

Perbandingan Dampak Bencana Angin Kencang Tahun 2020 Dan 2021 Daerah Istimewa Yogyakarta Berdasarkan Metode *K-means Clustering*

Thalia Nanda Khoffah^{1*}, Rohmatul Fajriyah¹

¹ Program Studi Statistika, Universitas Islam Indonesia, Jl. Kaliurang KM 14,5, Kabupaten Sleman Daerah Istimewa Yogyakarta, 55584, Indonesia

*Corresponding author: 20611166@students.uii.ac.id



P-ISSN: 2986-4178
E-ISSN: 2988-4004

Riwayat Artikel

Dikirim: 02 September 2023
Direvisi: 26 Maret 2024
Diterima: 29 Maret 2024

ABSTRAK

Penelitian ini bertujuan untuk mengklasifikasikan kecamatan-kecamatan yang terdapat di Daerah Istimewa Yogyakarta berdasarkan data jumlah kejadian angin kencang beserta dampaknya yang meliputi pohon tumbang, akses jalan, dan jaringan listrik sehingga bisa terlihat mana saja daerah di Yogyakarta yang rawan bencana angin kencang dan mana daerah yang tidak rawan terhadap bencana angin kencang. Metode yang digunakan yaitu *K-means Clustering* dengan menggunakan matriks yang relevan dan *Silhouette score*. Pada tahun 2020 *cluster* pertama terdapat 3 kecamatan dengan indikator rawan angin kencang tinggi pada setiap kecamatan tersebut, *cluster* dua terdapat 19 kecamatan dengan indikator rawan angin kencang sedang pada setiap kecamatan, dan *cluster* ketiga terdapat 56 kecamatan dengan indikator rawan bencana angin kencang rendah. Sedangkan pada tahun 2021 *cluster* pertama terdapat 3 kecamatan dengan indikator rawan angin kencang tinggi pada setiap kecamatan tersebut, *cluster* dua terdapat 10 kecamatan dengan indikator rawan angin kencang sedang pada setiap kecamatan, dan *cluster* ketiga terdapat 65 kecamatan dengan indikator rawan bencana angin kencang rendah.

Kata Kunci: Angin Kencang, Klaster, *K - means*

ABSTRACT

This study aims to classify the sub-districts in the Yogyakarta Special Region based on the data on the number of strong wind events and their impacts, including fallen trees, road access, and power lines, to identify which areas in Yogyakarta are prone to strong wind disasters and which areas are not. The method used is K-means Clustering with relevant matrices and Silhouette scores. In 2020, the first cluster consisted of 3 sub-districts with a high indicator of strong wind vulnerability in each sub-district, the second cluster consisted of 19 sub-districts with a moderate indicator of strong wind vulnerability in each sub-district, and the third cluster consisted of 56 sub-districts with a low indicator of strong wind disaster vulnerability. Meanwhile, in 2021, the first cluster consisted of 3 sub-districts with a high indicator of strong wind vulnerability in each sub-district, the second cluster consisted of 10 sub-districts with a moderate indicator of strong wind vulnerability in each sub-district, and the third cluster consisted of 65 sub-districts with a low indicator of strong wind disaster vulnerability.

Keywords: Strong winds, Clustering, *K- Means*

1. Pendahuluan

Indonesia terletak di atas lempeng bumi dan dikelilingi oleh dua samudra, sehingga negara ini sangat rentan terhadap bencana yang disebabkan oleh faktor geologis. Bencana yang terkait dengan siklus-siklus alam di bumi, seperti pergerakan lempeng tektonik dan erupsi gunung api, serta faktor geologis seperti jenis tanah dan batuan, pola aliran sungai, topografi wilayah, lipatan dan patahan, dan struktur tektonik merupakan contoh bencana yang disebabkan oleh faktor geologis. [1]. Oleh karena itu, Indonesia sering dilanda oleh bencana alam yang banyak sekali menyebabkan kerusakan-kerusakan bagi Indonesia, mulai dari infrastruktur sampai mengancam masyarakat. Menurut Undang-undang Nomor 24 Tahun 2007 Tentang Penanggulangan Bencana menyebutkan definisi bencana yaitu suatu peristiwa atau rangkaian peristiwa yang mengancam dan mengganggu kehidupan dan penghidupan masyarakat yang disebabkan, baik oleh faktor alam dan/atau faktor nonalam maupun faktor manusia sehingga mengakibatkan timbulnya korban jiwa manusia, kerusakan lingkungan, kerugian harta benda, dan dampak psikologis [2].

Angin kencang adalah fenomena yang sering terjadi di Indonesia, terutama pada saat transisi dari musim kemarau ke musim hujan. Menurut aturan yang dikeluarkan oleh Badan Meteorologi Klimatologi dan Geofisika (BMKG) pada tahun 2010, angin kencang adalah angin dengan kecepatan di atas 25 knot (sekitar 45 km/jam). Angin puting beliung adalah jenis angin kencang yang berputar dan keluar dari awan cumulonimbus, dengan kecepatan lebih dari 34,8 knot (sekitar 64,4 km/jam) dan terjadi dalam waktu singkat [3]. Kejadian ini dapat menyebabkan kerusakan pada bangunan rumah dengan angin yang sangat kencang bisa merobohkan rumah yang dihuni dalam sekejap, lalu dampak lain yang dihasilkan dari bencana angin kencang juga bisa merusak jalanan karena bisa merobohkan pohon yang sangat besar dan akan merusak jalanan yang ada di area pepohonan tersebut sehingga akan sangat meresahkan bagi masyarakat yang tinggal di daerah yang rawan angin kencang.

Wilayah Daerah Istimewa Yogyakarta memiliki potensi bencana yang cukup tinggi karena kondisi geografis, geologis, hidrologis, dan demografinya. Faktor yang dapat menyebabkan bencana di wilayah tersebut antara lain faktor alam, faktor non alam, dan faktor manusia. Salah satu contoh bencana yang sering terjadi di wilayah tersebut adalah angin kencang, yang awalnya hanya terjadi di daerah Sleman dan Yogyakarta, tetapi kini telah menyebar ke daerah lain seperti Kabupaten Bantul dan Gunung Kidul. Dampak dari bencana angin kencang ini sangat merugikan masyarakat yang tinggal di wilayah rawan bencana tersebut, seperti kerusakan rumah, akses jalan yang terhambat akibat pohon tumbang, serta gangguan listrik dan internet karena banyaknya tiang listrik yang tumbang. [4].

Berdasarkan pandangan Budi Santosa dan Ardian Umam, *clustering* merupakan proses pembagian obyek dari suatu set data ke dalam beberapa klaster yang memiliki kesamaan karakteristik. Tujuan utama dari metode *clustering* adalah untuk mengelompokkan sejumlah data atau obyek ke dalam klaster sehingga setiap klaster berisi data yang memiliki kemiripan sebanyak mungkin. Dalam *clustering*, upaya dilakukan untuk menempatkan obyek yang serupa atau memiliki jarak yang dekat di dalam satu klaster, sementara jarak antar klaster diupayakan sejauh mungkin. Hal ini mengindikasikan bahwa obyek dalam satu klaster memiliki kemiripan yang tinggi satu sama lain dan berbeda dengan obyek dalam klaster-klaster lainnya yang lain [5]. Dalam konteks penelitian ini, analisis klaster akan membantu mengidentifikasi daerah-daerah yang rawan terjadinya bencana angin kencang di Provinsi Yogyakarta. Informasi tersebut dapat memberikan manfaat yang besar bagi Badan Penanggulangan Bencana Daerah (BPBD) Daerah Istimewa Yogyakarta dalam upaya mitigasi dan penanggulangan bencana angin kencang. Dengan adanya identifikasi daerah rawan, BPBD dapat memberikan perhatian khusus dan meningkatkan kesiapsiagaan dalam menghadapi potensi bencana tersebut.

2. Metodologi Penelitian

Dalam penelitian ini, digunakan data sekunder dalam bentuk data runtun waktu bencana alam angin kencang yang terjadi di Daerah Istimewa Yogyakarta sejak Januari 2020 hingga Desember 2021. Data tersebut diperoleh dari kantor Pusat Pengendalian Operasi Badan Penanggulangan Bencana Daerah. Penelitian ini menggunakan seluruh kecamatan yang ada di Daerah Istimewa Yogyakarta sebagai objek penelitian. Terdapat empat variabel yang digunakan dalam penelitian ini yaitu kecamatan, jumlah kejadian, pohon tumbang, akses jalan, dan jaringan listrik. Dalam melakukan analisis, digunakan perangkat lunak *Microsoft Excel* dan *Rstudio*.

Dalam penelitian ini, digunakan metode analisis deskriptif dan *K-means* sebagai metode *clustering non-hierarchical*. Metode analisis deskriptif digunakan untuk memberikan gambaran yang komprehensif mengenai karakteristik data, sementara *K-means* digunakan untuk mengelompokkan data ke dalam klaster berdasarkan kemiripan karakteristik. Proses standarisasi data dilakukan menggunakan rumus Z-Score untuk menyeimbangkan nilai data yang memiliki besaran angka dengan rentang yang cukup jauh. Standarisasi ini penting karena perbedaan skala data dapat memengaruhi hasil analisis secara signifikan. Dengan menggunakan Z-Score, setiap nilai dalam distribusi data diubah ke dalam nilai standar yang menggambarkan seberapa jauh nilai tersebut dari rerata populasi dalam satuan standar deviasi. Ini memungkinkan perbandingan yang lebih akurat antar variabel dalam analisis yang dilakukan. Dengan demikian, kombinasi antara metode analisis deskriptif, *K-means clustering*, dan standarisasi data dengan Z-Score diharapkan dapat memberikan pemahaman yang lebih mendalam mengenai pola-pola yang ada dalam data dan memfasilitasi pengambilan keputusan yang lebih baik [6]. Tahapan yang dilakukan dalam penelitian ini terdiri dari beberapa langkah, yaitu:

1. Mulai
2. Mengumpulkan data mentah data bencana alam angin kencang berdasarkan dampaknya pada tahun 2020 dan 2021.
3. Melakukan *cleaning* data yang akan diteliti.
4. Melakukan penginputan data bencana alam angin kencang berdasarkan dampaknya pada tahun 2020 dan 2021 ke *software* yaitu *Rstudio*.
5. Dilakukan analisis deskriptif untuk memperoleh gambaran data bencana alam angin kencang tahun 2020 dan 2021.
6. Melakukan pengujian asumsi meliputi uji multikolinearitas dengan menggunakan nilai VIF dan deteksi *outlier*.
7. Melakukan standarisasi data dikarenakan terdapat perbedaan variansi yang cukup jauh pada data, dengan melakukan persamaan sebagai berikut.

$$Z = \frac{(x - \mu)}{\sigma} \quad (1)$$

Z: standarisasi data, x : nilai dari data yang akan di standarisasi, μ : nilai rata – rata seluruh data, σ : nilai standar deviasi dari seluruh data,

8. Menentukan jumlah *cluster* k dengan metode *sillhouette*.
9. Menentukan nilai *centroid cluster*.
10. Mengambil beberapa sampel secara acak, untuk tahapan iterasi dapat menggunakan rumus persamaan seperti berikut:

$$\bar{V}_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \quad (2)$$

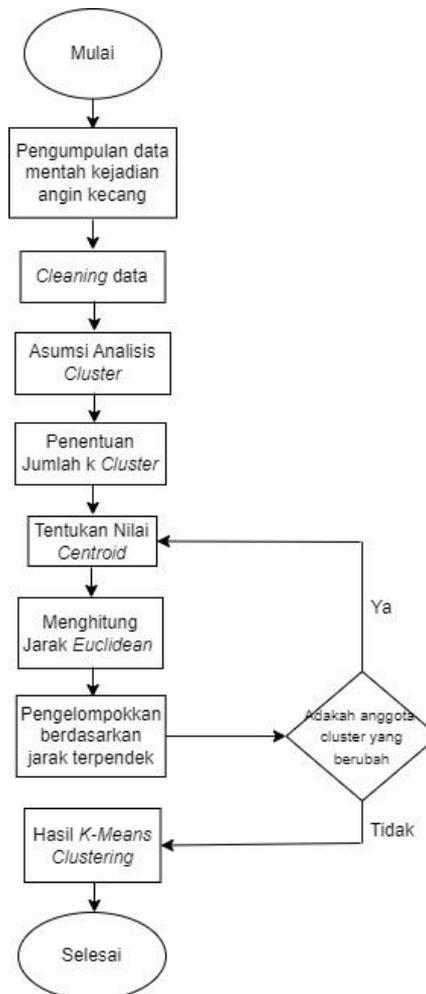
\bar{V}_{ij} : *centroid* rata – rata, N_i : jumlah *cluster*, i, k : indikator dari *cluster*, j : indikator dari variabel, X_{kj} : nilai data yang ada dalam *cluster*.

11. Melakukan penghitungan jarak antar data dari setiap pusat *cluster* menggunakan rumus jarak *Euclidean* seperti berikut:

$$d_{ij} = \sqrt{\sum_{k=1}^n ((x_{ik} - x_{jk})^2)} \quad (3)$$

Keterangan, d_{ij} : jarak data dan *centroid cluster*, x_{ik} = perhitungan jarak data ke- i , x_{jk} = *centroid cluster* j , n = jumlah variabel

12. Melakukan pengelompokan berdasarkan jarak terpendek, jika tidak terdapat *centroid* yang berbeda, maka diperoleh *cluster* dan anggota setiap *cluster*.
 13. Melakukan *profiling* setiap *cluster* yang telah didapatkan.

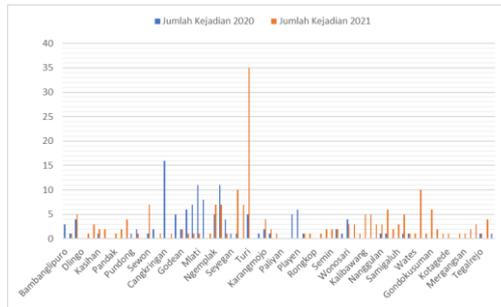


Gambar 1 Diagram Alir

3. Hasil dan Pembahasan

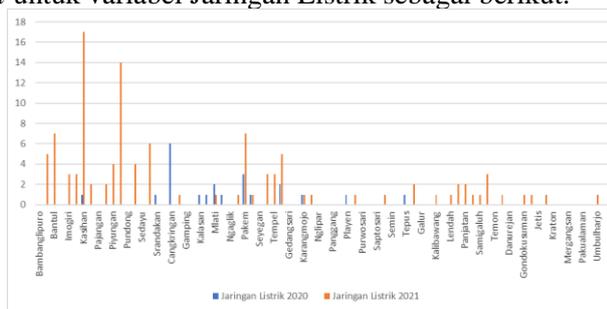
3.1. Analisis Deskriptif

Tujuan dari analisis deskriptif adalah untuk memberikan deskripsi yang jelas dan rinci mengenai karakteristik suatu informasi agar informasi yang disajikan mudah untuk dipahami dan bersifat informatif. Analisis ini menggunakan data bencana alam angin kencang berdasarkan dampaknya yaitu Pohon Tumbang, Akses Jalan dan Jaringan Listrik. Berikut analisis deskriptif menggunakan *barplot* untuk variabel Jumlah Kejadian.



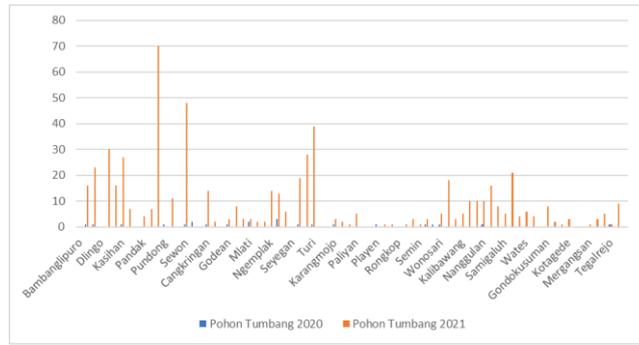
Gambar 2 Perbandingan Jumlah Kejadian Angin Kencang Tahun 2020 dan 2021

Berdasarkan **Gambar 2**, dapat dilihat bahwasannya untuk jumlah kejadian bencana angin kencang pada tahun 2021 lebih banyak terjadi daripada tahun 2020. Terdapat perbedaan signifikan dalam jumlah kejadian bencana angin kencang pada 2 tahun tersebut. Pada tahun 2020 jumlah kejadian bencana angin kencang terbanyak terjadi pada Kecamatan Cangkringan yaitu sebanyak 16 kejadian, sedangkan pada tahun 2021 kejadian bencana angin kencang terbanyak terjadi pada Kecamatan Turi yaitu sebanyak 35 kejadian. Hal ini menunjukkan bahwa jumlah kejadian bencana angin kencang pada tahun 2021 mengalami peningkatan dibandingkan tahun sebelumnya. Kenaikan ini dapat disebabkan oleh faktor cuaca yang kurang stabil pada tahun 2021. Selanjutnya, analisis deskriptif menggunakan *barplot* untuk variabel Jaringan Listrik sebagai berikut.



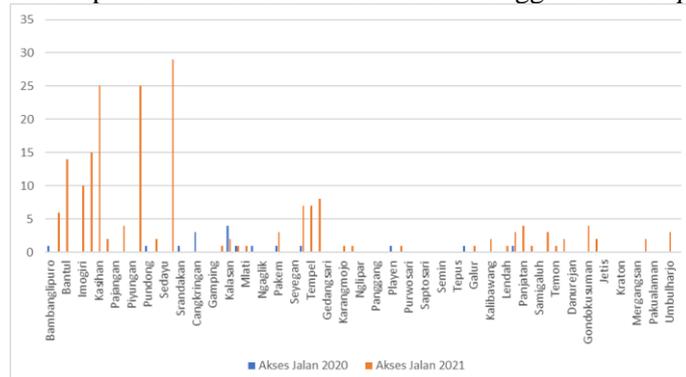
Gambar 3 Perbandingan Jaringan Listrik Tahun 2020 dan 2021

Pada **Gambar 3**, didapatkan informasi bahwasannya banyak dampak yang terjadi dari bencana angin kencang salah satunya untuk jaringan listrik. Pada tahun 2020 dampak dari bencana angin kencang untuk jaringan listrik terbanyak terjadi pada Kecamatan Cangkringan yaitu sebanyak 6 gangguan jaringan listrik, sedangkan pada tahun 2021 jumlah dampak kejadian untuk angin kencang terbanyak terdapat pada Kecamatan Kasihan yaitu melebihi dari 16 gangguan listrik. Hal ini menunjukkan bahwa dampak bencana angin kencang pada jaringan listrik pada tahun 2021 mengalami peningkatan yang cukup besar. Kenaikan ini dapat disebabkan oleh meningkatnya jumlah kejadian bencana angin kencang pada tahun 2021. Perlu dilakukan evaluasi terhadap sistem pemulihan jaringan listrik selama bencana dan dilakukan upaya perbaikan pada sistem agar dapat mengurangi dampak dari bencana angin kencang pada jaringan listrik di wilayah tersebut. Selanjutnya, analisis deskriptif menggunakan *barplot* untuk variabel Pohon Tumbang sebagai berikut.



Gambar 4 Perbandingan Pohon Tumbang tahun 2020 dan 2021

Berdasarkan grafik **Gambar 4**, terlihat bahwa pada tahun 2020 terdapat 3 pohon yang tumbang akibat bencana angin kencang, sedangkan pada tahun 2021 terdapat 70 pohon yang tumbang yang terjadi pada Kecamatan Pleret. Hal ini menunjukkan bahwa dampak bencana angin kencang pada pohon – pohon pada tahun 2021 meningkat cukup tinggi dibandingkan tahun sebelumnya. Kenaikan ini dapat disebabkan oleh meningkatnya kecepatan angin atau jumlah pohon yang lebih banyak yang terdampak pada tahun 2021. Diperlukan evaluasi terhadap kondisi pohon-pohon di wilayah tersebut dan tindakan perawatan atau penebangan pohon-pohon, terutama ranting-ranting yang tumbuh di tepi jalan, agar jalan tetap aman. Tujuannya adalah untuk meminimalisir bahaya saat terjadi hujan deras yang disertai angin agar para pengguna jalan dapat tetap aman. Berikut merupakan analisis deskriptif untuk variabel Akses Jalan menggunakan *barplot*.



Gambar 5 Perbandingan Akses Jalan tahun 2020 dan 2021

Berdasarkan informasi **Gambar 5**, terlihat bahwa pada tahun 2020 terdapat sedikit kasus kerusakan akses jalan akibat bencana angin kencang di wilayah - wilayah tersebut, sedangkan pada tahun 2021 terdapat hampir 30 kasus yang terletak pada Kecamatan Sewon. Hal ini menunjukkan bahwa dampak bencana angin kencang pada akses jalan mengalami peningkatan yang signifikan pada tahun 2021 dibandingkan tahun sebelumnya. Peningkatan ini dapat berdampak pada mobilitas masyarakat dan aksesibilitas terhadap wilayah tersebut. Perlu dilakukan upaya-upaya perbaikan dan pemeliharaan infrastruktur jalan serta perencanaan tata ruang yang lebih baik untuk mengurangi risiko dampak bencana pada akses jalan di wilayah tersebut.

Karakteristik data yang digunakan dalam penelitian ini dapat dilihat melalui statistika deskriptif yang telah disajikan dalam **Tabel 1** sebagai berikut.

Tabel 1 Statistika Deskriptif Data Kejadian Bencana Angin Kencang di Provinsi Daerah Istimewa Yogyakarta

		Jumlah Kejadian	Pohon Tumbang	Jaringan Listrik	Akses Jalan
2020	Mean	1.615	0.295	0.282	0.218
	Min	0.000	0.000	0.000	0.000
	Max	16.00	3.000	6.000	4.000

		Jumlah Kejadian	Pohon Tumbang	Jaringan Listrik	Akses Jalan
2021	Mean	2.487	7.615	1.423	2.487
	Min	0.000	0.000	0.000	0.000
	Max	35.00	70.00	17.00	29.00

Tabel 1 tersebut menunjukkan bahwa rentang data jumlah kejadian bencana angin kencang di Provinsi DI Yogyakarta pada tahun 2020 adalah antara 0 dan 16 dengan nilai rata-rata sebesar 1.615, sedangkan pada tahun 2021 rentang data tersebut adalah antara 0 dan 35 dengan nilai rata-rata sebesar 2.487. Nilai rentang tersebut menunjukkan variasi data yang cukup signifikan pada kedua tahun tersebut, sedangkan nilai rata-rata menunjukkan kecenderungan kejadian bencana angin kencang yang sedikit lebih tinggi pada tahun 2021 dibandingkan dengan tahun 2020.

Hal tersebut menunjukkan bahwa perlu memperhatikan risiko kejadian bencana angin kencang yang cukup signifikan di daerah tersebut dan terus melakukan evaluasi dan mitigasi risiko bencana. Selain itu, data statistika deskriptif tersebut dapat digunakan sebagai acuan dalam pengambilan kebijakan dan perencanaan program penanggulangan bencana angin kencang di masa yang akan datang. Pada tabel tersebut juga terdapat dampak – dampak dari angin kencang yang meliputi Pohon Tumbang, gangguan Jaringan Listrik dan kerusakan akses jalan. Pada dampak pohon tumbang, di tahun 2020 terdapat rata – rata 0.295 pohon yang tumbang akibat bencana angin kencang, dengan jumlah minimum 0 dan maksimum 3. Sedangkan ditahun 2021, pohon yang tumbang akibat angin kencang cukup tinggi meningkat menjadi 7.615, dengan jumlah minimum 0 dan maksimum 70 pohon yang tumbang.

Pada dampak gangguan jaringan listrik yang disebabkan angin kencang, pada tahun 2020 terdapat rata – rata 0.282 gangguan, dengan jumlah minimum 0 dan maksimum 6. Sedangkan pada tahun 2021, rata – rata 1.423 gangguan yang sedikit meningkat, dengan jumlah minimum 0 dan maksimum 17 gangguan. Dampak terakhir adalah kerusakan akses jalan, pada tahun 2020 terdapat rata – rata 0.218 kerusakan akses jalan, dengan jumlah minimum 0 dan maksimum 4 kerusakan. Sedangkan untuk tahun 2021, rata – rata kerusakan 2.487 dengan jumlah minimum 0 dan maksimum 29 kerusakan akses jalan. Dari hasil interpretasi tersebut, dapat disimpulkan bahwa dampak – dampak angin kencang di Provinsi Daerah Istimewa Yogyakarta cenderung mengalami peningkatan dari tahun 2020 ke tahun 2021. Pada dampak gangguan jaringan listrik dan kerusakan akses jalan peningkatannya tidak terlalu signifikan, sedangkan pada dampak pohon tumbang peningkatannya cukup lebih signifikan.

3.2. Uji Asumsi Analisis *Cluster*

3.2.1. Uji Multikolinearitas

Uji Multikolinearitas pada *Clustering* bertujuan untuk memastikan bahwa variabel – variabel yang digunakan dalam analisis *clustering* tidak berkorelasi terlalu tinggi. Jika terjadi multikolinearitas, artinya ada korelasi tinggi antara dua atau lebih, yang dapat mempengaruhi hasil analisis *clustering* dan menurunkan validitas hasil. Menurut Widarjono (2010), terjadinya multikolinearitas apabila nilai VIF atau *Variance Inflation Factor* (VIF_j) ≥ 10 [7]. Berdasarkan hasil perhitungan uji multikolinearitas terhadap data kejadian bencana angin kencang di Provinsi Daerah Istimewa Yogyakarta didapatkan *output* pada **Tabel 2** dan **Tabel 3**.

Tabel 2 Nilai VIF Antar Variabel Pada Tahun 2020

Nama Variabel	Tahun	Nilai VIF
Jumlah Kejadian		1.406807
Pohon Tumbang	2020	1.633455
Jaringan Listrik		1.404844
Akses Jalan		1.451984

Berdasarkan **Tabel 2**, didapatkan informasi bahwa tidak ada indikat kuat adanya multikolinearitas antar variabel, karena nilai VIF yang di peroleh kurang dari 10 dan menunjukkan bahwa variasi dalam setiap variabel hampir tidak dijelaskan oleh variabel lain. Oleh karena itu, untuk Uji multikolinearitas pada data tahun 2020 dapat disimpulkan bahwa variabel Jumlah Kejadian, Pohon Tumbang, Jaringan Listrik dan Akses Jalan relatif bebas dari multikolinearitas.

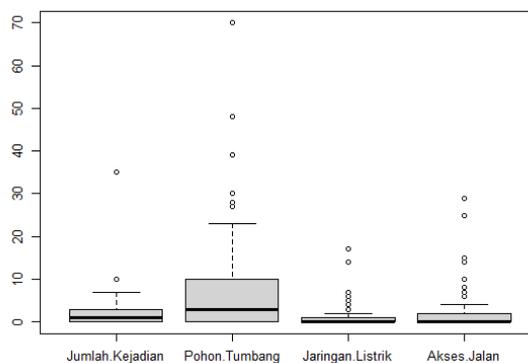
Tabel 3 Nilai VIF Antar Variabel Pada Tahun 2021

Nama Variabel	Tahun	Nilai VIF
Jumlah Kejadian	2021	3.700877
Pohon Tumbang		3.454601
Jaringan Listrik		3.941776
Akses Jalan		3.150141

Berdasarkan **Tabel 3**, didapatkan informasi bahwa tidak ada indikat kuat adanya multikolinearitas antar variabel, karena nilai VIF yang di peroleh kurang dari 10. Oleh karena itu, untuk Uji multikolinearitas dapat disimpulkan bahwa variabel Jumlah Kejadian, Pohon Tumbang, Jaringan Listrik dan Akses Jalan pada tahun 2021 relatif bebas dari multikolinearitas.

3.2.2. Uji Deteksi *Outlier*

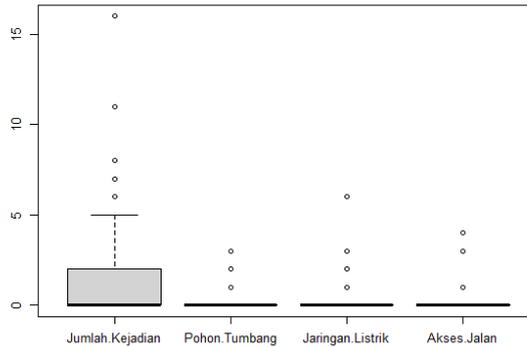
Menurut Triola (2018), kita dapat mengidentifikasi outlier dengan berbagai cara. Salah satunya adalah menggunakan metode grafis, seperti box plot, di mana kita melihat apakah suatu observasi data terletak jauh dari batas-batas yang ditentukan oleh whisker data. Selain itu, kita juga dapat memperhatikan berbagai ukuran pusat dan variasi data, seperti rata-rata dan median, untuk menentukan apakah nilai suatu observasi data terlalu ekstrem dibandingkan dengan data lainnya. Dengan demikian, terdapat beberapa pendekatan sederhana yang dapat digunakan untuk mengidentifikasi outlier dalam kumpulan data [8]. Berikut hasil dari grafik *boxplot* dari data kejadian bencana angin kencang di Povinsi Daerah Istimewa Yogyakarta dan didapatkan *output* pada **Gambar 5** dan **Gambar 6**.



Gambar 6 Grafik *Boxplot* Deteksi *Outlier* Data Tahun 2020

Secara grafis *boxplot* pada **Gambar 6** mewakili distribusi variabel kuantitatif dengan secara visual menampilkan lima ringkasan lokasi umum (maksimum, minimum, median, kuartil tiga dan kuartil pertama) dan setiap pengamatan yang diklasifikasikan dengan *outlier* yang dicurigai menggunakan kriteria rentang interkuartil (IQR). Berdasarkan **Gambar 6**, masih ada beberapa data *outlier* di setiap variabelnya pada data kejadian angin kencang tahun 2020. Pada variabel jumlah kejadian terdapat 2 data *outlier* yang mana melebihi dari nilai kuartil 3 dari data tersebut karena berada di atas *whisker boxplot*, lalu pada variabel pohon tumbang terdapat 6 data *outlier* yang melebihi nilai kuartil 3 karena berada di atas *whisker boxplot*, variabel ketiga yakni jaringan listrik diketahui terdapat 7 data *outlier* yang melebihi nilai kuartil 3 karena juga berada di atas *whisker boxplot*, dan untuk variabel terakhir yaitu akses jalan diketahui terdapat 8 data *outlier* yang melebihi

nilai kuartil 3 karena juga berada di atas *whisker boxplot*. *Outlier – outlier* ini kemungkinan merupakan data yang sangat rendah atau di luar nilai yang diharapkan dalam suatu variabel.



Gambar 7 Grafik *Boxplot* Deteksi *Outlier* Data Tahun 2021

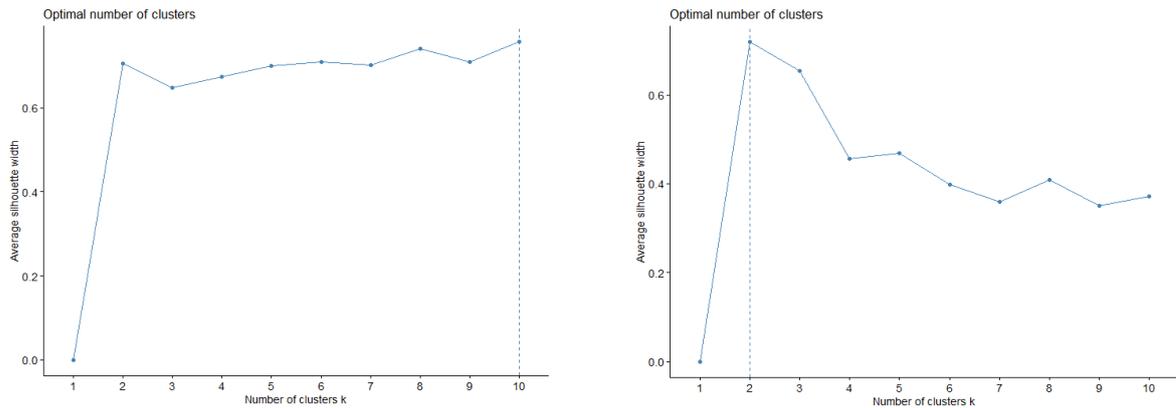
Secara grafis *boxplot* pada **Gambar 7** mewakili distribusi variabel kuantitatif dengan secara visual menampilkan lima ringkasan lokasi umum (minimum, maksimum, median, kuartil pertama dan kuartil ketiga) dan setiap pengamatan yang diklasifikasikan dengan *outlier* yang dicurigai menggunakan kriteria rentang interkuartil (IQR). Berdasarkan **Gambar 7**, masih ada beberapa data *outlier* di setiap variabelnya pada data kejadian angin kencang tahun 2021. Pada variabel jumlah kejadian terdapat 5 data *outlier* yang mana melebihi dari nilai kuartil 3 dari data tersebut karena berada di atas *whisker boxplot*, lalu pada variabel pohon tumbang terdapat 3 data *outlier* yang melebihi nilai kuartil 3 karena berada di atas *whisker boxplot*, variabel ketiga yakni jaringan listrik diketahui terdapat 4 data *outlier* yang melebihi nilai kuartil 3 karena juga berada di atas *whisker boxplot*, dan untuk variabel terakhir yaitu akses jalan diketahui terdapat 3 data *outlier* yang melebihi nilai kuartil 3 karena juga berada di atas *whisker boxplot*. *Outlier-outlier* ini kemungkinan merupakan data yang sangat rendah atau di luar nilai yang diharapkan dalam suatu variabel. *Boxplot* yang condong ke arah bawah dapat menunjukkan bahwa data cenderung memiliki nilai yang lebih tinggi, namun juga memiliki *outlier* yang signifikan pada nilai yang lebih rendah. Namun, pada penelitian ini peneliti tetap menggunakan data *outlier* dikarenakan apabila data ditiadakan maka akan mempengaruhi hasil akhir dalam pengelompokan sehingga menyebabkan Kecamatan tersebut tidak menjadi prioritas Badan Penanggulangan Bencana Daerah Daerah Istimewa Yogyakarta untuk memaksimalkan penanggulangan daerah rawan bencana angin kencang.

3.3. Analisis *K – Means Clustering*

K-means merupakan metode *clustering* non-hierarki yang bertujuan untuk mempartisi data ke dalam satu atau lebih cluster berdasarkan karakteristik yang dimilikinya. Dalam proses ini, data yang memiliki kesamaan karakteristik dikelompokkan ke dalam satu cluster, sedangkan data yang memiliki perbedaan karakteristik ditempatkan dalam kelompok yang berbeda. Dengan demikian, *K-means* bertujuan untuk mengelompokkan data ke dalam kelompok yang saling serupa, sehingga mempermudah analisis dan pemahaman terhadap struktur data yang ada [9].

K-means merupakan salah satu metode *clustering* yang berbasis jarak, dimana tujuannya adalah untuk membagi data ke dalam sejumlah cluster. Metode ini hanya dapat digunakan pada atribut numerik karena algoritma *K-means* mengukur jarak antara titik data menggunakan metrik jarak, seperti Euclidean distance. Dalam prosesnya, algoritma *K-means* menggunakan pendekatan partisi *clustering*, yang berarti memisahkan data menjadi *k* kluster yang berbeda. Keunggulan utama dari algoritma *K-means* adalah kemampuannya untuk mengelompokkan data secara efisien, terutama dalam menangani data yang besar, serta mampu mengidentifikasi data *outlier* dengan cepat. Karena alasan ini, *K-means* menjadi salah satu algoritma *clustering* yang sangat terkenal dan sering digunakan dalam berbagai aplikasi analisis data [10].

3.3.1. Penentuan Jumlah Cluster



Gambar 8 (a) Penentuan Jumlah Cluster Tahun 2020

(b) Penentuan Jumlah Cluster Tahun 2021

Berdasarkan **Gambar 8(a)** dan **8(b)** dapat dilihat bahwasannya nilai rata-rata *silhouette* tertinggi berada k bernilai 10 untuk data kejadian angin kencang 2020 dan k bernilai 2 untuk data kejadian angin kencang 2021. Namun, dalam menentukan banyaknya *cluster* yang terbentuk sifatnya subjektif atau tergantung kepada peneliti. Maka, banyaknya *cluster* atau *k* yang digunakan dalam penelitian ini yaitu $k = 3$, dikarenakan tingkatan resiko bencana di Indonesia terbagi menjadi 3 yaitu rendah, sedang, dan tinggi.

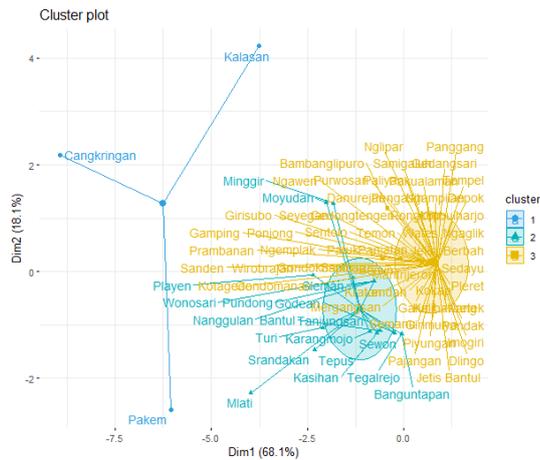
3.3.2. Analisis K - means Clustering 2020

Untuk menentukan kelompok dalam *Clustering* Non Hirarki yang akan dikelompokkan ke dalam suatu kelompok dengan karakteristik yang semirip mungkin. Dengan menggunakan k optimal sebesar 3, maka didapatkan hasil pengelompokkan dari standarisasi data dan dihasilkan titik pusat *cluster* (*centroid*) seperti pada **Tabel 4**.

Tabel 4 Titik Pusat Cluster (*Centroid* s) Tahun 2020

Cluster	Banyak Kejadian	Pohon Tumbang	Jaringan Listrik	Akses Jalan
1	3.1756170	0.3515373	3.7803093	5.1481849
2	1.7374790	1.6368455	1.2839497	0.8345690
3	-0.2978098	-0.2029094	-0.2622385	-0.2496018

Tabel 4 merupakan hasil titik pusat *cluster* pada setiap *cluster* untuk masing-masing variabel berdasarkan kesamaan terdekat dalam data hasil standarisasi menggunakan Z-Score. Berdasarkan **Tabel 4**, dapat dilihat bahwasannya terdapat tiga *cluster* yang terbentuk. Setiap *cluster* memiliki titik pusat *cluster* (*centroid*) yang berbeda – beda pada setiap variabelnya. Misalnya, pada *cluster* 1 memiliki titik pusat *cluster* sebesar 3.1756170 pada variabel Banyak Kejadian, 0.3515373 pada variabel Pohon Tumbang. Sedangkan untuk *cluster* 2 memiliki titik pusat *cluster* sebesar 1.7374790 pada variabel Banyak Kejadian, 1.6368455 pada variabel Pohon Tumbang. Hal ini menunjukkan bahwa setiap *cluster* memiliki karakteristik atau pola yang berbeda – beda dalam nilai rata – rata variabel. Oleh karena itu, nilai *centroid cluster* dapat digunakan untuk mengidentifikasi *cluster* mana yang memiliki karakteristik yang paling cocok dengan tujuan analisis yang dilakukan. Selanjutnya, hasil visualisasi pengelompokan ke dalam 3 kelompok ditampilkan pada



Gambar 9 Output Plot *K-means Clustering* Tahun 2020

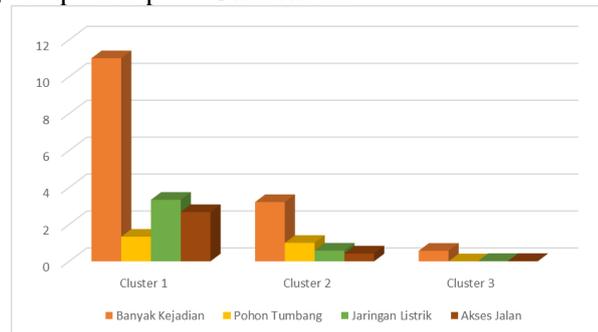
Berdasarkan **Gambar 10**, diatas, terdapat tiga *cluster* yang dihasilkan dari pengelompokan data bencana angin kencang berdasarkan dampak kejadian di Provinsi Daerah Istimewa Yogyakarta. *Cluster 1* yang ditandai berwarna biru muda, *cluster 2* dengan warna hijau muda, dan *cluster 3* dengan warna kuning yang mana memiliki anggota *cluster* paling banyak. Di setiap pengelompokannya akan ditarik berdasarkan titik pusat *centroid* nya dan akan menjadi satu kelompok ketika memiliki jarak ke *centroid* saling mendekati. Untuk lebih jelasnya maka disajikan tabel yang akan memuat seluruh anggota di kelompok *cluster* masing-masing dan ditampilkan pada **Tabel 5**.

Tabel 5 Hasil *Cluster K-means* Tahun 2020

<i>Cluster</i>	Jumlah	Anggota <i>Cluster</i>
1	3	“Cangkringan”, “Kalasan” dan “Pakem”.
2	19	“Banguntapan”, “Bantul”, “Kasihan”, “Pundong”, “Sewon”, “Sradakan”, “Godean”, “Minggir”, “Mlati”, “Moyudan”, “Sleman”, “Turi”, “Karangmojo”, “Playen”, “Tanjungsari”, “Tepus”, “Wonosari”, “Nanggulan”, dan “Tegalrejo”.
3	56	“Bambanglipuro”, “Dlingo”, “Imogiri”, “Jetis Bantul”, “Kretek”, “Pajangan”, “Pandak”, “Piyungan”, “Pleret”, “Sanden”, “Sedayu”, “Berbah”, “Depok”, “Gamping”, “Nganglik”, “Ngeplak”, “Prambanan”, “Seyegan”, “Tempel”, “Gedangsari”, “Girisubo”, “Ngawen”, “Nglipar”, “Paliyan”, “Panggang”, “Patuk”, “Ponjong”, “Purwosari”, “Rongkop”, “Saptosari”, “Semanu”, “Semin”, “Galur”, “Girimulyo”, “Kalibawang”, “Kokap”, “Lendah”, “Panjatan”, “Pengasih”, “Samigaluh”, “Sentolo”, “Temon”, “Wates”, “Danurejan”, “Gedongtengen”, “Gondokusuman”, “Gondomanan”, “Jetis”, “Kotagede”, “Kraton”, “Matrijeron”, “Mergangsan”, “Ngampilan”, “Pakualaman”, “Umbulharjo”, dan “Wirobrajan”.

Berdasarkan **Tabel 5**, didapatkan informasi bahwa data bencana angin kencang berdasarkan dampak kejadian di Provinsi Daerah Istimewa Yogyakarta pada tahun 2020 terdapat 3 *cluster* yang telah dibentuk. Pada *cluster 1* terdapat 3 kecamatan di Daerah Istimewa Yogyakarta, *cluster 2* terdapat 19 kecamatan di Daerah Istimewa Yogyakarta dan *cluster 3* terdapat 56 kecamatan di Daerah Istimewa Yogyakarta. Sehingga didapatkan total keseluruhan kecamatan di Daerah Istimewa Yogyakarta terdapat 78 kecamatan. Dari hasil pengelompokan ini, dapat diketahui bahwa data bencana alam dapat dikelompokkan ke dalam tiga *cluster* berdasarkan dampak kejadian dari bencana angin kencang. *Cluster-cluster* tersebut dapat menjadi acuan bagi pihak yang terkait untuk menentukan strategi penanganan bencana alam yang lebih tepat dan efektif. Selanjutnya peneliti melakukan

profilisasi data agar bisa mengathui karakteristik masing-masing *cluster* menggunakan rata-rata nilai yang didapatkan pada **Gambar 11**.



Gambar 10 Profilisasi Cluster K-means Tahun 2020

Berdasarkan **Gambar 11**, terlihat hasil profilisasi yang menggambarkan kejadian bencana angin kencang di setiap kecamatan di Yogyakarta. Terdapat perbedaan yang signifikan antara cluster-cluster yang terbentuk.

Cluster 1 mencakup kecamatan-kecamatan yang memiliki karakteristik indikator rawan bencana angin kencang yang tinggi dibandingkan dengan cluster lainnya. Hal ini terlihat dari variabel-variabel seperti jumlah kejadian, pohon tumbang, jaringan listrik, dan akses jalan yang memiliki nilai tertinggi dibandingkan dengan cluster lainnya. *Cluster 2* mencakup kecamatan-kecamatan yang memiliki karakteristik indikator rawan bencana angin kencang yang sedang dibandingkan dengan cluster lainnya. Variabel-variabel seperti jumlah kejadian, pohon tumbang, jaringan listrik, dan akses jalan memiliki nilai yang lebih rendah dibandingkan dengan cluster 1, namun masih di atas cluster 3. *Cluster 3* mencakup kecamatan-kecamatan yang memiliki karakteristik indikator rawan bencana angin kencang yang rendah dibandingkan dengan cluster lainnya. Variabel-variabel seperti jumlah kejadian, pohon tumbang, jaringan listrik, dan akses jalan memiliki nilai terendah dibandingkan dengan cluster-cluster lainnya.

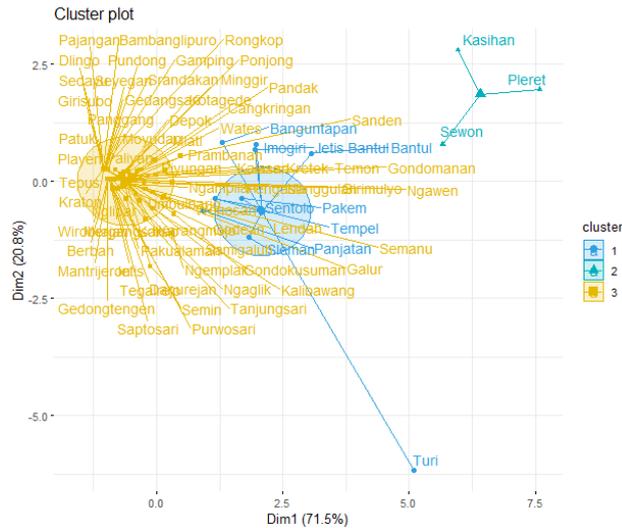
3.3.3. Analisis K – Means Clustering 2021

Dengan menggunakan k optimal sebesar 3, maka didapatkan hasil pengelompokan dari standarisasi data dan dihasilkan titik pusat *cluster* (*centroid*) seperti pada **Tabel 7**.

Tabel 6 Titik Pusat Cluster (*Centroids*) Tahun 2021

Cluster	Banyak Kejadian	Pohon Tumbang	Jaringan Listrik	Akses Jalan
1	0.419754	3.446981	3.8293666	4.2279137
2	1.253432	1.226196	0.9395673	0.9242310
3	-0.212209	-0.347737	-0.3212888	-0.3373239

Berdasarkan **Tabel 7**, dapat dilihat bahwasannya terdapat tiga *cluster* yang terbentuk. Setiap *cluster* memiliki titik pusat *cluster* (*centroid*) yang berbeda – beda pada setiap variabelnya. Misalnya, pada *cluster 1* memiliki titik pusat *cluster* sebesar 0.419754 pada variabel Banyak Kejadian, 3.446981 pada variabel Pohon Tumbang. Sedangkan untuk *cluster 2* memiliki titik pusat *cluster* sebesar 1.253432 pada variabel Banyak Kejadian, 1.226196 pada variabel Pohon Tumbang. Hal ini menunjukkan bahwa setiap *cluster* memiliki karakteristik atau pola yang berbeda – beda dalam nilai rata – rata variabel. Oleh karena itu, nilai *centroid cluster* dapat digunakan untuk mengidentifikasi *cluster* mana yang memiliki karakteristik yang paling cocok dengan tujuan analisis yang dilakukan. Selanjutnya, hasil visualisasi pengelompokan ke dalam 3 kelompok ditampilkan pada **Gambar 11**.



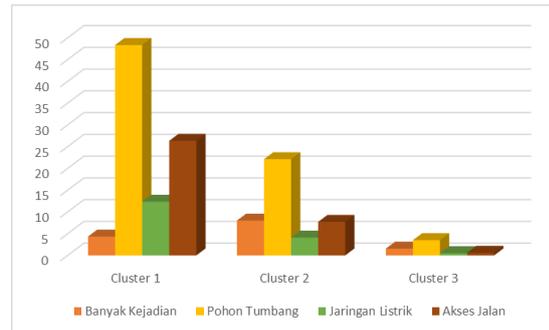
Gambar 11 Output Plot K – Means Clustering Tahun 2021

Berdasarkan **Gambar 11**, terdapat tiga *cluster* yang dihasilkan dari pengelompokan data bencana angin kencang berdasarkan dampak kejadian di Provinsi Daerah Istimewa Yogyakarta. *Cluster 1* yang ditandai berwarna biru muda, *cluster 2* ditandai dengan warna hijau muda, dan *cluster 3* ditandai dengan warna kuning yang mana memiliki anggota *cluster* paling banyak. Di setiap pengelompokannya akan ditarik berdasarkan titik pusat *centroid* nya dan akan menjadi satu kelompok ketika memiliki jarak ke *centroid* saling mendekati. Untuk lebih jelasnya maka disajikan tabel yang akan memuat seluruh anggota di kelompok *cluster* masing-masing dan ditampilkan pada **Tabel 8**.

Tabel 7 Hasil *Cluster K – Means* Tahun 2021

Cluster	Jumlah	Anggota Cluster
1	3	“Kasihan”, “pleret”, dan “sewon”.
2	10	“Banguntapan”, “Bantul”, “Imogiri”, “Jetis Bantul”, “Pakem”, “Sleman”, “Tempel”, “Turi”, “Panjatan”, Dan “Sentolo”.
3	65	“Bambanglipuro”, “Dlingo”, “Kretek”, “Pajangan”, “Pandak”, “Piyungan”, “Pundong”, “Sanden”, “Sedayu”, “Srandakan”, “Berbah”, “Cangkringan”, “Depok”, “Gamping”, “Godean”, “Kalasan”, “Minggir”, “Mlati”, “Moyudan”, “Nganglik”, “Ngemplak”, “Prambanan”, “Seyegan”, “Gedangsari”, “Girisubo”, “Karangmojo”, “Ngawen”, “Nglipar”, “Paliyan”, “Panggang”, “Patuk”, “Playen”, “Ponjong”, “Purwosari”, “Rongkop”, “Saptosari”, “Semanu”, “Semin”, “Tanjungsari”, “Tepus”, “Wonosari”, “Galur”, “Girimulyo”, “Kalibawang”, “Kokap”, “Lendah”, “Nanggulan”, “Pengasih”, “Samigaluh”, “Temon”, “Wates”, “Danurejan”, “Gedongtengen”, “Gondokusuman”, “Gondomanan”, “Jetis”, “Kotagede”, “Kraton”, “Matrijeron”, “Mergangsan”, “Ngampilan”, “Pakualaman”, “Tegalrejo”, “Umbulharjo”, dan “Wirobrajan”.

Berdasarkan **Tabel 8** terlihat bahwasanya terdapat 3 *cluster* yang sebelumnya telah terbentuk dari data jumlah kejadian angin kencang yang terjadi di Daerah Istimewa Yogyakarta pada tahun 2020 berdasarkan dampak kejadian. Pada *cluster 1* terdapat 3 kecamatan di Yogyakarta, *Cluster 2* terdapat 19 kecamatan di Daerah Istimewa Yogyakarta, dan untuk *Cluster 3* terdapat 56 kecamatan di Daerah Istimewa Yogyakarta. Setelah mendapatkan jumlah kecamatan di setiap *cluster*nya maka peneliti melakukan profilisasi data agar bisa mengathui karakteristik masing-masing *cluster* menggunakan rata-rata nilai yang didapatkan pada **Gambar 13**.



Gambar 12 Profilisasi Cluster K-means Tahun 2021

Berdasarkan **Gambar 13**, terlihat hasil profilisasi yang menggambarkan kejadian bencana angin kencang di setiap kecamatan di Yogyakarta. Terdapat perbedaan yang signifikan antara cluster-cluster yang terbentuk.

Cluster 1 memiliki karakteristik indikator rawan bencana angin kencang yang tinggi dibandingkan dengan *cluster* lainnya. Hal ini dapat dilihat dari jumlah dampak yang paling signifikan dibandingkan dengan *cluster* lainnya. Meskipun jumlah kejadian bencana angin kencang pada kecamatan yang masuk dalam *cluster* ini bukan yang tertinggi, namun dampak yang dihasilkan dari kejadian tersebut memiliki nilai tertinggi dibandingkan dengan *cluster* lainnya. Tetapi jumlah kejadian pada kecamatan dalam *cluster 1* berada pada urutan kedua setelah jumlah kejadian pada kecamatan dalam *cluster 2*. *Cluster 2* memiliki karakteristik indikator rawan bencana angin kencang yang sedang dibandingkan dengan *cluster* lainnya. Hal ini dapat dilihat dari jumlah dampak yang cukup signifikan dibandingkan dengan *cluster* lainnya. Meskipun jumlah kejadian bencana angin kencang pada kecamatan yang masuk dalam *cluster* ini bukan yang tertinggi, namun dampak yang dihasilkan dari kejadian tersebut memiliki nilai tertinggi kedua setelah *cluster 1*. Perlu dicatat bahwa jumlah kejadian pada kecamatan dalam *cluster 2* memiliki jumlah tertinggi dibandingkan dengan kecamatan dalam *cluster* lainnya. *Cluster 3* mencakup kecamatan-kecamatan yang memiliki karakteristik indikator rawan bencana angin kencang yang rendah dibandingkan dengan *cluster* lainnya. Variabel-variabel seperti jumlah kejadian, pohon tumbang, jaringan listrik, dan akses jalan memiliki nilai terendah dibandingkan dengan *cluster-cluster* lainnya.

4. Kesimpulan

Data kejadian angin kencang yang terjadi di Daerah Istimewa Yogyakarta pada tahun 2020 dikelompokkan menjadi 3 *cluster* dengan menggunakan algoritma *K-means Clustering* yang mana *cluster 1* yang berjumlah 3 kecamatan dengan indikator rawan bencana angin kencang tinggi, *cluster 2* yang berjumlah 19 kecamatan dengan indikator rawan bencana angin kencang sedang, lalu terakhir *cluster 3* yang berjumlah 56 kecamatan dengan indikator rawan bencana angin kencang rendah. Pada tahun 2021 juga terbentuk menjadi 3 *cluster* yang berbeda yang mana *cluster 1* berjumlah 3 kecamatan dengan indikator tawan bencana angin kencang tinggi, *cluster 2* berjumlah 10 kecamatan dengan indikator rawan bencana angin kencang sedang, lalu terakhir *cluster 3* berjumlah 65 kecamatan dengan indikator rawan bencana angin kencang tinggi.

5. Daftar Pustaka

- [1] N. N. Sari, *K-Affinity Propagation (K-AP) Clustering* untuk Klasifikasi Gempa Bumi, Yogyakarta: Universitas Islam Indonesia, 2018.
- [2] P. Indonesia, "Undang Undang Republik Indonesia," *Undang Undang Tentang Penanggulangan Bencana Nomor 24 Tahun 2007*, 2007.

- [3] F. Asfahanif and S. Amri, "Analisis Peristiwa Angin Kencang dengan Citra Satelit Himawari-8 (Studi Kasus : Bangkalan, 17 Oktober 2021)," *Journal Geografi*, pp. 84-90, 2022.
- [4] A. Kusuma and M. S. Roychansyah, "Strategi Permukiman Tangguh Bencana berdasarkan Studi Morfologi Kampung Pengok Kidul, Yogyakarta," *Jurnal Lingkungan Binaan Indonesia*, pp. 47-55, 2022.
- [5] A. Bahauddin, A. Fatmawati and F. P. Sari, "Analisis *Clustering* Provinsi di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma *K-Means*," *Jurnal Manajemen Informatika & Sistem Informasi*, vol. 4, 2021.
- [6] N. Wahidah, O. Juwita and F. N. Arifin, "Pengelompokan Daerah Rawan Bencana di Kabupaten Jember Menggunakan Metode *K-means Clustering*," *Informatics Journal*, vol. 8, pp. 22 - 29, 2023.
- [7] D. R. Ningrat, D. A. I. Maruddani and T. Wuryandari, "Analisis Cluster dengan Algoritma *K-Means* dan Fuzzy *C-Means Clustering* untuk Pengelompokan Data Obligasi Korporasi," *Jurnal Gaussian*, pp. 642-650, 2016.
- [8] P. R. Sihombing, S. D. A. Sunarjo and Y. P. A. C. Yuda, "Identifikasi Data Outlier (Pencilan) dan Kenormalan Data Pada Data Univariat serta Alternatif Penyelesaiannya," *Jurnal Ekonomi dan Statistik Indonesia*, vol. 2, 2022.
- [9] A. N. Khomarudin, "Teknik Data Mining : Algoritma *K-means Clustering*," *Ilmu Komputer*, 2018.
- [10] Y. Darmi and A. Setiawan, "Penerapan Metode *Clustering K-Means* dalam Pengelompokan Penjualan Produk," *Jurnal Media Infotama*, vol. 12, 2016.