

# Penerapan Algoritma K-Means Untuk Klasterisasi Kekerasan Berdasarkan Tempat Kejadian Di Jawa Barat

Fernando Kurniawan<sup>1)</sup>, Teti Desyani<sup>2\*)</sup>

1. Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang, Indonesia
2. Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang, Indonesia

## Article Info

**Kata Kunci:** Jawa Barat; Kasus Kekerasan; Klasterisasi; *K-Means*; Tempat Kejadian

**Keywords:** *West Java; Violence Cases; Clustering; K-Means; Incident Location*

## Article history:

Received 21 Desember 2025

Revised 22 Desember 2025

Accepted 24 Desember 2025

Available online 1 Mei 2026

## DOI :

[10.48144/suryainformatika.v16i1.2291](https://doi.org/10.48144/suryainformatika.v16i1.2291)

\* Corresponding author.

Teti Desyani

E-mail address:

[dosen00839@unpam.ac.id](mailto:dosen00839@unpam.ac.id)

## ABSTRAK

Fenomena kekerasan masih menjadi permasalahan sosial yang memerlukan perhatian serius, khususnya di Provinsi Jawa Barat yang memiliki jumlah kasus relatif tinggi. Kasus kekerasan dapat terjadi di berbagai lingkungan sosial, sehingga diperlukan analisis berbasis data untuk memahami pola distribusinya secara lebih terstruktur. Penelitian ini bertujuan untuk mengelompokkan kasus kekerasan di Provinsi Jawa Barat berdasarkan kategori tempat kejadian dengan menerapkan algoritma *K-Means*. Data penelitian diperoleh dari *Open Data Jabar* dengan periode pengamatan tahun 2020–2024. Proses analisis dilakukan menggunakan kerangka *Knowledge Discovery in Database (KDD)* yang mencakup tahap pemilihan data, pra-proses, transformasi data, data mining, dan interpretasi. Implementasi algoritma *K-Means* dilakukan menggunakan perangkat lunak *Altair AI Studio*. Penentuan jumlah kluster optimal dilakukan menggunakan *Elbow Method*, sedangkan kualitas kluster dievaluasi menggunakan *Davies-Bouldin Index (DBI)*. Hasil penelitian menunjukkan bahwa jumlah kluster optimal yang terbentuk adalah 4 kluster dengan nilai *DBI* sebesar 0,622, yang menunjukkan kualitas klasterisasi cukup baik. Setiap kluster merepresentasikan tingkat intensitas kasus kekerasan yang berbeda, mulai dari intensitas rendah hingga sangat tinggi. Kluster dengan intensitas sangat tinggi didominasi oleh kasus kekerasan pada kategori rumah tangga dan kategori lainnya, sedangkan kluster dengan intensitas rendah menunjukkan jumlah kasus yang relatif kecil pada seluruh kategori tempat kejadian. Hasil klasterisasi ini memberikan pemahaman yang lebih sistematis mengenai pola distribusi kasus kekerasan berdasarkan konteks tempat kejadian dan dapat dimanfaatkan sebagai dasar pendukung dalam perumusan strategi pencegahan dan penanganan kekerasan yang lebih terarah.

## ABSTRACT

*Violence remains a social problem that requires serious attention, particularly in West Java Province, which records a relatively high number of cases. Acts of violence occur across various social environments, making data-driven analysis essential to obtain a structured understanding of their distribution patterns. This study aims to cluster violence cases in West Java Province based on incident location categories using the K-Means algorithm. The data were obtained from Open Data Jabar covering the period from 2020 to 2024. The analysis process followed the Knowledge Discovery in Database (KDD) framework, including data selection, preprocessing, data transformation, data mining, and interpretation stages. The implementation of the K-Means algorithm was carried out using Altair AI Studio (RapidMiner). The optimal number of*

clusters was determined using the Elbow Method, while cluster quality was evaluated using the Davies-Bouldin Index (DBI). The results indicate that four optimal clusters were formed, with a DBI value of 0.622, indicating a fairly good clustering quality. Each cluster represents different levels of violence intensity, ranging from low to very high. The cluster with very high intensity is predominantly characterized by violence cases occurring in domestic settings and other locations, whereas the low-intensity cluster shows relatively small numbers of cases across all incident categories. Overall, the clustering results provide a more systematic insight into the distribution patterns of violence cases based on incident locations and can serve as supporting information for formulating more targeted violence prevention and intervention strategies.

## 1. PENDAHULUAN

Fenomena kekerasan hingga kini masih menjadi persoalan sosial yang menimbulkan dampak serius bagi korban, baik secara fisik, psikologis, maupun sosial. Di Indonesia, kasus kekerasan dilaporkan terjadi pada masyarakat tanpa memandang usia maupun status sosial. Laporan KPPPA menunjukkan bahwa pada tahun 2024 terdapat lebih dari 31 ribu kasus kekerasan, dan Jawa Barat merupakan salah satu provinsi dengan jumlah kasus yang sangat tinggi. Tingginya angka kasus tersebut menandakan bahwa kekerasan masih menjadi isu krusial yang membutuhkan upaya penanganan yang tepat serta analisis berbasis data untuk memahami pola kasus kekerasan secara lebih komprehensif [1].

Kekerasan tidak hanya terjadi dalam konteks tertentu, tetapi dapat berlangsung di berbagai lingkungan sosial. Kekerasan sering terjadi di lingkungan rumah tangga dan memberikan dampak serius terhadap perkembangan psikologis dan sosial korban. Dampak kekerasan dapat memicu gangguan emosional, mengasingkan diri dari lingkungan sosial, dan kecenderungan perilaku agresif pada tahap perkembangan selanjutnya [2]. Selain itu, lingkungan pendidikan juga tidak terlepas dari praktik kekerasan, baik dalam bentuk fisik, verbal, maupun simbolik, yang berdampak negatif terhadap rasa aman, psikologis, dan kualitas interaksi sosial korban [3]. Temuan tersebut menunjukkan bahwa kekerasan merupakan fenomena yang terjadi di berbagai jenis lingkungan, termasuk lingkungan kerja, fasilitas umum, lembaga pendidikan kilat, dan ruang sosial lainnya. Sehingga analisis berdasarkan konteks tempat kejadian menjadi penting untuk memperoleh gambaran pola kekerasan yang komprehensif.

Pendekatan *data mining* telah banyak dimanfaatkan untuk menganalisis data untuk mengidentifikasi pola dan karakteristik yang sulit diamati secara manual. Salah satu teknik yang umum digunakan adalah clustering dengan algoritma *K-Means*, karena mampu mengelompokkan berdasarkan tingkat kemiripan data secara efektif. seperti pada analisis kejadian bencana alam, algoritma *K-Means* terbukti menghasilkan pengelompokan wilayah dengan tingkat kejadian tinggi, sedang dan

rendah yang dapat dimanfaatkan sebagai dasar perencanaan dan pengambilan keputusan berbasis data [4]. Selain itu *data mining* juga diterapkan dalam konteks kebijakan publik untuk mengekstraksi informasi penting dari suatu data, sehingga mendukung perumusan kebijakan yang lebih responsif [5].

Berbagai penelitian terkait kekerasan di Jawa Barat telah dilakukan dengan pendekatan klasterisasi, meskipun masing-masing memiliki fokus yang berbeda. Penelitian terdahulu terkait kekerasan berdasarkan jenis kelamin dan kelompok umur, diperoleh enam klaster korban menggunakan *Davies-Bouldin Index* sebagai ukuran evaluasi metrik [6]. Penelitian lain menitikberatkan pada kelompok usia anak dan perempuan, dan hasilnya menunjukkan variasi rentang usia yang paling sering menjadi korban kekerasan di Jawa Barat [7]. selain penelitian yang berfokus pada karakteristik individu korban, penelitian lain juga menerapkan algoritma *K-Means* untuk menganalisis kekerasan berdasarkan jenis kekerasan, hasilnya menunjukkan pengelompokan daerah rawan kekerasan terhadap anak dan perempuan [1].

Dari berbagai penelitian terdahulu, terlihat masih terdapat kesenjangan penelitian. Penelitian terdahulu menerapkan klasterisasi untuk menganalisis kasus kekerasan berdasarkan karakteristik individual korban [6] [7], maupun berdasarkan jenis kekerasan [1]. Meskipun demikian, penelitian terdahulu tersebut belum mengaitkan kategori tempat kejadian kekerasan sebagai variabel kontekstual dalam analisis, padahal informasi tersebut penting untuk memahami konteks dan karakteristik kekerasan di suatu wilayah secara komprehensif. Selain itu, data yang digunakan dalam penelitian terdahulu umumnya berada pada rentang tahun 2017-2021, sehingga belum sepenuhnya merepresentasikan kondisi kekerasan terbaru.

Berdasarkan kesenjangan tersebut, penelitian ini bertujuan untuk klasterisasi kasus kekerasan di Provinsi Jawa Barat berdasarkan jenis tempat kejadian menggunakan algoritma *K-Means* dengan memanfaatkan data periode tahun 2020-2024. Hasil analisis diharapkan mampu mengelompokkan wilayah dengan

tingkat kekerasan tertentu, sehingga dapat digunakan sebagai dasar dalam merancang strategi pencegahan kekerasan yang lebih terarah dan memudahkan masyarakat memahami pola kekerasan di Jawa Barat.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan data mining dengan algoritma *K-Means* untuk menganalisis dan mengelompokkan kasus kekerasan yang terjadi di Provinsi Jawa Barat. Proses pengolahan data dilakukan dengan mengacu pada kerangka *Knowledge Discovery in Database (KDD)*, digunakan karena mampu memberikan alur sistematis dalam mengekstraksi pengetahuan dari suatu data yang kompleks [8].

2.1 Sumber Data

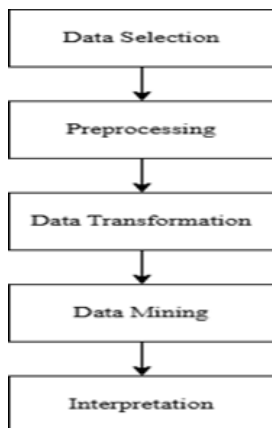
Data yang digunakan dalam penelitian ini bersumber dari *Open Data Jabar* sebagai sumber data resmi dari pemerintah [9]. Dataset mencakup beberapa variabel yaitu id, kode provinsi, nama provinsi, kode kabupaten dan kota, nama kabupaten dan kota, tempat kejadian, jumlah kekerasan, satuan, dan tahun.

Data dapat diakses melalui <https://opendata.jabarprov.go.id/id/dataset/jumlah-kasus-kekerasan-berdasarkan-tempat-kejadian-kekerasan-di-jawa-barat>

nama_provinsi	nama_kabupaten_kota	tempat_kejadian	jumlah_kekerasan	satuan	tahun
JAWA BARAT	KABUPATEN BOGOR	RUMAH TANGGA	5	KASUS	2020
JAWA BARAT	KABUPATEN BOGOR	TEMPAT KERJA	0	KASUS	2020
JAWA BARAT	KABUPATEN BOGOR	LAINNYA	2	KASUS	2020
JAWA BARAT	KABUPATEN BOGOR	SEKOLAH	4	KASUS	2020
JAWA BARAT	KABUPATEN BOGOR	FASILITAS UMUM	1	KASUS	2020
JAWA BARAT	KABUPATEN BOGOR	LEMBAGA PENDIDIKAN KILAT	0	KASUS	2020
JAWA BARAT	KABUPATEN SUKABUMI	RUMAH TANGGA	23	KASUS	2020
JAWA BARAT	KABUPATEN SUKABUMI	TEMPAT KERJA	5	KASUS	2020
JAWA BARAT	KABUPATEN SUKABUMI	LAINNYA	93	KASUS	2020
JAWA BARAT	KABUPATEN SUKABUMI	SEKOLAH	4	KASUS	2020

Gambar 1. Dataset kasus kekerasan dari *Open Data Jabar*

2.2 Knowledge Discovery in Database



Gambar 2. Tahapan *Knowledge Discovery in Database*

Berikut adalah penjelasan dari tahapan *Knowledge Discovery in Database (KDD)*.

*Data Selection*, merupakan proses pemilihan atribut yang relevan dan penting dipilih dari suatu data untuk dianalisis [10]. Pada tahap ini, data kasus kekerasan berdasarkan kategori tempat kejadian di Provinsi Jawa Barat dipilih sebagai data yang akan dianalisis. Variabel yang dipilih yaitu nama kabupaten kota, tempat kejadian, jumlah kekerasan, dan tahun.

*Preprocessing*, dilakukan untuk meningkatkan kualitas data sebelum proses analisis lebih lanjut. Proses ini meliputi pemeriksaan kelengkapan data, pembersihan data, konsistensi nilai, dan penanganan data yang tidak sesuai [11].

*Data Transformation*, data disusun dan disesuaikan ke dalam format numerik yang sesuai untuk proses klusterisasi [8]. Transformasi dilakukan agar data dapat diproses secara optimal oleh algoritma *K-Means*. Pada tahap ini peneliti melakukan *pivot table* dikarenakan *K-Means* hanya dapat menganalisis atribut bernilai numerik. Selain itu, dilakukan penggabungan variabel nama kabupaten/kota dan tahun untuk menciptakan suatu id unik.

*Data Mining*, merupakan tahap inti penelitian dengan menerapkan algoritma *K-Means* untuk klusterisasi berdasarkan kemiripan tingkat kasus kekerasan pada setiap kategori tempat kejadian [12]. Proses klusterisasi dilakukan dengan menentukan jumlah kluster, menghitung jarak antara data dan pusat kluster menggunakan *Euclidean Distance*, serta memperbarui nilai pusat kluster secara iteratif hingga mencapai kondisi konvergen [13]. Rumus Perhitungan jarak antar data menggunakan *Euclidean Distance* [14] dengan indeks seperti rumus 1.

$$d(i, j) = \sqrt{(x_{1i} - x_{1j})^2 + \dots + (x_{ki} - x_{kj})^2} \quad (1)$$

di mana  $d(i, j)$  menyatakan jarak antara data ke- $i$  dan *centroid* ke- $j$ . komponen  $x_{1i}$  hingga  $x_{ki}$  merepresentasikan nilai atribut ke-1 sampai atribut ke- $k$  pada data ke- $i$ , sedangkan  $x_{1j}$  hingga  $x_{kj}$  merupakan nilai atribut ke-1 hingga atribut ke- $k$  pada *centroid* ke- $j$ . adapun  $k$  menunjukkan jumlah atribut atau fitur yang digunakan dalam proses klusterisasi

*Interpretation*, dilakukan dengan menganalisis karakteristik setiap kluster yang terbentuk [15]. Tahap ini bertujuan untuk menafsirkan makna klusterisasi yang terbentuk, sehingga dapat diperoleh pemahaman mengenai kecenderungan penyebaran intensitas kasus kekerasan pada masing-masing kluster. Dengan demikian, interpretasi kluster dapat memberikan gambaran mengenai variasi distribusi kasus kekerasan antar wilayah dan rentang waktu berdasarkan tempat kejadian.

2.1 Elbow Method

Penentuan jumlah kluster optimal dilakukan menggunakan *Elbow Method*, dengan menganalisis perubahan nilai *Sum of Squared Error (SSE)* pada beberapa percobaan jumlah kluster. Jumlah kluster optimal ditentukan berdasarkan titik di mana penurunan nilai *SSE* mulai melambat, yang menunjukkan keseimbangan antara jumlah kluster dan kualitas pengelompokan [16].

2.2 Davies-Bouldin Index

Untuk mengevaluasi kualitas kluster yang terbentuk, penelitian ini menggunakan *Davies-Bouldin Index (DBI)*. Indeks ini mengukur rasio kedekatan antar kluster dan tingkat penyebaran data di dalam kluster. Nilai *DBI* yang rendah menunjukkan hasil klusterisasi yang lebih baik karena kluster yang terbentuk memiliki tingkat homogenitas yang tinggi dan pemisahan yang jelas [17].

3. HASIL DAN PEMBAHASAN

Data yang digunakan dalam proses klusterisasi merupakan hasil dari tahap transformasi data, di mana data kasus kekerasan disusun ke dalam bentuk numerik sesuai dengan kebutuhan algoritma *K-Means*. Unit analisis pada penelitian ini ditetapkan berdasarkan kombinasi antara wilayah kabupaten/kota dan tahun kejadian. Penyusunan tersebut dilakukan agar model kluster mampu menangkap variasi intensitas kasus kekerasan berdasarkan perbedaan waktu pada wilayah yang sama, sehingga analisis tidak hanya menggambarkan perbedaan antar wilayah, tetapi juga memungkinkan identifikasi perubahan pola kekerasan antar tahun pada setiap wilayah. Transformasi data dilakukan melalui penyajian kategori tempat kejadian dalam bentuk atribut numerik dengan enam variabel, yaitu RT (Rumah Tangga), TK (Tempat Kerja), SKL (Sekolah), FASUM (Fasilitas Umum), LPK (Lembaga Pendidikan Kilat), dan LAIN (Kategori Lainnya), sehingga setiap nilai pada atribut tersebut merepresentasikan jumlah kasus kekerasan berdasarkan kategori lokasi dan tahun pada masing-masing wilayah.

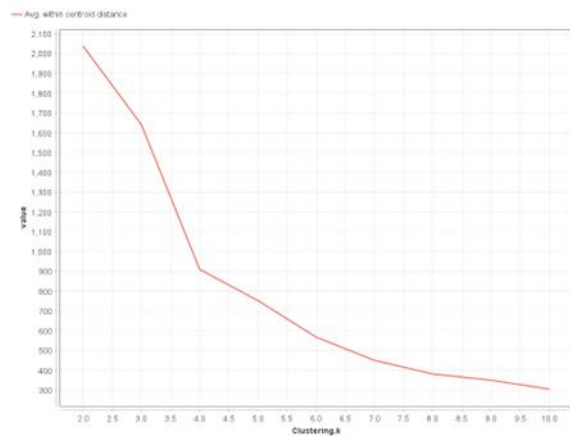
Tabel 1 Data transformasi

No.	ID	FASUM	LAIN	LPK	RT	SKL	TK
1	KABUPATEN BANDUNG BARAT-2020	0	13	0	9	1	0
2	KABUPATEN BANDUNG BARAT-2021	0	21	0	12	3	0
3	KABUPATEN BANDUNG BARAT-2022	1	4	0	38	0	1
4	KABUPATEN BANDUNG BARAT-2023	0	4	0	54	0	2
5	KABUPATEN BANDUNG BARAT-2024	0	3	0	63	0	0
...	...	...	...	...	...	...	...

132	KOTA TASIKMALAYA-2021	15	49	0	54	7	1
133	KOTA TASIKMALAYA-2022	2	1	0	3	2	0
134	KOTA TASIKMALAYA-2023	0	4	0	3	0	1
135	KOTA TASIKMALAYA-2024	10	7	0	26	2	1

3.1. Elbow Method

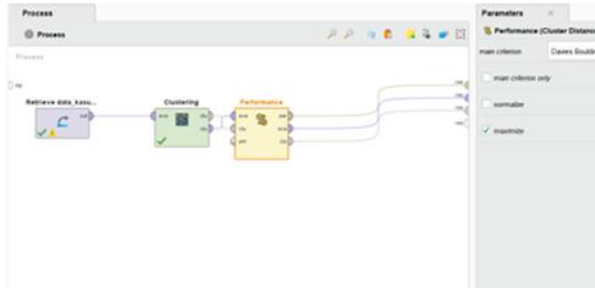
Penentuan jumlah kluster optimal dilakukan menggunakan *Elbow Method* dengan menganalisis perubahan *Sum of Squared Error (SSE)* pada beberapa percobaan jumlah kluster. Hasil pengujian menunjukkan titik siku (*elbow point*) terbentuk pada jumlah 4 kluster, yang ditandai dengan penurunan nilai *SSE* yang signifikan sebelum cenderung stabil pada jumlah kluster berikutnya.



Gambar 3. Grafik *Elbow Method*

3.2. Penerapan Algoritma *K-Means*

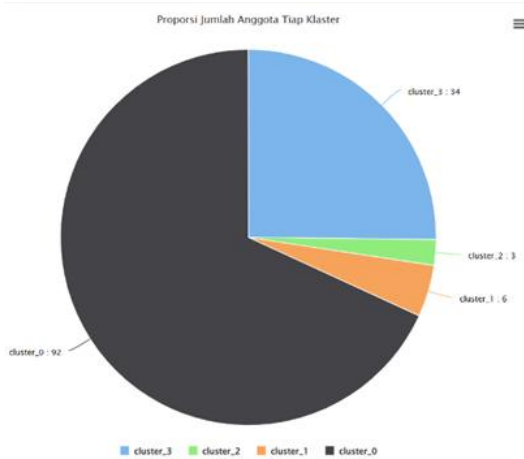
Proses klusterisasi dalam penelitian ini dilakukan menggunakan *Altair AI Studio* sebagai alat bantu analisis data. Dataset kasus kekerasan diproses melalui algoritma *K-Means*, kemudian hasil klusterisasi dievaluasi menggunakan *Davies-Bouldin Index (DBI)* untuk menilai kualitas kluster yang terbentuk. Alur kerja ini memastikan bahwa hasil pengelompokan tidak hanya berdasarkan kemiripan data, tetapi juga telah melalui proses evaluasi kualitas kluster secara kuantitatif sebelum dilakukan interpretasi lebih lanjut.



Gambar 4. Alur kerja K-Means di Altair Ai Studio

### 3.3. Proporsi Jumlah Anggota Tiap Kluster

Berdasarkan hasil *cluster model*, *Cluster 0* memiliki jumlah anggota terbanyak yaitu 92 data, *Cluster 3* sebanyak 34 data, *Cluster 1* sebanyak 6 data, sedangkan *Cluster 2* memiliki jumlah anggota sebanyak 3 data, dari total keseluruhan 135 data. Perbedaan jumlah anggota pada setiap kluster menunjukkan ketimpangan distribusi kasus kekerasan antar wilayah dan periode waktu.



Gambar 5. Visualisasi proporsi jumlah anggota tiap kluster

### 3.4. Davies-Bouldin Index

Hasil evaluasi menunjukkan nilai *Davies-Bouldin Index (DBI)* sebesar 0,622, yang mengindikasikan bahwa kluster yang terbentuk cukup baik. Nilai *DBI* yang relatif rendah menunjukkan bahwa model klusterisasi mampu merepresentasikan struktur data secara memadai dan layak untuk analisis lanjutan.

```
PerformanceVector:
Davies Bouldin: 0.622
Avg. within centroid distance: 909.450
Avg. within centroid distance cluster
0: 909.450
Avg. within centroid distance cluster
1: 909.450
Avg. within centroid distance cluster
2: 909.450
Avg. within centroid distance cluster
3: 909.450
```

### 3.5. Interpretation

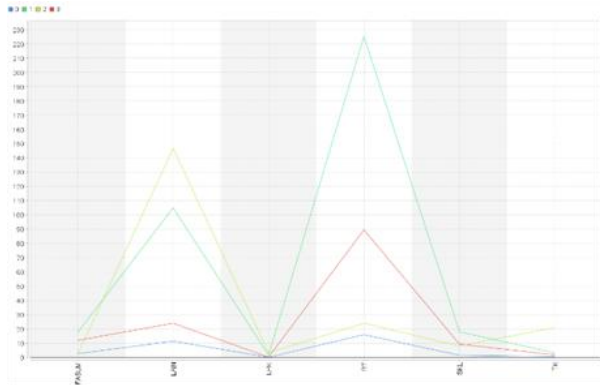
Hasil klusterisasi menghasilkan pengelompokan unit analisis berupa id kabupaten/kota dan tahun (2020-2024) ke dalam kluster dengan karakteristik tingkat kasus kekerasan yang berbeda. Setiap kluster dibentuk berdasarkan 6 kategori tempat kejadian kekerasan, yaitu rumah tangga, tempat kerja, sekolah, fasilitas umum, lembaga pendidikan kilat, dan kategori lainnya.

Kluster yang terbentuk dapat diklasifikasikan ke dalam kelompok dengan intensitas kasus rendah, sedang, tinggi, dan sangat tinggi. *Cluster 0* merepresentasikan wilayah dengan intensitas kasus kekerasan rendah, yang ditandai dengan jumlah kasus yang relatif kecil pada seluruh kategori tempat kejadian. *Cluster 1* menunjukkan intensitas kasus sangat tinggi, yang didominasi oleh tingginya jumlah kasus pada kategori rumah tangga dan lainnya. *Cluster 2* dengan intensitas kasus tinggi, yang terutama dipengaruhi oleh tingginya jumlah kasus kekerasan pada kategori lainnya. sementara *Cluster 3* menunjukkan intensitas kasus sedang, di mana nilai rata-rata pada kategori rumah tangga dan kategori lainnya terlihat cukup menonjol, namun masih berada di bawah tingkat intensitas yang ditunjukkan oleh *Cluster 1* dan *Cluster 2*.



Gambar 6. Scatter plot menggunakan atribut RT dan LAIN

*Line chart* menunjukkan jumlah rata-rata kekerasan dalam rumah tangga dan kategori lainnya menjadi kategori yang paling dominan. sedangkan kategori fasilitas umum, sekolah, dan tempat kerja berada pada rentang rata-rata rendah hingga sedang, sementara kategori lembaga pendidikan kilat merupakan kategori dengan rata-rata paling rendah.



Gambar 7. Line chart centroid

Hasil klusterisasi berdasarkan tingkat intensitas kasus kekerasan memberikan informasi penting terkait persebaran kasus kekerasan pada tiap wilayah. Informasi tersebut dapat dimanfaatkan sebagai dasar dalam menetapkan prioritas wilayah dalam pelaksanaan program pencegahan dan penanganan kekerasan secara lebih terarah oleh pemangku kebijakan [18]. Pembagian wilayah ke dalam klaster berisiko sangat tinggi, tinggi, sedang, dan rendah memungkinkan alokasi sumber daya, pengawasan, serta intervensi kebijakan dilakukan secara proporsional sesuai karakteristik klaster.

Tabel 2. Distribusi wilayah tiap klaster

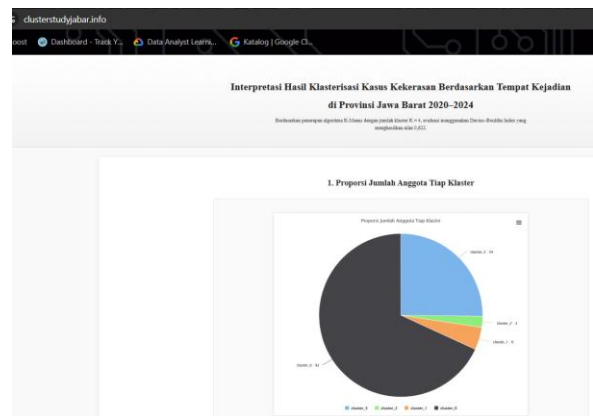
No.	ID	CLUSTER
1	KABUPATEN BANDUNG BARAT-2020	0
2	KABUPATEN BANDUNG BARAT-2021	0
3	KABUPATEN BANDUNG BARAT-2022	0
4	KABUPATEN BANDUNG BARAT-2023	0
	KABUPATEN BANDUNG BARAT-2024	3
...	...	...
79	KABUPATEN SUKABUMI-2023	2
80	KABUPATEN SUKABUMI-2024	2
81	KABUPATEN SUMEDANG-2020	0
82	KABUPATEN SUMEDANG-2021	0
...	...	...
131	KOTA TASIKMALAYA-2021	3
132	KOTA TASIKMALAYA-2022	0
134	KOTA TASIKMALAYA-2023	0
135	KOTA TASIKMALAYA-2024	0

Berdasarkan distribusi keanggotaan klaster dengan unit analisis kombinasi nama kabupaten/kota dan tahun, dapat diamati bahwa sebagian besar wilayah berada pada Cluster 0 intensitas rendah, namun sejumlah wilayah berpindah ke klaster lain pada periode tertentu. Beberapa wilayah menunjukkan peningkatan intensitas kasus yang tercermin dari perpindahan ke klaster dengan intensitas yang lebih tinggi pada tahun berikutnya, sedangkan wilayah lainnya justru berpindah ke klaster

dengan intensitas lebih rendah seiring penurunan jumlah kasus.

Pola perubahan ini mengindikasikan bahwa tingkat kekerasan pada kategori tempat kejadian bersifat dinamis baik antar wilayah maupun antar tahun periode 2020-2024, karena intensitas kasus dapat meningkat atau menurun pada periode tertentu. Dengan demikian, klusterisasi dengan unit analisis kombinasi wilayah dan tahun memberikan pemahaman yang lebih tepat mengenai perubahan intensitas kekerasan dari tahun ke tahun, tanpa menutupi variasi yang mungkin hilang apabila data digabungkan selama 5 tahun observasi.

Sebagai upaya meningkatkan keterbukaan dan pemahaman masyarakat. Hasil analisis disajikan dalam bentuk visualisasi berbasis web yang dapat diakses melalui platform *clusterstudyjabar.info*. pemanfaatan visualisasi data dalam konteks pemerintahan dan kebijakan publik telah terbukti mampu meningkatkan transparansi dan kualitas pengambilan keputusan berbasis data [19].



Gambar 8. Web interpretasi

#### 4. KESIMPULAN

Berdasarkan hasil penelitian, dari penerapan algoritma K-Means pada data kasus kekerasan di Provinsi Jawa Barat menunjukkan bahwa pendekatan klusterisasi mampu untuk mengelompokkan data kasus kekerasan berdasarkan kemiripan karakteristik. Melalui tahapan *Knowledge Discovery in Database*, jumlah klaster optimal sebanyak 4 klaster. Hasil evaluasi menggunakan *Davies-Bouldin Index* sebesar 0,622 mengindikasikan klaster terbentuk cukup baik.

Hasil klusterisasi menunjukkan adanya pola distribusi kasus kekerasan yang berbeda setiap klaster. Cluster 0 merepresentasikan wilayah dengan intensitas kasus kekerasan rendah pada semua kategori tempat kejadian. Cluster 1 menunjukkan intensitas kasus sangat tinggi yang didominasi oleh kategori rumah tangga dan kategori lainnya. Cluster 2 menggambarkan intensitas kasus tinggi yang dipengaruhi oleh tingginya jumlah kasus pada kategori lainnya. Cluster 3 menunjukkan intensitas kasus sedang, cenderung menonjol pada kasus kategori rumah tangga dan lainnya, tetapi tidak

setinggi Cluster 1 dan Cluster 2. Pola ini menegaskan bahwa kategori tempat kejadian berkontribusi secara signifikan dalam membedakan pola kekerasan antar wilayah dan periode waktu, serta membentuk karakteristik cluster yang beragam.

Secara keseluruhan, hasil klusterisasi memberikan pemahaman yang lebih terstruktur mengenai distribusi dan intensitas kasus kekerasan di Provinsi Jawa Barat berdasarkan konteks tempat kejadian. Informasi yang dihasilkan dapat dimanfaatkan sebagai dasar pendukung dalam penentuan prioritas wilayah, perumusan strategi pencegahan, dan penanganan kekerasan yang lebih terarah. Penyajian hasil analisis melalui visualisasi data juga berpotensi meningkatkan keterbukaan informasi dan membantu masyarakat maupun pemangku kepentingan dalam memahami pola kekerasan secara lebih komprehensif.

Penelitian selanjutnya dapat mempertimbangkan penambahan variabel lain yang berpotensi memengaruhi pola kekerasan serta membandingkan algoritma klusterisasi yang berbeda untuk menguji konsistensi hasil klusterisasi.

## REFERENSI

- [1] R. Rahma and R. Mufidah, "Pengelompokan Daerah Rawan Kekerasan Terhadap Perempuan dan Anak Di Jawa Barat Menggunakan Algoritma K-Means," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 7, no. 3, pp. 850–857, 2022, doi: 10.29100/jupi.v7i3.3144.
- [2] D. C. Damayanti, S. Musa, and S. Mulyadi, "Dampak Kekerasan dalam Rumah Tangga terhadap Permasalahan Perkembangan Sosial Anak Usia Dini," *J. Kaji. Gend. dan Anak*, vol. 7, no. 1, 2023.
- [3] A. Nur Fadhilah and Munjin, "Kekerasan dalam Pendidikan di Sekolah: Bentuk, Sebab, Dampak, dan Solusi," *J. KEPENDIDIKAN*, vol. 10, no. 2, pp. 325–344, 2022.
- [4] D. Rohman, R. Annisa, D. Indriyana Efendi, and D. Solahudin, "CLUSTERING BENCANA ALAM MENGGUNAKAN K-MEANS PADA WILAYAH JAWA BARAT," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, 2024.
- [5] R. A. Atmiko, M. Priangga, N. Panca, B. Hizbullah, M. Anis, and N. Fauzi, "Pemanfaatan Data X Untuk Analisis Sentimen Publik Pada Kebijakan PPN 12 %," *J. Inform. dan Teknol. Interaktif*, vol. 2, no. 1, pp. 274–283, 2025.
- [6] M. Riskandi, M. Martanto, and U. Hayati, "Klusterisasi Korban Kekerasan Menggunakan Algoritma K-Means Di Jawa Barat," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 820–826, 2024, doi: 10.36040/jati.v8i1.8457.
- [7] R. Fauziah and A. I. Purnamasari, "Implementasi Algoritma K-Means pada Kasus Kekerasan Anak dan Perempuan Berdasarkan Usia," *Hello World J. Ilmu Komput.*, vol. 2, no. 1, pp. 34–41, 2023, doi: 10.56211/helloworld.v2i1.232.
- [8] S. Nachwa *et al.*, "PENDEKATAN KLASIFIKASI DALAM KNOWLEDGE DISCOVERY UNTUK SULTAN MAHMUD BADARUDDIN II DI GOOGLE MAPS," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 3, pp. 4782–4789, 2025.
- [9] N. P. Dharshinni, G. Singh, J. A. T. Naibaho, R. D. Lumban Tobing, and A. Simamora, "Penerapan algoritma k-means pada data pengangguran di jawa barat," *DSI J. Data Sci. Indones.*, vol. 3, no. 1, pp. 23–34, 2023.
- [10] D. A. Farizal and F. A. Alijoyo, "Optimasi Pemilihan Metode Pengajaran Dosen Menggunakan Data Mining, dan Algoritma K-Means dalam Proses Bisnis Pendidikan," *INFORMATICS Digit. Expert*, vol. 2, pp. 105–109, 2024.
- [11] R. Karunia Ibadirachman, Y. Herry Chrisnanto, and P. Nurul Sabrina, "OPTIMASI PARAMETER DBSCAN MENGGUNAKAN METODE DIFFERENTIAL EVOLUTION UNTUK DETEKSI ANOMALI PADA DATA TRANSAKSI BANK," *J. Teknol. Terpadu*, vol. 10, no. 1, pp. 22–31, 2024.
- [12] S. Pujiono, R. Astuti, and F. M. Basysyar, "IMPLEMENTASI DATA MINING UNTUK MENENTUKAN POLA PENJUALAN PRODUK MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, 2024.
- [13] M. Maesaroh, T. N. Padilah, J. H. Jaman, F. I. Komputer, and U. S. Karawang, "PENERAPAN ALGORITMA K-MEANS CLUSTERING PADA PENGELOMPOKAN," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 4, pp. 2783–2787, 2023.
- [14] M. Iqbal, Syaripuddin, and M. Nurul Huda, "Implementasi Algoritma K-Means Clustering dengan Jarak Euclidean dalam Mengelompokkan Daerah Penyebaran," *Basis J. Ilm. Mat.*, vol. 2, no. 1, pp. 47–56, 2023.
- [15] B. A. Prayudha *et al.*, "ALGORITMA K-MEANS UNTUK MENINGKATKAN MODEL KLASTERISASI DATA SISWA SMK SAMUDRA NUSANTARA KABUPATEN CIREBON," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 1, pp. 1314–1321, 2025.
- [16] N. A. Maori, "METODE ELBOW DALAM OPTIMASI JUMLAH CLUSTER PADA K-MEANS CLUSTERING," *SIMETRIS*, vol. 14, no. 2, pp. 277–287, 2023.
- [17] Y. Hasan, U. B. Darma, K. Medan, and S. Utara, "PENGUKURAN SILHOUETTE SCORE DAN DAVIES-BOULDIN INDEX PADA HASIL CLUSTER K-MEANS DAN," *JITET (Jurnal Inform. dan Tek. Elektro Ter.*, vol. 12, no. 3, 2024.
- [18] G. Natalie Krista Anindita and Nurhayati, "PENGARUH PROGRAM PENGENTASAN KEMISKINAN," *Media Ekon.*, vol. 32, no. 1, pp. 61–76, 2024.
- [19] D. A. Syafiq and Y. Sembiring, "Pengembangan Model Visualisasi Data Kependudukan Berbasis Teknologi Informasi untuk Mendukung Pengambilan Keputusan di Kecamatan Medan Amplas," *J. Penelit. Inov.*, vol. 5, no. 2, pp. 1329–1338, 2025.