

CLUSTER ANALYSIS

- Gunawan, I., 2016, Pengantar Statistika Inferensial, Rajawali Press
- Kuntoro, et al., 2011, Analisis data dengan SPSS, Departemen Biostatistika & Kependudukan Fakultas Kesehatan Masyarakat Universitas Airlangga, Surabaya
- Trihendradi, C., 2010, Step by step SPSS 18 Analisis Data Statistik, Andi Press



CLUSTERING

PROSES PENGELOMPOKAN, SEHINGGA SEMUA ANGGOTA KELOMPOK DARI SETIAP PARTISI MEMILIKI PERSAMAAN BERDASARKAN MATRIKS TERTENTU

CLUSTER ANALYSIS

TEKNIK ANALISIS DATA YANG BERTUJUAN UNTUK MENGELOMPOKKAN INDIVIDU/OBYEK KE DALAM BEBERAPA KELOMPOK YANG MEMILIKI SIFAT BERBEDA, SEHINGGA INDIVIDU/OBYEK YANG TERLETAK DALAM 1 KELOMPOK MEMILIKI SIFAT RELATIF HOMOGEN



METODA CLUSTER ANALYSIS

HIERARCHICAL CLUSTERING METHODS

HASIL PENGELOMPOKANNYA DISAJIKAN SECARA BERJENJANG DARI N , $N-1$, SAMPAI 1 CLUSTER. DARI YANG PALING MIRIP SAMPAI YANG PALING TIDAK MIRIP

NON HIERARCHICAL CLUSTERING METHODS

BERTUJUAN UNTUK MENGELOMPOKKAN N OBYEK KE DALAM K KELOMPOK ($K < N$). OBYEK DIKELOMPOKKAN SEDEMIKIAN RUPA SEHINGGA JARAK TIAP-TIAP OBYEK KE PUSAT KELOMPOK DI DALAM SATU KELOMPOK ADALAH MINIMUM. (K-MEANS CLUSTER)

HIERARCHICAL CLUSTER

Metoda hierarchical cluster menggunakan hirarkhi yang didasari struktur pohon. Proses dimulai dengan menggabungkan 2 obyek yang paling mirip, lalu gabungan tsb akan bergabung lagi dengan 1 atau lebih obyek yang paling mirip lainnya, dst.

Latihan 1: Peneliti akan melakukan identifikasi terhadap berbagai merk susu bubuk yang dijual di pasaran

Variabel:

- Merk susu
- Kandungan lemak (g/100 gr susu)
- Karbohidrat (g/100gr susu)
- Mineral (g/100 gr susu)
- Energi (kilokalori/100 gr susu)

| | MerkSusu | Lemak | Karbohidrat | Mineral | Energi |
|----|-------------------------|--------|-------------|---------|--------|
| 1 | Oat Quaker | 11.00 | 57.00 | .96 | 370.00 |
| 2 | Nestle Carnation | 6.00 | 9.70 | 1.60 | 119.00 |
| 3 | Logi Havermout | 9.16 | 67.33 | .00 | 386.67 |
| 4 | Ovaltine | 8.40 | 75.00 | 1.28 | 416.00 |
| 5 | Milo | 10.00 | 66.00 | 4.50 | 400.00 |
| 6 | Dancow Balita | 23.67 | 44.67 | 4.67 | 476.67 |
| 7 | Frisian Flag Instant | 26.00 | 40.00 | 5.80 | 496.00 |
| 8 | Frisian Flag Full Cream | 28.00 | 38.40 | 5.60 | 506.00 |
| 9 | Frisian Flag Coklat | 13.00 | 69.70 | 3.00 | 444.00 |
| 10 | Frisian Flag Madu | 18.00 | 52.80 | 5.00 | 459.00 |
| 11 | Dancow Coklat | 110.00 | 65.25 | 4.50 | 425.00 |
| 12 | Dancow Full Cream | 25.90 | 40.33 | 5.92 | 495.00 |
| 13 | Indomilk Full Cream | 28.00 | 36.90 | 5.60 | 505.60 |
| 14 | Indomilk Coklat | 14.00 | 62.50 | 5.00 | 446.50 |
| 15 | Presteo Rendah Lemak | 1.00 | 50.00 | 8.00 | 357.00 |
| 16 | Anlene Kalsium Tinggi | .80 | 48.80 | 5.70 | 340.00 |
| 17 | Tropicana Slim | .00 | 50.50 | 1.92 | 347.00 |
| 18 | Protifar | 1.00 | 27.80 | 7.20 | 366.00 |

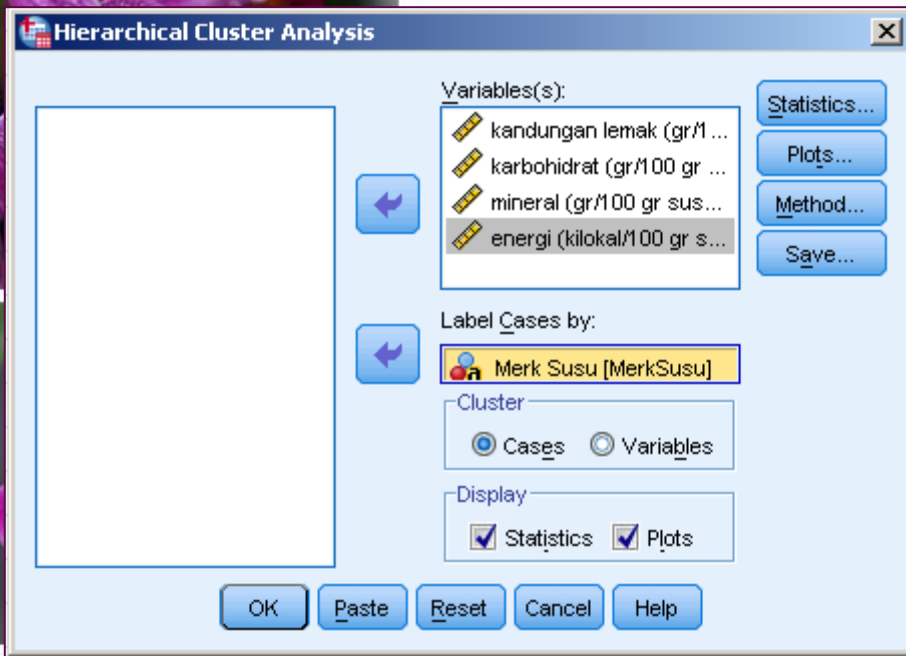
| | Name | Type | Width | Decimals | Label | Values | Missing | Columns | Align | Measure | Role |
|---|-------------|---------|-------|----------|----------------------------------|--------|---------|---------|---------|-----------|---------|
| 1 | MerkSusu | String | 23 | 0 | Merk Susu | None | None | 23 | ≡ Left | ● Nominal | ↘ Input |
| 2 | Lemak | Numeric | 12 | 2 | Kandungan lemak (gr/100 gr susu) | None | None | 12 | ≡ Right | ▨ Scale | ↘ Input |
| 3 | Karbohidrat | Numeric | 12 | 2 | Karbohidrat (gr/100 gr susu) | None | None | 12 | ≡ Right | ▨ Scale | ↘ Input |
| 4 | Mineral | Numeric | 12 | 2 | Mineral (gr/100 gr susu) | None | None | 12 | ≡ Right | ▨ Scale | ↘ Input |
| 5 | Energi | Numeric | 12 | 2 | Energi (kilo kalori/100 gr susu) | None | None | 12 | ≡ Right | ▨ Scale | ↘ Input |

Langkah-langkah:

1. Transformasi data (standarisasi dengan Z-Score)

Bila satuan data sangat bervariasi, akan menyebabkan bias, sehingga data asli harus ditransformasi dengan cara distandarisasi sebelum dilakukan analisis. Transformasi dilakukan terhadap variabel ke bentuk *z-score*. Pada *Hierarchical Cluster* transformasi bisa dilakukan langsung pada *Method* di menu utama Cluster

- **Analyze => Classify => Hierarchical Cluster...**
- Masukkan semua variabel interval-rasio ke kotak **Variable(s): lemak, karbohidrat, mineral dan energi**
- Masukkan variabel **susu** ke dalam **Label Cases by:**
- Pada **Cluster** aktifkan **Cases**
- Pada **Display** aktifkan **Statistics** dan **Plots**

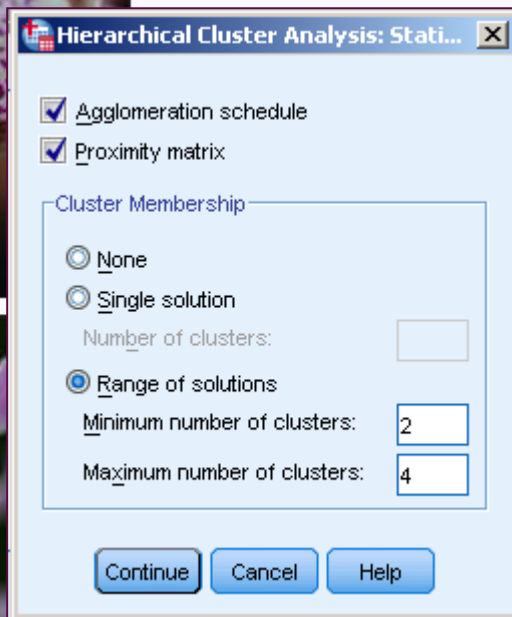


- Klik kotak **Statistics**

- Aktifkan **Agglomeration schedule** dan **proximity matrix** (menampilkan jarak antar variabel)

- Pada **Cluster Membership**, klik pada **Range of solution**, isi bilangan 2 pada **Minimum number of clusters** dan 4 pada **Maximum number of cluster**

- Klik **Continue**

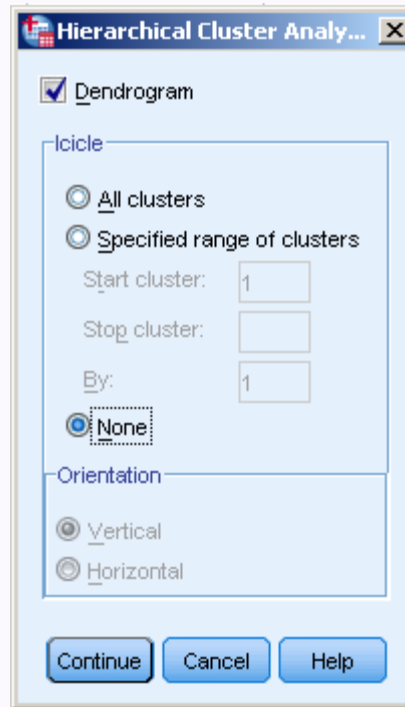


- Klik kotak **Plots...**

- Aktifkan **Dendrogram**

- Pada **Icicle** aktifkan **None**

- Klik **Continue**

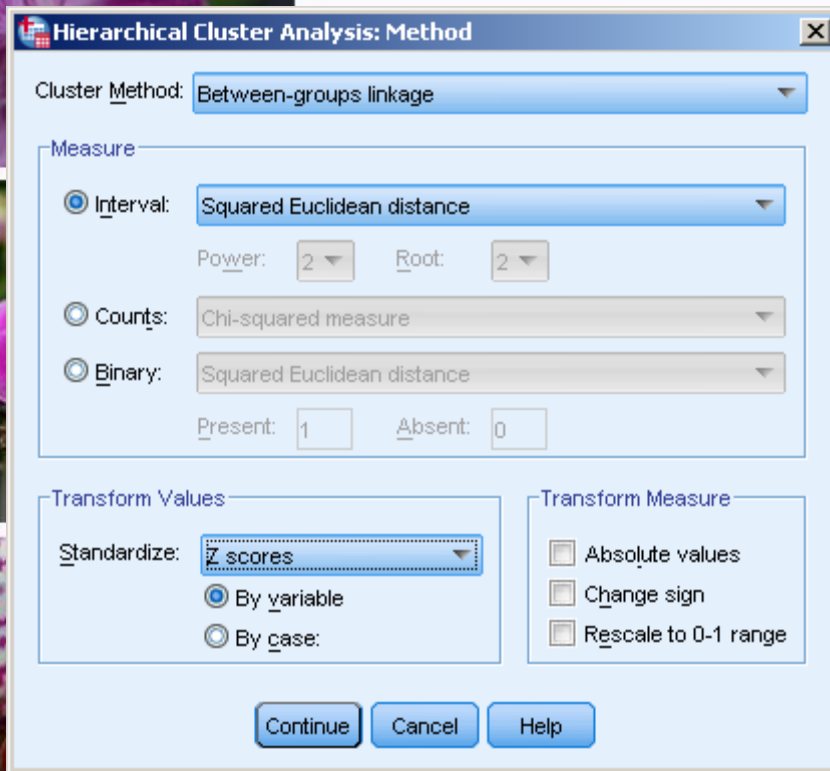


- Klik kotak **Method...**

- Pada **Cluster Method**: pilih **Between-groups linkage**

- Klik kotak **Method...**

- Pada **Cluster Method**: pilih **Between-groups linkage**



- Klik tampilan menu utama **Cluster**, klik **OK**

2. Output

Cluster

Proximity Matrix

| Case | Squared Euclidean Distance | | | | | | | | | | | | | | | | | |
|---------------------------|----------------------------|--------------------|------------------|------------|--------|-----------------|------------------------|---------------------------|-----------------------|----------------------|------------------|----------------------|------------------------|--------------------|------------------------|--------------------------|-------------------|-------------|
| | 1:Oat Quaker | 2:Nestle Carnation | 3:Ligo Havermout | 4:Ovaltine | 5:Milo | 6:Dancow Balita | 7:Frisian Flag Instant | 8:Frisian Flag Full Cream | 9:Frisian Flag Coklat | 10:Frisian Flag Madu | 11:Dancow Coklat | 12:Dancow Full Cream | 13:Indomilk Full Cream | 14:Indomilk Coklat | 15:Presto Rendah Lemak | 16:Anlene Kalsium Tinggi | 17:Tropicana Slim | 18:Protifar |
| 1:Oat Quaker | .000 | 15.766 | .608 | 1.464 | 2.857 | 4.861 | 7.891 | 8.142 | 2.060 | 4.280 | 18.954 | 8.046 | 8.341 | 4.015 | 10.062 | 4.919 | .594 | 10.895 |
| 2:Nestle Carnation | 15.766 | .000 | 21.143 | 26.061 | 22.614 | 22.011 | 24.337 | 24.716 | 26.142 | 23.006 | 41.565 | 24.515 | 24.373 | 25.278 | 20.712 | 14.713 | 12.326 | 14.623 |
| 3:Ligo Havermout | .608 | 21.143 | .000 | .638 | 3.993 | 7.445 | 11.189 | 11.458 | 2.197 | 6.412 | 20.644 | 11.368 | 11.772 | 5.442 | 13.835 | 7.984 | 2.078 | 16.005 |
| 4:Ovaltine | 1.464 | 26.061 | .638 | .000 | 2.360 | 6.420 | 9.729 | 10.124 | .809 | 4.875 | 19.125 | 9.836 | 10.524 | 3.440 | 11.621 | 7.108 | 2.949 | 15.371 |
| 5:Milo | 2.857 | 22.614 | 3.993 | 2.360 | .000 | 2.666 | 4.306 | 4.875 | .735 | 1.202 | 16.293 | 4.280 | 5.176 | .376 | 3.683 | 1.926 | 2.675 | 7.018 |
| 6:Dancow Balita | 4.861 | 22.011 | 7.445 | 6.420 | 2.666 | .000 | .383 | .445 | 3.142 | .352 | 13.953 | .423 | .519 | 1.441 | 4.809 | 3.337 | 4.511 | 4.580 |
| 7:Frisian Flag Instant | 7.891 | 24.337 | 11.189 | 9.729 | 4.306 | .383 | .000 | .036 | 5.347 | .989 | 14.697 | .003 | .060 | 2.496 | 4.621 | 4.206 | 7.083 | 3.948 |
| 8:Frisian Flag Full Cream | 8.142 | 24.716 | 11.458 | 10.124 | 4.875 | .445 | .036 | .000 | 5.718 | 1.251 | 14.549 | .055 | .008 | 2.927 | 5.438 | 4.870 | 7.460 | 4.422 |
| 9:Frisian Flag Coklat | 2.060 | 26.142 | 2.197 | .809 | .735 | 3.142 | 5.347 | 5.718 | .000 | 1.892 | 15.814 | 5.394 | 6.063 | .974 | 7.442 | 4.547 | 2.965 | 10.812 |
| 10:Frisian Flag Madu | 4.280 | 23.006 | 6.412 | 4.875 | 1.202 | .352 | .989 | 1.251 | 1.892 | .000 | 14.477 | .988 | 1.413 | .388 | 3.495 | 2.317 | 3.892 | 4.723 |
| 11:Dancow Coklat | 18.954 | 41.565 | 20.644 | 19.125 | 16.293 | 13.953 | 14.697 | 14.549 | 15.814 | 14.477 | .000 | 14.711 | 14.844 | 15.076 | 23.062 | 21.465 | 22.441 | 26.222 |
| 12:Dancow Full Cream | 8.046 | 24.515 | 11.368 | 9.836 | 4.280 | .423 | .003 | .055 | 5.394 | .988 | 14.711 | .000 | .083 | 2.467 | 4.456 | 4.147 | 7.199 | 3.876 |
| 13:Indomilk Full Cream | 8.341 | 24.373 | 11.772 | 10.524 | 5.176 | .519 | .060 | .008 | 6.063 | 1.413 | 14.844 | .083 | .000 | 3.193 | 5.559 | 4.977 | 7.585 | 4.301 |
| 14:Indomilk Coklat | 4.015 | 25.278 | 5.442 | 3.440 | .376 | 1.441 | 2.496 | 2.927 | .974 | .388 | 15.076 | 2.467 | 3.193 | .000 | 3.557 | 2.411 | 3.876 | 6.383 |
| 15:Presto Rendah Lemak | 10.062 | 20.712 | 13.835 | 11.621 | 3.683 | 4.809 | 4.621 | 5.438 | 7.442 | 3.495 | 23.062 | 4.456 | 5.559 | 3.557 | .000 | 1.075 | 7.250 | 1.932 |
| 16:Anlene Kalsium Tinggi | 4.919 | 14.713 | 7.984 | 7.108 | 1.926 | 3.337 | 4.206 | 4.870 | 4.547 | 2.317 | 21.465 | 4.147 | 4.977 | 2.411 | 1.075 | .000 | 2.814 | 2.129 |
| 17:Tropicana Slim | .594 | 12.326 | 2.078 | 2.949 | 2.675 | 4.511 | 7.083 | 7.460 | 2.965 | 3.892 | 22.441 | 7.199 | 7.585 | 3.876 | 7.250 | 2.814 | .000 | 7.381 |
| 18:Protifar | 10.895 | 14.623 | 16.005 | 15.371 | 7.018 | 4.580 | 3.948 | 4.422 | 10.812 | 4.723 | 26.222 | 3.876 | 4.301 | 6.383 | 1.932 | 2.129 | 7.381 | .000 |

This is a dissimilarity matrix

Contoh:

- jarak antara OAT Quaker dengan Nestle Carnation (=15,766) lebih besar daripada jarak antara OAT Quaker dengan LIGO Havermout (0,608).
- Artinya karakteristik susu OAT Quaker lebih mirip dengan susu Ligodari pada dengan Nestle Carnation.

Average Linkage (Between Groups)

Agglomeration Schedule

| Stage | Cluster Combined | | Coefficients | Stage Cluster First Appears | | Next Stage |
|-------|------------------|-----------|--------------|-----------------------------|-----------|------------|
| | Cluster 1 | Cluster 2 | | Cluster 1 | Cluster 2 | |
| 1 | 7 | 12 | .003 | 0 | 0 | 3 |
| 2 | 8 | 13 | .008 | 0 | 0 | 3 |
| 3 | 7 | 8 | .059 | 1 | 2 | 8 |
| 4 | 6 | 10 | .352 | 0 | 0 | 8 |
| 5 | 5 | 14 | .376 | 0 | 0 | 9 |
| 6 | 1 | 17 | .594 | 0 | 0 | 11 |
| 7 | 3 | 4 | .638 | 0 | 0 | 11 |
| 8 | 6 | 7 | .801 | 4 | 3 | 14 |
| 9 | 5 | 9 | .855 | 5 | 0 | 13 |
| 10 | 15 | 16 | 1.075 | 0 | 0 | 12 |
| 11 | 1 | 3 | 1.775 | 6 | 7 | 13 |
| 12 | 15 | 18 | 2.031 | 10 | 0 | 14 |
| 13 | 1 | 5 | 3.057 | 11 | 9 | 15 |
| 14 | 6 | 15 | 4.338 | 8 | 12 | 15 |
| 15 | 1 | 6 | 6.610 | 13 | 14 | 16 |
| 16 | 1 | 11 | 17.896 | 15 | 0 | 17 |
| 17 | 1 | 2 | 22.582 | 16 | 0 | 0 |

Cluster Membership

| Case | 4 Clusters | 3 Clusters | 2 Clusters |
|---------------------------|------------|------------|------------|
| 1:Oat Quaker | 1 | 1 | 1 |
| 2:Nestle Carnation | 2 | 2 | 2 |
| 3:Ligo Havermout | 1 | 1 | 1 |
| 4:Ovaltine | 1 | 1 | 1 |
| 5:Milo | 1 | 1 | 1 |
| 6:Dancow Balita | 3 | 1 | 1 |
| 7:Frisian Flag Instant | 3 | 1 | 1 |
| 8:Frisian Flag Full Cream | 3 | 1 | 1 |
| 9:Frisian Flag Coklat | 1 | 1 | 1 |
| 10:Frisian Flag Madu | 3 | 1 | 1 |
| 11:Dancow Coklat | 4 | 3 | 1 |
| 12:Dancow Full Cream | 3 | 1 | 1 |
| 13:Indomilk Full Cream | 3 | 1 | 1 |
| 14:Indomilk Coklat | 1 | 1 | 1 |
| 15:Presteo Rendah Lemak | 3 | 1 | 1 |
| 16:Anlene Kalsium Tinggi | 3 | 1 | 1 |
| 17:Tropicana Slim | 1 | 1 | 1 |
| 18:Protifar | 3 | 1 | 1 |

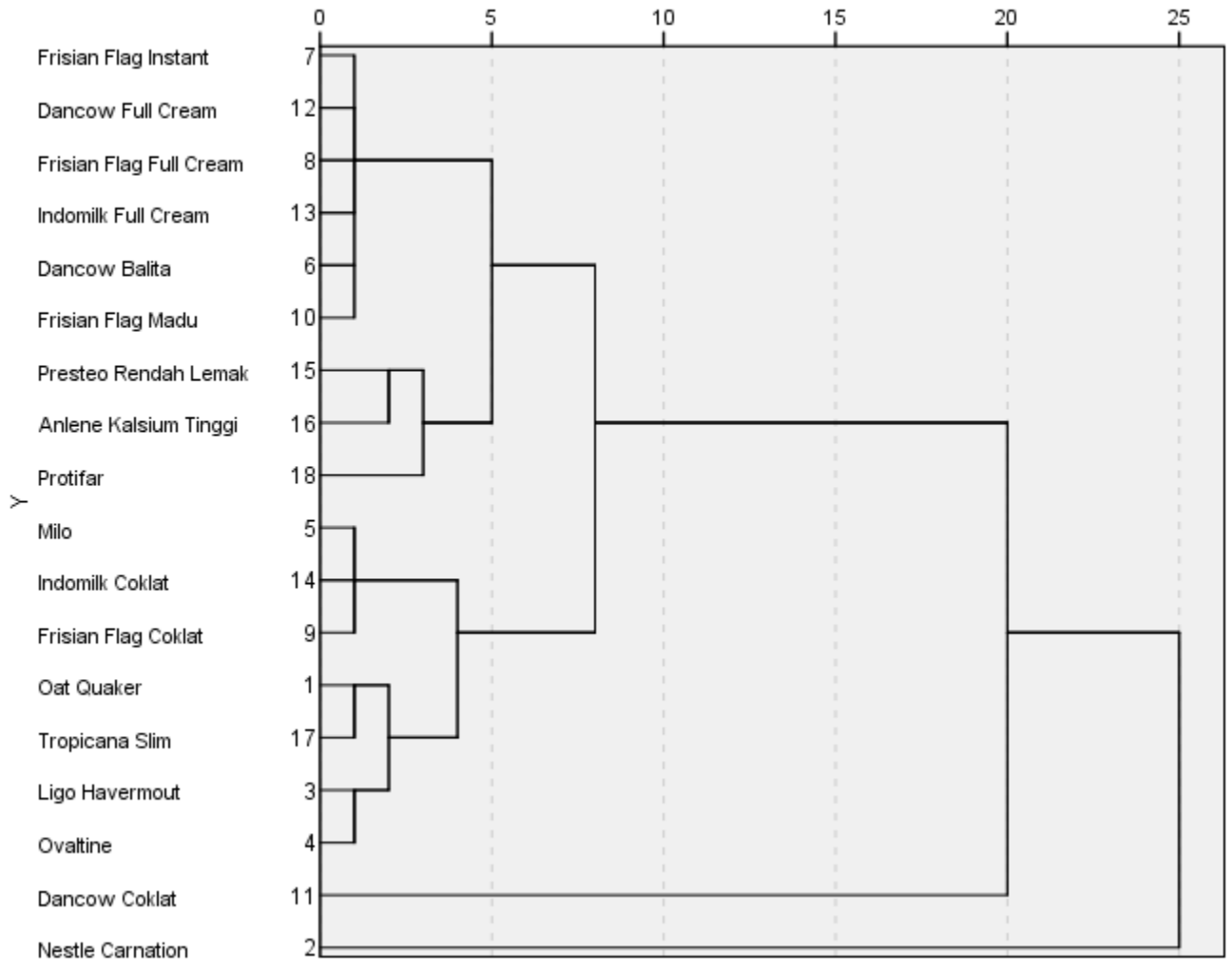
Jika ditentukan 4 cluster:

- Anggota cluster 1 : OAT, Ligo, Ovaltine, Milo, Frisian Flag Coklat, Indomilk coklat dan Tropicana Slim.
- Anggota cluster 2 : Nestle
- Anggota cluster 3: Dancow Balita, Frisian Flag Instan, Frisian Flag Full Cream, Frisian Flag Madu, Dancow Full Cream, Indomilk Full Cream, Prosteo, Anlene dan Protifar.
- Anggota cluster 4: Dancow Coklat



Dendrogram using Average Linkage (Between Groups)

Rescaled Distance Cluster Combine



K-MEANS CLUSTER

Metoda K-Means Cluster memproses semua obyek secara sekaligus. Proses dimulai dengan menentukan jumlah cluster terlebih dahulu

Latihan 2: Peneliti akan mengelompokkan pelanggan rumah sakit berdasarkan profil pelanggan, yaitu:

- Usia (tahun)
- Jumlah anak (orang)
- Penghasilan (Rp/bulan)
- Kegiatan membaca koran (jam/minggu)
- kegiatan menonton TV (jam/minggu)
- Jumlah motor yang dimiliki (buah)
- Jumlah mobil yang dimiliki (buah)
- Jumlah kartu kredit yang dimiliki (buah)
- Tingkat pembelian barang (frekuensi/minggu)
- Konsumsi membeli barang (Rp./bulan)
- Jumlah jam kerja(jam/minggu)
- Jumlah kegiatan belanja (jam/minggu)

| | Kotatempat inggal | Daerah | Status | Usia | Anak | Income | Koran | TV | Motor | Mobil | Kartu kredit | Beli | Konsumsi | Kerja | Shopping |
|----|----------------------|--------|--------|------|------|---------|-------|----|-------|-------|-----------------|------|----------|-------|----------|
| 1 | Jakarta utar | 1 | 1 | 25 | 0 | 2000000 | 10 | 20 | 1 | 1 | 3 | 4 | 600000 | 40 | 10 |
| 2 | Bandung | 1 | 2 | 26 | 0 | 750000 | 11 | 18 | 2 | 1 | 3 | 4 | 225000 | 40 | 12 |
| 3 | Tegal | 3 | 3 | 25 | 2 | 300000 | 5 | 16 | 0 | 0 | 1 | 2 | 90000 | 15 | 7 |
| 4 | Semarang | 1 | 3 | 40 | 2 | 750000 | 9 | 15 | 2 | 1 | 4 | 6 | 225000 | 55 | 20 |
| 5 | Yogya | 2 | 2 | 35 | 0 | 500000 | 7 | 11 | 1 | 1 | 2 | 3 | 150000 | 30 | 10 |
| 6 | Solo | 2 | 2 | 30 | 0 | 450000 | 6 | 14 | 1 | 0 | 2 | 3 | 135000 | 30 | 9 |
| 7 | Banjame... | 3 | 2 | 25 | 0 | 250000 | 6 | 17 | 0 | 0 | 1 | 2 | 75000 | 15 | 5 |
| 8 | Madiun | 2 | 2 | 24 | 0 | 400000 | 5 | 20 | 1 | 1 | 2 | 2 | 120000 | 30 | 11 |
| 9 | Pekalongan | 2 | 3 | 26 | 2 | 450000 | 5 | 19 | 1 | 0 | 3 | 2 | 135000 | 45 | 14 |
| 10 | Jepara | 3 | 3 | 21 | 3 | 300000 | 3 | 15 | 1 | 0 | 1 | 1 | 90000 | 14 | 5 |
| 11 | Blora | 3 | 1 | 20 | 0 | 350000 | 4 | 11 | 1 | 0 | 0 | 1 | 100000 | 10 | 6 |
| 12 | Kerawang | 3 | 2 | 35 | 0 | 350000 | 3 | 14 | 1 | 0 | 1 | 1 | 100000 | 15 | 5 |
| 13 | Jakarta s... | 1 | 2 | 30 | 0 | 1500000 | 8 | 16 | 2 | 0 | 5 | 3 | 450000 | 70 | 21 |
| 14 | Bandung | 1 | 2 | 35 | 0 | 1000000 | 9 | 12 | 1 | 1 | 3 | 5 | 300000 | 45 | 11 |
| 15 | Surabaya | 1 | 3 | 42 | 2 | 1500000 | 9 | 10 | 1 | 1 | 4 | 6 | 450000 | 60 | 18 |
| 16 | Magelang | 3 | 3 | 29 | 1 | 350000 | 5 | 11 | 1 | 0 | 1 | 1 | 125000 | 14 | 9 |
| 17 | Parakan | 3 | 1 | 28 | 0 | 300000 | 3 | 17 | 1 | 1 | 1 | 2 | 90000 | 15 | 6 |
| 18 | Tuban | 3 | 2 | 25 | 0 | 250000 | 4 | 12 | 1 | 0 | 1 | 2 | 75000 | 11 | 3 |
| 19 | Jakarta s... | 1 | 3 | 42 | 4 | 1500000 | 5 | 15 | 2 | 1 | 3 | 8 | 450000 | 42 | 14 |
| 20 | Semarang | 1 | 2 | 31 | 0 | 1000000 | 7 | 19 | 2 | 1 | 4 | 9 | 300000 | 60 | 18 |
| 21 | Surabaya | 1 | 1 | 32 | 0 | 950000 | 8 | 20 | 2 | 0 | 4 | 9 | 180000 | 55 | 20 |
| 22 | Surabaya | 1 | 1 | 34 | 0 | 900000 | 5 | 21 | 2 | 1 | 3 | 9 | 270000 | 45 | 15 |
| 23 | Ciamis | 3 | 1 | 40 | 0 | 1000000 | 5 | 22 | 1 | 1 | 1 | 2 | 300000 | 14 | 6 |
| 24 | Pati | 3 | 1 | 29 | 0 | 300000 | 2 | 15 | 2 | 0 | 0 | 2 | 90000 | 12 | 8 |
| 25 | Cepu | 3 | 3 | 27 | 2 | 250000 | 2 | 16 | 1 | 0 | 1 | 3 | 75000 | 15 | 8 |
| 26 | Wonogiri | 3 | 2 | 25 | 0 | 300000 | 3 | 19 | 1 | 0 | 1 | 3 | 90000 | 16 | 8 |



| | Name | Type | Width | Decimals | Label | Values | Missing | Colum... | Align | Measure | R |
|----|---------------|---------|-------|----------|--|---------------|---------|----------|-------|---------|---|
| 1 | Kotatempat... | String | 12 | 0 | Kota tempat tinggal | None | None | 7 | Left | Nominal | |
| 2 | Daerah | Numeric | 12 | 0 | Daerah tempat tinggal | {1, kota b... | None | 5 | Right | Scale | |
| 3 | Status | Numeric | 12 | 0 | Status pernikahan | {1, belum... | None | 4 | Right | Scale | |
| 4 | Usia | Numeric | 12 | 0 | Usia | None | None | 3 | Right | Scale | |
| 5 | Anak | Numeric | 12 | 0 | Jumlah anak | None | None | 4 | Right | Scale | |
| 6 | Income | Numeric | 12 | 0 | Penghasilan rata-rata per bulan | None | None | 6 | Right | Scale | |
| 7 | Koran | Numeric | 12 | 0 | Jumlah jam membaca koran setiap minggu | None | None | 4 | Right | Scale | |
| 8 | TV | Numeric | 12 | 0 | Jumlah jam menonton TV setiap minggu | None | None | 2 | Right | Scale | |
| 9 | Motor | Numeric | 12 | 0 | Jumlah motor yang dimiliki | None | None | 3 | Right | Scale | |
| 10 | Mobil | Numeric | 12 | 0 | Jumlah mobil yang dimiliki | None | None | 3 | Right | Scale | |
| 11 | Kartukredit | Numeric | 12 | 0 | Jumlah kartu kredit atau ATM yang dimiliki | None | None | 3 | Right | Scale | |
| 12 | Beli | Numeric | 12 | 0 | Tingkat pembelian barang setiap minggu | None | None | 3 | Right | Scale | |
| 13 | Konsumsi | Numeric | 12 | 0 | Tingkat pengeluaran bulanan | None | None | 6 | Right | Scale | |
| 14 | Kerja | Numeric | 12 | 0 | Jumlah jam kerja setiap minggu | None | None | 3 | Right | Scale | |
| 15 | Shopping | Numeric | 12 | 0 | Jumlah jam berbelanja setiap minggu | None | None | 6 | Right | Scale | |



Langkah-langkah:

1. Transformasi data (standarisasi dengan Z-Score)

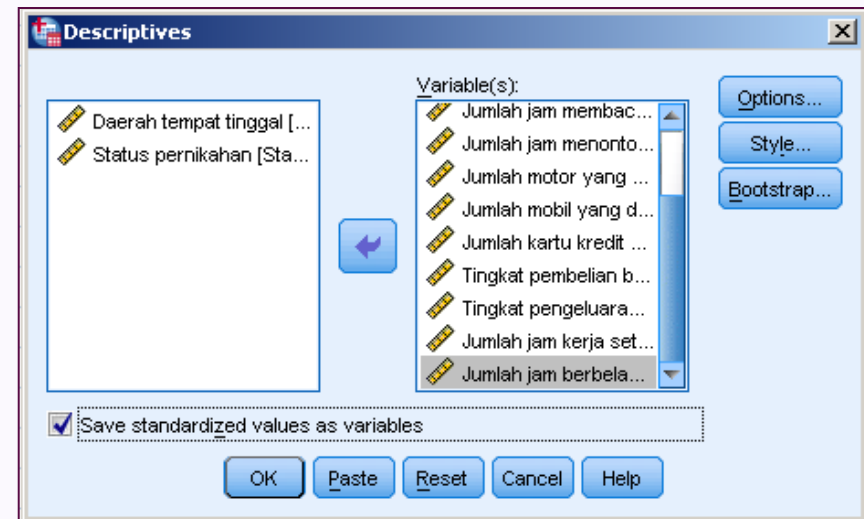
Bila satuan data sangat bervariasi, akan menyebabkan bias, sehingga data asli harus ditransformasi dengan cara distandarisasi sebelum dilakukan analisis. Transformasi dilakukan terhadap variabel ke bentuk z-score.

- **Analyze => Descriptive Statistics=> Descriptive...**

- Masukkan semua variabel interval-rasio ke kotak **Variable(s):** usia, anak, income, koran, tv, motor, mobil, k_kredit, beli, konsumsi, kerja dan shopping.

- Aktifkan (klik pada kotak) pilihan **Save Standardized values as variables**.

- Klik **OK**.



- Muncul 12 variabel baru dengan imbuhan 'Z' di depan setiap variabel (Zusia, Zanak, Zincome, Zkoran, Ztv, Zmotor, Zmobil, Zk_kredit, Zbeli, Zkonsumsi, Zkerja dan Zshopping)

| | Shopping | ZUsia | ZAnak | ZIncome | ZKoran | ZTV | ZMotor | ZMobil | ZKartukre... | ZBeli | ZKonsums i | ZKerja | ZShoppi... |
|----|----------|----------|---------|---------|----------|----------|----------|---------|--------------|----------|---------------|----------|------------|
| 1 | 10 | -.80764 | -.59772 | 2.66988 | 1.69225 | 1.17175 | -.39311 | 1.05915 | .63411 | .13279 | 2.71299 | .46532 | -.13879 |
| 2 | 12 | -.64735 | -.59772 | .12154 | 2.08863 | .59146 | 1.31036 | 1.05915 | .63411 | .13279 | .14736 | .46532 | .24105 |
| 3 | 7 | -.80764 | 1.12903 | -.79587 | -.28966 | .01116 | -2.09657 | -.90784 | -.79953 | -.63444 | -.77627 | -.86710 | -.70854 |
| 4 | 20 | 1.59679 | 1.12903 | .12154 | 1.29586 | -.27899 | 1.31036 | 1.05915 | 1.35093 | .90001 | .14736 | 1.26477 | 1.76040 |
| 5 | 10 | .79531 | -.59772 | -.38813 | .50310 | -1.43958 | -.39311 | 1.05915 | -.08271 | -.25082 | -.36577 | -.06765 | -.13879 |
| 6 | 9 | -.00617 | -.59772 | -.49007 | .10672 | -.56914 | -.39311 | -.90784 | -.08271 | -.25082 | -.46839 | -.06765 | -.32871 |
| 7 | 5 | -.80764 | -.59772 | -.89780 | .10672 | .30131 | -2.09657 | -.90784 | -.79953 | -.63444 | -.87889 | -.86710 | -1.08838 |
| 8 | 11 | -.96794 | -.59772 | -.59200 | -.28966 | 1.17175 | -.39311 | 1.05915 | -.08271 | -.63444 | -.57102 | -.06765 | .05113 |
| 9 | 14 | -.64735 | 1.12903 | -.49007 | -.28966 | .88160 | -.39311 | -.90784 | .63411 | -.63444 | -.46839 | .73181 | .62089 |
| 10 | 5 | -1.44883 | 1.99240 | -.79587 | -1.08243 | -.27899 | -.39311 | -.90784 | -.79953 | -1.01805 | -.77627 | -.92039 | -1.08838 |
| 11 | 6 | -1.60912 | -.59772 | -.69393 | -.68605 | -1.43958 | -.39311 | -.90784 | -1.51635 | -1.01805 | -.70785 | -1.13358 | -.89846 |
| 12 | 5 | .79531 | -.59772 | -.69393 | -1.08243 | -.56914 | -.39311 | -.90784 | -.79953 | -1.01805 | -.70785 | -.86710 | -1.08838 |
| 13 | 21 | -.00617 | -.59772 | 1.65055 | .89948 | .01116 | 1.31036 | -.90784 | 2.06776 | -.25082 | 1.68674 | 2.06423 | 1.95032 |
| 14 | 11 | .79531 | -.59772 | .63121 | 1.29586 | -1.14943 | -.39311 | 1.05915 | .63411 | .51640 | .66049 | .73181 | .05113 |
| 15 | 18 | 1.91738 | 1.12903 | 1.65055 | 1.29586 | -1.72973 | -.39311 | 1.05915 | 1.35093 | .90001 | 1.68674 | 1.53126 | 1.38056 |
| 16 | 9 | -.16646 | .26565 | -.69393 | -.28966 | -1.43958 | -.39311 | -.90784 | -.79953 | -1.01805 | -.53681 | -.92039 | -.32871 |
| 17 | 6 | -.32676 | -.59772 | -.79587 | -1.08243 | .30131 | -.39311 | 1.05915 | -.79953 | -.63444 | -.77627 | -.86710 | -.89846 |
| 18 | 3 | -.80764 | -.59772 | -.89780 | -.68605 | -1.14943 | -.39311 | -.90784 | -.79953 | -.63444 | -.87889 | -1.08029 | -1.46822 |
| 19 | 14 | 1.91738 | 2.85577 | 1.65055 | -.28966 | -.27899 | 1.31036 | 1.05915 | .63411 | 1.66724 | 1.68674 | .57192 | .62089 |
| 20 | 18 | .15413 | -.59772 | .63121 | .50310 | .88160 | 1.31036 | 1.05915 | 1.35093 | 2.05085 | .66049 | 1.53126 | 1.38056 |

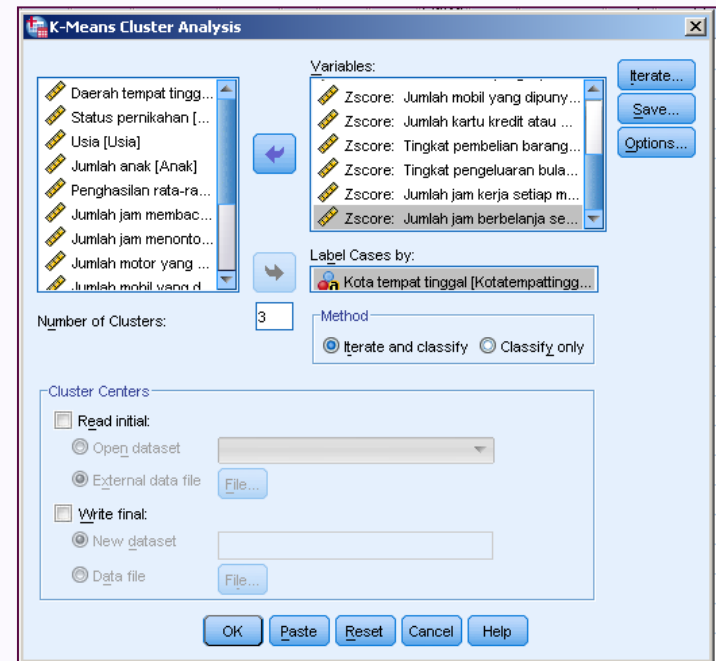
2. Analisis gugus

- **Analyze => Classify => K-Means Cluster**

- Masukkan semua variabel yang sudah distandarisasi (**Zusia, Zanak, Zincome, Zkoran, Ztv, Zmotor, Zmobil, Zk_kredit, Zbeli, Zkonsumsi, Zkerja dan Zshopping**).

- Masukkan variabel **kota tempat tinggal** ke **Label Cases by**:

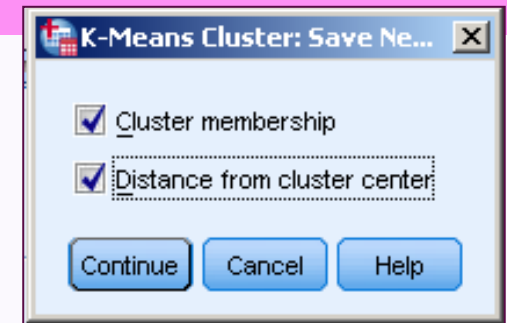
Isi jumlah cluster sebesar 3 yang akan dibentuk ke **Number of Cluster**. Jumlah cluster biasanya 2 – 5, tergantung tujuan penelitian dan faktor subyektif lain dari peneliti



- Klik kotak **Save....**

- Kotak dialog **Save...** memungkinkan hasil analisis cluster disimpan dalam bentuk variabel baru di SPSS Data Editor.

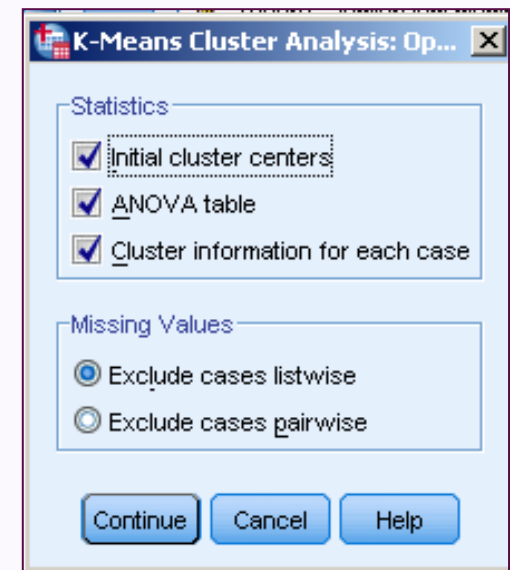
- Aktifkan kotak **Cluster membership** dan **Distance from cluster center**. Klik tombol **Continue**.



- Klik kotak **Options...**

- Pada **Statistics** aktifkan **Initial cluster centers** dan **Anova table**. Klik **Continue**.

- Pada menu utama **Cluster**, klik **OK**



3. Output

- **Output** disimpan dengan nama **K_Means Cluster.spv**

- **Output** Final Cluster Centers menunjukkan bahwa ke-3 cluster memiliki karakteristik yang berbeda.

| | N | Minimum | Maximum | Mean | Std. Deviation |
|--|----|---------|---------|-----------|----------------|
| Usia | 26 | 20 | 42 | 30.04 | 6.238 |
| Jumlah anak | 26 | 0 | 4 | .69 | 1.158 |
| Penghasilan rata-rata per bulan | 26 | 250000 | 2000000 | 690384.62 | 490513.859 |
| Jumlah jam membaca koran setiap minggu | 26 | 2 | 11 | 5.73 | 2.523 |
| Jumlah jam menonton TV setiap minggu | 26 | 10 | 22 | 15.96 | 3.447 |
| Jumlah motor yang dipunyai | 26 | 0 | 2 | 1.23 | .587 |
| Jumlah mobil yang dipunyai | 26 | 0 | 1 | .46 | .508 |
| Jumlah kartu kredit atau ATM yang dimiliki | 26 | 0 | 5 | 2.12 | 1.395 |
| Tingkat pembelian barang setiap minggu | 26 | 1 | 9 | 3.65 | 2.607 |
| Tingkat pengeluaran bulanan | 26 | 75000 | 600000 | 203461.54 | 146162.712 |
| Jumlah jam kerja setiap minggu | 26 | 10 | 70 | 31.27 | 18.763 |
| Jumlah jam berbelanja setiap minggu | 26 | 3 | 21 | 10.73 | 5.265 |
| Valid N (listwise) | 26 | | | | |

Initial Cluster Centers

| | Cluster | | |
|--|---------|---------|----------|
| | 1 | 2 | 3 |
| Zscore: Usia | -.80764 | 1.91738 | -1.60912 |
| Zscore: Jumlah anak | -.59772 | 2.85577 | -.59772 |
| Zscore: Penghasilan rata-rata per bulan | 2.66988 | 1.65055 | -.69393 |
| Zscore: Jumlah jam membaca koran setiap minggu | 1.69225 | -.28966 | -.68605 |
| Zscore: Jumlah jam menonton TV setiap minggu | 1.17175 | -.27899 | -1.43958 |
| Zscore: Jumlah motor yang dimiliki | -.39311 | 1.31036 | -.39311 |
| Zscore: Jumlah mobil yang dimiliki | 1.05915 | 1.05915 | -.90784 |
| Zscore: Jumlah kartu kredit atau ATM yang dimiliki | .63411 | .63411 | -1.51635 |
| Zscore: Tingkat pembelian barang setiap minggu | .13279 | 1.66724 | -1.01805 |
| Zscore: Tingkat pengeluaran bulanan | 2.71299 | 1.68674 | -.70785 |
| Zscore: Jumlah jam kerja setiap minggu | .46532 | .57192 | -1.13358 |
| Zscore: Jumlah jam berbelanja setiap minggu | -.13879 | .62089 | -.89846 |

Iteration History^a

| Iteration | Change in Cluster Centers | | |
|-----------|---------------------------|-------|-------|
| | 1 | 2 | 3 |
| 1 | 2.699 | 3.115 | 2.097 |
| 2 | .000 | .000 | .000 |

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is 5.773.

Final Cluster Centers

| | Cluster | | |
|--|---------|---------|---------|
| | 1 | 2 | 3 |
| Zscore: Usia | .18619 | 1.08919 | -.49774 |
| Zscore: Jumlah anak | -.59772 | .55344 | -.02214 |
| Zscore: Penghasilan rata-rata per bulan | 1.14088 | .83507 | -.71432 |
| Zscore: Jumlah jam membaca koran setiap minggu | 1.13731 | .56916 | -.60677 |
| Zscore: Jumlah jam menonton TV setiap minggu | .47540 | .20459 | -.24030 |
| Zscore: Jumlah motor yang dipunyai | .28828 | 1.02645 | -.50667 |
| Zscore: Jumlah mobil yang dipunyai | .66575 | .73132 | -.51444 |
| Zscore: Jumlah kartu kredit atau ATM yang dimiliki | .63411 | 1.11199 | -.65617 |
| Zscore: Tingkat pembelian barang setiap minggu | -.02066 | 1.60330 | -.63444 |
| Zscore: Tingkat pengeluaran bulanan | 1.17361 | .74601 | -.68961 |
| Zscore: Jumlah jam kerja setiap minggu | .56126 | 1.14930 | -.64680 |
| Zscore: Jumlah jam berbelanja setiap minggu | .24105 | 1.28560 | -.59459 |

Contoh variabel Usia:

- Cluster 1 mempunyai usia $z = 0,18619$,
- usia pada cluster 2 ($z = 1,08919$),
- usia pada cluster 3 ($z = 0,49774$)

Jika $x = \mu + z \cdot \theta$ dan

| Descriptive Statistics | | | | | |
|------------------------|----|---------|---------|-------|----------------|
| | N | Minimum | Maximum | Mean | Std. Deviation |
| Usia | 26 | 20 | 42 | 30.04 | 6.238 |

Maka

- Rerata usia C1 = $30,04 + (0,18619 \cdot 6,238) = 31,2$ th
- Rerata usia C2 = $30,04 + (1,08919 \cdot 6,238) = 36,8$ th
- Rerata usia C3 = $30,04 + (0,49774 \cdot 6,238) = 33,1$ th

Analisis perbedaan dengan Anova, didapatkan bahwa:

- Tingkat pembelian merupakan yang paling berbeda antar cluster (mempunyai F terbesar sebesar 69,790 dengan $p = \text{Sig.} = 0,000 < 0,05$).
- Jumlah jam menonton TV tidak berbeda antar cluster ($F = 1,136$ dengan $p = \text{Sig.} = 0,339 > 0,05$).

ANOVA

| | Cluster | | Error | | F | Sig. |
|--|-------------|----|-------------|----|--------|------|
| | Mean Square | df | Mean Square | df | | |
| Zscore: Usia | 5.504 | 2 | .608 | 23 | 9.047 | .001 |
| Zscore: Jumlah anak | 1.816 | 2 | .929 | 23 | 1.954 | .164 |
| Zscore: Penghasilan rata-rata per bulan | 9.173 | 2 | .289 | 23 | 31.706 | .000 |
| Zscore: Jumlah jam membaca koran setiap minggu | 6.967 | 2 | .481 | 23 | 14.480 | .000 |
| Zscore: Jumlah jam menonton TV setiap minggu | 1.124 | 2 | .989 | 23 | 1.136 | .339 |
| Zscore: Jumlah motor yang dipunyai | 5.294 | 2 | .627 | 23 | 8.448 | .002 |
| Zscore: Jumlah mobil yang dipunyai | 4.697 | 2 | .678 | 23 | 6.923 | .004 |
| Zscore: Jumlah kartu kredit atau ATM yang dimiliki | 7.944 | 2 | .396 | 23 | 20.052 | .000 |
| Zscore: Tingkat pembelian barang setiap minggu | 10.732 | 2 | .154 | 23 | 69.790 | .000 |
| Zscore: Tingkat pengeluaran bulanan | 8.680 | 2 | .332 | 23 | 26.128 | .000 |
| Zscore: Jumlah jam kerja setiap minggu | 7.888 | 2 | .401 | 23 | 19.668 | .000 |
| Zscore: Jumlah jam berbelanja setiap minggu | 7.755 | 2 | .413 | 23 | 18.796 | .000 |

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.



Distances between Final Cluster Centers

| Cluster | 1 | 2 | 3 |
|---------|-------|-------|-------|
| 1 | | 2.769 | 4.182 |
| 2 | 2.769 | | 5.262 |
| 3 | 4.182 | 5.262 | |

Number of Cases in each Cluster

| | | |
|---------|---|--------|
| Cluster | 1 | 5.000 |
| | 2 | 6.000 |
| | 3 | 15.000 |
| Valid | | 26.000 |
| Missing | | .000 |

Cluster Membership

| Case Number | Kota tempat tinggal | Cluster | Distance |
|-------------|---------------------|---------|----------|
| 1 | Jakarta utara | 1 | 2.699 |
| 2 | Bandung | 1 | 2.222 |
| 3 | Tegal | 3 | 2.089 |
| 4 | Semarang | 2 | 1.788 |
| 5 | Yogya | 3 | 2.897 |
| 6 | Solo | 3 | 1.529 |
| 7 | Banjarnegara | 3 | 2.072 |
| 8 | Madiun | 3 | 2.499 |
| 9 | Pekalongan | 3 | 2.831 |
| 10 | Jepara | 3 | 2.421 |
| 11 | Blora | 3 | 2.097 |
| 12 | Kerawang | 3 | 1.722 |
| 13 | Jakarta selata | 1 | 3.412 |
| 14 | Bandung | 1 | 2.129 |
| 15 | Surabaya | 2 | 3.109 |
| 16 | Magelang | 3 | 1.495 |
| 17 | Parakan | 3 | 1.883 |
| 18 | Tuban | 3 | 1.571 |
| 19 | Jakarta selata | 2 | 3.115 |
| 20 | Semarang | 2 | 1.820 |
| 21 | Surabaya | 2 | 2.674 |
| 22 | Surabaya | 2 | 2.262 |
| 23 | Ciamis | 1 | 3.567 |
| 24 | Pati | 3 | 2.359 |
| 25 | Cepu | 3 | 1.615 |
| 26 | Wonogiri | 3 | 1.515 |

Data bisa juga dilihat di **Data View**, kolom terakhir

| | ZBeli | ZKonsumsi | ZKerja | ZShopping | QCL_1 | QCL_2 |
|----|----------|-----------|----------|-----------|-------|---------|
| 1 | .13279 | 2.71299 | .46532 | -.13879 | 1 | 2.69893 |
| 2 | .13279 | .14736 | .46532 | .24105 | 1 | 2.22208 |
| 3 | -.63444 | -.77627 | -.86710 | -.70854 | 3 | 2.08897 |
| 4 | .90001 | .14736 | 1.26477 | 1.76040 | 2 | 1.78801 |
| 5 | -.25082 | -.36577 | -.06765 | -.13879 | 3 | 2.89705 |
| 6 | -.25082 | -.46839 | -.06765 | -.32871 | 3 | 1.52943 |
| 7 | -.63444 | -.87889 | -.86710 | -1.08838 | 3 | 2.07236 |
| 8 | -.63444 | -.57102 | -.06765 | .05113 | 3 | 2.49924 |
| 9 | -.63444 | -.46839 | .73181 | .62089 | 3 | 2.83133 |
| 10 | -1.01805 | -.77627 | -.92039 | -1.08838 | 3 | 2.42051 |
| 11 | -1.01805 | -.70785 | -1.13358 | -.89846 | 3 | 2.09662 |
| 12 | -1.01805 | -.70785 | -.86710 | -1.08838 | 3 | 1.72225 |
| 13 | -.25082 | 1.68674 | 2.06423 | 1.95032 | 1 | 3.41201 |
| 14 | .51640 | .66049 | .73181 | .05113 | 1 | 2.12884 |
| 15 | .90001 | 1.68674 | 1.53126 | 1.38056 | 2 | 3.10903 |
| 16 | -1.01805 | -.53681 | -.92039 | -.32871 | 3 | 1.49535 |
| 17 | -.63444 | -.77627 | -.86710 | -.89846 | 3 | 1.88276 |
| 18 | -.63444 | -.87889 | -1.08029 | -1.46822 | 3 | 1.57130 |
| 19 | 1.66724 | 1.68674 | .57192 | .62089 | 2 | 3.11509 |
| 20 | 2.05085 | .66049 | 1.53126 | 1.38056 | 2 | 1.81975 |
| 21 | 2.05085 | -.16052 | 1.26477 | 1.76040 | 2 | 2.67373 |
| 22 | 2.05085 | .45524 | .73181 | .81081 | 2 | 2.26211 |
| 23 | -.63444 | .66049 | -.92039 | -.89846 | 1 | 3.56705 |
| 24 | -.63444 | -.77627 | -1.02699 | -.51862 | 3 | 2.35876 |

1

Data View Variable View

Contoh soal:

Sebuah penelitian dilakukan terhadap 12 kota, ingin diketahui pengelompokan kota-kota tersebut berdasarkan instrumen 5 variabel yaitu:

- Jumlah pendapatan kota (triliun)
- Jumlah pinjaman pemerintah kota (milyar)
- Jumlah dana hibah yang dimiliki kota (milyar)
- Konsumsi penduduk (milyar)
- Jumlah penduduk kota (juta jiwa).

Analisis data pada Tabel berikut menjadi 3 kelompok:

| Kota | Pendapatan | Pinjaman | Dana hibah | Konsumsi | Penduduk |
|------|------------|----------|------------|----------|----------|
| A | 55 | 5.6 | 9 | 50 | 25 |
| B | 61 | 8 | 7 | 62 | 41 |
| C | 58 | 3.9 | 7 | 60 | 32 |
| D | 67 | 5.5 | 7 | 64 | 51 |
| E | 71 | 5.7 | 6 | 70 | 42 |
| F | 76 | 7.6 | 8 | 80 | 29 |
| G | 81 | 8.7 | 9 | 80 | 57 |
| H | 56 | 7.1 | 6 | 86 | 29 |
| I | 84 | 7.6 | 7 | 82 | 46 |
| J | 88 | 6.5 | 8 | 86 | 52 |
| K | 84 | 6.8 | 9 | 88 | 61 |
| L | 90 | 8 | 9 | 90 | 66 |