

KLASTERISASI PENJUALAN PRODUK MINUMAN MENGGUNAKAN K-MEANS DAN VISUALISASI DENGAN POWER BI

Siska Aulia, April Lia Hananto, Bayu Priyatna, Shofa Shofiah Hilabi

Sistem Informasi, Universitas Buana Perjuangan Karawang
Jalan HS. Ronggo Waluyo, Telukjambe Timur, Kabupaten Karawang, Indonesia
si22.siskaaulia@mhs.ubpkarawang.ac.id

ABSTRAK

Pengelolaan data penjualan produk minuman menjadi aspek penting dalam menjaga keseimbangan persediaan agar sesuai dengan kebutuhan konsumen. Variasi jenis produk dan fluktuasi tingkat permintaan seringkali menyulitkan pengambilan keputusan terkait pengelolaan stok apabila tidak didukung oleh analisis yang sistematis. Dengan demikian, tujuan penelitian ini adalah untuk melakukan klusterisasi data penjualan Kedai Cozy Drink periode April hingga Oktober 2025. Algoritma K-Means Clustering dengan pendekatan *Knowledge Discovery in Database* (KDD) digunakan pada penelitian ini. Dataset awal berjumlah 1.315 transaksi dan setelah proses *preprocessing* diperoleh 1.309 transaksi yang valid untuk dianalisis. Setelah metode Elbow diterapkan untuk menentukan jumlah cluster yang optimal, hasil analisis menunjukkan bahwa jumlah cluster terbaik adalah tiga. Evaluasi melalui perhitungan Silhouette Score sebesar 0,87 menunjukkan bahwa kualitas pengelompokan termasuk dalam kategori sangat baik. Ketiga cluster tersebut merepresentasikan kategori produk *best seller*, *profit driver*, dan *slow moving*. Hasil klusterisasi selanjutnya divisualisasikan menggunakan Power BI untuk mempermudah interpretasi pola penjualan produk sehingga dapat dijadikan dasar dalam pengelolaan stok yang lebih optimal dan efisien.

Kata kunci: Data Mining, K-Means, Clustering, Power BI

1. PENDAHULUAN

Kemajuan teknologi informasi mendorong penggunaan data sebagai landasan utama dalam proses pengambilan keputusan bisnis [1]. Pada usaha penjualan produk minuman, data transaksi yang tersimpan setiap hari mengandung informasi penting mengenai pola permintaan konsumen [2]. Akan tetapi, data tersebut kerap kali hanya dimanfaatkan sebagai arsip administratif semata dan belum diolah lebih lanjut untuk menunjang strategi pengelolaan stok secara optimal. Kondisi ini menyebabkan potensi informasi yang terkandung dalam data historis penjualan belum dimanfaatkan secara maksimal [3].

Variasi jenis produk serta perbedaan tingkat penjualan antar produk dapat menyebabkan ketidakseimbangan persediaan. Produk dengan tingkat permintaan tinggi berisiko mengalami kekurangan stok, sedangkan produk dengan tingkat permintaan rendah berpotensi menumpuk dan menimbulkan kerugian [4]. Kondisi ini menegaskan pentingnya penerapan metode analisis yang mampu mengidentifikasi pola penjualan dan mengelompokkan produk berdasarkan karakteristik performanya sebagai dasar pengelolaan persediaan yang lebih optimal [5].

Pendekatan clustering digunakan untuk mengidentifikasi kelompok data berdasarkan tingkat kemiripan karakteristik antar objek. Salah satu metode yang umum diterapkan adalah K-Means, yang mengelompokkan data dengan mempertimbangkan kedekatan nilai atribut tertentu [6]. Penerapan metode ini pada data penjualan memungkinkan pengelompokkan produk berdasarkan performa penjualannya, sehingga karakteristik masing-masing kelompok dapat dianalisis lebih lanjut [7].

Algoritma K-Means telah digunakan dalam beberapa penelitian sebelumnya untuk mengelompokkan data penjualan. Namun, kebanyakan penelitian belum mengintegrasikan proses analisis secara sistematis melalui tahapan *Knowledge Discovery in Database* (KDD) serta belum menyajikan hasil klusterisasi dalam bentuk visualisasi interaktif yang mendukung interpretasi data secara lebih komprehensif. Meskipun demikian, KDD sangat penting untuk memastikan bahwa proses pengolahan data dilakukan secara terstruktur dan menghasilkan pola yang valid.

Berdasarkan hal tersebut penelitian ini bertujuan untuk memecahkan masalah tersebut dan mengklasifikasikan data penjualan produk minuman menggunakan algoritma K-Means melalui pendekatan KDD serta menyajikan hasilnya dalam bentuk visualisasi menggunakan Power BI. Penyelesaian permasalahan dilakukan melalui proses seleksi dan pembersihan data transaksi, penerapan algoritma K-Means untuk membentuk cluster produk, evaluasi kualitas hasil klusterisasi, serta visualisasi hasil analisis dalam bentuk yang lebih informatif. Hasil yang diharapkan akan memberikan gambaran yang lebih baik tentang cara distribusi penjualan dalam membantu pengambilan keputusan.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Istilah *data mining* merujuk pada proses analisis data untuk menemukan pola tersembunyi dalam kumpulan data berukuran besar [8]. Proses ini memanfaatkan berbagai pendekatan seperti teknik statistik, pengolahan basis data, dan *machine learning*

[9]. Melalui *data mining*, data yang awalnya hanya berupa kumpulan catatan transaksi dapat diolah menjadi informasi yang bernilai. Informasi tersebut kemudian dapat dimanfaatkan untuk memahami kecenderungan tertentu, mengidentifikasi pola perilaku, serta mendukung perumusan strategi berbasis data [10]. Dalam penelitian ini, *data mining* digunakan untuk menganalisis data penjualan produk minuman sehingga dapat diperoleh gambaran pola performa penjualan secara lebih terstruktur dan sistematis [11].

2.2. Knowledge Discovery in Database (KDD)

KDD menggambarkan rangkaian tahapan untuk mengekstraksi informasi bernilai data mentah [12]. Proses ini meliputi *data selection*, *data preprocessing*, *data transformation*, *data mining*, dan evaluasi hasil. Melalui tahapan tersebut, data yang belum terstruktur dapat diolah menjadi pengetahuan yang relevan untuk mendukung pengambilan keputusan [13].

2.3. Clustering

Teknik clustering digunakan untuk membagi data ke dalam beberapa kelompok berdasarkan kedekatan karakteristik antar objek [14]. Pengelompokan ini terbentuk secara otomatis dari pola yang muncul dalam data, sehingga kelompok data akan terdiri dari data dengan atribut yang sebanding, dan kelompok data dengan atribut yang berbeda akan terdiri dari kelompok yang berbeda [15]. Pada penelitian ini, clustering diterapkan mengelompokkan produk minuman berdasarkan tingkat penjualannya [16].

2.4. Algoritma K-Means

Metode K-Means melakukan pengelompokan data ke dalam beberapa cluster titik pusat cluster (centroid) [17]. Dimulai dengan menghitung jumlah cluster (K), selanjutnya hitung jarak antara setiap data dan centroid. Setelah seluruh data ditempatkan pada cluster terdekat, posisi centroid akan diperbarui sesuai rata-rata anggota tiap cluster [18]. K-Means banyak digunakan karena sederhana, efisien, dan efektif dalam mengelompokkan data numerik dalam jumlah besar [19].

2.5. Elbow Method

Metode Elbow membantu dalam menentukan berapa banyak cluster yang cocok untuk algoritma K-Means. Pendekatan ini dilakukan dengan membandingkan nilai *Within Cluster Sum of Squares* (WCSS) pada berbagai variasi jumlah cluster [20]. Perubahan nilai WCSS kemudian dianalisis untuk melihat titik penurunan yang melambat, menunjukkan jumlah cluster yang paling sesuai [21].

2.6. Visualisasi Data dan Power BI

Visualisasi data merupakan tahap penting dalam proses analisis karena membantu menyajikan informasi dalam bentuk yang lebih mudah dipahami [22]. Melalui visualisasi, pola, tren, dan perbandingan

antar data dapat terlihat dengan lebih jelas dibandingkan hanya dalam bentuk tabel angka. Dalam proses analisis data, visualisasi berperan penting untuk mendukung pengambilan keputusan [23]. Dalam penelitian ini, hasil klasterisasi divisualisasikan menggunakan Microsoft Power BI. Perangkat lunak ini memungkinkan penyajian data dalam bentuk *dashboard* interaktif. Fitur grafik dinamis dan filter memungkinkan Pengguna mengeksplorasi hasil analisis secara lebih efektif, termasuk dalam penyajian hasil klasterisasi [24].

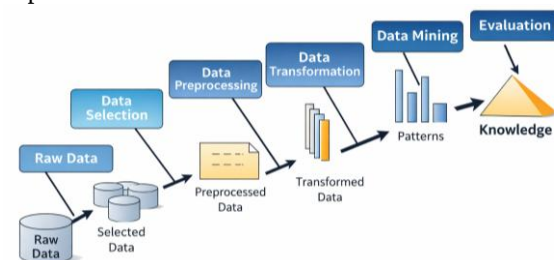
2.7. Penelitian Terdahulu

Algoritma K-Means telah digunakan sebagai dasar pengambilan keputusan bisnis dalam beberapa penelitian sebelumnya untuk menganalisis data penjualan. Penelitian yang dilakukan oleh Awaljan dkk. [25] menerapkan metode K-Means untuk mengelompokkan penjualan minuman ringan berdasarkan jenis produk guna mengidentifikasi pola kontribusi masing-masing barang. Hasil penelitian menunjukkan bahwa K-Means dapat membagi produk ke dalam berbagai kelompok sesuai tingkat performa penjualannya.

Selanjutnya, Nugraha dkk. [10] mengimplementasikan K-Means clustering dalam analisis penjualan pada sebuah toko ritel. Proses pengelompokan dilakukan untuk membedakan produk dengan kinerja penjualan tinggi dan rendah sehingga dapat menjadi dasar dalam penyusunan strategi pemasaran yang lebih tepat. Temuan penelitian menunjukkan bahwa teknik klasterisasi efektif dalam memberikan gambaran distribusi dan karakteristik penjualan produk.

3. METODE PENELITIAN

Dalam penelitian ini, pendekatan *Knowledge Discovery in Database* (KDD) digunakan untuk mengolah dan menganalisis data penjualan. Adapun tahap interpretasi hasil klasterisasi dan visualisasi menggunakan Power BI sebagai bentuk pemanfaatan hasil analisis untuk mendukung proses pengambilan keputusan.



Gambar 1. Tahapan Knowledge Discovery in Database

Berikut adalah tahapan-tahapan yang dilakukan dalam penelitian ini berdasarkan pendekatan KDD:

3.1. Data Selection

Tahap ini dilakukan dengan mengumpulkan transaksi penjualan produk minuman yang digunakan sebagai dataset utama penelitian. Seluruh data

disimpan dalam format Microsoft Excel (.xlsx). Sebelum dilakukan analisis lebih lanjut, dataset terlebih dahulu melalui proses *preprocessing* untuk memastikan kualitas dan kesiapan data.

3.2. Data Preprocessing

Data dibersihkan pada tahap *preprocessing* untuk memastikan kualitas dataset sebelum dianalisis. Proses ini meliputi perbaikan kesalahan penulisan, penghapusan spasi ganda, serta penghapusan transaksi dengan nilai kosong.

3.3. Data Transformation

Transformasi dilakukan dengan mengubah data transaksi menjadi bentuk agregat. Menghitung total penjualan setiap produk akan menghasilkan nilai yang akan digunakan sebagai variabel dalam proses klusterisasi.

3.4. Data Mining

Penerapan K-Means dalam penelitian ini dilakukan setelah data melalui tahap proses pembersihan dan transformasi. Sebelum proses klusterisasi dijalankan, metode Elbow digunakan untuk menghitung jumlah cluster yang ideal dengan menghitung nilai *Within-Cluster of Squares* (WCSS). Hasil analisis tersebut menjadi dasar dalam proses pengelompokan produk berdasarkan total nilai penjualan yang dinormalisasi.

3.5. Evaluasi

Silhouette Score digunakan dalam evaluasi hasil klusterisasi untuk mengevaluasi kualitas pemisahan antar cluster. Metode ini digunakan untuk mengetahui seberapa baik suatu objek berada dalam clusternya dibandingkan dengan cluster lainnya. Semakin tinggi

nilai silhouette yang diperoleh, semakin baik tingkat kohesi dalam cluster serta semakin jelas pemisahan antar cluster.

4. HASIL DAN PEMBAHASAN

4.1. Data Selection

Memilih data transaksi penjualan produk minuman periode April-Oktober 2025 untuk digunakan dalam analisis. Dataset terdiri dari 1.315 transaksi dengan atribut utama berupa nama produk dan kuantitas terjual. Data yang digunakan dapat dilihat pada Tabel 1.

Tabel 1. Data penjualan

No	Tanggal	Menu	Harga	Qty	Total
1	20-Apr	Matcha Frappe	13.000	16	208.000
2	20-Apr	Matcha	10.000	15	150.000
3	20-Apr	Matcha Latte	13.000	11	143.000
4	20-Apr	Matchachoco	15.000	7	105.000
5	20-Apr	Original Tea	5.000	12	60.000
...	...	...	...	...	...
1311	20-Oct	Red Velvet	10.000	8	80.000
1312	20-Oct	Oreo Milk	15.000	15	225.000
1313	20-Oct	Thai Tea	10.000	10	100.000
1314	20-Oct	Lemon Tea	8.000	10	80.000
1315	20-Oct	Taro Milk	10.000	15	150.000

4.2. Data Preprocessing

Pada tahap *preprocessing* ditemukan beberapa inkonsistensi penulisan nama produk yang berpotensi menyebabkan duplikasi data pada tahap agregasi. Permasalahan yang ditemukan meliputi kesalahan ejaan, perbedaan penggunaan huruf kapital, serta penggunaan spasi ganda pada nama produk seperti pada Tabel 2 berikut:

Tabel 2. Data preprocessing

No	Data Sebelum	Data Sesudah	Masalah	Tindakan
1	Matchaa	Matcha	Kesalahan ejaan	Standarisasi
2	Matcha frappe	Matcha Frappe	Perbedaan huruf kapital	Standarisasi
3	Matcha Latte	Matcha Latte	Spasi ganda	Standarisasi
4	Matca ricano	Matcharicano	Kesalahan ejaan	Standarisasi
5	Oreoo Milk	Oreo Milk	Kesalahan ejaan	Standarisasi
6	1/1 (kuantitas)	-	Format numerik tidak valid	Dihapus

Proses pembersihan data dilakukan untuk menyeragamkan format penulisan serta memastikan seluruh atribut numerik berada dalam format yang valid. Kesalahan penulisan nama menu diperbaiki melalui standarisasi sehingga data dengan penulisan berbeda namun merujuk pada produk yang sama dapat digabungkan pada saat proses agregasi. Selain itu, satu data dengan format kuantitas tidak valid dihapus karena tidak dapat dikonversi menjadi nilai numerik. Setelah tahap *preprocessing*, dataset yang digunakan dalam analisis berjumlah 1.309 transaksi yang telah memenuhi aspek konsistensi dan validitas data.

4.3. Data Transformation

Setelah tahap *preprocessing*, data transaksi diagregasi berdasarkan nama produk dengan menghitung total nilai penjualan selama periode penelitian. Hasil agregasi menunjukkan adanya perbedaan kontribusi pendapatan antar produk, dengan nilai tertinggi sebesar Rp 37.365.000 dan terendah Rp 848.000.

Normalisasi dilakukan menggunakan metode *Min-Max Scaling* dengan rentang 0-1 untuk memastikan perbedaan skala tidak memengaruhi perhitungan jarak pada algoritma K-Means. Hasil agregasi dan normalisasi tersebut selanjutnya digunakan sebagai input pada tahap penentuan jumlah

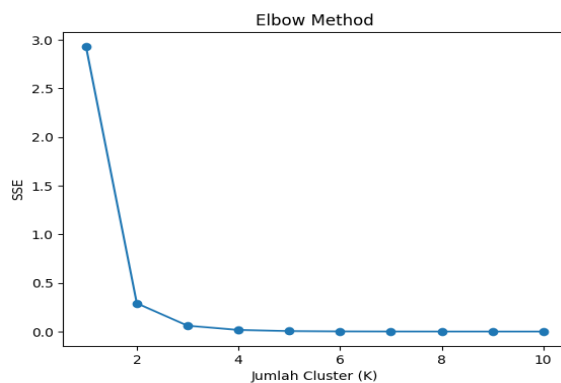
cluster dan proses klasterisasi. Rincian hasilnya ditampilkan pada Tabel 3.

Tabel 3. Data transformation

No	Menu	Total	Total_Scaled
1	Oreo Milk	37365000.0	1,000000
2	Mango Yakult	37128000.0	0,993510
3	Brown Sugar Boba	33612000.0	0,897226
4	Matcha Frappe	32916000.0	0,878166
5	Matcha	22300000.0	0,587452
6	Matcha Milkshake	13275000.0	0,340307
7	Thai Tea	12710000.0	0,324835
8	Matcha Aren Brulee	2686000.0	0,050333
9	Matcha Es Teler	2652000.0	0,048799
10	Taro Milk	2630000.0	0,048799
11	Matchachoco	2190000.0	0,036750
12	Dark Choco	2070000.0	0,029110
13	Matcha Latte	1911000.0	0,022373
14	Matcharicano	1665000.0	0,022373
15	Es Teler Creamy Original	1580000.0	0,020045
...			
24	Apple Tea	848000.0	0,000000

4.4. Data Mining

Pada tahap pengolahan data, algoritma K-Means digunakan untuk mengklusterisasikan data yang telah dinormalisasi sebelumnya. Sebelum algoritma dijalankan, dilakukan pengujian untuk menentukan berapa banyak cluster yang memiliki karakteristik data yang paling cocok. Metode Elbow digunakan untuk menentukan nilai WCSS.



Gambar 2. Penentuan jumlah cluster dengan Grafik Elbow

Hasil grafik Elbow menunjukkan titik siku pada K=3, sehingga terdapat tiga cluster yang ditetapkan. Selanjutnya, algoritma K-Means diterapkan untuk mengelompokkan produk berdasarkan total nilai penjualan yang dihasilkan dari proses scaling.

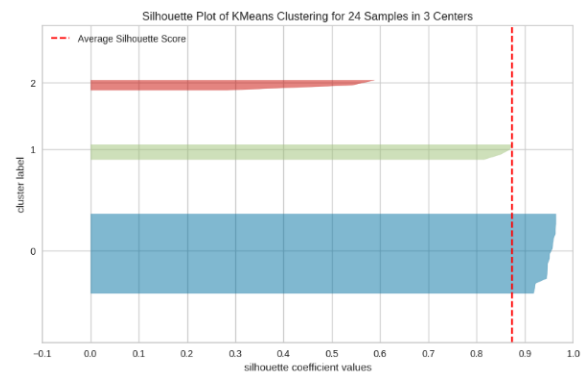
4.5. Evaluasi

Untuk mengevaluasi kualitas pembentukan cluster, Silhouette Score digunakan untuk menilai hasil klasterisasi. Nilai-nilai yang dihasilkan dari perhitungan ini disajikan dalam tabel 4 berikut.

Tabel 4. Perbandingan nilai Silhouette Score

Clustering K	Silhouette Score
2	0,85
3	0,87
4	0,86

Jumlah cluster yang digunakan adalah tiga karena nilai K=3 lebih besar. Nilai 0,87 termasuk dalam kategori sangat baik, yang menunjukkan bahwa struktur klasterisasi memiliki tingkat pemisah yang jelas dan minim tumpang tindih antar cluster.



Gambar 3. Grafik Silhouette Coefficient pada K-Means dengan K=3

Berdasarkan grafik Silhouette pada K=3, terlihat bahwa sebagian besar data memiliki nilai koefisien mendekati angka 1 dan tidak terdapat nilai negatif yang signifikan. Hal tersebut menunjukkan bahwa sebagian besar data telah tergolong dengan baik pada cluster masing-masing sesuai karakteristiknya. Garis vertikal pada grafik menunjukkan bahwa model klasterisasi yang digunakan memiliki kualitas pemisah yang optimal, dengan nilai rata-rata Silhouette sebesar 0,87. Oleh karena itu, hasil evaluasi tersebut menunjukkan bahwa algoritma K-Means dengan K=3 mampu menghasilkan pengelompokan data penjualan yang stabil dan representatif.

4.6. Interpretasi

Hasil klasterisasi algoritma K-Means dengan tiga cluster menunjukkan bahwa produk minuman terbagi menjadi tiga kategori utama berdasarkan total nilai penjualan. Produk dalam setiap kategori ini diuraikan pada tabel-tabel berikut.

Tabel 5. Data cluster 1 dengan kategori *best seller*

No	Menu	Total	Cluster	Kategori
1	Oreo Milk	37365000	C1	Best Seller
2	Mango Yakult	37128000	C1	Best Seller
3	Brown Sugar Boba	33612000	C1	Best Seller
4	Matcha Frappe	32916000	C1	Best Seller

Cluster 1 merupakan kelompok produk dengan kontribusi penjualan tertinggi (*best seller*). Produk

dalam kategori ini menjadi penyumbang utama pendapatan usaha karena dibandingkan dengan cluster lainnya, tingkat permintaannya paling tinggi. Tingginya volume penjualan menunjukkan bahwa produk pada kelompok ini memiliki daya tarik pasar yang kuat serta tingkat perputaran stok yang cepat. Oleh karena itu, pengelolaan persediaan pada cluster ini perlu menjadi prioritas dengan memastikan ketersediaan bahan baku secara konsisten agar tidak terjadi kekosongan stok yang berpotensi menyebabkan kehilangan peluang penjualan dan menurunkan kepuasan pelanggan.

Tabel 6. Data cluster 2 dengan kategori *profit driver*

No	Menu	Total	Cluster	Kategori
1	Matcha	22300000	C2	Profit Driver
2	Matcha Milkshake	13275000	C2	Profit Driver
3	Thai Tea	12710000	C2	Profit Driver

Cluster 2 dikategorikan sebagai produk dengan performa penjualan menengah (*profit driver*), yaitu produk yang memiliki tingkat permintaan yang relatif stabil dan memberikan kontribusi pendapatan yang konsisten, namun belum setinggi produk pada Cluster 1. Dari sisi pengelolaan persediaan, produk dalam kelompok ini memerlukan pengendalian stok yang proporsional, tidak sebanyak produk unggulan tetapi tetap dijaga ketersediaannya agar tidak terjadi kekosongan yang dapat menghambat potensi peningkatan penjualan. Pola permintaan yang cenderung stabil memungkinkan perencanaan stok dilakukan berdasarkan rata-rata penjualan historis. Sehingga risiko kelebihan maupun kekurangan persediaan dapat diminimalkan. Selain itu, strategi promosi seperti *bundling* atau penawaran khusus dapat diterapkan untuk mendorong peningkatan volume penjualan tanpa menimbulkan risiko *overstock* yang signifikan.

Tabel 7. Data cluster 3 dengan kategori *slow moving*

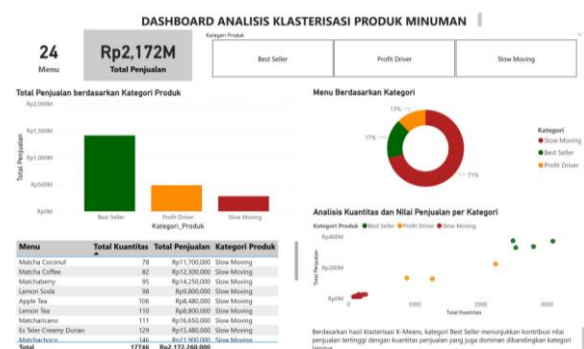
No	Menu	Total	Cluster	Kategori
1	Matcha Aren Brulee	2686000	C3	Slow Moving
2	Matcha Es Teler	2652000	C3	Slow Moving
3	Taro Milk	2630000	C3	Slow Moving
4	Matchachoco	2190000	C3	Slow Moving
5	Dark Choco	2070000	C3	Slow Moving
6	Matcha Latte	1911000	C3	Slow Moving
7	Matcharicano	1665000	C3	Slow Moving
8	Es Teler Creamy Original	1580000	C3	Slow Moving
9	Red Velvet	1580000	C3	Slow Moving
10	Es Teler Creamy Durian	1548000	C3	Slow Moving
11	Matchaberry	1425000	C3	Slow Moving

No	Menu	Total	Cluster	Kategori
12	Matcha Coffee	1230000	C3	Slow Moving
13	Matcha Coconut	1170000	C3	Slow Moving
14	Lemon Soda	980000	C3	Slow Moving
15	Lemon Tea	880000	C3	Slow Moving
16	Original Tea	875000	C3	Slow Moving
17	Apple Tea	848000	C3	Slow Moving

Cluster 3 dikategorikan sebagai kelompok produk dengan volume penjualan relatif rendah (*slow moving*). Produk dalam cluster ini menunjukkan tingkat permintaan yang lebih rendah sehingga perputaran stok cenderung lambat. Kondisi tersebut berpotensi meningkatkan risiko penumpukan persediaan apabila tidak dikelola dengan baik. Hasil pengelompokan ini dapat digunakan oleh manajemen sebagai dasar evaluasi untuk meninjau kembali strategi terhadap produk tersebut, baik melalui inovasi varian rasa, penyesuaian harga, peningkatan promosi, maupun pertimbangan pergantian menu dengan produk yang lebih sesuai dengan preferensi pasar.

4.7. Visualisasi Hasil Klasterisasi menggunakan Power BI

Hasil klasterisasi selanjutnya divisualisasikan menggunakan Power BI untuk mempermudah analisis dan interpretasi data. Dashboard menampilkan ringkasan jumlah produk dan total penjualan secara keseluruhan, distribusi berdasarkan kategori cluster, serta perbandingan kuantitas dan nilai penjualan antar kategori. Penyajian dalam bentuk grafik dan diagram interaktif ini membantu memahami pola kontribusi setiap kelompok produk secara lebih jelas sebelum dilakukan pembahasan lebih lanjut terhadap hasil visualisasi seperti pada Gambar 4 di bawah ini.



Gambar 4. Dashboard visualisasi hasil klasterisasi produk menggunakan Power BI

Berdasarkan hasil visualisasi, kategori *best seller* menunjukkan kontribusi pendapatan terbesar dengan total penjualan sebesar Rp 2.172.000 atau sekitar 71% dari keseluruhan penjualan. Sementara itu, kategori *profit driver* menyumbang sekitar 17% dan *slow moving* sebesar 12%. Proporsi ini memperlihatkan bahwa sebagian besar pendapatan usaha terpusat pada kelompok produk tertentu. Grafik scatter yang ditampilkan menunjukkan bahwa produk dalam

kategori *best seller* memiliki kuantitas dan total penjualan yang tinggi, sehingga mencerminkan tingkat permintaan pasar yang kuat. Sebaliknya, produk dalam kategori *slow moving* berada pada rentang kuantitas dan nilai penjualan yang rendah, yang mengindikasikan perlunya evaluasi strategi pemasaran maupun pengendalian persediaan. Fitur filter kategori yang tersedia juga memungkinkan pengguna melakukan analisis secara lebih spesifik terhadap masing-masing kelompok produk. Dengan demikian, visualisasi menggunakan Power BI tidak hanya menyajikan hasil klasterisasi yang informatif, tetapi juga mendukung proses pengambilan keputusan berbasis data yang lebih terarah.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, data penjualan produk minuman periode April-Oktober 2025 berhasil dikelompokkan menjadi tiga cluster menggunakan algoritma K-Means. Evaluasi menggunakan Silhouette menghasilkan nilai sebesar 0,87 yang menunjukkan kualitas pemisahan cluster yang sangat baik. Cluster pertama terdiri dari produk *best seller* dengan kontribusi penjualan tertinggi yaitu Oreo Milk, Mango Yakult, Brown Sugar Boba, dan Matcha Frappe. Cluster kedua mencakup produk *profit driver* dengan tingkat penjualan menengah dan relatif stabil, seperti Matcha, Matcha Milkshake, dan Thai Tea. Sementara itu cluster ketiga merupakan kelompok produk *slow moving* dengan tingkat penjualan relatif rendah, diantaranya Matcha Aren Brulee, Matcha Es Teler, Taro Milk, Matchachoco dan beberapa varian lainnya. Hasil pengelompokkan ini menunjukkan adanya perbedaan karakteristik kontribusi penjualan antarproduk, sehingga hasilnya dapat dimanfaatkan sebagai dasar dalam penyusunan strategi promosi dan pengelolaan persediaan secara lebih terarah dan efektif. Sangat disarankan untuk memasukkan variabel potensial tambahan ke dalam penelitian yang akan datang, seperti jumlah transaksi, frekuensi pembelian atau margin keuntungan, sehingga hasil pengelompokkan dapat memberikan gambaran yang lebih komprehensif. Selain itu, penelitian berikutnya dapat membandingkan algoritma K-Means dengan metode clustering lain untuk mengevaluasi konsistensi dan kualitas hasil pengelompokkan. Pengembangan visualisasi interaktif yang lebih mendalam juga dapat dilakukan guna mendukung analisis yang lebih detail dalam pengambilan keputusan bisnis berbasis data.

DAFTAR PUSTAKA

- [1] Mh. Siregar, I. Kuantan Singingi Jl Gatot Subroto Km, And T. Kuantan, "Klasterisasi Penjualan Alat-Alat Bangunan Menggunakan Metode K-Means (Studi Kasus Di Toko Adi Bangunan)".
- [2] D. Rusvinasari And L. H. Annisa, "Klasterisasi Pola Penjualan Menu Makanan Pada Rumah Makan Menggunakan Metode K-Means Clustering," Jurnal Informatika: Jurnal Pengembangan It, Vol. 10, No. 2, Pp. 398–409, Apr. 2025, Doi: 10.30591/Jpit.V10i2.8511.
- [3] Awaljan, Tukino, Novalia Elfina, And Ahmad Sandi, "Klasifikasi Hasil Penjualan Minuman Ringan Pada Koperasi Berdasarkan Jenis Barang Menggunakan Algoritma K-Means Clustering," 2022.
- [4] A. D. Juwari, "Pengelompokan Produk Penjualan Menggunakan K-Means Sebagai Pendukung Strategi Bisnis Kafe Omah Kopi," Jurnal Profesi Insinyur Universitas Lampung, Vol. 6, No. 2, Sep. 2025, Doi: 10.23960/Jpi.V6n2.175.
- [5] N. S. Aji, F. Natsir, And S. Istianah, "Penentuan Penjualan Barang Berdasarkan Pengelompokan Produk Dengan K-Means Clustering Metode Crisp-Dm," 2023.
- [6] M. Djaka Permana, A. Lia Hananto, E. Novalia, B. Huda, And T. Paryono, "Klasterisasi Data Jamaah Umrah Pada Tanurmutmainah Tour Menggunakan Algoritma K-Means," Jurnal Komtekinfo, Pp. 15–20, Feb. 2023, Doi: 10.35134/Komtekinfo.V10i1.332.
- [7] M. Noval, F. Dona Marleny, T. Informatika, U. Muhammadiyah Banjarmasin, And K. Selatan, "Implementasi Algoritma K-Means Untuk Analisis Pola Penjualan Pada Toko Monisa," 2025.
- [8] A. Yani, Z. Azmi, D. Suherdi, S. Informasi, And S. Triguna Dharma, "Implementasi Data Mining Menganalisa Data Penjualan Menggunakan Algoritma K-Means Clustering," Maret, Vol. 2, No. 2, Pp. 315–323, 2023, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jsi>
- [9] N. Noptapia, "Penerapan Data Mining Untuk Menentukan Strategi Penjualan Sparepart Mobil Pada Pt. Tri Satya Manggala Medan Menggunakan Metode K-Means Clustering," Jikteks : Jurnal Ilmu Komputer Dan Teknologi Informasi, Vol. 02, No. 02, Pp. 10–22, 2024.
- [10] A. Nugraha, O. Nurdian, And G. Dwilestari, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Yana Sport," 2022.
- [11] S. Pujiono, R. Astuti, And F. Muhamad Basysyar, "Implementasi Data Mining Untuk Menentukan Pola Penjualan Produk Menggunakan Algoritma K-Means Clustering," 2024.
- [12] M. Atalya, A. Leza, W. Utami, P. Anugrah, And C. Dewi, "Prediksi Prestasi Siswa Smas Katolik Santo Yoseph Denpasar Berdasarkan Kedisiplinan Dan Tingkat Ekonomi Orang Tua Menggunakan Metode Knowledge Discovery In Database Dan Algoritma Regresi Linier Berganda," 2024.
- [13] D. Athallah Ubaid And M. Ichsan Farel Rachmad, "Klasifikasi Risiko Strok Menggunakan Algoritma Random Forest Dengan Teknik Knowledge Discovery In Database," Journal Computer Science, Vol. 4,

- No. 1, 2025, [Online]. Available: <https://www.kaggle.com/datasets/mahatiratushe>
- [14] C. Michael Lauw, M. Wisnu Alfiansyah, F. Nathania, A. Asmina Nurul Adha, And A. Nursila, "Penerapan Algoritma K-Means Dalam Aplikasi Business Intelligence Untuk Segmentasi Pelanggan," 2023. [Online]. Available: <http://E-Journal.Potensi-Utama.Ac.Id/Ojs/Index.Php/Itjournal/Index>
- [15] D. Dwi Susilo, S. Shofiah Hilabi, B. Priyatna, And E. Novalia, "Implementasi Data Mining Dalam Pengelompokan Data Pembelian Menggunakan Algoritma K-Means Pada Pt.Otomotif 1",
- [16] D. Rusvinasari And L. H. Annisa, "Klasterisasi Pola Penjualan Menu Makanan Pada Rumah Makan Menggunakan Metode K-Means Clustering," *Jurnal Informatika: Jurnal Pengembangan It*, Vol. 10, No. 2, Pp. 398–409, Apr. 2025, Doi: 10.30591/Jpit.V10i2.8511.
- [17] A. Lia Hananto Et Al., "Dirgamaya Jurnal Manajemen Dan Sistem Informasi Strategi Promosi Penerapan Data Mining Mahasiswa Baru Dengan Metode K-Means Clustering."
- [18] D. S. Nugraha, I. Thoib, N. Sururi, F. Bayu F, And B. P. Candra, "Segmentasi Pelanggan Berbasis Analisis Rfm Menggunakan Algoritma K-Means Clustering," *Jurnal Komputer Teknologi Informasi Sistem Informasi (Juktisi)*, Vol. 4, No. 2, Pp. 1361–1369, Sep. 2025, Doi: 10.62712/Juktisi.V4i2.641.
- [19] A. Prayoga Rizkyandri And Y. Arvita, "Jurnal Informatika Dan Rekayasa Komputer (Jakakom) Implementasi Metode K-Means Clustering Untuk Menentukan Persediaan Barang Pada Toko Ss Babyshop," Vol. 3, No. 2, 2023, Doi: 10.33998/Jakakom.V3i2.
- [20] R. S. Wicaksana, D. Heksaputra, T. A. Syah, And F. F. Nur'aini, "Pendekatan K-Means Clustering Metode Elbow Pada Analisis Motivasi Pengunjung Festival Halal Jhff#2," *Jurnal Ilmiah Ekonomi Islam*, Vol. 9, No. 3, P. 4162, Nov. 2023, Doi: 10.29040/Jiei.V9i3.10591.
- [21] I. Wahyudi, M. Burhanis Sulthan, And L. Suhartini, "Analisa Penentuan Cluster Terbaik Pada Metode K-Means Menggunakan Elbow Terhadap Sentra Industri Produksi Di Pamekasan," 2021.
- [22] A. A. Jabar, R. Farta Wijaya, And S. Wahyuni, "Dashboard Visualisasi Data Kecerdasan Bisnis Untuk Mendukung Pengambilan Keputusan Bisnis Pada Pt Bmpt," *Jurnal Teknologi Informasi*, Vol. 6, No. 1, 2025, Doi: 10.46576/Djtechno.
- [23] A. Irfanullah, L. Janzen Sinay, And C. Fathia Palembang, "Peran Analisis Data Dalam Pengembangan Dashboard Visualisasi Data Prodams Pemerintah Kota Kediri," *Jurnal Pengabdian Kepada Masyarakat*, Vol. 14, No. 2, Pp. 323–330, Jun. 2023, [Online]. Available: <http://Journal.Upgris.Ac.Id/Index.Php/E-Dimas>
- [24] E. G. Limbor, T. A. Putri, S. Aurannisa, R. Qadriah, D. Arisandi, And J. T. Beng, "Perancangan Dashboard Penjualan Pakaian Di Toko Cleo Dengan Power Bi," 2024.
- [25] Awaljan, Tukino, E. Novalia, And S. Ahmad, "Klasifikasi Hasil Penjualan Minuman Ringan Pada Koperasi Berdasarkan Jenis Barang Menggunakan Algoritma K-Means Clustering," Dec. 2022