Volume 13 Number 01 (2026)

JURNAL REKAYASA SISTEM DAN INDUSTRI
e-ISSN: 2579-9142 p-ISSN: 2356-0843

https://journals.telkomuniversity.ac.id/jrsi

# Organizational Performance-Based Segmentation of Batik SMEs Using K-Means Clustering

Reza Aditya*[1], Bernieka Desyaranti[2], Fandi Achmad[2]

[1] Department of Science Data, Universitas Terbuka, Bandung, Jawa Barat, Indonesia
[2] Department of Industrial Engineering, Telkom University, Bandung, Jawa Barat, Indonesia

ARTICLE INFO

Article history:
Received 23-01-2026
Fixed 15-02-2026
Approved 21-02-2026

Keywords :
Batik SMEs, Organizational Performance, K-Means Clustering, Segmentation, Unsupervised Learning.

ABSTRACT

Batik Small and Medium Enterprises (SMEs) play a strategic role in the economy and cultural preservation in Indonesia, yet exhibit varying levels of organizational performance. This heterogeneity demands a more objective and data-driven segmentation approach so that development strategies and policies can be implemented effectively. This study aims to segment Batik SMEs based on organizational performance using the K-Means clustering method. Primary data were obtained through a structured questionnaire survey of 56 Batik SMEs, with organizational performance indicators covering operational and financial performance. All data were normalized using StandardScaler to ensure scale equality between variables. The optimal number of clusters was determined using the Elbow method and the silhouette coefficient. The analysis results showed that a two-cluster configuration was the optimal solution with the highest silhouette coefficient value. The K-Means model resulted in two segments of Batik SMEs with significantly different organizational performance characteristics: SMEs with low organizational performance and SMEs with high organizational performance. Centroid value analysis and cluster visualization confirmed clear cluster separation and a good level of internal homogeneity. These findings indicate that segmenting Batik SMEs based on organizational performance using an unsupervised learning approach is effective in uncovering the structure of performance heterogeneity. This research contributes to providing a data-driven segmentation framework that can be utilized as a tool to support managerial decision-making and policy formulation to sustainably improve the performance and competitiveness of Batik SMEs.

## 1. Introduction

Small and Medium Enterprises (SMEs) play a strategic role in the global economy due to their contribution to job creation, economic growth, and strengthening local industrial structures. In many developing countries [1], SMEs also serve as key drivers of creative and culture-based industries. In the Indonesian context, Batik SMEs function not only as economic entities but also as guardians of cultural heritage, possessing strong social, historical, and national identity values. However, the dynamics of market competition, changing consumer preferences, and resource constraints require Batik SMEs to maintain adaptive and sustainable organizational performance to survive and thrive.

Academically, organizational performance is viewed as a multidimensional concept reflecting an organization's ability to effectively achieve operational and financial goals. Operational performance reflects the extent to which SMEs are able to maintain productivity, product quality, customer satisfaction, timely delivery, and market position, while financial performance reflects the organization's ability to manage profit growth, revenue, and investment. In practice, the organizational performance of Batik SMEs shows considerable variation, with some SMEs achieving relatively high performance, while others still face various operational and financial limitations. This variation indicates heterogeneity in performance that requires a more systematic understanding.

However, approaches to developing and fostering batik SMEs are generally uniform and do not fully consider the

Email: reza.dityarr@gmail.com (Aditya, R.)*, berniekadesyaranti@student.telkomuniversity.ac.id (Desyaranti, R.), fandiac.achmad@gmail.com (Achmad, F.)

* Writer correspondence

DOI: https://doi.org/10.25124/jrsi.v13i01.10597

differences in organizational performance characteristics between SMEs [2], [3]. This situation has the potential to cause performance improvement strategies and implemented policies to be less than targeted. From a research perspective, previous studies on SMEs have largely used causal analysis approaches to identify factors influencing organizational performance, but relatively few have utilized organizational performance as the primary basis for SME segmentation. Consequently, empirical and data-driven mapping of batik SMEs based on operational and financial performance profiles remains limited [4].

Developments in data analytics and machine learning offer an opportunity to address these limitations through unsupervised learning approaches, specifically clustering methods [5]. The K-Means algorithm allows grouping SMEs into relatively homogeneous clusters based on similar performance characteristics, without requiring initial assumptions about the data structure. This approach is relevant for uncovering latent patterns and structural heterogeneity in organizational performance that cannot be fully captured by conventional analysis methods.

Based on this background, this study aims to segment batik SMEs based on organizational performance using the K-Means clustering method [6]. By utilizing operational and financial performance indicators, this study attempts to identify the cluster profiles of batik SMEs with distinct performance characteristics [7]. The main contribution of this study lies in providing a data-driven, performance-based segmentation framework for organizations, which not only enriches academic studies on SMEs and organizational performance but also provides an empirical basis for formulating more targeted managerial strategies and policies to improve the performance and competitiveness of batik SMEs in a sustainable manner.

Unlike previous clustering based studies on Batik SMEs, which primarily focused on innovation adoption, cleaner production practices, or marketing strategies, this study introduces a performance-based segmentation framework centered explicitly on Organizational Performance (OP). The novelty of this research lies in integrating two core dimensions Operational Performance (OPF) and Financial Performance (FP) as the primary clustering attributes rather than treating performance merely as an outcome variable in causal analysis. By positioning organizational performance as the main segmentation basis, this study shifts the analytical perspective from factor-driven explanations toward structural performance mapping. This approach provides a more comprehensive understanding of heterogeneity among Batik SMEs and offers a distinct contribution to both SME performance research and data-driven segmentation methodologies.

## 2. Research methods

### *2.1 Data Collection*

This study aims to segment Batik Small and Medium Enterprises (SMEs) based on organizational performance using the K-Means clustering approach [8]. The data collection method used was primary data obtained through a structured questionnaire survey, developed based on a literature review related to organizational performance in SMEs.

The questionnaire was designed to capture Batik SMEs' perceptions of their organization's operational and financial performance, as shown in Table 1. Variable Definition. The Organizational Performance (OP) construct is measured through two main dimensions: Operational Performance (OPF) and Financial Performance (FP). The operational performance dimension reflects the ability of Batik SMEs to achieve productivity targets, maintain customer satisfaction, ensure product quality, timely delivery, and position products in the market [9]. Meanwhile, the financial performance dimension represents the organization's ability to monitor and achieve profit and revenue growth, as well as evaluate investment benefits. The operational and financial performance indicators in Table 1 were adapted from established organizational and SME performance literature by Achmad et al [3].

Table 1.
Variable Definition

| Construct | Indicator | Code | Description |
|---|---|---|---|
| Organizational Performance (OP) | Operational Performance (OPF) | OPF1 | Our organization consistently achieves its productivity targets over the past year |
| | | OPF2 | Our organization ensures customer satisfaction with the products and services provided |
| | | OPF3 | The quality of products produced by our organization meets established standards |
| | | OPF4 | Our organization delivers products and services to customers on time |
| | | OPF5 | Our organization’s products dominate the market |
| | Financial Performance (FP) | FP1 | Our organization monitors net profit trends as a key performance indikator |
| | | FP2 | Our organization evaluates revenue growth as a critical measure of success |
| | | FP3 | Our organization has achieved consistent net profit growth over the past year |
| | | FP4 | Our organization has consistently experienced revenue growth over the past year |
| | | FP5 | Our organization compares investment values with the benefits gained from those investments |

In the context of the Batik industry, the selection of OPF1 - OPF5 and FP1 - FP5 indicators is highly relevant due to the labor-intensive, craftsmanship-based, and quality-driven nature of Batik SMEs. Operational indicators such as productivity, customer satisfaction, product quality, timely delivery, and market positioning reflect the ability of Batik SMEs to compete through authenticity, service reliability, and market differentiation rather than price alone. Meanwhile, financial indicators including profit monitoring, revenue growth, and investment evaluation capture key aspects of financial sustainability, particularly in managing working capital and ensuring business continuity. Collectively, these indicators provide a comprehensive representation of operational resilience and financial sustainability that accurately reflects the performance characteristics of Batik SMEs.

The subjects of this study were Batik SMEs operating in the research area. The primary respondents in this study were the owners or managers of Batik SMEs, who were deemed to have a comprehensive understanding of the operational and financial conditions of their businesses. The sampling technique used was purposive sampling, with the criteria being Batik SMEs that were still actively producing and had been operating for at least two years. This resulted in 56 respondents who were deemed capable of providing objective assessments of organizational performance indicators.

The research instrument used a Likert scale, where respondents were asked to rate each indicator based on its level of implementation within their organization. All indicators in this study were numeric, thus meeting the requirements for analysis using data-driven clustering methods, specifically the K-Means algorithm. This approach allows for the grouping of batik SMEs into segments with relatively homogeneous organizational performance characteristics.

The use of operational and financial performance indicators in this study refers to a previous study, which emphasized the importance of organizational performance as a basis for mapping and strategic decision-making in SMEs. Therefore, the data collected is expected to provide a comprehensive picture of the performance profile of batik SMEs as a basis for organizational performance-based segmentation.

### *2.2 K-Means clustering*

This study used the K-Means clustering method, a type of unsupervised learning, to segment batik SMEs based on organizational performance. This method was chosen because of its ability to group numerical data into clusters based on the degree of similarity in characteristics between objects. All organizational performance indicators used in this study are quantitative, making them suitable for analysis using the K-Means algorithm.

The analysis process was conducted using the Python programming language on the Google Colab platform, which provides a collaborative computing environment and comprehensive data analysis libraries, such as NumPy, Pandas, and Scikit-learn. Using this platform allows for systematic, replicable, and efficient data analysis [10].

In general, the K-Means algorithm works by determining the initial number of clusters (K) and then randomly assigning an initial centroid. Each observation is then allocated to the cluster closest to the centroid using Euclidean distance. This process is carried out iteratively until the centroid position reaches a stable state and does not experience significant changes.

Mathematically, the centroid value for each cluster is determined based on the average of all data in that cluster, which can be formulated as follows:

$$C_i = \frac{1}{M}\sum_{j=1}^{M} x_j$$

where C*i* is the centroid of the i-th cluster, M is the number of data points in the cluster, and *xj* is the value of the j-th observation.

The distance between each data point and the centroid is calculated using the Euclidean distance, which is formulated as:

$$d(x, C_i) = \sqrt{\sum_{k=1}^{n}(x_k - C_{ik})^2}$$

Each data point is then allocated to a cluster with the minimum distance from the centroid, resulting in the most homogeneous data grouping within the cluster and the least heterogeneous between clusters. The main objective of the K-Means algorithm is to minimize the total distance within the cluster (within-cluster sum of squares), which is formulated as follows:

$$\min \sum_{i=1}^{K}\sum_{x\in C_i} d(x, C_i)^2$$

The cluster analysis stages in this study include data exploration, data preprocessing, determining the optimal number of clusters, building a cluster model, cluster visualization, and evaluating the clustering results [11]. In the preprocessing stage, data was normalized using the StandardScaler method to ensure all variables had comparable scales and did not dominate the clustering process.

The optimal number of clusters was determined using the Elbow method and the silhouette coefficient as internal validation techniques. The silhouette value is used to assess the level of cohesion within the cluster and the separation between clusters, ensuring that the selected number of clusters optimally represents the data structure.

After the clusters were formed, the characteristics of each cluster were analyzed by comparing the average values of operational and financial performance indicators [12]. The cluster interpretation results were used to map the position of Batik SMEs into several segments with different levels of organizational performance. This mapping serves as the basis for developing managerial and policy implications, particularly in designing strategies for sustainably improving the performance and competitiveness of Batik SMEs.

To improve clarity and avoid redundancy, the methodological explanation has been streamlined by summarizing standard K-Means procedures and normalization steps while emphasizing aspects directly relevant to the research context. General algorithmic descriptions are presented concisely, allowing the focus to remain on the segmentation framework and its empirical application to Batik SMEs.

## 3. Results and Discussion

### *3.1 Exploratory Data Frame*

The Exploratory Data Analysis (EDA) stage was conducted to gain an initial understanding of the structure, distribution, and variation of Batik SME organizational performance data before clustering using the K-Means method [13]. This exploratory analysis aimed to identify general patterns, value trends, levels of heterogeneity, and potential extreme values for each operational and financial performance indicator.

Based on the descriptive statistics, all variables used in this study had the same number of observations, namely 56 Batik SMEs, indicating no missing data. The minimum and maximum values for all indicators ranged from 1 to 6, consistent with the measurement scale used. This indicates that the data is consistent and worthy of further analysis.

The mean value of the Operational Performance (OPF) indicator indicates that most Batik SMEs have relatively moderate to high levels of operational performance, particularly in productivity, customer satisfaction, product quality, and on-time delivery. Meanwhile, the Financial Performance (FP) indicator also showed a relatively high average value, particularly in aspects of profit monitoring, revenue growth, and investment evaluation [14]. This condition reflects the variation in financial performance levels among Batik SMEs.

The standard deviation for each indicator indicates sufficient data variability, indicating differences in organizational performance characteristics among Batik SMEs. This variation forms an important basis for cluster analysis, as the K-Means method works optimally when the data has a sufficient level of heterogeneity [15]. Furthermore, the quartile values (Q1, median, and Q3) indicate a relatively even distribution of data, without excessive dominance of extreme values.

Through this EDA stage, it can be concluded that the dataset meets the basic assumptions for applying distance-based clustering methods, specifically K-Means [16]. All variables are numeric, have a uniform value range, and exhibit sufficient variation to form meaningful clusters. Therefore, EDA serves as an empirical foundation for the next stages: data preprocessing, determining the optimal number of clusters, and developing a K-Means model for segmenting Batik SMEs based on organizational performance.

| | OPF1 | OPF2 | OPF3 | OPF4 | OPF5 | FP1 | FP2 | FP3 | FP4 | FP5 |
|---|---|---|---|---|---|---|---|---|---|---|
| **count** | 56.000000 | 56.000000 | 56.000000 | 56.000000 | 56.000000 | 56.000000 | 56.000000 | 56.000000 | 56.000000 | 56.000000 |
| **mean** | 5.875000 | 5.750000 | 5.803571 | 5.785714 | 5.803571 | 5.732143 | 5.803571 | 5.678571 | 5.607143 | 5.625000 |
| **std** | 0.333712 | 0.513455 | 0.400892 | 0.414039 | 0.400892 | 0.521922 | 0.443935 | 0.575473 | 0.561769 | 0.589684 |
| **min** | 5.000000 | 4.000000 | 5.000000 | 5.000000 | 5.000000 | 4.000000 | 4.000000 | 4.000000 | 4.000000 | 4.000000 |
| **25%** | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 5.000000 | 5.000000 | 5.000000 |
| **50%** | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 |
| **75%** | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 |
| **max** | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 | 6.000000 |

Figure 1 Data Frame Stage

## *3.2 Data Pre-processing: Normalization*

The data preprocessing stage in this study focused on data normalization to ensure that all organizational performance indicators were on a comparable scale before conducting cluster analysis using the K-Means method. Normalization is a crucial step because the K-Means algorithm is a distance-based method, so variables with a wider range of values have the potential to dominate the clustering process if not standardized first.

The normalization technique used in this study is Z-score standardization, implemented using the StandardScaler function from the scikit-learn library in the Python programming language [17]. This method transforms each variable into a standardized value with a mean of zero and a standard deviation of one. This process allows for fair comparisons between the Operational Performance (OPF) and Financial Performance (FP) indicators used in the study.

Technically, the normalization process is performed by applying StandardScaler to all variables in the data frame. The normalized data is then stored back in a structured data frame to facilitate further analysis, including cluster modeling and evaluation of clustering results [18]. This implementation ensures that the dataset is ready to be used consistently throughout all stages of the analysis.

Table 2.
Standard Scaler Syntax

| **StandardScaler Algorithm (sklearn library in python)** |
|---|
| from sklearn.preprocessing import StandardScaler |
| scale=StandardScaler() |
| std_atr=scale.fit_transform(df) |
| std_atr=pd.DataFrame(std_atr, columns=df.columns) |
| print(std_atr) |

The results of data normalization are visualized using boxplots to compare the data distribution before and after standardization. The visualization shows that the normalized variables have a relatively symmetrical distribution with a consistent spread around the zero value. This indicates that the normalization process has been effective and has successfully reduced the influence of scale differences between variables.

Although several potential outliers were identified in Figure 2, particularly in OPF2, FP1, and FP2, these values remained within the valid Likert scale range (1–6) and represented genuine variability in respondents' perceptions rather than data entry errors. Therefore, no observations were removed to preserve the integrity and representativeness of the sample. To mitigate the potential influence of extreme values on centroid positioning in the K-Means algorithm, Z-score standardization was applied. This transformation reduces scale disparities and minimizes the dominance of relatively extreme observations in the distance calculation process, thereby ensuring more stable and reliable clustering results.

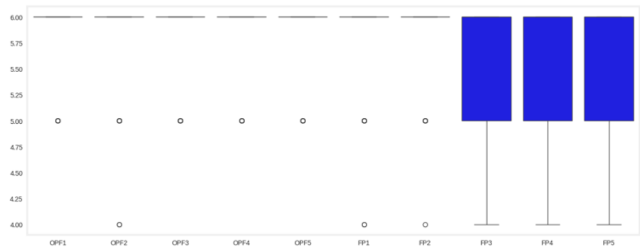


Figure 2 Box Plot Before Normalized

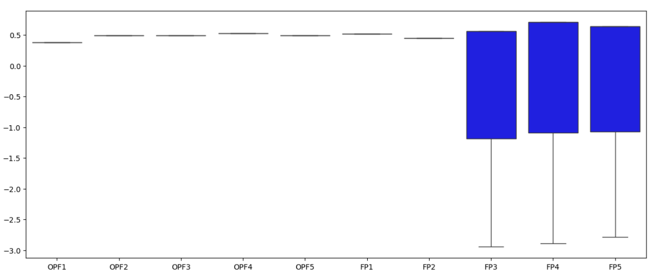


Figure 3 Box Plot After Normalization

By applying Z-score normalization, numerical stability in the clustering process can be improved, resulting in more accurate clustering results and more reliable interpretation. Therefore, this preprocessing stage ensures that the dataset has been systematically prepared and meets the requirements for distance-based cluster analysis, specifically the K-Means algorithm, for segmenting Batik SMEs based on organizational performance.

### *3.3 Model Tuning*

The model tuning stage was conducted to determine the optimal number of clusters that best represent the data structure of Batik SMEs based on organizational performance, including operational and financial performance [19]. Determining the appropriate number of clusters is a crucial step in K-Means analysis, as an inappropriate number of clusters can reduce segmentation clarity and degrade the interpretation of clustering results.

In this study, the Elbow method was used by analyzing the Within-Cluster Sum of Squares (WCSS) values for various numbers of clusters. The Elbow diagram shows the relationship between the number of clusters and the level of compactness of the data within the cluster. The analysis results show that the WCSS value decreases significantly up to a certain number of clusters, then tends to level off with subsequent cluster additions. The point at which the curve changes, known as the elbow point, indicates the number of clusters that most efficiently represents the data without excessively increasing model complexity.

To strengthen the results of the Elbow method, additional evaluations were conducted using the silhouette coefficient as a quantitative measure of cluster quality [20]. The silhouette coefficient measures the degree of cohesion within a cluster and separation between clusters, with higher values indicating a clearer and more well-separated cluster structure. Based on the test results, the highest silhouette value was obtained in the two-cluster configuration (k = 2), indicating that grouping Batik SMEs into two segments provides the most optimal separation.

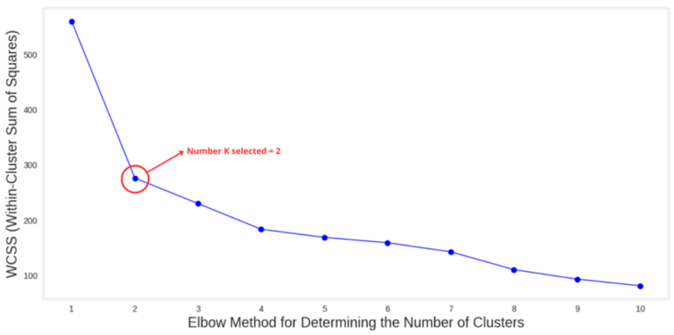

Figure 4 Elbow Diagram

The silhouette coefficient results are consistent with the findings in the Elbow diagram, where the decline in WCSS begins to stabilize after two clusters. The combination of these two methods indicates that increasing the number of clusters beyond two does not significantly improve cluster quality. Therefore, a two-cluster configuration was chosen as the optimal solution because it balances model simplicity and clustering accuracy.

Based on the model tuning results, the Batik SMEs in this study can be classified into two main organizational performance profiles, reflecting differences in operational and financial performance characteristics. Determining the optimal number of clusters serves as the basis for the next stage of analysis, namely cluster interpretation and a discussion of managerial and policy implications based on the results of the organizational performance segmentation.

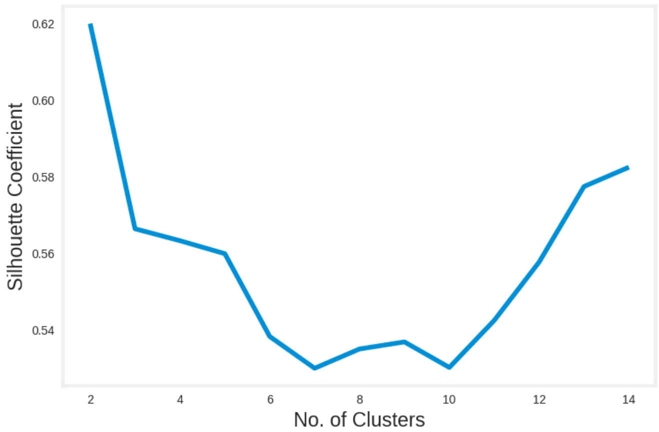

Figure 5 Silhouette Coefficient

Tabel 3.
Silhouette Coefficient

| n_clusters | Silhouette Coefficient |
|---|---|
| 2 | 0.6198966663282094 |
| 3 | 0.5663406529659459 |
| 4 | 0.5632278338799337 |
| 5 | 0.5598390608253085 |
| 6 | 0.5382267637103054 |
| 7 | 0.5299332476915452 |
| 8 | 0.5349640153893515 |
| 9 | 0.5368141944576965 |
| 10 | 0.5301458987676693 |
| 11 | 0.5424119626728424 |
| 12 | 0.5576369079589554 |
| 13 | 0.5773703513375172 |
| 14 | 0.5824049343868066 |

### *3.4 Model Building*

Based on the model tuning results, which indicated that the optimal number of clusters was two, the next stage was to build a final K-Means model using the normalized data. The model was built by applying the K-Means algorithm to a standardized dataset to ensure that each organizational performance indicator contributes equally to the clustering process.

The K-Means algorithm works by allocating each Batik SME to the cluster with the closest Euclidean distance to the centroid, then iteratively updating the centroid position until convergence is achieved. This process produces clusters with a high degree of internal homogeneity and clear separation between clusters, allowing for meaningful interpretation of the resulting segmentation.

The K-Means model results in two main clusters representing the different organizational performance profiles of Batik SMEs. The following table presents the standardized centroid values (Z-scores) for each Operational Performance (OPF) and Financial Performance (FP) indicator in each cluster.

Tabel 4.
Centroid Value Each Variable

| Variable | Cluster 0 | Cluster 1 |
|---|---|---|
| OPF1 | -1.25019018 | 0.37796447 |
| OPF2 | -1.62508142 | 0.49130368 |
| OPF3 | -1.63536685 | 0.49441323 |
| OPF4 | -1.53991773 | 0.46555652 |
| OPF5 | -1.44175048 | 0.43587805 |
| FP1 | -0.96932303 | 0.29305115 |
| FP2 | -0.95227447 | 0.28789693 |
| FP3 | -1.32470523 | 0.40049228 |
| FP4 | -1.09055051 | 0.32970132 |
| FP5 | -0.80622458 | 0.24374231 |

Based on these centroid values, Cluster 0 is characterized by negative values across all indicators, indicating that Batik SMEs in this cluster have relatively low operational and financial performance compared to the population average. Conversely, Cluster 1 exhibits positive centroid values across all indicators, indicating higher organizational performance, both operationally and financially.

The results of this model building confirm the heterogeneity of organizational performance among Batik SMEs, making segmentation based on organizational performance relevant and informative. The two clusters formed provide a clear segmentation framework for understanding the differences in Batik SME profiles and serve as a basis for developing more targeted managerial and policy implications, particularly in efforts to improve the performance and competitiveness of Batik SMEs.

### *3.5 Model Visualization*

The model visualization stage was conducted to provide a clear and intuitive visual representation of the clustering results and to improve the interpretability of the cluster structure generated by the K-Means model [21]. Visualization plays a crucial role in depicting the distribution patterns of Batik SMEs and the degree of separation between clusters, thus supporting a comprehensive evaluation of the effectiveness of the clustering model.

In this study, visualization was used to complement previously conducted quantitative validation, such as the Elbow method and the silhouette coefficient. Through a visual approach, the grouping of Batik SMEs based on similarities in operational and financial performance can be observed more empirically, thereby strengthening confidence in the segmentation results. Visualization also helps identify dominant behavioral patterns between clusters that cannot always be fully captured through numerical indicators alone.

The cluster visualization results are presented as a two-dimensional mapping (2D scatterplot), depicting the relative position of each Batik SME based on normalized organizational performance variables. This visualization shows a clear separation between the two clusters, indicating that the K-Means model is capable of distinguishing Batik SMEs into two groups with significantly different organizational performance characteristics. The relatively tight grouping of data points within each cluster also reflects a good level of internal cohesion.

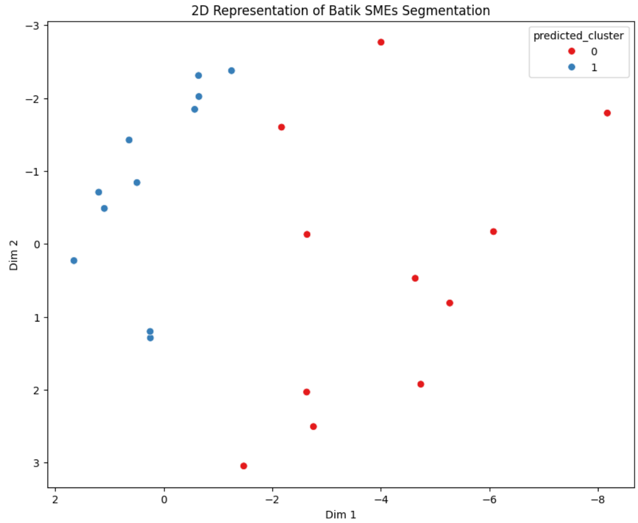

Figure 6 2D Mapping of Batik SMEs

In addition to visual mapping, cluster interpretation was strengthened through centroid analysis for each operational and financial performance indicator. The centroid-based analysis confirmed the existence of two well-defined clusters, each representing distinct patterns of organizational behavior among batik SMEs. The contrasting centroid values between the clusters indicate substantial heterogeneity in organizational performance within the batik SME sector.

The performance differences between clusters can be interpreted through organizational performance theory and the resource-based view (RBV), which emphasize the role of internal capabilities in shaping competitive outcomes. The findings suggest that operational indicators such as customer satisfaction and product quality tend to be more problematic for low-performing SMEs because they require structured processes, quality control systems, and consistent managerial oversight [22], [23]. In contrast, financial indicators are relatively less problematic, as most SMEs maintain basic profit and revenue monitoring practices for business survival. This pattern indicates that operational capability gaps, rather than purely financial constraints, are a key differentiator between low- and high-performing Batik SMEs.

Tabel 5.
Cluster Definition

| Cluster | Average Centroid Value | Cluster Name |
|---|---|---|
| Cluster 0 | -1.2635 | Low Organizational Performance SMEs |
| Cluster 1 | 0.3820 | High Organizational Performance SMEs |

Beyond the numerical distinction, the two clusters differ substantially in managerial and operational characteristics. Cluster 0 is characterized by low customer satisfaction (OPF2) and weak product quality compliance (OPF3), indicating limited quality control systems and less structured customer relationship management. SMEs in this cluster tend to operate with informal managerial practices and minimal performance monitoring. In contrast, Cluster 1 demonstrates stronger operational discipline and financial awareness, particularly reflected in more systematic investment evaluation (FP5). These SMEs typically implement clearer quality standards, monitor customer feedback more consistently, and apply more structured financial decision-making processes. Therefore, the cluster labels represent not only performance levels but also distinct managerial capability profiles.

Tabel 6.
Cluster Result

| Cluster | Objects |
|---|---|
| 0 | 0, 1, 11, 23, 26, 27, 29, 30, 36, 43, 49, 50, 52 |
| 1 | 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 28, 31, 32, 33, 34, 35, 37, 38, 39, 40, 41, 42, 44, 45, 46, 47, 48, 51, 53, 54, 55 |

In general, clusters with positive centroid values for most indicators reflect batik SMEs with higher organizational performance, both operationally and financially. Conversely, clusters with negative centroid values depict batik SMEs with relatively lower organizational performance, which still face various limitations in achieving productivity, service quality, and financial growth. This difference is consistently visualized in the two-dimensional mapping, which shows a fairly clear gap between the clusters.

Tabel 7.
Least Performing Indicator

| Cluster Name | Centroid Velue | Code | Description |
|---|---|---|---|
| Early-Stage Sustainability Batik SMEs | -1.62508142 | OPF2 | Customer satisfaction with products and services |
| | -1.63536685 | OPF3 | Product quality compliance with standards |
| Sustainability-Ready Batik SMEs | 0.24374231 | FP5 | Investment value versus benefits |
| | 0.28789693 | FP2 | Revenue growth as a success indicator |

Based on the least performing indicators identified in each cluster, several targeted strategic recommendations can be proposed. For Cluster 0 (Low Organizational Performance SMEs), improvement efforts should prioritize enhancing product quality standards and customer satisfaction through structured quality control systems, standard operating procedures (SOPs), and customer service training programs. Government support may include technical assistance, certification facilitation, and quality management workshops to strengthen production consistency and service reliability. For Cluster 1 (High Organizational Performance SMEs), strategic focus should shift toward optimizing financial performance, particularly in investment evaluation and revenue growth monitoring. SMEs in this cluster are encouraged to adopt more structured financial planning tools and digital accounting systems, while policymakers may provide access to financial literacy programs and investment advisory services. These differentiated strategies ensure that performance improvement initiatives are aligned with the specific weaknesses of each cluster.

By integrating visual and numerical approaches, this visualization model stage strengthens the robustness of the clustering results. Visualization not only serves as a supporting analytical tool but also as an effective communication tool to explain the segmentation results to non-technical stakeholders. Therefore, the visualization model provides intuitive evidence that the performance-based segmentation of Batik SMEs using K-Means clustering produces a stable, well-separated, and managerially meaningful cluster structure.

Although the sample size of 56 Batik SMEs may be considered relatively modest for clustering analysis with ten indicators, this study adopts an exploratory segmentation approach within a specific regional context. K-Means clustering does not require large samples in the same way as inferential statistical models; however, cluster robustness and generalizability should be interpreted cautiously. To enhance analytical stability, data were standardized, and the optimal number of clusters was validated using both the Elbow method and the silhouette coefficient. The consistency between centroid patterns and visualization results further supports the internal coherence of the clusters. Nevertheless, the findings should be viewed as context-specific, and future research with larger and more diverse samples is recommended to validate and generalize the identified segmentation structure.

## 4. Conclusion

This study aims to segment batik SMEs based on organizational performance using the K-Means clustering approach. Based on primary data analysis from 56 batik SMEs measured through operational and financial performance indicators, the results indicate that the data exhibits sufficient variation and is suitable for analysis using distance-based clustering. The normalization process using StandardScaler successfully equalized the scale of all variables, thereby increasing the reliability of the clustering results. Model tuning results using the Elbow method and the silhouette coefficient indicate that a two-cluster configuration (k = 2) is the optimal number of clusters. The K-Means model produces two groups of batik SMEs with clearly distinct organizational performance characteristics: Low Organizational Performance SMEs and High Organizational Performance SMEs. Cluster visualization and centroid analysis confirm strong cluster separation and heterogeneity in organizational performance among batik SMEs. These findings demonstrate that organizational performance-based segmentation using K-Means clustering is an effective approach to support managerial and policy decision-making, particularly in designing more targeted strategies to improve the performance and competitiveness of batik SMEs.

From a managerial perspective, the segmentation results provide practical guidance for designing targeted intervention strategies. SMEs classified in the low-performance cluster should prioritize strengthening operational capabilities, particularly in quality management, customer satisfaction monitoring, and structured performance evaluation. Meanwhile, high-performance SMEs can focus on strategic expansion, innovation, and market development. For policymakers and SME support institutions, this segmentation framework can serve as a basis for differentiated assistance programs, training initiatives, and performance-based policy design to enhance the competitiveness and sustainability of batik SMEs.

Future research should consider expanding the sample size and geographic coverage to improve the robustness and generalizability of the findings. Further studies may also compare different clustering algorithms or integrate additional variables, such as innovation capability, digital adoption, or organizational culture, to develop a more comprehensive segmentation model. Longitudinal analysis could also provide deeper insights into performance dynamics and cluster transitions over time.

## 5. Reference


[1] A. Enaifoghe, "Enhancing the role and contributions of small and medium enterprises (SMEs) in a globalized and innovative economy: Challenges and opportunities," 2023.

[2] D. A. Rahmat, A. A. Rumanti, M. A. Pulungan, A. S. Rizaldi, and M. Amelia, "Evaluating the role of open innovation and circular economy in enhancing organizational performance: Insights from batik small and medium enterprises in Banyuwangi, Indonesia," *Sustainability*, vol. 16, no. 24, p. 11194, 2024. DOI: https://doi.org/10.3390/su162411194

[3] F. Achmad and I. I. Wiratmadja, "Organizational performance and competitive advantage in SMEs: The role of green innovation and knowledge management," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 11, no. 2, p. 100532, 2025. DOI: https://doi.org/10.1016/j.joitmc.2025.100532

[4] M. Karmagatri, M. H. Widianto, R. K. Widyasari, A. Darisman, H. Hasan, and I. A. Jumbri, "Data-driven strategies for enhancing Indonesian batik export," in *Proc. 2024 IEEE 4th Int. Conf. on ICT in Business Industry & Government (ICTBIG)*, 2024, pp. 1–5. DOI: https://doi.org/10.1109/ICTBIG64922.2024.10911648

[5] S. Naeem, A. Ali, S. Anam, and M. M. Ahmed, "An unsupervised machine learning algorithms: Comprehensive review," *International Journal of Computing and Digital Systems*, 2023.

[6] F. Achmad, I. Zulkarnain, A. A. Rumanti, I. I. Wiratmadja, and S. M. Shuhidan, "Insights from clustering algorithm analysis," in *Proc. 7th Asia Pacific Conf. on Manufacturing Systems and 6th Int. Manufacturing Engineering Conf. (iMEC-APCOMS 2024)*, Melaka, Malaysia, 2025, p. 151.

[7] Z. F. Rosyada, N. U. Handayani, F. S. Minako, and A. Mustikasari, "Strategy to improve the innovation performance of batik industrial cluster," in *Proc. 3rd Asia Pacific Int. Conf. on Industrial Engineering and Operations Management*, Johor Bahru, Malaysia, 2022.

[8] Y. Kustiyahningsih *et al.*, "Integration K-means clustering and AHP for recommendations batik MSMEs," in *E3S Web of Conferences*, vol. 499, p. 01006, 2024. DOI: https://doi.org/10.1051/e3sconf/202449901006

[9] K. Dermayanti and M. Dayat, "Leveraging total quality management to enhance competitiveness of batik MSMEs in Gunting Pasuruan," *Journal of Islamic Economics (JoIE)*, vol. 4, no. 1, pp. 52–63, 2024. DOI: https://doi.org/10.21154/joie.v4i1.9328

[10] M. A. Kabir and M. D. Ahmed, "Python for data analytics: A systematic literature review of tools, techniques, and applications," *Academic Journal on Science, Technology, Engineering & Mathematics Education*, vol. 4, no. 04, pp. 10–69593, 2024. DOI: http://dx.doi.org/10.69593/ajsteme.v4i04.146

[11] I. Shafi, M. Chaudhry, E. C. Montero, E. S. Alvarado, I. de la Torre Diez, M. A. Samad, and I. Ashraf, "A review of approaches for rapid data clustering: Challenges, opportunities, and future directions," *IEEE Access*, vol. 12, pp. 138086–138120, 2024. DOI: http://doi.org/10.1109/ACCESS.2024.3461798

[12] E. Herman, K. E. Zsido, and V. Fenyves, "Cluster analysis with k-mean versus k-medoid in financial performance evaluation," *Applied Sciences*, vol. 12, no. 16, p. 7985, 2022. DOI: https://doi.org/10.3390/app12167985

[13] F. Achmad, I. Zulkarnain, A. A. Rumanti, I. I. Wiratmadja, and S. M. Shuhidan, "Enhancing performance of batik SMEs through cleaner production adoption: Insights from clustering algorithm analysis," in *Asia Pacific Conf. on Manufacturing Systems and Int. Manufacturing Engineering Conf.*, Singapore: Springer Nature Singapore, 2024, pp. 151–163. DOI: https://doi.org/10.1007/978-981-96-4353-0_14

[14] N. N. Vibhakar, K. K. Tripathi, S. Johari, and K. N. Jha, "Identification of significant financial performance indicators for the Indian construction companies," *International Journal of Construction Management*, vol. 23, no. 1, pp. 13–23, 2023. DOI: https://doi.org/10.1080/15623599.2020.1844856

[15] H. Hu, J. Liu, X. Zhang, and M. Fang, "An effective and adaptable K-means algorithm for big data cluster analysis," *Pattern Recognition*, vol. 139, p. 109404, 2023. DOI: https://doi.org/10.1016/j.patcog.2023.109404

[16] M. Chaudhry, I. Shafi, M. Mahnoor, D. L. R. Vargas, E. B. Thompson, and I. Ashraf, "A systematic literature review on identifying patterns using unsupervised clustering algorithms: A data mining perspective," *Symmetry*, vol. 15, no. 9, p. 1679, 2023. DOI: https://doi.org/10.3390/sym15091679

[17] Z. L. Thakker and S. H. Buch, "Effect of feature scaling pre-processing techniques on machine learning algorithms to predict particulate matter concentration for Gandhinagar, Gujarat, India," *Int. J. Sci. Res. Sci. Technol.*, vol. 11, no. 1, pp. 410–419, 2024.

[18] S. Shafi *et al.*, "Differential drought responses in deep and shallow-rooted rice genotypes: Enzymatic and non-enzymatic insights," *Plant Physiology Reports*, vol. 29, no. 2, pp. 294–307, 2024. DOI: https://doi.org/10.1007/s40502-024-00788-2

[19] F. Achmad, A. Kurniawati, F. H. Hasanudin, and I. I. Wiratmadja, "Mapping performance of batik SMEs: A perspective on green innovation, green culture, and green employee behavior through clustering algorithm," *Journal Industrial Servicess*, vol. 11, no. 2, 2025. DOI: http://dx.doi.org/10.62870/jiss.v11i2.33093

[20] T. Sumallika, V. Alekya, P. V. M. Raju, M. R. Rao, D. G. Shiney, and M. V. Sudha, "Exploring optimal cluster quality in health care data (HCD): Comparative analysis utilizing k-means elbow and silhouette analysis," *International Journal of Chemical and Biochemical Sciences (IJCBS)*, vol. 25, no. 16, pp. 48–60, 2024.

[21] L. Hu, M. Jiang, J. Dong, X. Liu, and Z. He, "Interpretable clustering: A survey," *ACM Computing Surveys*, 2024. DOI: https://doi.org/10.1145/3789495

[22] F. Achmad, H. D. Anggana, A. A. Rumanti, M. D. Rial, A. P. Suryaneda, M. A. Sena, and R. Aditya, "Development of SMEs promotion media through SECI-based digital catalog in Bandung Regency," *Jurnal Inkofar*, vol. 9, no. 1, 2025. DOI: https://doi.org/10.46846/jurnalinkofar.v9i1.471

[23] F. Achmad and I. I. Wiratmadja, "Driving sustainable performance in SMEs through frugal innovation: The nexus of sustainable leadership, knowledge management, and dynamic capabilities," *IEEE Access*, vol. 12, pp. 103329–103347, 2024. DOI: https://doi.org/10.1109/ACCESS.2024.3433474