



INNOVATIVE: Journal Of Social Science Research

Volume 3 Nomor 4 Tahun 2023 Page 8859-8872

E-ISSN 2807-4238 and P-ISSN 2807-4246

Website: <https://j-innovative.org/index.php/Innovative>

Pengelompokan Provinsi di Indonesia Menurut Indikator Indeks Kebahagiaan Menggunakan Metode *Average Linkage*

Ni Wayan Rita Damayanthi^{1✉}, Ni Luh Putu Suciptawati², Ketut Jayanegara³,
Eka N. Kencana⁴

Universitas Udayana, Bukit Jimbaran, Bali

Email: ritadamayanthi05@gmail.com^{1✉}

Abstrak

Analisis kluster merupakan teknik peubah ganda yang bertujuan untuk mengelompokkan objek-objek menjadi beberapa kelompok berdasarkan karakteristik yang dimiliki oleh objek tersebut. Penelitian ini bertujuan untuk mengetahui hasil pengklasteran 34 provinsi Indonesia berdasarkan indikator indeks kebahagiaan tahun 2021 menggunakan metode average linkage. Penelitian ini menggunakan ukuran jarak Minkowski dan tujuh variabel indikator indeks kebahagiaan. Berdasarkan hasil penelitian diketahui bahwa pengklasteran dengan metode average linkage membentuk empat kluster dengan nilai akurasi sebesar 53,27 persen. Kluster 1 terdiri dari 24 provinsi, kluster 2 beranggotakan tujuh provinsi, kluster 3 terdiri dari dua provinsi, dan kluster 4 beranggotakan satu provinsi yaitu Provinsi Gorontalo.

Kata Kunci: *analisis kluster; average linkage; indeks kebahagiaan.*

Abstract

Cluster analysis is a multivariate technique aimed at grouping objects into several clusters based on the characteristics possessed by these objects. This study aims to determine the clustering results of the 34 provinces of Indonesia based on the indicators of the happiness for the year 2021 using the average linkage method. The study employs the Minkowski distance measure and seven indicator variables of the happiness index. According to the research results, it is revealed that clustering with the average linkage method forms four clusters with an accuracy value of 53,27 percent. Cluster 1 consists of 24 provinces, cluster 2 includes seven provinces, cluster 3 consists of two provinces, and cluster 4 with one province, namely Gorontalo Province.

Keywords: *cluster analysis; average linkage; happiness index.*

PENDAHULUAN

Analisis multivariat merupakan pengembangan dari analisis univariat dan bivariat yang mengacu pada berbagai teknik statistik yang secara bersamaan mengamati beberapa pengukuran pada individu atau objek yang diteliti. Analisis multivariat dapat digunakan untuk menyederhanakan atau mereduksi data serta mengelompokkan data, salah satunya menggunakan analisis kluster (*cluster analysis*) (Hair et al., 2018). Analisis kluster melakukan pengelompokan terhadap objek-objek berdasarkan kesamaan ciri-ciri yang dimiliki oleh objek tersebut. Objek yang memiliki tingkat kemiripan tinggi akan berada dalam satu kluster, sedangkan objek yang tingkat kemiripannya rendah akan berbeda kluster atau tidak berada dalam satu kluster (Mattjik & Sumertajaya, 2011).

Hair et al. (2018) menyatakan terdapat dua metode klusterisasi, yaitu metode hierarki dan metode tak berhierarki. Metode hierarki digunakan untuk mengelompokkan objek-objek secara terstruktur berdasarkan kemiripan dari karakteristik objek tersebut serta banyaknya kluster yang ingin dibentuk belum diketahui. Terdapat dua jenis metode pengklasteran berhierarki, yaitu metode penggabungan (*agglomerative*) dan metode pemisahan (*divisive*). Metode penggabungan dimulai dengan menganggap n objek sebagai n kluster yang terpisah kemudian menggabungkan kluster secara bertahap sampai akhirnya diperoleh satu kluster besar yang memuat semua objek. Dalam mengelompokkan data menggunakan metode *agglomerative*, terdapat beberapa jenis teknik pautan yaitu pautan tunggal (*single linkage*), pautan rata-rata (*average linkage*), dan pautan lengkap (*complete linkage*). Selain itu, terdapat pula metode *ward* dan metode *centroid*. Sebaliknya, metode pemisahan diawali dengan membentuk satu kluster besar yang beranggotakan seluruh objek. Kemudian kluster besar tersebut dipisah menjadi kluster yang lebih kecil sampai terbentuk satu kluster yang beranggotakan satu objek.

Metode *single linkage*, *complete linkage*, dan *average linkage* merupakan metode yang sering digunakan dalam penelitian karena sederhana dan mudah dipahami. Metode *single linkage* dan *complete linkage* cenderung menghasilkan kluster yang tidak seimbang secara ukuran dan hasil kluster yang tidak konsisten atau berubah-ubah, serta kriteria pengklasteran didasarkan pada jarak ekstrim (jarak terpendek dan terpanjang) antar objek. Sedangkan metode *average linkage* cenderung menghasilkan kluster yang seimbang secara ukuran dan bentuk serta hasil kluster yang lebih konsisten, sehingga dianggap lebih stabil dibandingkan dengan metode *complete* dan *single linkage* (Mattjik & Sumertajaya, 2011).

Penelitian terdahulu yang mengaplikasikan metode *average linkage* untuk mengelompokkan kabupaten/kota di Provinsi Jawa Timur menurut faktor-faktor kemiskinan. Dari 38 kabupaten/kota yang ada di Jawa Timur diperoleh hasil metode *average*

linkage sebanyak tiga klaster, yaitu klaster tingkat rendah, klaster tingkat sedang, dan klaster tingkat tinggi (Nafisah & Chandra, 2017). Penelitian lainnya dilakukan Reinaldi et al. (2021) untuk membandingkan metode *complete linkage*, *single linkage*, dan *average linkage* pada klasterisasi kabupaten/kota di Jawa Timur menurut indikator kesejahteraan masyarakatnya. Hasil klasterisasi menggunakan metode *average linkage* memperoleh hasil yang lebih baik dibandingkan metode *complete* dan *single linkage* ditinjau dari *silhouette index* dengan nilai sebesar 0,6054.

Penerapan analisis klaster salah satunya dapat digunakan untuk mengelompokkan provinsi di Indonesia berdasarkan indikator kebahagiaan. Kebahagiaan merupakan suatu ukuran yang mampu mencerminkan tingkat kesejahteraan individual, sehingga mempunyai peranan penting dalam mendukung pembangunan suatu daerah. Tingkat kebahagiaan di suatu daerah dapat tercermin dari nilai indeks kebahagiaannya. Indeks kebahagiaan digunakan untuk memandang persepsi masyarakat terhadap pengalaman yang dirasakan dalam menjalani rutinitas kehidupan bermasyarakat. Berdasarkan laporan kebahagiaan dunia, pada tahun 2021 Indonesia berada di posisi ke-87 dari 146 negara di dunia dengan indeks kebahagiaan sebesar 71,49. Hal ini menunjukkan masyarakat di Indonesia belum cukup bahagia dibandingkan negara-negara lainnya di dunia (Helliwell et al., 2022). Berdasarkan permasalahan tersebut, maka riset ini dilakukan dengan tujuan mengelompokkan provinsi-provinsi di Indonesia berdasarkan indikator indeks kebahagiaan tahun 2021 menggunakan metode *average linkage*. Hasil pengelompokan diharapkan dapat memberikan gambaran yang bermanfaat bagi masyarakat serta pemerintah dalam upaya bersama untuk meningkatkan angka indeks kebahagiaan dan mengatasi disparitas antar provinsi-provinsi di Indonesia.

METODE PENELITIAN

Data yang dianalisis merupakan jenis data kuantitatif yang berasal dari Badan Pusat Statistik (BPS) Indonesia tahun 2021, sehingga termasuk dalam kategori data sekunder. Data yang digunakan adalah data menurut provinsi di Indonesia pada tahun 2021 yang mencakup 34 provinsi dan tujuh indikator indeks kebahagiaan yang terdiri dari variabel pengembangan diri, kepuasan terhadap pendidikan dan keterampilan, pendapatan rumah tangga, keadaan lingkungan, kondisi keamanan, keharmonisan keluarga, dan kesehatan (BPS, 2021).

Teknik analisis data yang digunakan untuk mengelompokkan provinsi di Indonesia menurut indikator indeks kebahagiaan dengan bantuan *software R* yaitu sebagai berikut.

1. Mengumpulkan data yang bersumber dari Badan Pusat Statistik Indonesia tahun 2021.

2. Melakukan eksplorasi data menggunakan *boxplot* untuk mengetahui karakteristik dalam data. Eksplorasi data merupakan proses memahami pola, hubungan, dan karakteristik pada sekumpulan data yang bertujuan untuk menemukan wawasan dan gambaran keseluruhan dari data. Diagram kotak garis (*boxplot*) adalah salah satu bentuk visualisasi data secara ringkas yang dapat menggambarkan distribusi statistik dari suatu peubah. *Boxplot* menyajikan ringkasan statistik yang terdiri dari nilai minimum, kuartil pertama, median, kuartil ketiga, nilai maksimum, deteksi pencilan dan nilai ekstrim.
3. Jika terdapat rentang nilai yang cukup besar atau terdapat perbedaan satuan pada data, lakukan standarisasi data. Standarisasi data bertujuan supaya setiap peubah memiliki kontribusi yang setara dalam proses pengklasteran, sehingga mampu menghasilkan kluster yang homogen dan terpisah dengan baik. Standarisasi data yang lazim digunakan yaitu dengan standarisasi *Z-score* yang mengkonversi setiap variabel terhadap nilai standar (Hair et al., 2018).

$$z = \frac{x_i - \bar{x}}{s} \quad (1)$$

dengan x_i yang berarti data ke- i , \bar{x} merupakan rata-rata dari data, dan s yang menunjukkan simpangan baku dari data (Walpole & Myers, 1995).

4. Melakukan pengklasteran menggunakan metode *average linkage*. Pada metode *average linkage*, untuk mengukur kedekatan antar dua kluster yaitu menggunakan perhitungan rata-rata jarak antar objek dalam satu kluster dengan objek pada kluster lain (Mattjik & Sumertajaya, 2011). Pengklasteran dengan metode *average linkage* diawali dengan membuat matriks jarak $D = \{d_{tu}\}$ untuk melihat kedekatan antar objek, misalnya objek X dan objek Y , kemudian objek tersebut digabungkan menjadi satu kluster sehingga membentuk kluster (XY) . Selanjutnya jarak kluster (XY) dan kluster Z lainnya dirumuskan dengan persamaan berikut.

$$d_{(XY)Z} = \frac{\sum_t \sum_u d_{ku}}{N_{(XY)}N_Z} \quad (2)$$

dengan d_{tu} merupakan jarak objek t pada kluster (XY) dan objek u pada kluster Z , kemudian $N_{(XY)}$ dan N_Z masing-masing merupakan jumlah objek pada kluster (XY) dan kluster Z (Johnson & Wichern, 2007).

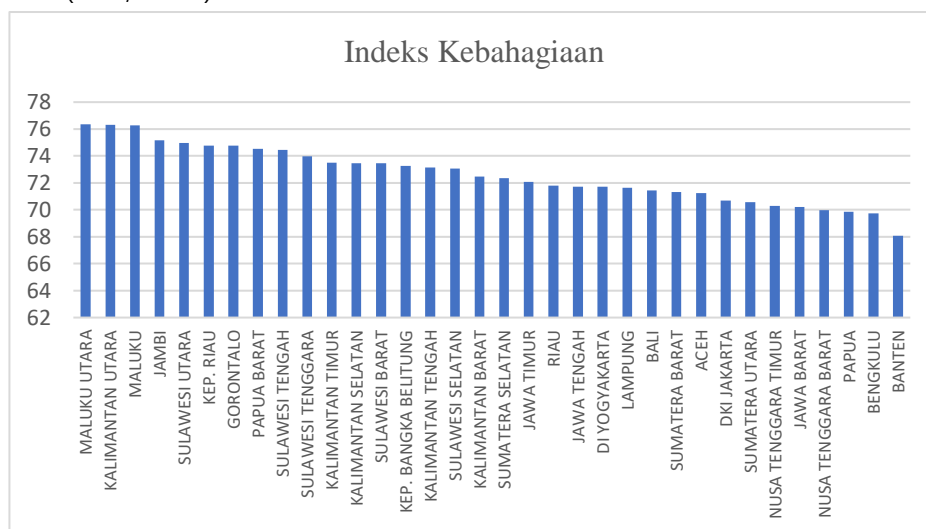
5. Tahapan pengklasteran menggunakan metode *average linkage* yaitu sebagai berikut.
 - i. Menghitung kedekatan antar objek menggunakan salah satu metrik jarak.
 - ii. Hasil dari perhitungan jarak pada tahap i dibuat ke dalam bentuk matriks jarak.
 - iii. Gabungkan objek menjadi satu kluster dilihat dari jarak yang minimum.
 - iv. Perbarui jarak antar objek dengan metode *average linkage* sesuai dengan persamaan 2. Proses terus berlanjut dan ulangi proses tersebut sampai jumlah

kluster $k = 1$ atau semua objek tergabung dalam satu kluster.

- v. Melakukan pemotongan dendrogram dan menentukan jumlah kluster beserta anggotanya. Kriteria pemotongan dendrogram dalam menentukan jumlah kluster yang tepat yaitu dengan melakukan pemotongan pada tingkat kemiripan tertentu sesuai yang diinginkan. Kemudian kriteria lainnya yaitu dengan memotong dendrogram pada titik yang mempunyai selisih nilai kemiripan terbesar antara penggabungan kluster ke- n dan $n-1$ (Manning et al., 2009).
- vii. Menginterpretasikan kluster hasil pengelompokan 34 provinsi di Indonesia menurut indikator indeks kebahagiaan.

HASIL DAN PEMBAHASAN

Pertama akan dilakukan proses mencari dan mengumpulkan data yang bersumber dari Badan Pusat Statistik Indonesia. Objek dalam penelitian ini yaitu 34 provinsi di Indonesia dengan peubah yang diamati yaitu sebanyak tujuh peubah yang menjadi indikator indeks kebahagiaan. Indeks kebahagiaan merupakan suatu ukuran yang dapat menggambarkan kesejahteraan masyarakat ditinjau dari persepsi subjektif setiap individu terhadap segala aspek dalam kehidupan individu tersebut. Indeks kebahagiaan Indonesia merupakan indeks komposit yang dihitung menggunakan dimensi, yaitu terdiri atas dimensi kepuasan hidup, perasaan, serta makna hidup dengan indikator yang diukur menggunakan skala 0 sampai 100 (BPS, 2021).



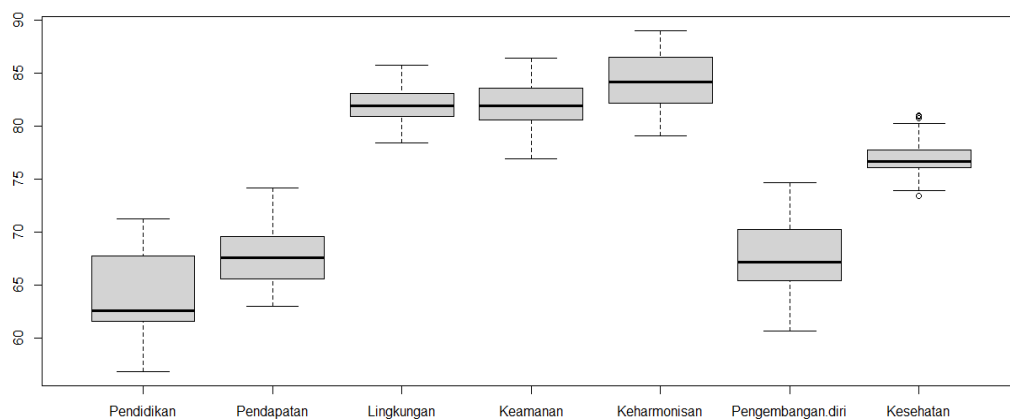
Gambar 1. Skor indeks kebahagiaan setiap provinsi di Indonesia tahun 2021

Berdasarkan Gambar 1 terlihat bahwa Provinsi Maluku Utara memiliki indeks kebahagiaan tertinggi di Indonesia. Hatib Abdul Kadir, seorang antropolog dari Universitas Brawijaya yang telah banyak meneliti masyarakat di Maluku mengatakan sekurangnya ada enam faktor yang membuat orang Maluku Utara dipandang paling bahagia di Indonesia.

Faktor pertama yaitu kondisi emosional dan kesejahteraan sosial, contohnya masyarakat Maluku Utara tidak mudah khawatir. Faktor kedua adalah renggangnya pengekan karena kultur masyarakat Maluku Utara cenderung terbuka serta tidak adanya keterikatan terhadap aturan yang kaku. Kemudian faktor ketiga yaitu Maluku Utara memiliki 1.474 pulau sehingga lebih dekat dengan alam serta jarak antara pantai dan pegunungan cukup berdekatan. Selanjutnya masyarakat lebih dekat dengan keluarga dan sering pergi ke tempat ibadah. Terakhir yaitu faktor-faktor sosial lainnya yang tidak berhubungan langsung dengan faktor ekonomi, seperti rendahnya angka perceraian yang membuat masyarakat Maluku Utara bahagia (Yudhistira, 2022).

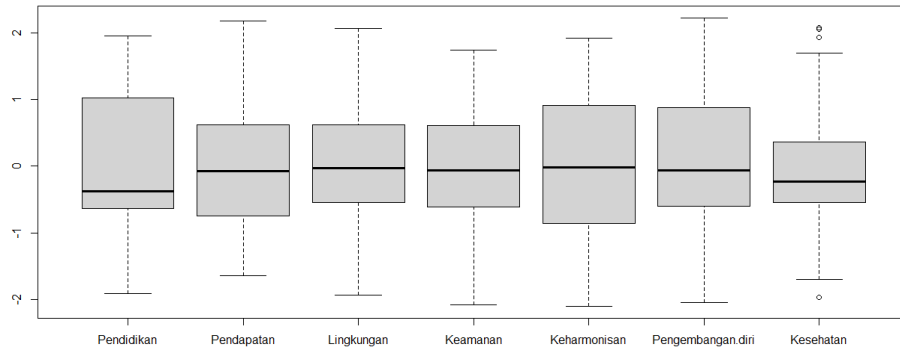
Sebaliknya, Provinsi Banten merupakan provinsi dengan indeks kebahagiaan terendah di Indonesia. Gubernur Banten, Wahidin Halim mengatakan masyarakat Banten tidak bahagia karena terjadi disparitas antar daerah atau terdapat perbedaan antara daerah di perkotaan dan perdesaan (Ridho, 2022).

Sebelum melakukan analisis kluster, akan dilakukan eksplorasi data menggunakan *boxplot* untuk mengetahui karakteristik dalam data. *Boxplot* menyajikan ringkasan statistik yang terdiri dari nilai minimum, kuartil pertama, median, kuartil ketiga, nilai maksimum, serta deteksi pencilan dan nilai ekstrim.



Gambar 2. Boxplot indikator indeks kebahagiaan

Berdasarkan Gambar 2 terlihat bahwa terdapat rentang nilai yang cukup besar pada data indikator indeks kebahagiaan. Dengan demikian, perlu dilakukan standarisasi data terlebih dahulu. Standarisasi data dengan *Z-score* dapat meminimalkan rentang nilai yang terlalu lebar dalam data. Hasil data yang telah distandarisasi menggunakan *Z-score* sesuai dengan persamaan 1 dapat divisualisasikan pada Gambar 3.



Gambar 3. Boxplot setelah standarisasi data

Selain itu, terlihat adanya pencilan pada salah satu variabel, yaitu variabel kesehatan, sehingga tidak cocok menggunakan jarak *Euclidean*. Diperlukan ukuran jarak yang lebih tahan terhadap pengaruh pencilan dalam data, salah satunya dapat menggunakan jarak *Minkowski*. Jarak *Minkowski* merupakan suatu metrik dalam ruang vektor bernorma yang digunakan untuk mengukur jarak antara objek dalam ruang geometri dengan mempertimbangkan perbedaan pada setiap dimensi komponen. Jarak *Minkowski* disebut sebagai generalisasi dari metrik jarak lainnya jika $h = 1$ sebagai jarak *Manhattan* dan jika $h = 2$ sebagai jarak *Euclidean*. Jika $i = (x_{i1}, x_{i2}, \dots, x_{ip})$ dan $j = (x_{j1}, x_{j2}, \dots, x_{jp})$, dua pengamatan dalam R^p , sehingga jarak *Minkowski* dapat dirumuskan dalam persamaan berikut (Han et al., 2012).

$$d(i, j) = \sqrt[h]{|x_{i1} - x_{j1}|^h + |x_{i2} - x_{j2}|^h + \dots + |x_{ip} - x_{jp}|^h} \quad (3)$$

Jarak *Minkowski* dapat membantu mengurangi pengaruh dari pencilan, dengan h lebih besar dari dua, jarak *Minkowski* akan memberikan bobot yang lebih besar pada perbedaan absolut yang besar antara koordinat titik-titik dalam data. Dengan demikian, pada penelitian ini digunakan ukuran jarak *Minkowski* dengan $h = 3$ sebagai panduan dalam melakukan pengelompokan. Sebagai contoh, dilakukan perhitungan jarak antara Provinsi Aceh terhadap Sumatera Utara dan Sumatera Barat menggunakan persamaan 3.

- i. Jarak *Minkowski* untuk Provinsi Aceh terhadap Sumatera Utara

$$\begin{aligned} d(aceh, sumut) &= \sqrt[3]{|x_{11} - x_{21}|^3 + |x_{12} - x_{22}|^3 + \dots + |x_{17} - x_{27}|^3} \\ &= \sqrt[3]{|-0.258 + 0.6382|^3 + |-0.1269 + 1.0319|^3 + \dots + |-0.8689 + 0.5765|^3} \\ &= \sqrt[3]{0.0549 + 0.7412 + \dots + 0.0249} = \sqrt[3]{3.6139} = 1.5346 \end{aligned}$$

- ii. Jarak *Minkowski* untuk Provinsi Aceh terhadap Sumatera Barat

$$\begin{aligned} d(aceh, sumbar) &= \sqrt[3]{|x_{11} - x_{31}|^3 + |x_{12} - x_{32}|^3 + \dots + |x_{17} - x_{37}|^3} \\ &= \sqrt[3]{|-0.258 + 0.3731|^3 + |-0.1269 + 0.5846|^3 + \dots + |-0.8689 + 0.6722|^3} \\ &= \sqrt[3]{0.0015 + 0.0958 + \dots + 0.0076} = \sqrt[3]{0.4796} = 0.7828 \end{aligned}$$

demikian pula untuk objek lainnya dilakukan perhitungan jarak yang sama sehingga diperoleh jarak antar 34 objek.

Berdasarkan perhitungan jarak dengan *Minkowski distance*, diperoleh objek yang memiliki jarak terdekat yaitu objek urutan ke-3 dan ke-12 atau Provinsi Sumatera Barat dan Jawa Barat dengan jarak *Minkowski* sebesar 0.5526276, sehingga Sumatera Barat dan Jawa Barat digabungkan menjadi satu kluster yang dapat dituliskan sebagai kluster (3,12).

Dalam metode *average linkage*, menghitung jarak antar objek disebut sebagai jarak rata-rata dimana jarak tersebut dihitung pada masing-masing kluster. Berdasarkan persamaan (2) dilakukan perhitungan jarak antara kluster (3,12) dengan kluster lain. Diberikan contoh perhitungan antara kluster (3,12) dengan kluster 1, 2, dan 4, yaitu

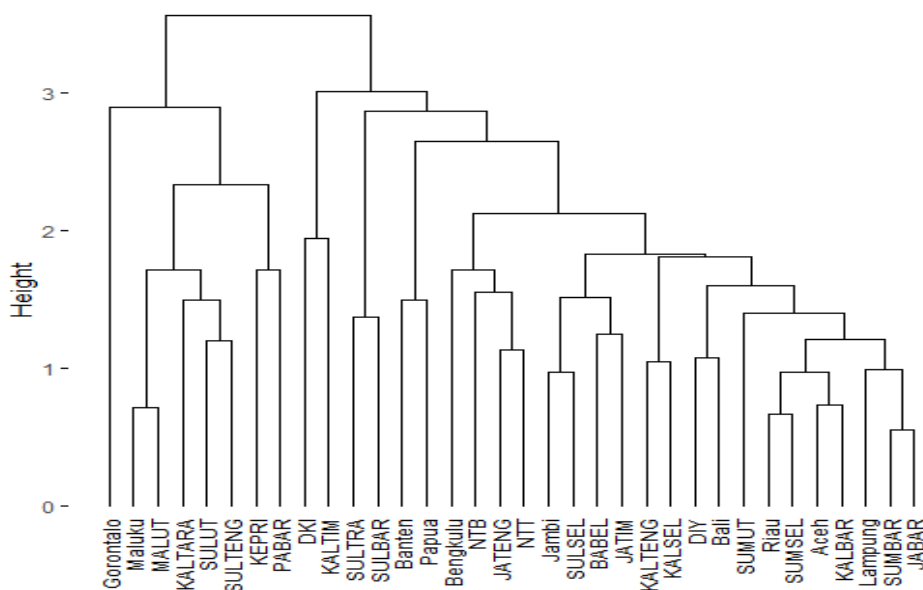
$$d_{(3,12)1} = \frac{d_{(3,1)} + d_{(12,1)}}{2 \times 1} = \frac{0.78 + 1.28}{2} = 1.03$$

$$d_{(3,12)2} = \frac{d_{(3,2)} + d_{(12,2)}}{2 \times 1} = \frac{1.24 + 1.04}{2} = 1.14$$

$$d_{(3,12)4} = \frac{d_{(3,4)} + d_{(12,4)}}{2 \times 1} = \frac{0.89 + 1.37}{2} = 1.13$$

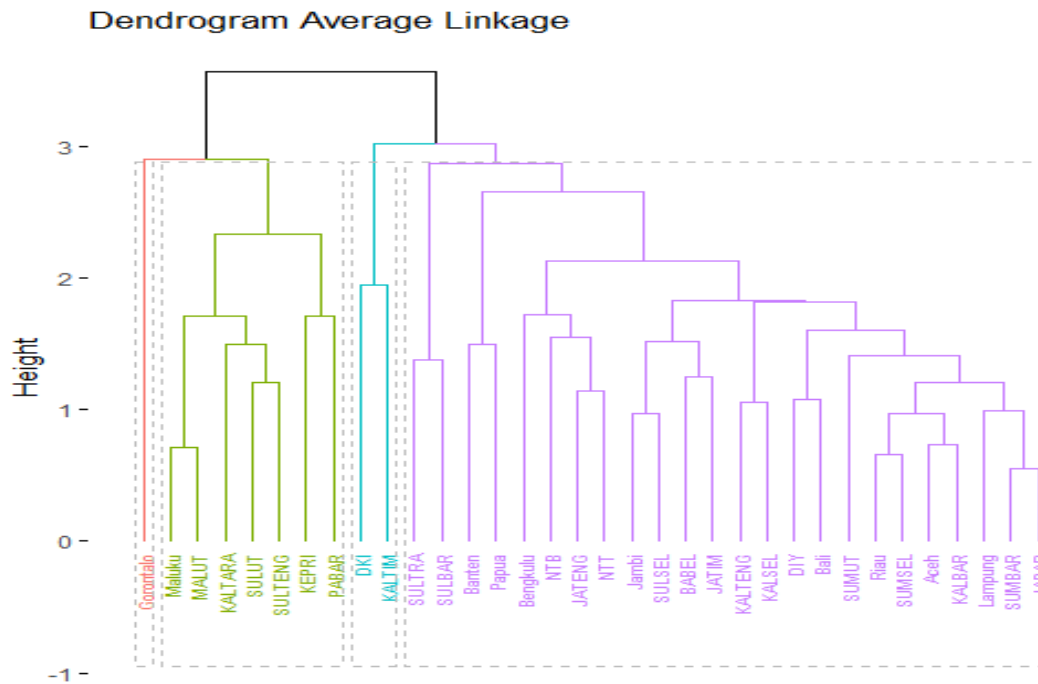
demikian seterusnya, proses perhitungan dilakukan dengan cara yang sama sampai semua kluster bergabung dalam kluster tunggal seperti terlihat pada Gambar 4.

Dendrogram Average Linkage



Gambar 4. Dendrogram pengelompokan average linkage

Setelah semua objek tergabung menjadi satu kluster besar, maka dilakukan pemotongan dendrogram. Pemotongan dendrogram dilakukan pada jarak 3 satuan dan pada penggabungan objek yang memiliki selisih jarak terbesar sehingga diperoleh empat kluster dengan anggota dari masing masing kluster sebanyak 1, 7, 2, 24 seperti yang disajikan dalam gambar dendrogram berikut.



Gambar 5. Pemotongan dendrogram *average linkage*

Berdasarkan Gambar 5 diperoleh empat kluster dengan anggota-anggota pada setiap klasternya disajikan dalam Tabel 2.

Tabel 1. Kluster Hasil Pengelompokan Average Linkage

Kluster	Provinsi Anggota Kluster
1	Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Bangka Belitung, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, Banten, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Sulawesi Selatan, Sulawesi Tenggara, Sulawesi Barat, Papua
2	Kepulauan Riau, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Maluku, Maluku Utara, Papua Barat
3	DKI Jakarta dan Kalimantan Timur
4	Gorontalo

Karakteristik dari setiap kluster dapat diketahui dengan menentukan rata-rata atribut dari setiap kluster. Rata-rata nilai dari masing-masing atribut pada setiap kluster dapat dituliskan dalam Tabel 3.

Klaster	Pendi- dikan	Penda- patan	Ling- kungan	Keamana n	Kehar- monisa n	Pengem- - bangan diri	Kesehata n
1	-0.5518	- 0.4263	-0.1898	-0.1737	-0.3890	-0.3535	-0.4892
2	1.4063	1.2830	1.0818	0.7336	1.2120	1.0915	1.5486
3	1.1892	0.5765	-1.8550	-1.1914	-0.4088	1.0872	0.3033
4	1.0219	0.0984	0.6935	1.4169	1.6706	-1.3298	0.2953

Tabel 2. Pusat Klaster Hasil Pengelompokan Average Linkage

Interpretasi hasil kluster menggunakan metode *average linkage*

- i. Klaster 1 beranggotakan 24 provinsi memiliki rata-rata kluster dengan atribut kepuasan terhadap pendidikan dan keterampilan, pendapatan rumah tangga, dan kesehatan terendah dibandingkan tiga kluster lainnya. Klaster ini memiliki paling banyak anggota, salah satunya Provinsi Bali dan sebagian besar provinsi di Pulau Jawa serta Pulau Sumatera.
- ii. Klaster 2 beranggotakan tujuh provinsi yang dicirikan oleh rata-rata kepuasan terhadap pengembangan diri, pendidikan, pendapatan rumah tangga, keadaan lingkungan, dan kesehatan tertinggi dibandingkan tiga kluster lainnya.
- iii. Klaster 3 beranggotakan dua provinsi yaitu Provinsi DKI Jakarta dan Kalimantan Timur, memiliki rata-rata keadaan lingkungan, kondisi keamanan, dan keharmonisan keluarga terendah dibandingkan kluster lainnya. Sedangkan atribut lainnya memiliki rata-rata nilai yang cukup tinggi.
- iv. Klaster 4 beranggotakan satu provinsi yaitu Provinsi Gorontalo mencirikan kluster dengan rata-rata kepuasan terhadap kondisi keamanan dan keharmonisan keluarga tertinggi, namun pengembangan diri terendah dibandingkan tiga kluster lainnya.

Hasil klasterisasi 34 provinsi di Indonesia ke dalam empat kluster menghasilkan nilai akurasi sebesar 0,5327. Nilai tersebut menunjukkan bahwa sebesar 53,27 persen dari data yang dikelompokkan ke dalam suatu kluster sesuai dengan kriteria atau karakteristik dari kluster tersebut.

SIMPULAN

Analisis kluster merupakan metode pengelompokan data yang tujuannya mereduksi jumlah objek dengan mengelompokkan objek tersebut ke dalam beberapa kelompok atau kluster berdasarkan peubah-peubah yang diamati. Dalam pengelompokan 34 provinsi di Indonesia menurut indikator indeks kebahagiaan menggunakan metode *average linkage* menghasilkan akurasi kluster sebesar 53,27 persen. Kluster 1 terdiri dari 24 provinsi merupakan kluster dengan rata-rata dari seluruh atribut yang rendah, kluster 2 beranggotakan tujuh provinsi yang mencirikan kluster dengan rata-rata dari seluruh atribut yang tinggi, kluster 3 beranggotakan dua provinsi dengan rata-rata kepuasan terhadap lingkungan, keamanan, dan keharmonisan yang rendah, serta kluster 4 beranggotakan Provinsi Gorontalo merupakan kluster dengan kepuasan terhadap keamanan dan keharmonisan yang tinggi, namun pengembangan diri sangat rendah.

Saran yang dapat diberikan bagi pemerintah yang daerahnya tergabung dalam kluster dengan tingkat kebahagiaan yang rendah, diperlukan membuat kebijakan maupun upaya yang dapat meningkatkan indeks kebahagiaan di provinsi-provinsi tersebut. Misalnya dengan meningkatkan kualitas pendidikan, menyediakan berbagai fasilitas kesehatan yang memadai, dan melakukan pembangunan infrastruktur yang merata ke seluruh provinsi di Indonesia. Selain itu, setiap individu masyarakat bersikap toleransi, selalu bersyukur, sering berkumpul dan menghabiskan waktu bersama keluarga.

Penelitian berikutnya disarankan untuk mencoba menggunakan metode kluster berhierarki lainnya, seperti metode *ward*. Selain itu, diharapkan menggunakan ukuran jarak lainnya seperti jarak mahalnobis dikarenakan terdapat pencilan dalam data indikator indeks kebahagiaan.

DAFTAR PUSTAKA

- BPS. (2021). *Indeks Kebahagiaan 2021*. Jakarta: Badan Pusat Statistik RI.
- Hair Jr, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2018). *Multivariate Data Analysis* (8th ed.). United Kingdom: Cengage Learning EMEA.
<https://doi.org/10.1002/9781119409137.ch4>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques. Data Mining: Concepts and Techniques* (3rd). Waltham: Morgan Kaufmann.
<https://doi.org/10.1016/C2009-0-61819-5>
- Helliwell, J. F., Layard, R., Sachs, J. D., De Neve, J.-E., & Aknin, L. B., & Wang, S. (2022). *World Happiness Report 2022*. New York: Sustainable Development Solutions Network. Diambil dari <http://worldhappiness.report/>

- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis* (6 th ed.). Upper Saddle River, New Jersey: Pearson Prentice Hall.
- Kaufman, L., & Rousseeuw, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. New Jersey: John Wiley & Sons, Inc.
- Manning, C. D., Raghavan, P., & Schütze, H. (2009). *An Introduction to Modern Information Retrieval* (second ed.). Cambridge, England: Cambridge University Press. <https://doi.org/10.1108/00242530410565256>
- Mattjik, A. A., & Sumertajaya, I. M. (2011). *Sidik Peubah Ganda Dengan menggunakan SAS*. Bogor: IPB PRESS.
- Nafisah, Q., & Chandra, N. E. (2017). Analisis Cluster Average Linkage Berdasarkan Faktor-Faktor Kemiskinan di Provinsi Jawa Timur. *Zeta - Math Journal*, 3(2), 31–36. <https://doi.org/10.31102/zeta.2017.3.2.31-36>
- Reinaldi, Y., Ulinnuha, N., Hartono, T., & Hafiyusholeh, M. (2021). Comparison of Single Linkage, Complete Linkage, and Average Linkage Methods on Community Welfare Analysis in Cities and Regencies in East Java. *Jurnal Matematika, Statistika dan Komputasi*, 18(1), 130–140. <https://doi.org/10.20956/j.v18i1.14228>
- Ridho, R. (2022). Provinsi Banten Jadi Daerah Paling Tidak Bahagia, Ini Respons Gubernur Wahidin Halim. *Kompas.com*. Diakses pada 19 Juli 2023, dari <https://regional.kompas.com/read/2022/01/08/120349878/provinsi-banten-jadi-daerah-paling-tidak-bahagia-ini-respons-gubernur>
- Walpole, R. E., & Myers, R. H. (1995). *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan* (4 th). Bandung: ITB.
- Yudhistira, A. W. (2022). *Mengapa Orang Maluku Utara Paling Bahagia di Indonesia?* *Katadata.co.id*. Diakses pada 19 Juli 2023, dari <https://katadata.co.id/ariayudhistira/analisisdata/61ee1b64b167e/mengapa-orang-maluku-utara-paling-bahagia-di-indonesia>
- Asdak, Chay, 2004, *Hidrologi dan Pengelolaan Daerah Aliran sungai*, Universitas Gaja Mada, Yogyakarta.
- Ashad, H. (2020a). Analisis Kalibrasi Koefisien Manning pada Saluran Tanah dan Pasangan Batu Saluran Sekunder Belawa Daerah Irigasi Saddang Kabupaten Sidrap. *Jurnal Teknik Sipil MACCA*, 5(3), 267–274. <https://jurnal.ft.umi.ac.id/index.php/jtasm/article/view/204/137>
- Ashad, H. (2020b). Kajian Pengaliran Banjir Sungai Terhadap Koefisien Kekasaran Manning Pada Sungai Larompong Kabupaten Luwu Provinsi Sulawesi Selatan. *Jurnal Teknik Sipil MACCA*, 5(2), 112–116.

<https://jurnal.ft.umi.ac.id/index.php/jtasm/article/view/85/62>

- Ashad, H., & Hadi, A. K. (2021). *Kajian Kebutuhan Debit Air pada Daerah Irigasi Bila Kabupaten Sidenreng Rappang*. *Jurnal Teknik Sipil MACCA*, 6(2), 141–150. <https://jurnal.ft.umi.ac.id/index.php/jtasm/article/view/339/216>
- Gemma Galgani Tanjung Dewandaru, Lasminto Umboro. 2014, *Studi Penanggulangan Banjir Kali Lamong Terhadap Genangan di Kabupaten Gresik*. Institut Teknologi Sepuluh Nopember, Surabaya.
- Haryono Putro dan Joetata Hadihardaja, 2013. *Variasi Koefisien kekasaran Manning (n) pada Flum Akrilic pada Variasi Kemiringan Saluran dan Debit Aliran*. Jurnal. Jurusan Teknik Sipil Universitas Gunadarma.
- Hadisusanto, Nugroho. 2011. *Aplikasi Hidrologi*, Malang, Jogja Media Umum.
- Indarto, S., & Dea, I. (2022). *Hidrologi : Teori Dan Penerapan Hidrologi Di Jawa Timur* . Terbitkan dalam-dalam.
- Kodoatie, R. J. & Syarief, R. 2005. *Pengelolaan Sumber Daya Air Terepadu*. Andi Offset, Yogyakarta.
- Rizaldy, A., Musa, R., & Mallombasi, A. (2021). Kalibrasi Koefisien Debit Model Buka Pintu Sorong Pada Saluran Terbuka (Uji Laboratorium). *Jurnal Teknik Sipil MACCA*, 6(1), 1–10. <https://jurnal.ft.umi.ac.id/index.php/jtasm/article/view/273/166>
- Kusuma, W. I. (2016). *Perencanaan Sistem Drainase Kawasan Perumahan Green Mansion Residence Sidoarjo*. Jurnal. Institut Teknologi Sepuluh Nopember.
- Lestari, U. S. (2016). *Kajian metode empiris untuk menghitung debit banjir Sungai Negara di ruas Kecamatan Sungai Pandan (Alabio)*. *Poros Teknik*, 8(2), 86–96.
- Maryono, A. 2007, *Restorasi Sungai*, Gadjadara University Press, Yogyakarta.
- Mulyanto, H. R. 2006, *Sungai Fungsi dan Sifat-Sifatnya*, Graha Ilmu, Yogyakarta.
- Pallu, Muh saleh, *Teori dasar Angkutan sedimen Didalam Saluran Terbuka*, CV. Telaga zamzam, Makassar, 2012
- Priyantoro, D. 1987. *Teknik pengangkutan Sedimen*. Malang : Himpunan Mahasiswa Pengairan Fakultas Teknik Universitas Brawijaya
- Sarwono Bambang, Lasminto Umboro, Ramanintyas Aninda. 2017. *Perencanaan Penanggulangan Banjir Akibat Luapan Sungai Petung*. Institut Teknologi Sepuluh Nopember Surabaya.
- Soemarto, CD. 1999. *Hidrologi Teknik*, Surabaya, Penerbit Usaha Nasional.
- Soemartono, CD. *Hidrologi Teknik*, Usaha Nasional, Surabaya 1987.
- Sowono Bambang, Ratnasari DA, Ansori Bagus Muhammad. 2015. *Studi Pengendalian Banjir Sungai Kalidawir*, Institut Teknologi Sepuluh Nopember, Surabaya.

- Soewarno, 1995. *Hidrologi, Aplikasi Metode Statistik untuk Analisa Data Jilid I*. Penerbit Nova, Bandung.
- Soewarno, (1995). *Hidrologi, Aplikasi Metode Statistik untuk Analisa Data Jilid II*. Penerbit Nova, Bandung.
- Sosrodarsono, S. dan Takeda, K. 1987. *Hidrologi Untuk Pengairan*. PT Pradanaya Paramita, PT. Jakarta.
- Tahir, A. (2011). *Kinerja campuran Split Mastic Asphalt (SMA) yang menggunakan serat selulosa alami dedak padi*. Jurnal Rekayasa Dan Manajemen Transportasi. <https://core.ac.uk/download/pdf/292007557.pdf>
- Tahir, M., & Musa, R. (2020). *Kajian Koefisien Kekasaran Manning (n) Pasangan Batu dan Beton Berdasarkan Kuantifikasi Kekasaran Hidrolis (Studi Kasus Daerah Irigasi Wawotobi Kab. Konawe Sultra)*. Jurnal Teknik Sipil MACCA, 5(2), 118–132. <https://jurnal.ft.umi.ac.id/index.php/jtسم/article/view/83/61>
- Suadnya, DP, JSF Sumarauw, and T. Mananoma. 2017. "Analisis Debit Banjir Dan Tinggi Muka Air." Jurnal Sipil Statik 5(3):143–50.
- Suripin, 2003. Sistem Drainase Perkotaan Yang Berkelanjutan, Andi, Yogyakarta.
- Triatmodjo, Bambang. 2016. *Soal Penyelesaian Hidrolika II*. Cetakan ke-16, Beta Offset, Yogyakarta.
- Triatmodjo, Bambang. 2019. *Hidrologi Terapan*. Cetakan ke-7, Beta Offset, Yogyakarta.
- Triatmodjo, Bambang. 2020. *Hidrolika II*. Cetakan ke-13, Beta Offset, Yogyakarta.
- Ven Te Chow. 1985, *Hidrolika Saluran Terbuka*. Diterjemahkan oleh E.V. Nensi Rosalina, Penerbit Erlangga, Jakarta
- Wahyuddin Qadri, Sholihin Muhammad, Sisingsih Dian. 2015. *Studi Penanganan Banjir Sungai Bila Kab. Sidrap*, Universitas Brawijaya, Malang.
- WAHYU INDRA KUSUMA (2016) 'Perencanaan Sistem Drainase Kawasan Perumahan Green Mansion Residence Sidoarjo', 2016, pp. 1–163.
- Yang,C.T. 1996, Sediment Transport Theory and Practice, The McGraw-Hill Companies,Inc, Singapura.
- Wellang, M., Hasim, M. F., & Simin, I. F. (2019). Analisa Koefisien Kekasaran Manning (n) dan Chezy (c) pada Saluran Terbuka dengan Variasi Debit Aliran dan Kemiringan. Jurnal Teknik Sipil MACCA,4(1),11–21. <https://jurnal.ft.umi.ac.id/index.php/jtسم/article/view/357/231>