

ANALISIS SEGMENTASI DAN NILAI PELANGGAN MENGGUNAKAN METODE *CLUSTER ANALYSIS*, RFM, PADA TOKO KEBUTUHAN BAYI

SEGMENTATION ANALYSIS AND CUSTOMER VALUE USING CLUSTER ANALYSIS METHOD, RFM, IN BABY NEEDS STORE

Panji Permadani, Muhammad Mansur Yafi*

Program Studi Teknik Industri, Fakultas Teknik, Universitas Nahdlatul Ulama Sidoarjo

mansuryafi.tin@unusida.ac.id

INFO ARTIKEL

Diterima: 15 Juni 2026

Direvisi: 22 Juni 2026

Disetujui: 10 Juli 2026

ABSTRAK

Persaingan bisnis ritel kebutuhan bayi menuntut pelaku usaha untuk mengelola data penjualan secara optimal. Toko XYZ masih menghadapi permasalahan berupa rendahnya perputaran penjualan pada beberapa produk yang mengakibatkan penumpukan stok dan kurang optimalnya pengelolaan persediaan. Penelitian ini bertujuan untuk melakukan segmentasi produk menggunakan pendekatan *Recency*, *Frequency*, dan *Monetary* (RFM) yang dikombinasikan dengan metode *K-Means Clustering*. Data yang digunakan merupakan data transaksi penjualan selama satu tahun dengan total 1.204 jenis produk. Sebelum proses *clustering*, data dinormalisasi menggunakan metode *Z-Score*, sedangkan jumlah cluster optimal ditentukan menggunakan metode *Elbow* dan *Silhouette*, yang menghasilkan tiga cluster sebagai jumlah terbaik. Hasil penelitian menunjukkan bahwa Cluster 1 terdiri dari 630 produk (52,33%) dengan performa penjualan yang stabil, Cluster 2 terdiri dari 90 produk (7,48%) yang merupakan produk unggulan karena memiliki nilai *Frequency* dan *Monetary* tertinggi, sedangkan Cluster 3 terdiri dari 484 produk (40,20%) dengan tingkat penjualan yang relatif rendah sehingga memerlukan evaluasi lebih lanjut. Hasil penelitian menunjukkan bahwa metode RFM dan *K-Means Clustering* mampu mengelompokkan produk berdasarkan karakteristik penjualannya sehingga dapat mendukung strategi pengelolaan persediaan dan pengambilan keputusan di Toko XYZ.

Kata Kunci:

Data Mining, *K-Means Clustering*, Retail, RFM, Segmentasi Produk.

ABSTRACT

The competitiveness of the baby retail business requires business owners to manage sales data optimally. XYZ Store still faces problems in the form of low sales turnover on several products which results in stock buildup and suboptimal inventory management. This study aims to segment products using the *Recency*, *Frequency*, and *Monetary* (RFM) approach combined with the *K-Means Clustering* method. The data used are sales transaction data for one year with a total of 1,204 product types. Before the clustering process, the data was normalized using the *Z-Score* method, while the optimal number of clusters was determined using the *Elbow* and *Silhouette* methods, which resulted in three clusters as the best number. The results show that Cluster 1 consists of 630 products (52.33%) with stable sales performance, Cluster 2 consists of 90 products (7.48%) which are superior products because they have the highest *Frequency* and *Monetary* values, while Cluster 3 consists of 484 products (40.20%) with a relatively low sales level that requires further evaluation. The results of the study show that the RFM and *K-Means Clustering* methods are able to group products based on their sales characteristics so that they can support inventory management strategies and decision making at XYZ Store.

Keywords:

Data Mining, *K-Means Clustering*, Retail, RFM, Product Segmentation.

*Corresponding author: mansuryafi.tin@unusida.ac.id

I. PENDAHULUAN

Bisnis retail merupakan salah satu sektor usaha yang mengalami perkembangan pesat seiring dengan kemajuan teknologi informasi dan meningkatnya kebutuhan masyarakat. Persaingan yang semakin ketat mendorong setiap pelaku usaha retail untuk menerapkan strategi yang mampu meningkatkan daya saing, mempertahankan pelanggan, dan mengoptimalkan penjualan. Kondisi tersebut menuntut perusahaan untuk tidak hanya menawarkan produk yang berkualitas, tetapi juga memahami perilaku dan kebutuhan pelanggan agar dapat memberikan pelayanan yang sesuai dengan harapan konsumen.

Persaingan yang tinggi juga dirasakan oleh Toko XYZ sebagai salah satu usaha retail yang bergerak di bidang perlengkapan bayi. Banyaknya toko yang menjual produk sejenis menyebabkan tingkat persaingan semakin meningkat, sehingga beberapa produk di Toko XYZ memiliki perputaran penjualan yang lambat dan membutuhkan waktu lebih lama untuk terjual. Kondisi tersebut dapat memengaruhi efektivitas pengelolaan persediaan, perputaran stok, serta pencapaian target penjualan [1]. Oleh karena itu, perusahaan perlu menerapkan strategi yang mampu menjaga hubungan baik dengan pelanggan [2].

Toko XYZ merupakan salah satu usaha retail kebutuhan bayi yang menyediakan berbagai produk seperti perlengkapan bayi, pakaian bayi, popok, susu, hingga perlengkapan ibu dan anak. Dalam aktivitas penjualannya, toko ini memiliki banyak data transaksi pelanggan yang tersimpan setiap harinya. Namun, data transaksi tersebut masih belum dimanfaatkan secara optimal sebagai dasar pengambilan keputusan bisnis. Strategi promosi dan pemberian diskon masih dilakukan secara umum kepada seluruh pelanggan tanpa mempertimbangkan karakteristik maupun tingkat loyalitas pelanggan. Kondisi tersebut menyebabkan promosi yang diberikan menjadi kurang efektif dan belum mampu meningkatkan loyalitas pelanggan secara maksimal [3].

Customer Relationship Management (CRM) merupakan strategi bisnis yang digunakan perusahaan untuk mengelola hubungan dengan pelanggan melalui pemanfaatan data pelanggan dan teknologi informasi [4]. CRM membantu perusahaan memahami kebutuhan pelanggan serta meningkatkan kualitas pelayanan secara berkelanjutan [5]. Selain itu, penerapan CRM juga mampu meningkatkan loyalitas pelanggan dan mendukung pertumbuhan bisnis perusahaan [6].

Dalam penerapan *Customer Relationship Management* (CRM), diperlukan pengolahan data pelanggan untuk memperoleh informasi yang bermanfaat sebagai dasar pengambilan keputusan perusahaan. Salah satu teknik yang dapat digunakan adalah *data mining*, yaitu proses menggali pola, hubungan, dan informasi tersembunyi dari data dalam jumlah besar dengan memanfaatkan berbagai teknik, seperti statistik, *machine learning*, dan pengelolaan basis data (*database*) [7].

Dalam dunia bisnis, *data mining* banyak dimanfaatkan untuk menganalisis perilaku pelanggan, memprediksi penjualan, memberikan rekomendasi produk, serta melakukan segmentasi pelanggan guna mendukung

penyusunan strategi pemasaran yang lebih efektif. Pada penelitian ini, proses segmentasi pelanggan dilakukan menggunakan metode *Cluster Analysis* dengan algoritma K-Means dan pendekatan *Recency, Frequency, Monetary* (RFM) untuk mengelompokkan pelanggan berdasarkan karakteristik transaksi yang dimiliki [8].

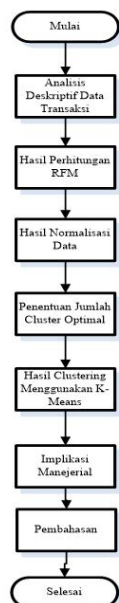
Metode *K-Means* digunakan untuk mengelompokkan pelanggan berdasarkan karakteristik transaksi menggunakan pendekatan *Recency, Frequency*, dan *Monetary* (RFM). *Recency* menunjukkan waktu terakhir pelanggan melakukan transaksi, *Frequency* menunjukkan seberapa sering pelanggan melakukan pembelian, sedangkan *Monetary* menunjukkan total nilai transaksi pelanggan. Melalui pengelompokan tersebut, perusahaan dapat mengetahui pelanggan aktif, pelanggan potensial, maupun pelanggan dengan loyalitas rendah [9]. Mengetahui pola pembelian pelanggan berdasarkan hubungan antar produk yang sering dibeli secara bersamaan. Informasi tersebut dapat dimanfaatkan perusahaan dalam menyusun strategi penjualan seperti bundling produk, rekomendasi produk, maupun strategi promosi yang lebih efektif [10].

Penelitian ini menggunakan data transaksi pelanggan Toko XYZ selama periode tahun 01-04-2025 sampai 31-03-2026. Tahapan penelitian dimulai dengan pengumpulan data transaksi pelanggan yang terdiri dari tanggal transaksi, nama pelanggan, jenis produk, jumlah pembelian, dan total transaksi. Selanjutnya dilakukan proses *Pre - Processing* untuk membersihkan data dari data kosong, data ganda, maupun format data yang tidak sesuai sehingga data siap digunakan dalam proses analisis. [11]. Tahap berikutnya yaitu *Normalisasi Data* dan pengelompokan pelanggan menggunakan metode *K-Means* berdasarkan atribut RFM.

Berdasarkan permasalahan tersebut, penelitian ini dilakukan untuk menganalisis segmentasi pelanggan dan nilai pelanggan menggunakan metode *Cluster Analysis* dan RFM Pada Toko XYZ. Dengan adanya penelitian ini diharapkan perusahaan dapat memahami karakteristik pelanggan, mengetahui pola pembelian pelanggan, serta menentukan strategi pemasaran yang lebih tepat sasaran untuk meningkatkan loyalitas pelanggan dan keuntungan perusahaan.

II. METODE PENELITIAN

Prosedur penelitian yang dilakukan pada penelitian ini disajikan pada gambar 1.



Gambar 1 Prosedur Penelitian

Penelitian ini menggunakan metode kuantitatif dengan pendekatan *Cluster Analysis* menggunakan algoritma *K-Means* dan metode *Recency, Frequency, Monetary* (RFM) untuk mengelompokkan produk berdasarkan karakteristik penjualannya. Data yang digunakan merupakan data transaksi penjualan Toko XYZ selama periode 1 April 2025 hingga 31 Maret 2026 yang meliputi tanggal transaksi, nama produk, jumlah pembelian, dan total penjualan. Tahap pertama penelitian adalah mengumpulkan data transaksi dari database perusahaan, kemudian dilakukan *Pre - Processing* untuk menghapus data yang kosong, duplikat, dan memperbaiki format data agar siap diolah. Setelah itu dilakukan perhitungan nilai *Recency, Frequency*, dan *Monetary (RFM)* menggunakan RStudio. Selanjutnya, nilai RFM dinormalisasi agar seluruh variabel memiliki skala yang sama sehingga hasil pengelompokan menjadi lebih akurat.

Tahap berikutnya adalah menentukan jumlah *cluster* yang paling optimal menggunakan metode *Elbow* dan *Silhouette*. Setelah diperoleh jumlah *cluster* terbaik, proses pengelompokan dilakukan menggunakan algoritma *K-Means*. Hasil pengelompokan kemudian dianalisis untuk mengetahui karakteristik setiap *cluster*, dilanjutkan dengan pembahasan serta penyusunan implikasi manajerial yang dapat dijadikan dasar dalam menentukan strategi pengelolaan produk dan meningkatkan efektivitas penjualan di Toko XYZ.

Setelah data dinormalisasi, dilakukan proses segmentasi pelanggan menggunakan metode *Cluster Analysis* dengan pendekatan algoritma *K-Means*. Algoritma *K-Means* merupakan metode *unsupervised learning* yang sering digunakan dalam proses segmentasi data karena mampu mengelompokkan objek berdasarkan kemiripan

68
karakteristiknya dengan proses yang relatif efisien dan efektif [12].

Metode ini digunakan untuk mengelompokkan pelanggan berdasarkan kemiripan karakteristik transaksi pelanggan sehingga diperoleh beberapa kelompok pelanggan dengan karakteristik yang berbeda. Hasil dari seluruh analisis digunakan sebagai dasar dalam penyusunan strategi *Customer Relationship Management* (CRM) yang dapat membantu perusahaan dalam meningkatkan loyalitas pelanggan dan efektivitas strategi pemasaran.

III. Hasil dan Pembahasan

A. Analisis Deskriptif Data Transaksi

Penelitian ini menggunakan metode *Recency, Frequency*, dan *Monetary* (RFM) sebagai dasar dalam menganalisis pola penjualan produk. Nilai RFM yang diperoleh kemudian diolah menggunakan algoritma *K-Means Clustering* untuk mengelompokkan produk berdasarkan karakteristik penjualannya [13].

[illegible]

Gambar 2 Beberapa Deskriptif Data Transaksi

Berdasarkan Gambar 2, data transaksi yang diperoleh dari Toko XYZ memuat informasi berupa kode transaksi, waktu transaksi (*timestamp*), kategori produk, kode barang, nama barang, jumlah pembelian, harga jual, total penjualan, dan diskon. Setiap kode transaksi dapat terdiri dari beberapa produk yang dibeli pada waktu yang sama, sehingga menunjukkan satu transaksi dapat mencakup lebih dari satu jenis barang. Data ini digunakan sebagai dasar dalam proses analisis, khususnya untuk menghitung nilai *Recency*, *Frequency*, dan *Monetary* (RFM) pada setiap produk sebelum dilakukan pengelompokan menggunakan metode *K-Means*. Data yang digunakan berjumlah 6.593 transaksi produk yang terdiri atas atribut kode transaksi, *timestamp*, kategori, kode barang, nama barang, jumlah pembelian, harga jual, total penjualan, dan diskon.

B. Hasil Perhitungan RFM

Setelah data transaksi penjualan diperoleh, tahap selanjutnya adalah menghitung nilai *Recency*, *Frequency*, dan *Monetary* (RFM) menggunakan perangkat lunak RStudio. Perhitungan dilakukan berdasarkan data transaksi penjualan selama periode 1 April 2025 hingga 31 Maret 2026 untuk setiap jenis produk. Hasil pengolahan menghasilkan data RFM sebanyak 1.204 jenis produk yang selanjutnya digunakan sebagai dasar dalam proses analisis berikutnya.

No	Nama Barang	Last Transaction	Recency	Frequency	Monetary
1	AIR MINERAL ALFA 330 ML	10/12/2025	163	5	16200
2	ARUMI DRESS	14/03/2026	69	5	312500
3	ASI BOOSTER CAPSUL MOM UUNG	28/12/2025	145	25	1750000
4	ATASAN BT BTB BABY	20/09/2025	244	1	12500
5	AVA SET	24/02/2026	87	1	72600
6	Angus Gurita	24/06/2025	332	5	88000
7	Animal Assembly Car	26/10/2025	208	2	135000
8	Atasan BT BTB Baby	23/05/2025	364	1	12500
9	Atlanta Rak Mandi	05/02/2026	106	4	240000
10	BABY BIB KAOS TANGAN & KAKI BABY	27/10/2025	207	9	128400
11	BABY CARE GENDONGAN RANSEL HIPSEAT HORSEY	17/02/2026	94	2	410000
12	BABY CARE GENDONGAN RANSEL HIPST BEE	23/12/2025	150	2	370000
13	BABY CARE GENDONGAN RANSEL HIPST RABBIT	27/01/2026	115	2	380000
14	BABY CARE GENDONGAN SAMPING + APRON SERIES BINTANG	16/11/2025	187	3	405000
15	BABY CARE GENDONGAN SAMPING G-APRON BEE	28/09/2025	236	2	258000
16	BABY CARE GENDONGAN SAMPING RING SLING	25/03/2026	58	1	110000
17	BABY CARE GENDONGAN SAMPING SAILOR	03/08/2025	292	2	258000
18	BABY CARE GENDONGAN SAMPING SLING HORSEY	31/03/2026	52	2	246000
19	BABY CARE KASUR KOLAM & KELAMBU SAILOR SERIES	29/11/2025	174	4	1120000
20	BABY CARE KASUR MATRAS BAYI & KELAMBU MOTIF	14/12/2025	159	7	1540000
21	BABY CARE KASUR MATRAS BAYI MOTIF	04/02/2026	107	1	117000
22	BABY CARE SOFA BAYI	24/05/2025	363	3	397500
23	BABY CARE SOFA BAYI BEE SERIES	30/11/2025	173	3	361500

Gambar 3 Beberapa Hasil Perhitungan RFM

Output perhitungan RFM terdiri atas beberapa atribut, yaitu nama barang, tanggal transaksi terakhir (*Last Transaction*), nilai *Recency*, *Frequency*, dan *Monetary*. Atribut *Last Transaction* menunjukkan tanggal terakhir suatu produk terjual selama periode penelitian [14].

Menghitung *recency* yakni dengan mengubah *TimeStamp* ke dalam Format *POSIXct* pada Rstudio, kemudian menentukan Tanggal Analisis pada tanggal (2026-05-22). Untuk periode nya mulai tanggal 01-04-2025 sampai 31-03-2026. Untuk menentukan *Last Transaction* dengan menghitung (*Max Tanggal*) mengambil transaksi terakhir setiap produk dan (*difftime*) menghitung selisih waktu *Units (days)* pada Rstudio. Menghitung *frequency* yakni dengan jumlah transaksi atau jumlah produk yang terjual untuk setiap jenis produk selama periode satu tahun. Menghitung *Monetary* yakni dengan menghitung jumlah barang, harga jual, total dan diskon setiap produk.

C. Hasil Normalisasi Data

Tahap berikutnya dalam penelitian ini adalah melakukan normalisasi terhadap nilai RFM. Normalisasi dilakukan agar skala data pada atribut *recency*, *frequency*, dan *monetary* tidak memiliki perbedaan yang terlalu besar, sehingga proses *Clustering* dapat berjalan lebih optimal. Proses normalisasi dilakukan dengan cara membagi setiap nilai pada atribut RFM dengan nilai terbesar pada masing-masing atribut tersebut [15]. Setelah dilakukan perhitungan nilai *Recency*, *Frequency*, dan *Monetary* (RFM) untuk setiap produk, tahap selanjutnya adalah melakukan normalisasi data. Normalisasi merupakan proses mengubah nilai data yang memiliki skala berbeda menjadi skala yang sama agar setiap variabel memiliki pengaruh yang seimbang pada saat dilakukan proses clustering menggunakan metode *K-Means*. Pada penelitian ini, variabel *Recency*, *Frequency*, dan *Monetary* memiliki satuan yang berbeda. Variabel *Recency* dinyatakan dalam satuan hari, *Frequency* dinyatakan dalam jumlah transaksi, sedangkan *Monetary* dinyatakan dalam nilai penjualan (rupiah). Perbedaan skala tersebut dapat menyebabkan variabel dengan nilai yang lebih besar, seperti *Monetary*, mendominasi proses pengelompokan. Oleh karena itu, normalisasi perlu dilakukan agar seluruh variabel berada pada rentang nilai yang sama [16].

Untuk mengatasi bias ini, standarisasi data (*Z-Score Normalization*) diterapkan pada ketiga variabel RFM (1).

$$Z = \frac{X - \mu}{\sigma} \quad (1)$$

No	Nama Barang	Last Transaction	Recency	Normalisasi	Frequency	Normalisasi	Monetary	Normalisasi	Cluster
1	AIR MINERAL ALFA 330 ML	10/12/2025	163	-0,27947015	5	-0,07018013	16200	-0,467570914	1
2	ARUMI DRESS	14/03/2026	69	-1,20360131	5	-0,07018013	312500	0,542004086	1
3	ASI BOOSTER CAPSUL MOM UUNG	28/12/2025	145	-0,44841101	25	0,85858586	1750000	3,26968544	2
4	ATASAN BT BTB BABY	20/09/2025	244	0,54858899	1	-0,66092518	12500	-0,47825213	3
5	AVA SET	24/02/2026	87	-0,60908402	1	-0,66092518	72600	-0,51121789	1
6	Angus Gurita	24/06/2025	332	1,38119971	5	-0,07018013	88000	-0,203668493	3
7	Animal Assembly Car	26/10/2025	208	-0,16301307	2	-0,33235811	135000	-0,180181812	3
8	Atasan BT BTB Baby	23/05/2025	364	0,66092518	1	-0,66092518	12500	-0,47825213	3
9	Atlanta Rak Mandi	05/02/2026	106	-0,27947015	4	-0,28235811	240000	0,37736281	1
10	BABY BIB KAOS TANGAN & KAKI BABY	27/10/2025	207	-0,15110139	9	0,30259110	128400	-0,176125228	1
11	BABY CARE GENDONGAN RANSEL HIPSEAT HORSEY	17/02/2026	94	-0,47825213	2	-0,33235811	410000	0,481110086	1
12	BABY CARE GENDONGAN RANSEL HIPST BEE	23/12/2025	150	-0,48727592	2	-0,33235811	370000	0,388180879	1
13	BABY CARE GENDONGAN RANSEL HIPST RABBIT	27/01/2026	115	-0,71146651	2	-0,33235811	380000	0,412541428	1
14	BABY CARE GENDONGAN SAMPING + APRON SERIES BINTANG	16/11/2025	187	-0,04841101	3	-0,45858586	405000	0,247433110	1
15	BABY CARE GENDONGAN SAMPING G-APRON BEE	28/09/2025	236	0,43810818	2	-0,33235811	258000	0,117408848	3
16	BABY CARE GENDONGAN SAMPING RING SLING	25/03/2026	58	-0,71146651	1	-0,66092518	110000	-0,466426466	1
17	BABY CARE GENDONGAN SAMPING SAILOR	03/08/2025	292	0,98475241	2	-0,33235811	258000	0,117408848	3
18	BABY CARE GENDONGAN SAMPING SLING HORSEY	31/03/2026	52	-1,37073141	2	-0,33235811	246000	-0,083795442	1
19	BABY CARE KASUR KOLAM & KELAMBU SAILOR SERIES	29/11/2025	174	-0,11127114	4	-0,17487056	1120000	2,02828151	1
20	BABY CARE KASUR MATRAS BAYI & KELAMBU MOTIF	14/12/2025	159	-0,18179488	7	0,22527108	1540000	3,128202718	2
21	BABY CARE KASUR MATRAS BAYI MOTIF	04/02/2026	107	-0,80016197	1	-0,66092518	117000	-0,242779657	1
22	BABY CARE SOFA BAYI	24/05/2025	363	0,66092518	3	-0,35545410	397500	0,414898745	3
23	BABY CARE SOFA BAYI BEE SERIES	30/11/2025	173	-0,18118302	3	-0,35545410	361500	-0,257040456	1
24	BABY CARE SOFA BAYI RABBIT SERIES	23/10/2025	211	0,19242602	3	-0,35545410	428000	0,133848006	3
25	BABY CARE SOFA BAYI RABBIT SLEET	19/10/2025	215	0,21771118	3	-0,35545410	397500	0,414898745	3
26	BABY CARE TAS BESAR MOON LIGHT	08/11/2026	134	-0,56457444	4	-0,27051401	825000	1,489112922	1
27	BABY CARE TAS BESAR MOON LIGHT	24/12/2025	149	-0,41710070	3	-0,35545410	450000	0,581910097	1
28	BABY CARE TAS BESAR KANGAR	19/10/2026	64	-1,25727732	3	-0,35545410	495000	-0,609775084	1
29	BABY CARE TAS KECIL CLAMOR SERIES	27/07/2024	299	1,65775669	3	-0,35545410	244500	0,084748629	3

Gambar 4 Beberapa Hasil Normalisasi Data

Sebagai contoh pada *recency* digunakan produk ARUMI DRESS yang memiliki nilai *Recency* sebesar 69 hari. Berdasarkan hasil perhitungan pada RStudio, nilai normalisasi (*Z-Score*) yang diperoleh adalah -1,20360131. Hasil perhitungan tersebut sama dengan nilai normalisasi pada tabel hasil RStudio, yaitu -1,20360131. Nilai negatif menunjukkan bahwa nilai *Recency* produk ARUMI DRESS berada di bawah rata-rata seluruh produk, yang berarti produk tersebut memiliki waktu transaksi terakhir yang lebih baru dibandingkan rata-rata produk lainnya.

Sebagai contoh digunakan produk ASI BOOSTER CAPSUL MOM UUNG yang memiliki *Frequency* sebanyak 25 transaksi. Berdasarkan hasil normalisasi pada RStudio diperoleh nilai *Z-Score* sebesar 2,88363562. Nilai tersebut menunjukkan bahwa *frekuensi* penjualan produk ASI BOOSTER CAPSUL MOM UUNG berada jauh di atas rata-rata seluruh produk sehingga produk tersebut termasuk produk yang sering terjual.

Sebagai contoh digunakan produk ASI BOOSTER CAPSUL MOM UUNG dengan nilai *Monetary* sebesar Rp1.750.000. Berdasarkan hasil normalisasi diperoleh nilai *Z-Score* sebesar 3,726968544. Hasil tersebut sesuai dengan nilai normalisasi pada tabel RStudio, yaitu 3,726968544. Nilai positif yang tinggi menunjukkan bahwa produk tersebut memberikan kontribusi penjualan yang jauh lebih besar dibandingkan rata-rata seluruh produk [17].

D. Penentuan Jumlah Cluster Optimal

Penentuan jumlah *cluster* (K) yang optimal pada algoritma *K-Means* umumnya dilakukan menggunakan dua metode yang saling mendukung, yaitu *Elbow Method* dan *Silhouette Score*. *Elbow Method* menentukan jumlah *cluster* berdasarkan nilai *Within-Cluster Sum of Squares* (WCSS), yaitu total kuadrat jarak antara setiap data dengan pusat *cluster* (*centroid*). Nilai K yang dipilih berada pada titik ketika penurunan WCSS mulai melambat atau membentuk pola menyerupai siku (*elbow*). Sementara itu, *Silhouette Score* digunakan untuk mengevaluasi kualitas hasil pengelompokan dengan mengukur tingkat kesamaan suatu data terhadap *cluster*-nya sendiri dibandingkan dengan *cluster* lain. Nilai *Silhouette Score* berada pada rentang -1 hingga 1, di mana nilai yang semakin mendekati 1 menunjukkan bahwa pemisahan antar *cluster* semakin baik dan hasil pengelompokan semakin optima [18].

Metode ini menjadi salah satu algoritma yang paling sering digunakan dalam proses pengelompokan data, terutama untuk klasifikasi pelanggan maupun produk. Huruf K pada *K-Means* menunjukkan jumlah *Cluster* yang

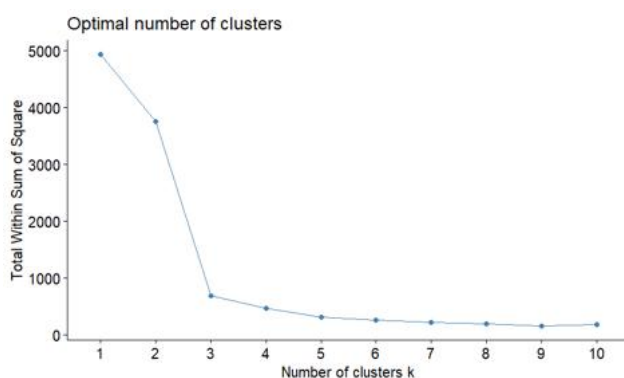
telah ditentukan sebelumnya, sedangkan *K-Means* merujuk pada nilai rata-rata data pada masing-masing *Cluster*

Pada metode *K-Means Clustering*, langkah pertama yang dilakukan adalah menentukan jumlah *cluster* yang akan dibentuk. Setelah jumlah *cluster* ditetapkan, sistem akan membentuk *centroid* sebagai titik pusat awal pada setiap *cluster*. Selanjutnya, setiap data dihitung jaraknya terhadap seluruh *centroid*, kemudian data akan dikelompokkan ke dalam *cluster* yang memiliki jarak paling dekat. Setelah seluruh data teralokasi, nilai *centroid* akan dihitung kembali berdasarkan rata-rata anggota pada masing-masing *cluster*. Proses ini dilakukan secara berulang (*iteratif*) hingga tidak terjadi lagi perpindahan data antar *cluster* atau posisi *centroid* telah stabil. Untuk menentukan kedekatan antara suatu data dengan *centroid*, penelitian ini menggunakan *Euclidean Distance*, yaitu metode perhitungan jarak yang mengukur kedekatan setiap data terhadap pusat *cluster* sehingga proses pengelompokan dapat dilakukan secara lebih akurat (2).

$$D = \sum_{j=1}^p (X_{ij} - X_{hj})^2 \quad (2)$$

TABEL I
TOTAL WITHIN SUM OF SQUARE

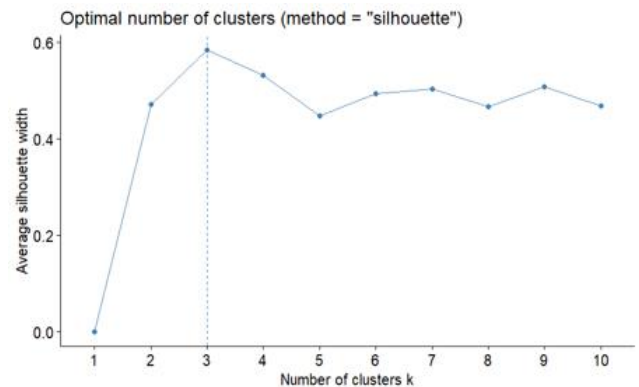
NUMBER OF CLUSTER (K)	TOTAL WITHIN SUM OF SQUARE (WSS)
1	4.950
2	3.750
3	700
4	500
5	350
6	290
7	250
8	220



Gambar 5 Grafik Nilai Elbow

Berdasarkan grafik *Elbow Method*, diperoleh jumlah *Cluster* optimal sebanyak 3 *Cluster*. Hal ini ditunjukkan oleh adanya titik siku pada nilai $k = 3$, di mana setelah titik tersebut penurunan nilai WSS cenderung melandai. Oleh karena itu, jumlah *Cluster* yang digunakan dalam penelitian ini adalah 3 *Cluster* karena dinilai mampu merepresentasikan karakteristik data dengan baik [19].

Selanjutnya dilakukan analisis *K-Means* menggunakan *Silhouette*. Metode *K-Means* memiliki keunggulan dalam kemudahan implementasi serta kompleksitas waktu dan ruang yang relatif rendah. Dengan efisiensi komputasi yang dimilikinya, algoritma ini mampu menghasilkan proses pengelompokan data yang cukup optimal [20].



Gambar 6 Grafik Nilai Silhouette

Berdasarkan grafik *Silhouette Method*, nilai *average silhouette width* tertinggi diperoleh pada $k = 3$ sebesar kurang lebih 0,59. Hal ini menunjukkan bahwa pengelompokan data menjadi 3 *Cluster* menghasilkan kualitas *Cluster* yang paling baik. Oleh karena itu, jumlah *Cluster* yang digunakan dalam penelitian ini adalah 3 *Cluster*, karena didukung oleh hasil metode *Elbow* dan *Silhouette* yang sama-sama menunjukkan nilai optimal pada $k = 3$ [21]. Ini membuat kita berasumsi bahwa $k = 3$ adalah pilihan yang tepat untuk data ini. Rata-rata $s(i)$ terbaik disebut koefisien *Silhouette* (SC) [22].

E. Hasil Cluster Menggunakan K-Means

Teknik *clustering* merupakan salah satu metode penting dalam *data mining* yang digunakan untuk mengelompokkan data berdasarkan tingkat kemiripan karakteristik yang dimiliki. Melalui metode ini, data yang memiliki karakteristik serupa akan ditempatkan pada kelompok (*cluster*) yang sama, sedangkan data dengan karakteristik yang berbeda akan dikelompokkan ke dalam *cluster* yang berbeda. Salah satu algoritma *clustering* yang paling banyak digunakan adalah *K-Means*. Algoritma ini bekerja dengan membagi data ke dalam sejumlah *cluster* yang telah ditentukan sebelumnya, kemudian mengelompokkan setiap data berdasarkan kedekatannya dengan titik pusat (*centroid*) sehingga jarak antar data dalam satu *cluster* menjadi sekecil mungkin. Berkat prosesnya yang sederhana, efisien dalam mengolah data berukuran besar, serta mampu menghasilkan pengelompokan yang baik, algoritma *K-Means* banyak diterapkan pada berbagai bidang, seperti segmentasi

pelanggan dalam bisnis dan pemasaran, analisis citra untuk pengenalan pola, serta pengolahan citra untuk pengenalan pola, serta pengolahan gambar [23].

Berdasarkan proses *Clustering* menggunakan metode *K-Means*, pada Tabel I Hasil *Clustering* Menggunakan *K-Means* seluruh data produk berhasil dikelompokkan ke dalam tiga *cluster* berdasarkan nilai *Recency*, *Frequency*, dan *Monetary* (RFM).

TABEL II
HASIL *CLUSTER* MENGGUNAKAN *K-MEANS*

CLUSTER	JUMLAH PRODUK	KARATERISTIK	DESKRIPSI
1	630	<i>BEST SELLER PRODUCT</i>	PRODUK YANG PALING SERING TERJUAL MASIH AKTIF DIBELI, DAN MEMBERIKAN KONTRIBUTSI PENDAPATAN TERBESAR. PRODUK INI MENJADI PRIORITAS UTAMA DALAM MENJAGA KETERSEDIAAN STOK
2	90	<i>SLOW MOVING PRODUCT</i>	PRODUK YANG SUDAH LAMA TIDAK TERJUAL, MEMILIKI FREKUENSI PENJUALAN KECIL. PRODUK INI MEMERLUKAN EVALUASI MELALUI PROMOSI, DISKON, BUDLING, ATAU PENGURANGAN STOK
3	484	<i>POTENTIAL PRODUCT</i>	PRODUK YANG MASIH CUKUP AKTIF TERJUAL DENGAN FREKUENSI DAN NILAI PENJUALAN SEDANG. PRODUK INI MEMILIKI POTENSI UNTUK BERKEMBANG MENJADI PRODUK UNGGULAN APABILA DIDUKUNG

Hasil pengelompokan menunjukkan bahwa *Cluster* 1 merupakan kelompok dengan jumlah anggota terbanyak, yaitu 630 produk atau sekitar 52,3% dari total data. Hal ini menunjukkan bahwa sebagian besar produk memiliki karakteristik yang serupa berdasarkan nilai RFM sehingga tergabung ke dalam *Cluster* 1.

Selanjutnya, *Cluster* 3 terdiri dari 484 produk atau sekitar 40,2% dari total data. Jumlah tersebut menunjukkan bahwa cukup banyak produk yang memiliki karakteristik berbeda dengan *Cluster* 1, tetapi masih membentuk

kelompok yang relatif besar. Sementara itu, *Cluster* 2 merupakan kelompok dengan jumlah anggota paling sedikit, yaitu 90 produk atau sekitar 7,5% dari keseluruhan data. Meskipun jumlah anggotanya lebih sedikit, *Cluster* 2 dapat menunjukkan kelompok produk dengan karakteristik yang lebih spesifik dibandingkan dua *cluster* lainnya [24].

Berdasarkan Tabel nilai rata-rata *recency*, *frequency*, dan *monetary* sebagai berikut

TABEL III
NILAI RATA-RATA *RECENCY*, *FREQUENCY*, DAN *MONETARY*

VARIABEL	RATA-RATA	INTERPRETASI
<i>RECENCY</i>	191,43 HARI	RATA-RATA SETIAP PRODUK TERKHIR TERJUAL SEKITAR 191 HARI SEBELUM TANGGAL 2026-05-22. SEMAKIN KECIL NILAI <i>RECENCY</i> , SEMAKIN BARU PRODUK TERSEBUT TERJUAL
<i>FREQUENCY</i>	5,48 KALI	RATA-RATA SETIAP PRODUK TERJUAL SEBANYAK 5-6 KALI SELAM PERIODE PENGAMATAN. SEMAKIN TINGGI NILAI <i>FREQUENCY</i> , SEMAKIN SERING PRODUK DIBELI PELANGGAN
<i>MONETARY</i>	RP. 209.469	RATA-RATA TOTAL NILAI PENJUALAN YANG DIHASILKAN SETIAP PRODUK ADALAH SEKITAR RP 209.469 SELAMA PERIODE PENELITIAN. SEMAKIN BESAR NILAI <i>MONETARY</i> , SEMAKIN BESAR KONTRIBUTSI PENDAPATAN DARI PRODUK TERSEBUT

Berdasarkan hasil perhitungan RFM, diperoleh rata-rata nilai *Recency* sebesar 191,43 hari, yang menunjukkan bahwa secara rata-rata produk terakhir kali terjual sekitar 191 hari sebelum tanggal 2026-05-22. Nilai *recency* yang lebih kecil menunjukkan bahwa suatu produk lebih baru terjual, sedangkan nilai yang lebih besar menunjukkan bahwa produk tersebut sudah cukup lama tidak mengalami transaksi.

Nilai rata-rata *Frequency* sebesar 5,48 kali menunjukkan bahwa setiap produk rata-rata terjual sebanyak lima hingga enam kali selama periode pengamatan. Produk dengan nilai *frequency* yang tinggi memiliki tingkat permintaan yang lebih baik dibandingkan produk dengan *frequency* rendah.

Sementara itu, rata-rata *Monetary* sebesar Rp209.469 menunjukkan bahwa setiap produk memberikan kontribusi pendapatan rata-rata sebesar Rp209.469 selama periode penelitian. Semakin tinggi nilai *monetary*, maka semakin besar kontribusi produk terhadap total penjualan toko.

F. Implikasi Manejerial Cluster

Hasil segmentasi produk memberikan informasi yang dapat dimanfaatkan oleh pihak Toko XYZ dalam menyusun strategi pengelolaan produk secara lebih efektif.

Produk pada *Cluster* 1 sebaiknya menjadi prioritas utama dalam pengelolaan persediaan karena memiliki tingkat penjualan dan kontribusi pendapatan yang tinggi. Ketersediaan stok perlu dijaga agar tidak terjadi *stock out* yang dapat mengurangi peluang penjualan.

Produk pada *Cluster* 2 memerlukan evaluasi lebih lanjut karena memiliki tingkat penjualan yang rendah dan sudah cukup lama tidak terjual. Toko dapat menerapkan strategi seperti pemberian diskon, promosi khusus, bundling dengan produk lain, atau mengurangi jumlah pemesanan agar tidak terjadi penumpukan stok.

Produk pada *Cluster* 3 memiliki performa penjualan yang cukup stabil sehingga masih berpotensi untuk ditingkatkan. Toko dapat melakukan promosi secara berkala, memperbaiki penempatan produk (*display*), atau memberikan rekomendasi produk kepada pelanggan agar penjualannya meningkat.

Selain itu, hasil penelitian ini dapat dijadikan sebagai dasar dalam menentukan kebijakan pengelolaan persediaan, perencanaan pembelian produk, serta penyusunan strategi pemasaran yang lebih tepat sasaran. Dengan memanfaatkan hasil segmentasi produk, Toko XYZ dapat mengoptimalkan pengelolaan stok, meningkatkan efektivitas promosi, dan mendukung peningkatan kinerja penjualan secara berkelanjutan.

G. Pembahasan

Hasil penelitian menunjukkan bahwa metode *Recency*, *Frequency*, dan *Monetary* (RFM) mampu menggambarkan karakteristik penjualan setiap produk berdasarkan waktu penjualan terakhir, *frekuensi* penjualan, dan nilai penjualan. Nilai RFM kemudian digunakan sebagai dasar dalam proses pengelompokan menggunakan algoritma *K-Means* sehingga produk yang memiliki karakteristik penjualan yang serupa dapat ditempatkan pada kelompok yang sama.

Penentuan jumlah *Cluster* dilakukan menggunakan metode *Elbow* dan *Silhouette*. Kedua metode tersebut menunjukkan bahwa jumlah *Cluster* yang paling optimal adalah 3 *Cluster*, sehingga pengelompokan mampu merepresentasikan karakteristik data dengan baik. Hasil ini menunjukkan bahwa penggunaan algoritma *K-Means* dapat menghasilkan segmentasi produk yang jelas berdasarkan pola penjualannya.

Berdasarkan hasil analisis *centroid*, *Cluster* 1 memiliki nilai *Frequency* dan *Monetary* yang paling tinggi serta nilai *Recency* yang rendah. Hal ini menunjukkan bahwa produk dalam *Cluster* tersebut merupakan produk yang sering terjual, masih aktif dibeli, dan memberikan kontribusi pendapatan terbesar bagi toko. Produk dalam *Cluster* ini dapat dikategorikan sebagai produk unggulan yang memiliki tingkat permintaan tinggi.

Sementara itu, *Cluster* 2 memiliki nilai *Recency* yang tinggi serta nilai *Frequency* dan *Monetary* yang rendah. Karakteristik tersebut menunjukkan bahwa produk dalam *Cluster* ini sudah cukup lama tidak terjual dan

memiliki tingkat penjualan yang rendah. Kondisi tersebut mengindikasikan bahwa produk memerlukan perhatian khusus agar tidak menjadi stok yang terlalu lama tersimpan.

Selanjutnya, *Cluster* 3 memiliki nilai *Recency* yang rendah dengan nilai *Frequency* dan *Monetary* yang berada di sekitar rata-rata. Hal ini menunjukkan bahwa produk pada *Cluster* tersebut masih memiliki aktivitas penjualan yang cukup baik meskipun kontribusi penjualannya belum sebesar *Cluster* 1. Produk pada *Cluster* ini masih memiliki peluang untuk dikembangkan melalui strategi pemasaran yang tepat sehingga dapat meningkatkan penjualannya. Secara keseluruhan, hasil penelitian menunjukkan bahwa kombinasi metode RFM dan *K-Means Clustering* mampu mengelompokkan produk berdasarkan karakteristik penjualannya sehingga memberikan informasi yang bermanfaat dalam proses pengambilan keputusan di Toko XYZ.

IV. Kesimpulan

Berdasarkan hasil penelitian, penerapan metode *Recency*, *Frequency*, dan *Monetary* (RFM) yang dikombinasikan dengan algoritma *K-Means Clustering* berhasil digunakan untuk melakukan segmentasi produk di Toko XYZ berdasarkan karakteristik penjualannya. Penelitian ini menggunakan data transaksi penjualan sebanyak 6.593 transaksi yang terdiri dari 1.204 jenis produk selama periode 1 April 2025 hingga 31 Maret 2026. Nilai RFM yang diperoleh dari setiap produk kemudian dinormalisasi sebelum dilakukan proses pengelompokan menggunakan algoritma *K-Means*.

Penentuan jumlah *cluster* dilakukan menggunakan metode *Elbow* dan *Silhouette*, yang sama-sama menunjukkan bahwa jumlah *cluster* yang paling optimal adalah tiga *cluster*. Hasil metode *Silhouette* menghasilkan nilai *average silhouette width* sebesar 0,59, yang menunjukkan bahwa kualitas pengelompokan cukup baik sehingga mampu merepresentasikan karakteristik data produk secara optimal.

Hasil pengelompokan menunjukkan bahwa *Cluster* 1 terdiri dari 630 produk (52,3%) dan merupakan kelompok *Best Seller Product*, yaitu produk yang paling sering terjual, masih aktif dibeli, serta memberikan kontribusi penjualan terbesar. *Cluster* 2 terdiri dari 90 produk (7,5%) dan termasuk *Slow Moving Product*, yaitu produk dengan frekuensi penjualan rendah, sudah lama tidak terjual, serta memerlukan evaluasi melalui strategi seperti promosi, diskon, atau pengurangan stok. Sementara itu, *Cluster* 3 terdiri dari 484 produk (40,2%) dan dikategorikan sebagai *Potential Product*, yaitu produk yang memiliki performa penjualan cukup baik dan masih berpotensi menjadi produk unggulan apabila didukung dengan strategi pemasaran yang tepat.

Secara keseluruhan, hasil penelitian menunjukkan bahwa kombinasi metode RFM dan *K-Means Clustering* mampu menghasilkan segmentasi produk yang jelas berdasarkan pola penjualannya. Informasi yang dihasilkan dapat dimanfaatkan oleh Toko XYZ sebagai dasar dalam pengambilan keputusan, seperti menentukan prioritas pengelolaan stok, menyusun strategi promosi yang lebih tepat sasaran, mengevaluasi produk dengan performa

rendah, serta mengoptimalkan penjualan produk secara berkelanjutan. Dengan demikian, hasil segmentasi tidak hanya membantu memahami karakteristik setiap kelompok produk, tetapi juga mendukung peningkatan efektivitas pengelolaan persediaan dan strategi pemasaran perusahaan.

DAFTAR RUJUKAN

- [1] Z. A. Rachman, S. A. Putra, A. R. Herdiyanto Darma, and A. Pratiwi Lintang, "ANALISIS FAKTOR YANG MEMPENGARUHI CUSTOMER RELATIONSHIP MANAGEMENT (CRM) TERHADAP KEPUASAN DAN LOYALITAS PELANGGAN," *Jurnal Teknologi Informasi*, vol. 4, no. 2, Dec. 2023, doi: 10.46576/djtechno.
- [2] J. Naimah, S. N. Saputri, M. F. Aunillah, and E. M. Hanesti, "Penerapan Customer Relationship Management (CRM) Sebagai Upaya Untuk Meningkatkan Loyalitas Pelanggan : Studi Kasus pada Toko SRC Mulia Mojokerto," Jul. 2024. doi: <https://doi.org/10.30996/jeb17.v9i01.10848>.
- [3] D. W. Mandagi, M. Manajemen, F. Ekonomi, and D. Bisnis, "PEMASARAN MEDIA SOSIAL, GESTALT MEREK DAN LOYALITAS PELANGGAN: STUDI EMPIRIS PADA TOMOHON INTERNATIONAL FLOWER FESTIVAL," 2023. [Online]. Available: <http://jurnal.ustjogja.ac.id>
- [4] Mutia Hannisyah Putri Lubis, Havid Syafwan, and Pristiyanilicia Putri, "Implementasi E-CRM dalam Mempertahankan Pelanggan Pada Toko Key Distro Kisaran," *Jurnal Ilmiah Multidisiplin Nusantara (JIMNU)*, vol. 1, no. 2, pp. 51–57, Jul. 2023, doi: 10.59435/jimnu.v1i2.108.
- [5] V. Guerola-Navarro, H. Gil-Gomez, R. Oltra-Badenes, and P. Soto-Acosta, "Customer relationship management and its impact on entrepreneurial marketing: a literature review," *International Entrepreneurship and Management Journal*, vol. 20, no. 2, pp. 507–547, Jun. 2024, doi: 10.1007/s11365-022-00800-x.
- [6] M. A. T. Wahyudi, A. Hermawan, and M. Ilham, "Penerapan Customer Relationship Management (CRM) bagi UMKM," Kediri,Malang,Surabaya, May 2024. doi: <https://doi.org/10.65255/jibma.v2i2.43>.
- [7] M. Risnasari, N. Aulia, and L. Cahyani, "Clustering Of Student Learning Styles in the Industry 4.0 Using KMeans Algorithm," *Jurnal Teknologi Pendidikan*, vol. 24, no. Vol. 24 No. 2 (2022): Jurnal Teknologi Pendidikan, pp. 246–257, Aug. 2022, doi: 10.21009/JTP2001.6.
- [8] A. M. Bleszynski, D. Alya, and K. Saadatunnisa, "Pemanfaatan Data Mining untuk Segmentasi Pelanggan pada Penjualan Produk Elektronik Menggunakan Metode RFM dan K-Means," Jun. 2026.
- [9] S. Ika Murpratiwi, I. Gusti Agung Indrawan, and A. Aranta, "ANALISIS PEMILIHAN CLUSTER OPTIMAL DALAM SEGMENTASI PELANGGAN TOKO RETAIL," *Jurnal Pendidikan Teknologi dan Kejuruan*, vol. 18, no. 2, Jul. 2021, doi: <https://doi.org/10.23887/jptk-undiksha.v18i2.37426>.
- [10] S. Aulia Miranda, F. Fahrullah, and D. Kurniawan, "Implementasi Association Rule Dalam Menganalisis Data Penjualan Sheshop dengan Menggunakan Algoritma Apriori," *METIK JURNAL*, vol. 6, no. 1, pp. 30–36, Jul. 2022, doi: 10.47002/metik.v6i1.342.
- [11] C. Lumbantoruan and D. A. Megawaty, "Penerapan Metode K-Means Clustering dan Principal Component Analysis (PCA) untuk Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Pendidikan," *Journal of Artificial Intelligence and Technology Information (JAITI)*, vol. 4, no. 2, pp. 120–131, Jun. 2026, doi: 10.58602/jaiti.v4i2.256.
- [12] A. Rahmawati, A. Hananto, and F. Nurapriani, "Analisis Segmentasi Pelanggan Menggunakan K-Means Clustering Untuk Optimalisasi Penjualan Sembako,"
- [13] Y. K. Febrina, R. Saputra, and K. F. G, "Segmentasi Pelanggan Toko Hanifah Berdasarkan Analisis RFM dengan Metode K-Means Clustering," *Jurnal Pustaka AI (Pusat Akses Kajian Teknologi Artificial Intelligence)*, vol. 5, no. 2, pp. 274–282, Aug. 2025, doi: 10.55382/jurnalpustakaai.v5i2.1084.
- [14] F. Sarimole Matheos and L. Hakim, "Klasifikasi Barang Menggunakan Metode Clustering K-Means Dalam Penentuan Prediksi Stok Barang," *Jurnal Sains dan Teknologi*, vol. 5, Feb. 2024, doi: 10.55338/saintek.v5i1.2709.
- [15] M. Fithriyah, M. A. Yaqin, and S. Zaman, "K-Means Clustering Untuk Segmentasi Produk Berdasarkan Analisis Recency, Frequency, Monetary (RFM) Pada Data Transaksi Penjualan," *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 3, no. 2, pp. 151–164, Aug. 2021, doi: 10.28926/ilkomnika.v3i2.284.
- [16] S. F. Djun, I. G. A. Gunadi, and S. Sariyasa, "Analisis Segmentasi Pelanggan pada Bisnis dengan Menggunakan Metode K-Means Clustering pada Model Data RFM," *JTIM : Jurnal Teknologi Informasi dan Multimedia*, vol. 5, no. 4, pp. 354–364, Feb. 2024, doi: 10.35746/jtim.v5i4.434.
- [17] L. C. Fahsya, C. Wijaya, F. M. Bintang, J. J. Mulyono, F. Ramadhan, and F. Amsury, "PENERAPAN CLUSTERING K-MEANS UNTUK SEGMENTASI PELANGGAN PADA BISNIS RETAIL," *HOAQ :Jurnal Teknologi Informasi*, vol. 17, no. Vol. 17 No. 1 (2026): Jurnal HOAQ-Teknologi Informasi, pp. 38–52, May 2026, doi: 10.52972/hoaq.voll7no1.

- [18] B. R. Fitra, G. Aprilian, and R. R. Pratama, "Segmentasi Pelanggan E-Commerce Menggunakan Pendekatan RFM (Recency, Frequency, Monetary) dan Algoritma K-Means Clustering," Kota Tangerang, Jun. 2026. [Online]. Available: <https://journal.jci.co.id/jisbt>
- [19] Duhita Maulida and M.Dachyar, "Customer Loyalty Strategy Using Customer Lifetime Value (A Case Study of Jamu and Herbal Products SME)," Depok, Indonesia, Mar. 2022.
- [20] F. Dwi Agustiar, B. Nurina Sari, and I. Maulana, "PENERAPAN DATA MINING UNTUK PENGELOMPOKAN PRODUK PENJUALAN MENGGUNAKAN ALGORITMA K-MEANS (STUDI KASUS : TOKO AGUNG MAKMUR JAYA)," Feb. 2025.
- [21] M.Dachyar and Lidya Vitasya, "Customer Loyalty Analysis Using Customer Lifetime Value (A Case Study of Baby Equipment SMEs)," Depok, Indonesia, Aug. 2021.
- [22] K. LEONARD and R. PETER J., "Finding Groups in Data An Introduction to Cluster Analysis," Belgium.
- [23] F. Khalish, N. M. Piranti, and O. Martadireja, "Implementasi Data Mining Menggunakan Teknik Clustering dengan Metode K-Means," May 2025. doi: <https://doi.org/10.54371/jiip.v8i5.7874>.
- [24] B. Hartono, S. Eniyati, and K. Hadiono, "Perbandingan Metode Perhitungan Jarak pada Nilai Centroid dan Pengelompokan Data Menggunakan K-Means Clustering," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 3, p. 503, Mar. 2023, doi: [10.30865/json.v4i3.6021](https://doi.org/10.30865/json.v4i3.6021).