**IMPLEMENTASI *K-MEANS* *CLUSTERING* DALAM**

**PEMETAAN DAERAH RAWAN PENCURIAN KENDARAAN**

**BERMOTOR DAN PENCURIAN DENGAN KEKERASAN**

**BERBASIS SISTEM INFORMASI GEOGRAFIS**

**(STUDI KASUS DI KABUPATEN PROBOLINGGO)**

**PROPOSAL SKRIPSI**



Oleh :

**Daffa Fauzi Rahman**

**NIM E41211408**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**JURUSAN TEKNOLOGI INFORMASI**

**POLITEKNIK NEGERI JEMBER**

**2024**

**IMPLEMENTASI *K-MEANS* *CLUSTERING* DALAM**

**PEMETAAN DAERAH RAWAN PENCURIAN KENDARAAN**

**BERMOTOR DAN PENCURIAN DENGAN KEKERASAN**

**BERBASIS SISTEM INFORMASI GEOGRAFIS**

**(STUDI KASUS DI KABUPATEN PROBOLINGGO)**

**PROPOSAL SKRIPSI**



Sebagai salah satu syarat untuk memperoleh gelar Sarjana Terapan (S.Tr.)

di Program Studi Teknik Informatika Jurusan Teknologi Informasi

Oleh :

**Daffa Fauzi Rahman**

**NIM E41211408**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**JURUSAN TEKNOLOGI INFORMASI**

**POLITEKNIK NEGERI JEMBER**

**2024**

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET DAN TEKNOLOGI**

POLITEKNIK NEGERI JEMBER

JURUSAN TEKNOLOGI INFORMASI

# HALAMAN PENGESAHAN PROPOSAL

1. Judul : Implementasi *K-Means* *Clustering* Dalam

Pemetaan Daerah Rawan Pencurian Kendaraan Bermotor dan Pencurian Dengan Kekerasan Berbasis Sistem Informasi Geografis (Studi Kasus di Kabupaten Probolinggo)

1. Identitas Pelaksanaan
   1. Nama Lengkap : Daffa Fauzi Rahman
   2. NIM : E41211408
   3. Jurusan/ Program Studi : Teknologi Informasi / Teknik Informatika
2. Lokasi : Politeknik Negeri Jember
3. Identitas Dosen Pembimbing
   1. Nama Lengkap : Bety Etikasari, S.Pd, M.Pd
   2. NIP : 19920528 201803 2 001
   3. Jurusan/ Program Studi : Teknologi Informasi / Teknik Informatika
4. Lama Kegiatan : Sepuluh (10) Bulan

Jember, 1 Juli 2024

|  |  |  |
| --- | --- | --- |
| Menyetujui, |  |  |
| Dosen Pembimbing | Dosen Pembahas | Pelaksana |
|  |  |  |
| Bety Etikasari, S.Pd, M.Pd NIP. 19920528 201803 2 001 | Dia Bitari Mei Yuana, S.ST., M.Tr.Kom. NIP. 19930508 202203 2 013 | Daffa Fauzi Rahman NIM. E41211408 |
|  |  |  |
| Mengetahui, | | |
| Koordinator Program Studi D-IV Teknik Informatika | | |
|  | | |
| Bety Etikasari, S.Pd, M.Pd NIP. 19920528 201803 2 001 | | |

# DAFTAR ISI

Halaman

[**HALAMAN PENGESAHAN PROPOSAL** iii](#_Toc170659030)

[**DAFTAR ISI** iv](#_Toc170659031)

[**DAFTAR GAMBAR** vi](#_Toc170659032)

[**DAFTAR TABEL** vii](#_Toc170659033)

[**DAFTAR LAMPIRAN** viii](#_Toc170659034)

[**BAB 1. PENDAHULUAN** 1](#_Toc170659035)

[**1.1 Latar Belakang** 1](#_Toc170659036)

[**1.2 Rumusan Masalah** 5](#_Toc170659037)

[**1.3 Tujuan Penelitian** 5](#_Toc170659038)

[**1.4 Manfaat** 5](#_Toc170659039)

[**1.5 Batasan Masalah** 6](#_Toc170659040)

[**BAB 2. TINJAUAN PUSTAKA** 7](#_Toc170659041)

[***2.1* *State Of The Art*** 7](#_Toc170659042)

[**2.2 Landasan Teori** 9](#_Toc170659043)

[2.2.1 Curanmor dan Curas 9](#_Toc170659044)

[2.2.2 Sistem Informasi Geografis (SIG) 10](#_Toc170659045)

[*2.2.3* *K-Means Clustering* 13](#_Toc170659046)

[*2.2.4* *Davies-Bouldin Index (DBI)* 15](#_Toc170659047)

[**BAB 3. METODE PENELITIAN** 18](#_Toc170659048)

[**3.1 Tempat dan Waktu Pelaksanaan** 18](#_Toc170659049)

[**3.2 Alat dan Bahan** 18](#_Toc170659050)

[3.2.1 Alat Penelitian 18](#_Toc170659051)

[3.2.2 Bahan Penelitian 19](#_Toc170659052)

[**3.3 Tahapan Penelitian** 19](#_Toc170659053)

[3.3.1 Studi Literatur 20](#_Toc170659054)

[3.3.2 Pengumpulan Data 20](#_Toc170659055)

[3.3.3 Pengolahan Data 20](#_Toc170659056)

[3.3.4 Pengembangan Sistem 21](#_Toc170659057)

[3.3.5 Pengujian 24](#_Toc170659058)

[3.3.6 Analisis dan Pembahasan 26](#_Toc170659059)

[**DAFTAR PUSTAKA** 27](#_Toc170659060)

[**LAMPIRAN** 43](#_Toc170659061)

# DAFTAR GAMBAR

Halaman

[**Gambar 2.1** Gambaran Data Spasial 13](#_Toc170659512)

[**Gambar 2.2** Tahapan Algoritma *K-Means Clustering* 14](#_Toc170659513)

[**Gambar 3.1** Tahapan Penelitian 19](#_Toc170659514)

[**Gambar 3.2** Use Case Sistem **Error! Bookmark not defined.**](#_Toc170659515)

[**Gambar 3.3** Flowchart Sistem Admin **Error! Bookmark not defined.**](#_Toc170659516)

[**Gambar 3.4** Flowchart Sistem User **Error! Bookmark not defined.**](#_Toc170659517)

# DAFTAR TABEL

Halaman

[**Tabel 2.1** *State Of The Art* **7**](#_Toc170659538)

[**Tabel 3.1** Tahapan Kegiatan 18](#_Toc170659539)

# DAFTAR LAMPIRAN

Halaman

[**LAMPIRAN 1** Kenaikan Kasus Kriminal di Probolinggo 43](#_Toc170659559)

[**LAMPIRAN 2** Data Kasus Curas dan Curanmor Tahun 2022 44](#_Toc170659560)

# PENDAHULUAN

## Latar Belakang

Kriminalitas atau kejahatan merupakan suatu kejadian umum yang mempunyai dampak luas terhadap seluruh lapisan masyarakat, dan kriminalitas atau kejahatan ini bisa terjadi dimana saja dan tidak mengenal waktu (Apriliana & Haris R, 2022). Seperti yang telah disampaikan sebelumnya, Riswandi juga berpendapat yang sama dalam (Rohman, 2023) tindakan kriminal merugikan seluruh lapisan masyarakat baik dalam segi ekonomis. Psikologi, dan juga merupakan tindakan yang melanggar hukum dan norma-norma agama maupun sosial yang ada pada masyarakat. Tindak kriminal terjadi tentunya disebabkan oleh beberapa faktor, seperti : kemiskinan, kesempatan kerja, karakter pelaku yang melakukan kasus kejahatan, kepadatan penduduk, keadaan lingkungan, dan jumlah patroli polisi. Jenis-jenis tindakan kriminal ada beberapa, seperti : pencurian, pembunuhan, tindak asusila, dan lain sebagainya (Apriliana & Haris R, 2022).

Badan Pusat Statistik (BPS) telah merilis hasil perhitungannya tentang kriminalitas di Indonesia pada publikasinya di tahun 2023. Pada publikasi tersebut dinyatakan bahwa tingkat kriminalitas di Indonesia mengalami kenaikan pada tahun 2022 dengan jumlah tindak kriminal 372.965 kasus. Pada tahun sebelumnya yaitu pada tahun 2020 – 2021 angka Tingkat kriminalitas Indonesia mengalami penurunan, pada tahun 2020 terjadi 247.218 dan pada tahun 2021 terjadi 239.481 kasus. Pada hal tersebut tampak peningkatan yang cukup signifikan terjadinya beberapa tindak kriminalitas di tahun 2022. Jika dalam interval waktu (*crime clock*), dapat disampaikan bahwa setiap 2 menit 11 detik terjadi satu tindak kriminal pada tahun 2021, sedangkan pada tahun 2022 tindak kriminal terjadi dalam rentang waktu yang lebih cepat, yaitu 1 menit 24 detik terjadi satu tindak kriminalitas di wilayah Indonesia. Jika dilihat lebih detail lagi pada publikasi BPS tersebut, provinsi dengan tingkat kriminalitas tertinggi terjadi pada provinsi Jawa Timur dengan jumlah kasus sebesar 51.905 kasus. (BPS, 2023)

Salah satu kabupaten pada provinsi dengan tingkat kriminalitas tertinggi di tahun 2022 (Jawa Timur) adalah Kabupaten Probolinggo. Kabupaten Probolinggo memiliki luas 1.696,17 dengan 24 kecamatan, 5 kelurahan dan 325 desa. Menurut data dari Badan Pusat Statistik (BPS) Kabupaten Probolinggo pada tahun 2022, jumlah penduduk Kabupaten Probolinggo tercatat sebanyak 1.159.965 jiwa (BPS Kab Probolinggo, 2024). Selain itu letak geografis Kabupaten Probolinggo yang berbatasan langsung dengan selat Madura dan ada beberapa tempat yang termasuk dalam dataran tinggi, membuat Kabupaten Probolinggo memiliki beberapa wisata yang terkenal, baik di tingkat nasional maupun mancanegara. Salah satu wisata di kabupaten Probolinggo yang terkenal hingga ke mancanegara yaitu Gunung Bromo. Dengan adanya banyak wisata yang terkenal di tingkat nasional maupun mancanegara, membuat Kabupaten Probolinggo banyak dikunjungi wisatawan. Berdasarkan data BPS Kabupaten Probolinggo, di tahun 2023 terdapat 854.956 wisatawan domestik dan 37.094 wisatawan mancanegara (BPS Kab Probolinggo, 2024). Dengan banyaknya jumlah penduduk Kabupaten Probolinggo dan banyaknya wisatawan yang berkunjung ke Kabupaten Probolinggo tentu dapat membuat ancaman terjadinya tindak kriminalitas semakin tinggi.

Pada Kabupaten sendiri tingkat kejahatan dari tahun 2021 - 2022 mengalami peningkatan. Menurut Badan Pusat Statistik Kabupaten Probolinggo tindak kriminalitas yang terjadi di wilayah hukum Polres Probolinggo pada tahun 2021 tercatat sebanyak 399 kasus, sedangkan pada tahun 2022 tindak kriminalitas yang tercatat sebanyak 442 kasus (BPS Kab Probolinggo, 2023). Dari data tersebut dapat dinyatakan terjadi kenaikan sebesar 9,7 %. Jenis tindak kriminalitas yang paling banyak terjadi pada jenis tindak kriminal pencurian. Tindak kriminal pencurian ini terbagi lagi dalam dua kategori, yaitu pencurian kendaraan bermotor (curanmor) dan pencurian dengan kekerasan (curas).

Dapat terlihat pada data BPS tentang kriminalitas di Kabupaten Probolinggo mengalami kenaikan dan tindak kriminal terbanyak adalah pencurian dengan kekerasan (curas) dan pencurian kendaraan bermotor (curanmor). Dengan peningkatan tersebut dapat memberikan kesan bahwa Kabupaten Probolinggo kurang aman, sehingga secara tidak langsung dapat mempengaruhi jumlah wisatawan yang ingin berkunjung ke Kabupaten Probolinggo. Salah satu upaya *preventif* dalam memberikan rasa aman bagi penduduk asli Kabupaten Probolinggo maupun wisatawan yang berkunjung ke Kabupaten Probolinggo atau tindakan untuk meminimalkan ancaman tindak kriminal dapat dengan memberikan pengamanan atau patroli pada beberapa titik daerah yang dapat dibilang rawan untuk terjadi tindakan-tindakan curas dan curanmor. Dalam menentukan suatu daerah tersebut termasuk ke dalam daerah rawan atau aman perlu dilakukan pemetaan untuk memastikannya, dan hasil dari pemetaan tersebut bisa dipublikasikan agar setiap orang dapat mengetahuinya. Pada Kabupaten Probolinggo sendiri belum ada pemetaan terkait tingkat kerawanan suatu kecamatan terhadap kasus tindak curas dan curanmor.

Sebelum ini terdapat beberapa penelitian yang membahas tentang pemetaan. Seperti penelitian yang dilakukan oleh Ruziq Nawaf Zulfahmi, dkk pada tahun 2023 yang menyatakan hasil dari pemetaan tingkat kriminalitas pada Kab/Kota di Jatim, Jateng, dan Yogyakarta menghasilkan dua *cluster* dengan rincian *cluster* 1 terdiri dari 15 kabupaten/kota, dan *cluster* 2 terdiri dari 63 kabupaten/kota, kemudian dengan melakukan pengujian *profiling cluster* menggunakan nilai rata - rata setiap *cluster*, Klaster 1 dapat dikatakan tinggi dibandingkan klaster 2 karena memiliki nilai dominan yang lebih besar untuk kategori dengan skor tinggi dari masing-masing variabel (Ruziq Nawaf Zulfahmi et al., 2023). Selain itu Preddy Marpaung, dkk melakukan penelitian pada tahun 2023 dan menyimpulkan bahwa dengan menerapkan algoritma *K-Means*, dapat mengelompokkan kepadatan penduduk di Kabupaten Deli Serdang ke dalam tiga kelompok (*cluster*) ,yaitu *Cluster* 1 daerah penduduk sangat padat sebanyak 3 kecamatan, *Cluster* 2 daerah Penduduk padat terdapat 4 kecamatan. *Cluster* 3 daerah Penduduk sedang 15 kecamatan (Preddy et al., 2023). Kemudian pada tahun 2022 Andrea Santana Adzani melakukan penelitian dan menyimpulkan bahwa tingkat akurasi perhitungan jarak antar data pada setiap klaster daerah rawan kecelakaan di Jember dalam sistem tersebut sebesar 83,87%. Kemudian berdasarkan hasil pengujian UAT dan blackbox yang telah dilakukan kepada 61 responden dan kepada pihak unit laka Satlantas Jember serta ahli IT, diperoleh hasil yang telah sesuai dengan target atau hasil yang diharapkan. Andrea pada penelitian tersebut menggunakan metode K-Means dengan perhitungan jarak antar data menggunakan persamaan *euclidean distance.* (Andrea Santana Adzani, 2022). Kemudian pada tahun 2021 Nofita Safira Anggraini juga melakukan penelitian mendapatkan hasil yaitu. analisis menggunakan metode clustering k-means terdapat 3 cluster daerah rawan begal, yaitu hijau untuk daerah aman dengan nilai rata rata cluster 2.889, orange untuk tingkat keamanan sedang dengan nilai rata rata cluster 1.333, dan merah untuk daerah rawan dengan nilai rata rata cluster 0.270833333. Pada penelitian itu Nofita juga menggunakan metode K-Means dengan perhitungan jarak antar data *euclidean distance* (Anggraini, 2021).

Berdasarkan pernyataan dan data-data yang telah di paparkan, dapat disimpulkan bahwa Kabupaten Probolinggo perlu menerapkan pemetaan daerah rawan curas dan curanmor pada setiap kecamatan yang berbasis Sistem Informasi Geografis, agar lebih mudah dipahami daerah mana saja yang termasuk kategori rawan, sehingga para penduduk Kabupaten Probolinggo dan wisatawan bisa lebih berhati hati ketika melewati suatu daerah, dan hasil dari pemetaan tersebut bisa digunakan aparat kepolisian sebagai acuan dalam melakukan patroli atau pengamanan. Selain itu dari beberapa penelitian terdahulu yang telah disebutkan, banyak dari penelitian tersebut menggunakan metode *K-Means clustering* yang telah dinyatakan oleh (Andrea Santana Adzani, 2022) memiliki akurasi 87,86 % yang dapat dinyatakan baik. Namun untuk perhitungan jarak antar datanya menggunakan persamaan *manhattan*, karena menurut Dinata dalam (Alifah & Fauzan, 2023) menyatakan bahwa persamaan *manhattan* lebih akurat dari pada persamaan *euclidean distance.*. Jadi Sistem Informasi Geografis yang akan dikembangkan dapat menggunakan metode K – Means *Clustering*, namun dengan perhitungan jarak datanya menggunakan persamaan *manhattan*.

## Rumusan Masalah

Dari latar belakang di atas, terdapat beberapa rumusan masalah dalam penelitian ini, sebagai berikut :

1. Bagaimana menerapkan algotitma *K - Means* dalam suatu website dan memvisualisasikan hasilnya ke dalam suatu peta Kabupaten Probolinggo ?
2. Bagaimana mendapatkan nilai *k* yang optimal untuk masing-masing data Curas dan Curanmor ?
3. Bagaimana hasil *clustering* untuk masing-masing data curas dan curanmor ?

## Tujuan Penelitian

Berdasarkan latar belakang dan rumusan masalah di atas terdapat beberapa tujuan dalam penelitian ini, sebagai berikut:

1. Mampu menerapkan algoritma *K – Means Clustering* ke dalam suatu website dan bisa memvisualisasikan hasilnya ke dalam suatu peta Kabupaten Probolinggo.
2. Mengetahui bagaimana mendapatkan nilai *k* yang optimal untuk masing-masing data Curas dan Curanmor.
3. Mengetahui hasil *clustering* final untuk masing masing data curas dan curanmor.

## Manfaat

Dengan adanya penelitian ini, diharapkan dapat memberikan manfaat kepada beberapa pihak, sebagai berikut :

1. Memberikan informasi bagi Kepolisian Resort Kabupaten Probolinggo tentang kecamatan yang berpotensi tinggi atau rendah untuk terjadi tindak kriminal pencurian kendaraan bermotor (curanmor) dan pencurian dengan kekerasan (curas) agar dapat dipertimbangkan untuk patroli rutin atau Upaya preventif lainnya.
2. Memberikan informasi bagi Masyarakat umum ataupun penduduk Kabupaten probolinggo tentang kecamatan yang memiliki Tingkat kerawanan tinggi untuk terjadi tindak kriminal pencurian kendaraan bermotor (curanmor) dan pencurian dengan kekerasan (curas) agar dapat lebih berhati hati ketika memasuki kecamatan tersebut.

## Batasan Masalah

1. Data yang digunakan pada penelitian ini merupakan data kasus curas dan curanmor pada tahun 2024 yang didapatkan dari Kepolisian Resort Kabupaten Probolinggo.
2. Tindak kriminal yang digunakan dalam penelitian ini hanya pada kategori pencurian kendaraan bermotor (curanmor) dan pencurian dengan kekerasan (curas)

# TINJAUAN PUSTAKA

## *State Of The Art*

Tabel 2.1 *State Of The Art*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No.** | **Judul** | **Penulis** | **Tahun** | **Hasil** |
| 1 | Pemetaan Kerentanan Tingkat Kriminalitas Menggunakan Metode *Self Organizing Map* | Ruziq Nawaf Zulfahmi , Maria Kristiana Daul , Muhammad Al Ayyubi, I Wayan Julianta Pradnyana, Rokhana Dwi Bekti | 2023 | *Cluster* 1 terdiri dari 15 kabupaten/kota, dan *cluster* 2 terdiri dari 63 kabupaten/kota. Berdasarkan pengujian profiling cluster menggunakan nilai ratarata setiap cluster, Klaster 1 dapat dikatakan tinggi dibandingkan klaster 2 karena memiliki nilai dominan yang lebih besar untuk kategori dengan skor tinggi dari masingmasing variabel. |
| 2 | Penerapan Data Mining Untuk Pengelompokan Kepadatan Penduduk Kabupaten Deli Serdang Menggunakan Algoritma K-Means | Preddy Marpaung , Ibnu Pebrian, Widia Putri | 2023 | Dengan menerapkan algoritma K-Means, dapat mengelompokkan kepadatan pendududk ke dalam tigah kelompok (cluster),yaitu Cluster 1/ daerah Penduduk sangat padat sebanyak 3 kecamatan, Cluster 2/ daerah Penduduk padat terdapat 4 kecamatan. Cluster 3/ daerah Penduduk sedang 15 kecamatan. Dengan adanya pengetahuan pengelompokan kepadatan penduduk ini, akan meminimalkan dampak akibat kepdatan penduduk bagi setiap orang yang mau bertempat tinggal di daerahdeli Serdang Dari hasil analisa peneliti, untuk kedepanya perlu dilakukan penelitian untuk pengelompokan kepadatan penduduk berdasarkan luas wilayah |
| 3 | Klastering Pemetaan Daerah Rawan Kecelakaan Menggunakan Metode K-Means Berbasis Sistem Informasi Geografis (Studi Kasus Di Kabupaten Jember) | Andrea Santana Adzani | 2022 | Tingkat akurasi perhitungan jarak antar data pada setiap klaster dalam sistem  tersebut sebesar 83,87%. Kemudian berdasarkan hasil pengujian UAT dan  blackbox yang telah dilakukan kepada 61 responden dan kepada pihak unit laka  Satlantas Jember serta ahli IT, diperoleh hasil yang telah sesuai dengan target  atau hasil yang diharapkan diawal sehingga sistem dapat dikatakan dapat  dioperasikan dengan baik. |
| 4 | Pemetaan Daerah Rawan Pencurian Dengan Kekerasan (Begal) Di Kabupaten Lumajang | Nofita Safira Anggraini | 2021 | Dari hasil analisis menggunakan metode clustering k-means terdapat 3 cluster yaitu Hijau untuk daerah aman dengan nilai rata rata cluster 2.889, Orange untuk tingkat keamanan sedang dengan nilai rata rata cluster 1.333, dan Merah untuk daerah rawan dengan nilai rata rata cluster 0.270833333. |

Dari beberapa penelitian yang telah dipaparkan pada *state of the art* di atas peneliti menyimpulkan bahwa penelitian ini sejalan dengan penelitian yang dilakukan oleh Nofita Safira Anggraini di tahun 2021, namun lebih melengkapi kategori tindak kriminalnya, tidak hanya terfokus pada kasus begal atau pencurian dengan kekerasan (curas) saja, melainkan menambahkan pada kasus pencurian kendaraan bermotor juga (curanmor). Selain itu peneliti juga menggunakan metode *K-Means Clustering* untuk pemetaannya, namun untuk persamaan pengukuran jarak antar data peneliti menggunakan persamaan *manhattan* guna mengetahui keakurasiannya.

## Landasan Teori

### Curanmor dan Curas

KUHP ( Kitab Undang-undang Hukum Pidana) pada pasal 362 menyatakan bahwa perbuatan pelaku kejahatan dengan mengambil suatu barang berupa kendaraan bermotor yang seluruhnya atau sebagian kepunyaan orang lain dengan maksud untuk memiliki kendaraan bermotor tersebut secara melawan hukum. Perbuatan kejahatan dengan mengambil kendaraan bermotor atau yang sering disebut curanmor berdasarkan pasal tersebut termasuk ke dalam tindak pidana pencurian. Sama seperti pencurian dengan kekerasan atau yang sering disebut curas merupakan bagian dari tindak pidana pencurian yang disertai kekerasan ataupun ancaman kekerasan. Peraturan terkait curas ini juga diatur pada KUHP ( Kitab Undang-undang Hukum Pidana) pada pasal 365. Kasus curas sering dianggap lebih serius dan berbahaya daripada jenis kasus pencurian lainnya, karena sudah melibatkan ancaman fisik dan psikologis korban, bahkan juga mengancam nyawa.

Kasus pencurian secara umum, baik curas maupun curanmor terjadi tentunya disebabkan oleh faktor internal dan eksternal. Faktor internal yang mempengaruhi terjadinya curas dan curanmor disebabkan oleh kemiskinan, kesempatan kerja, karakter pelaku yang melakukan kasus kejahatan. Sedangkan faktor eksternal yang mempengaruhi terjadinya curas dan curanmor disebabkan oleh kepadatan penduduk, keadaan lingkungan, dan jumlah patroli polisi (Apriliana & Haris R, 2022).

### Sistem Informasi Geografis (SIG)

Nugroho menjelaskan tentang pengertian dan tujuan Sistem Informasi geografis dalam (Umar, 2021) bahwa Geograpic Information System atau Sistem Informasi Geografis (SIG) merupakan aplikasi pengolahan data spasial dengan menggunakan sistem terkomputerisasi dengan menggabungkan antara data grafis dengan data atribut objek menggunakan peta dasar digital (basic map) geoerensi bumi. SIG adalah sebuah sistem atau teknologi berbasis komputer yang dibangun dengan tujuan untuk mengumpulkan, menyimpan, mengolah dan menganalisa, serta menyajikan data informasi dari suatu objek atau fenomena yang berkaitan dengan letak atau keberadaannya di permukaan bumi yang bertujuan untuk membantu masyarakat mencari lokasi yang sedang dicari. SIG juga dapat diartikan sebagai sistem informasi yang digunakan untuk memasukkan, menyimpan, memanggil kembali, mengolah, menganalisis dan menghasilkan data bereferensi geografis atau data geospasial, untuk mendukung pengambilan keputusan dalam perencanaan dan pengelolaan penggunaan lahan, sumber daya alam, lingkungan transportasi, fasilitas kota, dan pelayanan umum lainnya. Teknologi SIG mengintegrasi operasi-operasi umum database, seperti query, dan analisa statistic, dengan kemampuan visualisasi dan analisa yang unik yang dimiliki oleh pemetaan. Kemampuan inilah yang membedakan SIG dengan Sistem Informasi lainnya (Umar, 2021).

Sistem Informasi Geografis (GIS) memiliki cicri ciri khusus seperti yang disampaikan oleh Susianto dan Guntoro dalam (Andrea Santana Adzani, 2022), sebagai berikut :

1. Masukan data yang mampu memuat dan memproses data spasial dari berbagai sumber merupakan subsistem dari SIG, sub sistem ini juga mampu memproses perubahan data spasial yang memiliki perbedaan jenis, seperti dari peta kontur menjadi titik ketinggian.
2. Subsistem SIG mampu menyimpan dan memanggil data yang memungkinkan data spasial untuk ditampilkan, diubah, dan dihapus.
3. Subsistem lain yang dimiliki oleh SIG yakni mampu memanipulasi dan menganalisis peran data, pengelompokan dan pemisahan, perkiraan parameter dan hambatan, serta fungsi permodelan dari data yang dimuat.
4. Pelaporan yang dimiliki subsistem SIG berbentuk peta, grafis, dan tabel.

Menurut Susianto dan Guntoro juga dalam (Andrea Santana Adzani, 2022) suatu sistem informasi geografis memiliki beberapa sub sistem, sebagai berikut :

1. Data *Input*

Data *input* berfungsi untuk mengumpulkan lalu mempersiapkan suatu data *spasial* beserta atributnya dari berbagai sumber. Subsistem ini juga bertanggung jawab dalam mengkonversi atau merepresentasikan format data yang asli ke dalam format SIG.

1. Data *Output*

Data *output* berfungsi untuk menampilkan atau menghasilkan luaran hasil dari suatu proses, baik dalam bentuk *softcopy* maupun *hardcopy* seperti: tabel, grafik, peta dan lain – lain.

1. Data *Management*

Data *management* berfungsi untuk memanajemen data, baik data *spasial* maupun atribut ke dalam penyimpanan seperti basis data dengan sedemikian rupa agar dipanggil dan diubah dengan mudah.

1. Data Manipulasi dan Analisis

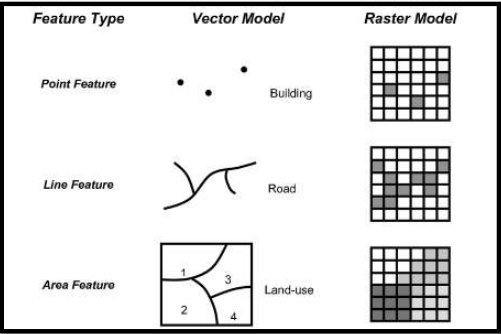
Data manipulasi dan analisis berfungsi untuk menentukan informasi mana saja yang dapat dihasilkan oleh SIG. Subsistem ini memiliki fungsi lain, yakni mampu memanipulasi dan melakukan permodelan data untuk menghasilkan luaran yang diharapkan.

Sistem Informasi Geografis tentunya membutuhkan data untuk di olah. Data yang di olah dan dihasilkan oleh suatu Sistem Informasi Geografis (SIG) terdapat dua model data, yaitu :

1. Data *Spasial*

Data *spasial* merupakan data yang memuat gambaran permukaan bumi. Model data ini dibagi menjadi dua, model data *raster* dan model data *vektor*,

1. Model data *raster* merupakan data yang sederhana, dimana setiap data atau informasi disimpan di *grid*, yang berupa bidang. *Grid* tersebut biasa disebut dengan *pixel*. Data tersebut merupakan hasil dari scanning seperti citra satelit digital
2. Model data *vektor* berupa simbol – simbol atau lebih dikenal dengan istilah *feature*, seperti *feature* garis (*line*), *feature* daerah (*area*), dan *feature* titik (*point*).



Gambar 2.1 Gambaran Data Spasial

1. Data *Atribut* / Data Non *Spasial* : merupakan data yang menyimpan suatu *atribut* dari gambaran yang ada di permukaan bumi.

### *K-Means Clustering*

Menurut Fina dalam (Rahayu, 2022) Algoritma K-Means *clustering* merupakan suatu metode penganalisaan data atau metode data mining yang melakukan proses pemodelan tanpa supervisi (unsupervised) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Metode K-Means berusaha mengelompokan data yang ada kedalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada didalam kelompok yang lain. Algoritma K-Means *Clustering*  ini pertama kali diterbitkan pada tahun 1955 dan terus digunakan sampai sekarang (Preddy et al., 2023). Rahmat juga menjelasakan dalam (Andrea Santana Adzani, 2022) bahwa algoritma K – Means merupakan algoritma pengelompokan data berdasarkan titik pusat cluster (centroid) paling dekat dengan data. Tujuan K – Means adalah pengelompokan data yang memaksimalkan kesamaan data yang dikelompokkan dan meminimalkan kesamaan data antara cluster. Persamaan fungsi jarak digunakan dalam cluster. Maksimalkan kesamaan data berdasarkan jarak terpendek antara data ke titik pusat. Sedangkan pengertian *clustering* yang dijelaskan oleh Rahmat dalam (Andrea Santana Adzani, 2022) menyatakan bahwa *clustering* merupakan proses dalam membagi data yang semulanya tidak berlabel menjadi sekumpulan data yang membentuk kelompok berdasarkan kemiripan yang dimiliki oleh data tersebut dengan data lainnya.

Dalam tahapan algoritma *K-Means Clustering­* terdapat tahap menghitung jarak dari masing-masing data yang ada terhadap masing-masing pusat *cluster* (*centroid*). Ada beberapa persamaan yang digunakan seperti *euclidean distance* dan *manhattan.* Menurut Dinata dalam (Alifah & Fauzan, 2023) persamaan *manhattan* lebih akurat daripadqa persamaan *euclidean distance.*

Menurut Ramadhani dalam (Alifah & Fauzan, 2023) untuk menerapkan algoritma K-Means *Clustering* ada beberapa tahapan yang digambarkan pada Gambar 2.2.



Gambar 2.2 Tahapan Algoritma *K-Means Clustering*

Tahapan Algoritma *K-Means Clustering* berdasarkan Gambar 2.1 dapat dijelaskan sebagai berikut :

1. Menentukan nilai k sebagai jumlah kluster yang akan dibentuk
2. Tentukan titik pusat (*centroid*) awal secara acak dari setiap kluster yang telah ditentukan
3. Hitung jarak dari setiap objek terhadap masing masing titik pusat (*centroid*) dari masing masing kluster dengan persamaan *Manhattan.* yang tertera pada persamaan (2.1) :

Keterangan :   
d = jarak antar x dan y  
x = data pada pusat kluster  
y = data pada atribut

1. Kelompokkan masing masing objek ke dalam titik pusat (*centroid*) terdekat
2. Lakukan iterasi, kemudian hitung pusat cluster (centroid) terbaru, dengan persamaan (2.2) :
3. Ulangi langkah c – e , hingga data tidak berpindah lagi ke *cluster* yang lain.

### Davies-Bouldin Index (DBI)

*Davies-Bouldin Index* (DBI) metode yang digunakan untuk mengevaluasi hasil cluster. Validasi yang dilakukan adalah seberapa baik clustering yang sudah dilakukan dengan menghitung kuantitas dan fitur turunan dari data set. Jumlah klaster terbaik ditunjukkan dengan index DB yang semakin kecil (Alifah & Fauzan, 2023). Tahapan pengujian *Davies-Bouldin Index* (DBI) adalah sebagai berikut :

1. Sum Of Square Within-Cluster (SSW).

Perhitungan Sum Of Square Within-Cluster (SSW) bertujuan untuk mengetahui seberapa keterikatan atau kemiripan dalam anggota satu cluster. Semakin kecil nilai yang di dapat semakin bagus karena semakin mirip. Persamaan SSW yang dimaksud, seperti persamaan (2.3):

Keterangan :  
m1 = jumlah data dalam *cluster* data ke-i  
cj = centroid cluster ke-i  
d( xj, cj) = jaka setiap data

1. Sum of Square Between-Cluster (SSB)

Perhitungan Sum of Square Between-Cluster (SSB) bertujuan untuk mengetahui separasi antar cluster atau seberapa besar perbedaan antar cluster sehingga terpisah ke dalam klompok lain. Semakin besar nilainya maka semakin bagus. Persamaan SSB dituliskan pada persamaan (2.4)

Keterangan :  
*d(xi, xj )* = jarak antara data ke-I dan data ke-j di *cluster* lain

1. Ratio (Rasio)

Perhitungan rasio bertujuan untuk mengetahui nilai perbandingan antara cluster ke-i dan cluster ke-j untuk menghitung rasio yang dimiliki masing-masing cluster. Persamaan rasio dituliskan oada persamaan (2.5)

Keterangan :   
(SSW)i: Sum Of Square Within-Cluster pada centroid i   
(SSB)ij : Sum of Square Between Cluster data ke i dengan j pada cluster yang berbeda.

1. *Davies Bouldin Index* (DBI).

Setelah mendapatkan nilai rasio kemudian menghitung DBI. Semakin kecil nilai DBI yang diperoleh (non negatif >= 0) maka cluster tersebut semakin baik.

Persamaan *Davies Bouldin Index* (DBI) dituliskan seperti persamaan (2.6)

Keterangan :

R(i,j ) : ratio dari nilai SSW dan S

# METODE PENELITIAN

## Tempat dan Waktu Pelaksanaan

Penelitian dengan judul “Implementasi K-Means *Clustering* Dalam Pemetaan Daerah Rawan Curanmor dan Curas Berbasis Sistem Informasi Geografis (Studi Kasus di Kabupaten Probolinggo) “ dilaksanakan di Politek Negeri Jember, Kepolisian Resort Probolinggo, Badan Pusat Statistik Kabupaten Probolinggo. Waktu yang dibutuhkan untuk penelitian ini 10 bulan, dimulai dari bulan Juni 2024 sampai bulan Mei. Detai waktu yang diperlukan pada penelitian ini digambarkan pada table berikut.:

Tabel 3.1 Tahapan Kegiatan

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Kegiatan |  | | | | | | | | | | |
| 2024 | | | | | 2025 | | | | | |
| 7 | 8 | 9 | 10 | 11 | | 12 | 1 | 2 | 3 | 4 |
| Studi Literatur |  |  |  |  |  | |  |  |  |  |  |
| Pengumpulan Data |  |  |  |  |  | |  |  |  |  |  |
| Pengolahan Data |  |  |  |  |  | |  |  |  |  |  |
| Pengembangan Sistem |  |  |  |  |  | |  |  |  |  |  |
| Pengujian |  |  |  |  |  | |  |  |  |  |  |
| Analisis dan pembahasan |  |  |  |  |  | |  |  |  |  |  |

## Alat dan Bahan

### Alat Penelitian

Pada penelitian ini menggunakan alat penelitian berupa perangkat keras dan perangkat lunak, sebagai berikut :

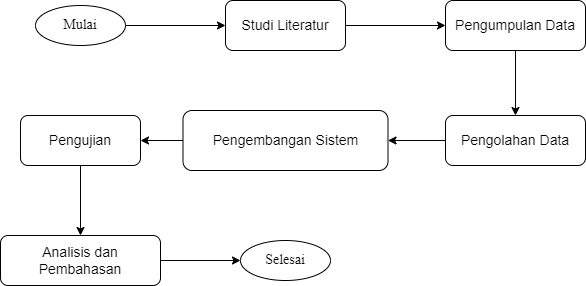
1. Perangkat Keras
2. Laptop Asus AMD E2
3. Smartphone Oppo F9
4. WiFi
5. Perangkat Lunak
6. OS Windows 10
7. *Visual Studio Code*
8. *MySQL*
9. *Framework Laravel*
10. *Library JS* ( *Leaflet* )
11. Microsoft Office

### Bahan Penelitian

Bahan yang digunakan pada penelitian ini adalah data pencurian kendaraan bermotor (curanmor) dan pencurian dengan kekerasan (curas) pada Kabupaten Probolinggo pada tahun 2024. Data tersebut didapat Kepolisian Resort Kabupaten Probolinggo

## Tahapan Penelitian

Pada tahapan penelitian ini terdapat beberapa tahapan yang digambarkan pada gambar 3.1.



Gambar 3.1 Tahapan Penelitian

Pada gambar 3.1 gambar dari tahapan pada penelitian ini yang dilakukan oleh penulis. Tahapan penelitian mencangkup studi literatur, pengumpulan data, pengolahan data, pengembangan sistem, pengujian, analisis dan pembahasan. Penjelasan lebih detail tentang masing-masing tahapan penelitian sebagai berikut :

### Studi Literatur

Studi literatur dilakukan agar peneliti memperoleh referensi maupun teori yang relevan dan sesuai dengan topik penelitian. Referensi yang telah penulis pelajari, antara lain :

1. Jurnal atau paper yang berkaitan dengan algoritma *K-Means Clustering*
2. Jurnal atau paper yang berkaitan dengan *elbow methods*
3. Jurnal atau paper yang berkaitan dengan sistem informasi geografis
4. Data sekunder yang diperoleh dari BPS Kabupaten Probolinggo

### Pengumpulan Data

Untuk mendukung berjalannya penelitian ini, diperlukan data terkait kasus pencurian dengan kekerasan (curas) dan pencurian kendaraan bermotor (curanmor) pada tahun 2024. Pengumpulan data curas dan curanmor pada penelitian ini dilakukan dengan mengajukan permohonan untuk mendapatkan data dari Kepolisian Resort Kabupaten Probolinggo. Pengajuan untuk mendapatkan data yang dilakukan kepada Polres Kabupaten Probolinggo diawali dengan mengirimkan Surat Ijin Survei dan Pengambilan Data, yang terlampir pada Lampiran \*CEK.

### Pengolahan Data

Dalam tahapan ini menjelaskan bagaimana alur atau tahapan data sekunder mentah yang didapat dari publikasi Badan Pusat Statistik (BPS) Kabupaten Probolinggo di olah sebelum memasuiki penerapan algoritma *K-Means Clustering* hingga setelah penerapan algoritma *K-Means Clustering* yang menghasilkan 3 *cluster* dengan masing-masing anggotanya, Berikut ini tahapan pengolahan data pada penelitian ini :

1. Menghitung jumlah banyaknya data
2. Menentukan nilai k sebagai jumlah kluster yang akan dibentuk
3. Tentukan titik pusat (*centroid*) awal secara acak dari setiap kluster yang telah ditentukan
4. Hitung jarak dari setiap objek terhadap masing masing titik pusat (*centroid*) dari masing masing kluster dengan persamaan *Manhattan.* yang tertera pada persamaan (3.1) :

(3.1)

Keterangan :   
d = jarak antar x dan y  
x = data pada pusat kluster  
y = data pada atribut

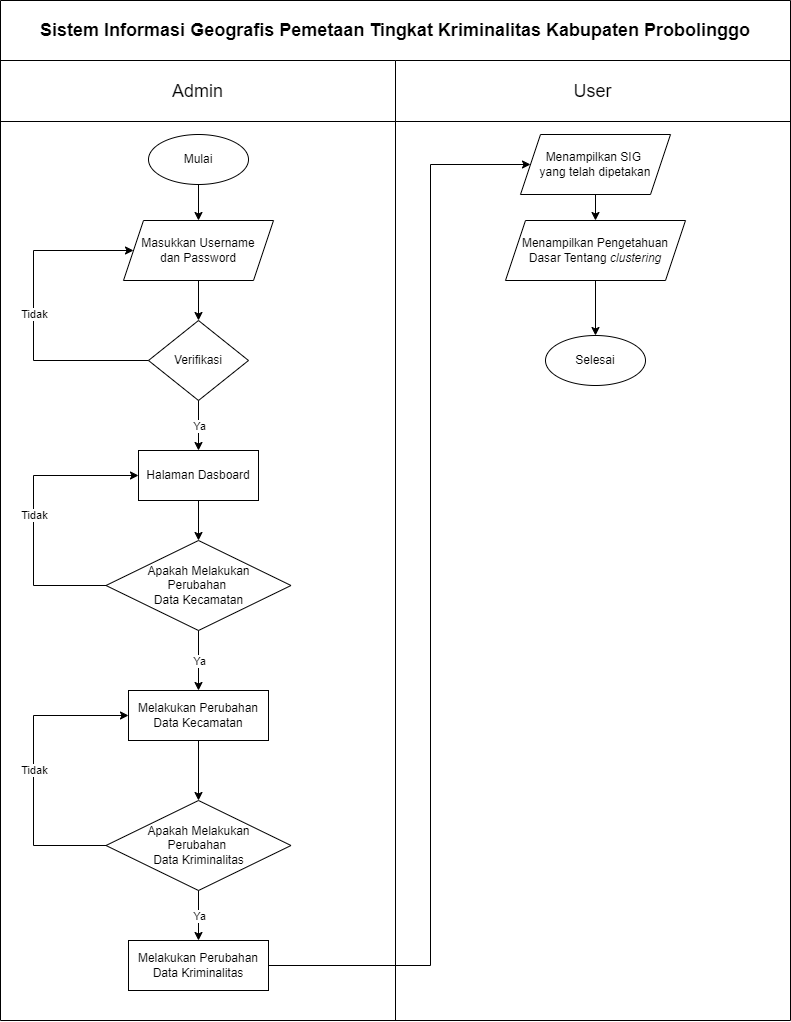
1. Kelompokkan masing masing objek ke dalam titik pusat (*centroid*) terdekat
2. Lakukan iterasi, kemudian hitung pusat cluster (centroid) terbaru, dengan persamaan (3.2) :
3. Ulangi langkah f – h, hingga data tidak berpindah lagi ke *cluster* yang lain

### Pengembangan Sistem

Pada tahap ini peneliti melakukan pengembangan Sistem Informasi Geografis (SIG) yang di dalamnya terdapat pemetaan daerah rawan curas dan curanmor dengan mengimplementasikan metode K-Means *clustering* dan pengukuran jarak antar datanya menggunakan persamaan *manhattan*. Sistem Informasi Geografis ini outputnya akan berbasis *website*, jadi secara umum bahasa pemrograman yang digunakan yaitu php dengan menggunakan juga *framework laravel*. Perhitungan K – Means yang digunakan untuk memetakan tingkat kerawanan suatu kecamatan akan di implementasikan menggunakan bahasa php dengan *function* tersendiri. Kemudian, hasil pemetaan yang telah diperoleh dari perhitungan K – Means akan ditampilkan dalam warna di setiap kecamatan menggunakan salah satu *library* JS yaitu *leaflet*.

Rancangan tentang sistem yang akan dikembangkan dalam penelitian ini digambarkan dalam beberapa diagram sebagai beritkut :

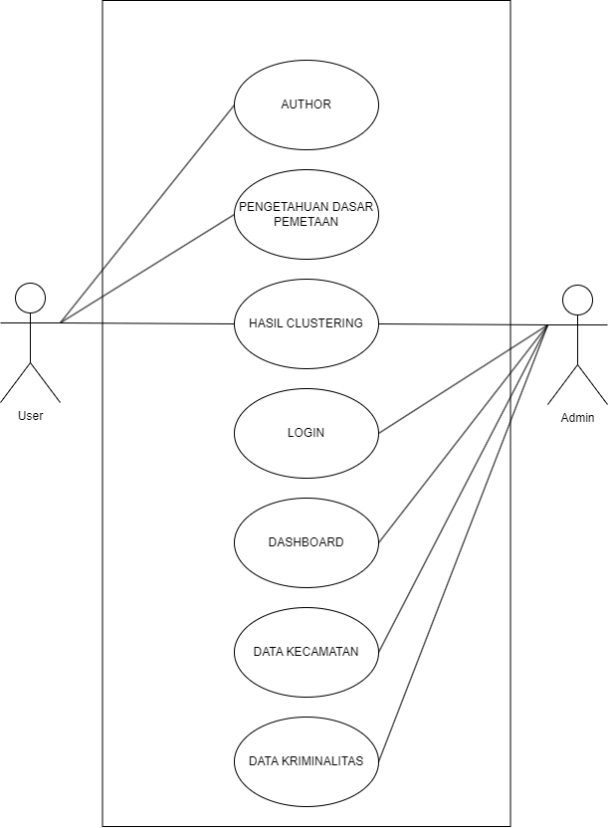
1. Flowchart Sistem



Gambar 3.2 Flowchart Sistem

Pada gambar 3.2 di atas dijelaskan bahwa admin harus masuk atau *login* ke sistem dengan memasukkan username dan password dengan benar. Setelah berhasil masuk ke sistem, admin dapat melakukan perubahan atau input data kecamatan dan juga melakukan perubahan data atau input data kriminalitas ( curas dan curanmor ) pada setiap kecamatan. Dari dua data utama tersebut, sistem akan memproses klusterisasi daerah dengan tingkat kerawanannya. Setelah proses *clustering* selesai, maka hasil *clustering*  tersebut akan di visualisasikan dalam bentuk peta Kabupaten Probolinggo dengan warna di setiap kecamatannya sebagai kategori atas tingkat kerawanan kasus curas dan curanmor pada kecamatan tersebut. Peta Hasil *clustering*  tersebut dapat dilihat secara langsung oleh user, tanpa harus *login* pada sistem.

1. Use Case Sistem



Gambar 3.3 Use Case Sistem

Pada gambar 3.3 dijelaskan terdapat dua pengguna dalam system yang akan dikembangkan. Dua pengguna tersebut merupakan admin dan user yang memiliki hak akses yang berbeda beda dalam system. Admin mempunyai akses pada halaman dashboard, halaman data kecamatan, dan halaman data kriminalitas melalui fitur login terlebih dahulu.. Dapat dinyatakan bahwa admin mempunyai akses yang lebih tinggi terhadap system, karena dengan memiliki akses ke halaman data kecamatan, dan halaman data kriminalitas, maka dapat juga mengubah data data tersebut yang dapat mempengaruhi hasil *clustering* yang terjadi. Kemudian untuk *user* memiliki hak akses pada halaman hasil *clustering,* halaman pengetahuan dasar pemetaan, dan halaman author, maka dapat dinyatakan bahwa user hanya dapat melihat data yang ditampilkan dan tidak dapat merubahnya.

### Pengujian

Pengujian adalah proses uji coba sistem, Tujuannya yaitu untuk memastikan kesesuaian perangkat lunak dengan rancangan yang telah ditetapkan serta untuk mengevaluasi kinerja fungsionalitas sistem agar dapat menentukan apakah semuanya berjalan dengan baik atau tidak. Proses pengujian dalam penelitian ini menggunakan dua metode yaitu:

1. Blackbox Testing

Menurut Hady dalam (Andrea Santana Adzani, 2022) blackbox testing merupakan salah satu metode pengujian yang berbeda dengan whitebox testing, dimana blackbox testing sendiri lebih sering dikatakan sebagai pengujian fungsionalitas sistem, tanpa harus mengetahui bentuk atau struktur program didalamnya. Blackbox testing menguji setiap fitur/menu yang disajikan oleh sistem, apakah setiap fitur/menu tersebut telah memberikan hasil yang sesuai dengan target awal atau hasil yang telah diharapkan diawal..

1. *User Acceptance Testing* (UAT)

*User Acceptance Testing* (UAT) merupakan salah satu pengujian sistem yang berfokus menguji interaksi antara *user* atau pengguna dengan sistem secara langsung yang berfungsi untuk memverifikasi bahwa fitur telah berjalan sesuai dengan kebutuhan user tersebut (Rumariana & Arifin, 2022). Hady menyatakan dalam (Andrea Santana Adzani, 2022) bahwa *User Acceptance Testing* (UAT) memiliki tujuan untuk mengetahui, apakah sistem yang telah dirancang telah memenuhi harapan pengguna, sehingga dapat mempermudah peneliti untuk mengetahui, bagian mana yang masih dirasa kurang. Output dari pengujuan *User Acceptance Testing* (UAT) yaitu dokumen hasil uji *software* dengan nilai yang baikdan sudah memenuhi kebutuhan yang diminta atau memenuhi kriteria (*acceptance criteria*) yang dibutuhkan pengguna.

1. Pengujian *David-Bouldien Index* (DBI)

Pada tahap ini dilakukan pengujian keakurasian dengan menggunakan metode *Davies-Bouldin Index* (DBI). Pada pengujian akurasi ini bertujuan untuk mengetahui hasilnya apakah tingkat keakurasian cluster yang terbentuk sudah menunjukkan nilai yang baik. Untuk memastikan hasil akurasinya dinyatakan baik perlu dibandingkan dan analisis antara hasil clustering pada data curas dan curanmor tahun 2021 – 2022 dengan tahun 2022 – 2023.

Tahapan pengujian *Davies-Bouldin Index* (DBI) adalah sebagai berikut (Alifah & Fauzan, 2023) :

1. Menentukan Sum Of Square Within-Cluster (SSW).

Perhitungan Sum Of Square Within-Cluster (SSW) bertujuan untuk mengetahui seberapa keterikatan atau kemiripan dalam anggota satu cluster. Semakin kecil nilai yang di dapat semakin bagus karena semakin mirip. Persamaan SSW yang dimaksud, seperti pada persamaan (3.3)

Keterangan :  
m1 = jumlah data dalam *cluster* data ke-i  
cj = centroid cluster ke-i  
d( xj, cj) = jaka setiap data

1. Mennentukan *Sum of Square Between-Cluster* (SSB)

Perhitungan *Sum of Square Between-Cluster* (SSB) bertujuan untuk mengetahui separasi antar cluster atau seberapa besar perbedaan antar cluster sehingga terpisah ke dalam klompok lain. Semakin besar nilainya maka semakin bagus. Persamaan SSB dituliskan pada persamaan (3.4)

Keterangan :  
*d(xi, xj )* = jarak antara data ke-I dan data ke-j di *cluster* lain

1. Menentukan *Ratio* (Rasio)

Perhitungan rasio bertujuan untuk mengetahui nilai perbandingan antara cluster ke-i dan cluster ke-j untuk menghitung rasio yang dimiliki masing-masing cluster. Persamaan rasio dituliskan pada persamaan (3.5).

Keterangan :   
(SSW)i: *Sum Of Square Within-Cluster* pada centroid i   
(SSB)ij : *Sum of Square Between Cluster* data ke i dengan j pada cluster yang berbeda.

1. Menentukan nilai Davies Bouldin Index (DBI).

Setelah mendapatkan nilai rasio kemudian menghitung DBI. Semakin kecil nilai DBI yang diperoleh (non negatif >= 0) maka cluster tersebut semakin baik. Persamaan DBI dituliskan pada persamaan (3.6)

Keterangan :

R(i,j ) : ratio dari nilai SSW dan SSB

### Analisis dan Pembahasan

Pada tahap penelitian ini, akan memberikan sebuah penjelasan akhir dari pembahasan pada tahap tahap sebelumnya. Kemudian pada tahap ini juga dilakukan analisis terhadap hasil *clustering* akhir dan juga hasil pengujian tingkat akurasi algoritma K – Means dalam penerapannya pada sistem informasi geografis tingkat kerawanan kasus curas dan curanmor di wilayah hukum Polres Kabupaten Probolinggo.

# HASIL DAN PEMBAHASAN

## Studi Literatur

## Pengumpulan Data

Sesuai dengan perencanaan, penelitian ini menggunakan data pencurian dengan kekerasan (curas) dan pencurian kendaraan bermotor (curanmor) di tahun 2024 yang didapatkan dari Kepolisian Resort Kabupaten Probolinggo. Proses pengajuan untuk mendapatkan data Kepada Polres Kabupaten Probolinggo dilakukan selama 6 bulan yang terhitung dari tanggal 25 Juni 2024 hingga 31 Januari 2025. Proses pengambilan data dari Polres Kabupaten Probolinggo dilampirkan pada Lampiran \*CEK

Data yang didapatkan dari Polres Kabupaten Probolinggo berupa data curas dan curanmor pada masing-masing kecamatan di wilayah Kabupaten Probolinggo. Data yang dikumpulkan terlihat pada gambar \*CEK

|  |  |  |  |
| --- | --- | --- | --- |
| **Kecamatan** | **2024** | | **Total Kasus  Per - Kecamatan** |
| **Curas** | **Curanmor** |
| Bantaran | 0 | 5 | 5 |
| Banyuanyar | 0 | 4 | 4 |
| Besuk | 0 | 2 | 2 |
| Dringu | 0 | 22 | 22 |
| Gading | 1 | 4 | 5 |
| Gending | 1 | 18 | 19 |
| Kotaanyar | 0 | 0 | 0 |
| Kraksaan | 0 | 37 | 37 |
| Krenjengan | 0 | 9 | 9 |
| Krucil | 0 | 3 | 3 |
| Kuripan | 0 | 2 | 2 |
| Leces | 0 | 13 | 13 |
| Lumbang | 0 | 1 | 1 |
| Maron | 0 | 21 | 21 |
| Paiton | 0 | 14 | 14 |
| Pakuniran | 0 | 4 | 4 |
| Pajarakan | 0 | 10 | 10 |
| Sukapura | 0 | 0 | 0 |
| Sumber | 0 | 1 | 1 |
| Sumberasih | 1 | 10 | 11 |
| Tegalsiwalan | 0 | 1 | 1 |
| Tiris | 0 | 2 | 2 |
| Tongas | 3 | 15 | 18 |
| Wonomerto | 2 | 4 | 6 |
| **Total** | **8** | **202** | **210** |

Pada tabel \*CEK merupakan data kasus curas dan curnamor yang terjadi pada setiap kecamatan di Kabupaten Probolinggo sepanjang tahun 2024. Untuk total kasus curas ( pencurian dengan kekerasan ) yang terjadi di seluruh wilayah Kabupaten Probolinggo pada tahun 2024 sebanyak 8 kasus. Kecamatan yang memiliki kasus curas tertinggi yaitu Kecamatan Tongas dengan 3 kasus curas. Kemudian, untuk total kasus curanmor ( pencurian kendaraaan bermotor ) yang terjadi di seluruh wilayah Kabupaten Probolinggo pada tahun 2024 sebanyak 202 kasus. Kecamatan yang memiliki kasus curanmor tertinggi pada tahun 2024 yaitu Kecamatan Kraksaan dengan 37 kasus.

## Pengolahan Data

Tahap pengolahan data pada penelitian ini, menjelaskan bagaimana implementasi algoritma *k-means* agar bisa mengolah data mentah, yang berupa data curas dan curanmor dari Polres, agar bisa memberikan kategori untuk setiap kecamatan berdasarkan kasus curas dan curanmor yang terjadi pada kecamatan tersebut. Pengolahan data yang dilakukan, tentunya mengikuti tahapan-tahapan dari algoritma *k-means*. Dalam hal ini penulis memaparkan proses perhitungan algoritma *k-means* dengan nilai k=3 untuk data curas dan curanmor. Alasan penentuan nilai k = 3, karena nilai tersebut merupakan nilai yang optimal untuk data curas dan curanmor, berdasarkan pengujian yang telah dilakukan pada poin 4.5. Berikut tahapan dari perhitungan algoritma *k-means* dengan nilai k = 3 untuk masing-masing data curas dan curanmor :

### Menghitung Jumlah Data

Pada tahap ini dilakukan perhitungan jumlah data yang didapatkan dari Polres. Jumlah data ini perlu dihitung karena juga sebagai dasar untuk menentukan nilai k atau jumlah klaster nanti. Pada data yang didapat dari Polres, menjelaskan data curas dan curanmor dari masing-masing kecamatan, dan ada 24 kecamatan di wilayah Kabupaten Probolinggo. Jadi dapat dinyatakan jumlah data untuk masing-masing kasus curas dan curanmor ada 24 data.

### Menentukan nilai k

Dalam menentukan nilai k atau jumlah klaster yang akan dibentuk, perlu memperhatikan jumlah data yang ada. Karena penentuan nilai k harus lebih kecil atau sama dengan jumlah data. Dalam studi kasus ini penulis menentukan nilai k = 3 untuk kedua data curas dan curanmor, dikarenakan nilai tersebut merupakan nilai k yang paling optimal berdasarkan hasil pengujian pada poin 4.5.

### Menentukan Centroid Awal Pada Setiap Klaster

Setelah mengetahui atau menentukan nilai k yang optimal untuk masing-masing data curas dan curanmor, selanjutnya perlu menentukan nilai centroid awal untuk setiap klaster. Pada penelitian ini nilai k yang telah ditetapkan yaitu 3, baik untuk data curas maupun data curanmor, sehingga perlu 3 nilai centroid untuk masing-masing data. Dalam menentukan nilai centroid awal, tidak ada aturan yang mengatur bagaimana pemilihan centroid awal, atau dengan kata lain, penentuan nilai centroid awal dipilih secara acak. Berikut nilai centroid awal yang telah dipilih secara acak untuk masing-masing data.

Centroid awal untuk data curas

|  |  |
| --- | --- |
| Centroid | Nilai Centroid |
| C1 | 0 |
| C2 | 1 |
| C3 | 3 |

Sesuai pada tabel \*CEK, nilai centroid awal yang ditetapkan untuk data curas Kabupaten Probolinggo yaitu C1 = 0, C2 =1, dan untuk C3 = 3.

Centroid awal untuk data curanmor

|  |  |
| --- | --- |
| Centroid | Nilai Centroid |
| C1 | 10 |
| C2 | 20 |
| C3 | 30 |

Sesuai pada tabel \*CEK, nilai centroid awal yang ditetapkan untuk data curanmor Kabupaten Probolinggo yaitu C1 = 0, C2 =1, dan untuk C3 = 3.

### Menghitung Jarak Setiap Data Terhadap Centroid

Setelah menentukan nilai centroid awal pada setiap klaster, sesuai dengan tahapan algoritma k-means langkah selanjutnya yaitu menghitung jarak dari masing masing data terhadap masing-masing centroid. Pada tahap perhitungan jarak ini dilakukan berulang kali hingga tidak ada perubahan pada anggota dari masing-masing klaster pada iterasi selanjutnya. Perhitungan jarak berulang tersebut disebut sebagai iterasi. Dalam perhitungan jarak ini digunakan persamaan *Euclidean Distance*, yang tertuang pada persamaan nomor 4.1

(4.1)

Keterangan :

: Jarak data curas pada kecamatan i terhadap centroid ke i

: Data curas atau curanmor pada kecamatan ke i

C : Centroid ke i

Jadi pada persamaan *euclidean distance* penulis menghitung jarak antar data pada masing-masing kecamatan di setiap kasus terhadap masing-masing centroid pada setiap kasusnya.

* 1. Perhitungan Jarak Data Kasus Curas

Dalam iterasi pertama ini penulis menggunakan data curas pada kecamatan Bantaran sebagai contoh perhitungan jarak data terhadap setiap centroid. Pada Kecamatan Bantaran terjadi kasus curas sebanyak 0 kasus.

= 0

= 1

= 3

Pada iterasi pertama ini seluruh data curas pada setiap kecamatan dihitung jaraknya terhadap masing-masing nilai centroid awal. Setelah setiap data curas dihitung jaraknya, kemudian dicari nilai minimum dari jarak data terhadap masing-masing centroid yang ada. Jarak minimum tersebut merupakan acuan dalam menentukan data tersebut masuk dalam klaster apa. Semakin kecil jarak data terhadap salah satu centroid pada suatu klaster, maka data tersebut merupakan anggota dari klaster tersebut. Hasil perhitungan jarak data curas terhadap masing-masing centroid bisa dilihat pada tabel \*CEK

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Kecamatan** | **C1** | **C2** | **C3** | **Klaster** |
| Bantaran | 0,000 | 1,000 | 3,000 | C1 |
| Banyuanyar | 0,000 | 1,000 | 3,000 | C1 |
| Besuk | 0,000 | 1,000 | 3,000 | C1 |
| Dringu | 0,000 | 1,000 | 3,000 | C1 |
| Gading | 1,000 | 0,000 | 2,000 | C2 |
| Gending | 1,000 | 0,000 | 2,000 | C2 |
| Kotaanyar | 0,000 | 1,000 | 3,000 | C1 |
| Kraksaan | 0,000 | 1,000 | 3,000 | C1 |
| Krenjengan | 0,000 | 1,000 | 3,000 | C1 |
| Krucil | 0,000 | 1,000 | 3,000 | C1 |
| Kuripan | 0,000 | 1,000 | 3,000 | C1 |
| Leces | 0,000 | 1,000 | 3,000 | C1 |
| Lumbang | 0,000 | 1,000 | 3,000 | C1 |
| Maron | 0,000 | 1,000 | 3,000 | C1 |
| Paiton | 0,000 | 1,000 | 3,000 | C1 |
| Pakuniran | 0,000 | 1,000 | 3,000 | C1 |
| Pajarakan | 0,000 | 1,000 | 3,000 | C1 |
| Sukapura | 0,000 | 1,000 | 3,000 | C1 |
| Sumber | 0,000 | 1,000 | 3,000 | C1 |
| Sumberasih | 1,000 | 0,000 | 2,000 | C2 |
| Tegalsiwalan | 0,000 | 1,000 | 3,000 | C1 |
| Tiris | 0,000 | 1,000 | 3,000 | C1 |
| Tongas | 3,000 | 2,000 | 0,000 | C3 |
| Wonomerto | 2,000 | 1,000 | 1,000 | C2 |

Pada iterasi pertama proses klasterisasi data curas, jumlah data yang masuk ke masing-masing klaster adalah : C1 sebanyak 19 data, C2 terdiri dari 4 data, dan untuk C3 berjumlah 1 data. Meskipun demikian, jumlah ini masih dapat berubah pada iterasi berikutnya karena proses *optimasi* akan memperbarui posisi *centroid*, sehingga pembagian data ke dalam klaster bisa mengalami penyesuaian berdasarkan perhitungan jarak yang terbaru. Untuk mempermudah dalam mengetahui hasil *clustering* pada iterasi pertama, dapat dilihat pada tabel \*CEK

|  |  |  |
| --- | --- | --- |
| C1 | C2 | C3 |
| Bantaran | Gading | Tongas |
| Banyuanyar | Gending |  |
| Besuk | Sumberasih |  |
| Dringu | Wonomerto |  |
| Kotaanyar |  |  |
| Kraksaan |  |  |
| Krenjengan |  |  |
| Krucil |  |  |
| Kuripan |  |  |
| Leces |  |  |
| Lumbang |  |  |
| Maron |  |  |
| Paiton |  |  |
| Pakuniran |  |  |
| Pajarakan |  |  |
| Sukapura |  |  |
| Sumber |  |  |
| Tegalsiwalan |  |  |
| Tiris |  |  |

Hasil *clustering* pada tabel \*CEK akan menjadi acuan dalam perhitungan *centroid* untuk iterasi selanjutnya.

* 1. Perhitungan Jarak Data Kasus Curanmor

Dalam iterasi pertama ini penulis menggunakan data curanmor pada kecamatan Bantaran sebagai contoh perhitungan jarak data terhadap setiap centroid. Pada Kecamatan Bantaran terjadi kasus curanmor sebanyak 0 kasus.

= 5

= 15

= 25

Pada iterasi pertama ini seluruh data curanmor pada setiap kecamatan dihitung jaraknya terhadap masing-masing nilai centroid awal. Setelah setiap data curanmor dihitung jaraknya, kemudian dicari nilai minimum dari jarak data terhadap masing-masing centroid yang ada. Jarak minimum tersebut merupakan acuan dalam menentukan data tersebut masuk dalam klaster apa. Semakin kecil jarak data terhadap salah satu centroid pada suatu klaster, maka data tersebut merupakan anggota dari klaster tersebut. Hasil perhitungan jarak data curanmor terhadap masing-masing centroid bisa dilihat pada tabel \*CEK

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Kecamatan** | **C1** | **C2** | **C3** | **Klaster** |
| Bantaran | 5,000 | 15,000 | 25,000 | C1 |
| Banyuanyar | 6,000 | 16,000 | 26,000 | C1 |
| Besuk | 8,000 | 18,000 | 28,000 | C1 |
| Dringu | 12,000 | 2,000 | 8,000 | C2 |
| Gading | 6,000 | 16,000 | 26,000 | C1 |
| Gending | 8,000 | 2,000 | 12,000 | C2 |
| Kotaanyar | 10,000 | 20,000 | 30,000 | C1 |
| Kraksaan | 27,000 | 17,000 | 7,000 | C3 |
| Krenjengan | 1,000 | 11,000 | 21,000 | C1 |
| Krucil | 7,000 | 17,000 | 27,000 | C1 |
| Kuripan | 8,000 | 18,000 | 28,000 | C1 |
| Leces | 3,000 | 7,000 | 17,000 | C1 |
| Lumbang | 9,000 | 19,000 | 29,000 | C1 |
| Maron | 11,000 | 1,000 | 9,000 | C2 |
| Paiton | 4,000 | 6,000 | 16,000 | C1 |
| Pakuniran | 6,000 | 16,000 | 26,000 | C1 |
| Pajarakan | 0,000 | 10,000 | 20,000 | C1 |
| Sukapura | 10,000 | 20,000 | 30,000 | C1 |
| Sumber | 9,000 | 19,000 | 29,000 | C1 |
| Sumberasih | 0,000 | 10,000 | 20,000 | C1 |
| Tegalsiwalan | 9,000 | 19,000 | 29,000 | C1 |
| Tiris | 8,000 | 18,000 | 28,000 | C1 |
| Tongas | 5,000 | 5,000 | 15,000 | C1 |
| Wonomerto | 6,000 | 16,000 | 26,000 | C1 |

Pada iterasi pertama proses klasterisasi data curanmo, jumlah data yang masuk ke masing-masing klaster adalah : C1 sebanyak 20 data, C2 terdiri dari 3 data, dan untuk C3 berjumlah 1 data. Meskipun demikian, jumlah ini masih dapat berubah pada iterasi berikutnya karena proses *optimasi* akan memperbarui posisi *centroid*, sehingga pembagian data ke dalam klaster bisa mengalami penyesuaian berdasarkan perhitungan jarak yang terbaru. Untuk mempermudah dalam mengetahui hasil *clustering* pada iterasi pertama, dapat dilihat pada tabel \*CEK

|  |  |  |
| --- | --- | --- |
| C1 | C2 | C3 |
| Bantaran | Dringu | Kraksaan |
| Banyuanyar | Gending |  |
| Besuk | Maron |  |
| Gading |  |  |
| Kotaanyar |  |  |
| Krenjengan |  |  |
| Krucil |  |  |
| Kuripan |  |  |
| Leces |  |  |
| Lumbang |  |  |
| Paiton |  |  |
| Pakuniran |  |  |
| Pajarakan |  |  |
| Sukapura |  |  |
| Sumber |  |  |
| Sumberasih |  |  |
| Tegalsiwalan |  |  |
| Tiris |  |  |
| Tongas |  |  |
| Wonomerto |  |  |

Hasil *clustering* pada tabel \*CEK akan menjadi acuan dalam perhitungan *centroid* untuk iterasi selanjutnya.

### Menentukan Nilai Centroid Baru

Guna mengetahui apakah hasil *clustering* dari iterasi pertama sudah benar, perlu dilakukan iterasi kedua. Iterasi kedua digunakan untuk memvalidasi, apakah anggota klaster yang telah dihasilkan oleh iterasi pertama masih sama pada iterasi kedua ataukah ada data yang berubah klasternya. Proses dalam perhitungan iterasi kedua sama dengan iterasi pertama. Namun centroidnya tidak menggunakan nilai acak lagi, melainkan berpedoman pada hasil *clustering* dari iterasi pertama. Proses menentukan centroid yang baru dengan menghitung rata-rata data pada setiap klaster yang sudah terbentuk di iterasi pertama atau dapat dirumuskan dalam persamaan \*CEK

Keterangan :

: Nilai Centroid pada klaster ke-i yang baru.

: Jumlah seluruh data pada anggota klaster ke-i

: Jumlah anggota klaster ke-i

Berikut merupakan perhitungan centroid baru dari masing-masing kasus curas dan curanmor.

1. Perhitungan Centroid Baru Data Curas

= 0

= 1,25

= 3

Dari perhitungan tersebut sudah menghasilkan nilai centroid baru yang berbeda dari centroid acak di awal. Hasil perhitungan yang berupa centroid baru dapat dilihat pada tabel \*CEK

|  |  |
| --- | --- |
| Centroid | Nilai Centroid |
| C1 | 0 |
| C2 | 1,25 |
| C3 | 3 |

Centroid baru untuk kasus curas pada tabel\*CEK akan digunakan untuk perhitungan jarak data pada iterasi kedua.

1. Perhitungan Centroid Baru Data Curanmor

Selanjutnya perlu juga menghitung centroid baru untuk kasus curanmor dengan berdasarkan hasil *clustering* dari ietrasi pertama.

Dari perhitungan tersebut sudah menghasilkan nilai centroid baru yang berbeda dari centroid acak di awal. Hasil perhitungan yang berupa centroid baru dapat dilihat pada tabel \*CEK

|  |  |
| --- | --- |
| Centroid | Nilai Centroid |
| C1 | 5,2 |
| C2 | 20,333 |
| C3 | 37 |

Centroid baru untuk kasus curanmor pada tabel\*CEK akan digunakan untuk perhitungan jarak data pada iterasi kedua.

### Melakukan Iterasi Selanjutnya

Setelah mendapatkan centroid baru dari hasil iterasi pertama, langkah selanjutnya melakukan iterasi kedua dengan menggunakan centroid baru tersebut. Pada iterasi kedua juga mengelompokkan data ke klaster berdasarkan jarak terdekat dari data tersebut ke masing-masing centroid.

Setelah mengetahui anggota klaster yang baru dari iterasi kedua, dilakukan pengecekan apakah anggota klaster iterasi kedua sama dengan anggota kaster dari iterasi pertama tadi. Jika tidak ada perubahan anggota klaster dari iterasi pertama ke iterasi kedua, maka proses *clustering* sudah bisa dianggap konvergen, sehingga hasil akhir dari iterasi kedua merupakan hasil akhir dari proses *K-Means* *Clustering.*

Namun ketika ada perbedaan anggota klaster antara hasil iterasi pertama dan iterasi kedua, maka proses iterasi akan dilanjutkan. Proses tersebut mengulangi tahap penentuan centroid baru dari iterasi sebelumnya, perhitungan jarak untuk iterasi selanjutnya, dan pengecekan anggota klaster dari iterasi sebelumnya dengan iterasi selanjutnya hingga mencapai konvergen atau tidak ada perubahan anggota klaster pada dua iterasi terakhir.

### Hasil Akhir K-Means Clustering

Setelah melanjutkan iterasi kedua dan seterusnya untuk masing-masing data kasus, proses iterasi *k-means* *clustering* berhenti sesuai ketentuan ketika tidak ada perubahan anggota klaster lagi. Iterasi terakhir pada masing-masing kasus berbeda. Berikut iterasi terakhir dan juga hasil final dari *k-means clustering* pada masing-masing kasus.

1. Iterasi Terakhir dan Hasil Final *Clustering* Kasus Curas

Proses iterasi pada kasus curas berhenti pada iterasi kedua, karena tidak ada lagi perubahan anggota klaster antara iterasi pertama ke iterasi kedua. Selain itu untuk centroid terakhirnya merupakan centroid yang dihasilkan dari hasil klasterisasi pada iterasi pertama. Centroid terakhir yang digunakan dan juga perhitungan jarak di iterasi kedua (iterasi terakhir) dicantumkan pada tabel \*CEK

|  |  |
| --- | --- |
| Centroid | Nilai Centroid |
| C1 | 0 |
| C2 | 1,25 |
| C3 | 3 |

Nilai centroid terakhir pada tabel \*CEK merupakan hasil perhitungan dari iterasi sebelumnya (iterasi pertama). Nilai centroid tersebut yang kemudian digunakan untuk menghitung jarak pada iterasi selanjutnya (iterasi kedua) pada tabel \*CEK.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Kecamatan** | **C1** | **C2** | **C3** | **Klaster** |
| Bantaran | 0,000 | 1,250 | 3,000 | C1 |
| Banyuanyar | 0,000 | 1,250 | 3,000 | C1 |
| Besuk | 0,000 | 1,250 | 3,000 | C1 |
| Dringu | 0,000 | 1,250 | 3,000 | C1 |
| Gading | 1,000 | 0,250 | 2,000 | C2 |
| Gending | 1,000 | 0,250 | 2,000 | C2 |
| Kotaanyar | 0,000 | 1,250 | 3,000 | C1 |
| Kraksaan | 0,000 | 1,250 | 3,000 | C1 |
| Krenjengan | 0,000 | 1,250 | 3,000 | C1 |
| Krucil | 0,000 | 1,250 | 3,000 | C1 |
| Kuripan | 0,000 | 1,250 | 3,000 | C1 |
| Leces | 0,000 | 1,250 | 3,000 | C1 |
| Lumbang | 0,000 | 1,250 | 3,000 | C1 |
| Maron | 0,000 | 1,250 | 3,000 | C1 |
| Paiton | 0,000 | 1,250 | 3,000 | C1 |
| Pakuniran | 0,000 | 1,250 | 3,000 | C1 |
| Pajarakan | 0,000 | 1,250 | 3,000 | C1 |
| Sukapura | 0,000 | 1,250 | 3,000 | C1 |
| Sumber | 0,000 | 1,250 | 3,000 | C1 |
| Sumberasih | 1,000 | 0,250 | 2,000 | C2 |
| Tegalsiwalan | 0,000 | 1,250 | 3,000 | C1 |
| Tiris | 0,000 | 1,250 | 3,000 | C1 |
| Tongas | 3,000 | 1,750 | 0,000 | C3 |
| Wonomerto | 2,000 | 0,750 | 1,000 | C2 |

Pada tabel \*CEK merupakan perhitungan jarak atau iterasi terakhir dalam proses *k-means* *clustering* karena sudah tidak ada perubahan anggota klaster antara iterasi sebelumnya (iterasi pertama) dengan iterasi selanjutnya (iterasi kedua). Hasil *clustering* dari iterasi kedua atau bisa disebut hasil akhir proses *k-means* *clustering* dapat dijelaskan pada tabel \*CEK

|  |  |  |
| --- | --- | --- |
| C1 | C2 | C3 |
| Bantaran | Gading | Tongas |
| Banyuanyar | Gending |  |
| Besuk | Sumberasih |  |
| Dringu | Wonomerto |  |
| Kotaanyar |  |  |
| Kraksaan |  |  |
| Krenjengan |  |  |
| Krucil |  |  |
| Kuripan |  |  |
| Leces |  |  |
| Lumbang |  |  |
| Maron |  |  |
| Paiton |  |  |
| Pakuniran |  |  |
| Pajarakan |  |  |
| Sukapura |  |  |
| Sumber |  |  |
| Tegalsiwalan |  |  |
| Tiris |  |  |

Untuk mempermudah pembaca dalam memahami hasil akhir clustering atau untuk mempermudah mengetahui kecamatan mana yang termasuk kategori rawan., maka hasil akhir clustering pada tabel \*CEK di ubah nama kategorinya yang sebelumnya C1, C2, C3 menjadi Rawan, Sedang, dan Tidak Rawan.

Nama klaster yang baru yaitu Rawan, Sedang, dan Tidak Rawan memiliki karakteristik masing-masing. Nama klaster rawan memiliki arti bahwa kecamatan pada klaster tersebut sering terjadi kasus curas dari pada kecamatan di klaster yang lain. Sedangkan klaster aman memiliki karakteristik kecamatan yang masuk pada klaster ini memiliki kejadian curas lebih sedikit dari kecamatan pada klaster lain. Kemudian untuk klaster sedang, memiliki karakteristik kecamatan yang masuk menjadi anggotanya merupakan kecamatan yang kejadian curasnya tidak lebih banyak kasus curasnya daripada kecamatan yang ada di klaster rawan dan tidak lebih sedikit dari kecamatan yang ada di klaster aman, atau dengan kata lain klaster yang dikategorikan sedang merupakan klaster yang letaknya berada di tengah-tengah dalam diagram persebaran klaster.

Dalam perubahan nama kategori tersebut perlu memperhatikan karakteristik dari masing-masing kategori dengan data yang masuk dalam kategori tersebut. Jadi penamaan kategori yang baru mengikuti centroid terakhir yang digunakan, karena centroid merupakan cerminan dari data yang ada dalam klaster tersebut. Dalam studi kasus ini klaster C3 bisa berubah nama menjadi klaster rawan, karena nilai centroidnya lebih tinggi dari klaster lainnya. Kemudian untuk klaster C1 bisa berubah nama menjadi klaster aman, karena nilai centroidnya lebih kecil dari klaster lainnya. Selanjutnya untuk C2 bisa berubah menjadi klaster sedang, karena nilai centroidnya tidak lebih besari dari C3 dan tidak lebih kecil dari C1. Jadi hasil *clustering* menggunakan nama klaster yang baru dapat dilihat pada tabel \*CEK

|  |  |  |
| --- | --- | --- |
| Aman | Sedang | Rawan |
| Bantaran | Gading | Tongas |
| Banyuanyar | Gending |  |
| Besuk | Sumberasih |  |
| Dringu | Wonomerto |  |
| Kotaanyar |  |  |
| Kraksaan |  |  |
| Krenjengan |  |  |
| Krucil |  |  |
| Kuripan |  |  |
| Leces |  |  |
| Lumbang |  |  |
| Maron |  |  |
| Paiton |  |  |
| Pakuniran |  |  |
| Pajarakan |  |  |
| Sukapura |  |  |
| Sumber |  |  |
| Tegalsiwalan |  |  |
| Tiris |  |  |

Dari tabel \*CEK dapat disimpulkan bahwa klaster rawan hanya memiliki 1 anggota yaitu kecamatan Tongas. Untuk klaster sedang memiliki 4 anggota, yaitu kecamatan Gading, Gending, Sumberasih, dan Wonomerto. Sedangkan untuk klaster aman memiliki 19 anggota yang terdiri dari Kecamatan Bantaran, Banyuanyar, Besuk, Dringu, Kotaanyar, Kraksaan, Krenjengan, Krucil, Kuripan, Leces Lumbang Maron, Paiton, Pakuniran, Pajarakan, Sukapura, Sumber, Tegalsiwalan, dan Tiris.

1. Iterasi Terakhir dan Hasil Final *Clustering* Kasus Curanmor

Proses iterasi pada kasus curanmor berhenti pada iterasi ketiga, karena tidak ada lagi perubahan anggota klaster antara iterasi kedua ke iterasi ketiga. Selain itu untuk centroid terakhirnya merupakan centroid yang dihasilkan dari hasil klasterisasi pada iterasi kedua. Centroid terakhir yang digunakan dan juga perhitungan jarak di iterasi ketiga (iterasi terakhir) dicantumkan pada tabel \*CEK

|  |  |
| --- | --- |
| Centroid | Nilai Centroid |
| C1 | 3,647 |
| C2 | 17,167 |
| C3 | 37,000 |

Nilai centroid terakhir pada tabel \*CEK merupakan hasil perhitungan dari iterasi sebelumnya (iterasi kedua). Nilai centroid tersebut yang kemudian digunakan untuk menghitung jarak pada iterasi selanjutnya (iterasi ketiga) pada tabel \*CEK.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Kecamatan** | **C1** | **C2** | **C3** | **Klaster** |
| Bantaran | 1,353 | 12,167 | 32,000 | C1 |
| Banyuanyar | 0,353 | 13,167 | 33,000 | C1 |
| Besuk | 1,647 | 15,167 | 35,000 | C1 |
| Dringu | 18,353 | 4,833 | 15,000 | C2 |
| Gading | 0,353 | 13,167 | 33,000 | C1 |
| Gending | 14,353 | 0,833 | 19,000 | C2 |
| Kotaanyar | 3,647 | 17,167 | 37,000 | C1 |
| Kraksaan | 33,353 | 19,833 | 0,000 | C3 |
| Krenjengan | 5,353 | 8,167 | 28,000 | C1 |
| Krucil | 0,647 | 14,167 | 34,000 | C1 |
| Kuripan | 1,647 | 15,167 | 35,000 | C1 |
| Leces | 9,353 | 4,167 | 24,000 | C2 |
| Lumbang | 2,647 | 16,167 | 36,000 | C1 |
| Maron | 17,353 | 3,833 | 16,000 | C2 |
| Paiton | 10,353 | 3,167 | 23,000 | C2 |
| Pakuniran | 0,353 | 13,167 | 33,000 | C1 |
| Pajarakan | 6,353 | 7,167 | 27,000 | C1 |
| Sukapura | 3,647 | 17,167 | 37,000 | C1 |
| Sumber | 2,647 | 16,167 | 36,000 | C1 |
| Sumberasih | 6,353 | 7,167 | 27,000 | C1 |
| Tegalsiwalan | 2,647 | 16,167 | 36,000 | C1 |
| Tiris | 1,647 | 15,167 | 35,000 | C1 |
| Tongas | 11,353 | 2,167 | 22,000 | C2 |
| Wonomerto | 0,353 | 13,167 | 33,000 | C1 |

Pada tabel \*CEK merupakan perhitungan jarak atau iterasi terakhir dalam proses *k-means* *clustering* untuk kasus curanmor, karena sudah tidak ada perubahan anggota klaster antara iterasi sebelumnya (iterasi kedua) dengan iterasi saat ini (iterasi ketiga). Hasil *clustering* dari iterasi ketiga atau bisa disebut hasil akhir proses *k-means* *clustering* dapat dijelaskan pada tabel \*CEK

|  |  |  |
| --- | --- | --- |
| C1 | C2 | C3 |
| Bantaran | Dringu | Kraksaan |
| Banyuanyar | Gending |  |
| Besuk | Leces |  |
| Gading | Maron |  |
| Kotaanyar | Paiton |  |
| Krenjengan | Tongas |  |
| Krucil |  |  |
| Kuripan |  |  |
| Lumbang |  |  |
| Pakuniran |  |  |
| Pajarakan |  |  |
| Sukapura |  |  |
| Sumber |  |  |
| Sumberasih |  |  |
| Tegalsiwalan |  |  |
| Tiris |  |  |
| Wonomerto |  |  |

Untuk mempermudah pembaca dalam memahami hasil akhir *clustering* kasus curanmor atau untuk mempermudah mengetahui kecamatan mana yang termasuk kategori rawan., maka hasil akhir clustering pada tabel \*CEK di ubah nama kategorinya yang sebelumnya C1, C2, C3 menjadi Rawan, Sedang, dan Tidak Rawan.

Nama klaster yang baru yaitu Rawan, Sedang, dan Tidak Rawan memiliki karakteristik masing-masing. Nama klaster rawan memiliki arti bahwa kecamatan pada klaster tersebut sering terjadi kasus curanmor dari pada kecamatan di klaster yang lain. Sedangkan klaster aman memiliki karakteristik kecamatan yang masuk pada klaster ini memiliki kejadian curanmor lebih sedikit dari kecamatan pada klaster lain. Kemudian untuk klaster sedang, memiliki karakteristik kecamatan yang masuk menjadi anggotanya merupakan kecamatan yang kasus curanmornya tidak lebih banyak daripada kecamatan yang ada di klaster rawan dan juga tidak lebih sedikit dari kecamatan yang ada di klaster aman, atau dengan kata lain, klaster yang dikategorikan sedang merupakan klaster yang letaknya berada di tengah-tengah dalam diagram persebaran klaster.

Dalam perubahan nama kategori tersebut perlu memperhatikan karakteristik dari masing-masing kategori dengan data yang masuk dalam kategori tersebut. Jadi penamaan kategori yang baru mengikuti centroid terakhir yang digunakan, karena centroid merupakan cerminan dari data yang ada dalam klaster tersebut. Dalam studi kasus ini klaster C3 bisa berubah nama menjadi klaster rawan, karena nilai centroidnya lebih tinggi dari klaster lainnya. Kemudian untuk klaster C1 bisa berubah nama menjadi klaster aman, karena nilai centroidnya lebih kecil dari klaster lainnya. Selanjutnya untuk C2 bisa berubah menjadi klaster sedang, karena nilai centroidnya tidak lebih besari dari C3 dan tidak lebih kecil dari C1. Hasil *clustering* menggunakan nama klaster yang baru dapat dilihat pada tabel \*CEK

|  |  |  |
| --- | --- | --- |
| Aman | Sedang | Rawan |
| Bantaran | Dringu | Kraksaan |
| Banyuanyar | Gending |  |
| Besuk | Leces |  |
| Gading | Maron |  |
| Kotaanyar | Paiton |  |
| Krenjengan | Tongas |  |
| Krucil |  |  |
| Kuripan |  |  |
| Lumbang |  |  |
| Pakuniran |  |  |
| Pajarakan |  |  |
| Sukapura |  |  |
| Sumber |  |  |
| Sumberasih |  |  |
| Tegalsiwalan |  |  |
| Tiris |  |  |
| Wonomerto |  |  |

Dari tabel \*CEK dapat disimpulkan bahwa klaster rawan untuk kasus curanmor hanya memiliki 1 anggota yaitu kecamatan Kraksaan. Untuk klaster sedang pada kasus curanmor memiliki 6 anggota, yaitu kecamatan Dringu, Gending, Leces, Maron, Paiton, dan Tongas. Sedangkan untuk klaster aman memiliki 17 anggota yang terdiri dari Kecamatan Bantaran, Banyuanyar, Besuk, Gading, Kotaanyar, Krenjengan, Krucil, Kuripan, Lumbang, Pakuniran, Pajarakan, Sukapura, Sumber, Sumberasih, Tegalsiwalan, Tiris, dan Wonomerto.

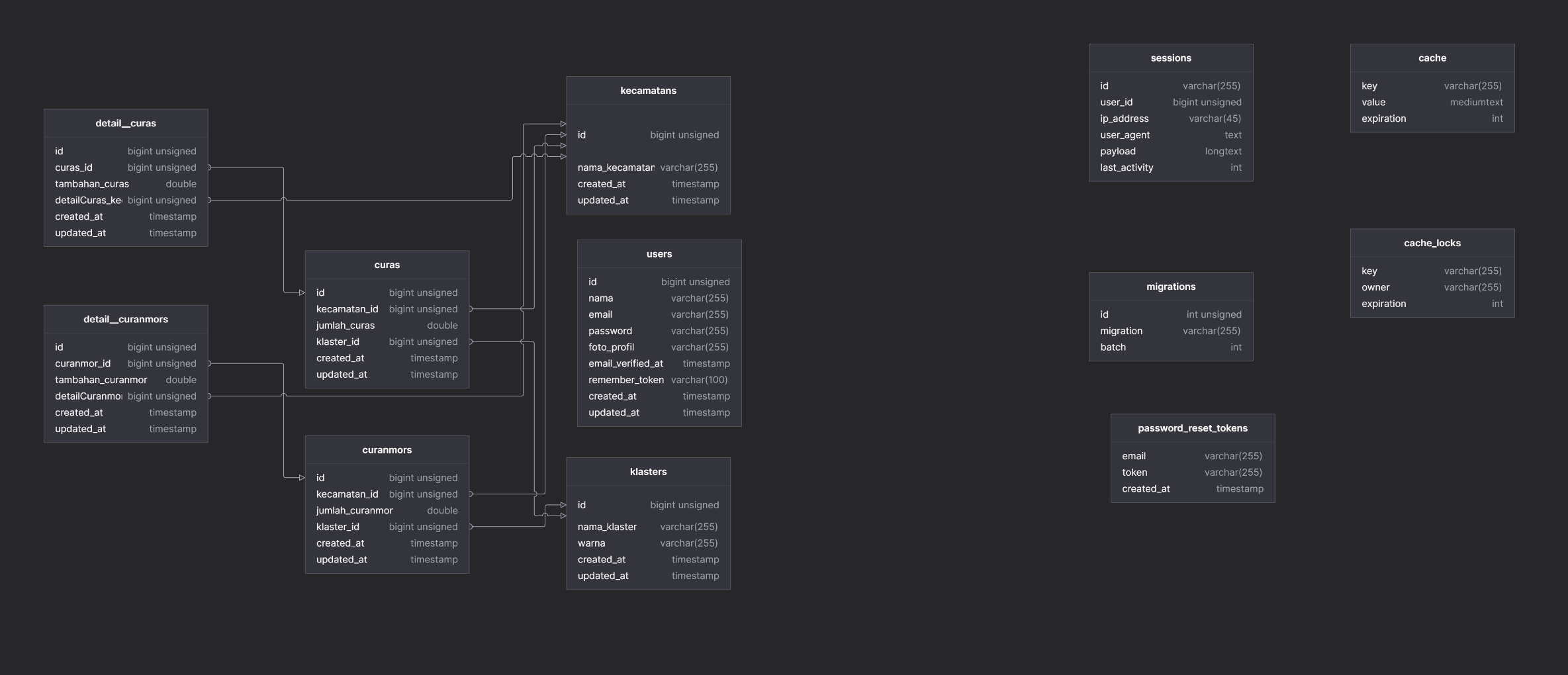
## Pengembangan Sistem

Pada tahapan pengembangan sistem, penulis mulai membuat produk untuk penelitian ini. Produk yang dibuat oleh penulis untuk penelitian ini berupa web sistem informasi geografis yang menampilkan pemetaan dari kasus curas dan juga kasus curanmor pada wilayah Kabupaten Probolinggo. Dalam web tersebut, juga terdapat implementasi dari algoritma k-means clustering untuk melakukan klasterisasi pada data curas dan data curanmor yang ada agar data tersebut lebih mudah untuk bisa dipahami, karena sudah terbagi atas beberapa klaster. Sesuai dengan poin 4.3.7 Hasil Akhir *K-Means Clustering*, yang menyatakan bahwa hasil klasterisasi terdiri dari 3 klaster, yaitu rawan, sedang, dan aman.

Dalam melakukan pengembangan sistem, penulis berpedoman dengan flowchart di gambar \*CEK dan *use case diagram* di gambar \*CEK, sehingga sistem yang dikembangkan sesuai dengan perencanaan. Hasil dari pengembangan sistem ini, penulis jelaskan dalam beberapa bagian utama sebagai berikut.

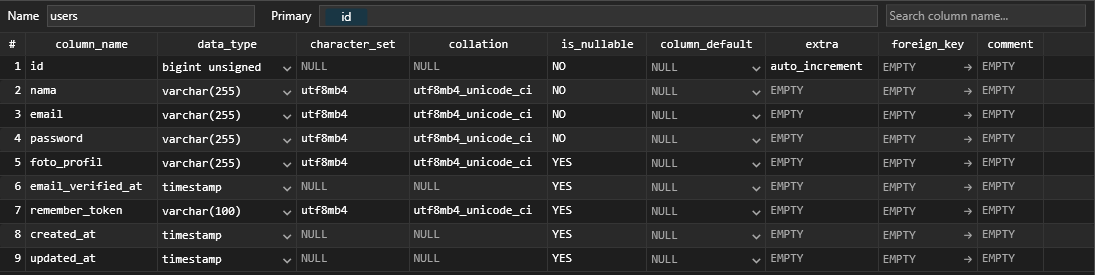
### Database Sistem

Pembuatan *database* dilakukan berdasarkan analisis yang telah dirancang pada tahap sebelumnya. Penulis merujuk pada *flowchart* dan diagram *use case* yang telah disusun pada Gambar \*CEK dan Gambar \*CEK sebagai acuan utama dalam merancang struktur *database*. *Flowchart* menggambarkan alur proses sistem secara keseluruhan, sedangkan *use case* menunjukkan interaksi antara aktor dan sistem. Informasi dari kedua diagram tersebut membantu penulis dalam mengidentifikasi kebutuhan data serta relasi antar entitas yang dibutuhkan dalam sistem. Setiap entitas dan atribut yang terdapat pada *database* dirancang agar sesuai dengan kebutuhan fungsional yang telah diuraikan pada *flowchart* dan *use case*. Selain itu, penulis juga memperhatikan normalisasi agar struktur *database* efisien dan terhindar dari redundansi. Hasil dari perancangan ini kemudian dibuat menjadi suatu *database* yang memiliki relasi antar tabelnya. *Database* sistem yang dibuat dapat dilihat pada Gambar \*CEK

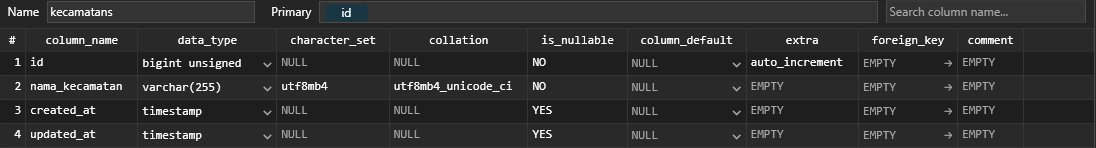


Pada Gambar \*CEK yang merupakan struktur database yang digunakan pada Web SIG. Pada database tersebut terdiri dari tujuh tabel dengan 3 tabel diantaranya merupakan tabel master. Tiga tabel yang dimaksud dalam tabel master yaitu, tabel users yang menyimpan data pengguna, tabel klasters yang menyimpan data klaster, dan tabel kecamatans, yang menyimpan data kecamatan. Pada tabel tersebut juga menggunakan teori normalisasi 2nf, yang terbukti dengan adanya tabel detail\_curas dan detail\_curanmors yang menyimpan tambahan data kasus curas dan curanmor pada setiap tanggalnya. Penjelasan lebih rinci terkait masing-masing tabel pada database akan diuraikan sebagai berikut.

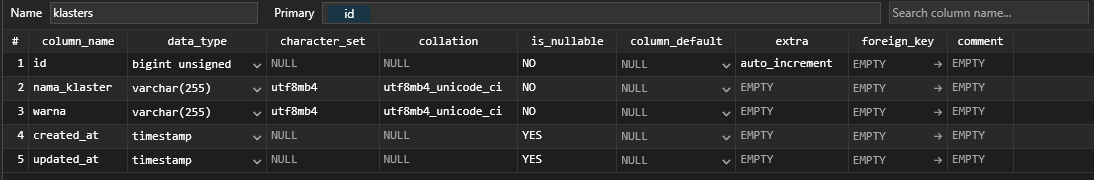
1. Tabel Users



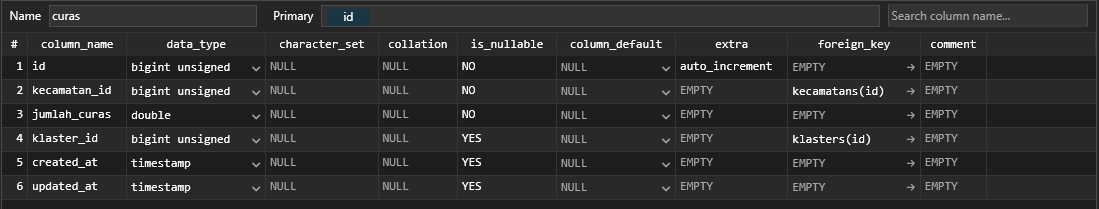
1. Tabel Kecamatans



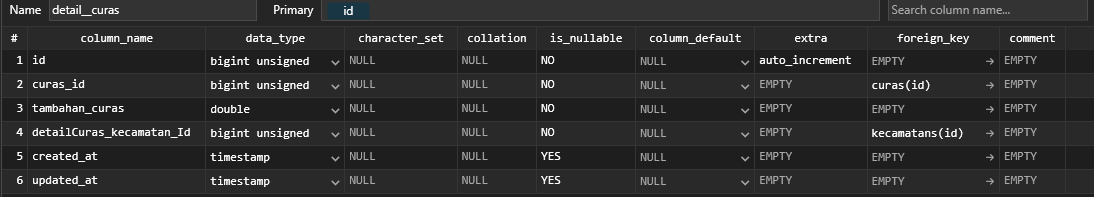
1. Tabel Klasters



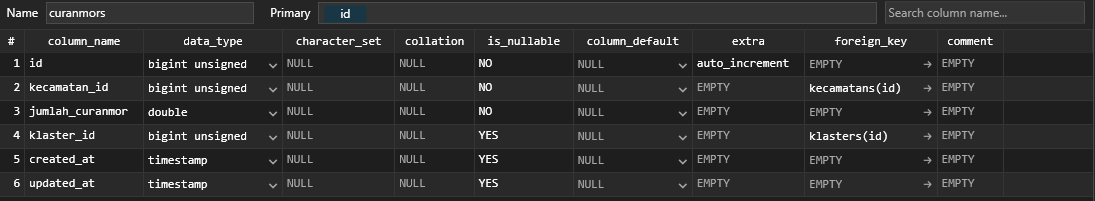
1. Tabel Curas



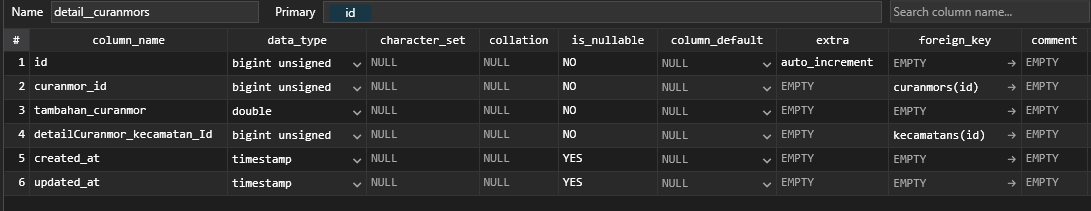
1. Tabel Detail\_Curas



1. Tabel Curanmors



1. Tabel Detail\_Curanmors



### Penerapan Algoritma K-Means

### Fitur Fitur Pada Sistem

## Pengujian

## Analisis dan Pembahasan

## 

# DAFTAR PUSTAKA

Alifah, R. F. N., & Fauzan, A. C. (2023). Implementasi Algoritma K-Means Clustering Berbasis Jarak Manhattan untuk Klasterisasi Konsentrasi Bidang Mahasiswa. *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, *5*(1), 31–41. https://doi.org/10.28926/ilkomnika.v5i1.542

Andrea Santana Adzani. (2022). Klastering Pemetaan Daerah Rawan Kecelakaan Menggunakan Metode K-Means Berbasis Sistem Informasi Geografis (Studi Kasus Di Kabupaten Jember). *Science*, *7*(1), 1–8.

Anggraini, N. S. (2021). *Pemetaan Daerah Rawan Pencurian Dengan Kekerasan (Begal) Di Kabupaten Lumajang Skripsi*. https://sipora.polije.ac.id/5567/

Apriliana, & Haris R, D. (2022). Pemetaan Daerah Rawan Kriminalitas pada Wilayah Hukum Polres Cirebon Kota Tahun 2018-2021. *Seminar Nasional Dan Diseminasi Tugas Akhir*, 2022.

BPS. (2023). Statistik Kriminal. *Badan Pusat Statistik*, *021*, 1–62. https://www.bps.go.id/id/publication/2023/12/12/5edba2b0fe5429a0f232c736/statistik-kriminal-2023.html

BPS Kab Probolinggo. (2024). *Kabupaten Probolinggo Dalam Angka 2024*.

Preddy, …, Marpaung, P., Pebrian, I., & Putri, W. (2023). Penerapan Data Mining Untuk Pengelompokan Kepadatan Penduduk Kabupaten Deli Serdang Menggunakan Algoritma K-Means. *Jurnal Ilmu Komputer Dan Sistem Informasi (JIKOMSI)*, *6*(2), 64–70.

Rahayu, R. (2022). Jurnal Informatika dan Teknologi Informasi Jurnal Informatika dan Teknologi Informasi. *Jurnal Informatika Dan Teknologi Informasi*, *1*(2), 98–103. https://doi.org/10.56854/jt.v1i2.80

Rohman, F. F. (2023). Sistem Informasi Geografis Tingkat Kriminalitas Kota Jember Menggunakan Metode K-Means. In *Politeknik Negeri Jember*.

Rumariana, A., & Arifin, M. (2022). Kepuasan Pengguna Aplikasi Geographic Information System (GIS) Stunting. *Prosiding University Research Colloqium*, 28–36. http://stunting.sipandawa.com

Ruziq Nawaf Zulfahmi, Maria Kristiana Daul, Muhammad Al Ayyubi, I Wayan Julianta Pradnyana, & Rokhana Dwi Bekti. (2023). Pemetaan Kerentanan Tingkat Kriminalitas Menggunakan Metode Self Organizing Map. *INSOLOGI: Jurnal Sains Dan Teknologi*, *2*(5), 872–881. https://doi.org/10.55123/insologi.v2i5.2566

Umar, T. L. (2021). Perancangan Sistem Informasi Geografi Tempat Bersalin Berbasis Mobile. *Jurnal Informatika Dan Rekayasa Perangkat Lunak (JATIKA)*, *2*(2), 221–229. http://jim.teknokrat.ac.id/index.php/informatika

# LAMPIRAN

|  |  |  |
| --- | --- | --- |
| **Kecamatan** | **Kasus Kriminal 2021** | **Kasus Kriminal 2022** |
|
| Bantaran | 7 | 11 |
| Banyuanyar | 10 | 0 |
| Besuk | 7 | 65 |
| Dringu | 28 | 0 |
| Gading | 12 | 0 |
| Gending | 8 | 28 |
| Kotaanyar | 9 | 7 |
| Kraksaan | 117 | 166 |
| Krenjengan | 8 | 11 |
| Krucil | 9 | 0 |
| Kuripan | 2 | 0 |
| Leces | 15 | 0 |
| Lumbang | 5 | 0 |
| Maron | 7 | 66 |
| Paiton | 15 | 21 |
| Pakuniran | 4 | 22 |
| Pajarakan | 7 | 11 |
| Sukapura | 4 | 0 |
| Sumber | 5 | 0 |
| Sumberasih | 26 | 0 |
| Tegalsiwalan | 5 | 22 |
| Tiris | 39 | 7 |
| Tongas | 39 | 5 |
| Wonomerto | 11 | 0 |
| **TOTAL** | **399** | **442** |

LAMPIRAN 1 Kenaikan Kasus Kriminal di Probolinggo

|  |  |  |
| --- | --- | --- |
| **Kecamatan** | **2022** | |
| **Curas** | **Curanmor** |
| Bantaran | 0 | 4 |
| Banyuanyar | 0 | 0 |
| Besuk | 0 | 18 |
| Dringu | 0 | 0 |
| Gading | 0 | 0 |
| Gending | 0 | 8 |
| Kotaanyar | 0 | 1 |
| Kraksaan | 1 | 99 |
| Krenjengan | 0 | 2 |
| Krucil | 0 | 0 |
| Kuripan | 0 | 0 |
| Leces | 0 | 0 |
| Lumbang | 0 | 0 |
| Maron | 1 | 21 |
| Paiton | 1 | 9 |
| Pakuniran | 0 | 5 |
| Pajarakan | 0 | 3 |
| Sukapura | 0 | 0 |
| Sumber | 0 | 0 |
| Sumberasih | 0 | 0 |
| Tegalsiwalan | 2 | 0 |
| Tiris | 0 | 2 |
| Tongas | 0 | 1 |
| Wonomerto | 0 | 0 |
| **TOTAL** | **5** | **173** |
| **Grand TOTAL** | **178** | |

LAMPIRAN 2 Data Kasus Curas dan Curanmor Tahun 2022