

Segmentasi Pelanggan Toko ABC Menggunakan Algoritma K-Means Untuk Optimasi Strategi Pemasaran

Zevanya Aurellia Cahyadewi¹, Ellya Nurfarida², Rinanza Zulmy Alhamry³

^{1,2,3}Manajemen Informatika, Politeknik Negeri Malang

E-mail: ¹cahyadevi@gmail.com, ²ellya.nurfarida@polinema.ac.id,

³rinanza.z.alhamri@polinema.ac.id

Abstrak – Dalam menghadapi persaingan bisnis yang semakin kompetitif, perusahaan dituntut untuk melakukan strategi penjualan yang tepat. Salah satu tantangan utama perusahaan saat ini adalah membangun hubungan yang lebih kuat dengan pelanggan. Penelitian ini bertujuan untuk memberikan solusi bagi perusahaan melalui penerapan strategi segmentasi pelanggan berdasarkan perilaku pembelian pelanggan dengan melakukan analisis terhadap *Recency*, *Frequency*, *Monetary*. Analisis *Recency*, *Frequency*, *Monetary* berguna mendukung strategi pemasaran yang lebih terarah. Klasterisasi dengan algoritma K-Means digunakan untuk mengelompokkan pelanggan ke dalam beberapa segmen berdasarkan perilaku pembelian mereka. Evaluasi model dalam penelitian ini menggunakan perhitungan *Silhouette Score* dan *Davies-Bouldin Index* dalam menentukan jumlah klaster yang optimal. Hasil evaluasi menunjukkan bahwa segmentasi terbaik diperoleh pada jumlah klaster sebanyak empat, dengan nilai *Silhouette Score* sebesar 0,5644 dan *Davies-Bouldin Index* sebesar 0,7236. Hasil segmentasi ini memberikan dasar yang kuat untuk penerapan program yang lebih tepat sasaran, seperti program loyalitas, promosi retensi, dan kampanye reaktivasi pelanggan.

Kata Kunci — algoritma K-Means, Klasterisasi, Segmentasi pelanggan, strategi penjualan

1. PENDAHULUAN

Dalam era persaingan bisnis yang semakin ketat, perusahaan dituntut untuk terus berinovasi dan mengoptimalkan seluruh sumber daya yang dimiliki guna mempertahankan eksistensinya serta meningkatkan keunggulan bersaing. Salah satu aspek penting yang perlu diperhatikan adalah kemampuan perusahaan dalam memahami perilaku pelanggan dan membangun hubungan jangka panjang yang bernilai. Pelanggan tidak lagi dipandang sebagai target semata, melainkan sebagai aset strategis yang dapat menentukan arah keberhasilan perusahaan. Salah satu strategi pemasaran yang melibatkan pendekatan dengan konsumen yaitu segmentasi pelanggan. Segmentasi pelanggan merupakan proses membagi pelanggan ke dalam beberapa kelompok berdasarkan karakteristik atau perilaku tertentu.

Masalah yang terjadi pada toko ABC, yaitu belum menerapkan segmentasi pelanggan. Masalah ini memunculkan permasalahan lain, seperti strategi pemasaran yang diterapkan masih bersifat umum sehingga upaya pemasaran menjadi kurang optimal. Selain itu, tidak adanya sistem segmentasi pelanggan berbasis data menyulitkan toko dalam merancang program pemasaran yang spesifik dan tepat sasaran, seperti program loyalitas, promosi retensi, atau reaktivasi pelanggan yang sudah tidak aktif karena toko tersebut belum mampu membedakan pelanggan berdasarkan kebiasaan pembelian mereka.

Dari masalah-masalah yang disebutkan, diperlukan solusi yang dapat memberikan dampak positif bagi toko ABC dalam strategi pemasaran. Berdasarkan pengumpulan data dari penelitian yang telah ada. Penelitian yang dilakukan Dihin Muriyatmoko, Dian Fikrianti, dan Ffid Rahma Ifalus tahun 2025 berjudul “Analisis Penjualan Produk Terlaris di Toko Bangunan Pekanbaru Jaya Menggunakan Metode Clustering K-Means” menggunakan parameter berupa produk, kuantiti, unit, dan total untuk mengelompokkan data produk berdasarkan hasil penjualannya. [1].

Penelitian yang dilakukan oleh Dyah Susilowati, Hairani Hairani, Indah Puji Lestari, Khairan Marzuki, Lalu Zazuli Azhar Mardedi (2022) berjudul “Segmentasi Lokasi Promosi Penerimaan Mahasiswa Baru Menggunakan Metode RFM dan K-Means”. Tujuan penelitian ini adalah melakukan segmentasi tingkat potensial sekolah sebagai strategi untuk memilih lokasi promosi penerimaan mahasiswa baru Universitas Bumigora menggunakan analisis model RFM dan metode K-means. Hasil penelitian ini adalah terbentuk 3 klaster tingkat potensial sekolah yang dapat dijadikan skala prioritas untuk lokasi promosi penerimaan mahasiswa baru Universitas Bumigora yaitu kurang potensial, potensial, dan sangat potensial. Klaster sangat potensial (C2) terdapat 28 sekolah, klaster potensial (C3) terdapat 90 sekolah, dan klaster kurang potensial (C1) terdapat 152 sekolah [2].

Beberapa penelitian sebelumnya menunjukkan jika mengelola data segmentasi pelanggan, model yang banyak digunakan dan terbukti memiliki tingkat akurasi tinggi adalah model *Recency*, *Frequency*, *Monetary* (RFM).

Model ini menilai berdasarkan perilaku pelanggan yang dilihat dari waktu transaksi terakhir pelanggan melakukan transaksi (*Recency*), jumlah transaksi (*Frequency*), dan jumlah uang yang dibelanjakan (*Monetary*) [3]. Dengan mempelajari transaksi pelanggan di masa lalu, perusahaan dapat melakukan penilaian terhadap 3 (tiga) aspek dimensi tersebut. Selanjutnya, data RFM tersebut diolah dengan melakukan klusterisasi menggunakan algoritma *K-Means* yang mampu mengidentifikasi pola-pola perilaku dalam data pelanggan secara tidak terawasi (*unsupervised*) [4]. Tahap selanjutnya, melakukan evaluasi terhadap hasil segmentasi menggunakan metrik *Silhouette Score* dan *Davies-Bouldin Index*, guna menentukan jumlah klaster yang optimal dan kualitas pemisahan antar klaster. Pendekatan ini diharapkan dapat mendukung pengambilan keputusan yang lebih tepat dalam strategi pemasaran.

2. METODE PENELITIAN

Metodelogi penelitian yang digunakan dalam penelitian ini adalah Metodelogi *CRISP-DM*. Keunggulan dari Metodelogi *CRISP-DM* adalah fleksibilitas dan perbaikan terus menerus sepanjang siklus proyek, serta memiliki sifat sekuensial atau runtut.



Gambar 1. Metode CRISP-DM

Gambar 1 merupakan gambar alur *CRISP-DM*. Beberapa tahapan yang dilakukan dalam metodologi *CRISP-DM*, adalah sebagai berikut:

1. Pemahaman Bisnis (*Business Understanding*)

Pada tahap ini, fokus utama bisnis adalah untuk mempertahankan pelanggan lama dan menarik pelanggan baru agar terus berbelanja di toko. Salah satu solusi yang dapat diterapkan adalah dengan memberikan pelayanan yang lebih personal kepada setiap pelanggan berdasarkan karakteristik transaksi mereka. Untuk itu, pendekatan RFM (*Recency*, *Frequency*, *Monetary*) dipilih sebagai dasar dalam melakukan segmentasi pelanggan.

2. Pemahaman Data (*Data Understanding*)

Tahapan ini bertujuan untuk mengidentifikasi, mengumpulkan, dan menganalisis kumpulan data yang dapat membantu untuk mencapai tujuan proyek [10]. Data yang digunakan adalah data transaksi penjualan pelanggan. Meskipun jumlah data relatif sedikit, namun masih dapat dimanfaatkan untuk eksplorasi awal. Tabel 1 merupakan daftar fitur yang terdapat dalam dataset, berikut:

Tabel 1. Fitur dan Tipe Data

No	Nama Kolom Dataset	Tipe Data
1	NODO	VARCHAR
2	TGLTRANSAKSI	DATE
3	CUSTOMER	VARCHAR
4	KODEBRG	VARCHAR
5	JUMLAH	INT
6	SUBTOTAL	FLOAT
7	SALES	VARCHAR

3. Persiapan Data (*Data Preparation*)

Tahap ini mencakup pembersihan, transformasi, dan pengolahan awal data agar siap digunakan dalam proses pemodelan[8]. Adapun langkah-langkah yang dilakukan adalah sebagai berikut:

- Mengubah nilai negatif menjadi positif. Nilai-nilai transaksi yang seharusnya bersifat positif tetapi tercatat negatif (misalnya karena kesalahan input atau sistem) dikonversi menggunakan fungsi absolut (*abs*) agar nilainya menjadi positif dan sesuai konteks.
- Menghapus data duplikat berdasarkan tanggal. Untuk menghindari duplikasi data saat proses import, dilakukan penyaringan data berdasarkan rentang tanggal yang diinginkan. Data yang memiliki tanggal

transaksi yang sama dengan data yang sudah ada sebelumnya akan dihapus atau dikecualikan. Proses ini menggunakan filter waktu agar hanya data yang sesuai dengan rentang waktu tertentu yang diproses.

C. Menghitung nilai RFM untuk setiap pelanggan:

- Recency*: Selisih hari antara tanggal terakhir transaksi dan tanggal referensi saat ini.
- Frequency*: Jumlah transaksi yang dilakukan pelanggan.
- Monetary*: Total pengeluaran pelanggan.

Adapun sumber data RFM diperoleh dari dataset transaksi ditunjukkan pada Tabel 2.

Tabel 2. Data RFM

No	Nama Kolom Dataset	Kegunaan sebagai
1	TGL TRANSAKSI	<i>Recency</i> (R)
2	JUMLAH	<i>Frequency</i> (F)
3	SUBTOTAL	<i>Monetary</i> (M)

Berikut contoh hasil perhitungan RFM pada beberapa pelanggan ditunjukkan dalam Tabel 3 berikut.

Tabel 3. Data Perhitungan RFM

<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>
64	57	392167349,945
64	7	22918087,500
64	36	530231223,600

Untuk memastikan bahwa setiap fitur RFM memiliki skala yang beragam dan menghindari dominasi satu fitur terhadap fitur lainnya, dilakukan proses normalisasi data menggunakan metode Standard Scaler atau standardisasi *Z-score*. Berikut rumus dari metode standardisasi *Z-score*:

$$Z = \frac{\text{nilai} - \text{Mean}}{\text{std dev}}$$

Sebagai contoh proses normalisasi pada baris pertama, kebutuhan data yang digunakan ditampilkan pada Tabel 4 berikut.

Tabel 4. Data Perhitungan RFM

Data	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>
Nilai	64	57	392167349,9
Std Dev.	85, 56	24,60	170038583,1
Mean	134,98	19,66	80463455,9

Berdasarkan Tabel 3, maka hasil normalisasi *Z-score* pada baris pertama sebagai berikut:

$$Z \text{ Recency} = \frac{64 - 134,980952}{85,56434292} \approx -0,82956$$

$$Z \text{ Frequency} = \frac{57 - 19,6571429}{24,59919804} \approx 1,518052$$

$$Z \text{ Monetary} = \frac{392167349,9 - 80463455,9}{170038583,1} \approx 1,833136$$

```
array([-8.29562291e-01,  1.51805181e+00,  1.83313627e+00])
```

Gambar 2. Perhitungan Normalisasi (Sistem)

Berdasarkan hasil perhitungan manual dan hasil yang ditampilkan oleh sistem pada Gambar 2, diketahui bahwa keduanya memberikan hasil yang sama. Hal ini menunjukkan bahwa proses normalisasi yang dilakukan oleh sistem telah berjalan dengan akurasi yang tinggi dan sesuai dengan rumus yang digunakan secara manual. Dengan demikian, dapat disimpulkan bahwa sistem telah berhasil mengimplementasikan proses normalisasi secara benar. Adapun hasil normalisasi untuk beberapa baris data dapat dilihat pada Tabel 5.

Tabel 5. Hasil normalisasi RFM

<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>
-0.829562291	1.51805181	1.83313627
-0.829562291	-0.514534776	-0.338425358
-0.829562291	0.664365445	2.64509242

4. Pemodelan (*Modelling*)

Pada tahap ini, teknik data mining dilakukan untuk pencarian pola dan mengeksplorasi data terpilih dengan menggunakan teknik dan metode tertentu [11]. Tahapan ini dilakukan untuk membangun model segmentasi pelanggan menggunakan klusterisasi dengan algoritma *K-Means*. Langkah-langkahnya yang dilakukan meliputi:

A. Menentukan jumlah kluster optimal menggunakan:

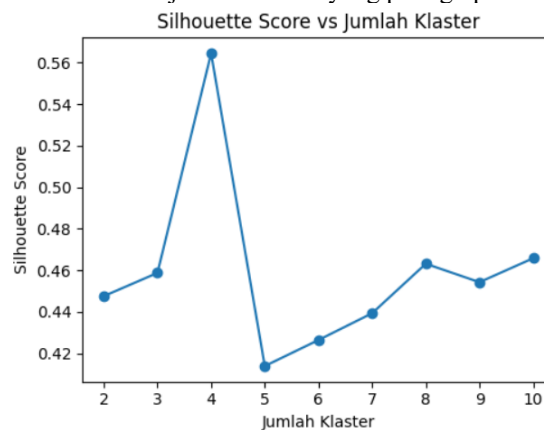
- Silhouette Score* untuk mengukur seberapa mirip suatu data dalam kluster dengan data lainnya dalam kluster yang sama.
- Davies-Bouldin Index* untuk mengevaluasi kualitas pemisahan antar kluster (semakin rendah, semakin baik).

B. Melatih model *K-Means* menggunakan data RFM yang telah di normalisasi.

C. Mengklasifikasikan pelanggan ke dalam beberapa segmen berdasarkan hasil *clustering*.

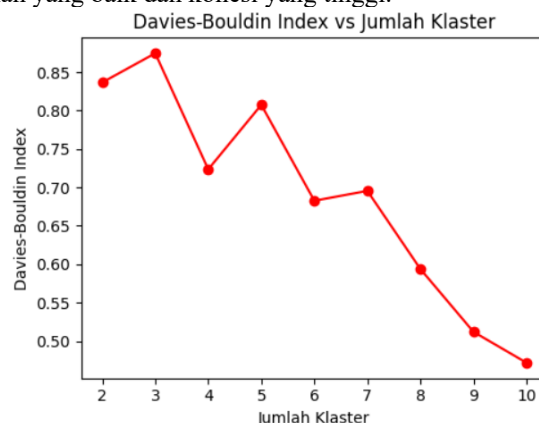
5. Evaluasi (*Evaluation*)

Evaluasi dilakukan untuk menentukan jumlah kluster yang paling optimal berdasarkan dua metrik:



Gambar 3. Hasil evaluasi jumlah kluster dengan Silhouette Score

- A. *Silhouette Score* (Gambar 3) menunjukkan bahwa nilai tertinggi diperoleh saat jumlah kluster $k = 4$ yang mengindikasikan pemisahan yang baik dan kohesi yang tinggi.



Gambar 4. Hasil evaluasi dengan DBI

- B. *Davies-Bouldin Index* (Gambar 4) menunjukkan nilai terendah pada $k = 10$.

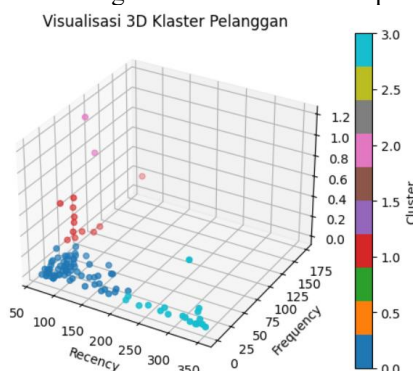
Jumlah klaster: 2, Silhouette Score: 0.4475, Davies-Bouldin Index: 0.8366
Jumlah klaster: 3, Silhouette Score: 0.4587, Davies-Bouldin Index: 0.8746
Jumlah klaster: 4, Silhouette Score: 0.5644, Davies-Bouldin Index: 0.7236
Jumlah klaster: 5, Silhouette Score: 0.4139, Davies-Bouldin Index: 0.8075
Jumlah klaster: 6, Silhouette Score: 0.4264, Davies-Bouldin Index: 0.6827
Jumlah klaster: 7, Silhouette Score: 0.4393, Davies-Bouldin Index: 0.6955
Jumlah klaster: 8, Silhouette Score: 0.4631, Davies-Bouldin Index: 0.5938
Jumlah klaster: 9, Silhouette Score: 0.4543, Davies-Bouldin Index: 0.5120
Jumlah klaster: 10, Silhouette Score: 0.4658, Davies-Bouldin Index: 0.4718

Gambar 5. Perhitungan Evaluasi

Gambar 5 merupakan perhitungan skor *Silhouette Score* dan DBI. Dengan mempertimbangkan kedua metrik tersebut secara bersamaan berdasarkan Gambar 5, jumlah klaster $k = 4$ dipilih sebagai solusi yang paling optimal. Keputusan ini didasarkan pada beberapa pertimbangan berikut:

1. *Silhouette Score* tertinggi $\rightarrow k = 4$ (0,5644)
2. DBI cukup rendah $\rightarrow k = 4$ (0,7236)
3. $k = 4$ dipilih sebagai jumlah klaster optimal karena memberikan keseimbangan terbaik antara kohesi dan pemisahan.

Dengan demikian, penggunaan empat segmen pelanggan dipandang sebagai pilihan yang paling seimbang dan sesuai untuk kebutuhan segmentasi dalam konteks penelitian ini.



Gambar 6. Hasil Visualisasi dari Implementasi $k=4$

Gambar 6 menampilkan hasil pengelompokan data pelanggan berdasarkan algoritma *K-Means Clustering* dengan jumlah klaster $k = 4$. Visualisasi ini menunjukkan penyebaran titik data dalam ruang berdimensi tiga, yaitu *Recency*, *Frequency*, dan *Monetary*, yang telah dinormalisasi. Rumus jarak *Euclidean* yang digunakan untuk tiga fitur (R, F, M) :

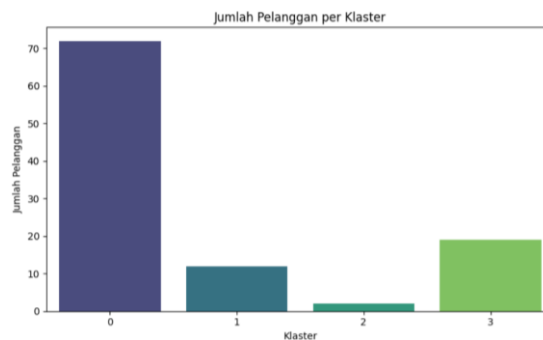
$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

Setiap klaster ditampilkan dalam warna berbeda untuk mempermudah identifikasi:

- A. Klaster 0 ditandai dengan warna biru tua
- B. Klaster 1 berwarna merah
- C. Klaster 2 berwarna pink
- D. Klaster 3 berwarna biru muda

Berdasarkan penyebaran titik pada visualisasi: Klaster 2 menunjukkan karakteristik pelanggan dengan nilai *Frequency* dan *Monetary* yang tinggi, serta *Recency* yang rendah. Hal ini mengindikasikan bahwa pelanggan dalam klaster ini merupakan pelanggan yang aktif dan memiliki pengeluaran besar, serta baru-baru ini melakukan transaksi. Sebaliknya, Klaster 3 menampilkan pelanggan dengan nilai *Frequency* dan *Monetary* yang rendah, serta *Recency* yang tinggi. Artinya, pelanggan dalam klaster ini sudah lama tidak bertransaksi, dan juga memiliki intensitas transaksi serta total pengeluaran yang rendah, yang menunjukkan tingkat keterlibatan pelanggan yang rendah.

Visualisasi ini memberikan gambaran yang jelas mengenai perbedaan perilaku pelanggan pada setiap klaster, serta memperkuat hasil evaluasi sebelumnya bahwa klaster yang terbentuk memiliki pemisahan yang cukup baik dan mendukung analisis karakteristik pelanggan secara mendalam.



Gambar 7. Jumlah Pelanggan dengan k : 4

Hasil segmentasi pada Gambar 7 menunjukkan bahwa pelanggan terbagi ke dalam empat klaster, yang terdiri dari klaster 0 sebanyak 72 pelanggan, klaster 1 sebanyak 12 pelanggan, klaster 2 sebanyak 2 pelanggan, dan klaster 3 sebanyak 19 pelanggan. Distribusi ini menunjukkan bahwa sebagian besar pelanggan berada pada klaster 0, sedangkan pelanggan pada klaster 2 sangat sedikit, namun memiliki nilai bisnis yang tinggi (premium).

6. Penerapan (*Deployment*)

Tahap ini merupakan implementasi akhir dari model segmentasi yang telah dibangun ke dalam proses bisnis nyata di toko. Hasil dari proses klusterisasi digunakan untuk memahami karakteristik pelanggan berdasarkan rata-rata nilai *Recency*, *Frequency*, dan *Monetary* dari masing-masing klaster. Analisis ini bertujuan untuk mendukung pengambilan keputusan dalam perencanaan strategi pemasaran yang lebih terarah. Langkah-langkah yang dilakukan pada tahap *deployment* meliputi:

- Menghitung rata-rata nilai RFM pada masing-masing klaster. Tabel 6 menyajikan hasil perhitungan rata-rata nilai *Recency*, *Frequency*, dan *Monetary* pada masing-masing klaster.
- Membandingkan rata-rata nilai setiap klaster dengan rata-rata keseluruhan dari dataset untuk mengidentifikasi tingkat rendah atau tinggi dari masing-masing dimensi.
- Memberikan label karakteristik pada setiap klaster berdasarkan hasil perbandingan tersebut.

Tabel 6. Rata-rata per klaster

Klaster	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>
0	103,42	14,40	35649657,81
1	73,25	64,42	262274438,4
2	71,5	82	997563066,5
3	300,26	4,74	38919374,48

Jika nilai rata-rata RFM pada suatu klaster lebih kecil dari rata-rata keseluruhan, maka dikategorikan sebagai rendah, jika sebaliknya akan dikategorikan sebagai tinggi. Berdasarkan karakteristik tersebut, setiap klaster dapat diperlakukan dengan strategi yang berbeda sesuai dengan perilaku pelanggan di dalamnya.

Tabel 7. Karakter per klaster

R	F	M	Klaster	Label
Rendah (Baik)	Rendah (Buruk)	Rendah (Buruk)	0	Pelanggan Baru
Rendah (Baik)	Tinggi (Baik)	Tinggi (Baik)	1	Sering Belanja
Rendah (Baik)	Tinggi (Baik)	Tinggi (Baik)	2	Pelanggan Setia
Tinggi (Buruk)	Rendah (Buruk)	Rendah (Buruk)	3	Pelanggan Beresiko

Tabel 7 merupakan hasil klasifikasi karakteristik masing-masing klaster berdasarkan perbandingan nilai rata-rata *Recency*, *Frequency*, dan *Monetary* terhadap rata-rata keseluruhan. Tindakan lanjutan berdasarkan labelnya yang bisa dilakukan tim pemasaran dan manajemen toko dijelaskan pada Tabel 8 sebagai berikut:

Tabel 8. Tindakan per Label

Label	Tindakan
Pelanggan Baru	Edukasi & promosi, rekomendasi produk
Sering Belanja	Upsell atau cross-sell
Pelanggan Setia	Jaga loyalitas (reward, VIP program)
Pelanggan Beresiko	Kirim diskon reaktivasi, survei kepuasan

Sistem segmentasi pelanggan siap digunakan oleh tim pemasaran dan manajemen toko.

3. HASIL DAN PEMBAHASAN

A. Hasil

1. Implementasi Halaman Login

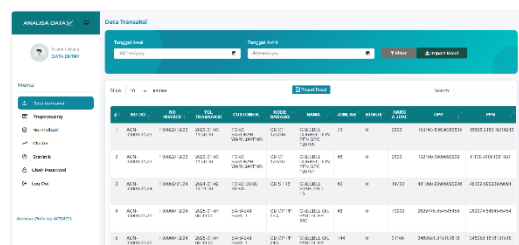
Halaman login adalah halaman pertama ketika sistem dibuka. Halaman ini menampilkan form username dan password serta informasi mengenai mekanisme login. Tampilan halaman login dapat dilihat pada Gambar 8.



Gambar 8. Implementasi Halaman Login

2. Implementasi Halaman Data Transaksi

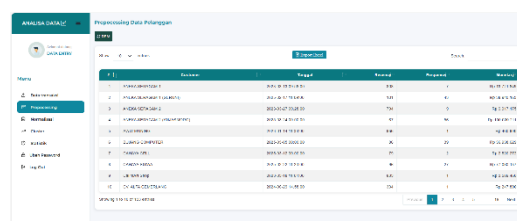
Halaman menu data transaksi adalah halaman yang memuat semua pengelolaan data import. Tampilan halaman data transaksi dapat dilihat pada Gambar 9.



Gambar 9. Implementasi Halaman Data Transaksi

3. Implementasi Halaman Preprocessing

Halaman menu *preprocessing* merupakan halaman yang digunakan untuk melakukan pengolahan data awal (*preprocessing*). Data yang ditampilkan pada halaman ini berasal dari proses import yang sebelumnya telah dilakukan pada menu Import. Selanjutnya, data tersebut akan melalui tahap *cleaning* atau pembersihan untuk mempersiapkannya sebelum masuk ke tahap selanjutnya yaitu proses normalisasi. Pengolahan data dilakukan berdasarkan kebutuhan data *Recency*, *Frequency*, *Monetary*. Tampilan halaman *preprocessing* dapat dilihat pada Gambar 10.



Gambar 10. Implementasi Halaman Preprocessing

4. Implementasi Halaman Normalisasi

Halaman menu normalisasi merupakan halaman yang digunakan untuk melakukan proses normalisasi data serta penentuan kluster data menggunakan algoritma *K-Means*. Data yang ditampilkan pada halaman ini berasal dari hasil proses *cleaning* pada tahap sebelumnya. Pada halaman ini, pengguna dapat melakukan pengecekan secara manual untuk memastikan bahwa proses normalisasi data serta perhitungan kluster dengan metode *K-Means* menggunakan jarak *Euclidean* telah dilakukan dengan benar. Tampilan halaman normalisasi dapat dilihat pada Gambar 11.

Nama	Alamat	No. HP	Email	Recency	Frequency	Monetary	Cluster	Label
1. ANANDA CHANDRA	123456789	08123456789	ananda.chandra@gmail.com	10	10	1000000000	1	1
2. ANANDA CHANDRA	123456789	08123456789	ananda.chandra@gmail.com	10	10	1000000000	1	1
3. ANANDA CHANDRA	123456789	08123456789	ananda.chandra@gmail.com	10	10	1000000000	1	1
4. ANANDA CHANDRA	123456789	08123456789	ananda.chandra@gmail.com	10	10	1000000000	1	1
5. ANANDA CHANDRA	123456789	08123456789	ananda.chandra@gmail.com	10	10	1000000000	1	1

Gambar 11. Implementasi Halaman Normalisasi

5. Implementasi Halaman Kluster

Halaman menu kluster merupakan halaman yang menampilkan seluruh data yang telah dikelompokkan menggunakan algoritma *K-Means*. Setelah proses klusterisasi selesai, masing-masing kluster diberikan karakteristik dan label tertentu berdasarkan hasil analisis nilai *Recency*, *Frequency*, dan *Monetary* (RFM) dari data tersebut. Proses pelabelan ini bertujuan untuk mempermudah interpretasi hasil klusterisasi serta menentukan strategi atau tindakan khusus yang sesuai untuk masing-masing kluster.

Nama	Alamat	No. HP	Email	Recency	Frequency	Monetary	Cluster	Label
1. ANANDA CHANDRA	123456789	08123456789	ananda.chandra@gmail.com	10	10	1000000000	1	1
2. ANANDA CHANDRA	123456789	08123456789	ananda.chandra@gmail.com	10	10	1000000000	1	1
3. ANANDA CHANDRA	123456789	08123456789	ananda.chandra@gmail.com	10	10	1000000000	1	1
4. ANANDA CHANDRA	123456789	08123456789	ananda.chandra@gmail.com	10	10	1000000000	1	1
5. ANANDA CHANDRA	123456789	08123456789	ananda.chandra@gmail.com	10	10	1000000000	1	1

Gambar 12. Implementasi Halaman Kluster

Gambar 12 menggambarkan tampilan dari halaman kluster. Karakteristik pada tampilan diatas memiliki pola tindakan yang berbeda dan akan dijelaskan pada Gambar 13.

Karakter	Tindakan per Cluster	Label
Rendah Kurang dari rata-rata setiap cluster	Pelanggan Setia Beri hadiah (reward, VIP program)	1
Tinggi Lebih dari rata-rata setiap cluster	Pelanggan Baru Edukasi & promosi, rekomendasi produk (produk frekuensi)	2
	Pelanggan Berisiko Kirim status kesehatan, surat kesehatan	3

Gambar 13. Penjelasan Karakter dan Tindakan

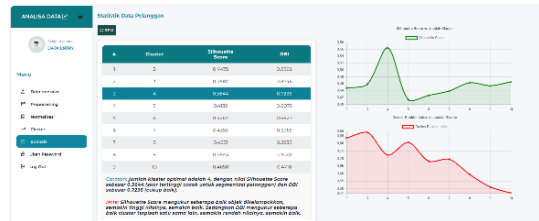
Gambar 13 merupakan analisis lebih lanjut terhadap hasil klusterisasi dan tindakan yang disarankan. *Recency* (R) dikategorikan Baik apabila nilainya lebih rendah dari rata-rata *Recency* keseluruhan, karena menunjukkan pelanggan baru-baru ini melakukan transaksi. Sebaliknya, *Frequency* (F) dan *Monetary* (M) dianggap Baik apabila nilainya lebih tinggi dari rata-rata masing-masing, karena menunjukkan frekuensi pembelian yang tinggi dan nilai belanja yang besar. Berdasarkan hasil klusterisasi dan pelabelan data, setiap kategori pelanggan memiliki karakteristik yang berbeda sehingga memerlukan pendekatan atau tindakan yang disesuaikan. Berikut adalah strategi yang dapat diterapkan untuk masing-masing kategori:

- Pelanggan Setia**
Pelanggan yang termasuk dalam kategori ini memiliki nilai *recency*, *frequency*, dan *monetary* yang tinggi. Tindakan yang dapat dilakukan adalah menjaga loyalitas mereka melalui pemberian *reward*, program loyalitas, atau keanggotaan eksklusif seperti program VIP.
- Pelanggan Sering Belanja**
Pelanggan ini memiliki frekuensi pembelian yang tinggi, namun mungkin nilai transaksi (*monetary*) belum maksimal. Strategi yang sesuai adalah melakukan pendekatan *upselling* atau *cross-selling* untuk meningkatkan nilai pembelian mereka.
- Pelanggan Baru**
Merupakan pelanggan yang baru pertama kali melakukan transaksi. Pendekatan yang tepat untuk kategori ini meliputi edukasi tentang produk atau layanan, pemberian promosi awal, serta rekomendasi produk yang relevan untuk mendorong pembelian lanjutan.
- Pelanggan Berisiko**

Pelanggan dalam kategori ini menunjukkan penurunan aktivitas atau memiliki nilai *recency* yang rendah. Untuk mencegah kehilangan pelanggan, tindakan yang bisa dilakukan adalah memberikan diskon reaktif atau penawaran khusus guna menarik mereka kembali bertransaksi.

6. Implementasi Halaman Statistik

Halaman menu statistik merupakan halaman yang menampilkan visualisasi dan hasil evaluasi terhadap kualitas klusterisasi yang telah dilakukan. Evaluasi ini ditunjukkan melalui grafik serta skor yang diperoleh dari dua metode pengukuran, yaitu: *Silhouette Score* dan DBI.



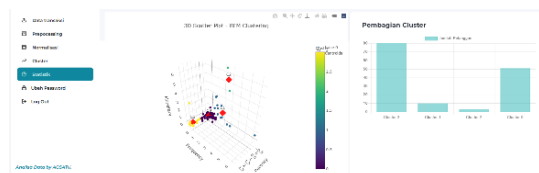
Gambar 14. Implementasi Halaman Statistik

Gambar 14 menampilkan visualisasi nilai *Silhouette Score* dan *Davies-Bouldin Index* (DBI) untuk berbagai jumlah kluster (k) yang diuji dalam proses evaluasi klusterisasi. Nilai *Silhouette Score* digambarkan dalam grafik berwarna hijau, di mana skor mendekati angka 1 menunjukkan hasil klusterisasi yang semakin baik, karena data dalam kluster lebih kohesif dan terpisah dengan baik dari kluster lain.

Sebaliknya, *DBI* divisualisasikan dalam grafik berwarna merah, dengan nilai ideal mendekati 0, yang menandakan kluster yang saling terpisah dengan baik dan memiliki tingkat tumpang tindih yang rendah. Berdasarkan hasil evaluasi tersebut, jumlah kluster optimal dipilih sebanyak empat kluster ($k = 4$), dengan pertimbangan sebagai berikut:

- Silhouette Score* tertinggi diperoleh pada $k = 4$, yaitu sebesar 0,5644.
- Nilai *DBI* pada $k = 4$ juga masih tergolong baik, yaitu sebesar 0,7236.

Dengan demikian, pemilihan $k = 4$ dianggap sebagai jumlah kluster yang optimal karena memberikan keseimbangan terbaik antara kohesi internal dan pemisahan antar kluster.



Gambar 15. Visualisasi Grafik Pengelompokan Kluster

Gambar 15 menampilkan visualisasi hasil pengelompokan data berdasarkan proses klusterisasi menggunakan algoritma *K-Means* dengan jumlah kluster sebanyak empat ($k = 4$). Visualisasi terdiri dari dua bagian:

- Gambar pertama menampilkan persebaran titik data berdasarkan kluster, disertai dengan posisi centroid masing-masing kluster.
- Gambar kedua memperlihatkan jumlah data per kluster dalam bentuk grafik batang.

Visualisasi ini memberikan gambaran umum mengenai struktur kluster dan distribusi data di setiap kelompok.

B. Pengujian

Pada bagian ini menjelaskan tentang Langkah-langkah pengujian yang telah dilakukan. Pengujian bertujuan untuk mengetahui apakah produk yang dihasilkan berfungsi dengan benar. Data yang dimuat dalam bentuk tabel dan gambar harus diberi judul dengan nomor urut. Di bagian bawah Tabel dan Gambar diberi sumber data tersebut lengkap dengan tahunnya. Berikut adalah skenario pengujian sistem "Segmentasi Pelanggan pada Toko ABC Menggunakan Analisis RFM dan *K-Means*" yang user yang diuraikan pada Tabel 9 sebagai berikut.

Tabel 9. Hasil Pengujian

Skenario	Prosedur pengujian	Hasil yang diharapkan	Hasil pengujian
Login dan logout	Melakukan otentikasi dengan memasukkan username dan password serta dapat melakukan logout	Sistem berhasil memverifikasi dan mengarahkan pengguna ke halaman dashboard sesuai dengan akses yang dimiliki serta dapat keluar dari sistem.	Sesuai
Menu Data Transaksi (Kelola RFM)	Melakukan import data sesuai dengan parameter yang telah ditentukan.	Data excel yang diimport berhasil tersimpan ke dalam database dan sistem dapat menampilkan data ke dalam tabel, serta dapat diolah menjadi data RFM.	Sesuai
Menu preprocessing	Melakukan refresh data RFM.	Sistem berhasil menyimpan data yang telah diperbarui ke dalam database, serta menampilkan data RFM dalam bentuk tabel.	Sesuai
Menu normalisasi	Melihat data normalisasi	Sistem menampilkan data yang telah dinormalisasi beserta hasil pengelompokan data klaster.	Sesuai
Menu Klaster	Melakukan refresh data RFM dan filter klaster.	Sistem telah berhasil menampilkan hasil klasterisasi terbaru, termasuk informasi karakteristik dan label tiap klaster, serta memungkinkan pengguna untuk memfilter data berdasarkan klaster tertentu.	Sesuai
Menu Statistik	Melihat grafik berdasarkan perhitungan <i>Silhouette Score</i> serta DBI, dan pengelompokan klaster.	Sistem berhasil menampilkan grafik berdasarkan perhitungan <i>Silhouette Score</i> dan DBI, serta hasil pengelompokan klaster.	Sesuai

Sumber: Data primer tahun 2025

4. SIMPULAN

Dari hasil dan pembahasan yang dilakukan dapat disimpulkan:

1. Segmentasi pelanggan berhasil dilakukan menggunakan metode *K-Means Clustering* berdasarkan analisis data *Recency*, *Frequency*, dan *Monetary* (RFM). Metode ini mampu mengelompokkan pelanggan ke dalam beberapa klaster yang berbeda berdasarkan perilaku belanjanya.
2. Hasil evaluasi menggunakan *Silhouette Score* dan *Davies-Bouldin Index* menunjukkan bahwa jumlah klaster optimal adalah 4, dengan nilai *Silhouette Score* sebesar 0,5644 dan *DBI* sebesar 0,7236. Nilai-nilai ini menandakan bahwa hasil klasterisasi cukup baik dalam hal kohesi dan pemisahan antar kelompok.
3. Setiap klaster memiliki karakteristik yang berbeda-beda, yang dapat dijadikan dasar dalam merancang strategi pemasaran yang lebih tepat sasaran, seperti program loyalitas untuk pelanggan setia, kampanye reaktivasi untuk pelanggan berisiko, serta promosi edukatif untuk pelanggan baru.

5. SARAN

Penelitian lebih lanjut dapat dilakukan dengan menambahkan faktor lain seperti membership dan jenis kelamin sebagai faktor tambahan dari pengelompokan sehingga menghasilkan kelompok yang lebih ideal untuk menentukan strategi penjualan guna meningkatkan keuntungan perusahaan ABC.

DAFTAR PUSTAKA

- [1] D. Muriatmoko, D. Fikrianti and F. R. Ifalus, "Analisis Penjualan Produk Terlaris di Toko Bangunan Pekanbaru Jaya Menggunakan Metode Clustering K-Means," in *Stains : Seminar Nasional Teknologi dan Sains*, Kediri, 2025.
- [2] H. Hairani, D. Susilowati, I. Puji Lestari, K. Marzuki, and L. Z. A. Mardedi, "Segmentasi Lokasi Promosi Penerimaan Mahasiswa Baru Menggunakan Metode RFM dan K-Means Clustering," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 2, pp. 275–282, 2022, doi: 10.30812/matrik.v21i2.1542.
- [3] K. Tsipsis and A. Chorianopoulos, *Data Mining Techniques in CRM: Inside Customer Segmentation*. 2010. doi: 10.1002/9780470685815.
- [4] P. Li, C. Wang, J. Wu, and R. Madlenak, "An E-commerce Customer Segmentation Method based on RFM Weighted K-means," *Proc. - 2022 Int. Conf. Manag. Eng. Softw. Eng. Serv. Sci. ICMSS 2022*, pp. 61–68, 2022, doi: 10.1109/ICMSS55574.2022.00017.
- [5] A. T. Widiyanto and A. Witanti, "Segmentasi Pelanggan Berdasarkan Analisis RFM Menggunakan Algoritma K-Means Sebagai Dasar Strategi Pemasaran (Studi Kasus PT Coversuper Indonesia Global)," *KONSTELASI Konvergensi Teknol. dan Sist. Inf.*, vol. 1, no. 1, pp. 204–215, 2021, doi: 10.24002/konstelasi.v1i1.4293.
- [6] K. Z. Wijaya, A. Djunaidi, and F. Mahananto, "Segmentasi Pelanggan Menggunakan Algoritma K-Means dan Analisis RFM di Ova Gaming E-Sports Arena Kediri," *J. Tek. ITS*, vol. 10, no. 2, 2021, doi: 10.12962/j23373539.v10i2.67707.
- [7] P. A. Wicaksana, I. B. A. Swamardika, and R. S. Hartati, "Literature Review Analisis Perilaku Pelanggan Menggunakan RFM Model," *Maj. Ilm. Teknol. Elektro*, vol. 21, no. 1, p. 21, 2022, doi: 10.24843/mite.2022.v21i01.p04.
- [8] P. Indra Pangestu, T. Iman Hermanto, and D. Irmayanti, "Analisis Segmentasi Pelanggan Berbasis Recency Frequency Monetary (Rfm) Menggunakan Algoritma K-Means," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 3, pp. 1486–1492, 2023, doi: 10.36040/jati.v7i3.7171.
- [9] S. F. Djun, I. G. A. Gunadi, and S. Sariyasa, "Analisis Segmentasi Pelanggan pada Bisnis dengan Menggunakan Metode K-Means Clustering pada Model Data RFM," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 5, no. 4, pp. 354–364, 2024, doi: 10.35746/jtim.v5i4.434.
- [10] M. A. Hasanah, S. Soim, and A. S. Handayani, "Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir," *J. Appl. Informatics Comput.*, vol. 5, no. 2, pp. 103–108, 2021, doi: 10.30871/jaic.v5i2.3200.
- [11] D. U. Iswavigra, L. E. Zen, and H. Hanim, "Marketing Strategy UMKM Dengan CRISP-DM Clustering & Promotion Mix Menggunakan Metode K-Medoids," *J. Inf. dan Teknol.*, vol. 5, pp. 45–54, 2023, doi: 10.37034/jidt.v5i1.260.