

Segmentasi Produk Berbasis Data Penjualan Menggunakan *K-Means Clustering* Pada *Coffee Shop Akarkata Kopi*

Budi Kurniawan ^{1,*}, Fatmawati ²

Program Studi Sistem Informasi; Universitas Bina Sarana Informatika; Jl. Kramat Raya No.98, RT.2/RW.9, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10450
Telp: (021) 8000063; e-mail: bukurwaelah@gmail.com, fatmawati.fmw@bsi.ac.id

* Korespondensi: e-mail: bukurwaelah@gmail.com

Diterima: 16 Agustus 2025 ; Review: 20 Desember 2025; Disetujui: 30 Desember 2025

Cara sitasi: Kurniawan B, Fatmawati. 2025. Segmentasi Produk Berbasis Data Penjualan Menggunakan *K-Means Clustering* Pada *Coffee Shop Akarkata Kopi*. Bina Insani ICT Journal. Vol 12 (2): 169 - 180.

Abstrak: Akarkata Kopi dihadapkan pada keterbatasan pasar karena lokasinya berada di kawasan perkantoran, sehingga optimalisasi strategi penjualan menjadi penting untuk meningkatkan daya saing usaha. Penelitian ini bertujuan melakukan segmentasi produk berbasis data penjualan menggunakan algoritma *K-Means Clustering* untuk mendukung pengambilan keputusan manajerial. Data yang digunakan berasal dari penjualan 30 produk selama periode Oktober hingga Desember 2024. Proses pengelompokan dilakukan menggunakan *Microsoft Excel* dengan penentuan jumlah kluster sebanyak tiga kelompok. Kualitas hasil clustering dievaluasi menggunakan nilai *Sum of Squared Error (SSE)* untuk mengukur homogenitas dalam kluster serta stabilitas centroid antar iterasi. Hasil penelitian menunjukkan terbentuknya tiga segmen produk, yaitu produk dengan tingkat penjualan tinggi, menengah, dan rendah. Hasil segmentasi ini memungkinkan manajemen untuk melakukan alokasi sumber daya promosi dan pengembangan produk secara lebih proporsional berdasarkan potensi penjualan pada setiap segmen.

Kata kunci: Akarkata Kopi, *k-means*, penjualan, segmentasi, strategi produk

Abstract: Akarkata Kopi faces market limitations due to its location in an office area, making sales strategy optimization essential to enhance business competitiveness. This study aims to perform product segmentation based on sales data using the *K-Means Clustering* algorithm to support managerial decision-making. The dataset consists of sales data from 30 products collected during the period of October to December 2024. The clustering process was conducted using *Microsoft Excel* with three clusters determined. The quality of the clustering results was evaluated using the *Sum of Squared Error (SSE)* to measure intra-cluster homogeneity and centroid stability across iterations. The results indicate the formation of three product segments, namely high-, medium-, and low-selling products. This segmentation enables management to allocate promotional resources and product development efforts more proportionally based on the sales potential of each segment.

Keywords: Akarkata Kopi, *k-means*, product strategy, sales, segmentation

1. Pendahuluan

Pertumbuhan bisnis *coffee shop* di Indonesia menunjukkan tren yang signifikan seiring dengan meningkatnya minat masyarakat terhadap konsumsi kopi. Fenomena ini mendorong munculnya berbagai usaha *coffee shop* di berbagai wilayah, termasuk kawasan perkantoran. Salah satu usaha yang tengah berkembang adalah Akarkata Kopi, yang berlokasi di kawasan Ruko Grogol Permai Blok B-11, Jakarta Barat. Perubahan gaya hidup masyarakat urban,

seperti meningkatnya budaya nongkrong serta preferensi terhadap variasi produk dan suasana tempat, menjadi faktor pendorong utama pertumbuhan bisnis *coffee shop* di wilayah perkotaan [1].

Meskipun memiliki potensi pasar dari kalangan pekerja dan masyarakat sekitar, Akarkata Kopi menghadapi tantangan berupa keterbatasan jangkauan pasar akibat kondisi geografis dan segmentasi konsumen yang relatif homogen. Dalam kondisi tersebut, optimalisasi strategi penjualan dan pengelolaan menu menjadi penting guna meningkatkan daya saing usaha dan efisiensi operasional.

Salah satu pendekatan yang dapat digunakan untuk mendukung pengambilan keputusan strategis adalah analisis data penjualan. Analisis ini memungkinkan pelaku usaha untuk memahami performa produk, mengidentifikasi produk dengan tingkat penjualan tinggi, menengah, maupun rendah, serta merancang strategi promosi dan pengembangan produk secara lebih terarah [2]. Sebagian besar penelitian sebelumnya dalam bidang data mining pada sektor ritel dan *food and beverage* berfokus pada segmentasi pelanggan, baik berdasarkan karakteristik demografis maupun perilaku pembelian. Sementara itu, penelitian yang secara khusus membahas segmentasi produk terutama pada usaha mikro dan kecil yang beroperasi di kawasan perkantoran masih relatif terbatas. Oleh karena itu, terdapat *research gap* dalam pemanfaatan data penjualan untuk melakukan segmentasi produk sebagai dasar pengambilan keputusan bisnis pada skala usaha kecil.

Penelitian ini bertujuan untuk mengisi celah tersebut dengan mengimplementasikan algoritma *K-Means Clustering* dalam melakukan segmentasi produk berdasarkan kinerja penjualan di Akarkata Kopi. Fokus penelitian diarahkan pada segmentasi produk, bukan pelanggan, sehingga diharapkan dapat memberikan perspektif yang berbeda dalam optimalisasi pengelolaan menu dan strategi promosi pada usaha *coffee shop* berskala mikro.

Dalam penelitian ini, jumlah kluster ditentukan sebanyak tiga kelompok yang merepresentasikan produk dengan tingkat penjualan tinggi, menengah, dan rendah. Penentuan jumlah kluster ini didasarkan pada pertimbangan praktis dan kebutuhan analisis manajerial. Namun demikian, penelitian ini memiliki keterbatasan metodologis karena belum menggunakan metode evaluasi jumlah kluster seperti *Elbow Method* atau *Silhouette Coefficient*, yang dapat menjadi peluang pengembangan penelitian selanjutnya.

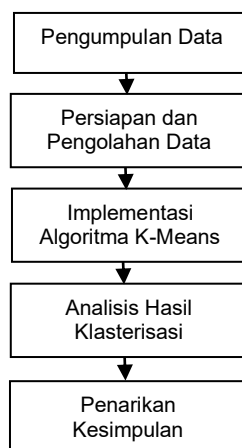
Adapun alur penelitian meliputi pengumpulan data penjualan selama tiga bulan, pengolahan data menggunakan algoritma *K-Means Clustering*, analisis hasil segmentasi produk, serta perumusan rekomendasi strategi berdasarkan kluster yang terbentuk. Hasil segmentasi diharapkan dapat menjadi dasar dalam merancang strategi promosi dan pengembangan produk yang lebih efektif dan proporsional sesuai dengan karakteristik performa masing-masing kelompok produk [3].

Beberapa penelitian sebelumnya menunjukkan bahwa algoritma *K-Means Clustering* efektif digunakan dalam analisis data penjualan untuk segmentasi produk. Studi terdahulu berhasil mengelompokkan produk fashion berdasarkan volume penjualan ke dalam beberapa kluster [4], serta mengklasifikasikan pola penjualan makanan dan minuman pada restoran ke dalam kategori tinggi, sedang, dan rendah [5]. Penelitian lain juga menerapkan *K-Means* untuk mengelompokkan kualitas produk dengan hasil nilai *Davies-Bouldin Index* kurang dari satu, yang menunjukkan kluster yang kompak dan efisien [6]. Temuan-temuan tersebut memperkuat relevansi penggunaan algoritma *K-Means Clustering* dalam penelitian ini.

2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan metode data mining melalui algoritma *K-Means Clustering*. Pendekatan ini bertujuan untuk mengelompokkan produk Akarkata Kopi berdasarkan kinerja penjualannya selama periode Oktober hingga Desember 2024. Implementasi algoritma dilakukan secara manual menggunakan *Microsoft Excel*, mengingat kesederhanaan metode serta kesesuaiannya dengan konteks usaha mikro yang tidak memerlukan perangkat lunak pemrograman yang kompleks.

Tahapan pelaksanaan penelitian dilakukan secara sistematis sebagai berikut:



Sumber: Hasil penelitian (2025)

Gambar 1. Kerangka Penelitian

1. Pengumpulan Data
Data diperoleh dari pencatatan transaksi penjualan Akarkata Kopi selama periode Oktober hingga Desember 2024.
2. Persiapan dan Pengolahan Data
Data penjualan disusun dalam bentuk tabular, mencakup nama produk, kategori produk, harga satuan, jumlah penjualan per bulan, serta total penjualan.
3. Implementasi Algoritma *K-Means Clustering*
Pengelompokan produk dilakukan secara manual di *Microsoft Excel* menggunakan fungsi-fungsi matematis dan logika.
4. Analisis Hasil Klasterisasi
Hasil pengelompokan dianalisis untuk mengidentifikasi segmen produk berdasarkan tingkat penjualan.
5. Penarikan Kesimpulan
Kesimpulan ditarik berdasarkan hasil klasterisasi guna memberikan rekomendasi strategi penjualan bagi Akarkata Kopi.

Kerangka metode ini mengadaptasi pendekatan yang telah digunakan dalam penelitian segmentasi produk furniture menggunakan *Microsoft Excel* secara manual [7].

Algoritma *K-Means Clustering*

Algoritma *K-Means Clustering* merupakan metode pengelompokan data yang bertujuan untuk meminimalkan jarak antara objek data dengan pusat kluster (centroid). Tahapan algoritma meliputi penentuan jumlah kluster (K), inialisasi centroid awal, perhitungan jarak menggunakan rumus *Euclidean*, pengelompokan data ke dalam kluster terdekat, serta pembaruan centroid secara iteratif hingga mencapai kondisi konvergen [8].

Adapun langkah-langkah implementasi dalam penelitian ini adalah sebagai berikut:

1. Menentukan jumlah kluster (K) sebanyak tiga.
2. Inialisasi centroid awal dengan memilih data secara acak.
3. Perhitungan jarak *Euclidean*:

$$d = \sqrt{(x_1 - c_1)^2 + (x_2 - c_2)^2 + \dots + (x_n - c_n)^2}$$

Implementasi dilakukan menggunakan fungsi *SQRT* dan *POWER* di *Excel*.

4. Mengelompokkan data ke dalam kluster dengan jarak minimum menggunakan fungsi *MIN*.
5. Memperbarui centroid menggunakan fungsi *AVERAGEIF*.
6. Melakukan iterasi hingga perubahan nilai centroid kurang dari 0,001, yang digunakan sebagai kriteria konvergensi.

Prosedur ini diadaptasi dari metode yang telah dijelaskan pada studi sebelumnya [9].

Akuisisi dan Struktur Data

Data penelitian diperoleh dari catatan penjualan Akarkata Kopi selama periode Oktober hingga Desember 2024. Data bersifat kuantitatif dan disusun dalam tabel yang mencakup nama produk, kategori produk, harga satuan, jumlah penjualan bulanan, serta total penjualan.

Variabel yang digunakan dalam proses clustering dibatasi pada indikator kinerja penjualan, yaitu total unit terjual dan total nilai penjualan selama tiga bulan. Harga satuan tidak digunakan sebagai variabel clustering, melainkan sebagai informasi kontekstual dalam interpretasi hasil klusterisasi.

Pengolahan Data

Pengolahan data yang sistematis sangat penting untuk mendukung efektivitas algoritma *K-Means* dalam menemukan pola konsumsi [10]. Tahapan pengolahan data meliputi:

1. Kompilasi data penjualan tiga bulan ke dalam satu tabel utama.
2. Normalisasi data menggunakan metode *Min-Max Normalization* untuk menyamakan skala antar variabel.
3. Pemilihan variabel kinerja penjualan yang relevan untuk proses clustering.
4. Implementasi algoritma *K-Means* secara manual di *Excel*:
 - a. Menghitung jarak antar data dan centroid.
 - b. Mengelompokkan data berdasarkan jarak minimum.
 - c. Memperbarui centroid untuk tiap kluster.
 - d. Melakukan iterasi hingga hasil stabil.

Penentuan Jumlah Kluster

Jumlah kluster ditetapkan sebanyak tiga tanpa menggunakan metode validasi statistik seperti *Elbow Method* atau *Silhouette Coefficient*. Pemilihan nilai $K = 3$ didasarkan pada pertimbangan praktis, yaitu kemudahan interpretasi hasil bagi pelaku UMKM serta kebutuhan untuk mengelompokkan produk ke dalam kategori penjualan tinggi, menengah, dan rendah. Meskipun demikian, pemilihan jumlah kluster ini menjadi keterbatasan metodologis penelitian dan dapat dikembangkan lebih lanjut pada penelitian selanjutnya dengan menambahkan metode evaluasi jumlah kluster yang lebih objektif.

3. Hasil dan Pembahasan

Data penjualan yang digunakan dalam penelitian ini berasal dari laporan transaksi Akarkata Kopi selama periode Oktober hingga Desember 2024. Data awal mencakup sebanyak 47 produk dengan variasi kategori, harga, dan jumlah penjualan. Setelah dilakukan proses data *cleansing* untuk menghapus entri produk yang tidak memiliki transaksi selama periode tersebut, hanya 30 produk yang memenuhi kriteria kelengkapan data dan digunakan dalam proses analisis.

Setiap produk dianalisis berdasarkan dua variabel utama, yaitu total unit terjual dan total nilai penjualan (dalam Rupiah) selama tiga bulan. Kedua variabel ini dianggap mewakili performa penjualan dari sisi kuantitas dan nilai ekonomi produk. Selain itu, data harga satuan juga digunakan sebagai variabel tambahan untuk memperkuat hasil segmentasi produk berdasarkan rentang harga dan volume penjualan. Berikut data penjualan selama periode oktober-desember 2024:

Tabel 1. Data Penjualan Sebelum Pembersihan

Nama Produk	Total Penjualan	Total Nilai
Americano	66	1,254,000
Biscoff Latte	53	1,325,000
Cadbury Melt	1	30,000
Caffé Latte	195	4,290,000
Cappuccino	68	1,496,000
Caramel Macchiato	4	128,000
Chicken Popcorn	18	360,000
Earl Grey Macchiato	0	-
Es Bandung	0	-
Flavoured Caffé Latte	2	64,000
Flavoured Tea	40	1,000,000
French Fries Shake	39	702,000
Frozen Chocolate	0	-
Hojicha Latte	0	-
Honey Lemon	4	100,000
Iced Coffee Mint	0	-
Iced Shaken Latte	0	-
Java Tea	71	1,562,000
Kopi Susu Gula Aren	611	15,275,000

Lemon Yakult	6	162,000
Lemongrass Honey Tea	0	-
Lychee Yakult Blend	4	116,000
Mango Freeze	0	-
Mango Yakult Blend	3	87,000
Matcha Espresso	0	-
Matcha Freeze	0	-
Matcha Latte	56	1,400,000
Mocha	6	168,000
Nasi Goreng	12	336,000
Nasi Putih	5	30,000
Ovaltine Macchiato	0	-
Pandan Coffee	3	84,000
Pink Lady	0	-
Pisang Goreng	37	185,000
Regally Blend	0	-
Salted Caramel	3	105,000
Sausage Platter	5	175,000
Signature Chocolate	81	2,025,000
Soto Ayam	0	-
Strawberry Cooler	0	-
Strawberry Freeze	6	180,000
Tahu Cabe Garam	9	180,000
Tahu Goreng	4	72,000
Taro Crunchy	0	-
Taro Espresso	0	-
Tea Selection	10	120,000
Truffle Fries	5	100,000

Sumber: Hasil Penelitian (2025)

Tabel 2. Data Penjualan Sesudah pembersihan

Nama Produk	Total Penjualan	Total Nilai
Americano	66	1,254,000
Biscoff Latte	53	1,325,000
Cadbury Melt	1	30,000
Caffé Latte	195	4,290,000
Cappuccino	68	1,496,000
Caramel Macchiato	4	128,000
Chicken Popcorn	18	360,000
Flavoured Caffé Latte	2	64,000
Flavoured Tea	40	1,000,000
French Fries Shake	39	702,000
Honey Lemon	4	100,000
Java Tea	71	1,562,000
Kopi Susu Gula Aren	611	15,275,000
Lemon Yakult	6	162,000
Lychee Yakult Blend	4	116,000
Mango Yakult Blend	3	87,000
Matcha Latte	56	1,400,000
Mocha	6	168,000
Nasi Goreng	12	336,000
Nasi Putih	5	30,000
Pandan Coffee	3	84,000
Pisang Goreng	37	185,000
Salted Caramel	3	105,000
Sausage Platter	5	175,000
Signature Chocolate	81	2,025,000
Strawberry Freeze	6	180,000
Tahu Cabe Garam	9	180,000
Tahu Goreng	4	72,000
Tea Selection	10	120,000
Truffle Fries	5	100,000

Sumber: Hasil Penelitian (2025)

Normalisasi dan Inisialisasi Centroid

Karena rentang nilai unit penjualan dan nilai penjualan berbeda jauh, maka dilakukan normalisasi menggunakan metode *Min-Max*. Setelah itu, centroid awal ditentukan secara acak untuk memulai proses iterasi.

Tabel 3. Proses Normalisasi Data dan Inisialisasi Centroid Awal Iterasi 1

Nama Produk	Total Penjualan	Total Nilai	Normalisasi total penjualan	Normalisasi total Nilai
Americano	66	1,254,000	0,11	0,08
Biscoff Latte	53	1,325,000	0,09	0,08
Cadbury Melt	1	30,000	0,00	0,00
Caffé Latte	195	4,290,000	0,32	0,28
Cappuccino	68	1,496,000	0,11	0,10
Caramel Macchiato	4	128,000	0,00	0,01
Chicken Popcorn	18	360,000	0,03	0,02
Flavoured Caffé Latte	2	64,000	0,00	0,00
Flavoured Tea	40	1,000,000	0,06	0,06
French Fries Shake	39	702,000	0,06	0,04
Honey Lemon	4	100,000	0,00	0,00
Java Tea	71	1,562,000	0,11	0,10
Kopi Susu Gula Aren	611	15,275,000	1,00	1,00
Lemon Yakult	6	162,000	0,01	0,01
Lychee Yakult Blend	4	116,000	0,00	0,01
Mango Yakult Blend	3	87,000	0,00	0,00
Matcha Latte	56	1,400,000	0,09	0,09
Mocha	6	168,000	0,01	0,01
Nasi Goreng	12	336,000	0,02	0,02
Nasi Putih	5	30,000	0,01	0,00
Pandan Coffee	3	84,000	0,00	0,00
Pisang Goreng	37	185,000	0,06	0,01
Salted Caramel	3	105,000	0,00	0,00
Sausage Platter	5	175,000	0,01	0,01
Signature Chocolate	81	2,025,000	0,13	0,13
Strawberry Freeze	6	180,000	0,01	0,01
Tahu Cabe Garam	9	180,000	0,01	0,01
Tahu Goreng	4	72,000	0,00	0,00
Tea Selection	10	120,000	0,01	0,01
Truffle Fries	5	100,000	0,01	0,00

Sumber: Hasil Penelitian (2025)

Berikut Adalah Tabel centroid awal yang ditentukan secara acak pada iterasi 1:

Tabel 4. Centroid Iterasi 1

C1		C2		C3	
0,03	0,02	0,09	0,09	0,13	0,13

Proses Iterasi dan Pembentukan Cluster

Proses iterasi dilakukan sebanyak enam kali hingga konvergen. Setiap iterasi melibatkan perhitungan jarak *Euclidean* antara data dan centroid, pengelompokan berdasarkan jarak terdekat, serta pembaruan centroid hingga tidak terjadi perubahan lagi dalam klusterisasi.

Tabel 5. Proses Iterasi Ke 6

Nama Produk	Total Penjualan	Total Nilai	Normalisasi total penjualan	Normalisasi total Nilai
Americano	66	1,254,000	0,11	0,08
Biscoff Latte	53	1,325,000	0,09	0,08
Cadbury Melt	1	30,000	0,00	0,00
Caffé Latte	195	4,290,000	0,32	0,28
Cappuccino	68	1,496,000	0,11	0,10
Caramel Macchiato	4	128,000	0,00	0,01
Chicken Popcorn	18	360,000	0,03	0,02
Flavoured Caffé Latte	2	64,000	0,00	0,00
Flavoured Tea	40	1,000,000	0,06	0,06
French Fries Shake	39	702,000	0,06	0,04
Honey Lemon	4	100,000	0,00	0,00
Java Tea	71	1,562,000	0,11	0,10

Kopi Susu Gula Aren	611	15,275,000	1,00	1,00
Lemon Yakult	6	162,000	0,01	0,01
Lychee Yakult Blend	4	116,000	0,00	0,01
Mango Yakult Blend	3	87,000	0,00	0,00
Matcha Latte	56	1,400,000	0,09	0,09
Mocha	6	168,000	0,01	0,01
Nasi Goreng	12	336,000	0,02	0,02
Nasi Putih	5	30,000	0,01	0,00
Pandan Coffee	3	84,000	0,00	0,00
Pisang Goreng	37	185,000	0,06	0,01
Salted Caramel	3	105,000	0,00	0,00
Sausage Platter	5	175,000	0,01	0,01
Signature Chocolate	81	2,025,000	0,13	0,13
Strawberry Freeze	6	180,000	0,01	0,01
Tahu Cabe Garam	9	180,000	0,01	0,01
Tahu Goreng	4	72,000	0,00	0,00
Tea Selection	10	120,000	0,01	0,01
Truffle Fries	5	100,000	0,01	0,00

Sumber: Hasil Penelitian (2025)

Berikut tabel 6 merupakan centroid yang ditentukan berdasarkan nilai rata-rata data yang sudah ditentukan cluster nya di iterasi pertama atau sebelumnya

Tabel 6. Centroid Iterasi 6

	C1	C2	C3
	0,02	0,01	0,14
		0,12	1,00
			1,00

Sumber: Hasil Penelitian (2025)

Berikut Nilai Jarak (*Euclidean Distance*) dari iterasi ke 6:

Tabel 7. Proses Iterasi ke 6 Nilai Jarak (*Euclidean Distance*)

C1	C2	C3	Penentuan kelas	N Error (Mencari perubahan kelas)
0,11442	0,0523	1,28223	2	Klaster Tetap
0,10156	0,06395	1,29387	2	Klaster Tetap
0,019	0,18387	1,41421	1	Klaster Tetap
0,40441	0,23951	0,99211	2	Klaster Tetap
0,12704	0,03796	1,26859	2	Klaster Tetap
0,01142	0,17592	1,40619	1	Klaster Tetap
0,01629	0,1487	1,37921	1	Klaster Tetap
0,01637	0,18116	1,41148	1	Klaster Tetap
0,07143	0,09388	1,32401	1	Klaster Tetap
0,05732	0,10846	1,33906	1	Klaster Tetap
0,01233	0,17714	1,40749	1	Klaster Tetap
0,1336	0,03143	1,26205	2	Klaster Tetap
1,39541	1,23063	0	3	Klaster Tetap
0,00752	0,17198	1,4023	1	Klaster Tetap
0,01179	0,17644	1,40675	1	Klaster Tetap
0,01417	0,17893	1,40925	1	Klaster Tetap
0,10852	0,05708	1,28691	2	Klaster Tetap
0,00739	0,17172	1,40202	1	Klaster Tetap
0,00912	0,15706	1,38727	1	Klaster Tetap
0,01431	0,17905	1,40958	1	Klaster Tetap
0,01428	0,17906	1,40939	1	Klaster Tetap
0,04383	0,13702	1,36573	1	Klaster Tetap
0,01357	0,17814	1,40842	1	Klaster Tetap
0,00885	0,17265	1,40285	1	Klaster Tetap
0,16648	0,00941	1,22894	2	Klaster Tetap
0,00718	0,1712	1,40146	1	Klaster Tetap
0,00261	0,16755	1,39798	1	Klaster Tetap
0,01344	0,17838	1,40879	1	Klaster Tetap
0,00552	0,16905	1,39962	1	Klaster Tetap
0,011	0,17593	1,40633	1	Klaster Tetap

Sumber: Hasil Penelitian (2025)

Hasil Akhir Segmentasi Produk

Setelah iterasi ke-6, hasil akhir menunjukkan pembagian produk ke dalam tiga cluster yang menggambarkan performa penjualan berbeda. Cluster 1 berisi produk dengan penjualan rendah, Cluster 2 menampilkan produk dengan penjualan sedang, dan Cluster 3 hanya terdiri dari satu produk dengan penjualan sangat tinggi

Tabel 8. Hasil Akhir Perhitungan dan Pembagian Cluster

Nama Produk	Total Penjualan	Total Nilai	Cluster Akhir
Americano	66	1,254,000	2
Biscoff Latte	53	1,325,000	2
Cadbury Melt	1	30,000	1
Caffé Latte	195	4,290,000	2
Cappuccino	68	1,496,000	2
Caramel Macchiato	4	128,000	1
Chicken Popcorn	18	360,000	1
Flavoured Caffé Latte	2	64,000	1
Flavoured Tea	40	1,000,000	1
French Fries Shake	39	702,000	1
Honey Lemon	4	100,000	1
Java Tea	71	1,562,000	2
Kopi Susu Gula Aren	611	15,275,000	3
Lemon Yakult	6	162,000	1
Lychee Yakult Blend	4	116,000	1
Mango Yakult Blend	3	87,000	1
Matcha Latte	56	1,400,000	2
Mocha	6	168,000	1
Nasi Goreng	12	336,000	1
Nasi Putih	5	30,000	1
Pandan Coffee	3	84,000	1
Pisang Goreng	37	185,000	1
Salted Caramel	3	105,000	1
Sausage Platter	5	175,000	1
Signature Chocolate	81	2,025,000	2
Strawberry Freeze	6	180,000	1
Tahu Cabe Garam	9	180,000	1
Tahu Goreng	4	72,000	1
Tea Selection	10	120,000	1
Truffle Fries	5	100,000	1

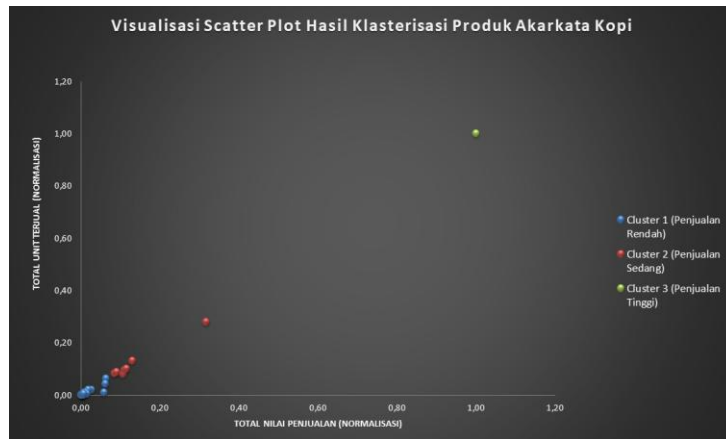
Sumber: Hasil Penelitian (2025)

Tabel 9. Jumlah Produk per Cluster Hasil K-Means

Cluster	Jumlah Produk	Keterangan
Cluster 1	22	Penjualan Rendah
Cluster 2	7	Penjualan Sedang
Cluster 3	1	Penjualan Tinggi

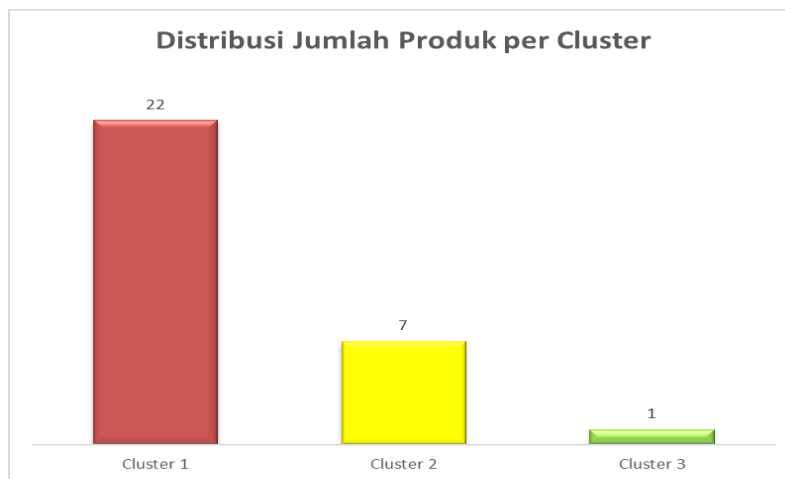
Sumber: Hasil Penelitian (2025)

Untuk memperjelas pemisahan antar klaster, dilakukan visualisasi data menggunakan *scatter plot* dua dimensi berdasarkan variabel total unit terjual dan total nilai penjualan yang telah dinormalisasi. Visualisasi ini menunjukkan bahwa sebagian besar produk pada Cluster 1 dan Cluster 2 memiliki distribusi yang relatif berdekatan, sedangkan Cluster 3 terpisah secara signifikan karena nilai penjualannya yang jauh lebih tinggi. Hal ini mengindikasikan adanya perbedaan karakteristik yang jelas antara produk unggulan dan produk lainnya.



Sumber: Hasil Penelitian (2025)

Gambar 1. Visualisasi *Scatter Plot* Hasil Klasterisasi Produk



Sumber: Hasil Penelitian (2025)

Gambar 2. Grafik Distribusi Jumlah Produk per *Cluster*

Analisis Per *Cluster*

Cluster 1 terdiri dari mayoritas produk (22 produk), dengan karakteristik penjualan rendah. Rata-rata produk di cluster ini memiliki unit penjualan kurang dari 10 per bulan. Faktor penyebab bisa berasal dari kurangnya promosi, harga tinggi, atau rendahnya popularitas. Untuk produk dalam *cluster* ini, strategi seperti diskon, *bundling*, atau promosi musiman dapat dipertimbangkan.

Cluster 2 berisi produk-produk dengan kontribusi penjualan menengah, misalnya *Americano*, *Caffé Latte*, dan *Signature Chocolate*. Produk ini menunjukkan potensi untuk ditingkatkan, baik dengan promosi tambahan atau penguatan identitas produk.

Cluster 3 hanya memuat satu produk, yaitu Kopi Susu Gula Aren. Dominasi produk Kopi Susu Gula Aren dalam *Cluster 3* menunjukkan adanya produk unggulan (*flagship product*) yang memiliki kontribusi penjualan jauh lebih tinggi dibandingkan produk lainnya. Fenomena ini dapat dipengaruhi oleh beberapa faktor, seperti preferensi konsumen terhadap cita rasa lokal yang manis dan familiar, kesesuaian harga dengan daya beli konsumen perkantoran, serta kemungkinan efek rekomendasi dari mulut ke mulut. Selain itu, produk ini berpotensi menjadi identitas utama (*signature product*) Akarkata Kopi yang mendorong pembelian berulang.

Meskipun beberapa produk seperti *Flavoured Tea* dan *French Fries Shake* memiliki jumlah unit penjualan yang relatif lebih tinggi dibandingkan produk lain di *Cluster 1*, keduanya tetap terklasifikasi dalam klaster penjualan rendah. Hal ini disebabkan oleh nilai penjualan total yang masih jauh lebih kecil dibandingkan produk pada *Cluster 2* dan terutama *Cluster 3*. Kondisi ini menunjukkan bahwa algoritma *K-Means Clustering* sangat sensitif terhadap keberadaan *outlier*, yaitu produk Kopi Susu Gula Aren yang memiliki nilai penjualan sangat

dominan. Akibatnya, posisi centroid kluster menjadi condong ke nilai ekstrem, sehingga produk dengan penjualan menengah namun nilai relatif rendah tetap terklasifikasi ke dalam kluster rendah. Fenomena ini merupakan karakteristik umum algoritma *K-Means* dan menjadi keterbatasan metodologis dalam penelitian ini.

Tabel 10. Produk di *Cluster* Pertama (Penjualan Rendah)

Nama Produk	Total Penjualan	Total Nilai	Cluster Akhir
Cadbury Melt	1	30,000	1
Caramel Macchiato	4	128,000	1
Chicken Popcorn	18	360,000	1
Flavoured Caffé Latte	2	64,000	1
Flavoured Tea	40	1,000,000	1
French Fries Shake	39	702,000	1
Honey Lemon	4	100,000	1
Lemon Yakult	6	162,000	1
Lychee Yakult Blend	4	116,000	1
Mango Yakult Blend	3	87,000	1
Mocha	6	168,000	1
Nasi Goreng	12	336,000	1
Nasi Putih	5	30,000	1
Pandan Coffee	3	84,000	1
Pisang Goreng	37	185,000	1
Salted Caramel	3	105,000	1
Sausage Platter	5	175,000	1
Strawberry Freeze	6	180,000	1
Tahu Cabe Garam	9	180,000	1
Tahu Goreng	4	72,000	1
Tea Selection	10	120,000	1
Truffle Fries	5	100,000	1

Sumber: Hasil Penelitian (2025)

Tabel 11. Produk di *Cluster* Kedua (Penjualan Sedang)

Nama Produk	Total Penjualan	Total Nilai	Cluster Akhir
Americano	66	1,254,000	2
Biscoff Latte	53	1,325,000	2
Caffé Latte	195	4,290,000	2
Cappuccino	68	1,496,000	2
Java Tea	71	1,562,000	2
Matcha Latte	56	1,400,000	2
Signature Chocolate	81	2,025,000	2

Sumber: Hasil Penelitian (2025)

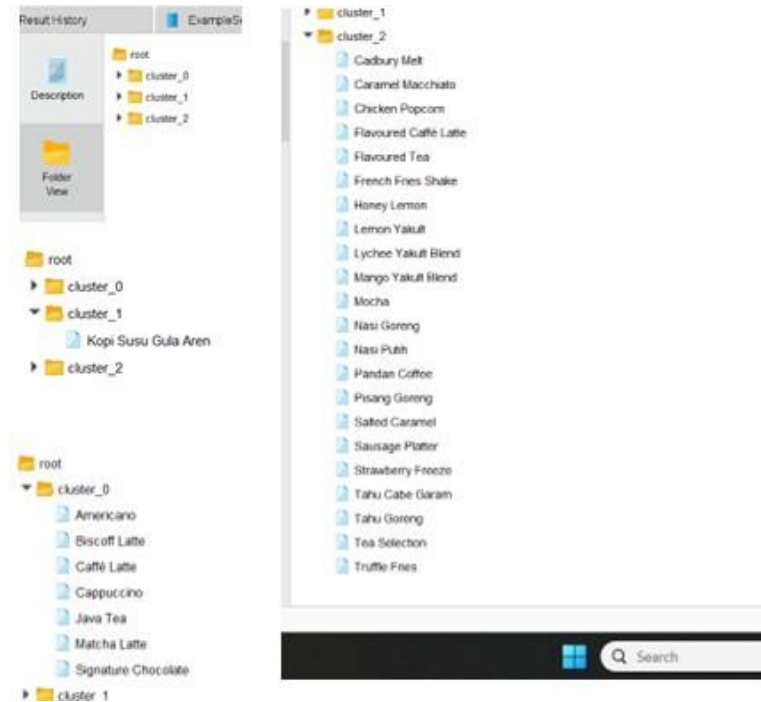
Tabel 12. Produk di *Cluster* Ketiga (Penjualan Tertinggi)

Nama Produk	Total Penjualan	Total Nilai	Cluster Akhir
Kopi Susu Gula Aren	611	15,275,000	3

Sumber: Hasil Penelitian (2025)

Pembandingan dengan *RapidMiner*

Sebagai validasi, hasil clustering manual dibandingkan dengan hasil dari aplikasi *RapidMiner*. Parameter yang digunakan sama: jumlah cluster ($K = 3$), metode *Euclidean Distance*, dan maksimal 6 iterasi. Hasil dari *RapidMiner* menunjukkan pembagian *cluster* identik dengan proses manual di *Excel*, meskipun label nama klasternya berbeda.



Sumber: Hasil Penelitian (2025)

Gambar 3. Proses dan Hasil di *Rapidminer*

Tabel 13. Penyesuaian Nama *Cluster*

<i>RapidMiner</i>	<i>Excel</i>
cluster_2	1
cluster_0	2
cluster_1	3

Sumber: Hasil Penelitian (2025)

Tabel 14. Total Item / *Cluster* (Perbandingan *RapidMiner* vs *Excel*)

<i>RapidMiner</i> (cluster)	Total Item	<i>Excel</i> (cluster)	Total Item
cluster_2	22	1	22
cluster_0	7	2	7
cluster_1	1	3	1

Sumber: Hasil Penelitian (2025)

4. Kesimpulan

Penelitian ini berhasil menjawab tujuan utama, yaitu melakukan segmentasi produk di Akarkata Kopi berdasarkan data penjualan menggunakan algoritma *K-Means Clustering*. Hasil penelitian menunjukkan bahwa metode ini mampu membagi produk ke dalam tiga klaster yang merepresentasikan tingkat performa penjualan yang berbeda, yaitu produk dengan penjualan tinggi, sedang, dan rendah.

Implementasi algoritma dilakukan secara manual menggunakan *Microsoft Excel*, dan hasil klasterisasi terbukti identik dengan pengolahan menggunakan perangkat lunak *RapidMiner*. Temuan ini menunjukkan bahwa algoritma *K-Means Clustering* dapat diadaptasi secara sederhana oleh pelaku UMKM dengan keterbatasan teknologi, selama didukung oleh pemahaman konsep analisis data dan ketersediaan data penjualan yang akurat. Secara teoretis, penelitian ini memperkaya literatur mengenai penerapan algoritma data mining dalam konteks UMKM, khususnya pada segmentasi produk yang masih relatif jarang dikaji dibandingkan segmentasi pelanggan.

Dari sisi praktis, hasil segmentasi produk memberikan manfaat langsung bagi manajemen Akarkata Kopi dalam mengidentifikasi produk unggulan, merancang strategi promosi yang lebih tepat sasaran, serta mengoptimalkan pengelolaan stok dan pengembangan menu. Produk dengan performa tinggi dapat dipertahankan dan dikembangkan, sementara

produk dengan performa rendah dapat dievaluasi melalui strategi promosi, bundling, atau inovasi produk.

Namun demikian, penelitian ini memiliki beberapa keterbatasan. Penelitian hanya menggunakan data penjualan tanpa mempertimbangkan faktor lain seperti margin keuntungan, biaya produksi, dan preferensi pelanggan, sehingga hasil segmentasi masih berfokus pada performa kuantitatif penjualan. Selain itu, hasil klusterisasi menunjukkan bahwa algoritma *K-Means* cukup sensitif terhadap distribusi data yang tidak seimbang, khususnya keberadaan produk dengan nilai penjualan yang sangat dominan (*outlier*).

Oleh karena itu, penelitian selanjutnya disarankan untuk mengintegrasikan variabel tambahan yang lebih komprehensif, menerapkan teknik transformasi data atau metode clustering alternatif seperti *K-Medoids*, serta menggunakan metode validasi kluster yang lebih objektif guna meningkatkan keandalan dan kedalaman analisis segmentasi produk.

Ucapan Terima Kasih

Terima kasih kepada pihak Akarkata Kopi atas kerjasama dan data yang diberikan, serta Universitas Bina Sarana Informatika atas dukungan akademik dan fasilitas yang menunjang proses penelitian ini.

Referensi

- [1] R. Y. Daulay, R. A. Passalaras, and J. Heikal, "Customer Segmentation Using K-Means Clustering with SPSS Program in a Case Study of Consumer Interest in Current Coffee Shop," *BUDGETING: Journal of Business, Management and Accounting*, vol. 5, no. 2, pp. 721–740, Apr. 2024, doi: 10.31539/budgeting.v5i2.9288.
- [2] S. Valencia and W. T. Atmojo, "Analisis Pola Pembelian pada Data Penjualan CanNgopi menggunakan Algoritma FP-Growth," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 8, no. 2, pp. 214–224, 2024, doi: 10.31603/komtika.v8i2.12672.
- [3] Y. Nuryaman, B. Sudarsono, U. Faddillah, and A. Asistyasari, "Klasifikasi Data Penjualan Menggunakan Algoritma K-Means Dan Analytic Hierarchy Process," *Jurnal Larik: Ladang Artikel Ilmu Komputer*, vol. 2, no. 1, pp. 23–28, 2022, doi: 10.31294/larik.v2i1.1371.
- [4] Normah; Nurajizah; Salbinda; Arinda, "Penerapan Data Mining Metode K-Means Clustering untuk Analisa Penjualan pada Toko Fashion Hijab Banten," *Jurnal Teknik Komputer AMIK BSI*, vol. 7, no. 2, pp. 158–163, 2021, doi: 10.31294/jtk.v7i2.10553.
- [5] E. S. Ana, S. Anwar, and M. H. Fuad, "Analisa Tingkat Penjualan Makanan Dan Minuman Dengan Klusterisasi Menggunakan Algoritma K-Means," *Jurnal INSAN Journal of Information System Management Innovation*, vol. 4, no. 1, pp. 45–54, 2024, doi: 10.31294/jinsan.v4i1.3763.
- [6] A. Rifqi and B. B. Butar, "IJCIT (Indonesian Journal on Computer and Information Technology) Analisis Kualitas Pisang Berdasarkan Tingkat Kematangan Dengan Algoritma K-Means (Analysis of Banana Quality Based on Ripeness Level Using the K-Means Algorithm)," vol. 7, no. 1, pp. 1–9, 2022.
- [7] M. R. Avivah, "Penerapan Algoritma K-Means untuk Mengelompokan Penjualan Mebel," vol. 1, pp. 61–65, 2020.
- [8] D. Zakiyah, N. Merlina, and N. A. Mayangky, "Penerapan Algoritma K-Means Clustering Untuk Mengetahui Kemampuan Karyawan IT," vol. 2, no. 1, 2022.
- [9] M. Annas and S. N. Wahab, "Data Mining Methods : K-Means Clustering Algorithms," vol. 3, no. 1, 2023.
- [10] F. K. Alvisan, "Clustering Minimarket Untuk Menentukan Jumlah Kebutuhan Pembelian Menggunakan Metode K-Means," *Jurnal NOE*, vol. 4, no. 2, pp. 160–168, 2021, [Online]. Available: <https://ojs.unpkediri.ac.id/index.php/noe>