



## Segmentasi Perilaku Belanja Pelanggan dengan K-Means Clustering

Nabila Amanda<sup>1</sup>, Steaven Yehezkiel Sidabutar<sup>2</sup>, Dominikus Zemika Pratama<sup>3</sup>, Audy Rizky Butar Butar<sup>4</sup>, Indra M Sarkis S<sup>5</sup>

<sup>1,2,3,4,5</sup>Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Methodist Indonesia

Jl. Hang Tuah No. 8, Madras Hulu, Kec. Medan Polonia, Medan, Indonesia

<sup>1</sup>nabilaamanda2701@gmail.com, <sup>2</sup>steavensidabutar@gmail.com, <sup>3</sup>dominikuszemika@gmail.com, <sup>4</sup>audybutarbutar9@gmail.com, <sup>5</sup>poetramora@gmail.com

### ARTICLE INFORMATION

Received: July 07, 2026

Revised: September 07, 2026

Available online: September 10, 2026

### KEYWORDS

Pengelompokan K-Means, Segmentasi Konsumen, Metode Elbow, Silhouette Score, Pola Belanja

### CORRESPONDENCE

Phone: +62 822-6759-7599

E-mail: nabilaamanda2701@gmail.com

### ABSTRAK

Penelitian ini bertujuan mengelompokkan konsumen berdasarkan pendapatan tahunan dan skor belanja menggunakan metode K-Means Clustering agar diperoleh segmen konsumen yang lebih beragam untuk merumuskan strategi pemasaran. Desain/metode/pendekatan: Dataset yang digunakan adalah Mall Customer Segmentation Data dari Kaggle dengan 200 data konsumen. Tahapan meliputi pra-proses data, normalisasi Z-score, penentuan jumlah *cluster* optimal dengan metode Elbow dan Silhouette Score untuk  $k = 2$  hingga  $k = 10$ , serta pembentukan model K-Means. Temuan/hasil:  $k = 5$  memberikan Silhouette Score tertinggi, yaitu 0,5547. Model lima *cluster* membentuk segmen Pelanggan Standar (40,5%), Premium (19,5%), Impulsif (11,0%), Hemat (17,5%), dan Pasif (11,5%). Kesimpulan: Hasil segmentasi dapat menjadi landasan bagi manajemen ritel dalam merumuskan strategi pemasaran yang lebih individual dan efisien.

## 1. PENDAHULUAN

Pertumbuhan yang cepat dalam bisnis ritel dan *e-commerce* selama satu dekade terakhir telah menyebabkan peningkatan signifikan dalam volume transaksi yang harus dikelola oleh pelaku usaha. Sebenarnya, data transaksi ini menyimpan informasi penting tentang pola dan perilaku belanja konsumen, tetapi dalam praktiknya, banyak pelaku usaha yang masih menerapkan pendekatan pemasaran yang seragam (satu strategi untuk semua konsumen) tanpa mempertimbangkan perbedaan karakteristik masing-masing kelompok konsumen. Pendekatan seperti ini berpotensi mengurangi efektivitas promosi, menyebabkan pemborosan dalam anggaran pemasaran, serta menurunkan loyalitas konsumen akibat tawaran yang tidak sesuai dengan kebutuhan mereka.

Segmentasi konsumen merupakan salah satu solusi untuk mengatasi masalah ini, yaitu dengan mengelompokkan konsumen ke dalam beberapa segmen yang memiliki karakteristik belanja yang relatif sama dalam satu kelompok, namun berbeda dengan kelompok yang lain. Salah satu metode pembelajaran tanpa pengawasan yang banyak digunakan untuk segmentasi adalah algoritma K-Means Clustering, karena memiliki konsep yang sederhana, proses komputasi yang tidak berat, dan hasil pengelompokan yang mudah dipahami untuk keperluan bisnis.

Beberapa penelitian sebelumnya telah mengkaji penggunaan K-Means Clustering dalam berbagai konteks segmentasi konsumen. Penelitian [1] dan [2] menerapkan algoritma *clustering* untuk mengategorikan konsumen *e-commerce* berdasarkan perilaku pembelian, sementara [3] Menggabungkan usia, pendapatan, dan model RFM (Recency, Frequency, Monetary) untuk segmentasi konsumen di sebuah toko ritel. Penelitian [4] berfokus pada

segmentasi konsumen *e-commerce* berdasarkan pola pembelian, sedangkan [5] Menerapkan K-Means pada data perilaku penggunaan kartu kredit.

Selain itu, [6] melakukan analisis segmentasi konsumen secara umum dengan algoritma K-Means, [7] mengaitkan segmentasi pelanggan dengan data kepribadian dalam konteks *data science*, dan [8] menggunakan K-Means pada transaksi retail *online* untuk mengoptimalkan strategi pemasaran. Penelitian [9] menggunakan studi kasus aplikasi Alfagift, [10] menerapkan metode yang serupa pada data penjualan *online*, dan [11] menekankan pemanfaatan hasil *clustering* untuk merancang strategi pemasaran yang efektif. Penelitian [12] mengombinasikan data kepribadian dan pola konsumsi, [13] berfokus pada optimalisasi penjualan sembako, dan [14] mengelompokkan perilaku pelanggan secara umum menggunakan K-Means.

Dari tinjauan terhadap keempat belas penelitian tersebut, dapat disimpulkan bahwa sebagian besar penelitian menggunakan kombinasi fitur dan studi kasus yang bervariasi, tetapi belum banyak yang secara langsung membandingkan beberapa nilai *k* menggunakan dua indikator sekaligus, yaitu metode *Elbow* dan *Silhouette Score*, sebelum menentukan jumlah *cluster* akhir. Kebanyakan penelitian juga langsung menetapkan jumlah *cluster* tanpa menjelaskan secara mendetail proses validasinya, sehingga dasar pemilihan jumlah *cluster* kurang dapat diukur secara kuantitatif.

Dengan adanya kesenjangan tersebut, penelitian ini bertujuan untuk menerapkan algoritma K-Means Clustering pada *dataset* Mall Customer Segmentation Data untuk mengelompokkan pelanggan berdasarkan pendapatan tahunan dan skor belanja, dengan menentukan jumlah *cluster* optimal yang divalidasi dengan jelas menggunakan metode *Elbow* dan *Silhouette Score*, serta menghasilkan profil setiap segmen beserta rekomendasi strategi pemasaran yang bisa langsung diterapkan oleh pelaku usaha ritel.

## 2. METODE

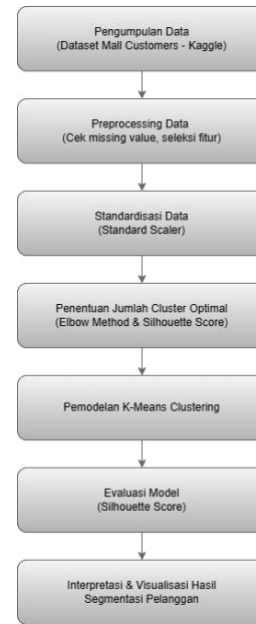
Penelitian ini dilakukan melalui beberapa langkah yang terstruktur, mulai dari pengumpulan informasi hingga pemahaman hasil segmentasi, seperti yang dijelaskan pada bagian berikut.

### 1. Dataset Penelitian

Dataset yang digunakan dalam studi ini adalah Data Segmentasi Pelanggan Mall yang diperoleh dari platform Kaggle. Dataset ini terdiri dari 200 data pelanggan di sebuah pusat perbelanjaan (*mal*) dengan lima atribut, yakni *CustomerID*, *Gender* (jenis kelamin), *Age* (usia), *Annual Income* (pendapatan tahunan dalam ribuan dolar), serta *Spending Score* (skor belanja antara 1-100 yang mencerminkan perilaku dan pola belanja pelanggan berdasarkan penilaian pusat perbelanjaan).

### 2. Tahapan Penelitian

Tahapan penelitian yang dilakukan ditampilkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

#### 1. Praproses Data

Langkah awal pada praproses data adalah memeriksa *missing value* di semua atribut *dataset*. Hasil pemeriksaan menunjukkan bahwa tidak ada nilai yang hilang dari 200 data pelanggan. Atribut *CustomerID* tidak digunakan dalam proses pengelompokan karena hanya berfungsi sebagai identitas dan tidak mencerminkan karakteristik perilaku belanja. Fitur utama dalam proses segmentasi adalah *Annual Income* dan *Spending Score*, karena kedua fitur ini secara langsung menggambarkan kapasitas ekonomi dan kecenderungan perilaku belanja pelanggan.

#### 2. Standardisasi Data

Karena *Annual Income* dan *Spending Score* memiliki rentang nilai yang berbeda, dilakukan normalisasi data dengan menggunakan *Z-score normalization* (*StandardScaler*) agar kedua fitur dapat memberikan kontribusi yang seimbang terhadap perhitungan jarak dalam algoritma K-Means. Rumus standardisasi tercantum dalam Persamaan (1).

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

#### 3. Algoritma K-Means Clustering

K-Means merupakan algoritma pembelajaran tanpa pengawasan yang bertujuan untuk mengelompokkan sejumlah *n* data ke dalam *k* kelompok berdasarkan kedekatan jarak (biasanya dengan menggunakan jarak Euclidean) terhadap *centroid* masing-masing kelompok. Proses yang dilakukan dalam algoritma K-Means adalah sebagai berikut: (1) menentukan jumlah kelompok (*k*) dan secara acak menginisialisasi *k* *centroid* awal; (2) menghitung jarak setiap data ke seluruh *centroid* dengan Persamaan (2) dan mengelompokkan data ke *centroid* yang terdekat; (3) memperbarui posisi *centroid* dengan menghitung rata-rata seluruh data dalam kelompok tersebut; (4) mengulangi langkah (2) dan (3) hingga posisi *centroid* tidak lagi mengalami perubahan yang signifikan (konvergen).

$$d(x, c) = \sqrt{\sum (x_i - c_i)^2} \quad (2)$$

Proses iterasi K-Means bertujuan untuk meminimalkan nilai *Within-Cluster Sum of Squares* (WCSS atau *inertia*) yang dinyatakan dalam Persamaan (3).

$$WCSS = \sum |x_i - c_j|^2 \quad (3)$$

#### 4. Penentuan Jumlah Cluster Optimal

Penetapan jumlah *cluster* (k) yang optimal dilakukan dengan menggunakan dua pendekatan yang saling berkaitan, yaitu *Elbow Method* dan *Silhouette Score*. *Elbow Method* memeriksa titik 'siku' pada grafik nilai WCSS terhadap jumlah *cluster*, yaitu titik di mana penurunan WCSS mulai melambat. *Silhouette Score* mengukur seberapa baik sebuah data sesuai dengan *cluster* yang dimilikinya dibandingkan dengan *cluster* lain, dengan rentang nilai dari -1 hingga 1; semakin mendekati 1, semakin baik struktur *cluster*-nya. Hal ini dijelaskan dalam Persamaan (4).

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (4)$$

### 3. Alat dan Bahan

Penelitian ini memanfaatkan bahasa pemrograman *Python 3* dengan dukungan dari pustaka *pandas* untuk pengolahan data, *scikit-learn* untuk penerapan algoritma K-Means dan perhitungan *Silhouette Score*, serta *matplotlib* untuk menggambarkan hasil. Proses komputer dilakukan di perangkat komputer dengan spesifikasi biasa tanpa harus menggunakan sumber daya komputasi khusus, mengingat ukuran *dataset* yang tergolong kecil (200 baris data).

## 3. HASIL DAN PEMBAHASAN

### 1. Deskripsi Data

Analisis deskriptif dilakukan pada 200 data pelanggan untuk mengetahui karakteristik umum *dataset* sebelum *clustering* dilakukan. Dari segi jenis kelamin, *dataset* ini terdiri dari 112 pelanggan perempuan (56%) dan 88 pelanggan laki-laki (44%). Ringkasan statistik deskriptif untuk atribut numerik ditampilkan pada Tabel 1.

Tabel 1. Statistik Deskriptif Dataset

| Variabel                 | Minimum | Maksimum | Rata-rata | Std. Deviasi |
|--------------------------|---------|----------|-----------|--------------|
| Usia (tahun)             | 18      | 70       | 38,85     | 13,97        |
| Pendapatan Tahunan (k\$) | 15      | 137      | 60,56     | 26,26        |
| Skor Belanja (1-100)     | 1       | 99       | 50,20     | 25,82        |

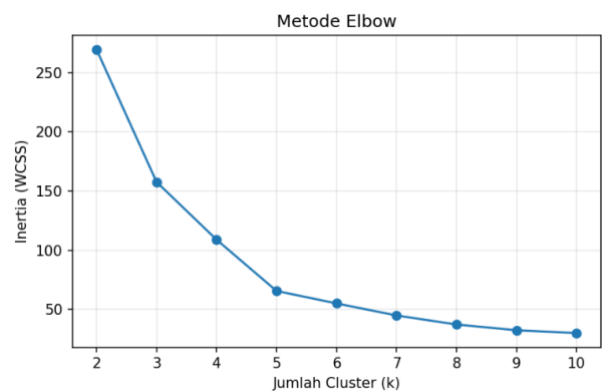
Tabel 1 menunjukkan bahwa rentang usia pelanggan berkisar antara 18 hingga 70 tahun dengan rata-rata 38,85 tahun, pendapatan tahunan berada di antara 15 hingga 137 (dalam ribuan dolar) dengan rata-rata 60,56, dan skor belanja bervariasi antara 1 hingga 99 dengan rata-rata 50,20. Sebaran nilai yang luas pada ketiga atribut ini menunjukkan adanya variasi dalam karakteristik pelanggan yang mungkin bisa dikelompokkan menjadi beberapa segmen.

### 2. Penentuan Jumlah Cluster Optimal

Proses penentuan jumlah *cluster* optimal dilakukan dengan menguji nilai k dari 2 sampai 10 menggunakan fitur *Annual Income* dan *Spending Score* yang sudah dinormalisasi. Hasil pengujian tersebut disajikan dalam Tabel 2 dan divisualisasikan pada Gambar 2.

Tabel 2. Nilai Inertia dan Silhouette Score untuk k = 2 hingga k = 10

| Jumlah Cluster (k) | Inertia (WCSS) | Silhouette Score |
|--------------------|----------------|------------------|
| 2                  | 269,69         | 0,3213           |
| 3                  | 157,70         | 0,4666           |
| 4                  | 108,92         | 0,4939           |
| 5                  | 65,57          | 0,5547           |
| 6                  | 55,06          | 0,5399           |
| 7                  | 44,86          | 0,5281           |
| 8                  | 37,23          | 0,4552           |
| 9                  | 32,39          | 0,4571           |
| 10                 | 29,98          | 0,4432           |

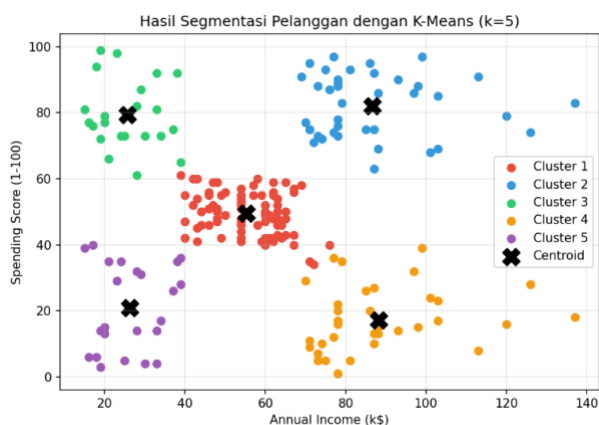


Gambar 2. Grafik Metode Elbow

Grafik dalam Gambar 2 menunjukkan adanya penurunan yang signifikan pada nilai *inersia* hingga k = 5, setelah itu mengalami pelemahan pada nilai k berikutnya. Hal ini mengindikasikan bahwa k = 5 bisa dianggap sebagai titik siku. Temuan ini didukung oleh nilai *Silhouette Score* yang terlihat pada Tabel 2, di mana k = 5 memberikan nilai tertinggi, yaitu 0,5547, di antara semua nilai k lainnya yang diuji. Berdasarkan dua indikator ini, penelitian ini menetapkan k = 5 sebagai jumlah *cluster* yang paling optimal untuk membangun model K-Means akhir.

### 3. Hasil Clustering

Model K-Means dengan k = 5 kemudian dilatih menggunakan fitur *Annual Income* dan *Spending Score* yang telah dinormalisasi. Hasil visualisasi segmentasi pelanggan disajikan dalam Gambar 3, sementara profil masing-masing *cluster* dapat dilihat di Tabel 3.



Gambar 3. Hasil Segmentasi Pelanggan dengan K-Means ( $k = 5$ )

Tabel 3. Profil Hasil Segmentasi Pelanggan

| Cluster | N  | Persentase | Rata-rata Usia | Rata-rata Pendapatan (k\$) | Rata-rata Skor Belanja | Label Segmen       |
|---------|----|------------|----------------|----------------------------|------------------------|--------------------|
| 1       | 81 | 40,5%      | 42,72          | 55,30                      | 49,52                  | Pelanggan Standar  |
| 2       | 39 | 19,5%      | 32,69          | 86,54                      | 82,13                  | Pelanggan Premium  |
| 3       | 22 | 11,0%      | 25,27          | 25,73                      | 79,36                  | Pelanggan Impulsif |
| 4       | 35 | 17,5%      | 41,11          | 88,20                      | 17,11                  | Pelanggan Hemat    |
| 5       | 23 | 11,5%      | 45,22          | 26,30                      | 20,91                  | Pelanggan Pasif    |

Berdasarkan Tabel 3 dan Gambar 3, lima *cluster* yang terbentuk dapat dijelaskan sebagai berikut:

**Cluster 1 – Pelanggan Standar (81 pelanggan, 40,5%):** merupakan kelompok pelanggan yang pendapatan dan skor belanjanya berada di tingkat menengah. Segmen ini merupakan yang terbesar dan menunjukkan pola belanja yang moderat, sehingga strategi pemasaran yang cocok adalah program loyalitas dan promosi rutin untuk meningkatkan frekuensi pembelian.

**Cluster 2 – Pelanggan Premium (39 pelanggan, 19,5%):** adalah kelompok dengan pendapatan dan skor belanja yang tinggi. Segmen ini menjadi sasaran utama karena memiliki daya beli dan kecenderungan berbelanja yang tinggi, sehingga produk premium, layanan khusus, atau program keanggotaan VIP sangat sesuai untuk mereka.

**Cluster 3 – Pelanggan Impulsif (22 pelanggan, 11,0%):** terdiri dari individu dengan pendapatan rendah namun memiliki skor belanja yang tinggi. Ini menunjukkan adanya kecenderungan untuk berbelanja impulsif. Strategi yang tepat untuk kelompok ini adalah penawaran diskon, cicilan, atau penawaran terbatas (*flash sale*) yang bisa mendorong transaksi tanpa membebani keuangan mereka.

**Cluster 4 – Pelanggan Hemat (35 pelanggan, 17,5%):** kelompok ini memiliki pendapatan yang tinggi, namun skor belanja yang rendah menunjukkan perilaku belanja yang sangat selektif. Pendekatan pemasaran yang efektif menekankan pada nilai produk, kualitas, dan manfaat jangka panjang untuk meyakinkan mereka untuk melakukan pembelian.

**Cluster 5 – Pelanggan Pasif (23 pelanggan, 11,5%):** kelompok ini terdiri dari individu dengan pendapatan dan tingkat pengeluaran yang rendah. Segmen ini menunjukkan sikap yang kurang aktif dalam berbelanja, sehingga upaya pemasaran sebaiknya diarahkan pada pendidikan mengenai produk, penawaran harga yang lebih terjangkau, atau paket hemat untuk meningkatkan ketertarikan mereka terhadap belanja.

#### 4. Evaluasi Model

Model K-Means dengan  $k = 5$  memberikan nilai *Silhouette Score* sebesar 0,5547. Berdasarkan skala yang disusun oleh Kaufman dan Rousseeuw, nilai *Silhouette Score* yang berada antara 0,51 dan 0,70 menandakan bahwa struktur *cluster* yang terbentuk adalah wajar, yang berarti setiap *cluster* memiliki tingkat pemisahan dan kerapatan yang cukup baik, meski ada sedikit tumpang tindih pada beberapa data di sekitar batas antar-*cluster*, seperti yang dapat dilihat pada Gambar 3.

#### 5. Perbandingan dengan Penelitian Terdahulu

Temuan dari penelitian ini sejalan dengan hasil yang terdapat dalam penelitian sebelumnya [1], [2], [4], [8], [10], dan [11]. Hal ini juga menunjukkan bahwa algoritma K-Means efektif dalam mengelompokkan pelanggan berdasarkan data transaksi maupun perilaku pembelian, dengan jumlah *cluster* optimal yang bervariasi antara 3 dan 5, tergantung pada *dataset* dan kombinasi fitur yang digunakan. Perbedaan utama antara penelitian ini dengan yang dilakukan oleh [3], [7], dan [12] Yang menambahkan variabel demografis, kepribadian, atau model RFM, adalah bahwa penelitian ini secara khusus menggunakan pendapatan dan skor belanja sebagai fitur utama, serta melakukan validasi jumlah *cluster* secara eksplisit melalui kombinasi metode *Elbow* dan *Silhouette Score* sebelum menetapkan model akhir, sehingga pemilihan  $k = 5$  berlandaskan pada dasar kuantitatif yang lebih jelas.

#### 6. Implikasi Manajerial

Hasil dari segmentasi ini memberikan sejumlah implikasi praktis bagi manajemen di sektor ritel dan *e-commerce*. Pertama, adanya segmen Pelanggan Premium yang, meskipun jumlahnya relatif kecil (19,5%), memiliki daya beli serta skor belanja tertinggi, menunjukkan perlunya peningkatan alokasi anggaran pemasaran untuk menjaga loyalitas kelompok ini, misalnya melalui program keanggotaan eksklusif. Kedua, segmen pelanggan impulsif yang memiliki pendapatan rendah tetapi skor belanja tinggi dapat menjadi sasaran promosi jangka pendek seperti penawaran diskon musiman, tetapi harus diimbangi dengan pendidikan mengenai pengelolaan keuangan untuk menjaga hubungan jangka panjang dengan pelanggan. Ketiga, segmen Pelanggan Hemat yang berpendapatan tinggi tetapi jarang berbelanja adalah peluang besar yang belum optimal, sehingga strategi pemasaran yang berbasis nilai dan edukasi tentang manfaat produk perlu ditingkatkan untuk mengubah kelompok ini menjadi pelanggan yang lebih aktif. Dengan memahami karakteristik setiap segmen secara kuantitatif, manajemen dapat mendistribusikan sumber daya pemasaran dengan lebih efisien dibandingkan dengan pendekatan pemasaran yang bersifat umum untuk semua pelanggan.

#### 4. KESIMPULAN

Algoritma K-Means Clustering berhasil digunakan untuk membagi pelanggan berdasarkan fitur pendapatan tahunan (*Annual Income*) dan skor belanja (*Spending Score*) dari *dataset* Data Segmentasi Pelanggan Mall. Berdasarkan hasil evaluasi menggunakan gabungan metode *Elbow* dan *Silhouette Score*, diperoleh bahwa jumlah *cluster* terbaik adalah  $k = 5$  dengan nilai *Silhouette Score* sebesar 0,5547, yang menunjukkan kualitas pengelompokan tergolong baik. Model tersebut menghasilkan lima segmen pelanggan, yaitu Pelanggan Standar (40,5%), Pelanggan Premium (19,5%), Pelanggan Impulsif (11,0%), Pelanggan Hemat (17,5%), dan Pelanggan Pasif (11,5%). Masing-masing segmen memiliki karakteristik yang berbeda berdasarkan tingkat pendapatan dan pola belanja, sehingga hasil pengelompokan ini dapat dijadikan dasar dalam menyusun strategi pemasaran yang lebih tepat sasaran dan efektif. Meskipun demikian, masih terdapat beberapa data yang berada pada area perbatasan antar-*cluster* sehingga berpotensi mengalami sedikit tumpang tindih dalam proses pengelompokan.

Berdasarkan hasil dan keterbatasan dari penelitian ini, beberapa saran untuk kajian selanjutnya antara lain: (1) menambahkan variabel lainnya seperti frekuensi pembelian, kategori produk, atau data RFM (Recency, Frequency, Monetary) untuk memperkaya profil segmentasi; (2) membandingkan kinerja algoritma K-Means dengan teknik analisis *cluster* lain seperti *DBSCAN*, *Hierarchical Clustering*, atau *Gaussian Mixture Model* untuk menemukan algoritma yang paling cocok dengan karakteristik data; (3) menerapkan penelitian ini pada data transaksi nyata dari suatu perusahaan ritel atau *e-commerce* untuk menguji validitas hasil segmentasi; serta (4) mengembangkan *dashboard* visualisasi interaktif sehingga hasil segmentasi dapat digunakan secara langsung oleh tim pemasaran dalam pengambilan keputusan.

#### 5. UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan bantuan, baik secara langsung maupun tidak langsung, sehingga artikel ini dapat diselesaikan dengan baik.

#### DAFTAR PUSTAKA

- [1] P. M. K. Mado and Hendry, "Implementasi Algoritma Clustering K-Means untuk," *J. Indones. Manaj. Inform. dan Komun.*, vol. 6, no. 3, pp. 1680–1686, 2025.
- [2] R. M. Juliyanto, F. Patria, H. D. Pamungkas, F. S. Nugraha, and S. Indonesia, "Penerapan Algoritma Clustering Untuk Segmentasi," no. November, pp. 117–127, 2024.
- [3] 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)," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4677.
- [4] H. Safitri, S. Putri Lenggo Geni, F. Merry, and M. Wati, "Penerapan K-Means Clustering untuk Segmentasi Konsumen E-Commerce Berdasarkan Pola Pembelian," *JUKI J. Komput. dan Inform.*, vol. 7, no. 1, pp. 89–99, 2025.
- [5] F. D. S. Alhamdani, A. A. Dianti, and Y. Azhar, "Segmentasi Pelanggan Berdasarkan Perilaku Penggunaan Kartu Kredit Menggunakan Metode K-Means Clustering," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 6, no. 2, pp. 70–77, 2021, doi: 10.14421/jiska.2021.6.2.70-77.
- [6] Sulistyowati, B. E. Ketherin, A. A. Arifiyanti, and A. Sodik, "Analisis Segmentasi Konsumen Menggunakan Algoritma K-Means Clustering Jurusan Sistem Informasi , Institut Teknologi Adhi Tama Surabaya," no. September, 2018.
- [7] M. Harahap, Y. Lubis, and Z. Situmorang, "Analisis Pemasaran Bisnis dengan Data Science : Segmentasi Kepribadian Pelanggan berdasarkan Algoritma K-Means Clustering," *Data Sci. Indones.*, vol. 1, no. 2, pp. 76–88, 2022, doi: 10.47709/dsi.v1i2.1348.
- [8] E. F. L. Awalina and W. I. Rahayu, "Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail," *J. Teknol. dan Inf.*, vol. 13, no. 2, pp. 122–137, 2023, doi: 10.34010/jati.v13i2.10090.
- [9] P. H. Suharti, A. S. Suryandari, and R. N. Amalia, "Analisis Kinerja Modul Pengendali Tekanan Udara Pct-14 Berbasis Plc Dengan Berbagai Metoda Tuning," *Sebatik*, vol. 26, no. 2, pp. 420–427, 2022, doi: 10.46984/sebatik.v26i2.2134.
- [10] C. H. Ardana, A. A. A. A. A. Khoyum, and M. Faisal, "Segmentasi Pelanggan Penjualan Online Menggunakan Metode K-means Clustering," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 9, no. 1, pp. 1–9, 2024, doi: 10.14421/jiska.2024.9.1.1-9.
- [11] R. Insan Pratama Siagian, E. Pratama, F. Azmi Lubis, S. Audy Priscilia, and F. Ramadhani, "Segmentasi Pelanggan Dengan Algoritma K-Means Untuk Strategi Pemasaran Yang Efektif," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 4, pp. 5615–5620, 2025, doi: 10.36040/jati.v9i4.13865.
- [12] I. Iqbal, N. Hidayat, D. P. Gevano, and A. P. R. Ilahi, "Segmentasi Pelanggan Menggunakan K-Means Clustering Berdasarkan Data Kepribadian dan Pola Konsumsi," *J. Tek. Inform.*, vol. 6, no. 5, pp. 3914–3924, 2025, doi: 10.52436/1.jutif.2025.6.5.5140.
- [13] A. Rahmawati, T., A. Hananto, and F. Nurapriani, "Analisis Segmentasi Pelanggan Menggunakan K-Means Clustering Untuk Optimalisasi Penjualan Sembako," *J. Ilm. Ilk. - Ilmu Komput. Inform.*, vol. 8, no. 2, pp. 233–243, 2025, doi: 10.47324/ilkominfo.v8i2.352.
- [14] H. Mukhtar, I. D. Pramaditya, W. S. Weisdiyanto, and S. H. Putra, "Jurnal Software Engineering and Information System ( SEIS ) ALGORITMA K-MEANS UNTUK PENGELOMPOKAN PERILAKU CUSTOMER," *J. Softw. Eng. Inf. Syst.*, vol. 4, no. 2, pp. 96–101, 2024.