

Pemanfaatan Data Mining untuk Identifikasi Pola Perilaku Pelanggan Menggunakan Model RFM dan Algoritma K-Means dalam Mendukung Strategi Bisnis

Sabrina Shoraida Fitriah^{1*}, Steven Wijaya², Revina Dwi Lestari³

^{1,2,3}Fakultas Ilmu Komputer, Program Studi Sistem Informasi, Universitas Pamulang, Tangerang Selatan, Indonesia

¹sabrinashoraidafitriah22@gmail.com ²stevenwijaya012@gmail.com ³revinasimamora18@gmail.com

(Email Correspondence: sabrinashoraidafitriah22@gmail.com)

Received: 4 April 2026 | Revised: 8 April 2026 | Accepted: 30 April 2026

Abstrak

Transformasi digital pada sektor bisnis menyebabkan peningkatan volume data transaksi penjualan yang dapat dimanfaatkan untuk mendukung pengambilan keputusan berbasis data. Namun, banyak perusahaan masih belum memanfaatkan data transaksi secara optimal untuk memperoleh informasi strategis. Penelitian ini bertujuan mengidentifikasi pola perilaku pelanggan berdasarkan data transaksi penjualan menggunakan teknik data mining guna mendukung strategi bisnis yang lebih efektif. Penelitian menggunakan dataset Online Retail II dengan pendekatan CRISP-DM yang meliputi tahap pemahaman bisnis, persiapan data, pemodelan, dan evaluasi. Data yang telah dibersihkan ditransformasikan menggunakan model Recency, Frequency, Monetary (RFM) untuk merepresentasikan perilaku pelanggan. Proses segmentasi dilakukan menggunakan algoritma K-Means dengan jumlah cluster optimal yang ditentukan melalui Elbow Method. Hasil penelitian menunjukkan bahwa jumlah cluster optimal adalah tiga cluster dengan nilai Silhouette Score sebesar 0,5321 yang menunjukkan kualitas cluster cukup baik. Cluster yang terbentuk terdiri atas pelanggan menengah, pelanggan tidak aktif, dan pelanggan loyal. Selain itu, segmentasi RFM menunjukkan bahwa sebagian besar pelanggan berada pada kategori pasif dan berisiko, sementara sebagian lainnya termasuk pelanggan setia dan pelanggan terbaik. Hasil penelitian membuktikan bahwa kombinasi model RFM dan algoritma K-Means mampu mengidentifikasi pola perilaku pelanggan secara efektif serta mendukung penyusunan strategi pemasaran yang lebih tepat sasaran dan berbasis data.

Kata kunci: data mining, klasterisasi, RFM, K-Means, segmentasi pelanggan, strategi bisnis.

Abstract

Digital transformation in the business sector has led to an increase in the volume of sales transaction data that can be leveraged to support data-driven decision-making. However, many companies have yet to fully utilize transaction data to gain strategic insights. This study aims to identify customer behavior patterns based on sales transaction data using data mining techniques to support more effective business strategies. The study uses the Online Retail II dataset with the CRISP-DM approach, which includes the business understanding, data preparation, modeling, and evaluation stages. The cleaned data was transformed using the Recency, Frequency, Monetary (RFM) model to represent customer behavior. The segmentation process was performed using the K-Means algorithm, with the optimal number of clusters determined via the Elbow Method. The results indicate that the optimal number of clusters is three, with a Silhouette Score of 0.5321, indicating reasonably good cluster quality. The resulting clusters consist of mid-tier customers, inactive customers, and loyal customers. Furthermore, the RFM segmentation reveals that the majority of customers fall into the passive and at-risk categories, while the remainder include loyal customers and top-tier customers. The research findings demonstrate that the combination of the RFM model and the K-Means algorithm effectively identifies customer behavior patterns and supports the development of more targeted and data-driven marketing strategies.

Keywords: data mining, clustering, RFM, K-Means, customer segmentation, business strategy.

1. PENDAHULUAN

Transformasi digital pada sektor bisnis telah mendorong peningkatan jumlah data yang dihasilkan oleh berbagai aktivitas bisnis, termasuk data transaksi penjualan pada sektor retail. Data transaksi yang tersimpan dalam jumlah besar sebenarnya mengandung informasi penting yang dapat dimanfaatkan untuk memahami perilaku pelanggan dan mendukung pengambilan keputusan bisnis. Namun, pada banyak perusahaan, data transaksi masih digunakan sebagai arsip operasional dan belum dimanfaatkan secara optimal untuk menghasilkan informasi yang bernilai strategis. Kondisi tersebut menyebabkan perusahaan kesulitan dalam memahami pola pembelian pelanggan dan menyusun strategi pemasaran yang tepat sasaran [1], [2].

Dalam lingkungan bisnis yang semakin kompetitif, pelanggan merupakan aset penting yang menentukan keberhasilan perusahaan. Pemahaman terhadap karakteristik dan perilaku pelanggan menjadi faktor utama dalam penyusunan strategi pemasaran yang efektif. Segmentasi pelanggan memungkinkan perusahaan untuk mengelompokkan pelanggan berdasarkan karakteristik dan perilaku yang serupa sehingga perusahaan dapat memberikan layanan, promosi, dan penawaran yang lebih sesuai dengan kebutuhan setiap kelompok pelanggan. Pendekatan ini membantu perusahaan meningkatkan loyalitas pelanggan sekaligus mengoptimalkan sumber daya pemasaran yang dimiliki [3], [4].

Salah satu pendekatan yang dapat digunakan untuk melakukan segmentasi pelanggan adalah data mining. Data mining merupakan proses eksplorasi dan analisis data untuk menemukan pola, hubungan, dan informasi tersembunyi yang dapat digunakan dalam mendukung pengambilan keputusan. Dalam bidang pemasaran, data mining banyak dimanfaatkan untuk menganalisis perilaku pelanggan, mengidentifikasi pola pembelian, serta mengembangkan strategi pemasaran berbasis data. Teknik data mining yang sering digunakan dalam segmentasi pelanggan adalah clustering karena mampu mengelompokkan data berdasarkan tingkat kemiripan karakteristik tanpa memerlukan label kelas sebelumnya [5], [6].

Algoritma K-Means merupakan salah satu metode clustering yang paling banyak digunakan karena memiliki proses yang sederhana, efisien, dan mampu menghasilkan kelompok data yang mudah diinterpretasikan. Dalam penerapannya, algoritma K-Means dapat dikombinasikan dengan model Recency, Frequency, Monetary (RFM) untuk menghasilkan segmentasi pelanggan yang lebih akurat berdasarkan perilaku transaksi. Model RFM mengukur nilai pelanggan berdasarkan waktu transaksi terakhir, frekuensi transaksi, dan total nilai pembelian sehingga mampu memberikan gambaran yang lebih komprehensif mengenai karakteristik pelanggan. Beberapa penelitian sebelumnya menunjukkan bahwa kombinasi model RFM dan algoritma K-Means efektif digunakan dalam segmentasi pelanggan serta mendukung penyusunan strategi pemasaran yang lebih tepat sasaran [7], [8].

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengidentifikasi pola perilaku pelanggan pada data transaksi penjualan menggunakan model RFM dan algoritma K-Means. Hasil segmentasi yang diperoleh diharapkan dapat membantu perusahaan dalam memahami karakteristik pelanggan, meningkatkan efektivitas strategi pemasaran, mengurangi risiko kehilangan pelanggan, serta mendukung pengambilan keputusan bisnis yang lebih tepat sasaran dan berbasis data.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan CRISP-DM (Cross Industry Standard Process for Data Mining) sebagai kerangka kerja dalam proses analisis data. Pendekatan ini dipilih karena memberikan tahapan yang sistematis dalam menyelesaikan permasalahan data mining mulai dari pemahaman masalah bisnis hingga interpretasi hasil yang diperoleh. Tahapan yang diterapkan dalam penelitian ini meliputi Business Understanding, Data Preparation, Modeling, dan Evaluation. Pendekatan serupa juga banyak digunakan dalam penelitian segmentasi pelanggan karena mampu menghasilkan alur analisis yang terstruktur dan mudah diimplementasikan [1], [2].

Pada tahap Business Understanding, dilakukan identifikasi permasalahan bisnis terkait belum adanya segmentasi pelanggan yang mampu menggambarkan perilaku pembelian pelanggan berdasarkan data transaksi. Kondisi tersebut menyebabkan perusahaan kesulitan dalam menentukan strategi pemasaran yang sesuai dengan karakteristik pelanggan yang berbeda-beda. Oleh karena itu, penelitian ini berfokus pada pemanfaatan data transaksi untuk mengidentifikasi pola perilaku pelanggan dan mengelompokkannya ke dalam beberapa segmen yang memiliki karakteristik serupa [3].

Tahap Data Preparation dilakukan menggunakan dataset Online Retail II yang berisi data transaksi pelanggan retail online. Dataset terdiri atas atribut Invoice, StockCode, Description, Quantity, InvoiceDate, Price, Customer ID, dan Country. Pada tahap ini dilakukan proses pembersihan data dengan menghapus nilai kosong (missing value), data duplikat, transaksi pembatalan (return transaction), serta data yang tidak relevan. Setelah proses pembersihan selesai, dilakukan transformasi data menggunakan model Recency, Frequency, Monetary (RFM). Variabel Recency digunakan untuk mengukur waktu sejak transaksi terakhir pelanggan, Frequency digunakan untuk mengukur jumlah transaksi pelanggan, sedangkan Monetary digunakan untuk mengukur total nilai transaksi pelanggan. Model RFM dipilih karena mampu menggambarkan nilai pelanggan secara objektif berdasarkan aktivitas transaksi yang dilakukan [4], [5].

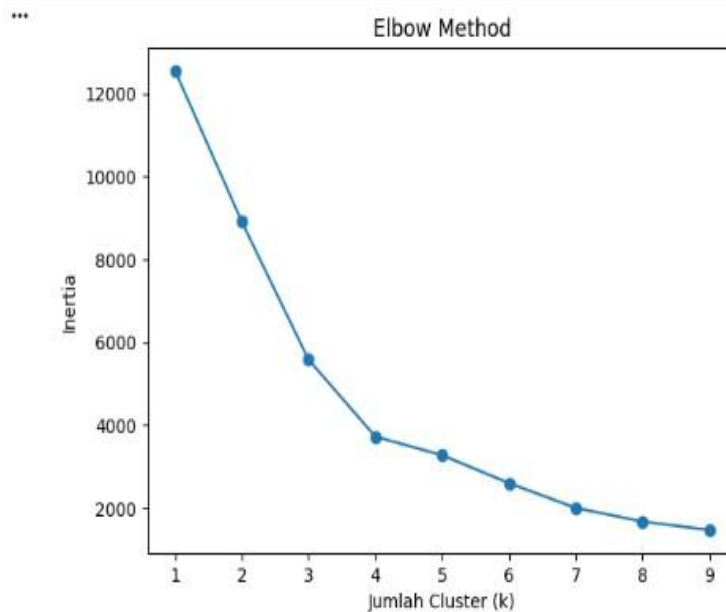
Tahap Modeling dilakukan dengan menerapkan algoritma K-Means Clustering pada data RFM yang telah dinormalisasi menggunakan StandardScaler. Sebelum proses clustering dilakukan, jumlah cluster optimal ditentukan menggunakan Elbow Method dengan menganalisis perubahan nilai inertia pada beberapa jumlah cluster. Setelah jumlah cluster optimal diperoleh, algoritma K-Means digunakan untuk mengelompokkan pelanggan berdasarkan kemiripan karakteristik RFM yang dimiliki. Penggunaan K-Means dipilih karena memiliki kemampuan yang baik dalam membentuk kelompok pelanggan yang homogen serta mudah diinterpretasikan dalam konteks segmentasi pelanggan [6], [7].

Tahap terakhir adalah Evaluation, yaitu mengevaluasi kualitas cluster yang terbentuk menggunakan Silhouette Score. Metode ini digunakan untuk mengukur tingkat kemiripan data dalam cluster yang sama serta tingkat perbedaan antar cluster. Semakin tinggi nilai Silhouette Score yang diperoleh, maka semakin baik kualitas segmentasi yang dihasilkan. Selanjutnya dilakukan interpretasi terhadap karakteristik setiap cluster berdasarkan nilai rata-rata Recency, Frequency, dan Monetary untuk menghasilkan rekomendasi strategi bisnis yang dapat digunakan dalam meningkatkan loyalitas pelanggan, mengoptimalkan aktivitas pemasaran, dan mendukung pengambilan keputusan berbasis data [8], [9].

3. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan dataset Online Retail II yang berisi data transaksi pelanggan retail online. Sebelum dilakukan proses analisis, data terlebih dahulu melalui tahap pembersihan untuk menghilangkan data yang tidak lengkap, data duplikat, transaksi pembatalan, dan data yang tidak relevan. Proses ini bertujuan untuk meningkatkan kualitas data sehingga hasil segmentasi yang diperoleh menjadi lebih akurat. Hasil data yang telah dibersihkan kemudian digunakan sebagai dasar dalam pembentukan variabel Recency, Frequency, dan Monetary (RFM).

Setelah data dinyatakan bersih, dilakukan pembentukan model RFM yang digunakan untuk menggambarkan perilaku pelanggan berdasarkan aktivitas transaksi. Variabel Recency digunakan untuk mengukur jarak waktu sejak transaksi terakhir pelanggan, Frequency digunakan untuk menghitung jumlah transaksi yang dilakukan pelanggan, sedangkan Monetary digunakan untuk menghitung total nilai transaksi pelanggan. Ketiga variabel tersebut digunakan sebagai dasar dalam proses segmentasi pelanggan karena mampu merepresentasikan nilai pelanggan secara objektif berdasarkan data historis transaksi.



Gambar 1. Hasil Elbow Method

Menunjukkan hasil penerapan Elbow Method untuk menentukan jumlah cluster optimal pada algoritma K-Means. Grafik memperlihatkan hubungan antara jumlah cluster (k) dengan nilai inertia. Terlihat bahwa penurunan inertia terjadi secara signifikan dari $k=1$ hingga $k=4$, kemudian laju penurunan mulai melambat pada cluster berikutnya. Berdasarkan grafik Elbow Method, titik siku terlihat pada rentang $k=3$ hingga $k=4$. Dalam penelitian ini dipilih tiga cluster karena memberikan segmentasi yang lebih sederhana dan mudah diinterpretasikan. Dengan demikian, proses clustering selanjutnya dilakukan menggunakan tiga cluster.

*** Hasil Cluster:

Customer ID	Recency	Frequency	Monetary	Kategori	Cluster
12347.00	3	7	2904.17	Tinggi	0
12348.00	249	3	111.08	Rendah	1
12349.00	19	1	1105.40	Tinggi	0
12350.00	311	1	294.40	Sedang	1
12352.00	37	7	1347.49	Tinggi	0

Gambar 2. Hasil Pengelompokan Pelanggan Menggunakan K-Means

Setelah jumlah cluster ditentukan, dilakukan proses clustering menggunakan algoritma K-Means.. Visualisasi tersebut menunjukkan bahwa pelanggan berhasil dikelompokkan ke dalam tiga kelompok utama yang memiliki karakteristik berbeda berdasarkan nilai Recency dan Monetary. Pemisahan cluster terlihat cukup jelas sehingga menunjukkan bahwa algoritma K-Means mampu mengidentifikasi pola perilaku pelanggan berdasarkan data transaksi yang dimiliki.

4. EVALUASI (SILHOUETTE SCORE)

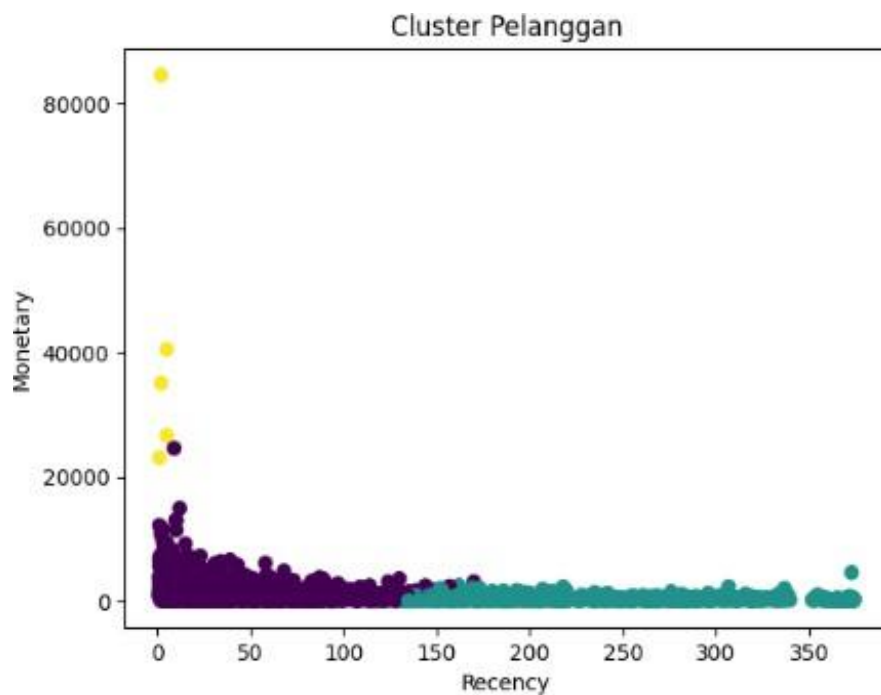
```
from sklearn.metrics import silhouette_score

score = silhouette_score(X_scaled, rfm['Cluster'])
print("Silhouette Score:", score)
```

Silhouette Score: 0.5321809745694568

Gambar 3. Silhouette Score

Untuk mengukur kualitas cluster yang terbentuk, dilakukan evaluasi menggunakan Silhouette Score. Hasil evaluasi menunjukkan nilai Silhouette Score sebesar 0,5321. Nilai tersebut menunjukkan bahwa cluster yang dihasilkan memiliki kualitas yang cukup baik karena anggota dalam satu cluster memiliki tingkat kemiripan yang tinggi dan memiliki perbedaan yang cukup jelas dengan anggota cluster lainnya. Dengan demikian, hasil segmentasi dapat digunakan sebagai dasar dalam proses analisis perilaku pelanggan dan penyusunan strategi bisnis.



Gambar 4. Cluster Pelanggan

Menunjukkan hasil clustering pelanggan menggunakan algoritma K-Means berdasarkan nilai Recency dan Monetary. Pelanggan berhasil dikelompokkan ke dalam tiga cluster yang memiliki karakteristik berbeda. Cluster dengan nilai Monetary tinggi dan Recency rendah menunjukkan pelanggan loyal, sedangkan cluster dengan Recency tinggi dan Monetary rendah menunjukkan pelanggan yang kurang aktif. Hasil ini menunjukkan bahwa K-Means mampu membedakan perilaku pelanggan berdasarkan data transaksi.

*** Profil Cluster:

	Recency	Frequency	Monetary
Cluster			
0	40.52	4.66	1155.35
1	242.25	1.55	316.72
2	3.00	122.40	41971.13

Gambar 5 .Profil Cluster

Karakteristik masing-masing cluster dianalisis menggunakan rata-rata nilai RFM. Hasil analisis menunjukkan bahwa Cluster 0 memiliki nilai rata-rata Recency sebesar 40,52 hari, Frequency sebesar 4,66 transaksi, dan Monetary sebesar 1.155,35. Karakteristik tersebut menunjukkan bahwa pelanggan pada cluster ini masih cukup aktif melakukan transaksi dan memberikan kontribusi pendapatan yang cukup baik. Oleh karena itu, kelompok ini dapat dikategorikan sebagai pelanggan menengah yang berpotensi ditingkatkan menjadi pelanggan loyal.

Cluster 1 memiliki nilai rata-rata Recency sebesar 242,25 hari, Frequency sebesar 1,55 transaksi, dan Monetary sebesar 316,72. Nilai Recency yang tinggi menunjukkan bahwa pelanggan pada kelompok ini sudah lama tidak melakukan transaksi. Selain itu, frekuensi transaksi dan nilai pembelian yang rendah menunjukkan bahwa pelanggan pada cluster ini memiliki aktivitas yang rendah dan berisiko mengalami churn. Dengan demikian, cluster ini dikategorikan sebagai pelanggan tidak aktif yang memerlukan strategi retensi pelanggan agar dapat kembali melakukan transaksi.

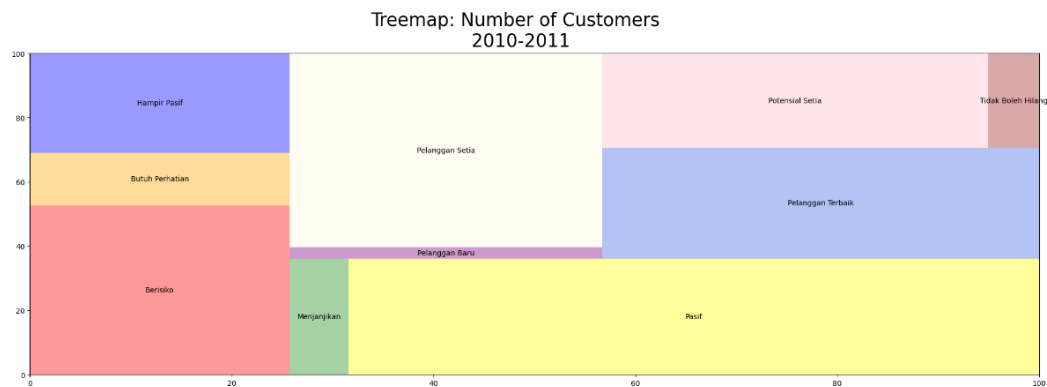
Sementara itu, Cluster 2 memiliki nilai rata-rata Recency sebesar 3 hari, Frequency sebesar 122,40 transaksi, dan Monetary sebesar 41.971,13. Nilai tersebut merupakan yang tertinggi dibandingkan cluster lainnya. Hasil ini menunjukkan bahwa pelanggan pada cluster ini sangat aktif melakukan transaksi dan memberikan kontribusi pendapatan terbesar bagi perusahaan. Oleh karena itu, cluster ini dikategorikan sebagai pelanggan loyal atau pelanggan bernilai tinggi yang perlu dipertahankan melalui program loyalitas, pemberian reward, serta layanan eksklusif.

```
rfmStats2 = rfm(["Segment", "Recency", "Frequency", "Monetary"]).groupby("Segment").agg(["mean", "median", "count", "std"])
rfmStats2.columns = rfmStats2.columns.map(lambda x: x.strip('_').strip('.'))
rfmStats2
```

Segment	Recency_mean	Recency_median	Recency_count	Recency_std	Frequency_mean	Frequency_median	Frequency_count	Frequency_std	Monetary_mean	Monetary_median	Monetary_count	Monetary_std
Berisiko	153.73	142.00	567	66.93	2.62	2.00	567	0.82	645.90	482.51	567	569.63
Butuh Perhatian	52.85	53.00	177	10.65	2.13	2.00	177	0.34	595.95	457.72	177	552.83
Hampir Pasif	54.22	54.00	333	10.71	1.09	1.00	333	0.29	339.83	236.18	333	370.71
Menjanjikan	24.85	25.00	88	5.62	1.00	1.00	88	0.00	242.35	167.76	88	218.33
Pasif	218.65	226.00	1034	91.78	1.05	1.00	1034	0.22	256.72	185.18	1034	256.68
Pelanggan Baru	8.50	9.50	46	4.46	1.00	1.00	46	0.00	227.37	197.10	46	135.59
Pelanggan Setia	34.61	31.00	781	15.51	5.67	5.00	781	3.33	1599.71	1215.82	781	1261.88
Pelanggan Terbaik	6.68	6.00	624	3.95	9.96	7.00	624	11.36	2697.31	1698.13	624	4644.95
Potensial Setia	17.83	18.00	471	9.39	1.85	2.00	471	0.58	477.76	384.18	471	380.81
Tidak Boleh Hilang	125.26	107.00	62	51.97	6.84	6.00	62	2.47	1246.86	1065.01	62	760.38

Gambar 6. Profil Segmentasi Pelanggan Berdasarkan Model RFM

menunjukkan karakteristik masing-masing segmen pelanggan berdasarkan nilai rata-rata Recency, Frequency, dan Monetary (RFM). Hasil analisis menunjukkan bahwa segmen Pelanggan Terbaik memiliki nilai Frequency dan Monetary tertinggi serta nilai Recency yang rendah, yang menunjukkan tingkat loyalitas dan kontribusi pendapatan yang tinggi. Segmen Pelanggan Setia juga menunjukkan aktivitas transaksi yang cukup tinggi dan berpotensi untuk dipertahankan melalui program loyalitas. Sebaliknya, segmen Pasif, Berisiko, dan Tidak Boleh Hilang memiliki nilai Recency yang relatif tinggi, yang mengindikasikan pelanggan sudah lama tidak melakukan transaksi sehingga memerlukan strategi retensi dan reaktivasi pelanggan.



Gambar 7. Visualisasi Distribusi Jumlah Pelanggan Pada Segmen RFM

Menunjukkan distribusi jumlah pelanggan pada masing-masing segmen RFM menggunakan visualisasi treemap. Luas area pada setiap kotak merepresentasikan jumlah pelanggan dalam segmen tersebut. Hasil visualisasi menunjukkan bahwa segmen Pasif merupakan kelompok terbesar dengan jumlah pelanggan terbanyak, diikuti oleh Pelanggan Setia, Pelanggan Terbaik, dan Berisiko. Sementara itu, segmen Pelanggan Baru dan Tidak Boleh Hilang memiliki jumlah pelanggan yang relatif lebih sedikit. Temuan ini menunjukkan bahwa sebagian besar pelanggan berada pada kelompok dengan tingkat aktivitas transaksi yang rendah sehingga perusahaan perlu memfokuskan strategi bisnis pada program retensi dan reaktivasi pelanggan. Di sisi lain, pelanggan yang termasuk dalam kategori Pelanggan Terbaik dan Pelanggan Setia perlu dipertahankan melalui program loyalitas agar tetap memberikan kontribusi terhadap pendapatan perusahaan.

Secara keseluruhan, hasil penelitian menunjukkan bahwa kombinasi model RFM dan algoritma K-Means mampu mengidentifikasi pola perilaku pelanggan secara efektif. Segmentasi yang dihasilkan memberikan informasi yang bermanfaat bagi perusahaan dalam memahami karakteristik pelanggan dan menyusun strategi pemasaran yang lebih tepat sasaran. Pelanggan loyal dapat dipertahankan melalui program loyalitas dan reward, pelanggan menengah dapat ditingkatkan melalui promosi yang dipersonalisasi, sedangkan pelanggan tidak aktif dapat menjadi target utama dalam kampanye re-engagement dan retensi pelanggan. Dengan demikian, hasil segmentasi pelanggan yang diperoleh dapat mendukung pengambilan keputusan bisnis yang lebih efektif dan berbasis data.

4. KESIMPULAN

Penelitian ini berhasil menerapkan teknik data mining menggunakan model Recency, Frequency, Monetary (RFM) dan algoritma K-Means untuk mengidentifikasi pola perilaku pelanggan berdasarkan data transaksi penjualan pada dataset Online Retail II. Hasil penentuan jumlah cluster menggunakan Elbow Method menunjukkan bahwa jumlah cluster optimal adalah tiga cluster. Evaluasi menggunakan Silhouette Score menghasilkan nilai sebesar 0,5321 yang menunjukkan bahwa kualitas cluster yang terbentuk cukup baik dan mampu menggambarkan perbedaan karakteristik pelanggan secara jelas. Hasil clustering menunjukkan bahwa pelanggan dapat dikelompokkan menjadi tiga kategori utama, yaitu pelanggan menengah, pelanggan tidak aktif, dan pelanggan loyal. Cluster pelanggan loyal memiliki nilai Frequency dan Monetary tertinggi serta nilai Recency terendah, sehingga memberikan kontribusi terbesar terhadap pendapatan perusahaan. Sebaliknya, pelanggan tidak aktif memiliki nilai Recency yang tinggi serta aktivitas transaksi yang rendah sehingga berpotensi mengalami churn. Selain itu, hasil segmentasi RFM menunjukkan bahwa sebagian besar pelanggan berada pada kategori pasif dan berisiko, sementara pelanggan terbaik dan pelanggan setia memberikan kontribusi transaksi yang lebih tinggi. Berdasarkan hasil tersebut, segmentasi pelanggan yang dihasilkan dapat dimanfaatkan sebagai dasar dalam penyusunan strategi bisnis yang lebih efektif, seperti program loyalitas untuk pelanggan bernilai tinggi, promosi yang dipersonalisasi untuk pelanggan menengah, serta strategi reaktivasi bagi pelanggan yang tidak aktif. Dengan demikian, penerapan model RFM dan algoritma K-Means terbukti mampu membantu perusahaan memahami perilaku pelanggan dan mendukung pengambilan keputusan bisnis yang lebih tepat sasaran berbasis data.

DAFTAR PUSTAKA

- [1]. Fajar, M., Adji, M., & Dwilestari, G. (2025). *ANALISIS DATA TRANSAKSI PENJUALAN BARANG MENGGUNAKAN TEKNIK K-MEANS CLUSTERING*. 9(1), 619–625.
- [2]. Sulisty, G. B., Hasan, N., Kiswati, S., Natalia, F., Muningsih, E., & Korespondensi, P. (2025). Segmentasi Pelanggan dan Optimalisasi Penjualan pada Data Retail Online Berbasis Model RFM. *Computer and Network Technology*, 5(1), 16–22. <http://jurnal.bsi.ac.id/index.php/conten16>
- [3]. Niko, N. S., Rahman, A., Marini Umi Atmaja, D., & Basri, A. (2023). Klasterisasi Stok Produk Retail Untuk Menentukan Pergerakan Kebutuhan Konsumen Dengan Algoritma K-Means. *Bulletin of Information Technology (BIT)*, 4(3), 306–312. <https://doi.org/10.47065/bit.v4i3.736>
- [4]. Nugraha, R. D., Adelia, D. D., & Rivaldi, D. (2025). Segmentasi Pelanggan Retail Berbasis Perilaku menggunakan Algoritma K-Means Clustering. *Digital Transformation Technology*, 5(2), 141–148. <https://itscience.org/jurnal/index.php/digitech/article/view/6340>
- [5]. Wibowo, A., & Handoko, A. R. (2020). Segmentasi Pelanggan Ritel Produk Farmasi Obat Menggunakan Metode Data Mining Klasterisasi Dengan Analisis Recency Frequency Monetary (RFM) Termodifikasi. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 7(3), 573. <https://doi.org/10.25126/jtiik.2020702925>
- [6]. Mirantika, N., Syamfithriani, T. S., & Trisudarmo, R. (2023). Implementasi Algoritma K-Medoids Clustering Untuk Menentukan Segmentasi Pelanggan. *Jurnal Nuansa Informatika*, 17(1), 196–204.
- [7]. Hendrastuty, N. (2024). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa. *Jurnal Ilmiah Informatika Dan Ilmu Komputer (JIMA-ILKOM)*, 3(1), 46–56. <https://doi.org/10.58602/jima-ilkom.v3i1.26>
- [8]. 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. <https://doi.org/10.34010/jati.v13i2.10090>
- [9]. Kumar, Sunil; Lochab, A.; Mishra, M. (2021). The Impact of Business Model Innovation (BMI) and Organisational Values on Firm Performance: Mediating Role of Corporate Sustainability. *Corporate Sustainability-A New Paradigm*, 32(3), 56–71.
- [10]. Baihaqi, W. M. (2024). *PENERAPAN DATA MINING DALAM MENGANALISIS*. 1(2), 73–79.
- [11]. Daniati, E. (n.d.). *Analisis Segmentasi Kebiasaan Belanja Berdasarkan Data Penjualan Menggunakan Metode K-Means*. 9, 1743–1752.
- [12]. Aini, N., Suarna, N., & Prihartono, W. (2023). *STRATEGI PROMOSI ONLINE SHOP DI MEDIA SOSIAL MELALUI PENERAPAN DATA MINING DENGAN METODE K-MEANS CLUSTERING*. 7(6), 3628–3632.
- [13]. Pratama, M., & Zulfiandry, R. (2025). *Implementasi Data Mining Untuk Analisis Prilaku Konsumen Dalam Meningkatkan Penjualan Di Shopee*. 21(2), 703–709.
- [14]. Fajar, M., Adam, S., Putra, B., Puteri, S. I., Fajrissiddiq, A., & Sani, L. (2025). *Eksplorasi dan Analisis Data Mining untuk Prediksi Pola Konsumen Menggunakan Teknik Klasifikasi dan Clustering*.
- [15]. Putra, E. R., & Septiani, S. (2026). *PEMANFAATAN DATA MINING UNTUK IDENTIFIKASI POLA PEMBELIAN PRODUK PLATFORM PERDAGANGAN ELEKTRONIK E-COMMERCE PLAZA BANTEN*. 9(1), 160–172.