



# Implementasi Algoritma K-Means Pada Sistem Persediaan Barang

Achsyarul Khaliq, Ismasari Nawangsih, Annisa Maulana Majid\*

Fakultas Teknik, Program Studi Teknik Informatika, Universitas Pelita Bangsa, Bekasi, Indonesia

Email: <sup>1</sup>khaliqachsyarul@gmail.com, <sup>2</sup>ismasari.n@pelitabangsa.ac.id, <sup>3,\*</sup>annisa.maulanamajid@pelitabangsa.ac.id

Email Penulis Korespondensi: annisa.maulanamajid@pelitabangsa.ac.id

**Abstrak**—Persediaan barang merupakan salah satu komponen penting dalam kegiatan operasional perusahaan, khususnya pada sektor perdagangan, karena berpengaruh langsung terhadap kelancaran penjualan dan tingkat kepuasan pelanggan. Pengelolaan persediaan yang tidak terstruktur dapat menyebabkan terjadinya penumpukan barang, kekurangan stok, serta pengambilan keputusan yang kurang tepat. Toko Bangunan Maisa yang berlokasi di Semerap, Kabupaten Kerinci, Jambi, hingga saat ini masih melakukan pencatatan persediaan secara manual, sehingga pemilik toko mengalami kesulitan dalam mengidentifikasi barang dengan tingkat penjualan tinggi maupun rendah. Penelitian ini bertujuan untuk mengimplementasikan algoritma K-Means Clustering dalam mengelompokkan persediaan barang berdasarkan tingkat penjualan guna mendukung pengelolaan stok yang lebih efektif dan efisien. Metode penelitian yang digunakan adalah data mining dengan tahapan pengumpulan data, preprocessing, perhitungan manual algoritma K-Means, serta implementasi menggunakan perangkat lunak RapidMiner. Data yang dianalisis berjumlah 507 item persediaan yang telah melalui proses pembersihan data sehingga layak digunakan dalam pengelompokan. Pengelompokan dilakukan dengan dua cluster, yaitu cluster barang dengan tingkat penjualan rendah dan cluster barang dengan tingkat penjualan tinggi. Hasil penelitian menunjukkan bahwa sebanyak 494 item barang atau sebesar 97,44 persen tergolong ke dalam cluster penjualan rendah, sedangkan 13 item barang atau sebesar 2,56 persen tergolong ke dalam cluster penjualan tinggi. Hasil ini menunjukkan bahwa sebagian besar produk memiliki perputaran penjualan yang relatif rendah, sementara hanya sebagian kecil produk yang menjadi penyumbang utama terhadap total penjualan toko. Informasi yang dihasilkan dari proses pengelompokan ini dapat digunakan sebagai dasar pengambilan keputusan dalam pengelolaan persediaan barang, khususnya dalam menentukan prioritas penyetokan, pengendalian stok, serta penyusunan strategi pemasaran yang tepat dan berbasis data.

**Kata Kunci:** Data Mining; Persediaan Barang; K-Means Clustering; Pengelompokan

**Abstract**—Inventory is an important component in a company's operational activities, especially in the trade sector, because it directly affects the smoothness of sales and the level of customer satisfaction. Unstructured inventory management can lead to stockpiling of goods, stock shortages, and inappropriate decision making. Maisa Building Materials Store located in Semerap, Kerinci Regency, Jambi, currently still records inventory manually, so the shop owner has difficulty in identifying items with high and low sales levels. This study aims to implement the K-Means Clustering algorithm in grouping inventory based on sales levels to support more effective and efficient stock management. The research method used is data mining with the stages of data collection, preprocessing, manual calculation of the K-Means algorithm, and implementation using RapidMiner software. The analyzed data amounted to 507 inventory items that have gone through a data cleaning process so that they are suitable for use in grouping. Grouping is done with two clusters, namely a cluster of goods with a low sales level and a cluster of goods with a high sales level. The results of the study indicate that 494 items, or 97.44 percent, fall into the low-sales cluster, while 13 items, or 2.56 percent, fall into the high-sales cluster. These results indicate that most products have relatively low sales turnover, while only a small proportion contribute significantly to total store sales. The information generated from this clustering process can be used as a basis for decision-making in inventory management, particularly in determining stocking priorities, stock control, and developing appropriate, data-driven marketing strategies.

**Keywords:** Data Mining; Inventory; K-Means Clustering; Clustering

## 1. PENDAHULUAN

Teknologi dan informasi terus mengalami kemajuan yang cukup signifikan dari waktu ke waktu. Sehingga pada dunia bisnis sekali pun membutuhkan informasi yang sangat jelas dan akurat untuk digunakan oleh pelaku bisnis besar, menengah maupun kecil sekalipun. Salah satu contoh informasi yang dibutuhkan pada dunia bisnis adalah data penjualan, yang mencakupi data transaksi, stok barang dan data lainnya. Data-data ini akan berguna bagi para pelaku untuk mengantisipasi hal buruk yang bisa terjadi kedepannya. Seperti dalam persaingan bisnis dimana para pelaku bisnis dituntut untuk selalu berpikir kritis dengan informasi yang mereka dapatkan dari data-data penjualan. Adapun, para pelaku bisnis mendapatkan informasi yang terjadi dengan menggunakan metode maupun suatu teknik dalam pengambilan atau pengolahan data [1].

Persediaan merupakan salah satu komponen penting dalam kegiatan operasional perusahaan, terutama yang bergerak di bidang perdagangan barang. Persediaan memiliki peran ganda; di satu sisi dapat dianggap sebagai beban jika berlebihan dan tidak terkelola dengan baik, namun di sisi lain sangat penting untuk menjamin kelancaran pemenuhan permintaan konsumen. Tanpa persediaan yang memadai, proses pelayanan kepada pelanggan dapat terhambat dan berisiko menurunkan tingkat kepuasan pelanggan [2].

Toko Bangunan Maisa adalah sebuah usaha yang bergerak dalam penjualan alat-alat bangunan, berlokasi di Semerap, Kabupaten Kerinci, Provinsi Jambi. Dalam aktivitas operasionalnya, toko ini belum menggunakan sistem pencatatan dan pengelolaan persediaan yang terkomputerisasi. Pemilik toko hanya melakukan pengecekan secara manual terhadap jumlah barang yang dikirim oleh distributor dan mencocokkannya dengan nota pembelian. Jika terdapat ketidaksesuaian, maka akan dilakukan komplain kepada pihak distributor. Prosedur ini belum mampu memberikan gambaran yang jelas mengenai stok barang yang tersedia, barang yang sering terjual, atau barang yang jarang diminati. Untuk meningkatkan efisiensi dalam manajemen persediaan, diperlukan penerapan teknologi berbasis data yang mampu mengelompokkan barang berdasarkan karakteristik tertentu, seperti frekuensi penjualan. Salah satu metode yang dapat



diterapkan adalah algoritma *K-Means Clustering*, yang merupakan bagian dari *unsupervised learning*. Algoritma ini mengelompokkan data ke dalam sejumlah kluster berdasarkan jarak terdekat ke titik pusat (*centroid*) [1]. Dengan pengelompokan ini, pemilik toko dapat mengidentifikasi barang mana yang perlu mendapatkan prioritas dalam penyetokan dan mana yang tidak. Untuk itu penulis menggunakan algoritma *K-Means data mining* untuk menghasilkan informasi yang baik dan dapat menyediakan informasi yang sangat strategis.

Penelitian ini bertujuan untuk mengembangkan metode persediaan barang menggunakan algoritma *K-Means Clustering*. Algoritma *K-Means* merupakan algoritma non hirarki yang berasal dari metode *Clustering* yang paling sederhana dan mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang cepat dan efisien metode *K-Means* ini mempartisi data dalam kelompok sehingga data berkarakteristik sama dimasukkan kedalam set kelompok yang sama. *K-Means* merupakan suatu metode pengujian komponen populasi data dan mengelompokkan data tersebut kedalam suatu *cluster* yang telah didefinisikan tergantung dari jarak minimum antar komponen populasi dengan masing-masing pusat *cluster* [2]. Secara keseluruhan penelitian ini memberikan kontribusi penting dalam penginputan data persediaan barang yang akurat dan terperinci.

Berdasarkan penelitian oleh Alfian Adiyanto, Y Arie Wijaya pada tahun 2023 yg membahas tentang Penerapan Algoritma *K-Means* Pada Pengelompokan Dataset Bahan Pangan Indonesia Tahun 2022-2023. Tujuan Penelitian ini adalah melakukan pengelompok/ kluster yang optimal menggunakan algoritma *k-means* pada dataset bahan pangan. Oleh karena itu, penelitian ini menggunakan metode analisa data *Knowledge Discovery in Database (KDD)* Sumber data penelitian ini diperoleh dari *website* Pusat Infomasi Harga Pangan Strategis Nasional. Data ini kumpulan data dari bulan Desember 2022-Januari 2023. Penelitian ini menggunakan *clustering* metode *K-Means* untuk mengelompokkan data tersebut dengan metode *K-Means* yang sederhana [3] .

Selanjutnya berdasarkan penelitian Febriansyah, S Muntari pada tahun 2023 yg membahas tentang Penerapan Algoritma *K-Means* Untuk Klasterisasi Penduduk Miskin Pada Kota Pagar Alam. Tujuan dari penelitian ini adalah untuk mendapatkan kluster data kemiskinan di Kota Pagar Alam. Pendataan penerima bantuan Program Keluarga Harapan (PKH) belum tepat, pemberian bantuan hanya memperhatikan kriteria kemiskinan secara umum, sehingga masih banyak masyarakat miskin yang merasa lebih pantas dapat bantuan PKH [4].

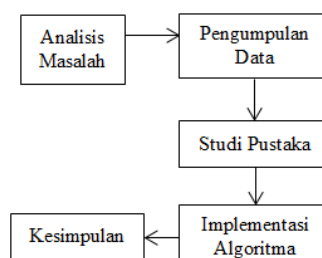
Selanjutnya berdasarkan penelitian yg diteliti oleh Andi Patappari, Nurul Muhlisa pada tahun 2023 yang membahas tentang Sistem Informasi Persediaan Barang Berbasis *Web* Pada Toko Throve Store Soppeng. Penelitian ini membahas tentang Rancang Bangun Sistem Informasi Persediaan Barang Berbasis *Web* pada Toko Throve Store Soppeng. Penelitian ini bertujuan meng-implementasikan Sistem Informasi Persediaan Barang Berbasis *Web* pada Toko Throve Store Soppeng [5].

## 2. METODE PENELITIAN

### 2.1 Kerangka Kerja Penelitian

Dalam penyusunan Penelitian ini, ada beberapa tahapan yang perlu dilakukan, yaitu:

- Analisis Permasalahan**  
Tahap analisis permasalahan bertujuan untuk mengkaji masalah secara menyeluruh, menemukan faktor penyebab utama, serta menetapkan rumusan masalah atau tujuan penelitian secara jelas.
- Pengumpulan Data**  
Pada tahap ini dilakukan proses pengumpulan data dan informasi yang relevan sesuai dengan permasalahan dan tujuan penelitian yang telah ditentukan.
- Kajian Pustaka**  
Kajian pustaka dilakukan dengan menelaah berbagai literatur, jurnal, dan referensi terdahulu yang berkaitan dengan topik dan permasalahan penelitian.
- Penerapan Algoritma**  
Tahap ini mencakup proses perancangan, pengembangan, dan penerapan algoritma atau metode yang dipilih untuk menyelesaikan permasalahan penelitian.
- Penarikan Kesimpulan**  
Tahap akhir ini berfokus pada penyusunan kesimpulan yang diperoleh berdasarkan hasil pengolahan data, kajian pustaka, serta penerapan algoritma yang telah dilakukan.



Gambar 1. Kerangka Kerja Penelitian



## 2.2 Data Mining

*Data mining* adalah suatu proses penambangan informasi penting dari suatu data. Informasi penting ini didapat dari suatu proses yang amat rumit seperti menggunakan *artificial intelligence*, teknik statistik, ilmu matematika, *machine learning*, dan lain sebagainya [6],[7],[8]. Teknik-teknik rumit tersebut nantinya akan mengidentifikasi dan mengekstraksi informasi yang bermanfaat dari suatu *database* besar. *Data mining* telah menjadi disiplin ilmu yang dibangun dalam domain kecerdasan buatan (AI), dan rekayasa pengetahuan (KE). *Data mining* berakar pada *machine learning* dan statistika, tetapi merambah bidang lain dalam ilmu komputer dan ilmu lainnya seperti biologi, lingkungan, finansial, jaringan, dan sebagainya [9],[10],[11]. *Data Mining* diartikan sebagai menambang data atau upaya untuk menggali informasi yang berharga dan berguna pada *database* yang sangat besar. Hal terpenting dalam teknik data mining adalah aturan untuk menemukan pola frekuensi tinggi antar himpunan *itemset* yang disebut fungsi *Association Rules* (Aturan Asosiasi). Pemanfaatan dari *data mining* sendiri bisa dilihat dari dua sudut pandang, baik sudut pandang komersial dan sudut pandang keilmuan. Dari sudut pandang komersial, *data mining* bisa digunakan untuk menangani adanya peledakan dari volume data. Dengan melihat bagaimana menyimpannya, mengekstraknya dan memanfaatkannya. Tentunya berbagai ilmu komputasi dapat untuk menghasilkan informasi yang dibutuhkan [12],[13]. Selain itu *data mining* juga bisa dimanfaatkan untuk menyelesaikan masalah dalam kebutuhan dibidang bisnis, misalnya:

- Mengetahui hilangnya pelanggan dikarenakan adanya pesaing.
- Mengetahui item suatu produk yang memiliki kesamaan karakteristik.
- Mengidentifikasi produk-produk yang sudah terjual dengan produk lainnya.
- Untuk memprediksi dari tingkat penjualan.
- Menilai tingkat resiko dalam menentukan jumlah produksi pada suatu item.
- Memprediksi perilaku bisnis dimasa depan.

## 2.3 Machine Learning

*Machine Learning* atau Pembelajaran Mesin merupakan teknik pendekatan dari *Artificial Intelligent* ( AI ) yang digunakan untuk menirukan hingga menggantikan peran manusia dalam melakukan aktivitas hingga memecahkan masalah [14]. Secara singkat *Machine Learning* adalah mesin yang dibuat supaya dapat belajar dan melakukan pekerjaan tanpa arahan dari penggunanya. Menurut Arthur Samuel yakni seorang pelopor Amerika di bidang permainan komputer dan kecerdasan buatan menyatakan bahwa *Machine Learning* merupakan cabang ilmu yang mempelajari tentang bagaimana cara untuk memerikan kemampuan bagi komputer untuk belajar tanpa secara eksplisit diprogram. Sedangkan Tom M. Mitchell membuat definisi bahwa "Sebuah program komputer yang belajar dari pengalaman (*experience*) E dari tugas yang dibebankan (*Task*) T dengan kinerjanya (*performance*) P yang terukur [15],[16],[17]. Metode dalam *Machine Learning* terdiri dari dua pendekatan yaitu *Supervised Learning* dan *Unsupervised Learning*. *Supervised Learning* mengadopsi konsep pendekatan fungsi, dimana pada dasarnya algoritma dilatih agar dapat memilih fungsi-fungsi yang paling menggambarkan input dimana X tertentu membuat estimasi terbaik dari Y. Algoritma *Supervised Learning* adalah jenis *Machine Learning* yang paling umum. Dalam bahasa indonesia, *Supervised Learning* diartikan sebagai pembelajaran diawasi. Istilah "diawasi" ini muncul karena algoritma ini dirancang untuk belajar melalui contoh. *Unsupervised Learning* adalah algoritma yang tidak membutuhkan data berlabel. Pada *Unsupervised Learning*, algorithm tidak membutuhkan data training. Algoritma ini digunakan dalam mendeteksi pola dan pemodelan deskriptif yang tidak membutuhkan kategori atau output berlabel yang menjadi dasar algoritma untuk mencari model yang tepat. Algoritma ini digunakan untuk *clustering* dan *association rule*. Berikut merupakan beberapa algoritma yang dapat digunakan dalam pengimplementasian metode *Supervised Learning* dan *Unsupervised Learning* [18].

## 2.4 Algoritma K-Means Clustering

Algoritma *K-Means* merupakan algoritma non hirarki yang berasal dari metode clustering yang paling sederhana dan mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang cepat dan efisien metode *K-Means* ini mempartisi data dalam kelompok sehingga data berkarakteristik sama dimasukkan kedalam set kelompok yang sama *K-Means* merupakan suatu metode pengujian komponen populasi data dan mengelompokkan data tersebut kedalam suatu *cluster* yang telah didefinisikan tergantung dari jarak minimum antar komponen populasi dengan masing-masing pusat *cluster*. Metode *K-Means Clustering* mengelompokkan datan memetakan suatu populasi data ke dalam beberapa kelompok *cluster* dimana tiap-tiap data dari kelompok *cluster* memiliki karakteristik yang sama dengan kelompoknya dan berbeda dengan kelompok lainnya [19],[20]. Algoritma dasar dari *K-Means Clustering* dapat kita tentukan dengan Langkah-langkah sebagai berikut.

- Menentukan jumlah *cluster* yang diinginkan.
- Memilih *cluster* secara random dan mengelompokkan data yang lainnya ke dalam *cluster-cluster* tersebut berdasarkan jarak terpendek.
- Menghitung *Centroid*/ rata-rata dari data yang ada di dihasilkan dari masing- masing cluster.
- Mengalokasikan Kembali masing-masing data kedalam *Centroid*/ rata-rata *cluster* terdekat.

$$d = \sqrt{(X_i - S_i)^2 + (Y_i - t_i)^2} \quad (1)$$

Keterangan:

d = *Euclidean Distance*

i = Banyak objek



$x,y$  = Titik koordinat objek  
 $s,t$  = Titik koordinat *Centroid*

- f. Ulangi Langkah ke-3 apabila masih ditemukan data yang berpindah *cluster* sehingga menimbulkan perubahan nilai *Centroid cluster*.

## 2.5 Persediaan Barang

Menurut Sunarto, “persediaan barang sebagai aset yang akan diperdagangkan dan dijual agar dapat memenuhi tingkat pengembalian yang diinginkan tergantung dari jenis perusahaan tersebut”. Sedangkan menurut Munawaroh menjelaskan bahwa dalam perusahaan manufaktur, persediaan lebih ke arah persediaan bahan baku yang dibutuhkan dalam proses produksi, sedangkan dalam perusahaan dagang, persediaan lebih ke arah persediaan barang yang akan dijual [21]. Berdasarkan pendapat tersebut dapat disimpulkan bahwa persediaan barang adalah aset yang diperdagangkan pada toko yang mana barang masuk merupakan persediaan dan barang keluaran merupakan penjualan atau pengeluaran barang.

## 3. HASIL DAN PEMBAHASAN

### 3.1 Hasil Penelitian

Penelitian ini menggunakan metode *Knowledge Discovery in Database* (KDD) dengan beberapa tahapan utama:

a. *Selected Data*

Data penjualan material di Toko Bangunan Maisa diambil selama periode Juli 2024–Juli 2025 dengan total 509 item untuk meningkatkan akurasi hasil data mining.

b. *Preprocessed Data*

Data mentah dibersihkan dari atribut yang tidak berisi nilai penjualan per bulan. Dari 509 data awal, tersisa 507 data yang layak digunakan dalam penelitian.

c. *Data Mining*

Pengolahan data dilakukan menggunakan algoritma *K-Means Clustering* pada *software* RapidMiner. Data dipartisi menjadi dua *cluster*, yaitu kelompok penjualan terendah dan tertinggi. Hasil *clustering* menghasilkan informasi mengenai pembagian data pada tiap kelompok serta titik pusat (*centroid*) masing-masing *cluster*.

d. *Interpretation/Evaluation*

Hasil *clustering* digunakan untuk menentukan stok produk dengan kategori terendah, sedang, dan terbanyak.

Proses *data mining* dalam penelitian ini dilakukan dengan bantuan *software* RapidMiner 10.1, yang digunakan untuk mengelompokkan data menggunakan algoritma klusterisasi *K-Means*. Data yang diolah meliputi nama material dan hasil penjualan per bulan di Toko Bangunan Maisa. Hasil pengolahan data menghasilkan informasi yang terbagi ke dalam dua *cluster*, yaitu kelompok barang dengan tingkat penjualan tinggi dan kelompok barang dengan tingkat penjualan rendah. Informasi ini sangat bermanfaat dalam mendukung keputusan penyetokan barang agar lebih tepat sasaran serta mengurangi risiko penumpukan atau kekosongan stok yang tidak diperlukan. Berikut dapat dilihat pada Tabel 1 merupakan Data Transaksi Penjualan Toko Bangunan Maisa.

**Tabel 1.** Data Transaksi Penjualan Toko Bangunan Maisa

No.	Nama Material	Harga Satuan	Harga Beli	Penjualan /Perbulan	Margin (Rp)	Margin (%)
1	Accessories Talang PVC	25000	20000	250000	5000	25
2	Alkasit Kalle	19500	15000	136500	4500	30
3	alluminium foil	4000	2000	88000	2000	100
4	amplas besi	5000	3000	55000	2000	66,67
5	amplas dinding	5000	3000	105000	2000	66,68
6	amplas kayu	5000	3000	95000	2000	66,69
7	amplas kayu	6000	4000	24000	2000	50
8	amplas SSN Tokio 400	9000	6000	27000	3000	50
9	amplasTpl willmer 100	5500	3500	22000	2000	57,14
10	amplasTpl willmer 60	4500	2500	13500	2000	80
11	Bak Cuci Piring 1 lb + syp	550000	520000	550000	30000	5,77
12	Bambu	6000	4000	60000	2000	50
13	Ban Dalam Swallow	40000	35000	200000	5000	14,29
14	Bata Ringan	7500	6000	750000	1500	25
15	Batako kapur	90000	70000	45000	20000	28,57
16	Batako semen	1700	1000	127500	700	70
17	Batu Bata	40000	30000	250000	1000	33,33
18	Batu Kali	185000	165000	92500	20000	12,12
19	Batu Koral	450000	430000	450000	20000	4,65
20	Batu Split	185000	165000	185000	20000	12,12



No.	Nama Material	Harga Satuan	Harga Beli	Penjualan /Perbulan	Margin (Rp)	Margin (%)
21	Batu Tempel Andesit Bakar	150000	130000	250000	20000	15,38

3.2 Pembahasan

Tahap awal dalam penelitian ini adalah proses pembersihan data (*data cleaning*). Pada tahap ini dilakukan eliminasi terhadap data yang kurang relevan atau tidak ada nilainya dengan menghapus nilai kosong (*missing value*) maupun data yang bersifat redundan. Pembersihan ini bertujuan untuk memastikan bahwa data yang akan diproses berada dalam kondisi normal dan konsisten sehingga tidak menimbulkan bias atau kesalahan pada tahap analisis selanjutnya. Selain itu, dilakukan pula penyesuaian atribut agar sesuai dengan kebutuhan proses data mining, sehingga data yang dipakai benar-benar representatif terhadap kondisi persediaan barang di Toko Bangunan Maisa.

a. *Cleaning Data*

Dari total 509 data awal, setelah melalui tahap pembersihan hanya tersisa 507 data yang layak digunakan sebagai basis penelitian. Data tersebut merupakan data yang sudah bebas dari ketidakkonsistenan dan telah dipastikan valid untuk diproses lebih lanjut dengan algoritma *K-Means*. Dengan demikian, hasil penelitian yang diperoleh dapat lebih akurat dan relevan, sekaligus mendukung tujuan penelitian dalam mengelompokkan barang berdasarkan tingkat penjualannya agar strategi pengelolaan persediaan menjadi lebih efektif.

b. *Proses Data Mining*

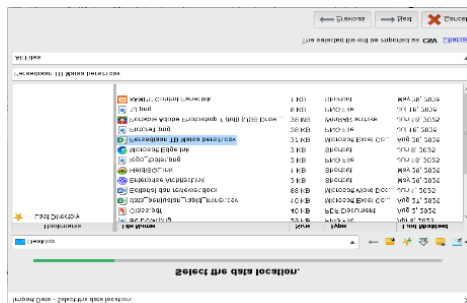
Pada tahap *data mining*, proses penggalian data dilakukan untuk menemukan pola-pola baru yang sebelumnya tidak terlihat dan kemudian diubah menjadi informasi yang bermanfaat. Dalam penelitian ini digunakan perangkat lunak *RapidMiner* sebagai alat bantu untuk mengolah data persediaan dan penjualan barang. *RapidMiner* menyediakan berbagai operator bawaan yang mendukung teknik *data mining*, termasuk algoritma *K-Means Clustering*, sehingga memudahkan peneliti dalam melakukan proses analisis secara sistematis. Model proses penggunaan *tools RapidMiner* dalam penelitian ini dirancang untuk mengelompokkan data penjualan dan persediaan barang di Toko Bangunan Maisa ke dalam beberapa *cluster* sesuai karakteristik penjualannya. Dengan demikian, hasil yang diperoleh dapat memberikan gambaran yang lebih jelas mengenai kelompok barang dengan tingkat penjualan tinggi maupun rendah, sehingga mendukung pengambilan keputusan yang lebih tepat dalam pengelolaan stok. Berikut hasilnya:

1. *Import Dataset*

Berikut ditampilkan pada Gambar 2 dan 3 yg merupakan Import Dataset.

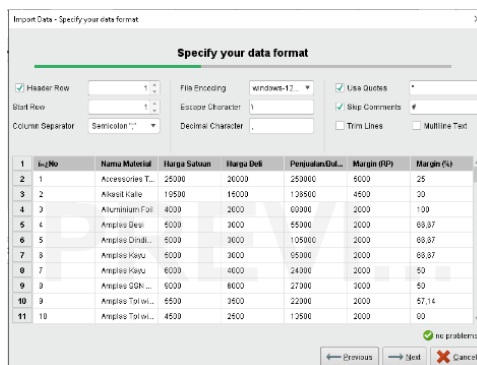
a) Buka *RapidMiner Studio*

b) Pilih *Read CSV operator*, lalu pilih file Persediaan TB Maisa bersih.csv



Gambar 2. Import Dataset

c) Pastikan delimiter ; dipilih agar data terbaca dengan benar.



Gambar 3. Proses Import Data

2. Atur tipe data

Berikut ditampilkan pada Gambar 4 merupakan Proses atur data yg terdiri dari :

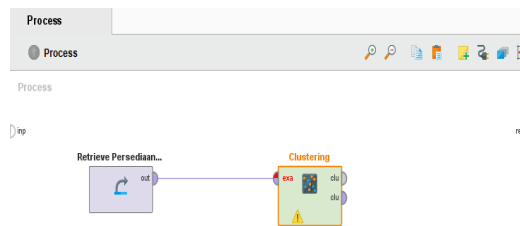


- a) Nama Material → polynominal (*string*/kategori)
- b) Penjualan/Bulan → *numeric*

IdNo	Nama Mate...	Harga Satu...	Harga Beli	Penjualan...	Margin (Rp)
integer	polynominal	integer	integer	integer	integer
1	Accessories Tele...	25000	20000	250000	5000
2	Alkasit Kalle	15000	15000	105000	4500
3	Aluminium Tali	4000	3000	88000	2600
4	Analisa Besi	5000	3000	55000	2600
5	Analisa Dinding	5000	3000	165000	2600
6	Analisa Kaca	5000	3000	85000	2600
7	Analisa Keras	4000	4000	34000	2600
8	Analisa RSH Tok	5000	4000	37000	3800
9	Analisa Tpl wlm...	5000	3500	22000	2600
10	Analisa Tpl wlm...	4000	2500	11500	2600
11	Dak Cast Pring 1...	650000	620000	650000	30000

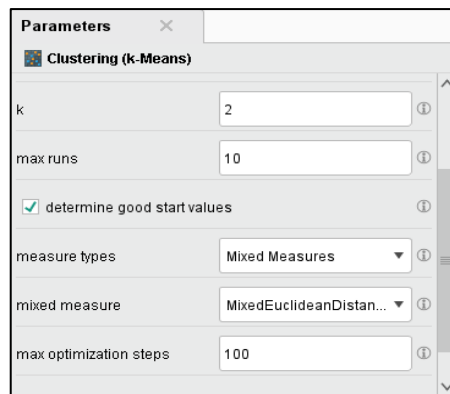
Gambar 4. Proses atur data

- 3. Pilih Atribut yang Relevan  
Untuk *clusterisasi*, bisa fokus pada kolom Penjualan/Bulan. Bisa juga menambahkan atribut lain seperti Harga Beli, Margin (Rp) apabila *cluster* ingin lebih kompleks.
- 4. Desain analisa  
Berikut ditampilkan pada Gambar 5 yang merupakan Proses desain analisa.
  - a) Masukkan data yang akan di *cluster*
  - b) Tambahkan operator *K-Means*



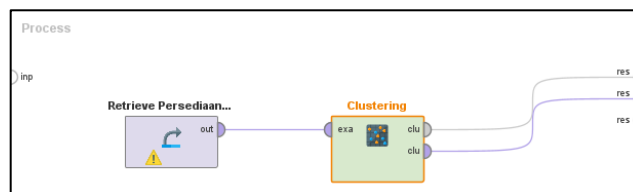
Gambar 5. Proses desain analisa

- c) *Setting* parameter sebagai berikut:  
 $k = 2$  (karena kita ingin 2 *cluster*: penjualan tinggi & rendah)  
 $max\ runs = 10$  (agar hasil lebih stabil). Berikut ditampilkan pada Gambar 6 yg merupakan Setting Parameter.



Gambar 6. Setting Parameter

- d) Hubungkan *Read CSV* → *K-Means*  
Berikut pada Gambar 7 ditampilkan hubungan antara *Read CSV* → *K-Means*.



Gambar 7. Menerapkan *K-Means*



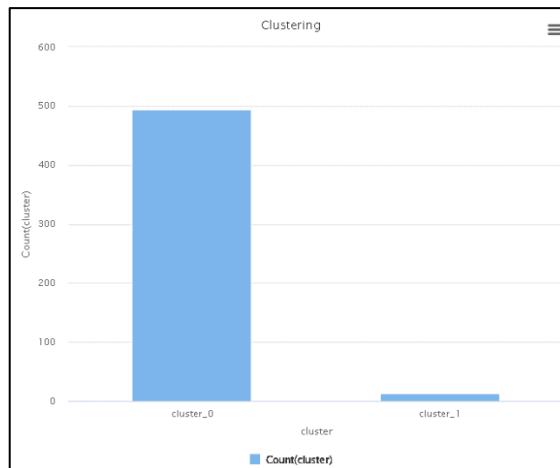
e) *Running* proses, berikut hasilnya

Row No.	Id	cluster	k/(No	Nama Mater...	Harga Satuan	Harga Beli	PenjualanB.	Margin R
1	1	cluster_0	1	Accessories	25000	20000	250000	5000
2	2	cluster_0	2	Alkali Kalle	19500	15000	136500	4500
3	3	cluster_0	3	Aluminium F...	4000	2000	88000	2000
4	4	cluster_0	4	Amplas Besi	5000	3000	55000	2000
5	5	cluster_0	5	Amplas Ondi...	5000	3000	105000	2000
6	6	cluster_0	6	Amplas Kayu	5000	3000	95000	2000
7	7	cluster_0	7	Amplas Kayu	6000	4000	24000	2000
8	8	cluster_0	8	Amplas GSH...	9000	6000	27000	3000
9	9	cluster_0	9	Amplas Tpl w...	5500	3500	22000	2000
10	10	cluster_0	10	Amplas Tpl w...	4500	2500	13500	2000
11	11	cluster_0	11	Bak Cuci Pli...	550000	520000	550000	30000
12	12	cluster_0	12	Bambu	6000	4000	60000	2000
13	13	cluster_0	13	Ban Dalam B...	40000	35000	200000	5000
14	14	cluster_0	14	Bata Ringan	7500	6000	750000	1500

Gambar 8. Hasil *Clustering*

5. Analisis Hasil

- a) Berikut pada Gambar 9 merupakan Histogram hasil *cluster* dengan Interpretasi yg terdiri dari :
  - Cluster 0* = kelompok penjualan rendah sebanyak 494 item barang
  - Cluster 1* = kelompok penjualan tinggi sebanyak 13 item barang



Gambar 9. Histogram hasil *cluster*

- b) Pada Gambar 10 menunjukkan Masuk ke menu *Visualization*, pilih *Plot Type* nya *Scatter*. Tentukan sumbu X-Axis adalah materialnya dan sumbu *Value Column* adalah penjualan per bulan



Gambar 10. Tampilan *Scatter Plot*

3.3 Interpretasi Hasil

Hasil visualisasi *clustering* menggunakan algoritma *K-Means* pada data penjualan memperlihatkan adanya pemisahan yang jelas antara dua kelompok, yaitu barang dengan tingkat penjualan tinggi dan barang dengan tingkat penjualan rendah. Pada grafik, sumbu X merepresentasikan variabel Nama Material, sedangkan sumbu Y menunjukkan besaran nilai Penjualan per Bulan. Sebagian besar material berada pada area bawah grafik dengan nilai penjualan relatif kecil dan ditandai dengan warna biru. Kondisi ini menunjukkan bahwa mayoritas barang termasuk dalam *cluster* penjualan rendah. Sebaliknya, terdapat sejumlah kecil titik data yang terletak jauh lebih tinggi pada sumbu Y dengan warna kuning hingga

merah, yang mengindikasikan terbentuknya cluster penjualan tinggi, meskipun jumlahnya tidak sebanyak *cluster* pertama. Distribusi tersebut menunjukkan bahwa pola penjualan barang di toko cenderung tidak merata, di mana hanya sebagian kecil produk yang memberikan kontribusi signifikan terhadap total penjualan bulanan, sedangkan sebagian besar produk berada pada kategori penjualan rendah. Temuan ini memberikan dasar analitis bagi pengambilan keputusan manajerial, khususnya dalam menentukan fokus strategi pemasaran, pengelolaan persediaan, serta penetapan prioritas terhadap produk unggulan yang berkontribusi besar pada pendapatan perusahaan.

### 3.4 Analisis Karakteristik Cluster

Analisis karakteristik *cluster* dengan menggunakan atribut Nama Material sebagai identifikasi produk dan Penjualan/Bulan sebagai variabel numerik menghasilkan dua kelompok utama. Pemisahan antara kedua *cluster* ini menunjukkan bahwa pola penjualan toko bersifat tidak merata, di mana hanya sedikit produk yang mendominasi nilai penjualan, sedangkan sebagian besar produk lainnya memiliki perputaran relatif lambat. Hasil ini memberikan dasar analitis bagi pihak manajemen toko untuk menentukan fokus strategi, misalnya memperkuat ketersediaan dan promosi pada produk *cluster* kedua, sambil tetap mengelola produk *cluster* pertama agar tidak menimbulkan biaya penyimpanan yang berlebihan.



Gambar 4.1 Analisis Cluster

Berdasarkan Gambar 11 yang merupakan hasil pemodelan K-Means clustering menggunakan atribut Nama Material dan Penjualan per Bulan, terbentuk dua kelompok data yang ditampilkan dalam grafik scatter plot. Cluster 0 (biru) mencakup sebagian besar material yang memiliki tingkat penjualan bulanan relatif rendah, umumnya berada di bawah 2 juta rupiah per bulan. Hal ini menunjukkan bahwa mayoritas produk termasuk dalam kategori slow moving atau produk dengan perputaran rendah, meskipun jumlah item dalam cluster ini cukup banyak. Cluster 1 (hijau) terdiri atas sejumlah kecil material, namun dengan nilai penjualan bulanan yang jauh lebih tinggi, yaitu berkisar antara 3 juta hingga lebih dari 10 juta rupiah per bulan. Produk-produk dalam cluster ini dapat dikategorikan sebagai fast moving items atau produk unggulan yang memiliki kontribusi signifikan terhadap total penjualan toko. Visualisasi pada grafik menunjukkan pola yang jelas: terdapat dominasi produk berpenjualan rendah dalam jumlah besar, dan hanya sebagian kecil produk yang memiliki penjualan sangat tinggi. Temuan ini mengindikasikan bahwa strategi manajemen persediaan sebaiknya difokuskan pada pengendalian jumlah stok untuk produk cluster 0 agar tidak menimbulkan biaya simpan berlebih, sekaligus memperkuat ketersediaan dan strategi pemasaran untuk produk cluster 1 sebagai penopang utama penjualan.

### 3.5 Implikasi Manajerial

Hasil analisis clustering memberikan gambaran yang penting bagi manajemen dalam merumuskan strategi bisnis yang lebih tepat sasaran. Dengan ditemukannya dua kelompok utama material, yaitu cluster 0 (produk dengan penjualan rendah) dan cluster 1 (produk dengan penjualan tinggi), perusahaan dapat mengambil langkah strategis untuk meningkatkan efektivitas pengelolaan persediaan maupun pemasaran. Bagi cluster 0, strategi manajerial yang dapat dilakukan adalah dengan menerapkan inventory control yang lebih ketat, misalnya dengan menurunkan jumlah stok minimum, melakukan promosi khusus, atau bahkan mengevaluasi ulang kelayakan produk tersebut untuk terus dipasarkan. Langkah ini penting agar tidak terjadi penumpukan barang yang justru meningkatkan biaya penyimpanan dan mengurangi efisiensi operasional. Sebaliknya, untuk cluster 1, yang merupakan produk dengan penjualan tinggi, perusahaan perlu memastikan ketersediaan barang tetap terjaga melalui safety stock yang memadai dan sistem distribusi yang lebih responsif. Selain itu, strategi pemasaran juga dapat difokuskan pada produk cluster ini dengan memberikan promosi tambahan atau paket bundling guna semakin meningkatkan volume penjualan. Secara keseluruhan, hasil clustering dapat dijadikan dasar dalam mengembangkan strategi diferensiasi manajemen persediaan serta fokus pemasaran, sehingga perusahaan tidak hanya mampu mengendalikan biaya operasional tetapi juga dapat memaksimalkan potensi keuntungan dari produk yang menjadi andalan. Data pada cluster 0 yang merupakan kelompok penjualan rendah sebanyak 494 item barang dan pada cluster 1 yang merupakan kelompok penjualan tinggi sebanyak 13 item barang. Data terlampir pada halaman lampiran.



#### 4. KESIMPULAN

Berdasarkan hasil penelitian, penerapan algoritma K-Means Clustering terbukti mampu mengelompokkan persediaan barang berdasarkan tingkat penjualan secara objektif dengan memanfaatkan data historis. Metode ini efektif dalam mengidentifikasi pola penjualan sehingga memudahkan analisis kondisi persediaan barang. Metode K-Means Clustering dengan bantuan RapidMiner membagi data penjualan menjadi 2 cluster utama, yaitu cluster\_0 (barang dengan tingkat penjualan rendah) dan cluster\_1 (barang dengan tingkat penjualan tinggi). Hasil pengelompokan persediaan barang menggunakan algoritma K-Means Clustering dapat digunakan sebagai dasar pendukung pengambilan keputusan dalam pengelolaan persediaan. Informasi yang dihasilkan membantu pemilik toko dalam menentukan prioritas penyetokan dan pengendalian stok agar pengelolaan persediaan menjadi lebih efektif dan efisien.

#### REFERENCES

- [1] R. Y. Simanullang, P. Wanny, and S. M. Rambe, "Implementation of the K-Means Algorithm on Smart Systems in Grouping Gadget Accessory Purchase Patterns," *JITCSE (Journal Inf. Technol. Komput. Sci. Electr. Eng.*, vol. 2, no. 3, pp. 22–31, 2026, doi: 0.61306/jitese.
- [2] D. Riswanda and A. T. Priandika, "Analisis dan Perancangan Sistem Informasi Manajemen," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 2, no. 1, pp. 94–101, 2021.
- [3] A. Adiyanto and Y. Arie Wijaya, "Penerapan Algoritma K-Means Pada Pengelompokan Data Set Bahan Pangan Indonesia Tahun 2022-2023," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 2, pp. 1344–1350, 2023, doi: 10.36040/jati.v7i2.6849.
- [4] F. Febriansyah and S. Muntari, "Penerapan Algoritma K-Means untuk Klasterisasi Penduduk Miskin pada Kota Pagar Alam," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 8, no. 1, pp. 66–77, 2023, doi: 10.14421/jiska.2023.8.1.66-77.
- [5] A. Patappari and N. Muhlisa, "Sistem Informasi Persediaan Barang Berbasis Web Pada Toko Throve Store Soppeng," *J. Ilm. Sist. Inf. dan Tek. Inform.*, vol. 6, no. 1, pp. 1–8, 2023, doi: 10.57093/jisti.v6i1.142.
- [6] A. Putera, U. Siahaan, N. A. Harahap, and R. Y. Simanullang, "Analysis of Inpatient Data Using Cluster Analysis on Simulation Dataset," *J. BIT (Bulletin Inf. Technol.*, vol. 6, no. 1, pp. 33–39, 2025, doi: 10.47065/bit.v5i2.1830.
- [7] P. S. Jetlie Kang, "Analisis Pola Penjualan Produk Elektronik Pada E-Commerce Menggunakan Algoritma FP-Growth," *J. Comasie*, vol. 01, 2025.
- [8] R. Wahyusari, "Penerapan Algoritma FP-Growth Untuk Menemukan Pola Peminjaman Alat Pada Workshop Teknik Mesin," *Log. J. Ilmu Komput. dan Pendidik.*, vol. 1, no. 3, pp. 406–411, 2023, [Online]. Available: <https://journal.mediapublikasi.id/index.php/logic/article/view/2745%0Ahttps://journal.mediapublikasi.id/index.php/logic/article/download/2745/1211>
- [9] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan, "Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner," *JBASE - J. Bus. Audit Inf. Syst.*, vol. 4, no. 1, pp. 13–21, 2021, doi: 10.30813/jbase.v4i1.2729.
- [10] A. Supriyadi, A. Triayudi, and I. D. Sholihati, "Perbandingan Algoritma K-Means Dengan K-Medoids Pada Pengelompokan Armada Kendaraan Truk Berdasarkan Produktivitas," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 6, no. 2, pp. 229–240, 2021, doi: 10.29100/jupi.v6i2.2008.
- [11] I. Maryani, O. Revianti, H. M. Nur, and S. Sunanto, "Implementasi Data Mining Pada Penjualan Di Toko GOC Kosmetik Dengan Menggunakan Metode Algoritma Apriori," *Indones. J. Softw. Eng.*, vol. 8, no. 1, pp. 92–98, 2022.
- [12] P. M. S. Tarigan, J. T. Hardinata, H. Qurniawan, M. Safii, and R. Winanjaya, "Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang: Studi Kasus: Toko Sinar Harahap," *J. Janitra Inform. dan Sist. Inf.*, vol. 2, no. 1, pp. 9–19, 2022.
- [13] E. Marlina *et al.*, "Implementasi Data Mining Untuk Memprediksi Status Gizi Anak Balita Pada Puskesmas Gedung Sari Menggunakan Polynomial Regression," *J. Teknol. Komput. dan Sist. Inf.*, vol. 6, no. 2, pp. 193–196, 2023.
- [14] H. Mustakim and S. Priyanta, "Aspect-Based Sentiment Analysis of KAI Access Reviews Using NBC and SVM," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 16, no. 2, p. 113, 2022, doi: 10.22146/ijccs.68903.
- [15] L. Al-Alawi, J. Al Shaqsi, A. Tarhini, and A. S. Al-Busaidi, "Using machine learning to predict factors affecting academic performance: the case of college students on academic probation," *Educ. Inf. Technol.*, vol. 28, no. 10, pp. 12407–12432, 2023, doi: 10.1007/s10639-023-11700-0.
- [16] J. Maulani and M. Sari, "Komparasi Metode K-Nearest Neighbor (Knn) Dengan Support Vector Machine (Svm) Terhadap Tingkat Akurasi Klasifikasi Kualitas Air," *Smart Comp Jurnalnya Orang Pint. Komput.*, vol. 12, no. 2, pp. 430–435, 2023, doi: 10.30591/smartcomp.v12i2.4205.
- [17] P. A. Myers *et al.*, "pyMAISE: A Python platform for automatic machine learning and accelerated development for nuclear power applications," *Prog. Nucl. Energy*, vol. 180, p. 105568, 2025.
- [18] F. Ekundayo, I. Atoyebi, A. Soyele, and E. Ogunwobi, "Predictive analytics for cyber threat intelligence in fintech using big data and machine learning," *Int J Res Publ Rev*, vol. 5, no. 11, pp. 1–15, 2024.
- [19] D. P. Indini, Mesran, and Dito Putro Utomo, "Penerapan Data Mining Dalam Pengelompokan Data Reseller di Telkomsel Authorized Partner (TAP) Deli Tua Dengan Algoritma K-Means," *J. Ilm. Media Sisfo*, vol. 17, no. 2, pp. 189–202, 2023, doi: 10.33998/mediasisfo.2023.17.2.1391.
- [20] R. Dila, S. Defit, and S. Arlis, "Analisis Algoritma K-Means Clustering Dalam Pengelompokan Prestasi Belajar Siswa Menengah Atas ( SMA )," vol. 5, no. 5, pp. 1113–1119, 2025, doi: 10.47065/bulletincsr.v5i5.751.
- [21] J. Yandi and K. Wijaya, "Rancang Bangun Aplikasi Persediaan Barang Pada Counter Karya Cellmenggunakan Vb Net," *JSK (Jurnal Sist. Inf. dan Komputerisasi Akuntansi)*, vol. 7, no. 1, pp. 1–6, 2023, doi: 10.56291/jsk.v7i1.109.