

**PENGGUNAAN METODE *GAUSSIAN KERNEL FUZZY C-MEANS*  
UNTUK PENGELOMPOKAN KABUPATEN/KOTA PROVINSI SULAWESI  
SELATAN BERDASARKAN DATA KRIMINALITAS**



**AYU LESTARI  
H051191046**



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

**PENGGUNAAN METODE *GAUSSIAN KERNEL FUZZY C-MEANS*  
UNTUK PENGELOMPOKAN KABUPATEN/KOTA PROVINSI SULAWESI  
SELATAN BERDASARKAN DATA KRIMINALITAS**

**AYU LESTARI  
H051191046**



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

**PENGGUNAAN METODE GAUSSIAN KERNEL FUZZY C-MEANS  
UNTUK PENGELOMPOKAN KABUPATEN/KOTA PROVINSI SULAWESI  
SELATAN BERDASARKAN DATA KRIMINALITAS**

**AYU LESTARI**

**H051191046**

Skripsi

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**  
**PENGGUNAAN METODE GAUSSIAN KERNEL FUZZY C-MEANS**  
**UNTUK PENGELOMPOKAN KABUPATEN/KOTA PROVINSI SULAWESI**  
**SELATAN BERDASARKAN DATA KRIMINALITAS**

**AYU LESTARI**

**H051191046**

Skripsi,

Telah dipertahankan di depan Panitia Ujian Sarjana Statistika pada 14 Agustus 2024  
dan dinyatakan telah memenuhi syarat kelulusan

pada

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

Mengesahkan:

Pembimbing tugas akhir,

Anisa, S.Si., M.Si.

NIP.1973022719988022001



Diketahui:  
Ketua Program Studi,

Dr. Anna Lemiyati, S.Si., M.Si.

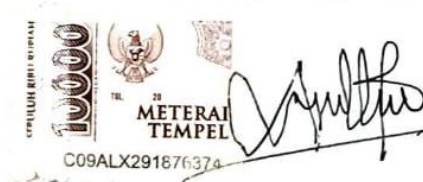
NIP.197708082005012002

**PERNYATAAN KEASLIAN SKRIPSI  
DAN PELIMPAHAN HAK CIPTA**

Dengan ini saya menyatakan bahwa, skripsi berjudul “Penggunaan Metode *Gaussian Kernel Fuzzy C-Means* Untuk Pengelompokan Kabupaten/Kota Provinsi Sulawesi Selatan Berdasarkan Data Kriminalitas” adalah benar karya saya dengan arahan dari Ibu Anisa, S.Si., M.Si. sebagai Pembimbing. 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, 14 Agustus 2024



Ayu Lestari

NIM H051191046


## 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. *Alhamdulillahil'alamiin*, 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 **Prof. Dr. Nurtiti Sunusi, S.Si., M.Si.** selaku penasehat akademik dan ibu **Anisa, S.Si., M.Si.** selaku pembimbing yang dengan penuh kesabaran telah meluangkan waktu dan pemikirannya. Ucapan terima kasih juga penulis sampaikan kepada Ibu **Prof. Dr. Nurtiti Sunusi, S.Si., M.Si.** dan ibu **Dra. Nasrah Sirajang, 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 keluarga terkhusus orang tua, Ayah **Zaenal Arifin** dan Ibu **Marwah** yang telah memberikan dukungan, pengorbanan kasih sayang serta doa yang tak henti-hentinya dipanjatkan kepada penulis.

Penulis juga berterima kasih kepada teman-teman terkhusus **Syamsul, Daus, Ferdi, Fadhel, Alya, Dian, Tamara, Wilda, Astri, Fitri, Fira, Nita**, yang telah memberikan semangat lebih kepada penulis dan cerita berharga dalam kehidupan bermahasiswa. Teman-teman **MIPA 2019**, 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, 14 Agustus 2024



AYU LESTARI

## ABSTRAK

AYU LESTARI. **Penggunaan metode *gaussian kernel fuzzy c-means* untuk pengelompokan kabupaten/kota provinsi sulawesi selatan berdasarkan data kriminalitas** (dibimbing oleh Anisa).

**Latar Belakang.** Analisis *cluster* adalah analisis untuk mengelompokkan data yang memiliki elemen dan karakteristik serupa ke dalam kelompok (*cluster*) yang berbeda dengan mengklasifikasi berdasarkan variabel-variabel yang dipertimbangkan dalam penelitian. *Fuzzy C-Means* (FCM) yang merupakan salah satu metode *clustering non-hierarki* dengan keanggotaan data dalam suatu *cluster* ditentukan oleh derajat keanggotaan. FCM menggunakan jarak antar data dalam menentukan nilai keanggotaan observasi, sehingga metode FCM sangat rentan terhadap pengaruh *noise* dan perbedaan skala nilai pada data. **Metode.** GKFCM bertujuan untuk meningkatkan kemampuan FCM dalam menangkap struktur dan perbedaan variabel yang lebih kompleks dalam data dengan memetakan data menggunakan *kernel*. **Tujuan.** Penelitian ini adalah untuk memperoleh jumlah *cluster* terbaik serta karakteristik setiap *cluster* pada data kriminalitas kabupaten/kota di Provinsi Sulawesi Selatan menggunakan GKFCM. **Hasil.** *Cluster* yang terbentuk sebanyak 4 kelompok wilayah berdasarkan nilai *Davies-Bouldin Index* terendah sebesar 0,82. Kelompok pertama merupakan wilayah dengan resiko kejahatan tertinggi diikuti oleh kelompok kedua dan keempat hingga kelompok ketiga yang memiliki resiko kejahatan paling rendah. Nilai evaluasi *Average Silhouette Value* metode GKFCM sebesar 0,74 menunjukkan bahwa kualitas *cluster* yang terbentuk termasuk ke dalam struktur *cluster* yang sangat baik. **Kesimpulan.** Metode GKFCM mampu mengelompokkan kabupaten/kota di Provinsi Sulawesi Selatan menjadi 4 *cluster* dengan kualitas yang sangat baik, yang menunjukkan perbedaan risiko kejahatan yang signifikan di antara kelompok-kelompok wilayah di Sulawesi Selatan.

**Kata Kunci:** *Average Silhouette Value, Fuzzy Clustering, Gaussian Kernel Fuzzy Clustering, Kriminalitas, Sulawesi Selatan*



## ABSTRACT

AYU LESTARI. **The Use of Gaussian Kernel Fuzzy C-Means for Clustering Regencies/Cities in South Sulawesi Province Based on Crime Data** (supervised by Anisa).

**Introduction.** Cluster analysis is a method to group data with similar elements and characteristics into different clusters by classifying them based on variables considered in the research. Fuzzy C-Means (FCM) is one of the non-hierarchical clustering methods where data membership in a cluster is determined by the degree of membership. FCM uses the distance between data points to determine membership values, making it very susceptible to noise and differences in data value scales. **Method.** GKFCM aims to enhance FCM's ability to capture more complex structures and variable differences in data by mapping the data using a kernel. **Objective.** This research aims to obtain the best number of clusters and the characteristics of each cluster in crime data of regencies/cities in South Sulawesi Province using GKFCM. **Results.** The formed clusters consist of 4 groups of regions based on the lowest Davies-Bouldin Index value of 0.82. The first group represents areas with the highest crime risk, followed by the second and fourth groups, with the third group having the lowest crime risk. The evaluation value of the GKFCM method's Average Silhouette Value of 0.74 indicates that the formed clusters are of very good quality. **Conclusion.** The GKFCM method effectively clusters regencies/cities in South Sulawesi Province into 4 clusters with very good quality, demonstrating significant differences in crime risk among the region groups in South Sulawesi.

**Keywords:** Average Silhouette Value, Crime, Fuzzy Clustering, Gaussian Kernel Fuzzy Clustering, South Sulawesi



## DAFTAR ISI

<b>HALAMAN JUDUL .....</b>	<b>i</b>
<b>PERNYATAAN PENGAJUAN .....</b>	<b>ii</b>
<b>HALAMAN PENGESAHAN.....</b>	<b>iii</b>
<b>PERNYATAAN KEASLIAN SKRIPSI .....</b>	<b>iv</b>
<b>UCAPAN TERIMA KASIH .....</b>	<b>v</b>
<b>ABSTRAK.....</b>	<b>vi</b>
<b>ABSTRACT .....</b>	<b>vii</b>
<b>DAFTAR ISI .....</b>	<b>viii</b>
<b>DAFTAR TABEL .....</b>	<b>x</b>
<b>DAFTAR GAMBAR .....</b>	<b>xi</b>
<b>DAFTAR LAMPIRAN.....</b>	<b>xii</b>
<b>DAFTAR NOTASI .....</b>	<b>xiv</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah .....	2
1.4 Tujuan Penelitian .....	3
1.5 Manfaat Penelitian .....	3
1.6 Landasan Teori .....	3
1.6.1 Analisis <i>Cluster</i> .....	3
1.6.2 <i>Fuzzy C-Means Clustering</i> .....	3
1.6.3 <i>Gaussian Kernel</i> .....	6
1.6.4 <i>Gaussian Kernel Fuzzy C-Means Clustering</i> .....	6
1.6.5 <i>Davies-Bouldin Index</i> .....	7
1.6.6 <i>Average Silhouette Value</i> .....	8
1.6.7 Kriminalitas.....	10
<b>BAB II METODOLOGI PENELITIAN .....</b>	<b>12</b>
2.1 Jenis dan Sumber Data .....	12
2.2 Struktur Data.....	12
2.3 Metode Analisis.....	12
<b>BAB III HASIL DAN PEMBAHASAN.....</b>	<b>14</b>

3.1 Penentuan Jumlah <i>Cluster</i> Menggunakan <i>Davies-Bouldin Index</i> .....	14
3.2 Proses Analisis <i>Gaussian Kernel Fuzzy C-Means Clustering</i> .....	43
3.3 Interpretasi Hasil <i>Cluster</i> .....	50
3.4 Evaluasi Hasil <i>Cluster Gaussian Kernel Fuzzy C-Means</i> .....	51
<b>BAB IV PENUTUP</b> .....	<b>55</b>
4.1 Kesimpulan .....	55
4.2 Saran .....	55
<b>DAFTAR PUSTAKA</b> .....	<b>56</b>
<b>LAMPIRAN</b> .....	<b>59</b>

**DAFTAR TABEL**

<b>Tabel</b>	<b>Halaman</b>
1. Kriteria Penilaian <i>Clustering</i> Berdasarkan <i>Average Silhouette Value</i> .....	9
2. Struktur Variabel Data .....	12
3. Nilai <i>Davies-Bouldin Index</i> Untuk Tiap Jumlah <i>Cluster</i> GKFCM .....	41
4. Titik Pusat <i>Cluster</i> Pada Iterasi Ke-1 .....	43
5. Matriks Keanggotaan <i>Fuzzy</i> GKFCM Pada Iterasi Ke-1 .....	46
6. Titik Pusat <i>Cluster</i> Pada Iterasi Ke-2 .....	47
7. Titik Pusat <i>Cluster</i> Pada Iterasi Ke-93.....	48
8. Matriks Keanggotaan <i>Fuzzy</i> GKFCM Pada Iterasi Ke-93 .....	48
9. Nilai Rata-rata untuk Masing-masing Kelompok Wilayah.....	51

**DAFTAR GAMBAR**

<b>Gambar</b>	<b>Halaman</b>
1. Peta <i>Cluster</i> GKFCM Kabupaten/Kota Sulawesi Selatan.....	50
2. Nilai <i>Silhouette Value</i> Hasil <i>Clustering</i> GKFCM pada Data Kriminalitas Kabupaten/Kota di Sulawesi Selatan .....	50

**DAFTAR LAMPIRAN**

<b>Lampiran</b>	<b>Halaman</b>
1. Data Kriminalitas Sulawesi Selatan Tahun 2022 .....	60
2. Nilai <i>Kernel Gaussian</i> Titik Pusat Dengan Data Pada Iterasi Ke-1 .....	67
3. Jarak <i>Euclidian</i> Antar Titik Data Kriminalitas Tiap Kabupaten/Kota.....	35
4. Nilai <i>Silhouette Value</i> Hasil GKFCM pada Keseluruhan Data.....	38

## 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>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 <i>clustering</i> yang menentukan <i>centroid</i> untuk setiap <i>cluster</i> 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 <i>cluster</i> dalam metode <i>clustering</i> 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
$m$	Degree of <i>fuzzification</i>
$n$	Jumlah banyaknya data
$c$	Jumlah banyaknya <i>cluster</i>
$x_i$	Data ke- $i$
$v_j$	Pusat <i>cluster</i> ke- $j$
$u_{ij}$	Nilai keanggotaan data ke- $i$ pada <i>cluster</i> ke- $j$
$\sigma^2$	Variansi dari data
$\psi(x_i, v_j)$	Nilai <i>gaussian kernel</i> data $x_i$ dengan pusat <i>cluster</i> $v_j$
<b>DBI</b>	<i>Davies-bouldin index</i>
<b>MAE<sub>j</sub></b>	Jarak intra- <i>cluster</i> ke- $j$
$d_j$	Nilai rata-rata <i>cluster</i> ke- $j$
$T_j$	Jumlah data pada <i>cluster</i> ke- $j$
$R_{j,k}$	Kesamaan antar <i>cluster</i> ke- $j$ dan ke- $k$
$\ d_j - d_k\ $	Jarak <i>euclidian mean cluster</i> ke- $j$ dan <i>cluster</i> ke- $k$
$\ x_i - d_j\ $	Jarak <i>euclidian mean cluster</i> ke- $j$ dan data ke- $i$
<b>ASV</b>	<i>Average silhouette value</i>
$s(i)$	Nilai <i>silhouette</i> masing-masing objek
$a(i)$	Jarak rata-rata sampel $i$ ke sampel lain dalam <i>cluster</i>
$b(i)$	Minimum jarak rata-rata sampel dari sampel $i$ ke <i>cluster</i> lain.
$nc$	Jumlah data dalam <i>cluster</i> $c$
$nk$	Jumlah data di luar <i>cluster</i> $c$



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Analisis *cluster* adalah analisis untuk mengelompokkan data yang memiliki elemen dan karakteristik serupa ke dalam kelompok (*cluster*) yang berbeda. Analisis *cluster* digunakan untuk mengeksplorasi dan mengelompokkan observasi-observasi yang serupa berdasarkan pola-pola dari kumpulan observasi tersebut. Analisis *cluster* berguna untuk meringkas data dengan cara mengelompokkan data berdasarkan kesamaan karakteristik data yang diteliti (Sitepu dan Gultom, 2011). Tujuan utama analisis *cluster* adalah mengklasifikasi objek (kasus atau elemen) kedalam kelompok-kelompok yang relatif homogen berdasarkan variabel-variabel yang dipertimbangkan dalam penelitian. Analisis *cluster* terbagi atas dua metode, yaitu metode hierarki dan metode *non*-hierarki (Goreti dkk., 2016).

*Fuzzy C-Means* (FCM) yang merupakan salah satu bagian dari metode *clustering non*-hierarki adalah suatu teknik *clustering* yang mana keanggotaan data dalam suatu *cluster* ditentukan oleh derajat keanggotaan. Derajat keanggotaan pada FCM memungkinkan suatu titik data bisa termasuk kedalam lebih dari satu kemungkinan *cluster* (Hardiyanto dkk., 2012). Algoritma FCM memiliki kelebihan dalam penempatan pusat *cluster* yang lebih tepat dibandingkan dengan metode *cluster* lainnya. FCM dapat mengatasi kelemahan dari metode seperti *K-Means* dengan memperbarui pusat *cluster* menggunakan nilai keanggotaan *fuzzy*

setiap data (Nurkholis dkk., 2022). FCM dikenal baik karena dapat mengatasi ketidakpastian dan fleksibilitas keanggotaan observasi dalam mendeteksi keanggotaan *cluster* setiap data (Nidyashofa dan Istiawan, 2017).

Metode FCM menggunakan jarak antar data dalam menentukan nilai keanggotaan observasi, sehingga metode FCM sangat rentan terhadap pengaruh *noise* dan perbedaan skala nilai pada data (Chowdhary dkk., 2020). Solusi yang dapat diterapkan adalah dengan menggunakan *kernel gaussian* pada metode FCM yang membantu dalam memetakan hubungan antar titik data. Metode FCM yang menggunakan *kernel gaussian* dikenal dengan sebutan *Gaussian Kernel Fuzzy*

*C-Means* (GKFCM). Metode GKFCM bertujuan untuk meningkatkan kemampuan FCM dalam menangkap struktur dan perbedaan variabel yang lebih kompleks dalam data dengan memetakan data menggunakan *kernel* (Liu dkk., 2019). Metode GKFCM memiliki parameter jumlah *cluster* yang perlu untuk dievaluasi untuk memperoleh jumlah *cluster* paling optimal dengan melihat seberapa baik data telah dikelompokkan oleh masing-masing jumlah *cluster* yang ditentukan. Davies Bouldin Index (DBI) merupakan salah satu metrik yang umum digunakan untuk evaluasi hasil *cluster* (Fikri dkk., 2023). DBI mengukur kualitas *cluster* dengan cara

membandingkan jarak antara *cluster* dan jarak dalam *cluster*. Semakin rendah nilai DBI, semakin baik kualitas *clustering* yang dihasilkan, menunjukkan bahwa *cluster* lebih seragam dan terpisah baik dari data pada *cluster* berbeda (Fajar dkk., 2024). DBI dapat digunakan untuk mengevaluasi seberapa baik pengelompokan yang dihasilkan oleh metode GKFCM.

Penelitian yang dilakukan oleh Syam tahun 2021 menggunakan metode FCM dalam pengelompokan kabupaten/kota Sulawesi Selatan berdasarkan jumlah tenaga kesehatan menghasilkan 6 *cluster* dengan nilai *pseudo F-statistic* sebesar 16,1843 namun kota Makassar dan Bone dikategorikan sebagai *outlier*. Penelitian lainnya dilakukan oleh Chang-Chien dkk. tahun 2021 menggunakan FCM dan GKFCM pada beberapa data berbeda, hasil dari penelitian ini menunjukkan bahwa metode GKFCM memperoleh nilai *average error* sebesar 0,29 lebih baik dibandingkan FCM dengan *average error* sebesar 0,34. Gosain dan Dahiya tahun 2020 juga telah melakukan penelitian pada data sintesis yang mengandung *outlier*, memperoleh hasil nilai *average error* sebesar 65,32 pada FCM lalu 6,2 pada FCM berbasis intuisi dan *gaussian kernel*, menunjukkan bahwa *gaussian kernel* membantu model menangani pengaruh dari data *outlier* pada proses *clustering*.

Berdasarkan uraian latar belakang yang telah dipaparkan serta penelitian sebelumnya, menunjukkan bahwa metode FCM mampu melakukan klasifikasi pada data dengan kompleksitas tertentu namun metode FCM sangat sensitif terhadap *outlier* dan *noise*. Pengembangan metode FCM yaitu GKFCM menjadi solusi pada kelemahan FCM terhadap *outlier* dan *noise* menggunakan pendekatan *kernel*, sehingga pada penelitian tugas akhir ini akan dilakukan pengelompokan kabupaten/kota di Sulawesi Selatan menggunakan metode GKFCM pada data kriminalitas setiap kabupaten/kota di Sulawesi Selatan.

## 1.2 Rumusan Masalah

Rumusan masalah dari penulisan tugas akhir ini adalah sebagai berikut:

1. Bagaimana menentukan jumlah *cluster* terbaik menggunakan nilai *Deviasi Bouldin Index* serta menentukan karakteristik setiap kelompok kabupaten/kota di Provinsi Sulawesi Selatan yang terbentuk berdasarkan data kriminalitas menggunakan metode *Gaussian Kernel Fuzzy C-Means*?

## 1.3 Batasan Masalah

Batasan Fokus penelitian ini akan dibatasi pada data kriminalitas tahun 2022 Provinsi Sulawesi Selatan. yang diperoleh dari buku Publikasi BPS "Provinsi Sulawesi Selatan Dalam Angka Tahun 2023". Indikator yang digunakan dalam penelitian ini adalah Jumlah Kejahatan yang Dilaporkan, Risiko Penduduk Terkena Kejahatan per 100.000 Penduduk, Persentase Penyelesaian Kejahatan dan Selang Waktu Terjadinya Tindak Kejahatan.

## 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Memperoleh jumlah *cluster* terbaik menggunakan nilai *Deviasi Bouldin Index* serta mendapatkan karakteristik setiap kelompok kabupaten/kota di Provinsi Sulawesi Selatan yang terbentuk berdasarkan data kriminalitas menggunakan metode *Gaussian Kernel Fuzzy C-Means*.

## 1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan dapat membantu dalam memahami lebih baik bagaimana penduduk di kabupaten/kota di Provinsi Sulawesi Selatan terbagi berdasarkan karakteristik tindak kriminalitas. Menggunakan metode *Gaussian Kernel Fuzzy C-Means*, penelitian ini dapat memberikan pengelompokan yang lebih akurat dari metode tradisional. Hasil penelitian ini akan membantu pemerintah dalam mengidentifikasi *cluster* populasi yang memiliki karakteristik yang serupa.

## 1.6 Landasan Teori

### 1.6.1 Analisis Cluster

Analisis *cluster* merupakan salah satu metode data *mining* yang bersifat unsupervised yang memiliki arti bahwa karakteristik tiap *cluster* tidak ditentukan sebelumnya melainkan berdasarkan kemiripan atribut-atribut dari suatu kelompok atau *cluster*. Analisis *cluster* membagi data menjadi kelompok-kelompok atau *cluster* berdasarkan suatu kemiripan atribut-atribut diantara sekumpulan data, atribut-atribut yang mirip tersebut dipresentasikan sebagai titik-titik dalam ruang multidimensi (Andini dan Farokhah, 2022).

Analisis *cluster* merupakan metode multivariat yang mempunyai tujuan untuk mengelompokkan objek berdasarkan kesamaan karakteristik. Analisis *cluster* mengklasifikasi objek sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam *cluster* yang sama (Ferezagia, 2018). Pembuatan *cluster* pada wilayah sebuah negara memungkinkan pemerintah atau pengurus wilayah untuk lebih memahami kinerjanya dalam mengelola aspek atau variabel yang mempengaruhi kehidupan masyarakat, sehingga pemerintah dapat menentukan beberapa strategi yang efektif dalam rangka meningkatkan kualitas kehidupan masyarakat daerah tersebut (Handoyono, 2022).

### 1.6.2 Fuzzy C-Means Clustering

Salah satu algoritma *fuzzy clustering* yang paling banyak digunakan adalah Algoritma *Fuzzy C-Means* (FCM). Algoritma FCM didasarkan pada konsep *fuzzy*,

dimana *fuzzy* merupakan suatu nilai yang dapat bernilai benar atau salah secara bersamaan, namun seberapa besar nilai kebenaran dan kesalahannya tergantung pada derajat keanggotaan yang dimilikinya (Sovina dan Harahap, 2022). Sistem logika *fuzzy* digunakan dalam sebuah sistem yang dibangun pada sebuah fenomena yang bersifat samar-samar. Logika *fuzzy* dapat dianggap sebagai pendekatan untuk memetakan suatu ruang *input* atau masukan ke dalam suatu ruang *output* atau keluaran (Setiawan dkk., 2018). Dasar dari logika *fuzzy* adalah teori himpunan *fuzzy*. Pada teori himpunan *fuzzy*, peranan derajat keanggotaan sebagai penentu keberadaan elemen dalam suatu himpunan sangatlah penting. Nilai keanggotaan biasa disebut derajat keanggotaan atau membership value menjadi ciri utama dari penalaran dengan logika *fuzzy* (Zahida, 2018).

Algoritma FCM mencoba mengelompokkan kumpulan terbatas  $n$  elemen  $X = (x_1, x_2, \dots, x_n)$  kedalam  $c$  kelompok data atau *cluster* dengan menggunakan nilai *fuzzy* yang didasarkan pada jarak data ke pusat *cluster*  $V = (v_1, v_2, \dots, v_c)$  dengan memperhatikan beberapa variabel tertentu. Algoritma FCM memberikan daftar pusat *cluster*  $c$  dan matriks partisi  $U$  dengan nilai  $u_{ij}$  menunjukkan kemungkinan elemen  $x_i$  termasuk dalam *cluster*  $y_j$  (Vignesh, 2019). Matriks partisi  $U$  merupakan matriks berukuran  $n \times c$ ,  $U = [u_{ij}]$  yang merupakan matriks partisi untuk menentukan nilai dari  $\{y_j\}$  atau nilai sebuah data termasuk kedalam kumpulan data dengan ketentuan sebagai berikut (Bezdek dkk., 1988):

$$u_j(x_i) = u_{ij} = \begin{cases} 1; & x_i \in y_j \\ 0 & \text{lainnya} \end{cases}$$

$$\sum_{i=1}^n u_{ij} > 0 \quad \text{untuk semua } i;$$

$$\sum_{j=1}^c u_{ij} = 1 \quad \text{untuk semua } j;$$

Algoritma FCM bertujuan untuk meminimalkan fungsi tujuan yang memiliki derajat *fuzzifier*  $m$ . *Fuzzifier*  $m$  menentukan tingkat ketidakjelasan *cluster* yang pada umumnya nilai  $m$  diatur pada nilai  $m = 2$ . Ketika  $m$  mendekati nilai 1, pusat *cluster* yang paling dekat dengan observasi diberikan nilai yang jauh lebih besar dibandingkan yang lain (Kononenko dan Matjazkucar, 2007). Algoritma FCM memberikan derajat kepemilikan terhadap *cluster* pada setiap titik data, seperti dalam logika *fuzzy*, sehingga setiap titik bukan sepenuhnya hanya terdapat pada satu *cluster*. Titik pengamatan di tepi *cluster* mungkin memiliki tingkat derajat kepemilikan *cluster* yang lebih rendah dibandingkan titik yang lebih dekat dengan pusat *cluster* (Vignesh, 2019).

Algoritma FCM *Clustering* bertujuan untuk membentuk kelompok data yang padat dengan cara meminimalkan fungsi objektifnya. Fungsi objektif ini melibatkan penjumlahan perbedaan berpasangan antara titik data dan pusat *cluster*, dengan

tingkat keanggotaan untuk setiap *cluster* pada matriks  $\mathbf{U}$  dikalikan dengan kuadrat jarak antara titik data dan pusat *cluster*  $\mathbf{V}$ . Proses meminimalkan fungsi pada FCM bertujuan untuk mencari tingkat keanggotaan yang menghasilkan kelompok yang kompak di sekitar pusat data yang signifikan, memungkinkan elemen data untuk termasuk dalam lebih dari satu kelompok dengan tingkat keanggotaan yang dapat diukur secara *fuzzy*. Fungsi objektif FCM ditampilkan pada persamaan (Askari, 2020):

$$J_m(\mathbf{U}, \mathbf{V}) = \sum_{i=1}^n \sum_{j=1}^c (u_{ij})^m (\|x_i - v_j\|)^2; \quad 1 \leq m \leq \infty \quad (1)$$

Keterangan:

- $u_{ij}$  : Nilai keanggotaan data ke- $i$  pada *cluster* ke- $j$
- $v_j$  : Pusat *cluster* ke- $j$
- $x_i$  : Data ke- $i$
- $m$  : Tingkat ketidakpastian (*degree of fuzziness*)
- $n$  : Jumlah banyaknya data
- $c$  : Jumlah banyaknya *cluster*

Nilai  $u_{ij}$  merupakan nilai fungsi keanggotaan dari data yang memiliki nilai diantara 0 dan 1. Persamaan untuk menghitung nilai  $u_{ij}$  untuk setiap data  $x_i$  pada *cluster*  $c_j$  adalah sebagai berikut (Alam dkk., 2019):

$$u_{ij} = \sum_{k=1}^c \left( \frac{(\|x_i - v_j\|)}{(\|x_i - v_k\|)} \right)^{\left( \frac{2}{m-1} \right)} \quad (2)$$

Dengan

$$v_j = \frac{\sum_{i=1}^n (u_{ij})^m x_i}{\sum_{i=1}^n u_{ij}^m} \quad (3)$$

$$\|x_i - v_i\|^2 = \sqrt{\sum (x_i - v_i)^2} \quad (4)$$

Pada persamaan diatas  $\|x_i - v_j\|$  dapat direpresentasikan sebagai jarak *Euclidean*. Algoritma *Fuzzy C-Means* (FCM) mengoptimalkan fungsi objektif  $J_m(\mathbf{U}, \mathbf{V})$  melalui pembaruan berkelanjutan dari  $\mathbf{U}$  dan  $\mathbf{V}$ . Iterasi dilakukan hingga perbedaan antara tingkat keanggotaan pada dua iterasi berturut-turut  $\|U_q - U_{q-1}\| \leq e$ , kurang dari atau sama dengan nilai toleransi  $e$ , dengan  $e$  berada dalam rentang 0 hingga 1, proses ini berlangsung hingga batas toleransi tercapai. (Alam dkk., 2019).

### 1.6.3 Gaussian Kernel

*Gaussian kernel* adalah fungsi matematika yang mengukur kesamaan antara dua titik data berdasarkan jarak *euclidean* antara titik-titik tersebut. *Gaussian kernel* digunakan untuk menentukan seberapa mirip dua titik data berdasarkan titik data pada ruang fitur, nilai fungsi *gaussian kernel* berkurang secara eksponensial seiring dengan bertambahnya jarak antara titik-titik tersebut, memungkinkan *gaussian kernel* ini untuk menangkap hubungan non-linear yang kompleks dalam data (Kowalczyk, 2017). Berikut merupakan persamaan fungsi *gaussian kernel* untuk melihat hubungan antara dua titik data (Jiang dkk., 2020)

$$\psi(x_i, v_j) = \exp\left(\frac{-\|x_i - v_j\|^2}{\sigma^2}\right) \quad (5)$$

Keterangan:

- $u_{ij}$  : Nilai keanggotaan data ke- $i$  pada *cluster* ke- $j$
- $v_j$  : Pusat *cluster* ke- $j$
- $x_i$  : Data ke- $i$
- $\|x_i - v_j\|$  : Jarak *euclidian* data  $x_i$  dari pusat *cluster*  $v_j$
- $\sigma^2$  : Varians dari data

Ketika data terdistribusi tidak merata atau beberapa bagian dari data sangat padat dengan banyak titik data yang berdekatan, sementara bagian lain dari data sangat jarang dengan titik data yang tersebar jauh. Situasi ini sering kali muncul dalam *dataset* nyata di mana pola atau *cluster* data tidak selalu linier atau seragam. *gaussian kernel*, dirancang untuk menangani kompleksitas ini dengan sangat baik. Berikut adalah beberapa alasan mengapa RBF *kernel* penting untuk data yang *gaussian kernel* dapat memetakan data ke ruang fitur yang lebih tinggi di mana pola *non-linear* menjadi lebih mudah diidentifikasi dan dipisahkan. Ini memungkinkan model untuk menangkap kompleksitas dan struktur yang ada dalam data yang tidak terdistribusi secara merata (Rochim et al., 2021).

### 1.6.4 Gaussian Kernel Fuzzy C-Means Clustering

Fungsi objektif standar dari algoritma *Fuzzy C-Means* (FCM) bertujuan untuk membagi himpunan data  $X$  menjadi  $V$  kelompok atau *cluster* dengan cara meminimalkan suatu fungsi matematis tertentu, yang disebut sebagai fungsi objektif. Fungsi objektif FCM ditampilkan pada persamaan (1). Tahapan proses *clustering* GKFCM adalah dengan meminimalkan fungsi objektif pada persamaan dapat dituliskan sebagai berikut:

$$J_m(\mathbf{U}, \mathbf{V}) = 2 \sum_{i=1}^n \sum_{j=1}^c (u_{ij})^m (1 - \psi(x_i, v_j)) \quad (6)$$

Dengan batasan

$$\sum_{i=1}^c u_{ij} = 1$$

Keterangan:

- $\psi(x_i, v_j)$  : Nilai *gaussian kernel* data  $x_i$  dengan pusat *cluster*  $v_j$   
 $u_{ij}$  : Nilai keanggotaan data ke- $i$  pada *cluster* ke- $j$   
 $v_j$  : Pusat *cluster* ke- $j$   
 $x_i$  : Data ke- $i$   
 $m$  : Tingkat ketidakpastian (*degree of fuzziness*)  
 $n$  : Jumlah banyaknya data  
 $c$  : Jumlah banyaknya *cluster*

Fungsi *gaussian*  $\psi(x_i, v_k)$  yang dimasukkan kedalam persamaan untuk mengevaluasi jarak merupakan fungsi *gaussian* pada Persamaan (6). Persamaan yang digunakan pada tiap iterasi untuk memperbarui nilai matriks partisi dan pusat *cluster* sehingga dapat meminimalkan fungsi objektif sebagai berikut:

$$u_{ij} = \frac{\left( (1 - \psi(x_i, v_j))^{-1} \right)^{\frac{1}{m-1}}}{\sum_{k=1}^c \left( \frac{1}{1 - \psi(x_i, v_k)} \right)^{\frac{1}{m-1}}} \quad (7)$$

$$v_j = \frac{\sum_{i=1}^n u_{ij}^m \psi(x_i, v_j) x_i}{\sum_{i=1}^n u_{ij}^m \psi(x_i, v_j)} \quad (8)$$

Dengan menggunakan fungsi *gaussian* ini, kita dapat mengevaluasi jarak antara data point ( $x_i$ ) dan pusat *cluster* ( $v_j$ ) dengan memasukkan nilai-nilai tersebut ke dalam fungsi *gaussian*. Ini memberikan representasi jarak yang mempertimbangkan distribusi *gaussian* antara *data point* dan pusat *cluster*, yang sesuai dengan pendekatan *gaussian* yang diadopsi (Jiang dkk., 2020).

### 1.6.5 Davies-Bouldin Index

David L. Davies dan Donald W. Bouldin memperkenalkan sebuah metode dan memberikan nama pada metode tersebut dengan nama kedua beliau, yaitu *Davies-Bouldin Index* (DBI) yang digunakan untuk mengevaluasi *cluster*. Evaluasi menggunakan DBI memiliki skema evaluasi *cluster* internal, di mana hasil *cluster* yang baik atau tidak dilihat dari kuantitas dan kedekatan antara hasil *cluster*. DBI adalah satu metode yang digunakan untuk mengukur validitas *cluster* dalam suatu metode pengelompokan, di mana hubungan antar data dalam kelompok didefinisikan sebagai jumlah kedekatan data ke titik pusat *cluster* diikuti oleh pemisahan berdasarkan jarak antara titik pusat *cluster* ke titik pusat *cluster* lainnya. Pengukuran dengan DBI memaksimalkan jarak antar *cluster* misalnya jarak antara *cluster*  $C_i$  dan  $C_j$  dan pada saat yang sama mencoba untuk meminimalkan jarak antara titik-titik



dalam suatu *cluster*. Jika jarak antar *cluster* maksimal, itu berarti bahwa kesamaan karakteristik antara setiap *cluster* bernilai kecil sehingga perbedaan antara *cluster* tampak lebih jelas. Jika jarak antar data dalam *cluster* minimum atau data dalam *cluster* yang sama saling berdekatan, itu berarti bahwa setiap objek dalam *cluster* memiliki tingkat kesamaan karakteristik yang tinggi (Mughnyanti dkk., 2020).

Prinsip pendekatan pengukuran DBI adalah memaksimalkan jarak antar *cluster* serta meminimalkan jarak dalam *cluster*. Semakin kecil nilai DBI menunjukkan skema *cluster* yang paling optimal. Perhitungan DBI dijabarkan dalam persamaan berikut (Gie dan Jollyta, 2020):

$$DBI = \frac{1}{c} \sum_{j=1}^c \max_{j \neq k} R_{j,k} \quad (9)$$

Dengan

$$R_{j,k} = \frac{MAE_j + MAE_k}{\|\mathbf{d}_j - \mathbf{d}_k\|} \quad (10)$$

$$MAE_j = \left( \frac{1}{T_j} \sum_{x \in C_j} \|\mathbf{x}_i - \mathbf{d}_j\| \right) \quad (11)$$

Keterangan

$DBI$	: <i>Davies-bouldin index</i>
$MAE_j$	: Jarak intra- <i>cluster</i> ke- $j$
$\mathbf{d}_j$	: Nilai rata-rata <i>cluster</i> ke- $j$
$T_j$	: Jumlah data pada <i>cluster</i> ke- $j$
$R_{j,k}$	: Kesamaan antar <i>cluster</i> ke- $j$ dan ke- $k$
$c$	: Jumlah <i>cluster</i>
$\ \mathbf{d}_j - \mathbf{d}_k\ $	: Jarak <i>euclidian mean cluster</i> ke- $j$ dan <i>cluster</i> ke- $k$
$\ \mathbf{x}_i - \mathbf{d}_j\ $	: Jarak <i>euclidian mean cluster</i> ke- $j$ dan data ke- $i$

### 1.6.6 Average Silhouette Value

*Average Silhouette Value* (ASV) diperkenalkan oleh Peter J. Rousseeuw pada tahun 1986, adalah metode yang digunakan untuk memvalidasi data *cluster*. ASV memberikan representasi yang menunjukkan seberapa baik setiap titik data sesuai dengan *cluster* nya (Subbalakshmi dkk., 2015). ASV adalah suatu nilai yang mampu mengukur tingkat kehomogenan anggota dalam internal *cluster* dan tingkat keheterogenan anggota antar *cluster*. Nilai ASV evaluasi *cluster* berada pada interval -1 hingga 1 (Septianingsih, 2022). Nilai *silhouette value* mendekati 1,

menunjukkan bahwa titik data secara signifikan lebih dekat dengan titik data dalam *cluster* yang sama dibandingkan dengan titik data *cluster* lain. Ketika nilai mendekati 0, menandakan adanya ketidakpastian dalam penempatan titik data tersebut, dengan kemungkinan bahwa data tersebut bisa masuk ke dalam *cluster* lain yang memiliki karakteristik serupa. Sementara itu, nilai *silhouette value* yang mendekati -1 mengindikasikan penempatan yang salah, dengan titik data lebih dekat dengan *cluster* lain daripada yang seharusnya, mengindikasikan ketidakcocokan antara titik data dan *cluster*nya (Shahapure dan Nicholas, 2020). Berikut merupakan Tabel 2.1 yang menampilkan standar kriteria evaluasi *cluster* menggunakan nilai ASV menurut (Guntara dan Suprawoto, 2022):

**Tabel 1.** Kriteria Penilaian *Clustering* Berdasarkan *Average Silhouette Value*

<b>Average Silhouette Value</b>	<b>Kriteria</b>
$0,7 < ASV \leq 1$	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>

Nilai ASV dapat dihitung setelah setiap data dikelompokan dan nilai tengah masing-masing *cluster* telah dihitung. Persamaan untuk menghitung ASV adalah sebagai berikut (Septianingsih, 2022):

$$ASV = \frac{1}{N} \sum_{i=1}^N s(i) \quad (12)$$

dengan:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (13)$$

$$a(i) = \frac{1}{nc - 1} \sum_{r=1}^{nc-1} d(x_{ic} - x_{rc}) \quad (14)$$

$$d(i, C) = \frac{1}{nk} \sum_{r=1}^{nk} d(x_i - x_{rc}) \quad (15)$$

$$b(i) = \min d(i, C) \quad (16)$$

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*

### 1.6.7 Kriminalitas

Tindak kriminalitas adalah suatu tindakan yang melanggar hukum, norma, dan peraturan sosial yang berlaku dalam suatu kelompok masyarakat, sehingga bagi yang melakukan pelanggaran harus diberikan sanksi atau penghukuman. Tindak kriminalitas sangat berdampak negatif terhadap kesejahteraan masyarakat di sebuah wilayah, karena dapat menimbulkan rasa tidak aman, kecemasan, kepanikan, dan ketakutan. Tingginya kasus kriminalitas yang terjadi disuatu wilayah dapat menggambarkan besarnya risiko setiap masyarakat menjadi korban kriminalitas, risiko tersebut dapat dinyatakan dalam parameter tingkat kejahatan atau kriminalitas (Sulistiani, 2022). Secara umum kriminalitas dapat dijelaskan dengan indikator sebagai berikut:

#### 1. Jumlah Kejadian yang Dilaporkan

Jumlah kejadian yang dilaporkan adalah setiap peristiwa yang dilaporkan masyarakat pada Polri, atau peristiwa dimana pelakunya tertangkap tangan oleh kepolisian. Laporan masyarakat ini akan dicatat dan ditindaklanjuti oleh Polri jika dikategorikan memiliki cukup bukti. Kejadian kejahatan yang dilaporkan terdiri atas kejahatan terhadap nyawa, fisik/badan, kesusilaan, hak milik/barang, narkoba, ketertiban umum, serta penipuan, penggelapan, dan korupsi (BPS, 2016).

#### 2. Risiko penduduk Terkena Kejahatan per 100.000 Penduduk (*Crime Rate*)

*Crime rate* merupakan angka yang dapat menunjukkan tingkat kerawanan terjadinya tindak kejahatan pada suatu wilayah tertentu dalam waktu tertentu. Semakin tinggi angka *crime rate* maka tingkat kerawanan akan kejahatan suatu daerah semakin tinggi pula, dan sebaliknya. Selama periode tahun 2018–2020, tingkat resiko terkena tindak kejahatan (*crime rate*) setiap 100.000 penduduk mengalami penurunan (BPS, 2021). Perhitungan Risiko penduduk Terkena Kejahatan per 100.000 Penduduk adalah sebagai berikut (Kepala Kepolisian Negara Republik Indonesia, 2009):

$$crime\ rate = \frac{Jumlah\ Kejadian\ yang\ Dilaporkan}{Jumlah\ Penduduk} \times 100.000$$

#### 3. Persentase Penyelesaian Tindak Pidana (*Clearance Rate*)

*Clearance rate* adalah proporsi kejahatan yang diketahui polisi yang diselesaikan (penangkapan dilakukan) versus semua kejahatan yang diketahui polisi. Dalam peradilan pidana, tingkat penyelesaian dihitung dengan membagi jumlah kejahatan yang "dibersihkan" (dikenakan dakwaan) dengan jumlah total kejahatan yang tercatat. Tingkat penyelesaian digunakan oleh berbagai kelompok sebagai ukuran

kejahatan yang diselesaikan oleh polisi. Tingginya presentase *clearance rate* oleh pihak kepolisian belum mampu menekan angka kejahatan. Ketika angka *clearance rate* meningkat hal tersebut menunjukkan tingkat keberhasilan dari pihak kepolisian dalam menyelesaikan kasus yang dilaporkan (BPS, 2016). Perhitungan Persentase Penyelesaian Tindak Pidana adalah sebagai berikut (Kepala Kepolisian Negara Republik Indonesia, 2009):

$$\text{Clearance Rate} = \frac{\text{Kasus yang diselesaikan}}{\text{Jumlah Kejadian yang Dilaporkan}} \times 100\%$$

#### 4. Selang Waktu Terjadinya Tindak Pidana (*Crime Clock*)

*Crime clock* menunjukkan pada beberapa kali kejahatan yang terjadi pada setiap jam. Selang waktu terjadinya kejahatan di Indonesia selama 2 menit 11 detik pada 2021. Semakin tinggi *crime clock* maka semakin aman pula daerah tersebut dari tindak kejahatan (BPS, 2021). Perhitungan Selang Waktu Terjadinya Tindak Pidana adalah sebagai berikut (Kepala Kepolisian Negara Republik Indonesia, 2009):

$$\text{Crime Clock} = \frac{365 \times 24}{\text{Jumlah Kejadian yang Dilaporkan}}$$

## BAB II METODOLOGI PENELITIAN

### 2.1 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang bersumber dari Badan Pusat Statistik (BPS) Provinsi Sulawesi Selatan yaitu data data Kriminalitas Sulawesi Selatan tahun 2022 pada publikasi Provinsi Sulawesi Selatan Dalam Angka 2023 yang dapat diakses melalui situs resmi BPS Provinsi Sulawesi Selatan pada tautan <https://sulsesel.bps.go.id>.

### 2.2 Struktur Data

Struktur data yang digunakan berupa data *cross section* dengan 24 observasi yang meliputi 24 kabupaten/kota di Sulawesi Selatan. Keseluruhan data yang digunakan terdapat pada Lampiran 1. Adapun variabel-variabel yang digunakan pada penelitian ini dapat dilihat pada Tabel 2.

**Tabel 2.** Struktur Variabel Data

Variabel	Keterangan	Satuan
X1	Jumlah Kejahatan yang Dilaporkan	Kasus
X2	Risiko Penduduk Terkena Kejahatan per 100.000 Penduduk	Penduduk
X3	Persentase Penyelesaian Kejahatan	Persen
X4	Selang Waktu Terjadinya Kejahatan	Jam

### 2.3 Metode Analisis

Pendekatan analisis yang digunakan dalam penelitian ini adalah Pengelompokan dengan metode GKFCM. Pengolahan data dilakukan dengan bantuan *Software Jupyter Notebook*. Tahapan analisis data yang dilakukan dalam penelitian ini yaitu sebagai berikut:

1. Menentukan jumlah *cluster* terbaik menggunakan nilai *Davies-Bouldin Index* yang diperoleh dengan menggunakan Persamaan (9) yang membutuhkan nilai pusat dan pengelompokan *cluster* dari tiap observasi. Penentuan pusat *cluster* dan pengelompokan *cluster* tiap observasi dilakukan sebagai berikut:
  - a. Menentukan jumlah *cluster* yang akan digunakan dalam proses *clustering*.
  - b. Menentukan nilai titik pusat *cluster* dengan menggunakan nilai acak sebagai inisiasi awal pusat *cluster* untuk iterasi pertama dan Persamaan (8) pada iterasi kedua dan seterusnya.
  - c. Menghitung nilai *kernel* setiap titik observasi dengan setiap titik pusat *cluster* menggunakan Persamaan (6) sehingga diperoleh matriks berisi nilai

- kernel gaussian* antara setiap titik pusat *cluster* dengan setiap observasi.
- d. Nilai *kernel* dan nilai pusat *cluster* yang telah diperoleh digunakan untuk memperbarui nilai matriks keanggotaan *fuzzy U* dengan Persamaan (7). Nilai matriks keanggotaan *fuzzy* yang telah diperbarui digunakan untuk proses iterasi selanjutnya.
  - e. Langkah b hingga d terus diiterasikan hingga batas iterasi telah tercapai atau nilai  $\|U_q - U_{q-1}\| \leq e$ .
  - f. Menghitung nilai *Davies-Bouldin Index* menggunakan Persamaan (9), lalu mengulangi seluruh proses dari langkah a hingga b menggunakan nilai *cluster* yang berbeda
2. Mengevaluasi kualitas *cluster* hasil pengelompokan GKFCM menggunakan nilai ASV.
  3. Setelah memperoleh jumlah *cluster* terbaik, maka selanjutnya pengelompokan yang telah diperoleh digunakan untuk menggambarkan distribusi pengelompokan kabupaten/kota berdasarkan data kriminalitas pada peta Sulawesi Selatan.
  4. Hasil yang diperoleh diinterpretasikan lalu dibuatkan kesimpulan.

## BAB III HASIL DAN PEMBAHASAN

### 3.1 Penentuan Jumlah *Cluster* Menggunakan *Gaussian Kernel Fuzzy C-Means Clustering* dan *Davies-Bouldin Index*

*Davies-Bouldin Index* (DBI) adalah salah satu metrik evaluasi yang digunakan untuk mengukur kualitas dan kesamaan antara *cluster* dalam proses analisis *cluster*. DBI dihitung dengan cara membandingkan sejauh mana setiap *cluster* mirip dengan *cluster* terdekat lainnya. Skor tersebut dapat digunakan untuk mengevaluasi seberapa baik sebuah algoritma *cluster* menempatkan pusat *cluster* dan seberapa baik setiap *cluster* terpisah satu sama lain. DBI digunakan untuk menentukan jumlah *cluster* dalam proses analisis *cluster*. Semakin rendah nilai DBI, maka semakin baik kualitas *cluster* yang diperoleh dari hasil pengelompokan.

Skor DBI mencapai nilai minimum saat setiap *cluster* memiliki kemiripan yang tinggi dengan anggota *cluster* terdekatnya dan memiliki jarak yang besar dengan *cluster* lainnya. Penentuan nilai DBI dari GKFCM dalam proses pengelompokan kabupaten/kota di Sulawesi Selatan berdasarkan data kriminalitas dilakukan dengan menggunakan Persamaan (9) sebagai berikut:

$$DBI = \frac{1}{c} \sum_{j=1}^c \max_{j \neq k} R_{j,k}$$

Dengan

$$R_{j,k} = \frac{MAE_j + MAE_k}{\|d_j - d_k\|}$$

$$MAE_j = \left( \frac{1}{T_j} \sum_{x \in C_j} \|x_i - d_j\| \right)$$

Nilai DBI digunakan untuk mengetahui kualitas hasil pembagian *cluster* dari metode GKFCM, sehingga dalam perhitungan DBI dibutuhkan nilai rata-rata setiap kelompok yang terbentuk dari metode GKFCM. Dalam penelitian ini perhitungan DBI diletakkan pada langkah awal untuk mengetahui jumlah *cluster* paling optimal yang akan digunakan dan dijabarkan untuk mengelompokkan data serta hasil *cluster* paling optimal yang akan diinterpretasikan kesimpulan penelitian. Perhitungan nilai DBI untuk setiap jumlah pembagian kelompok *cluster* pada metode GKFCM ditampilkan pada bagian berikutnya.

#### 3.1.1 *Gaussian Kernel Fuzzy C-Means* dengan Dua *Cluster*

Analisis *cluster* dengan metode *Gaussian Kernel Fuzzy C-Means* (GKFCM) merupakan merupakan variasi dari algoritma *Fuzzy C-Means* (FCM) yang lebih umum, tetapi dengan modifikasi distribusi *Gaussian* pada perhitungan matriks keanggotaan. Proses *clustering* GKFCM dilakukan dengan memperbarui titik pusat



*cluster*  $v_j$  dan nilai keanggotaan *fuzzy*  $u_{ij}$  dalam setiap langkah iterasi hingga diperoleh hasil yang maksimal. Iterasi algoritma GKFCM dihentikan ketika nilai  $\|U_q - U_{q-1}\| \leq e$ , pada penelitian ini nilai  $e = 10^{-9}$ .

Persamaan (8) digunakan untuk memperbarui nilai  $v_j$  untuk tiap *cluster* dan Persamaan (7) digunakan untuk memperbarui nilai keanggotaan  $u_{ij}$  dari setiap data. Pada Persamaan (7) dibutuhkan nilai  $v_{ij}$  dan nilai *kernel gaussian* dari data dan pusat *cluster*  $\psi(x_i, v_j)$ . Sehingga sebelum memperbarui nilai matriks keanggotaan  $u_{ij}$  terlebih dahulu dihitung nilai  $v_{ij}$  menggunakan Persamaan (8) dan  $\psi(x_i, v_j)$  menggunakan Persamaan (5). Langkah awal untuk memulai iterasi GKFCM adalah dengan melakukan penentuan nilai acak untuk nilai  $v_j$  pada iterasi Ke-1 karena Persamaan (8) tidak dapat digunakan sebelum diperoleh nilai keanggotaan  $u_{ij}$  untuk setiap data. Nilai acak untuk  $v_j$  pada iterasi Ke-1 ditampilkan dalam Tabel 3 sebagai berikut:

**Tabel 3.** Titik Pusat *Cluster* Pada Iterasi Ke-1 GKFCM Dua *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	474,39	182,58	77,37	24,24
2	502,21	167,57	72,93	27,69

Nilai pusat *cluster*  $v_j$  pada Tabel 3 kemudian digunakan untuk menghitung nilai *kernel gaussian* antara titik pusat *cluster*  $v_j$  dan data. Persamaan (5) digunakan untuk menghitung nilai *kernel gaussian* titik pusat dengan data GKFCM pada iterasi ke-1, berikut adalah proses perhitungan nilai *kernel gaussian* pada metode GKFCM.

$$\psi(x_i, v_j) = \exp\left(\frac{-\|x_i - v_j\|^2}{\sigma^2}\right)$$

$$\psi(x_i, v_j) = \exp\left(\frac{-\left(\sqrt{(x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2}\right)^2}{\sigma^2}\right)$$

$$\psi(x_i, v_j) = \exp\left(\frac{-\left((x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2\right)}{\sigma^2}\right)$$

Nilai  $\sigma^2$  merupakan nilai varians dari data, berdasarkan data penelitian pada Lampiran 1, maka diperoleh nilai  $\sigma^2 = 1330877$ . Dengan Persamaan (5) diatas diperoleh nilai *gaussian kernel* yang ditampilkan pada Lampiran 2. Hasil Persamaan (5) yang telah diperoleh digunakan untuk menghitung nilai keanggotaan *fuzzy* untuk setiap data dengan Persamaan (7) sebagai berikut:

**Tabel 4.** Matriks Keanggotaan *Fuzzy* GKFCM Dua *Cluster* Pada Iterasi Ke-1

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>	
	1	2
Bantaeng	0,53	0,47
Barru	0,54	0,46
Bone	0,49	0,51
Bulukumba	0,49	0,51
Enrekang	0,53	0,47
Gowa	0,50	0,50
Jeneponto	0,47	0,53
Kepulauan Selayar	0,57	0,43
Kota Makassar	0,50	0,50
Kota Palopo	0,49	0,51
Kota Parepare	0,57	0,43
Luwu	0,43	0,57
Luwu Timur	0,57	0,43
Luwu Utara	0,46	0,54
Maros	0,45	0,55
Pangkajene Dan Kepulauan	0,54	0,46
Pinrang	0,48	0,52
Sidenreng Rappang	0,10	0,90
Sinjai	0,61	0,39
Soppeng	0,56	0,44
Takalar	0,48	0,52
Tana Toraja	0,55	0,45
Toraja Utara	0,58	0,42
Wajo	0,38	0,62

Iterasi dilanjutkan ke langkah selanjutnya setelah memperoleh nilai matriks keanggotaan *fuzzy*  $u_{ij}$  menggunakan Persamaan (7). Nilai  $u_{ij}$  yang diperoleh pada iterasi ke-1 akan digunakan untuk menghitung nilai  $v_j$  pada iterasi ke-2. Menggunakan Persamaan (8) sehingga diperoleh nilai titik pusat *cluster*  $v_j$  GKFCM iterasi ke-2 yang ditampilkan pada Tabel 5 yang memperlihatkan nilai titik pusat dari setiap *cluster* dengan  $v_1$  sebagai titik pusat cluster 1 dan seterusnya, sebagai berikut:

**Tabel 5.** Titik Pusat *Cluster* Pada Iterasi Ke-2 GKFCM Dua *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	515,36	186,43	74,45	25,56
2	594,03	194,24	72,37	21,37

Setelah diperoleh nilai titik pusat *cluster* pada Tabel 5, langkah-langkah perhitungan berikutnya sama dengan langkah perhitungan pada iterasi sebelumnya yaitu memperbarui nilai *kernel gaussian* antar data dengan titik pusat *cluster*, memperbarui nilai keanggotaan *fuzzy* dan memperbarui titik pusat *cluster* untuk iterasi selanjutnya. Iterasi dilanjutkan hingga diperoleh hasil maksimal pada iterasi langkah ke-100 dengan nilai  $\|U_{100} - U_{99}\| = 8,53 \times 10^{-10}$  karena nilai  $\|U_{100} - U_{99}\| \leq 10^{-9}$  maka dapat disimpulkan bahwa algoritma GKFCM telah mencapai konvergensi sehingga nilai  $v_j$  dan  $u_{ij}$  yang diperoleh merupakan nilai yang paling optimal. Tabel 6 menampilkan nilai hasil akhir titik pusat *cluster* dari algoritma GKFCM yang diperoleh pada iterasi ke-100 sebagai berikut:

**Tabel 6.** Titik Pusat *Cluster* Pada Iterasi Ke-100 GKFCM Dua *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	360,24	146,79	75,82	30,10
2	1157,01	328,29	68,09	8,18

Pada iterasi ke-100 diperoleh tabel nilai keanggotaan *fuzzy* paling optimal yang menunjukkan derajat keanggotaan masing-masing data pada tiap *cluster*. Data termasuk ke dalam sebuah *cluster* yang nilai derajat keanggotaan *fuzzy* paling tinggi. Tabel 7 menampilkan nilai derajat keanggotaan *fuzzy* serta nilai *cluster* untuk masing-masing data.

**Tabel 7.** Matriks Keanggotaan *Fuzzy* GKFCM Dua *Cluster* Pada Iterasi Ke-100

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>		<i>Cluster</i>
	1	2	
Bantaeng	0,93	0,07	1
Barru	0,97	0,03	1
Bone	0,20	0,80	2
Bulukumba	0,05	0,95	2
Enrekang	0,94	0,06	1
Gowa	0,38	0,62	2
Jeneponto	0,27	0,73	2

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>		Cluster
	1	2	
Kepulauan Selayar	0,98	0,02	1
Kota Makassar	0,50	0,50	1
Kota Palopo	0,12	0,88	2
Kota Parepare	0,91	0,09	1
Luwu	0,67	0,33	1
Luwu Timur	1,00	0,00	1
Luwu Utara	0,54	0,46	1
Maros	0,53	0,47	1
Pangkajene Dan Kepulauan	0,98	0,02	1
Pinrang	0,14	0,86	2
Sidenreng Rappang	0,95	0,05	1
Sinjai	1,00	0,00	1
Soppeng	0,99	0,01	1
Takalar	0,05	0,95	2
Tana Toraja	0,99	0,01	1
Toraja Utara	1,00	0,00	1
Wajo	0,83	0,17	1

Tabel 7 merupakan nilai matriks keanggotaan *fuzzy* algoritma GKFCM. Matriks keanggotaan *fuzzy* merupakan representasi dari tingkat sejauh mana setiap titik data tergabung dengan setiap *cluster*. Matriks keanggotaan *fuzzy* memungkinkan GKFCM untuk menangani ketidakpastian atau ketidakjelasan dalam pemilihan *cluster*, dengan menentukan derajat keanggotaan setiap titik data yang beragam ke dalam setiap *cluster* yang ada. Pembagian *cluster* yang terdapat pada Tabel 7 kemudian dihitung rata-rata tiap clusternya yang akan digunakan dalam menghitung nilai DBI GKFCM dua *cluster* dengan Persamaan (9). Sebagai berikut:

$$DBI = \frac{1}{c} \sum_{j=1}^c \max_{j \neq k} R_{j,k}$$

Dengan

$$R_{j,k} = \frac{MAE_j + MAE_k}{\|d_j - d_k\|}$$

$$MAE_j = \left( \frac{1}{T_j} \sum_{x \in C_j} \|x_i - d_j\| \right)$$

Sehingga diperoleh nilai DBI nilai DBI GKFCM dua *cluster*. Sebagai berikut:

**Tabel 8.** Nilai DBI GKFCM Dengan Dua *cluster*

Jumla <i>Cluster</i>	Nilai <i>Davies Bouldin Index</i>
2	4,744

Nilai DBI yang diperoleh pada Tabel 8 akan dibandingkan oleh nilai DBI pada pembagian *cluster* GKFCM yang lain. Nilai DBI yang paling rendah akan menunjukkan bahwa GKFCM dengan nilai *cluster* paling rendah tersebut merupakan jumlah kemungkinan *cluster* yang paling optimal digunakan pada analisis GKFCM untuk Data Kriminal Provinsi Sulawesi Selatan Tahun 2022.

### 3.1.2 Gaussian Kernel Fuzzy C-Means dengan Tiga *Cluster*

Proses GKFCM dilanjutkan dengan menggunakan pembagian tiga *cluster*. Langkah awal untuk memulai iterasi GKFCM sama seperti pada bagian sebelumnya yaitu dengan melakukan penentuan nilai acak untuk nilai  $v_j$  pada iterasi Ke-1 karena Persamaan (8) tidak dapat digunakan sebelum diperoleh nilai keanggotaan  $u_{ij}$  untuk setiap data. Nilai acak  $v_j$  pada iterasi Ke-1 ditampilkan pada Tabel 9 sebagai berikut:

**Tabel 9.** Titik Pusat *Cluster* Pada Iterasi Ke-1 GKFCM Tiga *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	415,53	174,27	81,61	29,83
2	484,55	173,98	72,66	25,36
3	521,29	172,64	71,13	24,48

Nilai pusat *cluster*  $v_j$  pada Tabel 9 kemudian digunakan untuk menghitung nilai *kernel gaussian* antara titik pusat *cluster*  $v_j$  dan data. Persamaan (5) digunakan untuk menghitung nilai *kernel gaussian* titik pusat dengan data GKFCM pada iterasi ke-1, berikut adalah proses perhitungan nilai *kernel gaussian* pada metode GKFCM.

$$\psi(x_i, v_j) = \exp\left(\frac{-\|x_i - v_j\|^2}{\sigma^2}\right)$$

$$\psi(x_i, v_j) = \exp\left(\frac{-\left(\sqrt{(x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2}\right)^2}{\sigma^2}\right)$$

$$\psi(x_i, v_j) = \exp\left(\frac{-\left((x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2\right)}{\sigma^2}\right)$$

Nilai  $\sigma^2$  merupakan nilai varians dari data, berdasarkan data penelitian pada Lampiran 1, maka diperoleh nilai  $\sigma^2 = 1330877$ . Dengan Persamaan (5) diatas diperoleh nilai *gaussian kernel* yang ditampilkan pada Lampiran 3. Hasil Persamaan (5) yang telah diperoleh digunakan untuk menghitung nilai keanggotaan *fuzzy* untuk setiap data dengan Persamaan (7) sebagai berikut:

**Tabel 10.** Matriks Keanggotaan *Fuzzy* GKFCM Tiga *Cluster* Pada Iterasi Ke-1

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>		
	1	2	3
Bantaeng	0,44	0,30	0,26
Barru	0,50	0,28	0,22
Bone	0,32	0,34	0,35
Bulukumba	0,30	0,34	0,36
Enrekang	0,45	0,30	0,25
Gowa	0,33	0,33	0,34
Jeneponto	0,25	0,34	0,41
Kepulauan Selayar	0,55	0,26	0,19
Kota Makassar	0,33	0,33	0,33
Kota Palopo	0,31	0,34	0,36
Kota Parepare	0,33	0,36	0,31
Luwu	0,19	0,33	0,48
Luwu Timur	0,66	0,21	0,13
Luwu Utara	0,22	0,34	0,45
Maros	0,22	0,34	0,45
Pangkajene Dan Kepulauan	0,51	0,28	0,21
Pinrang	0,27	0,34	0,39
Sidenreng Rappang	0,03	0,58	0,39
Sinjai	0,81	0,12	0,07
Soppeng	0,57	0,25	0,18
Takalar	0,28	0,34	0,38
Tana Toraja	0,53	0,27	0,20
Toraja Utara	0,68	0,20	0,13
Wajo	0,12	0,30	0,58

Iterasi dilanjutkan ke langkah selanjutnya setelah memperoleh nilai matriks keanggotaan *fuzzy*  $u_{ij}$  menggunakan Persamaan (7). Nilai  $u_{ij}$  yang diperoleh pada iterasi ke-1 akan digunakan untuk menghitung nilai  $v_j$  pada iterasi ke-2. Menggunakan Persamaan (8) sehingga diperoleh nilai titik pusat *cluster*  $v_j$  GKFCM iterasi ke-2 yang ditampilkan pada Tabel 11 yang memperlihatkan nilai titik pusat dari

setiap *cluster* dengan  $v_1$  sebagai titik pusat *cluster* 1 dan seterusnya, sebagai berikut:

**Tabel 11.** Titik Pusat *Cluster* Pada Iterasi Ke-2 GKFCM Tiga *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	388,21	154,54	75,25	28,99
2	612,55	204,81	73,14	21,28
3	707,74	216,47	69,93	16,89

Setelah diperoleh nilai titik pusat *cluster* pada Tabel 11, langkah-langkah perhitungan berikutnya sama dengan langkah perhitungan pada iterasi sebelumnya yaitu memperbarui nilai *kernel gaussian* antar data dengan titik pusat *cluster*, memperbarui nilai keanggotaan *fuzzy* dan memperbarui titik pusat *cluster* untuk iterasi selanjutnya. Iterasi dilanjutkan hingga tercapai batas iterasi maksimal pada iterasi langkah ke-100 dengan nilai  $\|U_{100} - U_{99}\| = 3,17 \times 10^{-9}$  sehingga nilai  $v_j$  dan  $u_{ij}$  yang diperoleh merupakan nilai yang paling optimal. Tabel 12 menampilkan nilai hasil akhir titik pusat *cluster* dari algoritma GKFCM yang diperoleh pada iterasi ke-100 sebagai berikut:

**Tabel 12.** Titik Pusat *Cluster* Pada Iterasi Ke-100 GKFCM Tiga *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	301,01	134,08	77,34	33,50
2	828,60	247,64	68,47	11,01
3	1694,33	279,03	70,41	5,35

Pada iterasi ke-100 diperoleh tabel nilai keanggotaan *fuzzy* paling optimal yang menunjukkan derajat keanggotaan masing-masing data pada tiap *cluster*. Data termasuk ke dalam sebuah *cluster* yang nilai derajat keanggotaan *fuzzy* paling tinggi. Tabel 13 menampilkan nilai derajat keanggotaan *fuzzy* serta nilai *cluster* untuk masing-masing data.

**Tabel 13.** Matriks Keanggotaan *Fuzzy* GKFCM Tiga *Cluster* Pada Iterasi Ke-100

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>			<i>Cluster</i>
	1	2	3	
Bantaeng	0,90	0,07	0,03	1
Barru	0,98	0,02	0,01	1
Bone	0,00	0,01	0,99	3
Bulukumba	0,09	0,29	0,62	3
Enrekang	0,91	0,06	0,02	1

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>			<i>Cluster</i>
	1	2	3	
Gowa	0,14	0,17	0,70	3
Jeneponto	0,01	0,99	0,00	2
Kepulauan Selayar	0,97	0,02	0,01	1
Kota Makassar	0,33	0,33	0,33	1
Kota Palopo	0,17	0,46	0,38	2
Kota Parepare	0,66	0,30	0,04	1
Luwu	0,13	0,84	0,03	2
Luwu Timur	0,97	0,02	0,00	1
Luwu Utara	0,04	0,95	0,01	2
Maros	0,04	0,95	0,01	2
Pangkajene Dan Kepulauan	0,98	0,01	0,00	1
Pinrang	0,04	0,93	0,03	2
Sidenreng Rappang	0,71	0,26	0,03	1
Sinjai	0,96	0,03	0,01	1
Soppeng	1,00	0,00	0,00	1
Takalar	0,07	0,84	0,08	2
Tana Toraja	0,99	0,00	0,00	1
Toraja Utara	0,99	0,01	0,00	1
Wajo	0,37	0,58	0,04	2

Tabel 13 merupakan nilai matriks keanggotaan *fuzzy* algoritma GKFCM. Matriks keanggotaan *fuzzy* merupakan representasi dari tingkat sejauh mana setiap titik data tergabung dengan setiap *cluster*. Matriks keanggotaan *fuzzy* memungkinkan GKFCM untuk menangani ketidakpastian atau ketidakjelasan dalam pemilihan *cluster*, dengan menentukan derajat keanggotaan setiap titik data yang beragam ke dalam setiap *cluster* yang ada. Pembagian *cluster* yang terdapat pada Tabel 13 kemudian dihitung rata-rata tiap clusternya yang akan digunakan dalam menghitung nilai DBI GKFCM dua *cluster* dengan Persamaan (9). Sebagai berikut:

$$DBI = \frac{1}{c} \sum_{j=1}^c \max_{j \neq k} R_{j,k}$$

Dengan

$$R_{j,k} = \frac{MAE_j + MAE_k}{\|d_j - d_k\|}$$



$$MAE_j = \left( \frac{1}{T_j} \sum_{x \in C_j} \|x_i - d_j\| \right)$$

Sehingga diperoleh nilai DBI nilai DBI GKFCM dua *cluster*. Sebagai berikut:

**Tabel 14.** Nilai DBI GKFCM Dengan Tiga *cluster*

Jumla <i>Cluster</i>	Nilai <i>Davies Bouldin Index</i>
3	4,896

Nilai DBI yang diperoleh pada Tabel 14 akan dibandingkan oleh nilai DBI pada pembagian *cluster* GKFCM yang lain. Nilai DBI yang paling rendah akan menunjukkan bahwa GKFCM dengan nilai *cluster* paling rendah tersebut merupakan jumlah kemungkinan *cluster* yang paling optimal digunakan pada analisis GKFCM untuk Data Kriminal Provinsi Sulawesi Selatan Tahun 2022.

### 3.1.3 Gaussian Kernel Fuzzy C-Means dengan Empat *Cluster*

Proses GKFCM dilanjutkan dengan menggunakan pembagian empat *cluster*. Langkah awal untuk memulai iterasi GKFCM sama seperti pada bagian sebelumnya yaitu dengan melakukan penentuan nilai acak untuk nilai  $v_j$  pada iterasi Ke-1 karena Persamaan (8) tidak dapat digunakan sebelum diperoleh nilai keanggotaan  $u_{ij}$  untuk setiap data. Nilai acak  $v_j$  pada iterasi Ke-1 ditampilkan pada Tabel 15 sebagai berikut:

**Tabel 15.** Titik Pusat *Cluster* Pada Iterasi Ke-1 GKFCM Empat *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	512,02	182,65	67,77	24,44
2	515,98	174,41	77,77	26,24
3	387,17	155,66	76,51	29,53
4	458,85	185,06	73,47	26,28

Nilai pusat *cluster*  $v_j$  pada Tabel 15 kemudian digunakan untuk menghitung nilai *kernel gaussian* antara titik pusat *cluster*  $v_j$  dan data. Persamaan (5) digunakan untuk menghitung nilai *kernel gaussian* titik pusat dengan data GKFCM pada iterasi ke-1, berikut adalah proses perhitungan nilai *kernel gaussian* pada metode GKFCM.

$$\psi(x_i, v_j) = \exp\left(\frac{-\|x_i - v_j\|^2}{\sigma^2}\right)$$

$$\psi(x_i, v_j) = \exp\left(\frac{-\left(\sqrt{(x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2}\right)^2}{\sigma^2}\right)$$

$$\psi(x_i, v_j) = \exp\left(\frac{-\left((x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2\right)}{\sigma^2}\right)$$

Nilai  $\sigma^2$  merupakan nilai varians dari data, berdasarkan data penelitian pada Lampiran 1, maka diperoleh nilai  $\sigma^2 = 1330877$ . Dengan Persamaan (5) diatas diperoleh nilai *gaussian kernel* yang ditampilkan pada Lampiran 4. Hasil Persamaan (5) yang telah diperoleh digunakan untuk menghitung nilai keanggotaan *fuzzy* untuk setiap data dengan Persamaan (7) sebagai berikut:

**Tabel 16.** Matriks Keanggotaan *Fuzzy* GKFCM Empat *Cluster* Pada Iterasi Ke-1

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>			
	1	2	3	4
Bantaeng	0,19	0,19	0,39	0,24
Barro	0,15	0,15	0,47	0,22
Bone	0,26	0,26	0,23	0,25
Bulukumba	0,27	0,27	0,22	0,24
Enrekang	0,18	0,18	0,40	0,24
Gowa	0,25	0,25	0,24	0,25
Jeneponto	0,30	0,30	0,17	0,23
Kepulauan Selayar	0,14	0,13	0,51	0,22
Kota Makassar	0,25	0,25	0,25	0,25
Kota Palopo	0,27	0,27	0,22	0,25
Kota Parepare	0,27	0,24	0,17	0,32
Luwu	0,34	0,34	0,12	0,20
Luwu Timur	0,05	0,05	0,79	0,11
Luwu Utara	0,32	0,32	0,14	0,22
Maros	0,32	0,33	0,14	0,22
Pangkajene Dan Kepulauan	0,14	0,14	0,51	0,21
Pinrang	0,29	0,29	0,18	0,24
Sidenreng Rappang	0,42	0,47	0,02	0,10
Sinjai	0,01	0,00	0,98	0,01
Soppeng	0,11	0,11	0,61	0,18
Takalar	0,28	0,28	0,19	0,24

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>			
	1	2	3	4
Tana Toraja	0,13	0,13	0,53	0,20
Toraja Utara	0,05	0,05	0,80	0,10
Wajo	0,36	0,40	0,08	0,16

Iterasi dilanjutkan ke langkah selanjutnya setelah memperoleh nilai matriks keanggotaan *fuzzy*  $u_{ij}$  menggunakan Persamaan (7). Nilai  $u_{ij}$  yang diperoleh pada iterasi ke-1 akan digunakan untuk menghitung nilai  $v_j$  pada iterasi ke-2. Menggunakan Persamaan (8) sehingga diperoleh nilai titik pusat *cluster*  $v_j$  GKFCM iterasi ke-2 yang ditampilkan pada Tabel 17 yang memperlihatkan nilai titik pusat dari setiap *cluster* dengan  $v_1$  sebagai titik pusat *cluster* 1 dan seterusnya, sebagai berikut:

**Tabel 17.** Titik Pusat *Cluster* Pada Iterasi Ke-2 GKFCM Empat *Cluster*

Cluster	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	2171,14	283,15	71,33	4,06
2	1319,94	391,77	65,60	6,77
3	291,40	130,61	77,38	34,09
4	765,20	221,58	68,11	11,85

Setelah diperoleh nilai titik pusat *cluster* pada Tabel 17, langkah-langkah perhitungan berikutnya sama dengan langkah perhitungan pada iterasi sebelumnya yaitu memperbarui nilai *kernel gaussian* antar data dengan titik pusat *cluster*, memperbarui nilai keanggotaan *fuzzy* dan memperbarui titik pusat *cluster* untuk iterasi selanjutnya. Iterasi dilanjutkan hingga diperoleh hasil maksimal pada iterasi langkah ke-93 dengan nilai  $\|U_{93} - U_{92}\| = 8,53 \times 10^{-10}$  karena nilai  $\|U_{93} - U_{92}\| \leq 10^{-9}$  maka dapat disimpulkan bahwa algoritma GKFCM telah mencapai konvergensi sehingga nilai  $v_j$  dan  $u_{ij}$  yang diperoleh merupakan nilai yang paling optimal. Tabel 18 menampilkan nilai hasil akhir titik pusat *cluster* dari algoritma GKFCM yang diperoleh pada iterasi ke-93 sebagai berikut:

**Tabel 18.** Titik Pusat *Cluster* Pada Iterasi Ke-93 GKFCM Empat *Cluster*

Cluster	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	2171,14	283,15	71,33	4,06
2	1319,94	391,77	65,60	6,77
3	291,40	130,61	77,38	34,09

<i>Cluster</i>	<b>Titik Pusat <i>Cluster</i></b>			
	$X_1$	$X_2$	$X_3$	$X_4$
4	765,20	221,58	68,11	11,85

Pada iterasi ke-93 diperoleh tabel nilai keanggotaan *fuzzy* paling optimal yang menunjukkan derajat keanggotaan masing-masing data pada tiap *cluster*. Data termasuk ke dalam sebuah *cluster* yang nilai derajat keanggotaan *fuzzy* paling tinggi. Tabel 19 menampilkan nilai derajat keanggotaan *fuzzy* serta nilai *cluster* untuk masing-masing data.

**Tabel 19.** Matriks Keanggotaan *Fuzzy* GKFCM Empat *Cluster* Pada Iterasi Ke-93

<b>Kabupaten/Kota</b>	<b>Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i></b>				<b><i>Cluster</i></b>
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	
Bantaeng	0,02	0,03	0,87	0,08	3
Barru	0,00	0,01	0,97	0,02	3
Bone	0,31	0,50	0,07	0,12	2
Bulukumba	0,01	0,96	0,01	0,02	2
Enrekang	0,02	0,03	0,89	0,06	3
Gowa	0,99	0,00	0,00	0,00	1
Jeneponto	0,01	0,05	0,03	0,91	4
Kepulauan Selayar	0,00	0,01	0,96	0,02	3
Kota Makassar	0,25	0,25	0,25	0,25	1
Kota Palopo	0,07	0,70	0,07	0,16	2
Kota Parepare	0,03	0,06	0,55	0,36	3
Luwu	0,01	0,02	0,04	0,93	4
Luwu Timur	0,00	0,01	0,95	0,04	3
Luwu Utara	0,00	0,00	0,00	0,99	4
Maros	0,00	0,00	0,00	1,00	4
Pangkajene Dan Kepulauan	0,00	0,00	0,98	0,01	3
Pinrang	0,03	0,17	0,07	0,74	4
Sidenreng Rappang	0,02	0,05	0,58	0,36	3
Sinjai	0,01	0,01	0,93	0,05	3
Soppeng	0,00	0,00	1,00	0,00	3
Takalar	0,04	0,37	0,08	0,51	4
Tana Toraja	0,00	0,00	0,99	0,00	3
Toraja Utara	0,00	0,00	0,97	0,02	3
Wajo	0,02	0,05	0,22	0,71	4

Tabel 19 merupakan nilai matriks keanggotaan *fuzzy* algoritma GKFCM. Matriks keanggotaan *fuzzy* merupakan representasi dari tingkat sejauh mana setiap titik data tergabung dengan setiap *cluster*. Matriks keanggotaan *fuzzy* memungkinkan GKFCM untuk menangani ketidakpastian atau ketidakjelasan dalam pemilihan *cluster*, dengan menentukan derajat keanggotaan setiap titik data yang beragam ke dalam setiap *cluster* yang ada. Pembagian *cluster* yang terdapat pada Tabel 19 kemudian dihitung rata-rata tiap clusternya yang akan digunakan dalam menghitung nilai DBI GKFCM dua *cluster* dengan Persamaan (9). Sebagai berikut:

$$DBI = \frac{1}{c} \sum_{j=1}^c \max_{j \neq k} R_{j,k}$$

Dengan

$$R_{j,k} = \frac{MAE_j + MAE_k}{\|d_j - d_k\|}$$

$$MAE_j = \left( \frac{1}{T_j} \sum_{x \in C_j} \|x_i - d_j\| \right)$$

Sehingga diperoleh nilai DBI nilai DBI GKFCM dua *cluster*. Sebagai berikut:

**Tabel 20.** Nilai DBI GKFCM Dengan Dua *cluster*

Jumla <i>Cluster</i>	Nilai <i>Davies Bouldin Index</i>
4	0,819

Nilai DBI yang diperoleh pada Tabel 20 akan dibandingkan oleh nilai DBI pada pembagian *cluster* GKFCM yang lain. Nilai DBI yang paling rendah akan menunjukkan bahwa GKFCM dengan nilai *cluster* paling rendah tersebut merupakan jumlah kemungkinan *cluster* yang paling optimal digunakan pada analisis GKFCM untuk Data Kriminal Provinsi Sulawesi Selatan Tahun 2022.

### 3.1.4 *Gaussian Kernel Fuzzy C-Means* dengan Lima *Cluster*

Proses GKFCM dilanjutkan dengan menggunakan pembagian lima *cluster*. Langkah awal untuk memulai iterasi GKFCM sama seperti pada bagian sebelumnya yaitu dengan melakukan penentuan nilai acak untuk nilai  $v_j$  pada iterasi Ke-1 karena Persamaan (8) tidak dapat digunakan sebelum diperoleh nilai keanggotaan  $u_{ij}$  untuk setiap data. Nilai acak  $v_j$  pada iterasi Ke-1 ditampilkan pada Tabel 21 sebagai berikut:

**Tabel 21.** Titik Pusat *Cluster* Pada Iterasi Ke-1 GKFCM Lima *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	512,77	175,98	72,75	24,25
2	441,30	147,82	76,17	30,41
3	409,56	154,72	73,04	29,13
4	486,28	181,17	75,07	25,94
5	559,29	193,20	74,00	22,47

Nilai pusat *cluster*  $v_j$  pada Tabel 21 kemudian digunakan untuk menghitung nilai *kernel gaussian* antara titik pusat *cluster*  $v_j$  dan data. Persamaan (5) digunakan untuk menghitung nilai *kernel gaussian* titik pusat dengan data GKFCM pada iterasi ke-1, berikut adalah proses perhitungan nilai *kernel gaussian* pada metode GKFCM.

$$\psi(x_i, v_j) = \exp\left(\frac{-\|x_i - v_j\|^2}{\sigma^2}\right)$$

$$\psi(x_i, v_j) = \exp\left(\frac{-\left(\sqrt{(x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2}\right)^2}{\sigma^2}\right)$$

$$\psi(x_i, v_j) = \exp\left(\frac{-\left((x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2\right)}{\sigma^2}\right)$$

Nilai  $\sigma^2$  merupakan nilai varians dari data, berdasarkan data penelitian pada Lampiran 1, maka diperoleh nilai  $\sigma^2 = 1330877$ . Dengan Persamaan (5) diatas diperoleh nilai *gaussian kernel* yang ditampilkan pada Lampiran 5. Hasil Persamaan (5) yang telah diperoleh digunakan untuk menghitung nilai keanggotaan *fuzzy* untuk setiap data dengan Persamaan (7) sebagai berikut:

**Tabel 22.** Matriks Keanggotaan *Fuzzy* GKFCM Lima *Cluster* Pada Iterasi Ke-1

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>				
	1	2	3	4	5
Bantaeng	0,16	0,24	0,29	0,18	0,13
Barru	0,14	0,25	0,34	0,17	0,10
Bone	0,21	0,19	0,19	0,20	0,22
Bulukumba	0,21	0,18	0,18	0,20	0,23
Enrekang	0,16	0,24	0,29	0,18	0,13
Gowa	0,20	0,20	0,20	0,20	0,21
Jeneponto	0,22	0,15	0,14	0,19	0,29

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>				
	1	2	3	4	5
Kepulauan Selayar	0,13	0,24	0,36	0,17	0,09
Kota Makassar	0,20	0,20	0,20	0,20	0,20
Kota Palopo	0,21	0,18	0,18	0,20	0,23
Kota Parepare	0,22	0,17	0,16	0,26	0,20
Luwu	0,22	0,11	0,09	0,17	0,40
Luwu Timur	0,07	0,26	0,54	0,09	0,04
Luwu Utara	0,22	0,13	0,11	0,18	0,35
Maros	0,22	0,13	0,11	0,18	0,35
Pangkajene Dan Kepulauan	0,13	0,26	0,36	0,16	0,09
Pinrang	0,22	0,16	0,15	0,20	0,27
Sidenreng Rappang	0,55	0,05	0,02	0,34	0,04
Sinjai	0,04	0,17	0,72	0,06	0,02
Soppeng	0,11	0,26	0,41	0,14	0,07
Takalar	0,22	0,17	0,16	0,20	0,26
Tana Toraja	0,13	0,26	0,37	0,15	0,09
Toraja Utara	0,08	0,24	0,53	0,11	0,05
Wajo	0,22	0,08	0,06	0,14	0,50

Iterasi dilanjutkan ke langkah selanjutnya setelah memperoleh nilai matriks keanggotaan *fuzzy*  $u_{ij}$  menggunakan Persamaan (7). Nilai  $u_{ij}$  yang diperoleh pada iterasi ke-1 akan digunakan untuk menghitung nilai  $v_j$  pada iterasi ke-2. Menggunakan Persamaan (8) sehingga diperoleh nilai titik pusat *cluster*  $v_j$  GKFCM iterasi ke-2 yang ditampilkan pada Tabel 23 yang memperlihatkan nilai titik pusat dari setiap *cluster* dengan  $v_1$  sebagai titik pusat *cluster* 1 dan seterusnya, sebagai berikut:

**Tabel 23.** Titik Pusat *Cluster* Pada Iterasi Ke-2 GKFCM Lima *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	636,92	206,35	71,73	18,34
2	439,22	164,25	76,22	29,78
3	364,34	146,28	75,24	28,62
4	608,99	211,15	73,82	21,64
5	749,74	219,99	67,42	14,15

Setelah diperoleh nilai titik pusat *cluster* pada Tabel 23, langkah-langkah perhitungan berikutnya sama dengan langkah perhitungan pada iterasi sebelumnya yaitu

memperbarui nilai *kernel gaussian* antar data dengan titik pusat *cluster*, memperbarui nilai keanggotaan *fuzzy* dan memperbarui titik pusat *cluster* untuk iterasi selanjutnya. Iterasi dilanjutkan hingga tercapai batas iterasi maksimal pada iterasi langkah ke-100 dengan nilai  $\|U_{100} - U_{99}\| = 1,22 \times 10^{-7}$  sehingga nilai  $v_j$  dan  $u_{ij}$  yang diperoleh merupakan nilai yang paling optimal. Tabel 24 menampilkan nilai hasil akhir titik pusat *cluster* dari algoritma GKFCM yang diperoleh pada iterasi ke-100 sebagai berikut:

**Tabel 24.** Titik Pusat *Cluster* Pada Iterasi Ke-100 GKFCM Lima *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	1408,2	332,16	66,91	6,30
2	661,37	197,64	64,97	13,70
3	279,22	125,85	77,35	35,01
4	952,91	287,40	70,28	9,32
5	2198,23	286,36	71,26	3,99

Pada iterasi ke-100 diperoleh tabel nilai keanggotaan *fuzzy* paling optimal yang menunjukkan derajat keanggotaan masing-masing data pada tiap *cluster*. Data termasuk ke dalam sebuah *cluster* yang nilai derajat keanggotaan *fuzzy* paling tinggi. Tabel 25 menampilkan nilai derajat keanggotaan *fuzzy* serta nilai *cluster* untuk masing-masing data.

**Tabel 25.** Matriks Keanggotaan *Fuzzy* GKFCM Lima *Cluster* Pada Iterasi Ke-100

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>					<i>Cluster</i>
	1	2	3	4	5	
Bantaeng	0,02	0,08	0,83	0,04	0,02	3
Barru	0,00	0,02	0,97	0,01	0,00	3
Bone	0,60	0,07	0,05	0,11	0,18	1
Bulukumba	0,96	0,01	0,00	0,02	0,01	1
Enrekang	0,02	0,07	0,86	0,03	0,01	3
Gowa	0,00	0,00	0,00	0,00	1,00	5
Jeneponto	0,03	0,17	0,02	0,77	0,01	4
Kepulauan Selayar	0,01	0,04	0,93	0,01	0,00	3
Kota Makassar	0,20	0,20	0,20	0,20	0,20	1
Kota Palopo	0,42	0,13	0,07	0,31	0,07	1
Kota Parepare	0,04	0,47	0,36	0,11	0,02	2
Luwu	0,00	0,98	0,01	0,01	0,00	2
Luwu Timur	0,01	0,08	0,89	0,02	0,01	3



Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>					Cluster
	1	2	3	4	5	
Luwu Utara	0,02	0,82	0,03	0,13	0,01	2
Maros	0,02	0,81	0,03	0,14	0,01	2
Pangkajene Dan Kepulauan	0,00	0,01	0,98	0,01	0,00	3
Pinrang	0,01	0,02	0,00	0,97	0,00	4
Sidenreng Rappang	0,02	0,57	0,31	0,08	0,01	2
Sinjai	0,01	0,12	0,83	0,03	0,01	3
Soppeng	0,00	0,01	0,99	0,00	0,00	3
Takalar	0,03	0,04	0,01	0,91	0,01	4
Tana Toraja	0,00	0,00	1,00	0,00	0,00	3
Toraja Utara	0,01	0,05	0,93	0,01	0,00	3
Wajo	0,01	0,90	0,05	0,04	0,00	2

Tabel 25 merupakan nilai matriks keanggotaan *fuzzy* algoritma GKFCM. Matriks keanggotaan *fuzzy* merupakan representasi dari tingkat sejauh mana setiap titik data tergabung dengan setiap *cluster*. Matriks keanggotaan *fuzzy* memungkinkan GKFCM untuk menangani ketidakpastian atau ketidakjelasan dalam pemilihan *cluster*, dengan menentukan derajat keanggotaan setiap titik data yang beragam ke dalam setiap *cluster* yang ada. Pembagian *cluster* yang terdapat pada Tabel 25 kemudian dihitung rata-rata tiap clusternya yang akan digunakan dalam menghitung nilai DBI GKFCM dua *cluster* dengan Persamaan (9). Sebagai berikut:

$$DBI = \frac{1}{c} \sum_{j=1}^c \max_{j \neq k} R_{j,k}$$

Dengan

$$R_{j,k} = \frac{MAE_j + MAE_k}{\|d_j - d_k\|}$$

$$MAE_j = \left( \frac{1}{T_j} \sum_{x \in C_j} \|x_i - d_j\| \right)$$

Sehingga diperoleh nilai DBI nilai DBI GKFCM dua *cluster*. Sebagai berikut:

**Tabel 26.** Nilai DBI GKFCM Dengan Dua *cluster*

Jumla <i>Cluster</i>	Nilai <i>Davies Bouldin Index</i>
5	1,597

Nilai DBI yang diperoleh pada Tabel 26 akan dibandingkan oleh nilai DBI pada pembagian *cluster* GKFCM yang lain. Nilai DBI yang paling rendah akan menunjukkan bahwa GKFCM dengan nilai *cluster* paling rendah tersebut merupakan jumlah kemungkinan *cluster* yang paling optimal digunakan pada analisis GKFCM untuk Data Kriminal Provinsi Sulawesi Selatan Tahun 2022.

### 3.1.5 Gaussian Kernel Fuzzy C-Means dengan Enam Cluster

Proses GKFCM dilanjutkan dengan menggunakan pembagian enam *cluster*. Langkah awal untuk memulai iterasi GKFCM sama seperti pada bagian sebelumnya yaitu dengan melakukan penentuan nilai acak untuk nilai  $v_j$  pada iterasi Ke-1 karena Persamaan (8) tidak dapat digunakan sebelum diperoleh nilai keanggotaan  $u_{ij}$  untuk setiap data. Nilai acak  $v_j$  pada iterasi Ke-1 ditampilkan pada Tabel 21 sebagai berikut:

**Tabel 27.** Titik Pusat *Cluster* Pada Iterasi Ke-1 GKFCM Enam *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	639,08	211,54	75,05	23,59
2	421,67	170,07	74,90	32,18
3	496,99	177,56	74,35	24,07
4	442,32	160,34	73,63	28,79
5	432,08	148,70	77,16	25,60
6	512,27	187,75	71,46	22,84

Nilai pusat *cluster*  $v_j$  pada Tabel 27 kemudian digunakan untuk menghitung nilai *kernel gaussian* antara titik pusat *cluster*  $v_j$  dan data. Persamaan (5) digunakan untuk menghitung nilai *kernel gaussian* titik pusat dengan data GKFCM pada iterasi ke-1, berikut adalah proses perhitungan nilai *kernel gaussian* pada metode GKFCM.

$$\psi(x_i, v_j) = \exp\left(\frac{-\|x_i - v_j\|^2}{\sigma^2}\right)$$

$$\psi(x_i, v_j) = \exp\left(\frac{-\left(\sqrt{(x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2}\right)^2}{\sigma^2}\right)$$

$$\psi(x_i, v_j) = \exp\left(\frac{-\left((x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2\right)}{\sigma^2}\right)$$

Nilai  $\sigma^2$  merupakan nilai varians dari data, berdasarkan data penelitian pada Lampiran 1, maka diperoleh nilai  $\sigma^2 = 1330877$ . Dengan Persamaan (5) diatas diperoleh nilai *gaussian kernel* yang ditampilkan pada Lampiran 6. Hasil Persamaan

(5) yang telah diperoleh digunakan untuk menghitung nilai keanggotaan *fuzzy* untuk setiap data dengan Persamaan (7) sebagai berikut:

**Tabel 28.** Matriks Keanggotaan *Fuzzy* GKFCM Enam *Cluster* Pada Iterasi Ke-1

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>					
	1	2	3	4	5	6
Bantaeng	0,08	0,22	0,15	0,20	0,21	0,14
Barru	0,06	0,25	0,13	0,21	0,23	0,12
Bone	0,20	0,16	0,17	0,16	0,16	0,17
Bulukumba	0,22	0,15	0,17	0,15	0,15	0,17
Enrekang	0,08	0,22	0,15	0,20	0,22	0,13
Gowa	0,18	0,16	0,17	0,16	0,16	0,17
Jeneponto	0,37	0,10	0,15	0,11	0,11	0,16
Kepulauan Selayar	0,05	0,28	0,13	0,21	0,23	0,11
Kota Makassar	0,17	0,17	0,17	0,17	0,17	0,17
Kota Palopo	0,22	0,15	0,17	0,15	0,15	0,17
Kota Parepare	0,10	0,18	0,20	0,17	0,14	0,22
Luwu	0,71	0,04	0,07	0,04	0,04	0,09
Luwu Timur	0,02	0,29	0,08	0,22	0,32	0,06
Luwu Utara	0,58	0,06	0,10	0,07	0,06	0,12
Maros	0,56	0,07	0,11	0,07	0,07	0,12
Pangkajene Dan Kepulauan	0,05	0,25	0,13	0,21	0,25	0,11
Pinrang	0,32	0,12	0,15	0,12	0,12	0,17
Sidenreng Rappang	0,01	0,02	0,68	0,04	0,03	0,21
Sinjai	0,01	0,39	0,06	0,20	0,29	0,04
Soppeng	0,04	0,28	0,11	0,22	0,26	0,10
Takalar	0,29	0,13	0,16	0,13	0,13	0,17
Tana Toraja	0,05	0,26	0,12	0,21	0,25	0,11
Toraja Utara	0,02	0,32	0,09	0,22	0,28	0,07
Wajo	0,41	0,06	0,17	0,08	0,07	0,21

Iterasi dilanjutkan ke langkah selanjutnya setelah memperoleh nilai matriks keanggotaan *fuzzy*  $u_{ij}$  menggunakan Persamaan (7). Nilai  $u_{ij}$  yang diperoleh pada iterasi ke-1 akan digunakan untuk menghitung nilai  $v_j$  pada iterasi ke-2. Menggunakan Persamaan (8) sehingga diperoleh nilai titik pusat *cluster*  $v_j$  GKFCM iterasi ke-2 yang ditampilkan pada Tabel 29 yang memperlihatkan nilai titik pusat dari setiap *cluster* dengan  $v_1$  sebagai titik pusat *cluster* 1 dan seterusnya, sebagai berikut:

**Tabel 29.** Titik Pusat *Cluster* Pada Iterasi Ke-2 GKFCM Enam *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	771,25	220,54	64,47	12,13
2	374,96	153,52	75,57	29,83
3	552,51	187,38	72,94	19,87
4	418,74	163,20	76,40	30,07
5	385,42	150,66	76,46	30,15
6	655,01	222,76	73,75	20,23

Setelah diperoleh nilai titik pusat *cluster* pada Tabel 29, langkah-langkah perhitungan berikutnya sama dengan langkah perhitungan pada iterasi sebelumnya yaitu memperbarui nilai *kernel gaussian* antar data dengan titik pusat *cluster*, memperbarui nilai keanggotaan *fuzzy* dan memperbarui titik pusat *cluster* untuk iterasi selanjutnya. Iterasi dilanjutkan hingga tercapai batas iterasi maksimal pada iterasi langkah ke-100 dengan nilai  $\|U_{100} - U_{99}\| = 6.86 \times 10^{-6}$  sehingga nilai  $v_j$  dan  $u_{ij}$  yang diperoleh merupakan nilai yang paling optimal. Tabel 24 menampilkan nilai hasil akhir titik pusat *cluster* dari algoritma GKFCM yang diperoleh pada iterasi ke-100 sebagai berikut:

**Tabel 30.** Titik Pusat *Cluster* Pada Iterasi Ke-100 GKFCM Enam *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	2200,7	286,47	71,26	3,99
2	215,00	101,63	85,61	45,16
3	967,41	295,10	69,07	9,12
4	707,32	204,18	64,39	12,55
5	367,61	154,65	72,90	24,59
6	1409,91	326,58	66,92	6,27

Pada iterasi ke-100 diperoleh tabel nilai keanggotaan *fuzzy* paling optimal yang menunjukkan derajat keanggotaan masing-masing data pada tiap *cluster*. Data termasuk ke dalam sebuah *cluster* yang nilai derajat keanggotaan *fuzzy* paling tinggi. Tabel 30 menampilkan nilai derajat keanggotaan *fuzzy* serta nilai *cluster* untuk masing-masing data.

**Tabel 31.** Matriks Keanggotaan *Fuzzy* GKFCM Enam *Cluster* Pada Iterasi Ke-100

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>						<i>Cluster</i>
	1	2	3	4	5	6	
Bantaeng	0,01	0,83	0,01	0,02	0,12	0,01	2
Barru	0,00	0,94	0,00	0,01	0,05	0,00	2
Bone	0,17	0,04	0,11	0,07	0,05	0,57	6
Bulukumba	0,01	0,00	0,02	0,01	0,00	0,95	6
Enrekang	0,00	0,86	0,01	0,02	0,10	0,01	2
Gowa	1,00	0,00	0,00	0,00	0,00	0,00	1
Jeneponto	0,01	0,02	0,61	0,30	0,03	0,03	3
Kepulauan Selayar	0,00	0,41	0,01	0,03	0,53	0,01	5
Kota Makassar	0,17	0,17	0,17	0,17	0,17	0,17	1
Kota Palopo	0,07	0,06	0,30	0,13	0,07	0,37	6
Kota Parepare	0,01	0,16	0,07	0,24	0,48	0,03	5
Luwu	0,00	0,00	0,01	0,98	0,01	0,00	4
Luwu Timur	0,00	0,05	0,00	0,01	0,94	0,00	5
Luwu Utara	0,00	0,01	0,03	0,94	0,01	0,00	4
Maros	0,00	0,01	0,03	0,94	0,01	0,00	4
Pangkajene Dan Kepulauan	0,00	0,75	0,01	0,02	0,22	0,00	2
Pinrang	0,00	0,01	0,93	0,04	0,01	0,01	3
Sidenreng Rappang	0,01	0,12	0,05	0,23	0,58	0,02	5
Sinjai	0,00	0,01	0,00	0,00	0,99	0,00	5
Soppeng	0,00	0,35	0,01	0,02	0,62	0,00	5
Takalar	0,00	0,01	0,93	0,03	0,01	0,02	3
Tana Toraja	0,00	0,75	0,01	0,01	0,23	0,00	2
Toraja Utara	0,00	0,03	0,00	0,00	0,97	0,00	5
Wajo	0,01	0,06	0,06	0,69	0,16	0,02	4

Tabel 31 merupakan nilai matriks keanggotaan *fuzzy* algoritma GKFCM. Matriks keanggotaan *fuzzy* merupakan representasi dari tingkat sejauh mana setiap titik data bergabung dengan setiap *cluster*. Matriks keanggotaan *fuzzy* memungkinkan GKFCM untuk menangani ketidakpastian atau ketidakjelasan dalam pemilihan *cluster*, dengan menentukan derajat keanggotaan setiap titik data yang beragam ke dalam setiap *cluster* yang ada. Pembagian *cluster* yang terdapat pada Tabel 31 kemudian dihitung rata-rata tiap clusternya yang akan digunakan dalam menghitung nilai DBI GKFCM dua *cluster* dengan Persamaan (9). Sebagai berikut:

$$DBI = \frac{1}{c} \sum_{j=1}^c \max_{j \neq k} R_{j,k}$$

Dengan

$$R_{j,k} = \frac{MAE_j + MAE_k}{\|d_j - d_k\|}$$

$$MAE_j = \left( \frac{1}{T_j} \sum_{x \in C_j} \|x_i - d_j\| \right)$$

Sehingga diperoleh nilai DBI nilai DBI GKFCM dua *cluster*. Sebagai berikut:

**Tabel 32.** Nilai DBI GKFCM Dengan Dua *cluster*

Jumla <i>Cluster</i>	Nilai <i>Davies Bouldin Index</i>
6	0,827

Nilai DBI yang diperoleh pada Tabel 32 akan dibandingkan oleh nilai DBI pada pembagian *cluster* GKFCM yang lain. Nilai DBI yang paling rendah akan menunjukkan bahwa GKFCM dengan nilai *cluster* paling rendah tersebut merupakan jumlah kemungkinan *cluster* yang paling optimal digunakan pada analisis GKFCM untuk Data Kriminal Provinsi Sulawesi Selatan Tahun 2022.

### 3.1.6 Gaussian Kernel Fuzzy C-Means dengan Tujuh *Cluster*

Proses GKFCM dilanjutkan dengan menggunakan pembagian tujuh *cluster*. Langkah awal untuk memulai iterasi GKFCM sama seperti pada bagian sebelumnya yaitu dengan melakukan penentuan nilai acak untuk nilai  $v_j$  pada iterasi Ke-1 karena Persamaan (8) tidak dapat digunakan sebelum diperoleh nilai keanggotaan  $u_{ij}$  untuk setiap data. Nilai acak  $v_j$  pada iterasi Ke-1 ditampilkan pada Tabel 33 sebagai berikut:

**Tabel 33.** Titik Pusat *Cluster* Pada Iterasi Ke-1 GKFCM Tujuh *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	469,3289	177,70	75,41	25,99
2	423,87	150,64	73,58	30,89
3	378,50	135,19	80,93	32,53
4	492,79	170,21	70,60	24,70
5	551,87	196,13	67,40	21,36
6	390,12	152,01	85,24	31,94
7	549,62	203,75	80,01	25,66

Nilai pusat *cluster*  $v_j$  pada Tabel 33 kemudian digunakan untuk menghitung nilai *kernel gaussian* antara titik pusat *cluster*  $v_j$  dan data. Persamaan (5) digunakan untuk menghitung nilai *kernel gaussian* titik pusat dengan data GKFCM pada iterasi ke-1, berikut adalah proses perhitungan nilai *kernel gaussian* pada metode GKFCM.

$$\psi(x_i, v_j) = \exp\left(\frac{-\|x_i - v_j\|^2}{\sigma^2}\right)$$

$$\psi(x_i, v_j) = \exp\left(\frac{-\left(\sqrt{(x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2}\right)^2}{\sigma^2}\right)$$

$$\psi(x_i, v_j) = \exp\left(\frac{-\left((x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2\right)}{\sigma^2}\right)$$

Nilai  $\sigma^2$  merupakan nilai varians dari data, berdasarkan data penelitian pada Lampiran 1, maka diperoleh nilai  $\sigma^2 = 1330877$ . Dengan Persamaan (5) diatas diperoleh nilai *gaussian kernel* yang ditampilkan pada Lampiran 7. Hasil Persamaan (5) yang telah diperoleh digunakan untuk menghitung nilai keanggotaan *fuzzy* untuk setiap data dengan Persamaan (7) sebagai berikut:

**Tabel 34.** Matriks Keanggotaan *Fuzzy* GKFCM Tujuh *Cluster* Pada Iterasi Ke-1

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>						
	1	2	3	4	5	6	7
Bantaeng	0,12	0,17	0,23	0,11	0,08	0,20	0,08
Barro	0,11	0,17	0,28	0,09	0,06	0,23	0,06
Bone	0,14	0,14	0,13	0,15	0,16	0,13	0,15
Bulukumba	0,14	0,13	0,12	0,15	0,17	0,13	0,16
Enrekang	0,12	0,17	0,24	0,11	0,08	0,21	0,08
Gowa	0,14	0,14	0,14	0,14	0,15	0,14	0,15
Jeneponto	0,14	0,11	0,09	0,15	0,21	0,10	0,21
Kepulauan Selayar	0,11	0,17	0,27	0,09	0,06	0,25	0,06
Kota Makassar	0,14	0,14	0,14	0,14	0,14	0,14	0,14
Kota Palopo	0,14	0,13	0,12	0,15	0,17	0,13	0,17
Kota Parepare	0,18	0,12	0,09	0,16	0,16	0,11	0,18
Luwu	0,11	0,08	0,06	0,14	0,29	0,06	0,27
Luwu Timur	0,02	0,07	0,69	0,02	0,01	0,18	0,01
Luwu Utara	0,12	0,09	0,07	0,14	0,25	0,07	0,25
Maros	0,12	0,09	0,07	0,15	0,25	0,08	0,24
Pangkajene Dan Kepulauan	0,09	0,16	0,32	0,08	0,05	0,24	0,05

Kabupaten/Kota	Nilai Keanggotaan Fuzzy Pada Cluster						
	1	2	3	4	5	6	7
Pinrang	0,14	0,12	0,10	0,15	0,20	0,10	0,20
Sidenreng Rappang	0,09	0,02	0,01	0,82	0,03	0,01	0,03
Sinjai	0,02	0,08	0,31	0,01	0,01	0,57	0,01
Soppeng	0,08	0,15	0,37	0,06	0,04	0,26	0,04
Takalar	0,14	0,12	0,11	0,15	0,19	0,11	0,19
Tana Toraja	0,09	0,16	0,33	0,08	0,05	0,25	0,05
Toraja Utara	0,04	0,11	0,50	0,03	0,01	0,29	0,02
Wajo	0,09	0,06	0,04	0,14	0,35	0,04	0,28

Iterasi dilanjutkan ke langkah selanjutnya setelah memperoleh nilai matriks keanggotaan fuzzy  $u_{ij}$  menggunakan Persamaan (7). Nilai  $u_{ij}$  yang diperoleh pada iterasi ke-1 akan digunakan untuk menghitung nilai  $v_j$  pada iterasi ke-2. Menggunakan Persamaan (8) sehingga diperoleh nilai titik pusat cluster  $v_j$  GKFCM iterasi ke-2 yang ditampilkan pada Tabel 35 yang memperlihatkan nilai titik pusat dari setiap cluster dengan  $v_1$  sebagai titik pusat cluster 1 dan seterusnya, sebagai berikut:

**Tabel 35.** Titik Pusat Cluster Pada Iterasi Ke-2 GKFCM Tujuh Cluster

Cluster	Titik Pusat Cluster			
	$X_1$	$X_2$	$X_3$	$X_4$
1	656,00	229,32	74,45	21,72
2	469,67	175,82	76,37	30,34
3	345,27	133,62	75,87	28,96
4	560,85	184,24	71,69	17,84
5	761,08	224,44	67,36	13,58
6	366,28	150,04	75,47	29,10
7	773,37	233,98	68,20	13,63

Setelah diperoleh nilai titik pusat cluster pada Tabel 35, langkah-langkah perhitungan berikutnya sama dengan langkah perhitungan pada iterasi sebelumnya yaitu memperbarui nilai kernel gaussian antar data dengan titik pusat cluster, memperbarui nilai keanggotaan fuzzy dan memperbarui titik pusat cluster untuk iterasi selanjutnya. Iterasi dilanjutkan hingga diperoleh hasil maksimal pada iterasi langkah ke-70 dengan nilai  $\|U_{70} - U_{69}\| = 8,11 \times 10^{-10}$  karena nilai  $\|U_{70} - U_{69}\| \leq 10^{-9}$  maka dapat disimpulkan bahwa algoritma GKFCM telah mencapai konvergensi sehingga nilai  $v_j$  dan  $u_{ij}$  yang diperoleh merupakan nilai yang paling optimal. Tabel 35 menampilkan nilai hasil akhir titik pusat cluster dari algoritma GKFCM yang diperoleh pada iterasi ke-70 sebagai berikut:



**Tabel 36.** Titik Pusat *Cluster* Pada Iterasi Ke-70 GKFCM Tujuh *Cluster*

<i>Cluster</i>	Titik Pusat <i>Cluster</i>			
	$X_1$	$X_2$	$X_3$	$X_4$
1	971,95	284,04	68,75	9,07
2	510,47	191,50	72,78	17,45
3	156,11	78,23	99,76	59,11
4	732,45	211,45	65,43	12,04
5	1227,22	608,36	62,48	7,16
6	315,31	136,54	71,15	28,43
7	1745,70	234,45	72,29	5,08

Pada iterasi ke-70 diperoleh tabel nilai keanggotaan *fuzzy* paling optimal yang menunjukkan derajat keanggotaan masing-masing data pada tiap *cluster*. Data termasuk ke dalam sebuah *cluster* yang nilai derajat keanggotaan *fuzzy* paling tinggi. Tabel 37 menampilkan nilai derajat keanggotaan *fuzzy* serta nilai *cluster* untuk masing-masing data.

**Tabel 37.** Matriks Keanggotaan *Fuzzy* GKFCM Tujuh *Cluster* Pada Iterasi Ke-70

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>							<i>Cluster</i>
	1	2	3	4	5	6	7	
Bantaeng	0.00	0.01	0.95	0.00	0.00	0.03	0.00	3
Barru	0.01	0.05	0.46	0.02	0.00	0.46	0.00	6
Bone	0.01	0.01	0.01	0.01	0.02	0.01	0.94	7
Bulukumba	0.23	0.06	0.04	0.09	0.33	0.04	0.21	5
Enrekang	0.00	0.00	0.97	0.00	0.00	0.02	0.00	3
Gowa	0.10	0.08	0.07	0.09	0.13	0.08	0.46	7
Jeneponto	0.52	0.05	0.02	0.35	0.03	0.02	0.01	1
Kepulauan Selayar	0.01	0.08	0.14	0.02	0.01	0.73	0.00	6
Kota Makassar	0.14	0.14	0.14	0.14	0.14	0.14	0.14	1
Kota Palopo	0.01	0.00	0.00	0.00	0.98	0.00	0.00	5
Kota Parepare	0.04	0.56	0.07	0.12	0.02	0.18	0.01	2
Luwu	0.03	0.07	0.01	0.86	0.01	0.02	0.00	4
Luwu Timur	0.01	0.10	0.05	0.02	0.00	0.82	0.00	6
Luwu Utara	0.01	0.01	0.00	0.97	0.00	0.00	0.00	4
Maros	0.01	0.01	0.00	0.98	0.00	0.00	0.00	4
Pangkajene Dan Kepulauan	0.01	0.05	0.24	0.02	0.00	0.68	0.00	6

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>							Cluster
	1	2	3	4	5	6	7	
Pinrang	0.93	0.01	0.00	0.04	0.01	0.00	0.00	1
Sidenreng Rappang	0.00	0.95	0.01	0.01	0.00	0.02	0.00	2
Sinjai	0.01	0.18	0.06	0.03	0.00	0.72	0.00	6
Soppeng	0.00	0.01	0.01	0.00	0.00	0.98	0.00	6
Takalar	0.89	0.01	0.01	0.04	0.03	0.01	0.01	1
Tana Toraja	0.00	0.03	0.14	0.01	0.00	0.81	0.00	6
Toraja Utara	0.00	0.03	0.02	0.01	0.00	0.93	0.00	6
Wajo	0.04	0.54	0.03	0.30	0.01	0.07	0.01	2

Tabel 37 merupakan nilai matriks keanggotaan *fuzzy* algoritma GKFCM. Matriks keanggotaan *fuzzy* merupakan representasi dari tingkat sejauh mana setiap titik data tergabung dengan setiap *cluster*. Matriks keanggotaan *fuzzy* memungkinkan GKFCM untuk menangani ketidakpastian atau ketidakjelasan dalam pemilihan *cluster*, dengan menentukan derajat keanggotaan setiap titik data yang beragam ke dalam setiap *cluster* yang ada. Pembagian *cluster* yang terdapat pada Tabel 37 kemudian dihitung rata-rata tiap clusternya yang akan digunakan dalam menghitung nilai DBI GKFCM dua *cluster* dengan Persamaan (9). Sebagai berikut:

$$DBI = \frac{1}{c} \sum_{j=1}^c \max_{j \neq k} R_{j,k}$$

Dengan

$$R_{j,k} = \frac{MAE_j + MAE_k}{\|d_j - d_k\|}$$

$$MAE_j = \left( \frac{1}{T_j} \sum_{x \in C_j} \|x_i - d_j\| \right)$$

Sehingga diperoleh nilai DBI nilai DBI GKFCM dua *cluster*. Sebagai berikut:

**Tabel 38.** Nilai DBI GKFCM Dengan Tujuh *cluster*

Jumla <i>Cluster</i>	Nilai <i>Davies Bouldin Index</i>
7	1,746

Nilai DBI yang diperoleh pada Tabel 38 akan dibandingkan oleh nilai DBI pada pembagian *cluster* GKFCM yang lain. Nilai DBI yang paling rendah akan menunjukkan bahwa GKFCM dengan nilai *cluster* paling rendah tersebut merupakan jumlah kemungkinan *cluster* yang paling optimal digunakan pada analisis GKFCM untuk Data Kriminal Provinsi Sulawesi Selatan Tahun 2022.

Perhitungan nilai DBI untuk setiap jumlah pembagian kelompok *cluster* pada metode GKFCM ditampilkan dalam Tabel 3 sebagai berikut:

**Tabel 39.** Nilai *Davies-Bouldin Index* Untuk Tiap Jumlah *Cluster* GKFCM

Jumlah Cluster	Nilai <i>Davies-Bouldin Index</i>
2	4,74
3	4,90
<b>4</b>	<b>0,82</b>
5	1,60
6	0,83
7	1,75
8	1,73
9	1,87
10	1,24

Berdasarkan Tabel 3 diperoleh bahwa jumlah *cluster* dengan nilai *Davies-Bouldin Index* paling minimum adalah sebanyak 4 *cluster*. Nilai minimum menunjukkan bahwa pembagian 4 *cluster* untuk analisis GKFCM menghasilkan pengelompokan dengan kemiripan yang paling tinggi pada setiap anggota *cluster* yang berdekatan dan memiliki jarak yang paling besar antar masing-masing titik pusat *cluster* yang terbentuk. Berdasarkan hasil yang diperoleh maka jumlah kelompok yang akan dibentuk dalam analisis *cluster* GKFCM untuk data kriminalitas di Sulawesi Selatan adalah sebanyak 4 kelompok kabupaten/kota. Proses dibawah merupakan contoh perhitungan untuk nilai DBI pada pembagian data ke dalam 4 *cluster*. Hasil pengelompokan GKFCM dengan 4 *cluster* menghasilkan nilai rata-rata tiap *cluster* sebagai berikut:

$$\begin{aligned} \mathbf{d}_1 &= [6583 \quad 495,76 \quad 60,56 \quad 2,37] \\ \mathbf{d}_2 &= [1408,67 \quad 394,44 \quad 66,80 \quad 6,32] \\ \mathbf{d}_2 &= [306,83 \quad 141,36 \quad 78,34 \quad 33,26] \\ \mathbf{d}_2 &= [801,43 \quad 228,70 \quad 67,19 \quad 11,24] \end{aligned}$$

Nilai rata-rata tiap *cluster*  $\mathbf{d}_j$  digunakan untuk menghitung nilai Jarak Intra-*cluster*.

$$MAE_j = \left( \frac{1}{T_j} \sum_{x \in C_j} \|x_i - \mathbf{d}_j\| \right)$$

Dengan

$$\|x_i - \mathbf{d}_j\| = \sqrt{(x_{i1} - d_{j1})^2 + (x_{i2} - d_{j2})^2 + \dots + (x_{il} - d_{jl})^2}$$

$$MAE_j = \left( \frac{1}{T_j} \sum_{x \in C_j} \|x_i - \mathbf{d}_j\| \right)$$

$$MAE_j = \left( \frac{1}{T_j} \|\mathbf{x}_1 - \mathbf{d}_j\| + \|\mathbf{x}_2 - \mathbf{d}_j\| + \dots + \|\mathbf{x}_3 - \mathbf{d}_j\| \right)$$

$$MAE_j = \left( \frac{1}{T_j} \left( \sqrt{(x_{i1} - d_{j1})^2 + \dots + (x_{il} - d_{jl})^2} + \sqrt{(x_{i1} - d_{j1})^2 + \dots + (x_{il} - d_{jl})^2} + \dots + \sqrt{(x_{i1} - d_{j1})^2 + \dots + (x_{il} - d_{jl})^2} \right) \right)$$

$$MAE_1 = \left( \frac{1}{2} \left( \sqrt{(2215 - 6583)^2 + \dots + (3,95 - 2,37)^2} + \sqrt{(10951 - 6583)^2 + \dots + (0,79 - 2,37)^2} \right) \right)$$

$$MAE_1 = 4372,98$$

⋮

$$MAE_4 = \left( \frac{1}{7} \left( \sqrt{(869 - 801,43)^2 + \dots + (10,08 - 11,24)^2} + \sqrt{(689 - 801,43)^2 + \dots + (12,71 - 11,24)^2} + \dots + \sqrt{(606 - 801,43)^2 + \dots + (14,45 - 11,24)^2} \right) \right)$$

$$MAE_4 = 128,31$$

Nilai intra *cluster* yang diperoleh kemudian digunakan pada Persamaan (10) untuk menentukan kesamaan antar *cluster* ke-*j* dan ke-*k* ( $R_{j,k}$ ) sebagai berikut:

$$R_{j,k} = \frac{MAE_j + MAE_k}{\|\mathbf{d}_j - \mathbf{d}_k\|}$$

$$R_{j,k} = \frac{MAE_j + MAE_k}{\sqrt{(d_{j1} - d_{k1})^2 + (d_{j2} - d_{k2})^2 + \dots + (d_{jl} - d_{kl})^2}}$$

$$R_{1,2} = \frac{4372,98 + 242,79}{\sqrt{(6583 - 1408,67)^2 + (495,76 - 394,44)^2 + \dots + (2,37 - 6,32)^2}}$$

$$R_{1,2} = 0,89$$

$$R_{1,3} = \frac{4372,98 + 106,74}{\sqrt{(6583 - 306,83)^2 + (495,76 - 141,36)^2 + \dots + (2,37 - 33,26)^2}}$$

$$R_{1,3} = 0,71$$

⋮

$$R_{4,3} = \frac{128,31 + 106,74}{\sqrt{(801,43 - 306,83)^2 + (228,70 - 141,36)^2 + \dots + (11,24 - 33,26)^2}}$$

$$R_{4,3} = 0,47$$

Nilai-nilai Kesamaan antar *cluster* ke-*j* dan ke-*k* kemudian digunakan untuk menentukan nilai DBI pada pembagian 4 *cluster* metode GKFCM.

$$DBI = \frac{1}{c} \sum_{j=1}^c \max_{j \neq k} R_{j,k}$$

$$DBI = \frac{1}{c} \left( \left( \max_{j \neq k} R_{j,k} \right) + \left( \max_{j \neq k} R_{j,k} \right) + \dots + \left( \max_{j \neq k} R_{j,k} \right) \right)$$

$$DBI = \frac{1}{4} \left( \max\{0,89; 0,71; 0,78\} + \max\{0,89; 0,71; 0,78\} + \dots + \max\{0,89; 0,71; 0,78\} \right)$$

$$DBI = \frac{1}{4} (0,89 + 0,89 + 0,71 + 0,78) = 0,82$$

### 3.2 Proses Analisis Gaussian Kernel Fuzzy C-Means Clustering

Analisis *cluster* dengan metode *Gaussian Kernel Fuzzy C-Means* (GKFCM) merupakan variasi dari algoritma *Fuzzy C-Means* (FCM) yang lebih umum, tetapi dengan modifikasi distribusi *Gaussian* pada perhitungan matriks keanggotaan. Proses *clustering* GKFCM dilakukan dengan memperbarui titik pusat *cluster*  $v_j$  dan nilai keanggotaan *fuzzy*  $u_{ij}$  dalam setiap langkah iterasi hingga diperoleh hasil yang maksimal. Iterasi algoritma GKFCM dihentikan ketika nilai  $\|U_q - U_{q-1}\| \leq e$ , pada penelitian ini nilai  $e = 10^{-9}$ .

Persamaan (8) digunakan untuk memperbarui nilai  $v_j$  untuk tiap *cluster* dan Persamaan (7) digunakan untuk memperbarui nilai keanggotaan  $u_{ij}$  dari setiap data. Pada Persamaan (7) dibutuhkan nilai  $v_{ij}$  dan nilai *kernel gaussian* dari data dan pusat *cluster*  $\psi(x_i, v_j)$ . Sehingga sebelum memperbarui nilai matriks keanggotaan  $u_{ij}$  terlebih dahulu dihitung nilai  $v_{ij}$  menggunakan Persamaan (8) dan  $\psi(x_i, v_j)$  menggunakan Persamaan (5). Langkah awal untuk memulai iterasi GKFCM adalah dengan melakukan penentuan nilai acak untuk nilai  $v_j$  pada iterasi Ke-1 karena Persamaan (8) tidak dapat digunakan sebelum diperoleh nilai keanggotaan  $u_{ij}$  untuk setiap data. Nilai acak untuk  $v_j$  pada iterasi Ke-1 ditampilkan dalam Tabel 4 sebagai berikut:

**Tabel 40.** Titik Pusat *Cluster* Pada Iterasi Ke-1

Cluster	Titik Pusat Cluster			
	$X_1$	$X_2$	$X_3$	$X_4$
1	512,02	182,65	67,77	24,44
2	515,98	174,41	77,77	26,24
3	387,17	155,66	76,51	29,53
4	458,85	185,06	73,47	26,28

Nilai pusat *cluster*  $v_j$  pada Tabel 4 kemudian digunakan untuk menghitung nilai *kernel gaussian* antara titik pusat *cluster*  $v_j$  dan data. Persamaan (6) digunakan untuk menghitung nilai *kernel gaussian* titik pusat dengan data GKFCM pada iterasi ke-1, berikut adalah proses perhitungan nilai *kernel gaussian* pada metode GKFCM.

$$\psi(\mathbf{x}_i, \mathbf{v}_j) = \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{v}_j\|^2}{\sigma^2}\right)$$

$$\psi(\mathbf{x}_i, \mathbf{v}_j) = \exp\left(\frac{-\left(\sqrt{(x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2}\right)^2}{\sigma^2}\right)$$

$$\psi(\mathbf{x}_i, \mathbf{v}_j) = \exp\left(\frac{-\left((x_{i1} - v_{j1})^2 + (x_{i2} - v_{j2})^2 + \dots + (x_{il} - v_{jl})^2\right)}{\sigma^2}\right)$$

Nilai  $\sigma^2$  merupakan nilai varians dari data, berdasarkan data penelitian pada Lampiran 1, maka diperoleh nilai  $\sigma^2 = 1330877$ .

$$\psi(\mathbf{x}_1, \mathbf{v}_1) = \exp\left(\frac{-((x_{11} - v_{11})^2 + \dots + (x_{1l} - v_{1l})^2)}{\sigma^2}\right)$$

$$\psi(\mathbf{x}_1, \mathbf{v}_1) = \exp\left(\frac{-((127 - 512,02)^2 + \dots + (68,97 - 24,44)^2)}{1330877}\right)$$

$$\psi(\mathbf{x}_1, \mathbf{v}_1) = 0,88$$

$$\psi(\mathbf{x}_2, \mathbf{v}_1) = \exp\left(\frac{-((x_{21} - v_{11})^2 + \dots + (x_{2l} - v_{1l})^2)}{\sigma^2}\right)$$

$$\psi(\mathbf{x}_2, \mathbf{v}_1) = \exp\left(\frac{-((224 - 512,02)^2 + \dots + (39,1 - 24,44)^2)}{1330877}\right)$$

$$\psi(\mathbf{x}_2, \mathbf{v}_1) = 0,94$$

⋮

$$\psi(\mathbf{x}_4, \mathbf{v}_4) = \exp\left(\frac{-((x_{41} - v_{41})^2 + \dots + (x_{4l} - v_{4l})^2)}{\sigma^2}\right)$$

$$\psi(\mathbf{x}_4, \mathbf{v}_4) = \exp\left(\frac{-((606 - 458,85)^2 + \dots + (14,45 - 26,28)^2)}{1330877}\right)$$

$$\psi(\mathbf{x}_4, \mathbf{v}_4) = 0,98$$

Keseluruhan nilai *kernel gaussian* dari setiap data dengan pusat *cluster* ditampilkan dalam bentuk tabel pada Lampiran 2. Hasil Persamaan (6) yang telah diperoleh digunakan untuk menghitung nilai keanggotaan *fuzzy* untuk setiap data dengan Persamaan (7) sebagai berikut:

$$u_{ij} = \frac{\left( (1 - \psi(x_i, v_j))^{-1} \right)^{\frac{1}{m-1}}}{\sum_{k=1}^c \left( \frac{1}{1 - \psi(x_i, v_k)} \right)^{\frac{1}{m-1}}}$$

Pada Persamaan (7) terdapat parameter *degree of fuzziness*, yang dalam penelitian ini digunakan nilai  $m = 2$ .

$$u_{11} = \frac{\left( (1 - \psi(x_1, v_1))^{-1} \right)^{\frac{1}{m-1}}}{\sum_{k=1}^4 \left( \frac{1}{1 - \psi(x_1, v_k)} \right)^{\frac{1}{m-1}}}$$

$$u_{11} = \frac{\left( (1 - \psi(x_1, v_1))^{-1} \right)^{\frac{1}{2-1}}}{\left( \frac{1}{1 - \psi(x_1, v_1)} \right)^{\frac{1}{2-1}} + \left( \frac{1}{1 - \psi(x_1, v_2)} \right)^{\frac{1}{2-1}} + \dots + \left( \frac{1}{1 - \psi(x_1, v_4)} \right)^{\frac{1}{2-1}}}$$

$$u_{11} = \frac{\left( \frac{1}{1 - 0,88} \right)}{\left( \frac{1}{1 - 0,88} \right) + \left( \frac{1}{1 - 0,88} \right) + \dots + \left( \frac{1}{1 - 0,91} \right)}$$

$$u_{11} = 0,19$$

$$u_{21} = \frac{\left( (1 - \psi(x_2, v_1))^{-1} \right)^{\frac{1}{2-1}}}{\left( \frac{1}{1 - \psi(x_2, v_1)} \right)^{\frac{1}{2-1}} + \left( \frac{1}{1 - \psi(x_2, v_2)} \right)^{\frac{1}{2-1}} + \dots + \left( \frac{1}{1 - \psi(x_2, v_4)} \right)^{\frac{1}{2-1}}}$$

$$u_{21} = \frac{\left( \frac{1}{1 - 0,94} \right)}{\left( \frac{1}{1 - 0,94} \right) + \left( \frac{1}{1 - 0,94} \right) + \dots + \left( \frac{1}{1 - 0,96} \right)}$$

$$u_{21} = 0,15$$

⋮

$$u_{244} = \frac{\left( (1 - \psi(x_{24}, v_4))^{-1} \right)^{\frac{1}{2-1}}}{\left( \frac{1}{1 - \psi(x_{24}, v_1)} \right)^{\frac{1}{2-1}} + \left( \frac{1}{1 - \psi(x_{24}, v_2)} \right)^{\frac{1}{2-1}} + \dots + \left( \frac{1}{1 - \psi(x_1, v_4)} \right)^{\frac{1}{2-1}}}$$

$$u_{244} = \frac{\left( \frac{1}{1 - 0,88} \right)}{\left( \frac{1}{1 - 0,99} \right) + \left( \frac{1}{1 - 0,99} \right) + \dots + \left( \frac{1}{1 - 0,98} \right)}$$

$$u_{244} = 0,16$$

Secara keseluruhan nilai keanggotaan yang diperoleh ditampilkan dalam tabel matriks keanggotaan *fuzzy* pada Tabel 5 sebagai berikut:

**Tabel 41.** Matriks Keanggotaan *Fuzzy* GKFCM Pada Iterasi Ke-1

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada Cluster			
	1	2	3	4
Bantaeng	0,19	0,19	0,39	0,24
Barro	0,15	0,15	0,47	0,22
Bone	0,26	0,26	0,23	0,25
Bulukumba	0,27	0,27	0,22	0,24
Enrekang	0,18	0,18	0,40	0,24
Gowa	0,25	0,25	0,24	0,25
Jeneponto	0,30	0,30	0,17	0,23
Kepulauan Selayar	0,14	0,13	0,51	0,22
Kota Makassar	0,25	0,25	0,25	0,25
Kota Palopo	0,27	0,27	0,22	0,25
Kota Parepare	0,27	0,24	0,17	0,32
Luwu	0,34	0,34	0,12	0,20
Luwu Timur	0,05	0,05	0,79	0,11
Luwu Utara	0,32	0,32	0,14	0,22
Maros	0,32	0,33	0,14	0,22
Pangkajene Dan Kepulauan	0,14	0,14	0,51	0,21
Pinrang	0,29	0,29	0,18	0,24
Sidenreng Rappang	0,42	0,47	0,02	0,10
Sinjai	0,01	0,00	0,98	0,01
Soppeng	0,11	0,11	0,61	0,18
Takalar	0,28	0,28	0,19	0,24
Tana Toraja	0,13	0,13	0,53	0,20
Toraja Utara	0,05	0,05	0,80	0,10
Wajo	0,36	0,40	0,08	0,16

Iterasi dilanjutkan ke langkah selanjutnya setelah memperoleh nilai matriks keanggotaan *fuzzy*  $u_{ij}$  menggunakan Persamaan (7). Nilai  $u_{ij}$  yang diperoleh pada iterasi ke-1 akan digunakan untuk menghitung nilai  $v_j$  pada iterasi ke-2. Menggunakan Persamaan (8) sebagai berikut:

$$v_j = \frac{\sum_{i=1}^n u_{ij}^m \psi(x_i, v_j) x_i}{\sum_{i=1}^n u_{ij}^m \psi(x_i, v_j)}$$

$$v_j = \frac{u_{1j}^m \psi(x_1, v_j) x_1 + u_{2j}^m \psi(x_2, v_j) x_2 + \dots + u_{nj}^m \psi(x_n, v_j) x_n}{u_{1j}^m \psi(x_1, v_j) + u_{2j}^m \psi(x_2, v_j) + \dots + u_{nj}^m \psi(x_n, v_j)}$$



$$v_1 = \frac{(0,19)^2(0,88)x_1 + (0,15)^2(0,94)x_2 + \dots + (0,36)^2(0,99)x_n}{(0,19)^2(0,88) + (0,15)^2(0,94) + \dots + (0,36)^2(0,99)}$$

$$v_1 = \frac{(0,03) \begin{bmatrix} 127 \\ 67,12 \\ 84,22 \\ 68,97 \end{bmatrix} + (0,02) \begin{bmatrix} 224 \\ 127,98 \\ 65,62 \\ 39,1 \end{bmatrix} + \dots + (0,13) \begin{bmatrix} 606 \\ 149,22 \\ 64,68 \\ 14,45 \end{bmatrix}}{1,17}$$

$$v_1 = \begin{bmatrix} 704,03 \\ 220,69 \\ 70,56 \\ 16,82 \end{bmatrix}$$

$$v_2 = \frac{(0,19)^2(0,88)x_1 + (0,15)^2(0,94)x_2 + \dots + (0,40)^2(0,99)x_n}{(0,19)^2(0,88) + (0,15)^2(0,94) + \dots + (0,40)^2(0,99)}$$

$$v_2 = \frac{(0,03) \begin{bmatrix} 127 \\ 67,12 \\ 84,22 \\ 68,97 \end{bmatrix} + (0,02) \begin{bmatrix} 224 \\ 127,98 \\ 65,62 \\ 39,1 \end{bmatrix} + \dots + (0,16) \begin{bmatrix} 606 \\ 149,22 \\ 64,68 \\ 14,45 \end{bmatrix}}{1,24}$$

$$v_2 = \begin{bmatrix} 699,80 \\ 215,90 \\ 70,24 \\ 16,63 \end{bmatrix}$$

⋮

$$v_4 = \frac{(0,24)^2(0,91)x_1 + (0,22)^2(0,96)x_2 + \dots + (0,16)^2(0,98)x_n}{(0,24)^2(0,91) + (0,22)^2(0,96) + \dots + (0,16)^2(0,98)}$$

$$v_4 = \frac{(0,05) \begin{bmatrix} 127 \\ 67,12 \\ 84,22 \\ 68,97 \end{bmatrix} + (0,05) \begin{bmatrix} 224 \\ 127,98 \\ 65,62 \\ 39,1 \end{bmatrix} + \dots + (0,02) \begin{bmatrix} 606 \\ 149,22 \\ 64,68 \\ 14,45 \end{bmatrix}}{0,83}$$

$$v_4 = \begin{bmatrix} 601,78 \\ 214,64 \\ 75,04 \\ 24,08 \end{bmatrix}$$

Berdasarkan proses perhitungan di atas, diperoleh nilai titik pusat *cluster*  $v_j$  GKFCM iterasi ke-2 yang ditampilkan pada Tabel 6 yang memperlihatkan nilai titik pusat dari setiap *cluster* dengan  $v_1$  sebagai titik pusat cluster 1 dan seterusnya hingga  $v_4$ , sebagai berikut:

**Tabel 42.** Titik Pusat *Cluster* Pada Iterasi Ke-2

Cluster	Titik Pusat Cluster			
	$X_1$	$X_2$	$X_3$	$X_4$
1	704,04	220,69	70,56	16,82
2	699,80	215,90	70,24	16,63

Cluster	Titik Pusat Cluster			
	$X_1$	$X_2$	$X_3$	$X_4$
3	353,11	142,50	75,02	28,76
4	601,78	214,64	75,04	24,08

Setelah diperoleh nilai titik pusat *cluster* pada Tabel 6, langkah-langkah perhitungan berikutnya sama dengan langkah perhitungan pada iterasi sebelumnya yaitu memperbarui nilai *kernel gaussian* antar data dengan titik pusat *cluster*, memperbarui nilai keanggotaan *fuzzy* dan memperbarui titik pusat *cluster* untuk iterasi selanjutnya. Iterasi dilanjutkan hingga diperoleh hasil maksimal pada iterasi langkah ke-93 dengan nilai  $\|U_{93} - U_{92}\| = 9,32 \times 10^{-10}$  karena nilai  $\|U_{93} - U_{92}\| \leq 10^{-9}$  maka dapat disimpulkan bahwa algoritma GKFCM telah mencapai konvergensi sehingga nilai  $v_j$  dan  $u_{ij}$  yang diperoleh merupakan nilai yang paling optimal. Tabel 7 menampilkan nilai hasil akhir titik pusat *cluster* dari algoritma GKFCM yang diperoleh pada iterasi ke-93 sebagai berikut:

**Tabel 43.** Titik Pusat Cluster Pada Iterasi Ke-93

Cluster	Titik Pusat Cluster			
	$X_1$	$X_2$	$X_3$	$X_4$
1	2171,11	283,15	71,33	4,06
2	1319,86	391,81	65,60	6,77
3	291,39	130,61	77,38	34,09
4	765,18	221,57	68,11	11,86

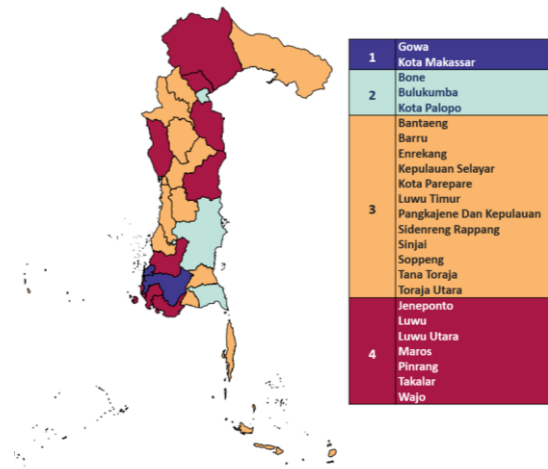
Pada iterasi ke-93 diperoleh tabel nilai keanggotaan *fuzzy* paling optimal yang menunjukkan derajat keanggotaan masing-masing data pada tiap *cluster*. Data termasuk ke dalam sebuah *cluster* yang nilai derajat keanggotaan *fuzzy* paling tinggi. Tabel 8 menampilkan nilai derajat keanggotaan *fuzzy* serta nilai *cluster* untuk masing-masing data.

**Tabel 44.** Matriks Keanggotaan Fuzzy GKFCM Pada Iterasi Ke-93

Kabupaten/Kota	Nilai Keanggotaan Fuzzy Pada Cluster				Cluster
	1	2	3	4	
Bantaeng	0,02	0,03	<b>0,87</b>	0,08	3
Barru	0,00	0,01	<b>0,97</b>	0,02	3
Bone	0,31	<b>0,50</b>	0,07	0,12	2
Bulukumba	0,01	<b>0,96</b>	0,01	0,02	2
Enrekang	0,02	0,03	<b>0,89</b>	0,06	3
Gowa	<b>0,99</b>	0,00	0,00	0,00	1

Kabupaten/Kota	Nilai Keanggotaan <i>Fuzzy</i> Pada <i>Cluster</i>				<i>Cluster</i>
	1	2	3	4	
Jeneponto	0,01	0,05	0,03	<b>0,91</b>	4
Kepulauan Selayar	0,00	0,01	<b>0,96</b>	0,02	3
Kota Makassar	<b>0,25</b>	0,25	0,25	0,25	1
Kota Palopo	0,07	<b>0,70</b>	0,07	0,16	2
Kota Parepare	0,03	0,06	<b>0,55</b>	0,36	3
Luwu	0,01	0,02	0,04	<b>0,93</b>	4
Luwu Timur	0,00	0,01	<b>0,95</b>	0,04	3
Luwu Utara	0,00	0,00	0,00	<b>0,99</b>	4
Maros	0,00	0,00	0,00	<b>1,00</b>	4
Pangkajene Dan Kepulauan	0,00	0,00	<b>0,98</b>	0,01	3
Pinrang	0,03	0,17	0,07	<b>0,74</b>	4
Sidenreng Rappang	0,02	0,05	<b>0,58</b>	0,36	3
Sinjai	0,01	0,01	<b>0,93</b>	0,05	3
Soppeng	0,00	0,00	<b>1,00</b>	0,00	3
Takalar	0,04	0,37	0,08	<b>0,51</b>	4
Tana Toraja	0,00	0,00	<b>0,99</b>	0,00	3
Toraja Utara	0,00	0,00	<b>0,97</b>	0,02	3
Wajo	0,02	0,05	0,22	<b>0,71</b>	4

Tabel 8 merupakan nilai matriks keanggotaan *fuzzy* algoritma GKFCM. Matriks keanggotaan *fuzzy* merupakan representasi dari tingkat sejauh mana setiap titik data tergabung dengan setiap *cluster*. Matriks keanggotaan *fuzzy* memungkinkan GKFCM untuk menangani ketidakpastian atau ketidakjelasan dalam pemilihan *cluster*, dengan menentukan derajat keanggotaan setiap titik data yang beragam ke dalam setiap *cluster* yang ada. Pembagian *cluster* yang terdapat pada Tabel 8 dapat dibentuk ke dalam bentuk peta untuk memperlihatkan sebaran pembagian kelompok kabupaten/kota di Sulawesi Selatan berdasarkan data kriminalitas pada setiap daerah kabupaten/kota sebagai berikut:



**Gambar 1.** Peta *Cluster* GKFCM Kabupaten/Kota Sulawesi Selatan

Berdasarkan Gambar 1, dilihat dari grafik peta hasil pengelompokan kabupaten/kota di Sulawesi Selatan berdasarkan data kriminalitas, terdapat beberapa daerah yang berbatasan dengan daerah lain yang termasuk ke dalam kelompok *cluster* yang sama. *Cluster* pertama beranggotakan Kota Makassar dan Kabupaten Gowa merupakan kelompok daerah dengan *cluster* sama dan saling berbatasan. *Cluster* dua dan empat merupakan kelompok daerah dengan *cluster* yang sama namun tidak semuanya saling berbatasan. Gambar 1 menunjukkan bahwa tidak semua daerah yang saling berbatasan memiliki kelompok *cluster* yang sama atau tidak memiliki karakteristik data kriminalitas yang serupa, berdasarkan metode GKFCM. *Cluster* yang terbentuk tidak membentuk kelompok daerah yang saling berdekatan menunjukkan bahwa karakteristik kriminalitas kabupaten/kota di Sulawesi Selatan tidak semuanya sama meskipun memiliki lokasi geografis yang saling berdekatan.

### 3.3 Interpretasi Hasil *Cluster*

Penentuan Berdasarkan proses *clustering* menggunakan algoritma GKFCM pada data kriminalitas kabupaten/kota Sulawesi Selatan pada tahun 2022, diperoleh jumlah *cluster* optimal sebanyak 4 *cluster*. *Cluster* pertama terdiri dari 2 daerah, *cluster* kedua terdiri dari 3 daerah, *cluster* ketiga dengan anggota terbanyak yang terdiri dari 12 daerah dan *cluster* keempat terdiri dari 7 daerah. Pengelompokan masing-masing kabupaten/kota Sulawesi Selatan berdasarkan data kriminalitas ditampilkan pada Tabel 8 dan Gambar 1. Berdasarkan pengelompokan wilayah kabupaten/kota tersebut diperoleh nilai rata-rata tiap variabel untuk masing-masing kelompok wilayah sebagai berikut:

**Tabel 45.** Nilai Rata-rata untuk Masing-masing Kelompok Wilayah

<i>cluster</i>	Titik Pusat <i>Cluster</i>			
	<i>X1</i>	<i>X2</i>	<i>X3</i>	<i>X4</i>
1	6583	495,76	60,56	2,37
2	1408,67	394,44	66,80	6,32
3	306,83	141,36	78,34	33,26
4	801,43	228,69	67,19	11,24

Berdasarkan nilai pada Tabel 9, diperoleh karakteristik untuk masing-masing kelompok wilayah sebagai berikut:

1. *Cluster* pertama merupakan kelompok daerah dengan rata-rata jumlah kejahatan yang dilaporkan (*X1*) paling tinggi dibandingkan dengan kelompok daerah lain dalam setahun yaitu rata-rata 6583 kejahatan dilaporkan, Risiko Penduduk Terkena Kejahatan per 100.000 Penduduk (*X2*) yang paling tinggi, Tingkat Persentase Penyelesaian Kejahatan (*X3*) terendah, dan merupakan kelompok daerah dengan selang waktu terjadinya kejahatan (*X4*) yang paling singkat dimana kejahatan terjadi setiap 2,37 jam.
2. *Cluster* kedua merupakan kelompok daerah dengan nilai variabel (*X1*) dan (*X2*) paling tinggi kedua setelah *cluster* pertama dan dengan nilai variabel (*X3*) dan (*X4*) paling rendah kedua setelah *cluster* pertama.
3. *Cluster* ketiga merupakan kelompok daerah dengan rata-rata jumlah kejahatan yang dilaporkan (*X1*) paling rendah dibandingkan kelompok daerah lain dalam setahun, Risiko Penduduk Terkena Kejahatan per 100.000 Penduduk (*X2*) yang paling rendah, Tingkat Persentase Penyelesaian Kejahatan (*X3*) tinggi, dan merupakan kelompok daerah dengan selang waktu terjadinya kejahatan (*X4*) yang paling lama.
4. *Cluster* ke-empat merupakan kelompok daerah dengan nilai variabel (*X1*) dan (*X2*) paling tinggi ketiga setelah *cluster* pertama dan dengan nilai variabel (*X3*) dan (*X4*) paling rendah ketiga setelah *cluster* pertama.

Berdasarkan karakteristik-karakteristik diatas dapat disimpulkan bahwa urutan kelompok kabupaten/kota dari daerah dengan tingkat kriminalitas paling rendah ke paling tinggi dan penanganan kasus kriminal paling baik ke paling buruk yaitu *cluster* 3, *cluster* 4, *cluster* 2 lalu *cluster* 1.

### 3.4 Evaluasi Hasil *Cluster Gaussian Kernel Fuzzy C-Means*

Setelah proses pembentukan *cluster* menggunakan metode GKFCM, langkah selanjutnya adalah mengevaluasi hasil *cluster* tersebut. Evaluasi dilakukan dengan menggunakan ASV untuk menilai seberapa baik objek-objek dalam *cluster* tersebut dikelompokkan. ASV memberikan gambaran mengenai konsistensi dan kesesuaian penempatan setiap objek dalam *cluster*, untuk memastikan bahwa *clustering* yang dihasilkan adalah optimal dan sesuai dengan struktur data yang ada. Hasil

pengelompokan yang diperoleh kemudian digunakan untuk menghitung nilai ASV menggunakan Persamaan (12). ASV diperoleh dengan merata-ratakan nilai *silhouette value*  $s(i)$  setiap data.

Misalkan untuk mencari nilai  $s(i)$  untuk daerah Bantaeng pada hasil *cluster* GKFCM untuk keseluruhan data, pertama dilakukan perhitungan jarak Bantaeng ke titik data lain dalam satu *cluster* yang sama:

$$d(x_i, x_j) = \sqrt{(X1_i - X1_j)^2 + (X2_i - X2_j)^2 + \dots + (X11_i - X11_j)^2}$$

Contoh untuk jarak titik data Bantaeng ke Barru

$$d(x_i, x_j) = \sqrt{(127 - 224)^2 + (67,12 - 127,98)^2 + \dots + (68,97 - 39,1)^2}$$

$$d(x_i, x_j) = 119,80$$

Hasil perhitungan jarak *euclidian* antar data kriminalitas tiap kabupaten/kota ditampilkan pada Lampiran 3. Nilai jarak yang diperoleh kemudian digunakan untuk menghitung rarak rata-rata sampel ke sampel lain dalam *cluster* menggunakan Persamaan (12) sebagai berikut:

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

Contoh jarak rata-rata titik data Bantaeng dengan titik data daerah lain dalam satu *cluster*

$$a(i) = \frac{1}{12 - 1} \sum_{r=1}^{16-1} d(x_{ic} - x_{rc})$$

$$a(i) = \frac{1}{12 - 1} (119,80 + 49,43 + \dots + 237,19)$$

$$a(i) = 222,18$$

Selanjutnya adalah mencari nilai rata-rata jarak titik data Bantaeng dengan daerah lain dengan *cluster* yang berbeda:

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

$$d(i, C) = \frac{1}{12} \sum_{r=1}^8 d(x_i - x_{rc})$$

$$d(i, C) = \frac{1}{12} (2100,63 + 10843 + \dots + 489.42)$$

$$d(i, C) = 1820,69$$

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

$$b(i) = \min d(1820,69)$$

$$b(i) = 1820,69$$

Nilai  $a(i)$  dan  $b(i)$  yang diperoleh kemudian digunakan pada Persamaan (11) untuk memperoleh nilai *silhouette value* untuk daerah Bantaeng.

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

$$s(i) = \frac{1820,69 - 222,18}{\max\{1820,69; 222,18\}}$$

$$s(i) = \frac{1820,69 - 222,18}{1820,69} = 0,88$$

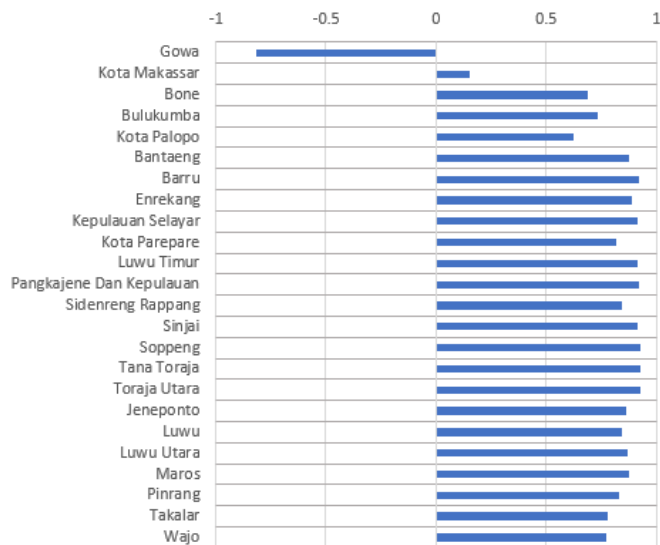
Hasil perhitungan *silhouette value* untuk setiap daerah terdapat pada Lampiran 4. Setelah memperoleh nilai *silhouette value* untuk setiap daerah maka nilai tersebut dirata-ratakan untuk memperoleh nilai ASV.

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

$$ASV = \frac{1}{24} \sum_{i=1}^{24} s(i)$$

$$ASV = \frac{1}{24} (-0,82 + 0,15 + \dots + 0,77) = 0,75$$

Untuk memudahkan interpretasi, maka nilai ASV untuk keseluruhan hasil *cluster* beserta nilai *silhouette value* untuk masing-masing ditampilkan dalam bentuk grafik sebagai berikut:



**Gambar 2.** Nilai *Silhouette Value* Hasil *Clustering* GKFCM pada Data Kriminalitas Kabupaten/Kota di Sulawesi Selatan

Berdasarkan Gambar 2, dapat ditarik kesimpulan bahwa nilai ASV pada pengelompokan kabupaten/kota di Sulawesi Selatan berdasarkan data kriminalitas dengan metode GKFCM menghasilkan nilai ASV sebesar 0,74 menunjukkan bahwa

kualitas *cluster* yang terbentuk termasuk ke dalam struktur *cluster* yang sangat baik menurut standar nilai ASV pada Tabel 1.



## BAB IV PENUTUP

### 4.1 Kesimpulan

Adapun kesimpulan yang diperoleh berdasarkan hasil dan pembahasan adalah sebagai berikut:

1. Proses penentuan *cluster* untuk kabupaten/kota di Sulawesi Selatan menggunakan metode *Gaussian Kernel Fuzzy C-Means* menghasilkan 4 kelompok wilayah sesuai dengan jumlah *cluster* optimal yang diperoleh dari nilai Davies-Bouldin Index. Jumlah masing-masing daerah tiap *cluster* adalah 2 daerah pada *cluster* pertama, 3 daerah pada *cluster* kedua, 12 daerah pada *cluster* ketiga dan 7 daerah pada *cluster* keempat.
2. Kelompok yang terbentuk memiliki karakteristik masing-masing, jumlah Kejahatan yang Dilaporkan (X1) tertinggi terdapat pada *cluster* pertama dan terendah pada *cluster* kedua. Risiko Penduduk Terkena Kejahatan per 100.000 Penduduk (X2) paling tinggi pada *cluster* kedua dan terendah pada *cluster* ketiga. Persentase Penyelesaian Kejahatan (X3) tertinggi pada *cluster* ketiga dan terendah pada *cluster* kedua. Selang Waktu Terjadinya Kejahatan (X4) tertinggi pada *cluster* ketiga dan terendah pada *cluster* kedua. *Cluster* Keempat merupakan *cluster* dengan nilai variabel yang lebih rendah dibandingkan *cluster* 3 namun lebih baik dibandingkan *cluster* 1 dan 2. Nilai evaluasi *Average Silhouette Value* metode *Gaussian Kernel Fuzzy C-Means* sebesar 0,74 menunjukkan bahwa kualitas *cluster* yang terbentuk termasuk ke dalam struktur *cluster* yang sangat baik.

### 4.2 Saran

Saran untuk penelitian lanjutan berdasarkan temuan dan pembahasan yang telah dilakukan adalah sebagai berikut:

1. Penggunaan jenis *kernel* lain yang dapat dijadikan sebagai perbandingan seperti *Polynomial Kernel Fuzzy C-Means* atau *Laplacian Kernel*.
2. Menambahkan pendekatan *Hierarchical Clustering* menggunakan metode *Fuzzy C-Means Hierarchical Clustering* (FCMHC). Metode ini memungkinkan analisis struktur hierarkis dalam data dan dapat memberikan representasi yang lebih luas terhadap hubungan antar *cluster*.

## DAFTAR PUSTAKA

- Alam, M. S., Rahman, M. M., Hossain, M. A., Islam, M. K., Ahmed, K. M., Ahmed, K. T., Singh, B. C., dan Miah, M. S. (2019). Automatic human brain tumor detection in mri image using template-based k means and improved *fuzzy c means clustering* algorithm. *Big Data and Cognitive Computing Journal*, 27(3), 1–18.
- Andini, T. D., dan Farokhah, L. (2022). Peningkatan Ketersediaan Darah Sesuai Segmentasi Umur Menggunakan K-Means *Clustering*. *Jurnal Manajemen Informatika (JAMIKA)*, 12(2), 126–136.
- Askari, S. (2020). *Fuzzy C-Means clustering* algorithm for data with unequal *cluster* sizes and contaminated with *noise* and *outliers*: Review and development. *Journal Expert Systems with Applications*, 148(1), 1–72.
- Bezdek, J. C., Ehrlich, R., dan Full, W. (1988). FCM: the *Fuzzy c-Means clustering* algorithm. *Journal of Computer and Geoscience*, 10(2), 191–203.
- BPS. (2016). *Statistik Kriminal Sulawesi Selatan 2016*. Badan Pusat Statistik Jakarta-Indonesia.
- BPS. (2021). *Statistik Kriminal 2021*. Badan Pusat Statistik Indonesia.
- Chang-Chien, S. J., Nataliani, Y., dan Yang, M. S. (2021). *Gaussian-Kernel C-Means Clustering Algorithms*. *Soft Computing*, 25(3), 1699–1716.
- Chowdhary, C. L., Mittal, M., Kumaresan, P., Pattanaik, P. A., dan Marszalek, Z. (2020). An efficient segmentation and classification system in medical images using intuitionist possibilistic *fuzzy C-mean clustering* and *fuzzy SVM* algorithm. *Sensors Journal MDPI*, 20(14), 1–20.
- Fajar, M., Rahaningsih, N., dan Dinar Dana, R. (2024). Analisis Pola Penjualan Obat Di Apotek an-Naafi Menggunakan Metode K-Means *Clustering*. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 486–492.
- Ferezagia, D. V. (2018). Analisis Tingkat Kemiskinan. *Jurnal Sosial Humaniora Terapan*, 1(1), 1–6.
- Fikri, A., Hutabarat, B. F., dan Khaira, U. (2023). Komparasi Antara Metode K-Means *Clustering* Dan Complete Linkage Dalam Pengelompokan Penyaluran Pinjaman Oleh Financial Technology. *Jurnal Ilmiah Media Sisfo*, 17(2), 228–239.
- Gie, W., dan Jollyta, D. (2020). Perbandingan Euclidean dan Manhattan Untuk Optimasi *Cluster* Menggunakan Davies Bouldin Index: Status Covid-19 Wilayah Riau. *Prosiding Seminar Nasional Riset Dan Information Science (SENARIS)*, 2(April), 187–191.

- Goreti, M., Novia N, Y., dan Wahyuningsih, S. (2016). Perbandingan hasil analisis *cluster* dengan menggunakan metode single linkage dan metode c-means (studi kasus: data tingkat kualitas udara ambien pada perusahaan perkebunan di Kabupaten Kutai Barat tahun 2014). *Jurnal Ekspansional*, 7(1), 9–16.
- Gosain, A., dan Dahiya, S. (2020). A New Robust *Fuzzy Clustering* Approach: DBKIFCM. *Neural Processing Letters*, 52(3), 2189–2210.
- Guntara, M., dan Suprawoto, T. (2022). Drop Out Student *Clusterization* Using the k-Medoids Algorithm. *JTKSI (Jurnal Teknologi Komputer Dan Sistem Informasi)*, 5(1), 61–66.
- Handoyono, N. A. (2022). Apakah Semakin Tinggi Ipm Akan Semakin Bahagia? Analisis Kluster Ditinjau Dari Kualitas Perekonomian. *Jurnal Akuntansi Dewantara*, 6(3), 1–11.
- Hardiyanto, I., Purwananto, Y., Kom, S., Kom, M., dan Soelaiman, R. (2012). Implementasi Segmentasi Citra dengan Menggunakan Metode Generalized *Fuzzy C- Means Clustering* Algorithm with Improved *Fuzzy Partitions*. *Teknik Pomits*, 1(1), 1–5.
- Jiang, W., Fang, X., dan Ding, J. (2020). *Gaussian kernel fuzzy C-means* algorithm for service resource allocation. *Scientific Programming Journal*, 1(1), 1–6.
- Kepala Kepolisian Negara Republik Indonesia. (2009). Peraturan Kepala Kepolisian Negara Republik Indonesia Nomor 7 Tahun 2009 Tentang Sistem Laporan Gangguan Keamanan Dan Ketertiban Masyarakat. In *Berita Negara Republik Indonesia* (Issue 115).
- Kononenko, I., dan Matjazkukar. (2007). *Machine Learning and Data Mining Introduction to Principle and Algorithms*. Horwood Publishing Limited.
- Kowalczyk, A. (2017). *Support Vector Machine Succinity*. Syncfusion, Inc.
- Liu, S., Dong, L., Liao, X., Cao, X., dan Wang, X. (2019). Photovoltaic array fault diagnosis based on *gaussian kernel fuzzy C-means clustering* algorithm. *Sensors (Switzerland)*, 19(7), 1–15.
- Mughnyanti, M., Efendi, S., dan Zarlis, M. (2020). Analysis of determining centroid *clustering* x-means algorithm with davies-bouldin index evaluation. *IOP Conference Series: Materials Science and Engineering*, 725(3), 1–6.
- Nidyashofa, N., dan Istiawan, D. (2017). Penerapan Algoritma *Fuzzy C-Means* untuk Pengelompokan Kabupaten / Kota di Jawa Tengah Berdasarkan Status Kesejahteraan Tahun 2015. *The 6th University Research Colloquium, September*, 23–30.
- Nurkholis, J., Oktafianto, K., dan Thohari, H. (2022). *Clustering* Data Kenaikan Kelas Siswa Madrasah Tsanawiyah (Mts) Menggunakan Metode *Fuzzy C-Means*

- (Studi Kasus Mts Plus Al Amin Banjarejo). *MathVision : Jurnal Matematika*, 4(1), 11–18.
- Rahim, R., Putri, N. D., dan Pertiwi, W. Y. (2023). *Provinsi Sulawesi Selatan Dalam Angka 2023*. Badan Pusat Statistik Sulawesi Selatan.
- Rochim, A. F., Widyaningrum, K., dan Eridani, D. (2021). *Comparison of Kernels Function between of Linear, Radial Base and Polynomial of Support Vector Machine Method Towards COVID-19 Sentiment Analysis*. 224–228.
- Septianingsih, A. (2022). Analisis K-Means *Clustering* Pada Pemetaan Provinsi Indonesia Berdasarkan Indikator Rumah Layak Huni. *Jurnal Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 3(1), 224–241.
- Setiawan, A., Yanto, B., dan Yasdomi, K. (2018). *Logika Fuzzy Dengan MATLAB Contoh Kasus Penelitian Penyakit Bayi Dengan Fuzzy Tsukamoto* (1st ed.). Jayapungas Press.
- Shahapure, K. R., dan Nicholas, C. (2020). Cluster Quality Analysis Using Silhouette Score Ketan. *International Conference on Data Science and Advanced Analytics (DSAA)*, 7(2), 747–748.
- Sitepu, R., dan Gultom, B. (2011). *Clustering Analysis for Air Pollution Level on Industrial Sector in South Sumatera*. *Jurnal Penelitian Sains*, 14(3), 11–17.
- Sovina, M., dan Harahap, F. A. (2022). Penentuan Status Gizi Dengan Indeks Massa Tubuh (IMT) Menggunakan Logika Fuzzy Determination of Nutritional Status With Body Mass Index (BMI) Using Fuzzy Logic. *InfoSys Journal*, 7(1), 105–116.
- Subbalakshmi, C., Rama Krishna, G., Krishna Mohan Rao, S., dan Venketeswa Rao, P. (2015). A method to find optimum number of clusters based on fuzzy silhouette on dynamic data set. *Procedia Computer Science*, 46(1), 346–353. <http://dx.doi.org/10.1016/j.procs.2015.02.030>
- Sulistiani, D. A. (2022). *Pemodelan Kriminalitas Di Sulawesi Selatan Menggunakan Model Geographically Weighted Regression (GWR)*. Universitas Islam Negeri Sunan Ampel Surabaya.
- Syam, S. C. (2021). *Penggunaan Metode Fuzzy C-Means Untuk Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Jumlah Tenaga Kesehatan*. Universitas Hasanuddin.
- Vignesh, R. H. (2019). *Fuzzy C-mean Clustering Using Data Mining*. BookRix.
- Zahida, Q. (2018). *Usulan Perbaikan Kualitas Layanan Rumah Sakit Dengan Pendekatan Lean Service Dan Fuzzy Fmea (Studi Kasus: Poli Anak Rumah Sakit "JIH")*. Universitas Islam Indonesia.

# LAMPIRAN

**Lampiran 1. Data Kriminalitas Sulawesi Selatan Tahun 2022**

<b>Kabupaten/Kota</b>	<b>X1</b>	<b>X2</b>	<b>X3</b>	<b>X4</b>
Bantaeng	127	67,12	84,22	68,97
Barru	224	127,98	65,62	39,1
Bone	1669	220,25	72,79	5,24
Bulukumba	1350	319,93	65,55	6,48
Enrekang	153	72,86	124,18	57,25
Gowa	2215	287,33	71,24	3,95
Jeneponto	869	236,68	81,47	10,08
Kepulauan Selayar	275	202,03	58,54	31,85
Kota Makassar	10951	704,2	49,88	0,79
Kota Palopo	1207	643,14	62,05	7,25
Kota Parepare	464	315,45	84,48	18,87
Luwu	689	186,25	50,79	12,71
Luwu Timur	367	120,37	78,74	23,86
Luwu Utara	741	232,97	65,04	11,82
Maros	750	209,89	74,93	11,68
Pangkajene Dan Kepulauan	275	80,98	94,18	31,85
Pinrang	948	248,74	68,45	9,24
Sidenreng Rappang	500	165,06	71,4	17,52
Sinjai	380	154,78	75,78	23,05
Soppeng	301	131,13	59,8	29,1
Takalar	1007	337,1	64,94	8,69
Tana Toraja	270	112,72	72,96	32,44
Toraja Utara	346	145,83	70,23	25,31
Wajo	606	149,22	64,68	14,45

**Lampiran 2.** Nilai *Kernel Gaussian* GKFCM Dua *Cluster* Iterasi Ke-1

<b>Data</b>	<b>Cluster</b>	
	<b>1</b>	<b>2</b>
1	0,90	0,89
2	0,95	0,94
3	0,34	0,36
4	0,55	0,57
5	0,91	0,90
6	0,10	0,11
7	0,89	0,90
8	0,97	0,96
9	0,00	0,00
10	0,57	0,58
11	0,99	0,98
12	0,97	0,97
13	0,99	0,98
14	0,95	0,95
15	0,94	0,95
16	0,96	0,96
17	0,84	0,86
18	1,00	1,00
19	0,99	0,99
20	0,98	0,97
21	0,79	0,81
22	0,97	0,96
23	0,99	0,98
24	0,99	0,99

**Lampiran 3.** Nilai *Kernel Gaussian* GKFCM Tiga *Cluster* Iterasi Ke-1

Data	Cluster		
	1	2	3
1	0,93	0,90	0,88
2	0,97	0,95	0,93
3	0,31	0,35	0,37
4	0,51	0,56	0,59
5	0,94	0,91	0,89
6	0,09	0,10	0,11
7	0,85	0,89	0,91
8	0,98	0,97	0,95
9	0,00	0,00	0,00
10	0,53	0,57	0,59
11	0,98	0,98	0,98
12	0,94	0,97	0,98
13	1,00	0,99	0,98
14	0,92	0,95	0,96
15	0,92	0,95	0,96
16	0,98	0,96	0,95
17	0,80	0,85	0,87
18	0,99	1,00	1,00
19	1,00	0,99	0,98
20	0,99	0,97	0,96
21	0,75	0,80	0,82
22	0,98	0,96	0,95
23	1,00	0,99	0,98
24	0,97	0,99	0,99



**Lampiran 4.** Nilai *Kernel Gaussian* GKFCM Empat *Cluster* Iterasi Ke-1

Data	Cluster			
	1	2	3	4
1	0,88	0,88	0,94	0,91
2	0,94	0,94	0,98	0,96
3	0,37	0,37	0,29	0,33
4	0,58	0,58	0,49	0,54
5	0,90	0,90	0,95	0,92
6	0,11	0,11	0,08	0,10
7	0,91	0,91	0,84	0,88
8	0,96	0,96	0,99	0,97
9	0,00	0,00	0,00	0,00
10	0,59	0,59	0,50	0,56
11	0,98	0,98	0,98	0,99
12	0,98	0,98	0,93	0,96
13	0,98	0,98	1,00	0,99
14	0,96	0,96	0,91	0,94
15	0,96	0,96	0,90	0,94
16	0,95	0,95	0,99	0,97
17	0,86	0,87	0,78	0,83
18	1,00	1,00	0,99	1,00
19	0,99	0,99	1,00	0,99
20	0,97	0,96	0,99	0,98
21	0,82	0,82	0,73	0,78
22	0,95	0,95	0,99	0,97
23	0,98	0,98	1,00	0,99
24	0,99	0,99	0,96	0,98

**Lampiran 5.** Nilai *Kernel Gaussian* GKFCM Lima *Cluster* Iterasi Ke-1

Data	Cluster				
	1	2	3	4	5
1	0,88	0,92	0,94	0,90	0,86
2	0,94	0,96	0,97	0,95	0,92
3	0,37	0,32	0,30	0,35	0,40
4	0,58	0,53	0,50	0,56	0,62
5	0,90	0,93	0,94	0,91	0,87
6	0,11	0,09	0,09	0,10	0,13
7	0,91	0,87	0,85	0,89	0,93
8	0,96	0,98	0,98	0,97	0,94
9	0,00	0,00	0,00	0,00	0,00
10	0,59	0,54	0,52	0,58	0,63
11	0,98	0,98	0,98	0,99	0,98
12	0,98	0,95	0,94	0,97	0,99
13	0,98	1,00	1,00	0,99	0,97
14	0,96	0,93	0,92	0,95	0,97
15	0,96	0,93	0,91	0,95	0,97
16	0,95	0,98	0,98	0,96	0,93
17	0,86	0,82	0,80	0,85	0,89
18	1,00	1,00	0,99	1,00	1,00
19	0,99	1,00	1,00	0,99	0,98
20	0,97	0,98	0,99	0,97	0,95
21	0,82	0,77	0,75	0,80	0,85
22	0,95	0,98	0,98	0,96	0,93
23	0,98	0,99	1,00	0,98	0,96
24	0,99	0,98	0,97	0,99	1,00

**Lampiran 6.** Nilai *Kernel Gaussian* GKFCM Enam *Cluster* Iterasi Ke-1

Data	Cluster					
	1	2	3	4	5	6
1	0,81	0,93	0,89	0,92	0,93	0,88
2	0,87	0,97	0,94	0,96	0,97	0,94
3	0,45	0,31	0,36	0,32	0,32	0,37
4	0,68	0,51	0,57	0,53	0,52	0,58
5	0,82	0,94	0,90	0,93	0,94	0,90
6	0,15	0,09	0,11	0,09	0,09	0,11
7	0,96	0,86	0,90	0,87	0,86	0,91
8	0,90	0,98	0,96	0,98	0,98	0,96
9	0,00	0,00	0,00	0,00	0,00	0,00
10	0,68	0,53	0,58	0,54	0,53	0,60
11	0,97	0,98	0,98	0,98	0,98	0,99
12	1,00	0,95	0,97	0,95	0,95	0,98
13	0,94	1,00	0,98	0,99	1,00	0,98
14	0,99	0,92	0,95	0,93	0,93	0,96
15	0,99	0,92	0,95	0,93	0,92	0,96
16	0,89	0,98	0,96	0,97	0,98	0,95
17	0,93	0,81	0,85	0,82	0,81	0,86
18	0,98	1,00	1,00	1,00	1,00	1,00
19	0,95	1,00	0,99	1,00	1,00	0,99
20	0,91	0,99	0,97	0,98	0,99	0,96
21	0,89	0,76	0,81	0,77	0,76	0,82
22	0,90	0,98	0,96	0,98	0,98	0,95
23	0,93	1,00	0,98	0,99	0,99	0,98
24	1,00	0,97	0,99	0,98	0,98	0,99

**Lampiran 7.** Nilai *Kernel Gaussian* GKFCM Tujuh *Cluster* Iterasi Ke-1

Data	Cluster						
	1	2	3	4	5	6	7
1	0,91	0,93	0,95	0,90	0,86	0,94	0,86
2	0,95	0,97	0,98	0,95	0,92	0,98	0,92
3	0,34	0,31	0,28	0,35	0,39	0,29	0,39
4	0,55	0,51	0,48	0,57	0,61	0,49	0,61
5	0,92	0,94	0,96	0,91	0,87	0,95	0,88
6	0,10	0,09	0,08	0,11	0,12	0,08	0,12
7	0,88	0,86	0,83	0,90	0,93	0,84	0,93
8	0,97	0,98	0,99	0,96	0,94	0,99	0,94
9	0,00	0,00	0,00	0,00	0,00	0,00	0,00
10	0,56	0,53	0,49	0,58	0,62	0,50	0,62
11	0,99	0,98	0,97	0,98	0,98	0,98	0,99
12	0,96	0,95	0,93	0,97	0,99	0,93	0,98
13	0,99	1,00	1,00	0,99	0,97	1,00	0,97
14	0,94	0,92	0,90	0,95	0,97	0,91	0,97
15	0,94	0,92	0,90	0,95	0,97	0,90	0,97
16	0,96	0,98	0,99	0,96	0,93	0,99	0,93
17	0,84	0,81	0,78	0,85	0,89	0,79	0,89
18	1,00	1,00	0,99	1,00	1,00	0,99	1,00
19	0,99	1,00	1,00	0,99	0,98	1,00	0,98
20	0,98	0,99	1,00	0,97	0,95	0,99	0,95
21	0,79	0,75	0,72	0,80	0,84	0,73	0,84
22	0,97	0,98	0,99	0,96	0,94	0,99	0,94
23	0,99	1,00	1,00	0,98	0,97	1,00	0,97
24	0,99	0,98	0,96	0,99	1,00	0,97	1,00

**Lampiran 8.** Nilai *Kernel Gaussian* Titik Pusat Dengan Data Pada Iterasi Ke-1

Data	Cluster			
	1	2	3	4
1	0,88	0,88	0,94	0,91
2	0,94	0,94	0,98	0,96
3	0,37	0,37	0,29	0,33
4	0,58	0,58	0,49	0,54
5	0,90	0,90	0,95	0,92
6	0,11	0,11	0,08	0,10
7	0,91	0,91	0,84	0,88
8	0,96	0,96	0,99	0,97
9	0,00	0,00	0,00	0,00
10	0,59	0,59	0,50	0,56
11	0,98	0,98	0,98	0,99
12	0,98	0,98	0,93	0,96
13	0,98	0,98	1,00	0,99
14	0,96	0,96	0,91	0,94
15	0,96	0,96	0,90	0,94
16	0,95	0,95	0,99	0,97
17	0,86	0,87	0,78	0,83
18	1,00	1,00	0,99	1,00
19	0,99	0,99	1,00	0,99
20	0,97	0,96	0,99	0,98
21	0,82	0,82	0,73	0,78
22	0,95	0,95	0,99	0,97
23	0,98	0,98	1,00	0,99
24	0,99	0,99	0,96	0,98

**Lampiran 9.** Jarak *Euclidian* Antar Titik Data Kriminalitas Tiap Kabupaten/Kota

<b>Kabupaten/Kota</b>	<b>Gowa</b>	<b>Makassar</b>	<b>Bone</b>	<b>Bulukumba</b>	<b>Kota Palopo</b>	<b>Bantaeng</b>	<b>Barru</b>	<b>Enrekang</b>
Gowa	0,00	8745,97	550,11	865,64	1069,00	2100,63	1997,68	2074,48
Kota Makassar	8745,97	0,00	9294,64	9608,70	9744,20	10843,00	10742,55	10816,84
Bone	550,11	9294,64	0,00	334,29	626,42	1550,94	1448,36	1524,90
Bulukumba	865,64	9608,70	334,29	0,00	353,45	1250,56	1142,71	1224,69
Kota Palopo	1069,00	9744,20	626,42	353,45	0,00	1225,77	1110,27	1201,04
Bantaeng	2100,63	10843,00	1550,94	1250,56	1225,77	0,00	119,80	49,43
Barru	1997,68	10742,55	1448,36	1142,71	1110,27	119,80	0,00	108,80
Enrekang	2074,48	10816,84	1524,90	1224,69	1201,04	49,43	108,80	0,00
Kep. Selayar	1942,12	10687,85	1394,45	1081,77	1031,42	205,29	90,48	191,11
Kota Parepare	1751,34	10494,28	1208,89	886,30	812,45	421,60	305,79	398,27
Luwu Timur	1855,65	10600,15	1305,97	1003,29	989,67	250,00	144,61	226,35
Pangkep	1951,28	10694,31	1401,36	1101,90	1089,17	153,54	75,35	128,43
Sidrap	1719,41	10464,93	1170,37	864,08	853,58	389,27	279,37	365,07
Sinjai	1839,89	10585,32	1290,79	984,15	960,66	271,80	159,42	248,50
Soppeng	1920,56	10665,45	1371,17	1066,11	1040,90	191,20	77,93	173,89
Tana Toraja	1953,03	10697,44	1403,39	1100,03	1077,06	154,89	49,47	136,08
Toraja Utara	1874,47	10619,74	1325,25	1019,17	994,50	237,19	124,15	215,65
Jeneponto	1347,01	10092,89	800,23	488,42	529,00	763,41	654,93	737,25
Luwu	1529,51	10275,07	980,86	674,57	690,82	578,20	469,61	554,55
Luwu Utara	1475,04	10220,89	928,14	615,20	620,83	638,86	528,26	613,95
Maros	1467,07	10213,01	919,08	610,10	629,87	641,78	533,13	616,19
Pinrang	1267,60	10013,38	721,59	408,27	471,89	843,12	734,61	817,54
Takalar	1209,05	9950,79	672,29	343,44	365,61	922,66	811,02	897,22
Wajo	1614,96	10359,90	1065,44	763,38	777,96	489,42	383,38	465,20

**Lampiran 3.** Jarak *Euclidian* Antar Titik Data Kriminalitas Tiap Kabupaten/Kota (Lanjutan)

<b>Kabupaten/Kota</b>	<b>Kep.Selayar</b>	<b>Kota Parepare</b>	<b>Luwu Timur</b>	<b>Pangkep</b>	<b>Sidrap</b>	<b>Sinjai</b>	<b>Soppeng</b>	<b>Tana Toraja</b>
Gowa	1942,12	1751,34	1855,65	1951,28	1719,41	1839,89	1920,56	1953,03
Makassar	10687,85	10494,28	10600,15	10694,31	10464,93	10585,32	10665,45	10697,44
Bone	1394,45	1208,89	1305,97	1401,36	1170,37	1290,79	1371,17	1403,39
Bulukumba	1081,77	886,30	1003,29	1101,90	864,08	984,15	1066,11	1100,03
Kota Palopo	1031,42	812,45	989,67	1089,17	853,58	960,66	1040,90	1077,06
Bantaeng	205,29	421,60	250,00	153,54	389,27	271,80	191,20	154,89
Barru	90,48	305,79	144,61	75,35	279,37	159,42	77,93	49,47
Enrekang	191,11	398,27	226,35	128,43	365,07	248,50	173,89	136,08
Kep. Selayar	0,00	222,32	124,92	126,19	228,83	116,76	75,58	90,61
Kota Parepare	222,32	0,00	218,00	301,60	155,20	181,56	247,50	281,16
Luwu Timur	124,92	218,00	0,00	101,58	140,64	36,91	69,70	97,85
Pangkep	126,19	301,60	101,58	0,00	241,70	129,95	66,19	38,51
Sidrap	228,83	155,20	140,64	241,70	0,00	120,65	202,54	236,36
Sinjai	116,76	181,56	36,91	129,95	120,65	0,00	84,22	118,17
Soppeng	75,58	247,50	69,70	66,19	202,54	84,22	0,00	38,53
Tana Toraja	90,61	281,16	97,85	38,51	236,36	118,17	38,53	0,00
Toraja Utara	91,54	207,22	34,11	99,31	155,40	35,67	48,62	83,25
Jeneponto	595,85	412,69	515,49	614,58	376,09	496,01	578,44	612,16
Luwu	414,81	261,71	330,04	429,80	191,36	311,77	392,34	426,44
Luwu Utara	467,50	289,76	391,01	491,43	250,53	369,70	451,99	486,61
Maros	475,78	305,09	393,53	492,97	254,08	374,26	456,44	490,18
Pinrang	675,07	488,93	595,28	694,44	455,83	575,93	657,96	691,91
Takalar	744,75	543,88	676,01	776,41	535,51	653,22	735,73	770,81
Wajo	335,69	219,56	241,33	339,69	107,43	226,50	305,93	338,56

**Lampiran 3.** Jarak *Euclidian* Antar Titik Data Kriminalitas Tiap Kabupaten/Kota (Lanjutan)

Kabupaten/Kota	Toraja Utara	Jeneponto	Luwu	Luwu Utara	Maros	Pinrang	Takalar	Wajo
Gowa	1874,47	1347,01	1529,51	1475,04	1467,07	1267,60	1209,05	1614,96
Makassar	10619,74	10092,89	10275,07	10220,89	10213,01	10013,38	9950,79	10359,90
Bone	1325,25	800,23	980,86	928,14	919,08	721,59	672,29	1065,44
Bulukumba	1019,17	488,42	674,57	615,20	610,10	408,27	343,44	763,38
Kota Palopo	994,50	529,00	690,82	620,83	629,87	471,89	365,61	777,96
Bantaeng	237,19	763,41	578,20	638,86	641,78	843,12	922,66	489,42
Barru	124,15	654,93	469,61	528,26	533,13	734,61	811,02	383,38
Enrekang	215,65	737,25	554,55	613,95	616,19	817,54	897,22	465,20
Kepulauan Selayar	91,54	595,85	414,81	467,50	475,78	675,07	744,75	335,69
Kota Parepare	207,22	412,69	261,71	289,76	305,09	488,93	543,88	219,56
Luwu Timur	34,11	515,49	330,04	391,01	393,53	595,28	676,01	241,33
Pangkep	99,31	614,58	429,80	491,43	492,97	694,44	776,41	339,69
Sidrap	155,40	376,09	191,36	250,53	254,08	455,83	535,51	107,43
Sinjai	35,67	496,01	311,77	369,70	374,26	575,93	653,22	226,50
Soppeng	48,62	578,44	392,34	451,99	456,44	657,96	735,73	305,93
Tana Toraja	83,25	612,16	426,44	486,61	490,18	691,91	770,81	338,56
Toraja Utara	0,00	531,17	346,15	404,76	409,30	610,95	688,34	260,31
Jeneponto	531,17	0,00	189,45	129,12	122,16	80,97	171,47	277,70
Luwu	346,15	189,45	0,00	71,35	69,74	267,04	352,27	91,96
Luwu Utara	404,76	129,12	71,35	0,00	26,67	207,64	285,67	158,89
Maros	409,30	122,16	69,74	26,67	0,00	201,89	286,95	156,62
Pinrang	610,95	80,97	267,04	207,64	201,89	0,00	106,31	356,24
Takalar	688,34	171,47	352,27	285,67	286,95	106,31	0,00	442,87
Wajo	260,31	277,70	91,96	158,89	156,62	356,24	442,87	0,00



Lampiran 4. Nilai *Silhouette Value* Hasil GKFCM pada Keseluruhan Data

<b>Kabupaten/Kota</b>	<b>cluster</b>	<b>SV</b>
Gowa	1	-0,82
Kota Makassar	1	0,15
Bone	2	0,69
Bulukumba	2	0,73
Kota Palopo	2	0,62
Bantaeng	3	0,88
Barru	3	0,92
Enrekang	3	0,89
Kepulauan Selayar	3	0,91
Kota Parepare	3	0,82
Luwu Timur	3	0,92
Pangkajene Dan Kepulauan	3	0,92
Sidenreng Rappang	3	0,84
Sinjai	3	0,91
Soppeng	3	0,93
Tana Toraja	3	0,93
Toraja Utara	3	0,92
Jeneponto	4	0,86
Luwu	4	0,84
Luwu Utara	4	0,87
Maros	4	0,87
Pinrang	4	0,83
Takalar	4	0,78
Wajo	4	0,77