International Journal Software Engineering and Computer Science (IJSECS)
6 (3), 2026, 900-907
Published Online December 2026 in IJSECS (http://www.journal.lembagakita.com/index.php/ijsecs)
P-ISSN: 2776-4869, E-ISSN: 2776-3242. DOI: https://doi.org/10.35870/ijsecs.v6i3.8324.

RESEARCH ARTICLE Open Access

Customer Segmentation Analysis Using the K-Means Method on Order Data from Aisyah Catering MSME

Aisha Yulia Putri [1*], **Sri Lestari** [2], **Mesra Betty Yel** [3]

[1*,2,3] Department of Information Systems, Faculty of Computer Technology, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, East Jakarta City, Special Capital Region of Jakarta, Indonesia.

Article History:

Received: August 7, 2026.
Revised: August 11, 2026.
Accepted: August 27, 2026.
Available online: September 20, 2026.
Published: December 1, 2026.

***Corresponding Author:**
aishayulia7@gmail.com

Abstract: Aisyah Catering is a micro, small, and medium enterprise (MSME) providing catering services that relies on manual order recording, which limits the business owner's ability to examine customer purchasing patterns and formulate targeted marketing strategies. This study aims to segment Aisyah Catering's customers based on order volume using the K-means clustering algorithm. The methodology follows the CRISP-DM framework, beginning with the collection of 197 transaction records digitized from physical receipts, followed by data preprocessing using Altair AI Studio to yield 187 valid records. In the modeling stage, K-means was applied to order portion volume as the primary quantitative attribute. Cluster selection was evaluated using the Elbow method and the Davies-Bouldin Index (DBI). The evaluation identified k=3 as the most interpretable cluster configuration, yielding a DBI value of 0.449. The three resulting segments comprise a small-scale segment (49.2% of customers, averaging 360.11 portions), a medium-scale segment (33.7% of customers, averaging 717.78 portions), and a large-scale segment (17.1% of customers, averaging 1,045.31 portions). Volume contribution analysis indicates that the large-scale segment, despite containing the fewest customers, accounted for 29.9% of total order volume, nearly matching the contribution of the small-scale segment. Consequently, these segmentation findings provide an empirical foundation for Aisyah Catering to design targeted marketing strategies, streamline raw material procurement, and customize service delivery according to distinct customer profiles.

Keywords: Customer Segmentation; K-Means; Clustering; Catering MSME; Data Mining.

1. Introduction

Micro, small, and medium enterprises (MSMEs)—productive economic entities managed independently by individuals or business groups—serve as a cornerstone of the Indonesian economy. According to official reports from the Ministry of Cooperatives and SMEs of the Republic of Indonesia, the contribution of MSMEs to the national gross domestic product consistently ranged between 58% and 61% throughout the 2020–2024 period (Kementerian Koperasi dan UKM RI, 2024). This substantial macroeconomic share indicates that MSMEs represent the primary foundation absorbing labor and sustaining local employment across diverse regions. Within this ecosystem, the culinary and catering service industry represents a rapidly evolving sector, providing food and beverage solutions for personal celebrations, weddings, corporate functions, and public institution gatherings. Aisyah Catering exemplifies this operational landscape, having operated continuously since 2007 while coordinating approximately 35 personnel and specialized vendor networks. The enterprise caters to heterogeneous consumer profiles, handling event orders that vary substantially from small private events to large gatherings requiring thousands of servings.

Despite managing substantial business volume, operational data management at Aisyah Catering remains entirely dependent on conventional methods, primarily paper receipts and handwritten invoice archives. This conventional record-keeping creates considerable operational vulnerabilities, including physical document degradation, manual calculation errors, and protracted retrieval times for past records. More fundamentally, physical transaction logs prevent business owners from obtaining an integrated, longitudinal understanding of historical purchasing behavior. Without systematic data consolidation, identifying transaction frequencies, peak ordering cycles, and distinctive customer volume segments becomes largely intuitive rather than empirical. This deficiency severely limits managerial effectiveness, particularly in formulating customized marketing outreach and anticipating resource allocations, even though granular behavioral insights are critical for executing targeted promotional strategies (Nurlathifa & Sucahyati, 2025).

This operational constraint is widely observed throughout the broader MSME sector in Indonesia. Empirical findings by Dewobroto and Shania (2023) indicate that among tens of millions of registered Indonesian MSMEs, the majority have yet to integrate digital technology effectively into daily operations, particularly in structured transaction recording. Consequently, enterprise owners frequently deploy uniform, generalized marketing programs to an otherwise heterogeneous consumer base, disregarding substantial differences in individual purchasing scales and consumption preferences. To address these limitations, customer segmentation through data mining techniques, specifically clustering algorithms, provides an established analytical approach. The K-means algorithm remains widely utilized in MSME customer analytics owing to its conceptual simplicity, linear scalability, and high computational efficiency. Previous studies, such as Ardi *et al.* (2023), confirmed that K-means effectively categorizes fashion retail consumers into cohesive behavioral groups without requiring pre-labeled training data. In culinary applications, Kristanti *et al.* (2024) implemented K-means within the recency, frequency, and monetary (RFM) framework to classify retail shoppers, while Djun *et al.* (2024) demonstrated the utility of combining the Elbow method, Silhouette score, and Davies-Bouldin Index to validate optimal cluster boundaries.

Nevertheless, existing clustering literature predominantly relies on multidimensional RFM frameworks that presuppose comprehensive, unbroken temporal tracking across recency, frequency, and monetary metrics. In actual field conditions, early-stage digitalizing MSMEs with paper-based archives rarely possess complete historical timelines, rendering full RFM parameterization practically unfeasible. This structural discrepancy creates a relevant research gap: establishing whether robust, actionable customer segmentation can be achieved using a streamlined single quantitative variable that directly reflects operational capacity and commercial contribution. In catering enterprises, order portion volume represents the most fundamental indicator of transaction magnitude, production load, and revenue generation. Selecting this primary metric aligns with feature selection principles, prioritizing domain relevance over parameter quantity.

Based on these considerations, this study aims to: 1) digitize and preprocess physical transaction records from Aisyah Catering to prepare a clean dataset for clustering, 2) apply the K-means algorithm to group customers based on order portion volume, 3) determine the optimal cluster configuration through the combined evaluation of the Elbow method and the Davies-Bouldin Index, and 4) analyze the volume contributions and behavioral characteristics of each resulting segment to formulate practical marketing and operational strategies for Aisyah Catering. This study contributes methodologically by demonstrating that empirical customer segmentation remains reliable and actionable for resource-constrained MSMEs utilizing a streamlined single-variable model, while practically delivering a data-driven framework for raw material procurement, kitchen capacity planning, and differentiated customer service.

2. Literature Review

2.1 Related Work

The application of the K-means clustering algorithm for MSME customer segmentation has received increasing empirical attention. Danuwinata and Tjen (2023) implemented K-means on transaction data from Lily Cakes in Pontianak, demonstrating that grouping customers by transaction patterns provided an empirical foundation for structured promotional campaigns. Ardi *et al.* (2023) utilized the algorithm within a fashion MSME in Solo, showing that K-means effectively clusters customer purchasing profiles based on behavioral similarities without requiring labeled training datasets. Within the culinary sector, Nurlathifa and Sucahyati (2025) applied K-means under the recency, frequency, and monetary (RFM) model to coffee shop patrons, deriving actionable segments to support customer retention programs. In an algorithmic comparison, Ariska *et al.* (2024) evaluated K-means against K-medoids on MSME data from Pagar Alam City, concluding that K-means maintained superior computational efficiency despite greater sensitivity to extreme values.

In the retail domain, Kristanti *et al.* (2024) applied K-means by combining the Elbow method, Silhouette score, and Gap Statistic to determine optimal cluster boundaries, identifying distinct purchasing characteristics across four customer groups. Similarly, Djun *et al.* (2024) deployed K-means on MSME transaction data using the RFM framework and evaluated cluster validity through a tripartite combination of the Elbow method, Silhouette score, and the Davies-Bouldin Index (DBI). While tripartite validation provides multifaceted statistical verification, literature demonstrates that paired complementary metrics—such as the Elbow method for intra-cluster compactness and the Davies-Bouldin Index for inter-cluster separation—offer robust evaluation when applied consistently and interpreted with domain awareness. Unlike preceding studies that relied on multivariate RFM models requiring dense transaction histories, the present study focuses on a single core attribute: order portion volume. This streamlined approach addresses the practical operational context of Aisyah Catering, where paper-based archives preclude complete longitudinal recency and frequency tracking. Consequently, this study demonstrates that business-relevant customer segmentation remains achievable using transaction volume alone, provided the chosen attribute directly reflects operational capacity and commercial contribution.

2.2 Customer Segmentation and the STP Framework

Customer segmentation involves dividing a heterogeneous market into distinct, homogeneous subgroups sharing similar purchasing behaviors or transactional scale, enabling organizations to tailor operational resources and marketing initiatives to specific customer demands. Segment profiles typically serve as the empirical foundation for executing the Segmenting, Targeting, and Positioning (STP) framework. As Shoib *et al.* (2024) demonstrated that applying the STP framework to ceramic craft MSMEs assisted business owners in establishing distinct target groups and positioning products competitively. Similarly, Manggu and Beni (2021) confirmed that structured STP adoption across MSMEs in Bengkayang City contributed substantially to promotional efficiency and market engagement.

2.3 K-Means Clustering Algorithm

K-means is a partitioning-based unsupervised clustering algorithm designed to classify unlabeled data into K predetermined clusters based on proximity to iteratively updated centroids. The algorithm initializes by selecting K arbitrary cluster centroids, assigns each data point to its nearest centroid using Euclidean distance metrics, and subsequently recalculates centroid locations as the mean coordinates of all assigned data points. This assignment and update cycle repeats iteratively until centroid coordinates stabilize and mathematical convergence is achieved (Syam, 2017). A fundamental requirement in K-means implementation is selecting an optimal number of clusters (K), as inappropriate parameterization risks producing unrepresentative partitions. Widiyanto and Witanti (2021) recommend the Elbow method, which tracks the reduction in within-cluster sum of squares (WCSS) or average within-centroid distance across sequential values of K to detect the marginal inflection point. Complementarily, the Davies-Bouldin Index evaluates the ratio between intra-cluster dispersion and inter-cluster separation, where lower values signify dense, well-isolated clusters (Djun *et al.*, 2024).

2.4 Conceptual Validity of Single-Attribute Clustering

Algorithmic K-means formulations do not require a minimum multidimensional feature vector to function properly. The mathematical mechanism operates by calculating Euclidean distances across data coordinates within an arbitrary n-dimensional space, which naturally encompasses one-dimensional real coordinate spaces (n=1). Under a single quantitative attribute, Euclidean distance calculations simplify to absolute one-dimensional intervals:

$$d(x,c) = \sqrt{(x-c)^2} = x - c \tag{1}$$

Consequently, one-dimensional K-means partitions continuous numerical values along the real number line into K discrete intervals optimized to minimize variance around local cluster centroids. Although single-attribute clustering restricts multidimensional behavioral profiling relative to multivariate models such as RFM, it remains mathematically valid provided the selected feature accurately represents the operational dimension under investigation. In the context of catering operations, order portion volume represents the definitive metric governing kitchen production load, perishable inventory planning, and transaction scale. This single-attribute formulation aligns directly with data mining feature selection principles, which mandate selecting parameters based on direct domain utility rather than feature quantity.

3. Research Methodology

3.1 Research Framework

This study adopted a descriptive quantitative data mining framework based on the Cross-Industry Standard Process for Data Mining (CRISP-DM). The framework encompasses six systematic phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. This structured approach was selected to provide rigorous guidance from problem identification through the extraction of actionable business insights. Primary data collection involved two approaches: semi-structured interviews with the proprietor of Aisyah Catering to comprehend core business processes, and the systematic collection of historical transaction records from physical order receipts. All physical order slips were optically captured using digital cameras and manually transcribed into structured tabular spreadsheets using Microsoft Excel. This process generated an initial dataset comprising 197 transaction records containing customer identification codes, service package types, ordered portion quantities, event dates, and operational remarks. To comply with ethical research standards and maintain commercial confidentiality, all customer identifying details were anonymized using unique alphanumeric identifiers.

3.2 Data Preprocessing

Data preprocessing was executed within Altair AI Studio utilizing a sequential workflow comprising Retrieve, Select Attributes, and Filter Examples operators. The Retrieve operator imported the 197 digitized records into the local analytical repository. Subsequently, the Select Attributes operator isolated order portion volume as the sole modeling variable, as it directly quantified customer transaction scale and kitchen production demand. Contextual attributes—namely customer identifier, service package type, and event date—were excluded from distance calculations because non-numeric and descriptive metadata do not contribute directly to geometric clustering spaces. The Filter Examples operator was configured with the is not missing condition to eliminate incomplete entries lacking portion metrics. This filtering step reduced the initial dataset from 197 records to 187 fully validated transactions. An outlier evaluation on portion quantities revealed several high-magnitude transactions, including single orders reaching 1,200 portions. These data points were retained under the empirical assumption that they represented legitimate large-scale institutional or formal banquet bookings. However, because these extreme observations were preserved based on operational plausibility rather than direct manual re-verification against original physical receipts, minor undocumented recording discrepancies might persist. This condition represents an acknowledged methodological boundary, and subsequent studies should incorporate direct cross-validation protocols for high-magnitude data points prior to modeling.

3.3 K-Means Clustering Implementation

The K-means algorithm was executed on the 187 validated records using the Clustering operator in Altair AI Studio. To determine the most robust partition, sequential experimental trials were conducted across cluster values ranging from K=2 to K=5. Partitions were evaluated using the Elbow method based on average within-centroid distance and the Davies-Bouldin Index (DBI) to assess structural cluster separation. Within Altair AI Studio, DBI values are presented as negative metrics by software convention to maintain a universal "higher-is-better" optimization standard; consequently, real DBI values were derived from absolute magnitudes. To ensure algorithmic reproducibility, the operational parameter configurations for the Clustering operator are detailed in Table 1. Distance calculations between data coordinates and cluster centroids utilized standard Euclidean distance. Because the feature space comprised exclusively a single numeric attribute, data normalization or standardization was omitted, as scale harmonization is strictly required only when handling multidimensional variables with disparate measurement scales.

Tabel 1. Clustering Operator Parameter Configuration (K-Means)

Parameter	Value	Description
K	2 / 3 / 4 / 5	Number of cluster partitions systematically evaluated
Measure Type	Numerical Measures	Metric space configuration for numerical variables
Numerical Measure	Euclidean Distance	Distance computation metric between data coordinates and centroids
Max Runs	10 (default)	Maximum randomized centroid initialization attempts
Max Optimization Steps	100 (default)	Maximum optimization iterations permitted per execution run

3.4 Hardware and Software Tools

The computational pipeline utilized Altair AI Studio as the primary visual analytics environment, executing data mining workflows without requiring custom algorithmic scripts, supported by Microsoft Excel for primary digitization and tabular structuring. The primary empirical dataset comprised digitized order records from Aisyah Catering, encompassing customer codes, package categories, order volumes, and event execution schedules. As established during preprocessing, only the numerical order portion volume was introduced into the clustering engine, while remaining attributes served purely descriptive purposes.

3.5 Evaluation Metrics

Because K-means operates as an unsupervised learning paradigm lacking ground-truth class labels, conventional classification metrics (such as precision, recall, or accuracy) are inapplicable. Methodological evaluation followed a three-tier framework: initial data validation to verify structural completeness; algorithmic performance evaluation using complementary Elbow curves and DBI metrics to verify intra-cluster compactness and inter-cluster separation; and domain-level post-clustering analysis to profile segment characteristics for strategic marketing deployment. Quantitative validation was deliberately bounded to the Elbow method and the Davies-Bouldin Index without incorporating the Silhouette score or alternative metrics. This deliberate constraint reflects the established analytical scope rather than an omission, and findings should be interpreted within this dual-metric context. Subsequent investigations seeking wider statistical validation may integrate the Silhouette score as an auxiliary index.

4. Results and Discussion

4.1 Results

Experimental trials were conducted across four variations of K (K=2 to K=5) using the Clustering and Performance operators in Altair AI Studio. The resulting average within-centroid distance metrics for each evaluated K value are presented in Table 2. The intra-cluster distance exhibited a monotonic decline as the number of clusters increased, decreasing from 20,637.799 at K=2 to 11,854.464 at K=3, followed by 6,522.207 at K=4, and reaching 3,266.538 at K=5.

Tabel 2. Evaluation of Average Within-Centroid Distance

K	Average within-centroid distance
2	20,637.799
3	11,854.464
4	6,522.207
5	3,266.538

Cluster separation and structural compactness were subsequently assessed using the Davies-Bouldin Index (DBI). The empirical DBI results across the experimental iterations are summarized in Table 3. The absolute DBI scores were 0.302 for K=2, 0.449 for K=3, 0.552 for K=4, and 0.503 for K=5.

Table 3. Evaluation of Davies-Bouldin Index

K	Davies-Bouldin Index
2	0.302
3	0.449
4	0.552
5	0.503

Based on these evaluative criteria and practical segmentation requirements, $k=3$ was chosen for detailed behavioral examination, partitioning the customer base of Aisyah Catering into three operational segments. The distribution of customers and the mean portion volume per cluster are presented in Table 4.

Table 4. Customer Segmentation Profiles

Cluster	Number of customers	Percentage (%)	Mean portion volume
Cluster 0	92	49.2	360.11
Cluster 2	63	33.7	717.78
Cluster 1	32	17.1	1,045.31

The estimated aggregate order volume contribution generated by each customer cluster is reported in Table 5. Total cumulative volume reached 111,800.18 portions, with proportional contributions distributed as 29.6% from Cluster 0, 40.4% from Cluster 2, and 29.9% from Cluster 1.

Table 5. Estimated Order Volume Contribution per Cluster

Cluster	Number of customers	Mean portions	Estimated total portions	Volume contribution (%)
Cluster 0	92	360.11	33,130.12	29.6
Cluster 2	63	717.78	45,220.14	40.4
Cluster 1	32	1,045.31	33,449.92	29.9
Total	187	—	111,800.18	100.0

4.2 Discussion

The determination of K balanced algorithmic performance metrics against commercial interpretability. In the evaluation of average within-centroid distance, the values exhibited a continuous downward slope across the parameter range without displaying a distinct, abrupt inflection point; consequently, identifying the optimal configuration could not rely solely on the Elbow curve. Model selection was therefore clarified by analyzing DBI results in conjunction with operational decision-making needs. From a strictly mathematical standpoint, the lowest DBI value occurred at K=2 (0.302), whereas K=3 yielded a higher index of 0.449. Nevertheless, K=3 was adopted as an optimal operational trade-off because it established a more actionable classification by isolating intermediate-volume orders from large-scale banquets, which would otherwise be conflated into an overly aggregated binary division.

Segment characterization was established through the mean order volume of each identified cluster. Cluster 0 (averaging 360.11 portions) represents the small-scale order category, encompassing the largest share of the customer base. Cluster 2 (averaging 717.78 portions) occupies an intermediate operational scale, serving as a transitional segment between individual household gatherings and large corporate events. Cluster 1 (averaging 1,045.31 portions) represents the highest order magnitude despite containing the fewest clients, reflecting institutional buyers or major wedding organizers. A key analytical finding emerges when evaluating segments by volume contribution rather than customer headcount. Cluster 1 accounts for only 32 customers (17.1%) yet generates 29.9% of total catering demand, nearly matching the cumulative output of Cluster 0, which comprises nearly three times as many clients. Concurrently, Cluster 2 serves as the primary operational volume driver, generating 40.4% of total output. This empirical distribution demonstrates that customer volume does not scale linearly with customer frequency, indicating that low-headcount segments with high mean transaction volumes hold disproportionate strategic importance. This finding corroborates Djun *et al.* (2024), who reported that compact customer clusters can contribute substantial cumulative transaction value to overall business operations.

These empirical profiles provide practical input for operationalizing the Segmenting, Targeting, and Positioning (STP) framework at Aisyah Catering, consistent with frameworks proposed by As Shoib *et al.* (2024) and Manggu and Beni (2021). Cluster 0 can be targeted with standardized, budget-conscious packages through regular promotional channels due to its recurring small-event nature. Cluster 2, representing the largest aggregate volume generator, should be prioritized in perishable ingredient procurement and kitchen workflow scheduling to ensure delivery stability. Cluster 1, representing a high-value institutional tier, warrants dedicated account management, customized menu consultations, and structured retention incentives given its substantial commercial contribution. However, this investigation remains bounded within analytical segmentation, and direct operational deployment within Aisyah Catering was beyond the empirical scope of this study. The strategic recommendations outlined herein represent actionable operational pathways that require systematic implementation and longitudinal field assessment to empirically evaluate their impact on revenue growth and client retention.

5. Conclusion

This study successfully implemented the K-means clustering algorithm to segment customer order patterns at Aisyah Catering based on order portion volume, utilizing 187 validated transaction records derived from 197 digitized physical receipts. The combined evaluation of the Elbow method and the Davies-Bouldin Index established K=3 as the most actionable cluster partition (DBI = 0.449). This configuration categorized the customer base into three distinct operational tiers: a small-scale segment (49.2% of customers, averaging 360.11 portions), a medium-scale segment (33.7% of customers, averaging 717.78 portions), and a large-scale segment (17.1% of customers, averaging 1,045.31 portions). The volume contribution analysis indicated that the large-scale tier, despite comprising the smallest customer headcount, generated 29.9% of total order volume—a contribution nearly equivalent to the small-scale tier, which encompassed nearly three times as many clients. Concurrently, the medium-scale tier served as the primary operational driver, generating 40.4% of total cumulative volume.

These empirical findings demonstrate that customer headcount does not correspond linearly with commercial volume contribution, underscoring the necessity of volume-based customer segmentation to guide MSME resource allocation and promotional targeting. Furthermore, this research demonstrates that meaningful, managerially actionable customer segmentation can be achieved using a streamlined single-variable model, providing a pragmatic analytical alternative for early-stage digitalizing MSMEs that lack the dense transactional histories required by multidimensional frameworks such as RFM. Future research should consider incorporating complementary parameters—such as order frequency, menu package preferences, and transaction intervals—once digital recording systems mature. Additionally, empirical follow-up studies are recommended to deploy these segmented strategies directly within enterprise operations and measure their longitudinal impact on sales growth and client retention.

References

Al-Fahmi, B. M., Rahmawati, E., & Sagirani, T. (2023). Penerapan K-Means clustering pada pariwisata Kabupaten Bojonegoro untuk mendukung keputusan strategi pemasaran. *Jurnal Nasional Teknologi dan Sistem Informasi*, *9*(2), 141–149. https://doi.org/10.25077/teknosi.v9i2.2023.141-149

Andrian, Y., & Susilawati, S. (2025). Penerapan algoritma K-Means dalam segmentasi pelanggan untuk meningkatkan strategi pemasaran di e-commerce. *INCODING: Journal of Informatics and Computer Science Engineering*, *5*(1), 95–108. https://doi.org/10.34007/incoding.v5i1.825

Ardi, R. B., Nastiti, F. E., & Sumarlinda, S. (2023). Algoritma K-Means clustering untuk segmentasi pelanggan (Studi kasus: Fashion Viral Solo). *INFOTECH Journal*, *9*(1), 124–131. https://doi.org/10.31949/infotech.v9i1.5214

Ariska, S., Puspita, D., & Anggraini, I. (2024). Comparison of K-Means and K-Medoids algorithm for clustering data UMKM in Pagar Alam City. *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, *13*(2), 193–199. https://doi.org/10.32736/sisfokom.v13i2.2090

Ashari, I. F., Banjarnahor, R., Farida, D. R., Aisyah, S. P., Dewi, A. P., & Humaya, N. (2022). Application of data mining with the K-means clustering method and Davies Bouldin index for grouping IMDB movies. *Journal of Applied Informatics and Computing*, *6*(1), 07–15. https://doi.org/10.30871/jaic.v6i1.3485

As Shoib, A., Dewi, A. P., Prista, D., Wahyu, F. D., & Sudarmiatin. (2024). Analisis strategi pemasaran STP (Segmenting, Targeting, Positioning) pada UMKM pengrajin keramik. *Jurnal Bintang Manajemen*, *2*(4), 91–98. https://doi.org/10.55606/jubima.v2i4.3419

Danuwinata, E. R., & Tjen, J. (2023). Segmentasi pelanggan menggunakan metode K-Means dalam data mining untuk strategi promosi UMKM Lily Cakes Pontianak. *INFORMASI: Jurnal Informatika dan Sistem Informasi*, *17*(2), 203–210. https://doi.org/10.37424/informasi.v17i2.438

Dewobroto, W. S., & Shania, W. (2023). Faktor pendorong dan dampak dari peralihan usaha mikro ke pemasaran digital sebagai solusi meningkatkan penjualan pasca pandemi Covid-19. *Jurnal Penelitian Pendidikan Indonesia*, *9*(1), 172–180. https://doi.org/10.29210/020221823

Djun, S. F., Gunadi, I. G. A., & Sariyasa, S. (2024). Analisis segmentasi pelanggan pada bisnis dengan menggunakan metode K-Means clustering pada model data RFM. *JTIM: Jurnal Teknologi Informasi dan Multimedia*, *5*(4), 354–364. https://doi.org/10.35746/jtim.v5i4.434

Haryandini, N. A. P., & Prihartono, W. (2025). Analisis segmentasi pelanggan untuk menyusun strategi promosi menggunakan K-Means: Studi kasus di PT XYZ. *Jurnal Informatika dan Teknik Elektro Terapan*, *13*(1), 70–78. https://doi.org/10.23960/jitet.v13i1.5697

Kementerian Koperasi dan Usaha Kecil dan Menengah Republik Indonesia. (2024). *Kontribusi UMKM terhadap PDB Indonesia periode 2020–2024*. Databoks Katadata.

Kristanti, B. T., Junaidi, A., & Mandyartha, E. P. (2024). Implementasi K-Means clustering dalam segmentasi pelanggan berdasarkan usia, pendapatan, dan model RFM (Studi kasus: Lantikya Store Jombang). *Jurnal Informatika dan Teknik Elektro Terapan*, *12*(3), 691–699. https://doi.org/10.23960/jitet.v12i3.4677

Manggu, B., & Beni, S. (2021). Analisis penerapan segmentasi, targeting, positioning (STP) dan promosi pemasaran sebagai solusi meningkatkan perkembangan UMKM Kota Bengkayang. *Sebatik*, *25*(1), 27–34. https://doi.org/10.46984/sebatik.v25i1.1146

Marcelina, D., Kurnia, A., & Terttiaavini. (2023). Analisis klaster kinerja usaha kecil dan menengah menggunakan algoritma K-Means clustering. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *3*(2), 293–301. https://doi.org/10.57152/malcom.v3i2.952

Nurlathifa, A., & Sucahyati, D. (2025). Segmentasi pelanggan menggunakan metode K-Means clustering berdasarkan model RFM pada bisnis food and beverage (Studi kasus: Coffee Shop X). *Jurnal Nasional Teknologi dan Sistem Informasi*, *11*(3), 218–225. https://doi.org/10.25077/teknosi.v11i3.2025.218-225

Rohman, N., & Wibowo, A. (2020). Perbandingan metode K-Medoids dan metode K-Means dalam analisis segmentasi pelanggan mall. *SINTECH (Science and Information Technology) Journal*, *7*(1), 7–14. https://doi.org/10.31598/sintechjournal.v7i1.1507

Setyaningtyas, S., Nugroho, B. I., & Arif, Z. (2022). Tinjauan pustaka sistematis: Penerapan data mining teknik clustering algoritma K-Means. *Jurnal Teknoif Teknik Informatika Institut Teknologi Padang*, *10*(2), 52–61. https://doi.org/10.21063/jtif.2022.V10.2.52-61

Syam, F. A. (2017). Implementasi metode klastering k-means untuk mengelompokan hasil evaluasi mahasiswa. *Jurnal Ilmu Komputer dan Bisnis*, *8*(1), 1857–1864.

Widiyanto, A. T., & Witanti, A. (2021). Segmentasi pelanggan berdasarkan analisis RFM menggunakan algoritma K-Means sebagai dasar strategi pemasaran (Studi kasus PT Coversuper Indonesia Global). *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, *1*(1), 204–215. https://doi.org/10.24002/konstelasi.v1i1.4293

Zebua, O. K., Halawa, A., Telaumbanua, E., & Mendrofa, M. P. (2026). Eksplorasi peran strategi segmentasi, targeting, dan positioning (STP) dalam membentuk perilaku pembelian konsumen pada UMKM kuliner. *Scientific Journal of Reflection: Economic, Accounting, Management and Business*, *9*(1), 181–188. https://doi.org/10.37481/sjr.v9i1.1345