



## Exploratory Analysis of Geolocational Data

<sup>1</sup>Rushabh Marathe, <sup>2</sup>Adarsh Jadhav

<sup>1</sup>Student, B.E. Mechanical, Dr. D Y Patil College of Engineering, Pune

<sup>2</sup>Student, B.E Information Technology, Dr. D Y Patil College of Engineering, Pune

DOI: <https://doi.org/10.55248/gengpi.2023.42005>

---

### ABSTRACT

This project uses the K-Means clustering method to find the best migrant shelters by ranking migrant shelters according to their amenities, budget and proximity preferences. Fetch, cleans, analyse and aggregates K-Means on geolocation data to recommend immigrant accommodation in the city.

**Key Concepts:** Data, Dataset, Recommendation, Map

---

### 1. Introduction

In this project, hotels, restaurants, swimming pools, super services and other needs are recommended to a user who has become accustomed to a new hosted area. It is difficult for the user to find all the places in the rehosted area. So it's easy when we fly near places. We are too tired to cook a home cooked meal often. Even if someone gets a home-cooked meal every day, it's not uncommon to want to dine out every once in a while to socialize. Anyway, the food you eat matters no matter where you live. When a person moves to a new place. You already have certain likes and tastes. This would save food consumers and suppliers many benefits by living close to their favorite outlets. This is convenient for the owners, leads to better sales and saves the user's time.

---

### 2. Related Topics

#### 2.1 Collection of Data and Setup an Environment

If we want to perform data analysis, we need to retrieve a data set and configure the required data analysis environment. You need a Jupyter notebook for the analysis.

#### 2.2 Data Cleaning and Visualization

After receiving the data and understanding what is written there. The best way to do this is to visualize data with charts. Charts help us understand data quickly and are much more intuitive.

#### 2.3 Pre-Process of Data

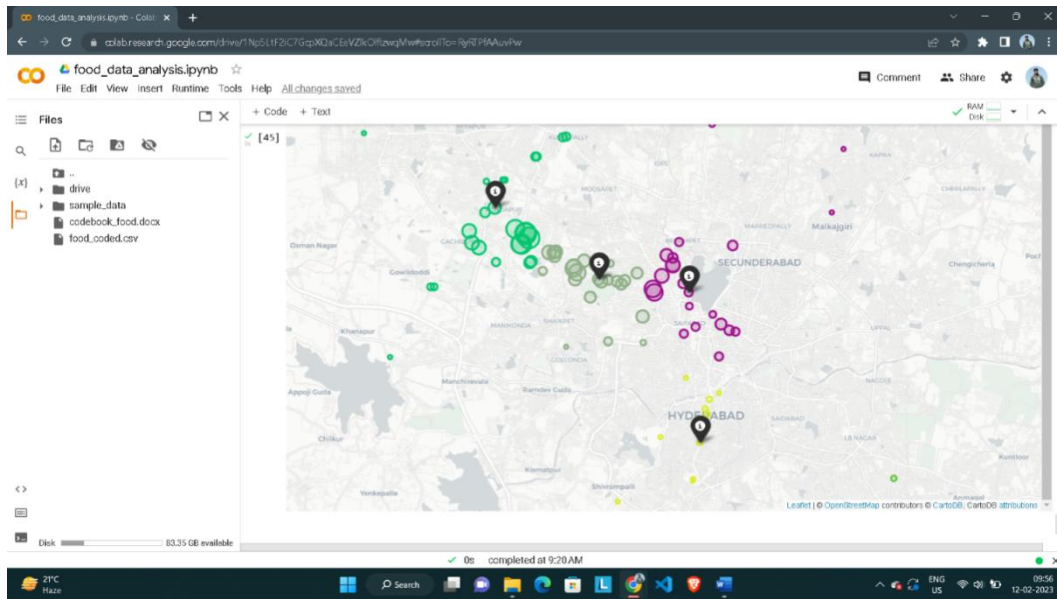
K means that clustering helps us group places based on the objects around them. A location with multiple lodgings nearby is flagged as "amenity rich." while a place with less gear is called "poor gear". Similar locations are grouped together. Run the K-Means clustering algorithm and find the best K-value to use on our website..

#### 2.4 Get Geo-Location Data using FourSquare API

Now that we've got the best cluster scores, we need to pull the geolocation data from the Foursquare API to find those homes for our users.

#### 2.5 Plotting Clusters on the Maps.

Atlast, present the results on a map where the user can view the locations.



**FIGURE 1: Clusters On The Map**

### 3. Existing Systems

The existing system includes a restaurant navigation app like Swiggy Zomato etc., supports ride sharing services like Uber Pool etc. Not recommended for our budget accommodations. There are rare cases of homes within our budget. Also does not recommend restaurants, gyms, etc., based on previous user budget research, lacks the accuracy of real recommendations.

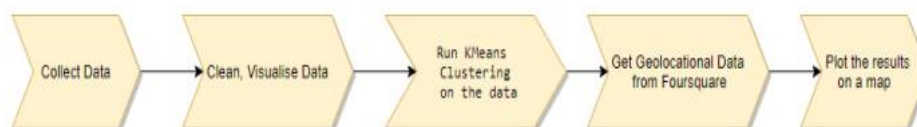
### 4. Proposed System

The proposed system recommends hostels, apartments and houses and also displays details of these houses, apartments and hostels. Recommend accommodation within our budget. It has big house boxes for our budget. Also recommend restaurants, hotels, over pool routes, etc. Based on user's budget. It offers genuine recommendations, of which there is no shortage. In this project we use the K-Means algorithm, but it has the disadvantage that two circular clusters centered on the same mean have different radii. K-Means uses the median to define the center of the cluster and does not distinguish between the two clusters.

#### 4.1 Ways to Approaches

- Fetch Datasets from the relevant locations (Data Collection)
- Clean the Datasets to prepare them for analysis. (Data Cleaning via Pandas)
- Visualise the data using boxplots. (Using Matplotlib /Seaborn /Pandas)
- Fetch Geolocal Data from the Foursquare API. (REST APIs)
- Use K-Means Clustering to cluster the locations (Using ScikitLearn)
- Present findings on a map. (Using Folium/Seaborn)

#### 4.2 Projects Stages



### 4.3 Modules

- Data Collection Module
- Searching Module
- Suggestion Module
- Communication Module

#### 4.3.1 Data Collection Module

Collect user data and store it in a database for later use.

#### 4.3.2 Searching Module

After entering the data, the user searches for a position with the desired services within his budget.

#### 4.3.3 Suggestion Module

After searching for the desired information, a recommendation will appear in the search bar. Depending on your budget and needs.

#### 4.3.4 Communication Module

From the recommendations displayed, the user selects the best accommodation and, thanks to the communication form, forwards it to the contact details of the owner.

---

## 5. Applications

- Academic performance (grouping students by their learning rate)
- Diagnostic systems (grouping system faults under various reasons)
- Search engines (grouping search results)
- Wireless sensor networks (Mapping networks)

---

## 6. Conclusion

This app/website is as user friendly as it is easy to use and affordable. A common problem of migrants is solved by this app/website. This app/website makes it easy to find accommodation, fits our budget and will be more useful for studying students.

### References

1. M. Sumithra and S. Malathi, "A survey on Medical Image Segmentation Methods with Different Modalities", International Journal of Engineering Research and Technology (IJERT) – Vol. 6 Issue 2, Mar 2018.
2. B.Buvaneswari and Dr.T. KalpalathaReddy, "ELSA- A Novel Technique to Predict Parkinson's Disease in Bio- Facial", International Journal of Advanced Trends in Computer Science and Engineering, ISSN 2278-3091, Vol.8, No.1, Pp. 12- 17, 2019
3. K. Sridharan , and Dr. M. Chitra , Proficient Information Retrieval Using Trust Based Search On Expert And Knowledge Users Query Formulation System, Australian Journal of Basic and Applied Sciences, 9(23) July 2015, Pages: 755-765.
4. Sharanyaa, S., and K. Sangeetha. "Blocking adult account in osn's using iterative social based classifier algorithm." International Journal of Scientific Engineering and Science 2, no. 1 (2018): 33-36.
5. B.Buvaneswari and Dr.T. Kalpalatha Reddy, "ACPT- An Intelligent Methodology for Disease Diagnosis", Journal of Advanced Research in Dynamical and Control Systems, ISSN : 0974-5572, Vol.11, No.4, Pp.2187-2194, 2019.
6. Sumithra, M., Shruthi, S., Ram, S., Swathi, S., Deepika, T., "MRI image classification of brain tumor using deep neural network and deployment using web framework", Advances in Parallel Computing, 2021, 38, pp. 614–617.
7. K. Sridharan , and Dr. M. Chitra "RSSE: A Paradigm for Proficient Information Retrieval using Semantic Web" , Life Science Journal 2013;10(7s), pp: 418-425

8. Sharanyaa, S., S. Vijayalakshmi, M. Therasa, U. Kumaran, and R. Deepika. "DCNET: A Novel Implementation of Gastric Cancer Detection System through Deep Learning Convolution Networks." In 2022 International Conference on Advanced Computing Technologies and Applications (ICACTA), pp. 1-5. IEEE, 2022.
9. B.Buvanswari and T.Kalpalatha Reddy, "A Review of EEG Based Human Facial Expression Recognition Systems in Cognitive Sciences" International Conference on Energy, Communication, Data analytics and Soft Computing (ICECDS), CFP17M55-PRJ:978-1-5386-1886-8", August 2017.
10. Chethana, C., Subbiah Swaminathan, S. Sharanyaa, E. Sathish, R. Prathipa, and Anuradha Thakare. "Application Of Reverse Engineering in the Process of Utilization of Human Brain in Artificial Intelligence." *Journal of Optoelectronics Laser* 41, no. 3 (2022): 89-93.
11. M. Sumithra and Dr. S. Malathi, "Modified Global Flower Pollination Algorithm-based image fusion for medical diagnosis using computed tomography and magnetic resonance imaging", *International Journal of Imaging Systems and Technology*, Vol. 31, Issue No.1, pp. 223-235, 2021
12. K. Sridharan, and Dr. M. Chitra "SBPE: A paradigm Approach for proficient Information Retrieval, *Jokull Journal*", Vol 63, No. 7; Jul 2013
13. Sharanyaa, S., P. N. Renjith, and K. Ramesh. "Classification of Parkinson's disease using speech attributes with parametric and nonparametric machine learning techniques." 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS). IEEE, 2020.
14. M. Sumithra and Dr. S. Malathi, "3D DenseNet Model with Back Propagation for Brain Tumor Segmentation", *International Journal Of Current Research and Review*, Vol. 13, Issue 12, 2021.
15. B.Buvaneswari and Dr.T. Kalpalatha Reddy, "EEG signal classification using soft computing techniques for brain disease diagnosis", *Journal of International Pharmaceutical Research*, ISSN : 1674-0440, Vol.46, No.1, Pp.525-528, 2019.
16. Sharanyaa, S., P. N. Renjith, and K. Ramesh. "An Exploration on Feature Extraction and Classification Techniques for Dysphonic Speech Disorder in Parkinson's Disease." In *Inventive Communication and Computational Technologies*, pp. 33-48. Springer, Singapore, 2022.
17. K. Sridharan, and Dr. M. Chitra "Web Based Agent And Assertion Passive Grading For Information Retrieval", *ARPN Journal of Engineering and Applied Sciences*, VOL. 10, NO. 16, September 2015 pp:7043-7048
18. M. Sumithra and Dr. S. Malathi, "Segmentation Of Different Modalities Using Fuzzy K-Means And Wavelet ROI", *International Journal Of Scientific & Technology Research*, Vol. 8, Issue 11, pp. 996-1002, November 2019.
19. Sharanyaa, S., S. Lavanya, M. R. Chandhini, R. Bharathi, and K. Madhulekha. "Hybrid Machine Learning Techniques for Heart Disease Prediction." *International Journal of Advanced Engineering Research and Science* 7, no. 3 (2020).
20. Butler, J. A. (2008). *Designing Geodatabases for Transportation*. Esri Press.
21. M. Sumithra and S. Malathi, "A Survey of Brain Tumor Segmentation Methods with Different Image Modalities", *International Journal of Computer Science Trends and Technology (IJCSST)* – Vol. 5 Issue 2, Mar – Apr 2017
22. B.Buvaneswari and Dr.T. Kalpalatha Reddy, "High Performance Hybrid Cognitive Framework for Bio-Facial Signal Fusion Processing for the Disease Diagnosis", *Measurement*, ISSN: 0263-2241, Vol. 140, Pp.89-99, 2019.
23. Sharanyaa, S., and M. Shubin Aldo. "Explore places you travel using Android." In 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), pp. 4796-4799. IEEE, 2016.
24. M. Sumithra and Dr. S. Malathi, "A Brief Survey on Multi Modalities Fusion", *Lecture Notes on Data Engineering and Communications Technologies*, Springer, 35, pp. 1031-1041, 2020.
25. Sharanyaa, S., and Madhumitha RP. "Eyeball Cursor Movement Detection Using Deep Learning." RP, Madhumitha and Rani. B, Yamuna, *Eyeball Cursor Movement Detection Using Deep Learning (July 12, 2021)* (2021).
26. Calixto, A., Diniz, F. B., and Zannin, P. (2003). The statistical modeling of road traffic noise in an urban setting. *Cities*, 20:1-74.
27. Arampatzis, G., Kiranoudis, C., Scaloubacas, P., and Assimacopoulos, D. (2004). A gis-based decision support system for planning urban transportation policies. *European Journal of Operational Research*, 152(2):465 – 475. *New Technologies in Transportation Systems*. Barczyszyn, G. L. (2015).
28. M. Sumithra and Dr. S. Malathi, "A Novel Distributed Matching Global and Local Fuzzy Clustering (DMGLFC) FOR 3D Brain Image Segmentation for Tumor Detection", *IETE Journal of Research*. doi.org/10.1080/03772063.2022.2027284, 2021