Elinvo (Electronics, Informatics, and Vocational Education)
Vol 9, No. 1, 2024, pp. 187-196
ISSN 2580-6424 (printed) | ISSN 2477-2399 (online)

Implementation of K-Means Clustering in Mapping Teacher Distribution Using Geographic Information System

Khairul Muttaqin[1*], Rif'ah Nurhidayah Nst[1], Novianda[1] , Ahmad Ihsan[1], Jumriani Sultan[2], Fina Rifqiyah[2]

[1]Universitas Samudra, Langsa, Indonesia
[2]Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

Article Info

Article history:
Received August 09, 2024
Revised September 10, 2024
Accepted September 23, 2024

Keywords:
Teachers' distribution; clustering; K-Means; GIS; QGIS

Abstract

The placement of teachers in Indonesia has not been evenly distributed across several regions due to inaccurate recruitment and placement processes. The quality of education, particularly in rural areas, is negatively impacted by this uneven distribution. Teachers play a crucial role in enhancing education, making it essential to address this issue. This study seeks to equilibrate the allocation of teachers in Langsa City using the K-Means Clustering method based on the number of teachers, students, and study groups at the Madrasah Ibtidaiyah, Madrasah Tsanawiyah, and Madrasah Aliyah levels. The clustering results are then mapped using the Quantum Geographic Information System. The study identifies 20 schools with a shortage of teachers, 7 schools with sufficient teachers, and 3 schools with a surplus. The utilization of the K-Means Clustering method demonstrated a high accuracy rate of 92.8%. The implications of these findings suggest that educational authorities can use the clustering results to strategically address teacher shortages by reallocating teaching resources more effectively, thus potentially improving educational outcomes in underserved areas. Moreover, the GIS mapping offers a practical tool for ongoing monitoring and decision-making regarding teacher distribution.

Corresponding Author:
Email: khairulmuttaqin@unsam.ac.id

INTRODUCTION

The proportion of teachers in Indonesia remains unequal, with a significant concentration in urban areas and a shortage in rural regions. Current teacher distribution policy instruments in Indonesia are inadequate to allocate teachers equitably, leading to an overflow in the urban areas and a scarcity in remote areas [1]. This imbalance is accentuated by disparities in teacher quality between rural and urban areas, which contributes significantly to the urban-rural disparity in student academic achievement. According to studies, if rural teachers had the equivalent competence as urban teachers, the student achievement disparity could be decreased [2]. The previous study confirmed that the uneven distribution is partly due to improper recruitment and placement strategies, which fail to account for the unique needs of different regions [3].

Indonesia is an archipelago with thousands of islands, and many regions, especially rural and remote areas, face severe shortages of qualified teachers. Urban areas, on the other hand, might be oversaturated with teachers [4]. Clustering can help to more effectively allocate teachers to areas where they are most needed, ensuring that every region, regardless of its remoteness, has access to quality education. The uneven distribution of teachers, especially in terms of qualifications, has a direct impact on educational outcomes [5]. Inaccurate placement processes often lead to teachers being placed in regions where their skills may not match the local needs or where they are unable to thrive. Clustering can identify regional educational needs and match teachers with appropriate skills and qualifications to those areas, improving overall teaching quality [6]. Ineffective recruitment and placement processes have led to mismatches in the supply and demand for teachers across various regions. Clustering can help streamline these processes by categorizing regions based on specific criteria such as student-teacher ratios, regional needs, infrastructure, and resource availability. This would allow for more informed and data-driven decisions, reducing inefficiencies [7].

Several alternatives have been explored to address this imbalance, such as incentive-based teacher placements, rural teacher training programs, and remote teaching solutions through digital platforms [8][9]. However, these alternatives often face logistical challenges, and funding constraints, or fail to provide a precise assessment of teacher shortages and surpluses at the local level [10]. These limitations highlight the need for a more data-driven approach to teacher distribution, such as Geographic Information Systems (GIS), which allows for detailed mapping and analysis [10][11]. Geographic Information System (GIS), assisted by the K-Means Clustering method for grouping and inputting data into the GIS, provides a more structured and accurate approach [12]. K-Means is a non-hierarchical data clustering approach that splits data into clusters where data with similar features are placed among the same cluster, and data with different features are organized in other clusters. [13]. By using the K-Means method, it is expected that the data can be processed to distinguish regions with a surplus, adequacy, and shortage of teachers, and then mapped using the geographic information system [14].

Other clustering methods, such as Gaussian Mixture Models, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), and hierarchical clustering, could also be considered for this type of analysis. However, hierarchical clustering is computationally intensive and not well-suited for large datasets, while DBSCAN is more effective for data with varying densities but is less intuitive for defining the number of clusters in advance [15]. Gaussian Mixture Models can provide probabilistic cluster assignments but require more complex parameter tuning [16]. K-Means, by comparison, is easy to understand, efficient with computation, and performs effectively with enormous datasets, making it an effective choice for clustering in Geographic Information Systems applications, such as mapping teacher distribution [17].

By clustering regions based on educational needs, infrastructure, and demographics, the Ministry of Education can optimize resources such as training programs, teaching aids, and support systems [6]. Teachers placed in clusters with similar challenges could benefit from tailored training, mentorship programs, and support networks, enhancing their effectiveness. Clustering enables policymakers to gain clearer insights into regional disparities and challenges in teacher distribution [8]. This information is crucial for long-term educational planning, regarding budgeting, recruitment policies, and professional development programs. In conclusion, clustering in teacher placement in Indonesia is an urgent necessity to rectify the current inefficiencies in recruitment and allocation. By leveraging clustering, the government can ensure more strategic, equitable, and effective deployment of teachers across the nation's diverse regions [7][9].

Given these advantages, the K-Means method is applied in this study to cluster data into three categories: surplus, adequate, and shortage of teachers in Langsa City [18]. By using K-Means, this research effectively processes data related to the number of teachers, students, and study groups to highlight regions with differing levels of teacher availability [19]. These clusters are then visualized through GIS mapping, providing a clear geographic representation of teacher distribution across Madrasah Ibtidaiyah, Madrasah Tsanawiyah, and Madrasah Aliyah levels [20]. This approach not only identifies areas with imbalances but also serves as a tool for decision-makers to reallocate resources efficiently [21].

Numerous studies have previously explored the K-Means Clustering method. A previous study examines clustering using the K-Means algorithm based on the quantity of students, teachers, and schools at the high, junior, and elementary school levels. [22]. Another research study covers a web-based geographical information system for the placement of vocational schools and high schools in Surakarta [23]. However, these studies primarily focus on different geographic regions and do not incorporate multiple parameters for analysis at different educational levels. A summary of the previous studies is presented in Table 1.

Table 1. Existing Research

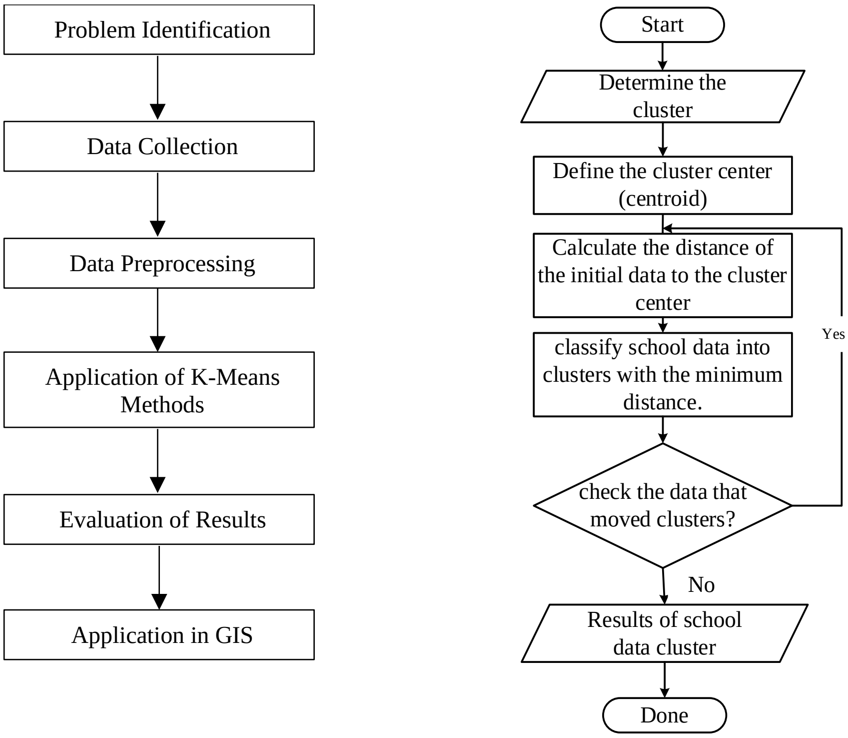
Author	Year	Method	Finding
F. F. Prasetyo [22]	2018	Spatial Clustering K-Means	The outcomes of this study involve generating maps of the Banten Province area determined by the level of education with a shortage, sufficiency, and surplus of teachers related to the district or city.
R. Renaldi et al. [23]	2020	GIS	The result of this research is a web GIS that can display the distribution of schools and school profiles in the city of Surakarta.
A. Aditya et al. [17]	2020	GIS & K-Means	The K-Means algorithm results show 14 provinces in Cluster 1, 5 in Cluster 2, and 15 in Cluster 3. Cluster 1 has a high national test score, cluster 2 has a low national test score, and Cluster 3 has a moderate national exam score.
Mustakim, Z., & Kamal, R [5]	2021	K-Means	The purpose of this research is to categorize Indonesian provinces based on secondary education management quality requirements. This study used eleven attributes to classify the provinces into three primary groups, implementing a quantitative approach defined as K-means cluster analysis.
This study	2024	GIS & K-Means	The implications of these findings suggest that educational authorities can use the clustering results to strategically address teacher shortages by reallocating teaching resources more effectively, thus potentially improving educational outcomes in underserved areas.

Based on previous research related to geographic information systems and the K-Means method, this study will focus on mapping the distribution of teachers in Langsa City. Unlike past research, this study uses the K-Means Clustering method to analyze the distribution of teachers based on data on the number of students, teachers, and study groups in Langsa City. The K-Means method will classify the data into three clusters: C1 (surplus of teachers), C2 (enough teachers), and C3 (shortage of teachers) [17]. This classification aims to understand the distribution of teachers at the MA (Islamic Senior High School), MTS (Islamic Junior High School), and MI (Islamic Elementary School) levels, and the results will be mapped using GIS [18]. This investigation emphasizes the usage of the K-Means method for classification and the application of GIS for mapping the distribution of teachers at different educational levels in Langsa City.

METHODS

In this study, there are several methods used to map teacher needs in schools, including data acquisition or data collection on the distribution of teachers in all schools in Langsa City. The next method is the classification process or grouping of schools based on teacher needs, then mapping using a geographic information system. The research framework is shown in Figure 1(a). The research process begins with the identification of problems, focusing on the uneven distribution of teachers at the MI, MTs, and MA levels in Langsa City. The next stage is data collection, which is aimed at gathering valid data through methods such as researcher observation, interviews, and literature review. Following data collection, system analysis is conducted to determine the system's requirements, including data on the number of schools and teachers, which will be grouped utilising the K-Means Clustering method. After the analysis, system design is carried out to create a structured layout that will ease the system's implementation. The system is then implemented by applying the transformed data using the K-Means method. Finally, system testing is performed to assess the functional quality of the system.

ISSN 2580-6424 (printed) | ISSN 2477-2399 (online)



(a) The Framework (b) K-Means Calculation

Figure 1. The Research Flowchart

Teacher distribution mapping

GIS is designed to work with spatial data sources [3]. GIS is a highly effective tool for presenting remote sensing (RS) data and converting it into information useful for various stakeholders and purposes [4]. Quantum GIS (QGIS) is free and open-source GIS software for processing geospatial data. QGIS is a mapping program that uses spatial databases, meaning it utilizes coordinate data [8]. Mapping is an integrated human-machine system designed to present information to support operational functions, management, and decision-making within an organization through a map [8]. Langsa City originated from the expansion of East Aceh Regency. It is located about 400 kilometres from Banda Aceh, the capital of Aceh Province. Geographically, Langsa City is situated between 04°24'35.68" - 04°33'47.03" North Latitude and 97°53'14.59" - 98°04'42.16" East Longitude [9].

Teachers are educational professionals who are expected not only to teach or transfer knowledge but also to provide guidance, set an example, train students, and contribute to the community while performing various administrative tasks [10]. Clustering is a process aimed at organizing data with commonalities into clusters to ensure that data in one cluster has maximum similarity while data between clusters has the smallest similarity [11]. A flowchart is a sequence of processes in a system that illustrates the input tools, output, and types of media storage used in data processing [12].

K-Means Method

K-Means is a non-hierarchical data clustering approach that divides data into clusters by grouping data with similar qualities together and data with dissimilar features into separate clusters. The method applied is the K-Means method. The process flow of applying this method in the system is illustrated in Figure 1(b).

The procedures to apply the K-Means method start with calculating the number of clusters according to the problem being studied to facilitate grouping data with similar characteristics. Next, determine the centroid of each cluster. Centroid points are selected randomly. Then, measure the distance between each data point and the previously determined centroids. Calculate the distance using the Euclidean Distance Formula (1). Group the objects based on the closest or minimum distance

ISSN 2580-6424 (printed) | ISSN 2477-2399 (online)

between the data points and the centroids. Do the iteration and update the position of the centroids using the equation in Formula (2). If the cluster centroids no longer change, the clustering process is complete. Otherwise, return to the step that calculates the centroid distance and repeat until the centroids remain unchanged.

$$D_{ik} = \sqrt{\sum_{j=1}^{M}(x_{ij} - C_{kj})^2} \quad \text{.....(1)}$$

Which is:

D_{ik} = distance of data point *x* to-i
M = number of variables
x_{ij} = data to be clustered
C_{kj} = centroid of the cluster

$$c_{ij} = \frac{\sum_{i=1}^{p} x_{ij}}{p} \quad \text{..... (2)}$$

Which are:

Cij = centroid of the cluster
Xij = the i-th object
P = number of objects / total number of objects in the cluster

RESULT AND DISCUSSION

Cluster Analysis Results

The K-Means clustering method begins with determining the number of clusters based on the problem at hand, where similar data points are grouped together. Next, the centroids, or the initial central points of each cluster, are selected randomly; once the centroids have been determined, the distance between each data point and the centroids is determined using a distance measure, such as Euclidean distance. The data points are then assigned to the clusters with the nearest centroids, ensuring that data points with the shortest distance to a centroid are grouped together.

After this initial assignment, the centroids of the clusters are updated by recalculating the mean of all the points within each cluster. This is an iterative process, and the distances of the data points and the new centroids are recalculated after each update. This process continues until the centroids stabilize and no longer change positions, indicating that the data points have been clustered optimally. The final step involves evaluating the clustering results to ensure the clusters are meaningful and distinct. The effectiveness of the clustering can be assessed using different metrics, such as the intra-cluster variance, which measures the compactness of clusters, or by visualizing the clusters in a graphical format. Once the clustering is complete, the results are used to draw insights or make decisions based on the problem being addressed.

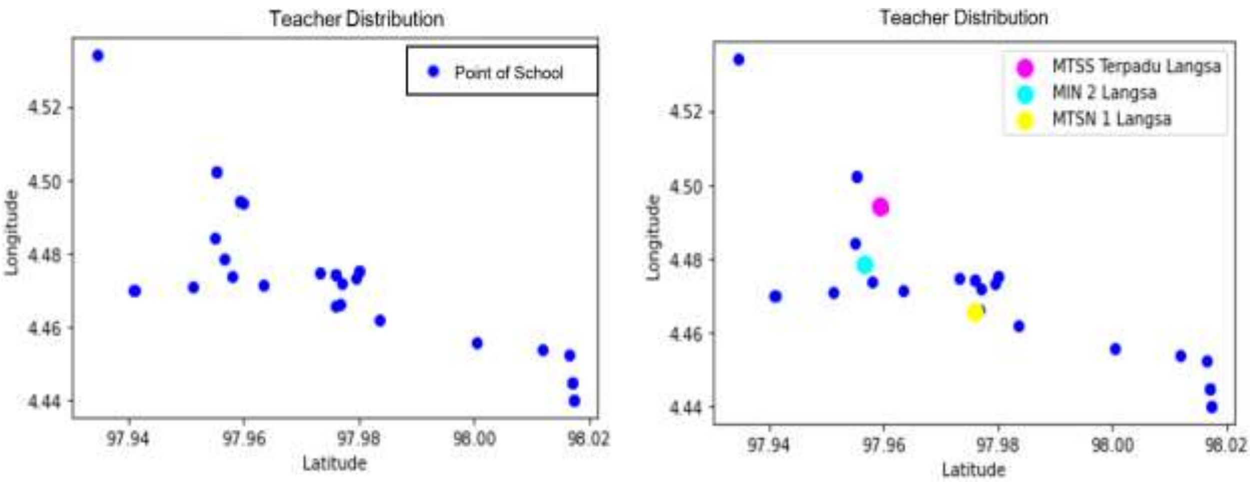
Table 2.The School Data

No.	School	Number of students	Number of teachers	Study groups
1	MIN 1 Langsa	979	44	30
2	MIN 2 Langsa	1119	53	37
3	MTSS Ulumul Qur'An	874	94	29
4	MTSN 1 Langsa	750	46	22
….	….	….	….	….
28	MAS Al Washliyah	20	12	3
29	MAS Gampong Teungoh	46	17	4
30	MAS Kencana Langsa	38	5	1
	Total	10550	829	393

The analysis of K-Means clustering on the school data table is based on parameters with predetermined values, including the number of students, teachers, and study groups at 30 school locations. The usage of the K-Means Clustering method resulted in three clusters representing schools with a surplus of teachers, those with an adequate number, and those experiencing a shortage. The clustering was based on data from 30 schools, including the number of students, teachers, and study groups. The initial centroids were randomly selected, and the clusters were iteratively refined until convergence. The final results showed that 20 schools were in the "shortage" cluster, 7 in the "adequate" cluster, and 3 in the "surplus" cluster (see Table 2). These clusters highlight distinct patterns of teacher distribution across Langsa City.

Teacher Distribution Patterns

A scatter diagram is used to depict the dispersion of teachers. The results reveal that schools classified as having a shortage of teachers are predominantly located in more remote or rural areas of Langsa City, while those with a surplus are in more urbanized locations. The distribution patterns across the MA, MTs, and MI levels suggest a consistent trend of inadequate teacher allocation in rural regions. Figure 2 (a) presents a GIS-based map visualizing these clusters, showing a concentration of teacher shortages in specific geographic zones, which is critical for targeted policy interventions.



(a) (b)

Figure 2. (a) Plot Texaher Distributions Pattern and (b) Plot Centroid

To cluster the data into several clusters, the K-Means clustering method can be utilized. The steps involved are two steps. Firstly, the number of clusters will be determined. In this study, the data will be divided into three clusters. Secondly, determine the initial centroids. The initial centroids are selected randomly and can be seen in Table 3. The cluster centroids and their plotting can be seen in Figure 2 (b). After acquiring the initial centroids, use the Euclidean Distance Formula to find the distance between them (1). The distances obtained are C1, C2, and C3. The detailed calculation is shown in Table 4. The calculated distances will be used to select the closest distance between the data and the cluster centroids. This distance indicates that the data point belongs to the cluster with the nearest centroid. By comparing the cluster results, the smallest value is chosen. From Table 4, the results of the first iteration are calculated, and the data is grouped into three clusters: teacher shortage, sufficient teachers, and teacher surplus. Table 5 displays the outcomes of this grouping.

Table 3. Initial Centroid

Centroid	School	Number of Students	Number of Teachers	Groups Study
K1	MIN 2 LANGSA	1119	53	37
K2	MTSN 1 LANGSA	750	46	22
K3	MTSS TERPADU LANGSA	497	34	15

Table 4. First Iteration Calculation Results

No	C1	C2	C3	Nearest Distance	Note
1	140,4635184	229,1484235	482,3370191	140,4635184	1
2	0	369,3710871	622,67889	0	1
3	754,7403527	385,5152396	132,1892583	132,1892583	3
4	553,6876376	184,5020325	69,07242576	69,07242576	3
5	752,0146275	382,8224131	129,5028957	129,5028957	3
….	….	….	….	….	….
28	1100,289962	731,0383027	477,6578273	477,6578273	3
29	1074,110795	704,8269291	451,4543166	451,4543166	3
30	1082,663844	713,4886124	460,128243	460,128243	3

Table 5. Iteration Grouping

Cluster	Numbers	The schools
C1	2	MIN 1 Langsa, MIN 2 Langsa
C2	3	MIS Al-Ashriyah, MIS Gampong Mutia, MI IT Nurshadrina
C3	25	MTSS Ulumul Qur'an, MIS Terpadu Langsa, MTSN 1 Langsa, MIN 4 Langsa, MAS Ulumul Qur'an, MTSS Terpadu Langsa, MAN 2 Langsa, MIN 3 Langsa, MIS Terpadu T. Al Mubarak, MIS Paya Bujok Tunong, MTSS Bustanul Huda, MTSS Raudhatun Najah, MIN 5 Langsa, MTSS Darul Huda, MTSS Al Washliyah, MTSS Geudubang Aceh, MTSS MIM Langsa, MTSS Timbang Langsa, MTSS Yapila Langsa, MAN 1 Langsa, MAS Raudhatun Najah, MAS Darul Huda, MAS Al Washliyah, MAS Gampong Teungoh, MAS Kencana Langsa

The results of the ratio values from Iteration 1 are as follows: *BCV*/*WCV* = 1245,778178 / 2421736 = 0,000514415, where *BCV* = the distance of each specific centroid and *WCV* = square of the closest distance. The last step is to recalculate the distance of the data to the new centroids using the Euclidean Distance formula until the values no longer change. After performing 4 iterations, the centroid values no longer change. Based on the calculations using the K-Means method, the data clustering results in three clusters for teacher distribution are shown in Table 6.

Table 6. The Final School Clusters

Cluster	**Definition**	**The schools**
Cluster 1	Schools with a surplus of teachers (3 schools)	MIN 1 Langsa, MIN 2 Langsa, and MTSS Ulumul Qur'an Langsa
Cluster 2	Schools with a sufficient number of teachers (7 schools)	MIS Terpadu Langsa, MTSN 1 Langsa, MIS Al-Ashriyah, MIN 4 Langsa, MAS Ulumul Qur'an Langsa, MTSS Terpadu Langsa, and MAN 2 Langsa.
Cluster 3	Schools with a shortage of teachers (20 schools)	MIN 3 Langsa, MIS Gampong Mutia, MIS Terpadu T. Al Mubarak, MIS Paya Bujok Tunong, MIS IT Nurshardina, MTSS Bustanul Huda, MTSS Raudhatun Najah, MIN 5 Langsa, MTSS Darul Huda, MTSS Al-Washliyah, MTSS Geudubang Aceh, MTSS MIM Langsa, MTSS Timbang Langsa, MTSS Yapila Langsa, MAN 1 Langsa, MAS Raudhatun Najah, MAS Darul Huda, MAS Al-Washliyah, MAS Gampong Teungoh, MAS Kencana Langsa

ISSN 2580-6424 (printed) | ISSN 2477-2399 (online)

To effectively follow up on recommendations related to clustering insufficient and excess teachers, a structured mechanism is needed, with specific roles for various stakeholders. Government and Policymakers as the central authority responsible for overseeing teacher placement and ensuring equitable distribution based on data-driven insights from clustering. Regional Education Departments take the bodies to manage the on-the-ground implementation of clustering recommendations by collaborating with local schools and adjusting teacher placements accordingly. School principals and local administrators are responsible for providing accurate data on teacher needs, such as student-teacher ratios, and working closely with the education departments to ensure smooth implementation of recommendations. Teachers' unions should advocate for fair and transparent processes when teachers are relocated. They ensure that teachers' rights are respected, especially when incentives or compulsory relocations are involved. Independent evaluation bodies or research institutions can monitor the long-term effects of clustering recommendations to ensure that they lead to equitable and improved educational outcomes.

By integrating these roles and mechanisms, the follow-up on clustering recommendations can be executed effectively, ensuring that both teacher shortages and excesses are addressed in a strategic and equitable manner. There is a need to explore effective strategies for addressing teacher shortages, particularly in rural or remote regions. Research could focus on the impact of incentive-based policies, teacher relocation programs, and innovative classroom management solutions in shortage areas. In conclusion, schools must engage in proactive measures to address teacher surpluses and shortages, with ongoing research playing a key role in optimizing these solutions.

Implications for Policy and Practice

After the design is completed, it is implemented into the WebGIS Teacher Distribution system, allowing users to visualize the clustering results. Below is the WebGIS display result, which shows the distribution of teachers across different regions in Langsa City based on the clustering analysis. This implementation is not only useful for visualization purposes but also has significant implications for educational policy and practice. By visually mapping teacher shortages, sufficiencies, and surpluses, educational authorities can make more informed decisions regarding teacher placement, recruitment, and resource allocation. For instance, targeted interventions can be designed for areas identified as having a shortage of teachers, while areas with a surplus can be considered for reallocation or sharing of teaching resources.

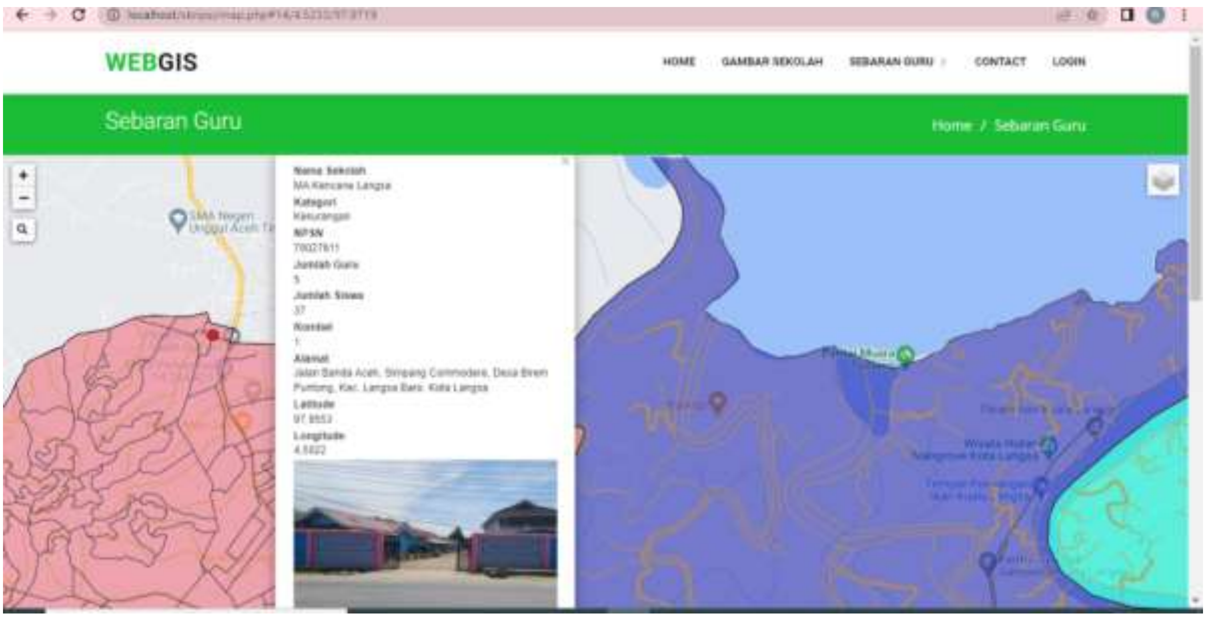


Figure 3. Teacher Distribution Map

Figure 3 shows an example of the display on WebGIS. The WebGIS display consists of a Home page, a Teacher distribution map page, a Login page, and a K-Means Clustering results page. The home page displays the menu options including home, school images, teacher distribution, contact, and login. This is the initial view when the system is launched. On the Teacher Distribution Map Page, the map displays the distribution of teachers across all education levels, including MA, MTS, and MI. Login Page allows the admin to access the WebGIS teacher distribution system. On the K-Means Clustering Results Page, the admin inputs all the data to be processed for analyzing teacher distribution using the K-Means clustering method.

The results of this study, using the K-Means clustering method to identify teacher distribution patterns in Langsa City, are in line with previous research that has utilized clustering algorithms to analyze resource distribution in educational contexts. For example, this study [5] compared the K-Means and K-Medoids clustering algorithms to analyze the distribution of high school teachers in Indonesia. Their findings indicated that K-Medoids performed better than K-Means in scenarios with fewer clusters due to the impact of outliers, whereas K-Means was more effective in scenarios with larger numbers of clusters. This suggests that the choice of clustering algorithm can significantly affect the results, depending on the nature of the dataset and the clustering scenarios.

However, the previous study focused on a comparative analysis of clustering algorithms for high school teacher distribution [5]; this study extends the analysis to all educational levels (MI, MTs, MA) and incorporates additional variables such as the number of students and study groups. This broader approach allows for a more detailed understanding of the distribution needs across different educational contexts. Additionally, the use of GIS in this study provides a visual mapping of these distribution patterns, which can further enhance decision-making processes for educational authorities.

The findings from both studies highlight the importance of selecting appropriate clustering algorithms and considering multiple variables when analyzing resource distribution in educational settings. While the K-Medoids algorithm may be more suitable for datasets with significant outliers, K-Means can be highly effective for larger datasets with well-defined clusters. The implication for policy is that a tailored approach, which considers the specific characteristics of the dataset and the educational context, can lead to more effective resource allocation and management strategies.

CONCLUSION

According to the research, the application of the K-Means approach in WebGIS leads to numerous findings. The K-Means clustering method, applied to teacher distribution based on the number of students, teachers, and study groups, successfully achieved clustering into teacher shortage, sufficient teachers, and teacher surplus for MI, MTs, and MA levels within the Ministry of Religious Affairs. The teacher distribution map can be viewed on the GIS (Geographic Information System) based on 3 clusters: C1 (Teacher Shortage) with 20 schools, C2 (Sufficient Teachers) with 7 schools, and C3 (Teacher Surplus) with 3 schools. The accuracy testing of the K-Means clustering method, based on 30 data points from 2022, yielded an accuracy of 92.8%. This high level of accuracy indicates the reliability and effectiveness of the method for clustering analysis in educational contexts.

The clustering results have significant implications for improving teacher allocation policies in Langsa City. By identifying schools with teacher shortages, sufficiency, and surpluses, education authorities can better target interventions and optimize resource distribution. For future research, it is recommended to add more analysis parameters for mapping teacher distribution to make the results more informative. Additionally, data synchronization between schools should be ensured for further research development.

ACKNOWLEDGMENT

The author expresses gratitude to Allah SWT for His blessings and grace, which enabled the completion of this research. The author also delivers big thanks to all family, friends, and lecturers for their support and assistance.

REFERENCES

[1] M. Novita, C. Wijaya, and A. Atmoko, “Dynamic of Indonesian Teacher Distribution Policies Implementation at Regional Level,” Journal of Education & Social Policy, vol. 7, no. 1, pp. 15–24, 2020.
[2] J. Zhang, S. Jin, M. Torero, and T. Li, “Teachers and Urban-Rural Gaps in Educational Outcomes,” American Journal of Agricultural Economics, vol. 100, no. 4, pp. 1207–1223, 2018.
[3] E. Purwanto and Y. Hayati, “Impact of teacher distribution on student academic performance in rural Indonesia,” J. Educ. Dev., vol. 6, no. 2, pp. 100–110, 2018.
[4] Drake, B. C. (1997). Geographic Dimensions of Indonesia’s Increasing Importance in the World. 19–26.
[5] Mustakim, Z., & Kamal, R. (2021). K-Means Clustering for Classifying the Quality Management of Secondary Education in Indonesia. Cakrawala Pendidikan, 40(3), 725–737. https://doi.org/10.21831/cp.v40i3.40150.
[5] T. Widiyaningtyas, U. Pujianto, and M. Prabowo, “K-Medoids and K-Means Clustering in High School Teacher Distribution,” 2019 Int. Conf. Electr. Electron. Inf. Eng. (ICEEIE), pp. 330–335, 2019, doi: 10.1109/ICEEIE47180.2019.8981466.
[6] Chen, Y., Tan, P., Li, M., Yin, H., & Tang, R. (2024). K-means clustering method based on nearest-neighbor density matrix for customer electricity behavior analysis. International Journal of Electrical Power and Energy Systems, 161(July). https://doi.org/10.1016/j.ijepes.2024.110165.
[7] Li, Y., & Zhang, H. (2024). Big data technology for teaching quality monitoring and improvement in higher education - joint K-means clustering algorithm and Apriori algorithm. Systems and Soft Computing, 6(July), 200125. https://doi.org/10.1016/j.sasc.2024.200125.
[8] S. N. Suryahadi and A. Sambodho, “Incentive-based teacher placements in remote Indonesia: A case study,” Southeast Asian Studies, vol. 54, no. 1, pp. 45–56, 2018.
[9] B. Susetyo, D. R. Kurniawan, and E. Hermawan, “Geospatial mapping and K-Means clustering for teacher shortage analysis in Indonesian schools,” Int. J. Geogr. Inf. Sci., vol. 33, no. 9, pp. 2001–2013, 2020.
[10] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, “Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan,” J. Nas. Teknol. dan Sist. Inf., vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
[11] Y. A. Priambodo and S. Y. J. Prasetyo, “Pemetaan Penyebaran Guru di Provinsi Banten dengan Menggunakan Metode Spatial Clustering K-Means (Studi kasus: Wilayah Provinsi Banten),” Indones. J. Comput. Model., vol. 1, no. 1, pp. 18–27, 2018.
[12] M. Ramadhan, D. D. Priyatna, and A. Hidayat, “Analysis of teacher placement efficiency using Geographic Information Systems (GIS),” J. Teknol. Inf., vol. 13, no. 4, pp. 123–130, 2020.
[13] M. Utami and H. Sutopo, “K-Means clustering implementation in analyzing teacher distribution data in Semarang,” 2019 Int. Conf. Sci. Technol., pp. 105–112, 2019.
[14] D. Susanto, “Application of GIS in Education Data Mapping,” Int. Conf. Comput. Eng. Inf. Technol., pp. 115–123, 2020.
[15] H. Rahmawati, F. A. Putri, and N. A. Azizah, “Optimizing teacher allocation using GIS-based DBSCAN clustering,” Int. J. Inf. Technol. Syst. Eng., vol. 6, no. 2, pp. 80–95, 2021.
[16] A. S. Fauzi and R. Wahyudin, “Gaussian Mixture Models in Clustering Teachers for Effective Resource Distribution,” J. Comput. Educ., vol. 5, no. 3, pp. 105–115, 2020.
[17] A. Aditya, I. Jovian, and B. N. Sari, “Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019,” J. Media Inform. Budidarma, vol. 4, no. 1, p. 51, 2020.
[18] K. L. BPS, Ed., Kota Langsa Dalam Angka 2021. Kota Langsa: BPS-Statistics of Langsa Municipality, 2021.
[19] E. Priyanto and H. Setiawan, “Mapping teacher distribution using K-Means clustering and GIS in West Java,” 2020 Int. Conf. Comput. Eng. Appl., pp. 145–150, 2020.
[20] M. Wicaksono and A. F. Nurcahya, “Clustering analysis of teacher distribution using GIS,” Int. Conf. Inf. Technol. Electr. Eng., pp. 85–90, 2019.
[21] F. Arifin and D. Purnamasari, “GIS-based teacher distribution modeling in East Java,” J. Teknol. Inf. dan Rekayasa, vol. 4, no. 3, pp. 215–225, 2021.
[22] F. F. Prasetyo, “Mapping the distribution of teachers in Banten Province using K-Means clustering,” Int. Conf. Adv. Comput. Sci., vol. 4, pp. 77–84, 2018.
[23] R. Renaldi and D. A. Anggoro, “Sistem Informasi Geografis Pemetaan Sekolah Menengah Atas/Sederajat di Kota Surakarta menggunakan Leaflet Javascript Library berbasis Website,” Emit. J. Tek. Elektro, vol. 20, no. 2, pp. 109–116, 2020.