

**PEMILIHAN *CLUSTER* OPTIMAL UNTUK
PENGELOMPOKKAN PROVINSI RAWAN KRIMINALITAS
DI INDONESIA BERDASARKAN 5 METODE EVALUASI
MENGUNAKAN *K-MEANS CLUSTERING***

SKRIPSI

**Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Matematika**

Oleh :

RISKA TRI ANGGARA

08011282126049



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS SRIWIJAYA**

2025

LEMBAR PENGESAHAN

**PEMILIHAN CLUSTER OPTIMAL UNTUK PENGELOMPOKKAN
PROVINSI RAWAN KRIMINALITAS DI INDONESIA BERDASARKAN 5
METODE EVALUASI MENGGUNAKAN K-MEANS CLUSTERING**

SKRIPSI


**Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Matematika**

Oleh

**RISKA TRI ANGGARA
NIM. 08011282126049**

Indralaya, 25 Februari 2025

Pembimbing Kedua



**Dr. Yuli Andriani, S.Si., M.Si
NIP. 197207021999032001**

Pembimbing Utama



**Dr. Ir. Herlina Hanum, M.Si
NIP. 196501081990032007**

**Mengetahui,
Ketua Jurusan Matematika**

**Dr. Dian Cahyawati Sukanda, S.Si., M.Si
NIP. 197303212000122001**



PERNYATAAN KEASLIAN KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama Mahasiswa : Riska Tri Anggara

NIM : 08011282126049

Fakultas/Jurusan : Matematika dan Ilmu Pengetahuan Alam/Matematika

Menyatakan bahwa skripsi ini adalah hasil karya ilmiah saya sendiri dan karya ilmiah ini belum pernah diajukan sebagai pemenuhan persyaratan untuk memperoleh gelar kesarjanaan strata satu (S1) dari Universitas Sriwijaya maupun perguruan tinggi lain.

Semua informasi yang dimuat didalam skripsi ini yang berasal dari penulis lain baik yang dipublikasi atau tidak telah diberikan penghargaan dengan mengutip nama sumber penulis baik yang secara benar. Semua isi dari skripsi ini sepenuhnya menjadi tanggung jawab saya sebagai penulis.

Demikianlah surat pernyataan ini saya buat dengan sebenarnya.

Indralaya, 10 Maret 2025

Penulis



Riska Tri Anggara

NIM. 08011282126049

HALAMAN PERSEMBAHAN

Motto

“I believed in myself, I tried, I was strong, I could, and I got it.”

Skripsi ini kupersembahkan kepada:

- 1. Allah SWT**
- 2. Kedua Orang Tuaku**
- 3. Keluarga Besarku**
- 4. Semua Dosen dan Guruku**
- 5. Saudaraku**
- 6. Sahabatku**
- 7. Almamaterku**

KATA PENGANTAR

Segala puji dan syukur saya panjatkan atas kehadiran Allah Subhanahu Wa Ta'ala karena berkat rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul “Pemilihan *Cluster* Optimal untuk Pengelompokkan Provinsi Rawan Kriminalitas di Indonesia Berdasarkan 5 Metode Evaluasi Menggunakan *K-Means Clustering*” dengan baik. Shalawat serta salam senantiasa tercurahkan kepada Nabi Muhammad SAW beserta keluarga dan para sahabat serta para pengikutnya hingga akhir zaman.

Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Matematika di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya. Penulis sepenuhnya menyadari bahwa dalam proses pembuatan skripsi ini bukanlah akhir dari proses pembelajaran karena menimba ilmu dapat dilakukan seumur hidup, kapanpun dan dimanapun kita tidak akan terlepas dari proses pembelajaran. Skripsi ini dapat terselesaikan pastinya membutuhkan bantuan dari berbagai pihak, baik secara langsung maupun tidak langsung. Penulis mengucapkan terima kasih kepada kedua orang tua tercinta. Ayahanda **Herman Aswandi** terima kasih telah bertahan dan berjuang hingga saat ini, beliau selalu memberikan semangat dan motivasi kepada anaknya untuk selalu menjunjung tinggi pendidikan. Terima kasih untuk Almarhuma Ibu **Herminopti** yang selalu menjadi alasan penulis dalam melakukan sesuatu serta memberikan kasih sayang, dukungan, serta doa dan nasihat selama hidupnya, walaupun ibu tidak lagi bersama penulis secara fisik, saya percaya doa dan cinta

ibu pasti menyertai setiap langkah penulis. Semoga ibu bangga dengan apa yang telah penulis capai dan semoga Allah SWT senantiasa melimpahkan rahmat dan kedamaian kepada ibu di sisi-Nya. Segala hormat dan kerendahan hati penulis mengucapkan terima kasih sekaligus penghargaan kepada:

1. Bapak **Prof. Hermansyah, S.Si., M.Si., Ph.D** selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
2. Ibu **Dr. Dian Cahyawati Sukanda, S.Si., M.Si** selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya sekaligus Dosen Pembimbing Akademik yang telah memberikan bimbingan serta arahan yang sangat berguna selama proses perkuliahan.
3. Ibu **Des Alwine Zayanti, S.Si., M.Si** selaku Sekretaris Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya yang telah memberikan masukan dan arahan selama proses perkuliahan ini.
4. Ibu **Dr. Ir. Herlina Hanum, M.Si** selaku Dosen Pembimbing Utama yang bersedia menerima penulis menjadi salah satu anak bimbingan dan bersedia selalu memberikan nasihat, bimbingan, saran, pengalaman, serta meluangkan waktu di tengah kesibukannya dalam pengerjaan skripsi ini.
5. Ibu **Dr. Yuli Andriani, S.Si., M.Si** selaku Dosen Pembimbing Pembantu yang bersedia memberikan nasihat, saran, serta meluangkan waktu di tengah kesibukannya dalam pengerjaan skripsi ini.
6. Bapak **Drs. Ali Amran, M.T** sebagai ketua seminar dan Ibu **Sri Indra Maiyanti, S.Si., M.Si** sebagai sekretaris seminar yang telah membantu

pelaksanaan seminar dan sidang sarjana penulis serta memberikan saran dan masukan kepada penulis dalam pengerjaan skripsi ini.

7. **Seluruh Dosen** Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya yang telah memberikan ilmu serta pengalaman selama proses penulis.
8. Bapak **Irwansyah** dan Ibu **Hamidah** selaku Pegawai Tata Usaha Jurusan Matematika yang membantu dalam proses administrasi selama masa perkuliahan.
9. Kakak saya **Verly Angelina**, bibi saya **Willy Astuti**, dan sepupu saya **Ica, Bela, dila** terima kasih karena selalu memberikan semangat, doa dan dukungan serta selalu menghibur penulis ketika penulis mengalami kesusahan.
10. Teman seperjuangan di masa kuliah **Ayu, Andini, Adis, Fajriah, Febri, Fidella, Mutik, Nabila, Rindu, Sapiah, Vira**, serta seluruh **Teman-Teman Angkatan 2021** yang selalu memberikan dukungan dan bantuan selama masa perkuliahan ini.
11. Semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah memberian banyak doa serta dukungan dalam menyelesaikan skripsi ini.

Semoga kebaikan semua pihak yang membantu dibalas oleh Allah SWT dan semoga skripsi ini bermanfaat bagi Mahasiswa/Mahasiswi Jurusan Matematika Fakultas Matematika Universitas Sriwijaya dan semua pihak.

Indralaya, 25 Februari 2025

Penulis

OPTIMAL CLUSTER SELECTION FOR CLUSTERING CRIME-PRONE PROVINCES IN INDONESIA BASED ON 5 EVALUATION METHODS USING K-MEANS CLUSTERING

By:
Riska Tri Anggara
08011282126049

ABSTRACT

Cluster analysis is one method that is usually used to group objects into groups that are similar to each other based on their characteristics. This study aims to cluster 34 provinces in Indonesia based on the number of crime cases in 2022 using the K-Means clustering method. Provinces are grouped based on crime cases, which include 8 variables, namely murder, domestic violence, rape, kidnapping, theft, narcotics, fraud, and public order violations. Determination of the optimal number of clusters with $k=2, 3, 4,$ and 5 clusters was done by evaluation using five methods, namely Calinski Harabasz Index, Davies Bouldin Index, Elbow, Silhouette Coefficient, and Standard Deviation ratio. The results showed that the optimal number of clusters using the Calinski Harabasz Index, elbow, and standard deviation ratio methods was 4 clusters, while the Davies Bouldin Index and Silhouette Coefficient method was 2 clusters. Selection of the best clusters from 5 evaluation methods based on the results that have the most similarity because it is concluded that the results obtained are more representative and more accurate so that the best cluster selected is 4 clusters. Clustering in cluster 1 consists of 2 provinces consisting of North Sumatra and East Java Provinces with the variables of rape and fraud cases at a high level, while the other 6 variables are at a very high level. Cluster 2 consists of 3 provinces consisting of DKI Jakarta, West Java and South Sulawesi Provinces with the variables of rape, kidnapping and fraud cases at a very high level, while the other 5 variables are at a high level. Cluster 3 consists of 6 provinces consisting of Aceh, Riau, South Sumatra, Lampung, Central Java, and North Sulawesi with all variables at a high level. The remaining provinces are in cluster 4, which consists of 23 provinces with all variables at a medium level.

Keywords: K-Means Clustering, Crime, Cluster Evaluation

PEMILIHAN *CLUSTER* OPTIMAL UNTUK PENGELOMPOKKAN PROVINSI RAWAN KRIMINALITAS DI INDONESIA BERDASARKAN 5 METODE EVALUASI MENGGUNAKAN *K-MEANS CLUSTERING*

Oleh:
Riska Tri Anggara
08011282126049

ABSTRAK

Analisis *cluster* salah satu metode yang biasanya digunakan untuk melakukan pengelompokan objek menjadi beberapa kelompok yang memiliki kemiripan satu sama lain berdasarkan karakteristiknya. Penelitian ini bertujuan untuk mengelompokkan 34 provinsi di Indonesia berdasarkan jumlah kasus kriminalitas tahun 2022 menggunakan metode *K-Means clustering*. Provinsi-provinsi dikelompokkan berdasarkan kasus kriminalitas, yang meliputi 8 variabel yaitu pembunuhan, kekerasan dalam rumah tangga (KDRT), perkosaan, penculikan, pencurian, narkoba, penipuan, dan pelanggaran ketertiban umum. Penentuan jumlah *cluster* optimal dengan $k=2, 3, 4,$ dan 5 *cluster* dilakukan dengan evaluasi menggunakan lima metode, yaitu Indeks *Calinski Harabasz*, Indeks *Davies Bouldin*, *Elbow*, *Silhouette Coefficient*, dan rasio Simpangan Baku. Hasil penelitian menunjukkan bahwa jumlah *cluster* optimal menggunakan metode Indeks *Calinski Harabasz*, *elbow*, dan rasio simpangan baku ialah 4 *cluster*, sedangkan metode Indeks *Davies Bouldin* dan *Silhouette Coefficient* ialah 2 *cluster*. Pemilihan *cluster* terbaik dari 5 metode evaluasi berdasarkan hasil yang mempunyai kesamaan terbanyak karena disimpulkan hasil yang didapat lebih representatif dan lebih akurat sehingga *cluster* terbaik yang dipilih yaitu 4 *cluster*. Pengelompokan pada *cluster* 1 beranggotakan 2 provinsi yang terdiri dari Provinsi Sumatera Utara dan Jawa Timur dengan variabel kasus perkosaan dan penipuan berada ditingkat tinggi, sedangkan 6 variabel lainnya berada di tingkat sangat tinggi. *Cluster* 2 beranggotakan 3 provinsi yang terdiri dari Provinsi DKI Jakarta, Jawa Barat, dan Sulawesi Selatan dengan variabel kasus perkosaan, penculikan, dan penipuan berada di tingkat sangat tinggi, sedangkan 5 variabel lainnya berada di tingkat tinggi. *Cluster* 3 beranggotakan 6 provinsi yang terdiri dari Provinsi Aceh, Riau, Sumatera Selatan, Lampung, Jawa Tengah, dan Sulawesi Utara dengan semua variabel berada di tingkat tinggi. Sisa provinsi lainnya berada pada *cluster* 4 yang berjumlah 23 Provinsi dengan semua variabel berada di tingkat sedang.

Kata kunci: *K-Means Clustering*, Kriminalitas, Evaluasi *cluster*

DAFTAR ISI

LEMBAR PENGESAHAN	ii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	v
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	3
1.3 Pembatasan Masalah	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Kriminalitas	5
2.2 Standarisasi Data	5
2.3 Analisis <i>Cluster</i>	6
2.4 Metode <i>Hierarki</i>	7
2.5 Metode <i>Non Hierarki</i>	7
2.5.1 <i>K-Means</i>	8
2.6 Indeks <i>Calinski Harabasz</i>	9
2.7 Indeks <i>Davies Bouldin</i>	11
2.8 Metode <i>Elbow</i>	13

2.9 <i>Silhouette Coefficient</i>	14
2.10 Simpangan Baku.....	16
BAB III METODOLOGI PENELITIAN	18
3.1 Tempat.....	18
3.2 Waktu	18
3.3 Alat	18
3.4 Metodologi Penelitian	18
3.4.1 Sumber Data	18
3.4.2 Objek dan Variabel Penelitian	18
3.4.3 Metode Analisis	20
BAB IV HASIL DAN PEMBAHASAN	22
4.1 Deskripsi Data	22
4.2 Standarisasi Data	25
4.3 Analisis <i>Cluster K-Means</i>	27
4.3.1 Pengelompokan dalam 2 <i>cluster</i>	27
4.3.2 Pengelompokan dalam 4 <i>cluster</i>	32
4.4 Indeks <i>Calinski Harabasz</i>	37
4.5 Indeks <i>Davies Bouldin</i>	41
4.6 Metode <i>Elbow</i>	44
4.7 Metode <i>Silhouette Coefficient</i>	45
4.8 Simpangan Baku.....	49
4.9 Pembahasan hasil	51
4.9.1 Karakteristik <i>Cluster</i> Terbaik	51
BAB V KESIMPULAN DAN SARAN	55
5.1 Kesimpulan.....	55

5.2 Saran.....	56
DAFTAR PUSTAKA	57
LAMPIRAN.....	59

DAFTAR TABEL

Tabel 2.1 Nilai Silhouette Coefficient	15
Tabel 3.1 Objek Penelitian	18
Tabel 3.2 Variabel Kasus Kriminalitas dalam Penelitian	19
Tabel 4.1 Data Kriminalitas Indonesia Tahun 2022	22
Tabel 4.2 Deskriptif Data Kriminalitas Provinsi di Indonesia Tahun 2022	24
Tabel 4. 3 Kategorisasi Data.....	24
Tabel 4. 4 Standarisasi Data	27
Tabel 4.5 Penentuan Centroid Awal.....	27
Tabel 4.6 Jarak semua objek pada iterasi 1	28
Tabel 4.7 Hasil Pengelompokan Objek Iterasi 1	29
Tabel 4. 8 Nilai Centroid Baru	30
Tabel 4.9 Hasil Pengelompokan Objek Iterasi 2	30
Tabel 4.10 Nilai Centroid Baru	31
Tabel 4. 11 Hasil Pengelompokan Objek Iterasi 3	31
Tabel 4. 12 Penentuan Centroid Awal.....	32
Tabel 4.13 Jarak semua objek pada iterasi 1	33
Tabel 4.14 Hasil Pengelompokan Objek Iterasi 1	34
Tabel 4.15 Nilai Centroid Baru	35
Tabel 4.16 Hasil Pengelompokan Objek Iterasi 2	35
Tabel 4.17 Nilai Centroid Baru	36
Tabel 4.18 Hasil Pengelompokan Objek Iterasi 4	36
Tabel 4. 19 Nilai Indeks Calinski Harabasz	40
Tabel 4.20 Nilai Sum of Square between Cluster.....	42
Tabel 4.21 Nilai Indeks Davies Bouldin	43
Tabel 4.22 Nilai Metode Elbow	45
Tabel 4.23 Nilai Silhouette Coefficient	48
Tabel 4.24 Nilai Simpangan Baku.....	51
Tabel 4.25 Penentuan Cluster Terbaik.....	51
Tabel 4.26 Hasil Pengelompokan Daerah Rawan Kriminalitas pada Tahun 2022.....	53
Tabel 4.27 Rata-Rata Data Setiap Variabel.....	54
Tabel 4.28 Karakteristik Data.....	54

DAFTAR GAMBAR

Gambar 4.1 Grafik Data Kriminalitas	23
Gambar 4.2 Plot <i>K-Means Clustering</i>	37
Gambar 4.3 Grafik Indeks <i>Calinski Harabasz</i>	40
Gambar 4.4 Grafik Indeks <i>Davies Bouldin</i>	43
Gambar 4.5 Grafik <i>Elbow</i>	45
Gambar 4. 6 Grafik <i>Silhouette Coefficient</i>	48

DAFTAR LAMPIRAN

Lampiran 1 Objek Penelitian	59
Lampiran 2 Data Kriminalitas Indonesia Tahun 2022	60
Lampiran 3 Standarisasi Data.....	61
Lampiran 4 Jarak semua objek pada iterasi 1 2 Cluster	62
Lampiran 5 Hasil Pengelompokan Objek Iterasi 1	63
Lampiran 6 Hasil Pengelompokan Objek Iterasi 2.....	64
Lampiran 7 Hasil Pengelompokan Objek Iterasi 3.....	65
Lampiran 8 Hasil Pengelompokan K-Means 2 Cluster menggunakan R-Studio	66
Lampiran 9 Hasil Pengelompokan K-Means 3 Cluster menggunakan R-Studio	67
Lampiran 10 Hasil Pengelompokan K-Means 4 Cluster menggunakan R-Studio	68
Lampiran 11 Hasil Pengelompokan K-Means 5 Cluster menggunakan R-Studio	69

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara kepulauan terbesar di dunia dengan lebih dari 17.001 pulau dan populasi sebanyak 283,5 juta jiwa. Posisi geografis di Indonesia yang strategis di jalur pelayaran serta kekayaan sumber daya alam membuatnya rentan terhadap berbagai ancaman keamanan (Samy & Kusumadewi, 2021). Populasi yang besar dan wilayah yang luas, Indonesia menghadapi beragam jenis kriminalitas di berbagai daerah, mulai dari kota-kota besar hingga pedesaan.

Kriminalitas ialah semua bentuk perilaku, tindakan, maupun lisan yang dapat membahayakan masyarakat, melanggar undang-undang serta merugikan masyarakat secara politis, psikologis, dan ekonomis (Pradana, 2019). Tindak kriminalitas memiliki dampak besar dalam kehidupan masyarakat dan terdapat banyak faktor yang mendorong seseorang untuk melakukan tindak kriminal. Saat ini, kriminalitas telah menjadi sesuatu yang dianggap lazim di masyarakat, di mana pelaku seringkali tidak berpikir panjang sebelum bertindak, sehingga tidak ragu untuk melukai atau bahkan menghilangkan nyawa korban. Tingkat kriminalitas di Indonesia mencakup berbagai jenis, mulai dari pelanggaran ringan seperti melanggar ketertiban umum dan pencurian, hingga kejahatan berat seperti pembunuhan, narkoba, dan pemerkosaan. Menurut Badan Pusat Statistik (BPS) Indonesia tahun 2023 mencatat sepanjang tahun 2020 sampai 2022 kasus kriminalitas meningkat pesat pada akhir-akhir ini. Berbagai kriminalitas yang sudah terjadi tersebar luas di Indonesia, sehingga sulit untuk masyarakat

menentukan tempat kejadian yang rawan kriminalitas dan lokasi yang cukup aman. Oleh karena itu, informasi terkait daerah yang rawan akan kriminalitas dapat membantu masyarakat agar lebih waspada ketika melakukan kegiatan di suatu daerah tersebut. Informasi daerah yang rawan akan kriminalitas juga penting bagi kepolisian Indonesia, serta kepolisian daerah guna mengantisipasi terjadinya kriminalitas (Hilman et al., 2015).

Salah satu solusi guna memenuhi informasi kriminalitas yang tepat, dapat dilakukan dengan cara mengelompokkan provinsi di Indonesia yang bertujuan untuk mengamati daerah-daerah yang rawan terhadap kriminalitas berdasarkan data kriminalitas pada tahun 2022. Pengelompokan provinsi dapat dilakukan menggunakan analisis *cluster*. *Clustering* merupakan salah satu metode yang dapat mengelompokkan data ke beberapa *cluster* dengan tingkat kemiripan datanya tinggi dalam satu kelas dan cukup efisien dalam mengelompokkan data dengan jumlah yang besar (Maori & Evanita, 2023).

Terdapat beberapa algoritma yang dapat digunakan dalam metode *cluster* salah satunya yang paling sering digunakan adalah *K-Means*. *K-Means* merupakan metode yang digunakan untuk menentukan cara yang efisien dalam membagi entitas ke- n menjadi suatu kelompok (Saputra & Nataliani, 2021). Metode *K-Means* dalam jika dibandingkan dengan metode lain memiliki kelebihan yaitu mampu mengelompokkan data secara otomatis berdasarkan tingkat kemiripan antar data dalam kluster yang dihasilkan (Mirantika, 2021). *K-Means* sangat banyak digunakan untuk melakukan penelitian di semua bidang salah satunya produksi, *K-Means* digunakan untuk mengelompokkan produk berdasarkan

data persediaan barang (Syahfitri et al., 2023). Bidang sosial *K-Means* digunakan untuk pengelompokan penerima bantuan sosial masyarakat berdasarkan pekerjaan, kepemilikan asset dan penghasilan (Putra & Anggrawan, 2021).

Berdasarkan penelitian Ananda (2024), pengelompokan kabupaten dan kota di Sumatera Selatan berdasarkan tingkat produksi tanaman padi dan palawija tahun 2019-2021 menggunakan metode *K-means Clustering*. Penelitian tersebut menggunakan dua metode sebagai evaluasi *clustering* dalam mencari jumlah *k* terbaik yaitu *Silhouette Coefficient* dan Simpangan Baku. Oleh karena itu, penelitian ini menggunakan metode *K-means clustering* dalam mengelompokan provinsi di Indonesia berdasarkan jumlah kasus kriminalitas pada tahun 2022 menggunakan 2 metode yang sama seperti penelitian sebelumnya dan ditambah 3 metode baru yaitu metode Indeks *Calinski Harabasz*, Indeks *Davies Bouldin*, dan *Elbow* sebagai evaluasi *clustering* dalam menentukan jumlah *k* terbaik, serta meningkatkan validitas hasil agar memperoleh hasil yang komprehensif, konsistensi, dan stabilitas.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan diatas, rumusan masalah dalam penelitian ini adalah sebagai berikut :

1. Bagaimana hasil *cluster* terbaik dalam pengelompokkan Provinsi di Indonesia menggunakan analisis Indeks *Calinski Harabasz*, Indeks *Davies Bouldin*, *Elbow*, *Silhouette Coefficient*, dan Simpangan Baku?
2. Bagaimana hasil *clustering* provinsi di Indonesia menggunakan metode *K-Means* berdasarkan *cluster* terbaik?

1.3 Pembatasan Masalah

Penelitian ini menggunakan data sekunder berupa data kriminalitas tahun 2022 yang diperoleh dari Badan Pusat Statistik Indonesia (BPS).

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah pada penelitian ini yang telah dijelaskan diatas, tujuan dari penelitian ini adalah:

1. Memperoleh jumlah *cluster* terbaik berdasarkan analisis Indeks *Calinski Harabasz*, Indeks *Davies Bouldin*, *Elbow*, *Silhouette Coefficient*, dan Simpangan Baku.
2. Mengetahui daerah rawan kriminalitas di Indonesia pada tahun 2022 menggunakan metode *K-Means* berdasarkan *cluster* terbaik.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Bagi peneliti, sebagai sarana pembelajaran dan pemahaman dalam mengetahui sistematika dan menganalisa data menggunakan metode-metode salah satunya *K-Means Clustering*.
2. Bagi pembaca, diharapkan dapat dijadikan referensi guna melakukan penelitian lebih lanjut dengan topik yang sama dan menambah informasi terkait daerah yang rawan akan kriminalitas di Indonesia.
3. Bagi pemerintah, setelah mendapatkan hasil pengelompokan provinsi di Indonesia diharapkan dapat dijadikan sebuah informasi tambahan terkait kriminalitas dan lebih memperketat keamanan di daerah yang berpotensi terjadinya kriminalitas.

DAFTAR PUSTAKA

- Ananda, N. P. (2024). *K-Means untuk menganalisis perubahan pengelompokan kabupaten dan kota di Sumatera Selatan berdasarkan tingkat produksi tanaman padi dan palawija tahun 2019-2021*. Skripsi Jurusan Matematika FMIPA Universitas Sriwijaya.
- Badan Pusat Statistik (BPS) Indonesia. (2023). *Statistik Kriminal 2023*. Badan Pusat Statistik Indonesia.
- Badan Pusat Statistik Indonesia. (2022). *BPS Indonesia*. bit.ly/BPS-Indonesia-Statistik-Kriminal
- Firmansyah, T., Poningsih, & Andani, S. R. (2022). Analisis Clustering Algoritma K-Means sebagai rekomendasi penambahan koleksi duku di perpustakaan madrasah tsanawiyah negeri 2 Simalungun. *ZAHRA: Buletin Big Data ...*, 1(1), 44–48.
- Hilman, G. Y., Sasmito, B., & Wijaya, A. P. (2015). Pemetaan daerah rawan kriminalitas di wilayah hukum poltabes Semarang tahun 2013 dengan menggunakan metode Clustering. *Jurnal Geodesi Undip*, 4(1), 32–42.
- Kaufman, L., & Rousseeuw, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. New Jersey: John Wiley and Sons Inc.
- Khomarudin, A. N. (2016). Teknik Data Mining: Algoritma K-Means Clustering. *Jurnal Komunitas ELearning IlmuKomputer*, 4(1), 7–8.
- Maori, N. A., & Evanita, E. (2023). Metode Elbow dalam optimasi jumlah Cluster pada K-Means Clustering. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 14(2), 277–288.
- Mirantika, N. (2021). Penerapan algoritma K-Means Clustering untuk pengelompokan penyebaran covid-19 di Provinsi Jawa Barat. *Nuansa Informatika*, 15(2), 92–98. <https://doi.org/10.25134/nuansa.v15i2.4321>
- Mubarak, H. (2023). *Perbandingan Algoritma K-Means dan K-Medoid dalam pengelompokan Kabupaten/Kota di Provinsi Jambi berdasarkan data penyakit menular dengan Calinski Harabasz Index*. Skripsi Program Studi Matematika FST Universitas Jambi.
- Muliono, R., & Sembiring, Z. (2019). Data mining Clustering menggunakan algoritma K-Means untuk klusterisasi tingkat tridarma pengajaran dosen. *Journal of Computer Engineering System and Science*, 4(2), 2502–2714.
- Nurkholik, D. (2022). *Analisis K-Medoids Clustering Metode Elbow pada kasus covid-19 di Provinsi DKI Jakarta*. Universitas Islam Negeri Maulanan Malik Ibrahim.

- Pradana, A. (2019). *Pemetaan daerah kerawanan kriminalitas menggunakan K-Means Clustering di wilayah hukum Polresta Sidoarjo*. Skripsi Program Studi Matematika FST Universitas Islam Negeri Sunan Ampel.
- Putra, L. Ga. R., & Anggrawan, A. (2021). Pengelompokan penerima bantuan sosial masyarakat dengan Metode K-Means. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 21(1), 205–214.
- Samy, M., & Kusumadewi, J. A. (2021). Diplomasi pertahanan maritim Indonesia dalam menghadapi ancaman keamanan non-tradisional upaya mewujudkan visi poros maritim dunia. *Jurnal Hubungan Internasional*, 14(1), 45.
- Saputra, E. A., & Nataliani, Y. (2021). Analisis pengelompokan data nilai siswa untuk menentukan siswa berprestasi menggunakan Metode Clustering K-Means. *Journal of Information Systems and Informatics*, 3(3), 424–439.
- Syahfitri, N., Budianita, E., Nazir, A., & Afrianty, I. (2023). Pengelompokan produk berdasarkan data persediaan barang menggunakan Metode Elbow dan K-Medoid. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 4(3), 1668–1675.
- Umagapi, I. T., Umaternate, B., Hazriani, & Yuyun. (2023). Uji kinerja K-Means Clustering menggunakan Davies Bouldin Index pada pengelompokan data prestasi siswa. *Seminar Nasional Sistem Informasi Dan Teknologi (SISFOTEK)*, 7(1), 303–308.
- Wahyudi, D. (2013). Perlindungan hukum terhadap korban kejahatan cyber crime di Indonesia. *Jurnal Ilmu Hukum Jambi*, 4(1), 98–113.
- Wang, X., & Xu, Y. (2019). An improved index for Clustering validation based on Silhouette Index and Calinski-Harabasz Index. *IOP Conference Series: Materials Science and Engineering*, 569(5).