

# Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Analisa Penjualan Parfume

**Restu Riadi, Mesran**

Fakultas Ilmu Komputer dan Teknologi Informasi, Prodi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia

Email: <sup>1,\*</sup>riadirestu783@gmail.com, <sup>2</sup>mesran.skom.mkom@gmail.com

**Abstrak**—Toko IM Parfume Rantauprapat adalah sebuah toko yang menawarkan berbagai jenis aroma parfum dengan merek IM. Meskipun menyediakan berbagai pilihan, tidak semua jenis parfum terjual dengan cepat, ada yang laris dan ada juga yang kurang diminati. Data penjualan, pembelian, dan pengeluaran pada toko tersebut tidak teratur, sehingga data tersebut hanya berfungsi sebagai arsip tanpa dapat dimanfaatkan untuk pengembangan strategi pemasaran. Seharusnya data yang telah terkumpul dapat digunakan sebagai sistem pengambilan keputusan untuk mengatasi permasalahan bisnis. Untuk mencapai hal ini, penulis merancang sebuah aplikasi data mining dalam penelitian ini dengan harapan dapat memberikan hasil yang efektif dan maksimal dalam menganalisis penjualan parfum di toko IM Parfume Rantauprapat. Penerapan Data Mining dengan Algoritma K-Means terbukti memberikan analisis terbaik dan menjadi solusi dalam mengembangkan bisnis parfum. Melalui pemodelan clustering dengan algoritma K-Means dan dengan membagi jumlah cluster menjadi 3, rapidminer berhasil membentuk tiga cluster, dimana cluster 1 terdiri dari 9 produk, cluster 2 memiliki 3 produk, dan cluster 3 memiliki 13 produk dari total 25 item produk yang diamati.

**Kata Kunci:** Toko IM Parfume, Penjualan, Data Mining, Algoritma, K-Means

**Abstract**—IM Perfume Rantauprapat store is a shop that offers various types of perfume scents under the IM brand. Even though it provides a wide range of choices, not all types of perfume sell quickly, some are in demand and some are less desirable. Data on sales, purchases and expenses at the store is irregular, so that the data only functions as an archive without being used for developing marketing strategies. The data that has been collected should be used as a decision-making system to solve business problems. To achieve this, the authors designed a data mining application in this study with the hope of providing maximum and effective results in analyzing perfume sales at the IM Perfume Rantauprapat store. The application of Data Mining with the K-Means Algorithm is proven to provide the best analysis and be a solution in developing the perfume business. Through clustering modeling with the K-Means algorithm and by dividing the number of clusters into 3, rapidminer succeeded in forming three clusters, where cluster 1 consisted of 9 products, cluster 2 had 3 products, and cluster 3 had 13 products out of a total of 25 product items observed.

**Keywords:** IM Perfume Store, Sales, Data Mining, Algorithm, K-Means

## 1. PENDAHULUAN

Perkembangan bisnis parfume yang semakin maju dan banyaknya peminat parfume, hal ini menjadikan persaingan dalam dunia bisnis menjadi semakin ketat. Untuk tetap eksis dalam persaingan ini, para pengusaha dituntut untuk mengembangkan strategi bisnis yang dapat meningkatkan penjualan dan pemasaran. Salah satu pendekatan yang dapat digunakan adalah memanfaatkan data penjualan untuk mencari pola dan informasi yang berguna dalam mengoptimalkan kinerja perusahaan[1]. Hal ini tidak dimanfaatkan oleh salah satu toko parfume yang peneliti temukan yaitu Toko IM Parfume Rantauprapat.

Toko IM Parfume Rantauprapat adalah sebuah tempat penjualan berbagai jenis parfum dengan merek IM. Meskipun menyediakan beragam aroma, tidak semua produknya laku di pasaran, ada yang terjual dengan sangat baik dan ada pula yang kurang diminati. Sayangnya, pengelolaan data penjualan, pembelian, dan pengeluaran di toko tersebut tidak teratur, sehingga data tersebut hanya berfungsi sebagai arsip tanpa memberikan manfaat untuk pengembangan strategi pemasaran. Sebenarnya, data yang ada dapat digunakan sebagai alat pengambilan keputusan untuk mengatasi berbagai tantangan bisnis serta mendukung pengembangan teknologi melalui teknik data mining, yaitu proses ekstraksi informasi penting dari data yang ada.

Data mining adalah suatu proses yang melibatkan penggunaan teknik pengenalan pola, metode statistik, dan matematika untuk menyelidiki dan menganalisis sejumlah besar data yang tersimpan dalam suatu repositori. Tujuan dari data mining adalah untuk mengidentifikasi korelasi, pola, dan tren baru yang memiliki makna atau arti penting[2]–[5] Pemanfaatan data mining memndukung setiap pebisnis membuat keputusan dengan cepat dan akurat. Melihat tren pembelian konsumen untuk menentukan pola penjualan, dianalisis dan diproses dengan benar dapat membantu menentukan produk mana yang paling laris, paling laris atau paling lambat, dan melakukan manajemen persediaan dapat digunakan sebagai masukan bagi bisnis saat merumuskan strategi pemasarannya. Anda dapat menggunakan hasil output data mining untuk membuat keputusan di masa mendatang[6].

K-means clustering merupakan sebuah teknik clustering data yang tidak bersifat hirarkis, yang bertujuan untuk mengelompokkan data ke dalam satu atau lebih kelompok (cluster). Data yang memiliki karakteristik serupa akan dikelompokkan menjadi satu cluster, sedangkan data dengan karakteristik yang berbeda akan ditempatkan dalam cluster lainnya. Hasilnya, data dalam satu cluster akan memiliki kesamaan yang tinggi dan perubahan yang minimal[7]. Algoritma K-means digunakan untuk mengidentifikasi minat pelanggan terhadap produk yang dijual, sehingga perusahaan membutuhkan sistem yang dapat membantu pengambilan keputusan secara cepat dan akurat. Untuk mengatasi permasalahan tersebut, analisis dilakukan dengan menerapkan clustering menggunakan algoritma K-Means. Dengan memanfaatkan sistem data mining dalam pengelompokan jenis aroma parfum, diharapkan dapat memberikan solusi terbaik bagi Toko IM Parfume Rantauprapat.

Beberapa penelitian sebelumnya telah dilakukan untuk menerapkan data mining dengan menggunakan algoritma K-Means Clustering dalam berbagai bidang bisnis. Sebagai contoh, penelitian oleh Rendy dan Leonard pada tahun 2019 membahas penerapan metode K-Means dalam perusahaan asuransi, dimana hasilnya menunjukkan bahwa algoritma K-Means dapat membantu dalam menilai kualitas agen asuransi[8]. Penelitian lain oleh Muhammad dan Prihandoko pada tahun 2017, menggunakan data mining untuk menganalisis data bencana milik BNPB dengan algoritma K-Means dan linear regression. Hasil penelitian ini menyarankan bahwa informasi tentang jumlah kejadian bencana alam sangat penting dalam upaya penanggulangan bencana, dan teknik data mining dapat digunakan untuk mengelompokkan dan memprediksi data bencana di masa depan[9]. Selanjutnya, penelitian oleh Gustientiedina dkk pada tahun 2018 fokus pada penerapan algoritma K-Means untuk mengelompokkan data obat-obatan di RSUD Pekanbaru. Hasilnya menunjukkan bahwa pengelompokan ini dapat menjadi referensi penting dalam pengambilan keputusan terkait perencanaan dan pengendalian pasokan medis di rumah sakit tersebut[10]. Penelitian lain oleh Randi Rian Putra dan Cendra Wadisman pada tahun 2018 berkaitan dengan implementasi data mining untuk pemilihan pelanggan potensial dengan algoritma K-Means. Penelitian ini menggunakan observasi dan wawancara dengan pemilik perusahaan yang bersangkutan untuk menganalisis data yang diperlukan dalam proses penentuan pelanggan potensial[11]. Selain itu, Fitri Yunita melakukan penelitian pada tahun 2019 mengenai penerapan data mining dengan algoritma K-Means Clustering dalam penerimaan mahasiswa baru di Universitas Islam Indragiri. Penelitian ini menggunakan atribut seperti asal sekolah, program studi, dan nilai UAN untuk mengelompokkan data mahasiswa baru menjadi tiga cluster berdasarkan karakteristik mereka[12].

Adanya sistem data mining dalam pengelompokan jenis aroma perfume, dengan harapan dapat memberikan solusi terbaik bagi pihak Toko IM Perfume Rantauprapat . Berdasarkan pembahasan diatas dan melihat keputusan serta metode yang digunakan, maka penulis mengangkat permasalahan ini kedalam skripsi dengan judul “Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Analisa Penjualan Perfume (Study Kasus : Toko IM Perfume Rantauprapat)”.

## 2. METODE PENELITIAN

### 2.1 Data Mining

Data mining adalah suatu metode pengolahan data yang bertujuan untuk menemukan pola-pola tersembunyi dalam data. Hasil dari pengolahan data dengan metode data mining ini dapat digunakan untuk pengambilan keputusan di masa mendatang. Istilah data mining juga sering disebut sebagai knowledge discovery atau pengenalan pola[13]–[17]. Konsep data mining dapat dipandang sebagai node-node yang masih memerlukan pemahaman yang mendalam, mengacu pada data mentah yang sedang dihadapi[18],[19].

### 2.2 Penjualan

Penjualan merupakan bagian dari kegiatan pemasaran yang bertujuan untuk melakukan pertukaran produk dari produsen kepada konsumen. Meskipun saat ini istilah penjualan sering dianggap memiliki arti yang sama dengan pemasaran, namun pemasaran tetap memiliki cakupan yang lebih luas daripada penjualan. Kegiatan penjualan merupakan salah satu aspek dari proses pemasaran yang bertujuan untuk mengadakan pertukaran produk dari produsen kepada konsumen. Pemasaran, di sisi lain, mencakup lebih dari sekadar penjualan karena prosesnya dimulai sebelum barang diproduksi dan dijual. Pemasaran berperan sejak tahap perencanaan dan pengembangan produk hingga ke tahap meyakinkan konsumen untuk menggunakan produk tersebut.

### 2.3 Algoritma K-Means

K-Means adalah sebuah metode pengelompokan berbasis clustering. Teknik ini bekerja dengan melakukan pengelompokan secara bertahap. Algoritma K-Means merupakan suatu algoritma clustering yang bersifat iteratif. Prosesnya dimulai dengan menentukan secara acak atau teracak nilai  $K$ , yang mewakili jumlah cluster yang akan terbentuk. Selanjutnya, algoritma ini secara acak menetapkan nilai awal untuk setiap pusat cluster, yang sering disebut sebagai centroid, mean, atau "mean". Metode ini bertujuan untuk membagi data menjadi beberapa kelompok atau cluster[20]. Berikut adalah langkah-langkah untuk melakukan pengelompokan menggunakan algoritma K-Means.

1. Pilih jumlah cluster  $k$ .
2. Inisialisasi pusat  $k$  cluster dapat dilakukan dengan berbagai cara. Tetapi sebagian besar waktu ini terjadi secara acak. Pusat cluster secara acak diberi nilai awal.
3. Memetakan semua data/objek ke cluster berikut: Kedekatan dua benda ditentukan berdasarkan jarak antara kedua benda tersebut. Untuk menghitung jarak dari semua data ke pusat setiap cluster, kita dapat menggunakan teori jarak Euclidean, yang dirumuskan sebagai berikut:

$$d_{ij} = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + (x_{ki} - x_{kj})^2} \quad (1)$$

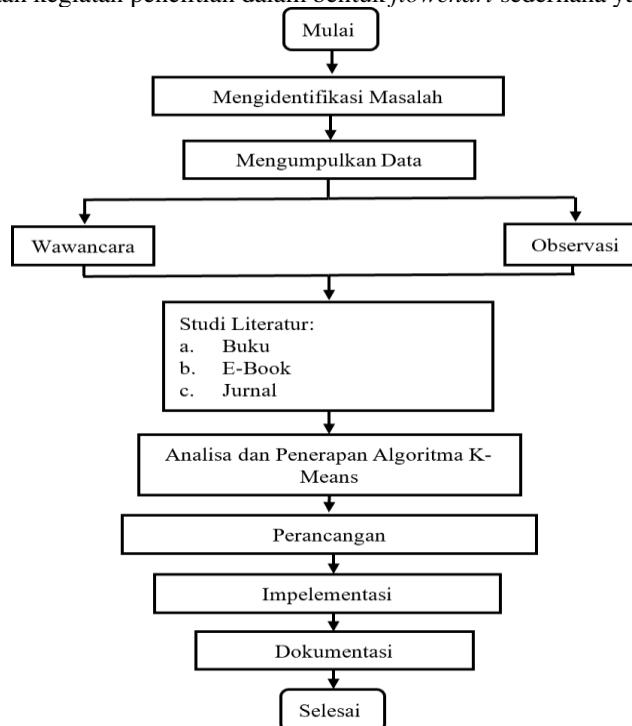
4. Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang. Pusat cluster adalah rata-rata dari semua data/objek dalam cluster tertentu. Jika dikehendaki bisa juga menggunakan median dari cluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.

$$R_k = \frac{1}{N_k} (x_{1k} + x_{2k} + \dots + x_{nk}) \quad (2)$$

5. Tugaskan lagi tiap objek memakai pusat cluster yang baru. Jika pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah nomor 3.

#### 2.4 Tahapan Penelitian

Berikut ini tahapan pelaksanaan kegiatan penelitian dalam bentuk *flowchart* sederhana yaitu seperti gambar di bawah ini:



**Gambar 1.** *Flowchart* Tahapan Penelitian

Berikut ini tahapan pelaksanaan kegiatan penelitian dalam pada penelitian ini yaitu:

1. Mengidentifikasi Masalah  
Tahapan ini adalah cara penulis untuk memperkirakan dan menjabarkan permasalahan yang terjadi dalam menganalisa penjualan parfume di Toko IM Parfume Rantauprapat.
2. Mengumpulkan Data  
Pada tahap ini penulis melakukan pengumpulan data apa saja yang dibutuhkan dalam penelitian dan membangun sistem, yaitu dengan cara melakukan observasi dan wawancara.
3. Studi Literatur  
Tahapan ini merupakan dimana penulis melakukan pemahaman objek yang diteliti atau membaca beberapa sumber yang bisa dijadikan sebagai panduan referensi seperti buku, e-book, jurnal dan sumber lainnya.
4. Analisa dan Penerapan Algoritma K-Means.  
Tahapan Analisa ini digunakan untuk mengetahui apa yang menjadi sumber masalah dalam menganalisa penjualan parfume, sehingga analisa yang dihasilkan nantinya dapat mengatasi permasalahan yang ada. Kemudian penulis melakukan penerapan algoritma K-Means. Penerapan algoritma ditujukan untuk menghitung nilai-nilai dari setiap jenis aroma parfume yang ada.
5. Perancangan  
Pada tahapan ini Penulis memberikan gambaran mengenai sebuah sistem untuk analisa penjualan parfume yang akan diusulkan. Tahapan perancangan sistem ini merupakan data yang telah dianalisis dengan tujuan agar mudah dan dimengerti oleh pemakai (user). Dimana data akan diproses ke dalam perangkat lunak sistem (source code) untuk mengetahui apakah sistem dapat berjalan sesuai kebutuhan.
6. Implementasi  
Tahap ini merupakan tahap pengimplementasian dari hasil proses analisa dan perancangan sistem, yang mana data akan diproses kedalam perangkat lunak sistem, untuk mengetahui apakah sistem berjalan dengan lancar dan berjalan sesuai yang di inginkan. Untuk penerapan data mining dalam analisa penjualan parfume membutuhkan perangkat pendukung yaitu berupa *software* dan *hardware*. Setelah selesai membuat perancangan system. Langkah selanjutnya adalah melakukan pengujian. Pengujian dilakukan untuk mengetahui apakah sistem berjalan dengan baik atau tidak dan harus melakukan perbaikan sistem.
7. Dokumentasi  
Tahapan ini merupakan tahapan akhir hasil evaluasi seluruh tahapan kegiatan penelitian yang dilakukan dan dibuat dalam bentuk laporan. Tujuan dari dibuatnya dokumentasi untuk memberikan masukan dalam pengembangan sistem terhadap analisa penjualan parfume kedepannya.

### 3. HASIL DAN PEMBAHASAN

Pada tahap ini akan dibahas proses analisa penjualan parfume di Toko IM Parfume Rantauprapat, untuk melakukan analisa penjualan parfume di Toko IM Parfume Rantauprapat dengan menerapkan Algoritma K-Means menggunakan sampel data untuk menyelesaikan masalah ini. Berikut ini sample data yang peneliti dapatkan dari Toko IM Parfum Rantauprapat.

**Tabel 1.** Daftar Nama-nama Aroma Parfume

No	Nama Aroma Parfume	Stok Awal	Stok Terjual	Stok Akhir
1	Taylor Swift	45	30	15
2	Bacarat	60	50	10
3	Britney Fantasi	40	15	25
4	Coffe	75	60	15
5	Jlo Platinum	60	45	15
6	Justin Bieber	45	33	12
7	Lovely	50	38	12
8	Maherzain	37	20	17
9	Silver	30	18	12
10	Rafi Ahmad	50	33	17
11	212 Man	30	20	10
12	212 VIP Man	50	24	26
13	2n Up Men	30	18	12
14	Ac scandal	45	34	11
15	Agnes monica	55	39	16
16	Algner Blue	20	15	5
17	Alphald	25	17	8
18	Amor	30	20	10
19	Anasui Dolly	45	28	17
20	212 VIP Man	36	24	12
21	D26 Imperative	35	30	5
22	D26 Man	35	30	5
23	Diamor	30	17	13
24	Dilan	30	20	10
25	Dior Sauvage	50	35	15

#### 3.1 Proses Penerapan Algoritma K-Means

Untuk dapat menganalisis data penjualan parfum menjadi beberapa cluster, lakukan langkah-langkah berikut.

1. Tentukan berapa banyak cluster yang Anda butuhkan. Untuk penelitian ini, kami mengelompokkan data yang ada menjadi tiga cluster.
2. Tentukan titik awal untuk setiap cluster. Dalam penelitian ini, titik tengah pertama ditentukan dengan mengambil semua data berdasarkan jumlah persediaan awal, persediaan terjual, dan persediaan akhir.

**Tabel 2.** Centroid Awal

Keterangan	Stock Awal	Stock Terjual	Stock Akhir
C1	75	60	15
C2	50	33	17
C3	30	20	10

3. Tahap berikutnya adalah menghitung jarak pusat cluster terhadap data penjualan parfume. Data yang diambil sebagai cluster awal adalah data ke 4, 10 dan 18 sehingga didapatkan data centeroid awal yaitu:

Jarak data penjualan ke-1 ke pusat cluster yaitu:

$$d(1,1) = \sqrt{(45 - 75)^2 + (30 - 60)^2 + (15 - 15)^2} = 42,426$$

$$d(1,2) = \sqrt{(45 - 50)^2 + (30 - 33)^2 + (15 - 17)^2} = 6,164$$

$$d(1,3) = \sqrt{(45 - 30)^2 + (30 - 20)^2 + (15 - 10)^2} = 18,708$$

Dari hasil perhitungan diperoleh jarak data penjualan ke-1 dengan pusat cluster pertama (C1) adalah 42.426 dengan pusat cluster kedua (C2) adalah 6.164 dan pusat cluster ketiga (C3) adalah 18.708. Oleh karena itu jarak terpendek dengan pusat cluster adalah 6.164.

Jarak data penjualan ke-2 ke pusat cluster:

$$d(2,1) = \sqrt{(60 - 75)^2 + (50 - 60)^2 + (10 - 10)^2} = 18,708$$

$$d(2,2) = \sqrt{(60 - 50)^2 + (50 - 33)^2 + (10 - 17)^2} = 20,928$$

$$d(3,3) = \sqrt{(60 - 30)^2 + (50 - 20)^2 + (10 - 10)^2} = 42,426$$

Dari hasil perhitungan di atas diperoleh bahwa jarak data penjualan ke-10 dengan pusat cluster pertama (C1) adalah 18,708 dengan pusat cluster kedua (C2) adalah 20,928 dan pusat cluster ketiga (C3) adalah 42,426. Oleh karena itu jarak terpendek dengan pusat cluster adalah 18,708.

Jarak data-data penjualan ke-3 ke pusat cluster yaitu:

$$d(2,1) = \sqrt{(30-75)^2 + (20-60)^2 + (10-10)^2} = 57,879$$

$$d(2,2) = \sqrt{(30-50)^2 + [(20-33)]^2 + [(10-17)]^2} = 22,090$$

$$d(3,3) = \sqrt{(30-20)^2 + [(40-20)]^2 + [(15-10)]^2} = 18,708$$

Dari perhitungan di atas dapat diketahui bahwa jarak antara data penjualan ke-10 dengan pusat cluster pertama (C1) adalah 57,879 dengan pusat cluster kedua (C2) sebesar 22,09072 dan pusat cluster ketiga (C3) adalah 18,708. Dengan demikian, jarak terpendek ke pusat cluster adalah 18,708.

Demikian pula, data penjualan lainnya dihitung dengan cara yang sama. Dalam penelitian ini menghitung menggunakan Microsoft Excel. Hasil perhitungan dan penentuan jarak terpendek dapat dilihat pada tabel 3 di bawah ini.

**Tabel 3.** Data Cluster 1 Iterasi 1

No	Nama Aroma Parfume	Jarak Cluster			C1	C2	C3	Jarak Terpendek
		SA	ST	Sak				
1	Taylor Swift	45	30	15	6,164	42,426	18,708	6,164
2	Bacarat	60	50	10	18,708	20,928	42,426	18,708
3	Britney Fantasi	40	15	25	57,879	22,090	18,708	18,708
4	Coffe	75	60	15	0	36,851	60,415	0
5	Jlo Platinum	60	45	15	21,213	15,748	39,370	15,748
6	Justin Bieber	45	33	12	7,071	40,472	19,949	7,071
7	Lovely	50	38	12	7,071	33,436	26,981	7,071
20	212 VIP Man	36	24	12	53,160	17,378	7,483	7,483
21	D26 Imperative	35	30	5	50,990	19,442	12,247	12,247
22	D26 Man	35	30	5	50,990	19,442	12,247	12,247
23	Diamor	30	17	13	62,273	25,922	4,242	4,242
24	Dilan	30	20	10	60,415	24,859	0	0
25	Dior Sauvage	50	35	15	2,828	35,355	25,495	2,828

Setelah semua data dihitung dan jarak terpendek telah ditentukan, langkah berikutnya adalah mengelompokkan data ke dalam masing-masing cluster. Jarak terpendek diberi nilai numerik bukan nol 1. Pengelompokan data untuk perhitungan di atas ditunjukkan pada tabel berikut.

**Tabel 4.** Kelompok Data ke-1

No	C1	C2	C3
1	1	0	0
2	0	1	0
3	0	0	1
4	0	1	0
5	0	1	0
6	1	0	0
7	1	0	0
20	0	0	1
21	0	0	1
22	0	0	1
23	0	0	1
24	0	0	1
25	1	0	0

Setelah mendapatkan anggota dari setiap klaster, langkah selanjutnya adalah menghitung pusat klaster baru dengan membagi jumlah total penjualan dalam klaster yang sama dengan total klaster, seperti terlihat pada Tabel 4.4.

**Tabel 5.** Pusat Cluster Baru

Pusat Cluster	Stock Awal	Stock Terjual	Stock Akhir
C1	0,138579903	0,055186687	0,129918659
C2	0,219557517	0,219557517	0,516874987
C3	0,61734272	0,262234076	0,61734272

Perhitungan untuk mencari centroid baru dilakukan dengan menggunakan rumus yang sama pada rumus iterasi 1. Iterasi dengan cara yang sama dilakukan sampai tidak ada perubahan data cluster seperti pada Tabel 4.5. Pada penelitian

ini dilakukan pengulangan sebanyak 2 kali. Pada iterasi ke-2 didapatkan cluster yang tidak berubah dari iterasi pertama dimana nilai C1 = 9, C2 = 3 dan C3 = 13 sama dengan nilai iterasi ke-2. Hasilnya seperti terlihat pada tabel di bawah ini.

**Tabel 6.** Data Cluster 1 terasi ke 2

No	Nama Aroma Parfume	Jarak Cluster			C1	C2	C3	Jarak Terpendek
		SA	ST	Sak				
1	Taylor Swift	45	30	15	18,050	5,2326	22,022	5,232
2	Bacarat	60	50	10	28,795	23,201	43,139	23,201
3	Britney Fantasi	40	15	25	28,736	19,442	17,738	17,738
4	Coffe	75	60	15	46,005	35,437	64,335	35,437
5	Jlo Platinum	60	45	15	24,938	18,461	43,151	18,461
6	Justin Bieber	45	33	12	7,052	18,501	21,872	7,052
7	Lovely	50	38	12	16,652	13,338	28,938	13,338
20	212 VIP Man	36	24	12	26,670	9,1766	8,427	8,427
21	D26 Imperative	35	30	5	28,474	12,064	10,724	10,724
22	D26 Man	35	30	5	28,474	12,064	10,724	10,724
23	Diamor	30	17	13	34,026	16,844	1,2688	1,268
24	Dilan	30	20	10	33,291	15,495	2,6476	2,647
25	Dior Sauvage	50	35	15	11,205	15,192	29,051	11,205

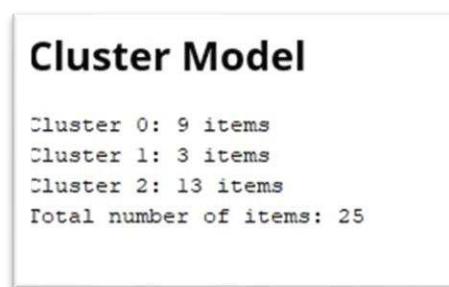
Data tersebut kemudian dikelompokkan ke dalam cluster masing-masing dan jarak terpendek diberi angka bukan nol 1. Pengelompokan data untuk perhitungan data di atas ditunjukkan pada tabel berikut.

**Tabel 7.** Kelompok Data ke-2

No	C1	C2	C3
1	1	0	0
2	0	1	0
3	0	0	1
4	0	1	0
5	0	1	0
6	1	0	0
7	1	0	0
20	0	0	1
21	0	0	1
22	0	0	1
23	0	0	1
24	0	0	1
25	1	0	0

### 3.2 Pengujian

Tahap ini dimulai dengan analisis sistem dan perancangan sistem yang akan dilakukan. Langkah selanjutnya adalah menyiapkan sistem dengan tujuan menghasilkan analisis penjualan yang sesuai berdasarkan hasil perhitungan yang telah dilakukan dengan menggunakan aplikasi quickminer. Pemodelan clustering dengan algoritma k-means dengan jumlah cluster sebanyak 3, maka dengan menggunakan rapidminer cluster yang terbentuk juga sebanyak 3, dengan cluster 1 sebanyak 9 produk, cluster 2 sebanyak 3 produk dan cluster 3 sebanyak 13 produk dari jumlah sampel data produk 25 item.

**Gambar 2.** Hasil Cluster K-Means

Clustering dengan algoritma kmeans dilakukan secara manual pada data transaksi penjualan parfum di toko IM Parfume Rantauprapat yang dilakukan 2 iterasi dari 3 cluster yaitu C1 dan C2, C3 menjadi produk terlaris, C2 menjadi produk terlaris dan C3 menjadi produk terlaris, produk yang kurang laku, kelompok ini terlihat pada nilai centroid iterasi

ke-1 dan ke-2, dimana nilai centroid didapat dari harga nilai rata-rata penjualan untuk cluster mulai gudang 1, stok terjual dan stok akhir di bawah nilai centroid gugus 2.

**Tabel 8.** Nilai Centroid ke 2

Keterangan	Stock Awal	Stock Terjual	Stock Akhir
C1	62,5	27	11,75
C2	44,18	26	11,72
C3	30,2	28,9	12,4

Proses clustering untuk sampel data sebanyak 25 produk yang dianalisis dengan algoritma K-means memperoleh hasil C1 sebanyak 9 item C2 sebanyak 3 item dan C3 sebanyak 13 item yang diuraikan sebagai berikut:

1. Cluster C1 sebanyak 9 produk dengan nomor produk: 1, 6, 7, 10, 12, 14, 15, 19, 25
2. Cluster C2 sebanyak 3 produk dengan nomor produk: 2, 4, 5
3. Cluster C3 sebanyak 3 produk dengan nomor produk: 3, 8, 9, 11, 13, 16-18, 20-24

Clustering yang dilakukan dengan implementasi quickminer didapatkan hasil seperti pada gambar di atas dengan jumlah cluster 3 yaitu cluster 1, cluster 2 dan cluster 3 dan disebutkan juga cluster 1 memiliki 9 item, cluster 2 memiliki 3 item dan cluster 3 memiliki 13 item. Nilai centroid yang dihasilkan menggunakan quickminer sama dengan yang diperoleh dengan tangan. Hal ini meningkatkan akurasi dalam mengklasifikasikan best seller, best seller dan bukan best seller di toko IM Parfume Rantauprapat.

Dengan mengagregasi produk best seller, best seller dan non best seller di toko IM Parfume Rantauprapat dalam penjualan parfum dengan berbagai macam aroma diharapkan dapat mengurangi kesulitan yang dihadapi terutama kehabisan stok. Produk terlaris atau terlaris dan mengumpulkan produk yang tidak laku. Jumlah produk di klaster 1 dan klaster 2 dapat bertambah, sedangkan untuk produk di klaster 3, jumlah persediaan dapat berkurang. Selain itu, manajemen toko IM Parfume Rantauprapat juga perlu menerapkan berbagai strategi untuk meningkatkan penjualan karena produk yang kurang laku, lebih dari produk yang paling laris, sehingga harus ada strategi khusus untuk mengatasi masalah tersebut. menjual segala macam produk di cluster 3.

## 4. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan, maka diperoleh beberapa kesimpulan yaitu data mining menggunakan algoritma K-means Clustering yang diimplementasikan dalam penjualan parfum, dapat digunakan oleh IM Parfume Stores sebagai metode untuk menentukan arah kebijakan pemasarannya. Hasil analisis penjualan parfum yang diperoleh dari sistem ini menunjukkan bahwa data paling banyak diminta oleh konsumen dan disambut baik oleh remaja maupun dewasa. Dari hasil klasterisasi diperoleh tiga klaster, klaster huruf pertama (C0) merupakan kelompok parfum terlaris, klaster kedua (C1) merupakan kelompok parfum terlaris dan klaster kedua huruf tiga (C2) adalah kelompok wewangian yang tidak terjual. Hasil klaster yang diperoleh pada algoritma K-Means menunjukkan bahwa kelompok C0 produk terlaris meliputi 9 produk, kelompok C1 produk terlaris meliputi 3 produk, dan kelompok C2 mencakup 13 produk. Produk tidak laris manis.

## REFERENCES

- [1] S. Nurajizah and A. Salbinda, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Fashion Hijab Banten," *J. Tek. Komput. AMIK BSI*, vol. 7, no. 2, pp. 158–163, 2021, doi: 10.31294/jtk.v4i2.
- [2] M. S. Lestari Sinaga, Abdullah Ahmad, "PENERAPAN DATA MINING PADA JUMLAH PELANGGAN PERUSAHAAN AIR BERSIH MENURUT PROVINSI MENGGUNAKAN METODE K-MEANS CLUSTERING," *RESISTOR*, vol. 2, no. 2, pp. 119–125, 2019.
- [3] S. A. Windania Purba, Willy Siawin, Hardih, Marlince NK Nababan, N P Dharshinni, "IMPLEMENTASI DATA MINING UNTUK PENGELOMPOKKAN DAN PREDIKSI KARYAWAN YANG BERPOTENSI PHK DENGAN," *JUSIKOM PRIMA*, vol. 2, no. 1, 2018.
- [4] P. S. Hasugian, "PENERAPAN DATA MINING UNTUK KLASIFIKASI PRODUK MENGGUNAKAN ALGORITMA K-MEANS," *Mantik Penusa*, vol. 2, no. 2, pp. 191–198, 2018.
- [5] Z. A. Zulfa Nabila, Auliya Rahman Isnain, Permata, "ANALISIS DATA MINING UNTUK CLUSTERING KASUS COVID-19," *Teknol. dan Sist. Inf.*, vol. 2, no. 2, pp. 100–108, 2021.
- [6] A. Darmawan, N. Kustian, W. Rahayu, T. Tabebuya, and K. Pengunjung, "IMPLEMENTASI DATA MINING MENGGUNAKAN MODEL SVM," *J. String*, vol. 2, no. 3, pp. 299–307, 2018.
- [7] A. Bastian *et al.*, "PENERAPAN ALGORITMA K-MEANS CLUSTERING ANALYSIS PADA PENYAKIT MENULAR MANUSIA (STUDI KASUS KABUPATEN MAJALENGKA)," *J. Sist. Inf.*, no. 1, pp. 26–32, 2018.
- [8] R. Novianto and L. Goeirmanto, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering untuk Menganalisa Bisnis Perusahaan Asuransi," *J. Tek. Inform. dan Sist. Inf.*, vol. 6, no. 1, pp. 85–95, 2019.
- [9] P. Muhamad Iqbal Ramadhan, "PENERAPAN DATA MINING UNTUK ANALISIS DATA BENCANA MILIK BNPB MENGGUNAKAN ALGORITMA," *J. Inform. dan Komput.*, vol. 22, no. 1, pp. 57–65, 2017.
- [10] Y. D. Gustientiedina, M. Hasmil Adiya, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru," *J. Nas. Teknol. dan Sist. Inf.*, vol. 01, pp. 17–24, 2019.
- [11] C. W. Randi Rian Putra, "IMPLEMENTASI DATA MINING PEMILIHAN PELANGGAN POTENSIAL MENGGUNAKAN ALGORITMA K-MEANS," *Intecoms*, vol. 1, no. 1, pp. 72–77, 2018.

- [12] F. Yunita, "PENERAPAN DATA MINING MENGGUNKAN ALGORITMA K-MEANS CLUSTRING PADA PENERIMAAN MAHASISWA BARU (STUDI KASUS : UNIVERSITAS ISLAM INDRAGIRI)," *J. Sist.*, vol. 7, no. 3, pp. 238–249, 2018.
- [13] N. Rofiqo *et al.*, "PENERAPAN CLUSTERING PADA PENDUDUK YANG MEMPUNYAI KELUHAN KESEHATAN DENGAN DATAMINING K-MEANS," *KOMIK*, vol. 2, no. 1, pp. 216–223, 2018.
- [14] S. N. Arofah, F. Marisa, P. Studi, T. Informatika, and U. W. Malang, "Penerapan Data Mining untuk Mengetahui Minat Siswa pada Pelajaran Matematika menggunakan Metode K-Means Clustering," *J. Inf. Technol. Comput. Sci.*, vol. 3, no. 2, pp. 85–90, 2018, doi: 10.31328/jointecs.v3i2.787.
- [15] M. R. Alhapizi, M. Nasir, and I. Effendy, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru Universitas Bina Darma Palembang," *J. Softw. Eng. Ampera*, vol. 1, no. 1, pp. 1–14, 2020.
- [16] E. W. Tuti Hartati, Odi Nurdian, "Analisis dan penerapan algoritma k-means dalam strategi promosi kampus akademi maritim suaka bahari," *J. Sains Teknol. Transp. Marit.*, vol. 3, no. 1, pp. 1–7, 2021.
- [17] R. W. Sari and D. Hartama, "Data Mining: Algoritma K-Means Pada Pengelompokan Wisata Asing ke Indonesia Menurut Provinsi," in *Seminar Nasional Sains dan Teknologi Informasi (SENSASI)*, 2018, vol. 1, no. 1.
- [18] R. F. P. D. Ikhsan Romli, "PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS UNTUK KLASIFIKASI PENYAKIT ISPA," *IJUBI*, vol. 4, no. 1, pp. 10–15, 2021.
- [19] A. P. W. Kiki Fatmawati, "PENERAPAN RAPIDMINER DENGAN K-MEANS CLUSTER PADA DAERAH TERJANGKIT DEMAM BERDARAH DENGUE ( DBD ) BERDASARKAN PROVINSI," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 3, no. 2, pp. 173–178, 2018.
- [20] M. A. W. K. MURTI, "UNTUK MENGELOMPOKAN POTENSI PRODUKSI BUAH – BUAHAN DI PROVINSI DAERAH ISTIMEWA YOGYAKARTA," UNIVERSITAS SANATA DHARMA YOGYAKARTA, Yogyakarta, 2017.