

Analisis Tingkat Kemiskinan di Indonesia Menggunakan Pendekatan PCA dan K-Medoids Clustering Tahun 2022-2023

Nurul Aulia Nowi¹, Edy Widodo^{2*}, Sani Mutia³, Nabila Ratna Dewati⁴, Lidya Palipi Saraswasti⁵

^{1,2,3,4,5}Program Studi Statistika Fakultas MIPA Universitas Islam Indonesia, Jl. Kaliurang Km. 14.5, Sleman, Yogyakarta, 55584, Indonesia

*Corresponding author: edywidodo@uii.ac.id



P-ISSN: 2986-4178
E-ISSN: 2988-4004

Riwayat Artikel

Dikirim: 21 Januari 2025
Direvisi: 29 Mei 2025
Diterima: 29 Mei 2025

ABSTRAK

Kemiskinan menjadi tantangan utama pembangunan di Indonesia, terutama akibat ketimpangan sosial-ekonomi antar wilayah. Penelitian ini bertujuan untuk mengidentifikasi pola sosial-ekonomi dan infrastruktur dasar di 34 provinsi Indonesia dengan pendekatan *Principal Component Analysis* (PCA) dan *K-Medoids clustering*. Data yang digunakan meliputi persentase penduduk miskin, Indeks Pembangunan Manusia (IPM), rata-rata lama sekolah, pengeluaran per kapita, dan akses sanitasi layak untuk tahun 2022–2023. Hasil PCA menunjukkan bahwa dua komponen utama dapat menjelaskan 84,7% variabilitas data pada 2022 dan 90,4% pada 2023. Metode *K-Medoids* menghasilkan tiga kelompok provinsi dengan karakteristik yang berbeda. Kelompok provinsi termiskin, seperti Papua dan Nusa Tenggara Timur, ditandai oleh rendahnya IPM, akses sanitasi yang terbatas, dan pengeluaran per kapita rendah, sehingga memerlukan perhatian lebih dalam pembangunan infrastruktur dasar dan akses pendidikan. Studi ini memberikan wawasan untuk merumuskan kebijakan pembangunan yang lebih efektif guna mengurangi ketimpangan dan mendukung pencapaian visi Indonesia Emas 2045.

Kata Kunci: Kemiskinan, Ketimpangan Sosial-Ekonomi, *Principal Component Analysis*, *K-Medoids Clustering*, Indeks Pembangunan Manusia.

ABSTRACT

Poverty remains a major challenge in Indonesia's development, primarily due to socio-economic disparities between regions. This study aims to identify socio-economic patterns and basic infrastructure in 34 Indonesian provinces using Principal Component Analysis (PCA) and K-Medoids clustering. The data include the percentage of the poor population, Human Development Index (HDI), average years of schooling, per capita expenditure, and access to proper sanitation for 2022–2023. PCA results showed that two principal components explained 84.7% of data variability in 2022 and 90.4% in 2023. The K-Medoids method categorized provinces into three clusters with distinct characteristics. The poorest provinces, such as Papua and East Nusa Tenggara, were characterized by low HDI, limited sanitation access, and low per capita expenditure, highlighting the need for greater attention in developing basic infrastructure and educational access. This study provided insights to formulate more effective development policies to reduce disparities and support the achievement of Indonesia's 2045 Vision.

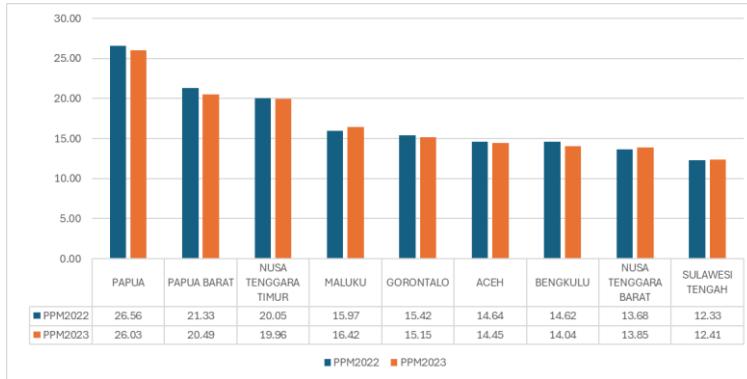
Keywords: Poverty, Socio-Economic Disparities, Principal Component Analysis, K-Medoids Clustering, Human Development Index.

1. Pendahuluan

Indonesia memiliki visi besar untuk menjadi negara maju yang berdaulat, adil, dan makmur pada tahun 2045, yang dikenal sebagai Indonesia Emas 2045. Visi ini sejalan dengan tujuan *Sustainable Development Goals* (SDGs) yang berfokus pada pengentasan kemiskinan, peningkatan kualitas pendidikan, dan pengurangan ketimpangan. *Sustainable Development Goals* (SDGs) merupakan suatu rencana aksi global yang disepakati oleh para pemimpin dunia, termasuk Indonesia, guna mengakhiri kemiskinan, mengurangi kesenjangan dan melindungi lingkungan. Salah satu tujuan utama SDGs yaitu menghapus kemiskinan dalam segala bentuknya di seluruh dunia, yang dimana sejalan dengan visi Indonesia Emas 2045 yang kedua dengan menargetkan tingkat kemiskinan yang sangat rendah, yakni antara 0.5% hingga 0.8%. Namun, realitas di Indonesia menunjukkan tantangan tersebut masih cukup besar. Pada tahun 2023, sekitar 9.36% penduduk Indonesia masih hidup di bawah garis kemiskinan, jauh dari target yang diharapkan [1]. Kondisi ini tidak hanya menggambarkan tingkat kesejahteraan masyarakat yang belum merata, tetapi juga mempertegas adanya kesenjangan sosial-ekonomi yang signifikan di berbagai wilayah di Indonesia.

Ketimpangan regional dalam distribusi kemiskinan di Indonesia menjadi salah satu isu utama yang memerlukan perhatian mendalam. Kemiskinan yang lebih terkonsentrasi di wilayah timur Indonesia dan sebagian daerah luar Jawa mencerminkan adanya ketimpangan pembangunan yang signifikan, khususnya Pulau Jawa yang menjadi pusat ekonomi nasional. Beberapa provinsi yang menjadi sorotan utama di Indonesia adalah Papua, Papua Barat, dan Nusa Tenggara Timur, yang secara konsisten mencatatkan angka kemiskinan tertinggi di Indonesia. Pada tahun 2023, PPM di Papua mencapai 26.03%, disusul Papua Barat dengan 20.49%, dan Nusa Tenggara Timur sebesar 19.96% [2]. Tingginya angka kemiskinan di provinsi-provinsi ini menunjukkan

bahwa tantangan pembangunan tidak hanya soal meningkatkan pendapatan masyarakat, tetapi juga memastikan akses terhadap infrastruktur dasar, pendidikan, kesehatan, dan lapangan kerja yang layak. Selain itu, daerah seperti Maluku, Gorontalo, dan Aceh juga termasuk dalam kategori wilayah dengan tingkat kemiskinan yang cukup tinggi. Pada tahun 2023, PPM di Maluku mencapai 16.42%, Gorontalo dengan 15.15%, dan Aceh sebesar 14.45%. Kondisi ini menandakan bahwa kemiskinan lebih terkonsentrasi di wilayah timur Indonesia dan sebagian daerah luar Jawa.



Gambar 1. Perbandingan Tingkat Kemiskinan di Indonesia Tahun 2022-2023 (%)

Tingginya angka kemiskinan di provinsi Papua, Papua Barat, dan Nusa Tenggara Timur mencerminkan ketimpangan pembangunan yang signifikan antara wilayah barat dan timur Indonesia. Hal ini menunjukkan bahwa tantangan pembangunan tidak hanya soal meningkatkan pendapatan masyarakat, tetapi juga memastikan akses terhadap pendidikan, kesehatan, lapangan kerja yang layak, dan infrastruktur dasar. Oleh karena itu, penelitian ini akan mengidentifikasi pola sosial-ekonomi dan infrastruktur dasar provinsi dengan menggunakan beberapa indikator utama, yaitu Persentase Penduduk Miskin (PPM) sebagai indikator kemiskinan, Indeks Pembangunan Manusia (IPM) untuk mengukur kualitas hidup, Rata-rata Lama Sekolah (RLS) sebagai ukuran tingkat pendidikan, pengeluaran per kapita (PK) untuk menunjukkan kemampuan ekonomi masyarakat, dan akses terhadap sanitasi layak sebagai indikator infrastruktur dasar.

Beberapa alat statistik yang dapat digunakan dalam penelitian ini adalah *K-Medoids clustering* dan *Principal Component Analysis* (PCA). Penelitian ini menerapkan *K-Medoids clustering* dan *Principal Component Analysis* (PCA) untuk menganalisis pola sosial-ekonomi provinsi dengan ketimpangan pembangunan signifikan. Algoritma *K-Medoids* adalah teknik pengelompokan data yang digunakan untuk mengelompokkan sejumlah objek menjadi satu atau lebih kelompok. Metode ini menggunakan satu objek dari sekumpulan data untuk mewakili kelompok tersebut [3]. Sedangkan *Principal Component Analysis* (PCA) adalah Teknik analisis multivariat yang bertujuan untuk mereduksi sejumlah variabel yang saling berkorelasi menjadi satu set variabel baru yang lebih kecil, disebut komponen utama, yang saling bebas dan dapat mewakili informasi yang terkandung dalam variabel asal [4].

Pada penelitian ini *K-Medoids clustering* untuk mengelompokkan provinsi berdasarkan karakteristik seperti kemiskinan, pendidikan, dan infrastruktur, dengan keunggulan lebih tahan terhadap *outlier* dibandingkan *K-Means*. Sedangkan PCA mereduksi variabel berkorelasi seperti IPM, RLS, pengeluaran per kapita, dan akses sanitasi menjadi komponen utama yang bebas dan mewakili informasi penting. Kombinasi kedua metode ini meningkatkan efisiensi analisis dan pemahaman terhadap ketimpangan pembangunan antar wilayah di Indonesia.

Berbagai penelitian telah dilakukan untuk menganalisis faktor-faktor yang mempengaruhi tingkat kemiskinan di Indonesia seperti IPM, RLS, TPT, dan pengeluaran per kapita menunjukkan hasil berpengaruh terhadap kemiskinan. Penelitian [5] menemukan bahwa infrastruktur seperti sanitasi, listrik, dan jalan berpengaruh negatif signifikan terhadap kemiskinan. Kemudian penelitian [6] dan [7] yang menyoroti peran IPM dan infrastruktur yang signifikan terhadap kemiskinan. Sementara [8] serta [9] menerapkan *clustering* dan PCA untuk mengelompokkan daerah miskin. Penelitian [10], [11], serta [12] juga menggunakan *K-Means* untuk mengelompokkan daerah miskin, sedangkan [13] menunjukkan bahwa infrastruktur jalan dan sanitasi secara parsial berdampak negatif dan signifikan terhadap kemiskinan di Sulawesi Selatan. Serta penelitian [14] mengklasifikasikan daerah tertinggal di Indonesia bagian timur berdasarkan persentase penduduk miskin, angka harapan hidup, puskesmas, dan alokasi dana khusus menggunakan metode *K-Means*, *K-Medoids*, dan *Fuzzy C-Means*.

Meskipun berbagai penelitian sebelumnya telah menggunakan beragam indicator-indikator tersebut, namun kajian yang secara khusus mengaitkan tingkat kemiskinan dengan infrastruktur masih terbatas. Oleh karena itu, penelitian ini berupaya memberikan perspektif yang lebih mendalam melalui analisis keterkaitan antarindikator ketimpangan pembangunan wilayah dengan memasukkan faktor sanitasi layak. Penelitian ini ditujukan untuk mengisi kekosongan analisis dinamika kemiskinan dan ketimpangan antarprovinsi di Indonesia pada periode 2022–2023, dalam konteks pasca-pandemi COVID-19, ketika upaya pemulihan ekonomi nasional menjadi prioritas dan isu ketimpangan pembangunan antar wilayah semakin menonjol. Fokus utama penelitian adalah mengidentifikasi pola sosial-ekonomi provinsi menggunakan indikator IPM, RLS, pengeluaran per kapita, dan akses sanitasi layak melalui metode PCA dan *K-Medoids*. Dengan pendekatan ini, penelitian diharapkan dapat memberikan kontribusi baru dalam memahami dinamika sosial-ekonomi serta pembangunan infrastruktur di provinsi dengan tingkat kemiskinan tinggi, sekaligus mendukung perumusan kebijakan yang lebih efektif guna mengurangi kemiskinan dan ketimpangan, serta menunjang visi Indonesia Emas 2045.

2. Metodologi Penelitian

2.1 Data

Populasi dalam penelitian ini meliputi 34 provinsi di Indonesia, dengan variabel penelitian disajikan dalam Tabel 1. Data yang digunakan adalah data tahun 2022 - 2023. Penelitian ini menggunakan data sekunder yang bersumber dari website Badan Pusat Statistik. Peneliti menggunakan beberapa *software* antara lain Ms. Excel dan R Studio dalam pengolahan data.

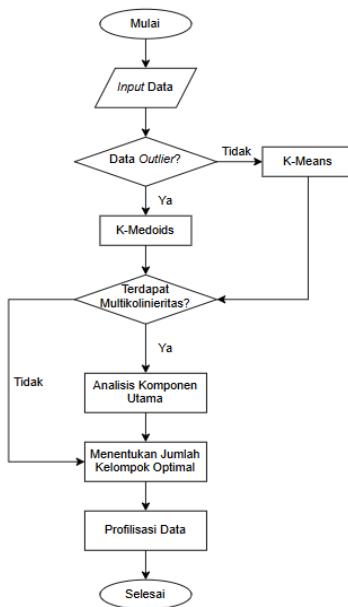
Tabel 1 Variabel Penelitian

Variabel	Deskripsi	Informasi
PPM	Persentase Penduduk Miskin sebagai ukuran tingkat kemiskinan di setiap provinsi.	Angka (%)
IPM	Indeks yang mengukur pembangunan manusia melalui dimensi kesehatan, pendidikan, dan standar hidup.	Angka (Indeks)
RLS	Rata-rata Lama Sekolah (RLS) untuk menggambarkan tingkat pencapaian pendidikan.	Angka (Tahun)
PK	Pengeluaran per Kapita Disesuaikan sebagai indikator tingkat kesejahteraan ekonomi.	Angka (Ribu Rupiah/Orang/Tahun)
Sanitasi Layak	Persentase penduduk yang memiliki akses ke sanitasi layak, yang mencakup fasilitas pembuangan limbah manusia yang aman dan memenuhi standar kesehatan.	Angka (%)

2.2 Tahapan Penelitian

Penelitian ini menggunakan metode analisis deskriptif, *Principal Component Analysis* (PCA), dan *K-Medoids clustering*. PCA diterapkan untuk mereduksi dimensi data dengan tujuan menyederhanakan analisis tanpa kehilangan informasi penting. Metode PCA juga dapat digunakan untuk menangani masalah multikolinearitas dengan mereduksi jumlah variabelnya [15]. Dalam analisis ini, komponen utama yang paling berkontribusi terhadap variasi data diidentifikasi untuk menentukan faktor-faktor kunci yang mempengaruhi tingkat kemiskinan. Sedangkan *K-Medoids clustering* digunakan untuk mengelompokkan provinsi berdasarkan karakteristik seperti kemiskinan, pendidikan, dan infrastruktur, dengan keunggulan lebih tahan terhadap *outlier* dibandingkan *K-Means*.

Selain itu, evaluasi kualitas hasil pengelompokan *K-Medoids* juga penting untuk memastikan validitas dan keandalan analisis. Validasi dapat dilakukan menggunakan metode *elbow*, yang membantu menentukan jumlah kelompok optimal berdasarkan perubahan dalam jumlah total deviasi kuadrat (*total within-cluster sum of squares*). Metode ini mengevaluasi penurunan deviasi kuadrat saat jumlah kelompok bertambah dan mengidentifikasi titik di mana penurunan mulai melambat (titik siku/*elbow*). Langkah-langkah penelitian ini dapat dilihat dalam diagram alir pada Gambar 2.



Gambar 2. Diagram Alir Analisis Penelitian

Tahapan analisis dalam penelitian ini adalah sebagai berikut:

- Memasukkan data penelitian yang akan digunakan.
- Melakukan pengecekan terhadap keberadaan *outlier* pada data. Jika data tidak mengandung *outlier* maka menggunakan metode *K-Means*, sedangkan apabila terdapat *outlier* maka menggunakan metode *K-Medoids* karena metode ini lebih *robust*, stabil, dan tidak terpengaruh oleh nilai ekstrim. Selain itu, *K-Medoids* lebih unggul dibandingkan metode *clustering* lain karena mampu menghasilkan *cluster* yang konsisten meskipun data mengandung *noise*, tidak sensitif terhadap distribusi data yang tidak simetris, serta lebih representatif untuk ukuran data yang relatif kecil hingga sedang.

- c. Melakukan pemeriksaan multikolinearitas antar variabel untuk memastikan tidak terjadi korelasi tinggi yang dapat mengganggu analisis.
- d. Jika ditemukan multikolinearitas, maka dilakukan Analisis Komponen Utama (PCA) untuk mereduksi dimensi dan mengatasi korelasi variabel, sedangkan jika tidak ditemukan multikolinearitas, proses langsung dilanjutkan ke tahap berikutnya.
- e. Menentukan jumlah kelompok optimal menggunakan metode validasi *cluster* agar hasil pengelompokan lebih akurat.
- f. Melakukan profilisasi data untuk menggambarkan karakteristik utama dari tiap kelompok yang terbentuk.
- g. Menyusun kesimpulan berdasarkan hasil pengelompokan data yang telah dilakukan.

2.3. Outlier

Outlier merujuk pada elemen yang terpisah atau berbeda dibandingkan dengan elemen lainnya dalam suatu kelompok data. Dalam konteks pengumpulan data, kelompok yang dimaksud adalah sekumpulan data yang akan diproses lebih lanjut. *Outlier* bisa muncul sebagai akibat dari kesalahan dalam proses pengumpulan data atau bisa juga menjadi indikasi adanya variasi dalam kelompok data tersebut [16].

2.4 Principal Component Analysis (PCA)

Pemilihan jumlah komponen utama dilakukan berdasarkan nilai *cumulative explained variance ratio*. Dalam penelitian ini, komponen utama yang dipertahankan adalah yang secara kumulatif menjelaskan lebih dari 80% variasi dalam data. Sebagai langkah pendukung, *scree plot* digunakan untuk memvisualisasikan distribusi nilai *eigen*.

Setiap komponen utama dianalisis untuk mengidentifikasi fitur asli mana yang memiliki kontribusi terbesar terhadap variasi data. Analisis ini dilakukan dengan mengamati bobot (*loading*) dari setiap fitur pada masing-masing komponen. Hasilnya menunjukkan bahwa fitur tertentu memberikan kontribusi terbesar pada komponen utama pertama, sedangkan fitur lainnya mendominasi komponen utama kedua. Reduksi dimensi ini bertujuan untuk mengurangi kompleksitas data dan mempercepat proses pengelompokan.

Asumsi yang harus dipenuhi dalam PCA adalah sebagai berikut.

a) Kecukupan data

Kelayakan analisis faktor diuji menggunakan *Kaiser-Meyer-Olkin* (KMO), yang mengukur kecukupan sampel. Nilai KMO yang direkomendasikan adalah ≥ 0.5 ; jika lebih rendah, PCA tidak disarankan. Rumus yang digunakan sebagai berikut.

$$KMO = \frac{\sum_{j=1}^p \sum_{l=1, l \neq j}^p r_{x_j x_l}^2}{\sum_{j=1}^p \sum_{l=1, l \neq j}^p r_{x_j x_l}^2 + \sum_{j=1}^p \sum_{l=1, k \neq j}^p \rho_{x_j x_k, x_m}^2} \quad (1)$$

dengan, p : banyaknya variabel, n : banyaknya objek, x_i : objek pada pengamatan ke- i , $r_{x_j x_l}$: korelasi antara variabel x_j dan x_l , \bar{x}_j : rata-rata variabel x_j , \bar{x}_l : rata-rata variabel x_l , dan $\rho_{x_j x_l, x_m}$: korelasi parsial antara variabel x_j dan x_l dengan menjaga agar x_m konstan.

b) Kelayakan variabel

Measure of Sampling Adequacy (MSA) mengevaluasi kecukupan setiap variabel. Variabel dengan nilai MSA < 0.5 sebaiknya dikeluarkan dari analisis. Rumus yang digunakan sebagai berikut.

$$MSA = \frac{\sum_{i=1}^p r_{ij}^2}{\sum_{i=1}^p r_{ij}^2 + \sum_{i=1}^p a_{ij}^2} \quad (2)$$

dengan, p : banyaknya variabel, r_{ij} : koefisien korelasi antara variabel i dan j , a_{ij} : koefisien korelasi parsial antara i dan j .

c) Multikolinearitas

Nilai VIF digunakan untuk mengetahui adanya multikolinearitas. Jika nilai VIF kurang dari 10 maka tidak menunjukkan adanya multikolinearitas. Rumus VIF yang digunakan sebagai berikut.

$$VIF = \frac{1}{R_i^{-2}} \quad (3)$$

dengan, VIF: *variance inflation factor*, R_i^{-2} : koefisien determinasi dari regresi antara variabel independen yang diuji terhadap semua variabel independen lainnya.

2.5 K-Medoids

Algoritma *K-Medoids* digunakan untuk menemukan *medoids* sebagai titik pusat dalam suatu kelompok. *K-Medoids* lebih kuat dibandingkan *K-Means*, terutama dalam mengurangi pengaruh data yang *noise* dan *outliers*. Hal ini disebabkan oleh penggunaan metrik jarak yang meminimalkan ketidaksamaan objek data, berbeda dengan *K-Means* yang menggunakan kuadrat jarak *euclidean*. Algoritma ini memilih objek sebenarnya untuk mewakili *cluster* daripada nilai rata-rata, yang membuatnya lebih akurat dalam penanganan data berbasis objek [17].

- Menentukan banyaknya kelompok yang akan dibentuk menggunakan *elbow*.
- Mengalokasikan setiap data (objek) ke kelompok terdekat menggunakan persamaan ukuran jarak *euclidean distance* metode dengan persamaan (4).

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (4)$$

dengan, $d(x, y)$: jarak *euclidean* antara dua titik x dan y , x_i : komponen ke- i dari vektor x , y_i : komponen ke- i dari vektor y , dan n : jumlah observasi.

- Pilih secara acak objek pada masing-masing kelompok sebagai kandidat *medoids* baru.
- Hitung jarak setiap objek yang berada pada masing-masing kelompok dengan kandidat *medoids* baru.
- Hitung total simpangan (S) dengan menghitung nilai total *distance* baru–total *distance* lama. Jika $S < 0$, maka tukar objek dengan data kelompok untuk membentuk sekumpulan k objek baru sebagai *medoids*.
- Ulangi langkah c) sampai e) hingga tidak terjadi perubahan *medoids*, sehingga didapatkan kelompok beserta anggota kelompok masing-masing [18].

3. Hasil dan Pembahasan

3.1. Analisis Deskriptif

Analisis deskriptif dilakukan untuk memberikan gambaran mengenai setiap variabel dalam penelitian ini, dengan menampilkan nilai minimum (*min*), maksimum

(*max*), rata-rata (*mean*), standar deviasi (*stdv*), dan jumlah data (*n*) untuk tahun 2022 dan tahun 2023. Tujuannya adalah untuk memahami statistik dasar yang dihasilkan, sehingga pola atau perbedaan antara kedua tahun dapat diidentifikasi secara jelas. Analisis ini menjadi langkah awal dalam menggambarkan karakteristik data secara ringkas dan informatif.

Tabel 2 Statistik Deskriptif PPM, IPM, RLS, PK, dan Sanitasi Layak Tahun 2022

	Percentase Penduduk Miskin (%)	Indeks Pembangunan Manusia (%)	Rata-Rata Lama Sekolah (Tahun)	Pengeluaran per Kapita (Ribu Rupiah/Orang/Tahun)	Sanitasi Layak (%)
<i>Min</i>	4.45	61.39	7.02	7.14	40.34
<i>Max</i>	26.56	81.65	11.31	18.92	96.21
<i>Mean</i>	10.24	71.97	8.83	11.08	81.00
<i>Stdv</i>	5.25	3.90	3.85	2246.72	9.784
<i>n</i>	34	34	34	34	34

Tabel 2 di atas menunjukkan ketimpangan antar daerah di Indonesia tahun 2022 dalam hal kesejahteraan, pendidikan, dan akses layanan dasar. Tingkat kemiskinan, IPM, dan lama sekolah bervariasi cukup besar, mencerminkan belum meratanya pembangunan. Pengeluaran per kapita juga menunjukkan kesenjangan daya beli antar wilayah. Meskipun akses sanitasi layak secara umum tinggi, masih ada daerah dengan akses yang rendah, menandakan infrastruktur dasar belum sepenuhnya merata.

Tabel 3 Statistik Deskriptif PPM, IPM, RLS, PK, dan Sanitasi Layak Tahun 2023

	Percentase Penduduk Miskin (%)	Indeks Pembangunan Manusia (%)	Rata-Rata Lama Sekolah (Tahun)	Pengeluaran per Kapita (Ribu Rupiah/Orang/Tahun)	Sanitasi Layak (%)
<i>Min</i>	4.25	62.25	7.15	7.56	43.00
<i>Max</i>	26.03	82.46	11.45	19.37	96.42
<i>Mean</i>	10.09	72.62	8.93	11.47	82.57
<i>Stdv</i>	5.18	3.85	0.91	2265.36	9.46
<i>N</i>	34	34	34	34	34

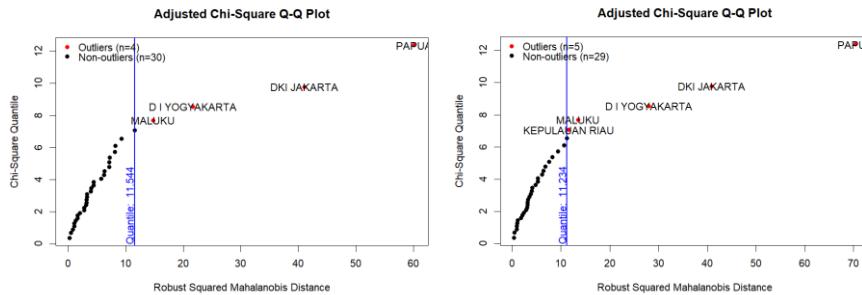
Tabel tahun 2023 menunjukkan perbaikan pada hampir seluruh indikator dibandingkan tahun 2022. Rata-rata tingkat kemiskinan menurun, IPM dan rata-rata lama sekolah meningkat, serta pengeluaran per kapita dan akses sanitasi layak juga mengalami kenaikan. Selain itu, variasi antar daerah dalam hal pendidikan dan sanitasi sedikit mengecil, menandakan adanya perbaikan dalam pemerataan. Meskipun demikian, ketimpangan antar wilayah masih terlihat, terutama pada indikator kemiskinan dan pengeluaran, sehingga pemerataan pembangunan tetap menjadi tantangan yang perlu diatasi.

Meskipun terjadi peningkatan dalam beberapa indikator, persentase penduduk miskin tahun 2023 yang mencapai 26.03% menunjukkan bahwa kemiskinan di daerah tertentu masih menjadi tantangan signifikan. Selain itu, akses terhadap sanitasi layak secara keseluruhan mengalami perbaikan, rentang nilai yang dimulai dari 43% dan standar deviasi yang tinggi pada kedua tahun tersebut mengindikasikan bahwa upaya peningkatan sanitasi layak belum merata di seluruh provinsi. Provinsi-provinsi tertentu mungkin tertinggal jauh dalam akses sanitasi layak dibandingkan yang lain. Hal ini menunjukkan bahwa kesenjangan sosial dan akses terhadap kebutuhan dasar masih memerlukan perhatian lebih lanjut.

3.2. Pengecekan Outlier

Berdasarkan hasil statistik deskriptif tersebut, terlihat adanya variasi yang cukup besar pada beberapa indikator antar daerah, khususnya pada tingkat kemiskinan dan pengeluaran per kapita. Variasi yang tinggi ini dapat mengindikasikan kemungkinan

adanya data ekstrim atau pencilan (*outlier*) yang dapat mempengaruhi hasil analisis secara keseluruhan. Oleh karena itu, diperlukan pengecekan lebih lanjut terhadap keberadaan *outlier* sebelum melanjutkan ke tahap analisis berikutnya dengan menggunakan metode kuantil berdasarkan jarak *mahalanobis*.



Gambar 3. Hasil *Outlier* Tahun 2022 dan Tahun 2023

Pada tahun 2022 terdapat 4 provinsi *outlier*, yaitu provinsi Maluku, DI Yogyakarta, DKI Jakarta, dan Papua. Pada tahun 2023 terdapat 5 provinsi yang *outlier*, yaitu provinsi Kepulauan Riau, Maluku, DI Yogyakarta, DKI Jakarta, dan Papua. Informasi mengenai provinsi-provinsi *outlier* ini menjadi penting dalam proses pengelompokan menggunakan *K-Medoids*. Keberadaan *outlier* yang diidentifikasi membantu *K-Medoids* dalam membentuk kelompok yang lebih akurat dan tahan terhadap pengaruh nilai ekstrim. Hal ini memastikan bahwa pola yang dihasilkan mencerminkan distribusi data yang lebih realistik dan mendukung interpretasi yang lebih baik mengenai ketimpangan pembangunan antar wilayah.

3.3 Uji Asumsi

3.3.1 Uji Kecukupan Data dan Uji Kelayakan Model

Pengujian kecukupan data dengan menggunakan metode *Kaiser-Meyer-Olkin* (*KMO*) dan uji kelayakan setiap variabel dianalisis faktor dengan menggunakan kriteria MSA sebagai berikut.

Tabel 4 Hasil Uji Kecukupan Data

Tahun	Overall MSA
2022	0.68
2023	0.66

Berdasarkan Tabel 4, dapat diketahui bahwa nilai *Kaiser-Meyer-Olkin factor adequacy* atau *overall MSA* sebesar 0.68 dan 0.66. Maka analisis faktor dalam kasus ini dapat dilanjutkan karena nilai KMO diatas 0.5.

3.3.2 Uji Multikolinearitas

Pengujian korelasi antar variabel dilakukan terhadap variabel-variabel independen untuk menghilangkan variabel yang berkorelasi tinggi. Uji asumsi ini menggunakan metode *Variance Inflation Factor* (VIF) untuk melihat ada tidaknya masalah multikolinearitas dalam data dengan menggunakan nilai batas VIF < 10.

Tabel 5 Uji Asumsi Multikolinieritas Tahun 2022

Variabel	VIF
IPM	11.07
RLS	3.04
Sanitasi Layak	2.58
PK	4.84

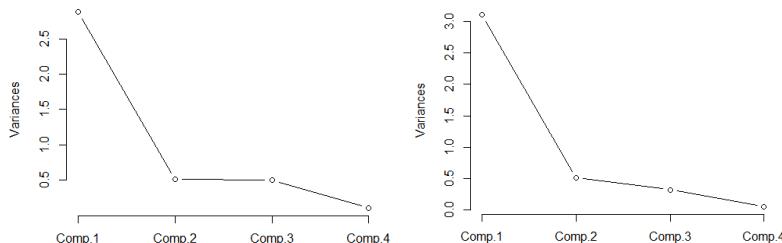
Tabel 6 Uji Asumsi Multikolinieritas Tahun 2023

Variabel	VIF
IPM	11.53
RLS	3.18
Sanitasi Layak	2.44
PK	5.10

Berdasarkan Tabel 5 dan Tabel 6, hasil analisis VIF menunjukkan bahwa salah satu variabel memiliki nilai $VIF > 10$, yaitu 11.07 dan 11.53, yang mengindikasikan adanya multikolinearitas serius, sementara variabel lainnya memiliki nilai VIF yang masih berada di bawah ambang batas kritis. Kondisi ini menunjukkan bahwa hubungan antar variabel dalam *dataset* saling berkorelasi, sehingga PCA perlu dilakukan untuk mereduksi dimensi data.

3.3. Principal Component Analysis

Dalam penelitian ini, ditemukan bahwa data yang digunakan memiliki masalah multikolinieritas. Oleh karena itu, data akan diolah terlebih dahulu menggunakan metode PCA yang didasarkan pada interpretasi *scree plot* (*plot* nilai *eigen*), yang menunjukkan perubahan signifikan dalam nilai *eigen*.



Gambar 4. Scree Plot Nilai Eigen Tahun 2022 dan Tahun 2023

Berdasarkan Gambar 4, *scree plot* menunjukkan pola serupa untuk tahun 2022 dan 2023. Pada kedua tahun, komponen 1 memiliki nilai *eigen* tertinggi, menjelaskan sebagian besar variansi data, diikuti oleh penurunan signifikan pada komponen 2 yang masih memberikan kontribusi yang cukup besar. Untuk komponen 3 dan 4, nilai *eigen* sangat kecil, menunjukkan kontribusi variansi yang tidak signifikan. Dengan demikian, komponen 1 dan komponen 2 pada kedua tahun dilakukan analisis lebih lanjut karena memiliki nilai *eigen* lebih dari 1.

Tabel 7 Vektor *Eigen* dan Komponen Utama Data Tahun 2022

Indeks Pembangunan Manusia	Rata-Rata Lama Sekolah	Sanitasi	Pengeluaran per Kapita
Comp. 1	0.55	0.45	0.50
Comp. 2		0.77	-0.19
Comp. 3	0.12	-0.30	0.71
Comp. 4	0.82	-0.30	-0.44

Tabel 8 Vektor *Eigen* dan Komponen Utama Data Tahun 2023

Indeks Pembangunan Manusia	Rata-Rata Lama Sekolah	Sanitasi	Pengeluaran per Kapita
Comp. 1	0.55	0.46	0.47
Comp. 2		0.74	-0.65
Comp. 3	0.13	-0.37	-0.58
Comp. 4	0.82	-0.31	0.18
			-0.45

Hasil PCA untuk data tahun 2022 menunjukkan bahwa komponen pertama (*Comp. 1*) didominasi oleh Indeks Pembangunan Manusia dan Pengeluaran per Kapita, yang mencerminkan dimensi pembangunan manusia dan kesejahteraan ekonomi. Komponen kedua (*Comp. 2*) didominasi oleh Rata-rata Lama Sekolah, sementara Sanitasi Layak memberikan kontribusi negatif yang signifikan. Komponen ketiga (*Comp. 3*) kembali didominasi oleh Pengeluaran per Kapita dengan Sanitasi Layak memberikan kontribusi negatif yang signifikan. Dan komponen keempat (*Comp. 4*) hanya didominasi Indeks Pembangunan Manusia, sedangkan komponen lainnya tidak memberikan kontribusi signifikan. Ini menunjukkan bahwa pada keempat komponen utama tersebut, komponen pertama (*Comp. 1*) dan komponen kedua (*Comp. 2*) adalah komponen yang paling berkontribusi terhadap variasi data dengan Indeks Pembangunan Manusia dan Rata-Rata Lama Sekolah memiliki pengaruh yang lebih besar.

Tabel 9 Ringkasan Analisis Komponen Utama Data Tahun 2022

	Std. Deviasi	% Variansi	Kumulatif %
Comp. 1	1.696	0.719	0.719
Comp. 2	0.715	0.128	0.847
Comp. 3	0.705	0.124	0.972
Comp. 4	0.334	0.027	1.000

Tabel 10 Ringkasan Analisis Komponen Utama Data Tahun 2023

	Std. Deviasi	% Variansi	Kumulatif %
Comp. 1	1.761	0.775	0.775
Comp. 2	0.717	0.128	0.904
Comp. 3	0.569	0.811	0.985
Comp. 4	0.243	0.014	1.000

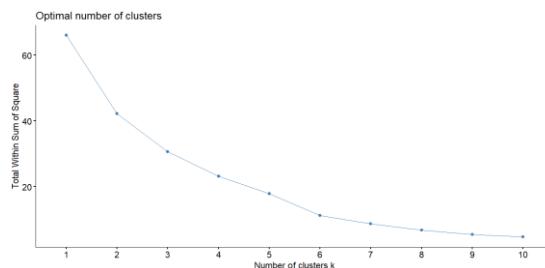
Dari hasil Tabel 9 dan Tabel 10, komponen utama pertama pada kedua tahun menjelaskan lebih dari 70% dari total varians, menjadikannya representasi utama dari variabel yang dianalisis. Namun analisis PCA ini sendiri memerlukan setidaknya dua dimensi komponen utama. Pada tahun 2022 hasil kumulatif dengan dua komponen utama dapat menjelaskan 84.7% dari total variabilitas data. Dengan komponen utama 1 yang menjelaskan 71.9% variabilitas, menangkap pola utama dalam data, sementara komponen utama 2 menambahkan 12.8% informasi tambahan yang relevan. Sedangkan pada data tahun 2023, hasil kumulatif dengan dua komponen utama dapat menjelaskan 90.4% dari total variabilitas data. Komponen utama 1 yang menjelaskan 77.5% variabilitas, menangkap pola utama dalam data, sementara komponen utama 2 menambahkan 12.8% informasi tambahan yang relevan.

Analisis PCA lebih lanjut menunjukkan bahwa komponen utama pertama didominasi oleh variabel IPM, RLS, Pengeluaran per Kapita, dan Sanitasi Layak dengan arah positif, serta Persentase Penduduk Miskin (PPM) dengan arah negatif, sehingga dapat ditafsirkan sebagai dimensi pembangunan dan kesejahteraan masyarakat. Sementara itu, komponen utama kedua lebih banyak menangkap variasi tambahan terutama dari aspek pendidikan, melalui kontribusi RLS. Dengan demikian, pola yang terbentuk menunjukkan bahwa perbedaan antar provinsi terutama dipengaruhi oleh

tingkat kesejahteraan umum, sementara aspek pendidikan berperan sebagai pembeda tambahan yang memperkaya interpretasi data.

3.4. Analisis K-Medoids

Pada penelitian ini, peneliti menggunakan metode *elbow* untuk menentukan k optimal.



Gambar 5. K Optimal Metode Elbow Tahun 2022

Berdasarkan Gambar 5, diperoleh bahwa k optimal adalah 3. Pada titik ini, penurunan nilai total *within sum of squares* mulai melambat. Maka, akan dilakukan pembentukan kelompok provinsi di Indonesia ke dalam 3 kelompok dengan anggota-anggota pengelompokan seperti pada Tabel 11.

Tabel 11 Anggota Pengelompokan Tahun 2022

Kelompok	Anggota	Jumlah
1	Papua, Papua Barat, Nusa Tenggara Timur, Maluku, Gorontalo, Aceh, Bengkulu, Nusa Tenggara Barat, Sulawesi Tengah, Lampung, DI Yogyakarta, Jawa Tengah, Jawa Timur, Sumatera Utara, Sulawesi Utara, Riau, Kalimantan Barat, Kalimantan Timur, Kepulauan Riau, Maluku Utara, Banten, Sumatera Barat, DKI Jakarta, Bali	24
2	Sumatera Selatan, Sumatera Barat, Sulawesi Tenggara, Sulawesi Selatan, Kalimantan Utara	5
3	Jawa Barat, Jambi, Kalimantan Tengah, Kalimantan Selatan, Kepulauan Bangka Belitung	5

Berikut adalah peta yang menunjukkan hasil pengelompokan provinsi di Indonesia berdasarkan hasil analisis tahun 2022.



Gambar 6. Peta Pengelompokan Tahun 2022

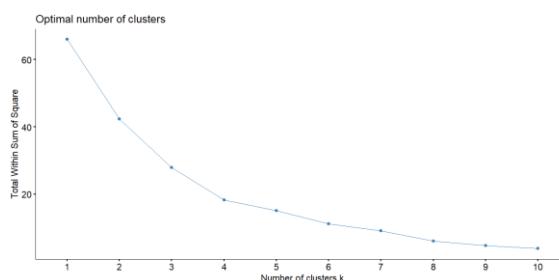
Berdasarkan Gambar 6, provinsi berwarna hijau (Kelompok 1) menunjukkan tingkat kemiskinan yang relatif rendah, provinsi berwarna oranye (Kelompok 2) memiliki tingkat kemiskinan sedang, sedangkan provinsi berwarna merah (Kelompok 3) mencerminkan tingkat kemiskinan yang relatif tinggi.

Hasil profilisasi data kemiskinan di Indonesia tahun 2022 digunakan untuk melihat karakteristik masing-masing kelompok, yang dapat diuraikan pada Tabel 12.

Tabel 12 Profilisasi Pengelompokan Data Tahun 2022

	Indeks Pembangunan Manusia (%)	Rata-Rata Lama Sekolah (Tahun)	Sanitasi Layak(%)	Pengeluaran per Kapita (Ribu Rupiah/Orang/Tahun)
Kelompok 1	72.13	8.93	80.53	11096.54
Kelompok 2	70.94	8.72	83.81	10191.00
Kelompok 3	72.19	8.54	80.41	11886.60

Berdasarkan hasil kelompok pada Tabel 11, terlihat perbedaan yang signifikan antara kelompok. Kelompok 3 merupakan kelompok dengan karakteristik yang paling baik, ditandai dengan pembangunan manusia yang sangat baik dan pengeluaran per kapita terbesar dibandingkan kelompok lainnya. Sementara, kelompok 2 adalah kelompok dengan karakteristik yang paling buruk, dengan pembangunan manusia dan pengeluaran per kapita paling rendah antar kelompoknya. Menunjukkan tantangan besar dalam pembangunan manusia, akses pendidikan, sanitasi, dan tingkat kesejahteraan. Kelompok 1 berada pada tingkat sedang, dengan kelompok 1 lebih unggul dalam pendidikan. Perbedaan antar kelompok ini mencerminkan variasi kondisi pembangunan di wilayah yang dikelompokkan.



Gambar 7. K Optimal Metode Elbow Tahun 2023

Berdasarkan Gambar 7, diperoleh bahwa k optimal yang digunakan untuk pengelompokan tahun 2023 sebanyak 3 kelompok. Berikut adalah masing-masing anggota 34 provinsi di Indonesia berdasarkan hasil kelompok sebanyak 3 kelompok.

Tabel 13 Anggota Pengelompokan Tahun 2023

Kelompok	Anggota	Jumlah
1	Acéh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kep. Bangka Belitung, Jawa Barat, Jawa Tengah, Jawa Timur, Banten, Nusa Tenggara Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Maluku, Maluku Utara	24
2	Kepulauan Riau, DKI Jakarta, DI Yogyakarta, Bali, Kalimantan Timur	5
3	Nusa Tenggara Timur, Kalimantan Barat, Sulawesi Barat, Papua, Barat Papua	5

Berikut adalah peta yang menunjukkan hasil pengelompokan provinsi di Indonesia berdasarkan hasil analisis tahun 2023.



Gambar 8. Peta Pengelompokan Tahun 2023

Hasil profilisasi data kemiskinan di Indonesia tahun 2023 digunakan untuk melihat karakteristik masing-masing kelompok, yang dapat diuraikan pada Tabel 14.

Tabel 14 Profilisasi Pengelompokan Data Tahun 2023

	Indeks Pembangunan Manusia (%)	Rata-Rata Lama Sekolah (Tahun)	Sanitasi Layak(%)	Pengeluaran per Kapita (Ribu Rupiah/Orang/Tahun)
Kelompok 1	72.53	8.90	82.66	11223.80
Kelompok 2	79.19	10.23	93.59	15375.80
Kelompok 3	66.51	7.75	71.12	8748.40

Berdasarkan hasil kelompok pada Tabel 14, terlihat perbedaan yang mencolok antara kelompok. Kelompok 2 merupakan kelompok dengan karakteristik paling baik, ditandai oleh pembangunan manusia tertinggi, pendidikan yang lebih maju, akses sanitasi yang sangat baik, dan pengeluaran per kapita terbesar dibandingkan kelompok lainnya yang menunjukkan aspek ekonomi yang baik. Sementara, Kelompok 3 adalah kelompok dengan karakteristik paling rendah, dengan semua indikator menunjukkan angka terendah antara kelompoknya. Ini menandakan bahwa wilayah pada kelompok 3 memiliki tantangan signifikan dalam pembangunan manusia, akses pendidikan yang rendah, sanitasi yang terbatas, dan pengeluaran per kapita paling kecil. Kelompok 1 berada pada tingkat sedang, dengan karakteristik yang cukup baik di semua aspek tetapi masih di bawah Kelompok 2.

Berdasarkan Tabel 11 dan Tabel 13, terlihat bahwa terdapat beberapa provinsi yang mengalami perubahan atau berpindah kelompok. Sebagai contoh yaitu Provinsi Kalimantan Barat yang mengalami perubahan posisi yang cukup mencolok, di mana pada tahun 2022 masuk dalam kelompok 1 yang memiliki karakteristik menengah dengan indikator pembangunan manusia, pendidikan, sanitasi, dan pengeluaran per kapita yang relatif baik, tetapi pada tahun 2023 bergeser ke kelompok 3 yang merupakan kelompok dengan karakteristik terendah. Pergeseran ini terjadi bukan karena indikator Kalimantan Barat menurun drastis, melainkan karena provinsi lain mengalami peningkatan yang lebih pesat sehingga posisi Kalimantan Barat secara relatif tertinggal dibandingkan daerah lain.

Kemudian berdasarkan hasil kelompok menggunakan PCA dan *K-Medoids*, tingkat kemiskinan secara dominan dipengaruhi oleh indikator Indeks Pembangunan Manusia (IPM) dan Rata-Rata Lama Sekolah (RLS) pada tahun 2022 dan 2023. Meskipun indikator sanitasi layak tidak memiliki pengaruh dominan, indikator ini tetap signifikan, terutama untuk wilayah yang memiliki tantangan dalam infrastruktur dasar, seperti provinsi Papua dan Papua Barat. Hasil ini sejalan dengan penelitian sebelumnya, misalnya [5], [6], dan [7], yang menekankan pentingnya infrastruktur dan Indeks Pembangunan Manusia sebagai faktor signifikan dalam menurunkan kemiskinan. Temuan ini juga mengkonfirmasi penelitian [13] bahwa infrastruktur sanitasi menjadi

salah satu pembeda penting antarwilayah, karena kelompok dengan sanitasi rendah cenderung masuk dalam kategori termiskin.

Dengan mempertimbangkan stabilitas pola pengelompokan wilayah berdasarkan indikator pembangunan, kebijakan pembangunan sebaiknya dirancang secara terfokus sesuai karakteristik spesifik masing-masing kelompok. Hal ini sejalan dengan kebijakan pemerintah terkait Sustainable Development Goals (SDGs) di Indonesia yang menekankan integrasi agenda pembangunan berkelanjutan ke dalam RPJMN melalui program-program spesifik di sektor sosial, ekonomi, dan lingkungan [19]. Sebagai contoh, peningkatan akses pendidikan di wilayah dengan Indeks Pembangunan Manusia (IPM) rendah terbukti berkontribusi signifikan pada percepatan pengentasan kemiskinan dan pemerataan kualitas sumber daya manusia [20]. Demikian pula, pengembangan infrastruktur sanitasi di daerah dengan akses terbatas mendukung perbaikan kesehatan masyarakat dan menurunkan beban penyakit, sejalan dengan hasil studi yang menekankan pentingnya infrastruktur dasar dalam mendorong inklusivitas sosial dan ekonomi [21]. Oleh karena itu, kebijakan berbasis kelompok ini diharapkan tidak hanya dapat mengoptimalkan sumber daya yang ada, tetapi juga mendukung pencapaian pembangunan yang lebih merata, efektif, dan konsisten dengan visi Indonesia Emas 2045.

4. Kesimpulan

Penelitian ini berhasil mengidentifikasi pola ketimpangan sosial-ekonomi di 34 provinsi Indonesia dengan menggunakan metode PCA dan *K-Medoids*. Analisis menunjukkan bahwa tingkat kemiskinan dipengaruhi secara dominan oleh IPM dan RLS, dengan sanitasi layak menjadi faktor tambahan yang signifikan. Hasil pengelompokan mencerminkan adanya perbedaan mencolok antar kelompok dalam aspek pendidikan, pembangunan manusia, dan pengeluaran per kapita, sementara ketimpangan ini cenderung stabil antara tahun 2022 dan 2023. Temuan ini menegaskan perlunya kebijakan yang lebih spesifik dan terarah untuk setiap kelompok, terutama dalam peningkatan akses pendidikan dan pemerataan infrastruktur dasar seperti sanitasi layak. Dengan demikian, penelitian ini memberikan kontribusi penting untuk perumusan strategi kebijakan yang mendukung pembangunan berkelanjutan dan pengurangan ketimpangan di Indonesia.

5. Daftar Pustaka

- [1] B. P. Statistik, "Profil Kemiskinan di Indonesia Maret 2023," Maret 2023. [Online]. Available: <https://www.bps.go.id/pressrelease/2023/07/17/2016/profil-kemiskinan-di-indonesia-maret-2023.html>. [Accessed 16 January 2025].
- [2] B. P. Statistik, "Persentase penduduk miskin (P0) menurut provinsi dan daerah (persen)," Augustus 2024. [Online]. Available: <https://www.bps.go.id>. [Accessed 21 January 2025].
- [3] G. B. Kaligis and S. Yulianto, "ANALISA PERBANDINGAN ALGORITMA K-MEANS, K-MEOIDS, DAN X-MEANS UNTUK PENGELOMPOKKAN KINERJA PEGAWAI," *itexplore*, vol. 1, no. 3, pp. 179-193, 2022.
- [4] N. R. Sari, C. Suryawati and N. Nandini, "Evaluasi Pelaksanaan Program Indonesia Sehat dengan Pendekatan Keluarga pada Indikator TB Paru di Puskesmas Tayu II Kabupaten Pati," *Jurnal Kesehatan Masyarakat*, vol. 7, no. 4, pp. 532-541, 2019.
- [5] F. Andrianus and K. Alfatih, "Pengaruh Infrastruktur terhadap Kemiskinan dengan Menggunakan Data Panel 34 Provinsi di Indonesia," *Jurnal Informatika Ekonomi Bisnis*, pp. 56-62, 2023.
- [6] I. S. A. Rohmah and J. A. Prakoso, "Pengaruh IPM, RLS, TPT, dan Pengeluaran Perkapita terhadap Kemiskinan di Provinsi Jawa Barat," *Transekonomika: Akuntansi, Bisnis dan Keuangan*, vol. 2, no. 6, pp. 255-266, 2022.

- [7] I. Azzahara, N. N. Debataraja and S. W. Rizki, "Pengelompokan Kabupaten/Kota di Kalimantan Barat Berdasarkan Indeks Pembangunan Manusia dengan Latent Class Cluster Analysis," *Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya*, vol. 11, no. 1, pp. 43-50, 2022.
- [8] K. H. Izzuddin and A. W. Wijayanto, "Pemodelan Clustering Ward, K-Means, Diana, dan PAM dengan PCA untuk Karakterisasi Kemiskinan Indonesia Tahun 2021," *emodelan Clustering Ward, K-Means, Diana, dan PAM dengan PCA untuk Karakterisasi Kemiskinan Indonesia Tahun 2021*, vol. 13, no. 1, pp. 41-53, 2024.
- [9] H. Sibarani, W. Saputra , I. Gunawan and Z. M. Nasution, "Penerapan Metode K-Means Untuk Pengelompokkan Kabupaten/Kota Di Provinsi Sumatera Utara Berdasarkan Indikator Indeks Pembangunan Manusia," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 6, no. 1, pp. 154-161, 2022.
- [10] Y. R. Sari, A. Sudewa, D. A. Lestari and T. I. Jaya, "Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 5, no. 2, p. 192, 2020.
- [11] S. N. Mayasari and J. Nugraha, "Implementasi K-Means Cluster Analysis untuk Mengelompokkan Kabupaten/Kota Berdasarkan Data Kemiskinan di Provinsi Jawa Tengah Tahun 2022," *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, vol. 3, no. 2, pp. 317-329, 2023.
- [12] F. N. Irawan, "Analisis Clustering K-Means Pada Data Informasi Kemiskinan di Jawa Barat Tahun 2018," *Jurnal Sisfokom*, vol. 8, no. 2, 2019.
- [13] N. I. Tahir, W. Alwi and K. Nurfadilah, "Aplikasi Metode Analisis Komponen Utama (AKU) dalam Mengidentifikasi Faktor yang Memengaruhi Kemiskinan di Kabupaten/Kota Provinsi Sulawesi Selatan," *Journal of Mathematics: Theory and Applications*, pp. 38-44, 2021.
- [14] B. A. Bachtiar and U. Erawanto, "'Comparative Analysis of Disadvantaged Areas in Regencies/Cities in Eastern Indonesia in 2021 Using the K-Means Clustering, K-Medoids Clustering, and Fuzzy C-Means Clustering Methods," *Politeknik Statistika STIS*, 2021.
- [15] D. R. P. Sari, "Metode Principal Component Analysis (PCA) sebagai Penanganan Asumsi Multikolinearitas," *PARAMETER: Jurnal Matematika, Statistika dan Terapannya*, vol. 2, no. 2, pp. 115-124, 2023.
- [16] D. Sanjaya and S. Budi, "Prediksi Pencapaian Target Kerja Menggunakan Metode Deep Learning dan Data Envelopment Analysis," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 6, no. 2, p. 288, 2020.
- [17] H. Herviany, M. Zarlis, E. Irawan, H. Okprana and Y. Pranayama, "Comparison of K-Means and K-Medoids Algorithms for Clustering," *MALCOM: Jurnal Ilmiah Informatika*, vol. 1, no. 1, pp. 34-40, 2021.
- [18] I. J. Panjaitan, T. Wulandari and B. Susetyo, "Perbandingan Hasil Analisis Cluster K-Means dan K-Medoids untuk Pemetaan Mutu SMK," *Jurnal Bayesian: Jurnal Ilmiah Statistika dan Ekonometrika*, vol. 4, no. 1, pp. 1-15, 2024.
- [19] N. H. Zen, L. Mageasti and Yulhendri, "Analisis Penerapan Sdgs Dalam Pembangunan Berkelanjutan Di Indonesia: Tinjauan Literatur dan Tantangan Implementasi," *Gudang Jurnal Multidisiplin Ilmu*, vol. 3, no. 1, pp. 775-785, 2025.
- [20] Republik Indonesia, Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2025–2029, Jakarta: Pemerintah Republik Indonesia, 2025.
- [21] K. K. B. Perekonomian, "Laporan Infrastruktur dan Inklusivitas Ekonomi 2024," Kemenko Perekonomian, Jakarta, 2024.