

## Penerapan Algoritma *Divisive Analysis* dalam Pengelompokan Provinsi di Indonesia Berdasarkan Prevalensi *Stunting*

### *Application of the Divisive Analysis Algorithm in Clustering Provinces in Indonesia Based on Stunting Prevalence*

Ari Krisna Suyono<sup>1</sup>, Memi Nor Hayati<sup>2 a)</sup>, Meiliyani Siringoringo<sup>3</sup>, Surya Prangga<sup>4</sup>,  
M. Fathurahman<sup>5</sup>

<sup>1, 2, 3, 4, 5</sup>Program Studi Statistika, Jurusan Matematika, FMIPA, Universitas Mulawarman, Indonesia

<sup>1, 2</sup>Laboratorium Statistika Terapan, Jurusan Matematika, FMIPA, Universitas Mulawarman, Indonesia

<sup>3</sup>Laboratorium Statistika Ekonomi dan Bisnis, Jurusan Matematika, FMIPA, Universitas Mulawarman, Indonesia

<sup>4, 5</sup>Laboratorium Statistika Komputasi, Jurusan Matematika, FMIPA, Universitas Mulawarman, Indonesia

<sup>a)</sup>Corresponding author: [meminorhayati@fmipa.unmul.com](mailto:meminorhayati@fmipa.unmul.com)

#### ABSTRACT

Cluster analysis is an analysis that aims to group data (objects) based only on the information contained in the data that describes objects and the relationships between the objects. Divisive analysis is a clustering method using a top-down approach which starts by placing all objects into one cluster or what is called a hierarchical root and then dividing the cluster root into several smaller clusters. This research aimed to group 34 provinces in Indonesia into 2, 3, and 4 clusters based on stunting prevalence data and factors causing stunting in 2022 using a divisive analysis algorithm. The results showed that for 2 clusters, cluster 1 consisted of 32 provinces with low stunting prevalence, and cluster 2 consisted of 2 provinces with high stunting prevalence. For 3 clusters, cluster 1 consisted of 26 provinces with moderate stunting prevalence, cluster 2 consisted of 6 provinces with low stunting prevalence, and cluster 3 consisted of 2 provinces with high stunting prevalence. For 4 clusters, cluster 1 consisted of 21 provinces with moderate stunting prevalence, cluster 2 consisted of 5 provinces with low stunting prevalence, cluster 3 consisted of 6 provinces with high stunting prevalence, and cluster 4 consisted of 2 provinces with very high stunting prevalence.

**Keywords:** cluster analysis, divisive analysis, stunting prevalence.

#### 1. Pendahuluan

Analisis kluster atau juga disebut sebagai *segmentation* adalah alat yang digunakan untuk mengelompokkan sejumlah objek berdasarkan variabel yang memiliki kesamaan karakteristik di antara objek-objek tersebut. Objek-objek ini bisa berupa barang, jasa, hewan, atau manusia (responden, konsumen, dan sebagainya). Objek-objek tersebut akan dikelompokkan ke dalam satu atau lebih kluster, sehingga objek-objek dalam satu kluster akan memiliki kemiripan atau kesamaan. Sebagian besar algoritma analisis kluster membangun kelompok melalui serangkaian pengulangan dan berhenti ketika batasan dari pengelompokan telah tercapai (Supranto, 2010; Vlandari, 2017).

Ciri dari kluster yang baik adalah lebih besar kemiripannya dalam satu kelompok dan lebih besar perbedaannya di antara kelompok lainnya (Prasetyo, 2012). Salah satu analisis kluster yang sering digunakan adalah analisis kluster hierarki. Analisis kluster hierarki dibagi menjadi dua jenis yaitu *agglomerative* dan *divisive*. Analisis hierarki *agglomerative* merupakan metode pengelompokan hierarki dengan pendekatan *bottom-up*, salah satu contohnya adalah *agglomerative nesting*. Analisis hierarki *divisive* merupakan metode pengelompokan hierarki dengan pendekatan *top-down*, salah satu contohnya adalah *divisive analysis* (Gorenescu, 2011). Algoritma *divisive analysis* dimulai dengan menempatkan semua objek ke dalam satu kluster atau disebut dengan akar hierarki kemudian membagi akar kluster menjadi beberapa kluster yang lebih kecil (Kumarahadi, Pratiwi, & Subanti, 2023). Kelebihan utama dari *divisive analysis* adalah dapat menangani data dengan struktur hierarki yang kompleks juga memiliki fleksibilitas dalam menentukan jumlah kluster akhir (Jain, Murty, & Flynn, 1999).

Pada penelitian Kasoqi, Hayati, dan Goejantoro (2021) mengenai algoritma *divisive analysis* tentang pengelompokan desa atau kelurahan berdasarkan potensi desa di Kutai Kartanegara menghasilkan 2 kluster. Kluster 1 beranggotakan 230 desa/kelurahan dan kluster 2 beranggotakan 2 kelurahan. Pada penelitian Satriawan, dan Styawan (2021), menjelaskan beberapa faktor penyebab *stunting* antara lain akses rumah tangga terhadap sanitasi dan sumber air minum, inisiasi menyusui dini, pemberian ASI eksklusif, pemberian imunisasi lengkap, pendapatan rumah tangga, serta akses rumah tangga terhadap makanan bergizi. Pengelompokan dilakukan menggunakan analisis *ward* dan menghasilkan 4 kluster. Kluster pertama beranggotakan 16 provinsi

merupakan kelompok dengan faktor *stunting* tinggi. Klaster kedua beranggotakan 8 provinsi merupakan kelompok dengan faktor *stunting* sedang. Klaster ketiga beranggotakan 6 provinsi merupakan kelompok dengan faktor *stunting* rendah. Klaster empat beranggotakan 4 provinsi merupakan kelompok dengan faktor *stunting* sangat tinggi.

Menurut Kementerian Kesehatan (2023), berdasarkan data Survei Status Gizi Indonesia (SSGI) tahun 2022, prevalensi *stunting* di Indonesia menurun sebesar 2,8% dibandingkan dengan tahun 2021. Meskipun ada penurunan, angka 2,8% ini masih di bawah target yang ditetapkan, yaitu 3,4% per tahun. Dengan penurunan sebesar 2,8% pada tahun 2022, maka untuk mencapai target pada tahun 2024, prevalensi *stunting* harus turun sebesar 7,6% dalam dua tahun ke depan. Hal ini menjadi tantangan besar dalam pelaksanaan program-program selama dua tahun mendatang. Keterbatasan sumber daya mengharuskan pemerintah untuk mengidentifikasi provinsi yang memerlukan perhatian dan penanganan khusus terlebih dahulu terkait kasus balita *stunting*.

Menurut Sandjojo & Madjid (2018), beberapa faktor yang menjadi penyebab *stunting* yaitu praktek pengasuhan yang kurang baik (kurangnya pengetahuan ibu, tidak mendapatkan Air Susu Ibu (ASI) secara eksklusif), kurangnya akses rumah tangga ke makanan bergizi (keterbatasan ekonomi dalam rumah tangga, kemampuan untuk mendapatkan makanan bergizi), kurangnya akses akses air minum bersih, serta masih terbatasnya layanan kesehatan. Keterbatasan sumber daya menyebabkan pemerintah harus mengetahui provinsi mana yang perlu mendapatkan perhatian dan penanganan khusus terlebih dahulu terkait kasus balita *stunting*. Berdasarkan uraian dari latar belakang, untuk membantu proses tercapainya target penanganan *stunting* dilakukan penelitian untuk pengelompokan Provinsi di Indonesia menjadi 2, 3, dan 4 klaster.

## 2. Tinjauan Pustaka

### 2.1 Analisis Klaster

Analisis klaster adalah teknik yang digunakan untuk mengelompokkan objek-objek ke dalam klaster atau kelompok yang relatif homogen. Tujuan dari analisis ini adalah untuk mengidentifikasi kelompok objek yang memiliki kemiripan tertentu sehingga dapat dibedakan dari kelompok lainnya (Supranto, 2010). Terdapat dua metode analisis klaster yaitu pengelompokkan hierarki dan non-hierarki.

Pengelompokkan hierarki adalah analisis klaster yang bertujuan membangun hierarki kelompok. Metode ini digunakan untuk mengelompokkan pengamatan secara terstruktur berdasarkan kemiripannya, dengan jumlah kelompok yang diinginkan belum diketahui. Strategi pengelompokkan hierarki umumnya dibagi menjadi dua jenis, yaitu *agglomerative* dan *divisive*. Pengelompokkan hierarki aglomeratif merupakan metode pengelompokkan hierarki dengan pendekatan bawah-atas (*bottom-up*). Metode hierarki *agglomerative* terdiri dari *single linkage*, *complete linkage*, *average linkage*, *unweighted pair-group average*, *weighted pair-group average*, *unweighted pair-group centroid*, *weighted pair-group*, dan *ward's method*. Dalam *divisive* ada dua metode yaitu *divisive analysis* dan *Automatic Interaction Detection (AID)* (Matjik dan Sumertajaya, 2011).

Berbeda dengan metode hierarki, metode non hierarki dimulai dengan menentukan terlebih dahulu jumlah klaster yang diinginkan. Setelah jumlah klaster diketahui, baru proses pengelompokkan dilakukan tanpa mengikuti proses hierarki. Terdapat tiga metode dalam analisis klaster non-hierarki yaitu *K-Means*, *K-Harmonic Means*, dan *K-Modes*. Hasil pengklasteran non-hierarki tergantung pada urutan observasi dalam data (Prasetyo, 2010).

### 2.2 Algoritma *Divisive Analysis*

Pengelompokkan *divisive analysis* merupakan pengelompokkan hierarki dengan menggunakan pendekatan *top-down*, pendekatan *top-down* menempatkan semua objek dalam satu klaster besar atau disebut akar hierarki, kemudian membagi akar hierarki menjadi beberapa sub-klaster yang lebih kecil (Han, Kamber, dan Pei, 2012). Menurut Rachmatin & Sawitri (2019), algoritma *divisive analysis* adalah proses pengelompokkan yang didasarkan pada nilai rata-rata antar objek. Jika suatu objek memiliki nilai rata-rata tertinggi, maka objek tersebut akan dipisahkan dan menjadi *splinter group*. Dalam algoritma ini, perhitungan dilakukan dengan melihat selisih antara nilai rata-rata dan nilai elemen matriks yang sudah menjadi kelompok splinter. Jika selisih antara nilai rata-rata dan nilai elemen matriks kelompok splinter negatif, maka perhitungan dihentikan dan matriks baru dibuat untuk membentuk klaster lainnya. Proses ini terus berlanjut hingga semua objek terpisah satu sama lain.

## 3. Metodologi Penelitian

Variabel yang digunakan berjumlah delapan variabel yaitu prevalensi *stunting* (%), dan faktor penyebab *stunting* diantaranya rerata lama sekolah perempuan (Tahun), pemberian ASI eksklusif (%), makanan pendamping ASI (%), kesenjangan ekonomi (rasio), ketidakcukupan pangan (rasio), akses air minum layak (%), dan posyandu aktif (%) di Indonesia pada tahun 2022. Tahapan analisis menggunakan bantuan *software R*, berikut tahapan analisis pada penelitian ini.

1. Melakukan analisis statistika deskriptif pada data prevalensi *stunting* di Indonesia tahun 2022 pada 34

provinsi berikut:

**Tabel 1.** Nama dan Kode Provinsi

| Provinsi             | Kode | Provinsi            | Kode |
|----------------------|------|---------------------|------|
| Aceh                 | 1    | Nusa Tenggara Barat | 18   |
| Sumatera Utara       | 2    | Nusa Tenggara Timur | 19   |
| Sumatera Barat       | 3    | Kalimantan Barat    | 20   |
| Riau                 | 4    | Kalimantan Tengah   | 21   |
| Jambi                | 5    | Kalimantan Selatan  | 22   |
| Sumatera Selatan     | 6    | Kalimantan Timur    | 23   |
| Bengkulu             | 7    | Kalimantan Utara    | 24   |
| Lampung              | 8    | Sulawesi Utara      | 25   |
| Kep, Bangka Belitung | 9    | Sulawesi Tengah     | 26   |
| Kep, Riau            | 10   | Sulawesi Selatan    | 27   |
| DKI Jakarta          | 11   | Sulawesi Tenggara   | 28   |
| Jawa Barat           | 12   | Gorontalo           | 29   |
| Jawa Tengah          | 13   | Sulawesi Barat      | 30   |
| DI Yogyakarta        | 14   | Maluku              | 31   |
| Jawa Timur           | 15   | Maluku Utara        | 32   |
| Banten               | 16   | Papua Barat         | 33   |
| Bali                 | 17   | Papua               | 34   |

- Melakukan standardisasi data dengan menggunakan Persamaan (1) (Han, Kamber, & Pei, 2012).

$$\hat{x}_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \tag{1}$$

- Mendeteksi multikolinieritas dengan menggunakan Persamaan (2) (Montgomery dan Peck, 1992).

$$VIF_j = \frac{1}{1 - R_j^2} \tag{2}$$

$$\text{dengan } R_j^2 = 1 - \frac{\sum_{i=1}^n (x_{ij} - \hat{x}_{ij})^2}{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2} \tag{3}$$

dimana:

$VIF_j$  = Variance Inflation Factor variabel ke- $j$

$R_j^2$  = koefisien determinasi dari variabel ke- $j$

$X_{ij}$  = data pengamatan ke- $i$  variabel ke- $j$

$\hat{X}_{ij}$  = nilai prediksi data pengamatan ke- $i$  variabel ke- $j$

$\bar{X}_j$  = rata-rata variabel ke- $j$

$R_j^2$  merupakan koefisien determinasi variabel  $X_j$  dimana  $j = 1, 2, \dots, q$  yang diregresikan dengan variabel lainnya. Jika terdapat nilai  $VIF \geq 10$ , maka dilakukan reduksi variabel dimana variabel dengan nilai  $VIF_j$  terbesar akan dikeluarkan terlebih dahulu.

- Melakukan pengelompokkan dengan menerapkan algoritma *divisive analysis* untuk setiap  $C = 2, 3$ , dan 4 dengan tahapan sebagai berikut:

- Menghitung jarak *Euclidean* untuk setiap objek pengamatan berdasarkan Persamaan (4) (Suyanto, 2017).

$$d(i, r) = \sqrt{(x_{i1} - x_{r1})^2 + (x_{i2} - x_{r2})^2 + \dots + (x_{iq} - x_{rq})^2} \tag{4}$$

- Menghitung rata-rata setiap objek berdasarkan matriks jarak dengan Persamaan (5) (Rachmatin & Sawitri, 2019).

$$\bar{d}_i = \frac{1}{n-1} \sum_{r=1}^n d(i, r), r \neq i \tag{5}$$

- c. Menentukan objek yang memiliki nilai rata-rata yang terbesar, objek yang memiliki nilai rata-rata yang terbesar akan dipisah dan berubah menjadi *splinter group*.
  - d. Menghitung selisih nilai antara elemen matriks *splinter group* dengan nilai rata-rata setiap objek yang tersisa.
  - e. Menentukan objek yang memiliki nilai selisih terbesar. Jika nilai selisih tersebut bernilai positif, maka objek terbesar bergabung dengan *splinter group*. Mengulang langkah (b) sampai langkah (e) sedemikian sehingga semua nilai selisih antara elemen matriks *splinter group* dengan nilai rata-rata bernilai negatif dan klaster terbagi menjadi dua klaster baru.
5. Profilisasi klaster yang terbentuk.

**4. Hasil dan Pembahasan**

Analisis statistika deskriptif dilakukan untuk melihat gambaran sederhana dari data prevalensi *stunting* dan faktor penyebab *stunting* pada 34 provinsi di Indonesia tahun 2022 dengan menampilkan rerata, simpangan baku, nilai minimum dan maksimum untuk semua variabel, ditampilkan pada Tabel 2.

**Tabel 2.** Statistika Deskriptif

| Nama Variabel                                   | Rerata | Simpangan Baku | Nilai Minimum | Nilai Maksimum |
|---|--------|----------------|---------------|----------------|
| Prevalensi <i>Stunting</i> (%) ( $X_1$ )        | 23,29  | 6,38           | 8,00          | 35,30          |
| Rerata Lama Sekolah Perempuan (Tahun) ( $X_2$ ) | 8,55   | 0,98           | 6,29          | 10,97          |
| Pemberian ASI Eksklusif (%) ( $X_3$ )           | 69,21  | 7,15           | 53,60         | 79,69          |
| Makanan Pendamping ASI (Rasio) ( $X_4$ )        | 83,90  | 8,41           | 66,90         | 98,70          |
| Kesenjangan Ekonomi (Rasio) ( $X_5$ )           | 0,34   | 0,05           | 0,24          | 0,44           |
| Ketidakcukupan Pangan (%) ( $X_6$ )             | 13,67  | 8,20           | 2,24          | 36,18          |
| Akses Air Minum Layak (%) ( $X_7$ )             | 87,64  | 7,74           | 65,39         | 98,42          |
| Posyandu Aktif (%) ( $X_8$ )                    | 56,41  | 29,76          | 3,4           | 100,00         |

Variabel yang digunakan pada penelitian ini didasarkan pada penelitian sebelumnya oleh Satriawan, dan Styawan (2021) yang melakukan pengelompokan provinsi di Indonesia berdasarkan faktor penyebab balita *stunting*.

Setelah melakukan analisis statistika deskriptif diketahui bahwa terdapat perbedaan satuan diantara variabel penelitian, maka perlu dilakukan standardisasi data untuk merubah nilai data asli menjadi bentuk *Z-Score* untuk membuat rentang nilai yang sama pada semua variabel. Hasil perhitungan standardisasi dapat dilihat pada Tabel 3.

**Tabel 3.** Standardisasi Data

| Provinsi       | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ | $X_8$ |
|----------------|-------|-------|-------|-------|-------|-------|-------|-------|
| Aceh           | 1,24  | 0,71  | -0,47 | -0,74 | -0,33 | -0,36 | 0,27  | -0,29 |
| Sumatera Utara | -0,34 | 0,97  | -1,68 | -0,74 | -0,61 | 0,14  | 0,58  | -1,18 |
| Sumatera Barat | 0,30  | 0,63  | 0,71  | -0,96 | -0,78 | -1,22 | -0,31 | -0,66 |
| ⋮              | ⋮     | ⋮     | ⋮     | ⋮     | ⋮     | ⋮     | ⋮     | ⋮     |
| Maluku Utara   | 0,44  | 0,39  | -0,05 | -1,40 | 2,08  | -0,17 | 0,06  | 1,46  |
| Papua Barat    | 1,05  | -0,85 | -1,04 | 0,56  | 1,92  | -0,78 | -0,78 | -1,64 |
| Papua          | 1,77  | -2,31 | 0,70  | 1,44  | 2,75  | -4,22 | -2,88 | -1,78 |

Setelah melakukan standardisasi selanjutnya perlu mendeteksi ada atau tidaknya multikolinieritas diantara variabel penelitian dengan menghitung nilai *VIF*. Nilai *VIF* untuk semua variabel ditampilkan pada Tabel 4.

**Tabel 4.** Nilai *VIF* Setiap Variabel

| Variabel | $R^2$ | <i>VIF</i> |
|----------|-------|------------|
| $X_1$    | 0,30  | 1,43       |
| $X_2$    | 0,49  | 1,95       |
| $X_3$    | 0,41  | 1,69       |
| $X_4$    | 0,44  | 1,79       |
| $X_5$    | 0,49  | 1,98       |
| $X_6$    | 0,30  | 1,45       |

|       |      |      |
|-------|------|------|
| $X_7$ | 0,41 | 1,70 |
| $X_8$ | 0,25 | 1,33 |

Hasil perhitungan menyatakan bahwa nilai *VIF* untuk semua variabel bebas bernilai di bawah 10 yang artinya tidak terjadi multikolinieritas antar variabel penelitian dan dapat dilanjutkan ke proses pengelompokan dengan algoritma *divisive analysis*. Algoritma *divisive analysis* dimulai dengan  $C = 2$ . Langkah pertama menghitung jarak *Euclidean* dengan menggunakan data hasil standardisasi kemudian menghitung rata-rata jarak *Euclidean* untuk setiap objek. Hasil perhitungan jarak *Euclidean* dan rata-rata jarak untuk setiap objek ditampilkan pada Tabel 5.

**Tabel 5.** Matriks Jarak *Euclidean* dan Rata-rata Jarak Objek

| $d(i,j)$ | 1    | 2    | 3    | ... | 15   | 16   | 17   | ... | 32   | 33   | 34   | $\bar{d}_i$ |
|----------|------|------|------|-----|------|------|------|-----|------|------|------|-------------|
| 1        | 0,00 | 3,06 | 1,77 | ... | 2,84 | 3,24 | 5,19 | ... | 3,90 | 4,42 | 6,50 | 3,55        |
| 2        | 3,06 | 0,00 | 3,22 | ... | 2,76 | 2,92 | 4,47 | ... | 4,31 | 4,02 | 7,02 | 3,81        |
| 3        | 1,77 | 3,22 | 0,00 | ... | 2,61 | 2,69 | 4,87 | ... | 4,14 | 4,68 | 6,36 | 3,36        |
| ⋮        | ⋮    | ⋮    | ⋮    | ⋮   | ⋮    | ⋮    | ⋮    | ⋮   | ⋮    | ⋮    | ⋮    | ⋮           |
| 15       | 2,84 | 2,76 | 2,61 | ... | 0,00 | 1,27 | 2,93 | ... | 3,90 | 4,25 | 6,47 | 3,03        |
| 16       | 3,24 | 2,92 | 2,69 | ... | 1,27 | 0,00 | 3,03 | ... | 4,25 | 4,71 | 6,79 | 3,23        |
| 17       | 5,19 | 4,47 | 4,87 | ... | 2,93 | 3,03 | 0,00 | ... | 5,03 | 5,73 | 7,67 | 4,50        |
| ⋮        | ⋮    | ⋮    | ⋮    | ⋮   | ⋮    | ⋮    | ⋮    | ⋮   | ⋮    | ⋮    | ⋮    | ⋮           |
| 32       | 3,90 | 4,31 | 4,14 | ... | 3,90 | 4,25 | 5,03 | ... | 0,00 | 4,16 | 6,10 | 4,15        |
| 33       | 4,42 | 4,02 | 4,68 | ... | 4,25 | 4,71 | 5,73 | ... | 4,16 | 0,00 | 3,41 | 4,42        |
| 34       | 6,50 | 7,02 | 6,36 | ... | 6,47 | 6,79 | 7,67 | ... | 6,10 | 3,41 | 0,00 | <b>6,24</b> |

Selanjutnya nilai objek dengan rata-rata terbesar akan dipisahkan dan menjadi *splinter group* 1. Jika dilihat dari Tabel 4, nilai rata-rata jarak terbesar ada pada objek-34 yaitu Provinsi Papua sebesar 6,24 sehingga objek-34 menjadi *splinter group* 1. Selanjutnya menghitung ulang rata-rata jarak dari objek yang tersisa. Hasil dari rata-rata jarak baru dikurangkan dengan elemen dari objek-34. Apabila selisih pengurangan bernilai positif maka objek dengan nilai selisih terbesar menjadi anggota *splinter group* 1. Apabila selisih pengurangan bernilai negatif maka perhitungan untuk mencari anggota *splinter group* 1 berhenti. Berikut hasil perhitungan dapat dilihat pada Tabel 6.

**Tabel 6.** Selisih antara Nilai Elemen Matriks *Splinter Group* 1 (Objek 34)

| $i$ | $\bar{d}_i$ | $d(i,34)$ | Selisih     |
|-----|-------------|-----------|-------------|
| 1   | 3,45        | 6,50      | -3,05       |
| 2   | 3,71        | 7,02      | -3,31       |
| 3   | 3,27        | 6,36      | -3,09       |
| ⋮   | ⋮           | ⋮         | ⋮           |
| 15  | 2,92        | 6,47      | -3,55       |
| 16  | 3,12        | 6,79      | -3,67       |
| 17  | 4,40        | 7,67      | -3,27       |
| ⋮   | ⋮           | ⋮         | ⋮           |
| 31  | 4,41        | 6,28      | -1,87       |
| 32  | 4,09        | 6,10      | -2,01       |
| 33  | 4,45        | 3,41      | <b>1,04</b> |

Objek-33 memiliki nilai selisih terbesar positif yaitu 1,04 sehingga perhitungan untuk pembentukan 2

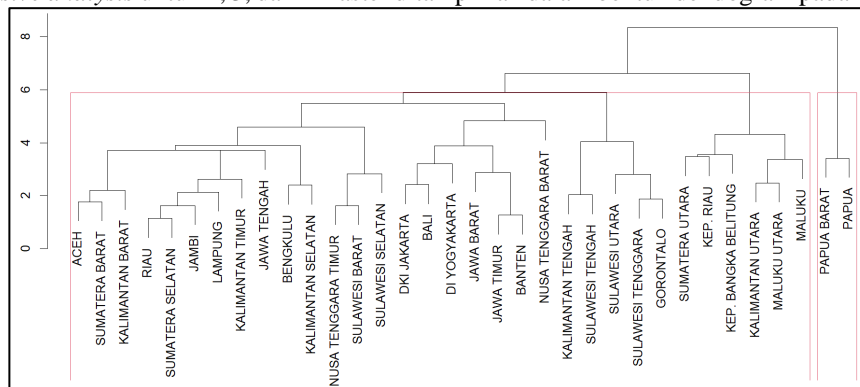
klaster berlanjut. Objek-33 menjadi anggota *splinter group* 1 bergabung dengan objek ke-34. Selanjutnya menghitung ulang rata-rata jarak dari objek yang tersisa. Hasil dari rata-rata jarak baru dikurangkan dengan rata-rata elemen dari objek-34 dan objek-33 yang telah menjadi anggota *splinter group* 1. Apabila selisih pengurangan bernilai positif maka objek dengan nilai selisih terbesar menjadi anggota *splinter group* 1. Apabila selisih pengurangan bernilai negatif maka perhitungan untuk mencari anggota *splinter group* 1 berhenti. Berikut hasil perhitungan dapat dilihat pada Tabel 7.

**Tabel 7.** Selisih antara Nilai Elemen Matriks *Splinter Group* 1 (Objek 34 dan 33)

| $i$ | $\bar{d}_i$ | $\frac{d(i,33) + d(i,34)}{2}$ | Selisih      |
|-----|-------------|-------------------------------|--------------|
| 1   | 3,42        | 5,46                          | -2,04        |
| 2   | 3,70        | 5,52                          | -1,82        |
| 3   | 3,22        | 5,52                          | -2,30        |
| ⋮   | ⋮           | ⋮                             | ⋮            |
| 15  | 2,88        | 5,36                          | -2,48        |
| 16  | 3,07        | 5,75                          | -2,68        |
| 17  | 4,36        | 6,70                          | -2,34        |
| ⋮   | ⋮           | ⋮                             | ⋮            |
| 30  | 3,78        | 3,89                          | <b>-0,11</b> |
| 31  | 4,44        | 4,85                          | -0,40        |
| 32  | 4,09        | 5,13                          | -1,04        |

Objek-30 memiliki nilai selisih terbesar bernilai negatif yaitu -0,11 maka perhitungan untuk pembentukan 2 klaster berhenti. Objek-34 dan Objek-33 menjadi anggota *splinter group* 1. Selanjutnya menghitung ulang rata-rata jarak yang tersisa, nilai rata-rata jarak terbesar akan menjadi anggota *splinter group* baru. Proses dilakukan untuk pembentukan 3 klaster hingga pembentukan 4 klaster dengan nilai selisih terbesar bernilai negatif pada *splinter group* 3.

Setelah algoritma *divisive analysis* berhenti maka didapatkan hasil pengelompokan menggunakan algoritma *divisive analysis* untuk 2, 3, dan 4 klaster ditampilkan dalam bentuk dendogram pada Gambar berikut.



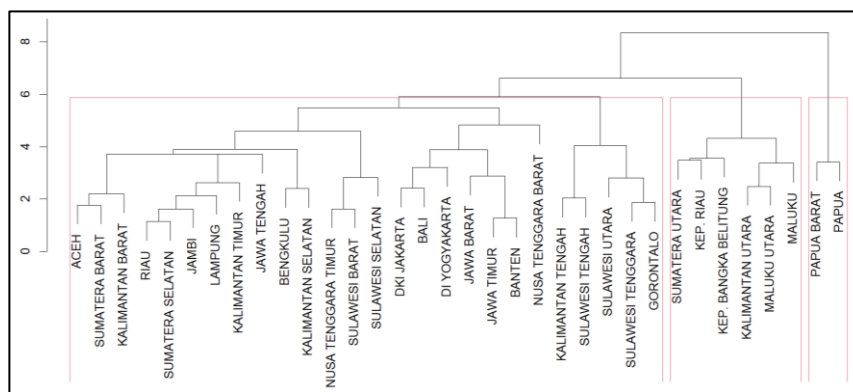
**Gambar 1.** Dendogram Pembentukan 2 Klaster

Gambar 1 menunjukkan hasil pada proses pembentukan 2 klaster yang mana anggota klaster 1 berjumlah 32 provinsi (Aceh, Sumatra Utara, dan lainnya kecuali Papua Barat dan Papua). Anggota klaster 2 berjumlah 2 provinsi (Papua Barat dan Papua). Hasil pembentukan 2 klaster dapat digambarkan wilayah dari setiap klaster dengan Peta Indonesia, ditampilkan pada Gambar 2.



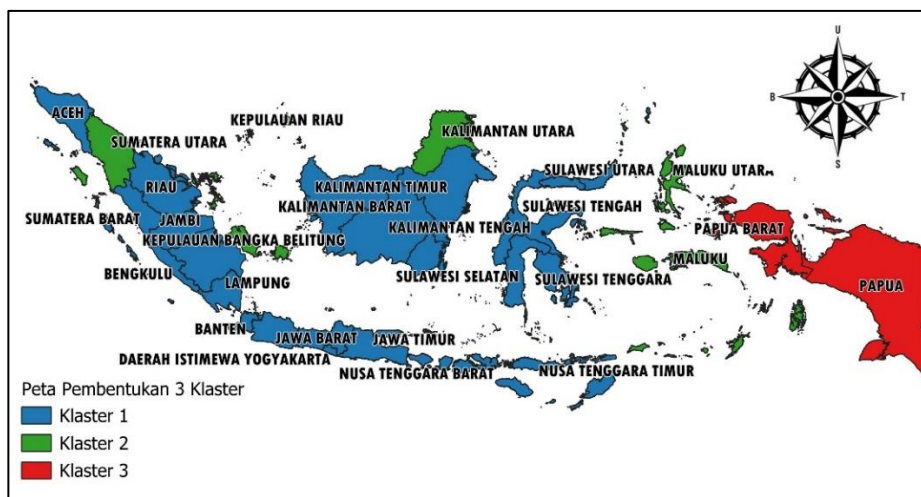
Gambar 2. Peta Pembentukan 2 Klaster

Selanjutnya melihat dendogram hasil pembentukan 3 klaster, ditampilkan pada Gambar 3.



Gambar 3. Dendogram Pembentukan 3 Klaster

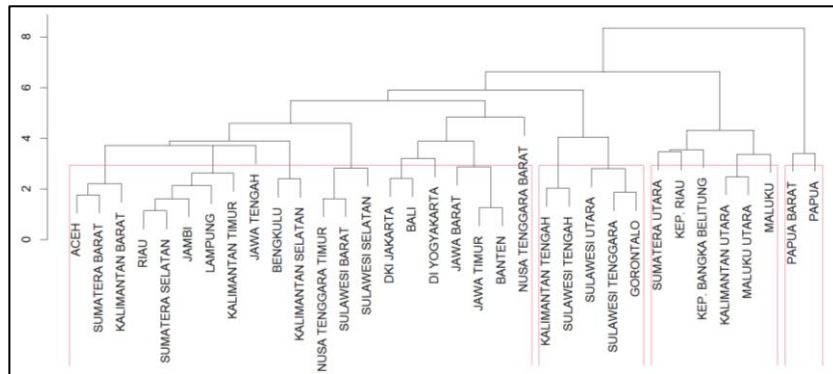
Pada proses pembentukan 3 klaster didapatkan hasil bahwa anggota klaster 1 berjumlah 26 provinsi (Aceh, Sumatra Barat, .... , Gorontalo, Sulawesi Barat), anggota klaster 2 berjumlah 6 provinsi (Sumatra Utara, Kep. Bangka Belitung, .... , Maluku, Maluku Utara), dan anggota klaster 3 berjumlah 2 provinsi (Papua Barat, Papua). Hasil pembentukan 3 klaster dapat digambarkan wilayah dari setiap klaster dengan Peta Indonesia, ditampilkan pada Gambar 4.



Gambar 4. Peta Pembentukan 3 Klaster

Selanjutnya melihat dendogram hasil pembentukan 4 klaster, ditampilkan pada Gambar 5.





Gambar 5. Dendrogram Pembentukan 4 Kluster

Pada proses pembentukan 4 kluster didapatkan hasil bahwa anggota kluster 1 berjumlah 21 provinsi (Aceh, Sumatra Barat, ..... , Sulawesi Selatan, Sulawesi Barat) , anggota kluster 2 berjumlah 5 provinsi (Sumatra Utara, Kep. Bangka Belitung, ..... , Maluku, Maluku Utara), anggota kluster 3 berjumlah 6 provinsi (Kalimantan Tengah, Sulawesi Utara, ..... , Sulawesi Tenggara, Gorontalo), dan anggota kluster 4 berjumlah 2 provinsi (Papua Barat, Papua). Hasil pembentukan 4 kluster dapat digambarkan wilayah dari setiap kluster dengan Peta Indonesia, ditampilkan pada Gambar 6.



Gambar 6. Peta Pembentukan 4 Kluster

Langkah selanjutnya adalah profilisasi hasil kluster, dengan menghitung rata-rata dari setiap variabel disetiap kluster. Hasil perhitungan di tampilkan pada Tabel 8.

Tabel 8. Profilisasi Setiap Pembentukan Kluster

| Variabel | Rata-rata             |           |                       |           |           |                       |           |           |           |
|----------|-----------------------|-----------|-----------------------|-----------|-----------|-----------------------|-----------|-----------|-----------|
|          | Pembentukan 2 Kluster |           | Pembentukan 3 Kluster |           |           | Pembentukan 4 Kluster |           |           |           |
|          | Kluster 1             | Kluster 2 | Kluster 1             | Kluster 2 | Kluster 3 | Kluster 1             | Kluster 2 | Kluster 3 | Kluster 4 |
| $X_1$    | 22,73                 | 32,30     | 23,00                 | 21,55     | 32,30     | 22,42                 | 21,55     | 25,42     | 32,30     |
| $X_2$    | 8,65                  | 7,01      | 8,36                  | 9,91      | 7,01      | 8,36                  | 9,91      | 8,34      | 7,01      |
| $X_3$    | 69,29                 | 67,98     | 70,29                 | 64,92     | 67,98     | 72,77                 | 64,92     | 59,91     | 67,98     |
| $X_4$    | 84,45                 | 75,15     | 85,67                 | 79,13     | 75,15     | 88,73                 | 79,13     | 72,82     | 75,15     |
| $X_5$    | 0,34                  | 0,39      | 0,35                  | 0,29      | 0,39      | 0,35                  | 0,29      | 0,36      | 0,39      |
| $X_6$    | 12,48                 | 32,78     | 10,73                 | 20,10     | 32,78     | 10,10                 | 20,10     | 13,35     | 32,78     |
| $X_7$    | 88,53                 | 73,48     | 88,37                 | 89,18     | 73,48     | 88,05                 | 89,18     | 89,74     | 73,48     |
| $X_8$    | 59,59                 | 5,55      | 58,53                 | 64,18     | 5,55      | 61,29                 | 64,18     | 46,92     | 5,55      |

Tabel 8 menunjukkan bahwa secara deskriptif dapat diketahui pada pembentukan 2 kluster, profil kluster 1 merupakan kluster dengan prevalensi *stunting* rendah sedangkan profil kluster 2 merupakan kluster dengan prevalensi *stunting* tinggi. Secara deskriptif dapat diketahui bahwa pada pembentukan 3 kluster, profil kluster



2 merupakan kluster dengan prevalensi *stunting* rendah diantara kluster lainnya, profil kluster 1 merupakan kluster dengan prevalensi *stunting* sedang, dan profil kluster 3 merupakan kluster dengan prevalensi *stunting* tinggi. Secara deskriptif dapat diketahui bahwa pada pembentukan 4 kluster, profil kluster 2 merupakan kluster dengan prevalensi *stunting* rendah diantara kluster lainnya, profil kluster 1 merupakan kluster dengan prevalensi *stunting* sedang, profil kluster 3 merupakan kluster dengan prevalensi *stunting* tinggi, dan profil kluster 4 merupakan kluster dengan prevalensi *stunting* sangat tinggi.

## 5. Kesimpulan

Hasil dari algoritma *divisive analysis* pada pembentukan 2, 3, dan 4 kluster. Pada proses pembentukan 2 kluster, ditemukan bahwa kluster 1 terdiri dari 32 provinsi dan kluster 2 terdiri dari 2 provinsi. Kluster 1 memiliki prevalensi *stunting* rendah, sedangkan kluster 2 memiliki prevalensi *stunting* tinggi. Pada proses pembentukan 3 kluster, ditemukan bahwa kluster 1 terdiri dari 26 provinsi, kluster 2 terdiri dari 6 provinsi, dan kluster 3 terdiri dari 2 provinsi. Kluster 2 memiliki prevalensi *stunting* paling rendah di antara kluster lainnya, kluster 1 memiliki prevalensi *stunting* sedang, dan kluster 3 memiliki prevalensi *stunting* tinggi. Pada proses pembentukan 4 kluster, ditemukan bahwa kluster 1 terdiri dari 21 provinsi, kluster 2 terdiri dari 5 provinsi, kluster 3 terdiri dari 6 provinsi, dan kluster 4 terdiri dari 2 provinsi. Kluster 2 memiliki prevalensi *stunting* paling rendah di antara kluster lainnya, kluster 1 memiliki prevalensi *stunting* sedang, kluster 3 memiliki prevalensi *stunting* tinggi, dan kluster 4 memiliki prevalensi *stunting* sangat tinggi. Hasil dari algoritma *divisive analysis* ini diharapkan dapat membantu tercapainya target pemerintah di tahun 2024.

Pada penelitian ini belum dilakukan pemilihan kluster optimal, sehingga untuk penelitian selanjutnya dapat menggunakan metode evaluasi untuk menentukan kluster optimal seperti menggunakan *Davies-Bouldin Index*, *Elbow Method*, dan *Gap Statistic*. Selain itu penelitian selanjutnya juga dapat menggunakan metode pengelompokan hierarki lain yaitu *Monothetic Divisive* atau dapat juga menggunakan *Agglomerative Nesting* sebagai salah satu alternatif untuk memecahkan permasalahan serupa.

## Daftar Pustaka

- Gorenescu, F. (2011). *Data mining: Concept, Model, And Technique*. Berlin: Springer.
- Govender, P., & Sivakumar, V. (2020). Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019). In *Atmospheric Pollution Research* (Vol. 11, Issue 1). Turkish National Committee for Air Pollution Research and Control. <https://doi.org/10.1016/j.apr.2019.09.009>
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and Techniques Third Edition*. San Fransisco: Morgan Kauffman Publisher.
- Kasoqi, I. A., Hayati, M. N., & Goejantoro, R. (2021). Pengelompokan Desa Atau Kelurahan di Kutai Kartanegara Menggunakan Algoritma Divisive Analysis. *Jurnal Statistika Unimus*, 9(2), 101-108.
- Kementerian Kesehatan RI. (2022). *Buku Saku Hasil Survei Status Gizi Indonesia*. Jakarta: Kementerian Kesehatan.
- Kumarahadi, B. M., Pratiwi, H., & Subanti, S. (2023). Penerapan Metode Hierarchical Clustering untuk Pengelompokan Kota/Kabupaten di Indonesia Berdasarkan Indikator Kemiskinan. *Jurnal TIKomSiN*, 11(2), 8 – 12.
- Matjik, A. A., & Sumertajaya, M. (2011). *Sidik Peubah Ganda dengan Menggunakan SAS*. Bogor: IPB Press.
- Montgomery, D. C. & Peck. (1992). *Introduction to Linear Regression Analysis*. New York: John Wiley & Sons.
- Prasetyo, E. (2010). *Data Mining dan Aplikasi Menggunakan MATLAB*. Yogyakarta: ANDI.
- Prasetyo, E. (2012). *Data mining: Konsep dan Aplikasi menggunakan Matlab*. Yogyakarta: Andi Offset.
- Rachmatin, D., & Sawitri, K. (2019). Perbandingan Antara Metode Agglomeratif, Divisive, dan K-Means Dalam Analisis Cluster. *Jurnal Pengajaran MIPA*, 24(1).
- Sandjojo, E. P., & Majid, T. (2017). *Buku Saku Desa dalam Penanganan Stunting*. Jakarta: Kementerian Desa Pembangunan Daerah Tertinggal dan Transmigrasi.
- Satriawan, D., & Styawan, D. A. (2021). Pengelompokan Provinsi di Indonesia Berdasarkan Faktor Penyebab Balita Stunting Menggunakan Analisis Cluster Hierarki. *Jurnal Statistika dan Aplikasinya*, 5 (1). 61-70.
- Supranto, J. (2010). *Analisis Multivariat: Arti dan Interpretasi*. Jakarta: Rineka Cipta.
- Suyanto. (2017). *Data mining untuk Klasifikasi dan Klasterisasi Data*. Bandung: Informatika.
- Vulandari, R. T. (2017). *Data mining Teori dan Aplikasi Rapidminer*. Yogyakarta: Penerbit Gava Media.