

Implementasi *K-Means Clustering* Untuk Mengelompokkan Kategori Penjualan Produk Toko Evanos

Muhammad Itba'ul Walidaini^{1*}, Hadi Sucipto¹

¹Program Studi S1 Teknik Informatika, Fakultas Teknologi Informasi, Universitas Hasyim Asy'ari Tebuireng Jombang, Jawa Timur, Indonesia

*Email: waaliwaalidayya@gmail.com

Info Artikel

Kata Kunci :

Data Mining,
K-Means Clustering,
Analisis Penjualan,
Pengelompokan Produk

Keywords :

Data Mining,
K-Means Clustering,
Sales Analysis,
Product Grouping

Tanggal Artikel

Dikirim : 24 Desember 2025

Direvisi : 2 Juni 2026

Diterima : 21 Oktober 2026

Abstrak

Dalam penjualan, faktor yang perlu diperhatikan dalam menganalisa laku tidaknya suatu produk adalah dengan memperhatikan ketersediaan stok guna mencegah penumpukan barang berlebih. Maka, dibutuhkan suatu proses untuk mengetahui tingkat penjualan tertinggi, sedang, dan terendah sebagai bahan pertimbangan dalam melakukan *restock* produk. Toko Evanos merupakan usaha UMKM yang bergerak dalam bidang jual beli jajanan khas daerah yang belum menerapkan pengelompokan produk dalam sistem informasi penjualannya. Kondisi tersebut menyebabkan pihak toko belum maksimal dalam menilai minat pembeli terhadap produk sehingga toko sering kehabisan stok ditengah banyaknya permintaan, terutama hari libur. Berdasarkan permasalahan tersebut, maka dibuat suatu sistem menggunakan *k-means clustering* untuk mengelompokkan tingkat penjualan produk ke dalam kategori tertinggi, sedang, dan terendah. Dalam sistem ini, penentuan *centroid* awal dilakukan menggunakan pendekatan statistik berupa nilai minimum, rata-rata, dan maksimum. Data penjualan diambil selama tiga bulan dengan 90 data produk. Hasil perhitungan *k-means clustering* menggunakan variabel stok keluar dan total penjualan menunjukkan kategori penjualan tertinggi sebanyak 6 item, sedang 28 item, dan terendah 56 item. Unjuk kerja sistem dalam melakukan *clustering* memperoleh nilai DBI sebesar 0,535 yang menunjukkan bahwa hasil *clustering* termasuk baik, sehingga dapat menjadi solusi terbaik bagi pemilik toko dalam mengontrol stok berdasarkan kategori penjualan produk.

Abstract

In sales activities, stock availability is an important things that should be considered to prevent excessive inventory accumulation. Therefore a method is needed to identify a products with high, medium, and low sales levels for restocking decisions. Evanos store is a UMKM market which sale a regional snacks but hasn't implemented a product classification in their sales information system. That condition makes difficult for the store to evaluate product interest, resulting in frequent stock shortages, especially on holidays. Based on these problems, a system with k-means clustering algorithm was developed to classify product sales into high, medium, and low categories. In this study, the initial centroids will determined using a statistical approach based on minimum, average, and maximum values. The dataset was collected in 3 months with 90 products. The result show that 6 products were classified for the high sales, 28 products for medium sales, and 56 products for low sales. The clustering performance evaluation for DBI value is 0,535, indicating the clustering result is good quality, it can be a best solution for store owner to manage the inventory more effectively based on product sales categories.

1. PENDAHULUAN

Kegiatan jual beli merupakan aktifitas dasar dalam menunjang kehidupan sehari-hari. Seiring dengan persaingan pasar yang semakin kompetitif, pelaku usaha dituntut untuk mampu menyusun strategi pemasaran yang baik. Akan tetapi, dilapangan banyak pelaku usaha yang belum maksimal dalam menganalisis data penjualan dengan optimal. Padahal, ini merupakan faktor yang penting untuk mempertimbangkan berbagai aspek yang dapat mempengaruhi keputusan dalam bisnis. Sebab dengan analisis tersebut, suatu perusahaan akan dapat membuat keputusan yang lebih matang dan juga meminimalisir resiko kerugian dalam pengembangan bisnisnya [1]. Salah satu faktor yang perlu diperhatikan dalam menganalisis penjualan adalah dengan memperhatikan ketersediaan stok guna mencegah penumpukan barang di gudang. Efek dari penumpukan ini sangat fatal. Banyak barang terlanjur kedaluarsa sebelum akhirnya terjual yang diakibatkan dari kurangnya kontrol dan pengawasan terhadap pembelian barang yang laku dan yang tidak laku. Ini merupakan masalah besar bagi pelaku usaha yang dimana jika hal ini terus terjadi, maka operasional akan terganggu.

Toko Evanos merupakan usaha UMKM (Usaha Mikro, Kecil, dan Menengah) yang bergerak dalam bidang jual beli jajanan khas daerah. Berdasarkan hasil observasi, masalah yang ditemukan dalam toko tersebut adalah tidak adanya sistem untuk mengelompokkan kategori penjualan pada sistem informasi penjualannya. Akibatnya, pihak toko mengalami kesulitan dalam manajemen stok dan mengelola ketersediaan produknya. Produk yang lebih laku harusnya lebih banyak dipajang di etalase daripada yang kurang laku. Karena tidak adanya sistem pengelompokan kategori penjualan, toko sering kehabisan stok ditengah banyaknya permintaan terutama saat hari libur.

Demi mengatasi permasalahan tersebut, maka dibuat sebuah sistem yang dapat membantu dalam mengolah dan menganalisis secara sistematis. Salah satunya menggunakan metode *k-means* untuk mengelompokkan data produk berdasarkan kategori tingkat penjualannya. Metode *k-means* dipilih karena mampu melakukan pengelompokan data besar dan kalkulasi outlier dengan cepat [2]. Pada penelitian ini, digunakan data simulasi yang disusun berdasarkan hasil observasi dan pola penjualan aktual Toko Evanos untuk menguji penerapan metode *k-means* dalam mengelompokkan tingkat penjualan produk. Dengan adanya pengelompokan ini, pemilik toko jadi lebih mudah dalam menilai mana produk yang laris dan mana yang kurang diminati. Hal ini tentu sangat berguna untuk menyusun strategi promosi serta mengontrol stok di gudang supaya lebih efisien [3].

Penelitian yang berkaitan dengan topik pengelompokan data penjualan menggunakan algoritma *k-means* diantaranya, penelitian pertama berjudul Penerapan *K-Means* Pada Pengelompokan Penjualan Produk *Smartphone* pada studi kasus HP Vivo. Penelitian tersebut menggunakan tiga *cluster* yang merepresentasikan kategori penjualan tinggi, sedang, dan rendah. Hasil pengelompokan menunjukkan 1 produk pada kategori penjualan tinggi, 4 produk pada kategori penjualan sedang, dan 8 produk pada kategori penjualan rendah [4]. Hal ini menunjukkan bahwa algoritma *k-means* dapat membantu dalam menganalisis data penjualan. Namun, dalam penelitian tersebut lebih berfokus pada pengelompokan produk berdasarkan tingkat penjualan dan belum membahas pemanfaatan hasil *clustering* sebagai dasar pengambilan keputusan dalam mengelola stok barang. Selain itu, objek penelitian yang digunakan berupa produk *smartphone* yang memiliki karakteristik berbeda dengan produk UMKM, khususnya produk makanan yang memiliki tingkat perputaran stok relatif cepat.

Penelitian kedua berjudul pengelompokan data penjualan sembako berdasarkan perilaku pembeli menggunakan algoritma *k-means clustering*. Penelitian tersebut menggunakan data penjualan selama satu tahun yang terdiri dari 1288 data transaksi dengan tujuan untuk mengidentifikasi pola pembelian konsumen. Dengan menggunakan dua *cluster* yaitu laku dan kurang laku, didapatkan hasil pengelompokan kategori laku sebanyak 1237 item, dan kurang laku sebanyak 51 item [5]. Hasil penelitian menunjukkan bahwa algoritma *k-means* mampu digunakan dalam identifikasi pola penjualan berdasarkan karakteristik data yang dimiliki. Selain itu, hasil *clustering* juga dapat memberikan gambaran mengenai kelompok produk mana yang memiliki tingkat permintaan berbeda, sehingga dapat dimanfaatkan sebagai bahan pertimbangan dalam pengambilan keputusan bisnis.

Berdasarkan penelitian terdahulu di atas, metode *k-means* terbukti mampu digunakan untuk mengelompokkan data penjualan berdasarkan tingkat penjualan maupun perilaku pembeli. Namun, penelitian tersebut umumnya masih berfokus pada hasil pengelompokan data dan belum membahas secara khusus pemanfaatan sebagai dasar pengambilan keputusan dalam pengendalian stok produk. Maka dari itu, tujuan penelitian ini adalah menerapkan algoritma *k-means* untuk mengelompokkan tingkat penjualan produk Toko Evanos ke dalam kategori tinggi, sedang, dan rendah agar dapat digunakan sebagai pendukung pengambilan keputusan dalam pengelolaan dan restock produk, sehingga dapat membantu pengelolaan stok secara lebih efektif.

2. METODE PENELITIAN

Metode dalam penelitian ini menggunakan algoritma *k-means* dengan mengelompokkan data produk kedalam tiga kategori penjualan dengan tujuan mengetahui minat pembeli terhadap produk serta mengontrol stok. Data produk akan diproses berdasarkan variabel stok keluar dan total penjualan untuk kemudian diketahui hasilnya berupa penjualan tertinggi, sedang, dan terendah.

2.1 Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif yang menekankan pada analisis data numerik. Pada studi kasus ini, digunakan data penjualan produk toko untuk dilakukan analisis menggunakan *k-means* untuk mengkategorikan produk berdasarkan tingkat penjualannya.

2.2 Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan dengan beberapa tahap. Pertama observasi, yaitu dengan mengunjungi lokasi penelitian untuk melihat dan mengamati kegiatan penjualan di toko. Kedua wawancara langsung dengan pemilik toko untuk memperoleh informasi yang berkaitan dengan penjualan. Ketiga melakukan studi kepustakaan dengan mencari sumber pustaka dari jurnal dan artikel ilmiah yang mendukung.

2.3 Data Penelitian

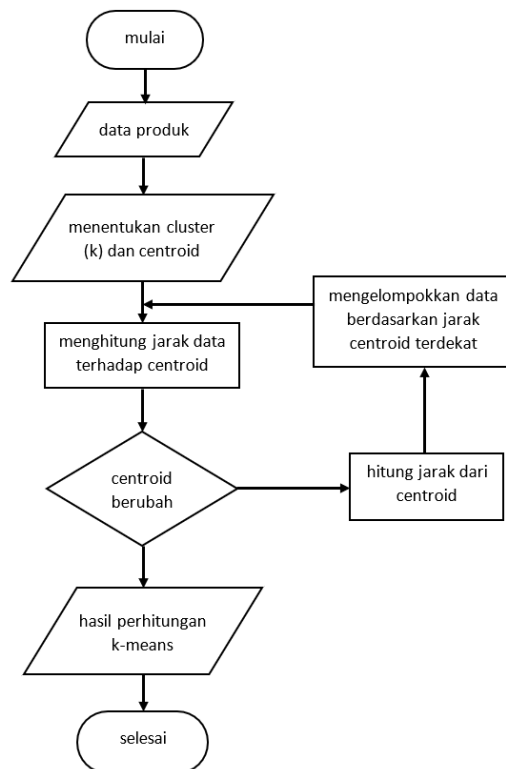
Setelah semua tahap sebelumnya dilakukan, maka diperoleh informasi mengenai pola penjualan produk berdasarkan observasi dan wawancara dengan pemilik toko. Data yang digunakan adalah data simulasi yang disusun berdasarkan kondisi nyata toko selama periode Juni, Juli, dan Agustus 2025 dengan jumlah 90 data produk. Penggunaan data simulasi dilakukan karena jumlah transaksi harian yang relatif sedikit sehingga kurang efektif untuk melakukan *clustering*. Oleh karena itu, data disusun mengikuti karakteristik dan pola penjualan yang diperoleh dari hasil observasi lapangan agar tetap dapat merepresentasikan kondisi penjualan toko. Data tersebut kemudian digunakan untuk menguji penerapan algoritma *k-means* dalam mengelompokkan tingkat penjualan produk ke dalam kategori tinggi, sedang, dan rendah.

Tabel 1. Data Produk

No	Kode Item	Nama Item	Harga Pokok	Harga Jual	Stok Item	Stok Keluar	Total Penjualan
1	ITM0001	Alen Bawang (keras)	8000	10000	31	580	5800000
2	ITM0002	Alen Jagung	8000	10000	33	571	5710000
3	ITM0003	Alen Balado	8000	10000	29	593	5930000
4	ITM0004	Alen Udang	11000	13000	28	452	5876000
5	ITM0008	Tempe Bulat	8000	10000	34	462	9340000
6	ITM0009	Manco Beras	8000	10000	28	459	12050000
7	ITM0010	Manco Ketan	10500	12500	32	449	16150000
8	ITM0011	Manco Wijen	13000	15000	31	476	4620000
9	ITM0012	Sale Pisang	10500	12500	32	475	4590000
...	...	...	...	...	...	...	...
90	ITM0090	Enting Jahe	16000	18000	32	375	6750000

2.4 Rancangan Sistem

Dalam merancang sistem ini seperti ditunjukkan pada Gambar 1, dimulai dengan menginput data produk, lalu menentukan jumlah *cluster*(k) dan centroid dengan pendekatan statistik (minimum, average, maximum), dilanjutkan dengan menghitung jarak data terhadap centroid hingga stabil dan didapatkan hasil berupa kategori tingkat penjualan masing-masing produk.



Gambar 1. Alur Sistem

2.5 Data Mining

Data mining merupakan metode untuk mengumpulkan dan memproses data dengan tujuan untuk menemukan informasi penting yang ada di dalamnya. Data mining memungkinkan pemakai menemukan informasi dalam suatu data di database yang tidak mungkin diketahui keberadaannya oleh pemakai [6]. Informasi yang diperoleh bisa berupa angka-angka, atau berupa data yang bisa digunakan untuk berbagai kebutuhan. Secara umum, data mining merupakan salah satu dari rangkaian *Knowledge Discovery in Database* (KDD). KDD merupakan cara untuk menggali suatu informasi melalui database yang telah ada [7].

2.6 Clustering

Clustering merupakan teknik data mining yang digunakan untuk mengelompokkan suatu data atau objek ke dalam kelompok-kelompok (*cluster*) hingga setiap data tersebut memiliki kemiripan tinggi dan berbeda dari objek di cluster lainnya [8]. Clustering bersifat tanpa arahan (*unsupervised*) yang berarti metode ini berjalan tanpa adanya latihan (*training*), guru (*teacher*), serta target output [9]. Tujuan dari *clustering* adalah untuk mengungkap struktur tersembunyi suatu data, baik yang bersifat semi-terstruktur maupun tidak terstruktur, sehingga ditemukan pola dan hubungan antar data yang tidak tampak secara langsung [10].

2.7 K-Means

K-Means merupakan algoritma untuk mengelompokkan data menjadi beberapa *cluster* atau kelompok dengan output berupa data tanpa label kelas [11]. Algoritma ini adalah bagian dari teknik *partitional clustering*, artinya berbasis pada titik pusat. Dimulai dengan menentukan *cluster* (k) dan juga centroid acak untuk ditetapkan sebagai centroid awal, lalu proses iterasi dijalankan hingga nilai centroidnya tetap, dengan indikasi data tidak berpindah-pindah dari satu kelompok ke kelompok lainnya [10]. Tahapan k-means dalam *clustering* data adalah sebagai berikut [12] :

1. Menentukan banyaknya *cluster* (k) yang akan digunakan.

- Menentukan centroid awal untuk langkah pertama iterasi dengan pendekatan statistik (*minimum, average, maximum*) dengan rumus :

$$c_{ij} = \frac{1}{N_a} \sum_{m=1}^{N_a} x_{mj} \quad (1)$$

Keterangan :

c_{ij} : centroid rata-rata cluster ke-i untuk variabel ke-j
 N_a : jumlah data anggota cluster ke-a
 i, m : indeks dari cluster
 j : indeks dari variabel
 x_{mj} : nilai data ke-m yang ada pada cluster tersebut untuk variabel ke-j

- Hitung jarak antara titik centroid dengan titik setiap objek dengan rumus *Euclidean Distance* :

$$d_{ik} = \sqrt{(x_{i1} - c_{k1})^2 + (x_{i2} - c_{k2})^2} \quad (2)$$

Keterangan :

d_{ik} : jarak data ke-i pada centroid ke-k
 x_{i1}, x_{i2} : atribut ke-1 dan ke-2 data ke-i
 c_{k1}, c_{k2} : atribut ke-1 dan ke-2 centroid ke-k

- Kelompokkan objek berdasarkan jarak ke centroid yang paling terdekat
- Kembali ke tahap nomor 2 hingga nilai centroid yang dihasilkan memiliki nilai tetap dan anggota cluster tidak berpindah ke kelompok lain.

3. HASIL DAN PEMBAHASAN

3.1 Hasil

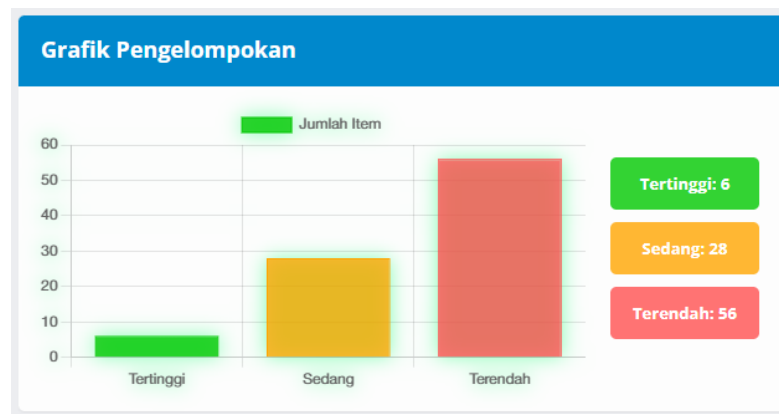
Berikut merupakan hasil dari penelitian yang telah dilakukan terhadap 90 data produk. Hasil pengelompokan menggunakan algoritma *k-means* menunjukkan penjualan produk tertinggi sebanyak 6 item, sedang sebanyak 28 item, dan terendah sebanyak 56 item seperti yang ditampilkan dalam Gambar 2. Jumlah produk pada kategori penjualan terendah menunjukkan nilai yang lebih banyak, artinya sebagian besar produk memiliki tingkat permintaan yang relatif rendah dibandingkan produk lainnya. Sementara itu, produk yang termasuk dalam kategori penjualan tinggi dapat menjadi prioritas utama dalam pengelolaan stok untuk menghindari kehabisan barang ketika terjadi peningkatan permintaan.

No	Kode Item	Nama Item	Harga	Cluster	Tingkat Penjualan
1	ITM0001	Alam Bawang	Rp. 10.000	C3	Terendah
2	ITM0002	Alam Jagung	Rp. 10.000	C3	Terendah
3	ITM0003	Alam Balaio	Rp. 10.000	C3	Terendah
4	ITM0004	Alam Ubiung	Rp. 13.000	C3	Terendah
5	ITM0005	Tempa Kripik 1/2 16	Rp. 20.000	C2	Sedang
6	ITM0006	Tempa Kripik 1/2 22	Rp. 25.000	C1	Tinggi
7	ITM0007	Tempa Kripik 1/2 41	Rp. 30.000	C1	Tinggi
8	ITM0008	Tempa Bubur	Rp. 10.000	C3	Terendah
9	ITM0009	Manico Beras	Rp. 10.000	C3	Terendah
10	ITM0010	Manico Katan	Rp. 12.500	C3	Terendah

Tertinggi : 6 | Sedang : 28 | Terendah : 56 | Total Data : 90

Gambar 2. Tampilan Hasil Pengelompokan

Pada Gambar 3, ditampilkan grafik hasil pengelompokan data produk berdasarkan kategori penjualan. Dalam grafik tersebut menunjukkan bahwa kategori penjualan terendah memiliki jumlah produk terbanyak yaitu 56 item, diikuti kategori penjualan sedang sebanyak 28 item, dan kategori penjualan tertinggi sebanyak 6 item. Hasil tersebut menunjukkan bahwa sebagian besar produk memiliki tingkat penjualan yang relatif rendah dibandingkan kategori lainnya.



Gambar 3. Tampilan Grafik Pengelompokan Data Produk

3.2 Pembahasan

1. Menentukan titik awal centroid berdasarkan pendekatan statistik dengan mencari nilai minimum, average, maximum dari 90 data produk dengan jumlah 3 *cluster* dengan menggunakan variabel stok keluar dan total penjualan. Nilai centroid awal yang digunakan dalam proses *clustering* ditampilkan dalam Tabel 2.

Tabel 2. Centroid Awal

Cluster	Stok Keluar	Total Penjualan
C1	593	16.150.000
C2	382	6.280.528
C3	211	3.718.000

2. Proses perhitungan menggunakan *Euclidean Distance* berdasarkan penentuan centroid awal yang telah ditentukan. Hasil perhitungan jarak pada iterasi pertama ditampilkan dalam Tabel 3. Nilai pada kolom Jarak C1, C2, dan C3 menunjukkan jarak masing-masing produk terhadap centroid setiap *cluster*. Semakin kecil nilai jarak yang diperoleh, maka produk akan semakin dekat dengan centroid *cluster* tersebut dan berpotensi menjadi anggota *cluster* yang bersangkutan.

Tabel 3. Hasil Iterasi 1

Kode Item	Jarak C1	Jarak C2	Jarak C3
ITM0001	10350000	480527,8186	2082000,033
ITM0002	10440000	570527,8091	1992000,033
ITM0003	10220000	350527,8413	2212000,033
ITM0004	10274000	404527,7838	2158000,013
ITM0005	6810000,001	3059472,223	5622000,006
ITM0006	4100000,002	5769472,223	8332000,004
ITM0007	270	9869472,222	12432000
ITM0008	11530000	1660527,78	902000,0349
ITM0009	11560000	1690527,78	872000,0353
ITM0010	10537500	668027,7811	1894500,015

3. Dari perhitungan iterasi 1, maka dihasilkan centroid baru yang ditampilkan pada Tabel 4. Nilai centroid baru diperoleh dari rata-rata data anggota pada masing-masing *cluster*. Perubahan nilai centroid menunjukkan adanya penyesuaian pusat *cluster* berdasarkan hasil pengelompokan data pada iterasi sebelumnya. Centroid baru ini kemudian digunakan sebagai acuan dalam proses perhitungan jarak pada iterasi berikutnya hingga diperoleh hasil *cluster* yang stabil.

Tabel 4. Centroid Baru

Cluster	Stok Keluar	Total Penjualan
C1	360	13.899.500
C2	385	6.570.175
C3	378	4.439.962

4. Proses perhitungan kembali iterasi berdasarkan centroid baru hingga diperoleh nilai centroid yang stabil. Pada Tabel 5 menampilkan nilai pada kolom Jarak C1, C2, dan C3 menunjukkan jarak masing-masing produk terhadap centroid pada setiap cluster. Berdasarkan hasil perhitungan, tidak terdapat perubahan anggota *cluster* dibandingkan iterasi sebelumnya, sehingga proses *clustering* dinyatakan konvergen, sehingga perhitungan berhenti pada iterasi ke 7.

Tabel 5. Hasil Iterasi 7

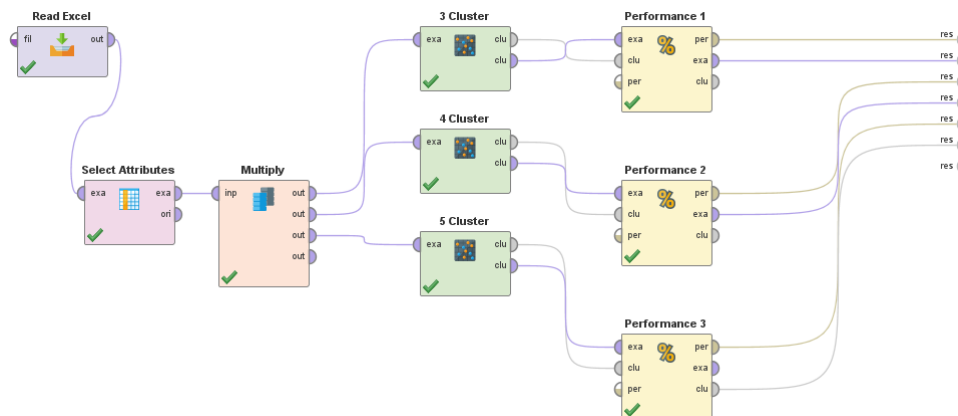
Kode Item	Jarak C1	Jarak C2	Jarak C3
ITM0001	6948333,336	1523660,729	734017,8823
ITM0002	7038333,336	1613660,727	644017,8832
ITM0003	6818333,337	1393660,732	864017,8815
ITM0004	6872333,334	1447660,717	810017,8597
ITM0005	3408333,334	2016339,288	4274017,858
ITM0006	698333,3405	4726339,287	6984017,858
ITM0007	3401666,667	8826339,286	11084017,86
ITM0008	8128333,334	2703660,716	445982,1491
ITM0009	8158333,334	2733660,716	475982,1482
ITM0010	7135833,334	1711160,716	546517,8606

5. Hasil akhir pengelompokan data ditampilkan dalam Tabel 6. Pada tabel tersebut diperoleh hasil pengelompokan 90 data produk ke dalam tiga cluster, yaitu cluster tertinggi (C1), cluster sedang (C2), dan cluster terendah (C3). Produk dengan label *cluster* tertinggi merupakan produk dengan tingkat penjualan paling tinggi sehingga perlu diprioritaskan dalam pengelolaan stok. Produk dengan label *cluster* sedang memiliki tingkat penjualan menengah, sedangkan produk dengan label *cluster* terendah menunjukkan tingkat penjualan yang relatif rendah dibandingkan *cluster* lainnya.

Tabel 6. Hasil

No	Kode Item	Nama Item	Cluster	Cluster Label
1	ITM0001	Alen Bawang (keras)	C3	Terendah
2	ITM0002	Alen Jagung	C3	Terendah
3	ITM0003	Alen Balado	C3	Terendah
4	ITM0004	Alen Udang	C3	Terendah
5	ITM0005	Tempe Kripik Isi 16	C2	Sedang
6	ITM0006	Tempe Kripik Isi 22	C1	Tertinggi
7	ITM0007	Tempe Kripik Isi 41	C1	Tertinggi
8	ITM0008	Tempe Bulat	C3	Terendah
9	ITM0009	Manco Beras	C3	Terendah
10	ITM0010	Manco Ketan	C3	Terendah

6. Pengujian kualitas *clustering* menggunakan DBI (*Davies-Bouldin Index*) melalui aplikasi RapidMiner Studio.



Gambar 4. Pengujian DBI (Davies-Bouldin Index)

Pada Gambar 4, ditampilkan proses pengujian DBI menggunakan RapidMiner Studio. Data yang telah melalui tahap seleksi atribut diduplikasi menggunakan operator *Multiply* untuk dilakukan pengujian pada tiga skenario jumlah cluster, yaitu 3 *cluster*, 4 *cluster*, dan 5 *cluster*. Lalu masing-masing hasil *clustering* dievaluasi menggunakan operator *Performance* untuk diperoleh nilai *Davies-Bouldin Index*-nya. Hasil pengujian digunakan sebagai dasar untuk membandingkan kualitas *cluster* pada setiap skenario jumlah *cluster* yang diuji. Untuk hasil pengujian dapat dilihat pada Tabel 7.

Tabel 7. Hasil DBI	
Jumlah Cluster	Nilai DBI
3	0,535
4	0,402
5	0,396

Berdasarkan hasil pengujian pada Tabel 7, dengan menggunakan 3 *cluster* mendapatkan skor DBI sebesar 0,535, untuk 4 *cluster* mendapatkan skor DBI sebesar 0,402, dan 5 *cluster* mendapatkan skor DBI sebesar 0,396. Dalam hal ini, algoritma *k-means* menghasilkan pembentukan *cluster* dengan sangat baik. Namun dalam penelitian ini jumlah *cluster* yang dipilih adalah 3 *cluster* yang memiliki nilai DBI 0,535 karena disesuaikan dengan kebutuhan pengelompokan kategori penjualan. Meskipun nilai DBI bukan yang paling rendah, namun hasil *clustering* tetap relevan dan dapat digunakan sebagai analisis data.

4. KESIMPULAN

Metode *k-means clustering* terbukti dapat mengelompokkan kategori penjualan produk. Peneliti berharap, dapat memberi kemudahan bagi pemilik toko dalam mengontrol stok produk yang ada pada toko, sehingga tidak ada lagi produk yang kedaluarsa akibat penumpukan stok di gudang. Hasil unjuk kerja algoritma *k-means* berdasarkan hasil perhitungan DBI menunjukkan skor sebesar 0,535. Meskipun nilai terbaik ada pada nilai 0,396 pada 5 *cluster*, dalam penelitian ini dipilih 3 *cluster* karena disesuaikan dengan kebutuhan analisa data untuk merepresentasikan tingkat penjualan tertinggi, sedang, dan terendah.

Penelitian ini masih memiliki beberapa keterbatasan, diantaranya data yang digunakan hanya berasal dari satu toko dengan periode pengamatan selama tiga bulan, serta masih menggunakan dua variabel yaitu stok keluar dan total penjualan. Oleh sebab itu, diharapkan pada penelitian selanjutnya dapat menggunakan periode pengamatan yang lebih panjang, jumlah data yang lebih besar, serta menambahkan variabel lain yang relevan untuk meningkatkan kualitas hasil *clustering*.

DAFTAR PUSTAKA

- [1] H. N. Safa'atin dan D. L. Denata, "Analisis Strategi Pemasaran Produk Lokal Pada Pasar Global," *Journal of Entrepreneurial Studies (JES)*, vol. 1, no. 1, pp. 28–35, May 2024. [Online]. Available: <https://doi.org/10.33508/jes.v1i1.5591>.
- [2] I. Padiku dan A. Lahinta, "Penerapan Clustering K-Means Untuk Mendukung Pengelolaan Koleksi Pada Perpustakaan Fakultas Teknik Universitas Negeri Gorontalo," *J. Tek.*, vol. 20, no. 1, hlm. 54–62, Jul 2022, doi: 10.37031/jt.v20i1.206.
- [3] A. D. Juwari, "Pengelompokan Produk Penjualan Menggunakan K-Means Sebagai Pendukung Strategi Bisnis Kafe Omah Kopi," *J. Profesi Ins. Univ. Lampung*, vol. 6, no. 2, Sep 2025, doi: 10.23960/jpi.v6n2.175.
- [4] F. P. A. Hasibuan, S. Sumarno, dan I. Parlina, "Penerapan K-Means pada Pengelompokan Penjualan Produk Smartphone," *SATESI J. Sains Teknol. Dan Sist. Inf.*, vol. 1, no. 1, hlm. 15–20, Sep 2021, doi: 10.54259/satesi.v1i1.3.
- [5] M. Husen Adlun, R. Astuti, dan F. M. Basysyar, "Pengelompokan Data Penjualan Sembako Berdasarkan Perilaku Pembeli Menggunakan Algoritma K-Means Clustering," *JATI J. Mhs. Tek. Inform.*, vol. 8, no. 4, hlm. 7699–7703, Agu 2024, doi: 10.36040/jati.v8i4.9806.
- [6] S. Suliman, "Implementasi Data Mining Terhadap Prestasi Belajar Mahasiswa Berdasarkan Pergaulan dan Sosial Ekonomi Dengan Algoritma K-Means Clustering," *SIMKOM*, vol. 6, no. 1, hlm. 1–11, Jan 2021, doi: 10.51717/simkom.v6i1.48.
- [7] F. Alghifari dan D. Juardi, "Penerapan Data Mining Pada Penjualan Makanan dan Minuman Menggunakan Metode Algoritma Naïve Bayes: Studi Kasus : Makan Barbeque Sepuasnya," *J. Ilm. Inform.*, vol. 9, no. 02, hlm. 75–81, Sep 2021, doi: 10.33884/jif.v9i02.3755.
- [8] W. Lestari, "Clustering Data Mahasiswa Menggunakan Algoritma K-Means Untuk Menunjang Strategi Promosi (Studi Kasus : STMIK Bina Bangsa Kendari)," *SIMKOM*, vol. 4, no. 2, hlm. 3548, Jul 2019, doi: 10.51717/simkom.v4i2.37.
- [9] Rindi Asti Ananda, Yani Maulita, dan Husnul Khair, "Clustering Menggunakan Algoritma K Means untuk Mengelompokan Data Perjudian Berdasarkan Wilayah di Kota Binjai (Studi Kasus: Pengadilan Negeri Binjai)," *Switch J. Sains Dan Teknol. Inf.*, vol. 2, no. 4, hlm. 167–180, Sep 2024, doi: 10.62951/switch.v2i4.226.
- [10] S. R. Agustin, I. Purnamasari, dan B. N. Sari, "Implementasi K-Means Untuk Pengelompokan Kategori Penjualan Barang Berbasis Web," vol. 5, no. 3, 2025.
- [11] L. Y. Hutabarat, I. Gunawan, I. Purnamasari, M. Safii, dan W. Saputra, "Penerapan Algoritma K-Means Dalam Pengelompokan Jumlah Penduduk Berdasarkan Kelurahan Di Kota Pematangsiantar," *J. Ilmu Komput. Dan Teknol.*, vol. 2, no. 2, hlm. 20–26, Jan 2022, doi: 10.35960/ikomti.v2i2.704.
- [12] N. N. Afidah, "Penerapan Metode Clustering dengan Algoritma K-Means untuk Pengelompokkan Data Migrasi Penduduk Tiap Kecamatan di Kabupaten Rembang," vol. 6, 2023.