

## ANALISIS KLASTER BERDASARKAN TINDAKAN KRIMINALITAS DI INDONESIA TAHUN 2019

Margareth Dwiyanti Simatupang<sup>1</sup>, Arie Wahyu Wijayanto<sup>2</sup>

<sup>1,2</sup> Politeknik Statistika STIS, Jl. Otto Iskandardinata No.64c Jakarta Timur  
Email: [1mds1708@gmail.com](mailto:1mds1708@gmail.com), [2ariewahyu@stis.ac.id](mailto:2ariewahyu@stis.ac.id)

### Abstract

*Crime is a problem that often occurs in a society. Currently, the Crime Index in Indonesia is 46.26 out of a scale of 100, so that Indonesia is in fourth place with the highest crime index in Asean countries. Although the number of crimes in Indonesia has decreased from 2017 to 2019, the decline in the number of crimes in Indonesia has slowed down in the past year. So it is necessary to group the criminal-prone areas in Indonesia so that they can provide information to the government and police to improve security in Indonesia. This study uses a variable of types of crime. The method used in this study is the K-Means and Fuzzy C-Means cluster analysis. Prior to grouping, the optimum number of clusters was determined. After that, validation of the method to be used between the K-Means and Fuzzy C-Means was carried out to obtain the best method. Validation is used by looking at connectivity, dunn index, and silhouette of each method. The results obtained are that there is no clustering algorithm that can be used universally to solve all problems regarding the grouping of criminal areas in Indonesia. Thus, both k-means and fuzzy c-means are able to group criminal areas in Indonesia.*

*Keywords: Crime, K-Means, Fuzzy C-Means, Connectivity, Dunn Index, and Silhouette.*

### Abstrak

Kriminalitas atau kejahatan merupakan masalah yang sering terjadi dalam suatu masyarakat. Saat ini indeks kejahatan di Indonesia sebesar 46.26 dari skala 100 sehingga Indonesia berada pada urutan keempat dengan indeks kejahatan tertinggi di negara Asean. Meskipun jumlah kejahatan di Indonesia mengalami penurunan dari tahun 2017 – 2019, namun penurunan jumlah kejahatan di Indonesia melambat dalam satu tahun terakhir. Sehingga perlu dilakukan pengelompokan daerah rawan kriminalitas yang ada di Indonesia agar dapat memberikan informasi kepada pemerintah dan kepolisian untuk meningkatkan keamanan di Indonesia. Penelitian ini menggunakan variabel jenis-jenis kejahatan. Metode yang digunakan dalam penelitian ini adalah dengan analisis kluster *K-Means* dan *Fuzzy C-Means*. Sebelum dilakukan pengelompokan, dilakukan penentuan jumlah kluster optimum. Setelah itu, dilakukan validasi metode yang hendak digunakan diantara *K-Means* dan *Fuzzy C-Means* untuk memperoleh metode yang terbaik. Validasi digunakan dengan melihat *connectivity*, *dunn index*, dan *silhouette* masing-masing metode. Hasil yang diperoleh yakni tidak ada algoritma klustering yang bisa digunakan secara universal untuk menyelesaikan seluruh permasalahan mengenai pengelompokan daerah kriminalitas di Indonesia. Sehingga, baik *k-means* dan *fuzzy c-means* tetap dapat melakukan pengelompokan daerah kriminalitas di Indonesia.

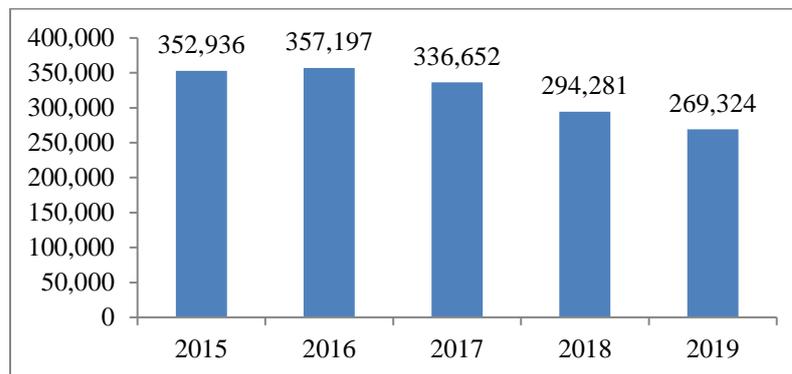
Kata kunci: Kriminalitas, *K-Means*, *Fuzzy C-Means*, *Connectivity*, *Dunn Index*, dan *Silhouette*.

## 1. Pendahuluan

Kriminalitas atau kejahatan merupakan masalah yang sering terjadi dalam suatu masyarakat. Kriminalitas merupakan suatu perbuatan yang dapat menimbulkan masalah-masalah dan keresahan bagi kehidupan didalam masyarakat [1]. Kejahatan memiliki dua macam pengertiannya yaitu secara yuridis dan secara sosiologi [14]. Secara yuridis formal, kejahatan adalah tingkah laku kejahatan yang melanggar hukum pidana yang ada. Pengertian secara sosiologi adalah meliputi segala tingkah laku manusia, walaupun tidak atau belumnya ditentukan dengan undang-undang [4].

Salah satu tujuan dari *Sustainable Development Goals* (SDGs) ialah Perdamaian, Keadilan dan Kelembagaan yang Tangguh. Tujuan ini mencakup target berupa mengurangi tindakan kejahatan, seperti tindakan kekerasan, eksploitasi, korupsi, terorisme, dan lainnya secara signifikan pada tahun 2030 [16].

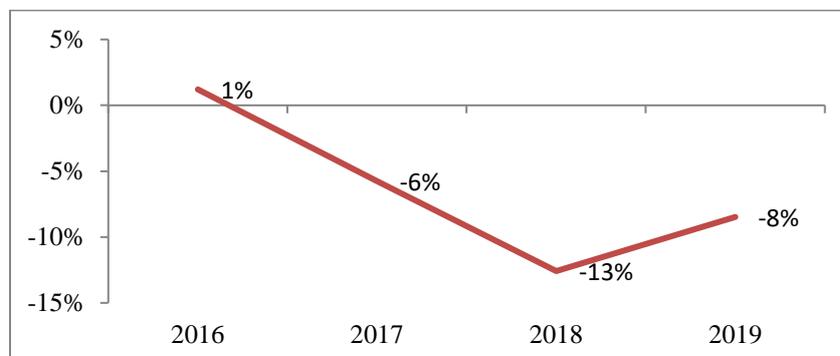
Menurut data yang diperoleh dari *Numbeo* [7], pada pertengahan tahun 2019 indeks kejahatan Indonesia sebesar 46.26 dari skala 100. Indeks ini memiliki makna, semakin mendekati angka 100 maka kejahatan di negara tersebut semakin tinggi. Indeks kejahatan di Indonesia berada di urutan ke-empat tertinggi di ASEAN. Indeks kejahatan di Indonesia lebih tinggi dibandingkan negara Asia Tenggara lainnya, seperti Filipina (41.09), Thailand (41.29), Brunei Darussalam (28.05), dan Singapura (27.70). Pada “Gambar 1”, berdasarkan data BPS yang diperoleh dari publikasi Statistik Kriminalitas Indonesia [5,6], pada tahun 2015 – 2019, jumlah kejadian kejahatan atau kriminalitas di Indonesia cenderung menurun. Pada tahun 2015, jumlah kejadian kejahatan sebanyak 352.936, meningkat menjadi 357.197 pada tahun 2016. Kemudian, tiga tahun selanjutnya mengalami penurunan setiap tahun sampai pada tahun 2019 jumlah kejahatan di Indonesia sebanyak 269.324.



Gambar 1. Jumlah Tindakan Kejahatan di Indonesia Tahun 2015-2016

Sumber : BPS, diolah

Meskipun jumlah kejahatan di Indonesia mengalami penurunan dari tahun 2017 – 2019, namun dapat dilihat penurunan jumlah kejahatan di Indonesia melambat dalam satu tahun terakhir. Pada Gambar 2, dapat dilihat pada tahun 2017 terjadi penurunan jumlah kejahatan sebesar 6%, kemudian mengalami penurunan sebesar 13%, dan pada tahun 2019 hanya mengalami penurunan sebesar 8%. Hal ini dapat dilihat pada “Gambar 2”



Gambar 2. Persentase Penurunan dan Kenaikan Jumlah Tindakan Kejahatan Tahun 2016-2019

Sumber : BPS, diolah

Pada penelitian sebelumnya telah dilakukan pengelompokan kriminalitas di beberapa negara dan wilayah dengan metode klastering yang berbeda-beda. Al-Janabi [3] mengusulkan sebuah kerangka analisis kejahatan dan kriminitas menggunakan algoritma *decision tree* untuk klasifikasi dan algoritma *k-means* untuk klastering data. Penelitian ini dilakukan untuk menemukan pola dan tren, memprediksi, dan memetakan jaringan kriminalitas dengan menggunakan data kepolisian Austin, Texas. Sivanagaleela dan Rajesh [13] melakukan penelitian untuk mengidentifikasi jenis-jenis kejahatan dengan menggunakan *fuzzy c-means* di India. Penelitian ini dapat memberikan hasil berupa identifikasi daerah-daerah rawan kejahatan dan jenis kejahatan yang sering terjadi di tempat tersebut. Hal ini berguna bagi pihak penyidik untuk memecahkan kasus kejahatan yang dapat diidentifikasi dalam waktu yang lebih sedikit. Suriani [15] juga melakukan analisis pengelompokan daerah rawan tindak kriminal yang ada di Sumatera Utara dengan *k-means clustering*. Hal ini dilakukan untuk membantu pihak kepolisian menentukan tingkat daerah yang rawan kriminal.

Dengan banyaknya kasus kriminalitas di Indonesia, maka perlu dilakukan pengelompokan daerah rawan kriminalitas. Peneliti membatasi metode yang digunakan dengan *k-means* dan *fuzzy c-means* untuk menemukan metode klastering yang terbaik. Tujuan dilakukannya penelitian ini adalah memberikan informasi kepada kebijakan pemerintah dan kepolisian yang ada di Indonesia untuk meningkatkan keamanan dan mejadi bahan evaluasi bagi setiap provinsi yang ada di Indonesia dalam menurunkan kejahatan sesuai dengan hasil pengelompokan berdasarkan metode klastering yang terbaik.

## 2. Metode

Sumber data yang digunakan berasal dari Publikasi Statistik Kriminal Indonesia tahun 2020 [6]. Variabel yang digunakan adalah jenis-jenis kejahatan yang dapat dilihat pada "Tabel 1"

No	Variabel
1	kejahatan terhadap nyawa
2	kejahatan terhadap fisik atau badan
3	kejahatan terhadap kesusilaan
4	kejahatan terhadap kemerdekaan
5	kejahatan terhadap hak milik atau barang tanpa kekerasan
6	kejahatan terhadap hak milik atau barang dengan kekerasan
7	kejahatan terkait narkoba
8	kejahatan terkait penipuan, penggelapan, dan korupsi
9	kejahatan terhadap ketertiban umum

### 2.1 Penentuan Jumlah Klaster Optimum

Metode analisis data yang digunakan adalah analisis *K-Means Clustering* dan *Fuzzy C-Means*. Analisis ini dapat digunakan untuk mengelompokkan daerah rawan kriminalitas yang ada di Indonesia berdasarkan kejahatan yang terjadi di Indonesia pada tahun 2019.

Penentuan jumlah klaster dilakukan dengan metode *Elbow*. Metode *Elbow* digunakan untuk menghasilkan informasi dalam menentukan jumlah klaster terbaik dengan cara melihat persentase hasil perbandingan antara jumlah klaster yang akan membentuk suatu titik [10]. Metode *Elbow* memberikan suatu gagasan berupa memilih nilai klaster dan kemudian menambah nilai klaster tersebut untuk dijadikan sebagai model dalam penentuan klaster terbaik. Persentase perhitungan yang dihasilkan menjadi pembanding antara jumlah klaster yang ditambah. Hasil persentase yang berbeda dari setiap nilai klaster dapat ditunjukkan dengan menggunakan grafik sebagai sumber informasinya. Jika nilai klaster pertama dengan nilai klaster kedua memberikan sudut dalam grafik atau nilainya mengalami penurunan paling besar maka nilai cluster tersebut yang terbaik.[11]

Perbandingannya dilakukan dengan menghitung SSE (*Sum of Square Error*) dari masing-masing nilai kluster. Karena semakin besar jumlah kluster  $K$ , maka nilai SSE akan semakin kecil. Rumus SSE pada *k-means* ada pada “Persamaan (1)” :

$$SSE = \sum_{K=1}^K \sum_{X_i \in S_k} \|X_i - C_k\|_2^2 \quad (1)$$

## 2.2 Algoritma *K-Means*

Langkah-langkah algoritma *K-Means* adalah sebagai berikut [8] :

1. Menentukan jumlah kluster yang ingin dibentuk
2. Mendistribusikan data secara acak ke dalam kluster
3. Menentukan pusat kluster (*centroid*) dari setiap data yang ada kluster seperti pada “Persamaan (2)” :

$$C_{kj} = \frac{x_{1j} + x_{2j} + \dots + x_{nj}}{n} \quad (2)$$

Keterangan :

$C_{kj}$  = pusat kluster ke- $k$  pada variabel ke- $j$  ( $j = 1, 2, \dots, p$ )

$n$  = banyak data pada kluster ke- $k$

4. Menentukan jarak setiap objek ke setiap *centroid* dengan menghitung jarak setiap objek dengan setiap *centroid* menggunakan *Euclidean distance* seperti pada “Persamaan (3)”

$$d(X_i, X_g) = \sqrt{\sum_{j=1}^p (X_{ij} - X_{gj})^2} \quad (3)$$

5. Menghitung fungsi objektif dengan formula seperti pada “Persamaan (4)” :

$$J = \sum_{i=1}^n \sum_{j=1}^k a_{ij} d(x_i, C_{kj})^2 \quad (4)$$

6. Mendistribusikan setiap data ke *centroid*/rata-rata terdekat yang dirumuskan seperti “Persamaan (7)” :

$$a_{ij} = \begin{cases} 1, & s = \min \{d(x_i, C_{kj})\} \\ 0, & \text{lainnya} \end{cases} \quad (5)$$

Keterangan :

$a_{ij}$  adalah nilai keanggotaan titik  $x_i$  ke pusat kluster  $C_{kj}$

$s$  adalah jarak terpendek dari data  $x_i$  ke pusat kluster  $C_{kj}$  setelah dibandingkan

7. Mengulangi langkah 3-6 sampai tidak ada lagi perpindahan objek atau tidak ada perubahan pada fungsi objektifnya.

## 2.3 Algoritma *Fuzzy C-Means*

Langkah-langkah algoritma *Fuzzy C-Means* adalah sebagai berikut [9] :

1. Meng-*input* data yang akan dikelompokkan ( $X$ ) dalam matriks berukuran  $n \times p$ . Dimana  $n$  adalah jumlah sampel data dan  $p$  adalah variabel setiap data.

2. Menentukan jumlah kluster ( $c$ ), pangkat pembobot ( $m > 1$ ), maksimum iterasi ( $MaxIter$ ), error terkecil yang diharapkan ( $\epsilon$ ), dan fungsi objektif awal ( $P_0 = 0$ ).
3. Membangkitkan bilangan random  $U_{ik}$ , dimana  $i = 1, 2, 3, \dots, n$  dan  $k=1, 2, \dots, c$
4. Menghitung pusat kluster ke- $k$  pada variabel ke- $j$ , yakni  $v_{kj}$ , dengan  $k = 1, 2, \dots, c$  dan  $j = 1, 2, \dots, p$
5. Menghitung nilai fungsi objektif pada iterasi ke- $t$  ( $P_t$ ) dengan formula pada “Persamaan (8)” :

$$P_t = \sum_{i=1}^n \sum_{k=1}^c (u_{ik})^m d_{ik}(x_i, v_{kj}) \quad (6)$$

6. Menghitung perubahan matriks keanggotaan  $u_{ik}$  dengan rumus yang ada pada “Persamaan (9)” :

$$u_{ik} = \left( \frac{\left( \sum_{j=1}^p d_{jk} \right)^{\frac{1}{m-1}}}{\sum_{k=1}^c \left( \sum_{j=1}^p d_{jk} \right)^{\frac{1}{m-1}}} \right)^{-1} \quad (7)$$

7. Memeriksa kondisi berhenti, yakni : apabila  $(|P_t - P_{t-1}|) < \epsilon$  atau  $(t > MaxIter)$  maka iterasi berhenti. Namun jika kondisi tersebut tidak terpenuhi, maka kembali ke langkah 4.

## 2.4 Validasi Model

### a. Connectivity index

*Connectivity index* berada pada rentang 0 sampai tak terhingga [12]. Nilai *connectivity* yang lebih rendah memberikan gambaran bahwa kluster tersebut memiliki kualitas yang lebih baik. Rumus *connectivity index* terdapat pada “Persamaan (8)”

$$conn = \sum_{i=1}^N \sum_{j=1}^L x_{i, nn_{i(j)}} \quad (8)$$

Keterangan :

$nn_{i(j)}$  adalah pengamatan *nearest neighbor*  $i$  ke  $j$  dan  $L$  dan  $nn_{i(j)}$  sebagai parameter yang menentukan jumlah tetangga yang berkontribusi pada pengukuran *connectivity*

### b. Dunn Index

Salah satu metode validasi kluster dari beberapa metode klustering adalah *Dunn Index*. *Dunn Index* merupakan fungsi validitas yang mampu memberikan penilaian yang efektif untuk pengaplikasian yang menggunakan beberapa metode klustering yang berbeda. [17]. Nilai *dunn index* berada pada rentang 0 sampai 1, yang memiliki arti semakin tinggi nilai *dunn Index* artinya semakin baik kluster yang terbentuk. Adapun rumus *dunn Index* adalah sebagai berikut [4]:

$$DU = \left\{ \frac{\min_{\substack{1 \leq i \leq k \\ i+1 \leq j \leq l}} (d(C_i, C_j))}{\max_{1 \leq l \leq q} d(C_l)} \right\} \quad (9)$$

dengan  $d(C_i, C_j) = \min_{i \in C_i, j \in C_j} d_{ij}$  dan  $d(C_l) = \max_{i, j \in C_l} d_{ij}$

Keterangan :

DU adalah *Dunn Index*

$q$  adalah jumlah kluster

$d(C_i, C_j)$  adalah jarak *squared euclidean* antar pasangan objek pada kluster  $i$  dan kluster  $j$  (*intercluster distance*)

$d(C_l)$  adalah jarak *square euclidean* antar anggota dalam kluster  $l$  (*intracluster distance*)

$d_{ij}$  adalah jarak *squared euclidean* antara objek  $i$  dan objek  $j$

**b. Koefisien Silhoutte**

Koefisien *silhoutte* dapat memberikan gambaran singkat mengenai seberapa baik objek pada klasternya. Nilai koefisien *silhoutte* setiap objek berada pada rentang -1 sampai dengan 1. Semakin besar nilai koefisien *silhoutte* akan semakin maik kualitas suatu kluster [2]. Rumus koefisien *silhoutte* terdapat pada “Persamaan (10)”

$$s(x_c) = \frac{1}{N} \sum_{i=1}^N s(x_i) \tag{10}$$

dan nilai koefisien *silhoutte* secara global terdapat pada “Persamaan (11)”

$$SC = \frac{1}{k} \sum_{i=1}^N S(x_c) \text{ dengan } s(x_i) = \frac{b_c - a_c}{\max[b_c, a_c]} \tag{11}$$

Keterangan :

$a_c$  adalah rata-rata kesamaan objek  $i$  dengan yang lainnya dalam satu kluster

$b_c$  adalah rata-rata kesamaan objek  $i$  dengan yang lainnya dalam setiap kluster

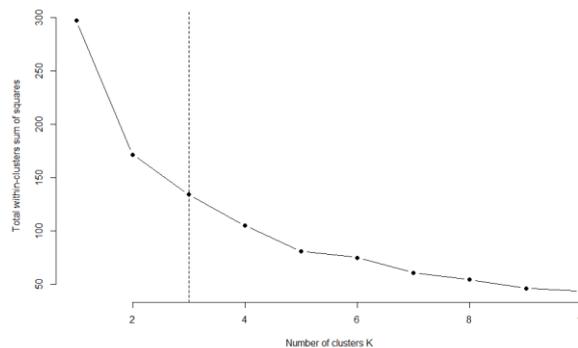
$s(x_i)$  adalah lebar silhoutte cluster ke- $i$

$N$  adalah banyak data

$K$  adalah banyaknya kluster

**3. Hasil dan Pembahasan**

Untuk mendapatkan hasil kluster yang optimum, maka perlu dilakukan validasi kluster terlebih dahulu. Penentuan jumlah kluster optimum dilakukan dengan metode *Elbow*. Berdasarkan “Gambar 3”, dapat dilihat bahwa jumlah kluster maksimum yang dapat dibentuk adalah tiga kluster. Hal ini dilihat dari penurunan *total within-clusters sum of squares* setelah  $k=3$  yang tidak membawa pengaruh banyak atas variabel di dalam kluster tersebut, sehingga dipilih jumlah kluster yang optimum sebanyak tiga kluster.

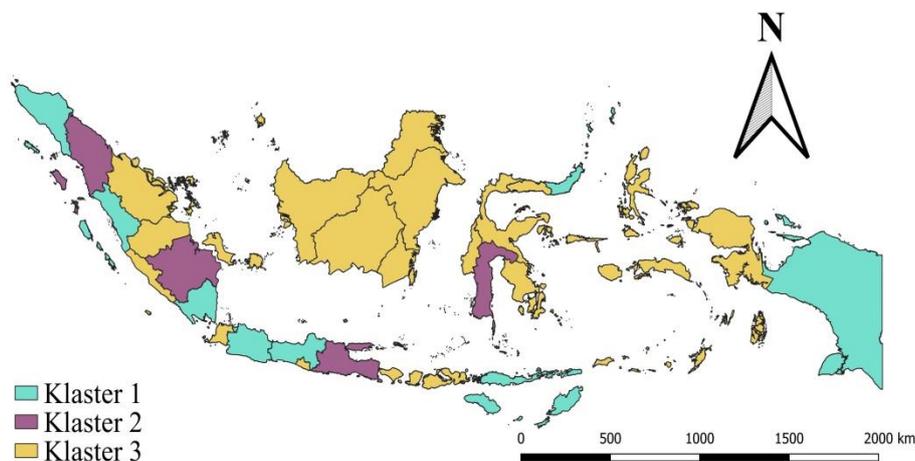


Gambar 3. *Optimal number of clusters*

Kemudian, dilakukan pengelompokan dengan metode *k-means* setiap provinsi ke dalam tiga kluster. Hasilnya diperoleh pada “Tabel 2” dan visualisasinya pada “Gambar 4” :

Tabel 2. Pengelompokan Tiap Provinsi dengan Metode *K-Means*

Kluster	Provinsi
1	Aceh, Sumatera Barat, Lampung, Metro Jaya, Jawa Barat, Jawa Tengah, NTT, Sulawesi Utara, Papua
2	Sumatera Utara, Sumatera Selatan, Jawa Timur, Sulawesi Selatan
3	Riau, Jambi, Bengkulu, Bangka Belitung, Kepulauan Riau, DI. Yogyakarta, Banten, Bali, NTB, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat



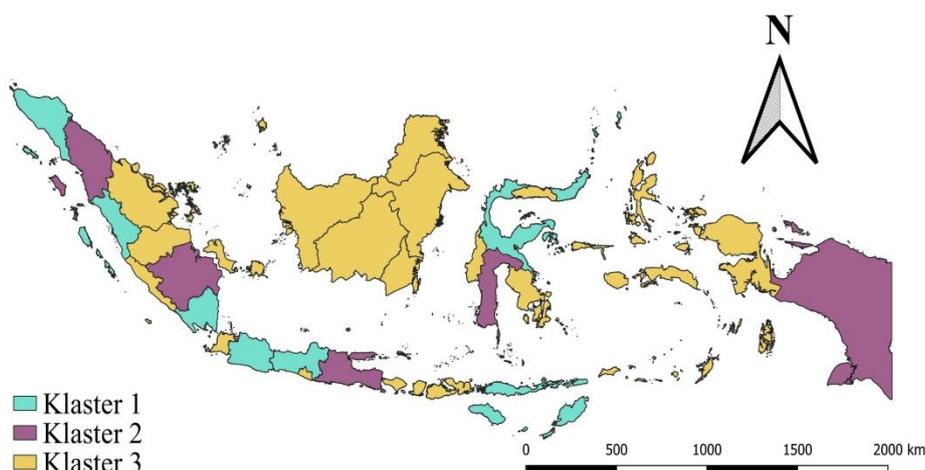
Gambar 4. Visualisasi Klastering Provinsi Berdasarkan Tindakan Kriminalitas Tahun 2019 dengan *K-means*

Kemudian, sebagai bentuk pembandingan, maka dilakukan pengelompokan dengan metode *fuzzy c-means* setiap provinsi ke dalam tiga kluster. Hasilnya diperoleh pada “Tabel 4” dan visualisasinya pada “Gambar 5” :

Tabel 4. Pengelompokan Tiap Provinsi dengan Metode *Fuzzy C-Means*

Kluster	Provinsi
1	Aceh, Sumatera Barat, Lampung, Jawa Barat, Jawa Tengah, NTT, Sulawesi Utara, Sulawesi Tengah,
2	Sumatera Utara, Sumatera Selatan, Metro Jaya, Jawa Timur, Sulawesi Selatan, Papua

3	Riau, Jambi, Bengkulu, Bangka Belitung, Kepulauan Riau, DI Yogyakarta, Banten, Bali, NTB, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat
---	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------



Gambar 5. Visualisasi Klastering Provinsi Berdasarkan Tindakan Kriminalitas Tahun 2019 dengan *Fuzzy C-Means*

Pengelompokkan dengan *k-means* dan *fuzzy c-means* memberikan hasil yang tidak terlalu berbeda. Hanya terdapat beberapa perbedaan pengelompokkan, yakni pada wilayah Metro Jaya, Sulawesi Tengah dan Papua Barat. Pada *k-means*, Metro Jaya dikelompokkan pada klaster 1 sedangkan dengan *fuzzy c-means* dikelompokkan pada klaster 2. Kemudian, wilayah Sulawesi Tengah dikelompokkan pada klaster 3 dengan metode *k-means* dan dikelompokkan pada klaster 1 dengan metode *fuzzy c-means*. Setelah itu, Papua Barat dikelompokkan pada klaster 1 dengan metode *k-means* dan dikelompokkan pada klaster 2 dengan metode *fuzzy c-means*.

Setelah itu, diperoleh validasi model klastering dengan *Connectivity*, *Dunn Index*, dan *Silhouette* seperti pada “Tabel 6” :

Tabel 6. Perbandingan Model Klastering

Metode	<i>Connectivity</i>	<i>Dunn Index</i>	<i>Silhouette</i>
<i>K-Means</i>	<b>15.5425</b>	0.1386	<b>0.4288</b>
<i>Fuzzy C-Means</i>	22.2333	<b>0.1507</b>	0.1955

Dari tabel diatas, dapat dilihat bahwa nilai *connectivity k-means* lebih rendah dibanding *fuzzy c-means*. Namun, apabila dilihat dari *dunn index*, *fuzzy c-means* memiliki nilai yang lebih tinggi dibanding *k-means*. Sedangkan jika dilihat dari *silhouette*, *k-means* memiliki nilai yang tinggi dibanding *fuzzy c-means*. Hal ini mengindikasikan bahwa pengelompokkan dengan *fuzzy c-means* memberikan hasil yang lebih baik apabila dilihat dari *dunn index* sedangkan *k-means* memberikan hasil yang lebih baik apabila dilihat dari *connectivity* dan *silhouette*.

#### 4. Kesimpulan

Berdasarkan hasil yang didapat dari proses dalam menentukan jumlah kluster terbaik dan metode kluster *k-means* serta *fuzzy c-means* dapat disimpulkan sebagai berikut :

- a. Penentuan jumlah kluster terbaik dengan metode *Elbow* dapat menghasilkan jumlah kluster sebanyak tiga. Hasil penentuan jumlah kluster terbaik dengan metode *Elbow* akan dijadikan *default* untuk proses karakteristik berdasarkan studi kasus yang dilakukan.
- b. Dengan melihat *dunn index*, maka metode klustering yang terbaik adalah *fuzzy c-means*. Namun, apabila dilihat dari *connectivity* dan *silhouette*, metode klustering yang terbaik adalah *k-means*.
- c. Oleh sebab itu, tidak ada algoritma klustering yang bisa digunakan secara universal untuk menyelesaikan seluruh permasalahan mengenai pengelompokan daerah kriminalitas di Indonesia. Sehingga, baik *k-means* dan *fuzzy c-means* tetap dapat melakukan pengelompokan daerah kriminalitas di Indonesia.
- d. Pemerintah dan pihak kepolisian dapat menangani tindak kejahatan yang ada di Indonesia berdasarkan pengelompokan yang dilakukan pada penelitian ini. Wilayah-wilayah yang berada pada kluster yang sama menunjukkan karakteristik jenis kejahatan yang mirip sehingga dapat mempermudah pemerintah dan kepolisian untuk menanggulangi ataupun mengurangi tindakan kejahatan di wilayah setiap kluster.
- e. Bagi peneliti selanjutnya dapat melakukan analisis kluster kriminalitas di Indonesia dengan menggunakan metode dan validasi model klustering yang lain serta dapat menambah variabel jenis-jenis kejahatan selain yang ada di penelitian ini.

### Ucapan Terima Kasih

Saya mengucapkan terima kasih kepada Tuhan Yesus yang terkasih, *reviewer*, orang tua, sahabat, dan para dosen jurusan statistika Politeknik Statistika STIS.

### Daftar Pustaka

- [1] Abdul Syani. 1987. Sosiologi Kriminalitas. Bandung: Remaja Rosda Karya.
- [2] Aini, F.N., Palgunadi. S., dan Anggrainingsih. R. 2014. Clustering Business Recess Model Petri Net dengan Complete Linkage. Jurnal ITSMART Vol 3. No. 2: Hal. 47-51.
- [3] Al-Janabi, K. B. S. (2011). A proposed framework for analyzing crime data set using decision tree and simple k-means mining algorithms. *Journal of Kufa for Mathematics and Computer*, 1(3), 8-24.
- [4] Ansari, Z., Babu, A.V., Azeem, M.F. dan Ahmed, W. 2011. *Quantitive Evaluation of Performance and Validity Indices for Clustering the Web Navigational Sessions*. World of [ Computer Science and Information Technology Journal (WCSIT) Vol. 1, No. 5: Hal. 217-226
- [5] BPS. (2016). Statistik Kriminal 2016. Jakarta: Badan Pusat Statistik.
- [6] BPS. (2020). Statistik Kriminal 2020. Jakarta: Badan Pusat Statistik.
- [7] Crime Index by Country 2020 Mid-Year. (n.d.). Retrieved Desember 12, 2020, from Numbeo: [https://www.numbeo.com/crime/rankings\\_by\\_country.jsp](https://www.numbeo.com/crime/rankings_by_country.jsp)
- [8] Eko Prasetyo (2012). "Data Mining Konsep dan Aplikasi Menggunakan Matlab". Yogyakarta. CV. Andi Offset
- [9] Kusumadewi, S., Purnomo, H., 2010, Aplikasi Logika Fuzzy untuk Pendukung Keputusan edisi 2, Graha Ilmu, Yogyakarta.
- [10] Madhulatha, T. S. (2012). An Overview on Clustering Methods. *IOSR Journal of Engineering*, 2(4)
- [11] Merliana, N. P. E., & Santoso, A. J. (2015). Analisa Penentuan Jumlah Cluster Terbaik pada Metode K-Means Clustering
- [12] S. M. Kim, M. I. Peña, M. Moll, G. Giannakopoulos, G. N. Bennett, and L. E. Kavraki, "An evaluation of different clustering methods and distance measures used for grouping metabolic

- pathways," in International Conference on Bioinformatics and Computational Biology, Kuala Lumpur, Malaysia, Feb. 2016, pp. 115-122.
- [13] Sivanagaleela, B., & Rajesh, S. (2019, April). Crime Analysis and Prediction Using Fuzzy C-Means Algorithm. In *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 595-599). IEEE.
- [14] Soesilo, R., 1985. *Kriminologi (Pengetahuan Tentang Sebab-Sebab Kejahatan)*, Bogor: Politeia.
- [15] Suriani, L. (2020). Pengelompokan Data Kriminal Pada Poldasu Menentukan Pola Daerah Rawan Tindak Kriminal Menggunakan Data Mining Algoritma K-Means Clustering. *Jurnal Sistem Komputer dan Informatika (JSON)*, 1(2), 151-157.
- [16] Sustainable Development Goals. (n.d.). Retrieved Desember 10, 2020, from Kementerian PPN/Bappenas: <http://sdgs.bappenas.go.id/tujuan-16/>
- [17] Uyun, S. dan Subanar. 2007. *Stock Data Clustering of Food and Beverage Company*. IJCCS Vol. 1, No. 2: Hal. 123-130