Teknika 20 (3): 575-586

JURNAL TEKNIKA
ISSN: 0854-3143
e-ISSN: 2622-3481
Journal homepage: http://jurnal.polsri.ac.id/index.php/teknika
Journal Email: teknika@polsri.ac.id

Klasterisasi Wilayah Provinsi Lampung Berdasarkan Indikator Sosial Ekonomi Menggunakan Metode *K-Means Clustering*

Rosni*[1], Moh. Bayu Mulya Ali K.M.[2], Muklas Rivai[3], Nasrullah Bachtiar[4]

[*1,2,3,4]Program Studi Sains Aktuaria, Fakultas Sains, Institut Teknologi Sumatera, Lampung, Indonesia
*Email Penulis Korespondensi: rosni@at.itera.ac.id

Abstrak

Pembangunan sosial ekonomi yang berkeadilan masih menjadi tantangan akibat adanya ketimpangan antarwilayah, termasuk di Provinsi Lampung. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Provinsi Lampung berdasarkan indikator sosial ekonomi guna mengidentifikasi pola kesenjangan dan karakteristik wilayah. Metode yang digunakan adalah K-Means Clustering dengan pendekatan kuantitatif deskriptif. Data yang digunakan merupakan data sekunder dari Badan Pusat Statistik (BPS) Provinsi Lampung tahun 2024, meliputi variabel PDRB ADHK, persentase penduduk miskin, Umur Harapan Hidup (UHH), dan Indeks Ketimpangan Gender (IKG) pada 15 kabupaten/kota. Penentuan jumlah klaster optimal dilakukan menggunakan metode elbow dan dievaluasi dengan silhouette. Hasil evaluasi menunjukkan bahwa jumlah klaster dua (k = 2) merupakan konfigurasi yang paling optimal dengan rata-rata nilai Silhouette sebesar 0,4, yang mengindikasikan kualitas pengelompokan cukup baik. Klaster pertama menunjukkan wilayah dengan PDRB tinggi, tingkat kemiskinan rendah, UHH tinggi, dan IKG rendah, sedangkan klaster kedua didominasi oleh wilayah dengan PDRB rendah, kemiskinan tinggi, UHH rendah, dan IKG tinggi. Hasil penelitian menunjukkan adanya disparitas sosial ekonomi yang signifikan antarwilayah di Provinsi Lampung. Temuan ini diharapkan dapat menjadi dasar bagi pemerintah daerah dalam merumuskan kebijakan pembangunan yang lebih terarah, inklusif, dan berkeadilan.

Kata kunci—*K-Means clustering, kemiskinan, PDRB ADHK, sosial ekonomi*

Abstract

Equitable socio-economic development remains a major challenge due to regional disparities, including in Lampung Province. This study aims to classify regencies and cities in Lampung Province based on socio-economic indicators to identify patterns of inequality and regional characteristics. The research employs K-means clustering as a descriptive quantitative approach. Secondary data were obtained from the 2024 publications of the Badan Pusat Statistik (BPS) of Lampung Province, covering four main variables: Gross Regional Domestic Product at Constant Prices (GRDP), the percentage of the population living below the poverty line, life expectancy, and the Gender Inequality Index (GII) across 15 regencies/cities. The optimal number of clusters was determined using the elbow method and evaluated via the silhouette

Received: March 13, 2026; Online Available: September 8, 2026

method. Evaluation results indicate that a two-cluster configuration (k = 2) is optimal, with an average silhouette score of 0,4, suggesting reasonably good clustering quality. The optimal number of clusters was determined using the elbow method and validated with silhouette analysis, resulting in two clusters. The first cluster comprises regions with high GRDP, low poverty rates, high life expectancy, and low gender inequality. In contrast, the second cluster comprises regions with low GRDP, high poverty rates, low life expectancy, and high gender inequality. The findings indicate significant socio-economic disparities among regions in Lampung Province. These results are expected to support local governments in formulating more targeted, inclusive, and equitable development policies.

Keywords—*K-Means clustering, poverty, PDRB ADHK, socio-economics*

1. PENDAHULUAN

Pembangunan sosial ekonomi merupakan salah satu pilar penting dalam meningkatkan kualitas hidup masyarakat dan menjamin keberlanjutan pembangunan suatu negara [1]. Keberhasilan pembangunan umumnya diukur melalui indikator sosial ekonomi, seperti pertumbuhan ekonomi, tingkat kemiskinan, kesehatan, dan kesetaraan gender [2]. Namun, pada tingkat global, ketimpangan antarwilayah masih menjadi tantangan utama dalam mewujudkan pembangunan yang merata, baik dari sisi distribusi pendapatan maupun akses terhadap layanan dasar [3]. Oleh karena itu, analisis berbasis data diperlukan untuk memahami pola sosial ekonomi agar kebijakan pembangunan dapat dirumuskan secara lebih tepat sasaran [4].

Indonesia sebagai negara kepulauan dengan keragaman sosial, budaya, dan ekonomi menghadapi permasalahan yang kompleks dalam pemerataan pembangunan [5]. Data Badan Pusat Statistik menunjukkan adanya disparitas antarprovinsi maupun antar-kabupaten/kota dalam indikator PDRB ADHK, tingkat kemiskinan, kesehatan masyarakat, dan kesenjangan gender [6]. Ketimpangan tersebut mengindikasikan bahwa hasil pembangunan belum sepenuhnya dirasakan secara merata di seluruh wilayah [7]. Oleh sebab itu, analisis pengelompokan wilayah berdasarkan indikator sosial ekonomi menjadi penting untuk memahami pola ketidakmerataan pembangunan tersebut.

Provinsi Lampung sebagai salah satu provinsi di Pulau Sumatera masih menghadapi tantangan pembangunan yang signifikan. Beberapa kabupaten/kota menunjukkan pencapaian PDRB yang relatif tinggi, namun masih diiringi oleh tingginya tingkat kemiskinan dan ketimpangan gender [8]. Di sisi lain, meskipun indikator kesehatan seperti umur harapan hidup cenderung meningkat, pemerataan kualitas layanan kesehatan antarwilayah belum sepenuhnya optimal [9]. Perbedaan kondisi tersebut menunjukkan bahwa pembangunan di Provinsi Lampung belum berjalan secara seimbang.

Perbedaan kondisi sosial ekonomi antar kabupaten/kota di Provinsi Lampung memberikan dampak terhadap arah pembangunan daerah [10]. Tidak jarang wilayah dengan PDRB tinggi masih menghadapi permasalahan kesejahteraan akibat tingginya angka kemiskinan [8], sementara wilayah dengan PDRB relatif rendah justru menunjukkan capaian sosial yang cukup baik [11]. Kondisi ini menegaskan pentingnya pemetaan wilayah berdasarkan indikator sosial ekonomi agar karakteristik masing-masing daerah dapat diidentifikasi secara lebih jelas.

Salah satu metode yang banyak digunakan dalam analisis karakteristik wilayah adalah *K-Means Clustering*. Metode ini mampu mengelompokkan objek berdasarkan tingkat kemiripan karakteristik dengan meminimalkan variasi dalam kelompok (*within-cluster variation*) dan memaksimalkan perbedaan karakteristik antarkelompok, sehingga objek yang berada dalam satu klaster memiliki tingkat homogenitas yang lebih tinggi [12], [13]. Berbagai penelitian menunjukkan bahwa *K-Means Clustering* efektif digunakan untuk mengidentifikasi pola dan tipologi wilayah berdasarkan berbagai indikator pembangunan, termasuk indikator ekonomi, kemiskinan multidimensi, kesehatan, dan kesejahteraan masyarakat. Pendekatan ini dapat memberikan gambaran karakteristik masing-masing kelompok wilayah sehingga mendukung penyusunan kebijakan pembangunan yang lebih berbasis data [13], [14], [15]. Selain itu,

penelitian sebelumnya menegaskan bahwa penentuan jumlah klaster yang optimal perlu dilakukan melalui metode validasi internal seperti metode *Elbow* dan *Silhouette* untuk memastikan struktur klaster yang terbentuk memiliki tingkat kekompakan (*cohesion*) dan keterpisahan (*separation*) yang baik [16], [17].

Meskipun demikian, penelitian terdahulu masih memiliki beberapa keterbatasan. Pertama, sebagian besar penelitian masih menggunakan indikator pembangunan yang bersifat parsial, seperti indikator ekonomi atau kesejahteraan tertentu, sehingga belum sepenuhnya menggambarkan kondisi sosial ekonomi wilayah secara multidimensi [13], [15]. Kedua, sebagian penelitian lebih berfokus pada aspek komparasi algoritma maupun performa metode *clustering*, sementara kajian mengenai interpretasi karakteristik masing-masing klaster sebagai dasar penyusunan strategi pembangunan wilayah masih relatif terbatas [14], [18], [19]. Ketiga, penelitian mengenai pengelompokan kabupaten/kota di Provinsi Lampung yang mengintegrasikan indikator ekonomi, kemiskinan, kesehatan, dan kesetaraan gender dengan menggunakan data terbaru Badan Pusat Statistik tahun 2024 masih belum banyak ditemukan. Kondisi tersebut menunjukkan adanya research gap dalam penyusunan tipologi sosial ekonomi wilayah yang mampu merepresentasikan karakteristik pembangunan daerah secara lebih komprehensif.

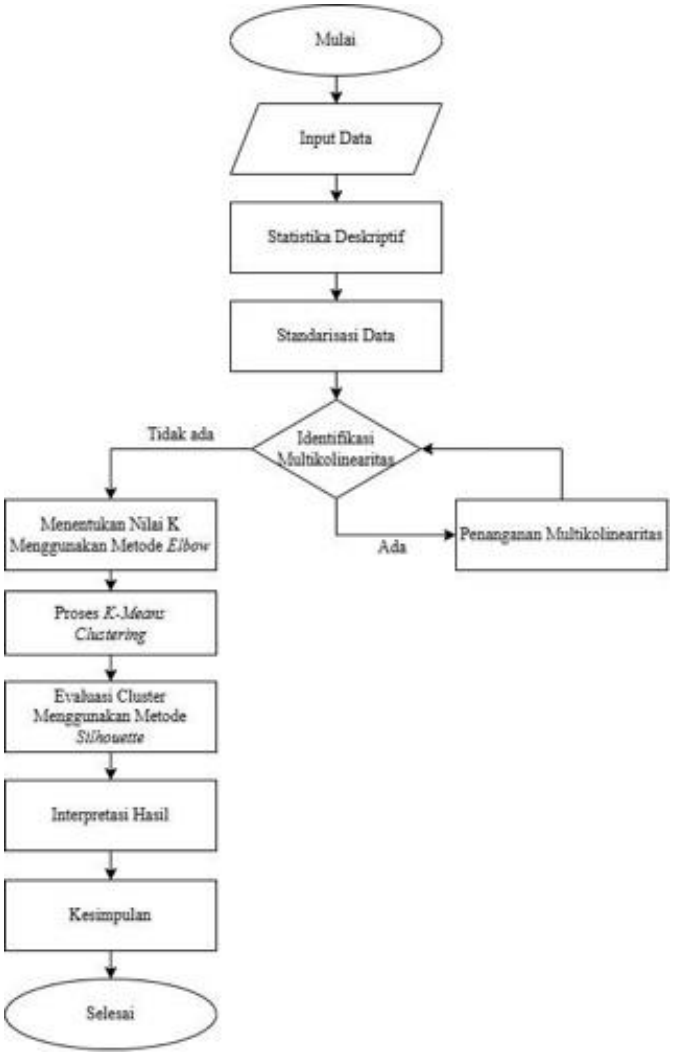
Berdasarkan uraian tersebut, penelitian ini bertujuan untuk mengelompokkan 15 kabupaten/kota di Provinsi Lampung menggunakan metode *K-Means Clustering* berdasarkan empat indikator utama, yaitu Produk Domestik Regional Bruto (PDRB) Atas Dasar Harga Konstan (ADHK), persentase penduduk miskin, Umur Harapan Hidup (UHH), dan Indeks Ketimpangan Gender (IKG). Jumlah klaster optimal ditentukan menggunakan *Elbow Method* dan divalidasi melalui *Silhouette Coefficient* sehingga hasil pengelompokan memiliki dasar objektif dan valid secara statistik. Kontribusi penelitian ini terletak pada penyusunan tipologi sosial ekonomi kabupaten/kota di Provinsi Lampung dengan mengintegrasikan dimensi ekonomi, kemiskinan, kesehatan, dan kesetaraan gender secara simultan menggunakan data terbaru tahun 2024. Hasil penelitian ini diharapkan dapat memberikan informasi empiris bagi pemerintah daerah dalam merumuskan kebijakan pembangunan yang lebih tepat sasaran, inklusif, dan berkelanjutan.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan metode clustering untuk mengelompokkan kabupaten/kota di Provinsi Lampung berdasarkan indikator sosial ekonomi. Data yang digunakan merupakan data sekunder yang diperoleh dari publikasi resmi Badan Pusat Statistik (BPS) Provinsi Lampung tahun 2024, dengan jumlah observasi sebanyak 15 kabupaten/kota. Variabel yang digunakan dalam penelitian ini terdiri dari empat indikator sosial ekonomi, yaitu Produk Domestik Regional Bruto Atas Dasar Harga Konstan (PDRB ADHK) (X_1), persentase penduduk miskin (X_2), Umur Harapan Hidup (UHH) (X_3), dan Indeks Ketimpangan Gender (IKG) (X_4).

2.1 Diagram Alir

Tahapan analisis data dalam penelitian ini disusun secara sistematis untuk menggambarkan proses pengolahan data hingga diperoleh hasil pengelompokan wilayah menggunakan metode *K-Means Clustering*. Diagram alir (*flowchart*) digunakan untuk memvisualisasikan tahapan penelitian. Adapun tahapan analisis data secara keseluruhan disajikan dalam bentuk diagram alir pada Gambar 1.



Gambar 1 Diagram alir penelitian

2.2 Prosedur Penelitian

Adapun langkah-langkah yang dilakukan dalam pengolahan data pada penelitian ini adalah sebagai berikut:

1. Melakukan pengumpulan data yang diperlukan untuk pengujian analisis. Kemudian, data ini disimpan dalam format Excel agar mudah diproses dan diolah lebih lanjut.
2. Mendeskripsikan karakteristik variabel independen Xi dengan statistika deskriptif. statistika deskriptif dilakukan untuk memberikan gambaran awal mengenai karakteristik masing-masing variabel.
3. Melakukan standarisasi data karena data memiliki perbedaan satuan dan skala, dilakukan standarisasi data menggunakan *z-score* agar setiap variabel memiliki kontribusi yang seimbang dalam proses klasterisasi. Transformasi *z-score* dilakukan menggunakan persamaan (1).

$$Z_i = \frac{X_i - \bar{X}}{S} \quad (1)$$

Keterangan:
Z_i: nilai standar variabel ke-i
X_i: nilai aktual variabel ke-i

$\bar{X}$: nilai rata-rata variabel
S: standar deviasi variabel

4. Mengidentifikasi multikolinearitas pada data menggunakan nilai VIF. Uji multikolinearitas dilakukan untuk memastikan bahwa antarvariabel tidak memiliki korelasi yang tinggi. Nilai VIF yang kurang dari 10 menunjukkan bahwa tidak terdapat masalah multikolinearitas, sehingga seluruh variabel layak digunakan dalam analisis klaster.
5. Menentukan jumlah *cluster* optimal menggunakan metode *elbow*. Penentuan jumlah klaster optimal dilakukan menggunakan metode *elbow* dengan menganalisis perubahan nilai *Sum of Squared Error* (SSE) pada berbagai jumlah klaster. Persamaan SSE yang digunakan seperti pada Formula (2).

$$SSE = \sum_{j=1}^{k} \sum_{z_i \in C_j} ||z_i - c_j||^2 \quad (2)$$

Keterangan:
z_i: objek ke-i dalam bentuk nilai hasil standardisasi *z-score*
c_j: *centroid* klaster ke-j

6. Setelah jumlah klaster ditentukan, melakukan proses *K-Means Clustering* yaitu mengelompokkan kabupaten/kota di Provinsi Lampung. Proses *K-Means Clustering* dengan tahapan inisialisasi *centroid* awal secara acak, penghitungan jarak menggunakan jarak *Euclidean*, pengelompokan data ke *centroid* terdekat, serta pembaruan *centroid* hingga mencapai kondisi konvergen. Formula (3) persamaan yang digunakan untuk menghitung jarak *Euclidean*.

$$d(z_i, c_j) = \sum_{k=1}^{p} (z_{ik} - c_{jk})^2 \quad (3)$$

Keterangan:
z_{ik}: nilai *z-score* objek ke-i pada variabel ke-k
c_{jk}: *centroid* klaster ke-j pada variabel ke-k
p : jumlah variabel penelitian (dalam penelitian ini p = 4)

7. Mengevaluasi hasil *cluster*. Hasil klasterisasi dievaluasi menggunakan metode *silhouette* untuk menilai kualitas pemisahan klaster yang terbentuk. Nilai koefisien *silhouette* digunakan untuk memastikan bahwa objek dalam satu klaster memiliki tingkat kemiripan yang tinggi dan terpisah dengan jelas dari klaster lainnya. Nilai *Silhouette* dihitung menggunakan persamaan (4).

$$S(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}} \quad (4)$$

Keterangan:
$S(i)$= nilai *silhouette* objek ke-i
$a(i)$= rata-rata jarak objek ke-i terhadap objek lain dalam klaster yang sama
$b(i)$= rata-rata jarak minimum objek ke-i terhadap objek pada klaster lain

8. Menginterpretasi hasil dengan meninjau rata-rata variabel pada tiap klaster serta membandingkan karakteristik antar klaster untuk memahami pola perbedaan wilayah.
9. Membuat kesimpulan.
10. Prosedur di atas dilakukan dengan menggunakan perangkat lunak R-Studio, yang memungkinkan pengolahan data statistik dan visualisasi secara efisien. Tahapan-tahapan ini memastikan bahwa pengelompokan wilayah dilakukan secara sistematis, objektif, dan dapat dipertanggungjawabkan secara ilmiah.

3. HASIL DAN PEMBAHASAN

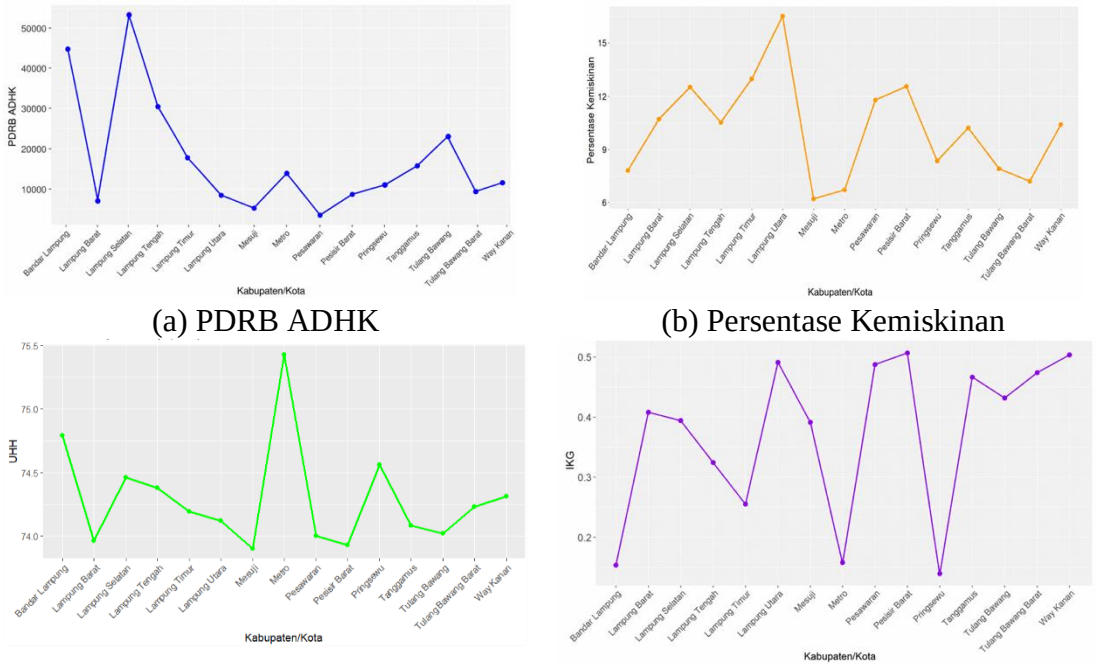
3.1 Statistika Dekriptif

Statistika deskriptif dilakukan untuk memberikan gambaran awal mengenai kondisi-kondisi sosial ekonomi di Provinsi Lampung tahun 2024, dan penyebaran datanya yang disajikan pada Tabel 1.

Tabel 1 Statistika Deskriptif

Variabel	Indikator			
	Minimum	Maximum	Median	Mean
PDRB ADHK (X_1)	3.418	55.830	12.398	18.853
Persentase Kemiskinan (X_2)	6,310	16,920	10,370	10,188
UHH (X_3)	73,900	75,430	74,190	74,290
IKG (X_4)	0,130	0,509	0,407	0,371

Berdasarkan Tabel 1, variabel PDRB ADHK memiliki nilai rata-rata dan maksimum tertinggi dibandingkan indikator sosial ekonomi lainnya. Sebaliknya, variabel Indeks Ketimpangan Gender (IKG) menunjukkan nilai rata-rata yang relatif lebih rendah, meskipun memiliki variasi antarwilayah yang cukup besar. Dilihat dari nilai median, lebih dari 50% kabupaten/kota di Provinsi Lampung memiliki nilai PDRB ADHK di bawah rata-rata, yang menunjukkan bahwa aktivitas ekonomi masih terkonsentrasi pada beberapa wilayah tertentu. Sementara itu, nilai Umur Harapan Hidup (UHH) cenderung relatif seragam antarwilayah, sedangkan persentase penduduk miskin menunjukkan perbedaan yang cukup nyata. Perbedaan nilai statistik ini mengindikasikan bahwa kondisi sosial ekonomi antarkabupaten/kota di Provinsi Lampung belum merata. Data untuk kekempat variabel X_1-X_4 disajikan secara visualisasi pada Gambar 2.



(a) PDRB ADHK (b) Persentase Kemiskinan

(c) UHH (d) IKG

Gambar 2 Grafik Persebaran Data

Gambar 2 bagian (a) menunjukkan kondisi PDRB ADHK, di mana terlihat adanya perbedaan cukup nyata antarwilayah. Beberapa daerah memiliki nilai PDRB yang tinggi,

sedangkan wilayah lainnya cenderung rendah. Hal ini mencerminkan adanya ketimpangan tingkat perekonomian antarkabupaten/kota. Bagian (b) menampilkan persentase penduduk miskin, grafik ini memperlihatkan bahwa jumlah penduduk miskin di setiap daerah tidak merata. Terdapat beberapa wilayah dengan persentase kemiskinan yang masih cukup besar, yang menunjukkan bahwa pembangunan ekonomi belum sepenuhnya dirasakan secara merata oleh seluruh masyarakat. Bagian (c) menggambarkan Umur Harapan Hidup (UHH) di setiap kabupaten/kota. Secara umum, nilai UHH tampak relatif stabil, meskipun terdapat sedikit perbedaan antarwilayah. Kondisi ini mengindikasikan bahwa akses terhadap layanan kesehatan di sebagian besar daerah sudah cukup baik, namun masih ada ruang untuk perbaikan di beberapa wilayah tertentu. Bagian (d) menampilkan Indeks Ketimpangan Gender (IKG). Grafik ini memperlihatkan variasi tingkat kesetaraan gender antarwilayah. Nilai IKG yang lebih tinggi menandakan bahwa kesetaraan antara laki-laki dan perempuan di wilayah tersebut lebih baik, sedangkan nilai yang lebih rendah mengindikasikan masih adanya kesenjangan dalam partisipasi maupun akses terhadap peluang pembangunan.

3.2 Standarisasi Data

Sebelum dilakukan uji multikolinearitas, data distandarisasi karena perbedaan satuan antarvariabel. Standarisasi dilakukan menggunakan metode *z-score* agar data berada pada skala yang sama. Hasil Standarisasi data dapat dilihat pada Tabel 2.

Tabel 2 Standarisasi Data

Kabupaten/Kota	X_1	X_2	X_3	X_4
Lampung Barat	-0.819	0.167	-0.817	0.264
Tanggamus	-0.403	0.031	-0.520	0.695
Lampung Selatan	1.013	0.809	0.418	0.169
⋮	⋮	⋮	⋮	⋮
Pesisir Barat	-0.964	0.833	-0.891	1.010
Kota Bandar Lampung	1.698	-0.957	1.234	-1.609
Kota Metro	-0.864	-1.157	2.815	-1.601

3.3 Identifikasi Multikolinearitas

Multikolinearitas merupakan kondisi ketika dua atau lebih variabel bebas memiliki hubungan yang sangat kuat atau saling berkorelasi tinggi. Hasil pemeriksaan multikolinearitas pada data yang digunakan dengan *Variance Inflation Factor* (VIF) dapat dilihat pada Tabel 3.

Tabel 3 Nilai VIF

Variabel	Nilai VIF	Keterangan
X_1	1,23	Tidak terjadi Multikolinearitas
X_2	1,35	Tidak terjadi Multikolinearitas
X_3	2,26	Tidak terjadi Multikolinearitas
X_4	2,70	Tidak terjadi Multikolinearitas

Berdasarkan Tabel 3, hasil pengujian multikolinearitas pada empat variabel yang digunakan menunjukkan bahwa seluruh variabel memiliki nilai VIF < 10. Dengan demikian, dapat disimpulkan bahwa tidak terdapat multikolinearitas pada variabel yang digunakan.

3.4 Penentuan Jumlah Cluster Optimal

Pada tahap ini, penentuan jumlah klaster yang optimal dilakukan sebelum proses pengelompokan data. Identifikasi jumlah klaster menggunakan metode elbow, yaitu pendekatan yang menentukan titik optimal berdasarkan pola penurunan nilai *Within-Cluster Sum of Squares* (WSS) atau *Sum of Squared Error* (SSE) pada setiap penambahan jumlah klaster. Jumlah klaster yang dipilih adalah titik ketika penurunan nilai SSE mulai melambat dan membentuk pola menyerupai siku (elbow), yang menunjukkan bahwa penambahan klaster selanjutnya tidak memberikan peningkatan yang signifikan terhadap kualitas pengelompokan. Setelah jumlah klaster optimal diperoleh, tahap berikutnya adalah menentukan nilai centroid awal pada masing-masing klaster, sebagaimana disajikan pada Tabel 4.

Tabel 3 Penentuan *Centroid* Awal Tiap *Cluster*

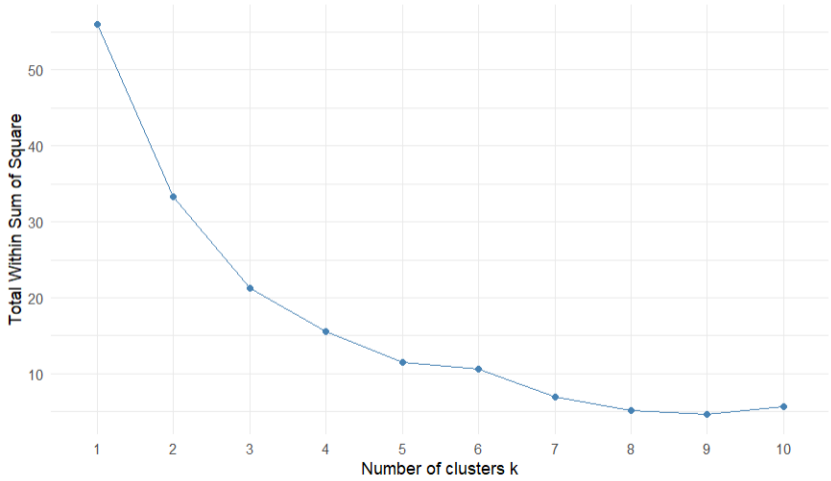
Centroid	X_1	X_2	X_3	X_4
Cluster 1	0.72	-0.14	0.85	-1.00
Cluster 2	-0.48	0.09	-0.56	0.67

Nilai *centroid* yang diperoleh pada Tabel 4 menunjukkan bahwa setiap klaster memiliki karakteristik yang berbeda dalam mengelompokkan kabupaten/kota di Provinsi Lampung. Setelah *centroid* masing-masing klaster ditentukan, langkah selanjutnya adalah menghitung nilai *Sum of Squared Error* (SSE) untuk setiap kabupaten/kota dalam klaster tersebut. Perhitungan SSE dilakukan untuk mengetahui seberapa jauh penyebaran data terhadap pusat klasternya. Hasil perhitungan nilai SSE tersebut disajikan pada Tabel 5.

Tabel 4 Nilai SSE Tiap Kabupaten/Kota dalam *Cluster*

Cluster	Kabupaten/Kota	Nilai SSE
Cluster 1	Lampung Selatan	0,34
	Lampung Timur	0,02
	Lampung Tengah	2,55
	Pringsewu	2,58
	Bandar Lampung	3,35
	Metro	5,09
Cluster 2	Lampung Barat	0,47
	Tanggamus	1,03
	Lampung Utara	0,29
	Way Kanan	2,65
	Tulang Bawang	2,46
	Pesawaran	1,41
	Mesuji	0,99
	Tulang Bawang Barat	2,11
	Pesisir Barat	7,77

Selanjutnya, nilai SSE dianalisis untuk berbagai jumlah klaster guna mengidentifikasi titik optimal pengelompokan. Visualisasi perubahan nilai SSE terhadap jumlah klaster ditampilkan pada Gambar 3.

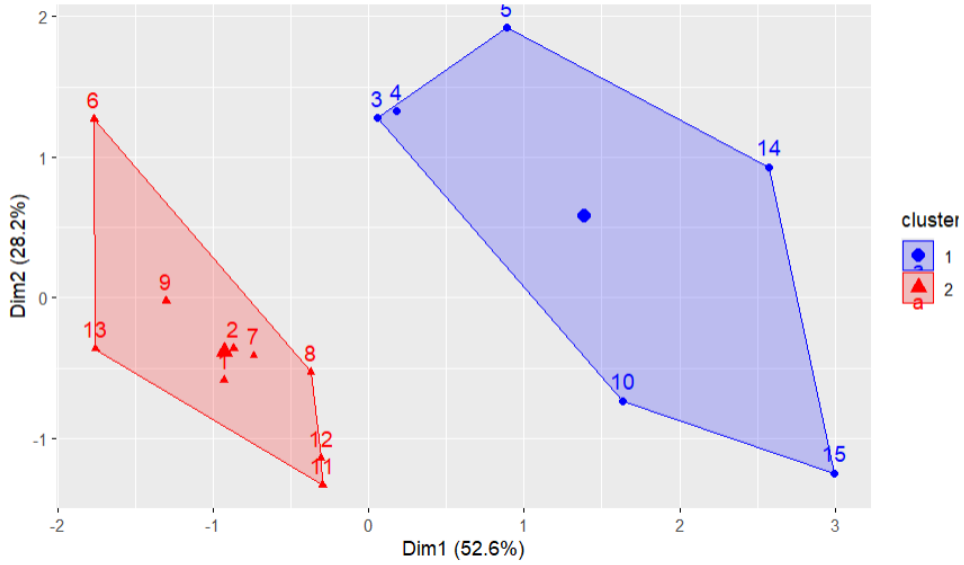


Gambar 3 Metode *Elbow*

Berdasarkan grafik metode elbow pada Gambar 3, terlihat adanya penurunan nilai SSE yang signifikan hingga jumlah klaster tertentu, kemudian diikuti oleh penurunan yang relatif landai. Pola tersebut menunjukkan titik siku (*elbow*) yang digunakan sebagai dasar dalam menentukan jumlah klaster optimal pada penelitian ini.

3.5 Hasil Klasterisasi

Setelah jumlah klaster optimal ditetapkan, tahap selanjutnya adalah melakukan proses *K-Means Clustering* untuk mengelompokkan kabupaten/kota di Provinsi Lampung berdasarkan kesamaan karakteristik indikator sosial ekonomi. Hasil pengelompokan divisualisasikan dalam bentuk cluster plot yang disajikan pada Gambar 4.



Gambar 4 Hasil *Cluster*

Gambar 4, visualisasi tersebut menunjukkan pemisahan wilayah ke dalam dua klaster berdasarkan kedekatan jarak terhadap *centroid* masing-masing klaster. Untuk memperjelas komposisi anggota pada setiap klaster, rincian jumlah dan daftar kabupaten/kota yang tergabung dalam masing-masing klaster disajikan pada tabel berikut.

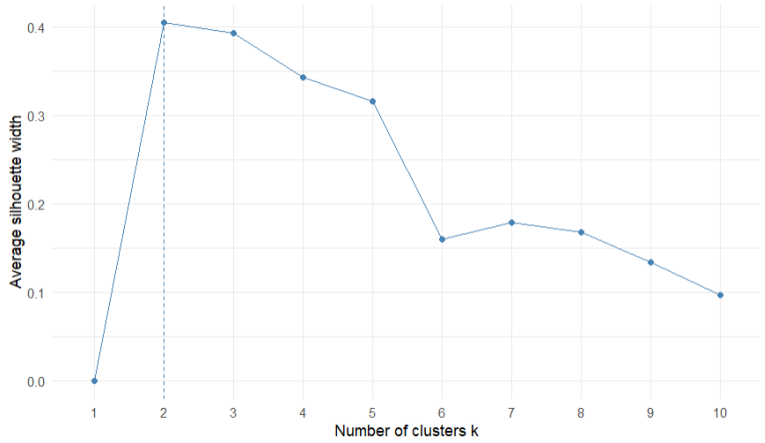
Tabel 5 Rincian Hasil *Cluster*

Cluster	Jumlah	Kabupaten/Kota
1	6	Lampung Selatan, Lampung Timur, Lampung Tengah, Pringsewu, Bandar Lampung dan Metro.
2	9	Lampung Barat, Tanggamus, Lampung Utara, Way Kanan, Tulang Bawang, Pesawaran, Mesuji, Tulang Bawang Barat, dan Pesisir Barat.

Berdasarkan hasil klasterisasi tersebut, terlihat bahwa kabupaten/kota di Provinsi Lampung terbagi ke dalam dua kelompok dengan karakteristik sosial ekonomi yang berbeda. Hasil ini menjadi dasar untuk analisis lebih lanjut mengenai karakteristik dan perbedaan kondisi sosial ekonomi antarklaster pada subbab pembahasan.

3.6 Evaluasi Klaster Silhouette

Evaluasi kualitas hasil klasterisasi dilakukan menggunakan metode *silhouette* untuk menilai tingkat kesesuaian objek data terhadap klaster yang terbentuk. Metode ini digunakan untuk mengukur seberapa baik setiap kabupaten/kota berada dalam klasternya dibandingkan dengan klaster lainnya. Hasil evaluasi *silhouette* divisualisasikan dalam bentuk grafik pada Gambar 5.



Gambar 5 Metode *Silhouette*

Berdasarkan Gambar 5, nilai koefisien *silhouette* yang diperoleh menunjukkan kualitas pengelompokan yang cukup baik. Nilai rata-rata *silhouette* yang positif mengindikasikan bahwa sebagian besar objek memiliki kedekatan yang lebih tinggi dengan klasternya sendiri dibandingkan dengan klaster lain. Dengan demikian, hasil pengelompokan wilayah menggunakan metode *K-Means* pada penelitian ini dapat dianggap valid dan layak digunakan sebagai dasar analisis lebih lanjut.

Berdasarkan hasil klasterisasi wilayah di Provinsi Lampung yang telah divalidasi melalui evaluasi klaster *silhouette*, pembahasan difokuskan pada perbedaan karakteristik sosial ekonomi antarklaster yang terbentuk. Analisis dilakukan dengan membandingkan nilai centroid masing-masing klaster pada indikator PDRB ADHK, persentase penduduk miskin, Umur Harapan Hidup (UHH), dan Indeks Ketimpangan Gender (IKG). Nilai centroid tersebut merepresentasikan kondisi rata-rata sosial ekonomi pada setiap klaster dan disajikan pada Tabel 7.

Tabel 6 Karakteristik *Centroid* Setiap *Cluster*

Cluster	X_1	X_2	X_3	X_4
1	Tinggi	Rendah	Tinggi	Rendah
2	Rendah	Tinggi	Rendah	Tinggi

Berdasarkan Tabel 7, klaster pertama menunjukkan karakteristik PDRB ADHK tinggi, tingkat kemiskinan rendah, umur harapan hidup tinggi, serta ketimpangan gender rendah. Kondisi ini mencerminkan wilayah dengan tingkat kesejahteraan sosial ekonomi yang relatif lebih baik dan pembangunan yang lebih merata. Daerah-daerah dalam klaster ini umumnya memiliki aktivitas ekonomi yang lebih berkembang, didukung oleh sektor perdagangan, jasa, dan industri yang relatif kuat. Tingginya nilai UHH pada klaster ini juga mengindikasikan kualitas kesehatan masyarakat yang lebih baik, yang kemungkinan dipengaruhi oleh akses layanan kesehatan dan fasilitas publik yang lebih memadai. Selain itu, rendahnya nilai IKG menunjukkan bahwa kesetaraan gender di wilayah ini cenderung lebih baik, sehingga partisipasi perempuan dalam kegiatan sosial dan ekonomi relatif lebih seimbang.

Sebaliknya, klaster kedua memiliki karakteristik PDRB ADHK rendah, tingkat kemiskinan tinggi, umur harapan hidup rendah, dan ketimpangan gender tinggi. Kondisi ini menunjukkan bahwa wilayah dalam klaster kedua masih menghadapi berbagai tantangan dalam pembangunan sosial ekonomi. Rendahnya PDRB ADHK mengindikasikan terbatasnya aktivitas ekonomi dan rendahnya produktivitas wilayah, yang berdampak langsung pada tingginya tingkat kemiskinan. Nilai UHH yang relatif lebih rendah juga mencerminkan bahwa kualitas kesehatan masyarakat di wilayah ini masih perlu ditingkatkan, yang kemungkinan dipengaruhi oleh keterbatasan akses terhadap layanan kesehatan dan infrastruktur dasar. Tingginya ketimpangan gender menunjukkan masih adanya kesenjangan peran dan kesempatan antara laki-laki dan perempuan dalam berbagai aspek pembangunan.

Perbedaan karakteristik antar klaster tersebut menegaskan adanya ketimpangan sosial ekonomi antarwilayah di Provinsi Lampung. Hasil ini menunjukkan bahwa pembangunan daerah belum berjalan secara merata, di mana sebagian wilayah telah mencapai tingkat kesejahteraan yang relatif baik, sementara wilayah lainnya masih tertinggal. Temuan ini sejalan dengan kondisi pembangunan wilayah yang bersifat terpusat pada daerah tertentu dan belum sepenuhnya menjangkau seluruh kabupaten/kota secara seimbang. Secara keseluruhan, hasil pembahasan ini memberikan gambaran yang jelas mengenai pola ketimpangan pembangunan di Provinsi Lampung. Informasi ini dapat dimanfaatkan sebagai dasar dalam perumusan kebijakan pembangunan daerah yang lebih terarah dan berbasis wilayah, khususnya dengan memprioritaskan peningkatan ekonomi, pengurangan kemiskinan, perbaikan kualitas kesehatan, serta penguatan kesetaraan gender pada wilayah-wilayah yang tergolong dalam klaster dengan kondisi sosial ekonomi yang lebih rendah.

4. KESIMPULAN

Berdasarkan hasil analisis menggunakan metode *K-Means Clustering* terhadap indikator sosial ekonomi di Provinsi Lampung, diperoleh jumlah klaster optimal sebanyak dua klaster yang ditentukan melalui metode *elbow* dan diperkuat dengan evaluasi *silhouette*. Hasil klasterisasi menunjukkan adanya perbedaan karakteristik sosial ekonomi yang jelas antarwilayah. Klaster pertama mencerminkan wilayah dengan tingkat kesejahteraan sosial ekonomi yang relatif lebih baik, ditandai oleh PDRB ADHK tinggi, tingkat kemiskinan rendah, umur harapan hidup tinggi, dan ketimpangan gender rendah. Sebaliknya, klaster kedua menunjukkan wilayah dengan kondisi sosial ekonomi yang masih tertinggal. Temuan ini menegaskan bahwa pembangunan di Provinsi Lampung belum merata antar kabupaten/kota dan masih terdapat ketimpangan sosial ekonomi yang signifikan.

5. SARAN

Berdasarkan hasil penelitian, disarankan agar penelitian selanjutnya menambahkan indikator sosial ekonomi lain, seperti tingkat pendidikan, tingkat pengangguran terbuka, dan Indeks Pembangunan Manusia, untuk memperoleh gambaran yang lebih komprehensif. Selain itu, cakupan wilayah penelitian dapat diperluas hingga tingkat kecamatan atau desa agar analisis ketimpangan lebih mendalam. Penelitian selanjutnya juga dapat mempertimbangkan penggunaan

metode klasterisasi lain sebagai pembanding guna menguji konsistensi hasil pengelompokan wilayah.

DAFTAR PUSTAKA

[1] Alriscki, D. G., and Fauzan, A., 2025, Peningkatan Distribusi Bantuan Sosial di Pangkalpinang dengan Pengelompokan Berbantuan Algoritma K-Means, *Statistika*, vol. 24, no. 2, hal. 191–199.

[2] Bahauddin, A., Fatmawati, A., and Sari, F. P., 2021, Analisis Clustering Provinsi di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means, *Jurnal Manajemen Informatika dan Sistem Informasi*, vol. 4, no. 1, hal. 1–11.

[3] Hariyadi, A. R., 2021, Dinamika Kebijakan Perencanaan Pembangunan Nasional Indonesia, *Jurnal Desentralisasi dan Kebijakan Publik*, vol. 2, no. 2, hal. 259–276.

[4] Rahayu, S. H. C., and P., 2023, Analisis Pengaruh Kesejahteraan Masyarakat terhadap Pertumbuhan Ekonomi, *Jurnal Informatika Ekonomi Bisnis*, vol. 5, no. 1, hal. 75–83.

[5] Santoso, E., and Anggraini, S., 2024, Disparitas Pembangunan Antarwilayah di Indonesia, *CERMIN: Jurnal Penelitian*, vol. 8, no. 2, hal. 355–367.

[6] Hanri, M., and Sholihah, N. K., 2025, Mengurangi Kemiskinan, Mengatasi Ketimpangan, *Labor Market Brief*, vol. 6, no. 1, hal. 1–9.

[7] Rinaldi, M., Pakpahan, G., Sihombing, L. V. F., Simanjuntak, R. P., and Manihuruk, S. D., 2025, Pengaruh Ketimpangan Ekonomi terhadap Tingkat Kemiskinan di Indonesia, *Journal of Sharia Economics Scholar*, vol. 3, no. 2, hal. 1–7.

[8] Pratama, M. I., and Nursari, A., 2025, Analisis Belanja Pendidikan, Belanja Kesehatan, Kemiskinan dan PDRB terhadap IPM Provinsi Lampung, *Jurnal Ekonomi dan Ekonomi Syariah*, vol. 8, no. 1, hal. 34–48.

[9] Fahrozi, A. A., Insani, F., Budianita, E., and Afrianty, I., 2023, Implementasi Algoritma K-Means dalam Menentukan Clustering pada Penilaian Kepuasan Pelanggan, *Indonesian Journal of Innovation Multidisipliner Research*, vol. 1, no. 4, hal. 474–492.

[10] Cahyani, A. R., Aida, N., and Yuliawan, D., 2024, Pengaruh Pengeluaran Pemerintah Sektor Kesehatan terhadap Indeks Pembangunan Manusia di Provinsi Lampung, *Jurnal Ekobistek*, vol. 13, no. 1, hal. 13–19.

[11] Hastari, D., Nurunnisa, F., Winanda, S., and Aprillia, D. D., 2023, Penerapan Algoritma K-Means untuk Mengelompokkan Data Negara Berdasarkan Faktor Sosial-Ekonomi, *SENTIMAS*, hal. 274–281.

[12] Dewi, K. S., Rivai, M., and Nurhasanah, A., 2025, Analisis Clustering Menggunakan Metode K-Means pada Data Produksi Tanaman Biofarmaka di Indonesia, *Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, vol. 6, no. 2, hal. 699–708.

[13] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, 2023, K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data, *Information Sciences*, vol. 622, hal. 178–210, doi: 10.1016/j.ins.2022.11.139.

[14] E. K. Korir, 2024, Comparative clustering and visualization of socioeconomic and health indicators: A case of Kenya, *Socio-Economic Planning Sciences*, vol. 95, no. 101961, doi: 10.1016/j.seps.2024.101961.

[15] C. Saraiva and J. Caiado, 2025, Global development patterns: A clustering analysis of economic, social and environmental indicators, *Sustainable Futures*, vol. 10, no. 100907, doi: 10.1016/j.sftr.2025.100907.

[16] A. M. Bagirov, R. M. Aliguliyev, and N. Sultanova, 2023, Finding compact and well-separated clusters: Clustering using silhouette coefficients, *Pattern Recognition*, vol. 135, no. 109144, doi: 10.1016/j.patcog.2022.109144.

[17] B. Yang and X. Xie, 2025, Analyzing and predicting global happiness index via integrated multilayer clustering and machine learning models, *PLOS ONE*, vol. 20, no. 4, doi: 10.1371/journal.pone.0322287.

[18] Fariha, D., Hartawan, D., Al-Abrour, M., Shafwan, M. R., and Ula, M., 2024, Pengelompokan Wilayah Indonesia Berdasarkan Indikator Sosio-Ekonomi Menggunakan Metode K-Means, SENASTIKA, hal. 1–6.

[19] Asyafiiyah, G. R. U., Widyastuti, A., and Andriani, F., 2025, Analisis Wilayah Prioritas Pembangunan di Provinsi Jawa Timur Berdasarkan Indikator Sosial Menggunakan Metode K-Means Clustering, *Journal of Computer Science and Artificial Intelligence*, vol. 5, no. 1, hal. 32–48.