

# Donor Segmentation Analysis Using the RFM Model and K-Means Clustering to Optimize Fundraising Strategies

Rezki <sup>1\*</sup>, Nouval Trezandy Lapatta <sup>2\*\*</sup>, Rizka Ardiansyah <sup>3\*</sup>, Wirdayanti <sup>4\*\*</sup>, Dwi Shinta Angreni <sup>5\*</sup>

<sup>1,2</sup> Teknik Informatika, Universitas Tadulako

<sup>3,4,5</sup> Sistem Informasi, Universitas Tadulako

[rezkiti19@gmail.com](mailto:rezkiti19@gmail.com) <sup>1</sup>, [nouval@untad.ac.id](mailto:nouval@untad.ac.id) <sup>2</sup>, [rizka@untad.ac.id](mailto:rizka@untad.ac.id) <sup>3</sup>, [wirda\\_arbie@untad.ac.id](mailto:wirda_arbie@untad.ac.id) <sup>4</sup>, [ds.angreni@untad.ac.id](mailto:ds.angreni@untad.ac.id) <sup>5</sup>

## Article Info

### Article history:

Received 2024-09-09

Revised 2024-09-17

Accepted 2024-09-18

### Keyword:

Clustering,  
Elbow method,  
K-means algorithm,  
Recency frequency monetary.

## ABSTRACT

This study aims to segment donors using the Recency, Frequency, Monetary (RFM) model and the K-Means algorithm to optimize fundraising strategies. The RFM model is used to measure donor engagement through three dimensions: Recency (the last time a donation was made), Frequency (the frequency of donations), and Monetary (the amount of donations). By utilizing RFM scores, donors are then grouped using the K-means algorithm to generate more specific donor segments. This study was conducted using donation data from a non-profit organization, focusing on strategies to improve donor loyalty and donation frequency. The segmentation results identified several key segments, including Loyal Donors, New Donors, Potential Donors, and Low-Priority Donors. Each segment exhibits different donation behavior characteristics and requires a different strategic approach. The implementation of these segmentation results is expected to help the organization design more effective communication strategies and donation programs, as well as improve donor retention and lifetime value. Additionally, this study identifies the potential for enhancing the analytical model for broader applications in the future. This research contributes to non-profit organizations by offering a more efficient approach to managing donor relationships.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

## I. PENDAHULUAN

Dalam dunia organisasi nirlaba dan amal, mengelola hubungan dengan donatur merupakan aspek penting untuk memastikan kesinambungan dan efisiensi operasional. Tanpa dukungan donatur, banyak program sosial yang dirancang untuk membantu Masyarakat kurang mampu tidak akan terwujud. Donatur menyediakan sumber daya yang diperlukan untuk melaksanakan berbagai kegiatan dan inisiatif yang mempunyai dampak sosial yang signifikan.

Donatur memiliki kualitas, kebiasaan, dan tuntutan yang bervariasi, sehingga perlu untuk mengenali setiap pelanggan satu per satu bagi organisasi agar tidak menjadi masalah, sehingga diperlukan strategi untuk memudahkan dalam mengidentifikasinya [1]. Dalam upaya meningkatkan hubungan dengan donatur, segmentasi memungkinkan organisasi untuk memahami perilaku donatur. Dengan melakukan segmentasi donatur, organisasi dapat

meningkatkan pendapatan donasi dan memastikan kepuasan dan loyalitas donatur krusial dalam hubungan jangka Panjang dengan organisasi [2]. Salah satu pendekatan yang sering digunakan dalam segmentasi donatur adalah kombinasi model RFM (*Recency, Frequency, Monetary*) dan algoritma *K-Means*.

Pada data transaksi customer umumnya memiliki banyak atribut dan ini bisa menyebabkan pengelolaan data nanti kurang baik [3]. Oleh karena itu untuk mengatasi masalah pemilihan atribut terbaik diusulkan menggunakan metode RFM. Teknik *Recency, Frequency, and Monetary* (RFM) mengkaji karakteristik pelanggan berdasarkan data transaksi untuk menghasilkan tiga faktor utama yaitu *recency*, *frequency*, dan *monetary*, dalam rangka mempelajari data historis peningkatan transaksi pelanggan [4]. Model RFM diterapkan dalam segmentasi data, agar hasil yang didapatkan lebih tepat [5],[6]. Model ini mampu memantau perubahan perilaku customer dari waktu ke waktu dan

memungkinkan dan memungkinkan adaptasi strategi organisasi sesuai dengan perubahan dalam pola transaksi customer [7]. Penelitian sebelumnya dalam melakukan segmentasi *customer* dengan model RFM banyak menggunakan metode *clustering* untuk mendukung proses segmentasi. Di mana *clustering* merupakan pengelompokan objek yang diperoleh dari suatu data yang menjelaskan hubungan antar objek satu dengan objek lainnya [8]. Metode *clustering* yang umum digunakan pada penelitian sebelumnya adalah *clustering K-Means*.

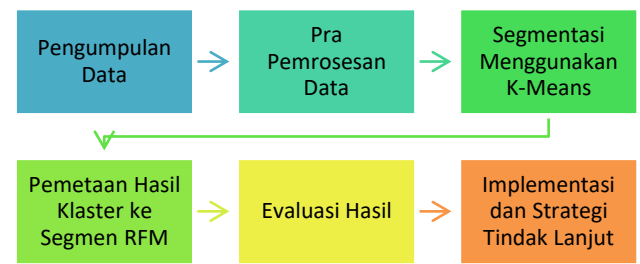
Metode *K-Means* adalah metode yang dikategorikan ke dalam metode klatering partisi, karena algoritma ini memiliki kesederhanaan dan kemudahan dalam menggunakannya dan pengguna bisa menentukan sendiri jumlah klasternya [9]. Sebelum menggunakan metode *K-Means* dalam penentuan nilai  $k$  menggunakan metode *Elbow*. Metode *Elbow* merupakan suatu metode yang dapat digunakan untuk menghasilkan informasi dalam menentukan *cluster* terbaik dengan cara melihat persentase hasil perbandingan antara jumlah *cluster* yang akan membentuk siku pada suatu titik tertentu [10]. Dengan menggunakan algoritma *K-means*, donatur dapat dikelompokkan ke dalam beberapa segmen yang memiliki karakteristik serupa, seperti “donatur setia”, “donatur potensial”, “donatur dengan kontribusi rendah” dan “donatur baru”. segmentasi ini memungkinkan organisasi untuk Menyusun strategi yang lebih personalisasi, seperti kampanye khusus untuk donatur yang kurang kontribusi atau apresiasi khusus bagi donatur setia atau potensial.

Oleh karena itu, penelitian ini bertujuan untuk mengeksplorasi penerapan model RFM dan algoritma *K-means* dalam segmentasi donatur, serta mengevaluasi efektivitasnya dalam meningkatkan manajemen hubungan donatur dan strategi pemasaran program donasi. Melalui analisis ini, kami berharap lembaga nirlaba atau lembaga non profit dapat menemukan wawasan berharga untuk mengoptimalkan interaksi mereka dengan donatur dan memaksimalkan dampak program yang mereka kelola.

## II. METODE PENELITIAN

Metode penelitian merupakan pendekatan sistematis yang digunakan untuk merencanakan, melaksanakan, dan menganalisis penelitian untuk mencapai tujuan yang ditetapkan [11]. Metode penelitian yang digunakan dalam jurnal terkait segmentasi donatur dengan model RFM (*Recency, Frequency, Monetary*) dan algoritma *K-Means* umumnya melibatkan beberapa tahapan utama yang diintegrasikan untuk mencapai hasil yang optimal dalam mengelompokkan donatur atau pelanggan.

Penelitian ini dilakukan dengan menerapkan metode penelitian kuantitatif, dimana metode ini melibatkan pengumpulan dan analisis data berupa angka atau variabel yang dapat diukur secara objektif. Tahapan penelitian yang digunakan pada penelitian ini seperti ditunjukkan pada gambar 1.



Gambar 1. Tahap Penelitian

### A. Pengumpulan data

Dataset ini didasarkan pada data transaksi dari donatur Yayasan Galang Bersama Kami yang bertempat di Palu Sulawesi Tengah. Pada tahap ini, data dikumpulkan berupa data donatur dari basis data Yayasan Galang Bersama Kami. Data ini mencakup tanggal donasi terakhir, frekuensi donatur, dan jumlah total donasi yang diberikan dalam priode tertentu.

### B. Pra Pemrosesan Data

Dataset Yayasan Galang Bersama Kami merupakan data mentah, sehingga dibutuhkan pra pemrosesan data. Pra pemrosesan data dilakukan dengan tujuan untuk mengubah data mentah ke suatu format yang prosesnya lebih mudah dan efektif untuk kebutuhan sehingga data siap digunakan untuk diproses lebih lanjut [12].

Berikut langkah-langkah dalam pra pemrosesan data antara lain:

#### 1) Pembersihan Data

Pembersihan data untuk menghilangkan duplikasi data dan mengoreksi kesalahan dalam pencatatan.

#### 2) Reduksi Data

Dengan mengurangi atribut dalam data set yang tidak diperlukan dan tidak berpengaruh pada proses Analisa guna untuk kompleksitas dan mempercepat analisis. Berikut atribut yang digunakan penelitian ini terlihat pada table 1.

TABEL 1.  
DATA REDUKSI

Data yang Tidak Digunakan	Data yang Digunakan
Nama_donatur	Donator_id
Nama_program	Tanggal
Donasi_id	Jumlah
Program_id	
Donasi_rekap_id	
Posisi_uang_id	
Channel_id	
Divisi_id	

#### 3) Pendekatan RFM

Model RFM ini diperkenalkan pertama kali oleh Arthur Huges dan telah banyak digunakan oleh industri manufaktur, retailer dan industry jasa [13]. Data set yang telah dipilih

kemudia pemodelan data menjadi RFM untuk mengukur nilai donatur, Dimana setiap donatur dievaluasi berdasarkan tiga dimensi utama: *Recency*, *Frequency*, dan *Monetary* [7]. RFM mempertahankan informasi tentang waktu pembelian (*recency*) paling akhir, berapa kali pelanggan melakukan pembelian (*frequency*), dan rata-rata uang yang dihabiskan [14]. Keuntungan model RFM terletak pada relevansinya selama beroperasi pada beberapa variabel yang dapat diamati dan bersifat objektif [15].

Setelah model data RFM didapatkan, selanjutnya membuat boxplot untuk mendeteksi data outlier atau data yang jauh dari nilai-nilai lainnya. Langkah terakhir dari prapemrosesan data yaitu membersihkan data outlier yang terdeteksi dan melakukan penormalisasian data.

### C. Segmentasi Menggunakan K-Means

Sebelum melakukan segmentasi menggunakan *K-Means* terlebih dahulu menentukan jumlah cluster menggunakan Teknik *Elbow*. Metode *Elbow* merupakan ide atau gagasan yang digunakan untuk menghasilkan informasi dalam memilih nilai cluster terbaik dengan cara melihat presentasi hasil perbandingan antara jumlah *cluster* yang membentuk siku pada suatu titik [16], [17]. Kurva *Elbow* merupakan metode yang digunakan untuk menentukan jumlah cluster optimal dalam algoritma *K-Means* berdasarkan nilai *WCSS* (*Within-Cluster Sum of Squares*). Tujuannya adalah menemukan titik di mana penambahan cluster tidak lagi memberikan perbaikan signifikan dalam hal kompaksi cluster, sehingga penambahan cluster setelah titik tersebut tidak meningkatkan kualitas pemisahan data.

*K-Means* merupakan metode *clustering* secara *partitioning* yang memisahkan data ke dalam kelompok yang berbeda [10]. Metode *K-Means* digunakan untuk mengelompokkan  $n$  vector berdasarkan atribut partisi  $k$ , di mana  $k < n$ , tergantung pada beberapa Tindakan. Pusat kluster adalah *mean* (nilai Tengah) semua vektor pada *cluster* tertentu. Algoritma ini dimulai dengan memilih *centroid*  $k$  awal secara acak, kemudian memberikan nilai vector ke *centroid* terdekat dengan *Euclidean distance* dan dihitung ulang *centroid* baru. proses ini berulang sampai *vector* tidak lagi mengubah *cluster* antar iterasi [15].

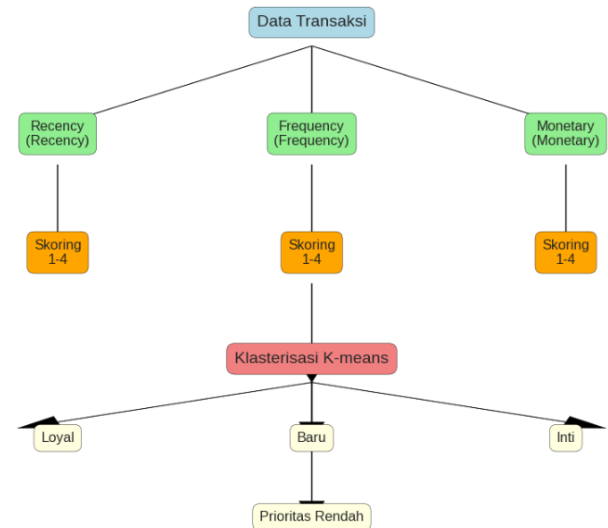
Algoritma *K-Means* kemudian diterapkan pada skor RFM yang telah dihitung untuk mengelompokkan donatur ke dalam sejumlah  $K$  kluster. Algoritma ini bekerja dengan meminimalkan jarak antara data dalam kluster yang sama dan memaksimalkan varian antar kluster.

### D. Pemetaan Hasil Kluster ke Segmen RFM

Pemetaan hasil *cluster* ke segmen RFM adalah langkah krusial dalam proses segmentasi donatur atau pelanggan, yang menghubungkan hasil dari algoritma *K-Means* dengan interpretasi bisnis yang bermakna.

Setelah algoritma *K-Means* dijalankan, setiap donatur akan ditempatkan ke dalam salah satu *cluster* berdasarkan skor RFM mereka. Hasil ini biasanya berupa sekumpulan

*cluster* yang mengelompokkan donatur dengan pola skor RFM yang serupa. Setiap dimensi diberi skor, biasanya dari 1 hingga 4, yang menggambarkan performa masing-masing donatur dalam aspek tersebut. Kombinasi dari skor RFM menghasilkan profil donatur yang dapat digunakan untuk segmentasi. Berikut gambar 2 sebagai ilustrasi untuk pembahasan RFM dan segmentasi *K-Means*.



Gambar 2. Alur Pemetaan Segmentasi

### E. Evaluasi Hasil

Tahap evaluasi merupakan tahap lanjutan Dimana tujuan data mining dievaluasi secara menyeluruh untuk mendapatkan pemodelan yang diinginkan [18]. Pada tahap ini dilakukan evaluasi untuk memastikan bahwa segmentasi tersebut valid dan berguna dalam konteks bisnis atau tujuan organisasi. Proses evaluasi ini dilakukan dengan cara pengujian apakah pola, informasi atau knowledge yang diperoleh sesuai dengan fakta atau hipotesa sebelumnya atau justru malah bertentangan [19]. Ini bisa melibatkan validasi dengan data eksternal atau melihat kinerja histori dari setiap segmen dalam hal penggalangan dana. Setelah segmen dibentuk, informasi yang diperoleh kemudian dipresentasikan ke dalam bentuk visualisasi yang memberikan interpretasi yang relevan dan membuat rekomendasi untuk Tindakan yang bisa diambil organisasi.

### F. Implementasi dan Strategi Tindak Lanjut

Pada tahap ini, informasi dan pengetahuan yang diperoleh disajikan dalam bentuk yang lebih mudah dipahami oleh Masyarakat umum [18]. Pada Langkah ini juga digunakan segmen yang telah terbentuk untuk mengarahkan strategi pemasaran yang lebih tepat sasaran, seperti kampanye donasi yang disesuaikan dengan karakteristik segmen. Segmen ini dipantau secara berkala, dan strategi disesuaikan berdasarkan kinerja segmen dari waktu ke waktu. Hasil penelitian ini berupa kelompok

donaturdan rekomendasi startegi pemasaran yang dapat diterapkan perusahaan berdasarkan data yang dihasilkan untuk meningkatkan jumlah donasi.

### III. HASIL DAN PEMBAHASAN

#### A. Pengumpulan Data

Dataset yang dikumpulkan pada penelitian ini bersumber dari *database* transaksi donatur Yayasan Galang Bersama Kami selama tujuh bulan dari 1 januari 2023 sampai 26 juli 2023, sejumlah 11 atribut dan berjumlah 2465 *record* atau data donatur yang masing-masing subdataset berformat excel. Data donatur di Yayasan Galang Bersama Kami memiliki atribut : nama\_donatur, nama\_program, doansi\_id, tanggal, donatur\_id, program\_id, donasi\_rekap\_id, jumlah, posisi\_uang\_id, channel\_id, dan divisi\_id.

	nama_donatur	nama_program	donasi_id	tanggal	program_id	donatur_id	donasi_rekap_id	jumlah	posisi_uang_id	channel_id	divisi_id
0	Hj Eri Purwati (kalem 011)	Safari Tahfidz Reguler 2022 (Non Program)	7137	2023-04-01 00:00:00	76	5933	104	1000000	11	41	17
1	Hj Saleah (kalem 070)	Safari Tahfidz Ramadhan (Non program)	7138	2023-05-01 00:00:00	69	5932	191	1000000	8	31	17
2	Salmawati (sobel 014)	Safari Tahfidz Reguler 2022 (Non Program)	7139	2023-08-01 00:00:00	76	5202	104	1000000	11	33	17
3	Saprudin (kalem 011)	Safari Tahfidz Reguler 2022 (Non Program)	7140	2023-08-01 00:00:00	76	5931	104	500011	8	41	17
4	Hamba Allah	Sedekah Bebas 2023	7141	2023-06-01 00:00:00	102	19	240	50000	4	17	8
2460	latah sudawati 303	Safari Tahfidz Reguler 2023	9780	26/07/2023	104	7067	109	250000	11	62	17
2461	agus widoro 303	Safari Tahfidz Reguler 2023	9781	26/07/2023	104	7535	109	100303	8	62	17
2462	andi ahfandi 301	Kisah Reguler 2023	9782	13/07/2023	105	6776	107	1000000	12	33	17
2463	Yose surya putra	Sedekah Bebas 2023	9783	26/07/2023	102	6572	115	100	6	17	8
2464	TADJUDIN NUJR	Hafidz Quran	9784	26/07/2023	17	6599	112	100000	3	17	8

Gambar 3. Data Transaksi Donatur

#### B. Pra Pemrosesan Data

Tahap ini dimulai dari membersihkan data duplikasi, menangani data kosong pada tabel, dan pendekatan data model RFM. Langkah pertama adalah mengumpulkan data yang diperlukan untuk menghitung RFM. Setiap donatur perlu memiliki informasi tentang:

- 1) Tanggal donasi (untuk menghitung *Recency*). Untuk menghitung *Recency*, perlu menghitung selisih waktu (biasanya dalam hari atau bulan) antara tanggal acuan dan tanggal donasi terakhir. Dimana tanggal acuan pada penelitian ini adalah 2023-07-26.

Rumus *Recency*:

$$Recency = \text{Tanggal acuan} - \text{Tanggal donasi terakhir} \text{ ('tanggal': lambda x: (anchorDate-x.max()).days)}$$

Semakin kecil angka *Recency*, semakin baik, karena donatur tersebut baru saja melakukan donasi.

- 2) Donatur\_id untuk menghitung jumlah donasi yang dilakukan (untuk menghitung *Frequency*). Seberapa sering donatur melakukan donasi dalam priode tertentu. Pada penelitian ini memiliki priode donasi dari 2024-01-01 sampai 2024-07-26.

Rumus *Frequency*:

$$Frequency = \text{'donatur\_id': lambda x: len(x)} \text{ (jumlah donasi yang dilakukan dalam periode tertentu).}$$

Donatur yang sering berdonasi memiliki nilai *frequency* yang lebih tinggi.

- 3) Jumlah uang yang didonasikan (untuk menghitung *Monetary*).

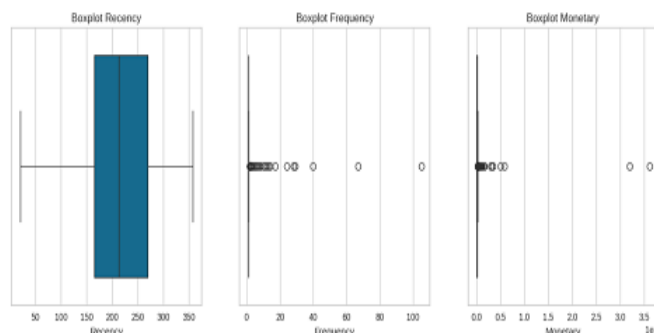
Rumus *Monetary*:

$$Monetary = \text{'jumlah': lambda x: x.sum()} \text{ (total nilai donasi dalam periode tertentu).}$$

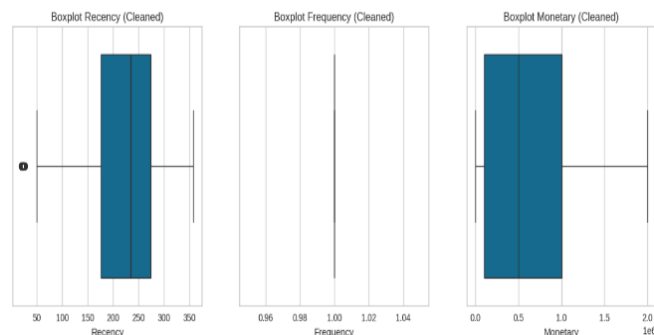
donatur_id	Recency	Frequency	Monetary
7	85	8	400000
11	84	3	400000
19	53	67	5439098
20	21	105	10965403
21	84	40	2391497

Gambar 4. Skor RFM

Setelah model data RFM didapatkan, selanjutnya membuat boxplot untuk memberikan gambaran tentang distribusi data dan mendeteksi outlier.



Gambar 5. Boxplot untuk Mendeteksi Outlier



Gambar 6. Boxplot Setelah Outlier Dibersihkan

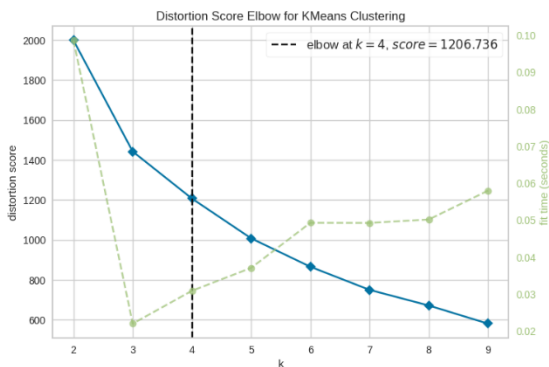
Dengan melakukan praproses data yang lebih detail, termasuk dengan mendeteksi dan membersihkan outlier menggunakan boxplot, dapat dipastikan bahwa data yang digunakan untuk analisis data lebih akurat dan bebas dari nilai ekstrem yang dapat mengganggu hasil sehemntasi.



Langkah-langkah ini meningkatkan kualitas data dan pada akhirnya, keakuratan model segmentasi yang digunakan. Dapat dilihat pada gambar 5 dan 6, merupakan hasil perbandingan data sebelum dan sesudah data outlier dibersihkan. Pada gambar 5 merupakan gambaran tentang distribusi data menggunakan boxplot guna untuk mendeteksi outlier. Sedangkan pada gambar 6 merupakan distribusi data menggunakan boxplot setelah data outlier berhasil dibersihkan.

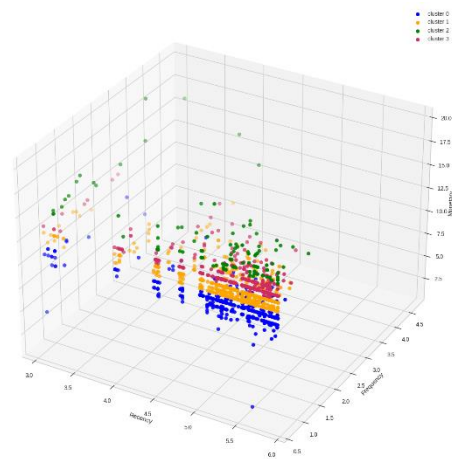
C. Segmentasi Menggunakan K-Means

Sebelum melakukan segmentasi menggunakan *K-Means*, terlebih dahulu menentukan nilai k menggunakan metode *Elbow*. Dapat dilihat pada gambar 7, terdapat titik siku yang terbentuk pada grafik di titik 5. Setelah titik lima, tidak terdapat lagi penurunan yang signifikan dalam nilai jumlah *cluster*. Oleh karena itu, dapat disimpulkan bahwa jumlah *cluster* optimal adalah lima. Namun, setelah melakukan uji performa, ditemukan bahwa jumlah *cluster* yang optimal sebenarnya adalah empat, sehingga dipilih titik ke empat untuk melakukan *clustering* Kembali.



Gambar 7. Distorsi Skor Elbow

Dari gambar 7 di atas memperlihatkan grafik *Elbow*. Grafik ini akan memplot jumlah cluster dari 2 hingga 10 dan nilai *WCSS* yang berkurang seiring bertambahnya jumlah cluster. Titik siku pada grafik ini menunjukkan jumlah cluster optimal. Ini adalah titik di mana penambahan cluster baru tidak menghasilkan penurunan yang signifikan dalam *WCSS*. Grafik yang dihasilkan akan menunjukkan bahwa *WCSS* turun tajam untuk 2, 3, atau 4 cluster, tetapi penurunan tersebut mulai melambat setelah di titik 4. Dan *WCSS* tidak banyak berubah setelah 4 cluster, maka 4 cluster adalah pilihan yang tepat untuk hasil *clustering* terbaik.



Gambar 8. Kluster Donatur

Pada gambar 8 di atas menunjukkan 4 kluster dari *K-Means*. Hasilnya tampak menjanjikan: kluster 2 dengan warna hijau adalah kelas dengan *recency* kecil, *frequency* besar, dana *monetary* tinggi. Kluster 3 dengan warna merah adalah kluster dengan *recency* besar, *frequency* kecil, dan *monetary* tinggi. Kluster 1 dengan warna kuning yang hampir sama dengan kluster 3 di mana *recency* besar, *frequency* besar, dan *monetary* rendah. Sedangkan kluster 0 dengan warna biru adalah kluster dengan *recency* rendah, *frequency* rendah, dan *monetary* kecil.

D. Pemetaan Hasil Kluster ke Segmen RFM

Untuk memetakan data RFM ke dalam segmen dapat dilakukan dengan membagi data ke dalam kuartil atau persentil. Gambar 8 menunjukkan klasifikasi setiap metrik (*recency*, *frequency*, *monetary*) ke dalam kategori seperti rendah, sedang, dan tinggi.

1. kuartil untuk *Recency*, membagi data *recency* ke dalam empat bagian berdasarkan nilai persentil 0%, 25%, 75%, 100%.
2. Kuartil untuk *Frequency*, mengingat bahwa nilai *frequency* sering kali memiliki distribusi yang sangat mirip atau nilai hamper sama, kadang-kadang digunakan persentil tinggi seperti, 97%, 98%, 99% untuk lebih memisahkan donator yang sangat aktif. Dapat dilihat hasilnya pada gambar 9.
3. Kuartil untuk *Monetary*, membagi data *monetary* ke dalam empat bagian berdasarkan nilai persentil 0%, 25%, 75%, 100%.

	Q_Recency	Q_Frequency	Q_Monetary
min	20.00	1.0	3.0
first_part	165.25	4.0	191875.0
second_part	212.50	5.0	500000.0
third_part	269.00	7.0	1000009.0
max	358.00	105.0	361892732.0

Gambar 9. Nilai Kuartil RFM

Setelah kuartil dihitung, selanjutnya memetakan setiap donator ke dalam skor RFM berdasarkan nilai mereka terhadap kuartil atau persentil yang sudah dihitung. Seperti pada gambar 10 di bawah, pemetaan skor untuk setiap dimensi:

- 1) Skor *recency* merupakan donator yang berada di kuartil pertama (0-25%) diberi skor 4 atau paling baru, sedangkan yang berada di kuartil terakhir (75%-100%) diberi skor 1 atau paling lama.
- 2) Skor *frequency* merupakan donator dengan nilai *frequency* tinggi di persentil 99-100% diberi skor 4, sedangkan yang lebih rendah di persentil bawah diberi skor 1.
- 3) Skor *monetary* sama seperti *recency*, nilai *monetary* tinggi di kuartil atas diberi skor 4, dan yang di kuartil bawah diberi skor 1.

donatur_id	Recency	Frequency	Monetary	log_Recency	log_Frequency	log_Monetary	r_score	f_score	m_score	rfm_score
5784	85	11	1500000	4.454347	2.484907	14.220976	4	4	4	444
21	84	40	2391497	4.442651	3.713572	14.687431	4	4	4	444
19	53	67	5439098	3.988984	4.219508	15.509124	4	4	4	444
52	22	29	361892732	3.135494	3.401197	19.706858	4	4	4	444
1753	22	13	2700000	3.135494	2.639057	14.808763	4	4	4	444
4201	22	28	4300000	3.135494	3.367296	15.274126	4	4	4	444
4938	22	12	1150000	3.135494	2.564949	13.955273	4	4	4	444
5948	22	10	1050000	3.135494	2.397895	13.864302	4	4	4	444
20	21	105	10965403	3.091042	4.663439	16.210256	4	4	4	444

Gambar 10. Skor RFM

Langkah terakhir adalah memetakan kombinasi skor RFM ke dalam segmen yang lebih mudah dipahami. Segmentasi ini sering kali berdasarkan pola umum dalam perilaku pelanggan atau donatur. Seperti pada gambar 11 menjabarkan segmen berdasarkan skor RFM:

- 1) “Core” dengan skor 444: donatur dengan nilai tertinggi disemua dimensi RFM.
- 2) “Loyal Large” dengan skor [3-4][3-4][3-4]: donatur yang sering dan besar kontribusinya, tetapi tidak melakukan donasi baru-baru ini.
- 3) “New Large” dengan skor [3-4][1-2][3-4]: donatur yang baru memberikan kontribusi besar.
- 4) “Lost Loyal Large” dengan skor [1-2][3-4][3-4]: donatur yang sebelumnya sering berdonasi tetapi sekarang jarang.
- 5) “Low Priority” dengan skor [1-2][1-2][1-2]: donatur yang memiliki skor rendah di semua dimensi.

donatur_id	Recency	Frequency	Monetary	log_Recency	log_Frequency	log_Monetary	r_score	f_score	m_score	rfm_score	RFM_Segment
7	85	8	400000	4.454347	2.197225	12.899222	4	4	2	442	02_loyal_small
11	84	3	400000	4.442651	1.386294	12.899222	4	1	2	412	03_new_small
19	53	67	5439098	3.988984	4.219508	15.509124	4	4	4	444	01_core
20	21	105	10965403	3.091042	4.663439	16.210256	4	4	4	444	01_core
21	84	40	2391497	4.442651	3.713572	14.687431	4	4	4	444	01_core
...	...	...	...	...	...	...	...	...	...	...	...
7531	155	1	41200	5.049856	0.693147	10.626216	4	1	1	411	03_new_small
7532	155	1	76500	5.049856	0.693147	11.245059	4	1	1	411	03_new_small
7533	155	1	100000	5.049856	0.693147	11.512935	4	1	1	411	03_new_small
7534	155	1	100000	5.049856	0.693147	11.512935	4	1	1	411	03_new_small
7535	154	1	100303	5.043425	0.693147	11.515961	4	1	1	411	03_new_small

1662 rows x 11 columns

Gambar 11. Pemetaan Skor RFM ke Segmen

E. Evaluasi Hasil

Evaluasi hasil ini melibatkan relevansi segmen-segmen yang lebih deskriptif dan mudah dipahami. Fungsi utama dari evaluasi hasil ini adalah untuk membuat hasil klustering menjadi lebih berarti bagi analisis strategi pemasaran program donasi. Dengan menggunakan segmen yang lebih jelas dapat mempermudah merancang kampanye yang spesifik untuk setiap kelompok donatur berdasarkan perilaku mereka yang diindikasikan oleh segmen tersebut. Terlihat pada gambar 12 merupakan hasil evaluasi dimana menghasilkan segmen yang mudah dipahami seperti donatur “Loyal”, “Baru”, “Inti”, dan donatur “Prioritas Rendah”.

donatur_id	Recency	Frequency	Monetary	log_Recency	log_Frequency	log_Monetary	r_score	f_score	m_score	rfm_score	RFM_Segment	WNew_Segment	WNew_seg_trans
7	85	8	400000	4.454347	2.197225	12.899222	4	4	2	442	02_loyal_small	1	Loyal
11	84	3	400000	4.442651	1.386294	12.899222	4	1	2	412	03_new_small	1	Loyal
19	53	67	5439098	3.988984	4.219508	15.509124	4	4	4	444	01_core	2	Inti
20	21	105	10965403	3.091042	4.663439	16.210256	4	4	4	444	01_core	2	Inti
21	84	40	2391497	4.442651	3.713572	14.687431	4	4	4	444	01_core	2	Inti
...	...	...	...	...	...	...	...	...	...	...	...	...	...
7531	155	1	41200	5.049856	0.693147	10.626216	4	1	1	411	03_new_small	0	Prioritas rendah
7532	155	1	76500	5.049856	0.693147	11.245059	4	1	1	411	03_new_small	0	Prioritas rendah
7533	155	1	100000	5.049856	0.693147	11.512935	4	1	1	411	03_new_small	0	Prioritas rendah
7534	155	1	100000	5.049856	0.693147	11.512935	4	1	1	411	03_new_small	0	Prioritas rendah
7535	154	1	100303	5.043425	0.693147	11.515961	4	1	1	411	03_new_small	0	Prioritas rendah

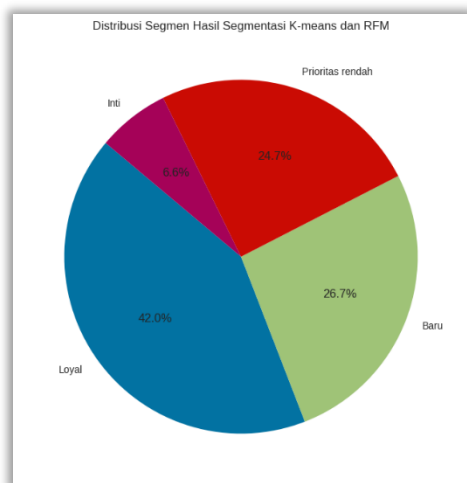
1662 rows x 13 columns

Gambar 12. Evaluasi Hasil

Pada gambar 13 dan 14 merupakan hasil dari analisis dan visualisasi setiap segmentasi. Menghitung persentasi setiap segmen, menghitung proporsi pelanggan dalam setiap segmen berdasarkan hasil klusterisasi. Setelah menghitung presentasi setiap segmen, selanjutnya mengvisualisasi hasil segmen menggunakan diagram pai atau batang untuk distribusi pelanggan dalam setiap segmen.

KMeans_seg_trans	RFM_segment	count
Baru	01_core	3
	02_loyal_large	4
	03_new_large	204
	04_lost_loyal_large	1
	05_promising	231
Inti	01_core	6
	02_loyal_large	2
	03_new_large	58
	05_promising	44
Loyal	02_loyal_large	1
	02_loyal_small	6
	03_new_large	116
	03_new_small	208
	05_promising	132
	06_low_priority	235
Prioritas rendah	02_loyal_small	3
	03_new_small	220
	06_low_priority	188

Gambar 13. Jumlah donator di setiap segmen



Gambar 14. Persentasi jumlah donator di setiap Segmen

#### F. Implementasi dan Strategi Tindak Lanjut

Untuk menerapkan hasil segmentasi donatur yang diperoleh melalui model RFM (*Recency, frequency, Monetary*) dan *clustering K-Means*, yayasan dapat mengikuti beberapa langkah praktis yang sudah didiskusikan bersama staf CRM yayasan selaku penyambung hubungan donatur dengan yayasan. Berikut langkah-langkah yang dapat dilakukan:

1) *Identifikasi Segmen Donatur Berdasarkan Hasil Clustering*. Setelah model RFM diterapkan dan data donatur dikelompokkan menggunakan *K-Means*, setiap donatur akan masuk ke dalam segmen tertentu. Selanjutnya, Analisis mendalam pada karakteristik setiap cluster untuk memahami profil donatur di setiap segmen. Tipe segmen yang terbentuk pada penelitian ini adalah:

- Donatur loyal yaitu donatur yang sering berdonasi dan dalam jumlah besar.
- Donatur inti yaitu donatur yang kadang-kadang memberikan donasi, tetapi jumlahnya tidak signifikan.

- Donatur baru yaitu donatur yang hanya sekali atau dua kali berdonasi.
- Donatur prioritas rendah yaitu donatur yang dulunya sering berdonasi tetapi sudah lama tidak berkontribusi.

2) *Membangun Strategi Komunikasi Khusus Berdasarkan Segmen*. Yayasan dapat menargetkan segmen donatur yang berbeda dengan pesan dan strategi komunikasi yang lebih personal, sesuai dengan perilaku dan nilai donasi mereka tunjukkan. Berikut strategi komunikasi berdasarkan segmen:

- Donatur loyal. Mengirimkan ucapan terima kasih personal atau laporan khusus bagaimana hasil donasi mereka disalurkan secara rinci, mengajak para donatur terlibat langsung ke kegiatan yayasan serta gunakan komunikasi one-on-one seperti panggilan telepon atau pertemuan pribadi untuk menjaga hubungan baik.
- Donatur prioritas rendah. Kirimkan kampanye pemulihan yang menekankan pentingnya kontribusi mereka di masa lalu dan undang untuk Kembali terlibat, berikan informasi tentang program baru atau inisiatif yang mungkin menarik minat mereka serta tawarkan kemudahan berdonasi di yayasan seperti menyediakan nomor rekening seluruh bank.
- Donatur inti. Gunakan pesan pengingat secara berkala yang memotivasi mereka untuk berdonasi lebih sering, buat kampanye yang menyoroti peristiwa atau kebutuhan khusus sehingga mereka terdorong untuk berpartisipasi serta tawarkan opsi donasi kecil tetapi rutin.
- Donatur baru. Kirimkan pesan yang memperkenalkan yayasan lebih dalam termasuk tujuan dan nilai-nilai yang diperjuangkan. Tunjukkan dampak donasi mereka secara langsung melalui laporan aktivitas. Serta ajak mereka untuk bergabung dalam program donasi berulang atau memberikan kontribusi yang lebih besar di masa depan.

3) *Penerapan Pengelolaan Donatur yang Lebih Efisien*. Yayasan dapat menggunakan hasil segmentasi untuk mengalokasikan sumber daya dan tenaga kerja lebih efisien. Gunakan segmentasi untuk menentukan prioritas dalam pengelolaan hubungan dengan donatur, seperti frekuensi komunikasi dan jenis layanan yang diberikan. Dengan mengetahui pola kontribusi dari setiap segmen, yayasan dapat mengantisipasi penurunan donasi dari segmen tertentu dan melakukan pemeliharaan hubungan lebih baik.

4) *Memantau dan Mengevaluasi Hasil Segmentasi*. Setelah menerapkan strategi komunikasi dan kampanye berdasarkan hasil segmentasi, yayasan perlu melakukan evaluasi secara berkala untuk mengukur efektivitasnya. Dengan memantau hasil, yayasan dapat terus menyempurnakan strategi dan memastikan setiap segmen donatur tetap terlibat secara optimal.

#### IV. DISKUSI DAN PEMBAHASAN

Penelitian ini berhasil menunjukkan bahwa kombinasi model RFM (*Recency Frequency Monetary*) dengan algoritma *K-Means* dapat memberikan segmentasi donatur yang lebih terarah dan bermanfaat bagi yayasan atau organisasi nirlaba. Hasil segmentasi memberikan hasil adanya 4 segmen donatur, yaitu donatur loyal (sering berdonasi dan memberikan kontribusi besar), donatur baru (baru-baru ini mulai berdonasi), donatur prioritas rendah (donatur yang jarang berdonasi), dan donatur inti (donatur dengan kontribusi tinggi tetapi frekuensi rendah).

Hasil segmentasi dari model RFM ini adalah pengimplementasiannya yang fleksibel dan mudah dalam adaptasi terhadap jenis dataset. Namun, salah satu kekurangan yang ditemukan dari model RFM ini adalah ketergantungan dengan algoritma *K-Means* yang dapat memberikan hasil yang berbeda tergantung pada pemilihan centroid, dan pada pemilihan  $k$  (jumlah kluster) tidak selalu menghasilkan segmentasi yang optimal.

Metode *Elbow* yang digunakan dalam penelitian ini memberikan indikasi baik untuk menentukan jumlah kluster, tetapi masih ada resiko overfitting dan underfitting jika data yang digunakan tidak cukup representatif. Segmentasi yang dihasilkan dari model ini dapat langsung diterapkan dalam strategi membangun hubungan dengan donatur dan penggalangan yayasan atau organisasi nirlaba. Misalnya, donatur prioritas rendah yang jarang berdonasi mungkin memerlukan pendekatan yang lebih personal atau insentif tambahan untuk meningkatkan partisipasi mereka dalam program donasi. Sementara itu, segmen donatur baru yang baru memulai berdonasi agar dimaintenance terus dan dapat dijadikan target untuk program donasi khusus atau donasi terjadwal dan donasi harian.

Untuk mengoptimalkan penggunaan model ini, peneliti lanjutan disarankan untuk menguji segmentasi ini dalam berbagai kondisi atau dengan dari yayasan atau organisasi nirlaba yang berbeda. Selain itu, eksplorasi penggunaan algoritma klustering lain, seperti *K-Medoids* dan DBSCAN, dapat dilakukan untuk membandingkan performa dan keakuratan segmentasi. Implementasi sistem otomatisasi untuk pemantauan dan evaluasi berkelanjutan dari segmen-segmen donatur juga disarankan untuk memastikan strategi yang diterapkan tetap relevan dan efektif.

#### V. KESIMPULAN

Dalam analisis segmentasi donatur yang menerapkan metode klustering *K-Means* dan segmentasi RFM pada dataset donasi untuk mengidentifikasi segmen donatur yang berbeda. Hasilnya menunjukkan adanya tumpang tindih yang signifikan oleh kedua metode tersebut. Tumpang tindih ini menunjukkan bahwa segmen yang dihasilkan dari klustering *K-Means* sejalan dengan yang didefinisikan oleh analisis RFM, yang menunjukkan konsistensi dan efektif dalam proses segmentasi. Pemetaan ke data RFM adalah proses sistematis yang melibatkan beberapa Langkah, mulai dari penentuan nilai atau skor RFM hingga perhitungan

kuartil dan pemetaan ke segmen yang lebih mudah dipahami. Hasilnya adalah segmentasi yang lebih tepat yang dapat digunakan untuk mengoptimalkan strategi pemasaran, penggalangan dana dan merancang pendekatan yang lebih efisien dalam pengelolaan hubungan dengan donatur.

#### DAFTAR PUSTAKA

- [1] R. Siagian, P. S. Pahala Sirait, and A. Halima, "E-Commerce Customer Segmentation Using K-Means Algorithm and Length, Recency, Frequency, Monetary Model," *J. Informatics Telecommun. Eng.*, vol. 5, no. 1, pp. 21–30, 2021, doi: 10.31289/jite.v5i1.5182.
- [2] M. Martono and M. Syafrullah, "Pengklasteran Dan Segmentasi Karakteristik Donatur Sedekah Daring Dengan Teknik Penambangan Data," *INOVTEK Polbeng - Seri Inform.*, vol. 9, no. 1, pp. 366–380, 2024, doi: 10.35314/isi.v9i1.4223.
- [3] S. E. Jamal, "Analisis RFM dan Algoritma K-Means untuk Clustering Loyalitas Customer RFM Analysis and K-Means Algorithm for Customer Loyalty Clustering," *Kom J. ENERGY (Jurnal Ilm. Ilmu-ilmu Tek.)*, vol. 9, no. 1, p. 18, 2019, [Online]. Available: <https://doi.org/10.51747/energy>
- [4] P. I. Pangestu, T. I. Hermanto, and D. Irmayanti, "Analisis Segmentasi Pelanggan Berbasis Model Recency Frequency Dan Monetary (Rfm) Menggunakan Algoritma K-Means," *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 3s1, pp. 1486–1492, 2023, doi: 10.23960/jitet.v11i3s1.3396.
- [5] Basri, W. Gata, and Risnandar, "Customer Loyalty Analysis Based on the Recency, Frequency and Monetary (RFM) and Decision Tree Models At Pt. Solo," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 5, pp. 943–950, 2020, doi: 10.25126/jtiik.202072284.
- [6] A. Madani, A. Rahmah, F. Nurunnisa, and A. Elia, "SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat Customer Segmentation at BC HNI 2 Pekanbaru by Applying the K-Medoids Algorithm and Recency, Frequency, Monetary (RFM) Model Segmentasi Pelanggan pada BC HNI 2 Pekanbaru dengan Menerapkan Algoritma K-Medoids dan Model Recency, Frequency, Monetary (RFM)." [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas>
- [7] F. M. Pranata, S. H. Wijoyo, and N. Y. Setiawan, "Analisis Performa Algoritma K-Means dan DBSCAN Dalam Segmentasi Pelanggan Dengan Pendekatan Model RFM," vol. 1, no. 1, pp. 1–9, 2017.
- [8] A. Pramudiansyah and H. Munte, "SEGMENTASI PELANGGAN MENGGUNAKAN ALGORITMA K-MEANS BERDASARKAN MODEL RECENCY FREQUENCY MONETARY," vol. 7, no. 2, 2021, [Online]. Available: <http://ejournal.fikom-unasman.ac.id>
- [9] A. Febriani and S. A. Putri, "Segmentasi Konsumen Berdasarkan Model Recency, Frequency, Monetary dengan Metode K-Means," *Jiems (Journal Ind. Eng. Manag. Syst.)*, vol. 13, no. 2, Sep. 2020, doi: 10.30813/jiems.v13i2.2274.
- [10] N. H. Harani, C. Prianto, and F. A. Nugraha, "Segmentasi Pelanggan Produk Digital Service Indihome Menggunakan Algoritma K-Means Berbasis Python," *J. Manaj. Inform.*, 2020, doi: 10.34010/jamika.v10i2.
- [11] A. R. Mulyawan, D. Gunawan, H. Basri, S. Alfarizi, and N. Ichsan, "Penerapan K-Medoids Clustering Dan Silhouette Method Untuk Strategi Pemasaran Program Donasi Pada Lembaga Amil Zakat," *Inf. Syst. Educ. Prof. J. Inf. Syst.*, vol. 8, no. 1, p. 107, 2023, doi: 10.51211/isbi.v8i1.2468.
- [12] D. P. Hidayatullah<sup>1</sup>, R. I. Rokhmawati<sup>2</sup>, and A. R. Perdanakusuma<sup>3</sup>, "Analisis Pemetaan Pelanggan Potensial Menggunakan Algoritma K-Means dan LRFM Model Untuk Mendukung Strategi Pengelolaan Pelanggan (Studi Pada



- Maninjau Center Kota Malang),” 2018. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [13] Putra M, Julhandri, Khoiru I, and Monalisa S, “Segmentasi pelanggan feandra cake,” 2020.
- [14] J. S. Lee and S. P. Jun, “Privacy-preserving data mining for open government data from heterogeneous sources,” *Gov. Inf. Q.*, vol. 38, no. 1, p. 101544, 2021, doi: 10.1016/j.giq.2020.101544.
- [15] B. E. Adiana, I. Soesanti, and A. E. Permasari, “Analisis Segmentasi Pelanggan Menggunakan Kombinasi RFM Model dan Teknik Clustering,” no. 2, 2018, doi: 10.21460/jutei.2017.21.76.
- [16] N. Ahsina, F. Fatimah, and F. Rachmawati, “Analisis Segmentasi Pelanggan Bank Berdasarkan Pengambilan Kredit Dengan Menggunakan Metode K-Means Clustering,” *J. Ilm. Teknol. Infomasi Terap.*, vol. 8, no. 3, 2022, doi: 10.33197/jitter.vol8.iss3.2022.883.
- [17] A. A. Firdaus, N. Iksan, D. N. Sadiyah, L. Sagita, and D. Setiawan, “Penerapan Algoritma Apriori untuk Prediksi Kebutuhan Suku Cadang Mobil,” *J. Sist. dan Teknol. Inf.*, vol. 9, no. 1, p. 13, Jan. 2021, doi: 10.26418/justin.v9i1.41151.
- [18] M. Jordy, A. Triayudi, and I. D. Sholihati, “Analisis Segmentasi Recency dan Customer Value Pada AVANA Indonesia Dengan Algoritma K-Means dan Model RFM (Recency, Frequency and Monetary),” *J. Inf. Syst. Res.*, vol. 4, no. 2, pp. 579–589, 2023, doi: 10.47065/josh.v4i2.2950.
- [19] A. Supriyadi, A. Triayudi, and I. D. Sholihati, “Perbandingan Algoritma K-Means Dengan K-Medoids Pada Pengelompokan Armada Kendaraan Truk Berdasarkan Produktivitas,” *JIPPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 6, no. 2, pp. 229–240, 2021, doi: 10.29100/jipi.v6i2.2008.