

Pendekatan Data Mining Dengan Algoritma K-Means Untuk Klasterisasi Faktor Perceraian Di Jawa Barat

Vina Andini¹, April Lia Hananto², Bayu Priyatna³, Agustia Hananto⁴

Prodi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Buana Perjuangan Karawang,
Karawang, Jawa Barat, Indonesia

Si22.vinaandini@mhs.ubpkarawang.ac.id¹, aprilialia@ubpkarawang.ac.id²,
bayu.priyatna@ubpkarawang.ac.id³, agustia.hananto@ubp.karawang.ac.id⁴

Abstrak: Penelitian ini dilatarbelakangi oleh tingginya angka perceraian di Provinsi Jawa Barat yang menunjukkan variasi faktor penyebab antarwilayah, sehingga diperlukan analisis kuantitatif berbasis data untuk mengidentifikasi pola pengelompokan determinannya. Tujuan penelitian ini adalah menganalisis konfigurasi kluster faktor penyebab perceraian serta mengidentifikasi faktor dominan pada masing-masing kelompok wilayah menggunakan pendekatan data mining. Metode yang digunakan adalah K-Means Clustering berbasis unsupervised learning terhadap data sekunder Open Data Jabar periode 2017–2024, dengan enam variabel utama yaitu ekonomi, KDRT, kawin paksa, zina, madat, dan cacat badan. Penentuan jumlah kluster dilakukan menggunakan *Elbow Method*, evaluasi model menggunakan *Silhouette Coefficient*, serta visualisasi pola dilakukan melalui *Principal Component Analysis* (PCA). Hasil penelitian menunjukkan terbentuknya tiga kluster dengan nilai *Silhouette* sebesar 0,61 yang mengindikasikan kualitas pemisahan kluster yang baik. Cluster pertama didominasi faktor ekonomi, cluster kedua menonjol pada faktor zina dan madat, sedangkan cluster ketiga menunjukkan kombinasi tekanan ekonomi, KDRT, dan kawin paksa. Temuan ini menegaskan bahwa perceraian di Jawa Barat dipengaruhi oleh pola determinan yang berbeda antarwilayah. Penelitian ini menyimpulkan bahwa pendekatan K-Means efektif dalam mengidentifikasi struktur laten faktor perceraian dan merekomendasikan kebijakan pencegahan yang disesuaikan dengan karakteristik kluster serta pengayaan variabel dan metode pada penelitian selanjutnya

Kata Kunci : K-Means, Klasterisasi, Faktor Perceraian,

Abstract: This study is motivated by the high divorce rate in West Java Province, which shows variation in causal factors across regions, thus requiring a quantitative data-driven approach to identify clustering patterns of its determinants. The objective of this study is to analyze the cluster configuration of divorce factors and to identify dominant variables in each regional group using a data mining approach. The method employed is K-Means Clustering based on unsupervised learning applied to secondary data from Open Data Jabar for the 2017–2024 period, focusing on six main variables: economic factors, domestic violence, forced marriage, adultery, drug abuse, and disability/chronic illness. The optimal number of clusters was determined using the Elbow Method, model evaluation was conducted using the Silhouette Coefficient, and cluster visualization was performed through Principal Component Analysis (PCA). The results indicate the formation of three clusters with a Silhouette value of 0.61, reflecting good cluster separation quality. The first cluster is dominated by economic factors, the second by adultery and drug abuse, while the third shows a combination of economic pressure, domestic violence, and forced marriage. The findings confirm that divorce patterns in West Java are shaped by varying determinants across regions. This study concludes that K-

Means effectively identifies latent structures of divorce factors and recommends cluster-based policy interventions and methodological enrichment in future research.

Keywords: *K-Means, Clustering, Divorce Factors,*

1. Pendahuluan

Perceraian merupakan gejala sosial yang menunjukkan kecenderungan perubahan dan menimbulkan konsekuensi signifikan terhadap kesejahteraan individu, ketahanan keluarga, serta tatanan masyarakat. Dalam konteks global, fenomena ini menjadi perhatian utama dalam kajian sosiologi dan demografi karena merefleksikan dinamika perubahan struktur sosial serta kondisi ekonomi keluarga (Khairuddin, 2024). Di Indonesia, angka perceraian menunjukkan pola yang fluktuatif dengan jumlah mencapai puluhan ribu kasus di setiap provinsi. Berdasarkan data yang dipublikasikan oleh Databoks Katadata, jumlah kasus perceraian di Jawa Barat mencapai sekitar 33,26 ribu kasus pada tahun 2024, yang menunjukkan tingginya dinamika permasalahan rumah tangga di wilayah tersebut (Agus Dwi Darmawan, 2026). Kondisi ini menegaskan urgensi penelitian untuk mengidentifikasi dan menganalisis faktor-faktor utama yang melatarbelakangi tingginya angka perceraian tersebut.

Sejumlah penelitian empiris di Indonesia telah mengungkap berbagai faktor penentu perceraian. Berdasarkan literatur permasalahan ekonomi, rendahnya kualitas komunikasi antar pasangan, serta campur tangan pihak ketiga diidentifikasi sebagai faktor dominan yang memicu perceraian di berbagai wilayah (Mujiono et al., 2025). Studi yang dilakukan di berbagai daerah juga menegaskan bahwa konflik domestik, kekerasan dalam rumah tangga, serta beragam masalah sosial lainnya merupakan faktor signifikan yang berkontribusi terhadap terjadinya perceraian (Suaeb et al., 2025). Selain itu tindakan perselingkuhan sering menjadi dasar utama gugatan perceraian di lingkungan Pengadilan Agama, karena pelanggaran terhadap komitmen perkawinan tersebut berdampak pada hilangnya kepercayaan serta terganggunya stabilitas dan keharmonisan rumah tangga (Achmad Husaini, 2024).

Kondisi disabilitas atau penyakit kronis yang timbul pascapernikahan juga teridentifikasi sebagai salah satu faktor penyebab perceraian, khususnya ketika pasangan tidak memiliki kesiapan emosional dan kemampuan sosial untuk menyesuaikan diri terhadap perubahan fisik yang terjadi (Mirsa Yainahu, 2025). Praktik pernikahan yang berlangsung tanpa persetujuan penuh dari kedua pihak berkontribusi pada meningkatnya potensi perceraian, khususnya pada pasangan berusia muda yang belum matang secara psikologis maupun ekonomi. Kondisi tersebut cenderung memicu konflik berkelanjutan dalam kehidupan rumah tangga (Izzah & Firdausi, 2021). Penyalahgunaan narkotika berkontribusi terhadap meningkatnya perceraian karena memicu tindak kekerasan, mengancam ekonomi, serta gangguan psikologis. Dampak tersebut secara simultan mencakup struktur dan fungsi keluarga secara komprehensif (Dwi Anggraeni et al., 2025). Hasil penelitian di Jawa Barat menunjukkan bahwa perceraian tidak dipicu oleh satu faktor tunggal, melainkan merupakan akumulasi berbagai permasalahan, seperti tekanan ekonomi, ketidakharmonisan dalam rumah tangga, serta rendahnya kematangan pasangan dalam mengelola dan menyelesaikan konflik (Hawa Rani et al., 2025).

Meskipun berbagai penelitian telah mengkaji determinan perceraian, sejumlah keterbatasan metodologi masih ditemukan dalam literatur terdahulu. Sebagian besar penelitian sebelumnya masih menggunakan pendekatan kualitatif atau deskriptif, sehingga

keterkaitan antar faktor penyebab perceraian belum dievaluasi secara komprehensif melalui analisis kuantitatif maupun penerapan metode *unsupervised learning* (Pasha, 2025). Studi kuantitatif berbasis *big data* yang fokus pada pengelompokan kluster faktor penyebab perceraian di Indonesia masih relatif minim. Oleh karena itu, penerapan pendekatan pembelajaran mesin diperlukan untuk mengidentifikasi determinan dominan secara lebih objektif berdasarkan data statistik (Rizkiyah et al., 2025).

Pendekatan kuantitatif yang mengintegrasikan teknik data mining, khususnya metode clustering, menjadi krusial karena mampu mengungkap pola determinan perceraian yang tidak teridentifikasi melalui analisis konvensional (Ofta Sari & Iqbal, 2025). Salah satu teknik yang relevan adalah K-Means, yang dapat mengelompokkan kombinasi faktor dominan berdasarkan data empiris antarwilayah di Jawa Barat secara lebih terukur dan objektif (Ferdiansyah et al., n.d.) Selain itu, metode ini memungkinkan analisis terhadap struktur kluster penyebab perceraian serta perbandingan karakteristik antar kelompok secara statistik. Dengan demikian, hasil penelitian dapat memperkuat landasan empiris dalam perumusan kebijakan keluarga berbasis bukti di masa mendatang.

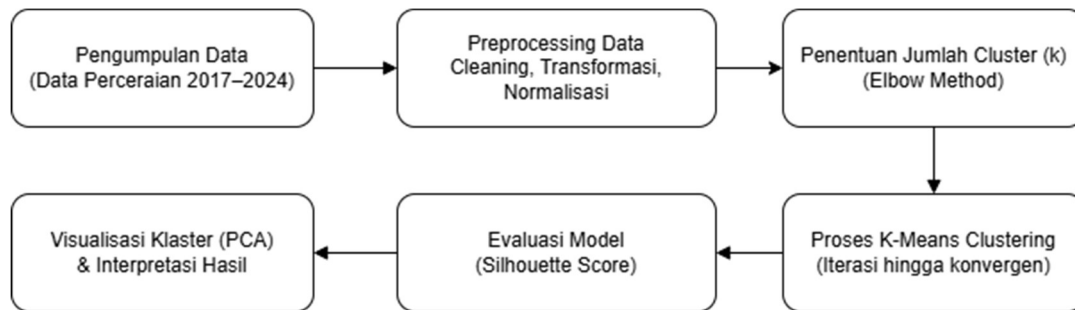
Meskipun demikian, sebagian besar penelitian sebelumnya masih menggunakan pendekatan deskriptif atau terbatas pada identifikasi faktor penyebab tanpa melakukan pengelompokan berbasis pola data secara komprehensif. Selain itu, penelitian yang mengintegrasikan metode K-Means dengan evaluasi kualitas kluster menggunakan *Silhouette Coefficient* serta visualisasi berbasis *Principal Component Analysis* (PCA) dalam konteks perceraian pada tingkat wilayah masih relatif terbatas. Oleh karena itu, penelitian ini mengintegrasikan metode K-Means Clustering, evaluasi menggunakan *Silhouette Coefficient*, serta visualisasi PCA untuk mengidentifikasi struktur laten faktor perceraian secara lebih sistematis pada tingkat kabupaten/kota di Provinsi Jawa Barat. Pendekatan ini tidak hanya mengelompokkan data, tetapi juga menghasilkan interpretasi berbasis pola yang lebih komprehensif dibandingkan penelitian sebelumnya.

Berdasarkan latar belakang tersebut, penelitian ini diarahkan untuk menganalisis pola pengelompokan faktor perceraian di Provinsi Jawa Barat melalui penerapan algoritma *K-Means*. Studi ini berfokus pada dua pertanyaan utama, yaitu: (1) bagaimana konfigurasi kluster faktor penyebab perceraian berdasarkan data statistik periode 2017–2024, dan (2) faktor dominan apa yang membentuk masing-masing kluster tersebut. Temuan yang dihasilkan diharapkan memperkaya bukti empiris dalam kajian perceraian di Indonesia, khususnya melalui pendekatan kuantitatif berbasis data mining, sekaligus menjadi rujukan strategis bagi perumusan kebijakan pencegahan perceraian yang lebih tepat sasaran dan berbasis data.

2. Metode Penelitian

Penelitian ini menerapkan desain kuantitatif dengan memanfaatkan teknik *data mining* berbasis *unsupervised learning*. Pendekatan tersebut dipilih untuk mengungkap pola serta struktur laten dalam data numerik terkait faktor penyebab perceraian tanpa menetapkan variabel dependen tertentu (Diningrat et al., 2025). Selanjutnya, metode klusterisasi digunakan untuk mengelompokkan wilayah berdasarkan tingkat kesamaan karakteristik determinan perceraian yang diamati. Algoritma yang diterapkan dalam penelitian ini adalah *K-Means Clustering*, yaitu metode klusterisasi non-hierarkis yang mengelompokkan data ke dalam sejumlah kluster berdasarkan tingkat kedekatan jarak antarobjek dalam ruang data (Teguh, 2025).

Data diperoleh dari portal resmi Open Data Jabar yang memuat statistik perceraian berdasarkan faktor penyebab di Provinsi Jawa Barat periode 2017–2024 (pembaruan November 2025). Unit analisisnya Wilayah Pengadilan Agama, dengan setiap observasi menunjukkan distribusi jumlah kasus perceraian per faktor. Dari 14 variabel yang tersedia, penelitian ini memfokuskan pada enam faktor utama, yaitu ekonomi, kekerasan dalam rumah tangga (KDRT), pernikahan paksa, zina, penyalahgunaan narkotika (madat), serta disabilitas atau penyakit kronis. Pemilihan variabel tersebut didasarkan pada dua pertimbangan utama. Pertama, secara teoritis, keenam variabel tersebut merupakan determinan yang paling sering diidentifikasi dalam berbagai penelitian sebelumnya sebagai faktor dominan penyebab perceraian, baik dari aspek ekonomi, sosial, maupun relasional. Kedua, dari sisi data, variabel-variabel tersebut memiliki tingkat kelengkapan dan konsistensi yang tinggi selama periode pengamatan 2017–2024, sehingga lebih representatif untuk dianalisis menggunakan metode klusterisasi. Sementara itu, variabel lainnya tidak digunakan karena memiliki keterbatasan data, seperti ketidakkonsistenan pencatatan atau frekuensi kemunculan yang rendah. Seluruh variabel berbentuk numerik (jumlah kasus), sehingga memungkinkan analisis klusterisasi dilakukan secara objektif melalui pengukuran jarak antarwilayah. Pengolahan data dilakukan secara sistematis melalui beberapa tahap. Tahapan penelitian yang dilakukan dalam studi ini disajikan pada Gambar 1.



Gambar 1 Tahapan pengolahan data

Pengolahan data dilakukan secara sistematis melalui tahapan pengumpulan data, *preprocessing* (cleaning, transformasi, dan normalisasi), penentuan jumlah kluster menggunakan metode Elbow, penerapan K-Means Clustering, evaluasi dengan *Silhouette Score*, serta visualisasi dan interpretasi hasil menggunakan PCA. Selanjutnya dilakukan standarisasi guna menyeimbangkan skala variabel agar tidak saling mendominasi dalam proses klusterisasi. Jumlah kluster optimal ditentukan menggunakan Elbow Method berdasarkan nilai WCSS, dengan memilih titik perubahan grafik yang paling signifikan (Rajsya et al., 2024). Setelah itu, algoritma K-Means dijalankan melalui inisialisasi centroid, perhitungan jarak Euclidean, pengelompokan data, dan pembaruan centroid secara iteratif hingga mencapai konvergensi.

K-Means Clustering dirancang untuk meminimalkan variasi di dalam kluster sekaligus memaksimalkan perbedaan antar kluster (Hafizah et al., 2025). Pengukuran kedekatan antarobjek dilakukan menggunakan jarak Euclidean yang dirumuskan sebagai berikut:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Proses pengelompokan berlangsung secara iteratif hingga tidak terdapat perubahan berarti pada posisi centroid maupun komposisi keanggotaan kluster, yang menandakan tercapainya kondisi konvergen (Iqbal & Nurul Huda, 2023).

Untuk mengukur kualitas hasil klusterisasi, digunakan Silhouette Coefficient. Metode ini mengevaluasi tingkat kesesuaian suatu objek terhadap cluster tempatnya berada dibandingkan dengan cluster lainnya (Agustia Hananto et al., 2024). Nilai Silhouette dihitung menggunakan rumus:

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Koefisien Silhouette berada pada interval -1 hingga 1 , di mana nilai yang semakin mendekati 1 merefleksikan kualitas pemisahan kluster yang semakin kuat (Khairani Ritonga & Hasibuan, n.d.). Selain pengujian kuantitatif, visualisasi dilakukan melalui *Principal Component Analysis* (PCA) untuk mereduksi dimensi data dan merepresentasikan sebaran kluster dalam ruang dua dimensi, sehingga pola pengelompokan dapat diamati secara lebih jelas dan interpretatif (Rahman et al., 2025)

3. Hasil dan Pembahasan

3.1 Dataset

Data yang digunakan dalam penelitian ini merupakan data sekunder yang bersumber dari portal resmi Open Data Jabar, yang memuat informasi perceraian berdasarkan faktor penyebab pada setiap wilayah Pengadilan Agama di Provinsi Jawa Barat. Dataset mencakup periode tahun 2017–2024, dengan proses pengambilan data terakhir dilakukan pada November 2025. Tabel berikut menyajikan cuplikan dataset mentah sebagai gambaran struktur data yang digunakan dalam penelitian:

Tabel 1 Cuplikan Dataset Perceraian Provinsi Jawa Barat (2017–2024)

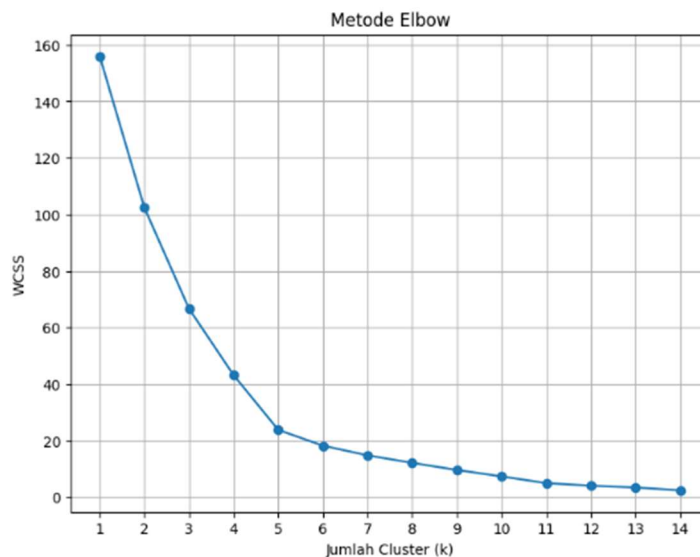
No	Wilayah Pengadilan Agama	Faktor Penyebab	Jumlah	Satuan	Tahun
1	BANDUNG	MABUK	11	KASUS	2017
2	BANDUNG	JUDI	1	KASUS	2017
3	BANDUNG	MENINGGALKAN SALAH SATU	504	KASUS	2017
4	BANDUNG	POLIGAMI	112	KASUS	2017
5	BANDUNG	KEKERASAN DALAM RUMAH TANGGA	39	KASUS	2017
...
858	TASIKMALAYA	EKONOMI	78	KASUS	2024

Secara keseluruhan, dataset terdiri dari 858 baris data dengan 14 variabel faktor penyebab perceraian. Namun, penelitian ini hanya menganalisis 6 variabel utama, yaitu

ekonomi, kekerasan dalam rumah tangga (KDRT), kawin paksa, zina, madat (penyalahgunaan narkoba), dan cacat badan. Pemilihan variabel dilakukan berdasarkan relevansi teoritis dan konsistensi ketersediaan data selama periode pengamatan. Data selanjutnya diproses melalui tahap normalisasi sebelum dilakukan klusterisasi menggunakan algoritma K-Means.

3.2 Penentuan Jumlah Kluster

Penentuan jumlah kluster optimal merupakan tahap penting dalam proses klusterisasi menggunakan algoritma K-Means, karena jumlah kluster yang tidak tepat dapat menghasilkan pengelompokan yang kurang representatif. Oleh karena itu, dalam penelitian ini digunakan metode Elbow untuk menentukan jumlah kluster yang paling optimal berdasarkan nilai Within Cluster Sum of Squares (WCSS). Metode ini bekerja dengan cara menghitung total variasi dalam setiap kluster dan mengamati pola penurunannya terhadap jumlah kluster yang berbeda.



Gambar 2 Elbow metode

Berdasarkan Gambar X, terlihat bahwa nilai WCSS mengalami penurunan yang signifikan dari $k = 1$ hingga $k = 3$. Setelah nilai tersebut, penurunan WCSS cenderung melandai, yang menunjukkan bahwa penambahan jumlah kluster tidak memberikan pengaruh yang signifikan terhadap pengurangan variasi dalam kluster. Titik perubahan yang menyerupai "siku" (elbow) terlihat pada $k = 3$, yang mengindikasikan bahwa jumlah kluster optimal dalam penelitian ini adalah sebanyak tiga kluster. Dengan demikian, nilai $k = 3$ digunakan dalam proses klusterisasi menggunakan algoritma K-Means pada tahap selanjutnya, karena dianggap mampu merepresentasikan struktur data secara efektif tanpa menambah kompleksitas model secara berlebihan.

3.3 Hasil Klusterisasi

Berdasarkan Tabel 2, terlihat bahwa distribusi cluster tidak merata, di mana sebagian besar wilayah tergolong dalam Cluster 1. Hal ini menunjukkan bahwa mayoritas kabupaten/kota di Provinsi Jawa Barat memiliki karakteristik perceraian yang relatif serupa, dengan dominasi faktor ekonomi sebagai penyebab utama.

Di sisi lain, terdapat beberapa wilayah yang menunjukkan karakteristik ekstrem atau berbeda dari pola umum. Indramayu merupakan wilayah dengan nilai faktor ekonomi tertinggi dibandingkan wilayah lainnya, yang mengindikasikan tingkat tekanan finansial yang sangat dominan dalam kasus perceraian. Selain itu, Depok juga menunjukkan nilai KDRT yang relatif tinggi dibandingkan wilayah lain dalam cluster yang sama, sehingga dapat dikategorikan sebagai wilayah dengan intensitas konflik domestik yang menonjol.

Sementara itu, Cirebon yang tergolong dalam Cluster 0 menunjukkan pola yang berbeda dengan dominasi pada variabel zina dan penyalahgunaan narkotika (madat), yang mencerminkan adanya pengaruh faktor sosial dan perilaku yang lebih kuat dibandingkan faktor ekonomi. Di sisi lain, Bandung yang termasuk dalam Cluster 2 memperlihatkan nilai yang tinggi pada beberapa variabel sekaligus, seperti ekonomi, KDRT, dan kawin paksa, sehingga dapat dikategorikan sebagai wilayah dengan karakteristik perceraian yang bersifat multidimensional.

Secara keseluruhan, pola distribusi ini menunjukkan bahwa meskipun sebagian besar wilayah memiliki karakteristik yang homogen, terdapat beberapa wilayah dengan kondisi yang lebih ekstrem atau unik yang membentuk cluster tersendiri. Hal ini menegaskan bahwa pendekatan klasterisasi mampu mengidentifikasi variasi pola perceraian yang tidak terlihat secara langsung dalam penyajian data mentah.

Tabel 2 Hasil Pengelompokan Wilayah Berdasarkan K-Means (k = 3)

Wilayah	Ekonomi	Kdrt	Kawin paksa	Zina	Madat	Cacat badan	Cluster
Bandung	2996	90	339	1	2	5	2
Bekasi	986	37	0	5	0	0	1
Bogor	381	15	2	1	1	0	1
Ciamis	7096	19	4	1	1	2	1
Cianjur	1689	5	0	0	0	1	1
Cibadak	316	13	0	2	3	1	1
Cibinong	2849	53	1	6	6	1	1
Cikarang	255	154	0	0	2	0	1
Cimahi	3527	79	0	5	1	0	1
Cirebon	255	64	2	121	68	4	0
Depok	1317	232	0	30	11	2	1
Garut	5839	49	0	0	0	0	1
Indramayu	10310	22	35	1	2	3	1
Karawang	2066	10	1	1	1	1	1
Kota banjar	422	22	1	0	0	0	1
Kota tasikmalaya	1220	15	1	1	1	0	1
Kuningan	3268	51	0	36	1	2	1
Majalengka	4852	19	2	0	0	2	1
Ngamprah	867	83	2	2	0	1	1
Purwakarta	1400	34	1	0	3	0	1
Soreang	31	2	0	0	0	1	1
Subang	2507	14	5	4	4	1	1
Sukabumi	109	15	1	1	0	0	1
Sumber	7538	13	1	1	1	1	1
Sumedang	4331	10	0	0	0	1	1
Tasikmalaya	5187	6	1	0	0	3	1

Untuk memahami karakteristik masing-masing cluster secara lebih mendalam, dilakukan perhitungan rata-rata nilai setiap variabel dalam tiap kelompok. Hasil perhitungan tersebut disajikan pada Tabel 3.

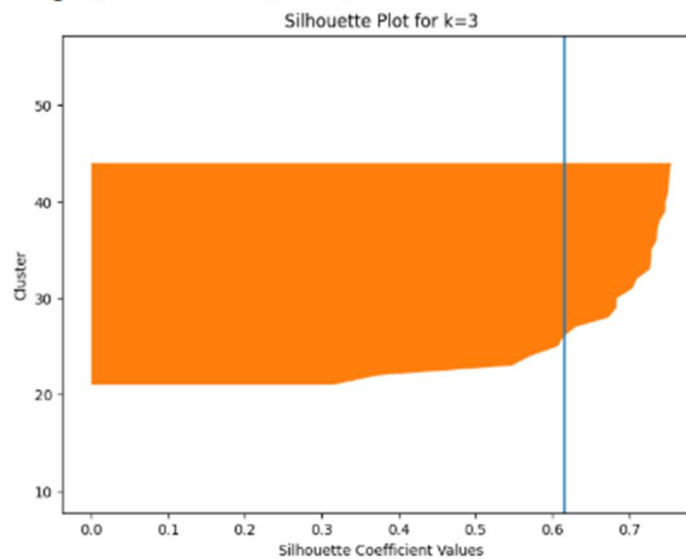
Tabel 3 Rata-rata Variabel Penyebab Perceraian per Cluster

Cluster	Ekonomi	KDRT	Kawin Paksa	Zina	Madat	Cacat Badan
0	255	64	2	121	68	4
1	2848.46	40.5	2.42	4.04	1.58	0.96
2	2996	90	339	1	2	5

Berdasarkan Tabel 3, Cluster 1 menunjukkan dominasi pada faktor ekonomi dengan nilai rata-rata yang jauh lebih tinggi dibandingkan variabel lainnya. Cluster 0 memiliki karakteristik relatif lebih menonjol pada variabel zina dan madat dibandingkan cluster lain. Sementara itu, Cluster 2 memperlihatkan kombinasi nilai yang tinggi pada faktor ekonomi dan KDRT serta nilai signifikan pada kawin paksa, yang menunjukkan adanya tekanan sosial-ekonomi yang lebih kompleks pada kelompok wilayah ini.

3.4 Evaluasi Model dan visualisasi

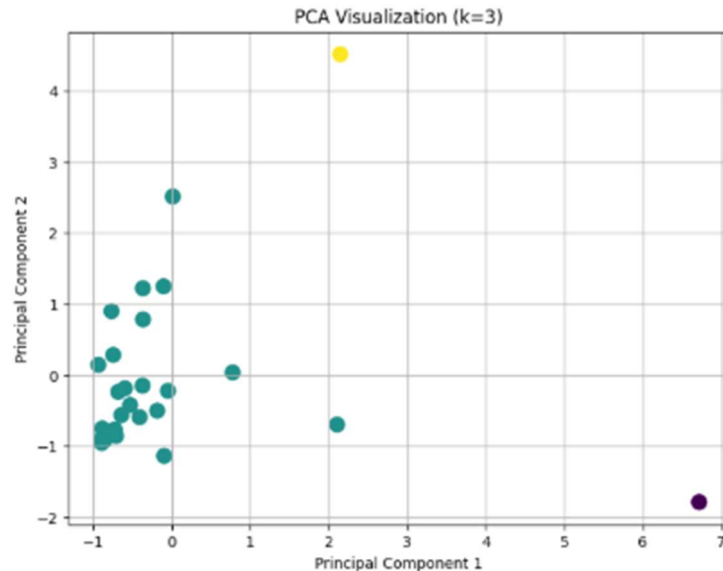
Evaluasi kualitas klasterisasi dilakukan menggunakan Silhouette Coefficient dengan jumlah cluster $k = 3$. Berdasarkan Gambar X, diperoleh nilai rata-rata Silhouette sebesar 0,61, yang ditunjukkan oleh garis vertikal pada grafik.



Gambar 3 Silhouette Plot (k = 3)

Sebagian besar nilai silhouette berada di atas 0,3 dan terkonsentrasi mendekati 0,6 hingga 0,7, yang menunjukkan bahwa mayoritas objek memiliki tingkat kedekatan yang baik dengan cluster tempatnya berada. Tidak terlihat dominasi nilai negatif yang signifikan, sehingga dapat disimpulkan bahwa struktur klaster yang terbentuk cukup stabil dan memiliki pemisahan yang jelas antar kelompok.

Untuk memperkuat evaluasi numerik tersebut, dilakukan visualisasi menggunakan Principal Component Analysis (PCA) dengan mereduksi data ke dalam dua komponen utama. Sebagaimana ditunjukkan pada Gambar 4, ketiga cluster membentuk pemisahan yang relatif jelas dalam ruang dua dimensi. Terlihat satu kelompok data yang terpisah cukup jauh pada sisi kanan grafik, serta satu kelompok lain yang berada pada posisi komponen kedua yang lebih tinggi. Sementara itu, sebagian besar wilayah terkonsentrasi pada area kiri bawah grafik. Pola ini menunjukkan adanya perbedaan karakteristik antar cluster yang cukup nyata dan konsisten dengan hasil evaluasi Silhouette.



Gambar 4 Visualisasi PCA (k = 3)

Secara keseluruhan, hasil evaluasi numerik dan visualisasi PCA menunjukkan bahwa model K-Means dengan tiga cluster mampu membentuk pengelompokan yang representatif dan layak untuk dianalisis lebih lanjut pada tahap pembahasan.

3.5 Interpretasi

Hasil klasterisasi menunjukkan bahwa perceraian di Provinsi Jawa Barat terbagi ke dalam tiga pola utama berdasarkan dominasi faktor penyebabnya. Cluster 1 didominasi oleh faktor ekonomi dengan nilai rata-rata yang jauh lebih tinggi dibandingkan variabel lainnya. Temuan ini mengindikasikan bahwa tekanan finansial menjadi determinan utama perceraian pada sebagian besar wilayah. Secara konseptual, ketidakstabilan ekonomi dapat meningkatkan konflik rumah tangga akibat ketidakseimbangan peran dan tekanan psikologis dalam keluarga, sehingga memperbesar risiko perceraian.

Cluster 0 memperlihatkan karakteristik yang lebih menonjol pada variabel zina dan penyalahgunaan narkotika (madat), yang menunjukkan adanya kontribusi faktor moral dan perilaku sosial terhadap terjadinya perceraian. Sementara itu, Cluster 2 menunjukkan kombinasi nilai tinggi pada faktor ekonomi, KDRT, serta pernikahan paksa. Pola ini mengindikasikan adanya tekanan multidimensional, di mana persoalan ekonomi beririsan dengan konflik domestik dan faktor struktural dalam pembentukan rumah tangga, sehingga memperlihatkan kompleksitas determinan perceraian pada kelompok wilayah tertentu.

Secara keseluruhan, temuan ini menegaskan bahwa perceraian di Jawa Barat tidak dipengaruhi oleh satu faktor tunggal, melainkan oleh kombinasi determinan sosial, ekonomi,

dan relasional yang berbeda antar wilayah. Pendekatan klusterisasi yang digunakan dalam penelitian ini mampu mengidentifikasi pola pengelompokan wilayah secara lebih sistematis dibandingkan analisis deskriptif biasa, sehingga memberikan pemahaman yang lebih komprehensif mengenai variasi faktor penyebab perceraian.

Hasil penelitian ini sejalan dengan beberapa penelitian sebelumnya yang menunjukkan bahwa faktor ekonomi merupakan determinan utama perceraian. Penelitian oleh Mujiono (2025) dan Suaeb (2025) juga mengungkap bahwa tekanan finansial menjadi pemicu dominan dalam konflik rumah tangga. Selain itu, temuan terkait faktor sosial seperti zina dan penyalahgunaan narkoba konsisten dengan penelitian (Achmad Husaini (2024) yang menekankan pentingnya aspek moral dan perilaku dalam stabilitas rumah tangga.

Di sisi lain, penelitian ini memberikan pendekatan yang berbeda dibandingkan studi sebelumnya, karena tidak hanya mengidentifikasi faktor penyebab secara deskriptif, tetapi juga mengelompokkan wilayah berdasarkan pola kombinasi faktor menggunakan metode K-Means Clustering. Pendekatan ini memungkinkan identifikasi struktur laten yang tidak terlihat dalam analisis konvensional, sehingga memberikan kontribusi yang lebih komprehensif dalam memahami dinamika perceraian berbasis data.

4. Kesimpulan dan Saran

Berdasarkan hasil analisis menggunakan algoritma K-Means terhadap data perceraian Provinsi Jawa Barat periode 2017–2024, penelitian ini berhasil mengidentifikasi tiga konfigurasi kluster utama yang merepresentasikan pola dominasi faktor penyebab perceraian pada tingkat wilayah Pengadilan Agama. Hasil evaluasi menunjukkan nilai Silhouette sebesar 0,61 yang mengindikasikan kualitas pemisahan kluster berada pada kategori baik, sehingga struktur pengelompokan dinilai stabil dan representatif. Sintesis antara permasalahan, tujuan, dan hasil penelitian memperlihatkan bahwa pertanyaan mengenai konfigurasi kluster dan faktor dominan telah terjawab secara empiris. Cluster 1 didominasi faktor ekonomi, Cluster 0 menonjol pada zina dan madat, sedangkan Cluster 2 menunjukkan kombinasi tekanan ekonomi, KDRT, dan kawin paksa. Keunggulan penelitian ini terletak pada penggunaan pendekatan kuantitatif berbasis data mining yang mampu mengungkap pola laten secara objektif, didukung evaluasi numerik dan visualisasi PCA. Namun demikian, keterbatasannya terletak pada penggunaan metode K-Means yang sensitif terhadap inisialisasi centroid dan asumsi bentuk kluster yang cenderung sferis, serta belum mempertimbangkan faktor eksternal lain seperti variabel demografis atau indeks kesejahteraan wilayah.

Implikasi penelitian ini menegaskan bahwa kebijakan pencegahan perceraian di Jawa Barat perlu dirancang secara diferensial sesuai karakteristik kluster wilayah. Wilayah dengan dominasi faktor ekonomi memerlukan intervensi berbasis penguatan ketahanan finansial keluarga, sementara wilayah dengan dominasi faktor moral dan perilaku sosial memerlukan pendekatan edukatif dan konseling keluarga. Adapun wilayah dengan tekanan multidimensional membutuhkan strategi terpadu lintas sektor, termasuk perlindungan terhadap kekerasan dalam rumah tangga dan peningkatan kesiapan pernikahan. Untuk penelitian selanjutnya, disarankan dilakukan pengayaan variabel, perbandingan dengan metode clustering lain seperti Hierarchical Clustering atau DBSCAN, serta integrasi analisis spasial agar diperoleh pemetaan yang lebih komprehensif dan mendalam. Dengan demikian, hasil penelitian tidak hanya relevan secara akademik, tetapi juga aplikatif dalam mendukung perumusan kebijakan keluarga berbasis data.

Daftar Pustaka

- Achmad Husaini. (2024). Dinamika Perceraian Karena Zina Dan Problematika Saksi: Studi Di Pengadilan Agama Kuala Tungkal. *Musawa: Journal for Gender Studies*, 11(1), 1–13. <https://doi.org/10.24239/msw.v11i1.439>
- Agus Dwi Darmawan. (2026). *Jumlah Perceraian di Jawa Barat Periode 2019-2024*. Databoks. <https://doi.org/https://databoks.katadata.co.id>
- Agustia Hananto, Tukino Tukino, & Elfina Novalia. (2024). Klasterisasi Kesiapan Digital Daerah. *Jurnal Ilmiah Teknik Informatika Dan Komunikasi*, 5(1), 28–41. <https://doi.org/10.55606/juitik.v5i1.979>
- Diningrat, C., Priyatna, B., Novalia, E., & Hilabi, S. S. (2025). Klasterisasi Kasus Kekerasan Berdasarkan Jenis Lokasi Kejadian di Jawa Barat Menggunakan Algoritma K-Means. *Jurnal Minfo Polgan*, 14(1), 518–528. <https://doi.org/10.33395/jmp.v14i1.14760>
- Dwi Anggraeni, N., Fadilah Dwi Winarni, N., Tosen, M., & Putri Meilinda, F. (2025). Analisis Faktor-Faktor Perceraian Tahun 2024 Di Pengadilan Agama Probolinggo (Perspektif Pengadilan Agama Probolinggo). *Al Fuadiy: Hukum Keluarga Islam*, 7(01). <https://doi.org/10.55606/af.v6i2>
- Ferdiansyah, I., Huda, B., Hananto, A., & Buana Perjuangan Karawang, U. (n.d.). *Analisis Clustering Menggunakan Metode K-Means Pada Kemiskinan Di Jawa Timur Tahun 2020*. <https://doi.org/https://doi.org/10.31004/innovative.v4i3.10360>
- Hafizah, N., Lia Hananto, A., Nurapriani, F., Novalia Sistem Informasi, E., Buana Perjuangan karawang Jalan Ronggo Waluyo, U. H., Timur, T., & Karawang, K. (2025). Segmentasi Nasabah Umkm Berdasarkan Kinerja Dan Keuntungan Menggunakan K-Means Clustering. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 9, Number 5). <https://doi.org/https://doi.org/10.36040/jati.v9i5.15056>
- Hawa Rani, A., Fernando, F., Bachrudin, M., & Novelya Herwin, Y. (2025). Analisis Faktor-Faktor Penyebab Kasus Cerai Gugat Di Indonesia. *Sriwijaya Journal of Private Law* ■, 2(2). <https://doi.org/10.28946/sjpl.v2i2.3369>
- Iqbal, M., & Nurul Huda, M. (2023). *Implementasi Algoritma K-Means Clustering dengan Jarak Euclidean dalam Mengelompokkan Daerah Penyebaran COVID-19 di Kabupaten Bogor*. 2(1), 47–56. <http://jurnal.fmipa.unmul.ac.id/index.php/basis>
- Izzah, H., & Firdausi, atul. (2021). Faktor dan Dampak Nikah Paksa Terhadap Putusnya Pernikahan Menurut Kompilasi Hukum Islam. In *The Indonesian Journal of Islamic Law and Civil Law* (Vol. 2, Number 2). <https://doi.org/https://doi.org/10.51675/jaksya.v2i1.174>
- Khairani Ritonga, P., & Hasibuan, M. S. (n.d.). *Analisis Perbandingan Silhouette dengan Elbow pada Algoritma K-Means dan DBSCAN*. 9, 2025. <https://doi.org/10.47002/metik.v9i1.1027>
- Khairuddin. (2024). Fenomena Tren Perceraian di Indonesia: Apa Penyebabnya? In *Abdurrauf Science and Society* (Vol. 1, Number 1). <https://doi.org/https://doi.org/10.70742/asoc.v1i1.95>
- Mirsa Yainahu. (2025). Analisis Perbandingan Undang-Undang Nomor 1 Tahun 1974 Tentang Perkawinan dan Kompilasi Hukum Islam dalam Perspektif Cacat Badan sebagai Alasan Perceraian. *Jurnal Ilmiah Wahana Pendidikan*, 11(B), 117–126. <https://doi.org/https://jurnal.peneliti.net/index.php/JIWP/article/view/11634>

- Mujiono, M., Romlan, R., & Bahri, S. (2025). Analisis yuridis faktor-faktor penyebab tingginya angka perceraian di kabupaten jombang. *Justicia Journal*, 14(1), 135–157. <https://doi.org/10.32492/jj.v14i1.14109>
- Ofta Sari, A., & Iqbal, M. (2025). Analisis Data Mining Terhadap Data Faktor Perceraian Di Sumatera Utara Dengan Metode K-Means Clustering. 4(1), 214–221. <https://ojs.trigunadharma.ac.id/index.php/jsi>
- Pasha, K. (2025). Prediksi Angka Perceraian Menggunakan Machine Learning. 11(1). <https://journal.fkom.uniku.ac.id/index.php/buffer>
- Rahman, N. F., Priyatna, B., Nurapriani, F., Paryono, T., Informasi, S., Buana, U., Karawang, P., Ronggo Waluyo, J. H., & Timur, T. (2025). Penerapan Algoritma K-Means Untuk Klasterisasi Film Pada Platform Amazon Prime Video. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 9, Number 5). <https://doi.org/https://doi.org/10.36040/jati.v9i5.15065>
- Rajsya, A., Purnawansyah, P., & Manga', A. R. (2024). Rancang Bangun Penerapan Metode Elbow Pada K-Means Untuk Clustering Data Persediaan Barang. *LINIER: Literatur Informatika Dan Komputer*, 1(4), 395–403. <https://doi.org/10.33096/linier.v1i4.2539>
- Rizkiyah, S., Zein Rizqin, I., Akbarany Bakhtiar Putri, M., Shella May Wara, S., Syaifullah Jauharis Saputra, W., Pembangunan Nasional, U., Timur, J., Rungkut Madya, J., Anyar, G., & Gn Anyar, K. (2025). Segmentasi Faktor Perceraian berdasarkan Provinsi di Indonesia Tahun 2024 dengan K-Means dan DBSCAN. In *Indonesian Journal on Data Science* (Vol. 3, Number 2). <https://doi.org/https://doi.org/10.30989/ijds.v3i2.1654>
- Suaeb, Lalu Supriadi Bin Mujib, & Khairul Hamim. (2025). Analysis of the Causes of Divorce Suits From A Gender Perspective in the Religious Court of Selong, East Lombok. *AL-AFKAR:Journal for Islamic Studies*. <https://doi.org/https://doi.org/10.31943/afkarjournal.v8i3.2386>
- Teguh, H. T. S. (2025). Implementasi Data Mining Clustering Dalam Mengelompokan Kasus Perceraian Yang Terjadi Di Provinsi Jawa Timur Menggunakan Algoritma K-Means. *JAMI: Jurnal Ahli Muda Indonesia*, 6(1), 68–83. <https://doi.org/10.46510/jami.v6i1.324>