

## **KLASTERISASI KOSMETIK SEPHORA MENGGUNAKAN K-MEANS UNTUK REKOMENDASI PEMBELIAN BERDASARKAN PREFERENSI ANGGARAN KONSUMEN**

**Rahmiawati<sup>1)</sup>, Wika Wardani<sup>2)</sup>**

<sup>1, 2)</sup> Universitas Darwan Ali, Jurusan Sistem Informasi Fakultas Ilmu Komputer  
Jl. Batu Berlian No.10 Kalimantan Tengah, Kotawaringin Timur  
e-mail: [rahmiawati2202@email.com](mailto:rahmiawati2202@email.com)<sup>1)</sup>, [wikawardani29@email.com](mailto:wikawardani29@email.com)<sup>2)</sup>

\* Korespondensi: e-mail: [wikawardani29@email.com](mailto:wikawardani29@email.com)

### **ABSTRAK**

Permintaan kosmetik yang terus meningkat menciptakan tantangan bagi konsumen dalam memilih produk yang sesuai dengan preferensi dan anggaran mereka. Untuk mengatasi masalah tersebut, dilakukan analisis cluster dengan mengelompokkan produk kosmetik ke dalam tiga kategori harga, yaitu affordable, mid-range, dan luxury.

Penelitian ini menggunakan data sekunder dari perusahaan Sephora dan menerapkan algoritma K-Means untuk mengklasifikasikan produk. Hasil pemodelan menghasilkan tiga cluster harga yang kohesif dengan masing-masing terdiri dari produk yang sesuai dengan kategori harga yang ditentukan. Visualisasi menggunakan Data View dan plot scatter/bubble membantu memahami pola dan hubungan antar produk berdasarkan harga mereka dalam cluster-cluster yang terbentuk. Hasil analisis klusterisasi ini memberikan rekomendasi pembelian produk kosmetik yang sesuai dengan preferensi anggaran konsumen, meningkatkan pengalaman belanja konsumen, dan kepuasan pelanggan. Dengan memberikan kontribusi yang berharga dalam pemahaman tentang preferensi konsumen terhadap produk kosmetik Sephora berdasarkan anggaran mereka, penelitian ini dapat memberikan panduan rekomendasi pembelian yang relevan dan sesuai dengan kebutuhan konsumen

**Kata Kunci:** Kosmetik, Sephora, Klusterisasi, K-Means

### **ABSTRACT**

*The ever-increasing demand for cosmetics creates challenges for consumers in choosing products that suit their preferences and budget. To solve this problem, a cluster analysis was conducted by grouping cosmetic products into three price categories, namely affordable, mid-range, and luxury.*

*This research uses secondary data from Sephora company and applies K-Means algorithm to classify the products. The modeling results produced three cohesive price clusters with each consisting of products that fit the specified price category. Visualization using Data View and scatter/bubble plots helped understand the patterns and relationships between products based on their prices within the clusters formed. The results of this clustering analysis provide recommendations for purchasing cosmetic products that match consumers' budget preferences, improving consumers' shopping experience and customer satisfaction. By making a valuable contribution to the understanding of consumers' preferences for Sephora cosmetics products based on their budgets, this research can provide relevant and customized purchase recommendation guidelines for consumers.*

**Keywords:** *Cosmetic, Sephora, Clustering, K-Means*

## I. PENDAHULUAN

Konsumen kosmetik telah menjadi bagian yang tidak terpisahkan dari kehidupan sehari-hari banyak orang, dan permintaan terhadap kosmetik terus meningkat seiring dengan peningkatan kesadaran akan penampilan dan perawatan diri.[1] Sebagai salah satu brand terkemuka di industri kosmetik, Sephora berperan sangat penting dalam memenuhi kebutuhan konsumen dengan menawarkan berbagai produk kosmetik berkualitas tinggi dari brand-brand ternama. Mengacu pada “2015 PinkReport: The Shopper”, Sephora berhasil menarik perhatian konsumen milenial sebagai kelompok terbesar yang berbelanja di toko maupun secara *online*. Keberadaan global Sephora juga menjadi keunggulan dalam mencapai konsumen di berbagai wilayah, sementara pengalaman berbelanja yang menarik dan kepercayaan terhadap merek yang dijual di Sephora merupakan faktor penting bagi konsumen.[2]

Namun, pesatnya pertumbuhan industri kosmetik juga menghadirkan tantangan bagi konsumen untuk memilih produk yang sesuai dengan preferensi dan anggaran mereka.[3] Di tengah pasar yang dipenuhi dengan beragam merek dan produk, konsumen sering kesulitan menavigasi kerumitan dalam memilih produk yang sesuai dengan kebutuhan mereka sekaligus tetap dalam batasan anggaran yang ditentukan.[4] Oleh karena itu, penting bagi konsumen untuk memperoleh informasi yang jelas dan akurat tentang produk yang mereka pertimbangkan agar mereka dapat membuat keputusan pembelian yang tepat.[5] Dalam konteks ini, keberadaan solusi efektif yang mendukung konsumen dalam mengambil keputusan pembelian menjadi semakin penting.

Pada penelitian ini mengusulkan penerapan analisis *cluster* menggunakan algoritma *K-Means* sebagai solusi untuk membuat rekomendasi pembelian produk kosmetik Sephora berdasarkan preferensi anggaran konsumen. *Clustering* adalah teknik pengelompokan data yang kuat dan efektif untuk mengidentifikasi pola dan kesamaan produk berdasarkan karakteristik tertentu.[6] Karakteristik harga digunakan sebagai variabel dalam proses *clustering*, dengan tujuan mengelompokkan produk kosmetik Sephora ke dalam kategori harga *affordable* (terjangkau), *mid-range* (menengah) dan *luxury* (mewah).

Dengan menggunakan algoritma *K-Means*, penelitian ini bertujuan untuk membentuk *cluster* yang kohesif dan bermakna sehingga konsumen dapat lebih memahami hubungan antara harga dan produk yang ditawarkan oleh Sephora. Dengan menggunakan hasil *clustering*, diharapkan dapat menemukan pola yang menjelaskan harga produk dan perbedaan antara harga yang berbeda. Tujuan dari penelitian ini tidak hanya untuk mempermudah proses pembelian produk kosmetik bagi konsumen, tetapi juga memberikan kontribusi yang signifikan terhadap kepuasan konsumen dengan memberikan rekomendasi produk berdasarkan preferensi anggaran mereka. Hasil studi ini diharapkan dapat membawa manfaat positif bagi sektor industri kosmetik, terutama bagi Sephora, dengan meningkatkan standar pelayanan dan menciptakan pengalaman berbelanja yang lebih memuaskan bagi pelanggan.

## II. METODE PENELITIAN

Dalam penelitian ini, pendekatan yang digunakan adalah metode klasterisasi dengan menerapkan algoritma *K-Means*. Klasterisasi adalah suatu teknik dalam *machine learning* yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok atau *cluster* yang relevan. Tujuan dari *clustering* adalah untuk menemukan struktur di data yang tidak diketahui sebelumnya.[7] Metode ini sangat berguna dalam analisis data dan pengelompokkan informasi yang kompleks, terutama data yang tidak memiliki label atau informasi sebelumnya tentang data tersebut.

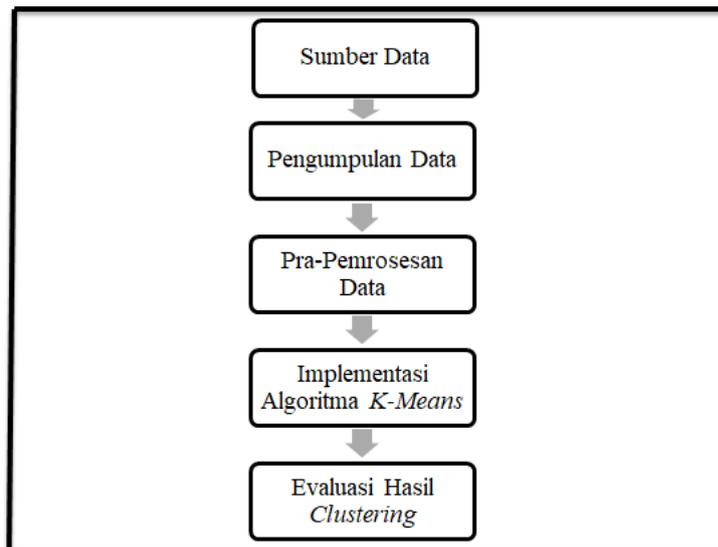
Algoritma *K-Means* adalah salah satu metode klasterisasi yang populer dan terbukti efisien dalam melakukan pengelompokan data. Algoritma *K-Means* merupakan metode klasterisasi yang mengelompokkan data berdasarkan titik pusat cluster (*centroid*) terdekat dengan data.[8] Tujuan utama dari *K-Means* adalah untuk memaksimalkan kemiripan data dalam satu *cluster* dan meminimalkan kemiripan data antara *cluster* yang berbeda. Kemiripan dalam *cluster* diukur menggunakan fungsi jarak, sehingga pemaksimalan kemiripan data diperoleh berdasarkan jarak terpendek antara data dan titik *centroid*.[9] Penghitungan jarak antara setiap data dan klaster dilakukan dengan menggunakan rumus *Euclidean Distance* (D) sebagaimana ditunjukkan dibawah ini.[10]

$$D_{(i,j)} = \sqrt{\sum_{k=1}^n (X_{ik} - C_{jk})^2} \dots\dots\dots(1)$$

Keterangan:

- D = Jarak Klaster
- X<sub>ik</sub> = Nilai Data (i,k)
- C<sub>jk</sub> = Nilai *Centroid* (j,k)
- n = Jumlah Klaster

Metode *clustering* menggunakan algoritma *K-Means* diharapkan dapat memberikan rekomendasi pembelian produk kosmetik yang lebih sesuai dengan preferensi anggaran konsumen dan memudahkan proses pengambilan keputusan bagi para konsumen dalam memilih produk dari yang dijual oleh Perusahaan Sephora. Langkah-langkah yang dilakukan dalam metode penelitian ini adalah sebagai berikut:



Gambar 1. Langkah-langkah Metode *clustering*

#### A. Sumber Data

Dalam penelitian ini, data yang digunakan bersumber dari data sekunder yang diperoleh dari perusahaan Sephora. Data tersebut mencakup berbagai produk kosmetik yang dijual di Sephora, termasuk merek-merek terkenal. Data ini mencakup informasi tentang atribut-atribut produk kosmetik seperti harga, nama produk, jenis produk, rating oleh konsumen, dan atribut lain yang relevan. Data ini akan digunakan sebagai dasar untuk melakukan analisis *clustering* menggunakan algoritma *K-Means*.

#### B. Pengumpulan Data

Data produk kosmetik dari Sephora dikumpulkan melalui proyek Data Camp dengan izin dari pihak Sephora. Terdapat 1.472 *record data* dengan 11 atribut reguler yang telah disusun dan diatur dalam format file *Excel*, sehingga dapat diimpor ke dalam *software* Rapidminer untuk analisis lebih lanjut.

#### C. Pra-Pemrosesan Data

Sebelum data dapat digunakan dalam algoritma *K-Means*, dilakukan pra-pemrosesan data untuk membersihkan dan mempersiapkan data. Pra-pemrosesan meliputi penghapusan data yang hilang atau tidak relevan, penanganan nilai-nilai yang hilang, dan pemformatan data agar sesuai dengan kebutuhan algoritma *K-Means*. [11]

#### D. Implementasi Algoritma *K-Means*

Setelah data dipersiapkan, algoritma *K-Means* diimplementasikan pada *software* Rapidminer. Algoritma *K-Means* akan melakukan analisis *clustering* dari produk kosmetik. *Software* Rapidminer akan mengelompokkan produk ke dalam tiga cluster berbeda berdasarkan kategori harga, yaitu terjangkau, menengah, dan mewah.

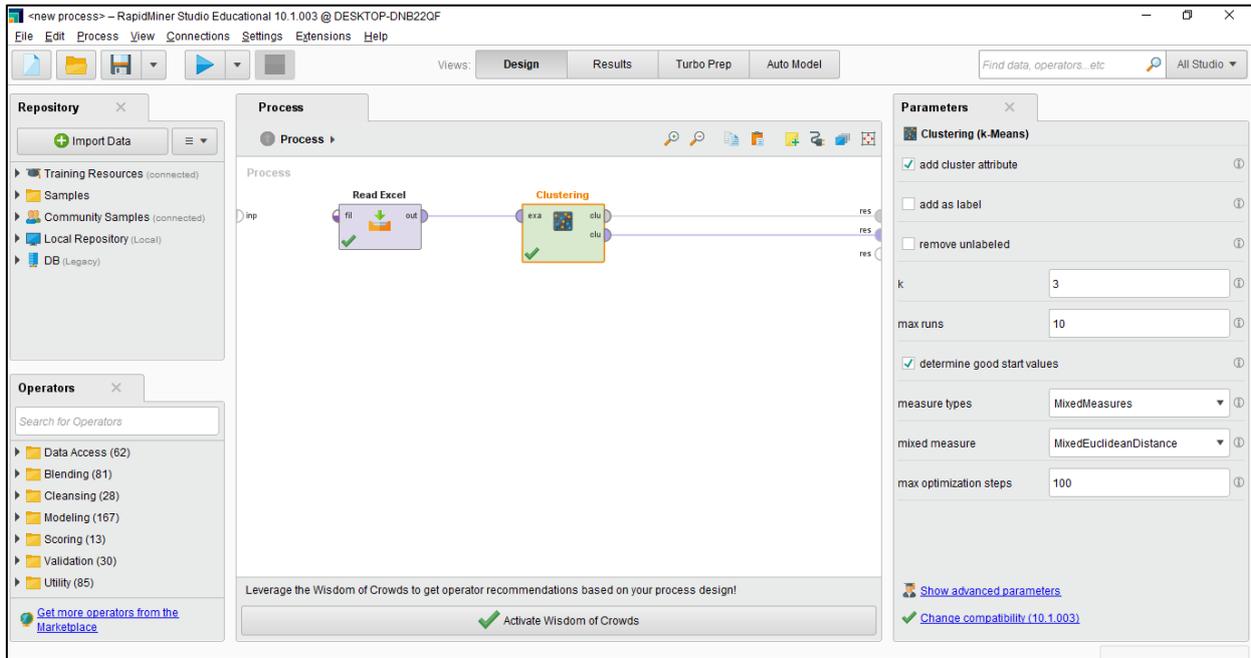
#### E. Evaluasi Hasil *Clustering*

Setelah proses *clustering* selesai, hasilnya dievaluasi untuk memahami pola dan karakteristik dari setiap kelompok produk. Evaluasi ini dapat membantu konsumen dengan jenis kulit tertentu untuk menemukan produk yang sesuai dengan kebutuhan mereka berdasarkan kategori harga.

### III. HASIL DAN PEMBAHASAN

Penelitian ini memiliki tujuan untuk menganalisis produk kosmetik yang ditawarkan oleh Sephora dengan menggunakan algoritma *K-Means*. Tujuan akhir dari penelitian ini adalah memberikan rekomendasi pembelian produk berdasarkan preferensi anggaran yang dimiliki oleh konsumen.

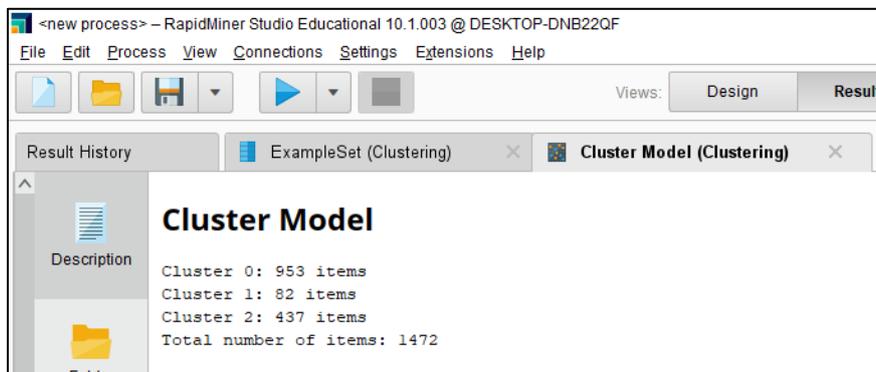
Pada fase awal, langkah yang diambil adalah memasukkan data dan melakukan validasi data untuk memastikan kualitas dataset. Kemudian, proses pemodelan dilakukan menggunakan perangkat lunak RapidMiner Studio Educational 10.1.003 dengan mengimplementasikan algoritma *K-Means*. Pemrosesan data dimulai dengan memasukkan dataset dalam format *Excel* menggunakan komponen "*Read Excel*" pada RapidMiner.



Gambar 2. Proses penginputan dataset kosmetik ke dalam RapidMiner

Setelah itu, data diproses dengan menggunakan metode *K-Means* dengan pengaturan parameter  $k=3$ , karena peneliti mengklasifikasikan produk kosmetik Sephora menjadi tiga kategori harga yaitu *affordable*, *mid-range* dan *luxury*. "*Measure Type = Mixed Measure*" diterapkan karena dataset kosmetik terdiri dari kode dan angka. Dalam pemrosesan ini, *Euclidean Distance* digunakan sebagai standar jarak antar titik *cluster*.

Selanjutnya, data diproses dan menghasilkan tiga model kluster dengan komposisi seperti yang ditunjukkan dalam Gambar 2 di bawah ini:



Gambar 3. Tampilan Cluster Model menggunakan Text View

Dari hasil pemodelan ini, peneliti berhasil mengidentifikasi tiga kluster yang terbentuk dari data produk kosmetik Sephora, dengan pengelompokan berdasarkan kategori harga.

- *Cluster 0* terdiri dari 953 item yang berisi produk dengan harga terjangkau (*affordable*)
- *Cluster 1* terdiri dari 82 item yang menampilkan produk dengan harga mewah (*luxury*)
- *Cluster 2* terdiri dari 437 item yang berisi produk dengan harga menengah (*mid-range*)

Proses analisis klasterisasi ini memberikan pemahaman yang lebih mendalam kepada peneliti tentang pola dan karakteristik dataset kosmetik yang telah berhasil dibagi ke dalam kelompok harga yang berbeda.

TABEL 1. DATA CENTROID

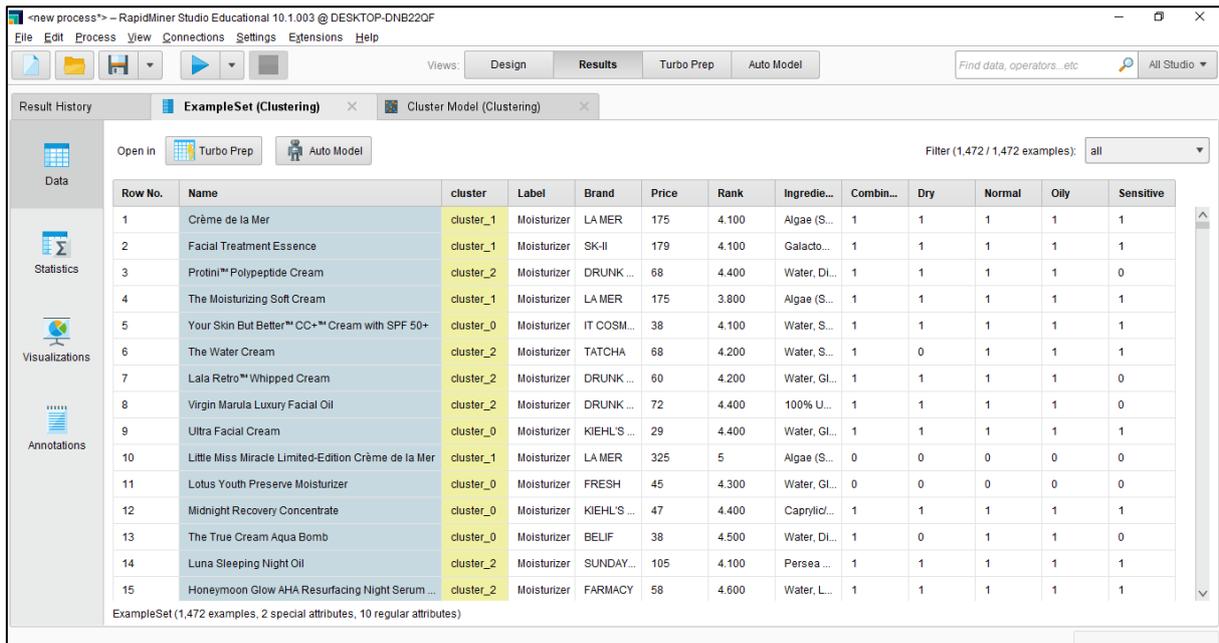
Atribut	Cluster_0	Cluster_1	Cluster_2
Brand	36.803	25.463	35.993
Combination	0.639	0.659	0.693
Dry	0.578	0.683	0.680
Ingredients	660.511	537.793	606.073
Label	3.274	2.878	3.151
Normal	0.622	0.695	0.709
Oily	0.598	0.622	0.625
Price	32.629	198.768	78.778
Rank	4.172	4.082	4.169
Sensitive	0.519	0.598	0.485

Tabel diatas menampilkan *centroid data* yang menggunakan *Centroid Table*. Tabel tersebut menggambarkan atribut-atribut yang terdapat dalam tiga kluster yang berbeda, yaitu *Cluster 0*, *Cluster 1*, dan *Cluster 2*. Setiap kluster memiliki karakteristik atribut yang berbeda-beda.

- Pada kluster pertama (*Cluster 0*) memiliki nilai rata-rata tertinggi pada atribut Brand dan Label
- Kluster kedua (*Cluster 1*) memiliki nilai kombinasi dan harga (*Combination* dan *Price*) yang lebih tinggi dibandingkan kluster lainnya.
- Sementara itu, kluster ketiga (*Cluster 2*) menonjol pada atribut *Ingredients* dan *Price*.

Dalam konteks ini, analisis kluster digunakan untuk mengelompokkan data ke dalam kelompok-kelompok yang memiliki karakteristik serupa. Dengan menggunakan pusat massa (*centroid*) dari setiap kluster, sehingga dapat mengidentifikasi perbedaan signifikan antara atribut-atribut dalam masing-masing kluster.

Pada *ExampleSet*, terdapat beberapa tampilan hasil kluster, salah satunya adalah *Data View*. *Data View* menampilkan hasil kluster data secara komprehensif sesuai dengan data yang telah dimasukkan. Tampilan ini memungkinkan peneliti untuk dengan jelas melihat bagaimana data telah dikelompokkan berdasarkan cluster harga yang dibentuk oleh algoritma *K-Means*. Dengan menggunakan *Data View*, peneliti dapat melakukan pemeriksaan dan analisis pola, karakteristik, serta perbedaan antar *cluster* untuk mendapatkan pemahaman yang lebih dalam tentang kelompok harga produk kosmetik yang terbentuk.



Gambar 4. Tampilan hasil klasterisasi pada Data View

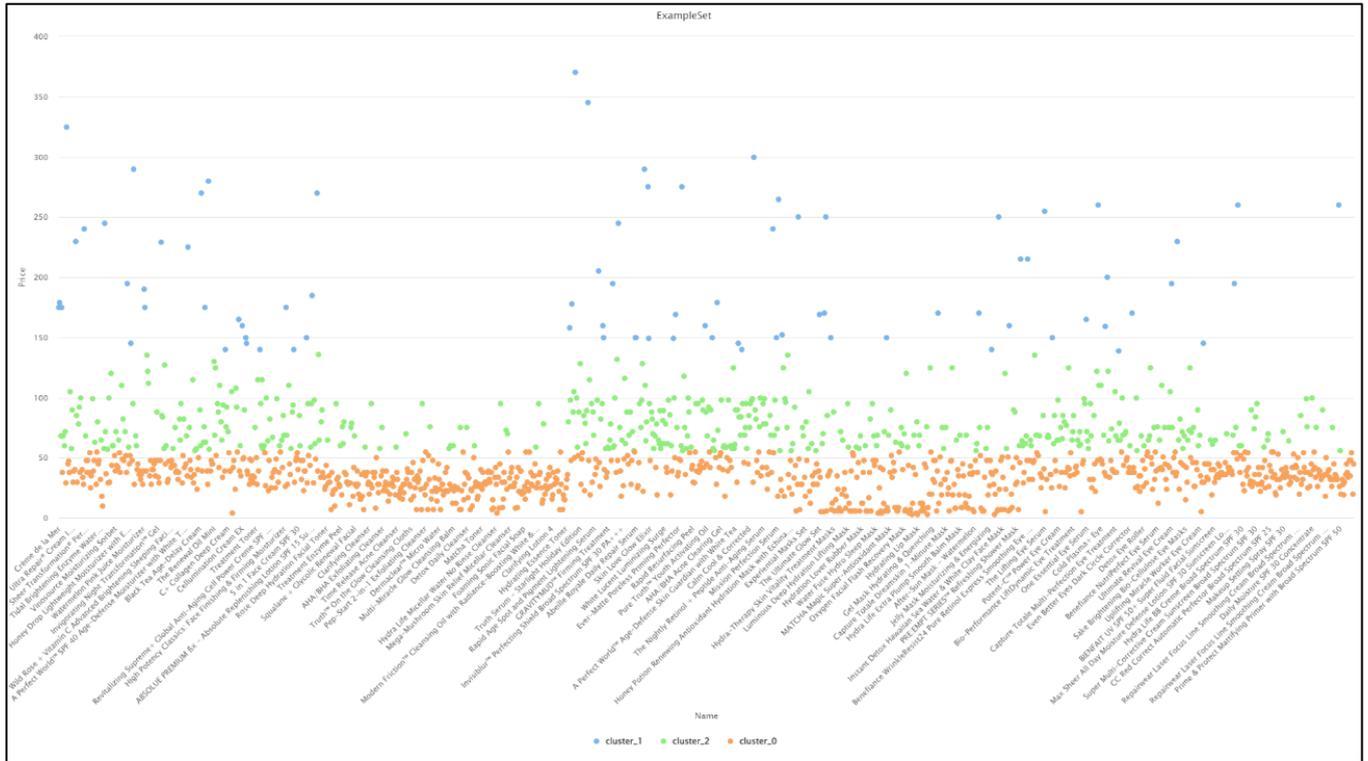
Untuk menganalisis pola klaster, dapat menggunakan *plot scatter* atau *bubble chart*. Dalam plot ini, sumbu x (sumbu horizontal) dapat diisi dengan variabel nama produk, sementara sumbu y (sumbu vertikal) dapat diisi dengan variabel harga produk. Selanjutnya, pada bagian warna plot, dapat dipilih variabel *cluster* untuk mengatur warna titik-titik pada scatter plot agar sesuai dengan hasil klasterisasi data. Contoh visualisasi seperti yang terlihat pada Gambar 4 dapat memberikan pemahaman yang lebih baik tentang bagaimana pengisian variabel plot type dapat mempengaruhi hasil visualisasi data.



Gambar 5. Tampilan plot type

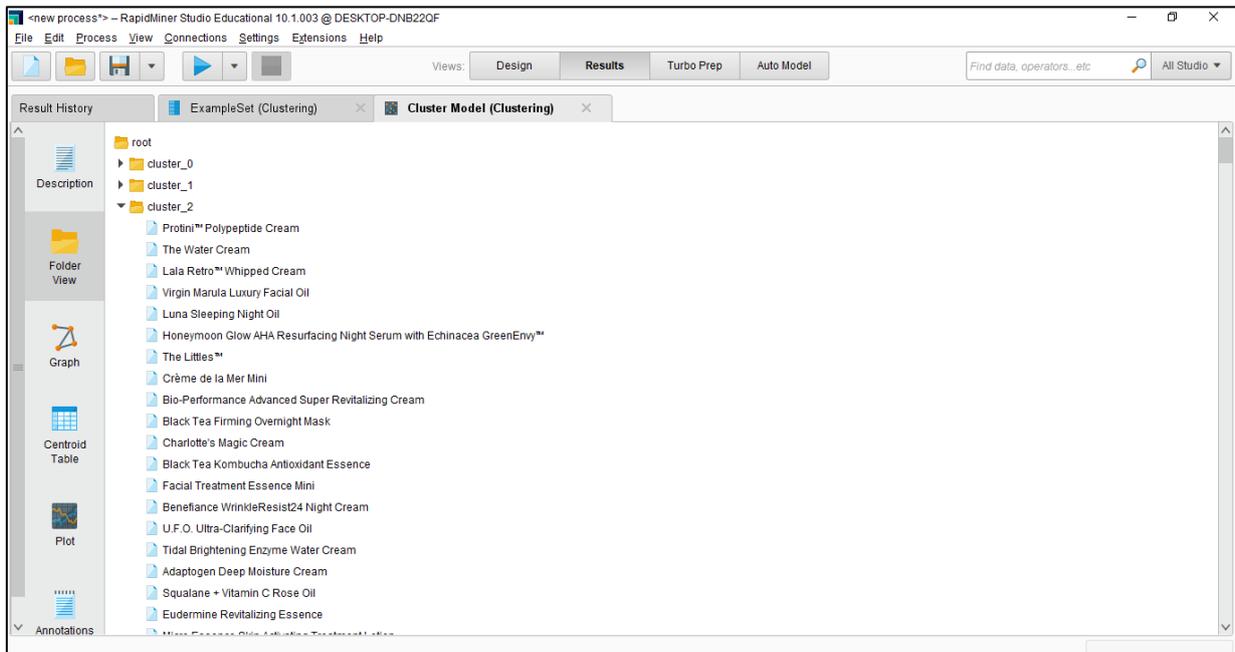
Dengan melakukan visualisasi seperti yang terlihat pada Gambar 5 di bawah ini, memungkinkan untuk dengan mudah mengidentifikasi pola dan hubungan antara produk kosmetik berdasarkan harga dalam kelompok-kelompok (*cluster*) yang telah terbentuk. Melalui penggunaan warna pada *scatter plot*, akan tampak dengan jelas bagaimana produk kosmetik dikelompokkan ke dalam *cluster* harga tertentu. Ini akan sangat membantu dalam eksplorasi dan

pemahaman karakteristik masing-masing *cluster* serta memberikan wawasan berharga dari hasil analisis klusterisasi data tersebut.



Gambar 6. Visualisasi menggunakan Scatter/Bubble berdasarkan nama produk dan harganya

Pada *Cluster Model (Clustering)*, bagian *Folder View* menampilkan seluruh data yang termasuk dalam setiap kelompok *cluster*. Dalam tampilan ini, setiap *cluster* menunjukkan nama produk yang termasuk dalam kelompok tersebut. *Folder View* menjadi tampilan yang sangat berguna untuk memahami kelompok produk kosmetik yang terbentuk berdasarkan pola dan karakteristik harga melalui penerapan algoritma *K-Means*.



Gambar 7. Visualisasi cluster menggunakan Folder View

Dengan menggunakan *Folder View* ini, peneliti dapat dengan mudah mengevaluasi dan menganalisis kelompok produk kosmetik yang sesuai dengan preferensi anggaran konsumen. Visualisasi ini juga memberikan pemahaman yang berharga dalam mengidentifikasi kelompok produk yang serupa dalam suatu *cluster* tertentu. Hasil analisis klasterisasi data kosmetik ini akan menjadi dasar untuk memberikan rekomendasi pembelian produk kosmetik Sephora yang sesuai dengan preferensi anggaran konsumen. *Folder View* menjadi alat yang berguna dalam menggali hasil analisis klasterisasi data produk kosmetik dari Sephora, dengan tujuan untuk meningkatkan pengalaman berbelanja konsumen.

Berikut adalah daftar beberapa produk kosmetik dari Sephora beserta harganya yang disusun berdasarkan hasil analisis klasterisasi pada tampilan *Folder View*:

A. *Daftar Produk dengan Harga Terjangkau (Affordable)*

TABEL 2. PRODUK HARGA TERJANGKAU (*AFFORDABLE*)

Nama Produk	Harga (\$)
Your Skin But Better™ CC+™ Cream with SPF 50+	38
Ultra Facial Cream	29
Lotus Youth Preserve Moisturizer	45
Midnight Recovery Concentrate	47
The True Cream Aqua Bomb	38
Ultra Repair® Cream Intense Hydration	30
Moisture Surge 72-Hour Auto-Replenishing Hydrator	39
Rose Deep Hydration Moisturizer	40
COMPLEXION RESCUE™ Tinted Moisturizer Broad Spectrum SPF 30	30
The True Cream Moisturizing Bomb	38
Virgin Marula Luxury Facial Oil Mini	72
...	...
DERMAPROTECT Daily Defense Broad Spectrum SPF 50+	45

Tabel diatas menampilkan beberapa produk kosmetik dari Sephora dengan rentang harga terjangkau (*affordable*). Produk-produk ini mencakup beragam pilihan seperti *cleanser*, *moisturizer*, *sun protect* dan produk perawatan kulit lainnya. Dari *Your Skin But Better™ CC+™ Cream* dengan *SPF 50+* hingga *DERMAPROTECT Daily Defense Broad Spectrum SPF 50+*, semua produk ini menawarkan kualitas perawatan kulit yang dapat dipercaya tanpa harus menguras kantong. Dengan harga mulai dari \$3 hingga \$55, konsumen dapat memilih produk yang sesuai dengan kebutuhan kulit, memberikan perlindungan dan kelembapan yang dibutuhkan. Dengan berbagai pilihan tersebut, konsumen dapat merawat kulit dengan baik tanpa harus mengorbankan anggaran.

B. *Daftar Produk dengan Harga Mewah (Luxury)*

Berikut ini adalah tabel yang memuat daftar sejumlah produk kosmetik mewah dari Sephora beserta harga-harganya.

TABEL 3. PRODUK HARGA MEWAH (*LUXURY*)

Nama Produk	Harga (\$)
Crème de la Mer	175
Facial Treatment Essence	179
The Moisturizing Soft Cream	175
Little Miss Miracle Limited-Edition Crème de la Mer	325
R.N.A. POWER Face Cream	230
GenOptics Aura Essence Serum	240
The Renewal Oil	245
FUTURE RESPONSE Age Defense Creme	195
Vintage Single Extract Essence	145
Crème Ancienne®	290
The Concentrate	370
...	...
Future Solution LX Total Protective Emulsion Broad Spectrum SPF 20 Sunscreen	260

Produk-produk ini menawarkan berbagai macam perawatan kulit yang eksklusif dan berkualitas tinggi. Di antara produk-produk yang tercantum dalam tabel di atas, beberapa di antaranya adalah *Crème de la Mer* seharga \$175, *Facial Treatment Essence* seharga \$179, *The Moisturizing Soft Cream* seharga \$175, *Little Miss Miracle Limited-Edition Crème de la Mer* yang dihargai sekitar \$325, *R.N.A. POWER Face Cream* seharga \$230, serta *GenOptics Aura Essence Serum* dengan harga \$240. Selain itu, terdapat pula produk lain seperti *The Renewal Oil* yang dijual seharga \$245, *FUTURE RESPONSE Age Defense Creme* seharga \$195, *Vintage Single Extract Essence* seharga \$145, *Crème Ancienne®* dengan harga \$290, *The Concentrate* yang dibanderol sekitar \$370, dan masih banyak lagi. Salah satu produk unggulan lainnya adalah *Future Solution LX Total Protective Emulsion Broad Spectrum SPF 20 Sunscreen* yang ditawarkan dengan harga \$260. Semua produk dalam daftar ini menawarkan solusi perawatan kulit yang kaya akan nutrisi dan perlindungan optimal bagi mereka yang mengutamakan kualitas dan keunggulan.

Sephora menawarkan rangkaian produk kosmetik mewah yang menghadirkan beragam pilihan menarik dengan rentang harga mulai dari \$139 hingga \$370. Dengan berbagai pilihan produk yang tersedia, para konsumen dapat menemukan beragam opsi kosmetik premium sesuai dengan preferensi dan kebutuhan mereka. Mulai dari perawatan kulit hingga koleksi riasan yang elegan, rentang harga tersebut memberikan fleksibilitas kepada para konsumen untuk memilih produk-produk yang sesuai dengan anggaran dan gaya pribadi mereka.

C. *Daftar Produk dengan Harga Menengah (Mid-Range)*TABEL 4. PRODUK HARGA MENENGAH (*MID-RANGE*)

Nama Produk	Harga (\$)
Protini™ Polypeptide Cream	68
The Water Cream	68
Lala Retro™ Whipped Cream	60
Virgin Marula Luxury Facial Oil	72
Luna Sleeping Night Oil	105
Honeymoon Glow AHA Resurfacing Night Serum with Echinacea GreenEnvy™	58
The Littles™	90
Crème de la Mer Mini	85
Bio-Performance Advanced Super Revitalizing Cream	78
Black Tea Firming Overnight Mask	92
...	...
Prime & Protect Mattifying Primer with Broad Spectrum SPF 50	56

Tabel di atas menampilkan beberapa produk kosmetik Sephora yang termasuk dalam kisaran harga menengah (*mid-range*). Produk-produk ini menawarkan berbagai pilihan untuk perawatan kulit dan kecantikan. Mulai dari krim *peptida Protini™*, krim ringan *The Water Cream*, hingga *Moisturizer* mewah *Virgin Marula Luxury Facial Oil*. Tersedia pula beragam serum, seperti *Honeymoon Glow AHA Resurfacing Night Serum with Echinacea GreenEnvy™*, serta opsi perawatan malam lainnya seperti *Luna Sleeping Night Oil*. Bagi yang mencari pilihan lengkap, *The Littles™* menyajikan rangkaian produk dalam satu set. Produk-produk tersebut dapat menjadi pilihan untuk meningkatkan rutinitas perawatan kulit konsumen, dengan harga yang berada dalam kisaran menengah.

#### IV. KESIMPULAN

Penelitian ini berhasil menganalisis produk kosmetik dari Sephora dengan menggunakan algoritma *K-Means*, menghasilkan tiga kategori harga yaitu *affordable*, *mid-range*, dan *luxury*. Algoritma *K-Means* berhasil membentuk tiga *cluster* harga yang terdiri dari produk-produk dengan karakteristik harga yang berbeda:

- *Cluster 0 (Affordable)*: *Cluster* ini terdiri dari 953 item produk kosmetik dengan harga terjangkau. Produk-produk dalam *cluster* ini cocok untuk konsumen dengan anggaran yang lebih ketat.
- *Cluster 1 (Luxury)*: *Cluster* ini terdiri dari 82 item produk kosmetik dengan harga mewah. Produk-produk dalam *cluster* ini cenderung ditujukan untuk konsumen yang lebih mementingkan kualitas dan siap untuk mengeluarkan budget lebih besar.
- *Cluster 2 (Mid-Range)*: *Cluster* ini terdiri dari 437 item produk kosmetik dengan harga menengah. Produk-produk dalam *cluster* ini menawarkan keseimbangan antara kualitas dan harga, cocok untuk konsumen dengan anggaran sedang.

Visualisasi menggunakan *scatter/bubble plot* dan *Folder View* membantu mengidentifikasi pola dan hubungan antara produk kosmetik berdasarkan harga serta memahami karakteristik masing-masing *cluster*. Metode *K-Means* berhasil mengelompokkan produk kosmetik berdasarkan harga dengan cukup baik, memberikan pemahaman yang lebih mendalam tentang karakteristik dataset. Namun, penggunaan algoritma *K-Means* memerlukan pemilihan jumlah *cluster* ( $k$ ) yang tepat, dan analisis hanya berfokus pada faktor harga saja, sementara produk kosmetik memiliki berbagai atribut lainnya yang juga berpengaruh pada preferensi konsumen.

Sebagai rekomendasi pengembangan selanjutnya, dapat dicoba beberapa nilai  $k$  untuk menemukan jumlah *cluster* yang lebih sesuai dengan karakteristik dataset. Selain itu, menggabungkan algoritma *K-Means* dengan metode lain, seperti analisis faktor, dapat lebih mempertimbangkan atribut lain yang memengaruhi preferensi konsumen. Dengan menambahkan fitur interaktif pada visualisasi, pengguna dapat lebih mendalam dalam eksplorasi setiap *cluster*. Dengan demikian, hasil ini dapat menjadi dasar pengembangan lebih lanjut untuk memberikan rekomendasi pembelian yang lebih baik kepada konsumen berdasarkan preferensi anggaran mereka.

#### DAFTAR PUSTAKA

- [1] D. RIANI, "PENGARUH CELEBRITY ENDORSER DAN WORD OF MOUTH TERHADAP MINAT UNTUK BERPINDAH MEREK ( BRAND SWITCHING ) DENGAN PERCEIVED VALUE SEBAGAI MEDIASI ( Studi pada Konsumen Produk Kosmetik Wardah )," 2015.
- [2] M. C. Hackett, *The Sephora Story: The Retail Success You Can't Makeup (The Business Storybook Series)*. Cincinnati: HarperCollins Leadership, 2020.
- [3] E. Cahyono, "STRATEGI PENGELOLA PASAR TRADISIONAL DALAM MENGHADAPI KEBERADAAN SUPER MARKET ( ANALISIS ETIKA BISNIS DALAM ISLAM )," vol. 2, no. 7, pp. 2049–2060, 2023.
- [4] N. F. Andriani and A. H. Setiawan, "Analisis preferensi konsumen terhadap penggunaan produk skincare korea selatan dan lokal," vol. 9, pp. 1–8, 2020.
- [5] Y. Arfah, *Keputusan Pembelian Produk*. Padang: Inovasi Pratama Internasional, 2022.
- [6] Saluky *et al.*, *ILMU KOMPUTER*. Malang: Literasi Nusantara Abadi Grub, 2023.
- [7] A. A. Permana, W. S, L. W. Santoso, and G. W. N. Wibowo, *Machine Learning*, 1st ed. Padang: Global Eksekutif Teknologi, 2003.
- [8] Y. D. Darmi and A. Setiawan, "Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk," *J. Media Infotama*, vol. 12, no. 2, pp. 148–157, 2017, doi: 10.37676/jmi.v12i2.418.
- [9] E. Buulolo, *Data Mining Untuk Perguruan Tinggi*. Yogyakarta: Budi Utama, 202AD.
- [10] M. M. K-means, A. Solichin, and K. Khairunnisa, "Klasterisasi Persebaran Virus Corona ( Covid-19 ) Di DKI Jakarta," vol. 5, no. 2, 2020.
- [11] J. Han, M. Kamber, and J. Pei, *Data Mining : Concepts and Techniques (3rd Edition)*, 3rd ed. Morgan Kaufmann, 2018.