

# Identification of Indonesian Provinces Based on Socioeconomic Indicators in 2024 Using K-Means

Erika Putri Permatasari<sup>1</sup>, Lathifa Aurellia Iriani<sup>2</sup>, Edy Widodo<sup>3\*</sup>

<sup>1,2,3</sup> Statistics, Islamic University of Indonesia, Yogyakarta, Indonesia,

<sup>1</sup> [23611080@students.uui.ac.id](mailto:23611080@students.uui.ac.id), <sup>2</sup> [23611081@students.uui.ac.id](mailto:23611081@students.uui.ac.id), <sup>3</sup> [edywidodo@uui.ac.id](mailto:edywidodo@uui.ac.id)

\*Corresponding author

---

## ABSTRAK

Kemiskinan merupakan permasalahan pembangunan yang masih menunjukkan ketimpangan antar wilayah di Indonesia. Perbedaan kondisi sosial dan ekonomi antar provinsi menuntut adanya pemetaan wilayah yang komprehensif sebagai dasar perumusan kebijakan yang tepat sasaran. Penelitian ini bertujuan untuk mengelompokkan provinsi-provinsi di Indonesia berdasarkan karakteristik kemiskinan dan pembangunan. Data yang digunakan merupakan data sekunder tahun 2024 yang bersumber dari Badan Pusat Statistik (BPS) dengan unit analisis 38 provinsi. Variabel yang digunakan meliputi Tingkat Kemiskinan, Indeks Pembangunan Manusia (IPM), Tingkat Pengangguran Terbuka (TPT), dan Produk Domestik Regional Bruto (PDRB) per Kapita. Metode analisis yang digunakan adalah *K-Means* dengan data yang telah distandarisasi menggunakan *Z-Score*. Jumlah kluster optimal ditentukan menggunakan metode *Elbow* dan dikonfirmasi dengan *Silhouette Score*. Hasil penelitian menunjukkan bahwa provinsi-provinsi di Indonesia terbagi ke dalam empat kluster dengan karakteristik sosial dan ekonomi yang berbeda. Setiap kluster mencerminkan Tingkat Kemiskinan, kualitas pembangunan manusia, dan kondisi ketenagakerjaan yang beragam. Temuan ini menegaskan bahwa permasalahan kemiskinan di Indonesia bersifat multidimensional, sehingga diperlukan kebijakan pembangunan dan pengentasan kemiskinan yang spesifik sesuai karakteristik masing-masing kluster.

### Kata Kunci:

Kemiskinan; Klusterisasi K-Means; IPM; TPT; PDRB per Kapita

## ABSTRACT

Poverty remains a persistent development challenge in Indonesia, characterized by substantial disparities across regions. Differences in social and economic conditions among provinces highlight the need for a comprehensive regional classification to support the formulation of targeted development policies. This study aims to classify Indonesian provinces based on their poverty and development characteristics. The data used are secondary data for the year 2024 obtained from Statistics Indonesia (Badan Pusat Statistik), with 38 provinces as the units of analysis. The variables include the poverty rate, Human Development Index (HDI), Open Unemployment Rate (OUR), and gross regional domestic product (GRDP) per capita. The analytical method employed is *K-Means* clustering, with variables standardized using *Z-Scores*. The optimal number of clusters was determined using the *Elbow* method and confirmed by the *Silhouette Score*. The results indicate that Indonesian provinces can be grouped into four clusters with distinct social and economic characteristics. Each cluster reflects different levels of poverty, human development quality, and labor market conditions. These findings emphasize that poverty in Indonesia is a multidimensional issue, underscoring the need for development and poverty alleviation policies that are tailored to the specific characteristics of each cluster.

### Keywords:

Poverty; *K-Means* Clustering; Human Development Index (HDI); Open Unemployment Rate (OUR); GRDP per Capita

---

## 1. PENDAHULUAN

Indonesia sebagai negara berkembang masih menghadapi berbagai persoalan mendasar dalam pembangunan, salah satunya adalah tingginya Tingkat Kemiskinan, adanya lonjakan

angka kemiskinan masih menjadi tantangan serius dalam kehidupan masyarakat Indonesia [1]. Kemiskinan telah berkembang menjadi masalah sosial multidimensi yang memerlukan perhatian serius dari pemerintah di Indonesia [2].

Secara konseptual, kemiskinan didefinisikan sebagai ketidakmampuan individu atau masyarakat dalam memenuhi kebutuhan dasar, baik makanan maupun non-makanan, yang diukur dari sisi pengeluaran [3]. Kemiskinan ini merupakan ketidakmampuan seseorang ataupun masyarakat wilayah sekitar untuk memenuhi kebutuhan dasar, sementara lingkungan pendukungnya kurang memberikan peluang untuk meningkatkan kesejahteraan masyarakat baik kebutuhan dalam hal makanan, minuman, pakaian, tempat tinggal yang menetap, pendidikan anak hingga kesehatan yang diukur dari sisi pengeluaran. Meskipun telah berjuang puluhan tahun untuk melepaskan diri dari kemiskinan, tetapi kenyataan memperlihatkan bahwa sampai saat ini Indonesia belum bisa melepaskan diri dari belenggu masalah Tingkat Kemiskinan [4].

Dalam menganalisis kemiskinan, beberapa indikator ekonomi memegang peranan penting. Pertumbuhan ekonomi yang diukur melalui Produk Domestik Regional Bruto (PDRB) menjadi salah satu indikator kunci. PDRB merupakan ukuran untuk menilai kondisi perekonomian di suatu wilayah yang diperoleh melalui perhitungan total nilai produk dan layanan yang diproduksi oleh pelaku ekonomi di suatu wilayah [5]. Pendapatan per Kapita yang berasal dari PDRB menjadi indikator moneter atas kegiatan ekonomi penduduk suatu wilayah. Semakin tinggi pendapatan per Kapita maka akan semakin baik kondisi perekonomian dan kesejahteraan masyarakat [6]. PDRB dapat diartikan sebagai jumlah nilai tambah yang dihasilkan oleh seluruh unit usaha atau merupakan jumlah seluruh nilai barang dan jasa oleh seluruh unit ekonomi di suatu wilayah [7]. Namun, pertumbuhan ekonomi saja tidak cukup untuk mengatasi kemiskinan. Tingkat pengangguran menjadi faktor penentu signifikan lainnya. Pengangguran merupakan keadaan yang keberadaanya tidak terelakkan, baik itu di negara berkembang maupun di negara maju. Pengangguran berpengaruh terhadap kerawanan kriminalitas, gejolak sosial, politik dan kemiskinan [8]. Indonesia dihadapkan pada keterbatasan lapangan kerja dibandingkan jumlah angkatan kerja yang tercipta. Beberapa masalah ketenagakerjaan yang dapat berdampak pada pengangguran antara lain: (1) *mismatch* antara penawaran dan permintaan tenaga kerja; (2) permasalahan upah tenaga kerja; (3) inflasi yang mendorong efisiensi tenaga kerja; (4) keterbatasan kemampuan pertumbuhan ekonomi dalam menciptakan kesempatan kerja [9].

Tingkat Pengangguran Terbuka (TPT) merupakan indikator kunci dalam menganalisis kondisi ketenagakerjaan suatu wilayah. Menurut teori ekonomi, pengangguran memiliki hubungan positif dengan kemiskinan, dimana ketiadaan pendapatan secara langsung mempengaruhi kemampuan rumah tangga dalam memenuhi kebutuhan dasarnya [10].

Indeks pembangunan manusia (IPM) menjadi indikator komprehensif lainnya yang berkaitan erat dengan kemiskinan. IPM digunakan untuk mengukur seberapa besar dampak yang ditimbulkan dari upaya peningkatan kemampuan modal dasar manusia [11]. Pembangunan manusia menitik beratkan pada peningkatan kualitas manusia yang diukur melalui pendidikan, kesehatan, dan daya beli dasar manusia. Semakin tinggi angka yang diperoleh maka semakin tercapai tujuan dari pembangunan. Pembangunan merupakan sebuah proses untuk melakukan perubahan ke arah yang lebih baik dan bila peningkatan IPM ini diharapkan dapat membawa dampak positif bagi 8 peningkatan kualitas sumber daya manusia yang nantinya akan mempengaruhi penurunan Tingkat Kemiskinan di kabupaten/kota di Indonesia [12].

Berbagai provinsi di Indonesia memiliki karakteristik ekonomi, sosial, dan pembangunan yang berbeda-beda. Data BPS Tahun 2023 [11] menunjukkan adanya variasi TPT yang cukup signifikan antarwilayah, serta perbedaan IPM yang mencerminkan ketimpangan kualitas

sumber daya manusia. Perbedaan tersebut mengindikasikan bahwa kondisi kemiskinan tidak dapat dianalisis hanya dengan satu indikator atau pendekatan tunggal. Namun demikian, kajian yang memetakan provinsi berdasarkan kombinasi beberapa variabel kunci secara simultan masih terbatas, sehingga diperlukan pendekatan baru yang mampu menggambarkan pola dan pengelompokan wilayah berdasarkan karakteristik multidimensional tersebut.

Untuk menjawab kesenjangan penelitian tersebut, pendekatan analisis kluster digunakan untuk mengelompokkan objek berdasarkan kemiripan karakteristik. Dengan menggunakan variabel Tingkat Kemiskinan, IPM, TPT, dan PDRB analisis kluster dapat memberikan gambaran yang lebih komprehensif mengenai struktur dan perbedaan kondisi antar provinsi di Indonesia.

Berdasarkan kompleksitas permasalahan tersebut, penelitian ini bertujuan untuk mengelompokkan provinsi di Indonesia berdasarkan empat variabel kunci menggunakan analisis kluster dengan metode K-Means. Pendekatan ini diharapkan dapat memberikan rekomendasi kebijakan yang lebih tepat sasaran sesuai karakteristik masing-masing provinsi.

## **2. METODE**

### **2.1. Kajian Teori**

Kemiskinan merupakan kondisi ketidakmampuan individu atau kelompok dalam memenuhi kebutuhan dasar hidup yang diukur berdasarkan pengeluaran per Kapita [3]. Kemiskinan bersifat multidimensional dan dapat muncul dalam berbagai pola, seperti kemiskinan kronis, siklus, musiman, maupun akibat kejadian tertentu seperti bencana alam atau kebijakan ekonomi [13]. Oleh karena itu, pemahaman kemiskinan memerlukan pendekatan yang mampu menangkap kompleksitas kondisi sosial dan ekonomi secara simultan.

Dalam pembangunan wilayah, PDRB per Kapita sering digunakan sebagai indikator kesejahteraan ekonomi, dimana peningkatan PDRB diharapkan dapat menurunkan kemiskinan melalui peningkatan pendapatan dan penciptaan lapangan kerja [14]. Namun, pertumbuhan ekonomi yang tidak diiringi pemerataan dapat menyebabkan ketimpangan antar wilayah. Selain itu, tingginya TPT berkontribusi langsung terhadap peningkatan kemiskinan karena keterbatasan pendapatan rumah tangga [10]. Di sisi lain, IPM mencerminkan kualitas sumber daya manusia melalui dimensi kesehatan, pendidikan, dan standar hidup layak, yang berperan penting dalam meningkatkan produktivitas dan menekan kemiskinan [11].

Untuk memetakan kondisi sosial-ekonomi antarwilayah, metode K-Means banyak digunakan karena mampu mengelompokkan wilayah berdasarkan tingkat kemiripan karakteristik. Metode ini bekerja dengan meminimalkan variasi dalam kluster melalui pengukuran jarak, umumnya menggunakan jarak Euclidean, serta memerlukan standarisasi data agar perbedaan skala variabel tidak mempengaruhi hasil pengelompokan [15]. Sejumlah penelitian terdahulu menunjukkan bahwa K-Means efektif dalam mengidentifikasi perbedaan karakteristik pembangunan antar provinsi di Indonesia, baik berdasarkan indikator kesejahteraan, pembangunan ekonomi, maupun PDRB wilayah.

Meskipun demikian, penelitian yang secara simultan mengintegrasikan indikator kemiskinan, IPM, TPT, dan PDRB per Kapita dalam satu kerangka analisis kluster pada tingkat provinsi masih terbatas. Oleh karena itu, penelitian ini bertujuan untuk mengisi celah tersebut dengan menggunakan metode K-Means untuk menghasilkan pemetaan kemiskinan dan pembangunan yang lebih komprehensif serta relevan sebagai dasar perumusan kebijakan pembangunan wilayah.

## 2.2. Penelitian Terdahulu

Sejumlah penelitian terdahulu telah membahas permasalahan kemiskinan, kesejahteraan, dan ketimpangan pembangunan wilayah di Indonesia dengan berbagai pendekatan metodologis. Salah satu pendekatan yang banyak digunakan adalah analisis klaster, khususnya algoritma K-Means, untuk mengelompokkan wilayah berdasarkan indikator pembangunan tertentu. Pendekatan ini dinilai efektif dalam memetakan karakteristik wilayah secara spasial dan memberikan gambaran heterogenitas pembangunan antarprovinsi.

Penelitian oleh Saputra dan Hidayat [9] menggunakan algoritma K-Means untuk mengelompokkan provinsi di Indonesia berdasarkan indikator Tujuan Pembangunan Berkelanjutan (SDGs) dan menghasilkan empat klaster provinsi dengan karakteristik pembangunan yang berbeda. Sementara itu, Etrisia dan Findi [16] mengkombinasikan analisis faktor dan analisis klaster untuk pengelompokan provinsi berdasarkan indikator pembangunan ekonomi, seperti IPM, Tingkat Kemiskinan, TPT, dan PDRB, yang menghasilkan lima klaster dengan permasalahan pembangunan yang beragam.

Beberapa penelitian lain menggunakan pendekatan klaster dengan fokus pada indikator sektoral tertentu. Sari dan Pratama [17] mengelompokkan provinsi berdasarkan kesiapan teknologi informasi dan komunikasi menggunakan metode K-Means dan klaster hierarkis, dan menemukan keterkaitan antara hasil klaster dengan tingkat PDRB per Kapita. Etrisia dan Findi [16] menggunakan metode K-Means untuk mengklasifikasikan tingkat inflasi di ibu kota provinsi di Indonesia sebelum dan saat pandemi COVID-19, yang menghasilkan tiga klaster dengan tingkat inflasi tinggi, sedang, dan rendah.

Pendekatan K-Means juga banyak digunakan untuk memetakan kesejahteraan masyarakat dan kondisi ekonomi wilayah. Wahyudi, Rahmadenni, dkk [18] mengelompokkan provinsi di Indonesia berdasarkan karakteristik kesejahteraan rakyat dan memperoleh tiga klaster utama dengan distribusi provinsi yang tidak merata. Wahyudi, Rahmadenni, dkk [18] melakukan pengelompokan kabupaten/kota di Indonesia berdasarkan PDRB dan menghasilkan empat klaster yang merepresentasikan tingkat kinerja ekonomi daerah dari sangat rendah hingga tinggi. Penelitian serupa juga dilakukan oleh Rachmad [18] yang mengelompokkan provinsi berdasarkan karakteristik kepemilikan tempat tinggal menggunakan algoritma K-Means.

Di sisi lain, beberapa penelitian menggunakan pendekatan regresi data panel untuk menganalisis faktor-faktor yang mempengaruhi kemiskinan. Amalina dan Hartono [6] menunjukkan bahwa IPM dan pertumbuhan ekonomi berpengaruh signifikan terhadap penurunan Tingkat Kemiskinan di Indonesia. Firdaus dan Setiawan [4] menemukan bahwa TPT merupakan salah satu faktor dominan yang mempengaruhi kemiskinan. Meskipun memberikan pemahaman kausal yang kuat, pendekatan regresi cenderung tidak memberikan gambaran pengelompokan wilayah secara eksplisit.

Berdasarkan kajian terhadap penelitian-penelitian terdahulu tersebut, dapat diidentifikasi adanya celah penelitian (*research gap*). Hingga saat ini, belum banyak penelitian yang secara simultan mengintegrasikan indikator ekonomi dan sosial utama, yaitu Tingkat Kemiskinan, IPM, TPT, dan PDRB per Kapita, dalam satu kerangka analisis klaster untuk memetakan kondisi kemiskinan secara komprehensif di tingkat provinsi secara nasional. Sebagian penelitian menggunakan terlalu banyak indikator, sementara penelitian lain lebih menekankan pada pendekatan regresi tanpa menghasilkan peta klaster wilayah.

Oleh karena itu, penelitian ini bertujuan untuk mengisi celah tersebut dengan menerapkan analisis klasterisasi K-Means terhadap seluruh provinsi di Indonesia. Kebaruan (*novelty*) penelitian ini terletak pada cakupan wilayah yang bersifat nasional, integrasi seimbang antara indikator sosial dan ekonomi utama, serta penekanan pada hasil pengelompokan wilayah yang

dapat digunakan sebagai dasar perumusan rekomendasi kebijakan pembangunan dan pengentasan kemiskinan yang lebih terarah di tingkat nasional maupun daerah.

### 2.3. Data

Penelitian ini menggunakan data sekunder berupa data periode tahun 2024. Sumber data diperoleh dari Badan Pusat Statistik (BPS) untuk data Tingkat Kemiskinan, IPM, TPT, dan PDRB per Kapita (Ribuan Rupiah).

### 2.4. Variabel Penelitian

Variabel yang digunakan pada penelitian ini dijelaskan pada **Tabel 1**. Sebagai berikut.

**Table 1.** Definisi Operasional Variabel Penelitian

| Variabel           | Satuan     | Definisi Operasional                                                                                                                                 |
|--------------------|------------|------------------------------------------------------------------------------------------------------------------------------------------------------|
| Tingkat Kemiskinan | Persen (%) | Tingkat kemiskinan merupakan persentase penduduk dengan pengeluaran per kapita per bulan di bawah garis kemiskinan terhadap total penduduk provinsi. |
| IPM                | Persen (%) | IPM merupakan indikator komposit yang mengukur capaian pembangunan manusia yang meliputi kesehatan, pengetahuan, dan standar hidup layak.            |
| TPT                | Persen (%) | TPT merupakan persentase angkatan kerja yang tidak bekerja dan sedang mencari pekerjaan terhadap total angkatan kerja provinsi.                      |
| PDRB per Kapita    | Persen (%) | PDRB per Kapita merupakan nilai Produk Domestik Regional Bruto atas dasar harga berlaku dibagi dengan jumlah penduduk pertengahan tahun provinsi.    |

### 2.5. Tahapan Penelitian

Prosedur penelitian ini disusun secara terstruktur mulai dari pengumpulan data hingga penarikan kesimpulan. Adapun tahapan-tahapan tersebut disajikan dalam bentuk *flowchart* pada **Gambar 1**. dan dapat dijelaskan sebagai berikut.

#### 1. Pengumpulan Data

Data yang digunakan adalah data sekunder yang diperoleh dari situs resmi Badan Pusat Statistik (BPS). Data yang dikumpulkan mencakup 38 provinsi di Indonesia dengan variabel Tingkat Kemiskinan, IPM, TPT, dan PDRB per Kapita pada periode tahun 2024.

#### 2. Preprocessing Data

Tahap *preprocessing* data dilakukan untuk memastikan kualitas data sebelum dilakukan analisis kluster. Pada tahap ini dilakukan:

- Pengecekan terhadap *missing value* dan *outlier* pada setiap variabel penelitian.
- Standarisasi data menggunakan metode *Z-Score* untuk menyamakan skala pengukuran antar variabel.

### 3. Penentuan Jumlah Kluster Optimal

Penentuan jumlah kluster optimal dilakukan dengan:

- Menggunakan *Elbow Method* dengan memplot nilai *Within Cluster Sum of Squares* (WCSS) terhadap jumlah kluster.
- Titik siku (*elbow*) pada grafik menunjukkan jumlah kluster optimal.
- Konfirmasi menggunakan *Silhouette Score* untuk memastikan kualitas pengelompokan.

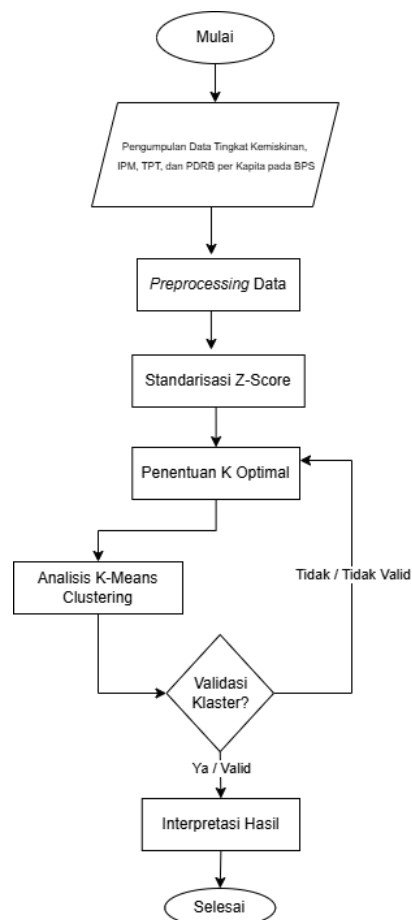
### 4. Pembentukan Kluster dengan K-Means

Pembentukan kluster dilakukan dengan menerapkan algoritma klusterisasi K-Means. Algoritma ini bekerja dengan meminimalkan variasi dalam kluster melalui proses iteratif hingga mencapai kondisi konvergen, yaitu ketika tidak terjadi perubahan keanggotaan kluster.

### 5. Validasi dan Interpretasi Kluster

Tahap akhir analisis meliputi:

- Analisis karakteristik masing-masing kluster berdasarkan nilai centroid yang terbentuk.
- Visualisasi hasil kluster untuk memudahkan interpretasi.



**Gambar 1.** Flowchart Tahapan Penelitian

## 2.6. Klasterisasi K-Means

Analisis kluster merupakan metode analisis yang tidak memiliki asumsi tentang jumlah kelompok atau struktur kelompok. Pengelompokan dilakukan atas dasar kesamaan atau jarak. *Input* yang diperlukan adalah ukuran kesamaan atau data yang persamaannya dapat dihitung. Dimana objek-objek pengamatan dibagi menjadi beberapa kelompok, objek-objek dalam satu kelompok memiliki banyak persamaan, sedangkan dengan kelompok lain memiliki banyak perbedaan sehingga terbentuk pengelompokan secara alami [15].

Selain itu, algoritma klasterisasi K-Means juga salah satu algoritma dengan *partitional*, karena metode K-Means berdasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai *centroid* awalnya. Dibutuhkan jumlah kluster awal yang diinginkan sebagai masukan dan menghasilkan titik *centroid* akhir sebagai *output*. Metode K-Means akan memilih pola  $K$  sebagai diawal *centroid* secara acak atau random. Jumlah iterasi untuk mencapai kluster *centroid* akan dipengaruhi oleh calon kluster *centroid* awal secara random. Sehingga didapat cara dalam pengembangan algoritma dengan menentukan *centroid* kluster yang dilihat dari kepadatan data awal yang tinggi agar mendapatkan kinerja yang lebih tinggi [20].

## 2.7. Standarisasi Data (Z-Score)

Standarisasi data digunakan untuk menyamakan skala antar variabel. Rumus untuk menentukan *Z-Score* sebagai berikut: [21].

$$Z = \frac{X - \mu}{\sigma} \quad (1)$$

Keterangan :

$Z$  = Nilai terstandarisasi

$X$  = Nilai asli

$\mu$  = Rata – rata

$\sigma$  = Standar deviasi

## 2.8. Jarak Euclidean

Jarak Euclidean digunakan dalam algoritma K-Means untuk mengukur jarak antar objek dengan menggunakan rumus sebagai berikut: [21].

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (2)$$

Keterangan :

$D_e$  = Euclidean Distance

$i$  = Banyaknya Data

$(s,t)$  = Koordinat *Centroid*

$(x,y)$  = Koordinat Objek

## 3. HASIL DAN PEMBAHASAN

### 3.1. Analisis Deskriptif

Analisis deskriptif dilakukan berdasarkan **Tabel 2.** untuk mendapatkan gambaran variabel-variabel yang akan digunakan dalam penelitian. Penelitian ini menggunakan data observasi dari 38 provinsi di Indonesia.

**Tabel 2.** Statistik Deskriptif

| Variabel       | Minimum | Maximum | Mean  | Std. Deviation |
|----------------|---------|---------|-------|----------------|
| X <sub>1</sub> | 4.00    | 32.97   | 11.17 | 6.747          |
| X <sub>2</sub> | 53.42   | 83.08   | 72.39 | 5.150          |
| X <sub>3</sub> | 1.250   | 6.850   | 4.374 | 1.396          |
| X <sub>4</sub> | 18.11   | 344.35  | 84.34 | 62.197         |

Data ini terdiri dari empat indikator utama pembangunan daerah di Indonesia, yaitu Tingkat Kemiskinan, IPM, TPT, dan PDRB per Kapita, yang masing-masing menggambarkan kondisi sosial-ekonomi di 38 provinsi di Indonesia. keempat variabel tersebut memiliki karakteristik yang berbeda-beda dan memberikan gambaran penting mengenai kualitas hidup serta kondisi kesejahteraan penduduk di tiap wilayah.

Indikator pertama adalah Tingkat Kemiskinan (X<sub>1</sub>). Rata-rata Tingkat Kemiskinan provinsi di Indonesia tahun 2024 berada di sekitar 11.17 persen. Nilainya bervariasi cukup lebar, mulai dari 4 persen sebagai yang terendah hingga 32.97 persen sebagai yang tertinggi. Rentang yang besar ini menunjukkan adanya kesenjangan kesejahteraan antar provinsi. Beberapa provinsi memiliki Tingkat Kemiskinan yang relatif rendah, terutama wilayah dengan ekonomi yang lebih maju, sementara beberapa daerah, terutama di Indonesia Timur, masih memiliki Tingkat Kemiskinan yang cukup tinggi sehingga memerlukan perhatian lebih dari sisi kebijakan sosial dan pembangunan ekonomi.

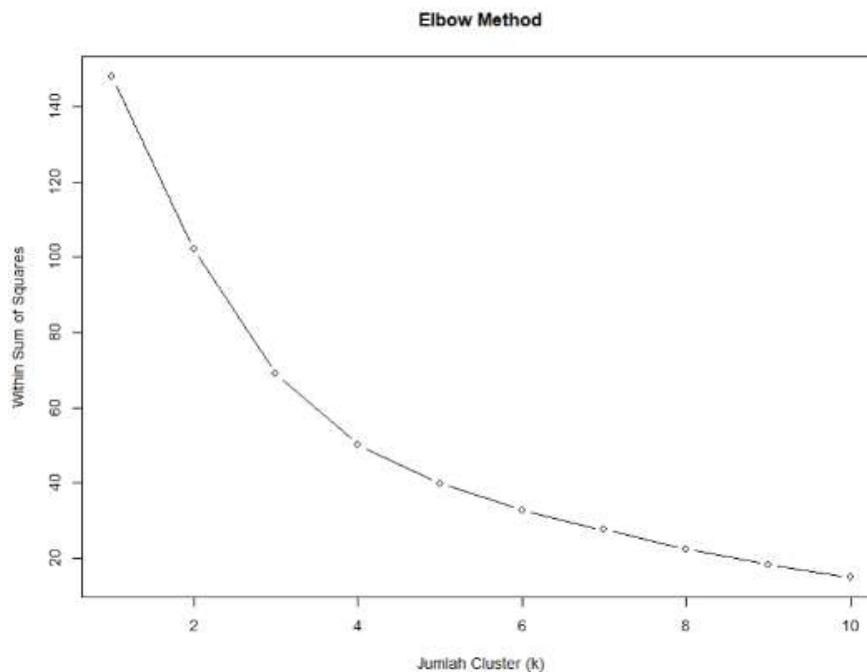
Indikator kedua adalah IPM (X<sub>2</sub>), yang mencerminkan capaian pembangunan dari aspek pendidikan, kesehatan, dan standar hidup layak. Nilai rata-rata IPM provinsi adalah 72.39, dengan minimum 53.42 dan maksimum 83.08. IPM nasional berada pada kategori “tinggi”, tetapi variasi antarprovinsi masih cukup besar. Provinsi dengan IPM rendah biasanya menghadapi tantangan pada pendidikan dan akses layanan kesehatan. Sementara itu, provinsi dengan IPM tinggi menunjukkan kualitas pembangunan manusia yang lebih baik, seperti akses pendidikan lebih merata dan layanan kesehatan yang lebih memadai. Selanjutnya adalah TPT (X<sub>3</sub>), nilai rata-rata TPT sebesar 4.37 persen, dengan nilai minimum 1.25 persen dan maksimum 6.85 persen. Secara umum, angka ini menunjukkan bahwa sebagian besar provinsi memiliki tingkat pengangguran yang relatif terkendali. Namun, adanya provinsi dengan TPT mendekati 7 persen menandakan bahwa lapangan kerja masih belum sepenuhnya merata. Daerah dengan TPT tinggi biasanya merupakan wilayah perkotaan besar yang padat penduduk atau wilayah industri yang terdampak dinamika ekonomi.

Indikator terakhir adalah PDRB per Kapita (X<sub>4</sub>), yang menggambarkan kemampuan daerah dalam menciptakan nilai tambah ekonomi. Data menunjukkan rata-rata PDRB per Kapita sebesar Rp 84.34 juta per tahun. Rentangnya sangat lebar, dari Rp 18.11 juta sebagai yang terendah hingga Rp 344.35 juta sebagai yang tertinggi. Perbedaan yang sangat besar ini menggambarkan ketimpangan ekonomi antar provinsi di Indonesia. Beberapa wilayah dengan industri besar atau kekayaan sumber daya alam tinggi memiliki PDRB per Kapita jauh di atas rata-rata nasional, sementara daerah dengan struktur ekonomi kecil memiliki kemampuan produksi yang lebih rendah. Perbedaan ini menjadi tantangan utama dalam pemetaan pembangunan nasional.



### 3.2. Inisiasi Klaster

Standarisasi data dilakukan sebelum melakukan analisis karena variabel memiliki denominasi yang berbeda. Langkah selanjutnya, algoritma K-Means dimulai dengan menentukan jumlah klaster. Penentuan jumlah klaster pada penelitian ini menggunakan *elbow method* dengan melihat kurva pada *scree plot* seperti pada **Gambar 2**.



**Gambar 2.** *Scree Plot*

Berdasarkan grafik *elbow method* diatas, terlihat bahwa nilai *Within Sum of Squares* (WSS) mengalami penurunan yang cukup tajam dari jumlah klaster  $k = 1$  hingga  $k = 4$ . Penurunan WSS yang curam pada rentang tersebut menunjukkan bahwa penambahan *cluster* memberikan peningkatan kualitas pengelompokan yang cukup signifikan.

Tetapi, setelah memasuki titik  $k = 4$ , penurunan nilai WSS mulai melandai dan grafik tidak lagi menunjukkan perubahan besar pada penambahan jumlah klaster berikutnya. Kurva yang mulai datar setelah  $k = 4$  mengindikasikan bahwa tambahan klaster diatas empat tidak memberikan penurunan WSS yang berarti dan cenderung hanya menambah kompleksitas model.

Titik  $k = 4$  dapat dianggap sebagai titik “siku” (*elbow*) pada grafik. Oleh karena itu, jumlah klaster yang paling optimal adalah empat klaster, karena pada titik tersebut model sudah mampu mengelompokkan data secara efisien tanpa perlu penambahan klaster tambahan yang tidak memberikan manfaat yang signifikan.

Berdasarkan penentuan jumlah klaster tersebut, maka pembagian anggota untuk masing-masing kelompok diinisiasikan ke dalam empat klaster yang disajikan pada **Tabel 3**. berikut ini.

**Tabel 3.** *Inisiasi Cluster*

| Klaster 1                                         | Klaster 2                                                      | Klaster 3                                                                   | Klaster 4                             |
|---------------------------------------------------|----------------------------------------------------------------|-----------------------------------------------------------------------------|---------------------------------------|
| Bengkulu,<br>Gorontalo, NTB,<br>NTT, Papua Barat, | Aceh, Banten, Jawa<br>Barat, Maluku, Papua,<br>Sulawesi Utara, | Bali, Bangka Belitung, DI<br>Yogyakarta, Jambi, Jawa<br>Tengah, Jawa Timur, | Jakarta,<br>Kalimantan<br>Timur, Kep. |

|                                    |                                   |                                                                                                                                                                              |      |
|------------------------------------|-----------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| Sulawesi Barat,<br>Sulawesi Tengah | Sumatera Barat,<br>Sumatera Utara | Kalimantan Barat, Kalimantan<br>Selatan, Kalimantan Tengah,<br>Kalimantan Utara, Lampung,<br>Maluku Utara, Riau, Sulawesi<br>Selatan, Sulawesi Tenggara,<br>Sumatera Selatan | Riau |
|------------------------------------|-----------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|

Berdasarkan pengelompokan diatas, pemetaan provinsi di Indonesia dapat disajikan pada **Gambar 3.** berikut ini:



**Gambar 3.** Pemetaan Provinsi Berdasarkan Dimensi Sosial Ekonomi, Tahun 2024

Berdasarkan hasil klasterisasi K-Means menggunakan data Kemiskinan, IPM, TPT, dan PDRB tahun 2024, seluruh provinsi di Indonesia berhasil dikelompokkan ke dalam empat klaster. Wilayah yang ditandai sebagai NA pada peta bukan disebabkan oleh ketidakikutsertaan provinsi tersebut dalam proses klasterisasi, melainkan akibat keterbatasan data spasial atau perbedaan penamaan wilayah pada proses pemetaan. Oleh karena itu, NA pada peta tidak memengaruhi hasil analisis klaster secara keseluruhan.

Klaster 1 (warna biru) merupakan kelompok provinsi dengan **Tingkat Kemiskinan** relatif tinggi, **IPM** rendah hingga menengah, serta **PDRB per Kapita** yang rendah dibandingkan klaster lain. **TPT** pada klaster ini umumnya relatif rendah hingga sedang, yang menunjukkan bahwa sebagian besar penduduk bekerja, namun bekerja pada sektor dengan produktivitas dan pendapatan rendah. Kondisi ini mencerminkan wilayah dengan keterbatasan kapasitas ekonomi dan kualitas sumber daya manusia. Oleh karena itu, kebijakan pembangunan pada klaster ini perlu difokuskan pada peningkatan kualitas pendidikan dan kesehatan, pengembangan sektor ekonomi lokal, serta penciptaan nilai tambah agar pertumbuhan ekonomi mampu menurunkan Tingkat Kemiskinan secara berkelanjutan.

Klaster 2 (warna hijau) menunjukkan provinsi dengan **Tingkat Kemiskinan** berada pada kisaran menengah, **IPM** menengah, serta **PDRB per Kapita** yang belum tinggi. **TPT** pada klaster ini cenderung lebih tinggi dibandingkan klaster 1, yang mengindikasikan adanya ketidakseimbangan antara pertumbuhan angkatan kerja dan penyerapan tenaga kerja. Wilayah dalam klaster ini memiliki potensi ekonomi yang cukup besar, namun belum sepenuhnya mampu mengkonversi pertumbuhan ekonomi menjadi peningkatan kesejahteraan Masyarakat secara merata. Oleh karena itu, diperlukan kebijakan peningkatan kualitas tenaga kerja dan penciptaan lapangan kerja yang lebih luas.

Klaster 3 (warna oranye) merupakan kelompok provinsi dengan kondisi ekonomi menengah ke atas. Wilayah dalam klaster ini memiliki **Tingkat Kemiskinan** relatif lebih rendah, **IPM** menengah hingga tinggi, serta **PDRB per Kapita** yang berada di sekitar rata-rata nasional. **TPT** relatif terkendali, mencerminkan pasar tenaga kerja yang cukup stabil. Klaster ini mencerminkan wilayah dengan struktur ekonomi yang relatif seimbang, meskipun masih menghadapi tantangan dalam meningkatkan produktivitas dan daya saing ekonomi.

Klaster 4 (warna merah) merupakan kelompok provinsi dengan kinerja sosial ekonomi terbaik. Wilayah dalam klaster ini memiliki **PDRB per Kapita** yang sangat tinggi, **IPM** tinggi, serta **Tingkat Kemiskinan** yang rendah. Namun, **TPT** relatif lebih tinggi dibandingkan klaster lain, yang mencerminkan struktur ekonomi modern dengan pasar tenaga kerja yang lebih kompetitif dan formal. Klaster ini menunjukkan wilayah dengan produktivitas ekonomi tinggi, namun masih menghadapi tantangan pemerataan kesejahteraan dan penyerapan tenaga kerja.

Langkah terakhir Algoritma K-Means adalah menentukan *centroid* akhir. *Centroid* akhir seperti pada **Tabel 4**. dapat digunakan untuk menganalisis karakteristik klaster:

**Tabel 4.** Pusat Klaster Akhir

|                               | Klaster |         |        |       |
|-------------------------------|---------|---------|--------|-------|
|                               | 1       | 2       | 3      | 4     |
| <b>Zscore (X<sub>1</sub>)</b> | 2.99    | -.00608 | -.0849 | -.813 |
| <b>Zscore (X<sub>2</sub>)</b> | -3.07   | .177    | -.0205 | 1.00  |
| <b>Zscore (X<sub>3</sub>)</b> | -1.75   | 1.23    | -.477  | .586  |
| <b>Zscore (X<sub>4</sub>)</b> | -.256   | -.387   | -.301  | 2.12  |

Analisis karakteristik setiap klaster dilakukan berdasarkan nilai pusat klaster akhir (*centroid*) dalam bentuk *Z-Score*. Nilai *Z-Score* positif menunjukkan bahwa nilai indikator pada klaster tersebut berada di atas rata-rata nasional, sedangkan nilai *Z-Score* negatif menunjukkan bahwa nilai indikator berada di bawah rata-rata nasional. Berdasarkan hasil pengelompokan menggunakan algoritma K-Means dengan empat klaster, diperoleh karakteristik masing-masing klaster sebagai berikut.

Klaster pertama merupakan kelompok provinsi yang memiliki karakteristik Tingkat Kemiskinan sangat tinggi, yang ditunjukkan oleh nilai *Z-Score* kemiskinan sebesar 2.99. Di sisi lain, klaster ini memiliki nilai *Z-Score* IPM sebesar -3.07 yang menunjukkan bahwa kualitas pembangunan manusia berada pada kondisi rendah dibandingkan dengan klaster yang lainnya. Selain itu, TPT pada klaster ini relatif lebih rendah dari rata-rata, dan PDRB per Kapita juga berada sedikit di bawah rata-rata nasional. Kombinasi karakteristik tersebut menunjukkan bahwa wilayah pada klaster ini menghadapi persoalan struktural, yaitu tingginya Tingkat Kemiskinan yang tidak diimbangi oleh kualitas sumber daya manusia dan kapasitas ekonomi yang memadai. Rendahnya IPM mengindikasikan keterbatasan pada aspek pendidikan, kesehatan, dan standar hidup layak, sehingga meskipun tingkat pengangguran relatif rendah, sebagian besar penduduk masih bekerja pada sektor dengan produktivitas dan pendapatan yang rendah. Kondisi ini mengisyaratkan perlunya kebijakan pembangunan yang berfokus pada peningkatan kualitas manusia melalui perbaikan akses pendidikan, kesehatan, serta penciptaan lapangan kerja produktif yang mampu meningkatkan kesejahteraan masyarakat.

Klaster kedua dicirikan oleh Tingkat Kemiskinan yang berada pada kisaran rata-rata nasional *Z-Score* -0.006, dengan nilai IPM yang sedikit lebih tinggi dari rata-rata *Z-Score* 0.177. Namun demikian, klaster ini memiliki tingkat pengangguran terbuka yang cukup tinggi, ditunjukkan oleh nilai *Z-Score* TPT sebesar 1.23, sementara PDRB per Kapita berada

di bawah rata-rata nasional *Z-Score*  $-0.387$ . Karakteristik ini menunjukkan bahwa wilayah pada klaster kedua memiliki kualitas pembangunan manusia yang relatif lebih baik, namun belum diikuti oleh kemampuan perekonomian daerah dalam menyerap tenaga kerja secara optimal. Tingginya tingkat pengangguran mengindikasikan adanya ketidakseimbangan antara penawaran dan permintaan tenaga kerja, meskipun kualitas sumber daya manusia cenderung lebih baik dibandingkan klaster pertama. Oleh karena itu, kebijakan yang diperlukan pada klaster ini adalah penguatan sektor-sektor produktif dan padat karya, peningkatan iklim investasi, serta pengembangan kewirausahaan untuk menciptakan lapangan kerja baru yang sesuai dengan kompetensi penduduk.

Klaster ketiga memiliki karakteristik yang relatif mendekati rata-rata nasional pada hampir seluruh indikator. Nilai *Z-Score* kemiskinan sebesar  $-0.0849$  dan IPM sebesar  $-0.0205$  menunjukkan bahwa Tingkat Kemiskinan dan pembangunan manusia pada klaster ini berada pada kondisi yang hampir seimbang dengan rata-rata nasional. Tingkat pengangguran terbuka relatif lebih rendah dari rata-rata *Z-Score*  $-0.477$ , sementara PDRB per Kapita juga masih berada di bawah rata-rata nasional *Z-Score*  $-0.301$ . Kondisi tersebut mengindikasikan bahwa wilayah pada klaster ketiga berada pada tahap pembangunan menengah, dengan permasalahan ekonomi dan sosial yang tidak terlalu ekstrem. Meskipun tingkat pengangguran relatif rendah, kapasitas ekonomi daerah masih terbatas sehingga belum mampu mendorong peningkatan kesejahteraan secara signifikan. Kebijakan pembangunan pada klaster ini perlu diarahkan pada peningkatan produktivitas sektor ekonomi unggulan daerah, penguatan nilai tambah, serta peningkatan kualitas tenaga kerja agar pertumbuhan ekonomi yang terjadi dapat lebih inklusif dan berkelanjutan.

Klaster keempat merupakan kelompok provinsi dengan kondisi pembangunan ekonomi paling maju, yang ditunjukkan oleh nilai *Z-Score* PDRB per Kapita sebesar 2.12, tertinggi di antara seluruh klaster. Selain itu, klaster ini juga memiliki kualitas pembangunan manusia yang tinggi dengan nilai *Z-Score* IPM sebesar 1.00. Tingkat Kemiskinan pada klaster ini tergolong rendah, tercermin dari nilai *Z-Score* kemiskinan sebesar  $-0.813$ . Namun demikian, tingkat pengangguran terbuka relatif sedikit di atas rata-rata nasional *Z-Score* 0.586. Karakteristik tersebut menunjukkan bahwa wilayah pada klaster keempat memiliki kapasitas ekonomi dan kualitas sumber daya manusia yang sangat baik, namun masih menghadapi tantangan dalam penyerapan tenaga kerja secara optimal. Tingginya PDRB per Kapita dan IPM mengindikasikan tingginya produktivitas ekonomi dan kualitas hidup penduduk, tetapi peningkatan pengangguran menunjukkan adanya kemungkinan mismatch antara kebutuhan pasar tenaga kerja dan kompetensi tenaga kerja. Oleh karena itu, kebijakan yang diperlukan pada klaster ini adalah penguatan inovasi, pengembangan ekonomi berbasis pengetahuan dan teknologi, serta peningkatan keterkaitan antara dunia pendidikan dan dunia kerja.

#### 4. KESIMPULAN

Berdasarkan indikator sosial ekonomi yang meliputi Tingkat Kemiskinan, IPM, TPT, dan PDRB, provinsi-provinsi di Indonesia berhasil dikelompokkan ke dalam empat klaster menggunakan metode K-Means. Hasil pengelompokan menunjukkan adanya homogenitas karakteristik dalam setiap klaster dan heterogenitas yang jelas antar klaster, yang mengindikasikan adanya ketimpangan sosial ekonomi antarwilayah. Hasil pengelompokan ini dapat menjadi dasar bagi pemerintah dalam merumuskan kebijakan pembangunan yang lebih efektif, terarah, dan sesuai dengan karakteristik masing-masing wilayah.

## 5. REFERENSI

- [1] M. Armini dan D. N. Azizah, *Dari Angka Menuju Makna*, Malang: Universitas Muhammadiyah Malang, 2025.
- [2] A. Suryahadi dan R. Al Izzati, *Kemiskinan Multidimensi di Indonesia*, Jakarta: SMERU Research Institute, 2021.
- [3] BPS, “Konsep dan Definisi Kemiskinan,” Badan Pusat Statistik, Jakarta, 2021.
- [4] M. Firdaus dan H. Setiawan, “Determinan Kemiskinan di Indonesia : Pendekatan Data Panel,” *Jurnal Ekonomi dan Studi Pembangunan*, vol. 23, 2022.
- [5] D. Leonita dan M. Sari, *Analisis PDRB dan Pertumbuhan Ekonomi Regional*, Yogyakarta: Andi Publisher, 2019.
- [6] R. Amalina dan D. Hartono, “Analisis Pengaruh Indeks Pembangunan Manusia dan Pertumbuhan Ekonomi terhadap Kemiskinan di Indonesia,” *Jurnal Ekonomi dan Pembangunan*, pp. 1-15, 2019.
- [7] Y. Febdianti, *Ekonomi Regional : Teori dan Aplikasi*, Bandung: Refika Aditama, 2017.
- [8] I. Suhendra dan A. Wicaksono, “Analisis Determinan Pengangguran di Indonesia,” *Jurnal Ekonomi dan Pembangunan*, 2020.
- [9] A. Saputra dan R. Hidayat, “Klasterisasi Provinsi di Indonesia Berdasarkan Indikator Sustainable Development Goals (SDGs) Menggunakan Algoritma K-Means,” *Jurnal Statistika dan Aplikasinya*, pp. 123-135, 2021.
- [10] S. Sukirno, *Makroekonomi : Teori Pengantar*, Jakarta: Rajawali Pers, 2018.
- [11] BPS, “Indeks Pembangunan Manusia,” Badan Pusat Statistik, Jakarta, 2023.
- [12] M. B. Dalimunthe dan B. E. Sitanggang, “The Relationship Between Economic Growth and Poverty In Indonesia,” *Journal of Social Science*, 2022.
- [13] M. S. Suliswanto, “Pengaruh Produk Domestik Bruto (PDRB) dan Indeks Pembangunan Manusia (IPM) terhadap Angka Kemiskinan di Indonesia,” *Jurnal Ekonomi Pembangunan*, 2020.
- [14] H. Y. Dama, “Pengaruh PDRB Terhadap Tingkat Kemiskinan di Kota Manado (Tahun 2005-2014),” *Jurnal Berkala Ilmiah Efisiensi*, 2016.
- [15] K. Ningsih dan A. W. Wijayanto, “Komparasi Metode Clustering Pada Provinsi di Indonesia Berdasarkan Pertumbuhan Ekonomi Tahun 2022,” *Kompitika : Jurnal Sistem Komputer*, 2023.
- [16] N. Etrisia dan M. Findi, “Klasifikasi Inflasi 34 Ibukota Provinsi di Indonesia Sebelum dan Saat Covid-19 Melalui Pengelompokan Wilayah dengan K-Means Clustering,” *Jurnal Ekonomi dan Studi Pembangunan*, 2023.
- [17] Sari dan Pratama, “Analisis Klaster Provinsi di Indonesia Berdasarkan Kesiapan

Teknologi Informasi dan Komunikasi,” *Jurnal Komputika*, 2022.

- [18] G. R. Wahyudi, D. E. Rahmadenni dan dkk, “Clustering Regencies in Indonesia for Regional Mapping Using The K-Means Algorithm,” *MALCOM*, vol. 5, 2025.
- [19] D. S. Rachmad, “Klasifikasi Pada Tempat Tinggal Menurut Provinsi dan Jenis Kepemilikan Berdasarkan Algoritma K-Means,” *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, 2020.
- [20] C. Prianto dan S. Bunyamin, Panduan Pembuatan Aplikasi Clustering Gangguan Jaringan Menggunakan Metode K-Means Clustering, Bandung: Kreatif Industri Nusantara, 2020.
- [21] L. Anggraeni dan R. R. Arum, “Aanalisis Cluster Menggunakan Algoritma K-Means Pada Provinsi Sumatera Barat Berdasarkan Indeks Pembangunan Manusia Tahun 2021,” *Prosiding Seminar Nasional UNIMUS*, pp. 639-640, 2022.