

## Implementasi Algoritma K-Means Guna Pengelompokan Data Penjualan Berdasarkan Pembelian

Siti Sahara Lubis, Sarjon Defit, Sumijan

Fakultas Ilmu Komputer, Universitas Putra Indonesia “YPTK” Padang, Kota, 25221, Indonesia

[saharalubis448@gmail.com](mailto:saharalubis448@gmail.com)

### Abstract

Information technology can help solve problems faced by humans by facilitating performance. Currently, the development of information technology is very rapid and fast, including in Indonesia itself. The development of information technology has opened a new chapter in the community, including in the business world. One form of information technology development can be seen in data mining. Data mining is a step in (KDD) that has a technique for analyzing data to dig up hidden information in large and complex amounts. Based on this, this study aims to provide a real solution to UD. Martua to find out which goods are selling well and which goods are not selling well so that objects can find out which goods need to be added to their stock and which goods need to be reduced in stock. The method used in this study is the K-Means method with stages, namely data selection, pre-processing, data transformation, information extraction, and evaluation of results. The data consists of 30 item data, there are 8 as members of C1 and are best-selling items and 22 as members of C2 and are not selling items. The conclusion that can be obtained from this study is that the K-Means method can group goods at UD. Martua. This study shows that the implementation of the K-Means method with the support of the RapidMiner application is effective in grouping goods data at UD Martua.

keyword : Information Technology, Data Mining, K-Means, Clustering, Goods

### Abstrak

Teknologi informasi mampu membantu menyelesaikan permasalahan yang dihadapi manusia dengan memudahkan kinerja. Saat ini perkembangan teknologi informasi sangatlah pesat dan cepat termasuk di Indonesia sendiri. Perkembangan teknologi informasi telah membuka babak baru dilingkungan masyarakat, termasuk didunia bisnis. Salah satu bentuk perkembangan teknologi informasi dapat dilihat dari data mining. Data mining merupakan suatu langkah dalam (KDD) yang memiliki teknik menganalisa data untuk digali informasi tersembunyi dalam jumlah besar dan kompleks. Berdasarkan hal tersebut maka penelitian ini bertujuan untuk memberikan solusi nyata kepada UD. Martua agar dapat mengetahui mana barang yang laris dan mana barang yang tidak laris agar objek dapat mengetahui barang yang perlu di tambah stoknya dan barang yang harus di kurangi stoknya. Metode yang digunakan pada penelitian ini adalah metode *K-Means* dengan tahapan yaitu pemilihan data, pra-pemrosesan, transformasi data, ekstraksi informasi dan evaluasi hasil. Data terdiri dari 30 data barang terdapat 8 sebagai anggota C1 dan merupakan barang laris dan 22 sebagai anggota C2 dan merupakan barang tidak laris. Kesimpulan yang dapat diperoleh dari penelitian ini metode *K-Means* dapat mengelompokkan barang pada UD.Martua. Penelitian ini menunjukkan bahwa implementasi metode *K-Means* dengan dukungan aplikasi *RapidMiner* efektif dalam mengelompokkan data barang pada UD Martua.

Kata Kunci :Teknologi Informasi , Data Mining, *K-Means*, *Clustering*, Barang



### 1. Pendahuluan

Teknologi informasi mampu membantu menyelesaikan permasalahan yang dihadapi manusia dengan memudahkan kinerja[1]. Saat ini perkembangan teknologi informasi sangatlah pesat dan cepat termasuk di Indonesia sendiri. Teknologi pada dasarnya berguna untuk mempermudah manusia dalam menjalankan sesuatu hal. Teknologi informasi ini sudah banyak digunakan untuk memproses, mengolah data,

menganalisis data untuk menghasilkan data atau informasi yang relevan, cepat, jelas dan akurat[2]. Teknologi informasi sudah banyak digunakan dilembaga pemerintahan atau pun perusahaan swasta dan institusi lainnya. Perkembangan teknologi informasi telah membuka babak baru dilingkungan masyarakat, termasuk didunia bisnis[3]. Didunia bisnis saat ini hampir semua transaksi menggunakan teknologi[4]. Bisnis sangat erat hubungannya dengan

penjualan, Penjualan adalah kegiatan pelengkap dari pembelian untuk terjadinya proses transaksi [5]. Dalam konteks bisnis dan analitik data, KDD (*Knowledge Discovery in Databases*) adalah proses yang digunakan untuk menemukan informasi atau pengetahuan yang berguna dari data yang besar dan kompleks [6]. Data mining merupakan suatu langkah dalam (KDD) yang memiliki teknik menganalisa data untuk digali informasi tersembunyi dalam jumlah besar dan kompleks [7]. Proses *Knowledge Discovery in Database* secara garis besar terdiri dari 5 tahapan yaitu data *selection*, *pre-processing/cleaning*, *transformation*, Data mining dan *interpretation/evaluation* [8]. *Clustering* merupakan suatu proses di mana data dikelompokkan ke dalam berbagai kelompok yang berbeda, dengan tujuan agar objek-objek serupa tergabung dalam satu kelompok atau *cluster*. [9]. *K-Means* merupakan sebuah algoritma clustering pada data mining untuk dapat menghasilkan kelompok dari data yang jumlahnya banyak dengan metode partisi yang berbasis titik dengan waktu komputasi yang cepat dan efisien. [10].

Penelitian yang dilakukan oleh Aprillia Anjani yang membahas tentang klasterisasi data penjualan terlaris pada perusahaan kosmetik brand Y.O.U yang didedikasikan untuk kecantikan wanita Indonesia. Metode yang sangat ideal sebagai algoritma yang bisa mengelola data dengan cara lebih efisien adalah algoritma K-Means. Penelitian ini bertujuan untuk melihat strategi penjualan pada periode berikutnya dapat mengetahui barang yang paling terlaris yang dijual dan barang yang kurang laris. Dengan data tersebut dapat dikelompokkan menjadi 3 Clustering yang didasarkan pada tingkat kesuksesan penjualan suatu jenis produk Cluster 0 yang terdiri dari 8 produk item yang memiliki tingkat penjualan laris, Sedangkan Dalam Cluster 1 yang mencakup 12 item produk dengan kategori paling laris dan Cluster 2 yang terdiri dari 5 item produk yang tidak laris, dalam Cluster ini menunjukkan tren tetap atau bahkan menurun dari bulan ke bulan dengan nilai Centroid menunjukkan ketidakpopuleran produk-produk tersebut di pasar. [10].

Penelitian yang dilakukan Silvia Lestari dkk, tentang penjualan alat perabot pada PT.XYZ, yang dimana data penjualan masih dilakukan dengan cara manual yang berakibat kurangnya peninjauan terhadap produk yang dibutuhkan konsumen. Untuk mengatasi permasalahan dilakukan penelitian menggunakan Algoritma K-Means. Penelitian ini menggunakan sampel sebanyak 10 dan mendapatkan hasil barang paling laris sebanyak 5, dan barang yang kurang laris sebanyak 5. Kesimpulan pada penelitian ini dapat memberikan solusi kepada Pihak perusahaan agar dapat mengetahui mana barang yang laris, dan tidak laris [11].

Penelitian juga dilakukan oleh Marina Adelina dkk, tentang mengelompokkan transaksi penjualan barang pada toko Arino. Penelitian yang digunakan Untuk

meningkatkan efisiensi manajemen toko adalah menerapkan metode algoritma *K-Means Clustering* dalam mengelompokkan data transaksi penjualan. Produk yang memiliki karakteristik konsumen serupa dikelompokkan bersama untuk meningkatkan aksesibilitas. Data yang di gunakan pada penelitian ini 1203 data dan 5 *cluster*, mendapatkan hasil *cluster* 0 memiliki 1044 anggota, *cluster* 1 memiliki 1 anggota, *cluster* 2 memiliki 18 anggota, *cluster* 3 memiliki 134 anggota dan *cluster* 4 memiliki 6 anggota. Kesimpulan pada penelitian ini adalah implementasi algoritma *clustering* memberikan wawasan berharga terkait pola penjualan, preferensi pelanggan, dan optimalisasi penempatan barang di toko [12].

Siti Farikah dkk, meneliti tentang. Permasalahan yang ada pada Toko Mahabbatain belum adanya klasifikasi terhadap menentukan jumlah produk terhadap pengelompokan katalog. Metode yang digunakan dalam penelitian ini adalah algoritma K-Means. Dari penelitian yang sudah dilakukan secara observasi dan dokumentasi menunjukkan bahwa dataset dari pembukaan Toko Mahabbatain Store yang dengan jumlah 499 data penjualan pada tahun 2023 dengan jumlah *cluster* 2 mendapatkan hasil *cluster* 0 sebanyak 185 anggota dan anggota *cluster* 1 memiliki 314 anggota. Kesimpulan dari penelitian ini adalah penerapan algoritma *K-means* dapat mengetahui hasil k optimum dalam penerapan menggunakan RapidMiner. Hasil uji coba cluster distance performance bahwa nilai k optimum adalah k=2 dengan nilai DBi=0,047 [13].

Penelitian yang dilakukan oleh Aprillia Anjani yang membahas tentang klasterisasi data penjualan terlaris pada perusahaan kosmetik brand Y.O.U yang didedikasikan untuk kecantikan wanita Indonesia. Metode yang sangat ideal sebagai algoritma yang bisa mengelola data dengan cara lebih efisien adalah algoritma K-Means. Penelitian ini bertujuan untuk melihat strategi penjualan pada periode berikutnya dapat mengetahui barang yang paling terlaris yang dijual dan barang yang kurang laris. Dengan data tersebut dapat dikelompokkan menjadi 3 Clustering yang didasarkan pada tingkat kesuksesan penjualan suatu jenis produk Cluster 0 yang terdiri dari 8 produk item yang memiliki tingkat penjualan laris, Sedangkan Dalam Cluster 1 yang mencakup 12 item produk dengan kategori paling laris dan Cluster 2 yang terdiri dari 5 item produk yang tidak laris, dalam Cluster ini menunjukkan tren tetap atau bahkan menurun dari bulan ke bulan dengan nilai Centroid menunjukkan ketidakpopuleran produk-produk tersebut di pasar. [14].

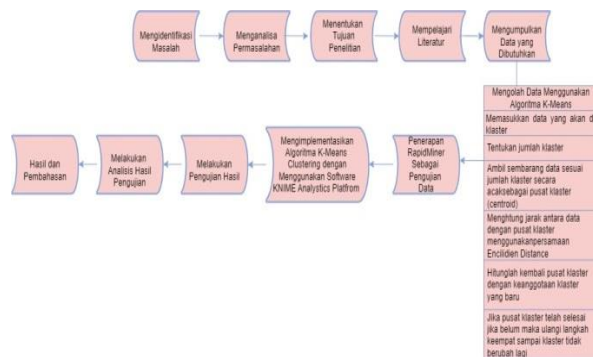
Penelitian yang dilakukan oleh Miftahul Fajar dkk, yang meneliti tentang analisis pola penjualan obat pada Apotek An-Naafi, mengalami minimnya informasi untuk pengembangan bisnis dan produk karena belum melakukan pengelompokan data penjualan obat berdasarkan perilaku konsumen. Penelitian ini menerapkan metode K-Means untuk menganalisis data

penjualan obat Metode K Means diterapkan dengan inisialisasi cluster yang sesuai, membentuk 2 cluster utama mencerminkan tingkat penjualan: rendah dan tinggi. Hasil cluster, dengan cluster 1 menonjol dalam penjualan dan pendapatan tinggi, sedangkan cluster 0 lebih beragam. Cluster 0 mencakup 381 hasil penjualan dengan pendapatan sebesar Rp. 668.767, yang tergolong rendah. Sementara itu, cluster 1 menunjukkan hasil unggul dengan 521 penjualan dan pendapatan sebesar Rp. 3.353.880, yang dikategorikan sebagai tinggi. [15].

Berdasarkan literatur diatas meskipun penelitian sebelumnya sudah mengeksplorasi penggunaan metode K-Means dengan banyak bidang penelitian sebelumnya mungkin tidak secara spesifik dalam membahas tentang pengelompokkan produk. Penelitian yang saat ini saya lakukan yaitu pengelompokkan produk adalah hal yang perlu diteliti karena dapat memudahkan dalam penentuan tingkat kelarisan suatu produk. Penelitian ini bertujuan untuk mengetahui produk mana saja yang laris dan kurang laris pada toko agar pihak toko dapat mengetahui produk mana saja yang harus di stok dan produk mana saja yang harus di kurangi stoknya. Oleh karena itu, penelitian ini diharapkan tidak hanya melengkapi hasil penelitian sebelumnya, tetapi juga memberikan kontribusi baru yang lebih tepat sasaran.

## 2. Metodologi Penelitian

Metodologi penelitian ini menjelaskan secara sistematis berguna sebagai pedoman untuk melakukan penelitian. Mulai dari pengumpulan data dan bahan penelitian yang sesuai dari jurnal untuk permasalahan yang diteliti. Bertujuan agar lebih terarah dan di harapkan bisa berjalan dengan baik sesuai dengan tujuan awal dari penelitian ini. Metodologi penelitian ini di gambarkan pada sebuah kerangka kerja dan peneliti akan menguraikan tahapan-tahapan dari kerangka kerja dari penelitian tersebut. Metodologi penelitian ini di gambarkan pada sebuah kerangka kerja yang dapat dilihat pada Gambar 1.



Gambar 1 Kerangka Kerja Penelitian

Metodologi penelitian ini di gambarkan pada sebuah kerangka kerja dan peneliti akan menguraikan tahapan-tahapan dari kerangka kerja dari penelitian tersebut. Kerangka kerja ini memperlihatkan hubungan antara variabel-variabel utama yang diteliti serta pengaruhnya

satu sama lain dalam konteks penelitian. Secara keseluruhan, gambar ini memberikan visualisasi yang jelas mengenai metodologi yang diterapkan dan tahapan-tahapan yang dilakukan selama penelitian. Tahapan-tahapan dari kerangka kerja penelitian adalah sebagai berikut :

### 1. Mengidentifikasi Masalah

Tahapan identifikasi masalah merupakan tahap awal dalam melakukan penelitian. Proses mengidentifikasi masalah bertujuan untuk meninjau objek yang akan di teliti, menentukan ruang lingkup dan pencarian permasalahan pada objek penelitian. Masalah dirumuskan dengan jelas dan spesifik agar dapat dijadikan fokus dalam penelitian atau penyelesaian.

### 2. Menganalisa Permasalahan

Memahami permasalahan dalam ruang lingkup dan batasan masalah yang telah ditentukan, sehingga permasalahan akan lebih mudah dipahami. Analisa masalah pada penelitian ini yaitu kurang tertatanya objek UD.Martua dalam mengelompokkan barang yang laris dan kurang laris. Dengan analisis yang mendalam, solusi atau alternatif penyelesaian yang efektif dapat dirumuskan.

### 3. Menentukan Tujuan dari penelitian

Setelah menganalisa permasalahan dan didapatkan rumusan permasalahan, selanjutnya menentukan tujuan dari penelitian yang ingin di capai. Penelitian ini di harapkan mampu menyelesaikan masalah dan bisa mengoptimalkan pengelompokkan barang di UD. Martua.

### 4 . Mempelajari Literatur

Mempelajari literatur melibatkan penelusuran dan kajian terhadap karya-karya ilmiah yang relevan dengan topik penelitian. Proses ini membantu peneliti memahami perkembangan terkini, menemukan celah penelitian, dan mengidentifikasi metode yang sesuai. Sumber literatur dalam penelitian ini dari jurnal-jurnal ilmiah terdahulu yang membahas Data Mining tentang Algoritma *K-means Clustering*.

### 5 . Mengumpulkan data-data yang dibutuhkan

Mengumpulkan data-data yang dibutuhkan melibatkan identifikasi sumber data yang relevan dan metode pengumpulan yang sesuai, seperti survei, wawancara, atau eksperimen. Peneliti kemudian mengumpulkan data secara sistematis, memastikan akurasi dan konsistensi informasi yang diperoleh. Data yang terkumpul ini menjadi dasar untuk analisis lebih lanjut dalam rangka menjawab pertanyaan penelitian. Beberapa metode pengumpulan data yang dilakukan pada peneliti antara lain sebagai berikut:

#### a . Observasi

Observasi adalah metode penelitian di mana peneliti mengamati dan mencatat kejadian yang terjadi pada objek penelitian. Teknik ini memungkinkan peneliti untuk mendapatkan data yang mencerminkan situasi nyata. Observasi yang dimaksud adalah melakukan pengamatan data pada UD. Martua, yang bertujuan mengetahui permasalahan yang ada untuk mengetahui jumlah stik barang dan jumlah barang yang terjual.

#### b. Wawancara

Wawancara penelitian adalah metode pengumpulan data di mana peneliti bertanya langsung kepada responden untuk mendapatkan informasi mendalam tentang topik yang sedang diteliti. Untuk mendapatkan informasi peneliti melakukan wawancara peneliti wawancara peneliti mewawancarai pemilik toko UD. Martua guna mencari tau penyebab terjadinya permasalahan untuk di angkat menjadi topik penelitian ini. Data yang diperoleh dari wawancara dapat memberikan wawasan yang kaya dan mendalam, mendukung analisis dan interpretasi penelitian.

#### c. Studi Pustaka

Studi pustaka adalah metode penelitian yang melibatkan peninjauan dan analisis terhadap sumber-sumber tertulis yang relevan dengan topik penelitian. Tujuannya adalah untuk memahami teori, konsep, dan temuan yang telah ada guna membangun kerangka teoritis yang solid. Dengan studi pustaka, peneliti dapat mengidentifikasi celah dalam literatur dan mengarahkan penelitian ke arah yang lebih spesifik dan berkontribusi pada bidang ilmu yang dipelajari.

#### 6. Mengolah Data menggunakan Algoritma K-Means Clustering

Mengolah data menggunakan algoritma K-Means Clustering melibatkan pembagian data ke dalam kelompok atau cluster berdasarkan kesamaan fitur yang dimiliki. Pada proses pengolahan data, data didapat langsung dari Toko UD. Martua. Data yang di peroleh sebanyak 30 data dengan *cluster* 1 dan *cluster* 2 dan di proses dengan menggunakan Algoritma K-Means Clustering [16] adalah sebagai berikut:

- Masukkan data yang akan di *klaster*
- Tentukan jumlah klaster
- Ambil sembarang data sejumlah klaster secara acak sebagai pusat klaster (*centroid*) dan selanjutnya untuk menentukan nilai *centroid* iterasi berikutnya digunakan rumus:

$$V = \frac{\sum_{i=1}^n x}{n}; I=1, 2, 3, \dots, n \quad (1)$$

Dimana:

V : *Centroid* baru

n : Banyak data dalam *cluster*

- Menghitung kembali setiap objek yang menggunakan pusat *cluster* (*centroid*) yang baru.

Jika proses ini pusat *cluster* tidak berubah maka proses *Clustering* selesai. Dan jika masih mengalami perubahan maka diulangi kembali ke langkah 3 sampai pusat *cluster* tidak berubah. Akan tetapi *Clustering* juga dapat dihentikan dengan menggunakan batasan maksimal terhadap suatu proses *Clustering*.

- Menghitung jarak antara data dengan pusat klaster menggunakan persamaan *Euclidean* yang dirumuskan sebagai berikut:
- Hitunglah kembali pusat klaster dengan keanggotaan klaster yang baru
- Jika pusat klaster tidak ada perubahan maka proses klaster telah selesai. Jika belum maka ulangi langkah ke 4 sampai klaster tidak berubah lagi.

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

Dimana:

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

X<sub>ki</sub> = Data ke i pada atribut data ke k

X<sub>kj</sub> = Titik pusat ke j pada atribut ke k

#### 7. Penerapan RapidMiner

Pengujian data dilakukan menggunakan *software RapidMiner*. Tujuan instalasi *software* untuk melakukan pengujian data yang telah di olah pada *Microsoft Excel*. Bertujuan untuk pengujian apakah data yang di olah mendapatkan hasil yang sama.

#### 8. Mengimplementasikan Algoritma K-means Clustering dengan menggunakan software RapidMiner

Mengimplementasikan algoritma K-Means Clustering dengan menggunakan operator KNIME Analytics Platform dimulai dengan mengimpor dataset ke dalam KNIME. Peneliti kemudian menambahkan operator K-Means Clustering ke dalam alur kerja dan mengonfigurasi parameter, seperti jumlah cluster dan fitur yang digunakan. Setelah menjalankan proses clustering, KNIME menghasilkan output yang memvisualisasikan hasil clustering dan memungkinkan analisis lebih lanjut untuk mengevaluasi kualitas dan struktur cluster yang terbentuk.

#### 9. Melakukan Pengujian Hasil

Setelah proses *clustering* dilakukan dengan menggunakan algoritma *K-means*. Pengujian hasil dilakukan untuk mengevaluasi pembentukan *cluster*. Beberapa tahapan untuk melakukan pengujian sistem antara lain:

- Pengujian dilakukan menggunakan *Microsoft Excel* dengan Algoritma *K-means Clustering*.
- Melakukan pengujian pada *software rapid miner* menggunakan Algoritma *K-means Clustering*.
- Membandingkan hasil yang di peroleh menggunakan *Microsoft Excel* dan *software rapid*

*miner* apakah hasil yang di peroleh memiliki tingkat akurasi yang sama sesuai dengan yang harapan.

## 10. Melakukan Analisa Hasil Pengujian

Melakukan analisa hasil pengujian yang diperoleh dari data yang telah diolah menggunakan Algoritma *K-means Clustering*. Tujuan untuk mengidentifikasi apakah hasilnya sama antara perhitungan manual dengan *software RapidMiner*.peneliti dapat menentukan kesimpulan dari hasil Analisa penelitian.

## 11. Hasil dan Pembahasan

Hasil dan Pembahasan menyajikan temuan utama dari penelitian dan menjelaskan implikasi dari hasil tersebut dalam konteks masalah yang diteliti. Peneliti menguraikan data yang diperoleh, membandingkannya dengan sebelum di lakukannya penelitian. Pada bagian ini juga mencakup mengenai keterbatasan penelitian dan kemungkinan arah penelitian selanjutnya.

### 3. Hasil dan Pembahasan

#### 3.1. Menentukan Jumlah *Cluster*

Menentukan jumlah *cluster* yang tepat memastikan bahwa data dikelompokkan secara efektif dan sesuai dengan pola yang mendasarinya. Pada tahapan yang pertama peneliti membagi data menjadi 2 bagian agar data yang digunakan dapat memudahkan peneliti dalam pengelompokan jenis barang yang akan diolah. Pada tahap ini peneliti membagi data menjadi 2 *cluster* yaitu *cluster C1* menyatakan barang laris dan *cluster C2* menyatakan barang tidak laris

#### 3.2. Menentukan nilai centroid secara acak

Menentukan nilai centroid secara acak melibatkan pemilihan titik awal dari dataset sebagai pusat *cluster* yang sementara. Proses ini dilakukan dengan memilih beberapa data secara acak dari dataset untuk menjadi centroid awal. Penentuan nilai centroid awal dapat diambil secara acak dari 30 data barang yang akan digunakan dalam penelitian. Centroid untuk *C1* diambil dari data sampel ke 23 yaitu (550, 340) dan *C2* diambil dari data sampel ke 8 yaitu (228, 195).

#### 3.3. Perhitungan Euclidean Distance

Perhitungan Euclidean Distance mengukur jarak antara dua titik dalam ruang multidimensi dengan menggunakan rumus yang menghitung akar kuadrat dari jumlah kuadrat selisih antara koordinat mereka. Dalam konteks *K-Means Clustering*, jarak ini digunakan untuk menentukan seberapa dekat setiap data dengan centroid cluster. Proses ini membantu dalam mengelompokkan data ke dalam cluster yang tepat dengan meminimalkan jarak antara data dan *centroid* yang relevan. Proses *clustering* pada data data stok

barang dan jumlah terjual ini menghasilkan enam iterasi. Hasil pada iterasi 1 dapat di lihat pada Tabel 1.

Tabel 1 Hasil Iterasi 1

No	C1	C2	Cluster
M1	710,8867702	1043,902773	C1
M2	9484,536046	9818,783682	C1
M3	6489,107797	6826,360158	C1
M4	3437,128162	3771,486312	C1
M5	365,0054794	679,0346088	C1
M6	12268,95089	12604,8572	C1
M7	381,8127813	41,76122604	C2
M8	353,1416147	0	C2
M9	4865,757906	5205,718874	C1
M10	4692,557938	5033,498684	C1
M11	5391,109441	5726,829839	C1
M12	4750,339462	5092,154161	C1
M13	7062,327449	7394,937998	C1
M14	12749,23135	13080,52021	C1
M15	347,3557254	675,6663378	C1
M16	161,1862277	514,1984053	C1
M17	707,4779148	1052,845668	C1
M18	7663,010179	7996,526371	C1
M19	2813,689748	3155,93235	C1
M20	21243,15713	21576,38719	C1
M21	10956,5323	11300,64197	C1
M22	187,8323721	520,096145	C1
M23	0	353,1416147	C1
M24	2553,00803	2884,061199	C1
M25	791,0556238	1118,144892	C1
M26	705,0049645	1033,104545	C1
M27	845,7168557	1175,598571	C1
M28	586,80576	911,6139534	C1
M29	555,0900828	878,9960182	C1
M30	957,7577982	1285,686198	C1

Proses iterasi dilakukan dimana *centroid* diperbarui berdasarkan rata-rata posisi dari objek-objek dalam klaster tersebut. Setelah pembaruan *centroid*, objek-objek dievaluasi ulang untuk memastikan mereka berada dalam klaster yang benar. Setelah menghitung rata-rata didapatkan titik *centroid* seperti yang ada pada Tabel 2.

Tabel 2 Titik Centroid 2

	C1	C2
	3772,571429	3333,321429
	222	175

#### 3.4 Mengelompokkan objek untuk menentukan anggota cluster berdasarkan jarak

Proses pengelompokan objek dilakukan dengan menentukan keanggotaan klaster berdasarkan jarak antara setiap objek dan titik pusat klaster (*centroid*). Objek-objek yang memiliki jarak paling dekat dengan *centroid* yang sama akan dikelompokkan ke dalam klaster yang sama. Hasil pengelompokkan iterasi terdapat pada Tabel 3.

Tabel 3 Pengelompokan Hasil Iterasi

C1	1,2,3,4,5,6,9,10,11,12,13,14,15,16,17,18,8,	28
C2	19,20,21,22,23,24,25,26,27,28,29,30,7,8	2

Pengelompokkan hasil iterasi terdapat 28 anggota untuk *C1* dan 8 anggota untuk *C2*. Proses iterasi berlanjut

hingga pengelompokan hasil iterasi mencapai kondisi konvergensi, di mana posisi centroid dan keanggotaan klaster tidak lagi mengalami perubahan signifikan.

3.5 Mengulangi ke tahap 3 hingga nilai centroid yang dihasilkan tetap dan anggota cluster tidak berpindah ke cluster lain.

Iterasi dilanjutkan dari tahap 3 hingga tercapai kondisi stabilitas, yaitu ketika nilai centroid yang dihasilkan pada iterasi berturut-turut tetap sama dan tidak ada perubahan dalam keanggotaan klaster. Ini menunjukkan bahwa algoritma telah mencapai persamaan. Berikut hasil iterasi 6 disajikan pada Tabel 4.

Tabel 4 Hasil Iterasi 6

No	C1	C2	Cluster
M1	10273,81291	865,7242633	C2
M2	1500,810536	7912,026447	C1
M3	4494,343586	4917,263314	C1
M4	7545,464599	1864,745693	C2
M5	10638,09557	1232,705358	C2
M6	1289,254151	10696,68026	C1
M7	11352,7362	1944,618857	C2
M8	11316,93882	1909,610039	C2
M9	6123,032726	3296,229133	C2
M10	6299,702752	3124,464194	C2
M11	5590,534524	3818,689369	C2
M12	6246,888107	3183,791215	C2
M13	3926,385776	5490,095491	C1
M14	1813,29321	11177,23832	C1
M15	10642,15151	1233,984805	C2
M16	10831,66775	1424,260685	C2
M17	10275,85823	868,8593097	C2
M18	3322,837419	6090,564547	C1
M19	8174,772842	1249,950599	C2
M20	10264,40657	19670,67555	C1
M21	1030,669572	9393,967559	C1
M22	10800,32586	1391,390851	C2
M23	10981,63625	1572,509618	C2
M24	8433,600122	984,543803	C2
M25	10198,82909	794,5832241	C2
M26	10283,84691	877,9535865	C2
M27	10141,37571	735,7000068	C2
M28	10405,33518	999,7732243	C2
M29	10437,96094	1032,393578	C2
M30	10031,34287	629,1712803	C2

Iterasi dihentikan ketika tidak ada perubahan dalam pembagian objek antara klaster dan centroid stabil pada posisi yang sama. Ini menandakan bahwa solusi pengelompokan telah dicapai dan tidak memerlukan iterasi lebih lanjut. Setelah di analisis, maka proses algoritma K-Means Clustering dapat disimpulkan pada Tabel 5.

Tabel 5 Hasil Analisis K-Means Clustering

Cluster	Jumlah Anggota	Nama Anggota	centroid Akhir
C1	8	Masako Ayam Renceng, Royco Ayam Renceng, Tepung Adabi Renceng, Kacang Kedelai, Kacang Tanah, Mie Lidi Cap KA, Gula Pasir Cap PTP, Gula Pasir Cap RMI	2,3,6,13, 14,18,20, 21
C1	22	Ajinamot 1000, MSG Cap 3 Tebu, Tepung Bumbu Sajiku Renceng, Tepung Racik Ayam,	1,4,5,7,8, 9,10,11,1 2,15,16,1 7,19,22,2

Tepung Serbaguna,	3,24,25,2
Tepung Terigu Segitiga Karung 25 KG, Tepung Tapioka Serasi Jaya 25 Kg, Tepung Lencana Merah 25 Kg, Tepung Sagu Jewara 25 Kg, Biji Pala, Lada Hitam, Cengek, Mie Lidi Cap Duku, Karet Sedang, Karet Gelang Besar, Saos Cap Kraton, Saos Sasa 1 Kg, Kecap Bango Sachet, Kecap ABC Sachet, Kantongan 24 Warna Cap Matahari, Kantongan Ukuran 15 Sparta, Krupuk Udang Larisa	6,27,28,2 9,30

Disimpulkan bahwa dari 30 data barang yang diberikan menghasilkan dua kelompok (*cluster*) dengan ketentuan *cluster* pertama (Barang Laris) menghasilkan 8 anggota dan *cluster* kedua (Barang tidak laris) menghasilkan 22 anggota. Serta dapat dinyatakan bahwa *barang* dengan *cluster* (Barang tidak laris ) merupakan *cluster* yang memiliki anggota terbanyak dalam *clustering* berdasarkan stok barang dan jumlah terjual.

#### 4. Kesimpulan

Penelitian ini menunjukkan bahwa penerapan algoritma *K-Means Clustering* dalam mengelompokkan barang di UD. Martua dapat mengetahui mana barang yang laris dan mana barang yang tidak laris agar objek dapat mengetahui barang yang perlu di tambah stoknya dan barang yang harus di kurangi stoknya. Data terdiri dari 30 data barang terdapat 8 sebagai anggota C1 dan merupakan barang laris dan 22 sebagai anggota C2 dan merupakan barang tidak laris. Kesimpulan yang dapat diperoleh dari penelitian ini metode *K-Means* dapat mengelompokkan barang pada UD.Martua. Penelitian ini menunjukkan bahwa implementasi metode *K-Means* dengan dukungan aplikasi RapidMiner efektif dalam mengelompokkan data barang pada UD.Martua.

#### Daftar Rujukan

- [1] Geni et al., "Analisa Algoritma K-Means Untuk Menentukan Strategi Marketing," vol. 8, pp. 211–220, 2024, doi: 10.30865/mib.v8i1.7085.
- [2] S. Arimbi Syakila Simangunsong1, Ibnu Qital Pasaribu2, Khoirun Nisa3, "An infinite-rank Lie algebra associated to  $SL(2, R)$  and  $SL(2, R)/U(1)$ ," *J. Math. Phys.*, vol. 65, no. 8, pp. 139–148, 2024, doi: 10.1063/5.0223755.
- [3] M. Febri, M. Sari, Y. Suherman, E. N. Azandra, B. Harto, and R. A. Steffanie, "Sistem Informasi Penyewaan Perlengkapan Pelaminan Pada Paris Pesta Di Kota Padang Berbasis Web," vol. 5, no. 3, pp. 895–900, 2024.
- [4] F. N. Ahadiyah, "Perkembangan Teknologi Infomasi Terhadap Peningkatan Bisnis Online," *INTERDISIPLIN J. Qual. Quant. Res.*, vol. 1, no. 1, pp. 41–49, 2024.
- [5] 2Popon Handayan Yuliana Anggraeni, "Jurnal Informatika dan Teknologi Informasi Penerapan Metode K-Means Untuk Menentukan Penjualan Tiket Renang Jurnal Informatika dan Teknologi Informasi," vol. 2, no. 1, pp. 173–181, 2023, doi: 10.56854/jt.v2i1.167.
- [6] E. Kosasih, T. Oktavia, O. Jaringan, H. M. Model, D. Mining,

- and P. K. Sinyal, "Integrasi Business Analytics dalam Manajemen Performa Jaringan Seluler 4G 1,2," pp. 537–548.
- [7] S. S. Lubis and B. Hendrik, "Implementasi Data Mining Pengelompokan Data Penjualan Berdasarkan Pembelian Dengan Menggunakan Algoritma K-Means Pada UD . Martua," vol. 1, no. 3, pp. 36–41, 2023.
- [8] D. F. Astri, M. Informatika, J. Barat, B. Langsung, A. Kecil, and A. S. Dasar, "CLUSTERING PENDUDUK MISKIN MENGGUNAKAN ALGORITMA K-MEANS PADA WILAYAH JAWA BARAT," vol. 8, no. 2, pp. 1548–1554, 2024.
- [9] M. Y. Nurzaman and B. N. Sari, "Implementasi K-Means Clustering Dalam Pengelompokan Banyaknya Jumlah Petani Berdasarkan Kecamatan Di Provinsi Jawa Barat," vol. 10, no. 3, 2023, doi: <https://doi.org/10.35957/jatisi.v10i3.4896>.
- [10] N. Kusmawanti *et al.*, "PENGELOMPOKKAN LAGU TRADISIONAL DI MEDIA SOSIAL TIKTOK MENGGUNAKAN ALGORITMA K-MEANS," vol. 8, no. 3, pp. 2962–2967, 2024.
- [11] S. L. 1; R. N. 2; L. A. U. 3; M. Zidane, "Implementasi Data Mining Dalam Menentukan Penjualan Alat Perabot," vol. 2, no. 1, pp. 271–278, 2024.
- [12] A. B. 2 Maria Adelina Bui 1, "IMPLEMENTASI METODE ALGORITMA K-MEANS CLUSTERING UNTUK MENGELOMPOKKAN TRANSAKSI PENJUALAN BARANG DI TOKO ARINO," vol. 8, no. 2, pp. 1451–1456, 2024.
- [13] S. Farikah *et al.*, "PENGELOMPOKAN KATALOG PADA TOKO MAHABBATAIN STORE MENGGUNAKAN ALGORITMA K-MEANS," vol. 8, no. 3, pp. 2890–2894, 2024.
- [14] A. Anjani, A. L. Hananto, and S. Shofiah, "KLASTERISASI DATA PENJUALAN TERLARIS PRODUK KOSMETIK YOU MENGGUNAKAN ALGORITMA K-MEANS," vol. 9, no. 1, pp. 17–25, 2024.
- [15] M. Fajar, N. Rahaningsih, R. D. Dana, T. Informatika, and K. Akuntansi, "ANALISIS POLA PENJUALAN OBAT DI APOTEK AN-NAAFI MENGGUNAKAN METODE K-MEANS CLUSTERING," vol. 8, no. 1, pp. 486–492, 2024.
- [16] I. Indra, N. Nur, M. Iqram, and N. Inayah, "Perbandingan K-Means dan Hierarchical Clustering dalam Pengelompokan Daerah Beresiko Stunting," *INOVTEK Polbeng - Seri Inform.*, vol. 8, no. 2, p. 356, 2023, doi: 10.35314/isi.v8i2.3612.