

## **Analisis Segmentasi Produk Parfum Pria Menggunakan Metode K-Means Clustering Untuk Mendukung Strategi Bisnis**

<sup>1</sup>Nabila Najma Aulia\*, <sup>2</sup>Elsha Sabilla.

<sup>12</sup>Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pamulang, Tangerang Selatan, Indonesia.

<sup>1</sup>[nanacjmaaa@gmail.com](mailto:nanacjmaaa@gmail.com), <sup>2</sup>[sabelsabilla@gmail.com](mailto:sabelsabilla@gmail.com).

### **Abstract**

*The development of digital technology has driven an increase in sales transactions, generating a large amount of data. This data contains a variety of information that can be used to aid business decision-making, particularly in understanding the characteristics of the products being sold. However, much of this data is often underutilized, necessitating analytical methods that can transform it into more meaningful information. This study aims to group men's perfume products based on price characteristics and sales volume using the K-Means Clustering method. The dataset used was obtained from the Kaggle platform and underwent data cleaning, attribute selection, determining the number of groups using the Elbow Method, and evaluating the clustering results using the Silhouette Score. The results showed that the most appropriate number of groups was three. The first group consisted of products with relatively high prices but low sales; the second group consisted of products with moderate prices and fairly stable sales; and the third group consisted of products with lower prices but the highest sales. Evaluation using the Silhouette Score yielded a value of 0.6065, indicating good cluster quality. The results of this grouping can be used as a basis for designing marketing strategies, managing inventory, and determining promotional focus according to the characteristics of each product group.*

**Keywords:** Data Mining, Clustering, K-Means, Product Segmentation, Marketing Strategy, Men's Fragrance.

### **Abstrak**

Perkembangan teknologi digital mendorong bertambahnya jumlah transaksi penjualan yang menghasilkan banyak data. Data ini mengandung berbagai informasi yang bisa digunakan untuk membantu pengambilan keputusan bisnis, terutama untuk memahami karakteristik produk yang dijual. Namun, banyak data yang ada sering kali tidak dimanfaatkan dengan baik sehingga dibutuhkan metode analisis yang bisa mengubah data tersebut menjadi informasi yang lebih berarti. Penelitian ini bertujuan untuk mengelompokkan produk parfum pria berdasarkan karakteristik harga dan jumlah penjualan menggunakan metode K-Means Clustering. Dataset yang digunakan diperoleh dari platform Kaggle dan melalui proses pembersihan data, pemilihan atribut, penentuan jumlah kelompok menggunakan Metode Elbow, serta evaluasi hasil pengelompokan menggunakan Skor Silhouette. Hasil penelitian menunjukkan bahwa jumlah kelompok yang paling tepat adalah tiga kelompok. Kelompok pertama terdiri dari produk dengan harga cukup tinggi tetapi memiliki penjualan yang rendah, kelompok kedua memiliki produk dengan harga sedang dan tingkat penjualan yang cukup stabil, sedangkan kelompok ketiga diisi oleh produk dengan harga lebih rendah tetapi penjualannya paling tinggi. Evaluasi menggunakan Skor Silhouette menghasilkan nilai sebesar 0,6065 yang menunjukkan bahwa kualitas kelompok yang terbentuk termasuk baik. Hasil pengelompokan ini bisa digunakan sebagai dasar dalam merancang strategi pemasaran, mengelola persediaan, serta menentukan fokus promosi sesuai dengan karakteristik dari masing-masing kelompok produk.

**Kata Kunci:** Data Mining, Clustering, K-Means, Segmentasi Produk, Strategi Pemasaran, Parfum Pria.

### **A. PENDAHULUAN**

Perkembangan teknologi digital telah mengubah banyak bidang bisnis, termasuk industri parfum. Saat ini, penjualan tidak hanya terjadi secara langsung, tetapi juga melalui banyak platform pasar online yang menghasilkan banyak data setiap hari. Data tersebut mencakup informasi tentang harga produk, angka penjualan, ketersediaan stok, dan fitur produk lainnya yang dapat berguna sebagai sumber informasi untuk membantu pengambilan keputusan bisnis. Namun, banyaknya data yang tersedia seringkali tidak diproses dengan cara terbaik, yang berarti informasi

penting yang terkandung di dalamnya tidak dimanfaatkan secara maksimal.

Dalam bisnis parfum, setiap produk memiliki ciri khasnya sendiri. Beberapa produk memiliki harga tinggi tetapi terjual relatif sedikit unit, sementara produk lain lebih terjangkau dan terjual jauh lebih banyak. Perbedaan karakteristik ini dapat memengaruhi strategi pemasaran, manajemen stok, dan identifikasi pasar sasaran. Jika proses identifikasi karakteristik produk hanya dilakukan secara manual, akan memakan waktu lebih lama, dan hasilnya

mungkin tidak sepenuhnya menunjukkan pola data secara keseluruhan.

Data penjualan yang tersimpan dalam jumlah besar sesungguhnya bisa dimanfaatkan untuk mendapatkan wawasan yang bermanfaat bagi perusahaan. Namun, tanpa analisis yang tepat, data tersebut hanya akan menjadi sekumpulan angka yang sulit diterapkan dalam proses pengambilan keputusan. Oleh sebab itu, diperlukan strategi yang dapat mengolah data menjadi informasi yang lebih mudah dimengerti sehingga perusahaan dapat mengidentifikasi pola penjualan dan karakteristik produknya.

Salah satu metode yang dapat diaplikasikan adalah teknik pemodelan kluster, yaitu proses pengelompokan data berdasarkan tingkat kesamaan karakteristik yang ada. Dengan metode ini, produk yang memiliki karakteristik serupa akan dikelompokkan dalam satu kluster, sementara produk dengan karakteristik yang berbeda akan ditempatkan dalam kluster lainnya. Hasil dari pengelompokan ini dapat memberikan wawasan tentang pola penjualan produk yang sebelumnya tidak nampak secara langsung.

Dalam penelitian ini, algoritma K-Means Clustering digunakan untuk mengelompokkan produk parfum pria berdasarkan atribut harga dan jumlah penjualannya. Pemilihan algoritma K-Means dilandaskan pada kemampuannya untuk membentuk kelompok data dengan efisien dan kemudahan dalam memahami hasil yang diperoleh. Melalui proses pengelompokan ini, diharapkan bisa teridentifikasi kelompok produk yang memiliki harga tinggi namun penjualannya terbatas, produk dengan penjualan yang konstan, serta produk yang memiliki permintaan tinggi di pasar.

Data yang digunakan diperoleh dari dataset publik yang diakses melalui platform Kaggle dan berisi informasi mengenai berbagai produk parfum pria yang dijual online. Hasil segmentasi diharapkan dapat memberikan gambaran karakteristik dari masing-masing kelompok produk sehingga dapat digunakan sebagai pertimbangan dalam merancang strategi pemasaran, pengelolaan stok, dan membuat keputusan bisnis yang berdasarkan data dengan lebih terarah.

## B. LANDASAN TEORI

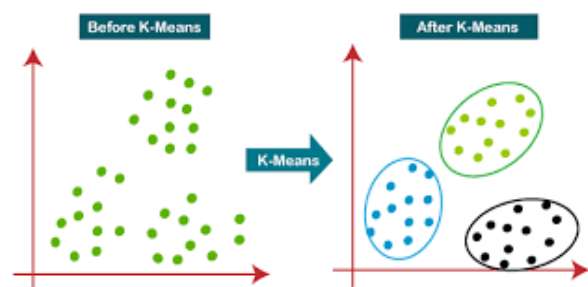
### Data Mining

Menurut (Setyaningtyas, Nugroho, & Arif, 2022), menjelaskan bahwa data mining dimanfaatkan untuk menggali informasi tersembunyi dari sekumpulan data sehingga dapat menghasilkan pengetahuan baru yang berguna dalam proses pengambilan keputusan. Pada penelitian ini, pendekatan tersebut digunakan untuk mengidentifikasi pola penjualan produk parfum berdasarkan karakteristik yang dimiliki setiap produk.

### Clustering

Menurut (Khalish, Piranti, & Martadireja, 2025), clustering adalah metode analisis data yang bertujuan untuk mengorganisir objek berdasarkan kesamaan fitur tertentu. Metode ini memfasilitasi pengelompokan data yang memiliki karakteristik mirip ke dalam satu kategori, sementara data yang berbeda akan ditempatkan dalam kategori lainnya. Melalui proses pengelompokan, pola dan sifat data dapat diidentifikasi dengan lebih jelas, sehingga mendukung analisis dan pengambilan keputusan. Dalam penelitian ini, teknik pengelompokan digunakan untuk mengidentifikasi kategori produk parfum pria berdasarkan harga dan volume penjualannya, agar dapat diperoleh pemahaman mengenai karakteristik masing-masing segmen produk.

### Algoritma K-Means



Menurut (Laksono, Syahidin, & Yunengsih, 2024), K-Means adalah salah satu teknik pengelompokan yang bertujuan untuk mengelompokkan data ke dalam beberapa kategori berdasarkan kesamaan karakteristik yang dimiliki. Proses pengelompokan dilakukan dengan memilih titik tengah kelompok (centroid) sebagai patokan, kemudian setiap data ditempatkan pada kelompok yang paling dekat dengan centroid itu. Prosedur ini dilakukan berulang kali hingga posisi centroid stabil dan susunan cluster yang terbentuk tidak lagi berubah secara signifikan. Biasanya, penghitungan kedekatan antara data dan centroid dilakukan dengan metode Jarak Euclidean yang dapat dilihat dalam persamaan berikut:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Keterangan:

- $d(x, y)$  = jarak antara data dan centroid
- $x_i$  = nilai data pada atribut ke- $i$
- $y_i$  = nilai centroid pada atribut ke- $i$
- $n$  = jumlah atribut yang digunakan

Algoritma K-Means sering diterapkan dalam berbagai studi karena proses pengelompokannya yang cukup sederhana, mudah untuk dilaksanakan, serta mampu mengelola jumlah data yang besar dengan efisiensi waktu yang baik. Dalam penelitian ini, K-Means digunakan untuk melakukan segmentasi produk parfum pria berdasarkan faktor harga, jumlah penjualan, dan harga per penjualan. Hasil dari pengelompokan ini diharapkan dapat memberikan gambaran tentang karakteristik setiap kelompok produk sehingga dapat digunakan sebagai bahan pertimbangan dalam merumuskan strategi pemasaran yang lebih efektif.

### Elbow Method

Menurut (Maori & Evanita, 2023), Elbow Method merupakan metode yang digunakan untuk menentukan jumlah cluster yang optimal pada algoritma K-Means. Metode ini dilakukan dengan menghitung nilai Within Cluster Sum of Squares (WCSS) pada beberapa jumlah cluster yang berbeda, kemudian divisualisasikan dalam bentuk grafik. Nilai WCSS dihitung menggunakan persamaan sebagai berikut:

$$WCSS = \sum_{i=1}^{\{k\}} \sum_{x \in C_i} (x - c_i)^2$$

Keterangan:

- WCSS = total variasi data dalam cluster
- $x$  = data pada cluster
- $c_i$  = centroid cluster ke- $i$

Jumlah klaster yang diambil dalam proses pengelompokan ditetapkan berdasarkan hasil analisis grafik Metode Elbow. Pemilihan ini dilakukan dengan mempertimbangkan pola penurunan nilai WCSS untuk setiap jumlah klaster yang diuji. Titik di mana penurunan nilai WCSS mulai terasa tidak signifikan dipilih sebagai jumlah klaster yang paling mewakili. Dengan cara ini, proses pengelompokan dapat menghasilkan klaster yang lebih efisien dan sesuai dengan karakteristik data yang dianalisis. Dalam penelitian ini, Metode Elbow digunakan sebagai langkah awal untuk menetapkan jumlah klaster yang optimal sebelum menerapkan algoritma K-Means.

### Silhouette Score

Menurut (Sudrajat, Hadiana, & Melina, 2025), Silhouette Score berfungsi sebagai salah satu cara untuk menilai seberapa efektif hasil pengelompokan yang muncul setelah tindakan clustering dilakukan. Metode ini menilai kedudukan masing-masing data dengan membandingkan seberapa dekat data tersebut dengan anggota-anggota dalam cluster yang sama serta seberapa dekat dengan cluster lain di sekitarnya. Dengan pendekatan ini, bisa diukur tingkat pemisahan antar cluster dengan lebih objektif. Perhitungan Silhouette Score dalam studi ini merujuk pada persamaan berikut:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Keterangan:

- $s(i)$  = nilai silhouette data ke- $i$
- $a(i)$  = rata-rata jarak data terhadap anggota dalam cluster yang sama
- $b(i)$  = rata-rata jarak data terhadap cluster terdekat lainnya

Nilai Silhouette Score berkisar antara -1 hingga 1. Semakin tinggi skor yang didapat, maka kualitas pengelompokan yang dihasilkan semakin baik, karena ini menunjukkan bahwa data dalam satu cluster memiliki karakteristik yang cukup mirip dan cukup terpisah dari cluster lainnya. Di sisi lain, nilai yang mendekati 0 menandakan adanya tumpang tindih antara cluster. Dalam studi ini, Silhouette Score

diterapkan untuk menilai hasil segmentasi produk parfum yang diperoleh melalui penggunaan algoritma K-Means, sehingga dapat dianalisis tingkat kualitas cluster yang terbentuk.

### Segmentasi Produk

Segmentasi produk merupakan proses pengelompokan produk berdasarkan karakteristik tertentu untuk memudahkan penyusunan strategi pemasaran yang lebih efektif. Melalui segmentasi, perusahaan dapat memahami perbedaan karakteristik setiap kelompok produk sehingga strategi promosi, pengelolaan stok, dan penentuan target pasar dapat dilakukan secara lebih tepat sasaran. Dalam penelitian ini, segmentasi produk dilakukan menggunakan algoritma K-Means untuk mengelompokkan produk parfum berdasarkan harga dan jumlah penjualan sehingga dapat menghasilkan rekomendasi strategi pemasaran yang sesuai dengan karakteristik masing-masing kelompok produk.

### C. METODE

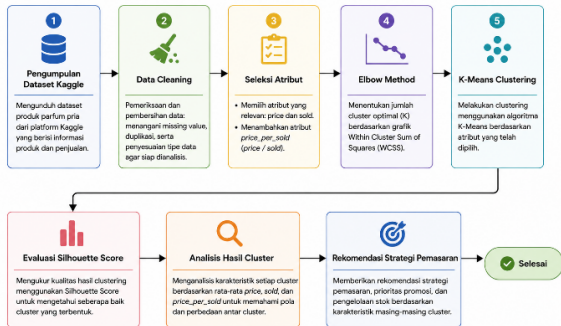
Pada penelitian ini, pendekatan yang digunakan adalah penggalan data dengan teknik pengelompokan untuk menyegmentasi produk parfum berdasarkan harga dan tingkat penjualannya. Dataset yang dipakai diambil dari platform Kaggle, yang mengandung informasi mengenai produk parfum untuk pria yang dipasarkan di berbagai marketplace. Data dalam penelitian ini mencakup atribut seperti merek produk, nama produk, jenis parfum, harga, jumlah penjualan, lokasi penjualan, serta sejumlah atribut tambahan lainnya.

Langkah awal dalam penelitian adalah pengumpulan dataset serta pemeriksaan struktur data yang ada. Setelah itu, tahap pembersihan data dilakukan untuk memastikan data tersebut siap untuk dianalisis. Pada fase ini, dilakukan identifikasi terhadap nilai yang hilang, penyesuaian tipe data, serta pemilihan atribut yang relevan dengan fokus penelitian. Atribut yang digunakan dalam pengelompokan adalah harga dan jumlah penjualan. Di samping itu, atribut harga per penjualan juga ditambahkan sebagai perbandingan antara harga produk dan penjualannya, guna menunjang proses segmentasi.

Setelah data dianggap layak untuk digunakan, langkah selanjutnya adalah menentukan jumlah cluster yang paling optimal dengan metode Elbow. Metode ini menghitung nilai WCSS pada berbagai jumlah cluster yang berbeda. Cluster yang paling sesuai ditentukan dengan melihat titik siku yang menunjukkan bahwa penambahan cluster tambahan tidak menyebabkan penurunan nilai WCSS yang signifikan.

Selanjutnya, proses pengelompokan dilaksanakan dengan menggunakan algoritma K-Means. Algoritma ini berfungsi untuk mengelompokkan data berdasarkan kedekatan karakteristik dengan pusat cluster. Iterasi dilakukan berulang kali sampai posisi pusat cluster stabil dan tidak mengalami pergeseran yang signifikan.

Untuk menilai kualitas dari hasil pengelompokan, evaluasi dilakukan dengan menggunakan Silhouette Score. Nilai yang diperoleh ini digunakan untuk menilai seberapa mirip data dalam satu cluster dibandingkan dengan cluster lainnya. Semakin tinggi nilai Silhouette Score, semakin baik kualitas dari cluster yang terbentuk.



Gambar 1. Alur Penelitian

Alur dari penelitian dapat dilihat pada Gambar 1. Proses penelitian dimulai dengan pengumpulan dataset, diikuti dengan pembersihan data, penentuan jumlah cluster dengan Elbow Method, pengelompokan menggunakan algoritma K-Means, evaluasi dengan Silhouette Score, hingga akhirnya menghasilkan segmentasi produk parfum dan strategi pemasaran yang direkomendasikan berdasarkan karakteristik dari masing-masing cluster.

#### D. HASIL DAN PEMBAHASAN

Penelitian ini memanfaatkan dataset mengenai parfum pria yang diambil dari platform Kaggle. Dataset ini mencakup 1.000 item produk dengan sepuluh atribut yang meliputi *brand*, *title*, *type*, *price*, *priceWithCurrency*, *available*, *availableText*, *sold*, *lastUpdated*, dan *itemLocation*. Namun, tidak semua atribut tersebut digunakan dalam analisis. Kajian ini berfokus pada atribut yang relevan dengan harga dan volume penjualan karena kedua elemen ini dianggap mampu menggambarkan karakteristik produk parfum di pasar.

```

Data Awal:
  brand      title \
0 Dior      Christian Dior Sauvage Men's EDP 3.4 oz Fragra...
1 AS SHOW   A-v-entus Eau de Parfum 3.3 oz 100ML Millesime...
2 Unbranded HOGO BOSS cologne For Men 3.4 oz
3 Giorgio Armani Acqua Di Gio by Giorgio Armani 6.7 Fl oz Eau D...
4 Lattafa   Lattafa Men's Hayaati Al Maleky EDP Spray 3.4 ...

  type      price priceWithCurrency available \
0 Eau de Parfum 84.99 US $84.99/ea 10.0
1 Eau de Parfum 109.99 US $109.99 8.0
2 Eau de Toilette 100.00 US $100.00 10.0
3 Eau de Toilette 44.99 US $44.99/ea 2.0
4 Fragrances 16.91 US $16.91 NaN

  availableText sold lastUpdated \
0 More than 10 available / 116 sold 116.0 May 24, 2024 10:03:04 PDT
1 8 available / 48 sold 48.0 May 23, 2024 23:07:49 PDT
2 More than 10 available / 27 sold 27.0 May 22, 2024 21:55:43 PDT
3 2 available / 159 sold 159.0 May 24, 2024 03:30:43 PDT
4 Limited quantity available / 156 sold 156.0 May 24, 2024 07:56:25 PDT

  itemLocation
0 Allen Park, Michigan, United States
1 Atlanta, Georgia, Canada
...
sold 6
lastUpdated 53
itemLocation 0
dtype: int64
  
```

Gambar 2. Data Awal Produk Parfum

Dari Gambar 2, terlihat bahwa dataset yang dikumpulkan memerlukan pengolahan lebih lanjut sebelum dapat diterapkan di tahap clustering. Data awal

menunjukkan masih adanya beberapa atribut yang tidak digunakan dalam riset serta beberapa nilai yang tidak lengkap pada atribut tertentu. Oleh karena itu, dilakukan pembersihan dan penyesuaian data agar sesuai dengan kebutuhan analisis.

```

Data setelah cleaning:
  brand      title \
0 Dior      Christian Dior Sauvage Men's EDP 3.4 oz Fragra...
1 AS SHOW   A-v-entus Eau de Parfum 3.3 oz 100ML Millesime...
2 Unbranded HOGO BOSS cologne For Men 3.4 oz
3 Giorgio Armani Acqua Di Gio by Giorgio Armani 6.7 Fl oz Eau D...
4 Lattafa   Lattafa Men's Hayaati Al Maleky EDP Spray 3.4 ...

  type      price priceWithCurrency available \
0 Eau de Parfum 84.99 US $84.99/ea 10.0
1 Eau de Parfum 109.99 US $109.99 8.0
2 Eau de Toilette 100.00 US $100.00 10.0
3 Eau de Toilette 44.99 US $44.99/ea 2.0
4 Fragrances 16.91 US $16.91 NaN

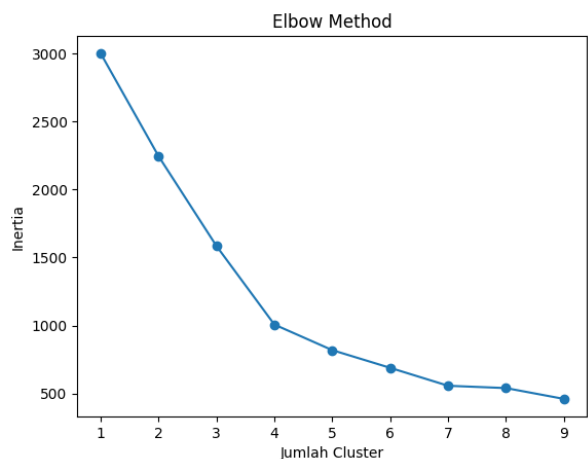
  availableText sold lastUpdated \
0 More than 10 available / 116 sold 116.0 May 24, 2024 10:03:04 PDT
1 8 available / 48 sold 48.0 May 23, 2024 23:07:49 PDT
2 More than 10 available / 27 sold 27.0 May 22, 2024 21:55:43 PDT
3 2 available / 159 sold 159.0 May 24, 2024 03:30:43 PDT
4 Limited quantity available / 156 sold 156.0 May 24, 2024 07:56:25 PDT

  itemLocation price_per_sold cluster
0 Allen Park, Michigan, United States 0.726410 0
...
itemLocation 0
price_per_sold 0
cluster 0
dtype: int64
  
```

Gambar 3. Data Setelah Cleaning

Gambar 3 memperlihatkan hasil data setelah dilakukan proses cleaning dan pengolahan. Pada tahap ini, dilakukan pemilihan atribut yang relevan, penyesuaian format data, serta penambahan atribut *price\_per\_sold* untuk menunjang proses pengelompokan data. Selain itu, hasil cluster yang diperoleh dari penerapan algoritma K-Means juga telah dilibatkan dalam dataset.

Setelah data disiapkan, langkah selanjutnya adalah menentukan jumlah cluster yang paling tepat dengan menggunakan metode Elbow Method. Metode ini dipakai untuk memantau perubahan nilai Within Cluster Sum of Squares (WCSS) pada beberapa jumlah cluster yang diuji.

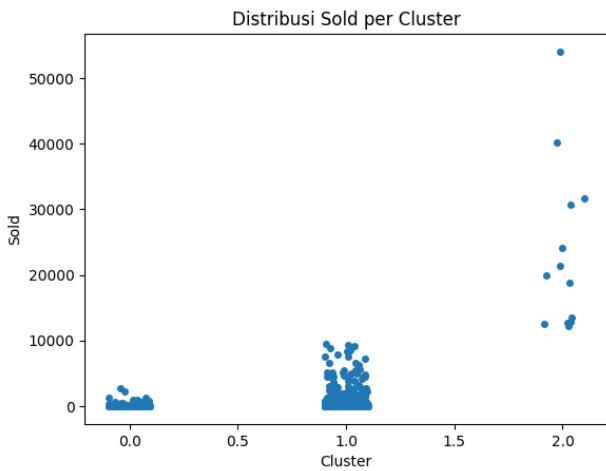


Gambar 4. Hasil Elbow Method

Dalam Gambar 4, terlihat bahwa penurunan nilai WCSS mulai melambat pada nilai K = 3, sehingga titik tersebut dipilih sebagai jumlah cluster yang optimal. Dengan demikian, clustering dilakukan dengan tiga kelompok data.

Proses clustering dilaksanakan dengan menggunakan algoritma K-Means yang memanfaatkan atribut harga

(price), volume penjualan (sold), dan atribut hasil pengolahan data (price\_per\_sold). Hasil pengelompokan menunjukkan bahwa data terdistribusi dalam tiga cluster yang memiliki ciri khas berbeda.



**Gambar 5.** Visualisasi Hasil Clustering

Visualisasi tersebut menunjukkan bahwa tiap cluster menampilkan pola yang berbeda berdasarkan karakteristik harga dan jumlah penjualan produk. Hal ini mengindikasikan bahwa algoritma K-Means dapat mengelompokkan data ke dalam kelompok yang menunjukkan tingkat kesamaan yang cukup baik.

Untuk mengevaluasi kualitas hasil clustering, dilakukan penilaian menggunakan Silhouette Score. Berdasarkan hasil pengujian, didapatkan nilai Silhouette Score sebesar 0,6065. Nilai ini menunjukkan bahwa hasil clustering memiliki kualitas yang memuaskan karena masing-masing cluster mampu membentuk kelompok data yang relatif terpisah satu sama lain.

**Tabel 1.** Karakteristik Rata-Rata Setiap Cluster

Cluster	Rata-rata Harga	Rata-rata Penjualan	Price per Sold
0	110,80	130,30	13,00
1	34,29	527,46	1,70
2	30,25	23.431,85	0,0016

Berdasarkan Tabel 1, Cluster 0 menunjukkan rata-rata harga yang tertinggi meskipun memiliki tingkat penjualan yang relatif rendah. Ciri khas ini menunjukkan bahwa produk dalam cluster ini merupakan kategori produk premium yang biasanya menasar segmen pasar tertentu.

Cluster 1 memiliki harga rata-rata yang lebih rendah daripada Cluster 0, namun mencatatkan tingkat penjualan yang lebih tinggi. Kelompok ini bisa dikategorikan sebagai produk dengan kinerja penjualan yang cukup konsisten dan berpotensi untuk terus berkembang dengan penerapan strategi promosi yang efektif.

Di sisi lain, Cluster 2 memperlihatkan tingkat penjualan tertinggi jika dibandingkan dengan cluster lain yang memiliki harga rata-rata yang cukup rendah. Hal ini

menunjukkan bahwa produk dalam cluster ini sangat diminati oleh konsumen, sehingga memerlukan perhatian ekstra dalam pengelolaan inventaris dan strategi pemasaran.

Berdasarkan karakteristik yang telah dianalisis, strategi pemasaran yang direkomendasikan untuk masing-masing cluster dapat ditemukan dalam Tabel 2.

**Tabel 2.** Rekomendasi Strategi Pemasaran

Cluster	Karakteristik Produk	Strategi Pemasaran
0	Harga tinggi, penjualan rendah	Memperkuat branding dan menonjolkan eksklusivitas produk
1	Harga menengah, penjualan stabil	Promosi berkala dan peningkatan kualitas layanan
2	Harga relatif rendah, penjualan sangat tinggi	Menjaga ketersediaan stok dan memperluas promosi

Temuan penelitian menunjukkan bahwa algoritma K-Means dapat mengkategorikan produk parfum ke dalam berbagai kelompok berdasarkan karakteristik harga dan volume penjualannya. Data ini bisa digunakan sebagai landasan untuk menetapkan strategi pemasaran yang lebih tepat, mulai dari pengelolaan stok, pemilihan target pasar, hingga pengembangan program promosi yang sesuai dengan karakteristik masing-masing kelompok produk.

## E. PENUTUP

### Kesimpulan

Penelitian ini berhasil menerapkan K-Means Clustering untuk mengelompokkan produk parfum berdasarkan atribut harga dan jumlah penjualan. Proses pengelompokan dilakukan pada dataset parfum pria yang diambil dari Kaggle dan menghasilkan tiga kategori produk dengan fitur yang berbeda.

Melalui pengujian menggunakan metode Elbow Method, ditemukan bahwa jumlah cluster yang paling efisien adalah tiga cluster. Penilaian menggunakan Silhouette Score menunjukkan nilai 0,6065 yang mengindikasikan bahwa hasil pengelompokan memiliki kualitas yang baik. Dari hasil clustering, Cluster 0 terdiri dari produk yang memiliki harga tinggi namun tingkat penjualan rendah, Cluster 1 berisi produk dengan harga sedang dan penjualan yang stabil, sementara Cluster 2 mencakup produk dengan harga yang lebih rendah dan penjualan tertinggi.

Hasil segmentasi ini menandakan bahwa masing-masing kelompok produk memiliki karakteristik unik sehingga memerlukan strategi pemasaran yang berbeda. Data dari proses pengelompokan dapat digunakan sebagai landasan untuk menentukan prioritas dalam promosi, manajemen persediaan, serta pengambilan keputusan pemasaran yang lebih tepat sesuai karakteristik dari setiap kelompok produk.

### Saran

Melihat hasil segmentasi yang didapat, produk di Cluster 0 bisa menjadi target untuk evaluasi dalam usaha meningkatkan penjualan melalui penguatan merek dan pendekatan promosi yang lebih efektif. Produk di Cluster 1 harus dipertahankan dengan melakukan promosi yang teratur dan pengelolaan persediaan yang baik. Di sisi lain, produk di Cluster 2 perlu dijadikan prioritas dalam ketersediaan dan distribusi karena memiliki tingkat penjualan tertinggi dibandingkan kelompok lainnya.

## F. DAFTAR PUSTAKA

- Han, J., Kamber, M., & Pei, J. (2022). *Concepts and Techniques*. Cambridge: Morgan Kaufmann.
- Kaggle. (2024, April 23). *Kaggle*. Diambil kembali dari Men's Perfume Dataset: <https://www.kaggle.com/>
- Khalish, F., Piranti, N. M., & Martadireja, O. (2025). Implementasi Data Mining Menggunakan Teknik Clustering dengan Metode K-Means. *JIIP (Jurnal Ilmiah Ilmu Pendidikan)*, 1254–1261.
- Laksono, B., Syahidin, Y., & Yunengsih, Y. (2024). Implementasi Data Mining Klasterisasi Data Pasien Rawat Inap Menggunakan Algoritma K-Means Clustering. *Jurnal Teknologi Sistem Informasi dan Aplikasi*, 12–20.
- Maori, N. A., & Evanita. (2023). Penerapan Elbow Method dalam Menentukan Jumlah Cluster Optimal pada Algoritma K-Means. *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, 85–93.
- Setyaningtyas, S., Nugroho, B. I., & Arif, Z. (2022). Tinjauan Pustaka Sistematis: Penerapan Data Mining Teknik Clustering Algoritma K-Means. *Jurnal Teknoif*, 52–61.
- Sudrajat, R., Hadiana, A. I., & Melina. (2025). Evaluasi Kualitas Klaster Menggunakan K-Means dengan Silhouette Score dan Elbow Method. *Jurnal Algoritma*, 34–42.
- Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2021). *Introduction to Data Mining*. New York: Pearson Education.