



KOLLAM TO VELLORE RELOCATION USING K - MEANS

By Shaun Oommen Alexander

PROBLEM STATEMENT

- This project is based on a problem that my friend faced in college. I currently study at Vellore Institute of Technology, Vellore as a 4th year Engineering student. My friend and I both moved from Kerala, India to Tamil Nadu, India for our higher studies. While moving to a new location it can be really tough to find out the right location on the basis of accessibility, distance, services etc. With this project I am looking to help my friend who is based in Kollam, Kerala to find the right location to shift to in Vellore, Tamil Nadu.

HOW DOES THIS PROJECT HELP?

- The reason I consider this to be a tough problem is because Vellore is quite underdeveloped and it can be hard to find a place to live, while Kollam is quite developed with lots of different venues and locations of different types. So for someone who is moving from an area like Kollam to Vellore, this project will try to find similar locations in Vellore to locations in Kollam using clustering. Here the k means clustering algorithm is used to achieve the task. Folium library can be used to visualize the clusters in both cities.

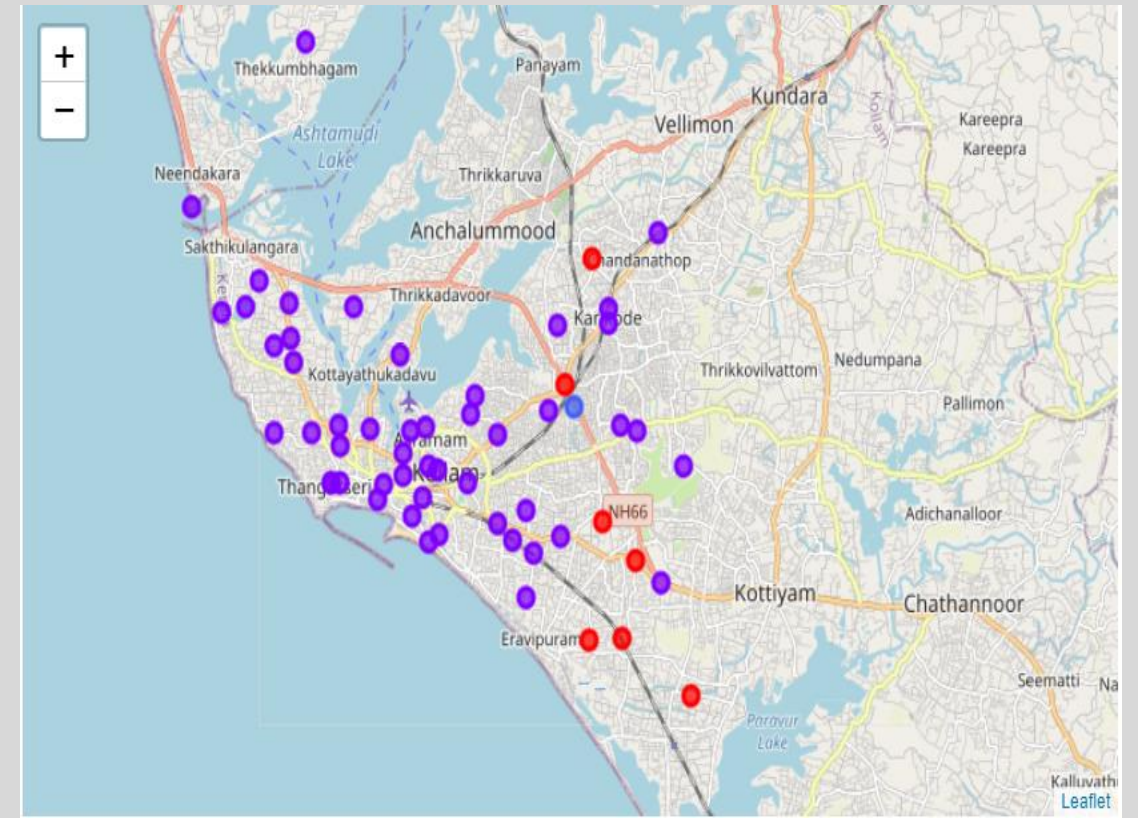
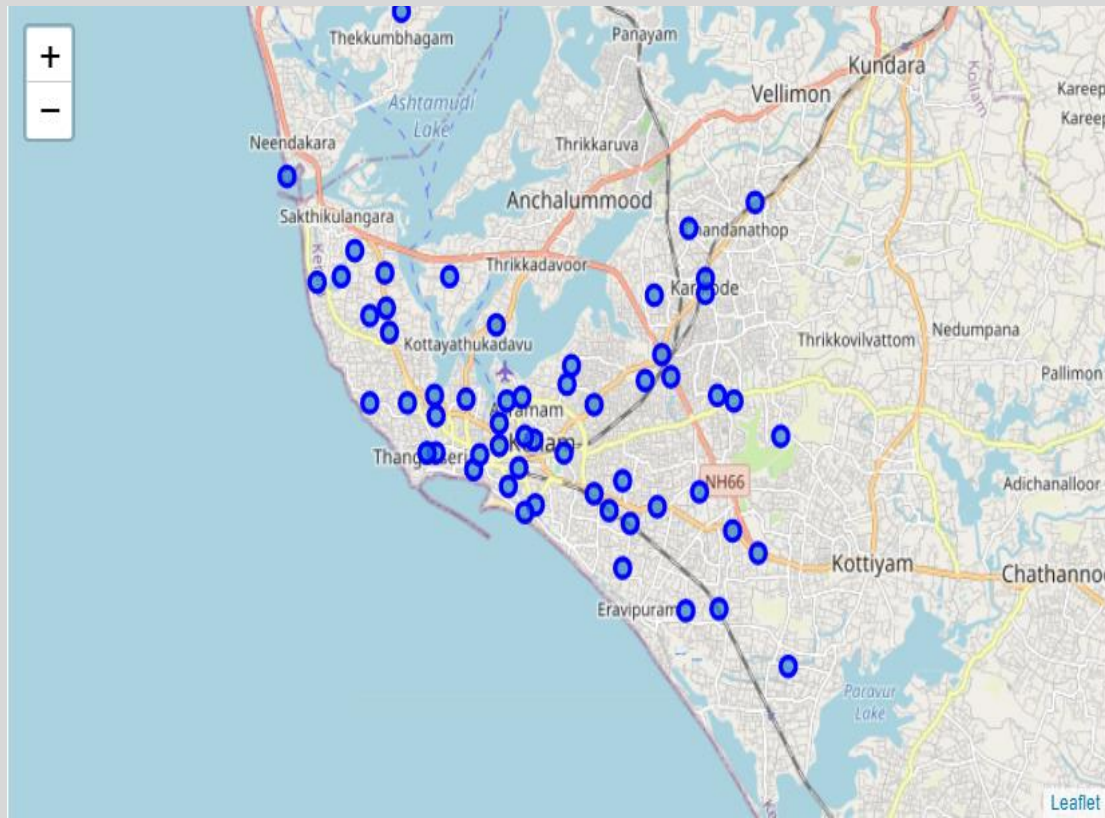
DATA DESCRIPTION

- Neighbourhood data is obtained by scraping the web using the BeautifulSoup library from the following links:
- https://en.wikipedia.org/wiki/List_of_areas_of_Vellore
- https://en.wikipedia.org/wiki/List_of_neighbourhoods_of_Kollam
- I also used the FourSquare API to obtain all the venues within 2000 metres of the location. Using this venue data we will be able to find locations in vellore with a similar profile

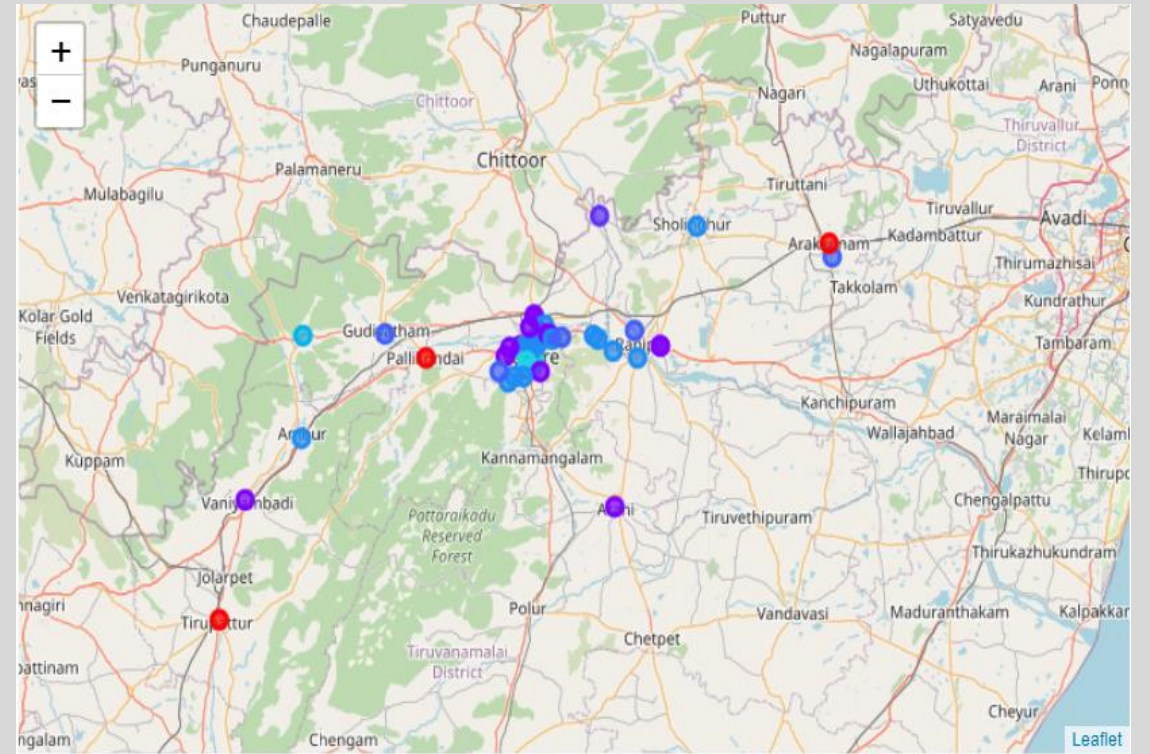
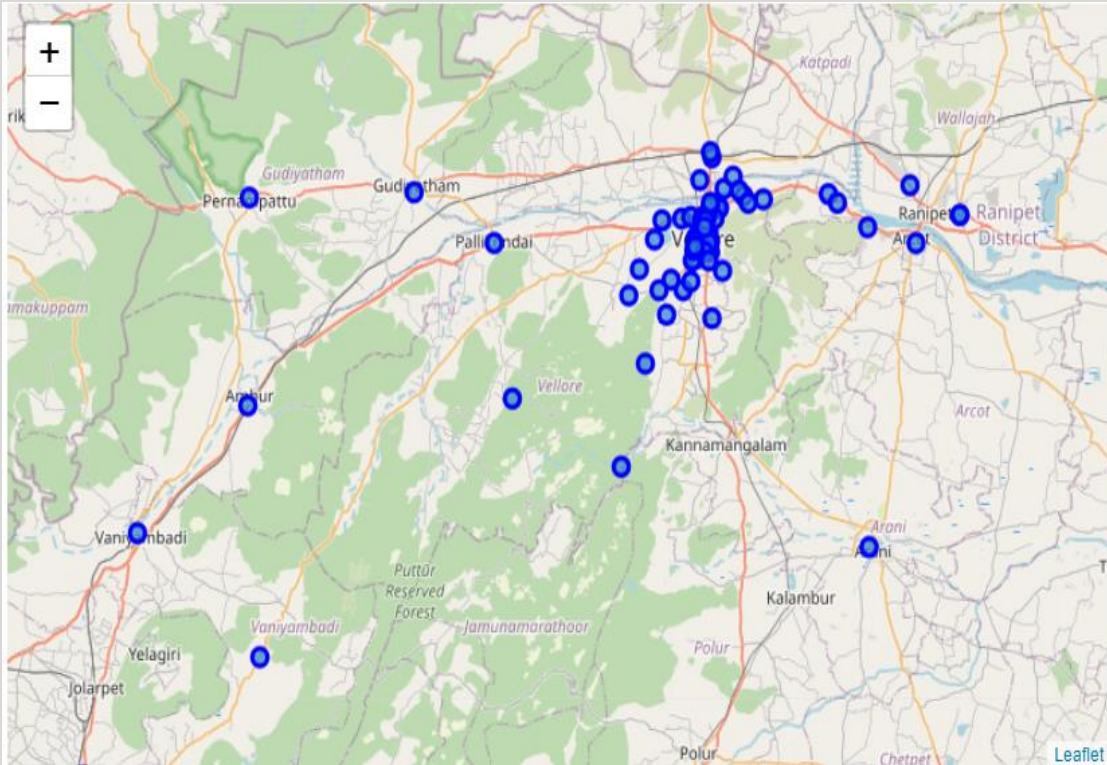
The following data are obtained from the Foursquare API:

1. Venue
2. Venue Latitude
3. Venue Longitude
4. Venue Category Data

Neighbourhoods at Kollam before and after clustering



Neighbourhoods at Vellore before and after clustering



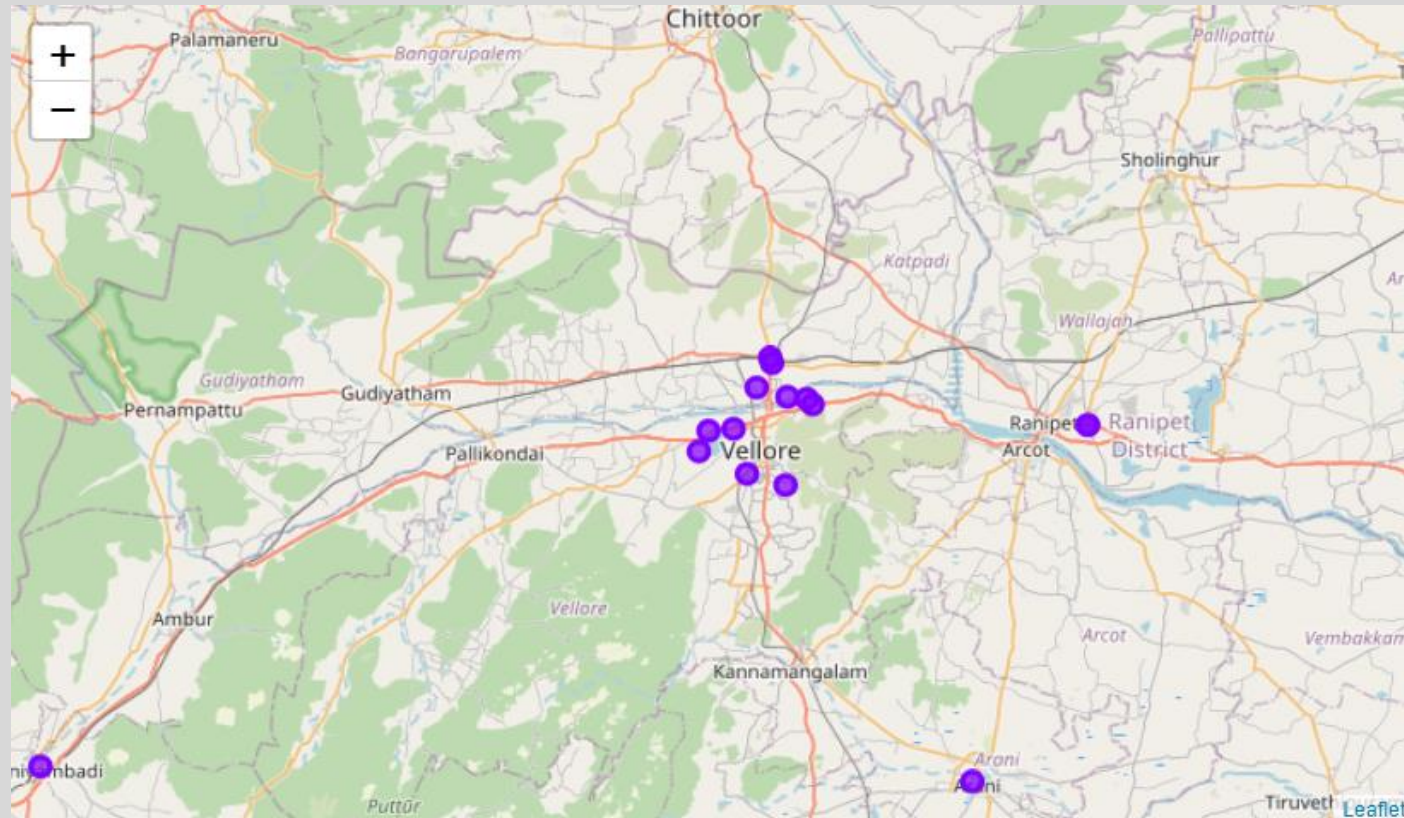
ANALYSIS

- As we can see there are a lot of different types of neighbourhoods in Vellore as compared to Kollam. To explain this, we need some background on Vellore. The reason being that Vellore is home to two major institutes Vellore Institute of Technology and Christian Medical College or CMC. The area around these institutes have much more venues than other remote locations. This results in neighbourhoods belonging to a more number of clusters than when compared to Kollam where all the neighbourhoods have similar nearby venues.

RESULTS

- In this Section we will be visualizing the various clusters obtained from our methodology and the neighbourhoods most likely to relocate to from Kilikollur, Kollam which is the neighbourhood my friend is currently living in.
- Our aim is to find the best location for my friend to stay outside college and hence we need to find places closer to his University. For this we have used the geolocator package to get the latitude and longitude of VIT university and for each neighbourhood in the cluster found the distance and added it to the dataframe

NEIGHBOURHOODS IN VELLORE SIMILAR TO KILIKOLLUR



TOP 6 SUGGESTED NEIGHBOURHOODS

	Neighbourhood	City	Latitude	Longitude	Cluster Labels	Distance to VIT
5	Dharapadavedu	Vellore	12.9695	79.1386	1.0	2.269835
17	Katpadi	Vellore	12.9734	79.1369	1.0	2.448409
1	Gandhi Nagar	Vellore	12.9460	79.1492	1.0	3.051000
36	Vallalar	Vellore	12.9443	79.1633	1.0	3.060659
2	Sathuvachari	Vellore	12.9399	79.1681	1.0	3.642485
39	Kazhinjur	Vellore	12.9527	79.1280	1.0	4.003187

CONCLUSION

- In this project I have successfully found relocation options for my friend who is considering moving from Kilikollur, Kollam to Vellore and filtered the options with respect to closeness to his University as well as similarity to his current location of residence. This can also be considered to be a proof of concept to finding similar location in an unknown region using K Means clustering and location data.