VARIANSI: Journal of Statistics and Its Application on Teaching and Research **ISSN 2684-7590 (Online)**
Vol. 4 No. 2 (2022), 95-108

DOI: *10.35580/variansiunm25*

Metode *Subtractive Fuzzy C-Means* (SFCM) dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Indikator Kemiskinan

Nurul Imania Kalla[1], Suwardi Annas[2], Muhammad Fahmuddin S[3]

Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Makassar, Indonesia

Keywords: Poverty, Subtractive Fuzzy C-Means, Partition Coefficient

Abstract:

The percentage of poor people in South Sulawesi Province has increased in the last 3 years. It is necessary to handle poverty that by adjusting to the poverty level of each district. It is necessary to have a description of socio-economic conditions by grouping regency/city based on poverty indicators. The purpose of this grouping is determining the poverty level of each district so that it can be used as material for evaluating poverty problems. The clustering method often used in research was Fuzzy C-Means (FCM). In the development, FCM was combined with the Subtractive Clustering (SC) method to obtain the hybrid Subtractive Fuzzy C-Means (SFCM) method. The SFCM method has the advantage of speed, iteration, resulting in a data partition that is more stable and accurate when it is compared to the previous method. On this research, the SFCM method was applied with 4 variables from the poverty indicator data for South Sulawesi Province in 2021. The results of the analysis show that based on the Partition Coefficient Index (PC), the percentage of poverty in South Sulawesi Province was divided into three clusters, namely the low, medium and high poverty percentage clusters. Clusters of low poverty percentage consist of Bulukumba, Bantaeng, Takalar, Gowa, Soppeng, Wajo, Sidenreng Rappang, East Luwu, Makassar City, Parepare City, and Palopo City. Clusters of medium poverty percentage consist of Sinjai, Maros, Barru, and Enrekang. Meanwhile, Clusters of high poverty percentage consist of the Selayar Islands, Jeneponto, Pangkajene and Islands, Bone, Pinrang, Luwu, Tana Toraja, North Luwu, and North Toraja.

1. Pendahuluan

Kemiskinan merupakan indikator yang sangat penting untuk melihat keberhasilan pembangunan ekonomi suatu negara dan merupakan salah satu masalah dalam kesejahteraan yang harus dituntaskan atau paling tidak mengalami penurunan tingkat kemiskinan (Fadli, 2018).

Salah satu daerah di Indonesia yang tidak lepas dari persoalan kemiskinan ialah Provinsi Sulawesi Selatan. Berdasarkan data BPS menunjukkan bahwa pada tahun 2019 persentase penduduk miskin di Provinsi Sulawesi Selatan sebesar 8,69%, pada tahun 2020 meningkat menjadi 8,72%, dan tahun 2021 kembali meningkat dengan persentase mencapai 8,78%, persentase penduduk miskin di Provinsi Sulawesi Selatan terus mengalami peningkatan sejak 3 tahun terakhir (Badan Pusat Statistik, 2021). Untuk mengatasi hal tersebut diperlukan gambaran kondisi sosial ekonomi di Sulawesi Selatan yaitu dengan mengelompokkan kabupaten/kota di Provinsi Sulawesi Selatan berdasarkan indikator kemiskinan.

* Corresponding author.
E-mail address: nurulimaniak@gmail.com

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Dalam statistika, metode yang digunakan untuk mengelompokkan variabel atau objek adalah metode *clustering* (Soemartini & Supartini, 2017). Adapun teknik pengelompokan yang sering digunakan ialah *Fuzzy C-Means. Fuzzy C-Means* adalah suatu algoritma untuk mengelompokkan data, yang dimana keberadaan tiap-tiap data dalam suatu klaster ditentukan oleh derajat keanggotaannya (Setiawan dkk, 2017).

Metode FCM memiliki beberapa kelemahan, antara lain membutuhkan banyaknya kelompok dan matriks keanggotaan kelompok yang ditetapkan sebelumnya (Le & Altman, 2011). Menurut Haqiqi dan Kurniawan (2015), biasanya matriks keanggotaan kelompok awal diinisialisasikan secara acak yang menyebabkan metode FCM memiliki masalah inkonsistensi. Alternatif metode pengelompokan lainnya yang dapat digunakan jika jumlah kelompok tidak diketahui sebelumnya adalah metode *Substractive Clustering* (SC). SC memperoleh hasil yang lebih konsisten dibandingkan dengan FCM (Bataineh dkk, 2011). Selain itu, SC memiliki kecepatan yang lebih baik dibandingkan FCM, namun SC memiliki akurasi yang lebih rendah dibandingkan dengan FCM (Hossen dkk, 2011). Untuk memperantarai kekurangan dan kelebihan kedua metode tersebut diusulkan sebuah metode baru yang merupakan penggabungan (*hybrid*) dari keduanya yang disebut *Subtractive Fuzzy C-Means* (SFCM). Metode ini digunakan oleh Liu dkk (2003) dalam penelitiannya dan menyimpulkan bahwa SFCM secara umum memberikan solusi yang lebih baik dibandingkan dengan FCM serta memberikan tingkat kecepatan yang lebih tinggi dalam hal konvergensi fungsi objektif. Kelebihan metode SFCM lainnya dengan menggunakan indeks validitas *Partition Coefficient* (PC) menyimpulkan bahwa SFCM dapat meningkatkan kecepatan, mengurangi jumlah iterasi, menghasilkan partisi data yang lebih stabil, dan lebih akurat (Hossen dkk, 2011).

Penelitian ini menggunakan metode SFCM pada data kemiskinan di Sulawesi Selatan. Hasil pengelompokan terbaik dipilih berdasarkan indeks validitas *Partition Coefficent* (PC). Tujuan dari pengelompokan tersebut sebagai informasi dan bahan evaluasi pemerintah dalam mengatasi kemiskinan sesuai dengan tingkat kemiskinan masing-masing daerah.

2. Tinjauan Pustaka

2.1. Analisis Pengelompokan

Secara umum ada dua metode di dalam analisis pengelompokan yaitu (Awaliah, 2018):

1. Metode hierarki (*hierarchical clustering methods*)

Metode *hierarchical clustering* digunakan saat jumlah *cluster* yang diinginkan belum diketahui. Adapun metode *hierarchical clustering* yaitu *Single Lingkage, Complete Linkage, Average Linkage, Ward's Method,* dan *Centroid Method* (Santosa, 2007).

2. Metode tak berhierarki (*non hierarchical clustering methods*)

Pada metode *non hierarchical clustering* banyaknya kelompok telah diketahui sebelumnya dan biasanya metode ini dipakai dalam pengelompokan data yang berukuran besar. Adapun metode *non hierarchical clustering* yaitu metode *C-Means* (CM), *Fu-zzy C-Means* (FCM), *Adaptive C-Means*, dan *K-Medoids* (Santosa, 2007).

2.2. Ukuran Jarak

Terdapat beberapa cara dalam mengukur jarak, namun yang digunakan dalam penelitian ini adalah jarak *Euclidean*. Jarak ini merupakan jarak yang paling umum digunakan dan dapat digunakan apabila semua peubahnya berskala kontinu dan harus memenuhi asumsi bahwa peubah-peubah yang diamati tidak berkorelasi dan antar peubah memiliki satuan yang sama (Windasari, 2020). Dalam metode ini, pengukuran jarak dilakukan dengan menghitung akar kuadrat dari penjumlahan kuadrat selisih dai nilai masing-masing peubah. Jarak *Euclidean* dapat dirumuskan sebagai berikut (Mongi, 2015):

$$d^2{}_{ij} = \sum_{k=1}^{p}(x_{ik} - x_{jk})^2 \tag{2.1}$$

dimana:

$d^2{}_{ij}$ = jarak kuadrat *Euclidean* antara objek ke-i dengan objek ke-j
p = jumlah peubah yang digunakan
x_{ik} = nilai dari objek ke-i pada peubah ke-k
x_{jk} = nilai dari objek ke-j pada peubah ke-k

2.3. Metode Fuzzy C-Means

Metode *Fuzzy C-Means* (FCM) merupakan pengembangan dari metode *C-Means Clustering* yang mengandung fungsi keanggotaan. Fungsi keanggotaan adalah suatu kurva yang menunjukkan pemetaan titik-titik input data ke dalam nilai keanggotaan (Kusumadewi & Hartati, 2006).

Algoritma *Fuzzy C-Means* adalah sebagai berikut (Kusumadewi dan Purnomo, 2010):

1. Input data yang akan dikelompok *X*, berupa matriks berukuran $n \times m$ ($n =$ jumlah sampel data, m = atribut setiap data). X_{ij} = data sampel ke-*i* (i = 1, 2, …, n), atribut ke-*j* (j = 1, 2, …, m).
2. Tentukan:
   Banyak kelompok = c;
   Fuzzifier = m;
   Maksimum iterasi = MaxIter;
   Perubahan nilai fungsi objektif terkecil yang diharapkan = ε;
   Fungsi objektif awal = P_0=0;
   Iterasi awal = t = 1
3. Membangkitkan bilangan random μ_{ik} dengan *i* merupakan banyak data dan *k* merupakan banyak kelompok sebagai elemen-elemen awal matriks keanggotaan awal U.
4. Menghitung pusat kelompok ke-*i* dengan persamaan:

$$p_i = \frac{\sum_{k=1}^{N}(\mu_{ik})^m X_k}{\sum_{k=1}^{N}(\mu_{ik})^m} \tag{2.2}$$

   dimana:
   μ_{ik} = nilai keanggotaan objek ke-*k* dengan pusat kelompok ke-*i*
   X_k = objek data ke-*k*
   N = banyaknya objek penelitian
   m = *fuzzifier*
5. Menghitung fungsi objektif pada iterasi ke-*t* dengan persamaan:

$$J\,(P,U,X,c,m) = \sum_{i=1}^{c}\sum_{k=1}^{N}(\mu_{ik})^m d_{ik}^2(x_k,p_i) \tag{2.3}$$

   dimana:
   c = banyak kelompok yang diinginkan
   N = banyaknya objek penelitian
   μ_{ik} = nilai keanggotaan objek ke-*k* dengan pusat kelompok ke-*i* yang merupakan bagian dari matriks U
   m = *fuzzifier*
   $d_{ik}^2(x_k,p_i)$ = jarak antara vektor pengamatan ke-*k* dengan pusat kelompok ke-*i*
6. Menghitung perubahan matriks keanggotaan dengan persamaan:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{d_{ik}^2}{d_{jk}^2}\right)^{\frac{1}{m-1}}} \tag{2.4}$$

   dimana μ_{ik} adalah nilai keanggotaan objek ke-*k* dengan pusat kelompok ke-*i*, d_{ik}^2 adalah jarak antara objek ke-*k* dengan pusat kelompok ke-*i*, d_{jk}^2 adalah jarak antara objek ke-*k* dengan pusat kelompok ke-*j*, dan m adalah *fuzzifier*.
7. Cek kondisi berhenti
   - Jika $|J_t - J_{t-1}| < \varepsilon \; atau \; t > MaxIter$ maka berhenti;
   - Jika tidak : $t = t + 1$, ulangi langkah ke-4.

2.4. Metode Subtractive Clustering

Metode *Subtractive Clustering* merupakan salah satu metode yang dikembangkan dalam analisis *fuzzy clustering*. Metode ini juga merupakan modifikasi dari *Mountain Method* yang diperkenalkan oleh Yager dan Filev pada tahun 1992. Dimisalkan terdapat *n* buah titik data $\{x_1, x_2, \ldots, x_n\}$ dalam sebuah ruang berdimensi *M*. Setiap titik data dipertimbangkan sebagai kandidat pusat kelompok yang nilai potensinya dihitung dengan rumus (Chiu, 1994):

$$P_i = \sum_{j=1}^{n} e^{-\frac{4}{r_a^2}\|x_i - x_j\|^2} \tag{2.5}$$

P_i adalah nilai potensi untuk titik data ke-*i* atau potensi dari nilai x_i, $\| \|$ menotasikan jarak *Euclidean*, dan r_a adalah konstanta positif. Konstanta r_a merupakan radius atau jari-jari yang mendefinisikan sebuah lingkungan tetangga. Data yang memiliki potensi yang tinggi adalah data yang memiliki jumlah tetangga paling banyak.

Setelah nilai potensi dari semua titik data dihitung, titik data dengan nilai potensi paling tinggi dipilih sebagai pusat kelompok pertama. Misalkan x_1^* adalah data yang terpilih sebagai pusat kelompok pertama dan P_1^* sebagai nilai potensinya. Nilai potensi untuk setiap titik data diperbaharui dengan rumus (Chiu, 1994):

$$P_i = P_i - P_1^* e^{-\frac{4}{r_\beta^2}\|x_i - x_1^*\|^2} \tag{2.6}$$

dimana r_β merupakan konstanta positif. Setelah nilai potensi tiap titik data diperbaharui, titik data dengan nilai potensi terbesar dipilih sebagai pusat kelompok kedua. Selanjutnya nilai potensi tiap titik data kembali diperbaharui. Proses ini akan terus berlanjut sampai diperoleh jumlah kelompok yang cukup. Pada metode ini jumlah *cluster* belum diketahui (Kusumadewi dan Purnomo, 2010).

Algoritma *Subtractive Clustering* adalah sebagai berikut (Suerni dkk, 2020):

1. Penginputan data (xi) dengan sampel ke-*i* (i = 1, 2, … , n) dan atribut ke-*j* (j = 1, 2, … , m).
2. Penetapan nilai jari-jari (r), *Squash factor* (q), *accept ratio* (ar), *reject ratio* (rr), nilai minimum dan maksimum data dari masing-masing atribut.
3. Penormalisasian setiap data dengan persamaan:
$$x_{ij_{norm}} = \frac{x_{ij} - x_{j_{min}}}{x_{j_{max}} - x_{j_{min}}} \tag{2.7}$$
4. Penentuan potensi awal tiap-tiap titik data
   a. i = 1
   b. pengerjaan hingga i = n
      - $T_j = x_{kj_{norm}}$; j = 1, 2, …, m
      - Perhitungan:
$$Dist_{ij} = \frac{T_j - x_{ij_{norm}}}{r}; i, k = 1,2, \dots, n; j = 1,2, \dots, m \tag{2.8}$$
      - Penentuan potensi awal tiap data dengan persamaan:
$$\text{Jika } m = 1, \text{ maka } D_k = \sum_{i=1}^{n} e^{-4(Dist_{ij}^2)} \tag{2.9}$$
$$\text{Jika } m > 1, \text{ maka } D_k = \sum_{i=1}^{n} e^{-4(\sum_{j=1}^{m} Dist_{ij}^2)} \tag{2.10}$$
5. Pencarian titik dengan potensi tertinggi
   M = max[$D_k|k$ = 1, 2, … , n]; untuk iterasi ke-1
   Z = max[$D_k|k$ = 1, 2, … , n]; untuk iterasi ke-2, 3, dan seterusnya
6. Penentuan pusat *cluster* dan kurangi potensinya terhadap titik-titik di sekitarnya dengan cara menetapkan vektor V sebagai calon pusat *cluster* dan menghitung nilai *ratio*.
$$R = \frac{Z}{M} \tag{2.11}$$
   khusus pada iterasi pertama nilai Z = M.
   Setelah nilai *ratio* diperoleh, ada 3 keadaan yang dapat terjadi:
   1) Jika nilai *ratio* > *accept ratio*, calon pusat *cluster* dapat diterima sebagai pusat *cluster* baru dan ditulis *Cl*. Prosedur kerja yang harus dilakukan selanjutnya setelah mendapat *cluster* baru adalah mengurangi potensi titik-titik data yang lain dengan cara:
      a. $S_{ij} = \frac{C_{ij_{norm}} - x_{ij_{norm}}}{r * q}; i = 1,2, \dots, n; j = 1,2, \dots, m$ (2.12)
      b. $D_{c_{li_{norm}}} = M \times e^{-4\left[\sum_{j=1}^{m} (S_{ij})^2\right]}; i = 1,2, \dots, n$ (2.13)
      c. $D_i^t = D_i^{t-1} - D_{c_{li_{norm}}}$ (2.14)
   2) Jika nilai *ratio* < *accept ratio* dan nilai *ratio* > *reject ratio*, calon baru akan diterima jika keberadaannya cukup jauh dari pusat *cluster* yang telah ada. Prosedur dalam keadaan ini:

a. Misalkan *Md* = -1;
b. Mengerjakan untuk *l* = 1 sampai *l* = *p*; *p* = banyaknya *cluster*

$$Sd_l = \sum_{j=1}^{m} \left(\frac{V_j - C_{lj_{norm}}}{r} \right)^2 ; j = 1,2, \dots, m \tag{2.15}$$

Jika (*Md* < *0*) atau (*Sd*<*Md*), maka *Md* = *Sdl*
Jika (*Sdl*>Md*)*, maka *Md* tidak berubah

$$Mds = \sqrt{Md} \tag{2.16}$$

dengan *Mds* adalah jarak terdekat data calon pusat *cluster* dengan pusat *cluster*.
Jika (r*atio* + *Mds*) ≥ 1, calon pusat *cluster* diterima sebagai pusat *cluster* baru.
Jika (*ratio* + *Mds*) < 1, calon pusat *cluster* tidak diterima dan tidak akan dipertimbangkan kembali sebagai pusat *cluster* dan potensi data tersebut diset menjadi 0.

3) Jika nilai *ratio* < *accept ratio* dan nilai *ratio* < *reject ratio*, sudah tidak ada lagi calon pusat *cluster* baru dan iterasi dihentikan.

7. Pengembalian pusat *cluster* dari bentuk ternormalisasi ke bentuk semula (denormalisasi)

$$C_{lj} = C_{lj_{norm}} * \left(x_{max_j} - x_{min_j} \right) + x_{min_j} \tag{2.17}$$

8. Hitung nilai fungsi keanggotaan kelompok

$$\sigma_j = r_j * (x_{max_j} - x_{min_j})/\sqrt{8} \tag{2.18}$$

2.5. Metode Subtractive Fuzzy C-Means

Subtractive Fuzzy C-Means (SFCM) merupakan penggabungan antara metode pengelompokan *Subtractive Clustering* (SC) dan *Fuzzy C-Means* (FCM). SC digunakan untuk menentukan jumlah kelompok dan matriks keanggotaan awal FCM sehingga inisialisasi secara random tidak perlu dilakukan. Perbedaan FCM dan SFCM ialah pada FCM matriks keanggotaan kelompok awal diinisialisasikan secara acak. Sedangkan pada SFCM matriks keanggotaan awal dihitung berdasarkan pusat kelompok yang didapatkan melalui metode SC.

Secara garis besar algoritma metode SFCM adalah sebagai berikut (Kurniawan dan Haqiqi, 2015):

1. Menghitung potensi awal tiap titik data dengan rumus

$$P_i = \sum_{j=1}^{n} e^{-\frac{4}{r_a^2} \|x_i - x_j\|^2} \tag{2.19}$$

2. Menetapkan titik data dengan potensi paling tinggi sebagai pusat kelompok pertama dan P_1^* sebagai nilai potensinya
3. Perbaharui potensi tiap titik data dengan rumus

$$P_i = P_i - P_1^* e^{-\frac{4}{r_\beta^2} \|x_i - x_1^*\|^2} \tag{2.20}$$

4. Jika rasio potensi titik data tertinggi terbaru dengan potensi titik data tertinggi sebelumnya lebih besar dari *accept ratio* maka titik data dengan potensi tertinggi terbaru ditetapkan menjadi pusat kelompok dan tahap 3 diulangi kembali. Jika rasio kurang dari *reject ratio* maka iterasi dihentikan berlanjut ke tahap 5.
5. Menghitung matriks keanggotaan awal FCM berdasarkan pusat kelompok yanng didapatkan melalui metode SC menggunakan persamaan

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}^2}{d_{jk}^2} \right)^{\frac{1}{m-1}}} \tag{2.21}$$

Jika terdapat jarak d_{ik}^2 yang bernilai nol maka μ_{ik} (nilai keanggotaan data *k* pada kelompok ke-*i*) akan bernilai 1 dan μ_{ik} (nilai keanggotaan data *k* pada kelompok lainnya) akan bernilai 0 (Liu dkk, 2003).

6. Melakukan tahapan 4 sampai dengan 7 pada algoritma FCM hingga ditemukannya matriks keanggotaan dan pusat kelompok terakhir.

2.6. Indeks Validitas

Indeks validitas yang digunakan dalam penelitian ini adalah *Partition Coefficient* (PC). Indeks ini mengukur jumlah *overlapping* antar kelompok. Persamaan indeks PC oleh Bezdek sebagai berikut:

$$PC\ (c) = \frac{1}{N}\sum_{i=1}^{C}\sum_{k=1}^{N}\mu_{ik}^{2} \tag{2.22}$$

dengan *N* adalah banyak objek penelitian, *c* adalah banyak kelompok, dan μ_{ik} adalah nilai keanggotaan objek ke-k dengan pusat kelompok ke-i. Indeks ini memiliki rentang 1/c sampai 1. Jumlah kelompok yang optimal ditunjukkan oleh nilai PC yang paling besar.

2.7. Kemiskinan

Badan Pusat Statistik mendefinisikan kemiskinan sebagai ketidakmampuan untuk memenuhi standar minimum kebutuhan yang meliputi makanan maupun *non* makanan. Adapun ukuran tingkat kemiskinan, dirumuskan Foster-Greer-Thorbecke, sebagai berikut:

$$P_a = \frac{1}{n}\sum_{i=1}^{q}\left[\frac{z - y_i}{z}\right]^a \tag{2.23}$$

dimana:

a = 0, 1, 2
z = garis kemiskinan
y_i = rata-rata pengeluaran per kapita sebulan penduduk yang berada di bawah garis kemiskinan (i = 1, 2, 3,...,q), $y_i < z$
q = banyaknya penduduk yang berada di bawah garis kemiskinan
n = jumlah penduduk

Jika a = 0, diperoleh *Head Count Index* (P_0), jika a = 1 diperoleh Indeks Kedalaman Kemiskinan (*Poverty Gap Index* – P_1) dan jika a = 2 disebut Indeks Keparahan Kemiskinan (*Poverty Severity Index* – P_2).

3. Metode Penelitian

3.1. Jenis Penelitian

Jenis penelitian yang digunakan adalah penelitian kuantitatif. Data numerik yang digunakan adalah data indikator kemiskinan di Provinsi Sulawesi Selatan tahun 2021 yang dianalisis menggunakan metode *Subtractive Fuzzy C-Means* (SFCM).

3.2. Jenis dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder yang dipublikasikan oleh Badan Pusat Statistik (BPS) Provinsi Sulawesi Selatan yaitu data indikator kemiskinan tahun 2021.

3.3. Definisi Operasional Peubah

Indikator kemiskinan yang digunakan berdasarkan Badan Pusat Statistik, yaitu: Persentase Penduduk Miskin (P_0), Indeks Kedalaman Kemiskinan (P_1), Indeks Keparahan Kemiskinan (P_2), dan Garis Kemiskinan (GK).

3.4. Prosedur Penelitian

Adapun prosedur penelitian yang akan digunakan adalah sebagai berikut:

1. Merumuskan permasalahan yang akan diteliti.
2. Mengkaji metode dan mencari referensi baik dari buku maupun jurnal terkait.
3. Mengumpulkan data sekunder yang akan digunakan pada penelitian.

4. Memaparkan gambaran umum permasalahan yang terjadi.
5. Melakukan pengelompokan dengan metode *Subtractive Fuzzy C-Means* (SFCM).
6. Menentukan jumlah kelompok optimum menggunakan indeks validitas *Partition Coefficient* (PC).
7. Menyusun laporan hasil penelitian.
8. Membuat kesimpulan.

3.5. Teknik Analisis Data

Langkah-langkah analisis data yang akan dilakukan adalah sebagai berikut:

1. Merumuskan masalah.
2. Melakukan analisis deskriptif terhadap masing-masing variabel.
3. Mengecek kelengkapan data dan menstandarisasi data.
4. Melakukan analisis pengelompokan dengan metode *Subtractive Fuzzy C-Means* (SFCM).
5. Melakukan perhitungan nilai indeks validitas menggunakan indeks validitas *Partition Coefficient* (PC).
6. Melakukan interpretasi hasil pengelompokan.
7. Kesimpulan.

4. Hasil dan Pembahasan

4.1 Analisis Deskriptif Indikator Kemiskinan

Sebelum dilakukan analisis data lebih lanjut, terlebih dahulu akan dibahas mengenai eksplorasi data secara deskriptif yang berkaitan dengan indikator kemiskinan Provinsi Sulawesi Selatan tahun 2021.

Tabel 4.1 Statistika Deskriptif Indikator Kemiskinan Sulawesi Selatan Tahun 2021

Indikator	Minimum	Maximum	Rata-rata	Simpangan Baku
P_0	4,82	14,28	9,47	2,91
P_1	0,64	2,78	1,43	0,50
P_2	0,13	0,78	0,33	0,15
GK	341484	475444	373436	30031

4.2 Standarisasi Data

Data yang digunakan dalam penelitian ini yakni data dengan satuan berbeda sehingga proses pengelompokan tidak bisa langsung dilakukan karena dapat mengakibatkan perhitungan pada analisis pengelompokan menjadi tidak valid. Maka dari itu, perlu dilakukan proses standarisasi dengan melakukan transformasi pada data asli sebelum dianalisis lebih lanjut.

Tabel 4.2 Hasil Nilai Standarisasi Data Indikator Kemiskinan Provinsi Sulawesi Selatan Tahun 2021

Kabupaten/Kota	P_0	P_1	P_2	GK
Kepulauan Selayar	1,0229	1,3131	0,8696	1,1087
Bulukumba	-0,6993	-0,8460	-0,8586	-0,1470
Bantaeng	-0,0200	-0,3904	-0,3933	-0,3361
Jeneponto	1,6507	0,4217	-0,2603	0,0323
Takalar	-0,4180	-0,8063	-1,0579	0,2374
Gowa	-0,6615	-0,0933	0,2049	1,1654
Sinjai	-0,2156	0,0454	-0,1939	-0,6975
Maros	0,0349	0,3623	0,5373	1,7050
Pangkajene dan Kepulauan	1,6507	2,6798	2,9966	-0,1778
Barru	-0,2704	0,1840	0,4044	-0,7163
Bone	0,3608	0,4812	0,6038	-0,4415
Soppeng	-0,6650	-0,3706	-0,3268	-1,0640
Wajo	-1,0320	-1,2619	-1,1909	-0,4081
Sidenreng Rappang	-1,5192	-1,5590	-1,3238	-0,4277
Pinrang	-0,2258	0,6792	1,1355	-0,9172

Enrekang	1,0297	-0,1527	-0,5927	-0,7089
Luwu	1,0503	0,6792	0,8696	-0,4499
Tana Toraja	0,9611	0,4614	0,2049	-0,7731
Luwu Utara	1,4139	1,2735	1,0690	-0,1572
Luwu Timur	-0,8674	-0,6677	-0,4597	-0,0496
Toraja Utara	0,8651	0,7981	0,4043	-0,8267
Kota Makassar	-1,5946	-1,3213	-1,1244	3,3968
Kota Parepare	-1,3957	-1,4006	-1,3238	0,3342
Kota Palopo	-0,4557	-0,5092	-0,1939	0,3187

4.3 Pengelompokan Menggunakan Metode SFCM

1) Penentuan Jumlah Kelompok Optimum

Langkah awal dalam penerapan SFCM menentukan nilai *accept ratio* sebesar 0,5, *reject ratio* sebesar 0,1, *fuzzifier* sebesar 2, iterasi maksimum sebanyak 1000 kali dan ε sebesar 10^{-5}. Sedangkan nilai radius (r) berbeda-beda karena jarak yang digunakan dalam menentukan jumlah kelompok. Jika nilai jarak (r) diperbesar, maka jumlah kelompok akan semakin kecil, begitupun sebaliknya. Dalam penelitian ini, digunakan r $\in$ {2,1; 2,2; 2,3; 2,4; 2,5}. Adapun indeks validitas PC yang dihasilkan untuk menentukan jumlah kelompok optimum seperti pada Tabel 4.3.

Tabel 4.3 Indeks Validitas PC Hasil Pengelompokan SFCM

R	Jumlah Kelompok	Banyaknya Iterasi	Fungsi Objektif	Indeks PC
2,5	3	27	28,45	0,374
2,4	4	96	21,23	0,282
2,3	5	41	16,95	0,229
2,2	15	43	3,96	0,145
2,1	18	117	3,16	0,143

Dari Tabel 4.3 pengelompokan yang dilakukan diperoleh jumlah kelompok, hasil indeks *Partition Coefficient* (PC), dan nilai fungsi objektif selama iterasi. Dalam penelitian ini, didapatkan indeks *Partition Coefficient* (PC) optimum adalah 0,374 pada radius 2,5, dengan jumlah kelompok yaitu 3 dan proses iterasinya berhenti pada iterasi ke-27. Nilai fungsi objektif pada iterasi terakhir yang diperoleh adalah 28,45. Banyaknya anggota kelompok pertama sampai dengan kelompok ketiga berturut-turut adalah 11, 4, dan 9.

2) Nilai Pusat Kelompok

Pada iterasi terakhir (iterasi ke-27), pusat kelompok V_{kj} yang dihasilkan dengan k = 1,2,3,4 dan j = 1,2,3 adalah:

a. Nilai Hasil Standarisasi

$$V_{kj} = \begin{pmatrix} -0{,}5032 & -0{,}5189 & -0{,}4574 & 0{,}0955 \\ 0{,}1104 & 0{,}1313 & 0{,}1143 & -0{,}0133 \\ 0{,}4513 & 0{,}4324 & 0{,}3884 & -0{,}1497 \end{pmatrix}$$

b. Nilai Indikator Kemiskinan Tahun 2021

$$V_{kj} = \begin{pmatrix} 8{,}00 & 1{,}17 & 0{,}26 & 376302{,}9 \\ 9{,}79 & 1{,}49 & 0{,}35 & 373036{,}5 \\ 10{,}78 & 1{,}65 & 0{,}39 & 368939{,}3 \end{pmatrix}$$

3) Derajat Keanggotaan Kabupaten/Kota untuk Setiap Kelompok pada Iterasi Terakhir (iterasi ke-27).

Dari derajat keanggotaan setiap kabupaten/kota pada iterasi terakhir dapat dilihat pada Tabel 4.4 dan diperoleh informasi mengenai kecenderungan kabupaten/kota untuk masuk ke kelompok mana. Derajat keanggotaan terbesar menunjukkan bahwa kecenderungan tertinggi kabupaten/kota untuk masuk menjadi anggota tersebut.

Tabel 4.4 Derajat Keanggotaan Kabupaten/Kota Untuk Setiap Kelompok Pada Iterasi Terakhir Menggunakan Metode SFCM

Wilayah Sulawesi Selatan	Derajat Keanggotaan (μ) Untuk Setiap Kelompok Pada Iterasi Terakhir			Kelompok
	μ_{i1}	μ_{i2}	μ_{i3}	
Kepulauan Selayar	0,241026	0,348245	0,410730	3
Bulukumba	0,602049	0,226704	0,171246	1
Bantaeng	0,424379	0,349034	0,226587	1
Jeneponto	0,239323	0,350235	0,410442	3
Takalar	0,572035	0,243608	0,184357	1
Gowa	0,373875	0,349861	0,276264	1
Sinjai	0,310684	0,396004	0,293312	2
Maros	0,299719	0,361686	0,338595	2
Pangkajene dan Kepulauan	0,272343	0,340414	0,387243	3
Barru	0,243457	0,398948	0,357595	2
Bone	0,125048	0,285124	0,589828	3
Soppeng	0,404994	0,327802	0,267204	1
Wajo	0,493812	0,278443	0,227745	1
Sidenreng Rappang	0,460002	0,292679	0,247318	1
Pinrang	0,235886	0,352472	0,411642	3
Enrekang	0,281343	0,361306	0,357351	2
Luwu	0,179679	0,313153	0,507169	3
Tana Toraja	0,190492	0,331633	0,477875	3
Luwu Utara	0,215551	0,331010	0,453438	3
Luwu Timur	0,657562	0,198692	0,143746	1
Toraja Utara	0,187666	0,322579	0,489755	3
Kota Makassar	0,381223	0,324574	0,294203	1
Kota Parepare	0,480496	0,284274	0,235229	1
Kota Palopo	0,628101	0,226431	0,145468	1

Hasil selengkapnya pengelompokan 24 kabupaten/kota ke dalam 3 kelompok dapat dilihat pada Tabel 4.5 yaitu:

Tabel 4.5 Hasil Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Menggunakan Metode SFCM

Kelompok	Anggota Kelompok
1	Bulukumba, Bantaeng, Takalar, Gowa, Soppeng, Wajo, Sidenreng Rappang, Luwu Timur, Kota Makassar, Kota Parepare, dan Kota Palopo.
2	Sinjai, Maros, Barru, dan Enrekang.
3	Kepulauan Selayar, Jeneponto, Pangkajene dan Kepulauan, Bone, Pinrang, Luwu, Tana Toraja, Luwu Utara, dan Toraja Utara.

4.4 Pembahasan

Untuk melihat karakteristik kabupaten/kota berdasarkan hasil pengelompokan yang terbentuk maka dari seluruh objek penelitian sebanyak 24 kabupaten/kota di Provinsi Sulawesi Selatan diambil rata-rata dari masing-masing variabel indikator kemiskinan yaitu P_0, P_1, P_2, dan GK ($\bar{X}$). Selanjutnya masing-masing kelompok diambil nilai rata-rata dari setiap variabelnya pula ($\bar{X}_c$). Sehingga diperoleh deskriptif kelompok setiap variabel yang diberikan pada Tabel 4.6.

Tabel 4.6 Deskriptif Kelompok Berdasarkan Rata-Rata

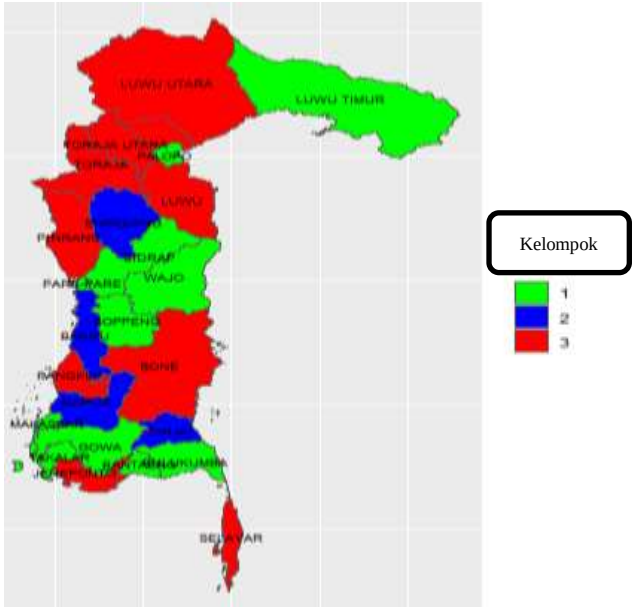
Variabel	$\bar{X}$	$\bar{X}_c$		
		C_1	C_2	C_3
P_0	9,47	6,99	9,89	12,3
P_1	1,43	1,00	1,48	1,92
P_2	0,33	0,22	0,33	0,46
GK	373436	381681	370300	364753

Selanjutnya, pada setiap variabel dalam kelompok diberi tanda, jika ($\bar{X} \geq \bar{X}_c$) maka diberi tanda positif (+), sedangkan jika ($\bar{X} < \bar{X}_c$) maka diberi tanda negatif (-). Karakteristik tiap kelompok yang diberikan pada Tabel 4.7.

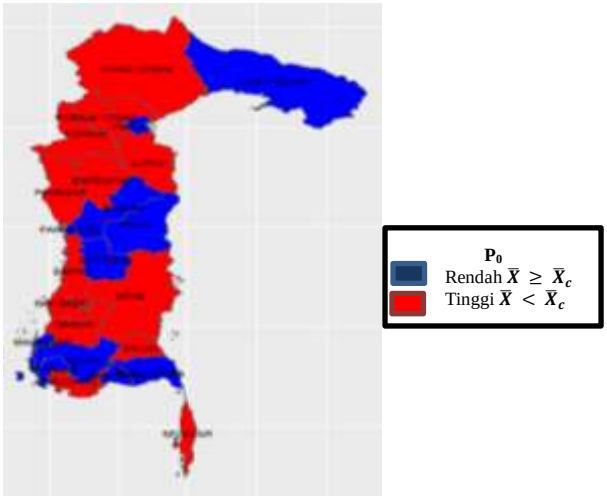
Tabel 4.7 Karakteristik Kelompok Berdasarkan Rata-Rata

Kelompok (C)	P_0	P_1	P_2	GK
1	+	+	+	-
2	-	-	+	+
3	-	-	-	+

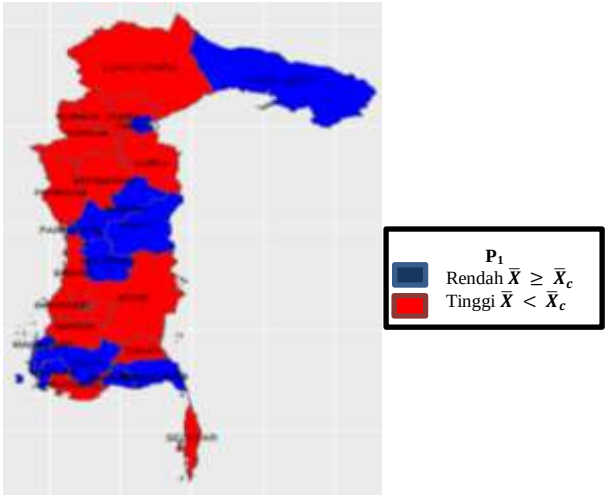
Berdasarkan Tabel 4.7 terlihat bahwa pada kelompok 1 variabel yang perlu mendapat perhatian lebih adalah variabel garis kemiskinan. Pada kelompok 2 variabel yang perlu mendapat perhatian lebih ialah variabel persentase penduduk miskin dan indeks kedalaman kemiskinan. Sementara untuk kelompok 3 variabel yang perlu mendapat perhatian lebih ialah variabel persentase penduduk miskin, indeks kedalaman kemiskinan dan indeks keparahan kemiskinan.



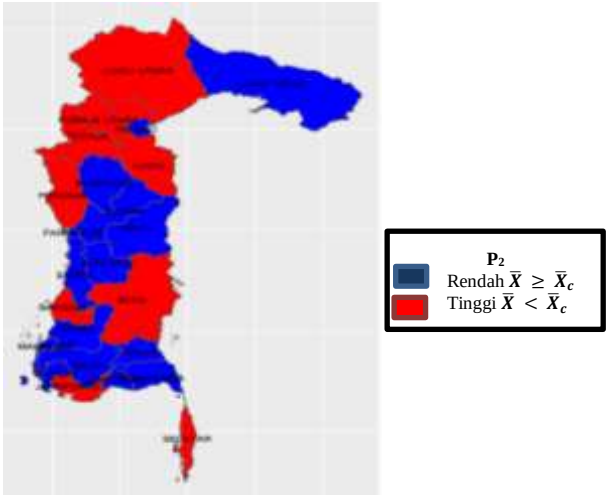
Gambar 4.1 Peta Provinsi Sulawesi Selatan Berdasarkan Persentase Kemiskinan Tahun 2021



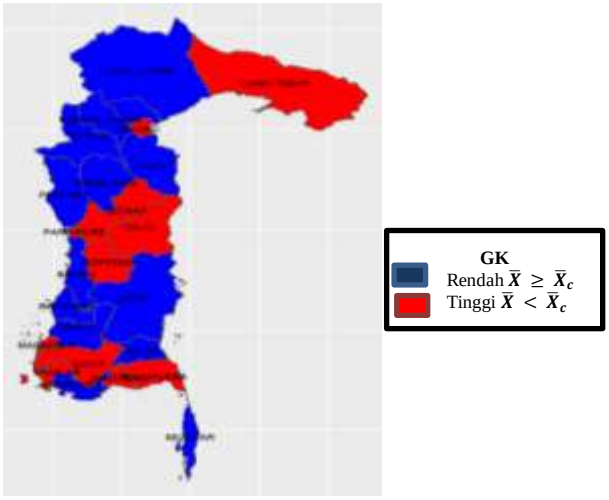
Gambar 4.2 Peta Provinsi Sulawesi Selatan Berdasarkan Indikator Persentase Penduduk Miskin Tahun 2021



Gambar 4.3 Peta Provinsi Sulawesi Selatan Berdasarkan Indikator Indeks Kedalaman Kemiskinan Tahun 2021



Gambar 4.4 Peta Provinsi Sulawesi Selatan Berdasarkan Indikator Indeks Keparahan Kemiskinan Tahun 2021



Gambar 4.5 Peta Provinsi Sulawesi Selatan Berdasarkan Indikator Garis Kemiskinan Tahun 2021

5. Kesimpulan

Berdasarkan hasil dan pembahasan yang dilakukan untuk mengelompokan kabupaten/kota di Provinsi Sulawesi Selatan berdasarkan Indikator Kemiskinan diambil kesimpulan:

1. Provinsi Sulawesi Selatan tahun 2021 memiliki rata-rata persentase penduduk miskin, rata-rata indeks kedalaman kemiskinan, rata-rata indeks keparahan kemiskinan dan rata-rata garis kemiskinan berturut-turut sebanyak 9,47, 1,43, 0,33, dan 373436.
2. Untuk penentuan jumlah kelompok yang optimal berdasarkan indeks validitas *Partition Coefficient* disimpulkan bahwa jumlah kelompok yang paling optimum adalah sebesar 3 kelompok yakni kelompok persentase kemiskinan rendah, sedang dan tinggi.
3. Berdasarkan hasil karakteristik setiap kelompok yang dibandingkan dengan nilai rata-rata di Provinsi Sulawesi Selatan disimpulkan bahwa kelompok 1 dengan persentase kemiskinan rendah meliputi Bulukumba, Bantaeng, Takalar, Gowa, Soppeng, Wajo, Sidenreng Rappang, Luwu Timur, Kota Makassar, Kota Parepare, dan Kota Palopo. Kelompok 2 dengan persentase kemiskinan sedang meliputi Sinjai, Maros, Barru, dan Enrekang. Sedangkan, kelompok 3 dengan persentase kemiskinan tinggi meliputi Kepulauan Selayar, Jeneponto, Pangkajene dan Kepulauan, Bone, Pinrang, Luwu, Tana Toraja, Luwu Utara, dan Toraja Utara.

References

Awaliah, R. (2018). Analisis Clustering untuk Mengelompokkan Tingkat Kesejahteraan Kabupaten/Kota Berdasarkan Sosial Ekonomi Rumah Tangga di Wilayah Provinsi Sulawesi Selatan. *Skripsi*. Makassar: Universitas Islam Negeri (UIN) Alauddin Makassar.

Badan Pusat Statistik. (2021). *Data dan Informasi Kemiskinan Kabupaten/Kota 2021*.

Bataineh, K. M., Naji, M., & Saqer, M. (2011). A Comparison Study Between Various Fuzzy Clustering Algorithms. *Jourdan Journal of Mechanical and Industrial Engineering, 5*(4), 335-343.

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The Fuzzy C-Means Clustering Algorithm. *Computers and Geoscience Printed in USA*, *10*(2), 191-203.

Chiu, S.L. (1994). Fuzzy Model Indentification Based on Cluster Estimation. *Journal of Intelligent and Fuzzy Systems, 2*, 267-278.

Fadli, Muzzammil. (2018). Clustering pada Data Kemiskinan di Indonesia dengan Metode Subtractive Fuzzy C-Means (SFCM). *Skripsi*. Surakarta : Universitas Sebelas Maret.

Haqiqi, B. N. & Kurniawan, R. (2015). Analisis Perbandingan Metode Fuzzy C-Means dan Subtractive Fuzzy C-Means. *Media Statistika, 8*(2), 59-67.

Hossen, J., Rahman, A., Sayeed, S., Samsuddin, K., & Rokhani, F. (2011). A Modified Hybrid Fuzzy Clustering Algorithm for Data Patitions. *Australian Journal of Basic and Applied Sciences, (*5)8. 674-681.

Kurniawan, R. & Haqiqi, B. N. (2015). Pengelompokan Menggunakan Metode Subtractive Fuzzy C-Means (SFCM), Studi Kasus Demam Berdarah di Jawa Timur. *Statistika, 3*(2), 22-30.

Kusumadewi, S., & Hartati, S. (2006). *Neuro Fuzzy: Integrasi Sistem Fuzzy & Jaringan Syaraf.* Yogyakarta: Graha Ilmu.

Kusumadewi, S., & Purnomo, H. (2010). *Aplikasi Logika Fuzzy untuk Pendukung Keputusan Edisi 2.* Yogyakarta: Graha Ilmu.

Le, T. & Altman, T. (2011). A New Initialization Method for The Fuzzy C-Means Algorithm Using Fuzzy Subtractive Clustering. *International Conference on Information and Knowledge Engineering.* USA: Las Vegas.

Liu, W. Y., Xiao, C. J., Wang, B. W., Shi, Y., & Fang, S. F. (2003). Study on Combining Subtractive Clustering with Fuzzy C-Means Clustering. *In International Conference on Machine Learning & Cybernetic*, *5*, 2659-2662.

Mongi, C. E. (2015). Penggunaan Analisis Two Step Clustering untuk Data Campuran. *JdC, 4*(1).

Santosa, B. (2007). *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis.* Yogyakarta: Graha Ilmu.

Setiawan, A., Susanto, B., & Mahatma, T. (2017). Inflation Data Clustering of Some Cities in Indonesia. *International Conference on Mathematics: Education, Theory and Application*, 1-8.

Soemartini, & Supartini, E. (2017). Analisis K-Means Cluster untuk Pengelompokan Kabupaten/Kota di Jawa Barat Berdasarkan Indikator Masyarakat. *Konferensi Nasional Penelitian Matematika dan Pembelajarannya II (KNPMP II)*, 144-154.

Suerni, W., Hayati, M, N., & Goejantoro, R. (2020). Penerapan Metode Subtractive Fuzzy C-Means pada Tingkat Partisipasi Pendidikan Jenjang Sekolah Menengah Atas/Sederajat di Kabupaten/Kota Pulau Kalimantan Tahun 2018. *Journal of Statistics and Its Applications. 2*(2). 63-74.

Windasari, R. (2020). Analisis Cluster Hirarki Metode Average Linkage Berdasarkan Jumlah Kriminalitas di Indonesia Tahun 2019. *Skripsi.* Malang: Universitas Islam Negeri Maulana Malik Ibrahim Malang.