



## PENGELOMPOKAN DAERAH BENCANA ALAM MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING

**Isni Rinjani**

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) IKMI Cirebon

**Saeful Anwar**

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) IKMI Cirebon

**Ruli Herdiana**

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) IKMI Cirebon

Alamat: Cirebon, Jawa Barat, Indonesia 45131

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

**Abstract.** *Natural disasters are events that significantly affect the human population. Landslides, earthquakes, floods, fires, droughts, earthquakes and other natural disasters often occur in West Java Province. Information and technology skills are developing quite fast nowadays. Thanks to modern technology, anyone can access and obtain information without restrictions. Information is very important for every aspect of life. One of them is information about natural disasters, because disaster management needs this kind of information. Data mining is a popular method for analyzing disaster data because it is considered a potential answer to disaster management challenges. Therefore, this study discusses the grouping of natural disaster areas for prediction of natural disaster areas in West Java with data mining techniques using the k-means clustering algorithm. The results of the study obtained 3 clusters including low clusters, medium clusters, and high clusters. The selected research source comes from the official website, namely West Java Open Data. The results of this research are expected to provide useful information in determining solutions to natural disaster management problems.*

**Keywords:** *Data Mining, Natural Disasters, K-Means, Clustering.*

**Abstrak.** Bencana alam adalah kejadian yang secara signifikan mempengaruhi populasi manusia. Tanah longsor, gempa bumi, banjir, kebakaran, kekeringan, puyuh dan bencana alam lainnya sering terjadi di Provinsi Jawa Barat. Keahlian informasi dan teknologi berkembang cukup cepat saat ini. Berkat teknologi modern, siapa pun dapat mengakses dan memperoleh informasi tanpa batasan. Informasi sangat penting untuk setiap aspek kehidupan. Salah satunya adalah informasi tentang bencana alam, karena penanggulangan bencana memerlukan informasi semacam ini. Data mining adalah metode populer untuk menganalisis data bencana karena dianggap sebagai jawaban potensial untuk tantangan pengelolaan bencana. Oleh karena itu, dalam penelitian ini membahas tentang pengelompokan daerah bencana alam untuk prediksi daerah bencana alam di Jawa Barat dengan teknik data mining menggunakan algoritma k-means clustering. Hasil penelitian didapatkan 3 cluster diantaranya cluster rendah, cluster sedang, dan cluster tinggi. Sumber penelitian yang dipilih berasal dari website resmi yaitu open data jabar. Hasil dari penelitian ini diharapkan dapat memberikan informasi yang berguna dalam menentukan solusi atas permasalahan penanggulangan bencana alam.

**Kata kunci:** Data Mining, Bencana Alam, K-Means, Clustering.

### LATAR BELAKANG

Secara geografis, Indonesia merupakan negara kepulauan yang berada pada pertemuan lempeng Eurasia, Indo-Australia, Filipina, dan Pasifik. Terdapat sabuk vulkanik di wilayah selatan dan timur

Indonesia yang membentang dari Sulawesi hingga pulau Sumatera, Jawa, dan Nusa Tenggara. Indonesia rentan terhadap gempa bumi, tsunami dan letusan gunung berapi sebagai akibat dari kerentanan ini. (Murdiaty et al., 2020). Jawa Barat memiliki frekuensi bencana alam tertinggi di Indonesia, menurut Badan Nasional Penanggulangan Bencana (BNPB) pada tahun 2020 terdapat 3 jenis bencana yang banyak terjadi di Jawa Barat yaitu: angin puting beliung 253 kejadian, tanah longsor 218 kejadian, dan banjir 170 kejadian. BNPB juga mencatat sebaran kejadian bencana alam di Jawa Barat pada tahun 2021 yaitu 768 kejadian bencana, sedangkan pada tahun 2022 tercatat 814 kejadian bencana (Firman et al., 2022).

Kemajuan informasi dan teknologi saat ini mengalami kemajuan yang sangat pesat. Siapapun, terlepas dari batasan lokasi atau waktu, dapat mengakses dan mendapatkan informasi dengan cepat dan benar berkat kecanggihan teknologi. Karena itu, pengetahuan ini sangat berguna dan penting untuk membuat keputusan. Warga saat ini membutuhkan informasi tentang berbagai topik, termasuk bencana alam. Namun, informasi yang diberikan tentang bencana alam terkadang kurang relevan, tidak mudah untuk dipahami, dan tidak berdasar. Meskipun bencana alam tidak dapat dihindari, efeknya dapat dikurangi dengan menentukan penyebabnya dan meneliti peristiwa bencana yang telah terjadi melalui analisis data bencana. Agar zona bencana alam dapat menjadi indikator penanggulangan bencana, diperlukan pengelompokan. Berbagai macam metode pengelompokan data menjadi sebuah informasi yang mudah dipahami sudah banyak dilakukan. Data mining algoritma K-means clustering merupakan salah satu teknik yang dapat dilakukan untuk pengelompokan data bencana alam (Firman et al., 2022).

Beberapa penelitian terdahulu menunjukkan bahwa data mining dapat digunakan dengan baik dalam mengelompokkan data bencana alam. Penelitian yang dilakukan oleh Muhamad Iqbal Ramadhan pada tahun 2017 dalam jurnal yang berjudul "Penerapan Data Mining Untuk Analisa Data Bencana Milik BNPB Menggunakan Algoritma K-means dan Linear Regression" menunjukkan hasil penggunaan agregasi data dan regresi linier untuk meramalkan bencana menggunakan data BNPB untuk lima tahun ke depan (Iqbal Ramadhan, 2017). Penelitian dengan judul "Pengelompokan Data Bencana Alam Berdasarkan Wilayah, Waktu, Jumlah Korban Dan Kerusakan Fasilitas Menggunakan Algoritma K-Means" yang dilakukan oleh Murdiaty dan Angela pada tahun 2020 mengelompokkan data bencana alam memanfaatkan teknik clustering menggunakan algoritma k-means menjadi beberapa kelompok, dilihat dari segi jenis bencana alam, waktu terjadinya bencana, jumlah korban, dan kerusakan berbagai fasilitas sebagai dampak bencana alam (Murdiaty et al., 2020). Berdasarkan penelitian yang dilakukan dapat disimpulkan bahwa metode K-means clustering dapat digunakan untuk mengelompokkan daerah bencana alam menjadi 3 kelompok yaitu rendah, sedang dan tinggi

Masalah yang mendasar dalam pengelompokan daerah bencana alam menggunakan metode K-Means Clustering adalah bagaimana mengelompokkan daerah bencana alam berdasarkan tingkat daerah yang rawan terhadap bencana alam. Hal ini dapat menyebabkan kesulitan dalam mengambil tindakan yang efektif untuk mencegah dan mengatasi bencana yang mungkin terjadi. Metode K-Means Clustering digunakan untuk mengelompokkan daerah bencana alam sehingga memudahkan dalam menentukan kebijakan dalam merencanakan tindakan pencegahan dan mitigasi bencana. Namun, metode ini

memerlukan data yang cukup banyak dan akurat serta perlu ada analisa yang baik untuk menentukan jumlah cluster yang sesuai sehingga hasil yang di dapatkan benar benar representatif.

Tujuan dari penelitian ini adalah untuk mengelompokan daerah-daerah yang rawan terhadap bencana alam sehingga hasil dari penelitian dapat dibagi ke dalam 3 kelompok yaitu rendah, sedang dan tinggi Dan menentukan cluster optimum dan nilai dbi dari mengelompokkan daerah bencana alam menggunakan algoritma K-Means clustering.

## **KAJIAN TEORITIS**

### **A. Data Mining**

Data mining merupakan sebuah proses dengan menggunakan Teknik dan alat analisis data untuk mencari pola dan hubungan yang tersembunyi dari sebuah data(Surya et al., 2021). Data mining adalah metode pencarian pola pada data yang belum terdapat dalam database dengan terlebih dahulu membuat model. Pengelompokan data juga dilakukan dalam penambangan data(Asroni et al., 2018).

### **B. Clustering**

Clustering adalah metode pengelompokan data ke dalam beberapa cluster maupun kelompok dimana data dalam cluster memiliki kemiripan yang maksimum dan data antar cluster memiliki kemiripan yang minimum(Fatmawati & Windarto, 2018). Tujuan clustering adalah untuk menempatkan data, pengamatan, atau kelompok ke dalam kelas-kelas yang berbagi objek. Pengklasteran berbeda dengan klasifikasi yang tidak adanya variable target dalam pengklasteran. Nilai target tidak dapat diklasifikasikan, diperkirakan, atau diprediksi menggunakan pengelompokan(Puji Rahayu et al., 2020).

### **C. Algoritma K-Means**

Algoritma K-Means merupakan algoritma yang berulang-ulang. Jumlah cluster yang ingin dibuat adalah N, dan metode K-Means dimulai dengan memilih N secara acak. Kemudian bagikan nilai N secara acak, nilai ini berfungsi sebagai centroid, mean, atau "means" dari cluster(Putu et al., n.d.). Algoritma K-means merupakan salah satu algoritma clustering (pengelompokan). Kmeans clustering merupakan metode clustering non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih cluster. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan cluster yang lain, sehingga data yang berada dalam satu cluster memiliki tingkat variasi kecil(Dhuhita, 2015).

### **D. Rapidminer**

RapidMiner adalah perangkat lunak yang mudah diakses dan tersedia secara konsisten (open source)(Mohede et al., 2020). RapidMiner adalah salah satu software untuk pengolahan data mining(Ramdhan et al., 2022).

## METODE PENELITIAN

Tahapan perancangan yang digunakan dalam penelitian ini menggunakan rancangan Knowledge Discovery in Database (KDD). KDD adalah metode untuk mengekstraksi atau membantu pengetahuan, informasi prospektif, dan pola yang ditemukan dalam kumpulan data yang sangat besar. Pengetahuan dan data yang dihasilkan oleh KDD dapat diandalkan, segar, mudah dipahami, dan praktis. Data jelas diperlukan untuk metode Knowledge Discovery in Database (KDD), yang menemukan pengetahuan dalam database.. Data yang digunakan yaitu data yang sudah terpisah dengan data operasional. Berikut adalah tahapan perancangan KDD yang akan dilakukan :



Gambar 1 Tahapan Perancangan

### 1. Data Preprocessing

Data Preprocessing adalah langkah dalam proses penambahan data. sebelum masuk ke tahap pengolahan. Kami akan mulai dengan memproses data mentah. Praprocessing data, juga dikenal sebagai prapemrosesan data, yaitu dilakukan dengan memilih data yang tepat dan menggunakan eliminasi untuk data yang tidak tepat.

### 2. Data Transformation

Data Transformation merupakan proses mengubah data yang dipilih menjadi prosedur data mining menggunakan berbagai teknik dan agresivitas data.

### 3. Data Mining

Data mining yaitu Langkah yang paling penting di mana teknik yang berbeda digunakan untuk mengekstraksi pola potensial yang berbeda untuk menghasilkan data yang relevan.

### 4. Interpretation/evolution

Memeriksa apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau teori yang diketahui merupakan bagian dari tahap ini.

## Teknik Pengumpulan Data

Teknik pengumpulan data yaitu dengan melakukan teknik online research, diperoleh dari mendownload atau mengunduh melalui website <https://opendata.jabarprov.go.id/> yang sudah disediakan oleh pemerintah Provinsi Jawa Barat. Adapun Sumber data yang digunakan yaitu dari Dinas Pemberdayaan Masyarakat dan Desa Jawa Barat, merupakan data Sekunder yang di publikasikan melalui website resmi Open Data Jabar ( <https://opendata.jabarprov.go.id> ) sebagai portal data terbuka yang menyediakan data

akurat dari seluruh organisasi perangkat daerah Jawa Barat. Dataset yang diambil yaitu data jumlah desa yang mengalami kejadian bencana alam berdasarkan jenis bencana di Provinsi Jawa Barat Periode tahun 2020-2022 yang diambil pada tanggal 21 Januari 2023.

**HASIL DAN PEMBAHASAN**

**A. Pengelompokan Daerah bencana Alam Menggunakan Algoritma K-Means Clustering**

**1. Pengumpulan Data**

Dataset yang digunakan adalah dataset daerah bencana alam periode tahun 2020-2022 yang di dapat secara online melalui website resmi yaitu Open Data Jabar. Jumlah record dataset ini yaitu berjumlah 570 data.

Tabel 1 Data Bencana Alam

id	Provinsi	Kode Kab/Kota	Nama Kab/Kota	Jenis bencana	jumlah	satuan	tahun
1	JAWA BARAT	3201	KABUPATEN BANDUNG	TANAH LONGSOR	122	DESA	2020
2	JAWA BARAT	3201	KABUPATEN BANDUNG	BANJIR	194	DESA	2020
3	JAWA BARAT	3201	KABUPATEN BANDUNG	GEMPA	114	DESA	2020
.....	.....	.....	.....	.....	.....	.....	.....
570	JAWA BARAT	3219	KABUPATEN BANJAR	BENCANA LAINNYA	2	DESA	2022

Tabel Diatas merupakan sebagian data daerah bencana di Jawa Barat yang memiliki 8 atribut yaitu atribut id, Provinsi, Kode Kab/Kota, Nama Kab/Kota, Jenis Bencana, Jumlah, Satuan dan tahun.

**2. Data Preprocessing**

Pada tahap ini akan dilakukan preprocessing data dengan mengurangi atribut yang tidak digunakan dalam proses data mining . Terdapat 4 atribut yang akan di gunakan yaitu kode Kab/Kota, Jenis Bencana, Jumlah dan Tahun.

Tabel 2 Data Preprocessing

Kode Kab/Kota	Jenis bencana	jumlah	tahun
3201	TANAH LONGSOR	122	2020
3201	BANJIR	194	2020
3201	GEMPA	114	2020

.....	.....	.....	.....
3219	BENCANA LAINNYA	2	2022

Tabel diatas menunjukkan data yang akan dipakai ke proses selanjutnya dengan mengurangi/mengeliminasi atribut atribut yang tidak sesuai hingga menghasilkan data yang sesuai.

### 3. Data Transformation

Pada tahap ini dilakukan transformasi data dengan bentuk yang sesuai pada proses data mining. Berikut transformasi data yang dilakukan. Dari 570 record data di transformasikan menjadi 19 data yang merupakan 19 Kabupaten/Kota dengan 1 kolom Kode Kab/Kota dan 10 kolom kejadian bencana yaitu Banjir, Gempa, Gunung Meletus, Kebakaran, Kekeringan, Puyuh, Tanah longsor, Tsunami, Gelombang Pasang dan Bencana lainnya yang di kalkulasikan pada periode 2020-2022 . Berikut adalah tabel data akhir untuk proses algoritma K-Means clustering.

Tabel 3 Data Transformation

Kode Kab/Kota	Banjir	Gempa	Gunung Meletus	Kebakaran	Kekeringan	Puyuh	Tanah Longsor	Tsunami	Gelombang pasang	Bencana Lainnya
3201	352	232	2	112	223	195	293	0	0	71
3202	61	133	7	50	202	93	313	2	3	73
3203	350	32	3	12	227	48	89	4	39	71
3204	517	405	2	60	311	483	800	3	3	127
3205	152	450	0	334	68	367	468	3	1	105
3206	158	484	9	150	530	201	561	18	32	56
3207	529	22	1	20	244	99	76	0	37	65
3208	382	646	9	191	574	162	987	6	63	106
3209	641	3	0	7	256	95	11	1	468	96
3210	429	19	3	40	234	201	54	4	56	59
3211	140	47	1	109	265	27	287	9	0	70
3212	111	40	0	75	172	60	190	0	0	65
3213	137	276	0	78	168	90	127	0	44	29
3214	40	39	0	66	190	56	219	1	1	35
3215	240	43	5	44	252	118	183	1	31	78
3216	256	824	3	130	442	259	858	1	85	110
3217	94	56	37	198	202	26	429	0	10	46
3218	160	911	20	101	500	167	767	11	20	89
3219	19	16	0	3	37	23	17	0	0	5

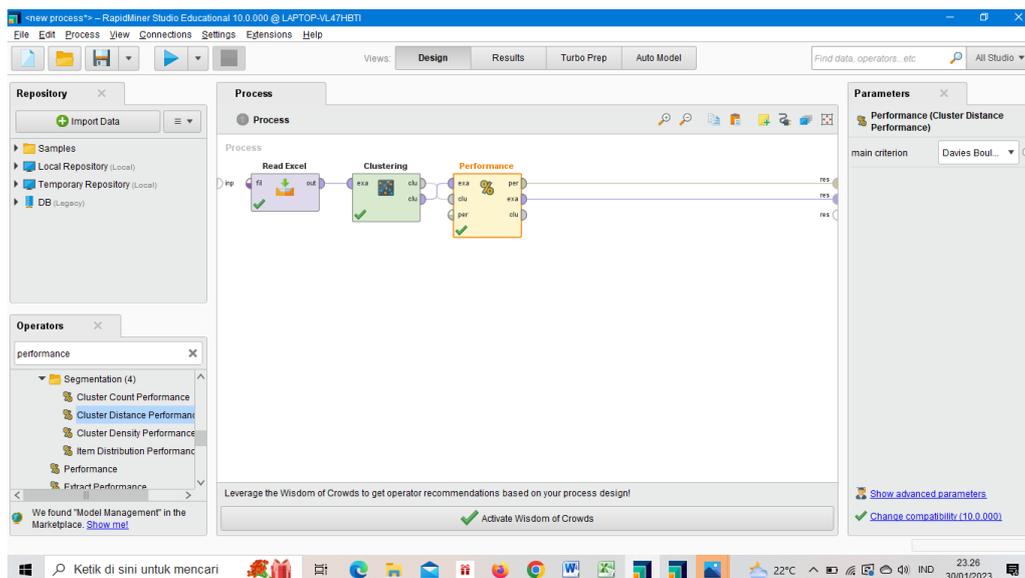
Tabel tersebut merupakan data akhir untuk proses algoritma K-means clustering. Terdapat 11 atribut pada tabel tersebut, dengan atribut Kode Kab/Kota yang merupakan kode dari nama Kab/Kota yang mengalami bencana alam, sedangkan 10 atribut lainnya merupakan jenis bencana alam yang berisikan jumlah bencana alam di Jawa Barat yang di kalkulasikan pada periode 2020-2022. Berikut adalah keterangan dari kode Kab.Kota :

- 3201 = Kabupaten Bandung
- 3202 = Kabupaten Bandung Barat
- 3203 = Kabupaten Bekasi
- 3204 = Kabupaten Bogor
- 3205 = Kabupaten Ciamis
- 3206 = Kabupaten Cianjur
- 3207 = Kabupaten Cirebon
- 3208 = Kabupaten Garut
- 3209 = Kabupaten Indramayu
- 3210 = Kabupaten Karawang
- 3211 = Kabupaten Kuningan
- 3212 = Kabupaten Majalengka
- 3213 = Kabupaten Pangandaran
- 3214 = Kabupaten Purwakarta
- 3215 = Kabupaten Subang
- 3216 = Kabupaten Sukabumi
- 3217 = Kabupaten Sumedang
- 3218 = Kabupaten Tasikmalaya
- 3219 = Kota Banjar

#### **4. Data Mining**

Setelah data di transformasikan pada tahap ini akan dibuat model pengolahan data dengan algoritma K-Means clustering menggunakan rapidminer versi 10.0. Berikut pada gambar 4.1 adalah

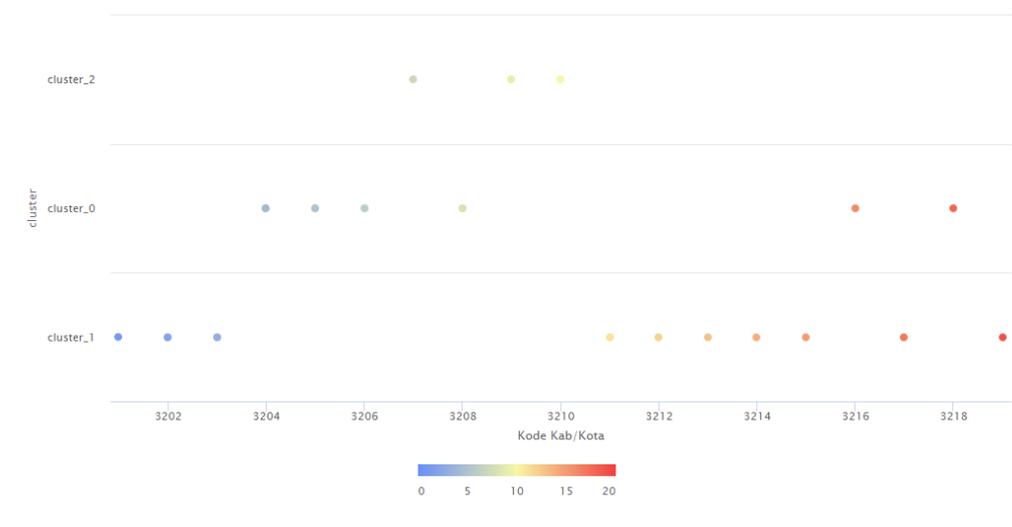
proses data mining K-Means clustering pada rapidminer 10.0.



Gambar 2 Proses Data Mining

Pada gambar diatas menunjukkan proses data mining menggunakan K-Means clustering. Operator read excel untuk memanggil atau membuka dataset dengan format xls, kemudian operator k-means clustering untuk mengelompokkan dan menguji kinerja hasil pengelompokkan dengan menetapkan 3 cluster yaitu cluster rendah, sedang dan tinggi. Operator cluster distance performance yang digunakan untuk menguji cluster 2-10 untuk mencari nilai optimum dari nilai dbi.

Berikut adalah gambar visualization titik cluster.

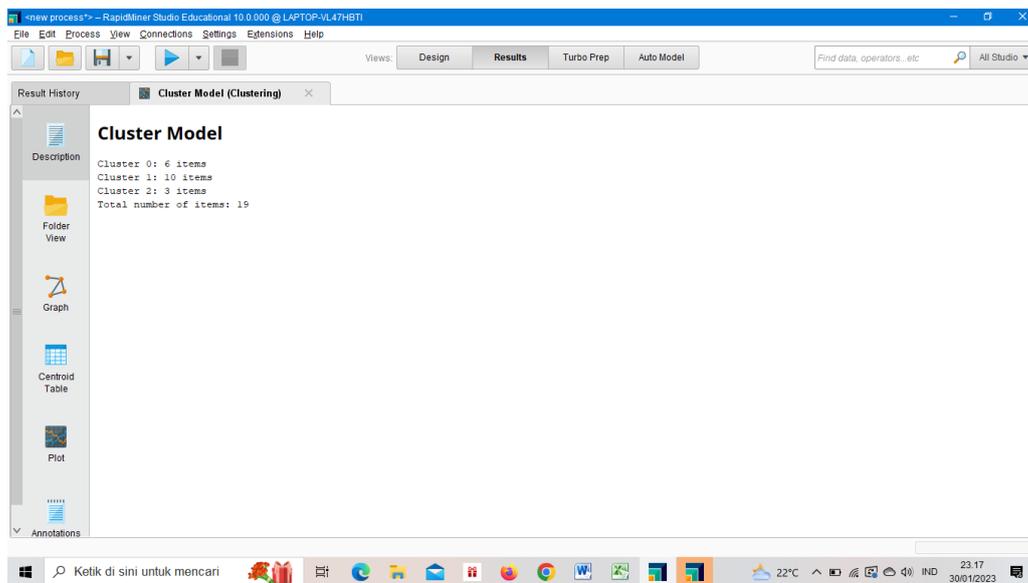


Gambar 3 Visualization titik cluster

Pada gambar tersebut daerah yang termasuk kedalam cluster \_0 adalah intensitas bencana tinggi, data daerah yang termasuk ke dalam cluster\_1 adalah daerah dengan intensitas bencana sedang dan cluster\_2 untuk intensitas bencana yang rendah.

### 5. Interpretation/evolution

Setelah proses data mining, selanjutnya adalah hasil dari pemrosesan k-means clustering. Berikut adalah cluster model jumlah anggota tiap cluster.



Gambar 4 Cluster Model

Dari Gambar diatas dapat dilihat bahwa terdapat 6 daerah pada cluster 0(tinggi), 10 daerah untuk cluster 1(sedang) dan 3 daerah pada cluster 2(rendah). Cluster 0(tinggi) beranggotakan 6 daerah yaitu Kabupaten Bogor, Kabupaten Ciamis, Kabupaten Garut, Kabupaten Sukabumi dan Kabupaten Tasikmalaya dengan kode 3204, 3205, 3208, 3216, 3218. Cluster 1(sedang) beranggotakan 10 daerah yaitu daerah Kabupaten Bandung, Kabupaten Bandung Barat, Kabupaten Bekasi, Kabupaten Kuningan, Kabupaten Majalengka, Kabupaten Pangandaran, Kabupaten Purwakarta, Kabupaten Subang, Kabupaten Sumedang dan Kota Banjar dengan kode 3201, 3202, 3203, 3211, 3212, 3213, 3214, 3215, 3216, 3217, 3219. Cluster 2(rendah) beranggotakan 3 daerah yaitu Kabupaten Cirebon, Kabupaten Indramayu, dan Kabupaten Karawang dengan kode 3207, 3209,3210. Dan untuk nilai

centroid masing-masing cluster dapat dilihat pada gambar di bawah.

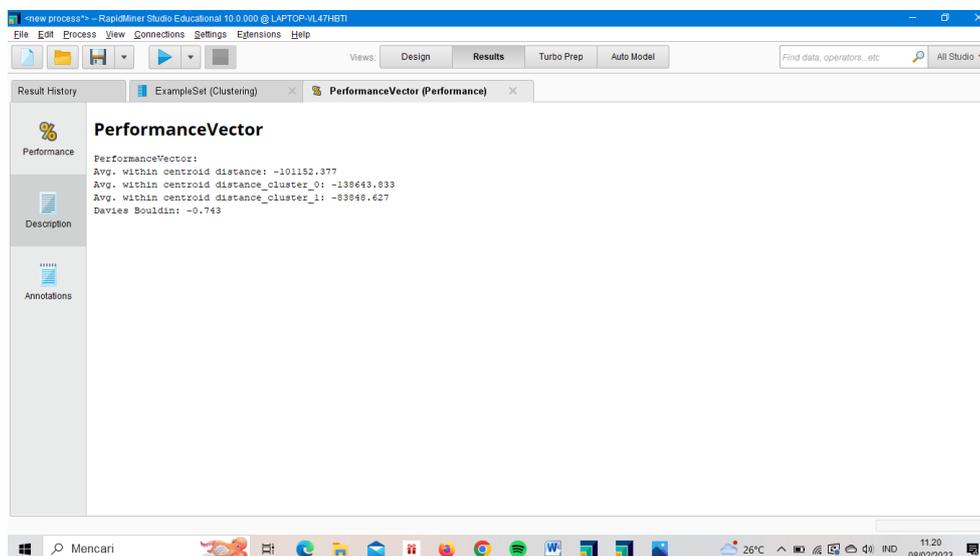
Attribute	cluster_0	cluster_1	cluster_2
Kode Kab/Kota	3209.500	3210.700	3208.667
Banjir	270.833	154.400	533
Gempa	620	91.400	14.667
Gunung meletus	7.167	5.500	1.333
Kebakaran	161	74.700	22.333
Kekeringan	404.167	193.800	244.667
Puyuh	273.167	73.600	131.667
Tanah Longsor	740.167	214.700	47
Tsunami	7	1.700	1.667
Gelombang Pasang	34	12.800	187
Bencana Lainnya	98.833	54.300	73.333

Gambar 5 Nilai Centroid

## B. Hasil Menentukan cluster optimum dan nilai dbi dari Mengelompokkan Daerah Bencana Alam Menggunakan Algoritma K-Means Clustering

Untuk menentukan cluster optimum dan nilai dbi dari mengelompokkan daerah bencana alam menggunakan algoritma k-means ini perlu dilakukan uji coba n nilai cluster 2-10. Dari tiap cluster yang dibuat dievaluasi dengan operator Cluster distance performance untuk mengetahui nilai dbi tiap cluster.

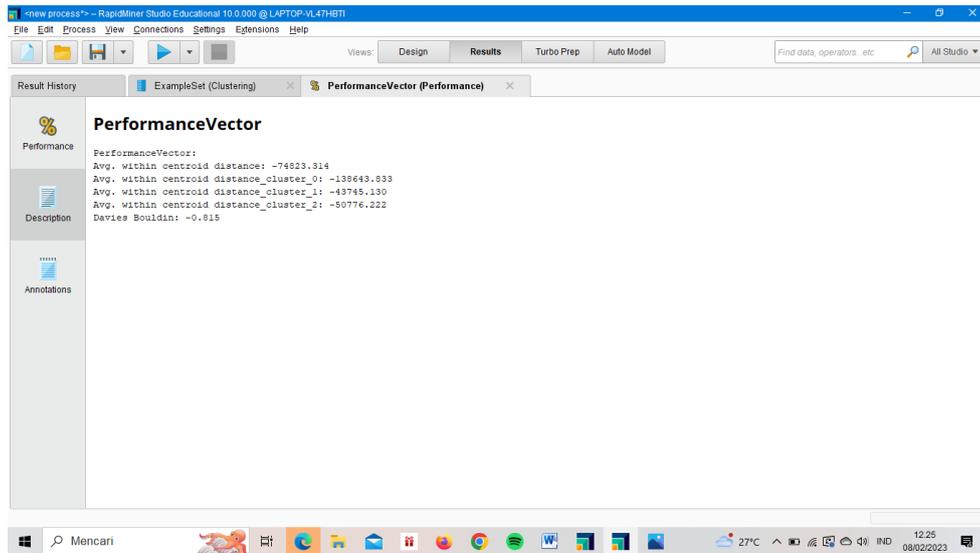
### 1. Nilai dbi cluster 2



Gambar 6 Nilai dbi cluster 2

Dari hasil pengklustran cluster 2 terdapat Nilai Davies Bouldin atau dbi yaitu:  $-0.743$ , Rata-rata dalam jarak pusat :  $-101152.377$ , Rata-rata dalam jarak pusat cluster 0 :  $-138643.833$ , Rata-rata dalam jarak pusat cluster 1 :  $-83848.627$ .

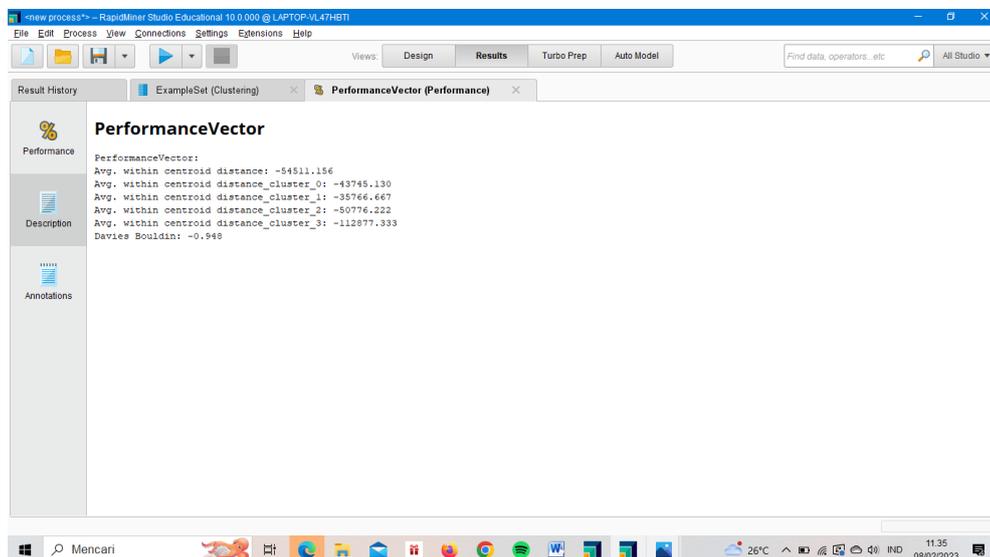
## 2. Nilai dbi cluster 3



Gambar 7 Nilai dbi cluster 3

Dari hasil pengklusteran cluster 3 terdapat Nilai Davies Buildin atau dbi yaitu: -0,815 , Rata-rata dalam jarak pusat : -74823.314, Rata-rata dalam jarak pusat cluster 0 : -138643.833, Rata-rata dalam jarak pusat cluster 1 : -43745.130, Rata-rata dalam jarak pusat cluster 2 : -50776.222.

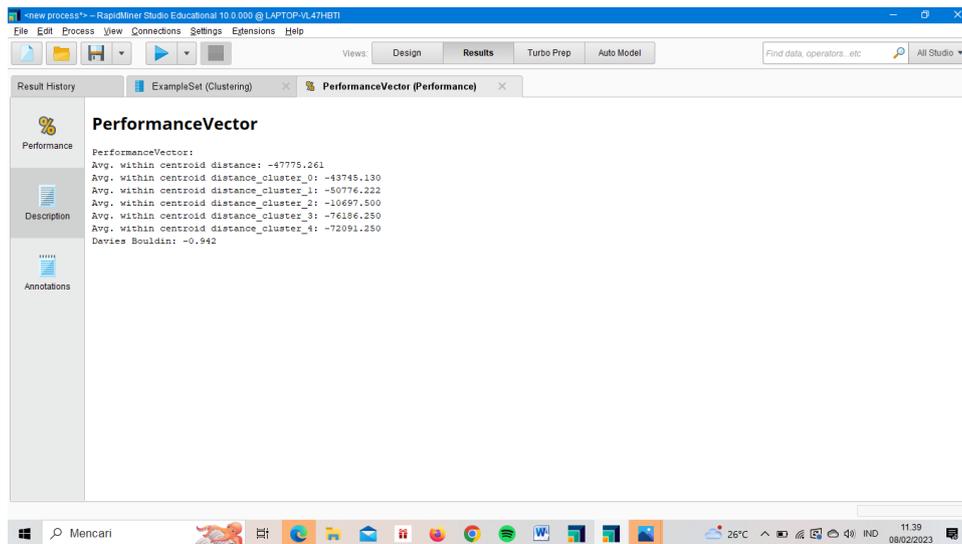
## 3. Nilai dbi cluster 4



Gambar 8 Nilai dbi cluster 4

Dari hasil pengklusteran cluster 4 terdapat Nilai Davies Bouldin atau dbi yaitu: -0.948 , Rata-rata dalam jarak pusat : -54511.156, Rata-rata dalam jarak pusat cluster 0 : -43745.130, Rata-rata dalam jarak pusat cluster 1 : -35766.667, Rata-rata dalam jarak pusat cluster 2 : -50776.222, Rata-rata dalam jarak pusat cluster 3 : -112877.333.

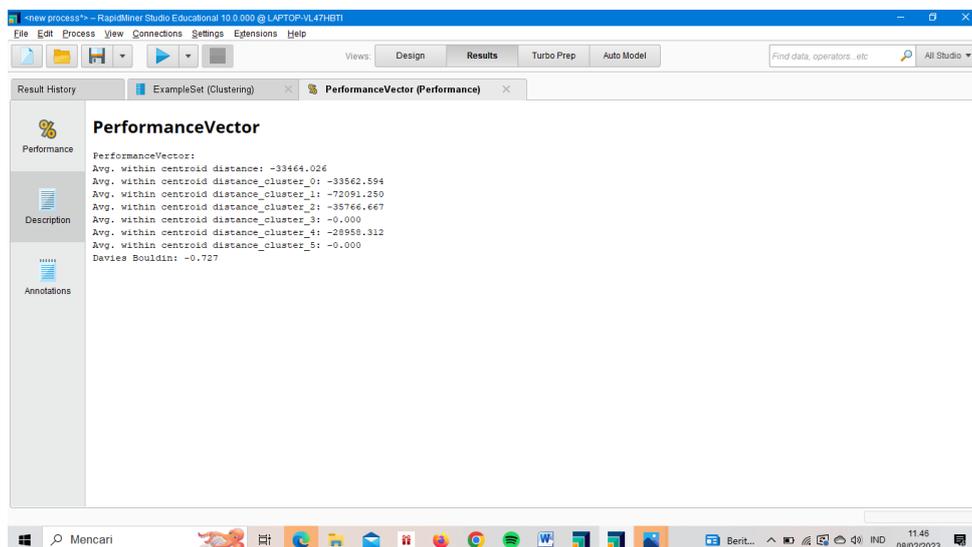
#### 4. Nilai dbi cluster 5



Gambar 9 Nilai dbi cluster 5

Dari hasil pengklusteran cluster 5 terdapat Nilai Davies Bouldin atau dbi yaitu: -0.942, Rata-rata dalam jarak pusat : -47775.261, Rata-rata dalam jarak pusat cluster 0 : -43745.130, Rata-rata dalam jarak pusat cluster 1 : -50776.222, Rata-rata dalam jarak pusat cluster 2 : -10697.500, Rata-rata dalam jarak pusat cluster 3 : -76186.250, Rata-rata dalam jarak pusat cluster 4 : -72091.250.

#### 5. Nilai dbi cluster 6

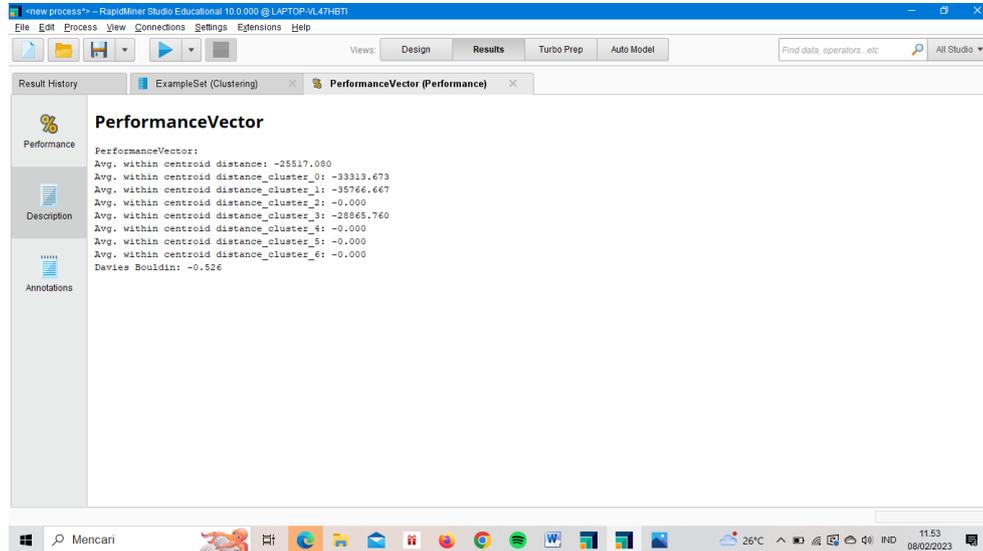


Gambar 10 Nilai dbi cluster 6

Dari hasil pengklusteran cluster 6 terdapat Nilai Davies Bouldin atau dbi yaitu: -0.727, Rata-rata dalam jarak pusat : -33464.026, Rata-rata dalam jarak pusat cluster 0 : -33562.594, Rata-rata dalam jarak pusat cluster 1 : -72091.250, Rata-rata dalam jarak pusat cluster 2 : -35766.667, Rata-

rata dalam jarak pusat cluster 3 : -0.000, Rata-rata dalam jarak pusat cluster 4 : -28958.312, Rata-rata dalam jarak pusat cluster 5 : -0.000.

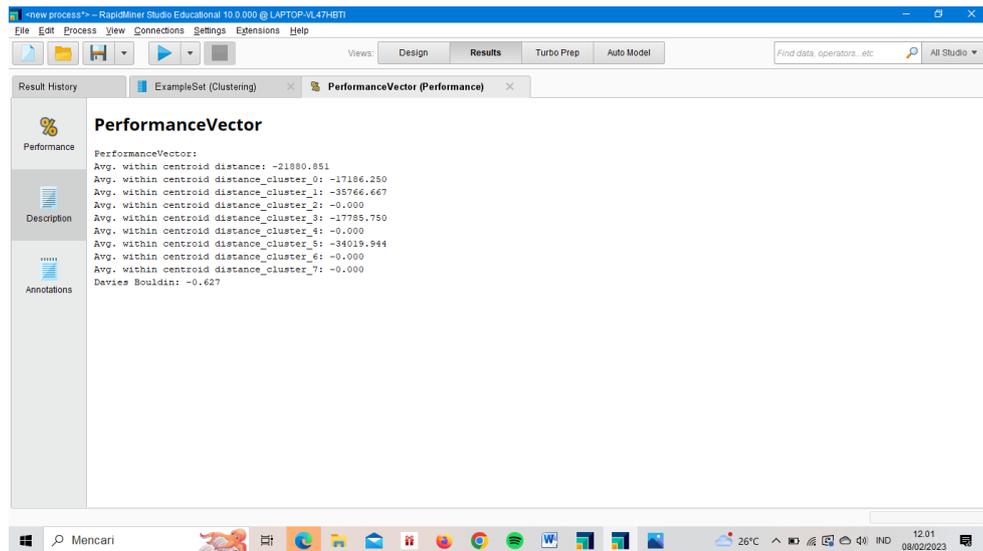
**6. Nilai dbi cluster 7**



Gambar 11 Nilai dbi cluster 7

Dari hasil pengklusteran cluster 7 terdapat Nilai Davies Bouldin atau dbi yaitu: -0.526, Rata-rata dalam jarak pusat : -25517.080, Rata-rata dalam jarak pusat cluster 0 : -33313.673, Rata-rata dalam jarak pusat cluster 1 : -35766.667, Rata-rata dalam jarak pusat cluster 2 : -0.000, Rata-rata dalam jarak pusat cluster 3 : -28865.760, Rata-rata dalam jarak pusat cluster 4 : -0.000, Rata-rata dalam jarak pusat cluster 5 : -0.000, Rata-rata dalam jarak pusat cluster 6 : -0.000.

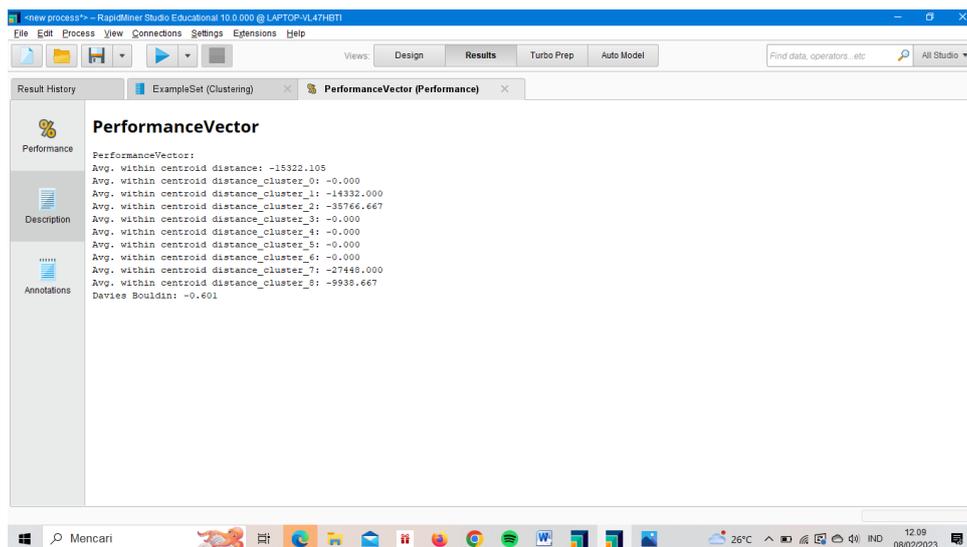
**7. Nilai dbi cluster 8**



Gambar 12 Nilai dbi cluster 8

Dari hasil pengklusteran cluster 8 terdapat Nilai Davies Bouldin atau dbi yaitu: -0.627, Rata-rata dalam jarak pusat : -21880.851, Rata-rata dalam jarak pusat cluster 0 : -17186.250, Rata-rata dalam jarak pusat cluster 1 : -35766.667, Rata-rata dalam jarak pusat cluster 2 : -0.000, Rata-rata dalam jarak pusat cluster 3 : -17785.750, Rata-rata dalam jarak pusat cluster 4 : -0.000, Rata-rata dalam jarak pusat cluster 5 : -34019.944, Rata-rata dalam jarak pusat cluster 6 : -0.000, Rata-rata dalam jarak pusat cluster 7 : -0.000.

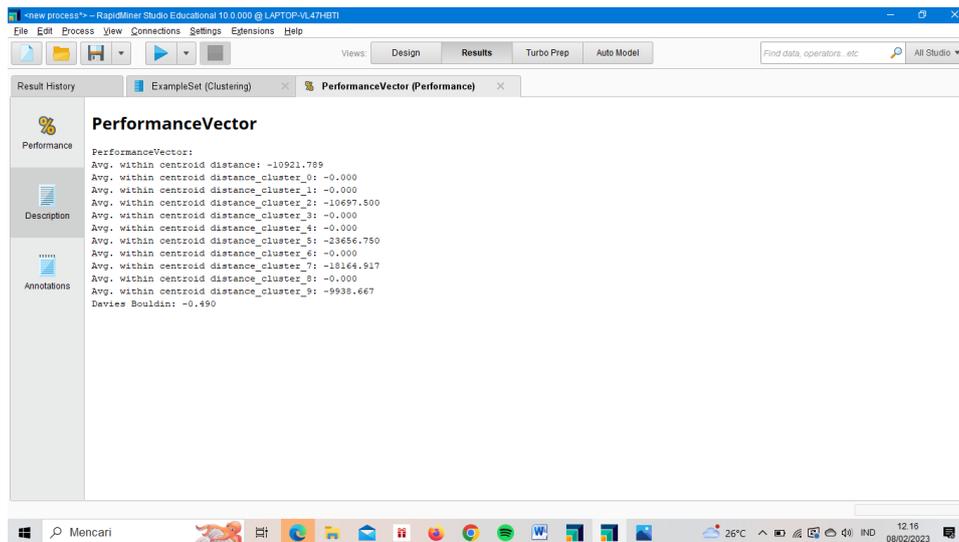
## 8. Nilai dbi cluster 9



Gambar 13 Nilai dbi cluster 9

Dari hasil pengklusteran cluster 9 terdapat Nilai Davies Bouldin atau dbi yaitu: -0.601, Rata-rata dalam jarak pusat : -15322.105, Rata-rata dalam jarak pusat cluster 0 : -0.000, Rata-rata dalam jarak pusat cluster 1 : -14332.000, Rata-rata dalam jarak pusat cluster 2 : -35766.667, Rata-rata dalam jarak pusat cluster 3 : -0.000, Rata-rata dalam jarak pusat cluster 4 : -0.000, Rata-rata dalam jarak pusat cluster 5 : -0.000, Rata-rata dalam jarak pusat cluster 6 : -0.000, Rata-rata dalam jarak pusat cluster 7 : -27448.000, Rata-rata dalam jarak pusat cluster 8 : -9938.667.

9. Nilai dbi cluster 10



Gambar 14 Nilai dbi cluster 10

Dari hasil pengklusteran cluster 10 terdapat Nilai Davies Bouldin atau dbi yaitu: -0.490, Rata-rata dalam jarak pusat : -10921.789, Rata-rata dalam jarak pusat cluster 0 : -0.000, Rata-rata dalam jarak pusat cluster 1 : -0.000, Rata-rata dalam jarak pusat cluster 2 : -10697.500, Rata-rata dalam jarak pusat cluster 3 : -0.000, Rata-rata dalam jarak pusat cluster 4 : -0.000, Rata-rata dalam jarak pusat cluster 5 : -23656.750, Rata-rata dalam jarak pusat cluster 6 : -0.000, Rata-rata dalam jarak pusat cluster 7 : -18164.917, Rata-rata dalam jarak pusat cluster 8 : -0.000, Rata-rata dalam jarak pusat cluster 9 : -9938.667.

Berdasarkan hasil uji coba mencari nilai n cluster K 2- 10 terdapat nilai optimum yaitu pada cluster 3 atau K 3 dengan nilai dbi yaitu -0,815, Rata-rata dalam jarak pusat : -74823.314, Rata-rata dalam jarak pusat cluster 0 : -138643.833, Rata-rata dalam jarak pusat cluster 1 : -43745.130, Rata-rata dalam jarak pusat cluster 2 : -50776.222.

**KESIMPULAN DAN SARAN**

**A. Kesimpulan**

Dari pengolahan data yang sudah dilakukan, diperoleh kesimpulan sebagai berikut:

1. Pengelompokan daerah bencana alam menggunakan algoritma k-means clustering dapat dikempokan menjadi 3 cluster yaitu cluster tinggi, sedang dan rendah. Cluster 0(tinggi) beranggotakan 6 daerah yaitu Kabupaten Bogor, Kabupaten Ciamis, Kabupaten Garut, Kabupaten Sukabumi dan Kabupaten Tasikmalaya, cluster 1(sedang) beranggotakan 10 daerah yaitu daerah Kabupaten Bandung, Kabupaten Bandung Barat, Kabupaten Bekasi, Kabupaten Kuningan, Kabupaten Majalengka, Kabupaten Pangandaran, Kabupaten Purwakarta, Kabupaten Subang, Kabupaten Sumedang dan Kota Banjar dan cluster 2(rendah)

beranggotakan 3 daerah yaitu Kabupaten Cirebon, Kabupaten Indramayu, dan Kabupaten Karawang.

2. Nilai dbi yang di dapat dari uji coba n nilai cluster 2-10 pada pengelompokan daerah bencana alam yang diambil yaitu pada cluster 3 dengan nilai dbi/ Davis Bouldin yaitu -0,815, Rata-rata dalam jarak pusat : -74823.314, Rata-rata dalam jarak pusat cluster 0 : -138643.833, Rata-rata dalam jarak pusat cluster 1 : -43745.130, Rata-rata dalam jarak pusat cluster 2 : -50776.222.

## B. Saran

Adapun saran peneliti usulkan untuk bahan pertimbangan bagi penelitian selanjutnya guna meningkatkan kualitas yang lebih baik yaitu :

1. Untuk penelitian selanjutnya perlu ditambahkan vaktor dan variable yang lebih baik sehingga lebih optimal dalam menganalisis cluster.
2. Untuk penelitian selanjutnya dapat menggunakan metode clustering atau klasifikasi lainnya sebagai perbandingan untuk hasil yang lebih akurat dan lebih baik.

## DAFTAR REFERENSI

- Asroni, A., Fitri, H., & Prasetyo, E. (2018). Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokan Data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus: Fakultas Kedokteran dan Ilmu Kesehatan, dan Fakultas Ilmu Sosial dan Ilmu Politik). *Semesta Teknika*, 21(1), 60–64. <https://doi.org/10.18196/st.211211>
- Dhuhita, W. (2015). Clustering Menggunakan Metode K-Mean Untuk Menentukan Status Gizi Balita. *Jurnal Informatika Darmajaya*, 15(2), 160–174.
- Fatmawati, K., & Windarto, A. P. (2018). Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi. *Computer Engineering, Science and System Journal*, 3(2), 173. <https://doi.org/10.24114/cess.v3i2.9661>
- Firman, M., Halik, A., & Septiana, L. (2022). Analisa Data Untuk Prediksi Daerah Rawan Bencana Alam Di Jawa Barat Menggunakan Algoritma K-Means Clustering. 6(4), 856–870. <https://doi.org/10.52362/jisamar.v6i4.939>
- Iqbal Ramadhan, M. (2017). Penerapan Data Mining Untuk Analisis Data Bencana Milik Bnpb Menggunakan Algoritma K-Means Dan Linear Regression. *Jurnal Informatika Dan Komputer*, 22(1), 57–65.
- Mohede, R. M., Rotinsulu, D. C., Tumangkang, S. Y. L., Pembangunan, J. E., Ekonomi, F., & Ratulangi, U. S. (2020). Analisis Kontribusi Serta Prediksi Pajak Dan Daerah Terhadap Peningkatan Pendapatan Asli Daerah (Pad) Di Kabupaten Kepulauan Sangihe. *Jurnal Berkala Ilmiah Efisiensi*, 20(01), 45–54.
- Murdiaty, M., Angela, A., & Sylvia, C. (2020). Pengelompokan Data Bencana Alam Berdasarkan Wilayah, Waktu, Jumlah Korban dan Kerusakan Fasilitas Dengan Algoritma K-Means. *Jurnal*

- Media Informatika Budidarma*, 4(3), 744. <https://doi.org/10.30865/mib.v4i3.2213>
- Puji Rahayu, Ika Anikah, Dias Bayu Saputra, Tri Anelia, & Martanto. (2020). Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Rotan. *KOPERTIP : Jurnal Ilmiah Manajemen Informatika Dan Komputer*, 4(2), 42–50. <https://doi.org/10.32485/kopertip.v4i2.118>
- Putu, N., Merliana, E., & Santoso, A. J. (n.d.). *Analisa Penentuan Jumlah Cluster Terbaik pada Metode K-Means*. 978–979.
- Ramdhan, D., Dwilestari, G., Dana, R. D., Ajiz, A., & Kaslani, K. (2022). Clustering Data Persediaan Barang Dengan Menggunakan Metode K-Means. *MEANS (Media Informasi Analisa Dan Sistem)*, 7(1), 1–9. <https://doi.org/10.54367/means.v7i1.1826>
- Surya, A. D., M.Sapriyaldi, Wanto, A., Windarto, A. P., & Damanik, I. S. (2021). Komparasi Algoritma Machine Learning untuk Penentuan Performance Terbaik Pada Prediksi Produksi Tanaman Jahe di Indonesia. *Seminar Nasional Ilmu Sosial Dan Teknologi (SANISTEK)*, 276–284.