Equivalent: Jurnal Ilmiah Sosial Teknik
Volume 8, Issue 3, 1324-1335
e_ISSN: 2775-0833
https://jurnalequivalent.id/index.php/jequi/index
DOI: doi.org/10.59261/jequi.v8i3.397

Integration of RFM and Dynamic Time Warping for Customer Segmentation in MSMEs Digital Printing

Erna Erviana*
Universitas Bina Nusantara, Indonesia

Tuga Mauritsius
Universitas Bina Nusantara, Indonesia

***Corresponding author:**
Erna Erviana, Universitas Bina Nusantara, Indonesia.
✉erna.erviana@binus.ac.id

Article Info:
Article history:
Received: July 03, 2026
Revised: July 28, 2026
Accepted: August 10, 2026
Available Online: September 01, 2026

Keywords:
DBSCAN; Customer Segmentation; Dynamic Time Warping; K-Means; Monetary Frequency Recency.

Abstract

Background: Customer segmentation is an important strategy for Micro, Small, and Medium Enterprises (MSMEs) to improve marketing effectiveness and maintain customer loyalty. However, conventional approaches based on static data are often unable to capture the dynamic nature of customer behavior.
Objective: This study aims to develop a more comprehensive customer segmentation approach by integrating static analysis based on Recency, Frequency, and Monetary (RFM) with dynamic analysis using Dynamic Time Warping (DTW) through a Hybrid Integration Matrix for *Agatek Print* customers.
Method: This study uses transaction data collected from May 2024 to October 2025, comprising 8,808 transactions from 2,194 customers. The data were analyzed using the K-Means and DBSCAN algorithms for static segmentation and Dynamic Time Warping (DTW) for time-series pattern analysis. Model performance was evaluated using the Silhouette Score.
Results: The results show K-Means outperforms DBSCAN, achieving a Silhouette Score of 0.5491 compared with 0.0103, indicating a more stable clustering structure. The DTW analysis successfully identifies dynamic transaction behavior patterns that are not captured by static approaches, as evidenced by four DTW clusters (2,135, 2, 1, and 56 customers) that do not align with the static RFM groups, including project-based and small-scale retail transaction patterns. The integration of both approaches reveals that most customers belong to segments characterized by low customer value and inactive purchasing behavior combined with project-based transaction patterns.
Conclusion: The hybrid RFM–DTW approach provides a more comprehensive understanding of customer behavior and can serve as a foundation for developing more adaptive marketing strategies in MSME contexts.

To cite this article: Erviana, E. & Mauritsius, T. (2026). Integration of RFM and Dynamic Time Warping for Customer Segmentation in MSMEs Digital Printing. *Equivalent: Jurnal Ilmiah Sosial Teknik*, *8*(3), 1324-1345. https://doi.org/10.59261/jequi.v8i3.397

INTRODUCTION

Customer retention and churn are immediate operational challenges for Micro, Small, and Medium Enterprises (MSMEs) in the digital printing sector, and understanding these dynamics is the central motivation for the customer segmentation approach developed in this study. Micro, Small, and Medium Enterprises (MSMEs) are fundamental pillars of national economic stability. Amid accelerating digitalization, the digital printing industry has emerged as one of the most dynamic sectors. As reviewed by Cioppi et al. (2023), digital transformation has fundamentally changed the marketing landscape, compelling businesses to adopt technology-based approaches to understand consumer interactions more dynamically. Market reports project that this industry in Indonesia will experience a compound annual growth rate (CAGR) of 5.11% during the 2025–2029 period. Technically, digital printing technology enables the printing process directly from digital data onto various media without requiring plate production, as in conventional offset printing, thereby improving production efficiency, increasing personalization flexibility, and reducing production time compared with traditional printing systems (Nagib & Dewanti, 2025). However, this rapid growth has created a competitive paradox: expanding market opportunities are accompanied by increasingly intense competition, requiring businesses to adopt more effective customer-retention strategies to maintain their market share.

Despite the promising market potential, the reality on the ground shows that many businesses in this sector face the dual challenge of revenue volatility and customer churn. This phenomenon is often described as the leaky bucket effect, in which customer losses consistently outpace the acquisition of new customers. This condition necessitates a shift in marketing strategy from a product-centric approach to a customer-centric approach, in which a deep understanding of customer behavior becomes the key source of competitive advantage. In this context, customer segmentation is no longer merely an administrative classification tool but a strategic instrument for efficiently allocating resources and building long-term customer loyalty. This retention pressure is not unique to Agatek Print; similar churn-driven revenue instability has been documented across Indonesian service businesses, where predictive churn analysis has been used to strengthen customer-retention decision-making (Admanegara & Handayani, 2024; Mamun, 2025; Sugumaran, 2025), indicating that the leaky bucket pattern observed here reflects a broader national phenomenon rather than an isolated case.

Empirical observations based on historical transactional data from a representative digital printing MSME revealed a chronic leaky bucket phenomenon. Analysis of data covering the period from October 2024 to September 2025 shows an alarming escalation, with the number of churned customers increasing dramatically from 43 at the beginning of the observation period to 234 at its conclusion. This increase of more than 400% over 12 months was accompanied by sharp fluctuations in monthly revenue. These findings indicate the failure of conventional one-size-fits-all marketing strategies, in which businesses fail to identify changes in customer behavior at an early stage and instead rely heavily on costly new-customer acquisition to compensate for revenue leakage.

The research landscape on transaction-driven customer segmentation has expanded rapidly, with the majority of studies focusing on the use of the Recency, Frequency, Monetary (RFM) model to assess customer value. RFM remains a fundamental framework for purchasing behavior analysis because of its ability to quantitatively capture customer interaction patterns. This approach has proven highly effective in extracting critical information from large transactional databases to identify customer loyalty and profitability without the complexity of incorporating demographic variables (Christy et al., 2021; Vianna Filho et al., 2026). The most dominant approach in the recent literature involves the use of static clustering algorithms to segment customers based on a single snapshot of customer value. For example, Kasem et al. (2024) demonstrated the effectiveness of the K-Means algorithm in generating actionable customer clusters for re-engagement and loyalty strategies. Similarly, Ullah et al. (2023) broadened the methodological evaluation by comparing algorithms such as K-Means, Agglomerative Hierarchical Clustering, and DBSCAN, concluding that majority-voting techniques based on K-Means often produce the greatest cluster stability for general retail datasets (Ullah et al., 2023). Methodological innovation in this area continues, as demonstrated by Kumar (2025), who developed Bisecting K-Means to address the initialization weaknesses of the standard K-Means algorithm. Likewise, Duan and Song (2025) proposed a depth-weighted K-Means algorithm to improve the accuracy of customer cluster identification in support of more adaptive marketing strategies.

Comparatively, the effectiveness of clustering algorithms is largely determined by the characteristics of the underlying data. K-Means clustering has long been recognized as an industry standard in database marketing because of its computational efficiency and its ability to maximize intra-cluster similarity through Euclidean distance optimization. Wilbert et al. (2023) confirmed that K-Means demonstrates greater stability and reliability than the K-medoids algorithm in external stability analyses, while Anitha and Patil (2022) highlighted its widespread adoption as an unsupervised learning technique for identifying customer purchasing patterns. However, K-Means has several fundamental limitations, including the assumption of spherical cluster structures, sensitivity to outliers, and the requirement to specify the number of clusters (k) in advance. Alternatively, DBSCAN provides a density-based clustering approach capable of identifying clusters with complex shapes while explicitly handling noise without requiring a predefined number of clusters (Bushra & Yi, 2021; Wang et al., 2026).

Although these static clustering methods successfully identify customer profiles at a single point in time, they leave a fundamental gap by failing to capture the evolving dynamics of customer behavior over time. Static approaches cannot explain how customers transition to a particular status or how their transaction patterns evolve throughout the customer lifecycle. This limitation is particularly important in the context of dynamic customer churn. Efforts to overcome this limitation have emerged through time-series-based approaches, particularly the application of Dynamic Time Warping (DTW), which has been shown to be more robust to temporal variation than conventional distance metrics. Incorporating the temporal dimension into customer segmentation has been demonstrated to significantly improve cluster validity compared with conventional RFM models (Vera & Angulo, 2023).

In line with these developments, Wang et al. (2024) proposed the DCSLM model, which integrates customer satisfaction with DTW-based distance measurement to analyze time-series data. However, purely time-series-based approaches often involve high computational complexity and are more difficult to interpret from a business perspective than the traditional RFM model. Conversely, other hybrid approaches, such as that proposed by Deldadehasl et al. (2025), integrate data mining with Multi-Criteria Decision-Making (MCDM) but still rely on static weighting schemes without accounting for temporal dynamics. Critically, although Wang et al. (2024) incorporate temporal dynamics through DTW-based distance measurement, their DCSLM model does not explicitly benchmark the resulting segments against conventional static clustering algorithms such as K-Means or DBSCAN. Consequently, it remains unclear how much additional segmentation value the dynamic component contributes beyond a purely static baseline. This comparative and integrative gap is precisely what the present study addresses through the proposed Hybrid Integration Matrix.

In addition to methodological limitations, important contextual gaps remain in the literature. Most previous studies have focused on general retail, global e-commerce, or industry-specific datasets such as hospitality. The unique characteristics of the digital printing industry including highly volatile order patterns, persistent leaky bucket behavior, and substantial revenue fluctuations have received limited scholarly attention. Furthermore, the high likelihood of outliers caused by incidental orders or order cancellations underscores the importance of evaluating robust clustering techniques, including density-based algorithms such as DBSCAN, as comparative methods. Collectively, the studies reviewed demonstrate considerable methodological sophistication in isolation; however, none simultaneously benchmarks static and dynamic clustering approaches within a unified integrative framework or evaluates the practical trade-off between computational complexity and segmentation richness. This unresolved gap, rather than merely the absence of prior research, provides the principal motivation for the integrative approach proposed in this study.

Based on the foregoing review, a clear gap exists in the development of an integration framework capable of comprehensively combining the strengths of static and dynamic customer analysis. Accordingly, this study addresses that gap by proposing a hybrid customer segmentation approach based on the Hybrid Integration Matrix. Unlike a simple side-by-side comparison of two clustering outputs, the Hybrid Integration Matrix applies an explicit cross-tabulation rule that maps each customer's static RFM-based cluster to its corresponding dynamic DTW-based trajectory cluster, thereby generating a joint segmentation label that neither method can produce independently. This rule-based integration, rather than the application of DTW alone, represents the principal methodological contribution of the study. The proposed framework not only rigorously compares the performance of static clustering algorithms (K-Means and DBSCAN) but also integrates them with dynamic trend analysis based on Dynamic Time Warping (DTW). Methodologically, this integration enables the simultaneous identification of customers' current economic value and the characterization of their transaction trajectories over time. Practically, the resulting segmentation framework is expected to support the development of more precise and context-specific customer-retention strategies for digital printing MSMEs.

The scientific contribution of this study lies in the development of an integrative customer segmentation framework that formally combines static value-based analysis (RFM with K-Means and DBSCAN) and dynamic behavioral-pattern analysis (Dynamic Time Warping [DTW]) through a rule-based Hybrid Integration Matrix. This framework extends the customer segmentation literature beyond conventional snapshot-based clustering toward a trajectory-aware segmentation approach suitable for transactional MSME data. The empirical setting of this research is Agatek Print, a digital printing MSME whose transaction records exhibit fluctuating order volumes and high customer-churn characteristics that are representative of the sector, making it an appropriate case for evaluating the proposed hybrid framework before broader application across other digital printing businesses.

METHOD

This study employed a quantitative data mining approach based on the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework as its primary methodology. The selection of the CRISP-DM framework was based on its ability to provide a structured and systematic workflow that connected data analysis with business needs, particularly in the context of customer segmentation for digital printing MSMEs. This approach enabled the analytical results to be not only computationally valid but also strategically applicable to business decision-making. Figure 1 illustrates the CRISP-DM workflow adopted in this study.

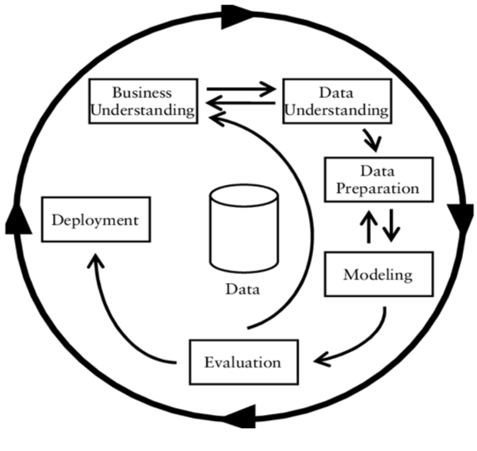

Figure 1. *CRISP-DM Workflow*

The object of this research was the historical transaction data of Agatek Print MSMEs collected from May 2024 to October 2025, comprising 10,000 transaction records. These data included customer identification variables, transaction frequency, and transaction values, which were used to understand customer behavior patterns within the context of the digital printing industry, characterized by fluctuating demand and high customer-churn rates. The research process began with the business understanding and data understanding phases, which focused on identifying customer-retention challenges and conducting an initial exploration of the transaction data structure to determine an appropriate analytical strategy. The May 2024–October 2025 observation period was selected because it represented the most recent 18-month interval for which complete and consistently recorded transaction logs were available in the Agatek Print Enterprise Resource Planning (ERP) system. This period was sufficiently long to capture seasonal fluctuations in customer orders while minimizing missing-data bias associated with older records. Because the dataset was obtained from a single digital printing MSME, the resulting customer segments and cluster profiles should be interpreted as case-specific rather than automatically generalizable to other businesses or industries. Accordingly, this scope limitation was acknowledged as a methodological constraint of the study.

During the data preparation stage, the raw transaction data were transformed into feature representations to support the proposed hybrid customer segmentation approach. This transformation included the extraction of Recency, Frequency, and Monetary (RFM) features as a static representation of customer behavior, where Recency was calculated as the elapsed time since the most recent transaction, Frequency represented the total number of transactions, and Monetary represented the cumulative value of customer transactions. In addition, the transaction records were transformed into time-series sequences to capture the temporal dynamics of customer behavior. This stage also included data cleansing through normalization and the treatment of missing values to ensure data quality before the modeling stage.

The modeling stage was conducted using two complementary approaches: static clustering and dynamic clustering. In the static clustering stage, two algorithms, K-Means and DBSCAN, were employed to compare customer segmentation performance. K-Means was used to partition the data by minimizing the Within-Cluster Sum of Squares (WCSS) objective function, thereby maximizing intra-cluster similarity while minimizing inter-cluster similarity (Tabianan et al., 2022), as expressed in the following equation:

$$J = \sum_{j=1}^{K} \sum_{x_i \in C_j} |x_i - c_j|^2 \qquad (1)$$

Meanwhile, DBSCAN was employed as a comparative algorithm because of its ability to handle datasets with complex cluster structures and identify outliers (noise) through the epsilon (ε) and minimum points (MinPts) parameters. Unlike K-Means, DBSCAN formed clusters based on data density without requiring the number of clusters to be specified in advance (Bushra & Yi, 2021; Wang et al., 2026).

During the dynamic clustering stage, the Dynamic Time Warping (DTW) approach was employed to measure the similarity between customer transaction time series. DTW enabled the alignment of two time series with different lengths or varying temporal speeds, allowing it to identify behavioral patterns such as increasing, stable, or declining transaction trends more accurately than conventional distance metrics. This approach was used to capture temporal customer behavior that could not be adequately represented by static clustering models.

The results of the static and dynamic clustering processes were subsequently integrated using a Hybrid Integration Matrix. This matrix mapped the relationship between customers' economic value (derived from RFM-based K-Means or DBSCAN clustering) and their transaction trajectory patterns (derived from DTW analysis), resulting in a more comprehensive and multidimensional customer segmentation framework.

Model performance was evaluated using the Silhouette Score, which measures clustering quality based on the degree of cohesion within clusters and separation between clusters. The Silhouette Score was used to assess how well each observation fit within its assigned cluster compared with neighboring clusters, thereby providing an overall measure of clustering validity. This evaluation metric has been shown to be particularly effective for determining the optimal number of clusters in partition-based algorithms such as K-Means (Hassan et al., 2022).

The final stage of the research was deployment, during which the results of the Hybrid Integration Matrix were translated into personalized marketing strategy recommendations. The primary focus was on customer-retention strategies, particularly for high-value customer segments exhibiting declining transaction activity, with the objective of reducing churn risk and improving the efficiency of marketing resource allocation.

This study did not involve human or animal participants directly and therefore did not require ethical approval. Nevertheless, all transaction data used in the analysis were anonymized and processed in accordance with the principles of data confidentiality and applicable research ethics standards.

RESULTS AND DISCUSSION

Results

This study processed Agatek Print's MSME customer transaction data using the CRISP-DM approach with two data representations, namely RFM-based static data and time-series-based dynamic data. The analyzed dataset covers the period May 2024 to October 2025 to support the customer segmentation process through static and dynamic clustering methods . The results of this section are presented based on the preprocessing stages, determination of the number of clusters, clustering results, and model evaluation.

Pre-processing and Construction Features

The research data is sourced from the Agatek Print ERP system for the period of May 2024 – October 2025 with a total of 10,002 transactions. After data cleansing (deletion of anonymous customers and merging of duplicate identities), a valid dataset of 8,808 transaction lines from 2,194 unique customers was obtained.

Features are constructed into two main models:

1) RFM (Static) model: consists of Recency, Frequency, and Monetary with a reference date of October 31, 2025.
2) Time-Series Model: uses weekly resampling to represent the sequence of customer transaction activity based on time.

RFM data is normalized using a StandardScaler (Z-Score) to equalize the scale between variables.

Optimal Cluster Determination

The determination of the number of clusters (k) is carried out using the Elbow Method based on the Within-Cluster Sum of Squares (WCSS) value.

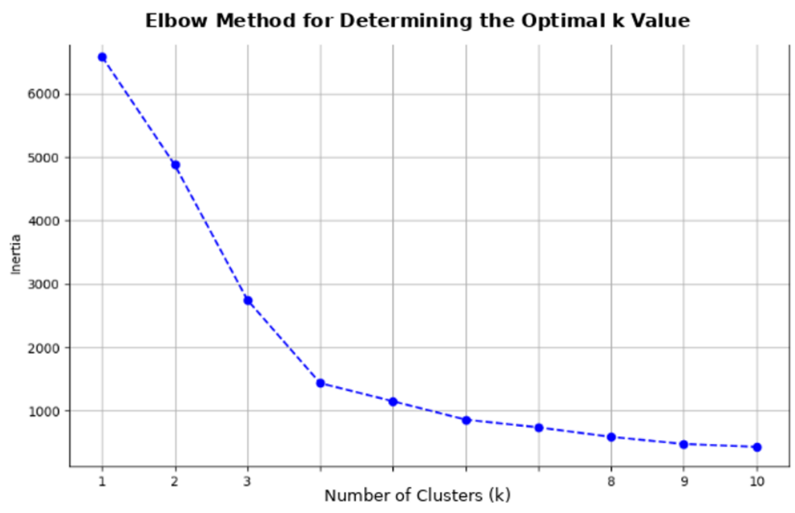

Figure 2. *Elbow Method for RFM Models*

The results show that the optimal value is at k = 4, where after that point the decrease in WCSS is not significant.

Comparison of *K-Means & DBSCAN Static* Clustering

K-Means Clustering Results

The K-Means algorithm is applied with the number of clusters k = 4 based on the results of the Elbow Method. The grouping results show that Cluster 0 consists of 892 customers, Cluster 1 consists of 1,269 customers, Cluster 2 consists of 1 customer, and Cluster 3 consists of 32 customers. The distribution is summarized in Table 1 which displays the numerical characteristics of each cluster based on the variables Recency, Frequency, and Monetary. It is worth noting that Cluster 2 contains only a single customer with an extreme Monetary value and Cluster 3 contains just 32 customers, a marked imbalance relative to Clusters 0 and 1; this pattern is consistent with the presence of a small number of high-value outlier transactions in the dataset rather than a genuine, well-populated segment, and it is treated as an outlier-driven cluster rather than a representative behavioral group in the interpretation that follows.

Table 1

Profile Distribution Per K-Means Cluster

Cluster	Recency	Frequency	Monetary (IDR)
0	324	1	100.000
1	87	1	126.000
2	0	106	114.977.060
3	29	33	4.316.685

DBSCAN Clustering Results

The DBSCAN algorithm was run using the epsilon parameter of 0.2 and MinPts of 5. The grouping results show that Cluster -1 (Noise/ Outlier) consists of 169 customers, Cluster 0 consists of 2,004 customers, Cluster 1 consists of 17 customers, and Cluster 2 consists of 4 customers. The distribution of the Clustering results is presented in Table 2 which contains the characteristics of each cluster based on the Recency, Frequency, and Monetary variables.

Table 2

DBSCAN Profile Distribution Per Cluster

Cluster	Recency	Frequency	Monetary (IDR)
-1	96	9	1.665.000
0	173	1	104.000
1	541	1	72.800
2	316	2	1.311.000

Model Evaluation

Model evaluation was carried out using Silhouette Score. The test results showed that the K-Means method obtained a Silhouette Score of 0.5491, while the DBSCAN method obtained a value of 0.0103. This low Silhouette Score was obtained using a single parameter setting (ε = 0.2, MinPts = 5) and should therefore be interpreted with caution: it may reflect either a genuine mismatch between DBSCAN's density-based assumptions and this dataset's structure, or simply a sub-optimal choice of ε and MinPts, since the parameter space was not exhaustively explored in this study. Consequently, the conclusion that DBSCAN is unsuitable for this dataset is treated as provisional, and a systematic ε–MinPts sensitivity search is recommended before ruling out density-based clustering altogether.

Dynamic Clustering (DTW) Results

Dynamic Time Warping (DTW) analysis is used to segment customers based on similarity in transaction sequence patterns. The grouping results show that Cluster 0 consists of 2,135 customers, Cluster 1 consists of 2 customers, Cluster 2 consists of 1 customer, and Cluster 3 consists of 56 customers. This is shown in figure 3.

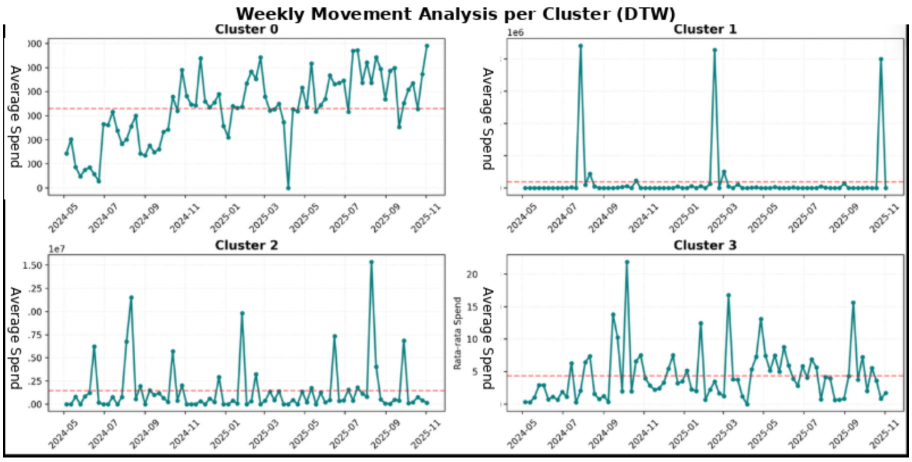

Figure 3. *DTW Cluster Deployment Representation*

Integrated Segmentation Analytics

Static Clustering (RFM) and Dynamic Clustering (DTW) results are combined into the Integration Matrix to see the distribution of customers between the two models. The merging results show the distribution of the number of customers in each combination of static and dynamic clusters, as summarized in Table 3. Compared with a conventional RFM segmentation alone, the added value of the Integration Matrix lies in its ability to split each static RFM cluster into behaviorally distinct sub-groups: for example, Table 3 shows that customers within the same static cluster are distributed across multiple DTW trajectory clusters, revealing transaction-pattern differences (steady, spiking, erratic, or pulsing behavior) that a static RFM label alone would treat as a single homogeneous segment. This finer-grained, two-dimensional labeling is what enables the more targeted retention actions described in Table 4, rather than the value distribution shown in Table 3 alone.

Table 3
Customer Distribution on the Integration Matrix

Dynamic Clustering/ Static	Cluster 0 (Lost)	Cluster 1 (Low-Value)	Cluster 2 (Champions)	Cluster 3 (Loyal)
Cluster 0 (Steady Upward)	885	0	0	0
Cluster 1 (Massive Spikers)	1266	0	1	0
Cluster 2 (Erratic High Value)	1	1	2	1
Cluster 3 (Small Scale Pulsing)	36	0	0	0

Discussion

The results of this study show that the integration of static segmentation based on Recency, Frequency, and Monetary (RFM) analysis with dynamic analysis based on Dynamic Time Warping (DTW) provides complementary customer insights and extends customer interpretation by incorporating temporal behavioral patterns that a single static approach cannot capture. Because this study does not present a quantitative comparison of segmentation quality before and after integration, this complementary framing is used in place of a stronger claim of superiority. Conceptually, these findings confirm that customer representation cannot rely solely on aggregate conditions at a single point in time but must also account for the dynamics of transaction behavior over time. Thus, customer segmentation in the context of micro, small, and medium-sized enterprises (MSMEs) is not static but rather exhibits an evolutionary character influenced by repetitive interaction patterns.

The results of the static clustering analysis showed that the K-means algorithm performed more consistently than DBSCAN based on the Silhouette Score. These findings are consistent with the study by Ufeli et al. (2025), which demonstrated the robustness of the K-means algorithm in identifying well-defined group partitions because of its ability to minimize inter-cluster ambiguity through within-cluster variance optimization. Theoretically, this outcome can be explained by the structure of the RFM data, which tends to form compact clusters that align well with the assumptions underlying K-means, namely minimizing the distance between observations and their assigned centroids. In contrast, the density-based DBSCAN algorithm is highly sensitive to the selection of the epsilon (ε) parameter when applied to unevenly distributed data, resulting in a substantial number of observations being classified as noise rather than assigned to meaningful clusters. In the context of MSME transaction data, which comprise active, inactive, and sporadic customers, centroid-based approaches such as K-means tend to produce more stable and interpretable customer segmentation.

In contrast to the static approach, the use of DTW in the dynamic analysis adds an important dimension to understanding customer similarity by comparing the shapes of transaction patterns rather than their absolute values. This finding suggests that customers with substantially different transaction values may still exhibit similar temporal purchasing patterns. Consequently, DTW expands the analytical perspective from merely evaluating customer value to examining the evolution of transaction behavior over time. This perspective is particularly relevant in the context of modern Customer Relationship Management (CRM), which emphasizes not only identifying high-value customers but also understanding how their behavioral patterns evolve over time.

The integration of the static and dynamic clustering results produces a multidimensional representation of customers, in which each customer is characterized according to both economic value and transaction movement patterns. To clarify the structure of the integrated results, Table 4 presents the combined characteristics of the static and dynamic segmentation based on customer transaction behavior patterns.

Table4 4

Characteristics of Static and Dynamic Clustering Integration Matrix

Dynamic Clustering/ Static	Cluster 0 (Lost)	Cluster 1 (Low-Value)	Cluster 2 (Champions)	Cluster 3 (Loyal)
Cluster 0 (Steady Upward)	Reactivation Target: Loyal former customers whose patterns used to be stable; it's worth figuring out why they quit.	Growing Potential: New or small customers who are starting to show regular loyalty.	The Dream: The most valuable customer who continues to grow consistently.	The Backbone: Loyal customers who are the foundation of the company's regular revenue.
Cluster 1 (Massive Spikers)	One-Off Projects: Project customers who have completed their contracts and are not coming back.	Opportunistic: Big shopping only when there is a big deal or urgent need, and then disappears.	Whale Accounts: Large corporate customers whose last transaction has just occurred (High Impact).	Occasional Partners: Customers who spend large on a regular basis but are not very close.
Cluster 2 (Erratic High Value)	Vanished VIPs: Used to be a lot of random big shopping, but now it's moved to competitors.	High-Risk Casuals: Shop big once in a while but have no brand loyalty.	Top Spenders: A major contributor to active cash flow, even if their spending schedule is unpredictable.	Unpredictable Fans: Loyal customers who are quite loyal but don't have a fixed shopping pattern.
Cluster 3 (Small Scale Pulsing)	Former Retailers: Small businesses that may have closed or no longer take supplies from you.	Micro-Buyers: Individuals who shop small only when there is a specific need (pulsing).	Active Micro-Sellers: SMEs that are very active in transacting even in small but consistent nominals.	Stable Retailers: Loyal customers who keep their stock with regular small-scale shopping.

This integration demonstrates that customer segmentation cannot be reduced to a single dimension but should instead be understood as a combination of current customer value and behavioral trajectory. This approach extends the classical RFM framework by incorporating a temporal dimension based on pattern similarity, resulting in a segmentation structure that is more adaptive to changes in customer behavior.

These findings are consistent with the research of Kasem et al. (2024), which demonstrated the effectiveness of RFM-based customer segmentation in identifying high-value customer groups. However, unlike their relatively static approach, the present study extends the analytical framework by incorporating the temporal dimension through Dynamic Time Warping (DTW). This approach also complements the findings of Ullah et al. (2023), who emphasized the importance of selecting clustering algorithms based on data characteristics, by demonstrating that algorithm performance is determined not only by static data structures but also by the dynamics of customer behavior over time.

Beyond this alignment with Kasem et al. (2024) and Ullah et al. (2023), it is important to consider why the digital printing context examined in this study may differ from the general retail settings investigated in previous research. Digital printing MSMEs such as Agatek Print combine recurring small-scale retail orders with irregular, project-based transactions (e.g., large one-off print jobs), creating a mixed order structure that is less common in general retail or e-commerce datasets. This heterogeneity provides a plausible theoretical explanation for why a purely static RFM perspective may underrepresent customers whose spending patterns are irregular but who are not necessarily disloyal, and why a temporal, trajectory-based perspective contributes proportionally greater value in this sector than in more homogeneous retail environments. This industry-specific rationale, rather than the general argument regarding clustering algorithms alone, supports treating static and dynamic analyses as complementary rather than substitutable in the digital printing context.

In terms of practical implications, this approach provides a stronger foundation for developing Customer Relationship Management (CRM) strategies for MSMEs, particularly in identifying customers who are not only valuable at present but also exhibit growth potential or churn risk based on the dynamics of their transaction histories. Consequently, the resulting segmentation can serve as a more adaptive basis for managerial decision-making in response to evolving market behavior.

However, this study has several limitations. First, its scope is confined to the digital printing industry, which possesses distinctive transaction characteristics. Second, the implementation of Dynamic Time Warping (DTW), which has relatively high computational complexity, may present challenges when applied to larger datasets or real-time systems. Therefore, further research is needed to evaluate the generalizability of this approach across different industry sectors while improving computational efficiency in time-series-based analyses. A practical direction for reducing this computational burden is to replace the exact $O(N^2)$ DTW computation with approximate or constrained variants, such as FastDTW or the Sakoe–Chiba band constraint, both of which have been shown to substantially reduce the computational time required for time-series clustering while maintaining comparable accuracy (Paparrizos & Bogireddy, 2025; Paparrizos & Gravano, 2017), thereby making the proposed hybrid approach more feasible for large-scale or near-real-time deployment.

CONCLUSION

This study concludes that integrating static customer segmentation methods (RFM combined with K-Means and DBSCAN) with a dynamic time-series approach (Dynamic Time Warping [DTW]) provides a complementary, multidimensional understanding of customers at Agatek Print MSME. K-Means proved to be more stable than DBSCAN based on the Silhouette Score, producing interpretable segments representing inactive, low-value, loyal, and high-value customer groups. Meanwhile, DTW independently identified temporal behavioral patterns (such as project-based and pulsing transaction rhythms) that could not be captured through static analysis alone. Together, these complementary perspectives form the proposed Integration Matrix, extending customer segmentation beyond a single snapshot of customer value toward a framework that also captures the trajectory of customer behavior over time.

This integrated framework provides a practical foundation for Customer Relationship Management (CRM) decisions, including customer reactivation, transaction-value enhancement, and loyalty development.

Nevertheless, this study has several limitations. The dataset was obtained from a single digital printing MSME, and therefore the findings should be generalized to other industries with caution. Furthermore, the computational complexity of Dynamic Time Warping (DTW) remains a challenge for large-scale or real-time implementation. Future research should evaluate the applicability of this integrated framework across multiple industry sectors, adopt computationally efficient DTW variants such as FastDTW or the Sakoe–Chiba band constraint, and incorporate additional non-transactional variables, such as product categories, to further improve the quality and practical usefulness of customer segmentation.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to Universitas Pembangunan Nasional Veteran Jakarta for providing academic support, research facilities, and a conducive environment throughout the completion of this study. Appreciation is also extended to scholars, researchers, and relevant stakeholders who contributed valuable insights and references related to cybersecurity, securitization discourse, and security studies in Indonesia.

AUTHOR CONTRIBUTION STATEMENT

Kamil Ghiffary A contributed to the conceptualization, research design, literature review, data analysis, and preparation of the original manuscript. Hartanto contributed to theoretical framework development, methodology validation, critical analysis, and manuscript review. Ivandra Solihin contributed to data interpretation, refinement of arguments, and final manuscript evaluation. All authors collaboratively contributed to the development of the research, reviewed and approved the final version of the manuscript, and take responsibility for the integrity, accuracy, and accountability of all aspects of the work.

REFERENCES

Admanegara, R. C., & Handayani, W. (2024). Customer Churn Analysis Using Machine Learning to Improve Customer Retention on Vissie Net. *International Journal of Scientific Research and Management (IJSRM)*, *12*(9), 7379–7387. https://doi.org/10.18535/ijsrm/v12i09.em05

Anitha, P., & Patil, M. M. (2022). RFM model for customer purchase behavior using K-Means algorithm. *Journal of King Saud University - Computer and Information Sciences*, *34*(5), 1785–1792. https://doi.org/10.1016/j.jksuci.2019.12.011

Bushra, A. A., & Yi, G. (2021). Comparative Analysis Review of Pioneering DBSCAN and Successive Density-Based Clustering Algorithms. *IEEE Access*, *9*, 87918–87935. https://doi.org/10.1109/ACCESS.2021.3089036

Christy, A. J., Umamakeswari, A., Priyatharsini, L., & Neyaa, A. (2021). RFM ranking – An effective approach to customer segmentation. *Journal of King Saud University - Computer and Information Sciences*, *33*(10), 1251–1257. https://doi.org/10.1016/j.jksuci.2018.09.004

Cioppi, M., Curina, I., Francioni, B., & Savelli, E. (2023). Digital transformation and marketing: a systematic and thematic literature review. *Italian Journal of Marketing*, *2023*(2), 207–288. https://doi.org/10.1007/s43039-023-00067-2

Deldadehasl, M., Hajian Karahroodi, H., & Nekah, P. H. (2025). Customer Clustering and Marketing Optimization in Hospitality: A Hybrid Data Mining and Decision-Making Approach from an Emerging Economy. *Tourism and Hospitality*, *6*(2), 80. https://doi.org/10.3390/tourhosp6020080

Duan, G., & Song, G. (2025). Corporate marketing based on improved depth-weighted k-mean arithmetic and improved extreme gradient boosting tree. *International Journal of Industrial Engineering Computations*, *16*(4), 1143–1154. https://doi.org/10.5267/j.ijiec.2025.6.010

Hassan, M. M., Mollick, S., & Yasmin, F. (2022). An unsupervised cluster-based feature grouping model for early diabetes detection. *Healthcare Analytics*, *2*, 100112. https://doi.org/10.1016/j.health.2022.100112

Kasem, M. S., Hamada, M., & Taj-Eddin, I. (2024). Customer profiling, segmentation, and sales prediction using AI in direct marketing. *Neural Computing and Applications*, *36*(9), 4995–5005. https://doi.org/10.1007/s00521-023-09339-6

Kumar, N. (2025). Intelligent customer segmentation: unveiling consumer patterns with machine learning. *Journal of Umm Al-Qura University for Engineering and Architecture*, *16*(3), 774–783. https://doi.org/10.1007/s43995-025-00180-7

Mamun, M. N. H. (2025). Advancements in machine learning for customer retention: A systematic literature review of predictive models and churn analysis. *Journal of Sustainable Development and Policy*, *1*(01), 250–284.

Nagib, R., & Dewanti, M. C. (2025). Analisis Pengaruh Teknologi Digital Printing Terhadap Efisiensi Produksi Dan Kualitas Hasil Cetak Di B21 Digital Printing. *EKBIS (Ekonomi & Bisnis)*, *13*(2), 255–261. https://doi.org/10.56689/ekbis.v13i2.2299

Paparrizos, J., & Bogireddy, S. P. T. R. (2025). Time-series clustering: A comprehensive study of data mining, machine learning, and deep learning methods. *Proceedings of the VLDB Endowment*, *18*(11), 4380–4395.

Paparrizos, J., & Gravano, L. (2017). Fast and Accurate Time-Series Clustering. *ACM Transactions on Database Systems*, *42*(2), 1–49. https://doi.org/10.1145/3044711

Sugumaran, M. (2025). *E-Commerce Customer Retention Analysis*. Dublin, National College of Ireland.

Tabianan, K., Velu, S., & Ravi, V. (2022). K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data. *Sustainability*, *14*(12), 7243. https://doi.org/10.3390/su14127243

Ufeli, C. P., Sattar, M. U., Hasan, R., & Mahmood, S. (2025). Enhancing Customer Segmentation Through Factor Analysis of Mixed Data (FAMD)-Based Approach Using K-Means and Hierarchical Clustering Algorithms. *Information*, *16*(6), 441. https://doi.org/10.3390/info16060441

Ullah, A., Mohmand, M. I., Hussain, H., Johar, S., Khan, I., Ahmad, S., Mahmoud, H. A., & Huda, S. (2023). Customer Analysis Using Machine Learning-Based Classification Algorithms for Effective Segmentation Using Recency, Frequency, Monetary, and Time. *Sensors*, *23*(6), 3180. https://doi.org/10.3390/s23063180

Vera, J. F., & Angulo, J. M. (2023). An MDS-based unifying approach to time series K-means clustering: application in the dynamic time warping framework. *Stochastic Environmental Research and Risk Assessment*, *37*(12), 4555–4566. https://doi.org/10.1007/s00477-023-02470-9

Vianna Filho, C., Luiz, A., de Lima, L., & Kleina, M. (2026). RFM model customer segmentation from a graph theory perspective. *Quality & Quantity*, 1–22.

Wang, H., Ou, Y., Liu, F., Li, X., Cao, Y., Wang, B., He, J., Liu, T., & Wang, Y. (2026). Research on neighbor grid-based adaptive spatiotemporal DBSCAN algorithm. *International Journal of Machine Learning and Cybernetics*, *17*(3), 123.

Wang, S., Sun, L., & Yu, Y. (2024). A dynamic customer segmentation approach by combining LRFMS and multivariate time series clustering. *Scientific Reports*, *14*(1), 17491. https://doi.org/10.1038/s41598-024-68621-2

Wilbert, H. J., Hoppe, A. F., Sartori, A., Stefenon, S. F., & Silva, L. A. (2023). Recency, Frequency, Monetary Value, Clustering, and Internal and External Indices for Customer Segmentation from Retail Data. *Algorithms*, *16*(9), 396. https://doi.org/10.3390/a16090396