POLA KONSUMSI PANGAN SKOR PPH WILAYAH DI INDONESIA BERDASARKAN GIZI TIAP DAERAH 2021–2024 PENERAPAN METODE HIERARCHICAL CLUSTERING

Evan Alfadihilah Cipta, Bayu Priyatna, Fitria Nurapriani, April Lia Hananto
Sistem Informasi, Universitas Buana Perjuangan Karawang
Jalan Ronggo Waluyo Sirnabaya, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat
si22.evancipta@mhs.ubpkarawang.aci.id

ABSTRAK

Pola konsumsi pangan merupakan indikator utama dalam menilai kualitas gizi dan ketahanan pangan suatu wilayah. Namun, variasi Skor Pola Pangan Harapan (PPH) antarwilayah di Indonesia masih menunjukkan ketimpangan yang signifikan, sehingga menyulitkan penentuan prioritas intervensi kebijakan pangan. Penelitian ini bertujuan untuk mengelompokkan 507 wilayah Kabupaten/Kota berdasarkan kemiripan profil kinerja Skor PPH periode 2021–2024. Prosedur penelitian dilakukan menggunakan pendekatan *Knowledge Discovery in Databases* (KDD) dengan menerapkan metode *Agglomerative Hierarchical Clustering*. Algoritma diimplementasikan menggunakan *Euclidean Distance* sebagai pengukur jarak dan *Ward's Method* untuk meminimalkan varians di dalam klaster. Hasil validasi melalui *Silhouette Coefficient* memberikan nilai optimal sebesar 0,4022 pada jumlah tiga klaster (K=3). Penelitian ini berhasil memetakan tiga zonasi risiko gizi nasional: Klaster 1 (Zona Hijau) dengan performa stabil di atas skor 85,0 yang didominasi Pulau Jawa; Klaster 0 (Zona Kuning) sebagai kelompok transisi; dan Klaster 2 (Zona Merah) dengan risiko tinggi di bawah skor 70,0 yang terkonsentrasi di wilayah Timur Indonesia. Hasil ini memberikan landasan struktural bagi pemangku kebijakan dalam merancang program ketahanan pangan yang lebih tepat sasaran.

Kata kunci : *Hierarchical Clustering, Skor PPH, Ketahanan Pangan, Ward's Method, Silhouette Coefficient.*

1. PENDAHULUAN

Pola konsumsi pangan merupakan indikator krusial untuk menilai keamanan pangan dan kualitas gizi masyarakat di suatu wilayah. Parameter ini diukur melalui Skor Pola Pangan Harapan (PPH), di mana skor ideal nasional ditetapkan sebesar 100 untuk menunjukkan komposisi pangan yang beragam dan seimbang. Namun, fakta di lapangan menunjukkan skor PPH di berbagai wilayah Indonesia masih mengalami variasi signifikan dan ketimpangan yang dipengaruhi faktor ekonomi serta geografis. Permasalahan utama muncul karena dataset skor PPH yang mencakup lebih dari 500 Kabupaten/Kota memiliki dimensi yang besar, sehingga sulit untuk mengidentifikasi pola tersembunyi jika hanya menggunakan analisis deskriptif standar [1].

Ketidakseimbangan ini jika dibiarkan akan berkontribusi pada prevalensi gizi kurang dan gizi lebih secara simultan, serta memperlebar kesenjangan ketahanan pangan antarwilayah . Selama ini, analisis pangan lebih banyak menggunakan pendekatan regresi yang hanya melihat pengaruh variabel tertentu tanpa mampu mengelompokkan wilayah berdasarkan kesamaan profil konsumsi secara struktural [2]

Oleh karena itu, diperlukan pendekatan berbasis *data mining* menggunakan metode *Hierarchical Clustering* untuk mengelompokkan wilayah berdasarkan kemiripan profil skor PPH. [3]

Metode ini dipilih karena kemampuannya dalam menemukan struktur tersembunyi melalui visualisasi hierarki atau *dendrogram* yang mudah diinterpretasikan oleh pemangku kebijakan. Rencana penyelesaian masalah dalam penelitian ini dimulai dengan mengolah data sekunder periode 2021–2024 menggunakan bahasa pemrograman Python, menerapkan teknik *Ward's Method* untuk menghasilkan klaster yang homogen dengan meminimalkan varians internal. Validasi hasil dilakukan menggunakan *Silhouette Coefficient* untuk menjamin akurasi dan stabilitas klaster yang terbentuk. Melalui pendekatan ini, diharapkan diperoleh identifikasi wilayah yang akurat sehingga pemerintah dapat merancang intervensi kebijakan pangan yang lebih tepat sasaran, terutama pada wilayah yang teridentifikasi memiliki risiko ketimpangan gizi tinggi [4]

2. TINJAUAN PUSTAKA

Menghadapi kompleksitas ketahanan pangan di Indonesia menuntut sebuah pendekatan multidimensi yang mampu mengawinkan kekuatan teknologi informasi dengan studi nutrisi yang mendalam. Dalam lanskap ini, *data mining* muncul sebagai instrumen krusial yang memberikan kemampuan bagi peneliti untuk menyelami lautan data berskala besar dalam tempo yang sangat efisien. [5]

Secara fundamental, metode ini beroperasi layaknya perangkat cerdas yang mengandalkan analisis statistik untuk mengekstraksi wawasan berharga yang selama ini mungkin terkubur dalam tumpukan data mentah. [6] Inti dari mekanisme ini adalah melakukan audit mendalam terhadap basis data berukuran raksasa guna mengidentifikasi pola-pola

baru yang memiliki nilai strategis bagi pengambilan keputusan kebijakan public [3]

Guna mendukung proses penggalian informasi tersebut, bahasa pemrograman Python menjadi tulang punggung utama karena kesederhanaan sintaksnya yang tetap dipersenjatai oleh ekosistem pustaka yang sangat kuat. Kehadiran *library* seperti NumPy untuk komputasi numerik dan Pandas untuk manipulasi data membuat proses pengolahan informasi menjadi jauh lebih gesit, akurat, dan efektif bagi siapa saja yang ingin mengubah data menjadi sebuah solusi nyata. Melalui sinergi antara teknologi dan data ini, tantangan mengenai disparitas gizi di berbagai pelosok daerah dapat dipotret dengan jauh lebih presisi [5]

Penelitian ini menerapkan *Hierarchical Clustering* telah banyak diterapkan dalam analisis wilayah berbasis gizi dan pangan. Penelitian oleh Syfriza Davies Raihannabil menggunakan teknik AGNES dan DIANA untuk mengelompokkan 38 provinsi berdasarkan indikator status gizi anak baduta, yang menghasilkan dua klaster utama. [3]

Sejalan dengan itu, Rahman dan Sari mengkaji kinerja pangan nasional dan menemukan empat kelompok wilayah dengan karakteristik skor PPH yang berbeda. Selain itu, Brigitta Melati Kumarahadi menerapkan teknik *Average Linkage* untuk memetakan disparitas kemiskinan antar daerah. Wardina Humayrah dkk. juga menggunakan *Ward's Method* untuk mengidentifikasi tiga pola konsumsi utama pada perempuan usia reproduktif di Indonesia. Secara keseluruhan, studi-studi tersebut membuktikan bahwa *Hierarchical Clustering* sangat efektif dalam menghasilkan struktur klaster yang informatif untuk mendukung pengambilan keputusan berbasis data di sektor ketahanan pangan [8]

2.1. *Hierarchical Clustering*

Struktur hierarkis digunakan untuk pengumpulan data. Metode umum adalah hierarki agregasi, di mana setiap titik data pertama-tama diperlakukan sebagai himpunan individual dan kemudian secara bertahap digabungkan menjadi himpunan yang lebih besar. Proses ini memanfaatkan teknik pengelompokan (clustering); yaitu, elemen dengan karakteristik serupa dikelompokkan ke dalam satu himpunan, dan elemen dengan karakteristik berbeda dikelompokkan ke dalam himpunan lain [9]

Teknik yang digunakan dalam proses pengelompokan adalah metode Ward linkage, yang bertujuan untuk meminimalkan distribusi keseluruhan klaster. Rumus yang digunakan adalah sebagai berikut.

$$d(Ci\ ,\ Cj) = |Ci\ |+|Cj\ |\ T\ ||\mu i - \mu j\ || \qquad (1)$$

Keterangan: $d(Ci\ ,\ Cj)$: jarak anatara dua klaster, $|Ci\ |$, $|Cj\ |$: jumlah data dalam klaster Ci dan Cj , T : total jumlah data dalam dataset, μi , μj : centroid dari masing-masing klister

2.2. Pola Konsumsi Pangan

Konsumsi pangan merupakan informasi jenis dan jumlah pangan yang dikonsumsi seseorang atau sekelompok orang pada waktu tertentu. Umumnya, konsumsi energi dan protein menjadi jenis zat gizi yang sering dianalisis untuk mengukur ketahanan pangan dan kesejahteraan Masyarakat

Pola konsumsi makanan didefinisikan sebagai komposisi makanan yang mencakup jenis dan jumlah komponen makanan yang dikonsumsi oleh individu rata-rata setiap hari [3] dan dikonsumsi oleh masyarakat secara umum selama periode waktu tertentu. Kualitas sumber pangan ditentukan oleh komposisi jenis Untuk mencapai profil nutrisi yang seimbang, berbagai jenis makanan digunakan. Pola konsumsi dievaluasi baik secara kualitatif, berdasarkan apa yang dikonsumsi, maupun secara kuantitatif, berdasarkan kuantitas, jenis, dan frekuensi konsumsi [5]

2.3. Data Mining

Data mining adalah teknik yang memungkinkan pengguna untuk mengakses sejumlah besar data dalam waktu yang relatif singkat. Dengan kata lain, ini adalah alat dan aplikasi yang melakukan analisis statistik pada data dengan mengekstrak informasi dan data yang sebelumnya tidak diketahui.Sederhananya, data mining adalah proses menemukan informasi baru dengan mencari pola[10] atau aturan tertentu dalam sejumlah besar data. Artinya, mekanisme data mining terletak pada eksplorasi basis data besar untuk menemukan pola atau bentuk baru yang mungkin berguna dalam proses pengambilan keputusan. Keputusan [11]

2.4. *Python*

Python adalah bahasa pemrograman yang mudah digunakan, bahasa pemrograman yang sangat populer, dan salah satu dari beberapa bahasa pemrograman modern dan fleksibel yang memajukan komputasi modern. Bahasa ini banyak dimanfaatkan dalam pengembangan perangkat lunak, kecerdasan buatan, pengembangan web, machine *learning*, hingga analisis data [12]

Python juga dilengkapi dengansan ekosistem pustaka yang sangat lengkap, seperti NumPy untuk komputasi numerik dan Pandas untuk manipulasi serta analisis data, sehingga membuat proses pengolahan data menjadi lebih cepat, efisien, dan mudah dilakukan bahkan oleh pemula.yang sanngat Kelebihan inilah yang menjadikan Python sebagai bahasa yang sangat relevan untuk penelitian berbasis data dan komputasi ilmia

2.5. *Preprocessing*

Tahapan *preprocessing* biasanya mencakup beberapa langkah utama, antara lain penanganan data hilang (*missing values*), penghapusan data duplikat, normalisasi atau standardisasi agar skala data seragam, serta transformasi variabel kategorikal menjadi format numerik yang dapat diproses oleh algoritma data mining. Selain itu, proses ini juga dapat melibatkan

deteksi data pencilan (*outlier*), pengkodean variabel, serta penyelarasan struktur data [11]

Preprocessing yang dilakukan secara benar akan memastikan bahwa data berada dalam kondisi optimal sebelum masuk ke tahap analisis utama. Dengan demikian, hasil data mining dapat diperoleh secara lebih akurat, stabil, dan relevan terhadap tujuan penelitian [12]

3. METODE PENELITIAN

Riset ini dijalankan melalui lima pilar transformasi data yang saling berkelindan, menciptakan sebuah ekosistem analisis yang presisi, sistematis, dan sepenuhnya dapat diuji ulang oleh peneliti lain. Objek penelitian difokuskan secara mendalam pada analisis data sekunder mengenai Skor Pola Pangan Harapan (PPH) di seluruh Kabupaten dan Kota dalam wilayah kedaulatan Republik Indonesia. Seluruh data yang diolah bersumber secara resmi dari portal Badan Pangan Nasional (BAPANAS) yang menjamin validitas serta otentisitas informasi yang mencakup rentang waktu longitudinal dari tahun 2021 hingga 2024 guna menangkap dinamika tren gizi daerah secara berkesinambungan.Perjalanan data dimulai dengan tahap Data Selection, di mana dilakukan penyaringan ketat untuk memisahkan variabel krusial seperti Nama Wilayah, Provinsi, Tahun, dan nilai Skor PPH dari dataset utama guna memastikan hanya informasi yang relevan yang masuk ke tahap analisis. Sebelum melangkah lebih jauh, fase Preprocessing dan Transformation menjadi jembatan penting untuk memurnikan data dari gangguan serta menyelaraskan dimensi angka melalui teknik Min-Max Scaling. Langkah normalisasi ini sangat vital dalam memetakan nilai Skor PPH ke rentang skala 0 hingga 1, sehingga pergerakan angka di setiap tahun memiliki bobot kontribusi yang setara tanpa didominasi oleh variabel dengan rentang nilai yang besar.

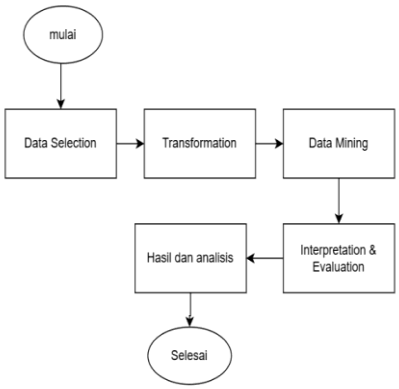
3.1. Objek Penelitian

Objek penelitian dalam studi ini difokuskan pada analisis data sekunder mengenai Skor Pola Pangan Harapan (PPH) Konsumsi pada tingkat Kabupaten dan Kota di seluruh wilayah kedaulatan Republik Indonesia. Pemilihan Skor PPH sebagai variabel utama didasarkan pada urgensi indikator ini dalam merepresentasikan kualitas, keberagaman, serta keseimbangan komposisi pangan yang dikonsumsi oleh masyarakat di suatu daerah

penelitian ini mencakup rentang waktu longitudinal selama empat tahun, dimulai dari tahun 2021 hingga 2024, guna menangkap dinamika dan tren perubahan kualitas gizi daerah secara berkesinambungan. Seluruh dataset yang digunakan bersumber secara resmi dari portal data Badan Pangan Nasional (BAPANAS), yang menjamin validitas dan otentisitas data yang diolah. Penarikan cakupan wilayah secara nasional (514 Kabupaten/Kota) bertujuan untuk mengidentifikasi disparitas kualitas konsumsi pangan secara makro dan memberikan pemetaan yang komprehensif mengenai kondisi.

3.2. Prosedur penelitian

Tahap awal ini merupakan proses penyaringan ketat untuk mengisolasi variabel yang benar-benar krusial, seperti identitas wilayah dan angka Skor PPH dari lautan data mentah. Tujuannya adalah memastikan bahwa hanya entitas data yang relevan yang masuk ke dalam mesin analisis.



Gambar 1. Alur Proses Penelitian

3.3. Data Selection

Tahap awal dimulai dengan proses seleksi data untuk memisahkan informasi yang relevan dengan tujuan penelitian. Peneliti mengambil data Skor PPH Kabupaten/Kota se-Indonesia dari dataset utama. Pada tahap ini, dipastikan bahwa setiap baris data memuat atribut penting seperti Nama Wilayah, Nama Provinsi, Tahun, dan nilai Skor PPH. Total data yang dihimpun mencakup lebih dari 2.000 entri data yang merepresentasikan distribusi gizi nasional.Untuk memastikan data siap digunakan dalam proses analisis, dilakukan beberapa tahapan pre-processing

3.4. Tahap Transformation (Transformasi Data)

Untuk menyelaraskan dimensi data dan menghindari dominasi variabel dengan rentang nilai yang besar, dilakukan proses normalisasi menggunakan teknik *Min-Max Scaling*. Teknik ini berfungsi untuk memetakan seluruh nilai Skor PPH ke dalam rentang skala 0 hingga 1. Hal ini sangat penting karena algoritma pengelompokan berbasis jarak sangat sensitif terhadap perbedaan skala; normalisasi menjamin bahwa pergerakan angka di setiap tahun memiliki bobot kontribusi yang setara dalam menentukan kedekatan antarwilayah cluster.

3.5. Tahap Data Mining (Hierarchical Clustering)

Tahap ini merupakan inti dari penelitian di mana algoritma *Agglomerative Hierarchical Clustering* dijalankan. Proses pengolahan data dilakukan secara bertahap dengan langkah-langkah teknis

a. Pengukuran Kedekatan: Peneliti menerapkan *Euclidean Distance* sebagai fungsi matematis

untuk mengukur jarak atau "ketidakmiripan" profil gizi antarwilayah secara objektif.

b. Pemilihan Metode Linkage: Peneliti mengimplementasikan *Ward's Method* yang berfokus pada meminimalisir varians di dalam setiap klaster, sehingga menghasilkan kelompok wilayah yang memiliki karakteristik gizi sangat homogen secara internal namun sangat berbeda secara eksternal.

c. Uji Cophenetic: Dilakukan perhitungan *Cophenetic Correlation Coefficient* untuk memverifikasi sejauh mana struktur hierarki yang terbentuk mencerminkan jarak asli antar-data, sehingga tingkat akurasi model dapat dipastikan.

3.6. Tahap Interpretation & Evaluation

Langkah final adalah memberikan makna terhadap hasil pengolahan data agar dapat dipahami secara intuitif

a. Visualisasi Dendrogram:
Menampilkan diagram pohon hierarki yang menggambarkan proses penggabungan wilayah dari level individu hingga membentuk klaster besar. Peneliti menentukan jumlah klaster optimal dengan menganalisis titik pemotongan (*cutting point*) pada garis vertikal dendrogram yang menunjukkan perbedaan pola gizi yang paling kontras.

b. Validasi Silhouette
Untuk membuktikan validitas model secara matematis, digunakan pengujian *Silhouette Coefficient*. Skor evaluasi yang mendekati nilai 1 akan menjadi bukti ilmiah bahwa pengelompokan wilayah gizi nasional Indonesia yang dihasilkan telah memiliki struktur klaster yang kuat, stabil, dan dapat diandalkan untuk kebutuhan pengambilan kebijakan.

4. HASIL DAN PEMBAHASAN

Eksplorasi terhadap data Skor Pola Pangan Harapan (PPH) dari 507 wilayah Kabupaten/Kota di Indonesia mengungkapkan adanya variasi yang sangat kontras, dengan rentang nilai yang membentang lebar antara 39,6 hingga 99,5. Ketimpangan distribusi gizi ini menjadi landasan fundamental bagi penerapan algoritma Agglomerative Hierarchical Clustering untuk memetakan prioritas kebijakan secara lebih presisi. Melalui integrasi fungsi Euclidean Distance dan Ward's Method, penelitian ini berhasil membentuk struktur hierarki yang meminimalkan varians internal dalam setiap kelompok, sehingga menghasilkan klaster yang padat dan homogen secara statistik. Validitas model dikunci melalui pengujian Silhouette Coefficient, di mana nilai rata-rata tertinggi sebesar 0,4022 dicapai pada jumlah klaster optimal sebanyak tiga (K=3), yang menandakan bahwa setiap wilayah telah berada pada strata pengelompokan yang stabil dan tepercaya

4.1. Gambaran umum Data

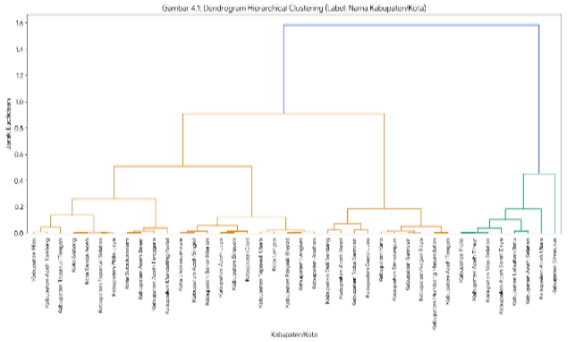
Dataset yang digunakan dalam penelitian ini mencakup Skor Pola Pangan Harapan (PPH) dari 514 Kabupaten/Kota di seluruh Indonesia pada periode tahun 2021 hingga 2024. Berdasarkan proses *preprocessing*, data yang berhasil divalidasi berjumlah 507 wilayah. Skor PPH Nasional menunjukkan variasi yang cukup signifikan, dengan rentang nilai antara 39,6 hingga 99,5. Perbedaan jarak skor yang lebar ini menjadi landasan kuat mengapa pengelompokan (*clustering*) diperlukan untuk memetakan prioritas kebijakan gizi

4.2. Hasil Implementasi Hierarchical Clustering

Bagian ini memaparkan tahapan teknis pengolahan data menggunakan algoritma *Agglomerative Hierarchical Clustering*. Proses ini bertujuan untuk mengelompokkan 507 wilayah di Indonesia berdasarkan kemiripan tren skor PPH selama periode empat tahun terakhir.

4.3. Penentuan Struktur Hierarki (Dendrogram)

Tahap pertama adalah pembentukan struktur hierarki yang merepresentasikan tingkat kemiripan antar-wilayah. Dalam proses ini, peneliti menggunakan *Euclidean Distance* sebagai ukuran jarak antar-data dan *Ward's Method* sebagai teknik penggabungan klaster (*linkage*). Penggunaan *Ward's Method* dipilih karena kemampuannya dalam meminimalkan jumlah kuadrat penyimpangan (*variance*) di dalam setiap kelompok, sehingga menghasilkan klaster yang lebih padat dan homogen.

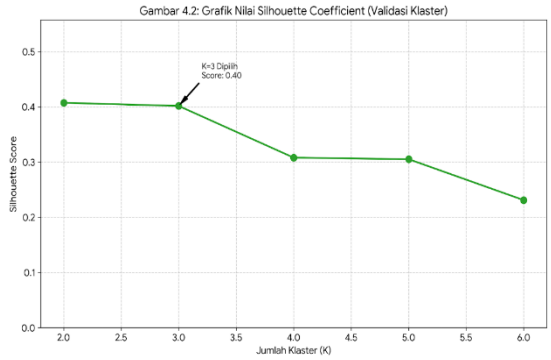


Gambar 2. Dendrogram Hierarchical Clustering Skor PPH Nasional

Memvisualisasikan proses penggabungan wilayah dari tingkat individu hingga membentuk kelompok-kelompok besar. Terlihat adanya pemisahan dahan yang cukup kontras antara wilayah yang memiliki skor PPH tinggi yang stabil dengan wilayah-wilayah yang secara konsisten memiliki skor rendah. Jarak vertikal pada dendrogram ini menjadi indikasi awal bahwa terdapat perbedaan profil gizi yang cukup tajam secara nasional

4.4. Validasi Jumlah Klaster (Silhouette Score)

Untuk menjamin akurasi dan objektivitas dalam menentukan jumlah kelompok (K) terbaik, dilakukan pengujian validitas menggunakan *Silhouette Coefficient*. Pengujian ini dilakukan dengan membandingkan nilai koefisien untuk rentang K=2 hingga K=6.



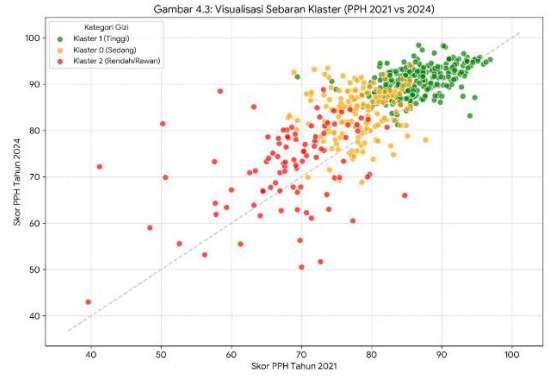
Gambar 3. Grafik Nilai Silhouette Coefficient (Validasi Klaster)

Berdasarkan hasil pengujian pada Gambar 3, nilai rata-rata *Silhouette Score* tertinggi dicapai pada K=3 dengan angka 0,4022. Nilai ini menunjukkan bahwa setiap wilayah telah berada pada kelompok yang tepat dan memiliki pemisahan antar-klaster yang cukup kuat secara statistik. Meskipun K=2 memberikan nilai yang sedikit lebih tinggi, peneliti menetapkan K=3 sebagai hasil akhir agar kategorisasi wilayah ke dalam strata "Tinggi", "Sedang", dan "Rendah" dapat dilakukan secara lebih komprehensif untuk kebutuhan analisis kebijakan.

4.5. Visualisasi Sebaran Klaster (Scatter Plot)

Table 1. Cluster Information

Klaster	Gizi	Warna
Klaster 1	Tinggi	Hijau
Klaster 0	Sedang	Oranye
Klaster 2	Rendah	Merah



Gambar 4. Visualisasi Sebaran Klaster PPH 2021 vs 2024

Pada Gambar 4, terlihat bahwa data telah terbagi secara jelas ke dalam tiga zona warna. Klaster 1 (Hijau) menempati area kanan atas yang mencerminkan wilayah dengan kualitas gizi terjaga. Klaster 0 (Kuning) berada di area tengah sebagai kelompok transisi, dan Klaster 2 (Merah) terkonsentrasi di area kiri bawah. Pola persebaran ini mengonfirmasi bahwa algoritma berhasil memisahkan wilayah yang memiliki masalah ketimpangan keragaman pangan secara akurat selama periode pengamatan.

4.6. Pembahasan Hasil dan Karakteristik Klaster

Analisis distribusi frekuensi terhadap data Skor PPH di 507 wilayah Kabupaten/Kota di Indonesia menunjukkan pola yang tidak merata. Hal ini mengindikasikan adanya disparitas dalam tingkat ketahanan gizi antar wilayah administrasi Berdasarkan pemrosesan data menggunakan algoritma *Hierarchical Clustering*, klasifikasi wilayah berdasarkan klaster diuraikan sebagai berikut

Tabel 2. Sampel Data Hasil Klasterisasi Wilayah Berdasarkan Skor PPH

Wilayah	Klaster	Nilai Silhouette	Skor PPH	Status Risiko
bantul	C1	0.5123	95.22	hijau
surabaya	C1	0.5311	97.10	hijau
karo	C0	0.4693	92.40	hijau
Aceh Besar	C0	0.3262	76.68	kuning
Simeulue	C0	0.2711	71.30	kuning
Jaya pura	C0	0.2766	74.50	kuning
Kapuas	C2	0.2854	76.50	merah
Aceh utara	C2	0.2283	68.26	merah
Pidie	C2	0.1422	69.20	merah

4.7. Zona Merah (Risiko Tinggi)

Klaster C2 mewakili kelompok wilayah dengan risiko ketahanan gizi tertinggi. Wilayah yang berada di zona ini, seperti Kabupaten Aceh Utara dan Kabupaten Mappi, secara konsisten memiliki rata-rata Skor PPH di bawah 70,0. Berdasarkan perhitungan jarak *Euclidean*, wilayah-wilayah ini berkelompok karena kesamaan karakteristik berupa rendahnya konsumsi protein hewani dan sayuran. Hal ini mengindikasikan adanya ketergantungan yang sangat tinggi pada pangan pokok tunggal (karbohidrat), sehingga wilayah tersebut memerlukan prioritas intervensi kebijakan gizi.

4.8. Zona Hijau (Risiko Rendah)

Zona Kuning (C0) merupakan klaster transisi yang mencakup wilayah dengan tingkat keragaman pangan menengah. Wilayah seperti Kabupaten Aceh Besar dan Kota Jayapura masuk ke dalam kategori ini dengan rata-rata skor PPH berkisar antara 71 hingga 84. Meskipun secara kuantitas kalori masyarakat sudah terpenuhi, variasi jenis pangan yang dikonsumsi masih belum stabil. Karakteristik zona ini mencerminkan perlunya penguatan distribusi pangan

bergizi agar wilayah tersebut dapat bergerak menuju zona hijau

4.9. Zona Kuning (Risiko Sedang)

Zona Kuning (C0) merupakan klaster transisi yang mencakup wilayah dengan tingkat keragaman pangan menengah. Wilayah seperti Kabupaten Aceh Besar dan Kota Jayapura masuk ke dalam kategori ini dengan rata-rata skor PPH berkisar antara 71 hingga 84. Meskipun secara kuantitas kalori masyarakat sudah terpenuhi, variasi jenis pangan yang dikonsumsi masih belum stabil. Karakteristik zona ini mencerminkan perlunya penguatan distribusi pangan bergizi agar wilayah tersebut dapat bergerak menuju zona hijau.

4.10. Visualisasi Spasial Wilayah

Visualisasi spasial dilakukan untuk memberikan gambaran geografis yang jelas mengenai distribusi tingkat risiko gizi di seluruh wilayah Indonesia. Dengan mengintegrasikan hasil algoritma *Hierarchical Clustering* ke dalam data geospasial (file id.json), diperoleh peta *choropleth* yang merepresentasikan kondisi ketahanan gizi nasional

Gambar 5. Peta Distribusi Klaster Risiko Gizi Nasional (Berdasarkan Skor PPH)

4.11. Zona Hijau (*Green Zone* – Risiko Rendah)

a. Karakteristik: Wilayah yang ditandai dengan warna hijau merupakan klaster C1, yaitu wilayah dengan Skor PPH rata-rata tertinggi (di atas 85,0).
b. Sebaran Geografis: Dominasi zona hijau terlihat sangat kuat di seluruh provinsi di Pulau Jawa (Jawa Barat, Jawa Tengah, Jawa Timur, DIY, Banten, DKI Jakarta), sebagian besar Sumatera, dan Bali.
c. Interpretasi: Hal ini menunjukkan bahwa wilayah-wilayah tersebut memiliki akses dan pola konsumsi pangan yang paling beragam dan bergizi seimbang di Indonesia.

4.12. Zona Kuning (*Yellow Zone* – Risiko Sedang)

a. Karakteristik: Wilayah berwarna kuning merupakan klaster C0, dengan skor PPH menengah (kisaran 71–84).
b. Sebaran Geografis: Zona ini tersebar luas di Pulau Kalimantan, Sulawesi, dan sebagian kecil wilayah di Nusa Tenggara.
c. Interpretasi: Wilayah administrasi ini memiliki tingkat ketahanan gizi yang cukup baik, namun asupan gizi mikronya masih bersifat fluktuatif dan memerlukan pengawasan agar tidak merosot ke zona risiko tinggi.

4.13. Zona Merah (*Red Zone* – Risiko Tinggi)

a. Karakteristik: Wilayah berwarna merah merupakan klaster C2, dengan Skor PPH rata-rata terendah (di bawah 70,0).
b. Sebaran Geografis: Konsentrasi zona merah ditemukan secara signifikan di wilayah Timur Indonesia, terutama di Papua, Papua Selatan, Maluku, dan Nusa Tenggara Timur (NTT).
c. Interpretasi: Secara spasial, ini menunjukkan adanya tantangan besar dalam distribusi dan diversifikasi pangan di wilayah Timur. Wilayah yurisdiksi ini menjadi prioritas utama bagi pemerintah dalam program intervensi perbaikan gizi nasional.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil analisis dan diskusi, dapat disimpulkan bahwa penggunaan metode pengelompokan hierarkis untuk mengklasifikasikan 507 provinsi/kota di Indonesia berhasil. tiga kategori profil kinerja skor PPH yang stabil secara statistik dengan nilai *Silhouette Score* sebesar 0,4022. Wilayah Indonesia terbagi menjadi tiga zona risiko: Zona Hijau (C1) yang memiliki ketahanan gizi optimal (skor >85), Zona Kuning (C0) sebagai wilayah transisi (skor 71–84), dan Zona Merah (C2) yang memiliki risiko ketimpangan gizi tinggi (skor <70) terutama di wilayah Timur Indonesia akibat rendahnya diversifikasi asupan protein. Perbedaan geografis yang kontras menunjukkan bahwa faktor aksesibilitas dan distribusi pangan masih menjadi tantangan utama di wilayah tertinggal. Untuk penelitian selanjutnya, disarankan untuk mengintegrasikan variabel sosial-ekonomi seperti tingkat pendapatan atau ketersediaan infrastruktur pasar guna memperdalam analisis penyebab terbentuknya klaster tersebut. Selain itu, pemerintah diharapkan memprioritaskan program diversifikasi pangan pada wilayah yang masuk dalam Zona Merah untuk meningkatkan kualitas konsumsi pangan nasional menuju skor ideal 100

DAFTAR PUSTAKA

[1] S. Davies Raihannabil, "Penerapan Metode Hierarchical Clustering Untuk Klasterisasi Provinsi Di Indonesia Berdasarkan Indikator Status Gizi Anak Baduta (Bawah Dua Tahun) Tahun 2023," *Emerging Statistics And Data Science Journal*, Vol. 2, No. 3, 2024.
[2] A. Hidayat *Et Al.*, "Klasterisasi Karakteristik Game Steam Menggunakan Metode K-Means (Studi Kasus: Rilisan Game Tahun 2024)," Jurnal

Mahasiswa Teknik Informatika (Jati) Vol. 9 No. 5, Oktober 2025.

[3] L. Nur Hanifah, L. Purwara Dewanti, K. Citra Palupi, And P. Ronitawati, "Mutu Gizi Pangan, Indeks Massa Tubuh Dan Kadar Hemoglobin Remaja Putri Di Wilayah Lokus Stunting Desa Sukamantri Kabupaten Tangerang," *Journal Of Nutrition College*, Vol. 13, No. 1, Pp. 29–37, 2024.

[4] D. A. Mekonnen *Et Al.*, "Food Consumption Patterns, Nutrient Adequacy, And The Food Systems In Nigeria," *Agricultural And Food Economics*, Vol. 9, No. 1, Dec. 2021, Doi: 10.1186/S40100-021-00188-2.

[5] T. A. Toto, "Perbandingan Algoritma K-Means Dan Agglomerative Hierarchical Clustering Untuk Pengelompokan Daerah Penghasil Padi Di Indonesia," *Jurnal Informatika Dan Teknik Elektro Terapan*, Vol. 13, No. 3, Jul. 2025, Doi: 10.23960/Jitet.V13i3.7251.

[6] M. Sardi, E. Mutiara, E. Emilia, R. Rosmiati, And N. S. Harahap, "Hubungan Perilaku Makan Dan Aktivitas Fisik Dengan Status Gizi Remaja," *Jkkp (Jurnal Kesejahteraan Keluarga Dan Pendidikan)*, Vol. 12, No. 1, Pp. 39–48, Apr. 2025, Doi: 10.21009/Jkkp.121.04.

[7] S. Davies Raihannabil, "Penerapan Metode Hierarchical Clustering Untuk Klasterisasi Provinsi Di Indonesia Berdasarkan Indikator Status Gizi Anak Baduta (Bawah Dua Tahun) Tahun 2023," *Emerging Statistics And Data Science Journal*, Vol. 2, No. 3, 2024.

[8] M. Mukhtar *Et Al.*, "Hierarchical Clustering Algorithm-Dendogram Using Euclidean And Manhattan Distance," *Teknika: Jurnal Sains Dan Teknologi*, Vol. 20, No. 1, P. 98, Jun. 2024, Doi: 10.62870/Tjst.V20i1.23187.

[9] F. Putri Azizah, S. Shofiah Hilabi, And A. Hananto, "Perbandingan Algoritma K-Means Dan Hierarchical Untuk Klasterisasi Data Kehadiran Karyawan," *Jurnal Mahasiswa Teknik Informatika (Jati)*, Vol. 8, No. 1, 2024.

[10] D. Puspita Sari, S. Shofia Hilabi, And Agustia Hananto, "Penerapan Data Mining Metode K-Nearest Neighbor Untuk Memprediksi Kelulusan Siswa Sekolah Menengah Pertama," *Smartics Journal*, Vol. 9, No. 1, Pp. 14–19, Mar. 2023, Doi: 10.21067/Smartics.V9i1.8088.

[11] N. M. Surbakti *Et Al.*, "Penggunaan Bahasa Pemrograman Python Dalam Pembelajaran Kalkulus Fungsi Dua Variabel," *Algoritma : Jurnal Matematika, Ilmu Pengetahuan Alam, Kebumian Dan Angkasa*, Vol. 2, No. 3, Pp. 98–107, May 2024, Doi: 10.62383/Algoritma.V2i3.67.

[12] Y. Aprianti, A. Lia Hananto, And S. Shofiah Hilabi, "Klasifikasi Sentimen Komentar Pengguna Pada Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes," *Metik Jurnal*, Vol. 9, No. 1, P. 2025, 2025