

# K-Medoids Clustering Analysis of Regional Development in West Sumatra Based on Socioeconomic Indicators

Kayla Faradina, Fadhilah Fitri\*

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

\*Corresponding author: [fadhilahfitri@fmipa.unp.ac.id](mailto:fadhilahfitri@fmipa.unp.ac.id)

Submitted : 27 April 2026

Revised : 30 Mei 2026

Accepted : 31 Mei 2026

## ABSTRACT

*Regional development disparities among districts and cities in West Sumatra Province remain a persistent challenge, reflected in significant differences across economic, social, and employment indicators. This study aims to cluster 19 districts/cities in West Sumatra Province based on socioeconomic indicators using the K-Medoids clustering method. The variables include GRDP per capita, economic growth rate, GRDP percentage distribution, Human Development Index (HDI), poverty rate, and open unemployment rate, using 2024 data obtained from the Central Bureau of Statistics (BPS) of West Sumatra Province. The optimal number of clusters was determined using the Elbow method, resulting in three clusters. Cluster 1 consists of 12 districts characterized by the lowest average GRDP per capita and HDI, along with the highest poverty rate. Cluster 2 comprises only Kota Padang, which recorded the highest values across most indicators including GRDP per capita, economic growth rate, and HDI, yet also exhibited the highest open unemployment rate. Cluster 3 includes 6 cities with relatively high HDI and the lowest poverty rate among the three clusters. Cluster validation using the Davies-Bouldin Index (DBI) produced a value of 0.8341, indicating that the clustering results are optimal. The findings are expected to provide a reference for local governments and the Regional Development Planning Agency (Bappeda) of West Sumatra Province in formulating more targeted regional development policies based on the characteristics of each cluster.*

**Keywords:** *Davies Bouldin Index, K-Medoids Clustering, Socioeconomic.*



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

## I. PENDAHULUAN

Pembangunan wilayah merupakan suatu proses perubahan yang direncanakan secara sistematis dan berkelanjutan untuk meningkatkan kesejahteraan masyarakat secara merata di seluruh wilayah. Pembangunan wilayah yang optimal ditandai dengan pemerataan pertumbuhan ekonomi, peningkatan kualitas sumber daya manusia yang berkelanjutan, serta penurunan tingkat kemiskinan dan pengangguran secara konsisten. Pemerataan pembangunan ekonomi sangat penting karena manfaatnya tidak hanya dirasakan melalui peningkatan pendapatan, tetapi juga melalui perluasan akses terhadap pendidikan, kesehatan, dan infrastruktur bagi seluruh lapisan masyarakat (Siregar dkk., 2025).

Ketimpangan pembangunan antarwilayah masih menjadi tantangan serius yang dihadapi Indonesia, tercermin dari kesenjangan yang signifikan pada indikator sosial ekonomi seperti pendapatan, kualitas sumber daya manusia, tingkat kemiskinan, dan kondisi ketenagakerjaan antardaerah (Yenny dkk., 2025). Dalam kerangka kebijakan nasional, Peraturan Menteri Bappenas Nomor 4 Tahun 2023 tentang Rancangan Rencana Kerja Pemerintah Tahun 2024 menegaskan bahwa arah pembangunan difokuskan pada percepatan transformasi ekonomi yang inklusif dan berkelanjutan, dengan tujuan menekan kesenjangan antar kelompok masyarakat maupun antarwilayah. Hal ini mengindikasikan bahwa pemerataan pembangunan menjadi prioritas utama yang memerlukan perhatian dari seluruh tingkat pemerintahan, baik di tingkat provinsi maupun kabupaten/kota.

Provinsi Sumatera Barat sebagai salah satu provinsi di Pulau Sumatera memperlihatkan adanya ketimpangan pembangunan yang bersifat multidimensi. Dari aspek ekonomi, terdapat perbedaan yang cukup mencolok pada nilai Produk Domestik Regional Bruto (PDRB) per kapita antarwilayah, di mana Kota Bukittinggi mencatatkan nilai tertinggi sebesar Rp57,44 juta sementara Kabupaten Pesisir Selatan hanya sebesar Rp20,67 juta, dengan selisih yang mencapai hampir tiga kali lipat. Dari dimensi sosial, Kota Padang memiliki Indeks Pembangunan Manusia (IPM) tertinggi sebesar 84,38, jauh di atas Kabupaten Kepulauan Mentawai yang hanya mencapai 66,67. Sementara dari dimensi kemiskinan,

Kabupaten Pesisir Selatan mencatatkan persentase penduduk miskin tertinggi sebesar 13,89 persen, berbanding jauh dengan Kota Padang Panjang yang hanya 2,33 persen (Badan Pusat Statistik Provinsi Sumatera Barat, 2025). Kondisi ini mempertegas bahwa pembangunan wilayah di Provinsi Sumatera Barat masih menunjukkan ketimpangan yang signifikan, tidak hanya dari dimensi ekonomi, tetapi juga dari dimensi sosial dan kemiskinan antarwilayah.

Ketimpangan yang bersifat multidimensi tersebut menuntut pendekatan analisis yang mampu memotret kondisi pembangunan wilayah secara menyeluruh dan simultan dari berbagai dimensi sekaligus, bukan hanya berdasarkan satu indikator secara parsial (Setiyanto dkk., 2025). Selama ini analisis kondisi pembangunan wilayah umumnya dilakukan secara deskriptif dan terpisah per indikator, sehingga belum mampu mengelompokkan kabupaten/kota berdasarkan kesamaan karakteristik sosial ekonomi secara komprehensif dan berbasis data. Akibatnya, pemerintah daerah menghadapi keterbatasan dalam mengidentifikasi wilayah-wilayah yang memiliki kondisi pembangunan serupa dan memerlukan pendekatan kebijakan yang sejenis.

Salah satu pendekatan yang relevan untuk menjawab kebutuhan tersebut adalah analisis kluster, yaitu metode pengelompokan objek berdasarkan kesamaan karakteristik dari sejumlah indikator secara bersamaan, sehingga diperoleh pemetaan wilayah yang lebih terukur dan berbasis data (Setiyanto dkk., 2025). Dari berbagai metode kluster yang tersedia, *K-Medoids* menjadi pilihan dalam penelitian ini karena pusat klasternya berupa titik data nyata yang disebut *medoid*, berbeda dengan *K-Means* yang menetapkan nilai rata-rata sebagai pusat kluster dan rentan terhadap pengaruh nilai ekstrem (Sucahyo dkk., 2025). Ketahanan terhadap *outlier* menjadi pertimbangan penting mengingat adanya variasi yang ekstrem pada data indikator sosial ekonomi kabupaten/kota di Provinsi Sumatera Barat. Relevansi metode ini juga didukung oleh penelitian Sari dkk. (2025) yang menerapkan *K-Medoids* dalam mengelompokkan wilayah-wilayah di Pulau Sumatera berdasarkan indikator sosial ekonomi, dengan hasil pengelompokan yang terbukti optimal dan dapat dimanfaatkan sebagai acuan dalam perencanaan kebijakan pembangunan.

Penelitian ini bertujuan untuk mengelompokkan 19 kabupaten/kota di Provinsi Sumatera Barat berdasarkan kemiripan karakteristik pembangunan wilayah menggunakan enam indikator sosial ekonomi, yaitu PDRB per kapita, laju pertumbuhan ekonomi, distribusi persentase PDRB, IPM, persentase penduduk miskin, dan Tingkat Pengangguran Terbuka (TPT), dengan data tahun 2024 yang bersumber dari BPS Provinsi Sumatera Barat. Hasil pengelompokan yang diperoleh diharapkan mampu memberikan gambaran yang komprehensif mengenai karakteristik pembangunan tiap wilayah, sehingga dapat dijadikan landasan dalam penyusunan kebijakan pembangunan yang lebih terarah dan sesuai dengan kebutuhan masing-masing kabupaten/kota oleh pemerintah daerah Provinsi Sumatera Barat.

## II. METODE PENELITIAN

### A. Sumber data dan Variabel Penelitian

Penelitian ini merupakan penelitian terapan yang memanfaatkan data sekunder dari publikasi Sumatera Barat Dalam Angka 2025 serta laman resmi Badan Pusat Statistik Provinsi Sumatera Barat. Unit amatan yang digunakan adalah data tahun 2024 yang mencakup 19 kabupaten/kota di Provinsi Sumatera Barat. Dalam Penelitian ini terdapat 6 variabel yang merupakan indikator sosial ekonomi. Adapun 6 variabel tersebut yaitu PDRB per kapita ( $X_1$ ), Laju pertumbuhan ekonomi ( $X_2$ ), Distribusi persentase PDRB ( $X_3$ ), Indeks Pembangunan Manusia ( $X_4$ ), Persentase penduduk miskin ( $X_5$ ), Tingkat Pengangguran Terbuka ( $X_6$ ).

### B. Tahapan Analisis Data

Seluruh proses pengolahan data dalam penelitian ini dilaksanakan dengan bantuan *software* RStudio. Adapun tahapan analisis yang diterapkan adalah sebagai berikut.

#### 1. Standardisasi Data

Sebelum analisis kluster dilakukan, standardisasi data terlebih dahulu diterapkan pada seluruh variabel mengingat setiap variabel memiliki satuan pengukuran yang tidak seragam. Kesenjangan satuan yang terlalu besar berpotensi menghasilkan perhitungan jarak yang tidak akurat dalam analisis kluster (Hair dkk., 2019). Proses standardisasi data dilakukan menggunakan metode *z-score* dengan rumus sebagai berikut.

$$Z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (1)$$

Dimana  $x_{ij}$  adalah nilai objek ke- $i$  pada variabel ke- $j$ ,  $\mu_j$  adalah rata-rata variabel ke- $j$ , dan  $\sigma_j$  adalah standar deviasi variabel ke- $j$ .

#### 2. Penentuan Jumlah Kluster Optimal dengan Metode *Elbow*

Metode *Elbow* digunakan untuk menentukan jumlah kluster terbaik dengan mengamati perubahan nilai kesalahan pada setiap penambahan jumlah kluster. Tujuan dari metode *Elbow* adalah memilih nilai *kyang* tidak terlalu besar

namun tetap menghasilkan nilai *Sum of Squared Errors* (SSE) yang kecil (Maori, 2023). Seiring meningkatnya jumlah kluster, nilai SSE cenderung menurun. Oleh karena itu, jumlah kluster optimal ditentukan pada titik ketika penurunan SSE mulai melandai dan membentuk pola menyerupai siku (*elbow*), karena penambahan kluster setelah titik tersebut tidak lagi memberikan penurunan SSE yang berarti (Han dkk., 2012). Dalam analisis kluster, SSE juga dikenal sebagai *Within Cluster Sum of Squares* (WSS), yang menggambarkan total variasi dalam masing-masing kluster. Adapun rumus WSS adalah sebagai berikut.

$$WSS = \sum_{k=1}^k \sum_{i \in C_k} (x_i - m_k)^2 \quad (2)$$

Dimana  $C_k$  adalah himpunan objek dalam kluster ke- $k$ ,  $x_i$  adalah nilai objek ke- $i$ ,  $m_k$  adalah nilai *medoid* kluster ke- $k$ , dan  $K$  adalah jumlah kluster.

### 3. Analisis Kluster *K-Medoids*

*K-Medoids* merupakan pengembangan dari metode *K-Means* yang dirancang untuk mengatasi kelemahan *K-Means* dalam menghadapi *outlier*. Perbedaan mendasar antara keduanya terletak pada cara penentuan pusat kluster; *K-Medoids* memanfaatkan objek perwakilan (*medoid*) sebagai pusat kluster, sedangkan *K-Means* menggunakan nilai rata-rata (*mean*) dari seluruh anggota kluster (Hoerunnisa dkk., 2024). Adapun tahapan analisis kluster *K-Medoids* adalah sebagai berikut.

- a. Bangkitkan pusat kluster (*medoid*) secara acak sebanyak *k* optimum yang telah ditentukan.
- b. Kelompokkan setiap objek ke *medoid* terdekat berdasarkan jarak *Euclidean* terkecil. Jarak *Euclidean* dihitung menggunakan rumus:

$$d_{i,m} = \sqrt{\sum_{j=1}^p (x_{ij} - m_{kj})^2}, i = 1,2,3, \dots, n; \text{ dan } k = 1,2,3, \dots, K \quad (3)$$

Dimana  $d_{i,k}$  adalah jarak antara objek ke- $i$  dengan *medoid* kluster ke- $k$ ,  $x_{ij}$  adalah nilai objek ke- $i$  pada variabel ke- $j$ ,  $m_{kj}$  adalah nilai *medoid* kluster ke- $k$  pada variabel ke- $j$ ,  $p$  adalah banyaknya variabel,  $n$  adalah banyaknya objek, dan  $K$  adalah jumlah kluster.

- c. Hitung nilai *Total Cost* menggunakan rumus:

$$Total\ Cost = \sum_{k=1}^K \sum_{i=1}^n d(i, k) \quad (4)$$

Dimana  $d(i, k)$  adalah jarak objek ke- $i$  ke *medoid* kluster ke- $k$ .

- d. Lakukan pemilihan *medoid* baru dari setiap kluster, kemudian hitung kembali jarak setiap objek ke *medoid* baru dan nilai *Total Cost* yang baru. Menggunakan rumus:

$$S = Total\ Cost\ baru - Total\ Cost\ lama \quad (5)$$

- e. Ulangi langkah (b) hingga (d) hingga tidak terjadi perubahan anggota *medoid*, yaitu kondisi ketika nilai  $S_{total\ cost} \geq 0$  yang berarti penggantian *medoid* tidak lagi menghasilkan *Total Cost* penurunan.

### 4. Validasi Hasil Kluster

Penilaian kualitas hasil analisis kluster dilakukan menggunakan *Davies-Bouldin Index* (DBI). DBI merupakan metode evaluasi kluster yang bekerja dengan cara memaksimalkan jarak antarkluster sekaligus meminimalkan jarak antar titik data dalam satu kluster (Kapita dkk., 2022). Hasil pengelompokan dinyatakan optimal apabila nilai DBI kecil atau mendekati nol. DBI dihitung menggunakan rumus:

$$DBI = \frac{1}{k} \sum_{j=1}^k \sum_{i=1}^k \max_{i \neq j} (R_{ij}) \quad (6)$$

Dimana nilai  $R_{ij}$  diperoleh melalui:

$$R_{ij} = \frac{SSW_k + SSW_l}{SSB_{kl}}$$

Dimana:

$$SSW_k = \frac{1}{n_k} \sum_{i \in C_k} d_{i,k}$$

$$SSW_k = \frac{1}{n_l} \sum_{i \in C_l} d_{i,l}$$

$$SSB_{k,l} = d(m_k, m_l)$$

Dengan  $R_{kl}$  adalah rasio antara kluster  $k$  dan kluster  $l$ ,  $SSW_k$  dan  $SSW_l$  adalah *sum of square within kluster*  $k$  dan  $l$ ,  $SSB_{kl}$  adalah *sum of square between* kluster  $k$  dan  $l$ ,  $d_{i,k}$  adalah jarak objek  $i$  ke  $i$  ke *medoid* ke- $k$ ,  $d(m_k, m_l)$  adalah jarak antar *medoid* kluster  $k$  dan  $l$ , dan  $K$  adalah jumlah kluster.

### III. HASIL DAN PEMBAHASAN

#### A. Statistika Deskriptif

Gambaran umum kondisi pembangunan wilayah 19 kabupaten/kota di Provinsi Sumatera Barat tahun 2024 disajikan pada Tabel 1.

**Tabel 1.** Statistik Deskriptif

Variabel	Minimum	Mean	Maksimum
PDRB per kapita ( $X_1$ )	20.670	35.780	57.435
Laju pertumbuhan ekonomi ( $X_2$ )	1,43	5,26	25,34
Distribusi persentase PDRB ( $X_3$ )	2,11	2,74	3,72
Indeks Pembangunan Manusia ( $X_4$ )	66,67	75,82	84,38
Persentase penduduk miskin ( $X_5$ )	2,33	6,02	13,89
Tingkat Pengangguran Terbuka ( $X_6$ )	13,89	5,01	9,88

Berdasarkan Tabel 1, terlihat bahwa seluruh variabel menunjukkan variasi yang cukup besar antarwilayah. PDRB per kapita memiliki rentang yang sangat lebar, dari Rp20,67 juta hingga Rp57,44 juta, yang mencerminkan ketimpangan kapasitas ekonomi antarwilayah yang signifikan. Laju pertumbuhan ekonomi juga menunjukkan variasi ekstrem dengan nilai minimum 1,43 persen dan maksimum 25,34 persen. IPM berkisar antara 66,67 hingga 84,38, sedangkan persentase penduduk miskin bervariasi dari 2,33 persen hingga 13,89 persen. Adanya variasi ekstrem pada beberapa variabel, terutama pada laju pertumbuhan ekonomi ( $X_2$ ) dan Tingkat Pengangguran Terbuka ( $X_6$ ), mengindikasikan keberadaan *outlier* yang menjadi pertimbangan utama dalam pemilihan metode *K-Medoids* sebagai metode yang *robust* terhadap *outlier*.

#### B. Standardisasi Data

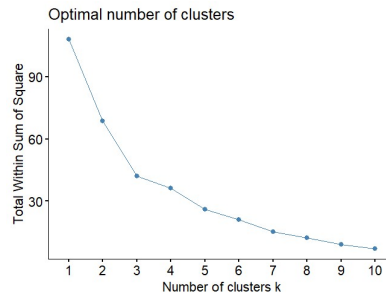
Sebelum dilakukan analisis kluster, seluruh data distandarisasi menggunakan metode *z-score* karena keenam variabel memiliki satuan pengukuran yang berbeda-beda. Hasil standarisasi menghasilkan data dengan rata-rata 0 dan standar deviasi 1 pada setiap variabel, sehingga perbedaan skala antarindikator tidak memengaruhi perhitungan jarak dalam analisis kluster. Hasil Standarisasi disajikan pada Tabel 2.

**Tabel 2.** Data Terstandarisasi

Kabupaten/Kota	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
1	0,1352	-0,6077	-0,6353	-1,9978	-0,0213	-2,0707
2	-1,4236	-0,8893	0,0953	-0,6446	3,2775	0,0271
⋮	⋮	⋮	⋮	⋮	⋮	⋮
18	-0,0132	1,0350	-0,4550	1,0773	-0,8085	-0,0830
19	0,6531	1,3401	-0,6033	1,0184	-0,3461	0,1778

#### C. Penentuan Jumlah Kluster

Penentuan jumlah kluster optimal dilakukan menggunakan metode *Elbow* berdasarkan nilai SSE. Hasil metode *Elbow* ditunjukkan pada Gambar 1.



Gambar 1. Plot Metode *Elbow*

Berdasarkan Gambar 1, terlihat bahwa penurunan nilai WSS yang paling signifikan terjadi pada k=1 hingga k=3, sedangkan setelah k=3 grafik mulai melandai dan tidak terjadi penurunan yang curam. Titik siku (*elbow*) terbentuk pada k=3, sehingga jumlah kluster optimal yang digunakan dalam penelitian ini adalah 3 kluster.

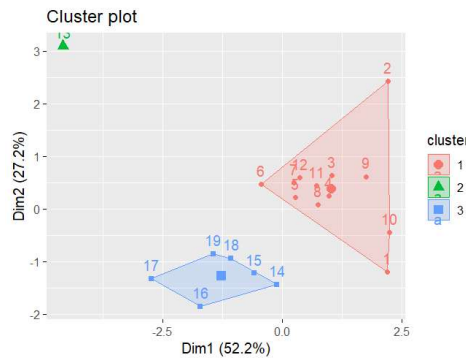
**D. Analisis Kluster *K-Medoids***

Berdasarkan jumlah kluster optimal sebanyak 3 kluster, selanjutnya dilakukan analisis kluster *K-Medoids* menggunakan data yang telah distandarisasi. Hasil pengelompokan 19 kabupaten/kota di Provinsi Sumatera Barat disajikan pada Tabel 3.

Tabel 3. Hasil Pengelompokan Analisis Kluster *K-Medoids*

Kluster	Banyak Anggota (Kabupaten/Kota)	Nama Kabupaten/Kota
1	12	Kepulauan Mentawai, Pesisir Selatan, Solok, Sijunjung, Tanah Datar, Padang Pariaman, Agam, Lima Puluh Kota, Pasaman, Solok Selatan, Dharmasraya, Pasaman Barat
2	1	Kota Padang
3	6	Kota Solok, Kota Sawah Lunto, Kota Padang Panjang, Kota Bukittinggi, Kota Payakumbuh, Kota Pariaman

Hasil pengelompokan tersebut divisualisasikan pada *cluster plot* berikut.



Gambar 2. Plot Hasil Analisis Kluster *K-Medoids*

Berdasarkan Gambar 2, ketiga kluster yang terbentuk dapat terpisah dengan cukup jelas. Kluster 1 (merah) mengelompok di sisi kanan atas bidang koordinat, Kluster 3 (biru) berada di sisi kiri bawah, sementara Kluster 2 (hijau) yang hanya terdiri dari Kota Padang (nomor 13) terlihat terpisah jauh di bagian kiri atas sebagai titik pencilan yang berdiri sendiri. Dimensi 1 (Dim1) menjelaskan 52,2% keragaman data, sedangkan Dimensi 2 (Dim2) menjelaskan 27,2%, sehingga secara kumulatif kedua dimensi mampu merepresentasikan 79,4% keragaman total data.

Karakteristik masing-masing kluster berdasarkan nilai rata-rata setiap variabel disajikan pada Tabel 4.

**Tabel 4.** Karakteristik Klaster Berdasarkan Nilai Rata-rata

Klaster	1	2	3
X1	29.205,08	54.090	45.879,17
X2	2,530833	3,42	3,041667
X3	5,133333	25,34	2,175
X4	72,9225	84,38	80,2
X5	6,946667	7	4,006667
X6	4,695833	9,88	4,836667

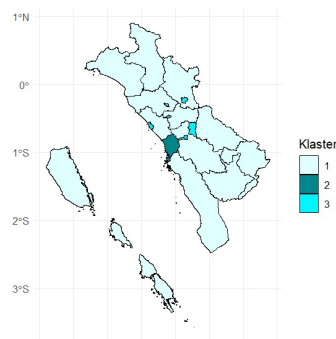
Berdasarkan Tabel 4, masing-masing klaster menunjukkan karakteristik yang berbeda berdasarkan nilai rata-rata indikator sosial ekonomi. Perbedaan tersebut mencerminkan variasi tingkat pembangunan wilayah yang dapat digunakan untuk mengidentifikasi pola kesenjangan antar daerah.

Klaster 1 didominasi oleh kabupaten yang memiliki rata-rata PDRB per kapita terendah sebesar Rp29.205.080, IPM terendah sebesar 72,92, serta persentase penduduk miskin tertinggi sebesar 6,95 persen di antara ketiga klaster. Meskipun distribusi persentase PDRB rata-rata mencapai 5,13 persen, laju pertumbuhan ekonomi rata-rata sebesar 2,53 persen belum mampu mendorong peningkatan kualitas SDM dan penurunan kemiskinan secara signifikan. Seluruh anggota klaster ini merupakan wilayah kabupaten yang umumnya memiliki basis ekonomi pada sektor primer seperti pertanian dan perkebunan, dengan akses terhadap layanan publik yang masih terbatas dibandingkan wilayah perkotaan.

Klaster 2 hanya beranggotakan Kota Padang sebagai ibu kota provinsi. Kota Padang mencatatkan nilai tertinggi pada hampir seluruh indikator, yaitu PDRB per kapita sebesar Rp54.090.000, distribusi persentase PDRB tertinggi sebesar 25,34 persen, laju pertumbuhan ekonomi sebesar 3,42 persen, serta IPM tertinggi sebesar 84,38. Namun di sisi lain, Kota Padang juga mencatatkan TPT tertinggi sebesar 9,88 persen, yang mengindikasikan bahwa tingginya aktivitas ekonomi belum sepenuhnya diimbangi dengan penyerapan tenaga kerja yang optimal. Keberadaan Kota Padang sebagai klaster tersendiri mencerminkan dominasi ekonomi yang sangat menonjol dibandingkan wilayah lainnya di Provinsi Sumatera Barat.

Klaster 3 beranggotakan seluruh kota selain Kota Padang. Keenam kota ini memiliki rata-rata PDRB per kapita sebesar Rp45.879.170, IPM rata-rata 80,20, serta persentase penduduk miskin terendah sebesar 4,01 persen di antara ketiga klaster. Rendahnya tingkat kemiskinan pada klaster ini mencerminkan kualitas kesejahteraan masyarakat yang lebih baik. Laju pertumbuhan ekonomi klaster ini sebesar 3,04 persen dengan distribusi persentase PDRB sebesar 2,18 persen, hal ini dapat dipahami karena kota-kota tersebut sudah berada pada tahap ekonomi yang lebih matang dibandingkan kabupaten-kabupaten di klaster 1.

Sebaran spasial ketiga klaster dapat dilihat pada Gambar 3.



**Gambar 3.** Peta Sebaran Klaster Kabupaten/Kota Berdasarkan Indikator Pembangunan Wilayah

Berdasarkan Gambar 3, terlihat bahwa Klaster 1 (biru muda) mendominasi hampir seluruh wilayah Provinsi Sumatera Barat, khususnya pada wilayah kabupaten yang tersebar di bagian utara, timur, selatan, dan kepulauan. Klaster 2 (hijau) hanya mencakup Kota Padang yang terletak di wilayah pesisir barat. Sementara itu, Klaster 3 (merah muda/magenta) tersebar di beberapa kota kecil yang umumnya terletak di kawasan dataran tinggi dan perkotaan di bagian tengah provinsi.

### E. Validasi Hasil Klaster

Validasi hasil analisis klaster *K-Medoids* dilakukan menggunakan Davies-Bouldin Index (DBI). Hasil perhitungan DBI untuk  $k=3$  diperoleh nilai sebesar 0,8341. Suatu hasil analisis klaster dikatakan optimal apabila memiliki nilai DBI yang kecil atau mendekati nol. Dengan nilai DBI sebesar 0,8341 yang mendekati nol. Untuk memastikan bahwa  $k=3$  merupakan pilihan terbaik, dilakukan perbandingan dengan nilai  $k$  lainnya, di mana  $k=2$  menghasilkan DBI sebesar 1,3845 dan  $k=4$  sebesar 0,8421. Nilai DBI terkecil diperoleh pada  $k=3$ , sehingga dengan nilai DBI sebesar 0,834, dapat disimpulkan bahwa hasil pengelompokan 19 kabupaten/kota di Provinsi Sumatera Barat menggunakan analisis klaster *K-Medoids* dengan 3 klaster adalah optimal.

## IV. KESIMPULAN

Pengelompokan 19 kabupaten/kota di Provinsi Sumatera Barat berdasarkan indikator sosial ekonomi tahun 2024 menggunakan analisis klaster *K-Medoids* menghasilkan 3 klaster optimal berdasarkan metode *Elbow*. Klaster 1 beranggotakan 12 kabupaten, yaitu Kepulauan Mentawai, Pesisir Selatan, Solok, Sijunjung, Tanah Datar, Padang Pariaman, Agam, Lima Puluh Kota, Pasaman, Solok Selatan, Dharmasraya, dan Pasaman Barat, yang merupakan wilayah dengan rata-rata PDRB per kapita dan IPM terendah, serta persentase penduduk miskin tertinggi di antara ketiga klaster. Klaster 2 hanya beranggotakan Kota Padang yang merupakan wilayah dengan rata-rata PDRB per kapita, laju pertumbuhan ekonomi, distribusi persentase PDRB, dan IPM tertinggi, meskipun juga memiliki TPT tertinggi. Klaster 3 beranggotakan 6 kota, yaitu Kota Solok, Kota Sawah Lunto, Kota Padang Panjang, Kota Bukittinggi, Kota Payakumbuh, dan Kota Pariaman, yang merupakan wilayah dengan persentase penduduk miskin terendah dan IPM yang relatif tinggi. Validasi hasil klaster menggunakan *Davies-Bouldin Index* (DBI) didapatkan nilai sebesar 0,8341 yang mendekati nol, ini menunjukkan bahwa hasil pengelompokan yang terbentuk sudah optimal. Hasil pengelompokan ini diharapkan dapat dimanfaatkan sebagai bahan pertimbangan bagi pemerintah daerah dan Bappeda Provinsi Sumatera Barat dalam merumuskan kebijakan pembangunan wilayah yang lebih terarah sesuai karakteristik tiap klaster.

## DAFTAR PUSTAKA

- Badan Pusat Statistik Provinsi Sumatera Barat. (2025). *Provinsi Sumatera Barat dalam Angka*. Badan Pusat Statistik Provinsi Sumatera Barat.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis Eighth Edition*. Cengage Learning EMEA.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques (Third Edition)*. Morgan Kaufmann.
- Hoerunnisa, A., Dwilestari, G., Dikananda, F., Sunana, H., & Pratama, D. (2024). "Komparasi Algoritma K-Means Dan K-Medoids Dalam Analisis Pengelompokan Daerah Rawan Kriminalitas Di Indonesia". *Jurnal Mahasiswa Teknik Informatika*, Vol. 8(1), hal.
- Kapita, S., Mubarak, A., Do Abdullah, S., & Fhadli, M. (2022). "Penerapan Algoritma Clustering Khonen-Som dengan Validasi Davies Bouldin Index pada Pengelompokan Potensi Uang di Indonesia". *IJIS Indonesian Journal on Information System*, Vol. 7(2), hal. 134–143.
- Maori, N. A. (2023). "Metode Elbow Dalam Optimasi Jumlah Cluster Pada K-Means Clustering". *Jurnal SIMETRIS*, Vol. 14(2), hal. 277–288.
- Sari, I. D. P., Pratiwi, S. M., & Al Rosyid, H. (2025). "Clustering Wilayah Pulau Sumatera Berdasarkan Indikator Sosial Ekonomi Menggunakan Metode K-Medoids". *RIGGS: Journal of Artificial Intelligence and Digital Business*, Vol. 4(4), hal. 7530–7536.
- Setiyanto, K. J. P. A. Z., Muhaimin, A., & Maulida Hindrayani, K. (2025). "Improved K-Means Clustering untuk Analisis Ketimpangan Pembangunan Kabupaten/Kota di Pulau Jawa". *In Driving Innovation through Data, Technology, and Engineering Seminar*, Vol. 1(1), hal. 210-219.
- Siregar, A. R. S., Lubis, N. H., Isfa, F., & Muzayyana. (2025). "Pentingnya Pemerataan Pembangunan Ekonomi dalam Meningkatkan Kesejahteraan Masyarakat Indonesia". *CEMERLANG: Jurnal Manajemen dan Ekonomi Bisnis*, Vol. 5, hal. 51–64.

- Sucahyo, F. R., Indyah Hartami Santi, Mohammad Faried Rahmat, & Diki Fahrizal. (2025). "Comparing K-Means and *K-Medoids* algorithms for clustering hamlet regions by tax liabilities in tax determination documents". *International Journal of Science and Technology Research Archive*, Vol. 8(1), hal. 69–78.
- Yenny, N., Mulhady Putra, Amanda Anggraini, & As-Syifa R Ramadhani Tanjung. (2025). "Analisis Perbandingan Sosial Ekonomi Desa dan Kota dalam Pembangunan Wilayah". *WISSEN : Jurnal Ilmu Sosial dan Humaniora*, Vol. 3(2), hal. 103–115.