

Segmentation Of West Java Province Based on Socio-Economic Indicators Using K-Means and Agglomerative Clustering Methods

Tiara Valentina^{1*}, Sefilah Naurah Sabela¹, Shindi Shella May War¹, Wahyu Syaifullah Jauharis Saputra¹

¹Universitas Pembangunan Nasional "Veteran" Jawa Timur, Surabaya, Indonesia

* Corresponding author: 23083010091@student.upnjatim.ac.id

Abstract

Regional segmentation based on socio-economic indicators is a crucial approach in data-driven development planning. Through accurate segmentation, governments can design more targeted policies aligned with the specific characteristics of each region. This study aims to compare two clustering methods, namely K-Means and Agglomerative Clustering, in grouping regions within West Java Province based on socio-economic indicators such as poverty rate, open unemployment rate, and Human Development Index (HDI). The analysis was conducted using the Python programming language on the Google Colab platform. Cluster performance was evaluated using the Elbow Method to determine the optimal number of clusters and the Silhouette Score to assess cluster quality. The results indicate that Agglomerative Clustering produces more consistent and interpretable segmentations, particularly in reflecting the socio-economic similarities between regions. However, in terms of computational efficiency, the K-Means method performs better due to its faster processing time. These findings offer valuable insights for regional policymakers in setting development priorities more effectively, grounded in the actual socio-economic conditions of each area.

Keywords: Agglomerative Clustering, Elbow Method, K-Means, Regional Segmentation, Silhouette Score, West Java

Article History:

Received : 09 June 2025

Revised : 30 June 2025

Accepted : 30 June 2025

© 2025 Valentina et al. Published by

Program Studi Statistika Fakultas MIPA

Universitas Lambung Mangkurat.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



How to Cite:

T. Valentina, S. N. Sabela, S. S. M. War, and W. S. J. Saputra. "Segmentation of West Java Province Based on Socio-Economic Indicators Using K-Means and Agglomerative Clustering Methods." *RAGAM: Journal of Statistics and Its Application*. 2025. 4(1), pp. 50–60. Doi: <https://doi.org/10.20527/ragam.v4i1.15589>

1. PENDAHULUAN

Pemerataan pembangunan antarwilayah merupakan prioritas strategis dalam perencanaan pembangunan nasional, khususnya di wilayah dengan kepadatan penduduk tinggi seperti Provinsi Jawa Barat. Ketimpangan sosial dan ekonomi antar kabupaten/kota di provinsi ini dapat menghambat pencapaian tujuan pembangunan berkelanjutan. Oleh karena itu, analisis segmentasi wilayah berbasis data sosial-ekonomi menjadi penting sebagai dasar perumusan kebijakan yang efektif dan berbasis bukti (*evidence-based policy*) (BPS Jawa Barat, 2023).

Jawa Barat memiliki keragaman karakteristik sosial-ekonomi antar daerah, yang terlihat dari perbedaan indikator seperti tingkat kemiskinan, pengangguran terbuka, angka partisipasi sekolah, serta indeks pembangunan manusia (IPM). Segmentasi wilayah berdasarkan indikator-indikator tersebut dapat membantu mengidentifikasi kelompok daerah dengan karakteristik serupa, sehingga memudahkan pengambilan keputusan dalam perencanaan pembangunan wilayah (Kemendagri, 2022).

Clustering merupakan salah satu metode dalam data mining yang sering digunakan untuk segmentasi wilayah. Di antara metode clustering yang banyak digunakan, K-Means dan *Agglomerative Clustering* menonjol karena keefektifannya dalam mengelompokkan data multivariat. K-Means adalah algoritma partisi yang bekerja dengan meminimalkan jarak rata-rata ke pusat kluster secara iteratif, sedangkan *Agglomerative Clustering* adalah metode hierarkis yang menggabungkan objek berdasarkan kedekatan hingga membentuk struktur pohon atau dendrogram (Zhao et al., 2021).

Dalam konteks geografis dan sosial-ekonomi, pemilihan metode clustering yang tepat sangat penting untuk mendapatkan hasil segmentasi yang akurat dan bermakna. Oleh karena itu, studi ini bertujuan untuk membandingkan performa metode K-Means dan *Agglomerative Clustering* dalam melakukan segmentasi wilayah di Provinsi Jawa Barat berdasarkan indikator sosial-ekonomi terkini. Hasil dari penelitian ini diharapkan dapat menjadi dasar dalam perencanaan pembangunan daerah yang lebih adaptif dan berbasis data.

Dalam studi ini, dilakukan analisis terhadap segmentasi wilayah kabupaten/kota di Provinsi Jawa Barat berdasarkan sejumlah indikator sosial-ekonomi. Tujuannya adalah untuk mengidentifikasi pola kesamaan karakteristik antarwilayah guna mendukung perencanaan pembangunan yang lebih terfokus dan berbasis data. Untuk mendukung segmentasi wilayah, digunakan dua metode unsupervised learning, yaitu K-Means Clustering dan *Agglomerative Clustering*. Sebelum penerapan kedua metode, dilakukan normalisasi data agar seluruh indikator memiliki skala yang sebanding.

Sebagai tambahan, dilakukan analisis visual terhadap hasil kluster menggunakan visualisasi plot dan dendrogram untuk melihat distribusi spasial antarwilayah hasil segmentasi. Hal ini memungkinkan pengamatan pola geografis yang mungkin tidak terlihat dari data numerik semata. Dengan pendekatan ini, perbandingan antara metode K-Means dan *Agglomerative Clustering* tidak hanya didasarkan pada hasil statistik, tetapi juga pada interpretabilitas dan konsistensi spasial hasil kluster yang dihasilkan.

2. TINJAUAN PUSTAKA

2.1. Segmentasi Wilayah Berbasis Data Sosial-Ekonomi

Segmentasi wilayah merupakan proses pengelompokan wilayah berdasarkan kesamaan karakteristik tertentu guna mendukung pengambilan kebijakan yang lebih tepat sasaran. Dalam konteks pembangunan daerah, segmentasi berbasis indikator sosial-ekonomi seperti IPM, tingkat kemiskinan, dan pengangguran dapat memberikan pemahaman yang lebih mendalam terhadap kesenjangan antarwilayah (Putra & Suryani, 2022). Pendekatan berbasis data ini semakin penting di era otonomi daerah, di mana pengelolaan pembangunan dilakukan secara lebih terdesentralisasi.

2.2. Clustering dalam Data Mining

Clustering adalah metode dalam data mining yang digunakan untuk mengelompokkan data ke dalam grup (klaster) yang memiliki kemiripan internal tinggi dan perbedaan antargrup yang signifikan (Han et al., 2021). Dua algoritma populer yang sering digunakan dalam segmentasi wilayah adalah K-Means Clustering dan *Agglomerative Hierarchical Clustering*.

2.3. K-Means Clustering

K-Means merupakan algoritma partisi yang membagi data ke dalam k klaster berdasarkan centroid (titik tengah klaster) yang diminimalkan berdasarkan jarak Euclidean. Metode ini terkenal karena efisiensinya, terutama untuk dataset berukuran besar. Namun, K-Means memiliki kelemahan dalam hal kepekaan terhadap pemilihan jumlah klaster awal dan *outlier* (Widodo et al., 2023).

2.4. Agglomerative Clustering

Agglomerative Clustering adalah salah satu jenis dari hierarchical clustering yang bekerja secara *bottom-up*, menggabungkan titik data paling mirip hingga terbentuk pohon hirarki (dendrogram). Keunggulan metode ini adalah tidak perlu menentukan jumlah klaster sejak awal, dan hasilnya dapat divisualisasikan dalam struktur hierarki yang informatif. Namun, metode ini lebih mahal secara komputasi dibanding K-Means (Nugroho & Safitri, 2021).

2.5. Evaluasi Hasil Clustering

Evaluasi hasil clustering penting untuk menentukan kualitas pemisahan antar klaster. Salah satu metrik yang umum digunakan adalah Silhouette Score, yang mengukur seberapa mirip suatu objek dengan klasternya sendiri dibandingkan dengan klaster lain. Nilai Silhouette Score yang tinggi menunjukkan bahwa objek terkelompok dengan baik (Sharma et al., 2022).

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode *unsupervised learning* berbasis clustering. Tujuan utama penelitian adalah melakukan segmentasi wilayah kabupaten/kota di Provinsi Jawa Barat berdasarkan indikator sosial-ekonomi.



Gambar 1. Diagram Alir

Langkah-langkah metodologi dalam penelitian ini adalah sebagai berikut:

3.1. Sumber Data

Data yang digunakan dalam penelitian ini berasal dari Publikasi Indikator Kesejahteraan Rakyat Provinsi Jawa Barat Tahun 2023, yang diperoleh dari sumber resmi BPS (Badan Pusat Statistik). Variabel yang digunakan dalam proses analisis terdiri dari:

- a) Indeks Pembangunan Manusia (IPM)
- b) Tingkat Pengangguran Terbuka (TPT)
- c) Produk Domestik Regional Bruto (PDRB) per Kapita
- d) Persentase Penduduk Miskin
- e) Rasio Gini
- f) Laju Pertumbuhan Ekonomi
- g) Laju Pertumbuhan Penduduk.

3.2. Pra-Pemrosesan Data

Tahapan pra-pemrosesan data dilakukan untuk menjamin kualitas dan keandalan data yang akan digunakan dalam analisis kluster. Pertama, data yang tidak lengkap atau mengandung nilai ekstrem yang tidak realistis dibersihkan dari dataset. Langkah ini penting untuk menghindari gangguan dan distorsi dalam proses pembentukan kluster sehingga hasil yang diperoleh dapat lebih akurat dan representatif. Selanjutnya, seluruh data numerik distandarisasi menggunakan teknik normalisasi *Z-score*. Proses normalisasi ini memastikan bahwa setiap variabel memiliki skala yang setara, sehingga metode clustering berbasis jarak dapat berjalan sesuai dengan asumsi dan memberikan hasil segmentasi yang valid serta dapat dibandingkan secara konsisten antar variabel.

3.3. Metode Analisis

Penelitian ini menggunakan dua metode *unsupervised learning* untuk mendukung segmentasi wilayah, yaitu K-Means Clustering dan *Agglomerative Hierarchical Clustering*. Sebelum menerapkan kedua metode tersebut, dilakukan normalisasi data agar seluruh indikator sosial-ekonomi memiliki skala yang sebanding. Indikator yang digunakan meliputi Indeks Pembangunan Manusia (IPM), Tingkat Pengangguran Terbuka (TPT), Produk Domestik Regional Bruto (PDRB) per kapita, tingkat kemiskinan, rasio Gini, laju pertumbuhan ekonomi, dan laju pertumbuhan penduduk.

K-Means Clustering adalah algoritma partisi yang membagi data menjadi sejumlah kluster (k) berdasarkan minimisasi jarak *Euclidean* antara setiap data dengan pusat kluster masing-masing. Algoritma ini bertujuan untuk membentuk kluster yang kompak dan homogen, sehingga anggota dalam kluster memiliki kemiripan yang tinggi. Sedangkan *Agglomerative Hierarchical Clustering* merupakan algoritma

bottom-up yang memulai proses dengan menganggap setiap data sebagai kluster tunggal. Kluster-kluster tersebut kemudian digabungkan secara bertahap berdasarkan kemiripan terdekat antar kluster hingga terbentuk hierarki kluster yang divisualisasikan dalam bentuk dendrogram. Pendekatan ini memungkinkan pemahaman yang lebih mendalam terhadap hubungan hierarkis antar wilayah.

Untuk mengevaluasi kualitas hasil kluster dari kedua metode tersebut, digunakan metrik *Silhouette Score* yang dirumuskan sebagai berikut:

$$Silhouette\ Score = \frac{b(i) - a(i)}{\max[a(i), b(i)]} \quad (1)$$

dimana:

$a(i)$ adalah rata-rata jarak antara data ke anggota kluster yang sama,

$b(i)$ adalah rata-rata jarak terkecil antara data ke kluster terdekat yang berbeda.

Nilai *Silhouette Score* berada dalam rentang $[-1,1]$, dengan nilai mendekati 1 menunjukkan pemisahan kluster yang baik dan konsistensi internal yang tinggi.

Selain evaluasi statistik, dilakukan juga analisis visual berupa plot dan dendrogram untuk melihat distribusi spasial hasil segmentasi. Hal ini mempermudah identifikasi pola geografis yang mungkin tidak terlihat hanya dari data numerik. Dengan demikian, perbandingan antara K-Means dan *Agglomerative Clustering* tidak hanya didasarkan pada hasil statistik, tetapi juga pada interpretabilitas dan konsistensi spasial hasil kluster yang diperoleh.

3.4. Interpretasi Hasil dan Kesimpulan

Setelah proses klusterisasi selesai, dilakukan interpretasi terhadap hasil, baik secara statistik deskriptif per kluster, maupun secara visual plot dan dendrogram untuk menunjukkan distribusi wilayah dalam setiap kluster.

3.5. Selesai

Hasil dari segmentasi ini dapat digunakan sebagai dasar rekomendasi kebijakan pembangunan wilayah berbasis data.

4. HASIL DAN PEMBAHASAN

4.1. Gambaran Umum Data

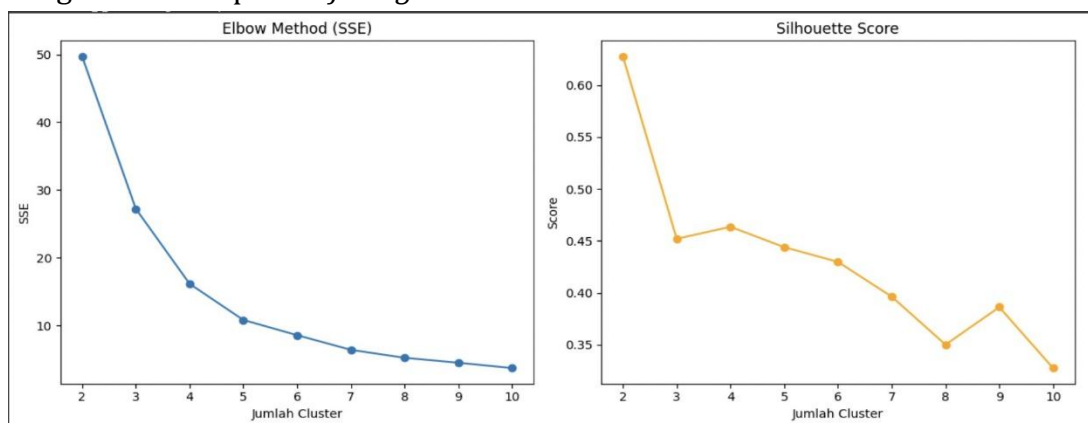
Dataset yang digunakan dalam penelitian ini terdiri dari 27 catatan yang merepresentasikan kabupaten/kota di Provinsi Jawa Barat. Variabel utama yang dianalisis meliputi sejumlah indikator sosial-ekonomi yang relevan dalam mengukur tingkat kesejahteraan dan pembangunan wilayah. Indeks Pembangunan Manusia (IPM) digunakan sebagai ukuran komposit yang menggambarkan kualitas pembangunan manusia berdasarkan dimensi kesehatan, pendidikan, dan standar hidup. Tingkat Pengangguran Terbuka (TPT) diukur dalam persentase dan

menunjukkan proporsi angkatan kerja yang aktif mencari pekerjaan namun belum memperoleh pekerjaan. PDRB per Kapita dihitung dalam satuan juta rupiah, merepresentasikan rata-rata nilai Produk Domestik Regional Bruto yang dihasilkan per individu dalam satu tahun. Selanjutnya, Persentase Penduduk Miskin menunjukkan proporsi penduduk yang berada di bawah garis kemiskinan terhadap total populasi di wilayah tersebut. Rasio Gini digunakan sebagai indikator ketimpangan distribusi pendapatan, di mana nilai mendekati nol menunjukkan pemerataan yang tinggi, sedangkan nilai mendekati satu menunjukkan tingkat ketimpangan yang besar. Laju Pertumbuhan Ekonomi menggambarkan pertumbuhan ekonomi wilayah berdasarkan peningkatan nilai PDRB atas dasar harga konstan. Terakhir, Laju Pertumbuhan Penduduk mengindikasikan seberapa cepat jumlah penduduk bertambah dalam kurun waktu satu tahun. Seluruh variabel ini menjadi dasar dalam proses segmentasi wilayah melalui metode klusterisasi untuk memahami perbedaan karakteristik antar kabupaten/kota di Provinsi Jawa Barat.

4.2. Penentuan Jumlah Kluster Optimal

Sebelum melakukan analisis cluster, dilakukan penentuan jumlah cluster optimal

- a. Berdasarkan grafik *Elbow Method* (Gambar 2), ditemukan titik siku pada $k = 4$, yang menunjukkan penurunan signifikan dalam nilai inertia. Titik ini mengindikasikan jumlah kluster optimal untuk meminimalkan kesalahan segmentasi tanpa *overfitting*.



Gambar 2. penentuan jumlah kluster

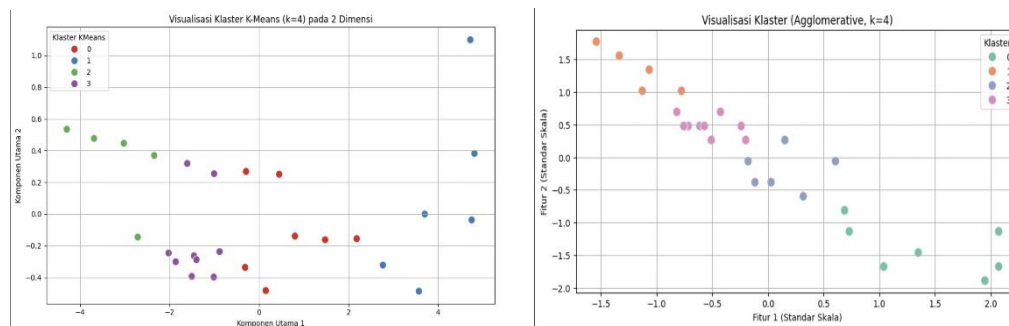
- b. Hasil Elbow tersebut diperkuat dengan nilai *Silhouette Score* yang juga relatif tinggi pada $k = 4$ (Gambar 3). Hal ini menunjukkan bahwa $k = 4$ menghasilkan kluster dengan kohesi internal dan pemisahan antar kluster yang cukup baik.



Gambar 3. Perbandingan *Silhouette Score* antara K-Means dan *Agglomerative Clustering*

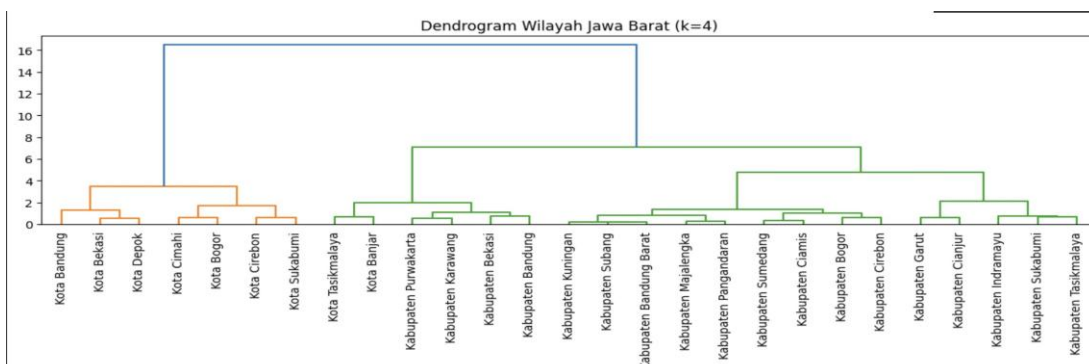
4.3. Hasil Analisis Klasterisasi

Setelah jumlah kluster optimal ditentukan sebanyak empat, dilakukan proses klasterisasi menggunakan metode K-Means dan *Agglomerative Clustering*. Hasil visualisasi dua dimensi dari metode K-Means dan *Agglomerative Clustering* ditunjukkan pada Gambar 4. Metode *Agglomerative Clustering* ini menghasilkan pengelompokan yang lebih kompak dan terstruktur secara hierarkis.



Gambar 4. Visualisasi Kluster K-Means dan *Agglomerative Clustering* dalam 2 Dimensi

Selain itu, dendrogram wilayah (Gambar 5) menunjukkan struktur hubungan antar wilayah di Provinsi Jawa Barat secara hierarkis.



Gambar 5. Dendrogram Wilayah Provinsi Jawa Barat

Berdasarkan hasil klasterisasi dengan metode *Agglomerative Clustering* pada $k = 4$, diperoleh distribusi wilayah sebagai berikut:

Tabel 1. Komposisi Anggota Klaster

Klaster	Jumlah Anggota	Daerah
0	7	Kota Bandung, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Bogor, Kota Cirebon, Kota Sukabumi
1	5	Kabupaten Sukabumi, Kabupaten Tasikmalaya, Kabupaten Garut, Kabupaten Cianjur, Kabupaten Indramayu
2	6	Kota Tasikmalaya, Kota Banjar, Kabupaten Bekasi, Kabupaten Bandung, Kabupaten Karawang, Kabupaten Purwakarta
3	9	Kabupaten Bogor, Kabupaten Cirebon, Kabupaten Sumedang, Kabupaten Pangandaran, Kabupaten Majalengka, Kabupaten Bandung Barat, Kabupaten Subang, Kabupaten Kuningan, Kabupaten Ciamis

4.4. Rata-rata Karakteristik Tiap Klaster

Untuk memahami profil masing-masing klaster, dilakukan perhitungan rata-rata dari indikator sosial-ekonomi pada setiap klaster. Hasilnya disajikan pada Tabel 2.

Tabel 2. Rata-rata Indikator Sosial-Ekonomi per Klaster

Indikator	Klaster 0	Klaster 1	Klaster 2	Klaster 3
Indeks Pembangunan Manusia (IPM)	80.31	67.80	74.12	70.86
Tingkat Pengangguran Terbuka (TPT) (%)	6.26	8.80	7.37	8.00
PDRB per Kapita (Rp Juta)	100.71	43.00	68.67	52.56
Persentase Penduduk Miskin (%)	4.29	9.00	6.28	7.26
Rasio Gini	0.32	0.204	0.258	0.223
Laju Pertumbuhan Ekonomi (%)	4.51	1.72	3.07	2.49
Laju Pertumbuhan Penduduk (%)	1.09	0.06	0.62	0.31

4.5. Interpretasi Hasil Klaster

Klaster 0 terdiri dari kota-kota besar dengan IPM dan PDRB per kapita yang sangat tinggi serta tingkat kemiskinan yang rendah. Klaster ini merepresentasikan wilayah dengan pembangunan manusia dan ekonomi yang paling maju di Jawa Barat. Sementara itu, Klaster 1 didominasi oleh kabupaten dengan kondisi sosial-ekonomi paling tertinggal. IPM rendah, kemiskinan tinggi, dan pertumbuhan ekonomi yang lambat mengindikasikan perlunya intervensi pembangunan yang lebih masif.

Klaster 2 berisi wilayah dengan indikator sosial-ekonomi menengah. PDRB dan IPM cukup baik namun belum sebaik klaster 0, menunjukkan potensi wilayah untuk berkembang dengan dukungan kebijakan yang tepat. Adapun Klaster 3 mencerminkan wilayah yang memiliki pembangunan manusia dan ekonomi yang relatif stabil namun masih menghadapi tantangan dalam hal ketimpangan dan pengangguran. Keempat klaster ini memberikan gambaran yang komprehensif

mengenai segmentasi sosial-ekonomi antar wilayah di Provinsi Jawa Barat. mencerminkan wilayah yang memiliki pembangunan manusia dan ekonomi yang relatif stabil namun masih menghadapi tantangan dalam hal ketimpangan dan pengangguran.

4.6. Insight dan Rekomendasi

Hasil klasterisasi menunjukkan bahwa terdapat ketimpangan sosial-ekonomi yang cukup mencolok antar wilayah di Provinsi Jawa Barat. Wilayah pada Klaster 1 memerlukan perhatian khusus karena mencerminkan daerah dengan capaian pembangunan yang rendah, baik dari sisi IPM, PDRB per kapita, maupun pertumbuhan ekonomi. Pemerintah provinsi disarankan untuk memberikan intervensi langsung dalam bentuk program peningkatan kualitas pendidikan dan kesehatan, pembangunan infrastruktur dasar, serta penciptaan lapangan kerja di sektor produktif.

Wilayah dalam Klaster 2 dan Klaster 3 dapat menjadi sasaran program akselerasi pembangunan. Pemerintah dapat memfokuskan kebijakan pada pengembangan usaha mikro, kecil, dan menengah (UMKM), serta mendorong investasi pada sektor-sektor yang memiliki daya ungkit tinggi terhadap pertumbuhan ekonomi. Di samping itu, peningkatan akses terhadap teknologi dan layanan digital di wilayah-wilayah ini juga dapat mempercepat proses transformasi ekonomi.

Adapun Klaster 0 yang mencakup kota-kota besar dengan capaian pembangunan tertinggi, perlu difokuskan pada pemeliharaan kualitas layanan publik, pengembangan inovasi, dan penerapan prinsip pembangunan berkelanjutan. Pemerintah daerah di wilayah ini dapat menjadi pusat pertumbuhan dan kolaborasi regional yang mendukung wilayah lain, misalnya melalui kerja sama antar daerah dalam bidang perdagangan, pendidikan, dan teknologi.

5. KESIMPULAN

Penelitian ini bertujuan untuk melakukan segmentasi wilayah Provinsi Jawa Barat berdasarkan indikator sosial-ekonomi menggunakan metode K-Means dan *Agglomerative Clustering*. Berdasarkan hasil analisis dengan pendekatan *Elbow Method* dan *Silhouette Score*, jumlah klaster optimal ditetapkan sebanyak empat. Metode *Agglomerative Clustering* dipilih karena memberikan visualisasi dan struktur klaster yang lebih informatif melalui dendrogram.

Hasil klasterisasi membagi wilayah menjadi empat kelompok dengan karakteristik yang berbeda. Klaster 0 terdiri dari kota-kota besar dengan tingkat pembangunan yang tinggi, Klaster 1 mewakili daerah yang masih tertinggal secara sosial-ekonomi, Klaster 2 merupakan wilayah dengan kondisi menengah yang potensial, dan Klaster 3 menggambarkan daerah yang stabil namun menghadapi tantangan ketimpangan. Perbedaan ini penting untuk merumuskan kebijakan yang lebih adaptif dan kontekstual.

Dengan demikian, pendekatan segmentasi wilayah berbasis klasterisasi memberikan

kontribusi strategis bagi perencanaan pembangunan daerah. Pemerintah dapat menggunakan hasil ini untuk menyusun kebijakan yang lebih spesifik, efisien, dan berkeadilan guna mempercepat pencapaian pembangunan sosial-ekonomi yang merata di seluruh wilayah Provinsi Jawa Barat.

DAFTAR PUSTAKA

- [1] BPS Jawa Barat. (2023). *Indikator Kesejahteraan Rakyat Provinsi Jawa Barat 2023*. Badan Pusat Statistik Provinsi Jawa Barat.
- [2] Kementerian Dalam Negeri (Kemendagri). (2022). *Profil Perkembangan Daerah: Provinsi Jawa Barat 2022*. Jakarta: Direktorat Jenderal Bina Pembangunan Daerah.
- [3] Zhao, Y., Zhang, Z., & Wang, J. (2021). *Comparative analysis of clustering algorithms on multidimensional socioeconomic data*. Journal of Big Data Analytics in Regional Science, 3(2), 112–126.
- [4] Han, J., Pei, J., & Kamber, M. (2021). *Data Mining: Concepts and Techniques (4th ed.)*. Morgan Kaufmann.
- [5] Nugroho, T., & Safitri, I. (2021). *Analisis Kluster Hierarki untuk Segmentasi Wilayah Berdasarkan Indikator Sosial*. Jurnal Informatika dan Sistem Informasi, 17(1), 33–41.
- [6] Putra, H. R., & Suryani, A. (2022). *Segmentasi Wilayah Kabupaten/Kota Berdasarkan Indikator Sosial-Ekonomi Menggunakan Clustering*. Jurnal Statistika dan Komputasi, 12(2), 45–55.
- [7] Sharma, A., Verma, R., & Gupta, N. (2022). *A Comparative Study of Clustering Algorithms on Socioeconomic Datasets*. Journal of Data Science and Analytics, 4(1), 19–28.
- [8] Widodo, S., Hidayat, A., & Lestari, D. (2023). *Implementasi K-Means Clustering pada Segmentasi Wilayah Berbasis Data Sosial Ekonomi*. Jurnal Ilmu Komputer dan Aplikasi, 11(3), 112–121.
- [9] Amrulloh, K., & Pudjiantoro, T. (2025). *Implementation of K-Means Clustering for Social Assistance Recipients*. Brilliance: Research and Innovation Journal, 6(2), 45–52.
- [10] Arifin, M. A., & Sari, D. P. (2024). *Segmentasi Wilayah Indonesia Berdasarkan IHK Menggunakan AHC dan Spectral Clustering*. Jurnal Teknologi Informasi dan Manajemen, 3(1), 12–20.
- [11] Fathia, N., & Rahman, A. (2023). *Grouping of Poverty in Indonesia Using K-Means with Silhouette Coefficient*. Journal of Data Science, 5(3), 88–95.
- [12] Hakim, R. A., & Lestari, N. (2024). *Penerapan Algoritma Agglomerative Clustering untuk Mengelompokkan Provinsi di Indonesia Berdasarkan Indikator Pendidikan*. Bit: Jurnal Fakultas Teknologi Informasi, 21(2), 109–117.
- [13] Hasan, Y. (2024). *Pengukuran Silhouette Score dan Davies-Bouldin Index pada Hasil Cluster K-Means dan DBSCAN*. KAKIFIKOM, 6(1), 55–63.

- [14] Hidayat, A., & Lestari, D. (2023). *Implementasi K-Means Clustering pada Segmentasi Wilayah Berbasis Data Sosial Ekonomi*. Jurnal Ilmu Komputer dan Aplikasi, 11(3), 112–121.
- [15] Nugroho, T., & Safitri, I. (2021). *Analisis Klaster Hierarki untuk Segmentasi Wilayah Berdasarkan Indikator Sosial*. Jurnal Informatika dan Sistem Informasi, 17(1), 33–41.
- [16] Putra, H. R., & Suryani, A. (2022). *Segmentasi Wilayah Kabupaten/Kota Berdasarkan Indikator Sosial-Ekonomi Menggunakan Clustering*. Jurnal Statistika dan Komputasi, 12(2), 45–55.
- [17] Sharma, A., Verma, R., & Gupta, N. (2022). *A Comparative Study of Clustering Algorithms on Socioeconomic Datasets*. Journal of Data Science and Analytics, 4(1), 19–28.
- [18] Widodo, S., Hidayat, A., & Lestari, D. (2023). *Implementasi K-Means Clustering pada Segmentasi Wilayah Berbasis Data Sosial Ekonomi*. Jurnal Ilmu Komputer dan Aplikasi, 11(3), 112–121.
- [19] Yuliana, S., & Prasetyo, B. (2023). *Analisis Tingkat Kemiskinan di Indonesia Menggunakan Metode K-Means Clustering*. Mahasiswa Indonesia, 4(2), 77–85.
- [20] Zahra, S. S. E. (2023). *Segmentation Analysis Using K-Means Clustering Model with SPSS: Case Study of Backpacker Jakarta Community Members*. Journal of Tourism and Hospitality, 5(1), 25–33.
- [21] Zulfa, M., & Hidayat, R. (2022). *Evaluation of K-Means Clustering Using Silhouette Score Method on Socioeconomic Data*. ILKOM: Jurnal Ilmiah, 14(2), 98–105.