Jurnal Sistem Komputer dan Informatika (JSON)
Volume 6, Nomor 4, Juni 2025
e-ISSN 2685-998X
DOI 10.30865/json.v6i4

A Comparative Study of K-Means and K-Medoids for Clustering Dengue Fever Risk Areas in Medan

Anisa Amelia Fitri[1*], Munirul Ula[2], Cut Agusniar[3]

Department of Informatics, Universitas Malikussaleh, Lhokseumawe, Indonesia
Email: [1]Anisa.210170205@mhs.unimal.ac.id, [2]Munirulula@unimal.ac.id, [3]Cutagusniar@unimal.ac.id
Email Penulis Korespondensi: Anisa.210170205@mhs.unimal.ac.id *
Submitted: **02/06/2025**; Accepted: **13/06/2025**; Published: **30/06/2025**

Abstract−Dengue Hemorrhagic Fever (DHF) is a localized disease that continues to contribute to a high number of cases in Medan City. The local health authority faces challenges in identifying priority areas for effective prevention and control. This study applies data clustering techniques to map DHF risk areas by comparing the performance of K-Means and K-Medoids algorithms. The optimal number of clusters was determined using the Silhouette Coefficient, while the clustering quality was assessed using the Davies-Bouldin Index (DBI). The findings indicate that K-Means performs best with four clusters and achieves a lower DBI value compared to K-Medoids. Based on this, the study recommends using K-Means to categorize DHF risk areas into four priority levels: high, medium, low, and very low. This approach is expected to support the Medan City Health Office in implementing more targeted and efficient DHF control strategies.

Keywords: Dengue Hemorrhagic Fever; K-Means; K-Medoids; Clustering; Silhouette Coefficient; Davies-Bouldin Index

1. INTRODUCTION

Dengue Hemorrhagic Fever (DHF) is a contagious disease due to infection with the dengue virus from part of the Flavivirus family and transmitted through Aedes aegypti mosquitoes. This disease primarily affects children and can cause severe symptoms such as high fever, bleeding, and even risk of death. In Indonesia, DHF remains a serious public health issue, with the number of cases increasing yearly [1]. The DHF cases in Medan City have shown a significant upward trend in recent years. In 2023, Medan recorded 965 DHF cases out of 4,687 cases in North Sumatra Province, making it the largest contributor in the region. This trend continued into 2024, with January to July, 558 DHF cases were reported in Medan. Specifically, in January 2024 alone, North Sumatra reported 405 cases with Medan as the largest contributor. These data indicate a high rate of DHF transmission in Medan City and highlight the need for more focused prevention and control efforts [2].

The escalation of DHF cases results from numerous influencing variables, including the large vector population, virus virulence, community immunity, demographics, population density, patient mobility, virus replication capability in mosquitoes, and human behavior[3]. The interaction among the host, vector, and environmental conditions are key factors in the spread and intensification of DHF, where environmental factors also affect the overall public health level [4]. Factors such as educational background, sanitation conditions, understanding of symptoms, and public perception of dengue fever (DF) also have a significant influence on the occurrence of DF cases[5]. One key epidemiological management strategy for DHF is mapping the treatment areas, which involves identifying high-risk zones requiring intensive intervention. However, a main challenge in controlling DHF in Medan City is the lack of structured information regarding the distribution of priority treatment areas. Available data mostly consist of case reports without in-depth analysis to identify patterns or clusters. This limitation hinders authorities from determining strategic measures for effective control and prevention.

To address this issue, data mining techniques, specifically clustering methods, can be employed. The process of uncovering hidden trends and valuable insights from large datasets is known as data mining. One of the primary goals of data mining is to detect patterns or trends that support future prediction and decision-making[6]. This technique classifies data points into clusters based on shared characteristics or similarities, similar observations are grouped into one cluster, while distinct data points are placed into separate clusters [7]. Two popular clustering techniques are K-Means and K-Medoids. K-Means divides data into k groups through the computation of the average (mean) of the data points to determine the cluster center, whereas K-Medoids identifies the data point nearest to the cluster center known as the medoid designated as the cluster centroid [8].

Several studies have examined the application of K-Means and K-Medoids algorithms. Syamfithriani et al. [9] evaluated these algorithms for mapping treatment areas of diarrhea cases in Kuningan Regency. Their analysis using Silhouette and Elbow methods showed that K-Means outperformed K-Medoids. Kamila et al. [10] applied K-Means and K-Medoids for clustering shipping transaction data in Riau Province. They found that K-Means was faster, taking about 1 second on average, compared to K-Medoids' 1 minute 38 seconds. Additionally, the Davies-Bouldin Index (DBI) values indicated better performance of K-Means due to lower DBI scores.

Most previous studies determined the optimal number of clusters mainly through the DBI method during evaluation, while some used Silhouette Coefficient or Elbow methods. Therefore, this study utilizes both Silhouette Coefficient and Elbow methods to determine the optimal cluster number. Using both methods is expected to enhance clustering accuracy for K-Means and K-Medoids algorithms, thereby producing a more precise mapping scheme for prioritizing areas in efforts to prevent and control dengue hemorrhagic fever.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 6, Nomor 4, Juni 2025
e-ISSN 2685-998X
DOI 10.30865/json.v6i4

Based on this analysis, this research aims to compare the K-Means and K-Medoids algorithms for clustering DHF cases in Medan City. The expected contributions include providing a more accurate spatial clustering model for DHF cases and assisting health authorities in resource allocation and intervention prioritization.[11]

2. RESEARCH METHODOLOGY

Conducting research involves several important stages that must be followed systematically to obtain valid and reliable results. Here is each stage of the research process.

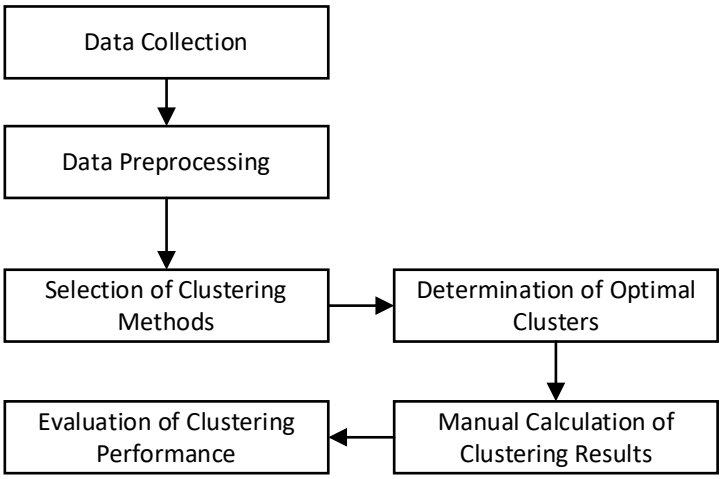

Figure 1. Research Stage

2.1 Research Stages

This research consists of several systematic stages designed to produce accurate clustering and evaluation results for mapping dengue fever handling areas in Medan City. The stages are explained as follows.

a. Data Collection: The main dataset for this research comprises figures on dengue incidence, population density, the size of each area, and available healthcare facilities within 21 sub-districts of Medan City.
b. Data Preprocessing: This involves organizing and standardizing the data so that it can be effectively processed by clustering algorithms.
c. Selection of Clustering Methods: K-Means and K-Medoids clustering techniques were chosen for comparison due to their popularity and reliability in partitioning datasets.
d. Determination of Optimal Clusters: To evaluate the ideal number of clusters, the study applied a metric called the Silhouette Score, which is used to assess how appropriately each data point belongs within Its designated cluster relative to other clusters.
e. Manual Calculation of Clustering Results: Each algorithm was manually implemented using the selected number of clusters to cluster the data.
f. Evaluation of Clustering Performance: The study employed the Davies-Bouldin Index (DBI) to ascertain which analytical method produced clusters of higher quality.

2.2 Clustering Method and Evaluation Techniques

There are two main clustering approaches:

a. Hierarchical Clustering forms a hierarchical structure (dendrogram) by grouping similar data points closer together and dissimilar ones further apart. Techniques include Average Linkage, Complete Linkage, and Single Linkage [12].
b. Non-Hierarchical Clustering starts by specifying the given number of clusters ahead of time. Methods under this approach include K-Means, Fuzzy K-Means, and Mixture Modeling [12].

In this research, a comparison was made between two cluster analysis method, K-Means and K-Medoids, which were applied to group sub-districts in Medan according to four specific attributes: dengue case counts, population density, area size, and the number of healthcare facilities.

2.3 Cluster Number Selection Using Silhouette Score

The Silhouette Coefficient serves as a metric to assess clustering quality by comparing how closely related a data point with respect to its allocated cluster compared to a distinct cluster. This measure aids in selecting the optimal cluster count prior to executing clustering algorithms. A higher Silhouette Coefficient value signifies that data points are appropriately grouped within their own cluster and distinctly separated from other clusters.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 6, Nomor 4, Juni 2025
e-ISSN 2685-998X
DOI 10.30865/json.v6i4

After estimating the appropriate number of clusters, the K-Means algorithm organizes data by calculating the average value (centroid) for each cluster. On the other hand, the K-Medoids technique picks actual observations (medoids) to serve as cluster centers, resulting greater resistance to the influence of anomalies [13].

2.4 K-Means Clustering Model

According to [9], the K-Means clustering algorithm involves a repeated sequence aimed at Dividing data into k specified groups. The procedure is outlined as follows:

a. Cluster Initialization: At the beginning, k, representing the number of clusters, is defined
b. Initial Centroid Selection: Select k initial centroids randomly. These centroids act as the initial reference points for assigning data to clusters.
c. Distance Calculation: Estimate the Euclidean distance from each data instance to central point found through the distance formula (1).
d. $d(x_i,c_k) = \sqrt{\sum_{j=i}^{n}(x_{ij} - c_{kj})^2}$ (1)
e. Cluster Assignment: Assign data points to the cluster whose centroid is nearest, determined by the least distance value.
f. Update the centroid through averaging position of all points part of the cluster
g. keep performing distance computations, segment assignments, and centroid recalculations until no changes occur in cluster memberships, signaling that the process has converged and the clusters are finalized.

2.5 K-Medoids Clustering Model

The K-Medoids clustering algorithm, similar to K-Means, is a partitioning technique that aims to group data into k clusters. However, unlike K-Means which uses the average of data points as the cluster center (centroid), K-Medoids selects actual data points, called medoids, as cluster centers. The K-Medoids algorithm begins by selecting an initial set of k medoids randomly from the dataset[14]. Each data object is then assigned to the nearest medoid based on a chosen distance metric, such as Euclidean Distance or Manhattan Distance [15]. The algorithm proceeds through the following steps:

a. Initialization: Select k data points randomly as initial medoids.
b. Step of Assigning Data Points: Allocate each data point to the cluster whose medoid is closest, using a distance metric such as Euclidean distance described in equation (1).
c. Update Step: For each cluster, evaluate the total distance between each point in the cluster and all other points. The point that minimizes the total distance is chosen as the new medoid.

Repeat the steps of assigning data points and updating medoids until there is no change in medoids or a predetermined iteration limit is reached.

This approach is generally more resistant to noise and outliers than K-Means because it reduces the total dissimilarity between each data point and its representative medoid, instead of relying on squared Euclidean distances.

2.6 Cluster Evaluation using Davies-Bouldin Index (DBI)

To assess the clustering effectiveness of the K-Means and K-Medoids algorithms, the Davies-Bouldin Index (DBI) is used. DBI is an internal metric that evaluates clustering quality by calculating the average similarity between each cluster and its most comparable counterpart. A smaller DBI score signifies superior clustering, indicating that clusters are more compact and distinctly separated. [16]

The DBI for k clusters is computed using the following formula (2) :

$$DBI = \frac{1}{K}\sum_{i=1}^{k} max_{i \neq j}(R_{i,j}) \quad (2)$$

3. RESULT AND DISCUSSION

3.1 Dataset

The data applied in the K-Means clustering experiment includes 21 sub-districts, reporting a total of 1102 dengue fever cases during the year 2024. Detailed information about the dataset is shown in the following table 1.

Table 1. Dataset

No	Kecamatan	Jumlah Kasus	Kepadatan Penduduk	Luas Wilayah (Km2)	Jumlah Fasilitas Kesehatan
1	Medan Tuntungan	119	100132	25,25	20
2	Medan Johor	86	154868	16,79	19
…	…	…	…	…	…
20	Medan Marelan	36	189469	30,16	14
21	Medan Belawan	7	110238	33,41	16

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 6, Nomor 4, Juni 2025
e-ISSN 2685-998X
DOI 10.30865/json.v6i4

3.2 Normalization Data

After preparing the dataset, the next step is normalization, which adjusts numerical attributes to a common, smaller scale. This step is crucial to make sure that no single variable disproportionately influences the clustering outcome, especially those with larger value ranges. Normalization helps balance the impact of all features, making the data more appropriate for distance-based clustering methods like K-Means and K-Medoids. The results of this normalization for each attribute are presented in Table 2.

Table 2. Normalization Data

Kecamatan	Jumlah Kasus	Kepadatan Penduduk	Luas wilayah (km2)	Jumlah Fasilitas Kesehatan
Medan Tuntungan	1	0,411058681	0,690158336	0,5
Medan Johor	0,705357143	0,76294101	0,427506985	0,428571429
…	…	…	…	…
Medan Marelan	0,258928571	0,985381094	0,842595467	0,071428571
Medan Belawan	0	0,476027309	0,943495809	0,214285714

3.3 Identify Optimal Cluster

Following data preparation, the modeling stage begins with applying K-Means and K-Medoids clustering algorithms. It is important first to identify the optimal cluster count to ensure accurate data segmentation. The Silhouette technique is used to evaluate this, measuring the similarity of each point to its assigned cluster versus others. Higher Silhouette values denote better cluster separation. Visualization of these scores, shown in Figures 1 and 2, guides the selection of the cluster number used in subsequent clustering steps.

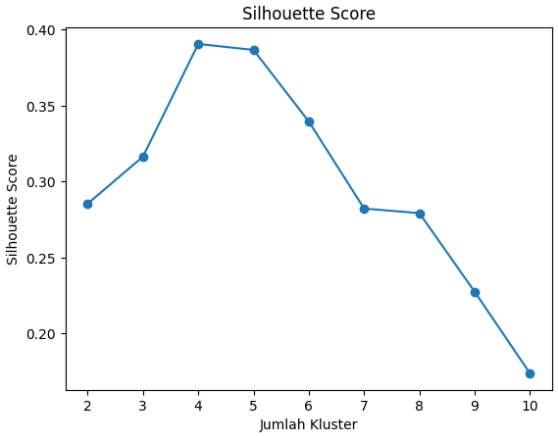

Figure 2. Silhouette K-Means

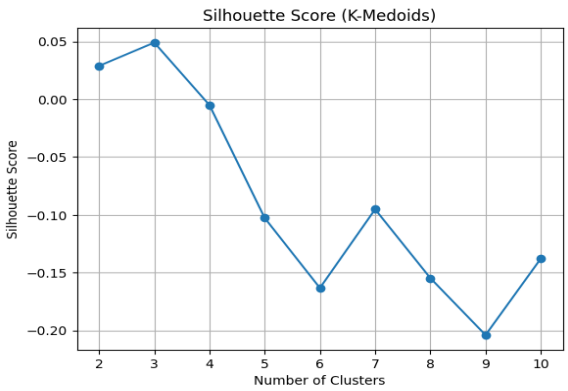

Figure 3. Silhouette K-Medoids

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 6, Nomor 4, Juni 2025
e-ISSN 2685-998X
DOI 10.30865/json.v6i4

From the figures 1 and 2, it was determined that the best number of clusters for the K-Means algorithm is 4, whereas for the K-Medoids algorithm, it is 3. Using these optimal cluster counts, clustering was carried out with both algorithms. The results of the K-Means clustering are visualized in Figures 3 and 4, while the outcomes of the K-Medoids clustering performed with Python are illustrated in Figures 5 and 6.

3.4 K-Means Clustering Results

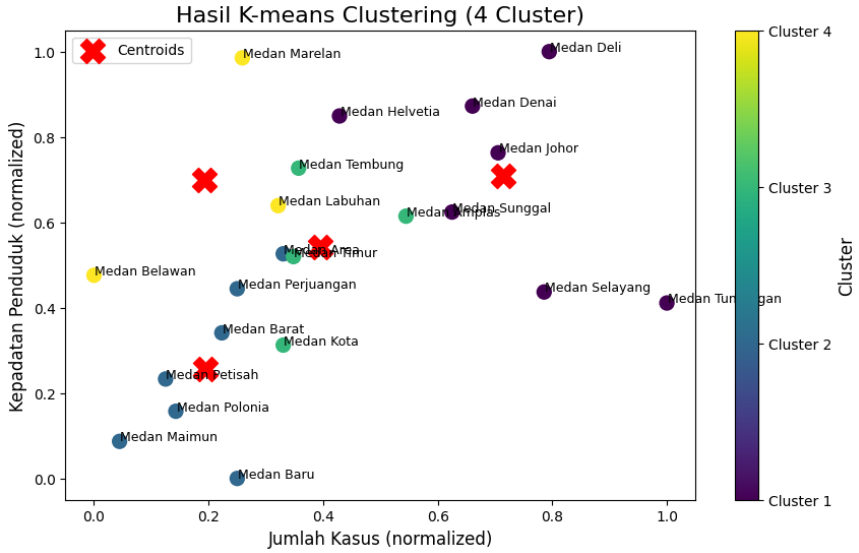

Figure 4. Plot Clustering 2D K-Means

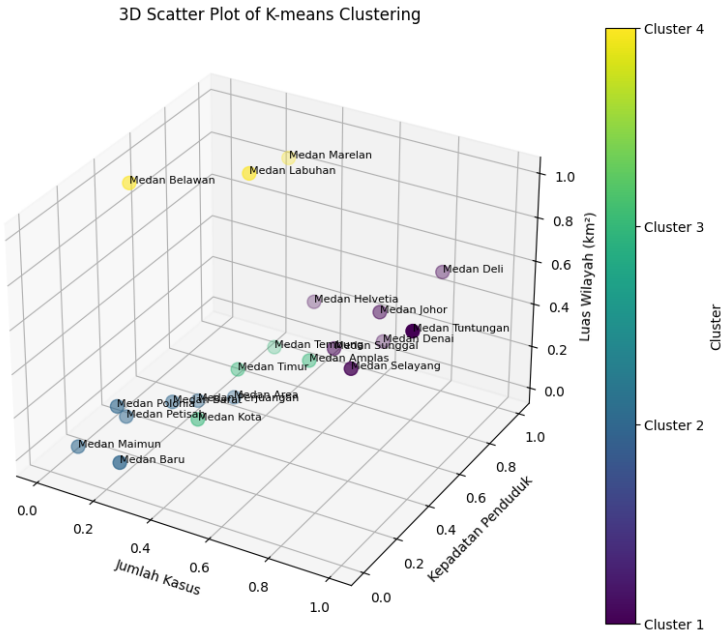

Figure 5. Plot Clustering 3D K-Means

Figures 3 and 4 illustrate the results of the K-Means clustering based on population density and dengue case counts, effectively classifying the 21 sub-districts of Medan City into four priority levels for dengue prevention and control. Each cluster is represented with a different color: purple for Cluster 1 (high priority), blue for Cluster 2 (medium priority), green for Cluster 3 (low priority), and yellow for Cluster 4 (very low priority). The red crosses (x) indicate the centroids of each cluster, serving as the average reference points for the grouped data. Cluster 1 (purple), representing the highest priority areas, consists of sub-districts with both high population density and a high number of dengue cases, highlighting the need for urgent intervention. In contrast, Cluster 4 (yellow) includes sub-districts with lower densities and fewer dengue cases, thus requiring minimal intervention. The clear

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 6, Nomor 4, Juni 2025
e-ISSN 2685-998X
DOI 10.30865/json.v6i4

separation of clusters illustrates how K-Means successfully grouped areas with similar risk levels, providing valuable insights for targeted resource allocation and preventive strategies.
By applying the K-Means algorithm with the selected optimal cluster number of 4, four priority zones were established: high, medium, low, and very low priority areas. These zones are detailed in Table 3.

Table 3. K-Means Results

No	Cluster	Jumlah	Persentase
1	(C1)	7	33,3%
2	(C2)	7	33.3%
3	(C3)	4	19,1%
4	(C4)	3	14,3%
Total		21	100%

Using the K-Means clustering approach, the dataset was divided into four distinct groups. Cluster 1 includes 7 sub-districts, Cluster 2 also has 7 sub-districts, Cluster 3 comprises 4 sub-districts, and Cluster 4 consists of 3 sub-districts. This grouping illustrates differences in dengue case numbers, population density, land area, and the number of healthcare facilities among the 21 sub-districts in Medan City.

3.5 K-Medoids Clustering Results

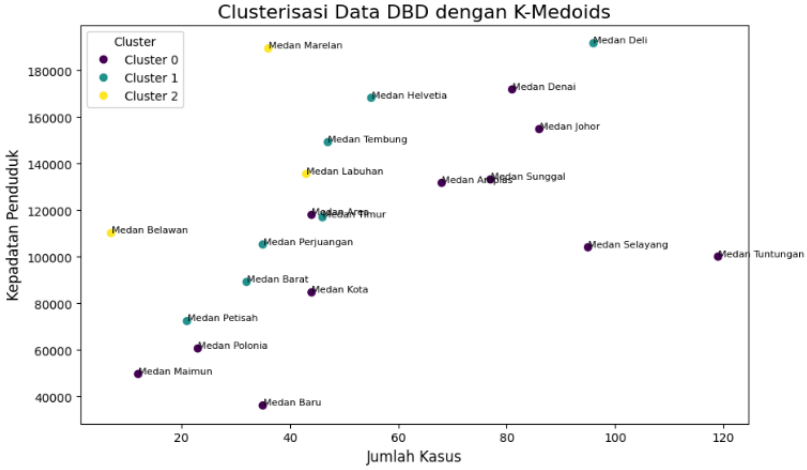

Figure 6. Plot Clustering 2D K-Medoids

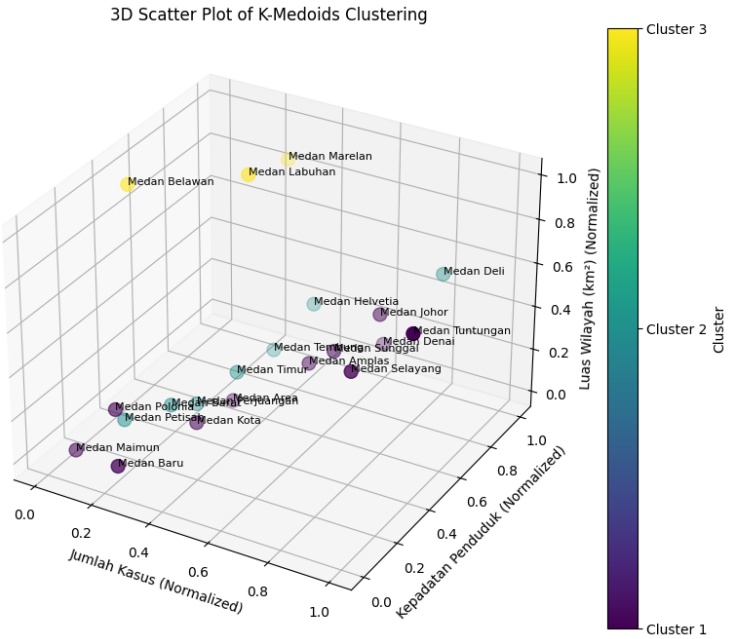

Figure 7. Plot Clustering 3D K-Medoids

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 6, Nomor 4, Juni 2025
e-ISSN 2685-998X
DOI 10.30865/json.v6i4

Figures 5 and 6 present the results of the K-Medoids clustering applied to population density and dengue case counts, dividing the 21 sub-districts of Medan City into three distinct clusters. Each cluster is color-coded: purple for Cluster 1 (high priority), green for Cluster 2 (medium priority), and yellow for Cluster 3 (low priority). The medoids—marked by red crosses (×)—are actual data points that serve as the center of each cluster, making the clustering more robust to outliers compared to K-Means. Cluster 1 (purple) groups sub-districts with the highest dengue case counts and population densities, indicating zones that require the most immediate and intensive attention. Cluster 2 (green) includes areas with moderate risk, suggesting a balanced but continuous monitoring strategy. Cluster 3 (yellow), with the lowest density and case levels, represents areas with the least urgency for intervention. The clustering pattern demonstrates how K-Medoids effectively identifies central representative areas, offering a practical alternative for public health decision-making in regions with potential data anomalies.
By applying the K-Medoids algorithm with the optimal cluster count of three, the data was categorized into three priority zones: high, medium, and low priority areas. Each cluster reflects a specific urgency level for managing dengue fever cases, determined by factors such as case numbers, population density, area size, and healthcare facility availability. These clustering results provide useful guidance for focused public health efforts and efficient distribution of resources.

Table 4. K-Medoids Results

No	Cluster	Jumlah	Persentase
1	(C1)	8	38%
2	(C2)	10	48%
3	(C3)	3	14%
Total		21	100%

Cluster 1, consisting of 8 sub-districts, is classified as a high-priority area due to its high number of dengue fever cases. These locations demand urgent and concentrated intervention efforts to effectively control and reduce disease transmission. Given the critical situation, health officials are recommended to prioritize resource allocation and immediate response activities in these areas.
Cluster 2 comprises 10 sub-districts categorized under medium priority. While dengue cases are present, their prevalence is lower compared to Cluster 1, indicating a need for continued attention but with less urgency.
Cluster 3 includes areas with minimal dengue cases and is designated as low priority. Although large-scale interventions may not be necessary here, ongoing surveillance and preventive actions remain important to avoid potential outbreaks.
To evaluate which algorithm produced the best clustering results, the Davies-Bouldin Index (DBI) was utilized. This metric assesses clustering quality by examining how compact clusters are internally and how well they are separated from each other. A smaller DBI score reflects superior clustering performance, implying distinct and cohesive groupings.

3.6 DBI Score

The optimal clustering method is identified as the one with the lowest DBI value, representing the most precise and effective data partitioning. The comparative DBI scores for both K-Means and K-Medoids algorithms are summarized in Table 5.

Table 5. DBI Score

Algorithm	DBI Score
K-Means	0.7958673
K-Medoids	0.9061

The clustering performance was assessed using the Davies-Bouldin Index (DBI), which evaluates how compact the clusters are internally and how distinct they are from one another. A smaller DBI score signifies better clustering, indicating that clusters are both tightly grouped and well separated. Based on the results presented in Table 5, the K-Means algorithm achieved a DBI value of 0.79586, while the K-Medoids algorithm yielded a higher DBI value of 0.9061. These results suggest that both algorithms are effective in mapping the dengue fever intervention zones in the city of Medan. However, K-Means demonstrates superior performance, as its lower DBI value indicates more clearly separated and compact clusters. In contrast, although K-Medoids also produced distinguishable clusters, the higher DBI value suggests that the internal compactness within its clusters could be further improved.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 6, Nomor 4, Juni 2025
e-ISSN 2685-998X
DOI 10.30865/json.v6i4

These findings are supported by several previous studies. For instance, a study conducted by [9] compared K-Means and K-Medoids for mapping diarrhea treatment zones in toddlers. The study found that K-Means produced three optimal clusters, while K-Medoids produced two. The evaluation using DBI showed a lower DBI value for K-Means, indicating better performance in clustering compared to K-Medoids. Similarly, [17] utilized landslide data from West Java Province in 2019 to identify disaster-prone areas. After processing 24 valid data points from 609 recorded landslide incidents in 2018, they tested optimal cluster numbers using DBI with both K-Means and K-Medoids algorithms for values of *k* from 2 to 6. Their results showed that K-Means was more optimal, yielding the lowest DBI of 0.265 at $k = 6$, compared to K-Medoids with a DBI of 0.342 at $k = 2$. In this case, Cluster 2 had the most regions, while Cluster 5 had the highest number of incidents (106) across four regions. Furthermore, a study by [18] demonstrated that both K-Means and K-Medoids were effective for clustering student data in determining study programs. However, after DBI evaluation, K-Means showed better clustering quality with a lower DBI value (1.19010) compared to K-Medoids (1.27833). Despite the different number of clusters produced, the reduced DBI in K-Means indicated more homogeneous and higher-quality clusters, making K-Means the more favorable option. It is important to note that although the number of clusters used in K-Means (four clusters) and K-Medoids (three clusters) differs, the comparison remains valid. Each algorithm determines the optimal number of clusters based on internal evaluation metrics and the distribution of the data. K-Means identified more granular groupings, which allowed for finer distinctions between sub-districts—ranging from very high to very low dengue risk—thereby enhancing the precision of intervention strategies. In contrast, K-Medoids, which generated fewer clusters, provided broader classifications (high, medium, and low risk) that may be less specific but still valuable. Therefore, the comparison highlights not only the clustering performance based on DBI values, but also the level of practical detail each model offers for public health decision-making. From a practical perspective, the clustering results can directly guide public health policies in Medan City. Specifically, Cluster 1 sub-districts, identified as high-priority zones with the highest dengue risk, are recommended for immediate intervention measures such as fogging, larval source reduction, and intensive public health campaigns. Cluster 2 areas, with moderate risk, should receive regular vector surveillance and periodic community awareness programs. Meanwhile, Cluster 3 and Cluster 4 zones, categorized as low and very low risk respectively, can be managed with routine monitoring and preventive health measures. This targeted approach enables more efficient allocation of resources, allowing health authorities to focus efforts where they are most needed and ultimately reduce the overall dengue burden in the city.

4. CONCLUSION

This study aimed to compare the performance of the K-Means and K-Medoids clustering algorithms in mapping priority intervention areas for dengue fever (DBD) in Medan City, using four main variables: the number of cases, population density, area size, and number of healthcare facilities. The findings indicate that both algorithms are effective in classifying regions based on the level of dengue risk. However, based on the evaluation using the Davies-Bouldin Index (DBI)**,** K-Means outperforms K-Medoids. K-Means achieved a lower DBI value of 0.7959, compared to 0.9061 for K-Medoids, suggesting that the clusters formed by K-Means are more compact and well-separated, resulting in better clustering quality. This indicates that K-Means is more reliable in identifying high-priority areas requiring urgent public health interventions. Although K-Medoids also produced meaningful clusters, the relatively higher DBI score suggests potential improvements in internal cluster cohesion. Overall, K-Means is recommended as the more suitable algorithm for dengue fever mapping in Medan due to its superior performance in generating distinct and cohesive clusters, which are crucial for effective resource allocation and targeted disease control efforts.

REFERENCES

[1] M. A. Roffa, “Pemetaan Sebaran Penyakit Demam Berdarah Dengue di Kota Lhokseumawe Tahun 2022-2023,” Universitas malikussaleh, Lhokseumawe, 2024. Accessed: Dec. 18, 2024. [Online]. Available: https://rama.unimal.ac.id/id/eprint/884

[2] R. P. Manurung, “Januari Hingga Juli 2024, Tercatat 558 Kasus DBD di Kota Medan,” Mistar.id. Accessed: Dec. 18, 2024. [Online]. Available: https://mistar.id/edukasi/kesehatan/januari-hingga-juli-2024-tercatat-558-kasus-dbd-di-kota-medan/

[3] M. G. Tansil, N. H. Rampengan, and R. Wilar, “Faktor Risiko Terjadinya Kejadian Demam Berdarah Dengue Pada Anak,” *Jurnal Biomedik:JBM*, vol. 13, no. 1, p. 90, Mar. 2021, doi: 10.35790/jbm.13.1.2021.31760.

[4] A. Hamid, A. Lestari, and I. Maliga, “Analisis Perbandingan Faktor Lingkungan Terkait Dengan Prevalensi Kejadian Demam Berdarah Dengue (DBD) Pada Daerah Sporadis Dan Daerah Endemis,” *Jurnal Kesehatan Lingkungan Indonesia*, vol. 22, no. 1, pp. 13–20, Feb. 2023, doi: 10.14710/jkli.22.1.13-20.

[5] I. G. W. K. Mahardika, M. Rismawan, and I. N. Adiana, “Hubungan Pengetahuan Ibu Dengan Perilaku Pencegahan DBD pada Anak Usia Sekolah di Desa Tegallinggah,” *Jurnal Riset Kesehatan Nasional*, vol. 7, no. 1, pp. 51–57, 2023, doi: https://doi.org/10.37294.

[6] I. J. Panjaitan, T. Wulandari, and B. Susetyo, "Perbandingan Hasil Analisis Cluster K-Means Dan K-Medoids Untuk Pemetaan Mutu SMK," *Jurnal Bayesian: Jurnal Ilmiah Statistika dan Ekonometrika*, vol. 4, no. 1, pp. 63–75, 2024, doi: 10.46306/bay.v4i1.

[7] A. Supriyadi, A. Triayudi, and I. D. Sholihati, "PERBANDINGAN ALGORITMA K-MEANS DENGAN K-MEDOIDS PADA PENGELOMPOKAN ARMADA KENDARAAN TRUK BERDASARKAN PRODUKTIVITAS," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 6, no. 2, pp. 229–240, 2021.

[8] R. Ramadhani and Y. Hendriyani, "PREDIKSI PRESTASI SISWA BERBASIS DATA MINING MENGGUNAKAN ALGORITMA DECISION TREE (STUDI KASUS: SMKN 2 PADANG)," 2021, [Online]. Available: http://ejournal.unp.ac.id/index.php/voteknika/

[9] T. S. Syamfithriani, N. Mirantika, and R. Trisudarmo, "Perbandingan Algoritma K-Means dan K-Medoids Untuk Pemetaan Daerah Penanganan Diare Pada Balita di Kabupaten Kuningan," *JURNAL SISTEM INFORMASI BISNIS*, vol. 12, no. 2, pp. 132–139, Mar. 2023, doi: 10.21456/vol12iss2pp132-139.

[10] I. Kamila, U. Khairunnisa, and Mustakim, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau," *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 5, no. 1, pp. 119–125, 2019, Accessed: Dec. 18, 2024. [Online]. Available: https://ejournal.uin-suska.ac.id/index.php/RMSI/article/view/7381/0

[11] R. Anggraini, E. Haerani, J. Jasril, and I. Afrianty, "Pengelompokkan Penyakit Pasien Menggunakan Algoritma K-Means," *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 6, p. 1840, Dec. 2022, doi: 10.30865/jurikom.v9i6.5145.

[12] A. Fitri, "PENERAPAN METODE K-MEANS CLUSTERING UNTUK PENGELOMPOKAN SISWA PADA APLIKASI SIAKAD GUNA MEWUJUDKAN SMART SCHOOL DI UPTD SMP NEGERI 1 PEUSANGAN ," Universitas Malikussaleh, Lhokseumawe, 2024. Accessed: Dec. 20, 2024. [Online]. Available: https://rama.unimal.ac.id/id/eprint/4973/1/SKRIPSI%20AIDILLA%20FITRI.pdf

[13] C. F. Palembang, M. Y. Matdoa, and S. P. Palembang, "Perbandingan Algoritma K-Means Dan K-Medoids Dalam Pengelompokkan Tingkat Kebahagiaan Provinsi Di Indonesia," *BULLET : Jurnal Multidisiplin Ilmu*, vol. 01, no. 5, pp. 830–839, 2022, Accessed: Dec. 18, 2024. [Online]. Available: https://journal.mediapublikasi.id/index.php/bullet/article/view/1135

[14] A. Talia, S. Suhartini, and R. Suprianto, "Rancang Bangun Aplikasi Pelayanan Sistem Rujukan Pada Puskesmas Sukajadi Berbasis Web," *Jurnal Minfo Polgan*, vol. 13, no. 2, pp. 1367–1376, Sep. 2024, doi: 10.33395/jmp.v13i2.14058.

[15] D. Hastari, F. Nurunnisa, S. Winanda, and D. Dwi Aprillia, "Penerapan Algoritma K-Means dan K-Medoids untuk Mengelompokkan Data Negara Berdasarkan Faktor Sosial-Ekonomi dan Kesehatan," *SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat*, vol. 1, no. 1, pp. 274–281, 2023, [Online]. Available: https://journal.irpi.or.id/index.php/sentimas

[16] P. Vania and B. Nurina Sari, "Perbandingan Metode Elbow dan Silhouette untuk Penentuan Jumlah Klaster yang Optimal pada Clustering Produksi Padi menggunakan Algoritma K-Means," *Jurnal Ilmiah Wahana Pendidikan*, vol. 9, no. 21, pp. 547–558, 2023, doi: 10.5281/zenodo.10081332.

[17] M. Herviany, P. D. Saleha, T. Nurhidayah, and Kasini, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokkan Daerah Rawan Tanah Longsor di Provinsi Jawa Barat," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 1, no. 1, pp. 34–40, 2021, Accessed: Dec. 18, 2024. [Online]. Available: https://journal.irpi.or.id/index.php/malcom

[18] Salamah, D. Abdullah, and Nurdin, "Comparative Analysis of K-Means and K-Medoids to Determine Study Programs," *International Journal of Engineering, Science and Information Technology*, vol. 5, no. 1, pp. 167–176, Dec. 2024, doi: 10.52088/ijesty.v5i1.673.