



PERBANDINGAN UKURAN JARAK PADA ANALISIS KLUSTER HIRARKI

**Muh. Zarkawi Yahya<sup>1,*), Sitti Nurhaliza<sup>2), Morina A Fathan<sup>3),
Muhammad Edy Rizal^{4), Andi Harismahyanti⁵⁾}</sup></sup></sup>**

^{1,2,3,4,5)} Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan
Alam, Universitas Tadulako
*email: zarkawi2206@gmail.com

Abstrak: Analisis klaster merupakan salah satu metode statistik untuk mengelompokkan objek berdasarkan kemiripan. Pada data kategorik, pemilihan ukuran jarak menjadi aspek penting karena memengaruhi struktur dan interpretasi klaster yang terbentuk. Penelitian ini bertujuan untuk membandingkan performa enam ukuran jarak Gower, Goodall1, Goodall2, Goodall3, Goodall4, dan Anderberg dalam analisis klaster hierarki menggunakan data kategorik dari Indonesian Family Life Survey (IFLS-5). Metode yang digunakan adalah hierarchical agglomerative clustering, dengan tahap awal pembersihan data dan konversi ke tipe faktor agar sesuai dengan karakteristik pengukuran jarak kategorik. Evaluasi hasil klaster dilakukan dengan dua indeks validasi internal, yaitu Silhouette dan Dunn, serta metrik eksternal Adjusted Rand Index (ARI) untuk menilai stabilitas klaster melalui proses bootstrapping. Ketiga metrik tersebut digunakan secara komplementer: Silhouette mengevaluasi konsistensi lokal anggota klaster (dengan nilai ≥ 0.5 umumnya dianggap baik), Dunn mengukur pemisahan antar-klaster secara global (semakin tinggi semakin baik), sementara ARI menunjukkan konsistensi struktur klaster terhadap variasi data (nilai mendekati 1 menunjukkan stabilitas tinggi). Hasil menunjukkan bahwa setiap ukuran jarak menghasilkan struktur klaster yang berbeda. Di antara semua ukuran yang diuji, Goodall4 memberikan hasil terbaik karena membentuk klaster yang mudah diinterpretasikan, memiliki nilai indeks Silhouette dan Dunn yang relatif tinggi, serta skor ARI mendekati sempurna. Hal ini mengindikasikan bahwa Goodall4 merupakan alternatif yang layak direkomendasikan dalam kasus serupa.

Kata Kunci: Analisis Klaster; Data Kategorik; IFLS 5; Hierarchical Clustering; Ukuran Jarak

Abstract: Cluster analysis is a statistical method used to group objects based on their similarity. In the context of categorical data, the choice of distance measure is crucial, as it directly affects the structure and interpretability of the resulting clusters. This study compares the performance of six distance measures Gower, Goodall1, Goodall2, Goodall3, Goodall4, and Anderberg in hierarchical clustering using categorical data from the Indonesian Family Life Survey (IFLS-5). The method employed is hierarchical agglomerative clustering, preceded by data cleaning and variable transformation into factor types to accommodate categorical distance calculations. Cluster evaluation was conducted using two internal validation indices, Silhouette and Dunn, along with an external metric, the Adjusted Rand Index (ARI), to assess clustering stability via bootstrapping. These three metrics are used complementarily: Silhouette evaluates local cohesion within clusters (with values ≥ 0.5 generally considered good), Dunn measures global separation between clusters



(higher values indicating better separation), and ARI assesses the consistency of cluster structures under varying data samples (values close to 1 indicate high stability). The results show that each distance measure yields distinct cluster structures. Among them, Goodall4 produced the most interpretable clusters, with relatively high Silhouette and Dunn scores and consistently near-perfect ARI values. These findings suggest that Goodall4 is a highly effective and reliable alternative for hierarchical clustering tasks involving categorical data.

Keywords: Cluster Analysis; Categorical Data; IFLS 5; Hierarchical Clustering; Distance Measure

PENDAHULUAN

Analisis klaster adalah metode statistik yang digunakan untuk mengelompokkan sekumpulan objek atau individu ke dalam beberapa kelompok (klaster) berdasarkan kemiripan karakteristik di antara mereka. Tujuannya adalah agar objek dalam satu klaster memiliki kesamaan yang tinggi, sementara objek antar klaster menunjukkan perbedaan yang signifikan. Pendekatan ini pertama kali diperkenalkan oleh Tyron pada tahun 1939, yang mengembangkan teknik ini untuk mengelompokkan individu ke dalam kelompok yang saling independen (Johnson, 1982).

Analisis klaster telah diterapkan secara luas di berbagai disiplin ilmu, termasuk kesehatan, pemasaran, pendidikan, dan ilmu sosial. Penerapan ini tercermin dalam sejumlah studi terkini seperti yang dilakukan oleh (Pansris, 2024), Gheorghe (2023), dan Laksono (2024). Dalam konteks kesehatan, metode klaster digunakan untuk mengidentifikasi pola distribusi penyakit, mengelompokkan pasien berdasarkan karakteristik klinis atau riwayat medis, serta mendukung pengambilan keputusan berbasis data. Studi oleh Wala (2024) dan Mayola (2025) menunjukkan bagaimana pendekatan ini dapat digunakan untuk merumuskan kebijakan kesehatan yang lebih tepat sasaran dan berbasis bukti.

Terdapat dua metode dalam analisis klaster, yaitu metode hierarki dan non-hierarki. Metode hierarki adalah metode yang dirancang untuk berhierarki (tingkatan) dari kumpulan data menurut karakteristik pada data tersebut. Sedangkan metode non-hierarki adalah metode yang digunakan untuk mengelompokkan data dan jumlah klaster yang akan dibuat dapat ditetapkan sebelumnya. Ada berbagai jenis analisis klaster yang digunakan dalam metode hierarki, salah satu yang paling umum digunakan adalah metode *complete linkage*. Metode ini cenderung menghindari penggabungan klaster yang anggotanya memiliki jarak yang sangat jauh secara ekstrem, sehingga klaster yang terbentuk lebih bulat dan homogen (Alamtaha et al., 2023).

Ukuran jarak atau similaritas maupun disimilaritas mengambil peran penting dalam analisis klaster karena menjadi dasar dalam pengukuran kemiripan antar objek. Sehingga, penentuan metode pengukuran jarak akan mempengaruhi performa dan hasil klaster (Aditya et al., 2021) (Johnson, 1982). Semakin kecil jarak antara satu individu dengan individu lainnya, maka tingkat kemiripan mereka semakin tinggi, sehingga individu tersebut cenderung dikelompokkan ke dalam klaster yang sama (Prabowo et al., 2020).



Ukuran similaritas seperti Euclidean dan Manhattan umum digunakan untuk data kuantitatif (Sha et al., 2023), namun tidak cocok untuk data kategorikal karena tidak ada cara universal untuk mengukur jarak antar kategori (Rios et al., 2021). Maka, mustahil untuk menentukan jarak antara dua kategori variabel secara langsung. Pendekatan yang paling umum saat ini adalah mengubah variabel tersebut menjadi variabel biner (dummy) dan kemudian menggunakan ukuran kesamaan yang sesuai untuk variabel biner. Namun, transformasi ke bentuk biner seringkali dapat menghilangkan informasi penting yang terkandung dalam variabel asli, hal ini telah ditunjukkan pada penelitian Sulc & Řezanková (2019) sehingga penggunaannya harus dilakukan dengan hati-hati. Dinh et al. (2024) secara eksplisit juga menyebutkan bahwa encoding biner menghilangkan informasi semantik dan menyebabkan masalah multi-dissimilarity, sehingga clustering kurang optimal. Berbagai ukuran similaritas untuk data kategorik telah dikembangkan, termasuk metode-metode seperti Goodall1 hingga Goodall4, serta Anderberg (Sulc & Řezanková, 2019). Sedangkan untuk data campuran jarak gower dikembangkan dengan metrik yang berbeda dalam menghitung jarak atribut bertipe numerik dan atribut bertipe kategorik (Fadilah & Wijayanto, 2023).

Berbagai studi telah dilakukan untuk membandingkan efektivitas ukuran jarak dalam analisis klaster, khususnya pada data numerik. Misalnya Mukhtar et al. (2024) mengevaluasi penggunaan jarak Euclidean dan Manhattan dalam algoritma klaster hierarkis, sementara Aditya et al. (2021) membandingkan jarak Euclidean dan Gower dalam konteks algoritma k-medoids. Selain itu Fadilah & Wijayanto (2023) meneliti performa One-Hot Encoding, jarak Gower, dan algoritma K-Prototype pada data campuran. Untuk data kategorikal Sulc & Řezanková (2019) telah membandingkan beberapa ukuran similaritas seperti Goodall1 hingga Goodall4 dalam analisis klaster hierarkis. Namun, penelitian tersebut belum mencakup ukuran jarak Gower dan Anderberg.

Penelitian ini akan mengevaluasi dan membandingkan hasil klaster dari berbagai metode pengukuran jarak dalam analisis klaster hierarki, yaitu Gower, Goodall1 hingga Goodall4, serta Anderberg, pada data kategorikal dengan studi kasus pada variabel-variabel kesehatan dari data IFLS-5. Tujuan dari penelitian ini adalah untuk mengidentifikasi metode pengukuran jarak yang paling efektif dalam menghasilkan struktur klaster yang representatif dan bermakna pada data kategorikal di bidang kesehatan. Melalui evaluasi komparatif antar metode, penelitian ini diharapkan dapat memberikan kontribusi dalam pemilihan metode jarak yang lebih tepat untuk analisis klaster hierarki pada data kategorikal, khususnya dalam konteks penelitian kesehatan.

METODE PENELITIAN

Penelitian ini merupakan studi kuantitatif eksploratif yang bertujuan untuk membandingkan performa berbagai ukuran jarak dalam analisis klaster hierarki terhadap data kategorik. Data yang digunakan adalah *Indonesian Family Life Survey* (IFLS) gelombang ke-5, yang bersumber dari RAND Corporation dan dapat diakses melalui <https://www.rand.org/>. Dataset awal terdiri atas 9.388 entri, namun setelah proses *cleaning* (menghapus entri dengan



nilai kosong pada variabel yang dianalisis) jumlah data tersisa sebanyak 2.060 entri valid. Atribut yang digunakan pada penelitian ini sebanyak tujuh atribut bertipe kategorik: status merokok, umur (dalam kelompok usia), jenis kelamin, tingkat ekonomi, serta riwayat hipertensi, diabetes, dan stroke.

Pra-pemrosesan data dimulai dengan semua atribut pada dataset dikodekan dalam bentuk numerik, sehingga dikonversi ke dalam bentuk faktor (kategori) untuk memungkinkan penggunaan ukuran jarak kategorik secara sah (Spada et al., 2025). Proses ini krusial karena metrik jarak seperti Goodall dan Gower sangat dipengaruhi oleh distribusi frekuensi kategori (Adha et al., 2025). Tahapan kedua adalah menghitung matriks jarak $D = [d_{ij}]$, di mana d_{ij} menunjukkan jarak antara pasangan objek x_i dan x_j . Tujuh jenis ukuran jarak digunakan sebagai berikut (Šulc, 2015):

1. Jarak Anderberg adalah ukuran jarak dalam menghitung dissimilaritas dengan mempertimbangkan jumlah kesamaan dan perbedaan antar kategori. Metode ini menekankan proporsi kesamaan terhadap total atribut yang dibandingkan. Semakin besar nilai a , semakin kecil jarak (lebih mirip).

$$d_{ij}^A = 1 - \frac{a}{(a + b + c)} \quad (1)$$

2. Jarak Goodall 1 adalah ukuran jarak memperhitungkan frekuensi relatif kategori, dimana G1 memberikan bobot pada frekuensi lebih rendah.

$$S_c(x_{ic}, x_{jc}) = \begin{cases} 1 - \sum_{q \in Q} p_c^2(q) ; x_{ic} = x_{jc}, Q \subseteq X_c; \forall q \in Q, p_c(q) \leq p_c(x_{ic}) \\ 0 ; \text{lainnya} \end{cases} \quad (2)$$

Dimana q merepresentasikan subset kategori pada variabk ke-c yang frekuensi relativanya lebih rendah atau sama dengan $p(x_{ic})$.

3. Jarak Goodall 2 adalah ukuran jarak memperhitungkan frekuensi relatif kategori, dimana G2 memberikan bobot pada frekuensi lebih tinggi atau sama.

$$S_c(x_{ic}, x_{jc}) = \begin{cases} 1 - \sum_{q \in Q} p_c^2(q) ; x_{ic} = x_{jc}, Q \subseteq X_c; \forall q \in Q, p_c(q) \geq p_c(x_{ic}) \\ 0 ; \text{lainnya} \end{cases} \quad (3)$$

4. Jarak Goodall 3, digunakan menghitung kesamaan dengan menekankan kategori langka. Semakin langka kategori tersebut, semakin tinggi nilai kesamaannya.

$$S_c(x_{ic}, x_{jc}) = \begin{cases} 1 - p_c^2(x_{ic}) ; x_{ic} = x_{jc} \\ 0 ; \text{lainnya} \end{cases} \quad (4)$$

5. Jarak Goodall 4, digunakan menghitung kesamaan dengan menekankan kategori umum.

$$S_c(x_{ic}, x_{jc}) = \begin{cases} p_c^2(x_{ic}) ; x_{ic} = x_{jc} \\ 0 ; \text{lainnya} \end{cases} \quad (5)$$

6. Jarak Gower, digunakan untuk data ordinal dan campuran, menghitung kesamaan kategori dengan skema penyesuaian sederhana.

$$d_{ij}^{Gower} = \left(\frac{1}{p}\right) \sum S_{ijk}, k = 1..p \quad (6)$$



dimana $s_{ijk} = 0$ jika $x_{ik} = x_{jk}$, 1 jika $x_{ik} \neq x_{jk}$

Berdasarkan perbedaan distribusi kumulatif kategori; digunakan untuk data ordinal. Setelah semua jarak dihitung, dilakukan pengklusteran hirarki dengan metode agglomerative. Proses ini membentuk dendrogram dan diuji untuk jumlah kluster $k = 2$ hingga $k = 10$. Rentang ini dipilih karena $k < 2$ tidak memungkinkan terbentuknya klaster, sementara $k > 10$ cenderung menghasilkan klaster kecil yang sulit diinterpretasikan secara sosial dan kesehatan. Batas ini ditetapkan untuk menjaga keseimbangan antara kompleksitas model dan kemudahan interpretasi.

Tedapat berbagai jenis metrik yang digunakan dalam evaluasi kualitas klustering, yaitu metrik validasi internal dan validasi eksternal. Metrik validasi internal yang digunakan yaitu sebagai berikut:

1. Silhouette Index, mengukur koherensi lokal antar anggota klaster.

$$s(i) = \frac{(b(i) - a(i))}{\max(a(i), b(i))} \quad (4)$$

2. Dunn Index, mengukur pemisahan antar klaster dan kepadatan dalam klaster.

$$\frac{D = \min \delta(c_i, c_j)}{\max \Delta(C_l)} \quad (5)$$

$\delta(c_i, c_j)$: jarak minimum antar kluster

$\Delta(C_l)$: diameter maksimum kluster

Silhouette Score dan Dunn Index merupakan dua metrik evaluasi internal yang sering digunakan dalam analisis klaster karena masing-masing menawarkan perspektif yang berbeda terhadap kualitas hasil klasterisasi. Silhouette Score mengukur konsistensi lokal dengan menilai kedekatan suatu data terhadap klaster tempatnya berada dibandingkan dengan klaster lain. Skor ini memiliki rentang nilai antara -1 hingga 1, di mana nilai yang lebih tinggi menunjukkan bahwa data lebih sesuai dengan klasternya. Keunggulan utama dari metrik ini adalah kestabilannya terhadap variasi ukuran klaster serta kemampuannya dalam menggambarkan kejelasan batas antar-klaster di sekitar tiap titik data. Namun demikian, Silhouette Score memiliki keterbatasan dalam menangkap struktur klaster yang kompleks atau saling tumpang tindih, karena hanya mempertimbangkan kedekatan lokal tanpa memperhitungkan distribusi global.

Dunn Index menilai kualitas klaster secara global dengan menghitung rasio antara jarak minimum antar-klaster dan diameter maksimum dalam-klaster. Nilai Dunn yang tinggi mengindikasikan bahwa klaster yang terbentuk bersifat kompak dan saling terpisah dengan baik. Meskipun demikian, metrik ini cukup sensitif terhadap variasi ukuran dan bentuk klaster, terutama ketika terdapat ketidakseimbangan jumlah anggota antar-klaster. Keberadaan klaster yang sangat kecil atau sangat tersebar dapat menyebabkan nilai Dunn menurun secara signifikan atau meningkat secara tidak proporsional, sehingga interpretasi hasil perlu dilakukan dengan hati-hati.

Beberapa metrik validasi eksternal yang sering digunakan adalah Adjusted Rand Indek (ARI), Normalized Mutual Information (NMI), dan Homogeneity Score (HS). Evaluasi stabilitas klasterisasi dilakukan dengan menggunakan Adjusted Rand Index (ARI) sebagai metrik untuk



mengukur konsistensi keanggotaan klaster terhadap variasi data. Meskipun metrik seperti Normalized Mutual Information (NMI) dan Homogeneity Score sering digunakan dalam evaluasi klasterisasi, keduanya memiliki keterbatasan ketika diterapkan pada data kategorik. NMI, yang berbasis pada teori informasi, mengukur kesamaan antara dua partisi dengan mempertimbangkan distribusi probabilistik dari label, sehingga lebih sesuai untuk data numerik atau teks yang memiliki struktur distribusi yang eksplisit. Namun, dalam konteks data kategorik yang tidak selalu memiliki distribusi probabilistik yang jelas, NMI dapat menunjukkan bias terhadap jumlah klaster dan menghasilkan interpretasi yang kurang stabil. Sebaliknya, Adjusted Rand Index (ARI) menawarkan keunggulan dalam mengevaluasi kesesuaian antara hasil klasterisasi dan label referensi dengan mengoreksi kemungkinan kesesuaian acak. ARI lebih robust terhadap variasi jumlah klaster dan memberikan interpretasi yang lebih konsisten, menjadikannya pilihan yang lebih tepat untuk data kategorik dalam studi evaluasi performa klasterisasi (Warrens & Hoef, 2022).

ARI digunakan untuk membandingkan hasil klasterisasi pada data lengkap dan hasil klasterisasi pada subset data lengkap. Nilai ARI berkisar antara -1 hingga 1, di mana nilai yang lebih tinggi menunjukkan stabilitas dan konsistensi struktur klaster. Dalam penelitian ini, proses bootstrapping dilakukan sebanyak 100 kali untuk masing-masing metode ukuran jarak (Gower, Goodall1–4, dan Anderberg), pada tiga proporsi data pelatihan: 60%, 70%, dan 80%. Penggunaan tiga proporsi data pelatihan, yaitu 60%, 70%, dan 80%, dimaksudkan untuk menguji stabilitas dan ketahanan metode klasterisasi terhadap variasi ukuran data yang tersedia. Dengan mengevaluasi performa metode pada beberapa proporsi data, peneliti dapat menilai apakah hasil klaster tetap konsisten meskipun jumlah data yang dianalisis berbeda. Ketiga proporsi tersebut dipilih untuk mewakili kondisi data terbatas (60%), sedang (70%), dan relatif lengkap (80%), sehingga memungkinkan simulasi berbagai skenario riil yang mungkin terjadi di lapangan. Pendekatan ini juga membantu memastikan bahwa metode yang digunakan tidak hanya efektif dalam kondisi ideal, tetapi juga cukup andal dalam menghadapi fluktuasi ukuran sampel yang umum dalam penelitian terapan. Proses analisis data pada penelitian ini menggunakan aplikasi statistik Rstudio dengan package cluster nomclust, dan tidyverse.

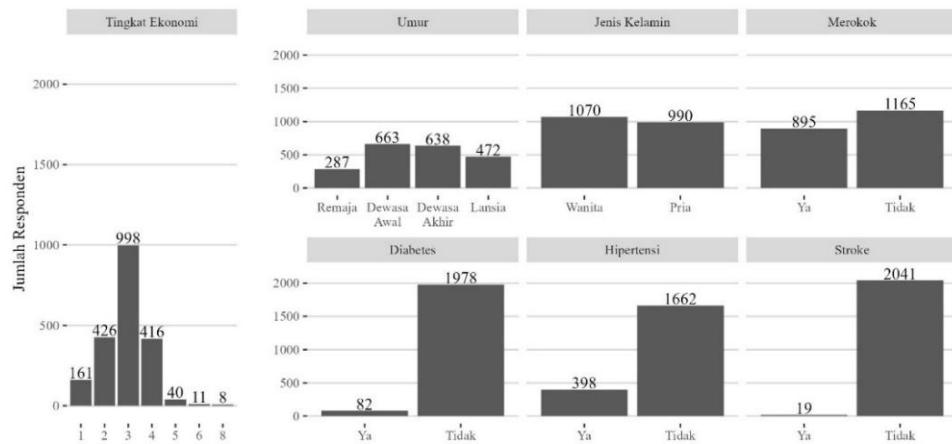
HASIL DAN PEMBAHASAN

Hasil Penelitian

Eksplorasi Data

Penelitian dimulai dengan melakukan eksplorasi data guna memahami karakteristik data yang akan diklusterkan.

Gambar 1 berikut menunjukkan data bersih yang terdiri atas 2060 responden.



Gambar 1. Karakteristik data kategorik yang dianalisis

Pada

Gambar 1, peubah “Tingkat Ekonomi” merupakan peubah ordinal, di mana kelas 1 berarti “sangat kurang mampu” dan kelas 8 berarti “sangat mampu”. Terdapat 4 kelas pada peubah “umur”, di mana “remaja” didefinisikan sebagai responden pada rentang umur 15 – 20 tahun, “dewasa awal” pada rentang 21 – 35 tahun, “dewasa akhir” pada rentang 36 – 50 tahun, dan “lansia” bagi responden dengan umur lebih dari 50 tahun.

Secara umum,

Gambar 1 menunjukkan bahwa mayoritas responden berasal dari keluarga dengan tingkat ekonomi menengah, dengan mayoritas berasal dari rentang usia dewasa dan lansia. Jumlah responden wanita dan pria relatif sama dengan jumlah perokok yang relatif sama. Sedangkan dari ketiga peubah penyakit, terlihat bahwa mayoritas responden tidak memiliki penyakit diabetes, hipertensi, atau stroke.

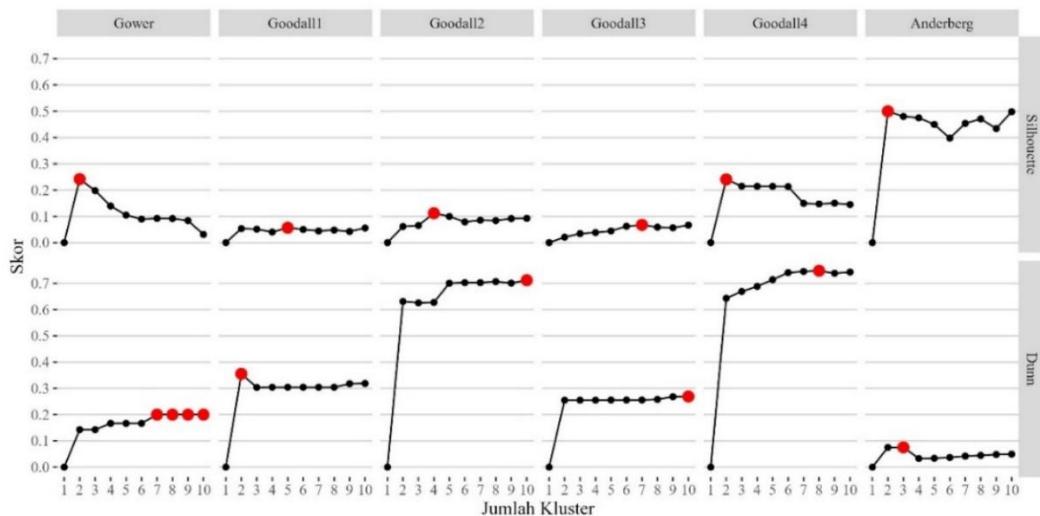
Ketujuh peubah dalam

Gambar 1 merupakan variabel kategorik, sehingga metode pengklasteran seperti K-Means dan K-Median tidak dapat digunakan, sebab menggunakan ukuran jarak numerik seperti Euclidean dan Manhattan. Oleh karena itu, pengklasteran pada penelitian ini menggunakan metode pengklasteran hierarki dengan ukuran jarak khusus data kategorik, yaitu ukuran jarak Gower, Goodall 1, 2, 3, dan 4, serta Anderberg.

Penentuan Jumlah Kluster (*k*) Optimal

Pengklasteran dimulai dengan menentukan jumlah kluster optimal untuk masing-masing matriks jarak, yaitu dengan membandingkan skor Silhouette dan Dunn untuk jumlah kluster $k = 2$ hingga $k = 10$.

Gambar 2 berikut ini menunjukkan perbandingan skor Silhouette dan Dunn dari keenam ukuran jarak yang digunakan di mana titik merah menunjukkan nilai maksimum dari masing-masing sub plot.



Gambar 2. Skor Silhouette dan Dunn bagi masing-masing ukuran jarak

Perbedaan orientasi evaluasi antara Silhouette dan Dunn tercermin dari pola hasil yang dihasilkan masing-masing ukuran jarak. Misalnya, Goodall2 dan Goodall4 menghasilkan skor Dunn tinggi namun Silhouette rendah, yang mengindikasikan bahwa kluster-kluster yang terbentuk terpisah secara global, tetapi terdapat inkonsistensi lokal berupa objek-objek yang kurang sesuai dengan klusternya. Sebaliknya, ukuran jarak Anderberg menunjukkan Silhouette tinggi namun Dunn rendah, menandakan bahwa objek-objek berada dalam kluster yang sesuai, tetapi kluster-kluster tersebut saling berdekatan dan tidak membentuk pemisahan yang tegas.

Dengan mempertimbangkan perbedaan ini, pemilihan jumlah kluster optimal dalam penelitian ini tidak hanya didasarkan pada satu indeks semata. Silhouette dan Dunn digunakan secara bersamaan untuk menyeimbangkan evaluasi lokal dan global, dan keputusan akhir juga mempertimbangkan aspek substantif, seperti kemudahan interpretasi hasil kluster. Sebagai contoh, pada ukuran jarak Gower, skor Silhouette tertinggi diperoleh pada $k = 2$, sedangkan skor Dunn cenderung stabil tinggi mulai dari $k = 7$ hingga $k = 10$. Dalam hal ini, dipilih $k = 2$ karena penurunan Silhouette setelah $k = 2$ cukup drastis, sedangkan penurunan Dunn dari $k = 7$ ke $k = 2$ tidak terlalu signifikan. Selain itu, jumlah kluster yang lebih kecil umumnya lebih mudah untuk ditafsirkan, khususnya ketika digunakan sebagai dasar segmentasi atau rekomendasi kebijakan.

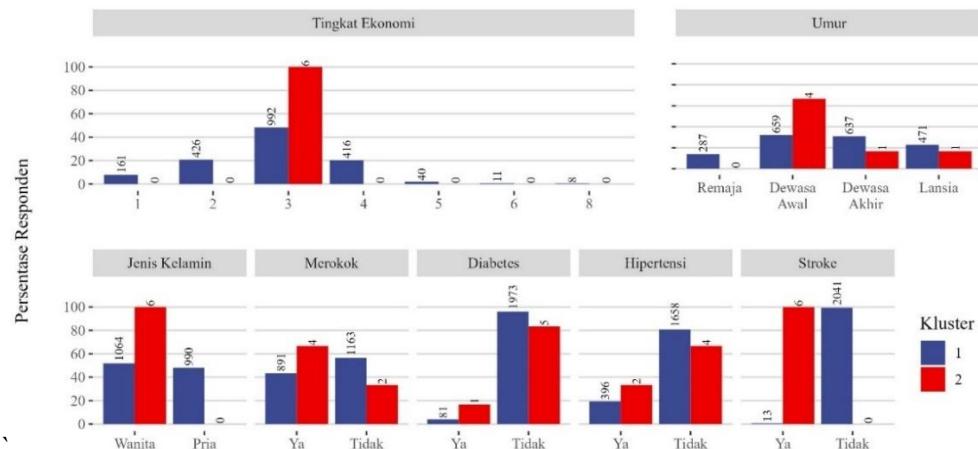
Strategi pemilihan k optimal untuk ukuran jarak lainnya juga dilakukan dengan pendekatan serupa, yaitu dengan mempertimbangkan bentuk grafik Silhouette dan Dunn secara bersamaan, serta memperhatikan konteks interpretasi hasil kluster. Tabel 1 merangkum jumlah kluster optimal yang dipilih untuk masing-masing ukuran jarak.

Tabel 1. Jumlah kluster (k) optimal

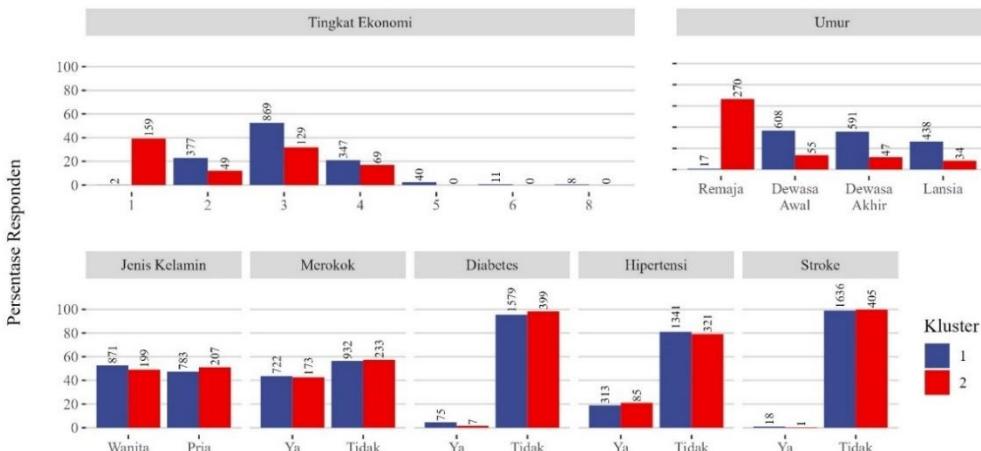
Ukuran Jarak	Gower	Goodall1	Goodall2	Goodall3	Goodall4	Anderberg
k optimal	2	2	5	7	3	2

Profiling Hasil Kluster

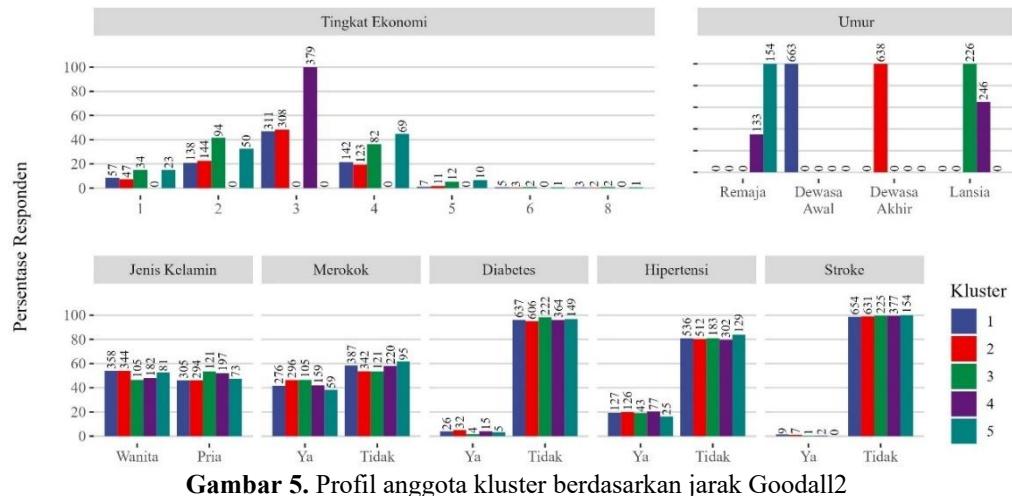
Setelah jumlah kluster optimal dari masing-masing ukuran jarak diperoleh, proses selanjutnya adalah proses pengklasteran dengan menggunakan k optimal pada Tabel 1. Gambar 3 hingga Gambar 8 menunjukkan profil keanggotaan dari keenam hasil kluster hierarki. Profil disajikan sebagai persentase responden (%) dari masing-masing kelas dan peubah, sementara angka di atas masing-masing bar adalah jumlah responden sebagai tambahan konteks.



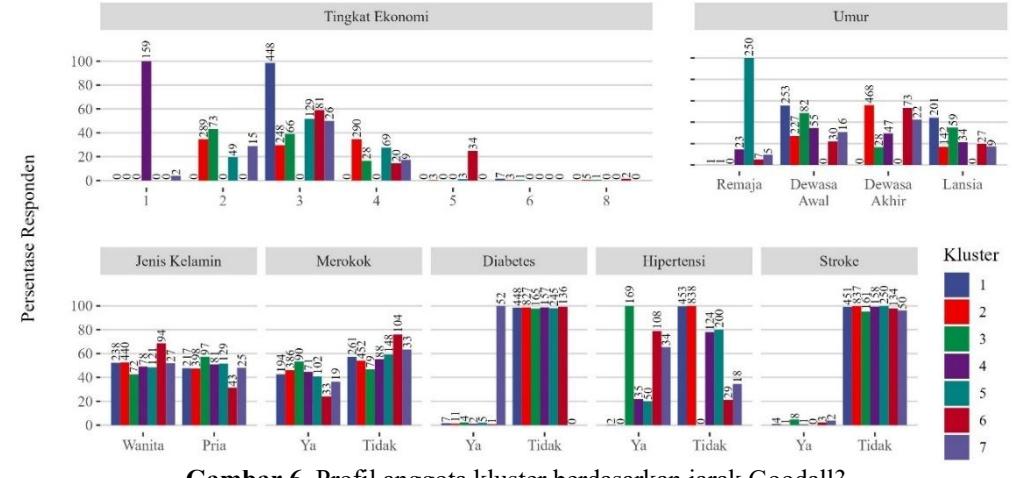
Gambar 3. Profil anggota kluster berdasarkan jarak Gower



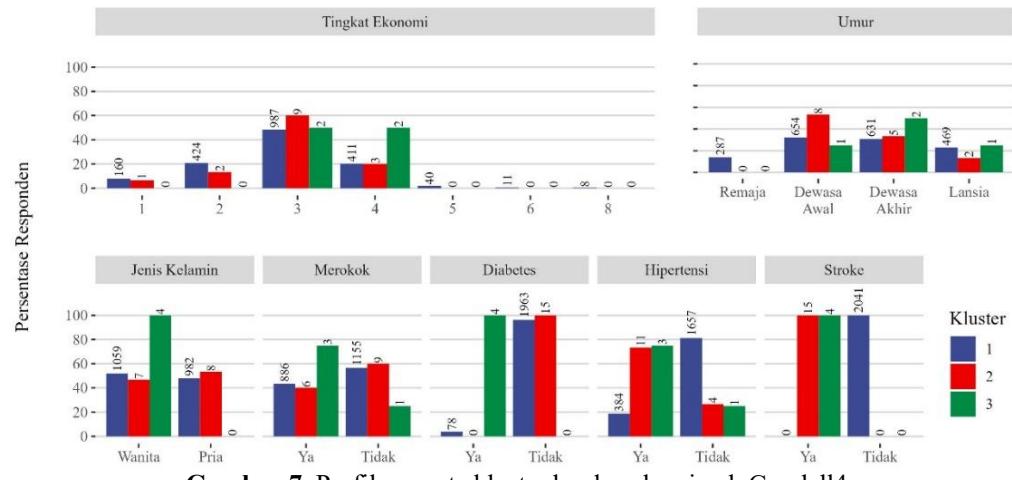
Gambar 4. Profil anggota kluster berdasarkan jarak Goodall1



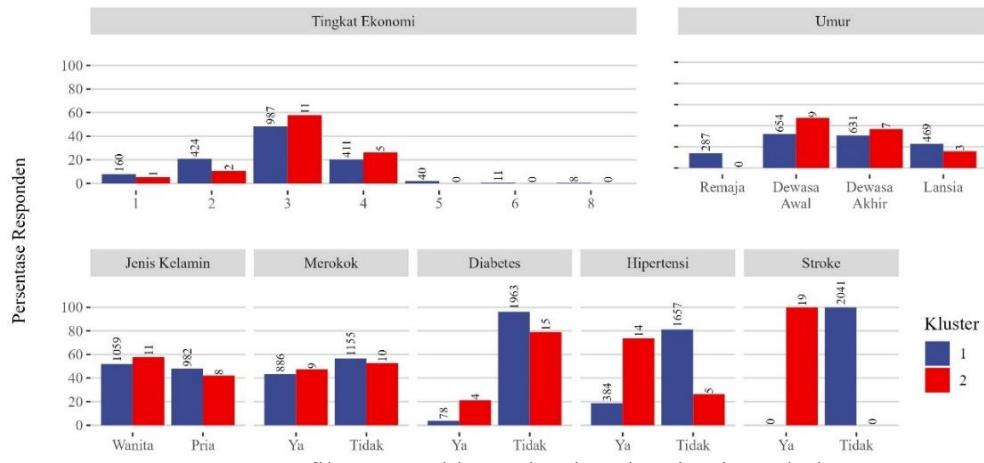
Gambar 5. Profil anggota kluster berdasarkan jarak Goodall2



Gambar 6. Profil anggota kluster berdasarkan jarak Goodall3



Gambar 7. Profil anggota kluster berdasarkan jarak Goodall4



Gambar 8. Profil anggota kluster berdasarkan jarak Anderberg

Gambar 3 hingga Gambar 8 menunjukkan profil keanggotaan yang berbeda-beda di antara keenam ukuran jarak. Dapat dilihat bahwa masing-masing hasil kluster didasarkan pada peubah diskriminatif yang berbeda-beda. Tabel 2 berikut meringkas peubah diskriminatif dan interpretasi dari hasil kluster yang ditunjukkan pada Gambar 3 hingga Gambar 8.

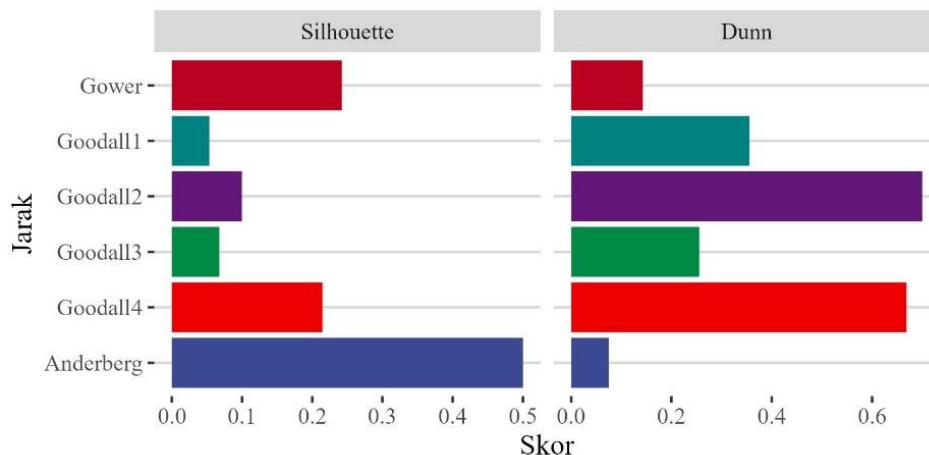
Tabel 2. Peubah diskriminatif dan ringkasan interpretasi dari masing-masing ukuran jarak

Ukuran Jarak	Peubah Diskriminatif	k	Ringkasan Hasil Kluster
Gower	Stroke, Tingkat Ekonomi	2	Enam responden wanita dengan riwayat stroke dan dengan tingkat ekonomi 3 membentuk satu kluster pencilan.
Goodall1	Umur	2	Responden pada kluster 2 cenderung lebih tua dibandingkan responden pada kluster 1, tetapi kedua kluster hampir tidak dapat dipisahkan.
Goodall2	Umur	5	Responden dibagi menjadi 5 kluster dengan pembagian kelompok umur yang lebih jelas dibandingkan Goodall1.
Goodall3	Umur, Tingkat Ekonomi, Diabetes, Hipertensi, Stroke	7	Terbentuk 7 kluster yang perbedaannya tidak cukup jelas satu sama lain.
Goodall4	Stroke, Diabetes	3	Terbentuk tiga kluster: responden tanpa riwayat stroke, responden dengan riwayat stroke tanpa diabetes, serta responden dengan riwayat stroke dan diabetes.
Anderberg	Stroke	2	Terbentuk dua kluster dengan 19 responden yang memiliki riwayat stroke membentuk satu kluster pencilan. Mirip dengan hasil kluster Gower, tetapi dengan definisi kluster yang relatif lebih jelas.

Metode Klustering Terbaik

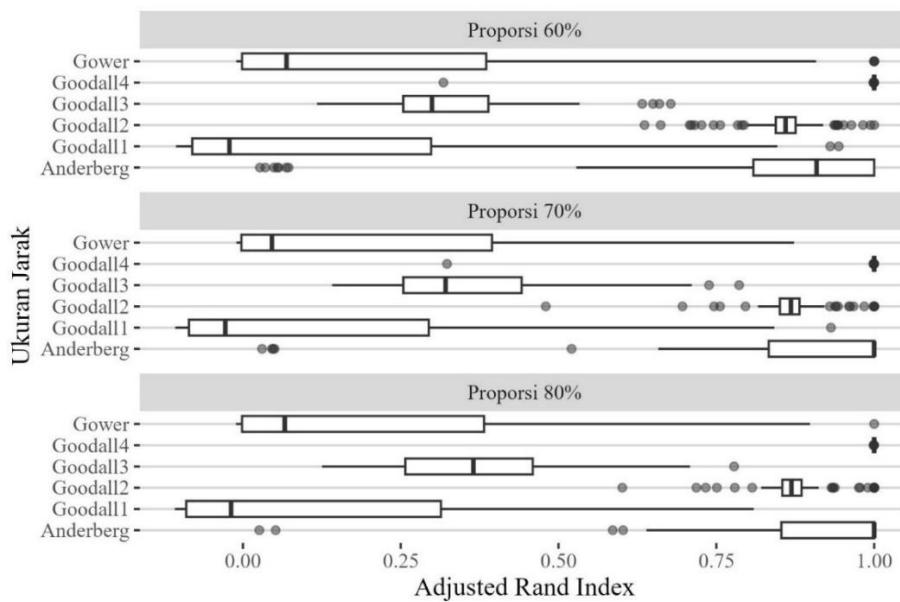
Penentuan ukuran jarak terbaik dilakukan dengan membandingkan empat aspek utama, yaitu kemudahan interpretasi hasil klustering, skor Silhouette, indeks Dunn, dan skor *Adjusted Rand Index* (ARI) berdasarkan 100 kali bootstrap.

Dari sisi interpretasi, Tabel 2 menunjukkan bahwa ukuran jarak Goodall2 dan Goodall3 menghasilkan kluster optimal dalam jumlah besar yang cenderung menyulitkan interpretasi. Sebaliknya, ukuran jarak lainnya menghasilkan dua hingga tiga kluster optimal saja, sehingga lebih mudah dipahami dan dianalisis.



Gambar 9. Skor Silhouette dan Dunn dari kluster optimal masing-masing ukuran jarak

Perbandingan dari aspek kualitas kluster secara internal ditampilkan pada Gambar 9. Ukuran jarak Gower dan Anderberg memiliki skor Silhouette yang cukup baik, namun nilai indeks Dunn-nya relatif rendah. Ini mengindikasikan bahwa meskipun klusternya cukup kompak, jarak antar-kluster tidak terlalu optimal. Sebaliknya, Goodall1, Goodall2, dan Goodall3 menunjukkan skor Dunn yang tinggi, tetapi skor Silhouette-nya rendah, menandakan adanya pemisahan kluster yang baik, namun dengan konsistensi internal yang lemah. Di antara keenam ukuran jarak, Goodall4 menunjukkan keseimbangan paling stabil antara kedua metrik tersebut, dengan nilai Silhouette dan Dunn yang sama-sama relatif baik.



Gambar 10. Distribusi *Adjusted Rand Index* hasil Bootstrapping 100 kali dengan proporsi 60%, 70% dan 80%

Evaluasi terakhir dilakukan melalui validasi stabilitas menggunakan ARI yang dihitung dari hasil bootstrapping sebanyak 100 kali pada proporsi data 60%, 70%, dan 80%. Sebaran yang lebih detail dapat pula dilihat pada Tabel 3.

Tabel 3. Distribusi skor ARI dari 100 kali bootstrap dengan proporsi 60%, 70%, dan 80%.

Proporsi	Jarak	Min.	Q1	Mean	Median	Q3	Maks.	Std Dev.
60%	Anderberg	0.03	0.81	0.84	0.91	1.00	1.00	0.24
	Goodall1	-0.11	-0.08	0.10	-0.02	0.30	0.94	0.25
	Goodall2	0.64	0.84	0.86	0.86	0.88	1.00	0.06
	Goodall3	0.12	0.25	0.33	0.30	0.39	0.68	0.11
	Goodall4	0.32	1.00	0.99	1.00	1.00	1.00	0.07
	Gower	-0.01	0.00	0.22	0.07	0.39	1.00	0.30
70%	Anderberg	0.03	0.83	0.88	1.00	1.00	1.00	0.20
	Goodall1	-0.11	-0.09	0.10	-0.03	0.29	0.93	0.26
	Goodall2	0.48	0.85	0.87	0.87	0.88	1.00	0.06
	Goodall3	0.14	0.25	0.36	0.32	0.44	0.79	0.14
	Goodall4	0.32	1.00	0.99	1.00	1.00	1.00	0.07
	Gower	-0.01	0.00	0.21	0.05	0.39	0.87	0.27
80%	Anderberg	0.03	0.85	0.91	1.00	1.00	1.00	0.16
	Goodall1	-0.11	-0.09	0.11	-0.02	0.31	0.81	0.25
	Goodall2	0.60	0.85	0.87	0.87	0.88	1.00	0.06
	Goodall3	0.13	0.26	0.38	0.37	0.46	0.78	0.14



	Goodall4	1.00	1.00	1.00	1.00	1.00	1.00	0.00
Gower		-0.01	0.00	0.21	0.07	0.38	1.00	0.27

Sebagaimana ditunjukkan pada Gambar 10 dan Tabel 3, skor ARI dari hasil klustering menunjukkan variasi stabilitas yang cukup signifikan antar metode. Ukuran jarak Gower dan Goodall1 memiliki rata-rata skor ARI yang rendah, berkisar antara 0,1 hingga 0,2, disertai dengan standar deviasi tinggi hingga hampir 0,3. Hal ini menunjukkan bahwa hasil klustering dari kedua metode tersebut sangat tidak konsisten terhadap variasi data, dengan fluktuasi besar pada hasil ulangan bootstrap.

Goodall3 menunjukkan performa yang sedikit lebih baik, dengan rata-rata skor ARI berada pada kisaran 0,3 hingga 0,4, serta standar deviasi sekitar 0,15. Ini mengindikasikan peningkatan konsistensi dibandingkan Gower dan Goodall1, meskipun masih menunjukkan ketidakstabilan dalam beberapa kasus. Adanya *outlier* pada Gambar 10 menunjukkan bahwa pada sebagian kecil ulangan, hasil Goodall3 dapat cukup baik, namun tidak konsisten secara keseluruhan.

Performa stabil mulai tampak pada Goodall2, yang menghasilkan rata-rata skor ARI pada kisaran 0,86 hingga 0,87, dengan standar deviasi hanya 0,06. Ini menandakan bahwa hasil klustering cukup stabil pada sebagian besar ulangan, meskipun masih terdapat beberapa pengecualian, sebagaimana ditunjukkan oleh kemunculan *outlier* pada Gambar 10. Hasil yang relatif baik juga diperoleh dari ukuran jarak Anderberg, dengan rata-rata dan median skor ARI lebih tinggi dibandingkan Goodall2, namun disertai standar deviasi yang sedikit lebih besar, yang menunjukkan bahwa meskipun metode ini cenderung stabil, masih terdapat fluktuasi yang perlu diperhatikan.

Akhirnya, hasil paling stabil ditunjukkan oleh Goodall4, dengan rata-rata skor ARI konsisten bernilai 1 pada seluruh proporsi data bootstrap, serta standar deviasi hanya sebesar 0,07. Artinya, meskipun dilakukan bootstrapping dengan berbagai proporsi, hasil klustering dari Goodall4 tetap identik dan tidak terpengaruh oleh variasi data. Hal ini menegaskan bahwa Goodall4 tidak hanya unggul secara numerik, tetapi juga sangat konsisten dan dapat diandalkan dalam penerapan nyata.

Pembahasan

Interpretasi Sosial dan Implikasi Kesehatan Masyarakat berdasarkan Hasil Klaster

Hasil klasterisasi dari enam ukuran jarak yang digunakan dalam penelitian ini mencerminkan segmentasi risiko sosial-kesehatan yang berbeda-beda. Pemahaman terhadap profil tiap klaster, sebagaimana ditunjukkan pada Gambar 3 hingga Gambar 8 serta Tabel 2, penting untuk mendukung desain kebijakan intervensi yang lebih berbasis data. Berikut ini adalah interpretasi sosial dan implikasi kesehatan masyarakat dari masing-masing hasil klaster:

- Gower dan Anderberg: kedua ukuran jarak ini sama-sama mengidentifikasi kelompok rentan dengan risiko kesehatan kronis dan keterbatasan sosial ekonomi. Pada ukuran



jarak Gower, enam responden perempuan dengan riwayat stroke dan tingkat ekonomi menengah membentuk suatu kluster pencilan, sementara pada ukuran jarak Anderberg, 19 responden dengan riwayat stroke membentuk satu kluster pencilan. Kedua kelompok ini berasal dari tingkat ekonomi menengah ke bawah. Dalam konteks kesehatan masyarakat, kelompok ini membutuhkan intervensi terpadu yang mencakup dukungan rehabilitatif dan akses terhadap perlindungan sosial, seperti subsidi transportasi ke fasilitas kesehatan atau program pemberdayaan ekonomi berbasis komunitas.

- Goodall1 dan Goodall2: kedua kluster ini membagi responden berdasarkan umur. Secara umum, perbedaan antar kluster Goodall1 kurang jelas, terutama jika dibandingkan dengan hasil kluster Goodall2 yang membentuk lima kluster. Meskipun begitu, pengklusteran berdasarkan umur tetap relevan dalam konteks pembentukan kebijakan berbasis siklus hidup. Misalnya, responden dalam kategori umur lansia memerlukan intervensi pada penyakit degeneratif dan perawatan jangka panjang, sementara usia muda dapat difokuskan pada program promosi gaya hidup sehat.
- Goodall3: terbentuk tujuh kluster yang tidak memiliki batas yang jelas. Meski memanfaatkan banyak variabel sekaligus (umur, tingkat ekonomi, riwayat diabetes, hipertensi, stroke), klaster yang terbentuk tidak stabil dan sulit diinterpretasikan. Secara praktis, pendekatan ini menyulitkan perumusan kebijakan yang efisien karena kurangnya kejelasan dalam penentuan kelompok sasaran.
- Goodall4: Menghasilkan tiga klaster yang jelas: tanpa stroke, stroke tanpa diabetes, dan stroke dengan diabetes. Klaster ketiga menandai kelompok dengan risiko komorbiditas tinggi, yang memerlukan kebijakan intervensi kesehatan terpadu. Misalnya, pengendalian gula darah terbukti tidak hanya menurunkan risiko komplikasi diabetes, tetapi juga menurunkan risiko stroke iskemik. Oleh karena itu, klaster ini sangat tepat dijadikan target utama program pencegahan dan rehabilitasi kronis yang terkoordinasi.

Evaluasi Teknis dan Performa Ukuran Jarak

Sebagaimana ditunjukkan pada bagian hasil, pemilihan ukuran jarak sangat berpengaruh terhadap kluster yang terbentuk. Sebagai contoh, ukuran jarak Gower menghasilkan klaster yang sangat tidak seimbang, dengan satu klaster besar dan satu klaster kecil yang sangat spesifik. Penyebab pola ini adalah perhitungan jarak data kategorik pada ukuran jarak Gower, di mana perbedaan kategori langsung dihitung sebagai jarak 1. Pola ini sejalan dengan penelitian terdahulu yang dilakukan oleh Pinyan (P. Liu et al., 2024) yang menyatakan bahwa pada data dengan tipe peubah campuran, peubah kategorik cenderung berpengaruh lebih besar dibandingkan tipe peubah numerik.

Pola yang serupa ditunjukkan oleh ukuran jarak Anderberg yang memisahkan seluruh responden dengan riwayat stroke ke dalam satu kluster dan terpisah dari mayoritas responden. Meskipun begitu, penyebab pola pada ukuran jarak Anderberg ini sedikit berbeda dengan penyebab pola pada ukuran jarak Gower, sebab ukuran jarak Anderberg tidak diperuntukkan untuk data campuran melainkan untuk data kategorik. Pola ini terjadi sebab jarak Anderberg tidak mempertimbangkan frekuensi kategori dalam dataset, sehingga pada data dengan



ketidakseimbangan kategori, metrik ini cenderung membentuk satu atau dua klaster dominan dan beberapa klaster kecil yang berisi kategori langka. Hal ini pulalah yang menyebabkan keseluruhan responden dengan riwayat stroke berada pada kluster yang sama, berbanding terbalik dengan hasil kluster Gower dimana hanya enam responden dengan riwayat stroke yang berada dalam satu klaster. Hal ini sesuai dengan temuan pada penelitian terdahulu yang dilakukan oleh (Sulc & Řezanková, 2019).

Adapun hasil klasterisasi dari empat varian jarak Goodall menunjukkan adanya perbedaan signifikan dalam jumlah klaster, karakteristik pembeda, serta performa metrik evaluasi. Pola perbedaan ini dapat ditelusuri melalui sensitivitas masing-masing ukuran jarak terhadap distribusi kategori dan keterkaitan antar variabel kategorik, khususnya dalam konteks data yang cenderung timpang dan terdiri dari tujuh variabel kategorik. Hal ini terutama karena ukuran jarak Goodall menghitung kesamaan kategori berdasarkan frekuensi kemunculan kategori, di mana kategori yang lebih sering dianggap kurang informatif (S. Liu et al., 2021).

Goodall1 menghasilkan dua klaster utama yang dibedakan oleh variabel umur. Dalam konteks ini, dominasi variabel umur yang memiliki sebaran kategori lebih merata membuatnya menjadi satu-satunya pembeda yang signifikan. Hal ini menjelaskan tidak rapatnya kluster yang terbentuk yang ditunjukkan oleh skor silhouette yang rendah, serta kurang stabilnya hasil kluster yang ditunjukkan oleh skor ARI yang rendah, meskipun jarak antar klaster relatif besar yang diindikasikan oleh tingginya nilai Dunn. Akibatnya, Goodall1 cenderung mereduksi kompleksitas struktur data dan tidak menangkap pola yang relevan dari variabel lain, terutama variabel yang timpang.

Goodall2 menunjukkan peningkatan jumlah klaster menjadi lima, masih berbasis variabel umur namun dengan variasi internal yang lebih kaya. Skor Dunn yang tertinggi menunjukkan bahwa jarak antar klaster cukup jelas, namun Silhouette yang tetap rendah mengindikasikan bahwa masing-masing klaster belum sepenuhnya homogen. Fenomena ini dapat dijelaskan oleh mekanisme Goodall2 yang memperhalus pengaruh frekuensi kategori, namun masih mewarisi bias terhadap variabel dengan distribusi kategori yang relatif seimbang. ARI yang relatif baik menunjukkan bahwa struktur klaster cukup stabil dibanding hasil pada data keseluruhan, meski belum optimal dalam interpretasi.

Goodall3, yang menghasilkan tujuh klaster berbasis kombinasi variabel umur, tingkat ekonomi, dan riwayat penyakit, justru menunjukkan performa evaluasi terendah. Meskipun secara eksplisit berhasil melibatkan lebih banyak variabel, skor Silhouette dan Dunn yang rendah serta ARI yang rendah menunjukkan bahwa klaster yang terbentuk tidak rapat, tidak terpisah jelas, dan tidak stabil pada berbagai subset data. Hal ini menyiratkan bahwa Goodall3 terlalu sensitif terhadap kategori yang jarang muncul secara bersamaan, sehingga menghasilkan klaster yang sulit diinterpretasikan dan tidak konsisten. Dalam konteks data kategorik dengan distribusi timpang, pendekatan ini justru memperbesar noise.

Goodall4 menunjukkan performa paling seimbang, menghasilkan tiga klaster yang terdefinisi dengan baik berdasarkan riwayat stroke dan diabetes. Skor Silhouette tertinggi dan Dunn yang juga relatif baik menunjukkan adanya struktur klaster yang kuat dan jelas. Sementara



itu, skor ARI yang mendekati sempurna pada hampir semua ulangan menunjukkan stabilitas yang tinggi terhadap variasi data. Ini mengindikasikan bahwa Goodall4 mampu menangkap kombinasi pola dari variabel dengan kategori dominan secara proporsional, tanpa terlalu mendiskreditkan kategori yang sering muncul atau memperkuat yang langka secara berlebihan.

Secara keseluruhan, pola hasil ini memperlihatkan bahwa pemilihan ukuran jarak sangat memengaruhi struktur dan kualitas klaster pada data kategorik, terutama ketika data mengandung variabel dengan distribusi kategori yang sangat timpang. Goodall1 dan Goodall2 cenderung menyederhanakan struktur dengan fokus pada variabel yang relatif seimbang, sedangkan Goodall3 memperumit struktur dengan klasterisasi yang terlalu sensitif terhadap variabel langka. Sebaliknya, Goodall4 memberikan kompromi terbaik antara kompleksitas, interpretabilitas, dan stabilitas klaster, menjadikannya ukuran jarak yang paling sesuai dalam konteks data ini.

Berdasarkan analisis teknis dan implikasi substantif, Goodall4 muncul sebagai ukuran jarak yang paling optimal. Secara metrik evaluasi, Goodall4 unggul dalam hal keterpisahan, kepadatan, dan stabilitas klaster. Secara substantif, ia menghasilkan segmentasi klinis yang relevan dan mudah ditindaklanjuti dalam kebijakan kesehatan masyarakat. Kombinasi keunggulan teknis dan konteks praktis inilah yang menjadikan Goodall4 paling layak digunakan dalam analisis klaster hierarki pada data kategorikal di bidang kesehatan.

SIMPULAN

Penelitian ini mengevaluasi enam ukuran jarak dalam analisis klaster hierarki menggunakan data IFLS-5 yang terdiri dari 2.060 responden dan tujuh variabel kategorik terkait faktor demografis dan kesehatan. Hasil menunjukkan bahwa pemilihan ukuran jarak sangat memengaruhi struktur klaster yang terbentuk.

Di antara semua ukuran, Goodall4 menghasilkan klaster yang paling stabil, terpisah jelas, dan mudah diinterpretasikan dengan skor Silhouette dan Dunn yang relatif tinggi serta ARI mendekati sempurna. Sebaliknya, Gower dan Anderberg cenderung menghasilkan klaster ekstrem atau tidak seimbang, sementara Goodall1–3 menunjukkan performa yang inkonsisten dan sulit ditafsirkan karena terlalu sensitif terhadap distribusi kategori.

Temuan ini menegaskan pentingnya pemilihan ukuran jarak yang sesuai untuk data kategorik. Namun demikian, hasil ini bersifat kontekstual terhadap struktur dan distribusi data IFLS-5. Performa masing-masing ukuran jarak dapat berbeda pada data dengan jumlah kategori lebih banyak, distribusi kategori yang lebih seimbang atau ekstrem, atau bila digunakan metode klasterisasi selain hierarki.

DAFTAR PUSTAKA

- Adha, M. F., Septian, M., Prananda, P. G. A., & FeryHerdiatmoko, H. (2025). Analisis Hierarchical Clustering pada Volume Trading disetiap Narrative Crypto Menggunakan Data Dune Analytics untuk Investasi Berbasis Data. *Jurnal Teknik Informatika Kaputama (JTIK)*, 9(1), 1–10.



- Aditya, A., Nurina Sari, B., Nur Padilah, T., Sitasi, C., Aditya, A., Sari, B. N., Padilah, T., & pengukuran jarak Euclidean dan Gower, P. (2021). Perbandingan pengukuran jarak Euclidean dan Gower pada klaster k-medoids. *Jurnal Teknologi dan Sistem Komputer*, 9(1), 1–7. <https://doi.org/10.14710/jtsiskom.2021.13747>
- Alamtha, Z., Djakaria, I., & Yahya, N. I. (2023). Implementasi Algoritma Hierarchical Clustering dan Non-Hierarchical Clustering untuk Pengelompokan Pengguna Media Sosial. *ESTIMASI: Journal of Statistics and Its Application*, 4(1), 33–34. <https://doi.org/10.20956/ejsa.vi.24830>
- Dinh, T., Hauchi, W., Fournier-Viger, P., Lisik, D., Ha, M.-Q., Dam, H.-C., & Huynh, V.-N. (2024). Categorical data clustering: 25 years beyond K-modes. *arXiv preprint*, 272, Article 126608. <https://doi.org/10.48550/arXiv.2408.17244>
- Fadilah, Z. R., & Wijayanto, A. W. (2023a). Penggunaan jarak Gower pada analisis klaster untuk data campuran. *Jurnal Statistika Terapan*, 15(2), 120–134.
- Fadilah, Z. R., & Wijayanto, A. W. (2023b). Perbandingan Metode Klasterisasi Data Bertipe Campuran: One-Hot-Encoding, Gower Distance, dan K-Prototype Berdasarkan Akurasi (Studi Kasus: Chronic Kidney Disease Dataset). *Journal of Applied Informatics and Computing*, 7(1), 57–67. <https://doi.org/10.30871/jaic.v7i1.5857>
- Gheorghe, G. (2023). Cluster analysis methods in public health research: A European perspective. *International Journal of Health Data Science*, 5(2), 101–117.
- Johnson, R. A. (1982). *Applied Multivariate Statistical Analysis* (SIXTH EDIT). Upper Saddle River.
- Laksono, B. (2024). Model pengelompokan pasien COVID-19 berdasarkan kondisi klinis di Indonesia. *Jurnal Informatika Medis*, 14(2), 88–103.
- Liu, P., Yuan, H., Ning, Y., Chakraborty, B., Liu, N., & Peres, M. A. (2024). A modified and weighted Gower distance-based clustering analysis for mixed type data: a simulation and empirical analyses. *BMC Medical Research Methodology*, 24(1). <https://doi.org/10.1186/s12874-024-02427-8>
- Liu, S., Zhang, H., & Liu, X. (2021). A study on two-stage mixed attribute data clustering based on density peaks. *International Arab Journal of Information Technology*, 18(5), 634–643. <https://doi.org/10.34028/iajit/18/5/2>
- Mayola, L. (2025). Data-driven approaches for health policy: Evidence from cluster-based analysis in Southeast Asia. *Southeast Asian Journal of Public Health*, 7(1), 12–27.
- Mukhtar, M., Ali, M. K. M., Arina, F., Wicaksono, A. S., Ikhsan, A., Budiaji, W., Abdullah, S., Pertiwi, D. D. A., Zidny, R., Oktarisa, Y., & Sukarna, R. H. (2024). Hierarchical clustering algorithm-dendrogram using Euclidean and Manhattan distance. *Teknika: Jurnal Sains dan Teknologi*, 20(01), 98–104.
- Pansris, B. (2024). Penerapan analisis klaster dalam segmentasi perilaku kesehatan masyarakat. *Jurnal Statistika dan Kesehatan*, 12(1), 55–70.
- Prabowo, R. A., Nisa, K., Faisol, A., & Setiawan, E. (2020). Simulasi Pemilihan Metode Analisis Cluster Hirarki Agglomerative Terbaik Antara Average Linkage Dan Ward Pada Data Yang Mengandung Masalah Multikolinearitas. *Jurnal Siger Matematika*, vol.1(No.2), 49–55.
- Rios, julio rivera, Martinez, C., & Delgado, P. A. (2021). Binary transformations and loss of structure in categorical clustering. *Computational Social Science Review*, 6(3), 142–156.
- Sha, X., Ma, Z., Sethuvenkatraman, S., & Li, W. (2023). A new clustering method with an



- ensemble of weighted distance metrics to discover daily patterns of indoor air quality. *Journal of Building Engineering*, 76(April), 107289. <https://doi.org/10.1016/j.jobe.2023.107289>
- Spada, F., Caruso, R., Notarnicola, I., Belloni, S., Maria, M. De, & Duka, B. (2025). Analyzing readiness for interprofessional education among health program students using hierarchical clustering. *Journal of Interprofessional Care*, 39(3), 1–8. <https://doi.org/https://doi.org/10.1080/13561820.2025.2452973>
- Šulc, Z. (2015). Application of Goodall's and Lin's similarity measures in hierarchical clustering. *Prací účastníků vědecké konference doktorského studia Fakulta informatiky a statistiky*, February, 112. <https://www.researchgate.net/publication/286928130>
- Šulc, Z., & Řezanková, H. (2019). Comparison of Similarity Measures for Categorical Data in Hierarchical Clustering. *Journal of Classification*, 36(1), 58–72. <https://doi.org/10.1007/s00357-019-09317-5>
- Wala, J. (2024). Penggunaan analisis klaster untuk perumusan kebijakan imunisasi di daerah terpencil. *Jurnal Informatika Medis*, 10(1), 33–49.
- Warrens, M. J., & Hoef, M. van der. (2022). Understanding the Adjusted Rand Index and Other Partition Comparison Indices Based on Counting Object Pairs. *Journal of Classification*, 39(3), 487–509.