

**TIPOLOGI PEMBANGUNAN MANUSIA DI PROVINSI JAWA TIMUR : PENDEKATAN PRINCIPAL COMPONENT ANALYSIS DAN ANALISIS KLASTER**Kevin Brema Saputra Sinulingga^{1*}, Hana Titania Sastrian², Aryaputra Jagaddatri³, Trimono⁴
^{1,2,3,4,5}Program Studi Sains Data , Fakultas Ilmu Komputer, Universitas Pembangunan

Nasional Veteran Jawa Timur

E-mail: 23083010081@student.upnjatim.ac.id***ABSTRACT**

Human development is assessed using the Human Development Index (HDI), which includes indicators for health, education, and a decent standard of living. This research aimed to pinpoint the different human development typologies across the 37 Regencies/Cities in East Java Province during 2023. We analyzed the HDI data—specifically Life Expectancy at Birth, Expected Years of Schooling, Mean Years of Schooling, and Adjusted Per Capita Expenditure—by integrating Principal Component Analysis (PCA) and K-Means Clustering. PCA effectively condensed the four original indicators into two main components, which together accounted for over 90% of the data's variance. These components were termed "General Human Development Level" and "Relative Life Expectancy Focus." Following this, K-Means Clustering was used to categorize the regions into three distinct HDI groups: very high, medium, and low. This grouping clearly reveals the existing disparities in human development among East Java's regions, underscoring the need for flexible and targeted policy approaches. We anticipate these findings will serve as a valuable resource for stakeholders as they formulate more impactful development initiatives.

Keywords: *Human Development Index (HDI), Principal Component Analysis (PCA), K-Means Clustering, Development Typology, East Java*

ABSTRAK

Pembangunan manusia dievaluasi melalui Indeks Pembangunan Manusia (IPM), yang mencakup dimensi kesehatan, pendidikan, dan standar hidup layak. Penelitian ini bertujuan untuk mengidentifikasi tipologi pembangunan manusia di 38 kabupaten/kota di Provinsi Jawa Timur pada tahun 2023. Dengan menggabungkan Principal Component Analysis (PCA) dan K-Means Clustering, data IPM yang terdiri dari Angka Harapan Hidup, Harapan Lama Sekolah, Rata-rata Lama Sekolah, dan Pengeluaran per Kapita Disesuaikan, dianalisis secara komprehensif. Hasil dari PCA menunjukkan bahwa empat indikator tersebut berhasil direduksi menjadi dua komponen utama, yang secara kumulatif menjelaskan lebih dari 90% varians data. Komponen-komponen ini diidentifikasi sebagai "Tingkat Pembangunan Manusia Umum" dan "Fokus pada Harapan Hidup Relatif". Selanjutnya, K-Means

Article History

Received: Juni 2025

Reviewed: Juni 2025

Published: Juni 2025

Plagiarism Checker No
235

Prefix DOI :

[10.8734/Kohesi.v1i2.36](https://doi.org/10.8734/Kohesi.v1i2.365)[5](https://doi.org/10.8734/Kohesi.v1i2.365)

Copyright : Author

Publish by : Kohesi



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



Clustering digunakan untuk mengelompokkan wilayah-wilayah tersebut ke dalam tiga kluster IPM: sangat tinggi, sedang, dan rendah. Pengelompokan ini secara jelas menyoroti adanya disparitas dalam pembangunan manusia antar wilayah di Jawa Timur, menunjukkan perlunya strategi kebijakan yang adaptif dan spesifik. Temuan dari penelitian ini diharapkan dapat menjadi panduan penting bagi pemangku kepentingan dalam merumuskan program pembangunan yang lebih efektif dan terarah.

Kata Kunci: *Indeks Pembangunan Manusia (IPM), Principal Component Analysis (PCA), K-Means Clustering, Tipologi Pembangunan, Jawa Timur*

1. PENDAHULUAN

Pembangunan merupakan langkah fundamental dalam mewujudkan kemajuan dan peningkatan kualitas hidup masyarakat. Tujuan utamanya adalah menciptakan lingkungan yang memungkinkan individu menikmati umur panjang, kesehatan, dan menjalani kehidupan yang produktif (Sutanto & Rahayu, 2024). Untuk mengukur sejauh mana keberhasilan pembangunan dan kesejahteraan manusia, Indeks Pembangunan Manusia (IPM) telah diakui secara luas sebagai tolok ukur yang komprehensif (United Nations Development Programme (UNDP), 2024). IPM menjadi indikator penting dalam mengevaluasi kualitas pembangunan manusia di suatu wilayah. Kualitas hidup, seperti yang tercermin pada Provinsi Jawa Timur yang memiliki kategori IPM tinggi, sangat dipengaruhi oleh indikator ini (Badan Pusat Statistik Provinsi Jawa Timur, 2024). IPM tersusun atas tiga dimensi utama: kesehatan, pendidikan, dan standar hidup layak. Setiap dimensi ini diwakili oleh beberapa indikator kunci, yaitu Angka Harapan Hidup (AHH), Harapan Lama Sekolah (HLS), Rata-rata Lama Sekolah (RLS), dan Pengeluaran per Kapita (PPC) (Kurniawan, Hasan & Wulandari, 2022). Banyaknya variabel yang memengaruhi IPM menimbulkan tantangan tersendiri bagi pemerintah dalam merumuskan kebijakan pembangunan yang efektif (Syahrani, Kusumaningdyah & Dewa, 2021). Untuk menghadapi tantangan tersebut dan menyederhanakan struktur data yang kompleks, analisis reduksi dimensi seperti Principal Component Analysis (PCA) dapat digunakan (Fauzi & Supriyadi, 2020). PCA merupakan teknik statistik yang bertujuan mereduksi jumlah variabel tanpa kehilangan informasi penting, dengan cara mengidentifikasi faktor-faktor utama yang mewakili hubungan antar variabel (Manullang et al., 2024; Sitompul, Leonidas & Sanggala, 2023). Teknik ini sangat bermanfaat ketika terdapat korelasi antar variabel dalam jumlah besar, seperti pada indikator IPM. PCA dilakukan berdasarkan perhitungan nilai eigen dan vektor eigen untuk menentukan arah penyebaran data (Fauzi & Supriyadi, 2020; Harahap, 2022).



Penelitian ini menggunakan PCA untuk mengetahui faktor utama yang mempengaruhi IPM di Provinsi Jawa Timur pada tahun 2023. Melalui PCA, variabel-variabel yang kompleks dapat disederhanakan menjadi sejumlah faktor utama yang lebih mudah diinterpretasikan (Nasri et al., 2022; Nurhasan, 2021). Hal ini sangat relevan mengingat indikator IPM memiliki keterkaitan yang erat dan mencerminkan beberapa dimensi penting pembangunan manusia.

Selanjutnya, untuk memahami tipologi pembangunan manusia dan mengelompokkan kabupaten/kota berdasarkan profil IPM-nya, penelitian ini akan melanjutkan hasil PCA dengan Analisis Klaster. Analisis Klaster bertujuan untuk mengelompokkan data berdasarkan kemiripan karakteristik multivariat (Sakinah et al., 2024; Trihandini, Purnamasari & Wasono, 2022). Kombinasi antara PCA dan Analisis Klaster diharapkan dapat memberikan gambaran yang lebih menyeluruh terhadap pembangunan manusia di masing-masing daerah di Jawa Timur.

Penelitian sebelumnya oleh Ansori et al. (2023) juga telah menerapkan gabungan metode PCA dan algoritma K-Means dalam pengelompokan kabupaten/kota di Jawa Timur berdasarkan indikator IPM. Variabel yang digunakan meliputi AHH, HLS, RLS, pengeluaran per kapita, ketimpangan pendapatan, serta variabel sosial-ekonomi lainnya. Hasilnya menunjukkan bahwa terdapat tiga klaster wilayah dengan karakteristik pembangunan yang berbeda, serta adanya disparitas antar wilayah yang ditunjukkan oleh nilai silhouette sebesar 0,43N. (A. V. Ansori, A. Muhaimin, and A. Terza., 2023). Hal ini memperkuat urgensi penggunaan metode gabungan PCA dan klaster untuk memahami pola pembangunan manusia secara lebih mendalam.

Melalui penelitian ini, diharapkan dapat diidentifikasi faktor-faktor utama yang memengaruhi IPM di Jawa Timur tahun 2023, serta terbentuk klasifikasi kabupaten/kota berdasarkan tipologi pembangunannya. Hasil analisis ini diharapkan dapat menjadi panduan strategis bagi pemerintah daerah dan pemangku kepentingan dalam merumuskan kebijakan pembangunan yang lebih tepat sasaran dan berkelanjutan (Hariyadi & Dewi, 2023; Wellem, Silva & Lela, 2023).

2. METODE PENELITIAN

Penelitian ini mengadopsi pendekatan kuantitatif dengan desain deskriptif, bertujuan untuk mengidentifikasi tipologi pembangunan manusia di Kabupaten/Kota Provinsi Jawa Timur. Pendekatan ini dipilih untuk memberikan analisis yang mendalam dan sistematis terhadap struktur data multivariat serta mengelompokkan wilayah berdasarkan karakteristik yang homogen.

2.1 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder kuantitatif untuk tahun 2023, yang diperoleh dari publikasi resmi Badan Pusat Statistik (BPS) Provinsi Jawa Timur. Jumlah observasi adalah 38, merepresentasikan seluruh Kabupaten dan Kota di Jawa Timur. Variabel-variabel penelitian yang digunakan sebagai indikator pembangunan manusia, yang juga merupakan komponen utama Indeks Pembangunan Manusia (IPM), disajikan dalam Tabel 1.

Tabel 1. Variabel Indikator Pembangunan Manusia

Variabel	Keterangan	Satuan
X_1	Angka Harapan Hidup (AHH)	Tahun
X_2	Harapan Lama Sekolah (HLS)	Tahun
X_3	Rata-rata Lama Sekolah (RLS)	Tahun
X_4	Pengeluaran Per Kapita Disesuaikan (PPC)	Ribu Rupiah

2.2 Pra-pemrosesan dan Standarisasi Data

Tahap awal analisis melibatkan pengumpulan data dari berbagai sumber BPS Jawa Timur dan konversinya ke dalam format digital (CSV). Proses ini penting untuk memastikan konsistensi data serta menyingkirkan baris agregat provinsi (Jawa Timur) yang tidak diperlukan untuk analisis per kabupaten/kota. Selanjutnya, variabel-variabel indikator pembangunan manusia perlu distandarisi. Standarisasi ini sangat krusial sebelum melakukan Analisis Komponen Utama (PCA) dan Analisis Klaster, guna menghilangkan bias yang mungkin timbul akibat perbedaan skala pengukuran antar variabel (Puspita, 2021).

Standarisasi dilakukan menggunakan metode Z-score, yang mengubah setiap nilai data menjadi unit standar deviasi dari rata-ratanya, dengan rumus sebagai berikut (Sitorus, 2019):

$$Z = \frac{(x - \mu)}{\sigma}$$

Dimana:

Z = Nilai yang distandarisi

x = Nilai observasi asli

μ = Rata-rata variabel

σ = Standar deviasi variabel

2.3 Analisis Komponen Utama (Principal Component Analysis - PCA)

Analisis Komponen Utama (PCA) merupakan teknik statistika multivariat yang berfungsi untuk mereduksi dimensi dari kumpulan variabel yang saling berkorelasi. Proses ini menghasilkan sekumpulan kecil variabel baru yang tidak saling berkorelasi, yang dikenal sebagai komponen utama (Fauzi & Supriyadi, 2020). Komponen-komponen utama ini mampu menangkap



sebagian besar varians dari data asli, sehingga informasi penting tidak hilang (Manullang et al., 2024). Cara kerja PCA adalah dengan menghitung nilai eigen dan vektor eigen dari matriks kovarians atau korelasi data yang telah distandarisasi (Sitompul, Leonidas & Sanggala, 2023). Vektor eigen berperan dalam menentukan arah komponen, sementara nilai eigen menunjukkan seberapa besar varians yang mampu dijelaskan oleh komponen terkait (Nasri et al., 2022).

Setiap komponen utama (PC_i) merupakan kombinasi linear dari variabel-variabel asli (X_i), dirumuskan sebagai berikut:

$$PC_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p$$

Di mana:

PC_i = Komponen utama ke- i

a_{ij} = *Loading* atau koefisien untuk variabel X_j pada komponen PC_i

X_j = Variabel asli ke- j

Penentuan jumlah komponen utama yang optimal untuk dipertahankan didasarkan pada kriteria nilai eigen lebih dari satu (Kaiser's Criterion) dan observasi Scree Plot yang menunjukkan "siku" atau titik di mana proporsi varians yang dijelaskan mulai menurun drastis. Selain itu, proporsi varians kumulatif yang dijelaskan oleh komponen-komponen yang dipilih juga dipertimbangkan (target minimal 80%). Hasil PCA berupa Principal Component Scores akan digunakan sebagai input untuk analisis kluster.

2.4 Analisis Kluster (K-Means)

Analisis Kluster adalah metode multivariat yang bertujuan mengelompokkan objek, seperti kabupaten/kota, ke dalam kluster-kluster berdasarkan kesamaan karakteristiknya. Tujuannya adalah agar objek dalam satu kluster lebih homogen dibandingkan dengan objek di kluster lain (Ediyanto, Mara & Satyahadewi, 2021). Dalam penelitian ini, analisis kluster menggunakan algoritma K-Means, sebuah metode klustering non-hierarki yang paling sering dipakai. Algoritma K-Means bekerja secara iteratif untuk meminimalkan jumlah kuadrat jarak antara setiap titik data dan pusat kluster (centroid) terdekatnya, yang dirumuskan sebagai berikut:

$$Inertia = \sum_{i=1}^k \sum_{x \in C_i} ||x - \mu_i||^2$$

Dimana :

k = Jumlah kluster

C_i = Anggota kluster ke- i

μ_i = *Centroid* dari kluster ke- i

x = Data pada kluster C_i



Input untuk analisis kluster adalah Principal Component Scores yang diperoleh dari PCA, yang telah mereduksi dimensi dan menghilangkan multikolinearitas antar variabel asli. Penentuan jumlah kluster optimal (K) dalam K-Means dievaluasi menggunakan dua metode utama:

- *Elbow Method*

Metode Elbow adalah teknik yang digunakan untuk menentukan jumlah cluster yang optimal dengan menghitung nilai Sum of Square Error (SSE) dari masing-masing cluster. Semakin tinggi nilai selisih SSE dengan cluster berikutnya sehingga membentuk titik sudut siku, maka semakin baik jumlah cluster [24]. Memilih k berdasarkan perubahan signifikan (titik tekuk) pada grafik inerti.

$$SSE = \sum_{k=1}^k \sum_{X_i} |X_i - C_k|^2$$

- *Silhouette Score*

Mengevaluasi seberapa baik tiap data cocok dengan klusternya sendiri dibandingkan dengan kluster lain. Nilai *Silhouette* dihitung untuk setiap k dengan rumus:

$$S(i) = \frac{b(i) - a(i)}{\text{Max}(a(i), b(i))}$$

Dimana :

$a(i)$ = Rata-rata jarak data ke anggota kluster yang sama

$b(i)$ = Rata-rata jarak data ke kluster terdekat yang berbeda

3. HASIL DAN PEMBAHASAN

Data IPM tahun 2023 yang digunakan dalam penelitian ini berasal dari Badan Pusat Statistik (BPS) Provinsi Jawa Timur. Data tersebut meliputi empat indikator utama: Angka Harapan Hidup (AHH), Harapan Lama Sekolah (HLS), Rata-rata Lama Sekolah (RLS), dan Pengeluaran per Kapita Disesuaikan (PPC).

3.1 Hasil *Principal Component Analysis* (PCA)

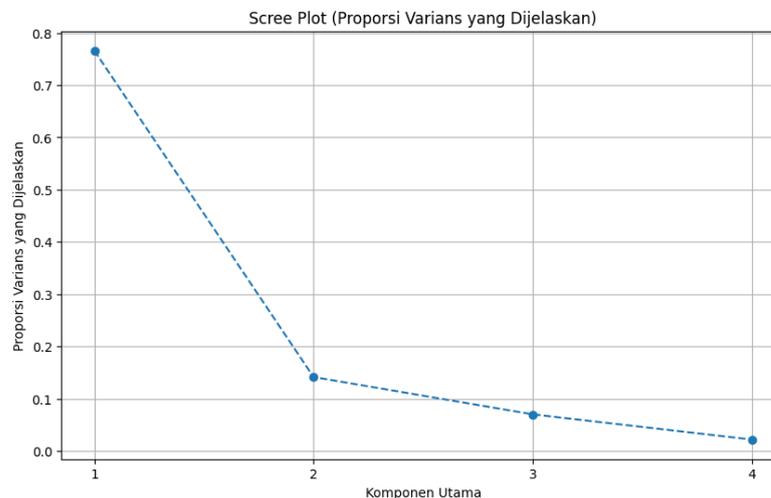
Hasil PCA menunjukkan bahwa dua komponen utama pertama (PC1 dan PC2) mampu menjelaskan sebagian besar varians dalam data. Secara spesifik, PC1 menjelaskan 76.60% varians, dan PC2 menjelaskan 14.16% varians. Secara kumulatif, kedua komponen ini menjelaskan 90.76% dari total varians data (Tabel 2). Angka ini menunjukkan bahwa reduksi dimensi dari empat variabel asli menjadi dua komponen utama sangat efektif dan mempertahankan informasi penting secara signifikan, mengingat target minimal 80% varians kumulatif yang dijelaskan.

Tabel 2. Explained Variance Ratio dan Kumulatif per Komponen Utama

Komponen Utama	<i>Explained Variance Ratio</i>	<i>Cumulative Explained Variance</i>
PC1	0.76596957	0.76596957
PC2	0.14164501	0.90761459
PC3	0.07027835	0.97789294
PC4	0.02210706	1.00000000

Penentuan jumlah komponen utama yang optimal juga didukung oleh *scree plot* (Gambar 1), yang menunjukkan "siku" yang jelas setelah komponen utama kedua. Hal ini konsisten dengan *Kaiser's Criterion*, di mana komponen dengan nilai eigen lebih dari satu atau titik di mana penurunan varians menjadi kurang signifikan. Oleh karena itu, PC1 dan PC2 dianggap cukup untuk mewakili struktur data IPM.

Gambar 1. Scree Plot (Proporsi Varians yang Dijelaskan)



Component loadings (Tabel 3) memberikan wawasan tentang kontribusi masing-masing variabel asli terhadap komponen utama.

Tabel 3. PCA Loadings Komponen Utama Terpilih

	AHH	HLS	RLS	PPC
PC1	0.433010	0.491037	0.550562	0.517945
PC2	0.854269	-0.424255	-0.011086	-0.300183

- *Principal Component 1 (PC1)*

Semua variabel (AHH, HLS, RLS, PPC) memiliki loading positif yang tinggi pada PC1. Variabel RLS (0.551) dan PPC (0.518) menunjukkan kontribusi yang paling dominan.

Dengan demikian, PC1 dapat diinterpretasikan sebagai "Indeks Kesejahteraan dan Pendidikan Agregat" atau "Tingkat Pembangunan Manusia Umum". Wilayah dengan nilai PC1 tinggi mengindikasikan capaian IPM yang secara keseluruhan baik pada dimensi kesehatan, pendidikan, dan standar hidup.

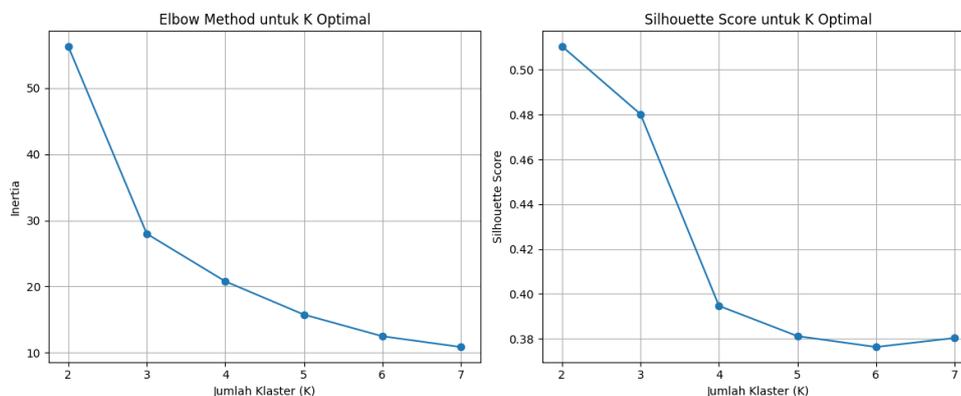
- *Principal Component 2 (PC2)*

Variabel AHH memiliki loading positif yang sangat tinggi (0.854), sementara HLS (-0.424) dan PPC (-0.300) memiliki loading negatif. RLS memiliki loading yang mendekati nol. Hal ini menunjukkan bahwa PC2 merepresentasikan "Fokus pada Harapan Hidup Relatif". Wilayah dengan nilai PC2 tinggi menonjol dalam aspek kesehatan (AHH), namun mungkin tidak sekuat dalam dimensi pendidikan (HLS) dan standar hidup (PPC) jika dibandingkan. Komponen ini menggambarkan adanya ketidakseimbangan atau prioritas relatif pada dimensi IPM tertentu di suatu wilayah.

3.2 Analisis Kluster (K-Means)

Principal Component Scores dari PC1 dan PC2 digunakan sebagai input untuk algoritma klustering K-Means. Untuk menentukan jumlah kluster optimal (K), digunakan Elbow Method dan Silhouette Score, dengan pengujian K dari 2 hingga 7.

Gambar 2. Elbow Method dan Silhouette Score untuk K Optimal



Dari *Elbow Method* (Gambar 2, kiri), titik "siku" yang menandakan penurunan inerti yang signifikan terlihat pada $K=3$. Sementara itu, *Silhouette Score* (Gambar 2, kanan) juga menunjukkan bahwa $K=3$ menghasilkan skor kohesi dan pemisahan kluster yang optimal. Berdasarkan kedua kriteria tersebut, jumlah kluster optimal (K) yang dipilih adalah 3. Pemilihan $K=3$ juga konsisten dengan temuan penelitian sebelumnya oleh Ansori et al. yang juga mengidentifikasi tiga kluster pembangunan di Jawa Timur.

Setelah $K = 3$ ditetapkan, algoritma K-Means dijalankan, dan label kluster ditambahkan ke dataframe gabungan. Distribusi Kabupaten/Kota di setiap kluster adalah sebagai berikut:



- Klaster 0: 9 Kabupaten/Kota
- Klaster 1: 10 Kabupaten/Kota
- Klaster 2: 18 Kabupaten/Kota

Mayoritas Kabupaten/Kota di Jawa Timur termasuk dalam Klaster 2, yang menunjukkan adanya kelompok pembangunan IPM yang cenderung homogen pada level sedang.

3.3 Karakteristik Klaster

Untuk memahami profil unik dari setiap klaster, rata-rata nilai variabel IPM asli (AHH, HLS, RLS, PPC) dihitung untuk setiap klaster (Tabel 4).

Tabel 4. Karakteristik Klaster (Rata-rata Variabel Asli)

Klaster	AHH (Tahun)	HLS (Tahun)	RLS (Tahun)	PPC (Ribu Rupiah)
0	74.11	14.74	10.74	15371.56
1	69.77	12.93	6.48	10297.70
2	73.02	13.28	8.28	11994.78

- Klaster 0 (Tingkat Pembangunan Sangat Tinggi)

Klaster ini memiliki nilai rata-rata tertinggi di semua indikator IPM. Harapan hidup sangat tinggi (74.11 tahun), harapan dan rata-rata lama sekolah berada di level sangat baik (14.74 dan 10.74 tahun), serta pengeluaran per kapita disesuaikan paling tinggi (Rp15.371.56). Klaster ini merepresentasikan wilayah-wilayah dengan capaian pembangunan manusia yang paling unggul di Jawa Timur, umumnya mencakup kota-kota besar dan daerah urban maju.

- Klaster 1 (Tingkat Pembangunan Rendah)

Klaster ini dicirikan oleh nilai rata-rata terendah pada semua indikator IPM. AHH (69.77 tahun) lebih rendah, demikian pula HLS (12.93 tahun), RLS (6.48 tahun), dan PPC (Rp10.297.700). Kabupaten/kota dalam klaster ini menghadapi tantangan signifikan dalam pembangunan manusia dan memerlukan perhatian khusus serta intervensi kebijakan yang komprehensif di semua dimensi IPM.

- Klaster 2 (Tingkat Pembangunan Sedang)

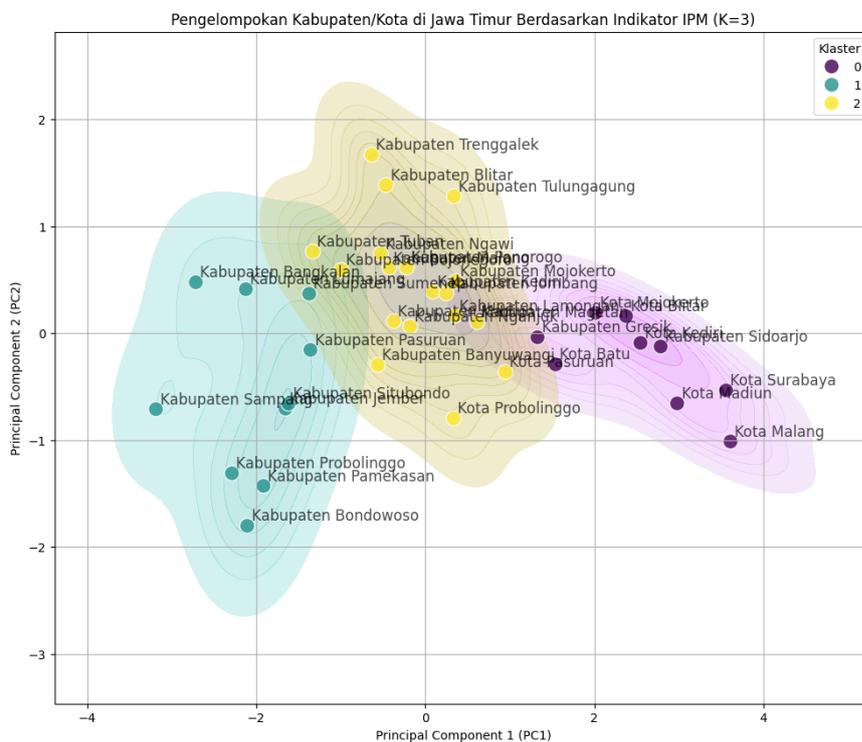
Klaster ini menunjukkan nilai rata-rata indikator IPM yang berada di antara Klaster 0 dan Klaster 1. AHH (73.02 tahun) cukup baik, HLS (13.28 tahun) dan RLS (8.28 tahun) pada level menengah, dan PPC (Rp11.994.78) juga sedang. Klaster ini merupakan

kelompok terbesar dan menunjukkan tingkat pembangunan manusia yang moderat. Wilayah-wilayah ini memiliki potensi untuk ditingkatkan, dan kebijakan dapat difokuskan pada penguatan area-area spesifik IPM untuk mencapai kemajuan lebih lanjut.

3.4 Visualisasi Klaster

Visualisasi klaster pada ruang Principal Component (PC1 dan PC2) ditampilkan pada Gambar 3. Plot ini menunjukkan sebaran Kabupaten/Kota yang diwarnai berdasarkan klaster yang terbentuk, dilengkapi dengan area densitas klaster menggunakan KDE plot untuk mempermudah identifikasi pengelompokan.

Gambar 3. Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator IPM



- **Klaster 0 (Warna Ungu)**

Terletak di bagian kanan atas plot, menunjukkan nilai PC1 dan PC2 yang relatif tinggi. Klaster ini mencakup kota-kota besar atau wilayah dengan karakteristik IPM sangat tinggi, yang konsisten dengan interpretasi Klaster 0 sebagai "Tingkat Pembangunan Sangat Tinggi" karena dominasi nilai tinggi pada semua indikator IPM.

- **Klaster 1 (Warna Hijau)**

Terletak di bagian kiri bawah plot, menunjukkan nilai PC1 dan PC2 yang relatif rendah. Klaster ini mewakili wilayah-wilayah dengan tingkat pembangunan IPM yang masih memerlukan peningkatan signifikan, sesuai dengan interpretasi Klaster 1 sebagai "Tingkat Pembangunan Rendah".



- Klaster 2 (Warna Kuning)

Terletak di bagian tengah plot, dengan rentang nilai PC1 dan PC2 yang lebih bervariasi. Klaster ini merupakan kelompok terbesar dan mencakup sebagian besar kabupaten/kota di Jawa Timur. Wilayah-wilayah ini memiliki tingkat pembangunan IPM sedang, yang mengindikasikan potensi untuk ditingkatkan melalui intervensi kebijakan yang tepat.

Visualisasi ini secara efektif mengkonfirmasi bahwa PCA berhasil mereduksi kompleksitas empat variabel IPM menjadi dua dimensi utama yang dapat menjelaskan perbedaan antar wilayah. Selanjutnya, analisis klaster berhasil mengidentifikasi tipologi pembangunan yang berbeda, memberikan gambaran yang jelas mengenai disparitas pembangunan manusia di Provinsi Jawa Timur.

KESIMPULAN

Penelitian ini mengidentifikasi struktur dimensi dan mengelompokkan Kabupaten/Kota di Jawa Timur berdasarkan Indeks Pembangunan Manusia (IPM) tahun 2023. Melalui Principal Component Analysis (PCA), empat indikator IPM berhasil direduksi menjadi dua komponen utama yang signifikan, menjelaskan lebih dari 90% varians data. PC1 merepresentasikan "Tingkat Pembangunan Manusia Umum", sedangkan PC2 menunjukkan "Fokus pada Harapan Hidup Relatif".

Selanjutnya, Analisis Klaster K-Means mengelompokkan wilayah menjadi tiga kategori IPM: sangat tinggi, sedang, dan rendah. Hasil ini menyoroti adanya disparitas pembangunan manusia di Jawa Timur, yang memerlukan kebijakan adaptif dan spesifik wilayah. Untuk pengembangan riset ke depan, disarankan untuk menyertakan variabel sosial-ekonomi lain seperti kemiskinan atau ketimpangan pendapatan, serta melakukan analisis longitudinal untuk memantau dinamika pembangunan.

DAFTAR PUSTAKA

- H. Sutanto, & Rahayu, I. (2024). Kajian disparitas pembangunan antar kabupaten/kota di Jawa Timur menggunakan indikator IPM. *Jurnal Kebijakan Pembangunan Daerah*, 14(3), 210-225.
- United Nations Development Programme (UNDP). (2024). *Human development report 2023/2024: Breaking the gridlock*. UNDP.



- Badan Pusat Statistik Provinsi Jawa Timur. (2024). Provinsi Jawa Timur dalam angka 2024. BPS Provinsi Jawa Timur.
- Kurniawan, L., Hasan, M., & Wulandari, S. (2022). Penerapan Principal Component Analysis untuk identifikasi faktor-faktor pembangunan daerah. *Prosiding Konferensi Nasional Statistika*, 112-120.
- Syahrani, E., Kusumaningdyah, A. A., & Dewa, D. D. (2021). Analisis faktor-faktor yang mempengaruhi tingkat kemiskinan kabupaten/kota di Jawa Tengah. *Geodika: Jurnal Kajian Ilmu dan Pendidikan Geografi*, 5(2), 247-258.
- Fauzi, A., & Supriyadi, R. (2020). Deteksi penyakit kanker payudara dengan seleksi fitur berbasis Principal Component Analysis dan Random Forest. *Jurnal Infortech*, 96. <https://ejournal.bsi.ac.id/ejurnal/index.php/infortech/article/view/8079/4111>
- Manullang, S., et al. (2024). Analisis faktor penyebab penyakit jantung menggunakan metode Principal Component Analysis (PCA). *Lebesgue: Jurnal Pendidikan Matematika, Matematika dan Statistika*, 5(3), 1568-1588.
- Sitompul, A. Y., Leonidas, D., & Sanggala, E. (2023). Analisis faktor pengendalian kualitas teh hitam pada PT. ABC dengan menggunakan metode Principal Component Analysis (PCA). *Jurnal Cahaya Mandalika*, 1161-1172.
- Harahap, D. A. (2022). Analisis faktor-faktor yang mempengaruhi Indeks Pembangunan Manusia (IPM) di Indonesia. *dspace UII*. <https://dspace.uii.ac.id/handle/123456789/40496>
- Nurhasan, P. D. (2021). Efektivitas media pembelajaran IPS berbasis Macromedia Flash untuk meningkatkan motivasi dan hasil belajar siswa SD. *Jurnal Civics & Social Studies*, 80-89.
- Nasri, E., et al. (2022). Faktor-faktor yang mempengaruhi rendahnya prestasi belajar ujian matematika siswa dengan menggunakan metode analisis faktor. *Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, 3(1), 12-28.
- Sakinah, B., et al. (2024). Penerapan analisis faktor dalam menentukan faktor-faktor yang mempengaruhi keputusan belanja online melalui aplikasi Shopee (Studi kasus: Mahasiswa Universitas Sumatera Utara). *Algoritma: Jurnal Matematika, Ilmu Pengetahuan Alam, Kebumihan dan Angkasa*, 2(6), 60-75.
- Trihandini, T., Purnamasari, I., & Wasono. (2022). Diagram kontrol multivariat berdasarkan jarak Chi-Kuadrat (Studi kasus: Produksi surat kabar Kaltim Post tahun 2017). *Jurnal Matematika dan Statistika serta Aplikasinya*, 10(1), 68-79.
- Ansori, N. A. V., Muhaimin, A., & Terza, A. (2023). Pembentukan kluster terhadap Indeks Pembangunan Manusia di wilayah Jawa Timur. *Jurnal Lebesgue*, 2(2), 97-106. <https://lebesgue.lppmbinabangsa.id/index.php/home/article/view/589>



- Wellem, I., Silva, Y. O., & Lela, A. M. (2023). Analisis faktor-faktor yang mempengaruhi disiplin kerja pegawai pada kantor Badan Perencanaan dan Litbang Kabupaten Sikka. *Projemen UNIPA*, 1-17.
- Hariyadi, K., & Dewi, I. I. (2023). Faktor dominan yang mempengaruhi partisipasi siswa dalam mengikuti kegiatan ekstrakurikuler olahraga. *SPRINTER: Jurnal Ilmu Olahraga*, 4(3), 423-428.
- Puspita, R. N. (2021). Analisis K-Means cluster pada kabupaten/kota di Provinsi Banten berdasarkan indikator Indeks Pembangunan Manusia. *Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, 2(3). <https://doi.org/10.46306/lb.v2i3>
- Sitorus, N. (2019). Hubungan status gizi dengan perkembangan anak usia 2 sampai 3 tahun di wilayah kerja Puskesmas Satria Kota Tebing Tinggi tahun 2018.
- Ediyanto, Mara, M. N., & Satyahadewi, N. (2021). Pengklasifikasian karakteristik dengan metode k-means cluster analysis. *Buletin Ilmiah Matematika, Statistika dan Terapannya (Bimaster)*, 2(2), 133-136.
- Sholeh, M., & Aeni, K. (2023). Perbandingan evaluasi metode Davies Bouldin, Elbow dan Silhouette pada model clustering dengan menggunakan algoritma K-Means. *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, 8(1), 56. <https://doi.org/10.30998/string.v8i1.16388>