



Analisis kluster kabupaten/kota di Jawa Tengah berdasarkan indikator kemiskinan dan ketimpangan menggunakan K-Harmonic Means

Nabbila Dyah Syaharani*, Universitas Muhammadiyah Semarang

Gita Rahmawati, Universitas Muhammadiyah Semarang

Selvi Ana Windia Sari, Universitas Muhammadiyah Semarang

M Al Haris, Universitas Muhammadiyah Semarang

*e-mail: nabbilads@gmail.com

Abstrak. Ketimpangan sosial dan ekonomi antarwilayah masih menjadi tantangan besar di Provinsi Jawa Tengah, terutama karena tingginya tingkat kemiskinan yang berdampak pada kualitas hidup masyarakat. Untuk mendukung perencanaan pembangunan yang lebih tepat sasaran dan berbasis data, diperlukan pemetaan kondisi sosial ekonomi antarwilayah. Penelitian ini bertujuan untuk mengelompokkan 35 kabupaten/kota di Jawa Tengah berdasarkan indikator kemiskinan dan ketimpangan dengan menggunakan metode *K-Harmonic Means*. *K-Harmonic Means* merupakan metode pengembangan dari *K-Means* yang lebih stabil terhadap inisialisasi pusat kluster dan menghasilkan pengelompokan yang lebih optimal. Tahapan analisis mencakup normalisasi data, penentuan jumlah kluster optimal menggunakan metode *Elbow*, dan proses klusterisasi dengan algoritma *K-Harmonic Means*. Hasil penelitian menunjukkan bahwa wilayah di Jawa Tengah dapat dikelompokkan ke dalam dua kluster utama. Kluster 1 terdiri dari 26 wilayah dengan karakteristik kemiskinan dan ketimpangan yang lebih tinggi, sedangkan Kluster 2 terdiri dari 9 wilayah dengan kondisi sosial ekonomi yang relatif lebih baik. Nilai *Silhouette Coefficient* sebesar 0.28 menunjukkan struktur kluster yang cukup baik. Hasil ini dapat memberikan gambaran yang jelas mengenai kondisi sosial ekonomi wilayah sebagai dasar penyusunan kebijakan pembangunan daerah.

Kata kunci: kemiskinan, ketimpangan, K-harmonic means, klusterisasi, jawa tengah.

Abstract. Social and economic disparities between regions remain a major challenge in Central Java Province, especially due to the high poverty rate that affects the population's quality of life. To support more targeted and data-driven development planning, it is necessary to map the socioeconomic conditions of each region. This study aims to classify 35 regencies/cities in Central Java using poverty and inequality indicators through the *K-Harmonic Means* method. *K-Harmonic Means* is an improved version of the *K-Means* method that offers greater stability in cluster center initialization and produces more optimal clustering results. The analysis involved data normalization, determining the optimal number of clusters using the *Elbow* method, and performing clustering with the *K-Harmonic Means* algorithm. The results show that the regions in Central Java can be grouped into two main clusters. Cluster 1 consists of 26 regions with higher poverty and inequality levels, while Cluster 2 includes 9 regions with relatively better socioeconomic conditions. The *Silhouette Coefficient* value of 0.28 indicates a fairly well-formed clustering structure. These findings provide a clearer overview of the

regional socioeconomic landscape and can serve as a foundation for formulating regional development policies.

Keywords: *poverty, inequality, K-harmonic means, clustering, central java.*

PENDAHULUAN

Kemiskinan dapat diartikan sebagai kondisi ketidakmampuan seseorang dalam memenuhi kebutuhan dasar untuk hidup. Fenomena ini menjadi tantangan serius bagi setiap negara, karena semakin tinggi angka kemiskinan, maka semakin rendah pula tingkat kesejahteraan masyarakat di negara tersebut. Sementara itu, ketimpangan merupakan masalah yang lazim terjadi di suatu daerah akibat pendapatan yang tidak merata antarindividu (Pratiwi, 2021). Berdasarkan Survei Sosial Ekonomi (Susenas) yang dilakukan oleh Badan Pusat Statistik (BPS), terdapat beberapa indikator kemiskinan yang digunakan untuk mengukur tingkat kesejahteraan masyarakat, yaitu persentase penduduk miskin, garis kemiskinan, jumlah penduduk miskin, indeks kedalaman kemiskinan, indeks keparahan kemiskinan, tingkat pengangguran terbuka, ipm, dan rasio gini (Aprillyana, 2020).

Tingginya angka kemiskinan di Indonesia memicu *instabilitas* ekonomi yang menyebabkan masyarakat kurang mampu terpuruk dibawah garis kemiskinan (Alyarahma et al., 2024). Hal ini menunjukkan bahwa keberhasilan pembangunan suatu negara sangat bergantung pada kemampuannya dalam menekan angka kemiskinan. Oleh karena itu, penanggulangan kemiskinan perlu dilakukan menyeluruh, mencakup berbagai aspek kehidupan, dan dilaksanakan secara merata di seluruh wilayah (Syahrani et al., n.d., 2021). Salah satu daerah yang mengalami tantangan tersebut adalah Provinsi Jawa Tengah, dimana kondisi sosial ekonominya menunjukkan dinamika yang kompleks antar kabupaten/kota (Ayumna & Sukardi, 2024). Untuk memahami variasi kondisi sosial ekonomi antar kabupaten/kota di provinsi tersebut, dibutuhkan pendekatan analisis yang mampu mengelompokkan wilayah berdasarkan karakteristik statistik kemiskinannya. Analisis kluster atau metode *clustering* digunakan untuk mengidentifikasi pola-pola dalam data, sehingga hasil pengelompokan dapat memberikan gambaran yang lebih jelas dan terstruktur mengenai kondisi kemiskinan di setiap daerah (Mayasari & Nugraha, 2023).

Analisis kluster merupakan teknik pengelompokan yang bertujuan untuk mengelompokkan objek-objek berdasarkan kesamaan karakteristik yang dimilikinya. Salah satu pendekatan yang umum digunakan dalam analisis kluster adalah metode *non-hierarki partitioning*. Dalam metode ini, data dibagi ke dalam sejumlah kluster tertentu yang telah ditentukan oleh peneliti. Salah satu metode *partitioning* yang paling populer adalah algoritma *K-Harmonic Means* (Salsabila et al., 2022).

Metode *K-Harmonic Means* merupakan metode pengembangan dari metode *K-Means*. Perbedaan utama antara kedua metode tersebut terletak pada penggunaan titik pusat/*centeroid* kluster. Pada metode *K-Harmonic Means*, penentuan posisi setiap titik data didasarkan pada fungsi keanggotaan. Pendekatan ini membuat metode *K-Harmonic Means* menjadi tidak sensitif terhadap inisialisasi awal titik pusat kluster serta mampu meningkatkan kualitas klusterisasi secara signifikan dibandingkan dengan metode *K-Means* (Kualitas et al., 2024). Validasi terhadap hasil kluster perlu dilakukan untuk mengukur kualitas dan kekuatan dalam kluster yang terbentuk. *Silhouette coefficient* merupakan salah satu metode validasi yang banyak digunakan. Metode ini memvalidasi kluster dengan menggabungkan konsep kohesi dalam kluster dan separasi antar kluster (Yunistya et al., 2022).

Penelitian terdahulu (Putri et al., 2023) menggunakan metode pendekatan *K-Harmonic Means* untuk mengelompokkan tingkat kemiskinan berdasarkan kabupaten/kota di Indonesia. Hasilnya, parameter $p = 2,25$ menghasilkan dua kluster optimal dengan nilai *dunn index*

sebesar 0,0349. Klaster pertama terdiri atas 54 kabupaten/kota dengan karakteristik tingkat kemiskinan tinggi, sedangkan klaster kedua terdiri atas 460 kabupaten/kota dengan tingkat kemiskinan rendah. Penelitian oleh (Yunistya et al., 2022) menerapkan metode *K-Harmonic Means* untuk mengelompokkan kabupaten/kota di Pulau Kalimantan berdasarkan indikator kemiskinan. Hasil penelitian menunjukkan pengelompokan paling optimal diperoleh pada 2 klaster dengan $p = 4$, menghasilkan *silhouette coefficient* sebesar 0,323. Klaster pertama terdiri dari 11 kabupaten/kota dan klaster kedua terdiri dari 45 kabupaten/kota. Selanjutnya, penelitian oleh (Solihat, 2024) menggunakan metode *K-Harmonic Means* untuk mengelompokkan kabupaten/kota di Jawa Barat berdasarkan indikator kemiskinan. Hasil penelitian menunjukkan bahwa pengelompokan paling optimal diperoleh pada 2 klaster dengan nilai *silhouette coefficient* sebesar 0,909. Klaster pertama mencakup 19 wilayah dengan tingkat kemiskinan lebih tinggi, sedangkan klaster kedua terdiri dari 8 wilayah dengan kondisi yang lebih baik.

Dengan demikian, penelitian ini akan menerapkan metode *K-Harmonic Means* untuk mengelompokkan kabupaten/kota di Provinsi Jawa Tengah berdasarkan indikator kemiskinan. Penelitian ini bertujuan untuk mengidentifikasi wilayah-wilayah dengan karakteristik kemiskinan yang serupa, sehingga dapat memberikan gambaran yang lebih terarah mengenai daerah yang memerlukan perhatian lebih lanjut. Hasil pengelompokan ini diharapkan dapat menjadi bahan pertimbangan dalam penyusunan strategi penanggulangan kemiskinan serta menjadi acuan bagi peneliti-peneliti selanjutnya.

METODE

Penelitian ini menggunakan data mengenai indikator kemiskinan dan ketimpangan di Provinsi Jawa Tengah untuk tahun 2024 yang diperoleh dari Badan Pusat Statistika (BPS). Data sekunder yang digunakan mencakup seluruh kabupaten dan kota di provinsi tersebut. Variabel-variabel yang dianalisis tercantum pada Tabel 1.

Tabel 1. Variabel penelitian

Variabel	Nama Variabel
X_1	Garis Kemiskinan
X_2	Jumlah Penduduk Miskin
X_3	Persentase Penduduk Miskin
X_4	Indeks Kedalaman Kemiskinan
X_5	Indeks Keparahan Kemiskinan
X_6	Gini Rasio
X_7	Tingkat Pengangguran Terbuka (TPT)
X_8	Indeks Pembangunan Manusia (IPM)

Dalam Penelitian ini, menggunakan pendekatan analisis klaster dengan menerapkan mengelompokkan objek-objek ke dalam beberapa klaster yang dimiliki. Objek objek yang berada dalam satu klaster memiliki tingkat kesamaan yang tinggi, sedangkan objek-objek yang tergabung dalam klaster berbeda menunjukkan perbedaan karakteristik yang lebih signifikan (Salsabila et al., 2022). Salah satu metode yang digunakan dalam analisis klaster adalah *K-Harmonic Means*, yang termasuk dalam kategori *center-based clustering* yang merupakan pengembangan dari *K-Means*, dimana kedua posisi titik pusat (*centroid*) dari setiap klaster hingga mencapai hasil pengelompokan yang optimal (Aya Dewanti Sofia & Kudus, 2023). Tidak seperti *K-Means* yang menggunakan rata-rata aritmetika, metode ini menggunakan rata-rata harmonik yang lebih sensitif terhadap nilai data kecil dan lebih tahan terhadap *outlier*. Rata-rata harmonik suatu klaster ke- g dinyatakan dengan:

$$HA(G_g) = \frac{t}{\sum_{a=1}^t \frac{1}{y_{ag}}} \quad (1)$$

dengan:

- $HA(G_g)$: rata-rata harmonik klaster ke- g
- t : jumlah data dalam klaster ke- g
- y_{ag} : nilai data ke- a dalam klaster ke- g

Dengan pendekatan harmonik ini, pengaruh nilai besar dalam perhitungan pusat klaster dapat diminimalkan karena nilai kecil akan lebih mendominasi hasil rata-rata. Fungsi objektif metode *K-Harmonic Means* disusun dengan mempertimbangkan kontribusi seluruh data terhadap semua pusat klaster, dan dirumuskan sebagai:

$$KHM = \sum_{q=1}^T \left(\frac{K}{\sum_{p=1}^K \frac{1}{||z_q - v_p||^r}} \right) \quad (2)$$

dengan:

- KHM : nilai fungsi objektif *K-Harmonic Means*
- T : jumlah total objek data
- K : jumlah klaster yang diinginkan
- z_q : objek data ke- q
- v_p : pusat klaster ke- p
- $||z_q - v_p||$: jarak antara data ke- q dan pusat klaster ke- p
- r : parameter pangkat (biasanya $r = 2$ untuk jarak *euclidean* kuadrat).

Penelitian ini terdiri dari beberapa tahapan utama yang meliputi pengumpulan data, pra-pemrosesan data, pelaksanaan algoritma *K-Harmonic Means*, serta analisis hasil klasterisasi. Berikut merupakan tahapan-tahapan yang dilakukan dalam penelitian ini (Yunistya et al., 2022):

- Standarisasi data dilakukan dengan tujuan menyeragamkan skala pengukuran, sehingga nilai-nilai data menjadi standar dan tidak lagi dipengaruhi oleh perbedaan satuan antar variabel (Widyadhana et al., 2021)
- Pengujian asumsi multikolinieritas dilakukan dengan menggunakan nilai *Variance Inflation Factor* (VIF). Apabila ditemukan indikasi multikolinieritas pada salah satu variabel independen, maka salah satu langkah penanganan yang dapat dilakukan adalah dengan mengeliminasi variabel tersebut dari model. Sebaliknya, jika tidak terdapat multikolinieritas, maka analisis dapat dilanjutkan ke tahap berikutnya (Salsabila et al., 2022).
- Penentuan banyaknya klaster yang akan digunakan (k)
- Penentuan pusat klaster (*centroid*) awal secara *random*
- Menghitung jarak setiap data terhadap *centroid* dari masing-masing klaster dengan menggunakan rumus *euclidean* berikut (Raja et al., 2024):

$$||z_q - v_p|| = \sqrt{\sum_{j=1}^n (z_{qj} - v_{pj})^2} \quad (3)$$

- Menghitung fungsi *objective* untuk mendapatkan hasil pengelompokan yang optimal dengan rumus pada persamaan (2).
- Menghitung proporsi keanggotaan untuk tiap objek data dengan rumus berikut (Yunistya et al., 2022):

$$m(v_p|z_q) = \frac{||z_q - v_p||^{-r-2}}{\sum_{p=1}^k ||z_q - v_p||^{-r-2}} \quad (4)$$

h. Penentuan nilai bobot

Nilai bobot ditentukan dengan memanfaatkan jarak data menggunakan metode *euclidean* yang telah dihitung sebelumnya. Di bawah ini disajikan perhitungan bobot untuk objek pertama:

$$w(z_q) = \frac{\sum_{p=1}^k ||z_q - v_p||^{-r-2}}{(\sum_{p=1}^k ||z_q - v_p||^{-r})^2} \quad (5)$$

i. Pemutakhiran posisi pusat kluster

Proses pembaruan titik pusat kluster dilakukan dengan mempertimbangkan nilai keanggotaan serta bobot yang dimiliki oleh masing-masing dari objek data. Berikut ini merupakan perhitungan untuk pusat kluster 1:

$$v_p = \frac{\sum_{q=1}^b a(v_1 - z_q) \cdot n(z_q) \cdot z_q}{\sum_{q=1}^b a(v_p - z_q) \cdot n(z_q)} \quad (6)$$

j. Validasi kluster dengan menggunakan metode *Silhouette Coefficient*. Perhitungan nilai *Silhouette Coefficient* ditunjukkan pada persamaan berikut:

$$s(i) = \frac{b(zi) - a(zi)}{\max(a(zi), b(zi))}, \text{ jika } |C_i| > 1 \quad (7)$$

Persamaan 7 menunjukkan bahwa nilai *mean* dari setiap data dalam suatu kluster merepresentasikan tingkat kekompakan atau kedekatan objek-objek dalam kluster tersebut, yang menjadi ukuran seberapa kuat suatu objek tergolong dalam kluster tertentu (Rahmah et al., 2022).

HASIL DAN PEMBAHASAN

Analisis Deskriptif

Analisis deskriptif dilakukan untuk memberikan gambaran umum mengenai karakteristik masing-masing variabel digunakan untuk memahami sebaran serta rentang data pada tiap indikator kemiskinan dan ketimpangan. Analisis Deskriptif indikator kemiskinan dan ketimpangan disajikan pada Tabel 2.

Tabel 2. Analisis deskriptif indikator kemiskinan dan ketimpangan

Variabel	X1	X2	X3	X4	X5	X6	X7	X8
Mean	507425,60	105,84	10,13	1,51	0,35	0,35	4,52	74,75
Std.Deviasi	70659,12	59,44	3,21	0,63	0,18	0,04	1,50	4,50
Min.	398344	7,25	4,03	0,59	0,10	0,28	2,35	68,46
Max.	671936	283,28	15,71	2,76	0,84	0,46	8,35	85,72

Variabel X1 yaitu garis kemiskinan memiliki rata-rata tertinggi sebesar 507425,60 dengan standar deviasi sebesar 70659,12, menunjukkan adanya perbedaan signifikan antar kabupaten dalam hal kebutuhan minimum dasar yang dibutuhkan untuk hidup layak. Nilai minimum sebesar 398344 dan maksimum 671936 menunjukkan rentang nilai yang lebar, mengindikasikan adanya ketimpangan garis kemiskinan antar wilayah. Variabel X2 (jumlah penduduk miskin) memiliki rata-rata sebesar 105,84 dan standar deviasi 59,44, yang menandakan tingkat variasi yang cukup tinggi dalam jumlah penduduk miskin antar kabupaten. Rentang nilainya dari 7,25 hingga 283,28 menunjukkan bahwa jumlah penduduk miskin sangat bervariasi antar wilayah, yang dapat disebabkan oleh perbedaan populasi total atau kondisi sosial-ekonomi setempat. Variabel X3 dengan rata-rata 10,13 dan standar deviasi 3,21, variabel ini mencerminkan tingkat kemiskinan relatif antar wilayah. Rentang dari 4,03 hingga 15,71

menunjukkan bahwa meskipun variasinya tidak sebesar jumlah absolut (X2), perbedaan tingkat kemiskinan relatif antar kabupaten tetap signifikan.

Variabel X4 memiliki rata-rata sebesar 1,51 dengan standar deviasi 0,63 menunjukkan bahwa rata-rata penduduk miskin berada sedikit di bawah garis kemiskinan, namun dengan variasi yang cukup moderat antar kabupaten. Nilai minimum sebesar 0,59 dan maksimum 2,76 memperlihatkan bahwa kedalaman kemiskinan tidak terlalu ekstrem tetapi tetap terdapat perbedaan yang perlu diperhatikan. Variabel indeks keparahan kemiskinan (X5) memiliki rata-rata 0,35 dan standar deviasi 0,18, menunjukkan bahwa tingkat ketimpangan di antara penduduk miskin tidak terlalu tinggi dan relatif seragam antar kabupaten. Ini menunjukkan bahwa meskipun ada kemiskinan, jurang ketimpangan antar yang miskin tidak terlalu dalam. Variabel gini raiso (X6) dengan rata-rata 0,35 dan standar deviasi 0,04, variabel ini menunjukkan tingkat ketimpangan pendapatan secara umum di setiap kabupaten relatif seragam. Nilai ini mencerminkan bahwa ketimpangan ekonomi antar wilayah berada pada level yang mirip, dengan variasi yang rendah (dari 0,28 hingga 0,46).

Variabel tingkat pengangguran terbuka (X7) rata-rata sebesar 4,52 dengan standar deviasi 1,50 menunjukkan bahwa tingkat pengangguran antar kabupaten memiliki perbedaan yang cukup berarti, dengan rentang nilai dari 2,35 hingga 8,35. Ini menandakan bahwa kondisi pasar tenaga kerja tidak merata antar wilayah. Variabel indeks pembangunan manusia (X8) dengan rata-rata 74,75 dan standar deviasi 6,50, variabel ini menunjukkan tingkat pembangunan manusia antar kabupaten cenderung tinggi, tetapi masih ada disparitas yang perlu diperhatikan. Rentang nilai dari 68,46 hingga 85,72 menunjukkan bahwa beberapa wilayah memiliki pencapaian pembangunan yang jauh lebih baik dibandingkan yang lain.

Hasil analisis deskriptif di atas menunjukkan bahwa data memiliki perbedaan skala dan variasi yang cukup besar antar variabel, seperti garis kemiskinan yang bernilai ratusan ribu dan gini rasio yang berada dalam skala desimal. Selain itu, terdapat ketidakseimbangan dalam sebaran nilai antar kabupaten. Kondisi ini membuat data kurang cocok untuk metode klasterisasi yang sensitif terhadap skala, seperti *K-Means*. Oleh karena itu, digunakan metode *K-Harmonic Means* yang lebih tahan terhadap perbedaan skala, inisialisasi centroid, dan sebaran yang tidak merata.

Preprocessing Data

Preprocessing data dilakukan untuk menyeragamkan skala pengukuran sehingga nilai antar variabel setara. Hasil *preprocessing data* ini diperoleh melalui proses standarisasi, yaitu dengan mengubah skala data asli menjadi skala baku yang memiliki rata-rata 0 dan standar deviasi 1. Dengan demikian, seluruh variabel memiliki skala yang sebanding, sehingga dapat dianalisis secara adil tanpa adanya dominasi variabel tertentu akibat perbedaan satuan pengukuran.

Tabel 3. *Preprocessing data*

Kabupaten	X1	X2	X3	X4	X5	X6	X7	X8
Kabupaten Cilacap	-0,95	1,37	0,17	0,13	0,01	0,11	2,25	-0,53
Kabupaten Banyumas	-0,09	1,74	0,58	0,94	0,85	0,86	1,13	-0,05
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Kota Pekalongan	1,41	-1,45	-1,08	-1,10	-0,99	-0,52	0,27	0,56
Kota Tegal	2,26	-1,48	-0,79	-0,27	-0,15	-0,62	0,92	0,60

Tabel 3 menunjukkan hasil *preprocessing data* pada Kabupaten/Kota di Provinsi Jawa Tengah dengan rentang nilai -1 hingga 2. Nilai tersebut mencerminkan perbedaan karakteristik antar wilayah. Nilai-nilai pada setiap kolom telah dinormalisasi sehingga dapat dibandingkan langsung antar wilayah tanpa terpengaruh oleh skala asli. Pada kota umumnya menunjukkan nilai positif tinggi pada X1 dan nilai negatif pada X2 hingga X6. Sebaliknya kabupaten memiliki nilai positif pada X2 hingga X7, sedangkan pada X1 dan X8 menunjukkan nilai yang negatif, hal ini menandakan wilayah ini masih menghadapi tantangan serius dalam pembangunan sosial ekonomi.

Clustering dengan *K-Harmonic Means*

1. Uji Asumsi Multikolinieritas

Pada tahap awal analisis, terdeteksi adanya multikolinearitas pada variabel X4, yang ditunjukkan dengan nilai VIF > 10 . Untuk mengatasi masalah ini, dilakukan penanganan dengan mengeliminasi variabel X4 dari analisis. Setelah eliminasi dilakukan, nilai VIF masing-masing variabel dihitung kembali dan hasilnya ditunjukkan pada Tabel 4. Dimana seluruh variabel yang tersisa memiliki nilai VIF di bawah 10, sehingga dapat disimpulkan bahwa tidak terdapat masalah multikolinearitas yang signifikan dalam data.

Tabel 4. Nilai VIF uji multikolinearitas

Variabel	X1	X2	X3	X5	X6	X7	X8
Nilai VIF	2,81	3,46	7,40	3,78	1,95	1,77	5,16

Berdasarkan Tabel 4 dapat dilihat bahwa seluruh variabel yang tersisa memiliki nilai VIF < 10 , sehingga dapat disimpulkan bahwa tidak terdapat masalah multikolinearitas yang signifikan dalam model.

2. Penentuan Banyaknya klaster yang digunakan

Penentuan jumlah klaster yang paling optimal dilakukan dengan menggunakan metode *Silhouette Coefficient*, yang mampu mengukur kualitas pemisahan antar klaster. Berdasarkan hasil analisis yang dilakukan dengan bantuan perangkat lunak Python, diperoleh nilai koefisien *Silhouette* untuk masing-masing jumlah klaster yang disajikan pada Tabel 5.

Tabel 5. Nilai *Silhouette Coefficients*

Banyak Klaster	<i>Silhouette Coefficient</i>
2	0,28
3	0,13
4	0,12
5	0,17

Berdasarkan Tabel 5, nilai *Silhouette Coefficient* tertinggi diperoleh pada saat jumlah klaster sebanyak 2, yaitu sebesar 0,28. Nilai ini menunjukkan bahwa struktur pengelompokan paling optimal terjadi ketika data dibagi ke dalam dua klaster, karena memberikan pemisahan antar klaster yang paling jelas dibandingkan jumlah klaster lainnya. Dengan demikian, dapat disimpulkan bahwa jumlah klaster terbaik untuk digunakan dalam analisis ini adalah dua klaster.

3. Penentuan pusat klaster Awal

Penentuan pusat awal untuk 2 klaster dilakukan berdasarkan nilai dari masing-masing variabel yang telah distandarisasi. Pusat awal ini menjadi acuan awal dalam perhitungan jarak harmonik setiap objek terhadap pusat klaster. Nilai yang digunakan sebagai pusat awal masing-masing klaster dapat dilihat pada Tabel 6.

Tabel 6. Nilai awal pembentukan 2 klaster

Pusat Awal	X_1	X_2	X_3	X_5	X_6	X_7	X_8
C_1	-1,57	0,54	1,45	1,46	0,16	0,71	-1,16
C_2	2,36	-0,48	-1,93	1,46	1,87	0,88	2,37

Berdasarkan Tabel 6, pusat klaster C_1 memiliki nilai garis kemiskinan ditandai dengan nilai x_1 sebesar -1.57 yang menunjukkan batas kemiskinan relatif rendah, mengidentifikasi kondisi ekonomi yang rentan. Jumlah penduduk miskin ($x_2 = 0,54$) dan persentase penduduk miskin ($x_3 = 1,45$) cukup tinggi, memperkuat gambaran adanya banyak warga yang hidup di bawah garis kemiskinan. Indeks keparahan kemiskinan ($x_5 = 1,46$) yang tinggi mengindikasikan kondisi kemiskinan yang cukup parah. Meskipun ketimpangan Gini Rasio ($x_5 = 1,46$) relatif rendah, ini kemungkinan menunjukkan bahwa hampir semua warga mengalami kondisi ekonomi yang serupa, yaitu kurang mampu. Tingkat pengangguran terbuka ($x_7 = 0,71$) yang sedang menunjukkan masih adanya persoalan lapangan kerja, sementara IPM ($x_8 = -1,16$) yang menggambarkan kualitas hidup, pendidikan, dan kesehatan yang kurang memadai di wilayah ini.

Sebaliknya, klaster C_2 menunjukkan nilai garis kemiskinan ($x_1 = 2,36$) yang tinggi, namun diiringi dengan nilai IPM ($x_8 = 2,37$) yang juga tinggi, yang berarti meski garis kemiskinan relatif tinggi, kualitas pembangunan manusia di wilayah ini lebih baik. Jumlah dan persentase penduduk miskin ($x_2 = -0,48$) dan ($x_3 = -1,93$) yang sangat rendah, menandakan kesejahteraan masyarakat yang lebih baik. Indeks keparahan kemiskinan ($x_5 = 1,46$) tetap tinggi, namun dengan ketimpangan pendapatan ($x_6 = 1,87$) yang jauh lebih besar, hal ini mengindikasikan bahwa distribusi pendapatan di wilayah ini tidak merata, dengan adanya kelompok masyarakat yang sangat kaya dan sangat miskin. Tingkat pengangguran terbuka ($x_7 = 0,88$) yang cukup tinggi menandakan masalah lapangan kerja yang masih ada, meskipun secara umum wilayah ini lebih maju dalam aspek sosial ekonomi.

4. Jarak setiap Data

Nilai jarak ini mencerminkan seberapa dekat karakteristik suatu wilayah dengan profil masing-masing klaster. Semakin kecil nilai jarak terhadap pusat klaster tertentu, semakin besar kemungkinan wilayah tersebut tergolong ke dalam klaster tersebut. Hasil perhitungan jarak masing-masing terhadap dua pusat klaster yang terbentuk disajikan didalam Tabel 7.

Tabel 7. Nilai jarak setiap data

Objek	C_1	C_2	Objek	C_1	C_2	Objek	C_1	C_2
0	2,74	5,82	12	3,48	4,58	24	3,15	5,97
1	2,58	5,35	13	2,58	6,50	25	3,78	5,16
2	1,62	6,62	14	2,65	5,85	26	2,07	7,17
3	8,70	7,12	15	2,62	5,54	27	3,97	5,03
4	1,85	6,69	16	3,11	7,01	28	4,02	7,85
5	3,44	4,92	17	3,92	4,47	29	6,67	2,16
6	1,50	6,98	18	4,75	4,18	30	5,85	2,24
7	2,95	5,91	19	4,97	5,39	31	6,86	2,17
8	3,13	5,37	20	2,73	5,60	32	7,12	7,84
9	3,95	4,27	21	4,39	3,89	33	5,36	3,47
10	4,36	3,60	22	4,03	5,62	34	5,49	3,59
11	3,26	5,57	23	3,64	4,19			

Berdasarkan Tabel 7, objek yang memiliki jarak lebih kecil ke pusat kluster C_1 cenderung tergolong ke dalam Klaster 1, seperti Objek 2 ($1,68 < 6,62$), sementara objek yang lebih dekat ke C_2 masuk Klaster 2, seperti Objek 30 ($2,24 < 5,85$). Hal ini menunjukkan bahwa metode *K-Harmonic Means* efektif dalam mengelompokkan wilayah berdasarkan kemiripan karakteristik kemiskinan.

5. Fungsi Objektif

Fungsi objektif pada analisis *K-Harmonic Means* merepresentasikan total rata-rata harmonik jarak antara seluruh data dengan pusat klaster yang terbentuk. Semakin kecil nilai fungsi objektif, semakin baik kualitas pemisahan antar klaster yang dihasilkan oleh model. Nilai fungsi objektif hasil analisis ditampilkan pada Tabel 8.

Tabel 8. Fungsi Objektif	
Fungsi	Nilai
Objektif KHM	237,10

Nilai fungsi objektif sebesar 237,10 menunjukkan bahwa hasil pengelompokan yang diperoleh telah menghasilkan distribusi data yang cukup optimal terhadap pusat klasternya. Hal ini mengindikasikan bahwa model *K-Harmonic Means* berhasil meminimalkan jarak rata-rata harmonik antar data dan klaster, sehingga klaster yang terbentuk dapat merepresentasikan pola kemiskinan di kabupaten/kota Jawa Tengah dengan baik.

6. Proporsi Keanggotaan

Proporsi keanggotaan masing-masing kabupaten/kota terhadap dua klaster yang terbentuk untuk menunjukkan kekuatan keterkaitan wilayah terhadap masing-masing klaster. Nilai proporsi tersebut disajikan pada Tabel 9.

Tabel 9. Proporsi Keanggotaan								
Objek	C_1	C_2	Objek	C_1	C_2	Objek	C_1	C_2
0	0,82	0,18	12	0,63	0,37	24	0,78	0,22
1	0,81	0,19	13	0,86	0,14	25	0,65	0,08
2	0,94	0,06	14	0,83	0,17	26	0,92	0,08
3	1,00	0,00	15	0,82	0,18	27	0,62	0,38
4	0,93	0,07	16	0,84	0,17	28	0,79	0,21
5	0,67	0,33	17	0,57	0,43	29	0,10	0,91
6	0,96	0,04	18	0,44	0,56	30	0,13	0,87
7	0,80	0,20	19	0,54	0,46	31	0,09	0,91
8	0,75	0,25	20	0,82	0,19	32	0,00	1,00
9	0,54	0,46	21	0,44	0,56	33	0,30	0,75
10	0,40	0,60	22	0,66	0,34	34	0,30	0,70
11	0,74	0,26	23	0,57	0,43			

Proporsi keanggotaan menunjukkan mayoritas kabupaten/kota lebih dominan dalam klaster C_1 , ditandai oleh nilai proporsi di atas 0,7 pada sebagian besar objek seperti objek 3 (1,00), objek 6 (0,96), dan objek 2 (0,94). Wilayah dalam klaster C_1 cenderung memiliki karakteristik tingkat kemiskinan rendah, IPM tinggi, dan indikator sosial ekonomi yang lebih baik. Sementara itu, objek 29, 30, dan 31 memiliki proporsi yang kuat dalam klaster C_2 (lebih dari 0,9), yang menunjukkan wilayah dengan tingkat kemiskinan relatif tinggi,

ketimpangan lebih besar, dan kualitas hidup lebih rendah. Wilayah yang proporsinya seimbang (misalnya objek 20 dan 21) mengindikasikan transisi antara kedua klaster.

7. Nilai Bobot (*Weights*) Tiap Kabupaten/Kota terhadap Klaster

Nilai bobot dihitung menggunakan jarak *euclidean* data yang sebelumnya telah dilakukan perhitungan. Nilai Bobot (*Weights*) Tiap Kabupaten/Kota terhadap Klaster ditampilkan pada Tabel 10.

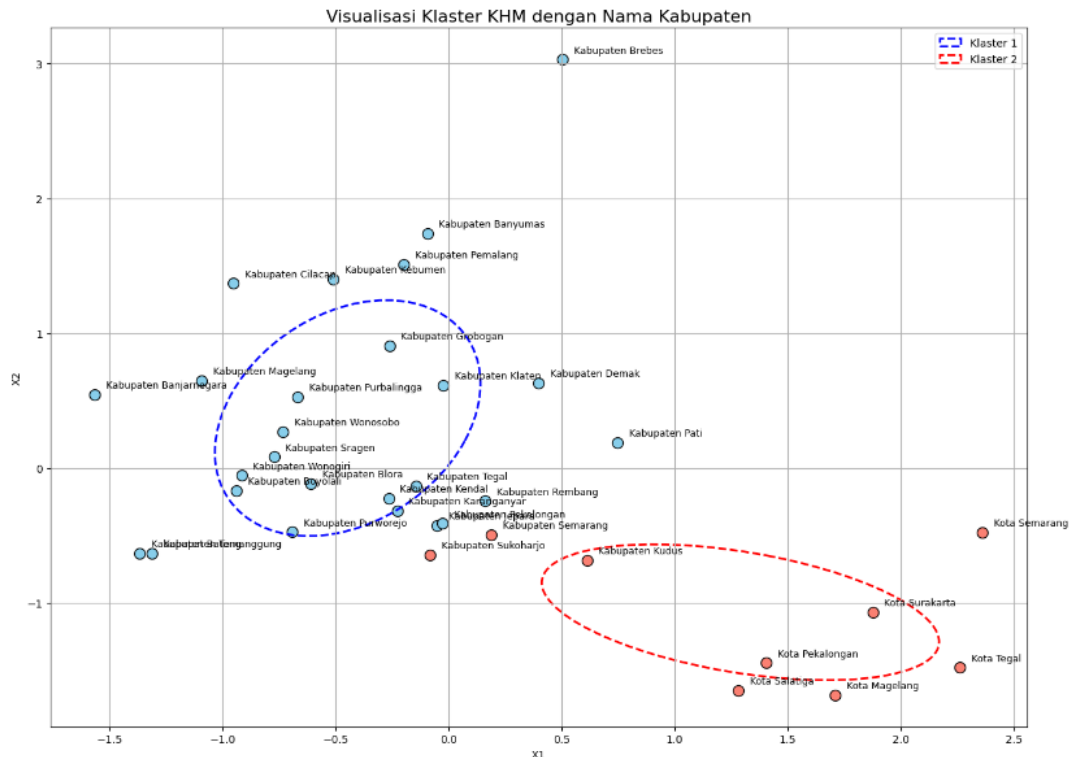
Tabel 10. Nilai Bobot 2 klaster

Objek	Bobot	Objek	Bobot	Objek	Bobot	Objek	Bobot	Objek	Bobot
0	6,15	7	6,98	14	5,82	21	8,48	28	12,78
1	5,38	8	7,30	15	5,61	22	10,73	29	4,23
2	2,47	9	8,41	16	8,09	23	7,55	30	4,37
3	0,00	10	7,69	17	8,69	24	7,76	31	4,27
4	3,19	11	7,93	18	9,84	25	9,29	32	0,00
5	7,95	12	7,67	19	13,35	26	3,97	33	8,49
6	2,15	13	5,75	20	6,04	27	9,71	34	9,02

Nilai bobot pada metode *K-Harmonic Means* menunjukkan kontribusi relatif setiap objek (kabupaten/kota) terhadap pusat klaster yang terbentuk. Semakin besar nilai bobot suatu objek terhadap suatu klaster, maka semakin besar pula pengaruhnya dalam menentukan posisi pusat klaster tersebut. Bobot ini juga mencerminkan tingkat kedekatan relatif objek terhadap semua pusat klaster secara harmonik, sehingga dapat membantu mengidentifikasi kabupaten/kota yang menjadi representatif utama dalam setiap klaster.

8. Hasil Pengelompokan Wilayah Berdasarkan Klaster *K-Harmonic Means*

Penentuan hasil pengelompokkan wilayah dilakukan dengan penentuan nilai K optimal menggunakan metode *Silhouette*, yang menghasilkan nilai K=2 dengan skor *Silhouette* sebesar 0,28. Berdasarkan hasil analisis menggunakan metode *K-Harmonic Means*, kabupaten/kota di provinsi Jawa Tengah terbagi menjadi 2 klaster. Pembagian ini didasarkan pada kemiripan karakteristik indikator kemiskinan dan ketimpangan yang telah dianalisis sebelumnya. Klaster 1 terdiri atas wilayah-wilayah yang cenderung memiliki tingkat kemiskinan yang tinggi dan indeks pembangunan manusia (IPM) yang relatif lebih rendah. Wilayah dalam klaster ini umumnya merupakan kabupaten dengan kondisi sosial ekonomi yang belum sebaik kota. Klaster 2 mencakup wilayah dengan tingkat kemiskinan yang lebih rendah, ketimpangan yang relatif kecil, dan IPM yang lebih tinggi. Wilayah dalam klaster ini didominasi oleh kota-kota besar yang telah berkembang, seperti Kota Semarang, Kota Surakarta, dan Kota Magelang. Visualisasi hasil pengelompokan wilayah berdasarkan klaster ditampilkan pada Gambar 1.



Gambar 1. Visualisasi Klaster Kabupaten/Kota di Jawa Tengah

Ciri dari klaster yang terbentuk dapat diidentifikasi melalui nilai rata-rata dari setiap variabel. Nilai rata-rata tiap variabel pada setiap klaster ditunjukkan pada Tabel 11.

Tabel 11. Rata-Rata Tiap Variabel Berdasarkan Kluster

Klaster	X1	X2	X3	X5	X6	X7	X8
Klaster 1	476305,50	127,52	11,37	0,40	0,34	4,54	72,76
Klaster 2	597328,00	43,19	6,54	0,21	0,38	4,45	80,48

Klaster 1 memiliki rata-rata persentase penduduk miskin (X2) dan keparahan kemiskinan (X3) yang lebih tinggi dibandingkan Klaster 2. Hal ini menunjukkan bahwa wilayah-wilayah dalam Klaster 1 mengalami kondisi kemiskinan yang lebih berat dan merata. Selain itu, gini rasio (X5) pada Klaster 1 juga lebih tinggi, menandakan ketimpangan distribusi pengeluaran yang lebih besar.

Sebaliknya, Klaster 2 memiliki nilai rata-rata indeks pembangunan manusia (X8) yang lebih tinggi, yaitu 80,48, dibandingkan Klaster 1 sebesar 72,76. Hal ini menunjukkan bahwa wilayah-wilayah dalam Klaster 2, yang didominasi oleh kota, memiliki kualitas hidup yang lebih baik, termasuk dalam aspek pendidikan, kesehatan, dan standar hidup. Perbedaan yang signifikan antara kedua klaster ini memperkuat hasil klasterisasi bahwa kabupaten dan kota di Jawa Tengah memiliki karakteristik sosial ekonomi yang berbeda secara nyata.

SIMPULAN

Penelitian ini mengelompokkan 35 kabupaten/kota di Provinsi Jawa Tengah ke dalam dua klaster berdasarkan indikator kemiskinan dan ketimpangan menggunakan metode *K-Harmonic Means*. Hasil pengelompokan menunjukkan bahwa Klaster 1 terdiri dari 26 wilayah yang didominasi oleh kabupaten, dengan karakteristik tingkat kemiskinan, indeks kedalaman dan keparahan kemiskinan, serta ketimpangan yang relatif lebih tinggi. Sementara itu, Klaster 2 terdiri dari 9 wilayah, yang sebagian besar merupakan kota, dengan karakteristik sosial-ekonomi yang lebih baik, mencakup tingkat kemiskinan dan ketimpangan yang lebih rendah serta capaian pembangunan yang lebih tinggi. Validitas hasil pengelompokan ini diperkuat dengan nilai *Silhouette Coefficient* sebesar 0,28, yang menunjukkan struktur klaster yang cukup baik dan terpisah secara jelas. Secara substantif, hasil ini menggambarkan adanya kesenjangan pembangunan antara wilayah kabupaten dan kota di Jawa Tengah, yang perlu mendapatkan perhatian dalam perencanaan kebijakan daerah. Pemerintah dapat menggunakan hasil klaster ini untuk menyusun strategi pengentasan kemiskinan yang lebih spesifik dan berbasis wilayah, terutama dengan memberikan perhatian lebih besar pada klaster 1. Untuk pengembangan ke depan, penelitian ini dapat diperluas dengan menambahkan variabel sosial-ekonomi lainnya serta pendekatan spasial atau temporer guna memperoleh gambaran ketimpangan yang lebih komprehensif dan dinamis.

DAFTAR PUSTAKA

- Alyarahma, N. N., Kholijah, G., & Sormin, C. (2024). Pengelompokan Provinsi di Indonesia Menggunakan Gaussian Mixture Model Berdasarkan Indikator Kemiskinan. *Journal of Mathematics: Theory and Applications*, 6(2), 158–167. <https://doi.org/10.31605/jomta.v6i2.4032>
- Aprillyana Nia. (2020). Estimasi Indikator Kemiskinan Tingkat Kecamatan Menggunakan Regresi Kekar M-Kuantil. *Jurnal Litbang Sukowati*. <https://www.academia.edu/download/74336691/45>
- Aya Dewanti Sofia, & Kudus, A. (2023). Pengelompokan Kabupaten/Kota berdasarkan Indikator Indeks Pembangunan Manusia 2022 Menggunakan K-Harmonic Means Clustering. *Jurnal Riset Statistika*, 163–172. <https://doi.org/10.29313/jrs.v3i2.3130>
- Ayumna, F. S., & Sukardi, A. S. (2024). Pengaruh kemiskinan, pendidikan, dan kesehatan terhadap pertumbuhan ekonomi di Jawa Tengah pada era Ganjar Pranowo. *Journal of Economics Research and Policy Studies*, 4(1), 42–54. <https://doi.org/10.53088/jerps.v4i1.787>
- Kualitas, K., Kopi, B., Penyusutan, B. T., Putra, A., Abdullah, D., & Daud, M. (2024). Menggunakan Metode K-Harmonic Means dengan Validasi Silhouette Index dan C-index. *Jurnal Janitra Informatika Dan Sistem Informasi*, 4(2), 74–86. <https://doi.org/10.59395/janitra.v4i2.203>
- Mayasari, S. N., & Nugraha, J. (2023). Implementasi K-Means Cluster Analysis untuk Mengelompokkan Kabupaten/Kota Berdasarkan Data Kemiskinan di Provinsi Jawa Tengah Tahun 2022 (Vol. 3, Issue 2). <https://doi.org/https://doi.org/10.24002/konstelasi.v3i2.7200>
- Pratiwi, M. C. Y. (2021). Analisis Ketimpangan Antarwilayah dan Pergeseran Struktur Ekonomi di Kalimantan. *Jurnal Borneo Administrator*, 17(1), 131–154. <https://doi.org/10.24258/jba.v17i1.779>
- Putri, N., Nonong Amalita, Dodi Vionanda, & Dony Permana. (2023). Grouping Level of Poverty Based on District/City in Indonesia Using K-Harmonic Means. *UNP Journal of Statistics and Data Science*, 1(3), 157–163. <https://doi.org/10.24036/ujsds/vol1-iss3/60>

- Rahmah, M., Candra, A., & Sembiring, R. W. (2022). Identifikasi Predikat Hasil Pengelompokan Data Kualitas Udara dengan Menggunakan Affinity Propagation dan Silhouette Coefficient. *InfoTekJar: Jurnal Nasional Informatika Dan Teknologi Jaringan*, 6(2), 2540–7597. <https://doi.org/10.30743/infotekjar.v6i2.4670>
- Raja, N. A., Tinungki, G. M., & Sirajang, N. (2024). Implementasi Algoritma Centroid Linkage dan K-Medoids dalam Mengelompokkan Kabupaten/Kota di Sulawesi Selatan Berdasarkan Indikator Pendidikan. *ESTIMASI: Journal of Statistics and Its Application*, 61–74. <https://doi.org/10.20956/ejsa.v5i1.13605>
- Salsabila, A., Widiari, T., Statistika, D., & Sains dan Matematika, F. (2022a). *Metode K-Harmonic Means Clustering Dengan Validasi Silhouette Coefficient (Studi Kasus : Empat Faktor Utama Penyebab Stunting 34 Provinsi di Indonesia Tahun 2018)*. 11(1), 11–20. <https://doi.org/https://doi.org/10.14710/j.gauss.v11i1.34003>
- Salsabila, A., Widiari, T., Statistika, D., & Sains dan Matematika, F. (2022b). *Metode K-Harmonic Means Clustering Dengan Validasi Silhouette Coefficient (Studi Kasus : Empat Faktor Utama Penyebab Stunting 34 Provinsi Di Indonesia Tahun 2018)*. 11(1), 11–20. <https://doi.org/https://doi.org/10.14710/j.gauss.v11i1.34003>
- Solihat Aprilianti. (2024). *Pengelompokan Kabupaten/Kota di Jawa Barat*. https://eprints.unpak.ac.id/8767/1/Skripsi_064120006_Aprilianti%20Solihat
- Syahrani, E., Ayu Kusumaningdyah, A., & Danar Dewa, D. (2021). Analisis Faktor-Faktor Yang Mempengaruhi Tingkat Kemiskinan Kabupaten/Kota Di Jawa Tengah. <https://doi.org/10.29408/geodika.v5i2.4033>
- Widyadhana, D., Hastuti, R. B., Kharisudin, I., & Fauzi, F. (2021). Perbandingan Analisis Kluster K-Means dan Average Linkage untuk Pengklasteran Kemiskinan di Provinsi Jawa Tengah. *PRISMA, Prosiding Seminar Nasional Matematika*, 4, 584–594. <https://journal.unnes.ac.id/sju/index.php/prisma/>
- Yunistya, D. I., Goejantoro, R., Deny, F., & Amijaya, T. (2022a). *The Application Of K-Harmonic Means Method In District/City Grouping (Case Study: Poverty in Kalimantan Island in 2020) Penerapan Metode K-Harmonic Means Dalam Pengelompokan Kabupaten/Kota (Studi Kasus: Kemiskinan di Pulau Kalimantan Tahun 2020)*. 19(1), 51–64. <https://doi.org/10.20956/j.v19i1.21116>
- Yunistya, D. I., Goejantoro, R., Deny, F., & Amijaya, T. (2022b). *The Application Of K-Harmonic Means Method In District/City Grouping (Case Study: Poverty in Kalimantan Island in 2020) Penerapan Metode K-Harmonic Means Dalam Pengelompokan Kabupaten/Kota (Studi Kasus: Kemiskinan di Pulau Kalimantan Tahun 2020)*. 19(1), 51–64. <https://doi.org/10.20956/j.v19i1.21116>