Jurnal Manajemen dan Teknologi
https://journal.stmiki.ac.id

E-ISSN: 3047-8251 | P-ISSN: 3047-826X
Vol. 3 No. 2 (2026) | November

OPEN ACCESS

https://doi.org/ 10.63447/jmt.v3i2.2128

Klasterisasi Karakteristik Pengguna E-Commerce Berdasarkan Profil Pengguna Menggunakan Algoritma K-Prototypes

Gala Di Akbar [1], Mira Afrina [2*], Purwita Sari [3], Ken Dhita Tania [4], Rizka Dhini Kurnia [5], Lucky Indra Kesuma [6]

[1,2*,3,4,5,6] Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Sriwijaya, Kota Palembang, Provinsi Sumatera Selatan, Indonesia.

Corresponding Email: mira_afrina@unsri.ac.id [2*]

Histori Artikel:
Dikirim 17 September 2026; *Diterima dalam bentuk revisi* 10 Oktober 2026; *Diterima* 25 Oktober 2026; *Diterbitkan* 25 November 2026. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

Abstrak

Segmentasi pelanggan yang akurat diperlukan untuk mendukung pemasaran e-commerce yang lebih personal. Namun, model Recency, Frequency, Monetary (RFM) berfokus pada riwayat transaksi sehingga belum sepenuhnya menggambarkan motivasi belanja dan proporsi pengeluaran konsumen untuk e-commerce (Share of Wallet/SoW). Penelitian ini bertujuan mengelompokkan pengguna e-commerce berdasarkan profil transaksi dan perilaku konsumen menggunakan algoritma K-Prototypes. Data diperoleh dari 198 pengguna e-commerce di Kota Palembang. Variabel klasterisasi meliputi RFM, SoW, motivasi belanja, dan preferensi kategori produk, sedangkan kategori usia digunakan dalam analisis lanjutan. Penentuan jumlah klaster menggunakan Elbow Method dan Silhouette Score menghasilkan empat persona, yaitu High-Value Promo Hunters, Routine Pragmatic Shoppers, Passive Micro-Spenders, dan Urgent Big-Ticket Buyers. Uji Chi-Square menunjukkan adanya asosiasi yang signifikan antara kategori usia dan persona, dengan nilai Cramér's V sebesar 0,274. Hasil ini menunjukkan bahwa usia berkaitan dengan pola belanja, tetapi tidak sepenuhnya menentukan profil konsumen. Temuan ini dapat membantu platform menyusun promosi dan pengalaman pengguna yang sesuai dengan karakteristik pelanggan. Penelitian ini terbatas pada 198 responden di Kota Palembang, sehingga hasilnya belum dapat digeneralisasi ke seluruh pengguna e-commerce.

Kata Kunci: E-Commerce; K-Prototypes; Share of Wallet; Kohort Generasi.

Abstract

Accurate customer segmentation supports more personalized e-commerce marketing. However, the Recency, Frequency, Monetary (RFM) model focuses on transaction history and may not fully capture shopping motivation or the share of a consumer's spending allocated to e-commerce (Share of Wallet/SoW). This study aims to cluster e-commerce users based on transaction profiles and consumer behavior using the K-Prototypes algorithm. Data were collected from 198 e-commerce users in Palembang, Indonesia. The clustering variables included RFM, SoW, shopping motivation, and product-category preferences, while age category was used for follow-up analysis. The Elbow Method and Silhouette Score supported a four-cluster solution. The resulting personas were High-Value Promo Hunters, Routine Pragmatic Shoppers, Passive Micro-Spenders, and Urgent Big-Ticket Buyers. The Chi-square test indicated a significant association between age category and persona, with Cramér's V of 0.274. This suggests that age is associated with shopping patterns but does not fully determine consumer profiles. The findings can inform more targeted promotions and user-experience design. The study is limited to 198 respondents from Palembang, so the findings may not represent e-commerce users in other regions.

Keyword: E-Commerce; K-Prototypes; Share of Wallet; Generational Cohort.

Jurnal Manajemen dan Teknologi
https://journal.stmiki.ac.id

E-ISSN: 3047-8251 | P-ISSN: 3047-826X
Vol. 3 No. 2 (2026) | November

OPEN ACCESS

https://doi.org/ 10.63447/jmt.v3i2.2128

1. Pendahuluan

Pertumbuhan e-commerce mendorong perusahaan beralih dari pemasaran massal menuju penargetan pelanggan yang lebih personal. Segmentasi berbasis *machine learning* dapat membantu mengenali pola belanja dan mendukung pengambilan keputusan pemasaran digital (Alie & Gustriansyah, 2024; Alves Gomes & Meisen, 2023). Dalam kerangka *Customer Relationship Management* (CRM), segmentasi tidak cukup hanya menggambarkan riwayat transaksi; platform juga perlu memahami motivasi dan alokasi dana pelanggan agar promosi serta layanan lebih relevan. Penerapan segmentasi berbasis komputasi juga dilaporkan dapat mendukung efisiensi operasional dan keberlanjutan profitabilitas platform (Kumar, 2025). Model *Recency, Frequency, Monetary* (RFM) banyak digunakan untuk merangkum waktu transaksi terakhir, frekuensi pembelian, dan nilai belanja pelanggan (Homepage *et al.*, 2025). Penelitian terdahulu telah mengembangkan RFM dengan menambahkan Volume melalui model LRFMV (Mahfuza *et al.*, 2022), menggabungkannya dengan matriks BCG untuk mengevaluasi portofolio pelanggan (Gholamveisy *et al.*, 2024), serta memasukkan informasi demografis ke dalam model RFMD (Ho *et al.*, 2023). Meski bermanfaat, indikator transaksi belum sepenuhnya menjelaskan motivasi belanja dan proporsi pengeluaran pelanggan untuk e-commerce atau *Share of Wallet* (SoW) (Blömker & Albrecht, 2024). Akibatnya, pelanggan dengan nilai transaksi yang sama dapat memiliki tingkat komitmen finansial yang berbeda. Penelitian segmentasi e-commerce menggunakan model RFM yang diperluas dan K-Medoids juga menunjukkan pentingnya pemilihan fitur serta algoritma yang sesuai dengan karakteristik data (Wu *et al.*, 2022), sedangkan K-Means telah diterapkan pada segmentasi pelanggan berdasarkan usia, pendapatan, dan RFM (Kristanti *et al.*, 2024).

Tinjauan (Alves Gomes & Meisen, 2023) turut menekankan pentingnya memahami kebutuhan dan motivasi pelanggan untuk mendukung penargetan yang lebih personal. Pemilihan algoritma juga perlu mempertimbangkan tipe data yang dianalisis. K-Means menggunakan jarak *Euclidean* dan tidak secara langsung mengolah atribut kategorikal; mengubah kategori seperti motivasi belanja menjadi angka dapat mengaburkan perbedaan makna antarkategori (Mahesa Putra *et al.*, 2025). K-Prototypes dirancang untuk mengolah atribut numerik dan kategorikal secara bersamaan melalui kombinasi jarak *Euclidean* dan ukuran ketidakmiripan kategori (Shi *et al.*, 2024). Penerapannya pada data pelanggan campuran telah diteliti dalam beberapa situasi. Beberapa referensi menggunakan K-Prototypes sebagai algoritma pembanding untuk membentuk segmen pelanggan pada data sebuah toko virtual dan menemukan hasil pengelompokan yang secara umum sebanding dengan *Hierarchical Agglomerative Clustering*. Dalam *e-commerce*, (Ho *et al.*, 2023) menerapkan K-Prototypes pada data demografis, transaksi, dan perilaku pelanggan sebagai bagian dari kerangka analisis prediksi *customer churn*. Studi-studi tersebut menunjukkan kegunaan K-Prototypes pada data pelanggan bertipe campuran, tetapi berfokus pada pembandingan metode atau identifikasi risiko pelanggan berhenti menggunakan layanan. Penelitian ini memperluas fokus dengan menggabungkan RFM, SoW, motivasi belanja, dan preferensi kategori produk untuk membentuk persona perilaku konsumen. Penelitian ini mengelompokkan seluruh responden secara serentak (*global clustering*), bukan membentuk klaster terpisah berdasarkan usia terlebih dahulu (*local clustering*). Kelompok pelanggan dibentuk menggunakan kriteria transaksi dan perilaku yang sama, sedangkan kategori usia dianalisis setelah proses klasterisasi selesai. Berdasarkan pendekatan tersebut, penelitian ini bertujuan mengidentifikasi persona pengguna e-commerce di Kota Palembang berdasarkan RFM, SoW, motivasi belanja, dan preferensi kategori produk, lalu menguji asosiasi antara kategori usia dan persona yang terbentuk (Vrhovac *et al.*, 2024). Hasilnya diharapkan memberikan gambaran segmentasi yang lebih realistis daripada asumsi yang hanya mengandalkan generasi, serta membantu platform merancang promosi dan pengalaman pengguna berdasarkan pola belanja pelanggan (Alves Gomes & Meisen, 2023).

Jurnal Manajemen dan Teknologi
https://journal.stmiki.ac.id

E-ISSN: 3047-8251 | P-ISSN: 3047-826X
Vol. 3 No. 2 (2026) | November

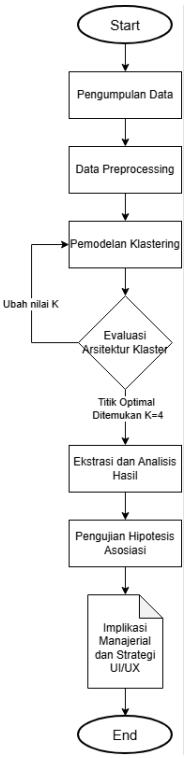
OPEN ACCESS

https://doi.org/ 10.63447/jmt.v3i2.2128

2. Metode Penelitian

2.1 Tahapan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan teknik *unsupervised machine learning* untuk mengelompokkan pengguna e-commerce berdasarkan data numerik dan kategorikal. Seperti ditunjukkan pada Gambar 1, tahapan penelitian meliputi pengumpulan data, pra-pemrosesan, pembentukan klaster menggunakan K-Prototypes, evaluasi jumlah klaster melalui *Elbow Method* dan *Silhouette Score*, serta ekstraksi dan interpretasi profil konsumen. Jika evaluasi belum menunjukkan pembagian yang memadai, nilai K disesuaikan dan proses klasterisasi diulang. Setelah jumlah klaster ditetapkan, analisis dilanjutkan dengan pengujian asosiasi antara kategori usia dan persona yang terbentuk. Tahap akhir adalah merumuskan implikasi manajerial bagi strategi pemasaran dan desain antarmuka pengguna. Berdasarkan evaluasi yang dilakukan, penelitian ini menggunakan empat klaster (K = 4).



Gambar 1. Flowchart Metodologi Penelitian

2.2 Pengumpulan Data dan Representasi Data

Data primer dikumpulkan melalui kuesioner daring menggunakan *Google Forms*. Tautan kuesioner disebarkan melalui media sosial, terutama *Instagram* dan *WhatsApp*, pada kanal atau komunitas yang secara khusus menjangkau wilayah Kota Palembang. Pengumpulan data mulai tercatat pada 8 Mei 2026, yaitu tanggal respons pertama. Formulir ditunggu selama satu bulan sejak pengisian pertama. Berdasarkan catatan waktu pengisian yang tersedia, respons terakhir tercatat pada 12 Mei 2026. Responden mengisi kuesioner melalui tautan *Google Forms*, dan jawaban dihimpun secara elektronik untuk selanjutnya diperiksa berdasarkan kriteria penelitian. Penelitian ini melibatkan 198 pengguna e-commerce yang berdomisili di Kota Palembang. Responden dipilih menggunakan pendekatan *non-probability sampling* dengan teknik *purposive sampling*. Kriteria inklusi meliputi (1) berdomisili di Kota Palembang dan (2) merupakan pengguna aktif yang telah melakukan transaksi pada platform e-commerce. Dengan kriteria tersebut, penelitian berfokus

Jurnal Manajemen dan Teknologi
https://journal.stmiki.ac.id
E-ISSN: 3047-8251 | P-ISSN: 3047-826X
Vol. 3 No. 2 (2026) | November
OPEN ACCESS
https://doi.org/ 10.63447/jmt.v3i2.2128

pada konsumen yang memiliki pengalaman transaksi, bukan hanya pengguna yang pernah mengakses platform. Untuk menilai kecukupan jumlah sampel, penelitian mengacu pada pedoman analisis multivariat yang dikutip oleh (Ali Memon *et al.*, 2020; Hair *et al.*, 2019), yaitu rasio acuan minimal 10 observasi untuk setiap variabel analisis. Penelitian ini mencakup tujuh variabel utama: *Recency*, *Frequency*, *Monetary*, *Share of Wallet*, motivasi belanja, preferensi kategori produk, dan kategori usia. Berdasarkan rasio tersebut, jumlah sampel acuan minimum adalah 70 responden, sedangkan data yang diperoleh berjumlah 198 responden. Jumlah ini melampaui acuan minimum, meskipun tidak dengan sendirinya menjamin bahwa sampel mewakili seluruh pengguna e-commerce di luar wilayah penelitian. Dataset disusun dalam bentuk data campuran (*mixed-data*), dengan setiap observasi mewakili satu responden. Empat atribut numerik yang digunakan untuk pembentukan klaster ialah *Recency* (jarak waktu sejak transaksi terakhir), *Frequency* (jumlah transaksi dalam periode pengamatan), *Monetary* (nilai pengeluaran), dan *Share of Wallet* (SoW), yaitu persentase pengeluaran e-commerce terhadap total pendapatan bulanan. Dua atribut kategorikal yang digunakan dalam klasterisasi adalah motivasi belanja dan preferensi kategori produk. Kategori usia tidak dimasukkan sebagai fitur pembentuk klaster; atribut ini disimpan untuk analisis lanjutan setelah klaster terbentuk (post-hoc). Pemisahan tersebut memungkinkan persona dibentuk berdasarkan pola transaksi dan perilaku, kemudian asosiasinya dengan kelompok usia diuji secara terpisah. Penggunaan karakteristik transaksi dan perilaku sebagai dasar segmentasi juga sejalan dengan penelitian tentang analisis pelanggan berbasis pembelajaran mesin (Ullah *et al.*, 2023).

2.3 Pra-pemrosesan Data (Data Preprocessing)

Pra-pemrosesan data dilakukan untuk menyetarakan skala antarfitur numerik sebelum proses klasterisasi. Langkah ini diperlukan karena perbedaan rentang nilai dapat memengaruhi perhitungan jarak pada algoritma berbasis jarak. Dalam penelitian ini, fitur numerik yang meliputi *Recency*, *Frequency*, *Monetary*, dan *Share of Wallet* (SoW) distandardisasi menggunakan *Standard Scaling* dengan metode *Z-score* (Tabianan *et al.*, 2022). Standardisasi mengubah nilai setiap fitur agar memiliki rata-rata 0 dan simpangan baku 1. Proses ini dirumuskan sebagai berikut:

$$z = \frac{x-\mu}{\sigma}$$

Pada Persamaan (1), (z) merupakan nilai setelah standardisasi, (x) adalah nilai awal, (μ) adalah rata-rata fitur, dan (σ) adalah simpangan baku fitur. Sementara itu, fitur kategorikal yang meliputi motivasi belanja dan preferensi kategori produk tetap dipertahankan dalam bentuk kategorinya. Pemisahan penanganan tersebut memungkinkan fitur numerik dan kategorikal diproses bersama menggunakan algoritma K-Prototypes tanpa mengubah makna kategori menjadi nilai numerik yang tidak mencerminkan urutan atau jarak sebenarnya.

2.4 Algoritma K-Prototypes dan Pendekatan Global

Algoritma *K-Prototypes* digunakan karena mampu mengolah data numerik dan kategorikal secara bersamaan. Berbeda dengan *K-Means* yang menggunakan jarak *Euclidean* untuk data numerik, *K-Prototypes* menggabungkan perhitungan jarak untuk atribut numerik dengan pengukuran ketidakcocokan pada atribut kategorikal (Aschenbruck *et al.*, 2025). Dengan demikian, variabel seperti *Recency*, *Frequency*, *Monetary*, dan *Share of Wallet* dapat dianalisis bersama variabel motivasi belanja dan preferensi kategori produk. Proses klasterisasi bertujuan mengelompokkan responden yang memiliki karakteristik serupa ke dalam kelompok yang sama. Algoritma menghitung perbedaan karakteristik numerik dan kategorikal, kemudian menyesuaikan pusat klaster hingga diperoleh pengelompokan dengan perbedaan dalam setiap klaster yang lebih kecil. Pengaruh kedua jenis data diseimbangkan melalui parameter pembobotan, sehingga atribut numerik maupun kategorikal dapat berkontribusi terhadap pembentukan klaster. Dalam penelitian ini, sebanyak 198 responden diproses secara bersamaan menggunakan pendekatan *global clustering*. Artinya, seluruh responden dikelompokkan berdasarkan pola transaksi dan perilaku yang sama tanpa terlebih dahulu dipisahkan menurut kategori usia. Sebaliknya, *local clustering* membagi responden ke dalam kelompok

E-ISSN: 3047-8251 | P-ISSN: 3047-826X
Vol. 3 No. 2 (2026) | November
OPEN ACCESS
https://doi.org/ 10.63447/jmt.v3i2.2128

usia terlebih dahulu, kemudian melakukan klasterisasi secara terpisah pada setiap kelompok. Pendekatan global dipilih agar profil konsumen terbentuk berdasarkan kesamaan karakteristik transaksi dan perilaku, bukan berdasarkan batasan usia sejak awal. Kategori usia selanjutnya digunakan dalam analisis lanjutan (post-hoc) untuk menguji hubungan antara kelompok usia dan persona yang terbentuk. Dengan cara ini, penelitian dapat mengidentifikasi apakah karakteristik persona tertentu cenderung ditemukan pada kelompok usia tertentu tanpa menjadikan usia sebagai dasar pembentukan klaster.

2.5 Evaluasi Metrik Klasterisasi

Penentuan jumlah klaster (K) dilakukan menggunakan *Elbow Method* dan *Silhouette Score*. Kedua metrik ini digunakan untuk mengevaluasi hasil pengelompokan dan membantu menentukan jumlah klaster yang sesuai dengan karakteristik data. *Elbow Method* mengevaluasi perubahan nilai cost function atau inersia *K-Prototypes* ketika jumlah klaster ditambah. Nilai tersebut cenderung menurun seiring bertambahnya jumlah klaster. Jumlah klaster dipilih pada titik ketika penurunan mulai melandai, karena penambahan klaster setelah titik tersebut memberikan perbaikan yang semakin kecil (John *et al.*, 2023; Kristanti *et al.*, 2024). Sementara itu, *Silhouette Score* digunakan untuk menilai kesesuaian setiap responden dengan klasternya dibandingkan dengan klaster terdekat. Nilai yang lebih tinggi menunjukkan bahwa responden cenderung lebih sesuai dengan klasternya sendiri dan memiliki pemisahan yang lebih baik dari kelompok lain. Nilai rata-rata *Silhouette Score* digunakan sebagai pertimbangan tambahan untuk mengevaluasi kualitas pengelompokan (John *et al.*, 2023; Kristanti *et al.*, 2024). Kedua metrik tersebut digunakan secara bersamaan agar penentuan jumlah klaster tidak hanya bergantung pada penurunan nilai cost function, tetapi juga mempertimbangkan kualitas pemisahan antarklaster. Hasil evaluasi dan dasar pemilihan jumlah klaster dalam penelitian ini selanjutnya disajikan pada bagian hasil dan pembahasan.

2.6 Pengujian Hipotesis Asosiasi

Pengujian hipotesis dilakukan untuk mengetahui apakah terdapat hubungan yang signifikan antara kategori usia responden dan persona perilaku e-commerce yang terbentuk melalui algoritma *K-Prototypes*. Pengujian menggunakan *Chi-Square Test of Independence* setelah proses klasterisasi selesai. Dengan demikian, kategori usia tidak memengaruhi pembentukan klaster, tetapi digunakan untuk menguji hubungan antara karakteristik demografis dan persona yang dihasilkan. Hipotesis yang diuji terdiri atas dua pernyataan. Hipotesis nol (H_0) menyatakan bahwa tidak terdapat hubungan yang signifikan antara kategori usia dan persona perilaku e-commerce. Sebaliknya, hipotesis alternatif (H_1) menyatakan bahwa terdapat hubungan yang signifikan antara kedua variabel tersebut. Keputusan pengujian didasarkan pada nilai signifikansi (p-value). Jika p-value lebih kecil dari taraf signifikansi 0,05, hipotesis nol ditolak. Hasil ini menunjukkan adanya hubungan statistik antara kategori usia dan persona, tetapi tidak membuktikan hubungan sebab-akibat. Sebelum hasil uji diinterpretasikan, kelayakan frekuensi harapan pada tabel kontingensi diperiksa berdasarkan aturan Cochran. Pemeriksaan ini memastikan bahwa jumlah sel dengan frekuensi harapan kurang dari 5 tidak melebihi 20% dari seluruh sel. Pemeriksaan tersebut diperlukan agar hasil uji *Chi-Square* dapat diinterpretasikan secara tepat. Apabila hubungan yang signifikan ditemukan, kekuatannya dievaluasi menggunakan *Cramér's V*. Ukuran ini menunjukkan seberapa kuat hubungan antara kategori usia dan persona, bukan sekadar apakah hubungan tersebut signifikan. Karena tabel penelitian terdiri atas enam kategori usia dan empat persona, interpretasi *Cramér's V* disesuaikan dengan dimensi tabel kontingensi. Penyesuaian ini mengacu pada kriteria ukuran efek Cohen agar kekuatan hubungan tidak dinilai hanya berdasarkan ambang batas umum tanpa mempertimbangkan ukuran tabel. Melalui tahapan tersebut, penelitian mengevaluasi signifikansi dan kekuatan hubungan antara kategori usia dan persona yang terbentuk. Hasil pengujian selanjutnya digunakan untuk menginterpretasikan distribusi persona pada kelompok usia yang berbeda tanpa menjadikan usia sebagai dasar pembentukan klaster.

Jurnal Manajemen dan Teknologi
https://journal.stmiki.ac.id

E-ISSN: 3047-8251 | P-ISSN: 3047-826X
Vol. 3 No. 2 (2026) | November

OPEN ACCESS

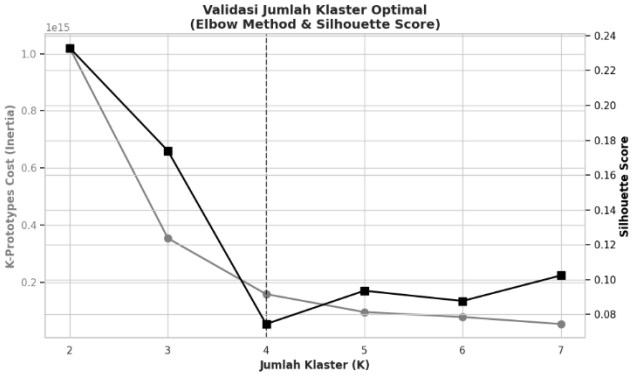
https://doi.org/ 10.63447/jmt.v3i2.2128

3. Hasil dan Pembahasan

3.1 Hasil

3.1.1 Penentuan Jumlah Klaster Optimal

Penentuan jumlah klaster dilakukan menggunakan *Elbow Method* dan *Silhouette Score* untuk memperoleh pembagian kelompok yang sesuai dengan karakteristik data. *Elbow Method* digunakan untuk mengamati perubahan nilai *cost function* ketika jumlah klaster bertambah, sedangkan *Silhouette Score* digunakan untuk menilai kesesuaian responden dengan klasternya serta pemisahan antar klaster.

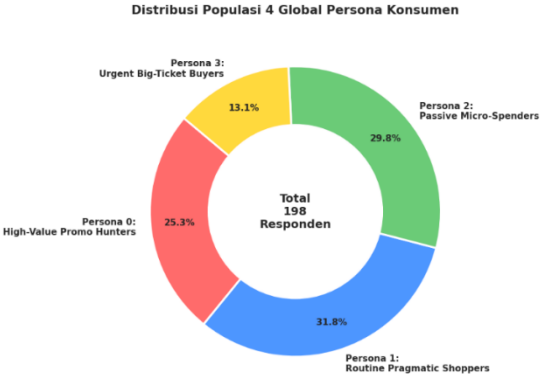


Gambar 2. Validasi Menggunakan Elbow dan Silhoutte Method

Hasil evaluasi pada Gambar 2 menunjukkan bahwa penambahan jumlah klaster menurunkan nilai *cost function*, tetapi penurunannya mulai melandai pada K = 4. Kondisi ini menunjukkan bahwa penambahan klaster setelah titik tersebut memberikan perbaikan yang semakin terbatas. Hasil *Silhouette Score* turut digunakan sebagai pertimbangan untuk menilai kualitas pemisahan kelompok. Berdasarkan kedua metrik tersebut, penelitian menetapkan empat klaster (K = 4) untuk tahap analisis selanjutnya. Pemilihan ini menghasilkan pembagian yang dapat diinterpretasikan menjadi empat persona konsumen berdasarkan karakteristik transaksi dan perilaku belanja.

3.1.2 Visualisasi Sebaran Data

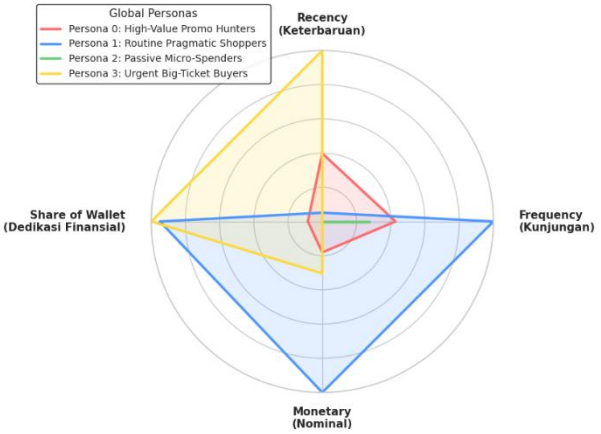
Setelah jumlah klaster ditetapkan sebanyak empat ((K=4)), algoritma K-Prototypes menghasilkan empat persona pengguna e-commerce dengan karakteristik transaksi dan perilaku belanja yang berbeda. Distribusi responden pada setiap persona ditampilkan pada Gambar 3. Karakteristik masing-masing persona selanjutnya dianalisis berdasarkan Recency, Frequency, Monetary, Share of Wallet (SoW), dan motivasi belanja.



Jurnal Manajemen dan Teknologi
https://journal.stmiki.ac.id
E-ISSN: 3047-8251 | P-ISSN: 3047-826X
Vol. 3 No. 2 (2026) | November
OPEN ACCESS
https://doi.org/ 10.63447/jmt.v3i2.2128

Gambar 3. Donut Chart

Berdasarkan Gambar 3, *Routine Pragmatic Shoppers* merupakan kelompok terbesar, yaitu 31,8% dari 198 responden. Kelompok ini diikuti oleh *Passive Micro-Spenders* sebesar 29,8%, *High-Value Promo Hunters* sebesar 25,3%, dan *Urgent Big-Ticket Buyers* sebesar 13,1%. Distribusi tersebut menunjukkan bahwa responden memiliki pola belanja yang beragam. Kelompok terbesar berorientasi pada pemenuhan kebutuhan rutin, sedangkan kelompok lainnya menunjukkan kecenderungan belanja yang berkaitan dengan promosi, pengeluaran bernilai kecil, atau kebutuhan mendesak. Perbedaan karakteristik keempat persona dapat diamati lebih lanjut melalui perbandingan indikator numerik pada Gambar 4.



Gambar 4. Radar Chart

Berdasarkan Gambar 4, setiap persona memiliki pola transaksi dan perilaku yang berbeda. Perbedaan tersebut dapat dijelaskan sebagai berikut.

1) Persona 0 (*High-Value Promo Hunters*): Kelompok ini mencakup 25,3% responden dan ditandai oleh nilai pengeluaran serta SoW yang tinggi, dengan SoW di atas 24%. Meskipun memiliki pengeluaran besar, kelompok ini sensitif terhadap harga dan cenderung tertarik pada promosi serta *flash sale*. Artinya, nilai transaksi yang tinggi tidak selalu menunjukkan bahwa pelanggan kurang memperhatikan harga. Dari sisi pemasaran, platform dapat menawarkan kupon yang relevan, promosi pada waktu tertentu, dan paket produk sesuai preferensi pelanggan. Informasi mengenai diskon, syarat promosi, dan harga akhir sebaiknya ditampilkan secara jelas agar pelanggan dapat membandingkan manfaat penawaran sebelum melakukan pembelian.
2) Persona 1 (*Routine Pragmatic Shoppers*): Sebagai kelompok terbesar dengan proporsi 31,8%, persona ini memiliki tingkat frekuensi transaksi dan dedikasi finansial yang relatif moderat. Konsumen dalam kelompok ini cenderung menggunakan e-commerce untuk memenuhi kebutuhan rutin dengan pertimbangan kegunaan dan kemudahan. Strategi pemasaran yang sesuai meliputi rekomendasi produk kebutuhan sehari-hari, pengingat pembelian ulang, dan penawaran berdasarkan riwayat belanja. Pada desain antarmuka, navigasi kategori, pencarian produk, informasi harga, dan proses pembayaran perlu dibuat sederhana agar transaksi rutin dapat dilakukan secara efisien.
3) Persona 2 (*Passive Micro-Spenders*): Kelompok ini mencakup 29,8% responden, dengan SoW di bawah 12% dan aktivitas transaksi yang relatif jarang. Pengeluaran mereka cenderung bernilai kecil dan berkaitan dengan kebutuhan pelengkap atau gaya hidup sekunder. Dengan demikian, keterlibatan kelompok ini dalam transaksi e-commerce relatif terbatas dibandingkan kelompok dengan pengeluaran lebih tinggi. Pendekatan pemasaran yang sesuai adalah menawarkan produk terjangkau, rekomendasi yang relevan, dan promosi dengan nilai pembelian yang sesuai anggaran. Antarmuka pengguna dapat menampilkan pilihan produk berdasarkan rentang harga sehingga pelanggan lebih mudah menemukan produk yang sesuai dengan kebutuhan mereka.

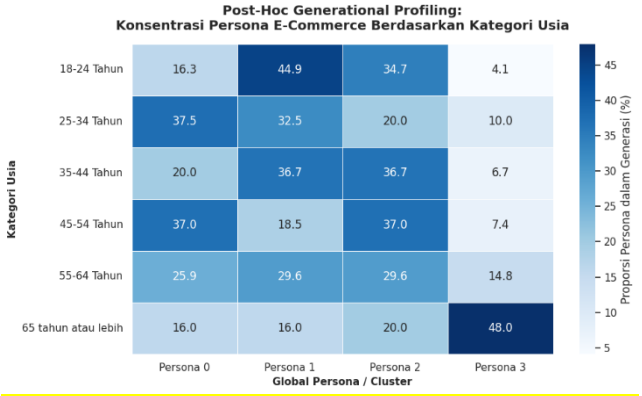
Jurnal Manajemen dan Teknologi
https://journal.stmiki.ac.id
E-ISSN: 3047-8251 | P-ISSN: 3047-826X
Vol. 3 No. 2 (2026) | November
OPEN ACCESS
https://doi.org/ 10.63447/jmt.v3i2.2128

4) Persona 3 (*Urgent Big-Ticket Buyers*): Kelompok ini merupakan segmen terkecil, yaitu 13,1% responden. Karakteristik utamanya adalah frekuensi transaksi yang rendah, tetapi nilai pengeluaran dan SoW tinggi ketika pembelian dilakukan. SoW kelompok ini berada di atas 24%. Pola tersebut menunjukkan bahwa transaksi bernilai tinggi dapat terjadi ketika konsumen menghadapi kebutuhan yang mendesak. Strategi pemasaran untuk kelompok ini sebaiknya menekankan ketersediaan produk, kejelasan spesifikasi, estimasi pengiriman, dan informasi garansi. Pada antarmuka pengguna, informasi penting perlu mudah ditemukan agar pelanggan dapat mengevaluasi produk dan menyelesaikan transaksi dengan lebih cepat dan yakin.

Secara keseluruhan, empat persona tersebut menunjukkan bahwa pengguna dengan pola pengeluaran yang tinggi sekalipun dapat memiliki motivasi yang berbeda. Sebagian sensitif terhadap promosi, sebagian berfokus pada kebutuhan rutin, sebagian bertransaksi dalam jumlah kecil, dan sebagian lainnya melakukan pembelian bernilai tinggi ketika menghadapi kebutuhan mendesak. Oleh karena itu, hasil klasterisasi dapat digunakan sebagai dasar untuk menyusun strategi promosi, rekomendasi produk, dan desain antarmuka yang lebih sesuai dengan karakteristik setiap kelompok pelanggan.

3.1.3 Pengujian Hipotesis dan Post-Hoc Generational Profiling

Pengujian Chi-Square dilakukan untuk mengetahui apakah terdapat hubungan antara kategori usia responden dan persona yang terbentuk melalui algoritma K-Prototypes. Hasil pengujian menunjukkan nilai Chi-Square sebesar 44,58 dengan *p-value* sebesar 0,0001. Karena nilai tersebut lebih kecil dari taraf signifikansi 0,05, hipotesis nol ditolak. Hasil ini menunjukkan bahwa terdapat hubungan yang signifikan secara statistik antara kategori usia dan distribusi persona konsumen. Kekuatan hubungan tersebut selanjutnya dievaluasi menggunakan Cramér's V dan menghasilkan nilai 0,274. Nilai ini menunjukkan adanya hubungan antara kategori usia dan persona, tetapi kekuatannya perlu diinterpretasikan berdasarkan kriteria yang digunakan dalam penelitian. Dengan demikian, usia berkaitan dengan pola segmentasi konsumen, tetapi tidak sepenuhnya menentukan perilaku belanja seseorang. Distribusi persona pada setiap kategori usia ditampilkan dalam bentuk *heatmap* pada Gambar 5. Visualisasi ini membantu memperlihatkan persona yang memiliki proporsi relatif besar pada setiap kelompok usia.



Gambar 5. Visualisasi Heatmap

Berdasarkan pemetaan proporsi pada Gambar 5, Berdasarkan Gambar 5, terdapat beberapa pola yang menunjukkan bahwa karakteristik konsumen tidak dapat dijelaskan hanya berdasarkan kelompok usia. Pada responden berusia 18–24 tahun, sebanyak 44,9% termasuk *Routine Pragmatic Shoppers* dan 34,7% termasuk *Passive Micro-Spenders*. Kedua persona tersebut mencakup 79,6% dari kelompok usia ini. Temuan tersebut menunjukkan bahwa dalam sampel penelitian, konsumen muda lebih banyak berada pada kelompok yang berorientasi pada kebutuhan rutin atau memiliki pengeluaran e-commerce relatif kecil, bukan pada kelompok dengan pengeluaran tinggi yang dipicu

Jurnal Manajemen dan Teknologi
https://journal.stmiki.ac.id
E-ISSN: 3047-8251 | P-ISSN: 3047-826X
Vol. 3 No. 2 (2026) | November
OPEN ACCESS
https://doi.org/ 10.63447/jmt.v3i2.2128

oleh promosi. Pada kelompok usia 25–34 tahun, *High-Value Promo Hunters* memiliki proporsi terbesar, yaitu 37,5%. Sementara itu, pada kelompok usia 35–44 tahun, *Routine Pragmatic Shoppers* dan *Passive Micro-Spenders* masing-masing mencakup 36,7%. Pada kelompok usia 45–54 tahun, *High-Value Promo Hunters* dan *Passive Micro-Spenders* masing-masing memiliki proporsi 37,0%. Pola ini memperlihatkan bahwa kelompok usia produktif tidak memiliki satu pola belanja yang seragam. Sebagian konsumen menunjukkan respons terhadap promosi, sementara sebagian lainnya tetap berorientasi pada kebutuhan rutin atau melakukan pembelian dengan nilai relatif kecil. Pada kelompok usia 55–64 tahun, *Routine Pragmatic Shoppers* dan *Passive Micro-Spenders* masing-masing mencakup 29,6%. Adapun pada kelompok usia 65 tahun ke atas, sebanyak 48,0% responden termasuk *Urgent Big-Ticket Buyers*. Proporsi tersebut menunjukkan bahwa persona dengan frekuensi transaksi rendah tetapi nilai pembelian tinggi juga ditemukan pada kelompok usia lanjut dalam sampel penelitian. Hasil pengujian dan distribusi persona menunjukkan bahwa kategori usia memiliki hubungan statistik dengan pola segmentasi konsumen. Namun, variasi persona dalam kelompok usia yang sama menegaskan bahwa pelanggan tidak seharusnya diperlakukan berdasarkan asumsi demografis semata. Strategi pemasaran perlu mempertimbangkan karakteristik transaksi, motivasi belanja, dan preferensi produk agar lebih sesuai dengan kebutuhan setiap kelompok konsumen.

3.2 Pembahasan

Penelitian terdahulu menunjukkan bahwa segmentasi pelanggan e-commerce telah dikembangkan melalui berbagai pemilihan fitur dan algoritma klasterisasi. (Awaliyah *et al.*, 2024; Wang, 2025) mengkaji segmentasi pelanggan toko virtual menggunakan data numerik dan kategorikal serta membandingkan beberapa pendekatan klasterisasi. Sementara itu, (Wu *et al.*, 2022) memperluas model RFM dengan menambahkan aktivitas pelanggan pada keranjang belanja dan daftar favorit, kemudian menerapkannya menggunakan algoritma K-Medoids yang ditingkatkan. Kedua penelitian tersebut menegaskan pentingnya pemilihan fitur dan metode yang sesuai dengan karakteristik data pelanggan. (Alves Gomes & Meisen, 2023), melalui tinjauan terhadap 105 publikasi, menunjukkan bahwa RFM dan K-Means banyak digunakan dalam penelitian segmentasi e-commerce. Perkembangan berikutnya juga terlihat pada penelitian (Ho *et al.*, 2023), yang menerapkan K-Prototypes untuk mengelompokkan pelanggan berdasarkan data demografis, transaksi, dan perilaku sebagai bagian dari kerangka prediksi *customer churn*. Studi tersebut memperlihatkan bahwa K-Prototypes telah digunakan dalam analisis pelanggan e-commerce dengan tujuan mendukung strategi retensi (Rydzewska *et al.*, 2024; Schwendtner *et al.*, 2024; Singh, 2024). Berbeda dari fokus penelitian terdahulu tersebut, penelitian ini menggabungkan RFM dan Share of Wallet (SoW) dengan motivasi belanja serta preferensi kategori produk untuk membentuk persona konsumen. Kategori usia tidak digunakan sebagai dasar pembentukan klaster, tetapi dianalisis setelah klaster terbentuk melalui uji Chi-Square. Pendekatan ini memungkinkan penelitian mengidentifikasi variasi perilaku dalam kelompok usia yang sama dan menilai hubungan antara usia dan persona tanpa mengasumsikan bahwa usia secara langsung menentukan perilaku belanja. Dengan demikian, kontribusi penelitian terletak pada integrasi karakteristik transaksi dan perilaku dalam pembentukan persona, disertai analisis hubungan demografis serta interpretasi praktis bagi strategi pemasaran e-commerce.

4. Kesimpulan

Penelitian ini menerapkan algoritma K-Prototypes untuk mengelompokkan pengguna e-commerce berdasarkan karakteristik transaksi dan perilaku belanja menggunakan 198 responden di Kota Palembang. Evaluasi menggunakan Elbow Method dan Silhouette Score menghasilkan empat persona, yaitu *High-Value Promo Hunters*, *Routine Pragmatic Shoppers*, *Passive Micro-Spenders*, dan *Urgent Big-Ticket Buyers*. Keempat persona tersebut menunjukkan bahwa konsumen memiliki pola transaksi, alokasi pengeluaran, dan motivasi belanja yang berbeda sehingga segmentasi pelanggan tidak cukup

E-ISSN: 3047-8251 | P-ISSN: 3047-826X
Vol. 3 No. 2 (2026) | November
OPEN ACCESS
https://doi.org/ 10.63447/jmt.v3i2.2128

hanya mengandalkan nilai transaksi. Pengujian Chi-Square menunjukkan adanya hubungan yang signifikan antara kategori usia dan persona, dengan nilai Cramér's V sebesar 0,274. Temuan ini menunjukkan bahwa usia berkaitan dengan pola belanja konsumen, tetapi tidak sepenuhnya menentukan karakteristik setiap pelanggan. Hasil segmentasi dapat dimanfaatkan sebagai dasar untuk menyusun promosi yang lebih relevan, rekomendasi produk, dan desain antarmuka yang sesuai dengan kebutuhan masing-masing persona. Penelitian ini memiliki keterbatasan pada jumlah sampel sebanyak 198 responden, wilayah pengumpulan data yang terbatas di Kota Palembang, serta penggunaan teknik *purposive sampling* yang membatasi generalisasi hasil ke populasi pengguna e-commerce yang lebih luas. Penelitian selanjutnya disarankan melibatkan responden dari berbagai wilayah dengan jumlah sampel yang lebih besar serta membandingkan K-Prototypes dengan algoritma klasterisasi lain yang mampu mengolah data campuran. Pengujian tersebut dapat membantu menilai konsistensi hasil klasterisasi dan meningkatkan pemahaman mengenai karakteristik serta perilaku konsumen e-commerce.

5. Ucapan Terima Kasih

Ucapan terima kasih terutama ditujukan kepada pemberi dana penelitian atau donatur. Ucapan terima kasih dapat juga disampaikan kepada pihak-pihak yang membantu pelaksanaan penelitian. Dosen / peneliti yang tercantum dalam daftar penulis tidak perlu diberikan ucapan terima kasih di bagian ini.

6. Daftar Pustaka

Ali Memon, M., Ting, H., Cheah, J.-H., Thurasamy, R., Chuah, F., & Huei Cham, T. (2020). Sample size for survey research: Review and recommendations. *Journal of Applied Structural Equation Modeling*, 4(2).

Alie, J., & Gustriansyah, R. (2024). Customer segmentation for digital marketing based on shopping patterns. *Jurnal Aplikasi Bisnis Dan Manajemen*. https://doi.org/10.17358/jabm.10.1.209.

Alves Gomes, M., & Meisen, T. (2023). A review on customer segmentation methods for personalized customer targeting in e-commerce use cases. *Information Systems and E-Business Management*, 21(3), 527–570. https://doi.org/10.1007/s10257-023-00640-4.

Aschenbruck, R., Szepannek, G., & Wilhelm, A. F. X. (2025). Initialization strategies for clustering mixed-type data with the k-prototypes algorithm. *Advances in Data Analysis and Classification*. https://doi.org/10.1007/s11634-025-00639-4.

Awaliyah, D. A., Budi Prasetiyo, Muzayanah, R., & Lestari, A. D. (2024). Optimizing customer segmentation in online retail transactions through the implementation of the K-Means clustering algorithm. *Scientific Journal of Informatics*, 11(2), 539–548. https://doi.org/10.15294/sji.v11i2.6137.

Blömker, J., & Albrecht, C. M. (2024). Psychographic segmentation of multichannel customers: Investigating the influence of individual differences on channel choice and switching behavior. *Journal of Retailing and Consumer Services*, 79. https://doi.org/10.1016/j.jretconser.2024.103806.

Gholamveisy, S., Homayooni, S., Shemshaki, M., Sheykhan, S., Boozary, P., Tanhaei, H. G., & Akbari, N. (2024). Application of data mining technique for customer purchase behavior via

E-ISSN: 3047-8251 | P-ISSN: 3047-826X
Vol. 3 No. 2 (2026) | November
OPEN ACCESS
https://doi.org/ 10.63447/jmt.v3i2.2128

extended RFM model with focus on BCG matrix from a data set of online retailing. *Journal of Infrastructure, Policy and Development*, 8(7). https://doi.org/10.24294/jipd.v8i7.4426.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage.

Ho, T., Nguyen, S., Nguyen, H., Nguyen, N., Man, D. S., & Le, T. G. (2023). An extended RFM model for customer behaviour and demographic analysis in the retail industry. *Business Systems Research*, 14(1), 26–53. https://doi.org/10.2478/bsrj-2023-0002.

Homepage, J., Fadhillah, M. F., Lovely, A., Suyoso, A., & Puspitasari, I. (2025). Customer segmentation with clustering algorithm based on recency, frequency, and monetary (RFM) attributes. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 5(1), 48–56. https://doi.org/10.57152/malcom.v5v1.1491.

John, J. M., Shobayo, O., & Ogunleye, B. (2023). An exploration of clustering algorithms for customer segmentation in the UK retail market. *Analytics*, 2(4), 809–823. https://doi.org/10.3390/analytics2040042.

Kristanti, B. T., Junaidi, A., & Mandyartha, E. P. (2024). Implementasi K-Means clustering dalam segmentasi pelanggan berdasarkan usia, pendapatan, dan model RFM (studi kasus: Lantiknya Store Jombang). *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(3). https://doi.org/10.23960/jitet.v12i3.4677.

Kumar, N. (2025). Intelligent customer segmentation: Unveiling consumer patterns with machine learning. *Journal of Umm Al-Qura University for Engineering and Architecture*, 16(3), 774–783. https://doi.org/10.1007/s43995-025-00180-7.

Mahesa Putra, A., Saputra, C., Hidayatullah, R., Irsandi, S., & Teknik Informatika, P. (2025). Penerapan algoritma K-Means clustering dalam menganalisis tren konsumen e-commerce. *Desember*, 6(2).

Mahfuza, R., Islam, N., Toyeb, M., Emon, M. A. F., Chowdhury, S. A., & Alam, M. G. R. (2022). LRFMV: An efficient customer segmentation model for superstores. *PLoS ONE*, 17(12). https://doi.org/10.1371/journal.pone.0279262.

Rydzewska, K., Nielek, R., Pawłowska, J., Wierzbicki, A., & Sedek, G. (2024). Performance of aging consumers in an e-commerce product choice task: The role of working memory and decision strategies. *PLoS ONE*, 19(6). https://doi.org/10.1371/journal.pone.0303806.

Schwendtner, T., Amsl, S., Teller, C., & Wood, S. (2024). Shopping behaviour of elderly consumers: Change and stability during times of crisis. *International Journal of Retail and Distribution Management*, 52(13), 1–15. https://doi.org/10.1108/IJRDM-01-2023-0029.

Shi, Y., Zhang, S., Wang, S., Xie, H., & Feng, J. (2024). Multiple-perspective consumer segmentation using improved weighted fuzzy k-prototypes clustering and swarm intelligence algorithm for fresh apricot market. *Italian Journal of Food Science*, 36(4), 38–56. https://doi.org/10.15586/ijfs.v36i4.2552.

Singh, P. (2024). What drives or decelerates Generation Z? An empirical study navigating consumer buying intentions in online shopping. *SAGE Open*, 14(3). https://doi.org/10.1177/21582440241263173.

E-ISSN: 3047-8251 | P-ISSN: 3047-826X
Vol. 3 No. 2 (2026) | November
OPEN ACCESS
https://doi.org/ 10.63447/jmt.v3i2.2128

Tabianan, K., Velu, S., & Ravi, V. (2022). K-Means clustering approach for intelligent customer segmentation using customer purchase behavior data. *Sustainability (Switzerland)*, 14(12). https://doi.org/10.3390/su14127243.

Ullah, A., Mohmand, M. I., Hussain, H., Johar, S., Khan, I., Ahmad, S., Mahmoud, H. A., & Huda, S. (2023). Customer analysis using machine learning-based classification algorithms for effective segmentation using recency, frequency, monetary, and time. *Sensors*, 23(6). https://doi.org/10.3390/s23063180.

Vrhovac, V., Orošnjak, M., Ristić, K., Sremčev, N., Jocanović, M., Spajić, J., & Brkljač, N. (2024). Unsupervised modelling of e-customers' profiles: Multiple correspondence analysis with hierarchical clustering of principal components and machine learning classifiers. *Mathematics*, 12(23). https://doi.org/10.3390/math12233794.

Wang, G. (2025). Customer segmentation in the digital marketing using a Q-learning based differential evolution algorithm integrated with K-means clustering. *PLoS ONE*, 20(2). https://doi.org/10.1371/journal.pone.0318519.

Wu, Z., Jin, L., Zhao, J., Jing, L., & Chen, L. (2022). Research on segmenting e-commerce customer through an improved K-Medoids clustering algorithm. *Computational Intelligence and Neuroscience*, 2022. https://doi.org/10.1155/2022/9930613.