

PENGGUNAAN METODE *POSSIBILISTIC FUZZY C-MEANS* UNTUK PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI SULAWESI SELATAN BERDASARKAN JUMLAH TENAGA KESEHATAN



AHMAD FIRDAUS IRIANTO
H051191007



PROGRAM STUDI STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024

**PENGGUNAAN METODE *POSSIBILISTIC FUZZY C-MEANS* UNTUK
PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI SULAWESI
SELATAN BERDASARKAN JUMLAH TENAGA KESEHATAN**

**AHMAD FIRDAUS IRIANTO
H051191007**



**PROGRAM STUDI STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024**

**PENGUNAAN METODE *POSSIBILISTIC FUZZY C-MEANS* UNTUK
PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI SULAWESI
SELATAN BERDASARKAN JUMLAH TENAGA KESEHATAN**

AHMAD FIRDAUS IRIANTO

H051191007



Diajukan sebagai salah satu syarat untuk mencapai gelar Sarjana Statistika

Program Studi Statistika

pada

**PROGRAM STUDI STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024**

SKRIPSI

**PENGUNAAN METODE *POSSIBILISTIC FUZZY*
C-MEANS UNTUK PENGELOMPOKAN KABUPATEN/KOTA DI
PROVINSI SULAWESI SELATAN BERDASARKAN JUMLAH
TENAGA KESEHATAN**

yang disusun dan diajukan oleh

AHMAD FIRDAUS IRIANTO
H051191007

Telah dipertahankan di depan Panitia Ujian Sarjana Statistika pada 11 Juni 2024 dan dinyatakan telah memenuhi syarat kelulusan

pada

Program Studi Statistika
Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Hasanuddin
Makassar

Mengesahkan:
Pembimbing Utama,



Anisa, S.Si., M.Si.
NIP. 19730227 199802 2 001

Mengetahui:
Ketua Program Studi,



Dr. (Ir.) Hamidah Amiyati, S.Si., M.Si.
NIP. 19770808 200501 2 002

**PERNYATAAN KEASLIAN SKRIPSI
DAN PELIMPAHAN HAK CIPTA**

Dengan ini saya menyatakan bahwa, skripsi berjudul "Penggunaan Metode Possibilistic Fuzzy C-Means Untuk Pengelompokan Kabupaten/Kota Di Provinsi Sulawesi Selatan Berdasarkan Jumlah Tenaga Kesehatan" adalah benar karya saya dengan arahan dari pembimbing skripsi saya (Anisa, S.Si., M.Si.). Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka skripsi ini. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan skripsi ini adalah karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut berdasarkan aturan yang berlaku.

Dengan ini saya melimpahkan hak cipta (hak ekonomis) dari karya tulis saya berupa skripsi ini kepada Universitas Hasanuddin.

Makassar, 11 Juni 2024

A 10,000 Rupiah Indonesian banknote is shown, partially obscured by a handwritten signature in black ink. The banknote features the Garuda Pancasila emblem and the text 'REPUBLIK INDONESIA', '10000', and 'METRA TEMPER'. The serial number 'C19C0ALX199101760' is visible at the bottom of the note.

Ahmad Firdaus Irianto
NIM H051191007

UCAPAN TERIMA KASIH

Puji dan syukur penulis panjatkan kepada Allah *Subhanahu Wa Ta'ala* atas segala limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini. Shalawat dan salam senantiasa tercurahkan kepada baginda Rasulullah *Shallallahu 'Alaihi Wa sallam*, yang telah membawa kita dari zaman kegelapan menuju zaman yang terang benderang. *Alhamdulillahilahi robbil'aalamiin*, berkat rahmat dan kemudahan yang diberikan oleh Allah *Subhanahu Wa Ta'ala*, penelitian ini dapat terlaksana dan terselesaikan atas bimbingan, diskusi dan arahan dari Ibu **Anisa, S.Si., M.Si.** selaku pembimbing sekaligus penasehat akademik yang dengan penuh kesabaran telah meluangkan waktu dan pemikirannya. Ucapan terima kasih juga penulis sampaikan kepada Ibu **Prof. Dr. Georgina Maria Tinungki, M.Si.** dan Bapak **Siswanto, S.Si., M.Si.** atas saran dan kritikan yang berharga serta waktu yang telah diberikan kepada penulis. Terima kasih juga kepada pimpinan Universitas Hasanuddin, Ketua Departemen Statistika, para dosen dan staff yang telah memberikan ilmu dan fasilitas kepada penulis.

Ucapan terima kasih serta penghargaan setinggi-tingginya penulis haturkan kepada orang tua, Abah Almarhum **Drs. Irianto Sulaiman** dan Ibunda **Sitti Maryam Bachtiar Bachtiar**, dan adik **Ahmad Furqan Irianto** yang telah memberikan dukungan, pengorbanan, kasih sayang serta doa yang tak henti-hentinya dipanjatkan kepada penulis.

Penulis juga berterima kasih kepada teman-teman **KM FMIPA**, khususnya **Akbar, Syahril, Ageng, Usamah, Kak Jalil, Kak Snuf, Kak Ail, Rahmat, Alif, Mahdis, Kak Juni, Isal, Jaya, Ferdi, Lesta, Agus, Amalia, Eply, Evi, Fathur, Fino, Melinda, Inna, Atti, Taufik, Adi, Fika, Inci**, khususnya **Uncung** yang telah memberikan semangat lebih kepada penulis dan cerita berharga dalam kehidupan bermahasiswa. Teman-teman **POIS20N, KURTO21S** dan **KM FMIPA 2022** yang menjadi tempat berbagi cerita, serta semua pihak yang tidak dapat disebutkan satu-persatu, semoga segala dukungan dan partisipasi yang diberikan bernilai ibadah disisi Allah *Subhanahu Wa Ta'ala*. Penulis menyadari bahwa masih banyak kekurangan, untuk itu dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga tulisan ini memberikan manfaat untuk pembaca.

Makassar, 11 Juni 2024



Ahmad Firdaus Irianto
NIM H051191007

ABSTRAK

AHMAD FIRDAUS IRIANTO. **PENGGUNAAN METODE *POSSIBILISTIC FUZZY C-MEANS* UNTUK PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI SULAWESI SELATAN BERDASARKAN JUMLAH TENAGA KESEHATAN** (dibimbing oleh Anisa)

Latar Belakang. Analisis *cluster* adalah pendekatan untuk menemukan kesamaan dalam data, metode tersebut dapat digunakan untuk mengelompokkan data berdasarkan karakteristik masing-masing begitu juga dalam kasus pengelompokan daerah kabupaten/kota. Metode analisis cluster yang dapat digunakan salah satunya adalah Metode *Possibilistic Fuzzy C-Means* (PFCM) yang dirancang untuk mengatasi kelemahan pada algoritma *Fuzzy C-Means* (FCM) dan *Possibilistic C-Means* (PCM). PFCM mengintegrasikan keunggulan FCM dalam akurasi pengelompokan dan PCM dalam menghadapi data yang tidak pasti. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Provinsi Sulawesi Selatan serta memperoleh hasil evaluasi *Average Silhouette Value* (ASV) dari *cluster* yang terbentuk dengan metode PFCM pada data Data Jumlah Tenaga Kesehatan Kabupaten/Kota di Provinsi Sulawesi tahun 2019. Hasil penelitian diperoleh Nilai ASV pada data dengan *outlier* sebesar 0,74 menunjukkan kualitas *cluster* yang sangat baik, sedangkan pada data tanpa *outlier* memiliki nilai ASV sebesar 0,42 menunjukkan struktur *cluster* yang lemah. Nilai ASV yang tinggi dipengaruhi adanya *outlier* yang meningkatkan jarak antar *cluster*, sehingga membuat setiap titik data lebih dekat ke pusat *cluster*-nya dibandingkan dengan *cluster* lain dan mempengaruhi nilai evaluasi *cluster*. Hasil yang diperoleh menunjukkan bahwa metode PFCM menghasilkan *cluster* yang serupa baik dengan maupun tanpa *outlier* serta PFCM terbukti mampu menangani *outlier* dalam pengelompokan data.

Kata Kunci: *Average Silhouette Value, Fuzzy Clustering, Possibilistic Fuzzy Clustering, Sulawesi Selatan, Tenaga Kesehatan*

ABSTRACT

AHMAD FIRDAUS IRIANTO. Robust Regression Modeling Using Least Median of Squares Estimation on Human Development Index in Indonesia (supervised by Anisa)

Background. *Cluster analysis is an approach used to identify similarities within data, allowing for the grouping of data based on their respective characteristics. This method is applicable in various contexts, including the clustering of districts or municipalities. One specific clustering method is the Possibilistic Fuzzy C-Means (PFCM), which addresses limitations observed in other algorithms such as Fuzzy C-Means (FCM) and Possibilistic C-Means (PCM). PFCM integrates the strengths of FCM for accurate clustering and PCM for handling uncertain data. The aim of the research was to cluster districts and municipalities within South Sulawesi Province, Indonesia, using data on the number of healthcare personnel in each district/municipality in 2019. The evaluation of these clusters was based on the Average Silhouette Value (ASV). The study revealed that the ASV was 0.74 when outliers were included in the data, indicating excellent cluster quality. Conversely, the ASV decreased to 0.42 when outliers were excluded, suggesting a weaker cluster structure. The higher ASV value in the presence of outliers is attributed to the increased distance between clusters, resulting in data points being closer to their respective cluster centers compared to other clusters, thereby influencing the cluster evaluation. These findings highlight that the PFCM method consistently produces similar clusters regardless of the presence of outliers and effectively manages outlier data points within the clustering process.*

Keywords: *Average Silhouette Value, Fuzzy Clustering, Health Workers, Possibilistic Fuzzy Clustering, South Sulawesi*

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGAJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN	iv
UCAPAN TERIMA KASIH	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
DAFTAR ISTILAH	xiv
DAFTAR LAMBANG/SINGKATAN	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Batasan Masalah	2
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Teori	3
1.5.1 Analisis <i>Cluster</i>	3
1.5.2 Jarak Mahalanobis	4
1.5.3 Logika <i>Fuzzy</i>	5
1.5.4 <i>Possibilistic Fuzzy C-Means Clustering</i>	5
1.5.5 <i>Calinski Harabasz Index</i>	7
1.5.6 <i>Average Silhouette Value</i>	7
1.5.7 Tenaga Kesehatan.....	9
BAB II METODOLOGI PENELITIAN	10
2.1 Jenis dan Sumber Data	10
2.2 Variabel Penelitian.....	10
2.3 Struktur Data.....	10
2.4 Metode Analisis	11
BAB III HASIL DAN PEMBAHASAN	13

3.1	Statistika Deskriptif	13
3.2	Deteksi Nilai <i>Outlier</i> Secara <i>Multivariat</i>	14
3.3	<i>Possibilistic Fuzzy C-Means Clustering</i>	16
3.3.1	Penentuan Jumlah <i>cluster</i> menggunakan <i>Calinski Harabasz Index</i>	16
3.3.2	Proses <i>clustering</i> dengan jumlah <i>cluster</i> terbaik.....	17
3.3.3	Interpretasi Hasil <i>Cluster Possibilistic Fuzzy C-Means</i>	23
3.4	<i>Possibilistic Fuzzy C-Means Clustering</i> Tanpa Nilai <i>Outlier</i>	24
3.4.1	Penentuan Jumlah <i>Cluster</i> Menggunakan <i>Calinski Harabasz Index</i>	24
3.4.2	Proses <i>clustering</i> dengan jumlah <i>cluster</i> terbaik.....	24
3.4.3	Interpretasi Hasil <i>Cluster Possibilistic Fuzzy C-Means</i> Tanpa <i>Outlier</i>	30
3.5	Evaluasi Hasil <i>Cluster Possibilistic Fuzzy C-Means</i>	31
BAB IV KESIMPULAN DAN SARAN		34
4.1	Kesimpulan	34
4.2	Saran	34
DAFTAR PUSTAKA		35
LAMPIRAN		38

DAFTAR TABEL

Tabel	Halaman
1. Kriteria Penilaian Clustering Berdasarkan Average Silhouette Value	8
2. Variabel Data Penelitian.....	10
3. Struktur Data Penelitian	10
4. Statistika Deskriptif Jumlah Tenaga Kesehatan	13
5. Nilai Jarak Mahalanobis Kabupaten/Kota Provinsi Sulawesi Selatan	16
6. Nilai <i>Calinski Harabasz Index</i> Untuk Tiap Jumlah <i>Cluster</i>	17
7. Titik Pusat <i>Cluster</i> PFCM pada Iterasi Pertama	17
8. Matriks Keanggotaan <i>Fuzzy</i> PFCM pada Iterasi Pertama	18
9. Nilai Parameter <i>Scaling</i> Iterasi Pertama	19
10. Titik Pusat <i>Cluster</i> PFCM pada Iterasi Kedua	19
11. Matriks Keanggotaan <i>Fuzzy</i> PFCM pada Iterasi Kedua.....	20
12. Titik Pusat <i>Cluster</i> PFCM pada Iterasi Ke-44.....	21
13. Matriks Keanggotaan <i>Fuzzy</i> PFCM pada Iterasi Ke-44.....	21
14. Pembagian <i>Cluster</i> PFCM Wilayah Kabupaten/Kota Sulawesi Selatan	22
15. Nilai Rata-rata Tiap Kelompok <i>Cluster</i> PFCM	23
16. Nilai <i>Calinski Harabasz Index</i> Untuk Tiap Jumlah <i>Cluster</i>	24
17. Titik Pusat <i>Cluster</i> PFCM Tanpa <i>Outlier</i> pada Iterasi Pertama	25
18. Matriks Keanggotaan <i>Fuzzy</i> PFCM Tanpa <i>Outlier</i> pada Iterasi Pertama	26
19. Nilai Parameter <i>Scaling</i> Iterasi Pertama PFCM Tanpa <i>Outlier</i>	27
20. Titik Pusat <i>Cluster</i> PFCM tanpa <i>Outlier</i> pada Iterasi Kedua	27
21. Matriks Keanggotaan <i>Fuzzy</i> PFCM Tanpa <i>Outlier</i> pada Iterasi Kedua	28
22. Titik Pusat <i>Cluster</i> PFCM tanpa <i>Outlier</i> pada Iterasi Ke-35.....	29

23. Matriks Keanggotaan <i>Fuzzy</i> PFCM Tanpa <i>Outlier</i> pada Iterasi Ke-35	29
24. Pembagian <i>Cluster</i> PFCM Tanpa <i>Outlier</i> Wilayah Kabupaten/Kota Sulawesi Selatan.....	30

DAFTAR GAMBAR

Gambar	Halaman
1. Peta Hasil Clustering PFCM Kabupaten/Kota Sulawesi Selatan.....	22
2. Peta Hasil Clustering PFCM Tanpa Outlier Kabupaten/Kota Sulawesi Selatan.....	30
3. Nilai Silhouette Value Hasil Clustering PFCM Pada Keseluruhan Data (a) dan Dengan Data Tanpa Outlier (b) Kabupaten/Kota Sulawesi Selatan.....	33

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Data Jumlah Tenaga Kesehatan Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2019	39
2. Boxplot Data Jumlah Tenaga Kesehatan Menurut Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2019.....	40
3. Nilai Kovarian Antar Variabel Pada Data Jumlah Tenaga Kesehatan di Provinsi Sulawesi Selatan Tahun 2019.....	43
4. Matriks Jarak Antara Data dengan Titik Pusat <i>Cluster</i> PFCM pada Iterasi Pertama	44
5. Matriks Keanggotaan <i>Possibilistic</i> metode PFCM pada Iterasi Pertama.....	45
6. Matriks Jarak Antara Data dengan Titik Pusat <i>Cluster</i> PFCM pada Iterasi Kedua	46
7. Matriks Jarak Antara Data dengan Titik Pusat <i>Cluster</i> PFCM tanpa <i>Outlier</i> pada Iterasi Pertama.....	47
8. Matriks Keanggotaan <i>Possibilistic</i> metode PFCM Tanpa <i>Outlier</i> pada Iterasi Pertama	48
9. Matriks Jarak Antara Data dengan Titik Pusat <i>Cluster</i> PFCM tanpa <i>Outlier</i> pada Iterasi Kedua	49
10. Nilai Jarak Antar Titik Data Jumlah Tenaga Kesehatan Kabupaten/Kota Provinsi Sulawesi Selatan Tahun 2019.....	50
11. Nilai <i>Silhouette Value</i> Hasil PFCM pada Keseluruhan Data	53
12. Nilai <i>Silhouette Value</i> Hasil PFCM pada Data Tanpa Nilai <i>Outlier</i>	54

DAFTAR ISTILAH

Istilah	Arti dan Penjelasan
<i>Cluster</i>	Kelompok data yang memiliki kesamaan atau kedekatan dalam suatu dataset tertentu
<i>Fuzzy</i>	Menunjukkan kepemilikan data yang tidak tegas, menggunakan nilai derajat kepemilikan
<i>Possibilistic</i>	Menunjukkan kemungkinan kepemilikan data tanpa batasan total menjadi satu keseluruhan
<i>Noise</i>	Data yang tidak relevan atau gangguan yang ada dalam dataset yang dianalisis
<i>Outlier</i>	Data yang berbeda jauh dari data lain, seringkali dianggap anomali
Algoritma	Prosedur atau langkah-langkah sistematis untuk menyelesaikan masalah dengan metode tertentu
<i>C-Means</i>	Metode clustering yang menentukan centroid untuk setiap cluster berdasarkan data tertentu.
<i>Machine Learning</i>	Cabang AI yang membuat sistem belajar dari data tanpa eksplisit diprogram.
<i>Centroid</i>	Titik tengah atau representatif dari suatu cluster dalam metode clustering tertentu.
Invers	Operasi matematika untuk menemukan kebalikan suatu matriks, berguna dalam berbagai analisis.
<i>kovarian</i>	Ukuran seberapa dua variabel berubah bersama, menunjukkan hubungan linier antara variabel
Dispersi	Sebaran atau variabilitas data dalam dataset, menunjukkan penyebaran nilai data
<i>Multivariat</i>	Melibatkan banyak variabel atau dimensi dalam analisis data atau model statistik
<i>Index</i>	Indikator atau penunjuk posisi dalam data, berguna untuk mengidentifikasi elemen
<i>Min</i>	Nilai terkecil dalam dataset, menunjukkan batas bawah dari data yang dianalisis
<i>Max</i>	Nilai terbesar dalam dataset, menunjukkan batas atas dari data yang dianalisis
<i>Mean</i>	Rata-rata dari nilai-nilai dalam dataset, memberikan representasi pusat data
Iterasi	Pengulangan langkah-langkah dalam proses algoritma untuk mencapai hasil optimal

DAFTAR LAMBANG/SINGKATAN

Lambang/Singkatan	Arti dan Penjelasan
m	Ekspone <i>fuzzification</i>
q	Ekspone <i>possibilistic</i>
w	Nilai konstan parameter <i>scaling</i>
b	Derajat Ketidakpastian <i>Fuzziness</i>
ε	Nilai Toleransi Error Terkecil
d^2	Jarak <i>Mahalanobis</i>
χ^2	<i>Chi Square</i>
μ	<i>Mean</i> atau Rata-rata
S	Matriks <i>Kovarian</i> Data
df	Derajat Kebebasan
$Cov(X_i, X_i)$	<i>Kovarian</i> Antar Variabel
X	Variabel Data
n	Jumlah Data
$\mu_A[x]$	Himpunan <i>Fuzzy</i>
x	Data atau Observasi
v_i	Pusat <i>Cluster</i>
u_{ij}	Nilai Keanggotaan <i>Fuzzy</i>
t_{ij}	Nilai Keanggotaan <i>Possibilistic</i>
η_i	Parameter <i>Scaling Cluster</i>
a	Parameter Derajat <i>Fuzzy</i>
b	Parameter Derajat <i>Possibilistic</i>
U	Matriks Keanggotaan <i>Fuzzy</i>
T	Matriks Keanggotaan <i>Possibilistic</i>
V	Matriks Pusat <i>Cluster</i>
B_c	Matriks Dispersi <i>Intercluster</i>
W_c	Matriks Dispersi <i>Intracluster</i>
$s(i)$	Nilai <i>Silhouette</i> Data
$a(i)$	Jarak Rata-Rata Sampel
$b(i)$	Minimum Jarak Rata-Rata Sampel
FCM	<i>Fuzzy C-Means</i>
PCM	<i>Possibilistic C-Means</i>
PFCM	<i>Possibilistic Fuzzy C-Means</i>
PPSDM	Pusat Pengembangan Sumber Daya Manusia
ASV	<i>Average Silhouette Value</i>

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis *cluster* merupakan pendekatan untuk mencari kesamaan dalam data dan menempatkan data yang sama ke dalam kelompok-kelompok. Analisis *cluster* membagi sekumpulan data ke dalam beberapa kelompok, berdasarkan kesamaan antar anggota kelompok (Afira & Wijayanto, 2021). *Fuzzy clustering* adalah salah satu metode dari analisis *cluster*, dalam pengelompokan data *fuzzy clustering* mengaplikasikan teori *fuzzy* pada proses pengelompokannya, dengan menggunakan konsep keanggotaan yang fleksibel, *fuzzy clustering* memungkinkan satu data poin dapat memiliki derajat keanggotaan pada beberapa *cluster* yang akan terbentuk secara bersamaan (Eliyanto, 2020).

Metode *Possibilistic Fuzzy C-Means* (PFCM) merupakan salah satu metode dari *fuzzy clustering* yang algoritmanya dirancang untuk mengatasi kelemahan yang ada pada algoritma *Fuzzy C-Means* (FCM) dan *Possibilistic C-Means* (PCM). Algoritma FCM menggunakan derajat keanggotaan untuk mengelompokkan data, tetapi seringkali tidak akurat ketika data mengandung *noise*. Untuk mengatasi masalah ini, algoritma PCM dikembangkan dengan menggunakan konsep teori kemungkinan, yang lebih efektif dalam menghadapi data yang tidak pasti dan noisy. Namun, meskipun PCM lebih baik dalam menangani *noise*, FCM tetap unggul dalam hal akurasi pengelompokan secara keseluruhan (Bahari, 2016).

Algoritma PFCM dikembangkan guna mengintegrasikan keunggulan dari kedua pendekatan ini. Dengan menggabungkan metode fuzzy dari FCM dan metode *possibilistic* dari PCM, PFCM mampu mengatasi masalah yang dihadapi oleh masing-masing algoritma secara individual. PFCM tidak hanya memperbaiki kelemahan FCM dalam menangani data *noisy* tetapi juga memberikan solusi yang lebih *robust* dalam proses pengelompokan dibandingkan dengan PCM. Dengan demikian, PFCM menawarkan pendekatan yang lebih kuat dan fleksibel untuk masalah clustering dalam berbagai kondisi data. (Bahari, 2016).

Metode tersebut dapat digunakan untuk mencari kesamaan dalam data dan menempatkan data yang sama ke dalam kelompok-kelompok. Analisis *cluster* tersebut membagi sekumpulan data ke dalam beberapa kelompok, dimana terdapat kesamaan karakteristik data (Afira & Wijayanto, 2021). Hasil dari pengelompokan dapat digunakan untuk mengambil kesimpulan dalam penentuan keputusan agar keputusan tersebut sesuai dengan bukti ilmiah yang solid (Hakim dkk., 2021).

Pengambilan keputusan yang tepat dan sesuai dengan proses ilmiah merupakan upaya dalam meningkatkan kualitas masyarakat Indonesia dalam sektor kesehatan khususnya dalam peningkatan tenaga kesehatan. Dalam penyediaan tenaga kesehatan di masa depan, beberapa aspek pentingnya adalah produksi tenaga kesehatan yang berkualitas dan kompeten, penempatan yang strategis berdasarkan kebijakan pemerintah dan kebutuhan fasilitas, serta distribusi yang merata sesuai dengan kepadatan penduduk (Lette, 2020).

Konteks pembangunan nasional dalam sektor kesehatan menjadi sangat krusial dan tidak dapat dipisahkan dari keseluruhan usaha pembangunan, karena keberhasilan sektor ini sangat berpengaruh terhadap peningkatan kualitas infrastruktur dalam menghadapi berbagai tantangan global (Luthfia & Alkhajar, 2019). Pembangunan nasional dalam sektor kesehatan sangat erat kaitannya dengan tenaga kesehatan maka dari itu permasalahan distribusi tenaga kesehatan masih menjadi salah satu isu dalam sistem kesehatan di Indonesia. Indonesia memiliki ciri geografis spesifik, sehingga menimbulkan perbedaan keadaan sosial ekonomi yang cukup tinggi dan desentralisasi yang belum mampu menyelesaikan permasalahan pemerataan tenaga kesehatan (Yuningsih, 2014).

Syam pada tahun 2021, telah melakukan penelitian untuk mengelompokkan kabupaten/kota di Provinsi Sulawesi Selatan berdasarkan data jumlah tenaga kesehatan menggunakan metode FCM dan memperoleh hasil bahwa *outlier* mempengaruhi pembentukan kelompok pada data dan pembentukan anggota kelompok yang tidak konsisten. Penelitian yang dilakukan oleh Chowdhary dkk. pada tahun 2020 mengenai segmentasi dalam gambar medis menggunakan *Fuzzy C-Means* dan *Possibilistic Fuzzy C-Means* memberikan hasil bahwa *clustering* menggunakan PFCM dapat menangani permasalahan *cluster* dengan karakteristik data yang tumpang tindih, yang menjadikan hasil *cluster* PFCM adalah 2.81% lebih baik dibandingkan *clustering* FCM.

Penelitian dan latar belakang yang telah dituliskan diatas menunjukkan bahwa metode PFCM mampu memberikan analisis yang lebih baik pada data yang mempunyai *outlier*, karakteristik tumpang tindih, dan mampu memberikan keputusan yang konsisten dalam hasilnya. Sehingga akan dilakukan penelitian untuk melihat bagaimana metode PFCM mengelompokkan kabupaten/kota Provinsi Sulawesi Selatan berdasarkan Data Jumlah Tenaga Kesehatan pada masing-masing kabupaten/kota.

1.2 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Data yang digunakan merupakan data jumlah tenaga Kesehatan menurut kabupaten/kota di Provinsi Sulawesi Selatan tahun 2019 dari *website* resmi Badan PPSDM Kementerian Kesehatan Republik Indonesia.
2. Indikator yang digunakan dalam penelitian ini adalah tenaga medis, keperawatan, kebidanan, kefarmasian, kesehatan masyarakat, kesehatan lingkungan, ahli gizi, keterampilan fisik, keteknisan medis, teknik biomedika, dan tenaga penunjang kesehatan.
3. Parameter model yang digunakan dalam penelitian ini ditetapkan sebelum penelitian dengan nilai Eksponen *fuzzification* (m) = 2, nilai Eksponen *possibilistic* (q) = 2, Nilai konstan parameter *scaling* (w) = 1, Nilai parameter derajat ketidakpastian *fuzziness* (a) = 0,5, Nilai parameter derajat *possibilistic* (b) = 0,5, Jumlah iterasi maksimum = 100 dengan nilai *Error* terkecil yang diharapkan (ϵ) = 10^{-9} .

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Memperoleh hasil pengelompokan data jumlah tenaga kesehatan kabupaten/kota di Provinsi Sulawesi menggunakan metode *Possibilistic Fuzzy C-Means* dibandingkan metode *Fuzzy C-Means*.
2. Memperoleh hasil evaluasi *cluster* yang terbentuk dengan metode *Possibilistic Fuzzy C-Means* pada data Data Jumlah Tenaga Kesehatan Kabupaten/ Kota di Provinsi Sulawesi menggunakan nilai *Average Silhouette Value*.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari hasil penelitian ini adalah sebagai berikut:

1. Penelitian ini diharapkan mampu memberikan kontribusi dalam optimalisasi kebijakan pembangunan di kabupaten/kota Provinsi Sulawesi Selatan. Hasil penelitian dapat menjadi landasan yang kuat untuk merancang kebijakan pembangunan yang lebih spesifik dan responsif terhadap kebutuhan masing-masing daerah.
2. Melalui penerapan metode *Possibilistic Fuzzy C-Means*, penelitian ini diharapkan mampu memberikan pemahaman yang mendalam tentang jumlah tenaga kesehatan di berbagai kabupaten/kota.
3. Memberikan wawasan tambahan berkaitan dengan hasil pengelompokan data jumlah tenaga kesehatan menggunakan *Possibilistic Fuzzy C-means* dibandingkan *Fuzzy C-Means* menggunakan *Calinski Harabasz Index*.

1.5 Teori

1.5.1 Analisis Cluster

Analisis *cluster*, atau lebih umum disebut sebagai *clustering* adalah teknik yang digunakan untuk mengelompokkan observasi, titik data, atau vektor fitur yang serupa berdasarkan karakteristik serupa (Mehta dkk., 2020). Analisis *cluster* adalah teknik *machine learning non-supervisi* yang digunakan untuk mengelompokkan sejumlah sampel data sehingga yang berada dalam kelompok yang sama. Tujuan dari *clustering* adalah membagi sejumlah n sampel data menjadi k *cluster* sehingga setiap sampel data termasuk dalam *cluster* dengan karakteristik yang hampir sama dan memiliki pusat *cluster* yang disebut sebagai *centroid* (Tsai dkk., 2019).

Menurut Kaufman dan Rousseeuw (1990), analisis *cluster* adalah seni menemukan kelompok dalam data. *Clustering* dapat digunakan untuk mengidentifikasi pola dan distribusi menarik dan memberikan wawasan berkaitan dengan struktur data yang mendasarinya. Oleh karena itu, analisis *cluster* adalah teknik yang berguna untuk menemukan dan mengekstrak informasi yang sebelumnya tidak menjadi perhatian. Meskipun konsep analisis *cluster* diusulkan sejak tahun 1930, namun aplikasinya mendapatkan popularitas baru pada tahun 1960-an. Teknik *clustering* telah diterapkan dalam berbagai disiplin ilmu seperti biologi, ilmu sosial hingga kedokteran (Govender & Sivakumar, 2020).

1.5.2 Jarak Mahalanobis

Jarak *mahalanobis* merupakan metode untuk menentukan kemiripan objek berdasarkan pada korelasi variabel dengan pola yang berbeda. Jarak *mahalanobis* adalah metode yang digunakan untuk mendapatkan jarak data tertentu terhadap *mean* data sehingga diperoleh suatu penyebaran data yang memiliki pola terhadap nilai *mean*. Metode ini didasarkan pada korelasi antara variabel dengan pola yang berbeda yang dapat diidentifikasi (Dayana dkk., 2019).

Jarak *mahalanobis* merupakan metode untuk mendeteksi *outlier* pada data *multivariat*. Deteksi *outlier* adalah suatu teknik untuk mencari observasi yang mempunyai perilaku yang berbeda dari observasi lainnya. Salah satu cara untuk mendeteksi *outlier* pada data pengamatan, khususnya data *multivariat* adalah dengan menggunakan konsep jarak. Jarak *mahalanobis* mempertimbangkan korelasi antar variabel serta melibatkan *kovarian* matriks dari data, sehingga dapat diketahui jarak suatu data terhadap sebaran keseluruhan data dalam tiap variabel (Sari dkk., 2021). Pemeriksaan *multivariat outlier* dapat dilakukan dengan statistik *Mahalanobis Distance* (d^2) yang berdistribusi *chi square* (χ^2) dengan derajat kebebasan (*df*) sejumlah variabel pengamatan (*p*). Nilai *Mahalanobis Distance* data pengamatan yang lebih dari nilai *chi square* (χ^2) tabel dengan derajat bebas variabel pengamatan *p* dan taraf signifikansi 0,05 dikategorikan *multivariate outlier* (Efendi dkk., 2022). Berikut adalah rumus Mahalanobis Distance pada persamaan (1):

$$d(x_i) = \sqrt{(x_i - \mu)^T S^{-1}(x_i - \mu)} \quad (1)$$

Keterangan:

$d(x_i)$: Jarak *mahalanobis* untuk data x_i

x_i : Data ke-*i*

μ : *Mean* dari setiap variabel

S^{-1} : Invers matriks *kovarian* data

Matriks *kovarian* mengandung nilai-nilai *kovarian* sebagai elemen dari matriks yang diperoleh dari nilai *kovarian* tiap variabel. Perhitungan *kovarian* antar variabel menggunakan persamaan (2) (Nasution, 2020):

$$Cov(X_i, X_j) = \frac{1}{n-1} \sum_{i=1}^n (X_{ij} - \mu_{X_{ij}})(X_{ij} - \mu_{X_{ij}}) \quad (2)$$

$$i = 1, 2, \dots, 11 ; j = 1, 2, \dots, 24$$

Keterangan:

n : Jumlah data

X_i : Nilai variabel ke-*i*

X_{ij} : Data ke-*j* dari variabel X_i

μ_x : Nilai rata-rata dari variabel X_i

1.5.3 Logika Fuzzy

Logika *fuzzy* pertama kali ditemukan oleh profesor Lotfi A. Zadeh, dari Universitas California, pada bulan Juni 1965. Logika *fuzzy* merupakan generalisasi dari logika klasik yang hanya memiliki dua nilai keanggotaan, yaitu 0 dan 1. Dalam logika *fuzzy*, nilai kebenaran suatu pernyataan berkisar dari sepenuhnya benar, sampai dengan sepenuhnya salah. Dengan teori himpunan *fuzzy*, suatu objek dapat menjadi anggota dari banyak himpunan dengan derajat keanggotaan yang berbeda dalam masing-masing himpunan (Wantoro dkk., 2019).

Himpunan *fuzzy*, keanggotaan suatu unsur dinyatakan dengan derajat keanggotaan (*membership values*), yang nilainya terletak dalam interval $[0,1]$ dan ditentukan dengan fungsi keanggotaan $\mu_A[x] \rightarrow [0,1]$. Untuk sebuah himpunan *fuzzy* A , sebuah elemen x adalah bukan anggota himpunan A jika $\mu_A[x] = 0$, elemen x adalah anggota penuh himpunan A jika $\mu_A[x] = 1$, dan elemen x tersebut adalah anggota himpunan A dengan derajat keanggotaan sebesar μ jika $\mu_A[x] = \mu$, dengan $0 < \mu < 1$ (Cahya, 2017).

1.5.4 Possibilistic Fuzzy C-Means Clustering

FCM adalah suatu teknik *clustering* data yang mencari perbedaan tiap titik data dalam sebuah *cluster* yang ditentukan oleh derajat keanggotaan. Dasar FCM, pertama kali yaitu menentukan pusat *cluster* dan menandai lokasi rata-rata untuk tiap-tiap *cluster*. Pada kondisi awal, pusat *cluster* ini masih belum akurat. Tiap titik data memiliki derajat keanggotaan yang akan diperbarui secara berulang, dapat dilihat bahwa pusat *cluster* akan bergerak menuju lokasi yang tepat. *Possibilistic C-Means* (PCM) menggunakan matriks yang menunjukkan kekhasan dari suatu data. Nilai keanggotaan pada setiap titik data dapat diinterpretasikan pada derajat kemungkinan. Metode *Possibilistic Fuzzy C-Means* merupakan algoritma yang menggabungkan kedua algoritma FCM dan PCM (Apsari dkk., 2020).

Hasil pengelompokan metode FCM lebih sensitif terhadap *noise*. Salah satu keterbatasannya adalah bahwa semua derajat keanggotaan untuk setiap titik data diintegrasikan ke dalam satu *cluster*, yang menyebabkan titik-titik abnormal menjadi anggota *cluster*. Keterbatasan FCM diatasi dengan menggabungkan pendekatan *possibilistic* dengan *Fuzzy C-Mean*, dan pendekatan tersebut dinamakan *Possibilistic Fuzzy C-Means* (PFCM) (Chowdhary dkk., 2020).

Possibilistic Fuzzy C-Means (PFCM) diperkenalkan oleh Pal dkk., (2005). Tujuan utama dari algoritma PFCM adalah mencari pusat *cluster*. Selain pusat *cluster*, algoritma PFCM menghasilkan nilai derajat keanggotaan dan nilai kekhasan atau *typicality value* untuk menentukan setiap titik data yang termasuk pada suatu *cluster* tertentu. *Possibilistic Fuzzy C-Means* (PFCM) merupakan gabungan dari algoritma *Fuzzy C-Means* (FCM) dan *Possibilistic C-Means* (PCM). Algoritma PFCM menghindari berbagai kelemahan pada algoritma FCM dan PCM. Pada algoritma PFCM dapat memecahkan permasalahan pada FCM berupa *noise* data dan juga memberi jawaban atas permasalahan ketika pengelompokan data dengan algoritma PCM (Bahari, 2016). Serupa dengan FCM, algoritma PFCM melakukan iterasi untuk

memperoleh nilai matriks U dan V yang paling optimal. Proses memperbarui nilai matriks keanggotaan dan pusat *cluster* pada PFCM pada persamaan (3):

$$u_{ij}^{(t+1)} = \left(\sum_{q=1}^c \left(\frac{\|x_j - v_i^{(t)}\|}{\|x_j - v_q^{(t)}\|} \right)^{\frac{2}{m-1}} \right)^{-1} \quad (3)$$

$$v_{il}^{(t+1)} = \frac{\sum_{j=1}^K \left(a(u_{ij}^{(t+1)})^m + b(t_{ij}^{(t+1)})^q \right) x_{jl}}{\sum_{j=1}^K \left(a(u_{ij}^{(t+1)})^m + b(t_{ij}^{(t+1)})^q \right)} \quad (4)$$

Keterangan:

- x_j : Nilai observasi ke- j dengan $x_j = (x_{j1}, x_{j2}, \dots, x_{jL})$.
- v_i : Pusat *cluster* ke- i .
- t_{ij} : Nilai tingkat keanggotaan *possibilistic*
- m : Nilai eksponen *fuzzification* ($m > 1$).
- q : Nilai eksponen *possibilistic* ($q > 1$)
- t : Menunjukkan urutan iterasi 0,1,2, ..., hingga iterasi maksimal
- l : Dimensi dari data $l = (1,2, \dots, L)$
- a : Nilai parameter derajat ketidakpastian atau *fuzziness*
- b : Nilai parameter derajat *possibilistic*

Nilai t_{ij} merupakan partisi dari matriks keanggotaan *possibilistic* T dengan $T = [t_{ij}]_{c \times K}$. Nilai t_{ij} diperbarui pada tiap iterasi menggunakan persamaan (5):

$$t_{ij}^{(t+1)} = \left(1 + \left(\frac{b \|x_j - v_i^{(t)}\|^2}{\eta_i^{(t+1)}} \right)^{\frac{1}{q-1}} \right)^{-1} \quad (5)$$

Dengan:

$$\eta_i^{(t+1)} = w \frac{\sum_{j=1}^K (u_{ij}^{(t+1)})^m \|x_j - v_i^{(t)}\|^2}{\sum_{j=1}^K (u_{ij}^{(t+1)})^m} \quad (6)$$

Keterangan:

- η_i : Parameter *scaling* untuk *cluster* ke- i
- w : Merupakan nilai konstan (umumnya $w = 1$)

Proses iterasi PFCM dilakukan hingga diperoleh hasil yang konvergen atau hasil perhitungan tidak lagi memberikan perbedaan yang signifikan. Nilai matriks matriks keanggotaan *fuzzy* U dijadikan sebagai titik acuan untuk melihat apakah proses iterasi telah mencapai nilai yang optimal menggunakan persamaan sebagai berikut (Chen dkk., 2021):

$$\|U^t - U^{t-1}\| < \varepsilon \quad (7)$$

Keterangan:

- $\|U^t - U^{t-1}\|$: Jarak *euclidian* antara matriks U iterasi sekarang dan sebelumnya
 U : Matriks keanggotaan *fuzzy*
 t : Jumlah iterasi terlewat
 ε : Nilai error yang diharapkan sebesar 10^{-9}

1.5.5 Calinski Harabasz Index

Calinski Harabasz Index merupakan ukuran evaluatif yang digunakan untuk mengukur performa hasil pengelompokan data (*clustering*). Indeks ini dihitung sebagai rasio antara dispersi rata-rata di antara setiap kelompok (*cluster*) dengan distribusi di antara kelompok-kelompok tersebut (Amalia dkk., 2021). *Calinski Harabasz Index* yang lebih tinggi mengindikasikan hasil model dengan pembagian *cluster* yang lebih baik (Sikana & Wijayanto, 2021). *Calinski Harabasz Index* adalah rasio jumlah dispersi antar *cluster* dan dispersi dalam *cluster* untuk semua *cluster* (dispersi didefinisikan sebagai jumlah kuadrat jarak). Setiap sekumpulan data X berukuran $K \times L$ yang telah dikelompokkan menjadi c *cluster*, *Calinski Harabasz Index* dinyatakan dalam Persamaan (8) sebagai berikut (Sikana & Wijayanto, 2021):

$$S = \frac{tr(B_c)}{tr(W_c)} \times \frac{K - c}{c - 1} \quad (8)$$

Dengan $tr(B_c)$ adalah *trace* dari matriks dispersi antar *cluster* dan $tr(W_c)$ adalah *trace* matriks dispersi dalam *cluster* yang dinyatakan dalam Persamaan (9) dan (10) berikut:

$$B_c = \sum_{q=1}^c K_q (\mathbf{v}_q - \mathbf{v}_x)(\mathbf{v}_q - \mathbf{v}_x)^T \quad (9)$$

$$W_c = \sum_{q=1}^c \sum_{x \in C_q} (x - \mathbf{v}_q)(x - \mathbf{v}_q)^T \quad (10)$$

Keterangan:

- C_q : Kumpulan titik dalam *cluster* q
 \mathbf{v}_q : Nilai rata-rata *cluster* q
 \mathbf{v}_x : Nilai rata-rata keseluruhan data
 K_q : Jumlah titik dalam *cluster* q

1.5.6 Average Silhouette Value

Average Silhouette Value merupakan salah satu metode validasi pengujian hasil *cluster* yang diperkenalkan oleh Rousseeuw pada tahun 1987. *Average Silhouette* adalah suatu nilai yang mengukur tingkat kehomogenan anggota dalam internal

cluster dan tingkat keheterogenan anggota antar *cluster*. Nilai *Average Silhouette Value* berada pada interval -1 hingga 1. Persamaan untuk menghitung *Average Silhouette Value* persamaan (11) (Septianingsih, 2022):

$$ASV = \frac{1}{N} \sum_{i=1}^N s(i) \quad (11)$$

Dengan:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (12)$$

$$a(i) = \frac{1}{nc - 1} \sum_{r=1}^{nc-1} d(x_{ic} - x_{rc}) \quad (13)$$

$$d(i, C) = \frac{1}{nk} \sum_{r=1}^{nk} d(x_i - x_{rc}) \quad (14)$$

$$b(i) = \min d(i, C) \quad (15)$$

Keterangan:

- ASV* : *Average Silhouette Value*
s(i) : Nilai *silhouette* masing-masing objek
a(i) : Jarak rata-rata sampel *i* ke sampel lain dalam *cluster*
b(i) : Minimum jarak rata-rata sampel dari sampel *i* ke *cluster* lain.
nc : Jumlah data dalam *cluster c*
nk : Jumlah data di luar *cluster c*
d(x_{ic} - x_{rc}) : jarak data ke - *i* dengan data ke - *r* dalam satu *cluster*
d(x_i - x_{rc}) : jarak data ke - *i* dengan data ke - *r* tidak satu *cluster*

Ketika *silhouette value* mendekati 1, ini menunjukkan bahwa jarak rata-rata antara titik data dan titik-titik dalam *cluster* yang sama mendekati nol, berarti jarak rata-rata antar data dalam satu *cluster* jauh lebih kecil dibandingkan jarak terkecil antara titik data dan titik-titik dari *cluster* lain, begitu pula sebaliknya. Ketika *silhouette value* mendekati nol, menunjukkan bahwa jarak rata-rata dalam *cluster* hampir sama dengan jarak rata-rata ke *cluster* lain terdekat (Dinh dkk., 2019). Standar kriteria evaluasi *cluster* menggunakan nilai *silhouette value* (Guntara & Suprawoto, 2022) :

Tabel 1. Kriteria Penilaian Clustering Berdasarkan Average Silhouette Value

Average Silhouette Value	Kriteria
$0,7 < ASV \leq 1$	Struktur <i>cluster</i> yang sangat baik
$0,5 < ASV \leq 0,7$	Struktur <i>cluster</i> yang cukup baik
$0,25 < ASV \leq 0,5$	Struktur <i>cluster</i> yang lemah
$ASV \leq 0,25$	Tidak ada struktur <i>cluster</i>

1.5.7 Tenaga Kesehatan

Menurut Undang-Undang No. 36 Tahun 2009 Tentang Kesehatan adalah segala bentuk dana, tenaga, perbekalan kesehatan, sediaan farmasi, dan alat kesehatan serta fasilitas pelayanan kesehatan dan teknologi yang dimanfaatkan untuk menyelenggarakan upaya kesehatan yang dilakukan oleh pemerintah, pemerintah daerah, dan/ atau masyarakat. Tenaga kesehatan meliputi tenaga medis (dokter umum dan spesialis), tenaga psikologi klinis, tenaga kesehatan Masyarakat, tenaga kesehatan lingkungan, keterampilan fisik, keteknisan medis, teknis biomedika, dan tenaga kesehatan lain (Al Isra & Ramli, 2020). Penjelasan untuk tiap variabel yang digunakan terdapat pada Undang-Undang No. 36 Tahun 2014 Tentang Tenaga Kesehatan Sebagai Berikut (Pemerintah Indonesia, 2014):

1. Jenis Tenaga Kesehatan yang termasuk dalam kelompok tenaga medis terdiri atas dokter, dokter gigi, dokter spesialis, dan dokter gigi spesialis.
2. Jenis Tenaga Kesehatan yang termasuk dalam kelompok tenaga keperawatan terdiri atas berbagai jenis perawat.
3. Jenis Tenaga Kesehatan yang termasuk dalam kelompok tenaga kebidanan adalah bidan.
4. Jenis Tenaga Kesehatan yang termasuk dalam kelompok tenaga kefarmasian terdiri atas apoteker dan tenaga teknis kefarmasian
5. Jenis Tenaga Kesehatan yang termasuk dalam kelompok tenaga kesehatan masyarakat terdiri atas epidemiolog kesehatan, tenaga promosi kesehatan dan ilmu perilaku, pembimbing kesehatan kerja, tenaga administrasi dan kebijakan kesehatan, tenaga biostatistik dan kependudukan, serta tenaga kesehatan reproduksi dan keluarga.
6. Jenis Tenaga Kesehatan yang termasuk dalam kelompok tenaga kesehatan lingkungan sebagaimana terdiri atas tenaga sanitasi lingkungan, entomolog kesehatan, dan mikrobiolog kesehatan.
7. Jenis Tenaga Kesehatan yang termasuk dalam kelompok tenaga gizi terdiri atas nutrisisionis dan dietisien.
8. Jenis Tenaga Kesehatan yang termasuk dalam kelompok tenaga keterampilan fisik terdiri atas fisioterapis, okupasi terapis, terapis wicara, dan akupunktur.
9. Jenis Tenaga Kesehatan yang termasuk dalam kelompok tenaga keteknisian medis terdiri atas perekam medis dan informasi kesehatan, teknik kardiovaskuler, teknisi pelayanan darah, refraksionis optisien/ optometris, teknisi gigi, penata anestesi, terapis gigi dan mulut, dan audiologis.
10. Jenis Tenaga Kesehatan yang termasuk dalam kelompok tenaga teknik biomedika terdiri atas radiografer, elektromedis, ahli teknoigi laboratorium medik, fisikawan medik, radioterapis, dan ortotik prostetik.
11. Tenaga Penunjang Kesehatan adalah sumber daya manusia kesehatan yang terlibat dan bekerja serta mengabdikan dirinya dalam upaya dan manajemen pada fasilitas kesehatan (Attriani, 2022).

BAB II METODOLOGI PENELITIAN

2.1 Jenis dan Sumber Data

Penelitian ini menggunakan data sekunder, yaitu Data Jumlah Tenaga Kesehatan menurut kabupaten/kota di Provinsi Sulawesi Selatan tahun 2019 yang diperoleh dari *website* resmi Badan PPSDM Kementerian Kesehatan Republik Indonesia pada situs <https://bppsdmk.kemkes.go.id>.

2.2 Variabel Penelitian

Variabel yang dianalisis dalam penelitian ini adalah jumlah setiap jenis tenaga kesehatan tiap kabupaten di Provinsi Sulawesi Selatan.

Tabel 2 Variabel Data Penelitian

Variabel	Keterangan
X1	Tenaga Medis
X2	Keperawatan
X3	Kebidanan
X4	Kefarmasian
X5	Kesehatan Masyarakat
X6	Kesehatan Lingkungan
X7	Ahli Gizi
X8	Keterampilan Fisik
X9	Keteknisan Medis
X10	Teknik Biomedika
X11	Tenaga Penunjang Kesehatan

2.3 Struktur Data

Data yang digunakan dalam penelitian terdiri atas 24 observasi yang mencakup 24 kabupaten/kota di Sulawesi Selatan. Tabel data penelitian secara lengkap dapat ditemukan pada Lampiran 1. Struktu data yang digunakan ditampilkan dalam tabel berikut:

Tabel 3 Struktur Data Penelitian

Kabupaten/Kota	X1	X2	...	X9	X10	X11
Bantaeng	$X1_1$	$X2_1$...	$X9_1$	$X10_1$	$X11_1$
Barru	$X1_2$	$X2_2$...	$X9_2$	$X10_2$	$X11_2$
Bone	$X1_3$	$X2_3$...	$X9_3$	$X10_3$	$X11_3$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
Toraja Utara	$X1_{23}$	$X2_{23}$...	$X9_{23}$	$X10_{23}$	$X11_{23}$
Wajo	$X1_{24}$	$X2_{24}$...	$X9_{24}$	$X10_{24}$	$X11_{24}$

2.4 Metode Analisis

Penelitian ini dilakukan dengan pengelompokan kabupaten/kota di provinsi Sulawesi Selatan dengan metode PFCM. Pengolahan data dilakukan dengan bahasa pemrograman *Python*. Tahapan analisis data yang dilakukan dalam penelitian ini yaitu sebagai berikut:

1. Melakukan eksplorasi data dengan menentukan statistika deskriptif dari setiap variabel kabupaten/kota di provinsi Sulawesi Selatan untuk melihat gambaran awal dan sebaran data.
2. Mendeteksi observasi yang termasuk sebagai nilai *outlier* secara *multivariat* dengan melihat nilai jarak *mahalanobis* yang diperoleh menggunakan Persamaan (1).
3. Menentukan jumlah *cluster* terbaik untuk proses pengelompokan *Possibilistic Fuzzy C-Means Clustering* menggunakan nilai *Calinski Harabasz Index* yang diperoleh dengan menggunakan Persamaan (8) dan membutuhkan nilai titik pusat *cluster* serta hasil pengelompokan dari tiap observasi. Penentuan titik pusat *cluster* dan pengelompokan tiap observasi dilakukan sebagai berikut:
 - a. Memasukkan data X yang akan dikelompokkan berupa matriks berukuran $K \times L$, dengan $K = 24$ dan $L = 11$.
 - b. Menetapkan jumlah *cluster* yang akan digunakan dalam proses *clustering*, untuk proses penentuan jumlah *cluster* terbaik menggunakan nilai *Calinski Harabasz Index* maka digunakan jumlah *cluster* (2, 3, 4, 5, 6).
 - c. Memasukkan nilai parameter yang telah ditetapkan dalam pada Batasan masalah penelitian.
 - d. Menentukan nilai titik pusat *cluster* V dengan menggunakan nilai acak sebagai inisialisasi awal pusat *cluster* V .
 - e. Menghitung nilai matriks keanggotaan *fuzzy* U untuk setiap observasi menggunakan Persamaan (3) serta nilai pusat masing-masing *cluster* pada matriks V .
 - f. Menghitung nilai parameter *scaling* η_i menggunakan Persamaan (6) nilai η_i akan digunakan pada perhitungan matriks *possibilistic*.
 - g. Menghitung nilai matriks keanggotaan *possibilistic* T menggunakan Persamaan (5) dan nilai parameter *scaling* η_i .
 - h. Memperbarui nilai titik pusat masing-masing *cluster* pada matriks V menggunakan Persamaan (4).
 - i. Mengulangi proses iterasi kembali ke langkah (e) untuk memperbarui nilai matriks keanggotaan *fuzzy* pada iterasi sebelumnya hingga terpenuhi kondisi $\|U^{(t+1)} - U^t\| < \varepsilon$ atau jumlah maksimal iterasi telah tercapai, maka iterasi dihentikan.
 - j. Memetakan *cluster* tiap observasi berdasarkan nilai pada matriks keanggotaan *fuzzy*.
 - k. Menghitung *Calinski Harabasz Index* menggunakan Persamaan (8) untuk masing-masing jumlah *cluster* yang digunakan.

4. Apabila pada hasil pemetaan kabupaten/kota di provinsi Sulawesi Selatan berdasarkan data kesehatan terdapat kabupaten/kota yang membentuk *cluster* sendiri dikarenakan memiliki karakteristik yang jauh berbeda dengan kabupaten/kota lainnya, maka kabupaten/kota yang termasuk kedalam nilai *outlier* pada langkah (2) dikeluarkan lalu proses analisis dilakukan kembali dari langkah (3).
5. Mengevaluasi hasil *cluster* dengan menghitung nilai *Average Silhouette Value* menggunakan Persamaan (11).
6. Hasil yang diperoleh diinterpretasikan lalu membuat kesimpulan.