

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Kepadatan penduduk merupakan salah satu aspek penting dalam perencanaan dan pengelolaan pembangunan daerah. Tingginya kepadatan di suatu wilayah dapat mencerminkan aktivitas ekonomi, sosial, dan urbanisasi yang dinamis. Namun, kepadatan yang tidak terkendali berpotensi menimbulkan berbagai permasalahan, seperti tekanan terhadap infrastruktur, kemacetan, penurunan kualitas lingkungan, serta kurangnya layanan publik yang memadai. Sebagai salah satu kota utama di Provinsi Aceh, Kota Lhokseumawe mengalami peningkatan jumlah penduduk yang signifikan dalam beberapa tahun terakhir. Pertumbuhan ini mendorong perlunya analisis yang lebih mendalam terkait distribusi dan segmentasi kepadatan penduduk, terutama di setiap kecamatan.

Pemahaman yang baik mengenai pola distribusi penduduk dapat membantu pemerintah daerah menyusun kebijakan berbasis data yang tepat sasaran. Informasi ini penting untuk perencanaan pembangunan, seperti pengalokasian anggaran infrastruktur, perbaikan layanan transportasi publik, penyediaan fasilitas kesehatan, dan pengembangan ruang terbuka hijau. Sebagai contoh, wilayah dengan kepadatan penduduk tinggi memerlukan intervensi cepat dalam bentuk fasilitas umum dan infrastruktur yang memadai, sementara wilayah berpenduduk jarang dapat dioptimalkan untuk meningkatkan aktivitas ekonomi dan pemberdayaan komunitas. Oleh karena itu, pemetaan dan analisis kepadatan penduduk sangat diperlukan untuk mendukung pembangunan kota yang berkelanjutan dan merata.

Dengan kemajuan teknologi informasi, pendekatan tradisional dalam menganalisis data kependudukan kini dianggap kurang efektif. Penerapan metode *data mining* menjadi solusi yang relevan karena mampu mengolah dan menggali informasi dari data dalam jumlah besar dan kompleks. Salah satu teknik yang umum digunakan dalam *data mining* adalah *clustering*, atau pengelompokan, yang bertujuan untuk mengidentifikasi kelompok-kelompok dengan karakteristik serupa.

*Clustering* melibatkan pengelompokan data atau objek ke dalam kluster-kluster, di mana setiap kluster terdiri dari data yang memiliki kesamaan tinggi dan berbeda dari objek di kluster lainnya. Dalam konteks ini, terdapat dua metode pengelompokan yang terkenal, yaitu partisi dan hierarkis.

*Clustering* bertujuan untuk mengidentifikasi kelas-kelas taksonomi atau batryologi, atau untuk melakukan analisis topologi terhadap data yang ada. Penting untuk dicatat bahwa dalam *data mining*, *clustering* bukanlah proses klasifikasi, karena dalam klasifikasi, data dikelompokkan ke dalam kelas-kelas yang sudah diketahui sebelumnya. Konsep *centroid* dalam *clustering* sangat penting untuk menilai keanggotaan dan koneksi lokal dalam kluster (Zufria & Damaiani Iskandar, 2024).

Salah satu teknik *clustering* yang paling sederhana dan umum digunakan adalah Algoritma *K-means*. Algoritma ini mengelompokkan objek berdasarkan jarak, dan memiliki ketelitian yang cukup tinggi terhadap ukuran objek. Hal ini membuat Algoritma *K-means* relatif lebih terukur dan efisien dalam proses pengelompokannya (Veronica et al., 2023). *K-means* juga menjadi salah satu metode *clustering* non-hirarki yang bertujuan untuk mempartisi data ke dalam satu atau lebih kluster atau kelompok. Metode ini mengelompokkan data yang memiliki karakteristik yang sama ke dalam kluster yang sama, sementara data dengan karakteristik yang berbeda dikelompokkan ke dalam kluster yang berbeda. Namun, *K-means* memiliki kelemahan yang disebabkan oleh penentuan pusat awal kluster. Hasil *clustering* dari metode *K-means* sangat bergantung pada nilai awal yang dipilih untuk pusat-pusat kluster (Butsianto & Mayangwulan, 2020).

Seperti penelitian yang dilakukan oleh Matius Tadi, Besse Arnawisuda Ningsi (2023), “Analisis Klaster Kemiskinan Kabupaten Kota Di Provinsi Banten Menggunakan Metode *K-means*”. Berdasarkan hasil analisis, pengklasteran kemiskinan kabupaten/kota di Provinsi Banten menggunakan metode *K-means* menghasilkan dua klaster dengan nilai struktur terbaik berdasarkan *silhouette index*. Klaster pertama terdiri dari Kota Tangerang, Kota Cilegon, dan Kota Tangerang Selatan, dengan sektor kemiskinan yang berpengaruh adalah persentase penduduk miskin usia 15 tahun ke atas yang bekerja di sektor informal (X6). Klaster kedua mencakup Kabupaten Pandeglang, Kabupaten Lebak, Kabupaten Tangerang, Kabupaten Serang, dan Kota Serang, dengan faktor kemiskinan yang berpengaruh adalah persentase penduduk miskin usia 15 tahun ke atas dengan pendidikan di bawah SD (X1), tamat SD/SMP (X2), tamat SMA ke atas (X3), angka melek huruf usia 15-55 tahun (X4), persentase penduduk miskin usia 15 tahun ke atas yang bekerja di sektor formal (X7), dan persentase pengeluaran per kapita untuk makanan (X10). Pembentukan dua klaster ini menunjukkan akurasi terbaik dibandingkan dengan pembentukan tiga atau empat klaster.

Penelitian yang serupa juga dilakukan oleh Indra Ferdiansyah, Baenil Huda, Agustia Hananto, Tukino (2024), "Analisis *Clustering* Menggunakan Metode *K-means* Pada Kemiskinan Di Jawa Timur Tahun 2020". Penggunaan metode *K-means* untuk *clustering* kemiskinan di Jawa Timur tahun 2020 mengidentifikasi dan mengklasifikasikan kota/kabupaten berdasarkan tingkat kemiskinan tertinggi, menengah, dan terendah. Dengan menggunakan *tools Orange*, hasil *clustering* menghasilkan tiga klaster: klaster 1 terdiri dari 6 kota/kabupaten dengan tingkat kemiskinan terendah sebesar 15%, klaster 2 mencakup 10 kota/kabupaten dengan tingkat kemiskinan menengah sebesar 26%, dan klaster 3 terdiri dari 23 kota/kabupaten dengan tingkat kemiskinan tertinggi sebesar 59%.

Dalam penelitian ini, algoritma *K-means Clustering* diterapkan untuk segmentasi kepadatan penduduk di Kota Lhokseumawe yang mencakup 4 Kecamatan diantaranya, Blang Mangat, Muara Dua, Muara Satu, Banda Sakti, dan Lhokseumawe, dengan menggunakan data geospasial berbasis *Geographic Information System* (GIS). Penggunaan GIS memberikan nilai tambah dalam visualisasi dan analisis data karena mampu memetakan dan menampilkan distribusi penduduk secara interaktif dan lebih informatif. Setiap kecamatan akan dikelompokkan berdasarkan tingkat kepadatan penduduk, sehingga menghasilkan *cluster-cluster* yang menggambarkan pola distribusi penduduk di seluruh kota.

Penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam mengenai karakteristik kepadatan penduduk di Kota Lhokseumawe, sekaligus mendukung pemerintah dalam mengambil keputusan yang lebih tepat dalam perencanaan pembangunan. Hasil segmentasi juga berperan dalam mempermudah pemerintah untuk menetapkan prioritas pembangunan infrastruktur dan layanan publik di setiap wilayah. Selain itu, penelitian ini dapat berkontribusi dalam penerapan teknik data mining dan GIS pada analisis kependudukan, serta menjadi referensi bagi kota lain yang menghadapi tantangan serupa dalam pengelolaan kepadatan penduduk.

Berdasarkan paparan di atas, penelitian ini akan menggunakan judul "**PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK SEGMENTASI KEPADATAN PENDUDUK DI KOTA LHOKSEUMAWE BERBASIS GIS**".

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang diuraikan di atas, maka permasalahan yang penulis rumuskan adalah :

1. Bagaimana menerapkan metode *K-means* dalam *clustering* untuk segmentasi kepadatan penduduk di Kota Lhokseumawe?
2. Bagaimana merancang sistem *clustering* menggunakan metode *K-means* untuk segmentasi kepadatan penduduk dengan penerapan GIS?

### **1.3 Tujuan Penelitian**

Berdasarkan latar belakang dan rumusan masalah di atas maka tujuan penelitian ini adalah sebagai berikut :

1. Menerapkan metode *K-means* dalam *clustering* untuk segmentasi kepadatan penduduk di Kota Lhokseumawe.
2. Merancang sistem *clustering* menggunakan metode *K-means* untuk segmentasi kepadatan penduduk dengan penerapan GIS.

### **1.4 Manfaat Penelitian**

Adapun manfaat dari penelitian ini adalah sebagai berikut :

1. Penelitian ini memperluas wawasan tentang penerapan teknik data mining dalam analisis kependudukan dan memberikan studi kasus nyata di Kota Lhokseumawe.
2. Hasil segmentasi dan visualisasi berbasis GIS dapat membantu pemerintah daerah Kota Lhokseumawe dalam merencanakan dan mengalokasikan sumber daya dengan lebih efektif, misalnya untuk pengembangan infrastruktur, layanan publik, dan fasilitas umum.
3. Penelitian ini dapat menjadi model atau referensi bagi kota-kota lain yang menghadapi tantangan serupa dalam mengelola kepadatan penduduk dan membutuhkan metode analisis yang berbasis data.
4. Sebagai sarana untuk menerapkan pengetahuan peneliti yang diperoleh selama menempuh masa studi Strata-1 Teknik Informatika Universitas Malikussaleh.

### **1.5 Ruang Lingkup Dan Batasan Masalah**

Adapun batasan masalah yang ditetapkan pada penelitian ini adalah sebagai berikut :

1. Penelitian ini hanya mencakup wilayah administratif Kota Lhokseumawe, yang terdiri dari kecamatan-kecamatan seperti Blang Mangat, Muara Dua, Muara Satu, Banda Sakti. Data di luar wilayah ini tidak termasuk dalam analisis.

2. Penelitian menggunakan data kependudukan tahun 2020-2023 yang diperoleh dari Badan Pusat Statistik (BPS) Kota Lhokseumawe. Variabel yang digunakan untuk segmentasi adalah jumlah penduduk, laju pertumbuhan penduduk, kepadatan penduduk per km<sup>2</sup>, dan persentase penduduk.
3. Metode pengelompokan yang diterapkan adalah *K-means Clustering*, tanpa membandingkan hasilnya dengan metode *clustering* lain.