Analisis Segmentasi Pelanggan Le Moment Bedasarkan Pola LRFM Dengan Menggunakan Metode K-Means Clustering

Sabrina Fatimah Az-Zahra[1], Andriyan Rizki Jatmiko[2]
Teknologi Informasi, Universitas Merdeka Malang .
e-mail: sabxfaz@gmail.com[1], adriyan.jatmiko@unmer.ac.id[2]

Kata Kunci:

E-commerce
K-Means Clustering
LRFM
Model RM
Segmentasi Pelanggan

ABSTRAK

Perkembangan industri e-commerce dan brand wewangian lokal di Indonesia mendorong perlunya strategi pemasaran berbasis data. Penelitian ini bertujuan melakukan segmentasi pelanggan brand parfum lokal Le Moment berdasarkan riwayat transaksi menggunakan model LRFM (Length, Recency, Frequency, Monetary) dan algoritma K-Means Clustering. Dari data transaksi Oktober 2024–April 2026, diperoleh 93 data pelanggan bersih setelah penanganan outlier (Tukey's Fences). Evaluasi variansi menunjukkan variabel Length dan Frequency bersifat konstan (zero variance) akibat pola pembelian tunggal, sehingga dilakukan reduksi dimensi menjadi model RM (Recency-Monetary). Penentuan jumlah klaster optimal mengombinasikan Metode Elbow dan Silhouette Analysis yang divalidasi dengan Principal Component Analysis (PCA), menghasilkan tiga klaster: Klaster 0 Active/Recent Customers (36,56%), Klaster 1 Dormant/At-Risk Customers (40,86%), dan Klaster 2 Churned/High-Value Customers (22,58%). Hasil segmentasi ini menjadi dasar formulasi strategi personalized marketing untuk meningkatkan retensi dan nilai pelanggan.

Keyword:

E-commerce
K-Means Clustering
LRFM
RM Model
Customer Segmentation

ABSTRACT

The rapid growth of the e-commerce industry and local fragrance brands in Indonesia has driven the need for data-driven marketing strategies. This study aims to segment customers of the local perfume brand Le Moment based on transaction history using the LRFM model (Length, Recency, Frequency, Monetary) and the K- Means Clustering algorithm. From transaction data spanning October 2024 to April 2026, 93 clean customer records were obtained after outlier handling (Tukey's Fences). Variance evaluation showed that the Length and Frequency variables were constant (zero variance) due to single-purchase patterns, leading to dimensionality reduction into an RM (Recency-Monetary) model. The optimal number of clusters, determined by combining the Elbow Method and Silhouette Analysis and validated with Principal Component Analysis (PCA), resulted in three segments: Cluster 0, Active/Recent Customers (36.56%); Cluster 1, Dormant/At-Risk Customers (40.86%); and Cluster 2, Churned/High-Value Customers (22.58%). These segmentation results form the basis for formulating personalized marketing strategies to improve customer retention and value.

Copyright © SENASIF 2026
ISSN: 2598-0076

I. PENDAHULUAN

Sektor perdagangan di Indonesia tengah mengalami transformasi struktural yang masif, didorong oleh penetrasi internet yang mencapai 221 juta jiwa pada periode 2025/2026. Pergeseran dari transaksi berbasis lokasi ke platform digital menciptakan ekosistem omnichannel yang menuntut efisiensi dan kemudahan akses. Salah satu industri yang berkembang pesat dalam platform *e-commerce* adalah industri wewangian lokal. Le Moment merupakan brand parfum lokal asal Indonesia yang mengusung konsep "*Capturing Moments in a Bottle*". Brand ini hadir dengan wewangian konsentrasi Eau de Parfum (EDP) melalui varian seperti *Homesick, Fall in Love, dan Pillow Talk.*

Pesatnya pertumbuhan transaksi *e-commerce* Le Moment menghasilkan volume data transaksi yang besar. Namun, data tersebut cenderung hanya menumpuk tanpa diolah secara maksimal untuk pengambilan keputusan strategis. Masalah utama yang dihadapi adalah penerapan strategi pemasaran massal (*one-size-fits-all*) tanpa mempertimbangkan perbedaan karakteristik dan kebiasaan belanja pelanggan. Pendekatan seragam ini berisiko memicu inefisiensi biaya promosi dan penurunan tingkat retensi konsumen.

Data mining merupakan proses ekstraksi pola dan pengetahuan penting dari himpunan data besar. Salah satu teknik *unsupervised learning* yang efektif untuk mengelompokkan data berdasarkan kemiripan karakteristik adalah clustering. Algoritma K-Means Clustering dipilih karena efisiensi komputasinya pada dataset transaksi dan kemampuannya memetakan segmen pelanggan secara objektif. Untuk mengekstrak perilaku transaksi secara komprehensif, digunakan kerangka kerja LRFM (*Length*, *Recency*, *Frequency*, *Monetary*). Penelitian ini bertujuan untuk melakukan rekayasa fitur data transaksi ke dalam kerangka LRFM, mengidentifikasi profil klaster pelanggan, dan merumuskan rekomendasi strategi pemasaran yang personal.

II. METODE

Penelitian ini menggunakan pendekatan kuantitatif eksploratif dengan tahapan yang sistematis. Tahapan penelitian meliputi pengumpulan data, pra-pemrosesan data, pembentukan variabel LRFM, penerapan algoritma K-Means Clustering, serta evaluasi dan profiling hasil klaster.

2.1 Pengumpulan Data

Data yang digunakan merupakan data sekunder riwayat transaksi e-commerce Le Moment periode Oktober 2024 hingga April 2026. Dataset awal terdiri dari 138 baris log transaksi dengan atribut: InvoiceNo, StockCode, Description, Quantity, InvoiceDate, UnitPrice, CustomerID, dan Country.

Copyright © SENASIF 2026
ISSN: 2598-0076

2.2 Pra-Pemrosesan Data

Tahap pra-pemrosesan meliputi data cleaning (penyelarasan Customer_ID, penanganan missing values, dan pembuangan transaksi bernilai Quantity <= 0 atau Unit_Price <= 0). Selanjutnya dilakukan *feature engineering* untuk menghitung total nominal belanja per item transaksi:

$$Total_Amount = Quantity \times Unit_Price \quad (1)$$

Untuk menangani data pencilan yang dapat mendistorsi posisi centroid pada algoritma K-Means, diterapkan metode Tukey's Fences dengan pengali ekstrem (3.0 × IQR):

$$Batas\ Atas = Q3 + 3.0 \times IQR \quad (2)$$

$$Batas\ Bawah = Q1 - 3.0 \times IQR \quad (3)$$

2.3 Pembentukan Variabel LRFM dan Reduksi Dimensi

Metrik LRFM dikonstruksi per individu pelanggan menggunakan formulasi berikut:

- **Length (L):** Selisih hari antara transaksi pertama dan transaksi terakhir.
- **Recency (R):** Selisih hari antara tanggal acuan analisis (T current) dengan tanggal transaksi terakhir.
- **Frequency (F):** Total hari unik pelanggan melakukan transaksi.
- **Monetary (M):** Akumulasi total nominal pengeluaran pelanggan.

Setelah pembersihan outlier, dataset dinormalisasi menggunakan MinMaxScaler ke skala [0, 1] agar seluruh variabel memiliki bobot seimbang dalam perhitungan jarak Euclidean:

$$Xscaled = (X - Xmin) / (Xmax - Xmin) \quad (4)$$

2.4 Penerapan K-Means dan Evaluasi Klaster

Algoritma K-Means mempartisi data ke dalam Kluster dengan meminimalkan *Within-Cluster Sum of Squares* (WCSS):

$$J = \sum (xj - \mu i)^2 \quad (4)$$

Penentuan jumlah Kluster optimal K dievaluasi menggunakan Metode Elbow dan Silhouette Analysis. Silhouette Score diukur melalui persamaan:

$$S(i) = (b(i) - a(i)) / max(a(i), b(i)) \quad (5)$$

III. HASIL DAN PEMBAHASAN

3.1 Hasil Pra-Pemrosesan Data

Tahap pembersihan data berhasil mereduksi variasi penulisan ID pelanggan dan membuang transaksi fiktif/batal. Agregasi awal pada 138 log transaksi menghasilkan 95 pelanggan unik. Penerapan metode *Tukey's Fences* (3.0 × IQR) mengidentifikasi 2 data pencilan ekstrem yang

Copyright © SENASIF 2026

ISSN: 2598-0076

dieliminasi sehingga menyisakan 93 pelanggan bersih. Eliminasi pencilan ini secara signifikan mencegah terjadinya *mean-shift* (*distorsi centroid*) dan meningkatkan kerapatan internal klaster.

3.2 Evaluasi Variansi Fitur dan Formulasi Model RM

Evaluasi statistik deskriptif terhadap 93 data pelanggan bersih disajikan pada Tabel 1.

Tabel 1. Evaluasi Variansi dan Distribusi Fitur LRFM

Fitur LRFM	Nilai Min	Nilai Maks	Rata-Rata	Standar Deviasi	Variansi (S^2)	Status Fitur
Length (L)	0,0 hari	0,0 hari	0,00 hari	0,0000	0,0000	Dieliminasi (Zero Variance)
Recency (R)	10,0 hari	580,0 hari	278,45 hari	124,3200	15.455,46	Dipertahankan
Frequency (F)	1,0 kali	1,0 kali	1,00 kali	0,0000	0,0000	Dieliminasi (Zero Variance)
Monetary (M)	Rp320.000	Rp335.000*	Rp320.892	2.894,1200	8.375.931,22	Dipertahankan

Berdasarkan Tabel 1, variabel *Length* dan *Frequency* bernilai konstan pada seluruh populasi (L = 0,0 hari; F = 1,0 kali) sehingga menghasilkan variansi $S^2 = 0{,}0000$. Fenomena ini mencerminkan karakteristik industri wewangian okasional/momen (*occasional purchase*), di mana mayoritas pelanggan bertindak sebagai pembeli tunggal (*one-time buyer*) dalam siklus analisis. Memasukkan variabel konstan ke dalam kalkulasi jarak *Euclidean* hanya menambah dimensi tanpa memberikan daya pembeda. Oleh karena itu, dilakukan reduksi dimensi menjadi Model RM (*Recency-Monetary*) yang lebih fokus dan responsif secara bisnis.

3.3 Penentuan Jumlah Klaster Optimal (K)

Pengujian K optimal dilakukan pada rentang K = 2 hingga K = 6. Metode Elbow menunjukkan penurunan *Within-Cluster Sum of Squares* (WCSS) yang sangat tajam dari K = 2 (WCSS ≈ 7,2) ke K = 3 (WCSS = 4,1, turun 43%). Setelah K = 3, kurva WCSS melandai secara bertahap yang mengindikasikan kondisi diminishing returns.

Tabel 2. Evaluasi Silhouette Score Berdasarkan Jumlah Klaster (K)

Jumlah Klaster (K)	Nilai Silhouette Score	Keterangan Evaluasi
K = 2	0,465	Struktur Sedang
K = 3	0,463	Titik Siku Optimal (Elbow)

Copyright © SENASIF 2026
ISSN: 2598-0076

Jumlah Klaster (K)	Nilai Silhouette Score	Keterangan Evaluasi
K = 4	0,484	Kenaikan Kualitas
K = 5	0,484	Stagnan
K = 6	0,498	Nilai Tertinggi

Meskipun K = 6 mencatatkan Silhouette Score sedikit lebih tinggi (0,498), selisihnya sangat tipis dibanding K = 3 (0,463). Memilih K = 3 didasari oleh azas kemudahan interpretasi bisnis (*principle of parsimony*), mencegah terjadinya *over-segmentation*, serta menyesuaikan dengan tekukan siku paling tajam pada Metode Elbow. Verifikasi menggunakan *Principal Component Analysis* (PCA) 2D mengonfirmasi bahwa K = 3 membentuk pemisahan spasial yang jelas tanpa adanya *overlapping.*

3.4 Profiling Segmen Pelanggan

De-normalisasi nilai centroid ke dalam satuan riil (hari dan Rupiah) menghasilkan profil karakteristik untuk masing-masing segmen sebagaimana disajikan pada Tabel 3.

Tabel 3. Profiling Karakteristik 3 Segmen Pelanggan Le Moment

Klaster	Label Segmen	Jumlah (N)	Proporsi	Rata-Rata Recency	Rata-Rata Monetary	Karakteristik Perilaku
Klaster 0	Active / Recent Customers	34	36,56%	189,21 hari (≈6,3 bln)	Rp381.470,59	Pelanggan paling aktif, berbelanja dalam 6 bulan terakhir dengan nilai transaksi reguler. Menjadi penyokong utama arus kas harian.
Klaster 1	Dormant / At-Risk Customers	38	40,86%	413,71 hari (≈13,8 bln)	Rp393.289,47	Kelompok mayoritas yang sudah pasif >1 tahun. Berada pada zona kritis berisiko tinggi beralih ke kompetitor.
Klaster 2	Churned / High-Value Customers	21	22,58%	300,19 hari (≈10,0 bln)	Rp1.080.333,33	Pelanggan VIP dengan nilai pengeluaran 2,83× lebih tinggi dari rata-rata, namun saat ini dalam kondisi tidak aktif.

Copyright © SENASIF 2026
ISSN: 2598-0076

3.5 Formulasi Strategi Pemasaran

Berdasarkan profil karakteristik klaster, disusun usulan strategi pemasaran personalisasi sebagai berikut:

1. **Active / Recent Customers (Klaster 0):** Menerapkan program *Referral & Social Proof Campaign* (voucher diskon 10% atau *free gift* untuk setiap ulasan produk), *Seasonal Reminder System* menjelang siklus momen tahunan, serta strategi *Cross-Selling* paket bundling.
2. **Dormant / At-Risk Customers (Klaster 1):** Menerapkan kampanye *Win-Back Direct Messaging* melalui pesan personalia bertema "Kami Rindu Momen Anda" disertai voucher potongan harga terbatas (*limited-time discount*, misalnya Rp50.000 berlaku 14 hari) untuk menciptakan *sense of urgency*.
3. **Churned / High-Value Customers (Klaster 2):** Menerapkan pendekatan *VIP Personal Consultation* serta pemberian *Exclusive Loyalty Appreciation Gift* berupa harga khusus atau free product pada pembelian porsi besar.
4.

IV. SIMPULAN

Penelitian ini berhasil mengimplementasikan segmentasi pelanggan pada Le Moment menggunakan algoritma K-Means Clustering. Reduksi dimensi dari LRFM menjadi model *Recency-Monetary* (RM) terbukti efektif mengatasi permasalahan *zero variance* pada variabel *Length* dan *Frequency*. Konfigurasi klaster optimal K = 3 membagi pelanggan ke dalam tiga segmen utama: *Active Customers* (36,56%), *Dormant / At-Risk Customers* (40,86%), dan *Churned / High-Value Customers* (22,58%). Hasil segmentasi ini memberikan landasan objektif bagi manajemen Le Moment untuk beralih dari pemasaran massal ke strategi pemasaran personalisasi yang tepat sasaran. Penelitian selanjutnya disarankan untuk mengintegrasikan atribut kualitatif seperti preferensi aroma wewangian serta membangun sistem segmentasi real-time berbasis CRM.

DAFTAR RUJUKAN

1. Anitha, P., & Patil, M. (2022). RFM model for customer purchase behavior using K-Means algorithm. *Journal of King Saud University*, 3(2): 17-31.
2. Awalina, E. F. L., & Rahayu, W. I. (2023). Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail. *Jurnal Teknologi dan Informasi*, 13(2): 122-137.
3. Fahmy, R., & Santosa, P. (2024). Analysis of Product Information Asymmetry and its Impact on Online Consumer Dissatisfaction. *Journal of Information Systems and Digital Economy*, 12(1): 34-50.

Copyright © SENASIF 2026
ISSN: 2598-0076

4. Guntara, M., & Lutfi, N. (2023). Optimasi Cacah Klaster pada Klasterisasi dengan Algoritma KMeans Menggunakan Silhouette Coeficient dan Elbow Method. *JuTI "Jurnal Teknologi Informasi"*, 2(1): 43.

5. Hasan, G. (2025). *Manajemen Hubungan Pelanggan (Customer Relationship Management - CRM)*. Padang: Menara Press Indonesia.

6. Hidayat, A., & Pratama, R. (2024). Analisis Transformasi Ekonomi Digital dan Perilaku Konsumen di Indonesia. *Jurnal Ekonomi dan Bisnis Digital*, 15(1): 45-62.

7. Hidayati, R., Zubair, A., Hidayat Pratama, A., & Indana, L. (2021). Analisis Silhouette Coefficient pada 6 Perhitungan Jarak K-Means Clustering. *Techno.COM: Jurnal Teknologi Informasi dan Komunikasi*, 20(2): 186–197.

8. Lestari, D. P. (2023). Logistics Challenges in Indonesian E-commerce: From Infrastructure to Last-Mile Delivery. *International Journal of Supply Chain Management*, 11(3): 201-218.

9. Nurputra, W., Witanti, W., & Komarudin, A. (2025). Principal Component Analysis (PCA) Untuk Meningkatkan Hasil Klasterisasi Penjualan Video Game Menggunakan Algoritma K-Means. *Jurnal Locus Penelitian Dan Pengabdian*, 4(8).

10. Pramesti, A., & Nuryana, I. (2026). Segmentasi Pelanggan Berbasis Clustering Menggunakan LRFM dan Variabel Transaksi Untuk Mendukung Strategi Pemasaran. *Jurnal Data Science Indonesia*, 6(1): 99-110.

11. Prasetyo, A., Maimunah, M., & Sukmasetya, P. (2023). K-Means Clustering Method for Determining Waste Transportation Routes to Landfill. *Jurnal Riset Informatika*, 5(3): 277-284.

12. Rafidah, H. (2025). Perancangan Sistem Aplikasi Penilaian Kedisiplinan Siswa Menggunakan Metode K-Means Elbow Berbasis Website. *Saturnus: Jurnal Teknologi Dan Sistem Informasi*, 3(4): 15-29.

13. Setiadi, H., & Wijaya, K. (2024). Omnichannel Strategy: Bridging the Gap Between Offline and Online Trade in Southeast Asia. *Indonesia Journal of Marketing Science*, 15(3): 301-318.

14. Siagian, R., Sirait, P., & Halim, A. (2022). Penerapan Algoritma K-Means dan KMedoids untuk Segmentasi Pelanggan pada Data Transaksi E-Commerce. *SISTEMASI: Jurnal Sistem Informasi*, 11(2): 2540-9719.

15. Tacharri, C., Rohmani, A., & Fahmi, A. (2025). Strategic Clustering of Poverty Areas in Central Java Using K-Means and Silhouette Evaluation. *Sinkron*, 9(2): 895-904.

Copyright © SENASIF 2026
ISSN: 2598-0076