



# Penerapan Metode K-Means Clustering dalam Pengelompokan Produk Kayu sebagai Sistem Pendukung Keputusan Produksi dan Penjualan (Studi kasus: CV Sutan Piko Mandiri)

Seli, Clara A. Nababan\*, Andri Wijaya

Universitas Katolik Musi Charitas Palembang

**Abstrak:** Pengelolaan persediaan merupakan salah satu aspek krusial dalam industri pengolahan kayu karena berpengaruh langsung terhadap kelancaran produksi, efisiensi biaya, dan kemampuan perusahaan dalam memenuhi permintaan pasar. Ketidaktepatan dalam menentukan jumlah stok sering menimbulkan permasalahan berupa kelebihan atau kekurangan persediaan, yang pada akhirnya dapat berdampak pada kerugian operasional. Oleh karena itu, diperlukan pendekatan berbasis data untuk mendukung pengambilan keputusan yang lebih akurat. Penelitian ini bertujuan untuk menerapkan metode K-Means Clustering dalam mengelompokkan produk kayu pada CV Sutan Piko Mandiri berdasarkan pola penjualan sebagai dasar pendukung keputusan produksi dan pengelolaan persediaan. Data yang digunakan terdiri dari 1.000 transaksi penjualan tahun 2025 dengan atribut harga per unit, volume unit, dan jumlah terjual. Tahapan penelitian meliputi pengumpulan data, preprocessing untuk menangani data hilang dan pemilihan atribut, normalisasi data, serta penentuan jumlah cluster optimal menggunakan elbow method melalui analisis nilai Sum of Squared Error (SSE). Hasil analisis menunjukkan bahwa nilai K optimal adalah 4, sehingga produk kayu dapat dikelompokkan ke dalam empat cluster dengan karakteristik penjualan yang berbeda. Analisis hasil clustering menunjukkan bahwa jenis kayu Mahoni, Kamper, dan Durian termasuk dalam kategori penjualan tinggi sehingga direkomendasikan sebagai prioritas produksi dan perlu dijaga ketersediaan stoknya. Kayu Balam berada pada kategori penjualan menengah, sedangkan kayu Jati tergolong dalam kategori penjualan rendah sehingga memerlukan pengendalian produksi yang lebih ketat untuk menghindari penumpukan stok. Temuan ini membuktikan bahwa penerapan metode K-Means Clustering mampu memberikan wawasan yang bermanfaat bagi perusahaan dalam menyusun strategi produksi dan pengelolaan persediaan secara lebih efektif, efisien, dan berbasis data.

**Kata Kunci:** K-Means Clustering, Sum of Squared Error (SSE), Elbow Method

DOI:

<https://doi.org/10.47134/jtsi.v3i1.5215>

\*Correspondence: Clara A. Nababan

Email: [nababanclayy03@gmail.com](mailto:nababanclayy03@gmail.com)

Received: 22-11-2025

Accepted: 22-12-2025

Published: 22-01-2026



**Copyright:** © 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Inventory management is a critical aspect of the wood processing industry as it directly affects production continuity, cost efficiency, and the company's ability to meet market demand. Inaccurate inventory planning often leads to overstock or stock shortages, which may cause operational inefficiencies and financial losses. Therefore, a data-driven approach is required to support more accurate and objective decision-making. This study aims to apply the K-Means Clustering method to classify wood products at CV Sutan Piko Mandiri based on sales patterns as a decision support tool for production and inventory management. The dataset consists of 1,000 sales transactions recorded in 2025, with attributes including unit price, unit volume, and quantity sold. The research stages include data collection, preprocessing to handle missing values and attribute selection, data normalization, and determination of the optimal number of clusters using the elbow method through Sum of Squared Error (SSE) analysis. The results indicate that the optimal number of clusters is  $K = 4$ , resulting in four distinct sales groups with different characteristics. Cluster analysis reveals that Mahogany, Camphor, and Durian wood types belong to the high-sales category and are recommended as production priorities with maintained stock availability. Balam wood is classified in the medium-sales category, while Teak wood exhibits low sales performance and therefore requires stricter production control to prevent excess inventory. These findings demonstrate that the K-Means Clustering method can effectively support companies in formulating production strategies and managing inventory more efficiently by providing insights based on historical sales data.

**Keywords:** K-Means Clustering, Sum of Squared Error (SSE), Elbow Method

## Pendahuluan

Pengelolaan persediaan merupakan aspek strategis dalam industri manufaktur karena berpengaruh langsung terhadap efisiensi biaya operasional, kelancaran proses produksi, dan kemampuan perusahaan dalam memenuhi permintaan pasar (Gupta & Singh, 2022) (Heizer et al, 2022) (Ramanathan, 2021). Ketidaktepatan dalam menentukan jumlah persediaan dapat menyebabkan kelebihan stok (*overstock*) maupun kekurangan stok (*stockout*) yang berdampak pada meningkatnya biaya penyimpanan serta penurunan tingkat pelayanan pelanggan (Rahman & Putra, 2021) (Wibowo & Setiawan, 2022). Oleh karena itu, perusahaan memerlukan pendekatan pengelolaan persediaan yang lebih sistematis dan berbasis data.

Perkembangan teknologi informasi dan analisis data mendorong perusahaan untuk memanfaatkan data historis sebagai dasar pengambilan keputusan. Sistem Pendukung Keputusan (SPK) yang mengintegrasikan teknik analisis data terbukti mampu meningkatkan efektivitas perencanaan produksi dan pengendalian persediaan (Turban et al, 2021) (Pratama & Suryani, 2022). Pendekatan berbasis data ini menjadi semakin penting seiring dengan meningkatnya kompleksitas rantai pasok dan dinamika permintaan pasar (Mirzaei & Rahmani, 2023).

Data mining merupakan salah satu pendekatan yang banyak digunakan untuk mengekstraksi informasi dan pola penting dari data dalam jumlah besar. Teknik *clustering* memungkinkan pengelompokan data berdasarkan tingkat kemiripan karakteristik sehingga dapat digunakan untuk segmentasi produk dan analisis perilaku penjualan (Han et al, 2022) (Tan et al, 2022). Pendekatan ini terbukti efektif dalam mendukung pengambilan keputusan operasional di bidang inventori dan manajemen rantai pasok (Arora & Dhiman, 2021).

Salah satu algoritma *clustering* yang paling umum digunakan adalah *K-Means Clustering*. Algoritma ini bekerja dengan mengelompokkan data berdasarkan jarak terdekat ke pusat cluster (*centroid*) dan memiliki keunggulan dalam hal kesederhanaan serta efisiensi komputasi (Kumar & Rani, 2023) (Safitri & Nugroho, 2023). Berbagai penelitian menunjukkan bahwa *K-Means* mampu mengelompokkan produk berdasarkan tingkat penjualan sehingga memudahkan identifikasi produk *fast moving*, *medium moving*, dan *slow moving* (Hidayat & Firmansyah, 2023) (Nugraha & Prakoso, 2024).

Penentuan jumlah *cluster* yang optimal menjadi faktor penting dalam penerapan algoritma *K-Means*. Metode *elbow* dengan analisis nilai *Sum of Squared Error* (SSE) sering digunakan untuk menentukan jumlah cluster yang memberikan keseimbangan antara akurasi model dan kompleksitas pengelompokan (Safira et al., 2022; Muasaroh & Fatah, 2024). Pemilihan nilai *K* yang tepat terbukti meningkatkan kualitas hasil *clustering* dan ketepatan rekomendasi pengelolaan persediaan (Novia & Rahayu, 2021) (Winarti et al, 2021).

Di Indonesia, penerapan *K-Means Clustering* dalam analisis penjualan dan persediaan telah banyak dilakukan pada sektor ritel dan manufaktur. Beberapa penelitian menunjukkan bahwa penggunaan perangkat lunak seperti *RapidMiner* mampu mempermudah proses analisis dan meningkatkan akurasi pengelompokan produk (Safira

et al, 2022) (Sari & Kurniawan, 2024). Pendekatan ini terbukti membantu perusahaan dalam menyusun kebijakan persediaan yang lebih efisien dan adaptif terhadap permintaan pasar (Sitorus & Hulu, 2025).

Berdasarkan uraian tersebut, penelitian ini menerapkan metode K-Means Clustering untuk mengelompokkan produk kayu pada CV Sutan Piko Mandiri berdasarkan pola penjualan. Pendekatan ini diharapkan dapat membantu perusahaan dalam menentukan prioritas produksi, mengoptimalkan pengelolaan persediaan, serta meningkatkan efisiensi operasional secara keseluruhan melalui pengambilan keputusan yang berbasis data dan analisis yang terstruktur (Arora & Malik, 2023) (Khedr, 2024).

## Metodologi

Penelitian ini menggunakan pendekatan kuantitatif dengan menerapkan metode *K-Means Clustering* untuk mengelompokkan produk kayu pada CV Sutan Piko Mandiri berdasarkan pola penjualan dan karakteristik produk. Data yang digunakan merupakan data sekunder yang diperoleh melalui dokumentasi internal perusahaan, mencakup data penjualan berbagai produk kayu selama satu tahun dengan total 1000 transaksi pada tahun 2024.

Tahap awal penelitian dimulai dengan pengumpulan data penjualan dan persediaan produk dari catatan operasional perusahaan. Data mentah yang diperoleh kemudian masuk ke tahap preprocessing, yang terdiri dari tiga langkah utama: (1) pengecekan data kosong untuk memastikan kelengkapan dan validitas seluruh atribut; (2) pemilihan atribut (*select attribute*) untuk menentukan variabel yang relevan digunakan pada proses *clustering* dan (3) normalisasi menggunakan metode *min-max* agar setiap variabel memiliki skala yang seragam. Normalisasi dilakukan karena algoritma *K-Means* sensitif terhadap perbedaan satuan seperti harga, volume unit, serta jumlah terjual.

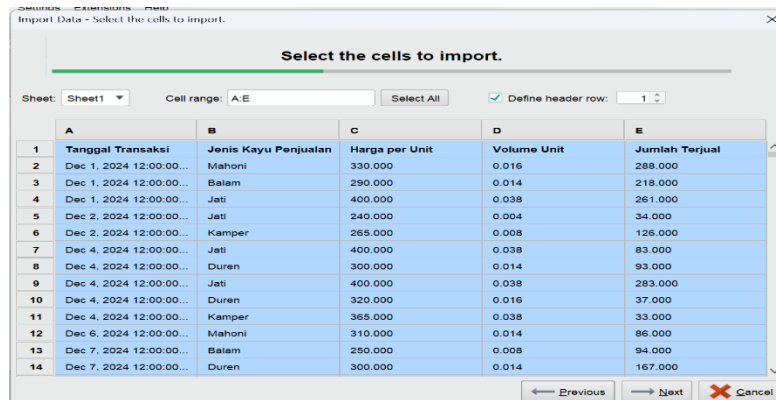
Setelah data siap diolah, langkah berikutnya adalah menentukan rentang nilai K yang akan diuji, yaitu dari  $K = 2$  hingga  $K = 10$ . Untuk setiap nilai K, dilakukan perhitungan *Sum of Squared Error* (SSE) guna mengukur kualitas pengelompokan. Seluruh nilai SSE tersebut kemudian dianalisis melalui tabel atau grafik untuk mencari titik siku (*elbow*), yaitu titik di mana penurunan SSE mulai melambat. Titik ini digunakan sebagai dasar untuk menentukan jumlah cluster (K) yang optimal.

Nilai K yang terpilih selanjutnya digunakan dalam proses *K-Means Clustering*. Pada tahap ini, data dikelompokkan berdasarkan atribut seperti harga produk, volume unit, dan jumlah terjual. Proses algoritma akan menghasilkan beberapa *cluster* dengan karakteristik yang berbeda, misalnya kelompok produk tingkat penjualan tinggi, sedang, maupun rendah.

Tahap lanjutan adalah evaluasi dan interpretasi *cluster*, di mana setiap kelompok dianalisis untuk melihat pola penjualan dan karakteristik produk yang terbentuk. Hasil interpretasi tersebut menjadi dasar dalam penyusunan rekomendasi keputusan, khususnya terkait strategi produksi, perencanaan permintaan, serta pengelolaan persediaan. Melalui pendekatan ini, perusahaan diharapkan dapat mengoptimalkan kapasitas produksi, meningkatkan efisiensi pengendalian persediaan, dan meminimalkan risiko kekurangan stok maupun kelebihan stok.

## Hasil dan Pembahasan

### 1. Import Data



Gambar 1. Import Data

Proses *import data* merupakan tahap awal dalam analisis data menggunakan *RapidMiner*. Pada langkah ini, peneliti memasukkan dataset penjualan dan file tersebut diimpor menggunakan format *xlsx (default modern)*. Tahapan import data penting dalam proses data mining karena memastikan bahwa dataset masuk dengan benar, dapat dibaca oleh sistem. Dengan memasukkan data melalui *fitur import*, *RapidMiner* kemudian mengenali struktur kolom seperti tanggal transaksi, jenis kayu penjualan, harga per unit, volume unit dan jumlah terjual, sehingga pengguna dapat menyesuaikan tipe data pada tahap selanjutnya.

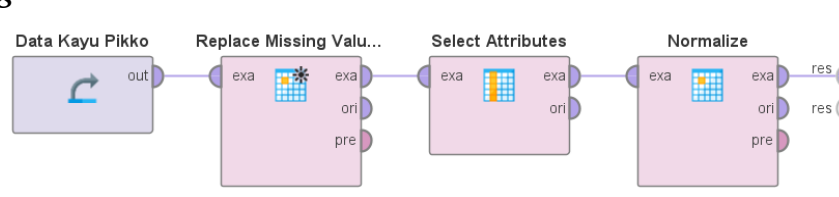
### 2. Mining Data

	Tanggal Transaksi	Jenis Kayu Penjualan	Harga per Unit	Volume Unit	Jumlah Terjual
	date	polynomial	integer	real	integer
1	Dec 1, 2024	Mahoni	330000	0.016	288
2	Dec 1, 2024	Balam	290000	0.014	218
3	Dec 1, 2024	Jati	400000	0.038	261
4	Dec 2, 2024	Jati	240000	0.004	34
5	Dec 2, 2024	Kamper	265000	0.008	126
6	Dec 4, 2024	Jati	400000	0.038	83
7	Dec 4, 2024	Duren	300000	0.014	93
8	Dec 4, 2024	Jati	400000	0.038	283
9	Dec 4, 2024	Duren	320000	0.016	37
10	Dec 4, 2024	Kamper	365000	0.038	33
11	Dec 6, 2024	Mahoni	310000	0.014	86
12	Dec 7, 2024	Balam	250000	0.008	94

Gambar 2. Proses Mining Data

Tahap *Mining Data* merupakan proses inti dalam alur *data mining*, di mana data yang telah berhasil di-import selanjutnya diolah dan dianalisis menggunakan algoritma tertentu. Peneliti menyesuaikan tipe data seperti: Tanggal Transaksi = date, Jenis Kayu Penjualan = polynomial, Harga Per Unit = integer, Volume Unit = real dan Jumlah Terjual = integer. Tahap ini memastikan bahwa data siap dianalisis dengan akurat, sehingga hasil clustering dapat merepresentasikan pola penjualan yang sebenarnya.

### 3. Preprocessing Data

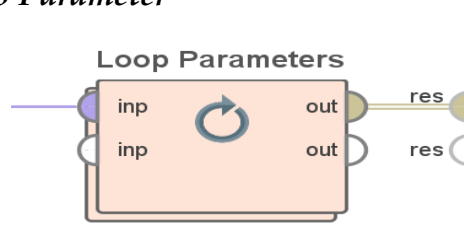


Gambar 3. Preprocessing Data

Tahap *preprocessing data* merupakan bagian penting dalam proses data mining untuk memastikan kualitas data sebelum dianalisis menggunakan algoritma K-Means. Dalam dataset penjualan yang digunakan, terdapat 6 *data missing* yang harus ditangani terlebih dahulu. Oleh karena itu dilakukan proses *Replace Missing Value* untuk memastikan tidak ada nilai kosong yang dapat mengganggu proses perhitungan jarak pada algoritma *K-Means*. Penanganan data hilang ini dilakukan dengan mengganti nilai kosong menggunakan metode yang sesuai (misalnya nilai rata-rata atau modus), sehingga struktur data menjadi konsisten.

*Select Attribute* yaitu memilih atribut yang relevan untuk analisis, sehingga mengurangi kompleksitas dan menghindari penggunaan variabel yang tidak diperlukan. Atribut yang dipilih dalam penelitian ini adalah Harga Per Unit, volume unit dan jumlah terjual. Pemilihan atribut ini dilakukan karena ketiganya berhubungan langsung dengan analisis pola penjualan. Tahap berikutnya adalah *Normalize Data* untuk menyesuaikan skala data numerik agar berada dalam rentang tertentu sehingga mempermudah proses analisis dan meningkatkan performa algoritma *machine learning*. Tipe kolom yang dilakukan normalisasi adalah Jumlah Terjual, sehingga perbedaan skala data tidak mengganggu proses penghitungan jarak *Euclidean* pada algoritma *clustering*.

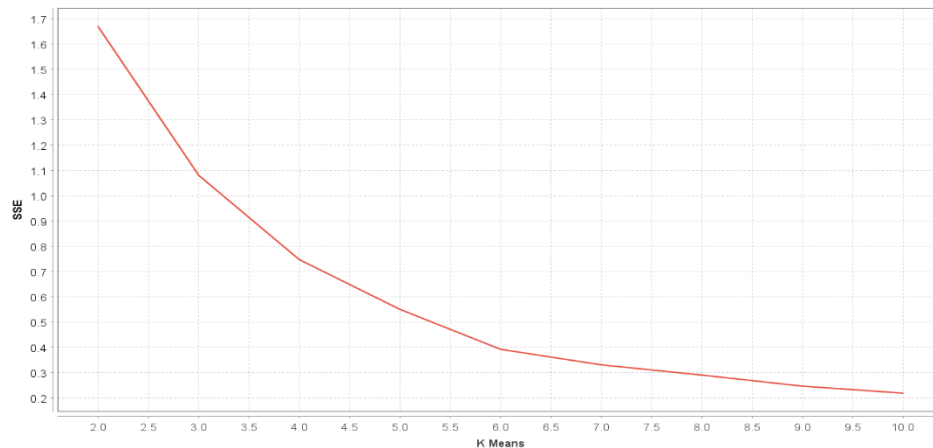
### 4. Tentukan Nilai K di Loop Parameter



Gambar 4. Atribut Loop Parameter

*Loop Parameters* digunakan pada tahap penentuan jumlah cluster (nilai K) untuk menguji berbagai kemungkinan nilai K dan menghitung nilai SSE (*Sum of Squared Error*) pada setiap percobaan. Melalui fitur ini, proses pemilihan K menjadi lebih efisien karena sistem dapat menjalankan pengujian secara otomatis dari K = 2 hingga K = 10, kemudian membandingkan hasil SSE yang dihasilkan. Dengan demikian, *Loop Parameters* membantu mengidentifikasi nilai K yang paling optimal berdasarkan nilai SSE terendah tanpa harus melakukan pengujian secara manual satu per satu.

## 5. Grafik SSE untuk Menentukan Titik Siku (Elbow)



**Gambar 5.** Grafik SSE untuk Menentukan Titik Siku (*Elbow*)

Nilai K optimal ditetapkan pada  $K = 4$  karena pada titik tersebut grafik mulai menunjukkan perubahan yang lebih landai dibandingkan nilai K sebelumnya. Penurunan SSE dari  $K=2$  ke  $K=3$  dan dari  $K=3$  ke  $K=4$  masih tergolong signifikan, namun mulai dari  $K=4$  ke  $K=5$  penurunan SSE menjadi jauh lebih kecil. Dengan demikian,  $K=4$  dipandang sebagai titik siku (*elbow*) yang paling tepat, karena mampu memberikan keseimbangan antara penurunan kesalahan dan efisiensi jumlah kluster. Selain itu, pemilihan  $K=4$  menghasilkan pembentukan kluster produk yang lebih seimbang, sehingga lebih representatif dalam menggambarkan variasi karakteristik produk pada dataset.

## 6. Proses *K-Means Clustering* Menggunakan K Optimal

Parameter	Value
Clustering (k-Means)	
add cluster attribute	<input checked="" type="checkbox"/>
add as label	<input type="checkbox"/>
remove unlabeled	<input type="checkbox"/>
k	4
max runs	10
determine good start values	<input checked="" type="checkbox"/>
measure types	Numerical Measures
numerical measure	Euclidean Distance
max optimization steps	100

**Gambar 6.** *K-Means Clustering* Menggunakan K Optimal

Pengaturan parameter yang digunakan dalam proses *K-Means Clustering*, di mana nilai K sebesar 4 (*metode elbow*). Opsi *add cluster attribute* diaktifkan untuk menambahkan atribut kluster pada setiap data hasil pengelompokan. Proses *K-Means* dijalankan sebanyak 10 kali (*max runs* = 10) untuk memastikan model mendapatkan hasil terbaik dengan variasi inisialisasi *centroid*, sementara fitur *determine good start values* digunakan agar algoritma

memilih titik awal *centroid* yang lebih stabil. Pengukuran jarak antar data menggunakan *Euclidean Distance*, yaitu ukuran jarak paling umum dalam analisis *clustering*, dan cocok untuk data numerik seperti yang dipilih pada bagian *measure types*. Sementara itu, jumlah maksimal iterasi optimasi adalah 100 langkah, memungkinkan algoritma untuk mencapai konvergensi dalam memperbarui *centroid* dan membentuk kluster yang paling representatif. Pengaturan ini memastikan bahwa proses *K-Means* berjalan optimal dan menghasilkan pembagian kluster yang akurat sesuai karakteristik data.

## 7. Hasil Cluster

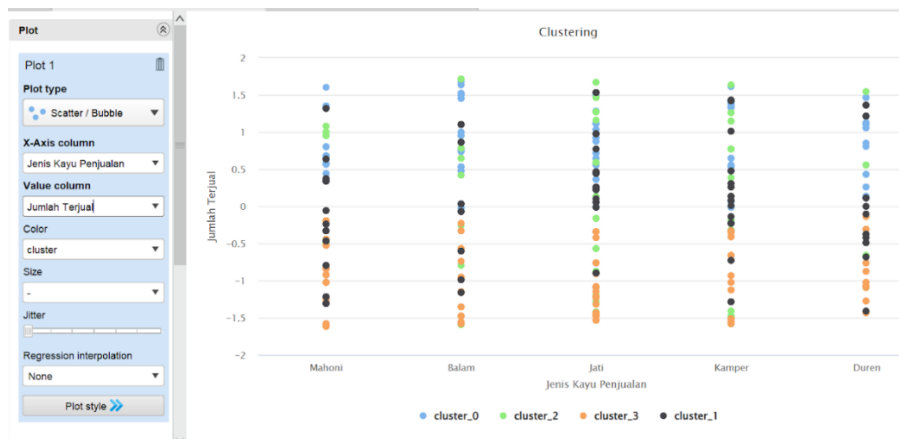
### Cluster Model

```
Cluster 0: 276 items
Cluster 1: 191 items
Cluster 2: 256 items
Cluster 3: 277 items
Total number of items: 1000
```

Gambar 7. Cluster Model

Hasil *Cluster Model* menunjukkan bahwa data berhasil terbagi ke dalam empat kluster dengan distribusi yang relatif seimbang, yaitu *Cluster 0* berisi 276 item, *Cluster 1* berisi 191 item, *Cluster 2* berisi 256 item, dan *Cluster 3* berisi 277 item, dengan total keseluruhan 1000 item. Penyebaran jumlah item yang tidak terpaut jauh antara kluster satu dengan yang lain mengindikasikan bahwa pemilihan  $K = 4$  sudah tepat, karena mampu menghasilkan kelompok data yang proporsional dan tidak terjadi dominasi kluster tertentu. Komposisi ini juga menunjukkan bahwa karakteristik data dalam dataset memiliki variasi yang cukup merata, sehingga setiap kluster dapat merepresentasikan kelompok dengan ciri khas masing-masing tanpa ketimpangan jumlah anggota. Hasil ini memperkuat bahwa proses *K-Means* yang diterapkan telah berjalan optimal dalam mengelompokkan data secara efektif.

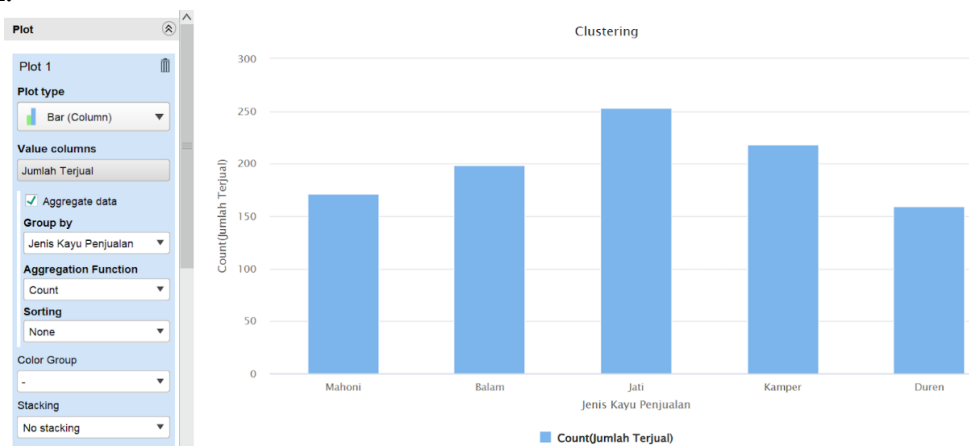
## 8. Visualisasi Hasil Cluster



Gambar 8. Visualisasi Hasil Cluster

Berikutnya juga disajikan visualisasi hasil *cluster* yang mana gambar 10. Menunjukkan hasil visualisasi klasterisasi pada data penjualan berdasarkan produk dan jumlah terjual yang telah didominasi. Pada sumbu horizontal (X-axis) terdapat berbagai jenis produk yang dijual (Mahoni, Balam, Jati, Kamper, Durian), sementara pada sumbu vertikal (Y-axis) ditampilkan nilai jumlah terjual setelah melalui proses normalisasi sehingga skala nilainya berbeda. Titik berwarna biru (*cluster 0*) umumnya menempati posisi atas pada sumbu Y, mengindikasikan pola penjualan tinggi atau performa penjualan yang lebih baik, sedangkan titik berwarna hijau (*cluster 2*) cenderung berada pada posisi tinggi-sedang, menunjukkan pola penjualan tinggi sampai menengah. Titik berwarna hitam (*cluster 1*) menempati area tengah sehingga merepresentasikan pola penjualan sedang, dan titik oranye (*cluster 3*) dominan pada area bawah yang mengindikasikan pola penjualan rendah.

Visualisasi juga memperlihatkan bahwa sebaran klaster relatif konsisten di seluruh jenis produk serta adanya tumpang tindih antar klaster memiliki karakteristik penjualan serupa di batas klaster. Hasil ini membantu mengidentifikasi segmen produk dengan performa tinggi untuk strategi pemasaran dan stok yang lebih agresif (*cluster* biru/hijau), sementara cluster orange menandai produk yang memerlukan intervensi, promosi, atau evaluasi ulang strategi penjualan; perlu dibuat untuk observasi *outlier* yang jauh dari konsentrasi titik lain karena dapat menunjukkan kasus penjualan tak biasa atau kesalahan pencatatan.



**Gambar 9.** Count Produk

Grafik diatas merupakan hasil visualisasi dari analisis *clustering* data penjualan yang dilakukan dengan aplikasi *RapidMiner*. Sumbu Horizontal (X) menunjukkan kategori produk mahoni, balam, jati, kamper dan durian. Sumbu Vertikal (Y) menunjukkan frekuensi atau jumlah penjualan dari produk yang ditunjukkan dalam bentuk jumlah unit penjualan.

Hasil analisis dalam grafik menunjukkan bahwa Jenis kayu Jati memiliki jumlah data terbanyak menjadikannya produk paling populer berdasarkan data yang dianalisis. Produk lain seperti Kamper dan Balam yang juga memiliki frekuensi tinggi, sementara itu, Mahoni dan Durian memiliki jumlah data yang lebih rendah, sehingga frekuensi transaksinya lebih sedikit dibandingkan produk lainnya.

## Simpulan

Berdasarkan hasil penerapan metode *K-Means Clustering* dengan jumlah *cluster* optimal  $K = 4$ , penelitian ini berhasil mengelompokkan 1.000 data transaksi penjualan produk kayu pada CV Sutan Piko Mandiri ke dalam empat kelompok dengan karakteristik penjualan yang berbeda. Hasil pengelompokan menunjukkan bahwa jenis kayu Mahoni, Kamper, dan Durian termasuk dalam kategori penjualan tinggi dengan tingkat perputaran stok yang cepat, sehingga direkomendasikan sebagai prioritas utama dalam perencanaan produksi dan pengelolaan persediaan agar ketersediaan stok tetap terjaga. Kayu Balam berada pada kategori penjualan menengah dan memerlukan pengelolaan produksi yang disesuaikan dengan pola permintaan, sedangkan kayu Jati tergolong dalam kategori penjualan rendah sehingga membutuhkan pengendalian produksi yang lebih ketat untuk menghindari penumpukan persediaan dan peningkatan biaya penyimpanan.

Penerapan metode *K-Means Clustering* terbukti mampu memberikan segmentasi produk yang lebih objektif dan sistematis dibandingkan pendekatan konvensional. Informasi hasil clustering dapat dimanfaatkan sebagai sistem pendukung keputusan bagi manajemen dalam mengevaluasi kinerja penjualan setiap jenis produk secara berkala. Dengan adanya pengelompokan ini, perusahaan dapat menyusun strategi produksi, pengendalian stok, dan distribusi secara lebih tepat sasaran serta responsif terhadap perubahan permintaan pasar. Jadi, penelitian ini membuktikan bahwa pendekatan berbasis data melalui *K-Means Clustering* dapat meningkatkan efisiensi operasional dan mendukung pengambilan keputusan yang lebih efektif dan berkelanjutan.

Untuk pengembangan penelitian selanjutnya, disarankan agar perusahaan mengintegrasikan hasil *clustering* ke dalam sistem operasional secara berkelanjutan, menambahkan variabel pendukung seperti biaya penyimpanan, margin keuntungan, dan waktu pemesanan ulang, serta melakukan perbandingan dengan *algoritma clustering* lain guna memperoleh hasil analisis yang lebih optimal dan komprehensif.

## Daftar Pustaka

- Arora, P., & Dhiman, G. (2021). Clustering techniques in data mining: A survey. *Materials Today: Proceedings*, 47, 4371–4376.
- Arora, P., & Malik, S. (2023). Machine learning applications in inventory optimization. *International Journal of Production Research*, 61(4), 1205–1218.
- Gupta, S., & Singh, R. (2022). Inventory management challenges in manufacturing industries. *Journal of Supply Chain Systems*, 11(2), 45–55.
- Han, J., Pei, J., & Kamber, M. (2022). *Data mining: Concepts and techniques* (4th ed.). Morgan Kaufmann.
- Heizer, J., Render, B., & Munson, C. (2022). *Operations management* (14th ed.). Pearson Education.
- Hidayat, T., & Firmansyah, R. (2023). Analisis penjualan berbasis data mining untuk pengendalian persediaan. *Jurnal Sistem Informasi Bisnis*, 13(2), 87–96.
- Hossain, M., & Rahman, S. (2022). Inventory classification using data mining techniques. *Expert Systems with Applications*, 190, 116178.

- Khedr, A. M. (2024). Machine learning applications for inventory and demand forecasting: A review. *International Journal of Production Economics*, 259, 108845.
- Kumar, A., & Rani, S. (2023). K-Means clustering for product segmentation in retail sectors. *International Journal of Data Analytics*, 9(1), 55–66.
- Lee, J., & Kim, H. (2023). Sales data clustering for inventory planning. *Applied Sciences*, 13(4), 2105.
- Mirzaei, M., & Rahmani, K. (2023). Applications of clustering in supply chain analytics. *Journal of Industrial Engineering and Management*, 16(3), 215–229.
- Muasaroh, Y. I., & Fatah, Z. (2024). Implementasi RapidMiner dalam optimasi pembentukan kelas unggulan menggunakan K-Means clustering. *Jurnal Advance Research Informatika*, 3(1), 66–72.
- Novia, C. P. E. A., & Rahayu, W. I. (2021). Perbandingan algoritma K-Means dan Naïve Bayes untuk prediksi prioritas pembayaran. *Kreatif Industri Nusantara*, 2(1), 12–20.
- Pratama, A. R., & Suryani, D. (2022). Data-driven inventory decision support system using clustering techniques. *Journal of Decision Support Systems*, 5(1), 21–30.
- Rahman, A., & Putra, D. (2021). Inventory segmentation using clustering approach. *Journal of Industrial Engineering Research*, 10(3), 145–154.
- Ramanathan, R. (2021). Inventory management and demand forecasting. *Operations and Logistics Review*, 5(2), 88–102.
- Safira, I., Salkiawati, R., & Priatna, W. (2022). Penerapan algoritma K-Means untuk mengetahui pola persediaan barang. *Jurnal Informatika dan Keamanan Informasi*, 3(1), 99–110.
- Safitri, R., & Nugroho, Y. (2023). Product clustering for inventory control using K-Means algorithm. *Journal of Information Systems Engineering*, 8(2), 134–142.
- Sari, M. P., & Kurniawan, A. (2024). Implementation of K-Means clustering for sales pattern analysis. *International Journal of Advanced Computer Science and Applications*, 15(1), 402–410.
- Sitorus, B. R., & Hulu, Y. (2025). Optimization of raw material inventory using clustering methods. *International Journal of Supply Chain Management*, 14(1), 33–41.
- Tan, P.-N., Steinbach, M., & Kumar, V. (2022). *Introduction to data mining* (2nd ed.). Pearson.
- Turban, E., Sharda, R., Delen, D., & King, D. (2021). *Decision support and business intelligence systems* (12th ed.). Pearson.
- Wibowo, A., & Setiawan, H. (2022). Penerapan data mining dalam sistem pendukung keputusan persediaan. *Jurnal RESTI*, 6(4), 612–620.
- Winarti, D., Revita, E., & Yandani, E. S. (2021). Penerapan data mining untuk analisis pola menggunakan algoritma K-Means. *Jurnal SIMTIKA*, 4(3), 8–22.