

Implementasi *Business Intelligence* Menggunakan ETL dan *K-MEANS* untuk Visualisasi Data Bencana Alam di Jawa Barat

Devie Firmansyah¹, Deni Ramdani², Awaluddin Dongoran³, Andi Moch Januriana⁴

^{1,2}Sistem Informasi, Institut Digital Ekonomi LPKIA, Indonesia

^{3,4}Pertanahan, Sekolah Tinggi Pertanahan Nasional, Indonesia
devief@lpkia.ac.id

Info Artikel

Sejarah artikel:

Diterima September 2024

Direvisi Desember 2024

Disetujui Desember 2024

Diterbitkan Desember 2024

ABSTRACT

In addition to its abundant natural wealth, West Java also harbors a high potential for disasters. The increasing frequency of natural disasters in recent years, such as earthquakes, landslides, and floods, has caused significant losses to the community. Climate change and environmentally unfriendly human activities are the main factors causing these disasters. Therefore, mitigation and preparedness efforts are key in facing future disaster threats. Business Intelligence can be used to process data and provide relevant information, which is a part of information technology. The purpose of this research is to visualize areas affected by natural disasters in West Java Province by implementing Business Intelligence. The method of this research is obtained in the form of a dataset sourced from the West Java Provincial Community and Village Empowerment Office, obtained through <https://opendata.jabarprov.go.id/id> for the period from 2020 to 2022. In this research, ETL (Extract, Transform, Load) is used to handle data, including extraction (Extract), transformation (Transform), and loading (Load), as well as K-Means calculations to cluster regions based on clusters. The result of this research is in the form of a dashboard report with the help of Google Data Studio, displaying data visualization results for the West Java region affected by natural disasters. Based on the visualization results, the areas in West Java Province with a high risk of natural disasters that fall into the first cluster in 2020 had 2 regions, in 2021 had 4 regions, and in 2022 had only 1 region.

Keywords : Natural Disasters; Business Intelligence; Extract, Transform, Load; K-Means; Google Data Studio.

ABSTRAK

Selain kekayaan alamnya yang melimpah, Jawa Barat juga menyimpan potensi bencana yang tinggi. Frekuensi bencana alam yang meningkat dalam beberapa tahun terakhir, seperti gempa bumi, longsor, dan banjir, telah menimbulkan kerugian yang signifikan bagi masyarakat. Perubahan iklim dan aktivitas manusia yang tidak ramah lingkungan menjadi faktor utama penyebab terjadinya bencana-bencana tersebut. Oleh karena itu, upaya mitigasi dan kesiapsiagaan menjadi kunci dalam menghadapi ancaman bencana di masa depan. *Business Intelligence* dapat digunakan untuk melakukan proses data dan menyediakan informasi relevan yang menjadi salah satu bagian dari teknologi informasi. Tujuan penelitian ini untuk memvisualisasikan wilayah yang terkena dampak bencana alam di Provinsi Jawa Barat dengan mengimplementasikan *Business Intelligence*. Metode dari penelitian ini diperoleh berupa dataset yang bersumber dari Dinas Pemberdayaan Masyarakat dan Desa Provinsi Jawa Barat yang diperoleh melalui <https://opendata.jabarprov.go.id/id> dengan jangka waktu tahun 2020 sampai dengan 2022. Dalam penelitian ini ETL (*Extract, Transform, Load*) digunakan untuk menangani data yang

meliputi ekstraksi (*Extract*), transformasi (*Transform*) dan pemuatan (*Load*), serta perhitungan *K-Means* untuk mengelompokkan wilayah berdasarkan *cluster*. Hasil penelitian ini yaitu berupa laporan berbentuk *dashboard* dengan bantuan *Google Data Studio* yang menampilkan hasil visualisasi data untuk wilayah Jawa Barat yang terkena bencana alam. Berdasarkan dari hasil visualisasi, daerah Provinsi Jawa Barat yang memiliki resiko tinggi terjadi bencana alam yang termasuk kedalam *cluster* pertama pada tahun 2020 terdapat 2 wilayah, pada tahun 2021 terdapat 4 wilayah dan pada tahun 2022 hanya terdapat 1 wilayah saja.

Kata Kunci : Bencana Alam; *Business Intelligence*; *Extract, Transform, Load*; *K-Means*; *Google Data Studio*.

PENDAHULUAN

Bencana adalah suatu kejadian yang dapat membahayakan dan dapat mengganggu penduduk setempat yang dapat disebabkan oleh alam maupun manusia. Salah satu bentuk bencana yang mungkin dapat terjadi adalah bencana alam. Bencana alam menurut Undang-Undang Nomor 24 Tahun 2007 tentang Penanggulangan Bencana diartikan sebagai bencana yang disebabkan oleh peristiwa alam. Gempa bumi, letusan gunung berapi, tsunami, banjir, angin puting beliung, kekeringan dan tanah longsor merupakan contoh bencana alam yang dapat terjadi. Bencana alam dapat terjadi di wilayah Indonesia, karena menurut Badan Penanggulangan Bencana Daerah (BPBD), secara geografis Indonesia merupakan negara yang memiliki kepulauan yang luas lautannya lebih besar dibandingkan dengan daratan. Letaknya di pertemuan tiga lempeng yaitu Lempeng Pasifik, Lempeng Eurasia dan Lempeng Indo-Australia. Salah satu provinsi di Indonesia yang paling rawan terhadap bencana alam adalah Jawa Barat. Jawa Barat dipilih sebagai fokus pada penelitian ini karena pada tahun 2022 Jawa Barat menjadi wilayah yang memiliki kejadian bencana alam terbanyak dibandingkan dengan wilayah lainnya. Penyebab wilayah Jawa Barat rawan terjadi bencana alam adalah karena lokasi topografinya, kontur tanah yang bersifat lembek dan mudah menyerap air.

Business Intelligence adalah salah satu jenis teknologi informasi yang dapat digunakan untuk memproses data dan informasi [2] Memperoleh, menyimpan dan mengkategorikan data dapat dilakukan dengan *Business Intelligence* [3]. Pada saat melakukan pengolahan data juga dibutuhkan suatu metode hingga data tersebut siap untuk digunakan.

ETL (*Extract, Transform, Load*) berperan penting dalam mengumpulkan data dari berbagai sumber, membersihkan data yang tidak konsisten, dan mengolahnya menjadi format yang siap analisis. Dalam konteks bencana, ETL dapat digunakan untuk mengumpulkan data dari sensor cuaca, laporan kerusakan, data demografi, dan sumber lainnya. Dalam metode ETL terdapat tiga tahapan proses, yaitu Ekstraksi (*Extract*), Transformasi (*Transform*) dan Pemuatan (*Load*) [4]. Dengan menggunakan metode ETL data akan diambil dari sumber data, setelah itu data akan diolah dan data akan disimpan ke bentuk yang baru. Algoritma *K-Means* digunakan untuk mengelompokkan data menjadi beberapa cluster berdasarkan kesamaan karakteristik. Dalam mitigasi bencana, *K-Means* dapat digunakan untuk mengidentifikasi wilayah-wilayah yang memiliki risiko bencana tinggi berdasarkan faktor-faktor seperti kepadatan penduduk, jenis tanah, sejarah

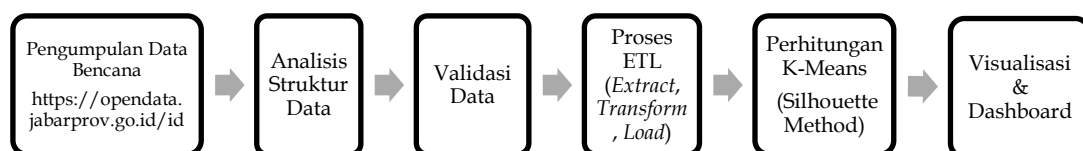
bencana, dan sebagainya. *K-Means* merupakan pendekatan pengelompokan data yang membagi data menjadi dua kategori atau lebih. Pendekatan ini membagi data menjadi satu kelompok, dimana data yang mempunyai karakteristik yang sama dikelompokkan menjadi satu dan data yang mempunyai karakteristik berbeda dikelompokkan ke dalam kelompok lain [5]. Agar data memiliki tampilan dalam bentuk visual maka diperlukan visualisasi data. Visualisasi data diperlukan agar penyajian data dapat dipahami secara menarik dan mudah dipahami [6].

Visualisasi data dapat diterapkan pada *dashboard* sehingga dapat membantu dalam melakukan analisis data. *Dashboard* dapat memberikan tampilan terpusat yang menampilkan informasi secara grafis. *Dashboard* mungkin mendapat manfaat dari pilihan grafis yang tepat untuk memastikan bahwa informasi yang diberikan jelas dan mudah dipahami. *Dashboard* menawarkan representasi visual dan mengirimkan informasi yang koheren, memungkinkan informasi untuk dipahami dan dieksplorasi dengan sekali pandang [7]. *Dashboard* sering kali menyertakan berbagai komponen visual seperti grafik, diagram, tabel, metrik dan representasi grafis lainnya [8]. Dengan tujuan memberikan informasi secara ringkas, mudah dipahami dan informatif. Platform *Google Data Studio* ini memungkinkan visualisasi data yang kompleks menjadi bentuk yang mudah dipahami. Dengan *Google Data Studio*, kita dapat membuat dashboard interaktif yang menampilkan informasi penting seperti peta risiko bencana, jumlah korban, distribusi bantuan, dan lain-lain.. *Google Data Studio* dapat digunakan untuk mengembangkan *dashboard* yang sesuai dengan kebutuhan [9]. *Google Data Studio* memudahkan pembuatan *dashboard* yang menarik secara visual dan berguna [10]. Karena sudah berbasis *cloud*, siapa saja dapat menggunakan *Google Data Studio* serta diakses dari mana saja.

Berdasarkan penelitian sebelumnya yang berjudul "*Intelligence Dashboard dengan Tableau Public untuk Visualisasi Propinsi Rawan Banjir di Indonesia*", terdapat gap penelitian terkait jangkauan penelitian dan alat yang digunakan untuk visualisasi data bencana, dimana pada penelitian ini alat yang digunakan adalah aplikasi *Tableau Public* dan bencana yang diteliti khusus bencana banjir dan jangkauan seluruh provinsi di Indonesia [1].

METODE

Terdapat beberapa tahap yang harus dilakukan pada penelitian ini seperti berikut.



Gambar 1. Tahapan Penelitian

Berdasarkan gambar diatas dapat dijelaskan sebagai berikut :

1. Mengumpulkan Data

Terdapat beberapa data yang dikumpulkan, data yang diperoleh berasal dari website <https://opendata.jabarprov.go.id/id>, data tersebut terdiri dari Data

Kejadian Bencana, Data Korban Bencana serta Data Jumlah Kejadian dan Korban Bencana.

2. Menganalisis Struktur Data

Tahap ini dilakukan analisis struktur data, setelah data berhasil didapat, yang bertujuan untuk mengetahui struktur data, karena setiap data memiliki struktur yang berbeda beda.

3. Melakukan Validasi Data

Tahap ini dilakukan validasi data yang bertujuan untuk memastikan data apakah sudah valid atau tidak, sehingga data siap untuk dilakukan pengolahan pada tahap selanjutnya.

4. Melakukan ETL

Pada tahap ETL ini dilakukan untuk mendapatkan, mengolah dan memuat data, pada ETL ini terdiri dari 3 tahapan yaitu *Extract*, *Transform* dan *Load*.

a. *Extract*

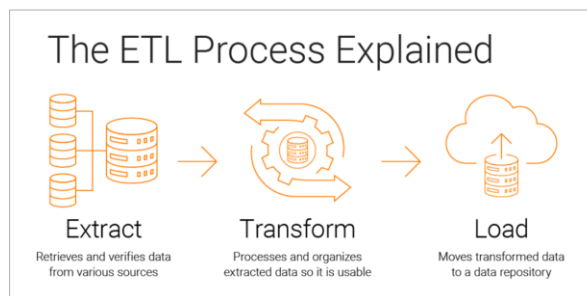
Data yang sudah dilakukan validasi pada tahap sebelumnya, dilakukan proses ekstraksi data menjadi suatu bentuk format data.

b. *Transform*

Pada tahap ini data yang telah diambil seringkali terdapat kesalahan yang harus diperbaiki atau duplikasi yang perlu dihilangkan, pembersihan data penting untuk memastikan bahwa data yang digunakan akurat.

c. *Load*

Pada tahap ini setelah melakukan tahapan sebelumnya yaitu tahap *transform*, tahap selanjutnya yaitu tahap *load*. Data dimuat kedalam suatu file untuk lanjut pada tahap selanjutnya.



Gambar 2. Tahapan *Extract*, *Transform*, *Load*

5. Melakukan Perhitungan *K-Means*

K-Means mencakup fungsi untuk melakukan pengelompokan data ke dalam *cluster*. Data diklasifikasikan ke beberapa kelompok, data dalam kategori yang sama memiliki sifat yang mirip. Berikut merupakan proses yang dilakukan dalam perhitungan *K-Means*.

a. Menentukan nilai *k* dengan metode *silhouette*, yaitu jumlah berapa *cluster* yang digunakan.

$$S_i = \frac{(b_i - a_i)}{\max(a_i, b_i)} \quad (1)$$

b. Memilih nomor acak untuk pusat awal *cluster* (pusat massa)

- c. Menghitung jarak dari setiap data ke setiap *centroid* untuk mendapatkan jarak terdekat setiap data terhadap *centroid* dengan menggunakan rumus jarak *Euclidean*, seperti berikut:

$$d(x_i, x_j) = \sqrt{(|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2)} \quad (2)$$

- d. Mengklasifikasikan dari setiap *cluster* data yang kedekatannya berdasarkan *centroid*.
- e. Memperbarui nilai *centroid* yang baru diperoleh dari rata-rata *cluster* dengan menggunakan rumus:

$$C_{m(q)} = \frac{1}{n_m} \sum_{i=1}^{n_m} x_{i(q)} \quad (3)$$

- f. Melakukan perhitungan kembali hingga data pada tiap masing masing *cluster* tidak ada yang mengalami perubahan *cluster*.
6. Membuat Visualisasi
Pada tahap ini data akan divisualisasikan menggunakan *Google Data Studio*, yang terdiri dari berbagai diagram grafik.
7. Membuat *Dashboard*
Pada tahap ini data yang sudah divisualisasikan dalam berbagai bentuk diagram grafik disatukan menjadi sebuah *dashboard*.

HASIL DAN PEMBAHASAN

Analisis Struktur Data

Struktur data yang terdapat pada masing masing data yang sudah dikumpulkan seperti pada berikut ini.

1. Data Kejadian Bencana terdiri dari id, nama_provinsi, kode_kabupaten_kota, nama_kabupaten_kota, jenis_bencana, jumlah_desa, satuan dan tahun.
2. Data Korban Bencana terdiri dari id, nama_provinsi, kode_kabupaten_kota, nama_kabupaten_kota, kondisi_korban, jumlah_korban, satuan dan tahun.
3. Data Jumlah Kejadian dan Korban Bencana terdiri dari id, nama_provinsi, kode_kabupaten_kota, nama_kabupaten_kota, tahun, jumlah_kejadian dan jumlah_korban.

Tabel 1. Struktur data

Nama Field	Deskripsi	Tipe Data
id	Merupakan id data	number
nama_provinsi	Merupakan nama provinsi	text
kode_kabupaten_kota	Merupakan kode dari kabupaten kota	number
nama_kabupaten_kota	Merupakan nama dari kabupaten kota	text
jenis_bencana	Merupakan jenis bencana yang terjadi	text
jumlah_desa	Merupakan jumlah desa yang terkena dampaknya	number
satuan	Merupakan satuan yaitu desa	text
tahun	Merupakan tahun terjadinya bencana	number

Validasi Data

Berdasarkan dari hasil observasi pada data kejadian bencana terdapat data yang tidak valid yaitu pada *field* satuan dengan data "DESA" seharusnya data pada *field* "satuan" adalah "KEJADIAN" sehingga harus dilakukan perubahan data dan ini berdampak pada *field* "jumlah_desa" yang harus dilakukan perubahan nama menjadi "jumlah_kejadian".

ETL (*Extract, Transform, Load*)

Tools aplikasi yang digunakan untuk tahapan ETL ini adalah aplikasi *cloud* dari google yaitu aplikasi *Google Data Flow*, aplikasi ini dipilih karena memudahkan proses pengambilan data untuk dilanjutkan ke visualisasi *dashboard* yang menggunakan aplikasi *Google Studio*. Hasil dari tahap ETL yang dilakukan ini terbagi menjadi 3 tahapan seperti berikut ini.

1. *Extract*

Data yang telah berhasil diambil yang berasal dari sumber data, selanjutnya diekstraksi menjadi bentuk format file *xlsx*.

2. *Transform*

Terdapat beberapa proses yang harus dilakukan agar data siap dan lebih mudah digunakan pada tahap *Load*, tahapan *transform* terdiri dari :

a. Penghapusan kolom yang tidak dibutuhkan

Kolom yang tidak dibutuhkan yaitu adalah kolom "id" dan kolom "nama_provinsi", dengan menggunakan *query*:

```
#"Removed Columns" = Table.RemoveColumns(#"Changed Type",{ "id", "nama_provinsi" }).
```

Hal ini bertujuan untuk menghilangkan kolom id dan kolom nama provinsi dikarenakan hanya satu provinsi yang dianalisis sehingga kolom id dan kolom nama provinsi tidak diperlukan dalam analisis data sehingga proses analisis data bisa dilakukan lebih cepat.

b. Mengganti data yang tidak valid

Mengganti data pada kolom "satuan" yang sebelumnya "DESA" menjadi "KEJADIAN" dengan menggunakan *query*:

```
#"Replaced Value" = Table.ReplaceValue(#"Removed Columns", "DESA", "KEJADIAN", Replacer.ReplaceText, {"satuan"}).
```

Hal ini bertujuan untuk memperbaiki kesalahan penulisan data kejadian bencana yang bertujuan untuk menghilangkan ketidakjelasan data satuan terjadinya bencana sehingga tidak terjadi kesalahpahaman ketika analisis data.

c. Merubah nama kolom

Merubah nama kolom pada kolom "jumlah_desa" menjadi "jumlah_kejadian" dengan menggunakan *query*:

```
#"Renamed Columns" = Table.RenameColumns(#"Replaced Value", {"jumlah_desa", "jumlah_kejadian"}).
```

Hal ini bertujuan untuk merubah penulisan kolom jumlah desa menjadi kolom jumlah kejadian yang bertujuan untuk menyesuaikan data kolom

yang berisi jumlah kejadian sehingga tidak terjadi kesalahpahaman ketika analisis data.

d. Mengurutkan data

Mengurutkan data pada kolom "kode_kabupaten_kota" dan kolom "tahun" dari paling terkecil hingga yang paling terbesar, dengan menggunakan *query*:

```
#"Sorted Rows" =Table.Sort("#Renamed Columns",
{"kode_kabupaten_kota",Order.Ascending}, {"tahun", Order.Ascending}).
```

3. Load

Data dimuat ke dalam bentuk suatu file untuk lanjut pada tahap selanjutnya.

id	nama_provinsi	kode_kabupaten_kota	nama_kabupaten_kota	jenis_bencana	jumlah_kejadian	satuan	tahun
1	JAWA BARAT	3201	KABUPATEN BOGOR	TANAH LONGSOR	428	KEJADIAN	2020
2	JAWA BARAT	3201	KABUPATEN BOGOR	BANJIR	311	KEJADIAN	2020
3	JAWA BARAT	3201	KABUPATEN BOGOR	GEMPA	238	KEJADIAN	2020
4	JAWA BARAT	3201	KABUPATEN BOGOR	TSUNAMI	1	KEJADIAN	2020
5	JAWA BARAT	3201	KABUPATEN BOGOR	GELOMBANG PASANG	1	KEJADIAN	2020
6	JAWA BARAT	3201	KABUPATEN BOGOR	PUYUH	216	KEJADIAN	2020
7	JAWA BARAT	3201	KABUPATEN BOGOR	GUNUNG MELETUS	2	KEJADIAN	2020
8	JAWA BARAT	3201	KABUPATEN BOGOR	KEBAKARAN	25	KEJADIAN	2020
9	JAWA BARAT	3201	KABUPATEN BOGOR	KEKERINGAN	120	KEJADIAN	2020
10	JAWA BARAT	3201	KABUPATEN BOGOR	BENCANA LAINNYA	35	KEJADIAN	2020
11	JAWA BARAT	3202	KABUPATEN SUKABUMI	TANAH LONGSOR	433	KEJADIAN	2020
12	JAWA BARAT	3202	KABUPATEN SUKABUMI	BANJIR	120	KEJADIAN	2020
13	JAWA BARAT	3202	KABUPATEN SUKABUMI	GEMPA	454	KEJADIAN	2020
14	JAWA BARAT	3202	KABUPATEN SUKABUMI	TSUNAMI	0	KEJADIAN	2020
15	JAWA BARAT	3202	KABUPATEN SUKABUMI	GELOMBANG PASANG	58	KEJADIAN	2020
16	JAWA BARAT	3202	KABUPATEN SUKABUMI	PUYUH	92	KEJADIAN	2020
17	JAWA BARAT	3202	KABUPATEN SUKABUMI	GUNUNG MELETUS	2	KEJADIAN	2020
18	JAWA BARAT	3202	KABUPATEN SUKABUMI	KEBAKARAN	53	KEJADIAN	2020
19	JAWA BARAT	3202	KABUPATEN SUKABUMI	KEKERINGAN	174	KEJADIAN	2020
20	JAWA BARAT	3202	KABUPATEN SUKABUMI	BENCANA LAINNYA	34	KEJADIAN	2020

Gambar 3. Data Kejadian Bencana Setelah Proses Transform

Data kejadian bencana yang telah dilakukan proses *transform* terdapat kolom yang dihilangkan, mengganti data yang tidak valid, merubah nama kolom dan mengurutkan data.

K-Means

Pada perhitungan *K-Means* ini data yang digunakan adalah data jumlah kejadian dan korban bencana. Sebelum dilakukan perhitungan *K-Means*, data yang sudah dikumpulkan terlebih dahulu melewati proses normalisasi data, dimulai dari membuat bentuk *Un-Normalized* sampai dengan bentuk hubungan tabel yang normal. Tahapan yang dilakukan untuk perhitungan *K-Means* adalah seperti berikut ini:

1. Menentukan jumlah *k*, *cluster* yang dibentuk adalah 3 *cluster*, seperti pada tabel dibawah ini.

Cluster	Keterangan
1	RESIKO TINGGI
2	RESIKO SEDANG
3	RESIKO RENDAH

2. Hasil nilai acak yang digunakan untuk pusat dari cluster awal atau (*centroid*) sebanyak nilai *k* seperti tabel dibawah ini.

Centroid	X	Y
C1	211	31141
C2	754	3188
C3	535	1232

Keterangan

X = Jumlah Kejadian

Y = Jumlah Korban

3. Berikut ini merupakan hasil dari perhitungan *Euclidean Distance* yang dapat menentukan jarak terdekat setiap data ke *centroid*.

Tabel 4. Hasil Perhitungan *Euclidean Distance*

Data	C1	C2	C3	Jarak Terdekat
1	1922,86	29488,58	31449,27	1922,86
2	27978,13	666,00	2147,81	666,00
3	27262,06	709,23	2676,55	709,23
4	232952,46	260905,01	262861,04	232952,46
5	20983,08	7047,64	9015,83	7047,64
6	30875,60	2969,35	1274,98	1274,98
7	6052,55	21936,29	23894,29	6052,55
8	30044,85	2114,89	166,82	166,82
...
54	24623,01	3369,89	5294,39	3369,89

4. Hasil klasifikasi *cluster* setiap data menurut kedekatannya nilai *centroid* dengan jarak terdekat seperti pada tabel di bawah ini.

Tabel 5. Hasil Klasifikasi Cluster

Data	Jarak Terdekat	Cluster
1	1922,86	1
2	666,00	2
3	709,23	2
4	232952,46	1
5	7047,64	2
6	1274,98	3
7	6052,55	1
8	166,82	3
...
54	3369,89	2

5. Berikut merupakan hasil dari perhitungan nilai dari *centroid* yang baru didapat berdasarkan rata rata nilai *cluster*.

Tabel 6. Hasil Perhitungan *Centroid* Baru

Centroid	X	Y
C1	534,05	83479,95
C2	585,92	6815,54
C3	466,43	749,81

6. Berikut merupakan hasil dari perhitungan *Euclidean Distance* kembali, klasifikasi *cluster* dan perbandingan dengan *cluster* sebelumnya, menggunakan *centroid* yang baru.

Tabel 7. Hasil Perhitungan Kembali dan Perbandingan Cluster Data Jarak Terdekat Cluster Perbandingan dengan Cluster Sebelumnya

Data	Jarak Terdekat	Cluster	Perbandingan dengan Cluster Sebelumnya
1	25866,56	2	BERUBAH
2	2618,96	3	BERUBAH
3	2942,67	2	TIDAK BERUBAH
4	180613,11	1	TIDAK BERUBAH
5	3500,49	2	TIDAK BERUBAH
6	1034,38	3	TIDAK BERUBAH
7	18310,60	2	BERUBAH
8	348,44	3	TIDAK BERUBAH
...
54	458,56	2	TIDAK BERUBAH

Tabel 8. Hasil Perhitungan Kembali Centroid Baru

Centroid	X	Y
C1	412,71	162812,29
C2	599,38	40762,54
C3	512,12	3069,06

Berdasarkan tabel 7, pada hasil perhitungan terdapat *cluster* yang berubah, maka diperlukan perhitungan kembali, dari perhitungan nilai *centroid* yang baru, perhitungan *Euclidean Distance*, klasifikasi *cluster* dan perbandingan dengan *cluster* sebelumnya. Berikut merupakan hasil perhitungan kembali dari nilai *centroid* yang baru. Berikut merupakan hasil dari perhitungan *Euclidean Distance* kembali, klasifikasi *cluster* dan perbandingan dengan *cluster* sebelumnya, menggunakan *centroid* yang baru.

Tabel 9. Hasil Perhitungan Kembali dan Perbandingan Cluster Data Jarak Terdekat Cluster Perbandingan dengan Cluster Sebelumnya

Data	Jarak Terdekat	Cluster	Perbandingan dengan Cluster Sebelumnya
1	8129,81	2	TIDAK BERUBAH
2	915,77	3	TIDAK BERUBAH
3	894,42	3	TIDAK BERUBAH
4	101281,06	1	TIDAK BERUBAH
5	7194,84	3	TIDAK BERUBAH
6	2916,93	3	TIDAK BERUBAH
7	15640,81	2	TIDAK BERUBAH
8	1973,49	3	TIDAK BERUBAH
...
54	3459,90	3	TIDAK BERUBAH

Berdasarkan tabel 9, hasil perhitungan *cluster* sudah tidak ada yang berubah, maka perhitungan telah selesai. Wilayah yang mengalami bencana alam dikelompokkan menjadi tiga *cluster*. Berikut merupakan hasil dari pengelompokkan *cluster* wilayah.

Tabel 10. Pengelompokan Wilayah

id	Nama_kabupaten_kota	Tahun	Cluster
1	KABUPATEN BOGOR	2020	2
2	KABUPATEN SUKABUMI	2020	3
3	KABUPATEN CIANJUR	2020	3
4	KABUPATEN BANDUNG	2020	1
5	KABUPATEN GARUT	2020	3
6	KABUPATEN TASIKMALAYA	2020	3
7	KABUPATEN CIAMIS	2020	2
8	KABUPATEN KUNINGAN	2020	3
...
54	KABUPATEN PANGANDARAN	2022	3

Berikut ini merupakan tabel *cluster* wilayah yang diurutkan berdasarkan tahun. Pada tahun 2020 wilayah yang termasuk kedalam *cluster* pertama terdapat 2 wilayah yaitu wilayah Kabupaten Bandung dan Kabupaten Karawang, wilayah yang termasuk kedalam *cluster* kedua sebanyak 5 wilayah dan wilayah yang termasuk kedalam *cluster* ketiga sebanyak 11 wilayah.

Tabel 11. Cluster Wilayah Tahun 2020

Tahun	Cluster 1 (Resiko Tinggi)	Cluster 2 (Resiko Sedang)	Cluster 3 (Resiko Rendah)
2020	KABUPATEN BANDUNG	KABUPATEN BOGOR	KABUPATEN SUKABUMI
	KABUPATEN KARAWANG	KABUPATEN CIAMIS	KABUPATEN CIANJUR
		KABUPATEN CIREBON	KABUPATEN GARUT
		KABUPATEN INDRAMAYU	KABUPATEN TASIKMALAYA
		KABUPATEN BEKASI	KABUPATEN KUNINGAN
			KABUPATEN MAJALENGKA
			KABUPATEN SUMEDANG
			KABUPATEN SUBANG
			KABUPATEN PURWAKARTA
			KABUPATEN BANDUNG BARAT
			KABUPATEN PANGANDARAN

Pada tahun 2021 wilayah yang termasuk kedalam *cluster* pertama terdapat 4 wilayah yaitu wilayah Kabupaten Indramayu, Kabupaten Subang, Kabupaten Karawang dan Kabupaten Bekasi, wilayah yang termasuk kedalam *cluster* kedua sebanyak 3 wilayah dan wilayah yang termasuk kedalam *cluster* ketiga sebanyak 11 wilayah.

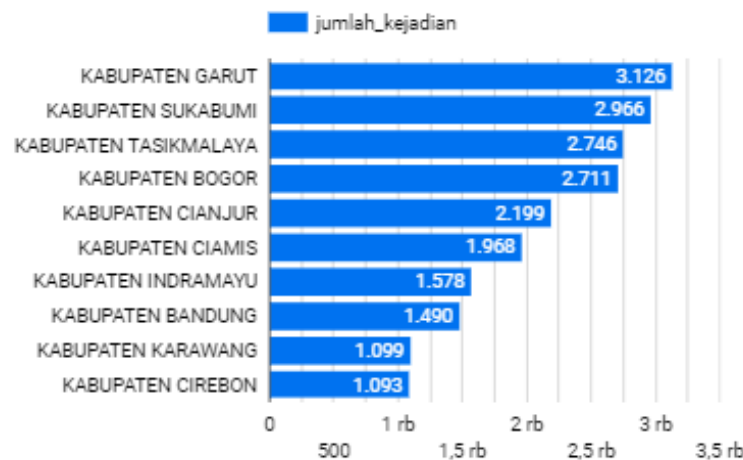
Tabel 22. Cluster Wilayah Tahun 2022

Tahun	Cluster 1 (Resiko Tinggi)	Cluster 2 (Resiko Sedang)	Cluster 3 (Resiko Rendah)
2022	KABUPATEN CIANJUR	KABUPATEN BOGOR	KABUPATEN SUKABUMI
		KABUPATEN BANDUNG	KABUPATEN TASIKMALAYA
		KABUPATEN GARUT	KABUPATEN CIAMIS
		KABUPATEN CIREBON	KABUPATEN KUNINGAN
		KABUPATEN KARAWANG	KABUPATEN MAJALENGKA
			KABUPATEN SUMEDANG
			KABUPATEN INDRAMAYU
			KABUPATEN SUBANG
			KABUPATEN PURWAKARTA
			KABUPATEN BEKASI
			KABUPATEN BANDUNG BARAT
			KABUPATEN SUKABUMI

Visualisasi

Berikut merupakan hasil visualisasi dari data dengan menggunakan *Google Data Studio*, setelah data diproses pada bagian sebelumnya.

1. Diagram Bagan Batang

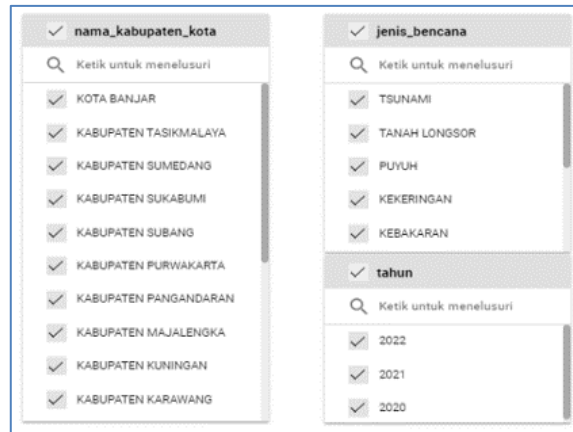


Gambar 4. Diagram Bagan Batang

Diagram 4 merupakan diagram batang yang menampilkan jumlah kejadian yang terjadi per kabupaten/kota, dengan data yang tampil mengurut berdasarkan yang paling banyak hingga paling sedikit, khususnya di wilayah Kabupaten Garut yang memiliki total 3.126 kejadian.

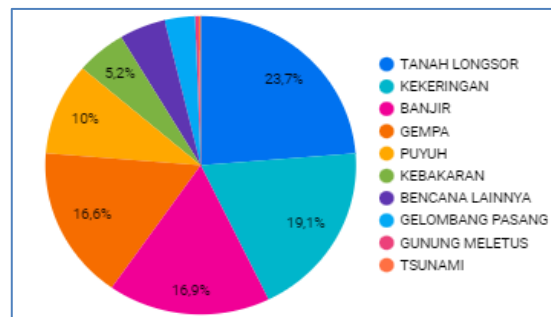
2. Kontrol Daftar Ukuran Tetap

Gambar 5 merupakan hasil dari kontrol daftar ukuran tetap yang terdiri dari nama kabupaten, kota, jenis bencana atau tahun, serta fungsinya untuk dapat memilih data yang akan ditampilkan, yang dapat ditampilkan semua data, banyak data, atau hanya satu pilihan data dengan mencentang sebagian kotak data yang akan dipilih.



Gambar 5. Kontrol Daftar Ukuran Tetap

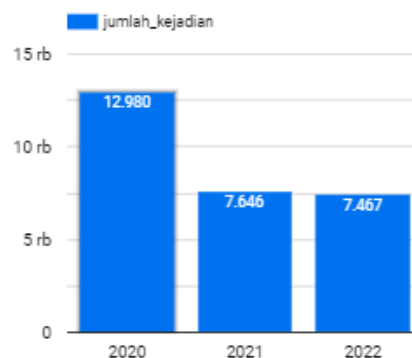
3. Diagram Lingkaran



Gambar 6. Diagram Lingkaran

Diagram lingkaran pada gambar 6 memungkinkan untuk mengamati data bencana berupa nilai % dan segala bentuk bencana di wilayah Jawa Barat. Diagram di atas menunjukkan bahwa jenis bencana dengan persentase terbesar adalah bencana Tanah Longsor yang memiliki persentase sebesar 23,7%, sedangkan jenis bencana yang memiliki persentase terendah adalah bencana Tsunami.

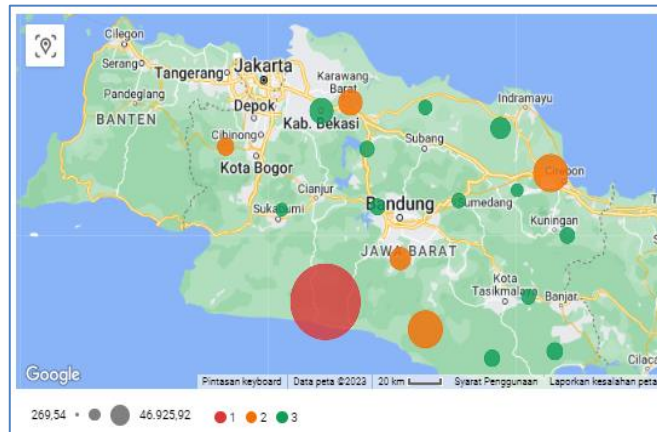
4. Diagram Bagan Kolom



Gambar 7. Diagram Bagan Kolom

Diagram 7 merupakan diagram kolom yang menggambarkan jumlah peristiwa yang terjadi menurut tahun, dengan kejadian terbesar terjadi pada tahun 2020 sebanyak 12.980 kejadian, tahun 2021 sebanyak 7.646 kejadian dan tahun 2022 sebanyak 7.467 kejadian.

5. Diagram Peta Balon

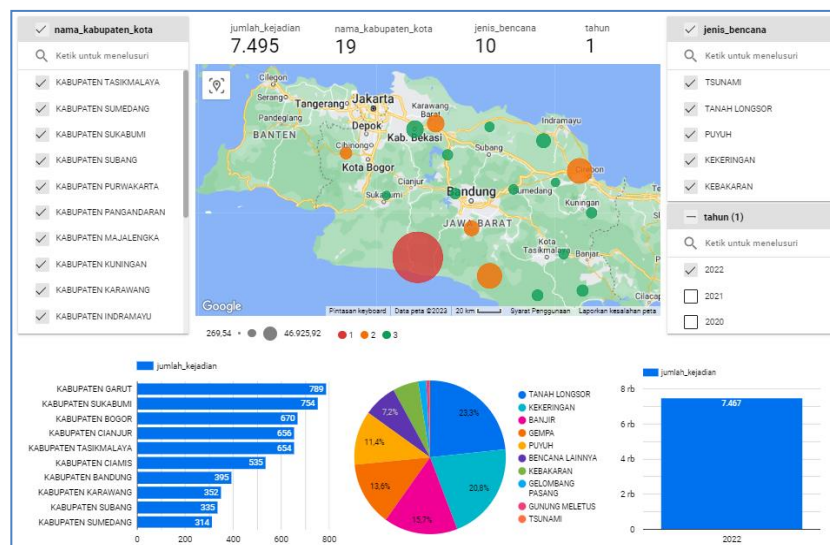


Gambar 8. Diagram Peta Balon

Gambar 8 menampilkan peta balon, yang dapat digunakan untuk mempermudah memfilter data di setiap wilayah di Jawa Barat cukup dengan mengarahkan kursor atau klik pada daerah dengan bentuk lingkaran yang memiliki warna untuk dapat melihat data bencana pada masing masing wilayah yang dipilih.

Dashboard

Hasil *dashboard* yang terdiri dari berbagai bentuk diagram grafik dari data yang sudah divisualisasikan.



Gambar 9. Dashboard Bencana Alam

Gambar 9 menunjukkan representasi *dashboard* yang dibuat dengan *Google Data Studio*. *Dashboard* ini memiliki berbagai macam diagram, seperti diagram bagan batang, diagram lingkaran, diagram bagan kolom, diagram peta balon dan daftar kontrol.

PENUTUP

Setelah dilakukan proses ETL (*Extract, Transform, Load*) pada data yang sudah didapatkan selanjutnya dilakukan proses perhitungan *K-Means* dengan

metode *Silhouette* didapatkan hasil untuk wilayah Provinsi Jawa Barat yang terdampak bencana alam yang memiliki resiko tinggi terjadi bencana alam yang termasuk kedalam *cluster* pertama pada tahun 2020 terdapat 2 wilayah yaitu wilayah Kabupaten Bandung dan Kabupaten Karawang, sedangkan pada tahun 2021 terdapat 4 wilayah yaitu wilayah Kabupaten Indramayu, wilayah Kabupaten Subang, wilayah Kabupaten Karawang serta wilayah Kabupaten Bekasi dan pada tahun 2022 hanya terdapat 1 wilayah saja yaitu wilayah Kabupaten Cianjur. Informasi dari wilayah yang terdampak bencana alam dibuatkan visualisasi data dalam bentuk *dashboard* menggunakan aplikasi *Google Data Studio* yang dapat menyajikan informasi terkait data secara terpusat dan mampu menampilkan informasi yang relevan sehingga dengan mudah dapat dipahami.

DAFTAR PUSTAKA

- [1] Sariasih, F. A. . (2022). Implementasi Business Intelligence Dashboard dengan Tableau Public untuk Visualisasi Propinsi Rawan Banjir di Indonesia. *Jurnal Pendidikan Tambusai*, 6(2), 14424–14431. doi.org/10.31004/jptam.v6i2.4715.
- [2] M. R. Atsani, G. Tyas Anjari, and N. Mega Saraswati, *Pengembangan Business Intelligence Di Rumah Sakit (Studi Kasus: RSUD Prof. Dr. Margono Soekarjo Purwokerto)*, vol. 12, no. 2. dspace.uui.ac.id, 2019. doi: 10.35671/telematika.v12i2.839.
- [3] J. C. Nugroho, I. N. Y. A. Wijaya, and A. A. N. Redioka, “Penerapan Aplikasi Business Intelligence Pada Manajemen Report Guna Menunjang Pengambilan Keputusan,” *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 10, no. 2, p. 335, 2021, doi: 10.35889/jutisi.v10i2.671.
- [4] I. P. W. Prasetya and I. N. H. Kurniawan, “Implementasi ETL (Extract, Transform, Load) pada Data warehouse Penjualan Menggunakan Tools Pentaho,” *TIERS Inf. Technol. J.*, vol. 2, no. 1, pp. 1–8, 2021, doi: 10.38043/tiers.v2i1.2844.
- [5] I. Nuryani and D. Darwis, “Analisis Clustering Pada Pengguna Brand Hp Menggunakan Metode K-Means,” *Pros. Semin. Nas. Ilmu Komput.*, vol. 1, no. 1, p. 2021, 2021.
- [6] W. Irmayani, “Visualisasi Data Pada Data Mining Menggunakan Metode Klasifikasi,” *J. Khatulistiwa Inform.*, vol. IX, no. I, pp. 68–72, 2021.
- [7] K. K. Dewi, A. Hermawan, and L. W. Kusuma, “Penerapan Dashboard Business Intelligence Untuk Menampilkan Fundamental Saham Lq45,” *Algor*, vol. 3, no. 1, pp. 60–70, 2021, doi: 10.31253/algor.v3i1.768.
- [8] I. Wahyudi and A. Syazili, “Dashboard Monitoring Website Dosen Studi Kasus Universitas Bina Darma,” *J. Pengemb. Sist. Inf. dan Inform.*, vol. 2, no. 3, pp. 188–197, 2021, doi: 10.47747/jpsii.v2i3.555.
- [9] Heri Purnadi, “Pemanfaatan Google Spreadsheet Dan *Google Data Studio* Sebagai Dashboard Suhu Dan Kelembaban Di Laboratorium,” *Insa. Metrol. PPSDK*, vol. 1, no. 1, pp. 28–33, 2021, doi: 10.55101/ppsdk.v1i1.639.
- [10] S. A. Hendrawan and D. Setyantoro, “Pemanfaatan Dashboard Business Intelligence untuk Laporan Penjualan pada Superstore,” *Tekinfo J. Bid. Tek. Ind. dan Tek. Inform.*, vol. 23, no. 1, pp. 46–52, 2022, doi: 10.37817/tekinfo.v23i1.1876.