

Cluster Analysis of Indonesian Provinces Based on Health Performance Indicators Using Fuzzy C-Means

Nurul Fiskia Gamayanti*, Mohammad Fajri, Hartayuni Sain, Fadjryani, Iman Setiawan

Departemen Of Statistics, Universitas Tadulako, Palu, Indonesia

*Corresponding author: nurulfiskia@gmail.com

Submitted : 31 Juni 2026
Revised : 12 Agustus 2026
Accepted : 31 Agustus 2026

ABSTRACT

In accordance with the Sustainable Development Goals (SDGs), with particular emphasis on Goal 3 which focuses on ensuring a healthy life and improving the welfare of all residents across all stages of life, ranging from infancy to older adulthood, it is closely related to looking at existing health indicators. The Indonesian region consisting of 34 provinces makes policies difficult to generalize due to the different characteristics of each region. So a cluster method is needed where regional objects will be clustered into several groups so that policies will be better implemented in a formed cluster. One clustering technique that may be employed is the Fuzzy C-Means (FCM) algorithm. FCM provides a membership value in the form of degrees ranging from 0 to 1, representing the degree of strength of the relationship of a data with each group, the analysis performed using the Fuzzy C-Means (FCM) method resulted in the formation of two clusters, which represent that cluster 1 is a cluster with indications of areas that have good health, and cluster 2 is a cluster with indications of poor, Cluster 1 comprises 12 provinces, whereas Cluster 2 includes 22 provinces.

Keywords: Fuzzy C-means, health indicator, cluster



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

I. PENDAHULUAN

Kesehatan menjadi salah satu unsur utama dalam pembangunan nasional yang berperan penting dalam meningkatkan kualitas hidup masyarakat. Selain itu, kesehatan menjadi bagian dari Tujuan Pembangunan Berkelanjutan (Sustainable Development Goals/SDGs), khususnya tujuan ketiga yang menitikberatkan pada upaya untuk menjamin kehidupan yang sehat serta meningkatkan kesejahteraan seluruh penduduk di setiap kelompok usia, mulai dari bayi hingga lanjut usia. Indonesia memiliki kewajiban untuk mewujudkan pembangunan kesehatan yang bersifat inklusif serta tersebar secara merata di seluruh wilayah. Dalam mengukur dan mengevaluasi kondisi kesehatan masyarakat, digunakan berbagai indikator kesehatan yang berfungsi sebagai alat ukur untuk menilai aspek-aspek tertentu dari kesehatan masyarakat. Indikator tersebut juga memberikan gambaran mengenai tingkat kesehatan masyarakat pada berbagai wilayah administratif, baik di tingkat kabupaten/kota maupun provinsi di Indonesia (Liana dkk., 2022).

Penyajian indikator kesehatan yang tepat dan sesuai dengan kebutuhan sangat diperlukan dalam menyediakan gambaran mengenai perkembangan kondisi kesehatan masyarakat dari waktu ke waktu. Informasi tersebut dapat dimanfaatkan sebagai dasar dalam perencanaan dan penyusunan program kesehatan yang lebih efektif guna meningkatkan derajat kesehatan masyarakat secara menyeluruh. Salah satu strategi yang dapat diterapkan adalah melakukan pengelompokan terhadap berbagai wilayah di Indonesia berdasarkan karakteristik dan tingkat kebutuhannya, sehingga penetapan prioritas pelayanan kesehatan dapat diarahkan secara lebih tepat pada daerah yang memerlukan perhatian lebih besar (Suhaeni dkk., 2018).

Sehingga analisis cluster digunakan sebagai salah satu pendekatan untuk mengatasi permasalahan tersebut dengan memberikan gambaran mengenai karakteristik masing-masing daerah berdasarkan indikator kesehatan yang dimiliki (Kontoro, 2022). Analisis cluster merupakan metode pengelempokan suatu objek ke dalam beberapa kategori berdasarkan tingkat kesamaan karakteristik yang dimilikinya. Salah satu metode cluster adalah *Fuzzy C-Means* (FCM). *Fuzzy C-Means* (FCM) merupakan salah satu metode pengelompokan yang dikembangkan berdasarkan algoritma C-Means dengan dengan mengintegrasikan konsep logika fuzzy. Melalui pendekatan ini,

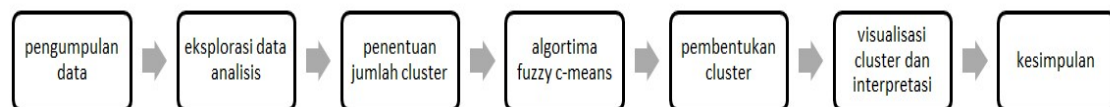
setiap objek data tidak hanya ditempatkan pada satu cluster tertentu, tetapi dapat memiliki tingkat keanggotaan pada beberapa kluster sekaligus. Pengelompokan dilakukan dengan memanfaatkan teori fuzzy untuk menentukan seberapa besar keterkaitan setiap data terhadap masing-masing kelompok (Gupta et al., 2021). FCM merupakan teknik pengelompokan data yang memungkinkan setiap objek untuk memiliki tingkat keanggotaan terhadap lebih dari satu cluster. (Singh, 2024). Keakuratan hasil pengelompokan diperoleh melalui proses iteratif yang memperbarui pusat cluster dan nilai keanggotaan setiap data hingga mencapai kondisi yang optimal (Ningrat et al., 2016).

Metode Fuzzy C-Means memanfaatkan derajat keanggotaan (*uik*), untuk menggambarkan tingkat keterkaitan setiap objek data dengan suatu kluster. Berbeda dengan metode klasterisasi tegas (*hard clustering*), FCM memungkinkan satu data memiliki keanggotaan pada beberapa kluster sekaligus dengan tingkat yang berbeda. Metode ini juga menggunakan parameter (*w*) sebagai *weighting exponent* yang berfungsi mengendalikan pengaruh derajat keanggotaan terhadap proses pengelompokan. Nilai parameter tersebut ditetapkan lebih besar dari satu ($w > 1$) guna menghasilkan pembagian kluster yang bersifat fuzzy (Desrianti dan Wijaya, 2020).

Berdasarkan kelebihan metode Fuzzy C-means tersebut, penelitian ini dilakukan dengan tujuan untuk mengclusterkan daerah provinsi yang tersebar di seluruh Indonesia dengan mempertimbangkan delapan indikator kesehatan yang digunakan, yaitu angka kematian ibu, angka kematian bayi, serta persentase balita dengan kondisi gizi kurang, angka kesakitan DBD, jumlah dokter umum, jumlah rumah sakit, presentase air minum layak, presentase jamban sehat untuk membentuk cluster sebagai bentuk tingkat kepentingan daerah berdasarkan hasil cluster yang nantinya terbentuk dengan memperhatikan 8 indikator tersebut, diharapkan penelitian ini dapat memberikan kontribusi yang baik kepada pihak-pihak terkait pengambilan keputusan untuk memberikan penanganan tepat sasaran jika wilayah provinsi dapat dibuat dalam cluster-cluster. .

II. METODE PENELITIAN

Cluster berdasarkan wilayah Provinsi di Indonesia dengan menggunakan Indikator kesehatan menggunakan analisis Fuzzy C-Means (FCM) memiliki beberapa alur penyelesaian yaitu :



Gambar 1. Skema Alur Pelaksanaan Penelitian

A. Variabel Penelitian

Penelitian ini menggunakan data berupa data indikator kesehatan di Indonesia berdasarkan 34 provinsi yang ada. Data diperoleh dari website resmi BPS. Indikator kesehatan yang digunakan sebanyak 8 indikator kesehatan yaitu :

Tabel 1. Variabel Penelitian

Variabel	Satuan
X1 (Angka Kematian Ibu)	Jiwa
X2 (Angka Kematian Bayi)	Jiwa
X3 (Persentase balita gizi kurang)	Persen (%)
X4 (Angka Kesakitan DBD)	Persen (%)
X5 (Jumlah Dokter Umum)	Orang (dokter)
X6 (Jumlah RS)	Unit rumah sakit
X7 (Persentase Air Minum Layak)	Persen (%)
X8 (Persentase Jamban Sehat)	Persen (%)

B. Fuzzy C-Means

Metode clustering secara umum dapat dibagi menjadi tiga kategori, salah satunya adalah Fuzzy Clustering (FC). Perbedaan Fuzzy Clustering dengan metode clustering pada umumnya adalah bagian penempatan setiap objek hanya

pada satu kelompok tertentu, sedangkan pada FC memberikan kemampuan kepada suatu objek untuk memiliki keterikatan terhadap lebih dari satu klaster secara simultan dengan tingkat keterikatan objek pada masing-masing cluster ditunjukkan melalui nilai keanggotaan yang berada dalam skala 0 sampai 1 (Abdullah, 2023). Nilai yang bernilai 0 mengindikasikan bahwa suatu objek tidak memiliki keterikatan atau keanggotaan terhadap suatu klaster, sementara nilai yang mencapai 1 menunjukkan tingkat keanggotaan penuh suatu objek terhadap klaster tertentu (Dubey, 2018). Konsep ini berlandaskan pada konsep teori himpunan fuzzy, yang menjelaskan bahwa setiap objek memiliki derajat keanggotaan yang berbeda pada masing-masing cluster. Dalam pendekatan fuzzy clustering, metode Fuzzy C-Means (FCM) menjadi salah satu algoritma yang paling banyak digunakan untuk pengelompokan yang menentukan posisi suatu data dalam cluster berdasarkan besarnya derajat keanggotaan yang dimiliki data tersebut (Qur'ani & Widiastuti, 2023).

Analisis yang digunakan dalam metode FCM, setiap objek pada dataset mempunyai nilai keanggotaan yang merepresentasikan keterkaitannya dengan seluruh cluster yang terbentuk. Nilai ini digunakan untuk menggambarkan tingkat keterkaitan suatu objek terhadap masing-masing cluster (Hashemi, 2023). Tingkat ketidakjelasan atau fleksibilitas keanggotaan tersebut dikendalikan oleh parameter yang dikenal sebagai fuzzifier (Zhou, 2020). Keberadaan parameter ini memungkinkan sebuah objek memiliki peluang untuk menjadi anggota lebih dari satu cluster secara bersamaan, sehingga metode FCM lebih mampu menangani data yang memiliki batas antar kelompok yang tidak tegas atau saling tumpang tindih (Amaliya et al., 2022).

Algoritma *Fuzzy C-Means* (FCM) pada pengukuran tingkat keterikatan keanggotaan setiap data dilakukan berdasarkan fungsi keanggotaan yang dipilih. Fungsi keanggotaan (*membership function*) berperan sebagai kurva yang merepresentasikan hubungan antara data masukan dengan nilai tingkat keanggotaannya pada suatu himpunan fuzzy (Lei, 2018). FCM adalah hasil pengembangan dari pendekatan C-Means yang mengintegrasikan konsep logika fuzzy dalam proses pengelompokan data (Lestari dkk., 2023). Dengan pendekatan ini, setiap objek dapat memiliki tingkat keanggotaan pada lebih dari satu klaster, sehingga proses pengalokasian data ke dalam kelompok dilakukan berdasarkan prinsip teori fuzzy (Gupta et al., 2021).

Algoritma analisis *Fuzzy C-Means* (FCM) (Desrianti dan Wijaya, 2020):

- Input data yang akan diclusterkan, berbentuk matriks X yang berukuran $k \times j$
Dimana :
 k = jumlah data
 j = jumlah variabel
- Menentukan inisialisasi awal, yakni jumlah *cluster* (c), *weighting exponent* ($w = 2$), maks iterasi ($MaxIter$), error kecil ($\epsilon = 10^{-9}$), fungsi objektif awal ($P_0 = 0$), dan iterasi awal ($t = 1$).
- Pembangkitan Bilangan Acak μ_{ik} , $i = 1, 2, \dots, c$; $k = 1, 2, \dots, n$ Sebagai komponen matriks partisi awal U dengan memenuhi persyaratan yang telah ditentukan $\sum_{k=1}^c \mu_{ik} = 1$.
- Menentukan pusat setiap cluster pada masing-masing kelompok berdasarkan persamaan berikut.

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w x_{ij})}{\sum_{i=1}^n (\mu_{ik})^w}$$

dimana,

- μ_{ik} : Nilai pada (matriks U) pada *cluster* ke- i dan data ke- k
 w : *Weighting exponent*
 x_{ij} : Nilai (pada matriks X) pada atribut ke- i dan data ke- j
 V_{kj} : Centroid *cluster* ke- k dan variabel ke- j

- Menghitung fungsi objektif pada iterasi ke- t , P_t .

Jika $(|P_t - P_{t-1}| < \epsilon)$ atau $(t > MaxIter)$ maka proses berhenti, jika tidak maka $t = t + 1$ dan ulangi langkah ke-4.

$$P_t = \sum_{k=1}^n \sum_{i=1}^c (\mu_{ik})^w (d_{ik})^w$$

dimana,

- μ_{ik} : Nilai partisi pada (matriks U) pada *cluster* ke- i dan data ke- k
 d_{ik} : jarak *euclidean* pada pusat *cluster* ke- i dan data ke- k

- P_t : Nilai objektif pada iterasi ke- n
 W : *Weighting exponent*
 f. Menghitung perubahan matriks partisi.

$$\mu_{ik} = \sum_{j=1}^c \left[\left(\frac{d_{ik}}{d_{ji}} \right)^{\frac{2}{w-1}} \right]^{-1}$$

dengan $d_{ik} = d(x_k - v_i) = \left[\sum_{j=1}^m (x_{kj} - v_{ij}) \right]^{1/2}$

dimana,

- μ_{ik} : Nilai partisi pada (matriks U) pada *cluster* ke- i dan data ke- k
 d_{ik} : Jarak *euclidean* pada pusat *cluster* ke- i dan data ke- k
 d_{jk} : Jarak *euclidean* pada pusat *cluster* ke- j dan data ke- k
 v_{kj} : Pusat *cluster* ke- k dan variabel ke- j
 w : *Weighting exponent*
 g. Iterasi kondisi berhenti, jika $(|P_t - P_{t-1}| < \epsilon)$ atau $(t > Max_{iter})$ terpenuhi. Jika tidak: $t = t + 1$, ulangi langkah ke-3.

III. HASIL DAN PEMBAHASAN

A. Statistik Deskriptif

Tabel 2. Statistik deskriptif data

Variabel Penelitian	Rata-Rata	Minimum	Maximum
Angka Kematian Ibu	129,5	20	792
Angka Kematian Bayi	938,029	143	5234
Persentase balita gizi kurang	40,882	1,2	7,9
Angka Kesakitan DBD	64,2038	14,23	186,42
Jumlah Dokter Umum	1484,765	118	7359
Jumlah Rumah sakit	76,2941	12	355
Persentase Air Minum Layak	88,189	66,49	99,42
Persentase Jamban Sehat	90,9	73,60	100

Berdasarkan Tabel 2 terlihat dengan memperhatikan 8 variabel indikator kesehatan di 34 provinsi di Indonesia bahwa setiap variabel memiliki gap yang cukup besar dari daerah terendah dengan nilai minimum jika dibandingkan dengan nilai tertinggi pada masing-masing variabel. Ini dapat memperjelas bahwa setiap provinsi memiliki karakteristik variabel yang bervariasi. Ini menjadi salah satu indikasi perlunya pengclusteran daerah untuk memetakan tingkat kepentingan berdasarkan variabel indikator kesehatan di Indonesia.

B. Penentuan Jumlah Cluster Optimum

Salah satu metode yang dapat digunakan dalam menentukan jumlah *cluster optimum* pada cluster adalah dengan menggunakan *Partition Coefficient Index* (PCI). Apabila nilai PCI mendekati nilai 1 maka *cluster* yang terbentuk akan semakin baik hasil yang diperoleh. Berikut merupakan hasil perhitungan manual dari penentuan jumlah *cluster optimum*. Berikut merupakan hasil pemilihan jumlah cluster optimal dengan membandingkan nilai PCI:

Tabel 3. Jumlah *Cluster* Optimum

Jumlah <i>Cluster</i>	<i>Partition</i> <i>Coefficient</i> <i>Index</i> (PCI)
2	0,6158
3	0,5918
4	0,4718
5	0,4272

Tabel 3 menyajikan hasil nilai PCI pada cluster 2, 3, 4, dan 5. Dengan memperhatikan nilai PCI maka terlihat cluster 2 merupakan cluster paling optimal karena memiliki nilai PCI yang paling mendekati nilai 1. Sehingga pada penelitian ini akan menggunakan jumlah *cluster* sebanyak 2.

C. Analisis Cluster Fuzzy C-Means

1. Menentukan Inisialisasi Awal

Pada penelitian ini jumlah *cluster* yang digunakan menggunakan uji *partition coefficient index*, menghasilkan 2 *cluster* yang optimal. Pada pembagian masing-masing *cluster* dalam mengelompokan provinsi di Indonesia dibagi menjadi 2 kelompok yakni *cluster* 1 dan *cluster* 2. *Cluster* 1 di definisikan sebagai tingkat kesehatan yang baik dan *cluster* 2 di definisikan sebagai tingkat kesehatan yang buruk. Dalam melakukan analisis *cluster Fuzzy C-Means* diperlukan inisialisasi awal untuk menentukan parameter awal yang akan digunakan. Berikut merupakan hasil nilai inisialisasi awal dapat dilihat pada Tabel 5.

Tabel 4. Nilai Inisialisasi Awal

w	$maxiter$	e	p_t	t	k
2	1000	10^{-9}	0	1	2

2. Menentukan Derajat Keanggotaan Awal

Dalam *Fuzzy C-Means* hal yang perlu dilakukan adalah memunculkan bilangan acak matriks U awal dengan komponen $i = 1, 2, 3, \dots, 34$; $k = 1, 2$, sebagai nilai μ_{ik} derajat keanggotaan awal. Nilai tersebut memiliki persyaratan bahwa setiap elemen baris berjumlah 1. Berikut merupakan nilai derajat keanggotaan awal yang dibangkitkan pada Tabel 5.

Tabel 5. Nilai Derajat Keanggotaan Awal

Data	μ_{i1}	μ_{i2}	$\sum \mu_{ik}$
1	0,6126	0,3874	1
2	0,6277	0,3723	1
3	0,7521	0,2479	1
4	0,6188	0,3813	1
5	0,4500	0,5500	1
6	0,5422	0,4578	1
7	0,4459	0,5541	1
8	0,5273	0,4727	1
9	0,5133	0,4867	1
...	...		
32	0,6418	0,3582	1

Data	μ_{i1}	μ_{i2}	$\sum \mu_{ik}$
33	0,7838	0,2162	1
34	0,3860	0,6140	1

3. Menghitung Pusat *Cluster* Awal

Menentukan pusat *cluster* awal untuk mengidentifikasi jarak antara data tersebut dengan pusat *cluster*. Jika suatu objek data memiliki jarak terpendek ke pusat *cluster*, maka dapat diidentifikasi bahwa objek data tersebut masuk dalam *cluster* tersebut.

$$V_{11} = \frac{\sum_{i=1}^{34} (\mu_{i1})^2 X_{i1}}{\sum_{i=1}^{34} (\mu_{i1})^2} = \frac{((0,6126)^2 0,0281) + ((0,6277)^2 0,0897) + \dots + ((0,3860)^2 - 0,6202)}{(0,6126)^2 + (0,6277)^2 + \dots + (0,3860)^2}$$

$$= -0,1005$$

$$V_{12} = \frac{\sum_{i=1}^{34} (\mu_{i2})^2 X_{i1}}{\sum_{i=1}^{34} (\mu_{i2})^2} = \frac{((0,3874)^2 0,0281) + ((0,3723)^2 0,0897) + \dots + ((0,6140)^2 - 0,6202)}{(0,3874)^2 + (0,3723)^2 + \dots + (0,6140)^2}$$

$$= 0,1104$$

Berdasarkan perhitungan di atas, maka pusat *cluster* awal V_{11} adalah -0,1005.

4. Menghitung Jarak *Cluster* Awal

Dalam mengukur tingkat kemiripan antar objek digunakan pendekatan ukuran jarak. Berikut merupakan hasil dari perhitungan pada *cluster* 1 objek 1.

$$d_{11} = \sqrt{(0,0281 - (-0,1005))^2 + \dots + (-1,322 - (-0,1868))^2}$$

$$= 1,5132$$

Berdasarkan perhitungan di atas, maka jarak *cluster* awal d_{11} adalah 1,5132. Hal ini dilakukan sebanyak jumlah objek yaitu 34 provinsi sehingga didapatkan 34 jarak *cluster* awal untuk masing-masing *cluster* 1 dan 2.

5. Menghitung Fungsi Objektif

$$P_1 = ((0,6126)^2 (1,5132)^2 + (0,6277)^2 (-1,211)^2 + \dots + (0,3860)^2 (3,2725)^2)$$

$$= 0,6136$$

Berdasarkan perhitungan di atas, maka fungsi objektif awal p_1 adalah 0,6136. Berikut merupakan hasil perhitungan fungsi objektif awal secara lengkap dapat dilihat pada Tabel 6 berikut.

Tabel 6. Nilai Fungsi Objektif Awal

Data	Fungsi Objektif
1	0,6136
2	0,7265
3	0,7158
4	0,3699
5	0,6161
6	0,6251
7	1,2607
8	0,5513
9	0,8296
....	
32	0,5353
33	1,4685
34	1,5530
total	45,9488.

6. Menghitung Perubahan Matriks Partisi

$$\mu_{i1} = \sum_{j=1}^2 \left[\left(\frac{d_{i1}}{d_{j1}} \right)^{\frac{2}{2-1}} \right]^{-1} = \frac{1}{\left(\frac{1,5132}{1,5132} \right)^2 + \left(\frac{1,5132}{1,9619} \right)^2} = \frac{1}{1,5949} = 0,6270$$

Berdasarkan perhitungan di atas, maka perubahan matriks partisi baru μ_{i1} adalah 0,6270.

Langkah ini dilakukan sebanyak 34 kali perhitungan untuk masing-masing objek.

7. Mengecek Kondisi Berhenti

Setelah diperoleh fungsi objektif awal maka dapat menggunakan rumus $(|P_t - P_{t-1}| < \varepsilon)$ atau $(|45,9488 - 0| > 10^{-9})$ maka proses iterasi selanjutnya dilanjutkan dengan mengulangi langkah 2 hingga memperoleh hasil akhir fungsi objektif $< 10^{-9}$.

Dari hasil di atas diperoleh hasil *cluster* sebagai berikut.

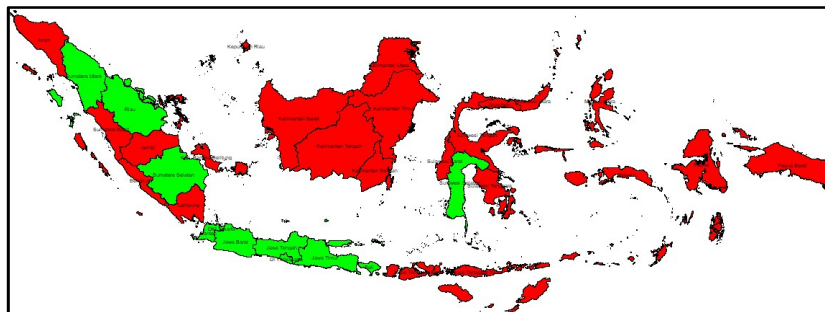
Tabel 7. Hasil Pengelompokan Fuzzy C-means

<i>Cluster</i>	Jumlah Anggota	Anggota Cluster
1	12	Aceh, Sumatera Utara, Riau, Sumatera Selatan, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, Banten, Bali, dan Sulawesi Selatan
2	22	Sumatera Barat, Jambi, Bengkulu, Lampung, Kep. Bangka Belitung, Kep. Riau, Busa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua, dan Papua Barat.

Dari tabel 7 disajikan hasil cluster yang terbentuk dari algortima Fuzzy C-means berdasarkan 8 indikator kesehatan di Indonesia terbentuk kedalam 2 cluster dengan 12 provinsi menjadi anggota *cluster* 1 dan 22 provinsi mejadi anggota *cluster* 2.

D. Visualisasi Hasil Cluster

Berdasarkan hasil pengelompokan indikator kesehatan di Indonesia menggunakan *Fuzzy C-Means* diperoleh visualisasi hasil *cluster* menggunakan peta Indonesia yang disajikan pada Gambar 2 sebagai berikut.



Gambar 2. Visualisasi Hasil Cluster Berdasarkan Indikator Kesehatan di Indonesia

Berdasarkan Gambar 2 diperoleh visualisasi hasil *cluster* dengan 2 kelompok yang masing-masing kelompok divisualisasikan dengan warna yang berbeda. Pada *cluster* 1 ditandai dengan warna hijau yang terindikasi bahwa provinsi yang masuk pada kelompok 1 merupakan daerah dengan tingkat kesehatan yang sudah baik. Pada *cluster* 2 ditandai dengan warna merah yang terindikasi bahwa provinsi yang masuk pada kelompok 2 merupakan daerah dengan tingkat kesehatan yang buruk.

Untuk mengetahui karakteristik dari masing-masing *cluster* dapat dilihat dari profilisasi *cluster* berdasarkan variabel bebas. Berikut merupakan hasil profilisasi *cluster* yang ditunjukkan pada Tabel 15.

Tabel 8. Karakteristik *Cluster*

Variabel	Cluster	
	1	2
X ₁	252,2727	70,8261
X ₂	1918,4545	469,1304
X ₃	2,9636	4,6261
X ₄	40,2091	75,6796
X ₅	3372,2727	582,0435
X ₆	161,7273	35,4348
X ₇	93,9064	85,4539
X ₈	95,2909	88,8000

Berdasarkan Tabel 8 di atas, diperoleh nilai rata-rata dari masing-masing *cluster* untuk setiap variabel. *Cluster* 1 menunjukkan nilai tertinggi dibandingkan dengan *Cluster* 2, yang mengindikasikan bahwa wilayah dalam *cluster* ini memiliki tingkat kesehatan yang lebih baik dibandingkan *cluster* lainnya. Sementara itu, *Cluster* 2 memiliki nilai terendah pada sebagian besar variabel dibandingkan dengan *cluster* 1. Namun, pada variabel X₃ dan X₄, nilai yang relatif tinggi, yang mencerminkan tingginya persentase balita dengan gizi kurang serta jumlah kasus demam berdarah dengue (DBD). Oleh karena itu, *cluster* ini dapat dikategorikan sebagai wilayah dengan tingkat kesehatan yang buruk.

IV. KESIMPULAN

Dari hasil analisis menggunakan algoritma Fuzzy C-means diperoleh hasil cluster dari 34 provinsi di Indonesia dengan menggunakan 8 indikator kesehatan yaitu angka kematian ibu, angka kematian bayi, Presentase balita gizi kurang, angka kesakitan DBD, jumlah dokter umum, jumlah rumah sakit, presentase air minum layak, presentase jamban sehat diperoleh 2 cluster. Cluster 1 merupakan cluster dengan kategori provinsi yang memiliki indikasi kesehatan baik sedangkan cluster 2 merupakan cluster dengan kategori provinsi yang memiliki indikasi kesehatan buruk jika memperhatikan 8 indikator tersebut. Cluster 1 terdiri dari 12 wilayah provinsi dan cluster 2 terdiri dari 22 wilayah provinsi.

DAFTAR PUSTAKA

- Abdullah, A. A., Ahmed, A. M., Rashid, T., & Veisi, H. (2024). Advanced clustering techniques for speech signal enhancement: A review and metanalysis of fuzzy c-means, k-means, and kernel fuzzy c-means methods. *arXiv preprint arXiv:2409.19448*.
- Amaliya, L. R., Rahim, A., & Septiani, A. (2022). Comparison of Fuzzy C-Means , Fuzzy Possibilistic C-Means and Possibilistic Fuzzy C-Means Algorithms on the Distribution of Contraceptive Users in NTB Province. *The 2nd International Seminar of Science and Technology*, 37–52.
- [BPS] Badan Pusat Statistik. (2024, Januari). Statistik Kesehatan. Diambil kembali dari <https://www.bps.go.id/id/statistics-table/2/MTQwMiMy/unmet-need-pelayanan-kesehatan-menurut-provinsi.html>
- Desrianti, R., & Wijaya, H. D. (2020). Implementasi Algoritma Fuzzy C-Means Pada Aplikasi Seleksi Karyawan Digital Talent di PT Telekomunikasi Indonesia. *Jurnal Media* 4, 879–888. <https://doi.org/10.30865/mib.v4i4.2267>

- Dubey, A. K., Gupta, U., & Jain, S. (2018). Comparative study of K-means and fuzzy C-means algorithms on the breast cancer data. *International Journal on Advanced Science, Engineering and Information Technology*, 8(1), 18-29.
- Gupta, A., Datta, S., & Das, S. (2021). Fuzzy Clustering to Identify Clusters at Different Levels of Fuzziness: An Evolutionary Multiobjective Optimization Approach. *IEEE Transactions on Cybernetics*, 51(5), 2601–2611. <https://doi.org/10.1109/TCYB.2019.2907002>
- Hashemi, S. E., Gholian-Jouybari, F., & Hajiaghahi-Keshteli, M. (2023). A fuzzy C-means algorithm for optimizing data clustering. *Expert Systems with Applications*, 227, 120377.
- Kontoro, P. I., Junaidi, J., Gamayanti, N. F., & Apusing, A. S. N. (2024, December). Clustering Indonesian Provinces Based on Poverty Levels Utilizing the Average Linkage Method with Principal Component Analysis. In *The 5th International Seminar on Science and Technology (ISST 2023)* (pp. 78-86). Atlantis Press.
- Lei, T., Jia, X., Zhang, Y., He, L., Meng, H., & Nandi, A. K. (2018). Significantly fast and robust fuzzy c-means clustering algorithm based on morphological reconstruction and membership filtering. *IEEE transactions on fuzzy systems*, 26(5), 3027-3041.
- Lestari, M., Kartini, D., Budiman, I., Faisal, M. R., & Muliadi, M. (2023). Comparison of Industrial Business Grouping Using Fuzzy C-Means and Fuzzy Possibilistic C-Means Methods. *Telematika*, 16(2), 91–102. <https://doi.org/10.35671/telematika.v16i2.2548>
- Liana, Y., Ratnasari, V., & Deskriptif, A. S. (2022). Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Kesehatan Masyarakat dengan Pendekatan Metode *Ensemble ROCK*. *Jurnal Sains Dan Seni ITS*, 11(2), 2337–3520.
- Ningrat, D. R., Asih, D., Maruddani, I., & Wuryandari, T. (2016). Analisis Cluster Dengan Algoritma K-Means Dan Fuzzy C-Means Clustering Untuk Pengelompokan Data Obligasi Korporasi. *Jurnal Gaussian*, 5(4), 641–650. <http://ejournal-s1.undip.ac.id/index.php/gaussian>
- Qur'ani, A. Y., & Widiastuti, R. S. (2023). Implementasi Fuzzy C-Means (FCM) dan Generalized Fuzzy C-Means (GFCM) dalam Clustering Produk Domestik Regional Bruto (PDRB) Menurut Pengeluaran. *ULIL ALBAB*.
- Singh, S., & Srivastava, S. (2024). Optimizing kernel possibilistic fuzzy C-means clustering using metaheuristic algorithms. *Evolving Systems*, 15(4), 1587-1606.
- Suhaeni, C., Kurnia, A., & Ristiyanti, R. (2018). Perbandingan Hasil Pengelompokan menggunakan Analisis *Cluster Berhierarchy*, *K-Means Cluster*, dan *Cluster Ensemble* (Studi Kasus Data Indikator Pelayanan Kesehatan Ibu Hamil). *Jurnal Media Infotama*, 14(1). <https://doi.org/10.37676/jmi.v14i1.469>
- Zhou, K., & Yang, S. (2020). Effect of cluster size distribution on clustering: a comparative study of k-means and fuzzy c-means clustering. *Pattern Analysis and Applications*, 23(1), 455-466.