**SEMINAR PROGRES**

**Tugas Akhir Semester Genap 2024/2025**



**PENERAPAN METODE K-MEANS UNTUK KLASTERISASI DAERAH DENGAN RISIKO OVERPOPULASI DI INDONESIA**

**Disusun oleh:**

**Ilham Maulana Prasetyo**

**(2118064)**

**PROGRAM STUDI TEKNIK INFORMATIKA S-1**

**FAKULTAS TEKNOLOGI INDUSTRI**

**INSTITUT TEKNOLOGI NASIONAL MALANG**

**2025**

# DAFTAR ISI

[DAFTAR ISI ii](#_Toc196292347)

[DAFTAR GAMBAR iv](#_Toc196292348)

[DAFTAR TABEL v](#_Toc196292349)

[BAB I PENDAHULUAN 1](#_Toc196292350)

[1.1 Latar Belakang 1](#_Toc196292351)

[1.2 Rumusan Masalah 2](#_Toc196292352)

[1.3 Batasan Masalah 3](#_Toc196292353)

[1.4 Tujuan 3](#_Toc196292354)

[1.5 Manfaat 3](#_Toc196292355)

[1.6 Sistematika Penulisan 4](#_Toc196292356)

[BAB II TINJAUAN PUSTAKA 5](#_Toc196292357)

[2.1 Penelitian Terkait 5](#_Toc196292358)

[2.2 Data Mining 7](#_Toc196292359)

[2.3 K-Means 8](#_Toc196292360)

[2.4 Overpopulasi 9](#_Toc196292361)

[2.5 Wilayah Administratif 9](#_Toc196292362)

[2.6 Website 10](#_Toc196292363)

[2.7 Database 10](#_Toc196292364)

[2.8 Laravel 10](#_Toc196292365)

[2.9 PHP 10](#_Toc196292366)

[BAB III ANALISIS DAN PERANCANGAN 11](#_Toc196292367)

[3.1 Analisis Kebutuhan 11](#_Toc196292368)

[3.1.1 Kebutuhan Fungsional 11](#_Toc196292369)

[3.1.2 Kebutuhan Non Fungsional 11](#_Toc196292370)

[3.2 Tabel Kriteria 12](#_Toc196292375)

[3.3 Diagram Blok Sistem 12](#_Toc196292376)

[3.4 Struktur Menu 13](#_Toc196292377)

[3.5 Use Case Diagram 13](#_Toc196292378)

[3.6 Flowchart 14](#_Toc196292379)

[3.6.1 Flowchart Sistem Klasterisasi Daerah Risiko Overpopulasi 14](#_Toc196292380)

[3.6.2 Flowchart Metode K-Means 15](#_Toc196292381)

[3.7 Data Klaster 16](#_Toc196292382)

[3.8 Desain Website 17](#_Toc196292383)

[BAB IV PROGRESS YANG TELAH DILAKUKAN 21](#_Toc196292384)

[4.1 Hasil Sampai Saat Ini (berisi gambar dan screenshot serta penjelasan progress yang sudah dilakukan) 21](#_Toc196292385)

[4.2 Kendala yang dihadapi 21](#_Toc196292386)

[4.3 Rencana Penyelesaian Masalah 21](#_Toc196292387)

[DAFTAR PUSTAKA 22](#_Toc196292388)

# DAFTAR GAMBAR

[Gambar 3.1 Diagram Blok Sistem 12](#_Toc196292389)

[Gambar 3.2 Struktur Menu 13](#_Toc196292390)

[Gambar 3.3 *Use* *Case* Diagram 13](#_Toc196292391)

[Gambar 3.4 *Flowchart* Sistem 14](#_Toc196292392)

[Gambar 3.5 Flowchart Metode K-Means 15](#_Toc196292393)

[Gambar 3.6 Halaman *Login* 17](#_Toc196292394)

[Gambar 3.7 Menu Dashboard 17](#_Toc196292395)

[Gambar 3.8 Menu Data Wilayah 18](#_Toc196292396)

[Gambar 3.9 Menu Data Populasi 18](#_Toc196292397)

[Gambar 3.10 Menu Proses *Clustering* 19](#_Toc196292398)

[Gambar 3.11 Menu Hasil *Clustering* 19](#_Toc196292399)

[Gambar 3.12 Menu Pemetaan 20](#_Toc196292400)

# DAFTAR TABEL

[Tabel 3.1 Tabel Kriteria 12](#_Toc196292401)

[Tabel 3.2 Tabel Klaster pada Sistem Klasterisasi Daerah Risiko Overpopulasi 16](#_Toc196292402)

# PENDAHULUAN

## Latar Belakang

Indonesia merupakan salah satu negara dengan jumlah penduduk terbesar di dunia, dengan angka pertumbuhan yang terus meningkat setiap tahun. Berdasarkan data Badan Pusat Statistik (BPS), jumlah penduduk Indonesia mencapai 281 juta jiwa pada tahun 2024, dengan konsentrasi populasi yang tidak merata (BPS, 2024). Beberapa daerah, terutama kota-kota besar seperti Jakarta, Surabaya, dan Bandung, menghadapi overpopulasi yang memicu berbagai permasalahan sosial, ekonomi, dan lingkungan. Overpopulasi dapat menyebabkan penurunan kualitas lingkungan, peningkatan polusi, serta tekanan pada sumber daya alam dan layanan publik (Talibi et al., 2022). Selain itu, kepadatan populasi yang tidak terkendali di kota-kota besar Indonesia meningkatkan risiko kesehatan dan memperburuk kondisi hidup masyarakat akibat sanitasi yang buruk dan akses layanan kesehatan yang terbatas (Temalagi et al., 2023). Pertumbuhan populasi yang tidak terkendali juga dapat meningkatkan angka pengangguran, kemiskinan, dan kesenjangan sosial (Milanda et al., 2024).

Untuk mengatasi permasalahan tersebut, peneliti menerapkan metode K-Means untuk mengelompokkan daerah di Indonesia yang memiliki risiko overpopulasi. Ini adalah solusi untuk membantu pemerintah seperti Badan Pusat Statistik (BPS) memahami pola persebaran penduduk dan merancang strategi mitigasi yang lebih efektif. Dalam beberapa penelitian terdahulu, metode K-Means *Clustering* telah digunakan untuk mengelompokkan data berdasarkan karakteristik yang sama ke dalam *cluster* yang sama. Beberapa penelitian terdahulu diantaranya, penelitian yang dilakukan oleh Iqbal et al. (2023) menunjukkan bahwa penerapan algoritma K-Means dengan jarak *Euclidean* efektif dalam mengelompokkan daerah penyebaran COVID-19 di Kabupaten Bogor, dengan mempertimbangkan variabel seperti kepadatan penduduk dan jumlah kasus terinfeksi (Iqbal et al., 2023). Penelitian lain oleh Bahauddin et al. (2021) menerapkan klasterisasi menggunakan algoritma K-Means untuk mengelompokkan provinsi di Indonesia ke dalam beberapa klaster berdasarkan tingkat kemiskinan guna membantu pemerintah dalam mengatasi kemiskinan yang terjadi (Bahauddin et al., 2021).

Penerapan metode K-Means *Clustering* memiliki beberapa kelebihan, di antaranya efisien dalam menangani data besar serta mengelompokkan wilayah dengan karakteristik serupa secara cepat. Terbukti dari hasil penelitian yang menunjukkan bahwa K-Means efektif dalam klasifikasi data pelanggan *e-commerce* di Indonesia (Munawar et al., 2021). Namun, metode K-Means juga memiliki kelemahan yaitu sensitivitas terhadap pemilihan jumlah klaster (*k*) dan inisialisasi awal pusat klaster (*centroid*), yang dapat mempengaruhi hasil klasterisasi. Penelitian oleh Sudarwati et al. (2022) menemukan bahwa penerapan K-Means dalam klasterisasi industri rumput laut di Indonesia diperoleh hasil yang bervariasi tergantung pada inisialisasi awal *centroid* (Sudarwati et al., 2022).

Berdasarkan pertimbangan tersebut, penelitian ini bertujuan untuk menerapkan metode K-Means *Clustering* dalam klasterisasi daerah dengan risiko overpopulasi di Indonesia. Hasil dari klasterisasi ini diharapkan dapat diperoleh pemetaan yang lebih jelas mengenai wilayah yang berpotensi mengalami kepadatan penduduk ekstrem, sehingga dapat menjadi dasar bagi pemerintah dalam merancang strategi mitigasi yang lebih efektif. Klasterisasi ini juga dapat digunakan sebagai alat bantu mencegah dampak negatif dari kepadatan penduduk yang berlebihan terhadap lingkungan dan kesejahteraan masyarakat.

## Rumusan Masalah

1. Bagaimana cara mengembangkan aplikasi berbasis website untuk klasterisasi daerah dengan risiko overpopulasi di Indonesia?
2. Bagaimana menerapkan algoritma K-Means dalam aplikasi berbasis website untuk klasterisasi daerah dengan risiko overpopulasi di Indonesia?
3. Bagaimana merancang dan mengimplementasikan sistem visualisasi hasil klasterisasi agar mudah dipahami oleh pengguna aplikasi?

## Batasan Masalah

1. Penelitian ini tidak akan membahas dampak sosial-ekonomi atau perubahan kebijakan akibat overpopulasi, melainkan lebih fokus pada identifikasi dan klasterisasi daerah yang berisiko.
2. Penelitian ini terbatas pada penggunaan algoritma K-Means untuk klasterisasi daerah dengan risiko overpopulasi.
3. Seluruh proses dalam penelitian ini, mulai dari pengolahan data hingga visualisasi hasil klasterisasi, akan dikembangkan dalam bentuk aplikasi berbasis web menggunakan bahasa pemrograman PHP dan *framework* Laravel.
4. Data yang digunakan dalam penelitian ini terbatas pada data wilayah administratif seperti provinsi atau kabupaten/kota di Indonesia, dengan rentang waktu antara tahun 2021 hingga 2024.

## Tujuan

1. Mengembangkan aplikasi berbasis website yang dapat digunakan untuk klasterisasi daerah dengan risiko overpopulasi di Indonesia.
2. Menerapkan algoritma K-Means secara efektif dalam sistem aplikasi web untuk mengelompokkan daerah berdasarkan tingkat risiko overpopulasi.
3. Merancang *user* *interface* yang interaktif dan mudah diakses untuk membantu pengguna memahami hasil klasterisasi daerah dengan risiko overpopulasi di Indonesia.

## Manfaat

1. Mengetahui daerah-daerah di Indonesia yang berisiko tinggi mengalami overpopulasi melalui klasterisasi menggunakan metode K-Means.
2. Memberikan informasi yang berguna bagi pemerintah dan pemangku kepentingan dalam merencanakan kebijakan terkait pengendalian pertumbuhan penduduk.
3. Menyediakan dasar analisis untuk tindakan yang dapat diambil guna mengurangi tekanan terhadap sumber daya alam dan kualitas hidup di daerah-daerah yang rentan terhadap overpopulasi.

## Sistematika Penulisan

Agar mempermudah pemahaman pada pembahasan penulisan skripsi ini, maka sistematika penulisan diperoleh sebagai berikut:

**BAB I**: Pendahuluan berisikan latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat, metode penelitian, dan sistematika penelitian.

**BAB II**: Tinjauan Pustaka berisikan dasar – dasar teori mengenai permasalahan yang berhubungan dengan penelitian ini.

**BAB III**: Analisis dan Perancangan pada Sistem berisikan perancangan pada sistem yang menggunakan *flowchart* dan desain struktur menu pada sistem.

**BAB IV**: Progress Yang Telah Dilakukan berisikan hasil kemajuan penelitian, kendala yang dihadapi, dan rencana penyelesaian masalah.

# TINJAUAN PUSTAKA

## Penelitian Terkait

Menurut Handayanna dan Sunarti (2024) dalam penelitiannya yang berjudul “Penerapan Algoritma K-Means Untuk Mengelompokkan Kepadatan Penduduk Di Provinsi DKI Jakarta”, bertujuan untuk mengelompokkan wilayah di DKI Jakarta berdasarkan jumlah kepadatan penduduk selama periode 2019-2022. Dengan menggunakan metode K-Means *Clustering* dan aplikasi RapidMiner, penelitian ini mengidentifikasi tiga wilayah dengan tingkat kepadatan tertinggi, yaitu Jakarta Selatan, Jakarta Timur, dan Jakarta Barat, yang terus mengalami peningkatan jumlah penduduk. Hasil klasterisasi ini memberikan wawasan bagi pemerintah dalam merancang kebijakan pengendalian populasi serta membantu masyarakat dalam mempertimbangkan lokasi tempat tinggal yang sesuai (Handayanna & Sunarti, 2024).

Menurut Aisy dan Kurniawan (2024) dalam penelitiannya yang berjudul “Klasterisasi Data Bencana Alam Di Kota Cirebon Menggunakan K-Means *Clustering* Berdasarkan Kawasan Dan Jenis Bencana”, bertujuan untuk mengidentifikasi dan mengelompokkan wilayah di Kota Cirebon yang rentan terhadap bencana berdasarkan jenis bencana yang terjadi. Dengan menggunakan metode K-Means *Clustering*, penelitian ini membagi wilayah Kota Cirebon ke dalam enam klaster dengan nilai *Davies-Bouldin Index* (DBI) sebesar 0.087, yang menunjukkan efektivitas pengelompokan tersebut. Hasil penelitian ini membantu Badan Penanggulangan Bencana Daerah (BPBD) dalam mengetahui kawasan yang paling rentan terhadap bencana serta menentukan strategi mitigasi yang lebih tepat (Aisy & Kurniawan, 2024).

Menurut Maulana et al. (2024) dalam penelitiannya yang berjudul “Implementasi Algoritma K-Means *Clustering* Dalam Pengelompokan Data Kerusakan Rumah Akibat Bencana Alam di Kabupaten Cirebon”, bertujuan untuk mengelompokkan tingkat kerusakan rumah akibat bencana alam berdasarkan jenis kerusakan. Dengan menggunakan pendekatan *Knowledge Discovery in Databases* (KDD) dan algoritma K-Means *Clustering*, penelitian ini menghasilkan dua klaster utama dengan nilai *Davies-Bouldin Index* (DBI) optimal sebesar 0,025. Hasil klasterisasi ini memberikan informasi mengenai distribusi tingkat kerusakan rumah, yang dapat membantu dalam perencanaan rekonstruksi dan mitigasi dampak bencana (Maulana et al., 2024).

Menurut Nabiilah et al. (2023) dalam penelitiannya yang berjudul “Penerapan Algoritma K-Means dan K-Medoids Dalam Pengelompokan Kepadatan Penduduk Provinsi Riau”, bertujuan untuk menganalisis distribusi penduduk yang tidak merata di Riau dan menentukan metode klasterisasi yang paling efektif dalam mengelompokkan wilayah berdasarkan tingkat kepadatan penduduk. Data yang digunakan dalam penelitian ini bersumber dari Badan Pusat Statistik (BPS) dengan rentang tahun 2010-2021 dan mencakup 12 kabupaten di Provinsi Riau. Penelitian ini membandingkan algoritma K-Means dan K-Medoids dalam proses klasterisasi, di mana hasil menunjukkan bahwa algoritma K-Means memiliki kinerja yang lebih baik dengan nilai *Davies-Bouldin Index* (DBI) sebesar 0.001 pada percobaan K=2. Temuan ini mengindikasikan bahwa K-Means lebih optimal dalam mengelompokkan data kepadatan penduduk dibandingkan K-Medoids (Nabiilah et al., 2023).

Menurut Sujak et al. (2025) dalam penelitiannya yang berjudul “Implementasi K-Means *Clustering* untuk Optimalisasi Anggaran Penyakit Tidak Menular”, bertujuan untuk mengelompokkan anggaran kesehatan yang dialokasikan untuk penyakit komorbid seperti diabetes melitus, hipertensi, obesitas, serta gangguan mental emosional. Dengan menggunakan metode K-Means *Clustering*, penelitian ini menghasilkan empat klaster dengan *silhouette score* sebesar 0,6156, yang memberikan wawasan mengenai pola pengeluaran anggaran kesehatan oleh pemerintah daerah. Hasil klasterisasi ini menunjukkan adanya potensi optimalisasi dana untuk dialokasikan ke sub-kegiatan lain atau sebagai dana darurat pandemi (Sujak et al., 2025).

Penelitian yang dilakukan oleh Nurhayah et al. (2024) dalam studinya yang berjudul “Analisis Jumlah Penduduk Menggunakan Algoritma K-Means Berdasarkan Kabupaten/Kota Di Indonesia”, penelitian ini bertujuan untuk mengelompokkan jumlah penduduk di berbagai Kabupaten/Kota di Indonesia menggunakan metode K-Means *Clustering*. Data yang digunakan dalam penelitian ini diperoleh dari Badan Pusat Statistik (BPS) yang mencakup informasi jumlah penduduk dari tahun 2020 sampai dengan 2022. Hasil analisis pengelompokan jumlah penduduk Kabupaten/Kota di Indonesia menggunakan metode K-Means diperoleh empat *cluster*, yaitu *cluster* dengan jumlah penduduk sangat padat, padat, sedang, dan rendah. Hasil penelitian ini dari pengujian menggunakan *Davies Bouldin Index* (DBI) menunjukkan bahwa *cluster* yang optimal diperoleh pada k=4 dengan nilai DBI sebesar 0.127, di mana *cluster* dengan jumlah penduduk sangat padat hanya mencakup 1 Kabupaten, sedangkan *cluster* dengan jumlah penduduk rendah mencakup 360 Kabupaten/Kota (Nurhayah et al., 2024).

## Data Mining

*Data Mining* adalah proses penggalian informasi yang berharga dari kumpulan data yang besar dan kompleks. Tujuan utama dari *data* *mining* adalah mengidentifikasi pola, hubungan, atau informasi yang tersembunyi dalam kumpulan data yang besar atau kompleks. Manfaat dari *data mining* adalah membantu mengambil keputusan yang lebih baik, mengidentifikasi pola dan tren yang tidak terlihat secara langsung, memprediksi masa depan, segmentasi pelanggan, efisiensi operasional, pengelolaan risiko, pemasaran yang lebih efektif, dan peningkatan pelayanan pelanggan (Rahayu et al., 2024). Menurut (Wahono, 2023), *data mining* merupakan salah satu proses di dalam *Knowledge Discovery in Database* (KDD), yaitu suatu proses pencarian pengetahuan yang bermanfaat dari data. KDD terdiri dari beberapa tahapan atau langkah yaitu:

1. *Data Selection*, tahap ini menyeleksi data yang dibutuhkan dari sekumpulan data, kemudian akan disimpan pada berkas yang berbeda dari data lainnya.
2. *Pre-processing* (*cleaning*), tahap ini menjaga keakuratan data yang disimpan dengan membuang data yang tidak lengkap atau tidak valid.
3. *Transformation*, memilih data yang relevan untuk analisis, sehingga hanya data yang sesuai yang digunakan pada tahapan selanjutnya.
4. *Data Mining*, merupakan tahap inti dari proses KDD. Tahap ini dibangun model data untuk menentukan pola dari data (Rahayu et al., 2024).
5. Evaluasi Pola (*Pattern Evaluation*), tahap ini mengindentifikasi data yang didapat dari proses *data mining* untuk mendapatkan pola yang akan dimasukkan ke dalam *knowledge based* untuk dilakukan analisis.
6. Presentasi Pengetahuan (*Knowledge Presentation*), menampilkan metode yang digunakan untuk memperoleh data yang berada di dalam *knowledge based*.

## K-Means

Algoritma K-Means adalah metode *clustering* non-hierarkis yang berusaha mempartisi data ke dalam satu atau lebih *cluster* atau kelompok. Metode ini mengelompokkan data dengan karakteristik yang sama ke dalam satu *cluster* yang sama, sementara data dengan karakteristik yang berbeda dikelompokkan ke dalam *cluster* yang lainnya (Maulana et al., 2024). Tujuan dari pengelompokan data adalah untuk meminimalkan fungsi objektif yang ditetapkan, dengan fokus pada mengurangi variasi di dalam setiap *cluster* dan meningkatkan variasi antar *cluster*.

Menurut Julyantari et al. (2021), algortima K-Means pada dasarnya melakukan 2 proses yaitu proses pendeteksian lokasi pusat *cluster* dan proses pencarian anggota dari tiap-tiap *cluster* (Julyantari et al., 2021). Berikut merupakan proses algoritma K-Means:

1. Tentukan k sebagai jumlah *cluster* yang dibentuk.
2. Tentukan k *centroid* (titik pusat *cluster*) awal secara *random*.

Penentuan *centroid* awal dilakukan secara random/acak dari objek-objek yang tersedia sebanyak k *cluster*, kemudian untuk menghitung *centroid* *cluster* ke-i berikutnya, digunakan rumus seperti pada persamaan (2.1):

(2.1)

Dimana :

v : *centroid* pada cluster

xi : objek ke-i

n : banyaknya objek/jumlah objek yang menjadi anggota *cluster*.

1. Hitung jarak setiap objek ke masing-masing *centroid* dari masing-masing *cluster*. Untuk menghitung jarak antara objek dengan *centroid* dapat menggunakan *Euclidian Distance*.

(2.2)

Dimana :

xi : objek x ke-i

yi : objek y ke-i

n : banyaknya objek

1. Alokasikan masing-masing objek ke dalam *centroid* yang paling dekat. Untuk melakukan pengalokasian objek kedalam masing-masing *cluster* pada saat iterasi secara umum dapat dilakukan dengan cara *hard* k-means dimana secara tegas setiap objek dinyatakan sebagai anggota *cluster* dengan mengukur jarak kedekatan sifatnya terhadap titik pusat *cluster* tersebut.
2. Lakukan iterasi, kemudian tentukan posisi *centroid* baru dengan menggunakan persamaan (2.1)
3. Ulangi langkah 3 jika posisi *centroid* baru tidak sama.

## Overpopulasi

Overpopulasi merupakan kondisi ketika jumlah penduduk suatu wilayah melebihi kapasitas lingkungan untuk menyediakan sumber daya yang cukup guna memenuhi kebutuhan dasar manusia, seperti makanan, air, dan tempat tinggal. Angka kelahiran yang tinggi, tingkat kematian yang rendah, dan migrasi penduduk yang tidak terkendali adalah penyebab overpopulasi terjadi. Menurut Ridwan et al. (2021), overpopulasi dapat menyebabkan kerusakan lingkungan yang serius, terutama di daerah perkotaan dengan tingkat kepadatan penduduk yang tinggi (Ridwan et al., 2021). Di Indonesia, overpopulasi merupakan tantangan utama dalam pembangunan berkelanjutan, di mana pertumbuhan penduduk yang tidak terkendali dapat memperlambat pencapaian kesejahteraan nasional. (Milanda et al., 2024). Oleh karena itu, untuk mengurangi dampak buruk overpopulasi terhadap lingkungan dan kesejahteraan masyarakat secara keseluruhan, diperlukan strategi pengendalian populasi yang efektif.

## Wilayah Administratif

Wilayah administratif merupakan bagian dari sistem pemerintahan suatu negara yang secara resmi memiliki batas-batas teritorial dan fungsi administrasi tertentu. Menurut Failaq dan Arelia (2022), wilayah administratif di Indonesia identik dengan satuan wilayah seperti provinsi, kabupaten/kota, kecamatan, dan kelurahan/desa, yang masing-masing memiliki peran dalam pelaksanaan pemerintahan, pelayanan publik, serta penegakan regulasi administratif (Failaq & Arelia, 2022). Hal ini menunjukkan bahwa meskipun suatu wilayah diakui secara administratif, belum tentu wilayah tersebut memiliki kewenangan otonom untuk mengatur urusannya sendiri. Dalam perspektif pembangunan wilayah, wilayah administratif berfungsi sebagai kerangka utama dalam menyusun rencana tata ruang, menentukan prioritas alokasi sumber daya, serta membentuk kesatuan wilayah kerja dalam konteks desentralisasi dan otonomi daerah (Pambudi & Sitorus, 2021). Oleh karena itu, pemahaman terhadap struktur wilayah administratif menjadi dasar penting dalam studi-studi spasial, demografi, maupun kebijakan publik di Indonesia.

## Website

## Database

## Laravel

## PHP

# ANALISIS DAN PERANCANGAN

## Analisis Kebutuhan

Analisis kebutuhan sistem merupakan proses untuk mengidentifikasi dan mendefinisikan fitur serta karakteristik yang harus dimiliki oleh sistem agar dapat memenuhi tujuan yang diharapkan. Dalam konteks sistem “Penerapan Metode K-Means untuk Klasterisasi Daerah dengan Risiko Overpopulasi di Indonesia”, kebutuhan dibagi menjadi dua jenis, yaitu kebutuhan fungsional dan non-fungsional.

### Kebutuhan Fungsional

1. *Website* memiliki fitur untuk mengelola data populasi seperti jumlah penduduk, kepadatan penduduk, luas wilayah, dan pertumbuhan penduduk. Data ini akan digunakan sebagai input untuk proses klasterisasi.
2. *Website* mampu melakukan proses klasterisasi menggunakan algoritma K-Means untuk mengelompokkan daerah berdasarkan tingkat risiko overpopulasi (rendah, sedang, tinggi).
3. *Website* dapat menampilkan hasil klasterisasi dalam bentuk peta wilayah dengan risiko overpopulasi. Selain itu, menampilkan hasil klasterisasi menggunakan metode K-Means.
4. *Website* harus menyediakan fitur untuk mengekspor data hasil klasterisasi dalam format seperti CSV atau PDF untuk keperluan analisis lebih lanjut.

### Kebutuhan Non Fungsional

1. *Website* harus dapat diakses 24 jam, sehingga pengguna dapat mengakses informasi kapan saja.
2. *Website* harus mampu melakukan proses klasterisasi dan pengolahan data populasi secara efisien, bahkan dengan volume data yang besar.
3. *Website* harus dapat diakses melalui koneksi internet tanpa memerlukan instalasi aplikasi tambahan pada perangkat pengguna.
4. *Website* harus dapat dijalankan pada web browser populer seperti Mozilla Firefox, Google Chrome, dan Microsoft Edge, dengan tampilan dan fungsionalitas yang konsisten di semua platform.

## Tabel Kriteria

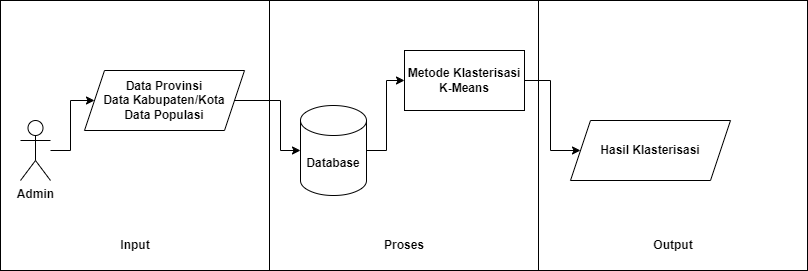
##### Tabel Kriteria

|  |  |
| --- | --- |
| **Kode Kriteria** | **Kriteria** |
| X1 | Jumlah Penduduk |
| X2 | Kepadatan Penduduk |
| X3 | Laju Pertumbuhan Penduduk |
| X4 | Luas Wilayah |

Tabel 3.1 merupakan tabel kriteria yang digunakan dalam klasterisasi daerah dengan risiko overpopulasi di Indonesia. Kode kriteria X1 (Jumlah Penduduk) menunjukkan bahwa semakin tinggi jumlah penduduk di suatu daerah, semakin besar potensi risiko overpopulasi yang dapat terjadi. Kode kriteria X2 (Kepadatan Penduduk) menggambarkan jumlah penduduk dalam suatu wilayah tertentu. Semakin tinggi kepadatan penduduk, semakin besar kemungkinan daerah tersebut mengalami overpopulasi. Kode kriteria X3 (Laju Pertumbuhan Penduduk) menunjukkan tingkat pertumbuhan populasi di suatu daerah dalam kurun waktu tertentu. Semakin tinggi laju pertumbuhan penduduk, semakin besar kemungkinan daerah tersebut mengalami peningkatan kepadatan penduduk secara signifikan. Kode kriteria X4 (Luas Wilayah) menjadi faktor penyeimbang dalam klasterisasi. Daerah dengan luas wilayah yang kecil tetapi memiliki jumlah penduduk yang tinggi lebih berisiko mengalami overpopulasi dibandingkan dengan daerah yang luas dengan jumlah penduduk yang sama.

## Diagram Blok Sistem

Berikut adalah diagram blok penerapan *data mining* untuk pengelompokan daerah risiko overpopulasi di Indonesia menggunakan metode K-Means.

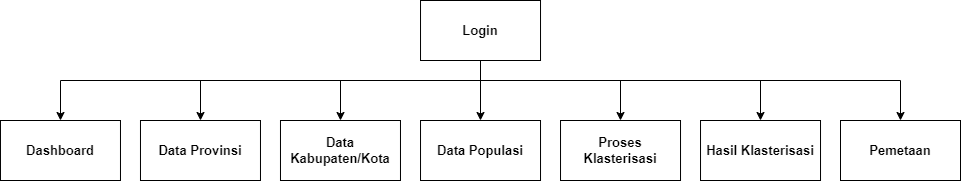


#### Diagram Blok Sistem

Gambar 3.1 merupakan blok diagram sistem, pada blok diagram terdapat 3 bagian yaitu *input*, proses, dan *output*. Pada *input* terdapat admin yang menginputkan data provinsi, data kabupaten/kota dan data populasi. Selanjutnya data yang telah diinputkan akan masuk ke bagian proses, data tersebut akan dimasukkan ke basis data dan dikelompokkan menggunakan algoritma K-Means. Setelah proses pengelompokan selesai akan diperoleh hasil pengelompokan.

## Struktur Menu

Pada website ini, membutuhkan beberapa menu didalamnya, menu tersebut digambarkan pada gambar 3.2:

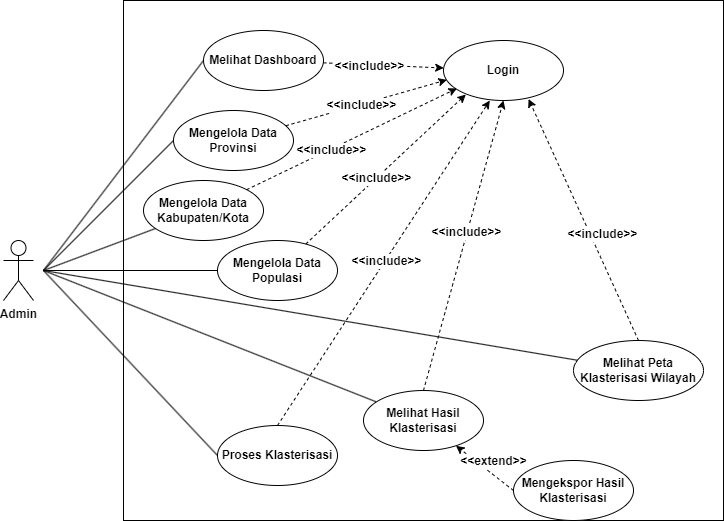


#### Struktur Menu

Gambar 3.2 merupakan struktur menu dari website, pengguna harus *login* terlebih dahulu sebelum mengakses sistem. Setelah berhasil *login*, pengguna dapat mengakses menu utama seperti *dashboard*, menu data provinsi, menu data kabupaten/kota, menu proses klasterisasi, menu hasil klasterisasi, dan menu pemetaan.

## Use Case Diagram

Berdasarkan penelitian ini, *use case* diagram dari program digambarkan sebagai berikut :

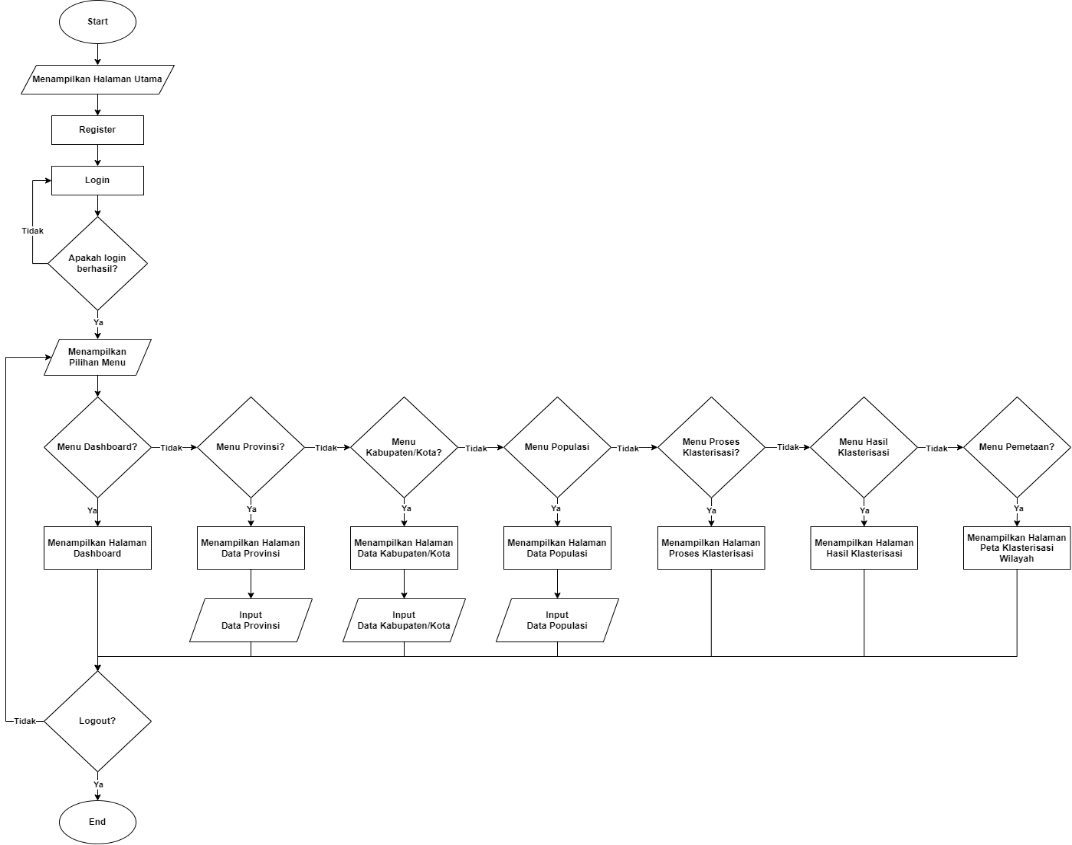


#### *Use* *Case* Diagram

Gambar 3.3 merupakan *use case* diagram dari program, Admin dapat melakukan *login* ke sistem, mengelola data provinsi, mengelola data kabupaten/kota, mengelola data populasi, menjalankan proses klasterisasi, melihat hasil klasterisasi, dan melihat peta klasterisasi. Selain itu, Admin juga dapat mengekspor hasil klasterisasi untuk keperluan analisis lebih lanjut.

## Flowchart

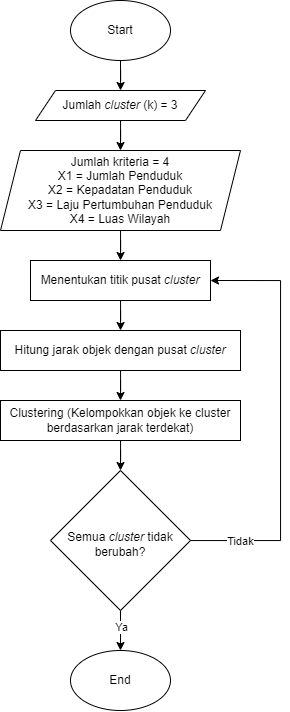
### Flowchart Sistem



#### *Flowchart* Sistem

Gambar 3.4 merupakan alur kerja dari sistem klasterisasi daerah dengan risiko overpopulasi. *Flowchart* sistem terlebih dahulu menampilkan halaman utama, apabila belum mendaftar harus mendaftar di halaman *register*. Setelah mendaftar dapat melakukan *login* dengan memasukkan kredential akun di halaman *login*, apabila *login* berhasil, maka diarahkan ke *dashboard*. Jika *login* gagal maka akan diarahkan kembali untuk menginputkan *email* dan *password*. Setelah *login* berhasil, terdapat menu data provinsi yang akan diisi inputan provinsi yang dapat diedit maupun dihapus, menu data kabupaten/kota yang akan diisi inputan kabupaten/kota yang dapat diedit maupun dihapus, menu data populasi yang akan diisi inputan jumlah penduduk, kepadatan penduduk, laju pertumbuhan. Selanjutnya terdapat menu proses klasterisasi yang akan menampilkan proses klasterisasi, terdapat juga menu hasil klasterisasi yang akan menampilkan data hasil klasterisasi. Selain itu, terdapat menu pemetaan yang akan menampilkan peta klasterisasi daerah berdasarkan hasil klasterisasi.

### Flowchart Metode K-Means



#### Flowchart Metode K-Means

Gambar 3.5 diatas menunjukkan alur kerja metode K-Means *Clustering*. Dimulai dengan memasukkan jumlah *cluster* atau k, dalam konteks ini, jumlah klaster ada 3 yaitu rendah, sedang dan tinggi. Setelah itu memasukkan kriteria, dalam konteks ini, terdapat 4 kriteria yaitu jumlah penduduk, kepadatan penduduk, laju pertumbuhan penduduk dan luas wilayah. Selanjutnya masuk pada menentukan titik pusat *cluster*, lalu masuk ke perhitungan jarak objek dengan pusat *cluster*. Setelah itu masuk ke tahap klasterisasi yaitu mengelompokkan objek ke *cluster* bedasarkan jarak terdekatnya. Lalu, terdapat pengecekan apakah data pada *cluster* tidak ada yang berubah, jika masih ada, maka akan mengulangi proses dari menentukan titik pusat *cluster*. Jika sudah tidak ada, maka akan berhenti.

## Data Klaster

Berikut klaster dari “Penerapan Metode K-Means untuk Klasterisasi Daerah dengan Risiko Overpopulasi di Indonesia” ditunjukkan pada Tabel 3.2.

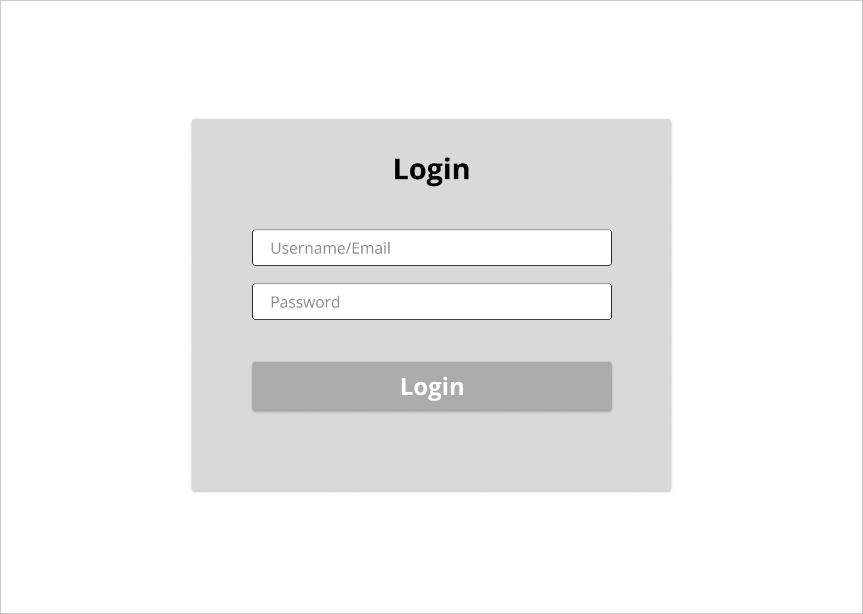
##### Tabel Klaster pada Sistem Klasterisasi Daerah Risiko Overpopulasi

|  |  |
| --- | --- |
| **Klaster** | **Nama Klaster** |
| C1 | Rendah |
| C2 | Sedang |
| C3 | Tinggi |

Tabel 3.2 menunjukkan pengelompokan wilayah berdasarkan tingkat overpopulasi ke dalam tiga klaster utama. Klaster 1 diberi label “Rendah”, yang menggambarkan wilayah dengan risiko overpopulasi yang relatif rendah berdasarkan kriteria jumlah penduduk, kepadatan penduduk, laju pertumbuhan penduduk dan luas wilayah. Klaster 2, dengan label “Sedang”, mencerminkan wilayah dengan tingkat risiko menengah berdasarkan kriteria jumlah penduduk, kepadatan penduduk, laju pertumbuhan penduduk dan luas wilayah. Sedangkan Klaster 3, yang dinamai “Tinggi”, mencakup wilayah dengan risiko overpopulasi paling besar berdasarkan kriteria jumlah penduduk, kepadatan penduduk, laju pertumbuhan penduduk dan luas wilayah.

## Desain Website

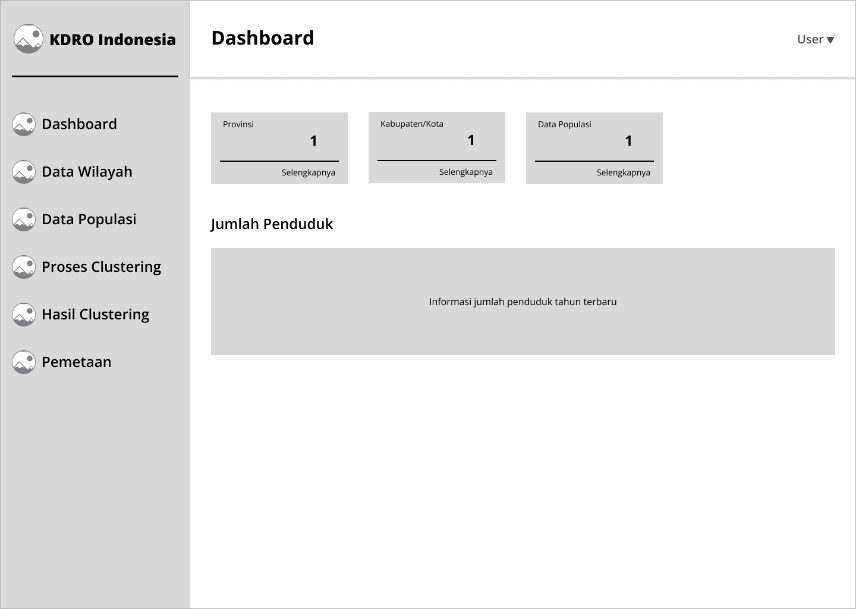
1. Halaman Login



#### Halaman *Login*

Gambar 3.5 merupakan halaman tampilan *login* dari *website* klasterisasi daerah dengan risiko overpopulasi. Pengguna harus memasukkan *username* dan *password* untuk mengakses menu lainnya dari *website*.

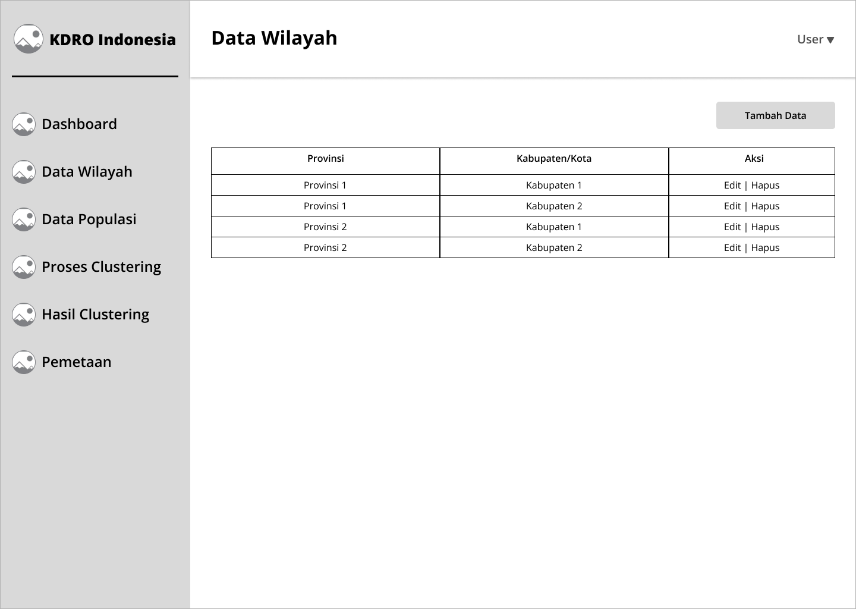
1. Menu Dashboard



#### Menu Dashboard

Gambar 3.6 merupakan halaman dashboard setelah pengguna berhasil *login*. Halaman ini berisi informasi jumlah penduduk tahun terbaru setiap provinsi, jumlah provinsi, jumlah kabupaten/kota, dan jumlah data populasi.

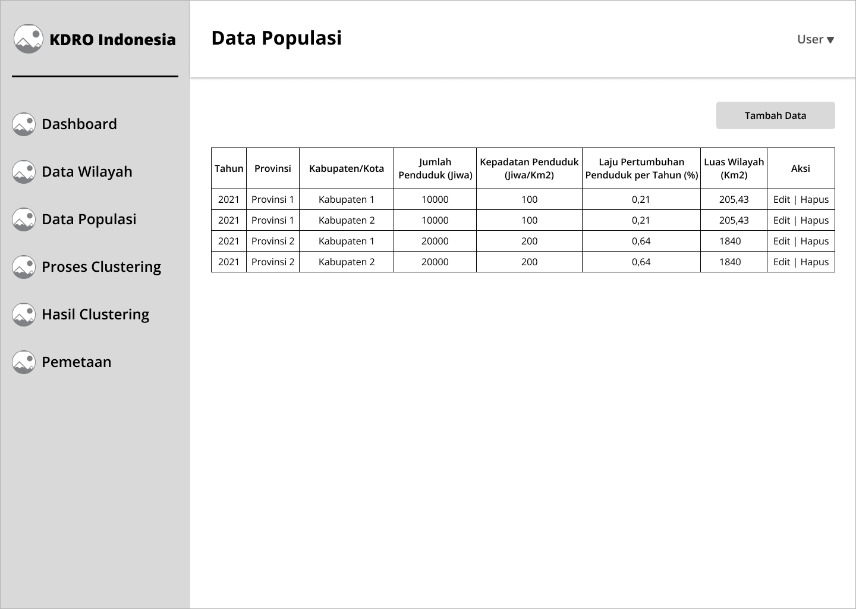
1. Menu Data Wilayah



#### Menu Data Wilayah

Gambar 3.7 merupakan halaman menu data wilayah. Menu ini digunakan untuk menginputkan data wilayah di Indonesia seperti provinsi dan kabupaten/kota. Data di menu ini dapat diedit dan juga dapat dihapus.

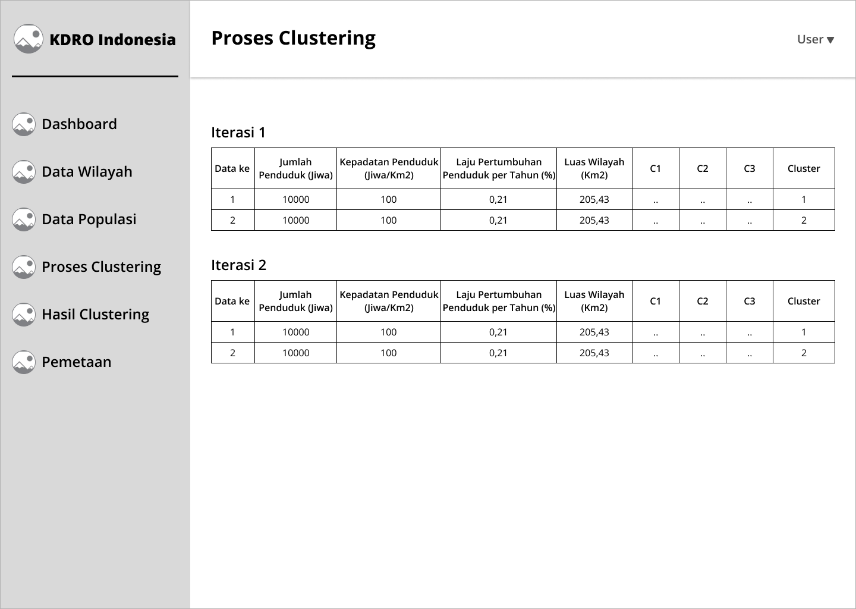
1. Menu Data Populasi



#### Menu Data Populasi

Gambar 3.8 merupakan halaman menu data populasi. Menu ini digunakan untuk menginputkan data kriteria klasterisasi seperti jumlah penduduk, kepadatan penduduk, laju pertumbuhan penduduk, dan luas wilayah. Data di menu ini dapat diedit dan juga dapat dihapus.

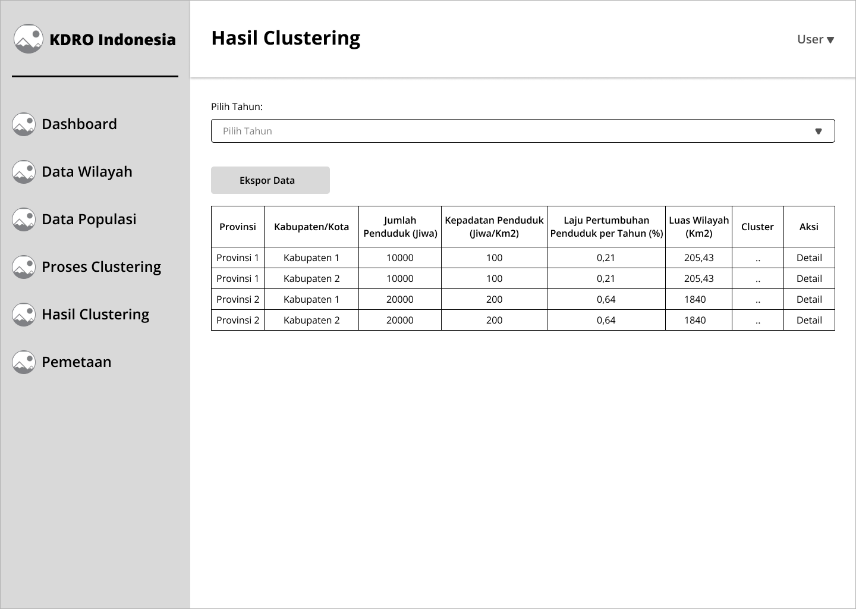
1. Menu Proses Clustering



#### Menu Proses *Clustering*

Gambar 3.9 merupakan halaman menu proses *clustering*. Menu ini digunakan untuk menampilkan iterasi dari proses klasterisasi.

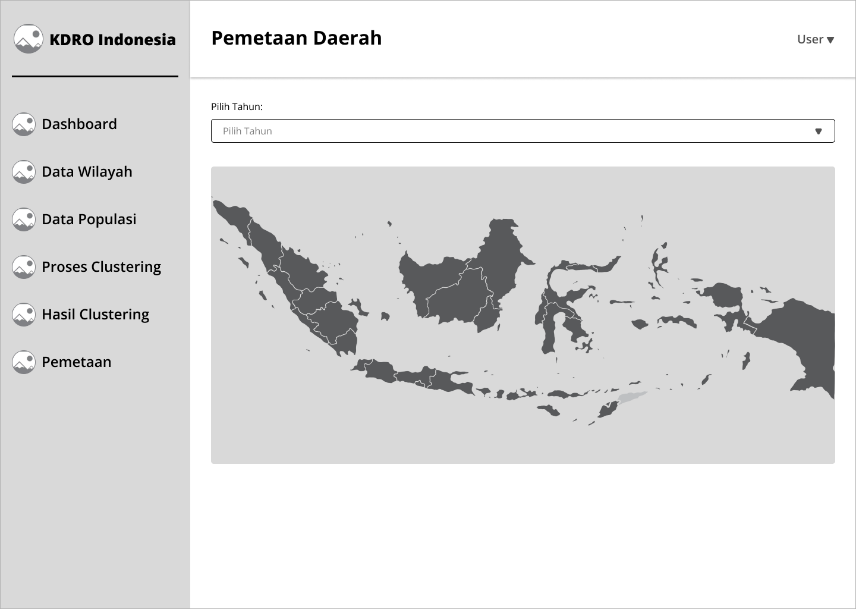
1. Menu Hasil Clustering



#### Menu Hasil *Clustering*

Gambar 3.10 merupakan halaman menu hasil *clustering*. Menu ini digunakan untuk menampilkan hasil akhir klasterisasi yang telah di hitung di halaman Proses *Clustering*.

1. Menu Pemetaan



#### Menu Pemetaan

Gambar 3.11 merupakan halaman menu pemetaan. Menu ini digunakan untuk menampilkan peta daerah klasterisasi dari data di halaman Hasil *Clustering*.

# PROGRESS YANG TELAH DILAKUKAN

## Hasil Sampai Saat Ini (berisi gambar dan screenshot serta penjelasan progress yang sudah dilakukan)

.

## Kendala yang dihadapi

.belum mengimplementasikan metode k-means

Mengambil data

Memfilter data

## Rencana Penyelesaian Masalah

.

# DAFTAR PUSTAKA

Aisy, R., & Kurniawan, R. (2024). Klasterisasi Data Bencana Alam Di Kota Cirebon Menggunakan K-Means Clustering Berdasarkan Kawasan Dan Jenis Bencana. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(2), 2166–2171. https://doi.org/10.36040/jati.v8i2.9096

Bahauddin, A., Fatmawati, A., & Permata Sari, F. (2021). Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means. *Jurnal Manajemen Informatika dan Sistem Informasi*, *4*(1), 1–8. https://doi.org/10.36595/misi.v4i1.216

BPS. (2024). *Jumlah Penduduk Pertengahan Tahun (Ribu Jiwa), 2022-2024*. https://www.bps.go.id/id/statistics-table/2/MTk3NSMy/jumlah-penduduk-pertengahan-tahun--ribu-jiwa-.html

Failaq, M. R. F., & Arelia, F. A. (2022). Diskrepansi Sistem Pemerintahan Daerah Khusus Ibu Kota Negara (IKN) Nusantara. *Jurnal Studi Kebijakan Publik*, *1*(1), 57–69. https://doi.org/10.21787/jskp.1.2022.57-69

Handayanna, F., & Sunarti. (2024). Penerapan Algoritma K-Means Untuk Mengelompokkan Kepadatan Penduduk Di Provinsi DKI Jakarta. *Journal of Applied Computer Science and Technology (JACOST)*, *5*(1), 50–55. https://doi.org/10.52158/jacost.v5i1.477

Iqbal, M., Syaripuddin, S., & Huda, M. N. (2023). Implementasi Algoritma K-Means Clustering dengan Jarak Euclidean dalam Mengelompokkan Daerah Penyebaran COVID-19 di Kabupaten Bogor. *BASIS (Jurnal Ilmiah Matematika)*, *2*(1), 47–56.

Julyantari, N. K. S., Budiarta, I. K., & Putri, N. M. D. K. (2021). Implementasi K-Means Untuk Pengelompokan Status Gizi Balita (Studi Kasus Banjar Titih). *Jurnal Janitra Informatika dan Sistem Informasi*, *1*(2), 92–101. https://doi.org/10.25008/janitra.v1i2.134

Maulana, A., Dana, R. D., & Nuris, N. D. (2024). Implementasi Algoritma K-Means Clustering Dalam Pengelompokan Data Kerusakan Rumah Akibat Bencana Alam Di Kabupaten Cirebon. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(2), 1417–1424. https://doi.org/10.36040/jati.v8i2.9024

Milanda, O., Almira, N., & Malik, G. A. (2024). Besarnya Pertumbuhan Angka Penduduk Indonesia (Overpopulation) Dalam Perspektif Undang-Undang Nomor 24 Tahun 2013 Tentang Administrasi Kependudukan. *Journal of Multidiscipline & Equality*, *1*(2), 43–53.

Munawar, A., Gendalasari, G. G., Kurniawan, I. M. G. A., Purnomo, D., Ependi, N. H., Rulinawaty, Indrawan, M. I., & Sadri, M. (2021). Cluster Application with K-Means Algorithm on the Population of Trade and Accommodation Facilities in Indonesia. *Journal of Physics: Conference Series*, *1933*(1). https://doi.org/10.1088/1742-6596/1933/1/012027

Nabiilah, Fauzan, F. J., Nurazizah, N., Hamid, A., & Octavia, S. F. (2023). Penerapan Algoritma K-Means dan K-Medoids Dalam Pengelompokan Kepadatan Penduduk Provinsi Riau. *SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat*, 223–229.

Nurhayah, Suarna, N., & Prihartono, W. (2024). Analisis Jumlah Penduduk Menggunakan Algoritma K-Means Berdasarkan Kabupaten/Kota Di Indonesia. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(1), 894–901. https://doi.org/10.36040/jati.v8i1.8863

Pambudi, A. S., & Sitorus, S. R. P. (2021). Omnimbus Law Dan Penyusunan Rencana Tata Ruang: Konsepsi, Pelaksanaan Dan Permasalahannya Di Indonesia. *Jurnal Ilmiah Wahana Bhakti Praja*, *11*(2), 198–216. https://doi.org/10.33701/jiwbp.v11i2.2216

Rahayu, P., Sudipa, I. G. I., Suryani, Surachman, A., Ridwan, A., Darmawiguna, I. G. M., Sutoyo, M., Slamet, I., Harlina, S., & May Sanjaya, I. M. (2024). *Buku Ajar Data Mining*. PT. Sonpedia Publishing Indonesia.

Ridwan, M., Hidayanti, S., & Nilfatri. (2021). Studi Analisis Tentang Kepadatan Penduduk Sebagai Sumber Kerusakan Lingkungan Hidup. *Jurnal IndraTech*, *2*(1), 25–36.

Sudarwati, W., Hardjomidjojo, H., Machfud, & Setyaningsih, D. (2022). K-means cluster approach in grouping of seaweed agroindustry development areas. *IOP Conference Series: Earth and Environmental Science*, *1063*(1). https://doi.org/10.1088/1755-1315/1063/1/012048

Sujak, G. M. M., Rofiq, H. N., & Tawakal, F. I. (2025). Implementasi K-Means Clustering untuk Optimalisasi Anggaran Penyakit Tidak Menular. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *5*(1), 67–74.

Talibi, T., Onyeneke, C., & Eguzouwa, C. (2022). *Natural and Human Impacts of Overpopulation on the Economy*. *10*(6), 217–223. https://doi.org/10.11648/j.ijsts.20221006.11

Temalagi, S., Sembiring, T. B., Nasution, J., & ... (2023). The Importance of Public Knowledge for a Sustainable Environment. *Journal of Survey in Fisheries Sciences*, *10*(1), 2869–2875. https://doi.org/10.53555/sfs.v10i1.1269

Wahono, R. S. (2023). Data Mining Data mining. In *Mining of Massive Datasets* (Vol. 2, Nomor January 2013).

**(BERISI FORMULIR BIMBINGAN SKRIPSI YANG TELAH DI TTD DOSEN 1 DAN DOSEN 2)**