Community Satisfaction Segmentation of Services at the Kudus Regency Population and Civil Registration Office Using K-Means Clustering

Dinda Aulia[1], Agus Subhan Akbar[2], Alzena Dona Sabilla[3]
[1,2,3] Department of Information Systems, Faculty of Science and Technology,
Nahdlatul Ulama Islamic University of Jepara, Indonesia

ABSTRACT

The Community Satisfaction Survey is an instrument used to evaluate the quality of public services. Although the Community Satisfaction Index of the Kudus Regency Population and Civil Registration Office (Disdukcapil Kabupaten Kudus) has consistently been classified as excellent, analyses based solely on average scores are insufficient to capture variations in satisfaction levels across different community groups. This study aims to segment community satisfaction using the K-Means Clustering algorithm and to generate evidence-based recommendations for service quality improvement. The dataset comprised 11,883 SKM records collected from the first quarter of 2024 to the fourth quarter of 2025, covering nine service quality indicators (U1–U9). The optimal number of clusters was determined using the Elbow Method and subsequently validated through the Silhouette Score and Davies-Bouldin Index. The results indicate that the optimal clustering solution is k = 2, yielding a Silhouette Score of 0.7231 and a Davies-Bouldin Index of 0.8265. The segmentation identified Cluster 1, consisting of 9,618 respondents (80.9%) with relatively high satisfaction levels, and Cluster 2, consisting of 2,265 respondents (19.1%) with relatively lower satisfaction levels. Service processing time, service requirements, service products, staff competency, and complaint handling were identified as the primary areas requiring improvement. The findings were subsequently implemented in a web-based analytics system to support evidence-based evaluation and decision-making for the continuous improvement of public service quality.

Keywords: K-Means Clustering, Community Satisfaction Survey, Satisfaction Segmentation, Public Service, Evidence-Based Recommendation.

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Corresponding Author:
Agus Subhan Akbar,
Department of Information System,
Universitas Islam Nahdlatul Ulama Jepara,
Jl. Taman Siswa, Pekeng, Kauman, Tahunan, Kec. Tahunan, Kabupaten Jepara, Jawa Tengah 59451
Email: agussa@unisnu.ac.id

1. INTRODUCTION

Public services encompass all forms of public goods and public service delivery provided by government institutions to meet the needs of society in accordance with applicable legal provisions [1]. Law No. 25 of 2009 on Public Services emphasizes the responsibility of public service providers to deliver high-quality, measurable services that are oriented toward community satisfaction. The primary objective of public service delivery is to effectively fulfill citizens' needs while strengthening public trust in government performance. Furthermore, the provision of high-quality public services reflects the principles of good governance by promoting transparency, accountability, and efficiency in service delivery [2].

To objectively assess and monitor service quality, the Indonesian Government, through Indonesia's Ministry of Administrative and Bureaucratic Reform (Kementerian Pendayagunaan Aparatur Negara dan Reformasi Birokrasi, hereafter referred to as KemenPAN-RB), has established the Community Satisfaction Survey (Survei Kepuasan Masyarakat) as the national standard instrument for evaluating public service quality. In accordance with Ministerial Regulation No. 14 of 2017 of the Ministry of Administrative and Bureaucratic Reform, which supersedes Ministerial Decree No. KEP/25/M.PAN/2/2004, community satisfaction data and information are obtained through both quantitative and qualitative assessments of citizens' perceptions of public services [3]. Community satisfaction assessment has been widely implemented across various public service institutions as a foundation for evaluating and improving service quality [4]. Consequently, the Community Satisfaction Survey provides government agencies with valuable insights into citizens' perceptions of service quality and serves as an evidence base for continuous service evaluation and improvement. According to [5],

ISSN: 2722-0001

evaluating public services constitutes a fundamental component of government performance management, as it assesses the effectiveness and efficiency of service delivery while providing an evidence-based foundation for policy formulation aimed at improving public service quality.

The Population and Civil Registration Office of Kudus Regency (Disdukcapil Kabupaten Kudus) has conducted the Community Satisfaction Survey on a quarterly basis. Based on Community Satisfaction Survey data collected from the first quarter of 2024 to the fourth quarter of 2025, the Community Satisfaction Index (IKM) consistently remained within the "Excellent" category, with scores ranging from 92.64 to 99.48. Nevertheless, considerable fluctuations were observed across reporting periods. In particular, the Community Satisfaction Index declined during the third quarter of both years, reaching 92.69 in 2024 and 92.64 in 2025. This recurring decline highlights the need for further analysis to identify satisfaction patterns across respondent groups and service indicators that may not be captured by aggregate scores alone.

Previous studies have applied K-Means clustering to analyze satisfaction with public services. Burelia et al. [1] employed K-Means to analyze community satisfaction with police services, particularly the issuance of police clearance certificates (SKCK), at Dumai Resort Police. Hariyanto et al. [2] investigated public service satisfaction using nine Community Satisfaction Survey indicators, providing a closely related methodological reference for the present study. Meanwhile, Sholeh and Aeni [6] examined the use of the Elbow Method, Silhouette Score, and Davies–Bouldin Index for selecting and evaluating the number of clusters in K-Means analysis. Similarly, Syahkur et al. [7] evaluated the number of clusters in K-Means++ using the Elbow and Silhouette methods, followed by validation with the Davies–Bouldin Index (DBI). Their study demonstrates the use of complementary evaluation methods to assess clustering configurations. These studies provide relevant empirical and methodological foundations for the present research, but the present study focuses on integrating multi-period satisfaction segmentation, cluster validation, institutional follow-up evidence, and web-based analytical implementation within a single public service evaluation framework.

The main research gap addressed in this study is the need to move beyond aggregate satisfaction scores by examining differences in satisfaction profiles across respondent groups and reporting periods, while connecting the results to documented institutional improvement measures. To address this gap, this study analyzes 11,883 Community Satisfaction Survey records collected over eight consecutive quarters, from the first quarter of 2024 to the fourth quarter of 2025. The Elbow Method is used to support the selection of the number of clusters, while the Silhouette Score and Davies–Bouldin Index are used to evaluate clustering quality. The resulting segments are examined alongside quarterly service evaluation reports and information obtained through interviews with personnel responsible for these reports. The findings are then integrated into a web-based analytics system to support cross-period analysis and evidence-based service improvement recommendations.

K-Means Clustering is one of the most widely adopted data mining techniques for segmentation analysis due to its capability to partition large numerical datasets based on feature similarity. The algorithm partitions data into several clusters by maximizing intra-cluster similarity while minimizing inter-cluster similarity. Through this approach, respondents can be grouped according to their satisfaction levels [8]. Such segmentation facilitates a more detailed examination of each cluster, allowing service indicators with relatively low performance to be identified. Consequently, service improvement recommendations can be tailored to the specific characteristics of each respondent group.

The quality of clustering results is commonly evaluated using internal validation metrics, particularly the Silhouette Score and the Davies-Bouldin Index (DBI). The Silhouette Score evaluates the extent to which each observation is appropriately assigned to its own cluster compared with the nearest neighboring cluster, whereas the Davies-Bouldin Index evaluates the degree of separation and compactness among clusters. Together, these metrics provide a comprehensive assessment of the quality and homogeneity of clusters generated by clustering algorithms [9] . In addition, the Elbow Method is widely used to determine the optimal number of clusters by identifying the inflection point in the cluster evaluation curve [6].

The main contributions of this study are: (a) segmenting community satisfaction using data collected over eight consecutive quarters; (b) evaluating the clustering results using the Elbow Method, Silhouette Score, and Davies–Bouldin Index; (c) examining changes in satisfaction patterns across reporting periods; and (d) implementing the analysis results in a web-based analytics system that integrates satisfaction segmentation and evidence-based service improvement recommendations. Accordingly,

ISSN: 2722-0001

this study aims to identify community satisfaction segments at Disdukcapil Kabupaten Kudus and translate the findings into targeted service improvement recommendations.

2. RESEARCH METHOD

A. Dataset and Research Variables

This study utilized Community Satisfaction Survey data collected by the Population and Civil Registration Office of Kudus Regency (Disdukcapil Kabupaten Kudus) from the first quarter of 2024 through the fourth quarter of 2025. The dataset comprised nine Community Satisfaction Survey assessment indicators (U1–U9), representing the key dimensions of public service quality. The data were obtained from periodic surveys conducted by the Population and Civil Registration Office of Kudus Regency and served as the basis for the community satisfaction segmentation process. These nine service indicators were adopted from the Community Satisfaction Survey framework stipulated in the Regulation of the Minister of Administrative and Bureaucratic Reform No. 14 of 2017 and are presented in Table 1.

Table 1. Description of Service Quality Indicators

Indicator	Description
U1	Service Requirements
U2	Service Procedures
U3	Service Processing Time
U4	Service Fees/Tariffs
U5	Service Products
U6	Staff Competency
U7	Staff Behavior
U8	Facilities and Infrastructure
U9	Complaint Handling

B. Research Flow

Fig 1 presents the overall research workflow adopted in this study. This study began with the collection of Community Satisfaction Survey data from the Population and Civil Registration Office of Kudus Regency (Disdukcapil Kabupaten Kudus), covering the period from the first quarter of 2024 to the fourth quarter of 2025. The collected data subsequently underwent preprocessing before determining the optimal number of clusters using the Elbow Method. The next stage involved applying the K-Means Clustering algorithm, followed by an evaluation of the clustering results using the Silhouette Score and the Davies-Bouldin Index to assess cluster quality. The resulting segmentation was then analyzed to identify the characteristics of each cluster and determine priority service indicators requiring improvement. Finally, the clustering results were integrated with quarterly Community Satisfaction Survey (SKM) reports from the Population and Civil Registration Office of Kudus Regency and information obtained through direct interviews with the personnel responsible for handling these reports. The quarterly reports documented improvement measures for service indicators with low scores, while the interviews provided additional information regarding the handling of the reports and related improvement efforts. These sources were considered alongside the clustering results to formulate evidence-based service improvement recommendations, which were subsequently incorporated into a web-based analytics system.

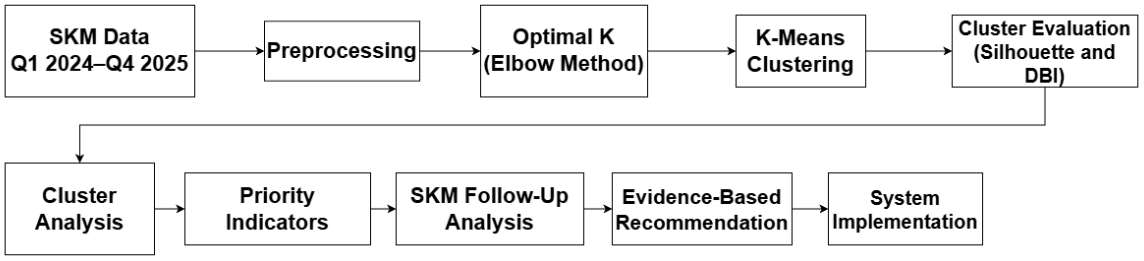


Fig 1. Research Framework

ISSN: 2722-0001

C. Data Preprocessing

The research process began with a data preprocessing stage to ensure data quality prior to clustering. During this stage, the dataset was examined and cleaned by verifying the absence of duplicate records, missing values, and outliers [10] . The examination identified no duplicate records or missing values, and all 11,883 records were retained for further analysis.

D. K-Means Clustering

The preprocessed data were then grouped using the K-Means Clustering algorithm, which partitions data into clusters based on the similarity of their characteristics, ensuring that observations with similar attributes are assigned to the same cluster[11]. As a partition-based clustering algorithm, K-Means represents each cluster by a centroid that serves as its center [12] . The clustering procedure begins by specifying the desired number of clusters (K) and initializing the cluster centroids. The Euclidean distance between every observation and each centroid is subsequently computed to determine the closest cluster assignment. After all observations have been allocated, new centroid positions are obtained by calculating the mean value of the observations within each cluster. These assignment and centroid update steps are performed repeatedly until the cluster memberships remain unchanged or the centroid locations stabilize. The primary objective of the K-Means algorithm is to minimize the total within-cluster sum of squared distances between observations and their respective centroids, as expressed in Equation (1).

$$E = \sum_{j=1}^{k} \sum_{i=1}^{n_j} \left\| x_{ij} - c_j \right\|^2 \tag{1}$$

Where E denotes the total error, representing the objective function to be minimized; $x_i(j)$ represents the i data point in the j, cluster; c_j denotes the centroid of the j, cluster; k denotes the total number of clusters; and n_j represents the number of data points assigned to the j cluster.

E. Determination of the Optimal Number of Clusters

The Elbow Method was employed to identify the most appropriate number of clusters by examining the variation in Within-Cluster Sum of Squares (WCSS) values across multiple cluster configurations. This method identifies the most appropriate number of clusters by locating the elbow point on the WCSS curve. The elbow point represents the stage at which increasing the number of clusters no longer results in a substantial reduction in the WCSS value, thereby providing a basis for selecting the optimal number of clusters for the K-Means Clustering process. The WCSS is calculated using Equation (2).

$$WCSS = \sum_{k=1}^{K} \sum_{i \in C_k} ||x_i - \mu_k||^2 \tag{2}$$

Where $WCSS$ denotes the total within-cluster sum of squared distance, k is the total number of clusters, C_k represents the set of data points assigned to cluster k, x_i denotes the i-th data point, and μ_k represents the centroid of cluster k.

F. Silhouette Score

The Silhouette Score is used to measure the degree of similarity between a data point and the cluster to which it belongs relative to other clusters. A Silhouette Score closer to 1 indicates better clustering quality, whereas a lower value suggests that the separation between clusters is not yet well defined [13]. The Silhouette Score is calculated using Equation (3).

$$S(i) = \frac{b_i - a_i}{max(a_i, b_i)} \tag{3}$$

ISSN: 2722-0001

Where $S(i)$ denotes the Silhouette Score of the $i-th$ data point, a_i represents the Euclidean distance between the $i-th$ data point and the centroid of its assigned cluster, while b_i denotes the minimum Euclidean distance between the $i-th$ data point and the centroid of the nearest neighboring cluster.

G. Davies-Bouldin Index

The Davies-Bouldin Index (DBI) is an internal cluster validation metric used to evaluate clustering quality based on cluster cohesion and separation. Cohesion refers to the degree of similarity between data points and the centroid of their assigned cluster [14]. whereas separation reflects the distinctiveness between different clusters. The Davies-Bouldin Index assesses clustering quality by considering both the similarity among clusters and the dispersion of data points within each cluster. A lower DBI value indicates better clustering performance, as it reflects more compact clusters with clearer separation. In addition, internal validity measures such as the Davies-Bouldin Index (DBI) are widely utilized to verify cluster separation and compactness, with lower DBI values confirming well-defined cluster boundaries. The inter-cluster similarity is calculated using Equation (4).

1). Inter-Cluster Similarity

$$R_{ij} = \frac{S_i + S_j}{M_{ij}} \tag{4}$$

Where Rij denotes the similarity ratio between cluster i and cluster j, Si and Sj represent the Euclidean distance of all data points in cluster i and cluster j, respectively, Mij denotes the distance between the centroids of cluster i and cluster j. The Davies-Bouldin Index is calculated using Equation (5).

2). Davies–Bouldin Index Calculation

$$DBI = \frac{1}{K}\sum_{i=1}^{K} max\ _{j\neq i}(R_{ij}) \tag{5}$$

Where DBI denotes the Davies-Bouldin Index, k is the total number of clusters, and R_{ij} represents the similarity ratio between cluster i and cluster j. Furthermore, S_i dan S_j denote the average distance of the data points in cluster i and cluster j to their respective centroids, while M_{ij} represents the distance between the centroids of the two clusters.

H. Evidence-Based Recommendation

Following the clustering and cluster evaluation processes, evidence-based recommendations were developed by integrating the segmentation results with quarterly Community Satisfaction Survey (SKM) reports and information obtained through direct interviews with personnel responsible for handling these reports at the Population and Civil Registration Office of Kudus Regency (Disdukcapil Kabupaten Kudus). The clustering results were used to identify differences in community satisfaction profiles and prioritize service indicators requiring improvement based on centroid differences and centroid values in Cluster 2. The quarterly reports were examined to identify documented improvement measures relevant to the selected indicators, while the interviews provided additional information regarding the handling of these reports and related improvement efforts. The documented improvement measures were considered when formulating recommendations for the priority indicators identified through clustering. This approach integrated the quantitative segmentation results with the improvement measures documented by the institution to formulate evidence-based service improvement recommendations.

I. Web-Based Analytics System Implementation

The community satisfaction segmentation results were implemented in a web-based analytics system to support the service evaluation process at the Population and Civil Registration Office of Kudus Regency (Disdukcapil Kabupaten Kudus). The system was developed using React.js with Vite as the build tool and Firebase Realtime Database as the database management platform, enabling integrated data management and real-time updates through a web browser. The core functionalities of the system include Community Satisfaction Survey data

ISSN: 2722-0001

management, community satisfaction segmentation using the K-Means Clustering algorithm, cluster evaluation using the Elbow Method, Silhouette Score, and Davies-Bouldin Index, and the presentation of analytical results through tabular reports and interactive visualizations. In addition, the system supports cluster characteristic analysis, monitoring of changes in service quality across reporting periods, and the generation of service improvement recommendations to facilitate more effective and efficient decision-making.

J. System Testing

System testing was conducted using the Black Box Testing method to ensure that all implemented functionalities operated according to user requirements and produced outputs consistent with the objectives of the study. The testing process focused on the Community Satisfaction Survey data management module, the K-Means Clustering process, cluster evaluation using the Elbow Method, Silhouette Score, and Davies-Bouldin Index, cluster characteristic analysis, the generation of evidence-based recommendations, and the monitoring of changes in service quality across reporting periods. The system was considered successful when each function produced the expected output and no errors were identified in either the data processing procedures or the presentation of analytical information.

3. RESULTS AND DISCUSSION

A. Data Preprocessing Results

The preprocessing stage was conducted to ensure data quality before performing the clustering analysis. Based on the preprocessing results, the dataset comprised 11,883 respondents from the Community Satisfaction Survey conducted by the Population and Civil Registration Office of Kudus Regency (Disdukcapil Kabupaten Kudus) between the first quarter of 2024 and the fourth quarter of 2025. The results of the preprocessing stage are summarized in Table 2.

Table 2. Data Preprocessing Results

Description	Count
Initial Number of Records	11.883
Duplicate Records	0
Missing Values	0
Removed Records	0
Final Number of Records	11.883

The preprocessing results confirmed that no missing values or duplicate records were identified, allowing the entire dataset to be retained for subsequent analysis The entire dataset was retained for subsequent analysis. The nine service quality indicators (U1–U9) were maintained on their original 1–4 rating scale, without applying Min-Max normalization, before implementing the K-Means Clustering algorithm. Descriptive statistics of the Community Satisfaction Survey indicators are presented in Table 3.

Table 3. Descriptive Statistics of Public Service Satisfaction Indicators

Community Satisfaction Survey Service Indicator	Mean	Median	Min	Max	Variance	Std. Deviation
U1 – Service Requirements Compliance with administrative requirements	3.827	4.00	1	4	0.1701	0.4125
U2 – Service Procedures Ease of service procedures	3.838	4.00	1	4	0.1542	0.3927
U3 – Service Processing Time Timeliness of document completion	3.803	4.00	1	4	0.1793	0.4235
U4 – Service Fees/Tariffs Transparency of service fees (IDR 0)	3.993	4.00	1	4	0.0147	0.1211
U5 – Service Products Conformity of the issued documents	3.833	4.00	1	4	0.1477	0.3843
U6 – Staff Competency Staff expertise and competence	3.824	4.00	1	4	0.1564	0.3954
U7 – Staff Behavior Staff courtesy and professionalism	3.896	4.00	1	4	0.0993	0.3152
U8 – Facilities and Infrastructure Adequacy of service facilities and infrastructure.	3.973	4.00	1	4	0.0414	0.2036
U9 – Complaint Handling Responsiveness to complaint resolution	3.821	4.00	1	4	0.1623	0.4029

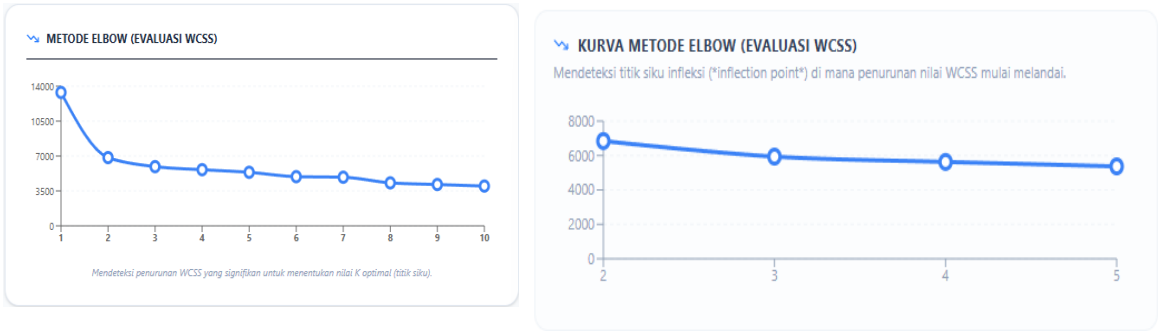
Descriptive statistical analysis was performed to provide an overview of the characteristics of the Community Satisfaction Survey data used in this study. As presented in Table 3, all service indicators exhibited mean values above 3.8, with a median of 4.00, indicating that overall community satisfaction

ISSN: 2722-0001

with the services provided by the Population and Civil Registration Office of Kudus Regency was classified as Excellent. Indicator U4 (Service Fees/Tariffs) achieved the highest mean score of 3.993 with the lowest standard deviation (0.1211), indicating a highly consistent evaluation among respondents. In contrast, Indicator U3 (Service Processing Time) recorded the lowest mean score (3.803) and the highest standard deviation (0.4235), suggesting greater variability in respondents' perceptions compared with the other indicators. Although the overall level of community satisfaction was high, the observed variability across several service indicators indicates differences in respondents' perceptions, which warrant further investigation through the clustering process.

B. Determination of the Optimal Number of Clusters

The Elbow Method was employed to determine the optimal number of clusters by identifying the elbow point on the Sum of Squared Errors (SSE) curve [15] , where the reduction in SSE begins to level off after a substantial initial decrease. In this study, the optimal number of clusters was determined by analyzing changes in the Within-Cluster Sum of Squares (WCSS). The WCSS evaluation results for K = 1 to K = 10 are presented in Fig. 2(a), while Fig. 2(b) provides a more detailed visualization for the range of K = 2 to K = 5.



(a) (b)

Fig 2. Elbow Curve for Determining the Optimal Number of Clusters
(a) WCSS Values for K = 1–10, (b) Enlarged view of the elbow curve for K = 2–5.

As illustrated in Fig. 2, the most substantial decrease in WCSS occurred when the number of clusters increased from K = 1 to K = 2. Beyond this point, the reduction in WCSS became less pronounced, suggesting an elbow point at K = 2. This finding was further supported by the internal validation metrics. Among the evaluated configurations (K = 2–5), K = 2 achieved the highest Silhouette Score (0.7231) and the lowest Davies–Bouldin Index (DBI) (0.8265). These results indicate that the two-cluster solution provides a more favorable balance between cluster compactness and separation than the other evaluated configurations. Therefore, K = 2 was selected as the optimal number of clusters for subsequent analysis. The comparative evaluation results are presented in Table 4.

Table 4. Comparison of Clustering Results for K = 2–5

K	WCSS	Silhouette	DBI	Status
2	6842.173	0.7231	0.8265	Optimal
3	5936.386	0.7059	1.5477	Sub-optimal
4	5627.731	0.7069	1.3406	Sub-optimal
5	5369.887	0.7037	1.5973	Sub-optimal

The clustering evaluation favored K = 2 among the configurations examined. Although WCSS decreased as the number of clusters increased, K = 2 achieved the highest Silhouette Score (0.7231) and the lowest Davies–Bouldin Index (DBI) (0.8265) among the evaluated alternatives. The Silhouette Score reflects how well observations fit their assigned clusters relative to neighboring clusters, while a lower DBI indicates better cluster compactness and separation. These results confirm that K = 2 provides the most favorable clustering structure among the evaluated configurations and was therefore selected for the subsequent interpretation of community satisfaction segments.

C. K-Means Clustering Results

ISSN: 2722-0001

The K-Means algorithm was applied using the optimal configuration of K = 2, as determined by the previous evaluation. The clustering process divided the 11,883 respondents into two clusters: Cluster 1, comprising 9,618 respondents (80.9%), and Cluster 2, comprising 2,265 respondents (19.1%). The final centroid values for each service indicator are presented in Table 5.

Table 5. Final Centroid Values and Optimal Cluster Positions (K = 2)

Community Satisfaction Survey Indicator	Cluster 1	Cluster 2
U1 – Service Requirements	3.9705	3.2168
U2 – Service Procedures	3.9753	3.2530
U3 – Service Processing Time	3.9571	3.1497
U4 – Service Fees/Tariffs	3.9984	3.9691
U5 – Service Products	3.9759	3.2265
U6 – Staff Competency	3.9652	3.2269
U7 – Staff Behavior	3.9917	3.4914
U8 – Facilities and Infrastructure	3.9963	3.8743
U9 – Complaint Handling	3.9485	3.2790
Average Cluster Index	3.9754	3.4096

The results indicate two distinct community satisfaction profiles. Cluster 1 represents respondents with consistently high satisfaction scores across the nine service indicators, whereas Cluster 2 represents respondents with comparatively lower scores on most indicators. This distinction suggests that, despite the generally high overall assessment of service quality, satisfaction varies among respondents. Cluster 2 therefore warrants further examination to identify the service dimensions associated with lower satisfaction and to support more targeted improvement efforts. The centroid patterns are visualized in the heatmap presented in Table 6.

Table 6. Heatmap Matrix of Centroid Coordinates

Cluster	U1	U2	U3	U4	U5	U6	U7	U8	U9
1	3.97	3.98	3.96	4.00	3.98	3.97	3.99	4.00	3.95
2	3.22	3.25	3.15	3.97	3.23	3.23	3.49	3.87	3.28

The centroid heatmap reveals differences in satisfaction profiles across the nine service indicators. The largest centroid gap was observed in U3 (Service Processing Time), at 0.8074, followed by U1 (Service Requirements, 0.7537), U5 (Service Products, 0.7494), U6 (Staff Competency, 0.7383), and U9 (Complaint Handling, 0.6695). These indicators distinguish the two clusters most clearly, particularly because their centroid values are lower in Cluster 2. In contrast, U4 (Service Fees/Tariffs) and U8 (Facilities and Infrastructure) showed smaller gaps of 0.0293 and 0.1220, respectively. These results identify service processing time as the most prominent dimension differentiating the clusters, followed by service requirements, service products, staff competency, and complaint handling. The findings provide a basis for prioritizing service improvements, which should also be considered alongside quarterly SKM follow-up reports and interview findings.

D. Cluster Evaluation

This section assesses the quality of the selected clustering structure using the Silhouette Score and the Davies–Bouldin Index (DBI). The evaluation results for K = 2 are presented in Table 7.

Table 7. Evaluation Results of the Optimal Clustering Model (K = 2)

Evaluation Metric	Value
Silhouette Score	0.7231
Davies-Bouldin Index	0.8265

The selected clustering model achieved a Silhouette Score of 0.7231, indicating that observations were generally well matched to their assigned clusters relative to neighboring clusters. The DBI value of 0.8265 further supports the clustering structure by indicating a favorable balance between within-cluster dispersion and separation between clusters. Together, these metrics support the quality of the resulting two-cluster structure for subsequent analysis of community satisfaction profiles.

E. Cluster Characteristic Analysis

The centroid profiles of the two clusters are compared in Fig. 3. The radar chart illustrates the differences in satisfaction scores across the nine service indicators.

ISSN: 2722-0001

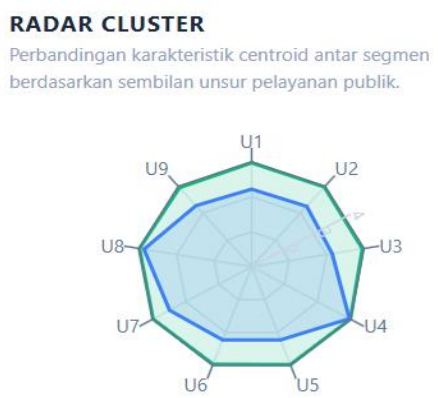


Fig 3. Radar Chart Comparing the Characteristics of Cluster 1 and Cluster 2

The largest centroid differences were observed in U3 (Service Processing Time, 0.8074), U1 (Service Requirements, 0.7537), U5 (Service Products, 0.7494), U6 (Staff Competency, 0.7383), and U9 (Complaint Handling, 0.6695). For these indicators, Cluster 2 consistently recorded lower centroid values than Cluster 1. In contrast, U4 (Service Fees/Tariffs) and U8 (Facilities and Infrastructure) showed smaller differences of 0.0293 and 0.1220, respectively. These findings identify service processing time as the most prominent dimension distinguishing the two satisfaction profiles, followed by service requirements, service products, staff competency, and complaint handling. These indicators are therefore relevant to the subsequent formulation of service improvement recommendations.

F. Analysis of Changes in Service Quality Across Reporting Periods

Changes in community satisfaction during 2024 are presented in Fig. 4. The Community Satisfaction Index fluctuated across the quarterly reporting periods, with the largest decline occurring between the second and third quarters, corresponding to a deviation of −0.233.

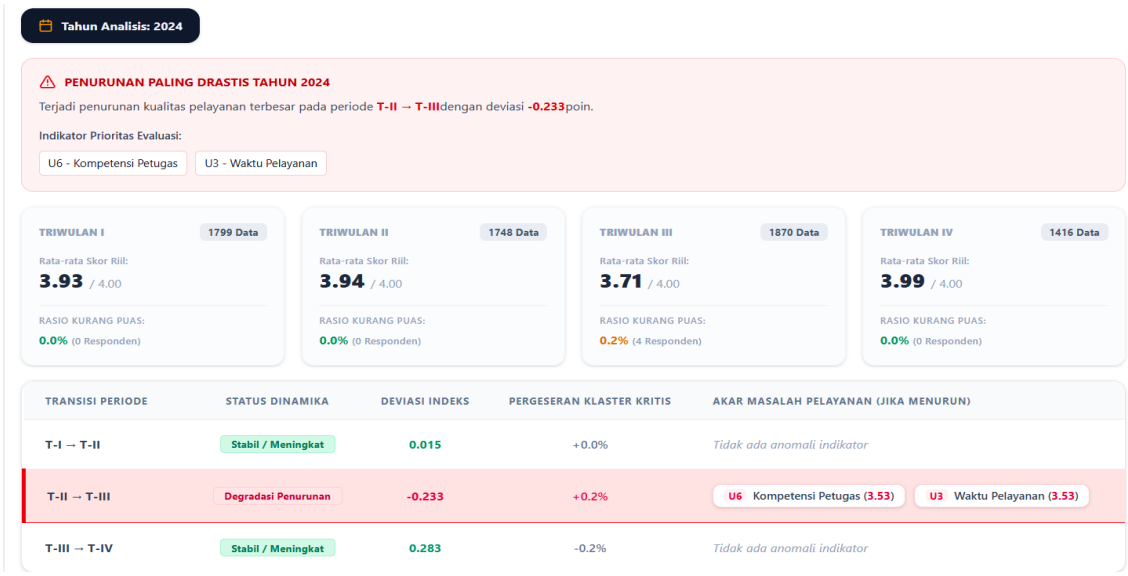


Fig 4. Service Quality Dynamics in 2024

As shown in Fig. 4, the decline from the second to the third quarter was accompanied by lower scores for U6 (Staff Competency) and U3 (Service Processing Time). An improvement was subsequently observed in the fourth quarter, indicating a recovery in the overall satisfaction index. A similar pattern was observed in 2025, as shown in Fig. 5. The largest decline in the Community Satisfaction Index occurred between the second and third quarters, with a deviation of −0.199.

ISSN: 2722-0001

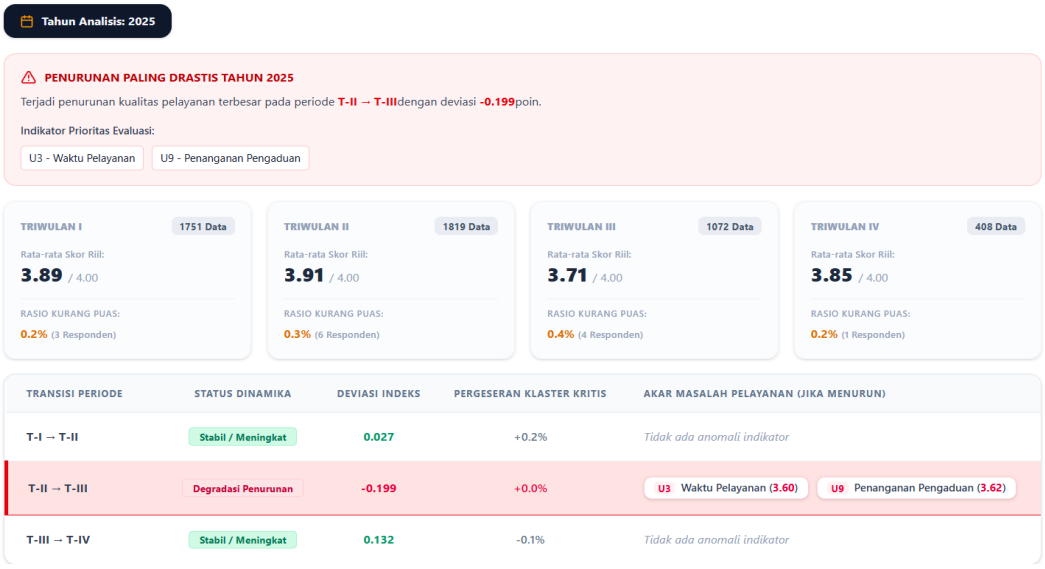


Fig 5. Service Quality Dynamics in 2025.

As shown in Fig. 5, the decline during the third quarter was accompanied by lower scores for U3 (Service Processing Time) and U9 (Complaint Handling). The index subsequently improved in the fourth quarter. The recurrence of a third-quarter decline in both years highlights the need to examine variations in service performance across reporting periods. In particular, U3 (Service Processing Time) showed a recurring association with the decline in satisfaction in both years. Therefore, this indicator warrants priority in service improvement planning, alongside other indicators identified through the clustering analysis and quarterly SKM follow-up reports.

G. Evidence-Based Recommendation

Service improvement recommendations were formulated by integrating the K-Means clustering results with the improvement measures documented in the quarterly Community Satisfaction Survey (SKM) reports of the Population and Civil Registration Office of Kudus Regency (Disdukcapil Kabupaten Kudus). The clustering results were used to identify indicators that differentiated the satisfaction profiles, particularly those with lower centroid values in Cluster 2. The quarterly SKM reports were examined to identify documented follow-up measures relevant to the selected indicators. Interviews with personnel responsible for the reports provided additional context regarding the reporting and follow-up process. The recommendations were formulated by considering the quantitative clustering results and the improvement measures documented in the institutional reports. The resulting recommendations are presented in Table 8.

Table 8. Service Improvement Priorities

No	Indicator	Key Findings	Recommended Improvement
1	U3 – Service Processing Time	Exhibited the largest centroid difference between clusters (0.8074) and the lowest centroid value in Cluster 2 (3.1497), indicating a prominent difference in satisfaction regarding service processing time.	Simplify service processes, optimize the population administration information system, strengthen staff capacity, and conduct regular monitoring and evaluation to improve service completion time.
2	U1 – Service Requirements	Showed a substantial centroid difference between clusters (0.7537), with a lower centroid value in Cluster 2 (3.22), indicating differences in satisfaction regarding service requirements.	Standardize and simplify service requirements by providing clear information through digital and offline channels and conducting public outreach to reduce document submission errors.
3	U5 – Service Products	Showed a substantial centroid difference between clusters (0.7494), with a lower centroid	Evaluate service products across service units and regularly update staff knowledge regarding available service

ISSN: 2722-0001

		value in Cluster 2 (3.23), indicating differences in satisfaction regarding service products.	products and applicable service standards.
4	U6 – Staff Competency	Showed a centroid difference of 0.7383, with a lower centroid value in Cluster 2 (3.2269), indicating differences in satisfaction regarding staff competency.	Provide regular training on information systems, regulations, service excellence, and communication skills, while improving the stability and integration of supporting systems.
5	U9 – Complaint Handling	Showed a centroid difference of 0.6695, with a lower centroid value in Cluster 2 (3.2790). Lower scores for this indicator were also observed alongside the decline in the Community Satisfaction Index during the third quarter of 2025.	Improve access to online and offline complaint channels, accelerate responses to submitted complaints, and regularly evaluate complaint types to support continuous service improvement.

As presented in Table 8, U3 (Service Processing Time) was prioritized because it exhibited the largest centroid difference between clusters (0.8074) and the lowest centroid value in Cluster 2 (3.1497). This finding is consistent with the decline in the Community Satisfaction Index (IKM) from the second to the third quarter of both 2024 and 2025, during which lower U3 scores were also observed. U1 (Service Requirements), U5 (Service Products), U6 (Staff Competency), and U9 (Complaint Handling) were included as additional priorities because they also showed substantial centroid differences between clusters. The recommendations for these indicators were aligned with the corresponding follow-up measures documented in the quarterly SKM reports. Although U2 (Service Procedures) also showed a substantial centroid difference (0.7223), the indicators presented in Table 8 were selected by considering the clustering results alongside the documented institutional follow-up measures.

The recommendations in Table 8 connect the quantitative results of the clustering analysis with the service improvement measures documented by Disdukcapil Kabupaten Kudus. The clustering results provide a basis for prioritizing indicators based on differences in satisfaction profiles, while the quarterly SKM reports provide institutionally documented follow-up measures relevant to service improvement planning.

H. Web-Based Analytics System Implementation

The community satisfaction segmentation results were implemented in a web-based analytics system designed to support service evaluation at the Population and Civil Registration Office of Kudus Regency (Disdukcapil Kabupaten Kudus). The system was developed using React.js with Vite as the build tool and Firebase Realtime Database for data storage and management. The dashboard interface is presented in Fig. 6.

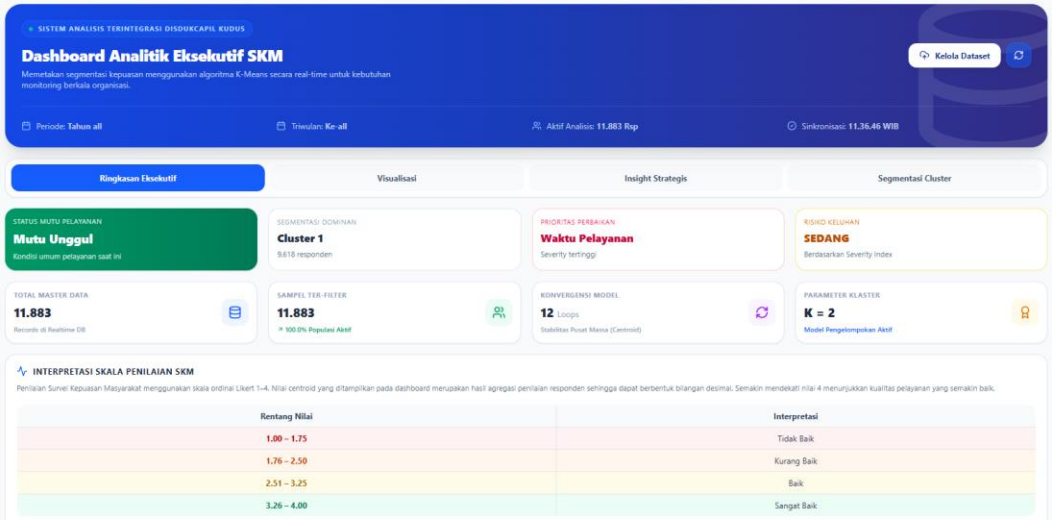


Fig 6. Web-Based Analytics Dashboard Interface

ISSN: 2722-0001

The system provides functionalities for managing Community Satisfaction Survey (SKM) data, performing data preprocessing, applying K-Means clustering, and evaluating cluster quality using the Elbow Method, Silhouette Score, and Davies–Bouldin Index (DBI). The analytical results are presented through interactive visualizations that support the interpretation of cluster characteristics and the monitoring of service quality across reporting periods. The system also presents service improvement recommendations based on the clustering results and supporting institutional information. These functionalities are intended to facilitate the interpretation of community satisfaction patterns and support evidence-informed decision-making for continuous public service improvement.

I. System Testing

The system was evaluated using the Black-Box Testing method to assess whether its core functionalities produced the expected outputs. Testing involved interacting with the web-based analytics system and comparing the observed outputs with the expected results. Seven core functionalities were examined: Community Satisfaction Survey (SKM) data management, data preprocessing, K-Means clustering, cluster evaluation, cluster characteristic analysis, analysis of changes in service quality across reporting periods, and evidence-based recommendation presentation. The test results are summarized in Table 9.

Table 9. Black-Box Testing Results

No	Functionality	Expected Result	Observed Result
1	Community Satisfaction Survey data management	Survey data can be displayed and managed through the web-based analytics system interface.	The system displayed the available survey data through its interface.
2	Data preprocessing	Survey data can be checked for duplicate records and missing values while retaining the original 1–4 rating scale for clustering.	The preprocessing results retained the original 1–4 rating scale, with 11,883 records available for clustering.
3	K-Means Clustering	The web-based analytics system can perform clustering and present the resulting clusters.	The system executed the K-Means clustering process and displayed the resulting clusters.
4	Cluster evaluation (Elbow Method, Silhouette Score, DBI)	The web-based analytics system can present the Elbow Method, Silhouette Score, and Davies–Bouldin Index results.	The system displayed the Elbow Method, Silhouette Score, and Davies–Bouldin Index results.
5	Cluster characteristic analysis	The web-based analytics system can present the characteristics of the resulting clusters.	The system displayed the characteristics of the resulting clusters.
6	Analysis of changes in service quality across reporting periods	The web-based analytics system can present changes in service quality across reporting periods.	The system displayed changes in service quality across the selected reporting periods.
7	Evidence-Based Recommendation	The web-based analytics system can present service improvement recommendations based on the analysis results.	The system displayed service improvement recommendations based on the analysis results.

The Black-Box Testing was conducted manually by comparing the expected outputs with the outputs observed during interaction with the web-based analytics system. The evaluation covered seven core functionalities, including data management, preprocessing, clustering, cluster evaluation, cluster characteristic analysis, reporting-period analysis, and evidence-based recommendation presentation. The results provide a functional assessment of the system's main features in supporting community satisfaction segmentation and public service quality evaluation at Disdukcapil Kabupaten Kudus.

4. CONCLUSION

This study demonstrates the applicability of the K-Means Clustering algorithm for segmenting community satisfaction with services provided by the Population and Civil Registration Office of Kudus Regency (Disdukcapil Kabupaten Kudus). Based on the Elbow Method, Silhouette Score, and Davies–Bouldin Index (DBI), the two-cluster solution was selected as the most suitable configuration among the evaluated alternatives, achieving a Silhouette Score of 0.7231 and a DBI of 0.8265. These evaluation results support the suitability of the resulting cluster structure for further analysis of community satisfaction profiles.

ISSN: 2722-0001

The analysis of 11,883 survey records produced Cluster 1, comprising 9,618 respondents (80.9%) with a mean cluster score of 3.9754, and Cluster 2, comprising 2,265 respondents (19.1%) with a mean cluster score of 3.4096. The cluster characteristics revealed differences in satisfaction profiles, particularly in Service Processing Time (U3), Service Requirements (U1), Service Products (U5), Staff Competency (U6), and Complaint Handling (U9). The analysis of quarterly reporting periods also identified declines in the Community Satisfaction Index from the second to the third quarter in both 2024 and 2025. Lower scores for Service Processing Time (U3) accompanied these declines, highlighting the importance of this indicator in service improvement planning. By integrating the clustering results with quarterly SKM reports and interview findings, the study provides an evidence-informed basis for prioritizing service improvement measures.

The developed web-based analytics system, implemented using React.js, Vite, and Firebase Realtime Database, integrates survey data management, K-Means clustering, cluster evaluation, analytical visualization, and the presentation of service improvement recommendations. The system provides a practical means of interpreting community satisfaction patterns and supporting data-informed evaluation of public services. Its functional performance should be interpreted in accordance with the scenarios and results documented in the Black-Box Testing process.

This study has two main limitations. First, the analysis was conducted using data from a single public service institution, which limits the generalizability of the findings to other institutional contexts. Second, the study employed only K-Means Clustering; therefore, the resulting segmentation has not been compared with alternative clustering approaches. Future research could compare K-Means with methods such as Fuzzy C-Means and Hierarchical Clustering to examine differences in cluster structure and performance. Further development could also incorporate predictive analytics, real-time data monitoring, and more adaptive recommendation mechanisms, provided that their effectiveness is evaluated. Applying the system to other public service institutions would further help assess its scalability and applicability across different organizational and data contexts.

REFERENCES

[1] U. Burelia, G. Urva, A. Sellyana, S. Tinggi Teknologi Dumai, and D. Sekolah Tinggi Teknologi Dumai, "MENGUKUR TINGKAT KEPUASAN MASYARAKAT PADA PELAYANAN KEPOLISIAN RESOR(POLRES) DUMAI MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING," 2022. doi: https://doi.org/10.52072/jutekinf.v10i1.354.

[2] Y. Hariyanto, A. Primadewi, and M. Hanafi, "Analisis Kepuasan Masyarakat terhadap Pelayanan Publik menggunakan K-Means Clustering," *Journal of Information System Research (JOSH)*, vol. 6, no. 2, pp. 1065–1074, Jan. 2025, doi: 10.47065/josh.v6i2.6577.

[3] Ade Ing Heldigard, Abd. Rohman, Dody Setyawan, Nanang Bagus, and Dekki Umamur Ra'is, "SURVEI KEPUASAN MASYARAKAT (IKM) SEBAGAI DASAR PENINGKATAN KUALITAS PELAYANAN PUBLIK DI KELURAHAN DADAPREJO KOTA BATU," *Jurnal Abdi Masyarakat (JAMAS)*, pp. 934–939, 2025.

[4] Eny Boedi Orbawati, Joko Tri Nugraha, and Fadlurrahman, "Mengukur Indeks Kepuasan Masyarakat (Studi Kasus Pelayanan SKCK dan SIM Pada Polres Kota Magelang)," *Jurnal Pemerintahan dan Politik*, vol. 8, no. 1, pp. 36–41, 2023, doi: https://doi.org/10.36982/jpg.v8i1.2882.

[5] P. R. Rosmila, R. Risqiati, and A. S. Darmawan, "KOMPARASI ALGORITMA NAÏVE BAYES DAN DECISION TREE UNTUK MENGANALISIS KEPUASAN MASYARAKAT DPMPTSP KABUPATEN BATANG," *Information System Journal*, vol. 8, no. 02, pp. 109–118, Dec. 2025, doi: 10.24076/infosjournal.2025v8i02.2364.

[6] M. Sholeh and K. Aeni, "PERBANDINGAN EVALUASI METODE DAVIES BOULDIN, ELBOW DAN SILHOUETTE PADA MODEL CLUSTERING DENGAN MENGGUNAKAN ALGORITMA K MEANS," 2023. doi: http://dx.doi.org/10.30998/string.v8i1.16388.

ISSN: 2722-0001

[7] Muhammad Raqib Syahkur, D. Hartama, and S. Solikhun, "Evaluasi Jumlah Cluster pada Algoritma K-Means++ Menggunakan Silhouette dan Elbow dengan Validasi Nilai DBI dalam Mengelompokkan Gizi Balita," *JST (Jurnal Sains dan Teknologi)*, vol. 13, no. 3, pp. 487–496, Oct. 2024, doi: 10.23887/jstundiksha.v13i3.86419.

[8] H. Nopriandi and D. Junika Putra, "PENERAPAN DATA MINING METODE K-MEANS CLUSTERING UNTUK ANALISA PENJUALAN PADA TOKO NUSA INDAH," 2025. Accessed: Jul. 11, 2026. [Online]. Available: https://journal.unilak.ac.id/index.php/Startup/article/view/27384

[9] Y. Hasan, "Pengukuran Silhouette Score dan Davies-Bouldin Index pada Hasil Cluster K-Means dan Dbscan," 2024. Accessed: Jul. 11, 2026. [Online]. Available: https://ejournal.ust.ac.id/index.php/KAKIFIKOM/article/view/3938

[10] F. A. Maresti, W. I. Rahayu, M. B. B. C. Lustin, and T. H. Pakpahan, "IMPLEMENTASI K-MEANS UNTUK MELAKUKAN SEGMENTASI PRODUK BERDASARKAN DATA TRANSAKSI RETAIL," *Jurnal Ilmiah Sains dan Teknologi*, vol. 9, no. 1, pp. 20–32, 2025, doi: https://doi.org/10.47080/saintek.v9i1.3856.

[11] M. Orisa, "Cluster pada Algoritma K-Means," *Seminar Nasional Inovasi dan Aplikasi Teknologidi Industri (SENATI)*, vol. 13, pp. 430–437, 2022, doi: https://doi.org/10.36040/seniati.v6i2.5034.

[12] N. Nur Afidah and Masrukan, "Penerapan Metode Clustering dengan Algoritma K-means untuk Pengelompokkan Data Migrasi Penduduk Tiap Kecamatan di Kabupaten Rembang," *PRISMA, Prosiding Seminar Nasional Matematika*, vol. 6, pp. 729–738, 2023, [Online]. Available: https://journal.unnes.ac.id/sju/index.php/prisma/

[13] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *Jurnal Ilmiah Informatika dan Ilmu Komputer (JIMA-ILKOM)*, vol. 3, no. 1, pp. 46–56, Mar. 2024, doi: 10.58602/jima-ilkom.v3i1.26.

[14] M. I. Ramadhan, A. Nazir, M. Irsyad, S. Sanjaya, and F. Syafria, "Clustering Analysis Using the K-Means Method to Identify Alumni Satisfaction Pattern," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 6, no. 1, pp. 31–41, 2026, doi: 10.57152/malcom.v6i1.2401.

[15] B. Hartono and V. Lusiana, "Analisis Metode Elbow SSE, Silhouette Score, dan Jaccard Stability dalam Pemilihan Jumlah Klaster Data yang Optimal," *Terapan Informatika Nusantara (TIN)*, vol. 6, no. 8, pp. 1521–1532, 2026, doi: 10.47065/tin.v6i8.9271.