



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 4 No. 4 (2025) pp: 7530-7536

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Clustering Wilayah Pulau Sumatera Berdasarkan Indikator Sosial Ekonomi Menggunakan Metode K-Medoids

Ika Desy Pramita Sari, Sefni Marcella Pratiwi, Harun Al Rosyid

Pendidikan Teknologi Informasi, Fakultas Teknik, Universitas Negeri Surabaya

ika.22081@mhs.unesa.ac.id, sefni.22051@mhs.unesa.ac.id, harunrosyid@unesa.ac.id

Abstrak

Ketidakseimbangan kondisi sosial ekonomi antarwilayah masih menjadi tantangan utama dalam pembangunan di Indonesia, khususnya di Pulau Sumatera yang memiliki keragaman kondisi geografis dan tingkat perkembangan wilayah. Perbedaan capaian indikator sosial ekonomi, seperti pendidikan, kesehatan, pengangguran, dan kemiskinan, menunjukkan perlunya pemetaan wilayah berbasis data untuk memetakan kemiripan karakteristik wilayah secara objektif. Penelitian ini bertujuan mengelompokkan 150 kabupaten/kota di Pulau Sumatera berdasarkan indikator sosial ekonomi menggunakan metode K-Medoids. Menggunakan data sekunder Badan Pusat Statistik (BPS) tahun 2023 yang mencakup Indeks Pembangunan Manusia (IPM), Angka Harapan Hidup (AHH), Rata-rata Lama Sekolah (RLS), Tingkat Pengangguran Terbuka (TPT), dan kemiskinan. Tahapan penelitian meliputi prapemrosesan data melalui penghapusan pada nilai kosong dan normalisasi menggunakan metode Min–Max Scaling, termasuk pembalikan skala pada indikator yang bersifat negatif terhadap kesejahteraan. Penentuan jumlah cluster optimal ditentukan melalui kombinasi metode Elbow dan Silhouette Score. Hasil analisis menunjukkan struktur data paling representatif terbentuk pada tiga cluster. Penerapan algoritma K-Medoids menghasilkan pengelompokan wilayah ke dalam tiga cluster dengan karakteristik sosial ekonomi relatif rendah, menengah, dan tinggi. Visualisasi menggunakan Principal Component Analysis (PCA) menunjukkan pola pemisahan cluster yang cukup jelas dan membantu interpretasi struktur pengelompokan wilayah. Pada hasil akhir, cluster pertama terdiri atas 34 kabupaten/kota, cluster kedua mencakup 84 kabupaten/kota, dan cluster ketiga beranggotakan 32 kabupaten/kota, yang menggambarkan distribusi ketimpangan sosial ekonomi antarwilayah di Pulau Sumatera.

Kata kunci: Clustering, Sumatera, K-Medoids, PCA

1. Latar Belakang

Ketimpangan sosial ekonomi antarwilayah masih menjadi tantangan utama dalam pembangunan di Indonesia. Perbedaan capaian indikator seperti tingkat kemiskinan, pengangguran, pendidikan, kesehatan, dan Indeks Pembangunan Manusia menunjukkan bahwa perkembangan wilayah belum berlangsung secara merata. Kondisi ini menuntut adanya pendekatan analisis berbasis data yang mampu mengelompokkan wilayah berdasarkan kesamaan karakteristik sosial ekonomi, sehingga proses perencanaan pembangunan dapat disusun secara lebih spesifik berdasarkan kebutuhan wilayah.

Metode *clustering* telah banyak digunakan untuk memetakan kondisi sosial ekonomi wilayah. Berbagai penelitian menunjukkan bahwa metode K-Means dan K-Medoids merupakan algoritma yang paling umum diterapkan dalam pengelompokan wilayah berdasarkan indikator sosial ekonomi dan kesehatan. Studi perbandingan antara K-Means dan K-Medoids pada data sosial ekonomi menunjukkan bahwa K-Medoids cenderung lebih stabil terhadap keberadaan pencilan karena pusat *cluster* ditentukan oleh objek aktual, bukan nilai rata-rata [1], [2]. Keunggulan ini menjadikan K-Medoids lebih sesuai digunakan pada data sosial ekonomi yang bersifat heterogen dan memiliki variasi ekstrem antarwilayah.

Penerapan K-Medoids juga telah dilakukan dalam analisis kerentanan sosial di Pulau Jawa dengan berbagai variasi ukuran jarak, dan hasilnya menunjukkan bahwa metode ini mampu menghasilkan pengelompokan wilayah yang representatif terhadap kondisi sosial masyarakat [3]. Penelitian lain mengkaji pengelompokan wilayah berdasarkan indikator inflasi dan ekonomi di Pulau Jawa dan Sumatra, namun analisis masih dilakukan secara agregat tanpa fokus mendalam pada karakteristik sosial ekonomi Pulau Sumatera secara spesifik [4]. Di sisi lain, beberapa penelitian hanya berfokus pada wilayah tertentu seperti Jawa Timur dan Jawa Barat dengan indikator kemiskinan terbatas [5], [6].

Clustering Wilayah Pulau Sumatera Berdasarkan Indikator Sosial Ekonomi Menggunakan Metode K-Medoids

Penelitian terbaru pada tingkat nasional juga menunjukkan perkembangan signifikan dalam penerapan K-Medoids. Studi tahun 2025 pengelompokan provinsi di Indonesia berdasarkan indikator kemiskinan dan kondisi ekonomi dengan menerapkan algoritma K-Medoids, dan membuktikan bahwa algoritma ini efektif dalam mengidentifikasi pola kemiripan antarprovinsi [7]. Namun, penelitian tersebut masih dilakukan pada level provinsi dan belum menyentuh variasi kondisi sosial ekonomi pada tingkat kabupaten/kota, khususnya di Pulau Sumatera yang memiliki karakteristik wilayah dan dinamika pembangunan yang berbeda antardaerah.

Selain metode *clustering*, penggunaan *Principal Component Analysis* (PCA) telah banyak dimanfaatkan untuk mereduksi aspek data dan membantu penggambaran hasil pengelompokan. Beberapa penelitian menunjukkan bahwa integrasi PCA dengan metode *clustering* dapat meningkatkan kejelasan struktur *cluster* dan mempermudah interpretasi hasil analisis pada data sosial ekonomi multivariat [8], [9], [10]. Meskipun demikian, penerapan PCA dalam penelitian sebelumnya umumnya hanya digunakan sebagai teknik pendukung visualisasi, tanpa penekanan pada keterkaitan antara hasil reduksi dimensi dan karakteristik *cluster* yang terbentuk, serta masih jarang dikombinasikan secara eksplisit dengan metode K-Medoids pada konteks wilayah Pulau Sumatera.

Berdasarkan uraian diatas, terdapat kesenjangan penelitian terkait pemetaan sosial ekonomi wilayah Pulau Sumatera secara komprehensif pada tingkat kabupaten/kota menggunakan metode K-Medoids yang robust, didukung oleh analisis PCA untuk memperjelas struktur *cluster*. Oleh karena itu, dilakukan penelitian untuk mengelompokkan wilayah Pulau Sumatra berdasarkan indikator sosial ekonomi dengan pendekatan K-Medoids menggunakan data resmi BPS tahun 2023, dengan penentuan jumlah *cluster* optimal serta visualisasi hasil menggunakan PCA. Penelitian ini diharapkan dapat memberikan gambaran yang lebih detail mengenai pola kesenjangan sosial ekonomi antarwilayah di Pulau Sumatra, sekaligus menjadi dasar pendukung dalam perumusan kebijakan pembangunan wilayah berbasis data.

2. Metode Penelitian

Penelitian ini menerapkan strategi kuantitatif dengan metode *unsupervised clustering* untuk mengelompokkan wilayah kabupaten/kota di Pulau Sumatera berdasarkan kemiripan indikator sosial ekonomi. Algoritma yang digunakan adalah K-Medoids, mengingat kemampuannya dalam menangani data berdimensi multivariat serta ketahanannya terhadap keberadaan pencilan (*outlier*), yang umum dijumpai pada data sosial ekonomi wilayah. Berbeda dengan K-Means yang menentukan pusat cluster berdasarkan nilai rata-rata, K-Medoids menentukan pusat *cluster* berdasarkan objek aktual (medoid), sehingga lebih stabil terhadap distribusi data yang tidak simetris [11], [12], [13].

Menggunakan data sekunder yang bersumber dari Badan Pusat Statistik (BPS) tahun 2023. Faktor yang dianalisis meliputi Indeks Pembangunan Manusia (IPM), Angka Harapan Hidup (AHH), Rata-rata Lama Sekolah (RLS), Tingkat Pengangguran Terbuka (TPT), serta kemiskinan. Setiap observasi merepresentasikan satu wilayah kabupaten/kota di Pulau Sumatera, sedangkan setiap atribut merepresentasikan indikator sosial ekonomi yang digunakan sebagai dasar pengelompokan.

Tahap prapemrosesan data dilakukan untuk menjamin bahwa dataset siap diproses dalam perhitungan jarak antar objek. Tahapan ini diawali dengan identifikasi dan penghapusan nilai kosong (*missing value*). Penghapusan dilakukan karena algoritma K-Medoids berbasis perhitungan jarak antar observasi, sehingga keberadaan nilai kosong dapat menyebabkan distorsi hasil klasterisasi. Hanya wilayah dengan data lengkap pada seluruh indikator yang dilibatkan dalam proses analisis. Selanjutnya, seluruh variabel dilakukan proses normalisasi menggunakan Min-Max Scaling guna setiap atribut berada pada rentang nilai yang sama, yaitu antara 0 dan 1. Normalisasi diperlukan untuk mencegah dominasi variabel tertentu akibat perbedaan skala pengukuran. Selain itu, indikator yang bersifat negatif terhadap kesejahteraan, yaitu TPT dan persentase penduduk miskin, ditransformasikan menggunakan pembalikan skala (*reverse transformation*). Transformasi ini bertujuan menyelaraskan arah interpretasi seluruh variabel, sehingga nilai yang lebih tinggi secara konsisten mencerminkan kondisi sosial ekonomi yang relatif lebih baik [11], [13], [14]. Sebagai ilustrasi proses pra pemrosesan, sebagian kecil data (lima sampel wilayah) ditampilkan untuk menunjukkan hasil normalisasi dan transformasi sebelum analisis dilakukan pada keseluruhan dataset.

Tabel 1. Dataset Sampel Wilayah

Kab/Kota (Pulau Sumatera)	IPM_norm	AHH_norm	RLS_norm	TPT_rev	Kemiskinan_rev
Kabupaten A	0.420474	0.219950	0.557971	0.476143	0.413265
Kabupaten B	0.244924	0.077307	0.381159	0.673956	0.239796
Kabupaten C	0.478003	0.309726	0.611594	0.26739	0.533163
Kabupaten D	0.328257	0.180549	0.37681	0.793241	0.477041
Kabupaten E	0.273266	0.043890	0.401449	0.609344	0.571429

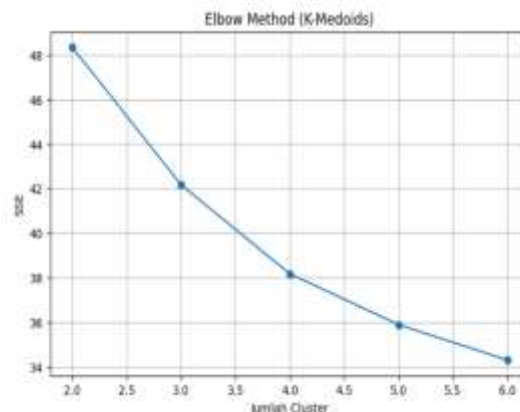
Jumlah cluster optimal ditentukan melalui penerapan gabungan metode Elbow dan Silhouette Score. Metode Elbow digunakan untuk mengevaluasi perubahan nilai *within-cluster dissimilarity* terhadap variasi jumlah *cluster*, sedangkan Silhouette Score digunakan untuk mengukur kualitas klasterisasi berdasarkan tingkat kohesi dan separasi antar *cluster*. Penggunaan kedua metode secara bersamaan bertujuan memperoleh jumlah *cluster* yang paling merepresentasikan struktur alami data [14], [15].

Setelah jumlah *cluster* optimal ditetapkan, proses klasterisasi dilakukan menggunakan algoritma K-Medoids dengan ukuran jarak Euclidean. Inisialisasi medoid awal dilakukan secara acak, kemudian algoritma dijalankan hingga mencapai kondisi konvergen, yaitu ketika tidak terdapat pertukaran medoid yang mampu menurunkan total jarak dalam *cluster*. Implementasi algoritma K-Medoids dilakukan menggunakan pustaka *pyclustering* pada bahasa pemrograman Python, sesuai dengan prosedur komputasi yang diterapkan dalam eksperimen.

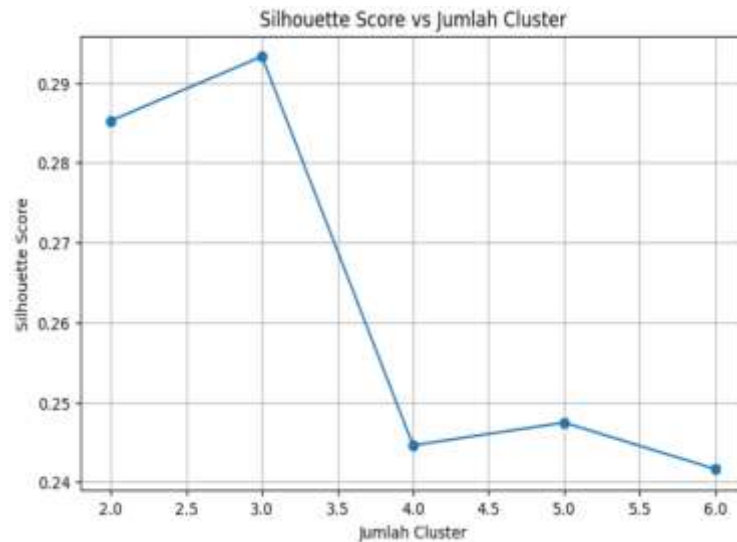
Untuk mendukung interpretasi hasil klasterisasi, dilakukan visualisasi menggunakan PCA untuk mereduksi dimensi data ke dalam dua komponen utama sehingga distribusi *cluster* dapat diamati secara visual. Perlu ditegaskan bahwa PCA dalam penelitian ini hanya digunakan sebagai alat bantu visualisasi dan tidak memengaruhi proses pembentukan *cluster*. Hasil klasterisasi selanjutnya disajikan dalam bentuk grafik dan tabel pemetaan wilayah, serta dianalisis pada tahap hasil dan diskusi.

3. Hasil dan Diskusi

Untuk memperoleh struktur pengelompokan yang representatif terhadap karakteristik data, digunakan kombinasi metode Elbow dan Silhouette Score dalam menentukan jumlah *cluster*. Grafik Elbow menunjukkan adanya titik perubahan kemiringan (*elbow point*) yang paling jelas pada jumlah tiga *cluster*, yang menandakan bahwa penurunan variasi dalam *cluster* mulai melambat setelah jumlah tersebut. Kondisi ini mengindikasikan bahwa penambahan *cluster* di atas tiga tidak memberikan peningkatan yang signifikan terhadap homogenitas dalam *cluster*. Hasil ini diperkuat oleh nilai Silhouette Score yang mencapai tingkat optimal pada jumlah *cluster* tiga, yang mencerminkan keseimbangan terbaik antara kohesi *intra-cluster* dan separasi antar *cluster*. Temuan ini konsisten dengan hasil penelitian klasterisasi sosial ekonomi dan kemiskinan di berbagai wilayah Pulau Sumatra yang umumnya menghasilkan struktur pengelompokan optimal dalam tiga kelompok utama [13], [15].



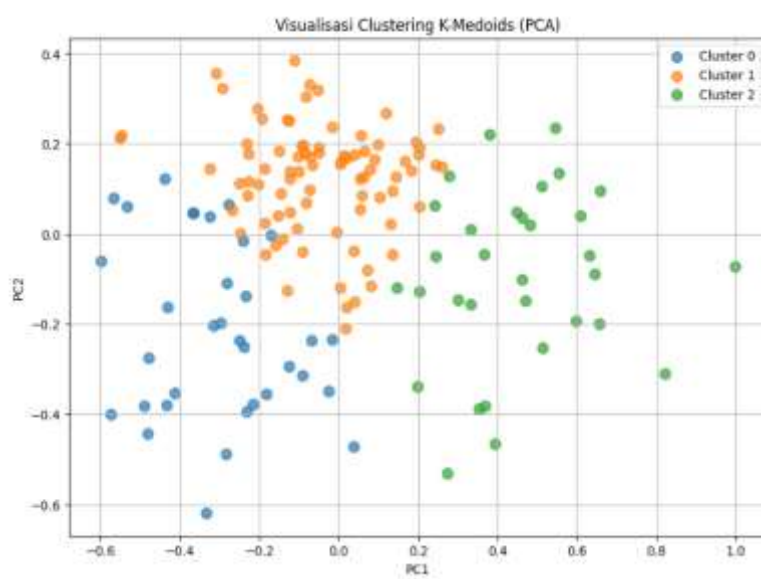
Gambar 1. Metode Elbow



Gambar 2. Silhouette Score

Berdasarkan hasil tersebut, algoritma K-Medoids kemudian diterapkan dengan jumlah *cluster* sebanyak tiga untuk mengelompokkan wilayah kabupaten/kota di Pulau Sumatra. Proses klasterisasi dilakukan berdasarkan kemiripan nilai indikator sosial ekonomi yang telah melalui tahap normalisasi dan transformasi. Setiap *cluster* direpresentasikan oleh satu medoid yang dipilih dari objek aktual dalam data, sehingga mencerminkan wilayah yang memiliki karakteristik paling sentral di dalam *cluster* tersebut. Hasil klasterisasi menunjukkan bahwa wilayah dengan pola indikator sosial ekonomi yang relatif serupa secara konsisten tergabung dalam *cluster* yang sama. Hal ini mengonfirmasi keandalan algoritma K-Medoids dalam mengelompokkan data sosial ekonomi yang bersifat heterogen dan memiliki potensi pencilan, sebagaimana dilaporkan dalam penelitian sebelumnya [11], [13].

Untuk mendukung interpretasi visual terhadap hasil klasterisasi, dilakukan reduksi dimensi PCA untuk memproyeksikan data multivariat ke dalam dua komponen utama yang menangkap variasi terbesar dalam data, sehingga struktur sebaran *cluster* dapat diamati secara visual. Hasil visualisasi menunjukkan bahwa ketiga *cluster* membentuk pola sebaran yang relatif terpisah, meskipun terdapat sejumlah wilayah yang berada pada area perbatasan antar *cluster*. Keberadaan wilayah pada area peralihan ini mencerminkan karakteristik indikator yang berada pada nilai tengah dan tidak menunjukkan perbedaan yang ekstrem terhadap *cluster* lain. Perlu ditegaskan bahwa PCA dalam penelitian ini digunakan semata-mata sebagai alat bantu visualisasi dan tidak memengaruhi proses pembentukan *cluster*, sehingga struktur pengelompokan sepenuhnya ditentukan oleh algoritma K-Medoids.



Tabel 2. Hasil Cluster

Cluster	Kabupaten/Kota
Cluster 0	Kabupaten Aceh Barat, Kabupaten Aceh Barat Daya, Kabupaten Aceh Jaya, Kabupaten Aceh Selatan, Kabupaten Aceh Singkil, Kabupaten Aceh Tengah, Kabupaten Aceh Timur, Kabupaten Aceh Utara, Kabupaten Bener Meriah, Kabupaten Gayo Lues, Kabupaten Nagan Raya, Kabupaten Pidie, Kabupaten Pidie Jaya, Kabupaten Simeulue, Kota Sabang, Kota Subulussalam, Kabupaten Nias Selatan, Kabupaten Nias Utara; Kabupaten Kepulauan Mentawai, Kabupaten Kepulauan Meranti, Kabupaten Tanjung Jabung Timur, Kabupaten Bengkulu Selatan, Kabupaten Kaur, Kabupaten Kepahiang, Kabupaten Lebong, Kabupaten Mukomuko, Kabupaten Rejang Lebong, Kabupaten Seluma, Kabupaten Lahat, Kabupaten Musi Rawas, Kabupaten Musi Rawas Utara, Kabupaten Ogan Ilir, Kabupaten Ogan Komering Ilir, Kabupaten Ogan Komering Ulu Selatan
Cluster 1	Kabupaten Aceh Tamiang, Kabupaten Aceh Tenggara, Kabupaten Bireuen, Kabupaten Asahan, Kabupaten Batu Bara, Kabupaten Dairi, Kabupaten Karo, Kabupaten Labuhanbatu, Kabupaten Labuhanbatu Selatan, Kabupaten Labuhanbatu Utara, Kabupaten Langkat, Kabupaten Mandailing Natal, Kabupaten Nias, Kabupaten Nias Barat, Kabupaten Padang Lawas, Kabupaten Padang Lawas Utara, Kabupaten Serdang Bedagai, Kabupaten Simalungun, Kabupaten Tapanuli Selatan, Kabupaten Tapanuli Tengah, Kabupaten Tapanuli Utara, Kota Gunungsitoli, Kota Tanjungbalai, Kabupaten Agam, Kabupaten Dharmasraya, Kabupaten Lima Puluh Kota, Kabupaten Padang Pariaman, Kabupaten Pasaman, Kabupaten Pasaman Barat, Kabupaten Pesisir Selatan, Kabupaten Sijunjung, Kabupaten Solok, Kabupaten Solok Selatan, Kabupaten Tanah Datar, Kabupaten Indragiri Hilir, Kabupaten Indragiri Hulu, Kabupaten Kampar, Kabupaten Kuantan Singingi, Kabupaten Pelalawan, Kabupaten Rokan Hilir, Kabupaten Rokan Hulu, Kabupaten Bintan, Kabupaten Karimun, Kabupaten Kepulauan Anambas, Kabupaten Lingga, Kabupaten Natuna, Kabupaten Batanghari, Kabupaten Bungo, Kabupaten Kerinci, Kabupaten Merangin, Kabupaten Muaro Jambi, Kabupaten Sarolangun, Kabupaten Tanjung Jabung Barat, Kabupaten Tebo, Kota Sungai Penuh, Kabupaten Bengkulu Tengah, Kabupaten Bengkulu Utara, Kabupaten Banyuasin, Kabupaten Empat Lawang, Kabupaten Muara Enim, Kabupaten Musi Banyuasin, Kabupaten Ogan Komering Ulu, Kabupaten Ogan Komering Ulu Timur, Kabupaten Penukal Abab Lematang Ilir (Pali), Kota Pagaralam, Kabupaten Bangka, Kabupaten Bangka Barat, Kabupaten Bangka Selatan, Kabupaten Bangka Tengah, Kabupaten Belitung, Kabupaten Belitung Timur, Kabupaten Lampung Barat, Kabupaten Lampung Selatan, Kabupaten Lampung Tengah, Kabupaten Lampung Timur, Kabupaten Lampung Utara, Kabupaten Mesuji, Kabupaten Pesawaran, Kabupaten Pesisir Barat, Kabupaten Pringsewu, Kabupaten Tanggamus, Kabupaten Tulang Bawang, Kabupaten Tulang Bawang Barat, Kabupaten Way Kanan
Cluster 2	Kabupaten Aceh Besar, Kota Banda Aceh, Kota Langsa, Kota Lhokseumawe, Kabupaten Deli Serdang, Kota Binjai, Kota Medan, Kota Padangsidimpuan, Kota Pematangsiantar, Kota Sibolga, Kota Tebing Tinggi, Kota Bukittinggi, Kota Padang, Kota Padang Panjang, Kota Pariaman, Kota Payakumbuh, Kota Sawahlunto, Kota Solok, Kabupaten Bengkalis, Kabupaten Siak, Kota Dumai, Kota Pekanbaru, Kota Batam, Kota Tanjungpinang, Kota Jambi, Kota Bengkulu, Kota Lubuk Linggau, Kota Palembang, Kota Prabumulih, Kota Pangkalpinang, Kota Bandar Lampung, Kota Metro

Berdasarkan perbandingan nilai relatif indikator sosial ekonomi yang telah dinormalisasi, masing-masing *cluster* menunjukkan pola karakteristik yang berbeda. Penamaan *cluster* 0, 1, dan 2 tidak merepresentasikan tingkatan normatif yang ditentukan sebelumnya, melainkan diberikan berdasarkan hasil pengelompokan algoritma K-Medoids dan urutan pola nilai indikator setelah proses klusterisasi.

Cluster 0 cenderung merepresentasikan wilayah dengan kondisi sosial ekonomi relatif lebih rendah. Pola ini ditunjukkan oleh nilai relatif IPM, AHH, dan RLS yang lebih kecil dibandingkan *cluster* lainnya, serta nilai TPT dan tingkat kemiskinan yang relatif lebih tinggi meskipun telah melalui transformasi pembalikan skala. Kombinasi nilai indikator tersebut menunjukkan bahwa wilayah dalam *cluster* 0 memiliki capaian kesejahteraan yang relatif tertinggal secara komparatif terhadap *cluster* lain dalam konteks Pulau Sumatra.

Cluster 1 menunjukkan karakteristik sosial ekonomi menengah, di mana nilai indikator berada pada kisaran moderat dan tidak menunjukkan kecenderungan ekstrem baik pada indikator kesejahteraan maupun indikator kerentanan. Wilayah dalam *cluster* ini umumnya memiliki capaian IPM, AHH, dan RLS yang lebih baik dibandingkan *cluster 0*, namun belum mencapai tingkat relatif tertinggi sebagaimana *cluster 2*. Pola ini mencerminkan kelompok wilayah dengan kondisi sosial ekonomi transisional, yang berada di antara kelompok dengan capaian rendah dan tinggi.

Sementara itu, *cluster 2* mencerminkan wilayah dengan kondisi sosial ekonomi relatif lebih baik. *Cluster* ini ditandai oleh nilai IPM, AHH, dan RLS yang relatif lebih tinggi, serta tingkat pengangguran dan kemiskinan yang relatif lebih rendah dibandingkan dua *cluster* lainnya. Konsistensi nilai indikator dalam *cluster* ini menunjukkan adanya pola kesejahteraan wilayah yang lebih mapan secara relatif dalam konteks pengelompokan yang dilakukan.

Interpretasi terhadap hasil klasterisasi ini bersifat deskriptif dan didasarkan pada perbandingan pola relatif antar *cluster*, tanpa dimaksudkan untuk menyimpulkan hubungan sebab-akibat antar indikator. Penamaan dan karakterisasi *cluster* dilakukan semata-mata untuk mempermudah interpretasi hasil klasterisasi dan tidak mengandung penilaian normatif. Secara keseluruhan, pola pengelompokan yang dihasilkan konsisten dengan temuan penelitian klasterisasi wilayah di Pulau Sumatra dan wilayah lain di Indonesia yang menggunakan pendekatan serupa[6], [11].

4. Kesimpulan

Penelitian ini berhasil menerapkan pendekatan klasterisasi untuk mengelompokkan kabupaten/kota di Pulau Sumatra berdasarkan indikator sosial ekonomi menggunakan algoritma K-Medoids. Penentuan jumlah *cluster* optimal melalui kombinasi metode Elbow dan Silhouette Score menghasilkan tiga *cluster* yang secara empiris mampu merepresentasikan struktur data secara memadai. Hasil ini menunjukkan bahwa pendekatan yang digunakan efektif dalam menangkap pola kemiripan dan perbedaan karakteristik sosial ekonomi antarwilayah. Pengelompokan yang dihasilkan memperlihatkan adanya variasi kondisi sosial ekonomi yang cukup jelas antar *cluster*. *Cluster* pertama merepresentasikan wilayah dengan capaian indikator kesejahteraan yang relatif lebih rendah, *cluster* kedua mencerminkan kondisi sosial ekonomi menengah, sedangkan *cluster* ketiga menggambarkan wilayah dengan kondisi sosial ekonomi yang relatif lebih baik. Pola ini konsisten dengan nilai relatif indikator IPM, AHH, dan RLS yang bersifat positif terhadap kesejahteraan, serta indikator TPT dan tingkat kemiskinan yang bersifat negatif, setelah melalui proses normalisasi dan penyesuaian arah interpretasi variabel. Dengan demikian, hasil klasterisasi dapat memberikan gambaran komparatif mengenai ketimpangan sosial ekonomi antarwilayah di Pulau Sumatra. Visualisasi hasil klasterisasi menggunakan Principal Component Analysis (PCA) berperan sebagai alat bantu untuk mempermudah interpretasi pola sebaran *cluster* dalam ruang dua dimensi. Meskipun PCA tidak digunakan dalam proses pembentukan *cluster*, visualisasi ini menunjukkan bahwa sebagian besar wilayah tergolong dalam kelompok yang relatif terpisah, dengan beberapa wilayah berada pada area peralihan antar *cluster*. Temuan ini mengindikasikan adanya heterogenitas kondisi sosial ekonomi yang bersifat gradual, bukan terpisah secara ekstrem. Secara praktis, hasil penelitian ini dapat dimanfaatkan sebagai dasar pemetaan kondisi sosial ekonomi wilayah yang bersifat deskriptif dan berbasis data. Informasi pengelompokan wilayah ini berpotensi mendukung perumusan kebijakan pembangunan yang lebih terarah dan berbasis karakteristik wilayah, terutama dalam konteks penanggulangan kemiskinan, pengurangan pengangguran, serta peningkatan kualitas sumber daya manusia. Selain itu, penggunaan algoritma K-Medoids yang lebih robust terhadap pencilan menjadikan hasil klasterisasi lebih stabil untuk data sosial ekonomi yang cenderung heterogen.

Referensi

- [1] D. Hastari, F. Nurunnisa, S. Winanda, and D. Dwi Aprillia, "Penerapan Algoritma K-Means dan K-Medoids untuk Mengelompokkan Data Negara Berdasarkan Faktor Sosial-Ekonomi dan Kesehatan," *SENTIMAS Semin. Nas. Penelit. dan Pengabd. Masy.*, pp. 274–281, 2023, [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas>
- [2] J. Homepage et al., "MALCOM: Indonesian Journal of Machine Learning and Computer Science Comparison of K-Means and K-Medoids on Poor Data Clustering in Indonesia Perbandingan K-Means dan K-Medoids Pada Pengelompokan Data Miskin di Indonesia," *Indones. J. Mach. Learn. Comput. Sci.*, vol. 2, no. 2, pp. 35–41, 2022.
- [3] I. M. Nur and A. Abdurakhman, "Analysis of Social Vulnerability in Java Island using K-Medoids Algorithm with Variation of Distance Measurements (Euclidean, Manhattan, Minkowski)," *Indones. J. Artif. Intell. Data Min.*, vol. 7, no. 2, p. 467, 2024, doi: 10.24014/ijaidm.v7i2.31111.
- [4] A. Eka Putra Haryanto, M. Ulfa Yanuar, D. Statistika Bisnis, and F. Vokasi, "Metode K-Means Clustering untuk Pengelompokan Kabupaten/Kota dalam Upaya Pengendalian Tingkat Inflasi di Pulau Jawa dan Sumatera K-Means Clustering Method for District/City Grouping in Effort to Control Inflation Rates in Java and Sumatera," vol. 1, pp. 29–42, 2022, [Online]. Available: <https://doi.org/10.21787/govstat.1.1.2022.29-42>

- [5] F. Alfiah, D. Al Farizi, and E. Widodo, "Analisis Clustering K-Medoids Berdasarkan Indikator Kemiskinan di Jawa Timur Tahun 2020," *J. Ilm. Sains*, vol. 22, no. April, pp. 1–7, 2020.
- [6] A. N. Damayanti and A. I. Mustakim, "Analisis Faktor yang Memengaruhi Indeks Pembangunan Manusia di Provinsi Jawa Barat Tahun 2023 Menggunakan Metode Analisis Komponen Utama dan Analisis Faktor," *J. Kaji. dan Penelit. Umum*, vol. 2, no. 6, pp. 14–23, 2024, doi: 10.47861/jkpu-nalanda.v2i6.1373.
- [7] H. Hafid, S. M. Meliyana, I. Muthahharah, and Z. Mar'ah, "Implementation K-Medoids Algorithm for Clustering Indonesian Provinces by Poverty and Economic Indicators," *Quant. Econ. Manag. Stud.*, vol. 6, no. 2, pp. 219–225, 2025, doi: 10.35877/454ri.qems3940.
- [8] dan R. R. M. R. Mubarak, A. T. J. Harjanto, "PENINGKATAN PERFORMA DBSCAN DENGAN REDUKSI DIMENSI PRINCIPAL COMPONENT ANALYSIS (PCA) DALAM KLASTERISASI TINGKAT KEMISKINAN DI INDONESIA," *J. Inform. Teknol. dan Sains*, vol. 7, pp. 1176–1184, 2025, doi: <https://doi.org/10.51401/jinteks.v7i3.6129>.
- [9] K. H. Izzuddin and A. W. Wijayanto, "Pemodelan Clustering Ward, K-Means, Diana, dan PAM dengan PCA untuk Karakterisasi Kemiskinan Indonesia Tahun 2021," *Komputika J. Sist. Komput.*, vol. 13, no. 1, pp. 41–53, 2024, doi: 10.34010/komputika.v13i1.10803.
- [10] U. Hasanah, M. R. Fauziah, A. Fitrianto, E. Erfiani, and L. M. R. D. Jumansyah, "Perbandingan Algoritma Klusterisasi dengan Principal Component Analysis pada Indikator Sosial Ekonomi Kesehatan Jawa Timur," *Techno.Com*, vol. 23, no. 4, pp. 847–863, 2024, doi: 10.62411/tc.v23i4.11534.
- [11] Bagus Candra Setiawan, Arum Nilawati, Rahmad Ferdian, Raihan Aditya Saputra, Yemima Putri Santoso, and Muhammad Riefky, "Analisis Cluster Hierarki Pada Tingkat Kemiskinan di Provinsi Sumatera Utara tahun 2023," *LogicLink*, vol. 2, no. 1, pp. 42–55, 2025, doi: 10.28918/logiclink.v2i1.10762.
- [12] A. Ikhran and Sani Mutia, "Penerapan Algoritma K-Medoids untuk Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Kemiskinan: Akses Perumahan dan Kesejahteraan," *Emerg. Stat. Data Sci. J.*, vol. 3, no. 2, pp. 583–596, 2025, doi: 10.20885/esds.vol3.iss.2.art13.
- [13] Retsya Lapiza, Syafriandi, N. Amalita, and D. Fitria, "Grouping The Districts in Sumatera Region Based on Economic Development Indicators Using K-Medoids and CLARA Methods," *UNP J. Stat. Data Sci.*, vol. 1, no. 1, pp. 16–22, 2023, doi: 10.24036/ujsds/vol1-iss1/13.
- [14] R. Setiana, D. W. Astuti, and M. Safir, "Pengelompokan Jumlah Pengangguran di Sumatera Utara Menggunakan Algoritma K-Means dan K-Medoids," *IT J.*, vol. 11, no. 2A, pp. 385–396, 2023.
- [15] L. Rhomaningtias, M. N. Kusharyadi, R. Westerdam Sean Jatindra, T. -, and M. Nasrudin, "Analisis Perbandingan Metode K-Medoid Dan Agglomerative Hierarchical Clustering Pada Data Konsumsi Rempah-Rempah Di Kabupaten / Kota," *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 3, 2025, doi: 10.23960/jitet.v13i3.7071.