E-Jurnal Matematika Vol. 15(4), November 2026, pp. 234-238 ISSN: 2303-1751
DOI: https://doi.org/10.24843/MTK.2026.v15.i04.p526

PENGELOMPOKAN PROVINSI DI INDONESIA BERDASARKAN INDIKATOR SDGS KESEJAHTERAAN RAKYAT

I Made Prima Mahardika[1§], I Komang Gde Sukarsa[2], Ketut Jayanegara[3]

[1]Prodi Matematika, Fakultas MIPA-Universitas Udayana [Email: prima.mahardika.234@gmail.com]
[2]Prodi Matematika, Fakultas MIPA-Universitas Udayana [Email: gedesukarsa@unud.ac.id]
[3]Prodi Matematika, Fakultas MIPA-Universitas Udayana [Email: ktjayanegara@unud.ac.id]
[§]*Corresponding Author*

ABSTRACT

Differences in the achievement of Sustainable Development Goals (SDGs) indicators for improving social welfare across provinces in Indonesia indicate the need to map the characteristics of each region as a basis for formulating more targeted development policies. This study aims to classify Indonesian provinces into groups based on SDGs social welfare indicators using the Ward hierarchical cluster analysis method and measuring distances using the Euclidean distance method. The data used are secondary data obtained from Statistics Indonesia (Badan Pusat Statistik/BPS), covering 38 provinces and 10 social welfare indicators. The results show that the levels of social welfare vary across provinces in Indonesia. Based on the dendrogram at an approximately 80% similarity level, 10 clusters of provinces with distinct characteristics were identified. The clustering results are expected to serve as a reference for formulating more effective development policies that are tailored to the characteristics of each province.

Keywords: *Sustainable Development Goals (SDGs), social welfare, cluster analysis, Ward's method, Euclidean distance*

1. PENDAHULUAN

Pembangunan merupakan proses yang bertujuan untuk meningkatkan kualitas hidup dan kesejahteraan masyarakat secara berkelanjutan. Dalam mewujudkannya, Perserikatan Bangsa-Bangsa (PBB) menetapkan Sustainable Develop-ment Goals (SDGs) sebagai agenda pem-bangunan global hingga tahun 2030 yang mengintegrasikan dimensi sosial, ekonomi, dan lingkungan (Wahyuningsih, 2017).

Indonesia berkomitmen melaksanakan SDGs melalui penyesuaian indikator global ke dalam indikator nasional sebagaimana diatur dalam Peraturan Presiden Republik Indonesia Nomor 111 Tahun 2022 tentang Pelaksanaan Pencapaian Tujuan Pembangunan Berkelanjutan. Indikator SDGs mampu memberikan gambaran kesejah-teraan masyarakat secara komprehensif karena mencakup berbagai dimensi pembangunan, seperti kesehatan, pendidikan, konsumsi pangan, perumahan, serta gender dan anak.

Sebagai negara dengan 38 provinsi yang memiliki karakteristik pembangunan yang beragam, Indonesia menunjukkan tingkat kesejah-teraan masyarakat yang berbeda antar-wilayah. Data Indikator SDGs Kesejahteraan Rakyat Tahun 2025 menunjukkan adanya variasi capaian kesejahteraan antarprovinsi sehingga diperlukan kebijakan pembangunan yang disesuaikan dengan karakteristik masing-masing wilayah.

Untuk mengidentifikasi kelompok provinsi yang memiliki karakteristik kesejahteraan yang serupa, penelitian ini menggunakan analisis cluster, yaitu metode statistik yang mengelompokkan objek berdasarkan tingkat kemiripan karakteristiknya. Metode yang digunakan adalah *Hierarchical Cluster Analysis* metode Ward dengan ukuran jarak Euclidean karena mampu menghasilkan cluster yang relatif homogen dan dapat memudahkan interpretasi hasil analisis melalui dendrogram (Sidebang, 2024).

Penelitian terdahulu telah banyak menerapkan analisis cluster dalam pengelompokan wilayah berdasarkan indikator sosial dan kependudukan. Fathia (2016) mengelompokkan kecamatan di Kabupaten Semarang berdasarkan potensi desa dengan

membandingkan metode *single linkage* dan *ward*. Kedua metode tersebut menghasilkan empat *cluster*, namun metode *single linkage* terbukti menghasilkan kelompok yang lebih heterogen berdasarkan nilai *R-Squared*. Sementara itu, Harnanto (2017) me-nerapkan analisis *cluster* dengan metode *ward* dengan jarak Euclidean untuk mengelompokkan kabupaten/kota di Jawa Tengah berdasarkan jenis alat kontrasepsi. Hasil analisis membagi wilayah Jawa Tengah ke dalam tiga *cluster*, dengan interpretasi *centroid* menunjukkan bahwa suntikan merupakan alat kontrasepsi yang paling dominan digunakan.

Namun penelitian mengenai pengelompokan provinsi di Indonesia berdasarkan indikator SDGs kesejahteraan rakyat masih terbatas. Oleh karena itu, penelitian ini bertujuan untuk mengelompokkan 38 provinsi di Indonesia ber-dasarkan indikator SDGs kesejahteraan rakyat tahun 2025 menggunakan metode Ward. Hasil dari penelitian ini diharapkan dapat memberikan gambaran karakteristik kesejahteraan masyarakat pada setiap kelompok provinsi sebagai bahan pertimbangan dalam penyusunan kebijakan pembangunan yang lebih tepat sasaran.

2. METODE PENELITIAN

Penelitian ini menggunakan data sekunder yang mencakup seluruh provinsi di Indonesia tahun 2025 dengan objek amatan sebanyak 38 provinsi. Variabel yang digunakan dalam penelitian kali ini adalah data indikator SDGs kesejahteraan rakyat yang meliputi ketidak-cukupan konsumsi pangan (X_1), kerawanan pangan (X_2), tempat persalinan (X_3), unmet need pelayanan kesehatan (X_4), angka melek aksara (X_5), angka anak tidak sekolah (X_6), kepemilikan rumah (X_7), akses sanitasi layak (X_8), usia kawin pertama (X_9), dan kepemilikan akta kelahiran (X_{10}). Proses pengolahan data dilakukan dengan menggunakan perangkat lunak Minitab versi 22.1.

1. Analisis Deskriptif

Analisis deskriptif dilakukan bertujuan untuk memberikan sebuah gambaran umum mengenai karakteristik data penelitian sebelum dilakukan proses pengelompokan menggunakan metode Analisis Hirarki metode Ward.

2. Analisis Cluster

Analisis cluster merupakan metode statistik multivariat yang digunakan untuk menge-lompokkan objek berdasarkan tingkat kemiripan karakteristiknya. Metode ini bertujuan meng-hasilkan kelompok yang homogen dalam cluster dan heterogen antarcluster sehingga memudah-kan identifikasi karakteristik masing-masing kelompok yang terbentuk (Simamora, 2005).

Secara umum, metode analisis cluster terbagi menjadi metode hierarki dan nonhierarki, di mana salah satu pendekatan hirarki yang umum digunakan adalah metode *agglomerative* yang menggabungkan objek secara bertahap ber-dasarkan jarak terdekat hingga seluruhnya menyatu dalam satu kelompok. Metode hirarki *agglomerative* memiliki lima teknik utama, yaitu *single linkage*, *complete linkage*, *average linkage*, *Ward's method*, dan metode *centroid* (Hardle dan Simar, 2007).

3. Ukuran Kedekatan

Ukuran kedekatan dalam analisis cluster digunakan untuk mengukur tingkat kemiripan antarobjek pengamatan. Semakin kecil jarak antara dua objek, semakin tinggi tingkat kemiripannya sehingga cenderung dikelompok-kan dalam cluster yang sama (Usman dan Nurdin, 2013). Salah satu ukuran jarak yang umum digunakan adalah jarak Euclidean, terutama untuk data berskala kontinu. Jarak Euclidean menghitung jarak antara dua objek berdasarkan akar kuadrat dari jumlah kuadrat selisih nilai masing-masing variabel, sehingga mampu meng-gambarkan tingkat kedekatan antarobjek secara kuantitatif. Jarak Euclidean digunakan karena data berupa variabel kuantitatif, sehingga per-bedaan antarprovinsi dapat dihitung langsung. Semakin kecil jarak antarobjeknya, semakin mirip karakteristik kedua provinsi. Rumus Jarak Euclidean dapat didefinisikan sebagai berikut

$$d(i,j) = d_{ij} = \sqrt{\sum_{k=1}^{p} \left(x_{ik} - y_{jk}\right)^2} \quad (1)$$

Keterangan:
d_{ij} = jarak antara objek ke-i dan objek ke-j
p = jumlah variabel klaster
x_{ik} = data dari objek ke-i pada variabel ke-k
y_{jk} = data dari objek ke-j pada variabel ke-k

4. Metode Ward

Metode Ward merupakan salah satu metode hierarchical clustering dengan pendekatan *agglomerative* yang melakukan pengelompokan objek secara bertahap berdasarkan tingkat kemiripannya. Metode ini bekerja dengan menggabungkan dua cluster yang menghasilkan peningkatan variasi dalam cluster (*Sum of Square Error*/SSE) paling kecil, sehingga mampu me-minimalkan variasi yang ada dalam cluster dan menghasilkan kelompok yang relatif homogen. Adapun rumus yang digunakan untuk meng-hitung nilai SSE adalah sebagai berikut (Wibowo, 2020):

$$SSE = \sum_{i=1}^{N} (y_i - \bar{y})'(y_i - \bar{y}) \qquad (2)$$

Keterangan:
y_i = Vektor kolom nilai rata-rata objek i
i = $1,2,3,\dots,n$
$\bar{y}$ = Vektor kolom rata-rata objek dalam cluster
N = Banyaknya objek

Hasil pengelompokan divisualisasikan dalam bentuk dendrogram yang dapat digunakan untuk melihat hubungan kedekatan antarobjek serta menentukan jumlah cluster yang terbentuk.

3. HASIL DAN PEMBAHASAN

3.1 Analisis Deskriptif

Analisis deskriptif dilakukan untuk memberikan Gambaran awal mengenai kondisi 38 provinsi di Indonesia berdasarkan indikator SDGs Kesejahteraan Rakyat tahun 2025. Hasil analisis menunjukkan adanya keragaman capaian indikator antarprovinsi yang ditunjukkan oleh nilai rata-rata, simpangan baku, nilai minimum, dan maksimum pada masing-masing variabel.

Indikator rumah tangga dengan akses sanitasi layak memiliki tingkat keragaman tertinggi, sedangkan *unmet need* pelayanan kesehatan menunjukkan keragaman yang relatif rendah. Selain itu, rentang nilai yang cukup besar pada beberapa indikator mengindikasikan adanya perbedaan kondisi kesejahteraan rakyat antar-provinsi, sehingga diperlukan adanya pengelompokan untuk mengidentifikasi provinsi dengan karakteristik yang serupa.

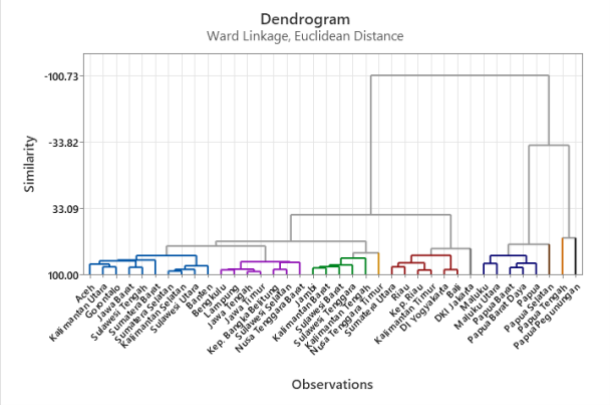
Tabel 1. Statistik Deskriptif Indikator SDGs Kesejahteraan Rakyat

ID	Mean	StDev	Min	Max
X1	12.2021	8.41077	2.67	32.3
X2	5.21395	2.96727	1.28	13.69
X3	88.7792	13.1191	38.13	100
X4	4.93974	1.56324	2.32	8.51
X5	96.5834	5.39518	68.32	99.8
X6	16.7876	9.95295	3.79	53.37
X7	84.9334	7.80784	54.54	96.61
X8	82.9582	14.9602	16.34	98.2
X9	5.70053	3.24491	0.79	12.94
X10	90.6474	12.3724	34.31	98.52

Dikarenakan sebagian besar indikator berbentuk persentase atau rasio, maka data dapat langsung dianalisis dengan menggunakan jarak Euclidean. Selain itu, penggunaan data asli dipertahankan agar variasi dan rentang capaian asli antarprovinsi tetap tercermin dalam hasil pengelompokan.

3.2 Analisis Cluster

Analisis cluster dilakukan menggunakan *Hierarchical Cluster Analysis* dengan metode Ward dan ukuran jarak Euclidean menggunakan perangkat lunak Minitab. Hasil analisis disajikan dalam bentuk dendrogram.



Gambar 1. Dendrogram Hasil *Hierarchical Cluster Analysis Metode Ward*

Penentuan jumlah cluster lalu dilakukan dengan memotong dendrogram pada tingkat kemiripan (*similarity level*) sekitar 80%. Pemotongan pada tingkat kemiripan 80% meng-hasilkan 10 kelompok yang dinilai ideal karena tidak terlalu sedikit sehingga menggabungkan provinsi dengan karakteristik yang sangat berbeda, dan tidak terlalu banyak

sehingga hasil pengelompokan tetap efisien untuk dianalisis.

Tabel 2. Pengelompokkan Cluster Provinsi di Indonesia

Cluster	Provinsi
1	Aceh, Kalimantan Utara, Gorontalo, Jawa Barat, Sulawesi Tengah, Sumatera Barat, Sumatera Selatan, Kalimantan Selatan, Sulawesi Utara, Banten
2	Sumatera Utara, Riau, Kep. Riau, Kalimantan Timur, DI Yogyakarta, Bali
3	Jambi, Kalimantan Barat, Sulawesi Barat, Sulawesi Tenggara, Kalimantan Tengah
4	Bengkulu, Lampung, Jawa Tengah, Jawa Timur, Kep. Bangka Belitung, Sulawesi Selatan, Nusa Tenggara Barat
5	DKI Jakarta
6	Nusa Tenggara Timur
7	Maluku, Maluku Utara, Papua Barat, Papua Barat Daya, Papua
8	Papua Selatan
9	Papua Tengah
10	Papua Pegunungan

Pengelompokan tersebut menunjukkan bahwa provinsi dalam satu cluster memiliki kondisi yang relatif homogen dibandingkan dengan provinsi pada cluster lainnya.



Gambar 2. Pemetaan Cluster Provinsi di Indonesia

3.3 Karakteristik Masing Masing Cluster

Karakteristik masing-masing cluster lalu dianalisis berdasarkan nilai pusat (*centroid*) yang dimiliki oleh setiap cluster yang kemudian dibandingkan dengan nilai pusat keseluruhan (*Grand Centroid*).

Dari Tabel 3, Cluster 1 sampai Cluster 4 memiliki karakteristik indikator SDGs yang mendekati rata-rata nasional, sedangkan Cluster 5 yang terdiri atas DKI Jakarta menunjukkan capaian kesejahteraan yang paling tinggi, sedangkan Cluster 6 hingga Cluster 10 menunjukkan karakteristik kesejahteraan yang relatif lebih rendah pada beberapa indikator SDGs. Perbedaan karakteristik tersebut terlihat pada indikator ketahanan pangan, pelayanan kesehatan, pendidikan, kepemilikan rumah, sanitasi layak, perkawinan usia dini, dan kepemilikan akta kelahiran.

Tabel 3. Tabel Cluster Centroid

ID	Cluster										GC
	1	2	3	4	5	6	7	8	9	10	
X1	8.1	7.9	9.8	8.1	3.2	12	25	28	32	29	12.2
X2	4.8	3.3	3.7	3.9	2.5	14	9.2	12	4.7	6.7	5.21
X3	95	97	85	98	100	90	72	81	72	38	88.8
X4	5.4	3.8	6.1	5.7	2.3	8.2	4	4.6	2.4	3.6	4.94
X5	99	98	97	96	100	96	99	95	86	68	96.6
X6	16	9.4	19	16	6.2	18	12	30	53	51	16.8
X7	84	79	90	91	55	92	85	78	90	97	84.9
X8	83	93	87	90	94	81	80	65	40	16	83
X9	5.5	2.1	8.4	6.6	0.9	5.1	6.7	13	2.6	6.4	5.7
X10	95	96	95	96	99	78	87	73	57	34	90.6

4. KESIMPULAN DAN SARAN

Penelitian ini membuat 10 kelompok provinsi di Indonesia berdasarkan indikator SDGs tentang kesejahteraan rakyat pada tahun 2025 dengan menggunakan metode *Hierarchical Cluster Analysis* dengan metode Ward, di mana jarak diukur dengan cara Euclidean. Hasil penge-lompokan menunjukkan ada perbedaan dalam tingkat kesejahteraan rakyat di antara provinsi. Cluster 1 hingga Cluster 4 memiliki tingkat ke-sejahteraan yang hampir sama dengan rata-rata nasional. DKI Jakarta menunjukkan hasil kesejahteraan yang lebih baik dibandingkan beberapa cluster lainnya. Namun, beberapa cluster masih memerlukan perhatian lebih terutama pada beberapa indikator kesejahteraan rakyat. Hasil ini diharapkan bisa dipakai sebagai pertimbangan dalam membuat kebijakan pembangunan yang lebih tepat sasaran sesuai dengan karakteristik masing-masing cluster. Penelitian berikutnya bisa juga memikirkan untuk menggunakan metode pengelompokan atau menambahkan indikator yang lebih luas agar hasil pengelompokan menjadi lebih jelas dan dalam.

DAFTAR PUSTAKA

Badan Pusat Statistik. (2025) *Indikator SDGs Kesejahteraan Rakyat 2025*. Jakarta: Badan Pusat Statistik.

Fathia, A.N., Rahmawati, R. and Tarno, (2016). ‘Analisis Klaster Kecamatan di Kabupaten Semarang Berdasarkan Potensi Desa Menggunakan Metode Ward dan Single Linkage’. *Jurnal Gaussian*, 5(4), pp.801–810.

Hardle, W. dan Simar, L. (2007) *Applied Multivariate Statistical Analysis*. 2nd

Edition. New York: Springer Berlin Heidelberg.

Harnanto, Y.I., Rusgiyono, A. dan Wuryandari, T. (2017) 'Penerapan Analisis Klaster Metode Ward Terhadap Kabupaten/Kota di Jawa Tengah Berdasarkan Pengguna Alat Kontrasepsi', *Jurnal Gaussian*, 6(4), pp. 528–537.

Republik Indonesia. (2022) *Peraturan Presiden Republik Indonesia Nomor 111 Tahun 2022 tentang Pelaksanaan Pencapaian Tujuan Pembangunan Berkelanjutan*. Jakarta: Sekretariat Negara Republik Indonesia.

Sidebang, A.M.A. (2024) 'Pemetaan Kemiskinan Digital Kabupaten/Kota di Sumatera Utara Menggunakan Ward Hierarchical Clustering', *Journal of Analytical Research, Statistics and Computation*, 3(2), pp. 23–39.

Simamora, B. (2005) *Analisis Multivariat Pemasaran*. Jakarta: Gramedia Pustaka Utama.

Usman, H. dan Nurdin, S. (2013) *Aplikasi Teknik Multivariate untuk Riset Pemasaran*. Jakarta: PT Raja Grafindo Persada.

Wahyuningsih. (2017) 'Millenium Development Goals (MDGs) dan Sustainable Development Goals (SDGs) dalam Kesejahteraan Sosial', *Bisma: Jurnal Bisnis dan Manajemen*, 11(3), pp. 390–39

Wibowo, R.A. (2020) 'Simulasi Pemilihan Metode Analisis Cluster Hirarki Agglomerative Terbaik antara Average Linkage dan Ward pada Data yang Mengandung Masalah Multikolinearitas'. *Skripsi*. Bandar Lampung: Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.