



# PETA MUTU PENDIDIKAN INDONESIA : ANALISIS KLASTER PROVINSI DENGAN K-MEANS DAN K-MEDOIDS

Presented by: Kelompok I

Nabila Syukri  
Yani Prihantini Hiola  
Mega Ramatika Putri

Course Project : Pemodelan Klasifikasi





# OVERVIEW

01

Pendahuluan

02

Tinjauan Pustaka

03

Tahapan Analisis

04

Kesimpulan

# 01 PENDAHULUAN

## Latar Belakang

Pendidikan merupakan salah satu faktor utama dalam pembangunan sumber daya manusia yang berkualitas.

Setiap sekolah di Indonesia memiliki peran penting dalam mencetak generasi yang mampu bersaing di tingkat nasional maupun global. Namun, kualitas (mutu) sekolah antar provinsi di Indonesia menunjukkan variasi cukup besar.

Untuk memahami pola perbedaan mutu sekolah, dapat digunakan **analisis clustering**, yaitu metode pengelompokan berdasarkan kesamaan karakteristik.

**K-Means** dan **K-Medoids** digunakan untuk mengelompokkan provinsi berdasarkan 4 peubah yang mewakili mutu sekolah



### Mutu Lulusan



### Proses Pembelajaran



### Mutu Guru



### Manajemen Sekolah/Madrasah



# 01 PENDAHULUAN

## Tujuan

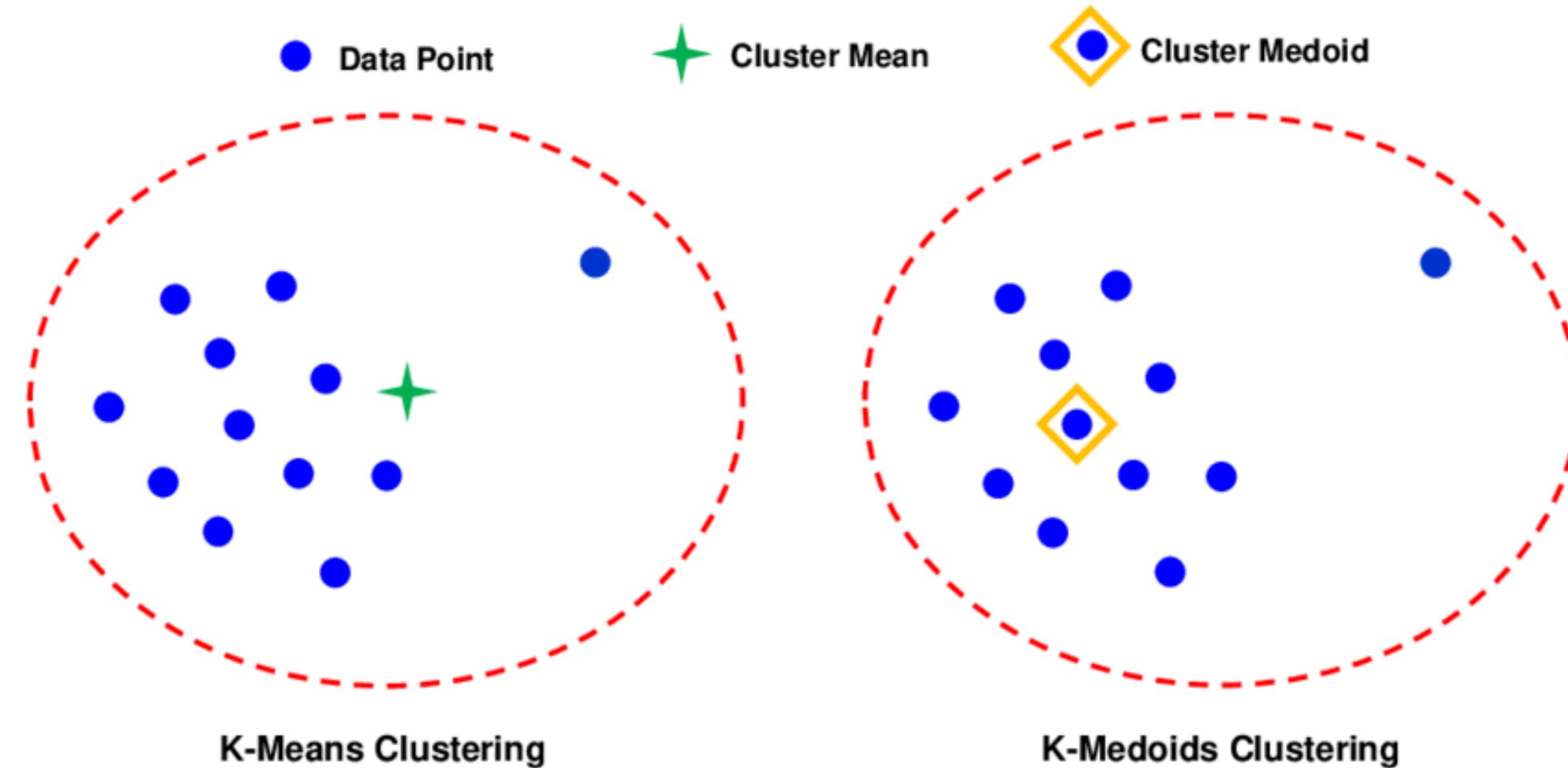
Membandingkan hasil penggerombolan provinsi berdasarkan rata-rata mutu SMA di Indonesia dengan menggunakan metode K-Means dan K-Medoids

## 02 TINJAUAN PUSTAKA

- Non-Hierarki Clustering
- Jarak
- Metrik Evaluasi Clustering

# NON-HIRARKI CLUSTERING

Metode clustering yang membagi data menjadi beberapa cluster tertentu ( $k$ )



Mengelompokkan data ke dalam beberapa kluster yang diwakili oleh **centeroid** yaitu rata-rata (mean) dari semua titik dalam cluster

mengelompokkan data ke dalam beberapa kluster dengan prinsip menggunakan objek perwakilan (**medoid**) sebagai pusat kluster (centroid) untuk setiap kluster



# JARAK

Pemilihan ukuran jarak mempengaruhi hasil penggerombolan, dilakukan perbandingan penggunaan jarak pada kedua metode clustering.

## EUCLIDEAN

$$d(x, y) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2}$$

Metode ini sesuai digunakan apabila setiap peubah berskala sama dan tidak terdapat korelasi yang kuat antar peubah

## MAHALANOBIS

$$d_M(x, y) \sqrt{(x - y)^T S^{-1} (x - y)}$$

Metode ini mempertimbangkan skala peubah dan korelasi antar peubah

# METRIK EVALUASI CLUSTERING

Dalam menilai kinerja hasil clustering, digunakan dua metrik evaluasi yaitu **Silhouette Coefficient** dan **Davies-Bouldin Index (DBI)**.

## SILHOUETTE COEFFICIENT

Silhouette digunakan untuk mengevaluasi keseimbangan antara **kohesi** (seberapa rapat objek dengan klasternya sendiri) dan **separasi** (seberapa jauh objek dari klaster terdekat).

- Nilai positif mendekat 1 berarti klaster **baik**
- Nilai mendekati 0 menunjukkan **ambiguitas**
- Nilai negatif menunjukkan **salah** klasterisasi.

## DAVIES-BOULDIN INDEX

DBI menilai **kompaknya** (seberapa rapat titik-titik data dalam satu cluster) suatu klaster sekaligus pemisahan antar klaster.

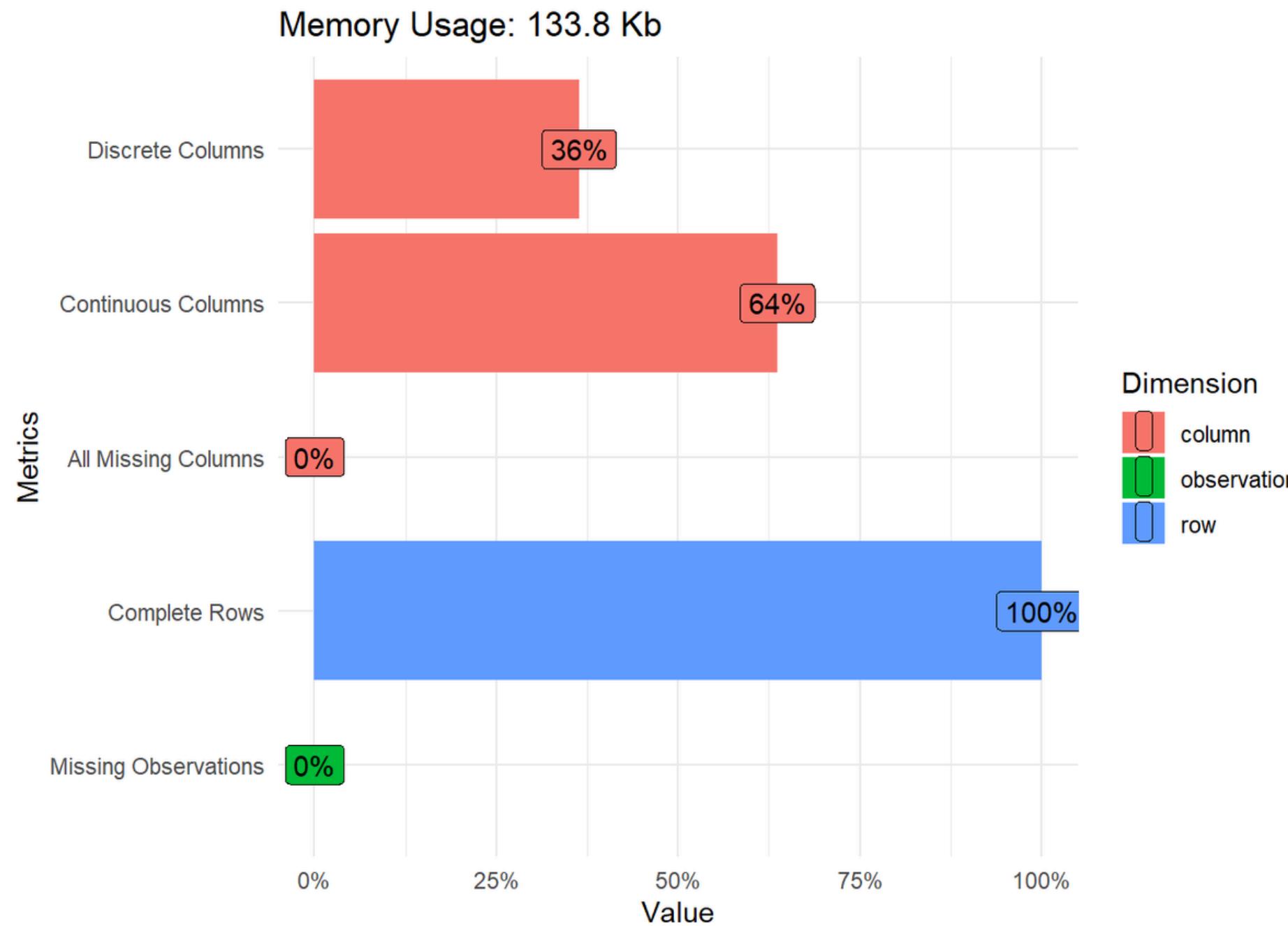
- Nilai DBI  $> 0$
- Semakin **rendah** DBI, semakin **baik** kualitas klaster.

## 03 TAHAPAN ANALISIS



01

# EKSPLORASI DATA



## Pengecekan Missing Value

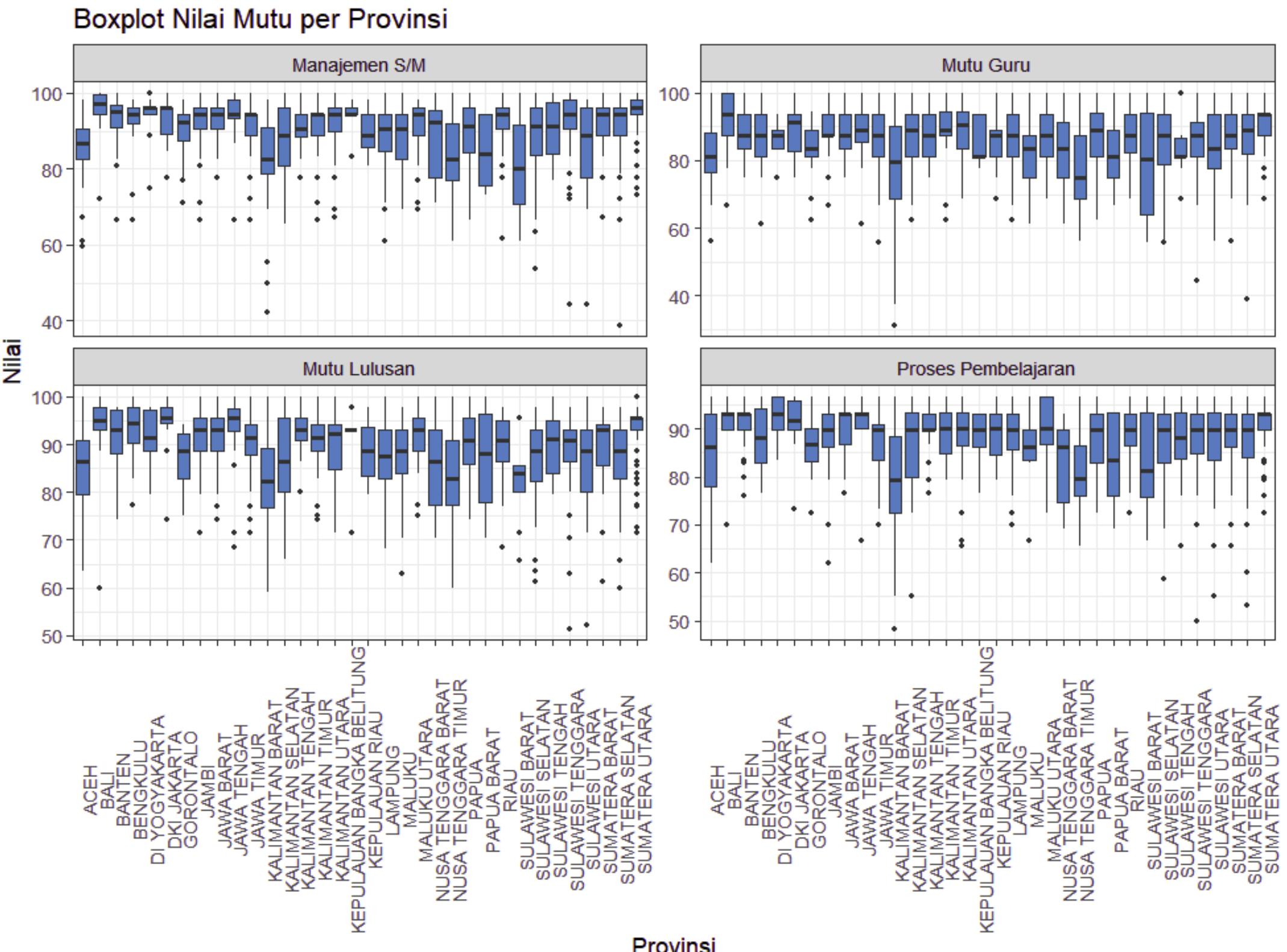
Tidak terdapat missing value  
pada kolom maupun observasi

01

# EKSPLORASI DATA

## Boxplot untuk masing-masing Provinsi

Berdasarkan boxplot masing-masing provinsi untuk setiap peubah, dapat diketahui bahwa secara umum, terdapat beberapa titik penculan di tiap Provinsi yang banyak tersebar di luar batas bawah boxplot.



01

# EKSPLORASI DATA

## Agregasi Data

- Terdapat 1.444 baris yang mewakili satu sekolah di berbagai provinsi
- Agregasi tiap sekolah berdasarkan provinsi dengan nilai **mean** untuk masing-masing peubah
- Sehingga Setiap provinsi diwakili oleh satu baris yang merupakan rata-rata seluruh sekolah di provinsi tersebut

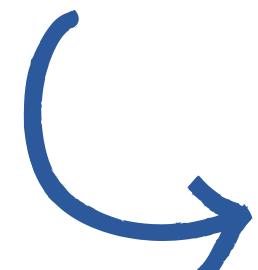
**Total Observasi = 34 Provinsi**

	Provinsi	Mutu Lulusan	Proses Pembelajaran	Mutu Guru	Manajemen S/M
1	JAWA TIMUR	88.57143	90.00000	88.88889	94.44444
2	SUMATERA BARAT	82.85714	86.66667	83.33333	88.88889
3	JAWA BARAT	74.28571	83.33333	77.77778	88.88889
4	BALI	94.28571	96.66667	100.00000	100.00000
5	SUMATERA SELATAN	97.14286	96.66667	94.44444	100.00000
6	JAWA BARAT	82.85714	90.00000	77.77778	77.77778
7	NUSA TENGGARA TIMUR	80.00000	80.00000	77.77778	83.33333
8	GORONTALO	91.42857	93.33333	83.33333	94.44444
9	JAWA BARAT	82.85714	90.00000	83.33333	94.44444
10	JAMBI	88.57143	86.66667	88.88889	94.44444

Showing 1 to 10 of 1,144 entries, 5 total columns

	Provinsi	Mutu Lulusan	Proses Pembelajaran	Mutu Guru	Manajemen S/M
1	ACEH	85.04377	84.45532	82.08983	85.40780
2	BALI	93.39481	91.85758	91.89646	95.51010
3	BANTEN	91.29524	90.38333	88.26620	92.14583
4	BENGKULU	93.38393	88.51042	87.03125	91.19097
5	DI YOGYAKARTA	91.76484	91.96667	87.13248	93.72222
6	DKI JAKARTA	94.50000	91.09524	88.80556	92.69048
7	GORONTALO	87.39464	85.84792	83.86111	89.18403
8	JAMBI	91.91962	89.17106	88.82479	93.10623
9	JAWA BARAT	91.57271	90.06468	88.27529	93.14096
10	JAWA TENGAH	92.48571	90.99825	88.52924	93.29532

Showing 1 to 10 of 34 entries, 5 total columns



01

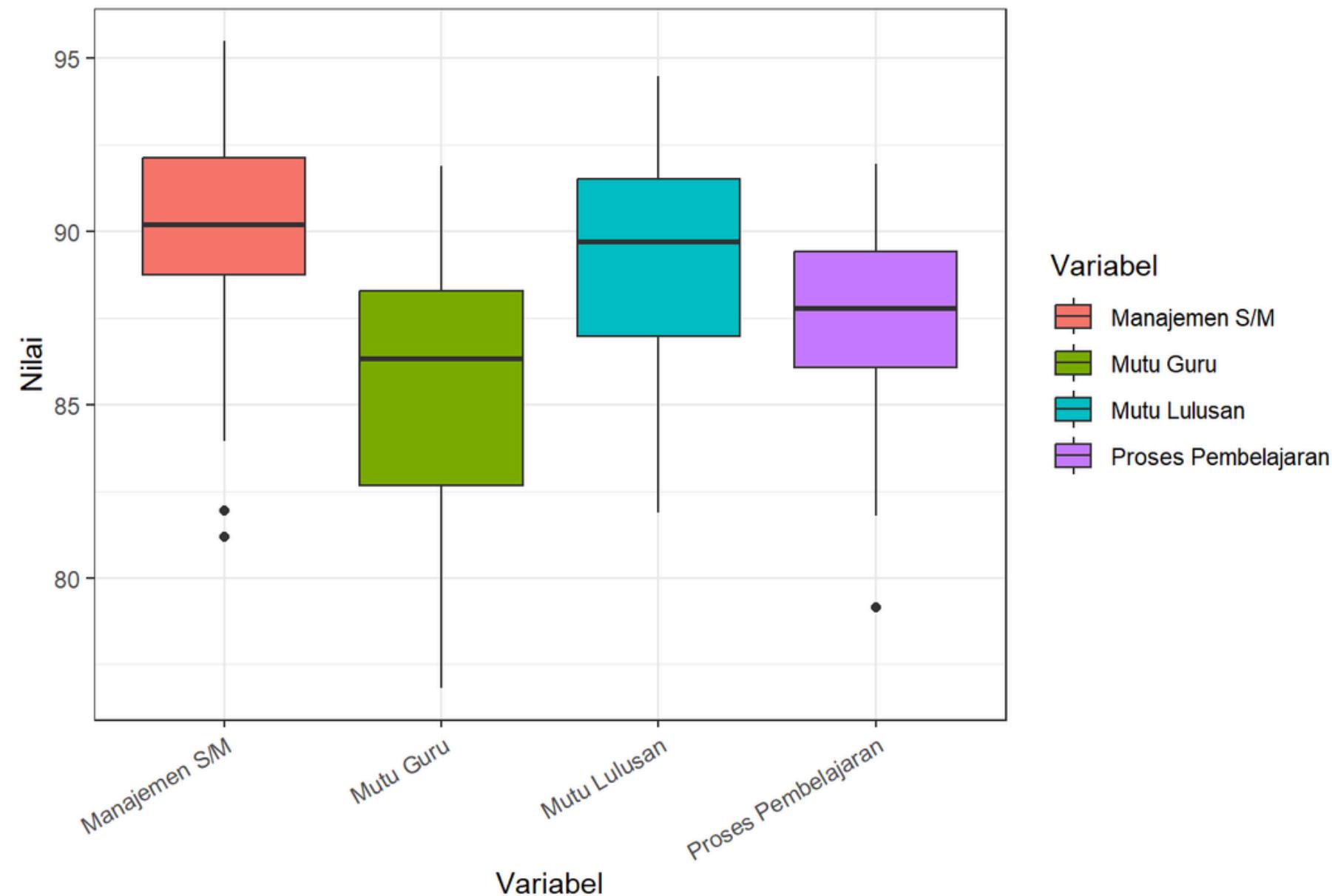
# EKSPLORASI DATA

## Statistik Deskriptif

	Manajemen S/M	Mutu Guru	Mutu Lulusan	Proses Pembelajaran
Minimum	81.17	76.82	81.88	79.14
Mean	89.75	85.48	88.94	87.47
Maksimum	95.51	91.9	94.50	91.97

Berdasarkan boxplot masing-masing peubah, terlihat bahwa terdapat tiga titik pencilan yakni **dua titik pada peubah Manajemen S/M dan satu titik pada peubah Proses Pembelajaran.**

Boxplot Masing-masing Peubah



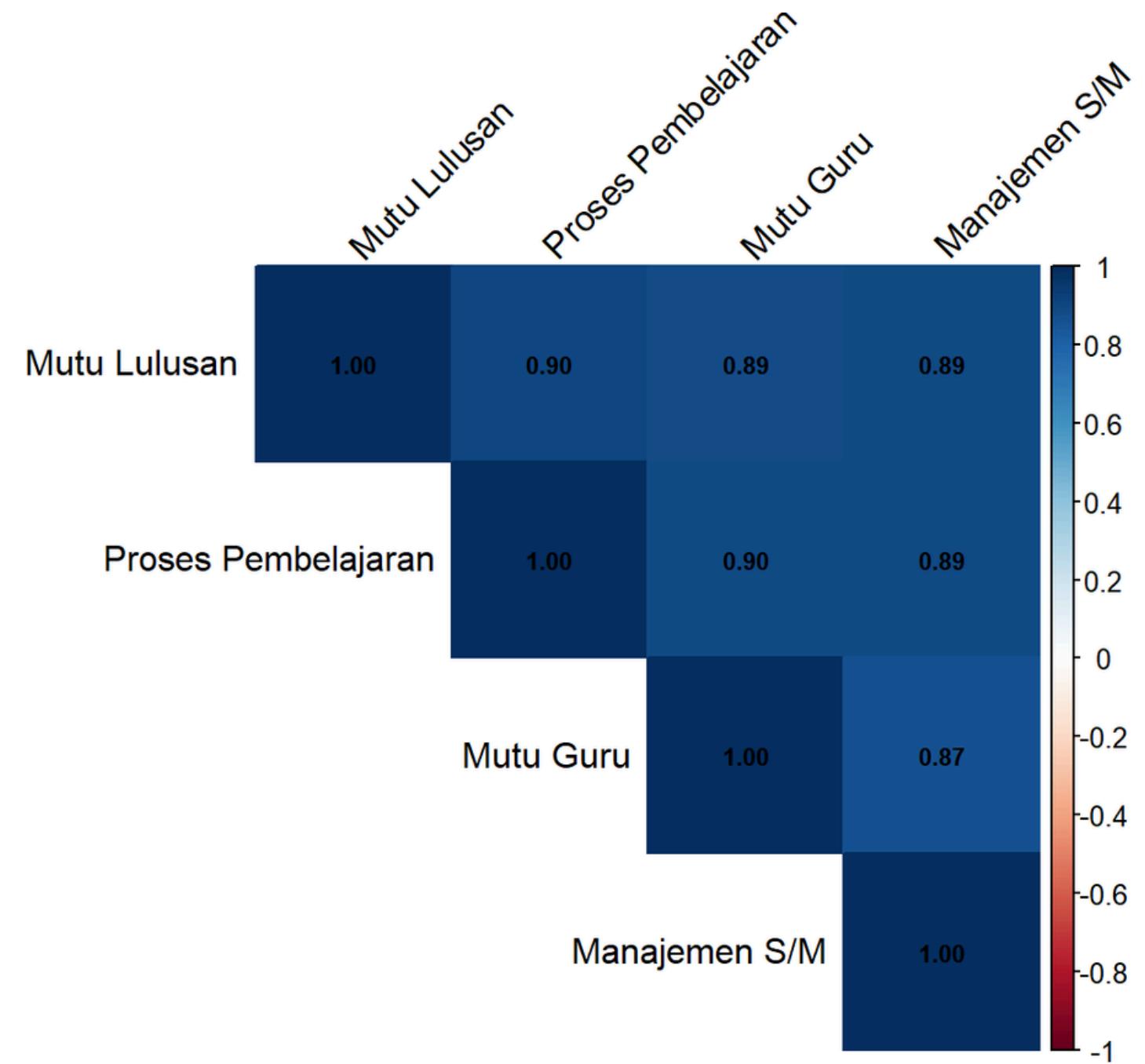
01

# EKSPLORASI DATA

## Korelasi

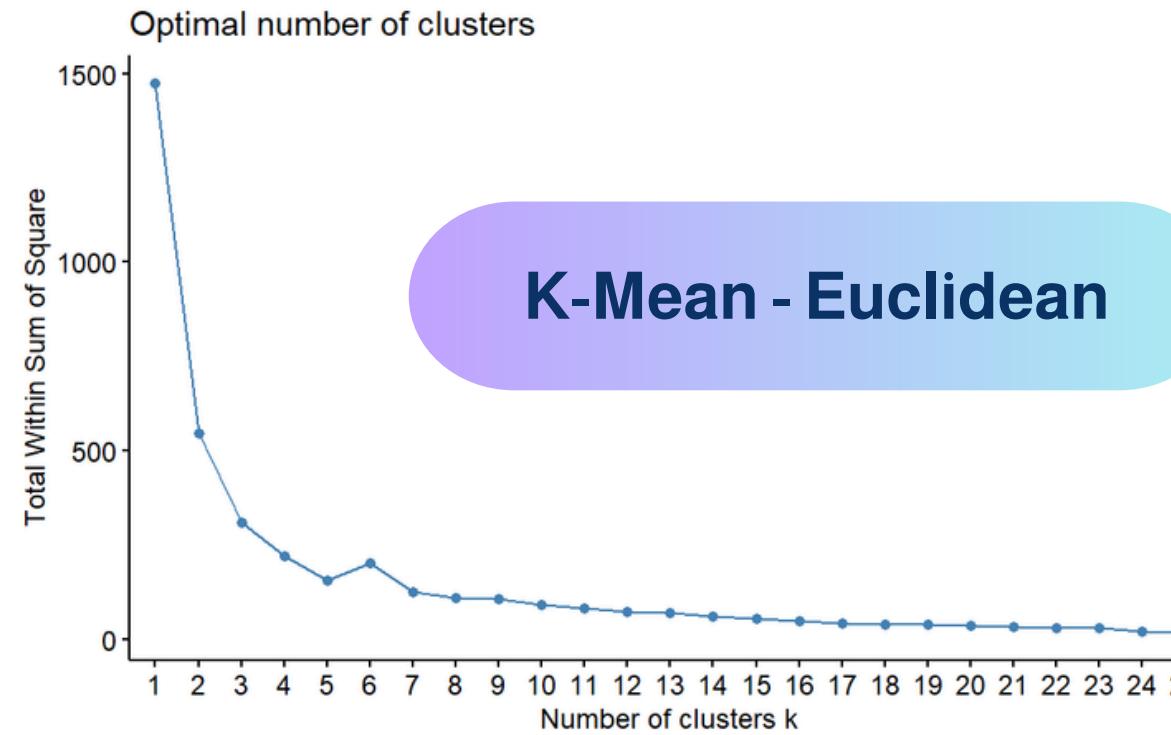
korelasi antar tiap peubah memiliki nilai yang cukup tinggi yaitu di atas 0.85.

Hal ini menandakan bahwa masing-masing peubah memiliki hubungan yang **searah** yaitu ketika suatu peubah mengalami kenaikan, maka peubah lainnya juga cenderung naik.

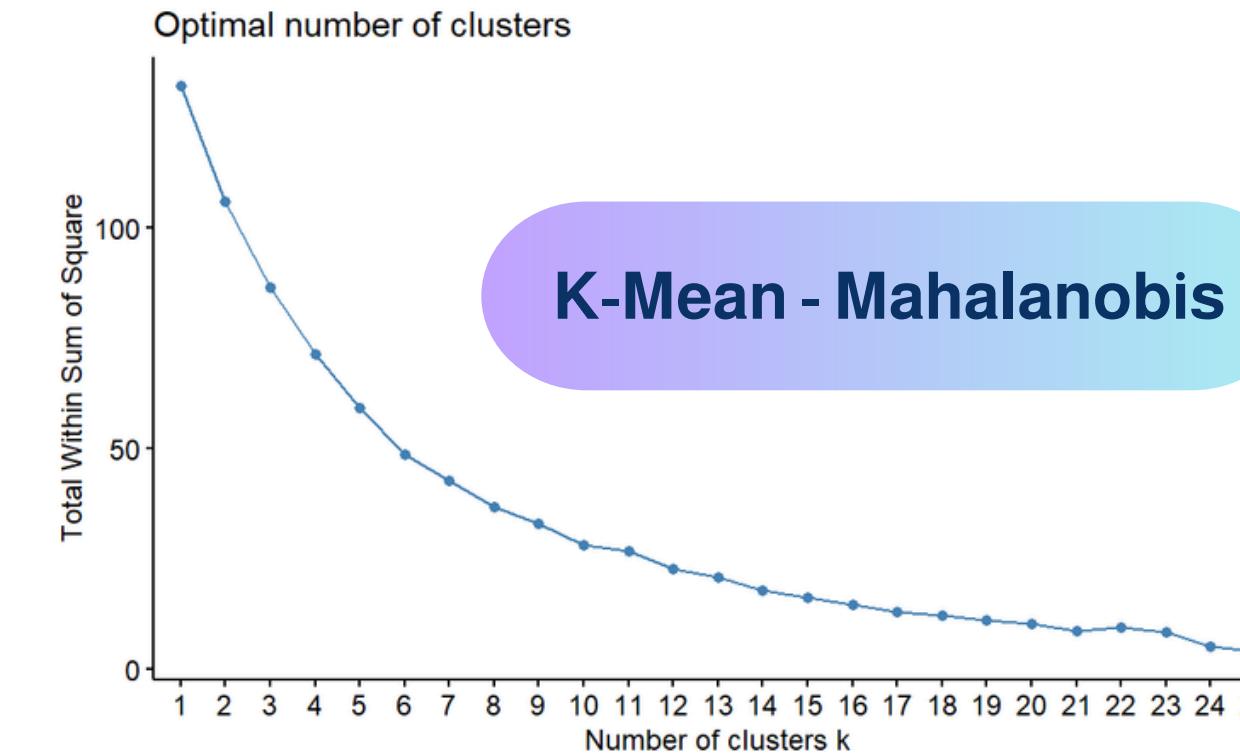


# 02

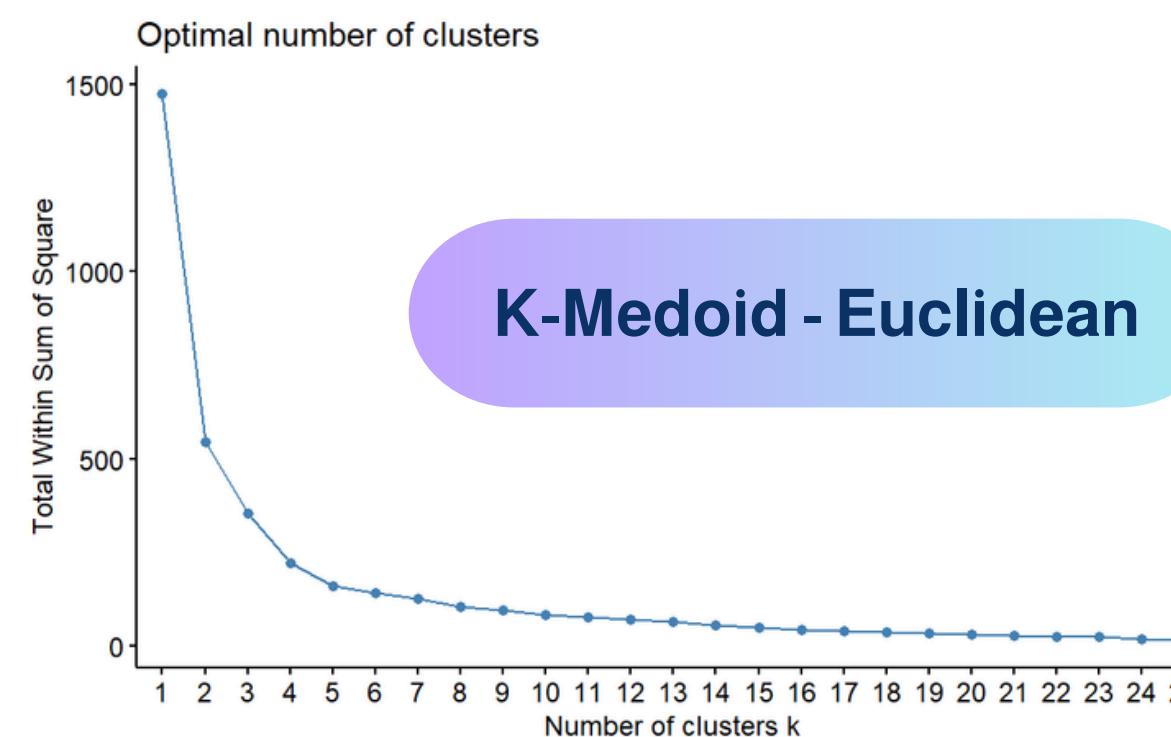
# MENENTUKAN JUMLAH CLUSTER (K)



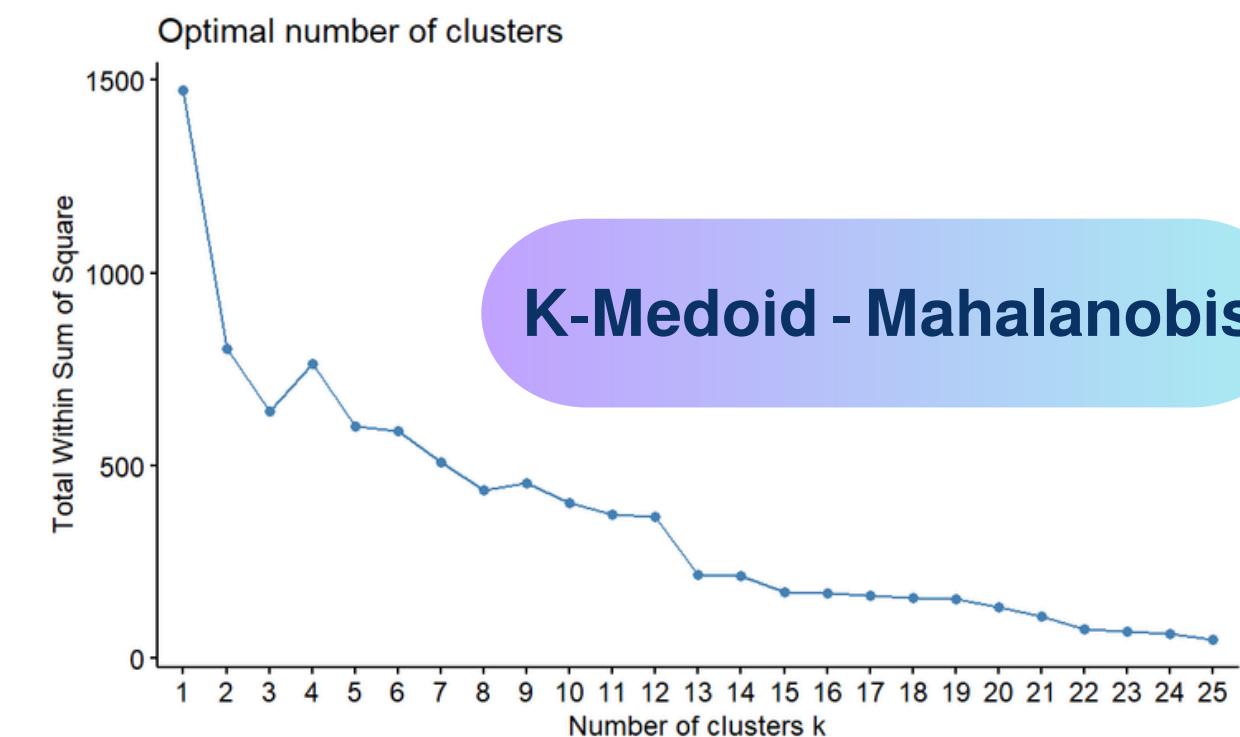
K-Mean - Euclidean



K-Mean - Mahalanobis



K-Medoid - Euclidean



K-Medoid - Mahalanobis

Penurunan WSS menjadi semakin kecil seiring dengan bertambahnya k.

Berdasarkan grafik **Elbow** yang dihasilkan, titik siku terlihat pada sekitar  $k = 4$  hingga  $k = 6$ , sehingga jumlah cluster optimal dipilih pada rentang tersebut.

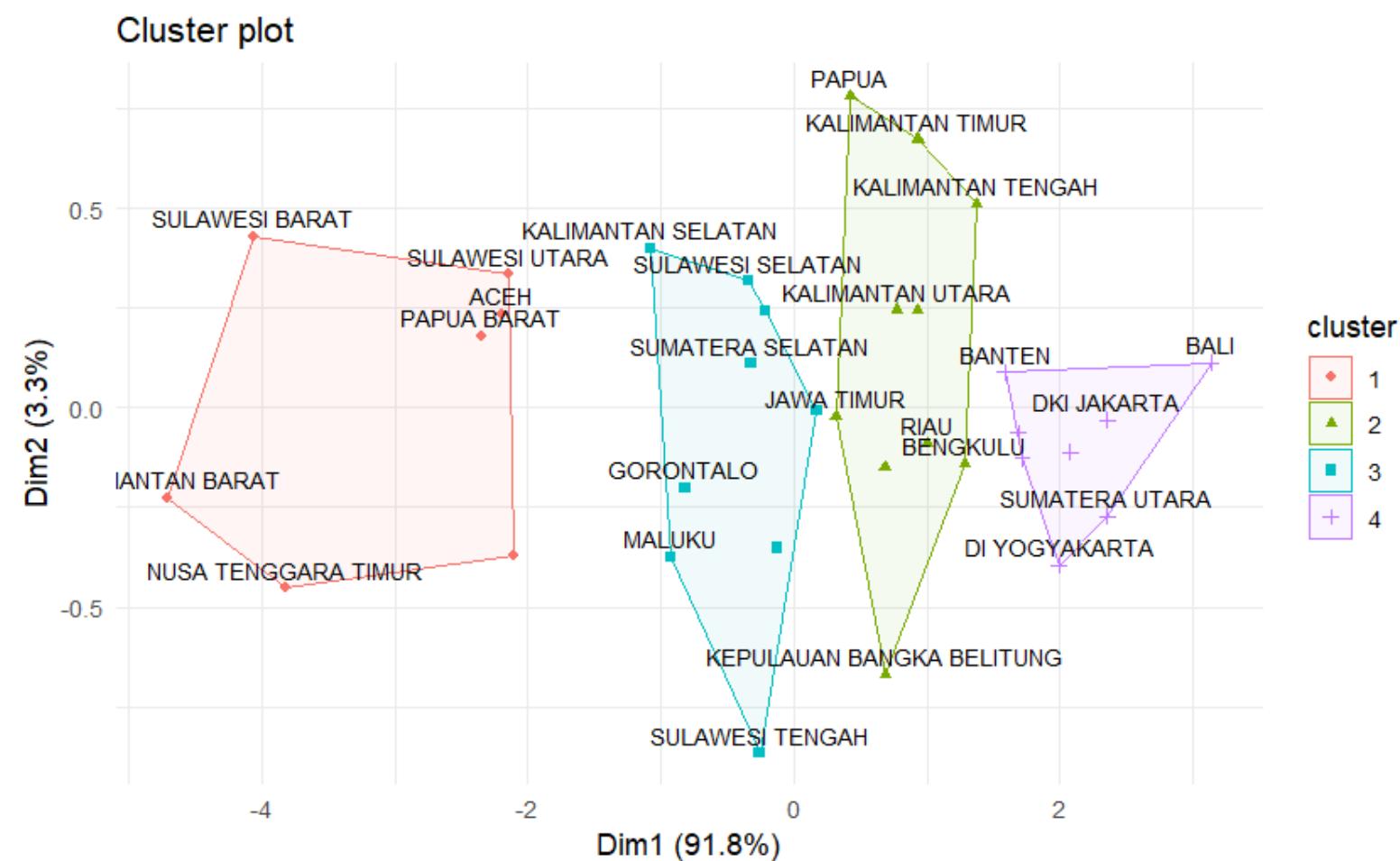
## 02

# MENENTUKAN JUMLAH CLUSTER

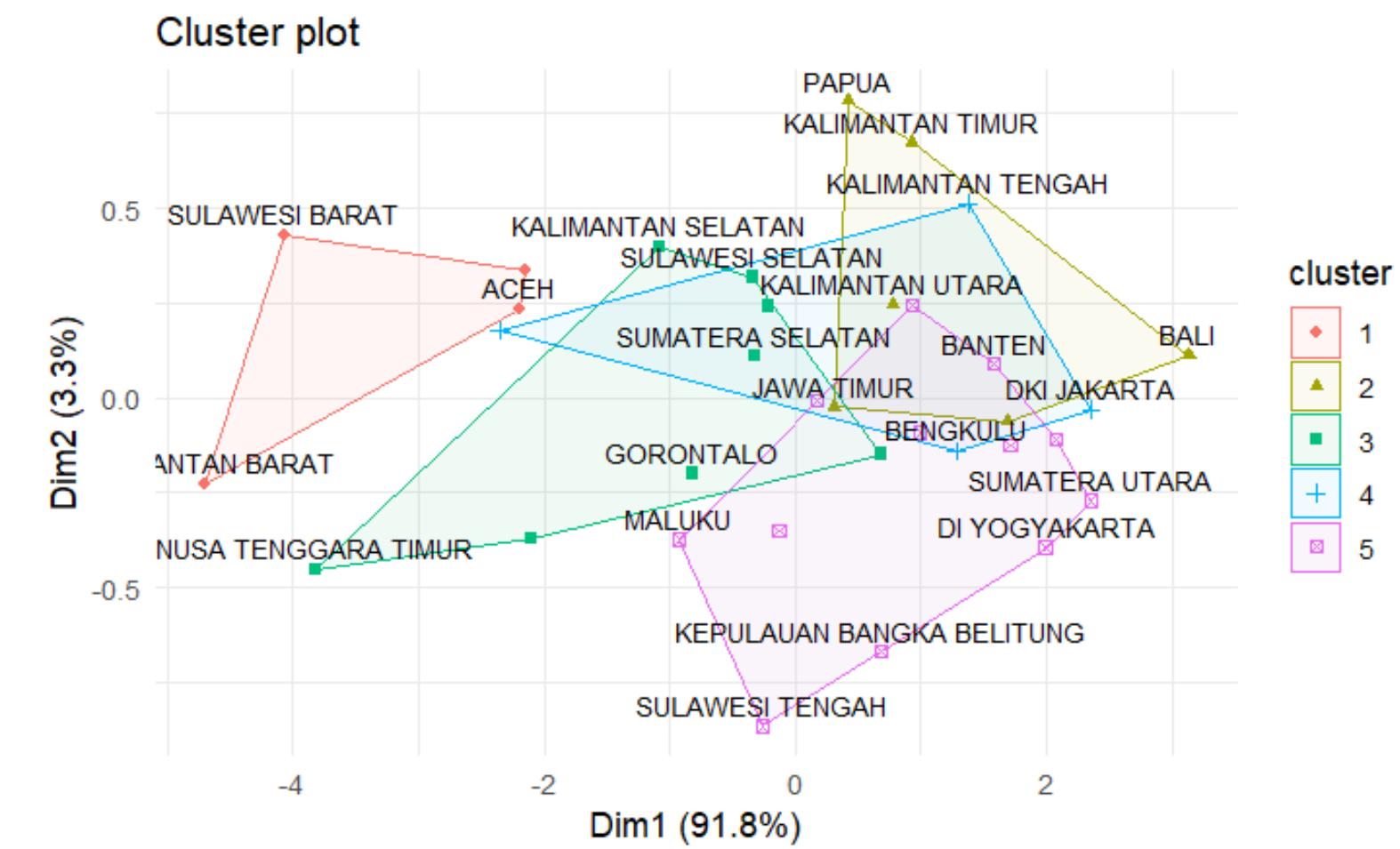
Evaluasi clustering digunakan adalah **Silhouette Coefficient** dan **Davies-Bouldin Index (DBI)**. Berdasarkan grafik elbow, Dipilih k = 4, 5, dan 6 yang akan digunakan sebagai kandidat jumlah cluster.

Metode	k	Silhouette	DBI	k terpilih
K-Mean Euclidean	4	0,3573	0,9151	4
	5	0,2757	1,1036	
	6	0,2912	1,0023	
K-Mean Mahalanobis	4	0,1397	2,745	5
	5	0,1443	2,3073	
	6	0,1569	2,3718	
K-Medoid Euclidean	4	0,3761	0,9503	4
	5	0,365	1,0305	
	6	0,3306	0,9912	
K-Medoid Mahalanobis	4	0,1744	2,4636	4
	5	0,2044	3,4667	
	6	0,2072	3,055	

## K-Mean - Euclidean



## K-Mean - Mahalanobis

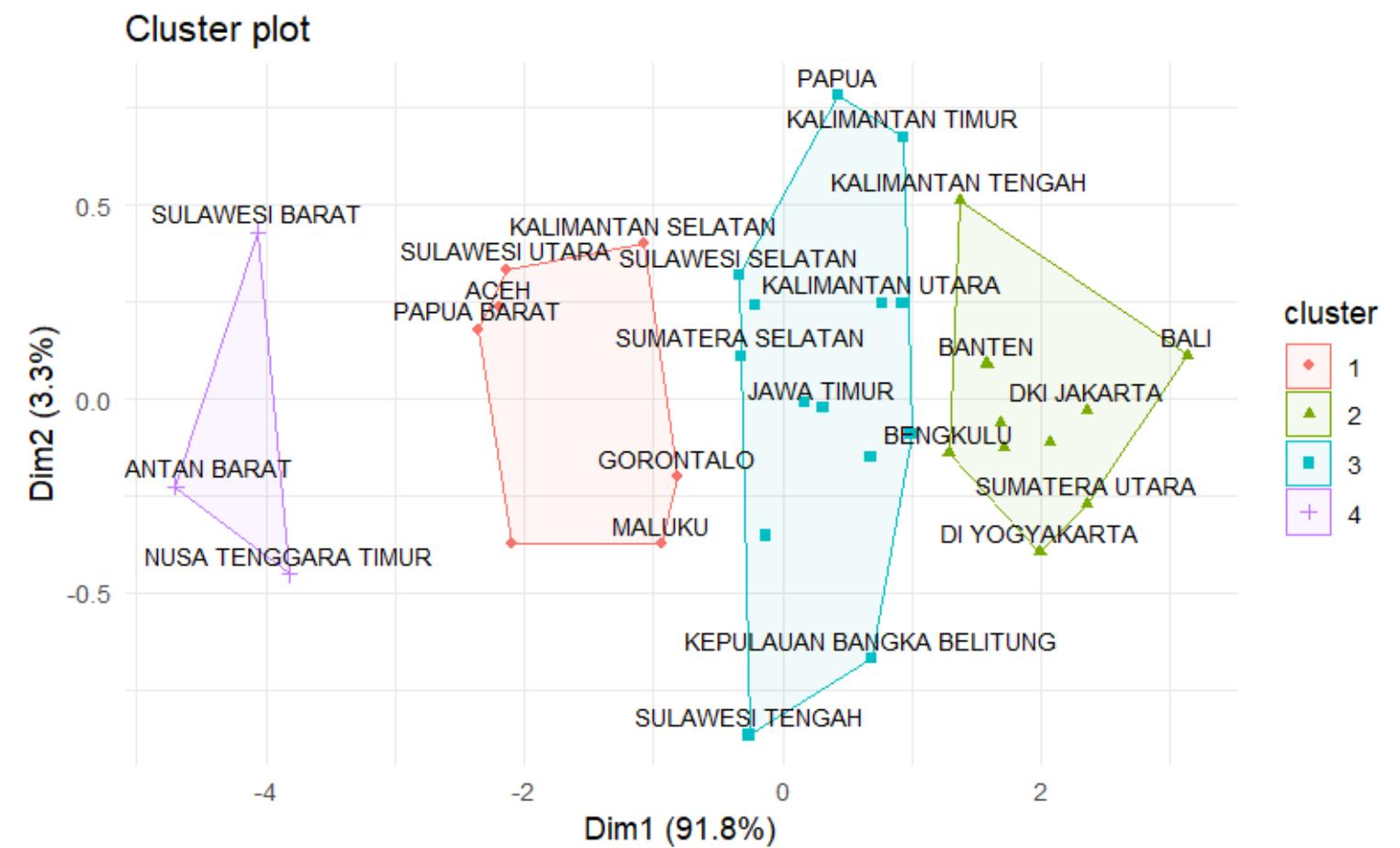


- Metode **K-Mean** dengan jarak **Euclidean** membentuk cluster yang jelas terpisah dari cluster yang lain tidak seperti pada metode k-mean dengan jarak mahalanobis.
- Penggunaan jarak **Mahalanobis** pada metode K-Mean menghasilkan beberapa cluster yang saling berdekatan sehingga tampak tumpang tindih (kurang jelas pemisahannya).

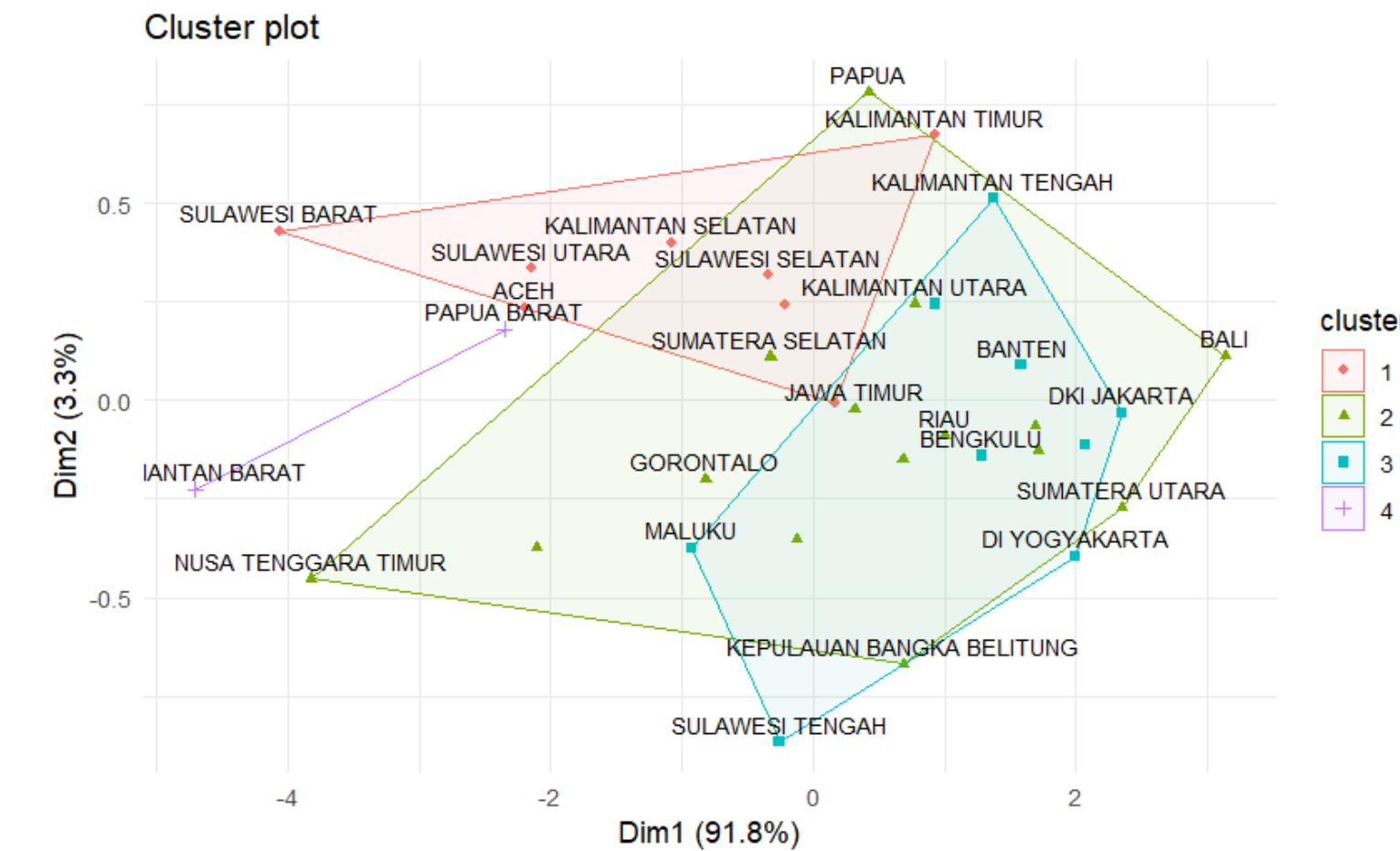
03

# ANALISIS CLUSTERING

## K-Medoid - Euclidean



## K-Medoid - Mahalanobis



- Metode **K-Medoid** dengan jarak **Euclidean** membentuk cluster yang jelas terpisah dari cluster yang lain tidak seperti pada metode k-medoid dengan jarak mahalanobis.
- Penggunaan jarak **Mahalanobis** pada metode K-Medoid menghasilkan beberapa cluster yang saling berdekatan sehingga tampak tumpang tindih (kurang jelas pemisahannya).

# EVALUASI MODEL

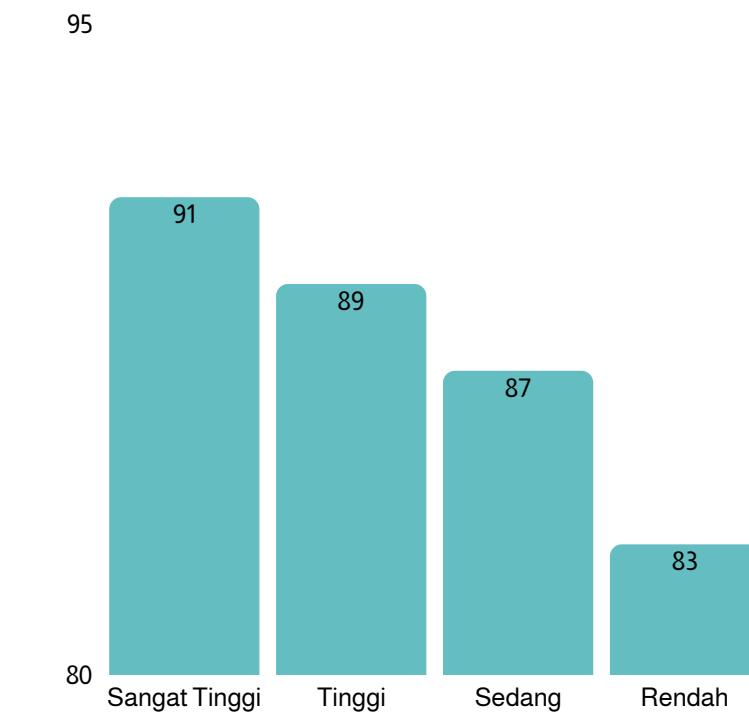
Metode Cluster	Jarak	Silhouette	DBI
K-Means	Euclidean (k=4)	0.3573	<b>0.9151</b>
K-Means	Mahalanobis (k=5)	0.1396	2.7449
K-Medoids	Euclidean (k=4)	<b>0.3761</b>	0.9503
K-Medoids	Mahalanobis (k=4)	0.1743	2.4636

- Nilai Silhouette berkisar -1 sampai 1 → Semakin tinggi semakin baik
- Nilai DBI memiliki nilai  $\geq 0$ , → Semakin kecil semakin baik

- K-Means Euclidean memiliki nilai DBI terendah yang berarti cluster terbentuk relatif kompak.
- K-Medoids Euclidean memiliki nilai Silhouette tertinggi berarti cluster antar provinsi terpisah lebih jelas.
- Secara umum, pada kasus ini penggunaan jarak euclidean lebih sesuai digunakan untuk data mutu SMA.

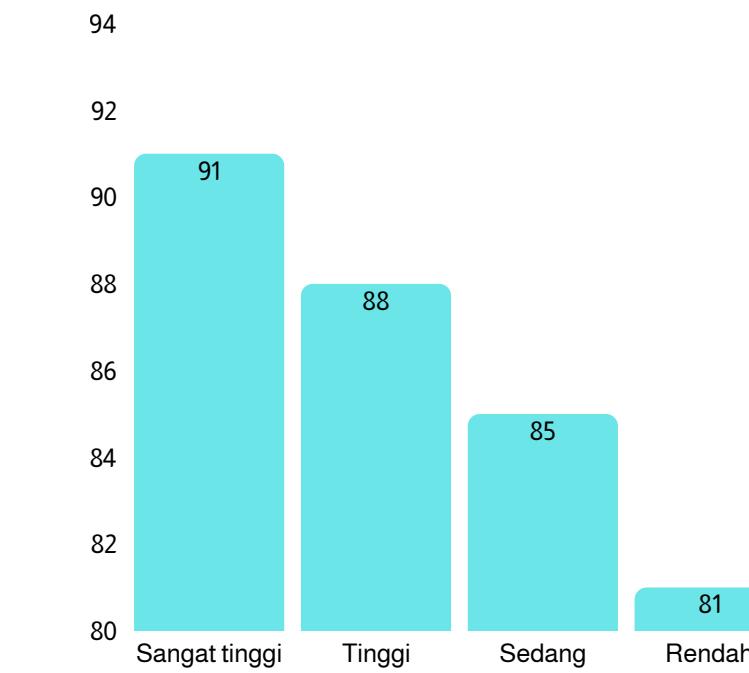
## K-MEAN EUCLIDEAN

Kluster	Mutu Lulusan	Proses Pembelajaran	Mutu Guru	Manajemen S/M	Rata-rata	Ketgori
1	84.3691	82.862	79.896	84.143	83	Rendah
2	90.503	88.388	87.504	90.898	89	Tinggi
3	87.580	87.128	84.544	89.481	87	Sedang
4	92.524	90.740	88.867	93.513	91	Sangat tinggi



## K-MEDOID EUCLIDEAN

Kluster	Mutu Lulusan	Proses Pembelajaran	Mutu Guru	Manajemen S/M	Rata-rata	Kategori
1	86.195	85.023	82.565	86.802	85	Sedang
2	92.584	90.407	88.670	92.934	91	Sangat tinggi
3	89.088	87.929	86.355	90.529	88	Tinggi
4	82.531	81.250	77.508	82.353	81	Rendah



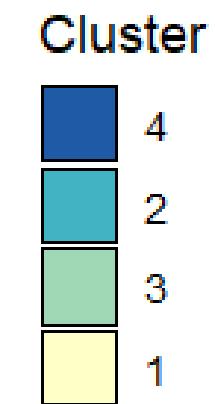
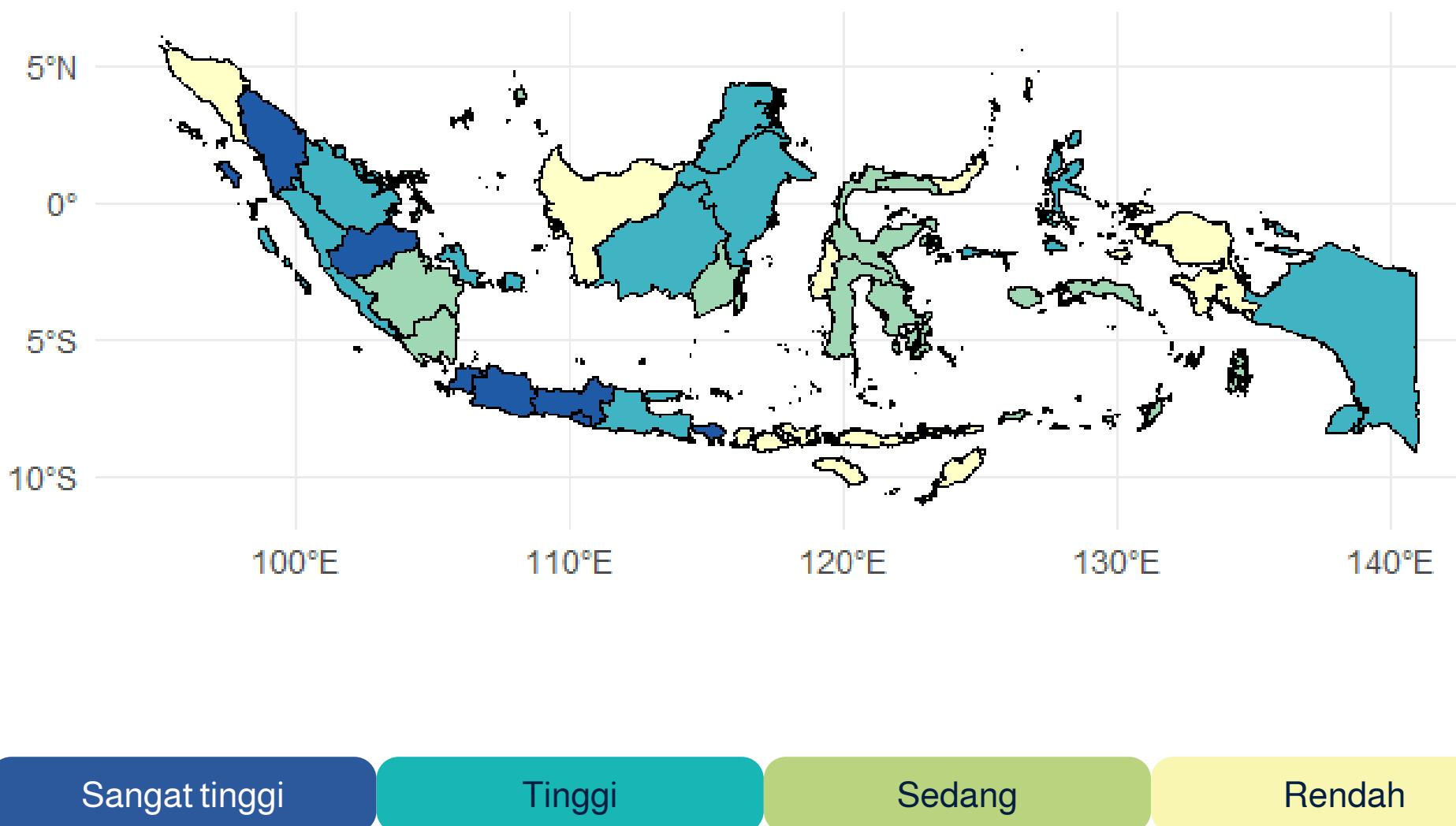
**Interpretasi :** Hasil kluster memberikan nilai-nilai tiap peubah yang searah, sehingga dibentuk 4 kategori dari hasil rata-rata 4 perubah

Jarak antar kategori pada K-Mean lebih sempit daripada K-Medoids

# VISUALISASI

## K-MEAN EUCLIDEAN

Peta Klasterisasi Provinsi di Indonesia  
Hasil clustering K-means (warna sesuai ranking)



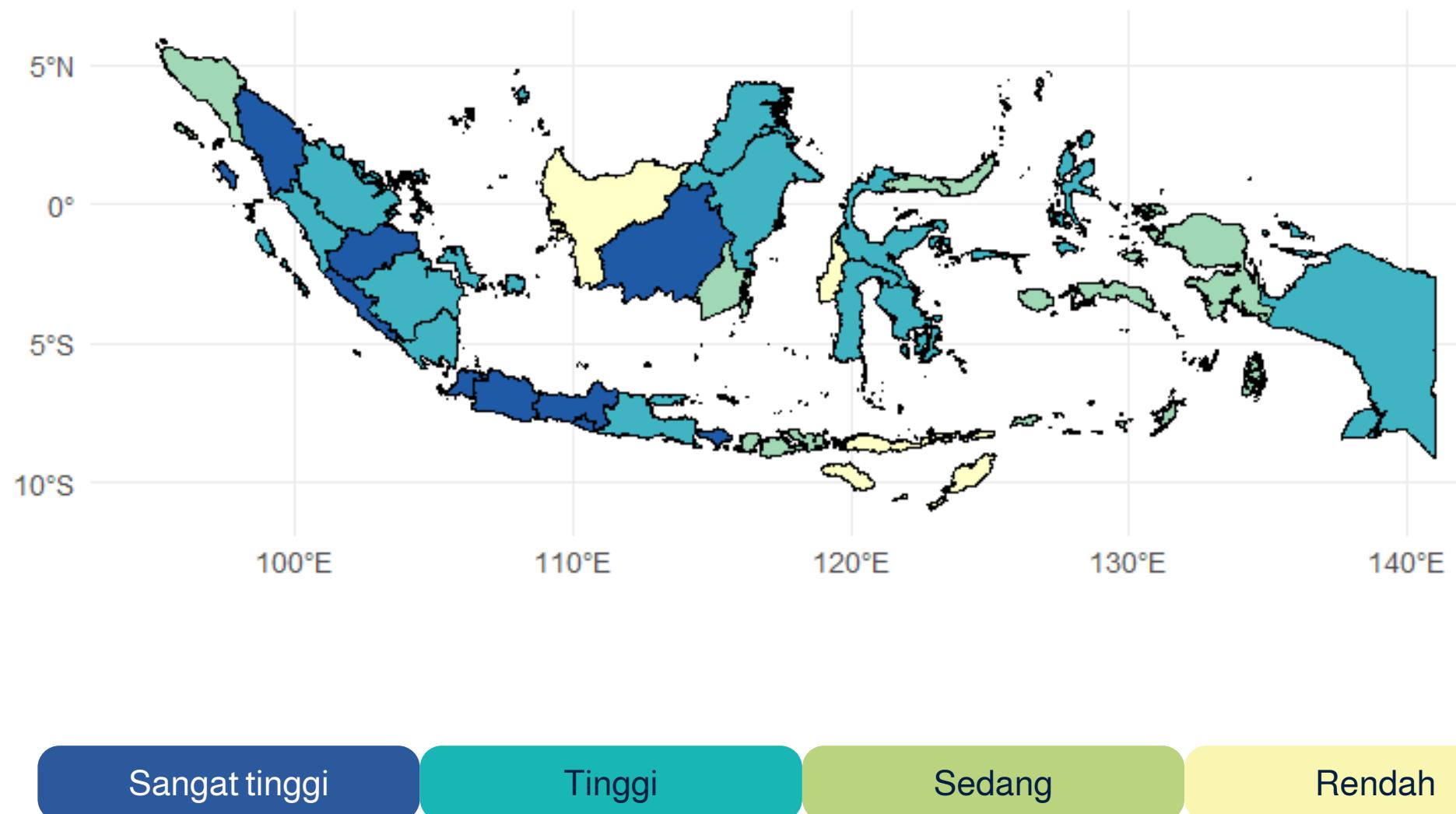
- Bali
  - Banten
  - DKI Jakarta
  - Jambi
  - Jawa Barat
  - Jawa Tengah
  - Sumatera Utara
  - DI Yogyakarta
- Gorontalo
  - Kalimantan Selatan
  - Kepulauan Riau
  - Lampung
  - Maluku
  - Sulawesi Selatan
  - Sulawesi Tengah
  - Sulawesi Tenggara
  - Sumatera Selatan
- Bangka Belitung
  - Bengkulu
  - Jawa Timur
  - Kalimantan Tengah
  - Kalimantan Timur
  - Kalimantan Utara
  - Maluku Utara
  - Papua
  - Riau
  - Sumatera Barat
- Aceh
  - Kalimantan Barat
  - NTB
  - NTT
  - Papua Barat
  - Sulawesi Barat
  - Sulawesi Utara

# VISUALISASI

## K-MEDOIDS EUCLIDEAN

Peta Klasterisasi Provinsi di Indonesia

Hasil clustering K-Medoid (warna sesuai ranking)

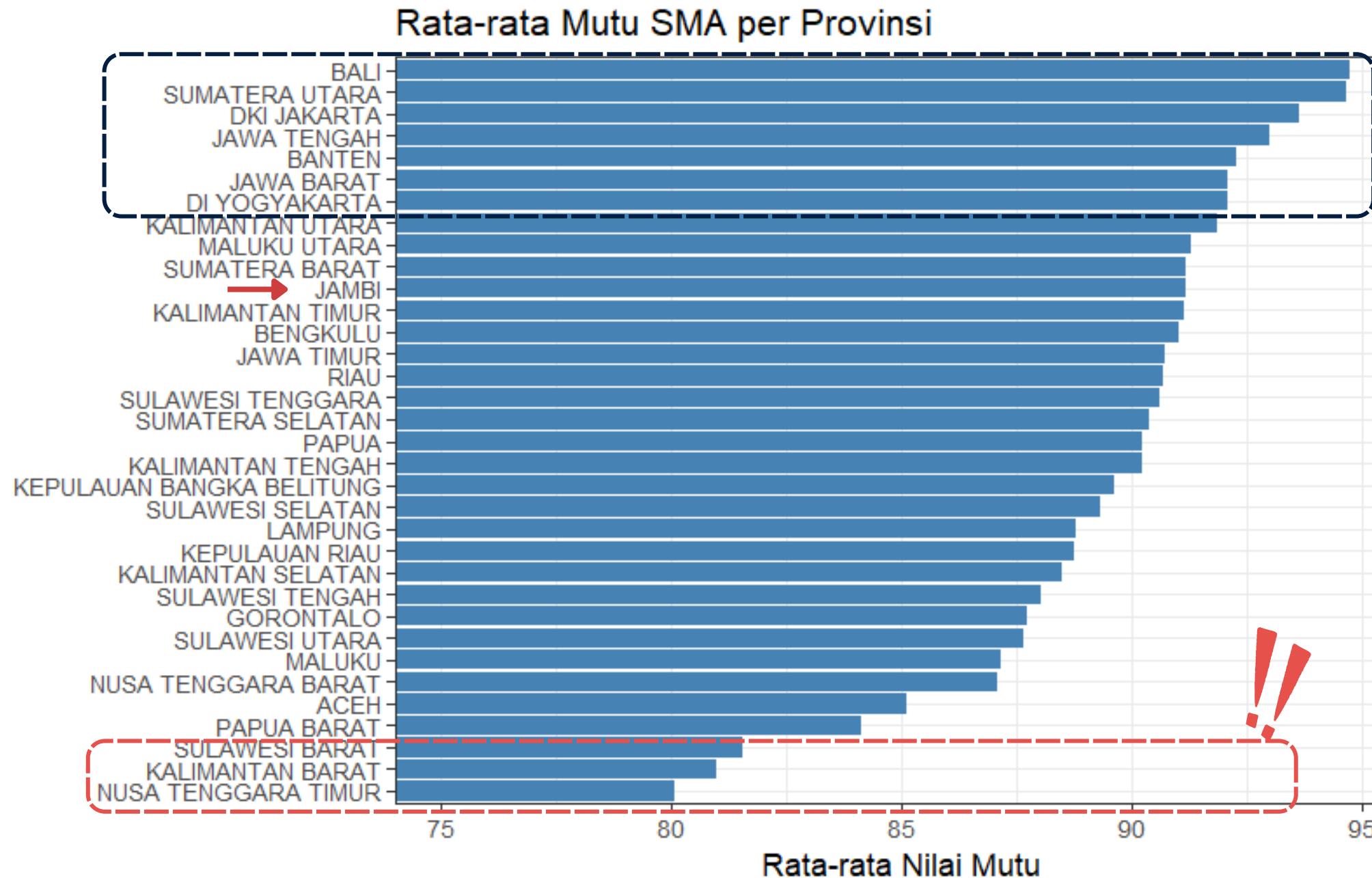


- Bali
  - Banten
  - Bengkulu
  - DKI Jakarta
  - Jambi
  - Jawa Barat
  - Jawa Tengah
  - Kalimantan Tengah
  - Sumatera Utara
  - DI Yogyakarta
- 
- Aceh
  - Gorontalo
  - Kalimantan Selatan
  - Maluku
  - NTB
  - Papua Barat
  - Sulawesi Utara
- 
- Bangka Belitung
  - Jawa Timur
  - Kalimantan Timur
  - Kalimantan Utara
  - Kepulauan Riau
  - Maluku Utara
  - Riau
  - Lampung
  - Papua
  - Sulawesi Tengah
  - Sulawesi Selatan
  - Sulawesi Tenggara
  - Sumatera Selatan
  - Sumatera Barat
- 
- Kalimantan Barat
  - NTT
  - Sulawesi Barat

## Provinsi yang **konsisten** pada kategorinya berdasarkan metode **K-Means** dan **K-Medoids**

Kategori	Provinsi
Sangat Tinggi	Bali, Banten, DKI-Jakarta, Jambi, Jawa Barat, Jawa Tengah, Sumatera Utara, DI Yogyakarta <small>Catatan : Provinsi yang di K-medoid Sangat tinggi juga ada di K-means → K-Mean lebih kompak</small>
Tinggi	Bangka Belitung, Jawa Timur, Kalimantan Timur, Kalimantan Utara, Maluku Utara, Riau, Sumatera Barat, Papua
Sedang	Gorontalo, Kalimantan Selatan, Maluku
Rendah	Kalimantan Barat, NTT, Sulawesi Barat <small>Catatan : Provinsi Rendah yang ada di K-Means juga ada di K-Medoid, artinya K-Medoid lebih sensitif dalam memisahkan</small>

- Kategori ekstrim (Sangat tinggi & Rendah) relatif konsisten
- Kategori menengah (Sedang & Tinggi) cenderung berbeda



3 Provinsi terbawah menunjukkan nilai rata-rata mutu jauh dari provinsi lainnya

## 04 KESIMPULAN

Hasil penggerombolan provinsi dengan K-Means dan K-Medoids menggunakan jarak Euclidean menunjukkan pola umum yang konsisten, terutama pada klaster sangat tinggi (Bali, Banten, DKI Jakarta, Jawa Barat, Jambi dan Jawa Tengah) sedangkan untuk klaster Rendah (Kalimantan Barat, NTT, Sulawesi Barat)



## SARAN

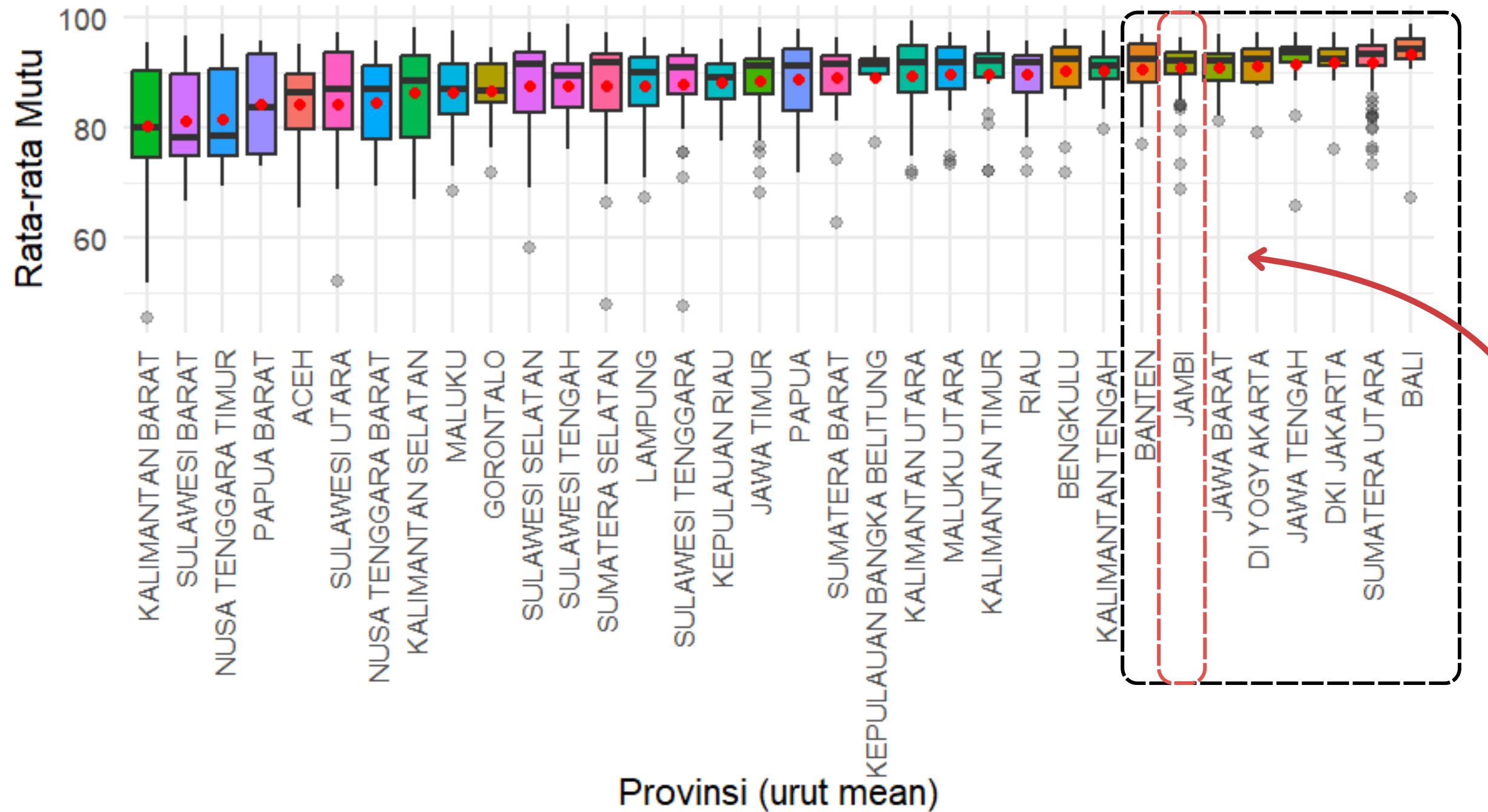
Analisis dilakukan dengan agregasi sekolah berdasarkan nilai **mean** sehingga perlu dilakukan perbandingan dengan agregasi berdasarkan nilai **median**

## RE-MODELLING

Hasil klaster “Sangat Tinggi” menunjukkan indikasi ambiguitas, khususnya pada Provinsi Jambi, sehingga diperlukan analisis lanjutan untuk mengevaluasi keandalan dan validitas segmentasi yang dihasilkan.

## EKSPLORASI RATA-RATA MUTU SEKOLAH PER PROVINSI SEBELUM AGREGASI

Boxplot Rata-rata Mutu Sekolah per Provinsi

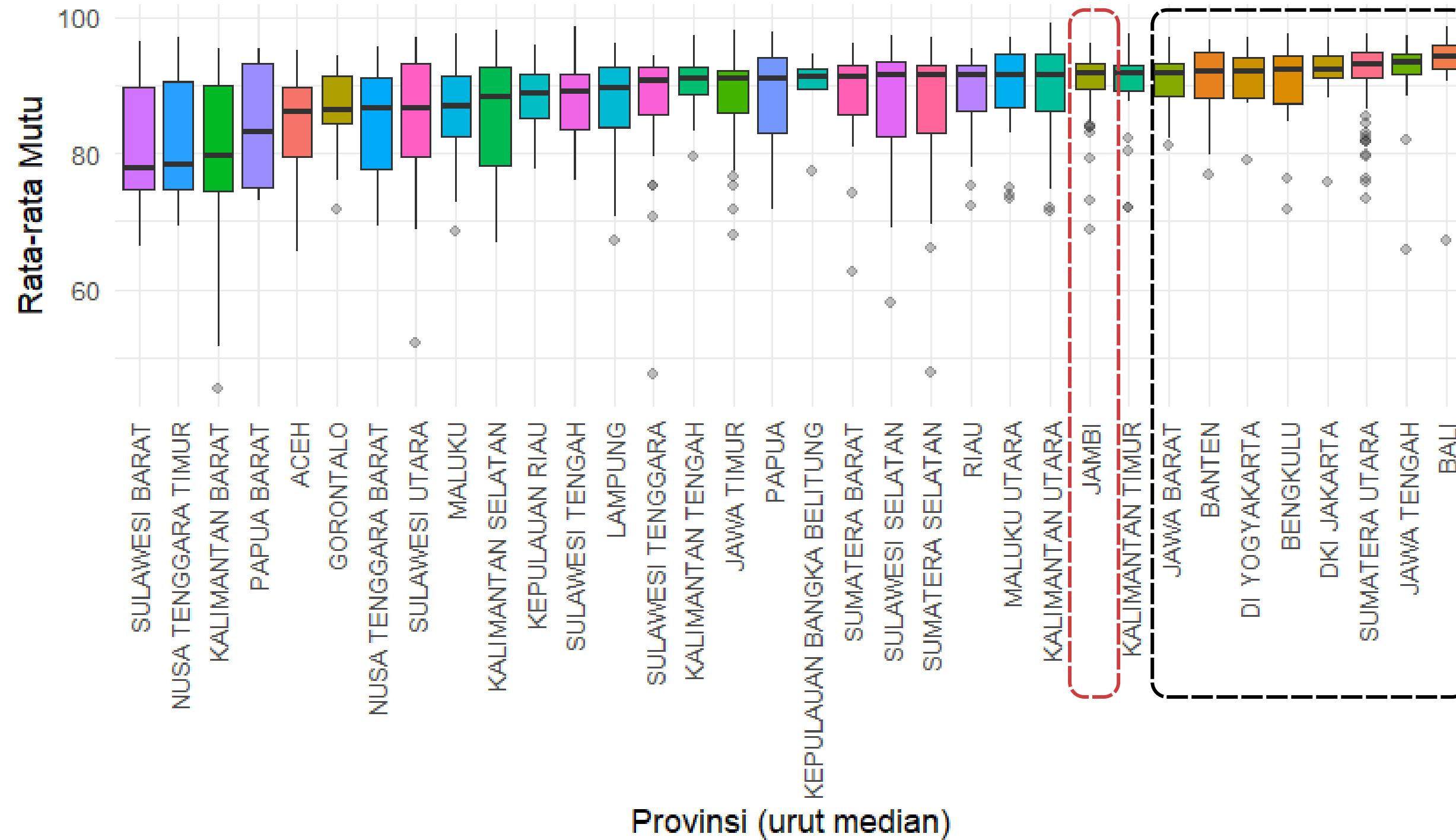


Berdasarkan urutan **mean**, delapan Provinsi teratas sejalan dengan cluster **Sangat Tinggi** pada hasil analisis K-Means dan K-Medoids

Namun, hal ini **tidak sesuai** dengan benchmark mutu pendidikan **Jambi** yang seharusnya tidak termasuk dalam kategori tersebut.

# EKSPLORASI RATA-RATA MUTU SEKOLAH PER PROVINSI SEBELUM AGREGASI

Boxplot Rata-rata Mutu Sekolah per Provinsi

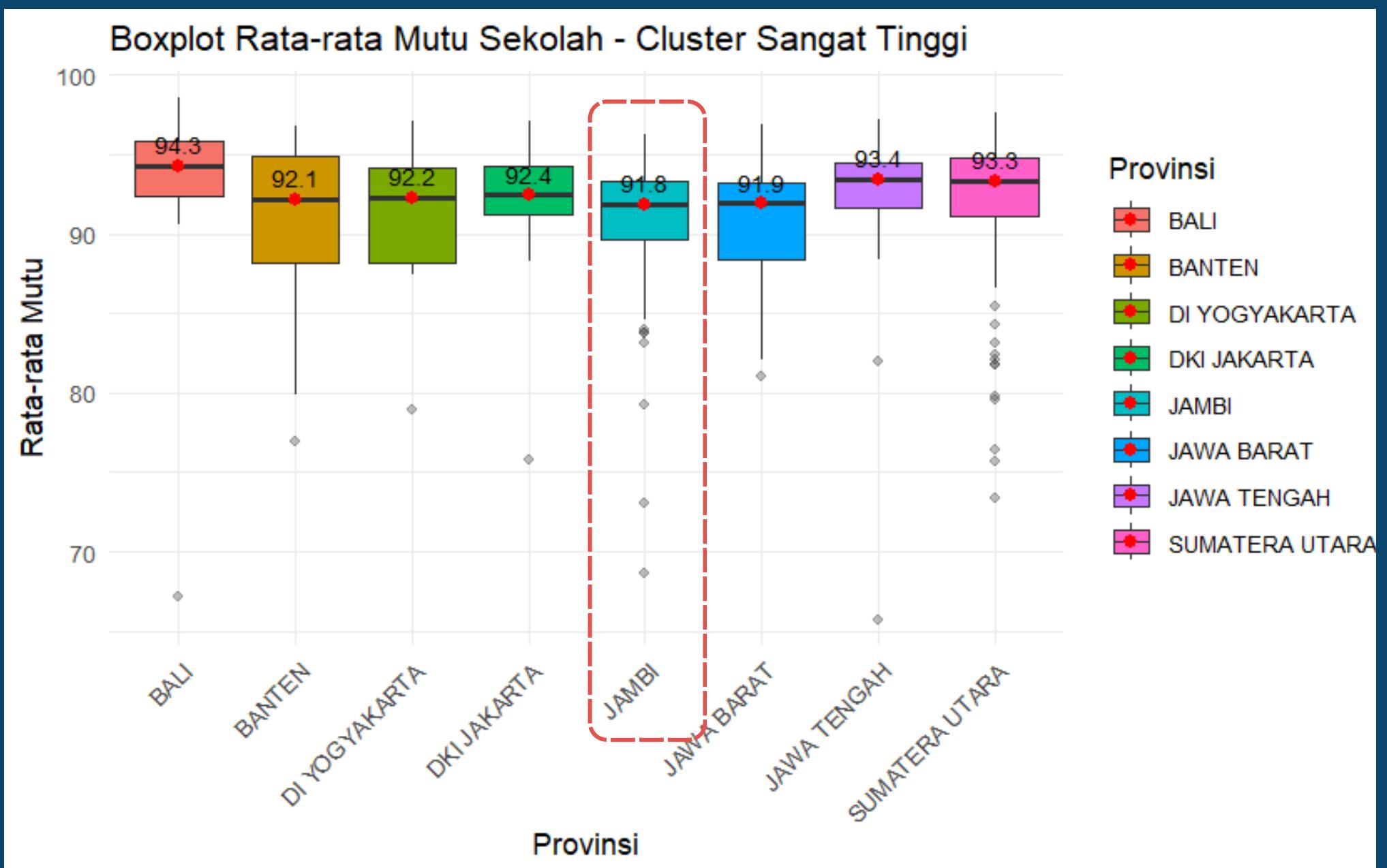


Jika diurutkan berdasarkan **median**, delapan Provinsi yang teratas diisi oleh **Bengkulu** menggantikan posisi Jambi.

Analisis eksplorasi data



## Pengecekan rata-rata mutu sekolah Cluster “Sangat Tinggi”



## Proporsi Outlier

A tibble: 8 × 4

Provinsi	total	outlier_count	prop_outlier
BALI	22	1	4.545455
BANTEN	24	1	4.166667
DI YOGYAKARTA	13	1	7.692308
DKI JAKARTA	14	1	7.142857
JAMBI	91	7	7.692308
JAWA BARAT	67	1	1.492537
JAWA TENGAH	19	2	10.526316
SUMATERA UTARA	111	12	10.810811

Proporsi < 10% → artinya sebagian besar sekolah tidak outlier.  
 Outlier ini hanya **minoritas** (mungkin sekolah terpencil atau kondisi khusus).

- Meskipun Jambi masuk Sangat Tinggi, terdapat banyak sekolah dengan mutu rendah (**outlier ke bawah**). Ini menunjukkan adanya ketimpangan mutu antar sekolah di Jambi.

## Pengecekan Jumlah Sampel yang Tersedia

Provinsi	Sampel yang tersedia	Jumlah Sekolah (2022)	Proporsi
JAMBI	91	239	38.08
KALIMANTAN TIMUR	23	67	34.33
KALIMANTAN UTARA	22	74	29.73
GORONTALO	16	70	22.86
PAPUA	48	256	18.75
JAWA TIMUR	80	459	17.43
KALIMANTAN BARAT	32	205	15.61
BALI	22	163	13.50
SULAWESI TENGGARA	39	313	12.46
LAMPUNG	60	519	11.56
SUMATERA BARAT	37	337	10.98
BENGKULU	16	146	10.96
SUMATERA UTARA	111	1075	10.33
MALUKU UTARA	21	217	9.68
KEPULAUAN RIAU	15	163	9.20
NUSA TENGGARA BARAT	31	343	9.04
ACEH	47	535	8.79
RIAU	39	455	8.57
SULAWESI SELATAN	50	599	8.35
SUMATERA SELATAN	50	612	8.17
MALUKU	23	284	8.10
NUSA TENGGARA TIMUR	46	598	7.69
PAPUA BARAT	10	131	7.63
DI YOGYAKARTA	13	173	7.51
SULAWESI UTARA	17	230	7.39
KALIMANTAN TENGAH	17	233	7.30
KEPULAUAN BANGKA BELITUNG	5	71	7.04
KALIMANTAN SELATAN	17	242	7.02
SULAWESI TENGAH	14	232	6.03
BANTEN	24	603	3.98
JAWA BARAT	67	1711	3.92
DKI JAKARTA	14	492	2.85
JAWA TENGAH	19	854	2.22

Sumber: Pusdatin Kemendikbud Ristek; Badan Pusat Statistik

Jumlah sampel sekolah di Jambi lebih banyak dari provinsi lainnya hingga mencapai 38%

## Proporsi sampel pada Provinsi Cluster “Sangat Tinggi”

Provinsi	Sampel	Jumlah Sekolah	Proporsi
JAMBI	91	239	38.08
BALI	22	163	13.50
SUMATERA UTARA	111	1075	10.33
DI YOGYAKARTA	13	173	7.51
BANTEN	24	603	3.98
JAWA BARAT	67	1711	3.92
DKI JAKARTA	14	492	2.85
JAWA TENGAH	19	854	2.22

- Beberapa provinsi di cluster Sangat Tinggi hanya memiliki sedikit sampel.
- Kondisi ini membuat representasi cluster **kurang stabil** dan mudah dipengaruhi outlier.
- Hal ini juga memungkinkan data sekolah yang tercakup **belum** sepenuhnya **merepresentasikan** seluruh kondisi di provinsi tersebut



Alternatif : Agregasi dengan **Median**

# HASIL CLUSTERING DENGAN AGREGASI MEDIAN



# Agregasi Data

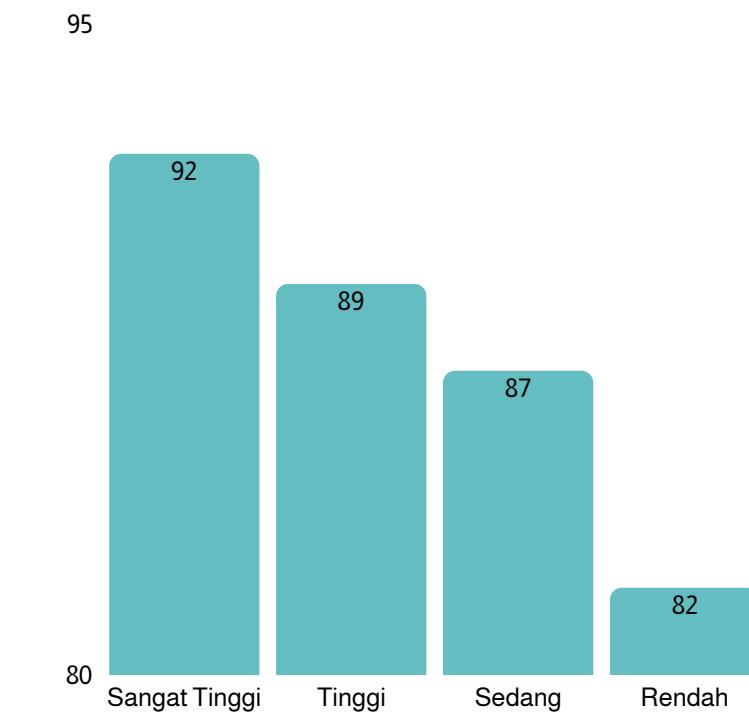
	Provinsi	Mutu Lulusan	Proses Pembelajaran	Mutu Guru	Manajemen S/M
1	ACEH	86.37143	86.30000	81.27778	86.55556
2	BALI	94.87143	93.20000	93.77778	97.11111
3	BANTEN	93.17143	93.20000	87.50000	95.19444
4	BENGKULU	94.31429	88.03333	87.50000	94.22222
5	DI YOGYAKARTA	91.42857	93.20000	87.50000	96.16667
6	DKI JAKARTA	95.45714	91.60000	91.33333	96.16667
7	GORONTALO	88.60000	86.66667	83.33333	92.30556
8	JAMBI	93.17143	89.76667	87.50000	94.22222
9	JAWA BARAT	93.17143	93.20000	87.50000	94.44444
10	JAWA TENGAH	95.45714	93.20000	88.88889	94.44444

Showing 1 to 10 of 34 entries, 5 total columns

- Agregasi tiap sekolah berdasarkan provinsi dengan nilai **median** untuk masing-masing peubah.
- Sehingga setiap provinsi diwakili oleh satu baris yang merupakan **median** seluruh sekolah di provinsi tersebut.

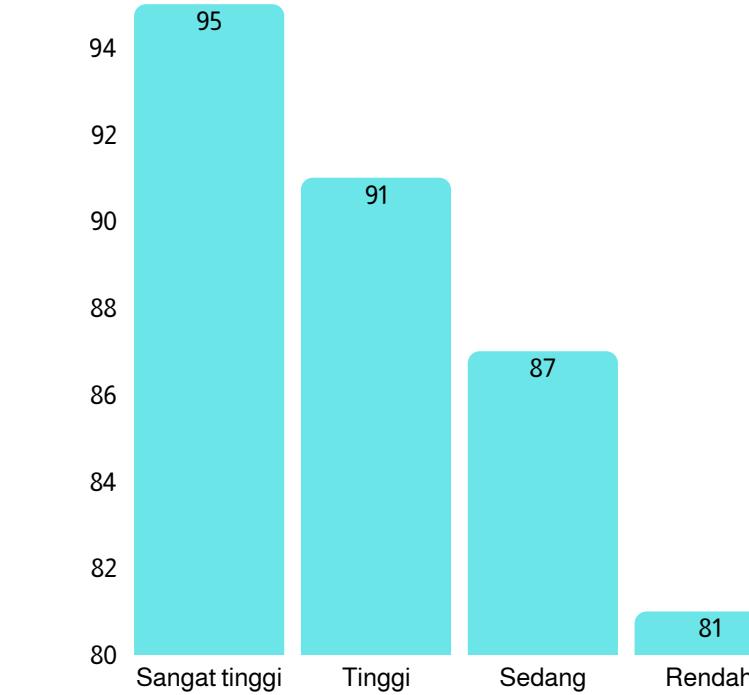
## K-MEAN EUCLIDEAN

Kluster	Mutu Lulusan	Proses Pembelajaran	Mutu Guru	Manajemen S/M	Rata-rata	Ketgori
1	84.3607	80.9500	79.0833	82.3611	82	Rendah
2	93.1143	91.1042	88.8941	94.8976	92	Sangat Tinggi
3	89.1082	89.8000	88.0952	90.7857	89	Tinggi
4	88.9837	87.6167	82.4524	90.8651	87	Sedang



## K-MEDOID EUCLIDEAN

Kluster	Mutu Lulusan	Proses Pembelajaran	Mutu Guru	Manajemen S/M	Rata-rata	Kategori
1	88.6286	86.3000	83.3333	90.3889	87	Sedang
2	95.4571	93.2000	93.7778	96.1667	95	Sangat tinggi
3	91.4286	89.7667	87.5000	94.2222	91	Tinggi
4	82.3429	79.4000	79.5278	82.6944	81	Rendah



**Interpretasi :** Hasil kluster memberikan nilai-nilai tiap peubah yang searah, sehingga dibentuk 4 kategori dari hasil rata-rata 4 peubah

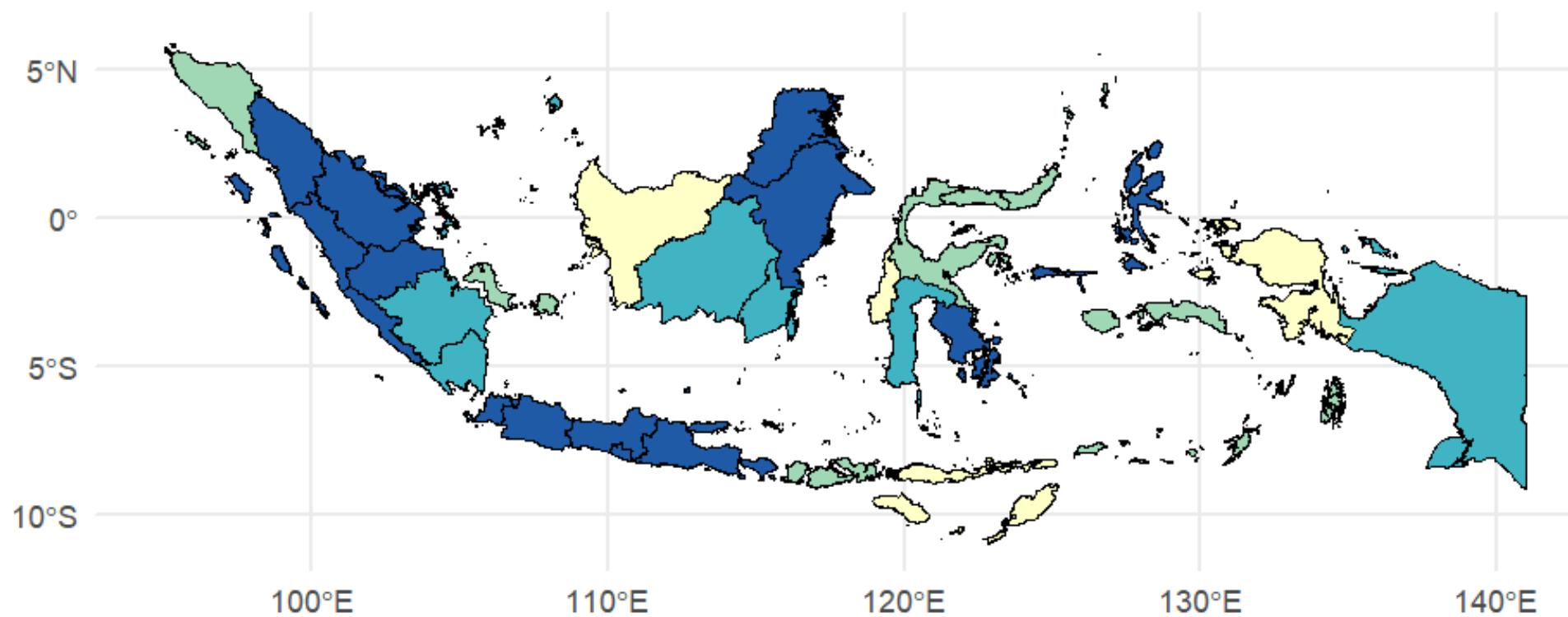
Jarak antar kategori pada K-Mean lebih sempit daripada K-Medoids

# VISUALISASI

## K-MEANS EUCLIDEAN

Peta Klasterisasi Provinsi di Indonesia

Hasil clustering K-means (warna sesuai ranking)

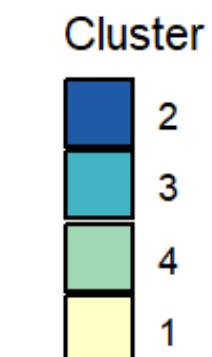


Sangat tinggi

Tinggi

Sedang

Rendah



- Bali
- Banten
- Bengkulu
- DI Yogyakarta
- DKI Jakarta
- Jambi
- Jawa Barat
- Jawa Tengah
- Jawa Timur
- Kalimantan Timur
- Kalimantan Utara
- Maluku Utara
- Riau
- Sulawesi Tenggara
- Sumatera Barat
- Sumatera Utara

- Kepulauan Riau
- Lampung
- Papua
- Sulawesi Selatan
- Sumatera Selatan

- Aceh
- Gorontalo
- Kep. Bangka Belitung
- Maluku
- NTB
- Sulawesi Tengah
- Sulawesi Utara

- Kalimantan Barat
- NTT

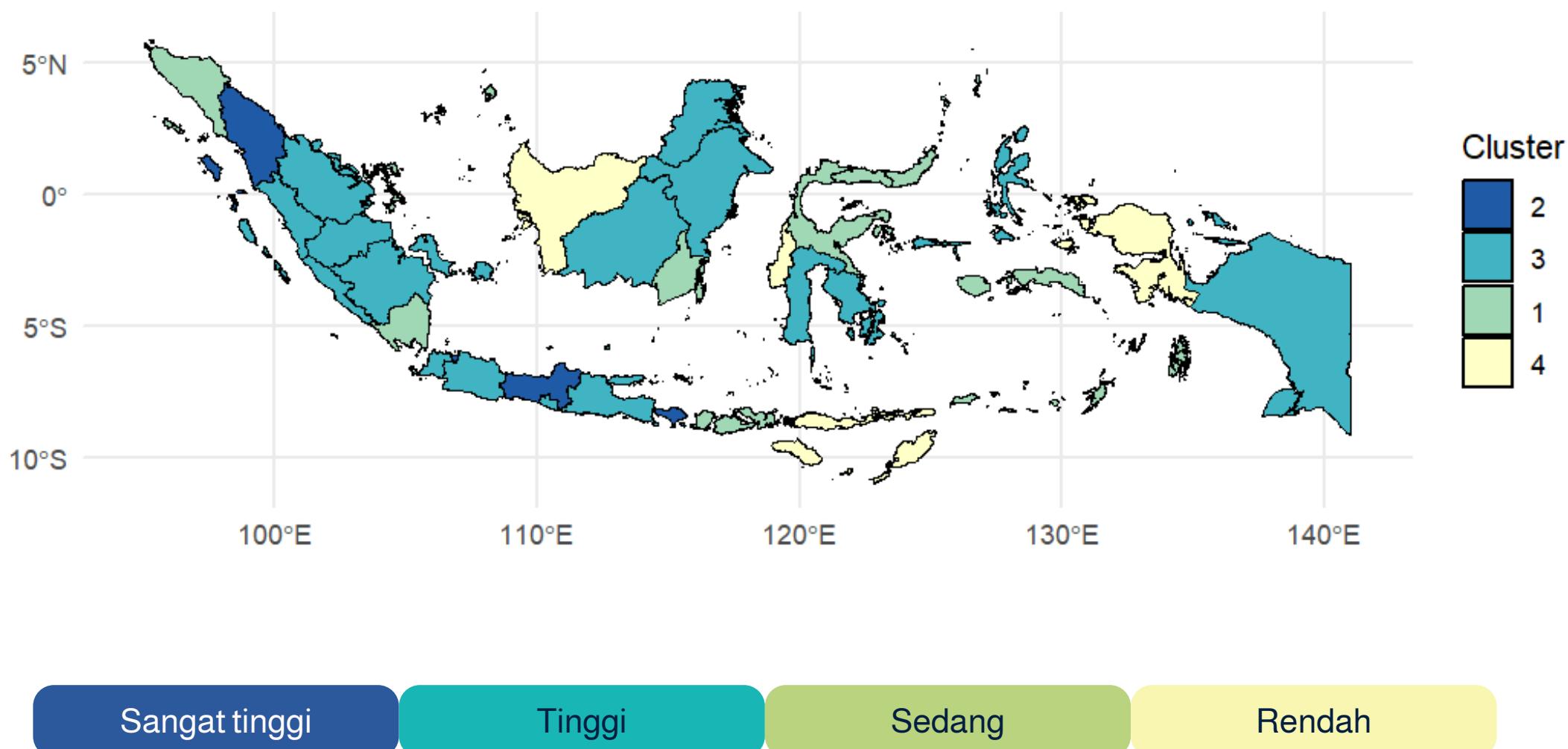
- Papua Barat
- Sulawesi Barat

# VISUALISASI

## K-MEDOIDS EUCLIDEAN

Peta Klasterisasi Provinsi di Indonesia

Hasil clustering K-medoids (warna sesuai ranking)



- Bali
- DKI Jakarta
- Jawa Tengah
- Sumatera Utara

- Aceh
- Gorontalo
- Kalimantan Selatan
- Kepulauan Riau
- Lampung
- Maluku
- NTB
- Sulawesi Tengah
- Sulawesi Utara

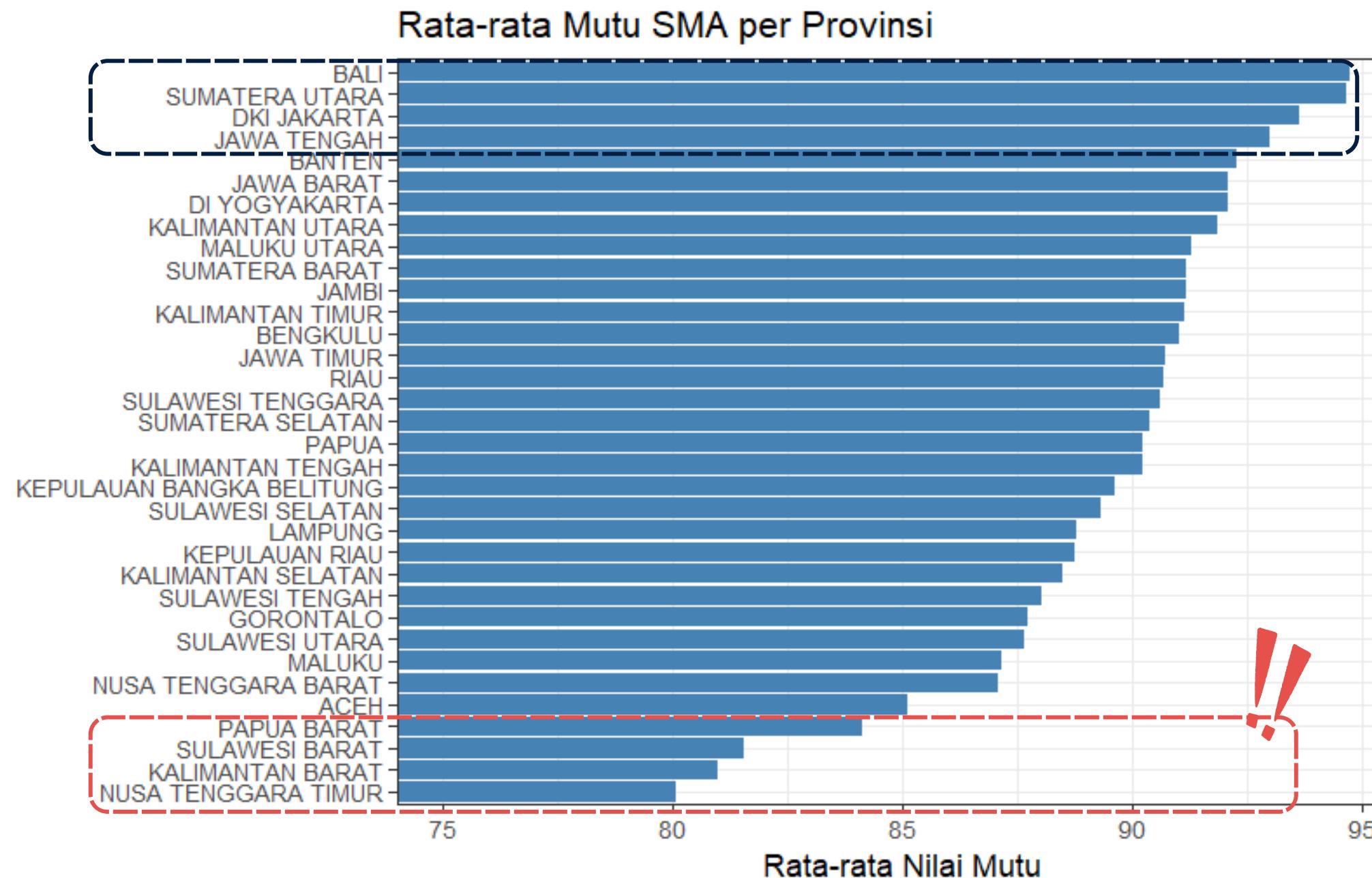
- Kalimantan Barat
- NTT
- Papua Barat
- Sulawesi Barat

- Banten
- Bengkulu
- DI Yogyakarta
- Jambi
- Jawa Barat
- Jawa Timur
- Kalimantan Tengah
- Kalimantan Timur
- Kalimantan Utara
- Kep. Bangka Belitung
- Kepulauan Riau
- Maluku Utara
- Papua
- Riau
- Sulawesi Selatan
- Sulawesi Tenggara
- Sumatera Barat
- Sumatera Selatan

## Provinsi yang **konsisten** pada kategorinya berdasarkan metode **K-Means** dan **K-Medoids**

Kategori	Provinsi
Sangat Tinggi	Bali, DKI-Jakarta, Jawa Tengah, Sumatera Utara
Tinggi	Kalimantan Tengah, Papua, Sulawesi Selatan, Sumatera Selatan
Sedang	Aceh, Gorontalo, Maluku, NTB, Sulawesi Tengah, Sulawesi Utara,
Rendah	Kalimantan Barat, NTT, Papua Barat, Sulawesi Barat

- Kategori ekstrim (Sangat tinggi & Rendah) relatif konsisten
- Kategori menengah (Sedang & Tinggi) cenderung berbeda



3 Provinsi terbawah menunjukkan nilai rata-rata mutu jauh dari provinsi lainnya

# EVALUASI MODEL

Metode Cluster	Jarak	Silhouette	DBI
K-Means	Euclidean (k=4)	0.3489	0.9516
K-Medoids	Euclidean (k=4)	0.3436	0.9025

- Nilai Silhouette berkisar -1 sampai 1 → Semakin tinggi semakin baik
- Nilai DBI memiliki nilai  $\geq 0$ , → Semakin kecil semakin baik

- K-Means Euclidean memiliki nilai Silhouette tertinggi berarti cluster antar provinsi terpisah lebih jelas.
- K-Medoids Euclidean memiliki nilai DBI terendah yang berarti cluster terbentuk relatif kompak.

## 04 KESIMPULAN

Terjadi pergeseran anggota klaster jika menggunakan agregasi dengan **median**. Provinsi dengan klaster “Sangat Tinggi” ditempati oleh Bali, DKI-Jakarta, Jawa Tengah, dan Sumatera Utara. Sedangkan klaster “Rendah” ialah Kalimantan Barat, NTT, Papua Barat, Sulawesi Barat.



## IMPLIKASI

Sekolah-sekolah yang berada pada klaster mutu rendah menjadi sinyal penting bahwa pemerintah perlu fokus pada daerah-daerah ini untuk mendorong pemerataan kualitas pendidikan.

# THANK YOU

-  [RPubs - K-Means and K-Medoids \(Euclidean\).](#)
-  [Nabila Syukri | LinkedIn](#)

