



This is an open access article under  
CC-BY-SA license

\* Corresponding Author

## ANALISIS KLASTER KESEJAHTERAAN MASYARAKAT KABUPATEN/KOTA DI JAWA TIMUR MENGGUNAKAN METODE HIERARKI DAN K-MEANS

Iloh Silvia Arfindi<sup>1</sup> Farida Rahmawati<sup>2</sup>  
Universitas Negeri Malang, Indonesia

Email: iloh.silvia.2104326@um.ac.id

**Abstract.** This study aims to classify districts/cities in East Java based on social welfare indicators in 2023 using hierarchical and K-Means clustering methods. The dataset consists of 38 districts/cities obtained from the Central Bureau of Statistics (BPS). The variables include poverty rate, unemployment rate, Human Development Index (HDI), per capita expenditure, sanitation access, and social assistance coverage. Data were standardized using z-score before clustering analysis. The optimal number of clusters was determined using Silhouette Coefficient, Davies-Bouldin Index (DBI), and Calinski-Harabasz Index (CHI). The results show that K-Means clustering with three clusters provides the most optimal grouping. Although hierarchical clustering (Ward method) produces a slightly higher Silhouette value, K-Means demonstrates better compactness and cluster separation based on DBI and CHI.

**Keyword:** Hierarchical Clustering, K-Means Clustering, Social Welfare.

**Abstrak.** Penelitian ini bertujuan untuk mengklasifikasikan kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator kesejahteraan sosial tahun 2023 dengan menggunakan metode kluster hierarki dan K-Means. Data yang digunakan terdiri dari 38 kabupaten/kota yang diperoleh dari Badan Pusat Statistik (BPS). Variabel yang dianalisis meliputi tingkat kemiskinan, tingkat pengangguran, Indeks Pembangunan Manusia (IPM), pengeluaran per kapita, akses sanitasi, dan cakupan bantuan sosial. Data distandarisasi menggunakan z-score sebelum dilakukan analisis kluster. Penentuan jumlah kluster optimal dilakukan menggunakan Silhouette Coefficient, Davies-Bouldin Index (DBI), dan Calinski-Harabasz Index (CHI). Hasil penelitian menunjukkan bahwa metode K-Means dengan tiga kluster menghasilkan pengelompokan yang paling optimal. Meskipun metode kluster hierarki (Ward) memiliki nilai Silhouette yang sedikit lebih tinggi, metode K-Means menunjukkan tingkat kompaksi kluster dan pemisahan antar kluster yang lebih baik berdasarkan nilai DBI dan CHI.

**Kata kunci:** Kluster Hierarki, Kluster K-Means, Kesejahteraan Sosial.



## PENDAHULUAN

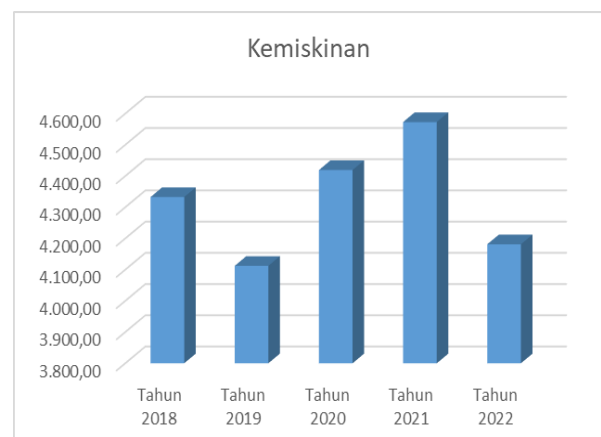
Negara di dunia seringkali menghadapi masalah yang sama yaitu berkaitan dengan kemiskinan, dimana kemiskinan sendiri secara umum didefinisikan dengan sebuah keadaan atau situasi atas ketidakmampuan seorang individu dalam memenuhi seluruh kebutuhan dasar seperti pangan, sandang dan lain sebagainya. Di Indonesia, kemiskinan menjadi isu yang sulit diatasi dan terus mengalami peningkatan di setiap tahunnya. Berbagai bantuan dan program telah diterapkan oleh pemerintah dengan tujuan mengurangi tingkat kemiskinan. Tetapi, sayangnya upaya tersebut masih belum cukup untuk menangani masalah kemiskinan yang ada di Indonesia. Di sini, peran besar pemerintah dalam mensejahterahkan masyarakat merupakan salah satu fokus utama agar dapat menjadi sebuah negara maju (Alfiah et al., 2021).

Kesejahteraan sosial dan ekonomi adalah aspek krusial yang mencerminkan kualitas hidup masyarakat di suatu daerah. Secara umum, kesejahteraan bisa diartikan sebagai ukuran yang menunjukkan bahwa suatu masyarakat telah mencapai kondisi yang lebih baik dibandingkan sebelumnya. Beragam faktor juga menunjukkan pentingnya pendekatan yang sistematis untuk memahami pola keterkaitan antara indikator-indikator kesejahteraan sosial dan ekonomi. Pendekatan ini bisa membantu pemerintah daerah merancang kebijakan yang lebih tepat sasaran (Ekonomi et al., 2024). Untuk mencapai kesejahteraan bagi masyarakat perlu beberapa hal yang dapat dilakukan, diantaranya meningkatkan kemampuan serta pemerataan akses kebutuhan dasar, memberikan perlindungan sosial, perbaikan kualitas hidup seperti memperbanyak lapangan kerja agar dapat meningkatkan pendapatan dan dengan pendapatan masyarakat yang tinggi mereka akan dapat memfasilitasi anak-anak mereka pada pendidikan yang lebih baik. Upaya terakhir yang dapat dilakukan adalah dengan adanya perluasan skala ekonomi yang diiringi dengan penambahan pilihan sosial bagi individu dan masyarakat secara keseluruhan. Dengan demikian, dapat menciptakan kehidupan yang lebih bermakna dan adil bagi setiap anggota masyarakat (Ni Putu Ambar pratiwi dan I Bagus Indrajaya, 2019).

Analisis kluster adalah kumpulan teknik pengolahan data berdasarkan golongan yang

dimiliki. Terdapat dua jenis analisis kluster, yaitu metode hirarki (yang terdiri dari agglomerative dan divisive) serta non-hirarki. Dalam penelitian ini, kami menerapkan analisis kluster hirarki serta non-hirarki untuk mengidentifikasi metode analisis kluster serta menentukan jumlah kluster yang paling optimal (Rahmawati et al., 2016). Dalam konteks ini, pengukuran dan analisis indikator kesejahteraan sosial sangat penting, terutama di daerah-daerah seperti Jawa Timur. Provinsi ini dikenal dengan keragaman sosial dan ekonomi yang signifikan serta memiliki populasi yang besar. Dalam beberapa tahun terakhir, tingkat kesejahteraan masyarakat di Jawa Timur menunjukkan tren peningkatan yang positif. Berdasar pada data BPS (Badan Pusat Statistik), rata-rata pendapatan masyarakat mengalami kenaikan dari Rp 55,41 juta pada tahun 2018 menjadi Rp 66,38 juta pada tahun 2022, yang berarti meningkat sebesar 10,52%. Meskipun demikian, ketimpangan dalam distribusi kesejahteraan masih menjadi tantangan penting, di mana beberapa kabupaten/kota menunjukkan kemajuan pesat sementara yang lain tertinggal. Pada Provinsi Jawa Timur mengalami peningkatan nilai ekonomi yang cukup drastis berbanding terbalik dengan kondisi provinsi lainnya, meskipun tantangan seperti pengangguran, kemiskinan, dan pemenuhan kebutuhan dasar masyarakat masih ada. Beberapa indikator pembangunan menunjukkan perbaikan yang jelas, contohnya, tingkat kemiskinan berkurang dari 12,28 persen satu dekade lalu menjadi 9,79 persen saat ini. Angka pengangguran juga mengalami penurunan, yang semula dengan angka 4,02 persen menjadi 3,74 persen dalam periode yang sama. menurut data BPS.

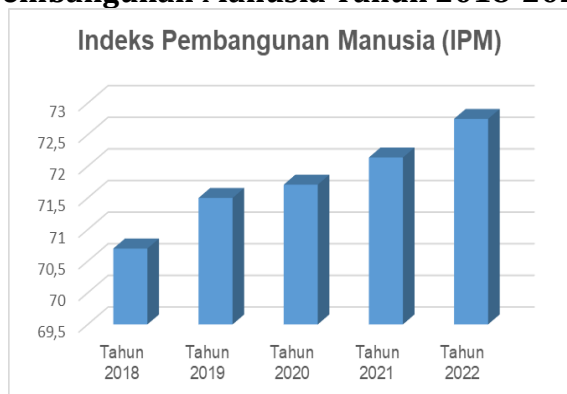
**Gambar 1.1 Perkembangan Jumlah Penduduk Miskin Tahun 2018-2022**



Sumber: Data BPS 2023, diolah

Pencapaian ini berdampak positif pada Indeks Pembangunan Manusia (IPM), yang kini mencapai 73,38 persen. Hasil ini menunjukkan performa lebih optimal daripada provinsi lain di Pulau Jawa, meskipun Jawa Timur masih sedikit tertinggal di belakang Jawa Barat. Selain itu, peningkatan IPM didukung oleh penurunan tingkat kemiskinan yang signifikan, dari 12,28 persen sejak awal dekade menjadi 9,79 persen saat ini. Hal tersebut memperlihatkan jika upaya pemerintah di dalam mengoptimalkan kesejahteraan masyarakat telah mulai menghasilkan dampak yang nyata, mendorong percepatan pembangunan manusia di Jawa Timur.

**Gambar 1.2 Perkembangan Indeks Pembangunan Manusia Tahun 2018-2022**



Sumber: Data BPS 2023, diolah

Walaupun persentase pengangguran telah menurun, jumlah pengangguran di Jawa Timur tetap cukup signifikan. Berdasarkan data dari BPS, terdapat sekitar 1,02 juta orang yang masih menganggur di provinsi ini. Angka tersebut turun dari 1,3 juta pada tahun 2020, meskipun target di bawah satu juta yang tercapai sebelum pandemi masih belum dapat diraih.

**Gambar 1.3 Perkembangan Indeks Pembangunan Manusia Tahun 2018-2022**



Sumber: Data BPS 2023, diolah

Wilayah dengan perekonomian yang relatif kuat justru menghadapi tingkat pengangguran yang tinggi. Contohnya, Kabupaten Sidoarjo mencatat angka pengangguran tertinggi di provinsi, yaitu 6,49 persen, yang melebihi rata-rata nasional maupun provinsi. Meskipun memiliki perputaran ekonomi yang baik dan menempati posisi kedua setelah Kota Surabaya, fenomena serupa juga terjadi di Kabupaten Gresik. Sebagai lokasi Kawasan Ekonomi Khusus (KEK) JIPE dan tempat beroperasinya smelter terbesar milik Freeport, Gresik memiliki potensi ekonomi besar, namun tetap mengalami angka pengangguran yang tinggi. Kondisi ini juga terlihat di wilayah industri lainnya, seperti Kabupaten Pasuruan, Kota Surabaya, Kabupaten Malang, dan Kabupaten Tuban, di mana skala ekonomi yang besar tidak serta-merta mengurangi tingkat pengangguran. Oleh sebab itu, dibutuhkan analisis mendalam mengenai beberapa faktor yang mempengaruhi kesejahteraan di berbagai daerah. Penelitian ini memiliki tujuan yakni guna mengidentifikasi serta mengelompokkan kabupaten/kota menurut indikator kesejahteraan, termasuk tingkat pendidikan, akses layanan kesehatan, angka kemiskinan, dan pendapatan per kapita. Metode cluster hierarki digunakan untuk memberikan gambaran hubungan antar daerah, sedangkan metode K-Means memungkinkan pengelompokan yang lebih fokus dan efisien. Dengan menerapkan kedua metode ini, diharapkan dapat diperoleh pemahaman lebih dalam mengenai karakteristik kesejahteraan sosial di setiap wilayah. Hasil penelitian ini tidak hanya memberikan wawasan bagi para pembuat kebijakan dalam menyusun program pembangunan yang lebih tepat sasaran, tetapi juga berkontribusi pada peningkatan kualitas hidup masyarakat secara keseluruhan. Analisis ini diharapkan dapat menghasilkan rekomendasi strategis untuk mengurangi ketimpangan dan meningkatkan kesejahteraan di seluruh wilayah Jawa Timur.

Meskipun berbagai penelitian telah menggunakan metode clustering untuk menganalisis kesejahteraan sosial, sebagian besar penelitian hanya menggunakan satu pendekatan tanpa melakukan perbandingan kinerja antar metode. Hal ini menyebabkan keterbatasan dalam menentukan metode yang

paling optimal dalam mengelompokkan wilayah berdasarkan indikator kesejahteraan. Oleh karena itu, penelitian ini bertujuan untuk mengisi kesenjangan tersebut dengan membandingkan metode hierarki dan K-Means serta mengevaluasi performanya menggunakan beberapa metrik validasi kluster.

## METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan teknik analisis kluster untuk mengelompokkan kabupaten/kota di Jawa Timur berdasarkan indikator kesejahteraan sosial. Data yang digunakan merupakan data sekunder yang bersumber dari Badan Pusat Statistik (BPS) dengan jumlah observasi sebanyak 38 kabupaten/kota.

Variabel yang digunakan meliputi: (1) persentase penduduk miskin, (2) tingkat pengangguran terbuka (TPT), (3) indeks pembangunan manusia (IPM), (4) rata-rata pengeluaran per kapita, (5) akses sanitasi layak, dan (6) penerima bantuan sosial (KKS).

Tahapan analisis dimulai dengan standarisasi data menggunakan z-score untuk menghindari bias skala. Selanjutnya dilakukan analisis menggunakan metode kluster hierarki (Ward linkage) dan metode non-hierarki K-Means. Penentuan jumlah kluster optimal dilakukan melalui evaluasi menggunakan Silhouette Coefficient, Davies-Bouldin Index (DBI), dan Calinski-Harabasz Index (CHI). Seluruh proses analisis dilakukan menggunakan perangkat lunak SPSS dan RStudio.

## HASIL DAN DISKUSI

Menurut (Durrofiq & Akbarita, 2022) dalam analisis data berperan penting dalam menstandarisasi variabel untuk menghindari potensi masalah akibat perbedaan skala nilai antar variabel dalam proses pengelompokan objek. Salah satu metode standarisasi yang umum digunakan adalah konversi setiap variabel ke dalam nilai standar atau z-score, di mana proses ini melibatkan perhitungan nilai rerata dan membaginya dengan standar deviasi dari setiap variabel. Standarisasi data menjadi langkah krusial dalam penelitian, terutama ketika data yang digunakan memiliki satuan yang beragam. Dalam studi ini, perbedaan satuan antar variabel membuat pengelompokan tidak bisa dilaksanakan secara langsung.

Sehingga, dibutuhkan proses standarisasi sebelum analisis lebih lanjut dapat dilakukan.

Hasil standarisasi dalam analisis kluster disajikan dalam bentuk z-score, tanpa disertai visualisasi plot, karena analisis ini hanya berfokus pada variabel independen tanpa adanya variabel terikat (dependen).

**Tabel 3.1 Standarisasi Data Variabel**

	n	Minimum	Maximum	Mean	Std. Deviation
X1	38	7.10	251.36	110.2326	68.31945
X2	38	1.71	8.05	4.6629	1.42883
X3	38	64.13	83.45	73.6803	5.05376
X4	38	456234	1052321	643854.53	135429.013
X5	38	7.13	34.16	19.4082	6.95233
X6	38	50.30	98.18	84.5871	12.18312
Valid	38				

*Sumber: SPSS, data diolah*

### Jarak Square Euclidean Distance

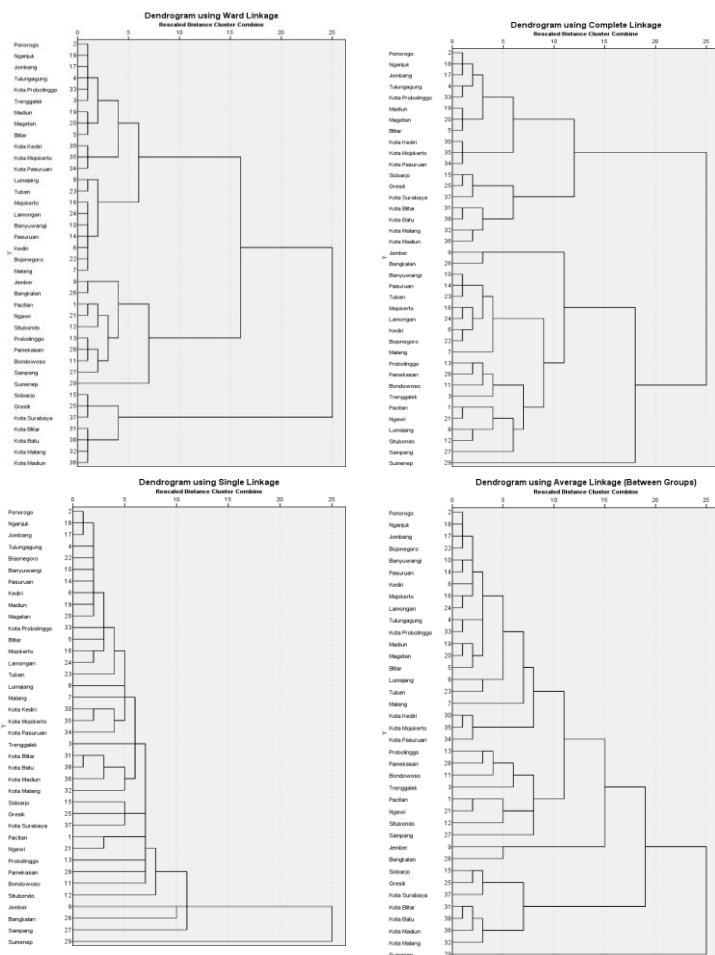
Square Euclidean Distance, atau jarak Euklides kuadrat, merupakan variasi dari jarak Euklides yang digunakan untuk mengukur kemiripan atau kedekatan antara dua titik dalam ruang berdimensi. Jarak ini dihitung dengan mengkuadratkan selisih antara koordinat titik-titik tersebut tanpa mengambil akar kuadrat, sehingga proses perhitungan menjadi lebih sederhana dan sering digunakan dalam algoritme pengelompokan data. Dengan metode ini, klustering dapat dilakukan secara lebih efektif, khususnya dalam konteks Hybrid Mutual Clustering, yang mengkombinasikan antara pendekatan Bottom-up dengan top-down guna mencapai hasil pengelompokan yang optimal.

### Proses Clustering Hierarki

Analisis kluster merupakan teknik statistik yang dipergunakan untuk mengklasifikasikan observasi kedalam kategori berdasarkan kesamaan karakteristik. Terdapat dua pendekatan utama di dalam analisis kluster yakni metode hierarki serta metode non-hierarki. Pemilihan metode yang sesuai bergantung pada peneliti dan konteks penelitian, dengan mempertimbangkan aspek substansi, teori dan konsep yang relevan. Setiap metode memiliki keunggulannya masing-masing. Di dalam studi ini, metode yang dipergunakan yaitu kluster hierarki, karena jumlah kluster tidak ditentukan di awal, melainkan terbentuk berdasarkan hasil analisis. Hasil dari metode kluster hierarki dapat

divisualisasikan dalam bentuk dendrogram, yang berfungsi sebagai alat bantu dalam proses pengelompokan objek dengan memberikan representasi visual yang lebih jelas. Salah satu kelebihan metode hierarki adalah efisiensinya dalam pengolahan data, sehingga dapat mempercepat proses analisis dalam penelitian (Wahyuni & Wulandari, 2022). Hierarchical clustering adalah teknik pengelompokan yang menyajikan hasilnya dalam bentuk struktur bertingkat, dimulai dari n objek dan akhirnya menyatu menjadi satu kluster tunggal (Karimah et al., 2025).

**Gambar 3.1 Dendrogram Hasil Kluster Hierarki Ward Linkage, Complete Linkage, Single Linkage dan Average Linkage**



Sumber: SPSS, data diolah

Hasil analisis kluster hierarki menunjukkan struktur pengelompokan yang bervariasi berdasarkan metode linkage yang digunakan. Metode Ward linkage menghasilkan solusi optimal dengan empat kluster, yang mengindikasikan adanya empat kelompok utama dalam data berdasarkan varians minimal antar kluster. Sementara itu, metode complete linkage mengidentifikasi tiga kluster yang berbeda,

menunjukkan kecenderungan untuk menggabungkan observasi berdasarkan jarak terjauh antar anggota kluster. Di sisi lain, metode single linkage menghasilkan empat kluster, mencerminkan sensitivitas terhadap keberadaan "rantai" observasi yang berdekatan. Metode average linkage juga menghasilkan empat kluster, yang menunjukkan keseimbangan antara jarak antar observasi dalam kluster dan jarak antar kluster secara keseluruhan. Perbedaan jumlah kluster yang dihasilkan oleh berbagai metode ini menggarisbawahi pentingnya mempertimbangkan karakteristik data dan tujuan analisis dalam memilih metode linkage yang paling sesuai. Meskipun metode hierarki menghasilkan empat kluster, hasil evaluasi menunjukkan bahwa penggunaan tiga kluster pada metode K-Means lebih optimal dan lebih mudah diinterpretasikan dalam konteks kebijakan.

**Agglomeratif Coefficient Metode Hierarki**

Analisis kluster adalah teknik statistik yang dipergunakan untuk mengelompokkan data dengan banyak variabel dengan tujuan agar objek-objek yang memiliki kesamaan dapat dikelompokkan jadi satu (Dewi & Ahadiyah, 2022). Di dalam studi ini, dipergunakan metode hierarki dengan 4 aglomeratif yaitu complete linkage, average linkage, wards method serta single linkage. Berikutnya akan ditentukan metode yang dinilai paling baik dan sesuai dengan meninjau nilai *agglomerative coefficient*.

**Tabel 3.2 Agglomerative Coefficient Metode Hierarki**

Metode	Nilai
Single Linkage	0,4397173
Complete Linkage	0,8158294
Average Linkage	0,6650151
Ward's Methods	0,8822733

Sumber: Rstudio, data diolah

Nilai Agglomerative Coefficient merupakan ukuran yang memperlihatkan seberapa baik metode hierarki di dalam mengelompokkan data (Kara, 2021). Metode yang memiliki nilai koefisien mendekati 1 maka pengelompokkan aglomeratifnya akan semakin kuat. Oleh sebab itu, tabel diatas memperlihatkan jika metode yang mendekati angka 1 yaitu ward's methods, sehingga metode tersebut memiliki aglomeratif yang kuat.

## K-Means Cluster

K-Means Clustering ialah non-hierarki yang bertujuan guna membagi data kedalam satu atau beberapa kelompok berdasarkan kesamaan tertentu (Sibarani, 2018). Setiap objek akan ditempatkan dalam kelompok tertentu sesuai dengan kedekatannya terhadap titik pusat (centroid) dari kelompok tersebut (Saputra & Nataliani, 2021).

**Tabel 3. Number of Cases in Each Cluster**

Number of Cases in each Cluster		
Cluster	1	9.000
	2	4.000
	3	25.000
Valid		38.000
Missing		0.000

*Sumber: SPSS, data diolah*

Berdasarkan hasil analisis menggunakan metode K-Means, data berhasil dikelompokkan menjadi tiga kluster dengan karakteristik yang berbeda. Kluster 1 terdiri dari 9 kasus, kluster 2 terdiri dari 4 kasus, dan kluster 3 mencakup 25 kasus serta merupakan kluster dengan jumlah anggota terbesar. Proses kluster menunjukkan konvergensi yang baik dalam tiga iterasi, di mana posisi pusat kluster stabil dan tidak mengalami perubahan yang signifikan. Pusat-pusat kluster menunjukkan perbedaan yang jelas antar kluster, terutama pada variabel Z-score (X1), (X2), (X3), (X4), (X5), dan (X6), yang berperan dalam membedakan karakteristik masing-masing kluster.

### Pembentukan Kluster Optimum

Mengevaluasi kualitas kluster yang dihasilkan, penelitian ini menggunakan tiga metrik validasi internal yang umum digunakan, yaitu Silhouette Coefficient, Calinski-Harabasz Index (CHI) dan Davies-Bouldin Index (DBI). Silhouette Coefficient memberikan ukuran seberapa mirip suatu objek dengan klasternya sendiri dibandingkan dengan kluster lain, dengan nilai yang lebih tinggi menunjukkan pemisahan kluster yang lebih baik. DBI mengukur rasio dispersi within-cluster terhadap separasi between-cluster, dimana nilai yang rendah mengindikasikan adanya kluster yang kompak dan optimal. Sementara itu, CHI mengukur rasio varians between-cluster

terhadap varians within-cluster, dengan nilai yang lebih tinggi menunjukkan kluster yang lebih terdefinisi dengan baik. Ketiga metrik ini digunakan secara komprehensif untuk memberikan validasi yang kuat terhadap struktur kluster yang ditemukan, memastikan bahwa hasil pengklasteran yang diperoleh memiliki kualitas yang optimal dan dapat diandalkan.

**Tabel 3.3 Pembentukan Cluster Optimum**

Metode	Silhouette	DBI	CHI
Ward Method	0,2352054	0,2704878	17.82681
K-Means Cluster	0,2082102	0,2224533	16,81130

*Sumber: Rstudio, data diolah*

Berdasarkan hasil evaluasi menggunakan Silhouette Coefficient, Calinski-Harabasz Index (CHI), dan Davies-Bouldin Index (DBI), metode Ward memiliki nilai Silhouette Coefficient yang sedikit lebih tinggi dibandingkan K-Means, yang menunjukkan pemisahan kluster yang lebih baik secara individual. Namun, metode K-Means menunjukkan nilai DBI yang lebih rendah, yang mengindikasikan bahwa kluster yang dihasilkan lebih kompak. Selain itu, nilai CHI yang relatif tinggi pada metode K-Means menunjukkan adanya pemisahan antar kluster yang cukup jelas. Oleh karena itu, pemilihan metode K-Means sebagai metode optimal didasarkan pada pertimbangan kombinasi beberapa metrik evaluasi yang memberikan keseimbangan antara kompaksi dan separasi kluster.

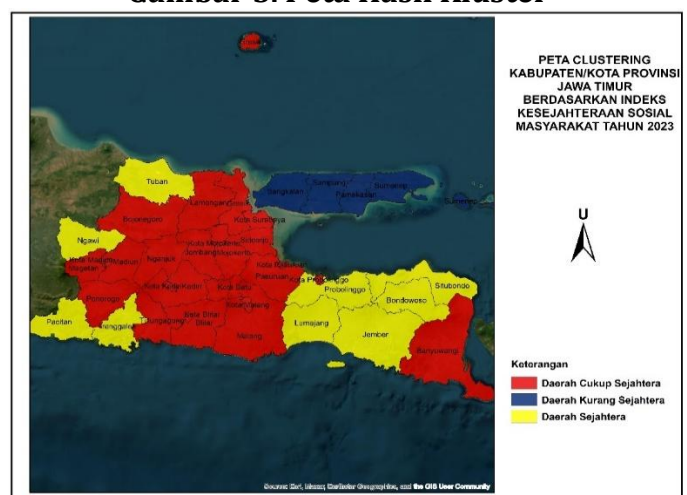
### Karakteristik dan Rekomendasi Kebijakan Pemerintah

Setelah pemodelan menggunakan algoritma kluster hierarki dengan metode Ward linkage yang menghasilkan tiga kluster pada dataset tahun 2023, hasil penelitian ini akan divisualisasikan dalam bentuk peta menggunakan perangkat lunak Quantum GIS (QGIS). Dengan demikian, informasi mengenai pengelompokan daerah yang terpengaruh oleh variabel indeks kesejahteraan sosial masyarakat di Provinsi Jawa Timur dapat disajikan secara jelas dan mudah dipahami. Hasil pengelompokan ini dapat dijadikan dasar dalam penilaian karakteristik kabupaten/kota di Jawa Timur, serta sebagai pembanding antar daerah berdasarkan kategori tingkat kesejahteraan, yaitu rendah, sedang, dan tinggi.

<b>Cluster 1 (Cluster Daerah Kesejahteraan Rendah)</b>			
Anggota Cluster	Kabupaten Pacitan Kabupaten Trenggalek Kabupaten Lumajang Kabupaten Jember	Kabupaten Bondowoso Kabupaten Situbondo	Kabupaten Probolinggo Kabupaten Ngawi Kabupaten Tuban
Karakteristik	Jumlah Penduduk Miskin yang tergolong cukup tinggi Tingkat Pengangguran Terbuka (TPT) yang rendah IPM yang tergolong rendah Rata-rata pengeluaran perkapita makanan sebulan yang tergolong rendah Persentase rumah tangga penerima Kartu Keluarga Sejahtera (KKS) yang cukup tinggi Persentase Rumah Tangga yang mempunyai akses terhadap layanan sanitasi layak yang cukup rendah		
<b>Cluster 2 (Cluster Daerah Kesejahteraan Sedang)</b>			
Anggota Cluster	Kabupaten Bangkalan Kabupaten Sampang	Kabupaten Pamekasan	Kabupaten Sumenep
Karakteristik	1. Jumlah Penduduk Miskin yang tergolong tinggi 2. Tingkat Pengangguran Terbuka (TPT) yang tinggi 3. IPM yang tergolong cukup rendah 4. Rata-rata pengeluaran perkapita makanan sebulan yang tergolong cukup tinggi 5. Persentase rumah tangga penerima Kartu Keluarga Sejahtera (KKS) yang tinggi 6. Persentase Rumah Tangga yang mempunyai akses terhadap layanan sanitasi layak yang rendah.		
<b>Cluster 3 (Cluster Daerah Kesejahteraan Tinggi)</b>			
Anggota Cluster	Kabupaten Ponorogo Kabupaten Tulungagung Kabupaten Blitar Kabupaten Kediri Kabupaten Malang Kabupaten Banyuwangi Kabupaten Pasuruan Kabupaten Sidoarjo Kabupaten Mojokerto Kabupaten Jombang Kabupaten Nganjuk	Kabupaten Madiun Kabupaten Magetan Kabupaten Bojonegoro Kabupaten Lamongan Kabupaten Gresik Kota Kediri Kota Blitar Kota Malang	Kota Probolinggo Kota Pasuruan Kota Mojokerto Kota Madiun Kota Surabaya Kota Batu
Karakteristik	1. Jumlah Penduduk Miskin yang tergolong rendah 2. Tingkat Pengangguran Terbuka (TPT) yang cukup tinggi 3. IPM yang tergolong tinggi 4. Rata-rata pengeluaran perkapita makanan sebulan yang tergolong tinggi 5. Persentase rumah tangga penerima Kartu Keluarga Sejahtera (KKS) yang rendah 6. Persentase rumah tangga yang mempunyai akses terhadap layanan sanitasi layak yang tinggi.		

Hasil ini menunjukkan bahwa daerah dengan tingkat kesejahteraan tinggi umumnya memiliki IPM yang tinggi, tingkat kemiskinan rendah, dan akses layanan dasar yang lebih baik. Sebaliknya, daerah dengan kesejahteraan rendah cenderung memiliki keterbatasan akses terhadap layanan dasar serta tingkat kemiskinan yang lebih tinggi. Hal ini mengindikasikan bahwa kesejahteraan merupakan hasil interaksi kompleks antara faktor ekonomi, sosial, dan infrastruktur. Temuan ini juga menunjukkan bahwa pendekatan kebijakan yang bersifat seragam kurang efektif, sehingga diperlukan intervensi berbasis karakteristik wilayah.

**Gambar 5. Peta Hasil Kluster**



Sumber: QGIS, data diolah

## KESIMPULAN

Penelitian ini menunjukkan bahwa metode K-Means lebih efektif dibandingkan metode hierarki dalam mengelompokkan kesejahteraan sosial di Jawa Timur berdasarkan enam indikator utama. Hasil clustering menghasilkan tiga kelompok wilayah dengan karakteristik kesejahteraan yang berbeda, yaitu rendah, sedang, dan tinggi. Temuan ini menegaskan bahwa pendekatan clustering dapat menjadi alat yang efektif dalam mendukung perumusan kebijakan berbasis data. Penelitian ini berkontribusi dalam memberikan pendekatan komparatif antar metode clustering serta memberikan dasar bagi perencanaan pembangunan yang lebih terarah dan berbasis karakteristik wilayah. Penelitian ini juga menunjukkan bahwa pemilihan metode clustering yang tepat sangat berpengaruh terhadap hasil pengelompokan dan interpretasi kebijakan yang dihasilkan.

## DAFTAR PUSTAKA

- Afira, N., & Wijayanto, A. W. (2021). Analisis cluster dengan metode partitioning dan hierarki pada data informasi kemiskinan provinsi di Indonesia tahun 2019. *Komputika: Jurnal Sistem Komputer*, 10(2), 101–109. <https://doi.org/10.34010/komputika.v10i2.4317>
- Alfiah, F., Almadayani, A., Al Farizi, D., & Widodo, E. (2021). Analisis clustering K-medoids berdasarkan indikator kemiskinan di Jawa Timur tahun 2020. *Jurnal Ilmiah Sains*, 22(1), 1–10. <https://doi.org/10.35799/jis.v22i1.35911>
- Dewi, A. F., & Ahadiyah, K. (2022). Agglomerative hierarchy clustering pada penentuan kelompok kabupaten/kota di Jawa Timur berdasarkan indikator pendidikan. *Zeta – Math Journal*, 7(2), 57–63. <https://doi.org/10.31102/zeta.2022.7.2.57-63>
- Durrofiq, I., & Akbarita, R. (2022). Algoritma K-means dalam mengelompokkan kota dan kabupaten di Jawa Timur berdasarkan indikator kesejahteraan rakyat. *Jurnal Ilmiah Matematika*, 9(2), 49–54.
- Ekonomi, J., Ebismen, M., Rohmatulillah, O. N., Nirmala, K. B., & Wulandari, S. P. (2024). Analisis kluster pada faktor-faktor yang mempengaruhi indikator kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur. *Jurnal Ekonomi*, 3(4), 307–325.
- Indarti, S. H. (2017). Pembangunan Indonesia dalam pandangan Amartya Sen. *The Indonesian Journal of Public Administration (IJPA)*, 3(1), 35–50. <https://doi.org/10.52447/ijpa.v3i1.727>
- Kara, F. (2021). Comparison of tree diameter distributions in managed and unmanaged Kazdağı fir forests. *Silva Balcanica*, 22(1), 31–43. <https://doi.org/10.3897/silvabalcanica.22.e58020>
- Karimah, A., Nur, I., Fitriana, L., & Universitas Terbuka. (2025). Pengelompokan provinsi di Indonesia berdasarkan indeks pembangunan literasi masyarakat 2023 menggunakan metode kluster hierarki dan non-hierarki. *Jurnal Statistika*, 2(1), 394–407.
- Pratiwi, N. P. A., & Indrajaya, I. G. B. (2019). Pengaruh pertumbuhan ekonomi dan pengeluaran terhadap kesejahteraan masyarakat. *Buletin Studi Ekonomi*, 24(2), 220–233.
- Pangke, I. N. P., Kumenaung, A. G., & Tumilaar, R. L. H. (2021). Efektivitas pemanfaatan dana desa terhadap kesejahteraan masyarakat di Kabupaten Sangihe (studi kasus Kecamatan Tamako). *Jurnal EMBA*, 9(2), 297–305.
- Puspitasari, W. I. (2017). Faktor-faktor yang mempengaruhi migrasi tenaga kerja ke luar negeri berdasarkan provinsi di Indonesia. *Jurnal Ilmu Ekonomi Terapan*, 2(1), 44–55. <https://doi.org/10.20473/jiet.v2i1.5505>
- Rahmawati, L., Sihwi, S. W., & Suryani, E. (2016). Analisa clustering menggunakan metode K-means dan hierarchical clustering. *Jurnal Teknologi & Informasi ITS smart*, 3(2), 66–73. <https://doi.org/10.20961/its.v3i2.654>
- Roza, D., & Putra, G. T. P. S. (2019). Perundang-undangan untuk mewujudkan Indonesia.

*Jurnal Cendekia Hukum*, 5(1), 12–20.  
<https://doi.org/10.3376/jch.v5i1.185>

Saputra, E. A., & Nataliani, Y. (2021). Analisis pengelompokan data nilai siswa untuk menentukan siswa berprestasi menggunakan metode clustering K-means. *Journal of Information Systems and Informatics*, 3(3), 424–439.  
<https://doi.org/10.51519/journalisi.v3i3.164>

Sibarani, R. (2018). Algoritma K-means clustering strategi pemasaran penerimaan mahasiswa baru Universitas Satya Negara Indonesia. *Jurnal ALU*, 1(2), 44–50.

Sriningsih, M., Hatidja, D., & Prang, J. D. (2018). Penanganan multikolinearitas dengan menggunakan analisis regresi komponen utama. *Jurnal Ilmiah Sains*, 18(1), 18–25.  
<https://doi.org/10.35799/jis.18.1.2018.19396>

Sultan, S., Rahayu, H. C., & Purwiyanta. (2023). Analisis pengaruh kesejahteraan masyarakat terhadap pertumbuhan ekonomi di Indonesia. *Jurnal Informatika Ekonomi Bisnis*, 5, 75–83.  
<https://doi.org/10.37034/infeb.v5i1.198>

Talakua, M. W., Leleury, Z. A., & Talluta, A. W. (2017). Analisis cluster menggunakan metode K-means pada indikator pembangunan manusia. *Jurnal Ilmu Matematika dan Terapan*, 11(2), 119–128.

Usna, W., & Aprilia, R. (2024). Comparison of agglomerative hierarchical clustering and K-means algorithm in poverty data clustering. *Data Journal*, 7(3), 489–500.  
<https://doi.org/10.24042/djm>

Wahyuni, I., & Wulandari, S. P. (2022). Pemetaan kabupaten/kota di Jawa Timur. *Jurnal Sains dan Seni*, 11(1), D70–D75.