

ARTICLE

Optimalisasi Strategi Pemasaran melalui Segmentasi Pelanggan dengan Analisis RFM dan Algoritma K-Means untuk Bisnis Ritel

Marketing Strategy Optimization through Customer Segmentation with RFM Analysis and K-Means Algorithm for Retail Businesses

Aliya Anisa Rahma,^{*} Ahmad Faqih, dan Ade Rizki Rinaldi

Teknik Informatika, Fakultas Ilmu Komputer, STMIK IKMI Cirebon, Kota Cirebon, Jawa Barat, Indonesia

*Penulis Korespondensi: raliya813@gmail.com

(Disubmit 09-01-25; Diterima 10-03-25; Dipublikasikan online pada 20-06-25)

Abstrak

Industri ritel yang kompetitif memerlukan pemahaman mendalam tentang kebutuhan pelanggan untuk menyusun strategi pemasaran yang relevan dan efektif. Penelitian ini bertujuan untuk melakukan segmentasi pelanggan di Toko Mitra 10 Cirebon menggunakan analisis Recency, Frequency, dan Monetary (RFM) yang dikombinasikan dengan algoritma K-Means. Data yang digunakan berasal dari catatan transaksi pelanggan dalam periode tertentu. Nilai RFM dihitung untuk setiap pelanggan berdasarkan Recency (waktu sejak transaksi terakhir), Frequency (jumlah transaksi), dan Monetary (total nilai transaksi). Metode K-Means digunakan untuk mengelompokkan pelanggan menjadi beberapa segmen, dengan jumlah kluster optimal ditentukan melalui metode elbow. Analisis menghasilkan tiga segmen utama: Lost Customers, dengan Recency tinggi, Frequency rendah, dan Monetary rendah; Potential Loyalists, dengan Frequency sedang dan Monetary bervariasi; serta Loyal Customers, dengan Frequency tinggi dan kontribusi Monetary signifikan. Hasil segmentasi ini mendukung penyusunan strategi pemasaran yang berbeda untuk setiap kluster: kampanye reaktivasi untuk Lost Customers, program loyalitas untuk Potential Loyalists, dan layanan eksklusif untuk Loyal Customers. Pendekatan berbasis data ini meningkatkan efektivitas pemasaran, loyalitas pelanggan, serta kontribusi pendapatan, sekaligus menegaskan pentingnya analisis data dalam pengambilan keputusan pemasaran yang relevan dan personal.

Kata kunci: Segmentasi Pelanggan; RFM; K-Means; Strategi Pemasaran; Toko Mitra 10 Cirebon

Abstract

This study segments customers at Mitra 10 Cirebon using RFM (Recency, Frequency, Monetary) analysis and K-Means clustering, based on transaction data. RFM values are calculated for each customer, and K-Means groups them into clusters, with the optimal number determined by the elbow method. Three segments are identified: Lost Customers (high Recency, low Frequency, low Monetary), Potential Loyalists (moderate Frequency, varying Monetary), and Loyal Customers (high Frequency, high Monetary). Targeted strategies include reactivation campaigns for Lost Customers, loyalty programs for Potential Loyalists, and exclusive services for Loyal Customers. This data-driven approach improves marketing effectiveness, customer loyalty, and revenue through personalized strategies.

KeyWords: Customer Segmentation; RFM; K-Means; Marketing Strategy; Mitra 10 Cirebon

This is an Open Access article - copyright on authors, distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International License (CC BY SA) (<http://creativecommons.org/licenses/by-sa/4.0/>)

How to Cite: A. A. Rahma *et al.*, "Optimalisasi Strategi Pemasaran melalui Segmentasi Pelanggan dengan Analisis RFM dan Algoritma K-Means untuk Bisnis Ritel", *JIKO (JURNAL INFORMATIKA DAN KOMPUTER)*, Volume: 9, No.2, Pages 338–351, Juni 2025, doi: 10.26798/jiko.v9i2.1737.

1. Pendahuluan

Perkembangan pesat dalam bidang teknologi informasi telah mengubah berbagai aspek kehidupan, termasuk bisnis ritel. Dalam era persaingan yang semakin ketat, pemahaman perilaku pelanggan menjadi kunci dalam menyusun strategi pemasaran yang efektif. Salah satu pendekatan yang banyak digunakan dalam segmentasi pelanggan adalah analisis RFM (Recency, Frequency, Monetary) yang dikombinasikan dengan algoritma K-Means. Penelitian sebelumnya telah membuktikan efektivitas metode ini dalam mengelompokkan pelanggan berdasarkan karakteristik pembelian mereka.

Segmentasi menggunakan RFM dan K-Means dapat membagi pelanggan ke dalam dua segmen utama, yaitu pelanggan dengan tingkat pembelian rendah dan tinggi. Namun, penelitian tersebut masih memiliki keterbatasan dalam mempertimbangkan faktor eksternal yang memengaruhi perilaku pelanggan, seperti tren pasar dan strategi pemasaran yang diterapkan sebelumnya [1]. Meskipun hasil ini memberikan wawasan tentang pengelompokan pelanggan, pendekatan yang digunakan belum secara spesifik menyesuaikan strategi pemasaran berdasarkan segmentasi yang dihasilkan.

Perbandingan dengan penelitian ini menunjukkan bahwa penelitian saat ini memberikan cakupan lebih luas dengan mempertimbangkan variabel tambahan dalam strategi pemasaran, sehingga lebih aplikatif dalam bisnis ritel fisik [2]. Hasil penelitian ini menunjukkan adanya tiga kluster pelanggan, yaitu kluster terbaik, sedang, dan terburuk, yang masing-masing memiliki karakteristik berbeda terkait dengan nilai R, F, dan M. Segmentasi ini digunakan untuk menyusun strategi pemasaran yang lebih efektif guna meningkatkan loyalitas pelanggan [3].

Dalam meningkatkan penjualan di tengah persaingan bisnis yang ketat, penelitian ini bertujuan untuk menemukan strategi pemasaran yang efektif berdasarkan analisis data penjualan serta membahas penerapan teknik pengembangan data, terutama berfokus pada algoritma pengelompokan K-Means, untuk menganalisis data penjualan dan mengklasifikasikan produk berdasarkan kinerjanya, menekankan pentingnya mengidentifikasi produk mana yang terjual dengan baik dan mana yang tidak, karena ini dapat mengarah pada manajemen inventaris dan strategi pemasaran yang lebih baik [4].

Tujuan utama penelitian ini adalah untuk mengelompokkan pelanggan ke dalam segmen yang berbeda guna meningkatkan strategi pemasaran berbasis data [5]. Penelitian ini berupaya mengisi gap dari studi sebelumnya dengan menawarkan pendekatan segmentasi pelanggan yang lebih kontekstual untuk bisnis ritel fisik, khususnya Toko Mitra 10 Cirebon. Berbeda dengan penelitian sebelumnya yang lebih banyak berfokus pada e-commerce atau hanya menggunakan atribut RFM, penelitian ini mengintegrasikan strategi pemasaran berbasis data untuk setiap segmen pelanggan yang dihasilkan. Keunggulan utama dari pendekatan ini adalah validasi segmentasi menggunakan indeks Silhouette guna memastikan kualitas kluster, serta penerapan strategi pemasaran yang lebih aplikatif. Dengan kombinasi RFM dan K-Means, penelitian ini mampu mengelompokkan pelanggan secara lebih akurat. RFM dipilih karena kemampuannya dalam mengidentifikasi perilaku pelanggan berdasarkan tiga parameter utama, sedangkan K-Means memungkinkan pembagian pelanggan ke dalam kelompok yang lebih homogen. Untuk menentukan jumlah kluster optimal, digunakan metode elbow agar segmentasi lebih objektif. Selain menyajikan hasil segmentasi, penelitian ini juga menghubungkan implikasi nyata dalam strategi pemasaran. Misalnya, pelanggan loyal dapat diberikan program eksklusif untuk meningkatkan retensi, sementara pelanggan yang berpotensi loyal dapat diberikan insentif berbasis pola belanja mereka. Kampanye reaktivasi juga dapat diterapkan untuk menarik kembali pelanggan yang tidak aktif. Dengan pendekatan ini, penelitian ini tidak hanya menghasilkan segmentasi pelanggan yang lebih akurat tetapi juga memberikan strategi pemasaran berbasis data yang langsung dapat diterapkan dalam bisnis ritel fisik. Hal ini menunjukkan bahwa pendekatan berbasis data dapat meningkatkan efektivitas pemasaran serta membantu alokasi sumber daya secara lebih optimal.

2. Metode

Studi ini menggunakan model RFM untuk mengevaluasi perilaku pembelian pelanggan ditinjau dari tiga dimensi utama kebaruan (seberapa baru pelanggan melakukan pembelian), frekuensi (seberapa sering pelanggan melakukan pembelian), dan nilai moneter (berapa banyak uang yang dihabiskan pelanggan) Python digu-

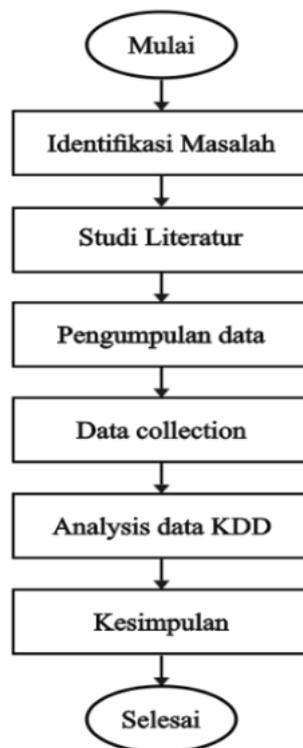
nakan untuk mengimplementasikan model RFM karena efisiensinya dalam penanganan dan analisis data. Studi ini menyoroti penggunaan pustaka Python, seperti Pandas untuk manipulasi data dan Matplotlib untuk visualisasi data, yang merampingkan proses analisis dan meningkatkan interpretabilitas hasil [6].

Langkah awal melibatkan pemilihan data berdasarkan kriteria RFM, yang berfokus pada perilaku pelanggan terkait dengan pola pembelian mereka. Data kemudian diubah menjadi nilai RFM, yang penting untuk analisis selanjutnya. Setelah pengelompokan, kualitas cluster dinilai menggunakan indeks siluet (SI), yang mengukur seberapa mirip suatu objek dengan klasternya sendiri dibandingkan dengan cluster lain. Langkah ini sangat penting untuk memvalidasi efektivitas proses pengelompokan. Nilai RFM rata-rata untuk setiap cluster dihitung untuk menentukan simbol peringkat setiap cluster. Peringkat ini membantu dalam memahami karakteristik segmen pelanggan yang berbeda [7].

K-Means adalah algoritma clustering yang digunakan untuk membagi data ke dalam sejumlah kluster berdasarkan kedekatan jarak kecentroid, yaitu titik pusat kluster. Proses ini dimulai dengan menetapkan jumlah kluster yang diinginkan oleh pengguna. Selanjutnya, centroid awal dipilih secara acak dari data yang ada. Data kemudian dikelompokkan berdasarkan jarak terdekat ke centroid menggunakan rumus jarak Euclidean. Setelah pengelompokan, posisi centroid dihitung ulang berdasarkan rata-rata dari semua titik dalam kluster. Proses ini diulang hingga tidak ada perubahan pada pengelompokan, menandakan bahwa kluster telah stabil. Dengan demikian, K-Means efektif dalam mengelompokkan data yang memiliki karakteristik serupa dan merupakan metode yang umum digunakan dalam analisis data dan segmentasi pasar [8].

2.1 Alur Metode Penelitian

Fokus utama penelitian ini adalah mengoptimalkan strategi pemasaran di Toko Mitra 10 Cirebon melalui pendekatan segmentasi pelanggan menggunakan analisis RFM serta algoritma K-Means. Oleh karena itu, data yang dipakai dalam penelitian ini harus relevan, memiliki kualitas yang baik, dan mudah diakses. Berikut adalah tahapan-tahapan metode yang diterapkan dalam penelitian ini:



Gambar 1. Alur Metode Penelitian

Berikut ini adalah deskripsi untuk setiap tahap dalam teknik analisis data pada Gambar 1:

1. Identifikasi Masalah

Tahapan awal penelitian dimulai dengan identifikasi permasalahan utama yang ingin dipecahkan. Dalam konteks penelitian ini, masalah yang diidentifikasi adalah penerapan model Algoritma K-means pada segmentasi pelanggan Mitra10 Cirebon menggunakan analisis RFM.

2. Studi Literatur

Sebelum memasuki tahap analisis, dilakukan kajian literatur untuk memahami pendekatan yang telah digunakan dalam penelitian sebelumnya. Beberapa literatur kunci yang digunakan meliputi:

- a. Klaster data pelanggan potensial menggunakan pendekatan K-Means dengan dukungan metode Elbow untuk menentukan jumlah kluster yang tepat [9].
- b. Membahas pentingnya memahami karakteristik pelanggan untuk mengembangkan strategi pemasaran yang efektif, terutama mengingat penurunan pelanggan [10].

3. Data collection

Data yang digunakan dalam penelitian ini berasal dari system Sales Detail Report Mitra10 Cirebon, Rentang waktu ini dapat mencakup periode historis dari kurun waktu mulai dari 1 Juli - 31 Oktober 2024. Data ini diunduh dalam format CSV atau Excel dan memberikan gambaran menyeluruh tentang Segmentasi pelanggan berdasarkan tiga poin penting RFM.

4. Analysis data metode KDD

Tahap Analysis data metode KDD dan pemodelan dilakukan menggunakan perangkat lunak Python yang memfasilitasi penerapan Algoritma K-Means dalam memprediksi serta membantu menganalisis karakteristik pelanggan Mitra10 Cirebon. Adapun langkahnya Data preprocessing, Data transformation, Data mining, dan Evaluasi.

5. Pembahasan hasil data

tahap di mana hasil dari analisis data ditinjau dan diinterpretasikan untuk memberikan makna dan wawasan lebih mendalam. Pada tahap ini, tidak hanya menyajikan angka atau grafik, tetapi juga menjelaskan apa arti dari hasil tersebut dan bagaimana relevansinya terhadap tujuan atau pertanyaan penelitian.

6. Kesimpulan dan Saran

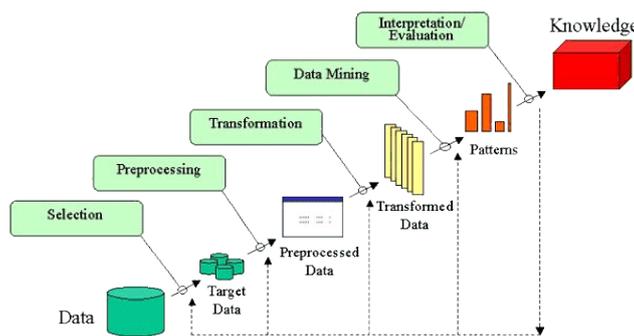
Kesimpulan dari penelitian ini akan mencakup tentang Analisis RFM (*Recency, Frequency, Monetary*) yang dilakukan pada data transaksi Toko Mitra10 Cirebon berhasil mengidentifikasi karakteristik unik dari setiap segmen pelanggan. Hasil penelitian ini menunjukkan bahwa pelanggan Toko Mitra10 Cirebon dapat dikelompokkan menjadi beberapa segmen berdasarkan pola pembelian mereka. Dengan memahami karakteristik setiap segmen, toko dapat menyusun strategi pemasaran yang lebih *targeted* dan efektif, sehingga meningkatkan loyalitas pelanggan dan penjualan.

2.2 Teknik Pengumpulan Data

Pengumpulan data dilakukan dengan Sistem Manajemen Basis Data (DBMS) Memanfaatkan sistem yang ada di Toko Mitra 10 untuk mengekstrak data transaksi historis.

2.2.1 Analisis Data

Penelitian ini menggunakan teknik analisis data kuantitatif kerangka kerja KDD (*Knowledge Discovery in Databases*) seperti terlihat pada Gambar 2 dengan penerapan Algoritma K-Means untuk memprediksi Segmentasi pelanggan berdasarkan analisis RFM pada Toko Mitra10 Cirebon. Berikut adalah teknik analisis data yang diterapkan: Setelah mengumpulkan data, langkah pertama dalam analisis adalah menghitung



Gambar 2. Knowledge Discovery in Database

nilai RFM untuk setiap pelanggan. Proses ini melibatkan penghitungan tiga komponen utama:

1. *Recency* (R) Mengukur waktu sejak pelanggan terakhir melakukan pembelian. Nilai yang lebih kecil menunjukkan bahwa pelanggan baru-baru ini berbelanja.
2. *Frequency* (F) Menghitung jumlah total pembelian yang dilakukan oleh pelanggan dalam periode waktu tertentu. Semakin tinggi nilainya, semakin sering pelanggan melakukan pembelian.
3. *Monetary* (M) Menentukan total pengeluaran pelanggan selama periode yang dianalisis. Nilai yang lebih tinggi menunjukkan pelanggan yang lebih berharga bagi bisnis.

3. Hasil

Penelitian ini menggunakan teknik analisis data kuantitatif kerangka kerja KDD (*Knowledge Discovery in Databases*) dengan penerapan Algoritma K-Means untuk memprediksi Segmentasi pelanggan berdasarkan analisis RFM pada Toko Mitra10 Cirebon.

3.1 Tahapan *Knowledge Discovery in Database* (KDD)

3.1.1 Data Selection

Pada tahap ini, data yang akan digunakan untuk proses KDD (*Knowledge Discovery in Databases*) adalah data yang dipilih dengan cermat dan relevan. Data pelanggan yang akan digunakan untuk analisis RFM ini merupakan data transaksi historis pelanggan Toko Mitra 10 Cirebon, data ini diambil dari system Sales Detail Report pada Toko Mitra10 Cirebon dengan kurun waktu mulai dari 1 Juli sampai dengan 31 Oktober 2024. Dataset ini terdiri dari 70251 data dengan jumlah atribut 9 buah yakni *StoreName*, *Receipt No*, *CustomerID*, *TransactionDate*, *Time*, *ItemNo*, *Description*, *Quantity*, *Sales*. Tampak dalam Table 1 berikut ini:

3.1.2 Data Preprocessing

Tujuan *preprocessing* data adalah memastikan data yang digunakan dalam analisis bersih, konsisten, dan siap untuk diolah lebih lanjut, sehingga dapat menghasilkan hasil analisis yang akurat dan valid. Dalam penelitian ini, data transaksi pelanggan dari Toko Mitra10 Cirebon diolah melalui beberapa langkah *preprocessing* yang sistematis untuk mempersiapkannya dalam proses *clustering* menggunakan *algoritma K-Means*. *Preprocessing* data mencakup tahapan yang dilakukan dalam proses preprocessing pada penelitian ini dengan penanganan *outlier* Menggunakan metode IQR untuk mengidentifikasi outlier pada variabel *Recency*, *Frequency*, dan *Monetary*. *Outlier* kemudian di-cap (bukan dihapus) agar analisis tetap representatif. *Scaling* data dinormalisasi menggunakan *MinMaxScaler* untuk memastikan nilai antar fitur berada dalam rentang yang seragam (0-1).

1. Transformasi Data melakukan transformasi data penggabungan 2 Dataset, *df_recency* dan *frequency_df*, berdasarkan kolom kunci yang sama, yaitu *CustomerID*. Operasi ini penting untuk menyatukan informasi *recency* dan *frequency* dari pelanggan ke dalam satu *table* (*rf_df*) agar data dapat diolah secara menyeluruh pada tahap berikutnya.

```
# Assuming df_recency has 'CustomerID' instead of 'CustomerID'
rf_df = df_recency.merge(frequency_df, left_on='CustomerID',
right_on='CustomerID')
```

2. Konversi Tipe Data menelaraskan tipe data agar kompetibel untuk menggabungkan atau operasi lainnya, kemudian menonversi data no-numerik (seperti *string* atau tanggal) menjadi data numerik yang dapat di proses lebih lanjut, misalnya mengubah tanggal pembelian terakhir menjadi *recency*.

```
# Alternatively, rename the column in df_recency before merging:
df_recency = df_recency.rename(columns={'CustomerID': 'CustomerID'})
rf_df = df_recency.merge(frequency_df, on='CustomerID')
rfm_df = rf_df.merge(monetary_df, on='CustomerID').drop(
columns='LastPurchaseDate')
rfm_df.head()
```

Tabel 1. Data Transaksi Pelanggan Mitra 10

No	Store Name	Receipt No.	Customer ID	Date	Time	Item No	Description	Qty	Sales
1	Mitra10 SGJ	SG 00463101	2010051856	01/07/24	9:19:24	1000017004	C/JAYA ABADI PVC PLAIN 40 X 60	2	38292.80
2	Mitra10 SGJ	SG 00463101	2010051856	01/07/24	9:19:24	1000025037	FANTASY ARTFC GRASS CROWN 25MM SM-CROWN02 ROLL=3M	1	535594.59
3	Mitra10 SGJ	SG 00463144	2010341605	01/07/24	12:09:38	400015138	DEXTONE ACRILIC SEALANT TBG WHITE 300ML	4	22702.70
4	Mitra10 SGJ	SG 00463144	2010341605	01/07/24	12:09:38	900008951	ACE OLDFIELDS SUN BRUSH 1 INC	2	3153.16
...
70248	Mitra10 SGJ	SG 00487987	2010753559	31/10/24	13:49:17	400017458	JOTUN JOTA-SHIELD INFINITY BASE-A 18L	1	2855812.61
70249	Mitra10 SGJ	SG 00487987	2010753559	31/10/24	13:49:17	100031439	SINCERE MBZM121 GLAZED TILE 60X120CM/E	83	232216.22
70250	Mitra10 SGJ	SG 00487987	2010753559	31/10/24	13:49:17	100033523	J/NIRO 60X120/1.44 MAGNIFICO GCO02 FANTASY WHITE	18	352342.34
70251	Mitra10 SGJ	SG 00487989	2010889772	31/10/24	14:07:37	500007514	BARDOLINE RECTANGULAR 3 TONE BROWN	2	619819.82

Setelah data RFM diperoleh, dari tahap hasil Python diatas dihasilkan data RFM yang terlihat seperti Table 2 sebagai berikut:

- Pelanggan dengan *Recency* rendah (0-10 hari) dan *Frequency* tinggi (≥ 10 transaksi) berpotensi sebagai pelanggan loyal dan perlu diberi program eksklusif atau *reward* untuk mempertahankan mereka.
- Pelanggan dengan *Frequency* rendah (1-3 transaksi) dan *Monetary* kecil bisa jadi pelanggan baru atau tidak terlalu loyal, bisa diberikan promosi atau diskon khusus untuk meningkatkan keterlibatan mereka.
- Pelanggan dengan *Recency* tinggi (≥ 60 hari) dan *Frequency* rendah termasuk dalam kategori *Lost Customers* yang perlu diberikan kampanye pemasaran kembali agar tertarik berbelanja lagi.

- d. Pelanggan dengan *Monetary* tinggi tetapi *Recency* tinggi Pelanggan bernilai tinggi yang mulai jarang berbelanja, perlu pendekatan khusus seperti penawaran eksklusif atau komunikasi personal. Dengan menggunakan *K-Means Clustering*, pelanggan dalam tabel ini bisa dikelompokkan ke dalam segmen yang berbeda untuk menerapkan strategi pemasaran yang lebih efektif.

Tabel 2. Hasil RFM

No	CustomerID	Recency	Frequency	Monetary
1	1010888090000	0	67	23170885,72
2	2010000105655	5	9	1652905,41
3	2010000364042	67	2	37747,75
4	2010000906673	12	11	4765533,35
...
...
10121	2019983115039	45	1	188684,69
10122	2019988845094	68	4	1143496,39
10123	2019997516182	36	6	526090,99
10124	2019998634605	102	7	3000008,12

3.1.3 Data Transformation

Tujuan utama dari transformasi ini adalah untuk mengubah data mentah menjadi skala yang sama sehingga dapat dibandingkan dengan lebih mudah. Selain itu, dengan membagi pelanggan ke dalam kuartil. Data transaksi pelanggan diolah menjadi skor RFM dengan menghitung dan memberikan bobot pada masing-masing variable. Kemudian RFM dihitung menggunakan perhitungan *rank-normalized* dan bobot masing-masing variabel (*Recency*=0.15, *Frequency*=0.28, *Monetary*=0.57) sesuai metode yang ditentukan. Adapun untuk penjelasan pada Table 3 sebagai berikut:

- CustomerID: berfungsi untuk Identifikasi unik untuk setiap pelanggan.
- Recency, Frequency, Monetary*: merupakan nilai asli dari masing-masing metrik RFM.
- R_quartile, F_quartile, M_quartile*: Kuartil dari masing-masing metrik. Kuartil membagi data menjadi empat bagian yang sama. Misalnya, pelanggan dengan *R_quartile* = 1 memiliki nilai *Recency* yang paling rendah (transaksi terakhir paling baru).
- RFM_Score*: Skor gabungan dari ketiga metrik RFM. Skor ini digunakan untuk mengurutkan pelanggan berdasarkan nilai keseluruhannya. Dari tabel 2 akan dimasukkan ke dalam Table 3 Hasil *RFM Score* yang sudah dihasilkan dari codingan pada *Python table RFM Score* yang terlihat sebagai berikut:

- *RFM Score* 444 atau 443 *Loyal Customers* strategi program eksklusif, VIP reward, diskon spesial.
- *RFM Score* 332 atau 311 *Potential Loyalists* strategi berikan insentif agar tetap aktif, promosi khusus, bonus belanja.
- *RFM Score* 211 atau 133 *Lost Customers* strategi kampanye reaktivasi, penawaran spesial untuk menarik kembali.

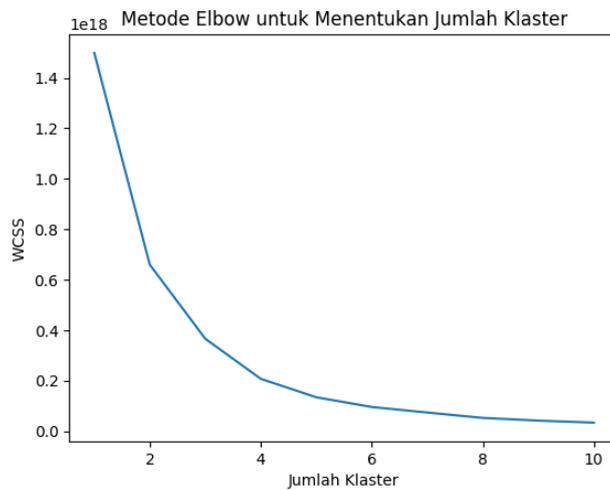
3.1.4 Data Mining

Elbow Method digunakan untuk menentukan jumlah *cluster* yang optimal. Pada tahap ini, dilakukan pengelompokan pelanggan (*clustering*) berdasarkan skor RFM menggunakan *algoritma K-Means* adapun Langkah-langkahnya sebagai berikut:

- Menentukan jumlah *cluster* optimal menggunakan *Elbow Method*.
- Menggunakan *algoritma K-Means* untuk mengelompokkan data ke dalam *cluster* sesuai hasil *Elbow Method*.
- Menambahkan kolom *Cluster* ke dalam *DataFrame* untuk mengidentifikasi segmen pelanggan. ini membentuk visualisasi Gambar 3 *Method Elbow*

Tabel 3. Hasil RFM Score

No	CustomerID	Recency	Frequency	Monetary	R_quartile	F_quartile	M_quartile	RFM_Score
1	1010888090000	0	67	23170885,72	4	4	4	444
2	2010000105655	5	9	1652905,41	4	4	3	443
3	2010000364042	67	2	37747,75	2	1	1	211
4	2010000906673	12	11	4765533,35	4	4	4	444
...
...
10121	2019983115039	45	1	188684,69	3	1	1	311
10122	2019988845094	68	4	1143496,39	2	2	3	223
10123	2019997516182	36	6	526090,99	3	3	2	332
10124	2019998634605	102	7	3000008,12	1	3	3	133



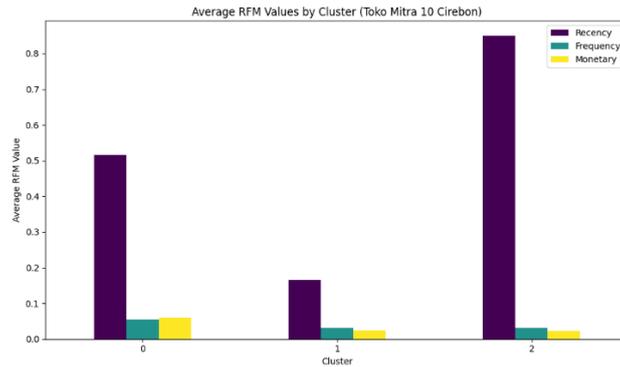
Gambar 3. Metode Elbow

- a. Sumbu X (Jumlah Kluster) menunjukkan jumlah kluster yang dicoba, dari 1 hingga 10.
- b. Sumbu Y (WCSS - Within-Cluster Sum of Squares) mengukur seberapa baik data dikelompokkan dalam setiap jumlah kluster. Semakin kecil WCSS, semakin baik kluster terbentuk.
- c. Kurva yang menurun WCSS berkurang saat jumlah kluster bertambah. Ini karena lebih banyak kluster berarti jarak antar titik dalam satu kluster semakin kecil.

3.1.5 Evaluasi dan Interpretasi

Mengevaluasi kualitas pola yang ditemukan dan menginterpretasikan hasilnya. Visualisasi menggunakan diagram batang untuk memvisualisasikan rata-rata nilai RFM untuk setiap cluster seperti terlihat pada Gambar 4. Pada bagian ini menggunakan visualisasi lain seperti *scatter plot* atau *heatmap*. Hasil *clustering* divisualisasikan dalam bentuk grafik untuk memahami pola segmen pelanggan, seperti hubungan antara *Frequency* dan *Monetary*. Cluster diinterpretasikan untuk memberikan wawasan strategis, misalnya:

- a. Cluster dengan nilai *Frequency* dan *Monetary* tinggi dianggap sebagai pelanggan loyal.
 - b. Cluster dengan nilai *Recency* rendah dan *Frequency* rendah dapat diidentifikasi sebagai pelanggan pasif atau hampir hilang (*churn*). Visualisasi dilakukan menggunakan *matplotlib* dan *seaborn* untuk memberikan interpretasi yang lebih mudah dipahami.
1. Klaster 0:
 - a. *Recency* tinggi, menunjukkan pelanggan ini sudah lama tidak bertransaksi.
 - b. *Frequency* dan *Monetary* rendah, artinya mereka jarang berbelanja dan nilai transaksi mereka kecil.
 - c. Kemungkinan kategori: *Lost Customers* (Pelanggan yang hilang).
 2. Klaster 1:



Gambar 4. RFM Cluster

- a. *Recency* sedang, berarti pelanggan ini masih cukup aktif dibandingkan Klaster 0.
 - b. *Frequency* dan *Monetary* lebih tinggi daripada Klaster 0.
 - c. Kemungkinan kategori: Potential Loyalists (Pelanggan potensial yang dapat ditingkatkan loyalitasnya).
3. Klaster 2:
- a. *Recency* sangat tinggi, menunjukkan pelanggan dalam klaster ini belum berbelanja dalam waktu yang sangat lama.
 - b. *Frequency* dan *Monetary* rendah, artinya pelanggan ini tidak sering berbelanja dan nilai transaksi mereka kecil.
 - c. Kemungkinan kategori: *Lost Customers* atau *Dormant Customers* (Pelanggan yang tidak aktif).

4. Pembahasan

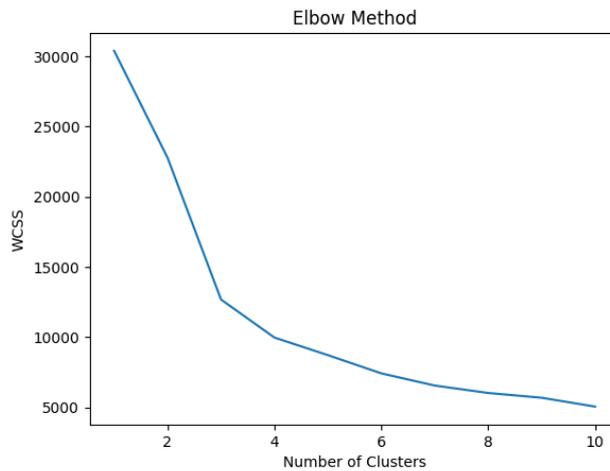
4.1 Analisis RFM dan Penerapan K-Means

Analisis RFM (*Recency*, *Frequency*, *Monetary*) adalah metode yang efektif untuk mengelompokkan perilaku pelanggan, berdasarkan tiga variabel utama *Recency* (waktu sejak pembelian terakhir), *Frequency* (jumlah transaksi dalam periode tertentu), dan *Monetary* (total uang yang dibelanjakan). Penelitian ini sejalan dengan penelitian terdahulu hasil penelitian menunjukkan bahwa metode *K-Means clustering* berhasil mengidentifikasi segmen-segmen pelanggan dengan akurasi tinggi. Pelanggan dikelompokkan ke dalam kategori berdasarkan nilai RFM mereka, yang memungkinkan perusahaan untuk menargetkan strategi pemasaran dengan lebih efektif [11].

Penelitian ini mendukung temuan sebelumnya bahwa RFM tidak hanya berfungsi sebagai alat segmentasi, tetapi juga sebagai sarana untuk memahami dinamika perilaku pelanggan, seperti pada Toko Mitra 10 Cirebon. Metode RFM menghitung nilai *Recency* (jarak waktu transaksi terakhir), *Frequency* (frekuensi transaksi), dan *Monetary* (total pembelanjaan) untuk setiap pelanggan. Data diolah menggunakan Python dengan normalisasi sebelum diterapkan algoritma K-Means untuk segmentasi. Jumlah klaster optimal ditentukan dengan metode *elbow*, yang mengidentifikasi titik optimal berdasarkan perubahan signifikan nilai *Within-Cluster Sum of Squares* (WCSS). Pendekatan ini sejalan dengan penelitian sebelumnya yang menggunakan metode *elbow* untuk menentukan jumlah klaster optimal dalam berbagai konteks, menunjukkan relevansi lintas sektor dengan adaptasi terhadap karakteristik data spesifik [12].

Penelitian ini menggunakan RFM dan K-Means untuk mengelompokkan pelanggan pada toko Mitra10 Cirebon. Data diproses dengan Python, dan normalisasi dilakukan sebelum implementasi *algoritma K-Means*. Penentuan jumlah *cluster* optimal dilakukan melalui metode *elbow*, menunjukkan adaptasi terhadap konteks bisnis [13]. Selanjutnya Perhitungan nilai *Recency*, *Frequency*, dan *Monetary* (RFM) untuk pelanggan Mitra10 Cirebon dilakukan berdasarkan data transaksi yang terdapat dalam *file Mitra10.csv*. Proses pengolahan data ini memanfaatkan bahasa pemrograman Python. Metode *elbow* digunakan untuk menentukan jumlah klaster optimal dalam *algoritma K-Means*. Hasil visualisasi, yang ditampilkan dalam Gambar 5 memperlihatkan hubungan antara jumlah klaster *k* dan nilai *within-cluster sum of squares* (WCSS). Titik optimal (dikenal sebagai *elbow*) diidentifikasi sebagai momen ketika penurunan WCSS melambat secara

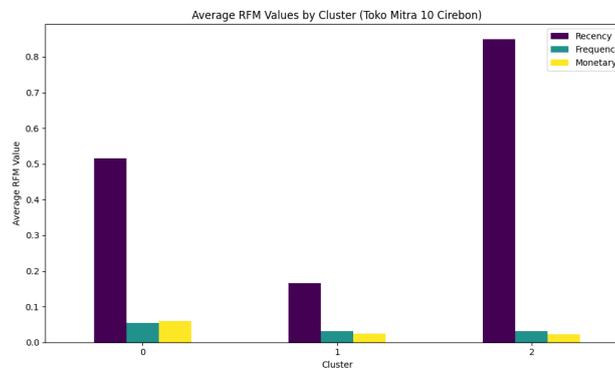
signifikan, menunjukkan jumlah kluster terbaik untuk segmentasi. Visualisasi ini memastikan pembagian kluster pelanggan yang akurat dan relevan sesuai karakteristik data. Setelah nilai RFM dihitung, skor setiap variabel dinormalisasi dan diberi bobot (*Recency* 15%, *Frequency* 28%, *Monetary* 57%) untuk menghasilkan RFM Score.



Gambar 5. Elbow Method

4.2 Karakteristik Setiap Cluster

Dari hasil *clustering*, pelanggan terbagi ke dalam beberapa kelompok dengan karakteristik berbeda ditunjukkan dalam Grafik rata-rata nilai *Recency*, *Frequency*, dan *Monetary* (RFM) untuk setiap kluster hasil segmentasi pelanggan Toko Mitra10 Cirebon menggunakan *algoritma K-Means*. Grafik ini menunjukkan perbedaan karakteristik antara kluster, di mana nilai *Recency*, *Frequency*, dan *Monetary* memberikan gambaran perilaku pelanggan yang berbeda-beda, sehingga dapat digunakan sebagai dasar untuk strategi pemasaran yang lebih personal dan efektif memberikan wawasan tentang bagaimana pelanggan dapat dikelompokkan secara lebih efisien berdasarkan nilai RFM mereka, sehingga strategi pemasaran dapat lebih terarah dan personal. Menerapkan *algoritma K-Means* untuk segmentasi pelanggan dalam *e-commerce*. Penelitian ini menemukan bahwa segmen dengan loyalitas tertinggi memiliki potensi terbesar untuk mempertahankan kontribusi pendapatan [14]. Penelitian ini memberikan wawasan mendalam tentang pola pembelian pelanggan untuk mendukung strategi pemasaran yang lebih terarah. Metode ini juga membantu meminimalkan kesalahan dalam menentukan jumlah *cluster*, yang sangat penting untuk meningkatkan akurasi dan relevansi dalam penentuan segmentasi pelanggan pada Toko Mitra10 Cirebon. Grafik tersebut disajikan dalam bentuk Gambar 6. Berdasarkan hasil analisis grafik tersebut diperoleh data yang telah



Gambar 6. Grafik Cluster RFM Mitra10.

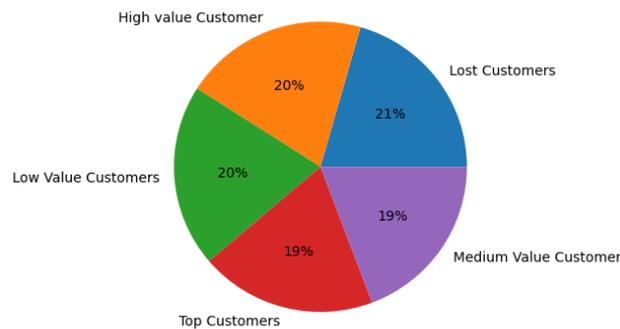
diproses, terbentuk beberapa *cluster* yang masing-masing memiliki karakteristik tertentu.

1. Pelanggan Loyal (*Recency* rendah, *Frequency* tinggi, *Monetary* tinggi):
 - a. Interpretasi: Pelanggan sering berbelanja dan memberikan kontribusi signifikan.

- b. Strategi: Penawaran eksklusif, program loyalitas, dan diskon khusus.
2. Pelanggan Pasif (*Recency* tinggi, *Frequency* rendah, *Monetary* rendah):
 - a. Interpretasi: Pelanggan jarang berbelanja dan kontribusi rendah.
 - b. Strategi: Strategi retensi seperti pengingat promosi atau penawaran menarik.
3. Pelanggan Potensial (*Recency* rendah, *Frequency* sedang, *Monetary* sedang):
 - a. Interpretasi: Berpotensi menjadi pelanggan aktif jika diberikan perhatian yang tepat.
 - b. Strategi: Promosi berbasis pengulangan pembelian atau komunikasi lebih intensif.
4. Pelanggan Pembelanjaan Tinggi (*Recency* sedang, *Frequency* tinggi, *Monetary* tinggi)
 - a. Interpretasi: Pelanggan yang memberikan kontribusi besar, terutama pada produk premium.
 - b. Strategi: Fokus pada penawaran eksklusif, *upselling*, dan *cross-selling*.

4.3 Strategi Pemasaran Berdasarkan Segmentasi

Perlunya segmentasi pelanggan yang efektif untuk meningkatkan strategi pemasaran. Serta Memahami karakteristik dan perilaku kelompok pelanggan yang berbeda, meningkatkan strategi pemasaran dengan memahami keragaman pelanggan melalui segmentasi. Hasilnya menunjukkan identifikasi empat kelompok pelanggan yang berbeda berdasarkan analisis, yang dapat menginformasikan strategi pemasaran yang ditargetkan. Studi ini menekankan pentingnya segmentasi pelanggan untuk menyediakan layanan yang disesuaikan dan meningkatkan strategi retensi dan akuisisi pelanggan [15]. Analisis ini menunjukkan bahwa kombinasi RFM dan K-Means efektif untuk memahami perilaku pelanggan, memungkinkan Toko Mitra10 menyusun strategi yang lebih tepat sasaran. Grafik ini menjadi landasan penting dalam menjawab rumusan masalah penelitian, yakni bagaimana menyusun strategi pemasaran yang tidak hanya optimal tetapi juga mampu memenuhi kebutuhan pelanggan secara spesifik berdasarkan karakteristik segmennya. Adapun untuk grafiknya disajikan dalam bentuk Gambar 7. Diagram *Pie* segmentasi pelanggan sebagai berikut.



Gambar 7. Diagram *Pie* segmentasi pelanggan

1. *Lost Customers* (21%) - Warna Biru
 - a. Pelanggan yang sudah lama tidak bertransaksi (*Recency* tinggi). *Frequency* dan *Monetary* mereka rendah.
 - b. Strategi: Perlu program reaktivasi pelanggan, seperti email promosi atau diskon khusus untuk menarik mereka kembali.
2. *Medium Value Customers* (19%) - Warna Ungu
 - a. Pelanggan dengan frekuensi dan nilai transaksi sedang. Mereka mungkin masih aktif, tetapi tidak terlalu sering berbelanja.
 - b. Strategi: Bisa diberikan program loyalty rewards atau diskon berbasis jumlah transaksi untuk mendorong pembelian lebih sering.
3. *Top Customers* (19%) - Warna Merah
 - a. Pelanggan yang memiliki frekuensi tinggi dan kontribusi nilai transaksi besar. Mereka merupakan pelanggan setia dengan *engagement* tinggi.
 - b. Strategi: Bisa diberikan layanan eksklusif, diskon VIP, atau personalisasi rekomendasi produk agar tetap loyal.

4. *Low Value Customers* (20%) - Warna Hijau
 - a. Pelanggan dengan frekuensi dan *monetary* rendah. Bisa jadi mereka hanya sesekali berbelanja tanpa loyalitas tinggi.
 - b. Strategi: Perlu insentif seperti *bundling* produk murah atau program *referral* untuk meningkatkan keterlibatan mereka.
5. *High Value Customers* (20%) - Warna Oranye
 - a. Pelanggan dengan *monetary* tinggi, tetapi mungkin frekuensi belanjanya tidak terlalu sering. Bisa jadi pelanggan yang membeli dalam jumlah besar tetapi hanya sesekali.
 - b. Strategi: Bisa diberikan program *membership* premium atau eksklusif untuk meningkatkan loyalitas mereka.

Penyusunan Strategi Pemasaran Berdasarkan Segmentasi Pelanggan Berdasarkan hasil *clustering* dengan *algoritma K-Means* yang diterapkan pada data RFM, ditemukan berbagai karakteristik pelanggan yang dapat digunakan untuk menyusun strategi pemasaran yang lebih personal dan optimal menentukan jumlah cluster yang optimal untuk analisis yang akurat [16]. pentingnya memahami kebutuhan dan preferensi pelanggan untuk mengoptimalkan pengiriman layanan dan mengatasi masalah yang ada di dalam perusahaan [17]. dengan tujuan untuk meningkatkan loyalitas pelanggan dan profitabilitas perusahaan. Metode yang digunakan dalam penelitian ini adalah model RFM (*Recency, Frequency, Monetary*) dan *algoritma K-Means* untuk melakukan segmentasi pelanggan [18]. Metode K-Means efektif dalam mengelompokkan data berdasarkan tingkat kepuasan pelanggan [19]. Dengan demikian K-Means dapat dijadikan metode yang efektif untuk mencari pola pengelompokan data berdasarkan kriteria tertentu yang hasilnya dapat dianalisis lebih jauh sesuai kebutuhan data [20]. Dengan segmentasi ini, Toko Mitra10 dapat meningkatkan loyalitas pelanggan, retensi, dan pendapatan melalui pendekatan pemasaran yang personal dan strategis.

5. Simpulan

Penelitian ini berhasil menerapkan analisis RFM (*Recency, Frequency, Monetary*) dan *algoritma K-Means* untuk mengidentifikasi dan mengelompokkan perilaku pelanggan. Hasil utama dari penelitian ini mencakup: Keberhasilan Segmentasi Pelanggan Metode RFM dan *algoritma K-Means* berhasil mengelompokkan pelanggan ke dalam klaster yang homogen, mencerminkan pola pembelian yang berbeda-beda. Hal ini mendukung identifikasi kelompok pelanggan seperti pelanggan loyal, pasif, dan potensial. Karakteristik Setiap Klaster pelanggan yang terbentuk memberikan informasi penting untuk strategi pemasaran. Contohnya, pelanggan loyal memerlukan pendekatan personalisasi, sementara pelanggan pasif membutuhkan strategi retensi. Strategi Pemasaran yang Optimal Integrasi analisis RFM dan *algoritma K-Means* memungkinkan perancangan strategi pemasaran berbasis data, meningkatkan kepuasan pelanggan, loyalitas, dan pendapatan toko.

Pustaka

- [1] W. A. Silamantha and K. Hadiono, "Analisis rfm dan k-means clustering untuk segmentasi pelanggan pada pt. sanutama bumi arto," *KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer dan Manajemen)*, vol. 5, no. 3, pp. 1297–1305, 2024. doi: <https://doi.org/10.30645/kesatria.v5i3.448>.
- [2] S. Wahyuni, T. T. Wulansari, and F. Fahrullah, "Segmentasi pelanggan berdasarkan analisis recency, frequency, monetary menggunakan algoritma k-means pada cv. toedjoe sinar group," *Jurnal Rekayasa Teknologi Informasi*, vol. 7, no. 2, pp. 29–36, Dec 2023. doi: [10.30872/jurti.v7i2.8748](https://doi.org/10.30872/jurti.v7i2.8748).
- [3] I. Maskanah, A. Primajaya, and A. Rizal, "Segmentasi pelanggan toko purnama dengan algoritma k-means dan model rfm untuk perancangan strategi pemasaran," *INOVTEK Polbeng - Seri Informatika*, vol. 5, no. 2, p. 218, Nov 2020. doi: [10.35314/isi.v5i2.1443](https://doi.org/10.35314/isi.v5i2.1443).
- [4] A. Nugraha, O. Nurdiawan, and G. Dwilestari, "Penerapan data mining metode k-means clustering untuk analisa penjualan pada toko yana sport," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 6, no. 2, pp. 849–855, 2022. doi: [10.36040/jati.v6i2.5755](https://doi.org/10.36040/jati.v6i2.5755).
- [5] O. N. Akande, E. O. Asani, and B. T. Dutare, "Customer segmentation through rfm analysis and k-means clustering: Leveraging data-driven insights for effective marketing strategy,"

- Ceddi Journal of Information System and Technology*, vol. 3, no. 1, pp. 14–25, 2024. doi: <https://doi.org/10.56134/jst.v3i1.81>.
- [6] A. Hermawan, N. R. Jayanti, A. Saputra, A. P. Parta, M. A. A. Thahir, and Taufiqurrahman, “Leveraging the rfm model for customer segmentation in a software-as-a-service (saas) business using python,” *Maeswara: Jurnal Riset Ilmu Manajemen dan Kewirausahaan*, vol. 2, no. 5, pp. 77–89, 2024. doi: <https://doi.org/10.61132/maeswara.v2i5.1283>.
- [7] S. Monalisa, Y. Juniarti, E. Saputra, F. Muttakin, and T. K. Ahsyar, “Customer segmentation with rfm models and demographic variable using dbscan algorithm,” *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, vol. 21, no. 4, pp. 742–749, 2023. doi: [10.12928/TELKOMNIKA.v21i4.22759](https://doi.org/10.12928/TELKOMNIKA.v21i4.22759).
- [8] S. D. K. Wardani, A. S. Ariyanto, M. Umroh, and D. Rolliawati, “Perbandingan hasil metode clustering k-means, db scanner & hierarchical untuk analisa segmentasi pasar,” *JIKO (Jurnal Informatika dan Komputer)*, vol. 7, no. 2, p. 191, 2023. doi: [10.26798/jiko.v7i2.796](https://doi.org/10.26798/jiko.v7i2.796).
- [9] F. Marisa, A. R. Wardhani, W. Purnomowati, A. V. Vitianingsih, A. L. Maukar, and W. Puspitarini, “K-means with elbow method,” *JIKO (Jurnal Informatika dan Komputer)*, vol. 7, no. 2, pp. 307–312, 2023. doi: [10.26798/jiko.v7i2.911](https://doi.org/10.26798/jiko.v7i2.911).
- [10] M. A. Satriawan, R. Andreswari, and O. N. Pratiwi, “Segmentasi pelanggan telkomsel menggunakan metode clustering dengan rfm model dan algoritma k-means telkomsel customer segmentation using clustering method with rfm model and k-means algorithm,” in *e-proceeding of Engineering*, vol. 8, no. 2, 2021, pp. 2876–2883.
- [11] K. Tabianan, S. Velu, and V. Ravi, “K-means clustering approach for intelligent customer segmentation using customer purchase behavior data,” *MDIP*, vol. 14, no. 12, pp. 1–15, 2022. doi: <https://doi.org/10.3390/su14127243>.
- [12] M. D. Rivaldo, G. W. N. Wibowo, and H. Mulyo, “Implementasi algoritma k-means untuk klasterisasi data hasil tangkapan ikan di karimunjawa,” *Jurnal Minfo Polgan*, vol. 13, no. 1, pp. 1045–1056, Jul 2024. doi: [10.33395/jmp.v13i1.13928](https://doi.org/10.33395/jmp.v13i1.13928).
- [13] S. Budilaksono *et al.*, “Customer profiling for precision marketing using rfm method, k-means algorithm and decision tree,” *Sinkron*, vol. 6, no. 1, pp. 191–200, Oct 2021. doi: [10.33395/sinkron.v6i1.11225](https://doi.org/10.33395/sinkron.v6i1.11225).
- [14] P. I. Pangestu, T. I. Hermanto, and D. Irmayanti, “Analisis segmentasi pelanggan berbasis recency frequency monetary (rfm) menggunakan algoritma k-means,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 3, pp. 1486–1492, 2023. doi: <https://doi.org/10.36040/jati.v7i3.7171>.
- [15] B. T. Kristanti, A. Junaidi, and E. P. Mandyartha, “Implementasi k-means clustering dalam segmentasi pelanggan berdasarkan usia, pendapatan, dan model rfm (studi kasus: Lantikya store jombang),” *JITET*, vol. 12, no. 3, pp. 2099–2112, 2024. doi: <http://dx.doi.org/10.23960/jitet.v12i3.4677>.
- [16] R. Safira and C. A. Sugianto, “Optimalisasi algoritma k-means menggunakan metode elbow dalam pengelompokan data stunting,” *BRAHMANA: Jurnal Penerapan Kecerdasan Buatan*, vol. 5, no. 2, pp. 257–264, 2024. doi: <https://doi.org/10.30645/brahmana.v5i2.396>.
- [17] W. Romadhona, B. I. Nugroho, and A. A. Murtopo, “Implementasi data mining pemilihan pelanggan potensial menggunakan algoritma k-means,” *Jurnal Minfo Polgan*, vol. 11, no. 2, pp. 100–104, Sep 2022. doi: [10.33395/jmp.v11i2.11797](https://doi.org/10.33395/jmp.v11i2.11797).
- [18] K. Z. Wijaya, A. Djunaidi, and F. Mahananto, “Segmentasi pelanggan menggunakan algoritma k-means dan analisis rfm di ova gaming e-sports arena kediri,” *Jurnal Teknik ITS*, vol. 10, no. 2, pp. 230–237, Dec 2021. doi: [10.12962/j23373539.v10i2.67707](https://doi.org/10.12962/j23373539.v10i2.67707).
- [19] I. Zulvia, F. A. Hidayatulloh, and E. Rahmawati, “Analisis kepuasan pasien terhadap pelayanan kesehatan di klinik alkindi herbal menggunakan metode k-means clustering,” *JIKO (Jurnal Informatika dan Komputer)*, vol. 6, no. 2, p. 261, 2022. doi: [10.26798/jiko.v6i2.612](https://doi.org/10.26798/jiko.v6i2.612).

- [20] A. L. Maukar, F. Marisa, and A. A. Widodo, "Analisis data penerimaan mahasiswa baru berbasis k-means," *JIKO (Jurnal Informatika dan Komputer)*, vol. 6, no. 2, p. 142, 2022. doi: 10.26798/jiko.v6i2.558.