

# Implementasi Data Mining Clustering Dalam Mengelompokkan Kasus Perceraian Yang Terjadi Di Provinsi Jawa Barat Menggunakan Algoritma K-Means

Ihsan Ahmad Fauzi  
STMIK IKMI Cirebon

Raditya Danar Dana  
STMIK IKMI Cirebon

Alamat: Jl. Perjuangan No. 10 B Majasem. Kota/Kabupaten, Kec. Kesambi. Kode Pos, 45131

Korespondensi penulis: [ihsanahmad.fauzi@gmail.com](mailto:ihsanahmad.fauzi@gmail.com)

**Abstract.** Divorce cases in West Java Province have increased every year. The profile map of divorce cases that occurred in each region is not yet known, so efforts to provide guidance to minimize divorce cases are not optimal. Divorce case data is also not equipped with a visualization feature that makes it easier for authorized officials to easily understand and analyze data. This study analyzes divorce cases in regencies/cities in West Java Province, using the K-Means Algorithm Clustering data mining method. The clustering method is grouping data based on the same characteristics. In determining the number of clusters, that is by using the value of the Davies Bouldin Index. The results of this study obtained the best cluster of grouping divorce cases, there were 2 clusters, namely cluster 0, there were 5 regencies and 9 cities, while in cluster 1 there were 13 regencies, with a Davies Bouldin index value of 0.168 and an avg.within centroid distance value of 5,870. Cluster 0 is the city/district with the lowest divorce cases and cluster 1 is the city/district with the highest divorce cases.

**Keywords:** Data Mining, Clustering, K-Means, Davies Bouldin Index.

**Abstrak.** Kasus perceraian di Provinsi Jawa Barat mengalami peningkatan setiap tahunnya. Peta profil kasus perceraian yang terjadi di setiap daerah belum diketahui, sehingga upaya memberikan pembinaan guna meminimalisir kasus perceraian belum optimal. Data kasus perceraian belum dilengkapi dengan fitur visualisasi yang memudahkan aparat yang berwenang mudah dalam memahami dan menganalisa data. Penelitian ini menganalisa kasus perceraian di Kabupaten/Kota yang berada di Provinsi Jawa Barat, analisa yang dilakukan yaitu dengan menggunakan data mining metode Clustering dengan menggunakan algoritma K-Means. Metode clustering yaitu mengelompokkan data berdasarkan karakteristik yang sama. Dalam menentukan jumlah cluster yaitu dengan menggunakan nilai dari Davies Bouldin Index. Hasil dari penelitian ini diperoleh Cluster terbaik dari pengelompokan kasus perceraian terdapat 2 cluster yaitu cluster 0 terdapat 5 kabupaten dan 9 kota sedangkan pada cluster 1 terdapat 13 kabupaten, dengan nilai davies bouldin index 0,168 dan nilai avg.within centroid distance 5.870. Cluster 0 merupakan kota/kabupaten dengan kasus perceraian rendah dan untuk cluster 1 merupakan kota/kabupaten dengan kasus perceraian tertinggi.

**Kata kunci:** Data Mining, Clustering, K-Means, Davies Bouldin Index.

## LATAR BELAKANG

Berdasarkan data dari Badan Pusat Statistik (BPS) jumlah kasus perceraian di Indonesia cenderung meningkat setiap tahunnya. Peningkatan angka perceraian juga terjadi di Provinsi Jawa Barat. Dengan tingginya angka perceraian yang ada perlu adanya penelitian guna mengelompokkan tingkat perceraian berdasarkan masing-masing kabupaten/kota di Provinsi Jawa Barat.

**Tabel 1. Data Jumlah Kasus Perceraian**

No	Nama Kabupaten/ Kota	Data Kasus Perceraian							
		2014	2015	2016	2017	2018	2019	2020	2021
1.	Kabupaten Bogor	43765	45749	49957	52381	55470	58908	59081	63731
2.	Kabupaten Sukabumi	50013	46967	53516	54935	56839	59599	58727	61537
3.	Kabupaten Cianjur	46250	46323	45951	43386	42884	44139	43146	45035
4.	Kabupaten Bandung	54765	47701	48751	48188	46750	48532	49598	54154
5.	Kabupaten Garut	25562	24948	28635	30362	32144	35641	38485	42941
6.	Kabupaten Tasikmalaya	35086	34526	36157	35975	37752	38856	36929	40634
7.	Kabupaten Ciamis	16712	17234	18887	19528	19689	22067	23057	24939
8.	Kabupaten Kuningan	13295	13667	14450	15262	16412	17739	18871	21052
9.	Kabupaten Cirebon	24745	26541	28940	30743	37635	42807	45653	51020
10.	Kabupaten Majalengka	22064	22711	23874	24636	27783	31702	30860	36262
11.	Kabupaten Sumedang	17042	17216	18892	21682	21777	23091	23792	25768
12.	Kabupaten Indramayu	22099	24802	23750	25616	27902	33688	37998	42887
13.	Kabupaten Subang	29282	28349	25855	24926	24352	25259	24619	27469
14.	Kabupaten Purwakarta	12806	14365	17161	18734	19252	19521	19408	20399
15.	Kabupaten Karawang	24913	27453	32035	33329	37079	40136	42186	47499
....	....	....	....	....	....	....	....	....	....
27.	Kota Banjar	4199	4800	5778	5828	6405	6829	7195	7697
<b>Jumlah Kasus</b>		<b>600.364</b>	<b>597.599</b>	<b>641.394</b>	<b>667.069</b>	<b>701.525</b>	<b>760.043</b>	<b>783.741</b>	<b>862.608</b>

Data kasus perceraian pada tahun 2014 menuju ke tahun 2015 menunjukkan angka penurunan dimana jumlah kasus awal yang mencapai 600.364 kasus turun menjadi 597.599 kasus, namun di tahun-tahun berikutnya data kasus perceraian di Provinsi Jawa Barat cenderung meningkat dimana pada tahun 2016 mencapai 641.394 kasus, tahun 2017 mencapai 667.069 kasus, tahun 2018 mencapai 701.525 kasus, tahun 2019 mencapai 760.043 kasus, tahun 2020 mencapai 783.741 kasus, dan tahun 2021 mencapai 862.608 kasus.

Menikah pada hakikatnya merupakan konsekuensi dari kebutuhan manusia sebagai makhluk sosial yang diciptakan berbeda jenis kelamin oleh Tuhan untuk saling melengkapi satu sama lain. Dengan adanya pernikahan hal yang diharapkan oleh setiap pasangan dalam membangun rumah tangga adalah dapat hidup bahagia bersama selamanya. Sehingga dalam menjalani rumah tangga tidak ada satupun pasangan yang mengharapkan perselisihan hingga memicu putusnya status pernikahan apalagi sampai berakhir pada sebuah perceraian (Elvi, 2021). Dalam sebuah pernikahan pertengkaran dan masalah merupakan hal yang lumrah muncul. Setiap pasangan pun dituntut untuk saling bekerjasama untuk menghadapi persoalan yang ada. Walaupun perceraian merupakan suatu hal yang tidak diharapkan timbul dalam sebuah hubungan, tetapi bagi pasangan yang merasa sudah tidak adanya kebahagiaan yang hadir hingga antar pasangan tidak sanggup lagi menjalani hubungan rumah tangga maka akan menganggap bahwa perceraian adalah keputusan terbaik yang bisa diambil.

Dari beberapa penelitian yang sudah dilakukan dapat disimpulkan bahwa dalam pengelompokan kasus perceraian dapat diterapkan pada kasus lain dengan atribut yang berbeda sehingga penelitian ini akan dilakukan teknik pengelompokan data perceraian berdasarkan kabupaten/kota dengan atribut yang digunakan yaitu nama provinsi, nama kabupaten/kota, status perkawinan, jumlah penduduk, satuan, dan tahun dengan metode yang digunakan yaitu teknik data mining clustering. Teknik data mining clustering merupakan sebuah teknik data yang dapat mengelompokkan berbagai jenis data berdasarkan karakteristiknya masing-masing sehingga kelompok yang nantinya terbentuk merupakan sebuah data dengan karakteristik yang sama (Triyansyah & Fitriana, 2018). Metode yang digunakan dalam penelitian ini menggunakan algoritma K-Means dimana metode ini dapat mengelompokkan data berdasarkan titik pusat cluster atau centroid terdekat dengan data (Azwanti, 2018).

Peta profil kasus perceraian yang terjadi di masing-masing daerah di Provinsi Jawa Barat masih belum diketahui, sehingga upaya pembinaan guna meminimalisir terjadinya kasus perceraian belum berjalan dengan optimal. Kemudian data kasus perceraian yang terjadi di Provinsi Jawa Barat belum dilengkapi dengan fitur visualisasi sehingga aparatur yang berwenang masih sulit dalam memahami dan menganalisa data.

Tujuan dari penelitian ini adalah untuk mencari cluster terbaik kasus perceraian yang terjadi di masing-masing kabupaten/kota di Provinsi Jawa Barat dimana data tersebut dapat divisualisasikan menjadi peta profil kasus perceraian sehingga memudahkan bagi aparat yang berwenang untuk memahami dan dapat membina kasus perceraian yang terjadi di masing-masing daerah. Dengan hal ini harapannya adalah kasus perceraian yang ada dapat diminimalisir.

## **KAJIAN TEORITIS**

Penelitian tentang kasus perceraian yang terjadi di masyarakat juga telah dilakukan oleh beberapa penelitian sebelumnya, misalnya penelitian yang telah dilakukan oleh Uci Dwi Rahayu dkk dalam sebuah artikel yang berjudul Analisis Kasus Perceraian Pada Pengadilan Negeri Bekasi Menggunakan Algoritma K-Means Clustering yang dipublikasikan pada jurnal IKRAITH-INFORMATIKA (2022). Jurnal ini memfokuskan pada permasalahan data-data perceraian yang masih belum terakurasi dengan baik sehingga dilakukan clustering menggunakan algoritma K-Mean supaya menghasilkan olahan data yang lebih komprehensif terhadap kecenderungan kasus perceraian. Atribut yang digunakan dalam proses data yaitu perbedaan usia pasangan, lama usia pernikahan, jenis cerai, jumlah anak, dan lokasi dengan

hasil yang diperoleh dari penelitian tersebut yaitu terdapat 2 cluster dimana terdapat perbedaan yang signifikan pada cluster 0 dengan cluster 1 yaitu pada lama usia pernikahan (Dwi Rahayu et al., 2022).

Penelitian lain terkait pengelompokan kasus perceraian juga telah dilakukan oleh N. Nurhayati dkk dengan judul Analisis Metode Klastering Pada Kasus Penyebab Perceraian Berdasarkan Provinsi Dengan Teknik K-Means pada jurnal KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer) tahun 2020 menjelaskan bahwa keputusan bercerai dapat disebabkan oleh sifat, perilaku, kepribadian, dan latar belakang yang berbeda-beda. Dari hal tersebut dapat dikelompokkan dengan menggunakan teknik ilmu komputer data mining sehingga hasil yang didapat nantinya bisa menjadi informasi bagi pemerintah dalam menangani kasus perceraian dan menyusun program guna mengurangi tingkat perceraian sesuai dengan masing-masing provinsi. Terdapat 4 variabel yang digunakan yaitu adanya perselisihan dan pertengkaran terus-menerus, masalah ekonomi, meninggalkan salah satu pihak, dan KDRT sehingga diperoleh 2 cluster yaitu cluster tingkat tinggi mencakup 3 provinsi yakni Jawa Barat, Jawa Tengah, Jawa Timur dan cluster tingkat rendah mencakup 26 provinsi yang tidak terdapat dalam cluster tingkat tinggi (Nurhayati et al., 2020).

## **METODE PENELITIAN**

### **1. Clustering**

Cluster merupakan kumpulan dari beberapa data yang tergabung karena memiliki sebuah kemiripan antar sebagian data dengan kelompok lain yang memiliki beberapa perbedaan. Berbeda dengan klasifikasi, pengelompokan mempunyai sebuah variabel objek dalam pengelompokannya. Pengelompokan tidak melakukan tugas klasifikasi seperti memprediksi angka sebuah variabel objek, namun algoritma clustering berusaha membagi kumpulan informasi yang ada ke masing-masing cluster yang seragam/sejenis. Dimana kesamaan data dalam sebuah cluster akan menghasilkan nilai yang semakin besar, sedangkan perbedaan cluster akan menghasilkan nilai yang semakin kecil (Anggreini Novita Lestari, 2019).

### **2. Rapid Miner**

RapidMiner merupakan perangkat lunak yang bersifat terbuka (open source). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. RapidMiner merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. RapidMiner ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi.

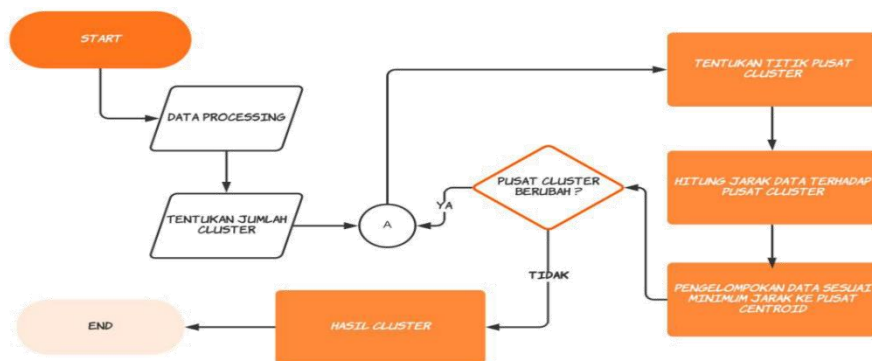
RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik.

### 3. Data Mining

Data mining atau yang umum disebut knowledge discovery in database (KDD) adalah kegiatan yang mencakup pengumpulan, penggunaan riwayat data untuk menjumpai kesesuaian, desain keterikatan dalam sebuah himpunan data yang berukuran besar. Data Mining menghasilkan output yang dapat digunakan untuk menentukan pilihan di masa yang akan datang. Teknik yang dikenal dalahsaturday adalah pengelompokan/clustering.

### 4. Algoritma K-Mean

Teknik clustering yang marak digunakan adalah Algoritma K-Mean karena mudah dalam pengoperasiannya. Algoritma K-Mean tergolong lebih cepat jika dibandingkan dengan pendekatan hierarkis, karena itu teknik clustering alternatif ini sering digunakan untuk pengolahan informasi dalam skala yang besar (Disi & Sumargo, 2019). K-Mean adalah satu dari beberapa metode cluster atau algoritma non hirarki yang berupaya memisahkan data objek yang tersedia kedalam satu jenis atau beberapa kelompok atau cluster objek menurut keunikannya dimana dalam menentukan titik terdekatnya dilakukan proses secara berulang-ulang (Mulyani et al., 2019).



**Gambar 1. Algoritma K-Means**

Algoritma K-Mean mengelompokkan seluruh titik data pada  $X$  kemudian masing-masing titik  $x_i$  terbagi dalam satu cluster  $K$ . Tujuan dari clustering adalah untuk meminimalkan peran sebuah entitas tertentu dalam metode clustering sehingga dapat ditemukan sebuah cara meminimumkan kejadian sembarang dalam suatu group untuk mengoptimalkan data sembarang antar kelompok (Parlina et al., 2018). Adapun langkah-langkah yang dilakukan dalam pemrosesan algoritma K-Mean sebagai berikut :

- a) Tentukan cluster secara sembarang sebagai centroid (titik sentral).
- b) Hitung menggunakan rumus Euclidian Distance untuk menentukan jarak data yang paling mendekati centroid. Rumus Euclidian Distance yaitu :

$$d(i, j) = \sqrt{\sum_j^m (C_{IJ} - C_{kj})^2}$$

keterangan :

$x_i$  = data kriteria

$\mu_j$  = Centroid Pada Cluster ke-j

- c) Kumpulkan dari masing-masing data pada cluster dengan jarak terpendek.

$$\text{Min} \sum_{k=1}^k dik = \sqrt{\sum_j^m (C_{IJ} - C_{kj})^2}$$

- d) Update nilai centroid data dengan cara menghitung nilai rata-rata cluster.

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p}$$

- e) Ulangi langkah 1, 2, 3 hingga seluruh anggota cluster hingga tidak mengalami perubahan.
- f) Tentukan nilai sentral cluster terakhir yang mana hal ini dapat memastikan kesesuaian data dengan sayarat pada langkah 1, 2, dan 3 telah dikerjakan dan memmperoleh nilai yang sesuai.

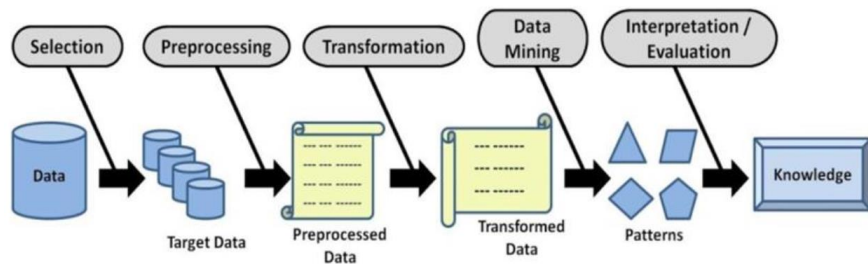
## 5. Davies Bouldin Index

Prinsip pendekatan pengukuran DBI adalah memaksimalkan jarak titik pust cluster dan meminimalkan jarak titik pusat cluster. Semakin kecil angka yang diperoleh dari kalkulasi nilai DBI menghasilkan skema cluster yang paling optimal. Adapun rumus DBI sebagai berikut :

$$DBI = \frac{1}{M} \sum_{j=1}^M \max_{j \neq k} R_{jk}$$

## 6. Tahap Perancangan

Knowledge Discovery in Databases (KDD) adalah sebuah pengelolaan yang dilakukan untuk mendapatkan berita/fakta yang bermanfaat yang berasal dari database. Yang mana dalam tahapan pengelolaannya meliputi interpretasi pada bidang aplikasi, membuat data acuan yang ditetapkan berasal dari data asli (data yang belum dilakukan proses pengolahan) yang berasal dari database, dan juga preprocessing data dan cleaning data (Aprianti & Permadi, 2018).



**Gambar 2. Tahapan Knowledge of Database**

Proses KDD terdiri dari beberapa tahapan sebagai berikut :

- a) Data Selection. Pada tahap ini hal yang dilakukan adalah menentukan kelompok data, membuat kelompok data tujuan, atau menitik beratkan pada subset variabel (contoh data) yang mana discovery (penemuan) yang akan dikerjakan. Data yang diperoleh selanjutnya di pilah dan disimpan pada sebuah berkas yang terpisah dari basis data operasional. Pada penelitian ini atribut yang digunakan pada data kasus perceraian sebanyak 5 atribut dan 484 jumlah data yang diambil dari data kasus perceraian pada tahun 2013 sampai tahun 2021. Data yang dipakai dalam penelitian ini diambil dari situs web Open Data Jabar, yang disajikan pada tabel dibawah ini :

**Tabel 2. Data Penelitian**

No	Nama Kabupaten/Kota	Status Perceraian	Jumlah	Tahun
1	Kabupaten Bogor	Cerai Hidup	61155	2013
2	Kabupaten Bogor	Cerai Mati	99447	2013
3	Kabupaten Sukabumi	Cerai Hidup	55780	2013
4	Kabupaten Sukabumi	Cerai Mati	79143	2013
5	Kabupaten Cianjur	Cerai Hidup	51948	2013
6	Kabupaten Cianjur	Cerai Mati	65796	2013
7	Kabupaten Bandung	Cerai Hidup	58085	2013
8	Kabupaten Bandung	Cerai Mati	70626	2013
9	Kabupaten Garut	Cerai Hidup	28264	2013
10	Kabupaten Garut	Cerai Mati	62229	2013
11	Kabupaten Tasikmalaya	Cerai Hidup	40739	2013
12	Kabupaten Tasikmalaya	Cerai Mati	48594	2013
13	Kabupaten Ciamis	Cerai Hidup	23951	2013
14	Kabupaten Ciamis	Cerai Mati	73735	2013
15	Kabupaten Kuningan	Cerai Hidup	14899	2013
...	...	...	...	...
484	Kota Banjar	Cerai Mati	11719	2021

- b) Pre-processing dan Cleaning Data. Pada tahap preprocessing dan cleaning data yang dilakukan adalah melepaskan data yang tidak tetap dan noise, data yang ganda atau terduplikasi, membetulkan data yang salah, dan menambah data yang saling berhubungan.

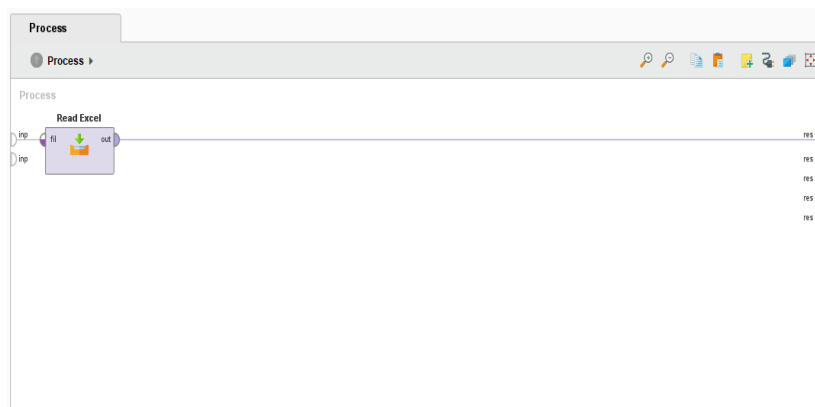
- c) Transformation. Tahap ini menyatukan data pada sesuatu yang lebih tepat guna mengerjakan mining dengan cara diringkaskan (agregasi).
- d) Data Mining. Tahap data mining merupakan tahap yang dilakukan untuk menentukan model atau info unik pada data yang ditentukan memakai metode atau teknik algoritma yang selaras dengan tujuan proses KDD secara menyeluruh.
- e) Evaluasi/Interpretation. Tahap ini memiliki tujuan untuk mengetahui model-model yang diperoleh. Evaluasi (menyelidiki) mungkinkah pola yang nanti dihasilkan akan selaras dengan hipotesis penelitian sebelumnya. Knowledge yang dihasilkan dari model-model tersebut akan ditampilkan dalam bentuk visualisasi.

## **HASIL DAN PEMBAHASAN**

### **1. Hasil Preprocessing dan Transformasi Data**

#### **a) Data Selection**

Proses yang dilakukan pertama kali adalah data selection, dalam tahap ini hal yang dilakukan adalah mengimport data menggunakan operator read excel. Operator read excel ini berfungsi untuk membaca data kasus perceraian yang terjadi di Jawa Barat dari tahun 2013 sampai 2021.



**Gambar 3. Operator Read Excel**

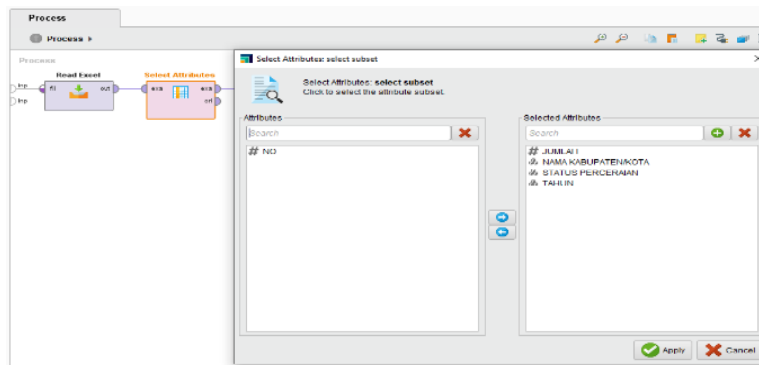
Ketika data sudah berhasil di import menggunakan operator read excel, selanjutnya akan tampil preview data. Preview data akan menampilkan data yang akan digunakan dalam proses data mining seperti pada Gambar 4.



NO	NAMA KABUPATEN/K...	STATUS PERCERAIAN	JUMLAH	TAHUN
1	KABUPATEN BOGOR	CERAI HIDUP	61155	2013
2	KABUPATEN BOGOR	CERAI MATI	99447	2013
3	KABUPATEN SUKABUMI	CERAI HIDUP	55780	2013
4	KABUPATEN SUKABUMI	CERAI MATI	79143	2013
5	KABUPATEN CIANJUR	CERAI HIDUP	51948	2013
6	KABUPATEN CIANJUR	CERAI MATI	65796	2013
7	KABUPATEN BANDUNG	CERAI HIDUP	58065	2013
8	KABUPATEN BANDUNG	CERAI MATI	70626	2013
9	KABUPATEN GARUT	CERAI HIDUP	28264	2013
10	KABUPATEN GARUT	CERAI MATI	62229	2013
11	KABUPATEN TASIKMALAYA	CERAI HIDUP	40739	2013
12	KABUPATEN TASIKMALAYA	CERAI MATI	48594	2013
13	KABUPATEN CIAMIS	CERAI HIDUP	23951	2013
14	KABUPATEN CIAMIS	CERAI MATI	73735	2013
15	KABUPATEN KUNINGAN	CERAI HIDUP	14899	2013
16	KABUPATEN KUNINGAN	CERAI MATI	49061	2013
17	KABUPATEN CIREBON	CERAI HIDUP	29503	2013
18	KABUPATEN CIREBON	CERAI MATI	74673	2013
19	KABUPATEN MAJALENGA	CERAI HIDUP	24816	2013

Gambar 4. Hasil Proses Data read Excel

Selanjutnya masukkan operator select atribut dimana operator ini berfungsi untuk menentukan atribut apa saja yang akan digunakan dalam proses data mining. Dalam proses ini atribut yang dipilih adalah nama kabupaten, status perceraian, jumlah dan tahun.



Gambar 5. Parameter Select Atribut

**b) Data Cleaning**

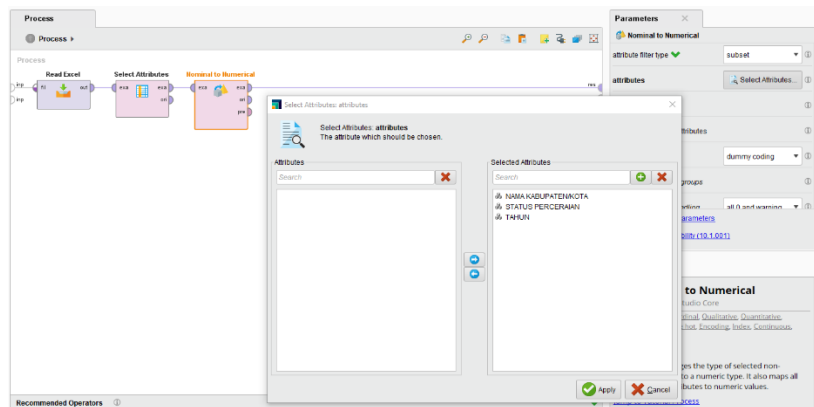
Pada penelitian ini data yang digunakan merupakan data yang relevan dan konsisten. Setelah dilakukan pengecekan pada hasil statistik data, terlihat bahwa dalam data penelitian ini tidak terdapat missing value, maka dari itu proses data cleaning tidak dilakukan.

Name	Type	Missing	Statistics
NO	Integer	0	Min: 1, Max: 484, Average: 242.500
NAMA KABUPATEN/KOTA	Nominal	0	Least: KOTA BANJAR (16), Most: KABUPATEN BANDUNG (18), Unique: KABUPATEN BANDUNG (18), KABUPATEN BANDUNG BARAT (16), ... [25 more]
STATUS PERCERAIAN	Nominal	0	Least: CERAI MATI (242), Most: CERAI HIDUP (242), Unique: CERAI HIDUP (242), CERAI MATI (242)
JUMLAH	Integer	0	Min: 4186, Max: 171154, Average: 41118.411
TAHUN	Integer	0	Min: 2013, Max: 2021, Average: 2017.017

Gambar 6. Review Missing Value

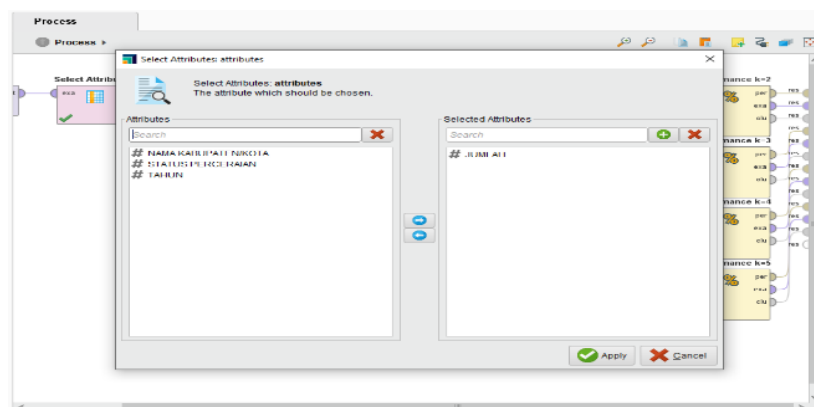
### c) Preprocessing Transformasi Data

Dalam tahap ini atribut yang telah ditentukan selanjutnya di ubah atau disesuaikan ke dalam bentuk algoritma yang akan digunakan, dalam hal ini adalah algoritma K-Mean. Pada tahap transformasi data terdapat 2 proses yang dilakukan, yang pertama adalah menggunakan operator nominal to numerical agar atribut yang ada pada data penelitian (atribut nama kabupaten, status perceraian, dan tahun) dapat diproses karena data mining hanya dapat memproses data yang berbentuk numerical atau angka.



Gambar 7. Transformasi Data (Nominal to Numerical)

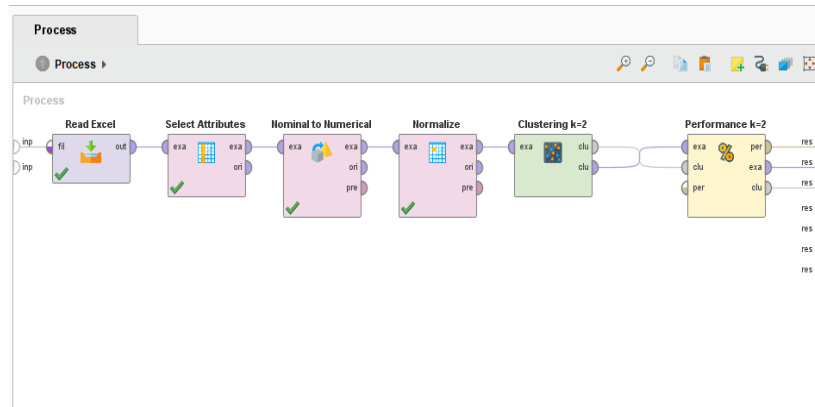
Setelah proses nominal to numerical dilakukan, langkah kedua yaitu menggunakan operator normalize. Normalize digunakan untuk menentukan rentang nilai dalam sebuah data. Menyesuaikan rentang nilai sangat penting saat berhadapan dengan Atribut unit dan skala yang berbeda. Atribut yang digunakan dalam proses Normalize hanya jumlah dan status perceraian.



Gambar 8. Transformasi Data (Normalize)

#### d) Komputasi K-Means

Setelah proses preprocessing dan transformasi data dilakukan, langkah selanjutnya adalah melakukan proses clustering menggunakan tools Rapidminer Studio versi 10.1.0 dengan penerapan algoritma K-Means. Data gambar penerapan algoritma K-Means menggunakan tools RapidMiner Studio versi 10.1.0 tersaji pada gambar berikut :



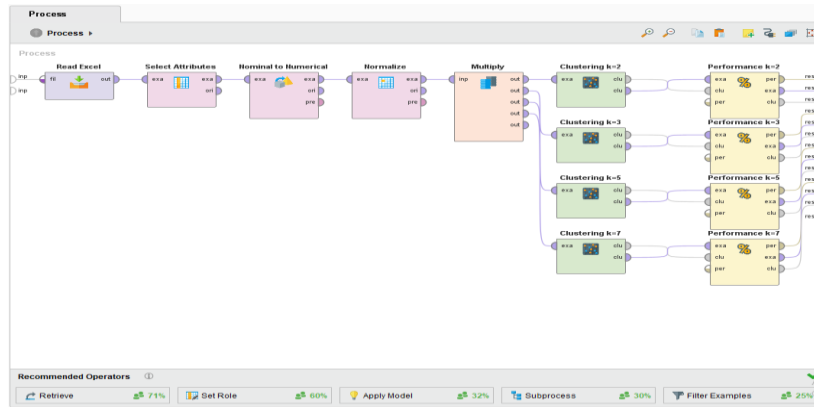
**Gambar 9. Komputasi K-Means**

Pada Gambar 9 menunjukan tahapan atau proses melakukan clustering menggunakan RapidMiner. Dalam melakukan clustering, operator yang digunakan pertama adalah Read Excel yang berfungsi untuk membaca data file microsoft excel. Operator yang digunakan ketiga yaitu Select Atribut berfungsi untuk menentukan atau memilih atribut apa saja yang akan dipakai. Operator yang digunakan keempat yaitu Nominal to Numerical yang berfungsi untuk mengubah data atribut menjadi tipe data angka/numerik. Operator yang digunakan kelima yaitu Normalize yang berfungsi untuk menyesuaikan rentang nilai atribut yang dipakai. Selanjutnya adalah Operator clustering K-Means yang digunakan untuk mengelompokan data berdasarkan algoritma K-Mean dan untuk Performance digunakan untuk menampilkan nilai atau mengevaluasi hasil kinerja nilai Davies Bouldin Index dan Centroid Distance.

## 2. Hasil Penentuan Jumlah Cluster Optimal

Penentuan jumlah cluster terbaik dapat dilakukan menggunakan tools RapiMiner dengan cara melihat hasil performa dari Index Davies Bouldin (IDB). Dimana metode IDB ini merupakan tahapan evaluasi pada cluster yang ada dengan memperhatikan nilai terendah, yang mana hasil dari sebuah cluster yang rendah ini dapat dikalkulasikan berdasarkan kohesi dan separasinya (Dwi Rahayu et al., 2022).

*Implementasi Data Mining Clustering Dalam Mengelompokkan Kasus Perceraian Yang Terjadi Di Provinsi Jawa Barat Menggunakan Algoritma K-Means*



**Gambar 10. Menentukan Cluster Terbaik**

Pada Gambar 10 menampilkan proses perhitungan nilai menggunakan Davies Bouldin Index menggunakan tools RapidMiner, yang mana dari K2, K3, K4, dan K5 cluster yang akan dihitung akan diketahui berapa nilai DBI yang akan dihasilkan.

**Tabel 3. Nilai DBI Performance Cluster**

No	Cluster	Nilai Davies Bouldin Index
1	K2	0,168
2	K3	0,209
3	K4	0,247
4	K5	0,270

Berdasarkan Tabel III menunjukkan hasil cluster yang terbaik yaitu k=2 dengan perolehan nilai Davies Bouldin Index 0,168.

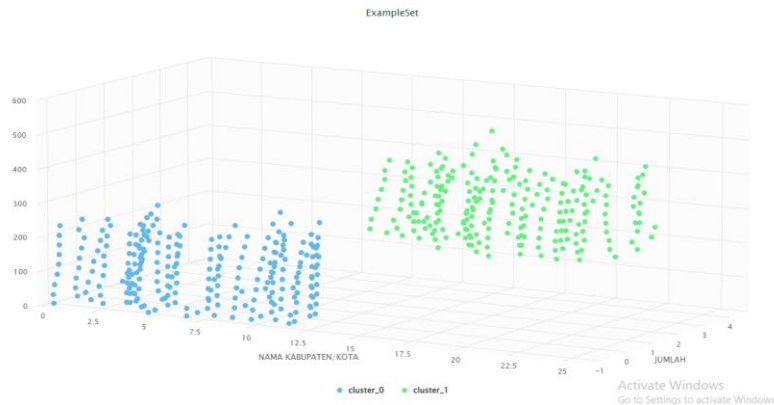
### 3. Hasil Komputasi K-Means Clustering

#### 1) Hasil Performance



**Gambar 11. Hasil Performance Vektor**

Berdasarkan Gambar pengujian Performance Vector tersebut diperoleh hasil 2 kelompok cluster yang memiliki rata-rata centroid 5.870, dengan rata-rata centroid cluster 0 adalah 6.068 dan rata-rata centroid cluster 1 adalah 5.658. K=2 dipilih karena berdasarkan kalkulasi yang dihasilkan dari nilai DBI, acuannya adalah nilai yang mendekati 0 dimana untuk K=2 diperoleh nilai DBI terkecil yaitu 0,168.



**Gambar 12. Visualisasi Scatter Plot Cluster**

Gambar diatas menunjukan bahwa scatter plot persebaran klaster yang diperoleh dari hasil pengelompokan perceraian. Scatter berwarna biru menunjukan cluster 0 dan scatter berwarna hijau menunjukan cluster 1 dengan nilai cluster pada sumbu x dan sumbu y. Kelompok pada cluster 0 dapat dikategorikan sebagai kasus perceraian rendah dan cluster 1 dapat dikategorikan sebagai kasus perceraian tinggi. Adapun hasil pengelompokan dari masing-masing cluster sebagai berikut :

**Tabel 4. Daftar Cluster 0**

No	Nama Kabupaten/Kota	Cluster
1.	Kabupaten Purwakarta	cluster_0
2.	Kabupaten Karawang	cluster_0
3.	Kabupaten Bekasi	cluster_0
4.	Kabupaten Bandung Barat	cluster_0
5.	Kabupaten Pangandaran	cluster_0
6.	Kota Bogor	cluster_0
7.	Kota Sukabumi	cluster_0
8.	Kota Bandung	cluster_0
9.	Kota Cirebon	cluster_0
10.	Kota Bekasi	cluster_0
11.	Kota Depok	cluster_0
12.	Kota Cimahi	cluster_0
13.	Kota Tasikmalaya	cluster_0
14.	Kota Banjar	cluster_0

**Tabel 5. Daftar Cluster 1**

No	Nama Kabupaten/Kota	Cluster
1.	Kabupaten Bogor	cluster_1
2.	Kabupaten Sukabumi	cluster_1
3.	Kabupaten Cianjur	cluster_1
4.	Kabupaten Bandung	cluster_1
5.	Kabupaten Garut	cluster_1
6.	Kabupaten Tasikmalaya	cluster_1
7.	Kabupaten Ciamis	cluster_1
8.	Kabupaten Kuningan	cluster_1
9.	Kabupaten Cirebon	cluster_1
10.	Kabupaten Majalengka	cluster_1
11.	Kabupaten Sumedang	cluster_1
12.	Kabupaten Indramayu	cluster_1
13.	Kabupaten Subang	cluster_1
14.	Kabupaten Bogor	cluster_1

Dari tabel yang ditampilkan diatas diperoleh bahwa Cluster terbaik dari pengelompokan kasus perceraian terdapat 2 cluster yaitu cluster 0 terdapat 5 kabupaten dan 9 kota sedangkan pada cluster 1 terdapat 13 kabupaten, dengan nilai davies bouldin index 0,168 dan nilai avg.within centroid distance 5.870. Cluster 0 merupakan kota/kabupaten dengan kasus perceraian rendah dan untuk cluster 1 merupakan kota/kabupaten dengan kasus perceraian tertinggi.

## **KESIMPULAN DAN SARAN**

Kesimpulan pada penelitian ini, berdasarkan dari data yang sudah diolah menggunakan RapidMiner dengan menggunakan metode algoritma K-Means memberikan hasil tingkat kasus perceraian rendah dan kasus perceraian tinggi. Berdasarkan penentuan kluster menggunakan 2 cluster karena hasil Davies Bouldin Index nya mendekati nilai 0.

Kelompok/cluster 0 merupakan jumlah kasus perceraian rendah dengan anggota cluster berjumlah 5 kabupaten dan 9 kota yang meliputi Kabupaten Purwakarta, Kabupaten Karawang, Kabupaten Bekasi, Kabupaten Bandung Barat, Kabupaten Pangandaran, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, Kota Banjar.

Kelompok/cluster 1 merupakan jumlah kasus perceraian yang tinggi dengan anggota cluster berjumlah 13 kabupaten yang meliputi Kabupaten Bogor, Kabupaten Sukabumi, Kabupaten Cianjur, Kabupaten Bandung, Kabupaten Garut, Kabupaten Tasikmalaya,

Kabupaten Ciamis, Kabupaten Kuningan, Kabupaten Cirebon, Kabupaten Majalengka, Kabupaten Sumedang, Kabupaten Indramayu, dan Kabupaten Subang.

## DAFTAR REFERENSI

- Anggreini Novita Lestari. (2019). Teknik Clustering Dengan Algoritma K-Medoids Untuk Menangani Strategi Promosi Di Politeknik Tedc Bandung. *Jurnal Teknologi Informasi Dan Pendidikan*, 12(2). <http://tip.ppj.unp.ac.id>
- Aprianti, W., & Permadi, J. (2018). K-Means Clustering Untuk Data Kecelakaan Lalu Lintas Jalan Raya Di Kecamatan Pelaihari. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIK)*, 5(5), 613–620. <https://doi.org/10.25126/jtiik2018551113>
- Azwanti, N. (2018). Segmentasi Tingkat Pemakaian Material dengan Data Mining Clustering. *Jurnal Komputer Terapan*, 4(2). <http://jurnal.pcr.ac.id>
- Disi, A. P., & Sumargo, B. (2019). Pengelompokan Pengguna Internet Dengan Metode K-Means Clustering. *Jurnal Statistika Dan Aplikasinya (JSA)*, 3(1).
- Dwi Rahayu, U., Chusna, N. L., & Fachri, M. (2022). Analisis Kasus Perceraian Pada Pengadilan Negeri Bekasi Menggunakan Algoritma K-Means Clustering. *IKRAITH-INFORMATIKA*, 6.
- Elvi. (2021). Analisis Algoritma K-Means Dalam Pengelompokan Perkara Perceraian Berdasarkan Kelurahan Di Kota Jambi. *Jurnal Processor*, 16(1), 9–19. <https://doi.org/10.33998/processor.2021.16.1.920>
- Mulyani, E. D. S., Yusup, A. M., Tisna, A. K., Fauzi, F. A., Seta, I. B., Khairunas, R., & Ardiansyah, W. (2019). Clustering Wilayah Dan Pelanggaran Berkendaraan Menggunakan Algoritma K-Means Pada Data Satlanta Polres Tasikmalaya Kota. 8(1).
- Nurhayati, N., Azzahra, F., Ramadani, S., Hastuti, S. D., & Irawan, E. (2020). Analisis Metode Klastering Pada Kasus Penyebab Perceraian Berdasarkan Provinsi Dengan Teknik K-Means. *Jurnal KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 4(1). <https://doi.org/10.30865/komik.v4i1.2699>
- Parlina, I., Perdana Windarto, A., Wanto, A., & Lubis, Mr. (2018). Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Asessment Center Untuk Clustering Program Sdp. *CESS (Journal of Computer Engineering System and Science)*, 3(1), 2502–2714.
- Triyansyah, D., & Fitriana, D. (2018). Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing. *Jurnal Telekomunikasi Dan Komputer*, 8. <https://doi.org/10.22441/incomtech.v8i2.4174>