



PENERAPAN DATA MINING UNTUK PEMETAAN DAERAH RAWAN BENCANA SEBAGAI UPAYA KESIAPSIAGAAN TERHADAP BENCANA

Vega Purwayoga¹⁾, Ali Astra Mikail²⁾, Salma Dewi Nur Faridah³⁾, Virra Retnowati A'izzah⁴⁾

^{1,2,3,4}Informatika, Universitas Siliwangi

^{1,2,3,4}Jl. Siliwangi No.24, Kahuripan, Tawang, Tasikmalaya, Jawa Barat 46115

Email: ¹vega.purwayoga@unsil.ac.id, ²217006103@student.unsil.ac.id, ³207006079@student.unsil.ac.id,
⁴207006020@student.unsil.ac.id

Abstract

Disasters have a major impact on several sectors, such as infrastructure, manufacturing, tourism and transportation. One way to prepare for or improve disaster preparedness is to implement preventive measures. Preventive actions can be taken by identifying disasters in each area from past data. This study aims to map areas affected by disasters to facilitate disaster preparedness programs. The data used in this research are areas of West Java that will be affected by the disaster in 2022 from January to October. The disaster data used in this study are floods, landslides, abrasion, tornadoes, droughts, fires, earthquakes and tsunamis. Research to use data mining techniques, namely grouping techniques. The clustering algorithm used in this study is the K-means cluster. The clustering process was carried out several times to find out the comparison of the quality of the grouping results which in this study used the Within Cluster Sum of Squares (WSS). The best WSS value is when the number of k or the number of clusters is 5, which is 89.8%. This research is expected to be a reference for disaster preparedness. This research also produced disaster grouping maps, where each cluster has different characteristics or types of disaster.

Keyword: Cluster, Disaster, K-Means, Visualization.

Abstrak

Bencana memberikan suatu dampak yang besar bagi beberapa sektor seperti, infrastruktur, manufaktur, pariwisata dan transportasi. Salah cara untuk mempersiapkan atau meningkatkan kesiapsiagaan bencana yaitu dengan menerapkan tindakan preventif dapat dilakukan dengan cara mengidentifikasi bencana setiap daerah dari data lampau. Penelitian ini bertujuan untuk memetakan daerah yang terdampak bencana untuk memudahkan dalam program kesiapsiagaan bencana. Data yang digunakan dalam penelitian ini daerah-daerah Jawa Barat yang terdampak bencana pada Tahun 2022 dari Bulan Januari sampai dengan Bulan Oktober. Data bencana yang digunakan dalam penelitian ini yaitu banjir, tanah longsor, abrasi, angin puting beliung, kekeringan, kebakaran, gempa bumi dan tsunami. Penelitian untuk menggunakan *teknik data mining* yaitu teknik pengelompokan. Algoritme pengelompokan yang digunakan dalam penelitian ini yaitu K-means cluster. Proses pengelompokan dilakukan beberapa kali percobaan untuk mengetahui perbandingan kualitas hasil pengelompokan yang dalam penelitian ini menggunakan *Within Cluster Sum of Squares* (WSS). Nilai WSS terbaik yaitu ketika jumlah k atau jumlah *cluster* berjumlah 5 yaitu sebesar 89.8 %. Penelitian ini diharapkan dapat menjadi referensi untuk kesiapsiagaan bencana. Penelitian ini juga menghasilkan peta pengelompokan bencana, dimana setiap *cluster* memiliki karakteristik atau jenis bencana yang berbeda.

Kata Kunci: Bencana, K-Means, Pengelompokan, Visualisasi.

1. PENDAHULUAN

Bencana memberikan dampak yang besar dan merugikan bagi masyarakat Indonesia. Tercatat pada periode 1815 sampai dengan 2019 jumlah kejadian bencana berjumlah banjir dengan total 10.438 kejadian, longsor sebanyak 6.050 kejadian, kekeringan 2.124 kejadian, serta kebakaran hutan dan lahan dengan total 1.914 kejadian. Tercatat bahwa setiap tahun kejadian bencana mengalami peningkatan [1]. Beberapa dampak bencana yang terjadi dapat mengakibatkan lumpuhnya perekonomian, infrastruktur, terganggunya jaringan komunikasi, wabah penyakit, gagal panen dan sebagainya



[2]. Sering terjadinya bencana dikarenakan salah satunya disebabkan oleh letak geografis negara Indonesia yang berada pada pertemuan 4 lempeng tektonik utama dunia yaitu Lempeng Eurasia, Indo-Australia, Pasifik dan Filipina [3].

Dalam mengurangi dampak yang di sebabkan oleh bencana, kesiapsiagaan merupakan salah satu faktor penting dalam menghadapi hal tersebut. Masyarakat sebagai elemen utama yang merasakan suatu bencana harus mempunyai kesiapsiagaan dan mitigasi dalam menghadapi bencana. Penanggulangan bencana memerlukan kecepatan dan ketetapan agar proses penanggulangan bencana berjalan sukses [4]. Kesiapsiagaan bencana menjadi strategi fundamental untuk mereduksi atau mengurangi semua kerentanan baik fisik maupun sosial ekonomi (Joni Sulistiawan) [4]. Bencana menjadi peristiwa yang sulit untuk diprediksi kedatangannya, sedangkan pengelolaan kesiapsiagaan yang hanya berbasis pada pendekatan pasca-bencana saja mulai kurang mendapat relevansinya, sehingga dibutuhkan manajemen bencana yang baik [5] [6].

Proses penanganan untuk kesiapsiagaan bencana telah dilakukan oleh beberapa penelitian sebelumnya. Pada penelitian telah melakukan analisis sebaran bencana yang terjadi di Kabupaten Purwakarta [7]. Dimana data kejadian lampau dapat digunakan untuk mengetahui sebaran bencana yang terjadi pada suatu wilayah [8]. Penelitian [9] telah mengelompokkan data bencana dengan menggunakan algoritme K-Means. Pengelompokkan data bencana diharapkan dapat berguna sebagai referensi atau dasar untuk prediksi kejadian bencana pada suatu daerah [9] [10]. Penelitian [11] telah mengelompokkan daerah resiko COVID-19 dengan K-Means. Pengelompokkan daerah yang memiliki potensi bencana dapat digunakan untuk rekomendasi bantuan kepada daerah yang berpotensi bencana tersebut [12].

Dengan memanfaatkan perkembangan teknologi informasi diharapkan dapat memprediksi titik bencana yang tersebar di seluruh wilayah Indonesia sebagai salah satu bentuk kesiapsiagaan terhadap potensi bencana di Indonesia. Salah satu teknologi yang bisa digunakan yaitu penggunaan metode penambahan data dalam mengumpulkan data atau informasi untuk pemetaan daerah rawan bencana [11]. Beberapa penelitian sebelumnya telah menerapkan K-Means untuk identifikasi bencana alam, namun belum ada yang memvisualisasikan peta bencana berdasarkan *cluster* atau kelompok bencana. Penelitian mencoba menerapkan K-Means dengan mengevaluasi jumlah *cluster* yang optimal dengan mengembangkan peta hasil pengelompokkan bencana. Peta bencana harapannya dapat digunakan untuk rekomendasi kesiapsiagaan bencana. Daerah yang termasuk *cluster* rawan bencana diharapkan dapat melakukan tindakan preventif untuk meminimalisir dampak yang diakibatkan bencana.

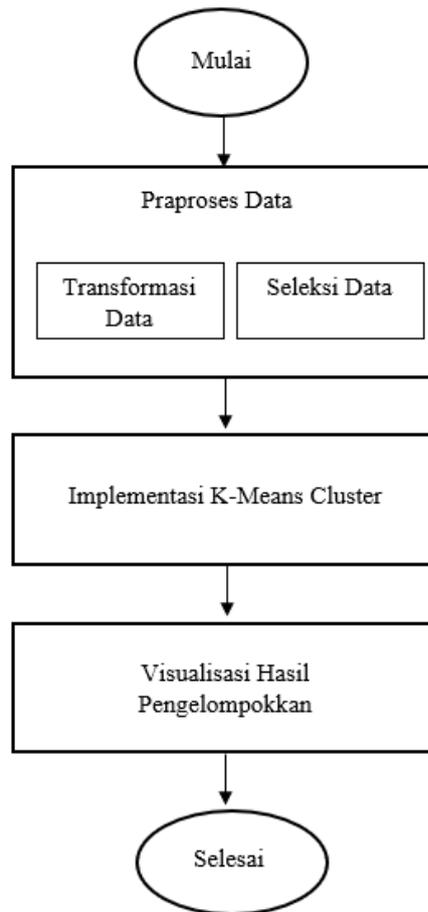
2. METODE PENELITIAN

2.1 Data dan Area Studi

Data yang digunakan dalam penelitian ini didapat dari laman data set katalog.data.go.id di mana data tersebut melingkupi banyak kejadian bencana di daerah Jawa Barat dengan id, nama provinsi, kode kabupaten kota, nama kabupaten kota, jenis bencana, jumlah kejadian dan tahun sebagai atribut pada data persebaran bencana tersebut. Dataset dan deskripsi atribut telah dijelaskan pada Tabel 1.

Tabel 1. Dataset

Nama atribut	Deskripsi atribut
id	Nilai numerik yang digunakan untuk kode pada suatu kejadian bencana yang berisikan nama provinsi, kode kabupaten atau kota, nama kabupaten atau kota, jenis bencana, jumlah kejadian dan tahun.
Nama provinsi	Nama provinsi yang mengalami atau terimbas kejadian bencana
Kode kabupaten kota	Kode dari suatu kabupaten atau kota
Nama kabupaten kota	Nama dari suatu kabupaten atau kota
jenisbencana	Ragam bencana yang melanda suatu daerah yang terdiri dari banjir, tanah longsor, abrasi, angin puting beliung, kekeringan, kebakaran hutan, gempa bumi dan tsunami.
Jumlah kejadian	Banyaknya kejadian bencana yang terjadi pada area atau daerah tersebut.



Gambar 1. Tahapan penelitian

2.2 Praproses Data

Praproses data merupakan salah satu bagian terpenting dalam proses analisis data termasuk penerapan *data mining* [13]. Proses meringkas suatu data adalah cara mengetahui informasi secara umum dari suatu data. Seperti dimensi, karakteristik nilai dari suatu atribut, nilai rata-rata, nilai minimum, nilai maksimum dan nilai tengah dari suatu atribut. Transformasi dalam penelitian ini yaitu mentranspose data yang awalnya data tersebut adalah nilai dari suatu atribut akan diubah menjadi suatu atribut. Pada penerapan teknik data mining terkadang dibutuhkan atribut yang dapat mempengaruhi pengambilan pengetahuan, sehingga bagi atribut yang tidak memiliki pengaruh tidak akan dilibatkan pada proses analisis dengan menggunakan teknik *data mining* [13]. Proses meringkas suatu data adalah cara mengetahui informasi secara umum dari suatu data. Seperti dimensi, karakteristik nilai dari suatu atribut, nilai rata-rata, nilai minimum, nilai maksimum dan nilai tengah dari suatu atribut.

2.3 Penerapan K-Means Cluster

Penerapan K-Means *cluster* dilakukan dengan menggunakan R Programming. R Programming merupakan bahasa pemrograman tingkat tinggi yang mudah diterapkan untuk teknik *data mining* [14]. K-Means adalah satu algoritma clustering yang populer. Cara kerja K-Means yaitu mempartisi data yang ke beberapa kelompok berdasarkan karakteristik data tersebut. Objek pertama dalam sebuah cluster dapat digunakan sebagai titik pusat cluster. Algoritma K-Means akan mengulangi tahapan-tahapan sebagai berikut sampai titik tengah suatu cluster tidak berubah [15]. Berikut tahapan K-Means *cluster*

- a. Menentukan titik tengah pada setiap cluster, titik tengah *cluster* ditentukan secara acak.
- b. Menghitung jarak setiap objek dengan titik tengah suatu cluster dengan *euclidean distance* [15].
- c. Memasukkan objek ke dalam cluster berdasarkan jarak minimumnya.



2.4 Visualisasi Pengelompokkan

Visualisasi diharapkan dapat memberikan penjelasan lebih rinci terkait cluster yang berisikan suatu daerah. Setiap kelompok terbentuk dari daerah-daerah yang memiliki kemiripan karakteristik bencana. Visualisasi dalam penelitian ini berbentuk grafik dan peta dari cluster bencana.

3. HASIL DAN PEMBAHASAN

3.1 Praproses Data

Proses perangkuman data pada penelitian ini yaitu untuk mengetahui dimensi dari data bencana yang telah didapat. Penghitungan dimensi dilakukan dengan menggunakan fungsi dim. Dengan menggunakan fungsi dim didapat bahwa terdapat 216 baris dan 4 kolom. Perangkuman data juga digunakan untuk melihat atribut apa saja yang terdapat pada data bencana di antaranya, id, nama provinsi, kode kabupaten, nama kabupaten, jenis bencana, jumlah kejadian dan tahun.

Tahapan selanjutnya adalah transformasi data. Sebelum proses transformasi data atribut yang tersaji terdiri dari id, nama provinsi, kode kabupaten, nama kabupaten, jenis bencana, jumlah kejadian dan tahun, sebagaimana yang disajikan pada Tabel 1. Setelah dilakukan transformasi data maka bentuk data sebagaimana yang dapat dilihat pada Tabel 2. Jenis bencana dirubah menjadi kolom untuk mengetahui jumlah setiap bencana pada setiap kabupaten/kota. Berdasarkan data yang didapat dari 680 Bencana yang melanda kabupaten dan kota yang berada di provinsi Jawa Barat yaitu 17% bencana banjir, 38% tanah longsor, 0% abrasi, 43% angin puting beliung, kekeringan 0%, 0% kebakaran, 2% gempa bumi, 0% tsunami sebagaimana yang telah disajikan pada Gambar 2.

Setelah transformasi data dilakukakan tahapan selanjutnya yaitu memilih atribut yang berpengaruh pada proses pengelompokkan daerah rawan bencana. Proses seleksi data dalam hal ini yaitu mengambil beberapa atribut diantaranya nama kabupaten dan jumlah kejadian dari setiap bencana yang telah disajikan pada Tabel 3.

Tabel 2. Hasil praproses data

id	Nama provinsi	Kode kab/kota	Nama kab/kota	a	b	c	d	e	f	g
1	Jawa Barat	3210	Kabupaten Bogor	9	53	0	67	2	0	0
2	Jawa Barat	3202	Kabupaten Sukabumi	12	49	0	38	0	0	6
3	Jawa Barat	3203	Kabupaten Cianjur	5	13	0	5	0	0	1
...
25	Jawa Barat	3277	Kota Cimahi	6	3	0	2	0	0	0
26	Jawa Barat	3278	Kota Tasikmalaya	1	1	0	3	0	0	0
27	Jawa Barat	3279	Kota Banjar	0	0	0	3	0	0	0

*Deskripsi atribut

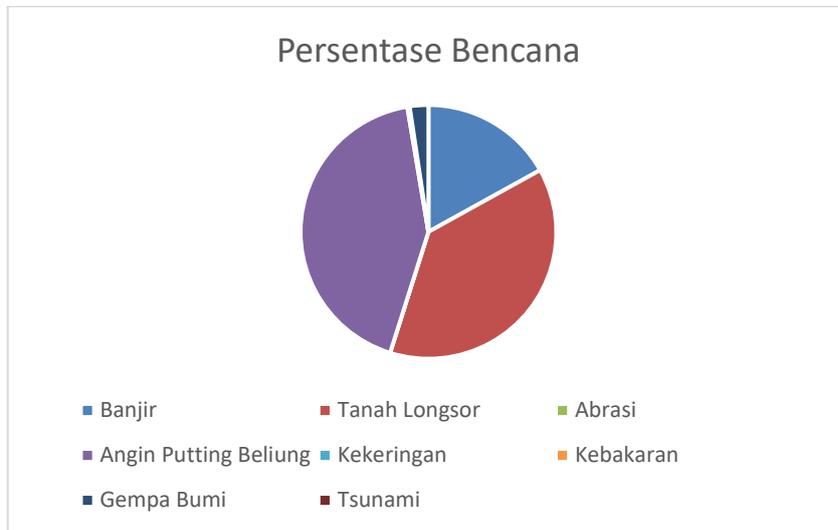
- a: Banjir
- b: Tanah Longsor
- c: Angin Puting Beliung
- d: Kekeringan
- e: Kebakaran
- f: Gempa Bumi
- g: Tsunami

Tabel 3. Hasil seleksi data

id	Nama kab/kota	a	b	c	d	e	f	g
1	Kabupaten Bogor	9	53	0	67	2	0	0
2	Kabupaten Sukabumi	12	49	0	38	0	0	6
3	Kabupaten Cianjur	5	13	0	5	0	0	1



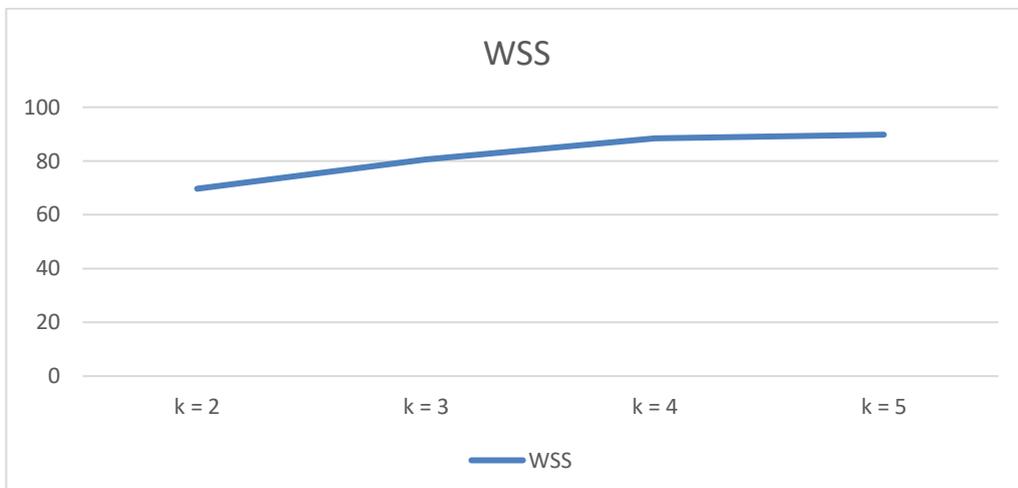
...
25	Kota Cimahi	6	3	0	2	0	0
26	Kota Tasikmalaya	1	1	0	3	0	0
27	Kota Banjar	0	0	0	3	0	0



Gambar 2. Persentase Bencana

3.2 Penerapan K-Means Cluster

Penerapan K-Means dilakukan dengan menggunakan *library cluster*. Percobaan pertama data akan dikelompokkan menjadi dua kelompok, selanjutnya jumlah kelompok akan ditingkatkan untuk membandingkan kualitas hasil pengelompokan. Ketika data dikelompokkan menjadi dua kelompok nilai *within cluster sum of squares* (WSS) nya adalah 69.7%. Nilai WSS dengan jumlah $k = 3$ sampai dengan $k = 5$ berturut-turut adalah 80.6 %, 88.4 % dan 89.8 %. Dapat dilihat pada Gambar 3 bahwa dalam kasus ini, semakin banyak jumlah k maka semakin besarn nilai WSSnya. Sebagian hasil pengelompokan yang telah dihasilkan ketika jumlah $k = 5$ dapat dilihat pada Tabel 4. Dimana setiap *cluster* memiliki kriteria yang dapat dilihat pada Tabel 5. Pada *cluster* 1 beranggotan daerah yang rata-rata bencana $a = 4.25$, bencana $b = 13.37$, bencana $c = 0$, bencana $d = 6.37$, bencana $e = 0$ dan bencana $f = 0$ dan bencana $g = 0.13$. Karakteristik *cluster* 2 sampai dengan 5 telah dijelaskan sebagaimana yang ditampilkan pada Tabel 5.



Gambar 3. Visualisasi nilai WSS



Tabel 4. Hasil pengelompokkan

id	Nama kab/kota	Cluster = 2	Cluster = 3	Cluster = 4	Cluster = 5
1	Kabupaten Bogor	2	1	1	3
2	Kabupaten Sukabumi	2	1	1	3
3	Kabupaten Cianjur	1	2	4	1
4	Kabupaten Bandung	1	2	4	4
5	Kabupaten Garut	1	2	4	1
6	Kabupaten Tasikmalaya	1	3	2	5
7	Kabupaten Ciamis	1	2	3	2
8	Kabupaten Kuningan	1	2	4	1
9	Kabupaten Cirebon	1	3	2	4
10	Kabupaten Majalengka	1	2	4	1
11	Kabupaten Sumedang	1	2	4	1
12	Kabupaten Indramayu	1	3	2	5
13	Kabupaten Subang	1	2	3	2
14	Kabupaten Purwakarta	1	3	2	1
15	Kabupaten Karawang	1	2	3	2
16	Kabupaten Bekasi	1	3	2	5
17	Kabupaten Bandung Barat	1	2	4	1
18	Kabupaten Pangandaran	1	3	2	5
19	Kota Bogor	1	3	2	1
20	Kota Sukabumi	1	3	2	5
21	Kota Bandung	1	3	2	5
22	Kota Cirebon	1	3	2	5
23	Kota Bekasi	1	3	2	5
24	Kota Depok	1	3	2	5
25	Kota Cimahi	1	3	2	5
26	Kota Tasikmalaya	1	3	2	5
27	Kota Banjar	1	3	2	3

Tabel 5. Karakteristik setiap cluster

Cluster	a	b	c	d	e	f	g
1	4.25	13.37	0	6.37	0	0	0.13
2	3.33	7	0	25.33	0	0	0
3	10.5	51	0	52.5	1	0	3
4	13.5	6.5	0	11	0	0	1
5	1.92	1.26	0	2.92	0	0	0.59

3.3 Visualisasi Hasil Pengelompokkan

Setelah proses pengelompokkan dengan K-Means telah dilakukan tahapan selanjutnya adalah visualisasi hasil pengelompokkan. Visualisasi dalam hal ini bertujuan untuk mempermudah pembacaan daerah mana yang termasuk ke dalam cluster 1 dan seterusnya. Telah disajikan visualisasi hasil pengelompokkan dari jumlah cluster = 2 sampai dengan cluster = 5, berurutan dari Gambar 4 sampai dengan Gambar 7.



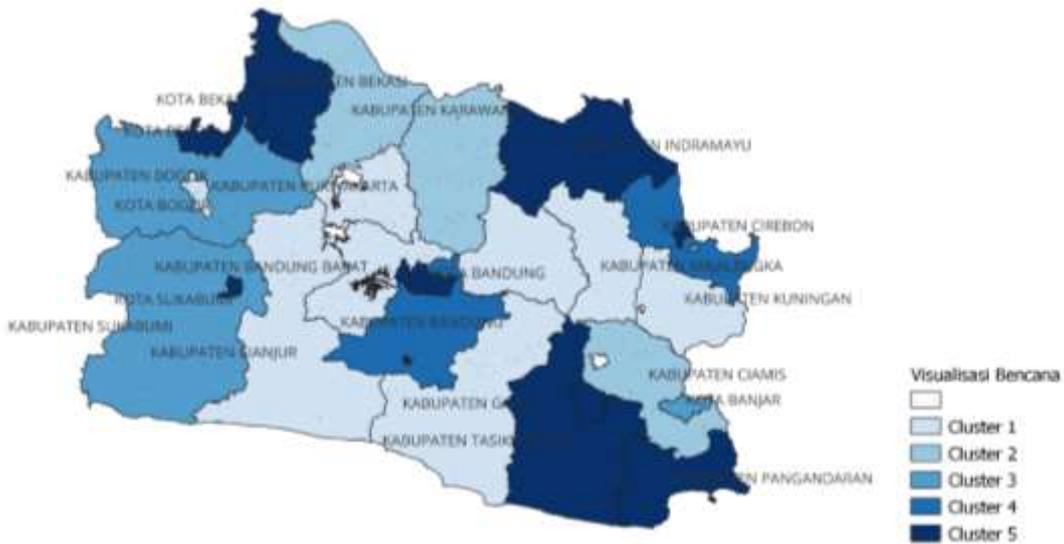
Gambar 4. Visualisasi hasil pengelompokan jumlah *cluster* 2



Gambar 5. Visualisasi hasil pengelompokan jumlah *cluster* 3



Gambar 6. Visualisasi hasil pengelompokan jumlah cluster 4



Gambar 7. Visualisasi hasil pengelompokan jumlah cluster 5

4. KESIMPULAN

Penerapan K-Means untuk pemetaan daerah rawan bencana telah berhasil dilakukan. Penelitian ini diselesaikan dengan beberapa tahapan diantaranya, praproses data, penerapan K-Means, dan yang terakhir pemetaan hasil pengelompokan. Jumlah cluster dengan nilai WSS terbaik yaitu berjumlah 5, dimana setiap cluster memiliki karakteristik yang berbeda. Penelitian ini juga telah menghasilkan suatu visualisasi hasil pengelompokan. Visualisasi ditampilkan ke dalam bentuk peta untuk memudahkan dalam proses identifikasi daerah berdasarkan cluster atau kelompoknya. Penelitian ini diharapkan dapat digunakan untuk sebagai referensi untuk penanganan dan kesiapsiagaan bencana. Peta pengelompokan dapat dijadikan sebagai bahan kajian untuk melihat tren bencana dari setiap daerah.



UCAPAN TERIMAKASIH

Terima kasih disampaikan kepada pihak-pihak yang telah mendukung terlaksananya penelitian ini.

DAFTAR PUSTAKA

- [1] S. Yulianto, R. K. Apriyadi, A. Aprilyanto, T. Winugroho, I. S. Ponangsera, and W. Wilopo, "Histori Bencana dan Penanggulangannya di Indonesia Ditinjau Dari Perspektif Keamanan Nasional," *PENDIPA Journal of Science Education*, vol. 5, no. 2, pp. 180–187, Jan. 2021, doi: 10.33369/pendipa.5.2.180-187.
- [2] D. Dwi Utomo, F. Yul Dewi Marta, and K. Sumatera Barat, "Dampak Bencana Alam Terhadap Perekonomian Masyarakat di Kabupaten Tanah Datar conditions of the Creative Commons Attribution-ShareAlike 4.0 International License-(CC-BY-SA)," *Jurnal Terapan Pemerintahan Minangkabau*, vol. 2, no. 1, pp. 92–97, 2022, doi: 10.33701/jtpm.v2i1.2395.
- [3] M. Ardhyansyah Agung, R. Achmad, and H. Pasongli, "KESIAPSIAGAAN MASYARAKAT DALAM MENGHADAPI BENCANA GEMPA BUMI DI DESA BOBANEHENA KECAMATAN JAILOLO KABUPATEN HALMAHERA BARAT."
- [4] J. Sulistiawan, "Sinergitas Tni Dan Bnpb Dalam Penanggulangan Bencana Gempa Bumi Dan Tsunami Di Palu Guna Meningkatkan Pertahanan Nirmiliter Dalam Rangka Memperkuat Pertahanan Negara", *Jurnal Inovasi Penelitian*, vol. 2, no. 8, pp. 2841-2848, 2022.
- [5] R. Djalante and F. Thomalla, "Disaster risk reduction and climate change adaptation in Indonesia: Institutional challenges and opportunities for integration," *International Journal of Disaster Resilience in the Built Environment*, vol. 3, no. 2, pp. 166–180, Jul. 2012. doi: 10.1108/17595901211245260.
- [6] S. Nurillah, D. Maulana, and B. Hasanah, "Manajemen Mitigasi Penanggulangan Bencana Banjir Oleh Badan Penanggulangan Bencana Daerah (BPBD) Kota Cilegon di Kecamatan Ciwandan," *JDKP Jurnal Desentralisasi dan Kebijakan Publik*, vol. 3, no. 1, pp. 334–350, Mar. 2022, doi: 10.30656/jdkp.v3i1.4613.
- [7] T. I. Hermanto *et al.*, "Analisis Sebaran Titik Rawan Bencana dengan K-Means Clustering dalam Penanganan Bencana," Teguh Iman Hermanto, 2021.
- [8] D. Isya Ramadhani, O. Damayanti, O. Thaushiyah, and A. Rahman Kadafi, "Penerapan Metode K-Means Untuk Clustering Desa Rawan Bencana Berdasarkan Data Kejadian Terjadinya Bencana Alam," *Jurnal Riset Komputer*, vol. 9, no. 3, pp. 2407–389, 2022, doi: 10.30865/jurikom.v9i3.4326.
- [9] M. Murdiaty, A. Angela, and C. Sylvia, "Pengelompokan Data Bencana Alam Berdasarkan Wilayah, Waktu, Jumlah Korban dan Kerusakan Fasilitas Dengan Algoritma K-Means," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 4, no. 3, p. 744, Jul. 2020, doi: 10.30865/mib.v4i3.2213.
- [10] F. Reviantika, C. N. Harahap, and Y. Azhar, "Analisis Gempa Bumi Pada Pulau Jawa Menggunakan Clustering Algoritma K-Means Analysis Earthquake in Java Island Using Clustering K-Means Algorithm," *Jurnal Dinamika Informatika*, vol. 9, no. 1, 2020, [Online]. Available: <https://twitter.com/infobmkg>
- [11] H. Gunawan and V. Purwayoga, "Data Mining Menggunakan Algoritma K-Means Clustering Untuk Mengetahui Potensi Penyebaran Virus Corona Di Kota Cirebon," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 11, no. 1, pp. 1–8, Jan. 2022, doi: 10.32736/sisfokom.v11i1.1316.
- [12] V. Purwayoga, "Modified skyline query to measure priority region for personal protective equipment recipient of COVID-19 health workers," *Jurnal Teknologi dan Sistem Komputer*, vol. 9, no. 3, pp. 167–173, Jul. 2021, doi: 10.14710/jtsiskom.2021.14003.
- [13] S. A. Alasadi and W. S. Bhaya, "Review of data preprocessing techniques in data mining," *Journal of Engineering and Applied Sciences*, vol. 12, no. 16, pp. 4102–4107, Sep. 2017, doi: 10.3923/jeasci.2017.4102.4107.
- [14] S. Praviilovic, "R Language in Data Mining Techniques and Statistics," *American Journal of Software Engineering and Applications*, vol. 2, no. 1, p. 7, 2013, doi: 10.11648/j.ajsea.20130201.12.
- [15] V. Purwayoga and B. Susanto, "Pengelompokan Daerah Berdasarkan Ketersediaan Masjid Muhammadiyah Dengan Algoritma K-Means," vol. 13, no. 1, 2021, doi: 10.24853/jurtek.13.1.75-80.