



# **PROYEK UAS KECERDASAN BUATAN**



**ROBBY FIRMANSYAH**  
**312210643**



# Latar Belakang



Proyek ini bertujuan menerapkan metode clustering pada data tanpa label menggunakan algoritma K-Means. Untuk meningkatkan kualitas analisis, dilakukan preprocessing data, visualisasi TSNE, dan penentuan jumlah cluster optimal menggunakan Elbow Method.



# Tujuan Penelitian

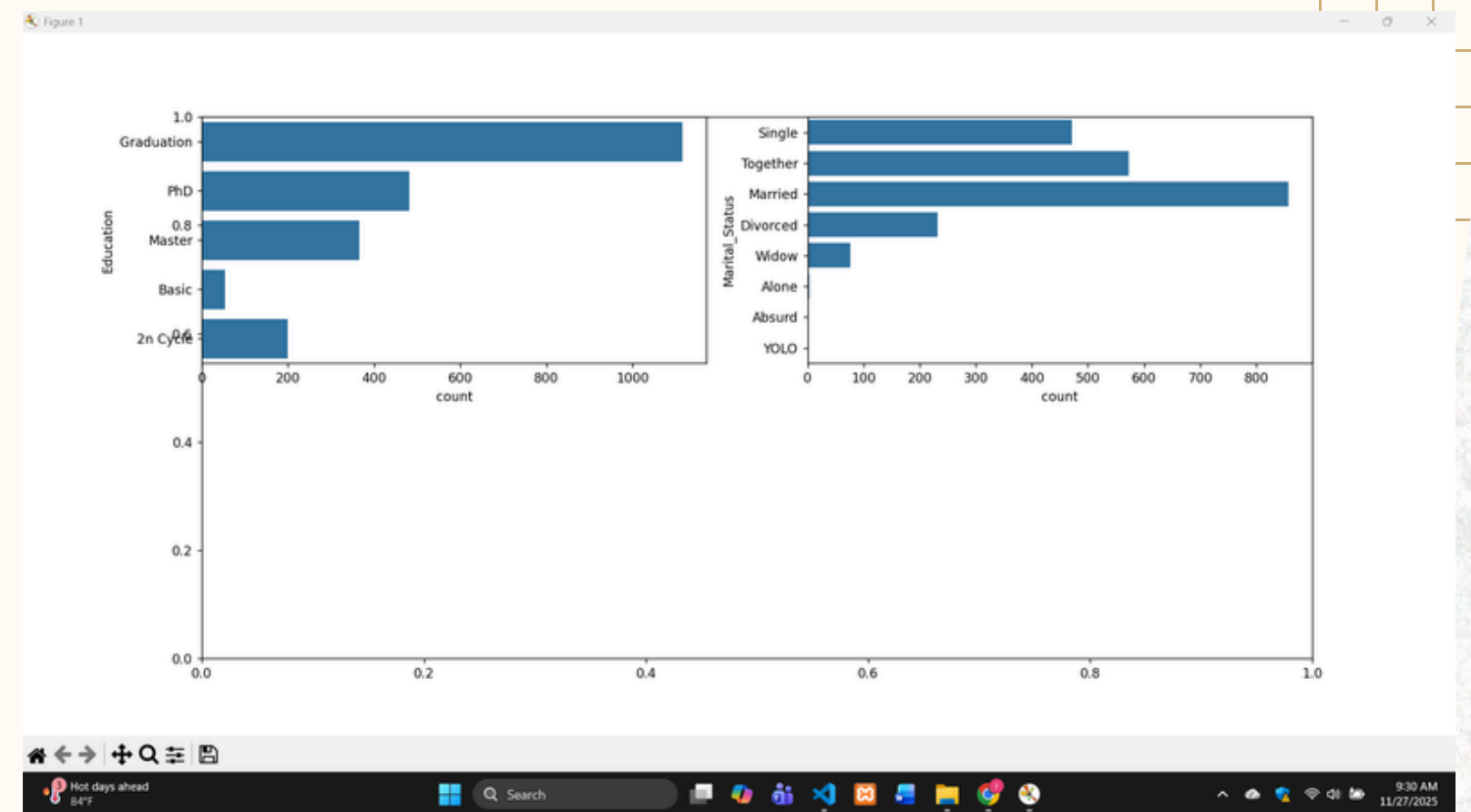
Menganalisis dan mengelompokkan data berdasarkan kemiripan karakteristik menggunakan metode unsupervised learning.





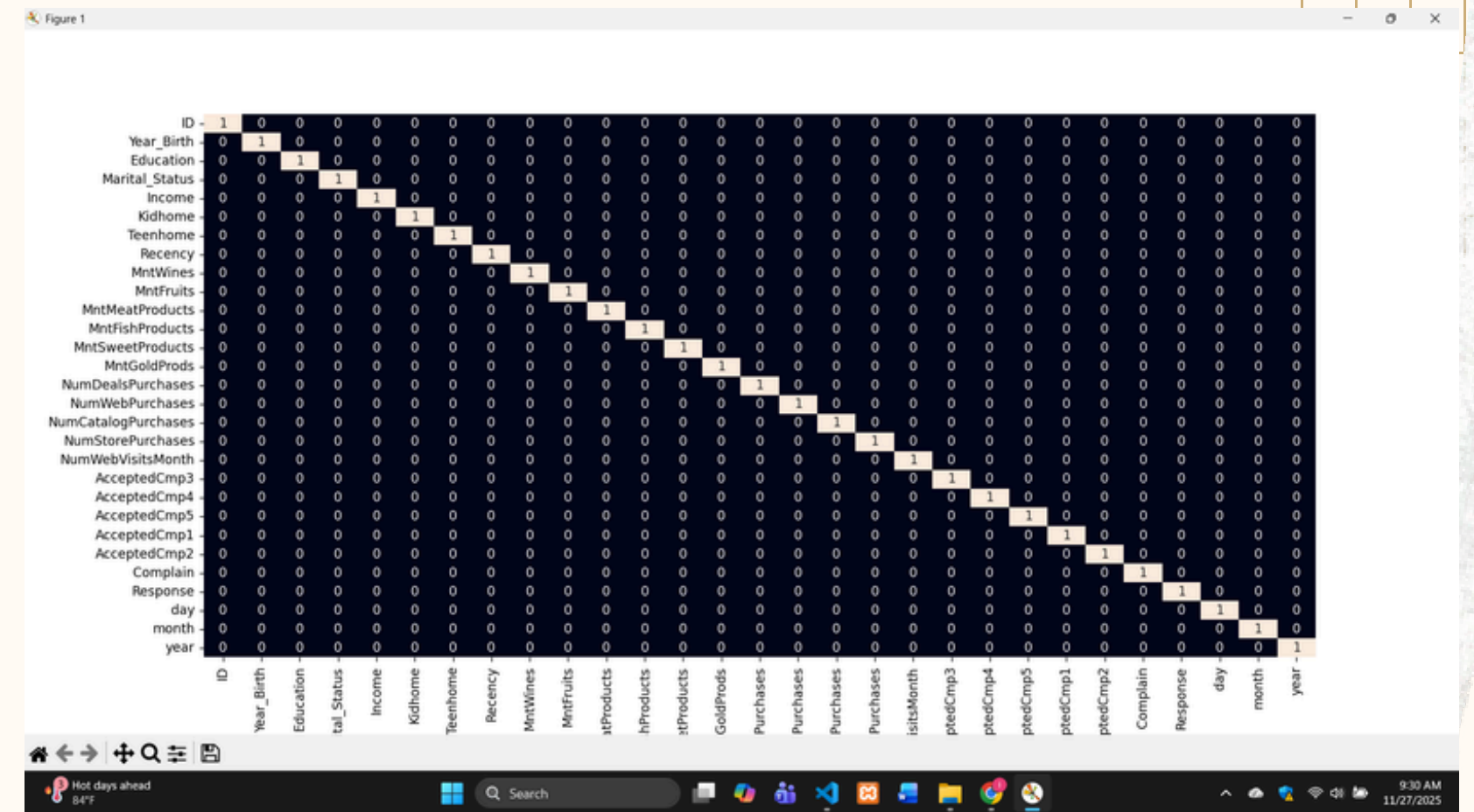
# Output 1

- Gambar ini menampilkan visualisasi distribusi data untuk atribut yang bersifat kategorikal menggunakan countplot. Tujuan visualisasi ini adalah untuk mengetahui jumlah kemunculan setiap kategori dalam dataset. Dari grafik ini dapat dilihat kategori mana yang paling dominan, sehingga membantu dalam memahami karakteristik awal data sebelum dilakukan proses analisis lebih lanjut.



# Output 2

- Gambar heatmap korelasi menunjukkan hubungan antar atribut numerik dalam dataset. Nilai korelasi yang tinggi (lebih dari 0,8) ditandai dengan warna tertentu, yang mengindikasikan adanya keterkaitan kuat antar variabel. Visualisasi ini digunakan untuk mengidentifikasi atribut yang memiliki hubungan erat dan berpotensi redundan dalam proses clustering.

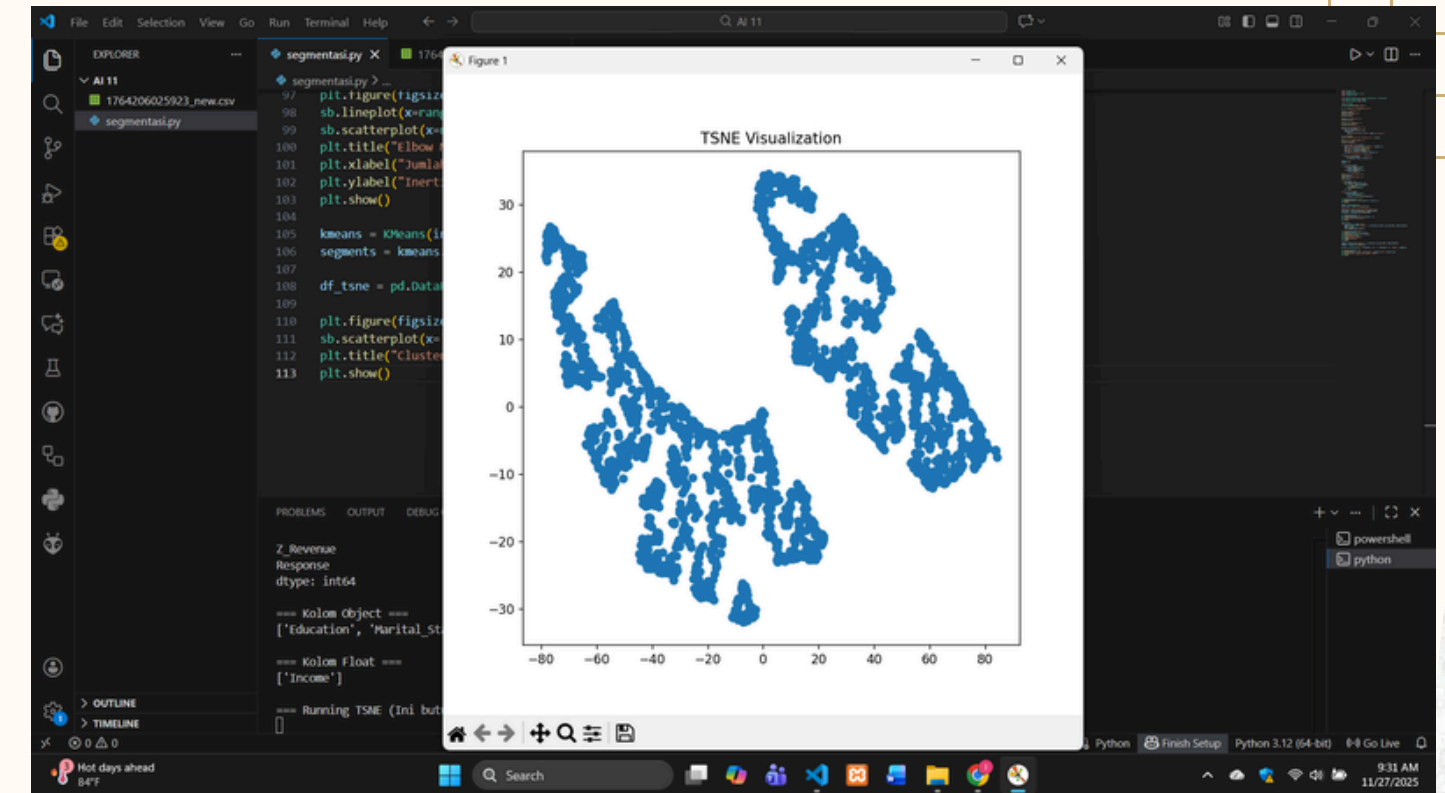




# Output 3



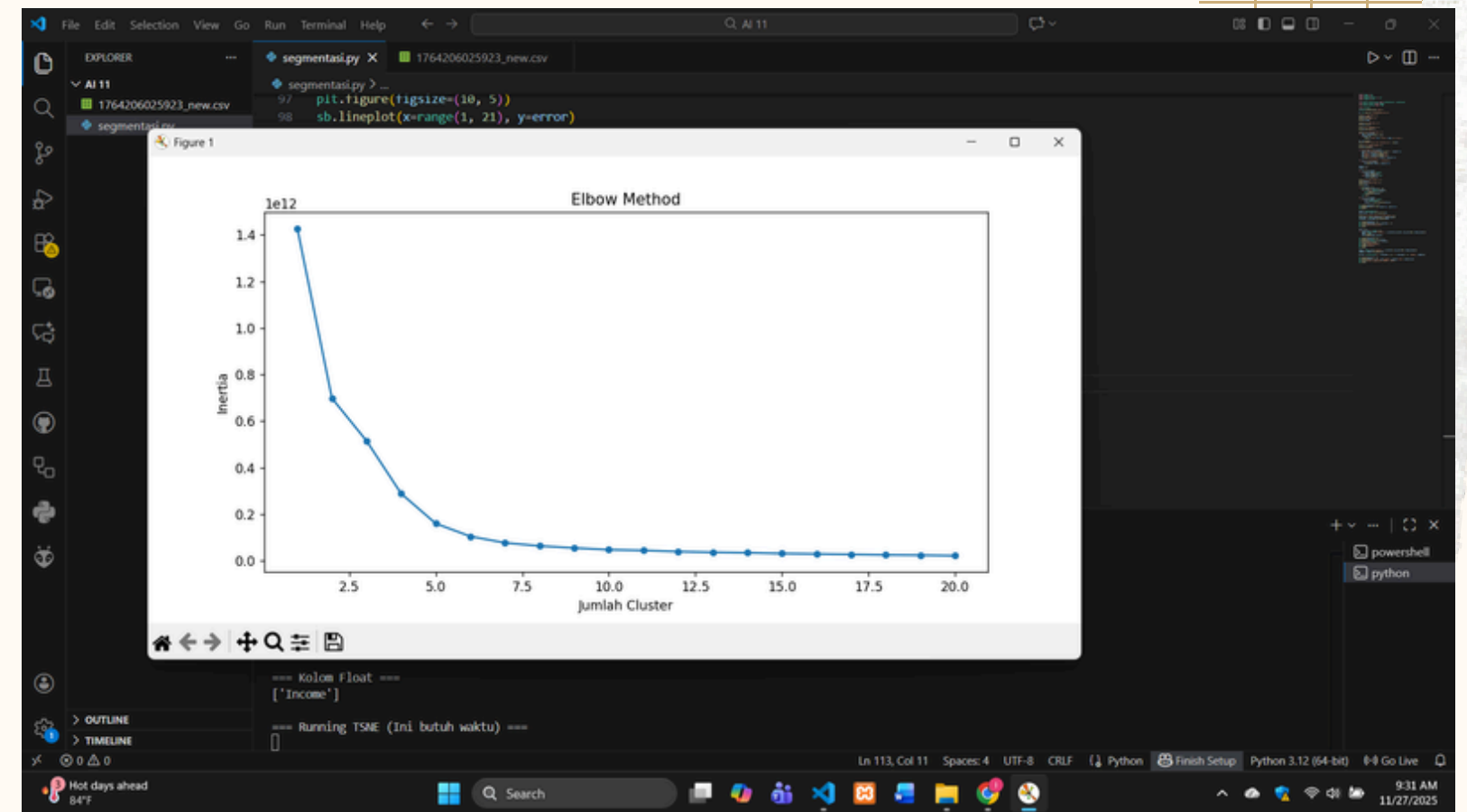
- Gambar ini merupakan hasil visualisasi menggunakan metode TSNE yang berfungsi untuk mereduksi data berdimensi tinggi menjadi dua dimensi. Visualisasi ini membantu dalam melihat pola penyebaran data serta kemungkinan terbentuknya kelompok secara visual sebelum diterapkan algoritma clustering.



# Output 4



- Gambar Elbow Method menunjukkan hubungan antara jumlah cluster dengan nilai inertia. Titik siku pada grafik mengindikasikan jumlah cluster optimal, yaitu kondisi ketika penurunan nilai inertia mulai melambat. Berdasarkan grafik ini, jumlah cluster yang dipilih untuk proses clustering adalah sebanyak lima cluster.

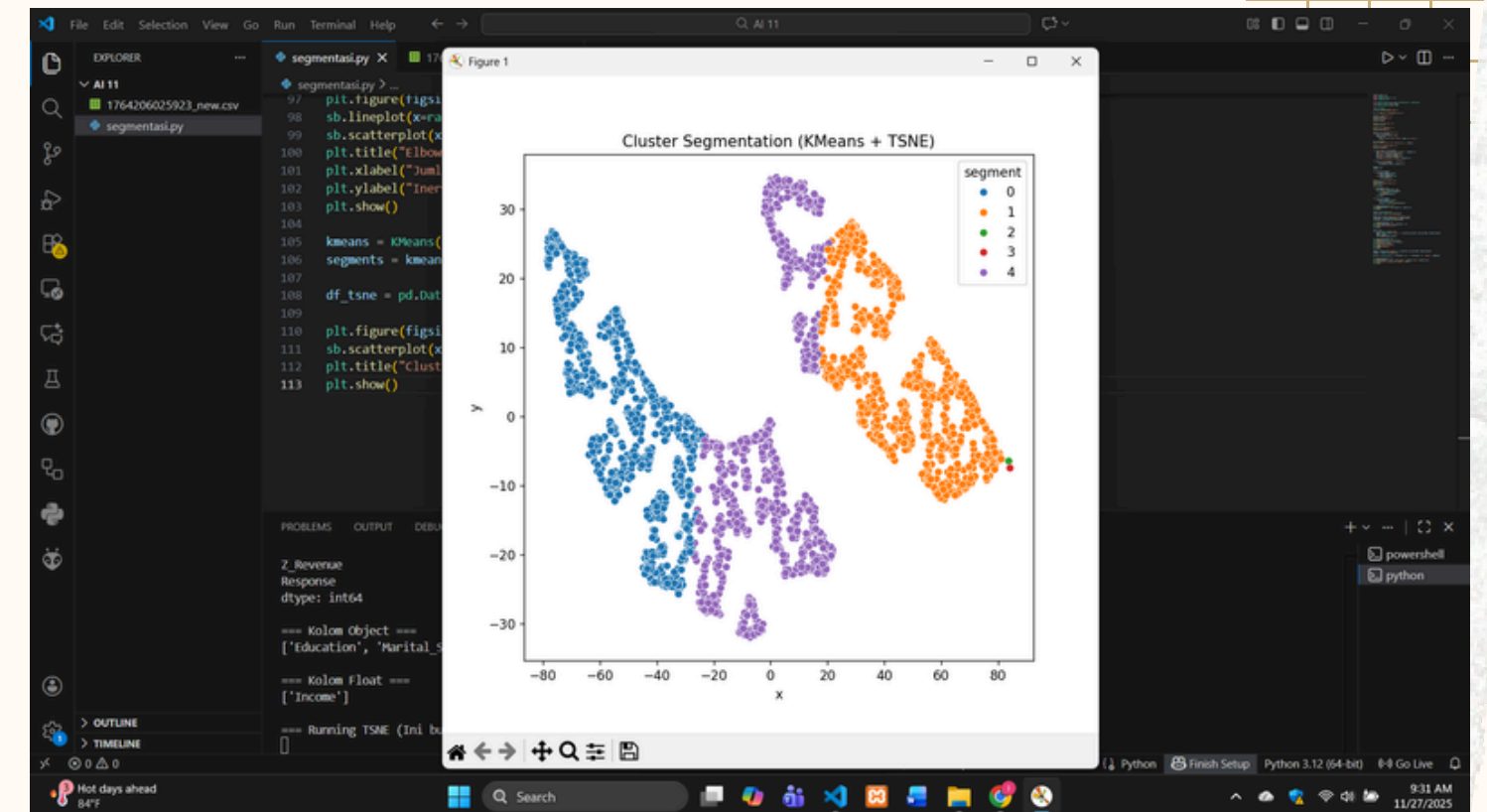




# Output 5



- Gambar ini menampilkan hasil akhir segmentasi data menggunakan algoritma K-Means dengan jumlah cluster sebanyak lima. Setiap warna pada grafik merepresentasikan satu cluster yang berbeda. Visualisasi ini menunjukkan bahwa data berhasil dikelompokkan berdasarkan kemiripan karakteristik antar data.





# Kesimpulan

- Proyek ini berhasil menerapkan metode unsupervised learning menggunakan algoritma K-Means Clustering untuk melakukan segmentasi data.
- Proses preprocessing data seperti pembersihan data, encoding, dan standardisasi terbukti penting untuk meningkatkan kualitas hasil clustering.
- Metode Elbow digunakan secara efektif untuk menentukan jumlah cluster optimal, yaitu sebanyak 5 cluster.
- Visualisasi TSNE membantu dalam memahami pola dan sebaran data berdimensi tinggi secara lebih intuitif.
- Hasil akhir menunjukkan bahwa data dapat dikelompokkan berdasarkan kemiripan karakteristik, sehingga memberikan wawasan yang lebih jelas terhadap struktur data.