



Transformasi Data Menjadi Kebijakan: Studi *Clustering* Hierarki pada Layanan Keagamaan Kementerian Agama

Yossy Candra^{1*}, Andrea Tri Rian Dani²

¹ *Statistisi, Sekretariat Jenderal, Kementerian Agama*

² *Program Studi Statistika, Jurusan Matematika, FMIPA, Universitas Mulawarman*

yondra2022@gmail.com

Abstract

Indonesia is a country rich in religious, cultural, and geographical diversity. To ensure fair and equitable religious service policies, this study analyzes the diversity of regional characteristics. The aim is to cluster provinces in Indonesia based on religious service data. The method used is hierarchical clustering with five algorithms and three distance measures. The combination of average linkage and chebyshev distance resulted in the highest cophenetic correlation coefficient (0.969), making it the most optimal algorithm. The silhouette coefficient analysis identified two clusters as the optimal number, with a silhouette coefficient value of 0.736, and ANOVA testing validated that the differences between clusters are significant. The first cluster consists of the provinces of Aceh, East Java, West Java, and Central Java, which have higher average religious service values. The fact that all other provinces, including Jakarta, are included in the second cluster suggests that religious services could still be of higher quality. These results offer crucial information that the Ministry of Religious Affairs can use to keep enhancing the standard of religious services throughout Indonesia, with an emphasis on expanding service enhancements.

Keywords: ANOVA; Cluster Analysis; Hierarchical Clustering; Religious Services; Similarity Distance

Abstrak

Indonesia adalah negara dengan keberagaman agama, budaya, dan kondisi geografis yang sangat kaya. Untuk memastikan kebijakan layanan keagamaan yang merata dan adil, penelitian ini menganalisis keberagaman karakteristik daerah. Tujuannya adalah untuk mengklasifikasikan provinsi-provinsi di Indonesia berdasarkan data layanan keagamaan. Metode yang digunakan adalah *hierarchical clustering* dengan lima algoritma dan tiga ukuran jarak. Kombinasi *average linkage* dan jarak *chebyshev* menghasilkan koefisien korelasi *cophenetic* tertinggi (0,969), menjadikannya algoritma paling optimal. Uji *silhouette coefficient* menunjukkan dua kluster sebagai jumlah optimal, dengan nilai *silhouette coefficient* sebesar 0,736, dan uji ANOVA memvalidasi bahwa perbedaan antar kluster signifikan. Kluster pertama terdiri dari Provinsi Aceh, Jawa Timur, Jawa Barat, dan Jawa Tengah, yang memiliki rata-rata setiap variabel lebih tinggi. Kluster kedua mencakup seluruh provinsi lainnya, termasuk DKI Jakarta, yang menunjukkan bahwa masih ada peluang untuk peningkatan kualitas layanan keagamaan. Temuan ini memberikan gambaran penting bagi Kementerian Agama untuk terus berupaya meningkatkan kualitas layanan keagamaan di seluruh wilayah Indonesia, dengan fokus pada perbaikan aspek layanan yang lebih luas.

Kata Kunci: Analisis Kluster; ANOVA; *Hierarchical Clustering*; Jarak Kemiripan; Layanan Keagamaan

1. PENDAHULUAN

Indonesia dikenal dengan keberagaman agama, budaya, dan geografis yang kaya (Khairiah & Walid, 2020). Pada konteks pelayanan publik, Kementerian Agama memiliki mandat untuk memastikan layanan keagamaan yang merata, adil, dan inklusif bagi seluruh umat beragama, mencakup layanan, infrastruktur, dan pendidikan berbasis agama (Siraj, 2018). Semakin kompleksnya kebutuhan masyarakat dan dinamika sosial keagamaan memerlukan kebijakan berbasis data yang dapat menangkap variasi dan kebutuhan spesifik tiap daerah.

Penyusunan kebijakan pelayanan publik di Indonesia sering mengabaikan keragaman karakteristik daerah, di mana setiap daerah memiliki kondisi geografis, sosial, dan demografis yang mempengaruhi kebutuhan dan kapasitas layanan (Arifin, 2024). Ketidakseimbangan antara permintaan dan ketersediaan layanan tidak bisa diselesaikan dengan pendekatan seragam, karena beberapa daerah memiliki sarana layanan yang cukup, sementara yang lain terbatas (Siregar et al., 2024). Ketimpangan ini menekankan pentingnya kebijakan yang disesuaikan dengan segmentasi dan kebutuhan lokal (Rizaldy & Asnani, 2024). Hal ini juga berlaku pada layanan keagamaan, yang harus disesuaikan dengan kondisi tiap daerah. Pendekatan seragam sering tidak efektif karena perbedaan kebutuhan dan ketersediaan layanan (Hariyah, 2016). Tantangan utama adalah bagaimana mengelompokkan daerah berdasarkan ciri layanan keagamaannya agar kebijakan lebih tepat sasaran.

Untuk mendukung kebijakan berbasis data, metode *clustering* merupakan pendekatan statistik yang efektif untuk mengidentifikasi pola dan segmentasi daerah berdasarkan data layanan keagamaan (Juliyanto et al., 2024). Metode ini memungkinkan pengelompokan wilayah yang memiliki karakteristik serupa, mendukung perencanaan kebijakan yang lebih tepat sasaran (Ismail, 2024). *Clustering hierarki* memungkinkan visualisasi hubungan antar wilayah melalui dendrogram, memberikan pemahaman yang lebih mendalam tentang kesamaan dan perbedaan antar wilayah (Candra, 2025). Pendekatan ini juga membantu dalam penilaian distribusi layanan keagamaan, sehingga kebijakan dapat disesuaikan dengan kebutuhan spesifik tiap daerah.

Sebagai inovasi, uji ANOVA digunakan untuk menguji perbedaan signifikan antar kluster dan memvalidasi hasil *clustering* (Lawuna et al., 2024). Uji ini memperkuat hasil *clustering* dengan memberikan dukungan empiris mengenai karakteristik tiap kluster. Penelitian sebelumnya, seperti yang dilakukan oleh Yuniar dan Heikal (2024), menggunakan *K-Means clustering* untuk segmentasi sertifikat halal berdasarkan 12 variabel, menghasilkan lima kluster wilayah yang mendasari kebijakan pengembangan produk halal. Penelitian ini memiliki kebaruan menggunakan *hierarchical clustering* dengan berbagai algoritma dan ukuran jarak, serta data lebih komprehensif yang mencakup delapan variabel layanan keagamaan. Uji ANOVA menunjukkan perbedaan

signifikan antar klaster, yang memperkuat validitas hasil *clustering* dan memberikan gambaran kualitas layanan untuk mendukung kebijakan yang lebih tepat sasaran.

Fokus utama penelitian ini adalah mengelompokkan wilayah di Indonesia berdasarkan karakteristik layanan keagamaan menggunakan metode *hierarchical clustering* dengan berbagai algoritma dan ukuran jarak, serta menguji signifikansi perbedaan antar klaster dengan uji ANOVA (Lawuna et al., 2024). Hasil penelitian diharapkan memberikan kontribusi penting dalam analisis statistik yang mendalam untuk mendukung pengambilan keputusan berbasis bukti di Kementerian Agama (Muhammad, 2024). Penelitian ini berpotensi membantu merumuskan kebijakan pelayanan keagamaan yang lebih tepat sasaran dan mendukung transformasi data menjadi kebijakan yang sesuai dengan kebutuhan daerah, menghasilkan kebijakan yang lebih inklusif dan berkelanjutan.

2. METODE PENELITIAN

2.1 Data dan Sumber Data

Penelitian ini menggunakan data sekunder sebagai sumber informasi yang diperoleh dari publikasi resmi "Statistik Kementerian Agama 2024" yang mencakup 34 provinsi di Indonesia. Delapan variabel utama yang menggambarkan berbagai aspek layanan keagamaan digunakan, dengan rincian variabel yang disajikan pada Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Keterangan	Variabel	Keterangan
x_1	Jumlah Unit Kerja Daerah	x_5	Jumlah Bangunan
x_2	Jumlah Masjid Terdaftar	x_6	Jumlah Guru Pada MI, MTs, dan MA
x_3	Jumlah Kantor Urusan Agama	x_7	Jumlah Pondok Pesantren
x_4	Jumlah Satuan Pendidikan	x_8	Jumlah PNS Kantor Wilayah

2.2 Tahapan Analisis

Data dianalisis dengan *hierarchical clustering* menggunakan kombinasi lima algoritma dan tiga ukuran jarak untuk mengelompokkan daerah berdasarkan layanan keagamaan. Uji ANOVA dilakukan untuk mengidentifikasi perbedaan signifikan antar klaster, memperkuat hasil *clustering* dan memberikan dukungan empiris. Langkah-langkah lebih lengkap dijelaskan berikut ini.

1. Melakukan analisis awal menggunakan statistik deskriptif untuk memberikan gambaran umum mengenai data.
2. Melakukan standardisasi Z_{score}
Untuk menghindari bias yang disebabkan oleh perbedaan skala variabel, dilakukan standardisasi data (Yafi et al., 2023).
3. Melakukan pengelompokan data dengan menerapkan kelima algoritma *hierarchical clustering*.

Secara umum, tahapan pengelompokan data menggunakan kelima algoritma memiliki kesamaan.

- a. Menghitung ukuran jarak kemiripan, dalam penelitian ini menggunakan tiga ukuran jarak kemiripan, yaitu *euclidean*, *manhattan*, dan *chebyshev*.
- b. Menggabungkan dua kluster yang memiliki jarak terkecil berdasarkan perhitungan kemiripan antar objek dalam kluster.
- c. Menghitung jarak antar kluster berdasarkan kelima algoritma. Menghitung jarak maksimum antar kluster berdasarkan algoritma *complete linkage* berdasarkan Persamaan 1.

$$d_{(AB)C} = \max(d_{AC}, d_{BC}) \quad (1)$$

Menghitung jarak minimum antar kluster berdasarkan algoritma *single linkage* berdasarkan Persamaan 2.

$$d_{(AB)C} = \min(d_{AC}, d_{BC}) \quad (2)$$

Menghitung jarak rata-rata antar kluster berdasarkan algoritma *average linkage* berdasarkan Persamaan 3.

$$d_{(AB)C} = \frac{d_{AC} + d_{BC}}{n_{AB}n_C} \quad (3)$$

(Dewi & Ahadiyah, 2022)

Menghitung jarak antar *centroid* (titik pusat) kedua kluster yang digabungkan berdasarkan Persamaan 4.

$$d_{(AB)} = \frac{n_A \bar{x}_A + n_B \bar{x}_B}{n_A + n_B} \quad (4)$$

(Candra, 2024)

Menghitung jarak antara kluster dihitung dengan minimasi varian (*variance minimization*) berdasarkan Persamaan 5.

$$SSE = \sum_{i=1}^n (X_i - \bar{X})' - (X_i - \bar{X}) \quad (5)$$

(Irwan et al., 2024)

- d. Menggabungkan kluster dengan ukuran jarak terkecil setelah perhitungan masing-masing algoritma. Pada algoritma *centroid linkage*, setelah langkah (d) dilakukan perhitungan ulang ukuran jarak kemiripan.
- e. Mengulang proses dari langkah (a) hingga (d) untuk menggabungkan kluster dengan jarak terdekat, sampai jumlah kluster yang diinginkan tercapai.
- f. Menginterpretasikan hasil kluster.

4. Menghitung koefisien korelasi *cophenetic* untuk mengukur hubungan antara elemen dalam matriks jarak asli dan matriks *cophenetic*, guna menentukan algoritma hierarki yang optimal. Rumusnya disajikan pada Persamaan 6.

$$r_{coph} = \frac{\sum_{i < k} (d_{ik} - \bar{d})(d_{cik} - \bar{d}_c)}{\sqrt{\left[\sum_{i < k} (d_{ik} - \bar{d})^2\right] \left[\sum_{i < k} (d_{cik} - \bar{d}_c)^2\right]}} \quad (6)$$

Koefisien korelasi *cophenetic* antara 0 dan 1, mengukur seberapa baik *clustering* mencerminkan jarak asli antar objek. Nilai menghampiri 1 menunjukkan hasil yang lebih baik.

(Azari et al., 2024)

5. Menentukan jumlah kluster optimal dengan menghitung nilai *silhouette coefficient* (*SC*) menggunakan algoritma yang terbaik. Rumus *SC* disajikan pada Persamaan 7.

$$SC(i) = \frac{p_i - r_i}{\max\{p_i - r_i\}} \quad (7)$$

dengan

- r_i : *mean* jarak objek ke- i dengan semua objek lainnya dalam kluster yang sama
 p_i : *mean* jarak objek ke- i dengan semua objek yang ada di kluster yang berbeda

Nilai *SC* memiliki rentang antara -1 hingga 1, dengan nilai menghampiri 1 menunjukkan kluster terstruktur, dan -1 menunjukkan kluster tidak terstruktur.

(Putra et al., 2023)

6. Melakukan uji *Analysis of Variance* (ANOVA) untuk menguji signifikansi dan memvalidasi hasil kluster yang terbentuk.

Uji ANOVA satu arah digunakan untuk mengidentifikasi pengaruh variasi satu faktor terhadap variabel dependen (Utami et al., 2023). Pada penelitian kluster ini, uji ANOVA dilakukan sebanyak 8 kali untuk 8 variabel, dengan perhitungan lengkap disajikan pada Tabel 2.

Tabel 2. Uji ANOVA Satu Arah

Sumber Keberagaman	Jumlah Kuadrat	Derajat Bebas	Rata-Rata Jumlah Kuadrat	F_{hitung}
Perlakuan	$SSB = \frac{1}{n} \sum_{j=1}^k \sum_{i=1}^n y_{ij}^2 - \frac{y_{..}^2}{N}$	$df_1 = k - 1$	$MSB = \frac{SSB}{k - 1}$	$f = \frac{MSB}{MSE}$
Galat	$SSE = SST - SSB$	$df_2 = N - k$	$MSE = \frac{SSE}{N - k}$	
Total	$SST = \sum_{j=1}^k \sum_{i=1}^n y_{ij}^2 - \frac{y_{..}^2}{N}$	$df_3 = N - 1$		

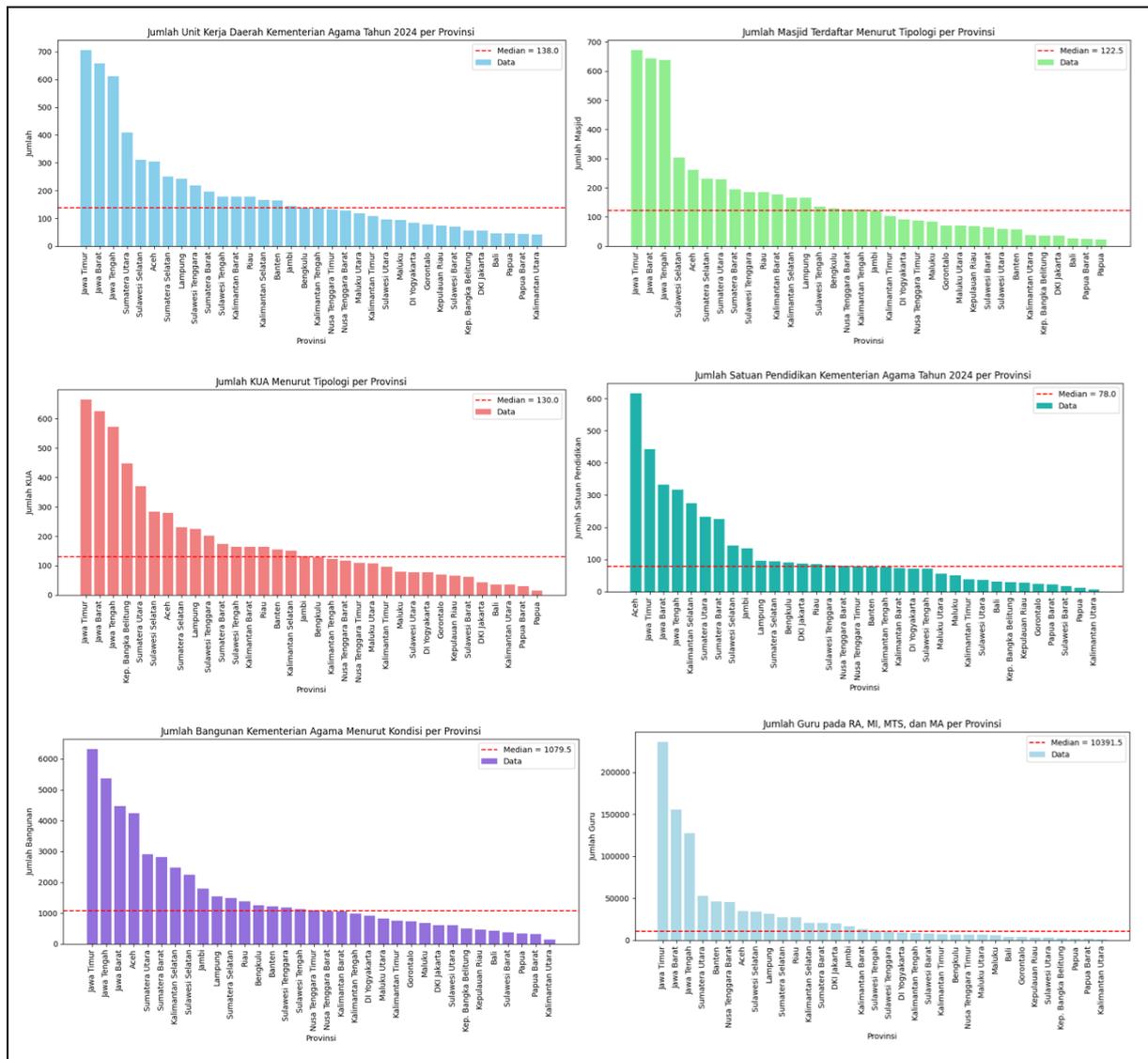
(Sugiyono, 2018)

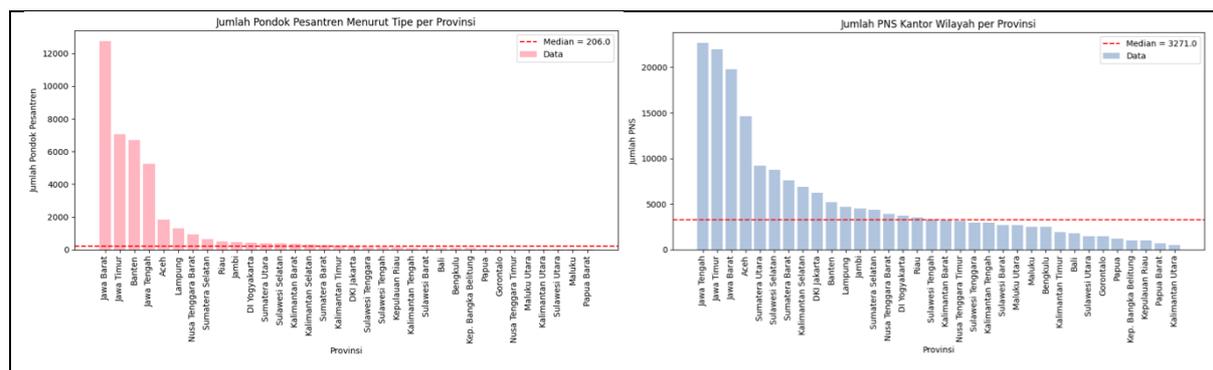
7. Pemetaan dan interpretasi hasil pengelompokan.

3. HASIL DAN PEMBAHASAN

3.1 Gambaran Umum Data Penelitian

Pada penelitian ini, data digambarkan secara deskriptif melalui diagram batang untuk setiap variabel yang diteliti. Setiap diagram dilengkapi dengan garis rata-rata guna menunjukkan kecenderungan umum dari data. Visualisasi ini memberikan gambaran awal terhadap pola distribusi masing-masing variabel. Diagram batang tersebut disajikan pada Gambar 1.





Gambar 1. Diagram Batang Setiap Variabel

Secara umum, provinsi-provinsi besar seperti Jawa Barat, Jawa Timur, dan Jawa Tengah menunjukkan kualitas layanan keagamaan yang lebih unggul dibandingkan dengan provinsi-provinsi lainnya. Hal ini terlihat dari tingginya jumlah unit kerja, masjid, Kantor Urusan Agama (KUA), serta satuan pendidikan dan jumlah PNS yang jauh lebih banyak. Kondisi ini mencerminkan tingginya kebutuhan akan layanan keagamaan di daerah dengan populasi besar dan tingkat urbanisasi yang tinggi.

Namun, disisi yang berbeda provinsi-provinsi seperti Papua dan Papua Barat memiliki angka yang jauh lebih rendah di hampir semua variabel tersebut. Hal ini mengindikasikan adanya ketimpangan yang signifikan dalam hal infrastruktur dan sumber daya. Tantangan geografis dan demografis yang unik di kedua provinsi tersebut mungkin menjadi faktor utama yang memengaruhi kemampuan mereka dalam menyediakan layanan keagamaan yang optimal. Secara keseluruhan, data ini menunjukkan kesenjangan yang besar antara provinsi-provinsi besar dan provinsi-provinsi kecil atau terpencil di Indonesia dalam hal pengelolaan dan distribusi layanan keagamaan.

3.2 Pengelompokan Menggunakan Kelima Algoritma *Hierarchical Clustering*

Pada tahap ini, dilakukan pengelompokan menggunakan lima algoritma, yaitu *complete linkage*, *single linkage*, *average linkage*, *centroid*, dan *ward*. Selain itu, dilakukan perbandingan antara tiga jenis jarak kemiripan, yaitu *euclidean*, *manhattan*, dan *chebyshev*, pada masing-masing algoritma, sehingga diperoleh total 15 hasil pengelompokan. Setelah melalui beberapa tahapan pengelompokan menggunakan bantuan *software* R, dihasilkan 15 dendrogram yang menggambarkan hasil-hasil tersebut.

3.3 Pemilihan Algoritma Terbaik Berdasarkan Koefisien Korelasi *Cophenetic*

Penelitian ini menggunakan perhitungan koefisien korelasi *cophenetic* untuk menentukan algoritma dan jarak kemiripan terbaik, berdasarkan pada Persamaan (6). Perhitungan koefisien korelasi *cophenetic* menggunakan bantuan *software* R secara lengkap disajikan pada Tabel 3.

Tabel 3. Nilai Koefisien Korelasi *Cophenetic*

Algoritma	Jarak Kemiripan		
	<i>Euclidean</i>	<i>Manhattan</i>	<i>Chebyshev</i>
<i>Single Linkage</i>	0,952	0,959	0,949
<i>Complete Linkage</i>	0,909	0,942	0,949
<i>Average Linkage</i>	0,967	0,963	0,969
<i>Centroid</i>	0,962	0,961	0,955
<i>Ward</i>	0,930	0,941	0,946

Catatan: Nilai yang dicetak tebal merupakan nilai tertinggi

Berdasarkan Tabel 3, kombinasi *average linkage* dengan jarak *chebyshev* (0,969) memiliki nilai koefisien korelasi *cophenetic* tertinggi, menunjukkan kluster yang paling optimal. *Single linkage* dengan jarak *manhattan* (0,959) dan *centroid* dengan jarak *euclidean* (0,962) juga memberikan hasil baik, namun sedikit lebih rendah. Sementara itu, algoritma dengan jarak kemiripan lainnya menunjukkan nilai yang juga lebih rendah, menandakan bahwa kombinasi ini kurang optimal. Dengan demikian, *average linkage* dengan *chebyshev* adalah kombinasi terbaik.

3.4 Uji Validitas Kluster dengan *Silhouette Coefficient* pada Kombinasi Terbaik

Perhitungan *silhouette coefficient* digunakan untuk menentukan kluster yang optimal dari algoritma dan jarak kemiripan terbaik (algoritma *average linkage* dengan jarak kemiripan *chebyshev*), berdasarkan pada Persamaan (7). Perhitungan nilai *silhouette coefficient* menggunakan bantuan *software R* dapat dilihat pada Tabel 4.

Tabel 4. Nilai *Silhouette Coefficient* Masing-Masing Kluster

Kluster	SC
2	0,736
3	0,659
4	0,411
5	0,394

Catatan: Nilai yang dicetak tebal merupakan nilai tertinggi

Berdasarkan Tabel 4, kluster 2 dengan SC tertinggi 0,736 menunjukkan kluster yang paling optimal. Kluster 3 masih cukup baik dengan nilai 0,659, sementara kluster 4 dan 5 menunjukkan penurunan kualitas kluster dengan nilai 0,411 dan 0,394. Oleh karena itu, kluster 2 dipilih sebagai kluster yang optimal.

Berdasarkan nilai SC, diperoleh 2 kluster yang optimal, dengan profilisasi anggota tiap kluster disajikan pada Tabel 5.

Tabel 5. Profilisasi Masing-Masing Kluster

Kluster	Anggota
1	Aceh, Jawa Barat, Jawa Timur, Jawa Tengah

2	Seluruh Provinsi di Pulau Sulawesi, Pulau Kalimantan, dan Pulau Papua. Provinsi-Provinsi di Pulau Sumatera Kecuali Aceh, dan Sebagian Provinsi di Pulau Jawa diantaranya DKI Jakarta, Banten, dan D. I Yogyakarta
---	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Hasil analisis menunjukkan bahwa kedua klaster yang terbentuk memiliki ciri khas masing-masing berdasarkan variabel-variabel yang digunakan dalam penelitian. Proses identifikasi karakteristik setiap klaster dilakukan menggunakan *software* R, dengan mengacu pada nilai rata-rata dari setiap variabel. Informasi lengkap mengenai karakteristik klaster tersebut disajikan pada Tabel 6.

Tabel 6. Karakteristik Variabel pada Setiap Klaster

Variabel	Klaster		Variabel	Klaster	
	Klaster 1	Klaster 2		Klaster 1	Klaster 2
x_1	569.750	139.333	x_5	5095.500	1110.000
x_2	553.500	113.467	x_6	138508.500	15293.100
x_3	535.750	137.867	x_7	6726.000	479.400
x_4	426.500	80.833	x_8	19762.250	3525.400

Catatan: Nilai yang dicetak tebal menunjukkan nilai tertinggi untuk masing-masing variabel di antara kedua klaster

3.5 Uji ANOVA untuk Validasi Hasil *Clustering*

Penelitian ini menambahkan inovasi dengan melakukan uji ANOVA untuk memvalidasi hasil *clustering* dan menguji perbedaan yang signifikan antar klaster yang terbentuk. Perhitungan uji ANOVA dilakukan menggunakan *software* R, yang dapat dilihat pada Tabel 8.

Tabel 8. Nilai *P-value* Uji ANOVA Masing-Masing Variabel

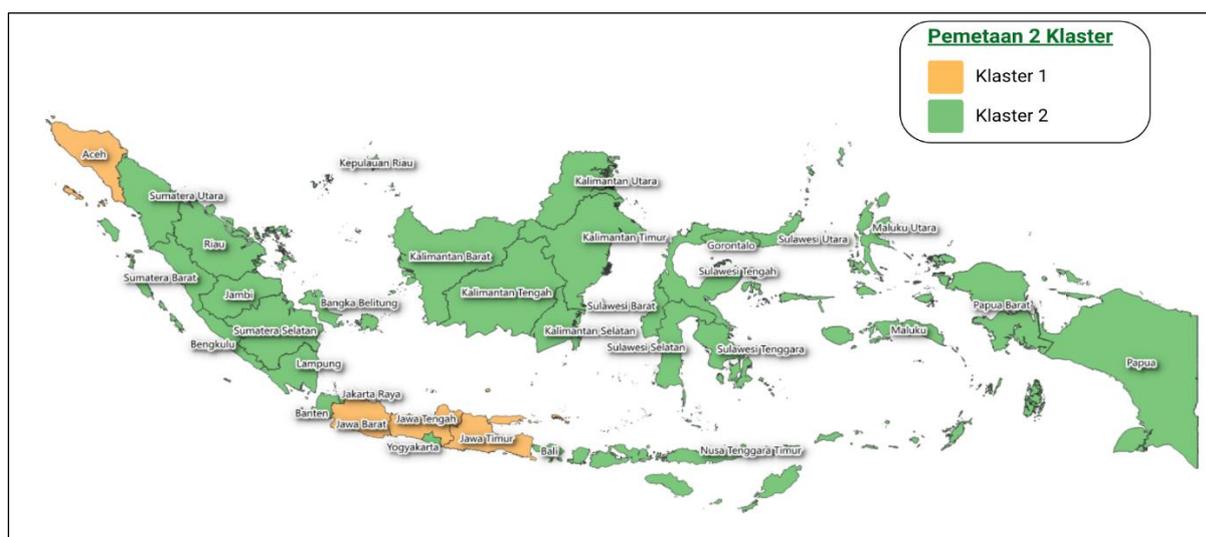
Variabel	<i>p-value</i>	Keputusan
x_1	$2,210 \times 10^{-9}$	Tolak H_0
x_2	$2,020 \times 10^{-10}$	Tolak H_0
x_3	$7,960 \times 10^{-8}$	Tolak H_0
x_4	$7,130 \times 10^{-10}$	Tolak H_0
x_5	$2,220 \times 10^{-11}$	Tolak H_0
x_6	$3,880 \times 10^{-9}$	Tolak H_0
x_7	$2,680 \times 10^{-7}$	Tolak H_0
x_8	$9,650 \times 10^{-14}$	Tolak H_0

Berdasarkan Tabel 8, hasil uji ANOVA untuk memvalidasi hasil *clustering* menunjukkan nilai *p-value* yang sangat kecil pada semua variabel yang diuji. Setiap nilai *p-value* untuk semua variabel jauh dibawah tingkat signifikansi $\alpha = 0,05$, yang menyebabkan H_0 (hipotesis nol) ditolak pada semua variabel. Oleh karena itu, dapat disimpulkan bahwa terdapat perbedaan yang signifikan antara klaster-klaster yang terbentuk., yang mengindikasikan bahwa hasil *clustering* valid dan dapat diterima.

3.6 Pemetaan Klaster Optimal

Berdasarkan keseluruhan tahapan analisis kluster yang telah dilakukan, termasuk uji validitas menggunakan ANOVA (*Analysis of Variance*), diperoleh bahwa pembentukan dua kluster merupakan hasil yang paling optimal. Hasil uji ANOVA menunjukkan bahwa perbedaan antara kluster signifikan, yang mengindikasikan bahwa masing-masing kluster memiliki karakteristik yang berbeda secara statistik.

Profil masing-masing kluster disajikan secara rinci pada Tabel 6, yang menunjukkan perbedaan pola karakteristik antar provinsi berdasarkan variabel yang dianalisis. Selanjutnya, untuk memberikan visualisasi geografis dari distribusi kluster tersebut, dilakukan pemetaan spasial yang ditampilkan pada Gambar 2.



Gambar 2. Pemetaan Profil Kluster Terbaik

Berdasarkan Gambar 2, diperoleh dua kluster optimal. Kluster pertama mencakup Provinsi Aceh, Jawa Timur, Jawa Barat, dan Jawa Tengah, sementara kluster kedua meliputi provinsi lainnya. Temuan dalam Tabel 8 menunjukkan bahwa kluster pertama memiliki rata-rata nilai lebih tinggi dalam semua variabel, termasuk layanan keagamaan, infrastruktur, dan pendidikan berbasis agama. Menariknya, DKI Jakarta, meskipun ibu kota, tergolong dalam kluster kedua dengan nilai lebih rendah dibandingkan kluster pertama. Pembentukan kluster ini memvalidasi perlunya kebijakan yang disesuaikan dengan segmentasi daerah, mengingat perbedaan karakteristik regional (Rizaldy & Asnani, 2024). Ini menunjukkan bahwa layanan keagamaan di DKI Jakarta dan beberapa provinsi dalam kluster dua memiliki potensi untuk diperbaiki. Temuan ini menjadi perhatian penting bagi Kementerian Agama untuk terus meningkatkan kualitas layanan keagamaan di seluruh Indonesia.

4. KESIMPULAN

Berdasarkan temuan analisis, penelitian ini menerapkan metode *hierarchical clustering* dengan algoritma *average linkage* dan jarak *chebyshev*, yang terbukti paling optimal dengan nilai koefisien korelasi cophenetic sebesar 0,969. Uji *silhouette coefficient* mengidentifikasi dua kluster optimal, yang validasinya diperkuat melalui uji ANOVA

yang menunjukkan perbedaan signifikan antar klaster. Klaster pertama terdiri dari Provinsi Aceh, Jawa Tengah, Jawa Timur, dan Jawa Barat, yang memiliki kualitas layanan keagamaan lebih baik, sementara klaster kedua mencakup provinsi lainnya, termasuk DKI Jakarta, yang masih menunjukkan bahwa masih ada peluang untuk peningkatan kualitas layanan keagamaan. Temuan ini memberikan gambaran penting bagi Kementerian Agama untuk terus berupaya meningkatkan dan memajukan kualitas layanan keagamaan di seluruh wilayah Indonesia.

5. UCAPAN TERIMA KASIH

Terima kasih kami ucapkan kepada Kementerian Agama Republik Indonesia yang telah memfasilitasi penulis dan memberikan dukungan dalam pengembangan ilmu pengetahuan

6. REKOMENDASI

Penelitian masa depan juga dapat fokus pada algoritma *clustering* yang lebih *robust*, seperti DBSCAN, K-Means++, atau *clustering* berbasis *deep learning*, untuk menangani data yang lebih kompleks dan meningkatkan akurasi hasil pengelompokan.

7. REFERENSI

- Arifin. (2024). Analisis Terhadap Kebijakan Pemerintah dalam Mengatur Urusan Pemerintah Daerah di Era Desentralisasi Asimetris. *PROGRESIF: Jurnal Hukum*, XVIII(2), 208-235.
- Azari, H., Hartanti, D., & Sari, A. A. (2024). Pengelompokan Produksi Padi dan Beras Provinsi Jawa Timur dengan Metode *Agglomerative Hierarchical Clustering*. *Infotek: Jurnal Informatika dan Teknologi*, 7(2), 379-389.
- Candra, Y., & Dani, A. T. R. (2025). Mapping Crime-Prone Areas Using Principal Component Analysis (PCA) – Centroid Linkage. *Multica Science and Technology (MST)*, 1(1), 1-8.
- Candra, Y., Goejantoro, R., & Dani, A. T. R. (2024). Pengelompokan Provinsi Berdasarkan Indikator Ekonomi, Pendidikan, Kesehatan, dan Kriminalitas dengan Algoritma *Centroid Linkage*. *Jurnal Ilmiah Matematika, Sains, dan Teknologi*, 12(1), 1-7.
- Dewi, A. F., & Ahadiyah, K. (2022). *Agglomerative Hierarchy Clustering* pada Penentuan Kelompok Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Pendidikan. *Zeta – Math Journal*, 7(2), 57-63.
- Hariyah. (2016). Tren Penelitian Studi Islam dalam Jurnal Badan Litbang dan Diklat Kemenag RI: Penggunaan *Co-Words*. *Record and Library Journal*, 2(2), 162-175.
- Irwan, Sanusi, W., & Hasanah, A. (2024) Perbandingan Analisis *Cluster* Metode *Complete Linkage* dan Metode *Ward* dalam Pengelompokan Indeks Pembangunan Manusia di Sulawesi Selatan. *Journal of Mathematics, Computations, and Statistics*, 7(1), 75-86.
- Ismail, G. (2024) Penerapan Analisis Klaster dan Prediksi Indeks Pembangunan Manusia untuk Mengevaluasi Kualitas Hidup Manusia dalam Pembangunan Nasional. *Seminar Nasional; Official Statistics*, 559-569.

- Juliyanto, R. M., Patria, F., Pamungkas, H. D., & Nugraha, F. S., (2024). Penerapan Algoritma *Clustering* untuk Segmentasi Pelanggan Berdasarkan Perilaku Pembelian. *Seminar Nasional AMIKOM Surakarta (SEMNAS)*, e-ISSN: 3031-5581.
- Khairiah, & Walid, A. (2020). Pengelolaan Keberagaman Budaya Melalui Multilingualisme di Indonesia. *Fikri: Jurnal Kajian Agama, Sosial, dan Budaya*, 5(1), 131-143.
- Lawuna, M. P., Zega, Y., & Mendrofa, R. N. (2024). Analisis *Cluster* Mahasiswa Pendidikan Matematika Universitas Nias. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, 8(2), 946-961.
- Muhammad. (2024). Strategi Manajemen Pendidikan dalam Meningkatkan Kualitas Pendidikan di Era Digital. *At-Ta'lim: Jurnal Pendidikan Agama Islam*, 6(1), 67-78.
- Putra, F. B., Dani, A. T. R., & Wigantono, S (2023). Penerapan Algoritma *Hierarchical Clustering* dalam Pengelompokan Kabupaten/Kota di Papua Berdasarkan Indikator Kemiskinan. *Mathematics and Applications Journal*, 5(2), 118-126.
- Rizaldi, M. F., & Asnani. (2024). Analisis Penentuan dan Kebijakan Strategi Pengembangan Sektor Basis Wilayah Kabupaten Pesisir Barat. *Sosiologi: Jurnal Ilmiah Kajian Ilmu Sosial dan Budaya*, 26(1), 93-114.
- Siraj, A. (2018). Refungsionalisasi Aparat Pelayanan Keagamaan di Kantor Kementerian Agama Kota Makassar. *Jurnal Ilmiah Ilmu Administrasi Publik: Jurnal Pemikiran dan Pelayanan Administrasi Publik*, 8(1), 45-52.
- Siregar, A. M., Siregar, H. N. I., Fitria, A., Sastia, F., Dinatingias, T. F., & Yusnaldi, E. (2024). Kondisi Geografis dan Penduduk. *Jurnal Inovasi Pendidikan*, 7(12), 150-156.
- Sugiyono. (2018). *Metodologi Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta
- Suraya, G. R., & Wijayanto, A. W. (2022). Comparison of Hierarchical Clustering, K-Means, K-Medoids, and Fuzzy C-Means Method in Grouping Provinces in Indonesia according to the special Index for Handling Stunting. *Indonesian Journal of Statistics and its Application*, 6(2), 180-121.
- Utami, R. K., Windarti, S., & Muslim, M. (2023). Analisis Data Penyakit DBD dengan *K-Means Clustering* di Kabupaten Bantul Menggunakan *Data Mining*. *Jurnal Manajemen Informasi dan Administrasi Kesehatan (JMIAK)*, 6(2), 43-49.
- Yafi, M., Goejantoro, R., & Dani, A. T. R. (2023). Pengelompokan Algoritma K-Medoids dengan Principal Component Analysis (PCA). *Pras Seminar Nasional Matematika, Statistika, dan Aplikasinya*, 3(1), 183-195.
- Yahya, M. Z., Nurhaliza, S., Fathan, M. A., Rizal, M. E., & Harismahyanti, A. (2025). Perbandingan Ukuran Jarak Pada Analisis Kluster Hirarki. *Leibniz: Jurnal Matematika*, 5(2), 93-111.
- Yuniar, N., & Heikal, J. (2024). Segmentasi sertifikat halal pada wilayah Indonesia menggunakan K-means clustering. *Halal Ecosystem Journal*, 1(1), 23-42.