

ANALISIS CLUSTER WILAYAH KECAMATAN BERDASARKAN FAKTOR PENYEBAB STUNTING DI PROVINSI KEPULAUAN BANGKA BELITUNG

Fadillah Luthfiyah

Program Studi Matematika, Fakultas Sains dan Teknik, Universitas Bangka Belitung, Indonesia

e-mail: diyah29102003@gmail.com*

Ririn Amelia

Program Studi Matematika, Fakultas Sains dan Teknik, Universitas Bangka Belitung, Indonesia

e-mail: ririn-amelia@ubb.ac.id

Izma Fahria

Program Studi Matematika, Fakultas Sains dan Teknik, Universitas Bangka Belitung, Indonesia

e-mail: fahriaizma@yahoo.com

Abstrak

Stunting merupakan permasalahan gizi yang sampai saat ini masih menjadi perhatian dunia, termasuk di Indonesia. Berdasarkan data Survei Kesehatan Indonesia Tahun 2023, pada 38 provinsi yang ada di Indonesia terdapat 15 provinsi yang mengalami kenaikan prevalensi *stunting*. Provinsi Kepulauan Bangka Belitung menjadi salah satu provinsi yang mengalami kenaikan prevalensi *stunting* di tahun 2023 dengan angka sebesar 20,6%. Kenaikan angka prevalensi *stunting* tersebut tidak terlepas dari berbagai faktor-faktor yang dapat mempengaruhi *stunting*. Beberapa wilayah kabupaten/kota juga masih menghadapi tantangan besar dalam upaya penurunan *stunting*. Berdasarkan uraian tersebut, maka akan dilakukan penelitian untuk menganalisis *cluster* wilayah kecamatan yang ada di Provinsi Kepulauan Bangka Belitung berdasarkan faktor penyebab terjadinya *stunting* dengan menggunakan metode *Agglomerative Nesting*. Melalui analisis *cluster*, diharapkan dapat diidentifikasi karakteristik faktor penyebab *stunting* di setiap wilayah sehingga lebih mudah dipahami. Hasil penelitian menunjukkan bahwa metode *Agglomerative Nesting* (AGNES) dengan *average linkage* merupakan metode yang paling optimal dengan koefisien korelasi *cophenetic* sebesar 0,7463. Selain itu, jumlah *cluster* yang optimal terbentuk sebanyak 3 *cluster*. Adapun *cluster* 1 memiliki karakteristik faktor penyebab *stunting* yang sudah cukup baik, namun masih perlu peningkatan pada indikator bayi < 6 bulan yang mendapat ASI eksklusif. Sedangkan pada *cluster* 2, karakteristik faktor penyebab *stunting* cukup buruk pada hampir semua indikator jika dibandingkan dengan *cluster* lainnya. Pada *cluster* 3, karakteristik faktor penyebab *stunting* sudah cukup baik. Namun tetap diperlukan perhatian pada indikator ibu hamil yang anemia serta akses sanitasi yang aman.

Kata Kunci: *Stunting*, Analisis Cluster, *Agglomerative Nesting*.

Abstract

Stunting is a nutritional problem that is still a global concern, including in Indonesia. Based on data from the 2023 Indonesian Health Survey, out of 38 provinces in Indonesia, 15 provinces experienced an increase in *stunting* prevalence. The Bangka Belitung Islands Province is one of the provinces that experienced an increase in *stunting* prevalence in 2023 with a figure of 20.6%. The increase in the *stunting* prevalence rate cannot be separated from various factors that can affect *stunting*. Several districts/cities also still face major challenges in efforts to reduce *stunting*. Based on this description, a study will be conducted to analyze the cluster of sub-districts in the Bangka Belitung Islands Province based on the factors causing *stunting* using the *Agglomerative Nesting* method. Through cluster analysis, it is hoped that the characteristics of the factors causing *stunting* in each region can be identified so that they are easier to understand. The results of the study showed that the *Agglomerative Nesting* (AGNES) method with *average linkage* is the most optimal method with a *cophenetic* correlation coefficient of 0.7463. In addition, the optimal number of clusters formed is 3 clusters. Cluster 1 has characteristics of factors causing *stunting* that are quite good, but still needs improvement in the indicator of infants <6 months who receive exclusive breastfeeding. While in cluster 2, the characteristics of factors causing *stunting* are quite bad in almost all indicators when compared to other clusters. In cluster 3, the characteristics of factors causing *stunting* are quite good. However, attention is still needed to the indicators of pregnant women with anemia and access to safe sanitation.

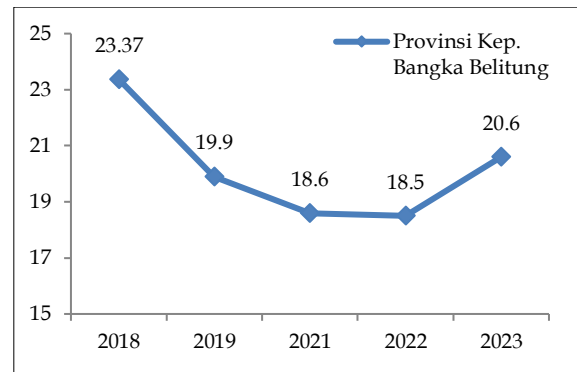
Keywords: *Stunting*, Cluster Analysis, *Agglomerative Nesting*.

PENDAHULUAN

Stunting adalah suatu kondisi gangguan pertumbuhan anak akibat kekurangan asupan gizi dan adanya infeksi kronis dalam waktu yang lama sehingga anak tumbuh lebih pendek dari anak normal seusianya (WHO, 2015). Sampai saat ini, permasalahan *stunting* masih menjadi perhatian dunia khususnya di negara-negara berkembang. Di Indonesia, permasalahan *stunting* tergolong kronis dan menjadi salah satu prioritas pemerintah dalam program pembangunan nasional (Pratiwi, 2023).

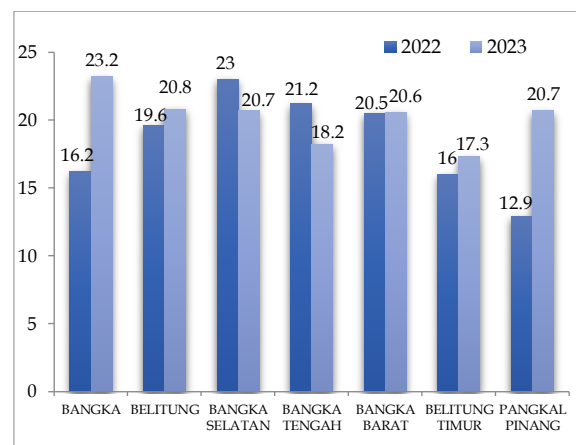
Menurut laporan Studi Status Gizi Indonesia (SSGI) Kementerian Kesehatan, prevalensi *stunting* di Indonesia turun dari 24,4% di tahun 2021, menjadi 21,6% di tahun 2022. Berdasarkan hasil Survei Kesehatan Indonesia (SKI) tahun 2023, prevalensi *stunting* turun hanya sebesar 0,01% yakni 21,5% dengan kasus terbanyak pada kelompok usia 2 sampai 3 tahun (Badan Kebijakan Pembangunan Kesehatan, 2024). Perkembangan angka prevalensi *stunting* sampai saat ini masih belum mampu mencapai target Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2020-2024, yakni menurunnya angka prevalensi *stunting* menjadi 14% pada tahun 2024 (Perpres RI, 2020). Oleh karena itu, pemerintah masih perlu meningkatkan berbagai upaya dan strategi untuk mempercepat penurunan angka *stunting*. Berdasarkan data SKI Tahun 2023, pada 38 provinsi yang ada di Indonesia terdapat 15 provinsi yang mengalami kenaikan prevalensi *stunting* pada tahun 2023.

Provinsi Kepulauan Bangka Belitung menjadi salah satu provinsi yang mengalami kenaikan prevalensi *stunting* di tahun 2023. Prevalensi *stunting* di Provinsi Kepulauan Bangka Belitung berdasarkan data Survei Kesehatan Indonesia (SKI) tahun 2023 sebesar 20,6%. Berikut ini merupakan grafik yang menunjukkan angka prevalensi *stunting* di Provinsi Kepulauan Bangka Belitung dari tahun 2018 sampai tahun 2023.



Gambar 1. Grafik Prevalensi *Stunting* Provinsi Kepulauan Bangka Belitung Tahun 2018 – 2023

Gambar 1 menunjukkan bahwa terjadi kenaikan prevalensi *stunting* pada tahun 2023 sebesar 2,1%. Kenaikan angka prevalensi *stunting* tersebut tidak terlepas dari berbagai faktor-faktor yang dapat mempengaruhi *stunting*. Adapun faktor penyebab *stunting* yang terekam melalui Survei Kesehatan Indonesia (SKI) 2023 adalah faktor ibu, anak dan rumah tangga (Badan Kebijakan Pembangunan Kesehatan, 2024). Faktor ibu dapat ditinjau melalui kondisi kesehatan dan gizi ibu sebelum kehamilan, saat kehamilan hingga setelah persalinan. Faktor anak berkaitan dengan kondisi anak pada saat lahir dan setelah kelahiran. Sedangkan faktor pada level rumah tangga antara lain akses air minum layak dan akses sanitasi aman. Selain itu, adanya ketimpangan dalam hal upaya penurunan *stunting* antar daerah di Provinsi Kepulauan Bangka Belitung, juga dapat menyebabkan hasil penurunan prevalensi *stunting* di setiap wilayah bisa berbeda-beda. Hal ini dapat dilihat melalui Gambar 2



Gambar 2. Grafik Prevalensi *Stunting* Kabupaten/Kota Tahun 2022-2023

Berdasarkan Gambar 2, dapat dilihat bahwa angka prevalensi *stunting* Kota Pangkalpinang mengalami kenaikan paling tinggi sebesar 7,8%,

Kabupaten Bangka sebesar 7%, Kabupaten Belitung Timur sebesar 1,3%, Kabupaten Belitung sebesar 1,2% dan Kabupaten Bangka Barat sebesar 0,1%. Sedangkan Kabupaten Bangka Tengah mengalami penurunan prevalensi *stunting* sebesar 3% dan Kabupaten Bangka Selatan sebesar 2,4%. Hal tersebut menunjukkan beberapa wilayah masih menghadapi tantangan besar dalam upaya penurunan *stunting*. Ketimpangan ini menunjukkan perlunya pendekatan yang lebih spesifik dalam upaya penurunan *stunting* dengan menyesuaikan karakteristik wilayah masing-masing. Oleh karena itu, diperlukan analisis pengelompokan wilayah berdasarkan faktor penyebab *stunting* untuk memahami karakteristik setiap wilayah secara lebih mendalam.

Salah satu analisis yang digunakan dalam mengelompokkan data untuk mengetahui karakteristik dalam data yakni analisis *cluster*. Analisis *cluster* memiliki konsep dasar dengan mengelompokkan objek ke dalam suatu *cluster* berdasarkan tingkat kemiripan yang tinggi, sedangkan objek-objek antar *cluster* yang berbeda memiliki tingkat ketidakmiripan yang tinggi (Farissa *et al.*, 2021). Penelitian ini akan menggunakan metode analisis *Hierarchical Clustering* dengan pendekatan *Agglomerative Nesting* (AGNES). Pada analisis menggunakan AGNES, jumlah *cluster* yang dibentuk dilakukan dengan membangun hierarki atau tingkatan data berdasarkan kesamaan karakteristik objek yang disebut *dendrogram* dengan pendekatan *bottom-up* (Ginatra *et al.*, 2021).

Berdasarkan uraian tersebut, maka akan dilakukan penelitian untuk menganalisis *cluster* wilayah kecamatan yang ada di Provinsi Kepulauan Bangka Belitung berdasarkan faktor penyebab terjadinya *stunting* dengan menggunakan metode *Agglomerative Nesting* (AGNES). Melalui analisis *cluster*, diharapkan dapat diidentifikasi karakteristik faktor penyebab *stunting* di setiap wilayah sehingga lebih mudah dipahami. Hasil analisis ini diharapkan dapat membantu pengalokasian sumber daya secara lebih efisien dan efektif, serta membantu dalam merumuskan kebijakan yang tepat sasaran dalam mengatasi kenaikan prevalensi *stunting* di Provinsi Kepulauan Bangka Belitung.

KAJIAN TEORI (GUNAKAN STYLE SECTION) STUNTING

Stunting merupakan salah satu masalah gizi yang sampai saat ini masih menjadi perhatian dunia dan umumnya banyak terjadi di negara-negara berkembang. *Stunting* merupakan kondisi ketika pertumbuhan anak balita terhambat, yang disebabkan karena kekurangan gizi kronis sehingga anak tumbuh terlalu pendek untuk usianya. Kekurangan gizi biasanya terjadi pada saat bayi masih dalam kandungan dan setelah bayi lahir, namun *stunting* baru terlihat setelah bayi berusia 2 tahun (Samsuddin *et al.*, 2023).

Balita yang mengalami *stunting* dapat dilihat melalui status gizi yang didasarkan pada indeks panjang badan menurut umur (PB/U) atau tinggi badan menurut umur (TB/U). Menurut standar antropometri penilaian status gizi anak, apabila hasil pengukuran yang berada pada ambang batas (*Z-Score*) < -2 SD sampai dengan -3 SD dikategorikan pendek (*stunted*). Sedangkan apabila hasil pengukuran yang berada pada ambang batas (*Z-score*) < -3 SD dikategorikan sangat pendek (*severely stunted*) (Hutabarat, 2023).

Faktor penyebab *stunting* dari faktor ibu dapat ditinjau melalui kondisi kesehatan dan gizi ibu. Ibu hamil yang menderita anemia dan kekurangan energi kronik dapat meningkatkan risiko terjadinya kelahiran bayi dengan Berat Badan Lahir Rendah (BBLR) (Suryati *et al.*, 2022). Selama kehamilan, ibu hamil mengalami peningkatan volume darah yang lebih besar dibandingkan dengan peningkatan jumlah sel darah merah dan hemoglobin di dalam sel darah merah (Kemenkes RI, 2021). Selain itu, ibu hamil mengalami KEK juga bisa berdampak pada pertumbuhan dan sirkulasi plasenta ibu menuju janin yang dapat menyebabkan masalah pertumbuhan pada bayi yang dilahirkan.

Adapun faktor penyebab *stunting* dari faktor bayi berkaitan dengan kondisi pada saat lahir dan setelah kelahiran. Kondisi anak yang lahir dengan Berat Badan Lahir Rendah (BBLR) yakni < 2.500 gram dapat berisiko mengalami *stunting* bahkan kematian pada bayi (Tarmizi, 2023). Ibu harus melakukan Inisiasi Menyusui Dini (IMD), pemberian ASI eksklusif selama 6 bulan dilanjutkan dengan penambahan pemberian ASI sampai usia 2 tahun dengan pemberian makanan pendamping yang bergizi, dan imunisasi dasar yang lengkap untuk mencegah *stunting* (Rusliani *et al.*, 2022). ASI mengandung berbagai hormon dan enzim, yang

berperan dalam mendukung perkembangan jaringan tubuh serta membantu mencegah infeksi. Melalui pemberian ASI eksklusif, risiko terjadinya infeksi dapat diminimalkan, sehingga anak lebih terlindungi dari masalah kesehatan yang dapat menghambat pertumbuhannya (Meliyana, 2024).

Faktor yang mempengaruhi *stunting* pada level rumah tangga berdasarkan hasil Survei Kesehatan Indonesia (SKI) 2023 antara lain akses air minum layak dan akses sanitasi aman dan layak (Badan Kebijakan Pembangunan Kesehatan, 2024). Hal ini didukung oleh Peraturan Presiden Republik Indonesia Nomor 72 Tahun 2021, bahwa Strategi Nasional Penurunan *Stunting* salah satunya bertujuan untuk meningkatkan akses air minum dan sanitasi (Perpres, 2021). Untuk memastikan kualitas air minum yang dikonsumsi masyarakat, penting untuk melakukan pengawasan kualitas air minum secara berkala. Selain itu, keluarga yang tidak memiliki fasilitas sanitasi yang memadai berisiko mengalami berbagai penyakit infeksi seperti diare dan cacingan, yang akan menghambat proses penyerapan nutrisi dan mengganggu tumbuh kembang balita (Kuewa *et al.*, 2021).

ANALISIS CLUSTER

Analisis *cluster* merupakan salah satu teknik analisis multivariat yang digunakan untuk mengelompokkan objek berdasarkan kesamaan dan ketidaksamaan karakteristiknya (Musfiani, 2019). Terdapat beberapa jenis analisis *cluster* antara lain *hierarchical clustering* dan *non-hierarchical clustering*. Metode *hierarchical clustering* terdiri dari dua pendekatan yaitu *Agglomerative* dan *Divisive*. Metode *agglomerative* merupakan pendekatan *bottom-up*, di mana setiap objek dianggap sebagai sebuah kelompok sendiri dan kemudian menggabungkannya menjadi satu kelompok besar (Ginantra *et al.*, 2021). Sedangkan pendekatan *Divisive Analysis* (DIANA) merupakan kebalikan dari metode *agglomerative* dalam *hierarchical clustering* yang bersifat *top-down*, di mana seluruh objek data dianggap sebagai satu kelompok besar dan kemudian dipisahkan menjadi kelompok yang lebih kecil (Ginantra *et al.*, 2021).

Sedangkan pada metode *non-hierarchical clustering* terdapat beberapa algoritma antara lain algoritma *K-Means* dan *K-Medoids*. *K-Means* merupakan metode untuk membagi sekumpulan

data menjadi k *cluster* yang berbeda, di mana setiap *cluster* ditentukan oleh pusatnya, yaitu rata-rata posisi titik-titik data dalam *cluster* (Permana *et al.*, 2023). Sedangkan *K-medoids* adalah metode yang menggunakan objek dalam kelompok untuk mewakili *cluster* sebagai pusatnya. Objek ini disebut *medoid*, yang terletak di tengah *cluster* dengan jarak terdekat ke titik lainnya, sehingga lebih tahan terhadap *outlier*. *Cluster* dibentuk berdasarkan kedekatan *medoid* dengan objek bukan *medoid* (Permana *et al.*, 2023).

PREPROCESSING DATA

Preprocessing data merupakan salah satu tahapan yang bertujuan untuk mengatasi permasalahan yang dapat memengaruhi hasil analisis data (Daniswara dan Nuryana, 2023). Salah satu tahapan dalam *preprocessing data* adalah melakukan standarisasi data. Standarisasi data dilakukan untuk mengatasi perbedaan satuan yang signifikan diantara variabel yang diteliti, karena dapat menyebabkan perhitungan pada analisis *cluster* menjadi tidak valid. Proses standarisasi dilakukan dengan transformasi pada data asli sebelum dianalisis lebih lanjut. Transformasi dilakukan terhadap variabel yang relevan ke dalam bentuk *z-score* dengan persamaan berikut :

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

Keterangan :

z = nilai *z-score* standarisasi data

x = nilai data

\bar{x} = nilai rata-rata data

s = standar deviasi

Melalui proses standarisasi data, variabel yang ditransformasi ke dalam bentuk *z-score* akan memiliki nilai rata-rata 0 (nol) dan standar deviasi yang bernilai 1 (satu). Selain melakukan standarisasi data, pada *preprocessing data* juga dilakukan pendeteksian *outlier*. Data *outlier* mengacu pada nilai yang sangat tinggi atau sangat rendah (nilai ekstrem) dibanding nilai yang lainnya, yang dapat mempengaruhi perhitungan rata-rata dalam kelompok data tersebut (Purnomo *et al.*, 2022).

Pendeteksian data *outlier* dapat dilakukan dengan mengubah data awal ke dalam bentuk *z-score* melalui proses standarisasi data. Suatu data dikatakan *outlier* apabila nilai *z-score* lebih besar dari

+2,5 atau lebih kecil dari -2,5 (Hasrul, 2018). Terdapat beberapa penanganan data *outlier* yakni :

- Transformasi data, membuat distribusi data lebih mendekati normal dengan transformasi logaritma, akar kuadrat atau invers.
- Menghapus data *outlier*, jika data *outlier* disebabkan akibat kesalahan pengukuran, kesalahan input dan sebagainya.
- Mempertahankan data *outlier*, jika data *outlier* mencerminkan variasi yang sebenarnya dalam data dan dianggap penting untuk hasil analisis.

ASUMSI ANALISIS CLUSTER

SAMPEL REPRESENTATIF

Sampel yang representatif merupakan sampel yang diambil sedemikian rupa dan dapat mewakili populasi yang ada. Tidak ada ketentuan dalam menentukan jumlah sampel yang representatif, namun penting untuk memiliki jumlah sampel yang cukup agar proses *clustering* dapat dilakukan dengan benar. Pemenuhan asumsi syarat kecukupan suatu sampel dapat menggunakan uji *Kaiser-Mayer-Olkin* (KMO). KMO merupakan indeks yang membandingkan nilai koefisien korelasi dengan korelasi parsial. Persamaan untuk menghitung nilai KMO adalah sebagai berikut :

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p \alpha_{ij}^2} \quad (2)$$

Keterangan :

r_{ij} = koefisien korelasi antara variabel i dan j

α_{ij} = koefisien korelasi parsial antara variabel i dan j

Sampel dikatakan mewakili populasi atau sampel representatif apabila nilai KMO berkisar antara 0,5 sampai 1 (Nahdliyah *et al.*, 2019). Jika nilai KMO < 0,5, maka sampel dianggap tidak dapat mewakili populasi. Sebaliknya, jika nilai KMO > 0,5, ini menunjukkan bahwa sampel dapat mewakili populasi, sehingga layak untuk dilakukan analisis (Puspitasari dan Susanti, 2016).

UJI MULTIKOLINEARITAS

Uji multikolinearitas bertujuan untuk melihat ada tidaknya hubungan (korelasi) linear yang signifikan antar variabel. Pengujian multikolinearitas dapat dilihat melalui nilai koefisien korelasi *Pearson* antar variabel bebas, jika nilai koefisien korelasi antar variabel > 0,80 maka terjadi multikolinearitas

(Hasrul, 2018). Penanganan multikolinearitas dapat dilakukan beberapa alternatif, antara lain :

- Mengganti atau menghilangkan salah satu variabel yang mempunyai korelasi yang kuat.
- Transformasi variabel, seperti transformasi ke dalam bentuk diferensi pertama, logaritma dan akar kuadrat.
- Menambah jumlah observasi.

JARAK EUCLIDEAN

Pengelompokan data dalam analisis *cluster* menggunakan suatu ukuran jarak (*Distance Measure*), untuk menjelaskan kedekatan antara objek atau data yang menggambarkan kesamaan atau ketidaksamaan dalam data yang kompleks (Hasanah dan Sofro, 2022). Untuk mengukur jarak antar dua objek dalam analisis *cluster*, dapat digunakan pengukuran jarak menggunakan jarak *Euclidean*. Persamaan jarak *Euclidean* adalah sebagai berikut :

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

Keterangan :

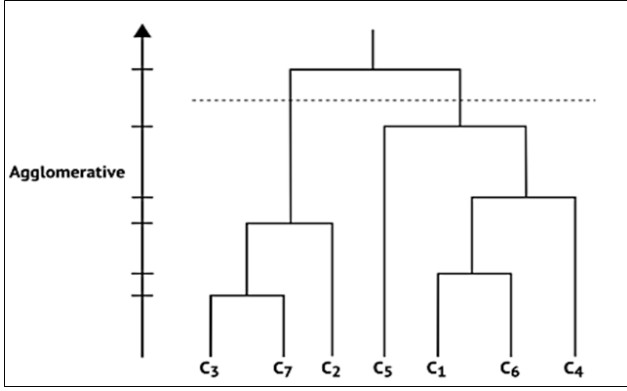
$d(x, y)$ = jarak antara x dan y

x_i = nilai atribut dari data ke- i

y_i = data pada atribut ke- i

AGGLOMERATIVE NESTING (AGNES)

Metode *agglomerative nesting* merupakan analisis *cluster* yang menggunakan pendekatan *bottom-up*, yang dimulai dengan menganggap setiap objek sebagai *cluster* tersendiri. Selanjutnya objek tersebut digabungkan secara bertahap menjadi *cluster* yang lebih besar hingga semua objek tergabung dalam satu *cluster* (Februariyanti dan Winarko, 2010). Hasil analisis *cluster* dapat direpresentasikan melalui sebuah *dendrogram* yang merupakan visualisasi struktur pengelompokan data dalam analisis *cluster hierarchical clustering*. Gambar 3 merupakan contoh hasil analisis *cluster* data menggunakan metode *agglomerative nesting* (Bachtiar *et al.*, 2017).



Gambar 3. Dendrogram Hasil Analisis Cluster
Agglomerative Nesting

Terdapat beberapa metode pengelompokan antar kelompok tersebut antara lain (Izzuddin dan Wijayanto, 2024) :

1. Single Linkage

Metode ini mendefinisikan kesamaan *cluster* berdasarkan jarak minimum atau jarak terdekat. Secara matematis perhitungan jarak antara (a, b) dengan *cluster* lainnya dapat dirumuskan dengan persamaan berikut :

$$d_{(ab)c} = \min \{d_{ac}, d_{bc}\} \quad (5)$$

2. Complete Linkage

Metode ini mendefinisikan kesamaan *cluster* berdasarkan jarak maksimum atau jarak terjauh. Secara matematis perhitungan jarak antara (a, b) dengan *cluster* lainnya dapat dirumuskan dengan persamaan berikut :

$$d_{(ab)c} = \max \{d_{ac}, d_{bc}\} \quad (6)$$

3. Average Linkage

Metode ini mendefinisikan kesamaan *cluster* berdasarkan jarak rata-rata antar seluruh individu. Secara matematis perhitungan jarak antara (a, b) dengan *cluster* lainnya dapat dirumuskan dengan persamaan berikut :

$$d_{(ab)c} = \text{ave}\{d_{ac}, d_{bc}\} \quad (7)$$

Untuk membandingkan setiap *linkage* yang paling optimal dapat menggunakan koefisien korelasi *cophenetic*. Koefisien korelasi *cophenetic* adalah suatu ukuran yang digunakan untuk menilai seberapa baik *dendrogram* yang dihasilkan oleh analisis *cluster* merepresentasikan jarak asli antara objek dalam analisis (Iis et al., 2022). Nilai koefisien korelasi *cophenetic* berkisar antara -1 hingga 1,

apabila semakin mendekati 1 berarti semakin baik *cluster* yang dihasilkan (Faujia et al., 2022).

EVALUASI CLUSTER

METODE ELBOW

Metode Elbow adalah salah satu teknik yang digunakan untuk menentukan jumlah *cluster* yang optimal dengan mengamati titik di mana grafik menunjukkan perubahan mendadak, yang terlihat seperti sudut siku (Ulya et al., 2023). Untuk menentukan nilai k yang optimal, dilakukan dengan membandingkan nilai *Sum of Square Error* (SSE) yang ditampilkan dalam bentuk grafik (Syahfitri et al., 2023). Persamaan SSE sebagai berikut :

$$SSE = \sum_{k=1}^K \sum_{x_i \in C_k} d(x_i - \varphi_k)^2 \quad (8)$$

Keterangan:

K = k *cluster* yang terbentuk

C_k = *cluster* ke- k

x_i = nilai atribut dari data ke- i

φ_k = nilai rata-rata/pusat *cluster* ke- k

Jika nilai dari *cluster* pertama dan *cluster* kedua menghasilkan sudut dalam grafik atau menunjukkan penurunan yang paling signifikan, maka *cluster* tersebut dianggap optimal (Madhulatha, 2015).

SILHOUETTE COEFFICIENT

Silhouette Coefficient merupakan salah satu metode untuk mengevaluasi *cluster* yang terbentuk dengan tujuan untuk membandingkan jarak rata-rata antara suatu titik yang tergabung dalam satu *cluster* dengan titik-titik dalam *cluster* lainnya (Butar, 2023). Nilai *silhouette* objek menggambarkan seberapa dekat objek tersebut dengan *cluster*-nya sendiri dibandingkan dengan *cluster* lainnya (Siregar, 2024). Untuk menghitung nilai *Silhouette Coefficient* menggunakan persamaan berikut :

$$SC_i = \frac{b_i - a_i}{\max(a_i, b_i)} \quad (9)$$

Keterangan:

SC_i = nilai *silhouette coefficient* untuk objek ke- i

a_i = nilai rata-rata objek ke- i dengan objek lainnya yang berada pada satu *cluster*

b_i = nilai minimum rata-rata objek ke- i dengan objek lainnya yang berada pada *cluster* yang berbeda

Jumlah *cluster* yang optimal adalah jumlah *cluster* yang menunjukkan nilai rata-rata koefisien *silhouette*

tertinggi, yang biasanya mendekati angka 1 (Simanjuntak dan Khaira, 2021).

METODE

Penelitian ini menggunakan jenis penelitian kuantitatif dengan pendekatan deskriptif. Penelitian kuantitatif deskriptif umumnya bertujuan untuk mengukur dan menggambarkan tingkat suatu variabel pada populasi atau sampel tertentu (Ali *et al.*, 2022). Hasil dari penelitian ini akan memberikan jawaban terhadap suatu masalah dan mendapatkan informasi lebih luas tentang suatu fenomena dengan menggunakan tahap-tahap pendekatan kuantitatif. Penelitian ini dilakukan terhadap lokasi kecamatan yang ada di Provinsi Kepulauan Bangka Belitung yang terdiri dari 47 kecamatan.

Data yang digunakan dalam penelitian merupakan data sekunder yang dikumpulkan berdasarkan wilayah per-kecamatan di Provinsi Kepulauan Bangka Belitung. Sumber data ini diperoleh dari publikasi Dinas Kesehatan yang ada di masing-masing kabupaten/kota di Provinsi Kepulauan Bangka Belitung. Adapun rentang waktu data yang digunakan adalah data dari tahun 2022 sampai dengan tahun 2023. Variabel yang digunakan dalam penelitian ini adalah faktor penyebab *stunting* dengan indikator j di kecamatan ke- i dapat dilihat pada Tabel 1. Jenis data dari seluruh variabel yang ada dalam penelitian ini berupa data numerik. Data numerik dalam penelitian ini berupa persentase dari masing-masing variabel di setiap lokasi ke- i .

Tabel 1 Variabel Penelitian

Variabel	Keterangan
X_{i1}	Ibu hamil dengan riwayat Anemia
X_{i2}	Ibu hamil dengan riwayat Kekurangan Energi Kronik (KEK)
X_{i3}	Bayi dengan Berat Badan Lahir Rendah (BBLR)
X_{i4}	Bayi diberi ASI eksklusif (usia < 6 bulan)
X_{i5}	Sarana air minum yang diawasi/diperiksa sesuai standar
X_{i6}	Sanitasi yang aman

Tahapan analisis data yang dilakukan pada penelitian ini sebagai berikut :

1. Melakukan *preprocessing* data

Pada tahap ini dilakukan proses standarisasi data dengan cara mengubah data ke dalam bentuk *z-score*

(lihat persamaan (1)) . Jika terdapat data yang *outlier*, maka akan dilakukan penanganan data *outlier*.

2. Uji asumsi dalam analisis *cluster*

Terdapat asumsi yang harus dipenuhi dalam analisis *cluster* yaitu menguji kecukupan sampel dan multikolinearitas. Apabila sampel yang digunakan tidak representatif dengan nilai uji *Kaiser-Meyer-Olkin* (KMO) < 0,5 maka dilakukan penambahan variabel baru. Apabila terjadi multikolinearitas dengan nilai koefisien korelasi > 0,8 maka akan dilakukan penanganan multikolinearitas.

3. Menentukan ukuran kesamaan/kedekatan objek

Menentukan kedekatan objek dalam penelitian ini digunakan pengukuran jarak *Euclidean* dengan menggunakan persamaan (4) sehingga diperoleh matriks jarak dari masing-masing objek/data.

4. Implementasi *Agglomerative Nesting* (AGNES)

Tahap algoritma *Agglomerative Nesting* (AGNES) melibatkan perhitungan jarak antar data berdasarkan perhitungan jarak *Euclidean* serta melibatkan iterasi dalam pembentukan *cluster* sampai dengan jumlah yang ditentukan. Metode *linkage* yang optimal ditentukan menggunakan nilai korelasi koefisien *cophenetic*. Hasil dari algoritma ini dapat ditampilkan dalam bentuk diagram pohon atau *dendogram*.

5. Evaluasi jumlah *cluster* optimal

Evaluasi jumlah *cluster* optimal dilakukan dengan membandingkan nilai indeks validasi dari setiap jumlah *cluster* yang berbeda. Indeks validasi yang digunakan dalam penelitian ini antara lain adalah metode *elbow* dan *silhouette coefficient*. Jumlah *cluster* yang optimal ditentukan berdasarkan yang paling sesuai dengan model evaluasi.

6. Interpretasi *cluster* optimal

Cluster optimal yang terpilih akan menjadi hasil akhir dari penelitian ini. Untuk menginterpretasikan *cluster* dapat dilakukan dengan membandingkan rata-rata (*centroid*) pada masing-masing *cluster*. Hal ini juga dilakukan untuk menjelaskan karakteristik dari setiap *cluster* tersebut.

HASIL DAN PEMBAHASAN

DATA PENELITIAN

Pengumpulan data dilakukan dengan memperoleh data melalui pihak lain yaitu melalui instansi Dinas Kesehatan Kabupaten/Kota di Provinsi Kepulauan Bangka Belitung. Adapun data yang digunakan dalam penelitian ini berupa data

sebanyak enam variabel dari X_1 sampai X_6 (lihat Tabel 1), dengan nilai berupa persentase dari setiap indikator tersebut di masing-masing kecamatan. Data tersebut mencakup 47 kecamatan di Provinsi Kepulauan Bangka Belitung dengan indikator yang diambil dari tahun 2022 sampai dengan tahun 2023. Selanjutnya dilakukan perhitungan rata-rata dari data tersebut yang akan digunakan untuk analisis *cluster* dalam penelitian ini. Proses perhitungan dalam penelitian ini dilakukan menggunakan perangkat lunak *R Studio* dan *Microsoft Excel*.

Setelah data dikumpulkan, data kemudian diintegrasikan sehingga didapatkan data sebanyak 47 kecamatan di Provinsi Kepulauan Bangka Belitung berdasarkan faktor penyebab *stunting*. Adapun tampilan dataset tersebut dapat dilihat pada Tabel 2 berikut.

Tabel 2 Dataset Penelitian

Kec.	X_1	X_2	X_3	X_4	X_5	X_6
Sungailiat	1.51	5.17	5.12	79.34	44.67	48.95
Pemali	2.21	2.49	1.13	75.38	59.72	50.07
...
Sungai Selan	9.02	10.28	3.40	65.80	50.00	0.00
Simpang Katis	1.67	5.22	5.47	71.96	77.27	93.81

PREPROCESSING DATA

STANDARISASI DATA

Meskipun data dalam penelitian ini sudah memiliki satuan yang sama pada setiap variabel-variabelnya, namun proses standarisasi data tetap dilakukan. Hal ini dilakukan untuk memastikan bahwa setiap variabel berada dalam skala yang sama sehingga kualitas data yang akan dianalisis lebih akurat. Dengan perhitungan menggunakan *R Studio* diperoleh hasil data yang telah distandarisasi. Adapun data yang telah distandarisasi dapat dilihat pada Tabel 3. Nilai Z_{X_1} , Z_{X_2} , Z_{X_3} sampai dengan Z_{X_6} merupakan nilai variabel yang datanya telah distandarisasi sehingga nilai pada data setiap kecamatan berada pada rentang $\pm 2,5$. Nilai Z_{X_i} yang positif pada setiap variabel menunjukkan bahwa objek-objek kecamatan tersebut memiliki nilai variabel di atas rata-rata. Sedangkan nilai Z_{X_i} yang negatif pada setiap variabel menunjukkan bahwa objek-objek kecamatan tersebut memiliki nilai variabel di bawah rata-rata.

Tabel 3 Standarisasi Dataset

Kec.	Z_{X_1}	Z_{X_2}	Z_{X_3}	Z_{X_4}	Z_{X_5}	Z_{X_6}
Sungailiat	-0.80	-0.80	-0.03	1.04	-0.28	1.49
Pemali	-0.65	-1.54	-1.41	0.76	0.26	1.54
...
Sungai Selan	0.82	0.62	-0.63	0.07	-0.08	-0.68
Simpang Katis	-0.76	-0.78	0.09	0.51	0.89	3.48

PENDETEKSIAN OUTLIER

Pendeteksian *outlier* dilakukan dengan memeriksa seluruh nilai data dari setiap variabel yang sudah distandarisasi. Jika terdapat data yang nilainya berada di luar rentang $\pm 2,5$ maka nilai tersebut dianggap sebagai *outlier*. Melalui pendeteksian menggunakan *R Studio*, hasil pengamatan terhadap data yang *outlier* dapat dilihat pada Tabel 4

Tabel 4 Data Outlier

Kecamatan	Variabel	Nilai Outlier
Simpang Teritip	Z_{X_1}	2,67
	Z_{X_1}	3,20
Rangkui	Z_{X_1}	3,24
Simpang Renggang	Z_{X_3}	2,69
Simpang Katis	Z_{X_6}	3,47

Berdasarkan Tabel 4 dapat dilihat bahwa terdapat sebanyak 5 data yang *outlier*. Oleh karena itu akan dilakukan penanganan terhadap data *outlier* tersebut. Penanganan data *outlier* dalam penelitian ini dilakukan dengan tetap mempertahankan nilai tersebut. Hal ini dikarenakan nilai *outlier* tersebut merupakan representasi dari populasi yang diteliti. Data yang *outlier* ini memberikan gambaran bahwa kecamatan tersebut pada variabel tertentu memiliki karakteristik yang sangat berbeda dibanding kecamatan lainnya.

UJI ASUMSI

SAMPEL REPRESENTATIF

Uji asumsi kecukupan sampel dilakukan untuk memastikan bahwa ukuran sampel yang digunakan dalam penelitian cukup representatif untuk analisis yang akan dilakukan. Pengujian sampel yang representatif dalam penelitian ini menggunakan indeks *Kaiser-Meyer-Olkin* (KMO). Perhitungan indeks KMO menggunakan *R Studio* diperoleh hasil pada Gambar 4


```
> KMO(Data_Standarisasi)
Kaiser-Meyer-olkin factor adequacy
Call: KMO(r = Data_Standarisasi)
Overall MSA = 0.58
MSA for each item =
  X1  X2  X3  X4  X5  X6
0.56 0.58 0.59 0.61 0.44 0.64
```

Gambar 4 Hasil Uji KMO

Jika nilai KMO > 0,5 maka data dalam penelitian sudah layak untuk dilakukan analisis. Oleh karena itu, berdasarkan Gambar 4 diketahui bahwa nilai indeks *Kaiser-Meyer-Olkin* (KMO) diperoleh sebesar 0.58. Nilai KMO tersebut sudah lebih dari 0,5 yang artinya sampel dalam penelitian ini telah cukup dan layak untuk dilakukan analisis. Namun dapat dilihat nilai KMO pada variabel X_5 masih < 0,5 sehingga dalam penelitian ini, variabel tersebut tidak diikutsertakan dalam proses analisis *cluster*.

TIDAK TERJADI MULTIKOLINEARITAS

Pengujian multikolinearitas dalam penelitian ini dilakukan dengan melakukan analisis terhadap nilai-nilai korelasi antar variabel pada matriks korelasi. Perhitungan matriks korelasi untuk mengetahui nilai-nilai korelasi antar variabel menggunakan *R Studio* dapat dilihat pada Tabel 5.

Tabel 5 Matriks Korelasi Antar Variabel

Variabel	X_1	X_2	X_3	X_4	X_5	X_6
X_1	1.00	0.55	0.27	-0.11	0.14	-0.32
X_2	0.55	1.00	0.64	-0.30	-0.09	-0.27
X_3	0.27	0.64	1.00	-0.38	-0.02	-0.20
X_4	-0.11	-0.30	-0.38	1.00	0.35	0.49
X_5	0.14	-0.09	-0.02	0.35	1.00	0.15
X_6	-0.32	-0.27	-0.20	0.49	0.15	1.00

Tabel 5 menunjukkan bahwa seluruh variabel dalam penelitian ini tidak memiliki nilai korelasi antar variabel yang lebih dari 0,80. Hal ini berarti bahwa tidak terjadi multikolinearitas antar variabel dalam penelitian ini, sehingga proses analisis *cluster* bisa dilanjutkan.

ANALISIS CLUSTER

Pada analisis *cluster* dengan metode *agglomerative nesting*, seluruh objek pada awalnya dianggap sebagai satu *cluster* sendiri yang terpisah satu sama lain. Setelah itu dilakukan perhitungan jarak dan penggabungan objek-objek yang memiliki kesamaan karakteristik ke dalam satu *cluster*. Perhitungan untuk memperbarui matriks jarak antar *cluster* baru

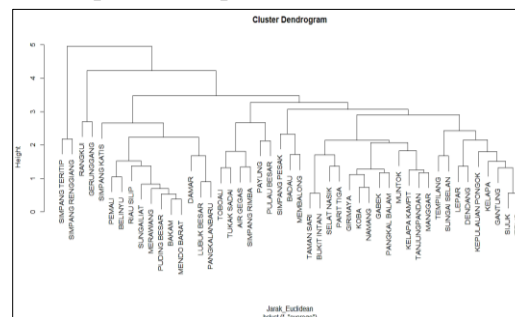
dengan *cluster* awal dalam penelitian ini menggunakan metode *linkage* yakni *single linkage*, *complete linkage*, dan *average linkage*. Ketiga metode ini digunakan dalam proses pengelompokkan dan akan ditentukan metode *linkage* yang paling optimal.

Nilai koefisien korelasi *cophenetic* yang mendekati 1 menunjukkan semakin baik *cluster* yang dihasilkan. Adapun perbandingan nilai koefisien korelasi *cophenetic* yang diperoleh untuk menentukan metode *linkage* yang paling optimal dihasilkan melalui perhitungan menggunakan *R Studio* dapat dilihat pada Tabel 6

Tabel 6 Koefisien Korelasi *Cophenetic*

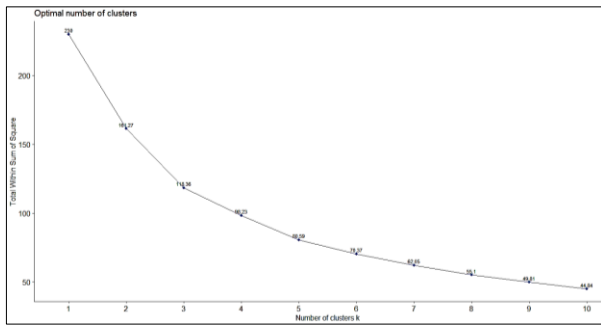
Metode <i>Linkage</i>	Koefisien Korelasi <i>Cophenetic</i>
<i>Single Linkage</i>	0,7333
<i>Complete Linkage</i>	0,6988
<i>Average Linkage</i>	0,7463

Tabel 6 menunjukkan bahwa nilai koefisien korelasi *cophenetic* yang paling tinggi dihasilkan oleh metode *average linkage* diantara metode *linkage* lainnya dengan nilai sebesar 0,7463. Oleh karena itu, penelitian ini akan menggunakan metode *Agglomerative Nesting* dengan *Average Linkage*. Adapun hasil *dendrogram* yang terbentuk dari metode tersebut dapat dilihat pada Gambar 5

Gambar 5 Hasil *Dendrogram Average Linkage*

EVALUASI CLUSTER OPTIMAL

Penentuan banyaknya jumlah *cluster* optimal dalam metode *agglomerative nesting* dengan *average linkage*, yaitu menggunakan metode *elbow* dan *silhouette coefficient*. Penentuan *cluster* optimal dapat dianalisis dengan melihat grafik dari kedua indeks validitas tersebut. Adapun hasil dari metode *elbow* dalam menentukan banyaknya jumlah *cluster* yang optimal dapat dilihat pada Gambar 6



Gambar 6 Grafik Metode Elbow

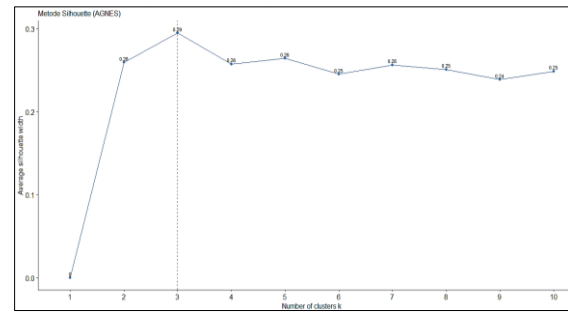
Gambar menunjukkan nilai *Sum of Square Errors* (SSE) dari metode *elbow* berdasarkan banyaknya jumlah *cluster* yang terbentuk melalui metode *agglomerative nesting*. Berdasarkan Gambar 6 dapat dilihat bahwa garis yang membentuk siku berada pada $k = 2$ dan $k = 3$. Selain melihat garis yang membentuk siku, hasil metode *elbow* juga dapat dilihat melalui penurunan nilai SSE yang signifikan. Adapun selisih nilai SSE antara dua jumlah *cluster* pada metode *agglomerative nesting* dapat dilihat pada Tabel 7.

Berdasarkan Tabel 7, penurunan nilai SSE yang sangat signifikan terjadi pada saat $k = 2$ sampai dengan $k = 3$. Oleh karena itu banyaknya jumlah *cluster* yang akan digunakan dalam analisis evaluasi *cluster* menggunakan *agglomerative nesting* sebanyak 2 sampai 3 *cluster*.

Tabel 7 Nilai SSE Evaluasi Cluster Metode Elbow

Jumlah Cluster	Nilai SSE	Selisih SSE
1	230	0
2	161,27	68,73
3	118,36	42,91
4	98,23	20,13
5	80,59	17,64
6	70,37	10,22
7	62,05	8,32
8	55,10	6,95
9	49,81	5,29
10	44,84	4,97

Adapun hasil dari *silhouette coefficient* dalam menentukan banyaknya jumlah *cluster* yang optimal dapat dilihat pada Gambar 7



Gambar 7 Grafik Silhouette Coefficient

Gambar 7 menunjukkan nilai *average silhouette width* dari *silhouette coefficient* berdasarkan banyaknya jumlah *cluster* yang terbentuk melalui metode *agglomerative nesting*. Berdasarkan Gambar dapat dilihat bahwa titik tertinggi berada pada $k = 3$. Oleh karena itu, berdasarkan evaluasi dengan metode *elbow* dan *silhouette coefficient* dapat diperoleh bahwa dalam penelitian ini, jumlah *cluster* sebanyak 3 merupakan jumlah *cluster* yang paling optimal terbentuk. Hal ini ditunjukkan oleh grafik dan nilai dari kedua indeks evaluasi *cluster* yang sesuai pada jumlah *cluster* sebanyak 3.

Adapun anggota dari masing-masing *cluster* tersebut dapat dilihat pada Tabel 8

Tabel 8 Anggota Cluster

Cluster 1	Koba, Girimaya, Gabek, Bukit Intan, Muntok, Namang, Simpang Rimba, Tukak Sadai, Simpang Katis, Merawang, Lubuk Besar, Pangkalanbaru, Puding Besar, Sungailiat, Riau Silip, Bakam, Mendo Barat, Pemali, Belinyu, Badau, Selat Nasik, Membalong, Kepulauan Pongok, Lepar, Dendang, Toboali, Air Gegas, Payung, Pulau Besar, Tempilang, Sungai Selan, Kelapa, Gantung, Sijuk, Jebus, Damar, Tanjungpandan, Kelapa Kampit, Manggar, Simpang Pesak, Pangkal Balam, Parit Tiga dan Taman Sari
Cluster 2	Simpang Renggang, Simpang Teritip
Cluster 3	Rangkui, Gerunggang

INTERPRETASI CLUSTER

Setelah *cluster* optimal terbentuk, kemudian akan dilakukan interpretasi masing-masing *cluster* untuk menjelaskan karakteristik dari setiap *cluster* yang terbentuk. Adapun karakteristik tersebut diwakili oleh pusat *cluster* (*centroid*) yang ditunjukkan oleh nilai rata-rata dari setiap variabel. Adapun pusat *cluster* tersebut dapat dilihat pada Gambar 8

```
# A tibble: 3 × 6
```

cluster	x1	x2	x3	x4	x6
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	4.26	7.65	5.12	64.4	16.1
2	15.4	17.8	10.8	49.6	5.18
3	15.6	6.73	1.52	89.6	7.10

Gambar 8 Pusat Cluster Optimal

Berdasarkan Gambar 8, dapat diketahui karakteristik dari masing-masing cluster, dimana cluster 1 yang terdiri dari 43 kecamatan memiliki karakteristik ibu hamil yang anemia (X_1), ibu hamil yang kekurangan energi kronik (X_2) dan bayi dengan berat badan lahir rendah (X_3), tergolong rendah. Selain itu, akses sanitasi yang aman (X_6) pada cluster 1 sudah cukup baik dibandingkan dengan cluster lainnya. Namun, bayi < 6 bulan yang mendapat ASI eksklusif sedikit lebih rendah jika dibandingkan dengan cluster 3.

Selain itu, cluster 2 yang terdiri dari 2 kecamatan memiliki karakteristik faktor penyebab *stunting* yang cukup buruk jika dibandingkan dengan cluster lainnya. Sedangkan cluster 3 yang juga terdiri dari 2 kecamatan sudah memiliki karakteristik faktor penyebab *stunting* yang cukup baik. Namun pada cluster 3 masih perlu diperhatikan pada indikator ibu hamil yang anemia (X_1) dan akses sanitasi yang aman (X_6).

UCAPAN TERIMA KASIH

Terima kasih kepada Ibu Ririn Amelia, S.T., M.Si., selaku Dosen Pembimbing Utama, dan Ibu Izma Fahria S.E., M.Sc. selaku Dosen Pembimbing Pendamping, atas bimbingan dan arahan yang sangat membantu dalam penelitian ini. Terima kasih juga kepada Dinas Kesehatan Kabupaten/Kota di Provinsi Kepulauan Bangka Belitung atas izin yang diberikan untuk memperoleh data yang diperlukan, sehingga proses penelitian dapat dilaksanakan secara efektif dan efisien.

PENUTUP

SIMPULAN

Berdasarkan hasil dan pembahasan sebelumnya, dapat disimpulkan bahwa analisis cluster kecamatan berdasarkan faktor penyebab *stunting* di Provinsi Kepulauan Bangka Belitung menggunakan metode *Agglomerative Nesting* (AGNES) menghasilkan jumlah cluster optimal sebanyak 3 cluster. Selain itu, metode *linkage* yang paling optimal dalam analisis cluster dalam penelitian ini yakni *average linkage*

dengan koefisien korelasi *cophenetic* sebesar 0,7463. Berdasarkan analisis juga dapat diketahui karakteristik dari masing-masing cluster. Adapun cluster 1 memiliki karakteristik faktor penyebab *stunting* yang sudah cukup baik, namun masih perlu peningkatan pada indikator bayi < 6 bulan yang mendapat ASI Eksklusif. Sedangkan pada cluster 2, karakteristik faktor penyebab *stunting* cukup buruk pada hampir semua indikator jika dibandingkan dengan cluster lainnya. Pada cluster 3, karakteristik faktor penyebab *stunting* sudah cukup baik. Namun tetap diperlukan perhatian pada indikator ibu hamil yang anemia serta akses sanitasi yang aman.

SARAN

Untuk peneliti selanjutnya, disarankan untuk memperluas cakupan analisis cluster dengan menerapkannya pada konteks yang berbeda. Selain itu, peneliti selanjutnya dapat menambahkan ataupun mengombinasikan variabel yang lebih beragam untuk menghasilkan pemahaman yang lebih mendalam terkait fenomena yang diteliti.

DAFTAR PUSTAKA

- Ali, M. M., Hariyati, T., Pratiwi, M. Y., & Afifah, S. (2022). Metodologi Penelitian Kuantitatif Dan Penerapan Nya Dalam Penelitian. *JPIB: Jurnal Penelitian Ibnu Rusyd*, 2(2), 1-5.
- Bachtiar, A. M., Dharmayanti, D., & Hamzah, R. L. (2017). Penerapan Metode Hierarchical Agglomerative Clustering Untuk Segmentasi Pelanggan Potensial Di Jeger Jersey Indonesia. *Komputa: Jurnal Ilmiah Komputer Dan Informatika*, 6(1), 35-42.
- Badan Kebijakan Pembangunan Kesehatan. (2024). Laporan Tematik Survei Kesehatan Indonesia (SKI) Tahun 2023: Potret Indonesia Sehat. In *Kementerian Kesehatan RI. Kementerian Kesehatan RI*.
- Butar, R. P. B. (2023). Analisis Hierarchical Dan Non-Hierarchical Clustering Untuk Pengelompokan Potensi Ekonomi Kelautan Indonesia 2021. *Jurnal Sistem Dan Teknologi Informasi (JustIN)*, 11(3), 543-553.

- Daniswara, A. A. A., & Nuryana, I. K. D. (2023). Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru. *JINACS: Journal of Informatics and Computer Science*, 5(1), 97-100.
- Farissa, R. A., Mayasari, R., & Umaidah, Y. (2021). Perbandingan Algoritma K-Means dan K-Medoids Untuk Pengelompokan Data Obat dengan Silhouette Coefficient di Puskesmas Karangsambung. *Journal of Applied Informatics and Computing*, 5(2), 109-116.
- Faujia, R. A., Setianingsih, E. S., & Pratiwi, H. (2022). Analisis Klaster K-Means Dan Agglomerative Nesting Pada Indikator Stunting Balita Di Indonesia. *Seminar Nasional Official Statistics*, 2022(1), 1249-1258.
- Februariyanti, H., & Winarko, E. (2010). Klastering Dokumen Menggunakan Hierarchical Agglomerative Clustering. *Seminar Nasional STIKOM Surabaya*.
- Ginantra, N. L. W. S. R., Arifah, F. N., Wijaya, A. H., Septarini, R. S., Ahmad, N., Ardiana, D. P. Y., Effendy, F., Iskandar, A., Hazriani, H., Sari, I. Y., Gustiana, Z., Prianto, C., Gustian, D., & Negara, E. S. (2021). *Data Mining dan Penerapan Algoritma*. Yayasan Kita Menulis.
- Hasanah, I. N., & Sofro, A. (2022). Analisis Cluster Berdasarkan Dampak Ekonomi Di Indonesia Akibat Pandemi Covid-19. *MATHunesa: Jurnal Ilmiah Matematika*, 10(2), 239-248.
- Hasrul, M. (2018). *Analisis Klaster untuk Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Indikator Kesejahteraan Rakyat* [UIN Alauddin]. <http://repositori.uin-alauddin.ac.id/13432/1/MUH.HASRUL.pdf>
- Hutabarat, E. N. (2023). Permasalahan Stunting dan Pencegahannya. *Jornal of Health and Medical Science*, 2(1), 158-163.
- Iis, Yahya, I., Wibawa, G. N. A., Baharuddin, Ruslan, & Laome, L. (2022). Penggunaan Korelasi Cophenetic Untuk Pemilihan Metode Cluster Berhierarki Pada Mengelompokkan Kabupaten/Kota Berdasarkan Jenis Penyakit Di Provinsi Sulawesi Tenggara Tahun 2020. *Prosiding Seminar Nasional Sains Dan Terapan (Sinta)*, VI(April), 1-16.
- Izzuddin, K. H., & Wijayanto, A. W. (2024). Pemodelan Clustering Ward, K-Means, DIANA, dan PAM dengan PCA untuk Karakterisasi Kemiskinan Indonesia Tahun 2021. *Jurnal Sistem Komputer*, 13(1), 41-53.
- Kemenkes. (2023). *Upaya Pencegahan Bayi Lahir Prematur*. MINISTRY OF HEALTH RELEASE. <https://www.kemkes.go.id/eng/rilis-kesehatan/upaya-pencegahan-bayi-lahir-prematur>
- Kemenkes RI. (2021). Pedoman Gizi Seimbang Ibu Hamil dan Ibu Menyusui. In *Kementerian Kesehatan Republik Indonesia*.
- Kuewa, Y., Herawati, H., Sattu, M., Otoluwa, A. S., Lalusu, E. Y., & Dwicahya, B. (2021). The relationship between environmental sanitation and the incidence of stunting in toddlers in Jayabakti village in 2021. *Public Health J*, 12(2), 112-118.
- Meliyana, E. (2024). Asi Eksklusif, MP Asi dan Stunting. In *BOOK CHAPTER STUNTING* (pp. 57-78).
- Merliana, N. P. E., Ernawati, & Santoso, A. J. (2015). Analisa Penentuan Jumlah Cluster Terbaik Pada Metode K-Means Clustering. *Prosiding Seminar Nasional Multi Disiplin Ilmu & Call For Papers Unisbank (SENDI_U)*, 978-979.
- Musfiani, M. (2019). Analisis Cluster Dengan Menggunakan Metode Partisi Pada Pengguna Alat Kontrasepsi Di Kalimantan Barat. *Bimaster: Buletin Ilmiah Matematika, Statistika Dan Terapannya*, 8(4), 893-902.
- Nahdliyah, M. A., Widiharih, T., & Prahutama, A. (2019). Metode K-Medoids Clustering Dengan Validasi Silhouette Index dan C-Index (Studi Kasus Jumlah Kriminalitas

- Kabupaten/Kota di Jawa Tengah Tahun 2018). *Jurnal Gaussian*, 8(2), 161–170.
- Permana, A. A., S. W., Santoso, L. W., Wibowo, G. W. N., Wardhani, A. K., Rahmadden, Wahidin, A. J., Yuliasuti, G. E., Elisawati, Wijayanti, R. R., & Abdurrasyid. (2023). Machine Learning. In *Machine Learning* (Vol. 45, Issue 13).
- Perpres. (2021). *Peraturan Presiden Nomor 72 Tahun 2021*. 1.
- Perpres RI. (2020). Rencana Pembangunan Jangka Menengah Nasional. *Peraturan Presiden Republik Indonesia Nomor 18 Tahun 2020*, 2020–2024.
- Pratiwi, I. G. (2023). Studi Literatur: Intervensi Spesifik Penanganan Stunting. *Indonesian Health Issue*, 2(1), 29–37.
- Purnomo, Sutadji, E., Utomo, W., Purnawirawan, O., Farich, R., A.S., S., M., R. F., Carina, A., & R., N. G. (2022). *Analisis Data Multivariat*. Omera Pustaka.
- Puspitasari, M. W., & Susanti, M. (2016). Pengelompokan Kabupaten/Kota Berdasarkan Faktor-Faktor yang Mempengaruhi Kemiskinan di Jawa Tengah Menggunakan Metode Ward dan Average Linkage. *Jurnal Pendidikan Matematika Dan Sains*, 5(6), 1–10.
- Rusliani, N., Hidayani, W. R., & Sulistyoningsih, H. (2022). Literature Review: Faktor-Faktor yang Berhubungan dengan Kejadian Stunting pada Balita. *Buletin Ilmu Kebidanan Dan Keperawatan*, 1(01), 32–40.
- Samsuddin, Festilia, S., Agusanty, Desmawati, Kurniatin, L. F., Bahriyah, F., Wati, I., Ulva, S. M., Abselian, U. P., Laili, U., Firdayana, M. M., Purwadi, H. N., & Ernawati, Y. (2023). *STUNTING*. EUREKA MEDIA AKSARA.
- Simanjuntak, K. P., & Khaira, U. (2021). Hotspot Clustering in Jambi Province Using Agglomerative Hierarchical Clustering Algorithm. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1(1), 7–16.
- Siregar, Y. I. H. (2024). *Analisis Algoritma K-medoids Untuk Pengelompokan Hasil Panen Sawit Berdasarkan Berat Dan Kualitas (Study Kasus: CV. Ram Bintang Motor)* [Universitas Medan Area]. <https://repositori.uma.ac.id/jspui/bitstream/123456789/24101/1/198160028> - Yusril Izza H Siregar - Fulltext.pdf
- Suryati, Ulfa, S. M., & Wulandatika, D. (2022). Pendampingan Ibu Hamil Dalam Pencegahan Anemia Defisiensi Besi (ADB) Dan KEK Di Kelas Prenatal. *AMMA : Jurnal Pengabdian Masyarakat*, 1(07), 818–822.
- Syahfitri, N., Budianita, E., Nazir, A., & Afrianty, I. (2023). Pengelompokan Produk Berdasarkan Data Persediaan Barang Menggunakan Metode Elbow dan K-Medoid. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 4(3), 1668–1675.
- Ulya, F. A., Abdullah, A. N., Hanan, T. A., & Nur, I. M. (2023). Pengelompokan Tingkat Pengangguran Terbuka Di Jawa Tengah Menggunakan Metode K - Means Clustering. *Journal Of Data Insights*, 1(2), 71–80.
- WHO. (2015). *Stunting in a nutshell*. World Health Organization. <https://www.who.int/news/item/19-11-2015-stunting-in-a-nutshell>