

Perbandingan K-Means dan Agglomerative Hierarchical Clustering dalam Pengelompokan Provinsi di Indonesia Berdasarkan Tingkat Kepemilikan JKN

¹Maulidina Cahaya Rani, ²Desmulyati

¹²Informatika, Universitas Bina Sarana Informatika, Jakarta

E-mail: ¹15220492@bsi.ac.id, ²desmulyati.dmy@bsi.ac.id

ABSTRAK

Sistem Jaminan Kesehatan Nasional (JKN) merupakan program pemerintah untuk mewujudkan *Universal Health Coverage*, namun tingkat kepemilikannya di setiap provinsi masih bervariasi. Penelitian ini bertujuan mengelompokkan 38 provinsi di Indonesia berdasarkan tingkat kepemilikan JKN periode 2023–2025 serta membandingkan metode K-Means dan Agglomerative Hierarchical Clustering (AHC) dengan pendekatan Ward Linkage. Data diperoleh dari Badan Pusat Statistik (BPS) dan diolah menggunakan Python pada Google Colab melalui tahapan *preprocessing*, penentuan jumlah kluster optimal, proses *clustering*, serta evaluasi menggunakan Silhouette Coefficient dan Davies-Bouldin Index (DBI). Hasil penelitian menunjukkan jumlah kluster optimal sebanyak tiga, yaitu kategori rendah, sedang, dan tinggi. K-Means memperoleh Silhouette Coefficient sebesar 0,595773 dan DBI 0,453632, sedangkan AHC memperoleh Silhouette Coefficient 0,587456 dan DBI 0,433335. Berdasarkan nilai Silhouette Coefficient yang lebih tinggi, K-Means direkomendasikan sebagai metode yang lebih optimal untuk mendukung evaluasi kebijakan dan pemerataan kepesertaan JKN di Indonesia.

Kata kunci : AHC, clustering, Davies-Bouldin Index, JKN, K-Means, Silhouette Coefficient

ABSTRACT

The National Health Insurance (JKN) is a government program designed to achieve Universal Health Coverage; however, its ownership rate varies across provinces in Indonesia. This study aims to cluster the 38 provinces based on JKN ownership levels during 2023–2025 and compare the performance of the K-Means and Agglomerative Hierarchical Clustering (AHC) methods using the Ward Linkage approach. Data were obtained from Statistics Indonesia (BPS) and processed using Python in Google Colab through *preprocessing*, optimal cluster determination, clustering, and evaluation using the Silhouette Coefficient and Davies-Bouldin Index (DBI). The results indicate that the optimal number of clusters is three, representing low, medium, and high categories. K-Means achieved a Silhouette Coefficient of 0.595773 and a DBI of 0.453632, while AHC obtained a Silhouette Coefficient of 0.587456 and a DBI of 0.433335. Based on the higher Silhouette Coefficient, K-Means is recommended as the more effective method to support policy evaluation and the equitable distribution of JKN participation across Indonesia.

Keyword : AHC, clustering, Davies-Bouldin Index, JKN, K-Means, Silhouette Coefficient

1. PENDAHULUAN

Sistem Jaminan Kesehatan Nasional (JKN) merupakan bagian penting dari pembangunan sistem kesehatan di Indonesia. Program yang dikelola BPJS Kesehatan sejak Januari 2014 ini bertujuan memberikan perlindungan kesehatan bagi seluruh masyarakat sebagai upaya mewujudkan *Universal Health Coverage* (UHC) (Pramono et al., 2025). Berdasarkan laporan Dewan Jaminan Sosial Nasional tahun 2025, cakupan kepesertaan JKN per 31 Desember 2025 telah mencapai 282,7 juta jiwa atau 98,1% dari total penduduk Indonesia, meningkat sekitar 4,6 juta jiwa (1,6%) dibandingkan tahun sebelumnya. Seiring dengan cakupan kepesertaan yang hampir menyeluruh, fokus pemerintah saat ini diarahkan pada upaya menjaga dan meningkatkan keaktifan peserta, khususnya pada sektor informal. Hingga akhir tahun 2025, jumlah peserta aktif tercatat sebanyak 233,52 juta jiwa atau 81,0% dari total penduduk, meningkat sekitar 10,85 juta jiwa (4,88%) dibandingkan tahun 2024 (Dewan Jaminan Sosial Nasional, 2025).

Meskipun cakupan JKN secara nasional terus meningkat, tingkat kepemilikan JKN di setiap provinsi masih menunjukkan perbedaan. Data Badan Pusat Statistik periode 2023–2025 pada 38 provinsi menunjukkan bahwa pemerataan kepemilikan JKN belum sepenuhnya tercapai. Kondisi tersebut menunjukkan perlunya pengelompokan provinsi berdasarkan karakteristik kepemilikan JKN sebagai bahan evaluasi dan penyusunan kebijakan yang lebih tepat sasaran. Sebagai bagian dari Sistem Jaminan Sosial Nasional (SJSN), pemerataan kepesertaan JKN perlu terus dipantau agar tujuan perlindungan kesehatan masyarakat dapat tercapai (Nugraha, 2024).

Penelitian Jannah et al. (2024) menerapkan algoritma K-Means untuk mengelompokkan peserta BPJS Kesehatan di Jawa Barat (Jannah et al., 2024). Namun, penelitian tersebut hanya mencakup satu provinsi dan belum membandingkan metode clustering. Sementara itu, Siswadi et al. (2025) membandingkan K-Means dan Agglomerative Hierarchical Clustering (AHC) pada data produksi dan harga cabai di Indonesia (Siswadi et al., 2025), sehingga belum membahas pemerataan kepemilikan JKN. Oleh karena itu, masih terdapat peluang

penelitian yang mengombinasikan cakupan nasional dengan perbandingan dua metode clustering.

K-Means merupakan metode clustering yang mengelompokkan data berdasarkan kedekatan terhadap centroid sehingga memiliki proses komputasi yang efisien, meskipun hasilnya dipengaruhi oleh pemilihan centroid awal (Agustin et al., 2025). Sebaliknya, Agglomerative Hierarchical Clustering (AHC) membentuk kelompok secara bertahap melalui pendekatan *bottom-up*. Dalam penelitian ini digunakan Ward Linkage karena mampu meminimalkan varians dalam kluster sehingga menghasilkan kelompok yang lebih homogen dan stabil (Maulana et al., 2025).

Berdasarkan latar belakang tersebut, penelitian ini bertujuan mengelompokkan 38 provinsi di Indonesia berdasarkan tingkat kepemilikan JKN periode 2023–2025 dengan membandingkan metode K-Means dan Agglomerative Hierarchical Clustering (AHC) menggunakan pendekatan Ward Linkage. Analisis dilakukan menggunakan Python pada Google Colab. Jumlah kluster optimal ditentukan melalui Metode Elbow dan Dendrogram, sedangkan kualitas hasil clustering dievaluasi menggunakan Silhouette Coefficient dan Davies-Bouldin Index (DBI). Hasil penelitian diharapkan dapat memberikan informasi mengenai kelompok provinsi dengan tingkat kepemilikan JKN rendah, sedang, dan tinggi sebagai dasar evaluasi serta perumusan kebijakan pemerataan JKN di Indonesia.

2. LANDASAN TEORI

2.1. Metode Elbow

Metode Elbow digunakan untuk menentukan jumlah kluster terbaik dengan melihat grafik penurunan error yang membentuk titik seperti siku. Tujuannya memilih nilai k yang optimal, yaitu yang menghasilkan SSE rendah. SSE (Sum of Squared Error) mengukur jarak antara data dan pusat kluster pada K-Means (Maori & Evanita, 2023). Persamaan SSE ditunjukkan pada Persamaan (1):

$$SSE = \sum_{k=1}^K \sum_{x_i} |x_i - c_k|^2 \quad (1)$$

K adalah jumlah kluster, c_k merupakan centroid pada kluster ke- k , dan x_i data atau observasi ke- i .

2.2. Dendrogram

Dendrogram adalah diagram pohon yang memvisualisasikan hubungan hierarkis antar-objek hasil klusterisasi berjenjang dari AHC. Melalui alat bantu visual ini, peneliti dapat lebih mudah menganalisis dan menentukan cara terbaik dalam mengelompokkan data berdasarkan karakteristik yang serupa pada ambang batas jarak tertentu (Abdulpatah et al., 2025).

2.3. K-Means Clustering

Metode K-Means Clustering digunakan untuk mengelompokkan data ke dalam beberapa kluster berdasarkan kemiripan karakteristik, di mana data yang memiliki kesamaan akan berada pada kelompok yang sama. Tingkat kedekatan antara data dan centroid diukur menggunakan metode Euclidean Distance sesuai Persamaan (2) (Julianti & Widiartha, 2024):

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

Di mana $d(x, y)$ adalah jarak *Euclidean* antara dua objek data, x_i dan y_i merupakan nilai atribut ke- i dari masing-masing objek, dan n adalah jumlah dimensi atau fitur data.

2.4. Agglomerative Hierarchical Clustering (Ward Linkage)

AHC diterapkan menggunakan pendekatan Ward Linkage, yaitu metode yang mengoptimalkan pengelompokan dengan meminimalkan variansi dalam kluster melalui pendekatan jumlah kuadrat jarak (Maulana et al., 2025). Untuk menghitung nilai kesalahan kuadrat total dari dua kluster, digunakan Persamaan (3):

$$SSE_{UV} = \frac{1}{2} \sum_{w=1}^p (x_{UW} - x_{VW})^2 \quad (3)$$

SSE_{UV} merupakan nilai *Sum of Squared Errors* dari penggabungan kluster U dan kluster V , p adalah jumlah total dimensi atau fitur data, sedangkan x_{UW} dan x_{VW} berturut-turut menunjukkan nilai koordinat atau rata-rata fitur ke- w dari kluster U dan kluster V .

2.5. Silhouette Coefficient

Silhouette Coefficient digunakan untuk mengukur kualitas dan kelayakan suatu kluster serta membantu menilai apakah data sudah berada pada kluster yang sesuai (Amri et al., 2025). Rumus yang digunakan ditunjukkan pada Persamaan (4):

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

Adapun $a(i)$ mewakili rata-rata jarak objek ke semua data lain dalam kluster yang sama (kohesi), dan $b(i)$ mewakili rata-rata jarak objek ke kluster terdekat lain (separasi).

2.6. Davies-Bouldin Index

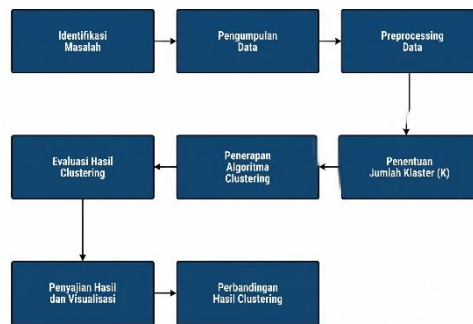
Davies-Bouldin Index (DBI) mengevaluasi kualitas clustering berdasarkan rasio jarak internal kluster terhadap jarak antar kluster (Yolandari et al., 2025). Nilai DBI dihitung melalui Persamaan (5)

$$DBI = \frac{1}{c} \sum_{i=1}^c R_i \quad (5)$$

Pada persamaan tersebut, c adalah jumlah kluster, R_i melambangkan nilai maksimal rasio kedekatan kluster ke- i , S_i merupakan dispersi internal kluster, dan $d(v_i, v_j)$ melambangkan jarak antarpusat klas

3. METODOLOGI

Penelitian ini menggunakan pendekatan *data mining* dengan teknik clustering, yaitu metode *unsupervised learning* yang mengelompokkan data berdasarkan tingkat kemiripan karakteristik tanpa menggunakan label kelas (Rustiyana et al., 2025). Tahapan penelitian meliputi identifikasi masalah, pengumpulan data, *preprocessing* data, penentuan jumlah kluster optimal, penerapan algoritma K-Means dan Agglomerative Hierarchical Clustering (AHC), evaluasi hasil menggunakan Silhouette Coefficient dan Davies-Bouldin Index (DBI), penyajian serta visualisasi hasil, dan analisis perbandingan kedua metode. Alur penelitian secara keseluruhan ditunjukkan pada Gambar 1.



Gambar 1. Kerangka Penelitian

3.1. Identifikasi Masalah

Tahap awal dilakukan dengan mengidentifikasi variasi tingkat kepemilikan JKN antarprovinsi di Indonesia serta perlunya pengelompokan wilayah dan perbandingan metode K-Means dengan Agglomerative Hierarchical Clustering (AHC).

3.2. Pengumpulan Data

Penelitian ini menggunakan data sekunder berupa persentase penduduk yang memiliki JKN menurut provinsi tahun 2023, 2024, dan 2025 yang bersumber dari publikasi resmi Badan Pusat Statistik melalui <https://www.bps.go.id>, mencakup 38 provinsi di Indonesia yang disajikan pada Tabel 1. Seluruh data diolah menggunakan bahasa pemrograman Python beserta pustaka *Pandas* dan *NumPy* untuk membantu proses pengolahan dan analisis data (Candra, 2025), yang dijalankan pada platform Google Colab karena kemudahan akses, dukungan GPU/TPU gratis, serta integrasi dengan Google Drive tanpa memerlukan instalasi tambahan (Alfian & Hidayat, 2024).

Tabel 1. Persentase Penduduk yang Memiliki Jaminan Kesehatan Nasional (JKN)

Provinsi	2023	2024	2025
Aceh	96.87	97.19	98.18
Sumatera Utara	61.86	65.75	68.69
Sumatera Barat	71.41	75.76	78.19
Riau	63.83	68.38	73.30
Jambi	53.49	58.06	61.04
Sumatera Selatan	64.43	71.49	73.33
Bengkulu	68.18	72.02	76.20
Lampung	66.48	70.18	71.90
Kep. Bangka Belitung	75.45	80.13	84.11
Kep. Riau	77.72	81.91	80.00
DKI Jakarta	88.48	89.91	91.69
Jawa Barat	64.85	68.84	70.62
Jawa Tengah	69.08	70.90	72.96
DI Yogyakarta	83.28	85.70	89.27
Jawa Timur	68.55	69.66	73.86

Banten	67.47	71.86	74.49
Bali	85.89	88.47	89.15
Nusa Tenggara Barat	63.92	68.28	71.73
Nusa Tenggara Timur	72.18	75.24	77.45
Kalimantan Barat	61.96	65.22	68.92
Kalimantan Tengah	57.96	64.53	67.89
Kalimantan Selatan	68.79	70.85	74.32
Kalimantan Timur	79.75	82.88	83.00
Kalimantan Utara	85.03	87.46	88.40
Sulawesi Utara	79.16	84.41	83.46
Sulawesi Tengah	75.41	77.39	80.86
Sulawesi Selatan	80.07	83.55	86.56
Sulawesi Tenggara	75.74	80.39	84.19
Gorontalo	84.11	87.74	87.21
Sulawesi Barat	88.52	89.53	89.32
Maluku	60.81	61.75	65.53
Maluku Utara	69.32	69.58	73.41
Papua Barat	70.98	72.41	76.74
Papua Barat Daya	73.29	81.37	84.13
Papua	76.93	81.04	79.90
Papua Selatan	88.43	90.11	92.19
Papua Tengah	67.06	67.34	70.83
Papua Pegunungan	89.91	97.28	96.31

3.3. Preprocessing Data

Data diperiksa untuk memastikan tidak terdapat nilai kosong maupun duplikasi, kemudian dilakukan penyesuaian tipe data, perhitungan rata-rata kepemilikan JKN periode 2023–2025, serta normalisasi menggunakan *StandardScaler*.

3.4. Penentuan Jumlah Kluster

Jumlah kluster optimal ditentukan menggunakan Metode Elbow dan Dendrogram. Berdasarkan kedua metode tersebut diperoleh jumlah kluster optimal sebanyak tiga ($K=3$).

3.5. Penerapan Algoritma Clustering

Data dikelompokkan menggunakan algoritma K-Means dan Agglomerative Hierarchical Clustering (AHC) dengan pendekatan Ward Linkage menggunakan jumlah kluster yang sama.

3.6. Evaluasi Hasil Clustering

Kualitas hasil pengelompokan dievaluasi menggunakan *Silhouette Coefficient* dan

Davies-Bouldin Index (DBI) serta dibandingkan pada beberapa alternatif jumlah klaster.

3.7. Penyajian Hasil dan Visualisasi

Hasil clustering diberi label kategori rendah, sedang, dan tinggi berdasarkan nilai rata-rata kepemilikan JKN, kemudian disajikan dalam bentuk tabel dan visualisasi peta tematik.

3.8. Perbandingan Hasil Clustering

Tahap akhir dilakukan dengan membandingkan performa K-Means dan AHC berdasarkan nilai Silhouette Coefficient dan DBI untuk menentukan metode yang memberikan hasil pengelompokan terbaik.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Preprocessing Data

Berdasarkan Hasil pengecekan menunjukkan tidak terdapat missing value maupun data duplikat pada seluruh kolom. Rata-rata kepemilikan JKN tertinggi terdapat pada Provinsi Aceh (97,41%), sedangkan terendah pada Provinsi Jambi (57,53%). Dataset lengkap beserta hasil standarisasi (Z-score) ditampilkan pada Tabel 2.

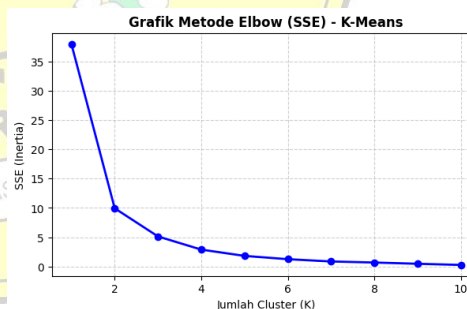
Tabel 2. Hasil Perhitungan Rata-rata dan Standarisasi Kepemilikan JKN

Provinsi	Rata_JKN	Z-Score
Aceh	97.413333	2.201099
Sumatera Utara	65.433333	-1.178274
Sumatera Barat	75.120000	-0.154670
Riau	68.503333	-0.853862
Jambi	57.530000	-2.013431
Sumatera Selatan	69.750000	-0.722125
Bengkulu	72.133333	-0.470275
Lampung	69.520000	-0.746430
Kep. Bangka Belitung	79.896667	0.350088
Kep. Riau	79.876667	0.347974
DKI Jakarta	90.026667	1.420540
Jawa Barat	68.103333	-0.896131
Jawa Tengah	70.980000	-0.592149
DI Yogyakarta	86.083333	1.003842
Jawa Timur	70.690000	-0.622794
Banten	71.273333	-0.561153
Bali	87.836667	1.189119
Nusa Tenggara Barat	67.976667	-0.909516
Nusa Tenggara Timur	74.956667	-0.171929
Kalimantan Barat	65.366667	-1.185319

Kalimantan Tengah	63.460000	-1.386799
Kalimantan Selatan	71.320000	-0.556221
Kalimantan Timur	81.876667	0.559317
Kalimantan Utara	86.963333	1.096833
Sulawesi Utara	82.343333	0.608631
Sulawesi Tengah	77.886667	0.137688
Sulawesi Selatan	83.393333	0.719586
Sulawesi Tenggara	80.106667	0.372279
Gorontalo	86.353333	1.032373
Sulawesi Barat	89.123333	1.325083
Maluku	62.696667	-1.467461
Maluku Utara	70.770000	-0.614340
Papua Barat	73.376667	-0.338890
Papua Barat Daya	79.596667	0.318386
Papua	79.290000	0.285980
Papua Selatan	90.243333	1.443435
Papua Tengah	6.841000	-0.863725
Papua Pegunungan	94.500000	1.893243

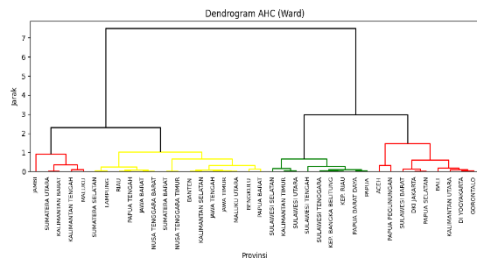
4.2. Hasil Penentuan Jumlah Klaster

Berdasarkan hasil pemrosesan metode Elbow, diperoleh nilai SSE untuk variasi K=1 hingga K=10. Penurunan nilai SSE paling tajam terjadi pada rentang K=1 ke K=3, setelah K=3, penurunan kurva melambat dan melandai secara konsisten sebagaimana ditunjukkan pada Gambar 2.



Gambar 2. Hasil Grafik Metode Elbow

Hasil ini diperkuat oleh Gambar 3, Dendrogram AHC menggunakan Ward Linkage juga mengonfirmasi terbentuknya tiga kelompok besar secara alami pada pemotongan ambang jarak threshold 1,5. Kombinasi kedua metode ini secara mutlak menetapkan jumlah klaster optimal adalah K=3.



Gambar 3. Hasil Dendrogram Berdasarkan Ward Linkage

Berdasarkan hasil analisis kedua metode, ditetapkan jumlah kluster optimal $K=3$ yang digunakan pada seluruh tahap analisis selanjutnya.

4.3. Hasil Penerapan Clustering

Untuk memudahkan proses perbandingan antaralgoritma, dilakukan sinkronisasi label kluster berdasarkan urutan nilai rata-rata Rata_JKN. Label kluster kemudian diinterpretasikan menjadi kategori 0 (Rendah), 1 (Sedang), dan 2 (Tinggi) sesuai dengan urutan nilai rata-rata dari yang terendah hingga tertinggi. Hasil pengelompokan menggunakan algoritma K-Means dan Agglomerative Hierarchical Clustering (AHC) setelah proses sinkronisasi tersebut disajikan pada Tabel 3.

Tabel 3. Hasil Clustering Sinkronisasi Label

Provinsi	(K) K-Means	(K) AHC
Aceh	2	2
Sumatera Utara	0	0
Sumatera Barat	1	0
Riau	0	0
Jambi	0	0
Sumatera Selatan	0	0
Bengkulu	0	0
Lampung	0	0
Kep. Bangka Belitung	1	1
Kep. Riau	1	1
DKI Jakarta	2	2
Jawa Barat	0	0
Jawa Tengah	0	0
DI Yogyakarta	2	2
Jawa Timur	0	0
Banten	0	0
Bali	2	2
Nusa Tenggara Barat	0	0
Nusa Tenggara Timur	1	0
Kalimantan Barat	0	0
Kalimantan Tengah	0	0
Kalimantan Selatan	0	0
Kalimantan Timur	1	1
Kalimantan Utara	2	2
Sulawesi Utara	1	1

Sulawesi Tengah	1	1
Sulawesi Selatan	1	1
Sulawesi Tenggara	1	1
Gorontalo	2	2
Sulawesi Barat	2	2
Maluku	0	0
Maluku Utara	0	0
Papua Barat	0	0
Papua Barat Daya	1	1
Papua	1	1
Papua Selatan	2	2
Papua Tengah	0	0
Papua Pegunungan	2	2

4.4. Hasil Evaluasi Clustering

Evaluasi Silhouette Coefficient dan DBI dilakukan pada rentang $K=2$ hingga $K=10$ sebagaimana Tabel 4. Pada $K=2$, nilai kedua metode identik karena data bersifat univariat sehingga struktur pemisahan yang jelas menghasilkan partisi yang sama meski pendekatan algoritma berbeda. Perbedaan mulai tampak pada $K=3$, dipengaruhi provinsi pada wilayah batas seperti Sumatera Barat dan Nusa Tenggara Timur.

Tabel 4. Hasil Evaluasi Clustering

K	K-Means		AHC (Ward)	
	Silhouette	DBI	Silhouette	DBI
2	0.637751	0.485994	0.637751	0.485994
3	0.595773	0.453632	0.587456	0.433335
4	0.593916	0.475673	0.593916	0.475673
5	0.615571	0.423363	0.615571	0.423363
6	0.578692	0.453103	0.556610	0.461183
7	0.585617	0.369219	0.577306	0.406917
8	0.561115	0.401283	0.590996	0.363370
9	0.590261	0.381442	0.603717	0.358425
10	0.628649	0.345599	0.616528	0.348914

Meskipun $K=2$ menghasilkan Silhouette Coefficient tertinggi (0,637751), jumlah kluster tersebut kurang informatif untuk merepresentasikan variasi kepemilikan JKN. Pada $K=3$, K-Means memperoleh Silhouette Coefficient 0,595773 dan DBI 0,453632, sedangkan AHC memperoleh Silhouette Coefficient 0,587456 dan DBI 0,433335. Kedua nilai tersebut berada pada kategori struktur kluster yang reasonable (di atas 0,5 untuk Silhouette Coefficient dan jauh di bawah 1 untuk DBI), sehingga $K=3$ dinilai memberikan keseimbangan terbaik antara kualitas kluster dan jumlah kelompok yang terbentuk.

4.5. Hasil Karakteristik

Analisis nilai rata-rata asli kepemilikan JKN pada masing-masing kluster dilakukan untuk mengidentifikasi karakteristik khas dari setiap kelompok, seperti disajikan pada Tabel 5.

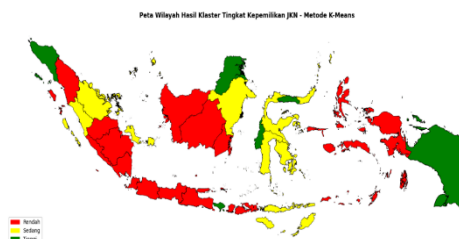
Tabel 5. Nilai Rata-rata Karakteristik Setiap Kluster

Metode	Rendah (%)	Sedang (%)	Tinggi (%)
K-Means	68.182963	79.485758	89.838148
AHC Ward	68.868500	80.474074	89.838148

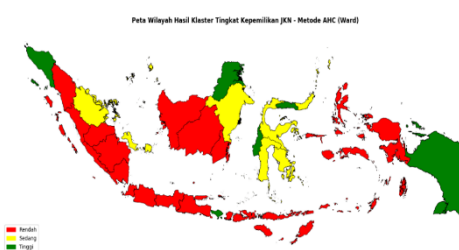
Kedua metode menghasilkan nilai rata-rata yang identik pada kategori Tinggi (89,84%), menunjukkan konsistensi dalam mengidentifikasi provinsi dengan kepemilikan JKN tertinggi, sementara perbedaan pada kategori Rendah dan Sedang disebabkan pergeseran Sumatera Barat dan Nusa Tenggara Timur pada metode AHC.

4.6. Hasil Visualisasi

Hasil *clustering* divisualisasikan dalam bentuk peta tematik (*choropleth map*) menggunakan GeoPandas, dengan warna merah untuk kategori Rendah, kuning untuk Sedang, dan hijau untuk Tinggi, sebagaimana Gambar 4 dan Gambar 5. Visualisasi menggunakan tiga kategori kluster, yaitu Rendah (merah), Sedang (kuning), dan Tinggi (hijau). Berdasarkan hasil visualisasi, metode K-Means dan Agglomerative Hierarchical Clustering (AHC) menunjukkan pola persebaran yang relatif serupa, dengan kategori Tinggi terkonsentrasi pada beberapa provinsi seperti Aceh, DKI Jakarta, DI Yogyakarta, Bali, Kalimantan Utara, Gorontalo, Sulawesi Barat, Papua Selatan, dan Papua Pegunungan. Perbedaan pengelompokan hanya ditemukan pada Provinsi Sumatera Barat dan Nusa Tenggara Timur yang berada pada kategori Sedang pada metode K-Means, namun masuk ke kategori Rendah pada metode AHC. Secara keseluruhan, kedua metode menghasilkan pola spasial yang konsisten dalam mengidentifikasi tingkat kepemilikan JKN antarprovinsi, sehingga visualisasi peta memperkuat hasil *clustering* yang telah diperoleh serta memudahkan interpretasi persebaran kategori Rendah, Sedang, dan Tinggi di Indonesia.



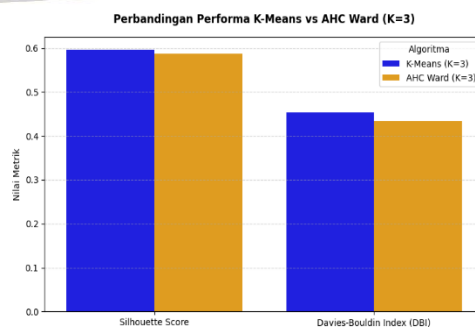
Gambar 4. Peta Hasil Clustering K-Means



Gambar 5. Peta Hasil Clustering AHC (Ward)

4.7. Hasil Perbandingan

Berdasarkan perbandingan metrik evaluasi akhir pada Tabel 4 kondisi K=3, metode K-Means memiliki keunggulan pada aspek Silhouette Coefficient ($0,595773 > 0,587456$), sedangkan AHC Ward unggul tipis pada nilai DBI yang lebih kecil ($0,433335 < 0,453632$). K-Means direkomendasikan sebagai metode terbaik karena Silhouette Coefficient mengukur kekompakan dan keterpisahan global antar-objek secara komprehensif, sehingga menghasilkan batasan kluster provinsi yang lebih tegas dan minim tumpang tindih. Berikut divisualisasikan dalam bentuk bar chart pada Gambar 6.



Gambar 6. Visualisasi Hasil Perbandingan

5. KESIMPULAN

Penelitian ini berhasil menyusun pengelompokan 38 provinsi di Indonesia berdasarkan tingkat kepemilikan JKN periode 2023-2025 menjadi tiga kluster logis (Rendah, Sedang, Tinggi) dengan kesesuaian klasifikasi antar-metode mencapai 94,74%. Jumlah kluster optimal sebanyak tiga ($K=3$) secara konsisten divalidasi oleh Metode Elbow dan Dendrogram. Hasil evaluasi menunjukkan metode K-Means lebih optimal untuk direkomendasikan bagi pemetaan wilayah ini karena menghasilkan nilai Silhouette Coefficient yang lebih tinggi sebesar 0,595773 dibandingkan AHC sebesar 0,587456, memberikan pemisahan wilayah yang lebih tegas untuk mendukung kebijakan pemerataan UHC di Indonesia. Penelitian selanjutnya disarankan menambahkan variabel lain seperti tingkat kemiskinan atau aksesibilitas fasilitas kesehatan, membandingkan dengan metode clustering lain seperti DBSCAN atau K-Medoids, serta mempertimbangkan pembaruan data secara berkala agar hasil pengelompokan selalu mencerminkan kondisi terkini.

6. UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Bina Sarana Informatika atas dukungan yang diberikan selama pelaksanaan penelitian ini. Penulis juga menyampaikan terima kasih kepada dosen pembimbing atas arahan, bimbingan, serta masukan yang sangat berharga selama proses penelitian. Ucapan terima kasih turut disampaikan kepada Badan Pusat Statistik (BPS) atas kesempatan yang diberikan kepada penulis untuk melaksanakan Praktik Kerja Lapangan (PKL) serta penyediaan data statistik yang digunakan dalam penelitian ini. Selain itu, penulis mengucapkan terima kasih kepada keluarga dan rekan-rekan yang senantiasa memberikan doa, motivasi, dan dukungan sehingga penelitian ini dapat diselesaikan dengan baik.

DAFTAR PUSTAKA

Abdulpatah, T., Sari, B. N., & Susilawati. (2025). Perbandingan Algoritma K-Means dan Agglomerative Hierarchical

Clustering untuk Pengelompokan Daerah Penghasil Padi di Indonesia.

Jurnal Informatika dan Teknik Elektro Terapan, 13(3).
<https://doi.org/10.23960/jitet.v13i3.7251>

Agustin, S. R., Purnamasari, I., & Sari, B. N. (2025). Implementasi K-Means Untuk Pengelompokan Kategori Penjualan Barang Berbasis Web. *Journal of Informatics Management and Information Technology*, 5(3), 167–176.

<https://doi.org/10.47065/jimat.v5i3.644>

Alfian, W., & Hidayat, T. (2024). Analisis Clustering Pegawai Berdasarkan Tingkat Kedisiplinan Menggunakan Algoritma K-Means dan Davies-Bouldin Index. *Journal of Electrical Engineering and Computer (JEECOM)*, 6(2), 437–448.
<https://doi.org/10.33650/jeeecom.v6i2.9556>

Amri, M. A., Bakri, M. F., Risal, A. A. N., & Surianto, D. F. (2025). Mini-Batch K-Means Clustering untuk Pengelompokan Kemandirian Daerah di Sulawesi Selatan. *JIP (Jurnal Informatika Polinema)*, 11(2), 235–244.

Candra, A. P. (2025). Analisis Data Menggunakan Python: Memperkenalkan Pandas dan NumPy. *Journal of Information System and Education Development*, 3(1), 11–16.
<https://doi.org/10.62386/jised.v3i1.118>

Dewan Jaminan Sosial Nasional. (2025). Monthly Report Monitoring JKN. <https://kesehatan.djsn.go.id/kesehatan/lap-bulanan2025/>

Jannah, Z., Irawan, A. S. Y., & Nurkifli, E. H. (2024). Penerapan Algoritma K-Means dalam Segmentasi Daerah Berdasarkan Jumlah Warga yang Terdaftar Menjadi Peserta BPJS Kesehatan di Jawa Barat. *Jurnal Mahasiswa Teknik Informatika*, 8(4).

<https://doi.org/10.36040/jati.v8i4.10019>

Julianti, S., & Widiartha, I. M. (2024). Algoritma K-Means untuk Clustering Provinsi di Indonesia Berdasarkan Kasus Stunting. *JNATIA: Jurnal Nasional Teknologi Informasi dan Aplikasinya*, 2(3).
<https://doi.org/10.24843/JNATIA.2024.v02.i03.p16>

- Maori, N. A., & Evanita. (2023). Metode Elbow dalam Optimasi Jumlah Cluster pada K-Means Clustering. *Jurnal SIMETRIS*, 14(2).
<https://doi.org/10.24176/simet.v14i2.9630>
- Maulana, R. H., Akbar, F. R., Yotenka, R., & Suparna. (2025). Analisis Agglomerative Hierarchical Clustering terhadap Jumlah Usaha Pertanian Perorangan Pengguna Lahan Pertanian. *Emerging Statistics and Data Science Journal*, 3(2).
<https://doi.org/10.20885/esds.vol3.iss.2.art15>
- Nugraha, A. (2024). Determinan dalam Kepesertaan Program Jaminan Kesehatan Nasional. Penerbit NEM.
- Pramono, J. S., Hasanah, S., Ardyanti, D., Tonapa, E., Bernadetha, Mayasari, A. D., Fatmala, A. N. P., Afrianti, P., Harlis, F. P., Yuniarti, C. A., Andreastuti, D., Palin, Y. T., Mardiana, N., Nurrahman, M. A., Rahayu, E. P., & Ansyari, F. (2025). Jaminan Kesehatan Nasional: Strategi, Implementasi dan Transformasi Kesehatan. Penerbit Adab.
- Rustiyana, Judijanto, L., Mahendra, G. S., Kamil, Z. A., Purba, D. N., Sutoyo, M. N., Hendrayana, I. G., Pasrun, Y. P., & Prayudani, S. (2025). Data Mining: Algoritma dan Penerapannya. PT. Sonpedia Publishing Indonesia.
- Siswadi, N., Jaya, J., Handhayani, T., & Lewenusa, I. (2025). Perbandingan Algoritma K-Means dan Hierarchical Clustering untuk Pengelompokan Wilayah Berdasarkan Produksi dan Harga Cabai di Indonesia. *Jurnal Mahasiswa Teknik Informatika*, 9(5).
<https://doi.org/10.36040/jati.v9i5.14777>
- Yolandari, N. A., Butarbutar, L. E., Rajagukguk, G. C. H., Zulfi, M. F., Arnita, & Ramadhani, F. (2025). Analisis Perbandingan K-Means dan DBSCAN dalam Pengelompokan Data Travel Review Ratings Menggunakan Evaluasi Silhouette Index dan Davies-Bouldin Index. *Jurnal Informatika dan Teknik Elektro Terapan*, 13(3).
<https://doi.org/10.23960/jitet.v13i3.6884>