

Optimalisasi Penentuan Klaster pada Indeks Khusus Penanganan Stunting Menggunakan Metode Agglomerative Hierarchical

Syahrul Aziz Pamungkas, Ria Dhea Layla Nur Karisma, Evawati Alisah

Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia

azizsyahrul96@gmail.com*, riadhea@uin-malang.ac.id, evawatialisah@mat.uin-malang.ac.id

Abstrak

Stunting merupakan masalah gizi yang masih menjadi fokus utama di negara berkembang, salah satunya Indonesia. Salah satu instrumen yang dirancang untuk mengukur performa penyelenggaraan program percepatan penurunan stunting di tingkat nasional yaitu, Indeks Khusus Penanganan Stunting (IKPS). Penelitian ini bertujuan untuk mengelompokkan provinsi di Indonesia berdasarkan indeks khusus penanganan *stunting* yang terdiri dari 6 indikator, yaitu indikator kesehatan, gizi, perumahan, pangan, pendidikan, dan perlindungan sosial menggunakan metode *agglomerative hierarchical clustering*. Metode *agglomerative hierarchical clustering* terbagi mejadi beberapa metode, diantaranya metode *single linkage*, *complete linkage*, *average linkage*, dan *ward*. Penelitian ini membandingkan keempat metode tersebut dengan tujuan untuk mendapatkan solusi klaster terbaik dalam pengelompokan provinsi di Indonesia berdasarkan Indeks Khusus Penanganan *Stunting* (IKPS). Penentuan metode terbaik pada *agglomerative hierarchical clustering* ditentukan dengan nilai koefisien korelasi *cophenetic*. Hasil penelitian menunjukkan bahwa metode *average linkage* memberikan solusi klaster yang lebih baik dibandingkan metode lainnya. Solusi klaster pada metode *average linkage* menghasilkan 8 klaster diantaranya, klaster 1 terdiri dari 1 provinsi, klaster 2 terdiri dari 9 provinsi, klaster 3 terdiri dari 12 provinsi, klaster 4 terdiri dari 6 provinsi, klaster 5 terdiri dari 1 provinsi, klaster 6 terdiri dari 1 provinsi, klaster 7 terdiri dari 3 provinsi, dan klaster 8 terdiri dari 1 provinsi di Indonesia.

Kata kunci: Stunting, Indeks Khusus Penanganan Stunting (IKPS); Agglomerative Hierarchical Clustering; Average Linkage; Koefisien Korelasi Cophenetic.

Abstract

Stunting is a nutrition problem that is still the main focus in developing countries, one of which is Indonesia. One of the instruments designed to measure the performance of the implementation of the stunting reduction acceleration program at the national level is the Stunting-Specific Intervention Index (*Indeks Khusus Penanganan Stunting/IKPS*). This study aims to group provinces in Indonesia based on a special index for handling stunting consisting of six indicators, namely health, nutrition, housing, food, education, and social protection indicators using the agglomerative hierarchical clustering method. The agglomerative hierarchical clustering method is divided into several methods, including single linkage, complete linkage, average linkage, and ward methods. This study compares the four methods with the aim of obtaining the best cluster solution in the grouping of provinces in Indonesia based on the stunting-specific intervention index. The determination of the best method in agglomerative hierarchical clustering is determined by the value of the cophenetic correlation coefficient. The results show that the average linkage method provides a better cluster solution than other methods. The cluster solution in the average linkage method produces eight clusters, including, cluster 1 consists of one province, cluster 2 consists of nine provinces, cluster 3 consists of twelve provinces, cluster 4 consists of six provinces, cluster 5 consists of one province, cluster 6 consists of one province, cluster 7 consists of three provinces, and cluster 8 consists of one province in Indonesia.

Keywords: Stunting, Stunting-Specific Intervention Index; Agglomerative Hierarchical Clustering; Average Linkage; Cophenetic

PENDAHULUAN

Stunting adalah kondisi pertumbuhan terhambat pada anak balita akibat kekurangan gizi kronis dan infeksi berulang, terutama dalam 1.000 Hari Pertama Kehidupan (HPK), yang mencakup periode sejak masa janin hingga usia 23 bulan. Masalah kekurangan gizi kronis yang disebabkan oleh asupan gizi yang tidak mencukupi dalam jangka waktu yang lama akibat pemberian makanan yang tidak sesuai dengan kebutuhan gizi [1]. *Stunting* tidak hanya berdampak pada pertumbuhan fisik, tetapi juga memengaruhi fungsi-fungsi penting lain dalam tubuh, seperti perkembangan otak dan sistem kekebalan tubuh. Balita yang mengalami *stunting* berpotensi memiliki kemampuan kognitif yang lebih rendah, lebih mudah terserang penyakit, dan di masa depan dapat memberikan dampak negatif terhadap produktivitas pada suatu negara. Program penanganan malnutrisi sudah dilakukan selama beberapa tahun terakhir, namun tampaknya belum fokus secara khusus pada malnutrisi kronis apa yang mengakibatkan *stunting*. Akibatnya, angka kejadian *stunting* tidak menunjukkan penurunan yang signifikan, meskipun angka kejadian *stunting* lain seperti *wasting* (kurus) sudah menurun [2]. Oleh karena itu, pemerintah menjadikan percepatan penurunan *stunting* sebagai salah satu program prioritas nasional [3].

Berdasarkan Survei Status Gizi (SSGI) Kementerian Kesehatan, prevalensi balita *stunting* di Indonesia mencapai 21,6% pada tahun 2022, angka tersebut turun 2,8% dari tahun 2021 sebesar 24,4%. Pada tahun 2022, Nusa Tenggara Timur (NTT) mempunyai prevalensi balita *stunting* sebesar 35,3%, angka tersebut turun 2,5% dari tahun 2021 sebesar 37,8%. Namun penurunan tersebut tidak merubah posisi NTT menjadi Provinsi tertinggi dengan prevalensi angka balita *stunting*. Provinsi dengan prevalensi balita *stunting* tertinggi kedua adalah Sulawesi Barat, dengan angka prevalensi mencapai 35%. Di Provinsi Sulawesi Barat, angka prevalensi *stunting* mengalami peningkatan yang signifikan dari tahun 2021 ke 2022, yaitu sebesar 1,2%. Selain Sulawesi Barat, Papua juga mencatat peningkatan signifikan dalam angka prevalensi *stunting*, yaitu sekitar 5,1%. Pada tahun 2021, prevalensi *stunting* di Papua mencapai 29,5%, sedangkan pada tahun 2022 meningkat menjadi 34,6%. Beberapa provinsi lain yang juga mencatat peningkatan prevalensi *stunting* meliputi Nusa Tenggara Barat (NTB), Papua Barat, Sumatera Barat, dan Kalimantan Timur [2]. Angka balita *stunting* di Indonesia mengalami penurunan menurut Survei Status Gizi Indonesia (SSGI) kementerian kesehatan. Walaupun prevalensi *stunting* telah menurun, angkanya masih belum mencapai batas yang ditetapkan oleh *World Health Organization* (WHO), yaitu 20%. [4].

Indeks Khusus Penanganan *Stunting* (IKPS) adalah salah satu instrumen yang dirancang untuk mengukur performa penyelenggaraan program percepatan penurunan *stunting* di tingkat nasional maupun tingkat kabupaten/kota. Selain IKPS, terdapat instrumen lain yang digunakan untuk menilai keberhasilan program ini, yaitu Survei Status Gizi tahunan yang dilaksanakan oleh Kementerian Kesehatan untuk mengukur tingkat *stunting* pada balita di Indonesia, serta evaluasi terhadap keberhasilan 8 Aksi Konvergensi yang dilakukan oleh Kementerian Dalam Negeri melalui Pemerintah Provinsi. Perhitungan IKPS pada umumnya mencakup 6 dimensi, yakni kesehatan, gizi, perumahan, pangan, pendidikan, dan dimensi perumahan [5].

Beberapa peneliti sebelumnya [6] yang membahas mengenai data *stunting* seperti penelitian yang berjudul Analisis Clustering Provinsi Berdasarkan Prevalensi Stunting Balita Menggunakan Algoritma *Single* Dan *Complete Linkage*. Yang menunjukkan hasil bahwa Gizi adalah salah satu parameter dalam menilai keberhasilan pembangunan kesehatan di suatu negara untuk menciptakan sumber daya manusia yang bermutu. Penelitian ini membandingkan keempat metode tersebut dengan tujuan untuk mendapatkan solusi kluster terbaik dalam pengelompokan

provinsi di Indonesia berdasarkan Indeks Khusus Penanganan *Stunting* (IKPS).

METODE

Data yang dianalisis dalam penelitian ini merupakan data indeks khusus penanganan *stunting* 34 Provinsi di Indonesia tahun 2022. Sumber data yang digunakan dalam penelitian ini merupakan data sekunder yang diambil dari *website* Badan Pusat Statistik Provinsi Jawa Timur <https://jatim.bps.go.id/> yang telah dipublikasi pada tahun 2023.

Analisis data menggunakan *agglomerative hierarchical* dengan membandingkan keempat metode, yaitu *single linkage*, *complete linkage*, *average linkage*, dan *ward method* [7] dengan tujuan untuk mendapatkan solusi klaster terbaik dalam pengelompokan provinsi di Indonesia berdasarkan Indeks Khusus Penanganan *Stunting* (IKPS). Dalam menganalisis data ini dibantu dengan bantuan *software RStudio* dan *excel*. Langkah-langkah analisis pertama yang dilakukan adalah pengecekan *outlier* pada data. Langkah selanjutnya melakukan pengujian asumsi multikolinieritas yang bertujuan untuk melihat apakah terdapat korelasi yang tinggi antar variabel independen. Setelah dipastikan tidak ada korelasi yang tinggi antar variabel selanjutnya menghitung jarak *Euclidean* dan *Manhattan* antar semua objek. Setelah mendapatkan nilai jarak antar variabel dilanjut dengan menganalisis menggunakan metode *agglomerative hierarchical*.

Menganalisis dengan *agglomerative hierarchical* dengan membandingkan metode *single linkage*, *complete linkage*, *average linkage*, dan *ward method*. Kemudian setelah mendapatkan masing-masing klaster kemudian mencari nilai klaster optimum setiap metode dengan melihat nilai *pseudo f-statistic* yang terbesar. Nilai *Pseudo F* tertinggi menunjukkan bahwa klaster tersebut memberikan hasil yang optimal, dimana keragaman kelompok sangat homogen sedangkan antar kelompok heterogen [8].

Setelah mendapatkan hasil klaster optimim setiap metode, langkah terakhir yaitu menguji kebaikan setiap metode. Uji kebaikan metode ini penting untuk menilai mutu dan kualitas hasil analisis klaster dengan mengevaluasi sejauh mana klaster dapat membedakan data berdasarkan variabel atau karakteristiknya [9]. Salah satu ukuran yang dapat digunakan untuk mengukur kebaikan klaster menggunakan analisis klaster *hierarki* adalah koefisien korelasi *cophenetic*. Nilai koefisien korelasi *cophenetic* berkisar antara -1 dan 1, nilai koefisien korelasi *chopenetic* mendekati 1 berarti solusi yang dihasilkan proses klaster baik dan sebaliknya [10].

HASIL DAN PEMBAHASAN

Analisis Statistik Deskriptif

Tabel 1 Statistika Deskriptif

Variabel	N	Mean	Maximum	Minimum	Std Dev
Imunisasi	34	69,09	93,20	25	15,03
Penolong persalinan oleh tenaga kesehatan	34	87,34	98,60	48,20	10,87
Keluarga Berencana (KB)	34	62,76	84,10	19,70	14,26
ASI eksklusif	34	86,50	99,60	67	9,07
Makanan Pendamping (MP) ASI	34	83,90	98,70	66,90	8,53
Air minum layak	34	87,64	98,40	65,40	7,85
Sanitasi layak	34	80,99	96,20	40,30	9,79
Ketidakcukupan Konsumsi Pangan	34	77,20	96,30	39,70	13,87
Pendidikan Anak Usia Dini (PAUD)	34	37	70	14,10	11
Kepemilikan JKN/Jamkesda	34	70,28	97	54	10,45

Berdasarkan Tabel 1 menjelaskan statistika deskriptif dari variabel-variabel dalam data indeks khusus penanganan *stunting* di Indonesia tahun 2022. Pada kolom N, kesebelas variabel memiliki 34 data, artinya tidak ada data yang hilang atau *missing value*. Berdasarkan pada dimensi kesehatan pada variabel imunisasi diketahui bahwa nilai terendah adalah 25%, sedangkan nilai tertinggi adalah 93,20%. Adapun rata-rata persentase imunisasi sebesar 69,09% dengan standar deviasi 15,03%. Persentase imunisasi yang semakin tinggi mencerminkan semakin lengkapnya pemberian imusisasi dasar pada anak. Ini akan membantu mencapai tujuan imunisasi, yaitu melindungi anak dari berbagai penyakit.

Berdasarkan pada dimensi kesehatan pada variabel penolong persalinan oleh tenaga kesehatan diketahui bahwa nilai terendah adalah 48,20%, sedangkan nilai tertinggi adalah 98,60%. Adapun rata-rata persentase penolong persalinan oleh tenaga kesehatan sebesar 87,34% dengan standar deviasi 10,87%. Persentase penolong persalinan oleh tenaga kesehatan yang semakin tinggi mencerminkan semakin baiknya akses dan kualitas layanan kesehatan yang diterima oleh ibu selama proses persalinan yang dapat meningkatkan keselamatan ibu dan mengurangi balita yang *stunting*.

Berdasarkan pada dimensi kesehatan pada variabel keluarga berencana (KB) diketahui bahwa nilai terendah adalah 19,70%, sedangkan nilai tertinggi adalah 84,10%. Adapun rata-rata persentase keluarga berencana (KB) sebesar 62,76% dengan standar deviasi 14,26%. Persentase keluarga berencana (KB) yang semakin tinggi mencerminkan semakin baiknya kesadaran masyarakat terhadap layanan KB, serta peningkatan kesejahteraan keluarga melalui perencanaan

jumlah dan jarak kelahiran.

Berdasarkan pada dimensi gizi pada variabel ASI eksklusif diketahui bahwa nilai terendah adalah 67%, sedangkan nilai tertinggi adalah 99,60%. Adapun rata-rata persentase ASI eksklusif sebesar 86,50% dengan standar deviasi 9,07%. Persentase ASI eksklusif yang semakin tinggi mencerminkan semakin meningkatnya kesadaran ibu tentang pentingnya memberikan ASI eksklusif yang mengandung nutrisi serta zat-zat penting yang dibutuhkan untuk perkembangan fisik dan mental anak, serta berperan dalam mencegah berbagai penyakit.

Berdasarkan pada dimensi gizi pada variabel makanan pendamping (MP) ASI diketahui bahwa nilai terendah adalah 66,90%, sedangkan nilai tertinggi adalah 98,70%. Adapun rata-rata persentase makanan pendamping (MP) ASI sebesar 83,90% dengan standar deviasi 8,53%. Persentase makanan pendamping (MP) ASI yang semakin tinggi mencerminkan semakin baiknya pemahaman orang tua tentang pentingnya memberikan makanan pendamping yang tepat dan bergizi untuk mendukung pertumbuhan dan perkembangan anak setelah masa menyusui.

Berdasarkan pada dimensi perumahan pada variabel air minum layak diketahui bahwa nilai terendah adalah 65,40%, sedangkan nilai tertinggi adalah 98,40%. Adapun rata-rata persentase air minum layak sebesar 87,64% dengan standar deviasi 7,85%. Persentase air minum layak yang semakin tinggi mencerminkan peningkatan kualitas hidup masyarakat, akses yang lebih baik terhadap sumber air bersih, dan upaya yang efektif dalam menjaga kesehatan serta mencegah penyakit yang disebabkan oleh konsumsi air yang tidak aman.

Berdasarkan pada dimensi perumahan pada variabel sanitasi layak diketahui bahwa nilai terendah adalah 40,30%, sedangkan nilai tertinggi adalah 96,20%. Adapun rata-rata persentase sanitasi layak sebesar 80,99% dengan standar deviasi 9,79%. Persentase sanitasi layak yang semakin tinggi mencerminkan upaya yang lebih baik dalam mengurangi resiko infeksi dan penyakit yang dapat mengganggu pertumbuhan dan perkembangan anak.

Berdasarkan pada dimensi pangan pada variabel ketidacukupan konsumsi pangan diketahui bahwa nilai terendah adalah 39,70%, sedangkan nilai tertinggi adalah 96,30%. Adapun rata-rata persentase ketidacukupan konsumsi pangan sebesar 77,20% dengan standar deviasi 13,87%. Persentase ketidacukupan konsumsi pangan yang semakin kecil mencerminkan upaya yang efektif dalam meningkatkan ketersediaan, aksesibilitas, dan pemanfaatan pangan yang memadai bagi masyarakat. Hal ini menunjukkan keberhasilan program intervensi seperti distribusi bantuan pangan, pemberdayaan ekonomi masyarakat, dan edukasi gizi, yang secara langsung berkontribusi dalam mengurangi kerentanan pangan dan meningkatkan kualitas asupan gizi, terutama di daerah yang sebelumnya rawan pangan.

Berdasarkan pada dimensi pendidikan pada variabel pendidikan anak usia dini (PAUD) diketahui bahwa nilai terendah adalah 14,10%, sedangkan nilai tertinggi adalah 70%. Adapun rata-rata persentase pendidikan anak usia dini (PAUD) sebesar 37% dengan standar deviasi 11%. Persentase pendidikan anak usia dini (PAUD) yang semakin tinggi mencerminkan meningkatnya akses dan partisipasi anak-anak dalam program Pendidikan yang berkualitas, serta kesadaran masyarakat tentang pentingnya pendidikan awal untuk perkembangan kognitif, sosial, dan emosional anak.

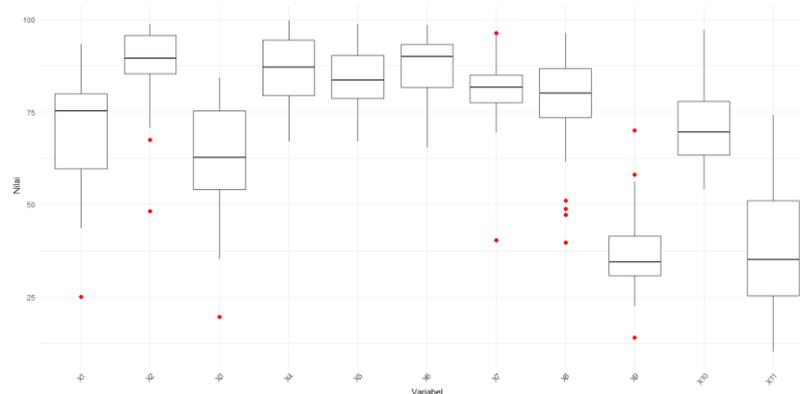
Berdasarkan pada dimensi perlindungan sosial pada variabel kepemilikan JKN/Jamkesda diketahui bahwa nilai terendah adalah 54%, sedangkan nilai tertinggi adalah 97%. Adapun rata-rata persentase kepemilikan JKN/Jamkesda sebesar 70,28% dengan standar deviasi 10,45%. Persentase kepemilikan JKN/Jamkesda yang semakin tinggi mencerminkan meningkatnya akses masyarakat terhadap layanan kesehatan.

Berdasarkan pada dimensi perlindungan sosial pada variabel penerima KPS/KKS diketahui bahwa nilai terendah adalah 10,30%, sedangkan nilai tertinggi adalah 74,20%. Adapun

rata-rata persentase penerima KPS/KKS sebesar 37,66% dengan standar deviasi 16,24%. Persentase penerima KPS/KKS yang semakin tinggi mencerminkan peningkatan akses masyarakat terhadap bantuan sosial yang bertujuan untuk mendukung kesejahteraan keluarga, mengurangi kemiskinan, dan memberikan perlindungan kepada kelompok rentan dalam memenuhi kebutuhan pokok mereka.

Deteksi *Outlier*

Outlier adalah data atau pengamatan yang memiliki nilai sangat berbeda dari nilai-nilai lainnya dalam suatu dataset. *Outlier* dapat muncul karena berbagai alasan, termasuk kesalahan pengukuran, variasi alami dalam data, atau adanya fenomena yang tidak biasa. Pada penelitian ini cara mengidentifikasi *outlier* menggunakan metode statistik yaitu, *Interquartile Range (IQR)* dengan visualisasi menggunakan *boxplot*. Dalam banyak kasus, jika *outlier* mewakili lebih dari 5% hingga 10% dari total data maka peneliti perlu mempertimbangkan untuk mengatasi *outlier* tersebut dengan menghapusnya atau menggantinya dengan nilai rata-rata atau median [11]. Diagram *boxplot* membantu memvisualisasikan IQR dan mengidentifikasi *outlier* berdasarkan batas yang ditentukan IQR. Diagram *boxplot* dapat dilihat sebagai berikut:



Gambar 1 *Outlier* Data IKPS

Dapat dilihat pada Gambar 1 terdapat *outlier* pada beberapa variabel. Dari 374 data terdapat 13 *outlier* atau hanya 3,2% dari keseluruhan jumlah data yang ada. Pada penjelasan pada buku [11] jika *outlier* mewakili lebih dari 5% hingga 10% dari total data maka peneliti perlu mempertimbangkan untuk mengatasi *outlier* tersebut dengan menghapusnya atau menggantinya dengan nilai *mean* atau *median*. Namun, karena persentase *outlier* hanya terdapat 3,2% sesuai dengan penjelasan pada Subbab 2.2, *outlier* tidak perlu diatasi karena tidak memiliki dampak yang signifikan terhadap interpretasi data.

Uji Multikolinieritas

Penentuan Uji multikolinieritas kadang perlu dilakukan karena untuk mengecek apakah ada korelasi antar variabel. Salah satu metode untuk mendeteksi adanya multikolinieritas pada variabel adalah dengan menghitung nilai *Variance Inflation Factors (VIF)*. Data dikatakan terbebas dari multikolinieritas jika nilai VIF-nya kurang dari sama dengan 10. Jika salah satu variabel terindikasi terdapat multikolinieritas maka penghapusan variabel yang terindikasi multikolinieritas dapat dilakukan. Berikut hipotesis dan cara mencari nilai VIF:

1. Hipotesis:

H_0 : Tidak terdapat multikolinieritas

H_1 : Terdapat multikolinieritas

2. Statistik uji VIF:

$$R^2 = \frac{SS_R}{SS_T}$$

$$VIF = \frac{1}{1 - R^2}$$

Tabel 2 Nilai VIF

Variabel	VIF
Imunisasi	2,12
Penolong persalinan oleh tenaga kesehatan	3,18
Keluarga Berencana (KB)	3,44
ASI eksklusif	1,82
Makanan Pendamping (MP) ASI	2,04
Air minum layak	2,57
Sanitasi layak	2,72
Ketidakcukupan Konsumsi Pangan	3,37
Pendidikan Anak Usia Dini (PAUD)	2,25
Kepemilikan JKN/Jamkesda	2,29
Penerima KPS/KKS	1,98

3. Daerah keputusan:

Dapat dilihat pada Tabel 2 bahwa nilai VIF semua variabel kurang dari 10 yang berarti tidak tolak H_0 yang artinya, kesepuluh variabel terbebas dari multikolinearitas dan semua variabel dapat di gunakan ke langkah berikutnya.

Perhitungan Jarak

Analisis klaster memerlukan beberapa ukuran untuk mengidentifikasi tingkat kemiripan antara objek-objek yang dianalisis. Salah satu ukuran yang dapat digunakan untuk mengukur kemiripan antar data dalam analisis klaster adalah ukuran jarak (*distance*). Jika jarak antara objek semakin kecil, maka tingkat kemiripan karakteristik antara kedua objek tersebut akan semakin besar. Sebaliknya, semakin jauh jarak antara objek, semakin rendah kemungkinan kemiripan di antara objek tersebut [12].

Beberapa ukuran ketidakmiripan (*dissimilarity measure*) atau ukuran jarak (*distance measurement*) yang digunakan untuk mengukur jarak antar dua objek. Terdapat berbagai metode untuk mengukur jarak antara dua objek dengan ukuran kedekatan yang berbeda, seperti jarak *Euclid*, *Minkowski*, *Manhattan*, dan *Mahalonobis* [13]. Pada penelitian ini pengukuran jarak dilakukan menggunakan metode *Euclidean* dan *Manhattan*. Hasilnya menggambarkan seberapa

jauh dua titik atau objek berada satu sama lain dalam ruang tersebut. Untuk menghitung jarak *Euclidean* dan jarak *Manhattan* sampai terbentuk sebuah matriks jarak berbentuk diagonal dengan dimensi 34×34 . Berikut adalah ringkasan Tabel matriks jarak *Euclidean* dan jarak *manhattan*:

Tabel 3 Matriks Jarak *Euclidean*

<i>I</i>	<i>I</i> ₁	<i>I</i> ₂	<i>I</i> ₃	<i>I</i> ₄	<i>I</i> ₅	...	<i>I</i> ₃₄
<i>I</i> ₁	0	54,16	37,92	49,94	68,65	...	91,96
<i>I</i> ₂	54,16	0	34,90	25,26	42,89	...	88,11
<i>I</i> ₃	37,92	34,90	0	28,73	42,14	...	83,78
<i>I</i> ₄	49,94	25,26	28,73	0	30,42	...	80,60
<i>I</i> ₅	68,65	42,89	42,14	30,42	0	...	92,98
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
<i>I</i> ₃₄	91,96	88,11	83,78	80,60	92,98	...	0

Tabel 4 Matriks Jarak *Manhattan*

<i>I</i>	<i>I</i> ₁	<i>I</i> ₂	<i>I</i> ₃	<i>I</i> ₄	<i>I</i> ₅	...	<i>I</i> ₃₄
<i>I</i> ₁	0	131,7	95,4	124,4	169,7	...	284
<i>I</i> ₂	131,7	0	97,7	65,9	120,4	...	251,5
<i>I</i> ₃	95,4	97,7	0	85	119,3	...	240,2
<i>I</i> ₄	124,4	65,9	85	0	83,5	...	225,4
<i>I</i> ₅	169,7	120,4	119,3	83,5	0	...	247,5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
<i>I</i> ₃₄	284	251,5	240,2	225,4	247,5	...	0

Pada matriks diatas dapat dilihat hasil perhitungan jarak menggunakan *Euclidean distance* dan *Manhattan distance* antara Provinsi Aceh dengan Provinsi Sumatera Utara berturut-turut sebesar 54,16 dan 131,7 sedangkan jarak antara Provinsi Aceh dengan Provinsi Sumatera Barat sebesar 37,92 dan 95,4, Semakin kecil nilai jarak, semakin tinggi tingkat kemiripan antara pasangan data.

Agglomerative Hierarchical

Setelah menghitung jarak masing-masing objek terhadap objek lain menggunakan *Euclidean distance* dan *Manhattan distance*, Maka langkah selanjutnya yaitu menganalisis menggunakan *agglomerative hierarchical clustering*, Proses klasterisasi ini menggunakan metode *single linkage*, *complete linkage*, *average linkage*, dan metode *ward*. Untuk memilih jumlah klaster yang paling sesuai dengan struktur data ada beberapa teknik untuk mengevaluasi jumlah klaster maksimum salah satunya yaitu dengan cara melihat nilai *Pseudo F-Statistic*. Nilai *Pseudo-F* tertinggi menunjukkan bahwa klaster tersebut memberikan hasil yang optimal, dimana keragaman kelompok sangat homogen sedangkan antar kelompok heterogen [8]. Metode ini mengukur variabilitas antar klaster dibandingkan dengan variabilitas dalam klaster, sehingga memberikan hasil yang lebih konsisten [14]. Selain itu, jumlah pengelompokan dengan jarak *Euclidean* dan jarak *Manhattan* pada metode *agglomerative hierarchical* ini menggunakan perkiraan jumlah kelompok 2 sampai 8 dan penentuan banyaknya kelompok yang paling

optimum dilihat berdasarkan nilai *Pseudo F-Statistic* terbesar dan hasilnya dapat dilihat pada Tabel 5:

Tabel 5 Pemilihan Kluster Terbaik

<i>Agglomerative Hierarchical</i>	Perhitungan Jarak	Nilai <i>Pseudo F-statistic</i>	Kluster Optimum
<i>Single Linkage</i>	<i>Euclidean</i>	6,4377	3
	<i>Manhattan</i>	6,4376	3
<i>Complete Linkage</i>	<i>Euclidean</i>	9,7570	2
	<i>Manhattan</i>	8,9007	7
<i>Average Linkage</i>	<i>Euclidean</i>	7,7998	8
	<i>Manhattan</i>	7,6893	6
<i>Ward</i>	<i>Euclidean</i>	9,7570	2
	<i>Manhattan</i>	8,6039	8

Berdasarkan Tabel 5 menghasilkan 8 kluster optimum untuk masing-masing metode *Agglomerative Hierarchical* menggunakan pendekatan metode *pseudo f-statistic* Untuk metode *single linkage* menggunakan perhitungan jarak *Euclidean* dan *Manhattan* masing masing mendapatkan kluster optimum berturut-turut $K = 3$ dan $K = 3$, Untuk metode *complete linkage* menggunakan perhitungan jarak *Euclidean* dan *Manhattan* masing masing mendapatkan kluster optimum berturut-turut $K = 2$ dan $K = 7$, Untuk metode *average linkage* menggunakan perhitungan jarak *Euclidean* dan *Manhattan* masing masing mendapatkan kluster optimum berturut-turut $K = 8$ dan $K = 6$, Untuk metode *ward* menggunakan perhitungan jarak *Euclidean* dan *Manhattan* masing masing mendapatkan kluster optimum berturut-turut $K = 2$ dan $K = 8$.

Ukuran Kebaikan Metode

Setelah mendapatkan hasil kluster optimim setiap metode, maka hal penting selanjutnya adalah menguji kebaikan setiap metode. Uji kebaikan metode ini penting untuk menilai mutu dan kualitas hasil analisis kluster dengan mengevaluasi sejauh mana kluster dapat membedakan data berdasarkan variabel atau karakteristiknya [9]. Salah satu ukuran yang dapat digunakan untuk mengukur kebaikan kluster menggunakan analisis kluster *hierarki* adalah koefisien korelasi *cophenetic*. Koefisien korelasi *cophenetic* adalah koefisien korelasi yang mengukur hubungan antara elemen-elemen yang dihasilkan oleh dendrogram dan elemen-elemen asli dari matriks ketidakmiripan [10].

Rumus untuk menghitung koefisien korelasi *cophenetic* sebagai berikut [15]:

$$r_{coph} = \frac{\sum_{i < j} (Y_{ij} - \bar{Y})(W_{ij} - \bar{W})}{\sqrt{[\sum_{i < j} (Y_{ij} - \bar{Y})^2][\sum_{i < j} (W_{ij} - \bar{W})^2]}}$$

dimana:

- r_{coph} = Koefisien korelasi *cophenetic*.
- Y_{ij} = Jarak asli antara titik ke- i dan ke- j dalam dataset.
- W_{ij} = Jarak *cophenetic* antara titik ke- i dan ke- j .
- \bar{Y} = Nilai rata-rata dari Y_{ij} (jarak *euclid*)
- \bar{W} = Nilai rata-rata dari W_{ij} (jarak *cophenetic*)

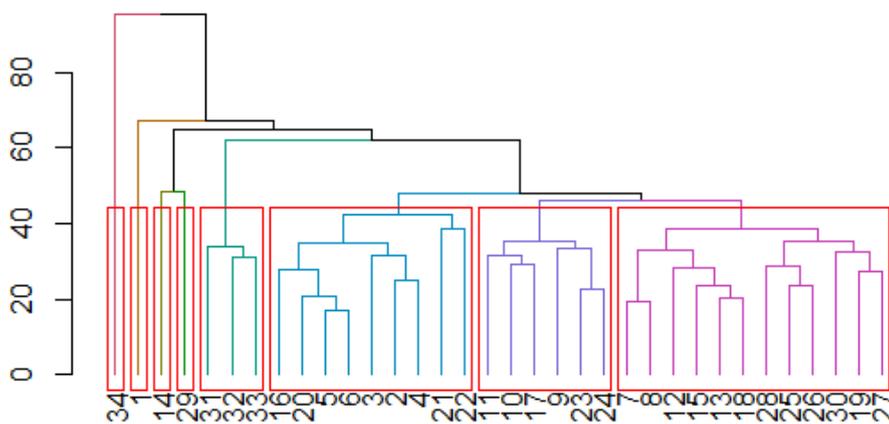
Berikut hasil perhitungan nilai koefisien korelasi *cophenetic* setiap metode:

Tabel 6 Nilai Koefisien Korelasi *Cophenetic*

Metode	Perhitungan Jarak	Koefisien Korelasi <i>Cophenetic</i>	Kluster Optimum
<i>Single</i>	<i>Euclidean</i>	0,7978832	3
<i>Linkage</i>	<i>Manhattan</i>	0,7963331	3
<i>Complete</i>	<i>Euclidean</i>	0,702774	2
<i>Linkage</i>	<i>Manhattan</i>	0,4626763	7
<i>Average</i>	<i>Euclidean</i>	0,8469745	8
<i>Linkage</i>	<i>Manhattan</i>	0,8123014	6
<i>Ward</i>	<i>Euclidean</i>	0,6129641	2
	<i>Manhattan</i>	0,4083691	8

Dari Tabel 6 dapat diketahui bahwa koefisien korelasi *cophenetic* yang tertinggi bernilai 0,8469745 terdapat pada metode *average linkage* dengan perhitungan jarak *Euclidean* yang mempunyai jumlah kelompok optimal yaitu, 8 kluster, Nilai koefisien ini menunjukkan bahwa metode *average linkage* dengan perhitungan jarak *Euclidean* lebih baik dalam mengukur hubungan antara elemen-elemen yang dihasilkan oleh dendrogram dan elemen-elemen asli dari matriks ketidakmiripan, Dengan kata lain, metode *average linkage* dengan perhitungan jarak *Euclidean* mampu merepresentasikan struktur data lebih akurat dibandingkan dengan metode kluster yang lain dengan nilai koefisien korelasi *cophenetic* lebih rendah. Dendrogram metode *average linkage* dengan perhitungan jarak *Euclidean* dengan pembagian kelompok menjadi 8 kluster dapat dilihat pada Gambar 2 berikut ini:

Dendrogram - Average Linkage (Euclidean)



Gambar 2 Dendrogram Hasil Kluster *Average Linkage (Euclidean)*

Perincian anggota kluster indeks khusus penanganan *stunting* di Indonesia dapat dilihat pada Tabel 7 berikut:

Tabel 7 Hasil Kluster Metode *Average Linkage* dengan Jarak *Euclidean*

Klaster	Banyak Anggota Klaster	Kabupaten/Kota
1	1	Aceh
2	9	Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Banten, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan.
3	12	Bengkulu, Lampung, Jawa Barat, Jawa Tengah, Jawa Timur, NTB, NTT, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Sulawesi Barat.
4	6	Kep. Bangka Belitung, Kep. Riau, DKI Jakarta, Bali, Kalimantan Timur, Kalimantan Utara.
5	1	DI Yogyakarta
6	1	Gorontalo
7	3	Maluku, Maluku Utara, Papua Barat
8	1	Papua

Interpretasi Hasil Klaster Terbaik

Dari hasil ukuran kebaikan metode, hasil dengan nilai koefisien korelasi *cophenetic* untuk melihat tertinggi terdapat pada metode *average linkage* dengan perhitungan jarak *Euclidean* dengan menghasilkan kelompok optimum sebanyak 8 klaster, Berikut profilisasi hasil klaster:

Tabel 8 Profilisasi Klaster Terbaik

Variabel	Klaster							
	1	2	3	4	5	6	7	8
Imunisasi	25	57,12	76,85	80,33	93,2	77	61,43	51,3
Penolong persalinan oleh tenaga kesehatan	92,4	84,48	90,87	95,68	98,6	94,9	62,1	72,5
Keluarga Berencana (KB)	53,4	70,23	67,18	59,7	60,1	68,4	45,2	19,7
ASI eksklusif	82,4	84,5	90,56	86,28	94,4	67	79,23	92,7
Makanan Pendamping (MP) ASI	95,9	86,72	85,65	79,77	92,2	68,1	77,57	76,8
Air minum layak	89,7	84,36	88,72	91,03	96,5	96,2	87,27	65,4
Sanitasi layak	77,5	79,24	81,4	90,08	96,2	79,8	76,47	40,3
Ketidakcukupan konsumsi pangan	81,7	83,41	82,28	78,65	77,5	69	49	39,7
Pendidikan Anak Usia Dini (PAUD)	34,2	31,51	40,66	34,68	70	56,2	34,6	14,1

Variabel	Klaster							
	1	2	3	4	5	6	7	8
Kepemilikan JKN/Jamkesda	97	60,69	69,91	77,5	81,5	79,4	65,8	84
Penerima KPS/KKS	48,4	28,73	50,09	21,68	74,2	70,6	29,07	10,3

Dari perhitungan *mean* pada Tabel 8 diketahui variabel apa saja yang menjadi fokus pemerintah dalam penanganan *stunting* tingkat provinsi di Indonesia menggunakan metode *average linkage* dengan diambil 3 variabel yang sangat berpengaruh pada setiap klaster. Pengambilan 3 variabel tersebut dapat dilihat pada Tabel 9 berikut:

Tabel 9 Variabel yang Berpengaruh dalam Klaster

Klaster	Kabupaten/Kota	Variabel
1	Aceh	Imunisasi, Keluarga Berencana (KB), ASI eksklusif, sanitasi layak, PAUD
2	Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Banten, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan.	Kepemilikan JKN/Jamkesda, ketidakcukupan konsumsi pangan, dan air minum layak
3	Bengkulu, Lampung, Jawa Barat, Jawa Tengah, Jawa Timur, NTB, NTT, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Sulawesi Barat.	Kepemilikan JKN/Jamkesda, penolong persalinan oleh tenaga kerja, dan sanitasi layak
4	Kep. Bangka Belitung, Kep. Riau, DKI Jakarta, Bali, Kalimantan Timur, Kalimantan Utara.	Penerimaan KPS/KKS, Keluarga Berencana (KB), dan Makanan pendamping (MP) ASI
5	DI Yogyakarta	Ketidakcukupan konsumsi pangan, Keluarga Berencana (KB), dan kepemilikan JKN/Jamkesda
6	Gorontalo	ASI eksklusif, Makanan Pendamping (MP) ASI, dan sanitasi layak
7	Maluku, Maluku Utara, Papua Barat	Penolong persalinan oleh tenaga kesehatan, sanitasi layak, dan Keluarga Berencana (KB)
8	Papua	Keluarga Berencana (KB), sanitasi layak, dan penerima KPS/KKS

KESIMPULAN

Berdasarkan hasil analisis data dan pembahasan diatas maka dapat diambil kesimpulan analisis yang dilakukan dengan metode *agglomerative hierarchical* dengan pengukuran jarak *Euclidean* dan *Manhattan* menghasilkan pengelompokan terbaik menggunakan metode *average linkage* dengan perhitungan jarak *Euclidean*, Pemilihan metode ini didukung oleh hasil dari nilai koefisien korelasi *cophenetic* tertinggi dari metode yang lainnya. Berdasarkan metode *average linkage* menghasilkan klaster terbaik sebanyak 8 klaster yaitu diantaranya, klaster 1 terdiri dari 1 provinsi, klaster 2 terdiri dari 9 provinsi, klaster 3 terdiri dari 12 provinsi, klaster 4 terdiri dari 6 provinsi, klaster 5 terdiri dari 1 provinsi, klaster 6 terdiri dari 1 provinsi, klaster 7 terdiri dari 3 provinsi, dan klaster 8 terdiri dari 1 provinsi.

DAFTAR PUSTAKA

- [1] Rahmadhita, K. (2020). Permasalahan Stunting dan Pencegahannya. *Jurnal Ilmiah Kesehatan Sandi Husada*, 11(1), 225–229. <https://doi.org/10.35816/jiskh.v10i2.253>
- [2] Candra, A. (2020). *Epidemiologi Stunting*. Fakultas Kedokteran Universitas Diponegoro.
- [3] BPS. (2022). *Laporan Indeks Khusus Penanganan Stunting 2020-2021*. Badan Pusat Statistik.
- [4] De Onis, M., Borghi, E., Arimond, M., Webb, P., Croft, T., Saha, K., De-Regil, L. M., Thuita, F., Heidkamp, R., Krasevec, J., Hayashi, C., & Flores-Ayala, R. (2019). Prevalence thresholds for wasting, overweight and stunting in children under 5 years. *Public Health Nutrition*, 22(1), 175–179. <https://doi.org/10.1017/S1368980018002434>.
- [5] BPS. (2021). *Laporan Indeks Khusus Penanganan Stunting 2020-2021*. Badan Pusat Statistik.
- [6] Pratiwi, S. I., Widiyah, T., & Hakim, A. R. (2019). Analisis Klaster Metode Ward Dan Average Linkage Dengan Validasi Dunn Index Dan Koefisien Korelasi Cophenetic (Studi Kasus: Kecelakaan Lalu Lintas Berdasarkan Jenis Kendaraan Tiap Kabupaten/Kota di Jawa Tengah Tahun 2018). *Jurnal Gaussian*, 8(4), 486–495. <https://doi.org/10.14710/j.gauss.v8i4.26747>
- [7] Nugraha, A., Asnawi, M. H., Rahmah, P. F., & Purwandari, T. (2021). Analisis Klaster Hirarki untuk Mengelompokkan Provinsi di Indonesia berdasarkan Indikator Kesejahteraan Rakyat. *E-Prosiding Seminar Nasional Statistika| Departemen Statistika FMIPA Universitas Padjadjaran*, 10, 1–12.
- [8] Wahyuni, I., & Wulandari, S. P. (2022). Pemetaan Kabupaten / Kota di Jawa Timur Berdasarkan Indikator Kesejahteraan Rakyat Menggunakan Analisis Cluster Hierarki. *Jurnal Sains Dan Seni*, 11(1), D70–D75.
- [9] Pratiwi, S. I., Widiyah, T., & Hakim, A. R. (2019). Analisis Klaster Metode Ward Dan Average Linkage Dengan Validasi Dunn Index Dan Koefisien Korelasi Cophenetic (Studi Kasus: Kecelakaan Lalu Lintas Berdasarkan Jenis Kendaraan Tiap Kabupaten/Kota di Jawa Tengah Tahun 2018). *Jurnal Gaussian*, 8(4), 486–495. <https://doi.org/10.14710/j.gauss.v8i4.26747>
- [10] Zen, M. A., Wahyuningsih, S., & Dani, A. T. R. (2022). Aplikasi Pendekatan Agglomerative Hierarchical Time Series Clustering untuk Peramalan Data Harga Minyak Goreng di Indonesia. *Seminar Nasional Official Statistics*, 2022(1), 293–302.

<https://doi.org/10.34123/semnasoffstat.v2022i1.1394>

- [11] F.Hair, J., C. Black, W., J.Babin, B., & E.Anderson, R. (2010). *Multivariate Data Analysis* (Seventh). Pearson Prentice Hall.
- [12] Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis* (Sixth). Pearson Education.
- [13] Lubis, I. A., Nugroho, S., & Swita, B. (2018). Kajian Metode Pengklasteran Hirarki dengan Berbagai Pengukuran Jarak. *Sigma Mu Rho (e-Jurnal Statistika)*, 42–53.
- [14] Amaliana, D. D., Wulandari, S. P., & Bisnis, D. S. (2024). *Pengelompokan Provinsi berdasarkan Aspek Pembangunan Pendidikan di Indonesia Tahun 2023 menggunakan Analisis Cluster*.
- [15] Jafarzadegan, M., Safi-Esfahani, F., & Beheshti, Z. (2019). Combining hierarchical clustering approaches using the PCA method. *Expert Systems with Applications*, 137, 1–10. <https://doi.org/10.1016/j.eswa.2019.06.064>