

CLUSTERING DAERAH RAWAN STUNTING DI INDONESIA MENGGUNAKAN ALGORITMA K-MEANS

Cahya Syalom¹, Ulfa Khaira², Daniel Arsa³

1,2,3) Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Jambi, Indonesia

Article Info

Article history:

Received: 28 Maret 2025

Revised: 27 April 2025

Accepted: 28 April 2025

ABSTRACT

Abstrak

Stunting merupakan salah satu masalah gizi utama di Indonesia yang prevalensinya masih tinggi, yaitu 21,5% pada tahun 2023, melebihi standar WHO yang seharusnya dibawah 20%. Penelitian ini bertujuan untuk mengidentifikasi daerah yang berisiko rawan stunting di Indonesia menggunakan metode K-Means berdasarkan faktor-faktor penyebab stunting. Data yang digunakan meliputi cakupan pelayanan esensial dan kasus stunting dari 38 provinsi selama 2022-2024, diperoleh dari *dashboard* Ditjen Bina Bangda Kemendagri sebanyak 3534 data. Penelitian diawali dengan menentukan variabel yang memiliki korelasi lebih kuat terhadap stunting diantara variabel lainnya pada masing-masing tahun menggunakan korelasi *Spearman*. Lalu, jumlah *cluster* optimal ditentukan menggunakan metode *Silhouette Coefficient*. Hasil menunjukkan bahwa jumlah *cluster* optimal untuk tiap tahun adalah K=2 dengan masing-masing nilainya sebesar 0,387 (2022), 0,436 (2023), dan 0,446 (2024). *Cluster* 1 mewakili daerah risiko rawan stunting tinggi dengan cakupan pelayanan esensial yang rendah, sedangkan *cluster* 2 mewakili daerah risiko rawan stunting rendah dengan pelayanan esensial yang lebih optimal. Pada 2022, *cluster* 1 terdiri dari 25 provinsi, dan *cluster* 2 terdiri dari 13 provinsi. Pada 2023, *cluster* 1 menurun menjadi 18 provinsi dan *cluster* 2 meningkat menjadi 20 provinsi. Pada 2024, *cluster* 1 semakin menurun menjadi 14 provinsi, sementara *cluster* 2 meningkat menjadi 24 provinsi.

Kata Kunci: Stunting; K-Means; Klasterisasi; *Silhouette Coefficient*.

Abstract

Stunting is one of the main nutritional problems in Indonesia whose prevalence is still high, which is 21.5% in 2023, exceeding the WHO standard which should be below 20%. This study aims to identify areas at risk of stunting in Indonesia using the K-Means method based on factors that cause stunting. The data used includes coverage of essential services and stunting cases from 38 provinces during 2022-2024, obtained from the Ditjen Bina Bangda Kemendagri dashboard as much as 3534 data. The research begins by determining the variables that have a stronger correlation with stunting among other variables in each year using Spearman correlation. Then, the optimal number of clusters was determined using the Silhouette Coefficient method. The results show that the optimal number of clusters for each year is K=2 with each value being 0.387 (2022), 0.436 (2023), and 0.446 (2024). Cluster 1 represents high stunting risk areas with low coverage of essential services, while cluster 2 represents low stunting risk areas with more optimized essential services. In 2022, cluster 1 consists of 25 provinces, and cluster 2 consists of 13 provinces. In 2023, cluster 1 decreases to 18 provinces and cluster 2 increases to 20 provinces. In 2024, cluster 1 declines further to 14 provinces, while cluster 2 increases to 24 provinces.

Keywords: Stunting; K-Means; Clustering; *Silhouette Coefficient*.

Djtechno: Jurnal Teknologi Informasi oleh Universitas Dharmawangsa Artikel ini bersifat open access yang didistribusikan di bawah syarat dan ketentuan dengan Lisensi Internasional Creative Commons Attribution NonCommercial ShareAlike 4.0 ([CC-BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/)).



1. PENDAHULUAN

Stunting merupakan salah satu masalah gizi utama di Indonesia yang diakibatkan oleh asupan gizi yang buruk dan infeksi kronis berulang, yang ditandai dengan tinggi badan ≤ -2 standar deviasi dibawah standar ketetapan pertumbuhan anak oleh *World Health Organization* (WHO) [1]. Stunting berdampak pada pertumbuhan fisik, perkembangan kognitif dan motorik, ukuran fisik tubuh serta peningkatan risiko penyakit kronis yang berhubungan dengan gizi di masa dewasa [2]. Kementerian Kesehatan mengumumkan prevalensi stunting di Indonesia pada tahun 2023 masih tergolong tinggi yaitu mencapai 21,5% [3]. Angka ini memerlukan perhatian serius dari pemerintah karena melebihi standar WHO, yang seharusnya dibawah 20% [4]. Untuk mendukung upaya pemerintah dalam penurunan stunting, penelitian ini akan menerapkan teknik *clustering* untuk mengidentifikasi daerah yang berisiko rawan stunting di Indonesia berdasarkan beberapa faktor penyebab stunting. Data yang digunakan bersumber dari website resmi Bangda Kemendagri, mencakup 29 variabel yang berpengaruh terhadap angka stunting di Indonesia

Clustering merupakan teknik yang digunakan untuk mengelompokkan sejumlah objek ke dalam beberapa *cluster* berdasarkan kemiripan pola yang serupa di *cluster* yang sama dan pola objek yang tidak mirip di *cluster* lainnya [5]. Metode *clustering* terdiri dari *Hierarchical Clustering* dan *Non-Hierarchical Clustering*. *Hierarchical Clustering* mengelompokkan objek yang memiliki kemiripan membentuk struktur hirarki berbentuk pohon atau *dendrogram* [6]. Algoritma *Hierarchical Clustering* terdiri dari *Agglomerative Hierarchy Clustering* dan *Divisive Hierarchy Clustering*. Sedangkan *Non-Hierarchical Clustering* mengelompokkan objek ke beberapa *cluster* secara langsung tanpa membentuk diagram pohon atau *dendrogram* [6]. Algoritma *Non-Hierarchical Clustering* terdiri dari K-Means, Fuzzy C-Means, dan K-Medoid.

Berdasarkan hasil penelitian oleh [6], metode *K-Means* terbukti lebih optimal dibandingkan metode *Fuzzy C-Means* dan *Hierarchical Clustering* dalam mengelompokkan provinsi di Indonesia berdasarkan indikator *women empowerment*. Metode *K-Means* memiliki nilai tertinggi pada *Silhouette Index* dan *Calinski-Harabasz Index*, menunjukkan bahwa metode ini paling efektif dalam membentuk *cluster*.

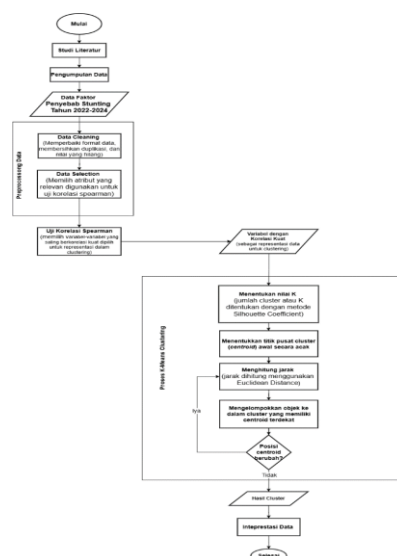
Penelitian lain oleh [7] mengelompokkan daerah rawan stunting di Kabupaten Asahan berdasarkan jumlah kasus stunting. Hasilnya menunjukkan bahwa jumlah *cluster* terbaik adalah dua *cluster*, yang ditentukan dengan metode *elbow*. Evaluasi kualitas *cluster* menghasilkan nilai Davies-Bouldin Index (DBI) sebesar 0,5129 dan skor *silhouette* sebesar 0,7143, yang menunjukkan bahwa kualitas *cluster* baik. Berdasarkan penelitian terdahulu, metode K-Means dapat diterapkan pada pengelompokan daerah berdasarkan indikator tertentu. Oleh karena itu, penelitian ini akan menggunakan algoritma K-Means untuk mengelompokkan daerah risiko rawan stunting di Indonesia.

K-Means adalah salah satu metode *clustering* non-hirarki untuk mengelompokkan data ke dalam satu atau lebih *cluster*, dimana data yang memiliki karakteristik yang sama dikelompokkan dalam *cluster* yang sama, dan data yang memiliki karakteristik yang berbeda dikelompokkan ke dalam *cluster* lain [8]. Kelebihan dari algoritma K-Means yaitu kesederhanaan, relatif cepat dalam proses komputasi, mudah untuk adaptasi, dan mudah diimplementasikan dibandingkan dengan algoritma *clustering* lainnya [9].

Dari latar belakang yang telah diuraikan tersebut, peneliti merasa perlu untuk melakukan pengelompokan daerah risiko rawan stunting di Indonesia berdasarkan faktor-faktor penyebab stunting dengan menggunakan K-Means. Hasil *clustering* akan divisualisasikan dalam bentuk peta di RStudio.

2. METODE PENELITIAN

2.1. Tahapan Penelitian



Gambar 1. Tahapan Penelitian

2.2. Pengumpulan Data

Data penelitian diperoleh dari *dashboard* Ditjen Bina Bangda Kemendagri dengan teknik *web scrapping*. *Web scrapping* adalah teknik mengekstrak data dari halaman website secara otomatis [10]. Proses ini menggunakan ekstensi Web Scraper. *Dataset* yang digunakan meliputi variabel cakupan layanan esensial dalam penurunan stunting dan kasus stunting pada tahun 2022-2024 sebanyak 3534 data.

Tabel 1. Variabel Penelitian

No.	Variabel	Kode Variabel	Tipe Data
1.	Provinsi	PROV	Kategorikal
2.	Remaja putri mengonsumsi Tablet Tambah Darah (TTD)	RETTD	Nurmerik
3.	Remaja putri menerima layanan pemeriksaan status anemia.	REANE	Numerik
4.	Calon pengantin atau calon ibu menerima TTD	IBTTD	Numerik
5.	Calon pasangan usia subur (PUS) memperoleh pemeriksaan kesehatan sebagai bagian dari pelayanan nikah	PUSPK	Numerik
6.	Cakupan calon PUS yang menerima pendampingan kesehatan reproduksi dan edukasi gizi sejak 3 bulan pranikah	PUSPKG	Numerik
7.	Pasangan calon pengantin yang mendapatkan bimbingan perkawinan dengan materi pencegahan stunting	CPMBPS	Numerik
8.	PUS dengan status miskin dan penyandang masalah kesejahteraan sosial yang menerima bantuan tunai bersyarat	PUSBTS	Numerik
9.	Cakupan PUS dengan status miskin dan penyandang masalah kesejahteraan sosial yang menerima bantuan pangan nontunai	PUSBPN	Numerik
10.	PUS fakir miskin dan orang tidak mampu yang menjadi Penerima Bantuan Iuran (PBI) Jaminan Kesehatan	PUSPBI	Numerik
11.	Ibu hamil Kurang Energi Kronik mendapat tambahan asupan gizi	IHKEK	Nurmerik
12.	Ibu hamil yang mengonsumsi TTD minimal 90 tablet	IHTTD	Numerik
13.	Persentase Unmeet Need pelayanan keluarga berencana	UNKB	Numerik
14.	Persentase kehamilan yang tidak diinginkan	KTD	Numerik
15.	Bayi usia kurang dari 6 bulan mendapat ASI eksklusif	ASIE	Numerik
16.	Anak usia 6-23 bulan yang mendapat MP-ASI	MPASI	Numerik
17.	Anak berusia di bawah lima tahun (balita) gizi buruk yang mendapat pelayanan tata laksana gizi buruk	BGBPEL	Numerik
18.	Anak berusia di bawah lima tahun (balita) yang dipantau pertumbuhan dan perkembangannya	BDP	Numerik
19.	Anak berusia di bawah lima tahun (balita) gizi kurang yang mendapat tambahan asupan gizi	BGKTAG	Numerik
20.	Balita yang memperoleh imunisasi dasar lengkap	IMDL	Nurmerik
21.	Keluarga yang Stop BABS	SBABS	Numerik
22.	Keluarga yang melaksanakan PHBS	PHBS	Numerik
23.	Keluarga berisiko stunting yang mendapatkan promosi peningkatan konsumsi ikan dalam negeri	KBSKI	Numerik
24.	Pelayanan Keluarga Berencana (KB) pasca persalinan	PKBPP	Numerik
25.	Keluarga berisiko stunting yang memperoleh pendampingan	KBSPE	Numerik
26.	Keluarga berisiko stunting yang mendapatkan manfaat sumber daya pekarangan untuk peningkatan asupan gizi	KBSSP	Numerik
27.	Rumah tangga yang mendapatkan akses air minum layak	AML	Numerik
28.	Rumah tangga yang mendapatkan akses sanitasi layak	ASL	Numerik
29.	Kelompok Keluarga Penerima Manfaat (KPM) Program Keluarga Harapan (PKH) yang mengikuti Pertemuan Peningkatan Kemampuan Keluarga (P2K2) dengan modul kesehatan dan gizi	P2K2	Numerik

30.	Keluarga Penerima Manfaat (KPM) dengan ibu hamil, ibu menyusui, dan baduta yang menerima variasi bantuan pangan selain beras dan telur	BAPA	Numerik
31.	Kasus Stunting	KS	Numerik

2.3. Preprocessing Data

Setelah tahap pengumpulan data, maka dapat dilakukan tahap *preprocessing data* agar sesuai dengan bentuk format yang diinginkan untuk tahap analisis lebih lanjut. Berikut beberapa tahapan yang dilakukan dalam *preprocessing data*:

- Pembersihan Data (*Data Cleaning*): proses memeriksa data dan membersihkan data dari inkonsistensi, duplikasi, nilai yang hilang dan kesalahan input atau penulisan pada data [11].
- Seleksi Data (*Data Selection*): proses pemilihan data yang relevan untuk digunakan dalam proses *data mining* [11]. Variabel yang dipilih adalah data yang bertipe numerik, yaitu seluruh data dalam kode variabel dari RETTD-KS, kecuali PROV yang bertipe kategorikal.

2.4. Korelasi Spearman

Korelasi *spearman* adalah ukuran nonparamterik yang digunakan untuk menggambarkan hubungan dua variabel tanpa membuat asumsi distribusi frekuensi terhadap variabel penelitian [12]. Uji korelasi *spearman* digunakan untuk memilih fitur atau variabel yang akan digunakan dalam proses *clustering*. Mengingat jumlah variabel yang tersedia sebanyak 29, analisis korelasi ini bertujuan untuk mengidentifikasi variabel yang memiliki hubungan paling kuat dengan kasus stunting pada masing-masing tahun. Adapun Interpretasi untuk nilai korelasi koefisien (r) yang menunjukkan tingkat hubungan antara variabel-variabel adalah sebagai berikut:

Tabel 2. Interpretasi Korelasi Koefisien

Interval Koefisien	Tingkat Hubungan
0,00-0,199	Sangat Lemah
0,20-0,399	Lemah
0,40-0,599	Cukup Kuat/Sedang
0,60-0,799	Kuat
0,80-1,000	Sangat Kuat

Sumber Tabel: [13]

2.5. K-Means Clustering

K-Means adalah salah satu metode *clustering* non-hirarki untuk mengelompokkan data ke dalam satu atau lebih *cluster*, dimana data yang memiliki karakteristik yang sama dikelompokkan dalam *cluster* yang sama, dan data yang memiliki karakteristik yang berbeda dikelompokkan ke dalam *cluster* lain [8]. Algoritma K-Means memiliki

kelebihan seperti mudah diimplementasikan, relatif cepat dalam proses komputasi, dan mudah untuk beradaptasi dibandingkan dengan algoritma *clustering* lainnya [9]. Selain itu, algoritma ini cukup akurat dalam menangani data dalam berbagai ukuran objek, sehingga relatif lebih terukur dan efisien untuk mengelola data dalam jumlah besar [14]. Secara umumnya, metode K-Means memiliki algoritma sebagai berikut [8]:

1. Menentukan jumlah *cluster* (k) dengan metode *Silhouette Coefficient*. *Silhouette Coefficient* adalah metode gabungan dari metode separasi dan kohesi yang digunakan untuk mengevaluasi kualitas dari *cluster* yang terbentuk [15]. Apabila nilai *Silhouette Coefficient* semakin mendekati 1, maka pengelompokan *cluster* semakin baik, begitu sebaliknya.
2. Menentukan titik pusat *cluster* (*centroid*) awal secara acak. Sedangkan pada tahap iterasi selanjutnya, digunakan rumus untuk menghitung nilai *centroid*:

$$v_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} x_{kj}$$

Keterangan:

v_{ij} = *centroid* atau *mean cluster* ke- i untuk variabel ke- j

N_i = jumlah data yang menjadi anggota *cluster* ke- i

i, k = indeks dari *cluster*

j = indeks dari variabel x

kj = nilai data ke- k yang ada dalam *cluster* tersebut untuk variabel ke- j

3. Menghitung jarak setiap objek ke *centroid* menggunakan *Euclidean Distance*.

$$d_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

Keterangan:

d_e = *Euclidean distance*

i = banyaknya objek

(x,y) = koordinat dari objek

(s,t) = koordinat dari *centroid*

4. Mengelompokkan objek ke dalam *cluster* yang memiliki *centroid* terdekat.
5. Ulangi langkah ketiga hingga kelima untuk posisi *centroid* baru yang tidak sama.

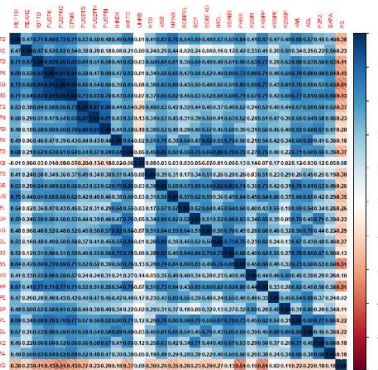
3. HASIL DAN PEMBAHASAN

Setelah melalui, proses *cleansing* data dan seleksi data, data numerik dipilih untuk digunakan dalam pengujian korelasi Spearman, kecuali PROV yang bertipe kategorikal.

Gambar 2. Dataset Penelitian

Sebelum masuk ke dalam proses *K-Means Clustering*, dilakukan seleksi fitur atau variabel dengan menggunakan uji korelasi *Spearman* pada data tahun 2022-2024.

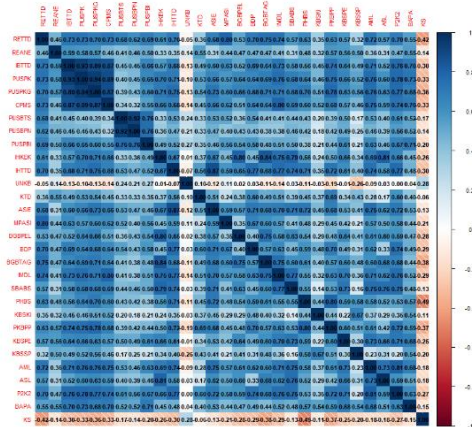
1. Uji korelasi Spearman untuk data tahun 2022



Gambar 3. Uji Korelasi Spearman pada Data Tahun 2022

Variabel yang memiliki nilai korelasi koefisien dengan rentang nilai $\geq 0,35$ terhadap kasus stunting (KS) pada data tahun 2022 berjumlah 10 variabel, yaitu RETTD, IBTTD, PUSPK, PUSPKG, CPMBPS, PUSBTS, IHTTD, MPASI, PHBS, dan PKBPP

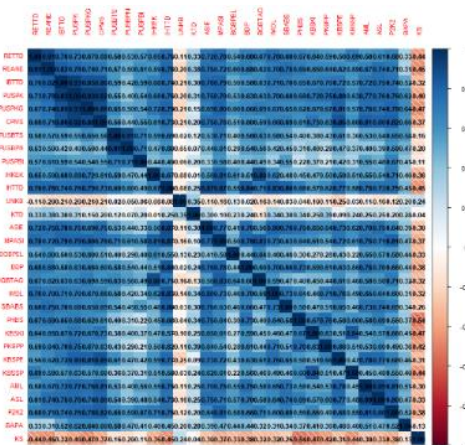
2. Uji korelasi Spearman untuk data tahun 2023



Gambar 4. Uji Korelasi Spearman pada Data Tahun 2023

Variabel yang memiliki nilai korelasi koefisien dengan rentang nilai $\geq 0,30$ terhadap kasus stunting (KS) pada data tahun 2023 berjumlah 9 variabel, yaitu RETTD, IBTTD, PUSPK, PUSPKG, CPMBPS, IHTTD, BGKTAG, PHBS, dan PKBPP.

3. Uji korelasi Spearman untuk data tahun 2024

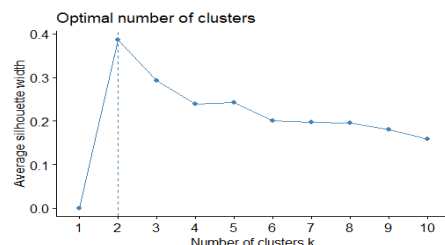


Gambar 5. Uji Korelasi *Spearman* pada Data Tahun 2024

Variabel yang memiliki nilai korelasi koefisien dengan rentang nilai $\geq 0,35$ terhadap kasus stunting (KS) pada data tahun 2024 berjumlah 14 variabel, yaitu RETTD, REANE, PUSPK, PUSPKG, CPMBPS, IHKEK, IHTTD, MPASI, BDP, PHBS, KBSKI, PKBPP, KBSSP, dan P2K2.

Setelah diperoleh variabel yang akan digunakan dalam proses *K-Means Clustering* pada masing-masing tahun, Selanjutnya, menentukan jumlah *cluster* (k) optimal menggunakan *Silhouette Coefficient*.

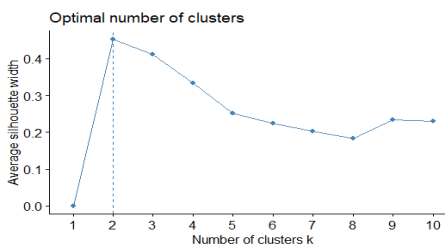
1. *Silhouette Coefficient* pada data tahun 2022



Gambar 6. *Silhouette Coefficient* pada Data Tahun 2022

Berdasarkan Gambar 6, Nilai K yang optimal untuk data tahun 2022 adalah 2 *cluster* dikarenakan nilai *Silhouette Coefficient* nya yang tertinggi dan paling mendekati angka 1. Nilai *Silhouette Coefficient* K=2 sebesar 0,387.

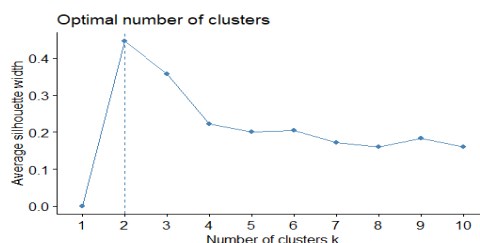
2. *Silhouette Coefficient* pada data faktor penyebab stunting tahun 2023



Gambar 7. *Silhouette Coefficient* pada Data Tahun 2023

Berdasarkan gambar 7, nilai K yang optimal untuk data tahun 2023 adalah 2 *cluster* dikarenakan nilai *Silhouette Coefficient* nya yang tertinggi dan paling mendekati angka 1. Nilai *Silhouette Coefficient* K=2 sebesar 0,436.

3. *Silhouette Coefficient* pada data faktor penyebab stunting tahun 2024



Gambar 8. *Silhouette Coefficient* pada Data Tahun 2024

Berdasarkan Gambar 8, nilai K yang optimal untuk data tahun 2024 adalah 2 *cluster* dikarenakan nilai *Silhouette Coefficient* nya yang tertinggi dan paling mendekati angka 1. Nilai *Silhouette Coefficient* K=2 sebesar 0,446.

Dengan begitu, nilai K optimal untuk tiap tahunnya adalah K=2 penelitian ini akan menghasilkan dua *cluster* berdasarkan hasil penentuan jumlah K optimal yang telah dilakukan pada data tahun 2022-2024. Berikut ini merupakan hasil *clustering* menggunakan algoritma K-Means pada data tahun 2022-2024:

1. Hasil *clustering* data tahun 2022

Tabel 3. Hasil *Cluster* K-Means K=2 Pada Data Tahun 2024

<i>Cluster</i>	<i>Label</i>	<i>Jumlah Anggota</i>	<i>Anggota Cluster</i>
1	Risiko Rawan Stunting Tinggi	25	Aceh, Sumatera Utara, Riau, Jawa Barat, Daerah Istimewa Yogyakarta, Banten, Bali, Nusa Tenggara Timur Kalimantan Barat, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, Papua, Papua Pegunungan, Papua Selatan, Papua Tengah, Papua Barat Daya.
2	Risiko Rawan Stunting Rendah	13	Sumatera Barat, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, Jawa Tengah, Jawa Timur, Nusa Tenggara Barat, Kalimantan Selatan, Sulawesi Tengah.

2. Hasil *clustering* data tahun 2023

Tabel 4. Hasil *Cluster* K-Means K=2 Pada Data Tahun 2023

<i>Cluster</i>	<i>Label</i>	<i>Jumlah Anggota</i>	<i>Anggota Cluster</i>
----------------	--------------	-----------------------	------------------------

1	Risiko Rawan Stunting Tinggi	18	Aceh, Bali, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tenggara, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, Papua, Papua Pengunungan, Papua Selatan, Papua Tengah, Papua Barat Daya.
2	Risiko Rawan Stunting Rendah	20	Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, Banten, Nusa Tenggara Barat, Kalimantan Selatan, Sulawesi Tengah, Sulawesi Selatan, Gorontalo.

3. Hasil *clustering* data tahun 2024

Tabel 5. Hasil *Cluster* K-Means K=2 Pada Data Tahun 2024

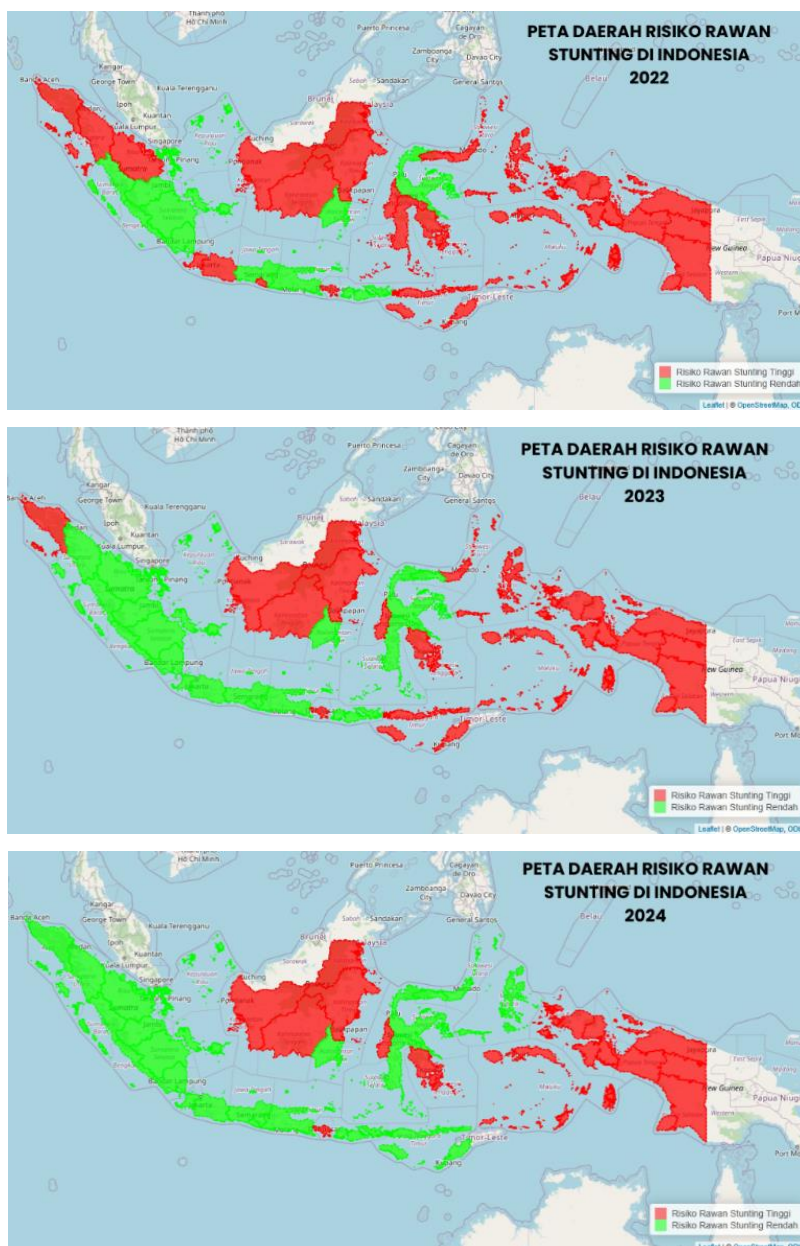
<i>Cluster</i>	Label	Jumlah Anggota	Anggota <i>Cluster</i>
1	Risiko Rawan Stunting Tinggi	14	Bali, Kalimantan Barat, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Sulawesi Tenggara, Sulawesi Barat, Maluku, Papua Barat, Papua, Papua Pengunungan, Papua Selatan, Papua Tengah, Papua Barat Daya.
2	Risiko Rawan Stunting Rendah	24	Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, Jawa Barat, Jawa Tengah, Daerah Istimewa Yogyakarta, Jawa Timur, Banten, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Selatan, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Gorontalo, Maluku Utara.

Berikut ini adalah interpretasi hasil *clustering* dari data tahun 2022-2024 ini secara keseluruhan:

- Cluster* 1 memiliki nilai pusat *cluster* yang lebih rendah untuk seluruh indikator pelayanan dibandingkan dengan *cluster* 2. *Cluster* 1 mewakili daerah dengan pelayanan esensial untuk penurunan stunting yang masih rendah, sehingga memiliki risiko tingkat kerawanan stunting yang tinggi. Sebaliknya, *cluster* 2 memiliki pelayanan esensial yang lebih optimal, mewakili daerah dengan risiko stunting yang rendah.
- Pada 2022, *cluster* 1 terdiri dari 25 provinsi, dan *cluster* 2 terdiri dari 13 provinsi. Pada 2023, *cluster* 1 menurun menjadi 18 provinsi dan *cluster* 2 meningkat menjadi 20 provinsi. Pada 2024, *cluster* 1 semakin menurun menjadi 14 provinsi, sementara *cluster* 2 meningkat menjadi 24 provinsi.

Pada Gambar 9, peta hasil *clustering* tahun 2022-2024 menunjukkan risiko tingkat kerawanan stunting menggunakan dua warna. Warna merah mewakili daerah dengan

risiko rawan stunting tinggi, dan warna hijau mewakili daerah dengan risiko rawan stunting rendah.



Gambar 9. Peta Daerah Risiko Rawan Stunting Tahun 2022-2024

4. SIMPULAN

Berdasarkan hasil penelitian yang dilakukan oleh penulis, maka dapat disimpulkan sebagai berikut:

1. Metode K-Means dapat diterapkan pada data cakupan pelayanan esensial penurunan stunting untuk mengidentifikasi daerah yang berisiko rawan stunting di Indonesia.

2. Jumlah *cluster* optimal untuk mengelompokkan daerah berisiko rawan stunting berdasarkan variabel pelayanan esensial dari tahun 2022-2024 dengan algoritma K-Means adalah $K=2$ berdasarkan hasil dari *Silhouette Coefficient*.
3. Karakteristik hasil *clustering* K-Means $K=2$ pada data tahun 2022-2024, yaitu *cluster* 1 mewakili daerah dengan risiko rawan stunting tinggi dengan pelayanan esensial penurunan stunting yang masih rendah. Sebaliknya, *cluster* 2 mewakili daerah dengan risiko rawan stunting rendah yang memiliki pelayanan esensial yang lebih optimal dibandingkan *cluster* 1.
4. Di setiap tahunnya, terjadi perubahan tingkat kerawanan stunting di beberapa provinsi dari tahun 2022 hingga 2024. Pada 2022, *cluster* 1 terdiri dari 25 provinsi, dan *cluster* 2 terdiri dari 13 provinsi. Pada 2023, *cluster* 1 menurun menjadi 18 provinsi dan *cluster* 2 meningkat menjadi 20 provinsi. Pada 2024, *cluster* 1 semakin menurun menjadi 14 provinsi, sementara *cluster* 2 meningkat menjadi 24 provinsi.

REFERENCES

- [1] WHO, "Stunting in a nutshell," World Health Organization. Accessed: Oct. 07, 2024. [Online]. Available: <https://www.who.int/news/item/19-11-2015-stunting-in-a-nutshell>
- [2] M. Ariani, "Determinan Penyebab Kejadian Stunting Pada Balita: Tinjauan Literatur," *Dinamika Kesehatan Jurnal Kebidanan dan Keperawatan*, vol. 11, no. 1, pp. 2549–4058, 2020, doi: 10.33859/dksm.v11i1.
- [3] Kementerian Kesehatan RI, "Laporan SKI TEMATIK 2023," Kementerian Kesehatan RI. Accessed: Dec. 20, 2024. [Online]. Available: <https://www.badankebijakan.kemkes.go.id/laporan-tematik-ski/>
- [4] Kementerian Kesehatan RI, "Prevalensi Stunting di Indonesia Turun ke 21,6% dari 24,4%," Biro Komunikasi dan Pelayanan Publik. Accessed: Oct. 11, 2024. [Online]. Available: <https://kemkes.go.id/id/%20prevalensi-stunting-di-indonesia-turun-ke-216-dari-244>
- [5] R. Rotul Muhima *et al.*, *Kupas Tuntas Algoritma Clustering: Konsep Perhitungan Manual Dan Program*. 2021.
- [6] P. Aseknino and A. W. Wijayanto, "Analisis Perbandingan Metode Hierarchical dan Non-Hierarchical dalam Pembentukan Cluster Provinsi di Indonesia Berdasarkan Indikator Women Empowerment," *Indonesian Journal of Applied Statistics*, vol. 6, no. 1, p. 57, Jan. 2024, doi: 10.13057/ijas.v6i1.68876.
- [7] M. Handayani and M. F. L. Sibuea, "Performance Analysis of Clustering Models Based on Machine Learning in Stunting Data Mapping," *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 9, no. 4, pp. 715–720, Sep. 2023, doi: 10.33330/jurteks.v9i4.2770.
- [8] N. Nur Afidah and M. Masrukan, "Penerapan Metode Clustering dengan Algoritma K-means untuk Pengelompokkan Data Migrasi Penduduk Tiap Kecamatan di Kabupaten Rembang," *PRISMA, Prosiding Seminar Nasional Matematika*, vol. 6, pp. 729–738, 2023, Accessed: Dec. 10, 2024. [Online]. Available: <https://journal.unnes.ac.id/sju/prisma/article/view/67038>
- [9] T. A. Cinderatama *et al.*, "Implementasi Metode K-Means, DbSCAN, dan MeanShift Untuk Analisis Jenis Ancaman Jaringan Pada Intrusion Detection System," *Jurnal INOVTEK Polbeng*, vol. 7, no. 1, pp. 169–184, 2022, doi: 10.35314/isi.v7i1.2336.
- [10] D. Fakta Sari, A. Kusjani, D. Kurniawati, and I. Setiawan, "PENCARIAN DATA QUICK COUNT PILPRES DENGAN TEKNIK WEB SCRAPING," *JIRK (Journal of Innovation Research and Knowledge)*,

-
- vol. 3, no. 5, pp. 1025–1034, 2023, Accessed: Jan. 14, 2025. [Online]. Available: <https://bajangjournal.com/index.php/JIRK/article/view/6695>
- [11] F. Marisa, A. L. Maukar, and T. Muhammad Akhriza, *Data Mining Konsep dan Penerapannya*. 2021.
- [12] A. Iswahyudi Yasril, F. Fatma, and D. Febrianti, "Penerapan Uji Korelasi Spearman Untuk Mengkaji Faktor yang Berhubungan Dengan Kejadian Diabetes Melitus di Puskesmas Sicincin Kabupaten Padang Pariaman," *Jurnal Human Care*, vol. 6, no. 3, pp. 527–533, 2021, doi: 10.32883/hcj.v6i3.1444.
- [13] K. Seljuna Monika, N. Anggraini, and W. Rajagukguk, "Hubungan Insentif dan Lingkungan Kerja Terhadap Komitmen Organisasi Pada Karyawan Pt. Kb Finansia Multi Finance," *Fundamental Management Journal*, vol. 7, no. 1, pp. 123–174, 2022, doi: 10.33541/fjm.v7i1p.3888.
- [14] M. Riszky Sulaeman, T. Nur Padilah, U. Singaperbangsa Karawang, J. HSRonggo Waluyo, T. Timur, and J. Barat, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Kecelakaan Berkendara Di Ruas Tol Jakarta-Cikampek," *Jurnal informasi dan Komputer*, no. 1, p. 11, 2023, doi: 10.35959/jik.v11i01.369.
- [15] D. A. I. C. Dewi and D. A. K. Pramita, "Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali," *Matrix: Jurnal Manajemen Teknologi dan Informatika*, vol. 9, no. 3, pp. 102–109, 2019, doi: 10.31940/matrix.v9i3.1662.