

PENGELOMPOKAN KABUPATEN/KOTA DI PULAU KALIMANTAN BERDASARKAN INDIKATOR PERUMAHAN DAN KESEHATAN LINGKUNGAN MENGGUNAKAN METODE *FUZZY GUSTAFSON-KESSEL*

Galuh Sukma Saraswati¹, Darnah Andi Nohe^{2*}, Memi Nor Hayati³

^{1,2,3} Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Mulawarman

*e-mail : darnah.98@gmail.com

DOI: 10.14710/j.gauss.13.1.156-167

Article Info:

Received: 2024-07-24

Accepted: 2024-10-07

Available Online: 2024-10-09

Keywords:

*Fuzzy Gustafson-Kessel, Housing
and Environmental Health, Validity
Index*

Abstract: *Group analysis is a method used to classify a range of items into various groups, ensuring that items with similar features are grouped together. Group analysis is divided into two methods, namely hierarchical methods and non-hierarchical methods. One example of a non-hierarchical grouping method is Fuzzy Gustafson Kessel (FGK). FGK is a development of the Fuzzy C-Means (FCM) method. The aim of this research is to determine the optimal grouping results for grouping districts/cities in Kalimantan based on housing and environmental health indicator data using the FGK method based on the validity indices of Partition Coefficient (PC), Coefficient Entropy (CE), and Modified Partition Coefficient (MPC). The conclusion of this research is that the optimal number of groups for grouping districts/cities in Kalimantan based on housing and environmental health indicators is 6 groups with a PC index value of 0.8470; CE index of 0.3365; and the MPC index is 0.8164.*

1. PENDAHULUAN

Analisis multivariat merupakan metode untuk mengidentifikasi pengaruh antara beberapa variabel dan variabel lainnya sekaligus. Analisis multivariat terbagi menjadi teknik dependensi dan interdependensi. Analisis dependensi meneliti hubungan variabel terikat dan variabel bebas. Beberapa metode yang termasuk pada metode dependensi yaitu regresi berganda dan diskriminan. Metode interdependensi meneliti hubungan antar variabel tanpa mempertimbangkan jenis variabel yang terlibat. Beberapa metode yang termasuk pada metode interdependensi yaitu analisis faktor, *clustering*, dan *multidimensional scalling* (Simamora, 2005).

Pengelompokan (*clustering*) merupakan teknik dimana objek dikelompokkan menjadi dua kelompok atau lebih, satu kelompok akan berisi objek yang memiliki kemiripan dibandingkan dengan objek di kelompok lain berdasarkan informasi yang ada pada objek tersebut (Aldino dkk., 2021). Pada umumnya, pengelompokan dibagi menjadi pengelompokan hierarki dan non hierarki. Pada proses pengelompokan metode hierarki dan non hierarki, kelompok dibuat sehingga setiap objek akan ada di dalam satu kelompok. Namun, terdapat kondisi ketika objek terletak diantara dua kelompok atau lebih, sehingga digunakan metode *fuzzy clustering* yang menggunakan derajat keanggotaan *fuzzy* sebagai dasar pembobotan (Jang, Sun, & Mizutani, 1997).

Fuzzy clustering merupakan salah satu metode analisis kelompok dengan mempertimbangkan tingkat keanggotaan yang mencakup himpunan *fuzzy* sebagai dasar pembobot bagi pengelompokan. *Fuzzy clustering* adalah pengembangan dari metode *partitioning data* dengan pembobotan *fuzzy*. Keunggulan utama *fuzzy clustering* adalah dapat memberikan hasil pengelompokan untuk data-data yang tersebar tidak teratur, karena

jika terdapat suatu data yang penyebarannya tidak teratur maka terdapat kemungkinan suatu titik data mempunyai sifat atau karakteristik dari kelompok lain, sehingga perlu adanya pembobotan kecenderungan titik data terhadap suatu kelompok (Sanusi dkk., 2019).

Salah satu contoh teknik *fuzzy clustering* adalah *Fuzzy Gustafson-Kessel* (FGK). Pengelompokan FGK adalah hasil pengembangan dari teknik *Fuzzy C-Means* (FCM). Metode FCM menggunakan jarak *euclidean* yang hanya mengukur rata-rata dari sebuah kelompok dan memiliki ketidakmampuan untuk memecahkan masalah dengan variabel yang berkorelasi atau variabel yang memiliki skala berbeda. Oleh karena itu, Gustafson & Kessel (1979) mengusulkan metode FGK yang menggunakan jarak *mahalanobis* yang mampu mengatasi kekurangan dari jarak *euclidean* dengan mempertimbangkan kovarian yang memungkinkan untuk menyesuaikan variabel yang berkorelasi (Rousseeuw & Leroy, 1987).

Pada pengelompokan dengan *fuzzy clustering*, diperlukan hasil pengelompokan yang optimal. Hasil pengelompokan yang optimal dapat diperoleh dengan menggunakan kriteria ukuran validitas. Validasi kelompok mengacu pada cara mengukur ketepatan partisi dan apakah partisi yang dibuat sudah tepat. Contoh indeks validasi yaitu indeks *Modified Partition Coefficient* (MPC), *Classification Entropy* (CE), dan *Partition Coefficient* (PC).

Penerapan metode FGK dan penggunaan indeks validitas telah dibahas pada penelitian sebelumnya. (Putri & Rochmawati, 2021) melakukan pengelompokan terhadap *tweet mention* akun @gojekindonesia dan @GrabID dengan metode FGK dan indeks *Xie-Beni* yang menghasilkan 2 kelompok *tweets* tentang keluhan terhadap sistem aplikasi pada akun @gojekindonesia dan pada akun @GrabID memberikan hasil 2 kelompok *tweets* tentang anjuran belanja dengan memanfaatkan ovo dan promo. Penelitian yang dilakukan oleh (Paradilla dkk., 2023) yang membahas pengelompokan pada data isu kemiskinan dengan menggunakan algoritma FGK. Penelitian tersebut menghasilkan 2 kelompok optimal berdasarkan indeks PC terbesar sebesar 0,5363 dan indeks CE terkecil sebesar 0,6500. Selain pada data *tweet mention* dan isu kemiskinan, analisis FGK juga dapat diaplikasikan untuk mengelompokkan data indikator perumahan dan kesehatan lingkungan.

Menurut Purnama (2018), perumahan dan permukiman merupakan struktur yang mencakup pembangunan, pengelolaan perumahan, pengelolaan daerah permukiman, perawatan dan perbaikan, upaya untuk mencegah dan meningkatkan kualitas perumahan dan permukiman yang tidak layak, penyediaan lahan, pendanaan dan metode pembiayaan, serta partisipasi penduduk. UU No. 1 tahun 2011 menyatakan salah satu tujuan perumahan dan permukiman yakni menjamin terwujudnya tempat tinggal yang layak untuk dihuni dan dapat dijangkau oleh lingkungan yang berada dalam keadaan sehat, serasi, terencana, aman, terpadu, teratur, dan berkelanjutan.

Badan Pusat Statistik (BPS) mencatat bahwa pada tahun 2022, Provinsi Kalimantan Barat merupakan provinsi di Indonesia yang memiliki persentase rumah tangga yang menempati rumah milik sendiri mencapai lebih dari 90 persen. Sedangkan Provinsi Kalimantan Timur termasuk Provinsi di Indonesia dengan rata-rata pengeluaran per kapita tertinggi untuk sewa/kontrak. Angka ini didukung dengan tingginya rumah tangga yang menempati rumah sewa/kontrak. Selanjutnya, Provinsi Kalimantan Barat dan Kalimantan Tengah memiliki persentase akses terhadap sanitasi layak kurang dari 80 persen, yaitu Kalimantan Barat sebesar 77,41% dan Kalimantan Tengah sebesar 74,33%. Hal ini menunjukkan bahwa terdapat perbedaan kondisi perumahan dan kesehatan lingkungan provinsi di Pulau Kalimantan. Oleh karena itu, prioritas peningkatan masalah perumahan dan kesehatan lingkungan pada tiap daerah juga berbeda, sehingga diperlukan suatu metode untuk mengelompokkan setiap Kabupaten/Kota di Kalimantan yang memiliki karakteristik yang sama berdasarkan indikator perumahan dan kesehatan lingkungan.

Penelitian ini dilakukan untuk mengelompokkan indikator perumahan dan kesehatan lingkungan di Kalimantan, sehingga hasil pengelompokan pada penelitian ini diharapkan dapat membantu pemerintah dalam mengoptimalkan upaya kesehatan lingkungan sesuai dengan karakteristik yang dimiliki masing-masing kabupaten/kota pada tiap kelompok.

2. TINJAUAN PUSTAKA

Analisis multivariat digunakan untuk meneliti hubungan atau pengaruh terhadap lebih dari dua variabel dalam waktu yang sama. Analisis multivariat merupakan alternatif dari kekurangan analisis univariat yang hanya difokuskan pada pengamatan satu variabel saja. Analisis multivariat melibatkan pengumpulan data dari berbagai pengukuran yang dilakukan secara bersamaan. Sebagian besar teknik dalam analisis multivariat ini menggunakan operasi matriks dalam setiap perhitungan yang dilakukan. (Johnson & Wichern, 2007). Menurut (Wijaya & Budiman, 2016), analisis multivariat terbagi menjadi dua kategori besar, yaitu metode dependensi dan metode interdependensi. Metode dependensi digunakan sebagai alat untuk menjelaskan variabel terikat dengan memanfaatkan dua atau lebih variabel bebas, contohnya yaitu regresi linier berganda. Metode interdependensi dibedakan berdasarkan jenis masukan variabel yang diukur dengan besaran metrik ataupun non metrik. Jika variabel masukan diukur dengan skala metrik, analisis kelompok dapat digunakan. Jika variabel masukan diukur dengan skala non metrik, teknik *multidimensional scalling* dapat digunakan.

Menurut (Supranto, 2010), analisis kelompok termasuk hubungan interdependensi, sehingga tidak terdapat perbedaan antara variabel bebas dan variabel terikat. Berdasarkan hak tersebut, analisis kelompok tidak efektif untuk menentukan dampak variabel bebas pada variabel terikat, namun hanya menggolongkan berbagai objek ke dalam kelompok yang seragam. Analisis kelompok berguna untuk menggolongkan rangkaian objek ke dalam kelompok yang saling bebas, sehingga objek yang memiliki ciri-ciri yang sama atau serupa akan berada di dalam satu kelompok, sedangkan kelompok lainnya terdiri dari individu-individu yang cenderung tidak serupa satu sama lain. Menurut (Hair dkk., 2014), kelompok yang baik ditandai dengan memiliki ciri-ciri berikut:

1. Tingkat kemiripan yang tinggi pada tiap anggota satu kelompok (*within cluster*).
2. Perbedaan yang signifikan pada tiap kelompok (*between cluster*).

Terdapat dua metode pada analisis kelompok, yaitu metode hierarki dan metode non hierarki. Langkah awal pada analisis hierarki yaitu mengelompokkan dua atau lebih objek yang mempunyai kemiripan terdekat, lalu dilanjutkan ke objek yang memiliki kesamaan kedua. Contoh pengelompokan hierarki yaitu *complete linkage*, *average linkage*, dan *single linkage*. Langkah awal metode non hierarki yaitu menentukan jumlah kelompok. Contoh pengelompokan non hierarki adalah *K-Means*.

Istilah *fuzzy* pada kamus *Oxford* diartikan sebagai *blurred*, *indistinct*, *imprecisely defined*, *confused*, dan *vague*. Maksud dari sistem *fuzzy* bukan mengacu pada sistem yang tidak jelas definisinya, namun merupakan sistem yang dibangun menggunakan definisi yang jelas berdasarkan teori logika *fuzzy* (Naba, 2009). Menurut (Kusumadewi & Purnomo, 2004), terdapat dua kategori himpunan dalam logika *fuzzy*, yaitu himpunan *crisp* dan himpunan *fuzzy*. Himpunan *crisp* hanya memiliki 2 nilai keanggotaan, yaitu 0 dan 1, sedangkan pada himpunan *fuzzy* memiliki nilai keanggotaan diantara rentang 0 hingga 1.

Fuzzy clustering merupakan metode pengelompokan yang memperhitungkan tingkat keanggotaan himpunan *fuzzy* sebagai dasar pembobot. *Fuzzy clustering* adalah evolusi dari metode partisi data dengan pembobotan *fuzzy*. Kelebihan *fuzzy clustering* adalah kemampuannya untuk menciptakan hasil pengelompokan untuk objek yang tersebar tidak

teratur. Pembobotan diperlukan untuk meminimalisir objek yang menyebar tidak teratur akan memiliki karakter dari kelompok lain (Sanusi dkk., 2019).

Terdapat beberapa metode *fuzzy clustering*, salah satu diantaranya adalah *Fuzzy Gustafson-Kessel* (FGK). FGK pertama kali diperkenalkan oleh Gustafson dan Kessel pada tahun 1979. Menurut (Balasko dkk., 2007), FGK merupakan evolusi dari metode FCM. FCM menggunakan nilai pembentuk yang sama pada setiap iterasi, sedangkan FGK mampu mengubah fungsi perhitungan jarak yang diperbarui pada tiap iterasi menjadi fungsi jarak adaptif dengan matriks *fuzzy covariance*. Meskipun FGK lebih unggul daripada FCM, perhitungan matriks \mathbf{A}_k pada metode FGK tidak dapat diterapkan ketika matriks *fuzzy covariance* merupakan matriks singular (Babuska, Veen, & Kaymak, 2002). Algoritma FGK yaitu:

1. Menentukan:
 - a. Banyaknya kelompok (c) yang akan dibentuk sesuai dengan kebutuhan pengelompokan,
 - b. Nilai pangkat (m) yaitu parameter *fuzzy* yang mengontrol kemungkinan adanya irisan antar kelompok yang terbentuk.
 - c. Batas iterasi tertinggi (*MaksIter*) yaitu pengulangan yang akan berhenti jika nilai maksimal iterasi sudah tercapai.
 - d. *Error*/kesalahan terkecil yang diharapkan (ϵ) yaitu batasan nilai yang membuat pengulangan akan berakhir setelah didapatkan nilai *error* yang diharapkan.
 - e. Fungsi objektif awal $P_0 = 0$ yaitu suatu fungsi yang akan dioptimumkan (maksimum atau minimum), nilai 0 berarti untuk mendapatkan nilai minimum.
2. Membangkitkan bilangan acak α_{ik} menggunakan bantuan *software R* sebagai unsur dari matriks partisi awal \mathbf{U}_0 . Adapun nilai setiap bilangan acak α_{ik} berkisar antara 0 hingga 1 dan jumlah tiap unsur kolom pada satu baris pada matriks \mathbf{U}_0 harus berjumlah 1.

$$\mathbf{U}_0 = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,c} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,c} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,c} \end{bmatrix}$$

dimana $i = 1, 2, 3, \dots, n$ adalah jumlah data dan $k = 1, 2, \dots, c$ adalah jumlah kelompok.

3. Menghitung pusat kelompok awal setiap variabel menggunakan persamaan berikut.

$$v_{kj} = \frac{\sum_{i=1}^n ((\alpha_{ik})^m x_{ij})}{\sum_{i=1}^n (\alpha_{ik})^m} \quad (1)$$

4. Menghitung matriks kovarian setiap kelompok menggunakan persamaan berikut.

$$\mathbf{F}_k = \frac{\sum_{i=1}^n (\mathbf{x}_{ij} - \mathbf{v}_{kj})^T (\alpha_{ik})^m (\mathbf{x}_{ij} - \mathbf{v}_{kj})}{\sum_{i=1}^n (\alpha_{ik})^m} \quad (2)$$

5. Menghitung nilai jarak *mahalanobis* menggunakan persamaan berikut.

$$\mathbf{D}_k^2 = (\mathbf{x}_{ij} - \mathbf{v}_{kj}) \mathbf{A}_k (\mathbf{x}_{ij} - \mathbf{v}_{kj})^T \quad (3)$$

Adapun matriks \mathbf{A}_k dapat dihitung dengan persamaan sebagai berikut.

$$\mathbf{A}_k = \left(\det(\mathbf{F}_k)^{\frac{1}{n}} \right) (\mathbf{F}_k^{-1}) \quad (4)$$

Matriks yang terbentuk dari \mathbf{D}_k^2 merupakan matriks bujur sangkar dengan elemen yang ditampilkan sebagai berikut.

$$\mathbf{D}_k^2 = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix}$$

dimana diagonal utama pada matriks \mathbf{D}_k^2 ialah nilai yang akan digunakan untuk jarak mahalanobis yang disimbolkan sebagai berikut.

$$\mathbf{S}_{ik}^2 = \begin{bmatrix} d_{11} \\ d_{22} \\ \vdots \\ d_{nn} \end{bmatrix} = \begin{bmatrix} S_{1k}^2 \\ S_{2k}^2 \\ \vdots \\ S_{nk}^2 \end{bmatrix}$$

6. Menghitung fungsi objektif pada iterasi ke- t dengan rumus berikut.

$$P_t = \sum_{i=1}^n \sum_{k=1}^c (\alpha_{ik})^m (S_{ik}^2) \quad (5)$$

7. Menghitung perubahan matriks keanggotaan yang baru dengan elemen α_{ik} menggunakan persamaan:

$$\alpha_{ik} = \left[\left(\frac{S_{ik}}{\sum_{k=1}^c S_{ik}} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (6)$$

8. Mengulangi langkah ke-3 hingga langkah ke-7 hingga kondisi $|P_t - P_{t-1}| < \varepsilon$ atau $t > MaksIter$ terpenuhi.

Pada analisis kelompok, diperlukan suatu ukuran untuk menentukan jumlah kelompok optimal yang terbentuk. Perhitungan jumlah kelompok optimal dapat ditentukan menggunakan indeks validitas. Indeks validitas bertujuan untuk menilai kinerja indikator pengelompokan, menilai mutu setiap kelompok, serta menentukan jumlah kelompok yang efisien (Zhu, Wang, & Wang, 2019). Terdapat beberapa indeks validitas untuk mengetahui jumlah kelompok optimal, yaitu:

1. *Partition Coefficient* (PC)

Indeks PC menentukan banyaknya tumpang tindih tiap kelompok. Indeks PC dihitung dengan rumus berikut.

$$\alpha_{ik} = \left[\left(\frac{S_{ik}}{\sum_{k=1}^c S_{ik}} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (7)$$

Apabila nilai PC semakin tinggi, maka akan semakin baik hasil pengelompokannya

2. *Classification Entropy* (CE)

Indeks CE digunakan untuk menilai tingkat kekaburan partisi kelompok. Indeks ini dihitung dengan rumus berikut.

$$\alpha_{ik} = \left[\left(\frac{S_{ik}}{\sum_{k=1}^c S_{ik}} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (8)$$

Apabila nilai CE semakin rendah, maka akan semakin baik hasil pengelompokannya

3. *Modified Partition Coefficient* (PC)

Nilai PC dan CE cenderung berubah seiring dengan perubahan nilai c . Dave (1996) melakukan modifikasi nilai PC dan menjadi MPC untuk kecenderungan monotonik tersebut. Indeks ini dihitung dengan rumus sebagai berikut.

$$MPC(c) = 1 - \frac{c}{c-1} (1 - PC) \quad (9)$$

Apabila nilai MPC semakin tinggi, maka akan semakin baik hasil pengelompokannya

3. METODE PENELITIAN

Data yang dimanfaatkan yaitu data sekunder yang berasal dari Survei Sosial Ekonomi Nasional (Susenas) oleh Badan Pusat Statistik (BPS) pada Maret 2022. Variabel yang digunakan yaitu 6 variabel indikator perumahan dan Kesehatan lingkungan yang terdiri dari persentase rumah tangga yang memiliki akses terhadap layanan sumber air minum layak (X_1), persentase rumah tangga yang menggunakan sumber penerangan dari Listrik (X_2), persentase rumah tangga yang memiliki akses terhadap layanan sanitasi layak (X_3), persentase rumah tangga dengan status kepemilikan rumah milik sendiri (X_4), persentase rumah tangga yang sumber bahan bakar memasaknya menggunakan gas (X_5), dan persentase rumah tangga yang jenis lantai luasnya bukan tanah (X_6).

Pengolahan data dilakukan menggunakan *software R* dan *Microsoft Excel* dengan proses sebagai berikut:

1. Menginput data yang akan dikelompokkan, yaitu x_{ij} dengan $i = 1, 2, \dots, 56$ dan $j = 1, 2, \dots, 6$.
2. Menggambarkan variabel penelitian menggunakan nilai minimum, nilai maksimum, rata-rata, dan simpangan baku.
3. Mengelompokkan data menggunakan metode *Fuzzy Gustafson-Kessel*.
4. Menentukan kelompok optimal menggunakan indeks validitas PC, CE, dan MPC.
5. Menginterpretasikan hasil pengelompokan optimal.

4. HASIL DAN PEMBAHASAN

Langkah pertama penelitian yaitu menggambarkan variabel penelitian secara umum. Analisis statistika deskriptif digunakan untuk melihat gambaran secara umum pada 6 variabel indikator perumahan dan kesehatan lingkungan berdasarkan Kabupaten/Kota di Pulau Kalimantan pada tahun 2022. Adapun hasil analisis statistika deskriptif disajikan pada Tabel 1.

Tabel 1. Statistika Deskriptif

Variabel (X_j)	Minimum	Maksimum	Rata-Rata	Simpangan Baku
X_1	45,53	99,86	78,24	14,26
X_2	90,20	100,00	98,58	2,15
X_3	52,75	98,13	80,22	11,43
X_4	20,72	93,64	80,55	11,87
X_5	69,52	99,74	93,02	5,60
X_6	98,18	100,00	99,62	0,44

Selanjutnya dilakukan pengelompokan terhadap 6 variabel indikator perumahan dan kesehatan lingkungan menggunakan metode FGK dengan tahapan sebagai berikut:

1. Menentukan nilai parameter FGK. Penelitian ini menggunakan jumlah kelompok (c) yaitu 2,3,4,5, dan 6; nilai pangkat (m) sebesar 2; maksimum iterasi sebesar 10000; *error* sebesar 10^{-2} ; dan P_0 sebesar 0.
2. Membangkitkan bilangan acak menggunakan bantuan *software R* sebagai elemen matriks keanggotaan awal \mathbf{U}_0 , dengan $i = 1, 2, \dots, 56$ dan $k = 1, 2$. Nilai α_{ik} dibangkitkan acak dengan syarat elemen pada setiap baris berjumlah 1. Adapun nilai keanggotaan dengan 2 kelompok ditampilkan pada Tabel 2.

Tabel 2. Nilai Keanggotaan Awal dengan 2 Kelompok

No.	Kabupaten/Kota	Nilai Keanggotaan	
		α_{i1}	α_{i2}
1	Sambas	0,6435	0,3565
2	Bengkayang	0,1471	0,8529
\vdots	\vdots	\vdots	\vdots
56	Kota Tarakan	0,4340	0,5660

Berdasarkan nilai keanggotaan awal dengan 2 kelompok pada Tabel 2, diperoleh matriks keanggotaan awal \mathbf{U}_0 berukuran 56×2 berikut.

$$\mathbf{U}_0 = \begin{bmatrix} 0,6435 & 0,3565 \\ 0,1471 & 0,8529 \\ \vdots & \vdots \\ 0,4340 & 0,5660 \end{bmatrix}_{(56 \times 2)}$$

- Setelah bilangan acak dibangkitkan, langkah selanjutnya yaitu menghitung pusat kelompok awal menggunakan persamaan (1). Hasil perhitungan pusat kelompok awal dengan jumlah kelompok 2 disajikan pada Tabel 3.

Tabel 3. Pusat Kelompok Awal dengan 2 Kelompok

Variabel	Pusat Kelompok	
	v_{1j}	v_{2j}
X_1	78,6039	78,8762
X_2	98,9315	98,1167
X_3	81,5862	79,6228
X_4	78,9344	81,6260
X_5	92,3277	93,5552
X_6	99,7061	99,5642

Adapun matriks yang terbentuk dari pusat kelompok awal dengan 2 kelompok adalah matriks \mathbf{V} berukuran 2×6 sebagai berikut:

$$\mathbf{V} = \begin{bmatrix} 78,6039 & 98,9315 & 81,5862 & 78,9344 & 92,3277 & 99,7061 \\ 78,8762 & 98,1167 & 79,6228 & 81,6260 & 93,5552 & 99,5642 \end{bmatrix}$$

- Setelah diperoleh nilai pusat kelompok awal dengan 2 kelompok, selanjutnya yaitu menghitung nilai matriks kovarian kelompok menggunakan persamaan (2). Hasil perhitungan matriks kovarian kelompok pertama dan kedua yaitu berupa matriks yang berukuran 6×6 sebagai berikut:

$$\mathbf{F}_1 = \begin{bmatrix} 218,0743 & 4,7630 & 29,7548 & -90,4522 & 21,8329 & 0,0279 \\ 4,7630 & 1,9959 & 4,3314 & -0,1899 & 3,6419 & -0,0201 \\ 29,7548 & 4,3314 & 122,9895 & -2,9173 & 3,2584 & 2,0475 \\ -90,4522 & -0,1899 & -2,9173 & 190,6527 & 14,6359 & 0,0444 \\ 21,8329 & 3,6419 & 3,2584 & 14,6359 & 47,0869 & -0,5033 \\ 0,0279 & -0,0201 & 2,0475 & 0,0444 & -0,5033 & 0,1782 \end{bmatrix}$$

$$\mathbf{F}_2 = \begin{bmatrix} 189,2908 & 18,4204 & 86,1397 & -50,6656 & 13,2344 & 1,5740 \\ 18,4204 & 7,1612 & 14,6693 & -10,0348 & -0,9458 & 0,1272 \\ 86,1397 & 14,6693 & 119,2061 & -42,2380 & 8,2049 & 1,1506 \\ -50,6656 & -10,0348 & -42,2380 & 97,9288 & -2,8171 & -0,0882 \\ 13,2344 & -0,9458 & 8,2049 & -2,8171 & 20,9749 & -0,1344 \\ 1,5740 & 0,1272 & 1,1506 & -0,0882 & -0,1344 & 0,1773 \end{bmatrix}$$

- Langkah selanjutnya dalam pengelompokan dengan menggunakan metode FGK yaitu menghitung nilai jarak mahalanobis. Langkah awal untuk perhitungan jarak

mahalanobis yaitu menghitung nilai A_k menggunakan persamaan (4). Hasil perhitungan dari matriks A_k pada kelompok pertama dan kedua yaitu berupa matriks yang berukuran 6×6 sebagai berikut:

$$A_1 = \begin{bmatrix} 0,0091 & -0,0087 & -0,0017 & 0,0046 & -0,0048 & 0,0022 \\ -0,0087 & 0,8858 & -0,0327 & 0,0007 & -0,0591 & 0,3095 \\ -0,0017 & -0,0327 & 0,0156 & -0,0006 & 0,0005 & -0,1816 \\ 0,0046 & 0,0007 & -0,0006 & 0,0098 & -0,0053 & -0,0114 \\ -0,0048 & -0,0591 & 0,0005 & -0,0053 & 0,0384 & 0,0986 \\ 0,0022 & 0,3095 & -0,1816 & -0,0114 & 0,0986 & 10,0508 \\ 0,0072 & -0,0110 & -0,0027 & 0,0013 & -0,0041 & -0,0408 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} -0,0110 & 0,1688 & -0,0126 & 0,0069 & 0,0209 & 0,0772 \\ -0,0027 & -0,0126 & 0,0110 & 0,0019 & -0,0032 & -0,0398 \\ 0,0013 & 0,0069 & 0,0019 & 0,0098 & -0,0000 & -0,0240 \\ -0,0041 & 0,0209 & -0,0032 & -0,0000 & 0,0406 & 0,0723 \\ -0,0408 & 0,0772 & -0,0398 & -0,0240 & 0,0723 & 4,7986 \end{bmatrix}$$

Setelah diperoleh nilai A_1 dan A_2 , selanjutnya dihitung nilai jarak mahalanobis menggunakan persamaan (3). Hasil perhitungan jarak mahalanobis yaitu berupa matriks berukuran 56×56 . Elemen diagonal utama yang terbentuk pada matriks jarak D_1^2 dan D_2^2 merupakan jarak mahalanobis pada iterasi pertama ($t = 1$). Adapun jarak mahalanobis dengan 2 kelompok pada tiap data disajikan pada Tabel 4.

Tabel 4. Nilai Jarak Mahalanobis dengan 2 Kelompok

No.	Kabupaten/Kota	Nilai Jarak	
		S_{i1}^2	S_{i2}^2
1	Sambas	7,8951	4,3211
2	Bengkayang	13,7650	1,9861
⋮	⋮	⋮	⋮
56	Kota Tarakan	4,0817	4,0128

- Setelah diperoleh nilai jarak mahalanobis pada kelompok pertama dan kelompok kedua, selanjutnya menghitung fungsi objektif iterasi pertama ($t = 1$) dengan bantuan persamaan (5). Nilai fungsi objektif pada $t = 1$ adalah sebesar 209,4483. Nilai fungsi objektif awal (P_0) adalah sebesar 0, sehingga nilai $|P_1 - P_0|$ adalah sebesar 209,4483 atau lebih dari $\varepsilon = 10^{-2}$ dan proses pengelompokan dilanjutkan ke iterasi kedua. Pada proses pengelompokan iterasi kedua, langkah awal yang dilakukan yaitu menghitung perubahan matriks keanggotaan.
- Pada iterasi kedua, perubahan matriks keanggotaan dihitung menggunakan persamaan (6). Hasil perhitungan nilai perubahan matriks keanggotaan secara lengkap disajikan pada Tabel 5.

Tabel 5. Nilai Keanggotaan yang Diperbarui dengan 2 Kelompok

No.	Kabupaten/Kota	Nilai Keanggotaan	
		α_{i1}	α_{i2}
1	Sambas	0,3537	0,6463
2	Bengkayang	0,1261	0,8739
⋮	⋮	⋮	⋮
56	Sanggau	0,2762	0,7238

Setelah diperoleh nilai perubahan matriks keanggotaan, selanjutnya dihitung kembali nilai pusat kelompok, matriks kovarian kelompok, nilai jarak mahalanobis, fungsi

objektif, dan perubahan matriks keanggotaan dengan kelompok sebanyak 2. Iterasi berhenti ketika nilai $|P_t - P_{t-1}|$ kurang dari 0,01 atau t lebih dari 10000. Adapun pada pengelompokan dengan 2 kelompok ini berhenti ketika iterasi ke-23. Hasil akhir nilai keanggotaan dari 56 Kabupaten/Kota di Kalimantan disajikan pada Tabel 6.

Tabel 6. Nilai Keanggotaan Pada Iterasi Terakhir dengan 2 Kelompok

No	Kabupaten/Kota	Nilai Keanggotaan		Kelompok yang Diikuti
		α_{i1}	α_{i2}	
1	Sambas	0,6921	0,3079	1
2	Bengkayang	0,0145	0,9855	2
⋮	⋮	⋮	⋮	⋮
56	Sanggau	0,0542	0,9458	2

Berdasarkan nilai keanggotaan pada iterasi terakhir, anggota kelompok dengan jumlah kelompok sebanyak 2 pada metode FGK dapat dilihat pada Tabel 7.

Tabel 7. Anggota Kelompok dengan Jumlah Kelompok 2 pada metode FGK

Kelompok	Kabupaten/Kota
1	Sambas, Mempawah, Kayong Utara, Kota Pontianak, Kota Singkawang, Kotawaringin Barat, Kotawaringin Timur, Seruyan, Pulang Pisau, Gunung Mas, Kota Palangka Raya, Tanah Laut, Kotabaru, Tapin, Hulu Sungai Tengah, Hulu Sungai Utara, Tabalong, Tanah Bumbu, Balangan, Kota Banjarmasin, Kota Banjar Baru, Paser, Kutai Barat, Kutai Kartanegara, Berau, Penajam Paser Utara, Mahakam Ulu, Kota Balikpapan, Kota Samarinda, Kota Bontang, Bulungan, Tana Tidung, Nunukan, Kota Tarakan
2	Bengkayang, Landak, Sanggau, Ketapang, Sintang, Kapuas Hulu, Sekadau, Melawi, Kubu Raya, Kapuas, Barito Selatan, Barito Utara, Sukamara, Lamandau, Katingan, Barito Timur, Murung Raya, Banjar, Barito Kuala, Hulu Sungai Selatan, Kutai Timur, Malinau.

8. Setelah diperoleh hasil pengelompokan dengan jumlah kelompok sebanyak 2, perhitungan dilanjutkan dengan langkah yang sama pada kelompok 3,4,5, dan 6. Adapun hasil pengelompokan pada kelompok 2,3,4,5, dan 6 disajikan pada Tabel 8.

Tabel 8. Hasil Pengelompokan dengan Metode FGK

Banyak Kelompok (c)	Banyak Iterasi	Fungsi Objektif
2	23	2.525,4120
3	109	1.608,6000
4	74	1.062,7700
5	162	775,2133
6	92	550,5877

Selanjutnya dilakukan pemilihan kelompok optimal berdasarkan indek PC, CE, dan MPC menggunakan persamaan (7), persamaan (8), dan persamaan (9). Hasil perhitungan indeks validitas yaitu sebagai berikut:

Tabel 9. Nilai Indeks Validitas PC, CE, dan MPC

Banyak Kelompok (c)	Indeks PC	Indeks CE	Indeks MPC
2	0,7566	0,3860	0,5133
3	0,7070	0,5141	0,5605

(*) : Nilai paling tinggi

Berdasarkan Tabel 10, diketahui bahwa kelompok 1 tidak memiliki nilai tertinggi pada keenam indikator perumahan dan kesehatan lingkungan, tetapi berada pada posisi kedua pada persentase rumah tangga yang memiliki akses terhadap sumber air minum layak, persentase rumah tangga yang memiliki akses terhadap layanan sanitasi layak, dan persentase rumah tangga yang jenis lantai terluasnya bukan tanah serta memiliki nilai persentase rumah tangga dengan status kepemilikan rumah milik sendiri paling rendah dibandingkan kelompok lain, sehingga pemerintah perlu mengembangkan suatu kebijakan untuk mengatasinya, misalnya yaitu dengan memberikan bantuan pembiayaan perumahan. Kelompok 2 adalah kelompok dengan persentase rumah tangga yang memiliki akses terhadap layanan sumber air minum layak dan persentase rumah tangga yang sumber bahan bakar memasaknya menggunakan gas tertinggi dibandingkan dengan kelompok lain serta memiliki nilai persentase rumah tangga dengan status kepemilikan rumah milik sendiri yang cukup rendah. Kelompok 3 adalah kelompok dengan persentase rumah tangga yang menggunakan sumber penerangan dari listrik dan persentase rumah tangga yang memiliki akses terhadap layanan sanitasi layak tertinggi dibandingkan kelompok lain serta memiliki nilai persentase rumah tangga yang memiliki akses terhadap layanan sumber air minum layak paling rendah dibandingkan dengan kelompok lain, sehingga kebijakan yang dapat diterapkan oleh pemerintah misalnya dengan membangun dan merehabilitasi sumber-sumber air bersih dan meningkatkan infrastruktur pengolahan dan distribusi air. Kelompok 4 tidak memiliki nilai tertinggi pada keenam indikator perumahan dan kesehatan lingkungan, tetapi berada pada posisi kedua pada persentase rumah tangga yang sumber bahan bakar memasaknya menggunakan gas serta memiliki nilai persentase rumah tangga yang memiliki akses terhadap sanitasi layak dan persentase rumah tangga yang jenis lantai terluasnya bukan tanah paling rendah dibandingkan dengan kelompok lain, sehingga upaya yang dapat dilakukan pemerintah untuk mengatasi hal tersebut misalnya dengan menyediakan pendanaan untuk sanitasi dan menyediakan infrastruktur sanitasi yang memadai. Kelompok 5 adalah kelompok dengan persentase persentase rumah tangga dengan status kepemilikan rumah milik sendiri dan persentase rumah tangga yang lantai terluasnya bukan tanah tertinggi dibandingkan dengan kelompok lain. Kelompok 6 tidak memiliki nilai tertinggi pada keenam indikator perumahan dan kesehatan lingkungan, tetapi berada pada posisi kedua pada persentase rumah tangga yang bahan bakar memasaknya menggunakan gas.

5. KESIMPULAN

Berdasarkan analisis dan pembahasan, jumlah kelompok optimal untuk mengelompokkan Kabupaten/Kota di Kalimantan sesuai dengan indikator perumahan dan kesehatan lingkungan dengan menggunakan metode *Fuzzy Gustafson-Kessel* adalah sebanyak 6 kelompok dengan nilai indeks PC sebesar 0,8470; indeks CE sebesar 0,3365; dan indeks MPC sebesar 0,8164.

Hasil pengelompokan optimal Kabupaten/Kota di Kalimantan sesuai dengan indikator perumahan dan kesehatan lingkungan dengan menggunakan metode *Fuzzy Gustafson-Kessel* adalah sebanyak 6 kelompok dengan anggota kelompok 1 yaitu Sambas, Kota Singkawang, Kotawaringin Timur, Kota Palangka Raya, Kotabaru, Kutai Barat, Kutai Kartanegara, Kota Balikpapan, Malinau, dan Tana Tidung. Anggota kelompok 2 yaitu Sanggau, Kayong Utara, Kota Pontianak, Seruyan, Kutai Timur, Kota Samarinda, Kota Bontang, dan Kota Tarakan. Anggota kelompok 3 yaitu Gunung Mas, Tanah Laut, Hulu Sungai Selatan, Tanah Bumbu, Kota Banjar Baru, Paser, Berau, Mahakam Ulu, Bulungan, dan Nunukan. Anggota kelompok 4 yaitu Kotawaringin Barat, Kapuas, Barito Selatan, Katingan, Pulang Pisau, Banjar, Barito

Kuala, Tapin, dan Tabalong. Anggota kelompok 5 yaitu Mempawah, Ketapang, Kapuas Hulu, Kubu Raya, Barito Utara, Barito Timur, Murung Raya, Hulu Sungai Tengah, Hulu Sungai Utara, dan Balangan. Anggota kelompok 6 yaitu Bengkayang, Landak, Sintang, Sekadau, Melawi, Sukamara, Lamandau, Kota Banjarmasin, dan Penajam Paser Utara.

Berdasarkan penelitian yang telah dilakukan, saran untuk penelitian selanjutnya yaitu dapat membandingkan metode FGK dengan metode pengelompokan lainnya seperti *K-Medians*, *Subtractive Fuzzy C-Means*, dan dilihat perbedaan tingkat akurasi.

DAFTAR PUSTAKA

- Aldino, A. A., Darwis, D., Prastowo, A. T., & Sujana, C. (2021). Implementation of K-Means Algorithm for Clustering Corn Planting Feasibility Area in South Lampung Regency. *Journal of Physics*, 1751(1).
- Badan Pusat Statistik. (2022). *Indikator Perumahan dan Kesehatan Lingkungan*. Badan Pusat Statistik.
- Balasko, B., Abonyoo, J., & Feil, B. (2007). *Fuzzy Clustering and Data Analysis Toolbox For Use With Matlab*. University of Vezprem.
- Gustafson, D., & Kessel, W. (1979). *Fuzzy Clustering with a Fuzzy Covariance Matrix*. IEEE Conference Publication.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate Data Analysis 7th Edition*. Person Education, Inc.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis Sixth Edition*. Pearson Education, Inc.
- Kusumadewi, S., & Purnomo, H. (2004). *Aplikasi Logika Fuzzy untuk Pendukung Keputusan*. Graha Ilmu.
- Naba, A. (2009). *Belajar Cepat Fuzzy Logic Menggunakan MATLAB*. Penerbit ANDI.
- Paradilla, Y. S., Hayati, M. N., & Sifriyani. (2023). Implementation of the Fuzzy Gustafson-Kessel Method on Grouping District/Cities in Kalimantan Island Based on Poverty Issues Factors. *Jurnal Ilmu Matematika dan Terapan*, 17(1), 125–134.
- Purnama, S. G. (2018). *Dasar Kesehatan Lingkungan*. Universitas Udayana.
- Putri, F. A. A., & Rochmawati, N. (2021). Penerapan Algoritma Fuzzy Gustafson-Kessel Untuk Clustering Tweets Mention Akun Go-Jek dan Grab Indonesia. *Journal of Informatics and Computer Science*, 2(3), 204–212.
- Rousseeuw, P. J., & Leroy, A. M. (1987). *Robust Regression and Outlier Detection*. John Wiley & Sons.
- Sanusi, W., Zaky, A., & Afni, B. N. (2019). Analisis Fuzzy C-Means dan Penerapannya dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Faktor-Faktor Penyebab Gizi Buruk. *Journal of Mathematics, Computations, and Statistics*, 2(1), 47–54.
- Simamora, B. (2005). *Analisis Multivariat Pemasaran*. PT Gramedia Pustaka.
- Supranto, J. (2010). *Analisis Multivariat Arti dan Interpretasi*. Rineka Cipta.
- Wijaya, T., & Budiman, S. (2016). *Analisis Multivariat Untuk Penelitian Manajemen*. Pohon Cahaya.