



Implementasi Metode K-Means *Clustering* Untuk Menganalisis Pola Pembelian Pelanggan Di *E-Commerce* (Studi Kasus: Pitapita Accessories)

Soleh Damarudin
Universitas Pamulang

Muhamad Santoso
Universitas Pamulang

Alamat: Jl. Raya Puspipetek No. 46, Buaran, Serpong, Kota Tangerang Selatan, Banten 15310

Korespondensi penulis: solehdamarudin4@gmail.com , dosen02593@unpam.ac.id

Abstract. *This study discusses the implementation of the K-Means Clustering method to analyze customer purchasing patterns in the PitaPita Accessories e-commerce. The main problem addressed is the absence of a structured customer segmentation, causing marketing strategies to remain generic and less targeted. This research aims to apply the K-Means algorithm to group customers based on their purchasing behavior, process transaction data into numerical attributes, and present the segmentation results as a basis for developing more personalized and effective marketing strategies. The dataset used in this study consists of 571 customer records and 571 transaction records. The analysis was conducted by constructing Recency, Frequency, and Monetary (RFM) attributes, where recency represents the time difference between the most recent transaction and a reference date, frequency represents the total number of transactions per customer, and monetary represents the total purchase value per customer. The RFM data were then clustered using the K-Means algorithm with $K = 3$ clusters. The results indicate that K-Means was successfully implemented in a web-based system and was able to generate customer segmentation into three clusters with distinct purchasing characteristics, reflected by the number of members and the average RFM values in each cluster. This segmentation can support decision-making for more targeted promotions, customer retention, and product offering strategies according to the characteristics of each customer group.*

Keywords: *K-Means Clustering, RFM, customer segmentation, e-commerce, purchasing patterns.*

Abstrak. Penelitian ini membahas implementasi metode K-Means Clustering untuk menganalisis pola pembelian pelanggan pada e-commerce PitaPita Accessories. Permasalahan yang dihadapi adalah belum tersedianya segmentasi pelanggan yang terstruktur, sehingga strategi pemasaran cenderung bersifat umum dan kurang tepat sasaran. Tujuan penelitian ini adalah menerapkan algoritma K-Means untuk mengelompokkan pelanggan berdasarkan pola pembelian, mengolah data transaksi menjadi atribut numerik, serta menyajikan hasil segmentasi sebagai dasar penyusunan

Received Januari 02, 2026; Accepted Februari 02, 2026; Published Maret 30, 2026

*Corresponding author, solehdamarudin4@gmail.com

strategi pemasaran yang lebih personal dan efektif. Data yang digunakan pada penelitian ini berjumlah 571 data pelanggan dan 571 data transaksi. Tahapan analisis dilakukan dengan membentuk atribut Recency, Frequency, dan Monetary (RFM), yaitu recency sebagai selisih waktu antara transaksi terakhir dengan tanggal acuan, frequency sebagai total jumlah transaksi per pelanggan, dan monetary sebagai total nilai pembelian pelanggan. Selanjutnya, data RFM dikelompokkan menggunakan algoritma K-Means dengan jumlah cluster (K) = 3. Hasil penelitian menunjukkan bahwa metode K-Means berhasil diimplementasikan pada sistem berbasis web dan mampu menghasilkan segmentasi pelanggan ke dalam tiga kelompok dengan karakteristik perilaku pembelian yang berbeda, ditunjukkan melalui ringkasan jumlah anggota cluster dan rata-rata nilai RFM pada setiap cluster. Segmentasi ini dapat dimanfaatkan untuk menentukan strategi promosi, retensi, dan penawaran produk yang lebih tepat sasaran sesuai karakteristik tiap kelompok pelanggan.

Kata kunci: *K-Means Clustering, RFM, segmentasi pelanggan, e-commerce, pola pembelian*

LATAR BELAKANG

E-Commerce merupakan salah satu bentuk kemajuan teknologi yang sangat terasa dalam kehidupan masyarakat saat ini. Hampir semua orang telah mengenal dan memanfaatkan *platform e-commerce* untuk berbelanja secara *online*. Perkembangan teknologi informasi yang semakin pesat, didukung oleh infrastruktur digital yang terus ditingkatkan oleh pemerintah dan pihak swasta, telah mendorong pertumbuhan perdagangan elektronik di Indonesia.

Dalam beberapa tahun terakhir, *e-commerce* tidak lagi sekadar menjadi alternatif, tetapi telah menjadi kebutuhan masyarakat. Kemudahan akses, kenyamanan berbelanja dari rumah, serta beragamnya pilihan produk membuat banyak konsumen beralih ke belanja *online*. Selain itu, kemajuan teknologi pembayaran digital dan meningkatnya penggunaan perangkat *mobile* turut mempercepat proses transaksi yang aman dan efisien.

PitaPita *Accessories* merupakan salah satu Usaha Mikro Kecil dan Menengah (UMKM) yang bergerak di bidang penjualan aksesoris wanita dan berfokus pada penjualan digital melalui platform Shopee. Perkembangan *e-commerce* memberikan peluang besar bagi UMKM seperti PitaPita untuk menjangkau pasar yang lebih luas dan bersaing dengan pelaku usaha lainnya. Namun, meskipun memiliki data transaksi pelanggan yang cukup besar, PitaPita belum memanfaatkan data tersebut secara optimal.

Saat ini, pengelolaan data pelanggan dan transaksi masih dilakukan secara konvensional. Data yang tersedia hanya digunakan sebagai catatan penjualan, seperti jumlah pesanan atau total pembayaran. Data tersebut umumnya disimpan dalam file Excel atau melalui sistem bawaan Shopee tanpa dilakukan analisis lebih lanjut. Kondisi ini membuat PitaPita sulit memahami perilaku pelanggan secara mendalam dan tidak memiliki dasar kuat dalam menyusun strategi pemasaran yang tepat sasaran.

Salah satu pendekatan yang dapat digunakan untuk mengatasi permasalahan tersebut adalah metode data mining, yaitu proses menggali informasi bernilai dari kumpulan data berukuran besar. Dalam konteks segmentasi pelanggan, algoritma K-Means Clustering merupakan salah satu metode yang efektif untuk mengelompokkan pelanggan berdasarkan kesamaan perilaku pembelian, seperti frekuensi belanja, total transaksi, dan rata-rata nilai pembelian.

Dengan penerapan K-Means, toko dapat mengetahui kelompok pelanggan aktif, pelanggan reguler, hingga pelanggan pasif. Informasi ini kemudian dapat dimanfaatkan untuk merumuskan strategi pemasaran yang lebih tepat, misalnya memberikan program loyalitas kepada pelanggan aktif atau promosi khusus untuk menarik kembali pelanggan pasif. Selain itu, hasil analisis dapat membantu efisiensi promosi, meningkatkan hubungan dengan pelanggan, dan mendukung peningkatan penjualan.

Metodologi Penelitian

Penelitian ini menggunakan dua pendekatan utama, yaitu pengembangan sistem informasi menggunakan metode *Agile* dan analisis data dengan algoritma *K-Means Clustering*. Metodologi ini dirancang untuk menghasilkan sistem yang fungsional sekaligus memberikan analisis pola pembelian pelanggan berdasarkan data transaksi.

Metode *Agile*



Gambar 1 Metode *Agile*

Menurut Trisnawati & Setiawan, 2022, *Agile* merupakan metode pengembangan perangkat lunak yang berfokus pada pengiriman produk secara bertahap melalui proses iteratif dan kolaboratif. Pendekatan ini menekankan fleksibilitas, adaptasi cepat terhadap perubahan kebutuhan, serta keterlibatan aktif pengguna pada setiap tahap pengembangan. Berbeda dengan model tradisional seperti *waterfall*, *Agile* tidak menunggu seluruh sistem selesai sepenuhnya untuk diuji. Sebaliknya, pengembangan dilakukan secara bertahap mulai dari *Minimum Viable Product* (MVP) hingga menghasilkan versi *final* yang optimal.

Setiap iterasi dalam *Agile* mencakup perencanaan, pengembangan, pengujian, dan evaluasi secara berulang. Pola ini memungkinkan tim memperoleh umpan balik secara cepat untuk melakukan perbaikan atau penyesuaian fitur di tahap berikutnya. Dengan demikian, risiko kesalahan dapat diminimalkan, sementara kualitas sistem dapat terus ditingkatkan dari waktu ke waktu.

Penerapan Algoritma K-Means Klustering

Tahap ini menjelaskan proses pengelompokan data pelanggan berdasarkan pola pembelian menggunakan metode *K-Means Clustering*. Tujuan utama dari penerapan algoritma ini adalah mengidentifikasi segmentasi pelanggan toko *online* PitaPita *Accessories* sehingga pemilik toko dapat mengetahui kelompok pelanggan berdasarkan perilaku belanja mereka. Dengan informasi tersebut, strategi pemasaran dapat disusun secara lebih terarah, seperti pemberian promosi khusus, retensi pelanggan tidak aktif, ataupun program loyalitas bagi pelanggan bernilai tinggi.

Proses klasterisasi dilakukan menggunakan data transaksi pelanggan yang diperoleh dari riwayat penjualan Shopee. Dari data mentah tersebut, dihitung tiga variabel utama, yaitu *Recency*, *Frequency*, dan *Monetary* (RFM), yang berfungsi sebagai fitur *input* dalam algoritma K-Means.

Dalam penelitian ini, beberapa studi terdahulu dijadikan sebagai referensi untuk memperkuat landasan teori serta menunjukkan posisi penelitian di antara riset-riset sebelumnya. Penelitian terdahulu yang dikaji umumnya berkaitan dengan penerapan algoritma *K-Means Clustering*, analisis pola pembelian pelanggan, serta pemanfaatan teknik data mining dalam segmentasi konsumen dan strategi pemasaran berbasis data.

Selain itu, penelitian yang ada juga memberikan gambaran mengenai penggunaan perangkat seperti *RapidMiner*, *Tanagra*, dan *Python* dalam proses klusterisasi.

PERANCANGAN

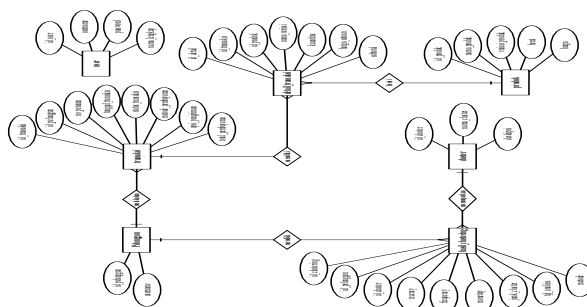
Perancangan Basis Data

Perancangan basis data merupakan tahap yang sangat penting dalam penelitian data mining, karena berfungsi sebagai fondasi penyimpanan data yang akan diolah pada tahap analisis. Pada penelitian ini, tujuan perancangan basis data adalah menyediakan struktur data yang mampu mendukung proses analisis pola pembelian pelanggan secara efisien, terintegrasi, dan mudah diolah. Mengingat fokus utama penelitian adalah penerapan metode K-Means Clustering dengan pendekatan RFM (*Recency, Frequency, Monetary*), maka rancangan basis data disusun sedemikian rupa agar dapat merekam riwayat transaksi secara historis sekaligus menyimpan informasi pelanggan secara rinci. Uraian lebih lanjut mengenai perancangannya dijelaskan pada subbab berikut.

Entity Relationship Diagram (ERD)

Entity Relationship Diagram (ERD) digunakan untuk menggambarkan hubungan antar data penting yang terdapat di dalam sistem. Dalam konteks penelitian ini, ERD dirancang untuk mempermudah pengelolaan data pelanggan, transaksi, produk, serta hasil analisis klusterisasi menggunakan pendekatan RFM (*Recency, Frequency, Monetary*). Sistem yang dibangun melibatkan enam entitas utama yang saling berelasi, yaitu: Pelanggan, Transaksi, Detail_Transaksi, Produk, Cluster, dan Hasil_Clustering.

Gambaran lengkap hubungan antar entitas tersebut dalam menganalisis pola pembelian pelanggan disajikan pada Gambar 3.3



Gambar 2 Entity Relationship Diagram (ERD) Usulan

Pada gambar di atas, entitas Pelanggan terhubung dengan entitas Transaksi melalui relasi melakukan_transaksi yang menunjukkan bahwa setiap transaksi penjualan dicatat atas

nama satu pelanggan tertentu. Dengan demikian, setiap data transaksi yang tersimpan di basis data selalu dapat ditelusuri kembali kepada pelanggan yang melakukan pembelian di toko *online PitaPita Accessories*.

Transformasi ERD ke *Logical Record Structure (LRS)*

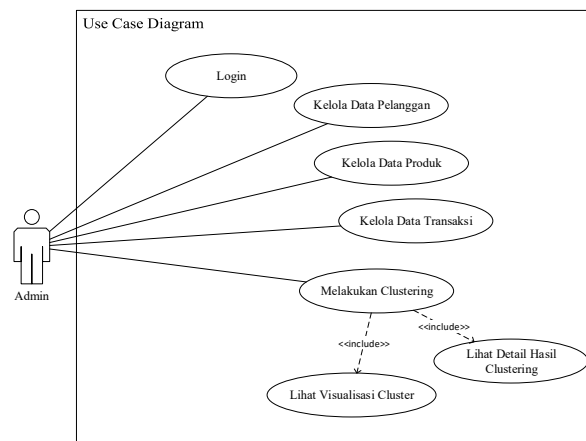
Setelah perancangan *Entity Relationship Diagram (ERD)* untuk sistem analisis pola pembelian pelanggan selesai dibuat, langkah berikutnya adalah mentransformasikan model tersebut ke dalam bentuk *Logical Record Structure (LRS)*.

Pada proses transformasi, setiap entitas yang ada pada ERD—seperti Pelanggan, Transaksi, Detail_Transaksi, Produk, Cluster, dan Hasil_Clustering—diidentifikasi atribut-atributnya, kemudian direpresentasikan sebagai tabel. Masing-masing tabel ditentukan *primary key* sebagai penanda unik dan *foreign key* sebagai penghubung antar tabel sesuai relasi yang telah didefinisikan pada ERD.

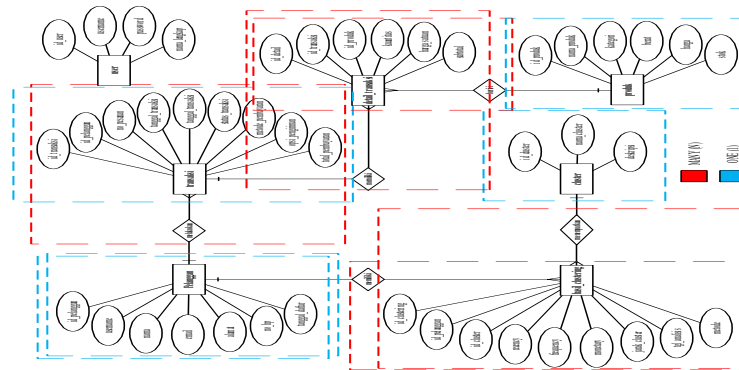
Struktur LRS yang dihasilkan inilah yang nantinya menjadi pedoman dalam pembentukan skema basis data untuk menyimpan data transaksi dan hasil klasterisasi RFM pada penelitian ini. Gambaran transformasi dari ERD ke LRS dapat dilihat pada Gambar 3.4.

Use Case Diagram

Use Case Diagram merupakan diagram yang digunakan untuk memodelkan fungsi-fungsi sistem dilihat dari sudut pandang pengguna. Diagram ini menampilkan hubungan antara aktor (pengguna) dengan *use case* atau fitur utama yang tersedia dalam sistem, sehingga memudahkan pengembang dalam memahami bagaimana pengguna berinteraksi dan memanfaatkan aplikasi yang dibangun.



Gambar 3 *Use Case Diagram*



Gambar 4 Transformasi ERD ke LRS

Pada gambar tersebut, ERD telah diubah ke dalam bentuk LRS, sehingga setiap entitas dan hubungan antar entitas direpresentasikan sebagai kumpulan tabel yang saling terhubung melalui penggunaan *primary key* dan *foreign key*.

IMPLEMENTASI DAN PENGUJIAN

Spesifikasi

Pada tahap spesifikasi, user model digunakan sebagai acuan awal dalam merancang struktur halaman maupun tampilan antarmuka sistem. Model ini memberikan gambaran sederhana mengenai fungsi utama serta alur navigasi, tanpa menekankan aspek desain visual. Tujuan utamanya adalah mempermudah proses perencanaan antarmuka sebelum masuk ke tahap pengembangan lebih lanjut.

Implementasi Program

Aplikasi ini dikembangkan menggunakan dengan bantuan *Bootstrap* untuk membangun tampilan web yang modern, responsif, serta mudah digunakan. Berikut disajikan uraian hasil implementasi beserta fungsinya:

- a. Halaman *Login*



Gambar 5 Halaman *Login*

Halaman *Login* ini merupakan antarmuka autentikasi untuk membatasi akses ke sistem, dimana pengguna memasukkan *username* dan *password* pada *form* yang telah disediakan, kemudian menekan tombol *Login* untuk melakukan validasi data. Tampilan dibuat sederhana dan terpusat dengan identitas aplikasi Pitapita *Accessories* di sisi kiri, sehingga pengguna dapat mengenali sistem dan melakukan proses masuk dengan cepat sebelum mengakses menu admin.

b. Halaman *Dashboard* Admin



Gambar 6 Halaman *Dashboard* Admin

Antarmuka *Dashboard* Admin ini berfungsi sebagai halaman utama untuk memantau kondisi sistem secara ringkas, dengan *sidebar* navigasi di sisi kiri yang memuat menu *Dashboard*, Pelanggan, Produk, Transaksi, Clustering, *User*, dan *Logout*. Pada bagian atas ditampilkan informasi akun yang sedang *login*, kemudian tersedia kartu ringkasan Total Pelanggan, Total Produk, Total Transaksi, dan Omzet Total sebagai indikator cepat. *Dashboard* juga menampilkan grafik tren omzet 14 hari terakhir untuk melihat pergerakan pendapatan, panel Top 5 Produk berdasarkan omzet untuk mengetahui produk dengan performa terbaik, serta tabel Transaksi Terbaru yang menampilkan ringkasan pesanan terkini dan tombol Lihat Semua untuk menuju daftar transaksi lengkap.

c. Halaman Data Pelanggan



Gambar 7 Halaman Data Pelanggan

Halaman Data Pelanggan ini digunakan untuk mengelola daftar pelanggan yang tersimpan dalam sistem, ditampilkan dalam bentuk tabel berisi kolom No, *Username*, dan Aksi. Admin dapat menambahkan pelanggan melalui tombol Tambah Pelanggan, mencari data dengan fitur *Search*, mengatur jumlah tampilan data per halaman, serta melakukan pengelolaan data melalui tombol Edit untuk memperbarui informasi dan Hapus untuk menghapus data pelanggan, lengkap dengan navigasi *pagination* di bagian bawah untuk berpindah halaman data.

d. Halaman Tambah Pelanggan



Gambar 8 Halaman Tambah Pelanggan

Antarmuka ini menampilkan *form* Tambah Pelanggan dalam bentuk *modal/popup* yang digunakan admin untuk menambahkan data pelanggan baru ke sistem. *Modal* menyediakan satu kolom *input Username* sebagai data utama yang harus diisi, serta tombol *Close* untuk membatalkan proses dan tombol Simpan untuk menyimpan data, sementara halaman tabel pelanggan di belakangnya tetap terlihat sebagai konteks data yang sedang dikelola.

e. Halaman Produk



Gambar 9 Halaman Produk

Halaman Data Produk ini digunakan untuk mengelola informasi produk yang tersedia pada sistem, ditampilkan dalam bentuk tabel berisi kolom No, Nama Produk, Variasi Produk, Berat (gram), Harga, dan Aksi. Admin dapat menambahkan produk baru melalui tombol Tambah Produk, melakukan pencarian dengan fitur *Search*, mengatur jumlah data per tampilan, serta melakukan pengelolaan data melalui tombol Edit untuk memperbarui informasi dan Hapus untuk menghapus produk, lengkap dengan fitur *pagination* di bagian bawah untuk menavigasi data produk yang banyak.

f. Halaman Tambah Produk



Gambar 10 Halaman Tambah Produk

Antarmuka ini menampilkan *form* Tambah Produk dalam bentuk modal/*popup* yang digunakan admin untuk menambahkan data produk baru ke sistem. Di dalam modal tersedia kolom input Nama Produk, Variasi Produk, Berat (gram), dan Harga sebagai informasi utama yang harus diisi, serta tombol *Close* untuk membatalkan proses dan tombol Simpan untuk menyimpan data, sementara halaman tabel produk tetap terlihat di latar belakang sebagai konteks pengelolaan data.

g. Halaman Transaksi



Gambar 11 Halaman Transaksi

Halaman Data Transaksi ini digunakan untuk menampilkan dan mengelola daftar transaksi pelanggan dalam bentuk tabel, yang memuat informasi No, No Pesanan, Pelanggan, Tanggal Transaksi, Status, Metode Pembayaran, Pengiriman, dan Total Pembayaran. Admin dapat menambahkan transaksi baru melalui tombol Tambah Transaksi, melakukan pencarian dengan fitur *Search*, mengatur jumlah data yang ditampilkan per halaman, serta menggunakan tombol *Detail* untuk melihat rincian transaksi dan tombol Edit untuk memperbarui data transaksi, dilengkapi navigasi *pagination* di bagian bawah untuk memudahkan penelusuran data.

h. Halaman Tambah Transaksi



Gambar 12 Halaman Tambah Transaksi

Antarmuka ini menampilkan *form* Tambah Transaksi dalam bentuk modal/*popup* yang digunakan admin untuk memasukkan data transaksi baru ke sistem. Di dalamnya tersedia input Pelanggan (*dropdown*), No Pesanan, Tanggal Transaksi (*date picker*), Status Transaksi, Metode Pembayaran, Opsi Pengiriman, serta Total Pembayaran (*Final Shopee*) sebagai data utama transaksi, kemudian dilengkapi tombol untuk menutup atau menyimpan agar data yang diinput dapat tercatat dan ditampilkan pada tabel transaksi.

i. Halaman Edit Transaksi



Gambar 13 Halaman Edit Transaksi

Pada gambar ditampilkan halaman Detail Barang Masuk yang menampilkan rincian transaksi pemasukan barang. Informasi mencakup nomor pemasukan, nama staf, waktu, serta daftar barang dengan harga, jumlah, dan subtotal.

j. Halaman Hapus Transaksi



Gambar 14 Halaman Hapus Transaksi

Antarmuka ini menampilkan konfirmasi Hapus Transaksi dalam bentuk modal/*popup* untuk memastikan admin tidak menghapus data transaksi secara tidak sengaja. Modal menampilkan pertanyaan konfirmasi, nomor pesanan transaksi yang akan dihapus, serta informasi bahwa *detail* transaksi terkait juga akan terhapus, kemudian menyediakan tombol Batal untuk membatalkan proses dan tombol Hapus untuk mengeksekusi penghapusan data pada sistem.

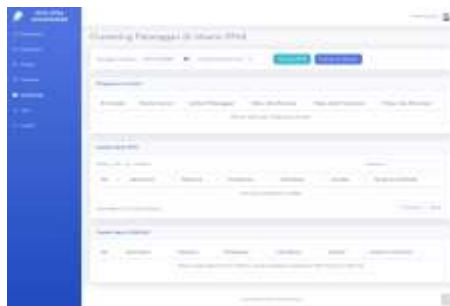
k. Halaman Detail Transaksi



Gambar 15 Halaman Detail Transaksi

Halaman Detail Transaksi ini digunakan untuk menampilkan rincian transaksi berdasarkan nomor pesanan, pelanggan, dan tanggal transaksi, serta menampilkan ringkasan Total Pembayaran (*Final Shopee*) dan Total Versi Detail (*cek internal*) sebagai pembandingan perhitungan. Di bawahnya tersedia tabel Item Produk yang memuat daftar produk dalam transaksi beserta variasi, kuantitas, harga satuan, subtotal, dan tombol aksi Edit/Hapus, dilengkapi fitur pencarian, pengaturan jumlah data, serta tombol Kembali dan Tambah Detail untuk navigasi dan penambahan item transaksi.

1. Halaman Clustering

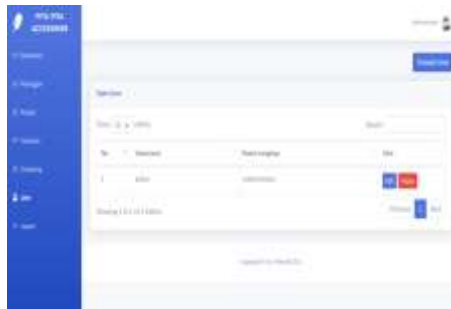


Gambar 16 Halaman Clustering

Halaman Clustering Pelanggan (K-Means RFM) ini digunakan untuk menjalankan analisis segmentasi pelanggan berbasis nilai RFM dengan metode K-Means, dimana admin dapat menentukan Tanggal Analisis dan Jumlah Cluster (K), lalu menekan tombol Hitung RFM untuk menghasilkan nilai *Recency*, *Frequency*, dan *Monetary* serta tombol Proses K-Means untuk melakukan pengelompokan. Hasil pengolahan ditampilkan pada tiga bagian, yaitu tabel Ringkasan Cluster yang memuat rekap tiap cluster (jumlah pelanggan serta rata-rata R, F, M), tabel Detail Hasil RFM yang menampilkan nilai RFM per pelanggan, dan tabel Detail Hasil KMEANS yang menampilkan hasil akhir *cluster*

beserta jarak ke *centroid*, lengkap dengan fitur pencarian, *pagination*, dan informasi status ketika data belum diproses.

m. Halaman *User*



Gambar 17 Halaman *User*

Halaman Data *User* ini digunakan untuk mengelola akun pengguna yang dapat mengakses sistem, ditampilkan dalam tabel berisi kolom No, *Username*, Nama Lengkap, dan Aksi. Admin dapat menambahkan akun baru melalui tombol Tambah User, melakukan pencarian dengan fitur *Search*, mengatur jumlah data yang ditampilkan, serta menggunakan tombol Edit untuk memperbarui data *user* dan Hapus untuk menghapus akun, dengan navigasi *pagination* di bagian bawah untuk memudahkan penelusuran data.

n. Halaman Tambah *User*



Gambar 18 Halaman Tambah *User*

Antarmuka ini menampilkan *form* Tambah User dalam bentuk modal/*popup* yang digunakan admin untuk membuat akun baru. Modal menyediakan input *Username*, Nama Lengkap, *Password* (minimal 6 karakter), serta Konfirmasi *Password* untuk memastikan kesesuaian kata sandi, kemudian dilengkapi tombol Tutup untuk membatalkan dan Simpan untuk menyimpan data user ke sistem, sementara halaman tabel user tetap terlihat di latar belakang sebagai konteks pengelolaan akun.

Pengujian Sistem

Tahap pengujian sistem merupakan bagian penting untuk memastikan seluruh fitur aplikasi berjalan sesuai kebutuhan sebelum digunakan secara menyeluruh. Pengujian dilakukan menggunakan dua pendekatan utama, yaitu *Black Box Testing* yang menilai fungsi aplikasi dari sudut pandang pengguna tanpa meninjau kode program, serta *White Box Testing* yang memeriksa struktur internal, alur proses, dan logika program guna memastikan efisiensi, keamanan, dan ketepatan implementasi. Melalui penerapan kedua metode tersebut, sistem diharapkan memiliki kinerja yang stabil, aman, dan tersusun secara baik sebelum diterapkan.

Hasil Pengujian Algoritma K-Means Menggunakan RapidMiner

Pengujian algoritma K-Means dilakukan menggunakan aplikasi RapidMiner/Altair AI Studio untuk memastikan proses pengelompokan pelanggan dapat berjalan sesuai tahapan analisis yang dirancang. Data yang digunakan merupakan dataset pelanggan yang telah diolah menjadi nilai Recency, Frequency, dan Monetary (RFM), sehingga proses clustering dapat dilakukan dengan memanfaatkan atribut numerik yang mewakili perilaku pembelian pelanggan. Tahapan pengujian terdiri dari proses input dataset, penyesuaian atribut, normalisasi data, hingga pembentukan cluster serta penampilan hasil model cluster.

Persiapan Dataset

Pada tahap ini, dataset yang digunakan untuk pengujian merupakan dataset RFM per pelanggan yang berisi atribut identitas pelanggan dan nilai RFM. Penggunaan dataset RFM diperlukan karena algoritma K-Means bekerja dengan menghitung jarak antar data sehingga membutuhkan atribut yang bersifat numerik. Dataset yang digunakan berjumlah 567 data pelanggan, dimana setiap baris merepresentasikan satu pelanggan dengan nilai *Recency*, *Frequency*, dan *Monetary* yang telah dihitung dari data transaksi toko online PitaPita.

Select the cells to import.

	A	B	C	D	E	F
	id_pelanggan	nama	umur	pendapatan	usia_keluarga	last_purchase
1	1.000	Adi	25	1.000	0.000	May 1, 2011 11:10
2	2.000	Budi	30	1.000	0.000	May 1, 2011 11:10
3	3.000	Citra	20	1.000	0.000	May 1, 2011 11:10
4	4.000	Dani	25	1.000	0.000	May 1, 2011 11:10
5	5.000	Eva	25	1.000	0.000	May 1, 2011 11:10
6	6.000	Fani	25	1.000	0.000	May 1, 2011 11:10
7	7.000	Gina	25	1.000	0.000	May 1, 2011 11:10
8	8.000	Hani	25	1.000	0.000	May 1, 2011 11:10
9	9.000	Ibu	25	1.000	0.000	May 1, 2011 11:10
10	10.000	Joni	25	1.000	0.000	May 1, 2011 11:10
11	11.000	Kiki	25	1.000	0.000	May 1, 2011 11:10
12	12.000	Lili	25	1.000	0.000	May 1, 2011 11:10
13	13.000	Mami	25	1.000	0.000	May 1, 2011 11:10
14	14.000	Nani	25	1.000	0.000	May 1, 2011 11:10

Gambar 19 Persiapan Dataset

Tahap Clustering dengan K-Means

Tahap utama pengujian dilakukan menggunakan *operator Clustering* (K-Means). Parameter yang digunakan pada pengujian ini yaitu jumlah cluster (k) sebesar 3, max runs sebesar 10, dan opsi *determine good start values* diaktifkan agar centroid awal yang dipilih lebih optimal. Pengukuran jarak yang digunakan adalah Euclidean Distance karena sesuai untuk data numerik hasil normalisasi. Selain itu, opsi *add cluster attribute* diaktifkan agar hasil cluster dapat ditambahkan langsung ke dataset, sehingga setiap pelanggan memiliki label cluster yang dapat dianalisis lebih lanjut.

Parameters

☒ Clustering (k-Means)

☒ add cluster attribute

☐ add as label

☐ remove unlabeled

k: 3

max runs: 10

☒ determine good start values

measure type: Numerical Measure

numerical measure: Euclidean Distance

max optimization steps: 100

[Show advanced parameters](#)

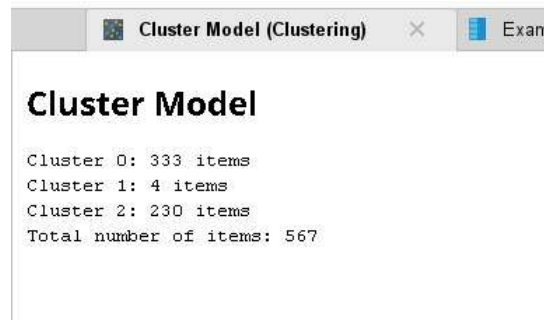
[Change command line](#)

Gambar 20 Tahap Clustering dengan K-Means

Hasil Clustering

Berdasarkan hasil pengujian K-Means di RapidMiner, sistem membentuk 3 kelompok pelanggan dari total 567 data pelanggan. Distribusi anggota cluster yang

terbentuk adalah Cluster 0 sebanyak 333 pelanggan, Cluster 1 sebanyak 4 pelanggan, dan Cluster 2 sebanyak 230 pelanggan. Perbedaan jumlah anggota cluster menunjukkan adanya variasi perilaku pembelian pelanggan, dimana sebagian besar pelanggan terkonsentrasi pada dua cluster besar dan terdapat satu cluster kecil yang kemungkinan berisi pelanggan dengan karakteristik pembelian yang sangat berbeda dibanding kelompok lainnya.



Gambar 21 Hasil Clustering

KESIMPULAN

Berdasarkan hasil perancangan, implementasi, dan pengujian sistem pada penelitian ini, dapat ditarik kesimpulan sebagai berikut:

- Algoritma K-Means Clustering berhasil diterapkan pada sistem berbasis web untuk menganalisis pola pembelian pelanggan toko *online* PitaPita. Proses analisis dilakukan melalui pengolahan data transaksi menjadi nilai *Recency*, *Frequency*, dan *Monetary* (RFM), kemudian dikelompokkan menggunakan parameter jumlah cluster (K) sesuai kebutuhan analisis.
- Hasil proses clustering dapat digunakan untuk melakukan segmentasi pelanggan berdasarkan perilaku pembelian, karena sistem mampu menghasilkan pengelompokan pelanggan ke dalam beberapa cluster yang memiliki karakteristik berbeda, ditunjukkan melalui ringkasan jumlah anggota cluster dan rata-rata nilai RFM pada setiap cluster.
- Hasil segmentasi pelanggan dapat dimanfaatkan sebagai dasar penyusunan strategi pemasaran yang lebih efektif, karena setiap cluster memberikan informasi tentang jenis pelanggan (misalnya pelanggan dengan frekuensi tinggi, pelanggan bernilai belanja tinggi, atau pelanggan yang sudah lama tidak bertransaksi), sehingga toko

dapat menentukan pendekatan promosi, retensi, dan penawaran produk yang lebih tepat sasaran.

DAFTAR PUSTAKA

- Abdurahman, F. R., Hurriyati, R., & Dirgantari, P. D. (2024). *E-Customer Behavior Mediates Promotion on E-Commerce Purchase Decision MSME Industry*. 2(01), 268–275.
- Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics (Switzerland)*, 9(8), 1–12. <https://doi.org/10.3390/electronics9081295>
- Amin, F., Anggraeni, D. S., & Aini, Q. (2022). Penerapan Metode K-Means dalam Penjualan Produk Souq.Com. *Applied Information System and Management (AISM)*, 5(1), 7–14. <https://doi.org/10.15408/aism.v5i1.22534>
- Andriyadi, A., Fikri, R. R. N., & Saputri, E. F. (2022). Evaluasi Sistem Informasi Perpustakaan Institut Informatika Darmajaya Dengan Whitebox Testing. *Journal of Innovation ...*, 3471(8), 743–746.
- Anggun Fergina, M. K., Syahid Abdullah, S. S. M. K., & Zaenal Alamsyah, S. K. (2021). *Buku Ajar Basis Data*. Nusa Putra Press.
- Anwar, M. A. H., & Kurniawan, Y. (2019). Dokumentasi Software Testing Berstandar Ieee 829-2008 Untuk Sistem Informasi Terintegrasi Universitas. *Kurawal - Jurnal Teknologi, Informasi Dan Industri*, 2(2), 118–125. <https://doi.org/10.33479/kurawal.v2i2.261>
- Ceder, N. (2025). *The Quick Python Book, Fourth Edition*. Manning.
- Darni, R., Anwar, M., & Hadi, A. (n.d.). *KARIR DAN DASAR-DASAR PENGEMBANGAN WEB*. UNP PRESS.
- Destriana, R., Husain, S. M., Handayani, N., & Siswanto, A. T. P. (2022). *Diagram UML Dalam Membuat Aplikasi Android Firebase “Studi Kasus Aplikasi Bank Sampah.”* Deepublish.
- Dina, D. F. M. (2024). Normalisasi Database Rancangan Sistem Penyewaan Buku Berbayar. *Computing Insight: Journal of Computer Science*, 4(1), 56–61. https://doi.org/10.30651/comp_insight.v4i1.15814
- Dr. Enjang Yusup Ali, M. K., Prof. Dr. Munir, M. I. T., A, P. D. J. P. M., Dr. Dedy Achmad Kurniady, M. P., Dr. H. Diding Nurdin, M. P., & Dr. Rahmat Fadhli, E. M. (2024). *Model Enterprise Architecture (EA) Untuk Perguruan Tinggi Bermutu*. Indonesia Emas Group.
- Dwi Krisbiantoro, M. K., & Prih Diantono Abda’u, M. K. (2021). *DASAR PEMROGRAMAN WEB dengan bahasa HTML, PHP, dan Database MySQL*. Zahira Media Publisher.
- Enterprise, J. (2023). *HTML, PHP, dan MySQL untuk Pemula (Update Version)*. Elex Media Komputindo.
- Fauzan, M. N., & Nurhidayah, S. (2020). *Membuat Sistem Approval Anggaran Pelatihan Dengan Php, Codeigniter, Dan Bootstrap*. Kreatif.
- Gillenson, M. L. (2023). *Fundamentals of Database Management Systems*. Wiley.
- Han, J., Pei, J., & Tong, H. (2022). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.
- Kurniawan, H., Apriliah, W., Kurnia, I., & Firmansyah, D. (2021). Penerapan Metode Waterfall Dalam Perancangan Sistem Informasi Penggajian Pada Smk Bina Karya

- Karawang. *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, 14(4), 13–23. <https://doi.org/10.35969/interkom.v14i4.78>
- Miftakul Amin, S. K. M. E. (2022). *Bahasa Query Menggunakan MySQL*. Penerbit Ahatek.
- Mujiarto, W., & Adiguna, M. A. (2023). Perancangan Sistem Informasi Inventory Barang Stroomnet Berbasis Website Menggunakan Metode Rapid Application Development. *Scientia Sacra: Jurnal Sains*, 3(1), 26–43.
- Namruddin, R., Basalamah, A., Ali, M. Z. A., Syarifuddin, A., Alam, S., Wardhani, N., & Abdurrahman, T. S. D. (2023). *BELAJAR DATABASE DENGAN MUDAH MENGGUNAKAN MYSQL*. TOHAR MEDIA.
- Neighbors, P. K. D. K., Ramadhani, R. A., & Wahyuniar, L. S. (n.d.). *Implementasi Data Mining Untuk Menentukan Pelanggan Potensial Menggunakan Algoritma*. 9, 1173–1182.
- No, V., Artiarno, A. M., Setiaji, P., & Nugraha, F. (2025). *Edumatic : Jurnal Pendidikan Informatika K-Means Clustering untuk Segmentasi Pelanggan : Mengungkap Pola Pembelian Strategi Pemasaran pada Sektor Ritel*. 9(2), 442–451. <https://doi.org/10.29408/edumatic.v9i2.30336>
- Novi, Y., Cuhwanto, A., & R, D. A. (2022). *Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means*. 15(1), 48–56.
- Nur, S. A., & S.R. Candra, N. (2022). Pengujian Kesetaraan Partisi pada White Box (Studi Kasus : Electronic City). *Journal of Informatics and Advanced Computing (JIAC)*, 3(2), 97–102.
- Nurshadrina, N., & Voutama, A. (2022). Penerapan Unified Modeling Language (UML) dalam membangun sistem pengenalan UMKM (Studi Kasus: Rafa Laundry). *Information Management for Educators and Professionals: Journal of Information Management*, 7(1), 21–30.
- Oracle Corporation. (2025). *MySQL 8.0 Reference Manual*.
- Perkasa, I. B., & Komputer, I. (2024). *STRATEGI DATA MINING UNTUK IDENTIFIKASI POLA*. 1(6), 1–16.
- Permana, Y., & Romadlon, P. (2019). PERANCANGAN SISTEM INFORMASI PENJUALAN PERUMAHAN MENGGUNAKAN METODE SDLC PADA PT. MANDIRI LAND PROSPEROUS BERBASIS MOBILE. *SIGMA – Jurnal Teknologi Pelita Bangsa*, 84(10), 1511–1518. <https://doi.org/10.1134/s0320972519100129>
- Pratiwi, B. (2024). *PELANGGAN*. 1(5), 1–18.
- Putra, A. P., Andriyanto, F., Karisman, K., Harti, T. D. M., & Sari, W. P. (2020). Pengujian Aplikasi Point of Sale Menggunakan Blackbox Testing. *Jurnal Bina Komputer*, 2(1), 74–78. <https://doi.org/10.33557/binakomputer.v2i1.757>
- Rizky Pangestu, P., & Voutama, A. (2024). Pemanfaatan Uml (Unified Modelling Language) Pada Sistem Pengelolaan Aspirasi Mahasiswa Berbasis Website. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(6), 11846–11851. <https://doi.org/10.36040/jati.v8i6.11732>
- Sutedi, S., Widyawati, D. K., Farkhan, M., Akhsa, A. T. P. D., Sesunan, M. F., Purwandari, N., Mukminna, H., Ariana, A. A. G. B., Efitra, E., & Sari, I. K. (2024). *Buku Ajar Sistem Basis Data*. PT. Sonpedia Publishing Indonesia.
- Tabianan, K., & Velu, S. (2022). *K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data*. 1–15.
- Trisnawati, L., & Setiawan, D. (2022). Sistem Monitoring Kegiatan Kemahasiswaan

- Menggunakan Metode Agile Development. *JOISIE Journal Of Information System And Informatics Engineering*, 6(1), 49–57.
- Uminingsih, Nur Ichsanudin, M., Yusuf, M., & Suraya, S. (2022). Pengujian Fungsional Perangkat Lunak Sistem Informasi Perpustakaan Dengan Metode Black Box Testing Bagi Pemula. *STORAGE: Jurnal Ilmiah Teknik Dan Ilmu Komputer*, 1(2), 1–8. <https://doi.org/10.55123/storage.v1i2.270>
- Uska, M., Wirasasmita, R., Usuluddin, U., & Arianti, B. (2020). Evaluation of Rapidminer-Aplication in Data Mining Learning using PeRSIVA Model. *Edumatic: Jurnal Pendidikan Informatika*, 4(2), 164–171. <https://doi.org/10.29408/edumatic.v4i2.2688>
- Widia, D. M., & Asriningtias, S. R. (2021). *Cara Cepat dan Praktis Membangun Web Dinamis dengan PHP dan MySQL*. Universitas Brawijaya Press.
- Widyatmoko, W., & Pamungkas, N. (2022). Pemodelan Unified Modeling Language pada Sistem Aplikasi Pariwisata (SiAP). *Jurnal Bumigora Information Technology (BITE)*, 4(1), 73–84. <https://doi.org/10.30812/bite.v4i1.1871>
- Wu, Z., Jin, L., Zhao, J., Jing, L., & Chen, L. (2022). *Research on Segmenting E-Commerce Customer through an Improved K-Medoids Clustering Algorithm*. 2022.