

Penerapan *Clustering K-Means* untuk Kategorisasi Data Penjualan: Studi Kasus Cendana Shooping

Application of K-Means Clustering for Sales Data Categorization: A Case Study of Cendana Shooping

¹Yuliana P. Talan, ²Krisantus J. Tey Seran*, ³Dian Grace Ludji, ⁴Hevi Herlina Ullu

^{1,2,3,4}Program Studi Teknologi Informasi, Fakultas Pertanian, Sains dan Kesehatan, Universitas Timor

^{1,2,3,4}Jl. Raymundus Sau Fernandez, Kelurahan Sasi, Kecamatan Kota kefamenanu, Kabupaten TTU, NTT

*e-mail:krisantusteyseran@unimor.ac.id

(received: 9 July 2026, revised: 10 September 2026, accepted: 11 September 2026)

Abstrak

Penelitian ini menerapkan metode *K-Means Clustering* untuk mengelompokkan data transaksi penjualan pada Toko Online Cendana Shooping yang berlokasi di Kefamenanu, Kabupaten Timor Tengah Utara. Toko memanfaatkan media sosial sebagai sarana promosi, komunikasi, dan transaksi antara penjual dan pelanggan. Data yang dianalisis terdiri dari 1.035 transaksi penjualan selama periode Januari 2024 hingga Maret 2025 dengan menggunakan variabel harga barang, jumlah pembelian, dan total pembayaran. Penentuan jumlah kluster dilakukan menggunakan metode *Elbow* dan diperoleh empat kluster ($k = 4$). Proses pengelompokan berlangsung selama tujuh iterasi hingga mencapai kondisi stabil atau konvergen. Kategori tingkat penjualan ditentukan berdasarkan urutan nilai rata-rata *Total_Pembayaran* pada centroid akhir, yaitu Kluster 3 sebagai kelompok sangat laris dengan nilai rata-rata *Total_Pembayaran* Rp163.333,30 dan 90 transaksi, Kluster 2 sebagai kelompok laris dengan nilai rata-rata Rp119.098,40 dan 61 transaksi, Kluster 1 sebagai kelompok cukup laris dengan nilai rata-rata Rp77.817,57 dan 296 transaksi, serta Kluster 4 sebagai kelompok kurang laris dengan nilai rata-rata Rp41.938,78 dan 588 transaksi. Evaluasi menggunakan *Silhouette Coefficient* menghasilkan nilai rata-rata 0,6246 yang menunjukkan kualitas kluster pada kategori layak/sesuai. Hasil penelitian dapat menjadi dasar pendukung dalam pengelolaan persediaan, penyusunan strategi pemasaran, dan pengambilan keputusan bisnis.

Kata kunci: data mining, *E-Commerce*, *K-Means Clustering*, *silhouette coefficient*

Abstract

This study applies the *K-Means Clustering* method to group sales transaction data from Cendana Shooping Online Store, located in Kefamenanu, North Central Timor Regency. The store utilizes social media as a means of promotion, communication, and transactions between sellers and customers. The analyzed dataset consists of 1,035 sales transactions recorded from January 2024 to March 2025, using item price, purchase quantity, and total payment as the clustering variables. The optimal number of clusters was determined using the *Elbow Method*, resulting in four clusters ($k = 4$). The clustering process converged after seven iterations, indicating that stable cluster centroids had been reached. Sales performance categories were determined based on the ranking of the average *Total_Payment* values of the final centroids. Cluster 3 was categorized as the very high-sales group, with an average *Total_Payment* of IDR 163,333.30 and 90 transactions; Cluster 2 as the high-sales group, with an average of IDR 119,098.40 and 61 transactions; Cluster 1 as the moderately high-sales group, with an average of IDR 77,817.57 and 296 transactions; and Cluster 4 as the low-sales group, with an average of IDR 41,938.78 and 588 transactions. Evaluation using the *Silhouette Coefficient* yielded an average score of 0.6246, indicating an acceptable clustering quality. The findings can serve as a basis for inventory management, marketing strategy development, and business decision-making.

Keywords: data mining, *E-Commerce*, *K-Means Clustering*, *silhouette coefficient*.

1 Pendahuluan

Perdagangan *online* melalui dunia internet atau “*E-Commerce*” bukan lagi merupakan konsep yang baru berkembang di Indonesia [1]. Penerapan sistem ini dapat menguntungkan banyak pihak yang terlibat, baik dari sisi konsumen, penjual maupun penyedia platform ini. *E-Commerce* telah menjadi layanan populer bagi konsumen untuk melakukan berbagai jenis transaksi jual beli secara *online*[2], [3]. Beberapa tahun terakhir, industri *E-Commerce* telah mengalami pertumbuhan yang pesat di Indonesia terutama pada masa pandemi Covid-19 [4], [5]. Saat ini sudah banyak platform *E-Commerce* yang berkembang di Indonesia seperti, tokopedia, shopee, bukalapak. Selain menggunakan platform dari pihak ketiga yang sudah disebutkan, kegiatan *E-Commerce* juga dapat dilakukan melalui media sosial pribadi setiap orang seperti Facebook dan Whatsapp [6]. Kegiatan penjualan ini juga dilakukan oleh Toko *Online* Cendana Shopping yang berada di Kota Kefamenanu Kabupaten Timor Tengah Utara. Dalam memasarkan produk, toko ini melakukan promosi, komunikasi, dan transaksi menggunakan Facebook dan Whatsapp [7] [8]. Melalui dua platform ini, pelanggan dapat memperoleh informasi produk, melakukan pemesanan, serta berkomunikasi langsung dengan penjual secara langsung atau cepat tanpa harus melibatkan pihak ketiga. Toko ini melayani penjualan barang seperti pakaian, aksesoris, dan alas kaki seperti sandal dan sepatu. Dengan tingginya jumlah barang transaksi harian yang sudah dilakukan, tentunya memerlukan analisis yang mendalam mengenai ketersediaan barang yang paling banyak dipesan [9], serta tempat penyimpanan barang berdasarkan tingkat penjualan [10].

Melihat permasalahan yang dihadapi Toko *Online* Cendana Shopping, penelitian ini dilakukan teknik data mining yang mengimplementasikan algoritma *machine learning* untuk pengelompokan yakni *K-Means* [11]. Algoritma *K-Means* merupakan metode *Clustering* non-hierarki [12], yang digunakan untuk mengelompokkan sekumpulan data ke dalam beberapa kluster berdasarkan tingkat kemiripan karakteristik yang dimiliki [13]. Proses kerja algoritma ini dimulai dengan menentukan jumlah kluster (*K*) yang akan dibentuk, kemudian memilih titik pusat awal (*centroid*) untuk setiap kluster. Selanjutnya, setiap data dihitung jaraknya terhadap seluruh *centroid* menggunakan metode pengukuran jarak, yang umumnya menggunakan *Euclidean Distance*, sehingga setiap data dapat ditempatkan pada kluster dengan *centroid* terdekat. Setelah proses pengelompokan selesai, *centroid* baru dihitung berdasarkan nilai rata-rata seluruh data yang berada pada masing-masing kluster. Tahapan tersebut dilakukan secara berulang hingga tidak terjadi lagi perpindahan anggota kluster atau posisi *centroid* telah stabil (*konvergen*). Melalui proses iteratif tersebut, algoritma *K-Means* mampu menghasilkan kelompok data yang memiliki tingkat kesamaan tinggi di dalam satu kluster dan perbedaan yang jelas dengan kluster lainnya [14]. Data dalam penelitian ini sebanyak 1.035 data transaksi penjualan yang diperoleh dari Toko *Online* Cendana Shopping selama periode Januari 2024 hingga Maret 2025. Data yang dianalisis meliputi informasi mengenai harga barang, jumlah pembelian, dan total pembayaran pada setiap transaksi. Seluruh data tersebut dimanfaatkan sebagai objek penelitian untuk proses pengelompokan menggunakan algoritma *K-Means Clustering*, sehingga dapat diketahui pola penjualan produk berdasarkan tingkat kemiripan karakteristik yang dimiliki [15].

Penerapan teknik data mining dalam penelitian ini diawali dengan tahap, persiapan data sebelum dilakukan analisis [16]. Tahap ini bertujuan untuk memastikan bahwa data yang digunakan telah memenuhi kriteria kelayakan, seperti tidak terdapat data yang hilang maupun duplikasi data. Selanjutnya dilakukan proses *Clustering* menggunakan algoritma *K-Means* [17], yang dimulai dengan penentuan jumlah kluster, pemilihan *centroid* awal, perhitungan jarak menggunakan *Euclidean Distance*, pengelompokan data ke dalam kluster terdekat, serta pembaruan *centroid* secara berulang hingga mencapai kondisi stabil atau *konvergen*. Setelah proses klusterisasi selesai, kualitas hasil pengelompokan dievaluasi menggunakan metode *Silhouette Coefficient* [18], untuk mengetahui tingkat ketepatan dan kualitas kluster yang terbentuk. Penerapan algoritma *K-Means Clustering* dalam menganalisis data penjualan barang pada Toko *Online* Cendana Shopping sehingga menghasilkan pengelompokan produk berdasarkan tingkat penjualannya. Hasil yang diperoleh diharapkan dapat memberikan informasi mengenai kategori produk yang memiliki tingkat penjualan tinggi maupun rendah sehingga dapat dimanfaatkan sebagai bahan pertimbangan dalam pengelolaan persediaan barang, penyusunan strategi promosi, serta pengambilan keputusan bisnis yang lebih efektif. Selain itu, penelitian ini juga bertujuan untuk mengukur kualitas hasil pengelompokan melalui metode

Silhouette Coefficient, sehingga tingkat validitas kluster yang dihasilkan dapat diketahui secara lebih objektif.

2 Tinjauan Literatur

Penerapan algoritma *K-Means* untuk menganalisis dan mengelompokkan data penjualan telah banyak dilakukan dalam berbagai penelitian sebelumnya. Penggunaan algoritma *K-Means* dalam bidang penjualan umumnya bertujuan untuk mengelompokkan produk atau data transaksi berdasarkan karakteristik tertentu sehingga dapat memberikan informasi yang mendukung pengambilan keputusan dalam pengelolaan penjualan.

Penelitian yang dilakukan oleh Afiasari, Suarna, dan Rahangningsih [19]. menerapkan algoritma *K-Means* pada data transaksi penjualan Toko Bill Lights. Penelitian tersebut menggunakan dataset yang terdiri dari 160 produk dan menghasilkan tiga kelompok, yaitu 88 produk dalam kategori terjual, 26 produk dalam kategori tidak terjual, dan 46 produk dalam kategori kurang terjual. Hasil pengelompokan tersebut digunakan untuk membantu toko dalam menentukan strategi penjualan berdasarkan karakteristik tingkat penjualan produk. Penelitian tersebut menunjukkan bahwa algoritma *K-Means* dapat digunakan untuk mengelompokkan data penjualan berdasarkan kemiripan karakteristik sehingga menghasilkan informasi yang dapat mendukung strategi bisnis.

Penelitian serupa dilakukan oleh Rasyid dkk [20]. Dengan menerapkan algoritma *K-Means Clustering* untuk mengelompokkan penjualan produk pada online shoop toko gizi. Penelitian tersebut menggunakan data penjualan periode april sampai mei 2022 dan bertujuan mengidentifikasi produk yang sangat diminati, cukup diminati, dan kurang diminati. Hasil penelitian terhadap 41 jenis data penjualan menghasilkan tiga kluster, yaitu sangat diminati, cukup diminati, dan kurang diminati. Hasil tersebut menunjukkan bahwa *K-Means* dapat membantu toko dalam mengetahui tingkat minat konsumen terhadap produk yang dijual. Meskipun demikian, penelitian tersebut menggunakan data pada periode relatif singkat sehingga belum menggambarkan pola penjualan dalam periode yang lebih panjang. Selain itu, hasil penelitian belum memberikan analisis yang mendalam mengenai faktor yang menyebabkan karakteristik pada masing-masing kelompok serta rekomendasi strategi yang lebih spesifik berdasarkan hasil *Clustering*

Selanjutnya, Adelina Bui dan Bahtiar [21] menerapkan algoritma *K-Means Clustering* untuk mengelompokkan transaksi penjualan barang di Toko Arino. Penelitian tersebut bertujuan meningkatkan efisiensi manajemen toko retail melalui pengelompokan data transaksi penjualan. Hasil penelitian menunjukkan bahwa penerapan *K-Means* mampu memberikan informasi mengenai pola penjualan dan preferensi pelanggan sehingga dapat mendukung pengambilan keputusan berbasis data serta meningkatkan efisiensi operasional toko. Namun, penelitian tersebut memiliki keterbatasan pada data dan atribut yang digunakan. Data yang dianalisis hanya mencakup transaksi dan atribut tertentu sehingga belum mencerminkan seluruh aspek yang mungkin relevan dalam menggambarkan karakteristik penjualan. Penelitian tersebut juga menyarankan penggunaan metode analisis lain pada penelitian selanjutnya untuk menggali potensi data dan variabel secara lebih dalam.

Penelitian lainnya dilakukan oleh dermawan dkk [22]. Melalui penerapan algoritma *K-Means* pada data penjualan Toko Daun Indah Shopee. Penelitian tersebut bertujuan menganalisis data penjualan untuk menentukan kluster terbaik berdasarkan beberapa parameter evaluasi, termasuk Davies-Bouldin-Index, jumlah iterasi, dan *Measure Type*. Hasil penelitian menghasilkan dua kluster, yaitu kelompok produk yang banyak diminati dan kelompok produk yang kurang diminati. Hasil tersebut menunjukkan bahwa *K-Means* dapat dimanfaatkan untuk membantu menentukan strategi pemasaran berdasarkan tingkat minat terhadap produk. Namun, penelitian tersebut menggunakan data transaksi dalam periode November 2022 sampai Mei 2023 sehingga masih terbatas dalam menggambarkan pola penjualan jangka panjang. Selain itu, penelitian tersebut menggunakan pendekatan evaluasi tertentu sehingga masih memungkinkan dilakukan perbandingan dengan metode evaluasi *Clustering* lainnya.

Penelitian yang dilakukan oleh Amalina dkk [23]. Menerapkan metode *K-Means Clustering* dalam pengelompokan penjualan produk frozen food. Penelitian tersebut bertujuan menganalisis minat konsumen terhadap berbagai produk berdasarkan tingkat kepopulerannya sehingga membantu mengidentifikasi produk dengan tingkat penjualan tinggi dan rendah serta mendukung penyusunan strategi pemasaran. Namun, penelitian tersebut memiliki beberapa keterbatasan, antara lain jumlah

klaster ditentukan secara manual, sensitifitas *K-Means* terhadap outlier, asumsi bentuk klaster, serta tidak dipertimbangkan faktor eksternal seperti tren pasar dan kegiatan promosi. Selain itu, data yang digunakan hanya terdiri dari 45 produk sehingga ruang generalisasi hasil penelitian masih terbatas.

Berdasarkan beberapa penelitian terdahulu tersebut dapat dilihat bahwa algoritma *K-Means* telah banyak digunakan untuk mengelompokkan data penjualan berdasarkan tingkat penjualan, minat konsumen, maupun karakteristik transaksi. Hasil penelitian terdahulu menunjukkan bahwa penerapan *K-Means* dapat memberikan informasi yang bermanfaat untuk mendukung pengelolaan stok, strategi pemasaran, dan pengambilan keputusan berbasis data. Namun, masih terdapat beberapa keterbatasan, seperti penggunaan data dalam periode yang relatif singkat, jumlah data yang terbatas, keterbatasan atribut yang digunakan, serta belum optimalnya analisis terhadap karakteristik masing-masing klaster.

Berdasarkan kondisi tersebut, penelitian ini dilakukan pada Toko Online Cendana Shopping dengan menggunakan data transaksi penjualan sebanyak 1.035 transaksi pada periode Januari 2024 sampai Maret 2025. Berbeda dengan beberapa penelitian terdahulu yang berfokus pada produk atau periode penjualan tertentu, penelitian ini memfokuskan pengelompokan pada data transaksi penjualan yang diperoleh dari aktivitas penjualan toko online. Variabel yang digunakan dalam proses *Clustering* adalah harga barang, jumlah pembelian, dan total pembayaran. Ketiga variabel tersebut digunakan untuk menggambarkan karakteristik nilai transaksi yang terjadi pada Cendana Shopping.

Penelitian ini juga menentukan jumlah klaster menggunakan metode *Elbow* sehingga jumlah klaster tidak ditentukan sendiri. Selanjutnya, kualitas hasil pengelompokan dievaluasi menggunakan *Silhouette Coefficient* dan divisualisasikan menggunakan *Principal Component Analysis* (PCA). Pendekatan tersebut digunakan untuk memberikan gambaran mengenai kualitas dan persebaran hasil *Clustering* secara lebih jelas. Dengan demikian, posisi penelitian ini terletak pada penerapan algoritma *K-Means* untuk mengelompokkan data transaksi penjualan Toko Online Cendana Shopping berdasarkan harga barang, jumlah pembelian, dan total pembayaran. Hasil pengelompokan menghasilkan empat kelompok yang kemudian diinterpretasikan sebagai kategori sangat laris, laris, cukup laris, dan kurang laris berdasarkan karakteristik centroid masing-masing klaster. Hasil tersebut diharapkan dapat memberikan informasi yang lebih terarah bagi pengelola toko dalam melakukan evaluasi penjualan, mengelola persediaan, dan menentukan strategi pemasaran.

Perbedaan penelitian ini terletak pada penerapan *K-Means* terhadap data transaksi penjualan Cendana Shopping dengan periode pengamatan yang lebih panjang dan jumlah transaksi yang lebih besar dibandingkan beberapa penelitian terdahulu yang digunakan sebagai rujukan. Selain itu, penelitian ini tidak hanya menghasilkan kelompok penjualan, tetapi juga melakukan penentuan jumlah klaster menggunakan metode *Elbow*, evaluasi menggunakan *Silhouette Coefficient*, serta visualisasi hasil *Clustering* menggunakan PCA. Dengan pendekatan tersebut, penelitian ini memberikan gambaran mengenai karakteristik setiap kelompok transaksi yang dapat digunakan sebagai dasar informasi dalam mendukung pengambilan keputusan pada Toko Online Cendana Shopping.

3. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan menerapkan algoritma *K-Means Clustering* untuk menganalisis data penjualan pada Toko Online Cendana Shopping. Data yang digunakan berupa 1.035 transaksi penjualan yang diperoleh selama periode Januari 2024 hingga Maret 2025 melalui proses observasi dan wawancara dengan pemilik toko. Variabel yang digunakan dalam proses analisis meliputi harga barang, jumlah pembelian, dan total pembayaran. Tahapan penelitian ini dilakukan secara sistematis yang terdiri dari identifikasi permasalahan, studi literatur, pengumpulan data, persiapan data, implementasi algoritma *K-Means*, evaluasi hasil klaster, dan penarikan kesimpulan. Tahapan penelitian tersebut digunakan sebagai alur dalam proses pengolahan dan analisis data transaksi penjualan Toko Online Cendana Shopping.



Gambar 1 Tahapan penelitian

1. **Identifikasi Masalah**
Identifikasi masalah dilakukan untuk menentukan permasalahan utama, yaitu bagaimana mengelompokkan transaksi penjualan berdasarkan karakteristik harga barang, jumlah pembelian, dan total pembayaran serta bagaimana hasil pengelompokan tersebut dapat dimanfaatkan untuk mengenali tingkat penjualan sebagai bahan pendukung strategi pemasaran dan pengelolaan persediaan.
2. **Studi Literatur**
Tahap studi literatur dilakukan dengan mengumpulkan berbagai referensi yang berkaitan dengan konsep data mining, algoritma *K-Means Clustering*, metode evaluasi *Silhouette Coefficient*, serta penelitian terdahulu yang relevan. Literatur diperoleh dari buku, jurnal ilmiah, artikel, dan sumber terpercaya lainnya yang digunakan sebagai landasan teori dalam pelaksanaan penelitian.
3. **Pengumpulan Data**
Data penelitian dikumpulkan melalui observasi dan wawancara secara langsung dengan pemilik Toko Online Cendana Shoooping. Dalam kegiatan penjualannya, toko memanfaatkan Facebook dan WhatsApp sebagai sarana promosi, komunikasi, dan transaksi. Informasi produk dipromosikan melalui unggahan gambar dan deskripsi produk pada media sosial. Pelanggan yang berminat menghubungi penjual untuk melakukan pemesanan dan konfirmasi transaksi. Pembayaran dilakukan melalui metode tunai atau non-tunai/digital sesuai kesepakatan antara penjual dan pelanggan. Data yang dikumpulkan berupa catatan transaksi penjualan pada periode

Januari 2024 hingga Maret 2025. Objek yang dianalisis adalah transaksi penjualan, bukan produk secara individual.

4. Persiapan Data

Tahap persiapan data dilakukan untuk memastikan data transaksi dapat digunakan dalam proses *Clustering*. Persiapan data meliputi pemeriksaan struktur dan kelengkapan data, pemeriksaan nilai yang hilang, pemeriksaan duplikasi, pengecekan kesesuaian format dan tipe data, serta pemilihan atribut yang digunakan dalam proses *Clustering*. Hasil pemeriksaan menunjukkan bahwa data tidak memiliki nilai yang hilang maupun data duplikat. Data yang digunakan terdiri dari tiga variabel numerik, yaitu harga barang, jumlah pembelian, dan total pembayaran. Harga barang digunakan untuk menggambarkan nilai per unit barang, jumlah pembelian menggambarkan kuantitas barang dalam transaksi, sedangkan total pembayaran menggambarkan nilai keseluruhan transaksi. Pada penelitian ini, proses *Clustering* menggunakan nilai asli ketiga variabel tanpa normalisasi atau standarisasi. Oleh karena itu, perhitungan jarak Euclidean dilakukan pada skala asli data. Perbedaan rentang nilai antarvariabel perlu diperhatikan karena variabel harga dan total pembayaran memiliki rentang numerik yang jauh lebih besar daripada jumlah pembelian, sehingga memiliki pengaruh lebih besar terhadap perhitungan jarak. Kondisi tersebut menjadi bagian dari keterbatasan metode pada penelitian ini dan perlu dipertimbangkan dalam interpretasi hasil. Tujuan dilakukan tahapan ini yaitu pemeriksaan terhadap kualitas data untuk memastikan bahwa data yang digunakan telah lengkap, konsisten, serta tidak mengandung data ganda maupun data yang hilang. Dengan kondisi data yang telah memenuhi syarat, proses *Clustering* dapat dilakukan secara optimal [24].

5. Implementasi Algoritma *K-Means Clustering*

Pada tahap ini, algoritma *K-Means* diterapkan untuk mengelompokkan data transaksi berdasarkan kemiripan karakteristik harga barang, jumlah pembelian, dan total pembayaran. *K-Means* merupakan salah satu metode *Clustering* yang banyak digunakan untuk mengelompokkan data berdasarkan kedekatan karakteristik antar objek. Proses *Clustering* dilakukan melalui beberapa tahapan, yaitu menentukan jumlah kluster, menentukan centroid awal, menghitung jarak euclidean, menentukan keanggotaan kluster, menghitung centroid baru, dan melakukan iterasi hingga diperoleh kondisi konvergen.

6. Evaluasi Hasil Kluster

Setelah dilakukan proses *Clustering*, hasil pengelompokan dievaluasi menggunakan Silhouette Coefficient. Evaluasi tersebut dilakukan untuk mengetahui seberapa baik pemisahan antar kluster yang terbentuk. Jika nilai Silhouette mendekati 1 artinya bahwa data semakin sesuai dengan kluster tersebut berada.

Dalam algoritma *K-Means* dilakukan proses pengelompokan data berdasarkan kemiripan karakteristik. Proses *Clustering* dilakukan dengan menghitung jarak antara data dan centroid menggunakan jarak Euclidean. Data dengan jarak terdekat akan dimasukkan ke dalam kluster, kemudian centroid diperbarui dan proses diulang hingga mencapai kondisi konvergen.

rumus untuk menghitung centroid:

$$sc_i = \frac{1}{M} \sum_{j=1}^M x_j \quad \dots\dots\dots(1)$$

Dimana:

sc_i = nilai rata-rata dari data ke-i

M = Jumlah total elemen (misalnya jumlah data atau fitur yang dijumlahkan)

x_j = nilai ke-j yang dijumlahkan

Hitung jarak titik terdekat (*Euclidean Distance*):

$$D(x_2, x_1) = \sqrt{\sum_{j=1}^P |x_{2j} - x_{1j}|^2} \quad \dots\dots\dots(2)$$

Dimana :

$D(x_2, x_1)$ = Notasi untuk jarak Euclidean kuadrat antar x_2 dan vector x_1

$\sum$ (Sigma = Operasi penjumlahan $j = 1$ hingga p = Indeks penjumlahan

$|x_{2j} - x_{1j}|$ = Selisih absolut antara komponen ke- j dari vector x_2 dan komponen ke- j dari vector x_1 .

4. Hasil dan Pembahasan

Data yang dianalisis sebanyak 1.035 data transaksi penjualan Toko Online Cendana Shooping yang terdiri atas variabel harga barang, jumlah pembelian, dan total pembayaran. Data tersebut diproses menggunakan algoritma *K-Means Clustering* melalui tahapan penentuan jumlah kluster, perhitungan jarak menggunakan *Euclidean Distance*, pengelompokan data berdasarkan jarak terdekat, serta pembaruan centroid secara berulang hingga mencapai kondisi konvergen pada iterasi ke-7. Hasil implementasi menunjukkan bahwa seluruh data berhasil dikelompokkan ke dalam empat kluster yang memiliki karakteristik penjualan berbeda, sehingga dapat digunakan sebagai dasar dalam mengidentifikasi kategori produk berdasarkan tingkat penjualannya. Hasil pengelompokan dalam dilihat pada gambar.

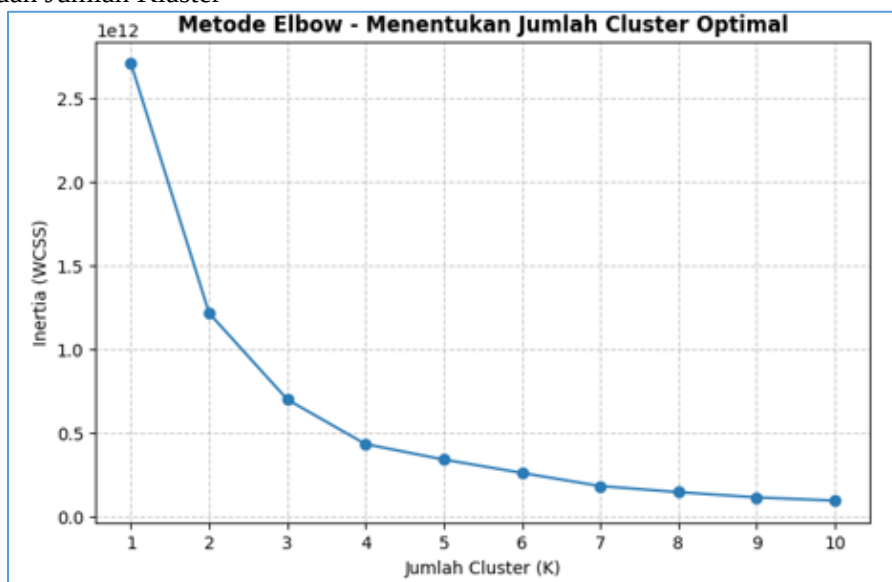
Sebelum masuk ke tahap perhitungan, berikut ditampilkan sebagian data transaksi penjualan yang menjadi data awal penelitian, yaitu data pertama dan data terakhir sebagai representasi dari keseluruhan 1.035 data transaksi yang digunakan.

Tabel 1 Data transaksi penjualan (data awal)

No	Kode Pembeli	Jenis Barang	Warna	Harga	Jumlah	Total Pembayaran
1	P001	Kemeja	Hitam	25000	1	25000
2	P002	Kemeja	Pink	25000	1	25000
3	P003	Blouse	Putih	25000	1	25000
4	P004	Kemeja	Hitam	35000	1	35000
5	P005	Dress	Hitam	35000	1	35000
6	P006	Tas	Hitam	35000	1	35000
7	P007	Baju Kaos	Hitam	35000	1	35000
8	P008	Setelan	Pink	65000	1	65000
9	P009	Baju Kaos	Hitam	35000	1	35000
10	P010	Dress	Hitam	45000	1	45000
...	...	...	...	...	...	...
1031	P753	Setelan	Mix	75000	3	225000
1032	-	Kemeja	m	45000	1	45000
1033	-	Baju Kaos	Pink	40000	1	40000
1034	-	Baju Kaos	Custo	60000	1	60000
1035	-	Sweater	m	60000	1	60000
	-	Hoodie	Abu-abu	60000	1	60000

Berdasarkan data awal tersebut, proses klusterisasi dilanjutkan dengan tahap perhitungan sebagai berikut.

a. Penentuan Jumlah Kluster



Gambar 2 Grafik Elbow

Berdasarkan Grafik *Elbow* diatas menunjukan bahwa penentuan jumlah kluster (K) yang optimal dilakukan menggunakan metode *Elbow*. Metode ini mengukur *Inertia* (*Within Cluster Sum of Squares* - *WCSS*) yang didefinisikan sebagai jumlah kuadrat jarak setiap titik data ke *centroid* kluster terdekatnya. Nilai K selanjutnya memberikan penurunan *WCSS* yang marjinal. Dengan demikian K = 4 dipilih karena mewakili titik terakhir pada kurva sebelum laju penurunan *inertia* melambat secara signifikan.

b. Proses *Clustering*

Setelah menentukan jumlah kluster optimal dilanjutkan dengan proses *K-Means*, ada 4 tahap dalam proses *K-Means Clustering* yaitu:

1) Menentukan titik centroid awal

Centroid awal dipilih sebagai titik pusat awal masing-masing kluster berdasarkan data transaksi yang mewakili variasi nilai harga barang, jumlah pembelian, dan total pembayaran.

Tabel 2 Centroid awal

Centroid Baru	V1	V2	V3
C1	75000	1	75000
C2	40000	2	80000
C3	80000	1	80000
C4	35000	1	35000

2) Menghitung jarak data ke centroid

Setelah centroid awal ditentukan, seluruh data transaksi dihitung jaraknya terhadap masing-masing centroid menggunakan *Microsoft Excel*. Selanjutnya setiap data ditempatkan pada kluster yang memiliki jarak terkecil. Hasil perhitungan iterasi pertama ditunjukkan pada Tabel 3.

Tabel 3 Jarak iterasi 1

No	C1	C2	C3	C4	Jarak Min	Kluster
1	70710,68	57008,77	77781,75	14142,14	14142,14	4
2	70710,68	57008,77	77781,75	14142,14	14142,14	4
3	70710,68	57008,77	77781,75	14142,14	14142,14	4
4	56568,54	45276,93	63639,61	0	0	4
5	56568,54	45276,93	63639,61	0	0	4

6	56568,54	45276,93	63639,61	0	0	4
7	56568,54	45276,93	63639,61	0	0	4
8	70710,68	57008,77	77781,75	14142,14	14142,14	4
9	14142,14	29154,76	21213,2	42426,41	14142,14	1
10	56568,54	45276,93	63639,61	0	0	4
...	...	...	...	...	...	...
1031	150000	149164	145086	194164	145086	3
1032	42426,4	35355,3	49497,5	14142,1	14142,1	4
1033	49497,5	40000	56568,5	7071,07	7071,07	4
1034	21213,2	28284,3	28284,3	35355,3	21213,2	1
1035	21213,2	28284,3	28284,3	35355,3	21213,2	1

3) Memperbarui centroid

Setelah data dikelompokkan kedalam klaster berdasarkan jarak minimum, Langkah selanjutnya adalah memperbarui posisi centroid untuk setiap klaster. Pada tahap ini, cara perhitungan centroid baru dilakukan dengan mengambil nilai rata-rata dari seluruh data yang berada dalam satu klaster untuk variabel yang digunakan.

Tabel 4 Centroid baru sesudah iterasi 1

Centroid Baru	V1	V2	V3
C1	66752,48	1	66752,48
C2	44140,63	2,515625	110390,6
C3	101794,9	1,188034	117692,3
C4	40224,3	1,003738	40317,76

Pada centroid baru mengalami perubahan, maka itu proses pengelompokan masih berlangsung pada iterasi berikut dengan menggunakan centroid baru tersebut.

Tabel 5 Jarak iterasi 2

No	C1	C2	C3	C4	Jarak Min	Cluster
1	59046,92	87509,56	120371,6	21596,6	21596,6	4
2	59046,92	87509,56	120371,6	21596,6	21596,6	4
3	59046,92	87509,56	120371,6	21596,6	21596,6	4
4	44904,78	75942,72	106299,4	7454,652	7454,652	4
5	44904,78	75942,72	106299,4	7454,652	7454,652	4
6	44904,78	75942,72	106299,4	7454,652	7454,652	4
7	44904,78	75942,72	106299,4	7454,652	7454,652	4
8	59046,92	87509,56	120371,6	21596,6	21596,6	4
9	2478,374	49954,2	64267,74	34972,11	2478,374	1
10	44904,78	75942,72	106299,4	7454,652	7454,652	4
...	...	...	...	...	...	...
1031	158462,3	118691,2	110602,5	187927,9	110602,5	3
1032	30762,65	65396,27	92248,73	6688,103	6688,103	4
1033	37833,71	70512,3	99270,85	388,9468	388,9468	4
1034	9549,442	52827,41	71240,53	27901,06	9549,442	1
1035	9549,442	52827,41	71240,53	27901,06	9549,442	1

Hasil iterasi kedua menunjukkan adanya perubahan nilai jarak serta beberapa perpindahan anggota klaster dibandingkan iterasi sebelumnya. Hal tersebut menandakan bahwa centroid belum stabil sehingga proses iterasi masih perlu dilanjutkan. Centroid kembali diperbarui berdasarkan hasil pengelompokan pada iterasi kedua.

Tabel 6 Centroid baru sesudah iterasi 2

Centroid baru	V1	V2	V3
C1	70265,23	1,021505	71125,45
C2	49000	2,377778	114333,3
C3	109816	1,355828	136257,7
C4	40036,5	1,031022	40985,4

Nilai centroid pada iterasi ketiga mengalami perubahan dibandingkan iterasi sebelumnya. Perubahan tersebut menunjukkan bahwa distribusi data pada masing-masing klaster masih mengalami penyesuaian. Selanjutnya dilakukan kembali proses perhitungan jarak terhadap centroid baru.

Tabel 7 Jarak iterasi 3

No	C1	C2	C3	C4	Jarak Min	Klaster
1	64625,83	92501,05	139900	21946	21946,05	4
2	64625,83	92501,05	139900	21946	21946,05	4
3	64625,83	92501,05	139900	21946	21946,05	4
4	50484,5	80559,16	125899	7822,5	7822,488	4
5	50484,5	80559,16	125899	7822,5	7822,488	4
6	50484,5	80559,16	125899	7822,5	7822,488	4
7	50484,5	80559,16	125899	7822,5	7822,488	4
8	64625,83	92501,05	139900	21946	21946,05	4
9	8077,363	51863,07	84179,1	34639	8077,363	1
10	50484,5	80559,16	125899	7822,5	7822,488	4
...	...	...	...	...	...	...
1031	153947,4	113679,9	95327,6	187307	95327,6	3
1032	36343,79	69448,62	111933	6383,8	6383,837	4
1033	43414,03	74876,19	118911	986,08	986,0771	4
1034	15137,72	55435,65	91087,1	27570	15137,72	1
1035	15137,72	55435,65	91087,1	27570	15137,72	1

Iterasi ketiga mengalami perubahan dibandingkan iterasi sebelumnya. Perubahan tersebut menunjukkan bahwa distribusi data pada masing-masing klaster masih mengalami penyesuaian. Selanjutnya dilakukan kembali proses perhitungan jarak terhadap centroid baru.

Tabel 8 Centroid baru sesudah iterasi 3

Centroid baru	V1	V2	V3
C1	74959,56	1,022059	75841,91
C2	52254,9	2,215686	113039,2
C3	116572,6	1,467742	150040,3
C4	41054,42	1,028912	41938,78

Nilai centroid pada iterasi keempat menunjukkan perubahan yang semakin kecil dibandingkan iterasi sebelumnya, menandakan posisi pusat klaster mulai stabil.

Tabel 9 Jarak iterasi 4

NO	C1	C2	C3	C4	Jarak Min	CLUSTER
1	71280,13	92161,45	154985,9	23338,09	23338,09	4
2	71280,13	92161,45	154985,9	23338,09	23338,09	4
3	71280,13	92161,45	154985,9	23338,09	23338,09	4
4	57138,67	79924,03	141026,1	9208,834	9208,834	4
5	57138,67	79924,03	141026,1	9208,834	9208,834	4
6	57138,67	79924,03	141026,1	9208,834	9208,834	4

7	57138,67	79924,03	141026,1	9208,834	9208,834	4
8	71280,13	92161,45	154985,9	23338,09	23338,09	4
9	14722,09	49701,14	99456,46	33244,71	14722,09	1
10	57138,67	79924,03	141026,1	9208,834	9208,834	4
...	...	...	...	...	...	...
1031	149158,1	114247,8	85716	186181,9	85716	3
1032	42997,66	68424,91	127106,7	4993,865	4993,865	4
1033	50068,09	74060,18	134060,6	2206,956	2206,956	4
1034	21788,86	53601,73	106337,7	26175,23	21788,86	1
1035	21788,86	53601,73	106337,7	26175,23	21788,86	1

Hasil pengelompokan pada iterasi keempat memperlihatkan bahwa sebagian besar data telah berada pada kluster yang tetap. Namun demikian masih terdapat beberapa perubahan sehingga iterasi berikutnya tetap dilakukan.

Tabel 10 Centroid baru sesudah iterasi 4

Centroid baru	V1	V2	V3
C1	77006,76	1,02027	77817,57
C2	52254,9	2,215686	113039,2
C3	120500	1,58	162000
C4	41054,42	1,028912	41938,78

Centroid baru after iterasi keempat kembali mengalami penyesuaian berdasarkan rata-rata anggota masing-masing kluster. Perubahan yang terjadi relatif kecil dibandingkan iterasi sebelumnya.

Tabel 11 Jarak iterasi 5

No	C1	C2	C3	C4	Jarak Min	Cluster
1	74124,21	92161,45	167000,7	23338,09	23338,09	4
2	74124,21	92161,45	167000,7	23338,09	23338,09	4
3	74124,21	92161,45	167000,7	23338,09	23338,09	4
4	59982,6	79924,03	153098,8	9208,834	9208,834	4
5	59982,6	79924,03	153098,8	9208,834	9208,834	4
6	59982,6	79924,03	153098,8	9208,834	9208,834	4
7	59982,6	79924,03	153098,8	9208,834	9208,834	4
8	74124,21	92161,45	167000,7	23338,09	23338,09	4
9	17562,81	49701,14	111755,3	33244,71	17562,81	1
10	59982,6	79924,03	153098,8	9208,834	9208,834	4
...	...	...	...	...	...	...
1031	59982,6	79924,03	153098,8	9208,834	9208,834	4
1032	147196,1	114247,8	77712,61	186181,9	77712,61	3
1033	45841,3	68424,91	139245,3	4993,865	4993,865	4
1034	52911,89	74060,18	146165,1	2206,956	2206,956	4
1035	24631,19	53601,73	118592,8	26175,23	24631,19	1

Hasil iterasi kelima menunjukkan jumlah perpindahan anggota kluster semakin sedikit. Hal ini menunjukkan bahwa proses *Clustering* mulai mendekati kondisi stabil.

Tabel 12 Centroid baru sesudah iterasi 5

Centroid baru	V1	V2	V3
C1	77772,41	1	77772,41
C2	54552,24	2,164179	115597
C3	125555,6	1,533333	163333,3
C4	41054,42	1,028912	41938,78

Nilai centroid baru setelah iterasi kelima mengalami perubahan yang sangat kecil dibandingkan iterasi kelima sehingga menunjukkan bahwa posisi pusat klaster hampir mencapai kondisi konvergen.

Tabel 13 Jarak Iterasi 6

No	C1	C2	C3	C4	Jarak Min	Cluster
1	74631,46	95295,09	171019,1	23338,09	23338,09	4
2	74631,46	95295,09	171019,1	23338,09	23338,09	4
3	74631,46	95295,09	171019,1	23338,09	23338,09	4
4	60489,33	82934,73	157066,1	9208,834	9208,834	4
5	60489,33	82934,73	157066,1	9208,834	9208,834	4
6	60489,33	82934,73	157066,1	9208,834	9208,834	4
7	60489,33	82934,73	157066,1	9208,834	9208,834	4
8	74631,46	95295,09	171019,1	23338,09	23338,09	4
9	18062,92	51664,43	115483,4	33244,71	18062,92	1
10	60489,33	82934,73	157066,1	9208,834	9208,834	4
...	...	...	...	...	...	...
1031	147253,7	111297,5	79741,09	186181,9	79741,09	3
1032	46347,19	71240,32	143150,2	4993,865	4993,865	4
1033	53418,26	76984,91	150102,8	2206,956	2206,956	4
1034	25133,99	55863,28	122373,6	26175,23	25133,99	1
1035	25133,99	55863,28	122373,6	26175,23	25133,99	1

Pada iterasi keenam hanya terjadi sedikit perubahan hasil pengelompokan. Oleh karena itu proses iterasi dilanjutkan satu kali lagi untuk memastikan bahwa seluruh centroid telah stabil.

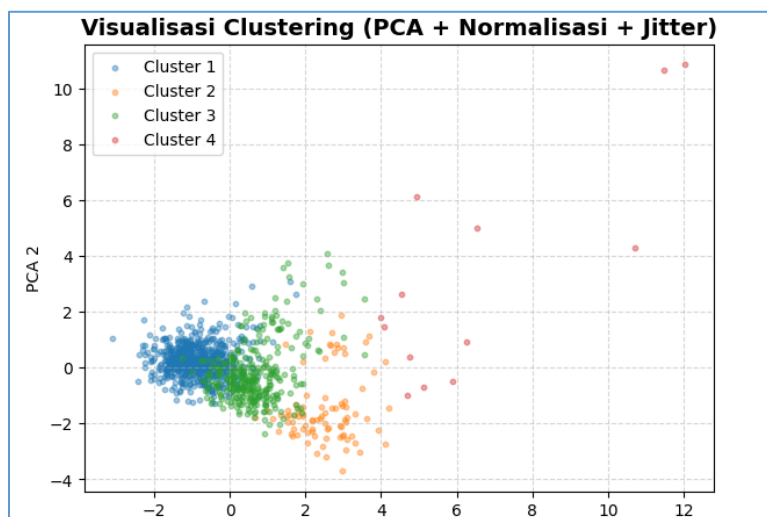
Tabel 14 Centroid Baru Sesudah Iterasi 6

Centroid baru	V1	V2	V3
C1	77006,76	1,02027	77817,57
C2	55983,61	2,180328	119098,4
C3	125555,6	1,533333	163333,3
C4	41054,42	1,028912	41938,78

Nilai centroid baru after iterasi 6 tidak mengalami perubahan yang signifikan dibandingkan iterasi sebelumnya sehingga menunjukkan bahwa pusat klaster telah stabil.

Tabel 15 Jarak Iterasi 7

No	C1	C2	C3	C4	Jarak Min	Cluster
1	74631,46	95295,09	171019,1	23338,09	23338,09	4
2	74631,46	95295,09	171019,1	23338,09	23338,09	4
3	74631,46	95295,09	171019,1	23338,09	23338,09	4
4	60489,33	82934,73	157066,1	9208,834	9208,834	4
5	60489,33	82934,73	157066,1	9208,834	9208,834	4
6	60489,33	82934,73	157066,1	9208,834	9208,834	4
7	60489,33	82934,73	157066,1	9208,834	9208,834	4
8	74631,46	95295,09	171019,1	23338,09	23338,09	4
9	18062,92	51664,43	115483,4	33244,71	18062,92	1
10	60489,33	82934,73	157066,1	9208,834	9208,834	4
...	...	...	...	...	...	...
1031	147253,7	111297,5	79741,09	186181,9	79741,09	3
1032	46347,19	71240,32	143150,2	4993,865	4993,865	4
1033	53418,26	76984,91	150102,8	2206,956	2206,956	4
1034	25133,99	55863,28	122373,6	26175,23	25133,99	1
1035	25133,99	55863,28	122373,6	26175,23	25133,99	1



Gambar 3 Hasil visualisasi PCA

Berdasarkan hasil perhitungan pada iterasi ketujuh, seluruh data tetap berada pada klaster yang sama seperti iterasi sebelumnya dan tidak terjadi lagi perubahan nilai centroid. Hal ini menunjukkan bahwa algoritma *K-Means* telah mencapai kondisi konvergen, sehingga proses iterasi dihentikan. Hasil *Clustering* akhir yang diperoleh kemudian digunakan sebagai dasar analisis karakteristik penjualan pada masing-masing klaster untuk mendukung pengambilan keputusan dalam pengelolaan stok dan penyusunan strategi pemasaran Toko Online Cendana Shooing

Berdasarkan hasil visualisasi PCA, klaster 4 memiliki jumlah anggota terbanyak, yaitu 588 data, sehingga tampak sebagai kelompok dengan sebaran titik yang paling dominan. Sementara itu, klaster 1 beranggotakan 296 data, klaster 3 beranggotakan 90 data, dan klaster 2 beranggotakan 61 data dengan pola penyebaran yang relatif lebih sedikit. Perbedaan distribusi titik pada setiap klaster menunjukkan bahwa algoritma *K-Means* mampu mengelompokkan data transaksi berdasarkan karakteristik yang serupa.

Evaluasi terhadap hasil *Clustering* dilakukan dengan menggunakan *Silhouette Score*[25], untuk mengukur sejauh mana objek-objek dalam satu klaster memiliki kemiripan dan seberapa baik klaster yang terbentuk saling terpisah. Hasil perhitungan pada gambar 5 menunjukkan bahwa rata-rata *Silhouette Score* sebesar 0,6246, yang termasuk dalam kategori layak/sesuai [26]. Hal ini dapat menjelaskan bahwa secara keseluruhan klaster yang terbentuk sudah cukup optimal, dimana sebagian besar data terkelompok dengan baik pada klasternya masing-masing.

Rata-rata Silhouette Score: 0.6246975914130294

Analisis Silhouette per Cluster:

	Nilai Silhouette							
	count	mean	std	min	25%	50%	75%	max
Cluster								
1	296.0	0.548564	0.231164	-0.001990	0.406972	0.677638	0.728574	0.737321
2	61.0	0.402141	0.152926	0.112359	0.317816	0.488577	0.517537	0.556915
3	90.0	0.356843	0.188016	-0.330435	0.281529	0.330451	0.555306	0.555306
4	588.0	0.727110	0.127272	0.280756	0.651792	0.770568	0.817014	0.817014

Gambar 4 Nilai Silhoutte coefficient

Karakteristik setiap klaster dianalisis berdasarkan nilai centroid akhir pada variabel harga barang, jumlah pembelian, dan total pembayaran. Penentuan kategori tingkat penjualan dilakukan berdasarkan peringkat nilai rata-rata total pembayaran pada centroid akhir setiap klaster. Klaster dengan nilai rata-rata total pembayaran tertinggi dikategorikan sebagai sangat laris, nilai berikutnya sebagai laris, kemudian cukup laris, dan nilai terendah sebagai kurang laris. Klaster 3 memiliki rata-

rata total pembayaran tertinggi, yaitu Rp163.333,30, dengan harga rata-rata Rp125.555,60 dan jumlah pembelian rata-rata 1,53 unit, sehingga dikategorikan sebagai sangat laris. Klaster 2 memiliki rata-rata total pembayaran Rp119.098,40, harga rata-rata Rp55.983,61, dan jumlah pembelian rata-rata 2,18 unit, sehingga dikategorikan sebagai laris. Klaster 1 memiliki rata-rata total pembayaran Rp77.817,57, harga rata-rata Rp77.006,76, dan jumlah pembelian rata-rata 1,02 unit, sehingga dikategorikan sebagai cukup laris. Klaster 4 memiliki rata-rata total pembayaran terendah, yaitu Rp41.938,78, dengan harga rata-rata Rp41.054,42 dan jumlah pembelian rata-rata 1,03 unit, sehingga dikategorikan sebagai kurang laris. Jumlah anggota klaster hanya menunjukkan banyaknya transaksi pada masing-masing kelompok dan bukan dasar penentuan kategori.

Berdasarkan hasil penelitian, algoritma *K-Means Clustering* berhasil mengelompokkan 1.035 transaksi penjualan Toko Online Cendana Shoooping menjadi empat klaster berdasarkan variabel harga barang, jumlah pembelian, dan total pembayaran. Kategori tingkat penjualan ditentukan berdasarkan peringkat nilai rata-rata total pembayaran pada centroid akhir, bukan berdasarkan jumlah anggota klaster. Klaster 3 merupakan kelompok sangat laris dengan rata-rata total pembayaran Rp163.333,30 dan 90 transaksi; Klaster 2 merupakan kelompok laris dengan rata-rata Rp119.098,40 dan 61 transaksi; Klaster 1 merupakan kelompok cukup laris dengan rata-rata Rp77.817,57 dan 296 transaksi; sedangkan Klaster 4 merupakan kelompok kurang laris dengan rata-rata Rp41.938,78 dan 588 transaksi. Visualisasi PCA digunakan untuk melihat pola persebaran klaster, sedangkan evaluasi menggunakan *Silhouette Coefficient* menghasilkan nilai 0,6246 yang menunjukkan kualitas klaster pada kategori layak/sesuai. Hasil ini dapat menjadi informasi pendukung bagi toko dalam pengelolaan persediaan dan penyusunan strategi pemasaran. Penelitian selanjutnya dapat menambahkan atribut seperti jenis barang, frekuensi pembelian, periode transaksi, keuntungan, atau karakteristik pelanggan apabila data tersedia, serta membandingkan *K-Means* dengan metode *Clustering* dan metrik evaluasi lainnya.

5. Kesimpulan

Penelitian ini berhasil menerapkan algoritma *K-Means Clustering* untuk mengelompokkan 1.035 transaksi penjualan Toko Online Cendana Shoooping selama periode Januari 2024 sampai Maret 2025 berdasarkan variabel harga barang, jumlah pembelian, dan total pembayaran. Tahapan penelitian meliputi identifikasi masalah, pengumpulan data, persiapan dan pemeriksaan data, penentuan jumlah klaster menggunakan metode *Elbow*, proses *Clustering* menggunakan *K-Means*, visualisasi menggunakan PCA, dan evaluasi menggunakan *Silhouette Coefficient*. Metode *Elbow* menghasilkan empat klaster ($k = 4$), sedangkan proses *K-Means* mencapai kondisi konvergen pada iterasi ke-7. Berdasarkan peringkat rata-rata total pembayaran pada centroid akhir, Klaster 3 dikategorikan sangat laris, Klaster 2 laris, Klaster 1 cukup laris, dan Klaster 4 kurang laris. Evaluasi menghasilkan nilai *Silhouette Coefficient* sebesar 0,6246 yang menunjukkan kualitas klaster pada kategori layak/sesuai. Hasil pengelompokan dapat digunakan sebagai informasi pendukung untuk mengevaluasi karakteristik transaksi, mengelola persediaan, dan menyusun strategi pemasaran. Namun, penelitian ini menggunakan skala asli tanpa normalisasi sehingga variabel dengan rentang nilai lebih besar, khususnya harga dan total pembayaran, memiliki pengaruh lebih besar terhadap jarak *Euclidean*.

Ucapan Terima Kasih

Ucapan terimakasih ini disampaikan kepada pemilik Toko Online Cendana Shoooping atas izin dan kesediaannya untuk memberikan data dan informasi yang dibutuhkan selama penelitian. Semoga hasil penelitian dapat bermanfaat bagi Toko Online Cendana Shoooping.

Referensi

- [1] B. Apriyanto dan S. L. M. Sitio, "Penerapan *K-Means* dalam Menganalisis Pola Pembelian Pelanggan pada Data Transaksi *E-Commerce*," *bit-Tech*, Vol. 7, No. 3, hal. 790–797, 2025, DOI: 10.32877/bt.v7i3.2195.
- [2] M. D. Pamungkas dan H. Februariyanti, "Penerapan Algoritma *K-Means Clustering* untuk Mengelompokkan Data Review Barang pada *E-Commerce* Lazada," *semantik*, Vol. 8, No. 2, hal. 99, 2022, DOI: 10.55679/semantik.v8i2.29058.
- [3] R. Sari Handayani, Gizka Alysah Diva, Eko Nur Hermansyah, dan Ari Siswati, "Transformasi <http://sistemasi.ftik.unisi.ac.id>

- Digital E-Commerce di Era Post-Pandemic: Analisis Tren dan Dampak pada Perilaku Konsumen Indonesia,” J. Inform. dan Kesehat., Vol. 2, No. 2, hal. 108–117, 2025, DOI: 10.35473/ikn.v2i2.3704.*
- [4] H. Hasanudin, “Dampak Pandemi Covid-19 pada Penjualan *E-Commerce* di Media Sosial,” *Ekon. Bisnis*, Vol. 20, No. 1, hal. 41–50, 2021, DOI: 10.32722/eb.v20i1.4058.
- [5] T. Hardiani, “Analisis *Clustering* Kasus Covid 19 di Indonesia menggunakan *Algoritma K-Means*,” *J. Nas. Pendidik. Tek. Inform.*, Vol. 11, No. 2, hal. 156–165, 2022, DOI: 10.23887/janapati.v11i2.45376.
- [6] N. Nurfidah, A. Armin, dan F. Farid, “Pengaruh Media Sosial *Facebook* dan *Whatsapp* dalam meningkatkan Volume Penjualan pada Toko Karya Baru (Study Kasus di Desa Tente Kecamatan Woha Kabupaten Bima) Tahun 2022,” *Bus. Manag.*, Vol. 2, No. 1, hal. 10–22, 2023, DOI: 10.58258/bisnis.v2i1.4137.
- [7] H. R. Ananda dan H. J. Salas, “Konvergensi Media *Facebook* dan *Whatsapp Business* dalam Strategi Komunikasi Pemasaran Digital : Analisis Model AIDA pada Toko Material dr . Interior 88 Media *Convergence of Facebook and Whatsapp Business in Digital Marketing Communication Strategies: An AIDA Model Analysis of the dr . Interior 88 Building Materials*,” hal. 34–48.
- [8] A. Amir, A. Gani, M. Sufri, dan M. H. Syahnur, “Strategi Komunikasi Pemasaran Bisnis dan Online dalam meningkatkan Penjualan pada Jejaring Media Sosial, *Facebook Marketplace*,” *Cent. Econ. Student J.*, Vol. 6, No. 3, hal. 2621–8186, 2023, [Daring]. Tersedia pada: <https://doi.org/10.56750/csej.v6i3.622>
- [9] B. Y. Putra, F. Y. Azzahra, dan I. A. Erlanda, “Klasterisasi Pengunjung Mall menggunakan *Algoritma K-Means* berdasarkan Pendapatan dan Pengeluaran,” *J. Inform. dan Tek. Elektro Terap.*, Vol. 11, No. 3s1, 2023, DOI: 10.23960/jitet.v11i3s1.3392.
- [10] R. Rustam, S. Rahmatullah, S. Supriyanto, dan S. Wahyuni, “Penerapan *Data Mining* untuk Prediksi Penjualan Produk Triplek pada PT Puncak Menara Hijau Mas,” *J. Inf. dan Komput.*, Vol. 8, No. 2, hal. 75–86, 2020, DOI: 10.35959/jik.v8i2.186.
- [11] M. A. Hakim, A. B. Prasetyo, dan D. Eridani, “Penerapan *Data Mining* dengan *Algoritma K-Means Clustering* Penyewaan Alat-Alat Event pada Studi Kasus CV. Dipo Rental Creativindo *Implementation of Data Mining using the K-Means Clustering Event Rental Algorithm in a Case Study of CV. Dipo Rental Creativindo*,” *J. Tek. Komput.*, Vol. 1, No. 4, hal. 148–155, 2023, DOI: 10.14710/jtk.v1i4.37011.
- [12] Indriyani, M. Handayani, dan Sumantri, “*Clustering* Minat Pembeli Produk *Furniture* menggunakan *Algoritma K-Means* untuk Strategi Penjualan,” *J. Algoritma*, Vol. 22, No. 2, hal. 518–526, 2025, DOI: 10.33364/algoritma/v.22-2.2625.
- [13] D. Handayani, A. Noeman, A. Hidayat, dan H. Lubis, “Implementasi *Algoritma K-Means* untuk,” Vol. 6, No. April, hal. 160–171, 2026.
- [14] B. Hartono, S. Eniyati, dan K. Hadiono, “Perbandingan Metode Perhitungan Jarak pada Nilai *Centroid* dan Pengelompokan Data menggunakan *K-Means Clustering*,” *J. Sist. Komput. dan Inform.*, Vol. 4, No. 3, hal. 503, 2023, DOI: 10.30865/json.v4i3.6021.
- [15] F. A. W. T. S. Wijaya, E. Prasetyo, dan R. F. Tias, “*Dynamic Time Warping* pada Metode *K-Means* untuk Pengelompokan Data Trend Penjualan Produk,” *J. Nas. Teknol. dan Sist. Inf.*, Vol. 10, No. 2, hal. 100–109, 2024, DOI: 10.25077/teknosi.v10i2.2024.100-109.
- [16] R. Fikria dan S. Sriani, “Analisis Metode *K-Means Clustering* dalam Pengelompokan Penjualan Sembako,” *J. Inf. Syst. Res.*, Vol. 5, No. 4, hal. 1464–1471, 2024, DOI: 10.47065/josh.v5i4.5699.
- [17] W. W. Kristianto dan C. Rudianto, “Penerapan *Data Mining* pada Penjualan Produk menggunakan Metode *K-Means Clustering* (Studi Kasus Toko Sepatu Kakikaki),” *J. Pendidik. Teknol. Inf.*, No. 5, hal. 90–98, 2020.
- [18] Y. Pitaloka Anggriani, A. Arif, dan F. Febriansyah, “Implementasi *Algoritma K-Means Clustering* dalam Menentukan Blok Tanaman Sawit Produktif pada PT Arta Prigel,” *JATI (Jurnal Mhs. Tek. Inform.)*, Vol. 8, No. 2, hal. 1820–1825, 2024, DOI: 10.36040/jati.v8i2.9225.
- [19] N. Afiasari, N. Suarna, dan N. Rahaningsi, “Implementasi *Data Mining* Transaksi Penjualan menggunakan *Algoritma Clustering* dengan Metode *K-Means*,” *J. SAINTEKOM*, Vol. 13, No. 1, hal. 100–110, 2023, DOI: 10.33020/saintekom.v13i1.402.

- [20] H. Al Rasyid, B. F. K. Soebari, dan D. S. Y. Kartika, “Implementasi Algoritma *K-Means Clustering* untuk Pengelompokan Penjualan Produk pada Online Shop Toko Gizi,” *Pros. Semin. Nas. Teknol. dan Sist. Inf.*, Vol. 2, No. 1, hal. 242–248, 2022, DOI: 10.33005/sitasi.v2i1.304.
- [21] M. Adelina Bui dan A. Bahtiar, “Implementasi Metode Algoritma *K-Means Clustering* untuk mengelompokkan Transaksi Penjualan Barang di Toko Arino,” *JATI (Jurnal Mhs. Tek. Inform.)*, Vol. 8, No. 2, hal. 1451–1456, 2024, DOI: 10.36040/jati.v8i2.8975.
- [22] Hibrizi Dzaky Dermawan Udi Kurniawan Yudhistira Arie Wijaya, “Analisis Data Penjualan menggunakan Algoritma *K-Means Clustering* pada Toko Daun Indah di Shopee *Analysis of Sales Data using K-Means Clustering Algorithm at Daun Indah Shop in Shopee*,” *J. Comput. Eng. Syst. SCI. e-ISSN*, Vol. 9, No. 1, hal. 175–191, 2024, [Daring]. Tersedia pada: www.jurnal.unimed.ac.id
- [23] T. Amalina, D. B. P. Adhi, dan B. N. Sari, “Metode *K-Means Clustering* dalam Pengelompokan Penjualan Produk *Frozen Food* Tifani Amalina 1 , Danendra Bima Adhi Pramana 2 , Betha Nurina Sari 3 Teknik Informatika, Universitas Singaperbangsa Karawang,” *J. Ilm. Wahana Pendidik.*, Vol. 8, No. September, hal. 574–583, 2022.
- [24] S. Sulistina, P. Eko Prasetyo Utomo, dan B. Ferdian Hutabarat, “Klasterisasi Wilayah Rawan Kriminalitas di Kota Jambi (2022-2024) menggunakan Algoritma *K-Means*,” *J. INSTEK (Informatika Sains dan Teknol.)*, Vol. 10, No. 2, hal. 546–560, 2025, DOI: 10.24252/instek.v10i2.62051.
- [25] R. Hidayati, A. Zubair, A. Hidayat Pratama, dan L. Indana, “*Silhouette Coefficient Analysis in 6 Measuring Distances of K-Means Clustering*,” *Techno.Com*, Vol. 20, No. 2, hal. 186–197, 2021.
- [26] N. Ulinnuha, “*Provincial Clustering in Indonesia based on Plantation Production using Fuzzy C-Means*,” *ITSMART J. Ilm. Teknol. dan Inf.*, Vol. 9, No. 1, hal. 8–12, 2020.