



# Pemetaan dan Pengelompokan Provinsi di Indonesia berdasarkan Statistik Infrastruktur Indonesia 2020

Andika Lumban Toruan<sup>1</sup>, Winner De Bofan<sup>2</sup>, Lazarus Rinaldi K. T.<sup>3</sup>, M. Ilham Hakiqi<sup>4</sup>,  
Fadhlan Musyaffa<sup>5</sup>, M. Arzei Z. Y.<sup>6</sup>, Irham Khairul K.<sup>7</sup>, Toni Toharudin<sup>8</sup>  
Universitas Padjadjaran<sup>1,.....8</sup>  
[andika19007@mail.unpad.ac.id](mailto:andika19007@mail.unpad.ac.id)

## Abstract

**Abstrak.** Pembangunan infrastruktur adalah perwujudan dari pembangunan nasional yang merupakan salah satu tujuan negara Indonesia. Dalam Undang-Undang Dasar Negara Republik Indonesia Tahun 1945 dirumuskan bahwa pembangunan nasional merupakan rangkaian usaha pembangunan yang berkepanjangan demi kehidupan masyarakat, bangsa, dan negara. Namun pada faktanya pembangunan infrastruktur di Indonesia belum merata untuk semua provinsi. Oleh karena itu perlu dilakukan pembenahan agar pembangunan nasional berhasil, salah satunya dengan mengatasi masalah terkait ketidakmerataan infrastruktur di Indonesia. Penelitian ini bertujuan untuk mengelompokkan dan memetakan provinsi di Indonesia berdasarkan data infrastruktur pada tahun 2020. Pada penelitian ini provinsi di Indonesia akan diposisikan dalam ruang 2 dimensi berdasarkan 8 variabel yang ada pada data infrastruktur pada tahun 2020 dengan menggunakan metode *multidimensional scaling*. Provinsi-provinsi di Indonesia juga dibagi menjadi beberapa kelompok dengan menggunakan analisis kluster. Hasil yang diperoleh dari analisis ini adalah pengelompokan provinsi di Indonesia ke dalam 3 kelompok sesuai dengan statistik infrastruktur setiap provinsi yang berperan sebagai karakteristik dari masing-masing kelompok tersebut. Klaster 1 termasuk kedalam kategori tinggi yang terdiri dari 3 provinsi, klaster 2 termasuk kedalam kategori sedang yang terdiri dari 7 provinsi, dan klaster 3 termasuk kedalam kategori rendah yang terdiri dari 24 provinsi. Didapatkan juga nilai stress yang diperoleh adalah 0.0000212% yang berarti konfigurasi objek sebagai titik dalam dimensi sudah sempurna

*Kata kunci: infrastruktur; analisis Kluster; Stress; Multidimensional Scalling*

## I. PENDAHULUAN

Pengertian Infrastruktur menurut *American Public Works Association* (Stone, 1974 dalam Kodoatie, R., 2005), adalah fasilitas-fasilitas fisik yang dikembangkan atau dibutuhkan oleh agen-agen public untuk fungsi-fungsi pemerintahan dalam penyediaan air, tenaga listrik, pembuangan limbah, transportasi dan pelayanan untuk memfasilitasi tujuan-tujuan social dan ekonomi. Dari pengertian tersebut dapat diartikan bahwa infrastruktur sangat penting disuatu negara.

Menurut *Global competitiveness Index 4.0* pada tahun 2019 menempatkan Indonesia pada peringkat 50 dari 141 negara dalam bidang infrastruktur. Dengan peringkat itu, Indonesia turun 5 peringkat dan peringkat tersebut masih kalah dengan negara tetangga seperti Singapura, Malaysia dan Thailand. Oleh karena itu pembangunan dalam infrastruktur perlu ditingkatkan agar peringkat menyusul dari negara tetangga.

Pembangunan infrastruktur adalah perwujudan dari pembangunan nasional yang dimana merupakan salah satu tujuan negara Indonesia yang telah dirancang dalam pembukaan UUD tahun 1945, bahwa pembangunan nasional merupakan rangkaian usaha pembangunan yang berkepanjangan pada aspek-aspek kehidupan masyarakat, bangsa, dan negara.

Dengan dikenalnya Indonesia sebagai negara yang luas, berkepulauan, dan penduduk yang banyak maka pemerintah Indonesia sedang masif dalam pembangunan infrastruktur dalam membangun pertumbuhan ekonomi dan lain-lain. Indonesia juga mempunyai rencana dalam pembangunan jangka Panjang. Hal ini sesuai dengan visi pembangunan nasional tahun 2005 sampai tahun 2025 yang wajib dicapai yakni "Indonesia mandiri, maju, makmur dan adil sesuai yang tertuang dalam di dalam pembukaan perundangan Dasar Negara Republik Indonesia Tahun 1945 (Waluyo, 2007:36)





Akan tetapi, pada kenyataannya pembangunan infrastruktur di Indonesia masih belum merata. Meskipun infrastruktur di Pulau Jawa dan Luar Jawa sudah cukup banyak, tetapi di beberapa daerah pembangunan di Sumatera, Kalimantan, Sulawesi, hingga Papua belum merata. Menurut Direktur Jenderal Perimbangan Keuangan Kementerian Keuangan, Astera Primanto Bhakti, salah satu penyebab tidak meratanya pembangunan infrastruktur ini adalah pemerintah pusat kesulitan melakukan kontrol terhadap penggunaan dana Anggaran Pendapatan dan Belanja Daerah (APBD). Primanto menjelaskan APBD menjadi tanggung jawab masing-masing kepala daerah. Berdasarkan pengamatannya, dana daerah lebih banyak digunakan untuk belanja pegawai bukan program kegiatan yang bisa membangkitkan pembangunan daerah. Jadi pembangunan infrastruktur di setiap daerah bergantung kepada keputusan kepala daerah dalam merancang dan mengelola APBD. Akibatnya, pembangunan daerah memiliki dampak yang berbeda di setiap wilayahnya.

Pemerintah Indonesia perlu mengklasifikasikan setiap daerah di Indonesia menjadi beberapa kluster menurut tingkat kemajuan infrastrukturnya. Pembagian kluster ini penting dalam memberikan perlakuan kebijakan, agar pemerintah tahu daerah mana saja yang menjadi prioritas dalam membangun infrastruktur. Pemerintah pusat harus memperhatikan dan mendukung daerah menengah serta tertinggal dalam memutuskan perencanaan dan penggunaan anggaran daerah.

Sejalan dengan itu, tujuan dari penelitian ini adalah untuk mengklasifikasikan setiap provinsi di Indonesia menjadi beberapa kluster berdasarkan tingkat kemajuan infrastrukturnya. Penelitian sejenis ini masih sangat jarang dilakukan di Indonesia sehingga diharapkan penelitian ini dapat memberikan manfaat kepada pihak-pihak yang membutuhkan, yaitu agar pemerintah tahu daerah-daerah yang diprioritaskan dalam pembangunan infrastruktur.

## II. METODE PENELITIAN

### Data Penelitian

Pada penelitian ini data yang digunakan adalah data sekunder yang diambil dari publikasi statistic Oleh BPS Indonesia, yaitu statistik Infrastruktur Indonesia Tahun 2020

### Variabel Penelitian

Pada penelitian ini variabel yang akan diteliti adalah variabel yang dapat mencerminkan kualitas dari infrastruktur pada provinsi tersebut, sehingga dari variabel tersebut dapat melihat baik atau buruknya kualitas infrastruktur di provinsi tersebut. Variabel yang diperoleh terdapat 8 variabel yaitu:

1. Wilayah Administrasi
2. Fasilitas Pendidikan
3. Fasilitas Kesehatan
4. Bank dan koperasi
5. Hotel dan Penginapan
6. Pasar
7. BUMDes
8. Embung

### Metode Analisis Data

#### 1. Analisis Multidimensional Scaling

Analisis Multidimensional Scalling (MDS) merupakan salah satu teknik peubah ganda yang dapat digunakan untuk menentukan posisi suatu obyek lainnya berdasarkan penilaian kemiripannya, juga untuk mengetahui hubungan interdependensi atau saling ketergantungan antar variabel atau data (Johnson, 1992). Hubungan ini tidak diketahui melalui reduksi ataupun pengelompokan variabel, melainkan dengan membandingkan variabel yang ada pada setiap obyek yang bersangkutan dengan menggunakan perceptual map. MDS berhubungan dengan pembuatan peta untuk menggambarkan posisi sebuah obyek dengan obyek lainnya berdasarkan kemiripan obyek-obyek tersebut. MDS juga merupakan teknik yang bisa membantu peneliti untuk mengidentifikasi dimensi kunci yang mendasari evaluasi objek dari responden.

Multidimensional scalling digunakan untuk mengetahui hubungan interdependensi atau saling ketergantungan antara variabel atau data. Hubungan ini tidak diketahui melalui reduksi maupun pengelompokan variabel, melainkan dengan membandingkan variabel yang ada pada setiap objek yang bersangkutan dengan menggunakan *perceptual map*. Konsep dasar Multidimensional Scalling adalah pemetaan.





Jenis – Jenis Analisis Multidimensional Scalling (MDS) sebagai berikut:

1. Multidimensional Scalling Metrik

Data jarak yang digunakan pada analisis ini adalah data rasio. Penskalaan ini digunakan untuk menemukan himpunan titik dalam ruang dimensi  $n$  di mana masing-masing titik mewakili satu objek sehingga jarak antar titik menjadi  $d_{rt} \approx f(d_{rt})$

2. Multidimensional Scalling Non Metrik

Data jarak yang digunakan dalam penskalaan ini adalah data berskala ordinal. Dalam analisis ini, fungsi transformasi hanya mempunyai batasan  $\delta_{rt} < \delta_{r't'} \rightarrow f(\delta_{rt}) \leq f(\delta_{r't'})$  untuk semua  $1 r, t, r', t', \leq n$ . Suatu fungsi STRESS (*Standardized Residual Sum of Square*) sebagai berikut:

$$Stress = \sqrt{\frac{\sum_{r,t} (d_{rt} - \hat{d}_{rt})^2}{\sum_{r,t} d_{rt}^2}}$$

**Tabel 1.** Kriteria model pada MDS

Stress (%)	Kriteria Model Multidimensional Scalling
>20%	Kurang Baik
20-10%	Cukup
10-5%	Baik
5-2,5%	Sangat Baik
<2,5%	Sempurna

Prosedur Analisis Multidimensional Scalling (MDS)

1. Merumuskan masalah
2. Memperoleh dan input data
3. Menentukan prosedur penskalaan multidimensional
4. Memberi label nama dimensi
5. Menentukan banyak dimensi

Untuk menentukan posisi atau koordinat stimulus dari objek objek yang diteliti dengan menggunakan algoritma multidimensional scalling adalah sebagai berikut:

1. Membentuk sebuah matriks jarak (D)
2. Menghitung kuadrat dari matriks D yang disebut menghitung  $D^2$
3. Menentukan matriks B dengan menggunakan proses *double centring* :

$B = -\frac{1}{2}JD^2J$  dengan  $J = 1 - \frac{1}{n}A$  yang mana A adalah matriks yang semua elemennya adalah 1 dan  $n$  adalah jumlah objek

4. Diperoleh mutlak nilai eigen 2 terbesar pada B serta  $m$  vector eigen
5. Diperoleh koordinat matriks  $F = E_m \Lambda_m^{\frac{1}{2}}$  dimana  $E_m$  adalah matriks dari vector eigen  $m$  dan  $\Lambda_m$  adalah matriks diagonal dari masing-masing  $m$  vector eigen matriks B





## 2. Analisis Kluster

Analisis kluster yang digunakan adalah metode *K-Means Clustering* yang mana merupakan salah satu metode dari analisis kluster non-hierarki. Metode *K-Means Clustering* adalah metode paling umum yang digunakan untuk mengelompokkan data pengamatan menjadi kluster dan pada tiap pengamatannya termasuk dalam kluster dengan rata-rata terdekat. Pada intinya, metode ini mengklasifikasikan objek dengan meminimalkan jumlah kuadrat jarak antara data dan sentroid kluster yang sesuai. Langkah pertama yang harus dilakukan adalah memilih variabel dalam data yang akan digunakan untuk pengelompokan, lalu melihat skala dan satuan dalam variabel yang diteliti. Oleh karena itu, perlu untuk melakukan proses standarisasi jika variabel yang diteliti satuannya berbeda. Selanjutnya, melakukan pengujian untuk memenuhi asumsi bahwa sampel harus mencerminkan populasi (representatif) yaitu dengan menggunakan uji Keizer Meyer Olkin (KMO), dan asumsi tidak terjadinya multikolinieritas yaitu dengan menggunakan uji bilangan kondisi (k). Apabila nilai KMO berada di antara 0.5-1, maka sampel dapat dikatakan mewakili populasi, dan apabila nilai  $k < 100$  pada bilangan kondisi, maka dapat dikatakan terjadi multikolinieritas yang cukup lemah hingga sangat lemah antar variabelnya, sehingga analisis dapat dilanjutkan. Namun, apabila terjadi multikolinieritas, maka harus melakukan penanganan dengan metode *Principal Component Analysis* (PCA) untuk mengatasi hal tersebut (Soemartini, 2008). Langkah selanjutnya adalah analisis kluster *K-Means* dengan menggunakan jarak Euclidean dan menentukan terlebih dahulu banyaknya kluster yang diinginkan dan mengalokasikan data sesuai dengan jumlah kluster yang ditentukan. Setelah itu, tentukan nilai sentroid pada tiap klusternya berdasarkan nilai *range* yang berada pada sumber data dengan melakukan pemilihan sesuai dengan sentroid yang dipilih. Setelah memilih nilai sentroid, langkah selanjutnya adalah menghitung jarak terdekat dengan menggunakan jarak Euclidean yang memiliki rumus sebagai berikut:

$$d(x_i, y_j) = \sqrt{\sum (x_i - \mu_j)^2}$$

Keterangan:

$x_i$  = data kriteria

$\mu_j$  = *centroid* pada kluster ke-j

Kemudian, tampilkan hasilnya berdasarkan jarak terendah dari hasil perhitungan tersebut dan apabila hasilnya masih belum sesuai, maka dilakukan iterasi kembali sampai hasil klusterisasinya sudah sesuai dengan iterasi sebelumnya. Langkah terakhir yang harus dilakukan adalah mempresentasikan hasil analisis deskriptif dan analisis klusternya berdasarkan data yang digunakan.

## III. HASIL DAN PEMBAHASAN

### a. Analisis Multidimensional Scaling

Tabel 2. Kordinat Stimulus

Provinsi	Dimensi	
	1	2
Aceh	-133.4405	-1094.4773
Sumatera Utara	-9767.0627	-2218.3986
Sumatera Barat	2011.9477	507.6288
Riau	1181.3157	-524.7730
Jambi	3962.5713	-177.2114
Sumatera Selatan	-283.2647	-1003.4854
Bengkulu	6390.5595	84.9501
Lampung	-585.4847	-1530.7822
Kep. Bangka Belitung	7812.3557	442.6165
Kepulauan Riau	7119.2090	1018.9682
DKI Jakarta	2619.2710	1288.5041
Jawa Barat	-32690.5650	-4885.7685





Jawa Tengah	-29150.4650	4772.8611
DI Yogyakarta	5468.1074	1402.4977
Jawa Timur	-39155.7830	2325.1264
Banten	-1370.3228	-1863.7260
Bali	3576.8856	6316.7651
Nusa Tenggara Barat	1601.7653	-539.2564
Nusa Tenggara Timur	-248.1502	-736.9134
Kalimantan Barat	1289.2232	-506.5181
Kalimantan Tengah	4398.5116	-276.9520
Kalimantan Selatan	3589.9941	-362.0765
Kalimantan Timur	5175.3525	826.4788
Kalimantan Utara	8326.2960	379.7355
Sulawesi Utara	5010.7588	-71.5530
Sulawesi Tengah	3760.1229	-462.9235
Sulawesi Selatan	-2822.3993	-1846.4464
Sulawesi Tenggara	4445.1984	-256.7565
Gorontalo	7462.3010	156.1334
Sulawesi Barat	6898.7398	-227.8197
Maluku	5802.3437	-279.9818
Maluku Utara	6490.4166	-196.0727
Papua Barat	7278.7466	53.9249
Papua	4534.9452	-514.2978

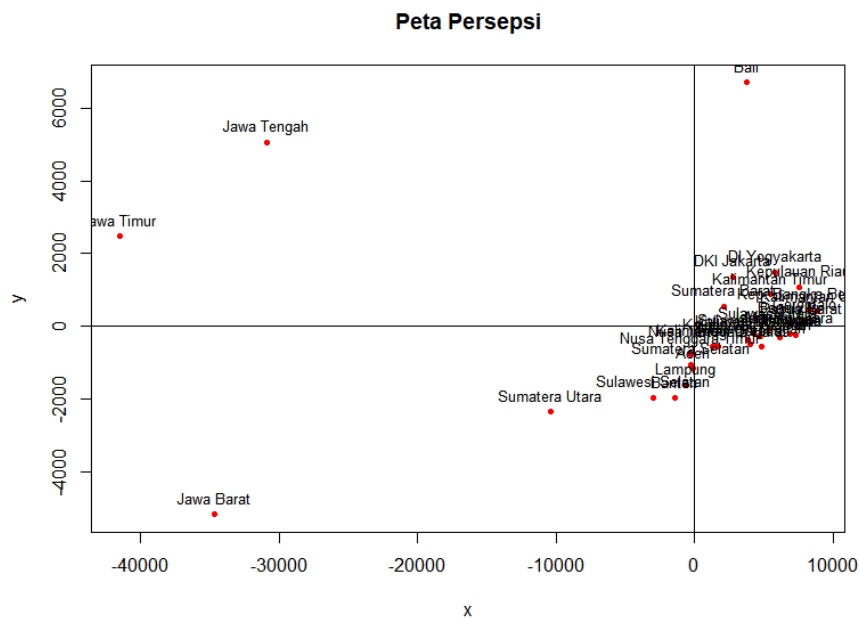
### **Koordinat Stimulus**

Koordinat stimulus adalah titik-titik koordinat dari setiap provinsi yang akan digambarkan dengan peta konfigurasi berdimensi 2. Koordinat stimulus diperoleh dengan mengkalikan vector eigen dengan akar dari nilai eigen.

### **Peta Persepsi**

Untuk membuat peta persepsi dapat menggunakan koordinat stimulus pada tabel 1 di atas, dari koordinat stimulus diperoleh peta persepsi sebagai berikut:





**Gambar 1.** Peta Persepsi hasil MDS

Dapat dilihat bahwa banyak provinsi yang memiliki jarak yang cukup dekat tetapi berada pada kuadran yang berbeda. Hal ini menunjukkan bahwa provinsi-provinsi yang berdekatan mempunyai karakteristik yang cukup mirip, tetapi dari peta persepsi tersebut belum dapat ditentukan kelompok untuk masing-masing provinsi. Selanjutnya dilakukan analisis kluster untuk melakukan pengelompokan.

**Tabel 3.** Nilai Stress dan Goodness of Fit

Stress (%)	GoF (%)
0.0000212	96.37

**Kebaikan Hasil MDS**

Sebelum melanjutkan ke analisis kluster sebaiknya melihat dahulu nilai stress untuk melihat kebaikan hasil MDS. Nilai stress adalah suatu ukuran yang digunakan untuk menilai suatu konfigurasi dari objek sebagai titik-titik dalam dimensi q sudah baik atau belum. Lalu diperoleh nilai stress yang dapat dilihat pada tabel 2 yaitu sebesar 0.0000212% yang dimana hasil tersebut merupakan kriteria sempurna berdasarkan tabel kriteria model MDS yang artinya konfigurasi dari objek sebagai titik-titik dalam dimensi sudah sempurna

**b. Analisis kluster**

**Asumsi Multikolinearitas**

Multikolinearitas dicari dengan menggunakan bilangan kondisi (k). Rumusnya adalah sebagai berikut:

$$k = \frac{\lambda_{max}}{\lambda_{min}}$$

Dimana  $\lambda$  adalah nilai eigen dari matriks kovarians atau korelasi variabel. Berikut merupakan kriteria batas bilangan kondisi (k) untuk melihat apakah terdapat multikolinearitas:

**Tabel 4.** Tabel kriteria multikolinearitas

Batas	Multikolinearitas
$k < 100$	ti Multikolinearitas lemah
$100 \leq k \leq 1000$	ti Multikolinearitas sedang hingga kuat
$k > 1000$	ti Multikolinearitas kuat





**Asumsi Kecukupan Sampel dengan Uji KMO (Kaiser Meyer Olkin)**

Uji KMO digunakan untuk mengetahui apakah data observasi layak dan dapat dianalisis dengan analisis komponen utama. Nilai statistik KMO yang digunakan untuk mengukur kecukupan samplingnya dapat dihitung dengan rumus sebagai berikut:

$$KMO = \frac{\sum_{i \neq j} r_{ij}^2}{\sum \sum_{i \neq j} r_{ij}^2 + \sum \sum_{i \neq j} a_{ij}^2}, i = 1, 2, \dots, p; k = 1, 2, \dots, p$$

Keterangan:

$r_{ij}$  : koefisien korelasi sederhana antara variabel ke-i dan ke-j

$a_{ij}$  : koefisien korelasi parsial antara variabel ke-i dan ke-j

Nilai KMO yang kecil mengindikasikan penggunaan analisis komponen utama harus dipertimbangkan kembali karena akorelasi antar variabel tidak dapat dijelaskan oleh variabel lain. Adapun kriteria keputusan dari KMO adalah sebagai berikut:

**Tabel 5.** Nilai KMO Kelayakan Komponen

Nilai KMO	Kelayakan Komponen
0.90 – 1.00	Sangat Baik
0.80 – 0.90	Baik
0.70 – 0.80	Lebih dari cukup
0.60 – 0.70	Cukup
0.50 – 0.60	Agak Cukup
0.00 – 0.50	Tidak Layak

**Analisis Komponen Utama**

Dengan analisis komponen utama, didapatkan bahwa hanya dengan menggunakan 2 komponen utama yang baru sudah dapat menjelaskan sebanyak 86.22% dari keseluruhan varians. Fungsi komponen utama tersebut adalah sebagai berikut:

$$KU_1 = -0.359X_1 - 0.389X_2 - 0.386X_3 - 0.375X_4 - 0.255X_5 - 0.391X_6 - 0.365X_7 - 0.282X_8$$

$$KU_2 = -0.286X_1 + 0.043X_2 - 0.026X_3 + 0.213X_4 + 0.824X_5 - 0.022X_6 - 0.213X_7 - 0.383X_8$$

Data hasil analisis komponen utama direduksi menjadi 2 komponen utama, yaitu adalah sebagai berikut:

**Tabel 6.** Data baru hasil reduksi variabel

Provinsi	KU1	KU2
Aceh	-9539.17	-2636.32
Sumatera Utara	-13568	135.574
Sumatera Barat	-5171.91	837.479
Riau	-5976.92	-48.778
Jambi	-3949.46	-98.824
Sumatera Selatan	-7602.1	-722.443
Bengkulu	-2609.86	-363.761
Lampung	-7107.29	-599.314
Kep. Bangka Belitung	-1243.45	73.354
Kepulauan Riau	-1683.57	613.374
DKI Jakarta	-4137.94	1712.85
Jawa Barat	-26550.2	3736.969
Jawa Tengah	-28433.9	2818.9
DI Yogyakarta	-3050.46	1910.105
Jawa Timur	-33617.1	4172.804
Banten	-6337.27	528.586
Bali	-6109.12	6972.52
Nusa Tenggara Barat	-5201.4	520.355
Nusa Tenggara Timur	-7891.16	-1437.07







Kalimantan Barat	-5440.48	256.236
Kalimantan Tengah	-3789.08	-29.369
Kalimantan Selatan	-4483.25	-254.004
Kalimantan Timur	-3448.07	562.452
Kalimantan Utara	-881.496	25.974
Sulawesi Utara	-3496.36	-109.998
Sulawesi Tengah	-4363.23	-83.793
Sulawesi Selatan	-8969.6	-697.001
Sulawesi Tenggara	-4211.62	-435.83
Gorontalo	-1557.95	-176.747
Sulbar	-1674.59	-80.066
Maluku	-2612.13	-29.315
Maluku Utara	-2101.62	-100.019
Papua Barat	-1872.69	-203.106
Papua	-4576.32	-1081.14

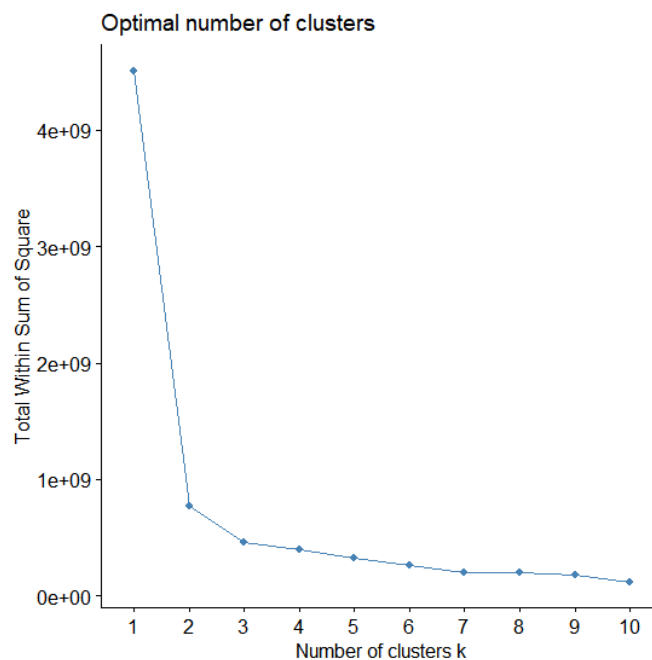
Dilakukan lagi uji multikolinieritas pada komponen utama lalu diperoleh nilai eigen dari matriks korelasi untuk masing-masing variabel komponen utama yaitu sebesar 1.4945545 dan 0.5054455. maka didapatkan bilangan kondisi (k) adalah:

$$k = \frac{1.4945545}{0.5054455} = 2.96$$

Didapatkan bilangan kondisi (k) sebesar 2.96 yang berarti tidak ada lagi masalah multikolinieritas pada data setelah analisis komponen utama. Maka selanjutnya dapat dilakukan clustering.

### Analisis K-Means Clustering

Analisis cluster k-means dimulai dengan menentukan jumlah kelompok targetnya terlebih dahulu. Dapat digunakan beberapa metode seperti metode Elbow dan Silhouette. Pada analisis ini digunakan metode Elbow.



Gambar 2. Banyaknya kluster dengan Metode Elbow







Dengan menggunakan metode Elbow, dapat dilihat bahwa kurva sudah cukup landai untuk  $k=2$  dan sangat landai untuk  $k=3$ . Maka dari gambar tersebut dapat disimpulkan bahwa jumlah optimal kelompok yang dibentuk adalah 3 klaster.

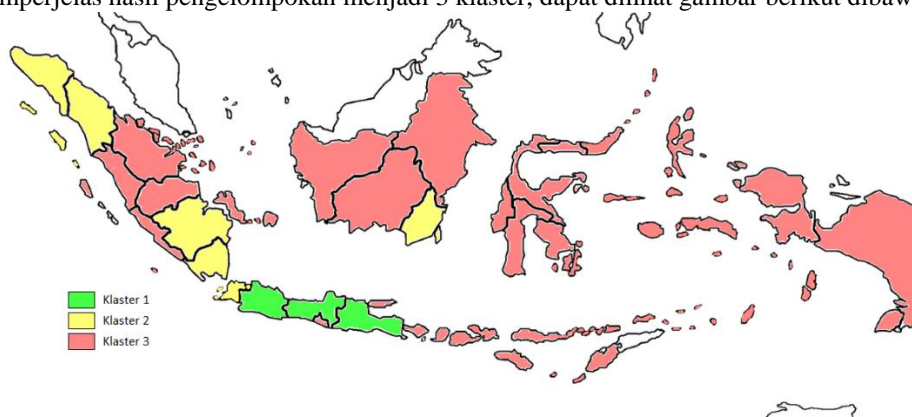
Didapatkan pengelompokan provinsi di Indonesia berdasarkan infrastrukturnya adalah sebagai berikut:

Kelompok 1 : Jawa Barat, Jawa Tengah, dan Jawa Timur

Kelompok 2 : Aceh, Sumatera Utara, Sumatera Selatan, Lampung, Banten, DKI Jakarta, dan Sulawesi Selatan

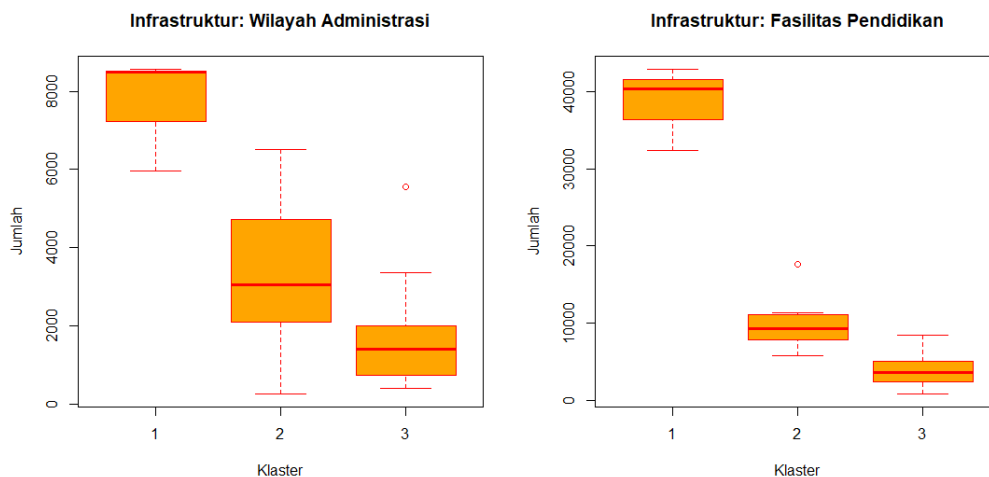
Kelompok 3 : Sumatera Barat, Riau, Jambi, Bengkulu, Kep. Bangka Belitung, Kep. Riau, DI Yogyakarta, NTT, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Bali, NTB, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, dan Papua

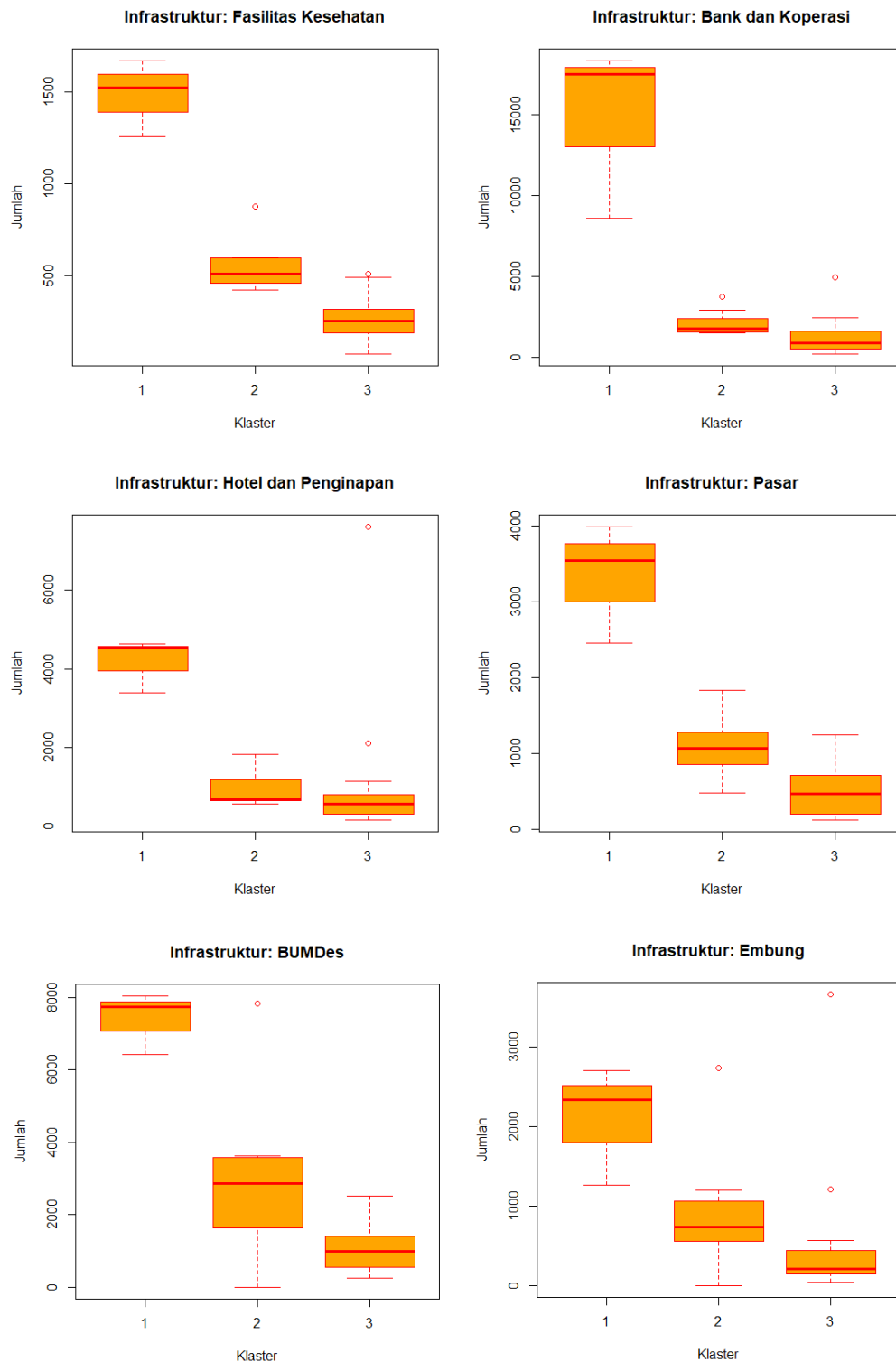
Untuk memperjelas hasil pengelompokan menjadi 3 klaster, dapat dilihat gambar berikut dibawah ini:



Gambar 3. Hasil Pengelompokan Provinsi di Indonesia

Selanjutnya karakteristik dari masing-masing klaster yang terbentuk dapat dilihat melalui boxplot, yaitu sebagai berikut:





**Gambar 4.** Boxplot untuk masing-masing variabel

Dapat dilihat dari boxplot bahwa untuk semua fasilitas infrastruktur, kluster 1 lebih baik daripada kluster 2 dan 3, dan kluster 2 lebih baik daripada kluster 3, dan kluster 3 merupakan kluster dengan infrastruktur paling sedikit.





#### IV. KESIMPULAN

1. Berdasarkan analisis Multidimensional Scaling dengan nilai kebaikan yang didapatkan, yaitu stress sebesar 0.0000212% dan goodness of fit sebesar 96.37% maka peta konfigurasi yang dihasilkan sangat baik.

2. Dengan menggunakan K-Means Clustering, dapat dibentuk 3 kelompok yang memiliki karakteristik masing-masing, dengan provinsi-provinsi yang termasuk kedalam masing-masing kelompok adalah sebagai berikut:

- Kelompok 1 : Jawa Barat, Jawa Tengah, dan Jawa Timur
- Kelompok 2 : Aceh, Sumatera Utara, Sumatera Selatan, Lampung, Banten, DKI Jakarta, dan Sulawesi Selatan
- Kelompok 3 : Sumatera Barat, Riau, Jambi, Bengkulu, Kep. Bangka Belitung, Kep. Riau, DI Yogyakarta, NTT, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Bali, NTB, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, dan Papua.

3. Pada karakteristik yang ditunjukkan oleh Boxplot untuk masing-masing variabel, dapat dilihat bahwa provinsi yang berada pada klaster 1 lebih baik daripada provinsi yang berada pada klaster lainnya, provinsi yang berada pada klaster 2 lebih baik daripada yang berada di klaster 3, dan provinsi yang berada pada klaster 3 merupakan provinsi-provinsi dengan jumlah fasilitas yang paling sedikit. Maka untuk menanggapi hal ini, pemerintah sebagai pengambil kebijakan sebaiknya melakukan pemerataan infrastruktur dengan mengutamakan daerah-daerah yang berada pada klaster 3, dan selanjutnya klaster 2.





#### DAFTAR PUSTAKA

- Badan Pusat Statistik. (2021). Publikasi Statistik: Statistik Infrastruktur Indonesia 2020. Jakarta: Badan Pusat Statistik
- Rencher, A. C. (2002). *Methods of Multivariate Analysis*. Second Ed. Canada: A John Wiley & Sons, Inc.
- Waluyo, J (2008). Hubungan Antara Tingkat Kesenjangan Pendapatan Dengan Pertumbuhan Ekonomi: Suatu Study Lintas Negara. Dalam *Jurnal Ekonomi Pembangunan Kajian Ekonomi Negara Berkembang*
- Rachelia, R. S. S., Rahayu W., Syabani, A. N. S., Senida., V. D., Puteri Z. I., Ginanjar, I.. (2021). PEMETAAN PROVINSI DI INDONESIA BERDASARKAN INDIKATOR KUALITAS LAYANAN PENDIDIKAN. Seminar Nasional Statistik Online (SNSO 2021). ISSN ONLINE. 2599-2546. ISSN Cetak. 2087-2590.
- Nagari, S. S & Inayati, L. (2019). Implementation of Clustering Using K-Means Method to Determine Nutritional Status, *Jurnal Biometrika dan Kependudukan*. DOI: 10.20473/jbk.v9i1.2020.62-68
- Siwi, C. P & Nurfirdaus, Y. (2019). K-Means Cluster Analysis of Sub-District in Sidoarjo Based on Long-Term Contraceptive Method. *Jurnal Biometrika dan Kependudukan*. DOI:10.20473/jbk.v9i1.2020.161-170
- Awan, A. G., and Anum, V. (2014). IMPACT OF INFRASTRUCTURE DEVELOPMENT ON ECONOMIC GROWTH: A CASE STUDY OF PAKISTAN. *International Journal of Development and Economic Development*. Vol. 2, No. 5. Pp 1-15
- Demurger, S. (2001). Infrastructure Development and Economic Growth: An Explanation for Regional Disparities in China?. *Journal of Comparative Economics* 20, 95-117
- Koner, J., Purandare, A., dan Dhume, A. (2012). An Empirical Study on Impact of Infrastructural Development on Social and Economic Growth in Indian States. *European Journal of Business and Management*. Vol 4. No. 17, 2002

