

## **PENGELOMPOKAN KABUPATEN/KOTA DI KALIMANTAN BARAT BERDASARKAN FAKTOR-FAKTOR KEMISKINAN MENGUNAKAN METODE DBSCAN**

**Purwanty, Shantika Martha, Siti Aprizkiyandari**

### **INTISARI**

*Salah satu teknik analisis cluster adalah Density Based Spatial Clustering Algorithm with Noise (DBSCAN). Teknik ini mengelompokkan data berdasarkan kepadatannya. Jumlah data atau minimum point (Minpts) dalam radius epsilon ( $\epsilon$ ) dari setiap kumpulan data adalah merupakan konsep kepadatan data yang dimaksud. Dalam penelitian ini, Kabupaten/kota di Kalimantan Barat dikelompokkan berdasarkan kriteria berdasarkan faktor-faktor kemiskinan tahun 2021 dengan menggunakan metode DBSCAN. Indeks pembangunan manusia ( $x_1$ ), tingkat pengangguran terbuka ( $x_2$ ), dan persentase penduduk usia 15 tahun ke atas yang melek huruf ( $x_3$ ) menjadi variabel yang digunakan dalam penelitian ini. Tahap pertama dalam penelitian ini adalah menggunakan z-score untuk standarisasi data. Selanjutnya dengan melihat kurva siku menggunakan plot K-Nearest Neighbor (KNN) untuk menentukan nilai epsilon dan Minpts. Selanjutnya menentukan titik centroid pertama yang dipilih secara acak. Selanjutnya gunakan jarak euclidean untuk menghitung jarak. Setelah itu pembentukan cluster berdasarkan jumlah titik dalam radius  $\epsilon$  dan lebih besar sama dengan Minpts. Kemudian validasi cluster dengan silhouette coefficient. Selanjutnya adalah interpretasi cluster. Dari hasil pengelompokan menggunakan metode DBSCAN diperoleh dua cluster dan lima noise dengan parameter  $\epsilon$  sebesar 0,6 dan Minpts sebesar dua. Tujuh Kabupaten yaitu Sambas, Bengkayang, Landak, Sanggau, Kapuas Hulu, Sekadau, dan Melawi termasuk dalam cluster satu. Dua Kabupaten yaitu Mempawah dan Kubu Raya termasuk dalam cluster dua. Dibandingkan dengan cluster satu, cluster dua memiliki nilai rata-rata IPM dan tingkat pengangguran terbuka yang lebih besar. Berdasarkan hasil tersebut maka program-program dari pemerintah dapat diarahkan pada cluster dua.*

**Kata Kunci:** Analisis Cluster, Epsilon, Minpts

### **PENDAHULUAN**

Permasalahan kemiskinan masih belum terselesaikan, terutama di negara-negara berkembang seperti Indonesia, dimana tingkat kemiskinan cukup tinggi dibandingkan dengan negara-negara tetangga lainnya [1]. Ketimpangan pembangunan menjadi penyebab kemiskinan di Indonesia, khususnya di wilayah pedesaan. Selain itu, berbagai faktor yang saling terkait, seperti lokasi geografis, kondisi iklim, pendapatan, kesehatan, pendidikan, dan akses terhadap produk dan layanan, juga berkontribusi terhadap masalah kemiskinan [2]. Kalimantan Barat merupakan salah satu Prov di Indonesia dengan tingkat kemiskinan yang relatif tinggi. Dengan jumlah penduduk miskin sebanyak 367,89 ribu jiwa atau 7,15% dari total penduduk, Provinsi Kalimantan Barat menempati peringkat kedua dari lima Provinsi di Pulau Kalimantan dalam hal kemiskinan pada tahun 2021 [3].

Kurangnya pendapatan atau anjloknya pendapatan merupakan salah satu variabel yang mempengaruhi angka kemiskinan. Rendahnya produktivitas kerja akibat rendahnya indeks pembangunan manusia (IPM) menjadi penyebab menurunnya pendapatan [4]. Persentase penduduk usia 15 tahun ke atas yang melek huruf dan tingkat pengangguran terbuka merupakan faktor lain yang mempengaruhi angka kemiskinan [5]. Artinya, agar upaya pengentasan kemiskinan menjadi lebih efisien, tepat sasaran, dan berhasil, diperlukan pengelompokan kemiskinan di Kabupaten/Kota berdasarkan variabel-variabel yang sudah ada. Di antara teknik pengelompokan yang tersedia adalah analisis *cluster*.

Dalam analisis statistik, analisis *cluster* adalah teknik multivariat yang mengelompokkan individu atau objek yang berkaitan menurut sejumlah karakteristik yang sebanding. Mengklasifikasikan sekumpulan individu atau objek yang terkait saja adalah tujuan dari analisis *cluster*. Ketidakteraturan

eksternal yang tinggi dan keragaman internal yang tinggi (di dalam *cluster*) merupakan prasyarat untuk analisis *cluster* [6].

*Density Based Spatial Clustering Algorithm with Noise* (DBSCAN) merupakan salah satu algoritma *cluster* yang digunakan untuk mengidentifikasi kelompok dalam suatu data dengan menggunakan kepadatan data. Metode DBSCAN digunakan untuk mengelompokkan data berdasarkan kepadatan data atau data yang saling terhubung. DBSCAN membentuk *cluster* data yang berdekatan satu sama lain sedangkan data yang berjauhan tidak dimasukkan [7]. Selain DBSCAN, pendekatan algoritma pengurutan lain yang memberikan hasil untuk pengelompokan datar atau hierarki disebut OPTICS (*Ordering Points to Identifikasi Struktur Clustering*). Algoritma ini dapat menjangkau objek yang mungkin diklasifikasikan sebagai *cluster* karena OPTICS menggunakan parameter yang mirip dengan DBSCAN ( $\epsilon$  dan Minpts) dan menerapkan konsep jarak jangkauan daripada memperhitungkan nilai epsilon [8]. Sebaliknya, DBSCAN memiliki keunggulan karena mampu mengidentifikasi outlier atau noise dan tidak akan memasukkan data noise ke dalam cluster [9]. Penelitian ini bertujuan untuk mengelompokkan Kabupaten/Kota di Kalimantan Barat berdasarkan faktor-faktor kemiskinan menggunakan metode DBSCAN agar program-program pengentasan kemiskinan dari pemerintah dapat lebih terarah dan tetap sasaran pada Kabupaten/Kota di Kalimantan Barat.

Pengelompokan menggunakan DBSCAN dimulai dengan menginput data faktor-faktor kemiskinan. Kemudian dilanjutkan melakukan eksplorasi data pada variabel yang digunakan untuk mengetahui karakteristik dari data. Selanjutnya melakukan standarisasi data dengan *z-score*. Kemudian menentukan nilai epsilon dan Minpts dengan plot *K-Nearest Neighbour* (KNN) dengan mengamati *knee of a curve* atau kurva siku. Setelah itu menentukan titik centroid awal yang dipilih secara acak dari data. Kemudian perhitungan jarak dengan menggunakan jarak *euclidean*. Setelah itu pembentukan *cluster* berdasarkan jumlah titik yang berada dalam radius  $\epsilon$  dan lebih besar sama dengan MinPts. Kemudian validasi *cluster* dengan *silhouette coefficient* untuk mendapatkan hasil *cluster* yang optimal. Langkah terakhir adalah melakukan interpretasi hasil *cluster* dengan menghitung nilai rata-rata tiap *cluster*.

### ALGORITMA K-NEAREST NEIGHBOUR (KNN)

Salah satu pendekatan klasifikasi adalah algoritma *K-Nearest Neighbor* (KNN). Untuk mengetahui hasil epsilon terbaik, metode KNN digunakan untuk menghitung epsilon pada data [10]. Algoritma *K-Nearest Neighbor* (KNN) dapat diselesaikan dengan langkah-langkah berikut :

1. Temukan parameter jumlah tetangga terdekat (K).
2. Gunakan perhitungan jarak Euclidean untuk mendapatkan jarak antar data.
3. Dari jarak terpendek, pilih K tetangga terdekat.
4. Tentukan nilai rata-rata K tetangga terdekat.
5. Buatlah representasi visual dari nilai rata-rata.
6. Perhatikan hasil grafik kurva siku-siku, terutama lokasi di mana kurva tampak membengkok, baik berlawanan arah maupun dari kemiringan tinggi ke kemiringan rendah.
7. Titik siku yang dihasilkan kemudian dijadikan epsilon.

### JARAK EUCLIDEAN

Jarak *euclidean* adalah metrik jarak yang paling sering digunakan untuk menghitung jarak antar observasi. Jarak antara titik inti dengan titik lainnya ditentukan dengan menggunakan rumus jarak Euclidean [11]. Persamaan jarak *euclidean* adalah sebagai berikut:

---

$$d_{ik} = \sqrt{\sum_{j=1}^p (X_{ij} - y_{kj})^2} \quad (1)$$

dengan

$d_{ik}$  : Jarak *euclidean* objek ke- $i$  ke objek ke- $k$

$x_{ij}$  : Nilai objek  $i$  pada variabel ke  $j, j = 1, 2, \dots, p$

$y_{kj}$  : Nilai objek  $k$  pada variabel ke  $j$

$p$  : Banyak variabel yang diamati

### **DENSITY BASED SPATIAL CLUSTERING ALGORITTM WITH NOISE (DBSCAN)**

Salah satu metode pengelompokan berbasis kepadatan data adalah *Density Based Spatial Clustering Algorithm with Noise* (DBSCAN). Jumlah data atau Minpts dalam radius epsilon ( $\epsilon$ ) dari setiap data merupakan kepadatan data dalam metode DBSCAN. Untuk kumpulan data apa pun, kepadatan ini menghasilkan tiga status berbeda ialah inti, batas, dan noise (outlier). *Clustering* dilakukan menggunakan DBSCAN berdasarkan parameter input yaitu epsilon dan Minpts [7].

Dirancang pada tahun 1996, metode DBSCAN melihat kepadatan lokal komponen database untuk menemukan pengelompokan dalam kumpulan data geografis yang sangat besar hanya dengan menggunakan satu parameter masukan. Informasi juga dapat dikategorikan sebagai noise atau outlier menggunakan DBSCAN [7].

Teknik Pengelompokan menggunakan *Density Based Spatial Clustering Algorithm with Noise* (DBSCAN) adalah sebagai berikut:

1. Standarisasi data dengan *z-score*

Adapun persamaan *z-score* adalah sebagai berikut:

$$Z_i = \frac{x_i - \bar{x}}{s}, s \neq 0 \quad (2)$$

dengan

$Z_i$  : Standarisasi Variabel

$x_i$  : Data ke  $i$

$\bar{x}$  : Rata-rata Keseluruhan Data Tiap Variabel

$s$  : Standar Deviasi

2. Menentukan titik centroid awal secara acak
3. Menentukan Minpts dan epsilon dengan menggunakan algoritma *K-Nearst Neighbour* (KNN)
4. Menghitung jarak *euclidean*
5. Pembentukan *cluster*
6. Validasi *cluster* dengan *silhouette coefficient*.
7. Interpretasi *cluster*.

### **VALIDASI CLUSTER**

*Silhouette coefficient* digunakan sebagai validasi *cluster*. *Silhouette coefficient* mengukur seberapa efektif suatu objek diposisikan dalam sebuah *cluster*, serta kekuatan dan kualitas cluster. Berikut langkah-langkah menghitung nilai *silhouette coefficient* [12]:

1. Menghitung rata-rata jarak objek ke- $i$  dengan semua objek yang berada di dalam satu *cluster* A dengan persamaan:

$$a(i) = \frac{1}{|A|-1} \sum_{j \in A, j \neq i} d(i, j) \quad (3)$$

dengan

$a(i)$  : Rata-rata jarak  $i$  terhadap semua objek di *cluster* A

$j$  : Objek lain dalam satu *cluster*

$i$  : Objek yang akan diteliti

A : Total data dalam *cluster* A

$d(i, j)$  : Jarak antara objek  $j$  dan  $i$

2. Menghitung rata-rata jarak objek ke- $i$  dengan semua objek yang berada pada *cluster* lain, dan diambil nilai terkecilnya dengan persamaan:

$$d(i, C) = \frac{1}{|A|} \sum_{j \in C} d(i, j) \quad (4)$$

dengan

$d(i, C)$  : Jarak rata-rata objek  $i$  dengan semua objek pada *cluster* lain  $C$  dimana  $A \neq C$

$C$  : *Cluster* lain

A : Total data dalam *cluster* A

3. Menentukan nilai minimum yaitu  $b(i)$  yang menunjukkan perbedaan rata-rata  $i$  untuk *cluster* yang terdekat dengan tetangganya dan dengan persamaan:

$$b(i) = \min_{C \neq A} d(i, C) \quad (5)$$

dengan

$b(i)$  : Rata-rata jarak  $i$  terhadap semua objek di *cluster* lain

4. Selanjutnya menghitung nilai *silhouette* dengan persamaan:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (6)$$

dengan

$a(i)$  : Rata-rata jarak  $i$  terhadap semua objek di *cluster* A

$b(i)$  : Rata-rata jarak  $i$  terhadap semua objek di *cluster* lain

5. Selanjutnya menghitung nilai *silhouette coefficient* atau nilai rata-rata *silhouette* di semua objek yang berada pada masing-masing *cluster* dengan persamaan:

$$SC = \frac{1}{n} \sum_{i=1}^n s(i) \quad (7)$$

dengan

$SC$  : *Silhouette coefficient*

$n$  : Banyak pengamatan

Berikut ini ialah kriteria pengukuran *silhouette coefficient* yang ditunjukkan pada Tabel 1.

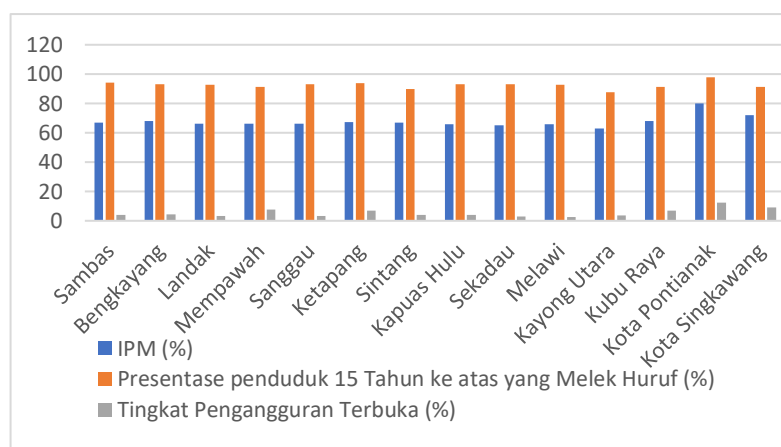
**Tabel 1.** Kriteria Pengukuran *Silhouette Coefficient*

Nilai <i>Silhouette Coefficient</i>	Kriteria
0,71 – 1,00	Struktur Kuat
0,51 – 0,70	Struktur Baik
0,26 – 0,50	Struktur Lemah
$\leq 0,25$	Struktur Buruk

## HASIL DAN PEMBAHASAAN

Data yang digunakan dalam penelitian ini adalah jenis data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Kalimantan Barat tahun 2021. Data yang digunakan berasal dari 14 Kabupaten/Kota di Provinsi Kalimantan Barat tahun 2021. Tiga variabel IPM ( $x_1$ ), persentase penduduk

usia 15 tahun ke atas yang melek huruf ( $x_2$ ), dan tingkat pengangguran terbuka ( $x_3$ ). Langkah selanjutnya yaitu melakukan eksplorasi data yang ditunjukkan pada Gambar 1.



**Gambar 1.** Eksplorasi Data

Gambar 1 merupakan eksplorasi data dengan variabel IPM, persentase penduduk usia 15 tahun ke atas yang melek huruf, dan tingkat pengangguran terbuka di 14 Kabupaten/Kota Kalimantan Barat. Selanjutnya Kabupaten Kayong Utara memiliki IPM terendah sebesar 62,90% dan Kabupaten Kubu Raya memiliki IPM tertinggi sebesar 68,16%. Variabel persentase penduduk usia 15 tahun ke atas yang melek huruf, Kabupaten Kayong Utara memiliki persentase terendah sebesar 87,48%, sedangkan Kabupaten Sambas memiliki persentase tertinggi sebesar 94,02%. Variabel tingkat pengangguran terbuka, Kabupaten Melawi memiliki tingkat pengangguran terbuka terendah sebesar 2,66%, sedangkan Kota Pontianak tertinggi sebesar 12,38%.

### 1. Standarisasi Data

*Z-score* digunakan untuk melakukan standarisasi data pada variabel-variabel yang berdampak pada kemiskinan di Kalimantan Barat pada tahun 2021. *Z-score* digunakan untuk menyesuaikan data apabila terdapat kesenjangan data antar variabel yang signifikan. Tabel 2 merupakan nilai standarisasi data yang dihitung menggunakan *z-score*.

**Tabel 2.** Standarisasi *Z-Score*

No.	Kab/Kota (i)	IPM (%)	Presentase penduduk 15 Tahun ke atas yang Melek Huruf (%)	Tingkat Pengangguran Terbuka (%)
1.	Sambas	-0,144	0,664	-0,508
2.	Bengkayang	0,087	0,194	-0,349
3.	Landak	-0,362	0,177	-0,772
4.	Mempawah	-0,406	-0,589	0,810
5.	Sanggau	-0,365	0,332	-0,691
6.	Ketapang	-0,063	0,560	0,539
7.	Sintang	-0,185	-1,106	-0,515
8.	Kapuas Hulu	-0,475	0,280	-0,434
9.	Sekadau	-0,677	0,341	-0,878
10.	Melawi	-0,446	0,065	-0,970
11.	Kayong Utara	-1,175	-2,152	-0,575
12.	Kubu Raya	0,117	-0,495	0,567

13.	Pontianak	3,008	2,265	2,456
14.	Singkawang	1,087	-0,538	1,321

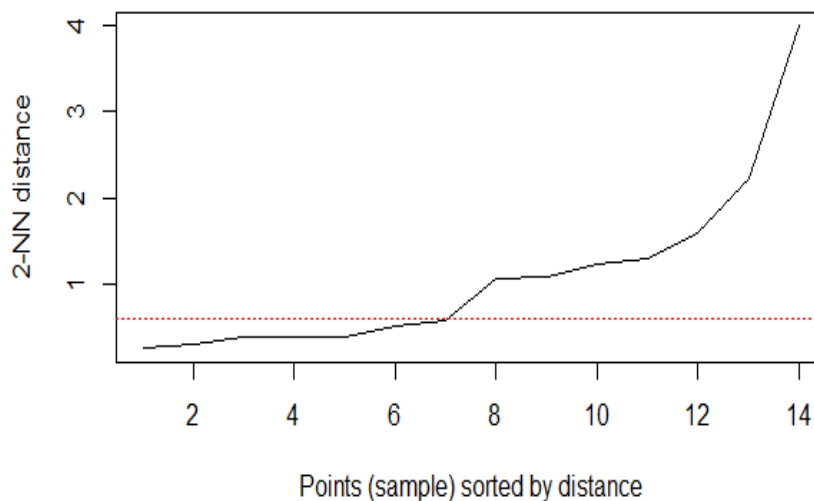
Berdasarkan Tabel 2 diperoleh nilai standarisasi yang menunjukkan bahwa jarak antara variabel data yang tidak jauh satu sama lain dimana nilai minimal sebesar -2,152 dan nilai maksimum sebesar 3,008.

## 2. Menentukan Titik *Centroid* Awal

Penentuan titik *centroid* awal diperlukan untuk tahap pengelompokan DBSCAN. *Centroid* awal dipilih secara acak. Kabupaten Sambas dengan IPM sebesar -0,144%, persentase penduduk melek huruf sebesar 0,664%, dan tingkat pengangguran terbuka sebesar -0,508%, dipilih sebagai *centroid* awal pada penelitian ini.

## 3. Menentukan Minpts dan Epsilon

Penelitian ini menggunakan parameter Minpts 2, dan perhitungan KNN menghasilkan  $\epsilon$  optimal dengan menganalisis K sama dengan Minpts. Untuk membantu menentukan nilai  $\epsilon$  adalah plot KNN. Berikut adalah plot KNN yang menampilkan Minpts 2:



**Gambar 2.** Plot KNN

Plot KNN terlihat pada Gambar 2. Dengan melihat kurva siku diketahui bahwa 0,6 merupakan nilai epsilon yang ideal. Titik dimana suatu kurva tampak membengkok, terutama dari kemiringan yang tinggi ke kemiringan yang rendah atau ke arah yang lain.

## 4. Menghitung Jarak *Euclidean*

Jarak *euclidean* digunakan untuk menentukan jarak antar Kabupaten/Kota. Berikut contoh adalah perhitungan pada Kabupaten Sambas dengan Kabupaten lainnya yang dimulai dari Kabupaten Sambas dan Kabupaten Bengkayang.

$$d_{(i,k)} = \sqrt{(-0,144 - 0,087)^2 + (0,664 - 0,194)^2 + (-0,508 - (-0,349))^2} = 0,5465$$

Perhitungan jarak *euclidean* dilakukan hingga 14 Kabupaten/Kota. Hasil penentuan jarak *euclidean* pada suatu Kabupaten/Kota di Kalimantan Barat disajikan pada Tabel 3.

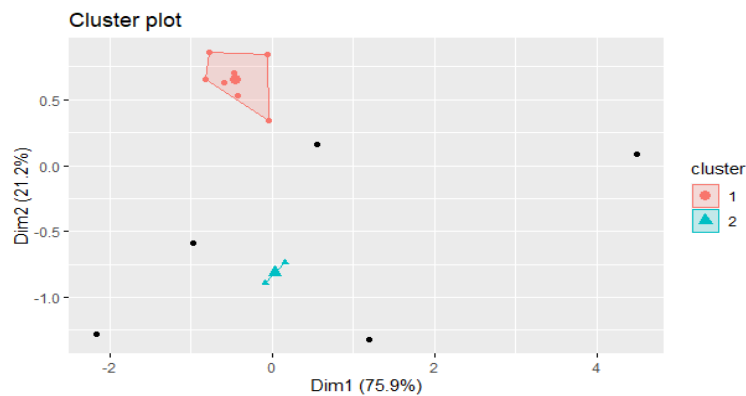
**Tabel 3.** Jarak *Euclidean*

Kabupaten/Kota	Sambas	Bengkayang	...	Singkawang
Sambas	0,0000	0,5465		
Bengkayang	0,5465	0,0000		
Landak	0,5953	0,6175	∴	
∴	∴	∴	∴	
Singkawang	2,5108	2,0801	...	0,0000

Jarak *euclidean* masing-masing Kabupaten/Kota di Kalimantan Barat tahun 2021 disajikan pada Tabel 3. Untuk pengolahan DBSCAN, digunakan jarak *euclidean* dengan memilih centroid secara acak hingga seluruh data dianalisis.

**5. Pembentukan Cluster**

Pembentukan *cluster* dengan metode DBSCAN ini digunakan nilai parameter *Minpts* sebesar 2 dan  $\epsilon$  sebesar 0,6. *Cluster* terbentuk jika suatu titik dalam radius  $\epsilon = 0,6$  ke atas sama dengan *Minpts* = 6 sehingga titik tersebut menjadi inti, tetangganya menjadi pembatas dan membentuk kepadatan yang dapat dijangkau atau jika jarak antar titik tidak lebih daripada nilai epsilon, maka *cluster* terbentuk dan jika dalam radius  $\epsilon$  lebih kecil dari *Minpts* = 2 atau jika titik tidak mencukupi nilai  $\epsilon = 0,6$  maka terjadi *noise*. Gambar 3 menampilkan hasil pembentukan cluster menggunakan metode DBSCAN.



**Gambar 3.** Hasil *Cluster*

Dari Gambar 3 terdapat dua *cluster* yang dihasilkan. Warna merah menunjukkan *Cluster* satu yang terdiri dari tujuh Kabupaten Sambas, Bengkayang, Landak, Sanggau, Kapuas Hulu, Sekadau, dan Melawi. *Cluster* dua yang terdiri dari dua Kabupaten Mempawah dan Kubu Raya digambarkan dengan warna biru dan terhubung dengan baik. Lima Kabupaten/Kota yaitu Keatapang, Sintang, Kayong Utara, Pontianak dan Singkawang merupakan noise yang digambarkan dengan warna hitam.

**6. Validasi Cluster**

Perhitungan *silhouette coefficient* digunakan untuk validasi *cluster*. Data yang digunakan dalam perhitungan *silhouette coefficient* merupakan data *cluster* dari Kabupaten Sambas. Mencari nilai  $a(i)$  dan  $b(i)$  menggunakan jarak *euclidean* adalah langkah selanjutnya. dimana  $b(i)$  adalah jarak rata-rata cluster yang lain dengan menggunakan persamaan 5 dan  $a(i)$  adalah jarak rata-rata cluster yang sama dengan persamaan 3.

Nilai *silhouette coefficient* yang ditunjukkan pada Tabel 4 sebagai berikut.

**Tabel 4.** *Silhouette coefficient*

No.	Cluster	Kabupaten/Kota (i)	a(i)	b(i)	s(i)
1.	Cluster 1	Sambas	0,615	3,819	0,839
2.		Bengkayang	0,700	2,845	0,754
3.		Landak	0,377	4,059	0,907
4.		Sanggau	0,346	3,968	0,913
5.		Kapuas Hulu	0,433	3,120	0,861
6.		Sekadau	0,534	4,744	0,887
7.		Melawi	0,529	4,689	0,887
8.	Cluster 2	Mempawah	0,585	0,585	0,001
9.		Kubu Raya	0,585	0,585	0,001
<b>Silhouette Coefficient</b>					<b>0,672</b>

Nilai koefisien siluet sebesar 0,67 pada Tabel 4 menunjukkan bahwa struktur tersebut baik karena mendekati 1 atau menghasilkan hasil pengelompokan yang baik.

## 7. Interpretasi Cluster

Berikut ini adalah interpretasi *cluster* dan nilai rata-rata variabel tiap *cluster* yang ditampilkan pada Tabel 5.

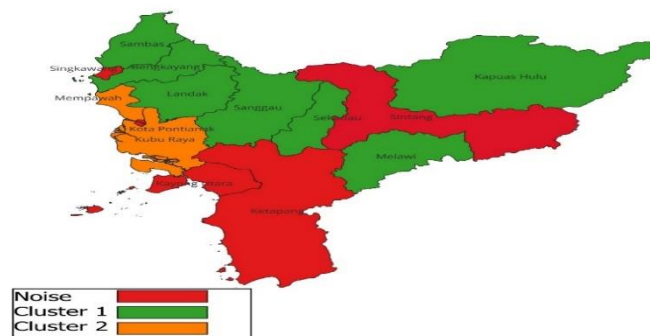
**Tabel 5.** Nilai Rata-Rata Tiap Cluster

	IPM (%)	Persentase penduduk 15 Tahun ke atas yang Melek Huruf (%)	Tingkat Pengangguran Terbuka (%)
Cluster 1	66,30	93,16	3,55
Cluster 2	67,10	91,22	7,37

Tabel 5 menunjukkan terdapat dua *cluster* dari pembentukan *cluster* dengan metode DBSCAN:

1. Cluster satu menjelaskan *cluster* yang mempunyai nilai rata-rata IPM dan tingkat pengangguran terbuka lebih rendah.
2. Cluster dua menjelaskan *cluster* yang mempunyai nilai rata-rata IPM dan tingkat pengangguran terbuka lebih tinggi.

Berikut adalah hasil pengelompokan Kabupaten/Kota di Kalimantan Barat dengan faktor-faktor kemiskinan membentuk dua *cluster* dan noise yang ditampilkan pada peta pada Gambar 4.



**Gambar 4.** Peta Cluster



## PENUTUP

### Kesimpulan

Setelah diterapkan metode DBSCAN pada pengelompokan Kabupaten/Kota di Kalimantan Barat, dengan  $\epsilon$  sebesar 0,6 dan Minpts sebesar dua, diperoleh hasil nilai *silhouette coefficient* sebesar 0,67 yang menunjukkan hasil pengelompokan yang baik dan dihasilkannya dua cluster yaitu *cluster* satu yang terdiri dari tujuh Kabupaten di Kalimantan Barat yaitu Sambas, Bengkayang, Landak, Sanggau, Kapuas Hulu, Sekadau, dan Melawi, serta *cluster* dua yang terdiri dari dua Kabupaten yaitu Mempawah dan Kubu Raya. Lima Kabupaten/Kota yang termasuk dalam Noise adalah Ketapang, Sintang, Kayong Utara, Pontianak, dan Singkawang. Pada *cluster* satu, rata-rata nilai IPM dan tingkat pengangguran terbuka lebih rendah, sedangkan rata-rata persentase penduduk usia 15 tahun ke atas yang melek huruf lebih besar. *Cluster* dua ditandai dengan persentase rata-rata penduduk berusia 15 tahun ke atas yang melek huruf lebih rendah, namun nilai IPM rata-rata dan tingkat pengangguran terbuka lebih tinggi. Dibandingkan dengan *cluster* satu, *cluster* dua memiliki tingkat pengangguran terbuka dan rata-rata IPM yang lebih besar.

### Saran

Beberapa saran terkait penelitian ini yang dapat peneliti berikan sebagai berikut.

1. Untuk menciptakan tenaga kerja terampil yang mampu bersaing bahkan menciptakan lapangan kerja, pemerintah di Kabupaten Kubu Raya dan Mempawah perlu meningkatkan kesadaran masyarakat terhadap program pelatihan kerja, seperti Kartu Prakerja. Hal ini akan membantu pencari kerja, pekerja atau buruh yang kehilangan pekerjaan, dan pekerja atau buruh untuk meningkatkan kemampuan dan meningkatkan kesempatan kerja.
2. Pemerintah meningkatkan proyek Indeks Pembangunan Desa (IDM) di Kabupaten Kubu Raya dan Mempawah. Selain mengetahui kondisi kemandirian dan kemajuan desa saat ini, IDM juga menyediakan data dan informasi mendasar untuk pembangunan desa sehingga setiap pemerintah daerah dapat fokus pada setiap nilai indikator IPM yang perlu ditingkatkan guna menentukan intervensi, anggaran, dan desa. kebijakan pembangunan yang lebih tepat sasaran dan fokus. Hal ini berpotensi meningkatkan IPM, yang akan menurunkan jumlah orang miskin seiring dengan peningkatan tersebut.
3. Pembangunan pendidikan di Kabupaten Kubu Raya dan Mempawah khususnya pendidikan usia 15 tahun keatas harus menjadi proyek unggulan untuk menekan angka putus sekolah dan meningkatkannya angka melek huruf melalui berbagai program baik prioritas maupun strategis.
4. Penelitian-penelitian di masa depan dapat menggabungkan temuan-temuan mereka dengan pendekatan-pendekatan lain agar lebih berharga dan berguna sebagai panduan bagi pemerintah dalam menerapkan kebijakan untuk menurunkan angka kemiskinan di Kabupaten/Kota di Provinsi Kalimantan Barat.

### DAFTAR PUSTAKA

- [1]. Priseptian, L., & Primandhana, W. P. (2022). Analisis faktor-faktor yang mempengaruhi kemiskinan. In *Forum Ekonomi*, 24(1), 45-53.
  - [2]. Nasmiwati, R., & Triani, M. (2019). Analisis faktor-faktor yang mempengaruhi kemiskinan rumah tangga di Kec Bayang Kab Pesisir Selatan. *Jurnal Kajian Ekonomi dan Pembangunan*, 1(1), 213-222.
  - [3]. BPS Prov Kalbar. *Kalbar Dalam Angka 2021*,(online) <https://www.kalbar.bps.go.id/>, (diakses pada 08 Juli 2022).
-

- [4]. Nurlita, C. A., Musa, A. H., & Suharto, R. B. (2017). Pengaruh Indeks Pembangunan Manusia (IPM) dan Pertumbuhan Ekonomi terhadap Pengangguran dan Jumlah Penduduk Miskin di Samarinda. *Jurnal Ilmu Ekonomi Mulawarman (JIEM)*, 2(1), 1-10.
- [5]. Nafisah, Q., & Chandra, N. E. (2017). Analisis Cluster Average Linkage Berdasarkan Faktor-Faktor Kemiskinan di Prov Jawa Timur. *Zeta-Math Journal*, 3(2), 31-36.
- [6]. Talakua, M. W., Leleury, Z. A., & Taluta, A. W. (2017). Analisis cluster dengan menggunakan metode k-means untuk pengelompokan Kab/Kota di Prov maluku berdasarkan indikator indeks pembangunan manusia tahun 2014. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 11(2), 119-128.
- [7]. Id, I. D. (2017). Modifikasi DBSCAN (Density-Based Spatial Clustering with Noise) Pada Objek 3 Dimensi. *Jurnal Komputer Terapan*, 3(1), 41-52.
- [8]. Rahman, R. R. A., & Wijayanto, A. W. (2021). Pengelompokan Data Gempa Bumi Menggunakan Algoritma DBSCAN. *Jurnal Meteorologi dan Geofisika*, 22(1), 31-38.
- [9]. Sari, B. N., & Primajaya, A. (2019). Penerapan clustering DBSCAN untuk pertanian padi di Kab Karawang. *JIKO (Jurnal Inform. dan Komputer) STMIK AKAKOM*, 4(1), 28-34.
- [10]. Dinata, R. K., Akbar, H., & Hasdyna, N. (2020). Algoritma K-Nearest Neighbor dengan Euclidean Distance dan Manhattan Distance untuk Klasifikasi Transportasi Bus. *ILKOM Jurnal Ilmiah*, 12(2), 104-111.
- [11]. Anton, H., 2014. *Elementary Linear Algebra Application Version*, Ed ke-11, Wiley, Canada.
- [12]. Anggara, M., Sujiani, H., & Nasution, H. (2016). Pemilihan distance measure Pada K-Means clustering Untuk Pengelompokan member di Alvaro fitness. *Jurnal Sistem dan Teknologi Informasi (JUSTIN)*, 1(1), 1-6.

PURWANTY : Jurusan Matematika, FMIPA UNTAN, Pontianak,  
[purwanty@student.untan.ac.id](mailto:purwanty@student.untan.ac.id)

SHANTIKA MARTHA : Jurusan Matematika, FMIPA UNTAN, Pontianak,  
[Shantika.martha@math.untan.ac.id](mailto:Shantika.martha@math.untan.ac.id)

SITI APRIZKIYANDARI : Jurusan Budidaya Pertanian, FAPERTA UNTAN,  
Pontianak, [siti.aprizkiyandari@faperta.ac.id](mailto:siti.aprizkiyandari@faperta.ac.id)

---