

**SEGMENTASI KEBIASAAN PENGELUARAN GENERASI Z
MENGUNAKAN ALGORITMA K-MEANS CLUSTERING:
STUDI EMPIRIS POLA PERILAKU KEUANGAN**

LAPORAN NON-REGULER

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



Disusun oleh :

GUNAWAN SYLVESTER

22.11.5268

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**SEGMENTASI KEBIASAAN PENGELUARAN GENERASI Z
MENGUNAKAN ALGORITMA K-MEANS CLUSTERING: STUDI
EMPIRIS POLA PERILAKU KEUANGAN**

LAPORAN NON-REGULER

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



Disusun oleh :

GUNAWAN SYLVESTER

22.11.5268

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

HALAMAN PERSETUJUAN

HALAMAN PERSETUJUAN

JALUR NON-REGULER

SEGMENTASI KEBIASAAN PENGELUARAN GENERASI Z
MENGUNAKAN ALGORITMA K-MEANS CLUSTERING: STUDI EMPIRIS
POLA PERILAKU KEUANGAN

yang disusun dan diajukan oleh

Gunawan Sylvester

22.11.5268

telah disetujui oleh Dosen Pembimbing
pada tanggal 16 Desember 2025

Dosen Pembimbing,



Mardiana Cahyani, S.Kom., M.Eng.
NIK. 190302393

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

JALUR NON-REGULER

**SEGMENTASI KEBIASAAN PENGELUARAN GENERASI Z
MENGUNAKAN ALGORITMA K-MEANS CLUSTERING; STUDI EMPIRIS
POLA PERILAKU KEUANGAN**

yang disusun dan diajukan oleh

Gunawan Sylvester
22.11.5268

Telah dipertahankan di depan Dewan Penguji
pada tanggal 16 Desember 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Bayu Setiaji, M.Kom.
NIK. 190302216



El Fajastuti, S.Kom., M.Kom.
NIK. 190302227



Maifid Rahardil, S.Kom., M.Eng.
NIK. 190302193



Laporan ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 16 Desember 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, S.Kom., M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN KARYA

HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : Gunawan Sylvester

NIM : 22.11.5268

Menyatakan bahwa Artikel dengan judul berikut

SEGMENTASI KEBIASAAN PENGELUARAN GENERASI Z MENGUNAKAN ALGORITMA K-MEANS CLUSTERING: STUDI EMPIRIS POLA PERILAKU KEUANGAN

Dosen Pembimbing : Mujid Rahardi, S.Kom., M.Eng.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rancangan dan kegiatan SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidak-benaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 16 Desember 2025



Gunawan Sylvester

HALAMAN PERSEMBAHAN

Artikel ini saya persembahkan kepada:

- Tuhan Yesus Kristus, sumber kekuatan, hikmat, dan anugerah dalam setiap langkah hidup saya, termasuk dalam proses penyusunan artikel ini.
- Kedua orang tua tercinta, Papa dan Mama saya, adik-adik kandung saya, Arnold dan Crysnatalio Sylvester, atas kasih, doa, dan dukungan tanpa batas yang selalu menjadi penyemangat utama dalam setiap perjuangan.
- Para dosen dan pembimbing yang telah membagikan ilmu serta memberikan arahan dan bimbingan dengan penuh kesabaran.
- Sahabat dan rekan seperjuangan yang selalu memberikan semangat, kebersamaan, dan motivasi selama masa perkuliahan; Edwin, Agung, Hevin, Rizki, dan Farrel.
- Zeal Ministry, sebagai keluarga rohani yang terus mendukung pertumbuhan iman dan menjadi tempat penguatan secara spiritual; khususnya Bang Daniel, Kak Lusi, BPC, Kak Menry, Babe-Mami, Nike, Dave, Jouvan, Nicholas, Imey, dan nama-nama lainnya yang tidak dapat saya sebutkan satu persatu.

Segala hormat dan syukur hanya bagi Tuhan Yesus Kristus. Semoga karya ini menjadi berkat bagi banyak orang.

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala limpahan rahmat, karunia, dan kemudahan yang telah diberikan sehingga penulis dapat menyelesaikan Artikel yang berjudul **“Segmentasi Kebiasaan Pengeluaran Generasi Z Menggunakan Algoritma K-Means Clustering: Studi Empiris Pola Perilaku Keuangan”** sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika.

Penyusunan artikel ini tentu tidak terlepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

- Ibu Prof. Dr. Kusriani, S.Kom., M.Kom., selaku Dekan Fakultas Ilmu Komputer.
- Ibu Eli Pujastuti, M.Kom., selaku Ketua Program Studi Informatika.
- Bapak Majid Rahardi, S.Kom., M.Eng., selaku Dosen Pembimbing yang telah dengan sabar membimbing, memberikan arahan, serta masukan yang sangat berarti selama proses penyusunan artikel ini.
- Bapak/Ibu Tim Dosen Penguji, yang telah memberikan kritik, saran, dan evaluasi yang membangun demi penyempurnaan karya ilmiah ini.
- Kedua orang tua tercinta yang selalu memberikan doa, dukungan moral, dan semangat yang tiada henti.
- Serta seluruh pihak yang telah membantu secara langsung maupun tidak langsung dalam penyusunan artikel ini.

Penulis menyadari bahwa artikel ini masih jauh dari sempurna. Oleh karena itu, penulis membuka diri terhadap segala kritik dan saran yang membangun untuk perbaikan di masa mendatang.

Yogyakarta, 16 Desember 2025

Penulis

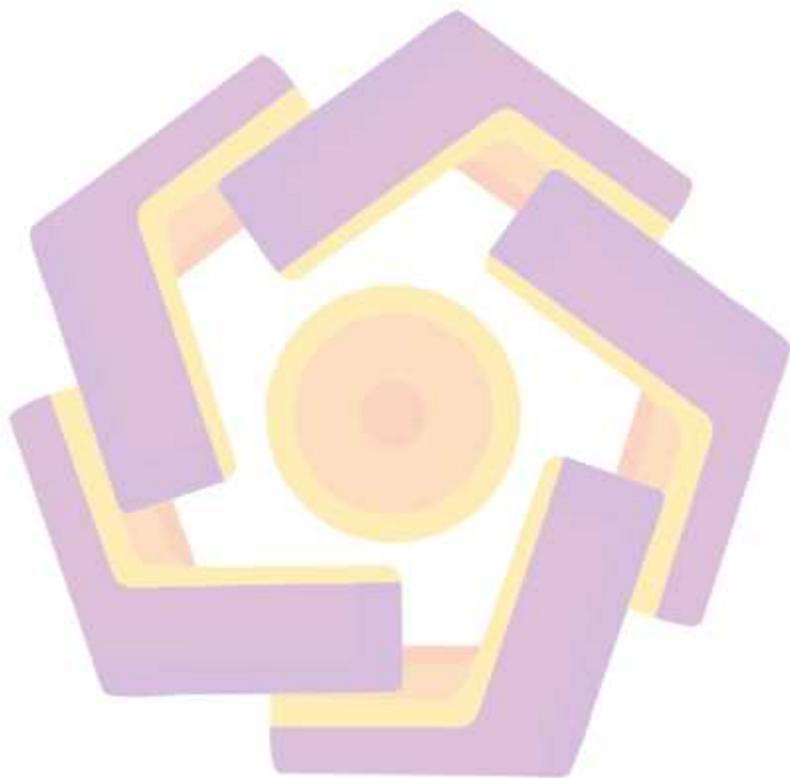
DAFTAR ISI

Halaman Judul.....	1
Halaman Persetujuan.....	2
Halaman Pengesahan	3
Halaman Pernyataan Keaslian Karya.....	4
Halaman Persembahan	5
Kata Pengantar	6
Daftar Isi (*).....	7
Daftar Tabel(*).....	9
Daftar Gambar (*).....	10
Daftar Lampiran.....	11
Daftar Lambang dan Singkatan.....	12
Daftar Istilah.....	13
Intisari	15
<i>Abstract</i>	16
Bab I Pendahuluan	1
1.1. Gambaran Umum.....	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah.....	2
1.4. Tujuan	3
Bab II Tinjauan Pustaka	4
2.1. Studi Literatur	4
2.2. Landasan Teori.....	6
BAB III Metode Penelitian	8
3.1 Akuisisi Data.....	8
3.2 Pra-pemrosesan Data.....	9
3.3 Eksplorasi Data	9
3.4 Penentuan Jumlah Cluster	10
3.5 Klasterisasi	10
3.6 Interpretasi Hasil.....	13
BAB IV Pembahasan	14

4.1 Deskripsi dan Visualisasi Data	14
4.2 Hasil Klasterisasi K-Means	19
4.2.1 Centroid Setiap Klaster	22
4.2.2 Ukuran Setiap Klaster	23
4.2.3 Perhitungan selisih (Euclidean Distance) antara Centroid dan menampilkan DBI untuk setiap K	24
4.2.4 Menampilkan nilai Davies-Bouldin Index (DBI) untuk setiap K	26
4.2.5 Informasi centroid dan klaster Centroid Setiap Klaster (Skala Asli) ...	28
4.2.6 Perbandingan antar metrik validasi (Davies-Bouldin Index (DBI), Silhouette Score, dan Calinski-Harabasz Index)	30
4.3 Visualisasi Hasil Klasterisasi	31
4.4 Implikasi Hasil Penelitian	38
BAB V Kesimpulan	40
5.1 Kesimpulan	40
5.2 Saran	40
Referensi	43
Curriculum Vitae	46
Lampiran dan Bukti Pendukung	48

DAFTAR TABEL

Tabel 3. 1 Contoh beberapa baris dengan Klaster	19
Tabel 3. 2 Ukuran Setiap Klaster	23
Tabel 3. 3 Matriks Selisih (Euclidean Distance) Antar Centroid	25

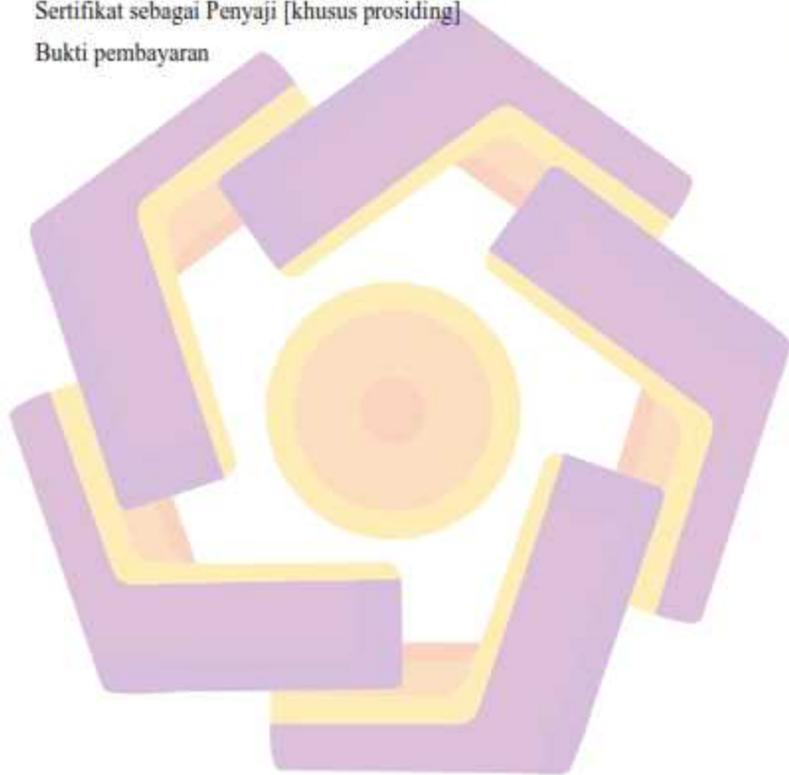


DAFTAR GAMBAR

Gambar 3.1 Flowchart/Alur Penelitian	8
Gambar 3.2 Dataset Generasi Z Money Spending yang diambil dari Kaggle	9
Gambar 4.1 Dataset "genz_money_spends.csv"	14
Gambar 4.2 Tampilan Data Setelah Dilakukan Normalisasi	15
Gambar 4.3 Visualisasi Histogram Data	15
Gambar 4.4 Visualisasi Heatmap Pengeluaran Gen-Z	17
Gambar 4.5 Visualisasi Boxplot Pengeluaran Terpilih	18
Gambar 4.6 Hasil Klasterisasi	20
Gambar 4.7 Selisih (Euclidean Distance) Antar Centroid	24
Gambar 4.8 Nilai DBI Pada Setiap K	26
Gambar 4.9 Informasi centroid dan kluster Centroid Setiap Kluster (Skala Asli)	28
Gambar 4.10 Perbandingan antar metrik validasi (Davies-Bouldin Index (DBI), Silhouette Score, dan Calinski-Harabasz Index)	30
Gambar 4.11 Penentuan Jumlah Kluster Dengan Elbow Method dan Silhouette Score	32
Gambar 4.12 Plot DBI untuk setiap K untuk visualisasi	33
Gambar 4.13 Visualisasi hasil klasterisasi menggunakan Scatter Plot	34
Gambar 4.14 Visualisasi hasil klasterisasi menggunakan Scatter Plot PCA	35

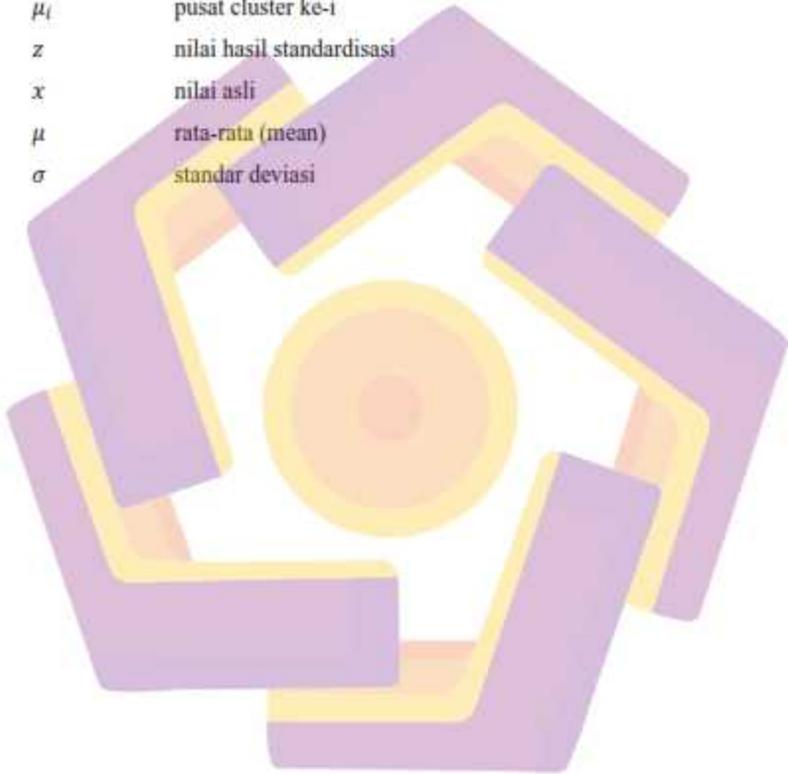
DAFTAR LAMPIRAN

Letter of Acceptance (LOA)	49
Lembar Review	49
Bukti Terbit/Terindex	53
Sertifikat sebagai Penyaji [khusus prosiding]	53
Bukti pembayaran	53



DAFTAR LAMBANG DAN SINGKATAN

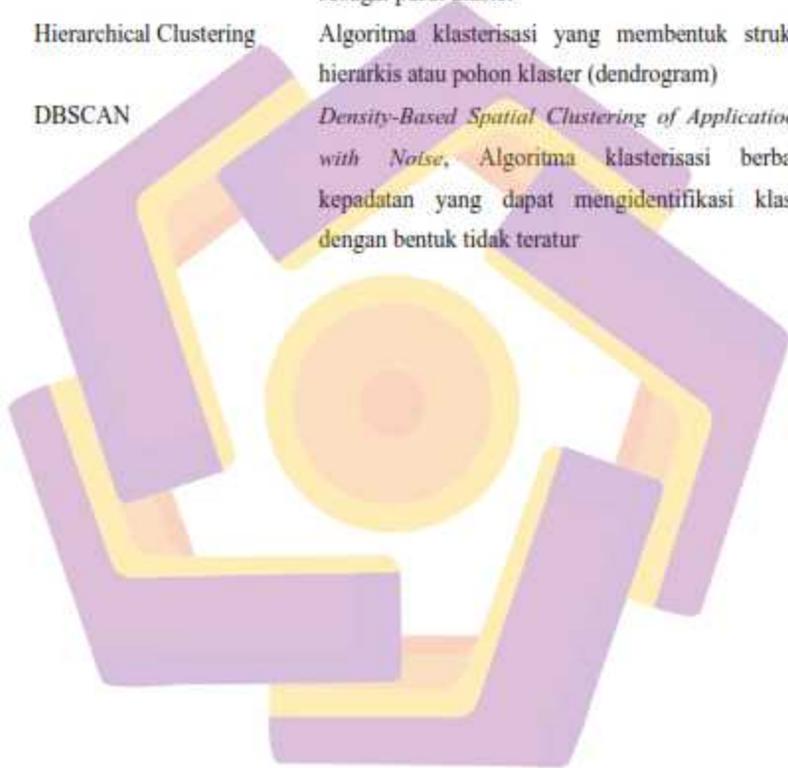
J	fungsi objektif (within-cluster sum of squares / WCSS)
C_i	cluster ke- i
x	anggota data dalam cluster
μ_i	pusat cluster ke- i
z	nilai hasil standardisasi
x	nilai asli
μ	rata-rata (mean)
σ	standar deviasi



DAFTAR ISTILAH

Clustering	Teknik pengelompokan data ke dalam kelompok (cluster) berdasarkan kemiripan
K-Means	Algoritma unsupervised learning yang membagi data ke dalam K cluster
Elbow Method	Metode untuk menentukan jumlah cluster optimal berdasarkan nilai WCSS
WCSS	<i>Within-Cluster Sim of Squares</i> Metode untuk menentukan jumlah cluster optimal berdasarkan nilai WCSS
MinMaxScaler	Teknik normalisasi untuk menyamakan skala fitur dalam dataset dengan merubah data ke rentang [0,1]
Centroid	Titik pusat dari setiap cluster yang terbentuk dalam algoritma K-Means
Preprocessing	Tahapan pembersihan dan transformasi data sebelum digunakan dalam analisis
Outlier	Data yang memiliki nilai jauh dari data lainnya
Feature	Atribut atau kolom data yang digunakan dalam analisis
Cluster Center	Posisi rata-rata dari semua titik dalam sebuah cluster
Interpretasi Cluster	Proses memahami karakteristik dari tiap cluster yang terbentuk
Davies-Bouldin Index (DBI)	Indeks yang mengukur kualitas kluster berdasarkan jarak antar centroid kluster dan sebaran dalam kluster
Silhouette Score	Skor yang mengukur seberapa baik kluster terpisah dengan menghitung keseragaman dan jarak antar kluster
Calinski-Harabasz Index	Indeks yang mengukur kualitas kluster berdasarkan kekompakan dalam kluster dan pemisahan antar kluster

PCA	<i>Principal Component Analysis</i> , Teknik reduksi dimensi untuk memvisualisasikan data dalam ruang dua dimensi
K-Medoids	Algoritma klasterisasi mirip K-Means tetapi lebih stabil terhadap outlier, menggunakan titik data aktual sebagai pusat kluster
Hierarchical Clustering	Algoritma klasterisasi yang membentuk struktur hierarkis atau pohon kluster (dendrogram)
DBSCAN	<i>Density-Based Spatial Clustering of Applications with Noise</i> , Algoritma klasterisasi berbasis kepadatan yang dapat mengidentifikasi kluster dengan bentuk tidak teratur



INTISARI

Generasi Z, yang lahir antara tahun 1997 dan 2012, menunjukkan perilaku konsumsi yang khas yang dipengaruhi oleh teknologi digital, gaya hidup modern, dan keputusan keuangan pribadi. Penelitian ini mengelompokkan perilaku keuangan mereka menggunakan algoritma K-Means pada dataset "Pengeluaran Uang Generasi Z" dari Kaggle. Selain K-Means, algoritma pengelompokan alternatif, seperti K-Medoids dan Pengelompokan Hierarkis, juga dievaluasi untuk menilai efektivitasnya dalam mengelompokkan perilaku keuangan. Dataset yang digunakan mencakup 1.700 individu dengan atribut pengeluaran untuk kategori seperti sewa, makanan, hiburan, pendidikan, tabungan, dan investasi. Dataset ini berisi 15 kolom, yang mencakup atribut numerik. Data telah dinormalisasi menggunakan Min-Max Scaling sebelum pengelompokan. Analisis ini mengidentifikasi enam kluster dengan karakteristik pengeluaran yang berbeda, mulai dari kelompok yang lebih berorientasi pada konsumsi (fokus pada hiburan dan belanja online) hingga kelompok yang lebih fokus pada tabungan dan investasi. Pendekatan kuantitatif diterapkan, yang mencakup analisis eksplorasi data, uji korelasi, dan Metode Elbow untuk menentukan jumlah kluster optimal. Hasilnya menunjukkan bahwa jumlah kluster optimal adalah 6, yang dikonfirmasi oleh Indeks Davies-Bouldin (DBI) sebesar 2,412, yang menunjukkan bahwa pemisahan kluster lebih lanjut mungkin diperlukan untuk meningkatkan kualitas segmentasi. Kluster 0 (rata-rata usia 20,6) fokus pada tabungan dan investasi sambil mengeluarkan uang secara moderat untuk sewa dan kebutuhan pokok. Kluster 1 (rata-rata usia 23,6) memprioritaskan pendidikan dan lebih banyak mengeluarkan uang untuk sewa sambil menabung untuk masa depan. Kluster 2 (rata-rata usia 20,3) sangat melek digital, menghabiskan uang untuk belanja online dan hiburan, dengan pengeluaran moderat untuk kebutuhan dasar. Kluster 3 (rata-rata usia 25,2) stabil secara finansial, menyeimbangkan pengeluaran dengan penekanan pada investasi dan tabungan. Kluster 4 (rata-rata usia 24,9) memprioritaskan tabungan dan investasi, dengan pengeluaran moderat untuk sewa dan kebutuhan pokok. Kluster 5 (rata-rata usia 24,96) berkomitmen pada tabungan dan investasi, dengan pengeluaran seimbang untuk kebutuhan pokok dan hiburan. Efektivitas model pengelompokan dibandingkan menggunakan kombinasi metrik validasi, termasuk Indeks Davies-Bouldin (DBI), Skor Siluet, dan Indeks Calinski-Harabasz, yang memberikan evaluasi komprehensif tentang pemisahan dan kekompakan kluster. Analisis ini menunjukkan bahwa pola pengeluaran Generasi Z bervariasi, dengan beberapa kluster menekankan hiburan dan konsumsi, sementara yang lainnya fokus pada stabilitas keuangan melalui tabungan dan investasi. Studi ini memberikan kontribusi pada pemahaman kebiasaan keuangan Generasi Z dan memberikan wawasan untuk strategi pemasaran dan kebijakan yang lebih terarah. Temuan ini memberikan wawasan berharga bagi bisnis dan penyedia layanan untuk menyesuaikan strategi dengan kebutuhan khusus dari setiap segmen.

Kata kunci: Pengambilan keputusan keuangan, Generasi Z, pengelompokan K-Means, segmentasi, perilaku pengeluaran.

ABSTRACT

Generation Z, born between 1997 and 2012, shows distinct consumption behaviors shaped by digital technology, modern lifestyles, and personal financial decisions. This research segments their financial behaviors using the K-Means clustering algorithm on the "Generation Z Money Spending" dataset from Kaggle. In addition to K-Means, alternative clustering algorithms, such as K-Medoids and Hierarchical Clustering, are also evaluated to assess their effectiveness in segmenting financial behaviors. The dataset used includes 1,700 individuals with spending attributes for categories such as rent, food, entertainment, education, savings, and investments. The dataset contains 15 columns, which include numerical attributes. The data has been normalized using Min-Max Scaling before clustering. The analysis identifies six clusters with distinct spending characteristics, ranging from groups that are more consumption-oriented (focusing on entertainment and online shopping) to groups that are more focused on savings and investments. A quantitative approach was applied, which included exploratory data analysis, correlation testing, and the Elbow Method to determine the optimal number of clusters. The results indicate that the optimal number of clusters is 6, as confirmed by the Davies-Bouldin Index (DBI) of 2.412, which further improvements in cluster separation may be necessary to enhance the quality of the segmentation. Cluster 0 (average age of 20.6) focuses on savings and investments while spending moderately on rent and groceries. Cluster 1 (average age 23.6) prioritizes education and spends more on rent while saving for the future. Cluster 2 (average age 20.3) is digitally savvy, spending on online shopping and entertainment, with moderate basic needs spending. Cluster 3 (average age 25.2) is financially stable, balancing spending with an emphasis on investments and savings. Cluster 4 (average age 24.9) prioritizes savings and investments, with moderate spending on rent and groceries. Cluster 5 (average age 24.96) is committed to savings and investments, with balanced spending on essentials and entertainment. The effectiveness of clustering models is compared using a combination of validation metrics, including the Davies-Bouldin Index (DBI), Silhouette Score, and Calinski-Harabasz Index, which provide a comprehensive evaluation of the cluster separation and compactness. The analysis shows that Generation Z's spending patterns vary, with some clusters emphasizing entertainment and consumption, while others focus on financial stability through savings and investments. This study contributes to the understanding of Generation Z's financial habits and provides insights for more targeted marketing strategies and policies. These findings provide valuable insights for businesses and service providers to tailor strategies to the distinct needs of each segment.

Keyword: financial decision-making, Generation Z, K-Means clustering, segmentation, spending behavior