

BAB I

PENDAHULUAN

1.1 Latar Belakang

Segmentasi wilayah adalah proses pengelompokan wilayah berdasarkan kesamaan karakteristik tertentu guna mempermudah perencanaan pembangunan yang lebih tepat sasaran dan adaptif (Kementerian PPN/Bappenas, 2024). Penelitian ini menggunakan variabel-variabel, yaitu kepadatan penduduk, luas wilayah, Indeks Pembangunan Manusia (IPM), dan Produk Domestik Regional Bruto (PDRB). Kombinasi variabel ini diharapkan mampu menangkap gambaran menyeluruh tentang karakteristik demografi, geografi, sosial, dan ekonomi di setiap daerah sehingga kelompok wilayah yang dibentuk bukan hanya homogen dalam satu aspek, tetapi juga relevan dari berbagai perspektif perencanaan pembangunan. Peraturan Presiden Nomor 18 Tahun 2020 tentang Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2020–2024, Pemerintah Republik Indonesia menegaskan pentingnya pendekatan berbasis wilayah dalam mendukung kebijakan berbasis bukti (*evidence-based policy*). Pendekatan ini diarahkan untuk mengakomodasi perbedaan kondisi antarwilayah, menentukan prioritas pembangunan secara lebih tepat, serta mengurangi ketimpangan melalui pemanfaatan data dan statistik yang akurat.

Provinsi Jawa Tengah terdiri atas 29 kabupaten dan 6 kota dengan luas wilayah mencapai 34.337,49 km² dan jumlah penduduk sebanyak 38.233,9 juta jiwa pada tahun 2025 (BPS Provinsi Jawa Tengah, 2025). Keberagaman indikator pembangunan wilayah di setiap kabupaten dan kota menjadikan Jawa Tengah

sebagai wilayah strategis yang memerlukan pendekatan segmentasi dalam perencanaan pembangunannya. Hal ini didukung data BPS Provinsi Jawa Tengah tahun 2024 bahwa terdapat variasi signifikan dalam kepadatan penduduk, luas wilayah, IPM, dan PDRB antarkabupaten dan kota di provinsi ini.

Variasi indikator pembangunan antardaerah di Provinsi Jawa Tengah dijelaskan melalui karakteristik masing-masing wilayah, yaitu Kota Semarang, sebagai pusat pemerintahan dan ekonomi memiliki kepadatan penduduk besar, luas wilayah yang relatif kecil, IPM tinggi, dan PDRB per kapita tinggi karena konsentrasi sektor jasa, industri, akses pendidikan, dan kesehatan yang lebih baik. Kabupaten Wonosobo berada di wilayah pegunungan memiliki kepadatan penduduk lebih rendah, luas wilayah yang lebih besar, serta IPM dan PDRB per kapita yang lebih rendah akibat struktur ekonomi yang didominasi pertanian. Pemerintah daerah dapat menggunakan segmentasi wilayah untuk mengelompokkan daerah dengan karakteristik serupa agar kebijakan pembangunan lebih tepat sasaran dan sesuai dengan potensi serta tantangan lokal (BPS Provinsi Jawa Tengah, 2023).

Penelitian ini menggunakan kepadatan penduduk, luas wilayah, IPM, dan PDRB sebagai variabel untuk mengukur tingkat pembangunan dan membandingkan karakteristik antarwilayah. Keempat variabel yang digunakan memiliki fungsi penting dalam menggambarkan karakteristik demografi, geografi, sosial, dan ekonomi di suatu daerah. Kepadatan penduduk menunjukkan tingkat konsentrasi populasi pada suatu wilayah yang berkaitan dengan kapasitas terhadap infrastruktur, kebutuhan layanan dasar, dan potensi urbanisasi. Luas wilayah menentukan kapasitas ruang fisik suatu daerah dalam penyediaan layanan,

pengembangan ekonomi, dan pemerataan akses pembangunan. IPM menggambarkan kualitas hidup masyarakat melalui aspek harapan hidup, pendidikan, dan standar hidup layak. PDRB mencerminkan kekuatan ekonomi dan produktivitas suatu wilayah berdasarkan kontribusi berbagai sektor usaha. Analisis yang dilakukan dalam segmentasi wilayah ini adalah analisis *cluster*. Analisis *cluster* merupakan suatu teknik untuk membentuk kelompok yang mirip satu sama lain, sedangkan objek antarkelompok memiliki perbedaan yang jelas (Kaufman and Rousseeuw, 1990).

Clustering merupakan metode pengelompokan wilayah menjadi beberapa kelompok berdasarkan kemiripan karakteristik masing-masing wilayah. Proses ini dilakukan dengan berbagai pendekatan yang disesuaikan dengan karakteristik data penelitian. Metode *clustering* terbagi menjadi dua, yaitu metode nonhierarki dan hierarki. Metode nonhierarki merupakan metode *clustering* yang mengelompokkan data menjadi sejumlah k *cluster* yang telah ditentukan sebelumnya oleh peneliti. Setiap *cluster* minimal berisi satu objek dan setiap objek hanya dapat menjadi anggota satu *cluster*. Metode hierarki merupakan metode *clustering* dengan membentuk pengelompokan yang saling berkaitan (Kaufman dan Rousseeuw, 1990). Dalam penelitian ini, digunakan metode non-hierarki melalui algoritma *K-Medoids* dan *Density-Based Spatial Clustering Algorithm with Noise (DBSCAN)*.

K-Medoids adalah algoritma *clustering* yang diperkenalkan oleh Leon Kaufman dan Peter J. Rousseeuw pada tahun 1987 dalam buku yang berjudul "*Finding Groups in Data: An Introduction to Cluster Analysis*". Algoritma ini termasuk dalam teknik *clustering* partisi dan dikembangkan sebagai

penyempurnaan dari *K-Means* agar lebih *robust* terhadap pengaruh *outlier*. Konsep *clustering K-Medoids* yaitu setiap objek dialokasikan ke salah satu dari *k cluster* dan tidak dapat menjadi anggota lebih dari satu *cluster*. Jika *K-Means* menggunakan nilai rata-rata sebagai pusat *cluster (centroid)*, maka *K-Medoids* menggunakan objek aktual dalam kumpulan data sebagai pusat *cluster* yang disebut *medoids* yaitu objek data aktual yang paling representatif sehingga lebih *robust* terhadap *outlier* serta memiliki keunggulan dalam stabilitas hasil *cluster* dan akurasi ketika diterapkan pada data nyata yang kompleks (Park dan Jun, 2009). Proses pengelompokan objek-objek menjadi beberapa *cluster* berdasarkan jarak terdekat dari objek pengamatan ke *medoids* yang dipilih secara acak (Kaufman dan Rousseeuw, 1990).

Density-Based Spatial Clustering Algorithm with Noise (DBSCAN) diperkenalkan oleh Martin Ester, Hans-Peter Kriegel, Jörg Sander, dan Xiaowei Xu pada tahun 1996 sebagai metode *clustering* berbasis kepadatan. Algoritma ini membentuk *cluster* berdasarkan kedekatan antartitik dalam suatu wilayah yang memiliki tingkat kepadatan tinggi. *DBSCAN* menggunakan dua parameter utama, yaitu *Eps (Epsilon)* yang menyatakan radius lingkungan dan *MinPts* atau minimum objek dalam *Eps* titik pusat untuk membentuk suatu *cluster*. Konsep kepadatan mengacu pada banyaknya titik-titik objek yang berada dalam radius *Epsilon* (ϵ) suatu titik pusat minimal berjumlah sebanyak atau lebih dari *MinPts*. *DBSCAN* mengidentifikasi seluruh objek penelitian dan membagi menjadi 3 kategori, yaitu *core point*, *border point*, dan *noise* (Han *et al.*, 2012).

Penelitian terdahulu telah mengkaji perbandingan berbagai algoritma dalam proses pengelompokan data. Penelitian yang dilakukan oleh Akbar *et al.* (2023)

mengenai perbandingan algoritma *clustering* antara *K-Medoids* dan *DBSCAN* dalam *clustering* tentang Indikator Pembangunan Manusia (IPM) di Provinsi Sulawesi Selatan tahun 2021 dengan validasi *Silhouette Coefficient* menunjukkan bahwa pengelompokan *cluster* terbaik adalah dengan algoritma *K-Medoids*. Penelitian yang dilakukan oleh Saputri dan Arianto (2023) mengenai perbandingan algoritma *K-Means*, *K-Medoids*, dan *DBSCAN* dalam penggerombolan provinsi di Indonesia berdasarkan indikator kesejahteraan masyarakat menunjukkan bahwa pengelompokan *cluster* terbaik adalah dengan algoritma *DBSCAN*.

Penelitian ini membahas tentang segmentasi wilayah Provinsi Jawa Tengah menggunakan algoritma *K-Medoids* dan *DBSCAN* berdasarkan indikator pembangunan wilayah tahun 2024. Hasil penelitian ini diharapkan dapat menjadi dasar dalam menentukan pendekatan *clustering* yang paling representatif untuk mendukung analisis pemerataan dan evaluasi kebijakan pembangunan antarwilayah di Provinsi Jawa Tengah.

1.2 Rumusan Masalah

Masalah yang ditemukan dari uraian latar belakang dapat dirumuskan sebagai berikut:

1. Bagaimana pengelompokan kabupaten/kota di Provinsi Jawa Tengah berdasarkan indikator pembangunan wilayah tahun 2024 menggunakan metode *K-Medoids Clustering*?
2. Bagaimana pengelompokan kabupaten/kota di Provinsi Jawa Tengah berdasarkan indikator pembangunan wilayah tahun 2024 menggunakan metode *DBSCAN*?

3. Bagaimana hasil pengelompokan optimal kabupaten/kota di Provinsi Jawa Tengah berdasarkan indikator pembangunan wilayah tahun 2024 dengan validasi *Silhouette Coefficient*?

1.3 Batasan Masalah

Rumusan masalah yang telah ditentukan menjadi dasar penetapan batasan masalah sebagai berikut:

1. Data yang digunakan merupakan data indikator pembangunan wilayah di Provinsi Jawa Tengah tahun 2024 yang diambil melalui laman resmi Badan Pusat Statistik.
2. Analisis *cluster* menggunakan algoritma *K-Medoids* dan *DBSCAN* serta pemilihan kelompok terbaik menggunakan validasi *Silhouette Coefficient*.
3. Penentuan jumlah *cluster* optimal pada algoritma *K-Medoids* menggunakan nilai validasi *Silhouette Coefficient* dengan jumlah *cluster* $k = 2, 3, \dots, 10$ dan penentuan nilai parameter yang optimal pada algoritma *DBSCAN* menggunakan nilai validasi *Silhouette Coefficient* dengan percobaan *trial and error* untuk $MinPts = 4, 5, 6, 7, \text{ dan } 8$ serta $Epsilon = 0,4$ hingga $3,0$.

1.4 Tujuan Penelitian

1. Menerapkan metode *K-Medoids Clustering* untuk mengelompokkan kabupaten/kota di Provinsi Jawa Tengah berdasarkan indikator pembangunan wilayah tahun 2024.
2. Menerapkan metode *DBSCAN* untuk mengelompokkan kabupaten/kota di Provinsi Jawa Tengah berdasarkan indikator pembangunan wilayah tahun 2024.

3. Menentukan pengelompokan optimal kabupaten/kota di Provinsi Jawa Tengah berdasarkan indikator pembangunan wilayah tahun 2024 dengan validasi *Silhouette Coefficient*.