

Analisis Customer Lifetime Value (CLV) pada Penjualan Sambal Bakar Indonesia Grand Depok City Menggunakan Metode RFM

Evyta Oktaviani¹, Siti Khotijah², Siwi Puji Astuti³

^{1,2,3} Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Indraprasta PGRI, Depok, Indonesia

Email: ¹oktaviani25103@gmail.com

Abstrak—Data transaksi pelanggan pada platform digital memiliki potensi besar untuk dimanfaatkan dalam analisis perilaku pelanggan, namun belum dimanfaatkan secara optimal oleh banyak pelaku usaha kuliner. Penelitian ini bertujuan membangun sistem berbasis web yang mengintegrasikan metode Recency, Frequency, Monetary (RFM) dan algoritma K-Means Clustering untuk segmentasi pelanggan serta analisis Customer Lifetime Value (CLV) pada Sambal Bakar Indonesia outlet Grand Depok City. Data yang digunakan merupakan 5.972 transaksi dari 2.989 pelanggan pada periode 1 September 2025 hingga 1 Maret 2026, bersumber dari platform GoFood, GrabFood, dan ShopeeFood. Sistem dibangun menggunakan Python dengan framework Streamlit. Berdasarkan Elbow Method diperoleh K=4 sebagai jumlah cluster optimal, menghasilkan empat segmen: Lost Customer (571 pelanggan, 19,1%), New Customer (1.000 pelanggan, 33,5%), Regular (102 pelanggan, 3,4%), dan Big Spender (1.316 pelanggan, 44,0%). Evaluasi menggunakan Silhouette Score menghasilkan nilai 0,4172 (kategori Cukup). Analisis CLV menunjukkan segmen Regular memiliki rata-rata CLV tertinggi sebesar Rp 910.172, mengindikasikan pelanggan dengan frekuensi konsisten memberikan kontribusi nilai jangka panjang terbesar.

Kata Kunci: Customer Lifetime Value; K-Means Clustering; RFM; Segmentasi Pelanggan; Streamlit

Abstract—Customer transaction data on digital platforms has significant potential for customer behavior analysis but has not been optimally utilized by many culinary business. This study aims to build a web-based system integrating the Recency, Frequency, Monetary (RFM) method and K-Means Clustering algorithm for customer segmentation and Customer Lifetime Value (CLV) analysis at Sambal Bakar Indonesia, Grand Depok City outlet. Data consists of 5,972 transactions from 2,989 customers during September 1, 2025 to March 1, 2026, sourced from GoFood, GrabFood, and ShopeeFood platforms. The systems was developed using python with Streamlit framework. Based on the Elbow Method K=4 was obtained as the optimal number of clusters, resulting in four segments: Lost Customer (571 customers, 19.1%), New Customer (1,000 customers, 33.5%), Regular (102 customers, 3.4%) and Big Spender (1,316 customers, 44.0%). Clustering quality evaluation using Silhouette Score yielded 0.4172 (Sufficient category). CLV analysis shows Regular segment has the highest average CLV of Rp 910,172, indicating customers with consistent transaction frequency provide the greatest long-term value contribution.

Keywords: Customer Lifetime Value, K-Means Clustering, RFM, Customer Segmentation, Streamlit.

1. PENDAHULUAN

Perkembangan teknologi informasi yang pesat membawa perubahan signifikan dalam industri kuliner, khususnya melalui meningkatnya penggunaan layanan pemesanan makanan secara online. Platform digital seperti GoFood, GrabFood, dan ShopeeFood telah mengubah pola konsumsi masyarakat sekaligus menghasilkan data transaksi pelanggan dalam jumlah besar. Data tersebut memiliki potensi besar untuk dijadikan dasar pengambilan keputusan bisnis yang lebih efektif (Hendarsyah, 2020).

Sambal Bakar Indonesia outlet Grand Depok City (GDC) merupakan salah satu bisnis kuliner yang aktif memanfaatkan platform digital. Meskipun data transaksi pelanggan tersedia dalam jumlah besar, data tersebut belum dimanfaatkan secara optimal. Perusahaan mengalami kesulitan dalam mengidentifikasi pelanggan loyal, pelanggan dengan kontribusi pendapatan terbesar, serta pelanggan yang berpotensi berhenti bertransaksi. Kondisi ini mengakibatkan strategi pemasaran yang diterapkan kurang efektif dan tidak berbasis data.

Metode Recency, Frequency, Monetary (RFM) merupakan pendekatan yang terbukti efektif dalam segmentasi pelanggan. Savitri et al. (2018) berhasil menerapkan K-Means berbasis RFM pada klinik kecantikan dan menghasilkan segmentasi yang bermakna bagi strategi pemasaran. Hadi et al. (2017) juga membuktikan efektivitas RFM dalam pemetaan pelanggan perusahaan herbal.

Sementara itu, Wijaya et al. (2021) menerapkan K-Means dan RFM pada bisnis gaming dan

memperoleh hasil segmentasi yang dapat ditindaklanjuti secara strategis. Nahjan et al. (2023) menggunakan K-Means untuk analisis penjualan dan menunjukkan bahwa metode ini efektif untuk pengelompokan data pelanggan. Penelitian-penelitian tersebut menunjukkan bahwa kombinasi RFM dan K-Means konsisten menghasilkan segmentasi pelanggan yang relevan untuk berbagai jenis bisnis.

Selain segmentasi, analisis Customer Lifetime Value (CLV) memberikan gambaran nilai pelanggan secara jangka panjang. Dengan mengetahui CLV, perusahaan dapat menentukan pelanggan prioritas dan mengalokasikan sumber daya pemasaran secara lebih efisien (RevoU, 2024). Namun, belum ada penelitian yang secara spesifik mengintegrasikan RFM, K-Means Clustering, dan analisis CLV dalam satu sistem berbasis web untuk bisnis kuliner online di Indonesia.

Berdasarkan gap tersebut, penelitian ini bertujuan membangun sistem web menggunakan Python dan Streamlit yang mengintegrasikan metode RFM, K-Means Clustering, dan CLV untuk segmentasi pelanggan Sambal Bakar Indonesia GDC, sehingga dapat memberikan rekomendasi strategi pemasaran yang lebih tepat sasaran.

2. METODE PENELITIAN

2.1 Data Penelitian

Data yang digunakan merupakan data transaksi online pelanggan Sambal Bakar Indonesia GDC dari platform GoFood, GrabFood, dan ShopeeFood periode 1 September 2025 hingga 1 Maret 2026. Dataset terdiri dari 5.972 transaksi dengan 2.989 pelanggan unik. Variabel meliputi Customer ID, Order Date, Product Name, dan Sales (IDR).

2.2 Tahapan Penelitian

Tahapan penelitian meliputi: (1) pengumpulan data transaksi dalam format CSV; (2) preprocessing mencakup konversi tipe data, penghapusan nilai null, dan normalisasi format; (3) perhitungan metrik RFM per pelanggan; (4) normalisasi menggunakan StandardScaler (z-score); (5) penentuan cluster optimal dengan Elbow Method ($k=1-10$); (6) clustering K-Means dan evaluasi Silhouette Score; (7) perhitungan CLV dan pemberian label segmen; serta (8) pengembangan sistem web berbasis Streamlit.

2.3 Metode RFM

Metode RFM menganalisis perilaku pelanggan berdasarkan tiga dimensi (Hadi et al., 2017):

Recency (R) = Tanggal Referensi – Tanggal Transaksi Terakhir Pelanggan

Frequency (F) = Σ jumlah transaksi pelanggan selama periode analisis

Monetary (M) = Σ total nilai belanja pelanggan selama periode analisis

Tanggal referensi ditetapkan sebagai tanggal transaksi terakhir dalam dataset, yaitu 1 Maret 2026.

Data RFM kemudian dinormalisasi menggunakan StandardScaler dengan rumus z-score:

$X_{scaled} = (X - \mu) / \sigma$, di mana μ adalah rata-rata dan σ adalah standar deviasi. Normalisasi dilakukan agar ketiga dimensi RFM memiliki skala yang setara sebelum proses clustering.

2.4 K-Means Clustering

Algoritma K-Means mengelompokkan data ke dalam K cluster dengan meminimalkan jarak antara setiap titik data dengan centroid cluster-nya. Jarak dihitung menggunakan Euclidean Distance (Metisen & Sari, 2016):

$$d = \sqrt{(R_1 - R_2)^2 + (F_1 - F_2)^2 + (M_1 - M_2)^2}$$

Jumlah cluster optimal ditentukan melalui Elbow Method dengan menghitung nilai inertia untuk $k=1$ hingga $k=10$. Kualitas clustering dievaluasi menggunakan Silhouette Score: skor $\geq 0,50$ (Baik), $0,26-0,50$ (Cukup), $< 0,26$ (Kurang Optimal).

2.5 Customer Lifetime Value (CLV)

CLV dihitung menggunakan rumus berbasis Monetary dan Recency Weight (Juniati et al., 2020):

Recency Weight = $(\text{Max Recency} - \text{Recency} + 1) / \text{Max Recency}$

CLV = Monetary $\times$ Recency Weight

Pendekatan ini memberikan bobot lebih tinggi kepada pelanggan yang lebih baru bertransaksi, sehingga nilai CLV mencerminkan kontribusi jangka panjang yang lebih akurat.

3. ANALISA DAN PEMBAHASAN

3.1 Hasil Perhitungan RFM

Perhitungan RFM dilakukan terhadap seluruh 2.989 pelanggan. Tabel 1 menyajikan parameter statistik yang diperoleh dari keseluruhan dataset sebagai acuan normalisasi.

Tabel 1. Parameter Statistik Data RFM (n=2.989 pelanggan)

Metrik	Mean (μ)	Std Deviasi (σ)
Recency (hari)	75,1141	52,0470
Frequency (kali)	1,9980	1,4967
Monetary (Rp)	387.217,50	334.193,43

3.2 Hasil Elbow Method

Elbow Method dijalankan untuk k=1 hingga k=10 guna menentukan jumlah cluster optimal. Tabel 2 menyajikan nilai inertia yang diperoleh pada setiap nilai k.

Tabel 2. Nilai Inertia Elbow Method

Nilai k	Nilai Inertia	Keterangan
1	8.967,00	
2	5.577,35	
3	3.664,97	
4	2.844,72	← Titik Siku (K Optimal)
5	2.393,29	
6	2.053,94	
7	1.783,29	
8	1.583,41	
9	1.445,58	
10	1.294,30	

Berdasarkan Tabel 2, penurunan inertia paling signifikan terjadi dari k=1 ke k=4. Setelah k=4, penurunan inertia mulai melambat secara signifikan sehingga titik siku (elbow point) teridentifikasi pada k=4. Dengan demikian, K=4 ditetapkan sebagai jumlah cluster optimal untuk proses K-Means Clustering.

3.3 Hasil K-Means Clustering

Proses K-Means Clustering dengan K=4 menghasilkan empat cluster. Tabel 3 menyajikan centroid akhir setelah proses iterasi konvergen, yang merupakan posisi pusat masing-masing cluster dalam ruang RFM yang telah dinormalisasi.

Tabel 3. Nilai Centroid Akhir K-Means (K=4)

Centroid	R_scaled	F_scaled	M_scaled
C1	0,0311	1,0322	0,9968
C2	1,0125	-0,3795	-0,3896
C3	0,0392	3,3747	3,4351
C4	-0,7859	-0,4211	-0,4027

Tabel 4 menyajikan rata-rata nilai RFM aktual untuk masing-masing cluster sebelum normalisasi.

Tabel 4. Rata-rata RFM per Cluster

Cluster	Avg Recency (hari)	Avg Frequency (kali)	Avg Monetary (Rp)
1	76,73	3,54	720.355,17
2	127,81	1,43	257.016,15

3	77,16	7,05	1.535.210,29
4	34,21	1,37	252.631,23

Evaluasi kualitas clustering menggunakan Silhouette Score menghasilkan nilai 0,4172 yang tergolong kategori Cukup (rentang 0,26–0,50). Nilai ini menunjukkan bahwa pengelompokan pelanggan ke dalam empat cluster cukup representatif. Hasil ini sejalan dengan penelitian Savitri et al. (2018) yang memperoleh Silhouette Score dalam kategori cukup pada kasus segmentasi pelanggan klinik kecantikan menggunakan K-Means berbasis RFM.

3.4 Distribusi Segmen Pelanggan

Setiap cluster diberi label segmen berdasarkan karakteristik rata-rata RFM-nya. Tabel 5 menyajikan distribusi pelanggan per segmen beserta nilai CLV.

Tabel 5. Distribusi Segmen Pelanggan dan Analisis CLV

Segmen	Jml Pelanggan	% Total	Avg CLV (Rp)	Total CLV (Rp)
Big Spender	1.316	44,0%	206.965	272.365.619
New Customer	1.000	33,5%	74.316	74.316.080
Lost Customer	571	19,1%	414.497	236.677.993
Regular	102	3,4%	910.172	92.837.593

3.5 Analisis dan Pembahasan

Berdasarkan Tabel 5, segmen Big Spender mendominasi dengan 1.316 pelanggan (44,0%). Hal ini mengindikasikan bahwa mayoritas pelanggan Sambal Bakar Indonesia GDC memiliki total nilai belanja yang tinggi, namun frekuensi kunjungannya masih tergolong rendah (rata-rata 1,37 kali, Tabel 4). Kondisi ini membuka peluang untuk meningkatkan frekuensi kunjungan segmen Big Spender melalui program promosi berbasis loyalitas.

Temuan yang menarik adalah segmen Regular yang hanya terdiri dari 102 pelanggan (3,4%) justru memiliki rata-rata CLV tertinggi sebesar Rp 910.172 — hampir 4,4 kali lebih tinggi dibandingkan segmen Big Spender. Hal ini disebabkan oleh kombinasi frekuensi transaksi tinggi (rata-rata 7,05 kali) dan recency yang baik, sehingga menghasilkan Recency Weight yang tinggi dalam perhitungan CLV. Temuan ini mengonfirmasi bahwa nilai pelanggan tidak semata-mata ditentukan oleh besaran nilai transaksi, melainkan oleh konsistensi dan keaktifan transaksi terkini.

Segmen Lost Customer (571 pelanggan, 19,1%) memiliki rata-rata Recency tertinggi (127,81 hari), mengindikasikan pelanggan yang sudah lama tidak bertransaksi. Meski avg CLV-nya sebesar Rp 414.497 cukup tinggi — didorong oleh nilai monetary historis yang besar — namun Recency Weight yang rendah menekan nilai CLV aktual mereka. Segmen ini memerlukan kampanye win-back yang agresif untuk mengaktifkan kembali potensi transaksi yang ada.

Segmen New Customer (1.000 pelanggan, 33,5%) memiliki rata-rata CLV terendah (Rp 74.316) karena frekuensi transaksi dan nilai belanja yang masih terbatas. Namun, segmen ini merupakan peluang pertumbuhan terbesar jika dikelola dengan program onboarding yang tepat untuk meningkatkan frekuensi transaksi.

3.6 Sistem Berbasis Web

Sistem analisis segmentasi pelanggan dibangun menggunakan Python dengan framework Streamlit. Pipeline analisis dirancang secara bertahap (Step 1–7) dan dapat diakses melalui navigasi sidebar. Sistem menghasilkan delapan jenis laporan PDF formal mencakup laporan analisis cluster, analisis CLV (statistik, per segmen, Top-N), profil pelanggan individual, analisis menu penjualan, dan ringkasan eksekutif. Semua hasil analisis tersimpan di database SQLite sehingga dapat diakses kembali tanpa menjalankan ulang pipeline.

4. KESIMPULAN

Penelitian ini berhasil membangun sistem analisis segmentasi pelanggan berbasis web menggunakan Python dan Streamlit yang mengintegrasikan metode RFM, K-Means Clustering, dan Customer Lifetime Value (CLV). Berdasarkan hasil analisis terhadap 2.989 pelanggan dengan data

periode 1 September 2025 – 1 Maret 2026, diperoleh $K=4$ sebagai jumlah cluster optimal menghasilkan empat segmen: Big Spender (44,0%), New Customer (33,5%), Lost Customer (19,1%), dan Regular (3,4%). Evaluasi kualitas clustering menghasilkan Silhouette Score 0,4172 (kategori Cukup). Analisis CLV menunjukkan segmen Regular memiliki rata-rata CLV tertinggi (Rp 910.172) meskipun jumlah pelanggannya paling sedikit, mengindikasikan bahwa konsistensi dan keaktifan transaksi merupakan faktor utama dalam menentukan nilai pelanggan jangka panjang. Hasil penelitian ini diharapkan dapat membantu manajemen Sambal Bakar Indonesia GDC dalam menentukan prioritas strategi pemasaran yang lebih efektif dan berbasis data.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Indraprasta PGRI, dosen pembimbing, serta manajemen Sambal Bakar Indonesia GDC yang telah memberikan izin penggunaan data transaksi pelanggan untuk keperluan penelitian ini.

REFERENCES

- Darmi, Y., & Setiawan, A. (2016). Penerapan Metode Clustering K-Means Dalam Pengelompokan Pengguna Internet. *Jurnal Pseudocode*, 12(2), 148–157.
- Hadi, F., Rahmadia, D. O., Nugraha, F. H., & Bulan, N. P. (2017). Penerapan K-Means Clustering Berdasarkan RFM Mofek Sebagai Pemetaan dan Pendukung Strategi Pengelolaan Pelanggan. *Jurnal Ilmu Komputer dan Informasi*, 15(1), 69–76.
- Hendarsyah, D. (2020). Analisis Perilaku Konsumen dan Keamanan Kartu Kredit Perbankan. *Jurnal Ilmiah Ekonomi Kita*, 1(1), 85–96.
- Juniati, F., Zafa, R., & Monalisa, M. S. (2020). Klasterisasi Customer Lifetime Value dengan Model LRFM Menggunakan Algoritma Fuzzy C-Means. *Jurnal Rekayasa dan Manajemen Sistem Informasi*, 6(1), 38–42.
- Method, M., Indriyani, F., & Irfiani, E. (2019). Clustering Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means. *Jurnal TECHSI*, 7(November), 109–113.
- Metisen, B. M., & Sari, H. L. (2016). Analisis Clustering Menggunakan Metode K-Means dalam Pengelompokan Penjualan Produk pada Swalayan Fadila. *Jurnal Media Infotama*, 11(2), 110–118.
- Nahjan, M. R., Heryana, N., & Voutama, A. (2023). Implementasi RapidMiner dengan Metode Clustering K-Means untuk Analisa Penjualan Pada Toko OJ Cell. *Jurnal Ilmiah KOMPUTASI*, 7(1), 101–104.
- RevoU. (2024). Customer Lifetime Value. <https://www.revou.co/kosakata/customer-lifetime-value>
- Savitri, A. D., Bachtar, F. A., & Setiawan, N. Y. (2018). Segmentasi Pelanggan Menggunakan Metode K-Means Clustering Berdasarkan Model RFM Pada Klinik Kecantikan. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(9), 2957–2966.
- Septiani, Y., Arribe, E., & Diansyah, R. (2020). Penerapan Metode K-Means untuk Pengelompokan Mahasiswa Berdasarkan Minat Belajar. *Jurnal CoSciTech*, 3(1), 131–143.
- Wakhidah, N. (2016). Clustering Menggunakan K-Means Algorithm. *Jurnal Fakultas Teknologi Informasi*, 21(1), 70–80.
- Wijaya, K. Z., Djunaidy, A., & Mahananto, F. (2021). Segmentasi Pelanggan Menggunakan Algoritma K-Means dan Analisis RFM di Ova Gaming ESports Arena Kediri. *Jurnal Repositor*, 10(2).