

Analisis Clustering Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Kesehatan Masyarakat dengan Metode K-Means dan Agglomerative Hierarchical Clustering.

Rivaldo Timoty Manihuruk^{1*}, Muhammad Khairii Sufyaan^{2*}, dan Zaky Izmi Syakura^{3*}

¹Statistika, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

²Statistika, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

³Statistika, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

*Corresponding author: 5003201037@student.its.ac.id, 5003201077@student.its.ac.id, dan 5003201148@student.its.ac.id

ABSTRAK – Kesehatan masyarakat memegang peran krusial dalam pencapaian tujuan pembangunan global dan nasional. Indonesia, melalui undang-undang dan keterlibatan dalam *Sustainable Development Goals* (SDGs), menegaskan pentingnya kesehatan sebagai elemen utama kesejahteraan umum. Penelitian ini bertujuan memahami kondisi kesehatan di tingkat kabupaten/kota di Jawa Timur dengan menggunakan metode *Agglomerative Hierarchical Clustering* (AHC). Metode ini dipilih karena kemampuannya membentuk kelompok homogen dan memberikan gambaran komprehensif terkait kesehatan di wilayah tersebut. Hasil pengelompokan menyoroti evaluasi metrik yang beragam antara metode *K-Means* dan AHC, *K-Means* menunjukkan *silhouette score* tertinggi pada 2 klaster, sementara AHC *Ward Linkage* dengan tiga kluster memberikan evaluasi yang komprehensif namun tidak optimal. Analisis *Ward Linkage* menghasilkan tiga kelompok berbeda dengan identifikasi sembilan variabel signifikan terhadap kondisi kesehatan masyarakat di Jawa Timur. Saran penelitian mendepankan variasi metode, evaluasi lebih lanjut terhadap variabel yang digunakan, dan memberikan fokus pada wilayah dengan indikasi kesehatan buruk untuk perumusan kebijakan yang lebih efektif di tingkat regional.

Kata kunci– Clustering, Kesehatan Masyarakat, Jawa Timur, K-Means, AHC

I. PENDAHULUAN

A. Latar Belakang

Pentingnya kesehatan masyarakat telah diakui secara luas sebagai elemen kunci dalam mencapai tujuan pembangunan nasional dan internasional. Undang-Undang No. 23 Tahun 1992 menggariskan bahwa kesehatan bukan hanya mencakup keadaan fisik, tetapi juga keberlanjutan kesejahteraan sosial dan ekonomi individu. Pemahaman ini sejalan dengan semangat Pembukaan Undang-Undang Dasar 1945 dan cita-cita bangsa Indonesia yang menegaskan pentingnya kesehatan sebagai unsur kesejahteraan umum.

Agenda kesehatan juga mendapat perhatian khusus dalam *Sustainable Development Goals* (SDGs) sebagai tujuan ketiga dari 17 agenda pada Tujuan Pembangunan Berkelanjutan (TPB). Tujuan tersebut mencakup jaminan kehidupan yang sehat dan peningkatan kesejahteraan hidup untuk seluruh lapisan masyarakat. Indonesia, sebagai negara yang mengesahkan SDGs pada sidang umum PBB tahun 2015, memiliki kewajiban untuk mewujudkan pembangunan kesehatan yang merata untuk seluruh penduduknya [1]. Melalui Undang-Undang No. 17 Tahun 2007, pembangunan kesehatan juga dinyatakan sebagai bagian integral dari Rencana Pembangunan Jangka Panjang Nasional (RPJPN) Tahun 2005 – 2025. Pembangunan kesehatan bukan hanya tanggung jawab pemerintah, tetapi juga melibatkan semua komponen masyarakat Indonesia untuk meningkatkan kesadaran, kemauan, dan kemampuan hidup sehat [2].

Untuk mencapai pemahaman yang mendalam mengenai kondisi kesehatan masyarakat di tingkat kabupaten/kota di Jawa Timur, diperlukan suatu metode yang mampu memberikan pengelompokan yang jelas dan terstruktur. Penelitian ini memiliki tujuan utama yaitu melakukan pengelompokan kabupaten/kota di Jawa Timur berdasarkan indikator kesehatan masyarakat dengan menggunakan metode *K-means* dan metode *Agglomerative Hierarchical Clustering* (AHC). Pemilihan metode AHC ini didasarkan pada kemampuannya menghasilkan kelompok-kelompok yang saling terkait secara hierarkis berdasarkan kesamaan karakteristik kesehatan masyarakat. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam pemahaman yang lebih baik mengenai keragaman kondisi kesehatan di tingkat kabupaten/kota di Jawa Timur, sekaligus memberikan dasar yang kuat untuk perumusan dan evaluasi kebijakan kesehatan yang lebih efektif dan efisien di tingkat regional.

B. Rumusan Masalah

1. Bagaimana pengelompokan wilayah Kab/Kota berdasarkan indikator kesehatan Masyarakat menggunakan metode *K-Means*.
2. Bagaimana pengelompokan wilayah Kab/Kota berdasarkan indikator Kesehatan Masyarakat menggunakan metode *Agglomerative Hierarchical Clustering* (AHC).
3. Bagaimana faktor-faktor yang berpengaruh signifikan terhadap analisis clustering menggunakan *One-Way*

MANOVA dan *One-Way* ANOVA.

C. Tujuan

1. Mengetahui penerapan metode *K-Means* untuk pengelompokan wilayah Kab/Kota berdasarkan indikator kesehatan masyarakat.
2. Mengetahui penerapan metode *Agglomerative Hierarchical Clustering* (AHC) untuk pengelompokan wilayah Kab/Kota berdasarkan indikator kesehatan masyarakat.
3. Mengetahui faktor-faktor apa saja yang berpengaruh signifikan terhadap analisis clustering menggunakan *One-Way* MANOVA dan *One-Way* ANOVA.

D. Manfaat

1. Bagi peneliti dapat mengetahui karakteristik pengelompokan wilayah Kab/Kota berdasarkan indikator kesehatan masyarakat dan faktor-faktor yang berpengaruh
2. Bagi pembaca untuk dapat melakukan penelitian yang serupa dengan landasan pada penelitian ini
3. Untuk pemerintah dapat mempertimbangkan pengelompokan untuk kepentingan Kesehatan dan faktor-faktor yang berpengaruh terhadap kesehatan masyarakat.

II. TINJAUAN PUSTAKA

A Statistika Deskriptif

Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Statistika deskriptif dapat disajikan dalam bentuk tabel, gambar, maupun grafik yang dapat digunakan untuk menjelaskan karakteristik suatu data [3]. Penyajian statistika deskriptif dapat dilakukan secara numerik melalui nilai tertentu ataupun secara visual. Variabel yang berskala data numerik biasanya disajikan melalui ukuran pemusatan data berupa rata-rata dan median, sedangkan melalui ukuran penyebaran data disajikan dalam nilai varians, minimum, dan maksimum. Kemudian untuk variabel yang berskala data kategorik disajikan melalui nilai modus yang merupakan nilai terbanyak dari kategori data terkait.

Jika secara visual, statistika deskriptif dapat disajikan melalui box plot untuk data numerik ataupun diagram lingkaran untuk data kategorik. Box plot merupakan salah satu cara untuk menggambarkan karakteristik data melalui ukuran yang meliputi nilai observasi terendah, kuartil pertama, median, kuartil ketiga, serta nilai tertinggi dari data numerik dan sangat efektif untuk menampilkan beberapa sampel secara bersama-sama dengan tujuan untuk membandingkan secara visual. Kemudian diagram lingkaran merupakan sebuah grafik yang mana segmen lingkaran menunjukkan frekuensi relatif dari masing-masing kategori data dan biasanya juga disajikan dalam bentuk persentase [4].

B. Analisis Multivariat

Multivariat berasal dari kata multi yang artinya banyak dan variat yang berarti variabel. Melalui definisi tersebut maka dapat dikatakan bahwa analisis multivariat merupakan analisis yang berkaitan dengan jumlah variabel yang banyak atau lebih dari satu yang dianalisis secara simultan pada masing-masing pengamatan dan antar variabel saling berhubungan [5]. Salah satu analisis statistik multivariat yang sering digunakan antara lain Principal Component Analysis (PCA), Analisis Faktor, Multiple Regression, Analisis Diskriminan, Analisis Korelasi Kanonikal, Analisis Kelompok/Klaster, Analisis Korespondensi, Multidimensional Scalling, dan Structural Equation Modelling (SEM) [6]. Adapun struktur data dari analisis multivariat ditampilkan pada **Tabel 1** berikut.

Tabel 1 Struktur Data Analisis Multivariat

Pengamatan ke- <i>i</i>	Variabel 1 (X_1)	Variabel 2 (X_2)	...	Variabel m (X_m)
1	X_{11}		...	X_{1m}
2	X_{21}		...	X_{2m}
3	X_{31}		...	X_{3m}
.
.
.
n	X_{n1}	X_{n2}	...	X_{nm}

C. Analisis Pengelompokan

Analisis kelompok (*cluster analysis*) adalah metode dalam analisis multivariat untuk mengelompokkan n pengamatan kedalam C kelompok ($C \leq n$) berdasarkan karakteristiknya. Analisis ini merupakan salah satu metode primitif sehingga tidak memerlukan adanya asumsi yang digunakan untuk melakukan pengelompokan data, karena analisis ini mengelompokkan objek-objek sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam satu kelompok yang sama serta mempunyai kemiripan satu dengan yang lain [5]. Hasil analisis kelompok dipengaruhi oleh

beberapa hal, antara lain yaitu objek yang dikelompokkan, variabel yang diamati, ukuran kemiripan dan ketidakmiripan, serta metode pengelompokan yang digunakan. Ukuran kemiripan dan ketidakmiripan merupakan hal yang sangat mendasar dalam kelompok analisis.

Teknik pengelompokan yang digunakan dalam penelitian ini adalah metode K-Means dan hirarki atau disebut juga Agglomerative Hierarchical Clustering (AHC). Penggunaan metode AHC digunakan karena banyaknya kelompok yang dibentuk tidak diketahui sebelumnya dan banyaknya amatan tidak terlalu besar [5]. Berikut merupakan beberapa metode dalam pengelompokan menggunakan AHC.

1. K-Means

Metode *K-Means* lebih sering digunakan karena prosesnya yang cepat dan sederhana. Algoritma *K-Means* akan membagi data menjadi beberapa kelompok, dimana setiap kelompok memiliki nilai tengah (*centroid*) dari data-data dalam kelompok tersebut yang dibentuk secara acak. Kemudian, dengan pendekatan jarak *Euclidean* akan dihitung jarak antara data dengan *centroid*. Selanjutnya, akan ditempatkan ke dalam kelompok terdekat yang dihitung dari titik tengah kelompok. Proses penentuan *centroid* dan penempatan data kelompok dilakukan hingga nilai *centroid* menjadi *konvergen*. Adapun langkah-langkah algoritma *K-Means Clustering* adalah sebagai berikut [7].

- Menentukan k sebagai jumlah *cluster* yang ingin dibentuk.
- Mengalokasikan data ke dalam *cluster* secara acak.
- Menentukan *centroid* dari masing-masing *cluster* dengan Persamaan (2.1) sebagai berikut.

$$C_{ij} = \frac{1}{n_i} \sum_{k=1}^{n_i} X_{kj} \quad (2.1)$$

- Menentukan jarak antara masing-masing objek dan *centroid* dengan perhitungan menggunakan jarak *Euclidean* dengan Persamaan (2.2) sebagai berikut.

$$d(x_i, x_k) = \sqrt{\sum_{j=1}^p k(x_{ij} - x_{kj})^2} \quad (2.2)$$

- Menghitung fungsi objektif dengan Persamaan (2.2) sebagai berikut.

$$J = \sum_{k=1}^{n_i} \sum_{i=1}^c a_{ik} d(x_i, x_k)^2 \quad (2.3)$$

- Mengalokasikan masing-masing data ke *centroid* dirumuskan dengan Persamaan (2.4) sebagai berikut.

$$a_{ik} = \begin{cases} 1, & s = \min\{d(x_k, C_i)\} \\ 0, & \text{lainnya} \end{cases} \quad (2.4)$$

- Mengulangi kembali langkah c-f hingga tidak ada lagi perpindahan objek atau perubahan pada fungsi objektifnya.

2. Single Linkage (Pautan Tunggal)

Single linkage merupakan pengelompokan yang didasarkan pada jarak terdekat atau kesamaan yang banyak. Jika dua objek terpisah oleh jarak yang dekat maka objek tersebut digabungkan menjadi satu kelompok dan demikian seterusnya. Ukuran jarak yang digunakan untuk metode single linkage adalah seperti pada persamaan (2.2).

$$d_{(UV)W} = \min \{d_{UV}, d_{VW}\} \quad (2.5)$$

3. Complete Linkage (Pautan Lengkap)

Complete linkage merupakan pengelompokan yang didasarkan pada jarak terjauh atau kesamaan yang sedikit. Jika dua objek terpisah oleh jarak yang jauh maka objek tersebut digabungkan menjadi satu kelompok dan demikian seterusnya. Ukuran jarak yang digunakan untuk metode complete linkage ditunjukkan pada persamaan (2.6).

$$d_{(UV)W} = \max \{d_{UV}, d_{VW}\} \quad (2.6)$$

4. Average Linkage (Pautan Rataan)

Average linkage adalah metode dimana kelompok dibentuk berdasarkan nilai rata-rata seluruh individu pada kelompok lain. Persamaan (2.7) menunjukkan ukuran jarak yang digunakan untuk metode average linkage.

$$d_{(UV)W} = \frac{1}{N_{UV}N_W} \sum_q \sum_r d_{qr} \quad (2.7)$$

Dimana,

d_{UV} : jarak kelompok U dan kelompok W

d_{VW} : jarak kelompok V dan kelompok W

$d_{(UV)W}$: jarak minimum antara kelompok UV dan kelompok W

N_{UV} : jumlah pengamatan pada kelompok UV

N_W : jumlah pengamatan pada kelompok W

d_{qr} : jarak antar observasi ke- q dalam kelompok UV dengan observasi ke- r dalam kelompok W

5. Ward Linkage

Metode ward merupakan pengelompokkan objek-objek didasarkan dengan memaksimalkan kehomogenan dalam kelompok. Dalam metode ini jarak antara dua cluster adalah jumlah kuadrat antara dua cluster untuk seluruh variabel. Pengelompokan metode ward adalah meminimalkan peningkatan kriteria error sum of square (ESS). Dua

cluster yang memiliki peningkatan ESS paling minimum akan berkelompok. Sehingga untuk menghitung jarak antara duacluster menggunakan metode ward diperoleh berdasarkan persamaan (2.8).

$$ESS = \sum_{j=1}^N (x_j - \bar{x})(x_j - \bar{x}) \quad (2.8)$$

dimana x_j merupakan pengukuran multivariat terkait dengan item j dan \bar{x} adalah rata-rata dari semua item. Metode ward's digunakan karena mampu meminimalkan varians dalam setiap cluster [8].

D. One-Way MANOVA

MANOVA merupakan perluasan dari teknik univariat ANOVA yang melibatkan lebih dari satu variabel dependen (Johnson & Wicherin, 2007). Pengujian hipotesis pada MANOVA adalah sebagai berikut.

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_c$$

$$H_1 : \text{minimal ada satu } \mu_i \text{ yang tidak sama, dimana } i = 1, 2, \dots, C$$

Statistik Uji Statistik uji yang digunakan dalam pengambilan keputusan pada perbedaan antar kelompok adalah wilk's lamda. Nilai statistik uji wilk's lamda berkisar antara 0 dan 1. Semakin rendah statistik uji, maka perbedaan antar kelompok semakin signifikan. Rumusan untuk statistik uji wilk's lamda, didefinisikan pada persamaan (2.9).

$$\Lambda^* = \frac{|W|}{|B + W|} \quad (2.9)$$

Dimana W dan B didefinisikan oleh persamaan (2.10) dan (2.11).

$$W = \sum_{c=1}^C \sum_{j=1}^{n_c} (x_{cj} - \bar{x}_c)(x_{cj} - \bar{x}_c)^T \quad (2.10)$$

$$B = \sum_{c=1}^C (\bar{x}_c - \bar{x})(\bar{x}_c - \bar{x})^T \quad (2.11)$$

Nilai statistik uji wilk's lamda ditransformasi menjadi statistik F sehingga dapat dilakukan perbandingan dengan tabel F. Bentuk transformasi disajikan dalam Tabel 2 berikut.

Tabel 2.2 Struktur Data Analisis Multivariat

Jumlah Variabel (m_{numerik})	Jumlah Kelompok yang Terbentuk	Distribusi Sampling
$m_{\text{numerik}} = 1$	$C \geq 2$	$\left(\frac{\sum n_c - C}{C - 1} \right) \left(\frac{1 - \Lambda^*}{\Lambda^*} \right) \sim F_{C-1; \sum n_c - C}$
$m_{\text{numerik}} = 2$	$C \geq 2$	$\left(\frac{\sum n_c - C - 1}{C - 1} \right) \left(\frac{1 - \sqrt{\Lambda^*}}{\Lambda^*} \right) \sim F_{2(C-1); 2(\sum n_c - C - 1)}$
$m_{\text{numerik}} \geq 1$	$C = 2$	$\left(\frac{\sum n_c - m_{\text{numerik}} - 1}{m_{\text{numerik}}} \right) \left(\frac{1 - \Lambda^*}{\Lambda^*} \right) \sim F_{m_{\text{numerik}}; \sum n_c - m_{\text{numerik}} - 1}$
$m_{\text{numerik}} \geq 1$	$C = 3$	$\left(\frac{\sum n_c - m_{\text{numerik}} - 2}{m_{\text{numerik}}} \right) \left(\frac{1 - \sqrt{\Lambda^*}}{\Lambda^*} \right) \sim F_{2m_{\text{numerik}}; 2(\sum n_c - m_{\text{numerik}} - 2)}$

Daerah Kritis Tolak H_0 jika $F > F_{C-1; \sum n_c - C; \alpha/2}$ atau jika $p\text{-value} < \alpha$, dimana α adalah taraf signifikansi yang digunakan. Tolak H_0 artinya minimal terdapat satu populasi yang menghasilkan rata-rata yang berbeda.

E. One-Way ANOVA

ANOVA (Analysis of Variance) digunakan untuk membandingkan rata-rata dari beberapa populasi yang diwakili oleh beberapa kelompok sampel secara bersama. Pengujian hipotesis yang digunakan dalam ANOVA adalah sebagai berikut:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_C$$

H_1 : minimal ada satu μ_i yang tidak sama, dimana $i = 1, 2, \dots, C$

Persamaan (2.12) menampilkan formulasi untuk statistik uji ANOVA.

$$F = \frac{MSB}{MSW} \quad (2.12)$$

Dimana MSB merupakan MSW yang diperoleh melalui persamaan (2.13) dan (2.14)

$$MSB = \frac{(SST - SSW)}{C - 1} \quad (2.13)$$

$$MSW = \frac{SSW}{n - C} \quad (2.14)$$

Adapun SST dan SSW diperoleh melalui persamaan (2.15) dan (2.16)

$$SST = \sum_{l=1}^{m_{\text{numerik}}} \sum_{i=1}^n (x_{il} - \bar{x}_l)^2 \quad (2.15)$$

$$SSW = \sum_{c=1}^C \sum_{l=1}^{m_{\text{numerik}}} \sum_{i=1}^n (x_{ilc} - \bar{x}_{lc})^2 \quad (2.16)$$

Daerah kritis Tolak H_0 jika $F > F_{C-1; n-C; \alpha/2}$ atau jika $p\text{-value} < \alpha$, dimana α adalah taraf signifikansi yang digunakan. Tolak H_0 artinya minimal terdapat satu populasi yang memiliki nilai berbeda atau dapat dikatakan juga rata-rata seluruh populasi tidak sama

F. Analisis Multikolinearitas

Pada penelitian ini harus memenuhi asumsi bahwa tidak adanya gejala multikolineritas atau adanya korelasi antara variabel independen yang menyebabkan sulitnya memberikan pengaruh individual. Uji korelasi antar variabel bebas untuk analisis ini dilakukan untuk mengetahui apakah ada multikolineritas pada data [9]. Jika variabel bebas mempunyai korelasi yang kuat, maka dapat dikatakan menjadi multikolineritas. Pengujian dapat dilakukan dengan melihat nilai korelasi pada masing-masing variabel independen. Kriteria pengambilan keputusan terkait uji multikolinearitas adalah sebagai berikut: [10].

- Jika nilai VIF < 10 atau nilai *Tolerance* > 0,10, maka dinyatakan tidak terjadi multikolinearitas.
- Jika nilai VIF > 10 atau nilai *Tolerance* < 0,10, maka dinyatakan terjadi multikolinearitas.
- Jika koefisien korelasi masing-masing variabel bebas > 0,8 maka terjadi multikolinearitas. Tetapi jika koefisien korelasi masing-masing variabel bebas < 0,8 maka tidak terjadi multikolinearitas

G. Kondisi Kesehatan Masyarakat Provinsi Jawa Timur

Provinsi Jawa Timur, terletak di timur Pulau Jawa, membentang dengan luas wilayah daratan mencapai 47.803,39 km². Dengan batas wilayahnya melibatkan Laut Jawa di utara, Samudera Hindia di selatan, Provinsi Jawa Tengah di barat, dan Selat Bali di timur. Pulau Madura, yang terhubung dengan daratan melalui jembatan Suramadu, menjadi pulau terbesar dari 508 pulau di provinsi ini. Kabupaten Banyuwangi memiliki wilayah terluas, dan secara administratif, provinsi ini terdiri dari 29 kabupaten, 9 kota, dan 8.502 desa/kelurahan.

Penduduk Provinsi Jawa Timur pada tahun 2022 mencapai 40.348.441 jiwa, terbagi menjadi 19.912.535 laki-laki dan 20.435.906 perempuan. Kota Surabaya menjadi pusat dengan jumlah penduduk tertinggi (2.928.058 jiwa), sementara Kota Mojokerto memiliki jumlah penduduk paling sedikit (132.045 jiwa). Kepadatan penduduk di kota, terutama Surabaya, lebih tinggi dibandingkan dengan kabupaten, mencerminkan dinamika populasi yang perlu diperhatikan.

Provinsi Jawa Timur juga memiliki sistem kesehatan yang berkembang. Pusat Kesehatan Masyarakat (Puskesmas) di tingkat kecamatan mencapai jumlah 969 pada tahun 2022, dengan 625 puskesmas rawat inap dan 344 puskesmas non-rawat inap. Meskipun terjadi penurunan jumlah puskesmas karena dua di antaranya beralih menjadi rumah sakit tipe D, gambaran pertumbuhan sejak 2017 hingga 2021 mencerminkan komitmen pemerintah dalam meningkatkan akses masyarakat terhadap pelayanan kesehatan primer. Rasio puskesmas per kecamatan, rata-rata 1,5 pada tahun 2022, mencerminkan aksesibilitas masyarakat terhadap pelayanan kesehatan primer, namun distribusinya di seluruh kecamatan masih memerlukan perhatian terkait faktor geografis, luas wilayah, sosial ekonomi, dan transportasi.

Selain itu, Provinsi Jawa Timur, dengan luas wilayah 47.803,39 km², terkenal bukan hanya akan geografinya, tetapi juga sistem kesehatannya yang berkembang. Salah satu wujud Upaya Kesehatan Bersumberdaya Masyarakat (UKBM) yang dekat dengan masyarakat adalah Pos Pelayanan Terpadu (Posyandu), terutama dalam mendekatkan upaya promotif dan preventif, khususnya terkait dengan peningkatan status gizi dan pelayanan kesehatan ibu dan anak. Untuk

menilai kelembagaan posyandu di Provinsi Jawa Timur, dilakukan telaah tingkat perkembangan posyandu Pratama, Madya, Purnama, dan Mandiri. Dalam Rencana Strategis Dinas Kesehatan Provinsi Jawa Timur Tahun 2022, diutamakan jumlah kabupaten/kota dengan $\geq 80\%$ Posyandu Purnama dan Mandiri (PURI). Meskipun pandemi Covid-19 berdampak pada kegiatan pembinaan posyandu dan pelayanan posyandu pada tahun 2021, situasi membaik pada tahun 2022. Kegiatan pelayanan posyandu telah berjalan sebanyak 12 kali/tahun, menyesuaikan kondisi pandemi dengan tetap menerapkan protokol kesehatan. Seiring dengan kembali normalnya pelayanan posyandu, terjadi tren kenaikan tingkat perkembangan Posyandu PURI. Selain itu, pelayanan yang diberikan juga sudah lengkap mencakup Kesehatan Ibu Anak (KIA), Keluarga Berencana (KB), imunisasi, gizi, dan pencegahan serta penanggulangan diare.

III. METODOLOGI

A. Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder Kesehatan Masyarakat di Jawa Timur pada tahun 2022 yang terdiri dari 15 Variabel. Data ini bersumber dari Laporan Profil Kesehatan pada tahun 2022 yang diproduksi oleh Dinas Kesehatan Provinsi Jawa Timur.

B. Variabel Data

Adapun variabel data yang digunakan dalam penelitian ini terdapat 15 variabel, yang disajikan dalam Tabel 3 berikut.

Tabel 3 Variabel Data Penelitian

Variabel	Keterangan
Kab/Kota	Kabupaten/Kota yang terdapat dalam Jawa Timur
Angka Kematian Bayi (AKB)	Indikator tingkat kematian bayi dalam suatu wilayah, menggambarkan kesehatan bayi pada periode tertentu.
Angka Kematian Ibu (AKI)	Indikator tingkat kematian ibu dalam suatu wilayah, mencerminkan kesehatan maternal di periode tertentu.
Persentase Balita Gizi Kurang	Pbalita dengan status gizi kurang dalam populasi.
Persentase Pemberian ASI Eksklusif	Proporsi bayi yang hanya menerima ASI sebagai sumber makanan utama.
Persentase Kematian Akibat Demam Berdarah Dengue (DBD)	Proporsi kematian yang disebabkan oleh demam berdarah dengue dalam populasi.
Jumlah Dokter Umum per 100.000 penduduk	Jumlah dokter umum dalam setiap 100.000 penduduk, mengukur ketersediaan tenaga medis umum.
Jumlah Bidan per 100.000 penduduk	Jumlah bidan dalam setiap 100.000 penduduk, mengukur ketersediaan tenaga medis spesialis ibu dan anak.
Jumlah Perawat per 100.000 penduduk	Jumlah perawat dalam setiap 100.000 penduduk, mengukur ketersediaan tenaga perawat.
Jumlah Puskesmas per 100.000 penduduk	Jumlah puskesmas dalam setiap 100.000 penduduk, mencerminkan ketersediaan fasilitas kesehatan masyarakat.
Jumlah Posyandu Aktif per 100.000 penduduk	Jumlah Posyandu yang aktif dalam setiap 100.000 penduduk, mencerminkan kegiatan pelayanan kesehatan masyarakat
Persentase Sumber Air Minum Layak	Proporsi penduduk yang memiliki akses ke sumber air minum yang memenuhi standar kesehatan.
Persentase Rumah Tangga yang Memiliki Akses Terhadap Sanitasi Layak	Proporsi rumah tangga dengan fasilitas sanitasi yang memenuhi standar kesehatan.
Persentase Desa/Kelurahan Universal Child Immunization (UCI)	Proporsi desa/kelurahan dengan pencapaian imunisasi anak universal.
Angka Keberhasilan Pengobatan TBC	Tingkat keberhasilan pengobatan penyakit Tuberkulosis (TBC) dalam suatu wilayah

C. Langkah Penelitian

Langkah pengerjaan yang dilakukan pada penelitian ini antara lain.

1. Mengidentifikasi data dengan melakukan statistika deskriptif pada masing-masing variabel
2. Melakukan analisis pengelompokkan dengan metode K-Means menggunakan penentuan jumlah kelompok dengan mempertimbangkan *Elbow Method* dan *Silhouette Score*.
3. Melakukan analisis pengelompokkan dengan metode Agglomerative Hierarchical Clustering (AHC)
4. Melakukan pengujian analisis multivariat menggunakan *One-way* MANOVA dan *One-way* ANOVA untuk mengetahui variabel-variabel yang signifikan pada pengelompokkan
5. Melakukan penarikan kesimpulan dengan menggambarkan karakteristik setiap kelompok

IV. HASIL DAN PEMBAHASAN

A. Statistika Deskriptif

Statistika deskriptif adalah suatu metode statistika yang digunakan untuk dapat memperoleh informasi dari data yang merupakan hasil pengamatan nyata. Statistika deskriptif dapat disajikan dalam bentuk ukuran pemusatan data

seperti nilai rata-rata (mean), nilai varians, dan lainnya. Statistika deskriptif juga dapat disajikan dalam bentuk tabel, diagram, dan lain sebagainya yang mampu mendeskripsikan data tersebut.

Tabel 4 Variabel Data Penelitian

Variabel	mean	std	min	max
AKB	83.447.368	59.140.098	2.000.000	282.000.000
AKI	13.131.579	10.203.791	1.000.000	58.000.000
Persentase Balita Gizi Kurang	7.009.474	2.798.330	2.380.000	16.530.000
Persentase Pemberian ASI Eksklusif	65.981.579	18.267.248	12.000.000	93.900.000
Persentase Kematian Akibat Demam Berdarah Dengue (DBD)	1.194.737	1.042.333	0.000000	4.500.000
Jumlah Dokter Umum per 100.000 penduduk	31.302.971	22.735.842	9.985.783	98.451.290
Jumlah Bidan per 100.000 penduduk	90.698.615	32.525.229	49.607.198	188.411.567
Jumlah Perawat per 100.000 penduduk	211.887.742	163.562.988	95.417.321	713.833.110
Jumlah Puskesmas per 100.000 penduduk	2.676.636	0.733326	1.146.295	4.543.905
Jumlah Posyandu Aktif per 100.000 penduduk	103.914.446	24.401.582	52.931.670	148.598.400
Persentase Sumber Air Minum Layak	70.677.632	20.755.763	25.970.000	100.000.000
Persentase Rumah Tangga yang Memiliki Akses Terhadap Sanitasi Layak	90.120.526	9.835.692	65.370.000	100.000.000
Persentase Desa/Kelurahan Universal Child Immunization (UCI)	85.694.737	13.976.988	38.100.000	100.000.000
Angka Keberhasilan Pengobatan TBC	88.423.684	3.460.935	80.100.000	95.100.000

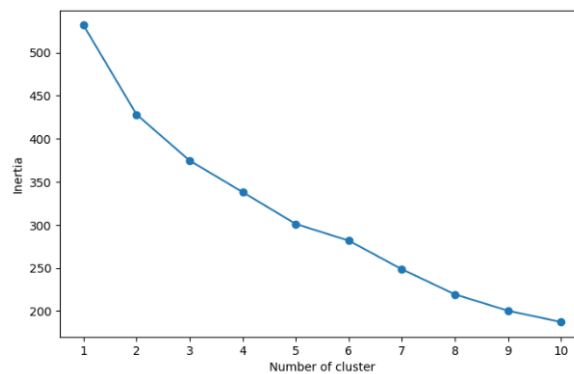
Tabel statistika deskriptif ini memberikan gambaran yang sangat penting tentang berbagai variabel terkait kesehatan dan kesejahteraan di suatu wilayah atau populasi. Dalam hal ini, nilai rata-rata (mean) seperti AKB yang mencapai 83.447.368 memberikan informasi tentang angka kematian balita secara keseluruhan. Standar deviasi yang tinggi, seperti pada AKB dengan nilai 59.140.098, mengindikasikan variasi yang signifikan dalam data angka kematian balita. Rentang nilai dari minimum, misalnya 2.000.000 untuk AKB hingga maksimum, misalnya 282.000.000 untuk AKB menyoroti variasi antara angka kematian balita terendah dan tertinggi dalam populasi tersebut. Analisis lebih lanjut dapat dilakukan untuk membandingkan karakteristik ini dengan populasi lain, memberikan wawasan mendalam mengenai kondisi kesehatan dan kesejahteraan suatu daerah.

B. Analisis Pengelompokan

Setelah dilakukan karakteristik variabel dari statistika deskriptif selanjutnya masuk ke tahap metode analisis pengelompokan menggunakan metode-metode yang telah dituliskan pada tinjauan pustaka. Sebelum masuk ke metode pengolahan analisis pengelompokan, dilakukan standarisasi pada data dengan tujuan agar skala data seimbang.

1. *K-Means*

Pada tahap awal pengelompokan menggunakan metode *K-Means* perlu diketahui untuk menentukan jumlah kelompok (*cluster*) yang akan dibuat berdasarkan wilayah Kab/Kota. Untuk mengetahui hal tersebut dilakukan dugaan terhadap jumlah kelompok dengan menggunakan metode *Elbow* dengan memperhatikan kurvanya.



Gambar 1 Elbow Curve Pada Metode K-Means Clustering

Dalam kurva *Elbow*, titik belokan yang terbentuk pada grafik umumnya dianggap sebagai petunjuk jumlah kelompok yang sesuai. Meskipun demikian, nilai optimal dari k yang diperoleh dari metode kurva elbow sering kali agak ambigu atau tidak sepenuhnya pasti dalam menentukan jumlah kelompok (k) yang terbaik. Pada grafik di atas, pilihan k diatur menjadi 2 atau 3 karena grafik yang dihasilkan menunjukkan adanya belokan. Untuk memastikan bahwa k yang dipilih memang optimal, analisis lebih lanjut akan dilakukan perhitungan nilai *Davies-Bouldin*, *Calinski-Harabasz*, *Silhouette* dan *Dunn*.

Tabel 5 Metrik Evaluasi K-Means

Jumlah Kelompok	DB	CH	Silhouette
2	1.6515	8.7385	0.2347
3	1.8933	7.5975	0.1599
4	1.5489	6.8948	0.1919
5	1.3864	6.5887	0.1999

Silhouette score dapat membantu menentukan jumlah kelompok yang optimal dalam analisis kelompok (*clustering*). Semakin tinggi *silhouette* score, semakin baik. Hal tersebut sama dengan nilai CH, tetapi untuk nilai DB semakin kecil maka semakin baik. Dalam kasus ini, score tertinggi adalah untuk jumlah kelompok 2 dengan nilai *silhouette* sebesar 0.2347. Jadi, berdasarkan nilai *silhouette* score, jumlah cluster terbaik adalah 2 untuk metode K-Means.

2. Agglomerative Hierarchical Clustering (AHC)

Pada bagian ini akan dilakukan pengelompokan indikator kesehatan masyarakat di Jawa Timur Adapun metode pengelompokan yang digunakan adalah metode *Agglomerative Hierarchical Clustering* (AHC) dengan ukuran jarak yang digunakan adalah jarak *euclidean* serta pautan yang digunakan adalah *single*, *complete*, *average linkage*, dan *ward linkage*. Pengelompokan dilakukan dengan membentuk 2, 3, dan 4 kelompok. Dipilihnya ketiga kelompok tersebut adalah karena jumlah kelompok tersebut dianggap sudah cukup untuk membandingkan kinerja dari kelompok yang akan terbentuk. Melalui masing-masing pautan dan jumlah kelompok tersebut akan dihitung nilai *Davies-Bouldin*, *Calinski-Harabasz*, *Silhouette* dan *Dunn*. Kelompok yang menghasilkan *Davies-Bouldin* terkecil, *Calinski-Harabasz* terbesar, *Silhouette* terbesar dengan *Dunn* terkecil yang menjadi kelompok optimum. Nilai *Davies-Bouldin*, *Calinski-Harabasz*, *Silhouette* dan *Dunn* untuk masing-masing pautan dan jumlah kelompok ditampilkan pada **Tabel 6**. sebagai berikut.

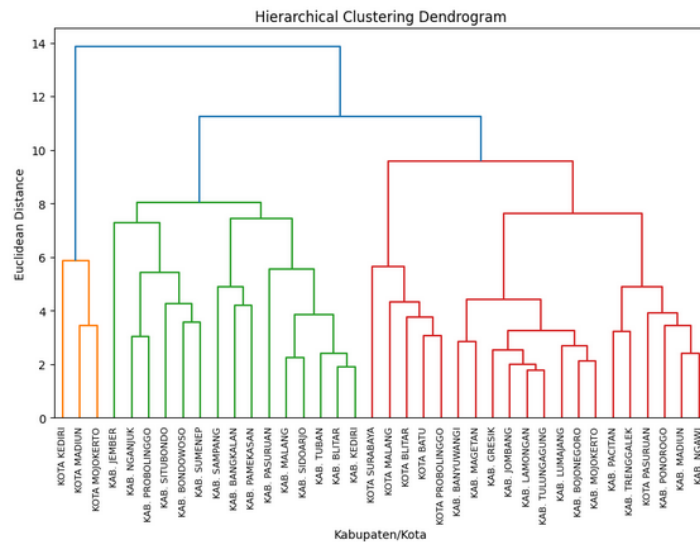
Tabel 6 Metrik Evaluasi AHC

Metode	Jumlah kelompok	DB	CH	Silhouette	Dunn
<i>Single linkage</i>	2	0.5488	2.8095	0.2842	inf
	3	0.8518	5.7629	0.2685	inf
	4	0.6459	4.4113	0.2414	inf
<i>Complete linkage</i>	2	1.0172	7.9096	0.3388	inf
	3	0.8518	5.7629	0.2685	inf
	4	0.9383	5.4201	0.2237	0.5267
<i>Average linkage</i>	2	1.0172	7.9096	0.3388	inf
	3	1.2333	6.7436	0.2251	inf
	4	1.9946	5.9296	0.0975	0.2855
<i>Ward linkage</i>	2	1.0172	7.9096	0.3388	inf
	3	1.8632	7.4562	0.1312	0.6014

4 1.6945 7.1028 0.1321 0.3826

Berdasarkan nilai metrik evaluasi, tidak ada satu metode pun yang secara konsisten unggul di semua metrik. Namun, jika kita harus memilih satu, terlihat bahwa kelompok 3 dari metode Ward Linkage memiliki nilai DB yang rendah, CH yang tinggi, serta *Dunn Index* yang relatif baik, meskipun *Silhouette Score* tidak optimal. Tetapi, pemilihan model terbaik sebaiknya juga mempertimbangkan konteks dan tujuan analisis yang ingin dicapai.

Dengan demikian pemilihan metode pada Analisis Pengelompokkan menggunakan AHC Metode *Ward Linkage* dengan jumlah kluster 3. Pemilihan ini lebih memilih metode AHC karena jumlah data yang cenderung sedikit.



Gambar 2 Dendrogram Hierarchical Clustering dengan Metode Ward Linkage

Pada Gambar 2, terlihat bahwa dendrogram dengan tiga kluster telah dihasilkan menggunakan metode euclidean distance dengan nilai sebesar 9. Dendrogram ini memberikan gambaran tentang pembentukan kelompok secara hierarkis dan aglomeratif dari 38 Kab/Kota di Jawa Timur, serta diimplementasikan melalui metode Ward yang berfokus pada minimisasi nilai varians.

C. Analisis Multivariat

1. Pengujian Signifikansi Perbedaan Kelompok yang Terbentuk Secara Multivariat

Pengujian ini dilakukan secara multivariat dengan acuan pada pengelompokan terbaik yang dihasilkan dari clustering. Digunakan One-Way MANOVA yang hasilnya ditampilkan pada Tabel 7 sebagai berikut.

Tabel 7 Hasil Uji One-Way MANOVA

Statistik Uji (Wilk's Lamda)	P-Value	Keputusan
7.0932	0.0329	Tolak H0

Sesuai dengan hasil uji wilk's lamda pada **Tabel 7** jika digunakan taraf signifikansi alpha sebesar 5%, maka keputusan Tolak H0 karena nilai p-value yang dihasilkan kurang dari 0,05. Sehingga dapat dikatakan bahwa dengan keyakinan 95% dapat disimpulkan minimal terdapat satu kelompok yang menghasilkan rata-rata yang berbeda signifikan dari ketiga kelompok kabupaten/kota yang terbentuk.

2. Pengujian Signifikansi Perbedaan Kelompok yang Terbentuk Secara Univariat

Setelah dilakukan pengujian signifikansi perbedaan kelompok secara multivariat dan diperoleh hasil yang signifikan, dilanjutkan dengan pengujian secara univariat untuk mengetahui variabel mana saja yang memberikan perbedaan yang berarti antar ketiga kelompok kabupaten/kota yang telah terbentuk sesuai dengan kelompok yang terbentuk melalui pengelompokan model terbaik. Pengujian secara univariat menggunakan One Way ANOVA. Tabel 8 menampilkan hasil dari One Way ANOVA.

Tabel 8 Hasil Uji One-Way ANOVA

Variabel	Fhitung	P-value
AKB	5.746525	6.95E-03
AKI	8.066057	1.32E-03

Persentase Balita Gizi Kurang	0.994478	3.80E-01
Persentase Pemberian ASI Eksklusif	2.557325	9.19E-02
Persentase Kematian Akibat Demam Berdarah Deng...	1.421561	2.55E-01
Jumlah Dokter Umum per 100.000 penduduk	16.287096	1.00E-05
Jumlah Bidan per 100.000 penduduk	42.693756	4.08E-10
Jumlah Perawat per 100.000 penduduk	45.983172	1.61E-10
Jumlah Puskesmas per 100.000 penduduk	7.641458	1.76E-03
Jumlah Posyandu Aktif per 100.000 penduduk	4.131363	2.45E-02
Persentase Sumber Air Minum Layak	9.211931	6.11E-04
Persentase Rumah Tangga yang Memiliki Akses Te...	11.953849	1.10E-04
Persentase Desa Kelurahan Universal Child Immu...	2.399111	1.06E-01
Angka Keberhasilan Pengobatan TBC	1.559677	2.24E-01

Informasi yang diperoleh melalui **Tabel 8** adalah bahwa dari 14 variabel, ternyata terdapat 9 variabel yang memberikan perbedaan rata-rata yang signifikan antar ketiga kelompok pada pengelompokan terbaik karena variabel pada **Tabel 8** yang bercetak tebal menghasilkan nilai Fhitung yang lebih besar dari nilai $F_{(0.05;2;36)}(3.2519)$ dan menghasilkan p-value yang lebih kecil dari 5% yang merupakan taraf signifikansi yang digunakan. Melalui hal tersebut dapat dikatakan bahwa masing-masing kelompok kabupaten/kota yang terbentuk memiliki perbedaan rata-rata antar kelompoknya pada keenam variabel yang bercetak tebal tersebut. Variabel yang memberikan pengaruh signifikan tersebut yaitu AKB, AKI, jumlah dokter umum, jumlah bidan, jumlah perawat, jumlah puskesmas, jumlah posyandu, persentase sumber air minum layak, dan jumlah rumah tangga yang memiliki akses terhadap sanitasi layak.

3. Perbedaan Karakteristik Kelompok yang Terbentuk

Melalui pengelompokan terbaik diketahui bahwa kelompok yang terbentuk adalah sebanyak 3 kelompok kabupaten/kota. ketiga kelompok kabupaten/kota tersebut akan dideskripsikan karakteristik masing-masing kelompok sehingga dapat diketahui kelompok mana yang memiliki kualitas kesehatan masyarakat yang sudah baik hingga yang masih memerlukan perhatian lebih dari pemerintah. Variabel yang digunakan adalah variabel yang telah dinyatakan sebagai variabel pembeda pada pengujian signifikansi perbedaan kelompok secara univariat pada masing-masing tipe data penelitian. Karakteristik kelompok tersebut ditampilkan melalui statistika deskriptif yang disajikan pada **Tabel 9**.

Tabel 9 Statistika Deskriptif Variabel Signifikan Berdasarkan Klaster

Variabel	Cluster Ward	0	1	2
AKB	mean	73.75	6	111.9
	std	42.59	4	67.04
	min	16	2	30
	max	174	10	282
AKI	mean	10.2	1	19.47
	std	6.07	0	11.76
	min	1	1	11
	max	25	1	58
Jumlah Dokter Umum per 100.000 penduduk	mean	34.59	77	17.79
	std	21.28	19.38	6.3
	min	17.93	60.77	9.99
	max	91.37	98.45	34.6
Jumlah Bidan per 100.000 penduduk	mean	87.06	181.7	77.36
	std	16.69	8.65	20.59
	min	64.79	171.9	49.61
	max	118.8	188.4	116.9
Jumlah Perawat per 100.000 penduduk	mean	211	659.2	123.7
	std	114.6	91.05	21.99

	min	119.3	554.1	95.42
	max	552.1	713.8	159.8
Jumlah Puskesmas per 100.000 penduduk	mean	2.84	3.66	2.26
	std	0.67	0.78	0.55
	min	1.81	3.07	1.15
	max	4.3	4.54	3.18
Jumlah Posyandu Aktif per 100.000 penduduk	mean	110.6	121.7	91.39
	std	22.89	31.04	20.6
	min	60.61	87.74	52.93
	max	139.4	148.6	119.3
Persentase Sumber Air Minum Layak	mean	77.11	95.79	57.07
	std	19.4	7.29	15.09
	min	25.97	87.37	32.49
	max	100	100	85.23
Persentase Rumah Tangga yang Memiliki Akses Terhadap Sanitasi Layak	mean	94.27	99.69	82.67
	std	4.91	0.54	10.92
	min	82.55	99.07	65.37
	max	100	100	100

Demi mempermudah dalam mengidentifikasi perbedaan karakteristik kualitas kesehatan masyarakat pada masing-masing kelompok yang terbentuk, maka dilakukan perbandingan karakteristik masing-masing kelompok kesehatan masyarakat di Jawa Timur

Tabel 10 Kategori Variabel Signifikan Tiap Klaster Berdasarkan Rerata

Variabel	Kelompok 0	Kelompok 1	Kelompok 2
AKB	Sedang	Rendah	Tinggi
AKI	Sedang	Rendah	Tinggi
Jumlah Dokter Umum per 100.000 penduduk	Sedang	Tinggi	Rendah
Jumlah Bidan per 100.000 penduduk	Sedang	Tinggi	Rendah
Jumlah Perawat per 100.000 penduduk	Sedang	Tinggi	Rendah
Jumlah Puskesmas per 100.000 penduduk	Sedang	Tinggi	Rendah
Jumlah Posyandu Aktif per 100.000 penduduk	Sedang	Tinggi	Rendah
Persentase Sumber Air Minum Layak	Sedang	Tinggi	Rendah
Persentase Rumah Tangga yang Memiliki Akses Terhadap Sanitasi Layak	Sedang	Tinggi	Rendah

4. Indikator Kesehatan Masyarakat yang Perlu Menjadi Perhatian pada Masing-Masing Kelompok Kualitas Kesehatan di Jawa Timur

Indikator kesehatan masyarakat terburuk dan terbaik yang perlu menjadi perhatian pada masing-masing kelompok kualitas kesehatan di Jawa Timur disajikan pada **Tabel 11**.

Tabel 11 Karakteristik Kelompok

Kualitas Kesehatan	Kabupaten/Kota	Indikator Kesehatan	
		Terburuk	Terbaik
Buruk (Kelompok 2)	Bangkalan, Blitar, Bondowoso, Jember, Kediri, Malang, Nganjuk, Pamekasan, Pasuruan, Probolinggo, Sampang, Sidoarjo, Situbondo, Sumenep, Tuban.	AKB, AKI, jumlah dokter umum, jumlah bidan, jumlah perawat, jumlah puskesmas, jumlah posyandu, persentase sumber air minum layak, dan jumlah rumah tangga yang memiliki akses terhadap sanitasi layak.	-
Sedang (Kelompok 0)	Banyuwangi, Bojonegoro, Gresik, Jombang, Kota Batu, Kota Blitar,	-	-

Berikut adalah Peta Jawa Timur yang dikelompokkan berdasarkan model pengelompokan terbaik berdasarkan indikator kesehatan masyarakat.



V. KESIMPULAN DAN SARAN

Setelah dilakukan analisis terhadap indikator kesehatan masyarakat di Jawa Timur, dapat ditarik beberapa kesimpulan sebagai berikut.

- Department of Statistics, Institut Teknologi Sepuluh Nopember

dengan tiga kluster menjadi pilihan utama, sebagian karena jumlah data yang terbatas.

3. Hasil pengelompokan menggunakan metode Ward Linkage menunjukkan tiga kelompok berbeda. Kelompok pertama mencakup tiga kota yang menunjukkan kondisi kesehatan masyarakat yang baik, sementara kelompok kedua terdiri dari 20 kabupaten/kota dengan kondisi kesehatan masyarakat yang sedang. Di sisi lain, kelompok ketiga terdiri dari 15 kabupaten/kota yang menunjukkan kondisi kesehatan masyarakat yang buruk. Selain itu, dari analisis ini teridentifikasi sembilan variabel yang memiliki pengaruh signifikan terhadap kondisi kesehatan masyarakat di Jawa Timur, termasuk di antaranya Angka Kematian Bayi (AKB), Angka Kematian Ibu (AKI), jumlah dokter umum, jumlah bidan, jumlah perawat, jumlah puskesmas, jumlah posyandu, persentase sumber air minum layak, dan jumlah rumah tangga dengan akses sanitasi yang layak.

B. Saran

Adapun saran dari penelitian ini sebagai berikut.

1. Untuk penelitian selanjutnya bisa mempertimbangkan metode yang digunakan dan lebih memperhatikan skor evaluasi dari metode yang dihasilkan.
2. Pada variabel yang digunakan dapat menggunakan variabel lain yang lebih signifikan yang dapat berupa variabel kategorik maupun numerik sehingga untuk penelitian selanjutnya dapat menggunakan metode yang berbeda juga.
3. Untuk pemerintah lebih memfokuskan wilayah kab/kota di Jawa Timur yang terindikasi indikator kesehatan tergolong buruk dengan memperhatikan faktor-faktor yang signifikan seperti pada hasil penelitian.

REFERENCES

- [1] R. E. Walpole, Pengantar Statistika Edisi ke-3, Jakarta: PT. Gramedia, 1993.
- [2] B. P. Santosa and Ashari, Analisis Statistik dengan Microsoft Excel & SPSS, Yogyakarta: Andi Offset, 2005.
- [3] E. Prasetyo, Data Mining Konsep dan Aplikasi Menggunakan Matlab, Yogyakarta: Andi Offset, 2012.
- [4] R. Johnson and G. Bhattacharyya, Statistics—Principles and Methods. 6th Edition, New York: Wiley, 2010.
- [5] R. Johnson and D. Wichern, Applied Multivariate Statistical Analysis. 6th Edition, Upper Saddle River: Pearson Prentice Hall, 2007.
- [6] J. F. Hair, W. C. Black and E. Andreson, Multivariate Data Analysis (seventh ed.), New Jersey: Prentice Hall Inc, 2009.
- [7] X. Gong and M. Richman, On the Application of Cluster Analysis to Growing Season Precipitation Data in North America East of the Rockies, North America: J. Climate, 1995.
- [8] I. Ghozali, Aplikasi Analisis Multivariate Dengan Program IBM SPSS 23 (Edisi 8). Cetakan ke VIII, Semarang: Universitas Diponegoro, 2016.
- [9] Pemerintah RI, Undang-Undang Republik Indonesia Nomor 17 Tahun 2007 Tentang Rencana Pembangunan Jangka Panjang Nasional Tahun 2005 - 2025, Jakarta: Badan Perencanaan Pembangunan Nasional, 2007.
- [10] UNDP, "Sustainable Development Goals," 2021. [Online]. Available: <https://www.undp.org/sustainable-development-goals>.