

Segmentasi Pelanggan Menggunakan Algoritma K-Means untuk Customer Intelligence dan Strategi Pemasaran pada Perusahaan Retail Online (Studi Kasus: Dataset Online Retail II)

¹Muhammad Naufal Skha Yusfa, ²Ahyat Musyawwa, ³Aldy Bifal Pratama

¹²³Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

¹muhammadnaufal607@gmail.com, ²ahyatmusyawwa123@gmail.com, ³aldysape@gmail.com

Abstract

Customer intelligence is crucial for online retail companies to design targeted marketing strategies and increase customer retention. This study aims to segment customers based on purchasing behavior using RFM analysis and K-Means clustering. The Online Retail II dataset containing more than 525,000 transactions from December 2010 to December 2011 was used. After data cleaning, outlier handling, and RFM feature engineering, the Elbow Method and Silhouette Score determined the optimal number of clusters ($K=3$). The results produced three distinct customer segments: Lost/At Risk (25.17%), Regular Customers (74.32%), and High-Value Loyal (0.51%). Actionable business recommendations were formulated for each segment. This segmentation provides deeper customer intelligence and supports more effective marketing strategies for online retail businesses.

Keywords: Customer Segmentation, RFM analysis, K-Means clustering, Customer Intelligence, Online retail

Abstrak

Customer intelligence sangat penting bagi perusahaan retail online untuk merancang strategi pemasaran yang tepat sasaran dan meningkatkan retensi pelanggan. Penelitian ini bertujuan untuk melakukan segmentasi pelanggan berdasarkan perilaku belanja menggunakan analisis RFM dan algoritma K-Means. Dataset Online Retail II yang berisi lebih dari 525.000 transaksi penjualan online periode Desember 2010 hingga Desember 2011 digunakan dalam penelitian ini. Setelah proses pembersihan data, penanganan outlier, dan rekayasa fitur RFM, metode Elbow dan Silhouette Score menentukan jumlah cluster optimal ($K=3$). Hasil penelitian menghasilkan tiga segmen pelanggan yang berbeda karakteristiknya: Lost/At Risk (25,17%), Regular Customers (74,32%), dan High-Value Loyal (0,51%). Rekomendasi strategi bisnis yang actionable disusun untuk setiap segmen. Segmentasi ini memberikan customer intelligence yang lebih dalam dan mendukung strategi pemasaran yang lebih efektif bagi perusahaan retail online.

Kata Kunci: Segmentasi Pelanggan, Analisis RFM, K-Means Clustering, Intelijen Pelanggan, Retail Online

A. PENDAHULUAN

Kemajuan teknologi digital telah mengubah lanskap bisnis retail di Indonesia menjadi semakin kompetitif. Pertumbuhan e-commerce yang pesat membuat perusahaan retail online harus bersaing tidak hanya dari segi harga dan produk, melainkan juga dari kemampuan memahami dan melayani pelanggan secara personal. Customer intelligence menjadi salah satu faktor penentu keberhasilan strategi pemasaran di era ini. Tanpa pemahaman yang mendalam terhadap perilaku pelanggan, perusahaan retail online berisiko kehilangan peluang untuk meningkatkan retensi dan optimalisasi pendapatan.

Di tengah persaingan yang ketat, banyak perusahaan retail online masih mengandalkan pendekatan pemasaran massal yang kurang tepat sasaran. Hal ini disebabkan oleh kurangnya analisis mendalam terhadap data transaksi

historis. Padahal, data transaksi yang besar dapat dimanfaatkan untuk memahami pola belanja pelanggan secara lebih akurat. Beberapa penelitian sebelumnya telah menunjukkan bahwa segmentasi pelanggan berbasis Recency, Frequency, Monetary (RFM) yang dikombinasikan dengan algoritma K-Means clustering mampu memberikan wawasan yang signifikan dalam mendukung strategi pemasaran (Fauzan & Alfian, 2024; Triantara et al., 2025; Yusak et al., 2024).

Meskipun banyak kajian mengenai segmentasi pelanggan menggunakan RFM dan K-Means telah dilakukan, sebagian besar penelitian tersebut masih berfokus pada data transaksi ritel skala kecil atau data simulasi. Belum banyak penelitian yang secara khusus menganalisis dataset transaksi retail online berskala besar seperti Online Retail II yang berisi lebih dari 525.000 catatan transaksi. Selain itu, jarang ditemukan penelitian yang tidak hanya

menghasilkan segmentasi, tetapi juga menerjemahkan hasil clustering tersebut menjadi rekomendasi strategi bisnis yang actionable dan siap diimplementasikan oleh perusahaan retail online di Indonesia.

Celah tersebut menjadi dasar penelitian ini. Penelitian ini difokuskan pada penerapan analisis RFM dan algoritma K-Means clustering terhadap Dataset Online Retail II untuk melakukan segmentasi pelanggan. Penelitian ini juga bertujuan untuk mengevaluasi kualitas segmentasi menggunakan Elbow Method dan Silhouette Score, serta menyusun rekomendasi strategi bisnis yang konkret berdasarkan karakteristik setiap segmen.

Tujuan penelitian ini adalah menghasilkan model segmentasi pelanggan yang dapat meningkatkan customer intelligence serta mendukung penyusunan strategi pemasaran yang lebih efektif. Diharapkan hasil penelitian ini dapat memberikan kontribusi praktis bagi perusahaan retail online dalam memahami perilaku pelanggan dan mengoptimalkan keputusan bisnis berbasis data.

B. METODE

Penelitian ini mengikuti kerangka kerja CRISP-DM (Cross Industry Standard Process for Data Mining) yang terdiri dari enam tahap utama. Pendekatan ini dipilih karena sangat sesuai untuk penelitian yang berorientasi pada penerapan data mining dalam mendukung pengambilan keputusan bisnis, khususnya segmentasi pelanggan menggunakan analisis RFM dan algoritma K-Means clustering.

1. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini adalah Online Retail II yang berisi lebih dari 525.000 catatan transaksi penjualan online selama periode Desember 2010 hingga Desember 2011. Dataset ini dipilih karena memiliki skala yang besar, berasal dari retail online nyata, serta mencakup berbagai jenis produk dan pelanggan, sehingga sangat relevan untuk menganalisis perilaku belanja di perusahaan retail online.

2. Prapemrosesan Data

Pada tahap ini dilakukan pembersihan data yang meliputi penghapusan data duplikat, penanganan missing value pada kolom CustomerID, serta filtering hanya transaksi yang valid (Quantity > 0 dan UnitPrice > 0). Selain itu, outlier pada variabel Quantity dan UnitPrice ditangani dengan metode capping pada rentang persentil 1% hingga 99% untuk mengurangi pengaruh nilai ekstrem tanpa menghapus terlalu banyak data (Febrina et al., 2025).

3. Rekayasa Fitur (Feature Engineering)

Dilakukan pembentukan fitur RFM (Recency, Frequency, Monetary) yang menjadi dasar utama segmentasi. Recency dihitung sebagai selisih hari antara tanggal transaksi terakhir dengan tanggal referensi, Frequency merupakan jumlah transaksi yang dilakukan oleh pelanggan, sedangkan Monetary adalah total nilai pembelian (monetary value) selama periode observasi. Rekayasa fitur

ini sangat penting karena mampu merangkum perilaku pelanggan secara efektif (Fauzan & Alfian, 2024).

4. Normalisasi Data

Karena ketiga variabel RFM memiliki skala yang berbeda, data dinormalisasi menggunakan StandardScaler. Proses normalisasi ini bertujuan untuk memastikan semua fitur memiliki bobot yang sama sehingga algoritma K-Means dapat bekerja secara optimal dan tidak didominasi oleh variabel dengan skala besar.

5. Penentuan Jumlah Cluster Optimal

Jumlah cluster (K) ditentukan dengan menggunakan dua metode evaluasi, yaitu Elbow Method (untuk melihat penurunan nilai inertia) dan Silhouette Score (untuk mengukur kualitas pemisahan antar cluster). Berdasarkan hasil kedua metode tersebut, dipilih K=3 sebagai jumlah cluster yang paling optimal, baik dari segi teknis maupun interpretasi bisnis (Achmad et al., 2024; Hairani et al., 2026).

6. Pemodelan dan Evaluasi

Algoritma K-Means Clustering diterapkan dengan pengaturan random state tetap untuk menjamin hasil yang reproducible. Setelah proses clustering selesai, dilakukan evaluasi model melalui analisis profil setiap cluster berdasarkan nilai rata-rata RFM, visualisasi scatter plot (Recency vs Monetary), serta perbandingan distribusi pelanggan. Evaluasi ini juga mencakup interpretasi karakteristik masing-masing segmen untuk menyusun rekomendasi strategi pemasaran yang actionable.

C. HASIL DAN PEMBAHASAN

Hasil dan pembahasan penelitian ini disajikan berdasarkan proses pemodelan yang telah dilakukan menggunakan algoritma K-Means setelah tahap persiapan data dan rekayasa fitur RFM selesai. Bagian ini menyajikan hasil eksperimen clustering, evaluasi kualitas model, interpretasi karakteristik setiap segmen pelanggan, serta rekomendasi strategi bisnis yang actionable. Semua temuan dianalisis secara mendalam untuk memberikan pemahaman yang komprehensif mengenai pola perilaku pelanggan pada Dataset Online Retail II.

1. Profil Setiap Cluster (K=3)

Tabel 1. Profil Setiap Cluster Hasil K-Means (K=3)

	Recency	Frequency	Monetary	Jumlah Customer	Persentase (%)
Cluster 0	41.1	4.71	1762.11	3224.0	74.32
Cluster 1	246.12	1.58	455.82	1092.0	25.17
Cluster 2	6.55	73.05	66785.99	22.0	0.51

Berdasarkan Tabel 1, diperoleh tiga kelompok pelanggan dengan karakteristik yang berbeda. Cluster 0 merupakan kelompok pelanggan terbesar dengan jumlah 3.224 pelanggan atau 74,32% dari total pelanggan. Cluster ini memiliki nilai Recency sebesar 41,1 hari, Frequency sebesar 4,71 transaksi, dan Monetary sebesar £1.762,11, sehingga dapat dikategorikan sebagai pelanggan reguler

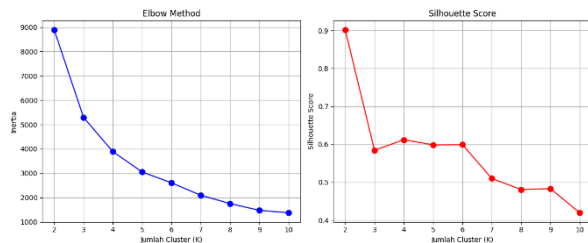
yang masih aktif bertransaksi dengan nilai pembelian menengah (Regular Customers).

Cluster 1 terdiri dari 1.092 pelanggan atau 25,17% dari total pelanggan. Kelompok ini memiliki nilai Recency tertinggi yaitu 246,12 hari, serta nilai Frequency (1,58) dan Monetary (£455,82) yang paling rendah dibandingkan cluster lainnya. Karakteristik tersebut menunjukkan bahwa pelanggan dalam cluster ini sudah lama tidak melakukan transaksi dan memiliki tingkat aktivitas pembelian yang rendah, sehingga dapat dikategorikan sebagai pelanggan yang berisiko churn (Lost/At Risk Customers).

Sementara itu, Cluster 2 merupakan kelompok pelanggan dengan jumlah paling sedikit, yaitu 22 pelanggan atau 0,51% dari total pelanggan. Meskipun jumlahnya sangat kecil, cluster ini memiliki nilai Frequency tertinggi sebesar 73,05 transaksi dan nilai Monetary sebesar £66.785,99, jauh di atas cluster lainnya. Selain itu, nilai Recency yang rendah (6,55 hari) menunjukkan bahwa pelanggan pada kelompok ini masih aktif melakukan pembelian. Oleh karena itu, Cluster 2 dapat dikategorikan sebagai pelanggan loyal bernilai tinggi (High-Value Loyal Customers) yang memberikan kontribusi pendapatan terbesar bagi perusahaan dan perlu dipertahankan melalui program loyalitas khusus.

Hasil segmentasi ini menunjukkan bahwa metode RFM dan algoritma K-Means berhasil mengelompokkan pelanggan berdasarkan perilaku transaksi mereka sehingga dapat menjadi dasar dalam penyusunan strategi pemasaran yang lebih tepat sasaran untuk setiap segmen pelanggan.

2. Elbow Method dan Silhouette Score



Gambar 1. Hasil Evaluasi Jumlah Cluster Menggunakan Elbow Method dan Silhouette Score

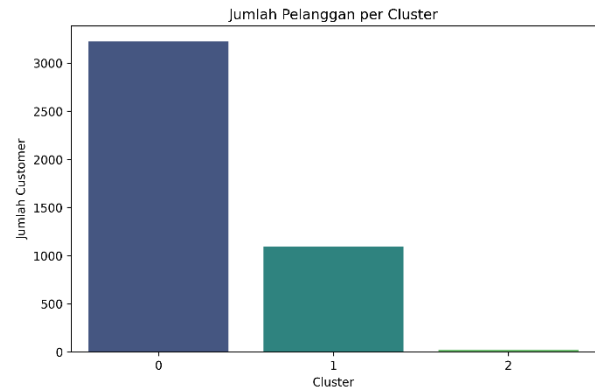
Penentuan jumlah cluster optimal dilakukan menggunakan kombinasi Elbow Method dan Silhouette Score sebagaimana ditunjukkan pada Gambar 1. Pada grafik Elbow Method terlihat bahwa nilai inertia mengalami penurunan yang sangat signifikan dari K=2 hingga K=5, kemudian cenderung melandai setelahnya. Pola tersebut menunjukkan adanya titik siku (elbow point) pada rentang K=3 hingga K=5 yang menandakan bahwa penambahan cluster setelah titik tersebut tidak lagi memberikan peningkatan kualitas yang signifikan.

Sementara itu, berdasarkan grafik Silhouette Score, nilai tertinggi diperoleh pada K=2 dengan skor sekitar 0,90. Meskipun demikian, penggunaan dua cluster dinilai kurang mampu memberikan segmentasi pelanggan yang detail

untuk kebutuhan customer intelligence dan penyusunan strategi pemasaran. Pada K=3 diperoleh nilai Silhouette Score sekitar 0,58 yang masih menunjukkan kualitas clustering yang baik dengan tingkat pemisahan antar cluster yang cukup jelas.

Selain mempertimbangkan kualitas statistik model, pemilihan K=3 juga memberikan interpretasi bisnis yang lebih bermakna karena mampu membedakan pelanggan ke dalam tiga kelompok utama, yaitu pelanggan berisiko churn (Lost/At Risk), pelanggan reguler (Regular Customers), dan pelanggan loyal bernilai tinggi (High-Value Loyal Customers). Oleh karena itu, K=3 dipilih sebagai jumlah cluster optimal karena mampu memberikan keseimbangan antara kualitas clustering dan kemudahan interpretasi hasil untuk kebutuhan pengambilan keputusan bisnis.

3. Jumlah Pelanggan per Cluster



Gambar 2. Distribusi Jumlah Pelanggan pada Setiap Cluster Hasil K-Means

Gambar 2. Menunjukkan distribusi jumlah pelanggan pada masing-masing cluster hasil segmentasi menggunakan algoritma K-Means. Terlihat bahwa Cluster 0 merupakan kelompok pelanggan terbesar dengan jumlah 3.224 pelanggan atau sekitar 74,32% dari total populasi pelanggan. Dominasi jumlah pelanggan pada cluster ini menunjukkan bahwa sebagian besar pelanggan memiliki pola transaksi yang relatif stabil dengan tingkat frekuensi pembelian dan nilai transaksi pada kategori menengah. Oleh karena itu, cluster ini dapat dianggap sebagai basis pelanggan utama yang berkontribusi secara konsisten terhadap aktivitas bisnis perusahaan.

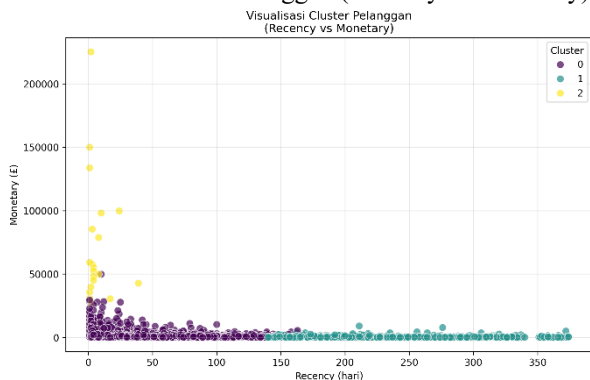
Cluster 1 menempati posisi kedua dengan jumlah 1.092 pelanggan atau sekitar 25,17% dari total pelanggan. Meskipun jumlahnya cukup besar, pelanggan pada cluster ini memiliki karakteristik aktivitas transaksi yang rendah dan waktu transaksi terakhir yang relatif lama. Kondisi tersebut mengindikasikan bahwa sebagian pelanggan dalam kelompok ini berpotensi berhenti bertransaksi sehingga memerlukan strategi retensi dan reaktivasi yang lebih intensif.

Sementara itu, Cluster 2 hanya terdiri dari 22 pelanggan atau sekitar 0,51% dari total populasi. Meskipun jumlahnya

sangat kecil, kelompok ini memiliki nilai Frequency dan Monetary yang jauh lebih tinggi dibandingkan cluster lainnya. Hal ini menunjukkan bahwa pelanggan pada Cluster 2 merupakan pelanggan loyal bernilai tinggi (High-Value Loyal Customers) yang memberikan kontribusi pendapatan terbesar bagi perusahaan. Oleh karena itu, perusahaan perlu memberikan perhatian khusus kepada segmen ini melalui program loyalitas eksklusif, layanan premium, maupun penawaran personal guna mempertahankan hubungan jangka panjang dengan pelanggan.

Distribusi pelanggan yang tidak seimbang antar cluster merupakan kondisi yang umum ditemukan pada data transaksi retail. Hasil ini menunjukkan bahwa sebagian besar pelanggan berada pada kategori pelanggan reguler, sementara hanya sebagian kecil pelanggan yang berkontribusi secara signifikan terhadap total pendapatan perusahaan. Temuan ini menjadi dasar penting dalam penyusunan strategi pemasaran yang lebih efektif dan tepat sasaran untuk setiap segmen pelanggan.

4. Visualisasi Cluster Pelanggan (Recency vs Monetary)



Gambar 3. Visualisasi Cluster Pelanggan (Recency vs Monetary)

Gambar 3. Menunjukkan hasil segmentasi pelanggan berdasarkan hubungan antara variabel Recency (waktu sejak transaksi terakhir) dan Monetary (total nilai pembelian). Setiap titik pada scatter plot merepresentasikan satu pelanggan, sedangkan warna yang berbeda menunjukkan cluster hasil pengelompokan menggunakan algoritma K-Means.

Berdasarkan visualisasi tersebut, terlihat bahwa Cluster 0 merupakan kelompok pelanggan dengan persebaran paling dominan. Pelanggan pada cluster ini umumnya memiliki nilai Recency yang relatif rendah hingga menengah serta nilai Monetary yang berada pada tingkat menengah. Karakteristik tersebut menunjukkan bahwa pelanggan dalam kelompok ini masih cukup aktif melakukan transaksi dan memberikan kontribusi pendapatan yang stabil bagi perusahaan. Oleh karena itu, Cluster 0 dapat dikategorikan sebagai Regular Customers.

Selanjutnya, Cluster 1 terkonsentrasi pada area dengan nilai Recency yang tinggi dan nilai Monetary yang rendah.

Posisi ini mengindikasikan bahwa pelanggan dalam cluster tersebut sudah cukup lama tidak melakukan transaksi serta memiliki total pembelian yang relatif kecil. Karakteristik ini menunjukkan adanya risiko pelanggan berhenti berbelanja atau berpindah ke kompetitor. Dengan demikian, Cluster 1 dapat diklasifikasikan sebagai Lost/At Risk Customers yang memerlukan strategi reaktivasi dan retensi pelanggan.

Sementara itu, Cluster 2 terlihat terpisah secara jelas dari cluster lainnya. Kelompok ini berada pada area dengan nilai Recency yang sangat rendah dan nilai Monetary yang sangat tinggi. Artinya, pelanggan pada cluster ini masih aktif bertransaksi dan memiliki kontribusi pembelian yang jauh lebih besar dibandingkan pelanggan lainnya. Meskipun jumlah pelanggan dalam cluster ini sangat sedikit, kontribusinya terhadap pendapatan perusahaan sangat signifikan. Oleh karena itu, Cluster 2 dikategorikan sebagai High-Value Loyal Customers yang perlu dipertahankan melalui program loyalitas, layanan eksklusif, serta penawaran yang lebih personal.

Secara keseluruhan, visualisasi ini menunjukkan bahwa algoritma K-Means berhasil memisahkan pelanggan ke dalam kelompok-kelompok yang memiliki karakteristik transaksi berbeda. Pemisahan cluster yang cukup jelas membuktikan bahwa variabel RFM, khususnya Recency dan Monetary, mampu memberikan gambaran yang efektif mengenai perilaku pelanggan dan dapat digunakan sebagai dasar dalam penyusunan strategi pemasaran yang lebih tepat sasaran.

5. Action Plan & Rekomendasi Strategis Bisnis

Tabel 2. Action Plan dan Rekomendasi Strategis Bisnis

Cluster	Segment	Recency(hari)	Frequency	Monetary (€)	Rekomendasi Strategis
0	Regular Customer	41.1	4.71	1762.11	Program Loyalty Points + Bundle produk + Promo frekuensi belanja
1	Lost / at Risk	246.12	1.58	455.82	Win-back Campaign (diskon 25-35% + email personal) + Survey alasan churn
2	High-Value Loyal	6.55	73.05	66785.99	VIP Program, Early Access, Personal Offer, Corporate Gifting

Pada Cluster 0 (Regular Customers) yang merupakan kelompok terbesar dengan 3.224 pelanggan atau sekitar 74,32% dari total pelanggan, nilai Recency sebesar 41,1 hari, Frequency sebesar 4,71 transaksi, dan Monetary sebesar £1.762,11 menunjukkan bahwa pelanggan masih aktif bertransaksi dengan kontribusi pendapatan yang stabil. Untuk mempertahankan dan meningkatkan loyalitas mereka, strategi yang direkomendasikan adalah Program Loyalty Points, bundling produk, serta promosi berbasis frekuensi belanja. Strategi ini bertujuan mendorong peningkatan frekuensi transaksi dan nilai pembelian pelanggan secara bertahap.

Pada Cluster 1 (Lost/At Risk Customers) yang terdiri dari 1.092 pelanggan atau sekitar 25,17% dari total pelanggan, nilai Recency yang tinggi sebesar 246,12 hari serta nilai Frequency (1,58) dan Monetary (£455,82) yang rendah mengindikasikan bahwa pelanggan sudah lama tidak melakukan transaksi dan berpotensi berhenti berbelanja. Oleh karena itu, strategi yang direkomendasikan adalah Win-back Campaign berupa diskon 25–35% disertai email personal, serta survei untuk mengetahui alasan pelanggan

berhenti bertransaksi. Strategi ini bertujuan mengaktifkan kembali pelanggan dan menekan tingkat churn.

Sementara itu, Cluster 2 (High-Value Loyal Customers) hanya terdiri dari 22 pelanggan atau sekitar 0,51% dari total pelanggan. Meskipun jumlahnya sangat kecil, kelompok ini memiliki nilai Frequency tertinggi sebesar 73,05 transaksi dan Monetary sebesar £66.785,99, jauh melampaui cluster lainnya, dengan nilai Recency yang sangat rendah (6,55 hari) yang menunjukkan keaktifan transaksi yang tinggi. Kontribusi pendapatan segmen ini sangat signifikan bagi perusahaan, sehingga strategi yang direkomendasikan adalah VIP Program, Early Access terhadap produk baru, Personal Offer, dan Corporate Gifting guna mempertahankan kepuasan dan loyalitas jangka panjang mereka.

D. PENUTUP

Simpulan

Berdasarkan hasil analisis, visualisasi, dan rekomendasi bisnis yang telah diuraikan pada bagian sebelumnya, dapat disimpulkan beberapa temuan penting dari penelitian ini. Berikut adalah simpulan dan saran yang dapat diambil sebagai penutup dari keseluruhan pembahasan.

Penelitian ini menunjukkan bahwa penerapan analisis RFM dan algoritma K-Means clustering memberikan dampak yang signifikan terhadap peningkatan customer intelligence pada data transaksi retail online. Melalui proses pembersihan data, rekayasa fitur RFM, serta pemodelan clustering dengan K=3 yang telah divalidasi menggunakan Elbow Method dan Silhouette Score, penelitian ini berhasil menghasilkan tiga segmen pelanggan yang memiliki karakteristik sangat berbeda. Segmentasi tersebut tidak hanya mampu menggambarkan pola perilaku belanja pelanggan secara jelas, tetapi juga memberikan wawasan strategis yang dapat langsung dimanfaatkan oleh perusahaan retail online untuk menyusun strategi pemasaran yang lebih tepat sasaran, meningkatkan retensi pelanggan, serta mengoptimalkan pendapatan.

Saran

Berdasarkan hasil penelitian dan keterbatasan yang ada, berikut beberapa saran yang dapat dijadikan bahan pertimbangan untuk pengembangan lebih lanjut:

1. Mengintegrasikan hasil segmentasi ini ke dalam dashboard sistem informasi perusahaan

Integrasi ini sangat diperlukan agar tim pemasaran dan manajemen dapat memantau profil pelanggan secara real-time. Dengan adanya dashboard yang interaktif, keputusan bisnis dapat diambil lebih cepat, tepat, dan berbasis data, sehingga strategi pemasaran menjadi lebih dinamis dan responsif terhadap perubahan perilaku pelanggan.

2. Penelitian lanjutan dapat menggabungkan K-Means dengan algoritma lain untuk prediksi churn

Penggabungan ini penting dilakukan karena K-Means hanya memberikan segmentasi statis. Dengan menambahkan algoritma klasifikasi seperti Random Forest atau XGBoost, perusahaan dapat memprediksi pelanggan mana yang berpotensi churn di masa depan. Prediksi ini memungkinkan intervensi retensi yang lebih dini dan tepat sasaran, sehingga tingkat kehilangan pelanggan dapat ditekan secara signifikan.

3. Menambahkan variabel demografi pelanggan jika data tersedia

Penambahan variabel demografi seperti usia, jenis kelamin, lokasi geografis, dan tingkat pendapatan akan menghasilkan segmentasi yang lebih kaya dan mendalam. Dengan demikian, strategi pemasaran dapat dibuat lebih personal, relevan, dan efektif sesuai karakteristik spesifik setiap segmen pelanggan.

Ucapan Terima Kasih

Puji dan syukur kami panjatkan ke hadirat Allah SWT atas limpahan rahmat, taufik, dan kemudahan-Nya sehingga penelitian dan penyusunan jurnal ini dapat diselesaikan dengan baik. Tanpa pertolongan dan izin-Nya, seluruh rangkaian kegiatan mulai dari pengolahan data, analisis clustering, hingga penyusunan laporan ilmiah ini tidak akan berjalan dengan lancar.

Penulis juga menyampaikan apresiasi yang sebesar-besarnya kepada seluruh anggota kelompok, yaitu Muhammad Naufal Skha Yusfa, Ahyat Musyawwa, dan Aldy Bifal Pratama, yang telah bekerja sama dengan penuh tanggung jawab, saling mendukung, serta berkontribusi secara aktif dalam setiap tahapan penyusunan jurnal ini, mulai dari pengumpulan data, pemrosesan, analisis, hingga penyusunan laporan.

Tidak lupa, rasa hormat dan terima kasih yang tulus kami sampaikan kepada Ibu Mufidah selaku dosen mata kuliah Data Mining yang telah memberikan bimbingan, arahan, serta masukan yang sangat berharga selama proses perkuliahan dan penyelesaian proyek akhir ini. Bimbingan dan masukan dari Ibu Mufidah menjadi landasan penting dalam penyusunan penelitian ini secara sistematis dan metodologis.

E. DAFTAR PUSTAKA

- Achmad, S. L., Fauzi, A., Rahmat, R., & Indra, J. (2024). SEGMENTASI PELANGGAN MENGGUNAKAN K-MEANS CLUSTERING DI TOKO RETAIL. *Jurnal Teknik Informasi Dan Komputer (Tekinkom)*, 7(2), 736. <https://doi.org/10.37600/tekinkom.v7i2.1226>
- Fadhilla, D. R., Ramadhan, R., Vedliyan, V., & Oktavian, D. (n.d.). *Segmentasi Pelanggan Berbasis RFM dengan Algoritma K-Means Pada Data Transaksi Online Retail* (Vol. 2, Number 3).
- Fauzan, R. M., & Alfian, G. (2024). Segmentasi Pelanggan E-Commerce Menggunakan Fitur Recency,

- Frequency, Monetary (RFM) dan Algoritma Klasterisasi K-Means. In *Jurnal Informatika Sunan Kalijaga* (Vol. 9, Number 3).
- Febrina, Y. K., Saputra, R., & G, K. F. (2025). Segmentasi Pelanggan Toko Hanifah Berdasarkan Analisis RFM dengan Metode K-Means Clustering. *Jurnal Pustaka AI (Pusat Akses Kajian Teknologi Artificial Intelligence)*, 5(2), 274–282. <https://doi.org/10.55382/jurnalpustakaai.v5i2.1084>
- Hairani, H., Wahyudi, R., Gede Yogi Pratama, & M.Khaerul Ihsan. (2026). Application of RFM Model and K-Means Algorithm in Customer Loyalty Segmentation of Indonesian Regional Water Utility Company (PDAM). *International Journal of Engineering and Computer Science Applications (IJECSA)*, 5(1), 1–8. <https://doi.org/10.30812/ijecca.v5i1.6087>
- Karimah, M., & Marwati, F. (2024). Sustainability of Quality Management by Implementing Data Mining to Predict Academic Achievement. *Journal of Social Science and Business Studies*, 2(3), 240–250. <https://doi.org/10.61487/jssbs.v2i3.90>
- Martiansah, R., Monalisa, S., Muttakin, F., & Fronita, M. (2025). Customer Segmentation Analysis Through RFM-D Model and K-Means Algorithm. In *Jurnal Sistem Cerdas*.
- Nurlathifa, A., & Sucahyati, D. (2025). Segmentasi Pelanggan menggunakan Metode K-means Clustering Berdasarkan Model RFM pada Bisnis Food and Beverage (Studi Kasus: Coffee shop X). *Jurnal Nasional Teknologi Dan Sistem Informasi*, 11(3), 218–225. <https://doi.org/10.25077/teknosi.v11i3.2025.218-225>
- Sitorus, E. R., & Nugraha, I. (2025). Analisis segmentasi pelanggan dengan model RFM (Recency, Frequency, Monetary) dan K-Means Clustering (Studi kasus: PT XYZ). *Jurnal Teknik Industri Terintegrasi*, 8(1), 266–278. <https://doi.org/10.31004/jutin.v8i1.39447>
- Triantara, I. K. Y., Raharja, M. A., & Sarasvananda, I. B. G. (2025). Segmentasi Pelanggan Berbasis RFMT Menggunakan K-Means dan Hierarchical Clustering I Komang Yosua Triantara a1, Made Agung Raharja a2, Ida Bagus Gede Sarasvananda a3. *JNATIA*, 3(4).
- Yusak, A., Rumapea, N., Pratiwi, D., & Sari, S. (2024). Analisis Segmentasi Pelanggan Ritel Online Menggunakan K-Means Clustering Berdasarkan Model Recency, Frequency, Monetary (RFM). *Jurnal Sains Dan Teknologi*, 6(3), 292–299. <https://doi.org/10.55338/saintek.v6i3.4607>