



PENERAPAN ALGORITMA *K-MEANS* DAN *K-MEDOIDS* PADA PEUBAH *RECENCY, FREQUENCY, MONETARY* (RFM) UNTUK SEGMENTASI PELANGGAN DI SWALAYAN XYZ

MUHAMMAD DHIYAAUL FIRDAUS



**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Penerapan Algoritma *K-Means* dan *K-Medoids* pada Peubah *Recency, Frequency, Monetary* (RFM) untuk Segmentasi Pelanggan di Swalayan XYZ” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2025

Muhammad Dhiyaul Firdaus
G14190021

Hak Cipta dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

ABSTRAK

MUHAMMAD DHIYAAUL FIRDAUS. Penerapan Algoritma *K-Means* dan *K-Medoids* pada Peubah *Recency*, *Frequency*, *Monetary* (RFM) untuk Segmentasi Pelanggan di Swalayan XYZ. Dibimbing oleh BAGUS SARTONO dan ERFIANI.

Persaingan yang semakin meningkat dalam industri ritel telah mendorong perusahaan ritel untuk memahami lebih baik preferensi dan perilaku pelanggan mereka. Penentuan segmentasi pelanggan yang akurat sangat penting dalam merancang strategi pemasaran yang efektif untuk meningkatkan kepuasan pelanggan dan memperoleh keuntungan yang lebih besar. Penelitian ini mengevaluasi penggunaan metode pengelompokan *K-Means* dan *K-Medoids* pada peubah RFM (*Recency*, *Frequency*, *Monetary*) untuk menentukan segmentasi pelanggan di swalayan XYZ. Penelitian ini menggunakan data transaksi dari pelanggan selama periode waktu tertentu untuk mengekstrak nilai RFM dari setiap pelanggan. Metode *K-Means* dan *K-Medoids* kemudian diterapkan untuk mengelompokkan pelanggan ke dalam segmen berdasarkan nilai RFM mereka. Tujuan dari penelitian ini adalah untuk membandingkan efektivitas kedua metode pengelompokan dalam menentukan segmentasi pelanggan. Hasil penelitian ini diharapkan dapat membantu Swalayan XYZ dalam memahami preferensi pelanggan dan memberikan rekomendasi strategi pemasaran. Hasil penggerombolan terbaik ditentukan berdasarkan parameter optimal yang dievaluasi menggunakan Davies-Bouldin Index, indeks Silhouette, dan analisis eksploratif. Metode analisis gerombol terbaik yang diperoleh adalah metode *K-Means* yang menghasilkan 4 gerombol.

Kata kunci: analisis RFM, *k-means*, *k-medoids*, manajemen hubungan pelanggan, segmentasi pelanggan



ABSTRACT

MUHAMMAD DHIYAAUL FIRDAUS. Implementation of K-Means and K-Medoids Algorithms on Recency, Frequency, Monetary (RFM) Variables for Customer Segmentation at XYZ Supermarket. Supervised by BAGUS SARTONO and ERFIANI.

The increasing competition in the retail industry has driven retail companies to better understand their customers' preferences and behaviors. Accurate customer segmentation is essential in designing effective marketing strategies to enhance customer satisfaction and generate greater profits. This study evaluates the use of K-Means and K-Medoids clustering methods on RFM (Recency, Frequency, Monetary) variables to determine customer segmentation at XYZ Supermarket. The research uses transaction data from customers over a specific period to extract the RFM values for each customer. The K-Means and K-Medoids methods are then applied to group customers into segments based on their RFM values. The purpose of this study is to compare the effectiveness of the two clustering methods in determining customer segmentation. The results of this research are expected to help XYZ Supermarket understand customer preferences and provide marketing strategy recommendations. The best clustering results were determined based on optimal parameters evaluated using the Davies-Bouldin Index, the Silhouette Index, and exploratory analysis. The best clustering method obtained was the K-Means method which produced four clusters.

Keywords: customer relationship management, customer segmentation, k-means, k-medoids, RFM analysis



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2025
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.



PENERAPAN ALGORITMA *K-MEANS* DAN *K-MEDOIDS* PADA PEUBAH *RECENCY, FREQUENCY, MONETARY* (RFM) UNTUK SEGMENTASI PELANGGAN DI SWALAYAN XYZ

MUHAMMAD DHIYAAUL FIRDAUS

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:

1 Cici Suhaeni, S.Si., M.Si., Ph.D



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

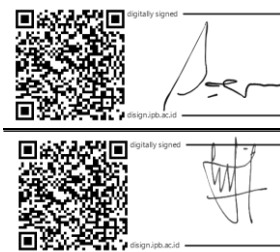
Judul Skripsi : Penerapan Algoritma *K-Means* dan *K-Medoids* pada Peubah
Recency, Frequency, Monetary (RFM) untuk Segmentasi
Pelanggan di Swalayan XYZ

Nama : Muhammad Dhiyaul Firdaus
NIM : G14190021

Disetujui oleh

Pembimbing 1:
Dr. Bagus Sartono, S.Si., M.Si.

Pembimbing 2:
Dr. Ir. Erfiani, M.Si.



Diketahui oleh

Ketua Program Studi Sarjana Statistika dan Sains Data:
Dr. Bagus Sartono, S.Si., M.Si.
NIP 19780411 200501 1 002



Tanggal Ujian:
12 Agustus 2025

Tanggal Lulus:



PRAKATA

Alhamdulillah, segala puji bagi Allah subhanaahu wa ta'ala atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan judul "Penerapan Algoritma *K-Means* dan *K-Medoids* pada Peubah *Recency, Frequency, Monetary* (RFM) untuk Segmentasi Pelanggan di Swalayan XYZ." Shawatil serta salam semoga senantiasa tercurah limpahkan kepada junjungan kita Nabi Muhammad shallallahu 'alaihi wasallam beserta keluarga dan para pengikutnya hingga akhir zaman.

Ucapan terima kasih penulis sampaikan kepada seluruh pihak yang telah memberikan dukungan, bimbingan, dan bantuan dalam proses penyusunan skripsi ini, antara lain:

1. Bapak Dr. Bagus Sartono, S.Si., M.Si. dan Ibu Dr. Ir. Erfiani, M.Si. selaku dosen pembimbing yang telah memberikan ilmu, bimbingan, dan saran kepada penulis dalam penyusunan skripsi ini.
2. Pimpinan perusahaan penyedia data Swalayan XYZ yang telah membantu penulis sehingga penulis dapat memulai penelitian dan menyusun skripsi ini.
3. Ayah, Ibu (almarhumah), dan Adik selaku penyemangat utama atas segala dukungan yang diberikan dan doa yang dipanjatkan.
4. Seluruh dosen pengajar dan tenaga pendidik Program Studi Sarjana Statistika dan Sains Data IPB atas ilmu, bimbingan, serta dukungan selama masa perkuliahan hingga penulis menyelesaikan skripsi ini.
5. Beasiswa YBM BRILiaN Bright Scholarship dan Pesantren Global Leadership yang telah memberikan dukungan berupa beasiswa penuh selama masa studi dan pembinaan komprehensif guna membentuk penulis menjadi pemimpin masa depan yang bervisi global, berkarakter, berdaya saing, dan mengaplikasikan nilai Islam dalam kehidupan.
6. Keluarga Statistika dan Sains Data IPB, Bright Scholarship IPB, Global Leadership, KKN Desa Gambirmanis, Q03 PPKU, dan sahabat-sahabat penulis yang tidak dapat disebutkan satu persatu atas segala dukungan, kebaikan, dan cerita yang telah ditorehkan.
7. Seluruh pihak yang telah membantu penulis dalam penyusunan skripsi ini yang tidak dapat dituliskan satu persatu.

Semoga skripsi ini dapat bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan, serta diberkahi oleh Allah subhanaahu wa ta'ala. Aamiin.

Bogor, Agustus 2025

Muhammad Dhiyaaul Firdaus



DAFTAR ISI

DAFTAR TABEL	x
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	x
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Manfaat	3
II TINJAUAN PUSTAKA	4
2.1 Analisis <i>Recency, Frequency, Monetary</i> (Analisis RFM)	4
2.2 Analisis Gerombol	4
2.3 <i>K-Means Clustering</i>	4
2.4 <i>K-Medoids Clustering</i>	5
2.5 <i>Davies-Bouldien Index</i>	5
2.6 Indeks <i>Silhouette</i>	6
2.7 <i>Z-score</i>	6
III METODE	8
3.1 Data	8
3.2 Prosedur Analisis Data	8
IV HASIL DAN PEMBAHASAN	10
4.1 Eksplorasi Data	10
4.2 Praproses Data	15
4.3 Penentuan Penggerombolan Terbaik	18
4.4 Analisis RFM dan Profil Demografi Anggota di Setiap Gerombol	21
4.5 Rekomendasi Strategi Bisnis	24
V SIMPULAN DAN SARAN	26
5.1 Simpulan	26
5.2 Saran	26
DAFTAR PUSTAKA	27
LAMPIRAN	29
RIWAYAT HIDUP	32



DAFTAR TABEL

1	Peubah yang digunakan dalam penelitian	8
2	Peubah setelah <i>feature engineering</i>	16
3	Statistik deskriptif peubah RFM	18
4	Statistik deskriptif peubah RFM setelah pembakuan data	18
5	Evaluasi hasil penggerombolan <i>K-Means</i> dan <i>K-Medoids</i>	19
6	Jumlah anggota dan median peubah RFM hasil penggerombolan terbaik	20
7	Evaluasi hasil penggerombolan setelah penarikan data pencilan	21
8	Karakteristik anggota di setiap gerombol berdasarkan nilai RFM	22
9	Proporsi anggota tiap gerombol berdasarkan lama anggota bergabung	22
10	Proporsi anggota tiap gerombol berdasarkan jenis kelamin	23
11	Proporsi anggota tiap gerombol berdasarkan kelompok umur	23
12	Proporsi anggota tiap gerombol berdasarkan generasi	23

DAFTAR GAMBAR

1	Proporsi anggota berdasarkan jenis kelamin	10
2	Distribusi anggota berdasarkan umur	10
3	Proporsi anggota berdasarkan generasi	11
4	Distribusi anggota berdasarkan kelompok lama keanggotaan	11
5	Diagram batang total transaksi, total pembeli, dan total nilai belanja anggota per hari	12
6	Total nilai belanja anggota per tanggal	12
7	Nilai belanja anggota per transaksi	13
8	Nilai belanja anggota per transaksi berdasarkan kondisi keuangan	14
9	Sepuluh kategori barang dengan jumlah pembelian terbanyak	15
10	Diagram kotak garis peubah RFM	17
11	Hubungan antar peubah RFM pada <i>K-Means</i> dengan $k=4$	19
12	Hubungan antar peubah RFM pada <i>K-Medoids</i> dengan $k=3$	19
13	Proporsi anggota di setiap gerombol	21
14	Diagram kotak garis peubah RFM di setiap gerombol	22

DAFTAR LAMPIRAN

1	Jumlah anggota di tiap kota/kabupaten	29
---	---------------------------------------	----