



***K-MEANS CLUSTERING* PADA PEMETAAN DAERAH
RAWAN KRIMINALITAS BERDASARKAN JENIS
KRIMINALITAS DI INDONESIA**

SKRIPSI

**Untuk memenuhi persyaratan
dalam menyelesaikan program sarjana Strata – 1 Matematika**

Oleh :

**USWATUN HASANAH
NIM. 2111011220006**

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
SEPTEMBER 2025**

HALAMAN PENGESAHAN

SKRIPSI

K-MEANS CLUSTERING PADA PEMETAAN DAERAH RAWAN KRIMINALITAS BERDASARKAN JENIS KRIMINALITAS DI INDONESIA

Oleh:
Uswatun Hasanah
NIM. 2111011220006



telah dipertahankan di depan Dosen Penguji pada tanggal 17 Juli 2025
Susunan Dosen Penguji:

Pembimbing I



Akhmad Yusuf, S.Si., M.Kom.
NIP. 198004022005011001

Dosen Penguji:

1. Oni Soesanto, S.Si., M.Si ()
2. Dr. Muhammad Ahsar Karim, S.Si., M.Sc. ()


Pembimbing II



Dr. Mochammad Idris, S.Si., M.Si.
NIP. 197702142005011001

Banjarbaru, 26 Agustus 2025
Jurusan Matematika FMIPA ULM
Ketua,




Dr. Na'imah Hijriati, S.Si, M.Si
NIP. 197911222008012013

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, 26 Agustus 2025



Uswatun Hasanah

2111011220006

ABSTRAK

K-MEANS CLUSTERING PADA PEMETAAN DAERAH RAWAN KRIMINALITAS BERDASARKAN JENIS KRIMINALITAS DI INDONESIA (Oleh: Uswatun Hasanah; Pembimbing: Akhmad Yusuf, Mochammad Idris; 2025, 71 halaman)

Tingginya angka kriminalitas di Indonesia menyulitkan aparat penegak hukum dan masyarakat dalam menentukan prioritas penanganan dan alokasi sumber daya keamanan secara efektif. Belum adanya pemetaan wilayah rawan kriminalitas berdasarkan jenis kejahatan menyebabkan kebijakan keamanan kurang tepat sasaran. Untuk mengatasi permasalahan tersebut, peneliti melakukan pengelompokan wilayah administratif di Indonesia berdasarkan tingkat kerawanan terhadap sembilan jenis kasus kriminal menggunakan metode *K-Means Clustering*. Peneliti menerapkan dua pendekatan inisialisasi *centroid* awal, yaitu secara acak (*random*) dan metode *Pillar*. Evaluasi jumlah *cluster* optimal dilakukan menggunakan *Elbow Method* dan *Silhouette Coefficient*. Hasil penelitian menunjukkan bahwa pada inisialisasi acak, jumlah *cluster* optimal menurut *Elbow Method* adalah tiga *cluster*, sedangkan *Silhouette Coefficient* tertinggi sebesar 0,5686 diperoleh pada dua *cluster*. Pada metode *Pillar*, *Elbow Method* mengindikasikan empat *cluster* dengan nilai *Silhouette Coefficient* tertinggi sebesar 0,6382. Struktur *cluster* yang dihasilkan kedua metode berbeda, di mana metode *random* menghasilkan kategori rendah, sedang, dan tinggi, sedangkan metode *Pillar* mampu mengidentifikasi kategori tambahan “sangat tinggi.” Peneliti menemukan bahwa kejahatan kesusilaan, terhadap nyawa, narkoba, dan kekerasan terhadap hak milik menjadi indikator dominan pembentukan *cluster*, sementara kejahatan terhadap fisik/badan, hak milik tanpa kekerasan, dan penipuan berkontribusi kecil. Secara keseluruhan metode *Pillar* lebih sensitif dalam mengidentifikasi wilayah dengan tingkat kriminalitas ekstrem.

Kata kunci: Kriminalitas, *K-Means*, *Elbow Method*, *Silhouette Coefficient*, Pemetaan Kejahatan

ABSTRACT

K-MEANS CLUSTERING FOR MAPPING CRIME-PRONE AREAS BASED ON CRIME TYPES IN INDONESIA (By: Uswatun Hasanah; Supervisors: Akhmad Yusuf, Mochammad Idris; 2025, 71 pages)

The high crime rate in Indonesia poses challenges for law enforcement officers and the community in effectively prioritizing crime prevention and allocating security resources. The absence of crime-prone area mapping based on crime types results in less targeted security policies. To address this issue, the researchers grouped administrative regions in Indonesia according to their vulnerability levels to nine types of criminal cases using the K-Means Clustering method. Two centroid initialization approaches were applied: random initialization and the Pillar method. The optimal number of clusters was evaluated using the Elbow Method and Silhouette Coefficient. The results show that with random initialization, the Elbow Method suggests three optimal clusters, while the highest Silhouette Coefficient value of 0.5686 was obtained with two clusters. Using the Pillar method, the Elbow Method indicates four optimal clusters, with the highest Silhouette Coefficient of 0.6382 also at four clusters. The cluster structures generated by the two methods differ; the random method produces clusters categorized as low, medium, and high, whereas the Pillar method identifies an additional "very high" category. The researchers found that crimes against decency, life, narcotics, and violent property crimes are dominant indicators in cluster formation, while crimes against the body, non-violent property crimes, and fraud contribute less. Overall, the Pillar method is more sensitive in identifying regions with extreme crime levels.

Keywords: Crime, K-Means, Elbow Method, Silhouette Coefficient, Crime Mapping

PRAKATA

Segala puji dan syukur penulis panjatkan ke hadirat Allah Subhanahu wa Ta'ala atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan tugas akhir berupa skripsi berjudul “*K-Means Clustering* pada Pemetaan Daerah Rawan Kriminalitas Berdasarkan Jenis Kriminalitas di Indonesia”. Skripsi ini disusun sebagai salah satu syarat untuk meraih gelar Sarjana dalam Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.

Penulis menyadari bahwa penyelesaian skripsi ini tidak terlepas dari dukungan berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Bapak Drs. Abdul Gafur, M.Si., M.Sc., Ph.D., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat Banjarbaru.
2. Ibu Dr. Na'imah Hijriati, S.Si., M.Si., selaku Ketua Jurusan Matematika, FMIPA Universitas Lambung Mangkurat Banjarbaru.
3. Bapak Drs. Faisal, M.Si., selaku dosen pembimbing akademik.
4. Bapak Akhmad Yusuf, S.Si., M.Kom., selaku dosen pembimbing pertama yang telah memberikan arahan dan dukungan dalam penentuan topik skripsi ini.
5. Bapak Dr. Mochammad Idris, S.Si., M.Si., selaku dosen pembimbing kedua yang telah membimbing penulis dalam proses penulisan skripsi.
6. Seluruh dosen Program Studi Matematika yang telah memberikan ilmu dan bimbingan selama masa perkuliahan.
7. Kedua orang tua, adik, kakak, serta seluruh keluarga besar penulis yang selalu memberikan dukungan moril maupun materil.
8. Teman-teman Matematika angkatan 2021 yang telah menjadi teman seperjuangan selama masa studi.
9. Rekan, sahabat, dan kerabat lain yang tak dapat disebutkan satu per satu.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis sangat terbuka terhadap kritik dan saran yang membangun. Akhir kata, semoga skripsi ini dapat memberikan manfaat, khususnya bagi mahasiswa Matematika FMIPA ULM.

Banjarbaru, 26 Agustus 2025

A handwritten signature in black ink, appearing to be 'Uswatun Hasanah', written in a cursive style.

Uswatun Hasanah

2111011220006

DAFTAR ISI

	Halaman
HALAMAN PENGESAHAN	ii
PERNYATAAN	iii
ABSTRAK	iv
ABSTRACT	v
PRAKATA	vi
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan Penelitian	2
1.3 Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA	5
2.1 <i>Knowledge Discovery in Database (KDD)</i>	5
2.2 <i>Data Mining</i>	6
2.3 <i>Clustering</i>	12
2.3.1 Jenis-Jenis Metode <i>Clustering</i>	14
2.4 <i>K- Means</i>	18
2.5 Metode <i>Elbow</i>	22
2.6 <i>Silhouette Coefficient</i>	23
2.7 Kriminalitas.....	25
BAB III PROSEDUR PENELITIAN	27
3.1 Jenis Penelitian.....	27
3.2 Jenis dan Sumber Data	27
3.3 Variabel Penelitian	27

3.4	Langkah Penelitian.....	27
BAB IV HASIL DAN PEMBAHASAN		31
4.1	Pemahaman Data (<i>Data Understanding</i>).....	31
4.1.1	Pengecekan dan Penanganan Data Hilang	31
4.1.2	Eksplorasi Data	31
4.2	Persiapan Data (<i>Data Preparation</i>)	34
4.2.1	Uji Asumsi Multikolinearitas	35
4.2.2	Standarisasi Data dan Uji <i>Outlier</i>	35
4.3	Penerapan Model (<i>Modelling</i>)	38
4.4	Evaluasi model (<i>Evaluation</i>).....	54
4.5	Hasil akhir <i>cluster</i> optimal.....	60
4.6	Interpretasi hasil <i>clustering</i> (<i>Interpretation</i>).....	61
BAB V KESIMPULAN DAN SARAN		66
5.1	Kesimpulan	66
5.2	Saran.....	67
DAFTAR PUSTAKA		69
LAMPIRAN.....		72
RIWAYAT HIDUP		128

DAFTAR TABEL

Tabel 3. 1 Variabel Penelitian	27
Tabel 4. 1 Statistika Deskriptif Data	34
Tabel 4. 2 Nilai VIF Pada Setiap Variabel.....	35
Tabel 4. 3 Nilai <i>z-score</i>	36
Tabel 4. 4 Hasil normalisasi <i>Min-Max</i>	38
Tabel 4. 5 <i>Centroid</i> awal untuk <i>cluster</i> 3.....	39
Tabel 4. 6 Hasil perhitungan jarak (<i>k</i>) = 3 iterasi satu	41
Tabel 4. 7 Hasil pengelompokan (<i>k</i>) = 3 iterasi satu	41
Tabel 4. 8 <i>Centroid</i> baru (<i>k</i>) = 3 iterasi satu	43
Tabel 4. 9 Hasil perhitungan jarak (<i>k</i>) = 3 iterasi dua.....	44
Tabel 4. 10 Hasil pengelompokan (<i>k</i>) = 3 iterasi dua	44
Tabel 4. 11 <i>Centroid</i> akhir (<i>k</i>) = 3 iterasi empat.....	45
Tabel 4. 12 Hasil akhir pengelompokan (<i>k</i>) = 3 iterasi keempat.....	45
Tabel 4. 13 Jarak data dengan <i>mean</i>	48
Tabel 4. 14 <i>Centroid</i> awal dengan metode <i>Pillar</i>	49
Tabel 4. 15 Hasil perhitungan jarak (<i>k</i>) = 4 iterasi satu	50
Tabel 4. 16 Hasil pengelompokan (<i>k</i>) = 4 iterasi satu	50
Tabel 4. 17 <i>Centroid</i> baru (<i>k</i>) =4 iterasi 2.....	52
Tabel 4. 18 Hasil pengelompokan (<i>k</i>) = 4 iterasi dua	52
Tabel 4. 19 <i>Centroid</i> akhir (<i>k</i>) = 4 iterasi tiga	53
Tabel 4. 20 Hasil akhir pengelompokan (<i>k</i>) = 4 iterasi tiga.....	53
Tabel 4. 21 SSE (<i>k</i>) = 2 sampai (<i>k</i>) = 8 <i>K-Means</i> inisiasi random	55
Tabel 4. 22 SSE (<i>k</i>) = 2 sampai (<i>k</i>) = 8 <i>K-Means</i> inisiasi <i>Pillar</i>	57
Tabel 4. 23 Nilai <i>Silhouette</i> (<i>k</i>) = 2 sampai (<i>k</i>) = 8 inisiasi random.....	59
Tabel 4. 24 Nilai <i>Silhouette</i> (<i>k</i>) = 2 sampai (<i>k</i>) = 8 inisiasi <i>Pillar</i>	59
Tabel 4. 25 Karakteristik <i>Cluster</i> Hasil <i>Clustering</i> Inisialisasi Random.....	62
Tabel 4. 26 Karakteristik <i>Cluster</i> Hasil <i>Clustering</i> Inisialisasi <i>Pillar</i>	62

DAFTAR GAMBAR

Gambar 2. 1 Proses CRIPS-DM	6
Gambar 2. 2 Dendogram <i>Hierarki Clustering</i>	15
Gambar 2. 3 Partisi dalam <i>non-hieraki</i>	17
Gambar 2. 4 Titik <i>cluster</i> terbaik	23
Gambar 3. 1 Diagram alur penelitian	30
Gambar 4. 1 Visualisasi data.....	32
Gambar 4. 2 Grafik deteksi <i>outlier</i>	37
Gambar 4. 3 Grafik <i>Elbow K-Means</i> inisiasi random	56
Gambar 4. 4 Grafik <i>Elbow K-Means</i> inisiasi <i>Pillar</i>	57
Gambar 4. 5 Grafik <i>Silhouette Coefficient K-Means</i> inisiasi random.....	59
Gambar 4. 6 Grafik <i>Silhouette Coefficient K-Means</i> inisiasi <i>Pillar</i>	60

DAFTAR LAMPIRAN

LAMPIRAN 1 : Data Mentah Kriminalitas di Indonesia	72
LAMPIRAN 2 : Nilai <i>Z-Score</i>	73
LAMPIRAN 3 : Nilai <i>Min-Max</i> (Data Input)	74
LAMPIRAN 4 : Hasil perhitungan jarak (k) = 3 iterasi satu.....	75
LAMPIRAN 5 : Hasil perhitungan jarak (k) = 3 iterasi dua.....	76
LAMPIRAN 6 : Jarak data dengan <i>mean</i>	77
LAMPIRAN 7 : Hasil perhitungan jarak (k) = 4 iterasi satu.....	78
LAMPIRAN 8 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi Random (k) = 2.....	79
LAMPIRAN 9 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi Random (k) = 3.....	82
LAMPIRAN 10 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi Random (k) = 4.....	85
LAMPIRAN 11 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi Random (k) = 5.....	88
LAMPIRAN 12 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi Random (k) = 6.....	92
LAMPIRAN 13 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi Random (k) = 7.....	95
LAMPIRAN 14 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi Random (k) = 8.....	99
LAMPIRAN 15 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi <i>Pillar</i> (k) = 2.....	102
LAMPIRAN 16 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi <i>Pillar</i> (k) = 3.....	106
LAMPIRAN 17 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi <i>Pillar</i> (k) = 4.....	109
LAMPIRAN 18 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi <i>Pillar</i> (k) = 5.....	113
LAMPIRAN 19 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi <i>Pillar</i> (k) = 6.....	117
LAMPIRAN 20 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi <i>Pillar</i> (k) = 7.....	120
LAMPIRAN 21 : <i>Script</i> Algoritma <i>K-Means</i> Inisialisasi <i>Pillar</i> (k) = 8.....	124