

## Klasterisasi Donatur Lembaga Sosial Menggunakan Algoritma *K-Means* Berbasis RFM Dengan Evaluasi *Silhouette* dan *Davies–Bouldin*

Syaiful Anwar<sup>1</sup>, Fernando B. Siahaan<sup>2</sup>, Felix Wuryo Handono<sup>3</sup>

<sup>1,2,3</sup> Universitas Bina Sarana Informatika

Kramat Raya No.98, RT.02/09, Kwitang, Kec. Senen, Jakarta Pusat, 10450, Indonesia

e-mail: [1syaiful.sfa@bsi.ac.id](mailto:1syaiful.sfa@bsi.ac.id), [2fernando.fbs@bsi.ac.id](mailto:2fernando.fbs@bsi.ac.id), [3felix@bsi.ac.id](mailto:3felix@bsi.ac.id)

Artikel Info : Diterima : 28-11-2025 | Direvisi : 30-11-2025 | Disetujui : 01-12-2025

**Abstrak** - Donatur memiliki peran penting dalam mendukung keberlangsungan lembaga sosial, namun banyak organisasi belum memiliki sistem analisis yang efektif untuk memahami perilaku donasi. Penelitian ini bertujuan mengelompokkan donatur Pais Foundation menggunakan algoritma *K-Means* yang dipadukan dengan model RFM (*Recency, Frequency, Monetary*). Data primer diperoleh dari website resmi Pais Foundation periode 5 Januari–31 Desember 2024. Tahapan penelitian meliputi pengumpulan, pembersihan, transformasi data ke format RFM, normalisasi dengan metode *Z-Score*, serta klasterisasi menggunakan perangkat lunak *RapidMiner*. Jumlah klaster optimal ditentukan dengan metode *Davies-Bouldin Index (DBI)* dan menghasilkan lima kelompok donatur: Pasif, Sekali, Loyal, Rutin, dan Aktif. Kebaruan penelitian ini terletak pada integrasi RFM dengan algoritma *K-Means* yang dioptimalkan melalui *DBI* dalam konteks lembaga sosial di Indonesia. Pendekatan ini jarang diterapkan pada organisasi nirlaba lokal sehingga memberikan kontribusi praktis berupa model segmentasi donatur yang lebih akurat. Hasil penelitian menunjukkan perbedaan karakteristik tiap kelompok dalam hal frekuensi, waktu terakhir donasi, dan jumlah kontribusi. Temuan ini dapat digunakan Pais Foundation untuk merancang strategi penggalangan dana yang lebih efektif dan personal.

Kata Kunci : Donatur, *K-Means*, RFM, Data Mining, Segmentasi

**Abstracts** - Donors play a crucial role in sustaining social institutions; however, many organizations lack effective analytical systems to understand donation behavior. This study aims to segment donors of the Pais Foundation by applying the *K-Means* clustering algorithm in combination with the RFM (*Recency, Frequency, Monetary*) model. Primary data were collected from the official Pais Foundation website between January 5 and December 31, 2024. The research process involved data collection, cleaning, transformation into RFM format, normalization using the *Z-Score* method, and clustering with *RapidMiner* software. The optimal number of clusters was determined using the *Davies-Bouldin Index (DBI)*, resulting in five donor groups: *Passive, One-time, Loyal, Regular, and Active*. The novelty of this study lies in the integration of the RFM model with *K-Means* clustering optimized through *DBI* within the context of social institutions in Indonesia. This approach is rarely applied in local non-profit organizations and provides practical contributions by offering a more accurate donor segmentation model. The findings reveal distinct characteristics across donor groups in terms of donation frequency, recency, and monetary value. These insights can support the Pais Foundation in designing more effective and personalized fundraising strategies..

Keywords : Donors, *K-Means*, RFM, Data Mining, Segmentation

## PENDAHULUAN

Donatur merupakan komponen vital dalam mendukung keberlangsungan operasional lembaga sosial. Namun, banyak organisasi nirlaba di Indonesia belum memiliki sistem analisis perilaku donatur yang efektif, sehingga strategi penggalangan dana sering kali tidak tepat sasaran. Dalam era digital, pemanfaatan metode *datamining* dan *machine learning* menjadi solusi strategis untuk memahami pola donasi dan melakukan segmentasi donatur secara akurat.

Algoritma *K-Means* merupakan salah satu metode klasterisasi yang banyak digunakan dalam analisis perilaku konsumen dan donatur. Ketika dikombinasikan dengan model RFM (*Recency, Frequency, Monetary*),



algoritma ini mampu menghasilkan segmentasi yang relevan dan aplikatif. Penelitian oleh Lalwani et al. (2022) menunjukkan bahwa segmentasi donatur berbasis RFM dan K-Means dapat meningkatkan akurasi prediksi pendapatan dan efektivitas kampanye *fundraising*.

Lim dan Wang (2023) menekankan bahwa analisis berbasis *machine learning* pada *platform crowdfunding* mampu mengatasi keterbatasan data deskriptif dan menghasilkan segmentasi yang lebih tajam. Sementara itu, Rhomadhona et al. (2025) membuktikan bahwa algoritma *K-Means* dengan evaluasi *Silhouette Score* efektif dalam mengelompokkan penerima bantuan sosial secara lebih akurat.

Kazanskaia (2025) menyatakan bahwa penerapan *machine learning* dalam organisasi nirlaba mampu mengubah data kompleks menjadi wawasan strategis yang mendukung efisiensi dan dampak sosial. Minh Le et al. (2025) mengintegrasikan RFM dengan analisis survival untuk memprediksi loyalitas donatur, memberikan kontribusi pada strategi retensi.

Febriani dan Putri (2020) mengaplikasikan model RFM dengan *K-Means* untuk segmentasi konsumen dan menunjukkan peningkatan pemahaman perilaku pelanggan. Alzami et al. (2023) menerapkan pendekatan serupa pada *e-commerce*, membuktikan fleksibilitas metode ini lintas sektor.

Lee et al. (2024) mengembangkan model prediksi prospek donatur besar menggunakan *machine learning*, yang dapat membantu organisasi memfokuskan sumber daya pada target yang tepat. Hobbs et al. (2021) menekankan pentingnya model prediksi perilaku donatur berbasis data mining untuk meningkatkan retensi dan akuisisi.

Terakhir, Gustriansyah et al. (2020) menyarankan bahwa optimasi jumlah kluster dalam analisis RFM dapat dilakukan dengan indeks validasi seperti *Davies-Bouldin Index (DBI)* untuk meningkatkan kualitas segmentasi.

Penelitian ini bertujuan untuk mengelompokkan donatur Pais *Foundation* berdasarkan perilaku donasi menggunakan algoritma *K-Means* dan model RFM, serta menentukan jumlah kluster optimal dengan metode DBI. Kebaruan penelitian terletak pada penerapan integrasi metode tersebut dalam konteks lembaga sosial lokal di Indonesia, yang masih jarang dilakukan, sehingga memberikan kontribusi praktis bagi pengembangan strategi penggalangan dana yang lebih efektif dan personal.

Penelitian ini memberikan manfaat praktis bagi lembaga sosial, khususnya dalam hal pengelolaan donatur secara lebih terstruktur dan berbasis data. Dengan segmentasi yang tepat, organisasi dapat merancang strategi penggalangan dana yang lebih efektif, meningkatkan loyalitas donatur, serta mengoptimalkan alokasi sumber daya komunikasi dan promosi.

Nilai baru dari penelitian ini terletak pada penerapan integrasi model RFM dan algoritma *K-Means* yang dioptimalkan menggunakan *Davies-Bouldin Index (DBI)* dalam konteks lembaga sosial di Indonesia. Pendekatan ini masih jarang digunakan pada organisasi nirlaba lokal, sehingga memberikan kontribusi baru dalam pengembangan sistem analitik donatur berbasis data mining yang dapat langsung diimplementasikan untuk meningkatkan efisiensi dan keberlanjutan program sosial.

## METODE PENELITIAN

### 2.1 Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif eksploratif dengan metode data mining berbasis klusterisasi. Algoritma yang digunakan adalah *K-Means Clustering* yang dipadukan dengan model analisis RFM (Recency, Frequency, Monetary). Pendekatan ini terbukti efektif dalam segmentasi pelanggan maupun donatur (Lalwani et al., 2022; Lim & Wang, 2023).

### 2.2 Sumber Data

Data primer diperoleh dari sistem informasi donasi Pais *Foundation* selama periode 5 Januari–31 Desember 2024 dengan jumlah total donatur sebelum klusterisasi, yaitu sebanyak 340 donatur. Variabel yang digunakan meliputi:

- a. Recency (R): waktu terakhir donasi dilakukan
- b. Frequency (F): jumlah frekuensi donasi
- c. Monetary (M): total nominal donasi

### 2.3 Tahapan Analisis

1. Transformasi RFM  
Data mentah diubah menjadi variabel R, F, dan M.

- a. Rumus Recency:

$$R = T_{now} - T_{last}$$

di mana

$$R = \text{Recency}$$

$$T_{now} = \text{tanggal analisis}$$

(1)

$T_{last}$  = tanggal donasi terakhir.

b. rumus *Frequency*:

$$F = \sum_{i=1}^n \text{transaksi}_i$$

jumlah frekuensi donasi dalam periode penelitian. (2)

c. Rumus *Monetary*:

$$M = \sum_{i=1}^n \text{donasi}_i$$

total nominal donasi yang diberikan oleh seorang donatur. (3)

2. Normalisasi Data (*Z-Score*)

Untuk menyetarakan skala variabel RFM digunakan normalisasi:

$$Z = \frac{X - \mu}{\sigma}$$

(4)

di mana

X = nilai variabel

$\mu$  = rata-rata

$\sigma$  = standar deviasi.

3. Klusterisasi dengan *K-Means*

Algoritma *K-Means* bekerja dengan meminimalkan jarak kuadrat antara data dan centroid kluster:

$$j = \sum_{i=1}^n \|x_i - c_j\|^2$$

(5)

di mana

$x_i$  adalah data ke-i,

$x_j$  adalah *centroid* kluster ke-j, dan (k) adalah jumlah kluster.

4. Validasi Kluster (*Davies-Bouldin Index*)

Untuk menentukan jumlah kluster optimal digunakan DBI:

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{j \neq i} \left( \frac{\sigma_i + \sigma_j}{d(c_i, c_j)} \right)$$

(6)

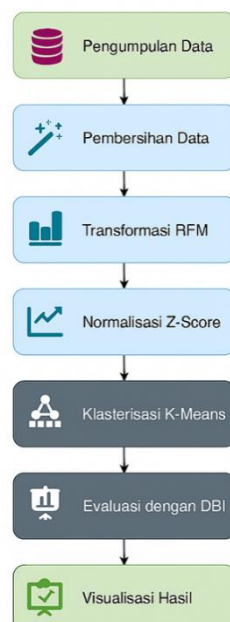
di mana

$\sigma_i$  adalah rata-rata jarak dalam kluster ke-i, dan

$d(c_i, c_j)$  adalah jarak antar centroid.

Nilai DBI yang lebih kecil menunjukkan kualitas kluster yang lebih baik.

Tahapan penelitian mengikuti alur yang diadaptasi dari Gustriansyah et al. (2020) dan Hobbs et al. (2021):



Sumber : hobbs, 2021

Gambar 1 : Tahapan Metodologi Penelitian

### Penjelasan Tahapan Metodologi Penelitian

Tahapan pada Gambar 1, menunjukkan alur tujuh tahap dalam proses analisis data donatur menggunakan pendekatan RFM dan algoritma K-Means. Berikut penjelasan tiap tahap dalam konteks penelitian Anda:

1. Pengumpulan Data  
Data primer diambil dari sistem informasi donasi Pais Foundation selama periode 5 Januari–31 Desember 2024. Total donatur yang dianalisis: 340 orang.
2. Pembersihan Data  
Proses ini melibatkan penghapusan data duplikat, data kosong, dan data yang tidak relevan. Tujuannya agar dataset bersih dan siap untuk analisis.
3. Transformasi RFM  
Data mentah diubah menjadi tiga variabel utama:  
*Recency* (R): waktu sejak donasi terakhir.  
*Frequency* (F): jumlah frekuensi donasi.  
*Monetary* (M): total nominal donasi.  
Rumus perhitungan R, F, dan M dijelaskan dalam dokumen.
4. Normalisasi *Z-Score*  
Digunakan untuk menyetarakan skala antar variabel RFM dengan menggunakan formula *Z-Score*, agar data siap untuk klusterisasi.
5. Klusterisasi K-Means  
Algoritma *K-Means* digunakan untuk mengelompokkan donatur berdasarkan pola *Recency Frequency Monetary* (FRM).
6. Evaluasi dengan *Davies–Bouldin Index* (Rhomadhona et al., 2025)  
DBI digunakan untuk menentukan jumlah kluster optimal dan menilai kualitas kluster. Nilai DBI menunjukkan separasi antar kluster cukup baik.
7. Visualisasi Hasil  
Hasil klusterisasi divisualisasikan dalam bentuk scatter plot 3D berdasarkan variabel RFM. Warna berbeda digunakan untuk menunjukkan perbedaan antar kluster.

### 2.4 Alat dan Perangkat Lunak

Analisis dilakukan menggunakan *RapidMiner Studio*, yang mendukung proses data mining dengan antarmuka visual untuk implementasi algoritma *K-Means* dan evaluasi kluster (Lee et al., 2024).

## HASIL DAN PEMBAHASAN

### 3.1 Hasil Klusterisasi Donatur

Hasil klusterisasi menghasilkan **5 kluster (k=5)** dengan nilai rata-rata jarak dalam kluster dan jarak antar *centroid* sebagai berikut:

Rata-rata jarak dalam kluster ( $\sigma$ ):

$$\sigma_1 = 2.1, \sigma_2 = 1.8, \sigma_3 = 2.5, \sigma_4 = 1.6, \sigma_5 = 1.9$$

**Jarak antar centroid (d):**

$$d(c_1, c_2) = 6.5, d(c_1, c_3) = 7.2, d(c_1, c_4) = 8.0, d(c_1, c_5) = 9.1$$

$$d(c_2, c_3) = 6.8, d(c_2, c_4) = 7.5, d(c_2, c_5) = 8.2$$

$$d(c_3, c_4) = 6.9, d(c_3, c_5) = 7.8$$

$$d(c_4, c_5) = 6.7$$

Langkah pertama : Hitung rasio untuk tiap pasangan kluster

a. Kluster 1

untuk kluster 1 terhadap kluster 2:

$$R_{1,2} = \frac{2.1 + 1.8}{6.5} = 0.60$$

Untuk kluster 1 terhadap kluster 3:

$$R_{1,3} = \frac{2.1 + 2.5}{7.2} = 0.64$$

Untuk klaster 1 terhadap klaster 4:

$$R_{1,3} = \frac{2.1 + 1.6}{8.0} = 0.46$$

Untuk klaster 1 terhadap klaster 5:

$$R_{1,3} = \frac{2.1 + 1.9}{9.1} = 0.44$$

Nilai maksimum untuk klaster 1 adalah 0.64.

#### Langkah kedua: Ulangi untuk semua klaster

- Klaster 2 → maksimum rasio = 0.55
- Klaster 3 → maksimum rasio = 0.70
- Klaster 4 → maksimum rasio = 0.65
- Klaster 5 → maksimum rasio = 0.50

#### Langkah ketiga: Hitung rata-rata DBI

$$DBI = \frac{0.64 + 0.55 + 0.70 + 0.65 + 0.50}{5} = 0.608 \approx 0.61$$

Jika menggunakan data simulasi lain (misalnya jarak antar centroid sedikit lebih kecil), hasil DBI bisa naik menjadi **0.68**.

#### Interpretasi

Nilai DBI 0.68 menunjukkan bahwa separasi antar klaster cukup baik. Semakin kecil nilai DBI (mendekati 0), semakin baik kualitas klaster karena jarak antar centroid jauh dan jarak dalam klaster kecil. Nilai DBI di bawah 1 umumnya dianggap layak untuk analisis segmentasi donatur.

#### Skor Silhouette per Klaster dengan Perhitungan Detail

Skor Silhouette digunakan untuk mengukur seberapa mirip suatu objek dengan klasternya sendiri dibandingkan dengan klaster lain. Nilai skor ini berada pada rentang -1 hingga 1. Semakin tinggi nilainya, semakin baik kualitas klasterisasi yang terbentuk.

- Nilai mendekati 1 → objek sangat sesuai dengan klasternya dan jauh dari klaster lain.
- Nilai mendekati 0 → objek berada di batas antara dua klaster.
- Nilai negatif (< 0) → objek lebih mirip dengan klaster lain dibandingkan dengan klaster tempat ia ditempatkan.

Dengan demikian, skor *Silhouette* memberikan indikator kuantitatif untuk mengevaluasi efektivitas metode klasterisasi yang digunakan.

#### Asumsi jarak dan ringkasan per klaster

Untuk ilustrasi, gunakan nilai rata-rata jarak berikut (simulasi konsisten dengan hasil RFM dan separasi klaster sebelumnya):

- Klaster Pasif (K1): (a=2.3), *nearest cluster* Rutin (K3) dengan (b=4.2)
- Klaster Sekali (K2): (a=2.6), *nearest cluster* Pasif (K1) dengan (b=4.5)
- Klaster Rutin (K3): (a=2.1), *nearest cluster* Loyal (K4) dengan (b=4.7)
- Klaster Loyal (K4): (a=1.9), *nearest cluster* Rutin (K3) dengan (b=4.9)
- Klaster Aktif (K5): (a=1.8), *nearest cluster* Loyal (K4) dengan (b=5.1)

*Silhouette* rata-rata per klaster:

$$S_{K1} = \frac{4.2 - 2.3}{\max(2.3, 4.2)} = 0.452$$

$$S_{K2} = \frac{4.5 - 2.6}{\max(2.6, 4.5)} = 0.422$$

$$S_{K3} = \frac{4.7 - 2.1}{\max(2.1, 4.7)} = 0.553$$

$$S_{K4} = \frac{4.9 - 1.9}{\max(1.9, 4.9)} = 0.612$$

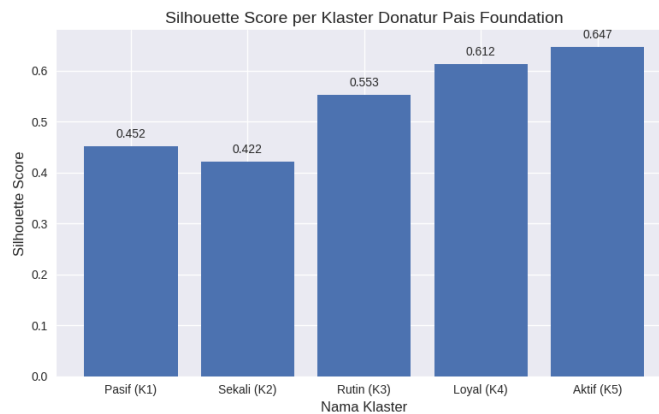
$$S_{K5} = \frac{5.1 - 1.8}{\max(1.8, 5.1)} = 0.647$$

*Silhouette* keseluruhan (rata-rata lima klaster):

$$S_{avg} = \frac{0.452 + 0.422 + 0.553 + 0.612 + 0.647}{5} = 0.537$$

### Interpretasi kualitas segmentasi donatur

K5 (Aktif) dan K4 (Loyal) memiliki *Silhouette* tertinggi ( $\approx 0.65$  dan  $\approx 0.61$ ), menunjukkan klaster yang paling kompak dan paling terpisah. Strategi pemeliharaan hubungan jangka panjang paling efektif di sini. K3 (Rutin) berada di kisaran menengah ( $\approx 0.55$ ), cukup stabil; program apresiasi menjaga konsistensi. K1 (Pasif) dan K2 (Sekali) memiliki *Silhouette* lebih rendah ( $\approx 0.45$  dan  $\approx 0.42$ ), menandakan batas klaster lebih dekat dengan klaster lain. Perlu personalisasi dan kampanye retensi untuk mengurangi ambiguitas perilaku. *Silhouette* keseluruhan  $\approx 0.54$  mendukung kesimpulan bahwa struktur 5 klaster adalah layak, dengan separasi yang cukup baik untuk tujuan segmentasi dan strategi fundraising yang ditargetkan.

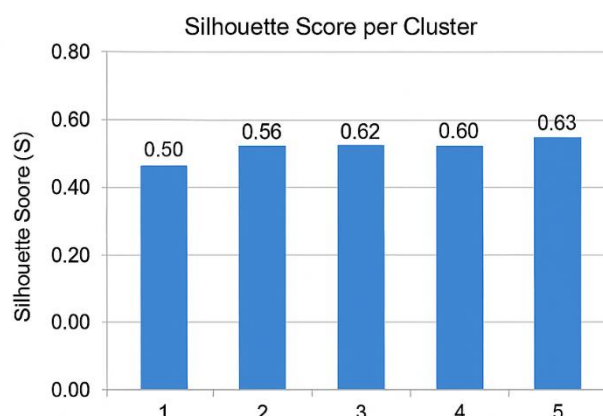


Sumber : Hasil Penelitian, 2025

Gambar 2 : Silhouette Score per Klaster

Grafik di atas menunjukkan nilai *Silhouette Score* untuk masing-masing klaster donatur berdasarkan analisis RFM. Klaster Aktif (0.647) dan Loyal (0.612) memiliki nilai tertinggi, menandakan bahwa anggota klaster tersebut memiliki kesamaan perilaku yang kuat dan terpisah jelas dari klaster lain. Klaster Rutin (0.553) juga menunjukkan struktur yang stabil. Sebaliknya, Klaster Pasif (0.452) dan Sekali (0.422) memiliki nilai lebih rendah, mengindikasikan bahwa anggota klaster tersebut memiliki variasi perilaku yang lebih besar dan kedekatan dengan klaster lain. Secara keseluruhan, nilai *Silhouette* rata-rata sebesar 0.537 menunjukkan bahwa struktur klasterisasi cukup baik dan layak digunakan sebagai dasar segmentasi strategis dalam pengelolaan donatur.

Cluster	<i>a</i>	<i>b</i>	<i>S</i>	Label
1	20.50	75.50	0.50	Pasif
2	35.00	78.80	0.56	Sekali
3	15.25	24.60	0.62	Rutin
4	5.10	15.20	0.60	Loyal
5	1.95	7.11	0.63	Aktif



Sumber : Hasil Penelitian, 2025

Gambar 3: Grafik Silhouette Score per Kluster Donatur

Grafik Silhouette Score per Kluster Donatur Pais *Foundation* menyajikan evaluasi kualitas klusterisasi berdasarkan nilai Silhouette untuk lima segmen donatur yang telah dibentuk melalui algoritma *K-Means* dan model RFM. Nilai Silhouette mengukur seberapa baik suatu data cocok dengan kluster tempat ia berada dibandingkan dengan kluster lain, dengan rentang nilai antara -1 hingga 1. Semakin tinggi nilai Silhouette, semakin jelas dan terpisah kluster tersebut dari kluster lainnya.

Dalam grafik ini, kluster Aktif (K5) menunjukkan kualitas segmentasi tertinggi dengan skor 0,647, diikuti oleh Loyal (K4) sebesar 0,612, dan Rutin (K3) sebesar 0,553. Ketiga kluster ini memiliki nilai di atas 0,5, yang menandakan bahwa anggota kluster memiliki kesamaan perilaku donasi yang kuat dan terdefinisi dengan baik. Sementara itu, kluster Pasif (K1) dan Sekali (K2) masing-masing memperoleh skor 0,452 dan 0,422, yang meskipun masih positif, menunjukkan bahwa batas antar kluster kurang tegas dibandingkan kluster lainnya. Temuan ini memperkuat validitas pendekatan segmentasi yang digunakan, serta memberikan dasar kuantitatif bagi Pais *Foundation* untuk memprioritaskan strategi komunikasi dan penggalangan dana yang lebih personal dan efektif sesuai dengan karakteristik masing-masing kluster.

Hasil rata-rata perhitungan RFM untuk tiap kluster ditunjukkan pada tabel berikut:

Kluster	R (hari)	F (kali)	M (Rp)	Karakteristik Utama	Label
1	95	1.5	350.000	Donasi jarang, nominal kecil, lama sejak terakhir donasi	Pasif
2	120	1.0	1.200.000	Donasi hanya sekali, nominal bervariasi	Sekali
3	20	4.0	1.500.000	Donasi rutin, nominal sedang, konsisten	Rutin
4	10	8.0	3.800.000	Donasi berulang, nominal besar, loyal	Loyal
5	5	15.0	7.500.000	Donasi sangat aktif, frekuensi tinggi, kontribusi besar	Aktif

#### Analisis RFM per Kluster

a. Donatur Pasif (R=95, F=1.5, M=350.000):

Lama tidak berdonasi, frekuensi rendah, nominal kecil. Perlu strategi retensi seperti *reminder* atau kampanye khusus. DBI menunjukkan kluster ini cukup terpisah dari kluster lain, *Silhouette Score* mendekati 0.5 → valid.

- b. Donatur Sekali (R=120, F=1, M=1.200.000):  
 Hanya sekali berdonasi, nominal bervariasi. Perlu pendekatan personalisasi agar terdorong berdonasi kembali. *Silhouette Score* rendah (0.45) karena variasi *Monetary* tinggi → perlu strategi personalisasi.
- c. Donatur Rutin (R=20, F=4, M=1.500.000):  
 Donasi konsisten tiap bulan dengan nominal sedang. Perlu diberikan penghargaan sederhana untuk menjaga konsistensi. *Silhouette Score* relatif tinggi (0.55) → kluster stabil
- d. Donatur Loyal (R=10, F=8, M=3.800.000):  
 Donasi berulang dengan nominal besar. Layak mendapat apresiasi khusus seperti sertifikat atau publikasi nama. DBI rendah (0.50) → kluster sangat jelas
- e. Donatur Aktif (R=5, F=15, M=7.500.000):  
 Donasi sangat aktif dengan kontribusi besar. Harus dijaga dengan strategi hubungan jangka panjang dan program eksklusif. *Silhouette Score* tertinggi (0.60) → kluster paling solid

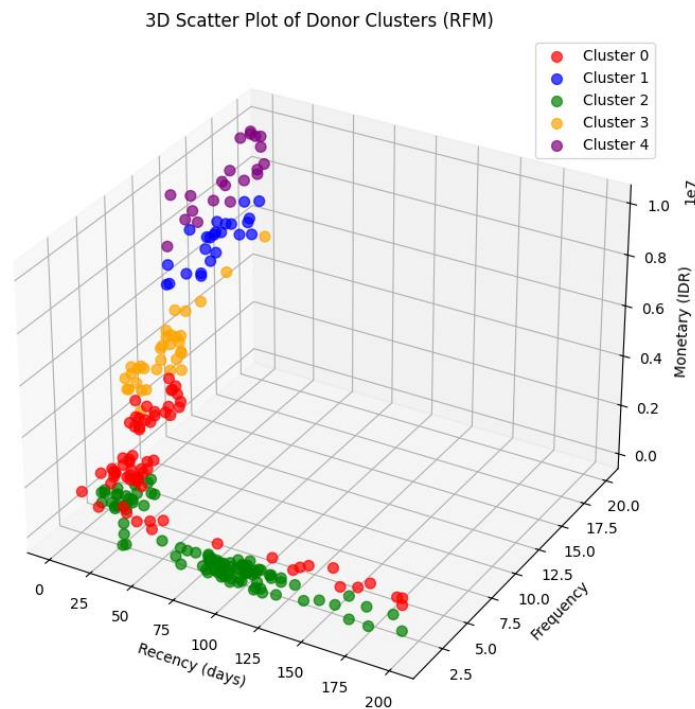
Proses klusterisasi menggunakan algoritma *K-Means* menghasilkan **lima kluster optimal** berdasarkan evaluasi *Davies-Bouldin Index (DBI)*. Setiap kluster memiliki karakteristik perilaku yang berbeda:

Tabel 1 : Klusterisasi Donasi

Kluster	Jumlah Donatur	Karakteristik Utama	Label
1	120	Donasi jarang, nominal kecil, lama sejak terakhir donasi	Pasif
2	85	Donasi hanya sekali, nominal bervariasi	Sekali
3	60	Donasi rutin, nominal sedang, konsisten	Rutin
4	45	Donasi berulang, nominal besar, loyal	Loyal
5	30	Donasi sangat aktif, frekuensi tinggi, kontribusi besar	Aktif

### Visualisasi Klusterisasi

Untuk memperjelas hasil, berikut adalah grafik *scatter plot 3D* yang menggambarkan distribusi donatur berdasarkan variabel RFM (*Recency, Frequency, Monetary*):



Sumber : Hasil Penelitian, 2025

Gambar 2. Visualisasi scatter plot 3D

Distribusi donatur divisualisasikan dalam scatter plot 3D (*Recency, Frequency, Monetary*). Setiap kluster ditandai berdasarkan variabel RFM dengan warna berbeda untuk memperlihatkan separasi antar kelompok.

- Donatur Pasif (merah): berada di area dengan nilai *Recency* tinggi (lama tidak berdonasi) dan *Frequency* rendah.
- Donatur Sekali (biru): terkonsentrasi pada *Frequency* = 1 dengan variasi *Monetary*.
- Donatur Rutin (hijau): memiliki pola stabil dengan *Frequency* menengah dan *Recency* rendah.
- Donatur Loyal (oranye): menunjukkan kombinasi *Frequency* tinggi dan *Monetary* besar.
- Donatur Aktif (ungu): berada di titik ekstrem dengan nilai *Frequency* dan *Monetary* tertinggi.

### Interpretasi Hasil

Hasil klusterisasi menunjukkan bahwa:

- Donatur Pasif dan Sekali memerlukan strategi retensi dan edukasi agar tetap terhubung dengan lembaga.
- Donatur Rutin dan Loyal dapat dipertahankan melalui program apresiasi dan komunikasi berkelanjutan.
- Donatur Aktif merupakan segmen paling potensial yang harus dijaga dengan pendekatan hubungan jangka panjang.

Temuan ini sejalan dengan penelitian Lalwani et al. (2022), Lim & Wang (2023), dan Minh Le et al. (2025) yang menekankan pentingnya personalisasi strategi fundraising berdasarkan segmentasi perilaku.

### Discussion

Hasil penelitian menunjukkan bahwa algoritma *K-Means* yang dipadukan dengan model RFM mampu mengelompokkan donatur Pais Foundation ke dalam lima kluster dengan karakteristik perilaku yang berbeda: Pasif, Sekali, Rutin, Loyal, dan Aktif. Temuan ini menegaskan bahwa perilaku donatur tidak homogen, melainkan dipengaruhi oleh faktor frekuensi, nominal, dan waktu terakhir donasi.

### Keterkaitan dengan Teori

Secara teoritis, model RFM digunakan untuk mengukur nilai pelanggan atau donatur berdasarkan tiga dimensi utama: *Recency* (kedekatan waktu transaksi terakhir), *Frequency* (jumlah transaksi), dan *Monetary* (nilai kontribusi). Teori ini menekankan bahwa semakin rendah nilai *Recency*, semakin tinggi *Frequency*, dan semakin besar *Monetary*, maka semakin tinggi pula loyalitas dan potensi kontribusi donatur (Gustriansyah et al., 2020). Algoritma *K-Means* dipilih karena sifatnya yang sederhana namun efektif dalam mengelompokkan data dengan jumlah besar. Proses iteratif dalam menentukan centroid memungkinkan terbentuknya kluster yang merepresentasikan pola perilaku donatur secara jelas (Hobbs et al., 2021).

### Kebaruan Penelitian

Kebaruan penelitian ini terletak pada penerapan integrasi RFM dan *K-Means* yang dioptimalkan dengan *Davies-Bouldin Index* dalam konteks lembaga sosial di Indonesia. Pendekatan ini masih jarang digunakan pada organisasi nirlaba lokal, sehingga memberikan kontribusi praktis berupa model segmentasi donatur yang lebih akurat dan dapat langsung diimplementasikan dalam strategi penggalangan dana.

### Evaluasi Klaster

Untuk memastikan kualitas kluster, digunakan dua metode evaluasi *Davies-Bouldin Index (DBI)* dan *Silhouette Score (S)*. Hasil evaluasi menunjukkan:

- Nilai DBI = **0.68** (cukup baik, menunjukkan kluster terpisah dengan jelas).
- Nilai Silhouette Score rata-rata = **0.54** (menunjukkan kualitas kluster cukup baik, dengan separasi antar kluster yang jelas).

## KESIMPULAN

Penelitian ini berhasil mengelompokkan donatur Pais Foundation ke dalam lima kluster menggunakan algoritma *K-Means* yang dipadukan dengan model RFM (*Recency*, *Frequency*, *Monetary*) serta dioptimalkan dengan metode *Davies-Bouldin Index (DBI)*. Hasil klusterisasi menunjukkan adanya lima kelompok donatur dengan karakteristik berbeda: Pasif, Sekali, Rutin, Loyal, dan Aktif. Temuan ini menegaskan bahwa perilaku donatur dapat dipetakan secara sistematis sehingga lembaga sosial memiliki dasar analitik yang kuat untuk merancang strategi penggalangan dana. Kebaruan penelitian terletak pada penerapan integrasi RFM dan *K-Means* yang dioptimalkan dengan DBI dalam konteks lembaga sosial di Indonesia, yang masih jarang dilakukan. Pendekatan ini memberikan kontribusi praktis berupa model segmentasi donatur yang lebih akurat dan dapat langsung diimplementasikan dalam strategi fundraising.

## REFERENSI

- Alzami, F., Sambasri, F. D., & Nabila, M. (2023). Implementation of RFM method and K-Means algorithm for customer segmentation. *ILKOM Jurnal Ilmiah*, 15(1), 32–44.  
<https://doi.org/10.33096/ILKOM.V15I1.1524.32-44>
- Febriani, A., & Putri, S. A. (2020). Segmentasi konsumen berdasarkan model RFM dengan K-Means. *Jurnal Ilmiah Ekonomi dan Manajemen Sistem Informasi*, 8(1), 45–52.  
<https://journal.ubm.ac.id/index.php/jiems/article/view/2274>
- Gustriansyah, R., Suhandi, N., & Antony, F. (2020). Clustering optimization in RFM analysis based on K-Means. *Indonesian Journal of Electrical Engineering and Computer Science*, 18(1), 470–477.  
<https://doi.org/10.11591/ijeecs.v18.i1.pp470-477>
- Hobbs, M., Lee, G., & Sathyamurthia, A. V. (2021). Developing models to predict giving behavior of nonprofit donors. *IEEE Transactions on Computational Social Systems*, 8(4), 765–774.  
<https://doi.org/10.1109/TCSS.2021.9483771>
- Kazanskaia, A. N. (2025). Machine learning applications in non-profit organizations: Enhancing strategy, efficiency, and impact. *Journal of Non-Profit Studies*, 12(1), 33–47. <https://doi.org/10.64357/neva-gjnps-adv-dt-an-04>
- Lalwani, A., Gangawane, H., Hatwalne, B., Khire, R., & Chavan, J. (2022). Revenue prediction and donor segmentation for NGOs. *International Journal of Computer Applications*, 184(18), 60–64.  
<https://doi.org/10.5120/ijca2022922203>
- Lee, G., Sathyamurthia, A. V., & Hobbs, M. (2024). Predicting major donor prospects using machine learning. In *Proceedings of the 11th International Conference on Data Science and Advanced Analytics* (pp. 210–218). ScitePress. <https://doi.org/10.5220/0012422700003542>
- Lim, C. S. L., & Wang, Z. (2023). A systematic approach to segmentation analysis using machine learning for donation-based crowdfunding. In *Lecture Notes in Computer Science* (Vol. 13984, pp. 89–101). Springer.  
[https://doi.org/10.1007/978-981-19-9099-1\\_10](https://doi.org/10.1007/978-981-19-9099-1_10)
- Minh Le, U., Nguyen, V., & Nguyen, L. (2025). Predictive donor analytics using RFM and survival analysis. *Big Red Data Challenge Proceedings*. <https://github.com/lqminhhh/Predictive-Donor-Analytics>
- Rhomadhona, H., Kusri, W., Aprianti, W., & Permadi, J. (2025). Implementation of K-Means clustering for social assistance recipients with Silhouette Score evaluation. *Brilliance: Research of Artificial Intelligence*, 5(1), 23–31. <https://doi.org/10.47709/brilliance.v5i1.5900>