

Analisis Klasterisasi Mengenai Tingkat Prevelensi Stunting di Jawa Barat Tahun 2023

Amalia Tiara*¹, Muhammad Ridho Mujahid², Nabilah Putri Salsabila³

^{1,2,3}Jurusan Administrasi Publik, UIN Sunan Gunung Djati Bandung, Indonesia

*Korespondensi: amltiara1901@gmail.com

KATA KUNCI	ABSTRAK
stunting, jawa barat, klustering K- means	<i>Penelitian ini menganalisis tingkat prevalensi stunting di 27 kabupaten/kota di Jawa Barat tahun 2023 menggunakan algoritma K-Means Clustering. Stunting merupakan permasalahan kesehatan serius yang memengaruhi kualitas sumber daya manusia, terutama di wilayah dengan akses sanitasi dan kondisi sosial-ekonomi rendah. Dengan data prevalensi stunting, akses sanitasi, dan tingkat kemiskinan, penelitian ini mengelompokkan wilayah menjadi empat klaster berdasarkan karakteristik yang serupa. Klasterisasi bertujuan untuk memetakan wilayah prioritas intervensi guna mendukung kebijakan berbasis data. Hasil penelitian menunjukkan bahwa wilayah dengan akses sanitasi rendah cenderung memiliki prevalensi stunting tinggi. Klasterisasi menghasilkan wawasan penting, seperti hubungan antara kondisi sosial-ekonomi dan pola distribusi stunting, serta akses sanitasi. Metode Elbow dan Silhouette Plot digunakan dalam R-Studio untuk menilai kualitas klasterisasi, menghasilkan rata-rata silhouette width sebesar 0,33 dengan 4 klaster. Dan data juga menggunakan Python untuk dianalisis dan divisualisasikan dalam ruang dua dimensi menggunakan metode Principal Component Analysis (PCA). Visualisasi ini memetakan data ke dalam lima klaster berbeda. Penelitian ini memberikan kontribusi signifikan bagi pemerintah daerah dalam merumuskan kebijakan intervensi berbasis wilayah. Dengan pendekatan ini, penelitian diharapkan dapat mendukung upaya penurunan prevalensi stunting di Jawa Barat secara efektif.</i>
KEYWORDS	ABSTRACT
stunting, west java, K-means clustering	<i>This study analyzes the prevalence rate of stunting in 27 districts/cities in West Java in 2023 using the K-Means Clustering algorithm. Stunting is a serious health problem that affects the quality of human resources, especially in areas with low sanitation access and socio-economic conditions. With data on stunting prevalence, sanitation access, and poverty levels, this study groups areas into four clusters based on similar characteristics. Clustering aims to map priority intervention areas to support data-based policies. The results show that areas with low sanitation access tend to have a high prevalence of stunting. Clustering produces important insights, such as the relationship between socio-economic conditions and stunting distribution patterns, as well as sanitation access. The Elbow and Silhouette Plot methods were used in R-Studio to assess the quality of clustering, resulting in an average silhouette width of 0.33 with 4 clusters. And the data also uses Python to be analyzed and visualized in two-dimensional space using the Principal Component Analysis (PCA) method. This visualization maps data into five different clusters. This study makes a significant contribution to local governments in formulating area-based intervention policies. With this approach, research is expected to support efforts to reduce the prevalence of stunting in West Java effectively.</i>

Latar Belakang

Di Indonesia, stunting adalah permasalahan kesehatan yang sangat serius, terutama di daerah dengan sosial ekonomi rendah seperti Jawa Barat. Asupan makanan yang tidak cukup serta faktor lingkungan jangka panjang seperti terbatasnya akses fasilitas sanitasi dan medis merupakan penyebab stunting pada balita. Stunting ini cukup umum di Jawa Barat, salah satu provinsi di Indonesia, terutama di daerah dengan kesejahteraan rendah dan akses ke fasilitas medis yang sedikit. Menurut data Kementerian Kesehatan RI, prevalensi stunting Jabar masih lebih tinggi dari rata-rata nasional, menekankan perlunya intervensi terfokus untuk mengatasi masalah ini.

Salah satu hambatan terbesar dalam meningkatkan sumber daya manusia adalah stunting, stunting merupakan masalah kesehatan yang ditandai dengan tinggi badan yang lebih rendah dari standar usia karena tingkat nutrisi yang terbatas dalam jangka panjang. Hal ini perlu diperhatikan karena menyebabkan disabilitas pada perkembangan fisik, kognitif, dan sosial yang dapat memangkas produktivitas di masa depan dan tingkat kualitas rentang hidup individu.

Masalah stunting bisa dialami oleh anak-anak di seluruh dunia, terutama di negara-negara berkembang seperti Indonesia. Penelitian ini berfokus pada wilayah Jawa Barat yang paling berisiko terkena dampak buruk stunting. Anak-anak dengan keterbatasan ekonomi adalah kelompok yang paling rentan terhadap kekurangan zat makanan dan layanan kesehatan yang kurang memadai. Selain itu, faktor lingkungan, dan terutama faktor-faktor sanitasi, selain pola asuh anak, juga meningkatkan risiko terjadinya stunting. Stunting ini sendiri merupakan kondisi gagal tumbuh pada anak balita (bayi dibawah lima tahun) akibat dari kekurangan gizi kronis, sehingga anak terlalu pendek untuk usianya. Kekurangan gizi terjadi sejak bayi dalam kandungan dan pada awal setelah bayi lahir, akan tetapi kondisi stunting baru terlihat ketika bayi berusia 2 tahun.

Di Indonesia, stunting adalah masalah kesehatan masyarakat utama. Indonesia dikhususkan untuk menurunkan prevalensi stunting hingga 14% pada 2024, sesuai dengan target Kementerian Kesehatan 2023. Dan sangat sulit bagi Jawa Barat, provinsi dengan populasi terbanyak. Provinsi ini menjadi prioritas dalam upaya nasional untuk meminimalkan stunting karena sejumlah masalah sosial-ekonomi, distribusi layanan kesehatan yang tidak merata, dan masalah lingkungan. Penegasan tingkat stunting Jabar yang bisa menghambat

pembangunan SDM daerah membuat penelitian ini mendesak. Menurut Kementerian Kesehatan, prevalensi stunting di Jawa Barat sebesar pada tahun 2023 sebanyak 21,7% yang menunjukkan adanya kesulitan dalam menurunkan tingkat stunting di Jawa Barat ini.. Selain itu, dibandingkan dengan pada tahun 2022 sebesar 20,2%, jumlah tersebut menunjukkan peningkatan prevalensi stunting Jawa Barat sebesar 1,5 persen dari tahun sebelumnya.

Penelitian Terdahulu

Berbagai kajian/penelitian tentang stunting di Jawa Barat telah dilakukan oleh peneliti, contohnya, dalam judul Clustering daerah rawan stunting di Jawa Barat menggunakan algoritma K-Means menyatakan bahwa Permasalahan pada stunting tidak boleh dibiarkan dan harus segera ditangani karena dapat menimbulkan efek negatif bagi kesehatan dalam jangka panjang ataupun jangka pendek (Taufik Hidayat et al., 2023)

Penelitian mengenai stunting di Indonesia, khususnya di Jawa Barat, telah banyak dilakukan. Namun, sebagian besar kajian sebelumnya masih berfokus pada tingkat nasional atau provinsi tanpa memberikan pemetaan yang mendalam hingga tingkat kabupaten/kota. Studi-studi seperti yang dilakukan sebelumnya lebih mengedepankan pendekatan spasial dan analisis statistik sederhana untuk memahami prevalensi stunting. Meskipun pendekatan ini memiliki nilai penting, mereka belum berhasil menggali variasi karakteristik antarwilayah secara substansial. Untuk mengisi celah tersebut, penelitian ini menerapkan metode klasterisasi dengan menggunakan algoritma K-Means Clustering. Metode ini memungkinkan pengelompokan wilayah berdasarkan kesamaan karakteristik terkait prevalensi stunting, akses sanitasi, dan kondisi sosial-ekonomi. Dengan pendekatan ini, penelitian ini memberikan kontribusi baru dalam mengidentifikasi pola distribusi stunting di 27 kabupaten/kota di Jawa Barat (Riznawati et al., 2023).

Penelitian ini juga memenuhi kebutuhan akan data yang lebih terperinci hingga tingkat lokal, yang ditekankan oleh Astuti et al. (2018). Sebelumnya, kajian-kajian hanya menyebutkan pentingnya data rinci tanpa memanfaatkan pendekatan kuantitatif yang mendalam untuk menyusun peta risiko atau mendukung perencanaan kebijakan berbasis data. Dalam artikel ini, data komprehensif dari 27 kabupaten/kota di Jawa Barat digunakan untuk menghasilkan analisis berbasis bukti yang mendukung perumusan kebijakan. Variabel yang dianalisis meliputi prevalensi stunting, akses sanitasi, dan Prevalensi penduduk miskin di Jawa Barat, menciptakan pendekatan holistik yang belum banyak diterapkan di penelitian lain.

Penelitian ini memiliki keunikan dalam penerapan algoritma K-Means Clustering menggunakan R-Studio dan Python yang memungkinkan pemetaan stunting balita hingga ke tingkat kabupaten/kota di Provinsi Jawa Barat. Pendekatan ini memberikan nilai tambah yang signifikan dibandingkan penelitian sebelumnya yang umumnya hanya berfokus pada analisis prevalensi di tingkat nasional atau provinsi tanpa memetakan variasi antarwilayah. Studi ini membantu pemerintah untuk lebih mengkoordinasikan upaya pengendalian stunting, membantu memberikan informasi lebih rinci tentang setiap kabupaten dan kota di Jawa Barat dengan faktor-faktor yang terjadi, termasuk pemetaan kluster dan jumlah kriteria. Penelitian yang menggunakan pendekatan pengelompokan ini diharapkan dapat secara langsung mendukung perencanaan kebijakan berbasis data, sehingga memungkinkan intervensi yang lebih luas di daerah dengan tingkat stunting yang tinggi.

Tinjauan Teoritis

Stunting adalah kondisi keterlambatan pertumbuhan pada anak yang disebabkan oleh kekurangan gizi kronis, terutama selama 1000 hari pertama kehidupan, yaitu dari masa kehamilan hingga usia dua tahun. Kondisi ini dapat memengaruhi perkembangan fisik dan kognitif anak serta berdampak pada kualitas sumber daya manusia di masa depan (Kementerian Kesehatan RI, 2018). Stunting masih perlu ditangani secara serius dalam skala global karena efeknya yang luas dan bertahan lama. Di Indonesia, penanganan dan pencegahan stunting telah muncul sebagai prioritas utama pembangunan kesehatan nasional. Tingkat prevalensi stunting tetap tinggi meskipun penerapan beberapa program intervensi, menunjukkan bahwa efektivitas program perlu ditingkatkan.

Penelitian ini menggunakan algoritma K-Means Clustering untuk mengelompokkan daerah rawan stunting di kabupaten/kota di Jawa Barat selama periode 2019-2021. Dengan menggunakan metode Elbow untuk menentukan jumlah kluster optimal, dan hasil evaluasi menggunakan Silhouette Coefficient juga sangat penting dalam mengkategorikan agar dapat sebagai kluster yang layak. Penelitian ini menunjukkan bahwa metode clustering dengan K-Means efektif untuk menganalisis data stunting di Jawa Barat, memberikan wawasan penting untuk mendukung program intervensi berbasis wilayah (Hidayat et al, 2023).

Penelitian dari Sudrajat et al. (2024) menunjukkan hasil klusterisasi menjadi empat kelompok: dua dengan tingkat stunting rendah, satu dengan tingkat tinggi, dan satu sangat

tinggi. Data yang digunakan mencakup informasi dari tahun 2018-2022, dengan metode evaluasi menggunakan

RapidMiner dan Elbow Method untuk menentukan jumlah cluster optimal. Pendekatan spasial dengan menggunakan Sistem Informasi Geografis (SIG) dan analisis autokorelasi spasial dapat memberikan gambaran pola distribusi stunting berdasarkan wilayah. Penelitian ini menggunakan metode Indeks Moran untuk melihat apakah terdapat pola pengelompokan kasus stunting di Jawa Barat. Hasil penelitian menunjukkan bahwa terdapat autokorelasi spasial positif, yang berarti prevalensi stunting cenderung mengelompok pada wilayah-wilayah tertentu (Riznawati et al, 2021).

Penelitian ini mengklasifikasikan data prevalensi balita stunting di tingkat kabupaten/kota di Provinsi Jawa Barat menggunakan pendekatan Analisis Klasterisasi (K-Means) data, karena kemampuan efektif dan efisien untuk mengelompokkan data dengan jumlah sampel yang tinggi dan beragam karakteristik, teknik Clustering K-Means dipilih. Agar bisa membagi data ke dalam beberapa cluster. Dalam penelitian ini, klaster yang menunjukkan pola pencegahan stunting rendah, menengah, dan tinggi dapat diidentifikasi melalui analisis data prevalensi stunting yang dikumpulkan dari 27 kabupaten/kota di Jawa Barat. Dan data prevalensi stunting Jawa Barat diperoleh dari data statistik yang dirilis oleh pemerintah, yang menyediakan informasi prevalensi stunting sebagai persentase dari jumlah total balita di setiap wilayah Jawa Barat.

Metode Penelitian

Penelitian ini menggunakan pendekatan analisis data kuantitatif dengan metode K-means clustering untuk mengelompokkan kabupaten/kota di Jawa Barat berdasarkan prevalensi stunting dan faktor-faktor terkait lainnya seperti kondisi sosial-ekonomi, dan kualitas sanitasi. Data yang digunakan berasal dari Open Data Jawa Barat yang mencakup prevalensi stunting di 27 kabupaten/kota di provinsi tersebut. Dataset ini memberikan informasi lengkap yang memungkinkan analisis distribusi stunting di berbagai daerah.

Metode K-means dipilih karena kemampuannya untuk mengelompokkan data berdasarkan kesamaan karakteristik, tanpa memerlukan label sebelumnya. Dalam penelitian ini, variabel yang digunakan untuk pengelompokan meliputi prevalensi stunting, kondisi ekonomi, dan kualitas sanitasi. Proses clustering dilakukan dengan menentukan jumlah klaster yang optimal menggunakan metode elbow untuk memilih jumlah klaster yang paling relevan.

Setelah pengelompokan dilakukan, analisis lebih lanjut akan dilakukan untuk menilai karakteristik setiap klaster dan hubungan antara faktor-faktor yang mempengaruhi prevalensi stunting di masing-masing kabupaten/kota. Hasil clustering ini diharapkan dapat memberikan gambaran yang lebih jelas mengenai daerah-daerah dengan prevalensi stunting tinggi dan faktor-faktor yang perlu ditargetkan dalam intervensi kebijakan. Semua data yang digunakan akan dianalisis secara deskriptif dan dianalisis menggunakan perangkat lunak statistik seperti R atau Python untuk mengimplementasikan algoritma K-means.

Hasil dan Pembahasan

Teknik klasterisasi adalah metode analisis data yang bertujuan untuk mengelompokkan objek berdasarkan karakteristik yang serupa. Dalam studi ini, algoritma K-Means diterapkan untuk melakukan analisis klasterisasi terkait variabel prevalensi stunting di kota dan kabupaten di Jawa Barat pada tahun 2023. Tujuan dari analisis ini adalah mengelompokkan wilayah Kabupaten/Kota di Jawa Barat ke dalam beberapa klaster berdasarkan indikator kepadatan penduduk, sehingga dapat mendukung pembuatan kebijakan yang berbasis data. Dataset yang digunakan memuat informasi terkait variabel prevalensi stunting di kota/kabupaten di Jawa Barat, persentase penduduk miskin di Kabupaten/Kota di Jawa Barat, dan persentase rumah tangga dengan akses terhadap sanitasi layak di Jawa Barat.

Klasterisasi Menggunakan R-Studio

Informasi mengenai Analisis Klasterisasi Tingkat Prevalensi Stunting di Jawa Barat Tahun 2023, yang melibatkan 27 Kota/Kabupaten di wilayah tersebut, telah dihimpun dan dikelompokkan menjadi empat kategori (cluster), yaitu tingkat prevalensi stunting rendah, sedang, tinggi, dan sangat tinggi. Data tersebut mencakup prevalensi stunting di Kota/Kabupaten di Jawa Barat, persentase penduduk miskin di Kabupaten/Kota di Jawa Barat, dan persentase rumah tangga dengan akses terhadap sanitasi layak di Jawa Barat, yang selanjutnya dianalisis menggunakan algoritma K-Means Clustering.

Kabupaten/Kota	x1	x2	x3	Cluster
Bogor	27,6	71,49	7,27	2
Sukabumi	27	53,21	7,01	1
Cianjur	11,4	63,83	10,22	4

Bandung	29,2	66,67	6,4	2
Garut	24,1	54,68	9,77	1
Tasikmalaya	20,7	51,33	10,28	1
Ciamis	25,4	71,38	7,42	2
Kuningan	23,4	83,03	12,12	3
Cirebon	22,9	89,02	11,2	3
Majalengka	24,1	82,67	11,21	3
Sumedang	14,4	72,55	9,36	4
Indramayu	18,4	95,9	12,13	3
Subang	18,7	89,4	9,52	3
Purwakarta	24	77,33	8,46	2
Karawang	17,1	83,02	7,87	4
Bekasi	23,2	91,54	4,93	2
Bandung Barat	25,1	56,93	10,52	1
Pangandaran	23,9	86,91	8,98	3
Kota Bogor	18,2	74,36	6,67	4
Kota Sukabumi	26,9	44,76	7,5	1
Kota Bandung	16,3	56,01	3,96	4
Kota Cirebon	19,9	93,27	9,16	3
Kota Bekasi	10,3	98,52	4,1	4
Kota Depok	14,3	97,43	2,38	4
Kota Cimahi	24,5	76,37	4,66	2
KotaTasikmalaya	27,1	56,25	11,53	1
Kota Banjar	23,6	89,06	6,14	2

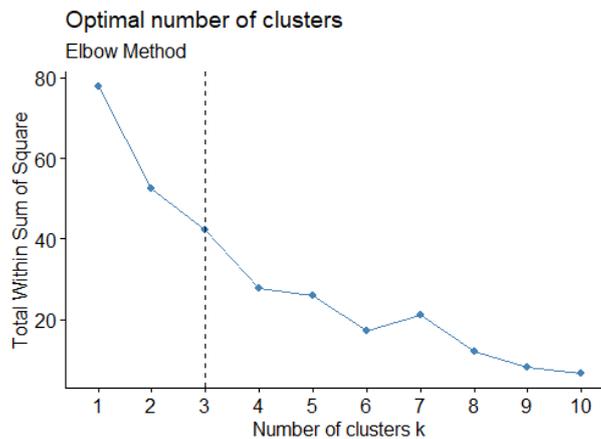
Tabel 1. Analisis Klasterisasi Mengenai Tingkat Prevelensi Stunting Di Jawa Barat Tahun 2023
 Sumber: Data Diolah Peneliti (2024)

Analisis ini menggunakan metode Elbow untuk menentukan jumlah kluster optimal serta visualisasi hasil klasterisasi dengan variabel utama yaitu persentase stunting dan akses terhadap sanitasi layak.

Penentuan Kluster Optimal dengan Metode Elbow

Metode Elbow digunakan untuk menentukan jumlah kluster optimal (k) dengan mengamati grafik Total Within Sum of Squares (TWSS). Grafik ini menunjukkan penurunan TWSS seiring dengan bertambahnya jumlah kluster (k). Berdasarkan grafik yang dihasilkan, titik elbow terlihat jelas pada k=3, di mana penurunan TWSS mulai melambat drastis setelah titik tersebut. Oleh karena itu, jumlah kluster optimal yang digunakan adalah tiga kluster. Pemilihan k=3 memberikan pembagian data yang efisien dan memadai untuk dianalisis lebih

lanjut. Dalam analisis ini, meskipun elbow point terdeteksi di kluster 3, algoritma K-Means diputuskan untuk dijalankan dengan 4 kluster karena menunjukkan variasi karakteristik wilayah yang lebih informatif terhadap data.



Gambar 1. Hasil K optimal dengan elbow visualizer

Setelah menentukan jumlah cluster yang optimal, pengelompokan data dapat dilakukan menggunakan Algoritma K-Means melalui pemrograman R(studio) dengan menerapkan console/sintaks berikut:



Gambar 2. Hasil Syntax K-means Clustering

Proses klusterisasi dimulai dengan menentukan jumlah kluster optimal menggunakan metode Elbow, yang menunjukkan titik optimal pada empat kluster. Setelah itu, data dianalisis menggunakan algoritma K-Means dengan parameter, dengan Jumlah kluster 4, Parameter nstart 20, untuk memastikan hasil yang stabil.

Hasil analisis menghasilkan 4 kluster, dengan distribusi ukuran masing-masing kluster sebagai berikut: (a) Kluster 1: 6 data; (b) Kluster 2: 7 data; (c) Kluster 3: 7 data; (d) Kluster 4: 7

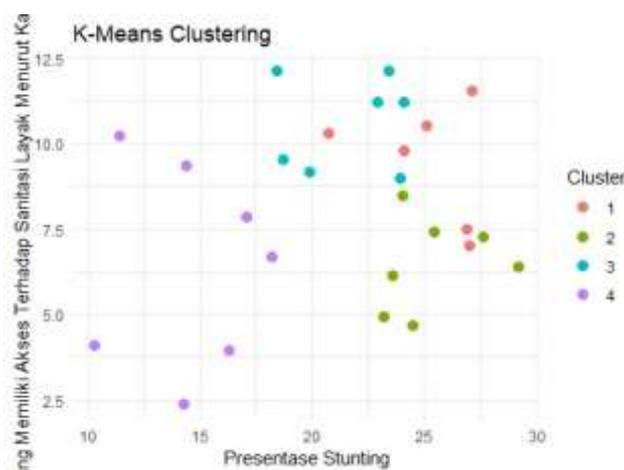
data. Rata-rata (mean) tiap klaster untuk variabel X1, X2, dan X3 adalah: (a) Klaster 1: [0.71000467, -1.3871678, 0.4703590]; (b) Klaster 2: [0.75083328, 0.1634867, -0.6384779]; (c) Klaster 3: [0.01735801, 0.8449302, 0.9122384] (d) Klaster 4: [-1.37199384, -0.4821020, -0.6769254].

Centroid adalah nilai rata-rata yang mewakili setiap variabel dalam sebuah klaster. Pada metode K-Means, centroid berfungsi sebagai titik pusat klaster yang menggambarkan karakteristik data dalam kelompok tersebut.

Nilai $\text{between_SS} / \text{total_SS}$ sebesar 64.5% menunjukkan bahwa 64.5% variasi total dalam data berhasil dipaparkan oleh klasterisasi ini. Hal ini mencerminkan kemampuan metode K-Means untuk mengelompokkan data secara efektif berdasarkan kesamaan karakteristik. Penentuan Prioritas Intervensi, wilayah dalam Klaster 2 dan Klaster 4, yang memiliki prevalensi stunting tinggi dan keterbatasan akses, dapat menjadi prioritas untuk intervensi lebih lanjut.

Visualisasi Klasterisasi

Proses pengelompokan data menggunakan metode K-Means berhasil membagi data menjadi empat kelompok berbeda (klaster), di mana setiap klaster menunjukkan karakteristik unik berdasarkan persentase prevalensi stunting dan persentase populasi yang memiliki akses ke fasilitas sanitasi layak. Hasil analisis ini ditampilkan dalam bentuk scatter plot, dengan setiap klaster diberi warna berbeda untuk mempermudah interpretasi. Hasil klasterisasi membagi data menjadi empat kelompok utama, yang masing-masing ditandai dengan warna berbeda.



Gambar 3. Hasil K-means Clustering

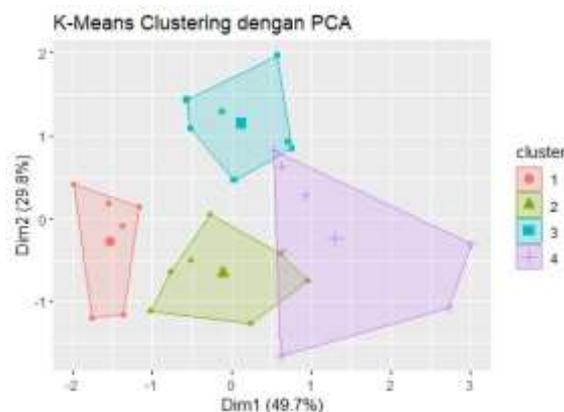
Proses pengelompokan data menggunakan metode K-Means berhasil membagi data menjadi empat kelompok berbeda (klaster), di mana setiap klaster menunjukkan karakteristik unik berdasarkan persentase prevalensi stunting dan persentase populasi yang memiliki akses ke fasilitas sanitasi layak. Hasil analisis ini ditampilkan dalam bentuk scatter plot, dengan setiap klaster diberi warna berbeda untuk mempermudah interpretasi. Hasil klasterisasi membagi data menjadi empat kelompok utama, yang masing-masing ditandai dengan warna berbeda.

Hasil klasterisasi menunjukkan hubungan yang signifikan antara prevalensi stunting dan akses terhadap sanitasi layak yaitu daerah dengan akses sanitasi yang lebih tinggi cenderung memiliki prevalensi stunting yang lebih rendah (contoh: Klaster 3) Sebaliknya, daerah dengan akses sanitasi yang buruk memiliki prevalensi stunting yang tinggi (contoh: Klaster 2).

Dengan mengelompokkan daerah-daerah berdasarkan karakteristik yang serupa, intervensi kebijakan dapat dirancang secara lebih terfokus dan efektif. Fokus utama adalah meningkatkan akses sanitasi dan intervensi gizi di daerah dengan kombinasi masalah stunting tinggi dan sanitasi buruk, khususnya di Klaster 2.

K-Means Clustering dengan PCA

Analisis klasterisasi dilakukan untuk memahami pola distribusi prevalensi stunting di Jawa Barat pada tahun 2023. Proses klasterisasi ini menggunakan algoritma K-Means yang hasilnya divisualisasikan melalui analisis komponen utama (PCA). Dari hasil tersebut, didapatkan empat klaster yang masing-masing mewakili wilayah dengan karakteristik tertentu.



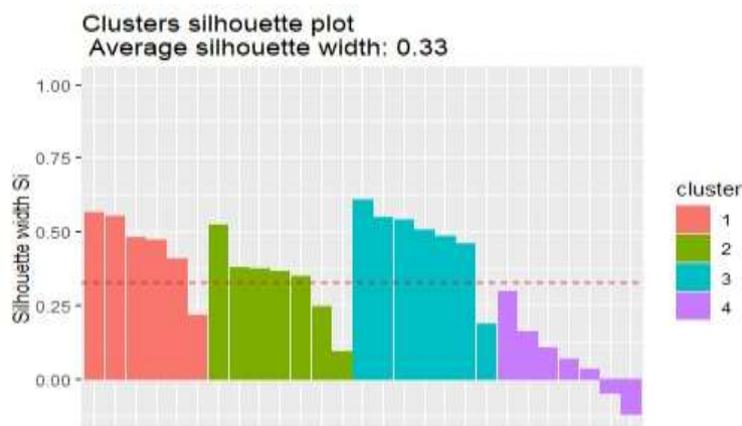
Gambar 4. Hasil K-means Clustering dengan PCA

Analisis klusterisasi dilakukan untuk memahami pola distribusi prevalensi stunting di Jawa Barat pada tahun 2023. Proses klusterisasi ini menggunakan algoritma K-Means yang hasilnya divisualisasikan melalui analisis komponen utama (PCA). Dari hasil tersebut, didapatkan empat kluster yang masing-masing mewakili wilayah dengan karakteristik tertentu.

Hasil Visualisasi Klusterisasi dengan PCA yaitu grafik pertama menunjukkan penyebaran data pada dua dimensi utama, yaitu Dimensi 1 (49,7%) dan Dimensi 2 (29,8%), yang secara keseluruhan menjelaskan sekitar 79,5% variasi dalam data. Keempat kluster ditampilkan dalam warna berbeda, yaitu kluster 1 (merah), kluster 2 (hijau), kluster 3 (biru muda), dan kluster 4 (ungu). Setiap kluster merepresentasikan kelompok wilayah dengan pola prevalensi stunting yang serupa.

Kualitas Klusterisasi Berdasarkan Silhouette Plot

Hasil analisis kualitas klusterisasi ditunjukkan melalui silhouette plot. Rata-rata silhouette width adalah 0,33, yang menandakan bahwa hasil klusterisasi kurang optimal. Idealnya, rata-rata silhouette width berada di atas 0,5 untuk menunjukkan klusterisasi yang baik.



Gambar 5. Hasil Silhouette Plot

Hasil analisis klusterisasi menunjukkan adanya perbedaan signifikan antara empat kluster yang terbentuk berdasarkan tingkat prevalensi stunting dan akses sanitasi. Kluster 1, yang terdiri dari 6 anggota, memiliki rata-rata nilai $x_1 = 0,71$, $x_2 = -1,39$, dan $x_3 = 0,47$, yang menunjukkan wilayah dengan prevalensi stunting rendah dan akses sanitasi yang cukup baik. Sementara itu, Kluster 4, dengan rata-rata $x_1 = -1,37$, $x_2 = 0,18$, dan $x_3 = -0,68$, mencerminkan

daerah dengan tingkat stunting tinggi dan akses sanitasi yang tidak optimal. Kluster 2 dan Kluster 3 masing-masing memiliki 7 anggota dengan karakteristik yang berbeda di antara keduanya.

Evaluasi kualitas klusterisasi dilakukan dengan menghitung rasio BSS/TSS, yang menunjukkan nilai sebesar 64,54%. Meskipun nilai ini cukup baik, hasilnya masih belum optimal karena idealnya berada pada kisaran 70-80%. Selain itu, total WSS yang mencapai 27,65592 menandakan adanya tingkat kompaksi data yang moderat di dalam masing-masing kluster. Evaluasi tambahan menggunakan Silhouette Plot menghasilkan rata-rata nilai silhouette width sebesar 0,33, yang mengindikasikan kualitas pemisahan antar kluster tergolong rendah. Beberapa anggota, khususnya di Kluster 4, memiliki nilai negatif yang menunjukkan kedekatan dengan kluster lain.

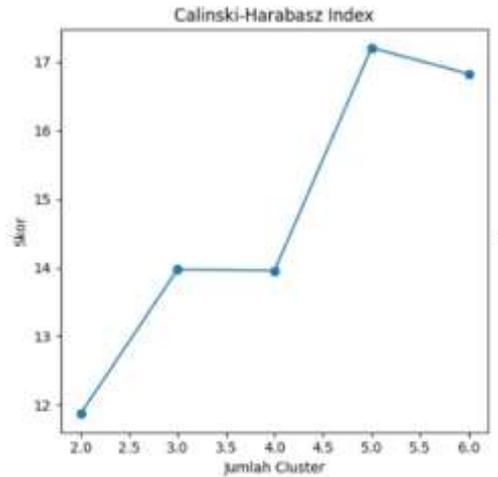
Hasil klusterisasi divisualisasikan dalam bentuk Scatter Plot dan PCA Cluster Plot. Scatter plot menggambarkan distribusi kluster berdasarkan variabel presentase prevalensi stunting (x_1) dan akses sanitasi layak (x_3), di mana Kluster 4 terlihat lebih tersebar dibandingkan kluster lainnya. PCA Cluster Plot, yang menggunakan analisis komponen utama untuk mereduksi dimensi data, menunjukkan beberapa tumpang tindih antar kluster, terutama Kluster 4, yang mendukung hasil dari silhouette plot.

Setiap kluster memiliki karakteristik spesifik berdasarkan analisis deskriptif. Kluster 1 mencakup daerah dengan prevalensi stunting rendah dan akses sanitasi baik. Kluster 2 terdiri dari wilayah dengan stunting moderat tetapi sanitasi cukup baik. Kluster 3 mencakup daerah dengan prevalensi stunting tinggi meskipun akses sanitasi relatif memadai. Sementara itu, Kluster 4 menggambarkan wilayah dengan prevalensi stunting tinggi serta akses sanitasi yang buruk.

Klusterisasi Menggunakan Phyton (Hasil Dinormalisasi)

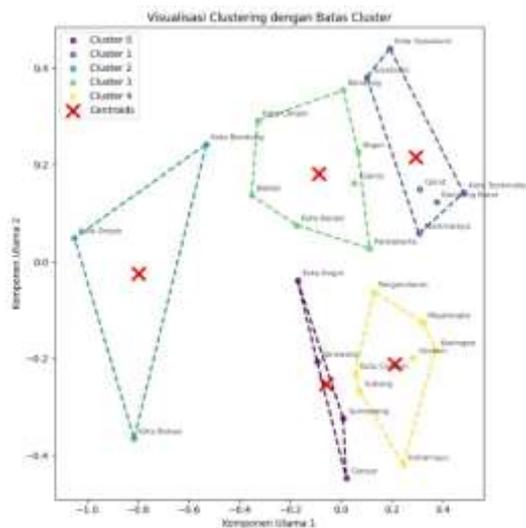
Berdasarkan visualisasi Calinski-Harabasz Index yang ditunjukkan pada grafik di atas, dapat disimpulkan bahwa jumlah kluster yang optimal adalah 5. Hal ini ditunjukkan oleh skor indeks tertinggi pada $k=5$ dengan nilai di atas 17. Kenaikan skor indeks yang signifikan pada jumlah kluster dari 4 ke 5 menegaskan bahwa proses klusterisasi menghasilkan kluster yang lebih baik pada konfigurasi 5 kluster dibandingkan jumlah kluster yang lebih sedikit. Setelah jumlah kluster mencapai 5, terjadi penurunan skor indeks saat $k=6$, yang

menunjukkan bahwa menambahkan jumlah kluster setelah titik tersebut tidak meningkatkan kualitas kluster secara signifikan.



Gambar 6. Hasil Calinski-Harabasz Index

Penentuan jumlah kluster yang optimal dengan metode Calinski-Harabasz ini sering digunakan dalam praktik analisis data karena memberikan keseimbangan antara kompleksitas model (jumlah kluster) dan kualitas pemisahan kluster yang dihasilkan. Nilai optimal dari indeks ini menunjukkan distribusi data yang paling sesuai berdasarkan kriteria evaluasi kompak dan terpisah (Arbelaitz et al., 2013). Dengan demikian, konfigurasi 5 kluster dapat dianggap sebagai solusi optimal untuk data ini berdasarkan hasil perhitungan Calinski-Harabasz Index.

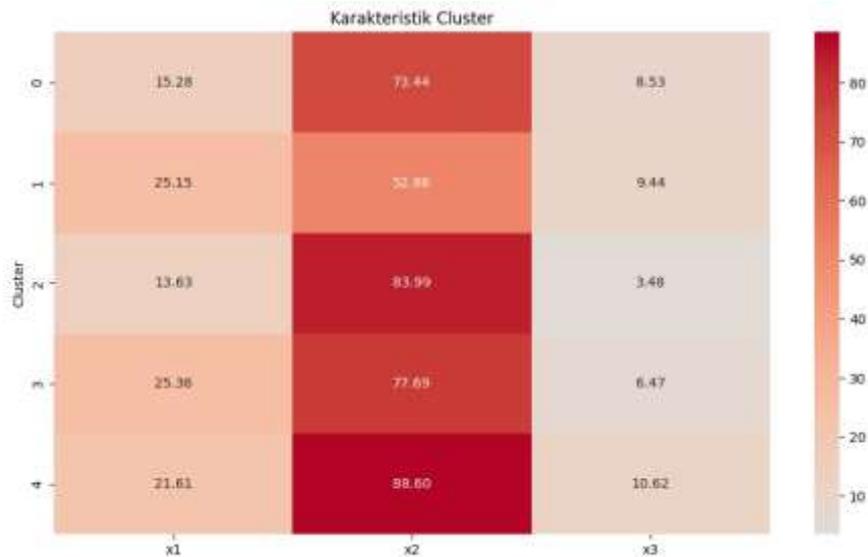


Gambar 7. Visualisasi Clustering

Visualisasi ini menunjukkan hasil analisis clustering berdasarkan Komponen Utama 1 dan Komponen Utama 2, yang merupakan hasil dari reduksi dimensi menggunakan metode Principal Component Analysis (PCA). Data ini dikelompokkan ke dalam lima kluster berbeda, yang ditandai dengan warna dan simbol unik, serta satu titik centroid per kluster yang ditandai dengan simbol "X" berwarna merah. Setiap kluster memiliki batasan yang dihubungkan dengan garis putus-putus sesuai dengan warna yang merepresentasikan masing-masing kluster.

Pertama, Cluster 0 (ungu) memiliki distribusi yang cenderung memanjang secara vertikal di sisi kiri bawah grafik. Wilayah ini mencakup Kota Bogor, Karawang, dan Cianjur. Titik centroid dari kluster ini berada di sekitar koordinat $(-0.1, -0.2)$. **Kedua**, Cluster 1 (biru tua) terdistribusi di bagian kanan atas visualisasi. Anggotanya meliputi Kota Sukabumi, Tasikmalaya, Bandung Barat, dan Purwakarta, dengan centroid terletak di sekitar $(0.2, 0.3)$. Pola ini menunjukkan adanya keterhubungan geografis antara wilayah-wilayah tersebut. **Ketiga**, Cluster 2 (hijau) mencakup wilayah yang relatif terpusat di bagian tengah grafik, dengan titik-titik seperti Kota Cimahi, Bandung, Bekasi, Banjar, dan Purwakarta. Centroid-nya berada di koordinat mendekati $(0.0, 0.2)$. Distribusi ini mengindikasikan hubungan kekerabatan data antara wilayah urban dan suburban di sekitar Bandung. **Keempat**, Cluster 3 (kuning) meliputi wilayah Cirebon, Subang, Majalengka, Kuningan, dan Indramayu, yang terletak di bagian kanan bawah visualisasi. Kluster ini memiliki centroid dekat dengan koordinat $(0.2, -0.1)$, dengan distribusi data yang membentuk pola melingkar. **Kelima**, Cluster 4 (biru kehijauan) memiliki persebaran yang lebih luas di sisi kiri atas grafik. Anggotanya meliputi Kota Depok dan Kota Bekasi, dengan centroid terletak di sekitar $(-0.8, 0.0)$. Pola distribusi ini menunjukkan perbedaan signifikan dalam variabel yang dianalisis dibandingkan dengan kluster lainnya, mengindikasikan bahwa wilayah tersebut memiliki karakteristik unik.

Secara umum, visualisasi ini menunjukkan bahwa analisis clustering berhasil mengelompokkan data menjadi lima kluster dengan distribusi yang jelas berdasarkan pola kedekatan geografis maupun karakteristik yang tidak disebutkan dalam visualisasi ini. Keberadaan centroid membantu untuk memahami pusat dari setiap kluster sebagai representasi rata-rata dari anggota-anggota di dalamnya. Visualisasi ini dapat menjadi dasar untuk memahami perbedaan karakteristik wilayah yang dianalisis dan memandu kebijakan berbasis klusterisasi.



Gambar 8. Karakteristik Kluster

Visualisasi ini menyajikan distribusi karakteristik variabel $x_1x_1x_1$, $x_2x_2x_2$, dan $x_3x_3x_3$ pada lima kluster berbeda (Cluster 0 hingga Cluster 4). Nilai-nilai pada matriks ditampilkan dalam bentuk warna yang bervariasi, dengan intensitas merah yang lebih pekat mengindikasikan nilai yang lebih tinggi. Hal ini memungkinkan analisis komprehensif mengenai dominasi variabel dalam setiap kluster.

Pertama, Cluster 0 memiliki karakteristik utama pada variabel $x_2x_2x_2$ dengan nilai 73.44, yang merupakan nilai dominan dibandingkan $x_1x_1x_1$ (15.28) dan $x_3x_3x_3$ (8.53). Hal ini menunjukkan bahwa variabel $x_2x_2x_2$ memiliki pengaruh yang signifikan dalam membentuk karakteristik kluster ini. **Kedua**, Cluster 1 memiliki kontribusi tertinggi pada variabel $x_2x_2x_2$ dengan nilai 52.86, sedangkan $x_1x_1x_1$ berada di angka 25.15 dan $x_3x_3x_3$ di angka 9.44. Nilai-nilai ini menegaskan bahwa $x_2x_2x_2$ tetap menjadi variabel yang dominan pada kluster ini, meskipun kontribusinya lebih rendah dibandingkan Cluster 0. **Ketiga**, Cluster 2 memperlihatkan nilai tertinggi pada variabel $x_2x_2x_2$ sebesar 83.99, yang menjadi nilai dominan tertinggi di antara seluruh kluster. Sementara itu, nilai $x_1x_1x_1$ hanya sebesar 13.63 dan $x_3x_3x_3$ relatif rendah di angka 3.48. Dominasi variabel $x_2x_2x_2$ yang signifikan pada kluster ini menunjukkan bahwa karakteristik kluster ini sangat ditentukan oleh variabel tersebut. **Keempat**, Cluster 3 memiliki distribusi variabel $x_2x_2x_2$ dengan nilai 77.69, yang juga menjadi karakteristik paling menonjol dibandingkan $x_1x_1x_1$ (25.36) dan $x_3x_3x_3$ (6.47). Kluster ini memperlihatkan kecenderungan pola yang serupa dengan Cluster 0 dan Cluster 1 dalam hal

dominasi variabel $x_2x_2x_2$. **Kelima**, Cluster 4 mencatat nilai tertinggi pada variabel $x_2x_2x_2$ dengan 88.60, menjadikannya klaster dengan nilai $x_2x_2x_2$ tertinggi di antara semua klaster. Sementara itu, $x_1x_1x_1$ memiliki nilai 21.61 dan $x_3x_3x_3$ sebesar 10.62. Distribusi ini mengindikasikan bahwa variabel $x_2x_2x_2$ merupakan pendorong utama pembentukan karakteristik Cluster 4.

Secara keseluruhan, analisis ini menunjukkan pola dominasi variabel $x_2x_2x_2$ pada semua klaster, dengan variasi intensitas yang berbeda-beda. Cluster 4 memiliki nilai $x_2x_2x_2$ tertinggi (88.60), diikuti oleh Cluster 2 (83.99), Cluster 3 (77.69), Cluster 0 (73.44), dan Cluster 1 (52.86). Hal ini menunjukkan bahwa variabel $x_2x_2x_2$ memiliki peran yang paling signifikan dalam membentuk pola clustering, sedangkan variabel $x_1x_1x_1$ dan $x_3x_3x_3$ memiliki kontribusi yang relatif lebih kecil dalam analisis ini. Visualisasi ini memberikan dasar yang kuat untuk memahami perbedaan karakteristik antar klaster berdasarkan ketiga variabel yang dianalisis.

Kesimpulan

Penelitian ini menggunakan algoritma K-Means Clustering untuk menganalisis prevalensi stunting di 27 kabupaten/kota di Jawa Barat pada tahun 2023. Hasil klasterisasi menunjukkan adanya hubungan yang signifikan antara prevalensi stunting, akses sanitasi, dan tingkat kemiskinan. Daerah dengan akses sanitasi yang buruk dan tingkat kemiskinan yang tinggi cenderung memiliki prevalensi stunting lebih tinggi, sehingga memerlukan perhatian khusus.

Dengan metode kuantitatif berbasis data, penelitian ini berhasil mengelompokkan wilayah menjadi beberapa klaster yang mencerminkan karakteristik sosial-ekonomi dan lingkungan yang beragam. Hasil ini memberikan dasar yang kuat bagi pemerintah untuk merancang kebijakan intervensi yang lebih terarah dan efektif dalam menangani masalah stunting, khususnya di wilayah-wilayah yang tergolong rawan.

Pendekatan berbasis data yang digunakan dalam penelitian ini juga memberikan kontribusi penting dalam memahami pola distribusi stunting secara lebih mendalam, mendukung perencanaan program kesehatan dan pembangunan yang berkelanjutan di Jawa Barat.

Rekomendasi Kebijakan

Berdasarkan temuan ini, beberapa rekomendasi kebijakan yang dapat diterapkan adalah: Fokus Intervensi pada Klaster 2 (Hijau) yaitu, program peningkatan akses sanitasi dan edukasi kesehatan harus difokuskan pada daerah-daerah dengan prevalensi stunting tinggi dan sanitasi rendah. Kemudian implementasi program gizi tambahan untuk anak-anak di wilayah ini juga sangat diperlukan. Pengembangan Infrastruktur Sanitasi untuk Klaster 4 (Ungu) meskipun prevalensi stunting tidak terlalu tinggi, daerah ini memerlukan perhatian dalam pembangunan infrastruktur sanitasi agar kualitas hidup masyarakat dapat meningkat. Pemeliharaan Kondisi pada Klaster 3 (Biru) yaitu, daerah ini mencerminkan keberhasilan dalam mengatasi stunting. Program-program yang sudah berjalan baik perlu dipertahankan dan dijadikan model bagi daerah lain.

Klaster 0: Implementasikan program perbaikan sanitasi di wilayah pedesaan, dukung pembangunan infrastruktur air bersih, dan sediakan bantuan sosial yang menasar keluarga miskin untuk meningkatkan akses sanitasi. Klaster 1: Memperkuat program edukasi masyarakat tentang pentingnya sanitasi dan kebersihan, serta dukung proyek-proyek sanitasi berkelanjutan melalui kerjasama pemerintah dan swasta. Klaster 2: Fokus pada peningkatan pengawasan dan perbaikan fasilitas sanitasi di daerah padat penduduk, serta implementasi kebijakan pemberdayaan ekonomi untuk mengurangi kemiskinan. Klaster 3: Prioritaskan pembangunan sanitasi dasar dan proyek air bersih yang menjangkau komunitas terpencil. Perluasan program bantuan pengentasan kemiskinan, seperti pelatihan keterampilan dan dukungan UMKM, akan membantu mengurangi beban ekonomi rumah tangga. Klaster 4: Perkuat program perbaikan perumahan dan sanitasi untuk meningkatkan kualitas hidup, serta edukasi masyarakat terkait kebersihan lingkungan.

Daftar Pustaka

- Adityaningrum, A., Arsad, N., Jusuf, H. (2021). FAKTOR PENYEBAB STUNTING DI INDONESIA: ANALISIS DATA SEKUNDER DATA SSGI TAHUN 2021 FACTORS CAUSING STUNTING IN INDONESIA: 2021 SSGI SECONDARY DATA ANALYSIS. *Jambura Journal of Epidemiology*, 3(1), 1-10.
- Fitri, R., Huljannah, N., Rochmah, T. (2022). PROGRAM PENCEGAHAN STUNTING DI INDONESIA: A SYSTEMATIC REVIEW. *National Nutrition Journal*, 17(3), 281-292.

- Hardinata, R., Oktaviana, L., Husain, F., Putri, S., Kartiasih, F. *Analisis Faktor-Faktor yang Memengaruhi Stunting di Indonesia Tahun 2021*.
- Riznawati, A., Yudhistira, D., Rahmaniati, M., Sipahutar, T., & Eryando, T. (n.d.). *Autokorelasi Spasial Prevalensi Stunting di Jawa Barat Tahun 2021*.
- Setiyawati, M. E., Ardhiyanti, L., Hamid, E. (n. d.). (2024). Studi Literatur: Keadaan Dan Penanganan Stunting Di Indonesia. *IKRAITH-HUMANIORA*, 8(2).
- Taufik Hidayat, Mohamad Jajuli, & Susilawati. (2023). Clustering daerah rawan stunting di Jawa Barat menggunakan algoritma K-Means. *INFOTECH : Jurnal Informatika & Teknologi*, 4(2), 137–146. <https://doi.org/10.37373/infotech.v4i2.642>
- Wibowo, A., Prasetyo, R., Nindya, T. (2022). Faktor-faktor yang Berhubungan dengan Kejadian Stunting pada Anak Usia di Bawah Lima Tahun di Indonesia: Tinjauan Literatur. *Jurnal Gizi Indonesia*, 10(2), 132-141.
- Widayatun. (2023). KEBERHASILAN DAN TANTANGAN PENURUNAN KASUS STUNTING DI INDONESIA: UPAYA MENCAPAI TARGET SDGS. *Jurnal Kependudukan dan Pembangunan Berkelanjutan*, 1(1), 33-43.