76.271083
Kandath	0.000000	0.000000	0.0	0.0	1.0	0.000000	0.00	0.0	0.000000	Cosmetics Shop	Cluster Labels	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	11.984557	75.380936
Kokkanisseri	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.00	0.0	0.000000	Basketball Court	Cluster Labels	Jewelry Store	Bakery	Burger Joint	Bus Station	Café	Clothing Store	11.867886	75.431233
Kottayad	0.100000	0.000000	0.2	0.1	0.0	0.000000	0.00	0.0	0.000000	Bakery	Clothing Store	Jewelry Store	Vegetarian / Vegan Restaurant	Shopping Mall	Bus Station	Coffee Shop	Outlet Mall	11.917065	75.335387
Chimangalam	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.00	0.0	0.000000	Train Station	Platform	Cluster Labels	Indian Restaurant	Bakery	Basketball Court	Burger Joint	Bus Station	11.917065	75.335387
uniyempuzha	0.142857	0.142857	0.0	0.0	0.0	0.000000	0.00	0.0	0.142857	Motorcycle Shop	Indian Restaurant	Bus Station	Café	Restaurant	Gym / Fitness Center	Cluster Labels	Bakery	11.908579	75.332472
andur kannur	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.00	0.0	0.000000	Restaurant	Cluster Labels	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	11.865832	75.497096
hamangalam Bazar	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.00	0.0	0.000000	ATM	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	Clothing Store	11.859131	75.368909

Kannur –Cluster 2

	Area	Bus Station	Café	Clothing Store	Coffee Shop	Cosmetics Shop	Currency Exchange	Fabric Shop	Fast Food Restaurant	Gym / Fitness Center	...	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	Latitude	Longitude
0	Alakode	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	Cluster Labels	Indian Restaurant	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	12.190978	75.467331
1	Alakode kuttaramba	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	Cluster Labels	Indian Restaurant	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	12.190978	75.467331
3	Aril	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	Cluster Labels	Indian Restaurant	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	11.532411	75.730335
5	Cherikode	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	Cluster Labels	Indian Restaurant	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	12.132956	75.202438
9	Kandankali	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	Cluster Labels	Indian Restaurant	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	11.350407	75.914712
18	Nareekamvally	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	Cluster Labels	Bus Station	Indian Restaurant	Jewelry Store	Bakery	Basketball Court	Burger Joint	Café	11.883045	75.348506
20	Pathampara	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	Cluster Labels	Indian Restaurant	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	12.111562	75.608442
23	Pazhassikari	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	Cluster Labels	Indian Restaurant	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	11.946222	75.414078
24	Pilathara	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	Cluster Labels	Bus Station	Indian Restaurant	Jewelry Store	Bakery	Basketball Court	Burger Joint	Café	12.089595	75.495640

9 rows × 38 columns

Kannur –Cluster 3

	Area	Bus Station	Café	Clothing Store	Coffee Shop	Cosmetics Shop	Currency Exchange	Fabric Shop	Fast Food Restaurant	Gym / Fitness Center	...	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	Latitude	Longitude
21	Payyan R.S.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	...	Cluster Labels	Fast Food Restaurant	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	11.958785	75.467972
22	Payyavur	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	...	Cluster Labels	Fast Food Restaurant	Jewelry Store	Bakery	Basketball Court	Burger Joint	Bus Station	Café	11.949375	75.447582

2 rows × 38 columns

Analyze Kozhikode

Cluster 1

	Area	...	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	Latitude	Longitude
	Arakinar	...	Hotel	Café	Harbor / Marina	Halal Restaurant	Gym / Fitness Center	Gym	Grocery Store	Furniture / Home Store	11.436896	76.058772
	Beypore	...	Grocery Store	Cluster Labels	Café	Harbor / Marina	Halal Restaurant	Gym / Fitness Center	Gym	Furniture / Home Store	11.200921	75.798190
	Beypore North	...	Grocery Store	Cluster Labels	Café	Harbor / Marina	Halal Restaurant	Gym / Fitness Center	Gym	Furniture / Home Store	11.383467	75.908301
	Calicut	...	Indian Restaurant	Hotel	Fast Food Restaurant	Juice Bar	Mobile Phone Shop	Pizza Place	Dessert Shop	Bowling Alley	11.173585	75.804002
	Calicut Beach	...	Ice Cream Shop	Indian Restaurant	Restaurant	Café	Cluster Labels	Fast Food Restaurant	Coffee Shop	Convenience Store	11.213722	75.798346
	Calicut Courts	...	Bus Stop	Kerala Restaurant	Middle Eastern Restaurant	Cluster Labels	Fish Market	Convenience Store	Department Store	Dessert Shop	11.284319	75.791957
	Calicut R.S.	...	Movie Theater	Indian Restaurant	Platform	Market	Asian Restaurant	Hotel	Train Station	Juice Bar	11.258753	75.780410
	Cheruvadi	...	Boutique	Cluster Labels	Fish Market	Coffee Shop	Convenience Store	Department Store	Dessert Shop	Fast Food Restaurant	11.240626	75.790925
	East Hill	...	Airport	Movie Theater	Cluster Labels	Fish Market	Coffee Shop	Convenience Store	Department Store	Dessert Shop	11.357631	75.807721
	Edakkad West Hill	...	Café	Cluster Labels	Harbor / Marina	Halal Restaurant	Gym / Fitness Center	Gym	Grocery Store	Furniture / Home Store	11.269570	75.825751
	Eliathur Kozhikode	...	Train Station	Cluster Labels	Fish Market	Clothing Store	Coffee Shop	Convenience Store	Department Store	Dessert Shop	11.450359	76.036290
	Eranhipalam	...	Café	Indian Restaurant	Bus Stop	Tea Room	Bakery	Fish Market	Coffee Shop	Convenience Store	11.295025	75.980794
	Farook College	...	Juice Bar	Cluster Labels	Café	Halal Restaurant	Gym / Fitness	Gym	Grocery Store	Furniture / Home Store	11.262577	75.921173

Cluster 2

	Area	Arts & Crafts Store	Asian Restaurant	Bakery	Beach	Boat or Ferry	Boutique	Bowling Alley	Bridal Shop	Burger Joint	...	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	Latitude	Longitude
	Adivaram Pudurupadi	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Indian Restaurant	Neighborhood	Fish Market	Coffee Shop	Convenience Store	Department Store	Dessert Shop	11.488039	76.013078
	Alli	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Indian Restaurant	Café	Halal Restaurant	Gym / Fitness Center	Gym	Grocery Store	Furniture / Home Store	10.850516	76.271083
	Calicut Medical College	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Indian Restaurant	Clothing Store	Cafeteria	Airport Terminal	Coffee Shop	Accessories Store	Halal Restaurant	11.275790	75.782236
	Chennamangallur	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Indian Restaurant	Café	Halal Restaurant	Gym / Fitness Center	Gym	Grocery Store	Furniture / Home Store	11.248657	75.779806
	Kutaranni	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Indian Restaurant	Café	Halal Restaurant	Gym / Fitness Center	Gym	Grocery Store	Furniture / Home Store	11.192389	75.850490
	Mailellampara	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Indian Restaurant	Café	Halal Restaurant	Gym / Fitness Center	Gym	Grocery Store	Furniture / Home Store	11.173455	75.835243
	Marikunnu	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Indian Restaurant	Café	Halal Restaurant	Gym / Fitness Center	Gym	Grocery Store	Furniture / Home Store	11.233603	75.822446
	Puduppadi	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Indian Restaurant	Café	Halal Restaurant	Gym / Fitness Center	Gym	Grocery Store	Furniture / Home Store	11.455389	75.997118

Cluster 3

	Area	Arts & Crafts Store	Asian Restaurant	Bakery	Beach	Boat or Ferry	Boutique	Bowling Alley	Bridal Shop	Burger Joint	...	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	Latitude	Longitude
	Calicut Arts & Science College	0.0	0.0	0.333333	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Market	Bakery	Furniture / Home Store	Gym	Grocery Store	Gym / Fitness Center	Halal Restaurant	11.173585	75.804002
	Calicut Civil Station	0.0	0.0	0.333333	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Restaurant	Furniture / Home Store	Bakery	Fast Food Restaurant	Coffee Shop	Convenience Store	Department Store	11.262623	75.767309
	Calicut Collectorate	0.0	0.0	0.333333	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Restaurant	Furniture / Home Store	Bakery	Fast Food Restaurant	Coffee Shop	Convenience Store	Department Store	11.283644	75.791106
	Chelavur	0.0	0.0	0.250000	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Indian Restaurant	Restaurant	Bakery	Café	Fast Food Restaurant	Coffee Shop	Convenience Store	11.272211	75.837198
	Cheruvannur	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	River	Furniture / Home Store	Fast Food Restaurant	Clothing Store	Coffee Shop	Convenience Store	Department Store	11.155988	75.811267
	Chevarambalam	0.0	0.0	0.500000	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Gym	Bakery	Fish Market	Coffee Shop	Convenience Store	Department Store	Dessert Shop	11.462840	75.943003
	Chevayur	0.0	0.0	0.250000	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Indian Restaurant	Restaurant	Bakery	Café	Fast Food Restaurant	Coffee Shop	Convenience Store	11.294967	75.914387
	Kakkoti	0.0	0.0	1.000000	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Bakery	Fried Chicken Joint	Coffee Shop	Convenience Store	Department Store	Dessert Shop	Fast Food Restaurant	11.277539	75.811267
	Kannancheri	0.0	0.0	0.285714	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Bakery	Gym / Fitness Center	Indian Restaurant	Supermarket	Juice Bar	Market	Fish Market	11.296241	75.940664
	Karaparamba	0.0	0.0	0.333333	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Men's Store	Fast Food Restaurant	Bakery	Coffee Shop	Convenience Store	Department Store	Dessert Shop	11.290320	75.774641
	Kundungal	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Furniture / Home Store	Café	Harbor / Marina	Halal Restaurant	Gym / Fitness Center	Gym	Grocery Store	11.279690	75.784386
	Parappanpoyil	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	--	Cluster Labels	Restaurant	Fast Food Restaurant	Clothing Store	Coffee Shop	Convenience Store	Department Store	Dessert Shop	11.358390	75.988775

Discussion

Based on cluster for each cities, we believe that classification for each cluster can be done better with calculation of venues categories (most common) in each cities. Referring to each cluster, we can't determine clearly what represent in each cluster by using Foursquare - Most Common Venue data.

We assumed each cluster as follow:

Cluster 1: Kannur: Tourism

Cluster 2: Kannur: Residential

Cluster 3: Kannur: Mix

Cluster 1: Kozhikode: Tourism

Cluster 2: Kozhikode: Residential

Cluster 3: Kozhikode: Mix

What is lacking at this point is a systematic, quantitative way to identify and distinguish different district and to describe the correlation most common venues as recorded in Foursquare. The reality is however more complex: similar cities might have or might not have similar common venues. A further step in this classification would be to find a method to extract these common venues and integrate the spatial correlations between different of areas or district.

We believe that the classification we propose is an encouraging step towards a quantitative and systematic comparison of the different cities. Further studies are indeed needed in order to relate the data acquired, and then observe it to more meaningful and objective results.

Conclusion

Using Foursquare API, we can captured data of common places all around the world. Using it, we refer back to our main objectives, which is to determine; the similarity or dissimilarity of either cities classification of area located inside the city whether it is residential, tourism places or others In conclusion, both cities Kannur and Kozhikode are the center of attraction among Kerala. However, to declare either cities are similar or dissimilar base on common venues visited is quite difficult. Both cities is similar in some venues also dissimilar in certain venues. Moreover, for classification based on common venues, again we must have more systematic or quantitative way to identify and declare this. Comparison can be made, but no such method or quantitative data to determine this. We hope in the future, a method to determine it can be establish and explore for references.

Thank you,

Sharika

Kannur