Green Intelligent Systems and Applications

Research Article
Volume 6(1), 2026, 65−80
https://doi.org/10.53623/gisa.v6i1.919

Recency, Frequency, and Monetary-Based Customer Segmentation Using K-Means for Analysing Transactional Behaviour in a Service-Based Micro, Small, and Medium Enterprises

Rizka Ardiansyah[1]*, Nouval Trezandy[1], Iskandar Iskandar[2], Meilani Ilman[1], Sahril Sahril[1]

[1]Information Technology Department, Faculty of Engineering, Tadulako University, Indonesia
[2]Mechanical Engineering Department, Faculty of Engineering, Tadulako University, Indonesia

*Correspondence: rizka@untad.ac.id

SUBMITTED: 24 November 2025; REVISED: 15 February 2026; ACCEPTED: 24 February 2026

ABSTRACT: Micro, Small, and Medium Enterprises (MSMEs) often faced challenges in designing effective promotional initiatives due to the limited use of systematic customer behavior analysis. This study examined the application of (Recency, Frequency, Monetary) RFM analysis combined with K-Means clustering to explore customer segmentation in a service-based MSME context. Transaction data from a local laundry service operating in Palu, Indonesia, consisting of 2,220 digital transaction records collected between 2022 and 2025, were processed and transformed into RFM variables using min–max normalization. The optimal number of clusters was determined using the Elbow method, resulting in four customer segments. Cluster quality was evaluated using internal validation metrics, yielding a Davies–Bouldin Index (DBI) of 0.61 and a Sum of Squared Errors (SSE) value of 1.73, indicating reasonably compact and well-separated clusters. The resulting segments exhibited distinct transactional profiles across recency, transaction frequency, and monetary contribution, reflecting heterogeneity in customer engagement within the studied MSME. Rather than prescribing specific marketing actions, the findings provided an interpretable analytical basis for considering differentiated promotional strategies aligned with observed customer behavior patterns. Overall, this study demonstrated that RFM-based segmentation offered a feasible and data-driven approach to supporting evidence-informed promotional planning in service-oriented MSMEs operating under data and resource constraints.

KEYWORDS: RFM analysis; K-Means clustering; customer segmentation; service-based MSME; transactional behaviour analysis.

1. Introduction

Micro, Small, and Medium Enterprises (MSMEs) play a central role in Indonesia's economic growth. However, despite their importance, many MSMEs have relied on intuition rather than data-driven approaches when planning promotional strategies, often resulting in inefficient marketing efforts and imprecise customer targeting [1]. In service-based MSMEs, such as laundry businesses, customer retention and transactional behavior are particularly critical due to high competition and relatively homogeneous service offerings [2]. Customer segmentation has therefore become an important analytical tool for understanding purchasing behavior and supporting marketing decision-making.

Among various approaches, Recency, Frequency, and Monetary (RFM) analysis has gained widespread attention as a simple yet effective method for capturing customer transaction characteristics. Prior studies have shown that combining RFM with clustering algorithms can successfully identify meaningful customer segments across retail, e-commerce, and hospitality sectors [3–11]. However, the existing research landscape is characterized by several key shortcomings. First, many RFM-based clustering studies primarily focus on the formation of customer segments, while providing limited analytical interpretation of how the resulting segments can be operationally linked to promotional decision-making, particularly in small service-oriented businesses. Consequently, clustering outcomes are often treated as descriptive labels rather than actionable insights.

Second, although RFM and clustering methods are frequently described as simple and practical, prior studies often lack sufficient methodological clarity regarding the analytical workflow, parameter selection, and evaluation criteria, making replication and adoption challenging for MSMEs with limited technical resources. This methodological gap is especially relevant for traditional service-based MSMEs, where data availability is typically longitudinal but not supported by advanced digital infrastructures.

Third, existing studies predominantly examine large-scale enterprises or digital platforms, while empirical evidence remains limited regarding the applicability and limitations of RFM-based segmentation in small-scale, traditional service MSMEs. In particular, there is a lack of context-aware analysis that explicitly discusses under what conditions simple RFM-based clustering remains effective and where its interpretative boundaries lie.

Motivated by these gaps, this study applied an RFM-based K-Means clustering approach to transaction data from a local MSME laundry service in Palu, Indonesia. The research aimed to provide an interpretable and replicable analytical framework that connects customer segmentation results with promotional strategy formulation, while explicitly acknowledging the contextual constraints of small service-based MSMEs. Accordingly, this study addressed the following research question: How can RFM-based K-Means clustering be analytically interpreted and operationalized to support promotional strategy development in resource-constrained, service-based MSMEs?

2. Related Work

Customer segmentation has been widely studied as a fundamental aspect of data-driven marketing and customer relationship management. One of the most established approaches in this domain is Recency, Frequency, and Monetary (RFM) analysis, which provides a parsimonious representation of customer transactional behavior. Previous studies have demonstrated that RFM variables can effectively capture differences in customer engagement and value contribution across various business contexts, particularly when combined with clustering techniques to uncover latent customer groups [3–8]. Several studies have applied RFM-based clustering in retail, e-commerce, and hospitality domains, reporting that algorithms such as K-Means can identify meaningful customer segments with distinct behavioral characteristics [9–11].

Recent studies continue to refine and apply RFM-based clustering frameworks in practical business environments. For instance, Syahra et al. [12] demonstrated the implementation of RFM combined with K-Means clustering to support customer relationship management in retail SMEs, highlighting the practical relevance of transaction-based segmentation for small and medium enterprises. Similarly, Agus and Hasibuan [13] proposed an enhanced RFM–K-Means framework to improve clustering robustness and segmentation quality, illustrating ongoing methodological refinement within the RFM segmentation stream.

Beyond RFM-specific applications, comparative studies on clustering algorithms further emphasize methodological considerations in customer segmentation research. John et al. [14] evaluated multiple clustering techniques in a retail context and observed that while more advanced algorithms may improve certain performance metrics, simpler methods such as K-Means remain competitive and offer advantages in interpretability and practical implementation. These findings reinforce the importance of balancing analytical adequacy with transparency, particularly when segmentation outcomes must be interpretable for managerial decision-making.

In the context of Micro, Small, and Medium Enterprises (MSMEs), customer analytics research remains comparatively limited. Existing studies on MSMEs and service-based businesses suggest that customer transaction data in these settings are often longitudinal but low-dimensional, making simple and interpretable analytical approaches particularly relevant [15–16]. However, many prior studies either adopt generic segmentation frameworks without contextual adaptation or prioritize methodological novelty over practical feasibility for small businesses under resource constraints. Based on these observations, this study positioned itself within the stream of RFM-based customer segmentation research while explicitly focusing on service-based MSMEs. This work adopted an interpretative perspective, examining how RFM-based K-Means clustering can be analytically interpreted to inform reflective promotional planning in a resource-constrained service environment. The study sought to address gaps related to contextual relevance, methodological clarity, and interpretability in existing customer segmentation research.

3. Materials and Methods

3.1. Research design.

This study adopted an RFM-based K-Means clustering approach guided by considerations of analytical sufficiency, interpretability, and contextual feasibility rather than methodological novelty. RFM metrics were particularly suitable for service-based MSMEs, where transaction data were typically limited in dimensionality but available longitudinally. By capturing recency, frequency, and monetary value, RFM provided an interpretable representation of customer transactional behavior that aligned with the analytical needs of small businesses.

K-Means clustering was employed to support exploratory customer segmentation with transparent centroid-based interpretation, allowing the resulting segments to be readily translated into actionable promotional strategies. While more complex clustering techniques exist, their additional complexity often reduces interpretability and practical adoption in resource-constrained MSME contexts. Consequently, K-Means was selected based on its fitness for purpose rather than algorithmic sophistication.

Cluster validity was assessed using the Sum of Squared Errors (SSE) and the Davies–Bouldin Index (DBI) to evaluate internal cohesion and separation. These indices were used to support the analytical adequacy of the clustering results rather than to claim optimal or universal segmentation. This methodological choice was consistent with the study's objective to provide a structured, interpretable, and replicable analytical framework for MSME promotional planning under contextual constraints.

The MSME selected for this study was chosen based on a set of inclusion and exclusion criteria summarized in Table 1. These criteria were designed to ensure the availability of sufficient behavioral data and the suitability of the business model for RFM-based customer analysis. The selection criteria were grounded in established data-mining and customer analytics research, which emphasizes the need for adequate transaction frequency, data completeness, and domain relevance to generate meaningful segments [7–9].

Table 1. MSME selection criteria.

Category	Inclusion Criteria	Exclusion Criteria	Justification
Type of Product/Service	Service-based MSMEs with repetitive or habitual customer interactions	One-time purchase MSMEs	RFM requires repeated customer activity; habitual purchase cycles yield more interpretable recency & frequency patterns [7].
Service Characteristics	MSMEs offering standardized, measurable services	Custom, highly variable services without consistent pricing	Standardized services produce stable monetary patterns necessary for meaningful segmentation [8].
Digital Data Availability	Digital transaction logs (POS system, Excel, digital receipts, mobile apps) covering ≥ 6 months	MSMEs using fully manual bookkeeping or incomplete records	RFM computation requires timestamped and complete transaction histories; incomplete data compromises clustering reliability [9].
Transaction Volume	A sufficient transaction volume (e.g., ≥ 500 transactions or ≥ 100 unique customers over study period)	MSMEs with very low transaction counts (< 100 transactions)	Clustering algorithms like K-Means require adequate sample size to form distinct customer segments [10].
Location & Context	Local Indonesian MSMEs with stable operations over ≥ 1 year	Seasonal, temporary, or newly established businesses	

To ensure the analytical suitability of the selected MSME for RFM-based customer segmentation, a set of inclusion and exclusion criteria was applied, as summarized in Table 1. These criteria were designed to confirm that the business context supported meaningful interpretation of recency, frequency, and monetary patterns, while providing sufficient and reliable transaction data for clustering analysis. In line with established customer analytics and data-mining literature, the selection emphasized service characteristics associated with regular customer interactions, the availability of digitally recorded transaction histories, and an adequate transaction volume to support stable and interpretable customer segments [7–11]. Collectively, these considerations ensured that the selected MSME reflected realistic customer behaviour patterns commonly observed in service-based micro-enterprise contexts, without overstating the generalizability of the findings.

3.2. Data collection.

This study employed a quantitative, descriptive approach using transaction data obtained from a laundry service MSME operating in Palu, Indonesia. All available digital transaction records from 2022 to 2025 were included, resulting in a dataset comprising 2,220 transaction records. As all eligible records within the study period were considered, the dataset represented a complete census rather than a sampled subset. The unit of analysis was the individual customer. Customer-level transaction histories were aggregated to compute Recency, Frequency, and Monetary (RFM) values, including purchase recency, transaction frequency, and total monetary expenditure over the observation period. The use of digitally recorded transaction logs ensured the availability of timestamped and complete data required for accurate RFM computation. All transaction data were anonymized prior to analysis to ensure customer privacy.

3.3. Data inclusion strategy.

This study employed a complete enumeration (census-based) data inclusion strategy, in which all available digital transaction records from the selected MSME were included in the analysis. This approach was adopted because the objective of the study was to analyze the entire population of recorded customer interactions. The selected MSME maintained consistent digital transaction logs, enabling comprehensive coverage of customer behaviour over the observation period. A total of 2,220 customer transaction records from 2022 to 2025 were utilized. The inclusion of the full dataset supported more reliable computation of Recency, Frequency, and Monetary (RFM) values and contributed to the stability of the clustering results. Prior studies in customer analytics have highlighted that using complete transaction histories enhances the validity of segmentation outcomes and reduces sampling bias, particularly in small and medium enterprise contexts [15–16]. The attributes included in the dataset are summarized in Table 2.

Table 2. Data attribute.

No	Data Attribute	Description
1	Date	Date the transaction was conducted
2	Object	Transaction object (clothes, curtains, bed, linen etc)
3	Service	Selected Service(complete, dry-clean, express service etc)
4	Quantity	Quantity of object (in kg or pcs)
5	Price	Price of service (in IDR)
6	Total Price	Total Price of Service (in IDR)
7	Customer	Customer name

3.4. Data preprocessing & clustering procedure.

The collected transaction records were processed through a multi-stage analytical workflow consisting of RFM variable construction, data normalization, and clustering. This workflow followed established practices in customer segmentation research, where behavioural indicators are standardized prior to cluster formation to ensure that each dimension contributes equally [14, 17]. All preprocessing and clustering procedures were performed using Python 3.x with the Numpy, Pandas, and Scikit-learn libraries.

3.4.1. RFM Variable Construction.

The RFM model was applied to quantify customer transactional behaviour using three metrics: Recency (R), Frequency (F), and Monetary (M). Recency represents the number of days since a customer's most recent transaction and was calculated as the difference between the reference date and the date of the last purchase:

$$R = \text{Current Date} - \text{Date of Last Purchase} \qquad (1)$$

Frequency captures the total number of transactions performed by a customer during the observation period:

$$F = \sum_{i=1}^{n} f_i \qquad (2)$$

where $f_i$denotes the count of each transaction event.

Monetary reflects the cumulative transaction value of a customer over the observation period:

$$M = \sum_{i=1}^{n} p_i \qquad (3)$$

where $p_i$represents the monetary value of each transaction. These metrics are widely used to assess customer engagement, spending intensity, and purchase recency in data-driven marketing and service analytics [3, 5, 17, 18].

3.4.2. RFM normalization.

Recency, Frequency, and Monetary metrics operate on different numeric scales; therefore, normalization was applied to ensure equal weighting during distance-based clustering. Min–Max normalization was employed to transform all RFM variables into a comparable range. Without normalization, features with larger numeric ranges, such as Monetary, could dominate distance calculations in K-Means [17, 19]. Since lower Recency values indicate more recent customer activity, the Recency scale was reversed so that higher normalized scores consistently represent more desirable customer behaviour. This adjustment aligns the directionality of Recency with Frequency and Monetary, following common practice in RFM-based modelling [18], and supports transparent implementation in resource-constrained MSME settings where methodological simplicity is important [20].

The min–max normalization for Recency was computed as:

$$R_{\text{normalized}} = \frac{R_{\max} - R}{R_{\max} - R_{\min}} \qquad (4)$$

For Frequency and Monetary, higher raw values indicate stronger engagement. The standard min–max normalization was applied as:

$$F_{\text{normalized}} = \frac{F - F_{\min}}{F_{\max} - F_{\min}} \qquad (5)$$

$$M_{\text{normalized}} = \frac{M - M_{\min}}{M_{\max} - M_{\min}} \quad (6)$$

3.4.3. K-Means clustering procedure.

K-Means clustering was applied to the normalized RFM dataset to group customers with similar transactional patterns. The algorithm partitions observations by iteratively assigning each data point to the nearest cluster centroid and updating centroid positions to minimize within-cluster variance. This objective was formalized through the Within-Cluster Sum of Squared Errors (WSS), which measures the total squared Euclidean distance between each data point and the centroid of its assigned cluster:

$$WSS = \sum_{i=1}^{k} \sum_{x \in C_i} \| x - \mu_i \|^2 \quad (7)$$

where $k$denotes the number of clusters, $C_i$represents the set of points assigned to cluster i, and $\mu_i$is the centroid of cluster i. By minimizing the WSS, K-Means produces clusters that are internally compact, such that data points within the same cluster exhibit high similarity in their RFM profiles. This formulation makes K-Means particularly suitable for numerical behavioural data and supports its widespread use in customer segmentation research due to its simplicity and computational efficiency [14, 18]. The clustering procedure was implemented using the K-Means algorithm provided by the Scikit-learn library.

3.5. Cluster validation method.

To evaluate the quality of the clusters produced by K-Means, the Davies–Bouldin Index (DBI) was employed as an internal cluster validation metric, assessing both cluster compactness and separation. DBI measures the average similarity between each cluster and its most similar neighbouring cluster, based on the ratio of within-cluster dispersion to between-cluster separation [17, 21]. Lower DBI values indicate clusters that are more compact and better separated.

Cluster compactness was represented by the average distance of each data point in a cluster to its centroid:

$$s_i = \frac{1}{n_i} \sum_{x \in C_i} d(x, c_i) \quad (8)$$

where $s_i$denotes the compactness of cluster i, $C_i$represents the set of data points in cluster i, $c_i$is the centroid of cluster i, $d(x, c_i)$is the distance between data point $x$and centroid c_i, and $n_i$is the number of points in cluster i.

Cluster separation was defined as the distance between the centroids of two clusters:

$$d_{ij} = d(C_i, C_j) \qquad (9)$$

Higher separation values indicate better distinction between groupings. The Davies–Bouldin Index was then computed as:

$$DBI = \frac{1}{N}\sum_{i=1}^{N} \max\left(\frac{s_i + s_j}{d_{ij}}\right) \qquad (10)$$

where Ndenotes the total number of clusters. DBI was selected because it does not require external class labels and is well-suited for evaluating RFM-based customer segmentation using distance-based clustering algorithms [17, 18].

4. Results

4.1. Determination of the number of clusters.

To determine the appropriate number of clusters for K-Means clustering, both the Elbow Method and the Davies–Bouldin Index (DBI) were applied to the normalized RFM dataset. The Elbow Method evaluated the Within-Cluster Sum of Squared Errors (WSS) across different values of $k$to assess changes in cluster compactness as the number of clusters increased. Figure 1 presents the WSS values for varying cluster counts. A substantial reduction in WSS was observed as the number of clusters increased from $k = 2$to $k = 3$, followed by a noticeably smaller rate of improvement beyond $k = 4$. This pattern suggested a point of diminishing returns, where adding additional clusters provided limited gains in cluster compactness.

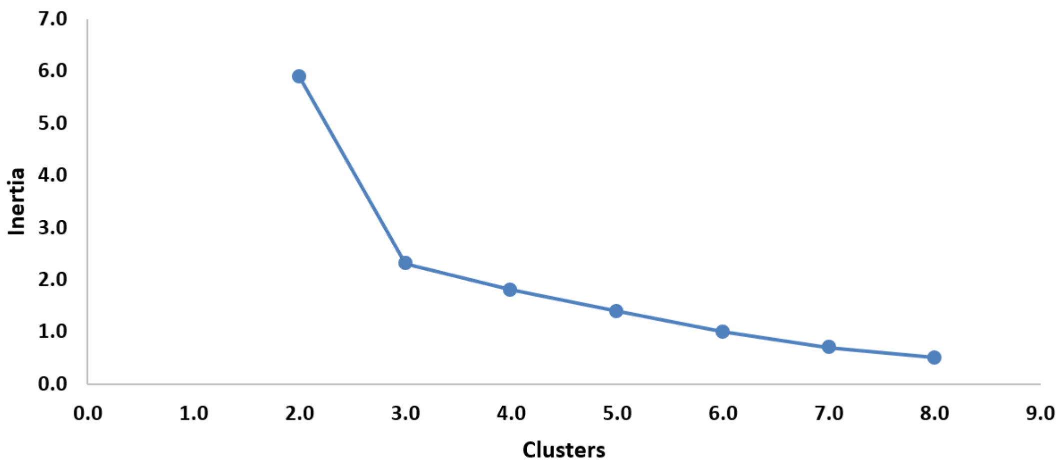

Figure 1. Within-cluster-sum of squared errors for each K.

The cluster solution was further evaluated using the Davies–Bouldin Index (DBI), which measures the balance between intra-cluster cohesion and inter-cluster separation. For the selected clustering configuration, a DBI value of 0.61 was obtained, indicating a relatively compact and well-separated cluster structure within the context of RFM-based customer segmentation. Based on the combined evidence from the Elbow Method and DBI evaluation, a four-cluster solution was chosen for subsequent analysis. This configuration balanced model parsimony with structural differentiation, enabling meaningful segmentation of customer transactional behaviour.

4.2. Cluster Distribution and Structure

Following the selection of the four-cluster solution, the distribution of customers across clusters was examined to assess the structural composition of the segmentation results. Figure 2 presents a scatter plot of the clustered data projected onto the Recency and Monetary dimensions, illustrating the relative positioning of customer groups and their centroids. The cluster distribution exhibited noticeable variation in cluster sizes: one cluster contained a substantially larger proportion of customers, whereas the remaining clusters were comparatively smaller. This uneven distribution reflects the heterogeneity in customer transactional behaviour captured by the RFM variables, with distinct groupings formed based on differences in recency and spending intensity.

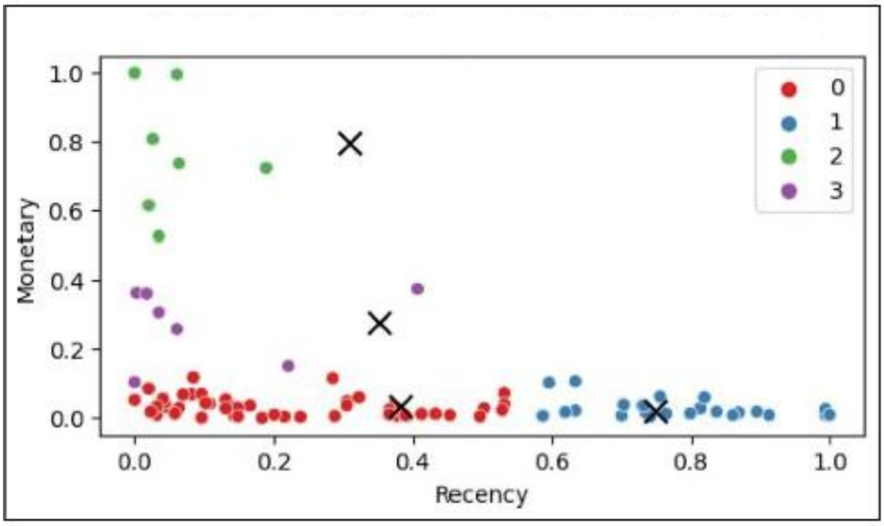

Figure 2. Cluster data scatter plot.

The spatial separation observed in Figure 2 indicated that the clusters occupied distinct regions of the RFM feature space. Customers with higher normalised recency and monetary values were concentrated in a specific region, whereas customers with lower engagement levels were grouped separately. The centroids of each cluster further highlighted these structural differences, representing the average transactional profile of customers within each group. Overall, the distribution and spatial arrangement of the clusters demonstrated that the K-Means algorithm effectively partitioned the normalised RFM dataset into structurally distinct customer groups, providing a solid foundation for the subsequent analysis of cluster characteristics.

4.3. RFM characteristics of each cluster.

To further examine the segmentation results, the characteristics of each cluster were analysed based on aggregated RFM values. Table 3 summarises the RFM profiles of the four clusters, highlighting relative differences in transaction recency, transaction frequency, and cumulative monetary value across customer groups. The clusters exhibited distinct RFM patterns.

Differences in the recency dimension indicated variation in how recently customers interacted with the service. Similarly, frequency values varied across clusters, reflecting differences in the number of transactions during the observation period. Monetary values also showed notable variation, highlighting differences in cumulative spending among customer groups. One cluster was characterised by comparatively higher frequency and monetary values, while other clusters displayed lower transaction frequency or lower cumulative spending. Variations in recency further distinguished the clusters, with some groups exhibiting more recent transactional activity than others. These differences indicate that the clustering process effectively separated customers into groups with distinct transactional profiles across all three RFM dimensions. The RFM characteristics reported in Table 3 demonstrated clear differentiation between clusters in terms of transaction timing, engagement intensity, and spending magnitude, providing a structured empirical basis for subsequent interpretation.

Table 3. RFM characteristics.

Clusters	Recency	Frequency	Monetary	Category
1	Moderately High to High	Low	Low	Fairly Loyal Consumers
2	Low	Low	Low	Not Loyal Consumers
3	High	High to Very High	High to Very High	Very Loyal Consumers
4	High to Very High	Low to Moderately High	Low to Moderately High	Loyal Consumers

5. Discussion

5.1. Interpretation of customer segments.

Based on the RFM characteristics reported in the Results, the identified customer clusters were further examined to derive meaningful interpretative profiles. This interpretation contextualised observed differences in Recency, Frequency, and Monetary values by linking them to commonly used engagement-based customer segment descriptors in the customer analytics literature. This step synthesised the clustering outcomes to support analytical understanding of customer behaviour patterns within the studied service-based MSME context. Table 4 presents an interpretative summary of the RFM-based customer segments, combining relative RFM profiles with descriptive labels that reflect differences in transaction recency, engagement intensity, and spending magnitude. The table functions as a conceptual bridge between the quantitative clustering results and the subsequent discussion of managerial implications, remaining grounded in the observed transactional data. These interpretative labels were designed to enhance analytical clarity and facilitate practical understanding of the customer segments.

Table 4. Interpretative summary of RFM-based customer segments.

Cluster	R (Recency)	F (Frequency)	M (Monetary)	Interpretation
Cluster 1 Fairly Loyal Consumers	Moderately high	Low	Low	Recent activity but low spending and infrequent visits. Shows latent potential for conversion through targeted engagement.
Cluster 2 Not Loyal Consumers	Low (long inactivity)	Very low	Low	Characterised by low engagement and high churn risk or reprioritization depending on resources.
Cluster 3 Very Loyal Consumers	High	Very high	Highest	Core customer base with strong loyalty and major revenue contribution. Represent high-value and consistently engaged customer group
Cluster 4 Loyal Consumers	High to very high	Moderate	Moderate	Consistently engaged segment. Moderate-value customers with upgrade potential through loyalty programs.

5.2. Implications for promotional strategy.

Building on the interpreted RFM-based customer segments, this study derived indicative implications for promotional planning in service-based MSMEs. These implications aimed to translate observed transactional patterns into structured managerial considerations, rather than to prescribe fixed or universally optimal strategies. Based on the segmentation synthesis summarized in Table 4, Table 5 presents a set of indicative promotional considerations aligned with each customer segment.

Table 5. Indicative promotional considerations for RFM-based customer segments.

Cluster	Strategic Focus	Recommended Strategies	Rationale
Cluster 3 Very Loyal Consumers (Retention Priority)	Retain and reward high-value customers	o Loyalty rewards (points, tier-based perks) o Exclusive discounts or service upgrades o Priority handling (express service) o Personalized appreciation messages	Personalization, recognition, and loyalty-based incentives strengthen emotional attachment and increase long-term retention among high-value customers, particularly in digitally mediated service environments [22−24].
Cluster 4 Loyal Consumers (Upsell / Engagement Focus)	Increase spending and engagement	o Targeted promotions to boost purchase frequency o Tier-based membership with escalating benefits o Value-added services (free pick-up upon reaching spending thresholds)	Moderate-value customers tend to respond positively to structured engagement programs, tier-based incentives, and convenience-oriented services that enhance perceived value and strengthen relational commitment [22, 23].
Cluster 1 Fairly Loyal Consumers (Activation Segment)	Stimulate more frequent transactions	o Service reminders / engagement nudges o Bundle-based promotional offers o Promote digital convenience (online booking/payment) o Introductory discounts for premium services	Customers exhibiting moderate engagement signals can be effectively stimulated through targeted reminders and behavioral nudges, as small promotional triggers are shown to increase transaction frequency and spending intensity [23].
Cluster 2 Not Loyal Consumers (Re-engagement or Deprioritization)	Low-cost re-engagement or selective targeting	o WhatsApp/SMS re-engagement messages o Personalized outreach (feedback, surveys) o Omni-channel communication (Facebook, Instagram, TikTok) o Offer multiple secure payment options	Low-engagement customers require cost-efficient, selectively targeted re-engagement strategies, as excessive marketing expenditure toward low-loyalty segments may not generate proportional returns [22, 24].

The proposed strategies were framed as illustrative examples of how MSMEs could differentiate promotional efforts according to customer engagement levels and value contribution. For customers exhibiting high frequency and monetary values, promotional considerations emphasised retention-oriented approaches that recognised existing engagement. For segments with moderate engagement, strategies focused on incremental stimulation of transaction frequency or service utilisation. In contrast, segments characterised by low recency and low transaction intensity were associated with more selective or low-cost engagement approaches, reflecting resource constraints commonly faced by MSMEs. Importantly, the promotional considerations outlined in Table 5 were not intended as prescriptive recommendations, but as context-dependent mappings between observed customer behaviour and potential managerial responses reported in prior studies. The feasibility and effectiveness of any specific promotional action remained contingent on organisational capacity, market conditions, and customer preferences. Collectively, these tables illustrated how RFM-based segmentation could inform reflective and differentiated promotional planning in service-oriented MSME environments, rather than providing a one-size-fits-all solution.

5.3. Methodological and contextual reflections.

This study adopts an RFM-based K-Means clustering approach with an emphasis on interpretability, analytical sufficiency, and contextual feasibility. From a methodological perspective, the use of RFM variables and K-Means clustering offers a transparent and replicable framework for customer segmentation in data-constrained service-based MSME environments. However, this approach relies on internal validity measures and distance-based assumptions, which may limit its sensitivity to complex or non-linear behavioural patterns. The clustering results should therefore be interpreted as exploratory groupings that summarise Transactional similarities rather than definitive representations of customer typologies. The selection of a census-based dataset strengthens the internal consistency of the analysis by capturing the full population of recorded customer transactions within the observation period. At the same time, this context-specific focus limits the external generalisability of the findings. Customer behaviour patterns observed in a single service-based MSME may differ from those in other sectors or geographic settings, particularly where digital maturity, service frequency, or pricing structures vary substantially. Contextual factors inherent to service-based MSMEs also shape the interpretation of the results. Transactional behaviour in routine services, such as laundry services, is often influenced by situational needs, proximity, and habitual usage rather than deliberate brand switching. Consequently, low recency or frequency values do not necessarily indicate customer dissatisfaction, but may reflect episodic consumption patterns common in local service markets. Finally, the promotional implications derived from the segmentation should be viewed as indicative rather than prescriptive. The effectiveness of any promotional action depends on organisational capacity, resource availability, and customer responsiveness, which were not explicitly evaluated in this study. Future research may extend this work by incorporating longitudinal evaluation of promotional interventions, additional behavioural features, or comparative analysis across multiple MSMEs to further examine the robustness and transferability of the proposed segmentation framework.

6. Limitation of The Study

This study has several limitations that should be acknowledged:

a) The analysis is based on transaction data from a single service-based MSME operating in the laundry sector. While the use of a complete census strengthens internal consistency within the studied context, the findings may not be directly generalisable to other service industries or MSMEs with different operational characteristics, customer bases, or levels of digital maturity. Future research involving multiple MSMEs across diverse service domains and geographic regions may enhance the external validity of the segmentation results.
b) The RFM model employed in this study relies exclusively on Recency, Frequency, and Monetary variables derived from Transactional records. Although RFM provides a parsimonious and interpretable representation of customer behaviour, it does not capture other potentially relevant dimensions such as service-type preferences, communication channels, customer demographics, or situational factors influencing service usage. The inclusion of richer multidimensional behavioural data may enable more nuanced segmentation and deeper behavioural insights.
c) The promotional considerations discussed in this study are analytically derived from observed RFM-based segment characteristics and established patterns reported in the literature. These considerations have not been empirically validated through real-world implementation or experimental evaluation. As a result, their effectiveness in improving customer engagement or business performance cannot be directly inferred. Future studies may employ longitudinal designs, field experiments, or A/B testing to assess the practical impact of segment-specific promotional interventions.
d) The K-Means clustering algorithm assumes relatively spherical cluster structures and requires prior specification of the number of clusters. While this approach supports interpretability and computational efficiency, it may not fully capture complex or overlapping behavioural patterns present in real-world customer data. Future research may explore alternative or complementary clustering techniques, such as density-based, hierarchical, or probabilistic models, to assess the robustness of the segmentation outcomes under different methodological assumptions.

7. Conclusions

This study examined the application of an RFM-based K-Means clustering approach for customer segmentation in a service-based MSME context, using 2,220 digital transaction records from a local laundry business collected between 2022 and 2025. The analysis identified four distinct customer clusters, each characterised by unique patterns of recency, transaction frequency, and monetary contribution, demonstrating that a parsimonious RFM framework could capture meaningful variation in customer transactional behaviour. The clustering results exhibited satisfactory internal structure, as indicated by a Davies–Bouldin Index of 0.61, suggesting that the clusters were reasonably compact and well-separated. While not claiming optimal or universal segmentation, these results confirmed the analytical adequacy of the approach for exploratory customer segmentation in data-constrained service-based MSME environments. Through interpretative analysis, the clusters were linked to engagement-based customer segment profiles commonly described in the customer analytics literature. This interpretation highlighted heterogeneity in customer behaviour and provided a structured empirical basis for reflecting on differentiated promotional considerations. The derived promotional implications were indicative, intended to guide context-dependent decision-making rather than serve as prescriptive managerial solutions. Overall, the findings demonstrated that RFM-based segmentation can provide a practical and interpretable tool for MSMEs seeking to move beyond intuition-driven marketing toward more evidence-informed practices. Although the approach does not capture all dimensions of customer behaviour, it offers a feasible starting point for data-driven customer analysis in service-oriented MSMEs with limited analytical resources. Future research may build on this work by incorporating additional behavioural variables, alternative clustering techniques, or empirical evaluation of segment-specific interventions to further assess the robustness and applicability of the proposed framework.

Acknowledgements

The authors would like to express their sincere appreciation to Faculty of Engineering, Tadulako University for supporting this research through Dana BLU Fakultas Teknik Universitas Tadulako Tahun Anggaran 2024, provided under the Perjanjian Penugasan Pelaksanaan Program Penelitian Nomor: 2653/UN28/KU/2024, dated 22 May 2024. This financial support played a crucial role in enabling the successful completion of this study.

Competing Interest

Although this research received funding support from the BLU Fakultas Teknik Universitas Tadulako (Perjanjian Penugasan Nomor 2653/UN28/KU/2024), the authors declare that the funding body had no role in the study design, data collection, analysis, interpretation, or manuscript preparation. The authors report no competing interests.

References

[1] Anatan, L.; Nur, A. (2023). Micro, Small, and Medium Enterprises' Readiness for Industry 4.0: Evidence from Indonesia. *Economies*, *11*, 156. https://doi.org/10.3390/economies11060156.

[2] Gao, J.; Li, G.; Chen, N.; Alzahrani, A.S.H. (2023). Impact of E-Commerce and Digital Marketing Adoption on Sustainability of SMEs. *Sustainability*, *15*, 1594. https://doi.org/10.3390/su15021594.

[3] Ernawati, E.; Baharin, S.S.K.; Kasmin, F. (2021). A Review of Data Mining Methods in RFM-Based Customer Segmentation. *Journal of Physics: Conference Series*, *1869*, 012085. https://doi.org/10.1088/1742-6596/1869/1/012085.

[4] Lang, L.; Liu, X.; Zhang, J. (2022). A Big Data-Based Dynamic Weight Approach for RFM Model in Customer Segmentation. *International Journal of Information Management Data Insights*, *2*, 100086. https://doi.org/10.1016/j.jjimei.2022.100086.

[5] Xu, L.; Li, Y.; Yu, H. (2021). Customer Segmentation Based on Improved RFM Model and K-Means Algorithm. *IEEE Access*, *9*, 13721–13728. https://doi.org/10.1109/ACCESS.2021.3051867.

[6] Rungruang, C.; Boongoen, S.; Athanon, T. (2024). RFM Model Customer Segmentation Based on Hierarchical Approach Using FCA. *Expert Systems with Applications*, *235*, 121206. https://doi.org/10.1016/j.eswa.2023.121206.

[7] Deng, Y.; Gao, Q. (2020). A Study on E-Commerce Customer Segmentation Management Based on Improved K-Means Algorithm. *Information Systems and e-Business Management*, *18*, 497–510. https://doi.org/10.1007/s10257-019-00444-8.

[8] Wu, T.; Liu, X. (2020). A Dynamic Interval Type-2 Fuzzy Customer Segmentation Model and Its Application in E-Commerce. *Applied Soft Computing*, *94*, 106366. https://doi.org/10.1016/j.asoc.2020.106366.

[9] Smaili, M.Y.; Hachimi, H. (2023). New RFM-D Classification Model for Improving Customer Analysis and Response Prediction. *Ain Shams Engineering Journal*, *14*, 102254. https://doi.org/10.1016/j.asej.2022.102254.

[10] Priyatharsini, L.; Umamakeswari, A.; Joy, J.C.L. (2018). RFM Ranking – An Effective Approach to Customer Segmentation. *Journal of King Saud University – Computer and Information Sciences*, *30*, 1–10. https://doi.org/10.1016/j.jksuci.2017.12.006.

[11] Oliveira, J.P.; Oliveira, L.E.S.; Cortez, P. (2021). Unsupervised Anomaly Detection of Retail Stores Using RFM and K-Means. *Procedia Computer Science*, *181*, 644–651. https://doi.org/10.1016/j.procs.2021.01.201.

[12] Syahra, Y.; Fadlil, A.; Yuliansyah, H. (2025). Customer Segmentation Using RFM and K-Means Clustering to Support CRM in Retail Industry. *Sinkron: Jurnal dan Penelitian Teknik Informatika*, *9*, 1120–1131. https://doi.org/10.33395/sinkron.v9i3.14907.

[13] Agus, I.; Hasibuan, M. (2025). Customer Segmentation Using an Enhanced RFM–K-Means Framework on the Online Retail Dataset. *International Journal of Informatics and Information Systems*, *8*, 221–233. https://doi.org/10.47738/ijiis.v8i4.289.

[14] John, J.M.; Shobayo, O.; Ogunleye, B. (2023). An Exploration of Clustering Algorithms for Customer Segmentation in the UK Retail Market. *Analytics*, *2*, 809–823. https://doi.org/10.3390/analytics2040042.

[15] Salminen, J.; Blomqvist, O.J.; Jung, S.-C.; Jansen, B.J. (2023). How Can Algorithms Help in Segmenting Users and Audiences? A Systematic Mapping Study of User Segmentation Approaches. *Journal of Marketing Analytics*, *11*, 311–332. https://doi.org/10.1057/s41270-023-00235-5.

[16] Nugraheni, P.; Darma, E.S.; Muhammad, R. (2025). Adoption of Digital Technology and Financial Knowledge: Strategies for Achieving Sustainable Performance of MSMEs. *Journal of Risk and Financial Management*, *18*, 646. https://doi.org/10.3390/jrfm18110646.

[17] Jiao, L.; Zhang, X.; Li, X. (2022). Interpretable Fuzzy Clustering Using Unsupervised Feature Selection and Validity Indexes. *Information Sciences*, *607*, 507–526. https://doi.org/10.1016/j.ins.2022.03.040.

[18] Sun, Y.; Li, Q.; Zhang, Y. (2023). Research on Customer Lifetime Value Based on Machine Learning and RFM Model. *Heliyon*, *9*, e14267. https://doi.org/10.1016/j.heliyon.2023.e14267.

[19] Wongoutong, C. (2024). The Impact of Neglecting Feature Scaling in K-Means Clustering. *PloS ONE*, *19*, e0310839. https://doi.org/10.1371/journal.pone.0310839.

[20] Pingali, S.R.; Leung, C.T.; Krishnan, P. (2023). Digital Readiness of Small and Medium Enterprises in Emerging Economies. *Journal of Business Research*, *160*, 113816. https://doi.org/10.1016/j.jbusres.2023.113816.

[21] Davies, D.L.; Bouldin, D.W. (1979). A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-1*, 224–227. https://doi.org/10.1109/TPAMI.1979.4766909.

[22] Pereira, M.S.; Castro, B.S.; Cordeiro, B.A.; Silva, E.C.; Gonçalves, M.C. (2025). Factors of Customer Loyalty and Retention in the Digital Environment. *Journal of Theoretical and Applied Electronic Commerce Research*, *20*, 71. https://doi.org/10.3390/jtaer20020071.

[23] Kamyabi, M.; Özgit, H.; Ahmed, J.N. (2025). Sustaining Digital Marketing Strategies to Enhance Customer Engagement and Brand Promotion: Position as a Moderator. *Sustainability*, *17*, 3270. https://doi.org/10.3390/su17073270.

[24] Lee, Y.; Na, K.; Rhim, J.; Kim, E. (2025). Primary Determinants and Strategic Implications for Customer Loyalty in Pet-Related Vertical E-Commerce: A Machine Learning Approach. *Systems*, *13*, 175. https://doi.org/10.3390/systems13030175.