

Pengelompokan Kemiskinan di Provinsi Sulawesi Selatan Tahun 2023 dengan Metode *K-Means Clustering*

A. Elisha Wulandari^{1*}, Andi M. Alfin Baso², Belia Nurul Fajri³, Anisa Kalondeng⁴, Anna Islamiyati⁵, Abdullah Pannu⁶, Muhammad Fadil⁷, Alfian Akbar Vallarino⁸, Anugrah Nur Isnaeni Rahman⁹

¹²⁴⁵⁷⁹Departemen Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Hadanuddin, Makassar, 60294, Indonesia

³Departemen Ekonomi Pembangunan, Fakultas Ekonomi dan Bisnis, Universitas Hadanuddin, Makassar, 60294, Indonesia

⁶⁸Badan Pusat Statistik Provinsi Sulawesi Selatan, Makassar, 90125, Indonesia

*Corresponding author, email: aelishawulandari@gmail.com

Abstract

Poverty remains a significant social and economic issue in South Sulawesi Province. This study aims to classify districts/cities in South Sulawesi based on poverty levels using the *K-Means Clustering* method. The data used were obtained from the Central Bureau of Statistics (BPS) for 2023, including indicators such as the percentage of poor population, education level, and employment sector. The *Silhouette Index* method was applied to determine the optimal number of clusters. The results indicate that South Sulawesi is divided into two clusters, representing high and low poverty levels. The scatter plot further reveals that cluster 1 is more varied, while cluster 2 is more concentrated. These findings can serve as a foundation for formulating more targeted policies to reduce poverty in South Sulawesi.

Keywords: Poverty, *K-Means Clustering*, Clustering, South Sulawesi, *Silhouette Index*.

Abstrak

Kemiskinan merupakan salah satu permasalahan sosial dan ekonomi yang masih menjadi tantangan di Provinsi Sulawesi Selatan. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Sulawesi Selatan berdasarkan tingkat kemiskinan menggunakan metode *K-Means Clustering*. Data yang digunakan diperoleh dari Badan Pusat Statistik tahun 2023 dengan indikator seperti persentase penduduk miskin, tingkat pendidikan, dan sektor pekerjaan. Metode *Silhouette Index* digunakan untuk menentukan jumlah kluster optimal. Hasil analisis menunjukkan bahwa wilayah Sulawesi Selatan terbagi menjadi dua kluster, yaitu kluster dengan tingkat kemiskinan tinggi dan rendah. Scatter plot memperjelas bahwa kluster 1 lebih bervariasi, sedangkan kluster 2 lebih terkonsentrasi. Temuan ini dapat digunakan sebagai dasar dalam perumusan kebijakan yang lebih tepat sasaran untuk pengentasan kemiskinan di Sulawesi Selatan.

Kata Kunci: Kemiskinan, *K-Means Clustering*, Klusterisasi, Sulawesi Selatan, *Silhouette Index*.

1. Pendahuluan

Kemiskinan masih menjadi salah satu tantangan terbesar yang dihadapi oleh negara berkembang di dunia, termasuk Indonesia. Kemiskinan merupakan kondisi seseorang atau sekelompok orang yang tidak dapat memenuhi kebutuhan dasarnya seperti makanan, obat-obatan dan tempat tinggal untuk mempertahankan dan mengembangkan

kehidupannya [1]. Menurut Badan Pusat Statistik (BPS) Sulawesi Selatan, pada Maret 2023 persentase penduduk miskin di Sulawesi Selatan mencapai 8.70%, meningkat 0.04% dibandingkan September 2022. Fenomena ini menunjukkan perlunya perhatian khusus dalam memahami dan menangani kemiskinan serta ketimpangan yang dialami oleh masyarakat di Sulawesi Selatan.

Menurut Badan Pusat Statistik (BPS) Sulawesi Selatan, tingkat kemiskinan di Sulawesi Selatan bervariasi pada setiap kabupaten dan kota. Kabupaten Jeneponto, Luwu Timur, dan Enrekang termasuk daerah dengan angka kemiskinan yang lebih tinggi dibandingkan daerah lain. Secara umum, penyebab peningkatan kemiskinan di Sulawesi Selatan meliputi keterbatasan akses terhadap pendidikan, minimnya kesempatan kerja, serta tingginya ketergantungan pada sektor pertanian yang rentan terhadap perubahan cuaca dan bencana alam. Selain itu, rendahnya akses terhadap layanan kesehatan dan keterbatasan infrastruktur di beberapa daerah pedesaan turut memperparah kondisi kemiskinan[2]. Sehingga, dampak yang diperoleh akan berkaitan erat dengan peningkatan kemiskinan di setiap daerah.

Kondisi ini tentunya dipengaruhi oleh perbedaan karakteristik di setiap daerah, seperti tingkat pembangunan infrastruktur, akses terhadap layanan dasar, serta struktur ekonomi yang dominan di wilayah tersebut. Setiap kabupaten/kota memiliki pola kemiskinan yang unik, sehingga diperlukan metode analisis yang mampu mengelompokkan daerah-daerah dengan karakteristik serupa. Salah satu pendekatan yang dapat digunakan adalah dengan metode klasterisasi[3].

Klasterisasi atau *clustering*, adalah metode dalam analisis data yang bertujuan mengelompokkan sekumpulan objek atau data berdasarkan kemiripan karakteristiknya. Objek-objek yang memiliki kesamaan tinggi akan dikelompokkan dalam satu klaster, sementara objek yang berbeda akan ditempatkan di klaster lain. Tujuan utama klasterisasi adalah memaksimalkan keseragaman dalam klaster dan meminimalkan kesamaan antar klaster[4]. Pendekatan klasterisasi telah banyak digunakan untuk memahami pola kemiskinan di berbagai wilayah. Penelitian oleh Latupeirissa dkk. (2022) mengembangkan metode *K-Means* untuk mengelompokkan kabupaten/kota di Provinsi Maluku berdasarkan tingkat kemiskinan dan menemukan tiga klaster utama[5]. Selain itu, penelitian oleh Tadi & Ningsi (2023) di Provinsi Banten menunjukkan bahwa metode *K-Means* dengan pengujian *Silhouette Index* menghasilkan klaster optimal yang dapat membantu dalam memahami distribusi kemiskinan di tingkat kabupaten/kota[6].

Metode *K-Means Clustering* merupakan teknik yang umum digunakan dalam analisis klasterisasi data kemiskinan. Algoritma ini membagi data ke dalam beberapa kelompok berdasarkan kesamaan karakteristiknya, sehingga dapat membantu dalam mengidentifikasi daerah dengan tingkat kemiskinan yang lebih tinggi atau lebih rendah[7]. Studi oleh Widyadhana dkk. (2021) mengembangkan perbandingan metode *K-Means* dan *Average Linkage* dalam pengelompokan kemiskinan di Jawa Tengah dan menemukan bahwa metode *K-Means* memiliki keunggulan dalam pengolahan data

berskala besar[8]. Oleh karena itu, metode yang tepat untuk Pengelompokan Kemiskinan di Provinsi Sulawesi Selatan Tahun 2023 adalah dengan metode *K-Means Clustering*.

2. Material dan Metode

Penelitian ini memanfaatkan sumber data sekunder, yaitu data yang diperoleh dari Badan Pusat Statistik Provinsi Sulawesi Selatan melalui situs <https://sulsel.bps.go.id>. Data dalam penelitian ini memanfaatkan data kemiskinan provinsi Sulawesi Selatan tahun 2023. Indikator kemiskinan yang digunakan pada penelitian ini yaitu Persentase Penduduk Miskin (X_1), Persentase Penduduk Miskin Usia 15 Tahun ke Atas Menurut dan Pendidikan yang Ditamatkan dibawah SD (X_2), Persentase Penduduk Miskin Usia 15 Tahun ke Atas Menurut dan Pendidikan yang Ditamatkan Tamat SD/SMP (X_3), Persentase Penduduk Miskin Usia 15 Tahun ke Atas Menurut dan Pendidikan yang Ditamatkan tamat SMA (X_4), Persentase Penduduk Miskin Usia 15 Tahun ke Atas Tidak Bekerja (X_5), Persentase Penduduk Miskin Usia 15 Tahun ke Atas Bekerja di Sektor Informal (X_6), Persentase Penduduk Miskin Usia 15 Tahun ke Atas Bekerja di Sektor Formal (X_7), Persentase Penduduk Miskin Usia 15 Tahun ke Atas Bekerja di Sektor Pertanian (X_8), Persentase Penduduk Miskin Usia 15 Tahun ke Atas Bekerja Bukan di Sektor Pertanian (X_9). Metode yang digunakan yaitu algoritma *K-means*.

Algoritma *K-Means* merupakan salah satu metode pengelompokan data yang membagi objek ke dalam k kluster berdasarkan kesamaan karakteristiknya. Dalam penerapannya, pengguna perlu menentukan nilai k terlebih dahulu, yang dapat dilakukan melalui metode *Silhouette Index* untuk menemukan jumlah kluster optimal. Prinsip dasar *K-Means* adalah mengelompokkan data dengan meminimalkan perbedaan dalam satu kluster dan memaksimalkan perbedaan antar kluster. Pengelompokan ini didasarkan pada jarak, di mana objek yang memiliki jarak lebih dekat dianggap lebih mirip dibandingkan dengan objek yang lebih jauh. Berikut ini adalah langkah-langkah dalam algoritma *K-Means*:

1. Tentukan jumlah kluster k .
2. Tentukan nilai pusat (*centroid*).

Penentuan *centroid* tahap awal dilakukan secara acak, sedangkan pada tahap iterasi digunakan rumus seperti pada persamaan (1):

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj}^2 \quad (1)$$

dengan:

V_{ij} : *centroid* rata-rata kluster ke- i untuk variabel ke- j

N_i : Jumlah anggota kluster ke- i

i, k : indeks kluster

j : indeks variabel

X_{kj} : nilai data ke- k variabel ke- j untuk kluster

3. Pada masing-masing record, hitung jarak terdekat dengan *centroid* menggunakan teori jarak *Euclidean Distance*, seperti pada persamaan (2):

$$De = \sqrt{(xi - si)^2 + (yi - ti)^2} \quad (2)$$

dengan:

De : *Eulidean Distance*

I : Banyak objek

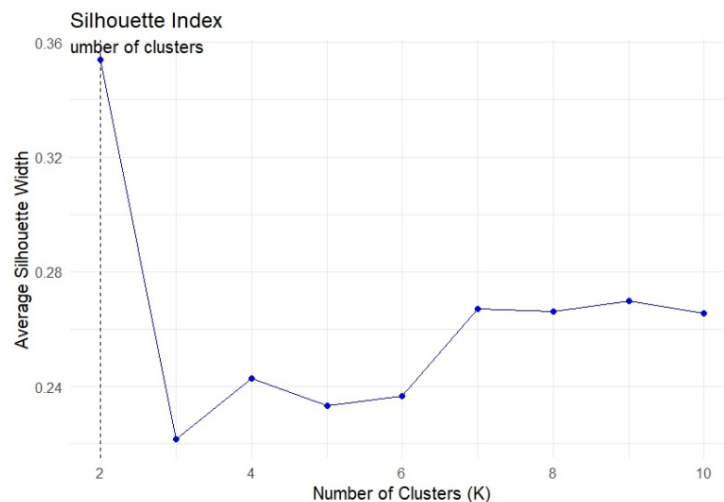
(x,y): Koordinat objek

(s,t) : Koordinat *centroid*

4. Kelompokkan objek berdasarkan jarak ke centroid terdekat. Ulangi langkah ke-2, lakukan iterasi hingga nilai *centroid* bernilai optimal.

3. Hasil dan Diskusi

Menentukan klaster terbaik dalam penelitian ini dilakukan dengan pengujian *Silhouette Index*. Metode ini membantu menilai seberapa baik setiap data berada dalam klasternya dibandingkan dengan klaster lain, sehingga ditemukan pembagian kelompok yang paling optimal. Hasil pengujian *Silhouette Index* ditunjukkan pada Gambar 1.



Gambar 1. *Silhouette Index*

Gambar 1 menunjukkan bahwa dari klaster 1 hingga 9, klaster 2 memiliki nilai tertinggi, yang berarti pembagian menjadi 2 klaster adalah yang paling optimal. Oleh karena itu, dilakukan pembentukan 2 klaster untuk melihat bagaimana provinsi dikelompokkan berdasarkan hasil ini.

Tabel 1. Nilai *Centroid* Klaster

K	X₁	X₂	X₃	X₄	X₅	X₆	X₇	X₈	X₉
1	6.93	20.07	53.98	25.96	44.57	39.34	16.1	37.22	18.22
2	6.73	30.7	48.83	20.47	53.42	33.62	12.97	34.38	12.21

Tabel 2. Nilai Kategori *Centroid* Klaster

K	X₁	X₂	X₃	X₄	X₅	X₆	X₇	X₈	X₉
1	8.18	20.15	50.45	29.38	55.41	24.23	20.36	17	27.6
2	10.06	27.2	48.23	24.69	46.11	40.77	13.09	37.53	16.22

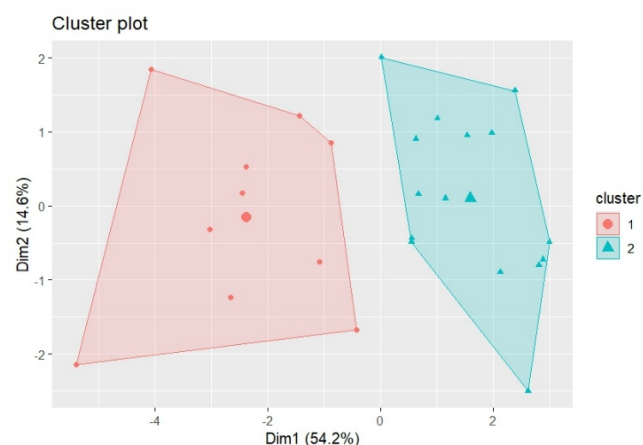
Tabel 3. Pembagian Klaster

No	Kabupaten/Kota	Klaster
1	Kepulauan Selayar	1
2	Bulukumba	1
3	Bantaeng	2
4	Jeneponto	2
5	Takalar	1
6	Gowa	2
7	Sinjai	2
8	Maros	2
9	Pangkajene dan Kepulauan	1
10	Barru	1
11	Bone	2
12	Soppeng	2
13	Wajo	2
14	Sidenreng Rappang	1
15	Pinrang	1
16	Enrekang	2
17	Luwu	2
18	Tana Toraja	2
19	Luwu Utara	2
20	Luwu Timur	2
21	Toraja Utara	2
22	Kota Makassar	1
23	Kota Parepare	1
24	Kota Palopo	1

Tabel 1 menunjukkan bahwa klaster kemiskinan terbagi menjadi 2 kategori, yaitu kategori kemiskinan rendah dan kategori kemiskinan tinggi. Hal ini dapat dilihat dari nilai

kategori Centroid klasternya dimana kategori 1 merupakan kemiskinan rendah dan kategori 2 merupakan kemiskinan yang tinggi. Sehingga, dari Tabel 3 diperoleh pembagian klaster berikut:

- a. Pada klaster 1 terdiri dari Bali, Bangka Belitung, Banten, DKI Jakarta, Jawa Barat, Kalimantan Timur, Kepulauan Riau, dan Sulawesi Utara. Berdasarkan tabel 2 klaster 1 termasuk ke dalam kategori kemiskinan yang rendah.
- b. Klaster 2 terdiri dari Kabupaten Bantaeng, Jeneponto, Gowa, Sinjai, Maros, Bone, Soppeng, Wajo, Enrekang, Luwu, Tana Toraja, Luwu Utara, Luwu Timur, dan Toraja Utara. Berdasarkan tabel 2 klaster 1 termasuk ke dalam kategori kemiskinan yang tinggi.



Gambar 2. Model *Scatter plot*

Gambar 2 menunjukkan bahwa *scatter plot* di atas terbagi menjadi dua klaster yang ditandai dengan warna merah (klaster 1) dan biru muda (klaster 2). Klaster 1 terletak di sisi kiri grafik dengan sebaran lebih bervariasi, sementara klaster 2 lebih terkonsentrasi di sisi kanan, menandakan pola karakteristik yang lebih seragam. Garis cembung (*convex hull*) di sekitar setiap klaster memperjelas pemisahan antar kelompok dalam pola distribusi data. Dalam pengelompokan tingkat kemiskinan di Sulawesi Selatan, klaster 1 menunjukkan tingkat kemiskinan yang rendah sedangkan klaster 2 menunjukkan tingkat kemiskinan yang lebih tinggi.

4. Kesimpulan

Berdasarkan hasil analisis pengelompokan kemiskinan dengan metode *K-Means*, 24 kabupaten/kota di Sulawesi Selatan tahun 2023 dikelompokkan ke dalam dua klaster. Klaster 1 mencakup daerah dengan tingkat kemiskinan rendah, sedangkan klaster 2 mencakup daerah dengan tingkat kemiskinan tinggi. Hasil ini didukung oleh pengujian *Silhouette Index* yang menunjukkan bahwa pembagian menjadi dua klaster adalah yang paling optimal. *Scatter plot* juga memperjelas bahwa klaster 1 memiliki sebaran lebih bervariasi, sedangkan klaster 2 lebih terkonsentrasi, menandakan pola kemiskinan yang

lebih seragam dalam kelompoknya. Pada klaster 1 merupakan kelompok kemiskinan kategori rendah yang terdiri dari Kabupaten Kepulauan Selayar, Bulukumba, Takalar, Pangkajene dan Kepulauan, Barru, Sidenreng Rappang, Pinrang, dan Kota Makassar, Pare-Pare, Palopo. Sedangkan, pada klaster 2 merupakan kelompok kemiskinan kategori tinggi yang terdiri dari Kabupaten Bantaeng, Jeneponto, Gowa, Sinjai, Maros, Bone, Soppeng, Wajo, Enrekang, Luwu, Tana Toraja, Luwu Utara, Luwu Timur, dan Toraja Utara.

Daftar Pustaka

- [1] Hardinandar, F. Determinan kemiskinan (studi kasus 29 kota/kabupaten di Provinsi Papua). *Jurnal REP (Riset Ekonomi Pembangunan)*, 4(1), 1-12, 2019.
- [2] Rizkiyanto, M. Analisis Pengelompokan Kemiskinan di Indonesia Menggunakan Metode K-Means. Prosiding Seminar Kecerdasan Artifisial, Sains Data, dan Pendidikan Masa Depan, *PROKASDADIK*, 1, 102-106, 2023.
- [3] Irfan, & Faizal, L. Implementasi Algoritma K-Means untuk Mengelompokkan Data Tingkat Kemiskinan di Sulawesi Selatan Berdasarkan Kota/Kabupaten. *Jurnal Ilmiah Sistem Informasi dan Teknik Informatika (JISTI)*, 7(2), 261-269, 2024.
- [4] Arwani, I. Optimasi Proses Klasterisasi di MySQL DBMS dengan Mengintegrasikan Algoritme MIC-Kmeans Menggunakan Bahasa SQL dalam Stored Procedure. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 7(2), 391-398, 2020.
- [5] Latupeirissa, S. J., Lewaherilla, N., & Hiariey, A. Pengelompokan Kabupaten/Kota di Provinsi Maluku Berdasarkan Data Kemiskinan Tahun 2021 Menggunakan Metode K-Means Cluster. *Jurnal Variance*, 4(1), 15-22, 2022.
- [6] Tadi, M., & Ningsi, B. A. Analisis Klaster Kemiskinan Kabupaten Kota di Provinsi Banten Menggunakan Metode K-Means. *Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, 4(1), 374-385, 2023.
- [7] Aulia, S. Klasterisasi Pola Penjualan Pestisida Menggunakan Metode K-Means Clustering (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja). *Djtechno: Jurnal Teknologi Informasi*, 1(1), 1-5, 2021.
- [8] Widyadhana, D., Hastuti, R. B., Kharisudin, I., & Fauzi, F. Perbandingan Analisis Klaster K-Means dan Average Linkage untuk Pengklasteran Kemiskinan di Provinsi Jawa Tengah. *PRISMA, Prosiding Seminar Nasional Matematika*, 4, 584-594, 2021.