

PENGELOMPOKAN PROVINSI DI INDONESIA BERDASARKAN DATA JUMLAH KEJADIAN DAN DAMPAK BENCANA BANJIR MENGUNAKAN METODE *FUZZY C-MEANS*

Memi Nor Hayati^{1*}, Rito Goejantoro², Meiliyani Siringoringo³, Ika Purnamasari³, Desi

Yuniarti³, Khairun Nida¹, Gerald Claudio Messakh¹

¹Laboratorium Statistika Terapan, FMIPA, Universitas Mulawarman

²Laboratorium Statistika Komputasi, FMIPA, Universitas Mulawarman

³Laboratorium Statistika Ekonomi dan Bisnis, FMIPA, Universitas Mulawarman

Keywords: Cluster Analysis, Flood Disaster, FCM, Modified Partition Coefficient

Abstract:

Cluster analysis is a technique used to find groups of similar data objects. The Fuzzy C-Means (FCM) method is a data grouping method where the existence of each data in a cluster is determined by the degree of membership. This study aims to determine the optimal number of clusters based on the Modified Partition Coefficient (MPC) validity index and to determine the optimal grouping results of 34 provinces in Indonesia based on data on the number of events and the impact of floods in 2017-2021. The optimal number of clusters using the FCM method is based on MPC value consists of 2 clusters, namely the first cluster consisting of 27 provinces in Indonesia and the second cluster consisting of 7 provinces in Indonesia.

1. Pendahuluan

Indonesia memiliki curah hujan yang tinggi, berkisar antara 2.000 hingga 3.000 mm/tahun. Hal ini memudahkan terjadinya banjir selama musim hujan yang berlangsung antara bulan Oktober hingga Januari. Banjir adalah peristiwa tergenangnya tanah yang diakibatkan oleh luapan air sungai (Findayani, 2015). Umumnya banjir yang hanya berupa genangan air, terlebih banjir bandang bersifat merusak dan dapat menimbulkan kerusakan yang cukup berarti (Bakornas PB, 2007).

Menanggapi berbagai kejadian bencana di Indonesia, pemerintah membentuk Badan Nasional Penanggulangan Bencana (BNPB). BNPB mencatat sejak tahun 2017 hingga 2021, jumlah bencana alam di Indonesia mencapai angka 18.658 kejadian. Banjir termasuk kategori bencana yang besar dengan persentase total kejadian 28% sejak tahun 2017 hingga 2021. Jumlah kejadian banjir dari tahun 2017 hingga 2021 berturut-turut adalah 992, 883, 796, 1.522, dan 1.175 kejadian. Salah satu provinsi yang terdampak bencana banjir yaitu Provinsi Aceh Utara (Zalmita dkk., 2021).

Bencana banjir diukur dengan melihat kemungkinan terjadinya kerusakan yang sering disebut sebagai risiko banjir atau dampak yang ditimbulkan bagi masyarakat setelah kejadian seperti korban jiwa maupun kerusakan material (Ramadhan & Handinata, 2019). Dalam kaitan ini, dampak bencana banjir dapat menimbulkan kerugian yang berpotensi mengganggu bahkan melumpuhkan kegiatan sosial-ekonomi masyarakat (Bakornas PB, 2007). Terjadinya

* Corresponding author.

E-mail address: memminorhayati@fmipa.unmul.ac.id



bencana banjir tidak dapat terus dihindari tetapi dapat dilakukan berbagai upaya pencegahan dan penanggulangan bencana. Meskipun upaya pencegahan dan penanggulangan bencana ini tidak dapat dilakukann secara seragam di setiap daerah karena kondisi geografis setiap wilayah yang berbeda. Oleh karena itu, perlu dilakukan pengelompokan provinsi di Indonesia berdasarkan dampak bencana banjir sebagai upaya mitigasi terhadap bencana banjir (Islamy dkk., 2022).

Salah satu metode statistik yang dapat digunakan untuk pengelompokan adalah dengan analisis *cluster*. *Clustering* adalah teknik yang digunakan untuk menemukan kelompok objek data yang sama dengan cara mengelompokkan beberapa objek menjadi beberapa *cluster* dimana *cluster* terbaik adalah *cluster* yang memiliki tingkat homogenitas yang tinggi antar objek di dalam suatu *cluster* dan tingkat homogenitas yang rendah dengan objek *cluster* lainnya (Silitonga & Morina, 2017). *Fuzzy clustering* merupakan salah satu teknik dalam analisis *cluster* yang digunakan untuk menentukan *cluster* optimal dalam suatu ruang vektor berdasarkan pada bentuk normal *euclidian* untuk jarak antar vektor. *Fuzzy clustering* sangat cocok digunakan dalam pemodelan *fuzzy*, terutama untuk menentukan aturan-aturan *fuzzy* (Sari & Suranti, 2016).

Menurut Kusumadewi & Purnomo (2010), terdapat beberapa algoritma *fuzzy clustering*, salah satunya adalah *Fuzzy C-Means* (FCM). FCM adalah suatu teknik pengelompokan data dimana keberadaan setiap titik data dalam suatu *cluster* ditentukan oleh derajat keanggotaannya. Konsep dasar dari FCM yaitu menentukan pusat *cluster* terlebih dahulu, di mana pusat *cluster* ini akan mengelompokkan setiap titik dalam suatu kelompok. Penentuan pusat *cluster* dan nilai keanggotaan akan dilakukan secara berulang sehingga didapat pusat klaster yang nantinya akan berpindah ke posisi yang tepat dikarenakan kondisi pada pusat klaster dan titik setiap datanya berdasarkan derajat keanggotaan belum terlalu akurat (Setiawan, dkk., 2018).

Penerapan metode FCM telah dibahas pada penelitian sebelumnya yang dilakukan Messakh dkk., (2023) yang mengelompokkan kabupaten/kota di Pulau Kalimantan berdasarkan indikator pendidikan menggunakan jumlah *cluster* awal sebanyak 2, 3, 4, 5, dan 6 *cluster*. Diperoleh hasil bahwa jumlah *cluster* yang optimal menggunakan nilai rasio simpangan baku dalam *cluster* dan simpangan baku antar *cluster* untuk mengelompokkan kabupaten/kota di Pulau Kalimantan berdasarkan indikator pendidikan adalah 6 *cluster*. Kemudian Putri dkk., (2022) mengimplentasikan algoritma FCM dan *Fuzzy Possibilistic C-Means* (FPCM) untuk melakukan klasterisasi data *tweets* pada akun Twitter Tokopedia. Penelitian ini menggunakan jumlah *cluster* awal sebanyak 3, 4, 5, 6, dan 7 *cluster*. Hasil perbandingan yang didapat dari kedua metode ini dengan menggunakan indeks validitas *Modified Partition Coefficient* (MPC) ialah metode FCM merupakan metode yang lebih baik dibandingkan dengan metode FPCM.

Berdasarkan latar belakang tersebut, maka penulis tertarik untuk melakukan penelitian ilmiah dengan judul “Pengelompokan Provinsi di Indonesia Berdasarkan Data Jumlah Kejadian dan Dampak Bencana Banjir Menggunakan Metode *Fuzzy C-Means*”.

2. Tinjauan Pustaka

2.1. Data Mining

Data mining merupakan bidang ilmu yang digunakan untuk mengatasi masalah pengambilan informasi dari database yang besar (Werdiningsih dkk., 2020). *Data mining* bertujuan untuk mengestrak (mengambil intisari) pengetahuan dari sekumpulan data untuk mendapatkan struktur yang mudah dipahami (Suyanto, 2017). Beberapa teknik yang sering digunakan dalam data mining antara lain prediksi, klasifikasi, asosiasi, dan klasterisasi (*clustering*) (Werdiningsih dkk., 2020).

2.2. Analisis Multivariat

Analisis multivariat adalah salah satu analisis statistika yang digunakan untuk menganalisis data dengan banyak variabel, baik variabel bebas maupun variabel tak bebas. Analisis multivariat dibagi menjadi dua yaitu analisis dependensi dan interdependensi. Analisis dependensi bertujuan untuk menerangkan atau memprediksi variabel tak bebas dengan menggunakan dua atau lebih variabel bebas. Metode analisis yang termasuk dalam analisis dependensi di antaranya analisis regresi linear berganda dan analisis diskriminan. Analisis interdependensi bertujuan untuk memberikan makna terhadap sekelompok variabel atau untuk membuat kelompok-kelompok secara bersamaan. Metode

analisis yang termasuk dalam analisis interdependensi antara lain analisis faktor dan analisis *cluster* (Wijaya & Budiman, 2016).

2.3. Analisis Cluster

Analisis *cluster* adalah salah satu teknik analisis multivariat yang digunakan untuk mengelompokkan objek berdasarkan karakteristiknya sehingga setiap objek yang memiliki kesamaan dengan objek lain yang ditempatkan dalam *cluster* yang sama. Prinsip yang digunakan dalam analisis *cluster* ini yaitu memaksimalkan homogenitas (kesamaan) dan meminimumkan homogenitas (kesamaan) antar *cluster* (Wijaya & Budiman, 2016).

Menurut Santoso (2015) ada 2 asumsi dalam analisis *cluster*, yaitu sampel yang digunakan dapat mewakili populasi dan tidak terjadi multikolinearitas, yaitu kemungkinan adanya hubungan linier yang kuat antar variabel bebas. Salah satu cara untuk mengetahui hubungan linier tersebut yaitu dengan melihat nilai *Variance Inflation Factor* (VIF). Nilai VIF diperoleh dengan menggunakan persamaan berikut:

$$VIF_j = \frac{1}{1 - R_j^2} \quad (1)$$

di mana R_j^2 merupakan koefisien determinasi antara variabel X_j yang diregresikan dengan variabel lainnya. Jika terdapat nilai VIF yang lebih besar dari 10 maka disimpulkan terdapat multikolinearitas. Untuk mengatasi multikolinearitas dapat dilakukan dengan mengeluarkan variabel yang memiliki nilai VIF terbesar kemudian meregresikan kembali variabel-variabel yang tersisa (Gujarati, 2003).

2.4. Logika Fuzzy

Logika *fuzzy* merupakan suatu bentuk logika dimana nilai kebenaran suatu variabel adalah berupa bilangan real antara 0 hingga 1. Logika *fuzzy* digunakan untuk menangani konsep kebenaran parsial. Logika *fuzzy* memiliki keunggulan yakni nilai kebenarannya didasarkan pada derajat kebenaran yang berkisar dari 0 (mutlak salah) sampai dengan 1 (mutlak benar) (Hakim dkk., 2021).

2.4.1 Teori Himpunan Fuzzy

Himpunan *fuzzy* merupakan himpunan objek yang memiliki nilai keanggotaan yang tumpang tindih antara satu himpunan dengan himpunan yang lainnya. Himpunan tersebut ditandai dengan fungsi keanggotaan yang menjadi karakteristik dan memberikan setiap anggota himpunan suatu nilai tingkat keanggotaan yang berkisar antara 0 hingga 1. Himpunan *fuzzy* adalah himpunan di mana terdapat salah satu anggota himpunan yang juga merupakan anggota himpunan lain atau memiliki *double* keanggotaan. Oleh sebab itu, kedua himpunan akan saling beririsan dengan anggota himpunan lain yang memiliki keanggotaan bersama (Hakim dkk., 2021)

2.4.2 Fuzzy Clustering

Dalam logika *fuzzy*, terdapat metode yang umum digunakan untuk mengelompokkan data, yaitu *fuzzy clustering*. *Fuzzy clustering* merupakan salah satu teknik yang digunakan untuk menentukan *cluster* optimal dalam suatu ruang vektor berdasarkan pada bentuk normal *euclidian* jarak antar vektor. *Fuzzy clustering* sangat cocok digunakan dalam pemodelan *fuzzy* terutama dalam mengidentifikasi aturan-aturan *fuzzy* (Sari & Suranti, 2016). Terdapat beberapa algoritma dalam *fuzzy clustering*, salah satunya adalah *Fuzzy C-Means* (FCM) (Kusumadewi & Purnomo, 2010).

2.5. Fuzzy C-Means (FCM)

FCM merupakan teknik pengelompokan data dimana keberadaan setiap titik data dalam *cluster* ditentukan oleh derajat keanggotaannya. Tahap awal *clustering* menggunakan metode FCM diawali dengan menentukan pusat *cluster* yang pada tahap ini pusat cluster masih belum akurat. Dengan memperbaiki secara berulang pusat cluster dan derajat keanggotaan tiap-tiap data, dapat dilihat bahwa pusat cluster akan bergerak menuju lokasi yang benar. Perhitungan

secara berulang ini dilakukan untuk meminimumkan nilai fungsi objektif yang menggambarkan jarak antara titik data dengan pusat cluster yang terbobot oleh derajat keanggotaan setiap data (Apsari dkk., 2020).

Adapun langkah-langkah metode FCM adalah sebagai berikut:

1. Memasukkan data berupa matriks berukuran $n \times m$ untuk menentukan jumlah data dan jumlah variabel dari tiap data yang akan digunakan.
 n = banyaknya data
 m = banyaknya variabel
 X_{ij} = data ke- i ($i=1,2, \dots, n$), variabel ke- j ($j=1, 2, \dots, m$)

$$\mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1m} \\ X_{21} & X_{22} & \dots & X_{2m} \\ \vdots & \vdots & \dots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nm} \end{bmatrix} \tag{2}$$

2. Menentukan:
 - a. Jumlah *cluster* yang akan dibentuk = c (≥ 2)
 - b. Bobot eksponen = w (> 1)
 - c. Maksimum iterasi (*MaxIter*)
 - d. *Error* terkecil yang diharapkan (ϵ)
 - e. Fungsi objektif awal (P_0) = 0
 - f. Iterasi awal $t = 1$
3. Membangkitkan bilangan random μ_{ik} , $i = 1,2, \dots, n$; $k = 1,2, \dots, c$ sebagai elemen-elemen matriks kekhasan relatif (\mathbf{U}) berukuran $n \times c$.

$$\mathbf{U} = \begin{bmatrix} \mu_{11} & \mu_{12} & \dots & \mu_{1c} \\ \mu_{21} & \mu_{22} & \dots & \mu_{2c} \\ \vdots & \vdots & \dots & \vdots \\ \mu_{n1} & \mu_{n2} & \dots & \mu_{nc} \end{bmatrix} \tag{3}$$

4. Menghitung pusat *cluster* ke- k : V_{kj} , dengan $k = 1,2, \dots, c$; dan $j = 1,2, \dots, m$

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w \cdot X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \tag{4}$$

5. Menghitung fungsi objektif pada iterasi ke- t , P_t dengan persamaan:

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \tag{5}$$

6. Menghitung perubahan matriks partisi:

$$\mu_{ik} = \frac{[\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}}{\sum_{k=1}^c [\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}} \tag{6}$$

7. Mengecek kondisi berhenti:
 - a. Jika $(|P_t - P_{t-1}| < \epsilon)$ atau $t > \text{MaxIter}$ maka berhenti
 - b. Jika tidak, $t = t + 1$, ulangi langkah ke-4

(Apsari dkk., 2020)

2.6. Standarisasi Data

Perbedaan rentang data yang cukup besar antar variabel dapat mengakibatkan perhitungan jarak antar objek menjadi tidak valid, sehingga dapat menyebabkan bias di dalam *cluster*. Oleh karena itu, perlu dilakukan proses standarisasi (Yulianto & Hidayatullah, 2014). Salah satu teknik standarisasi data yang dapat digunakan yaitu mengubah data ke dalam skor *standardized* atau yang biasa disebut *z-score* dimana nilai X_i pada variabel ke- j akan menjadi nilai baru X'_{ij} berdasarkan persamaan berikut.

$$X'_{ij} = \frac{X_{ij} - \bar{X}_j}{S_j} \tag{7}$$

$$\bar{X}_j = \frac{1}{n} \times \sum_{i=1}^n X_{ij} \tag{8}$$

$$S_j = \sqrt{\frac{1}{n-1} \times \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2} \tag{9}$$

dimana:

- X'_{ij} = standardisasi data ke- i variabel ke- j
- n = banyaknya data
- X_{ij} = data ke- i variabel ke- j
- \bar{X}_j = rata-rata variabel ke- j
- S_j = deviasi standar variabel ke- j

(Suyanto, 2017)

2.7. Indeks Validitas

Indeks validasi *Modified Partition Coefficient* (MPC) merupakan bentuk modifikasi dari indeks *Partition Coefficient* (PC) yang mampu mengurangi perubahan monoton pada PC. Sehingga jika diketahui nilai indeks MPC, maka hasil evaluasi yang diperoleh terhadap jumlah *cluster* akan lebih akurat (Putri, 2022). Persamaan indeks MPC adalah sebagai berikut:

$$MPC(c) = 1 - \frac{c}{c-1} (1 - PC) \tag{10}$$

$$PC(c) = \frac{1}{n} \sum_{i=1}^n \sum_{k=1}^c (\mu_{ik})^2 \tag{11}$$

dimana :

- c : banyaknya *cluster*
- n : banyaknya data
- μ_{ik} : elemen matriks kekhasan relatif data ke- i *cluster* ke- k

Indeks MPC memiliki rentang antara $0 < MPC(c) \leq 1$, semakin besar nilai MPC maka akan menghasilkan *cluster* yang lebih baik (Prasetyo, 2014).

2.8. Bencana Banjir

Banjir merupakan fenomena dimana aliran air sungai memiliki ketinggian melebihi permukaan air normal, sehingga mengakibatkan air meluap melampaui palung sungai dan menyebabkan terjadinya genangan pada lahan yang lebih rendah di sisi sungai. Aliran air yang semakin tinggi akan terus mengalir dan menggenangi permukaan tanah. Untuk negara tropis seperti Indonesia, berdasarkan sumber airnya, bencana banjir dapat diklasifikasikan menjadi banjir yang disebabkan oleh hujan deras yang melebihi kapasitas penyaluran sistem pengaliran air, banjir yang disebabkan meningkatnya permukaan air sungai, banjir yang disebabkan kegagalan bangunan air buatan manusia, dan banjir yang diakibatkan oleh kegagalan bendungan alam atau penyumbatan aliran sungai akibat runtuhnya/longsornya tebing sungai.

Bencana banjir mengakibatkan kerugian berupa korban jiwa dan harta benda, baik milik perorangan maupun umum yang dapat mengganggu bahkan melumpuhkan kegiatan sosial ekonomi masyarakat. Adapun rincian mengenai dampak dari bencana banjir dapat diuraikan sebagai berikut:

1. Manusia
 - a Jumlah korban meninggal dunia adalah orang yang dilaporkan tewas atau meninggal dunia akibat bencana.
 - b Jumlah korban yang hilang adalah orang yang dilaporkan hilang atau tidak ditemukan atau tidak diketahui keberadaanya setelah terjadi bencana.
 - c Jumlah korban luka-luka adalah orang yang mengalami luka-luka atau sakit, dalam keadaan luka ringan, maupun luka parah/berat, baik yang berobat jalan maupun rawat inap.
 - d Jumlah korban menderita adalah orang/sekelompok orang yang menderita akibat dampak buruk bencana namun masih dapat menempati tempat tinggalnya.
 - e Jumlah korban mengungsi merupakan korban yang terpaksa atau dipaksa keluar dari tempat tinggalnya ke tempat yang lebih aman dalam upaya menyelamatkan diri/jiwa untuk jangka waktu yang belum pasti sebagai akibat dampak buruk bencana.
2. Harta benda perorangan yakni rumah yang tergenang, rusak, dan hanyut akibat dampak bencana banjir
3. Fasilitas
 - a Fasilitas pendidikan yang tergenang, rusak, dan hanyut meliputi sekolah, madrasah, atau pesantren.
 - b Fasilitas kesehatan yang tergenang, rusak, dan hanyut meliputi rumah sakit, puskesmas, puskesmas pembantu/pustu.
 - c Fasilitas peribadatan yang tergenang, rusak, dan hanyut meliputi masjid, gereja, vihara dan pura

(Bakornas PB, 2007)

3. Metodologi Penelitian

3.1. Rancangan Penelitian

Penelitian ini merupakan penelitian non-eksperimen yaitu tanpa memberikan perlakuan apapun terhadap objek penelitian, yakni hanya mengambil data yang sudah tersedia (data sekunder). Penelitian ini bersifat *ex post facto*, di mana data dikumpulkan setelah semua peristiwa terjadi dan tidak ada manipulasi atau intervensi terhadap variabel bebas (Ibrahim dkk., 2018).

3.2. Sumber Data dan Variabel Penelitian

Penelitian ini menggunakan data sekunder dari diperoleh dari pendataan jumlah kejadian, korban, dan kerusakan menurut Provinsi di Indonesia tahun 2017-2021 oleh Badan Nasional Penanggulangan Bencana melalui *website* <https://dibi.bnpb.go.id/>. Selanjutnya variabel-variabel yang digunakan dalam penelitian ini adalah:

1. Jumlah kejadian bencana sebagai variabel X_1
2. Korban meninggal sebagai variabel X_2
3. Korban hilang sebagai variabel X_3
4. Korban terluka sebagai variabel X_4
5. Korban menderita sebagai variabel X_5
6. Korban mengungsi sebagai variabel X_6
7. Kerusakan rumah sebagai variabel X_7
8. Kerusakan fasilitas pendidikan sebagai variabel X_8
9. Kerusakan fasilitas kesehatan sebagai variabel X_9
10. Kerusakan fasilitas peribadatan sebagai variabel X_{10}

3.3. Metode Analisis

Analisis data pada penelitian ini yaitu:

1. Melakukan analisis statistika deskriptif
2. Melakukan standarisasi data menggunakan persamaan (7), (8) dan (9).
3. Mendeteksi multikolinearitas data dengan menghitung nilai VIF menggunakan persamaan (1)
4. Melakukan pengelompokan pada objek pengamatan dengan menggunakan metode *Fuzzy C-Means* sebagai berikut.
 - a. Menentukan
 - a.) jumlah *cluster* (c) : 2, 3 dan 4
 - b.) Pangkat (m) : 2
 - c.) Maksimum iterasi ($MaxIter$) : 1.000
 - d.) *Error* terkecil yang diharapkan (ϵ) : 10^{-5}
 - e.) Fungsi objektif awal (P_0) : 0
 - b. Membangkitkan bilangan acak μ_{ik} , sebagai elemen matriks partisi awal \mathbf{U}
 - c. Menghitung pusat *cluster* ke- k menggunakan persamaan (4)
 - d. Menghitung fungsi objektif pada iterasi ke- t , dengan menggunakan persamaan (5)
 - e. Menghitung perubahan matriks partisi menggunakan persamaan (6)
 - f. Mengulangi langkah c sampai dengan langkah e hingga $|P_t - P_{t-1}| < \epsilon$ atau $t > MaxIter$ terpenuhi
 - g. Mengulangi langkah b sampai dengan langkah g untuk setiap jumlah *cluster* yang digunakan
5. Menghitung indeks validitas untuk setiap *cluster* menggunakan persamaan (10) dan (11).
6. Menginterpretasikan hasil pengelompokan berdasarkan indeks validitas.

4. Hasil dan Pembahasan

4.1. Statistika Deskriptif Data Penelitian

Analisis statistika deskriptif dilakukan untuk mendapatkan gambaran karakteristik data penelitian. Kemudian nilai-nilai statistika deskriptif untuk variabel prediktor ditampilkan dalam Tabel 1 berikut:

Tabel 1. Statistika Deskriptif

Variabel	Rata-rata	Maks	Min
X_1	157,82	924	12
X_2	31,15	250	0
X_3	6,50	101	0
X_4	85	1.113	0
X_5	399.698	2.702.033	770
X_6	48.128	538.343	59
X_7	4899,5	54.430	0
X_8	98,71	603	1
X_9	17,59	129	0
X_{10}	78	646	1

Berdasarkan Tabel 1, untuk jumlah kejadian banjir sendiri memiliki rata-rata yaitu sebesar 158 kejadian, dengan jumlah kejadian banjir terkecil yaitu sebanyak 12 kejadian dan jumlah kejadian terbesar yaitu sebanyak 924 kejadian. Rata-rata jumlah korban meninggal dunia adalah sebanyak 31 orang, adapun jumlah korban meninggal dunia terkecil yaitu sebanyak 0 orang dan jumlah korban meninggal dunia terbesar yaitu sebanyak 1.323 orang. Dampak lain yang juga diakibatkan oleh bencana banjir yaitu terdapat sejumlah korban yang dilaporkan hilang, terluka, menderita dan harus mengungsi setelah terjadi bencana. jumlah korban hilang terkecil yaitu sebanyak 0 orang, sedangkan jumlah korban hilang terbesar yaitu sebanyak 101 orang. Rata-rata korban hilang adalah sebanyak 7 orang. Selain menyebabkan sejumlah korban meninggal dunia dan hilang, bencana banjir juga menyebabkan sejumlah korban terluka. Rata-rata jumlah korban terluka adalah sebanyak 85 orang, dengan jumlah korban terluka terkecil yaitu sebanyak 0 orang, sedangkan jumlah korban terluka terbesar yaitu sebanyak 1.113 orang. Jumlah korban menderita terkecil yaitu sebanyak 770 orang, jumlah korban menderita terbesar yaitu sebanyak 2.702.033 orang. Rata-rata jumlah korban menderita adalah sebanyak 400 orang. Rata-rata jumlah korban mengungsi adalah sebanyak 48.128 orang dengan jumlah korban mengungsi terkecil yaitu sebanyak 59 orang dan jumlah korban mengungsi terbesar yaitu sebanyak 538.343 orang.

Dampak lain yang ditimbulkan bencana banjir adalah kerusakan fasilitas-fasilitas seperti rumah penduduk, fasilitas pendidikan, fasilitas kesehatan dan fasilitas peribadatan. Jumlah kerusakan rumah terkecil sebanyak 0 rumah, sedangkan jumlah kerusakan rumah terbesar sebanyak 54.430 rumah dengan rata-rata kerusakan rumah adalah sebanyak 4.900 rumah. Rata-rata jumlah fasilitas pendidikan yang rusak adalah sebanyak 99 fasilitas, dengan kerusakan fasilitas pendidikan terkecil sebanyak 1 fasilitas, sedangkan jumlah kerusakan fasilitas pendidikan terbesar yaitu sebanyak 603 fasilitas. Kerusakan fasilitas kesehatan memiliki rata-rata sebesar 18 fasilitas. Dengan kerusakan fasilitas kesehatan terkecil sebanyak 0 fasilitas, sedangkan jumlah kerusakan fasilitas kesehatan terbesar yaitu sebanyak 129 fasilitas. Kerusakan fasilitas peribadatan terkecil adalah sebanyak 1 fasilitas sedangkan kerusakan fasilitas peribadatan terbesar sebanyak 646 fasilitas. Rata-rata kerusakan fasilitas peribadatan adalah sebesar 78 fasilitas.

4.2. Standarisasi Data

Perhitungan standarisasi data dilakukan menggunakan persamaan (7). Langkah awal sebelum melakukan perhitungan standarisasi data adalah menghitung nilai deviasi standar dari variabel ke- j (S_j) menggunakan persamaan (9). Perhitungan deviasi standar dari variabel ke-1 (S_1) yaitu sebagai berikut:

$$S_1 = \sqrt{\frac{(347 - 157,82)^2 + (240 - 157,82)^2 + \dots + (198 - 157,82)^2}{33}}$$

$$= 192,6991$$

Dilakukan juga perhitungan deviasi standar variabel ke-2 hingga ke-8. Adapun hasil perhitungan deviasi standar dari setiap variabel adalah sebagai berikut:

Tabel 2. Nilai Deviasi Standar

Variabel	Deviasi Standar
X_1	192,6991
X_2	51,8448
X_3	18,7297
X_4	242,6955
X_5	554.303,221
X_6	96.238,5839
X_7	9.880,1624
X_8	146,0435
X_9	36,0951
X_{10}	148,55

Setelah mendapatkan nilai deviasi standar setiap variabel, langkah selanjutnya adalah melakukan perhitungan standarisasi data. Contoh perhitungan standarisasi data menggunakan data provinsi yang pertama yaitu Provinsi Aceh adalah sebagai berikut.

$$X'_{1,1} = \frac{347 - 157,82}{192,6991} = 0,9817$$

$$\vdots$$

$$X'_{1,10} = \frac{16 - 78}{148,55} = -0,4174$$

Perhitungan standarisasi juga dilakukan pada data pengamatan ke-2 (Provinsi Sumatera Utara) hingga data pengamatan ke-34 (Provinsi Papua) untuk setiap variabel. Hasil standarisasi data secara lengkap dapat dilihat pada table berikut.

Tabel 3. Standarisasi Data

Data	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
1	0,9817	- 0,4658	- 0,1869	- 0,3502	0,5219	0,4289	- 0,3795	0,5184	- 0,4596	- 0,4174

2	0,4264	0,6337	0,4004	- 0,1195	0,5065	- 0,1394	0,6523	- 0,0665	- 0,3765	- 0,0943
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
34	- 0,5699	1,5016	5,0454	4,2358	0,3183	- 0,2841	- 0,2857	- 0,4293	- 0,3210	- 0,3365

4.3. Pendeteksian Multikolinearitas

Pendeteksian multikolinearitas dilakukan dengan melihat nilai VIF. Berikut ini merupakan hasil perhitungan nilai VIF pada setiap variabel.

$$VIF_1 = \frac{1}{1 - R_1^2} = \frac{1}{1 - 0,8846} = 8,6666$$

⋮

$$VIF_{11} = \frac{1}{1 - R_{11}^2} = \frac{1}{1 - 0,9264} = 13,5941$$

Nilai VIF secara lengkap dapat dilihat pada Tabel 4 sebagai berikut:

Tabel 4. Nilai VIF

Variabel	VIF
X_1	12,7012
X_2	9,5109
X_3	10,4489
X_4	7,7349
X_5	30,2803
X_6	10,3063
X_7	4,0003
X_8	6,5933
X_9	7,6938
X_{10}	25,6508

Berdasarkan Tabel 4, terdapat 5 variabel yang memiliki nilai VIF lebih besar dari 10, sehingga dapat dikatakan bahwa terdapat multikolinearitas antar variabel dan belum dapat dilanjutkan ke dalam proses pengelompokan. Untuk mengatasi multikolinearitas dapat dilakukan dengan mengeluarkan variabel yang memiliki nilai VIF paling besar kemudian meregresikan kembali variabel yang tersisa. Setelah mengeluarkan beberapa variabel dan melakukan perhitungan regresi didapatkan nilai VIF akhir sebagai berikut:

Tabel 5. Nilai VIF setelah mengeluarkan variabel X_5 dan X_{10}

Variabel	VIF
X_1	3,0556
X_2	9,5038
X_3	8,5167
X_4	5,7620
X_6	2,0984
X_7	3,8144
X_8	5,3743
X_9	5,3293

Berdasarkan Tabel 5 terlihat bahwa semua variabel yang tersisa memiliki nilai VIF lebih kecil dari 10, sehingga dapat dikatakan bahwa tidak terdapat multikolinearitas antar variabel dan dapat dilanjutkan ke proses pengelompokan dengan menggunakan variabel yang tidak dikeluarkan

4.4. Pengelompokan Fuzzy C-Means

Adapun langkah-langkah pengelompokan menggunakan metode *Fuzzy C-Means* sebagai berikut:

4.4.1 Menentukan Nilai-Nilai Paramater FCM

Sebelum melakukan proses pengelompokan dengan menggunakan metode FCM, langkah awal yang perlu dilakukan adalah menentukan banyak *cluster* yang akan digunakan (c), nilai bobot eksponen (w), maksimum iterasi ($MaxIter$), nilai *error* terkecil yang diharapkan (ϵ) dan fungsi objektif awal (P_o). Pada Penelitian ini menggunakan banyak *cluster* (c) = 2,3 dan 4, $w = 2$, $MaxIter = 1.000$, $\epsilon = 10^{-5}$ dan $P_o = 0$. Sebagai contoh perhitungan pada penelitian ini, dilakukan dengan menggunakan $c = 2$.

4.4.2 Membangkitkan Bilangan Acak

Setelah menentukan nilai-nilai parameter, langkah selanjutnya adalah melakukan pembangkitan bilangan acak μ_{ic} sebagai elemen matriks keanggotaan awal U . Nilai keanggotaan awal disajikan pada Tabel 6 sebagai berikut:

Tabel 6. Nilai Keanggotaan Awal

No	Provinsi	Cluster 1 (μ_{i1})	Cluster 2 (μ_{i2})
1	Aceh	0,1277	0,8723
2	Sumatera Utara	0,1250	0,8750
⋮	⋮	⋮	⋮
34	Papua	0,5833	0,4167

4.4.3 Menghitung Pusat Cluster

Setelah membangkitkan bilangan acak μ_{ic} , langkah selanjutnya adalah menghitung pusat *cluster* awal menggunakan persamaan (4). Adapun contoh perhitungan pusat *cluster* awal pada pusat *cluster* 1 adalah sebagai berikut:

$$\begin{aligned}
 v_{1,1} &= \frac{\sum_{i=1}^{34} ((\mu_{ik})^2 X'_{i,1})}{\sum_{i=1}^{34} (\mu_{ik})^2} \\
 &= \frac{((0,2394)^2 \times 0,3724) + ((0,1061)^2 \times 0,2372) + \dots + ((0,1988)^2 \times 0,7296)}{(0,2394)^2 + (0,1061)^2 + \dots + (0,1988)^2} \\
 &= 0,3786
 \end{aligned}$$

Dilakukan perhitungan yang sama pada *cluster* 2 untuk mendapatkan pusat *cluster* selanjutnya. Sehingga pusat *cluster* yang telah diperbarui dapat dilihat pada Tabel 7 sebagai berikut

Tabel 7. Pusat Cluster Awal

Variabel	Pusat Cluster	
	1	2
X_1	-0,0463	0,1257
X_2	-0,0076	-0,0441
X_3	-0,0612	-0,0216
X_4	-0,1086	0,0085
X_6	-0,2070	0,2392
X_7	-0,0035	-0,0649
X_8	0,0522	-0,1109
X_9	0,0889	-0,1752

Berdasarkan Tabel 7 dapat diketahui bahwa nilai -0,0463 menunjukkan nilai pusat *cluster* variabel pertama di *cluster* pertama, nilai -0,0076 menunjukkan nilai pusat *cluster* variabel kedua di *cluster* pertama, nilai -0,0612

menunjukkan nilai pusat *cluster* variabel ketiga di *cluster* pertama dan demikian seterusnya hingga pada variabel kedelapan di *cluster* kedua.

4.4.4 Menghitung Fungsi Objektif

Setelah menghitung pusat *cluster* awal, langkah selanjutnya adalah menghitung nilai fungsi objektif menggunakan nilai keanggotaan pada Tabel 7 untuk $t = 1$ dengan menggunakan persamaan (5) untuk $j = 1, 2, 3, 4, 6, 7, 8$ dan 9, hasil perhitungannya adalah sebagai berikut:

$$\begin{aligned} \sum_j (X_{1j} - V_{1j})^2 &= (0,9817 - (-0,0463))^2 + (-0,4658 - (-0,0076))^2 + \dots + (-0,4596 - 0,0889)^2 \\ &= 2,5131 \\ \vdots \\ \sum_j (X_{2j} - V_{2,j})^2 &= (0,9817 - 0,1257)^2 + (-0,4658 - (-0,0441))^2 + \dots + (-0,4596 - (-0,1752))^2 \\ &= 46,8648 \\ P_1 &= \sum_i \sum_k \left(\left[\sum_j (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^2 \right) \\ &= (2,5131 \times (0,1277)^2) + (1,4484 \times (0,8723)^2) + \dots + (46,8648 \times (0,4167)^2) \\ &= 151,9975 \end{aligned}$$

Berdasarkan hasil perhitungan, diperoleh nilai fungsi objektif untuk $i = 1, 2, 3, \dots, 34$; $c = 1$ dan 2; $j = 1, 2, 3, 4, 6, 7, 8, 9$ pada $t = 1$ sebesar 151,9975. Nilai fungsi objektif awal (P_0) adalah 0 sehingga $|P_1 - P_0| = 151,9975 > \epsilon = 10^{-5}$ karena perubahan fungsi objektif masih lebih besar dari nilai ϵ , maka proses dilanjutkan ke iterasi berikutnya.

4.4.5 Menghitung Perubahan Matriks Keanggotaan

Setelah menghitung nilai fungsi objektif pada $t = 1$, langkah selanjutnya adalah menghitung perubahan nilai keanggotaan menggunakan persamaan (6). Contoh perhitungan menggunakan data nilai keanggotaan 1 (Provinsi Aceh) sebagai berikut:

$$\begin{aligned} \mu_{11} &= \frac{[\sum_j (X_{1j} - V_{1j})^2]^{-\frac{1}{2-1}}}{\sum_k [\sum_j (X_{1j} - V_{kj})^2]^{-\frac{1}{2-1}}} \\ &= \frac{(2,5131)^{-1}}{(2,5131)^{-1} + (1,4484)^{-1}} \\ &= 0,6736 \end{aligned}$$

Perhitungan perubahan nilai keanggotaan dilakukan pada data nilai keanggotaan ke-2 (Provinsi Sumatera Utara) hingga data nilai keanggotaan ke-34 (Provinsi Papua) untuk setiap *cluster*. Hasil nilai keanggotaan yang telah diperbarui dapat dilihat pada Tabel 8 sebagai berikut:

Tabel 8. Nilai Keanggotaan yang Telah Diperbarui

No	Provinsi	Cluster 1 (μ_{i1})	Cluster 2 (μ_{i2})
1	Aceh	0,6736	0,3264
2	Sumatera Utara	0,7318	0,2682
⋮	⋮	⋮	⋮
34	Papua	0,7132	0,2868

Setelah menghitung perubahan matriks keanggotaan, langkah selanjutnya adalah menghitung kembali pusat *cluster*, fungsi objektif dengan pusat *cluster* yang telah diperbarui dan menghitung perubahan matriks kembali. Iterasi berhenti ketika $|P_t - P_{t-1}| < 10^{-5}$ atau $t > 1000$. Berdasarkan hasil perhitungan, iterasi berhenti pada iterasi ke-16. Diperoleh nilai keanggotaan akhir 34 Provinsi sebagai berikut:

Tabel 9. Nilai Keanggotaan Akhir

No	Provinsi	Cluster 1 (μ_{i1})	Cluster 2 (μ_{i2})
1	Aceh	0,8888	0,1112
2	Sumatera Utara	0,7728	0,2272
\vdots	\vdots	\vdots	\vdots
34	Papua	0,4327	0,5673

Berdasarkan nilai keanggotaan pada Tabel 9, data pengamatan ke-1 (Provinsi Aceh) mempunyai nilai keanggotaan sebesar 0,8888 pada *cluster* pertama dan 0,1112 pada *cluster* kedua. Diperoleh nilai keanggotaan terbesar yaitu 0,8888 sehingga data pengamatan ke-1 paling tepat menjadi anggota *cluster* pertama. Data pengamatan ke-2 (Kabupaten Sumatera Utara) mempunyai nilai keanggotaan sebesar 0,7728 pada *cluster* pertama dan 0,2272 pada *cluster* kedua. Diperoleh nilai keanggotaan terbesar yaitu 0,7728 sehingga data pengamatan ke-2 paling tepat menjadi anggota *cluster* pertama. Penentuan *cluster* yang diikuti dilakukan dengan cara yang sama pada data pengamatan ke-3 hingga data pengamatan ke-34. Sehingga dapat diketahui anggota *cluster* dengan $C = 2$ adalah sebagai berikut:

Tabel 10. Hasil Pengelompokan FCM

Cluster	Jumlah Anggota	Anggota Cluster	
		Kode	Provinsi
1	27	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 14, 16, 17, 18, 20, 21, 23, 24, 25, 26, 28, 29, 30, 31, 32, 33	Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Bangka Belitung, Kepulauan Riau, DKI Jakarta, DI Yogyakarta, Banten, Bali, Nusa Tenggara Barat, Kalimantan Barat, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat
2	7	12, 13, 15, 19, 22, 27, 34	Jawa Barat, Jawa Tengah, Jawa Timur, Nusa Tenggara Timur, Kalimantan Selatan, Sulawesi Selatan, Papua

Selain perhitungan pada $C = 2$, dilakukan juga perhitungan dengan langkah yang sama terhadap banyak *cluster* (C) 3 dan 4.

4.5 Menghitung Indeks Validitas

Langkah awal sebelum menghitung nilai MPC adalah menghitung nilai PC dari *cluster* ke- c menggunakan persamaan (11). Contoh perhitungan nilai MPC dengan $c = 2$ adalah sebagai berikut:

$$\begin{aligned}
 PC(2) &= \frac{1}{34} \sum_{i=1}^{34} \sum_{k=1}^2 (\mu_{ik})^2 \\
 &= \frac{1}{34} ((0,8888^2) + (0,1112^2) + \dots + (0,5673^2)) \\
 &= 0,8268
 \end{aligned}$$

Dari hasil perhitungan didapatkan nilai PC dengan $c = 2$ sebesar 0,8268

$$\begin{aligned}
 MPC(2) &= 1 - \frac{2}{2-1} (1 - 0,8268) \\
 &= 0,6536
 \end{aligned}$$

Dari hasil perhitungan didapatkan nilai MPC dengan $c = 2$ sebesar 0,6506. Adapun hasil perhitungan nilai indeks validitas untuk $c = 2, 3$, dan 4 dapat dilihat pada Tabel 11 sebagai berikut:

Tabel 11. Nilai MPC

Banyak Cluster	MPC
2	0,6536

3	0,6030
4	0,4854

Berdasarkan Tabel 11 dapat dilihat bahwa nilai MPC dengan $c = 2$ merupakan nilai MPC terbesar dibandingkan dengan $c = 3$ dan 4 sehingga diperoleh *cluster* optimal adalah sebanyak 2 *cluster*.

4.6 Interpretasi Hasil Cluster

Setelah dilakukan pengelompokan menggunakan metode FCM kemudian dilakukan validasi *cluster* dengan melihat nilai MPC pada masing-masing hasil pengelompokan yang dapat dilihat pada Tabel 12.

Tabel 12. Nilai Rata-rata Variabel Masing-masing *Cluster*

Variabel	Cluster	
	1	2
X_1	103,9259	365,7143
X_2	11,4074	107,2857
X_3	1,7778	24,7143
X_4	7,6296	383,4286
X_6	26.229,5185	132.591,4286
X_7	2.539,2963	14.003,2857
X_8	50,8519	283,2857
X_9	7,7407	55,5714

Berdasarkan Tabel 12 dapat dilihat rata-rata variabel dari masing-masing *cluster* di mana terdapat perbedaan rata-rata untuk setiap variabel dari kedua *cluster* yang terbentuk. Dari segi jumlah kejadian bencana banjir, korban meninggal, korban hilang, korban terluka, korban mengungsi, kerusakan rumah, kerusakan fasilitas pendidikan, dan kerusakan fasilitas kesehatan pada *cluster* 2 memiliki rata-rata yang lebih tinggi dibandingkan dengan *cluster* 1.

5. Kesimpulan

Berdasarkan dari hasil analisis dan pembahasan, maka didapatkan kesimpulan sebagai berikut:

1. Berdasarkan hasil indeks validitas MPC jumlah *cluster* yang optimal untuk mengelompokkan Provinsi di Indonesia berdasarkan data jumlah kejadian dan dampak bencana banjir menggunakan metode FCM adalah sebanyak 2 *cluster* dengan nilai MPC terbesar adalah 0,6536.
2. Hasil pengelompokan optimal menggunakan metode FPCM pada data jumlah kejadian dan dampak bencana banjir menurut Provinsi di Indonesia dengan nilai MPC sebesar 0,6536 pada jumlah *cluster* sebanyak 2 yaitu terdapat 27 Provinsi yang masuk dalam anggota *cluster* 1 dan 7 Provinsi yang masuk dalam anggota *cluster* 2. Adapun provinsi yang termasuk dalam anggota *cluster* 1 yaitu Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Bangka Belitung, Kepulauan Riau, DKI Jakarta, DI Yogyakarta, Banten, Bali, Nusa Tenggara Barat, Kalimantan Barat, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat. Sedangkan provinsi yang termasuk dalam anggota *cluster* 2 yaitu Jawa Barat, Jawa Tengah, Jawa Timur, Nusa Tenggara Timur, Kalimantan Selatan, Sulawesi Selatan, Papua.

References

Apsari, G. R., Pradana, M. S., & Chandra, N. E. (2020). Implementasi Fuzzy C- Means dan Possibilistik C-Means Pada Data Performance Mahasiswa. *Unisa Journal of Mathematics and Computer Science (UJMC)*, 6(2), 39–48.

Bakornas PB. (2007). *Pengenalan karakteristik bencana dan upaya mitigasinya di Indonesia : Vol. II*. Jakarta : Direktorat Mitigasi, Lahar Bakornas PB.

- Findayani, A. (2015). Kesiap Siagaan Masyarakat Dalam Penanggulangan Banjir Di Kota Semarang. *Jurnal Geografi : Media Informasi Pengembangan Dan Profesi Kegeografian*, 12(1), 102–114.
- Gujarati, D. (2003). *Ekonometri Dasar*. Terjemahan: Sumarno Zain. Jakarta: Erlangga.
- Hakim G. P. N., Diah S, Ahmad F, Fajar R. I. M., & Setiyo B. (2021). *Sistem Fuzzy : Panduan Lengkap Aplikatif*. Yogyakarta : Andi (Anggota IKAPI).
- Ibrahim, A., Asrul, H. A., Madi, Baharuddin, Muhammad, A. A., Darmawati. (2018). *Metodologi Penelitian*. Makassar : Gunadarma Ilmu.
- Islamy, U., Nursidah, D. R., Narendra, I. S., Anshori, M. L., & Edi, W. (2022). Pengelompokan Provinsi Di Indonesia Berdasarkan Indikator Dampak Bencana Banjir Tahun 2017-2020. *Bimaster*, 11(2), 381–388.
- Kusumadewi, S., & Purnomo, H. (2010). *Aplikasi Logika Fuzzy untuk Pendukung Keputusan Edisi 2*. Yogyakarta: Graha Ilmu.
- Messakh, G. C., Hayati, M. N., & Sifriyani. (2023). Comparison K-Means and Fuzzy C-Means in Regencies/Cities Grouping Based on Educational Indicators. *Jurnal Varian*, 7(1), 99-114.
- Prasetyo, E. (2014). *Data Mining Mengelolah Data Menjadi Informasi Menggunakan Matlab Edisi Baru*. Yogyakarta : Penerbit CV. ANDI.
- Ramadhan, A., & Handinata, R. (2019). Implementasi Algoritma Fuzzy C-Means untuk Pengelompokan Wilayah Bencana Banjir. *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI) 11*, 171-177.
- Santoso, S. (2015). *Menguasai Statistik Multivariat*. Jakarta : PT Elex Media Komputindo.
- Sari, H. L., & Suranti, D. (2016). Perbandingan Algoritma Fuzzy C-Means (FCM) dan Algoritma Mixture Dalam Pencusteran Data Curah Hujan Kota Bengkulu. *Proceeding SNATI 2016*, 7–15.
- Setiawan, R., Ernawati, & Efendi, R. (2018). Klasifikasi Kawasan Permukiman Tingkat Kelurahan Untuk Pembangunan Sistem Berbasis Data Kualitas Permukiman (Studi Kasus: 67 Kelurahan di Kota Bengkulu). *Jurnal Pseudocode*, 5(1), 45-55.
- Silitonga, P. D. P. S., & Morina, I. S. (2017). Klusterisasi Pola Penyebaran Penyakit Pasien Berdasarkan Usia Pasien Dengan Menggunakan K-Means Clustering. *Jurnal TIMES*, VI(2), 22–25.
- Suyanto. (2017). *Data Mining Untuk Klasifikasi dan Klasterisasi Data*. Bandung: Informatika.
- Werdiningsih, I., Nuqoba, B., & Muhammadun. (2020). *Data Mining Menggunakan Android, Weka, dan SPSS*. Surabaya: Airlangga University Press.
- Wijaya, T & Budiman S. (2016). *Analisis Multivariat Untuk Penelitian Manajemen*. Yogyakarta : Pohon Cahaya.
- Yulianto, S & Hidayatullah, K. H. (2014). Analisis Klaster untuk Pengelompokan Kabupaten/Kota di Provinsi Jateng berdasarkan Indikator Kesejahteraan Rakyat. *Jurnal Akd. Statistika Muhammadiyah Semarang*, 2(1), 56-63.
- Zalmita, N., Fitria, A., & Taher, A. (2021). Tingkat Kerugian Ekonomi Pada Bencana Banjir di Aceh Utara Tahun 2014-2019. *JURNAL GEOGRAFI Geografi Dan Pengajarannya*, 19(2), 61–68.