

Segmentasi Pelanggan *E-Commerce* Menggunakan Fitur *Recency, Frequency, Monetary* (RFM) dan Algoritma Klasterisasi K-Means

Reyhan Muhammad Fauzan ⁽¹⁾, Ganjar Alfian ^{(2)*}

Teknik Elektro dan Informatika, Sekolah Vokasi, Universitas Gadjah Mada, Yogyakarta, Indonesia

e-mail : reyhanmuhammad@mail.ugm.ac.id, ganjar.alfian@ugm.ac.id.

* Penulis korespondensi.

Artikel ini diajukan 25 Januari 2024, direvisi 1 Mei 2024, diterima 2 Mei 2024, dan dipublikasikan 25 September 2024.

Abstract

The rapid growth in the e-commerce industry demands the development of smarter and more focused marketing strategies. One approach that can be applied is customer segmentation using various features such as Recency, Frequency, and Monetary (RFM), along with machine learning-based clustering methods. The objective of this study is to design and develop a web-based e-commerce customer segmentation application using a combination of RFM features and clustering methods. The study proposes the K-Means algorithm and compares it with K-Medoids and Fuzzy C Means using publicly available e-commerce datasets. Experimental results showed that the K-Means algorithm outperformed K-Medoids and Fuzzy C Means (FCM) based on the Silhouette Score of 0.67305, Davies Bouldin Index of 0.51435, and Calinski Harabasz Index of 5647.89. Through analysis and testing, the designed application has proven effective in grouping customers into relevant segments. These segments are divided into three categories: *Loyal*, *Need Attention*, and *Promising*, visualized in a web-based application dashboard using Streamlit. The developed application allows e-commerce business owners and users from the business, management, and marketing divisions to categorize customers based on transaction data. The results of this study are expected to provide valuable insights to e-commerce management and marketing professionals who are facing increasingly fierce competition.

Keywords: *E-Commerce*, *Customer Segmentation*, *RFM*, *K-Means*, *Web Application*

Abstrak

Peningkatan pesat dalam industri *e-commerce* menuntut pengembangan strategi pemasaran yang lebih cerdas dan terfokus. Salah satu pendekatan yang dapat diterapkan adalah segmentasi pelanggan menggunakan berbagai fitur, seperti *Recency*, *Frequency*, dan *Monetary* (RFM), serta metode klasterisasi berbasis *machine learning*. Tujuan dari penelitian ini adalah untuk merancang dan membangun aplikasi segmentasi pelanggan *e-commerce* berbasis web yang menggunakan kombinasi fitur RFM dan metode klasterisasi. Penelitian ini mengusulkan algoritma K-Means dan membandingkannya dengan K-Medoids, serta Fuzzy C Means pada *dataset e-commerce* yang tersedia secara publik. Hasil penelitian menunjukkan bahwa algoritma K-Means lebih unggul dibanding algoritma K-Medoids dan Fuzzy C Means (FCM) berdasarkan nilai Silhouette Coefficient sebesar 0,67305, Davies Bouldin Index sebesar 0,51435, dan Calinski Harabasz Index sebesar 5647,89. Melalui analisis dan pengujian, aplikasi yang dirancang telah terbukti efektif dalam mengelompokkan pelanggan ke dalam segmen yang relevan. Segmen tersebut dibagi menjadi tiga kategori yaitu *Loyal*, *Need Attention*, dan *Promising*, kemudian divisualisasikan dalam bentuk dashboard pada aplikasi berbasis web menggunakan *Streamlit*. Aplikasi yang dikembangkan dalam penelitian ini memungkinkan pemilik bisnis *e-commerce* ataupun pengguna dari bidang bisnis, divisi manajemen, dan pemasaran untuk mengelompokkan pelanggan berdasarkan data transaksi. Hasil dari penelitian ini diharapkan dapat memberikan informasi berharga kepada manajemen *e-commerce* maupun bidang pemasaran dalam menghadapi persaingan yang semakin ketat.

Kata Kunci: *E-Commerce*, *Segmentasi Pelanggan*, *RFM*, *K-Means*, *Aplikasi Web*



1. PENDAHULUAN

Di era digital yang dipenuhi oleh industri *e-commerce* yang pesat, persaingan antarperusahaan untuk menarik, mempertahankan, dan memahami pelanggan semakin ketat. Aktivitas perdagangan melalui penerapan *e-commerce* sangat praktis hanya menggunakan perangkat elektronik seperti laptop, komputer, atau *smartphone* dan menggunakan internet sebagai perantara (Molla & Licker, 2005). Namun, dengan pertumbuhan pesat ini, juga muncul tantangan besar dalam mengelola dan memahami pelanggan *e-commerce*.

Mengingat kompleksitas bisnis dan pasar *e-commerce* yang terus berkembang, keberhasilan bisnis dalam sektor ini sangat bergantung pada kemampuan mereka untuk memahami dan merespon perubahan perilaku pelanggan dengan cepat dan efektif. Memahami kebutuhan dan perilaku pelanggan merupakan hal yang krusial dalam merancang strategi pemasaran yang efektif dan meningkatkan kualitas layanan. Salah satu tantangan utama dalam *e-commerce* adalah bagaimana mengelompokkan pelanggan ke dalam segmen yang relevan. Segmentasi pelanggan diperlukan untuk mengelompokkan pelanggan yang memiliki kesamaan karakteristik (Kim et al., 2006). Selanjutnya, hasil kelompok pelanggan akan mendapatkan perlakuan yang berbeda-beda dalam strategi pemasaran (Li et al., 2010; Shirole et al., 2021).

Analisis RFM adalah salah satu cara untuk melakukan segmentasi pelanggan berdasarkan *recency*, *frequency*, dan *monetary*. Faktor *recency* menilai seberapa baru pelanggan melakukan transaksi. *Frequency* menilai seberapa sering pelanggan bertransaksi, sedangkan *monetary* menilai seberapa besar total pengeluaran yang dilakukan oleh pelanggan saat bertransaksi (Anitha & Patil, 2022). Selanjutnya nilai RFM juga dapat digunakan sebagai fitur untuk segmentasi pelanggan menggunakan metode klusterisasi yang berbasis *unsupervised learning*.

Penelitian sebelumnya menganalisis pola penggunaan Mass Rapid Transit (MRT) dan layanan berbagi sepeda atau *bike sharing* (YouBike) di Taipei, Taiwan, dengan menggunakan data historis. Dengan menggunakan algoritma RFM dan K-Means *clustering*, penelitian ini mengidentifikasi tiga kelompok pengguna MRT-YouBike yang berbeda: *potential*, *vulnerable*, dan *loyal* (Chen et al., 2022). Penelitian selanjutnya berfokus pada data *e-retailer* sebagai studi kasusnya. Pada penelitian ini, RFM (*recency*, *frequency*, *monetary*) digunakan sebagai solusi pembuatan kluster. Penelitian ini menganalisis informasi pembelian pelanggan selama delapan bulan. Kemudian kluster dievaluasi menggunakan *metrics* Silhouette Coefficient untuk algoritma K-Means dengan jumlah kluster yang berbeda. Hasil yang didapatkan menunjukkan bahwa kluster dengan jumlah 3 lebih baik daripada kluster yang berjumlah 5 (Anitha & Patil, 2022). Implementasi model K-Means untuk segmentasi pelanggan menjadi 3 kluster yaitu *loyal*, *promising*, dan *need attention* sudah dilakukan oleh penelitian sebelumnya (Hilmy et al., 2023). Hasilnya menunjukkan bahwa model K-Means lebih baik dibanding dengan model K-Medoids dalam segmentasi data pelanggan.

Dengan menggunakan metode klusterisasi K-Means, dapat menentukan kategori dan strategi yang baik untuk pelanggan. Hasil penelitian sebelumnya menunjukkan bahwa metode klusterisasi K-Means dengan skor atribut RFM yang berbeda berhasil mengelompokkan 14.979 data pelanggan kedalam 5 kluster (Sutresno et al., 2018). Pada penelitian selanjutnya menganalisis model RFM dan klusterisasi K-Means pada *online bookstore*. Masalah dalam penelitian ini adalah menurunnya jumlah transaksi dalam rentang bulan Januari 2019 hingga bulan November 2020. Sehingga dilakukan analisis untuk menciptakan kluster yang optimal menggunakan *elbow method* dan dilakukan pengujian performa menggunakan *metrics* Silhouette Coefficient dan Calinski Harabasz Index. Hasil jumlah kluster yang optimal untuk strategi pemasaran dari 23.152 pelanggan adalah berjumlah 3 kluster (Juhari & Juarna, 2022).

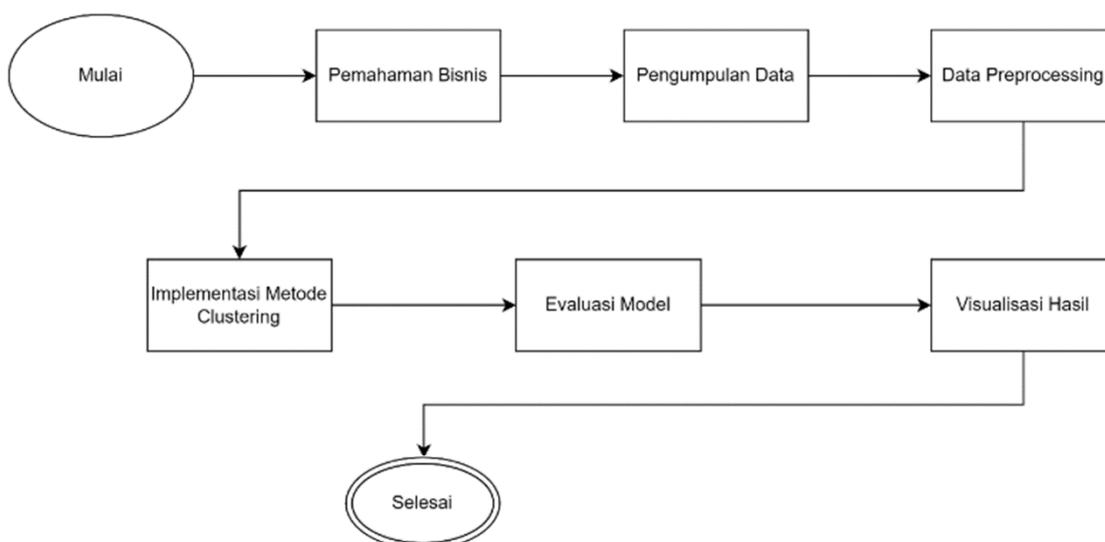
Selain berfokus terhadap analisis, penelitian selanjutnya memvisualisasikan hasil dari analisis yang didapat menjadi *dashboard* berbasis web. Pada penelitian kali ini data pelanggan diperoleh dari *platform* Kaggle. Penelitian ini menggunakan model RFM dan algoritma K-Means untuk menganalisis segmentasi pelanggan. Dalam penelitian ini juga digunakan metode Silhouette



untuk mengoptimalkan kluster yang berjumlah 4 kluster. Hasil yang diperoleh kemudian divisualisasikan menjadi bentuk *dashboard* menggunakan *platform* Streamlit (Alzami et al., 2023). Penelitian sebelumnya telah mengeksplorasi penggunaan metode RFM dan klusterisasi K-Means dalam segmentasi pelanggan, tetapi masih terdapat kekurangan dalam membandingkan metode ini dengan algoritma klusterisasi lain seperti K-Medoids dan Fuzzy C Means, serta dalam mengevaluasi efektivitas model klusterisasi tersebut. Dalam penelitian ini, akan dikembangkan sebuah aplikasi web yang mengintegrasikan RFM dan K-Means untuk segmentasi pelanggan. Efisiensi algoritma K-Means akan diukur dan dibandingkan dengan K-medoids dan Fuzzy C Means menggunakan tiga metrik validasi kluster: Skor Silhouette, Indeks Calinski-Harabasz, dan Indeks Davies-Bouldin. Tujuan akhir dari penelitian ini adalah untuk menyediakan sebuah alat yang efektif bagi perusahaan, khususnya di sektor bisnis dan *e-commerce*, untuk meningkatkan strategi pemasaran, mempertahankan pelanggan, dan membina hubungan jangka panjang yang lebih kuat dengan pelanggan.

2. METODE PENELITIAN

Langkah-langkah dalam proses analisis mencakup beberapa tahapan yang dimulai dengan memahami secara menyeluruh kebutuhan bisnis. Proses tersebut mencakup pengumpulan data yang relevan, pemrosesan awal untuk memastikan kualitasnya, pembuatan model yang sesuai, evaluasi kinerjanya, dan visualisasi hasil (Anitha & Patil, 2022). Semua langkah ini dapat dilihat detail dalam Gambar 1.



Gambar 1 Alur Penelitian

Alur dimulai dari memahami secara menyeluruh tentang bagaimana suatu bisnis atau organisasi beroperasi, tujuan bisnisnya, serta konteks eksternal yang mempengaruhi keberhasilannya. Pemahaman bisnis juga sangat penting dalam pengembangan perangkat lunak atau solusi teknologi, karena membantu para pengembang membangun solusi yang sesuai dengan kebutuhan bisnis dan memberikan nilai tambah. Selanjutnya adalah pengumpulan data yang dibutuhkan untuk penelitian. *Dataset* yang digunakan pada penelitian ini diperoleh melalui *platform online data.world* dengan nama Global Superstore. *Dataset* ini merupakan kumpulan data yang berasal dari komunitas Tableau dengan total 51.290 baris. Kumpulan data dimulai dari tahun 2011 hingga tahun 2014 yang berpusat pada transaksi pelanggan dari berbagai vendor dan pasar yang berbeda. Penelitian dengan *dataset* ini sudah dilakukan sebelumnya oleh (Mahfuza et al., 2022) dengan tujuan untuk melakukan segmentasi pelanggan. Adapun detail *dataset* dapat dilihat pada Gambar 2.



Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	...	Product ID	Category	Sub-Category
CA-2012-124891	7/31/2012	7/31/2012	Same Day	RH-19495	Rick Hansen	Consumer	New York City	New York	...	TEC-AC-10003033	Technology	Accessories
IN-2013-77878	2/5/2013	2/7/2013	Second Class	JR-16210	Justin Ritter	Corporate	Wollongong	New South Wales	...	FUR-CH-10003950	Furniture	Chairs
IN-2013-71249	10/17/2013	10/18/2013	First Class	CR-12730	Craig Reiter	Consumer	Brisbane	Queensland	...	TEC-PH-10004664	Technology	Phones
ES-			First	KM-	Kelbarian	Home				TEC-PH-		

Gambar 2 Ilustrasi Dataset Penelitian

Sebelum melakukan analisis diperlukan pembersihan data dengan melakukan *data preprocessing* agar hasil analisis dapat sesuai dengan keinginan. Melalui *data preprocessing dataset* dapat dibersihkan seperti mengisi nilai yang hilang, mengkoreksi data ganda, ataupun mengkoreksi ketidakcocokan antar data (Anitha & Patil, 2022). Pada tahap selanjutnya yaitu menentukan metode analisis dengan menggunakan RFM (*recency, frequency, monetary*) dan K-Means. RFM memberikan kerangka kerja yang kuat dengan mempertimbangkan faktor seberapa baru pelanggan berbelanja (*recency*), seberapa sering mereka berbelanja (*frequency*), dan seberapa besar total belanjaan mereka (*monetary*). Integrasi metode klusterisasi seperti K-Means membawa dimensi tambahan dalam proses segmentasi, memungkinkan identifikasi kelompok pelanggan yang memiliki pola perilaku serupa. Pembuatan model dimulai dengan menghitung skor *recency, frequency, dan monetary* yang diperoleh dari metode RFM. Untuk memberikan gambaran perhitungan tersebut, contoh data pelanggan disajikan dalam Tabel 1.

Tabel 1 Contoh Data Pelanggan

Nama Pelanggan	Tanggal Pembelian	Jumlah Pembelian
A	12 November 2023	\$150
B	13 November 2023	\$200
C	15 November 2023	\$50
B	17 November 2023	\$300
D	19 November 2023	\$200
D	20 November 2023	\$100

Dengan mengasumsikan bahwa hari ini adalah tanggal 23 November 2023, skor *Recency* dihitung berdasarkan perbedaan antara tanggal saat ini dan tanggal transaksi terakhir pelanggan. Skor *Frequency* diperoleh dengan menghitung total transaksi yang dilakukan, sementara skor *Monetary* dihitung dengan menjumlahkan total pengeluaran dari setiap transaksi. Hasil dari ketiga perhitungan skor tersebut dapat dilihat pada Tabel 2.

Tabel 2 Contoh Hasil Perhitungan RFM

Nama Pelanggan	Recency	Frequency	Monetary
A	11	1	\$150
B	6	2	\$500
C	8	1	\$50
D	3	2	\$300

Skor yang telah dihitung dapat berfungsi sebagai fitur untuk segmentasi pelanggan pada algoritma K-Means. Algoritma K-Means bekerja dengan pertama menentukan jumlah kluster (*K*) yang diinginkan (Adiyanto & Arie Wijaya, 2023). Kemudian, secara acak memilih titik-titik awal sebagai *centroid* atau titik tengah. Setiap titik dalam *dataset* kemudian dikelompokkan ke kluster



yang paling dekat berdasarkan jarak Euclidean ke *centroid* tersebut. Jarak (d) antara titik data x dan pusat massa y dihitung menggunakan rumus jarak Euclidean, sebagaimana tertulis pada Pers. (1). Setelah pengelompokan, posisi *centroid* diperbarui dengan menghitung rata-rata dari semua titik dalam kluster tersebut. Proses ini diulang – pengelompokan data dan pembaruan *centroid* – hingga kluster stabil, yaitu ketika tidak ada perubahan signifikan dalam pengelompokan atau sampai batas iterasi maksimum tercapai. Hasil akhirnya adalah pembentukan kluster dengan titik-titik data yang memiliki kesamaan ciri di dalamnya dan titik-titik *centroid* yang stabil.

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

Langkah berikutnya melibatkan evaluasi model pada fase penilaian hasil kluster. Silhouette Coefficient (SC) adalah metrik evaluasi yang digunakan untuk mengukur sejauh mana titik data dalam suatu kluster cocok dengan kluster tempatnya berada dan seberapa dekat atau jauh dari kluster lainnya. Metrik ini merupakan metode evaluasi kluster gabungan antara separasi dan kohesi (Paembonan & Abduh, 2021). Adapun rumus SC dapat dilihat pada Pers. (2).

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2)$$

Evaluasi selanjutnya dilakukan dengan menggunakan Davies Bouldin Index (DBI), yang mengukur perbandingan antara jarak di dalam suatu kluster dan jarak antar kluster. Pers. (3) menjelaskan DBI, dengan k sebagai jumlah total kluster yang digunakan, dan $(R_{i,j})$ sebagai rasio antara kluster i dan kluster j . Nilai DBI yang lebih rendah menandakan bahwa kualitas klusterisasi yang dihasilkan lebih baik (Agustino & Budaya, 2023).

$$DBI = \frac{1}{K} \sum_i \max_{i \neq j} (R_{i,j}) \quad (3)$$

Metrik evaluasi terakhir adalah Calinski Harabasz Index (CHI) yang digunakan untuk mengukur kualitas suatu klusterisasi pada data. Keuntungan dari CHI adalah bahwa nilai yang lebih tinggi mencerminkan kluster yang lebih jelas (Sikana & Wijayanto, 2021). Perhitungan CHI dapat dilihat pada Pers. 4.

$$CHI = \frac{tr(B_k)}{tr(W_k)} \times \frac{(N - k)}{(k - 1)} \quad (4)$$

Dalam melakukan segmentasi pelanggan, digunakan model *machine learning*. Model klusterisasi diterapkan menggunakan bahasa pemrograman Python dengan menggunakan pustaka Scikit-learn, dan parameter bawaan dari Scikit-learn yang digunakan. Setelah melalui tahap pengujian, langkah selanjutnya adalah fase implementasi model, yang melibatkan pemanfaatan pustaka Streamlit. Pustaka ini mempermudah pembuatan dasbor yang menampilkan hasil analisis secara visual. Untuk menjalankan *dashboard* menggunakan Streamlit, diperlukan koneksi ke GitHub untuk mengakses kode sumber yang sebelumnya telah diunggah ke GitHub. Melalui proses ini, program yang sebelumnya berjalan secara lokal dapat diakses secara publik.

3. HASIL DAN PEMBAHASAN

3.1 Perbandingan Model Klusterisasi

Perbandingan antara metode klusterisasi K-Means, K-Medoids, dan Fuzzy C Means dapat dievaluasi berdasarkan tiga metrik utama, yaitu Silhouette Coefficient (SC), Davies Bouldin Index (DBI), dan Calinski Harabasz Index (CHI). Silhouette Coefficient memberikan indikasi seberapa



baik objek dalam kluster terpisah dan saling berdekatan, dengan nilai mendekati 1 menandakan pembentukan kluster yang baik. Davies Bouldin Index menilai sejauh mana batas kluster terdefinisi dan seberapa homogen kluster tersebut. Nilai rendah pada DBI mengindikasikan pembentukan kluster yang baik. Sementara itu, Calinski Harabasz Index mengevaluasi homogenitas dan pemisahan kluster, dengan nilai tinggi menandakan pembentukan kluster yang baik. Hasil perhitungan yang telah dilakukan menunjukkan bahwa algoritma K-Means lebih unggul daripada algoritma K-Medoids dan Fuzzy C Means. Nilai ketiga indeks tersebut menunjukkan nilai 0,67305 pada Silhouette Coefficient, nilai 0,51435 Davies Bouldin Index, dan nilai 5647,89 pada Calinski Harabasz Index, sebagaimana tergambar pada Tabel 3.

Tabel 3 Perbandingan Model Klusterisasi

Matriks	K-Means	K-Medoids	Fuzzy C Means
Silhouette Coefficient (SC)	0,67305	0,66130	0,67303
Davies Bouldin Index (DBI)	0,51435	0,52575	0,51467
Calinski Harabasz Index (CHI)	5647,89	5271,67	5644,48

Hasil segmentasi menggunakan fitur RFM melalui penerapan algoritma K-Means menunjukkan bahwa pelanggan dapat diklasifikasikan ke dalam tiga kategori, yaitu *loyal*, *promising*, dan *need attention* (Hilmy et al., 2023). Jumlah masing-masing pelanggan/konsumen dalam kelompok *loyal*, *promising*, dan *need attention* adalah 222, 546, dan 822. Dari jumlah tersebut, 222 pelanggan dianggap *loyal* karena memiliki frekuensi tinggi dan nilai *monetary* yang tinggi. Kategori *promising* mencakup 546 pelanggan dengan frekuensi dan *monetary* sedang. Sementara itu, kategori *need attention* terdiri dari 822 pelanggan dengan frekuensi dan *monetary* yang rendah. Hasil pengelompokan menggunakan K-Medoids menunjukkan bahwa 316 pelanggan dikategorikan sebagai *loyal*, 478 pelanggan dalam kategori *promising*, dan 796 pelanggan diklasifikasikan *need attention*. Algoritma Fuzzy C Means menghasilkan 225 pelanggan pada kategori *loyal*, 546 pelanggan pada *promising*, dan 819 pelanggan pada kategori *need attention*. Rangkuman jumlah pelanggan dari hasil klusterisasi ini dapat dilihat pada Tabel 4.

Tabel 4 Hasil Pengelompokan Klusterisasi

Kategori	K-Means	K-Medoids	Fuzzy C Means
<i>Loyal</i>	222	316	225
<i>Promising</i>	546	478	546
<i>Need Attention</i>	822	796	819

3.2 Visualisasi Hasil

Penelitian ini bertujuan untuk merancang dan mengembangkan sistem segmentasi pelanggan *e-commerce* berbasis web untuk membantu memberikan dukungan pada proses pengambilan keputusan manajemen. Sistem ini dikonstruksi dengan menggunakan bahasa pemrograman Python, memanfaatkan perpustakaan Streamlit untuk antarmuka web dan memanfaatkan Google Sheets untuk meletakkan *file dataset* yang telah dimiliki. Proses pengelompokan data pelanggan dilakukan dengan menggunakan perpustakaan Scikit-learn. Setelah berhasil melakukan *login*, pengguna akan diarahkan ke halaman *dashboard*, sebagaimana terlihat pada Gambar 3, yang menampilkan hasil klusterisasi.



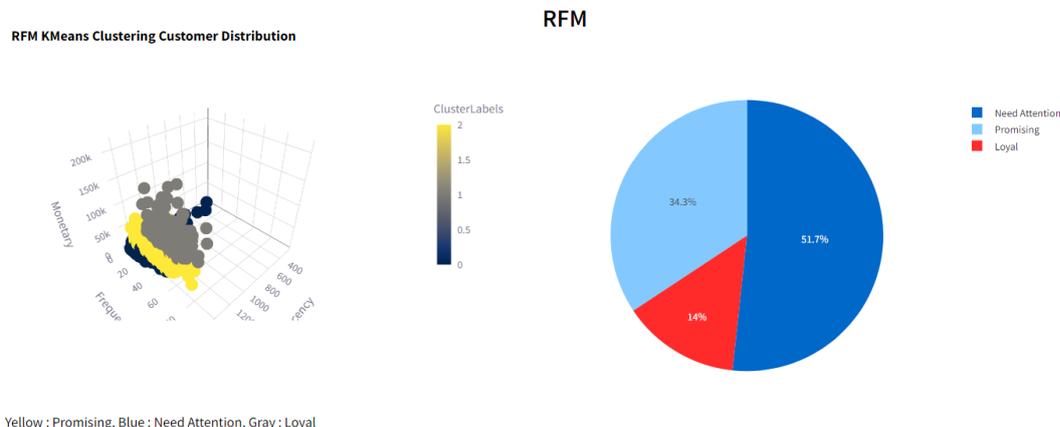
Data RFM (Recency, Frequency, Monetary) With KMeans Clustering

Total Customer: 1590

Loyal: 222

Promising: 546

Need attention: 822



Gambar 3 Hasil Visualisasi

4. KESIMPULAN

Segmentasi pelanggan menggunakan model RFM telah berhasil dijalankan, dan dari hasil analisis tersebut, beberapa kesimpulan dapat diambil. Pertama, fitur RFM (*Recency, Frequency, Monetary*) bersama dengan algoritma K-Means, K-Medoids, dan Fuzzy C Means berhasil menampilkan 3 segmen pelanggan utama, yaitu *Loyal, Promising, dan Need Attention*. Kedua, performa algoritma K-Means, K-Medoids, dan FCM dievaluasi menggunakan metrik seperti Silhouette Coefficient (SC), Davies Bouldin Index (DBI), dan Calinski Harabasz Index (CHI). Hasil evaluasi menunjukkan bahwa algoritma K-Means mencapai hasil paling optimal dengan nilai 0,67305 pada SC, 0,51435 pada DBI, dan 5647,89 pada CHI. Ketiga, *output* dari analisis dapat disajikan secara visual melalui aplikasi web menggunakan pustaka Streamlit, sehingga dapat memberikan dukungan untuk pengambilan keputusan yang efektif. Visualisasi menggunakan *scatter plots, pie charts, dan line charts* memudahkan penyajian hasil analisis dengan jelas, membantu pemahaman pola dan tren pelanggan secara intuitif. Kedepannya, optimalisasi model dan komparasi dengan model klasterisasi lain akan kami sajikan dalam penelitian selanjutnya.

DAFTAR PUSTAKA

- Adiyanto, A., & Arie Wijaya, Y. (2023). PENERAPAN ALGORITMA K-MEANS PADA PENGELOMPOKAN DATA SET BAHAN PANGAN INDONESIA TAHUN 2022-2023. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(2), 1344-1350. <https://doi.org/10.36040/jati.v7i2.6849>
- Agustino, D. P., & Budaya, I. G. B. A. (2023). Evaluasi Performa Segmentasi Pelanggan Tenant Inkubator Bisnis dengan Menggunakan Model Consensus Clustering. *Prosiding CORISINDO* 2023, 236-240. <https://www.stmikpontianak.org/ojs/index.php/corisindo/article/view/62>
- Alzami, F., Sambasri, F. D., Nabila, M., Megantara, R. A., Akrom, A., Pramunendar, R. A., Prabowo, D. P., & Sulistiyawati, P. (2023). Implementation of RFM Method and K-Means Algorithm for Customer Segmentation in E-Commerce with Streamlit. *ILKOM Jurnal Ilmiah*, 15(1), 32-44. <https://doi.org/10.33096/ilkom.v15i1.1524.32-44>
- Anitha, P., & Patil, M. M. (2022). RFM model for customer purchase behavior using K-Means algorithm. *Journal of King Saud University - Computer and Information Sciences*, 34(5), 1785-1792. <https://doi.org/10.1016/j.jksuci.2019.12.011>



- Chen, A. H. L., Liang, Y.-C., Chang, W.-J., Siau, H.-Y., & Minanda, V. (2022). RFM Model and K-Means Clustering Analysis of Transit Traveller Profiles: A Case Study. *Journal of Advanced Transportation*, 2022(1), 1–14. <https://doi.org/10.1155/2022/1108105>
- Hilmy, F. M., Nurhaliza, R. A., Huzyan Octava, M. Q., & Alfian, G. (2023). Web-based E-Commerce Customer Segmentation System Using RFM and K-Means Model. *2023 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, 83–87. <https://doi.org/10.1109/3ICT60104.2023.10391650>
- Juhari, T., & Juarna, A. (2022). IMPLEMENTATION RFM ANALYSIS MODEL FOR CUSTOMER SEGMENTATION USING THE K-MEANS ALGORITHM CASE STUDY XYZ ONLINE BOOKSTORE. *EXPLORE*, 12(1), 107–118. <https://doi.org/10.35200/explore.v12i1.548>
- Kim, S.-Y., Jung, T.-S., Suh, E.-H., & Hwang, H.-S. (2006). Customer segmentation and strategy development based on customer lifetime value: A case study. *Expert Systems with Applications*, 31(1), 101–107. <https://doi.org/10.1016/j.eswa.2005.09.004>
- Li, W., Wu, X., Sun, Y., & Zhang, Q. (2010). Credit Card Customer Segmentation and Target Marketing Based on Data Mining. *2010 International Conference on Computational Intelligence and Security*, 73–76. <https://doi.org/10.1109/CIS.2010.23>
- Mahfuza, R., Islam, N., Toyeb, Md., Emon, M. A. F., Chowdhury, S. A., & Alam, Md. G. R. (2022). LRFMV: An efficient customer segmentation model for superstores. *PLOS ONE*, 17(12), e0279262. <https://doi.org/10.1371/journal.pone.0279262>
- Molla, A., & Licker, P. S. (2005). eCommerce adoption in developing countries: a model and instrument. *Information & Management*, 42(6), 877–899. <https://doi.org/10.1016/j.im.2004.09.002>
- Paembonan, S., & Abduh, H. (2021). Penerapan Metode Silhouette Coefficient untuk Evaluasi Clustering Obat. *PENA TEKNIK: Jurnal Ilmiah Ilmu-Ilmu Teknik*, 6(2), 48–54. https://doi.org/10.51557/pt_jiit.v6i2.659
- Shirole, R., Salokhe, L., & Jadhav, S. (2021). Customer Segmentation using RFM Model and K-Means Clustering. *International Journal of Scientific Research in Science and Technology*, 591–597. <https://doi.org/10.32628/IJSRST2183118>
- Sikana, A. M., & Wijayanto, A. W. (2021). Analisis Perbandingan Pengelompokan Indeks Pembangunan Manusia Indonesia Tahun 2019 dengan Metode Partitioning dan Hierarchical Clustering. *Jurnal Ilmu Komputer*, 14(2), 66–78. <https://doi.org/10.24843/JIK.2021.v14.i02.p01>
- Sutresno, S. A., Iriani, A., & Sedyono, E. (2018). Metode K-Means Clustering dengan Atribut RFM untuk Mempertahankan Pelanggan. *Jurnal Teknik Informatika Dan Sistem Informasi*, 4(3), 433–440–433–440. <https://doi.org/10.28932/jutisi.v4i3.878>

