

## Analisis Cluster Non Hirarki dengan SPSS

Sebelumnya kita telah mempelajari Interpretasi Analisis Cluster Hirarki dengan SPSS. Selanjutnya perlu diingat kembali bahwasanya ada dua macam analisis cluster, yaitu analisis cluster hirarki dan analisis cluster non hirarki. Analisis cluster non hirarki salah satunya dan yang paling populer adalah analisis cluster dengan K-Means Cluster. Ciri khas dari analisis cluster yang satu ini adalah kita menentukan sendiri dari awal berapa cluster yang akan kita bentuk. Oleh karena itu dalam tutorial ini, kita akan coba membuat 3 cluster pada sampel dan variabel seperti artikel sebelumnya yaitu analisis cluster hirarki dengan SPSS. Langsung saja kita pelajari tutorial uji atau Analisis Cluster Non Hirarki dengan SPSS.

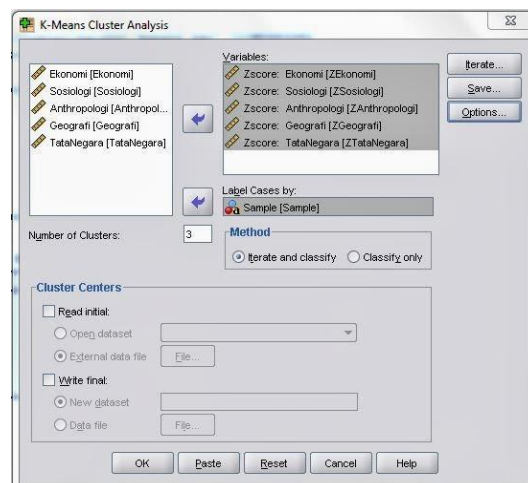
Download file kerja dalam tutorial ini: Dataset dan Output pada link sebagai berikut:

[https://drive.google.com/drive/folders/10pSu0zDc1cdSmhSPks1FAo7ercdD030Q?usp=share\\_link](https://drive.google.com/drive/folders/10pSu0zDc1cdSmhSPks1FAo7ercdD030Q?usp=share_link)

### Standarisasi Analisis Cluster

Sebelum melangkah lebih jauh, silahkan anda lakukan proses standarisasi variabel seperti yang telah kami jelaskan sebelumnya pada analisis cluster hirarki. Setelah proses standarisasi selesai, maka kita bulai tutorial ini, yaitu:

Pada menu, klik “Analyze”, “Classify”, “K-Means Cluster”.



Analisis K-Means Cluster

### Langkah Analisis Cluster K Means dengan SPSS

Perhatikan jendela di atas, anda masukkan **variabel** yang telah distandarisi ke dalam kotak “Variables” dan masukkan variabel “Sampel” ke dalam kotak “Label Cases by:”. Kemudian pada “Method” pilih “iterate and classify” serta pada “Number of cluster”

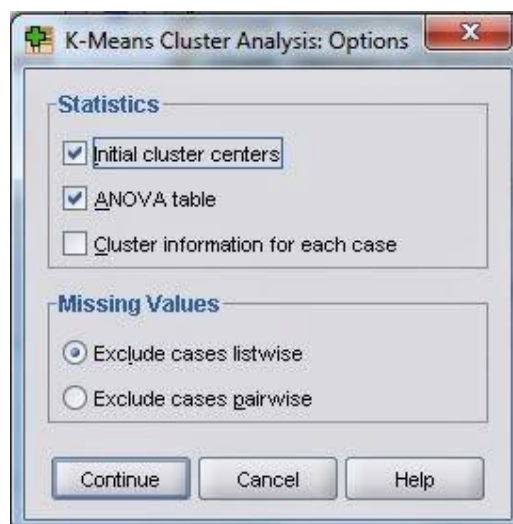
anda masukkan angka 3. Angka 3 di sini berarti anda akan membentuk 3 cluster. Selanjutnya klik tombol “Continue”.

Klik tombol “Save” kemudian centang “Cluster Membership” dan “Distance from cluster center”. Proses “Save” berguna untuk membentuk variabel baru secara otomatis untuk keperluan tahap akhir clustering yaitu tahap profiling. “Cluster Membership” artinya pada dataset, kita akan memberi tanda tiap sampel akan masuk ke salah satu cluster yang telah kita bentuk. Sedangkan “Distance from cluster center” untuk membentuk variabel baru yang merupakan jarak dengan nilai tengah cluster. Selanjutnya klik tombol “Continue”.



K-Means Save

Klik tombol “Options” kemudian pada “Statistics” centang “initial cluster center” dan “Anova Table”.



K-Means Initial Cluster

Setelah anda klik “Continue” maka selanjutnya anda berada pada jendela utama, maka klik “OK” dan lihat “Output”.

# Interprestasi Analisis Cluster Non Hirarki dengan SPSS

Setelah kita berhasil melakukan proses Analisis Cluster Non Hirarki dengan SPSS, maka langkah selanjutnya adalah menginterpretasikan hasilnya. Mari kita bersama-sama pelajari tutorial interpretasi analisis cluster non hirarki dengan SPSS. Langsung saja anda buka output view anda yang sudah anda hasilkan dari artikel sebelumnya.

## Initial Cluster

	Cluster		
	1	2	3
Zscore: Ekonomi	-.78155	2.28620	-.78155
Zscore: Sosiologi	-.54904	.52124	.03475
Zscore: Anthropologi	.55629	.05785	-.68980
Zscore: Geografi	1.47887	.54431	-1.18104
Zscore: TataNegara	1.43002	-.93086	-1.02530

Tabel Initial Cluster

Tabel “Initial Cluster Centers” di atas merupakan tampilan awal proses clustering sebelum dilakukan proses iterasi.

## Iterasi Analisis Cluster

Agar kita dapat mengetahui berapa kali dilakukan proses iterasi yang dilakukan pada objek sebanyak 14 sampel, dapat anda lihat di bawah ini:

Iteration	Change in Cluster Centers		
	1	2	3
1	1.379	1.308	.381
2	.482	.438	.000
3	.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 3. The minimum distance between initial centers is 3.632.

Iteration History

Dari tabel “Iteration History” di atas, dapat diketahui bahwasanya proses iterasi dilakukan sebanyak 3 kali. Proses ini dilakukan untuk mendapatkan cluster yang tepat. Dapat diketahui bahwa jarak minimum antar pusat cluster yang terjadi dari hasil iterasi adalah 3,632.

## Output Analisis Cluster dengan SPSS

Hasil akhir dari proses clustering dapat anda lihat seperti di bawah ini:

Final Cluster Centers			
	Cluster		
	1	2	3
Zscore: Ekonomi	-.43387	.99775	-.70485
Zscore: Sosiologi	-.86039	.57962	.35097
Zscore: Anthropologi	.64352	-.00445	-.79883
Zscore: Geografi	.06984	.86062	-1.16307
Zscore: TataNegara	1.01451	-.06206	-1.19058

Tabel Final Cluster

Pada output di tabel “Final Cluster Centers” dapat anda perhatikan, bahwasanya data di atas masih terkait dengan proses standarisasi yang mengacu pada z-score dengan ketentuan sebagai berikut:

- Nilai negatif (-) berarti data berada di bawah rata-rata total
- Nilai positif (+) berarti data berada di atas rata-rata total.

## Rumus Cluster

Rumus yang digunakan sebagai berikut:

$$\chi = \mu + Z \cdot \sigma$$

Rumus Rata-rata Sampel dalam Cluster

Di mana:

X = Rata-rata sampel dalam cluster

$\mu$  = Rata-rata populasi

Z = Nilai standarisasi

$\sigma$  = Standar Deviasi

Misalnya, apabila ingin mengetahui rata-rata nilai Geografi dalam cluster 1, yaitu:

(rata-rata nilai Geografi seluruh siswa) + (0,06984 x Standar Deviasi rata-rata Nilai Geografi)

= 69,43 + (0,06984 x 13,91)

= 70,4015

Jadi rata-rata nilai geografi yang berada pada cluster 1 adalah 70,4015.

Demikian seterusnya dapat diketahui rata-rata nilai masing-masing variabel dalam tiap cluster.

# Interprestasi Analisis Cluster Non Hirarki dengan SPSS

Interprestasi Analisis Cluster Non Hirarki dengan SPSS adalah: Berdasarkan tabel “Output Final Cluster Centers”, dengan ketentuan yang telah dijelaskan di atas, dapat didefinisikan sebagai berikut:

## Cluster 1:

Dalam cluster 1 ini berisi siswa dengan nilai ekonomi dan sosiologi yang rendah, nilai geografi yang sedang serta nilai anthropologi dan tata negara yang tinggi.

## Cluster 2

Dalam cluster 2 ini berisi siswa dengan nilai ekonomi dan sosiologi yang rendah, nilai anthropologi dan tata negara yang sedang serta nilai geografi yang tinggi.

## Cluster 3

Dalam cluster 3 ini berisi siswa dengan nilai ekonomi, anthropologi, geografi dan tata negara yang rendah serta nilai sosiologi yang sedang.

Perlu diingat kembali, penamaan masing-masing cluster sangat bersifat subjektif tergantung pada peneliti dengan mengacu pada tujuan penelitian.

Tahapan selanjutnya yang perlu dilakukan yaitu melihat perbedaan variabel pada cluster yang terbentuk. Dalam hal ini dapat dilihat dari nilai F dan nilai probabilitas (sig) masing-masing variabel, seperti yang dapat anda lihat dalam tabel berikut:

ANOVA						
	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Zscore: Ekonomi	3.953	2	.463	11	8.536	.006
Zscore: Sosiologi	2.937	2	.648	11	4.533	.037
Zscore: Anthropologi	2.312	2	.762	11	3.035	.089
Zscore: Geografi	4.569	2	.351	11	13.017	.001
Zscore: TataNegara	5.418	2	.197	11	27.528	.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.

## Interprestasi Analisis Cluster Non Hirarki

## Rumus ANOVA

Sedikit review, bahwa rumus dari uji F Anova adalah:

$$F = \frac{MS \text{ Between}}{MS \text{ Within}}$$

Rumus Anova

Dimana dalam tabel ANOVA di atas “MS Between” ditunjukkan oleh nilai “Means Square” dalam kolom “Cluster”, sedangkan “MS Within” ditunjukkan oleh nilai “Means Square” dalam kolom “Error”.

Kesimpulannya adalah:

Semakin besar nilai F dan ( $\text{sig} < 0,05$ ), maka semakin besar perbedaan variabel pada cluster yang terbentuk.

### ***Kesimpulan Analisis Cluster***

Maka berdasarkan tabel yang kita dapatkan dalam tutorial ini, yaitu bahwa untuk instrumen “Tata Negara” adalah variabel yang paling menunjukkan adanya perbedaan diantara siswa-siswa pada ketiga cluster yang terbentuk. Hal ini dengan ditunjukkannya nilai  $F = 27,528$  dan  $\text{sig} = 0,000$ . Dan untuk variabel yang lain dapat anda didefinisikan lebih lanjut.

Langkah selanjutnya adalah untuk mengetahui jumlah anggota masing-masing cluster yang terbentuk, yaitu dengan melihat pada tabel output di bawah ini:

Cluster	1	5.000
	2	5.000
	3	4.000
Valid		14.000
Missing		.000

Number Each Cluster

### ***Pembagian Cluster***

Nampak jelas bahwa cluster-1 beranggotakan 5 siswa, cluster-2 berisi 5 siswa, dan pada cluster-3 terdapat 4 siswa yang mengelompok. Dan untuk mengetahui siswa-siswa mana saja yang masuk dalam kategori tiap-tiap cluster dapat kembali dibuka tampilan “data view” dari dataset **SPSS** anda, yaitu pada kolom terakhir akan nampak seperti berikut ini:

ZGeografi	ZTataNegara	QCL_1	QCL_2
0.18486	1.43002	1	0.82973
0.54431	-0.93086	2	1.58817
-0.17459	0.67454	1	1.83593
-1.25293	-1.40304	3	0.25547
1.47887	1.43002	1	1.54385
-0.82159	0.76898	1	1.12122
1.26320	0.48567	2	1.75266
-1.18104	-1.02530	3	0.38131
-0.31837	0.76898	1	0.89678
1.40698	0.48567	2	2.42619
-1.25293	-1.40304	3	0.25547
0.40053	-0.26982	2	1.79717
-0.96537	-0.93086	3	0.37196
0.68809	-0.08094	2	1.73051

Z-Score

Perhatikan 2 kolom terakhir pada tabel di atas. “qcl\_1” menunjukkan nomor cluster dari keberadaan siswa, dan “qcl\_2” merupakan jarak antara obyek dengan pusat cluster. Dengan demikian, dapat ditafsirkan sebagai berikut:

**Cluster-1** : berisikan Siswa A, C, E, F, dan I dengan masing-masing jarak terhadap pusat cluster-1 adalah 0,82973; 1,58817; 1,83593; 0,25547; dan 1,54385.

**Cluster-2** : berisikan siswa B, G, J, L dan N, dengan masing-masing jarak terhadap pusat cluster-2 seperti yang anda lihat pada kolom “QCL\_2”.

**Cluster-3** : berisikan Siswa D, H, K, dan M, dengan masing-masing jarak terhadap pusat cluster-3 seperti yang anda lihat pada kolom “QCL\_2”.

**NB:** Apabila pada Analisis Cluster Hirarki dengan SPSS, pemberian tanda sampel menjadi anggota cluster mana, dapat dilihat pada “Output View” dan “Dataset”, tetapi pada Analisis Cluster Non Hirarki dengan SPSS, tanda itu hanya dapat dilihat pada “Dataset” saja.