

PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS DALAM KLASTERISASI DATA PENJUALAN PADA TOKO FROZEN FOOD XYZ

Nadiyah Syaidatus Shofa Abdul Hayat^{1*}

¹Program Studi Sistem Informasi, Universitas Pembangunan Nasional “Veteran” Jawa Timur, Surabaya,
Indonesia

*22082010038@student.upnjatim.ac.id

ABSTRAK

Toko Frozen Food XYZ menghadapi permasalahan dalam pengelolaan stok akibat banyaknya variasi produk dan perbedaan pola penjualan antarproduk, yang berpotensi menyebabkan penumpukan stok dan meningkatnya risiko kedaluwarsa. Penelitian ini bertujuan untuk membandingkan kinerja algoritma K-Means dan K-Medoids dalam melakukan klasterisasi data penjualan guna membantu pengelolaan persediaan secara lebih efektif. Data yang digunakan merupakan data transaksi penjualan periode Januari hingga Maret 2025 sebanyak 16.447 transaksi, yang kemudian diproses melalui tahapan preprocessing meliputi restrukturisasi, agregasi, dan normalisasi data. Variabel yang digunakan dalam klasterisasi adalah Jumlah Terjual dan Frekuensi Penjualan. Penentuan jumlah cluster optimal dilakukan menggunakan Elbow Method dan menghasilkan nilai $K = 5$ pada kedua algoritma. Kualitas hasil clustering dievaluasi menggunakan Silhouette Coefficient dan Davies-Bouldin Index (DBI). Hasil penelitian menunjukkan bahwa kedua algoritma menghasilkan pola klasterisasi yang konsisten, dengan dominasi produk berpenjualan rendah pada cluster terbesar dan hanya sedikit produk yang memiliki performa penjualan tinggi. Nilai Silhouette Coefficient sebesar 0,669 untuk K-Means dan 0,668 untuk K-Medoids menunjukkan kualitas cluster yang hampir sama, sedangkan nilai DBI sebesar 0,477 pada K-Means dan 0,480 pada K-Medoids menunjukkan bahwa K-Means sedikit lebih unggul dalam pemisahan cluster. Dengan demikian, kedua metode dinilai relevan untuk klasterisasi data penjualan, namun K-Means direkomendasikan sebagai metode yang lebih optimal untuk membantu pengambilan keputusan pengelolaan stok pada Toko Frozen Food XYZ.

Kata kunci: K-Means, K-Medoids, Klasterisasi, Data penjualan, Frozen Food

1 PENDAHULUAN

Industri *frozen food* di Indonesia terus menunjukkan peningkatan signifikan. Menurut Asosiasi Produsen Makanan Beku Indonesia (APMBI), permintaan terhadap produk makanan beku meningkat sekitar 35% dalam dua tahun terakhir, sebagaimana diberitakan oleh Kombi.id. Kenaikan ini dipengaruhi oleh perubahan gaya hidup masyarakat yang semakin mengutamakan makanan praktis serta perluasan distribusi melalui e-commerce dan armada frozen truck yang memperluas jangkauan pemasaran (Redaksi, 2025). Pertumbuhan permintaan tersebut tercermin pula dari ukuran pasar secara keseluruhan, Laporan dari Ken Research turut memperkuat gambaran pertumbuhan tersebut, dengan menyebutkan bahwa ukuran pasar frozen food di Indonesia telah mencapai sekitar USD 2 miliar. Faktor pendorongnya meliputi meningkatnya preferensi terhadap makanan siap saji dan berkembangnya infrastruktur *cold chain* yang mempermudah distribusi produk beku dengan kualitas terjaga (Sanjna, 2024).

Perubahan perilaku konsumen juga mendukung pertumbuhan pasar ini. Survei GoodStats tahun 2024 menunjukkan bahwa 86% responden memilih frozen food karena praktis, 72% karena

mudah disimpan, dan 65% karena memiliki masa simpan yang panjang. Temuan ini memperlihatkan bahwa frozen food semakin menjadi bagian dari gaya hidup konsumsi masyarakat modern di Indonesia (Wafa, 2025). Studi pada UMKM “Amalia Frozen Food” juga menemukan bahwa risiko operasional, termasuk kerusakan produk karena suhu, stok kadaluarsa, dan pencatatan persediaan, merupakan salah satu kategori risiko signifikan, selain risiko keuangan, pasar, dan pemasaran (Shalihah et al., 2024). Kondisi ini menyebabkan ketimpangan penjualan antarproduk, yang berdampak pada pengelolaan stok dan efektivitas strategi pemasaran. Permasalahan serupa juga dialami oleh Toko Frozen Food XYZ, di mana ketimpangan penjualan antarproduk dan risiko operasional seperti kerusakan produk serta stok kadaluarsa turut memengaruhi akurasi pengelolaan persediaan dan efektivitas strategi pemasaran toko.

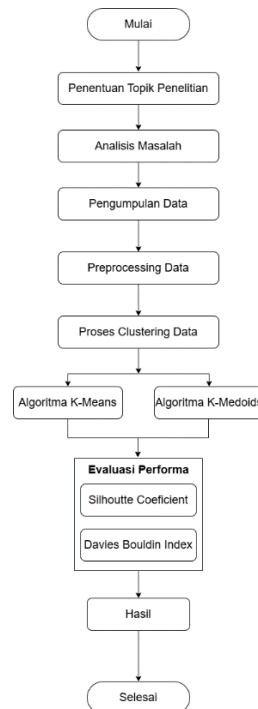
Dalam konteks pengelolaan data, penggunaan metode klusterisasi terbukti efektif untuk mengelompokkan data berdasarkan pola tertentu. Penelitian oleh (Rohman & Wibowo, 2024) membandingkan algoritma K-Means dan K-Medoids dalam analisis segmentasi pelanggan di sebuah mall. Hasil penelitian menunjukkan bahwa K-Means menghasilkan kluster yang lebih representatif dibanding K-Medoids, dengan nilai Silhouette Coefficient 0,553 untuk K-Means dan 0,485 untuk K-Medoids, sehingga algoritma K-Means dianggap lebih optimal.

Temuan ini relevan bagi Toko Frozen Food XYZ karena metode klusterisasi dapat digunakan untuk mengelompokkan produk berdasarkan pola penjualannya, sehingga pemilik toko dapat memahami kategori produk yang memiliki pola penjualan serupa dan merencanakan strategi restock, promosi, serta pengendalian stok yang lebih sesuai. Dengan kondisi tersebut, penerapan klusterisasi data penjualan menggunakan algoritma K-Means dan K-Medoids menjadi salah satu solusi yang tepat. Hasil klusterisasi dapat membantu pemilik Toko Frozen Food XYZ dalam merencanakan strategi restock, promosi, dan pengendalian stok secara lebih efisien, sekaligus membandingkan kinerja kedua algoritma untuk menemukan metode yang paling representatif.

2 METODE

2.1 Metode Penelitian

Dalam melaksanakan suatu penelitian, memerlukan metode penelitian sebagai kerangka konseptual yang sistematis. Metode ini berfungsi untuk mengarahkan setiap tahap penelitian agar dilaksanakan secara terstruktur dan konsisten, sehingga proses penelitian menjadi lebih efisien. Dengan demikian, penelitian diharapkan dapat menghasilkan data dan temuan yang valid, relevan, serta tepat waktu (Malik et al., 2018). Tahap awal dari penelitian ini meliputi Analisis Masalah, Pengumpulan Data, Pra-pemrosesan Data, Proses Pengelompokan menggunakan algoritma K-Means dan K-Medoids, serta Pengujian Akurasi menggunakan metrik kinerja DBI dan Silhouette Coefficient. Berikut merupakan alur tahapannya:



Gambar 1. Tahapan Penelitian

2.1.1 Penentuan Topik Penelitian

Tahap pertama dalam penelitian ini adalah menentukan fokus dan topik penelitian yang akan dilakukan. Topik penelitian ini dipilih karena diketahui bahwa industri frozen food di Indonesia mengalami pertumbuhan signifikan, didorong oleh permintaan akan makanan praktis dan ekspansi infrastruktur cold chain. Namun, di tengah pertumbuhan ini, pengelolaan data penjualan dan persediaan yang efisien menjadi krusial untuk menjaga kualitas produk dan profitabilitas. Oleh karena itu, topik penelitian yang dipilih adalah Penerapan Algoritma Klasterisasi K-Means dan K-Medoids untuk Analisis Pola Penjualan Guna Optimalisasi Manajemen Persediaan pada Sektor UMKM Frozen Food.

2.1.2 Analisis Masalah

Pada tahap ini dilakukan proses analisis awal terhadap permasalahan yang berkaitan dengan pengelolaan stok produk pada Toko Frozen Food XYZ. Analisis difokuskan pada upaya mengidentifikasi aspek-aspek yang berpotensi memengaruhi keseimbangan persediaan produk. Dalam tahap ini, data penjualan dianalisis untuk menemukan pola dan karakteristik penjualan tiap produk, yang selanjutnya digunakan sebagai dasar dalam proses pengelompokan (klasterisasi). Proses klasterisasi ini dirancang untuk memperoleh informasi yang dapat mendukung penyusunan strategi pengelolaan stok yang lebih efektif sesuai kebutuhan penelitian.

2.1.3 Pengumpulan Data

Tahap pengumpulan data pada penelitian ini dilakukan melalui wawancara dengan pemilik Toko Frozen Food XYZ untuk memperoleh informasi mengenai sumber dan bentuk data transaksi yang digunakan dalam penelitian. Melalui wawancara tersebut, peneliti mendapatkan penjelasan mengenai struktur data penjualan yang tersimpan dalam sistem internal toko, termasuk jenis atribut yang dicatat dalam setiap transaksi. Informasi hasil wawancara kemudian digunakan sebagai dasar

untuk melakukan proses ekstraksi data dari sistem penjualan toko. Data yang telah diperoleh selanjutnya dipersiapkan untuk proses analisis, khususnya pada tahap klasterisasi, dengan menyesuaikan format dan atribut data agar sesuai dengan kebutuhan metode yang digunakan.

2.1.4 Preprocessing Data

Tahap preprocessing data dilakukan untuk menyiapkan data penjualan agar sesuai untuk proses klasterisasi. Proses ini mencakup restrukturisasi data agar lebih terorganisasi, pemilihan atribut yang relevan, serta agregasi data untuk membentuk variabel yang diperlukan dalam analisis. Selain itu, dilakukan pembersihan data guna mengatasi duplikasi atau ketidakkonsistenan, serta normalisasi pada variabel numerik agar memiliki skala yang seimbang saat digunakan dalam algoritma klasterisasi. Seluruh langkah ini diterapkan untuk memastikan data siap diproses pada tahap pemodelan.

2.1.5 Proses Clustering Data

Pada penelitian ini proses clustering data dilakukan dengan menggunakan algoritma K-Means dan K-Medoids. *Clustering* merupakan salah satu teknik dalam data mining yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok (*cluster*) berdasarkan kesamaan karakteristik yang dimiliki (Kiwasi et al., 2024). Clustering membantu dalam mengidentifikasi pola atau struktur tersembunyi, menyederhanakan analisis data, dan mendukung pengambilan keputusan dengan memberikan wawasan tentang hubungan antar objek dalam dataset (Wibawa et al., 2024).

Algoritma *K-Means* merupakan metode *clustering* yang digunakan untuk mengelompokkan data ke dalam *K cluster* berdasarkan tingkat kesamaan antar titik data, dan populer dalam analisis data *eksploratif* karena kesederhanaannya serta kemampuannya menangani *dataset* besar (Maulani et al., 2024). Adapun langkah-langkah melakukan *clustering* dengan metode *K-Means* dimulai dengan (Rahmah Aulia, 2020):

1. Menentukan jumlah *cluster* *k* yang akan dibentuk.
2. Dilakukan inisialisasi *k* pusat *cluster*, yang umumnya dilakukan secara acak (*random*), di mana setiap pusat diberikan nilai awal *random*.
3. Setiap data atau objek dialokasikan ke *cluster* terdekat berdasarkan jarak ke pusat *cluster*. Kedekatan ini dihitung antara tiap titik data dan pusat *cluster*, dan jarak antar titik data biasanya dihitung menggunakan *Euclidean distance* (Rahmah Aulia, 2020), yang dirumuskan sebagai:

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

Dimana:

$D(i, j)$ = Jarak data ke *i* ke pusat *cluster* *j*

X_{ki} = Data ke *i* pada *atribut* data ke *k*

X_{kj} = Titik pusat ke *j* pada *atribut* ke *k*

4. Setelah semua data ditempatkan pada *cluster* terdekat, dilakukan perhitungan ulang pusat *cluster* berdasarkan anggota *cluster* saat ini. Pusat *cluster* biasanya dihitung dengan mengambil rata-rata (*mean*) dari semua data dalam *cluster*, namun dapat juga menggunakan *median* sebagai alternatif.

5. Proses pengelompokan data dan perhitungan ulang pusat *cluster* ini diulang sampai pusat *cluster* tidak berubah lagi, Jika pusat *cluster* berubah lagi maka proses kembali ke nomor 3. Proses ini dilakukan sampai pusat *cluster* tidak berubah lagi dan menandakan bahwa algoritma *K-Means* telah selesai.

Algoritma K-Medoids merupakan teknik *clustering* berbasis partisi yang digunakan untuk mengelompokkan sejumlah n objek ke dalam beberapa cluster. Dalam algoritma ini, proses partisi dilakukan dengan memilih titik acuan (medoid) yang paling mewakili setiap cluster dan meminimalkan tingkat ketidaksamaan antara setiap objek dengan medoid tersebut (Kaligis & Yulianto, 2022). Adapun langkah-langkah melakukan *clustering* dengan metode *K-Medoids* dimulai dengan (Marlina et al., 2018):

1. Inisialisasi pusat cluster sebanyak k (jumlah cluster).
2. Alokasikan setiap data (objek) ke cluster terdekat menggunakan persamaan ukuran jarak Euclidian Distance
3. Pilih secara acak objek pada masing-masing cluster sebagai kandidat medoids baru.
4. Hitung jarak setiap objek yang berada pada masing-masing cluster dengan kandidat medoids baru.
5. Hitung total simpangan (S) dengan menghitung nilai total distance baru – total distance lama. Jika $S < 0$, maka tukar objek dengan data cluster untuk membentuk sekumpulan k objek baru sebagai medoid.
6. Ulangi langkah 3 sampai 5 hingga tidak terjadi perubahan medoids, sehingga didapatkan cluster beserta anggota clustermasing-masing.

2.1.6 Evaluasi Performa

Setelah proses pengelompokan menggunakan K-Means dan K-Medoids, hasil pengelompokan di evaluasi performanya menggunakan metode Silhouette Coefficient dan DBI (Davies Bouldin Index). *Silhouette Coefficient* merupakan salah satu metrik evaluasi yang digunakan untuk menilai kualitas klasterisasi. Metrik ini mengukur sejauh mana suatu titik data sesuai dengan klaster tempat ia berada sekaligus terpisah dengan baik dari klaster lain (Hartono et al., 2023). Metode ini bekerja dengan membandingkan tingkat kesesuaian setiap titik data terhadap klasternya sendiri dengan kedekatannya terhadap klaster lain. Semakin tinggi nilai *Silhouette*, semakin baik kualitas pengelompokan yang dihasilkan (Kiwasi et al., 2024). Selanjutnya, Evaluasi performa juga dilakukan dengan menggunakan metode DBI (Davies Bouldin Index), *Davies-Bouldin Index (DBI)* merupakan metode evaluasi klaster yang pertama kali diperkenalkan oleh David L. Davies dan Donald W. Bouldin. Metode ini digunakan untuk mengukur kualitas pemisahan antar klaster dengan menghitung rata-rata rasio antara komponen kohesi (kedekatan data dengan pusat klasternya) dan separasi (jarak antar pusat klaster) (Mughnyanti et al., 2020). Nilai DBI yang rendah menunjukkan bahwa klaster terpisah dengan baik dan tidak saling tumpang tindih, sedangkan nilai DBI yang tinggi mengindikasikan adanya *overlap* antar klaster (Kiwasi et al., 2024).

2.1.7 Hasil

Setelah seluruh tahapan penelitian dilaksanakan, mulai dari pengumpulan data, preprocessing, hingga proses klasterisasi menggunakan algoritma K-Means dan K-Medoids, penelitian ini akan menghasilkan keluaran berupa pembentukan klaster produk frozen food berdasarkan pola penjualannya. Klaster yang terbentuk akan mengelompokkan produk ke dalam kategori dengan tingkat penjualan yang berbeda, seperti penjualan tinggi, sedang, dan rendah.

Selain itu, hasil penelitian ini juga akan menampilkan perbandingan performa kedua algoritma melalui pengujian metrik validitas Silhouette Coefficient dan Davies-Bouldin Index (DBI). Dari pengujian tersebut, akan ditentukan algoritma mana yang menghasilkan kualitas klasterisasi lebih optimal, sehingga dapat memberikan rekomendasi metode terbaik untuk analisis pola penjualan dan optimalisasi manajemen persediaan pada Toko Frozen Food XYZ.

3 HASIL DAN PEMBAHASAN

3.1 Analisis Masalah

Pada tahap analisis masalah yang dilakukan terhadap Toko Frozen Food XYZ, ditemukan bahwa toko menghadapi permasalahan utama dalam pengelolaan stok akibat banyaknya variasi produk yang dijual. Berdasarkan hasil pengamatan dan analisis data, beberapa produk mengalami penumpukan stok yang cukup signifikan, sementara produk lainnya memiliki pergerakan yang lebih cepat. Ketimpangan ini mengakibatkan risiko meningkatnya jumlah produk yang mendekati tanggal kedaluwarsa, terutama karena pola penjualan sebelumnya belum dianalisis secara sistematis. Untuk memahami karakteristik penjualan setiap produk dan mendukung pengambilan keputusan terkait pengelolaan persediaan, data penjualan kemudian dikelompokkan menggunakan metode klasterisasi. Proses klasterisasi ini memberikan gambaran mengenai pola penjualan yang berbeda antar produk sehingga dapat digunakan sebagai dasar strategi pengelolaan stok yang lebih baik.

3.2 Pengumpulan Data

Tahap pengumpulan data dalam studi ini dilakukan melalui wawancara dengan pemilik toko. Data yang digunakan melibatkan data transaksi penjualan Toko Frozen Food XYZ selama periode Januari 2025 hingga Maret 2025. Data ini diperoleh langsung dari sistem internal toko dan menghasilkan dataset besar sebanyak 16.447 entri transaksi. Setiap entri mencakup 18 atribut rinci, termasuk No Transaksi, Tanggal, Nama Item, Kode Item, Jumlah (Jml), Harga, Potongan, dan Total Akhir, yang sangat penting untuk menganalisis pola pembelian. Proses pengumpulan dilakukan dengan ekstraksi data digital dari basis data transaksi toko, yang kemudian akan dipersiapkan untuk proses klasterisasi guna mengidentifikasi produk yang memiliki pola penjualan serupa.

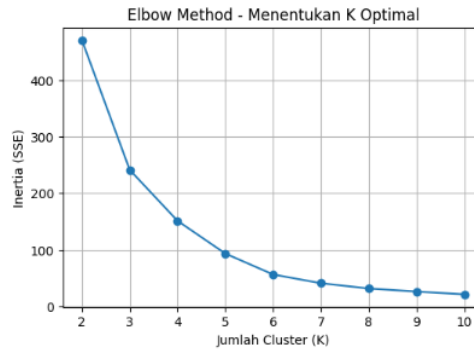
3.3 Preprocessing Data

Tahap preprocessing ini dimulai dengan restrukturisasi data (data restructuring), di mana 58.933 transaksi yang awalnya berbentuk nested diubah menjadi format flat sebanyak 161.304 baris, memisahkan setiap item pembelian ke baris tersendiri. Selanjutnya, dilakukan pemilihan data (data selection) untuk membuang atribut yang tidak relevan, dilanjutkan dengan agregasi data (data aggregation) yang merangkum data penjualan per produk untuk menghasilkan dua variabel klasterisasi utama: Jumlah Terjual dan Frekuensi Terjual. Setelah pembersihan data, variabel numerik hasil agregasi kemudian menjalani normalisasi (data normalization). Normalisasi ini penting untuk memastikan bahwa kedua variabel memiliki kontribusi yang seimbang dalam perhitungan jarak klaster, sehingga data siap dimanfaatkan secara optimal pada tahap pemodelan menggunakan K-Means dan K-Medoids.

3.4 Clustering Data

3.4.1 Clustering menggunakan Algoritma K-Means

Sebelum melakukan clustering data, dilakukan pencarian nilai K optimal untuk mengetahui berapa jumlah cluster yang paling optimal. Pencarian nilai K ini dilakukan dengan menggunakan Elbow Method dengan menguji 2 hingga 11 cluster. Jumlah cluster terbaik dipilih berdasarkan besarnya nilai sudut perbandingan setiap cluster. Pada gambar 2 menunjukkan hasil evaluasi klasterisasi menggunakan method elbow pada algoritma K-Means.



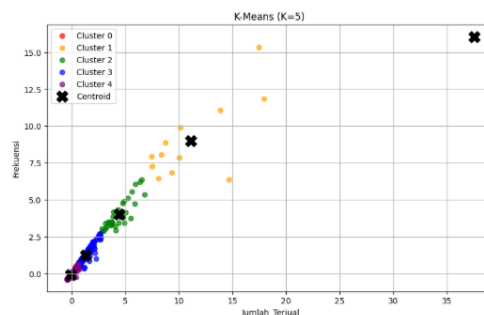
Gambar 2. Hasil Elbow Method K-Means

Kemudian dapat dilihat nilai mana yang terus berkurang atau bertambah, dan seterusnya. Terlihat pada Gambar 2, nilai antara cluster 4 dan cluster 5 tampak stabil. Data rinci dari Gambar 6 di bawah ini disajikan pada Tabel 1.

Tabel 1. Nilai Elbow Algoritma K-Means

Cluster	Nilai	Selisih
2	470.770844	0.000000
3	241.190387	229.580457
4	151.093164	90.097223
5	93.172882	57.920283
6	56.177029	36.995852
7	40.696771	15.480258
8	31.356921	9.339850
9	25.941626	5.415295
10	21.025062	4.916564

Berdasarkan Tabel 1, nilai total cost menurun tajam hingga K = 5, tetapi setelah itu penurunannya semakin kecil dan stabil. Hal ini menunjukkan titik “elbow” berada pada K = 5, sehingga jumlah cluster optimal untuk K-Means adalah 5 cluster.

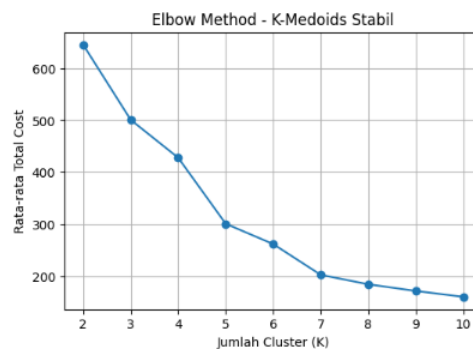


Gambar 3. Hasil K-Means

Berdasarkan hasil clustering menggunakan K-Means dengan $K = 5$, distribusi jumlah produk pada setiap cluster menunjukkan perbedaan yang cukup signifikan. Cluster 0 hanya memiliki 1 produk, Cluster 1 terdiri dari 12 produk, Cluster 2 mencakup 35 produk, Cluster 3 berjumlah 111 produk, dan Cluster 4 merupakan cluster terbesar dengan 409 produk. Hasil ini memberikan gambaran awal mengenai sebaran produk berdasarkan pola penjualan sebelum dilakukan analisis lebih lanjut terkait karakteristik masing-masing cluster.

3.4.2 Clustering menggunakan Algoritma K-Medoids

Hal yang sama dilakukan pada algoritma K-Medoids, yaitu menentukan jumlah cluster optimal menggunakan Elbow Method. Titik “elbow” ditentukan saat penurunan total cost mulai melambat, sehingga jumlah cluster terbaik dapat dipilih untuk menghasilkan pembagian data yang representatif. Hasil evaluasi klasterisasi K-Medoids ditampilkan pada Gambar 4.



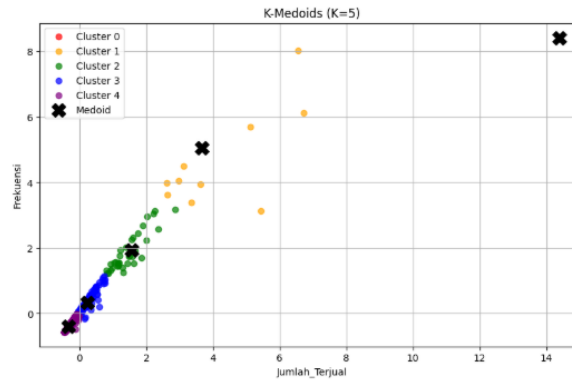
Gambar 4. Hasil Elbow Method K-Medoids

Selanjutnya, dapat diamati perubahan nilai total cost antar cluster. Pada Gambar 6 terlihat bahwa nilai mulai stabil pada peralihan dari cluster 4 ke cluster 5. Data rinci yang mendukung hal tersebut disajikan pada Tabel 2.

Tabel 2. Nilai Elbow Algoritma K-Medoids

Cluster	Nilai	Selisih
2	288.461365	0.000000
3	229.229631	59.231734
4	199.275321	29.954310
5	141.088850	58.186471
6	121.940632	19.148217
7	98.624345	23.316287
8	85.272190	13.352156
9	78.585946	6.686243
10	67.003002	11.582944

Berdasarkan Elbow Method, nilai grafik mengalami penurunan yang besar hingga $K = 5$, tetapi setelah itu penurunannya semakin kecil dan cenderung stabil. Hal ini menunjukkan bahwa penambahan cluster setelah $K = 5$ tidak memberikan peningkatan signifikan. Oleh karena itu, $K = 5$ dipilih sebagai nilai K terbaik untuk K-Medoids.



Gambar 5. Hasil K-Medoids

Hasil pengelompokan data produk dengan metode K-Medoids pada lima cluster menunjukkan adanya perbedaan jumlah anggota di setiap cluster. Cluster 0 hanya mencakup 1 produk, Cluster 1 berisi 11 produk, Cluster 2 terdiri dari 36 produk, Cluster 3 memiliki 111 produk, dan Cluster 4 menjadi cluster terbesar dengan total 409 produk. Temuan ini memberikan gambaran awal mengenai sebaran produk berdasarkan pola penjualan sebelum dilakukan analisis lebih lanjut terhadap karakteristik masing-masing cluster.

3.5 Evaluasi Performa

Kualitas hasil clustering menggunakan metode K-Means dan K-Medoids dievaluasi berdasarkan nilai Silhouette Coefficient dan Davies-Bouldin Index (DBI). Hasil pengujian menunjukkan adanya perbedaan kualitas pengelompokan antara kedua metode tersebut. Selanjutnya, perbandingan hasil evaluasi K-Means dan K-Medoids disajikan pada Tabel 3.

Tabel 3. Perbandingan Hasil Evaluasi Metode K-Means dan K-Medoids

Metode	K-Means	K-Medoids
Silhouette Coefficient	0.669	0.668
Davies-Bouldin Index (DBI)	0.477	0.480

Berdasarkan Tabel 3, diperoleh nilai Silhouette Coefficient pada metode K-Means sebesar 0,669 dan pada metode K-Medoids sebesar 0,668 menunjukkan bahwa perbedaan kualitas clustering antara kedua metode sangat kecil. Hal ini menandakan bahwa baik K-Means maupun K-Medoids sama-sama mampu membentuk cluster dengan struktur yang baik dan hasil yang relatif setara. Pada nilai Davies-Bouldin Index (DBI), metode K-Means menghasilkan nilai 0,477 sedangkan metode K-Medoids sebesar 0,480. Perbedaan ini juga tergolong kecil, namun nilai DBI yang lebih rendah pada K-Means menunjukkan bahwa metode ini sedikit lebih baik dalam pemisahan antar cluster. Dengan demikian, dapat disimpulkan bahwa kedua metode masih relevan digunakan dalam penelitian ini, namun secara keseluruhan metode K-Means dinilai lebih unggul meskipun selisihnya sangat tipis.

3.6 Analisis dan Hasil

3.6.1 Hasil Clustering K-Means

Berdasarkan hasil pengolahan data menggunakan metode K-Means, produk berhasil dikelompokkan ke dalam lima cluster (Cluster 0 sampai Cluster 4) berdasarkan variabel Jumlah Terjual dan Frekuensi Penjualan. Cluster 0 merepresentasikan produk dengan jumlah terjual dan

frekuensi penjualan paling tinggi sehingga dapat dikategorikan sebagai produk dengan performa terbaik. Cluster 1 berisi produk dengan tingkat penjualan yang masih tinggi, namun berada sedikit di bawah Cluster 0. Cluster 2 menunjukkan kelompok produk dengan tingkat penjualan dan frekuensi pada kategori menengah. Cluster 3 merepresentasikan produk dengan jumlah penjualan rendah dan frekuensi yang kecil. Sementara itu, Cluster 4 terdiri dari produk dengan tingkat penjualan dan frekuensi terendah, sehingga dapat dikategorikan sebagai produk dengan pergerakan paling lambat. Contoh hasil cluster dengan menggunakan K-Means pada beberapa produk dapat dilihat pada tabel 4.

Tabel 4. Hasil Clustering K-Means

Nama Item	Jumlah Terjual	Frekuensi	Cluster
SEMAR KULIT PANGSIT (PCS)	29700	1185	0
DANDY / SEMAR / DLL KULIT LUMPIA	14350	882	1
SALAM SOSIS MERAH PENDEK 500GR	6140	600	1
CEDEA DUMPLING AYAM 500GR	5630	414	2
DEL EXTRAHOT 200GR	2520	248	2
DAGING SLICE MIX FAT GRADE A 500GR	2450	196	3
TUSUK PENTOL BANYAK	780	73	3
MC LEWIS MAYONAIIS SWEET 1KG	840	48	4
BERNARDI BAKSO GOLD SAPI ISI 4	10	1	4

Pada K-Means clustering terdapat hanya 1 produk pada Cluster 0, 12 produk pada Cluster 1, 35 produk pada Cluster 2, 111 produk pada Cluster 3 dan Cluster 4 merupakan cluster dengan produk terbanyak yaitu 409 produk

3.6.2 Hasil Clustering K-Medoids

Pada metode K-Medoids, hasil clustering menunjukkan pola yang hampir sama. Cluster 0 tetap menggambarkan produk dengan jumlah terjual dan frekuensi tertinggi. Cluster 1 mencerminkan produk dengan performa penjualan tinggi, Cluster 2 berada pada tingkat penjualan menengah, Cluster 3 menunjukkan produk dengan penjualan rendah, dan Cluster 4 berisi produk dengan tingkat penjualan serta frekuensi paling rendah. Hasil Clustering dengan menggunakan metode K-Medoids untuk beberapa produk dapat dilihat pada table 5

Tabel 5. Hasil Clustering K-Medoids

Nama Item	Jumlah Terjual	Frekuensi	Cluster
SEMAR KULIT PANGSIT (PCS)	29700	1185	0
DANDY / SEMAR / DLL KULIT LUMPIA	14350	882	1
SALAM SOSIS MERAH PENDEK 500GR	6140	600	1
PANDAWA CIRENG SALJU BIRU / ORANGE	6650	493	2
DEL EXTRAHOT 200GR	2520	248	2
DAGING SLICE MIX FAT GRADE A 500GR	2450	196	3
MC LEWIS MAYONAIIS ORI 500GR	740	74	3
INDOMINA BOLA UDANG	870	65	4
BERNARDI BAKSO GOLD SAPI ISI 4	10	1	4

Pada K-Medoids clustering terdapat hanya 1 produk pada Cluster 0, 11 produk pada Cluster 1, 36 produk pada Cluster 2, 111 produk pada Cluster 3 dan Cluster 4 merupakan cluster dengan produk terbanyak yaitu 409 produk

4 KESIMPULAN

Berdasarkan hasil penelitian mengenai Perbandingan Algoritma K-Means dan K-Medoids dalam Klasterisasi Data Penjualan pada Toko Frozen Food XYZ, dapat disimpulkan bahwa kedua metode mampu mengelompokkan data produk secara efektif berdasarkan variabel Jumlah Terjual dan Frekuensi Penjualan. Hasil penentuan jumlah cluster menggunakan Elbow Method menunjukkan bahwa baik algoritma K-Means maupun K-Medoids menghasilkan nilai K optimal yang sama, yaitu $K = 5$, sehingga struktur pengelompokan pada kedua metode dapat dibandingkan secara adil. Pola hasil clustering menunjukkan konsistensi, dimana cluster dengan jumlah produk paling sedikit merepresentasikan produk dengan performa penjualan tertinggi, sedangkan cluster dengan jumlah anggota terbanyak didominasi oleh produk dengan tingkat penjualan terendah. Hal ini menegaskan adanya ketimpangan kontribusi produk terhadap total penjualan, dimana hanya sebagian kecil produk memberikan pengaruh besar terhadap omzet, sementara sebagian besar lainnya memiliki pergerakan stok yang lambat.

Berdasarkan hasil Evaluasi Performa, nilai Silhouette Coefficient pada metode K-Means sebesar 0,669 dan K-Medoids sebesar 0,668 menunjukkan bahwa kualitas pembentukan cluster pada kedua metode berada pada kategori baik dengan perbedaan yang sangat tipis. Hal serupa juga ditunjukkan oleh nilai Davies-Bouldin Index (DBI), dimana K-Means memperoleh nilai 0,477 dan K-Medoids sebesar 0,480. Meskipun selisih nilai evaluasi tergolong sangat kecil, namun K-Means memiliki nilai DBI yang lebih rendah dan Silhouette yang sedikit lebih tinggi, yang mengindikasikan bahwa pemisahan cluster yang dihasilkan K-Means secara keseluruhan sedikit lebih baik dibandingkan K-Medoids.

Dengan demikian, dapat disimpulkan bahwa kedua metode sama-sama relevan dan layak digunakan untuk klasterisasi data penjualan, namun algoritma K-Means dinilai lebih unggul secara kinerja meskipun perbedaannya sangat tipis. Hasil clustering ini diharapkan dapat membantu pihak Toko Frozen Food XYZ dalam menentukan strategi pengelolaan stok, mengidentifikasi produk unggulan, serta mengevaluasi produk dengan perputaran rendah sehingga keputusan bisnis dapat diambil secara lebih tepat dan berbasis data.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh pihak yang telah memberikan dukungan dan kontribusi dalam penyusunan penelitian ini, khususnya kepada institusi pendidikan dan dosen pembimbing yang telah memberikan arahan, bimbingan, serta masukan yang sangat berarti selama proses penelitian. Penulis juga mengucapkan terima kasih kepada pemilik dan pihak Toko Frozen Food XYZ yang telah bersedia memberikan data serta informasi yang diperlukan dalam penelitian ini. Selain itu, penulis menyampaikan apresiasi kepada seluruh pihak lain yang turut membantu, baik secara langsung maupun tidak langsung, sehingga penelitian ini dapat diselesaikan dengan baik. Dukungan dan kerja sama yang diberikan sangat berperan dalam kelancaran penyusunan dan penyempurnaan penelitian ini.

DAFTAR PUSTAKA

- Hartono, N., Informasi, S., dan Teknologi, S., & Alauddin Makassar, U. (2023). Implementasi Metode Clustering Sebagai Penunjang Strategi dalam Manajemen Pelanggan. *JURNAL FASILKOM*, 13(3), 605–613.
- Kaligis, G. B., & Yulianto, S. (2022). ANALISA PERBANDINGAN ALGORITMA K-MEANS, K-MEDOIDS, DAN X-MEANS UNTUK PENGELOMPOKKAN KINERJA PEGAWAI (Studi Kasus: Sekretariat DPRD Provinsi Sulawesi Utara). *Jurnal Penerapan Teknologi Informasi Dan Komunikasi*, 01(03), 179–193.
- Kiwasi, J., Reba, F., Aleksander, S., M., A., & Edy, H. (2024). *DATA MINING (MEMAHAMI POLA DI BALIK ANGKA)* (N. Rusiardi, Ed.). EUREKA MEDIA AKSARA.
- Malik, R. A., Defit, S., & Yuhandri, Y. (2018). Comparison of K-Means Clustering Algorithm with Fuzzy C-Means In Measuring Satisfaction Level Of Television Da'wah Surau TV. *Rabit : Jurnal Teknologi Dan Sistem Informasi Univrab*, 3(1), 10–21. <https://doi.org/10.36341/rabit.v3i1.387>
- Marlina, D., Fauzer Putri, N., Fernando, A., & Ramadhan, A. (2018). Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokkan Wilayah Sebaran Cacat pada Anak. *Jurnal CoreIT*, 4(2).
- Maulani, G., Cipta, R., Ditha, K., Eka, M., Jurnaidi, A., Astika, A., Budy, C., Hudianti, E., Lestari, S., Kurnia, W., & Rifai, A. (2024). *PENERAPAN DATA MINING DI BERBAGAI BIDANG* (Wisandra, Ade). CV HEI PUBLISHING INDONESIA.
- Mughnyanti, M., Efendi, S., & Zarlis, M. (2020). Analysis of determining centroid clustering x-means algorithm with davies-bouldin index evaluation. *IOP Conference Series: Materials Science and Engineering*, 725(1). <https://doi.org/10.1088/1757-899X/725/1/012128>
- Rahmah Aulia, S. (2020). KLASTERISASI POLA PENJUALAN PESTISIDA MENGGUNAKAN METODE K-MEANS CLUSTERING (STUDI KASUS DI TOKO JUANDA TANI KECAMATAN HUTABAYU RAJA). *Djtechno : Journal of Information Technology Research*, 1.
- Redaksi. (2025, September 25). *Peluang & Tren Bisnis Frozen Food Indonesia, dari Ready Meals hingga Produk Bakery*. Kombi.Id. <https://kombi.id/read/2025/09/25/peluang-tren-bisnis-frozen-food-indonesia-dari-ready-meals-hingga-produk-bakery/>
- Rohman, N., & Wibowo, A. (2024). Perbandingan Metode K-Medoids dan Metode K-Means Dalam Analisis Segmentasi Pelanggan Mall. *Science and Information Technology Journal*, 7(1), 49–58. <https://doi.org/10.31598>
- Sanjna. (2024, December). *Indonesia Frozen Food Market Outlook to 2030*. Kenresearch.Com. <https://www.kenresearch.com/industry-reports/indonesia-frozen-food-market>

- Shalihah, S. A., Ayu Dwi Nurmavita, D., & Alika, A. (2024). Analisis Manajemen Risiko Pada UMKM “Amalia Frozen Food” Bandung. In *Jurnal Serambi Ekonomi dan Bisnis pISSN* (Vol. 8, Issue 1). <https://ojs.serambimekkah.ac.id/serambi-ekonomi-dan-bisnis/p456>
- Wafa, I. (2025, January 14). 7 Alasan Orang Indonesia Gemar Konsumsi Frozen Food. Goodstats.Id. <https://data.goodstats.id/statistic/7-alasan-orang-indonesia-gemar-konsumsi-frozen-food-mjBAs>
- Wibawa, P., Iwan, I. G., Suryani, Surachman, A., Ridwan, A., Mahendra, I. G., Nurtanzis, M., Slamet, I., Harlina, S., & Dendi, I. M. (2024). *BUKU AJAR DATA MINING* (Fitra, Ed.). PT. Sonpedia Publishing Indoensia. <https://www.researchgate.net/publication/377415198>