

IDENTIFIKASI POLA DAN FAKTOR UTAMA JENIS KEJAHATAN DI INDONESIA DENGAN ANALISIS KOMPONEN UTAMA (PCA) DAN K-MEANS CLUSTERING

Erlangga Ria Krishna^{1*}, Nuramaliyah²

^{1,2}*Program Studi Statistika, Fakultas Sains dan Teknologi, Universitas Terbuka.*

**Penulis Korespondensi: 043572947@ecampus.ut.ac.id*

ABSTRAK

Penelitian ini bertujuan untuk mengidentifikasi pola dan faktor utama jenis kejahatan di Indonesia menggunakan Analisis Komponen Utama (PCA) dan *K-means clustering* sebagai analisis lanjutan dengan data Statistik Kriminal Indonesia dari Badan Pusat Statistik (BPS) periode 2019-2023. Analisis dilakukan terhadap 18 variabel yang mencakup kejahatan kekerasan, kejahatan terhadap properti, dan pelanggaran sosial. Hasil PCA menunjukkan dua komponen utama yang mampu menjelaskan 85,18% variasi data, mengonfirmasi bahwa sebagian besar dinamika kriminalitas Indonesia dapat direduksi ke dua dimensi inti yaitu PC₁ diidentifikasi sebagai faktor risiko finansial dan pelanggaran kepercayaan, dibentuk kuat oleh kejahatan seperti kekerasan dalam rumah tangga (X₁₄), penggelapan (X₁₇), dan penipuan (X₁₆). Sementara itu, PC₂ diidentifikasi sebagai faktor konflik ekstrem vs. regulatif, menunjukkan polarisasi antara isu pelanggaran ketertiban umum (X₁₉) dengan kejahatan kekerasan material. Selanjutnya, *K-Means Clustering* menghasilkan dua klaster yang merepresentasikan pergeseran struktural pola kejahatan yaitu klaster awal (2019-2021) yang stabil dan klaster transformasi (2022-2023). Klaster Transformasi ini ditandai oleh lonjakan faktor regulatif pada 2022 (PC₂ tertinggi) dan bergeser tajam ke dominasi faktor risiko finansial pada 2023 (PC₁ tertinggi). Temuan ini menegaskan bahwa pola kriminalitas di Indonesia tidak statis, melainkan bergerak dinamis mengikuti tekanan sosial-ekonomi, sehingga kebijakan penanggulangan harus adaptif dan fokus pada intervensi sosial-ekonomi untuk mengendalikan faktor risiko finansial (PC₁).

Kata kunci: PCA, K-means clustering, kejahatan, pola kriminalitas, Indonesia.

1 PENDAHULUAN

Kriminalitas merupakan salah satu permasalahan sosial yang terus menjadi perhatian di berbagai negara, termasuk Indonesia. Tingkat kriminalitas tidak hanya mencerminkan kondisi sosial-ekonomi masyarakat, tetapi juga menunjukkan efektivitas kebijakan keamanan serta ketahanan sosial suatu wilayah. Berdasarkan data Statistik Kriminal Indonesia yang diterbitkan oleh Badan Pusat Statistik (BPS) untuk periode 2019 hingga 2023 tren angka kejahatan di Indonesia menunjukkan fluktuasi yang signifikan di berbagai daerah (Badan Pusat Statistik, 2023). Pola ini umumnya berkaitan dengan ketimpangan sosial, tingkat pengangguran, urbanisasi, hingga perkembangan teknologi informasi yang memunculkan bentuk kejahatan baru. Artinya, kriminalitas di Indonesia tidak bisa lagi dilihat sebagai fenomena tunggal, karena jenis-jenis kejahatan saling berkaitan dan cenderung berubah polanya dari waktu ke waktu.

Fenomena sosial dalam beberapa tahun terakhir juga menunjukkan bahwa kriminalitas di Indonesia tidak lagi sepenuhnya bersifat konvensional. Misalnya, laporan Jakarta *Feminist Association* mencatat adanya lonjakan femisida pada 2024, dengan jumlah sekitar 184 hingga 187 kasus dan mayoritas terjadi dalam relasi intim atau keluarga (Jakarta *Feminist Association*, 2024; Kompas, 2024). Kondisi ini memperlihatkan bahwa kekerasan berbasis gender masih menjadi ancaman kuat di ranah domestik. Di sisi lain, pemerintah juga menghadapi maraknya perjudian

daring, lebih dari dua juta situs judi online ilegal telah ditutup, dan aktivitas ini menimbulkan dampak ekonomi maupun sosial yang luas (Reuters, 2024). Dua contoh tersebut mempertegas bahwa kriminalitas berkembang semakin kompleks, baik dari sisi jenis, pola, maupun dampaknya terhadap masyarakat.

Perubahan ini menunjukkan bahwa kriminalitas di Indonesia mengalami transformasi struktural, bukan hanya meningkat atau menurun secara kuantitas, tetapi juga berubah kualitas dan karakteristiknya. Sejumlah penelitian terdahulu menyebutkan bahwa faktor sosial-ekonomi seperti pengangguran, kemiskinan, dan ketimpangan sosial masih sering menjadi determinan penting kriminalitas di Indonesia. Septriani (2024) menegaskan bahwa kenaikan pengangguran berhubungan erat dengan meningkatnya kejahatan properti. Temuan serupa juga ditunjukkan oleh Fitriyyah & Pratama (2025) dan Ardefa dkk (2025) yang menyimpulkan bahwa kemiskinan dan pengangguran memiliki dampak positif signifikan terhadap *crime rate* di Indonesia. Selain itu, Watianthos dkk. (2023) menjelaskan bahwa kriminalitas di Indonesia bersifat spasial dan cenderung membentuk pola sebaran tertentu antardaerah. Namun, sebagian besar penelitian nasional masih dominan memakai analisis deskriptif atau clustering wilayah, sehingga keterkaitan antar-jenis kejahatan dan pola dominannya belum banyak diurai melalui pendekatan reduksi dimensi. Padahal, data kriminal bersifat multivariat dan saling berkorelasi, sehingga ada peluang besar untuk melihat “struktur utama” yang tersembunyi di balik banyaknya variabel kejahatan.

Dalam konteks tersebut, penggunaan Analisis Komponen Utama (*Principal Component Analysis/PCA*) menjadi relevan. PCA berfungsi mereduksi data multivariat yang kompleks menjadi beberapa komponen inti yang tetap mempertahankan sebagian besar informasi, sehingga variabel yang paling dominan membentuk variasi kriminalitas dapat terlihat lebih jelas. Penelitian Abda'u & Faiz (2024) menunjukkan bahwa kombinasi PCA dan K-means efektif untuk merangkum pola dominan kejahatan sekaligus mengelompokkan observasi berdasarkan kemiripan karakteristiknya. Pendekatan serupa juga dipakai pada studi analisis tren kriminal di berbagai wilayah lain, yang menyimpulkan bahwa PCA membantu mengatasi banyaknya variabel yang saling berkorelasi sebelum dilakukan *clustering* (Krishnendu dkk., 2020). Setelah komponen utama terbentuk, penelitian ini memanfaatkan scree plot dan biplot untuk membaca kontribusi variabel serta kecenderungan pola antarperiode, lalu dilanjutkan pengelompokan menggunakan K-means clustering berdasarkan skor komponen utama.

Berdasarkan uraian sebelumnya, penelitian ini bertujuan untuk mengidentifikasi komponen/faktor utama yang membentuk variasi jenis kejahatan di Indonesia, menjelaskan keterkaitan antarjenis kejahatan berdasarkan hasil PCA dan biplot, serta mengelompokkan pola kriminalitas Indonesia selama 2019-2023 menggunakan *K-means clustering* agar pergeseran karakter kejahatan antarperiode dapat terlihat lebih spesifik. Dengan demikian, penelitian ini diharapkan dapat menjadi dasar penyusunan kebijakan penanggulangan kejahatan yang lebih adaptif dan berbasis data.

2 METODE

2.1. Jenis dan Sumber Data

Penelitian ini menggunakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) melalui publikasi Statistik Kriminal Indonesia periode 2019 hingga 2023, yang dapat diakses melalui laman resmi <https://www.bps.go.id>. Banyak data yang digunakan sebanyak 18 data dengan data sebagai berikut.

Tabel 1. Variabel Penelitian

Variabel	Keterangan
X1	Tahun
X2	Jumlah Kasus Pembunuhan
X3	Jumlah Kasus Penganiayaan Ringan
X4	Jumlah Kasus Kekerasan dalam Rumah Jumlah Kasus Tangga
X5	Jumlah Kasus Penculikan
X6	Jumlah Kasus Mempekerjakan Anak di Jumlah Kasus Bawah Umur
X7	Jumlah Kasus Perkosaan
X8	Jumlah Kasus Pencabulan
X9	Jumlah Kasus Pencurian dengan Kekerasan
X10	Jumlah Kasus Pencurian
X11	Jumlah Kasus Pencurian dengan Pemberatan
X12	Jumlah Kasus Pencurian Kendaraan Bermotor
X13	Jumlah Kasus Pengrusakan/Penghancuran Barang
X14	Jumlah Kasus Pembakaran dengan Sengaja
X15	Jumlah Kasus Narkotika dan Psikotropika
X16	Jumlah Kasus Penipuan/Perbuatan Curang
X17	Jumlah Kasus Penggelapan
X18	Jumlah Kasus Korupsi
X19	Jumlah Kasus Kejahatan Terhadap Ketertiban Umum

Sumber: *Badan Pusat Statistik 2019-2023*

2.2. Metode Pengolahan dan Analisis

Penelitian ini menggunakan dua metode utama, yaitu Analisis Komponen Utama (PCA) dan *K-means* clustering. PCA digunakan untuk mereduksi dimensi data kriminal yang memiliki banyak variabel dan saling berkorelasi dan *K-means* clustering diterapkan pada skor komponen utama untuk mengelompokkan pola kriminalitas antartahun berdasarkan kemiripan karakteristik.

2.3. Normalisasi Data

Sebelum PCA dilakukan, data dinormalisasi agar setiap variabel berada pada skala yang sama. Untuk menggunakan variabel analisis yang sama posisi skala antar variabel di normalisasi ke skala 0-1 menggunakan *Min-Max Normalization*. Penggunaan Min-Max sesuai praktik umum dalam data mining, di mana nilai seluruh fitur disesuaikan ke dalam rentang yang sama tanpa merubah pola korelasi (Rizky & Fauziyah, 2025) yang dihitung dengan rumus berikut.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

dengan:

X_{norm} = Nilai data yang dinormalisasi

X = Nilai asal data

X_{max} dan X_{min} = Nilai minimum dan maksimum pada dataset per kolom

2.4. Analisis Komponen Utama (PCA)

PCA adalah metode statistik untuk mereduksi data multivariat menjadi beberapa komponen utama agar struktur dominan pada data lebih mudah terlihat. Pendekatan ini banyak digunakan dalam analisis data kriminalitas untuk mengurangi kompleksitas data tanpa kehilangan informasi penting (Maharrani dkk, 2024; Retno & Hakimi, 2025; Destia & Kartikasari, 2023). Tahapan PCA dalam penelitian ini adalah sebagai berikut.

1) Menghitung Matriks Kovarians

Untuk memastikan bahwa semua variabel berada dalam skala yang sama, langkah berikutnya adalah menghitung matriks kovarians pada dua variabel X dan Y dihitung menggunakan rumus berikut.

$$Cov(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n - 1}$$

dengan:

$X_i ; Y_i$ = Nilai observasi pada variabel X dan Y

$\bar{Y} ; \bar{X}$ = Rata-rata variabel X dan Y

n = Jumlah observasi

Sehingga, didapatkannya Matriks Kovarians (C) berikut.

$$C = \begin{bmatrix} Var(X_1) & Cov(X_1, X_2) & Cov(X_1, X_p) \\ Cov(X_2, X_1) & Var(X_2) & Cov(X_2, X_p) \\ \vdots & \vdots & \ddots \\ Cov(X_p, X_1) & Cov(X_p, X_2) & Var(X_p) \end{bmatrix}$$

2) Menghitung *Eigenvalue* dan *Eigenvector*

Setelah menghitung matriks kovarians, langkah berikutnya adalah menghitung nilai eigen dan vektor eigen. Matriks kovarians C dikalikan dengan vektor eigen V untuk menghasilkan nilai eigen λ . Hubungan matriks Kovarians dan vektor eigen diuraikan dengan rumus berikut.

$$C.V = \lambda.V$$

dengan:

C = Matriks kovarians

V = *Eigenvector*, arah komponen utama

λ = *Eigenvalue*, kontribusi komponen terhadap varians

Komponen utama yang dipilih adalah komponen dengan nilai eigen lebih dari 1, karena komponen tersebut menjelaskan variansi yang signifikan dalam data. Rumus variansi sebagai perbandingan adalah sebagai berikut.

$$Explained\ Variance\ Ratio = \frac{\lambda_i}{\sum \lambda}$$

dengan:

λ_i = *Eigenvalue* dari komponen utama

$\sum \lambda$ = Jumlah seluruh *eigenvalues*

3) Menentukan Komponen Utama

Setelah memilih komponen utama, data asli ditransformasikan ke dalam ruang komponen utama menggunakan eigenvector yang telah dihitung. Rumus yang dimaksud adalah sebagai berikut:

$$Z = XW$$

dengan:

Z = Matriks komponen utama

X = Matriks standar data asli

W = Matriks *eigenvector*

4) Scree Plot dan Biplot

Scree plot digunakan untuk menentukan jumlah komponen utama optimal. Grafik ini menampilkan *eigenvalue* tiap komponen secara berurutan, dan komponen dipilih ketika grafik mulai membentuk ‘*elbow*’ atau titik siku (Maharrani dkk 2024). Pada penelitian ini, scree plot menunjukkan bahwa dua komponen utama sudah cukup menjelaskan mayoritas variasi data, sesuai pendekatan yang umum digunakan dalam analisis PCA (Retno & Hakimi, 2025). Biplot menampilkan titik observasi (tahun) sekaligus vektor variabel kejahatan, sehingga dapat terlihat:

- Kedekatan antarperiode (tahun) berdasarkan karakteristik kriminalitasnya
- Arah dan kekuatan kontribusi variabel terhadap komponen utama
- Korelasi antarjenis kejahatan (vektor searah berarti korelasi positif, berlawanan berarti negatif).

2.5. *K-means Clustering*

Setelah melakukan analisis PCA untuk mereduksi dimensi data, langkah selanjutnya adalah menggunakan *K-means clustering* untuk mengelompokkan data berdasarkan skor komponen utama. *K-means clustering* adalah metode pembelajaran tidak terawasi yang digunakan untuk mengelompokkan data ke dalam K kelompok berdasarkan kemiripan atau jarak antar titik data. Setiap kelompok memiliki pusat (*centroid*) yang dihitung berdasarkan rata-rata posisi titik data dalam kelompok tersebut. Tujuan dari *K-means* adalah meminimalkan jarak antar data dalam satu cluster dan memaksimalkan jarak antar *cluster* (Maharrani dkk 2024). Tahapan *K-means* pada penelitian ini adalah sebagai berikut.

- Menentukan jumlah cluster (K). Nilai K ditentukan menggunakan metode *elbow*, yaitu melihat titik ketika penurunan total variansi mulai melandai.
- Menentukan centroid awal secara acak. *Centroid* adalah pusat *cluster* yang mewakili rata-rata posisi titik data.
- Menghitung jarak Euclidean. Jarak antara setiap titik data x_i dan pusat cluster *centroid* c_k menggunakan rumus jarak *Euclidean*.

$$\text{Jarak } (x_i, c_k) = \sqrt{\sum_{j=1}^n (x_{ij} - c_{kj})^2}$$

dengan:

x_i = Titik data ke- i ; c_k = Titik data ke- k

x_{ij} = Nilai variabel ke- j pada data ke- i

c_{kj} = Nilai variabel ke- j pada data ke- k

n = jumlah variabel (dimensi data)

- 4) Mengelompokkan data ke cluster terdekat. Setelah menghitung jarak, setiap titik data x_i akan dimasukkan ke dalam *cluster* yang memiliki *centroid* dengan jarak terdekat. Setiap titik data akan dipetakan ke cluster yang minimalkan jarak *Euclidean* dengan rumus berikut.

$$Cluster(X_i) = \arg \min Jarak(X_i, C_k)$$

- 5) Memperbarui centroid. Jika data dikelompokkan, lakukan perhitungan ulang posisi centroid dengan menghitung rata-rata dari semua titik data dalam cluster tersebut dengan rumus berikut.

$$c_k = \frac{1}{N_k} \sum_{x_i \in C_k} x_i$$

dengan:

c_k = Posisi *centroid* untuk cluster ke-k

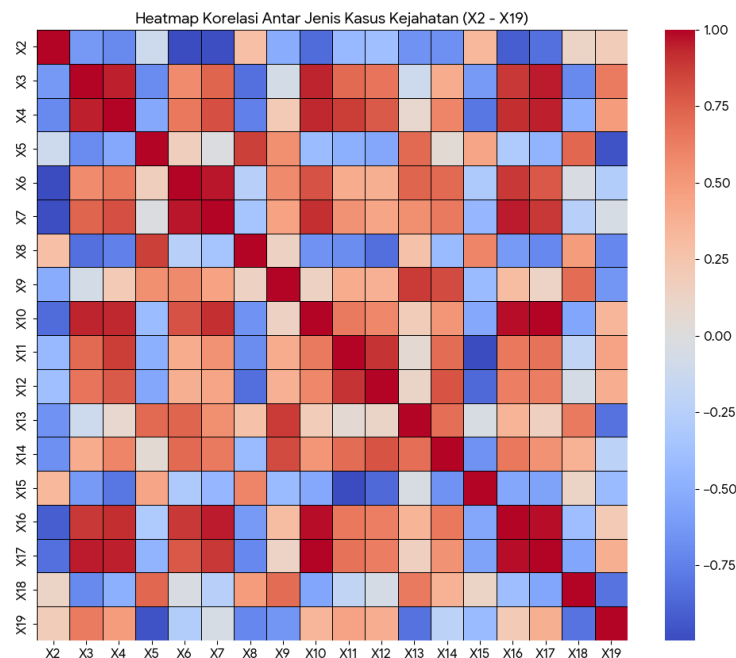
N_k = Jumlah titik data dalam cluster ke-k

C_k = semua titik data yang termasuk dalam cluster ke-k

- 6) Iterasi hingga konvergen. Langkah terakhir, lakukan pengulangan dengan memperbarui *centroid* dan mengelompokkan data ulang hingga posisi centroid tidak berubah lagi atau perubahan dalam jarak total antar *centroid* minimal (konvergen).

3 HASIL DAN PEMBAHASAN

3.1. Analisis Deskriptif Data



Gambar 1. Heatmap Korelasi Antar Jenis Kasus Kejahatan (X2 - X19)

Pada Gambar 1, memperlihatkan bahwa pola kriminalitas di Indonesia memiliki hubungan struktural yang tegas antar jenis kejahatan. Terdapat korelasi positif yang kuat (warna merah cerah) antara kejahatan properti yaitu berupa pencurian (X_{10}), penipuan (X_{16}), dan penggelapan (X_{17}) dengan kejahatan kekerasan non-fatal yaitu berupa penganiayaan ringan (X_3) dan kekerasan dalam Rumah Tangga (X_4). Sementara itu, jenis kasus Narkotika/Psikotropika (X_{15}) menunjukkan

korelasi yang lemah atau negatif terhadap sebagian besar kejahatan lain, mengindikasikan adanya dinamika pendorong yang berbeda dan terpisah dari kluster properti/kekerasan.

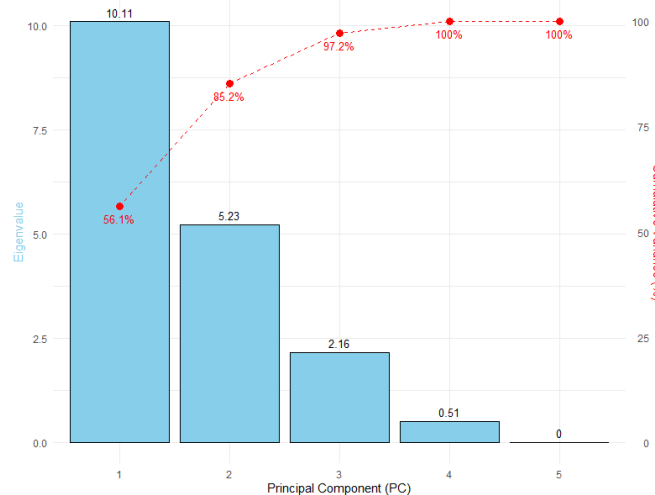
Jika dilihat dari konteks sosial yang lebih luas, periode ini memang diwarnai meningkatnya kompleksitas kejahatan. Di ranah kekerasan berbasis gender, misalnya, laporan femisida 2024 menunjukkan kasus kekerasan fatal terhadap perempuan masih tinggi dan banyak terjadi di relasi domestik (Jakarta *Feminist Association*, 2024; Kompas, 2024). Sementara di ranah pelanggaran sosial modern, lonjakan judi online juga makin terlihat nyata, bahkan pemerintah menutup lebih dari dua juta situs ilegal karena skalanya sudah masif (Reuters, 2024). Jadi, data deskriptif ini memberikan gambaran awal bahwa pola kriminal di Indonesia bergerak dinamis dan dipengaruhi beberapa dimensi berbeda, bukan sekadar “tinggi atau rendahnya kasus”.

3.2. Analisis Komponen Utama (PCA)

Tabel 2. Nilai Eigen, Persentase Variansi, dan Akumulasi Variansi

Parameter	PC ₁	PC ₂	PC ₃	PC ₄	PC ₅		PC ₁₈
Eigenvalue	12.63	6.53	2.70	0.63	0		0
% of Variance	56.14	29.04	11.99	2.83	0		0
Cumulative (%)	56.14	85.18	97.17	100	100		100

Pada Tabel 2, hasil PCA menunjukkan bahwa dua komponen utama sudah menjelaskan 85,18% variasi data. Komponen pertama (PC₁) menyumbang 56,14%, sedangkan komponen kedua (PC₂) menyumbang 29.04%. Hal tersebut juga dijelaskan pada Gambar 2 berikut.



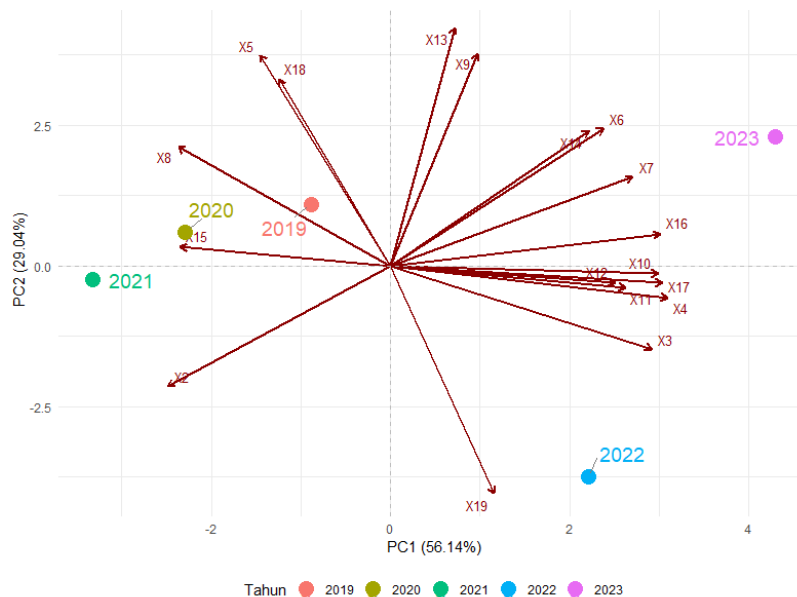
Gambar 2. Scree Plot Eigen Value vs Cumulative Variance

Scree plot pada Gambar 2 menunjukkan bahwa setelah PC₂, garis pada scree plot menurun drastis dan mendatar, menandakan bahwa komponen berikutnya hanya menyumbang variansi yang sangat kecil (<3%). Hal ini sesuai dengan kriteria Kaiser (eigenvalue > 1) dan prinsip *elbow*, sehingga pemilihan dua komponen utama sudah optimal.

Komponen pertama (PC₁) merepresentasikan kelompok risiko finansial dan pelanggaran kepercayaan. Tiga variabel dengan loading tertinggi adalah kekerasan dalam rumah tangga (X₄ = 0.3087), penggelapan (X₁₇ = 0.3031), dan penipuan/perbuatan curang (X₁₆ = 0.3016). Pola ini menunjukkan bahwa PC₁ menggambarkan intensitas kejahatan yang bersifat ekonomi dan pelanggaran kepercayaan. Fenomena ini masuk akal mengingat tekanan ekonomi dan

ketimpangan sosial sering menjadi pemicu tindakan kriminal agresif, sebagaimana diungkap oleh Septriani (2024) yang menekankan korelasi antara pengangguran dan kejahatan properti.

Komponen kedua (PC_2) menangkap dimensi kejahatan berdasarkan sifat pelanggaran regulatif dengan kekerasan material. Variabel dengan loading negatif terbesar adalah pengrusakan/penghancuran barang ($X_{13} = -0.4236$), pencurian dengan kekerasan ($X_9 = -0.3767$), dan penculikan ($X_5 = -0.3774$). Di sisi lain, PC_2 juga ditarik kuat oleh variabel sosial yang bernilai positif, terutama pelanggaran ketertiban umum ($X_{19} = 0.4037$), pembunuhan ($X_2 = 0.2132$), serta penganiayaan berat ($X_3 = 0.1476$). Ini menunjukkan bahwa perubahan pada jenis kejahatan tertentu bisa menggeser karakter kriminalitas antarperiode, sejalan dengan temuan Abdissiam & Pratomo (2024) tentang ketidakstabilan sosial pascapandemi.



Gambar 3. Grafik Biplot Analisis Komponen Utama (PCA) Jenis Kejahatan di Indonesia

Biplot pada Gambar 3 memperlihatkan distribusi tahun terhadap dua komponen utama (PC_1 dan PC_2). Dimensi PC_1 (Horizontal), yang merupakan faktor paling dominan, memisahkan pola kejahatan berdasarkan risiko finansial dan pelanggaran kepercayaan. Hal ini ditunjukkan oleh kluster vektor yang sangat kuat di sisi kanan Biplot, didorong oleh variabel dengan loading positif tertinggi seperti penggelapan ($X_{17} = 0.3031$), penipuan/perbuatan curang ($X_{16} = 0.3016$), dan kekerasan dalam rumah tangga ($X_4 = 0.3087$). Vektor-vektor ini secara kolektif mendefinisikan tahun-tahun yang dicirikan oleh tingginya insiden kejahatan bermotif ekonomi non-fisik. Kutub ini berlawanan dengan Narkotika ($X_{15} = -0.2348$) dan pembunuhan ($X_2 = -0.2484$) yang menarik pola ke sisi kiri PC_1 .

Sementara itu, Dimensi PC_2 (Vertikal) memisahkan kejahatan berdasarkan sifat pelanggaran regulatif dengan kekerasan material. Sumbu positif PC_2 di bagian atas didominasi secara jelas oleh terhadap ketertiban umum ($X_{19} = 0.4037$), mencerminkan adanya variasi yang terkait dengan isu tata kelola sosial. Sebaliknya, sumbu negatif PC_2 di bagian bawah ditarik oleh vektor kekerasan dan pemaksaan seperti pengrusakan/penghancuran barang ($X_{13} = -0.4236$) dan pencurian dengan kekerasan ($X_9 = -0.3767$). Pemisahan ini mengindikasikan bahwa dinamika isu regulatif dan kekerasan material bergerak secara independen dalam menentukan karakteristik kejahatan tahunan.

Tabel 3. PC scores

Tahun	PC1	PC2	PC3	PC4
2019	-0.9955	-1.2264	2.6482	-0.4311
2020	-2.5618	-0.6779	-0.2653	1.2824
2021	-3.7156	0.2741	-1.6385	-0.8368
2022	2.4706	4.1892	0.2389	0.079
2023	4.8023	-2.559	-0.9833	-0.0934

Pada tabel 3, pola pergeseran karakter kejahatan dari tahun ke tahun terlihat sangat jelas melalui skor komponen utama pada PC₁ dan PC₂. Periode awal 2019 dan 2020 ditandai oleh skor PC₁ dan PC₂ yang rendah dan negatif (misalnya PC₁ tahun 2020 adalah -2.56). Angka ini menegaskan bahwa profil kejahatan pada periode tersebut jauh dari faktor risiko finansial dan didominasi oleh variabel di kutub berlawanan seperti Narkotika (X₁₅) dan pembunuhan (X₂). Namun, perubahan mulai terlihat pada Tahun 2021, meskipun PC₁ tetap negatif (-3.72), skor PC₂ mulai bergerak positif (0.27). Pergerakan PC₂ ini mengindikasikan munculnya dimensi sosial dan isu regulatif sebagai variabel yang mulai berpengaruh. Pergeseran ini sejalan dengan laporan Abdissiam & Pratomo (2024) tentang dampak pandemi terhadap ketidakstabilan sosial.

Puncak dorongan Regulatif terjadi pada Tahun 2022, yang mencatat skor PC₂ tertinggi (4.19). Skor ekstrem ini menunjukkan dominasi pelanggaran ketertiban umum (X₁₉) yang memiliki loading positif sangat besar (0.4037) pada tahun tersebut. Namun, puncak transformasi pola kejahatan terlihat pada Tahun 2023, yang mencatat skor PC₁ tertinggi (4.80) dan PC₂ negatif (-2.56). Angka PC₁ yang sangat tinggi ini adalah bukti paling kuat adanya pergeseran fokus kejahatan ke arah risiko finansial dan properti. skor ini didorong oleh variabel dominan pc₁ seperti kekerasan dalam rumah tangga (X₄ = 0.3087), penggelapan (X₁₇ = 0.3031), dan penipuan/perbuatan curang (X₁₆ = 0.3016). Sebaliknya, skor PC₂ negatif menunjukkan bahwa profil kejahatan tahun 2023 juga memiliki kedekatan dengan kejahatan kekerasan material, yang ditarik oleh loading negatif besar pada pengrusakan/penghancuran barang (X₁₃ = -0.4236) dan pencurian dengan kekerasan (X₉ = -0.3767).

Berdasarkan hasil PCA, persamaan untuk dua komponen utama dapat dituliskan sebagai berikut.

$$\begin{aligned}
 PC1 &= (-0.2484) X_2 + (0.2916) X_3 + (0.3087) X_4 + (-0.1453) X_5 + (0.2377) X_6 \\
 &\quad + (0.2702) X_7 + (-0.2368) X_8 + (0.0967) X_9 + (0.2988) X_{10} + (0.2620) X_{11} \\
 &\quad + (0.2508) X_{12} + (0.0714) X_{13} + (0.2219) X_{14} + (-0.2348) X_{15} \\
 &\quad + (0.3016) X_{16} + (0.3031) X_{17} + (-0.1245) X_{18} + (0.1155) X_{19}
 \end{aligned}$$

$$\begin{aligned}
 PC2 &= (0.2132) X_2 + (0.1476) X_3 + (0.0565) X_4 + (-0.3744) X_5 + (-0.2447) X_6 \\
 &\quad + (-0.1589) X_7 + (-0.2112) X_8 + (-0.3767) X_9 + (0.0136) X_{10} + (0.0379) X_{11} \\
 &\quad + (0.0298) X_{12} + (-0.4236) X_{13} + (-0.2404) X_{14} + (-0.0338) X_{15} \\
 &\quad + (-0.0570) X_{16} + (0.0300) X_{17} + (-0.3313) X_{18} + (0.4037) X_{19}
 \end{aligned}$$

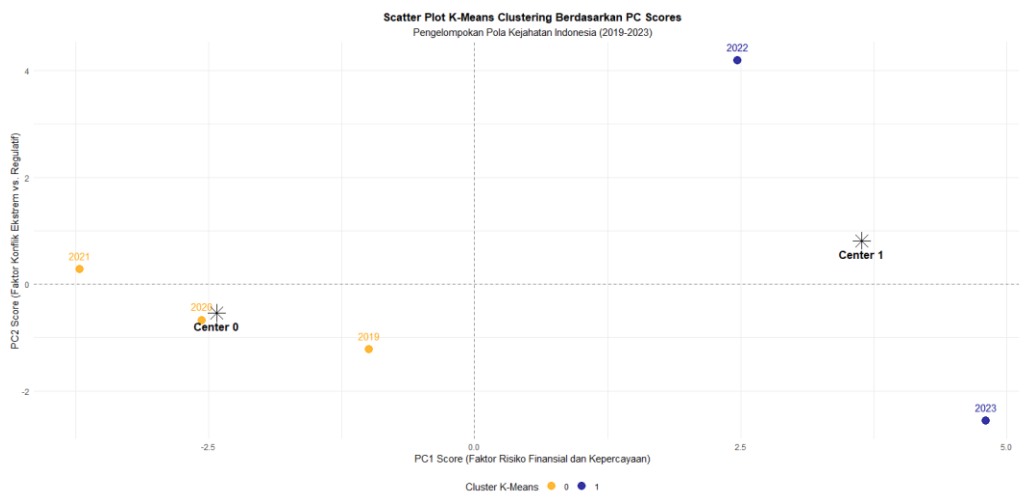
Nilai positif menunjukkan hubungan searah, sedangkan nilai negatif menunjukkan hubungan berlawanan. Jadi, makin besar variabel yang loading-nya positif, skor komponennya ikut naik. Jika nilainya negatif, maka akan menurunkan skor komponen. Contohnya, pelanggaran ketertiban umum (X₁₉) punya loading positif besar di PC₂ (0,4037), jadi ketika kasusnya naik, skor PC₂ ikut meningkat. Sebaliknya, pengrusakan/penghancuran barang (X₁₃) punya loading negatif besar di

PC_2 (-0,4236), artinya peningkatan kasus ini justru menurunkan skor PC_2 . Ini nunjukkan bahwa PC_2 memang baca dua arah pola yaitu sisi sosial-regulatif di satu sisi, dan sisi konflik/ekstrem di sisi lain.

3.3. Analisis *K-Means Clustering*

Setelah dilakukan reduksi dimensi menggunakan PCA, langkah berikutnya adalah pengelompokan data dengan metode *K-Means Clustering*. Analisis ini menggunakan skor komponen utama (PC_1 dan PC_2) sebagai dasar pengelompokan karena kedua komponen tersebut menjelaskan 85.18% variansi data, sehingga cukup representatif untuk menggambarkan pola kejahatan di Indonesia.

Proses K-Means dilakukan dengan $K = 2$ dan iterasi maksimum 20 kali untuk memastikan stabilitas hasil. Pemilihan dua cluster didasarkan pada metode elbow dan pertimbangan jumlah observasi (5 tahun), yang menunjukkan bahwa dua kelompok sudah cukup untuk memisahkan pola kriminalitas yang berbeda.



Gambar 4. Scatter Plot K-Means Berdasarkan Skor PCA

Pada gambar 4, memperlihatkan distribusi tahun terhadap dua komponen utama dengan warna yang menunjukkan cluster. Cluster 0 (periode karakteristik rendah risiko finansial) mencakup tahun 2019-2021, sedangkan Cluster 1 (periode karakteristik transformasi ekstrem) mencakup tahun 2022-2023. Pola ini menunjukkan pergeseran yang jelas yaitu awal periode berada di sisi negatif PC_1 dan rendah PC_2 , sedangkan akhir periode berada di sisi positif PC_1 dan PC_2 .

Tabel 4. Cluster Report

Tahun	PC_1 Score	PC_2 Score	Cluster
2019	-0.9955	-1.2264	0
2020	-2.5618	-0.6779	0
2021	-3.7156	0.2741	0
2022	2.4706	4.1892	1
2023	4.8023	-2.559	1

Tabel 4 menampilkan skor PC_1 dan PC_2 untuk setiap tahun beserta pengelompokan cluster. Pada Cluster 0 (periode karakteristik rendah risiko finansial) berada pada tahun 2019-2021 yang

memiliki skor PC_1 yang seluruhnya negatif (-0.99 hingga -3.72) dan PC_2 yang relatif rendah. Kondisi ini menunjukkan bahwa profil kejahatan pada awal periode secara statistik menjauhi Faktor Risiko Finansial, dengan intensitas kejahatan yang masih terpengaruh oleh variabel di kutub negatif PC_1 (Narkotika dan pembunuhan). Sedangkan, pada Cluster 1 (periode karakteristik transformasi ekstrem) berada pada tahun 2022-2023 ditandai dengan skor PC_1 yang sangat positif (2.47 hingga 4.80), menegaskan dominasi faktor risiko finansial. Walaupun kedua tahun berada dalam satu cluster, PC_2 skornya sangat ekstrem dan berlawanan (4.19 di 2022 dengan -2.56 di 2023), menandakan cluster ini menampung dua fase transformasi yang sangat kuat, tetapi arah perubahannya berbeda (dominasi regulatif di 2022 dan dominasi kekerasan material di 2023).

Tabel 5. Cluster Centers

Cluster	PC_1	PC_2	Karakteristik Utama
0	-2.4243	-0.5434	Karakteristik Awal (Rendah Risiko Finansial)
1	3.6365	0.8151	Karakteristik Transformasi (Tinggi Risiko Finansial)

Tabel 5 menunjukkan pusat cluster yang sangat tegas dan terpisah jauh. Cluster 0 berpusat pada PC_1 negatif (-2.42), menggambarkan periode awal yang didominasi oleh pola kejahatan yang kurang dipengaruhi oleh isu finansial. Sebaliknya, Cluster 1 berpusat pada PC_1 yang sangat positif (3.64), menandai periode akhir dengan intensitas faktor risiko finansial yang jauh lebih tinggi. Hasil ini menegaskan secara statistik bahwa pola kriminalitas di Indonesia 2019-2023 tidak statis, tetapi bergeser secara signifikan mengikuti dinamika sosial-ekonomi.

4 KESIMPULAN

Berdasarkan hasil analisis, dapat disimpulkan bahwa pola kriminalitas di Indonesia selama periode pengamatan memiliki struktur tertentu dan mengalami perubahan dari waktu ke waktu. Hasil PCA menunjukkan bahwa berbagai jenis kejahatan dapat diringkas ke dalam faktor-faktor utama yang mencerminkan kondisi sosial dan ekonomi, sementara K-Means memperlihatkan adanya pergeseran pola kejahatan antarperiode. Temuan ini menegaskan bahwa kriminalitas di Indonesia bersifat dinamis, sehingga kebijakan penanggulangan harus adaptif dan fokus pada faktor risiko finansial serta isu regulative.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh pihak yang telah memberikan dukungan, bantuan, dan kerja sama selama pelaksanaan penelitian ini, khususnya dalam proses perizinan, pengumpulan data, serta pemberian masukan yang konstruktif terhadap penyusunan naskah. Dukungan tersebut sangat berarti dalam penyelesaian penelitian ini.

DAFTAR PUSTAKA

- Abdissiam, D. A., & Pratomo, R. A. (2024). Dinamika jumlah dan pola sebaran kriminalitas sebelum hingga setelah COVID-19 di Kelurahan Damai, Kota Balikpapan. *REGION: Jurnal Pembangunan Wilayah dan Perencanaan Partisipatif*, 19(2), 123–135. <https://doi.org/10.20961/region.v19i2.85249>
- Ardefa, S. P., Naeruz, M., & Apriandi, A. (2025). Analisis pengaruh kemiskinan, pengangguran dan tingkat pendidikan terhadap tingkat kriminalitas di Indonesia. *Jurnal Pengabdian*

- Masyarakat dan Riset Pendidikan (JERKIN), 3(4), 4173–4177.
<https://doi.org/10.31004/jerkin.v3i4.1014>
- Badan Pusat Statistik. (2023). *Statistik Kriminal Indonesia 2019-2023*. Badan Pusat Statistik.
<https://www.bps.go.id>
- Destia, B., & Kartikasari, M. D. (2023). Comparison of fuzzy C-means and fuzzy Gustafson-Kessel clustering methods in provincial grouping in Indonesia based on criminality-related factors. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 17(2), 1093–1102.
<https://doi.org/10.30598/barekengvol17iss2pp1093-1102>
- Fitriyyah, N. R., & Pramana, S. (2022). Green spaces and crime: Spatial modeling of socioeconomic influences in Jakarta's urban areas. *Journal of Infrastructure and Sustainable Development Planning (JISDEP)*, 6(1), 1–15. <https://doi.org/10.46456/jisdep.v6i1.609>
- Jakarta Feminist Association. (2024). *Laporan kasus femisida di Indonesia*. Jakarta Feminist Association.
- Kompas. (2024). Jakarta Feminist Association: 184 hingga 187 kasus femisida sepanjang 2024. *Kompas*. <https://www.kompas.com>
- Krishnendu, S. G., Lakshmi, P. P., & Leeladharan, N. (2020, March). Crime analysis and prediction using optimized K-means algorithm. In *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 169–174). IEEE.
<https://doi.org/10.1109/ICCMC48092.2020.ICCMC-000169>
- Maharrani, R. H., Abda'u, P. D., & Faiz, M. N. (2024). Clustering method for criminal crime acts using K-means and principal component analysis. *International Journal of Electrical and Computer Engineering (IJECE)*, 34(1), 224–232.
<https://doi.org/10.11591/ijeecs.v34.i1.pp224-232>
- Retno, S., & Hakimi, M. (2025). Analysis of clustering results for crime incident data in Indonesia using fuzzy C-means. *Journal of Advanced Computer Knowledge and Algorithms*, 2(3), 22565–22572. <https://doi.org/10.29103/jacka.v2i3.22565>
- Rizkya, DD., & Fauziyah, M. (2025). Evaluasi Efek Normalisasi pada Algoritma Clustering untuk Data Multivariat. *Jurnal Teknologi dan Sistem Informasi*, 6(1), 45–53.
- Reuters. (2024). Indonesia shuts down over two million illegal gambling sites. *Reuters News*.
<https://www.reuters.com>
- Septriani, S. (2024). The impact of economic conditions on criminality in Indonesia. *European Journal of Development Studies*, 4(3), 345–356.
<https://doi.org/10.24018/ejdevelop.2024.4.3.345>
- Watrianthos, R., Suryadi, S., Kusmanto, & Samsir. (2023). Pemetaan tingkat kriminalitas di Indonesia: Analisis spasial dengan pendekatan SIG pada tingkat provinsi. *Bulletin of Information Technology (BIT)*, 4(3), 353–360. <https://doi.org/10.47065/bit.v4i3.861>