Jurnal Informasi dan Teknologi
https://jidt.org/jidt
2025 Vol. 7 No. 4 Hal: 25-32 e-ISSN: 2714-9730

Application of K-Means Clustering for Customer Segmentation on Sales Data in a Sheet Plastic Manufacturing Company

Kristiano Moniaga[1]✉, April Lia Hananto[2], Tukino[3], Shofa Shofiah Hilabi[4]

[1,2,3,4]Department of Information Systems, Faculty of Computer Science, Universitas Buana Perjuangan Karawang, Indonesia

si22.kristianomoniaga@mhs.ubpkarawang.ac.id

Abstract

Customer segmentation is applied to sales transaction data from a sheet plastic manufacturing company covering the 2019–2025 period and obtained from the company's Enterprise Resource Planning (ERP) system. This study aims to identify heterogeneous customer characteristics and generate actionable market segments based on Recency, Frequency, and Monetary (RFM) values using the K-Means Clustering algorithm. The methodology comprises systematic data cleaning, transaction aggregation, RFM calculation, feature normalization, cluster modeling, and determination of the optimal number of clusters through the Elbow Method and Silhouette Score. The final dataset consists of 46,372 transactions involving 1,223 active customers, with a cumulative transaction value of IDR 722.4 billion. The findings reveal five optimal clusters, validated by a Silhouette Score of 0.513, indicating reasonably good clustering quality and meaningful separation among customer segments. The segmentation identifies five distinct customer groups: Low Engagement (50%), characterized by limited transaction activity and requiring targeted reactivation strategies; Churn (37%), representing long-inactive customers at significant risk of disengagement and requiring structured re-engagement programs; Potential (11%), comprising active customers with substantial transaction values and strong development opportunities; Key Account (less than 1%), representing strategically important customers with the highest business contribution and requiring prioritized relationship management; and High Value (2%), consisting of loyal, profitable customers who should be retained through personalized loyalty initiatives. These findings demonstrate that integrating RFM analysis with K-Means Clustering provides a practical, data-driven approach to understanding customer heterogeneity in manufacturing markets. The resulting segmentation framework offers a strategic foundation for targeted retention initiatives, personalized promotional campaigns, improved customer relationship management, optimized allocation of sales resources, and more effective managerial decision-making based on measurable customer behavior and long-term value.

Keywords: *Customer Segmentation, RFM, K-Means Clustering, Data Mining, Manufacturing.*

1. Introduction

The rapid advancement of information technology and data analytics has encouraged companies to utilize transactional data as a strategic decision-making tool [1]. In the sheet plastic manufacturing industry, increasing market competition and dynamic customer demand require companies to better understand customer purchasing behavior. However, variations in purchase frequency, transaction value, and recency among customers often create challenges in determining service priorities and designing effective retention strategies [2]. Therefore, a systematic customer segmentation approach is necessary to support data-driven marketing decisions.

Customer segmentation is widely recognized as an effective technique for grouping customers based on purchasing behavior characteristics [3]. Previous studies have shown that the Recency, Frequency, and Monetary (RFM) model provides a quantitative framework for evaluating customer value, while clustering techniques such as K-Means Clustering are capable of forming homogeneous customer groups efficiently [4]. The main advantage of this approach lies in its simplicity, scalability for large datasets, and ease of interpretation. Nevertheless, its application in the sheet plastic manufacturing sector, particularly using long-term ERP-based transaction data, remains limited [4]. Several previous studies have applied RFM and K-Means for customer segmentation. Zhang et al. (2020) conducted customer segmentation using RFM and K-Means in the retail sector, successfully forming clear customer segments. Anitha and Patil (2022) implemented K-Means on sales data and effectively identified customer purchasing patterns. Hosseini et al. (2010) applied cluster analysis in the context of customer relationship management and divided customers based on transaction values. Namvar et al. (2010) performed transaction-based customer grouping using a two-phase clustering method with positive results. These studies confirm the effectiveness of RFM and K-Means for customer segmentation. However, most were conducted in retail or general contexts, while its application in the sheet plastic manufacturing industry using long-term ERP-based transaction data remains limited.

Receipt: 26-05-2025 | Revision: 29-10-2025 | Publish: 30-12-2025 | doi: 10.60083/jidt.vi0.740

2. Research Methods

This study employs a quantitative approach using computational analysis to address the research objectives of calculating RFM values and performing customer segmentation through clustering techniques. The research method is designed to systematically transform raw sales transaction data into meaningful customer segments that support marketing decision-making . The research framework consists of several main stages: data collection, data preprocessing, RFM calculation, data normalization, clustering process, and evaluation. The overall research framework is illustrated in Figure 1.

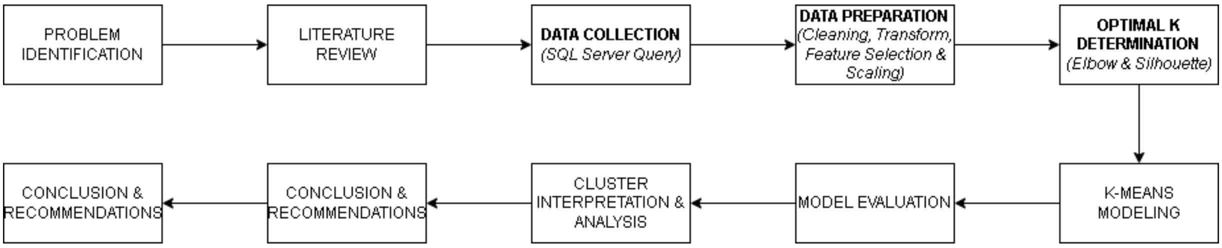

Figure 1. Research Framework

The dataset used in this study consists of sales transactions from 2019–2025 obtained from the company's ERP system. The preprocessing stage includes data cleaning, removal of incomplete records, and aggregation of transactions at the customer level. The RFM model is applied to measure customer behavior based on three variables: Recency (time since the last transaction), Frequency (total number of transactions), and Monetary (total transaction value) [5][6][7]. Recency is calculated as the difference between the reference date and the customer's most recent transaction date. Frequency represents the total number of purchase transactions during the observation period, while Monetary reflects the total purchase value. Before clustering, the RFM variables are normalized using Min-Max Scaling to ensure comparable value ranges. Customer segmentation is performed using the K-Means Clustering algorithm [4]. Unlike traditional implementations that rely on Euclidean Distance, this study applies an alternative distance measurement approach as discussed in the literature review to improve clustering accuracy for transactional data. The optimal number of clusters is determined using the Elbow Method and Silhouette Score to ensure model validity [8]. The clustering process is conducted step-by-step, including initialization of centroids, assignment of customers to the nearest centroid based on distance calculation, centroid updating, and iterative optimization until convergence is achieved [4][9]. The resulting clusters are then analyzed to interpret customer characteristics and support marketing strategy formulation.

3. Results and Discussion

3.1 Data Collection

The dataset used in this study was obtained from the Enterprise Resource Planning (ERP) system of a sheet plastic manufacturing company, specifically from the sales module stored in a Microsoft SQL Server database. Data extraction was performed using SQL queries on sales transactions with release status during the period of February 2019 to November 2025. The extraction process yielded 46,372 transaction records involving 1,360 unique customers. After data cleaning and validation, 1,223 active customers were identified for further analysis. The cumulative total sales value during the observation period reached Rp 722,416,248,310, indicating a high economic significance and providing a robust foundation for customer segmentation analysis.

3.2 RFM Variables Calculation

The RFM (Recency, Frequency, Monetary) model was applied to quantify customer behavior based on three key dimensions:

1. Recency: The number of days between the customer's last transaction and the reference date of November 6, 2025
2. Frequency: The total number of transactions made by each customer throughout the observation period
3. Monetary: The cumulative total purchase value for each customer from 2019 to 2025

Table 1 presents the top five customers with the highest transaction frequency, demonstrating the most intensive purchasing activity.

Table 1. Top Five Customers Based on Frequency

CUST_C ODE	CUST_NAME	FREQ	MONETARY		VARIATI ON	LAST ORDER	RECENCY (DAYS)
2000667	PT. FU*******	646	IDR	20,871,196,380	57	2025-11-06	0
2003805	PT. GA*******	462	IDR	40,548,638,120	119	2025-11-05	1

2000709	PT. IN*******	399	IDR	18,588,960,950	109	2025-10-22	15
2005029	CV. TI*******	385	IDR	5,671,014,721	190	2025-11-06	0
2000725	PT. JF*******	368	IDR	7,167,291,472	106	2025-10-23	14
2000935	PT. UN*******	359	IDR	12,721,675,000	41	2025-11-04	2
…	…	…	…		…	…	…

PT. Fu******* recorded the highest frequency with 646 transactions and a monetary value of Rp 20.87 billion, along with a recency of 0 days, indicating very recent activity. PT. Ga******* ranked second with 462 transactions and contributed the highest monetary value of Rp 40.55 billion. Table 2 highlights the top five customers based on monetary value, representing the company's primary revenue contributors.

Table 2. Top Five Customers Based on Monetary Value

CUST_C ODE	CUST_NAME	FREQ	MONETARY		VARIAT ION	LAST ORDER	RECENC Y (DAYS)
2003805	PT. GA*******	462	IDR	40,548,638,120	119	2025-11-05	1
2000738	PT. KE********	64	IDR	23,213,195,266	53	2024-10-01	401
2000667	PT. FU*******	646	IDR	20,871,196,380	57	2025-11-06	0
2000936	PT. UN*******	176	IDR	18,645,458,160	17	2025-10-24	13
2000709	PT. IN*******	399	IDR	18,588,960,950	109	2025-10-22	15
2001063	PT. BE*******	156	IDR	17,911,813,952	74	2025-11-05	1
…	…	…	…		…	…	…

While PT. Ga******* leads with the highest monetary contribution, PT. Ke******* shows a concerning recency of 401 days, indicating a prolonged period of inactivity despite a high historical monetary value. Table 3 presents customers with the lowest recency values (0 days), representing the most recently active customers as of the reference date.

Table 3. Top Five Customers Based on Lowest Recency (Most Active)

CUST_COD E	CUST_NAME	FREQ	MONETARY		VARIATIO N	LAST ORDER	RECENC Y (DAYS)
2000667	PT. FU*******	646	IDR	20,871,196,380	57	2025-11-06	0
2003299	PT. VA*******	135	IDR	9,035,150,215	57	2025-11-06	0
2004109	PT. MA*******	199	IDR	7,389,177,500	88	2025-11-06	0
2005029	CV. TI*******	385	IDR	5,671,014,721	190	2025-11-06	0
2003210	PT. TO*******	76	IDR	3,995,100,000	4	2025-11-06	0
2000607	PT. BU*******	150	IDR	3,245,885,700	18	2025-11-06	0
…	…	…	…		…	…	…

All customers in this group show strong recent engagement, with PT. Fu******* demonstrating both high frequency and low recency, reinforcing its position as a key account.

3.3 Data Normalization

Due to significant differences in measurement scales among RFM variables (Recency in days, Frequency in transaction counts, and Monetary in billions of rupiah), normalization was required to prevent any single variable from dominating the Euclidean distance calculations in the K-Means algorithm [4][10]. The StandardScaler method from the scikit-learn library was applied, transforming each variable to have a mean of zero and a standard deviation of one. This ensures balanced contribution from all variables during the clustering process. Table 4 shows sample normalized values for selected customers.

Table 4. Sample of Normalized RFM Values Using Standard Scaler

CUST_ CODE	CUST_NAME	SCALED_RECE NCY_DAYS	SCALED_TOTAL_ TRANSAKSI	SCALED_TOTAL_ AMOUNT	SCALED_VARI ASI_PRODUK
2000112	PT. AB*******	-1.043620887	1.985817817	0.13130049	0.107622357
2000113	BA*******	1.953982749	-0.41027233	-0.270006725	-0.435629409
2000114	PT. IN*******	0.057722793	2.726427499	3.450845897	4.919280851
2000117	PT. PE*******	-0.884637631	2.160078919	0.21127737	2.047807233
2000118	PT.DA*******	0.228245156	-0.366707054	-0.26943096	-0.202807224
2000123	PT. SI*******	1.903979951	0.199641526	0.095916493	0.030014961

2000124	PT. EH*******	1.294971514	-0.344924417	-0.22293216	-0.435629409
…	…	…	…	…	…

1. Negative scaled recency indicates customers with more recent transactions than average
2. Positive scaled frequency and monetary indicates above-average transaction intensity and spending
3. Positive scaled product variation indicates customers purchasing a wider variety of products

For example, PT. In******* shows exceptionally high scaled frequency (2.73) and monetary (3.45), positioning it as a high-value customer with diverse product purchases.

3.4 Determination of the Optimal Number of Clusters

The normalized dataset was subsequently used as input for the K-Means clustering algorithm to ensure unbiased and accurate customer segmentation [8]. To determine the optimal number of clusters, the Elbow Method was applied by evaluating the Within-Cluster Sum of Squares (WCSS) across different cluster values. The visualization of this evaluation is presented in Figure 2.

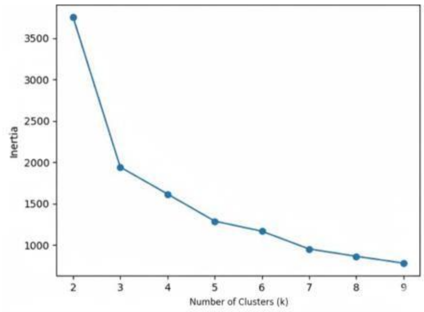

Figure 2. Elbow Method for Optimal Cluster Determination

The Elbow curve in Figure 2 shows a sharp decline in the Within-Cluster Sum of Squares (WCSS) value from cluster 1 to cluster 3, followed by a more gradual decrease afterward. This pattern indicates that adding more clusters beyond three does not significantly improve model compactness. Therefore, three clusters were selected as the optimal number for customer segmentation in this study. This result ensures a balance between model simplicity and segmentation effectiveness.

3.5 Customer Segmentation Results

After identifying the optimal number of clusters, the K-Means algorithm was implemented on the normalized RFM dataset. The distribution of customers across the resulting clusters is illustrated in Figure 3.

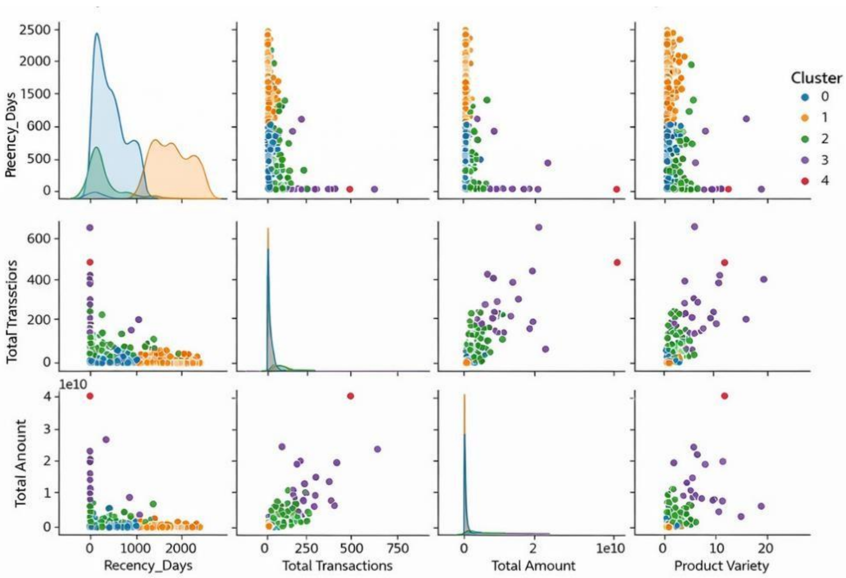

Figure 3. Elbow Method for Optimal Cluster Determination

The visualization in Figure 3 illustrates the distribution of customers across clusters based on Recency, Frequency, Monetary, and Product Variation variables. The plot demonstrates clear differentiation among the identified clusters, indicating that the K-Means algorithm successfully grouped customers with similar transactional behavior patterns. Customers in certain clusters exhibit high frequency and high monetary values combined with low recency, representing loyal and high-value customers who actively contribute to company revenue. Other clusters display moderate transaction frequency and spending levels, indicating potential customers who may increase their contribution with appropriate marketing strategies. Additionally, some clusters are characterized by high recency and low transaction intensity, suggesting customers with declining engagement or potential churn risk.The separation patterns observed across multiple variable combinations confirm that the clustering model effectively captures behavioral differences among customers. These results provide a strong analytical foundation for designing targeted retention programs, promotional strategies, and sales optimization efforts.

3.6 Luster Characteristics Analysis

To further examine the behavioral differences among customer segments, the average values of Recency, Frequency, Monetary, and Product Variation were calculated for each cluster. This analysis provides a clearer understanding of the purchasing patterns that define each segment.

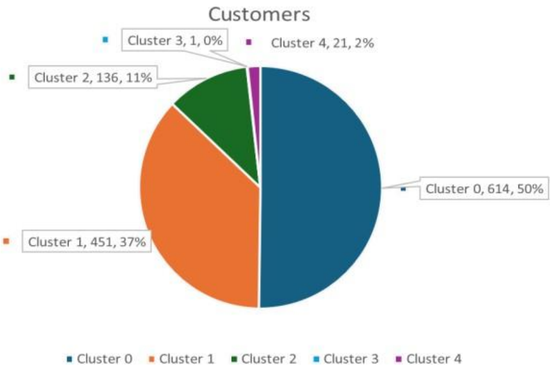

Figure 4. Customer Distribution by Cluster

Figure 4 illustrates the comparative average values of Recency, Frequency, Monetary, and Product Variation across all identified clusters. The differences in these values highlight distinct behavioral characteristics among customer groups. Clusters with low average Recency and high Frequency and Monetary values indicate highly engaged and loyal customers who make frequent purchases and contribute significantly to overall revenue. These customers represent the high-value segment and should be prioritized in retention strategies.Clusters showing moderate Frequency and Monetary values reflect potential customers who demonstrate stable purchasing behavior but still offer opportunities for growth through targeted marketing and promotional programs. In contrast, clusters characterized by high Recency and low Frequency and Monetary values indicate customers with declining transaction activity. This group represents at-risk customers who may require reactivation efforts to prevent churn. Overall, the variation in RFM patterns across clusters confirms that the segmentation model effectively differentiates customer behavior and provides actionable insights for strategic marketing decision-making.

3.7 Cluster Characteristics Analysis

To further examine the behavioral differences among customer segments, the average values of Recency, Frequency, Monetary, and Product Variation were calculated for each cluster. This analysis provides a clearer understanding of the purchasing patterns that define each segment.

3.7.1 Mean Values of RFM by Cluster

To gain a deeper understanding of the behavioral patterns that distinguish each customer segment, the average values of Recency, Frequency, Monetary, and Product Variation were calculated for every cluster. This analysis provides quantitative insight into the distinctive characteristics of each group and serves as the foundation for interpreting the meaning of each cluster in a business context. Table 5 presents the mean values of all RFM variables for each of the five clusters identified through the K-Means algorithm.

Table 5. Mean Values of RFM Variables by Cluster

Cluster 0	Average Recency (Days)	Average Total Transactions		Average Total Amount (Revenue)
	354.59	9.4	IDR 197,415,523	3.66
1	1,758.61	5.18	IDR 100,496,300	3.84
2	210.46	77.16	IDR 2,076,049,599	18.07
3	1.00	462.00	IDR 40,548,638,120	119.00

4	118.86	247.38	IDR 11,094,607,903	73.10
Cluster	Average Recency (Days)	Average Total Transactions	Average Total Amount (Revenue)	Average Product Variety

1. High-Value Segments (Clusters 3 and 4) These clusters demonstrate low recency (1-118 days), high frequency (247-462 transactions), and high monetary value (Rp 11-40 billion). Despite representing only 2% of total customers, they contribute most of the company revenue and exhibit strong loyalty with diverse product purchases.
2. Potential Segment (Cluster 2) This cluster shows moderate recency (210 days) with relatively high frequency (77 transactions) and monetary value (Rp 2 billion). These 136 customers (11%) have good engagement and significant growth potential.
3. Low-Engagement Segments (Clusters 0 and 1) These clusters account for 1,065 customers (87%) but show concerning patterns. Cluster 0 has high recency (355 days) with low frequency (9 transactions) and low monetary value (Rp 197 million). Cluster 1 exhibits extremely high recency (1,759 days) with the lowest frequency (5 transactions) and monetary value (Rp 100 million), indicating long-term inactive customers at high risk of churn.

3.7.2 Cluster Interpretation and Labeling

Based on the mean values analysis in Table 5, each cluster was interpreted and assigned a label that reflects its behavioral characteristics. This labeling helps translate the statistical segmentation into meaningful business categories that can guide marketing strategies and customer relationship management.

Table 6. Cluster Interpretation and Labeling

Cluster	Key Characteristics	Segmentation Label	Brief Description
0	High Recency, Low Frequency, Low Monetary	Low Engagement Customer	Customers with low activity levels who transact infrequently.
1	Very High Recency, Very Low Frequency & Monetary	Churned Customer	Customers who have been inactive for a long time and are likely to leave.
2	High Frequency, Large Monetary, High product variety	Potential Customer	Active customers with significant growth potential.
3	Very Low Recency, Highest Frequency & Monetary	Key Account / Loyal Customer	Prime customers who provide the largest revenue contribution.
4	High Frequency, Large Monetary, High product variety	High-Value Customer	Loyal customers with consistently high transaction values.

1. High-Value Segments (Clusters 3 and 4): These 22 customers (2%) are the most loyal with low recency, high frequency, and high monetary values. They contribute the majority of revenue and require priority retention strategies
2. Potential Segment (Cluster 2): These 136 customers (11%) show good engagement with growth potential. They should be targeted for development programs to increase their value.
3. Low-Performing Segments (Clusters 0 and 1): These 1,065 customers (87%) have high recency and low transaction activity. They need reactivation strategies or re-engagement programs to prevent churn.

3.7.3 Strategic Implications

The segmentation results in Table 6 provide actionable insights for developing targeted marketing strategies. Based on the characteristics of each cluster, specific strategic approaches can be formulated to optimize customer relationship management and resource allocation. Table 7 summarizes the strategic implications for each customer segment.

Table 7. Strategic Implications by Customer Segment

Cluster	Label	Strategic Focus	Objective
0	Low Engagement	Re-activation	Increase transaction frequency
1	Churn	Reactivation	Reduce customer churn / attrition
2	Potential	Development	Increase customer lifetime value
3	Key Account	VIP Retention	Secure primary revenue stream
4	High Value	Loyalty	Maintain high-value customer base

1. Cluster 3 (Key Account): Provide dedicated account manager and priority service to maintain this single highest-revenue customer.

2. Cluster 4 (High Value): Implement loyalty programs and volume-based rewards, plus enhanced after-sales service to retain these 21 valuable customers.
3. Cluster 2 (Potential): Offer product bundling and consultation services with loyalty incentives to increase transaction value for these 136 customers.
4. Cluster 0 (Low Engagement): Apply reactivation strategies including light promotions and order reminders to increase transaction frequency for these 614 customers.
5. Cluster 1 (Churn): Conduct special re-engagement programs with direct communication and purchase incentives to reactivate these 451 customers or evaluate for removal from active marketing databases.

4. Conclusion

This study successfully implemented the K-Means clustering algorithm combined with RFM analysis to segment customers in a sheet plastic manufacturing company. Based on 46,372 sales transactions from 1,223 active customers over the 2019–2025 period, with a total value of IDR 722.4 billion, five optimal clusters were identified and validated by the Elbow method and a Silhouette Score of 0.513. The resulting segments consist of Cluster 0 (Low Engagement Customer) with high recency, low frequency and low monetary value, comprising 50% of customers. Cluster 1 (Churn Customer) shows very high recency with very low frequency and monetary value, accounting for 37% of customers. Cluster 2 (Potential Customer) exhibits moderate recency with high frequency, high monetary value, and high product variation, representing 11% of customers. Cluster 3 (Key Account) has very low recency with the highest frequency and monetary value, representing less than 1% of customers. Cluster 4 (High Value Customer) shows low recency with high frequency, high monetary value, and high product variation, comprising 2% of customers.

These segments provide a foundation for targeted marketing strategies. Low engagement and churn customers require reactivation and re-engagement programs, potential customers need development initiatives, while key account and high value customers deserve priority retention efforts. By implementing differentiated approaches, the company can optimize resource allocation, improve customer retention, and maximize revenue.For future research, it is recommended to compare K-Means with other clustering algorithms such as hierarchical clustering or DBSCAN, and incorporate additional variables including profit margins, seasonal patterns, or product categories to enrich the analysis.

Acknowledgments

The authors would like to express their sincere gratitude to PT Pl******* for providing access to the sales transaction data used in this research.Deepest appreciation is extended to the supervising lecturers, Prof. Ir. April Lia Hananto, Ph.D, Tukino, M.M.S.I, and Dr.Shofa Shofiah Hilabi., M.Kom for their invaluable guidance, support, and constructive feedback throughout the study.Special thanks are also due to Mr. Geo Septian, an alumnus of the Information Engineering Department, Universitas Buana Perjuangan Karawang, who currently serves as Head of IT Department at PT Pl*******. His generous assistance in data extraction and technical support was instrumental to the successful completion of this research.The authors also acknowledge the Faculty of Computer Science, Universitas Buana Perjuangan Karawang, for the academic support and facilities provided.

References

[1] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-Means Clustering Algorithms: A Comprehensive Review," IEEE Access, vol. 11, pp. 50864–50903, 2023. doi: 10.1109/ACCESS.2023.3278345.

[2] D. Y. Ryu, Y. K. Ko, and Y. D. Ko, "RFM analysis for profiling profitable customers based on characteristics of the hotel industry," International Journal of Hospitality Management, vol. 129, p. 104176, 2025. doi: 10.1016/j.ijhm.2025.104176.

[3] A. L. Hananto et al., "Analysis of Drug Data Mining with Clustering Technique Using K-Means Algorithm," Journal of Physics: Conference Series, vol. 1908, no. 1, p. 012022, 2021. doi: 10.1088/1742-6596/1908/1/012022.

[4] J. T. Wei, S. Y. Lin, and H. H. Wu, "A review of the application of RFM model," African Journal of Business Management, vol. 4, no. 19, pp. 4199–4206, 2010. [Online]. Available: https://academicjournals.org/journal/AJBM/article-abstract/8F0A3F221234

[5] M. Khajvand, K. Zolfaghar, S. Ashoori, and S. Alizadeh, "Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study," Procedia Computer Science, vol. 3, pp. 57–63, 2011. doi: 10.1016/j.procs.2010.12.011.

[6] A. A. Raorane, R. V. Kulkarni, and B. D. Jitkar, "Association rule – Extracting knowledge using market basket analysis," Research Journal of Recent Sciences, vol. 1, no. 2, pp. 19–27, 2012. [Online]. Available: http://www.isca.in/rjrs/archive/v1/i2/4.ISCA-RJRS-2012-009.pdf

[7] S. S. Singh and N. C. Chauhan, "K-means v/s K-medoids: A comparative study," in National Conference on Recent Trends in Engineering & Technology, 2011. [Online]. Available: https://www.researchgate.net/publication/267693481

[8] A. K. Jain, "Data clustering: 50 years beyond K-means," Pattern Recognition Letters, vol. 31, no. 8, pp. 651–666, 2010. doi: 10.1016/j.patrec.2009.09.011.
[9] P. Anitha and M. M. Patil, "RFM model for customer purchase behavior using K-Means algorithm," Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 5, pp. 1785–1792, 2022. doi: 10.1016/j.jksuci.2019.12.011.
[10] S. M. S. Hosseini, A. Maleki, and M. R. Gholamian, "Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty," Expert Systems with Applications, vol. 37, no. 7, pp. 5259–5264, 2010. doi: 10.1016/j.eswa.2009.12.070.
[11] M. J. Shaw, C. Subramaniam, G. W. Tan, and M. E. Welge, "Knowledge management and data mining for marketing," Decision Support Systems, vol. 31, no. 1, pp. 127–137, 2001. doi: 10.1016/S0167-9236(00)00123-8.
[12] E. W. T. Ngai, L. Xiu, and D. C. K. Chau, "Application of data mining techniques in customer relationship management: A literature review and classification," Expert Systems with Applications, vol. 36, no. 2, pp. 2592–2602, 2009. doi: 10.1016/j.eswa.2008.02.021.
[13] S. F. Zhang, J. L. Wang, and X. Y. Liu, "A new customer segmentation model based on RFM and improved K-means," Advanced Materials Research, vol. 765–767, pp. 1483–1487, 2013. doi: 10.4028/www.scientific.net/AMR.765-767.1483.
[14] C. H. Cheng and Y. S. Chen, "Classifying the segmentation of customer value via RFM model and RS theory," Expert Systems with Applications, vol. 36, no. 3, pp. 4176–4184, 2009. doi: 10.1016/j.eswa.2008.04.003.
[15] D. L. Olson and D. D. Wu, "RFM and data mining for predictive customer relationship management," in Enterprise Risk Management Models, Springer, 2010, pp. 117–130. doi: 10.1007/978-3-642-11480-5_8.
[16] M. Namvar, M. R. Gholamian, and S. KhakAbi, "A two phase clustering method for intelligent customer segmentation," in 2010 International Conference on Intelligent Systems, Modelling and Simulation, 2010, pp. 215–219. doi: 10.1109/ISMS.2010.49.
[17] G. Punj and D. W. Stewart, "Cluster analysis in marketing research: Review and suggestions for application," Journal of Marketing Research, vol. 20, no. 2, pp. 134–148, 1983. doi: 10.1177/002224378302000204.
[18] S. S. Hosseini, M. R. Gholamian, and A. Maleki, "Customer segmentation based on RFM model and clustering techniques," International Journal of Industrial Engineering & Production Research, vol. 21, no. 3, pp. 135–143, 2010. [Online]. Available: http://ijiepr.iust.ac.ir/article-1-69-en.html
[19] J. Wu and Z. Lin, "Research on customer segmentation model based on improved K-means algorithm," Journal of Physics: Conference Series, vol. 1650, p. 032075, 2020. doi: 10.1088/1742-6596/1650/3/032075.
[20] Y. Zhang, Y. Zhang, and A. Swidan, "Customer segmentation in retail based on RFM and K-means," in Proceedings of the 2020 3rd International Conference on Big Data Technologies, 2020, pp. 45–50. doi: 10.1145/3422713.3422723.