

Pengelompokan Kabupaten/Kota di Provinsi Sumatera Barat Berdasarkan Sebaran Tingkat Pengangguran Menggunakan Metode *Density Based Spatial Clustering Applications with Noise*

Difitria Oktaviani¹, Dewi Murni²

^{1,2} Program Studi Matematika, Fakultas Matematika Ilmu Pengetahuan dan Alam, Universitas Negeri Padang (UNP)

Article Info

Article history:

Received November 10, 2024

Revised November 25, 2024

Accepted December 05, 2024

Keywords:

Unemployment Rate

Clustering

DBSCAN

Silhouette Coefficient

Kata Kunci:

Tingkat Pengangguran

Clustering

DBSCAN

Silhouette Coefficient

ABSTRACT

The population and workforce participation in West Sumatra Province increased in 2022. However, there is also an imbalance between the number of job seekers and job vacancies. This has resulted in an increase in unemployment in West Sumatra Province. This study aims to group districts/cities in West Sumatra based on unemployment rates using the Density Based Spatial Clustering Applications with Noise method, and validate the clusters using the silhouette coefficient evaluation method. Based on the grouping results, two clusters and one noise were obtained. Overall, with this method, the cluster validation results were 0.50, which means a good cluster structure category.

ABSTRAK

Jumlah penduduk dan partisipasi angkatan kerja di Provinsi Sumatera Barat mengalami peningkatan pada tahun 2022. Namun, juga terjadi ketidakseimbangan antara jumlah pencari kerja dan lowongan kerja. Hal ini mengakibatkan naiknya angka pengangguran di Provinsi Sumatera Barat. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Sumatera Barat berdasarkan tingkat pengangguran menggunakan metode *Density Based Spatial Clustering Applications with Noise*, dan validasi *cluster* dengan metode evaluasi *silhouette coefficient*. Berdasarkan hasil pengelompokannya diperoleh dua *cluster* dan satu *noise*. Secara keseluruhan, dengan metode ini diperoleh hasil validasi *cluster* sebesar 0,50 yang berarti kategori struktur *cluster* yang baik.

This is an open access article under the [CC BY-SA](#) license.



Penulis pertama

(Difitria Oktaviani)

Program Studi Matematika, Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Padang, Jl. Prof. Dr. Hamka, Air Tawar barat, Padang Utara, Padang, Indonesia. Kode Pos: 25131

Email: difitriaoktaviani9@gmail.com

1. PENDAHULUAN

Provinsi Sumatera Barat pada tahun 2022, mengalami peningkatan jumlah penduduk sebanyak 60.397 jiwa dibandingkan tahun sebelumnya. Sementara itu, pada tahun 2021 jumlah penduduk yang bekerja tercatat sebanyak 2,58 juta orang, dan mengalami kenaikan menjadi 2,69 juta orang pada tahun 2022 [1]. Peningkatan jumlah penduduk juga mengakibatkan naiknya jumlah angkatan kerja [2].

Angkatan kerja adalah penduduk berusia 15 tahun ke atas, mencakup individu yang sedang bekerja, mempunyai pekerjaan tetapi sedang tidak bekerja, serta pengangguran [3]. Pada Agustus 2022, Tingkat Partisipasi Angkatan Kerja (TPAK) di Provinsi Sumatera Barat juga mengalami sebanyak 1,58% dibandingkan dengan Agustus pada tahun sebelumnya [1]. Pada tahun 2022 terjadi ketidakseimbangan antara jumlah pencari kerja dan jumlah lowongan kerja. Jumlah pencari kerja mencapai 19.229 jiwa, sementara jumlah lowongan kerja hanya berjumlah 3.814. Hal ini berarti satu



lowongan pekerjaan hanya tersedia untuk lima orang pencari kerja. Ini menunjukkan bahwa meskipun banyak orang yang terlibat dalam mencari pekerjaan (TPAK meningkat), tapi jika jumlah lowongan pekerjaan tidak mencukupi untuk menyerap semua angkatan kerja yang tersedia, maka dapat menyebabkan peningkatan pengangguran [4].

Pada tahun 2022, jumlah pengangguran di Provinsi Sumatera Barat meningkat dibandingkan tahun sebelumnya, mencapai 180.106 jiwa. [5]. Tingginya angka pengangguran disebabkan oleh angkatan kerja yang terus meningkat, lapangan pekerjaan yang terbatas, laju pertumbuhan penduduk yang tinggi, dan besaran upah yang diterima. Apabila hal tersebut dibiarkan, maka akan menyebabkan kemiskinan dan juga berpotensi meningkatnya kriminalitas.

Berdasarkan kondisi tersebut, diperlukan langkah strategis untuk mengatasi tingkat pengangguran dengan melakukan pengelompokan kabupaten/kota yang memiliki tingkat pengangguran relatif tinggi. Pengelompokan ini dapat digunakan pemerintah untuk menentukan daerah yang perlu diprioritaskan dalam penanganan pengangguran, sehingga jumlah pengangguran dapat dikurangi. Melalui pengelompokan ini, pemerintah dapat merancang strategi penanganan pengangguran yang sesuai dengan kondisi spesifik masing-masing daerah. Proses pengelompokan dilakukan menggunakan analisis *cluster*.

Analisis *cluster* digunakan untuk membagi data ke dalam beberapa kelompok kecil, di mana setiap kelompok memiliki kesamaan atau kemiripan antar objek di dalamnya di mana setiap kelompok terdiri atas objek-objek yang memiliki karakteristik serupa [6]. Salah satu metode analisis klaster yang digunakan adalah metode berbasis kepadatan, yaitu *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) [7]. Metode DBSCAN mengacu pada jumlah minimum titik (*MinPts*) yang berada dalam batas jarak maksimum tertentu (*epsilon*) [8]. DBSCAN dapat mengelompokkan data berdasarkan densitas, sehingga mampu mengidentifikasi wilayah dengan tingkat pengangguran tinggi, sedang, dan rendah.

DBSCAN memiliki dua parameter utama yang menentukan jumlah yang terbentuk, yaitu *epsilon* dan *MinPts*. Parameter *epsilon* adalah batas jarak maksimum antara titik pusat dan titik anggota dalam satu klaster, sedangkan parameter *MinPts* adalah jumlah minimum anggota yang diperlukan untuk membentuk sebuah *cluster* [9]. Terdapat tiga jenis titik dalam metode ini yaitu *core*, *border*, dan *noise*. Suatu titik dikatakan sebagai *core point* apabila dalam *epsilon* tersebut terdapat jumlah titik $\geq$ *MinPts*. Jika dalam titik tersebut terdapat satu *core point* tapi tidak melebihi nilai *MinPts* maka disebut dengan *border point*. Suatu titik disebut sebagai *noise point* jika dalam radius *epsilon* hanya terdapat titik tersebut tanpa ada titik lain di sekitarnya [10]. Keunggulan metode DBSCAN terletak pada kemampuannya untuk membentuk klaster tanpa membutuhkan informasi awal mengenai jumlah cluster yang harus terbentuk, sehingga metode ini dapat secara otomatis menentukan jumlah dan struktur *cluster* berdasarkan pola yang ada dalam data., serta kemampuannya dalam mendeteksi adanya noise dalam data [11].

2. METODE PENELITIAN

Penelitian ini merupakan penelitian terapan yang menggunakan data sekunder. Data tersebut bersumber dari publikasi resmi Badan Pusat Statistik (BPS) Provinsi Sumatera Barat, yang berasal dari hasil Survei Angkatan Kerja Nasional (Sakernas) tahun 2023. Data tersebut adalah jumlah penduduk (X_1), angkatan kerja (X_2), Tingkat Partisipasi Angkatan Kerja (TPAK) (X_3), Pengangguran (X_4), dan Indeks Pembangunan Manusia (IPM) (X_5). Kelima indikator tersebut adalah variabel yang digunakan dalam proses pengelompokan kabupaten/kota. Berikut langkah-langkah analisis dalam penelitian ini mencakup beberapa tahapan penting sebagai berikut

- A. Mengumpulkan data mengenai indikator tingkat pengangguran.
 - a. Melakukan standarisasi data (*z-score*) [12]
 - b. Menghitung jarak manhattan untuk tiap variabel. Jarak ini yang akan digunakan untuk melakukan proses pengelompokan nantinya [13]
 - c. Menentukan nilai parameter *epsilon* dan *MinPts*. Penentuan nilai parameter dilakukan melalui uji coba menggunakan plot KNN *distance* di R Studio.



- d. Memeriksa jumlah titik dalam *eps* dengan membandingkannya terhadap *MinPts*. Apabila jumlah titik dalam *eps* $\leq$ *MinPts*, maka titik tersebut dikategorikan sebagai *noise*. Sebaliknya, jika jumlah titik dalam *eps* $\geq$ *MinPts*, maka akan terbentuk sebuah *cluster*.
- e. Memeriksa kembali apakah semua titik telah dikunjungi.
- f. Menghitung nilai *silhouette* untuk melihat seberapa baik kluster yang terbentuk [14]
Berikut adalah kriteria pengukuran *silhouette coefficient* [15]:

Tabel 1. Kriteria Nilai *Sillhoeutte Coefficient*

Nilai SC	Keterangan
$0,7 < SC \leq 1$	Struktur Kuat
$0,5 < SC \leq 0,7$	Struktur Baik
$0,25 < SC \leq 0,5$	Struktur Lemah
$SC \leq 0,25$	Tidak Ada Struktur

- g. Menghitung rata-rata variabel tiap kelompok atau *cluster*.
- h. Menginterpretasikan hasil yang sudah didapat dan menarik kesimpulan.

3. HASIL DAN PEMBAHASAN

Pada bagian ini, penjelasan hasil penelitian akan dibagi menjadi beberapa bagian yang mencakup hal-hal berikut:

3.1. Statistika Deskriptif

Data indikator tingkat pengangguran diperoleh dari buku publikasi BPS Provinsi Sumatera Barat. Setelah proses pengumpulan data selesai, langkah berikutnya adalah mendeskripsikan data berdasarkan masing-masing variabel dan mengidentifikasi karakteristiknya. Penentuan karakteristik data dilakukan dengan menghitung nilai *mean*, *maximum*, dan *minimum* sebagai berikut:

Tabel 2. Statistika Deskriptif Indikator Tingkat Pengangguran di Provinsi Sumatera Barat

Variabel	N	Mean	Maximum	Minimum
X_1	19	296.875	919.145	57.850
X_2	19	150.962	480.376	26.172
X_3	19	70,24	78,25	62,81
X_4	19	9.479	56.163	716
X_5	19	73,31	83,29	62,19

Berdasarkan Tabel 2, rata-rata jumlah penduduk di Provinsi Sumatera Barat adalah 296.875 orang, dengan jumlah penduduk tertinggi mencapai 919.145 orang dan terendah sebesar 57.850 orang. Sementara itu, rata-rata jumlah pengangguran adalah 9.479 orang, dengan jumlah pengangguran tertinggi sebesar 56.163 orang dan terendah sebesar 716 orang. Karena terdapat perbedaan rentang nilai yang cukup signifikan antar data, maka standarisasi data perlu dilakukan sebelum analisis lebih lanjut.

3.2. Standarisasi Data (Z-score)

Standarisasi data dilakukan ketika data memiliki satuan yang berbeda dan terdapat perbedaan rentang nilai yang signifikan antar data. Hasil perhitungannya yaitu:



Tabel 3. Standarisasi Data

No	X_1	X_2	X_3	X_4	X_5
1	-0,94068	-0,89050	1,73476	-0,71838	-2,05913
2	0,99585	0,70867	-0,71349	0,09204	-0,45708
3	0,45772	0,47604	0,90929	0,20853	-0,60895
4	-0,25643	-0,25372	-0,44917	-0,28768	-0,85528
5	0,36000	0,30615	0,14231	0,11975	-0,00331
6	0,63137	0,56184	0,04265	0,37936	-0,31076
7	1,10642	1,06561	0,60813	0,31272	-0,00331
8	0,41486	0,59116	0,93095	-0,11520	-0,56079
9	0,04783	0,01775	1,02412	-0,10298	-1,09234
10	-0,49069	-0,44608	1,67843	-0,46884	-0,66636
11	-0,28184	-0,07091	0,30481	-0,04691	-0,18667
12	0,66016	0,54036	-1,37864	0,31952	-0,69229
13	2,82136	2,95379	-1,61047	3,82695	1,84877
14	-1,00212	-1,00817	-0,16968	-0,65403	1,09683
15	-1,04491	-1,04901	0,13148	-0,63788	0,07818
16	-1,08373	-1,11896	-1,34831	-0,67313	1,01348
17	-0,79147	-0,75017	-1,25948	-0,50696	1,50243
18	-0,69619	-0,69064	-0,03968	-0,46458	1,15239
19	-0,90750	-0,94319	-0,53800	-0,58230	0,80420

Berdasarkan Tabel 3, data telah memiliki skala dan satuan yang sama, di mana setelah proses standarisasi, nilai rata-rata dari kelima variabel menjadi nol. Karakteristik dengan rentang nilai yang jauh sebelumnya dapat memengaruhi hasil perhitungan jarak antar titik data. Langkah selanjutnya adalah menganalisis data menggunakan metode DBSCAN.

3.3. Penentuan Parameter *Epsilon* dan *MinPts*

Langkah awal dalam pengelompokan dengan metode DBSCAN adalah menentukan parameter *epsilon* dan *MinPts*. Penentuan parameter optimal dilakukan melalui beberapa percobaan untuk mendapatkan hasil pengelompokan yang paling baik. Kombinasi nilai *MinPts* dan *epsilon* diuji secara uji coba dengan memperhatikan nilai *sillhouette coefficient* sebagai indikator kualitas *cluster*. Hasil percobaan pengelompokan dengan berbagai kombinasi nilai *MinPts* dan *epsilon* sebagai berikut:

Tabel 4. Hasil Percobaan Pengelompokan

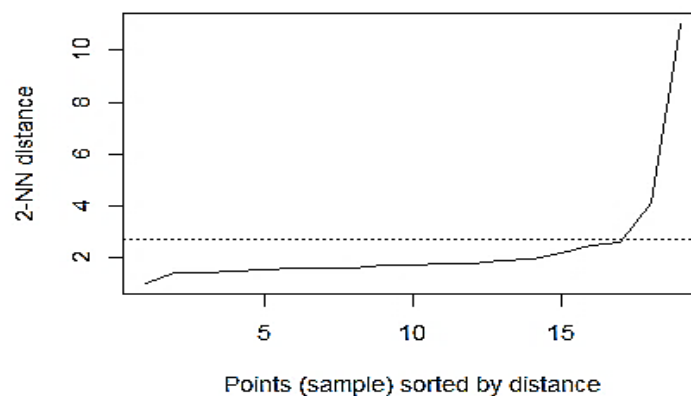
Percobaan	Epsilon	MinPts	Cluster	Noise	Sillhouette
1	1,5	2	3	9	0,10
2	1,7	2	3	5	0,22
3	2	2	2	3	0,41
4	2,7	2	2	1	0,50
5	4	2	1	1	0,66
6	3	3	1	1	0,66
7	2,4	3	2	3	0,41
8	1,7	3	2	7	0,25
9	1,5	3	1	13	0,36
10	3	3	1	1	0,66



Percobaan	Epsilon	MinPts	Cluster	Noise	Sillhouette
11	1,9	4	2	6	0,28
12	2,1	4	2	3	0,41
13	2,8	4	1	2	0,47
14	3	4	1	1	0,66
15	5,4	4	1	1	0,66
16	2,1	5	2	4	0,40
17	2,4	5	2	3	0,41
18	2,7	5	2	2	0,42
19	3,1	5	1	1	0,66
20	3,5	5	1	1	0,66
21	2,4	6	2	4	0,40
22	2,8	6	1	2	0,47
23	3,4	6	1	1	0,66
24	4,1	6	1	1	0,66
25	5,3	6	1	1	0,66

Berdasarkan hasil yang terdapat pada Tabel 4, nilai *epsilon* 2,7 dan *MinPts* 2 dipilih karena menghasilkan nilai *silhouette* sebesar 0,50. Meskipun nilai *silhouette* tertinggi sebesar 0,66, konfigurasi tersebut tidak dipilih karena hanya menghasilkan 1 *cluster*. Oleh karena itu, nilai *silhouette* tertinggi kedua, yaitu 0,50, dipilih karena mampu membentuk 2 *cluster* dan 1 *noise*.

Berdasarkan hasil uji coba pemilihan pengelompokan dengan DBSCAN pada Tabel 4 nilai *epsilon* didapatkan melalui plot KNN menggunakan jarak manhattan. Nilai K disamakan dengan *MinPts* untuk memastikan jarak yang dihitung sesuai dengan syarat minimal tetangga yang dibutuhkan untuk membentuk titik inti (*core point*). Pemilihan parameter *MinPts* dilihat dari *silhouette* yang tertinggi. Hasil plot KNN *distance* sebagai berikut:



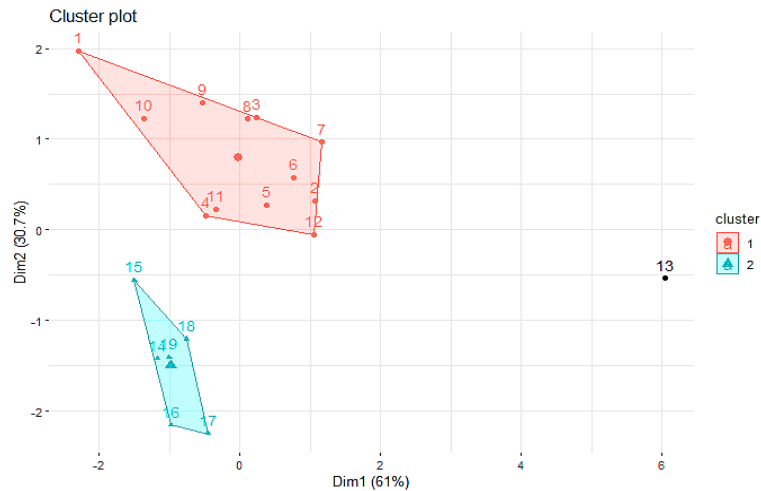
Gambar 1. Plot KNN *distance*

Berdasarkan Gambar 1, nilai *epsilon* yang dipilih adalah 2,7, karena pada rentang tersebut terlihat adanya patahan (siku) pada kurva, yaitu titik di mana grafik mengalami perubahan dari lereng yang curam ke lereng yang lebih landai atau ke arah lain. Oleh karena itu, titik 2,7 dipilih sebagai nilai *epsilon*.



3.4. Clustering

Berdasarkan hasil percobaan, diperoleh nilai parameter *epsilon* 2,7 dan *MinPts* 2 yang menghasilkan dua *cluster* dan satu *noise*. Visualisasi hasil pengelompokan pada masing-masing klaster, berdasarkan parameter *MinPts* dan *epsilon* adalah sebagai berikut:



Gambar 2. Visualisasi Hasil *Clustering* metode DBSCAN dengan *Epsilon* = 2,7 dan *MinPts* = 2

Berdasarkan Gambar 2 visualisasi hasil pengelompokan menghasilkan dua *cluster* dan satu *noise* dengan titik yang menjadi *noise* (pencilan) adalah titik 13 karena posisi titik 13 berada jauh dengan titik-titik yang lain, sehingga titik 13 disebut sebagai *noise*. Titik 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 adalah titik yang termasuk dalam *cluster* 1, dan titik 14, 15, 16, 17, 18, 19 adalah titik yang termasuk dalam *cluster* 2. Adapun untuk hasil lebih lengkapnya dapat dilihat pada tabel sebagai berikut:

Tabel 5. Hasil *Clustering*

Cluster	Jumlah Kabupaten/Kota	Nama Kabupaten/Kota
Noise	1	Kota Padang
Cluster 1	12	Kabupaten Kepulauan Mentawai, Kabupaten Pesisir Selatan, Kabupaten Solok, Kabupaten Sijunjung, Kabupaten Tanah Datar, Kabupaten Padang Pariaman, Kabupaten Agam, Kabupaten Lima Puluh Kota, Kabupaten Pasaman, Kabupaten Solok Selatan, Kabupaten Dharmasraya, Kabupaten Pasaman Barat
Cluster 2	6	Kota Solok, Kota Sawahlunto, Kota Padang Panjang, Kota Bukittinggi, Kota Payakumbuh, Kota Pariaman

3.5. Pengukuran Validitas *Cluster*

Validasi hasil klaster dilakukan dengan menggunakan *silhouette coefficient* untuk mengevaluasi seberapa baik dan kuat struktur data dalam klaster yang telah terbentuk. Perhitungan *silhouette coefficient* dilakukan berdasarkan data jarak yang telah diperoleh, sehingga dapat dihitung nilai $a(i)$ dan $b(i)$. Hasil perhitungan *silhouette coefficient* sebagai berikut:

Tabel 6. Hasil Perhitungan *Sillhouette Coefficient*

Cluster	Kabupaten/Kota	$a(i)$	$b(i)$	$s(i)$
Cluster 1	Kab. Kepulauan Mentawai	6,25399	5,67651	-0,09234
	Kab. Pesisir Selatan	3,51240	6,19886	0,43338
	Kab. Solok	2,61539	6,57321	0,60211
	Kab. Sijunjung	3,41885	3,94417	0,13319
	Kab. Tanah Datar	2,82320	4,84421	0,41720
	Kab. Padang Pariaman	2,89732	5,86829	0,50628
	Kab. Agam	3,75952	7,00888	0,46361
	Kab. Lima Puluh Kota	2,62829	6,29525	0,58250
	Kab. Pasaman	3,02954	5,99175	0,49438
	Kab. Solok Selatan	4,01319	4,85328	0,17310
	Kab. Dharmasraya	2,92120	4,00451	0,27052
	Kab. Pasaman Barat	3,86379	6,42911	0,39902
Cluster 2	Kota Solok	1,37898	5,72470	0,75912
	Kota Sawahlunto	2,25943	4,64045	0,51310
	Kota Padang Panjang	1,87439	6,69465	0,72002
	Kota Bukittinggi	2,11713	6,35831	0,66703
	Kota Payakumbuh	1,75825	5,00871	0,64896
	Kota Pariaman	1,45624	5,41720	0,73118
Noise	Kota Padang	0,00000	13,67196	1,00000
Sillhouette Coefficient				0,49591

Berdasarkan Tabel 6 diperoleh hasil dari *silhouette coefficient* sebesar 0,49591 atau 0,50. Dengan demikian, berdasarkan Tabel 1, hasil yang ditunjukkan termasuk dalam kategori struktur kluster yang baik, karena berada pada interval $0,5 < SC \leq 0,7$.

3.6. Interpretasi *Profiling Cluster*

Setelah anggota dari masing-masing *cluster* diperoleh, langkah selanjutnya adalah melakukan interpretasi profiling kluster. Tujuan dari tahap ini adalah untuk mengidentifikasi karakteristik unik dari setiap kluster. Berikut merupakan hasil perhitungan nilai rata-rata masing-masing kluster:

Tabel 7. Nilai Rata-rata Masing-masing *Cluster*

Variabel	Cluster 1	Cluster 2	Noise	Rata-rata Provinsi Sumatera Barat
Jumlah Penduduk	346.585	93.745	919.145	296875
Angkatan Kerja	175.184	47.614	480.376	150962
TPAK	72,10	67,76	62,81	70,24
Pengangguran	9.166	2.325	56.163	9479
IPM	69,94	78,39	83,29	73,31
Rata-rata	106.215	28.766	291.166	

Berdasarkan Tabel 7, nilai yang berada di atas rata-rata pada masing-masing *cluster* ditandai dengan tulisan tebal. Pada *cluster* pertama memiliki tiga variabel dengan nilai melebihi rata-rata Provinsi Sumatera Barat, yaitu sebesar 106.215. Nilai ini lebih kecil dibandingkan dengan *noise*, namun lebih besar dari *cluster* kedua, sehingga *cluster* ini dikategorikan sebagai daerah dengan tingkat

(Difitria Oktaviani)



pengangguran sedang. *Cluster* kedua memiliki satu variabel dengan nilai yang lebih dari rata-rata Provinsi Sumatera Barat, yaitu sebesar 28.766. Nilainya berada di bawah *noise* dan *cluster* pertama, sehingga *cluster* ini dikategorikan sebagai daerah dengan tingkat pengangguran rendah. Sementara itu, *noise* memiliki semua variabel dengan nilai lebih dari rata-rata Provinsi Sumatera Barat, yaitu sebesar 291.166. Nilainya jauh lebih tinggi dibandingkan dengan *cluster* pertama dan kedua, sehingga *noise* dikategorikan sebagai daerah dengan tingkat pengangguran yang sangat tinggi dibandingkan daerah lainnya.

4. KESIMPULAN

Berdasarkan pembahasan sebelumnya, pengelompokan menggunakan metode DBSCAN terhadap 19 kabupaten/kota di Provinsi Sumatera Barat menghasilkan dua *cluster* dan satu *noise*. *Cluster* pertama merupakan daerah dengan tingkat pengangguran sedang yang terdiri dari 12 anggota, yaitu Kabupaten Kepulauan Mentawai, Kabupaten Pesisir Selatan, Kabupaten Solok, Kabupaten Sijunjung, Kabupaten Tanah Datar, Kabupaten Padang Pariaman, Kabupaten Agam, Kabupaten Lima Puluh Kota, Kabupaten Pasaman, Kabupaten Solok Selatan, Kabupaten Dharmasraya, dan Kabupaten Pasaman Barat. *Cluster* kedua merupakan daerah dengan tingkat pengangguran rendah yang terdiri dari 6 anggota, yaitu Kota Solok, Kota Sawahlunto, Kota Padang Panjang, Kota Bukittinggi, Kota Payakumbuh, dan Kota Pariaman. Sementara itu, *noise* hanya memiliki satu anggota, yaitu Kota Padang, yang dikategorikan sebagai daerah dengan tingkat pengangguran yang jauh lebih tinggi dibandingkan daerah lainnya. Secara keseluruhan, metode DBSCAN dengan parameter *epsilon* 2,7 dan *MinPts* 2 menghasilkan validasi *cluster* dengan nilai *sillhouette coefficient* sebesar 0,50, yang termasuk dalam kategori struktur *cluster* yang baik.

REFERENSI

- [1] BPS Provinsi Sumatera Barat, *Provinisi Sumatera Barat dalam Angka 2023*. Padang: Padang, 2023.
- [2] F. S. N. Ramadhani, "Pengaruh Kondisi Demografi, Ketenagakerjaan, dan Ekonomi Terhadap pengangguran Terdidik di Indonesia," *J. Ilm. Mhs. FEB*, vol. 10, no. 2, 2022.
- [3] G. M. Prasetya and A. Sumanto, "Pengaruh Tingkat Pengangguran dan Tenaga Kerja Terhadap Kemiskinan Melalui Pertumbuhan Ekonomi," *IKINERJA J. Ekon. dan Manaj.*, vol. 19, no. 2, pp. 467–477, 2022.
- [4] Yuliani, M. Saleh, and A. Busari, "Faktor-faktor yang Mempengaruhi Penawaran Tenaga Kerja di Kota Samarinda," *J. Ilmu Ekon. Mulawarman*, vol. 3, no. 4, 2018.
- [5] M. F. Talokon *et al.*, "Faktor-faktor yang Mempengaruhi tingkat Pengangguran di Kota Tomohon," *J. Agribussines Rural Dev. (Jurnal Agribisnis dan Pengemb. Pedesaan)*, vol. 1, no. 2, 2019.
- [6] J. Supranto, *Analisis Multivariat: Arti dan Interpretasi*. Jakarta: Rineka Cipta, 2010.
- [7] P. Lopez, J. Carlos, and J. S. Perez, "An unsupervised pattern recognition methodology based on factor analysis and a genetic-DBSCAN algorithm to infer operational conditions from strain measurements in structural applications," *Chinese J. Aeronaut.*, vol. 34, no. 2, pp. 165–181, 2021, doi: 10.1016/j.cja.2020.09.035.
- [8] M. Hahsler, M. Piekenbrock, and D. Doran, "DBSCAN: Fast Density-based Clustering with R," *J. Stat. Softw.*, vol. 91, pp. 1–30, 2019.
- [9] M. M. Putri, C. Dewi, E. P. Siam, G. A. Wijayanti, N. Aulia, and R. Nooraeni, "Comparison of DBSCAN and K-Means Clustering for Grouping the Village Status in Central Java 2020," *J. Mat. Stat. dan Komputasi*, vol. 17, no. 3, pp. 394–404, 2021.
- [10] A. Yuwono, Y. Oslan, and D. Dwijono, "Implementasi Metode Density Based Spatial Clustering of Applications with Noise untuk Mencari Arah Penyebaran Wabah Demam Berdarah (Studi Kasus : Data Dinas Kesehatan Kodya Jogjakarta)," *J. EKSIS*, vol. 2, no. 1, 2015.
- [11] D. Safitri, T. Wuryandari, and R. Rahmawati, "Metode DBSCAN untuk Pengelompokkan Kabupaten/Kota di Provinsi Jawa Tengah Berdasarkan Produksi Padi Sawah dan Padi Ladang," *J. Stat. Univ. Muhammadiyah Semarang*, vol. 5, no. 1, 2017.
- [12] I. Hasan, *Pokok-Pokok Materi Statistik 1 (Statistika Deskriptif)*. Jakarta: PT Bumi Aksara, 2001.
- [13] Buaton, Sundari, and Maulita, "Clustering Tindak Kekerasan Terhadap Anak Menggunakan Algoritma K-Means dengan Perbandingan Jarak Kedekatan Manhattan City dan Euclidean," *Media Inf. Anal. dan Sist.*, vol. 1, no. 2, pp. 47–53, 2016.
- [14] R. Hidayati, A. Zubair, A. H. Pratama, and L. Indana, "Analisis Sillhouette Coefficient Pada 6 Perhitungan jarak K-Means Clustering," *Techno.Com*, vol. 20, no. 2, pp. 186–197, 2021.
- [15] L. Qadrini, "Metode K-Means dan DBSCAN pada Pengelompokan Data Dasar Kompetensi Laboratorium ITS Tahun 2017," *J Stat. J. Ilm. Teor. dan Apl. Stat.*, vol. 13, no. 2, pp. 5–11, 2020.