



Analisis Klasterisasi Wilayah Berdasarkan Tingkat Kepadatan Penduduk Menggunakan Algoritma K-Means Berbasis Sistem Informasi Geografis

Mustafa Iffat Shafi Athallah*, Wahyu Tjahjo Saputro, Ike Yunia Pasa

Fakultas Teknik, Program Studi Teknologi Informasi, Universitas Muhammadiyah Purworejo, Purworejo, Indonesia

Email: ^{1,*}iffatattallah1710@gmail.com, ²wahjusaputro@umpwr.ac.id, ³ikeypasa@umpwr.ac.id

Email Penulis Korespondensi: iffatattallah1710@gmail.com

Abstrak—Penelitian ini bertujuan untuk menganalisis dan memetakan tingkat kepadatan penduduk kabupaten/kota di Provinsi Jawa Tengah menggunakan pendekatan analisis spasial berbasis Sistem Informasi Geografis (SIG) dan algoritma K-Means. Permasalahan yang diangkat adalah belum tersedianya pemetaan kepadatan penduduk yang terklasifikasi secara sistematis dan informatif untuk mendukung analisis wilayah dan pengambilan keputusan. Data sekunder diperoleh dari Badan Pusat Statistik (BPS) meliputi jumlah penduduk, laju pertumbuhan penduduk, persentase penduduk, kepadatan penduduk per kilometer persegi, serta data spasial batas administrasi. Sebelum proses klasterisasi, seluruh variabel dinormalisasi menggunakan metode Min-Max untuk menghindari bias akibat perbedaan rentang nilai antarvariabel dalam perhitungan jarak Euclidean. Metode penelitian bersifat deskriptif kuantitatif dengan penerapan K-Means ($K=3$) untuk mengelompokkan wilayah ke dalam kategori kepadatan rendah, sedang, dan tinggi. Hasil analisis menunjukkan bahwa dari 35 kabupaten/kota, sebanyak 7 wilayah (20%) termasuk klaster tinggi, 22 wilayah (62,86%) klaster sedang, dan 6 wilayah (17,14%) klaster rendah. Implementasi hasil klasterisasi dalam bentuk peta tematik dengan skema warna merah (tinggi), kuning (sedang), dan hijau (rendah) berhasil memperjelas pola spasial persebaran penduduk, sehingga mendukung perencanaan pembangunan berbasis data demografis dan spasial secara lebih komprehensif.

Kata Kunci: Kepadatan Penduduk; Sistem Informasi Geografis; K-Means; Clustering; Pemetaan Spasial

Abstract—This study aims to analyze and map the population density of regencies and municipalities in Jawa Tengah using a spatial analysis approach based on Geographic Information Systems (GIS) and the K-Means clustering algorithm. The main issue addressed is the lack of systematically classified and informative population density mapping to support spatial analysis and regional decision-making. Secondary data were obtained from the Central Bureau of Statistics (BPS), including total population, population growth rate, population percentage, population density per square kilometer, and administrative boundary spatial data. Prior to clustering, all variables were normalized using the Min-Max scaling method to prevent bias caused by differences in variable ranges in Euclidean distance calculations. The research employed a quantitative descriptive method with K-Means ($K=3$) to classify regions into low, medium, and high population density clusters. The results indicate that out of 35 regencies/municipalities, 7 regions (20%) fall into the high-density cluster, 22 regions (62.86%) into the medium-density cluster, and 6 regions (17.14%) into the low-density cluster. The implementation of the clustering results into a thematic map using a color scheme of red (high), yellow (medium), and green (low) effectively visualizes spatial distribution patterns, thereby supporting data-driven regional planning and spatial-based policy formulation.

Keywords: Population Density; Geographic Information System; K-Means; Clustering; Spatial Mapping

1. PENDAHULUAN

Kepadatan penduduk merupakan salah satu indikator demografis yang memiliki peranan strategis dalam perencanaan dan pengelolaan pembangunan wilayah. Indikator ini tidak hanya mencerminkan jumlah penduduk yang mendiami suatu wilayah dalam satuan luas tertentu, tetapi juga menggambarkan tingkat tekanan terhadap sumber daya alam, infrastruktur, fasilitas umum, serta daya dukung lingkungan. Wilayah dengan kepadatan tinggi cenderung menghadapi tantangan berupa kemacetan, keterbatasan ruang terbuka hijau, peningkatan kebutuhan perumahan, serta tingginya beban pelayanan publik. Sebaliknya, wilayah dengan kepadatan rendah sering kali menghadapi permasalahan keterbatasan akses infrastruktur, rendahnya intensitas kegiatan ekonomi, dan kurang optimalnya pemanfaatan ruang [1], [2]. Oleh karena itu, analisis kepadatan penduduk menjadi bagian penting dalam perumusan kebijakan pembangunan yang berorientasi pada pemerataan, keberlanjutan, dan efisiensi tata ruang.

Provinsi Jawa Tengah sebagai salah satu provinsi dengan jumlah penduduk besar di Indonesia memiliki karakteristik wilayah yang beragam, baik dari segi geografis, sosial, maupun ekonomi. Perbedaan luas wilayah, tingkat urbanisasi, struktur perekonomian, dan kondisi topografi menyebabkan variasi kepadatan penduduk antar kabupaten dan kota cukup signifikan. Wilayah perkotaan seperti ibu kota provinsi dan kota-kota penyangga umumnya memiliki tingkat kepadatan yang jauh lebih tinggi dibandingkan wilayah perdesaan atau kabupaten dengan dominasi sektor pertanian. Ketimpangan distribusi ini memerlukan analisis yang tidak hanya bersifat numerik, tetapi juga mampu menggambarkan pola persebaran secara keruangan agar dapat dipahami secara lebih komprehensif [3], [4]. Namun demikian, penyajian data kependudukan yang masih didominasi oleh tabel statistik dan grafik konvensional belum sepenuhnya mampu memberikan gambaran spasial yang utuh mengenai dinamika kepadatan penduduk antarwilayah.

Dalam konteks analisis wilayah, pendekatan spasial menjadi sangat penting karena setiap fenomena kependudukan selalu memiliki dimensi lokasi. Sistem Informasi Geografis (SIG) hadir sebagai teknologi yang mampu mengintegrasikan data spasial dan data atribut untuk dianalisis serta divisualisasikan dalam bentuk peta tematik yang informatif. SIG memungkinkan pengguna untuk mengidentifikasi pola, hubungan, dan distribusi fenomena berdasarkan lokasi geografis secara lebih sistematis [5], [6]. Dengan memanfaatkan SIG, data kepadatan penduduk tidak hanya disajikan dalam bentuk angka, tetapi juga dapat dipetakan sehingga memperlihatkan pola konsentrasi, sebaran, serta kecenderungan wilayah



tertentu. Sejumlah penelitian menunjukkan bahwa pemanfaatan SIG dalam pemetaan kepadatan penduduk mampu meningkatkan kualitas analisis dibandingkan metode statistik konvensional yang bersifat nonspasial [3], [7]. Selain itu, penggunaan perangkat lunak open-source seperti QGIS dinilai efektif karena mendukung integrasi data spasial dan atribut secara fleksibel, efisien, dan mudah direplikasi [8], [9].

Meskipun SIG mampu menyajikan visualisasi spasial yang baik, proses klasifikasi wilayah berdasarkan tingkat kepadatan tetap memerlukan metode analisis kuantitatif yang objektif. Salah satu pendekatan yang banyak digunakan dalam analisis data adalah teknik clustering. Algoritma K-Means merupakan metode clustering non-hierarkis yang mengelompokkan data berdasarkan kedekatan karakteristik ke dalam sejumlah kluster tertentu [10], [11]. Prinsip dasar K-Means adalah meminimalkan jarak antar data dalam satu kluster dan memaksimalkan perbedaan antar kluster, sehingga setiap kelompok yang terbentuk memiliki tingkat homogenitas internal yang tinggi. Dalam konteks kependudukan, metode ini dapat digunakan untuk mengelompokkan wilayah berdasarkan variabel-variabel tertentu seperti jumlah penduduk, laju pertumbuhan penduduk, persentase penduduk, dan kepadatan penduduk per kilometer persegi.

Penerapan algoritma K-Means dalam analisis kependudukan telah menunjukkan hasil yang cukup efektif dalam menghasilkan segmentasi wilayah yang sistematis dan berbasis data [12], [13]. Integrasi antara SIG dan K-Means juga telah banyak digunakan dalam penelitian sebelumnya untuk menghasilkan peta tematik berbasis kluster yang mampu memberikan gambaran klasifikasi wilayah secara lebih jelas [14]. Dengan pendekatan ini, wilayah dapat dikelompokkan ke dalam kategori tertentu seperti kepadatan rendah, sedang, dan tinggi secara lebih objektif dibandingkan dengan metode klasifikasi manual atau berbasis interval sederhana. Namun demikian, sebagian penelitian masih berfokus pada analisis nonspasial atau hanya menggunakan visualisasi dasar tanpa mengintegrasikan teknik clustering multidimensi, sehingga belum sepenuhnya mampu menggambarkan heterogenitas wilayah secara menyeluruh [15], [16], [17].

Pendekatan clustering multidimensi menjadi relevan karena fenomena kepadatan penduduk tidak hanya dipengaruhi oleh satu variabel tunggal. Kombinasi antara jumlah penduduk, laju pertumbuhan penduduk, persentase penduduk terhadap total wilayah, serta tingkat kepadatan per kilometer persegi memberikan gambaran yang lebih komprehensif mengenai karakteristik demografis suatu daerah [3], [14]. Penggunaan beberapa variabel secara simultan memungkinkan proses klasifikasi yang lebih representatif terhadap kondisi nyata di lapangan. Penelitian Aryanto dkk. tahun 2024 menunjukkan bahwa algoritma K-Means mampu meningkatkan akurasi segmentasi wilayah berbasis statistik dan membantu proses interpretasi hasil analisis secara lebih objektif [18]. Sementara itu, Rahmawati dkk. tahun 2024 menegaskan bahwa integrasi SIG dalam analisis spasial memberikan kontribusi signifikan dalam mendukung pengambilan keputusan berbasis wilayah, khususnya dalam konteks perencanaan pembangunan dan pengelolaan sumber daya [19].

Selain sebagai alat visualisasi, Sistem Informasi Geografis memiliki peran strategis dalam mendukung analisis spasial dan pengambilan keputusan berbasis wilayah. SIG memungkinkan integrasi data spasial dan atribut secara terstruktur sehingga mampu mengungkap pola, hubungan, dan tren keruangan yang tidak dapat diperoleh melalui data tabular semata. Menurut [20], SIG merupakan sistem yang dirancang untuk mengumpulkan, menyimpan, mengelola, menganalisis, dan menampilkan data bereferensi geografis guna mendukung proses analisis dan perencanaan wilayah secara komprehensif. Hal senada juga dikemukakan oleh [21], yang menyatakan bahwa analisis spasial berbasis SIG sangat efektif dalam memahami dinamika kependudukan dan pemanfaatan ruang karena mampu mengaitkan data statistik dengan konteks lokasi secara langsung. Oleh karena itu, penerapan SIG yang dikombinasikan dengan metode analisis berbasis data mining, seperti algoritma K-Means, menjadi pendekatan yang relevan dan kuat dalam menghasilkan informasi kepadatan penduduk yang objektif, akurat, dan aplikatif untuk mendukung perencanaan pembangunan wilayah di Provinsi Jawa Tengah.

Berdasarkan berbagai temuan tersebut, dapat disimpulkan bahwa kombinasi antara analisis spasial berbasis SIG dan metode clustering K-Means memiliki potensi besar dalam menghasilkan pemetaan kepadatan penduduk yang lebih informatif dan aplikatif. Namun, penerapan pendekatan ini pada skala Provinsi Jawa Tengah dengan mempertimbangkan variabel multidimensi masih memerlukan pengembangan lebih lanjut agar menghasilkan klasifikasi yang objektif serta visualisasi spasial yang mudah dipahami. Oleh karena itu, penelitian ini berkontribusi dalam mengintegrasikan analisis kepadatan penduduk berbasis SIG dan algoritma K-Means pada skala Provinsi Jawa Tengah untuk menghasilkan klasifikasi wilayah secara objektif dan divisualisasikan dalam bentuk peta tematik. Berbeda dengan penelitian sebelumnya yang cenderung parsial, penelitian ini menggabungkan analisis spasial dan clustering multidimensi sehingga menghasilkan informasi kepadatan penduduk yang lebih komprehensif, akurat, dan aplikatif untuk mendukung perencanaan pembangunan wilayah.

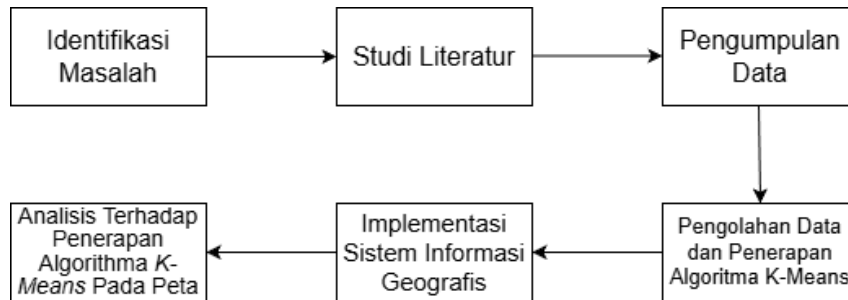
2. METODOLOGI PENELITIAN

Penelitian ini menggunakan metode deskriptif kuantitatif dengan pendekatan analisis spasial berbasis Sistem Informasi Geografis (SIG) untuk menganalisis kepadatan penduduk di Provinsi Jawa Tengah. Data yang digunakan merupakan data sekunder berupa jumlah penduduk dan luas wilayah kabupaten/kota yang bersumber dari Badan Pusat Statistik (BPS). Data tersebut diolah untuk memperoleh nilai kepadatan penduduk, kemudian dianalisis menggunakan algoritma *K-Means* guna mengelompokkan wilayah ke dalam tiga kluster, yaitu kepadatan rendah, sedang, dan tinggi. Hasil klusterisasi selanjutnya diintegrasikan dengan data spasial wilayah dan divisualisasikan dalam bentuk peta tematik menggunakan perangkat lunak *QGIS* untuk menampilkan pola sebaran kepadatan penduduk secara spasial dan informatif.



2.1 Tahapan Penelitian

Penelitian ini dilaksanakan melalui serangkaian tahapan yang disusun secara sistematis untuk memastikan proses analisis berjalan terstruktur dan terarah. Tahapan penelitian meliputi identifikasi masalah terkait kepadatan penduduk, studi literatur untuk memperoleh landasan teoritis dan metodologis, pengumpulan data kependudukan dan data spasial wilayah, pengolahan data serta penerapan algoritma *K-Means*, implementasi *Sistem Informasi Geografis (SIG)* menggunakan perangkat lunak *QGIS*, serta tahap akhir berupa analisis hasil dan penarikan kesimpulan. Gambar 1 merupakan tahapan penelitian yang dipakai pada jurnal ini.



Gambar 1. Tahapan Penelitian

2.2.1 Identifikasi Masalah

Tahap awal dalam penelitian ini adalah melakukan identifikasi masalah sebagai dasar dalam perumusan arah dan tujuan penelitian. Identifikasi masalah dilakukan dengan mengamati kondisi kepadatan penduduk di seluruh kabupaten dan kota di Provinsi Jawa Tengah yang menunjukkan tingkat kepadatan yang berbeda-beda antarwilayah. Perbedaan tersebut mencerminkan adanya ketimpangan distribusi penduduk yang berpotensi menimbulkan berbagai permasalahan, seperti tekanan terhadap fasilitas publik, ketidakseimbangan pembangunan wilayah, serta permasalahan tata ruang dan lingkungan.

Permasalahan utama yang dihadapi adalah belum tersedianya informasi kepadatan penduduk yang tersaji secara spasial dan terklasifikasi secara jelas berdasarkan tingkat kepadatannya, sehingga menyulitkan proses analisis dan pengambilan keputusan oleh pemerintah daerah. Data kependudukan yang tersedia umumnya masih disajikan dalam bentuk tabel statistik, yang kurang mampu menggambarkan pola sebaran penduduk dan hubungan spasial antarwilayah secara komprehensif. Oleh karena itu, diperlukan suatu sistem yang mampu mengelompokkan wilayah berdasarkan tingkat kepadatan penduduk secara objektif sekaligus memvisualisasikannya dalam bentuk peta tematik yang informatif. Pemanfaatan Sistem Informasi Geografis (SIG) dipandang sebagai solusi yang tepat karena mampu mengintegrasikan data kependudukan dengan data spasial wilayah, sehingga menghasilkan informasi kepadatan penduduk yang lebih efektif, mudah dipahami, dan aplikatif untuk mendukung perencanaan pembangunan berbasis wilayah.

2.2.2 Studi Literatur

Studi literatur dilakukan dengan mengkaji berbagai sumber referensi yang relevan dengan topik penelitian. Studi literatur mencakup pembahasan mengenai konsep Sistem Informasi Geografis (SIG), algoritma *K-Means*, metode *clustering*, serta karakteristik dan pengolahan data kependudukan. Selain itu, penelitian-penelitian terdahulu yang memiliki keterkaitan dengan pemetaan dan pengelompokan kepadatan penduduk juga dianalisis untuk memahami perkembangan metode yang telah diterapkan sebelumnya. Tujuan dari studi literatur ini adalah untuk memperoleh landasan teori yang kuat dan komprehensif sebagai dasar dalam perancangan sistem, pemilihan metode analisis yang tepat, serta penerapan algoritma yang digunakan dalam penelitian agar sesuai dengan permasalahan yang dikaji.

2.2.3 Pengumpulan Data

Pengumpulan data pada penelitian ini menggunakan data nonspasial dan data spasial yang diperoleh dari sumber resmi. Data nonspasial bersumber dari Badan Pusat Statistik (BPS) Provinsi Jawa Tengah, yang meliputi jumlah penduduk (ribu jiwa), laju pertumbuhan penduduk per tahun, persentase penduduk, serta kepadatan penduduk per kilometer persegi (km²). Data nonspasial tersebut digunakan sebagai dasar dalam perhitungan dan analisis kepadatan penduduk yang menjadi objek utama dalam proses *clustering*.

Selain itu, penelitian ini juga menggunakan data spasial berupa file shapefile (SHP) yang memuat batas administrasi wilayah kabupaten dan kota di Provinsi Jawa Tengah. Data spasial ini berfungsi sebagai representasi geografis wilayah penelitian yang akan diintegrasikan dengan data nonspasial. Teknik pengumpulan data dilakukan dengan metode dokumentasi, yaitu dengan mengunduh dan mengolah data resmi yang dipublikasikan oleh BPS serta data spasial wilayah administrasi, sehingga seluruh data yang digunakan memiliki validitas dan keakuratan yang dapat dipertanggung jawabkan.

Gambar 2 merupakan data non spasial yang digunakan untuk analisis pada penelitian ini, data tersebut bersumber dari Badan Pusat Statistik.



Kabupaten/Kota	Jumlah Penduduk (Ribu)	Laju Pertumbuhan Penduduk per Tahun	Persentase Penduduk	Kepadatan Penduduk per km persegi (Ribu)	Rasio Jenis Kelamin Penduduk
Cilacap	2.922,4	1,03	5,35	901	101,4
Banyuwangi	1.847,3	1,08	4,87	1.391	101,2
Purbalingga	1.037,6	1,05	2,74	1.334	102,4
Garjanegara	1.057,6	1,04	2,79	993	102,9
Kebumen	1.414,8	1,27	3,73	1.060	102,8
Purworejo	795,0	0,89	2,30	768	100,3
Wonorejo	920,5	1,25	2,43	935	103,7
Magelang	1.341,5	0,85	3,54	1.236	101,1
Biyakal	1.099,9	0,94	2,90	1.017	100,9

Gambar 2. Data BPS Jumlah Penduduk Jawa Tengah Tahun 2024

2.2.4 Pengolahan Data dan Penerapan Algoritma K-Means

Data yang telah dikumpulkan kemudian dilakukan proses pengolahan data. Tahapan pengolahan data meliputi pembersihan data (data cleaning) untuk memastikan data bebas dari kesalahan, duplikasi, dan data yang tidak lengkap. Selanjutnya dilakukan proses transformasi data agar sesuai dengan format yang dibutuhkan dalam penerapan algoritma K-Means. Mengingat penelitian ini menggunakan beberapa atribut dengan rentang nilai yang berbeda, yaitu jumlah penduduk (JP), laju pertumbuhan penduduk (LPP), persentase penduduk (PP), dan kepadatan penduduk (KP), maka sebelum proses clustering dilakukan normalisasi data. Normalisasi diperlukan untuk menghindari dominasi variabel dengan skala besar (misalnya jumlah penduduk dalam jutaan jiwa) terhadap variabel dengan skala kecil (misalnya laju pertumbuhan dalam persen). Tanpa proses normalisasi, perhitungan jarak pada algoritma K-Means dapat menghasilkan kluster yang bias karena dipengaruhi oleh variabel dengan rentang nilai terbesar. Pada penelitian ini, metode normalisasi yang digunakan adalah Min-Max Scaling, yaitu teknik transformasi data yang mengubah nilai atribut ke dalam rentang tertentu, umumnya antara 0 hingga 1.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Nilai normalisasi (X') diperoleh dari hasil transformasi data asli (X) menggunakan metode normalisasi Min-Max. Proses ini dilakukan dengan cara mengurangi setiap nilai data asli dengan nilai minimum atribut (X_{min}), kemudian membaginya dengan selisih antara nilai maksimum atribut (X_{max}) dan nilai minimum atribut (X_{min}). Melalui proses ini, setiap nilai pada atribut akan ditransformasikan ke dalam rentang 0 hingga 1. Nilai X_{min} merepresentasikan nilai terendah dari suatu atribut dalam keseluruhan dataset, sedangkan X_{max} menunjukkan nilai tertinggi dari atribut tersebut. Normalisasi ini bertujuan untuk menyelaraskan skala antarvariabel sehingga tidak terjadi dominasi variabel tertentu dalam proses perhitungan jarak pada algoritma K-Means. Melalui proses ini, seluruh variabel berada dalam skala yang sebanding sehingga setiap atribut memiliki kontribusi yang proporsional dalam perhitungan jarak pada algoritma K-Means.

Pada tahap ini juga dilakukan perhitungan kepadatan penduduk berdasarkan jumlah penduduk dan luas wilayah masing-masing Kabupaten/Kota. Pada tahap ini, data kepadatan penduduk Kabupaten/Kota di Provinsi Jawa Tengah dikelompokkan menggunakan algoritma K-Means. Algoritma K-Means merupakan metode clustering yang bertujuan untuk mengelompokkan data ke dalam sejumlah kluster berdasarkan tingkat kemiripan data. Dalam penelitian ini, jumlah kluster yang digunakan adalah tiga kluster (K3), yang merepresentasikan tingkat kepadatan penduduk, yaitu kepadatan rendah, kepadatan sedang, dan kepadatan tinggi.

$$D(p, c)_n = \sqrt{\sum_{i=0}^n (p_i - c_i)^2} \tag{2}$$

Langkah awal dalam penerapan algoritma K-Means adalah menentukan nilai centroid awal untuk masing-masing kluster. Penentuan centroid awal dilakukan secara acak dari data kepadatan penduduk. Selanjutnya, dilakukan perhitungan jarak antara setiap data kepadatan penduduk dengan masing-masing centroid. Rumus tersebut digunakan untuk menghitung jarak antara data kepadatan penduduk (p) dengan pusat kluster (c). Nilai p_i merupakan nilai atribut kepadatan penduduk pada Kabupaten/Kota ke- i , sedangkan c_i merupakan nilai centroid pada kluster ke- i . Data kemudian dimasukkan ke dalam kluster yang memiliki jarak paling kecil. Setelah seluruh data dikelompokkan ke dalam kluster terdekat, dilakukan perhitungan ulang centroid baru dengan cara menghitung nilai rata-rata dari seluruh data yang berada dalam masing-masing kluster. Proses perhitungan jarak dan pembaruan centroid ini dilakukan secara berulang hingga nilai centroid tidak mengalami perubahan atau telah mencapai kondisi konvergen.

Penentuan jumlah kluster ($K=3$) dalam penelitian ini tidak didasarkan pada metode optimasi statistik seperti Elbow atau Silhouette Coefficient, melainkan pada pendekatan konseptual yang bertujuan untuk mengelompokkan wilayah ke dalam tiga kategori interpretatif, yaitu kepadatan rendah, sedang, dan tinggi. Klasifikasi tiga tingkat ini umum digunakan dalam analisis wilayah untuk mendukung interpretasi kebijakan dan perencanaan pembangunan. Dengan demikian, pemilihan K3 bersifat aplikatif dan disesuaikan dengan kebutuhan analisis spasial dalam konteks perencanaan wilayah.



2.2.5 Implementasi Sistem Informasi Geografis

Hasil pengelompokan kepadatan penduduk menggunakan algoritma K-Means selanjutnya diimplementasikan ke dalam Sistem Informasi Geografis. Setiap Kabupaten/Kota ditampilkan dalam bentuk peta digital dengan warna yang berbeda, untuk kepadatan rendah berwarna hijau, kepadatan sedang berwarna kuning, dan kepadatan tinggi berwarna merah, sesuai dengan kluster kepadatan penduduk yang dihasilkan. Implementasi SIG ini bertujuan untuk mempermudah visualisasi, analisis, dan pemahaman terhadap persebaran kepadatan penduduk di Provinsi Jawa Tengah.

2.2.6 Analisis Terhadap Penerapan Algoritma K-Means Pada Peta

Tahap analisis terhadap penerapan algoritma *K-Means* pada peta dilakukan dengan mengevaluasi hasil clustering yang telah diintegrasikan ke dalam Sistem Informasi Geografis (SIG). Analisis ini bertujuan untuk menilai kesesuaian hasil pengelompokan wilayah kabupaten dan kota berdasarkan tingkat kepadatan penduduk dengan kondisi spasial yang ditampilkan pada peta tematik. Pada tahap ini dilakukan pengamatan terhadap pola persebaran kluster kepadatan penduduk, baik kluster rendah, sedang, maupun tinggi, untuk mengetahui konsistensi hasil clustering dengan karakteristik wilayah secara geografis. Selain itu, analisis juga difokuskan pada identifikasi wilayah yang menunjukkan konsentrasi kepadatan tinggi serta wilayah dengan kepadatan rendah, sehingga dapat diketahui keterkaitan antara hasil algoritma K-Means dengan faktor spasial seperti lokasi perkotaan, pusat kegiatan ekonomi, dan wilayah perdesaan. Hasil analisis ini digunakan untuk menilai efektivitas algoritma *K-Means* dalam mengelompokkan data kependudukan secara objektif serta memastikan bahwa visualisasi peta yang dihasilkan mampu memberikan informasi spasial yang akurat, informatif, dan mudah dipahami sebagai dasar pengambilan keputusan berbasis wilayah.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengolahan Data dan Penerapan Algoritma K-Means

Data yang telah dikumpulkan selanjutnya melalui tahap pengolahan data yang diawali dengan proses pembersihan data (data cleaning). Tahap ini bertujuan untuk memastikan kualitas, keakuratan, dan konsistensi data kependudukan yang digunakan sehingga layak untuk dianalisis lebih lanjut. Proses pembersihan data meliputi pemeriksaan kelengkapan data pada setiap variabel, identifikasi dan penghapusan data duplikat, serta penyeragaman penulisan nama wilayah kabupaten/kota agar sesuai dan dapat terintegrasi dengan data spasial yang digunakan dalam penelitian. Selain itu, dilakukan pengecekan kesesuaian format data, seperti tipe numerik dan satuan pengukuran, serta koreksi terhadap data yang tidak valid atau tidak wajar untuk meminimalkan potensi kesalahan pada tahap analisis selanjutnya. Melalui proses data cleaning ini, diharapkan data kependudukan yang digunakan memiliki tingkat keandalan yang tinggi sehingga dapat menghasilkan hasil clustering dan analisis spasial yang akurat dan dapat dipertanggungjawabkan secara ilmiah. Tabel 1 merupakan data Badan Pusat Statistik yang telah melalui proses pembersihan data.

Table 1. Data BPS yang sudah dibersihkan

Kabupaten/Kota	Jumlah Penduduk (Ribuan)	Laju Pertumbuhan Penduduk per Tahun	Persentase Penduduk	Kepadatan Penduduk per km persegi (Km ²)
Cilacap	2027.4	1.03	5.35	901
Banyumas	1847.1	1.06	4.87	1391
Purbalingga	1037.6	1.05	2.74	1334
Banjarnegara	1057.6	1.04	2.79	993
Kebumen	1414.8	1.27	3.73	1060
Purworejo	795	0.89	2.1	768
Wonosobo	920.5	1.25	2.43	935
Magelang	1341.5	0.85	3.54	1236
Boyolali	1099.9	0.94	2.9	1017
Klaten	1292.8	0.7	3.41	1846
Sukoharjo	941.7	1.01	2.49	2018
Wonogiri	1054.2	0.31	2.78	588
Karanganyar	961.9	0.75	2.54	1253
Sragen	1004.8	0.77	2.65	1010
Grobogan	1506.4	0.96	3.98	762
Blora	908	0.75	2.4	464
Rembang	665.2	0.83	1.76	641
Pati	1370.8	0.89	3.62	871
Kudus	883.3	1.06	2.33	2077
Jepara	1232.9	1.02	3.25	1228
Demak	1253	1.07	3.31	1396
Semarang	1089.8	0.9	2.88	1147
Temanggung	814.9	0.84	2.15	936



Kabupaten/Kota	Jumlah Penduduk (Ribu)	Laju Pertumbuhan Penduduk per Tahun	Persentase Penduduk	Kepadatan Penduduk per km persegi (Km2)
Kendal	1064.8	1.2	2.81	1062
Batang	838.2	1.2	2.21	978
Pekalongan	1019	1.21	2.69	1219
Pemalang	1541.7	1.25	4.07	1382
Tegal	1674.8	1.28	4.42	1702
Brebes	2065.5	1.16	5.45	1167
Kota Magelang	122.4	0.24	0.32	6600
Kota Surakarta	528	0.26	1.39	11283
Kota Salatiga	201.3	1.24	0.53	3661
Kota Semarang	1708.8	0.89	4.51	4572
Kota Pekalongan	321.1	1.19	0.85	7096
Kota Tegal	285.8	1.15	0.75	7200

Setelah proses data cleaning, dilakukan tahap normalisasi data sebelum penerapan algoritma K-Means. Normalisasi merupakan proses transformasi nilai data ke dalam skala tertentu untuk menghilangkan perbedaan rentang antarvariabel. Tahap ini penting karena variabel yang digunakan dalam penelitian memiliki skala yang sangat berbeda, seperti jumlah penduduk (JP) yang bernilai ribuan hingga jutaan jiwa, sedangkan laju pertumbuhan penduduk (LPP) berada pada rentang desimal yang relatif kecil. Tanpa normalisasi, algoritma K-Means yang berbasis perhitungan jarak Euclidean akan cenderung dipengaruhi oleh variabel dengan rentang nilai terbesar, sehingga variabel dengan skala kecil dapat terabaikan dalam proses pembentukan kluster. Kondisi tersebut berpotensi menghasilkan klusterisasi yang bias dan tidak merepresentasikan kontribusi seluruh variabel secara proporsional. Normalisasi data pada penelitian ini menggunakan rumus Min-Max Scaling (1) yang mentransformasikan setiap nilai atribut ke dalam rentang 0 hingga 1. Proses normalisasi dilakukan terhadap seluruh variabel yang digunakan dalam analisis, yaitu jumlah penduduk (JP), laju pertumbuhan penduduk (LPP), persentase penduduk (PP), dan kepadatan penduduk (KP). Tabel 2 merupakan data sebelumnya yaitu data cleaning yang sudah dilakukan normalisasi.

Tabel 2. Data yang sudah dinormalisasi

No	Kab/Kota	JP	LPP	PP	KP
1	Cilacap	0.9804	0.7596	0.9805	0.0404
2	Banyumas	0.8876	0.7885	0.8869	0.0857
3	Purbalingga	0.4710	0.7788	0.4717	0.0804
4	Banjarnegara	0.4813	0.7692	0.4815	0.0489
5	Kebumen	0.6651	0.9904	0.6647	0.0551
6	Purworejo	0.3461	0.6250	0.3470	0.0281
7	Wonosobo	0.4107	0.9712	0.4113	0.0435
8	Magelang	0.6274	0.5865	0.6277	0.0714
9	Boyolali	0.5031	0.6731	0.5029	0.0511
10	Klaten	0.6023	0.4423	0.6023	0.1277
11	Sukoharjo	0.4216	0.7404	0.4230	0.1436
12	Wonogiri	0.4795	0.0673	0.4795	0.0115
13	Karanganyar	0.4320	0.4904	0.4327	0.0729
14	Sragen	0.4541	0.5096	0.4542	0.0505
15	Grobogan	0.7123	0.6923	0.7135	0.0275
16	Blora	0.4043	0.4904	0.4055	0.0000
17	Rembang	0.2793	0.5673	0.2807	0.0164
18	Pati	0.6425	0.6250	0.6433	0.0376
19	Kudus	0.3916	0.7885	0.3918	0.1491
20	Jepara	0.5715	0.7500	0.5712	0.0706
21	Demak	0.5819	0.7981	0.5828	0.0861
22	Semarang	0.4979	0.6346	0.4990	0.0631
23	Temanggung	0.3564	0.5769	0.3567	0.0436
24	Kendal	0.4850	0.9231	0.4854	0.0553
25	Batang	0.3684	0.9231	0.3684	0.0475
26	Pekalongan	0.4614	0.9327	0.4620	0.0698
27	Pemalang	0.7304	0.9712	0.7310	0.0849
28	Tegal	0.7989	1.0000	0.7992	0.1144
29	Brebes	1.0000	0.8846	1.0000	0.0650
30	Kota Magelang	0.0000	0.0000	0.0000	0.5672
31	Kota Surakarta	0.2087	0.0192	0.2086	1.0000
32	Kota Salatiga	0.0406	0.9615	0.0409	0.2955



No	Kab/Kota	JP	LPP	PP	KP
33	Kota Semarang	0.8164	0.6250	0.8168	0.3797
34	Kota Pekalongan	0.1023	0.9135	0.1033	0.6130
35	Kota Tegal	0.0841	0.8750	0.0838	0.6226

Selanjutnya melalui tahap normalisasi data, langkah selanjutnya dalam penelitian ini adalah melakukan pengelompokan data menggunakan metode algoritma *K-Means*. Algoritma *K-Means* dipilih karena memiliki kemampuan yang efektif dalam mengelompokkan data berdasarkan tingkat kemiripan karakteristik tertentu, dalam hal ini adalah tingkat kepadatan penduduk. Penerapan algoritma ini bertujuan untuk mengelompokkan seluruh kabupaten dan kota di Provinsi Jawa Tengah ke dalam beberapa kelompok yang memiliki karakteristik kepadatan penduduk yang relatif serupa, sehingga perbedaan tingkat kepadatan antarwilayah dapat dianalisis secara lebih sistematis dan objektif.

Jumlah kluster yang digunakan dalam penelitian ini ditetapkan sebanyak tiga kluster, yang merepresentasikan kategori kepadatan penduduk rendah, sedang, dan tinggi. Kluster kepadatan tinggi dinyatakan sebagai C1, kluster kepadatan sedang sebagai C2, dan kluster kepadatan rendah sebagai C3. Penentuan jumlah kluster ini didasarkan pada kebutuhan analisis untuk memperoleh klasifikasi wilayah yang sederhana, mudah dipahami, serta relevan dengan tujuan perencanaan dan pengambilan keputusan pembangunan wilayah. Dengan pembagian tiga kluster tersebut, diharapkan dapat diperoleh gambaran yang jelas mengenai distribusi kepadatan penduduk di Provinsi Jawa Tengah.

Proses clustering menggunakan algoritma *K-Means* diawali dengan penentuan nilai pusat kluster (centroid) awal secara acak untuk masing-masing kluster. Selanjutnya dilakukan perhitungan jarak antara setiap data kepadatan penduduk kabupaten/kota dengan seluruh centroid yang telah ditentukan. Metode pengukuran jarak yang digunakan dalam penelitian ini adalah *Euclidean Distance*, karena metode ini umum digunakan dalam algoritma *K-Means* dan mampu mengukur kedekatan antar data secara geometris. Nilai jarak yang dihasilkan menunjukkan tingkat kedekatan antara data kepadatan penduduk dengan masing-masing kluster.

Setiap data kemudian dikelompokkan ke dalam kluster yang memiliki jarak terdekat terhadap centroid. Setelah seluruh data dialokasikan ke dalam kluster masing-masing, dilakukan pembaruan nilai centroid dengan cara menghitung rata-rata nilai kepadatan penduduk dari seluruh data yang berada dalam kluster tersebut. Proses perhitungan jarak dan pembaruan centroid ini dilakukan secara iteratif dan berulang hingga mencapai kondisi konvergen, yaitu ketika nilai centroid tidak lagi mengalami perubahan yang signifikan atau hasil pengelompokan data sudah stabil. Tabel 3 merupakan tabel perhitungan *Centroid*.

Tabel 3. Tabel Centroid

Centroid	JP	LPP	PP	KP
C1	0.823888	0.83895	0.824075	0.106588
C2	0.466786	0.666523	0.4673	0.062214
C3	0.08714	0.55384	0.08732	0.61966

Tabel 4 menunjukkan hasil pengelompokan kabupaten/kota menggunakan metode *K-Means* berdasarkan variabel JP, LPP, PP, dan KP. Penentuan cluster dilakukan berdasarkan jarak minimum antara data dan pusat cluster (C1, C2, dan C3). Hasil analisis membagi wilayah penelitian ke dalam tiga cluster yang mencerminkan perbedaan karakteristik antar kabupaten/kota.

Tabel 4. Tabel Perhitungan *K-Means*

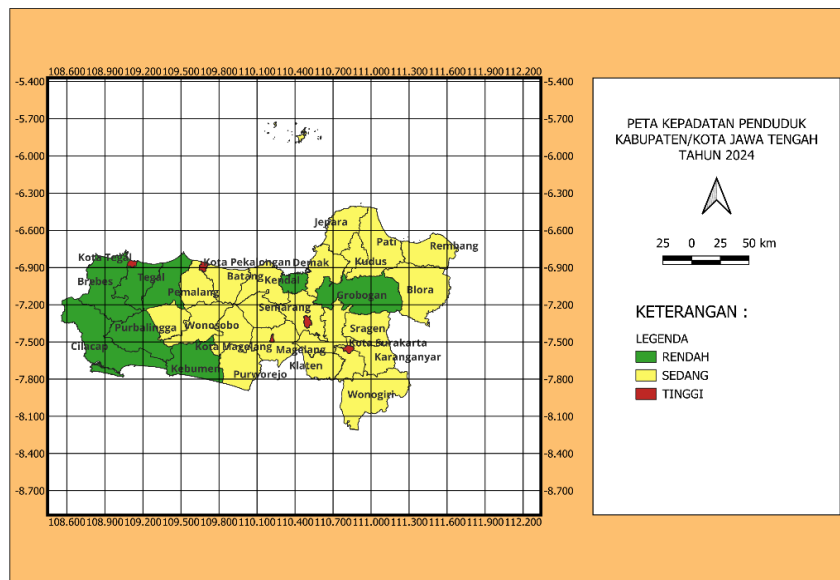
No	Kab/Kota	JP	LPP	PP	KP	C1	C2	C3	Nilai Min	Cluster
1	P1	0.980	0.760	0.981	0.040	0.244	0.732	1.405	0.244	1
2	P2	0.888	0.789	0.887	0.086	0.105	0.607	1.273	0.105	1
3	P3	0.471	0.779	0.472	0.080	0.503	0.114	0.798	0.114	2
4	P4	0.481	0.769	0.482	0.049	0.493	0.106	0.826	0.106	2
5	P5	0.665	0.990	0.665	0.055	0.276	0.428	1.085	0.276	1
6	P6	0.346	0.625	0.347	0.028	0.713	0.179	0.700	0.179	2
7	P7	0.411	0.971	0.411	0.044	0.602	0.315	0.846	0.315	2
8	P8	0.627	0.587	0.628	0.071	0.377	0.241	0.941	0.241	2
9	P9	0.503	0.673	0.503	0.051	0.486	0.052	0.827	0.052	2
10	P10	0.602	0.442	0.602	0.128	0.506	0.302	0.886	0.302	2
11	P11	0.422	0.740	0.423	0.144	0.578	0.127	0.697	0.127	2
12	P12	0.480	0.067	0.480	0.012	0.918	0.602	0.956	0.602	2
13	P13	0.432	0.490	0.433	0.073	0.655	0.183	0.736	0.183	2
14	P14	0.454	0.510	0.454	0.051	0.621	0.158	0.771	0.158	2
15	P15	0.712	0.692	0.714	0.028	0.229	0.350	1.074	0.229	1
16	P16	0.404	0.490	0.406	0.000	0.696	0.206	0.768	0.206	2
17	P17	0.279	0.567	0.281	0.016	0.821	0.286	0.662	0.286	2
18	P18	0.643	0.625	0.643	0.038	0.341	0.253	0.981	0.253	2
19	P19	0.392	0.789	0.392	0.149	0.615	0.184	0.680	0.184	2



No	Kab/Kota	JP	LPP	PP	KP	C1	C2	C3	Nilai Min	Cluster
20	P20	0.572	0.750	0.571	0.071	0.370	0.170	0.899	0.170	2
21	P21	0.582	0.798	0.583	0.086	0.345	0.211	0.914	0.211	2
22	P22	0.498	0.635	0.499	0.063	0.506	0.055	0.809	0.055	2
23	P23	0.356	0.577	0.357	0.044	0.714	0.181	0.691	0.181	2
24	P24	0.485	0.923	0.485	0.055	0.489	0.258	0.878	0.258	2
25	P25	0.368	0.923	0.368	0.048	0.652	0.292	0.789	0.292	2
26	P26	0.461	0.933	0.462	0.070	0.522	0.266	0.852	0.266	2
27	P27	0.730	0.971	0.731	0.085	0.188	0.482	1.135	0.188	1
28	P28	0.799	1.000	0.799	0.114	0.165	0.578	1.211	0.165	1
29	P29	1.000	0.885	1.000	0.065	0.256	0.785	1.443	0.256	1
30	P30	0.000	0.000	0.000	0.567	1.508	1.066	0.570	0.570	3
31	P31	0.209	0.019	0.209	1.000	1.492	1.197	0.678	0.678	3
32	P32	0.041	0.962	0.041	0.296	1.130	0.711	0.525	0.525	3
33	P33	0.816	0.625	0.817	0.380	0.347	0.589	1.061	0.347	1
34	P34	0.102	0.914	0.103	0.613	1.141	0.794	0.360	0.360	3
35	P35	0.084	0.875	0.084	0.623	1.167	0.807	0.321	0.321	3

3.2 Hasil Implementasi Sistem Informasi Geografis

Hasil pengelompokan tingkat kepadatan penduduk yang diperoleh melalui penerapan algoritma K-Means selanjutnya diimplementasikan ke dalam Sistem Informasi Geografis (SIG) menggunakan perangkat lunak QGIS Desktop versi 3.40.5. Tahap implementasi ini diawali dengan proses integrasi data atribut hasil clustering dengan data spasial berupa batas administrasi kabupaten dan kota di Provinsi Jawa Tengah. Proses penggabungan data dilakukan melalui penyesuaian atribut wilayah sehingga setiap satuan wilayah administrasi memiliki informasi tingkat kepadatan penduduk sesuai dengan kluster yang telah dihasilkan. Integrasi data atribut dan data spasial ini memungkinkan penyajian informasi kepadatan penduduk secara spasial dan terstruktur dalam satu sistem pemetaan digital. Gambar 3 menunjukkan peta kepadatan penduduk kabupaten/kota di Provinsi Jawa Tengah tahun 2024, yang menggambarkan variasi tingkat kepadatan penduduk antarwilayah. Perbedaan intensitas warna pada peta mencerminkan konsentrasi penduduk yang berbeda, sehingga memudahkan identifikasi wilayah dengan kepadatan tinggi dan rendah.



Gambar 3. Peta Kepadatan Penduduk Kabupaten/Kota Jawa Tengah Tahun 2024

Setiap kluster kepadatan penduduk divisualisasikan menggunakan simbolisasi dan gradasi warna yang berbeda untuk membedakan tingkat kepadatan antarwilayah. Kluster kepadatan penduduk rendah direpresentasikan dengan warna hijau, kluster kepadatan sedang dengan warna kuning, dan kluster kepadatan tinggi dengan warna merah. Pemilihan simbol dan warna tersebut bertujuan untuk meningkatkan kontras visual serta memudahkan pengguna dalam mengidentifikasi perbedaan tingkat kepadatan penduduk secara cepat dan intuitif. Hasil visualisasi disajikan dalam bentuk peta tematik yang informatif dan mudah dipahami, serta dilengkapi dengan elemen kartografi pendukung, seperti judul peta, legenda, skala, dan orientasi utara. Penambahan elemen kartografi tersebut bertujuan untuk meningkatkan keterbacaan, kejelasan informasi spasial, serta akurasi interpretasi peta, sehingga peta tematik yang dihasilkan tidak hanya berfungsi sebagai media visualisasi, tetapi juga sebagai alat analisis spasial yang mendukung pemahaman dan pengambilan keputusan berbasis wilayah.



3.3 Hasil Analisis Terhadap Penerapan Algoritma *K-Means* Pada Peta

Hasil analisis kepadatan penduduk di Provinsi Jawa Tengah diperoleh melalui penerapan algoritma *K-Means* yang sebelumnya sudah dilakukan normalisasi data dengan tiga kluster, yaitu kluster kepadatan penduduk tinggi, sedang, dan rendah, yang selanjutnya diimplementasikan ke dalam peta tematik berbasis Sistem Informasi Geografis (SIG). Integrasi hasil clustering dengan data spasial menggunakan perangkat lunak QGIS menghasilkan visualisasi yang mampu menggambarkan pola persebaran kepadatan penduduk antar kabupaten dan kota secara spasial dan sistematis. Tabel 4 menyajikan hasil akhir perhitungan *K-Means* yang menunjukkan pembagian kabupaten/kota ke dalam *cluster* berdasarkan kemiripan karakteristik data. Pengelompokan ini mencerminkan perbedaan pola antarwilayah sesuai dengan pusat cluster yang terbentuk.

Tabel 5. Hasil Perhitungan *K-Means*

No	Kab/Kota	JP	LPP	PP	KP	C1	C2	C3	Nilai Min	Cluster
1	P1	0.980	0.760	0.981	0.040	0.244	0.732	1.405	0.244	1
2	P2	0.888	0.789	0.887	0.086	0.105	0.607	1.273	0.105	1
3	P3	0.471	0.779	0.472	0.080	0.503	0.114	0.798	0.114	2
4	P4	0.481	0.769	0.482	0.049	0.493	0.106	0.826	0.106	2
5	P5	0.665	0.990	0.665	0.055	0.276	0.428	1.085	0.276	1
6	P6	0.346	0.625	0.347	0.028	0.713	0.179	0.700	0.179	2
7	P7	0.411	0.971	0.411	0.044	0.602	0.315	0.846	0.315	2
8	P8	0.627	0.587	0.628	0.071	0.377	0.241	0.941	0.241	2
9	P9	0.503	0.673	0.503	0.051	0.486	0.052	0.827	0.052	2
10	P10	0.602	0.442	0.602	0.128	0.506	0.302	0.886	0.302	2
11	P11	0.422	0.740	0.423	0.144	0.578	0.127	0.697	0.127	2
12	P12	0.480	0.067	0.480	0.012	0.918	0.602	0.956	0.602	2
13	P13	0.432	0.490	0.433	0.073	0.655	0.183	0.736	0.183	2
14	P14	0.454	0.510	0.454	0.051	0.621	0.158	0.771	0.158	2
15	P15	0.712	0.692	0.714	0.028	0.229	0.350	1.074	0.229	1
16	P16	0.404	0.490	0.406	0.000	0.696	0.206	0.768	0.206	2
17	P17	0.279	0.567	0.281	0.016	0.821	0.286	0.662	0.286	2
18	P18	0.643	0.625	0.643	0.038	0.341	0.253	0.981	0.253	2
19	P19	0.392	0.789	0.392	0.149	0.615	0.184	0.680	0.184	2
20	P20	0.572	0.750	0.571	0.071	0.370	0.170	0.899	0.170	2
21	P21	0.582	0.798	0.583	0.086	0.345	0.211	0.914	0.211	2
22	P22	0.498	0.635	0.499	0.063	0.506	0.055	0.809	0.055	2
23	P23	0.356	0.577	0.357	0.044	0.714	0.181	0.691	0.181	2
24	P24	0.485	0.923	0.485	0.055	0.489	0.258	0.878	0.258	2
25	P25	0.368	0.923	0.368	0.048	0.652	0.292	0.789	0.292	2
26	P26	0.461	0.933	0.462	0.070	0.522	0.266	0.852	0.266	2
27	P27	0.730	0.971	0.731	0.085	0.188	0.482	1.135	0.188	1
28	P28	0.799	1.000	0.799	0.114	0.165	0.578	1.211	0.165	1
29	P29	1.000	0.885	1.000	0.065	0.256	0.785	1.443	0.256	1
30	P30	0.000	0.000	0.000	0.567	1.508	1.066	0.570	0.570	3
31	P31	0.209	0.019	0.209	1.000	1.492	1.197	0.678	0.678	3
32	P32	0.041	0.962	0.041	0.296	1.130	0.711	0.525	0.525	3
33	P33	0.816	0.625	0.817	0.380	0.347	0.589	1.061	0.347	1
34	P34	0.102	0.914	0.103	0.613	1.141	0.794	0.360	0.360	3
35	P35	0.084	0.875	0.084	0.623	1.167	0.807	0.321	0.321	3

Berdasarkan hasil klusterisasi menggunakan algoritma *K-Means* terhadap data kependudukan Provinsi Jawa Tengah tahun 2024, diperoleh tiga kelompok utama kepadatan penduduk, yaitu kluster kepadatan tinggi (merah), kepadatan sedang (kuning), dan kepadatan rendah (hijau). Kluster kepadatan tinggi terdiri atas 7 kabupaten/kota yang memiliki nilai normalisasi relatif tinggi pada variabel jumlah penduduk (JP) dan persentase penduduk (PP), serta didukung oleh laju pertumbuhan penduduk (LPP) yang cenderung berada pada kategori menengah hingga tinggi. Secara spasial, wilayah dalam kluster ini merepresentasikan kawasan dengan konsentrasi aktivitas ekonomi, pusat pelayanan, dan tingkat urbanisasi yang lebih intensif. Tingginya kontribusi terhadap total populasi provinsi menunjukkan adanya tekanan terhadap ruang dan infrastruktur, sehingga memerlukan perhatian dalam perencanaan tata ruang, penyediaan fasilitas publik, dan pengendalian pertumbuhan kawasan terbangun.

Kluster kepadatan sedang merupakan kelompok terbesar dengan proporsi 62,9% dari total kabupaten/kota. Wilayah dalam kluster ini menunjukkan nilai variabel yang relatif moderat pada seluruh indikator (JP, LPP, PP, dan KP), sehingga mencerminkan kondisi demografis yang lebih seimbang. Secara struktural, kluster ini menggambarkan wilayah transisi yang tidak tergolong sangat padat maupun jarang penduduk. Dominasi kluster sedang mengindikasikan bahwa



sebagian besar wilayah Jawa Tengah memiliki tingkat kepadatan yang relatif homogen dan tidak ekstrem, sehingga kebijakan pembangunan dapat difokuskan pada pemerataan infrastruktur dan peningkatan kualitas pelayanan dasar secara berkelanjutan.

Sementara itu, klaster kepadatan rendah terdiri atas 6 kabupaten/kota dengan karakteristik rasio kepadatan penduduk yang relatif kecil serta distribusi penduduk yang lebih tersebar. Wilayah dalam klaster ini cenderung merepresentasikan daerah dengan karakteristik geografis tertentu, seperti kawasan perdesaan yang luas atau wilayah dengan tingkat aksesibilitas dan konsentrasi aktivitas ekonomi yang lebih rendah. Kondisi tersebut menunjukkan potensi ketersediaan ruang yang lebih besar untuk pengembangan wilayah, namun tetap memerlukan strategi peningkatan konektivitas dan penguatan pusat-pusat pertumbuhan baru.

Secara keseluruhan, hasil klasterisasi berbasis data ter-normalisasi menunjukkan bahwa pendekatan K-Means mampu menghasilkan segmentasi wilayah yang bersifat multidimensi dengan mengintegrasikan empat indikator kependudukan secara simultan. Distribusi klaster yang didominasi kategori sedang, dengan konsentrasi kepadatan tinggi yang terlokalisasi dan sebagian kecil wilayah berkepadatan rendah, mencerminkan struktur demografis Jawa Tengah yang relatif seimbang namun tetap memiliki diferensiasi spasial yang jelas. Visualisasi dalam bentuk peta tematik merah-kuning-hijau semakin mempertegas pola persebaran tersebut dan mempermudah identifikasi wilayah prioritas dalam perencanaan pembangunan berbasis karakteristik kependudukan.

4. KESIMPULAN

Berdasarkan hasil analisis dan penerapan algoritma K-Means berbasis Sistem Informasi Geografis (SIG), penelitian ini berhasil mengelompokkan 35 kabupaten/kota di Provinsi Jawa Tengah ke dalam tiga klaster kepadatan penduduk secara sistematis dan berbasis data yang telah dinormalisasi. Hasil klasterisasi menunjukkan bahwa sebanyak 7 wilayah (20%) termasuk dalam kategori kepadatan tinggi, 22 wilayah (62,86%) berada pada kategori kepadatan sedang, dan 6 wilayah (17,14%) tergolong dalam kategori kepadatan rendah. Distribusi ini mengindikasikan bahwa struktur kepadatan penduduk di Jawa Tengah didominasi oleh wilayah dengan karakteristik transisional, sementara konsentrasi kepadatan tinggi hanya terjadi pada sebagian wilayah tertentu yang memiliki kombinasi jumlah penduduk, persentase penduduk, dan kepadatan yang relatif besar secara simultan, sehingga berpotensi menimbulkan tekanan terhadap ruang dan kebutuhan infrastruktur yang lebih kompleks. Sebaliknya, klaster kepadatan rendah mencerminkan wilayah dengan distribusi penduduk yang lebih tersebar serta potensi pengembangan ruang yang masih terbuka untuk mendukung pemerataan pembangunan. Secara metodologis, penggunaan K-Means dalam penelitian ini tidak hanya mempertimbangkan satu indikator kepadatan penduduk, tetapi mengintegrasikan variabel multidimensi seperti jumlah penduduk, laju pertumbuhan penduduk, dan persentase penduduk yang telah dinormalisasi menggunakan metode Min-Max, sehingga setiap variabel memiliki kontribusi yang proporsional dalam perhitungan jarak Euclidean dan menghasilkan klasterisasi yang lebih objektif secara matematis. Implementasi hasil klasterisasi ke dalam peta tematik berbasis SIG dengan skema warna merah (tinggi), kuning (sedang), dan hijau (rendah) memberikan visualisasi spasial yang informatif, sistematis, dan mudah diinterpretasikan, sehingga mampu mendukung perencanaan pembangunan, penataan ruang, serta penentuan prioritas kebijakan berbasis karakteristik demografis dan pola persebaran penduduk secara komprehensif.

REFERENCES

- [1] P. N. Syukri, L. Pratiwi, M. Arif, E. Suci, A. Kusumawati, and A. Jabir, "Analisis Distribusi Kepadatan Penduduk Di Provinsi Sumatera Utara : Studi Berdasarkan Data Sensus Badan Pusat Statistik Provinsi Sumatera Utara Tahun 2020," *J. Sains Geogr.*, vol. 3, no. 1, pp. 24–32, 2025, doi: 10.2210/jsg.vv1ix.xxx.
- [2] K. Maghfiroh Ainatul, Apriyantika Merli, "Pemetaan Kepadatan Permukiman untuk Mendukung Konservasi Lahan Berkelanjutan Menggunakan SIG (Sistem Informasi Geografis) Tahun 2024," *Indones. J. Conserv.*, vol. 14, no. 2, pp. 1–6, 2025, doi: <https://doi.org/10.15294/ijc.v14i02.37138>.
- [3] R. A. Putri and Z. Fitri, "Penerapan Algoritma K-Means Clustering untuk Segmentasi Kepadatan Penduduk Berbasis GIS," *JURIKOM (Jurnal Ris. Komputer)*, vol. 12, no. 3, pp. 298–308, 2025, doi: 10.30865/jurikom.v12i3.8668.
- [4] W. Widodo, "Peta Digital Kependudukan Berbasis QGIS : Inovasi Teknologi untuk Visualisasi Data kependudukan dan Populasi di Dusun Mrisi , Bantul," *J. Surya Masy.*, vol. 6, no. 2, pp. 140–149, 2024, doi: <https://doi.org/10.26714/jsm.6.2.2024.140-149>.
- [5] R. F. M. Maftuh Naufal, Nursanto Ari Gunawan, "Systematic Literature Review: Population Density Mapping Using Data Mining," *J. Dinda Data Sci. Inf. Technol. Data Anal.*, vol. 5, no. 2, pp. 5–8, 2025, doi: <https://doi.org/10.20895/dinda.v5i2.1805>.
- [6] A. Lareina, N. Permata, and F. S. Rahman, "Analisis Spasial Persebaran Kepadatan Penduduk, Usia Produktif, dan Kasus HIV/AIDS di Kabupaten Gresik Tahun 2021-2023," *J. Kesmas Untika Luwuk Public Heal. J.*, vol. 16, 2025, doi: <https://doi.org/10.51888/phj.v16i2.368>.
- [7] A. Surachman, "Analisis Spasial Permukiman Berbasis Sistem Informasi Geografis menggunakan QGIS dan OpenStreetMap," *J. Esensi Infokom*, vol. 8, no. 1, pp. 49–54, 2024, doi: 10.55886/infokom.v8i1.852.
- [8] A. M. Nurfitri Andayani, Wimmy Hartawan, "Perancangan Sistem Pemetaan Wilayah Calon Pelanggan Dengan Menggunakan QGIS Pada PT.Indonesia Comnets Plus (ICON+) SBU Bengkulu," *J. Inform.*, vol. 1, no. 2, pp. 1–12, 2022, doi: <https://doi.org/10.57094/ji.v1i2.357>.
- [9] A. K. Sari *et al.*, "Penerapan Sistem Informasi Geografis Menggunakan QGIS dalam Menganalisis Sebaran Pemukiman di Samarinda Tahun 2024," *Kreat. Teknol. dan Sist. Inf.*, vol. 3, no. 1, pp. 32–37, 2025, doi: <https://doi.org/10.30872/kretisi.v3i1.2123>.
- [10] A. H. A. Karsa and A. R. Hidayat, "Metode Algoritma K-Means Untuk Clustering Data Produk Paling Laku Pada Toko Tono Grosir Plumbon Cirebon," *Syntax Lit. ; J. Ilm. Indones.*, vol. 7, no. 9, pp. 15984–15996, 2024, doi: 10.36418/syntax-



- literate.v7i9.15144.
- [11] L. L. Latifah, S. A. Hudjimartsu, and I. Yanuarsyah, "Analisis Daerah Rawan Kecelakaan Lalu Lintas Menggunakan Cluster Analysis Di Kota Bogor Berbasis Webgis," *J. Ilm. Teknol. Infomasi Terap.*, vol. 8, no. 2, pp. 235–244, 2022, doi: 10.33197/jitter.vol8.iss2.2022.760.
 - [12] S. Frisma Handayanna, "Penerapan Algoritma K-Means Untuk Mengelompokkan Kepadatan," *J. Appl. Comput. Sci. Technol.*, vol. 5, no. 1, pp. 50–55, 2024, doi: <https://doi.org/10.52158/jacost.v5i1.477>.
 - [13] P. Marpaung, I. Pebrian, and W. Putri, "Penerapan Data Mining Untuk Pengelompokan Kepadatan Penduduk Kabupaten Deli Serdang Menggunakan Algoritma K-Means," *J. Ilmu Komput. dan Sist. Inf.*, vol. 6, no. 2, pp. 64–70, 2023, doi: <https://doi.org/10.55338/jikoms.v6i2>.
 - [14] F. Delia *et al.*, "Penerapan Algoritma K-Means Clustering Dalam Pengelompokan Kepadatan Penduduk," *Technomedia J.*, vol. 9, no. 3, pp. 373–386, 2025, doi: <https://doi.org/10.33050/tmj.v9i3.2270>.
 - [15] W. A. M. Malonga, "Pemetaan Distribusi Penduduk Kabupaten Sumbawa Dengan Metode Sistem Grid Skala Ragam," *J. TAMBORA*, vol. 7, no. 2, pp. 1–7, 2023, doi: <https://doi.org/10.36761/jt.v7i2.2985>.
 - [16] A. P. Risti Wirantika and I. Veriansyah, "Pemanfaatan Sistem Informasi Geografis (SIG) Untuk Pemetaan Kepadatan Penduduk Di Kecamatan Sungai Kakap Tahun 2015-2020," *Geo Khatulistiwa J. Pendidik. Geogr. dan Pariwisata*, vol. 3, no. 1, pp. 10–17, 2023, doi: <https://doi.org/10.31571/v3i1.174>.
 - [17] O. S. Putri, L. L. Sitohang, and S. P. Prasetya, "Visualisasi dan Analisis Distribusi Spasial Jumlah Penduduk Miskin Jawa Timur Berbasis SIG dengan Teknologi QGIS," *J. Pendidik. Ris. dan Konseptual*, vol. 9, no. 4, pp. 824–834, 2025, doi: https://doi.org/10.28926/riset_konseptual.v9i4.1339.
 - [18] R. P. Aryanto, A. Nilogiri, and A. E. Wardoyo, "Klasterisasi Jumlah Penduduk Provinsi Jawa Timur Tahun 2021- 2023 Menggunakan Algoritma K-Means," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 9, no. 2, pp. 134–146, 2024, doi: <https://doi.org/10.14421/jiska.2024.9.2.134-146>.
 - [19] L. Rahmawati *et al.*, "Pengembangan Sistem Informasi Geografis (SIG) Untuk Analisis Spasial Dalam Pengambilan Keputusan," *J. Rev. Pendidik. dan Pengajaran*, vol. 7, no. 2, pp. 4058–4068, 2024, doi: <https://doi.org/10.31004/jrpp.v7i2.26929>.
 - [20] D. Moh. Erkamini, Iqbal Ramadhani Mukhlis, Putra, *Sistem Informasi Geografis (SIG)*. Daerah Istimewa Yogyakarta: PT. Green Pustaka Indonesia, 2023.
 - [21] J. T. Santoso, *Sistem informasi geografis*. Semarang: Yayasan Prima Agus Teknik Universitas Sains & Teknologi Komputer (Universitas STEKOM), 2021.