

**CLUSTERING BENCANA ALAM DI INDONESIA MENGGUNAKAN
ALGORITMA K-MEANS**

SKRIPSI



Disusun Oleh:

ERA OKTAVIANA

NIM.H02215003

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS NEGERI SUNAN AMPEL SURABAYA**

2022

PERNYATAAN KEASLIAN

Saya yang bertanda tangan dibawah ini,

Nama : Era Oktaviana

NIM : H02215003

Program Studi : Matematika

Angkatan : 2015

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul: “*CLUSTERING BENCANA ALAM DI INDONESIA MENGGUNAKAN ALGORITMA K-MEANS*”. Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 20 Juni 2022

Yang Menyatakan,



Era Oktaviana

NIM. H02215003

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

NAMA : Era Oktaviana

NIM : H02215003

JUDUL : *CLUSTERING* BENCANA ALAM DI INDONESIA
MENGUNAKAN ALGORITMA *K-MEANS*

Ini telah diperiksa dan disetujui untuk diujikan.

Surabaya, 23 Juni 2022

Dosen Pembimbing 1



(Aris Fanani, M.Kom)

NIP : 198701272014031002

Dosen Pembimbing 2



(Putroue Keumala Intan, M.Si)

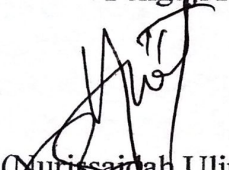
NIP : 198805282018012001

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi Era Oktaviana ini telah dipertahankan di depan
tim penguji skripsi di Surabaya, 12 Juli 2022

Mengesahkan,
Dewan Penguji

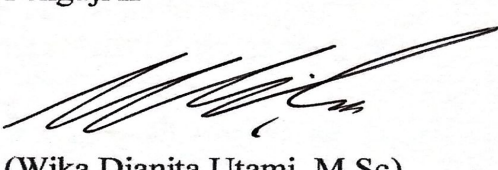
Penguji I



(Nuri Saidah Ulinnuha, M.Kom)

NIP : 199011022014032004

Penguji II



(Wika Dianita Utami, M.Sc)

NIP : 199206102018012003

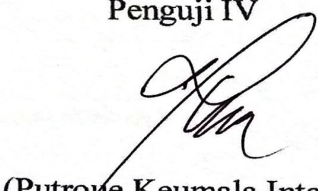
Penguji III



(Aris Fanani, M.Kom)

NIP : 198701272014031002

Penguji IV



(Putrone Keumala Intan, M.Si)

NIP: 198805282018012001

Mengetahui,

Dekan Fakultas Sains dan Teknologi

UIN Sunan Ampel Surabaya



Hamdani, M.Pd

NIP. 196507312000031002

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : Era Oktaviana
NIM : H02215003
Fakultas/Jurusan : Sains dan Teknologi/Matematika
E-mail address : eraoktavia@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

☒ Sekripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)

yang berjudul :

CLUSTERING BENCANA ALAM DI INDONESIA MENGGUNAKAN ALGORITMA

K-MEANS

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara **fulltext** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 22 Juli 2022

Penulis



(Era Oktaviana)
nama terang dan tanda tangan

ABSTRAK

CLUSTERING BENCANA ALAM DI INDONESIA MENGGUNAKAN

ALGORITMA K-MEANS

Bencana alam merupakan kejadian luar biasa yang disebabkan oleh faktor alam ataupun faktor dari ulah manusia yang berdampak pada lingkungan dan manusia itu sendiri. Indonesia menjadi salah satu negara yang menjadi rawan bencana alam seperti tanah longsor, banjir, banjir bandang, gempa bumi, tsunami, kekeringan, kebakaran, gunung meletus, puting beliung dan gelombang pasang laut karena keadaan iklim, tanah, hidrologi, geologi, serta geomorfologinya. Letak astronomis menjadikan Indonesia mempunyai iklim tropis yang berdampak pada curah hujan yang tinggi. Hasil pencatatan Badan Pusat Statistik (BPS) Indonesia tahun 2018 membuktikan bahwa sejak tiga tahun terakhir sudah lebih dari 20.000 kasus bencana alam terjadi diseluruh wilayah yang ada di Indonesia. BNPB adalah salah satu lembaga pemerintah yang bertugas menjadi pelaksana penanggulangan bencana alam di Indonesia yang memerlukan data yang kuat, tepat, dan akurat secara cepat ketika terjadi suatu bencana alam. Sehingga diperlukan pengelompokan provinsi-provinsi mana saja yang sering terdampak bencana alam. Metode yang digunakan dalam pengelompokan tersebut adalah K-Means *Clustering* yang di analisis menggunakan *Silhouette Index*. Berdasarkan pengelompokan tersebut, Indonesia terbagi menjadi 3 kelompok daerah rawan bencana alam dengan karakteristik rawan, sedang (cukup rawan), dan aman (tidak rawan). 3 kelompok tersebut sudah tervalidasi keakuratannya. *Cluster* yang dibentuk cukup optimal karena telah dilakukan uji validasi *cluster* menggunakan *Silhouette Index* dimana hasilnya menunjukkan 3 *cluster* yang optimal dengan nilai *Silhouette Index* sebesar 0.501.

Kata Kunci : *Clustering, K-Means, Silhouette Index*

UIN SUNAN AMPEL
S U R A B A Y A

ABSTRACT

NATURAL DISASTER CLUSTERING IN INDONESIA USING K-MEANS ALGORITHM

Natural disasters are extraordinary events caused by natural factors or human-made factors that have an impact on the environment and humans themselves. Indonesia is one of the countries prone to natural disasters such as landslides, floods, flash floods, earthquakes, tsunamis, droughts, fires, volcanic eruptions, cyclones and tidal waves due to climatic conditions, soil, hydrology, geology, and geomorphology. The astronomical location makes Indonesia have a tropical climate which has an impact on high rainfall. The results of the 2018 Indonesian Central Statistics Agency (BPS) records prove that since the last three years more than 20,000 cases of natural disasters have occurred in all regions in Indonesia. BNPB is one of the government agencies tasked with being the executor of natural disaster management in Indonesia which requires strong, precise, and accurate data quickly when a natural disaster occurs. So it is necessary to classify which provinces are often affected by natural disasters. The method used in the grouping is K-Means Clustering which is analyzed using the Silhouette Index. Based on these groupings, Indonesia is divided into 3 groups of areas prone to natural disasters with the characteristics of prone, moderate (quite vulnerable), and safe (not vulnerable). The 3 groups have been validated for accuracy. The cluster formed is quite optimal because a cluster validation test has been carried out using the Silhouette Index where the results show 3 optimal clusters with a Silhouette Index value of 0.501.

Keywords: Clustering, K-Means, Silhouette Index

UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR ISI

PENGESAHAN TIM PENGUJI SKRIPSI	iii
PERNYATAAN KEASLIAN	iv
MOTTO	v
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR	vii
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah	5
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	6
BAB II KAJIAN PUSTAKA	7
2.1 Bencana Alam	7
2.1.1 Tsunami	8
2.1.2 Gempa Bumi	8
2.1.3 Tanah Longsor	8
2.1.4 Angin Puting Beliung	9
2.1.5 Banjir	9
2.1.6 Banjir Bandang	10
2.1.7 Kekeringan	10
2.1.8 Kebakaran Hutan	11
2.1.9 Gelombang Pasang Laut	12

2.1.10	Gunung Meletus.....	12
2.2	Laju Pertumbuhan Penduduk	13
2.3	Curah Hujan	14
2.4	Normalisasi Data	15
2.5	Uji Multikolinearitas.....	16
2.6	<i>Clustering</i>	17
2.7	Algoritma K- Means	19
2.8	<i>Silhoutte Index (SI)</i>	20
BAB III METODE PENELITIAN.....		22
3.1	Sumber Data.....	22
3.2	Struktur Data dan Variabel Penelitian	23
3.3	Langkah Analisis	24
3.4	Perancangan <i>system</i>	24
BAB IV HASIL DAN PEMBAHASAN		27
4.1	Deskripsi Data.....	27
4.2	Hasil Normalisasi Data	33
4.3	Hasil Uji Multikolinearitas.....	35
4.4	Hasil <i>Clustering</i>	37
4.5	Hasil Analisis Algoritma K-Means	38
BAB V PENUTUP		41
5.1	Simpulan.....	41
5.2	Saran	41
DAFTAR PUSTAKA		42
LAMPIRAN		

DAFTAR GAMBAR

Gambar 3.1 Diagram Alir Penelitian.....	23
Gambar 4.1 Grafik Bencana Tanah Longsor.....	25
Gambar 4.2 Grafik Bencana Banjir.....	26
Gambar 4.3 Grafik Bencana Banjir Bandang.....	27
Gambar 4.4 Grafik Bencana Gempa Bumi.....	27
Gambar 4.5 Grafik Bencana Tsunami.....	28
Gambar 4.6 Grafik Bencana Gelombang Pasang Laut.....	28
Gambar 4.7 Grafik Bencana Puting Beliung.....	29
Gambar 4.8 Grafik Bencana Gunung Meletus.....	30
Gambar 4.9 Grafik Bencana Kebakaran Hutan.....	31
Gambar 4.10 Grafik Bencana Kekeringan.....	31
Gambar 4.11 Hasil <i>Cluster</i>	36

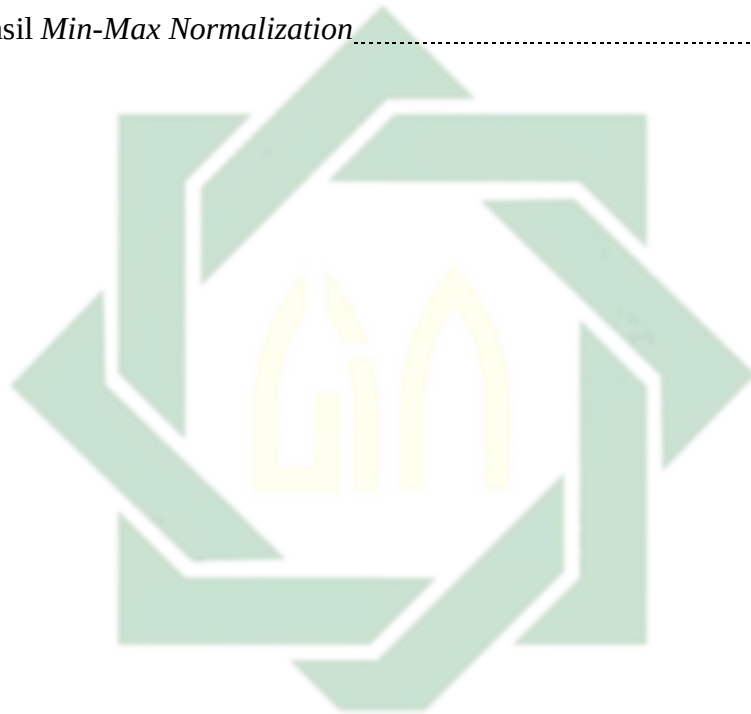
DAFTAR TABEL

Tabel 2.1 Interval <i>Silhouette</i>	18
Tabel 3.1 Data Bencana Alam di Indonesia.....	19
Tabel 3.2 Definisi Operasional Bencana Alam, Curah Hujan, dan Laju Pertumbuhan Penduduk.....	20
Tabel 4.1 Nilai Minimum dan Maksimum Data.....	32
Tabel 4.2 Hasil <i>Min-Max Normalization</i>	32
Tabel 4.3 Hasil Uji Multikolinearitas Coefficients.....	33
Tabel 4.4 Hasil Uji Multikolinearitas Coefficients Setelah Proses Penghapusan 3 Variabel.....	34
Tabel 4.5 Nilai <i>Silhouette</i>	35
Tabel 4.6 Hasil Perhitungan <i>Centroid</i>	35
Tabel 4.7 Hasil <i>Cluster</i> 34 Provinsi di Indonesia.....	36

UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR LAMPIRAN

Lampiran 1 Sampel Data Bencana Alam, Laju Pertumbuhan Penduduk (LPP), dan data Curah Hujan pada Tahun 2021.....	44
Lampiran 2 Hasil <i>Min-Max Normalization</i>	48



UIN SUNAN AMPEL
S U R A B A Y A

BAB I

PENDAHULUAN

1.1 Latar Belakang

Bencana alam merupakan kejadian luar biasa yang dapat mengancam dan mengganggu kehidupan serta penghidupan masyarakat yang disebabkan oleh faktor alam ataupun faktor dari ulah manusia yang berdampak pada lingkungan dan manusia itu sendiri. Sehingga menimbulkan kerusakan lingkungan, kerugian harta benda, dampak psikologis, dan bahkan menimbulkan korban jiwa. Seperti gempa bumi yang terjadi akibat tekanan yang disebabkan oleh lempengan yang bergerak. Banjir akibat dari penebangan pohon tanpa izin dan tingginya curah hujan. Serta tanah longsor akibat dari hutan yang di alih fungsikan menjadi lahan perumahan atau gedung lain. Sesuai dengan ayat Al- Qur'an surah Ar-Rum: 41

ظَهَرَ الْفَسَادُ فِي الْبَرِّ وَالْبَحْرِ بِمَا كَسَبَتْ أَيْدِي النَّاسِ لِيُذِيقَهُمْ بَعْضَ الَّذِي عَمِلُوا لَعَلَّهُمْ يَرْجِعُونَ (41)

“Telah nampak kerusakan di darat dan di laut disebabkan karena perbuatan tangan manusia, supaya Allah merasakan kepada mereka sebagian dari (akibat) perbuatan mereka, agar mereka kembali (ke jalan yang benar)”.

Manusia dituntut agar lebih memperhatikan, menyayangi, merawat dan menghormati lingkungannya. Karena pada dasarnya sumber daya alam dan lingkungan diciptakan untuk manusia, diberikan hak untuk memanfaatkan sumber daya alam tersebut dengan sewajarnya. Sebagaimana sabda Nabi Muhammad SAW.

"...ارْحَمُوا مَنْ فِي الْأَرْضِ يَرْحَمَكُم مَّنْ فِي السَّمَاءِ..."

“sayangilah makhluk yang ada di bumi, niscaya makhluk yang ada di langit akan menyayangi kalian...” (*HR At Tirmidzi*).

Indonesia merupakan salah satu negara yang banyak memiliki keindahan dan keberagaman alam seperti bukit, gunung, pantai, danau, dan laut yang dapat dinikmati oleh semua makhluk hidup khususnya manusia. Namun, keindahan dan keragaman tersebut dapat menyebabkan terjadinya bencana alam di Indonesia. Indonesia menjadi salah satu negara yang menjadi rawan bencana alam karena keadaan iklim, tanah, hidrologi, geologi, serta geomorfologinya. Letak astronomis menjadikan Indonesia mempunyai iklim tropis yang berdampak pada curah hujan yang tinggi. Apabila keadaan alam mulai rusak, maka curah hujan yang tinggi akan menyebabkan bencana banjir, ataupun longsor. Sedangkan secara geologi, Indonesia berada di antara tiga lempeng tektonik, yaitu lempeng Eurasia, lempeng Pasifik, dan lempeng Indo-Australia. Sehingga pertemuan antara tiga lempeng tersebut dapat menyebabkan Indonesia rawan akan terjadinya gempa bumi, tsunami, dan bencana lainnya (Muhammad & Prihandoko, 2017).

Tingginya curah hujan di Indonesia karena memiliki iklim tropis yang memiliki posisi lintang antara $0^{\circ} - 23\frac{10}{2}^{\circ}$ LU dan $0^{\circ} - 23\frac{10}{2}^{\circ}$ LS sehingga pengaruh dari iklim tropis menjadikan Indonesia memiliki curah hujan yang tinggi atau relatif banyak. Curah hujan yang tinggi mengakibatkan jumlah air yang meningkat. Air yang meluap atau menggenang dapat menyebabkan banjir dan banjir bandang. Erosi atau rusaknya tanah di daerah gunung semakin memudahkan terjadinya banjir juga dapat menyebabkan tanah longsor (Kemdikbud, 2021). Selain curah hujan yang tinggi berpengaruh terhadap bencana, laju pertumbuhan penduduk yang tinggi juga mempengaruhi terjadinya bencana alam.

Kepala badan Kependudukan dan Keluarga Berencana Nasional (BKKBN) mengatakan penduduk yang padat biasanya mengakibatkan rumah penduduk yang berhimpitan tanpa memikirkan resapan air. Kepadatan penduduk cenderung berdampak pada permukiman yang tidak ramah lingkungan. Akibat dari kepadatan penduduk juga menyebabkan fungsi lahan seperti hutan yang dialihkan menjadi lahan pemukiman, perkantoran dan sebagainya tanpa diikuti dengan upaya pelestarian lingkungan yang baik (Putra, 2021).

Badan Nasional Penanggulangan Bencana (BNPB) adalah salah satu lembaga pemerintah yang bertugas menjadi pelaksana penanggulangan bencana alam di Indonesia. BNPB memerlukan data yang kuat, tepat, dan akurat secara cepat ketika terjadi suatu bencana alam. Hal ini diperlukan agar lembaga yang terkait mampu melakukan penanggulangan bencana secara cepat namun tetap terkoordinasi dengan baik. Untuk merealisasikan hal tersebut dibutuhkan suatu alat atau analisis data yang dapat membantu proses pengumpulan informasi salah satunya yaitu merealisasikan penelitian ini dengan mitigasi bencana. Mitigasi bencana merupakan upaya untuk mengurangi dampak yang terjadi akibat suatu bencana termasuk kesiapan dalam menghadapi resiko jangka panjang. Penelitian kebencanaan tersebut dapat membantu BNPB dalam pengelompokan provinsi-provinsi mana saja yang sering terdampak bencana alam (BNPB, 2019).

Clustering merupakan suatu metode pengelompokan suatu dokumen dimana dokumen ini dikelompokkan dengan konten untuk mengurangi ruang pencarian yang diperlukan dalam merespon suatu *query* (Grossman, David A. , dan Ophir Frieder, 2004). Algoritma *K-Means* salah satu metode *clustering* untuk menyelesaikan permasalahan pada penjelasan diatas. Algoritma *K-Means* adalah algoritma *clustering* yang paling sederhana dibanding dengan algoritma yang lain. Kelebihan algoritma *K-Means* ini mudah untuk diterapkan dan dijalankan, relatif cepat, mudah diadaptasi dan dipraktekkan.

Penelitian terdahulu dilakukan oleh Ravena pada tahun 2019 yang berjudul “Zonasi Daerah Terdampak Bencana Angin Puting Beliung Menggunakan K-Means *Clustering* dengan Analisis *Silhouette*, *Davies Bouldin* Dan *Purity*”. Penelitian ini dilakukan agar pemerintah mengetahui daerah-daerah mana saja yang perlu diwaspadai akan bencana alam. Hasil penelitian menunjukkan bahwa daerah rawan bencana alam puting beliung terbagi menjadi 4 kelompok dengan karakteristik aman, cukup rawan, rawan dan sangat rawan. Dengan nilai *Silhouette Coefficient* sebesar 0.9116, *Davies Bouldin Index* sebesar 0.3633 dan *Purity* sebesar 1.

Hafiz Yusuf H,dkk melakukan penelitian pada tahun 2019 yang berjudul “Analisis *Cluster* Intensitas Kebencanaan di Indonesia Menggunakan Metode K-Means”. Penelitian ini dilakukan agar dimanfaatkan oleh pemerintah pusat maupun daerah agar lebih fokus dalam mempersiapkan langkah-langkah yang harus diambil untuk menanggulangi dampak bencana alam pada suatu daerah sesuai dengan jenis bencana yang sering terjadi pada daerah tersebut. Hasil dari penelitian menunjukkan bahwa daerah rawan bencana banjir terbagi 4 kelompok, tanah longsor terbagi 4 kelompok dan bencana puting beliung terbagi menjadi 4 kelompok.

Penelitian oleh Choirun N. dan Wiyli Y. juga dilakukan pada tahun 2021 tentang “Studi Perbandingan Algoritma Klastering Dalam Pengelompokan Persediaan Produk”. Penelitian ini dilakukan studi perbandingan ketiga algoritma *Clustering* yaitu *K-Means Clustering*, *K-Medoids Clustering* dan *Agglomerative Hierarchical Clustering* yang bertujuan untuk membandingkan ketiga metode *Clustering* dalam mengelompokkan data persediaan produk dengan optimal. Hasil penelitian menunjukkan bahwa data persediaan produk membentuk dua *cluster* dengan nilai hasil uji validitas *Silhouette Index* sebesar 0.52.

Berdasarkan pemaparan dari penelitian-penelitian diatas, algoritma K-Means dapat diterapkan untuk mengelompokkan provinsi-provinsi mana saja terdampak bencana alam yang ada di Indonesia berdasarkan data bencana alam yang paling sering terjadi, data rata-rata laju pertumbuhan penduduk dan data curah hujan. Selain itu penelitian ini diharapkan dapat bermanfaat untuk pemerintah pusat maupun pemerintah daerah agar lebih fokus dalam merancang langkah-langkah yang harus diambil untuk mencegah atau menanggulangi dampak suatu bencana yang akan terjadi dan juga diharapkan mampu dimanfaatkan oleh masyarakat Indonesia agar lebih mempersiapkan diri dan mempelajari cara penanggulangan bencana alam sesuai dengan yang sering terjadi ditempat tinggalnya. Sehingga peneliti akan melakukan penelitian dengan judul “*CLUSTERING BENCANA ALAM DI INDONESIA MENGGUNAKAN ALGORITMA K-MEANS*”.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang penulis paparkan diatas, maka dilakukan penelitian dengan rumusan pokok permasalahan penulisan skripsi sebagai berikut:

1. Bagaimana hasil evaluasi *clustering* bencana alam di Indonesia menggunakan Algoritma *K-Means*?
2. Bagaimana hasil pengelompokan terbaik terhadap *clustering* bencana alam di Indonesia menggunakan Algoritma *K-Means* ?

1.3 Batasan Masalah

Dalam sebuah penelitian sangat diperlukan batasan masalah untuk menghindari pembahasan yang meluas. Adapun batasan masalah dalam penulisan adalah sebagai berikut:

1. Penelitian ini menggunakan data sekunder tentang bencana alam di Indonesia dengan variabel yaitu tanah longsor, banjir, banjir bandang, gempa bumi, tsunami, kekeringan, kebakaran, gunung meletus, puting beliung dan gelombang pasang laut, data rata-rata laju pertumbuhan penduduk, dan data curah hujan.
2. Data penelitian yang digunakan, diperoleh dari sumber BPS pada tahun 2021.
3. Pengelompokan yang diteliti *cluster 2, cluster 3, cluster 4, cluster 5, cluster 6, dan cluster 7.*

1.4 Tujuan Penelitian

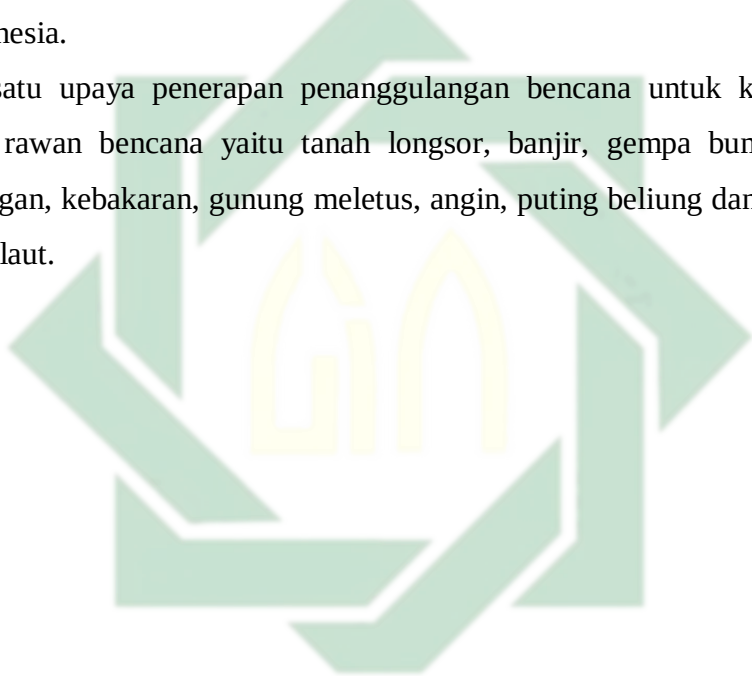
Berdasarkan penjelasan latar belakang dan rumusan masalah diatas, maka terdapat beberapa tujuan sebagai berikut:

1. Mengetahui hasil evaluasi *clustering* bencana alam di Indonesia menggunakan Algoritma *K-Means*.
2. Mengetahui hasil pengelompokan terbaik terhadap *clustering* bencana alam di Indonesia menggunakan Algoritma *K-Means*.

1.5 Manfaat Penelitian

Adapun manfaat yang diharapkan dari hasil penelitian ini diantara lain :

1. Pemerintah pusat atau daerah dapat mengetahui daerah rawan bencana yaitu tanah longsor, banjir, gempa bumi, tsunami, kekeringan, kebakaran, gunung meletus, angin, puting beliung dan gelombang pasang laut yang sering terjadi di Indonesia.
2. Salah satu upaya penerapan penanggulangan bencana untuk kewaspadaan daerah rawan bencana yaitu tanah longsor, banjir, gempa bumi, tsunami, kekeringan, kebakaran, gunung meletus, angin, puting beliung dan gelombang pasang laut.



UIN SUNAN AMPEL
S U R A B A Y A

BAB II

KAJIAN PUSTAKA

2.1 Bencana Alam

Berdasarkan Undang-Undang Nomor 24 Tahun 2007 Tentang Penanggulangan Bencana menjelaskan bencana adalah rangkaian peristiwa yang mengancam dan mengganggu kehidupan dan penghidupan masyarakat yang disebabkan oleh faktor alam, faktor non-alam maupun faktor manusia. Sehingga mengakibatkan adanya korban jiwa, kerusakan lingkungan, kerugian harta benda, dan dampak psikologis. Selain itu, BPS juga menyebutkan bahwa bencana alam adalah serangkaian peristiwa yang kejadiannya tidak terduga, yang mengancam masyarakat disebabkan oleh faktor alam yang menimbulkan banyaknya kerusakan lingkungan sekitarnya dan rasa khawatir untuk sebagian besar penduduk (BPS, 2022).

Bencana alam ada banyak jenisnya, namun secara umum terbagi menjadi 3 jenis yaitu sebagai berikut (BPBD, DISKOMINFO KABUPATEN BOGOR, 2021):

1. Bencana Alam Geologi yaitu bencana alam yang terjadi di permukaan bumi seperti tsunami, gempa bumi, gunung meletus, dan tanah longsor.
2. Bencana Alam Meteorologi/hidrometeorologi yaitu bencana alam yang berhubungan dengan iklim bencana alam ini umumnya tidak terjadi pada suatu tempat yang khusus. Bencana alam yang bersifat meteorologis ini paling banyak terjadi diseluruh dunia seperti banjir dan kekeringan.
3. Bencana Alam Ekstrateretial yaitu bencana alam yang terjadi di luar angkasa. Seperti datangnya berbagai benda langit misalnya asteroid, atau gangguan badai matahari.

2.1.1 Tsunami

Dalam Kamus Besar Bahasa Indonesia (KBBI), tsunami diartikan sebagai gelombang laut dahsyat yang terjadi karena gempa bumi atau letusan gunung api di dasar laut. Dilansir dalam buku Ensiklopedia Bencana 1: Tsunami (2016) karya Rani Siti Fitriani, dkk menjelaskan tsunami adalah gelombang air yang sangat besar dan terbentuk secara tiba-tiba. Gelombang yang umumnya disebabkan oleh gempa bumi, namun biasa juga disebabkan oleh gunung meletus dan longsor. Berikut beberapa penyebab terjadinya tsunami, yaitu: longsor laut yang memiliki bentuk muka bumi berupa tebing maupun bukit, seperti halnya daratan. Akibat gempa bumi dan dorongan arus laut yang terus-menerus menyebabkan tebing dan bukit menjadi labil sehingga terjadi longsor. Bagian dari tebing dan bukit yang longsor tersebut akan memicu terjadinya tsunami.

2.1.2 Gempa Bumi

Menurut BMKG, gempa bumi adalah peristiwa bergetarnya bumi akibat pelepasan energi di dalam bumi secara tiba-tiba yang ditandai dengan retaknya lapisan batuan pada kerak bumi. Akumulasi energi penyebab terjadinya gempa bumi yang dihasilkan dari pergerakan lempeng-lempeng tektonik. Energi yang dihasilkan tersebut di pancarkan ke segala arah berupa gelombang gempa bumi. Sehingga efeknya dapat dirasakan sampai ke permukaan bumi. Memahami tentang gempa bumi tentunya sangat penting agar kamu bias bertindak dengan tepat saat terjadinya bencana alam ini.

2.1.3 Tanah Longsor

Tanah longsor sering terjadi dipicu akibat dari curah hujan yang tinggi, lereng terjal, tanah yang kurang padat serta tebal, terjadinya

pengikisan, berkurangnya tutupan vegetasi, dan adanya getaran. Bencana longsor biasanya terjadi begitu cepat sehingga mengakibatkan terbatasnya waktu untuk melakukan evakuasi mandiri. Material longsor dapat menimbun apa saja yang berada di jalur yang terkena longsor tersebut. Ada dua faktor utama penyebab tanah longsor yaitu faktor pendorong yang mempengaruhi kondisi material sendiri dan faktor pemicu yang menyebabkan bergerak material tersebut. curah hujan yang tinggi, erosi tanah, lereng yang terjal, aktivitas pertambangan, dan beban yang berlebihan di suatu area (Abdi, 2021).

2.1.4 Angin Puting Beliung

Bencana puting beliung terjadi akibat dari peristiwa hidrometeorologis meningkat intensitas kejadiannya pada masa peralihan musim. Bencana ini menjadi bagian dari proses pertumbuhan awan hujan cumulus nimbus yang terbentuk akibat pemanasan intensif. Ancaman puting beliung sulit diprediksi karena merupakan fenomena atmosfer skala lokal. Bencana puting beliung berakibat seperti kerusakan rumah-rumah dan banyaknya pohon tumbang (BNPB, 2019).

2.1.5 Banjir

Banjir adalah peristiwa atau keadaan dimana terendamnya suatu daerah atau daratan karena volume air yang meningkat yang biasanya tidak digenangi air dalam jangka waktu tertentu. Banjir biasa terjadi karena curah hujan turun terus menerus dan mengakibatkan meluapnya air sungai, danau, laut atau drainase dengan jumlah air yang melebihi daya tampung media penampung air dari curah hujan tersebut. Selain faktor alam penyebab terjadinya banjir yaitu curah hujan yang tinggi, banjir juga terjadi akibat ulah manusia. Seperti berkurangnya kawasan resapan air karena alih fungsi lahan, penggundulan hutan yang meningkatkan erosi dan mendangkalan sungai, serta perilaku tidak bertanggung jawab yaitu

membuang sampah di sungai dan mendirikan rumah di bantaran sungai (BNPB, 2019).

2.1.6 Banjir Bandang

Banjir bandang merupakan banjir bervolume besar yang menggenangi rumah-rumah, alat transportasi, lahan dan lainnya, datang tiba-tiba disebabkan oleh intensitas hujan yang terjadi terus-menerus di daerah dataran rendah dengan debit air yang besar karena ter bendungnya aliran sungai pada alur sungai tersebut (Andryanto, 2021). Menurut Korps Insinyur Angkatan Darat AS, *National Weather Service* (NWS) dan Badan Manajemen Darurat Federal mengatakan bahwa banjir bandang adalah banjir dengan jangka waktu sebentar dengan aliran kecepatan yang tinggi disebabkan oleh hujan lebat dalam waktu singkat antara tiga sampai enam jam. NWS juga menyebutkan banjir bandang dapat terjadi bila dilihat dari kadar air tanah, ketinggian, dan limpasan hujan di wilayah dan waktu tertentu. Contohnya, jika kadar air tanah tinggi, maka hujan yang hanya satu inci saja dalam waktu singkat dapat mengakibatkan banjir bandang.

2.1.7 Kekeringan

Kekeringan merupakan kondisi kekurangan pasokan air dari curah hujan dalam jangka waktu tertentu biasanya satu musim atau lebih yang mengakibatkan beberapa sektor kegiatan, kelompok, atau lingkungan kekurangan air (UNISDR, 2019). Kekeringan muncul bila suatu wilayah mengalami curah hujan dibawah rata-rata secara terus-menerus. Kemarau yang berkepanjangan juga dapat menyebabkan satu wilayah kering, karena cadangan air tanah yang habis akibat penguapan, transpirasi dan digunakan untuk keperluan manusia.

Kondisi kekeringan yang parah dapat dikategorikan sebagai bencana alam apabila wilayah yang mengalami kekurangan air telah kehilangan

sumber pendapatan, akibat gagal panen atau ekosistem lingkungan yang rusak. Kekeringan dapat memberikan dampak menyeluruh terhadap komponen kehidupan. Tidak hanya tumbuhan yang akan mati, tetapi juga mengancam kematian manusia akibat kekurangan cairan dan kekurangan makanan karena gagal panen.

2.1.8 Kebakaran Hutan

Kebakaran secara umum adalah peristiwa nyala api, baik kecil atau besar pada tempat yang tidak kita kehendaki yang umumnya sukar untuk dikendalikan dan merugikan lingkungan sekitarnya. Pemicu kebakaran biasanya berasal dari sumber korsleting listrik, kebocoran tabung/kompor gas, puntung rokok yang dibuang sembarangan, cuaca panas, dan lainnya. Kebakaran dikelompokkan berdasarkan pada jenis benda atau bahan yang terbakar. Pengelompokan tersebut diharapkan akan lebih mudah, cepat dan tepat dalam pemilihan media pemadaman yang akan digunakan untuk proses pemadaman seperti air, busa, serbuk kimia kering dan lainnya. Teknik pemadaman kebakaran dapat dilakukan dengan cara menghilangkan salah satu atau beberapa unsur yang menyebabkan nyala api dengan cara yaitu: pendinginan (*cooling*) dan penyalutan (*smothering*) (BPBD, 2018).

Kebakaran hutan atau lahan adalah keadaan pada lahan dan hutan yang dilanda api sehingga mengakibatkan kerusakan serta dampak yang sangat merugikan. Beberapa penyebab terjadinya kebakaran hutan adalah sebagai berikut:

1. Terbatasnya keterangan dan pengetahuan tentang kebakaran hutan
2. Kesengajaan dalam pembakaran hutan untuk membuka lahan, membakar sampah sembarangan
3. Kebakaran hutan akibat gesekan antar batang, sambaran petir, atau letusan gunung api.

2.1.9 Gelombang Pasang Laut

Gelombang adalah pergerakan naik dan turunnya air dengan arah tegak lurus permukaan air laut yang membentuk grafik sinusoidal. Gelombang laut terjadi disebabkan oleh angin di atas lautan yang mentransfer energinya ke perairan, menyebabkan riak-riak, alun/bukit. Gelombang dipengaruhi oleh banyak faktor:

- a. Angin (kecepatan angin, panjang/jarak hembusan angin, lamanya hembusan angin)
- b. Geometri laut (topografi atau profil laut dan bentuk pantai)
- c. Gempa (apabila terjadi tsunami) (Wimala L. Dhanista, 2017).

Gelombang pasang merupakan gelombang besar yang ketinggian gelombang melebihi pada keadaan batas normal. Gelombang pasang ini dapat menimbulkan bahaya baik di lautan maupun di daratan terutama daerah pesisir pantai. Gelombang pasang biasanya disebabkan oleh perubahan musim yang mendatangkan cuaca ekstrim dan disertai angin kencang (Masdin, 2021).

2.1.10 Gunung Meletus

Bahaya erupsi gunung meletus memiliki dua jenis bahaya berdasarkan waktu kejadian, yaitu bahaya primer dan bahaya sekunder. Berikut beberapa bahaya dari erupsinya gunung meletus.

1. Awan panas yaitu aliran material vulkanik panas yang terdiri atas batuan berat, ringan (berongga) dengan lava massif dan butiran klastik yang pergerakannya dipengaruhi gravitasi dan cenderung mengalir melalui lembah.
2. Aliran lava yaitu magma yang meleleh ke permukaan bumi melalui rekahan dengan suhu $> 10.000^{\circ}\text{C}$ dan dapat merusak segala bentuk infrastruktur.

3. Gas beracun yaitu gas vulkanik yang dapat mematikan seketika apabila terhirup dalam tubuh.
4. Lontaran material (pijar) terjadi ketika letusan magmatic berlangsung dengan suhu mencapai 200⁰C dan diameter lebih dari 10 cm serta daya lontar ratusan kilometer.
5. Hujan abu yang material abu tersebut tampak halus dan bergerak sesuai arah angin.
6. Lahar letusan terjadi pada gunung berapi yang mempunyai danau kawah, terjadi bersamaan dengan letusan. Air danau yang bercampur material lepas gunung berapi mengalir dan membentuk banjir lahar (BNPB, 2019).

2.2 Laju Pertumbuhan Penduduk

Laju Pertumbuhan Penduduk (LPP) adalah angka rata-rata yang menunjukkan tingkat pertambahan penduduk pertahun dalam jangka waktu tertentu. Angka ini dinyatakan sebagai presentase dari penduduk dasar. Laju Pertumbuhan Penduduk dapat dihitung menggunakan tiga metode, yaitu: aritmatik, geometrik, dan eksponensial. Metode yang sering digunakan di BPS adalah metode geometrik dengan persamaan 2.1 berikut.

Rumus :
$$r = \left(\frac{P_t}{P_0} \right)^{\frac{1}{t}} - 1 \quad 2.1$$

Keterangan :

r : Laju Pertumbuhan Penduduk

P_t : jumlah penduduk tahun t

P_0 : jumlah penduduk tahun awal

t : periode waktu antara tahun dasar dan tahun t (dalam tahun)

Tujuan perhitungan Laju Pertumbuhan Penduduk (LPP) ini, untuk mengetahui perubahan jumlah penduduk antar dua periode waktu.

- a. Jika $LPP > 0$ maka terjadi penambahan penduduk. Pada tahun t dibandingkan dengan tahun sebelumnya.
- b. jika $LPP = 0$ maka tidak terjadi perubahan jumlah penduduk pada tahun t dibandingkan dengan tahun sebelumnya.
- c. Sedangkan jika $P_{t,0} < 100$ maka terjadi pengurangan jumlah penduduk pada tahun t dibandingkan tahun sebelumnya.

2.3 Curah Hujan

Indonesia merupakan negara yang berada di garis khatulistiwa yang memiliki musim kemarau dan musim penghujan. Dua musim ini sangat berpengaruh terhadap kelangsungan hidup makhluk hidup di Indonesia. Sehingga masyarakat Indonesia harus mengenali faktor dari kedua musim tersebut agar dapat melakukan aktivitasnya dengan lancar. Salah satu faktor penting dari musim-musim tersebut adalah curah hujan.

Curah hujan adalah jumlah air hujan yang jatuh selama periode waktu tertentu yang pengukurannya menggunakan satuan tinggi di atas permukaan tanah horizontal yang diasumsikan tidak terjadi infiltrasi, run off, maupun evaporasi.

1. Infiltrasi adalah proses masuknya air permukaan di bumi kedalam lapisan batuan. Air hujan yang jatuh ke bumi akan diserap ke dalam tanah yang mengalir.
2. *Run off* adalah proses keluarnya air dari dalam tanah maupun air permukaan seperti sungai dan danau menuju kembali ke laut sebagai tempat akhir aliran air.
3. Evaporasi adalah proses penguapan air dari permukaan bumi ke atmosfer akibat paparan sinar matahari. Penguapan ini dapat terjadi pada air di darat dan di laut.

Curah hujan juga sering disebut dengan presipitasi dapat diartikan jumlah air hujan yang turun di daerah tertentu dalam satuan waktu tertentu. Jumlah air yang jatuh ke permukaan tanah datar selama periode tertentu diukur dengan satuan tinggi millimeter (mm) diatas permukaan horizontal (Anwar, 2017).

2.4 Normalisasi Data

Normalisasi data merupakan salah satu teknik yang penting untuk dipahami dalam pra-proses data. Dalam analisis dan eksplorasi data sering kali kita menemukan banyak *features* data atau variabel di dalam dataset yang akan di analisis. Normalisasi data adalah proses membuat beberapa variabel memiliki rentang nilai yang sama, tidak ada yang terlalu besar maupun terlalu kecil sehingga dapat membuat analisis statistic menjadi lebih mudah. Salah satu metode yang dapat untuk normalisasi data yaitu *Min-Max*.

Min-Max adalah suatu metode yang setiap nilai pada sebuah fitur dikurangi dengan nilai minimum fitur tersebut, kemudian dibagi dengan rentang nilai atau nilai maksimum dikurangi nilai minimum fitur tersebut untuk menghasilkan nilai baru hasil dari normalisasi data antara 0 sampai 1. Seperti pada persamaan 2.2 berikut:

$$X_{new} = \frac{X_{old} - X_{min}}{X_{max} - X_{min}} \quad 2.2$$

Dimana:

X_{new} = data atribut yang akan di normalisasi

X_{old} = nilai data

X_{min} = nilai terkecil atribut tersebut

X_{max} = nilai tertinggi atribut tersebut

Metode normalisasi dengan *min-max normalization* akan melakukan transformasi linier terhadap data asli sehingga menghasilkan keseimbangan nilai

perbandingan antar data saat sebelum dan sesudah proses (Nasution, Khotimah, & Chamidah, 2019).

2.5 Uji Multikolinearitas

Menurut Ghazali (2016) pengujian multikolinearitas bertujuan untuk mengetahui apakah model regresi ditemukan adanya korelasi antar variabel *independent* atau variabel bebas. Uji multikolinearitas dapat menyebabkan tingginya variabel pada suatu sampel. Hal tersebut sebagai standar error besar, yang artinya ketika koefisien di uji, t-hitung akan bernilai kecil dari t-tabel. Hal ini menunjukkan tidak adanya hubungan linear antara variabel bebas yang dipengaruhi dengan variabel terikat. Beberapa hal yang perlu diketahui untuk menemukan ada atau tidaknya multikolinearitas pada model regresi, yaitu sebagai berikut:

1. Terjadi perubahan pada koefisien model regresi (nilainya menjadi lebih besar atau kecil) apabila dilakukan penambahan atau pengeluan sebuah variabel bebas dari model regresi.
2. Koefisien determinasi (R^2) yang tinggi, dan nilai signifikansi t rendah pada setiap variabel X (variabel bebas).

$$R^2 = 1 - \frac{ESS}{TSS} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad 0 \leq R^2 \leq 1 \quad 2.3$$

Dimana:

ESS : Error Sum Square (Jumlah Kuadrat Galat)

TSS : Total Sum Square (Jumlah Kuadrat Regresi)

y_i : variabel dependen ke-i

$\hat{y}_i$: variabel pengganggu (*error term*) dari model diestimasi

$\bar{y}$: variabel dependen rata-rata

3. Koefisien korelasi diantara variabel X tinggi (diatas 0.7).
4. Nilai *variance inflation factor* (VIF)

$$VIF = (1/1 - R^2) \quad 2.4$$

Jika nilai VIF >10, maka terdeteksi terjadi multikolinearitas. Sedangkan jika nilai VIF < 10 maka model terbebas dari multikolinearitas yang dapat disimpulkan bahwa variabel terbebas dari asumsi klasik multikolinearitas karena hasilnya lebih kecil dari angka 10.

5. Nilai standar error untuk koefisien regresi menjadi lebih besar dari yang sebenarnya.

Berikut cara mengatasi masalah dalam uji multikolinearitas yaitu:

- a. Membuang salah satu variabel yang berkolinearitas. Untuk mengetahui variabel mana yang akan dibuang, dilakukan coba-coba dan pilih persamaan yang memiliki nilai adjusted R^2 yang lebih tinggi.
- b. Menambah jumlah sampel, karena mungkin saja multikolinearitas terjadi hanya untuk sampel yang saat ini saja.

2.6 Clustering

Data mining adalah suatu proses pengumpulan informasi penting dari suatu data yang besar. Data mining memiliki dua fungsi utama yaitu deskriptif yang berfungsi untuk memahami lebih jauh tentang data yang diamati dengan harapan bias mengetahui perilaku dari sebuah data tersebut. data tersebut yang nantinya dapat digunakan untuk mengetahui karakteristik dari data yang dimaksud. Fungsi prediktif berfungsi untuk bagaimana sebuah proses yang nantinya akan menemukan pola tertentu dari suatu data. Pola-pola tersebut dapat diketahui dari berbagai variabel-variabel yang ada pada data. Ketika pola tersebut didapatkan, maka dapat digunakan untuk memprediksi variabel lain yang belum diketahui nilai atau jenisnya.

Data mining merupakan metode penelitian untuk menentukan hasil tertentu berdasarkan total data yang dikumpulkan dirancang untuk mengekstrak aturan dari data dalam jumlah besar. Sementara *Machine Learning* mengajarkan cara belajar dan memahami parameter yang diberikan untuk melatih system dalam melakukan tugas yang kompleks dan menggunakan data dan pengalaman yang dikumpulkan untuk menjadi lebih cerdas. *Machine Learning* (ML) merupakan salah satu cabang dari *artificial intelligence* (AI) yang dapat membantu manusia dalam berbagai bidang. *Machine Learning* dapat menganalisa data berdasarkan data yang diberikan saat awal pengembangan dan saat *Machine Learning* sudah digunakan.

Unsupervised Learning merupakan salah satu tipe algoritma dari *Machine Learning* yang digunakan untuk menarik kesimpulan dari dataset. metode ini hanya akan mempelajari suatu data berdasarkan kedekatannya saja atau yang biasa disebut dengan *clustering*. Metode *unsupervised Learning* yang paling umum digunakan untuk analisis *cluster* yaitu untuk mencari pola-pola tersembunyi atau pengelompokan dalam data (Salsabila, 2021).

Clustering merupakan salah satu cabang ilmu data mining. *Clustering* digunakan untuk mengidentifikasi dan mengelompokkan data sesuai dengan ciri maupun karakteristiknya. Berbeda dengan klasifikasi, teknik *clustering* yaitu teknik mengelompokkan data secara otomatis sebelum mengetahui label kelasnya. Tetapi *clustering* juga dapat digunakan untuk menghimpun data dengan label kelas yang diketahui.

Analisis *cluster* mengelompokkan objek sehingga setiap objek yang paling dekat kesamaanya dengan objek lain berada dalam *cluster* yang sama. *Cluster-cluster* yang terbentuk memiliki homogenitas (kesamaan) internal antar anggota *cluster* yang tinggi dan heterogenitas (perbedaan) eksternal antar *cluster* yang tinggi. Pada dasarnya solusi *cluster* bergantung pada variabel-variabel yang digunakan sebagai dasar untuk menilai kesamaan. Beberapa solusi yang berbeda dapat diperoleh dengan mengubah satu elemen atau lebih dan anggota *cluster*

untuk setiap penyelesaian/solusi tergantung pada beberapa elemen prosedur (Sari N. N., 2010).

2.7 Algoritma K- Means

K-Means merupakan suatu algoritma yang digunakan dalam pengelompokan secara partisi yang memisahkan data ke dalam kelompok yang berbeda-beda. Algoritma ini mampu meminimalkan jarak antara data ke klasternya. Pada dasarnya penggunaan algoritma ini, dalam proses *clustering* tergantung pada data yang didapatkan dan konklusi yang ingin dicapai di akhir proses (Nengsih, 2017). Tahapan inti dari algoritma *K-Means* adalah klasterisasi pengelompokan data berdasarkan titik pusat klaster (*centroid*) untuk memaksimalkan kemiripan data yang didapatkan berdasarkan jarak terpendek data terhadap titik *centroid* awal (Sari & Hartama, 2018).

Berikut adalah langkah–langkah yang dilakukan dalam mengelompokkan data dengan menggunakan algoritma K-Means (Everit, S, Landau, Leese, & Stahl, 2011).

1. Menentukan banyaknya *k*, dimana *k* adalah jumlah *cluster* yang akan dibentuk.
2. Menentukan titik pusat atau biasa disebut *centroid* dari setiap *cluster*. *Centroid* ini diambil secara acak.
3. Alokasikan semua data ke *centroid* terdekat dengan menghitung jarak setiap objek atau data ke titik pusat (*centroid*). Jarak *centroid* yang digunakan yaitu *Euclidean* dengan persamaan 2.5 berikut.

$$D(x, y) = \sqrt{\sum_{i=0}^n (x_i - y_i)^2} \quad 2.5$$

Dimana :

$D(x, y)$ = jarak data *x* ke pusat klaster *y*

x_i = data x pada observasi ke- i
 y_i = titik pusat ke y observasi ke- i
n = banyaknya observasi

4. Mengelompokkan setiap data berdasarkan kedekatannya (jarak terkecil) dengan *centroid* menggunakan persamaan 2.6 berikut.

$$\text{Min } \sum_{k=1}^k D(x, y) = \sqrt{\sum_{i=0}^n (x_i - y_i)^2} \quad 2.6$$

5. Mencari nilai titik pusat atau *centroid* terbaru. Nilai tersebut diperoleh dari rata *cluster* setiap *cluster* dengan menggunakan persamaan 2.7 berikut.

$$C_{ij} = \frac{\sum_{i=1}^p x_{ij}}{p} \quad 2.7$$

Dimana :

C_{ij} = centroid terbaru pada iterasi k

x_{ij} = anggota *cluster* k

p = banyaknya anggota *cluster* ke-k

6. Melakukan perulangan langkah 2-5 hingga tidak ada lagi anggota *cluster* yang berpindah ke *cluster* lain.

2.8 Silhoutte Index (SI)

Silhoutte Index merupakan salah satu ukuran validasi yang berbasis kriteria internal. *Silhoutte Index* akan mengevaluasi penempatan setiap objek dalam setiap *cluster* dengan membandingkan jarak rata-rata objek dalam satu *cluster* dan jarak antara bojek dngan *cluster* yang berbeda (F.N., S., & R., 2014). Menghitung koefisien *Silhoutte* yang didefinisikan sebagai rata – rata s(i) pada persamaan 2.8 berikut.

$$SC = \frac{1}{n} \sum_{i=1}^n s(i) \quad 2.8$$

Dengan,

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}, b(i) = \min d(i, C), \text{ dan } a(i) = \frac{1}{|A|-1} \sum_{j \in A, j \neq i} d(i, j)$$

b(i) : nilai minimum dari jarak rata-rata objek I dengan semua objek pada *cluster* lain C

a(i) : rata-rata jarak objek ke-I dengan semua objek yang berada didalam satu *cluster* A

Hasil perhitungan nilai koefisien *silhouette* berada pada range -1 sampai 1. Semakin besar nilai koefisien *silhouette* maka semakin baik kualitas suatu kelompok. Berikut tabel interval dari *silhouette*.

Tabel 2.1 Interval *Silhouette*

Interval <i>Silhouette</i>	Interpretasi
0.71 -1.0	Sangat baik
0.51 -0.70	Baik
0.26 - 0.50	Cukup
< 0.25	Jelek

BAB III

METODE PENELITIAN

3.1 Sumber Data

Hasil pencatatan BPS Indonesia tahun 2018 membuktikan bahwa sejak tiga tahun terakhir sudah lebih dari 20.000 kasus bencana alam terjadi diseluruh wilayah yang ada di Indonesia. Berdasarkan data tersebut, pemerintah harus lebih giat dalam mensosialisasikan upaya pengurangan resiko bencana kepada masyarakat. Selain itu perlu mengetahui daerah-daerah rawan bencana alam seperti yang diketahui, berikut ini merupakan data bencana alam yang sering terjadi per provinsi pada tahun 2021.

Tabel 3.1 Sampel Data Bencana Alam, Laju Pertumbuhan Penduduk (LPP),
dan Data Curah Hujan pada Tahun 2021

No.	Provinsi	Tanah Longsor	Banjir	Banjir Bandang		Kekeringan
0	ACEH	198	1435	81		173
1	SUMATERA UTARA	483	732	52		127
2	SUMATERA BARAT	222	342	65		43
3	RIAU	21	455	1		51
4	JAMBI	57	476	17		16
5	SUMATERA SELATAN	103	380	36		98
6	BENGKULU	81	171	15		19
7	LAMPUNG	70	328	23		30
8	KEP. BANGKA BELITUNG	1	59	0		0
9	KEP. RIAU	25	61	1		27
10	DKI JAKARTA	7	109	0		0
11	JAWA BARAT	1288	1193	100		442
12	JAWA TENGAH	1190	1249	69		278
13	DI YOGYAKARTA	59	50	1		31
14	JAWA TIMUR	511	1176	96		156
15	BANTEN	96	419	41		77

16	BALI	105	39	5		2
17	NUSA TENGGARA					
	BARAT	44	187	54		88
18	NUSA TENGGARA					
	TIMUR	610	570	151		454
19	KALIMANTAN					
	BARAT	66	864	22		75
20	KALIMANTAN					
	TENGAH	24	735	9		37
21	KALIMANTAN					
	SELATAN	74	929	31		35
22	KALIMANTAN					
	TIMUR	74	343	5		32
23	KALIMANTAN					
	UTARA	42	200	0		6
24	SULAWESI					
	UTARA	218	288	37		24
25	SULAWESI					
	TENGAH	133	567	57		42
26	SULAWESI					
	SELATAN	251	620	52		60
27	SULAWESI					
	TENGGARA	46	240	10		16
28	GORONTALO	65	259	24		57
29	SULAWESI					
	BARAT	144	125	7		22
30	MALUKU	58	153	2		26
31	MALUKU					
	UTARA	80	270	20		12
32	PAPUA BARAT	72	138	0		5
33	PAPUA	146	204	9		9

(sumber: <http://bps.go.id/>)

3.2 Struktur Data dan Variabel Penelitian

Data yang digunakan pada penelitian ini merupakan data sekunder bencana alam yang sering terjadi di Indonesia yaitu tsunami, tanah longsor, banjir, banjir bandang, gempa bumi, puting beliung, kebakaran, kekeringan, gunung meletus, dan gelombang pasang laut. Data yang diambil pada tanggal 26 Mei 2022 dari website resmi Badan Pusat Statistik (BPS) yaitu Data Informasi Bencana Alam

di Indonesia dengan alamat sebagai berikut: www.bps.go.id . berikut merupakan variabel-variabel yang akan digunakan pada penelitian ini:

Tabel 3.2 Definisi Operasional Bencana Alam, Curah Hujan,
dan Laju Pertumbuhan Penduduk

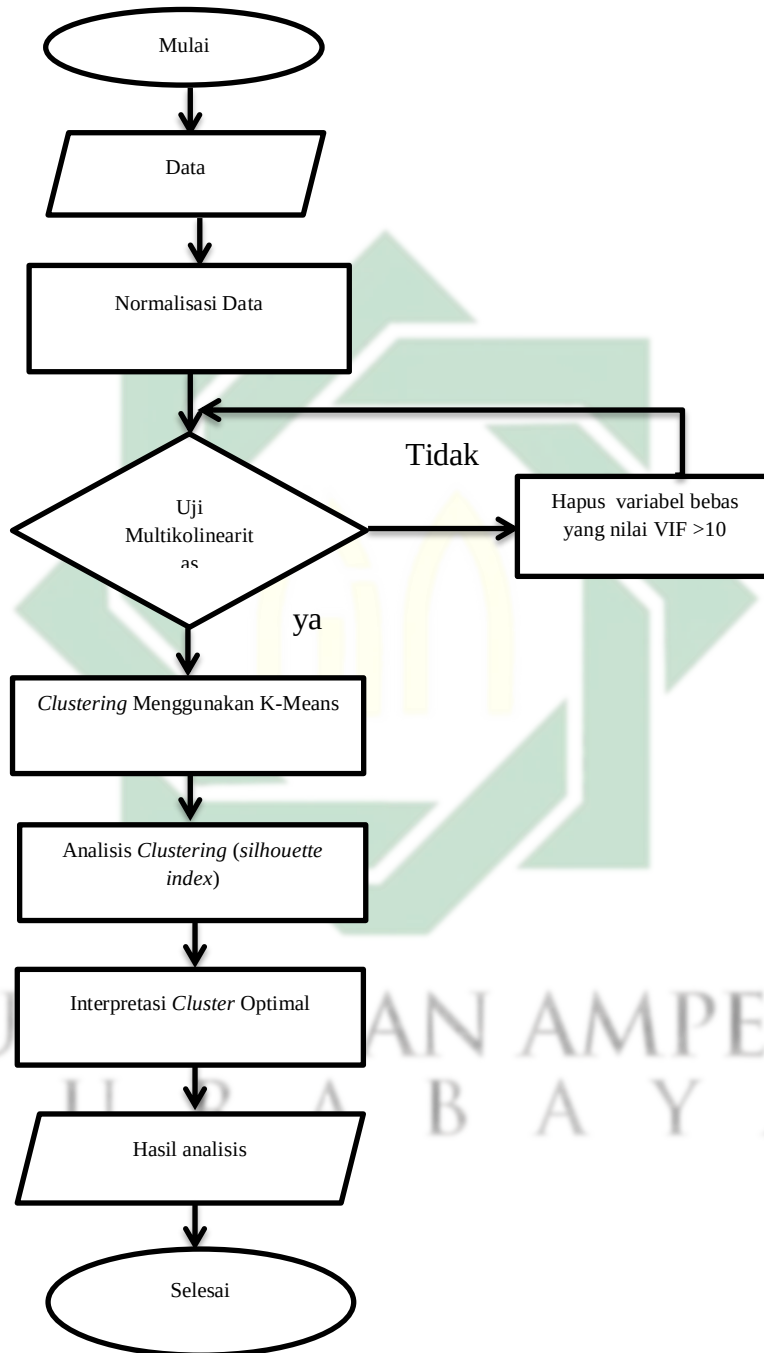
No.	Variabel	Definisi	Satuan
1	Banyaknya Desa/Kelurahan Menurut Jenis Bencana Alam	Data berupa jumlah desa yang terkena bencana tsunami, tanah longsor, banjir, banjir bandang, gempa bumi, puting beliung, kebakaran, kekeringan, gunung meletus, dan gelombang pasang laut di Indonesia	Jumlah desa/kelurahan
2	Curah Hujan	Data curah hujan yang tinggi	Jumlah data curah hujan tahun 2021
3	Laju Pertumbuhan Penduduk	Data rata-rata laju pertumbuhan penduduk	jumlah orang

3.3 Langkah Analisis

1. Pengambilan data
2. *Clustering* menggunakan K-Means dan Analisis *cluster* menggunakan *silhouette index*
3. Analisis daerah rawan bencana alam yang sering terjadi
4. Analisis hasil *clustering*

3.4 Perancangan system

Langkah analisis sebagaimana telah dijelaskan pada sub bab sebelumnya dapat digambarkan dengan diagram alir pada Gambar 3.1 berikut.

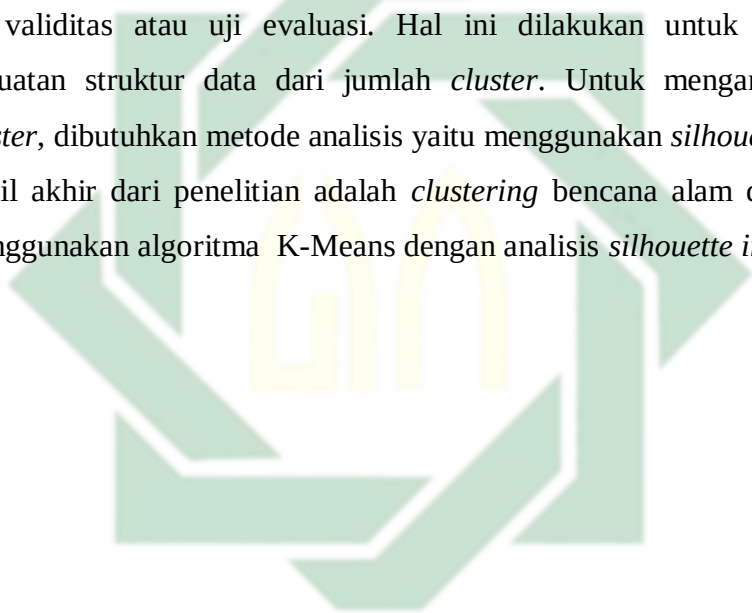


Gambar 3.1 Diagram Alir Penelitian

- a. Langkah awal penelitian ini yaitu pengambilan data. Data diambil dari website resmi BPS yaitu Banyaknya Desa/Kelurahan Menurut Jenis

Bencana Alam di Indonesia. Data yang diambil antara lain banyaknya desa/kelurahan menurut jenis bencana alam, curah hujan, dan rata-rata laju pertumbuhan penduduk.

- b. Setelah data diperoleh, maka dilakukan *clustering* menggunakan algoritma K-Means yang dapat diketahui kelompok-kelompok dari setiap provinsi manakah yang rawan akan bencana alam.
- c. Setelah terbentuk kelompok, pada penelitian *clustering* harus dilakukan uji validitas atau uji evaluasi. Hal ini dilakukan untuk mengetahui kekuatan struktur data dari jumlah *cluster*. Untuk menganalisis hasil *cluster*, dibutuhkan metode analisis yaitu menggunakan *silhouette index*.
- d. Hasil akhir dari penelitian adalah *clustering* bencana alam di Indonesia menggunakan algoritma K-Means dengan analisis *silhouette index*



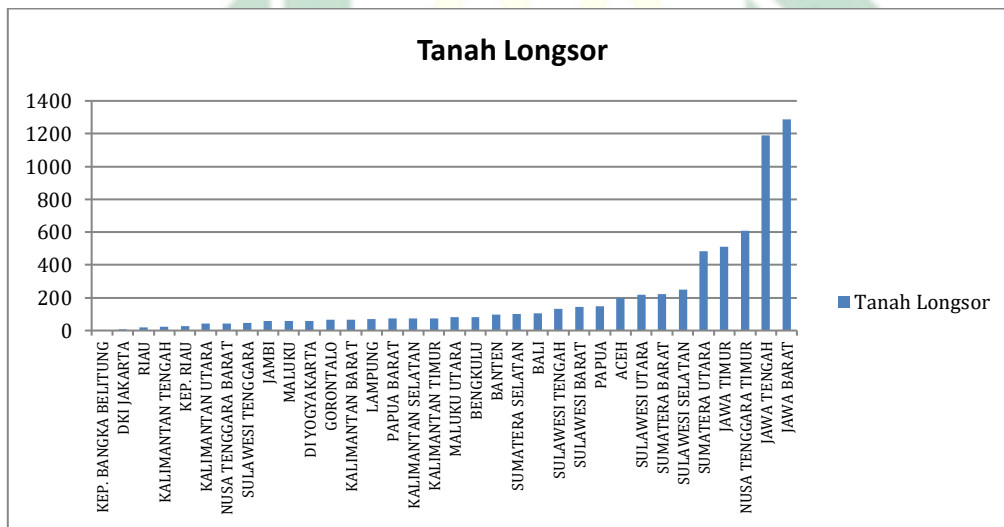
UIN SUNAN AMPEL
S U R A B A Y A

BAB IV

HASIL DAN PEMBAHASAN

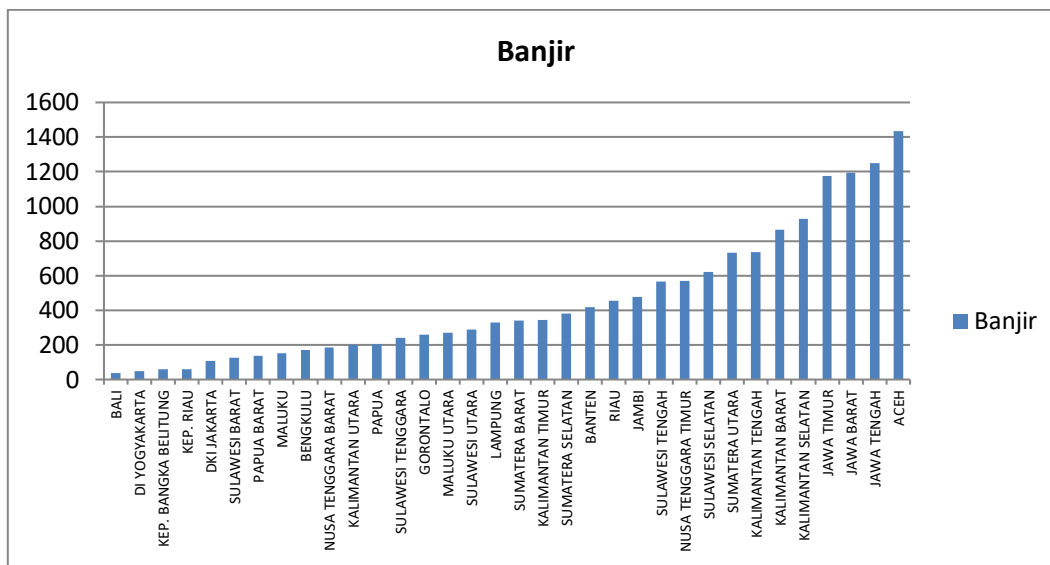
4.1 Deskripsi Data

Pada data bencana alam di 34 provinsi yang ada di Indonesia, dilakukan tahap *pre-processing (cleaning)*. Pada *pre-processing* ini dapat dilakukan dengan mengimport terlebih dahulu dataset bencana alam ke ORANGE, setelah import dataset tersebut dilakukan proses *cleaning* maka didapatkan hasil *cleaning* sebagai berikut:



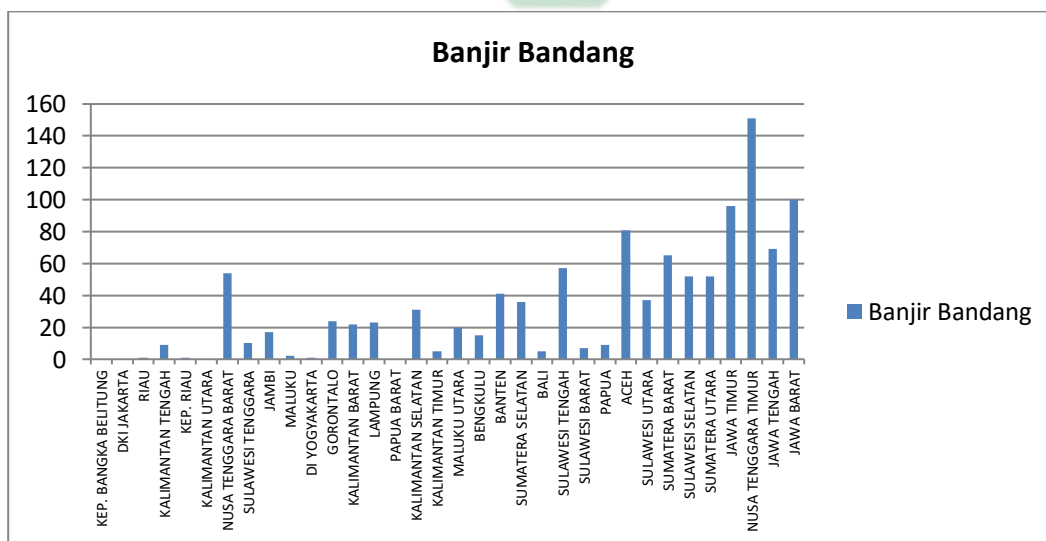
Gambar 4.1 Grafik Bencana Tanah Longsor

Jumlah desa yang sering terdampak bencana tanah longsor pada 2021 yaitu di provinsi Sumatera Utara sekitar 483 desa, Jawa Timur 511 desa, Nusa Tenggara Timur 610 desa, Jawa Tengah 1190 desa, dan Jawa Barat 1288 desa. Sedangkan daerah yang jarang terdampak bencana tanah longsor yaitu provinsi Kep. Bangka Belitung 1 desa, DKI Jakarta 7 desa, dan Riau 21 desa.



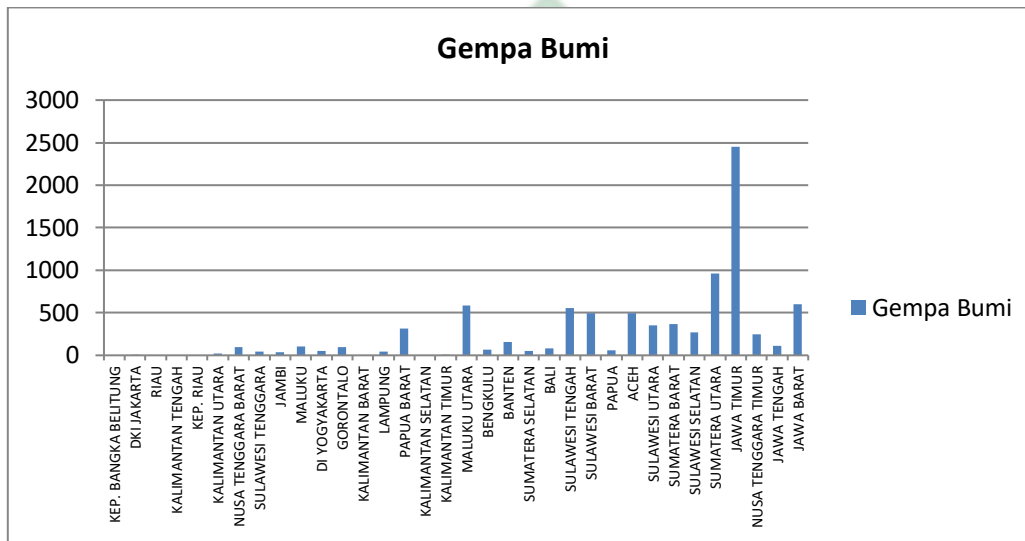
Gambar 4.2 Grafik Bencana Banjir

Jumlah desa yang sering terdampak bencana banjir pada 2021 yaitu di provinsi Kalimantan Selatan sekitar 929 desa, Jawa Timur 1176 desa, Jawa Barat 1193 desa, Jawa Tengah 1249 desa, dan Aceh 1435 desa. Sedangkan daerah yang jarang terdampak bencana tanah longsor yaitu provinsi Bali 39 desa, DIY 50 desa, dan Kep. Bangka Belitung 59 desa.



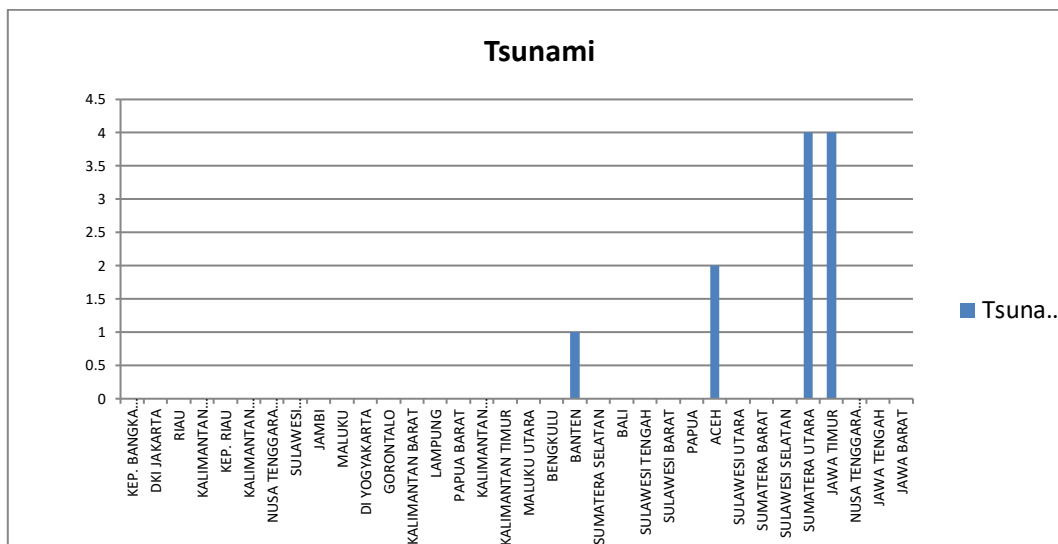
Gambar 4.3 Grafik Bencana Banjir Bandang

Jumlah desa yang sering terdampak bencana banjir bandang pada 2021 yaitu di provinsi Jawa Tengah sekitar 69 desa, Aceh 81 desa, Jawa Timur 96 desa, Jawa Barat 100 desa, dan Nusa Tenggara Timur 151 desa. Sedangkan daerah yang jarang terdampak bencana tanah longsor yaitu provinsi Riau, Kep. Riau, dan DIY yang terdampak bencana banjir bandang 1 desa.



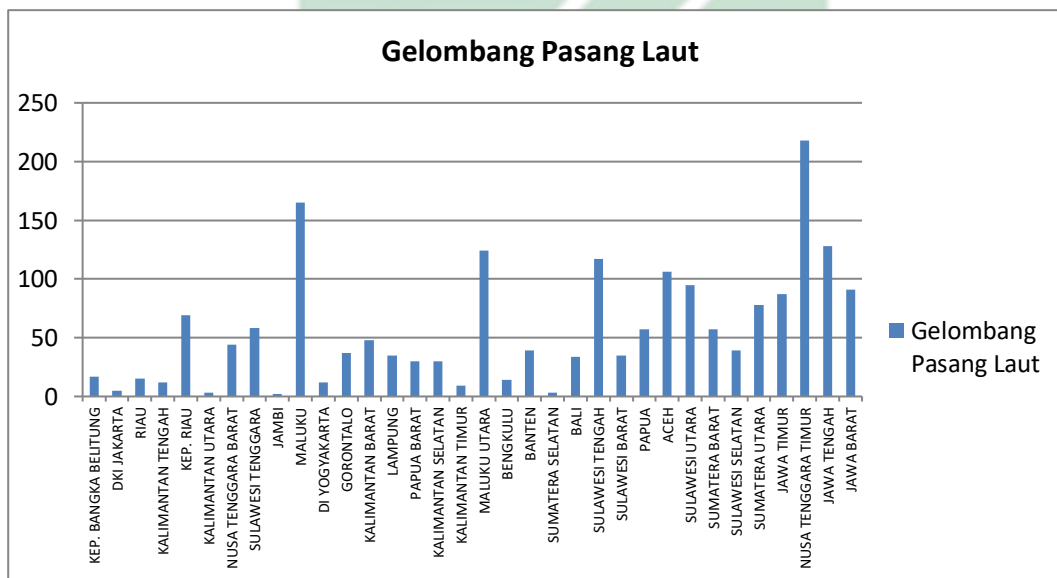
Gambar 4.4 Grafik Bencana Gempa Bumi

Jumlah desa yang sering terdampak bencana gempa bumi pada 2021 yaitu di provinsi Sulawesi Tengah sekitar 556 desa, Maluku Utara 589 desa, Jawa Barat 601 desa, Sumatera Utara 964 desa dan Jawa Timur 2449 desa. Sedangkan provinsi Riau, Kep. Bangka Belitung, Kep. Riau, dan Kalimantan (Barat, Tengah, dan Selatan) diperkirakan jarang terdampak bencana gempa bumi.



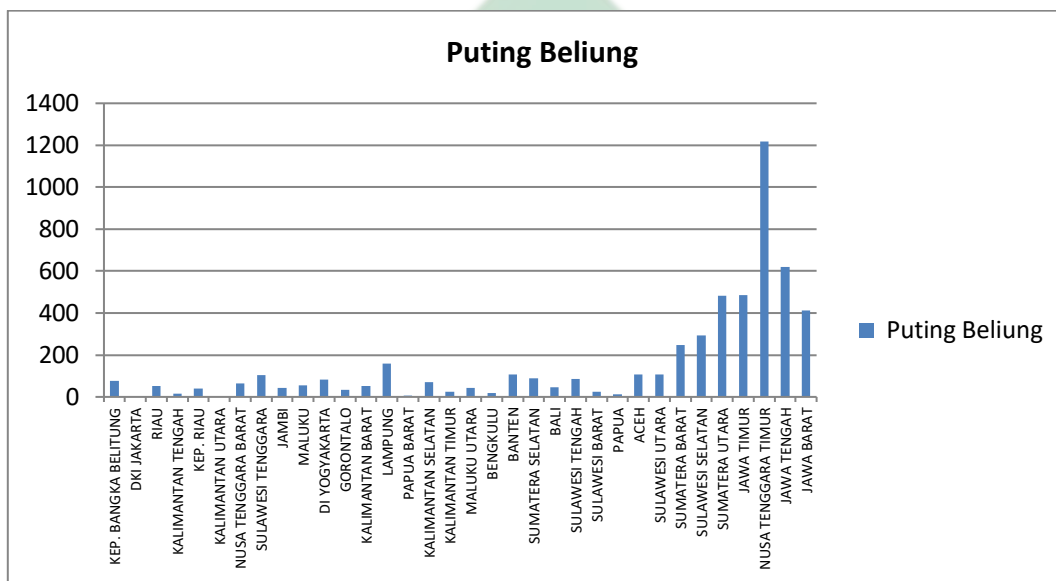
Gambar 4.5 Grafik Bencana Tsunami

Jumlah desa bencana yang terdampak tsunami pada 2021 yaitu provinsi Banten yang terdampak 1 desa, Aceh 2 desa, Sumatera Utara 4 desa, dan Jawa Timur ada 4 desa. Sedangkan provinsi-provinsi yang lain jarang sekali terdampak tsunami.



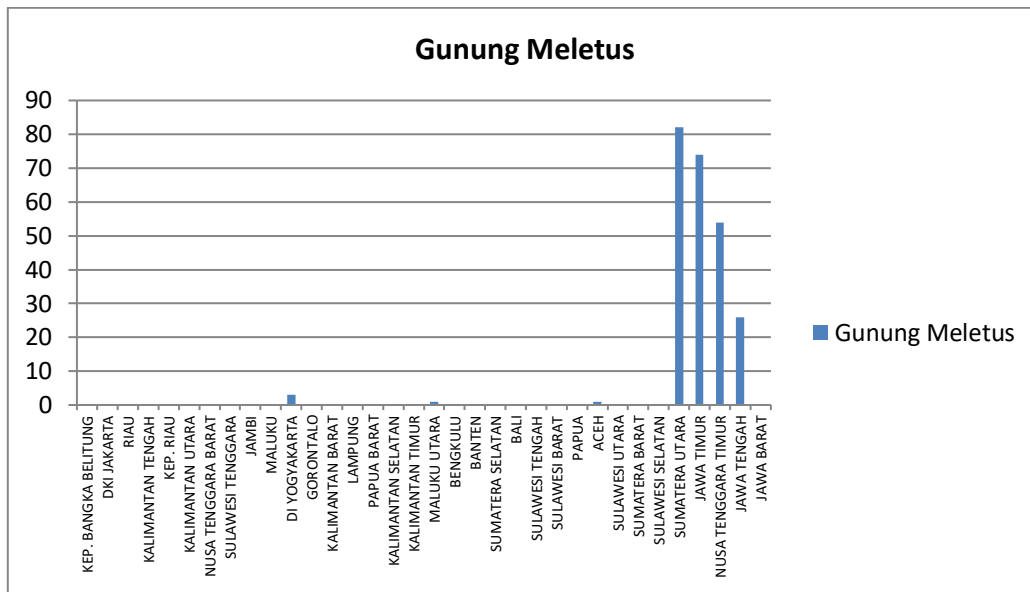
Gambar 4.6 Grafik Bencana Gelombang Pasang Laut

Jumlah desa yang terdampak yang sering terdampak bencana gelombang pasang laut pada 2021 yaitu di provinsi Sulawesi Tenggara sekitar 117 desa, Maluku Utara 124 desa, Jawa Tengah 128 desa, Maluku 165 desa, dan Nusa Tenggara Timur 218 desa. Sedangkan daerah yang jarang terdampak bencana tanah longsor yaitu provinsi Jambi 2 desa, Sumatera Selatan dan Kalimantan Utara 3 desa.



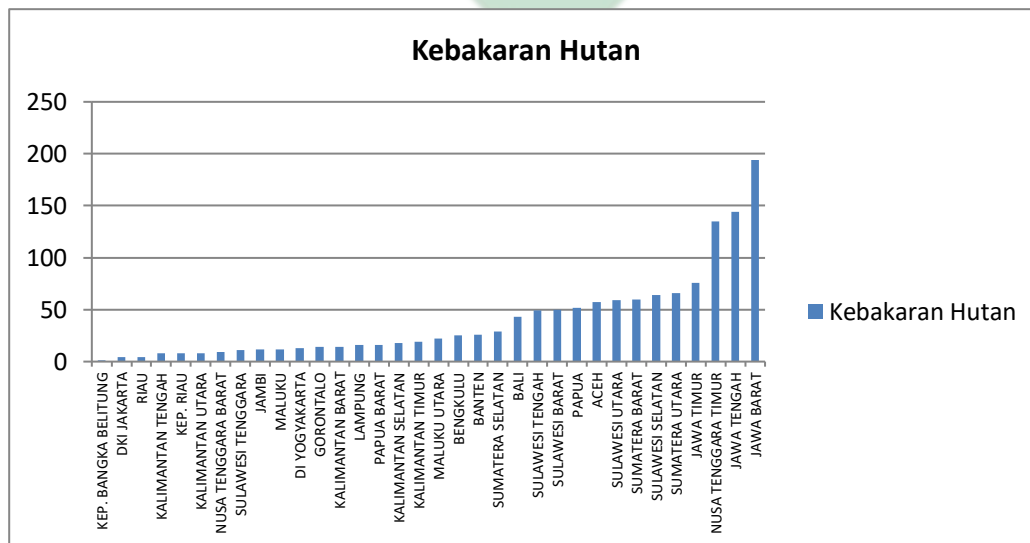
Gambar 4.7 Grafik Bencana Puting Beliung

Jumlah desa yang sering terdampak bencana puting beliung pada 2021 yaitu di provinsi Jawa Barat sekitar 412 desa, Sumatera Utara 483 desa, Jawa Timur 486 desa, Jawa Tengah 619 desa, dan Nusa Tenggara Timur 1219 desa. Sedangkan daerah yang jarang terdampak bencana tanah longsor yaitu provinsi Kalimantan Utara 3 desa, Papua Barat 6 desa, dan Papua 11 desa.



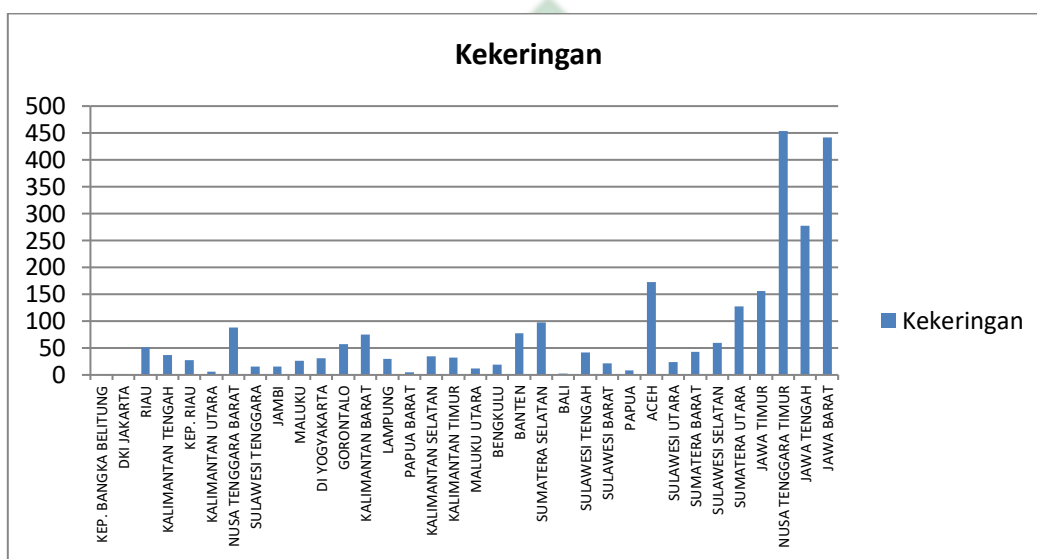
Gambar 4.8 Grafik Bencana Gunung Meletus

Jumlah desa yang sering terdampak bencana gunung meletus pada 2021 yaitu di provinsi DI Yogyakarta sekitar 3 desa, Jawa Tengah 26 desa, Nusa Tenggara Timur 54 desa, Jawa Timur 74 desa, dan Sumatera Utara 82 desa. Sedangkan provinsi-provinsi lainnya jarang terjadi Bencana tersebut.



Gambar 4.9 Grafik Bencana Kebakaran Hutan

Jumlah desa yang sering terdampak bencana kebakaran hutan pada 2021 yaitu di provinsi Kalimantan Tengah sekitar 66 desa, Jawa Barat 76 desa, Kalimantan Barat 135, Nusa Tenggara Timur 144, dan Riau 194 desa. Sedangkan daerah yang jarang terdampak bencana tanah longsor yaitu provinsi DKI Jakarta 1 desa, Bengkulu dan DIY 4 desa.



Gambar 4.10 Grafik Bencana Kekeringan

Jumlah desa yang sering terdampak bencana kekeringan pada 2021 yaitu di provinsi Jawa Timur sekitar 156 desa, Aceh 173 desa, Jawa Tengah 278 desa, Jawa Barat 442 desa, dan Nusa Tenggara Timur 454 desa. Sedangkan daerah yang jarang terdampak bencana tanah longsor yaitu provinsi Bali 2 desa, Papua Barat 5 desa, dan Kalimantan Utara 6 desa.

4.2 Hasil Normalisasi Data

Berdasarkan transformasi data tersebut, selanjutnya yaitu melakukan normalisasi data dengan menggunakan *min-max normalization* berharap hasil akurasi menjadi lebih tepat daripada tanpa menggunakan normalisasi data. Untuk menyelesaikan proses normalisasi *min-max* perlu menentukan nilai minimum dan maksimum data yang digunakan terdapat pada Tabel 4.1 berikut:

Tabel 4.1 Nilai Minimum dan Maksimum Data

	Nilai MIN	Nilai MAX
Tanah Longsor	1	1288
Banjir	39	1435
Banjir Bandang	0	151
Gempa Bumi	0	2449
Tsunami	0	4
Gelombang Pasang Laut	2	218
Puting Beliung	0	1219
Gunung Meletus	0	82
Kebakaran Hutan	1	194
Kekeringan	0	454
Data Laju Pertumbuhan Penduduk	0.6	3.48
Data Curah Hujan	0	2661

Dari data nilai minimum dan nilai maksimum tersebut dapat diperoleh hasil penyelesaian perhitungan *min-max normalization* yang dijelaskan pada Tabel 4.2 berikut:

Tabel 4.2 Hasil *Min-Max Normalization*

No.	Rumus Min-Max Tanah Longsor	Hasil		Rumus Min-Max Curah Hujan	Hasil
0	$(0.153069-1)/(1288-1)$	0		$(0.40699-0)/(2661-0)$	0.067644
1	$(0.374514-1)/(1288-1)$	0.004662		$(0.590004-0)/(2661-0)$	0.161593
2	$(0.171717-1)/(1288-1)$	0.01554		$(0.462232-0)/(2661-0)$	0.24953
3	$(0.01554-1)/(1288-1)$	0.017871		$(0.24953-0)/(2661-0)$	0.620068

4	$(0.043512-1)/(1288-1)$	0.018648		$(0.213078-0)/(2661-0)$	0.507704
5	$(0.079254-1)/(1288-1)$	0.031857		$(0.087937-0)/(2661-0)$	0
6	$(0.06216-1)/(1288-1)$	0.033411		$(0.453965-0)/(2661-0)$	0.148065
7	$(0.053613-1)/(1288-1)$	0.034965		$(0.135663-0)/(2661-0)$	0.500939
.....					
32	$(0.055167-1)/(1288-1)$	0.923854		$(0.235626-0)/(2661-0)$	0.214581
33	$(0.112665-1)/(1288-1)$	1		$(1-0)/(2661-0)$	0.316422

4.3 Hasil Uji Multikolinearitas

Uji multikolinearitas dilakukan untuk memastikan bahwa tidak ada korelasi sempurna antara variabel bebas satu dengan variabel bebas lainnya dengan melihat nilai *variance inflation factor* (VIF). Jika nilai VIF < 10 menunjukkan bahwa tidak adanya gejala multikolinearitas sedangkan jika nilai VIF > 10 menunjukkan bahwa terjadi multikolinearitas dalam model regresi. Hasil pengujian multikolinearitas yang diperoleh sebagai berikut.

Tabel 4.3 Hasil Uji Multikolinearitas Coefficients

Model	Collinearity Statistics	
	VIF	
0 Tanah Longsor	12.406498	
1 Banjir	10.048750	
2 Banjir Bandang	25.925886	
3 Gempa Bumi	6.654106	
4 Tsunami	14.156734	
5 Gelombang Pasang Laut	4.537942	
6 Puting Beliung	33.839479	

7	Gunung Meletus	25.581490
8	Kebakaran Hutan	5.040272
9	Kekeringan	23.559860
10	Laju Pertumbuhan Penduduk	2.602247
11	Data Curah Hujan	3.207321

Berdasarkan Tabel 4.3 dapat diketahui bahwa nilai VIF lebih dari 10 maka dapat disimpulkan bahwa variabel menunjukkan adanya gejala multikolinearitas karena hasilnya lebih besar dari angka 10. Berdasarkan tabel diatas, diketahui pada variabel-variabel tersebut nilai VIF yang lebih dari 10 banyak. Sehingga perlu di uji coba satu satu dilanjut dengan uji multikolinearitas lagi sampai nilai VIF memenuhi. Didapatkan 3 variabel dengan nilai VIF tertinggi yang perlu dihapus atau dieliminasi. Hasil penyelesaian tersebut ditunjukkan dalam Tabel 4.4 berikut:

Tabel 4.4 Hasil Uji Multikolinearitas Coefficients

Setelah Proses Penghapusan 3 Variabel

Model	Collinearity Statistics
	VIF
0 Tanah Longsor	7.588297
1 Banjir	5.154381
2 Gempa Bumi	4.392539
3 Tsunami	3.582691
4 Gelombang Pasang Laut	3.720577
5 Kebakaran Hutan	3.321458
6 Kekeringan	9.966563
7 Laju Pertumbuhan Penduduk	2.430550

8	Data Curah Hujan	3.551251
---	------------------	----------

Berdasarkan Tabel 4.4 dapat diketahui bahwa nilai VIF kurang dari 10 ($VIF < 10$) maka dapat disimpulkan bahwa variabel menunjukkan tidak adanya gejala multikolinearitas karena hasilnya lebih kecil dari angka 10.

4.4 Hasil *Clustering*

Pada proses *clustering* menggunakan K-Means dengan parameter yang digunakan adalah penentuan jumlah *cluster* (k) dan maksimum iterasi. Parameter ini akan berpengaruh pada penentuan *centroid* yang akan digunakan pada proses perhitungan jarak data dengan *centroid* menggunakan *Euclidean* hingga pada penentuan anggota *cluster*. Pengujian hasil *clustering* menggunakan metode *silhouette index* yang dilakukan dengan jumlah *cluster* $k=2$ sampai $k=7$.

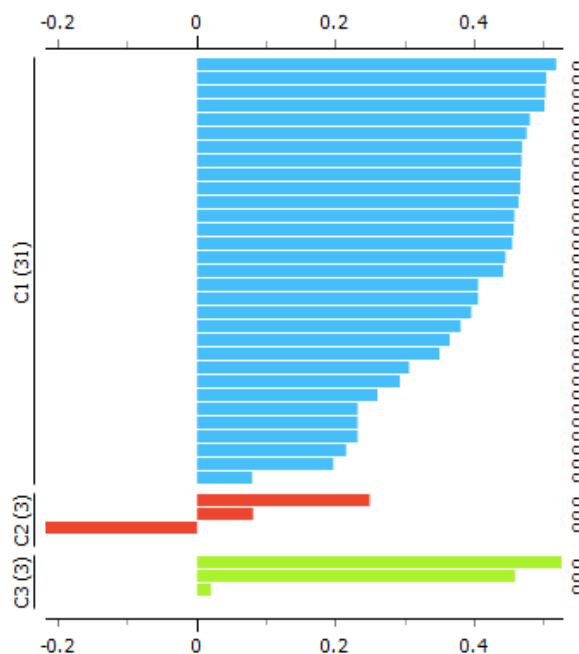
Dari hasil proses perhitungan *silhouette index* terhadap data bencana alam di 34 provinsi yang ada di Indonesia, didapatkan hasil *silhouette index* yang maksimum adalah pada $k=3$ dengan nilai *silhouette* = 0.501. nilai rata-rata *silhouette index* untuk semua *cluster* seperti terlihat pada Tabel 4.5 berikut:

Tabel 4.5 Nilai *Silhouette*

No.	Cluster	Nilai Silhouette
1	K2	0.478
2	K3	0.501
3	K4	0.202
4	K5	0.281
5	K6	0.213
6	K7	0.227

4.5 Hasil Analisis Algoritma K-Means

Setelah dilakukan implementasi dapat diamati hasil dari *clustering* menggunakan algoritma K-Means yang digunakan untuk pembentukan *cluster*. Hasil proses *clustering* dapat dilihat dalam bentuk visualisasi grafik *silhouette* plot yang menampilkan informasi data dengan warna biru menunjukkan keanggotaan *cluster* 1, data dengan warna merah menunjukkan keanggotaan *cluster* 2, dan pada keanggotaan *cluster* 3 ditunjukkan dengan warna hijau. Kemudian hasil tersebut diuji menggunakan *silhouette index* untuk semua *cluster* dengan nilai = 0.501 yaitu pada jumlah $k=3$.



Gambar 4.11 Hasil *Cluster*

Pada Gambar 4.11 diketahui bahwa hasil *cluster* dari seluruh data variabel-variabel tersebut menunjukkan nilai *silhouette* positif kecuali 1 data yaitu pada *cluster* 2. Hal ini dikarenakan jarak dengan anggota dalam cluster lebih besar dari pada jarak dengan anggota *cluster* yang berbeda. Sehingga hanya 1 data dengan nilai *silhouette* negatif.

Tabel 4.6 Hasil Perhitungan *Centroid*

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
C1	89.0	360.2	21.6	150.1	0.1	45.2	71.3	0.2	33.1	38.4	1.6	787.6
C2	899.5	1184.5	98.0	1525.0	2.0	89.0	449.0	37.0	63.0	299.0	1.1	1122.0
C3	1029.3	1004.0	106.7	320.3	0.0	145.7	750.0	26.7	89.7	391.3	1.3	640.7

Pada Tabel 4.6 diasumsikan bahwa *clustering* C1 merupakan data provinsi yang memiliki intensitas bencana alam yang tinggi atau rawan. *Clustering* C2 merupakan data provinsi yang memiliki intensitas bencana alam yang sedang atau cukup rawan. Sedangkan *clustering* C3 merupakan data provinsi yang memiliki intensitas bencana alam yang aman atau tidak rawan.

Berikut hasil pengelompokan provinsi yang terdampak bencana alam di Indonesia pada tahun 2021 sebanyak 3 *cluster* menggunakan algoritma K-Means:

Tabel 4.7 Hasil *Cluster* 34 Provinsi di Indonesia pada Tahun 2021

Kelompok	Jumlah Anggota	Provinsi yang ada di Indonesia
1	29	DKI Jakarta, Papua Barat, Nusa Tenggara Barat, Maluku, Jambi, Banten, Riau, Bengkulu, Lampung, Bali, Gorontalo, Kep. Bangka Belitung, DI Yogyakarta, Kalimantan Timur, Sulawesi Tenggara, Kep. Riau, Sumatera Selatan, Sulawesi Utara, Kalimantan Utara, Sulawesi Barat, Maluku Utara, Sulawesi Tengah, Sumatera Barat, Kalimantan Tengah, Sulawesi Selatan, Kalimantan Selatan, Kalimantan Barat, dan Papua
2	2	Aceh, Jawa Timur, dan Sumatera Utara

3	3	Jawa Tengah, Jawa Barat, dan Nusa Tenggara Timur
---	---	--

Berdasarkan Tabel 4.7 diatas, dapat diketahui bahwa hasil pengelompokan provinsi yang terdampak bencana alam di Indonesia pada tahun 2021 sebanyak 3 *cluster* dimana kelompok 3 yaitu daerah provinsi Jawa Tengah, Jawa Barat, Nusa Tenggara Timur merupakan daerah yang rawan bencana, kelompok 2 yaitu daerah provinsi Jawa Timur, Sumatera Utara merupakan daerah sedang, dan kelompok 1 yaitu selain 5 daerah provinsi tersebut (Jawa Tengah, Jawa Barat, Nusa Tenggara Timur, Jawa Timur, dan Sumatera Utara) merupakan daerah yang tidak rawan atau aman dari bencana.

Provinsi Nusa Tenggara Timur termasuk daerah rawan bencana dan Jawa Tengah menjadi daerah tertinggi rawan terhadap bencana yang menduduki peringkat kedua setelah Jawa Barat yang menjadi wilayah terbanyak terjadi bencana alam selama tahun 2021. BNPB mencatat ada sekitar 5402 kejadian bencana yang melanda Indonesia sepanjang tahun 2021 dengan komposisi mayoritas 99,5 persen dari kejadian itu merupakan bencana hidrometeorologi (Abdul, 2021). Secara morfologi wilayahnya lereng-lereng, tanahnya gempur yang bersifat lembek dan yang sensitif atau mudah menyerap air.

Penyebab utama bencana alam tersebut juga terjadi bukan tanpa sebab, ada campur tangan manusia di sektor alam dengan mengubah fungsi lahan yang seharusnya untuk resapan air namun dijadikan tempat pemukiman atau untuk bercocok tanam seiring waktu kekuatan daya tampung yang menurun. Secara umum, dengan bertambahnya frekuensi dan intensitas bencana setiap tahunnya maka pengetahuan masyarakat tentang potensi risiko menjadi mutlak untuk lebih ditingkatkan lagi.

BAB V

PENUTUP

5.1 Simpulan

Berdasarkan analisis yang telah dilakukan maka diperoleh beberapa kesimpulan sebagai berikut:

1. Hasil evaluasi uji validasi *cluster* algoritma K-Means dengan *Silhouette index* yang dilakukan dengan jumlah *cluster* $k=2$ sampai $k=7$. didapatkan jumlah *cluster* yang optimal sebanyak 3 *cluster* dengan nilai *silhouette* = 0.501 yang baik pada $k=3$.
2. Berdasarkan hasil evaluasi *cluster* dengan algoritma K-Means, didapatkan hasil 3 *cluster* terbaik dengan kategori rawan, sedang, dan tidak rawan (aman). Daerah yang tergolong rawan dan sering terjadi bencana alam di Indonesia adalah provinsi Jawa Tengah, Jawa Barat, dan Nusa Tenggara Timur karena daerah provinsi tersebut sering terjadi bencana tanah longsor, banjir, gempa bumi, tsunami, gelombang pasang laut, gunung meletus, kebakaran hutan, dan kekeringan.

5.2 Saran

Saran yang dapat diajukan pada penelitian ini adalah:

1. Untuk penelitian selanjutnya dapat ditambahkan faktor atau variabel seperti info geografis, luas lahan, kondisi cuaca dan waktu terjadinya sehingga bisa lebih optimal dalam menganalisa hasil *cluster*.
2. Untuk penelitian selanjutnya dapat menggunakan metode *cluster* yang lebih baik atau baru agar mendapat hasil *cluster* yang lebih optimal.
3. Pemerintah atau dinas yang terkait disarankan untuk melakukan pendataan ulang terkait daerah-daerah yang rawan terjadi bencana alam untuk membantu dalam penanggulangan suatu bencana yang akan terjadi suatu saat.

DAFTAR PUSTAKA

- Abdi, H. (2021). *Penyebab Tanah Longsor dan Cara Mencegahnya yang Perlu Dipahami*. liputan 6.
- Andryanto, S. D. (2021). *Penyebab Banjir Bandang*. Jakarta: Tempo.com.
- Anwar, I. C. (2017). *Tahapan Siklus Air dan Penjelasannya: dari Evapora hingga Run Off*. tirto.id.
- BNPB. (2019). *Tanggap Tangkas Tangguh Menghadapi Bencana*. Jakarta.
- BPBD. (2018). *Mengenal Bencana Kebakaran*. bpbd.kulonprogokab.go.id.
- BPBD. (2021, February 22). *DISKOMINFO KABUPATEN BOGOR*. Retrieved June 9, 2021, from <https://bogorkab.go.id>
- BPS. (2010). *Laju Pertumbuhan Penduduk*. Jakarta: sirusa.bps.go.id.
- BPS. (2022). *Bencana Alam*. Jakarta: sirusa.bps.go.id.
- Everit, S. B., Landau, S., Leese, M., & Stahl, a. D. (2011). *Cluster Analysis 5th Addition 2011*.
- F.N., A., S., P., & R., d. A. (2014). Clustering Business Rocess Model Petri Net dengan Complete Linkage. *Jurnal ITSMART* , Vvol.34 No.2 Hal. 127-138.
- Ghozali, I. (2016). *Aplikasi Analisis Multivariete Dengan Program IBM SPSS 23*. Semarang: Badan Penerbit Universitas Diponegoro, Edisi 8.
- Grossman, David A. , dan Ophir Frieder. (2004). *Information Retrieval Algorithms and Heuritics Second Edition*. Springer, The Netherlands.
- Hafiz, Y. H., Nabila, C. A., & dan Hasih, P. (2019). Analisis Cluster Intensitas Kebencanaan di Indonesia Menggunakan Metode K-Means. *Indonesia Journal of Applied Statistics*, Volume 2 No.2.
- Masdin. (2021, 01 17). *Kenali Perbedaan Gelombang Pasang Surut Air Laut dan Tsunami*. Retrieved 06 02, 2022, from unsulbarnews.com
- Muhammad, I. R., & Prihandoko. (2017). PENERAPAN DATA MINING UNTUK ANALISIS DATA BENCANA MILIK BNPB MENGGUNAKAN ALGORITMA K-MEANS DAN LINEAR REGRESION. *Fakultas Ilmu Komputer dan Teknologi*

- Nasution, D., Khotimah, H., & Chamidah, N. (2019). Perbandingan Normalisasi Data Untuk Klasifikasi Wine Menggunakan Algoritma K-NN. *CESS(Journal of Computer Engineering System and Science)*, 4(1),78-82.
- Nengsih, W. (2017). Descriptive Modelling Menggunakan K-Means untuk Pengklasteran Tingkat Kemiskinan di Provinsi Riau. 147-148.
- Ravena, I. R. (2019). Zonasi Daerah Terdampak Bencana Angin Puting Beliung Menggunakan K-Means Clustering dengan Analisis Silhouette, Davies Bouldin Dan Purity. *SKRIPSI*.
- Rimba, R. (n.d.). *Kekeringan – Pengertian, Faktor Penyebab, Dampak & Cara Mengatasi*. Retrieved June 7, 2022, from RimbaKita.com: <https://rimbakita.com/kekeringan/>
- Salsabila, M. R. (2021, February 19). *algoritma-supervised-vs-unsupervised-learning#:~:text=Unsupervised%20learning%20adalah%20salah%20satu,yang%20biasa%20disebut%20dengan%20clustering*. Retrieved June 6, 2022, from DQLab: <https://www.dqlab.id/>
- Sari, N. N. (2010). Pola Spasial Kerentanan Bencana Alam (Studi Kasus Cianjur).
- Sari, R., & Hartama, A. D. (2018). Datamining : Algoritma K-Means pada Pengelompokan Wisata Asing ke Indonesia Menurut Provinsi. *Sem. Nas. Sains Teknologi Informasi*, pp.322-326.
- Shelladita, F. S., Asep, J., & Intan, P. (2020). Pengelompokan Desa Menggunakan K-Means untuk Penyelenggaraan Penanggulangan Bencana Banjir. *Journal of Information System Vol.5, No. 2*, 156-167.
- Solmin, P., & Hisma, A. (2021). Penerapan Metode Silhouette Coefficient Untuk Evaluasi Clustering Obat. *JURNAL ILMIAH ILMU-ILMU TEKNIK*, Vol.6, No.2.
- UNISDR. (2019). *Lembaga Perserikatan Bangsa Bangsa (PBB) yang bergerak dalam pengurangan resiko bencana*.
- Wimala L. Dhanista, S. (2017). Gelombang Laut. *Teknik Kelautan ITS*.