TEKNOBIS : Teknologi, Bisnis Dan Pendidikan
Volume 4, No. 3 Tahun 2026
ISSN 3024-8256 (media online)
Hal 694-702

Penerapan RFM Analysis dan K-Means Clustering untuk Segmentasi Pelanggan pada Dataset Online Retail II

Nathan Noval Jivi Earnestine[1], Rangga Irawan[2], Bergio Fabian[3], Mufidah Karimah[4]

[1-4] Fakultas Ilmu Komputer, Program Studi Sistem Informasi, Universitas Pamulang, Tangerang Selatan, Indonesia
Email: [1]nathannoval123@gmail.com, [2]ranggairwns11@gmail.com, [3]Ishidam821@gmail.com, [4]dosen02829@unpam.ac.id

Abstrak−Persaingan di industri e-commerce yang semakin ketat mendorong perusahaan untuk lebih mengenali karakteristik pelanggan demi merancang strategi pemasaran yang lebih efektif. Penelitian ini bertujuan untuk melakukan klasifikasi segmen pelanggan pada dataset Online Retail II dengan mengintegrasikan metode Analisis RFM (Recency, Frequency, Monetary) dan algoritma K-Means Clustering. Dataset ini mencakup 1.067.371 transaksi yang tercatat antara Desember 2009 hingga Desember 2011. Proses pembersihan data menghapus baris yang tidak memiliki ID Pelanggan, nilai Kuantitas dan Harga yang tidak valid, serta data yang terduplikasi, sehingga menghasilkan 779.425 transaksi bersih dari 5.878 pelanggan yang berbeda. Indikator RFM dihitung dan diubah dengan logaritmik untuk mengurangi ketidakseimbangan sebelum dinormalisasi dan dikelompokkan. Penentuan jumlah cluster yang optimal menggunakan Metode Elbow dan Skor Silhouette menunjukkan bahwa k=4 adalah jumlah cluster yang tepat, dengan Skor Silhouette sebesar 0,3650. Hasil pengelompokan mengidentifikasi empat jenis pelanggan: Champion/VIP (20,3%), Loyal Customer (24,8%), Potential/New (21,3%), dan At Risk/Churned (33,6%). Penelitian ini merekomendasikan strategi pemasaran yang spesifik untuk masing-masing segmen guna mendukung program manajemen hubungan pelanggan (CRM) yang lebih efisien.

Kata Kunci: Analisis RFM; K-Means Clustering; Segmentasi Pelanggan; Customer Intelligence; Data Mining

Abstract−*The increasingly fierce competition in the e-commerce industry is pushing companies to better understand customer characteristics in order to design more effective marketing strategies. This study aims to classify customer segments in the Online Retail II dataset by integrating the RFM (Recency, Frequency, Monetary) Analysis method and the K-Means Clustering algorithm. This dataset includes 1,067,371 transactions recorded between December 2009 and December 2011. The data cleaning process removed rows with missing Customer IDs, invalid Quantity and Price values, and duplicated data, resulting in 779,425 net transactions from 5,878 different customers. The RFM indicator was calculated and transformed logarithmically to reduce imbalance before normalization and clustering. Determining the optimal number of clusters using the Elbow Method and Silhouette Score showed that k=4 was the appropriate number of clusters, with a Silhouette Score of 0.3650. The clustering results identified four customer types: Champion/VIP (20.3%), Loyal Customer (24.8%), Potential/New (21.3%), and At Risk/Churned (33.6%). This study recommends specific marketing strategies for each segment to support a more efficient customer relationship management (CRM) program.*

Keywords: *RFM Analysis; K-Means Clustering; Customer Segmentation; Customer Intelligence; Data Mining*

1. PENDAHULUAN

Perkembangan dalam bidang teknologi informasi telah menghasilkan peningkatan yang signifikan terhadap jumlah data transaksi yang dihasilkan oleh sektor ritel online. Informasi transaksi yang tersimpan dalam sistem perusahaan tidak hanya berfungsi sebagai catatan administratif, tetapi juga dapat digunakan untuk mendapatkan wawasan strategis yang dapat membantu dalam pengambilan keputusan bisnis (Irawan, Amaluddin, dan Wijayanti, 2025). Penggunaan data dalam skala besar melalui teknik data mining memungkinkan perusahaan untuk menemukan pola perilaku konsumen yang dapat menjadi dasar untuk merancang strategi pemasaran yang lebih efektif.

Salah satu penerapan data mining di sektor ritel adalah dalam segmentasi pelanggan. Segmentasi pelanggan merupakan proses pengelompokan konsumen berdasarkan karakteristik perilaku transaksi, sehingga setiap kelompok memiliki atribut tertentu yang sama (Christy, Umamakeswari, Priyatharsini, dan Neyaa, 2021). Metode RFM (Recency, Frequency, Monetary) adalah teknik segmentasi yang berfokus pada perilaku, yang umum digunakan untuk memperkirakan aktivitas pelanggan melalui pola pembelian, dengan mengukur waktu sejak

TEKNOBIS : Teknologi, Bisnis Dan Pendidikan
Volume 4, No. 3 Tahun 2026
ISSN 3024-8256 (media online)
Hal 694-702

transaksi terakhir (Recency), frekuensi transaksi (Frequency), dan total nilai belanja (Monetary) (Rahim, Mushafiq, Khan, dan Arain, 2021). Dengan menggunakan profil RFM, perusahaan dapat memahami perilaku pelanggan di masa lalu dan menetapkan target keterlibatan untuk masa depan.

Dalam pengelompokan pelanggan berdasarkan fitur RFM, penelitian ini menggunakan algoritma K-Means Clustering, teknik unsupervised learning yang sering dipilih karena kemudahan dalam melakukan perhitungan dan efisiensinya (Nasyuha, Zulham, dan Rusydi, 2022). Kombinasi RFM dan K-Means telah banyak digunakan dalam berbagai penelitian tentang manajemen hubungan pelanggan (CRM) karena mampu menghasilkan segmen pelanggan yang dapat ditindaklanjuti untuk strategi pemasaran, seperti program loyalitas dan kampanye re-engagement (Smaili dan Hachimi, 2023). Penelitian sebelumnya pada data ritel B2C dengan skala besar telah menunjukkan bahwa pendekatan RFM-K-Means dapat dengan efektif mengidentifikasi pelanggan yang bernilai tinggi serta pelanggan yang berpotensi churn (Irawan, Amaluddin, dan Wijayanti, 2025).

Menentukan jumlah cluster yang optimal merupakan langkah penting dalam proses clustering. Penilaian menggunakan Silhouette Score sering kali diterapkan untuk mengukur kualitas pemisahan antar cluster serta kohesi di dalam masing-masing cluster (Yulisasih, Herman, Sunardi, dan Yuliansyah, 2024). Penelitian ini menerapkan analisis RFM dan K-Means Clustering pada dataset Online Retail II untuk mengidentifikasi segmen pelanggan yang representatif, menilai kualitas cluster dengan Silhouette Score, dan merumuskan rekomendasi strategi pemasaran yang spesifik untuk setiap segmen.

2. METODE PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini menerapkan metode data mining melalui teknik pengelompokan untuk memisahkan segmen pelanggan. Proses penelitian ini mengikuti metode CRISP-DM yang terdiri dari identifikasi masalah bisnis, pengolahan dan eksplorasi data, penghitungan fitur RFM, penyesuaian data, penetapan jumlah kelompok yang ideal, pelaksanaan pengelompokan dengan K-Means, evaluasi menggunakan Silhouette Score, serta analisis hasil untuk memberikan saran bisnis (Pynadath, Rofin, dan Thomas, 2023).

2.2 Dataset Penelitian

Dataset yang digunakan adalah Online Retail II yang tersedia secara publik di UCI Machine Learning Repository dan Kaggle. Dataset ini mencakup seluruh transaksi sebuah perusahaan ritel daring berbasis di Inggris yang menjual barang-barang hadiah, periode 1 Desember 2009 hingga 9 Desember 2011 (Madhiraju, Reddy, & Sasikala, 2024).

Tabel 1. Atribut Dataset Online Retail II

No	Atribut	Keterangan
1	Invoice	Nomor transaksi (awalan C = pembatalan)
2	StockCode	Kode unik produk
3	Description	Nama produk
4	Quantity	Jumlah barang per transaksi
5	InvoiceDate	Tanggal & waktu transaksi
6	Price	Harga satuan (GBP)
7	Customer ID	Nomor identitas pelanggan
8	Country	Negara asal pelanggan

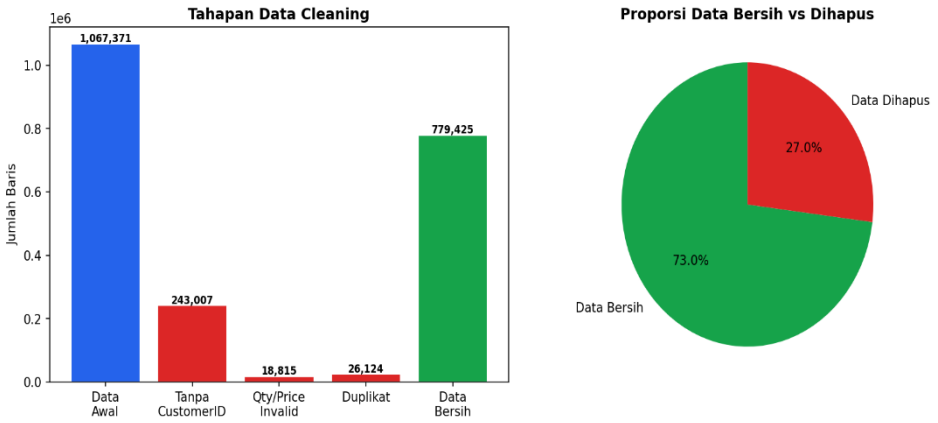
Variabel target **y** tidak digunakan dalam penelitian ini karena metode clustering termasuk ke dalam kategori *unsupervised learning* yang tidak memerlukan label kelas dalam proses pengelompokan data.

TEKNOBIS : Teknologi, Bisnis Dan Pendidikan
Volume 4, No. 3 Tahun 2026
ISSN 3024-8256 (media online)
Hal 694-702

2.3 Data Cleaning (Pembersihan Data)

Tahap pembersihan data dilakukan guna menjamin mutu data sebelum pelaksanaan perhitungan RFM. Prosedur ini melibatkan penghapusan baris yang tidak memiliki Customer ID (tidak bisa dianalisis per pelanggan), penghapusan transaksi dengan Quantity atau Price yang bernilai negatif/nol (tanda adanya retur atau kesalahan dalam input), serta penghilangan data yang duplikat (Saha, Tripathy, Nayak, Bhoi, dan Barsocchi, 2021).

Gambar 1. Ringkasan Data Cleaning



Gambar 1. Ringkasan Data Cleaning

Setelah proses cleaning, dataset awal sebanyak 1.067.371 baris berkurang menjadi 779.425 transaksi bersih (5.878 pelanggan unik), siap digunakan untuk perhitungan RFM.

2.4 Perhitungan RFM (Recency, Frequency, Monetary)

Analisis RFM menghitung tiga aspek dari perilaku konsumen: Recency (jumlah hari sejak transaksi terakhir hingga tanggal acuan), Frequency (jumlah transaksi yang berbeda), dan Monetary (jumlah total pengeluaran) (Christy et al., 2021; Rahim et al., 2021). Rumus yang digunakan adalah sebagai berikut:

$$Recency = Snapshot\_Date - Tanggal\_Transaksi\_Terakhir$$

$$Frequency = \Sigma\ Invoice\_Unik\_per\_Customer$$

$$Monetary = \Sigma\ (Quantity \times Price)$$

Snapshot date pada penelitian ini ditetapkan sehari setelah transaksi terakhir dalam dataset, yaitu 10 Desember 2011.

2.5 Transformasi dan Normalisasi Data

Distribusi karakteristik RFM biasanya menunjukkan skewness yang signifikan karena terdapat pelanggan dengan nilai yang sangat tinggi. Untuk menangani masalah ini, diterapkan transformasi logaritmik pada ketiga karakteristik sebelum melakukan normalisasi dengan StandardScaler, sehingga tiap karakteristik memiliki skala yang seimbang dan tidak ada atribut yang mendominasi dalam perhitungan jarak antar data (Abbasimehr & Bahrini, 2022; Tabianan, Velu, & Ravi, 2022).

$$.X\_log = log(1 + X)$$

2.6 Algoritma K-Means Clustering

K-Means merupakan metode pembelajaran tanpa pengawasan yang mengelompokkan data menjadi k kelompok berdasarkan kedekatan jarak Euclidean ke titik pusat (centroid) (Tabianan, Velu, & Ravi, 2022; Nasyuha et al., 2022). Proses algoritma ini dilakukan secara bertahap: memulai dengan memilih k centroid awal, mengukur jarak setiap titik data ke centroid, mengelompokkan data

TEKNOBIS : Teknologi, Bisnis Dan Pendidikan
Volume 4, No. 3 Tahun 2026
ISSN 3024-8256 (media online)
Hal 694-702

ke centroid yang paling dekat, dan akhirnya memperbarui posisi centroid hingga terjadi konvergensi, seperti yang dinyatakan dalam persamaan berikut:

$$J = \Sigma_i \Sigma_j ||x_i - c_j||^2$$

Keterangan: J merupakan total inersia (jumlah kuadrat dalam kluster), xi adalah data yang ke-i, dan cj adalah centroid dari kluster yang ke-j.

2.7 Evaluasi Cluster Menggunakan Silhouette Score

Evaluasi kualitas cluster dengan memanfaatkan Silhouette Score untuk mengukur keterikatan di dalam cluster dan pemisahan antar cluster (Yulisasih, Herman, Sunardi, & Yuliansyah, 2024). Skor ini berada dalam rentang -1 sampai 1; semakin mendekati angka 1 menandakan bahwa kualitas cluster semakin meningkat.

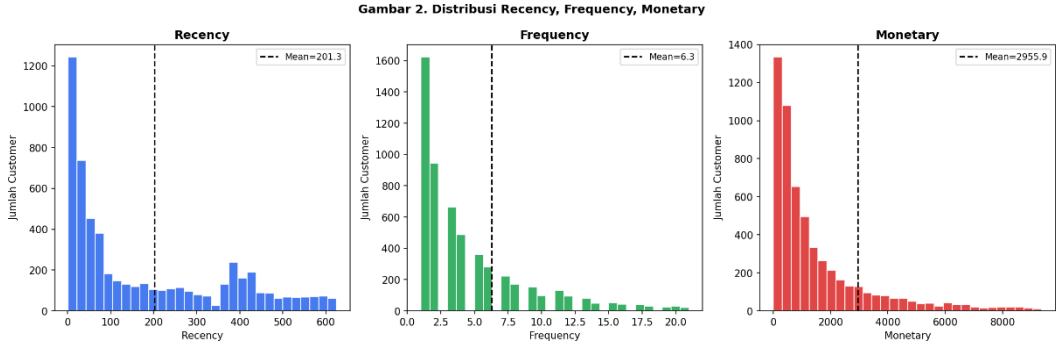
$$S(i) = [b(i) - a(i)] / max\{a(i), b(i)\}$$

Keterangan: a(i) menunjukkan rata-rata jarak data ke-i dalam kelompok yang sama, sementara b(i) menunjukkan rata-rata jarak data ke-i dengan kelompok terdekat lainnya. Uji coba dilakukan untuk k=2 hingga k=8, dengan pemilihan nilai k yang mempertimbangkan faktor Silhouette Score dan juga interpretasi dalam konteks bisnis (Rungruang, Riyapan, Intarasit, Chuarkham, & Muangprathub, 2024).

3. ANALISA DAN PEMBAHASAN

3.1 Hasil Preprocessing dan Distribusi RFM

Setelah data cleaning, dihitung fitur RFM untuk 5.878 pelanggan. Gambar 2 menampilkan distribusi ketiga fitur RFM.



Gambar 2. Distribusi Recency, Frequency, dan Monetary

Distribusi menunjukkan mayoritas pelanggan memiliki frekuensi transaksi rendah dengan sebagian kecil pelanggan melakukan transaksi sangat sering (heavy-tailed distribution), pola umum pada data transaksi ritel (Lewaaelhamd, 2024).

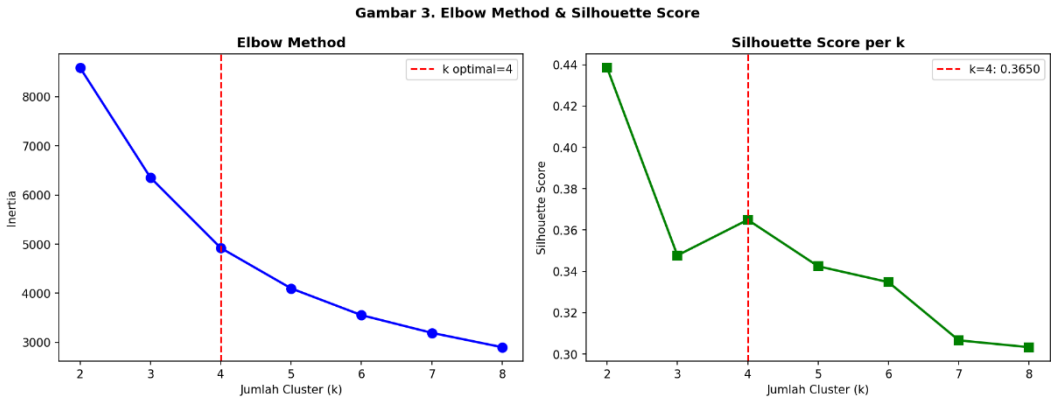
Tabel 2. Statistik Deskriptif RFM

Variabel	Mean	Min	Max
Recency (hari)	201,3	1	739
Frequency (kali)	6,3	1	398
Monetary (GBP)	2.955,9	2,95	580.987,0

3.2 Penentuan Jumlah Cluster Optimal

Penentuan jumlah cluster optimal dilakukan menggunakan Elbow Method dan Silhouette Score pada k=2 hingga k=8, sebagaimana ditampilkan pada Gambar 3 dan Tabel 3.

TEKNOBIS : Teknologi, Bisnis Dan Pendidikan
Volume 4, No. 3 Tahun 2026
ISSN 3024-8256 (media online)
Hal 694-702



Gambar 3. Elbow Method dan Silhouette Score

Tabel 3. Hasil Pengujian Silhouette Score

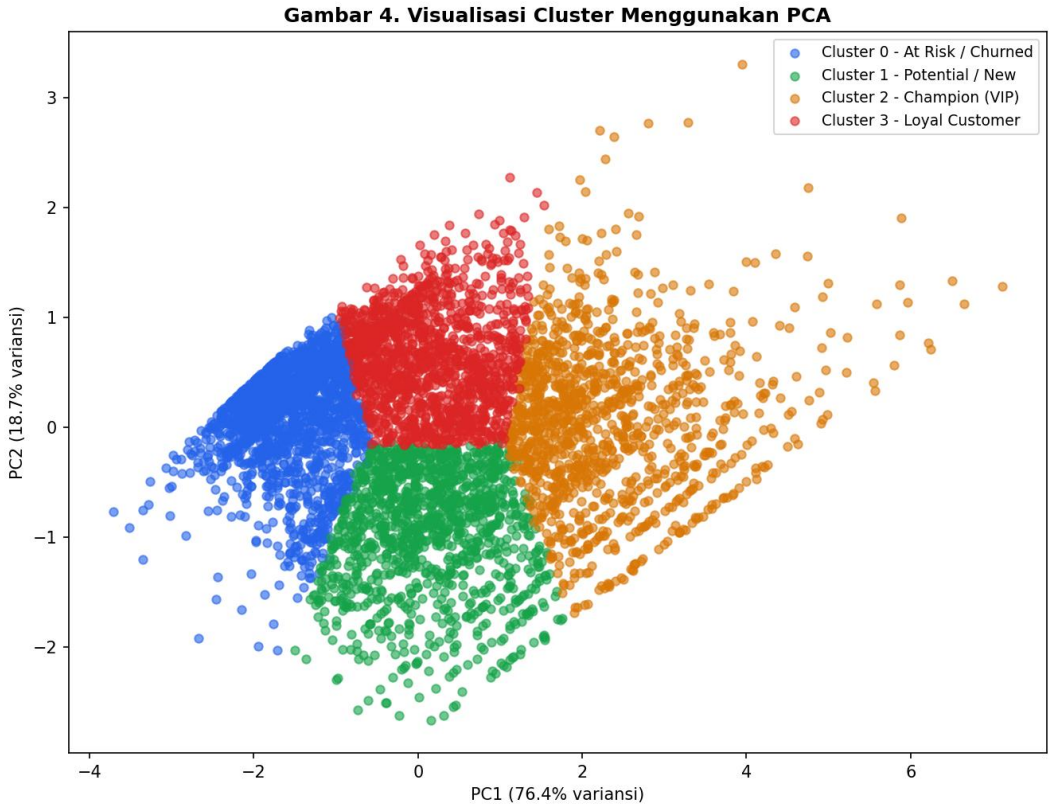
Jumlah Cluster (k)	Silhouette Score
2	0,4386
3	0,3477
4 (terpilih)	0,3650
5	0,3425
6	0,3348
7	0,3066
8	0,3033

Meskipun k=2 menghasilkan Silhouette Score tertinggi (0,4386), nilai ini hanya menghasilkan dua kelompok yang kurang granular untuk strategi pemasaran. Penelitian ini memilih k=4 dengan Silhouette Score 0,3650 karena menghasilkan segmentasi yang lebih representatif dan actionable bagi tim pemasaran, sejalan dengan pendekatan penelitian terdahulu yang mengutamakan keterjelasan bisnis di samping nilai statistik (Irawan et al., 2025; Smaili & Hachimi, 2023).

3.3 Hasil Clustering dan Visualisasi

Hasil clustering divisualisasikan menggunakan Principal Component Analysis (PCA) untuk mereduksi dimensi data ke dalam dua komponen utama, sebagaimana ditampilkan pada Gambar 4.

TEKNOBIS : Teknologi, Bisnis Dan Pendidikan
Volume 4, No. 3 Tahun 2026
ISSN 3024-8256 (media online)
Hal 694-702

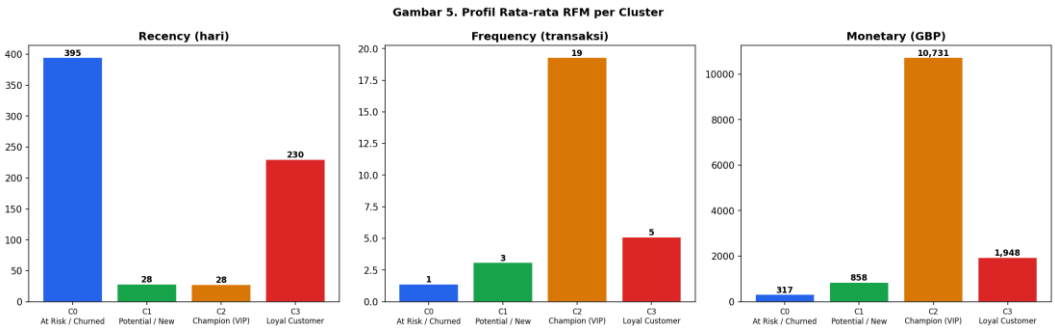


Gambar 4. Visualisasi Cluster Menggunakan PCA

Visualisasi PCA menunjukkan keempat cluster terpisah dengan cukup jelas, mengkonfirmasi kualitas segmentasi yang dihasilkan oleh algoritma K-Means (Tabianan, Velu, & Ravi, 2022).

Tabel 4. Profil Rata-Rata RFM per Cluster

Segmen	**Recency**	**Freq.**	**Monetary**
Champion (VIP)	28 hari	19,3x	£10.731
Loyal Customer	230 hari	5,1x	£1.949
Potential/New	28 hari	3,1x	£858
At Risk/Churned	395 hari	1,4x	£317

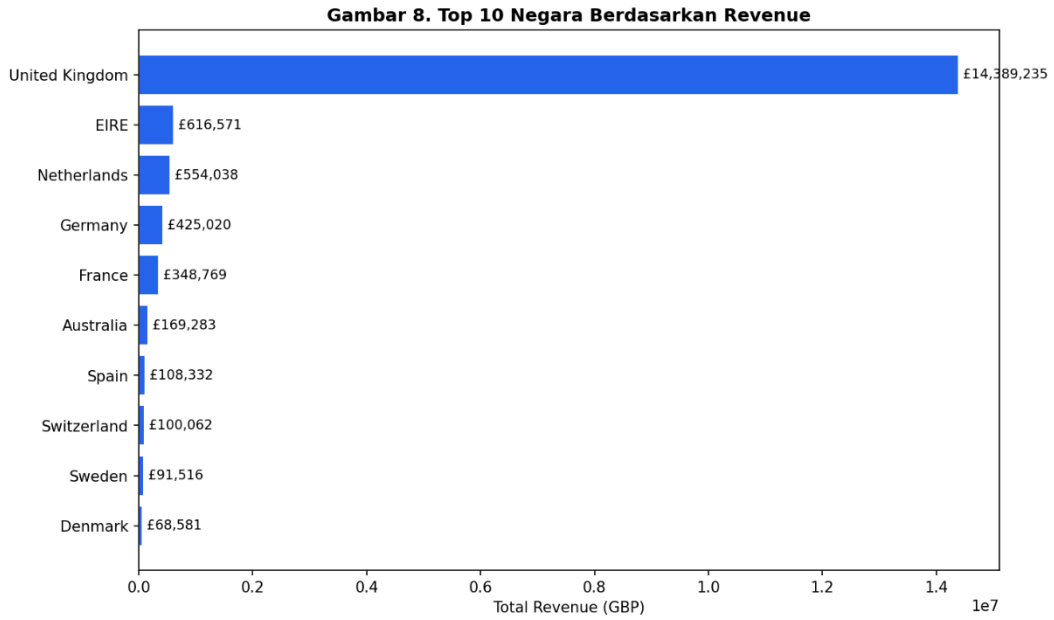


Gambar 5. Profil Rata-rata RFM per Cluster

TEKNOBIS : Teknologi, Bisnis Dan Pendidikan
Volume 4, No. 3 Tahun 2026
ISSN 3024-8256 (media online)
Hal 694-702

Tabel 4 dan Gambar 5 menunjukkan karakteristik rata-rata setiap segmen berdasarkan tiga metrik RFM. Segmen Champion (VIP) menunjukkan karakteristik terbaik dengan Recency terendah (28 hari), Frequency tertinggi (19,3 kali transaksi), dan Monetary tertinggi (£10.731), mengindikasikan pelanggan yang baru saja bertransaksi, sangat sering berbelanja, dan memberikan kontribusi nilai belanja terbesar. Sebaliknya, segmen At Risk/Churned menunjukkan karakteristik paling lemah dengan Recency tertinggi (395 hari), Frequency terendah (1,4 kali), dan Monetary terendah (£317), mengindikasikan pelanggan yang sudah lama tidak bertransaksi dan berisiko tinggi untuk hilang (churn). Segmen Loyal Customer memiliki Frequency dan Monetary sedang namun Recency yang cukup tinggi (230 hari), sementara Potential/New menunjukkan Recency rendah namun Frequency dan Monetary masih kecil, mengindikasikan pelanggan baru yang berpotensi dikembangkan menjadi pelanggan loyal (Rahim et al., 2021).

3.4 Analisis Geografis Pendukung



Gambar 8. Top 10 Negara Berdasarkan Revenue

Inggris (United Kingdom) sebagai negara asal perusahaan mendominasi kontribusi revenue, sejalan dengan karakteristik dataset yang berasal dari perusahaan ritel berbasis UK (Madhiraju, Reddy, & Sasikala, 2024).

3.5 Pembahasa Dan Rekomendasi Bisnis

Hasil segmentasi menunjukkan bahwa kombinasi RFM Analysis dan K-Means Clustering efektif mengidentifikasi empat segmen pelanggan dengan karakteristik yang jelas berbeda, konsisten dengan temuan penelitian terdahulu pada konteks ritel B2C (Irawan et al., 2025; Rahim et al., 2021). Rekomendasi strategi pemasaran untuk masing-masing segmen sebagai berikut:

a. Champion (VIP): memberikan program loyalitas eksklusif, early access produk baru, dan personal account manager untuk menjaga retensi pelanggan bernilai tinggi.
b. Loyal Customer: meningkatkan frekuensi pembelian melalui program referral, bundling produk, dan email marketing yang dipersonalisasi.
c. Potential/New: mendorong pembelian berulang dengan diskon onboarding dan rekomendasi produk berbasis riwayat pembelian.
d. At Risk/Churned: melakukan win-back campaign dengan diskon khusus, survei kepuasan, dan strategi re-engagement untuk mencegah churn lebih lanjut.

Temuan ini sejalan dengan penelitian Rahim et al. (2021) dan Irawan et al. (2025) yang menegaskan bahwa kombinasi RFM dan K-Means menghasilkan profil pelanggan yang actionable

TEKNOBIS : Teknologi, Bisnis Dan Pendidikan
Volume 4, No. 3 Tahun 2026
ISSN 3024-8256 (media online)
Hal 694-702

bagi strategi CRM, meskipun keterbatasan utama metode ini adalah ketergantungan pada data transaksi statis yang belum menangkap perubahan perilaku pelanggan secara dinamis (Abbasimehr & Bahrini, 2022).

4. KESIMPULAN

Penelitian ini berhasil menerapkan RFM Analysis dan K-Means Clustering untuk segmentasi pelanggan pada dataset Online Retail II. Proses data cleaning berhasil menghasilkan 779.425 transaksi bersih dari 1.067.371 transaksi awal, melibatkan 5.878 pelanggan unik. Transformasi logaritmik dan normalisasi StandardScaler terbukti efektif mengatasi skewness data RFM sebelum proses clustering. Jumlah cluster optimal yang dipilih adalah k=4 dengan Silhouette Score 0,3650, menghasilkan empat segmen: Champion/VIP (20,3%), Loyal Customer (24,8%), Potential/New (21,3%), dan At Risk/Churned (33,6%). Segmen Champion/VIP memberikan kontribusi nilai belanja tertinggi (rata-rata £10.731) meskipun jumlahnya relatif kecil, sementara segmen At Risk/Churned merupakan kelompok terbesar yang berisiko hilang. Hasil segmentasi memberikan landasan bagi perumusan strategi pemasaran yang lebih tepat sasaran dan mendukung program customer relationship management (CRM) yang lebih efektif.

Penelitian lanjutan disarankan untuk mengeksplorasi pendekatan dynamic RFM yang memperhitungkan perubahan perilaku pelanggan dari waktu ke waktu, serta membandingkan algoritma clustering lain seperti K-Medoids atau hierarchical clustering yang lebih tahan terhadap outlier (Rungruang et al., 2024; Tabianan et al., 2022). Integrasi dengan model prediktif churn berbasis machine learning juga dapat meningkatkan akurasi identifikasi pelanggan berisiko di masa mendatang.

Penelitian lanjutan disarankan untuk mengeksplorasi pendekatan dynamic RFM yang memperhitungkan perubahan perilaku pelanggan dari waktu ke waktu, serta membandingkan algoritma clustering lain seperti K-Medoids atau hierarchical clustering yang lebih tahan terhadap outlier [16][23]. Integrasi dengan model prediktif churn berbasis machine learning juga dapat meningkatkan akurasi identifikasi pelanggan berisiko di masa mendatang.

UCAPAN TERIMA KASIH

Penulis mengucapkan puji syukur kepada Tuhan Yang Maha Esa atas selesainya jurnal yang berjudul "Penerapan RFM Analysis dan K-Means Clustering untuk Segmentasi Pelanggan pada Dataset Online Retail II". Penulis menyampaikan terima kasih kepada dosen pengampu mata kuliah Data Mining atas arahan dan bimbingannya selama proses penyusunan jurnal ini.

REFERENCES

Abbasimehr, H., & Bahrini, A. (2022). An analytical framework based on the recency, frequency, and monetary model and time series clustering techniques for dynamic segmentation. Expert Systems with Applications, 192, 116373. https://doi.org/10.1016/j.eswa.2021.116373

Christy, A. J., Umamakeswari, A., Priyatharsini, L., & Neyaa, A. (2021). RFM ranking - An effective approach to customer segmentation. Journal of King Saud University - Computer and Information Sciences, 33(10), 1251-1257. https://doi.org/10.1016/j.jksuci.2018.09.004

Irawan, F., Amaluddin, F., & Wijayanti, A. (2025). E-commerce customer segmentation based on RFM analysis using DBSCAN algorithm to improve marketing strategy. Jurnal EMT KITA, 9(4), 1850-1859. https://doi.org/10.35870/emt.v9i4.5375

Lewaaelhamd, I. (2024). Customer segmentation using machine learning model: An application of RFM analysis. Journal of Data Science and Intelligent Systems, 2(1), 29-36. https://doi.org/10.47852/bonviewJDSIS32021293

Madhiraju, B., Reddy, S., & Sasikala, G. (2024). Customer segmentation using RFM analysis. EPRA International Journal of Economic and Business Review, 12(7), 15-22. https://doi.org/10.36713/epra17685

Nasyuha, A. H., Zulham, & Rusydi, I. (2022). Implementation of K-means algorithm in data analysis. TELKOMNIKA (Telecommunication Computing Electronics and Control), 20(2), 307-313. https://doi.org/10.12928/TELKOMNIKA.v20i2.21986

TEKNOBIS : Teknologi, Bisnis Dan Pendidikan
Volume 4, No. 3 Tahun 2026
ISSN 3024-8256 (media online)
Hal 694-702

Pynadath, M. F., Rofin, T. M., & Thomas, S. (2023). Evolution of customer relationship management to data mining-based customer relationship management: A scientometric analysis. Quality & Quantity, 57(4), 3241-3272. https://doi.org/10.1007/s11135-022-01500-y

Rahim, M. A., Mushafiq, M., Khan, S., & Arain, Z. A. (2021). RFM-based repurchase behavior for customer classification and segmentation. Journal of Retailing and Consumer Services, 61, 102566. https://doi.org/10.1016/j.jretconser.2021.102566

Rungruang, C., Riyapan, P., Intarasit, A., Chuarkham, K., & Muangprathub, J. (2024). RFM model customer segmentation based on hierarchical approach using FCA. Expert Systems with Applications, 237, 121449. https://doi.org/10.1016/j.eswa.2023.121449

Saha, L., Tripathy, H. K., Nayak, S. R., Bhoi, A. K., & Barsocchi, P. (2021). Amalgamation of customer relationship management and data analytics in different business sectors-A systematic literature review. Sustainability, 13(9), 5279. https://doi.org/10.3390/su13095279

Smaili, M. Y., & Hachimi, H. (2023). New RFM-D classification model for improving customer analysis and response prediction. Ain Shams Engineering Journal, 14(12), 102254. https://doi.org/10.1016/j.asej.2023.102254

Tabianan, K., Velu, S., & Ravi, V. (2022). K-Means clustering approach for intelligent customer segmentation using customer purchase behavior data. Sustainability, 14(12), 7243. https://doi.org/10.3390/su14127243

Yulisasih, B. N., Herman, H., Sunardi, S., & Yuliansyah, H. (2024). Evaluation of K-Means clustering using silhouette score method on customer segmentation. ILKOM Jurnal Ilmiah, 16(3), 330-342. https://doi.org/10.33096/ilkom.v16i3.2325.330-342