

Analisis Segmentasi Transaksi Penjualan Produk Apple secara Global Menggunakan Algoritma K-Means Clustering

Rizky Eka Adinagoro^{1*}, Dhiwa Irsyad Sabillarrasyid², Muh. Lutfi Al-Hafizh³, Arsa Lumaksana Aji⁴, Aprilisa Arum Sari⁵

¹Teknik Informatika
(Universitas Duta Bangsa Surakarta)
^{1*}230103150@udb.ac.id

²Teknik Informatika
(Universitas Duta Bangsa Surakarta)
²230103159@udb.ac.id

³Teknik Informatika
(Universitas Duta Bangsa Surakarta)
³230103171@udb.ac.id

⁴Teknik Informatika
(Universitas Duta Bangsa Surakarta)
⁴230103155@udb.ac.id

⁵Teknik Informatika
(Universitas Duta Bangsa Surakarta)
⁵aprilisa_arumsari@udb.ac.id

Abstrak— Industri teknologi konsumen global menghasilkan volume data transaksi yang sangat besar, namun potensi data tersebut belum sepenuhnya dimanfaatkan untuk mendukung pengambilan keputusan bisnis yang strategis. Penelitian ini bertujuan untuk melakukan segmentasi terhadap data transaksi penjualan produk Apple secara global menggunakan algoritma K-Means Clustering, guna mengidentifikasi kelompok-kelompok transaksi yang memiliki kesamaan karakteristik dan menghasilkan wawasan yang dapat diterapkan dalam strategi penjualan. Dataset yang digunakan terdiri dari 11.500 baris dan 27 kolom yang mencakup periode tahun 2022 hingga 2024, meliputi 47 negara dan enam kategori produk Apple. Setelah proses pra-pemrosesan berupa penghapusan nilai kosong pada fitur clustering dan normalisasi menggunakan StandardScaler, diperoleh 8.140 data yang valid. Fitur yang digunakan sebagai input clustering meliputi *revenue_usd*, *units_sold*, *discount_pct*, dan *customer_rating*. Penentuan jumlah kluster optimal dilakukan menggunakan Elbow Method yang mengidentifikasi $K = 4$ sebagai nilai optimal, dan kualitas kluster divalidasi menggunakan Silhouette Score dengan nilai 0,3120. Hasil clustering menghasilkan empat segmen transaksi yang berbeda secara karakteristik, yaitu segmen transaksi bernilai tinggi dengan volume pembelian besar, segmen sensitif terhadap diskon, segmen dengan kepuasan pelanggan tinggi, serta segmen dengan kepuasan pelanggan rendah. Temuan penelitian menunjukkan bahwa tingkat kepuasan pelanggan tidak berkorelasi langsung dengan nilai transaksi maupun besaran diskon. Hasil segmentasi ini dapat dimanfaatkan sebagai dasar perancangan strategi pemasaran yang lebih tertarget sesuai karakteristik masing-masing segmen.

Kata kunci— K-Means Clustering, Data Mining, Segmentasi Penjualan, Produk Apple, Customer Rating

Abstract— The global consumer technology industry generates enormous volumes of transactional data, yet the potential of such data remains underutilized in supporting strategic business decision-making. This study aims to segment global sales transaction data of Apple products using the K-Means Clustering algorithm, in order to identify groups of transactions with similar characteristics and generate actionable insights for sales strategy. The dataset consists of 11,500 rows and 27 columns covering the period from 2022 to 2024, encompassing 47 countries and six Apple product categories. Following preprocessing that included removal of missing values in clustering features and normalization using StandardScaler, 8,140 valid records were retained. The features used as clustering inputs were *revenue_usd*, *units_sold*, *discount_pct*, and *customer_rating*. The optimal number of clusters was determined using the Elbow Method, which identified $K = 4$ as the optimal value, and cluster quality was validated using the Silhouette Score, yielding a value of 0.3120. The clustering results produced four characteristically distinct transaction segments: a high-value segment with large purchase volumes, a discount-sensitive segment, a segment with high customer satisfaction, and a segment with low customer satisfaction. The findings reveal that customer satisfaction levels do not directly correlate with transaction value or discount amount. These segmentation results can serve as a foundation for designing more targeted marketing strategies tailored to the characteristics of each identified segment.

Keywords— K-Means Clustering, Data Mining, Sales Segmentation, Apple Products, Customer Rating

I. PENDAHULUAN

Industri teknologi konsumen global mengalami pertumbuhan yang sangat pesat dalam satu dekade terakhir. Apple Inc. sebagai salah satu perusahaan

teknologi terbesar di dunia berhasil mempertahankan dominasinya di pasar smartphone global dengan meraih pangsa pasar sebesar 20,1% pada tahun 2023, melampaui Samsung yang berada di posisi 19,4% [1]. Lebih jauh, Apple mencatat

pendapatan total sebesar USD 383,29 miliar pada tahun fiskal 2023, dengan kontribusi utama dari penjualan iPhone, layanan digital, wearable, Mac, dan iPad [2]. Selain itu, Apple berhasil menguasai 46% dari total pendapatan pasar smartphone global pada tahun 2024, sekaligus mencetak harga jual rata-rata tertinggi sepanjang sejarah perusahaan [3]. Skala distribusi global yang mencakup puluhan negara dengan ragam saluran penjualan, segmen pelanggan, dan karakteristik transaksi yang kompleks menciptakan volume data yang sangat masif dan kaya, namun sekaligus menimbulkan tantangan besar dalam pengambilan keputusan bisnis yang berbasis data secara efektif dan efisien.

Data mining merupakan proses penemuan pola, korelasi, anomali, dan pengetahuan bermakna dari kumpulan data berukuran besar melalui penggunaan algoritma komputasional [4]. Salah satu teknik utama dalam data mining adalah clustering, yaitu metode pengelompokan data berdasarkan kesamaan karakteristik tanpa menggunakan label yang telah ditentukan sebelumnya (*unsupervised learning*) [4]. Dalam konteks bisnis, K-means clustering bekerja dengan cara meminimalkan variasi internal dalam setiap kluster sekaligus memaksimalkan keterpisahan antar kluster, menjadikannya pendekatan *unsupervised learning* yang banyak diadopsi karena kesederhanaan komputasi dan interpretasinya [5].

Di antara berbagai algoritma clustering yang tersedia, K-Means merupakan algoritma yang paling banyak digunakan karena kesederhanaannya, efisiensinya dalam menangani dataset berskala besar, serta kemampuannya menghasilkan segmen yang dapat diinterpretasikan secara bisnis. Tabianan et al. menerapkan K-Means Clustering untuk mensegmentasi pelanggan e-commerce berdasarkan perilaku pembelian, dan hasilnya menunjukkan bahwa segmentasi berbasis clustering tidak hanya mampu mengelompokkan pelanggan, tetapi juga dapat dikombinasikan dengan model klasifikasi untuk memprediksi perilaku pembelian dengan lebih akurat [6]. Ufeli et al. mengembangkan pendekatan Factor Analysis of Mixed Data (FAMD) yang dikombinasikan dengan K-Means dan Hierarchical Clustering untuk menangani data pelanggan bertipe campuran (numerik dan kategorikal), dan terbukti meningkatkan presisi segmentasi dibandingkan

metode konvensional [7]. Tang menerapkan kombinasi RFM analysis dan K-Means clustering pada data ritel untuk mengklasifikasikan pelanggan ke dalam segmen yang actionable seperti loyal customers, champions, dan at-risk, sekaligus memvalidasi efektivitas pendekatan tersebut untuk mendukung strategi pemasaran [8]. Ogunleye membandingkan beberapa algoritma clustering untuk segmentasi pelanggan di pasar *retail* dan menyimpulkan bahwa K-Means secara konsisten menghasilkan performa yang baik, dengan penerapan Elbow Method untuk estimasi jumlah kluster serta Silhouette Score sebagai ukuran kualitas kluster yang dihasilkan [9]. Wang menegaskan bahwa segmentasi pelanggan yang efektif dan terstruktur menjadi fondasi utama dari strategi pemasaran digital yang adaptif dan berbasis data di era transformasi digital [10].

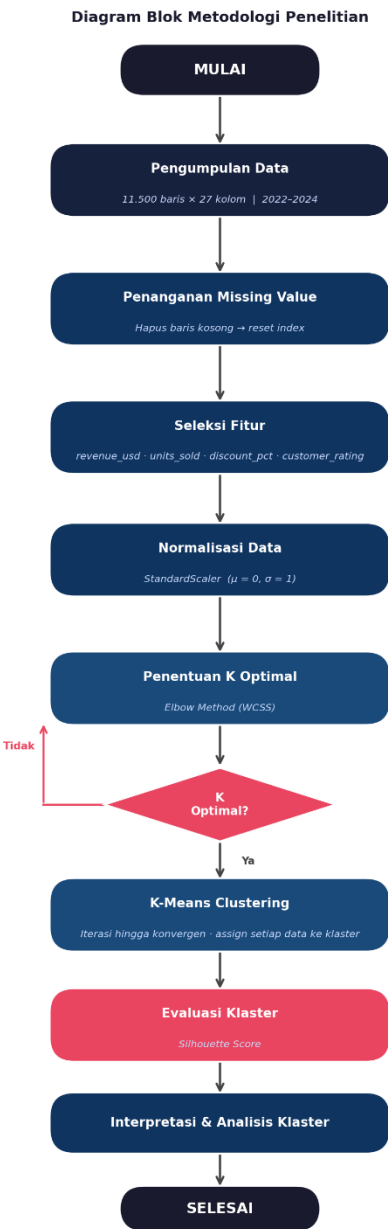
Meskipun berbagai penelitian terdahulu telah membuktikan efektivitas K-Means dalam segmentasi pelanggan dan pasar, penerapannya secara spesifik pada data transaksi penjualan global produk Apple yang bersifat multidimensional—mencakup karakteristik produk, cakupan geografis lintas negara, saluran penjualan, segmen pelanggan, nilai diskon, dan rating pelanggan—masih belum banyak dieksplorasi dalam literatur yang ada. Padahal, pemahaman terhadap pola segmentasi transaksi yang bervariasi di berbagai negara dan kategori produk dapat memberikan wawasan strategis yang signifikan bagi pengambilan keputusan di industri teknologi konsumen global [5][6]. Selain menghasilkan segmentasi transaksi, kualitas kluster yang terbentuk juga perlu dievaluasi agar hasil pengelompokan dapat dipertanggungjawabkan secara ilmiah. Oleh karena itu, penelitian ini menggunakan Elbow Method untuk menentukan jumlah kluster optimal serta Silhouette Score sebagai metrik evaluasi kualitas kluster. Evaluasi tersebut memberikan dasar kuantitatif dalam menilai tingkat kohesi dan separasi antar kluster sebelum dilakukan interpretasi terhadap karakteristik setiap segmen.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menerapkan algoritma K-Means Clustering dalam melakukan segmentasi terhadap data transaksi penjualan produk Apple secara global. Secara lebih spesifik, penelitian ini bermaksud: (1)

mengidentifikasi kelompok-kelompok transaksi yang memiliki kesamaan karakteristik berdasarkan variabel seperti *revenue*, jumlah unit terjual, persentase diskon, dan *customer rating*; (2) menentukan jumlah kluster optimal menggunakan Elbow Method dan Silhouette Score; serta (3) menginterpretasikan profil setiap kluster guna menghasilkan wawasan yang dapat mendukung strategi penjualan dan pemasaran produk Apple di pasar global.

II. METODOLOGI PENELITIAN

Penelitian ini dilakukan melalui serangkaian tahapan sistematis yang dirancang untuk menghasilkan segmentasi transaksi penjualan produk Apple yang valid dan dapat diinterpretasikan secara bisnis. Alur metodologi secara keseluruhan digambarkan pada diagram blok di atas, yang meliputi pengumpulan data, pra-pemrosesan, penerapan algoritma K-Means Clustering, hingga evaluasi kluster yang terbentuk.



Gambar 1. Diagram Blok Metodologi Penelitian

A. Dataset

Dataset yang digunakan dalam penelitian ini adalah *apple_global_sales_dataset*, sebuah dataset transaksi penjualan produk Apple secara global yang mencakup periode tahun 2022 hingga 2024. Dataset ini terdiri dari 11.500 baris data dan 27 kolom atribut yang merepresentasikan berbagai dimensi transaksi, mulai dari identitas transaksi, informasi produk, lokasi geografis penjualan (mencakup 47 negara dan 8 wilayah regional), saluran penjualan (*sales channel*), hingga profil pelanggan dan status pengembalian produk. Variabel numerik yang relevan untuk proses clustering antara lain

revenue_usd, *units_sold*, *discount_pct*, dan *customer_rating*, yang secara bersama-sama merepresentasikan karakteristik finansial dan perilaku dari setiap transaksi yang terjadi.

B. Pra-pemrosesan Data

Pra-pemrosesan data merupakan tahap krusial yang menentukan kualitas hasil clustering. Pada dataset ini, terdapat tiga kolom yang memiliki missing value, yaitu *storage* (4.804 nilai kosong), *previous_device_os* (8.056 nilai kosong), dan *customer_rating* (3.360 nilai kosong). Mengacu pada praktik umum dalam data mining [4], baris yang mengandung nilai kosong pada kolom fitur yang akan digunakan dalam proses clustering dihapus (*dropna*) dan indeks data di-reset untuk memastikan konsistensi dataset. Setelah proses pembersihan, dilanjutkan dengan seleksi fitur di mana hanya empat variabel numerik yang dipilih sebagai input clustering, yaitu *revenue_usd*, *units_sold*, *discount_pct*, dan *customer_rating*. Selanjutnya, seluruh fitur tersebut dinormalisasi menggunakan *StandardScaler* sehingga setiap variabel memiliki rata-rata nol ($\mu = 0$) dan standar deviasi satu ($\sigma = 1$). Normalisasi ini penting dilakukan karena algoritma K-Means bersifat sensitif terhadap perbedaan skala antar fitur dalam perhitungan jarak Euclidean [4][6].

C. Penerapan Algoritma K-Means Clustering

K-Means adalah algoritma clustering berbasis partisi yang bekerja dengan cara meminimalkan jumlah kuadrat jarak (*Within-Cluster Sum of Squares/WCSS*) antara setiap titik data terhadap centroid klasternya [4]. Algoritma ini bekerja secara iteratif melalui tahap: (1) inisialisasi centroid secara acak sebanyak K, (2) penugasan setiap titik data ke centroid terdekat berdasarkan jarak Euclidean, (3) pembaruan posisi centroid sebagai rata-rata seluruh anggota klaster, dan (4) pengulangan langkah 2–3 hingga tidak ada perubahan penugasan (*konvergen*) [7]. Sebelum proses clustering dijalankan, terlebih dahulu dilakukan penentuan nilai K optimal menggunakan *Elbow Method*, yaitu dengan memplot nilai WCSS terhadap berbagai nilai K dan mengidentifikasi titik "siku" di mana penurunan WCSS mulai tidak signifikan [7][9].

D. Evaluasi Klaster

Untuk mengukur kualitas klaster yang dihasilkan, penelitian ini menggunakan *Silhouette Score* sebagai metrik evaluasi. *Silhouette Score* mengukur

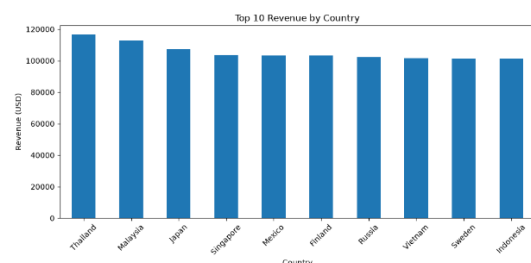
seberapa mirip suatu titik data terhadap klasternya sendiri dibandingkan dengan klaster lain, dengan nilai berkisar antara -1 hingga $+1$. Nilai yang mendekati $+1$ mengindikasikan bahwa titik data tersebut ditempatkan pada klaster yang tepat, nilai mendekati 0 mengindikasikan posisi di antara dua klaster, dan nilai negatif mengindikasikan kemungkinan penugasan yang keliru [9]. Hasil evaluasi ini digunakan sebagai konfirmasi kuantitatif atas jumlah klaster optimal yang telah dipilih melalui *Elbow Method*.

III. HASIL DAN PEMBAHASAN

Hasil penelitian

A. Hasil Pra-pemrosesan Data

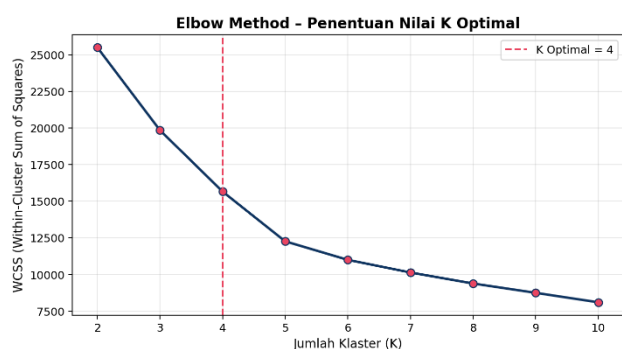
Proses pra-pemrosesan dilakukan dengan menghapus baris yang memiliki nilai kosong hanya pada kolom fitur clustering, yaitu *customer_rating*, sehingga dataset yang semula berjumlah 11.500 baris berkurang menjadi 8.140 baris yang valid. Berbeda dengan pendekatan *dropna* menyeluruh, strategi ini memastikan seluruh enam kategori produk Apple tetap terwakili, yaitu iPhone, Mac, iPad, Apple Watch, AirPods, dan Accessories. Setelah seleksi fitur dan normalisasi menggunakan *StandardScaler*, keempat fitur clustering (*revenue_usd*, *units_sold*, *discount_pct*, dan *customer_rating*) telah memiliki skala yang setara dan siap diproses oleh algoritma K-Means. Sebagai gambaran distribusi pendapatan awal sebelum clustering, Gambar 2 menunjukkan sepuluh negara dengan total pendapatan tertinggi. Gambar 2 memperlihatkan bahwa Thailand dan Malaysia berada di posisi teratas dengan pendapatan masing-masing melampaui USD 110.000, diikuti oleh Jepang, Singapura, dan Meksiko, yang mengindikasikan konsentrasi transaksi bernilai tinggi di kawasan Asia Tenggara dan Asia Timur.



Gambar 2. Top 10 Negara Berdasarkan Total Revenue

B. Penentuan Jumlah Kluster Optimal

Penentuan nilai K optimal dilakukan menggunakan Elbow Method dengan memplot nilai WCSS terhadap rentang K = 2 hingga K = 10, sebagaimana ditunjukkan pada Gambar 3. Gambar 3 memperlihatkan bahwa penurunan nilai WCSS yang signifikan terjadi dari K = 2 hingga K = 4, setelah itu kurva mulai melandai dan tidak memberikan pengurangan yang berarti. Berdasarkan titik "siku" yang teridentifikasi pada K = 4, nilai tersebut ditetapkan sebagai jumlah kluster optimal dalam penelitian ini.



Gambar 3. Elbow Method — Penentuan Nilai K Optimal

C. Hasil K-Means Clustering

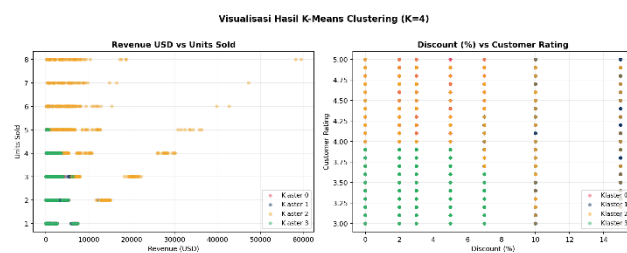
Dengan K = 4, algoritma K-Means berhasil mengelompokkan 8.140 transaksi ke dalam empat kluster yang berbeda. Profil rata-rata setiap kluster berdasarkan keempat fitur clustering disajikan pada Tabel 1.

Tabel 1. Profil Rata-rata Setiap Kluster Hasil K-Means Clustering

Kluster	Karakteristik Transaksi			Distribusi Data		
	Revenue (USD)	Units Sold	Diskon (%)	Rating	Jumlah	Presentase
0	1107	1,63	1,72	4,49	2987	36,70%
1	1002	1,76	11,65	4,01	1630	20,02%
2	6203	5,23	2,95	4,01	820	10,07%
3	1046	1,65	1,69	3,46	2730	33,21%

Tabel 1 menunjukkan bahwa Kluster 0 merupakan kelompok dengan anggota terbanyak (36,70%), sementara Kluster 2 merupakan kelompok paling kecil namun memiliki nilai revenue dan volume pembelian tertinggi. Secara visual, distribusi keempat kluster terhadap dua pasang fitur utama dapat dilihat pada Gambar 4. Gambar 4 memperlihatkan bahwa Kluster 2 (oranye) terpisah secara jelas dari kluster lainnya pada dimensi revenue dan units sold, sementara Kluster 0, 1, dan 3

menempati rentang yang berdekatan namun dapat dibedakan melalui dimensi discount dan rating.



(Gambar 4. Scatter Plot Hasil K-Means Clustering)

D. Evaluasi Kluster

Validasi kualitas kluster menggunakan Silhouette Score menghasilkan nilai sebesar 0,3120. Nilai tersebut menunjukkan bahwa hasil clustering memiliki kualitas pemisahan kluster pada kategori moderat, sehingga masih terdapat beberapa area dengan kedekatan antar kluster. Meskipun demikian, struktur kluster yang terbentuk masih cukup representatif untuk tujuan analisis eksploratif dan segmentasi transaksi. Penelitian selanjutnya dapat membandingkan jumlah kluster yang berbeda maupun menggunakan algoritma clustering alternatif, seperti DBSCAN atau Gaussian Mixture Model, untuk memperoleh kualitas segmentasi yang lebih optimal.[9]. Meskipun pemisahan antar kluster tidak bersifat sempurna — yang wajar terjadi pada data transaksi riil dengan distribusi yang saling tumpang tindih — nilai Silhouette Score tersebut mengkonfirmasi bahwa keempat kluster memiliki kohesi internal yang memadai dan dapat diinterpretasikan secara bermakna.

E. Pembahasan

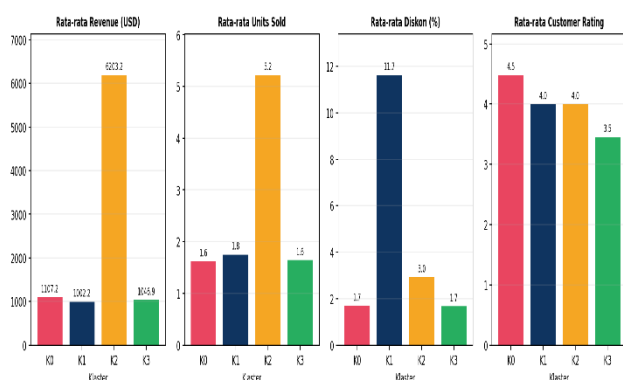
Hasil clustering mengungkap empat segmen transaksi yang secara strategis berbeda satu sama lain. Kluster 2, meskipun hanya mencakup 10,07% transaksi, menghasilkan rata-rata revenue enam kali lipat lebih tinggi dibandingkan kluster lainnya dengan volume pembelian terbesar (rata-rata 5,23 unit per transaksi). Pola ini mengindikasikan adanya segmen transaksi korporat atau pembelian massal (bulk). Kluster 2 memiliki rata-rata revenue dan jumlah unit terjual paling tinggi dibandingkan kluster lainnya sehingga dapat dikategorikan sebagai kelompok transaksi bernilai tinggi. Karakteristik tersebut menunjukkan adanya transaksi dengan skala pembelian yang lebih besar dibandingkan kluster lain. Namun, karena dataset tidak memuat informasi

mengenai jenis pelanggan, hasil ini tidak dapat digunakan untuk menyimpulkan bahwa transaksi tersebut berasal dari pelanggan korporat atau B2B., yang menjadi kontributor nilai penjualan paling signifikan meski secara jumlah transaksi kecil. Temuan ini sejalan dengan prinsip distribusi Pareto dalam analisis penjualan [5].

Klaster 1 memperlihatkan karakteristik yang berbeda secara diametral: rata-rata diskon tertinggi (11,65%) dengan revenue terendah. Segmen ini mencerminkan kelompok pembeli yang sangat responsif terhadap insentif harga, sehingga strategi promosi berbasis diskon lebih efektif diterapkan pada segmen ini dibandingkan segmen lainnya [6][10].

Yang paling menarik perhatian adalah perbedaan tajam antara Klaster 0 dan Klaster 3. Keduanya memiliki profil revenue, volume, dan diskon yang hampir identik, namun berbeda secara signifikan pada dimensi *customer rating* — Klaster 0 mencatat rata-rata 4,49 sementara Klaster 3 hanya 3,46. Perbedaan *customer rating* antara Klaster 0 dan Klaster 3 menunjukkan adanya variasi tingkat kepuasan pelanggan meskipun kedua klaster memiliki karakteristik revenue, jumlah unit terjual, dan diskon yang relatif serupa. Namun demikian, penelitian ini tidak dapat mengidentifikasi faktor penyebab perbedaan rating tersebut karena dataset tidak memuat atribut seperti kualitas layanan purna jual, waktu pengiriman, maupun pengalaman pelanggan. Oleh sebab itu, faktor-faktor tersebut dapat menjadi variabel tambahan pada penelitian berikutnya.. Visualisasi perbandingan rata-rata keempat fitur antar klaster secara ringkas dapat dilihat pada Gambar 5.

Profil Rata-rata Setiap Klaster



(Gambar 5. Profil Rata-rata Setiap Klaster)

IV. KESIMPULAN

Penelitian ini berhasil mencapai seluruh tujuan yang telah ditetapkan. Algoritma K-Means Clustering terbukti dapat diterapkan secara efektif untuk melakukan segmentasi terhadap data transaksi penjualan produk Apple secara global, menghasilkan empat klaster yang bermakna dan dapat dibedakan secara karakteristik. Penentuan jumlah klaster optimal melalui Elbow Method mengidentifikasi $K = 4$ sebagai nilai yang paling representatif, dan validasi menggunakan Silhouette Score mengkonfirmasi bahwa struktur klaster yang terbentuk memiliki kohesi internal yang memadai untuk diinterpretasikan secara bisnis.

Keempat klaster yang dihasilkan masing-masing merepresentasikan segmen transaksi yang berbeda secara strategis: segmen transaksi bernilai tinggi dengan volume pembelian besar, segmen yang responsif terhadap insentif diskon, segmen dengan tingkat kepuasan pelanggan tinggi meski tanpa diskon signifikan, serta segmen dengan profil transaksi serupa namun kepuasan pelanggan yang rendah. Temuan terpenting dari penelitian ini adalah bahwa perbedaan kepuasan pelanggan tidak selalu berkorelasi dengan nilai transaksi atau besaran diskon yang diberikan. Hasil penelitian menunjukkan bahwa *customer rating* tidak selalu berkorelasi langsung dengan nilai transaksi maupun besaran diskon. Namun demikian, penelitian ini belum dapat mengidentifikasi faktor penyebab variasi *customer rating* karena variabel tersebut tidak tersedia dalam dataset yang digunakan. Penelitian selanjutnya dapat menambahkan variabel yang berkaitan dengan kualitas layanan, pengiriman, maupun karakteristik pelanggan untuk memperoleh interpretasi yang lebih komprehensif.

Dari sisi penerapan, hasil segmentasi ini dapat dimanfaatkan oleh pemangku kepentingan di industri teknologi konsumen sebagai dasar perancangan strategi pemasaran yang lebih tertarget. Segmen transaksi bernilai tinggi dapat diprioritaskan untuk program loyalitas korporat, sementara segmen sensitif harga dapat didekati dengan program promosi yang terstruktur. Di sisi lain, segmen dengan kepuasan rendah memerlukan intervensi pada dimensi layanan, bukan pada dimensi harga,

sehingga alokasi sumber daya dapat dilakukan secara lebih efisien.

Untuk pengembangan penelitian selanjutnya, disarankan agar analisis diperluas dengan mengintegrasikan variabel kategorikal seperti saluran penjualan dan segmen pelanggan ke dalam proses clustering melalui teknik encoding yang sesuai, atau dengan mengeksplorasi algoritma clustering alternatif seperti DBSCAN atau Gaussian Mixture Model yang mampu mendeteksi klaster dengan bentuk yang lebih kompleks. Perluasan cakupan temporal dan geografis pada dataset yang lebih baru juga akan meningkatkan relevansi dan generalisasi temuan penelitian ini.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Ibu Aprilisa Arum Sari, S.T., M.Kom. selaku dosen pembimbing sekaligus dosen pengampu mata kuliah Data Mining yang telah memberikan arahan, bimbingan, dan masukan yang sangat berarti selama proses penelitian dan penulisan paper ini berlangsung. Penulis juga mengucapkan terima kasih kepada Universitas Duta Bangsa Surakarta atas dukungan akademik dan fasilitas yang telah diberikan sehingga penelitian ini dapat terlaksana dengan baik.

REFERENSI

- [1] G. S. Shipments, A. Dethroned, and S. After, "Global Smartphone Shipments 2023 : Apple Dethroned Samsung Smartphone Shipments 2023 : Main Highlights," pp. 2023–2025, 2024.
- [2] M. One, "SECURITIES AND EXCHANGE COMMISSION," vol. 12, no. d, 2023.
- [3] A. M. Global, S. Revenues, R. Growth, A. T. Years, A. S. P. Hits, and R. High, "Global Smartphone Revenues Resume Growth in 2024 After Two Years , ASP Hits Record High," pp. 24–27, 2025.
- [4] M. M. Degree, C. Science, and A. C. Lecture, "Data Mining : Concepts and," 2012.
- [5] K. P. Sinaga and M. Yang, "Unsupervised K-Means Clustering Algorithm," vol. 8, 2020, doi: 10.1109/ACCESS.2020.2988796.
- [6] K. Tabianan, S. Velu, and V. Ravi, "K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data Segmentation Using Customer Purchase Behavior Data".
- [7] C. P. Ufeli, M. U. Sattar, R. Hasan, and S. Mahmood, "Enhancing Customer Segmentation Through Factor Analysis of Mixed Data (FAMD) -Based Approach Using K-Means and Hierarchical Clustering Algorithms".
- [8] J. Tang and J. Tang, "Unlocking Retail Insights : Predictive Modeling and Customer Segmentation Through Data Analytics Segmentation Through Data Analytics".
- [9] J. M. John, O. Shobayo, B. Ogunleye, J. M. John, and O. Shobayo, "An Exploration of Clustering Algorithms for Customer

Segmentation in the UK Retail Market An Exploration of Clustering Algorithms for Customer Segmentation in the UK Retail Market".

[10]

G. W. Id, "Customer segmentation in the digital marketing using a Q -learning based differential evolution algorithm integrated with K -means clustering," pp. 1–21, 2025, doi: 10.1371/journal.pone.0318519.