

# Analisis Perilaku Belanja dan Pengelompokan Pelanggan di Toko Ritel ABC Menggunakan Metode *Clustering*

<sup>1,2,3,4,5)</sup>Fajra Aqila Adesti, Sastyaviani Irami P, Ar Royyan Utama T, Nisrina Faiza Mufid,  
dan Samaya Dhiya Salindri

<sup>1,2,3,4,5)</sup>Teknik Industri, Fakultas Teknologi Industri, Universitas Islam Indonesia, Sleman, Yogyakarta,  
55584, Indonesia.

Email: [22522226@students.uii.ac.id](mailto:22522226@students.uii.ac.id), [21522203@students.uii.ac.id](mailto:21522203@students.uii.ac.id), [21522312@students.uii.ac.id](mailto:21522312@students.uii.ac.id),  
[21522290@students.uii.ac.id](mailto:21522290@students.uii.ac.id), [22522107@students.uii.ac.id](mailto:22522107@students.uii.ac.id)

## ABSTRAK

Ritel ABC merupakan salah satu ritel sepatu di Indonesia yang menyediakan perlengkapan olahraga dan gaya hidup, dengan lebih dari 500 cabang yang tersebar di seluruh negeri. Penelitian ini menggunakan metode Analisis *Cluster* untuk mengidentifikasi segmen pelanggan dan mengembangkan strategi pemasaran yang optimal. Penelitian ini mengungkapkan bahwa faktor-faktor seperti usia, jenis kelamin, pekerjaan, pendapatan, dan preferensi konsumen memengaruhi keputusan pembelian di ABC. Analisis *Cluster* digunakan untuk mengelompokkan konsumen berdasarkan karakteristik yang dimiliki, menghasilkan lima *cluster* dengan tingkat potensi yang berbeda untuk strategi pemasaran. *Cluster* 1 diidentifikasi sebagai *cluster* terbaik dengan komponen usia, jenis kelamin, pekerjaan, pendapatan, dan niat beli konsumen di atas rata-rata, meskipun kepuasan konsumen berada di bawah rata-rata. Hasil penelitian menunjukkan bahwa pemahaman mendalam terhadap karakteristik dan preferensi masing-masing segmen dapat membantu toko ritel merancang strategi pemasaran yang lebih efektif, meningkatkan manajemen persediaan, dan memperbaiki pelayanan pelanggan.

**Kata kunci:** *Cluster*, *K-Means*, Perilaku Konsumen, Ritel,

## 1. Pendahuluan

Saat ini, sektor ritel di Indonesia sedang mengalami perkembangan pesat. Hal ini terlihat dari kemunculan berbagai pasar modern yang mulai menggantikan pasar tradisional. Pusat Berkembangnya persaingan dalam bisnis ritel telah mendorong perusahaan-perusahaan untuk mencari cara yang lebih cepat dan efektif untuk menarik pelanggan dan mendorong pembelian. Perusahaan-perusahaan terus berupaya menggunakan berbagai metode dan strategi untuk memengaruhi keputusan pembelian konsumen. Keputusan pembelian ini menjadi sangat krusial karena persaingan semakin sengit dalam era pasar modern saat ini, dengan banyak perusahaan menawarkan beragam produk dan layanan kepada masyarakat. Seiring dengan peningkatan pilihan yang tersedia, masyarakat juga semakin banyak melakukan pertimbangan sebelum membuat keputusan pembelian. Kegagalan perusahaan dalam bersaing untuk memengaruhi keputusan pembelian konsumen dapat berdampak besar, termasuk penurunan penjualan hingga risiko kebangkrutan (Hidayatullah, 2020)

Saingan yang sengit dalam dunia bisnis menyebabkan pasar menjadi dibanjiri dengan produk yang hampir serupa, tetapi dengan harga yang lebih terjangkau. Situasi ini tentunya menarik bagi konsumen karena mereka memiliki lebih banyak pilihan dan produk yang lebih terjangkau. Keberagaman pilihan dan kemampuan finansial konsumen yang meningkat telah menyebabkan pembelian dan konsumsi produk kadang-kadang didasarkan pada keinginan semata, bukan hanya kebutuhan. Bahkan terkadang, meskipun daya beli seorang konsumen mungkin tidak begitu kuat, berbagai fasilitas pembayaran seperti pembelian kredit memberikan fleksibilitas tambahan, memungkinkan terpenuhinya keinginan konsumen yang tidak didasari oleh kebutuhan (S & Cahyo, 2017)

Perilaku konsumen merujuk pada tindakan di mana individu memilih, membeli, dan memanfaatkan produk atau layanan dengan tujuan memenuhi kebutuhan atau keinginan mereka. Saat seorang konsumen melakukan keputusan pembelian ke *sport station*, berbagai faktor memengaruhi keputusan mereka. Faktor-faktor ini mencakup aspek budaya, aspek sosial, aspek pribadi, dan aspek psikologis. Setelah memahami faktor-faktor tersebut, seorang konsumen perlu mempertimbangkan semua aspek ini sebelum membuat keputusan berbelanja (Mauludin, Saputra, Sari, Munawaroh, & Regita, 2022).

Oleh karena itu peneliti melakukan penelitian di ritel sepatu ABC. Karena peneliti merasa ritel ini merupakan salah satu ritel yang paling potensial, dengan berbagai macam sepatu dari berbagai merek, selain itu, ritel ini adalah ritel yang paling terkenal dan paling banyak konsumen dari berbagai macam kalangan. Dan hal ini sesuai dengan objek dan subjek yang akan diteliti oleh peneliti dengan menggunakan analisis *cluster*.

Analisis *Cluster* adalah metode multivariat yang bertujuan untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimiliki oleh masing-masing objek. Dalam analisis ini, objek-objek dikelompokkan sedemikian rupa sehingga objek yang memiliki kesamaan karakteristik paling tinggi akan berada dalam kelompok yang sama. Kelompok-kelompok yang terbentuk memiliki tingkat kesamaan yang tinggi di dalam kelompoknya dan perbedaan yang tinggi antar kelompok. Berbeda dengan metode multivariat lainnya, analisis ini tidak menggunakan estimasi empiris untuk variabel-variabel, melainkan variabel-variabel yang digunakan ditentukan oleh peneliti sendiri. Fokus utama dari Analisis *Cluster* adalah membandingkan objek-objek berdasarkan sekumpulan variabel yang digunakan sebagai dasar analisis. Oleh karena itu, para ahli menganggap penggunaan variabel tersebut sebagai langkah kritis dalam proses Analisis *Cluster*. Variabel-variabel yang digunakan dalam analisis *cluster* adalah sekumpulan variabel yang mencerminkan karakteristik objek-objek tersebut.

Solusi dalam Analisis *Cluster* tidak bersifat unik, karena anggota-anggota dalam kelompok *cluster* tergantung pada beberapa elemen dalam prosedur analisis. Sebagai hasilnya, berbagai solusi yang berbeda dapat diperoleh dengan mengubah satu atau lebih elemen dalam analisis. Keseluruhan solusi dalam analisis *cluster* juga sangat tergantung pada variabel-variabel yang digunakan untuk menentukan tingkat kesamaan antara objek-objek. Penambahan atau pengurangan variabel yang relevan dapat mempengaruhi hasil analisis *cluster* secara substansial (Ediyanto, Mara, & Satyahadewi, 2013).

Analisis *cluster* sendiri dibagi menjadi dua, yakni metode hierarki dan metode non-hierarki atau yang lebih dikenal sebagai *K-Means*. *K-Means* adalah algoritma *unsupervised learning* yang digunakan untuk mengelompokkan data ke dalam *cluster*. Algoritma ini dapat beroperasi tanpa memerlukan label kategori pada data yang dimasukkan. *K-Means Clustering* adalah metode non-hierarki yang bertujuan mengelompokkan data ke dalam kelompok-kelompok dengan karakteristik serupa, yang berbeda dari kelompok lain. Dalam pengambilan sampel menggunakan *Cluster Sampling*, unit-unit populasi dipilih secara acak dari kelompok yang sudah ada, yang disebut *cluster*. *Clusterisasi* atau *clustering* pada dasarnya merupakan salah satu masalah yang menggunakan teknik *unsupervised learning* (Sandova, Kurniawan, & Supratati, 2024).

Dalam penelitian Segmentasi Pelanggan Menggunakan Metode *K-Means Clustering* Berdasarkan Model RFM (Recency, Frequency, Monetary), metode *clustering K-Means* digunakan untuk analisis ini. Dengan menggunakan metode *Elbow* dan nilai koefisien *Silhouette*, ditentukan bahwa nilai  $K=2$  adalah yang paling optimal untuk *clustering*, menghasilkan dua segmen pelanggan dengan karakteristik berbeda. Segmen pelanggan ini merupakan kelanjutan dari model RFM, yang didasarkan pada rentang waktu terakhir transaksi,

frekuensi transaksi, dan jumlah uang yang dikeluarkan untuk transaksi. Dari 4226 pelanggan yang disegmentasi, 3591 pelanggan dikategorikan sebagai *Potential Customers*, sementara 635 pelanggan dikategorikan sebagai *Non-Potential Customers*. *Potential Customers* menunjukkan karakteristik sering melakukan pembelian dengan jumlah pengeluaran yang tinggi dalam transaksi (Anshary, Soesanto, & Ayatullah, 2022).

Dalam penelitian Analisis Perilaku Konsumen pada Usaha Ritel dengan menggunakan Metode *Association Rule - Market Basket Analysis* dan *Clustering* sebagai Usulan Strategi Peningkatan Penjualan (Studi Kasus: Intimart Gedongan), Berdasarkan grafik *elbow method* dan *silhouette score*, diperoleh jumlah *cluster* sebanyak empat untuk mengelompokkan data transaksi pelanggan. Analisis hasil *clustering* menggunakan algoritma *K-Means* menunjukkan proporsi *cluster* 1 sebanyak 40.4% dengan 12.825 transaksi, *cluster* 2 sebesar 13.2% dengan 4.191 transaksi, *cluster* 3 sebesar 44.7% dengan 14.172 transaksi, dan *cluster* 4 sebesar 1.7% dengan 547 transaksi. *Cluster* 1, atau "*Mid-Morning Moderates*," terdiri dari pelanggan yang bertransaksi antara pukul 8 pagi hingga 12 siang, biasanya membeli produk yang sama atau dengan sedikit variasi, dengan nilai transaksi rata-rata Rp. 19.866,- dan rentang total transaksi antara Rp. 8.925,- hingga Rp. 26.540,-. *Cluster* 2, atau "*Diverse Afternoon Buyers*," terdiri dari pelanggan yang bertransaksi antara pukul 1 siang hingga 7 malam, biasanya membeli berbagai variasi produk dengan nilai transaksi rata-rata Rp. 49.271,- dan rentang total transaksi antara Rp. 31.000,- hingga Rp. 63.434,-. *Cluster* ini mencerminkan pelanggan yang aktif dan beragam dalam pembelian produk, menjadi target strategis untuk promosi produk bervariasi dan peningkatan nilai transaksi. *Cluster* 3, atau "*Evening Moderates*," terdiri dari pelanggan yang bertransaksi antara pukul 5 sore hingga 8 malam, biasanya membeli produk dengan sedikit variasi, dengan nilai transaksi rata-rata Rp. 19.591,- dan rentang total transaksi antara Rp. 9.500,- hingga Rp. 26.400,-. *Cluster* ini memiliki kemiripan dengan *cluster* 1. *Cluster* 4, atau "*High-Value Customers*," terdiri dari pelanggan yang bertransaksi antara pukul 10 pagi hingga 6 sore, biasanya membeli berbagai variasi produk dengan nilai transaksi rata-rata Rp. 177.475,- dan rentang total transaksi antara Rp. 128.050,- hingga Rp. 209.883,-. *Cluster* ini merupakan kelompok pelanggan yang sangat berharga karena kontribusinya yang signifikan pada pendapatan toko (Baihaqie, 2023).

Dalam penelitian Pengelompokan Data Transaksi Dalam Menentukan Strategi Penjualan Menggunakan Algoritma *K-Means* Hasil pengujian data menggunakan *tools* RapidMiner pada klasterisasi data penjualan di Toko Ritel *Accessories* dengan algoritma *K-Means Clustering* menunjukkan bahwa nilai Davies-Bouldin Index (DBI) dari  $k=2$  hingga  $k=10$  terdapat pada tabel sebelumnya. Berdasarkan hasil tersebut, *cluster* dengan nilai terkecil adalah  $k=5$ , dengan DBI sebesar 0,286. Karena nilai DBI untuk  $k=5$  adalah yang paling rendah dibandingkan dengan nilai  $k$  lainnya, dapat disimpulkan bahwa  $k=5$  dengan nilai 0,286 adalah hasil *cluster* terbaik. Beberapa saran untuk peneliti lain yang ingin melakukan penelitian serupa berdasarkan pengalaman pribadi selama penelitian ini adalah pentingnya pemahaman yang baik terhadap data yang akan diolah dan algoritma yang digunakan. Penelitian ini diharapkan dapat membantu pihak toko dalam mengambil keputusan untuk menentukan strategi penjualan (Asih, Irawan, & Bahtiar, 2024).

Dalam penelitian Klasterisasi Stok Produk Retail Untuk Menentukan Pergerakan Kebutuhan Konsumen Dengan Algoritma *K-Means* yang telah dilakukan, dapat disimpulkan bahwa penelitian ini memanfaatkan data produk retail dengan beberapa atribut, seperti Data Kebutuhan Awal dan Data Kebutuhan Akhir. Dengan pendekatan metode klasterisasi, data yang ada dapat digunakan untuk menganalisis pengelompokan produk terkait ketersediaan dan persediaan barang di gudang, sehingga memberikan wawasan dan informasi berharga. Proses

pengolahan data produk retail menggunakan algoritma *K-Means* meliputi tahap seleksi data (atribut yang digunakan), pengujian algoritma (*K-Means*), dan validasi (menggunakan *k-fold cross-validation*). Hasil dari penerapan algoritma *K-Means* menunjukkan pengelompokan produk ke dalam tiga *cluster*: *Cluster 1*, yang mencakup produk dengan ketersediaan rendah (*Low*) dengan 939 dari 1000 kategori produk yang diuji; *Cluster 2*, yang mencakup produk dengan ketersediaan sedang (*Medium*) dengan 51 dari 1000 kategori; dan *Cluster 3*, yang mencakup produk dengan ketersediaan tinggi (*High*) dengan 10 dari 100 kategori produk yang diuji. Pengujian menggunakan *tools* RapidMiner menghasilkan wawasan serupa, dengan *Cluster 0* memiliki 51 anggota klaster yang mewakili kategori *Medium*, *Cluster 1* memiliki 939 anggota mewakili kategori *Low*, dan *Cluster 2* memiliki 10 anggota yang sesuai dengan kategori *High* (Suwaryo, Rahman, Atmaja, & Basri, 2023).

Dalam penelitian *clustering* data penjualan produk makanan pada Toko Toserba Yogya Siliwangi menggunakan metode *K-Means* dilakukan dengan data penjualan kategori snack pada bulan Desember 2021, yang terdiri dari 470 data. Dengan menggunakan metode pengelompokan *K-Means* melalui aplikasi RapidMiner, diperoleh hasil bahwa data terbagi menjadi tiga *cluster*: *Cluster 1* dengan 3 item, *Cluster 2* dengan 13 item, dan *Cluster 0* dengan 454 item. Selain itu, hasil *running* dari *Davies-Bouldin Index* adalah 0,487, yang mendekati nilai 1 (Noviati, Mulyawan, Kurnia, & Rinaldi, 2022).

Dalam penelitian implementasi Metode *K-Means Clustering* Untuk Pengelompokan Data Stok Produk Toko Online Perdagangan Kaos, dari hasil pembahasan mengenai klasifikasi produk menggunakan *data mining* di toko khusus *T-shirt* dengan metode *K-Means*, kesimpulan yang dapat diambil adalah sebagai berikut: diperlukan sistem *data mining* berbasis *K-Means* untuk mengatasi masalah klasifikasi produk di industri kaos. Penelitian ini merekomendasikan pengembangan sistem klasifikasi pola penjualan produk di toko dan menyarankan pembuatan aplikasi untuk mendukung transaksi jual beli kaos. Meskipun penelitian ini memberikan wawasan berharga, masih ada kekurangan yang perlu diperbaiki untuk meningkatkan kinerja sistem. Kriteria penelitian yang ada belum sepenuhnya memenuhi kebutuhan sistem kategorisasi persediaan menggunakan *K-Means clustering*. Penetapan *centroid* awal dapat menghasilkan *centroid* yang berbeda, sehingga disarankan untuk meningkatkan akurasi proses *clustering*. Untuk hasil yang lebih akurat, penelitian selanjutnya dapat membandingkan algoritma *K-Means* dengan algoritma lain atau mengembangkan metode *clustering K-Means* yang lebih efektif (Wahyudi & Wibowo, 2022).

Dalam penelitian Penerapan *Data Mining* Dalam Pengelompokan Bahan Sembako Laris Menggunakan *K-Means Clustering* (Studi Kasus: Toko Gunung Bumi) Berdasarkan hasil pemodelan menggunakan algoritma *K-Means Clustering* dan data yang telah diolah dengan *Knowledge Discovery in Database (KDD)*, diperoleh dua kategori penjualan sembako: "tidak laris" dan "laris." Penentuan nilai *centroid* dilakukan secara acak dengan menggunakan rumus *Euclidean Distance*. Proses ini menghasilkan dua iterasi dengan nilai rata-rata sebagai berikut: C0 dengan rata-rata 397.950 dan 349.400, serta C1 dengan rata-rata 1.248.300 dan 1.272.700. Hasilnya menunjukkan bahwa C0 mencakup 20 barang yang tidak laris dan C1 mencakup 10 barang yang laris. Evaluasi dari *cluster* yang dibentuk menunjukkan dua *cluster* dengan nilai 0,605. Dengan demikian, penelitian ini sangat relevan untuk digunakan dalam penyetokan barang guna meningkatkan penjualan di Toko Gunung Bumi. (Permana & Sanjaya, 2023).

## 2. Metode

Pada penelitian kali ini peneliti mengumpulkan data dengan melakukan observasi secara langsung terhadap variabel yang sudah ditentukan, setelah dihasilkan rumusan masalah studi

kasus, langkah selanjutnya peneliti mulai mengumpulkan informasi, menyusun kuesioner untuk disebarkan kepada responden dari berbagai kalangan usia, latar belakang, minat dan gaya hidup yang berjumlah 50 orang. Selanjutnya data direkap menggunakan excel untuk melakukan transformasi data sebelum diolah di SPSS menggunakan metode *cluster*, metode ini merupakan metode multivariat yang bertujuan untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimiliki oleh masing-masing objek (Ediyanto et al., 2013). Metode *cluster* ini sendiri dipecah lagi menjadi dua, yakni hierarki dan non hierarki (*K-Means*). *K-Means* adalah algoritma yang digunakan untuk pengelompokan secara partisi. Setiap kelompok dipisahkan datanya ke dalam kelompok yang berbeda-beda. Dalam *K-Means*, diterapkan metode data *clustering* non-hierarki. Metode ini melakukan partisi data ke dalam satu atau lebih *cluster*/kelompok (Virgo, Defit, & Yunus, 2020). Setelah diolah, hasil didapatkan berupa dendogram yang digunakan untuk menentukan seberapa banyak *cluster* terbentuk.

Proses *knowledge discovery in database* (KDD) mengandung beberapa proses. Gambar 1 menunjukkan bagaimana sebuah metode KDD dijalankan secara urut mulai dari pemilihan data, melakukan *pre-processing* pada data tersebut agar kemudian dapat digunakan, melakukan transformasi, kemudian penambangan data, evaluasi, dan pada akhirnya dapat diperoleh informasi berupa kesimpulan terhadap suatu populasi data. Rangkaian proses ini dilakukan sebagai bagian dari penelitian yang dilakukan.



Gambar 1. Proses KDD

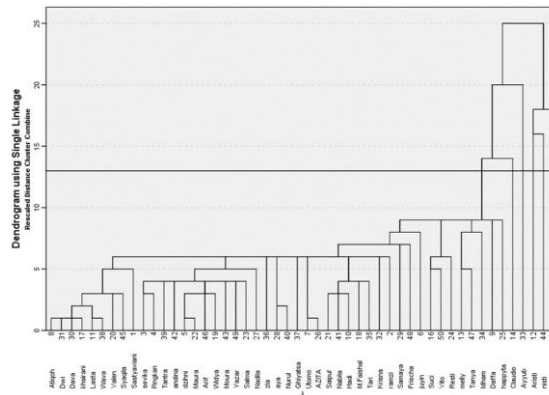
- Selection*: Menyaring data mentah untuk menentukan variabel data yang akan diproses.
- Preprocessing*: Tahap ini melibatkan pembersihan data dari *noise*, nilai yang hilang, dan data yang tidak relevan.
- Transformation*: Mengubah data ke dalam format yang sesuai untuk pengolahan oleh algoritma *data mining* yang akan digunakan, karena setiap algoritma membutuhkan format data tertentu.
- Data mining*: Merupakan tahap inti dari KDD (*Knowledge Discovery in Databases*), yaitu mengekstrak pengetahuan dari data menggunakan berbagai metode dan algoritma. Metode *data mining* termasuk *clustering*, *classification*, *association*, dan lainnya, dengan algoritma seperti C4.5, *K-Means*, dan Apriori yang telah berkembang menjadi FP-Growth. Penelitian ini menerapkan metode *clustering* dan algoritma *K-Means*.
- Interpretation/Evaluation*: Menginterpretasikan pengetahuan atau pola yang ditemukan dari proses *data mining* dan mengevaluasi apakah pola tersebut sesuai dengan tujuan yang diinginkan (Syahputra, Mayola, & Guswandi, 2022).

### 3. Hasil dan Pembahasan

Metode *clustering* yang digunakan pada penelitian kali ini menggunakan dua pendekatan yaitu menggunakan pendekatan hierarki dan non-hierarki,

a. Metode Hierarki

Hasil yang didapatkan dari pengelompokan dan pengujian menggunakan analisis *clustering* dapat dilihat pada hasil dendrogram sebagai berikut:



Gambar 2. Hasil Dendrogram

Hasil dendrogram yang ditunjukkan oleh Gambar 2 telah melewati proses pemotongan pada titik 13. Dari proses ini, didapatkan hasil bahwa terdapat beberapa *outliers* yaitu data yang bernomor urut 14, 33, 12, 44, dan 5. Titik pemotongan pada dendrogram ini dipilih berdasarkan subjektivitas dari peneliti yang ingin melihat jumlah data yang sangat berbeda jauh dibandingkan dengan rata-rata data yang telah dikumpulkan. Terlihat pula bahwa hasil dari pemotongan ini telah membentuk sebuah *cluster* besar yang berisikan anggota berlabel 8 hingga 25. Ini menunjukkan bahwa data-data yang tidak tergabung ke dalam *cluster* tersebut merupakan *outliers* yang berarti data tersebut memiliki perbedaan yang signifikan dibanding dengan data lainnya.

Tabel 1. Agglomeration Schedule

Agglomeration Schedule						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	8	31	,000	0	0	2
2	8	30	,570	1	0	6
3	7	26	,570	0	0	34
4	11	38	,666	0	0	8
5	5	22	,666	0	0	9
6	8	17	,745	2	0	8
7	28	40	1,550	0	0	30
8	8	11	1,550	6	4	11
9	5	46	1,940	5	0	15
10	21	41	1,981	0	0	13
11	8	20	2,120	8	0	14
12	3	4	2,120	0	0	16
13	10	21	2,215	0	10	22
14	8	45	2,685	11	0	26
15	5	19	2,685	9	0	19
16	3	39	2,780	12	0	18

17	18	35	2,780	0	0	22
18	3	42	2,865	16	0	23
19	5	43	2,919	15	0	20
20	5	49	2,960	19	0	21
21	5	23	3,350	20	0	23
22	10	18	3,350	13	17	31
23	3	5	3,585	18	21	27
24	16	50	3,692	0	0	32
25	13	47	3,978	0	0	39
26	1	8	3,978	0	14	28
27	3	27	4,023	23	0	28
28	1	3	4,593	26	27	29
29	1	36	4,667	28	0	30
30	1	28	4,681	29	7	33
31	10	32	4,689	22	0	34
32	16	24	4,768	24	0	41
33	1	37	4,782	30	0	35
34	7	10	5,061	3	31	35
35	1	7	5,237	33	34	36
36	1	2	5,412	35	0	37
37	1	29	5,527	36	0	38
38	1	48	6,146	37	0	40
39	13	34	6,647	25	0	42
40	1	6	6,810	38	0	41
41	1	16	7,361	40	32	42
42	1	13	7,440	41	39	43
43	1	9	7,452	42	0	44
44	1	25	8,065	43	0	45
45	1	14	12,267	44	0	48
46	12	44	14,148	0	0	47
47	12	15	16,212	46	0	49
48	1	33	17,430	45	0	49
49	1	12	22,599	48	47	0

Kemudian, hasil yang ditunjukkan pada Tabel 1 menunjukkan tentang bagaimana perilaku penggabungan oleh setiap data. Analisis *clustering* merupakan sebuah metode yang memiliki beberapa cara dalam pelaksanaannya. Penelitian ini mengadopsi teknik hierarki menggunakan metode aglomerasi, yang dalam pelaksanaannya, proses dimulai dari penggabungan dua objek yang paling mirip satu sama lainnya.

Berdasarkan Tabel 1, peneliti dapat mengetahui data mana yang paling mirip satu sama lainnya berdasarkan urutan bergabungnya data tersebut. Dapat dilihat bahwa pada urutan pertama, data 8 dan data 31 merupakan yang paling mirip satu sama lainnya. Dengan koefisien kesamaannya adalah 0.000, ini menunjukkan bahwa kedua data tersebut adalah sama persis. Berdasarkan analisa ini, orang yang bertanggung jawab pada toko ritel ABC dapat melakukan penelusuran pada pelanggan yang memiliki label tertentu, dan melihat kesamaan antara satu pelanggan dan pelanggan lainnya berdasarkan profil dari pelanggan-pelanggan yang

dibandingkan. Hal ini menjadi sangat berguna ketika sebuah toko ritel ingin melihat profil yang dimiliki oleh pelanggan yang datang. Toko ritel ABC dapat melakukan penelusuran profil pelanggan dengan menambahkan dan menggunakan fitur *membership* pada toko tersebut. Hal ini akan membuat pelanggan akan memasukkan beberapa data diri yang akan berguna bagi toko ritel ABC dalam analisa profil pelanggan.

Pada pengelompokan *cluster* yang dilakukan menggunakan metode hierarki, didapat sebanyak 5 *cluster* dengan karakter unik masing-masing. Hasil yang akan didapat adalah berupa profilisasi dari setiap *cluster*.

Tabel 2. Rangkuman Hasil Crosstab

Variabel	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Usia (↑)	0,08110 (100,0%) ↑	2,10860 (50,0%) ↑	2,10860 (50,0%) ↑	-1,94640 (16,7%) ↓	4,13609 (100,0%) ↑
Jenis Kelamin (↑)	-1,51217 (100%) ↓	0,64807 (2,9%) ↑	0,64807 (2,9%) ↑	0,64807 (2,9%) ↑	0,64807 (2,9%) ↑
Pekerjaan (↑)	-0,23497 (97,9%) ↓	2,70216 (50,0%) ↑	2,70216 (50,0%) ↑	-0,23497. (2,1%) ↓	5,63930 (100,0%) ↑
Pendapatan (↑)	0,03021 (100,0%) ↑	2,29575 (25,0%) ↑	2,29575 (25,0%) ↑	0,78539 (20,0%) ↑	2,29575 (25,0%) ↑
Pengeluaran (↓)	-1,39819 (100,0%) ↓	2,91721 (50,0%) ↑	2,91721. (50,0%) ↑	0, 32797 (9,1%) ↑	2,05413 (50,0%) ↑
Perilaku Konsumen (↑)	-1,16333. (95,2%) ↓	0,84241 (3,4%) ↑	-1,16333. (4,8%) ↓	-0,23497 (3,4%) ↓	5,63930 (3,4%) ↑
Niat Beli Konsumen (↑)	0,03021 (100,0%) ↑	-0,72555 (3,2%) ↓	-0,72555 (3,2%) ↓	1,61494 (8,3%) ↑	-0,72555 (3,2%) ↓
Jarak Lokasi (↓)	-2,16604 (100,0%) ↓	-0,92119 (10,0%) ↓	-0,72555 (4,8%) ↓	-0,23050 (10,0%) ↓	5,53212 (10,0%) ↑
Harga (↑)	-1,02795 (100,0%) ↓	1,41955 (20,0%) ↑	-1,02795 (5,9%) ↓	3,86704 (100,0%) ↑	0,60371 (6,7%) ↑
Kepuasan Konsumen (↑)	-0,49497 (92,5%) ↓	-0,49497 (2,5%) ↓	-0,49497 (2,5%) ↓	1,97990 (10,0%) ↑	-0,49497 (2,5%) ↓

Berdasarkan Tabel 2, dapat dilihat bahwa arah panah yang ada pada setiap variabel menunjukkan harapan dari peneliti terhadap variabel tersebut. Arah panah ke atas menunjukkan bahwa variabel tersebut diharapkan memiliki nilai positif pada koefisiennya. Harapan ini berdasar pada kriteria terbaik yang dapat dimiliki oleh sebuah *cluster*. Dapat dilihat bahwa *cluster* 1 memiliki usia di atas rata-rata total, jenis kelamin di bawah rata-rata total, pendapatan di atas rata-rata total, pengeluaran di bawah rata-rata total, perilaku konsumen di bawah rata-rata total, niat beli konsumen di atas rata-rata total, jarak lokasi di bawah rata-rata total, harga di bawah rata-rata total, dan kepuasan juga termasuk di bawah rata-rata total. Hasil ini dapat dilihat sebagai sebuah *cluster* yang memiliki potensi cukup bagus untuk dijadikan sebagai target utama pelanggan dari toko ritel ABC. Hal ini didukung oleh usia pelanggan yang termasuk dalam



*cluster* 1 merupakan usia yang diharapkan yaitu berada di atas rata-rata total, kemudian memiliki pendapatan yang juga diharapkan yaitu berada di atas rata-rata total, serta memiliki nilai beli yang tinggi. Variabel-variabel yang disebutkan dapat menjadi pendukung bagi pelanggan tersebut untuk melakukan *repeat buying* dan berpotensi menjadi pelanggan tetap dari toko ritel ABC.

#### b. Metode Non-Hierarki

Metode non-hierarki yang digunakan pada penelitian kali ini adalah metode *K-Means clustering*. *K-Means* adalah algoritma yang digunakan untuk pengelompokan secara partisi. Setiap kelompok dipisahkan datanya ke dalam kelompok yang berbeda-beda. Dalam *K-Means*, diterapkan metode data *clustering* non-hierarki. Metode ini melakukan partisi data ke dalam satu atau lebih *cluster*/kelompok (Virgo et al., 2020). Pada metode non-hierarki jumlah *cluster* yang akan terbentuk telah ditentukan sejak awal oleh peneliti, yaitu sejumlah 3 *cluster*. Dalam penentuan jumlah *cluster* ini, peneliti ingin melihat anggota yang sebelumnya mungkin tersebar ke dalam 5 *cluster*, akan dapat dilihat kecenderungan untuk bergabung ke dalam *cluster* yang lebih luas karakteristiknya.

**Final Cluster Centers**

	Cluster		
	1	2	3
Zscore(Usia)	2,78443	-,32440	-,13809
Zscore(Jenis_Kelamin)	,64807	,43205	-,16932
Zscore(Pekerjaan)	3,68121	-,23497	-,23497
Zscore(Pendapatan)	2,29575	,40780	-,29636
Zscore(Pengeluaran)	2,62952	,41428	-,32517
Zscore (Perilaku_Konsumen)	,17383	,24069	-,07915
Zscore (Niat_Beli_Konsumen)	-,72555	,56172	-,09299
Zscore(Jarak_Lokasi)	,73861	,57263	-,21465
Zscore(Harga)	,33177	1,09321	-,32236
Zscore (Kepuasan_Konsumen)	-,49497	1,23744	-,29431

**Gambar 3. Final Cluster Centers**

Hasil yang ditunjukkan oleh Gambar 3 memperlihatkan titik tengah dari *cluster* dari setiap variabel per *cluster*. Pada *cluster* 1, kita dapat lihat bahwa koefisien yang dimiliki oleh variabel usia, jenis kelamin, pekerjaan, pendapatan, pengeluaran, perilaku konsumen, jarak, dan harga yang diharapkan oleh pelanggan adalah positif. Ini menjadi sebuah informasi yang sangat berguna bagi toko ritel ABC karena dapat digunakan sebagai acuan untuk memikirkan strategi bisnis yang dapat diterapkan kepada calon pelanggan yang memiliki karakteristik seperti pada *cluster* 1.

#### 4. Simpulan

Berdasarkan hasil perhitungan, *cluster* terbaik yang dapat direkomendasikan adalah *cluster* 1. *Cluster* ini memiliki usia, jenis kelamin, pekerjaan, pendapatan, dan pengeluaran di atas rata-rata, dengan perilaku konsumen dan niat beli di posisi yang cukup tinggi. Meskipun kepuasan konsumen berada di bawah rata-rata, *cluster* ini sangat potensial untuk dipertimbangkan dalam analisis strategi pemasaran baru. Sebaliknya, *cluster* 1 dan *cluster* 4 juga menunjukkan potensi yang cukup untuk pertimbangan strategi pemasaran, sementara *cluster* 3 dan *cluster* 5 kurang potensial dalam hal tersebut.

### Daftar Pustaka

- Anshary, M. H., Soesanto, O., & Ayatullah. (2022). Berdasarkan grafik elbow method dan silhouette score, didapatkan jumlah kluster sebanyak. *RAGAM: Journal of Statistics and Its Application*, 09.
- Asih, D. P., Irawan, B., & Bahtiar, A. (2024). Pengelompokan Data Transaksi Dalam Menentukan Strategi Penjualan Menggunakan Algoritma K-Means . *Jurnal Mahasiswa Teknik Informatika*, 146.
- Baihaqie, M. R. (2023). *Analisis Perilaku Konsumen pada Usaha Ritel dengan menggunakan Metode Association Rule - Market Basket Analysis dan Clustering sebagai Usulan Strategi Peningkatan Penjualan (Studi Kasus: Intimart Gedongan)*. Yogyakarta: Universitas Islam Indonesia.
- Ediyanto, Mara, M. N., & Satyahadewi, N. (2013). Pengklasifikasian Karektiristik Dengan Metode K-Means Cluster Analysis. *Buletin Ilmiah Mat. Stat. dan Terapannya (Bimaster)*, 134.
- Hidayatullah, R. (2020). *Pengaruh Store Atmosphere Dan Price Discount Terhadap Purchase Decision Dengan Hedonic Shopping Motivation Sebagai Variabel Intervening*. Jakarta: Universitas Esa Unggul.
- Mauludin, M. S., Saputra, A. D., Sari, A. Z., Munawaroh, I., & Regita, E. P. (2022). Analisis Perilaku Konsumen Dalam Transaksi Di e-Commerce. *Dynamic Management Journal*, 110.
- Noviati, Mulyawan, Kurnia, D. A., & Rinaldi, A. R. (2022). Clustering Data Penjualan Produk Makanan pada Toko Toserba. *MEANS (Media Informasi Analisa dan Sistem)*, 83.
- Permana, A. S., & Sanjaya, R. (2023). Penerapan Data Mining Dalam Pengelompokan Bahan Sembako Laris Menggunakan K-Means Clustering(Studi Kasus: Toko Gunung Bumi). *Jurnal Pendidikan Tambusai*, 2423.
- Ramadhani, L., I. P., & Amijaya, F. D. (2018 ). Penerapan Metode Complete Linkage dan Metode Hierarchical Clustering Multiscale Bootstrap (Studi Kasus: Kemiskinan Di Kalimantan Timur Tahun 2016). *Jurnal EKSPONENSIAL* .
- S, A. J., & Cahyo, H. (2017). Perilaku Konsumen Materialistik: Perspektif Pembelian, Complusive Buying dan Environmental Attitudes. *Universitas Wijayakusuma Purwokerto*, 178.
- Sandova, F., Kurniawan, R., & Supratati, T. (2024). Penerapan Data Mining Menggunakan Metode K-Means Clustering Pada Penjualan Tas Di Asia Toserba Cirebon. *Jurnal Mahasiswa Teknik Informatika*, 245.
- Suwaryo, N., Rahman, A., Atmaja, D. M., & Basri, A. (2023). Klasterisasi Stok Produk Retail Untuk Menentukan Pergerakan Kebutuhan Konsumen Dengan Algoritma K-Means. *Bulletin of Information Technology (BIT)*, 311.
- Syahputra, H., Mayola, L., & Guswandi, D. (2022). Clustering Tingkat Penjualan Menu (Food and Beverage)Menggunakan Algoritma K-Means. *Jurnal KomtekInfo*, 30-31.
- Virgo, I., Defit, S., & Yunus, Y. (2020). Klasterisasi Tingkat Kehadiran Dosen Menggunakan Algoritma K-Means Clustering. *Jurnal Sistem Informasi dan Teknologi*, 24.
- Wahyudi, S. I., & Wibowo, A. (2022). mplementasi Metode K-Means Clustering UntukPengelompokan Data Stok Produk Toko OnlinePerdagangan Kaos. *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi*, 477.