

**IMPLEMENTASI FUZZY C-MEANS DAN FUZZY  
POSSIBILISTIC C-MEANS UNTUK PENGELOMPOKKAN  
KABUPATEN/KOTA DI PROVINSI BANTEN**

(Studi Kasus: Fasilitas Pelayanan Kesehatan dan Penyakit Menular di Provinsi  
Banten Tahun 2019)

**TUGAS AKHIR**



Widya Sri Mulyaningsih

16611021

**PROGRAM STUDI STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS ISLAM INDONESIA  
YOGYAKARTA  
2021**

**HALAMAN PERSETUJUAN PEMBIMBING**

**TUGAS AKHIR  
ISLAM**

Judul : Implementasi Fuzzy C-Means dan Fuzzy Possibilistic C-Means Untuk Pengelompokan Kabupaten/Kota di Provinsi Banten

Nama Mahasiswa : Widya Sri Mulyaningsih

NIM : 16611021

**TUGAS AKHIR INI TELAH DIPERIKSA DAN DISETUJUI UNTUK  
DIUJIKAN**

Yogyakarta, Januari 2021

**Pembimbing**



(Muhammad Hasan Sidiq Kurniawan, S.Si., M.Sc.)

الجامعة الإسلامية  
الابستدالاندو

## HALAMAN PENGESAHAN

### TUGAS AKHIR

#### IMPLEMENTASI FUZZY C-MEANS DAN FUZZY POSSIBILISTIC C-MEANS UNTUK PENGELOMPOKKAN KABUPATEN/KOTA DI PROVINSI BANTEN

Nama Mahasiswa : Widya Sri Mulyaningsih

NIM : 16611021

TUGAS AKHIR INI TELAH DIUJIKAN  
PADA TANGGAL: 6 Januari 2021

Nama Penguji:

Tanda Tangan

1. Dr. Edy Widodo, S.Si., M.Si.
2. Sekti Kartika Dini, S.Si, M.Si.
3. Muhammad Hasan Sidiq Kurniawan, S.Si., M.Sc.



Mengetahui,

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



(Prof. Riyanto, S.Pd., M.Si., Ph.D.)

## KATA PENGANTAR

*Assalamu'alaikum Wr.Wb*

Puji syukur kehadiran Allah SWT yang telah senantiasa melimpahkan rahmat dan hidayahnya, tidak lupa shalawat serta salam senantiasa tercurahkan kepada Nabi Muhammad SAW, sehingga penulis dapat menyelesaikan skripsi dengan baik.

Tugas akhir yang berjudul “Implementasi Fuzzy C-Means dan Fuzzy Possibilistic C-Means Untuk Pengelompokan Kabupaten/Kota di Provinsi Banten” ini sebagai salah satu syarat untuk memperoleh gelar sarjana Jurusan Statistika di Universitas Islam Indonesia. Selama penyusunan tugas akhir ini penulis tidak lepas dari bantuan berbagai pihak, baik secara moril maupun materil. Pada kesempatan ini penulis bermaksud menyampaikan ucapan terima kasih kepada:

1. Bapak Prof. Riyanto, S.Pd., M.Si., Ph.D. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam beserta jajarannya.
2. Bapak Dr. Edy Widodo, S.Si., M.Si., selaku Ketua Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Indonesia.
3. Bapak Muhammad Hasan Sidiq Kurniawan, S.Si., M.Sc., selaku Dosen Pembimbing atas bimbingan, kritik dan saran yang bersifat membangun.
4. Orang tua dan keluarga yang selalu memberikan do'a, dukungan, dan semangat kepada penulis
5. Teman-teman bimbingan tugas akhir Bapak Hasan, yang telah berjuang bersama menyelesaikan tugas akhir.
6. Teman-teman saya Tata, Yola, Rahma, Vivid dan Nadya yang telah memberikan semangat kepada saya dan membantu dalam penyusunan tugas akhir ini.
7. Serta semua pihak yang telah membantu penulis dalam penyusunan tugas akhir ini.

Penulis menyadari bahwa tugas akhir ini masih terdapat kekurangan didalamnya. Oleh karena itu segala kritik dan saran yang bersifat membangun selalu penulis harapkan. Akhir kata, semoga tugas akhir ini bermanfaat bagi penulis khususnya dan bagi semua yang membutuhkan.

*Wassalamualaikum Wr.Wb*

Yogyakarta, Januari 2021

(Widya Sri Mulyaningsih)



## DAFTAR ISI

HALAMAN SAMPUL .....	i
HALAMAN PERSETUJUAN PEMBIMBING .....	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR .....	iv
DAFTAR ISI.....	vi
DAFTAR TABEL.....	viii
DAFTAR GAMBAR .....	ix
DAFTAR LAMPIRAN.....	x
PERNYATAAN.....	xi
ABSTRAK.....	xii
ABSTRACT.....	xiii
BAB I.....	1
PENDAHULUAN .....	1
1.1    Latar Belakang Masalah .....	1
1.2    Rumusan Masalah .....	3
1.3    Batasan Masalah .....	3
1.4    Jenis Penelitian dan Metode Analisis .....	4
1.5    Tujuan Penelitian.....	4
1.6    Manfaat Penelitian.....	4
BAB II.....	5
TINJAUAN PUSTAKA .....	5
BAB III .....	10
LANDASAN TEORI.....	10
3.1    Data Mining .....	10
3.2 <i>Clustering</i> .....	10
3.3    Himpunan <i>Fuzzy</i> .....	12
3.4 <i>Fuzzy C-Means</i> .....	13
3.5 <i>Possibilistic C-Means</i> (PCM).....	17
3.6 <i>Fuzzy Possibilistic C-Means</i> .....	17

3.7	Validitas Cluster .....	21
3.8	Kesehatan Masyarakat .....	22
3.9	Fasilitas Pelayanan Kesehatan .....	23
3.9.1	Pusat Kesehatan Masyarakat (Puskesmas) .....	24
3.9.2	Rumah Sakit .....	24
3.10	Penyakit Menular .....	25
3.10.1	Tuberkulosis .....	25
3.10.2	Diare .....	25
3.10.3	Pneumonia .....	26
BAB IV	.....	27
METODOLOGI PENELITIAN	.....	27
3.1	Populasi Penelitian .....	27
3.2	Tempat dan Waktu Penelitian .....	27
3.3	Variabel Penelitian .....	27
3.4	Tahap Penelitian .....	28
BAB V	.....	30
HASIL DAN PEMBAHASAN	.....	30
4.1	Analisis Deskriptif .....	30
4.2	Perbandingan <i>Fuzzy C-Means</i> dan <i>Fuzzy Possibilistics C-Means</i> .....	35
4.3	Penentuan Metode Terbaik .....	37
4.4	Penentuan Jumlah <i>Cluster Optimal</i> .....	39
4.5	Profilisasi <i>Fuzzy C-Means</i> terhadap Hasil <i>Cluster</i> .....	40
BAB VI	.....	44
PENUTUP	.....	44
6.1	Kesimpulan .....	44
6.2	Saran .....	44
DAFTAR PUSTAKA	.....	45
RINGKASAN TUGAS AKHIR	.....	49
LAMPIRAN	.....	65

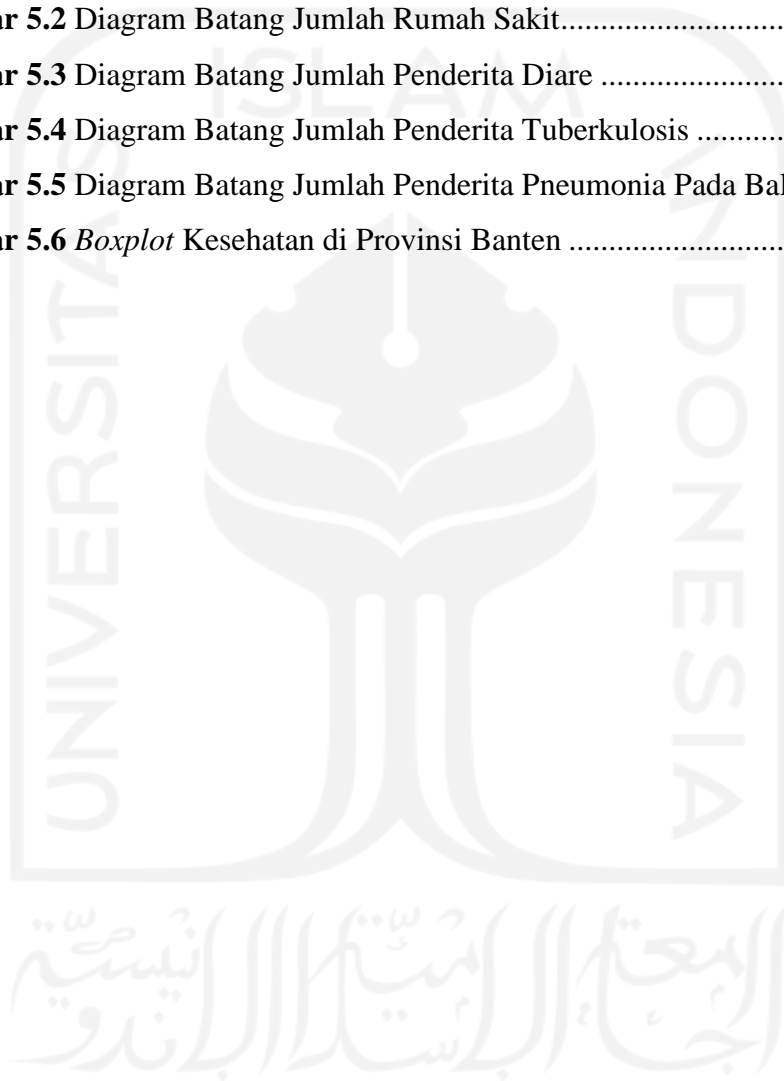
## DAFTAR TABEL

<b>Tabel 2.1</b> Rangkuman Peneliti Terdahulu .....	6
<b>Tabel 3.1</b> Tabel Contoh Studi Kasus .....	15
<b>Tabel 4.1</b> Penjelasan Variabel .....	27
<b>Tabel 5.1</b> Statistika Deskriptif Kesehatan di Provinsi Banten .....	30
<b>Tabel 5.2</b> Jumlah Iterasi dan Fungsi Objektif .....	36
<b>Tabel 5.3</b> Perbandingan <i>Partition Entrophy</i> .....	37
<b>Tabel 5.4</b> Perbandingan <i>Partition Coefficient</i> .....	38
<b>Tabel 5.5</b> Perbandingan <i>Modified Partition Coefficient</i> .....	38
<b>Tabel 5.6</b> Penentuan Jumlah <i>Cluster</i> Optimal.....	39
<b>Tabel 5.7</b> Hasil <i>Cluster 5</i> .....	40
<b>Tabel 5.8</b> Profilisasi Hasil Jumlah <i>Cluster 5</i> .....	42
<b>Tabel 5.9</b> Tingkat Urgensi Kesehatan di Provinsi Banten .....	42



## DAFTAR GAMBAR

<b>Gambar 4.1</b> <i>Flowchart</i> Tahap Penelitian.....	28
<b>Gambar 5.1</b> Diagram Batang Jumlah Puskesmas.....	31
<b>Gambar 5.2</b> Diagram Batang Jumlah Rumah Sakit.....	31
<b>Gambar 5.3</b> Diagram Batang Jumlah Penderita Diare .....	32
<b>Gambar 5.4</b> Diagram Batang Jumlah Penderita Tuberkulosis .....	33
<b>Gambar 5.5</b> Diagram Batang Jumlah Penderita Pneumonia Pada Balita.....	33
<b>Gambar 5.6</b> <i>Boxplot</i> Kesehatan di Provinsi Banten .....	34



## DAFTAR LAMPIRAN

<b>Lampiran 1</b> Data Kesehatan di Provinsi Banten .....	66
<b>Lampiran 2</b> Program R.....	67
<b>Lampiran 3</b> Hasil <i>Output Fuzzy C-Means</i> (FCM).....	69
<b>Lampiran 4</b> Hasil <i>Output Fuzzy Possibilistic C-Means</i> (FPCM).....	72



## PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang sebelumnya pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan orang lain, kecuali yang diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, Januari 2021



Penulis



**IMPLEMENTASI FUZZY C-MEANS DAN FUZZY POSSIBILISTIC C-MEANS UNTUK PENGELOMPOKKAN KABUPATEN/KOTA DI  
PROVINSI BANTEN**

(Studi Kasus: Fasilitas Pelayanan Kesehatan dan Penyakit Menular di Provinsi Banten Tahun 2019)

Widya Sri Mulyaningsih  
Program Studi Statistika, Fakultas MIPA  
Universitas Islam Indonesia

**ABSTRAK**

Kesehatan adalah suatu kondisi agar jiwa dan raga seseorang sejahtera, sehingga menjadikan seseorang produktif dalam menjalankan aktivitas sehari-hari. Upaya pemerintah guna mendukung peningkatan kualitas manusia Indonesia adalah dengan meningkatkan pengendalian penyakit dan meningkatkan pelayanan kesehatan yang mencakup penguatan sistem kesehatan, pengawasan obat dan makanan. Provinsi Banten menempati urutan lima besar di Indonesia untuk kasus penyakit menular secara langsung yaitu penyakit Tuberkulosis, Diare, dan Pneumonia. Pengelompokan Kabupaten/Kota berdasarkan pelayanan kesehatan dan penyakit yang diderita masyarakat diperlukan untuk mengetahui apakah pemerataan kesehatan masyarakat di Indonesia terutama di Provinsi Banten sudah lebih baik. Metode *clustering* yang digunakan pada penelitian ini adalah *Fuzzy C-Means* (FCM) dan *Fuzzy Possibilistic C-Means* (FPCM). Kedua metode tersebut akan dibandingkan keakurasiannya berdasarkan indeks validitas *Partition Entropy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC). Berdasarkan nilai indeks validitas *cluster Partition Entropy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC) didapatkan hasil bahwa metode *Fuzzy C-Means* (FCM) merupakan metode yang terbaik dibandingkan metode *Fuzzy Possibilistic C-Means* (FPCM). Jumlah *cluster* yang optimal pada metode *Fuzzy C-Means* (FCM) adalah 5 *cluster*. Dari kelima *cluster*, *cluster* 1 merupakan *cluster* yang memiliki tingkat urgensi kesehatan tertinggi di Provinsi Banten. Hal tersebut dikarenakan adanya ketimpangan antara jumlah pelayanan kesehatan dengan jumlah penderita penyakit menular.

**Kata Kunci:** *Fuzzy C-Means, Fuzzy Possibilistic C-Means, Modified Partition Coefficient, Partition Coefficient, Partition Entropy*

**IMPLEMENTATION FUZZY C-MEANS AND FUZZY POSSIBILISTIC  
C-MEANS FOR DISTRICT / CITY GROUPING IN BANTEN PROVINCE  
(Case Study: Health Service Facilities and Infectious Diseases of Banten  
Province in 2019)**

Widya Sri Mulyaningsih  
Program Studi Statistika, Fakultas MIPA  
Universitas Islam Indonesia

**ABSTRACT**

*Health is a condition for a person's body and soul to be prosperous, thus making someone productive in carrying out their daily activities. The government's efforts to support the improvement of the quality of Indonesian people are by increasing disease control and improving health services which include strengthening the health system, controlling drugs and food. Banten Province ranks in the top five in Indonesia for cases of direct infectious diseases, namely Tuberculosis, Diarrhea and Pneumonia. District / city grouping based on health services and diseases suffered by the community is needed to determine whether the distribution of public health in Indonesia, especially in Banten Province. The clustering method used in this research is Fuzzy C-Means (FCM) and Fuzzy Possibilistic C-Means (FPCM). The accuracy of the two methods will be compared based on the validity index of the Partition Entrophy (PE), Partition Coefficient (PC) and Modified Partition Coefficient (MPC). Based on the value of the cluster validity index for Partition Entrophy (PE), Partition Coefficient (PC) and Modified Partition Coefficient (MPC), the results show that the Fuzzy C-Means (FCM) method is the best method compared to the Fuzzy Possibilistic C-Means (FPCM) method. The optimal number of clusters in the Fuzzy C-Means (FCM) method is 5 clusters. Of the five clusters, cluster 1 is the cluster that has the highest level of health urgency in Banten Province. This is due to the imbalance between the number of health services and the number of people with infectious diseases.*

**Keywords:** *Fuzzy C-Means, Fuzzy Possibilistic C-Means, Modified Partition Coefficient, Partition Coefficient, Partition Entrophy*

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang Masalah**

Kesehatan adalah suatu kondisi agar jiwa dan raga seseorang sejahtera, sehingga menjadikan seseorang produktif dalam menjalankan aktivitas sehari-hari. Menurut Undang - Undang Nomor 23 Tahun 1992, kesehatan adalah keadaan sejahtera dari badan, jiwa, dan sosial yang memungkinkan setiap orang hidup produktif secara sosial dan ekonomis.

Sesuai Edaran Menteri PPN/Bappenas No. B.899/M.PPN/Ses/PP.03.02/12/2019 tanggal 20 Desember 2019 guna mendukung peningkatan kualitas manusia Indonesia, kementerian kesehatan memiliki misi yaitu memperkuat upaya kesehatan yang bermutu dan menjangkau seluruh penduduk Indonesia, memberdayakan masyarakat dan mengutamakan pembangunan kesehatan dan meningkatkan ketersediaan, pemerataan dan mutu sumber daya kesehatan. Maka upaya pemerintah adalah dengan meningkatkan pengendalian penyakit dan meningkatkan pelayanan kesehatan yang mencakup penguatan sistem kesehatan, pengawasan obat dan makanan (KementerianKesehatan, 2020).

Pemerataan kualitas kesehatan secara nasional di Indonesia cukup sulit dilakukan karena Indonesia merupakan negara kepulauan dengan jumlah penduduk yang besar. Oleh karena itu, pemerataan kualitas masyarakat dimulai dari kesehatan individu setiap masyarakat yang ditunjang oleh kesehatan keluarga, kemudian menuju tingkat yang lebih luas seperti desa/kelurahan, kabupaten dan provinsi.

Provinsi Banten menempati urutan lima besar di Indonesia untuk kasus penyakit menular secara langsung yaitu penyakit Tuberkulosis, Diare, dan Pneumonia (KementerianKesehatan, 2020). Provinsi Banten yang terdiri dari 8 kabupaten/kota yaitu Kabupaten Pandeglang, Kabupaten Lebak, Kabupaten Tangerang, Kota Tangerang, Kota Tangerang Selatan, Kabupaten Serang, Kota Serang, dan Kota Cilegon dengan 155 kecamatan serta 1552 desa/kelurahan

memiliki beberapa permasalahan kesehatan. Menurut Dinas Kesehatan Provinsi Banten beberapa permasalahan kesehatan tersebut adalah penyakit Tuberkulosis yang memiliki *Case Notification Rate* (CNR) seluruh kasus sebesar 150,53 per 100.000 penduduk, penyakit Diare yang memiliki jumlah kasus sebesar 258.642 per 1.000, dan penyakit Pneumonia yang memiliki jumlah kasus sebesar 36.205 yang mengalami kenaikan dari tahun 2017. Fasilitas pelayanan kesehatan yang menyelenggarakan upaya kesehatan masyarakat juga sangat penting dalam menunjang kesehatan masyarakat. Fasilitas pelayan kesehatan meliputi puskesmas, rumah sakit, dan lain lain. Di provinsi Banten jumlah Puskesmas ada sebanyak 238 dan jumlah rumah sakit sebanyak 112. Jumlah rumah sakit khusus dan rumah sakit umum di Provinsi Banten mengalami kenaikan dari tahun 2018 karena banyak rumah sakit yang baru beroperasi dan rasio puskesmas per kecamatan di Provinsi Banten sudah di atas rasio nasional yaitu sebesar 1,57 (DinasKesehatan, 2019).

Pengelompokkan Kabupaten/Kota berdasarkan pelayanan kesehatan dan penyakit yang diderita masyarakat diperlukan untuk mengetahui apakah pemerataan kesehatan masyarakat di Indonesia terutama di Provinsi Banten sudah lebih baik. Oleh karena itu, dalam penelitian ini akan dilakukan penelitian untuk mengetahui pengelompokkan Kabupaten/Kota di Provinsi Banten. Variabel-variabel yang akan diamati adalah jumlah puskesmas dan rumah sakit, jumlah penderita penyakit tuberkulosis, diare, dan pneumonia.

Analisis *clustering* merupakan proses pembagian data ke dalam suatu kelompok berdasarkan tingkat kesamaannya. Data yang memiliki kemiripan karakteristik akan berkumpul dalam satu kelompok, sementara data yang memiliki perbedaan karakteristik akan berkumpul dalam kelompok yang berbeda (Siyamto, 2017).

Salah satu jenis metode *clustering* adalah metode *partition clustering* yang merupakan suatu metode penyusunan partisi dari objek ke dalam suatu *cluster*. Dengan kata lain, jumlah *cluster* dalam hal ini sudah ditentukan sebelumnya. Pada penelitian ini, analisis difokuskan pada metode *partition clustering* yakni *fuzzy* karena metode tersebut dapat mengelompokkan suatu data yang memiliki kemungkinan berada pada lebih dari satu *cluster*. Metode *fuzzy* yang digunakan



untuk dibandingkan pada penelitian ini adalah *Fuzzy C-Means* (FCM) dan *Fuzzy Possibilistic C-Means* (FPCM). *Fuzzy C-Means* (FCM) merupakan suatu teknik pengelompokan data dimana keberadaan tiap-tiap titik data dalam suatu kelompok ditentukan oleh derajat keanggotaan data tersebut (Kusumadewi, dkk., 2006). *Fuzzy Possibilistic C-Means* (FPCM) merupakan pengembangan dari dua algoritma yaitu *Fuzzy C-Means* (FCM) dan *Possibilistic C-Means* (PCM). Kelebihan *Fuzzy Possibilistic C-Means* (FPCM) adalah dapat memecahkan permasalahan dua algoritma berupa data yang masih mengandung *error (noisy data)* pada *Fuzzy C-Means* (FCM) dan permasalahan ketika pengelompokan data pada algoritma *Possibilistic C-Means* (PCM) (Grover, 2014). Kedua metode tersebut akan dibandingkan keakurasiannya berdasarkan indeks validitas *Partition Entropy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC) sehingga akan didapatkan metode yang tepat dalam pengelompokan kabupaten/kota di Provinsi Banten berdasarkan fasilitas pelayanan kesehatan dan penderita penyakit menular. Pentingnya profilisasi hasil *cluster* adalah untuk mengetahui karakteristik dan sifat-sifat dari masing-masing *cluster* dan memberi nama atau label yang tepat.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka dapat dirumuskan beberapa permasalahan yang akan diteliti yaitu:

1. Bagaimana hasil perbandingan indeks validitas *cluster* pada metode *Fuzzy C-Means* dan *Fuzzy Possibilistic C-Means* pada data fasilitas pelayanan kesehatan dan penderita penyakit menular di Provinsi Banten?
2. Bagaimana hasil profilisasi pada pengelompokan kabupaten/kota di Provinsi Banten berdasarkan fasilitas pelayanan kesehatan dan penderita penyakit menular?

## 1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah:

1. Data yang digunakan dalam penelitian ini adalah data jumlah penderita penyakit tuberkulosis, diare dan pneumonia serta jumlah PUSKESMAS (Pusat Kesehatan Masyarakat) dan rumah sakit tahun 2019.



2. Sumber data yang digunakan dalam penelitian ini adalah data sekunder yang didapatkan dari *e-book* Dinas Kesehatan Provinsi Banten dan Badan PPSDM Kesehatan.
3. Metode yang digunakan dalam penelitian ini adalah metode *Fuzzy C-Means* (FCM) dan *Fuzzy Possibilistic C-Means* (FPCM).
4. Data diolah dengan menggunakan bantuan *software* R dan *Microsoft Excel*.

#### **1.4 Jenis Penelitian dan Metode Analisis**

Jenis penelitian pada tugas akhir ini adalah penelitian aplikatif. Metode yang digunakan adalah metode *Fuzzy C-Means* dan *Fuzzy Possibilistic C-Means* pada data fasilitas pelayanan kesehatan dan penderita penyakit menular berdasarkan keakurasiannya ketika meng*cluster*.

#### **1.5 Tujuan Penelitian**

Adapun tujuan yang ingin dicapai dalam penelitian ini yaitu:

1. Untuk mengetahui hasil perbandingan indeks validitas *cluster* pada metode *Fuzzy C-Means* dan *Fuzzy Possibilitic C-Means* pada data fasilitas pelayanan kesehatan dan penderita penyakit menular di Provinsi Banten.
2. Untuk mengetahui hasil profilisasi pada pengelompokkan kabupaten/kota di Provinsi Banten berdasarkan fasilitas pelayanan kesehatan dan penderita penyakit menular.

#### **1.6 Manfaat Penelitian**

Manfaat dari penelitian ini adalah sebagai salah satu sarana bagi pembaca dan pelengkap referensi bagi pihak-pihak yang memerlukan dalam menyelesaikan kasus untuk pengelompokkan yang menggunakan metode *clustering* terutama metode *Fuzzy C-Means* dan *Fuzzy Possibilitics C-Means*. Selain itu juga dapat menjadi dasar pengambilan keputusan dan kebijakan pemerintah daerah yang mengarah pada upaya peningkatan pembangunan fasilitas kesehatan di seluruh wilayah Banten berdasarkan studi kasus yang digunakan.

## BAB II

### TINJAUAN PUSTAKA

Terdapat beberapa penelitian yang berkaitan dengan metode *Fuzzy C-Means* dan *Fuzzy Possibilitic C-Means* sebelumnya yang ditinjau dari berbagai sudut pandang para penelitiannya. Diantara penelitian-penelitian sebelumnya adalah penelitian yang dilakukan oleh Christian Correa, Pillar Barreiro, Constantino Valero, Maria-Paz Diago, dan Javier Tardaguila (2014) dengan judul “*A Comparison of Fuzzy Clustering Algorithms Applied to Feature Extraction on Vineyard*” dengan hasil penelitian bahwa metode terbaik berdasarkan kecepatan dan performanya adalah metode *Fuzzy C-Means* (FCM) dibandingkan dengan metode *Possibilistic C-Means* (PCM), *Fuzzy Possibilistic C-Means* (FPCM), *Robust Fuzzy Possibilistic C-Means* (RFCM) dan *Fuzzy C-Means with Gustafson-Kessel algorithm* (FCM-GK).

Pada penelitian yang dilakukan T.Thilagaraj dan Dr. N Sengottaiyan (2019) dengan judul “*Implementation of Fuzzy C-Means and Fuzzy Possibilistic C-Means Algorithms to Find the Low Performers Using R-Tool*” memperoleh hasil bahwa metode terbaik berdasarkan performanya menggunakan *R-Tool* adalah metode *Fuzzy Possibilistic C-Means* (FPCM) dibandingkan metode *Fuzzy C-Means* (FCM).

Selanjutnya, penelitian dengan judul “*Perbandingan Metode C-Means dan Fuzzy C-Means dalam Pengelompokan Wilayah Desa/Kelurahan di Kabupaten Kutai Kartanegara*” yang dilakukan oleh Nissa Irabawati, Sri Wahyuningsih, dan Rudy Ramadani Syoer (2016) menyatakan bahwa metode terbaik berdasarkan fungsi objektif dan waktu komputasinya adalah metode *Fuzzy C-Means* (FCM) dengan pengelompokan 5 *cluster*.

Pada penelitian yang dilakukan oleh Wahidah Sanusi, Ahmad Zaky, dan Besse Nur Afni (2018) dengan judul “*Analisis Fuzzy C-Means dan Penerapannya Dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Faktor-Faktor Penyebab Gizi Buruk*” menyatakan bahwa dengan

menggunakan metode *Fuzzy C-Means* (FCM) didapatkan pengelompokan 2 *cluster* dengan fungsi objektif sebesar 1079141912,2224.

Terdapat juga beberapa penelitian yang berkaitan dengan kesehatan, seperti penelitian terkait penyakit tuberkulosis, penyakit diare, penyakit pneumonia dan pelayanan kesehatan. Penelitian-penelitian tersebut antara lain adalah penelitian yang dilakukan oleh Novianti Puspitasari, Joan Angelina Widians dan Pohny (2018) dengan judul “*A Clustering of Generative and Infectious Diseases Using Fuzzy C-Means*” didapatkan hasil pengelompokan 2 *cluster* dengan perhitungan validitas *cluster* menggunakan *Partition Coefficient* (PC) sebesar 0,685426108.

Selanjutnya, penelitian dengan judul “Penerapan Data Mining Untuk Pengelompokan Penyakit dengan Algoritma *Fuzzy C-Means*” yang dilakukan oleh Shinta Siti Sundari dan Nida Ariani (2019) didapatkan hasil bahwa pengujian sistem menghasilkan 4 *cluster* dengan nilai akurasi sebesar 76% dan nilai *Partition Coefficient Index* (PCI) sebesar 0,6154.

Penelitian dengan judul “*The Implementation of Hybrid Fuzzy Clustering on The Public Health Facility Data*” oleh Samingun Handoyo, Agus Widodo, Waego Hadi Nugroho, dan Imam N Purwanto (2019) didapatkan hasil bahwa metode terbaik untuk pengelompokan data fasilitas kesehatan adalah metode *Fuzzy C-Means* (FCM) dibandingkan dengan metode *Subtractive Fuzzy C-Means* (SFCM).

Pada Penelitian yang dilakukan oleh Siti Syahidatul Helma, Mustakim, Risma Rustiyan R, dan Eva Normala (2019) dengan judul “*Clustering Pada Data Fasilitas Pelayanan Kesehatan Kota Pekanbaru Menggunakan Algoritma K-Means*” didapatkan hasil bahwa *cluster* optimal yang didapatkan menggunakan metode *K-Means* adalah 9 *cluster*.

**Tabel 2.1** Rangkuman Peneliti Terdahulu

No	Nama	Tahun	Judul	Metode	Hasil
1.	Christian Correa, Pillar Barreiro, Constantino Valero, Maria-Paz Diago, dan Javier Tardaguila	2011	<i>A Comparison of Fuzzy Clustering Algorithms Applied to Feature Extraction on Vineyard</i>	<i>Fuzzy C-Means</i> (FCM), <i>Possibilistic C-Means</i> (PCM), <i>Fuzzy Possibilistic</i>	Metode terbaik berdasarkan kecepatan dan performanya adalah metode <i>Fuzzy C-Means</i> (FCM) dibandingkan

No	Nama	Tahun	Judul	Metode	Hasil
				<i>C-Means</i> (FPCM), <i>Robust Fuzzy Possibilistic C-Means</i> (RFCM) dan <i>Fuzzy C-Means with Gustafson-Kessel algorithm</i> (FCM-GK).	dengan metode lainnya
2.	T.Thilagaraj dan Dr. N Sengottaiyan	2019	<i>Implementation of Fuzzy C-Means and Fuzzy Possibilistic C-Means Algorithms to Find the Low Performers Using R-Tool</i>	<i>Fuzzy C-Means</i> (FCM) dan <i>Fuzzy Possibilistic C-Means</i> (FPCM)	Metode terbaik berdasarkan performanya menggunakan <i>R-Tool</i> adalah metode <i>Fuzzy Possibilistic C-Means</i> (FPCM)
3.	Nissa Irabawati, Sri Wahyuningsih, dan Rudy Ramadani Syoer	2016	Perbandingan Metode <i>C-Means</i> dan <i>Fuzzy C-Means</i> dalam Pengelompokkan Wilayah Desa/Kelurahan di Kabupaten Kutai Kartanegara	<i>C-Means</i> dan <i>Fuzzy C-Means</i> (FCM)	Metode terbaik berdasarkan fungsi objektif dan waktu komputasinya adalah metode <i>Fuzzy C-Means</i> (FCM) dengan pengelompokkan 5 cluster
4.	Wahidah Sanusi, Ahmad Zaky, dan Besse Nur Afni	2018	Analisis <i>Fuzzy C-Means</i> dan Penerapannya Dalam Pengelompokkan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Faktor-Faktor Penyebab Gizi Buruk	<i>Fuzzy C-Means</i> (FCM)	Didapatkan pengelompokkan 2 cluster dengan fungsi objektif sebesar 1079141912,2224
5.	Novianti Puspitasari,	2018	<i>A Clustering of Generative and Infectious</i>	<i>Fuzzy C-Means</i> (FCM)	Didapatkan hasil pengelompokkan 2 cluster dengan

No	Nama	Tahun	Judul	Metode	Hasil
	Widians dan Pohny		<i>Diseases Using Fuzzy C-Means</i>		perhitungan validitas <i>cluster</i> menggunakan <i>Partition Coefficient</i> (PC) sebesar 0,685426108
6.	Shinta Siti Sundari dan Nida Ariani	2019	Penerapan Data Mining Untuk Pengelompokan Penyakit dengan Algoritma <i>Fuzzy C-Means</i>	<i>Fuzzy C-Means</i> (FCM)	Didapatkan hasil bahwa pengujian sistem menghasilkan 4 <i>cluster</i> dengan nilai akurasi sebesar 76% dan nilai <i>Partition Coefficient Index</i> (PCI) sebesar 0,6154
7.	Samingun Handoyo, Agus Widodo, Waego	2019	<i>The Implementation of Hybrid Fuzzy Clustering on The Public Health Facility Data</i>	<i>Fuzzy C-Means</i> (FCM) dan <i>Subtractive Fuzzy C-Means</i> (SFCM)	Didapatkan hasil bahwa metode terbaik untuk pengelompokan data fasilitas kesehatan adalah metode <i>Fuzzy C-Means</i> (FCM) dibandingkan dengan metode <i>Subtractive Fuzzy C-Means</i> (SFCM)
8.	Siti Syahidatul Helma, Mustakim, Risma Rustiyan R, dan Eva Normala	2019	<i>Clustering Pada Data Fasilitas Pelayanan Kesehatan Kota Pekanbaru Menggunakan Algoritma K-Means</i>	<i>Algoritma K-Means</i>	Didapatkan hasil bahwa <i>cluster</i> optimal yang didapatkan menggunakan metode <i>K-Means</i> adalah 9 <i>cluster</i>

Perbedaan penelitian ini dengan penelitian sebelumnya adalah dibandingkannya 2 metode yaitu *Fuzzy C-Means* dan *Fuzzy Possibilitic C-Means* dan studi kasus penelitian. Pada penelitian ini, studi kasus yang digunakan adalah data fasilitas pelayanan kesehatan dan penderita penyakit menular. Data yang

digunakan adalah jumlah fasilitas pelayanan kesehatan yaitu Puskesmas dan Rumah sakit dan jumlah penderita penyakit menular secara langsung yaitu tuberkulosis, diare, dan pneumonia di Provinsi Banten pada tahun 2019.



## **BAB III**

### **LANDASAN TEORI**

#### **3.1 Data Mining**

*Data mining* adalah istilah yang digunakan untuk menggambarkan penemuan informasi yang bermanfaat dan pengetahuan dalam *database* yang prosesnya menggunakan teknik matematika, statistik, *machine learning*, dan kecerdasan buatan. Sebelumnya istilah *data mining* digunakan untuk menggambarkan proses identifikasi pola yang tidak ditemukan dalam data, namun seiring waktu definisi asli telah dimodifikasi untuk menyertakan sebagian besar jenis analisis data otomatis (Turban, dkk., 2005).

*Data mining* sering disebut sebagai *Knowledge Discover in Database* (KDD), yaitu suatu kegiatan yang meliputi pemakaian data, pengumpulan, dan historis untuk menemukan hubungan atau pola, dan keteraturan dalam set data berukuran besar. *Output* dari *data mining* bisa dipakai untuk memperbaiki pengambilan keputusan di masa depan (Santoso, 2007). Secara sederhana *data mining* adalah penemuan atau penambahan informasi baru dengan mencari aturan atau pola tertentu dari sejumlah data yang berukuran besar.

Karakteristik *data mining* yaitu *data mining* berhubungan dengan penemuan pola data yang tersembunyi dan tidak terduga, *data mining* biasa menggunakan data yang sangat besar untuk menghasilkan kesimpulan yang baik dalam kaitannya dengan pola data, dan *data mining* juga berguna dalam membuat keputusan yang kritis terutama yang bersifat strategis (Davies, 2004). Fungsi *data mining* terbagi menjadi enam yaitu, fungsi deskripsi, fungsi pengelompokkan, fungsi klasifikasi, fungsi prediksi, fungsi estimasi, dan fungsi asosiasi (Larose, 2005). Beberapa metode yang termasuk dalam data mining adalah *association rules*, *decision tree*, *clustering*, dan lain-lain.

#### **3.2 Clustering**

*Clustering* atau biasa disebut klaterisasi adalah suatu metode pengelompokkan data tanpa harus didasarkan pada kelas data yang telah ditentukan



sebelumnya. Tujuan *clustering* adalah memaksimalkan kemiripan antar anggota dalam suatu kelas dan meminimalkan kemiripan antar *cluster* (David, dkk., 2020). *Clustering* adalah teknik yang pengelompokannya dilakukan atas dasar kesamaan atau jarak, input yang diperlukan adalah ukuran kesamaan atau data dari mana persamaan dapat dihitung (Johnson & Wichern, 2007).

*Clustering* adalah teknik analisis untuk mengelompokkan individu atau objek kedalam sejumlah kecil kelompok yang memiliki kesamaan. Tujuan utama dari analisis *cluster* adalah untuk membagi sekelompok objek menjadi dua atau lebih kelompok berdasarkan kesamaan objek untuk satu set karakteristik tertentu. Dalam analisis *cluster* grup tidak ditentukan sebelumnya, sebaliknya teknik digunakan untuk mengidentifikasi kelompok. Analisis *cluster* melibatkan setidaknya tiga langkah. Pertama adalah pengukuran beberapa bentuk persamaan antar objek untuk menentukan berapa banyak kelompok yang ada pada sampel. Kedua adalah proses dimana objek dibagi menjadi beberapa kelompok. Ketiga adalah membuat profilisasi variabel untuk menentukan komposisinya (Anderson, dkk., 2014).

Tujuan *clustering* adalah untuk mengelompokkan sejumlah objek atau data kedalam *cluster* sehingga setiap *cluster* akan terisi objek atau data yang memiliki karakteristik yang sama. Prinsip yang digunakan dalam *clustering* adalah memaksimalkan homogenitas (kesamaan) dan memaksimalkan heterogenitas (ketidaksamaan) antar kelompok (Kusuma & Agani, Prototipe Komparasi Model Clustering Menggunakan Metode K-Means Dan FCM untuk Menentukan Strategi Promosi, 2015). Analisis *clustering* merupakan proses pembagian data ke dalam suatu kelompok berdasarkan tingkat kesamaannya. Data yang memiliki kemiripan karakteristik akan berkumpul dalam satu kelompok, sementara data yang memiliki perbedaan karakteristik akan berkumpul dalam kelompok yang berbeda (Siyamto, 2017).

*Clustering* menggunakan tiga konsep partisi yang dapat digunakan, yaitu partisi klasik, partisi *fuzzy* dan partisi *possibilistic*. Pada partisi klasik, suatu data akan menjadi anggota hanya pada satu *cluster* saja. Pada partisi *fuzzy*, nilai keanggotaan suatu data pada suatu *cluster* terletak pada interval 0 sampai 1 dan jumlah derajat keanggotaan setiap data pada semua *cluster* adalah 1. Sedangkan



pada partisi *possibilistics*, untuk menjamin suatu data menjadi anggota dari paling tidak satu *cluster* maka diharuskan jumlah nilai keanggotaannya lebih dari 0 tidak harus 1 (Kusumadewi, dkk., 2006).

### 3.3 Himpunan *Fuzzy*

Konsep dasar teori himpunan *fuzzy* pertama kali diperkenalkan pada tahun 1965 oleh Zadeh. Himpunan *fuzzy* merupakan fungsi yang derajat keanggotaannya dinyatakan dalam suatu bilangan riil dalam interval tertutup (Susilo, 2006). Fungsi keanggotaan himpunan *fuzzy*  $A$  pada himpunan semesta  $X$ , dinotasikan dengan  $\mu_A$  yaitu:

$$\mu_A : X \rightarrow [0,1] \quad (3.1)$$

Teori himpunan *fuzzy* merupakan kerangka matematis yang digunakan untuk menjelaskan kekurangan informasi, ketidakpastian, ketidaktepatan, dan ketidakjelasan (Kusumadewi & Purnomo, 2010). Ada beberapa hal yang perlu diketahui dalam memahami himpunan *fuzzy*, yaitu:

1. Variabel *fuzzy*

Variabel *fuzzy* merupakan suatu kata atau lambang yang menunjukkan kepada suatu yang tidak menentu dalam sistem *fuzzy*, contohnya: permintaan (naik dan turun), persediaan (banyak dan sedikit), dan sebagainya.

2. Himpunan *fuzzy*

Himpunan *fuzzy* merupakan suatu kumpulan yang mewakili suatu kondisi atau keadaan tertentu dalam suatu variabel *fuzzy*. Himpunan *fuzzy* memiliki 2 atribut, yaitu:

- a. Linguistik, yaitu penamaan suatu grup yang memiliki suatu kondisi atau keadaan tertentu dengan menggunakan Bahasa, seperti: muda, paruh baya, dan tua
- b. Numeris, yaitu suatu nilai yang berbentuk angka yang menunjukkan ukuran dari suatu variabel, seperti: 5, 10, dan 15

3. Semesta Pembicaraan

Semesta pembicaraan adalah keseluruhan nilai yang diperoleh untuk dioperasikan dalam suatu variabel *fuzzy*, contohnya: semesta pembicaraan untuk variabel umur  $X = [0,100]$

#### 4. Domain

Domain adalah keseluruhan nilai yang diizinkan dalam semesta pembicaraan dan boleh dioperasikan dalam suatu himpunan *fuzzy*, contohnya: himpunan *fuzzy* muda =  $[0,30]$  artinya seseorang dinyatakan muda jika umur diantara 0 sampai 30 tahun.

### 3.4 *Fuzzy C-Means*

Metode *fuzzy clustering* memungkinkan objek menjadi bagian dari beberapa kelompok secara bersamaan, dengan derajat keanggotaan yang berbeda. Derajat keanggotaan berada di interval 0 sampai 1 (Balasko, dkk., 2007). Salah satu *clustering* non hirarki yang termasuk dalam metode *fuzzy clustering* adalah *Fuzzy C-Means* (FCM). *Fuzzy C-Means* (FCM) merupakan model pengelompokan *fuzzy* dengan menggunakan *euclidean distance* sehingga objek atau data dapat menjadi anggota dari semua kelas suatu kelompok atau *cluster* yang terbentuk dari derajat keanggotaan yang berada antara 0 sampai 1 (Kusuma & Sangadji, 2016). Metode *Fuzzy C-Means* (FCM) memiliki kekurangan yaitu membutuhkan banyaknya kelompok dan matriks keanggotaan kelompok awal diinisialisasikan secara acak, sehingga metode FCM memiliki masalah inkonsistensi (Haqiqi & Kurniawan, 2015).

Konsep dasar FCM pertama kali adalah menentukan pusat *cluster* yang akan menandai lokasi rata-rata setiap *cluster*. Pada kondisi awal untuk pusat *cluster* ini kemungkinan belum akurat. Selain itu, FCM memberikan setiap titik data nilai keanggotaan untuk setiap *cluster*. Dengan cara memperbaharui pusat *cluster* dan nilai keanggotaan secara berulang untuk setiap titik data, maka pusat *cluster* akan menuju lokasi yang tepat dalam kumpulan data. Perulangan iterasi ini didasarkan pada meminimalkan fungsi obyektif yang mewakili jarak dari titik data tertentu ke pusat *cluster* yang dibobotkan oleh tingkat keanggotaan titik data tersebut (Ned &

Roger, 2017). Adapun langkah-langkah metode *Fuzzy C-Means* (FCM) adalah sebagai berikut:

1. Memasukkan data berupa matriks berukuran  $n \times m$  yang memiliki fungsi untuk menentukan jumlah data dan jumlah atribut atau variabel setiap data yang akan digunakan.

$n$  = jumlah sampel data

$m$  = atribut setiap data

$X_{ij}$  = data sampel ke- $i$  ( $i=1,2,\dots,n$ ), atribut ke- $j$  ( $j=1,2,\dots,m$ )

$$X_{ij} = \begin{bmatrix} X_{11} & \dots & X_{1m} \\ \vdots & \ddots & \vdots \\ X_{n1} & \dots & X_{nm} \end{bmatrix} \quad (3.2)$$

2. Menentukan:

a. Jumlah *cluster* =  $c$  ( $\geq 2$ );

b. Bobot pangkat =  $w$  ( $> 1$ );

c. Maksimum iterasi = MaxIter

d. Error terkecil yang diharapkan =  $\xi$

e. Fungsi objektif awal =  $P_0 = 0$

f. Iterasi awal =  $t = 1$

3. Membangkitkan bilangan *random* untuk mengisi elemen-elemen matriks partisi awal  $U$  atau matriks kekhasan relatif dengan rentang nilai 0 sampai 1. Berikut adalah bentuk matriks partisi awal:

$$\mu_{ik} = \begin{bmatrix} \mu_{11}(X_1) & \dots & \mu_{1m}(X_1) \\ \vdots & \ddots & \vdots \\ \mu_{n1}(X_n) & \dots & \mu_{nm}(X_n) \end{bmatrix}; 1 \leq i \leq n; 1 \leq k \leq c; \quad (3.3)$$

Jumlah setiap nilai elemen kolom dalam satu baris adalah 1.

4. Menghitung pusat *cluster* ke- $k$  dengan persamaan sebagai berikut

$$V_{kj} = \frac{\sum_{i=1}^n (\mu_{ik})^w X_{kj}}{\sum_{i=1}^n (\mu_{ik})^w}; 1 \leq i \leq n; 1 \leq j \leq m; 1 \leq k \leq c \quad (3.4)$$

$$V = \begin{bmatrix} V_{11} & \dots & V_{1m} \\ \vdots & \ddots & \vdots \\ V_{c1} & \dots & V_{cm} \end{bmatrix} \quad (3.5)$$

5. Menghitung fungsi objektif pada iterasi ke =  $t$  dengan persamaan sebagai berikut

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left( \left[ \sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (3.6)$$

6. Menghitung perubahan matriks partisi dengan persamaan sebagai berikut

$$\mu_{ik} = \frac{\left[ \sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-1}}{\sum_{k=1}^c \left[ \sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-1}} ; 1 \leq i \leq n ; 1 \leq j \leq m ; 1 \leq k \leq c \quad (3.7)$$

7. Mengecek kondisi berhenti:

- Jika  $(|P_t - P_{t-1}| < \xi)$  atau  $(t > \text{MaxIter})$  maka berhenti;
- Jika tidak,  $t = t + 1$ , ulangi langkah ke-4

Hasil  $\mu_{ik}$  (matriks partisi) berpengaruh terhadap hasil fungsi objektif, jika hasil  $\mu_{ik}$  kecil maka hasil fungsi objektif juga akan kecil. Hasil fungsi objektif akan memberikan *clustering* yang optimum.

Contoh studi kasus (Permatasari, 2012) : Pengelompokkan ibu hamil berdasarkan tingkat resiko preeklamsia

**Tabel 3.1** Tabel Contoh Studi Kasus

Nama	Umur Ibu	Umur Kehamilan	Sistole	Diastole	Proteinuria
Pasien 1	25	7	140	90	0
Pasien 2	30	38	170	120	3
Pasien 3	28	26	160	100	2
Pasien 4	26	20	130	80	0
Pasien 5	27	24	150	90	1
Pasien 6	26	16	130	90	0
Pasien 7	18	28	130	80	0
Pasien 8	20	20	130	80	0
Pasien 9	24	30	150	90	1
Pasien 10	24	22	140	80	0

Adapun langkah-langkah metode *Fuzzy C-Means* (FCM) adalah sebagai berikut:

- Memasukkan data berupa matriks berukuran  $n \times m$  yang memiliki fungsi untuk menentukan jumlah data dan jumlah atribut atau variabel setiap data yang akan digunakan.

$$X_{ij} = \begin{bmatrix} 25 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 24 & \dots & 0 \end{bmatrix}$$

- Menentukan:

- a. Jumlah *cluster* = 2
  - b. Bobot pangkat = 2
  - c. Maksimum iterasi = 100
  - d. Error terkecil yang diharapkan = 0.00001
  - e. Fungsi objektif awal = 0
  - f. Iterasi awal = 1
3. Membangkitkan bilangan *random* untuk mengisi elemen-elemen matriks partisi awal U atau matriks kekhasan relatif dengan rentang nilai 0 sampai 1. Berikut adalah bentuk matriks partisi awal:

$$\mu_{ik} = \begin{bmatrix} 0.453 & 0.376 & 0.764 & 0.617 & 0.551 & 0.711 & 0.312 & 0.419 & 0.762 & 0.343 \\ 0.547 & 0.624 & 0.236 & 0.383 & 0.449 & 0.289 & 0.688 & 0.581 & 0.238 & 0.657 \end{bmatrix}$$

Jumlah setiap nilai elemen kolom dalam satu baris adalah 1.

4. Menghitung pusat *cluster*

$$V_{11} = \frac{\sum_{i=1}^{10} (\mu_{i1})^2 X_{i1}}{\sum_{i=1}^{10} (\mu_{i1})^2} = \frac{((0.453)^2 25) + \dots + ((0.343)^2 24)}{(0.453)^2 + \dots + (0.343)^2} = 25.54795$$

$$V_{kj} = \begin{bmatrix} 25.54795 & 23.03823 & 144.2592 & 90.76531 & 0.800884 \\ 23.8922 & 23.69332 & 142.0064 & 89.33172 & 0.621246 \end{bmatrix}$$

5. Menghitung fungsi objektif pada iterasi ke-1

$$\begin{aligned} P_1 &= \sum_{i=1}^{10} \sum_{k=1}^2 \left( \left[ \sum_{j=1}^5 (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^2 \right) \\ &= ([276.893](0.453)^2 + \dots + [92.582](0.657)^2) \\ &= 2161.071 \end{aligned}$$

6. Menghitung perubahan matriks partisi

$$\begin{aligned} \mu_{11} &= \frac{\left[ \sum_{j=1}^5 (X_{1j} - V_{1j})^2 \right]^{-1}}{\sum_{k=1}^2 \left[ \sum_{j=1}^5 (X_{1j} - V_{kj})^2 \right]^{-1}} = \frac{1 / \left[ \sum_{j=1}^5 (X_{1j} - V_{1j})^2 \right]}{\sum_{k=1}^2 \left( 1 / \left[ \sum_{j=1}^5 (X_{1j} - V_{kj})^2 \right] \right)} \\ &= \frac{1 / ((25 - 25.54795)^2) + \dots + ((0 - 0.800884)^2)}{[1 / ((25 - 25.54795)^2) + \dots + ((0 - 0.800884)^2)] + [1 / ((25 - 23.8922)^2) + \dots + ((0 - 0.621246)^2)]} \\ &= 0.487 \end{aligned}$$

$$\mu_{ik} = \begin{bmatrix} 0.487 & 0.530 & 0.571 & 0.427 & 0.672 & 0.439 & 0.420 & 0.416 & 0.572 & 0.401 \\ 0.513 & 0.470 & 0.429 & 0.573 & 0.328 & 0.561 & 0.580 & 0.584 & 0.428 & 0.599 \end{bmatrix}$$

7. Mengecek kondisi berhenti:

- a. Jika  $(|P_t - P_{t-1}| < \xi)$  atau  $(t > \text{MaxIter})$  maka berhenti;
- b. Jika tidak,  $t = t + 1$ , ulangi langkah ke-4

### 3.5 Possibilistic C-Means (PCM)

*Possibilistic C-Means* (PCM) merupakan metode pengelompokan yang diperkenalkan oleh Krishnapuram dan Keller pada tahun 1993 untuk mengatasi data yang terdapat *outlier*. Berbeda dengan algoritma FCM, nilai keanggotaan yang dihasilkan oleh algoritma PCM dapat diartikan sebagai “derajat kepemilikan atau kesesuaian atau kekhasan”. Derajat kekhasan ditentukan untuk membangun bentuk asli yang mengkarakterisasi sub kategori suatu data, dengan mempertimbangkan ciri khas anggota kategori (Ozdemir & Kaya, 2019).

Daya Tarik utama *Possibilistic C-Means* (PCM) adalah dimana nilai-nilai keanggotaan merupakan ukuran kekuatan mutlak keanggotaan suatu kelas (Ganbold & Chasia, 2016). Kelebihan pada metode *Possibilistic C-Means* (PCM) adalah memungkinkan pengelompokan dari data yang terdapat *outlier*. Sedangkan kekurangan dari metode *Possibilistic C-Means* (PCM) adalah PCM sangat sensitif terhadap inisialisasi yang baik dan algoritma metode tersebut dapat menghasilkan cluster yang serupa karena kolom dan baris dari matriks yang tipikal tidak berkaitan satu sama lain, jika inisialisasi dari setiap baris tidak cukup jelas maka dapat menyebabkan *cluster* serupa (Grover, 2014). Adapun fungsi objektifnya, yaitu:

$$P_t = \sum_{i=1}^n \sum_{k=1}^c t_{ik}^w d_{ik}^2 + \sum_{i=1}^n \eta_i \sum_{k=1}^c (1 - t_{ik})^w \quad (3.8)$$

Dimana:

$$t_{ik} = \left[ 1 + \left( \frac{d_{ik}}{\eta_i} \right)^{\frac{1}{w-1}} \right]^{-1} \quad (3.9)$$

$$\eta_i = K \frac{\sum_{i=1}^n \mu_{ik}^w d_{ik}^2}{\sum_{i=1}^n \mu_{ik}^w}; K > 0 \quad (3.10)$$

Keterangan :

$t_{ik}$  = Matriks kekhasan absolut

$d_{ik}$  = Jarak

### 3.6 Fuzzy Possibilistic C-Means

*Fuzzy Possibilistics C-Means* (FPCM) merupakan suatu algoritma *clustering* pengembangan dari *Fuzzy C-Means* (FCM) dan *Possibilistic C-Means* (PCM). Pada algoritma *Fuzzy C-Means* (FCM), nilai  $\mu_{ik}$  dipengaruhi oleh  $X_{ij}$  dan semua pusat

cluster. Sementara pada algoritma *Possibilistic C-Means* (PCM), nilai  $t_{ik}$  dipengaruhi oleh  $X_{ij}$ , pusat cluster ke  $k$  serta  $\gamma_k$ . Sehingga, algoritma *Fuzzy Possibilistics C-Means* (FPCM) menghasilkan pusat *cluster* juga nilai derajat keanggotaan, dan nilai kekhasan untuk menentukan setiap titik data pada suatu *cluster* tertentu (Kusumadewi, dkk., 2006). Adapun langkah-langkah metode *Fuzzy Possibilistic C-Means* (FPCM) adalah sebagai berikut:

1. Memasukkan data untuk dikelompokkan  $X$ , berupa matriks berukuran  $n \times m$  yang memiliki fungsi untuk menentukan jumlah data dan atribut setiap data yang akan digunakan.

$n$  = jumlah sampel data

$m$  = atribut setiap data

$X_{ij}$  = data sampel ke- $i$  ( $i=1,2,\dots,n$ ), atribut ke- $j$  ( $j=1,2,\dots,m$ )

$$X_{ij} = \begin{bmatrix} X_{11} & \dots & X_{1m} \\ \vdots & \ddots & \vdots \\ X_{n1} & \dots & X_{nm} \end{bmatrix} \quad (3.11)$$

2. Menentukan:

- a. Jumlah *cluster* =  $c$  ( $\geq 2$ );
- b. Bobot pangkat =  $w$  ( $> 1$ ) dan  $\eta$  ( $> 1$ );
- c. Maksimum iterasi = MaxIter
- d. Error terkecil yang diharapkan =  $\xi$
- e. Fungsi objektif awal =  $P_o = 0$
- f. Iterasi awal =  $t = 1$

3. Memanggil matriks kekhasan relatif ( $\mu_{ik}$ ) dan pusat cluster ( $V_{kj}$ ) pada algoritma *Fuzzy C-Means* (FCM) untuk menghitung matriks kekhasan absolut dimana  $i = 1, 2, \dots, n$ ;  $k = 1, 2, \dots, c$ ; dan  $j = 1, 2, \dots, m$ ; dengan persamaan sebagai berikut

$$t_{ik} = \begin{bmatrix} t_{11} & \dots & t_{1c} \\ \vdots & \ddots & \vdots \\ t_{n1} & \dots & t_{nc} \end{bmatrix} \quad (3.12)$$

Dimana elemen matriknya yaitu:

$$t_{ik} = \frac{[\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{\frac{-1}{\eta-1}}}{\sum_{i=1}^n [\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{\frac{-1}{\eta-1}}} \quad (3.13)$$

4. Memperbaiki pusat *cluster* dimana  $i = 1, 2, \dots, n$ ;  $k = 1, 2, \dots, c$ ; dan  $j = 1, 2, \dots, m$ ; dengan persamaan sebagai berikut

$$V_{kj} = \frac{\sum_{i=1}^n (\mu_{ik}^w + t_{ik}^\eta) X_{ij}}{\sum_{i=1}^n (\mu_{ik}^w + t_{ik}^\eta)} \quad (3.14)$$

5. Menghitung fungsi objektif pada iterasi ke  $t$  dengan persamaan sebagai berikut

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left( \left[ \sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik}^w + t_{ik}^\eta) \right) \quad (3.15)$$

6. Memperbaiki matriks kekhasan relatif ( $\mu_{ik}$ ) dimana  $i = 1, 2, \dots, n$ ;  $k = 1, 2, \dots, c$ ; dan  $j = 1, 2, \dots, m$ ; dengan persamaan sebagai berikut

$$\mu_{ik} = \frac{[\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}}{\sum_{i=1}^c [\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}} \quad (3.16)$$

7. Memperbaiki matriks kekhasan absolut ( $t_{ik}$ ) dengan persamaan sebagai berikut

$$t_{ik} = \frac{[\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{\eta-1}}}{\sum_{i=1}^n [\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{\eta-1}}} \quad (3.17)$$

8. Mengecek kondisi berhenti:

- c. Jika  $(|P_t - P_{t-1}| < \xi)$  atau  $(t > \text{MaxIter})$  maka berhenti;
- d. Jika tidak,  $t = t + 1$ , ulangi langkah ke-4

Contoh studi kasus sama seperti metode *Fuzzy C-Means*. Adapun langkah-langkah metode *Fuzzy Possibilistic C-Means* (FPCM) adalah sebagai berikut:

1. Memasukkan data berupa matriks berukuran  $n \times m$  yang memiliki fungsi untuk menentukan jumlah data dan jumlah atribut atau variabel setiap data yang akan digunakan.

$$X_{ij} = \begin{bmatrix} 25 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 24 & \dots & 0 \end{bmatrix}$$

2. Menentukan:

- a. Jumlah *cluster* = 2
- b. Bobot pangkat ( $w = 2$ ) dan ( $\eta = 2$ );
- c. Maksimum iterasi = 100



- d. Error terkecil yang diharapkan = 0.00001  
e. Fungsi objektif awal = 0  
f. Iterasi awal = 1
3. Memanggil matriks kekhasan relatif ( $\mu_{ik}$ ) dan pusat cluster ( $V_{kj}$ ) pada algoritma *Fuzzy C-Means* (FCM) untuk menghitung matriks kekhasan absolut

$$t_{11} = \frac{[\sum_{j=1}^5 (X_{1j} - V_{1j})^2]^{-1}}{\sum_{i=1}^{10} [\sum_{j=1}^5 (X_{1j} - V_{kj})^2]^{-1}} = \frac{1/[\sum_{j=1}^5 (X_{1j} - V_{1j})^2]}{\sum_{i=1}^{10} (1/[\sum_{j=1}^5 (X_{1j} - V_{kj})^2])}$$

$$= \frac{1/((25-151.195)^2) + \dots + ((0-4.750)^2)}{[1/((25-151.195)^2) + \dots + ((0-4.750)^2)] + [1/((25-137.769)^2) + \dots + ((0-2859)^2)]}$$

$$= 0.055$$

$$t_{ik} = \begin{bmatrix} 0.055 & 0.009 & 0.044 & 0.046 & 0.416 & 0.060 & 0.038 & 0.042 & 0.180 & 0.110 \\ 0.068 & 0.009 & 0.038 & 0.073 & 0.238 & 0.090 & 0.062 & 0.070 & 0.158 & 0.193 \end{bmatrix}$$

4. Memperbaiki pusat *cluster*

$$V_{11} = \frac{\sum_{i=1}^{10} (\mu_{i1}^2 + t_{i1}^2) X_{i1}}{\sum_{i=1}^{10} (\mu_{i1}^2 + t_{i1}^2)} = \frac{((0.487+0.055)^2 25) + \dots + ((0.401+0.110)^2 24)}{(0.487+0.055)^2 + \dots + (0.401+0.110)^2}$$

$$= 83.003$$

$$V_{kj} = \begin{bmatrix} 83.003 & 75.547 & 450.861 & 312.553 & 2.484 \\ 93.591 & 92.970 & 634.012 & 308.185 & 1.406 \end{bmatrix}$$

5. Menghitung fungsi objektif pada iterasi ke = 1

$$P_1 = \sum_{i=1}^{10} \sum_{k=1}^2 \left( \left[ \sum_{j=1}^5 (X_{ij} - V_{kj})^2 \right] (\mu_{ik}^2 + t_{ik}^2) \right)$$

$$= ([154233.456](0.241)) + \dots + ([305997.779](0.396))$$

$$= 1247474.613$$

6. Memperbaiki matriks kekhasan relatif ( $\mu_{ik}$ )

$$\mu_{11} = \frac{[\sum_{j=1}^5 (X_{1j} - V_{1j})^2]^{-1}}{\sum_{k=1}^2 [\sum_{j=1}^5 (X_{1j} - V_{kj})^2]^{-1}} = \frac{1/[\sum_{j=1}^5 (X_{1j} - V_{1j})^2]}{\sum_{k=1}^2 (1/[\sum_{j=1}^5 (X_{1j} - V_{kj})^2])}$$

$$= \frac{1/((25-151.195)^2) + \dots + ((0-4.750)^2)}{[1/((25-151.195)^2) + \dots + ((0-4.750)^2)] + [1/((25-137.769)^2) + \dots + ((0-2859)^2)]}$$

$$= 0.663$$

$$\mu_{ik} = \begin{bmatrix} 0.663 & 0.682 & 0.672 & 0.659 & 0.666 & 0.662 & 0.659 & 0.659 & 0.666 & 0.661 \\ 0.337 & 0.318 & 0.328 & 0.341 & 0.334 & 0.338 & 0.341 & 0.341 & 0.334 & 0.339 \end{bmatrix}$$

7. Memperbaiki matriks kekhasan absolut ( $t_{ik}$ )

$$\begin{aligned}
t_{11} &= \frac{[\sum_{j=1}^5 (X_{1j} - V_{1j})^2]^{-1}}{\sum_{i=1}^{10} [\sum_{j=1}^5 (X_{ij} - V_{2j})^2]^{-1}} = \frac{1/[\sum_{j=1}^5 (X_{1j} - V_{1j})^2]}{\sum_{i=1}^{10} (1/[\sum_{j=1}^5 (X_{ij} - V_{2j})^2])} \\
&= \frac{1/((25 - 151.195)^2) + \dots + ((0 - 4.750)^2)}{[1/((25 - 151.195)^2) + \dots + ((0 - 4.750)^2)] + \dots + [1/((25 - 137.769)^2) + \dots + ((0 - 2859)^2)]} \\
&= 0.097 \\
t_{ik} &= \begin{bmatrix} 0.097 & 0.124 & 0.111 & 0.091 & 0.102 & 0.094 & 0.091 & 0.091 & 0.103 & 0.095 \\ 0.098 & 0.115 & 0.107 & 0.094 & 0.102 & 0.095 & 0.094 & 0.094 & 0.102 & 0.097 \end{bmatrix}
\end{aligned}$$

8. Mengecek kondisi berhenti:

- Jika  $(|P_t - P_{t-1}| < \xi)$  atau  $(t > \text{MaxIter})$  maka berhenti;
- Jika tidak,  $t = t + 1$ , ulangi langkah ke-4

### 3.7 Validitas Cluster

Validitas cluster merupakan suatu cara untuk mengevaluasi kualitas dari cluster dan menentukan apakah cluster tersebut dapat menjelaskan keseluruhan data atau tidak. Validasi cluster dapat dilakukan menggunakan *Cluster Validity Index* (CVI). *Cluster Validity Index* (CVI) merupakan ukuran validitas yang digunakan untuk menemukan jumlah cluster optimal yang sepenuhnya dapat menjelaskan struktur data dan mengukur derajat kekompakan dan separasi struktur data pada seluruh cluster. *Cluster Validity Index* (CVI) meliputi metode *Partition Entropy Index* (PEI) dan *Partition Coefficient Index* (PCI) (Mashfuufah & Istiawan, 2018). Indeks validitas yang pertama kali berhubungan dengan algoritma *Fuzzy C-Means* (FCM) adalah *Partition Entropy Index* (PEI). Metode *Partition Entropy Index* (PEI) didefinisikan dengan persamaan:

$$PEI = -\frac{1}{n} \sum_{k=1}^c \sum_{i=1}^n \mu_{ik} \ln(\mu_{ik}) \quad (3.18)$$

Metode *Partition Coefficient Index* (PCI) didefinisikan dengan persamaan:

$$PCI = \frac{1}{n} \sum_{k=1}^c \sum_{i=1}^n \mu_{ik}^2 \quad (3.19)$$

Dimana:

$n$  = banyak objek penelitian

$c$  = banyak cluster

$\mu_{ik}$  = nilai keanggotaan objek ke- $i$  pada cluster ke- $j$

*Partition Entropy* merupakan evaluasi yang mengukur tingkat keaburan dari partisi cluster, kluster optimal didapatkan jika nilai yang diperoleh semakin

kecil (mendekati angka 0). Sedangkan *Partition Coefficient* hanya melakukan evaluasi terhadap nilai membership degree, kluster optimal didapatkan jika nilai yang diperoleh semakin besar (mendekati angka 1) (David, dkk., 2020).

*Modified Partition Coefficient* (MPC) merupakan indeks validitas perbaikan dari metode *Partition Coefficient* (PC) karena metode tersebut cenderung mengalami perubahan terhadap nilai  $c$  (banyak cluster). *Modified Partition Coefficient* (MPC) didefinisikan dengan persamaan:

$$\begin{aligned} \text{MPC} &= 1 - \frac{c}{c-1} (1 - \text{PCI}) \\ &= 1 - \frac{c}{c-1} \left(1 - \left(\frac{1}{n} \sum_{k=1}^c \sum_{i=1}^n \mu_{ik}^2\right)\right) \end{aligned} \quad (3.20)$$

Dimana:

$n$  = banyak objek penelitian

$c$  = banyak *cluster*

$\mu_{ij}$  = nilai keanggotaan objek ke- $i$  pada cluster ke- $j$

Kluster optimal didapatkan jika nilai *Modified Partition Coefficient* (MPC), yang diperoleh semakin besar (mendekati angka 1). Semakin kecil atau semakin dekat nilai MPC ke 0 maka semakin kabur keakuratannya (Amirah, dkk., 2017).

Contoh studi kasus sama seperti metode *Fuzzy C-Means*.

$$\begin{aligned} \text{PEI} &= -\frac{1}{10} \sum_{k=1}^2 \sum_{i=1}^{10} \mu_{ik} \ln(\mu_{ik}) \\ &= -\frac{1}{10} \left( (0.453 \ln(0.453)) + \dots + (0.657 \ln(0.657)) \right) \\ &= 0.515258 \end{aligned}$$

$$\begin{aligned} \text{PCI} &= \frac{1}{10} \sum_{k=1}^2 \sum_{i=1}^{10} \mu_{ik}^2 \\ &= \frac{1}{10} \left( (0.453^2) + \dots + (0.657^2) \right) \\ &= 0.556658 \end{aligned}$$

$$\text{MPC} = 1 - \frac{2}{2-1} (1 - 0.556658) = 0.113316$$

### 3.8 Kesehatan Masyarakat

Kesehatan masyarakat merupakan seni dan ilmu melindungi, memelihara, dan meningkatkan kesehatan masyarakat melalui usaha-usaha pengorganisasian

masyarakat. Kesehatan masyarakat juga diartikan sebagai aplikasi keterpaduan antara sanitasi, ilmu kedokteran, dan ilmu sosial dalam mencegah penyakit yang terjadi di masyarakat. Ada 24 indikator kesehatan yaitu prevalensi balita gizi buruk dan kurang, prevalensi balita sangat pendek dan pendek, prevalensi balita sangat kurus dan kurus, prevalensi balita gemuk, prevalensi diare, prevalensi pneumonia, prevalensi hipertensi, prevalensi gangguan mental, prevalensi asma, prevalensi penyakit gigi dan mulut, prevalensi disabilitas, prevalensi cedera, prevalensi penyakit sendi, prevalensi ISPA, proporsi perilaku cuci tangan, proporsi merokok tiap hari, akses air bersih, akses sanitasi, cakupan persalinan oleh nakes, cakupan pemeriksaan neonatal-1, cakupan imunisasi lengkap, cakupan penimbangan balita, ratio Dokter/Puskesmas, dan ratio bidan/desa. Faktor-faktor yang mempengaruhi derajat kesehatan masyarakat adalah lingkungan, perilaku, pelayanan kesehatan, dan keturunan. (KementerianKesehatan, 2016)

### **3.9 Fasilitas Pelayanan Kesehatan**

Fasilitas pelayanan kesehatan merupakan istilah yang merujuk pada sarana dan prasarana yang dinyatakan dalam bentuk pelayanan yang disenggarakan oleh pemerintah, pemerintah daerah, atau swasta kepada masyarakat dengan tujuan untuk menjaga dan meningkatkan kesehatan. Menurut Undang-Undang Nomor 36 Tahun 2009 tentang Kesehatan, fasilitas pelayanan kesehatan adalah suatu alat dan/atau tempat yang digunakan untuk menyelenggarakan upaya pelayanan kesehatan, baik promotif, preventif, kuratif, maupun rehabilitatif yang dilakukan oleh pemerintah, pemerintah daerah, dan/atau masyarakat. Derajat kesehatan masyarakat suatu negara dipengaruhi oleh keberadaan fasilitas pelayanan kesehatan.

Fasilitas pelayanan kesehatan terdiri dari Fasilitas Kesehatan Tingkat Pertama (FKTP) dan Fasilitas Kesehatan Tingkat Rujukan Lanjut (FKTRL). Fasilitas Kesehatan Tingkat Pertama meliputi Puskesmas, klinik pratama, dan praktik dokter/dokter gigi perseorangan, sedangkan Fasilitas Kesehatan Tingkat Rujukan Lanjut meliputi rumah sakit umum dan rumah sakit khusus (KementerianKesehatan, 2020).

### **3.9.1 Pusat Kesehatan Masyarakat (Puskesmas)**

Dalam Peraturan Menteri Kesehatan Nomor 75 Tahun 2014 tentang Puskesmas, Puskesmas adalah fasilitas pelayanan kesehatan yang menyelenggarakan upaya kesehatan masyarakat dan upaya kesehatan perseorangan tingkat pertama, dengan lebih mengutamakan upaya promotif dan preventif, untuk mencapai derajat kesehatan masyarakat yang setinggi-tingginya di wilayah kerjanya (Kementerian Kesehatan, 2020). Upaya kesehatan tersebut diselenggarakan dengan memfokuskan kepada pelayanan untuk masyarakat luas untuk mencapai derajat kesehatan yang optimal, tanpa mengabaikan kualitas pelayanan kepada individu perorangan. Puskesmas dipimpin oleh seorang kepala puskesmas yang bertanggung jawab kepada Dinas Kesehatan Kabupaten/Kota.

### **3.9.2 Rumah Sakit**

Dalam rangka meningkatkan derajat kesehatan masyarakat selain dilakukan upaya promotif dan preventif, diperlukan juga upaya kuratif dan rehabilitatif. Upaya kesehatan yang bersifat kuratif dan rehabilitatif dapat diperoleh melalui rumah sakit yang juga berfungsi sebagai penyedia pelayanan kesehatan rujukan. Berdasarkan Peraturan Menteri Kesehatan Nomor 56 Tahun 2014 tentang Klasifikasi dan Perizinan Rumah Sakit, pengelompokan rumah sakit berdasarkan penyelenggaraan, yaitu rumah sakit pemerintah, rumah sakit pemerintah daerah, dan rumah sakit swasta.

Rumah sakit pemerintah adalah rumah sakit yang pelaksanaan teknisnya dari instansi pemerintah seperti Kementerian Kesehatan, Kepolisian, Tentara Nasional Indonesia (TNI), BUMN, dan Kementerian Lainnya. Rumah sakit daerah adalah rumah sakit yang pelaksanaan teknisnya dari daerah seperti pemerintah provinsi, pemerintah kabupaten dan pemerintah kota. Sedangkan rumah sakit swasta adalah badan hukum yang kegiatan usahanya hanya bergerak di bidang perumahsakit, pengecualian bagi rumah sakit publik yang diselenggarakan oleh badan hukum (Kementerian Kesehatan, 2020).

### **3.10 Penyakit Menular**

Penyakit menular meliputi penyakit menular langsung, penyakit yang dapat dikendalikan dengan imunisasi dan penyakit yang ditularkan melalui binatang. Penyakit menular langsung yaitu, Tuberkulosis, HIV/AIDS, Pneumonia, Hepatitis, Diare, dan Kusta. Penyakit yang dapat dikendalikan dengan imunisasi yaitu, Tetanus Neonatrum, Campak, Difteri, Polio, dan Lumpuh Layu Akut (*Acute Flaccid Paralysis* (AFP)). Sedangkan penyakit yang ditularkan melalui binatang yaitu, Demam Berdarah *Dengue* (DBD), Chikungunya, Filariasis, Malaria, Rabies, dan Leptospirosis (KementerianKesehatan, 2020).

#### **3.10.1 Tuberkulosis**

Tuberkulosis adalah suatu penyakit menular yang disebabkan oleh kuman atau bakteri *Mycobacterium tuberculosis* (TB). Kuman TB pada umumnya menyerang organ paru-paru, tetapi dapat juga menyerang organ tubuh lainnya seperti kelenjar getah bening, kulit saluran pencernaan (usus), selaput otak, dan lainnya. Sumber penularan adalah pasien dengan TB BTA positif melalui batuk atau bersin, pasien menyebarkan kuman ke udara dalam bentuk percikan dahak (*droplet nuclei*) (KementerianKesehatan, 2011).

#### **3.10.2 Diare**

Diare merupakan gejala infeksi pada saluran pencernaan yang dapat disebabkan oleh organisme seperti virus, bakteri, dan parasit. Organisme tersebut menginfeksi saluran pencernaan manusia melalui makanan dan minuman yang telah tercampur atau tercemar oleh organisme tersebut (*food borne disease*). Menurut WHO, diare adalah kejadian buang air besar dengan konsistensi lebih cair dari biasanya dengan 2 frekuensi atau lebih selama 1 hari atau lebih.

Ada tiga jenis diare berdasarkan gejala klinisnya, yaitu diare cair akut, diare akut berdarah, dan diare persisten. Diare cair akut adalah diare yang dimana penderitanya akan kehilangan cairan tubuh dalam jumlah yang besar sehingga mampu menyebabkan dehidrasi dalam waktu yang cepat. Diare akut berdarah sering juga disebut dengan disentri, diare ini ditandai dengan adanya darah dalam tinja yang disebabkan akibat kerusakan usus. Penderita yang menderita diare

berdarah akan menyebabkan kehilangan zat gizi yang berdampak pada penurunan status gizi. Diare persisten adalah diare yang berlangsung lebih dari 14 hari, diare jenis ini sering terjadi pada penderita dengan status gizi rendah, AIDS, dan anak dalam kondisi infeksi (KementerianKesehatan, 2014).

### 3.10.3 Pneumonia

Pneumonia adalah penyakit yang menginfeksi jaringan paru-paru (alveoli) yang disebabkan oleh mikroorganisme seperti virus, jamur dan bakteri. Pengendalian pneumonia pada balita menjadi program prioritas pengendalian pneumonia sampai saat ini. Pneumonia pada balita ditandai dengan batuk dan atau tanda kesulitan bernapas yaitu adanya nafas cepat, kadang disertai tarikan dinding dada bagian bawah kedalam (TDDK) (KementerianKesehatan, 2020).

Pneumonia merupakan salah satu masalah kesehatan yang menjadi penyumbang terbesar penyebab kematian anak usia dibawah lima tahun (balita). Pneumonia dapat disebabkan karena infeksi berbagai, virus, jamur dan bakteri. Penyebab utama virus adalah *Respiratory Syncytial Virus* (RSV) yang mencakup 15-40% kasus diikuti virus influenza A dan B, parainfluenza, human metapneumovirus dan adenovirus. Sedangkan penyebab utama bakteri adalah *Streptococcus pneumoniae/pneumococcus* (30-50% kasus), *Hemophilus influenzae type b* (10-30% kasus), *Staphylococcus aureus*, *Klebsiella pneumoniae*, dan lainnya. Pneumonia pada neonatus banyak disebabkan oleh bakteri Gram negatif seperti *Klebsiella spp*, *E coli* di samping bakteri Gram positif seperti *S pneumoniae*, grup b streptokokus dan *S aureus*. Berdasarkan hasil penelitian, 70% penyakit pneumonia disebabkan oleh bakteri. Kematian pada pneumonia berat, terutama disebabkan karena infeksi bakteri (KementerianKesehatan, 2014).



## BAB IV

### METODOLOGI PENELITIAN

#### 3.1 Populasi Penelitian

Populasi yang digunakan dalam penelitian ini yaitu seluruh Kabupaten/Kota di Provinsi Banten. Sampel yang digunakan dalam penelitian ini yaitu seluruh Kabupaten/Kota di Provinsi Banten yang diukur berdasarkan fasilitas pelayanan kesehatan dan penderita penyakit menular pada tahun 2019.

#### 3.2 Tempat dan Waktu Penelitian

Penelitian ini dilakukan pada bulan September 2020 sampai dengan Desember 2020 di Universitas Islam Indonesia Yogyakarta.

#### 3.3 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini adalah jumlah puskesmas, jumlah rumah sakit, jumlah penderita penyakit tuberkulosis, diare dan pneumonia pada tahun 2019. Banyaknya Kabupaten/Kota sesuai dengan jumlah Kabupaten/Kota di Provinsi Banten yaitu 8. Untuk variabel yang digunakan, dijelaskan pada **Tabel 4.1**

**Tabel 4.1** Penjelasan Variabel

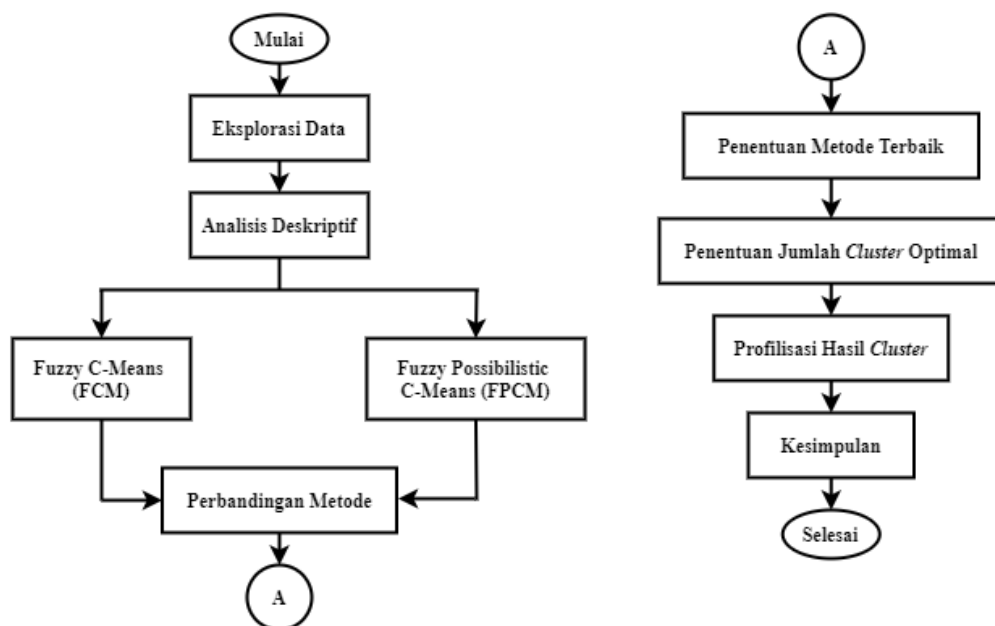
No	Variabel	Definisi Operasional	Satuan/ Skala	Contoh
1	Jumlah Puskesmas	Jumlah fasilitas pelayanan kesehatan yang menyelenggarakan upaya kesehatan masyarakat dan upaya kesehatan perseorangan tingkat pertama	Rasio	Jumlah puskesmas di Kabupaten Lebak tahun 2019 adalah 42
2	Jumlah Rumah Sakit	Jumlah fasilitas pelayanan kesehatan yang menyelenggarakan upaya kesehatan masyarakat dan upaya kesehatan perseorangan tingkat rujukan lanjut	Rasio	Jumlah rumah sakit di Kabupaten Lebak tahun 2019 adalah 2



No	Variabel	Definisi Operasional	Satuan/ Skala	Contoh
3	Jumlah Penderita Tuberkulosis	Jumlah penderita penyakit menular yang disebabkan oleh kuman atau bakteri <i>Mycobacterium tuberculosis</i> (TB)	Rasio	Jumlah penderita penyakit tuberkulosis di Kabupaten Lebak tahun 2019 adalah 2,833 orang
4	Jumlah Penderita Diare	Jumlah penderita penyakit menular pada saluran pencernaan yang disebabkan oleh organisme seperti virus, bakteri, dan parasit	Rasio	Jumlah penderita penyakit diare di Kabupaten Lebak tahun 2019 adalah 31,674 orang
5	Jumlah Penderita Pneumonia	Jumlah penderita penyakit menular yang menginfeksi jaringan paru-paru yang disebabkan oleh mikroorganisme seperti virus, jamur dan bakteri	Rasio	Jumlah penderita penyakit pneumonia di Kabupaten Lebak tahun 2019 adalah 3,092 orang

### 3.4 Tahap Penelitian

Berikut adalah diagram alir metode penelitian:



**Gambar 4.1** Flowchart Tahap Penelitian

**Gambar 4.1** menjelaskan tahapan peneliti yang dilakukan oleh penulis. Berikut tahapannya :

1. Eksplorasi data

Tahap ini merupakan tahap mencari dan pengambilan data

2. Analisis deskriptif

Pada tahap ini gambaran umum akan ditampilkan dengan *median* (nilai tengah), *std. devition* (simpangan baku), *minimum* (nilai terkecil), dan *maximum* (nilai terbesar)

3. Perbandingan metode

Pada tahap ini perbandingan metode dilakukan pada 2 metode yaitu *Fuzzy C Means* (FCM) dan *Fuzzy Possibilistic C Means* (FPCM) berdasarkan jumlah iterasi dan fungsi objektif.

4. Penentuan metode terbaik dan jumlah *cluster* optimal

Penentuan metode terbaik dan jumlah *cluster* optimal didasarkan pada nilai indeks *Partition Entrophy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC).

5. Profilisasi hasil *cluster*

Pada tahap ini profilisasi hasil *cluster* dilakukan untuk menggambarkan karakteristik masing-masing *cluster* dari *cluster* yang telah terbentuk.

6. Kesimpulan

Pada tahap penulis mengambil kesimpulan dari analisis yang telah dilakukan.

## BAB V

### HASIL DAN PEMBAHASAN

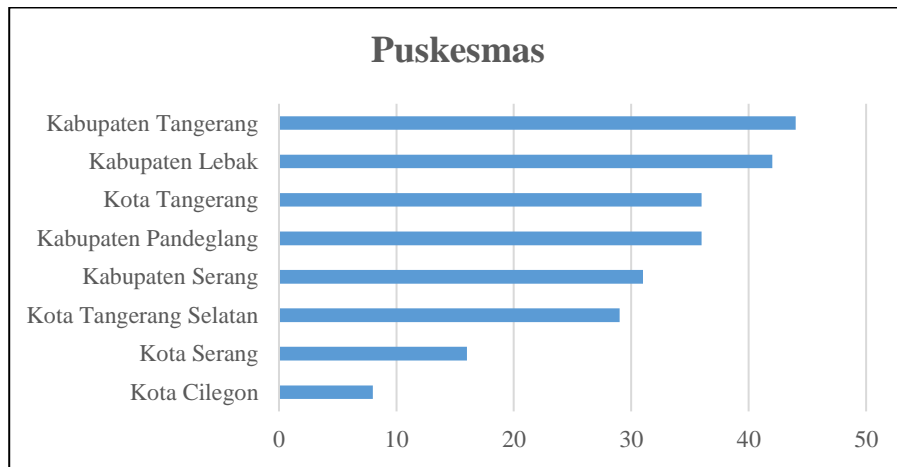
#### 4.1 Analisis Deskriptif

Analisis deskriptif digunakan untuk memberikan gambaran umum terkait data yang di analisis. Adapun gambaran umum yang akan ditampilkan adalah *median* (nilai tengah), *std. devition* (simpangan baku), *minimum* (nilai terendah), dan *maximum* (nilai tertinggi)

**Tabel 5.1** Statistika Deskriptif Kesehatan di Provinsi Banten

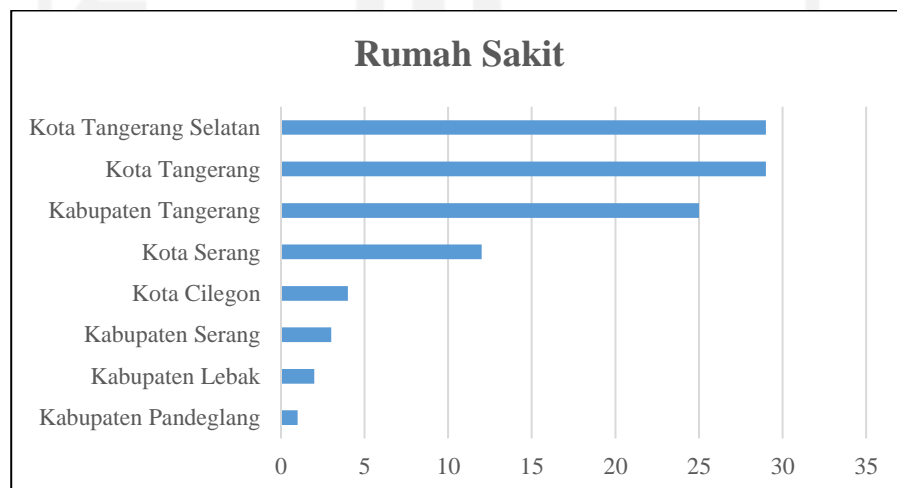
	Puskesmas	Rumah Sakit	Diare (Orang)	Tuberkulosis (Orang)	Pneumonia (Orang)
<i>Median</i>	33.50	8	32,295	2,844	3,360
<i>Std. Devition</i>	12.498	12.552	19,909.75	2,054.395	4,053.394
<i>Minimum</i>	8	1	8,703	1,604	611
<i>Maximum</i>	44	29	70,107	6,737	12,842

Berdasarkan pada **Tabel 5.1** *Median* atau nilai tengah dari jumlah puskesmas yaitu sebesar 33.50, nilai tengah dari jumlah rumah sakit yaitu sebesar 8, nilai tengah dari jumlah penderita penyakit diare yaitu sebesar 32,295 orang, nilai tengah dari jumlah penderita penyakit tuberkulosis yaitu sebesar 2,844 orang, dan nilai tengah dari jumlah penderita penyakit pneumonia yaitu sebesar 3,360 orang. Standar deviasi atau simpangan baku dari variabel puskesmas yaitu sebesar 12.498, variabel rumah sakit yaitu sebesar 12.552, variabel diare yaitu sebesar 19,909.75, variabel tuberkulosis yaitu sebesar 2,054.395, dan variabel pneumonia yaitu sebesar 4,053.394.



**Gambar 5.1** Diagram Batang Jumlah Puskesmas

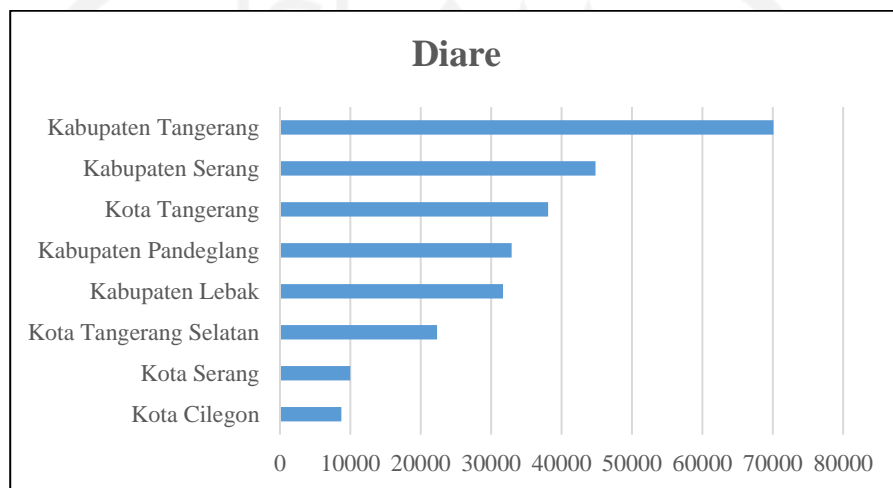
Berdasarkan **Tabel 5.1** dan **Gambar 5.1** nilai minimum dari variabel puskesmas yaitu sebesar 8. Kabupaten/Kota yang memiliki jumlah puskesmas terendah yaitu Kota Cilegon. Kota Cilegon merupakan salah satu Kabupaten/Kota di Provinsi Banten yang memiliki luas wilayah terendah dan merupakan Kabupaten/Kota di Provinsi Banten yang memiliki jumlah penduduk terendah. Sementara nilai maksimum dari variabel puskesmas yaitu sebesar 44. Kabupaten/Kota yang memiliki jumlah puskesmas tertinggi yaitu Kabupaten Tangerang. Kabupaten Tangerang merupakan Kabupaten/Kota di Provinsi Banten yang memiliki jumlah penduduk tertinggi.



**Gambar 5.2** Diagram Batang Jumlah Rumah Sakit

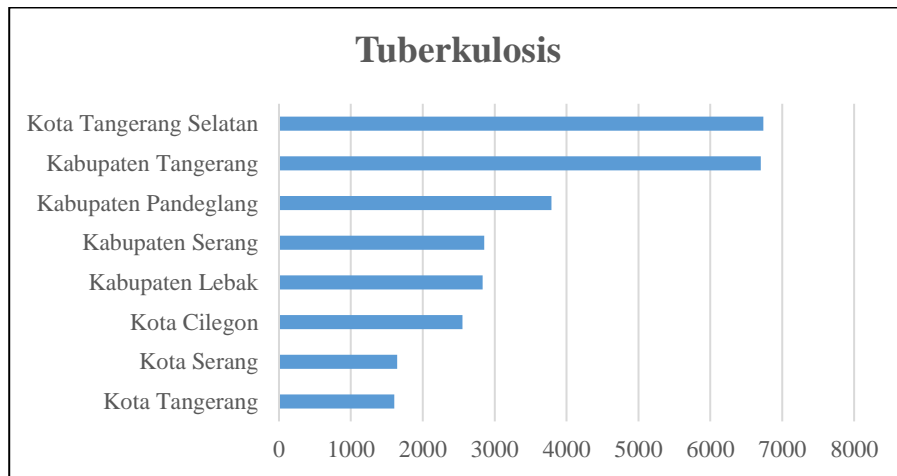
Berdasarkan **Tabel 5.1** dan **Gambar 5.2** nilai minimum dari variabel rumah sakit yaitu sebesar 1. Kabupaten/Kota yang memiliki jumlah rumah sakit terendah

yaitu Kabupaten Pandeglang. Sementara nilai maksimum dari variabel rumah sakit yaitu sebesar 29. Kabupaten/Kota yang memiliki jumlah rumah sakit tertinggi yaitu Kota Tangerang dan Kota Tangerang Selatan. Kota Tangerang dan Kota Tangerang Selatan merupakan Kabupaten/Kota di Provinsi Banten yang paling dekat dengan ibu kota Jakarta. Berdasarkan **Gambar 5.2** terlihat adanya ketimpangan jumlah sakit di Provinsi Banten.



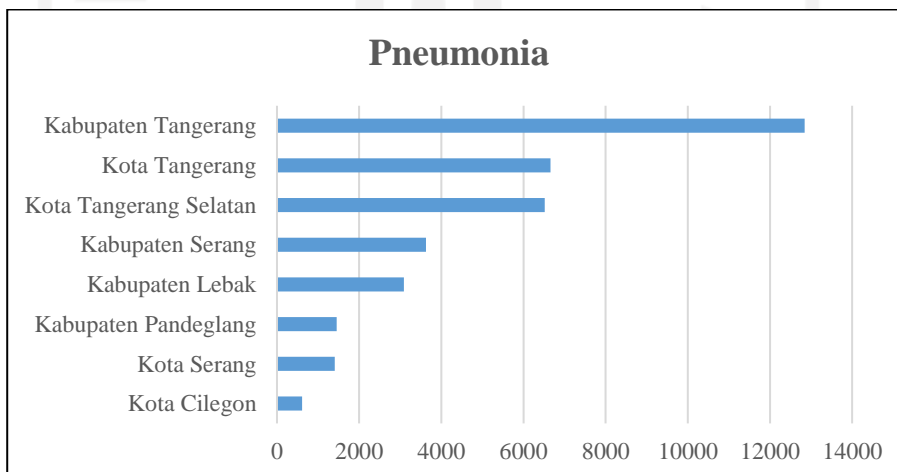
**Gambar 5.3** Diagram Batang Jumlah Penderita Diare

Berdasarkan **Tabel 5.1** dan **Gambar 5.3** nilai minimum dari variabel diare yaitu sebesar 8,703 orang. Kabupaten/Kota yang memiliki jumlah penderita penyakit diare terendah yaitu Kota Cilegon. Sementara nilai maksimum dari variabel diare yaitu sebesar 70,107 orang. Kabupaten/Kota yang memiliki jumlah penderita penyakit diare tertinggi yaitu Kabupaten Tangerang. Faktor utama resiko diare adalah faktor lingkungan seperti sarana air bersih, cara penyajian makanan dan pola hidup bersih dan sehat (PHBS).



**Gambar 5.4** Diagram Batang Jumlah Penderita Tuberkulosis

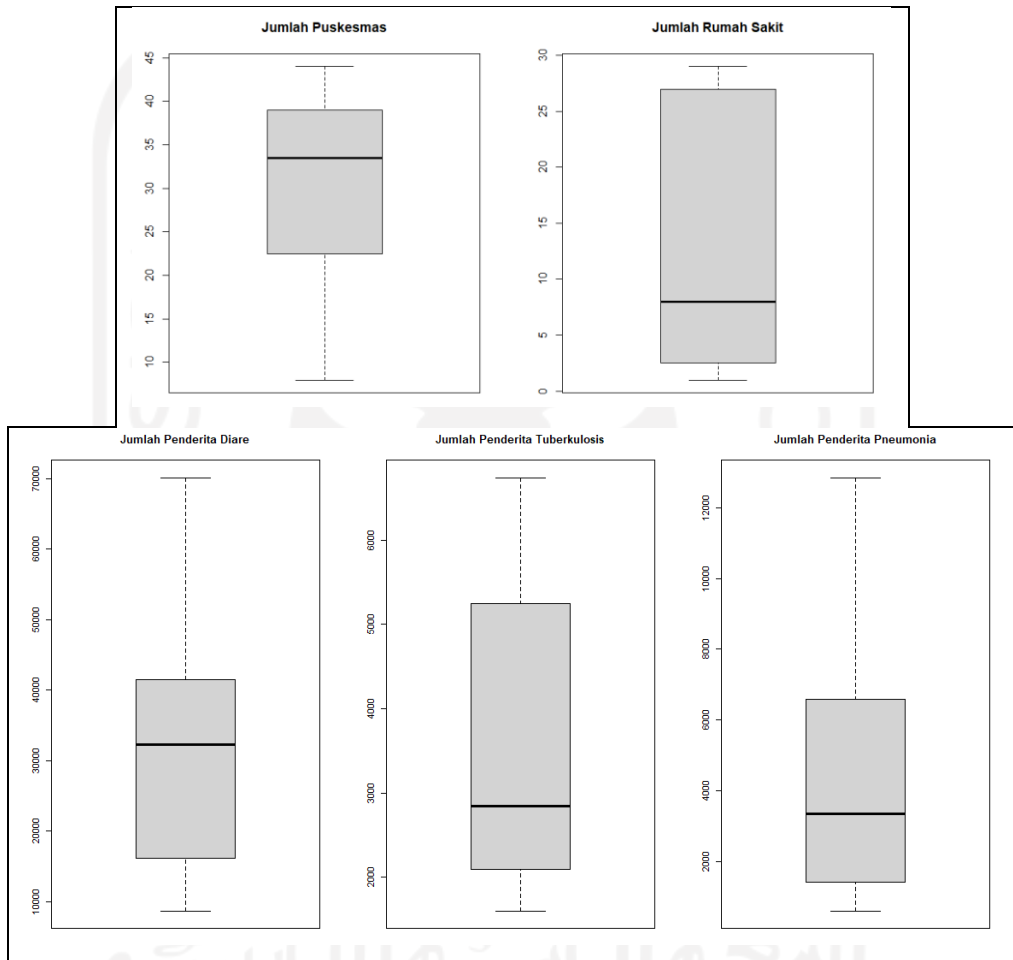
Berdasarkan **Tabel 5.1** dan **Gambar 5.4** nilai minimum dari variabel tuberkulosis yaitu sebesar 1,604 orang. Kabupaten/Kota yang memiliki jumlah penderita penyakit tuberkulosis terendah yaitu Kota Tangerang. Sementara nilai maksimum dari variabel tuberkulosis yaitu sebesar 6,737 orang. Kabupaten/Kota yang memiliki jumlah penderita penyakit tuberkulosis tertinggi yaitu Kota Tangerang Selatan. Faktor penyebab tingginya tuberkulosis adalah besarnya masalah kesehatan lain yang bisa berpengaruh terhadap resiko terjadinya tuberkulosis secara signifikan seperti merokok, HIV dan lainnya yang menyebabkan penurunan daya tahan tubuh.



**Gambar 5.5** Diagram Batang Jumlah Penderita Pneumonia Pada Balita

Berdasarkan **Tabel 5.1** dan **Gambar 5.5** nilai minimum dari variabel pneumonia pada balita yaitu sebesar 611 orang. Kabupaten/Kota yang memiliki

jumlah penderita penyakit pneumonia terendah yaitu Kota Cilegon. Kota Cilegon merupakan Kabupaten/Kota di Provinsi Banten yang memiliki jumlah balita terendah. Sementara nilai maksimum dari variabel pneumonia pada balita yaitu sebesar 12,842 orang. Kabupaten/Kota yang memiliki jumlah penderita penyakit pneumonia tertinggi yaitu Kabupaten Tangerang.



**Gambar 5.6** *Boxplot* Kesehatan di Provinsi Banten

*Boxplot* dapat menunjukkan ada atau tidaknya data *outlier* pada data yang akan dianalisis dan juga dapat menunjukkan persebaran data. Berdasarkan **Gambar 5.6** tidak terdapat data yang *outlier* dari kelima variabel. Dari gambar tersebut dapat dilihat juga jumlah puskesmas banyak yang dibawah nilai tengah atau rata-ratanya, sedangkan jumlah rumah sakit, jumlah penderita diare, jumlah penderita tuberkulosis, dan jumlah penderita pneumonia banyak yang diatas nilai tengah atau rata-ratanya. Sehingga ketiga penyakit menular secara langsung tersebut harus

menjadi perhatian khusus pemerintah daerah. Metode *Fuzzy Possibilistic C-Means* (FPCM) memiliki kelebihan mampu menganalisis data yang masih mengandung *error* seperti data yang *outlier*.

#### 4.2 Perbandingan *Fuzzy C-Means* dan *Fuzzy Possibilistics C-Means*

Pada metode *Fuzzy C-Means* (FCM) nilai pembobot ( $w$ ) yang digunakan adalah 2 dan nilai maksimum iterasi yang digunakan adalah 1000. Pada metode *Fuzzy Possibilistic C-Means* (FPCM), nilai pembobot untuk *Fuzzy C-Means* ( $w$ ) adalah 2, nilai pembobot untuk *Possibilistic C-Means* ( $\eta$ ) adalah 2 dan nilai maksimum iterasi adalah 1000. Perbandingan metode *Fuzzy C-Means* (FCM) dan *Fuzzy Possibilistic C-Means* (FPCM) dapat dilakukan dalam berbagai cara, salah satunya yaitu membandingkan jumlah iterasi dan fungsi objektifnya. Berikut adalah hasil perhitungan fungsi objektif pada iterasi ke = 1 secara manual pada *cluster 2* metode *Fuzzy C-Means* (FCM) sebagai berikut

$$\begin{aligned}
 P_t &= \sum_{i=1}^n \sum_{k=1}^c \left( \left[ \sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \\
 P_1 &= \sum_{i=1}^8 \sum_{k=1}^2 \left( \left[ \sum_{j=1}^5 (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^2 \right) \\
 &= [(X_{11} - V_{11})^2 + \dots + (X_{15} - V_{15})^2] (\mu_{11})^2 + \dots + [(X_{81} - V_{21})^2 + \dots + \\
 &\quad (X_{85} - V_{25})^2] (\mu_{82})^2 \\
 &= [(42 - (-0.244))^2 + \dots + (3092 - (-0.536))^2] (0.835)^2 + \dots + \\
 &\quad [(29 - 0.498)^2 + \dots + (6517 - 1.091)^2] (0.768)^2 \\
 &= 19.322
 \end{aligned}$$

Menghitung fungsi objektif pada iterasi ke = 1 secara manual pada *cluster 2* metode *Fuzzy Possibilistic C-Means* (FPCM) sebagai berikut

$$\begin{aligned}
 P_t &= \sum_{i=1}^n \sum_{k=1}^c \left( \left[ \sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik}^w + t_{ik}^\eta) \right) \\
 P_1 &= \sum_{i=1}^8 \sum_{k=1}^2 \left( \left[ \sum_{j=1}^5 (X_{ij} - V_{kj})^2 \right] (\mu_{ik}^2 + t_{ik}^2) \right) \\
 &= [(X_{11} - V_{11})^2 + \dots + (X_{15} - V_{15})^2] (\mu_{11}^2 + t_{11}^2) + \dots + [(X_{81} - V_{21})^2 + \\
 &\quad \dots + (X_{85} - V_{25})^2] (\mu_{82}^2 + t_{82}^2) \\
 &= [(42 - (-0.010))^2 + \dots + (3092 - (-0.490))^2] (0.835)^2 + (0.843)^2 + \\
 &\quad \dots + [(29 - 0.339)^2 + \dots + (6517 - 0.169)^2] (0.768)^2 + (0.315)^2 \\
 &= 30.871
 \end{aligned}$$



Jika ( $|P_t - P_{t-1}| < \xi$ ) atau ( $t > \text{MaxIter}$ ) yang artinya jika fungsi objektif pada iterasi sekarang dengan iterasi sebelumnya memiliki selisih lebih kecil dari nilai error yang dikehendaki atau iterasi sudah mencapai maksimum iterasi yang ditentukan oleh peneliti, maka berhenti. Jika tidak, ulangi langkah dari menghitung pusat *cluster* ( $V_{kj}$ ).

Berikut adalah hasil perbandingan kedua metode berdasarkan jumlah iterasi dan fungsi objektifnya:

**Tabel 5.2** Jumlah Iterasi dan Fungsi Objektif

Jumlah <i>Cluster</i>	Jumlah Iterasi		Fungsi Objektif	
	FCM	FPCM	FCM	FPCM
<i>Cluster 2</i>	48	25	14.88332	22.18462
<i>Cluster 3</i>	56	55	7.046465	22.69952
<i>Cluster 4</i>	38	55	3.510449	14.79498
<i>Cluster 5</i>	25	44	1.301461	4.617783
<i>Cluster 6</i>	64	20	0.7187028	3.420215
<i>Cluster 7</i>	46	17	0.4474671	2.224091

Berdasarkan **Tabel 5.2** jumlah iterasi yang dilakukan oleh metode *Fuzzy C-Means* (FCM) untuk meminimumkan fungsi objektif adalah 48, 56, 38, 25, 64, dan 46. Sementara pada metode *Fuzzy Possibilistic C-Means* (FPCM) adalah 25, 55, 55, 44, 20, dan 17. Berdasarkan jumlah iterasi dapat disimpulkan bahwa metode *Fuzzy Possibilistic C-Means* (FPCM) memiliki jumlah iterasi yang paling sedikit pada hampir keseluruhan *cluster* untuk meminimumkan fungsi objektif, sehingga metode *Fuzzy Possibilistic C-Means* (FPCM) memiliki waktu komputasi lebih cepat dibandingkan metode *Fuzzy C-Means* (FCM). Fungsi objektif pada metode *Fuzzy C-Means* (FCM) secara berturut-turut adalah 14.88332, 7.046465, 3.510449, 1.301461, 0.7187028, dan 0.4474671. Sementara pada metode *Fuzzy Possibilistic C-Means* (FPCM) adalah 22.18462, 22.69952, 14.79498, 4.617783, 3.420215, dan 2.224091. Berdasarkan fungsi objektif kedua metode tidak bisa dibandingkan karena perhitungan fungsi objektif yang berbeda pada setiap metode.

Hasil  $\mu_{ik}$  (matriks partisi) berpengaruh terhadap hasil fungsi objektif, jika hasil  $\mu_{ik}$  kecil maka hasil fungsi objektif juga akan kecil. Hasil fungsi objektif akan memberikan *clustering* yang optimum.

### 4.3 Penentuan Metode Terbaik

Penentuan metode terbaik dapat dilihat berdasarkan nilai indeks validitas *cluster* yaitu *Partition Entrophy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC). Berikut adalah Perhitungan manual indeks *Partition Entrophy* (PE) dijelaskan dengan persamaan dibawah ini

$$\begin{aligned} PE &= -\frac{1}{n} \sum_{k=1}^c \sum_{i=1}^n \mu_{ik} \ln(\mu_{ik}) = -\frac{1}{8} \sum_{k=1}^2 \sum_{i=1}^8 \mu_{ik} \ln(\mu_{ik}) \\ &= -\frac{1}{8} (\mu_{11} \ln(\mu_{11}) + \dots + \mu_{82} \ln(\mu_{82})) \\ &= -\frac{1}{8} ((0.835) \ln(0.835) + \dots + (0.768) \ln(0.768)) \\ &= 0.455311 \end{aligned}$$

Berikut adalah hasil indeks validitas *cluster* yaitu *Partition Entrophy* (PE) pada metode *Fuzzy C Means* (FCM) dan *Fuzzy Possibilistics C Means* (FPCM).

**Tabel 5.3** Perbandingan *Partition Entrophy*

Jumlah Cluster	<i>Partition Entrophy</i> (PE)	
	FCM	FPCM
Cluster 2	0.455311	0.518594
Cluster 3	0.486528	0.685020
Cluster 4	0.456031	0.674949
Cluster 5	0.312383	0.346145
Cluster 6	0.302283	0.307689
Cluster 7	0.168691	0.199277

Berdasarkan **Tabel 5.3** nilai indeks *Partition Entrophy* (PE) dinyatakan baik jika nilai yang diperoleh mendekati kecil atau mendekati 0. Berdasarkan indeks *Partition Entrophy* (PE), metode *Fuzzy C-Means* (FCM) merupakan metode yang terbaik dibandingkan metode *Fuzzy Possibilistic C-Means* (FPCM). Hal tersebut didasarkan pada nilai indeks *Partition Entrophy* (PE) yang mendekati kecil atau mendekati 0 pada keseluruhan *cluster*.

Berikut adalah Perhitungan manual indeks *Partition Coefficient* (PC) dijelaskan dengan persamaan dibawah ini

$$\begin{aligned} PC &= \frac{1}{n} \sum_{k=1}^c \sum_{i=1}^n \mu_{ik}^2 = \frac{1}{8} \sum_{k=1}^2 \sum_{i=1}^8 \mu_{ik}^2 \\ &= \frac{1}{8} ((\mu_{11}^2) + \dots + (\mu_{82}^2)) \end{aligned}$$

$$= \frac{1}{8}(((0.835)^2) + \dots + ((0.768)^2))$$

$$= 0.714512$$

Berikut adalah hasil indeks validitas *cluster* yaitu *Partition Coefficient* (PC) pada metode *Fuzzy C Means* (FCM) dan *Fuzzy Possibilistics C Means* (FPCM).

**Tabel 5.4** Perbandingan *Partition Coefficient*

Jumlah <i>Cluster</i>	<i>Partition Coefficient</i> (PC)	
	FCM	FPCM
<i>Cluster 2</i>	0.714512	0.683978
<i>Cluster 3</i>	0.739263	0.513175
<i>Cluster 4</i>	0.785091	0.522305
<i>Cluster 5</i>	0.867781	0.643257
<i>Cluster 6</i>	0.866774	0.759055
<i>Cluster 7</i>	0.931793	0.878887

Berdasarkan **Tabel 5.4** nilai indeks *Partition Coefficient* (PC) dinyatakan baik jika nilai yang diperoleh mendekati besar atau mendekati 1. Berdasarkan indeks *Partition Coefficient* (PC), metode *Fuzzy C-Means* (FCM) merupakan metode yang terbaik dibandingkan metode *Fuzzy Possibilistic C-Means* (FPCM). Hal tersebut didasarkan pada nilai indeks *Partition Coefficient* (PC) yang mendekati besar atau mendekati 1 pada keseluruhan *cluster*.

Berikut adalah Perhitungan manual indeks *Modified Partition Coefficient* (MPC) dijelaskan dengan persamaan dibawah ini

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

$$= 1 - \frac{2}{2-1} (1 - 0.714512)$$

$$= 1 - 2(0.285488)$$

$$= 0.429024$$

Berikut adalah hasil indeks validitas *cluster* yaitu *Modified Partition Coefficient* (MPC) pada metode *Fuzzy C Means* (FCM) dan *Fuzzy Possibilistics C Means* (FPCM).

**Tabel 5.5** Perbandingan *Modified Partition Coefficient*

Jumlah <i>Cluster</i>	<i>Modified Partition Coefficient</i> (MPC)	
	FCM	FPCM
<i>Cluster 2</i>	0.429024	0.367956

Jumlah Cluster	Modified Partition Coefficient (MPC)	
	FCM	FPCM
Cluster 3	0.608894	0.269763
Cluster 4	0.713455	0.403073
Cluster 5	0.834727	0.554071
Cluster 6	0.840129	0.710866
Cluster 7	0.920425	0.858702

Berdasarkan **Tabel 5.5** nilai indeks *Modified Partition Coefficient* (MPC) dinyatakan baik jika nilai yang diperoleh mendekati besar atau mendekati 1. Berdasarkan indeks *Modified Partition Coefficient* (MPC), metode *Fuzzy C-Means* (FCM) merupakan metode yang terbaik dibandingkan metode *Fuzzy Possibilistic C-Means* (FPCM). Hal tersebut didasarkan pada nilai indeks *Modified Partition Coefficient* (MPC) yang mendekati besar atau mendekati 1 pada keseluruhan cluster.

Berdasarkan ketiga nilai indeks validitas cluster yaitu *Partition Entrophy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC) metode *Fuzzy C-Means* (FCM) merupakan metode yang terbaik dibandingkan metode *Fuzzy Possibilistic C-Means* (FPCM).

#### 4.4 Penentuan Jumlah Cluster Optimal

Pada metode *Fuzzy C-Means* (FCM), jumlah cluster yang optimal dapat dilihat berdasarkan nilai indeks *Partition Entrophy* (PE) yang terendah, *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC) yang tertinggi serta jumlah iterasi dan fungsi objektifnya. Berikut adalah hasil indeks *Partition Entrophy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC) pada metode *Fuzzy C-Means* (FCM).

**Tabel 5.6** Penentuan Jumlah Cluster Optimal

Jumlah Cluster	Fuzzy C-Means (FCM)		
	PE	PC	MPC
Cluster 2	0.455311	0.714512	0.429024
Cluster 3	0.486528	0.739263	0.608894
Cluster 4	0.456031	0.785091	0.713455
Cluster 5	0.312383	0.867781	0.834727
Cluster 6	0.302283	0.866774	0.840129

Jumlah <i>Cluster</i>	<i>Fuzzy C-Means (FCM)</i>		
	PE	PC	MPC
<i>Cluster 7</i>	0.168691	0.931793	0.920425

Berdasarkan **Tabel 5.6** nilai indeks *Partition Entrophy* (PE) yang terendah, *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC) yang tertinggi pada metode *Fuzzy C-Means* (FCM) adalah keseluruhan *cluster* memenuhi kriteria. Jika jumlah *cluster* mendekati jumlah variabel yang akan di clustering maka tujuan *clustering* akan kurang maksimal. Maka berdasarkan **Tabel 5.2** jumlah iterasi yang meminimumkan fungsi objektif pada metode *Fuzzy C-Means* (FCM) terkecil yaitu jumlah *cluster* 5.

#### 4.5 Profilisasi *Fuzzy C-Means* terhadap Hasil *Cluster*

Anggota dari hasil *cluster* Kabupaten/Kota di Provinsi Banten menggunakan metode *Fuzzy C-Means* (FCM) dijelaskan pada **Tabel 5.7**

**Tabel 5.7** Hasil *Cluster* 5

<i>Cluster</i>	Kabupaten/Kota
<i>Cluster 1</i>	Kabupaten Lebak, Kabupaten Pandeglang, dan Kabupaten Serang
<i>Cluster 2</i>	Kota Tangerang Selatan
<i>Cluster 3</i>	Kota Cilegon dan Kota Serang
<i>Cluster 4</i>	Kabupaten Tangerang
<i>Cluster 5</i>	Kota Tangerang

Berdasarkan **Tabel 5.7** *cluster* 1 memiliki 3 anggota yaitu Kabupaten Lebak, Kabupaten Pandeglang, dan Kabupaten Serang. Ketiga Kabupaten/Kota memiliki kesamaan yaitu memiliki jumlah puskesmas yang cukup tinggi, jumlah rumah sakit yang terendah, jumlah penderita penyakit diare yang cukup rendah, dan jumlah penderita penyakit pneumonia yang cukup rendah. Kabupaten Lebak, Kabupaten Pandeglang, dan Kabupaten Serang merupakan Kabupaten/Kota di Provinsi Banten yang memiliki luas wilayah tertinggi, kepadatan penduduk terendah, indeks pembangunan manusia (IPM) terendah, dan laju pertumbuhan penduduk terendah.

*Cluster* 2 hanya memiliki 1 anggota yaitu Kota Tangerang Selatan. Kota Tangerang Selatan memiliki jumlah puskesmas yang cukup rendah, jumlah rumah sakit yang tertinggi, dan jumlah penderita penyakit tuberkulosis yang tertinggi.

Kota Tangerang Selatan merupakan Kabupaten/Kota di Provinsi Banten yang memiliki luas wilayah terendah, kepadatan penduduk tertinggi, indeks pembangunan manusia (IPM) tertinggi, angka harapan hidup (AHH) tertinggi, persentase penduduk miskin terendah, dan laju pertumbuhan penduduk tertinggi.

*Cluster 3* memiliki 2 anggota yaitu Kota Cilegon dan Kota Serang. Kedua Kabupaten/Kota memiliki kesamaan yaitu memiliki jumlah puskesmas yang terendah, jumlah rumah sakit yang cukup rendah, jumlah penderita penyakit diare yang terendah, jumlah penderita penyakit tuberkulosis yang cukup rendah, dan jumlah penderita penyakit pneumonia yang terendah. Kota Cilegon dan Kota Serang merupakan Kabupaten/Kota di Provinsi Banten yang memiliki jumlah penduduk terendah.

*Cluster 4* hanya memiliki 1 anggota yaitu Kabupaten Tangerang. Kabupaten Tangerang memiliki jumlah puskesmas yang tertinggi, jumlah rumah sakit yang cukup tinggi, jumlah penderita penyakit diare yang tertinggi, jumlah penderita penyakit tuberkulosis yang cukup rendah, dan jumlah penderita penyakit pneumonia yang tertinggi. Kabupaten Tangerang merupakan Kabupaten/Kota di Provinsi Banten yang memiliki jumlah penduduk tertinggi, jumlah penduduk miskin tertinggi dan merupakan Kabupaten/Kota kedua di Provinsi Banten yang memiliki laju pertumbuhan tertinggi.

*Cluster 5* hanya memiliki 1 anggota yaitu Kota Tangerang. Kota Tangerang memiliki jumlah puskesmas yang cukup tinggi, jumlah rumah sakit yang tertinggi, jumlah penderita penyakit diare yang cukup tinggi, jumlah penderita penyakit tuberkulosis yang terendah, dan jumlah penderita penyakit pneumonia yang cukup tinggi. Kota Tangerang merupakan Kabupaten/Kota kedua di Provinsi Banten yang memiliki luas wilayah terendah, kepadatan penduduk tertinggi, indeks pembangunan manusia (IPM) tertinggi, dan angka harapan hidup (AHH) tertinggi.

Profilisasi hasil *cluster* adalah untuk mengetahui karakteristik dan sifat-sifat dari masing-masing *cluster*. Profilisasi hasil *cluster* ini dilakukan dengan menghitung rata-rata puskesmas, rumah sakit, diare, tuberkulosis, serta pneumonia pada masing-masing *cluster*. Hasil Profilisasi *cluster* dijelaskan pada **Tabel 5.8**



**Tabel 5.8** Profilisasi Hasil Jumlah *Cluster 5*

<i>Cluster</i>	Puskesmas	Rumah Sakit	Diare	Tuberkulosis	Pneumonia	Banyak Anggota
<i>Cluster 1</i>	36	2	36,473	3,159	2,725	3
<i>Cluster 2</i>	29	29	22,318	6,737	6,517	1
<i>Cluster 3</i>	12	8	9,349	2,098	1,008	2
<i>Cluster 4</i>	44	25	70,107	6,701	12,842	1
<i>Cluster 5</i>	36	29	38,102	1,604	6,655	1

Berdasarkan **Tabel 5.8** *cluster 1* memiliki rata-rata puskesmas sebesar 36, rumah sakit sebesar 2, diare sebesar 36,473, Tuberkulosis sebesar 3,159, serta pneumonia sebesar 2,725. Jumlah anggota *cluster 1* adalah 3 kabupaten/kota. Pada *cluster 2* memiliki rata-rata puskesmas sebesar 29, rumah sakit sebesar 29, diare sebesar 22,318, Tuberkulosis sebesar 6,737, serta pneumonia sebesar 6,517. Jumlah anggota *cluster 2* adalah 1 kabupaten/kota. Pada *cluster 3* memiliki rata-rata puskesmas sebesar 12, rumah sakit sebesar 8, diare sebesar 9,349, Tuberkulosis sebesar 2,098, serta pneumonia sebesar 1,008. Jumlah anggota *cluster 3* adalah 2 kabupaten/kota. Pada *cluster 4* memiliki rata-rata puskesmas sebesar 44, rumah sakit sebesar 25, diare sebesar 70,107, Tuberkulosis sebesar 6,701, serta pneumonia sebesar 12,842. Jumlah anggota *cluster 4* adalah 1 kabupaten/kota. Pada *cluster 5* memiliki rata-rata puskesmas sebesar 36, rumah sakit sebesar 29, diare sebesar 38,102, Tuberkulosis sebesar 1,604, serta pneumonia sebesar 6,655. Jumlah anggota *cluster 5* adalah 1 kabupaten/kota. Tabel tingkat urgensi kesehatan di Provinsi Banten dijelaskan pada **Tabel 5.9**

**Tabel 5.9** Tingkat Urgensi Kesehatan di Provinsi Banten

Tingkat Urgensi	<i>Cluster</i> (Kabupaten/Kota)	Keterangan
1	<i>Cluster 1</i> (Kabupaten Lebak, Kabupaten Pandeglang, Kabupaten Serang)	Tingkat urgensi pertama atau yang paling tinggi karena ketiga Kabupaten/Kota tersebut memiliki jumlah fasilitas pelayanan kesehatan yang rendah, terutama rumah sakit. Sedangkan jumlah penderita penyakit menular yang tinggi, terutama penyakit diare.
2	<i>Cluster 4</i> (Kabupaten Tangerang)	Tingkat urgensi kedua atau cukup tinggi karena Kabupaten/Kota tersebut memiliki jumlah fasilitas pelayanan kesehatan yang tinggi, tetapi jumlah penderita penyakit

<b>Tingkat Urgensi</b>	<b>Cluster (Kabupaten/Kota)</b>	<b>Keterangan</b>
		menular juga sangat tinggi, terutama penyakit diare dan pneumonia.
3	<i>Cluster 3</i> (Kota Cilegon, Kota Serang)	Tingkat urgensi ketiga atau sedang karena kedua Kabupaten/Kota tersebut memiliki jumlah fasilitas pelayanan kesehatan dan jumlah penderita penyakit menular yang seimbang atau tidak terlalu rendah maupun terlalu tinggi.
4	<i>Cluster 5</i> (Kota Tangerang)	Tingkat urgensi keempat atau cukup rendah karena Kabupaten/Kota tersebut memiliki jumlah penderita penyakit menular cukup tinggi, terutama penyakit diare dan pneumonia. Tetapi jumlah fasilitas pelayanan kesehatan juga cukup tinggi.
5	<i>Cluster 2</i> (Kota Tangerang Selatan)	Tingkat urgensi kelima atau rendah karena Kabupaten/Kota tersebut memiliki jumlah penderita penyakit menular cukup rendah, kecuali penyakit tuberkulosis. Tetapi sebanding dengan jumlah fasilitas pelayanan kesehatan yang tinggi.

Tingkat urgensi kesehatan pada **Tabel 5.9** didukung oleh AHH (Angka Harapan Hidup) dan IPM (Indeks Pembangunan Manusia) masing-masing Kabupaten/Kota di Provinsi Banten.



## **BAB VI**

### **PENUTUP**

#### **6.1 Kesimpulan**

Berdasarkan hasil analisis dan pembahasan, diperoleh kesimpulan bahwa:

1. Hasil perbandingan nilai indeks validitas *cluster Partition Entrophy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC) pada metode *Fuzzy C-Means* (FCM) dan *Fuzzy Possibilistic C-Means* (FPCM) diperoleh bahwa metode *Fuzzy C-Means* (FCM) merupakan metode yang terbaik dalam studi kasus ini.
2. Hasil profilisasi pada pengelompokkan Kabupaten/Kota di Provinsi Banten berdasarkan fasilitas pelayanan kesehatan dan penderita penyakit menular adalah *cluster* 1 merupakan Kabupaten/Kota yang memiliki tingkat urgensi kesehatan tertinggi di Provinsi Banten. Hal tersebut dikarenakan adanya ketimpangan antara jumlah fasilitas pelayanan kesehatan dengan jumlah penderita penyakit menular.

#### **6.2 Saran**

Berdasarkan kesimpulan dari analisis, maka diberikan saran sebagai berikut:

1. Untuk penelitian selanjutnya, diharapkan dapat menerapkan metode *Fuzzy C-Means* pada pengelompokkan berdasarkan fasilitas pelayanan kesehatan dan penderita penyakit menular dan diharapkan dapat menggunakan indeks validitas *cluster* lainnya.
2. Untuk instansi pemerintah, terutama pemerintah daerah diharapkan dapat memberikan perhatian khusus pada Kabupaten Lebak, Pandeglang dan Serang karena ketiga Kabupaten/Kota tersebut memiliki tingkat urgensi kesehatan yang tinggi. Perhatian khusus juga untuk pemerataan jumlah rumah sakit dan pengendalian jumlah penderita penyakit menular secara langsung.

## DAFTAR PUSTAKA

- Amirah, M. M., Widodo, A. W., & Dewi, C. (2017). Pengelompokan Lagu Berdasarkan Emosi Menggunakan Algoritma Fuzzy C-Means. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*.
- Anderson, R. E., Babin, B. J., Black, W. C., & Jr., J. F. (2014). *Multivariate Data Analysis*. London: Pearson Education Limited.
- Balasko, B., Abonyi, J., & Feil, B. (2007). *Fuzzy Clustering and Data Analysis Toolbox For Use With Matlab*. Veszprem: University of Veszprem.
- Correa, C., Barreiro, P., Valero, C., Diago, M.-P., & Tardaguila, J. (2014). A Comparison of Fuzzy Clustering Algorithms Applied To Feature Extraction on Vineyard.
- David, Lauro, M. D., & Herwindiati, D. E. (2020). Sistem Prediksi Customer Loyalty Dengan Metode RFM dan Fuzzy C-Means. *Journal of Computer Science and Information Systems*.
- Davies, P. B. (2004). *Database System Third Edition*. New York: Palgrave Macmillan.
- Dinaskesehatan. (2019). *Profil Kesehatan Provinsi Banten*. Diakses dari <https://dinkes.bantenprov.go.id/id/archive/profil-kesehatan-provinsi-bant/1.html>
- Ganbold, G., & Chasia, S. (2016). Comparison Between Possibilistic C-Means (PCM) and Artificial Neural Network (ANN) Classification Algorithms in Land Use/ Land Cover Classification. *International Journal Of Knowledge Content Development & Technology*.
- Grover, N. (2014). A Study of Various Fuzzy Clustering Algorithms. *International Journal of Engineering Research*.
- Handoyo, S., Widodo, A., Nugroho, W. H., & Purwanto, I. N. (2019). The Implementation of A Hybrid Fuzzy Clustering on The Public Health Facility Data. *International Journal of Advanced Trends in Computer Science and Engineering*.

- Haqiqi, B. N., & Kurniawan, R. (2015). Analisis Perbandingan Metode Fuzzy C-Means dan Subtractive Fuzzy C-Means. *Media Statistika*.
- Helma, S. S., Mustakim, R. R. R., & Normala, E. (2019). Clustering Pada Data Fasilitas Pelayanan Kesehatan Kota Pekanbaru Menggunakan Algoritma K-Means. *Jurnal Teknologi Informasi, Komunikasi dan Industri*.
- Irabawati, N., Wahyuningsih, S., & Syoer, R. R. (2016). Perbandingan Metode C-Means dan Fuzzy C-Means Dalam Pengelompokan Wilayah Desa/Kelurahan di Kabupaten Kutai Kartanegara. *Jurnal Eksponensial*.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis Sixth Edition*. New Jersey: Pearson Prentice Hall.
- Kementerian kesehatan. (2011). *Pedoman Nasional Pengendalian Tuberkulosis*. Jakarta: Kementerian Kesehatan Republik Indonesia Direktorat Jenderal Pengendalian Penyakit Dan Penyehatan Lingkungan.
- Kementerian kesehatan. (2014). *Diare*. Diakses dari <https://www.kemkes.go.id/article/view/13010200028/diare.html>
- Kementerian kesehatan. (2014). *Pneumonia Balita*. Diakses dari <https://www.kemkes.go.id/folder/view/01/structure-publikasi-pusdatin-buletin.html>
- Kementerian kesehatan. (2020). *Profil Kesehatan Indonesia 2019*. Jakarta: Kementerian Kesehatan Republik Indonesia. Diakses dari <https://pusdatin.kemkes.go.id/resources/download/pusdatin/profil-kesehatan-indonesia/Profil-Kesehatan-Indonesia-2019.pdf>
- Kementerian kesehatan. (2020). *Rencana Strategis Kementerian Kesehatan Tahun 2020-2024*.
- Kusuma, D. T., & Agani, N. (2015). Prototipe Komparasi Model Clustering Menggunakan Metode K-Means dan FCM Untuk Menentukan Strategi Promosi. *Jurnal TIKOM*.
- Kusuma, D. T., & Sangadji, I. B. (2016). Segmentasi Penilaian Kompetensi Alumni STT-PLN Menggunakan Model Klaster Fuzzy Clustering Means (FCM). *Jurnal Kilat*.

- Kusumadewi, S., & Purnomo, H. (2010). *Aplikasi Logika Fuzzy Untuk Pendukung Keputusan*. Yogyakarta: Graha Ilmu.
- Kusumadewi, S., Hartati, S., Harjoko, A., & Wardoyo, R. (2006). *Fuzzy Multi-Attribute Decision Making (Fuzzy MADM)*. Yogyakarta: Graha Ilmu.
- Larose, D. T. (2005). *Discovering Knowledge In Data An Introduction To Data Mining*. New Jersey: John Wiley & Sons.
- Mashfuufah, S., & Istiawan, D. (2018). Penerapan Partition Entrophy Index, Partition Coefficient Indexs dan Xie Beni Index Untuk Penentuan Jumlah Cluster Optimal Pada Algoritma Fuzzy C-Means Dalam Pemetaan Tingkat Kesejahteraan Penduduk Jawa Tengah.
- Ned, G., & Roger, J. (2017). *Fuzzy Logic Toolbox*. Usa: The Mathworks.
- Ozdemir, O., & Kaya, A. (2019). Comparison Of FCM, PCM, FPCM And PFCM Algorithms In Clustering Methods. *Journal Of Science and Engineering*.
- Permatasari, Y. (2012). Klasifikasi Resiko Bahaya Kehamilan Dengan Metode Fuzzy C-Means. *Skripsi*
- Puspitasari, N., Widians, J. A., & Pohny. (2018). A Clustering Of Generative and Infectious Diseases Using Fuzzy C-Means. *Jurnal Ilmiah Teknologi dan Informasi*.
- Santoso, B. (2007). *Data Mining Teknik Pemanfaatan Data Untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Sanusi, W., Zaky, A., & Afni, B. N. (2018). Analisis Fuzzy C-Means dan Penerapannya Dalam Pengelompokkan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Faktor-Faktor Penyebab Gizi Buruk.
- Siyamto, Y. (2017). Pemanfaatan Data Mining Dengan Metode Clustering Untuk Evaluasi Biaya Dokumen Ekspor di PT Winstar Batam. *Media Informatika Budidarman*.
- Sundari, S. S., & Ariani, N. (2019). Penerapan Data Mining Untuk Pengelompokkan Penyakit Dengan Algoritma Fuzzy C-Means.
- Susilo, F. (2006). *Himpunan dan Logika Kabur Serta Aplikasinya*. Yogyakarta: Graha Ilmu.

- Thilagaraj, T., & Sengottaiyan, D. N. (2019). Implementation Of Fuzzy C-Means and Fuzzy Possibilistic C-Means Algorithms To Find The Low Performers Using R-Tool. *International Journal Of Scientific & Technology Research*.
- Turban, E., Aronson, J., & Liang, T.-P. (2005). *Decision Support Systems and Intelligent Systems Seventh Edition*. New Delhi: Prentice-Hall.





**RINGKASAN TUGAS AKHIR  
(MAKALAH)**

# IMPLEMENTASI FUZZY C-MEANS DAN FUZZY POSSIBILISTIC C-MEANS UNTUK PENGELOMPOKKAN KABUPATEN/KOTA DI PROVINSI BANTEN

Widya Sri Mulyaningsih<sup>1a</sup>, Muhammad Hasan Sidiq Kurniawan<sup>1b</sup>

<sup>1</sup>Program Studi Statistika, Fakultas MIPA, UII

e-mail:

<sup>a</sup>16611021@students.uui.ac.id, <sup>b</sup>hasan.sidiq@uui.ac.id

## Abstrak

Kesehatan adalah suatu kondisi agar jiwa dan raga seseorang sejahtera, sehingga menjadikan seseorang produktif dalam menjalankan aktivitas sehari-hari. Upaya pemerintah guna mendukung peningkatan kualitas manusia Indonesia adalah dengan meningkatkan pengendalian penyakit dan meningkatkan pelayanan kesehatan yang mencakup penguatan sistem kesehatan, pengawasan obat dan makanan. Provinsi Banten menempati urutan lima besar di Indonesia untuk kasus penyakit menular secara langsung yaitu penyakit Tuberkulosis, Diare, dan Pneumonia. Pengelompokan Kabupaten/Kota berdasarkan pelayanan kesehatan dan penyakit yang diderita masyarakat diperlukan untuk mengetahui apakah pemerataan kesehatan masyarakat di Indonesia terutama di Provinsi Banten sudah lebih baik. Metode *clustering* yang digunakan pada penelitian ini adalah *Fuzzy C-Means* (FCM) dan *Fuzzy Possibilistic C-Means* (FPCM). Kedua metode tersebut akan dibandingkan keakurasiannya berdasarkan indeks validitas *Partition Entrophy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC). Berdasarkan nilai indeks validitas *cluster Partition Entrophy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC) didapatkan hasil bahwa metode *Fuzzy C-Means* (FCM) merupakan metode yang terbaik dibandingkan metode *Fuzzy Possibilistic C-Means* (FPCM). Jumlah *cluster* yang optimal pada metode *Fuzzy C-Means* (FCM) adalah 5 *cluster*. Dari kelima *cluster*, *cluster* 1 merupakan *cluster* yang memiliki tingkat urgensi kesehatan tertinggi di Provinsi Banten. Hal tersebut dikarenakan adanya ketimpangan antara jumlah pelayanan kesehatan dengan jumlah penderita penyakit menular.

**Kata kunci:** *Fuzzy C-Means, Fuzzy Possibilistic C-Means, Modified Partition Coefficient, Partition Coefficient, Partition Entrophy*

## Abstract

*Health is a condition for a person's body and soul to be prosperous, thus making someone productive in carrying out their daily activities. The government's efforts to support the improvement of the quality of Indonesian people are by increasing disease control and improving health services which include strengthening the*



health system, controlling drugs and food. Banten Province ranks in the top five in Indonesia for cases of direct infectious diseases, namely Tuberculosis, Diarrhea and Pneumonia. District / city grouping based on health services and diseases suffered by the community is needed to determine whether the distribution of public health in Indonesia, especially in Banten Province. The clustering method used in this research is Fuzzy C-Means (FCM) and Fuzzy Possibilistic C-Means (FPCM). The accuracy of the two methods will be compared based on the validity index of the Partition Entropy (PE), Partition Coefficient (PC) and Modified Partition Coefficient (MPC). Based on the value of the cluster validity index for Partition Entropy (PE), Partition Coefficient (PC) and Modified Partition Coefficient (MPC), the results show that the Fuzzy C-Means (FCM) method is the best method compared to the Fuzzy Possibilistic C-Means (FPCM) method. The optimal number of clusters in the Fuzzy C-Means (FCM) method is 5 clusters. Of the five clusters, cluster 1 is the cluster that has the highest level of health urgency in Banten Province. This is due to the imbalance between the number of health services and the number of people with infectious diseases.

**Keywords:** Fuzzy C-Means, Fuzzy Possibilistic C-Means, Modified Partition Coefficient, Partition Coefficient, Partition Entropy.

## 1. PENDAHULUAN

Kesehatan adalah suatu kondisi agar jiwa dan raga seseorang sejahtera, sehingga menjadikan seseorang produktif dalam menjalankan aktivitas sehari-hari. Upaya pemerintah guna mendukung peningkatan kualitas manusia Indonesia adalah dengan meningkatkan pengendalian penyakit dan meningkatkan pelayanan kesehatan yang mencakup penguatan sistem kesehatan, pengawasan obat dan makanan (Kementerian Kesehatan, 2020).

Provinsi Banten menempati urutan lima besar di Indonesia untuk kasus penyakit menular secara langsung yaitu penyakit Tuberculosis, Diare, dan Pneumonia (Kementerian Kesehatan, 2020). Menurut Dinas Kesehatan Provinsi Banten beberapa permasalahan kesehatan tersebut adalah penyakit Tuberculosis yang memiliki *Case Notification Rate* (CNR) seluruh kasus sebesar 150,53 per 100.000 penduduk, penyakit Diare yang memiliki jumlah kasus sebesar 258.642 per 1.000, dan penyakit Pneumonia yang memiliki jumlah kasus sebesar 36.205 yang mengalami kenaikan dari tahun 2017. Di provinsi Banten jumlah Puskesmas ada sebanyak 238 dan jumlah rumah sakit sebanyak 112. Jumlah rumah sakit khusus dan rumah sakit umum di Provinsi Banten mengalami kenaikan dari tahun 2018 karena banyak rumah sakit yang baru beroperasi dan rasio puskesmas per kecamatan di Provinsi Banten sudah di atas rasio nasional yaitu sebesar 1,57 (Dinas Kesehatan, 2019).

Pengelompokan Kabupaten/Kota berdasarkan pelayanan kesehatan dan penyakit yang diderita masyarakat diperlukan untuk mengetahui apakah pemerataan kesehatan masyarakat di Indonesia terutama di Provinsi Banten sudah baik. Salah satu metode *clustering* adalah metode *Fuzzy C-Means* (FCM) dan *Fuzzy Possibilistic C-Means* (FPCM). Kedua metode tersebut akan dibandingkan



keakurasiannya berdasarkan indeks validitas *Partition Entrophy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC).

### **2.1 Rumusan Masalah**

Berdasarkan latar belakang yang telah dipaparkan, maka dapat dirumuskan beberapa permasalahan yang akan diteliti yaitu:

1. Bagaimana hasil perbandingan indeks validitas *cluster* pada metode *Fuzzy C-Means* dan *Fuzzy Possibilitic C-Means* pada data fasilitas pelayanan kesehatan dan penderita penyakit menular di Provinsi Banten?
2. Bagaimana hasil profilisasi pada pengelompokkan kabupaten/kota di Provinsi Banten berdasarkan fasilitas pelayanan kesehatan dan penderita penyakit menular?

### **2.2 Batasan Masalah**

Batasan masalah dalam penelitian ini adalah:

1. Data yang digunakan dalam penelitian ini adalah data jumlah penderita penyakit tuberkolosis, diare dan pneumonia serta jumlah PUSKESMAS (Pusat Kesehatan Masyarakat) dan rumah sakit tahun 2019.
2. Sumber data yang digunakan dalam penelitian ini adalah data sekunder yang didapatkan dari *e-book* Dinas Kesehatan Provinsi Banten dan Badan PPSDM Kesehatan.
3. Metode yang digunakan dalam penelitian ini adalah metode *Fuzzy C-Means* (FCM) dan *Fuzzy Possibilistic C-Means* (FPCM).
4. Data diolah dengan menggunakan bantuan *software* R dan *Microsoft Excel*.

### **2.3 Tujuan Penelitian**

Adapun tujuan yang ingin dicapai dalam penelitian ini yaitu:

1. Untuk mengetahui hasil perbandingan indeks validitas *cluster* pada metode *Fuzzy C-Means* dan *Fuzzy Possibilitic C-Means* pada data fasilitas pelayanan kesehatan dan penderita penyakit menular di Provinsi Banten.
2. Untuk mengetahui hasil profilisasi pada pengelompokkan kabupaten/kota di Provinsi Banten berdasarkan fasilitas pelayanan kesehatan dan penderita penyakit menular.

### **2.4 Manfaat Penelitian**

Manfaat dari penelitian ini adalah sebagai salah satu sarana bagi pembaca dan pelengkap referensi bagi pihak-pihak yang memerlukan dalam menyelesaikan kasus untuk pengelompokkan yang menggunakan metode *clustering* terutama metode *Fuzzy C-Means* dan *Fuzzy Possibilitics C-Means*. Selain itu juga dapat menjadi dasar pengambilan keputusan dan kebijakan pemerintah daerah yang mengarah pada upaya peningkatan pembangunan fasilitas kesehatan di seluruh wilayah Banten berdasarkan studi kasus yang digunakan.

## **2. TINJAUAN PUSTAKA**

Terdapat beberapa penelitian yang berkaitan dengan metode *Fuzzy C-Means* dan *Fuzzy Possibilitics C-Means* sebelumnya. Diantara penelitian-penelitian

sebelumnya adalah penelitian yang dilakukan oleh Christian Correa, Pillar Barreiro, Constantino Valero, Maria-Paz Diago, dan Javier Tardaguila (2014) dengan judul “*A Comparison of Fuzzy Clustering Algorithms Applied to Feature Extraction on Vineyard*” dengan hasil penelitian bahwa metode terbaik berdasarkan kecepatan dan performanya adalah metode *Fuzzy C-Means* (FCM) dibandingkan dengan metode *Possibilistic C-Means* (PCM), *Fuzzy Possibilistic C-Means* (FPCM), *Robust Fuzzy Possibilistic C-Means* (RFCM) dan *Fuzzy C-Means with Gustafson-Kessel algorithm* (FCM-GK). Pada penelitian yang dilakukan oleh Wahidah Sanusi, Ahmad Zaky, dan Besse Nur Afni (2018) dengan judul “Analisis *Fuzzy C-Means* dan Penerapannya Dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Faktor-Faktor Penyebab Gizi Buruk” menyatakan bahwa dengan menggunakan metode *Fuzzy C-Means* (FCM) didapatkan pengelompokan 2 *cluster* dengan fungsi objektif sebesar 1079141912,2224.

Terdapat juga beberapa penelitian yang berkaitan dengan kesehatan, seperti penelitian terkait penyakit tuberkulosis, penyakit diare, penyakit pneumonia dan pelayanan kesehatan. Penelitian-penelitian tersebut antara lain adalah penelitian yang dilakukan oleh Novianti Puspitasari, Joan Angelina Widians dan Pohny (2018) dengan judul “*A Clustering of Generative and Infectious Diseases Using Fuzzy C-Means*” didapatkan hasil pengelompokan 2 *cluster* dengan perhitungan validitas *cluster* menggunakan *Partition Coefficient* (PC) sebesar 0,685426108. Penelitian dengan judul “*The Implementation of Hybrid Fuzzy Clustering on The Public Health Facility Data*” oleh Samingun Handoyo, Agus Widodo, Waego Hadi Nugroho, dan Imam N Purwanto (2019) didapatkan hasil bahwa metode terbaik untuk pengelompokan data fasilitas kesehatan adalah metode *Fuzzy C-Means* (FCM) dibandingkan dengan metode *Subtractive Fuzzy C-Means* (SFCM).

### **3. METODOLOGI PENELITIAN**

#### **3.1 Populasi Penelitian**

Populasi yang digunakan dalam penelitian ini yaitu seluruh Kabupaten/Kota di Provinsi Banten. Sampel yang digunakan dalam penelitian ini yaitu seluruh Kabupaten/Kota di Provinsi Banten yang diukur berdasarkan fasilitas pelayanan kesehatan dan penderita penyakit menular pada tahun 2019.

#### **3.2 Tempat dan Waktu Penelitian**

Penelitian ini dilakukan pada bulan September 2020 sampai dengan Desember 2020 di Universitas Islam Indonesia Yogyakarta.

#### **3.3 Variabel Penelitian**

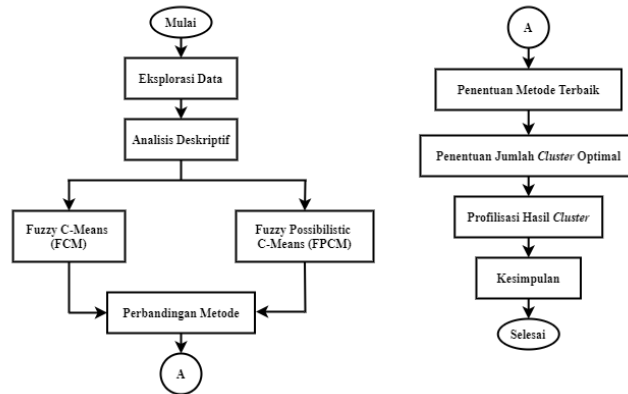
Variabel yang digunakan dalam penelitian ini adalah jumlah puskesmas, jumlah rumah sakit, jumlah penderita penyakit tuberkulosis, diare dan pneumonia pada tahun 2019. Banyaknya Kabupaten/Kota sesuai dengan jumlah Kabupaten/Kota di Provinsi Banten yaitu 8. Untuk variabel yang digunakan, dijelaskan pada **Tabel 3.1**

**Tabel 3.1** Penjelasan Variabel

No	Variabel	Definisi Operasional	Satuan/ Skala	Contoh
1	Jumlah Puskesmas	Jumlah fasilitas pelayanan kesehatan yang menyelenggarakan upaya kesehatan masyarakat dan upaya kesehatan perseorangan tingkat pertama	Rasio	Jumlah puskesmas di Kabupaten Lebak tahun 2019 adalah 42
2	Jumlah Rumah Sakit	Jumlah fasilitas pelayanan kesehatan yang menyelenggarakan upaya kesehatan masyarakat dan upaya kesehatan perseorangan tingkat rujukan lanjut	Rasio	Jumlah rumah sakit di Kabupaten Lebak tahun 2019 adalah 2
3	Jumlah Penderita Tuberkulosis	Jumlah penderita penyakit menular yang disebabkan oleh kuman atau bakteri <i>Mycobacterium tuberculosis</i> (TB)	Rasio	Jumlah penderita penyakit tuberkulosis di Kabupaten Lebak tahun 2019 adalah 2,833 orang
4	Jumlah Penderita Diare	Jumlah penderita penyakit menular pada saluran pencernaan yang disebabkan oleh organisme seperti virus, bakteri, dan parasit	Rasio	Jumlah penderita penyakit diare di Kabupaten Lebak tahun 2019 adalah 31,674 orang
5	Jumlah Penderita Pneumonia	Jumlah penderita penyakit menular yang menginfeksi jaringan paru-paru yang disebabkan oleh mikroorganisme seperti virus, jamur dan bakteri	Rasio	Jumlah penderita penyakit pneumonia di Kabupaten Lebak tahun 2019 adalah 3,092 orang

### 3.4 Tahap Penelitian

Berikut adalah diagram alir metode penelitian:



**Gambar 3.1** Flowchart Tahap Penelitian

**Gambar 3.1** menjelaskan tahapan peneliti yang dilakukan oleh penulis. Langkah pertama yaitu eksplorasi data yaitu tahap mencari dan pengambilan data. Analisis awal yang dilakukan adalah analisis deskriptif berupa gambaran umum dari data yang akan dianalisis. Selanjutnya, perbandingan metode dilakukan pada 2 metode yaitu *Fuzzy C-Means* (FCM) dan *Fuzzy Possibilistic C-Means* (FPCM) berdasarkan jumlah iterasi dan fungsi objektif. Penentuan metode terbaik dan jumlah *cluster* optimal didasarkan pada nilai indeks *Partition Entrophy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC). Kemudian profilisasi hasil *cluster* dan kesimpulan.

## 4. PEMBAHASAN

### 4.1 Analisis Deskriptif

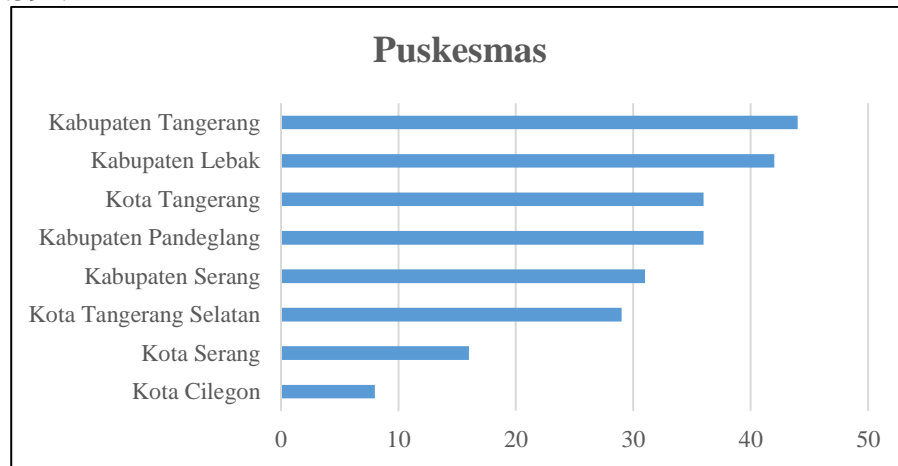
Analisis deskriptif digunakan untuk memberikan gambaran umum terkait data yang di analisis. Adapun gambaran umum yang akan ditampilkan adalah *median* (nilai tengah), *std. deviation* (simpangan baku), *minimum* (nilai terendah), dan *maximum* (nilai tertinggi)

**Tabel 4.1** Statistika Deskriptif Kesehatan di Provinsi Banten

	Puskesmas	Rumah Sakit	Diare (Orang)	Tuberkulosis (Orang)	Pneumonia (Orang)
<i>Median</i>	33.50	8	32,295	2,844	3,360
<i>Std. Devition</i>	12.498	12.552	19,909.75	2,054.395	4,053.394
<i>Minimum</i>	8	1	8,703	1,604	611
<i>Maximum</i>	44	29	70,107	6,737	12,842

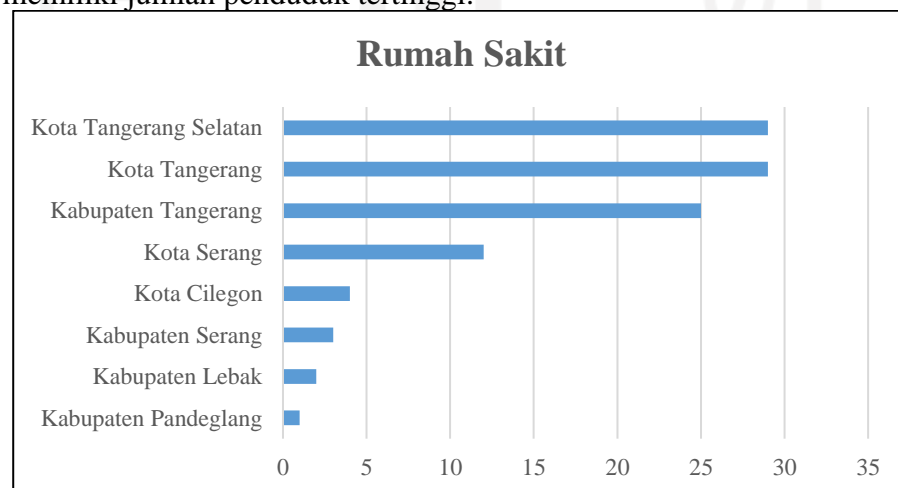
Berdasarkan pada **Tabel 4.1** *Median* atau nilai tengah dari jumlah puskesmas yaitu sebesar 33.50, nilai tengah dari jumlah rumah sakit yaitu sebesar 8, nilai tengah dari jumlah penderita penyakit diare yaitu sebesar 32,295 orang, nilai tengah dari jumlah penderita penyakit tuberkulosis yaitu sebesar 2,844 orang, dan nilai tengah dari jumlah penderita penyakit pneumonia yaitu sebesar 3,360 orang. Standar deviasi atau simpangan baku dari variabel puskesmas yaitu sebesar 12.498,

variabel rumah sakit yaitu sebesar 12.552, variabel diare yaitu sebesar 19,909.75, variabel tuberkulosis yaitu sebesar 2,054.395, dan variabel pneumonia yaitu sebesar 4,053.394.



**Gambar 5.1** Diagram Batang Jumlah Puskesmas

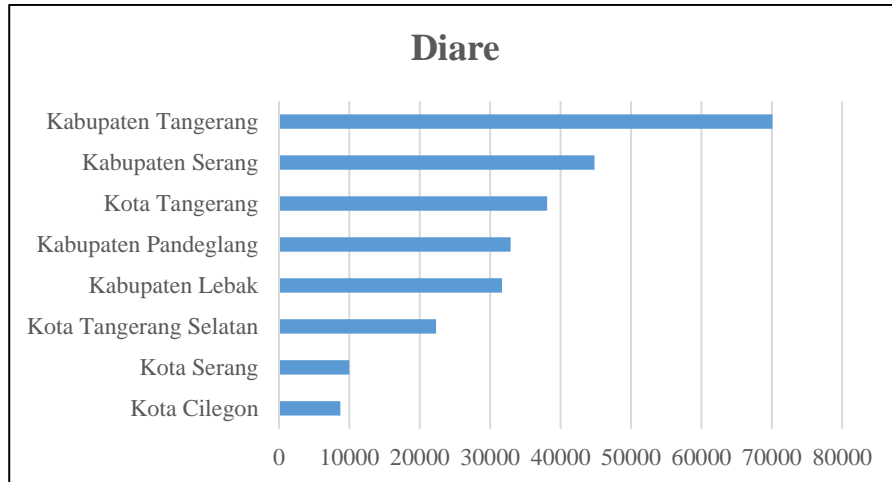
Berdasarkan **Tabel 5.1** dan **Gambar 5.1** nilai minimum dari variabel puskesmas yaitu sebesar 8. Kabupaten/Kota yang memiliki jumlah puskesmas terendah yaitu Kota Cilegon. Kota Cilegon merupakan salah satu Kabupaten/Kota di Provinsi Banten yang memiliki luas wilayah terendah dan merupakan Kabupaten/Kota di Provinsi Banten yang memiliki jumlah penduduk terendah. Sementara nilai maksimum dari variabel puskesmas yaitu sebesar 44. Kabupaten/Kota yang memiliki jumlah puskesmas tertinggi yaitu Kabupaten Tangerang. Kabupaten Tangerang merupakan Kabupaten/Kota di Provinsi Banten yang memiliki jumlah penduduk tertinggi.



**Gambar 5.2** Diagram Batang Jumlah Rumah Sakit

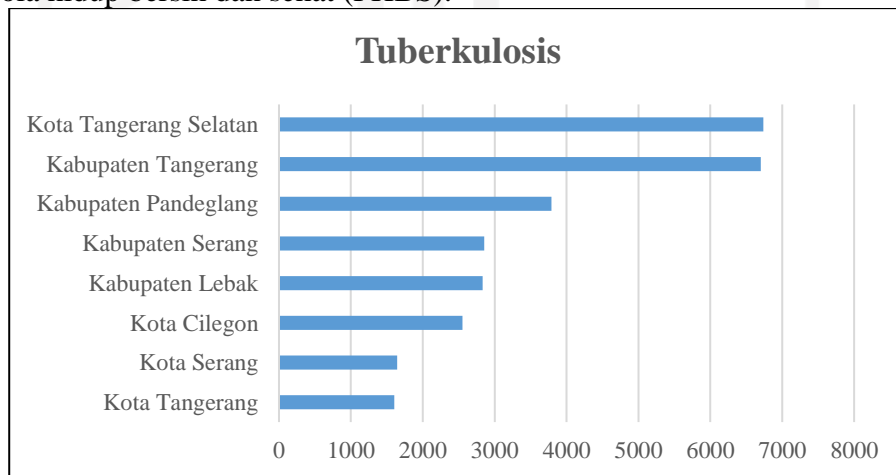
Berdasarkan **Tabel 5.1** dan **Gambar 5.2** nilai minimum dari variabel rumah sakit yaitu sebesar 1. Kabupaten/Pandeglang yang memiliki jumlah rumah sakit terendah yaitu Kabupaten Pandeglang. Sementara nilai maksimum dari variabel rumah sakit yaitu sebesar 29. Kabupaten/Kota yang memiliki jumlah rumah sakit tertinggi yaitu Kota Tangerang dan Kota Tangerang Selatan. Kota Tangerang dan Kota Tangerang

Selatan merupakan Kabupaten/Kota di Provinsi Banten yang paling dekat dengan ibu kota Jakarta. Berdasarkan **Gambar 5.2** terlihat adanya ketimpangan jumlah sakit di Provinsi Banten.



**Gambar 5.3** Diagram Batang Jumlah Penderita Diare

Berdasarkan **Tabel 5.1** dan **Gambar 5.3** nilai minimum dari variabel diare yaitu sebesar 8,703 orang. Kabupaten/Kota yang memiliki jumlah penderita penyakit diare terendah yaitu Kota Cilegon. Sementara nilai maksimum dari variabel diare yaitu sebesar 70,107 orang. Kabupaten/Kota yang memiliki jumlah penderita penyakit diare tertinggi yaitu Kabupaten Tangerang. Faktor utama resiko diare adalah faktor lingkungan seperti sarana air bersih, cara penyajian makanan dan pola hidup bersih dan sehat (PHBS).

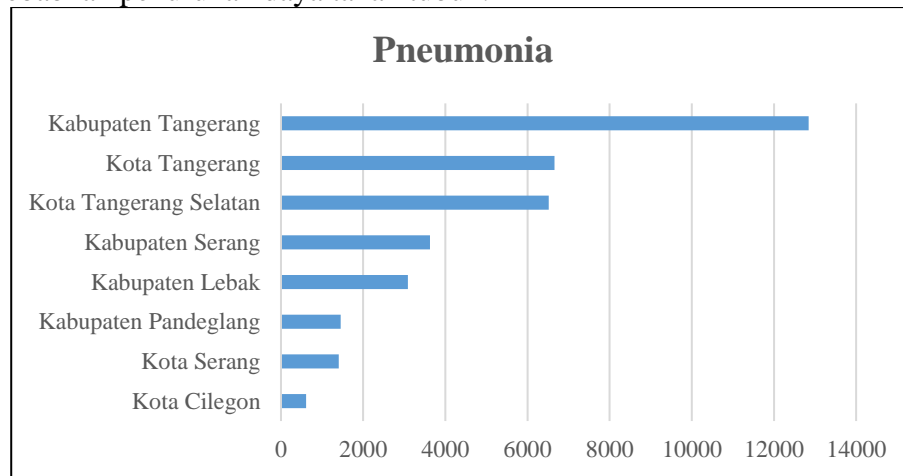


**Gambar 5.4** Diagram Batang Jumlah Penderita Tuberkulosis

Berdasarkan **Tabel 5.1** dan **Gambar 5.4** nilai minimum dari variabel tuberkulosis yaitu sebesar 1,604 orang. Kabupaten/Kota yang memiliki jumlah penderita penyakit tuberkulosis terendah yaitu Kota Tangerang. Sementara nilai maksimum dari variabel tuberkulosis yaitu sebesar 6,737 orang. Kabupaten/Kota yang memiliki jumlah penderita penyakit tuberkulosis tertinggi yaitu Kota Tangerang Selatan. Faktor penyebab tingginya tuberkulosis adalah besarnya masalah kesehatan lain yang bisa berpengaruh terhadap resiko terjadinya



tuberkulosis secara signifikan seperti merokok, HIV dan lainnya yang menyebabkan penurunan daya tahan tubuh.



**Gambar 5.5** Diagram Batang Jumlah Penderita Pneumonia Pada Balita

Berdasarkan **Tabel 5.1** dan **Gambar 5.5** nilai minimum dari variabel pneumonia pada balita yaitu sebesar 611 orang. Kabupaten/Kota yang memiliki jumlah penderita penyakit pneumonia terendah yaitu Kota Cilegon. Kota Cilegon merupakan Kabupaten/Kota di Provinsi Banten yang memiliki jumlah balita terendah. Sementara nilai maksimum dari variabel pneumonia pada balita yaitu sebesar 12,842 orang. Kabupaten/Kota yang memiliki jumlah penderita penyakit pneumonia tertinggi yaitu Kabupaten Tangerang.

#### 4.2 Perbandingan *Fuzzy C-Means* dan *Fuzzy Possibilistics C-Means*

Pada metode *Fuzzy C-Means* (FCM) nilai pembobot ( $w$ ) yang digunakan adalah 2 dan nilai maksimum iterasi yang digunakan adalah 1000. Pada metode *Fuzzy Possibilistic C-Means* (FPCM), nilai pembobot untuk *Fuzzy C-Means* ( $w$ ) adalah 2, nilai pembobot untuk *Possibilistic C-Means* ( $\eta$ ) adalah 2 dan nilai maksimum iterasi adalah 1000. Perbandingan metode *Fuzzy C-Means* (FCM) dan *Fuzzy Possibilistic C-Means* (FPCM) dapat dilakukan dalam berbagai cara, salah satunya yaitu membandingkan jumlah iterasi dan fungsi objektifnya. Berikut adalah hasil perbandingan kedua metode berdasarkan jumlah iterasi dan fungsi objektifnya:

**Tabel 4.2** Jumlah Iterasi dan Fungsi Objektif

Jumlah Cluster	Jumlah Iterasi		Fungsi Objektif	
	FCM	FPCM	FCM	FPCM
Cluster 2	48	25	14.88332	22.18462
Cluster 3	56	55	7.046465	22.69952
Cluster 4	38	55	3.510449	14.79498
Cluster 5	25	44	1.301461	4.617783
Cluster 6	64	20	0.7187028	3.420215
Cluster 7	46	17	0.4474671	2.224091

Berdasarkan **Tabel 4.2** jumlah iterasi yang dilakukan oleh metode *Fuzzy C-Means* (FCM) untuk meminimumkan fungsi objektif adalah 48, 56, 38, 25, 64, dan 46. Sementara pada metode *Fuzzy Possibilistic C-Means* (FPCM) adalah 25, 55, 55, 44, 20, dan 17. Berdasarkan jumlah iterasi dapat disimpulkan bahwa metode *Fuzzy Possibilistic C-Means* (FPCM) memiliki jumlah iterasi yang paling sedikit pada hampir keseluruhan *cluster* untuk meminimumkan fungsi objektif, sehingga metode *Fuzzy Possibilistic C-Means* (FPCM) memiliki waktu komputasi lebih cepat dibandingkan metode *Fuzzy C-Means* (FCM). Fungsi objektif pada metode *Fuzzy C-Means* (FCM) secara berturut-turut adalah 14.88332, 7.046465, 3.510449, 1.301461, 0.7187028, dan 0.4474671. Sementara pada metode *Fuzzy Possibilistic C-Means* (FPCM) adalah 22.18462, 22.69952, 14.79498, 4.617783, 3.420215, dan 2.224091. Berdasarkan fungsi objektif kedua metode tidak bisa dibandingkan karena perhitungan fungsi objektif yang berbeda pada setiap metode.

Hasil  $\mu_{ik}$  (matriks partisi) berpengaruh terhadap hasil fungsi objektif, jika hasil  $\mu_{ik}$  kecil maka hasil fungsi objektif juga akan kecil. Hasil fungsi objektif akan memberikan *clustering* yang optimum.

### 4.3 Penentuan Metode Terbaik

Penentuan metode terbaik dapat dilihat berdasarkan nilai indeks validitas *cluster* yaitu *Partition Entrophy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC). Berikut adalah hasil indeks validitas *cluster* yaitu *Partition Entrophy* (PE) pada metode *Fuzzy C Means* (FCM) dan *Fuzzy Possibilistics C Means* (FPCM).

**Tabel 4.3** Perbandingan *Partition Entrophy*

Jumlah Cluster	<i>Partition Entrophy</i> (PE)	
	FCM	FPCM
Cluster 2	0.455311	0.518594
Cluster 3	0.486528	0.685020
Cluster 4	0.456031	0.674949
Cluster 5	0.312383	0.346145
Cluster 6	0.302283	0.307689
Cluster 7	0.168691	0.199277

Berdasarkan **Tabel 4.3** nilai indeks *Partition Entrophy* (PE) dinyatakan baik jika nilai yang diperoleh mendekati kecil atau mendekati 0. Berdasarkan indeks *Partition Entrophy* (PE), metode *Fuzzy C-Means* (FCM) merupakan metode yang terbaik dibandingkan metode *Fuzzy Possibilistic C-Means* (FPCM). Hal tersebut didasarkan pada nilai indeks *Partition Entrophy* (PE) yang mendekati kecil atau mendekati 0 pada keseluruhan *cluster*.

Berikut adalah hasil indeks validitas *cluster* yaitu *Partition Coefficient* (PC) pada metode *Fuzzy C Means* (FCM) dan *Fuzzy Possibilistics C Means* (FPCM).

**Tabel 4.4** Perbandingan *Partition Coefficient*

Jumlah Cluster	<i>Partition Coefficient</i> (PC)	
	FCM	FPCM
Cluster 2	0.714512	0.683978



Jumlah Cluster	Partition Coefficient (PC)	
	FCM	FPCM
Cluster 3	0.739263	0.513175
Cluster 4	0.785091	0.522305
Cluster 5	0.867781	0.643257
Cluster 6	0.866774	0.759055
Cluster 7	0.931793	0.878887

Berdasarkan **Tabel 4.4** nilai indeks *Partition Coefficient* (PC) dinyatakan baik jika nilai yang diperoleh mendekati besar atau mendekati 1. Berdasarkan indeks *Partition Coefficient* (PC), metode *Fuzzy C-Means* (FCM) merupakan metode yang terbaik dibandingkan metode *Fuzzy Possibilistic C-Means* (FPCM). Hal tersebut didasarkan pada nilai indeks *Partition Coefficient* (PC) yang mendekati besar atau mendekati 1 pada keseluruhan cluster.

Berikut adalah hasil indeks validitas cluster yaitu *Modified Partition Coefficient* (MPC) pada metode *Fuzzy C Means* (FCM) dan *Fuzzy Possibilistics C Means* (FPCM).

**Tabel 4.5** Perbandingan *Modified Partition Coefficient*

Jumlah Cluster	Modified Partition Coefficient (MPC)	
	FCM	FPCM
Cluster 2	0.429024	0.367956
Cluster 3	0.608894	0.269763
Cluster 4	0.713455	0.403073
Cluster 5	0.834727	0.554071
Cluster 6	0.840129	0.710866
Cluster 7	0.920425	0.858702

Berdasarkan **Tabel 4.5** nilai indeks *Modified Partition Coefficient* (MPC) dinyatakan baik jika nilai yang diperoleh mendekati besar atau mendekati 1. Berdasarkan indeks *Modified Partition Coefficient* (MPC), metode *Fuzzy C-Means* (FCM) merupakan metode yang terbaik dibandingkan metode *Fuzzy Possibilistic C-Means* (FPCM). Hal tersebut didasarkan pada nilai indeks *Modified Partition Coefficient* (MPC) yang mendekati besar atau mendekati 1 pada keseluruhan cluster.

Berdasarkan ketiga nilai indeks validitas cluster yaitu *Partition Entrophy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC) metode *Fuzzy C-Means* (FCM) merupakan metode yang terbaik dibandingkan metode *Fuzzy Possibilistic C-Means* (FPCM).

#### 4.4 Penentuan Jumlah Cluster Optimal

Pada metode *Fuzzy C-Means* (FCM), jumlah cluster yang optimal dapat dilihat berdasarkan nilai indeks *Partition Entrophy* (PE) yang terendah, *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC) yang tertinggi serta jumlah iterasi dan fungsi objektifnya. Berikut adalah hasil indeks *Partition*

Entropy (PE), Partition Coefficient (PC) dan Modified Partition Coefficient (MPC) pada metode *Fuzzy C-Means* (FCM).

**Tabel 4.6** Penentuan Jumlah *Cluster* Optimal

Jumlah <i>Cluster</i>	<i>Fuzzy C-Means</i> (FCM)		
	PE	PC	MPC
<i>Cluster 2</i>	0.455311	0.714512	0.429024
<i>Cluster 3</i>	0.486528	0.739263	0.608894
<i>Cluster 4</i>	0.456031	0.785091	0.713455
<i>Cluster 5</i>	0.312383	0.867781	0.834727
<i>Cluster 6</i>	0.302283	0.866774	0.840129
<i>Cluster 7</i>	0.168691	0.931793	0.920425

Berdasarkan **Tabel 4.6** nilai indeks *Partition Entropy* (PE) yang terendah, *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC) yang tertinggi pada metode *Fuzzy C-Means* (FCM) adalah keseluruhan *cluster* memenuhi kriteria. Jika jumlah *cluster* mendekati jumlah variabel yang akan di clustering maka tujuan clustering akan kurang maksimal. Maka berdasarkan **Tabel 4.2** jumlah iterasi yang meminimumkan fungsi objektif pada metode *Fuzzy C-Means* (FCM) terkecil yaitu jumlah *cluster 5*.

#### 4.5 Profilisasi *Fuzzy C-Means* terhadap Hasil *Cluster*

Anggota dari hasil *cluster* Kabupaten/Kota di Provinsi Banten menggunakan metode *Fuzzy C-Means* (FCM) dijelaskan pada **Tabel 4.7**

**Tabel 4.7** Hasil *Cluster 5*

<i>Cluster</i>	Kabupaten/Kota
<i>Cluster 1</i>	Kabupaten Lebak, Kabupaten Pandeglang, dan Kabupaten Serang
<i>Cluster 2</i>	Kota Tangerang Selatan
<i>Cluster 3</i>	Kota Cilegon dan Kota Serang
<i>Cluster 4</i>	Kabupaten Tangerang
<i>Cluster 5</i>	Kota Tangerang

Berdasarkan **Tabel 4.7** *cluster 1* memiliki 3 anggota yaitu Kabupaten Lebak, Kabupaten Pandeglang, dan Kabupaten Serang. Ketiga Kabupaten/Kota memiliki kesamaan yaitu memiliki jumlah puskesmas yang cukup tinggi, jumlah rumah sakit yang terendah, jumlah penderita penyakit diare yang cukup rendah, dan jumlah penderita penyakit pneumonia yang cukup rendah. Kabupaten Lebak, Kabupaten Pandeglang, dan Kabupaten Serang merupakan Kabupaten/Kota di Provinsi Banten yang memiliki luas wilayah tertinggi, kepadatan penduduk terendah, indeks pembangunan manusia (IPM) terendah, dan laju pertumbuhan penduduk terendah.

*Cluster 2* hanya memiliki 1 anggota yaitu Kota Tangerang Selatan. Kota Tangerang Selatan memiliki jumlah puskesmas yang cukup rendah, jumlah rumah sakit yang tertinggi, dan jumlah penderita penyakit tuberkulosis yang tertinggi. Kota Tangerang Selatan merupakan Kabupaten/Kota di Provinsi Banten yang memiliki luas wilayah terendah, kepadatan penduduk tertinggi, indeks

pembangunan manusia (IPM) tertinggi, angka harapan hidup (AHH) tertinggi, persentase penduduk miskin terendah, dan laju pertumbuhan penduduk tertinggi.

*Cluster 3* memiliki 2 anggota yaitu Kota Cilegon dan Kota Serang. Kedua Kabupaten/Kota memiliki kesamaan yaitu memiliki jumlah puskesmas yang terendah, jumlah rumah sakit yang cukup rendah, jumlah penderita penyakit diare yang terendah, jumlah penderita penyakit tuberkulosis yang cukup rendah, dan jumlah penderita penyakit pneumonia yang terendah. Kota Cilegon dan Kota Serang merupakan Kabupaten/Kota di Provinsi Banten yang memiliki jumlah penduduk terendah.

*Cluster 4* hanya memiliki 1 anggota yaitu Kabupaten Tangerang. Kabupaten Tangerang memiliki jumlah puskesmas yang tertinggi, jumlah rumah sakit yang cukup tinggi, jumlah penderita penyakit diare yang tertinggi, jumlah penderita penyakit tuberkulosis yang cukup rendah, dan jumlah penderita penyakit pneumonia yang tertinggi. Kabupaten Tangerang merupakan Kabupaten/Kota di Provinsi Banten yang memiliki jumlah penduduk tertinggi, jumlah penduduk miskin tertinggi dan merupakan Kabupaten/Kota kedua di Provinsi Banten yang memiliki laju pertumbuhan tertinggi.

*Cluster 5* hanya memiliki 1 anggota yaitu Kota Tangerang. Kota Tangerang memiliki jumlah puskesmas yang cukup tinggi, jumlah rumah sakit yang tertinggi, jumlah penderita penyakit diare yang cukup tinggi, jumlah penderita penyakit tuberkulosis yang terendah, dan jumlah penderita penyakit pneumonia yang cukup tinggi. Kota Tangerang merupakan Kabupaten/Kota kedua di Provinsi Banten yang memiliki luas wilayah terendah, kepadatan penduduk tertinggi, indeks pembangunan manusia (IPM) tertinggi, dan angka harapan hidup (AHH) tertinggi.

Profilisasi hasil *cluster* adalah untuk mengetahui karakteristik dan sifat-sifat dari masing-masing *cluster*. Profilisasi hasil *cluster* ini dilakukan dengan menghitung rata-rata puskesmas, rumah sakit, diare, tuberkulosis, serta pneumonia pada masing-masing *cluster*. Hasil Profilisasi *cluster* dijelaskan pada **Tabel 4.8**

**Tabel 4.8** Profilisasi Hasil Jumlah *Cluster 5*

<i>Cluster</i>	Puskesmas	Rumah Sakit	Diare	Tuberkulosis	Pneumonia	Banyak Anggota
<i>Cluster 1</i>	36	2	36,473	3,159	2,725	3
<i>Cluster 2</i>	29	29	22,318	6,737	6,517	1
<i>Cluster 3</i>	12	8	9,349	2,098	1,008	2
<i>Cluster 4</i>	44	25	70,107	6,701	12,842	1
<i>Cluster 5</i>	36	29	38,102	1,604	6,655	1

Berdasarkan **Tabel 4.8** *cluster 1* memiliki rata-rata puskesmas sebesar 36, rumah sakit sebesar 2, diare sebesar 36,473, Tuberkulosis sebesar 3,159, serta pneumonia sebesar 2,725. Jumlah anggota *cluster 1* adalah 3 kabupaten/kota. Pada *cluster 2* memiliki rata-rata puskesmas sebesar 29, rumah sakit sebesar 29, diare sebesar 22,318, Tuberkulosis sebesar 6,737, serta pneumonia sebesar 6,517. Jumlah anggota *cluster 2* adalah 1 kabupaten/kota. Pada *cluster 3* memiliki rata-rata puskesmas sebesar 12, rumah sakit sebesar 8, diare sebesar 9,349, Tuberkulosis sebesar 2,098, serta pneumonia sebesar 1,008. Jumlah anggota *cluster 3* adalah 2 kabupaten/kota. Pada *cluster 4* memiliki rata-rata puskesmas sebesar 44, rumah

sakit sebesar 25, diare sebesar 70,107, Tuberkulosis sebesar 6,701, serta pneumonia sebesar 12,842. Jumlah anggota *cluster* 4 adalah 1 kabupaten/kota. Pada *cluster* 5 memiliki rata-rata puskesmas sebesar 36, rumah sakit sebesar 29, diare sebesar 38,102, Tuberkulosis sebesar 1,604, serta pneumonia sebesar 6,655. Jumlah anggota *cluster* 5 adalah 1 kabupaten/kota. Tabel tingkat urgensi kesehatan di Provinsi Banten dijelaskan pada **Tabel 4.9**

**Tabel 4.9** Tingkat Urgensi Kesehatan di Provinsi Banten

Tingkat Urgensi	Cluster (Kabupaten/Kota)	Keterangan
1	Cluster 1 (Kabupaten Lebak, Kabupaten Pandeglang, Kabupaten Serang)	Tingkat urgensi pertama atau yang paling tinggi karena ketiga Kabupaten/Kota tersebut memiliki jumlah fasilitas pelayanan kesehatan yang rendah, terutama rumah sakit. Sedangkan jumlah penderita penyakit menular yang tinggi, terutama penyakit diare.
2	Cluster 4 (Kabupaten Tangerang)	Tingkat urgensi kedua atau cukup tinggi karena Kabupaten/Kota tersebut memiliki jumlah fasilitas pelayanan kesehatan yang tinggi, tetapi jumlah penderita penyakit menular juga sangat tinggi, terutama penyakit diare dan pneumonia.
3	Cluster 3 (Kota Cilegon, Kota Serang)	Tingkat urgensi ketiga atau sedang karena kedua Kabupaten/Kota tersebut memiliki jumlah fasilitas pelayanan kesehatan dan jumlah penderita penyakit menular yang seimbang atau tidak terlalu rendah maupun terlalu tinggi.
4	Cluster 5 (Kota Tangerang)	Tingkat urgensi keempat atau cukup rendah karena Kabupaten/Kota tersebut memiliki jumlah penderita penyakit menular cukup tinggi, terutama penyakit diare dan pneumonia. Tetapi jumlah fasilitas pelayanan kesehatan juga cukup tinggi.
5	Cluster 2 (Kota Tangerang Selatan)	Tingkat urgensi kelima atau rendah karena Kabupaten/Kota tersebut memiliki jumlah penderita penyakit menular cukup rendah, kecuali penyakit tuberkulosis. Tetapi sebanding dengan jumlah fasilitas pelayanan kesehatan yang tinggi.

Tingkat urgensi kesehatan pada **Tabel 5.9** didukung oleh AHH (Angka Harapan Hidup) dan IPM (Indeks Pembangunan Manusia) masing-masing Kabupaten/Kota di Provinsi Banten.

## 5. KESIMPULAN

Berdasarkan hasil analisis dan pembahasan, diperoleh kesimpulan bahwa:

1. Hasil perbandingan nilai indeks validitas *cluster Partition Entrophy* (PE), *Partition Coefficient* (PC) dan *Modified Partition Coefficient* (MPC) pada metode *Fuzzy C-Means* (FCM) dan *Fuzzy Possibilistic C-Means* (FPCM) diperoleh bahwa metode *Fuzzy C-Means* (FCM) merupakan metode yang terbaik dalam studi kasus ini.
2. Hasil profilisasi pada pengelompokan Kabupaten/Kota di Provinsi Banten berdasarkan fasilitas pelayanan kesehatan dan penderita penyakit menular adalah *cluster 1* merupakan Kabupaten/Kota yang memiliki tingkat urgensi kesehatan tertinggi di Provinsi Banten. Hal tersebut dikarenakan adanya ketimpangan antara jumlah fasilitas pelayanan kesehatan dengan jumlah penderita penyakit menular.

## 6. DAFTAR PUSTAKA

- Correa, C., Barreiro, P., Valero, C., Diago, M.-P., & Tardaguila, J. (2014). A Comparison of Fuzzy Clustering Algorithms Applied To Feature Extraction on Vineyard.
- Dinaskesehatan. (2019). *Profil Kesehatan Provinsi Banten*. Diakses dari <https://dinkes.bantenprov.go.id/id/archive/profil-kesehatan-provinsi-bant/1.html>
- Handoyo, S., Widodo, A., Nugroho, W. H., & Purwanto, I. N. (2019). The Implementation of A Hybrid Fuzzy Clustering on The Public Health Facility Data. *International Journal of Advanced Trends in Computer Science and Engineering*.
- Kementeriankesehatan. (2020). *Profil Kesehatan Indonesia 2019*. Jakarta: Kementerian Kesehatan Republik Indonesia. Diakses dari <https://pusdatin.kemkes.go.id/resources/download/pusdatin/profil-kesehatan-indonesia/Profil-Kesehatan-Indonesia-2019.pdf>
- Kementeriankesehatan. (2020). *Rencana Strategis Kementerian Kesehatan Tahun 2020-2024*.
- Puspitasari, N., Widians, J. A., & Pohny. (2018). A Clustering Of Generative and Infectious Diseases Using Fuzzy C-Means. *Jurnal Ilmiah Teknologi dan Informasi*.
- Sanusi, W., Zaky, A., & Afni, B. N. (2018). Analisis Fuzzy C-Means dan Penerapannya Dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Faktor-Faktor Penyebab Gizi Buruk.



# LAMPIRAN

**Lampiran 1** Data Kesehatan di Provinsi Banten

<b>Jumlah Penderita Penyakit dan Pelayanan Kesehatan</b>					
<b>Kabupaten/Kota</b>	<b>Puskesmas</b>	<b>Rumah Sakit</b>	<b>Diare</b>	<b>Tuberkulosis</b>	<b>Pneumonia</b>
Kabupaten Lebak	42	2	31,674	2,833	3,092
Kabupaten Pandeglang	36	1	32,915	3,790	1,457
Kabupaten Serang	31	3	44,829	2,855	3,627
Kabupaten Tangerang	44	25	70,107	6,701	12,842
Kota Tangerang	36	29	38,102	1,604	6,655
Kota Cilegon	8	4	8,703	2,552	611
Kota Serang	16	12	9,994	1,644	1,404
Kota Tangerang Selatan	29	29	22,318	6,737	6,517



## Lampiran 2 Program R

```
## Fuzzy C-Means and Fuzzy Possibilistics C-Means##

#Mengaktifkan package
library(ppclust)
library(factoextra)
library(cluster)
library(fclust)
library(ggplot2)
library(ggpubr)

#Menginput Data
kes <- read.delim("clipboard")
kes

w=kes[,-1]
w
x <- scale(w)
x

#Statistika Deskriptif
summary(w)
sd(w$Puskesmas)
sd(w$Rumah.Sakit)
sd(w$Diare)
sd(w$Tuberkulosis)
sd(w$Pneumonia)

par(mfrow=c(1,2))
boxplot(w$Puskesmas, main="Jumlah Puskesmas")
boxplot(w$Rumah.Sakit, main="Jumlah Rumah Sakit")
par(mfrow=c(1,3))
boxplot(w$Diare, main="Jumlah Penderita Diare")
boxplot(w$Tuberkulosis, main="Jumlah Penderita Tuberkulosis")
boxplot(w$Pneumonia, main="Jumlah Penderita Pneumonia")

##FCM
#Mencari Jumlah Cluster yang optimal
kes.fcm <- fcm(x, centers=2)
kes.fcm$iter
kes.fcm$func.val
kes.fcm.a <- ppclust2(kes.fcm, "fclust")
idxpe <- PE(kes.fcm.a$U)
cat("Partition Entropy: ", idxpe)
idxpc <- PC(kes.fcm.a$U)
cat("Partition Coefficient: ", idxpc)
idxmpc <- MPC(kes.fcm.a$U)
cat("Modified Partition Coefficient: ", idxmpc)
#Hasil Cluster
kes.fcm$cluster

##FPCM
#Mencari Jumlah Cluster yang optimal
kes.fpcm <- pcm(x, centers=kes.fcm$v, memberships=kes.fcm$u)
```



```
kes.fpcm$iter
kes.fpcm$func.val
kes.fpcm.a <- ppclust2(kes.fpcm, "fclust")
idxpe <- PE(kes.fpcm.a$U)
cat("Partition Entropy: ", idxpe)
idxpc <- PC(kes.fpcm.a$U)
cat("Partition Coefficient: ", idxpc)
idxmpc <- MPC(kes.fpcm.a$U)
cat("Modified Partition Coefficient: ", idxmpc)
#Hasil Cluster
kes.fpcm$cluster
```



### Lampiran 3 Hasil *Output Fuzzy C-Means* (FCM)

#### 1. 2 Cluster

```
> ##FCM
> kes.fcm <- fcm(x, centers=2)

> #Mencari Jumlah Cluster yang optimal
> kes.fcm$iter
[1] 48
> kes.fcm$func.val
[1] 14.88332
> kes.fcm.a <- ppclust2(kes.fcm, "fclust")
> idxpe <- PE(kes.fcm.a$U)
> cat("Partition Entropy: ", idxpe)
Partition Entropy: 0.4553109
> idxpc <- PC(kes.fcm.a$U)
> cat("Partition Coefficient: ", idxpc)
Partition Coefficient: 0.7145121
> idxmpc <- MPC(kes.fcm.a$U)
> cat("Modified Partition Coefficient: ", idxmpc)
Modified Partition Coefficient: 0.4290242

> #Hasil Cluster
> kes.fcm$cluster
1 2 3 4 5 6 7 8
1 1 1 2 2 1 1 2
```

#### 2. 3 Cluster

```
> ##FCM
> kes.fcm <- fcm(x, centers=3)

> #Mencari Jumlah Cluster yang optimal
> kes.fcm$iter
[1] 56
> kes.fcm$func.val
[1] 7.046465
> kes.fcm.a <- ppclust2(kes.fcm, "fclust")
> idxpe <- PE(kes.fcm.a$U)
> cat("Partition Entropy: ", idxpe)
Partition Entropy: 0.4865279
> idxpc <- PC(kes.fcm.a$U)
> cat("Partition Coefficient: ", idxpc)
Partition Coefficient: 0.7392628
> idxmpc <- MPC(kes.fcm.a$U)
> cat("Modified Partition Coefficient: ", idxmpc)
Modified Partition Coefficient: 0.6088942

> #Hasil Cluster
> kes.fcm$cluster
1 2 3 4 5 6 7 8
3 3 3 2 3 1 1 2
```

### 3. 4 Cluster

```
> ##FCM
> kes.fcm <- fcm(x, centers=4)
.
> #Mencari Jumlah Cluster yang optimal
> kes.fcm$iter
[1] 38
> kes.fcm$func.val
[1] 3.510449
> kes.fcm.a <- ppclust2(kes.fcm, "fclust")
> idxpe <- PE(kes.fcm.a$U)
> cat("Partition Entropy: ", idxpe)
Partition Entropy: 0.4560306
> idxpc <- PC(kes.fcm.a$U)
> cat("Partition Coefficient: ", idxpc)
Partition Coefficient: 0.7850909
> idxmpc <- MPC(kes.fcm.a$U)
> cat("Modified Partition Coefficient: ", idxmpc)
Modified Partition Coefficient: 0.7134546
.
> #Hasil Cluster
> kes.fcm$cluster
1 2 3 4 5 6 7 8
1 1 1 4 3 2 2 3
```

### 4. 5 Cluster

```
> ##FCM
> kes.fcm <- fcm(x, centers=5)
.
> #Mencari Jumlah Cluster yang optimal
> kes.fcm$iter
[1] 25
> kes.fcm$func.val
[1] 1.301461
> kes.fcm.a <- ppclust2(kes.fcm, "fclust")
> idxpe <- PE(kes.fcm.a$U)
> cat("Partition Entropy: ", idxpe)
Partition Entropy: 0.312383
> idxpc <- PC(kes.fcm.a$U)
> cat("Partition Coefficient: ", idxpc)
Partition Coefficient: 0.8677812
> idxmpc <- MPC(kes.fcm.a$U)
> cat("Modified Partition Coefficient: ", idxmpc)
Modified Partition Coefficient: 0.8347265
.
> #Hasil Cluster
> kes.fcm$cluster
1 2 3 4 5 6 7 8
1 1 1 4 5 3 3 2
```

## 5. 6 Cluster

```
> ##FCM
> kes.fcm <- fcm(x, centers=6)

> #Mencari Jumlah Cluster yang optimal
> kes.fcm$iter
[1] 64
> kes.fcm$func.val
[1] 0.7187028
> kes.fcm.a <- ppclust2(kes.fcm, "fclust")
> idxpe <- PE(kes.fcm.a$U)
> cat("Partition Entropy: ", idxpe)
Partition Entropy: 0.3022829
> idxpc <- PC(kes.fcm.a$U)
> cat("Partition Coefficient: ", idxpc)
Partition Coefficient: 0.8667742
> idxmpc <- MPC(kes.fcm.a$U)
> cat("Modified Partition Coefficient: ", idxmpc)
Modified Partition Coefficient: 0.840129

> #Hasil Cluster
> kes.fcm$cluster
1 2 3 4 5 6 7 8
1 1 5 4 3 6 6 2
```

## 6. 7 Cluster

```
> ##FCM
> kes.fcm <- fcm(x, centers=7)

> #Mencari Jumlah Cluster yang optimal
> kes.fcm$iter
[1] 46
> kes.fcm$func.val
[1] 0.4474671
> kes.fcm.a <- ppclust2(kes.fcm, "fclust")
> idxpe <- PE(kes.fcm.a$U)
> cat("Partition Entropy: ", idxpe)
Partition Entropy: 0.1686913
> idxpc <- PC(kes.fcm.a$U)
> cat("Partition Coefficient: ", idxpc)
Partition Coefficient: 0.931793
> idxmpc <- MPC(kes.fcm.a$U)
> cat("Modified Partition Coefficient: ", idxmpc)
Modified Partition Coefficient: 0.9204252

> #Hasil Cluster
> kes.fcm$cluster
1 2 3 4 5 6 7 8
5 4 7 3 6 1 1 2
```

## Lampiran 4 Hasil Output Fuzzy Possibilistic C-Means (FPCM)

### 1. 2 Cluster

```
> ##FPCM
> kes.fpcm <- pcm(x, centers=kes.fcm$v, memberships=kes.fcm$u)
.
> #Mencari Jumlah Cluster yang optimal
> kes.fpcm$iter
[1] 25
> kes.fpcm$func.val
[1] 22.18462
> kes.fpcm.a <- ppclust2(kes.fpcm, "fclust")
> idxpe <- PE(kes.fpcm.a$U)
> cat("Partition Entropy: ", idxpe)
Partition Entropy: 0.518594
> idxpc <- PC(kes.fpcm.a$U)
> cat("Partition Coefficient: ", idxpc)
Partition Coefficient: 0.6839781
> idxmpc <- MPC(kes.fpcm.a$U)
> cat("Modified Partition Coefficient: ", idxmpc)
Modified Partition Coefficient: 0.3679562
.
> #Hasil Cluster
> kes.fpcm$cluster
1 2 3 4 5 6 7 8
2 2 2 2 2 2 2 2
```

### 2. 3 Cluster

```
> ##FPCM
> kes.fpcm <- pcm(x, centers=kes.fcm$v, memberships=kes.fcm$u)
.
> #Mencari Jumlah Cluster yang optimal
> kes.fpcm$iter
[1] 55
> kes.fpcm$func.val
[1] 22.69952
> kes.fpcm.a <- ppclust2(kes.fpcm, "fclust")
> idxpe <- PE(kes.fpcm.a$U)
> cat("Partition Entropy: ", idxpe)
Partition Entropy: 0.6850202
> idxpc <- PC(kes.fpcm.a$U)
> cat("Partition Coefficient: ", idxpc)
Partition Coefficient: 0.513175
> idxmpc <- MPC(kes.fpcm.a$U)
> cat("Modified Partition Coefficient: ", idxmpc)
Modified Partition Coefficient: 0.2697625
.
> #Hasil Cluster
> kes.fpcm$cluster
1 2 3 4 5 6 7 8
3 3 3 2 2 1 1 2
```

### 3. 4 Cluster

```
> ##FPCM
> kes.fpcm <- pcm(x, centers=kes.fcm$v, memberships=kes.fcm$u)

> #Mencari Jumlah Cluster yang optimal
> kes.fpcm$iter
[1] 55
> kes.fpcm$func.val
[1] 14.79498
> kes.fpcm.a <- ppclust2(kes.fpcm, "fclust")
> idxpe <- PE(kes.fpcm.a$U)
> cat("Partition Entropy: ", idxpe)
Partition Entropy: 0.6749489
> idxpc <- PC(kes.fpcm.a$U)
> cat("Partition Coefficient: ", idxpc)
Partition Coefficient: 0.552305
> idxmpc <- MPC(kes.fpcm.a$U)
> cat("Modified Partition Coefficient: ", idxmpc)
Modified Partition Coefficient: 0.4030734
.

> #Hasil Cluster
> kes.fpcm$cluster
1 2 3 4 5 6 7 8
1 1 1 4 3 2 2 3
```

### 4. 5 Cluster

```
> ##FPCM
> kes.fpcm <- pcm(x, centers=kes.fcm$v, memberships=kes.fcm$u)

> #Mencari Jumlah Cluster yang optimal
> kes.fpcm$iter
[1] 44
> kes.fpcm$func.val
[1] 4.617783
> kes.fpcm.a <- ppclust2(kes.fpcm, "fclust")
> idxpe <- PE(kes.fpcm.a$U)
> cat("Partition Entropy: ", idxpe)
Partition Entropy: 0.3461454
> idxpc <- PC(kes.fpcm.a$U)
> cat("Partition Coefficient: ", idxpc)
Partition Coefficient: 0.643257
> idxmpc <- MPC(kes.fpcm.a$U)
> cat("Modified Partition Coefficient: ", idxmpc)
Modified Partition Coefficient: 0.5540713

> #Hasil Cluster
> kes.fpcm$cluster
1 2 3 4 5 6 7 8
1 1 1 4 5 3 3 2
```

## 5. 6 Cluster

```
> ##FPCM
> kes.fpcm <- pcm(x, centers=kes.fcm$v, memberships=kes.fcm$u)

> #Mencari Jumlah Cluster yang optimal
> kes.fpcm$iter
[1] 20
> kes.fpcm$func.val
[1] 3.420215
> kes.fpcm.a <- ppclust2(kes.fpcm, "fclust")
> idxpe <- PE(kes.fpcm.a$U)
> cat("Partition Entropy: ", idxpe)
Partition Entropy: 0.3076887
> idxpc <- PC(kes.fpcm.a$U)
> cat("Partition Coefficient: ", idxpc)
Partition Coefficient: 0.7590546
> idxmpc <- MPC(kes.fpcm.a$U)
> cat("Modified Partition Coefficient: ", idxmpc)
Modified Partition Coefficient: 0.7108656

> #Hasil Cluster
> kes.fpcm$cluster
1 2 3 4 5 6 7 8
1 1 5 4 3 6 6 2
```

## 6. 7 Cluster

```
> ##FPCM
> kes.fpcm <- pcm(x, centers=kes.fcm$v, memberships=kes.fcm$u)

> #Mencari Jumlah Cluster yang optimal
> kes.fpcm$iter
[1] 17
> kes.fpcm$func.val
[1] 2.224091
> kes.fpcm.a <- ppclust2(kes.fpcm, "fclust")
> idxpe <- PE(kes.fpcm.a$U)
> cat("Partition Entropy: ", idxpe)
Partition Entropy: 0.1992769
> idxpc <- PC(kes.fpcm.a$U)
> cat("Partition Coefficient: ", idxpc)
Partition Coefficient: 0.8788871
> idxmpc <- MPC(kes.fpcm.a$U)
> cat("Modified Partition Coefficient: ", idxmpc)
Modified Partition Coefficient: 0.8587017

> #Hasil Cluster
> kes.fpcm$cluster
1 2 3 4 5 6 7 8
5 4 7 3 6 1 1 2
```