

Analisis Kluster pada Faktor-Faktor yang Mempengaruhi Indikator Kesejahteraan Sosial dan Ekonomi di Provinsi Jawa Timur Tahun 2023

Oktaviana Nur Rohmatulillah^{1*}, Karisma Bunga Nirmala², Sri Pingit Wulandari³

¹⁻³Institut Teknologi Sepuluh Nopember, Indonesia

Alamat: Jalan Raya ITS, Keputih, Sukolilo, Surabaya, Jawa Timur 60117, Indonesia

*Korespondensi penulis: oktaviana12002@gmail.com

Abstract. *Social and economic welfare reflects the quality of life in a region and is influenced by local social, economic, and environmental factors. East Java, as the second most populous province in Indonesia, faces challenges in improving the welfare of its residents, particularly due to varying regional characteristics such as employment, education, and population demographics. To understand the patterns of interrelationships among factors affecting welfare, this study conducted a cluster analysis to group regions based on similar characteristics. The cluster analysis employed both hierarchical (complete linkage) and non-hierarchical (K-means) approaches to determine the optimal number of clusters. The results revealed that the level of diversity across regions in East Java tends to be homogeneous in social and economic aspects, with average values exceeding standard deviations. Assumption tests for the cluster analysis confirmed that the data met the assumptions of multivariate normal distribution and dependency. Through hierarchical (complete linkage) and non-hierarchical (K-means) cluster analysis, two main clusters were formed, dividing districts/cities in East Java based on welfare characteristics. Using the complete linkage method, 27 regions were grouped into cluster 1, and 11 regions into cluster 2, while K-means grouped 26 regions into cluster 1 and 12 regions into cluster 2. Out of the six variables used, one variable was found to be insignificant in influencing the clustering results. Based on the mapping results, the grouping aligns with similar criteria, where urban areas predominantly fall into one cluster, and the other cluster is dominated by rural areas.*

Keywords: *Kluster Analysis, Complete Linkage, K-means, Social and Economic Welfare*

Abstrak. Kesejahteraan sosial dan ekonomi mencerminkan kualitas hidup masyarakat di suatu wilayah dan dipengaruhi oleh faktor sosial, ekonomi, dan lingkungan setempat. Jawa Timur sebagai provinsi dengan penduduk terbesar kedua di Indonesia, menghadapi tantangan dalam meningkatkan kesejahteraan warganya, terutama karena perbedaan karakteristik wilayah seperti ketenagakerjaan, pendidikan, dan kependudukan. Untuk memahami pola keterkaitan antar faktor yang memengaruhi kesejahteraan, penelitian ini melakukan analisis kluster yang mengelompokkan wilayah berdasarkan karakteristik serupa. Analisis kluster ini menggunakan pendekatan hirarki (complete linkage) dan non-hirarki (K-means) untuk menentukan jumlah kluster yang paling optimal. Hasil penelitian menunjukkan bahwa tingkat keragaman antar wilayah di Jawa Timur cenderung homogen dalam aspek sosial dan ekonomi, dengan nilai rata-rata yang lebih besar dari standar deviasi. Uji asumsi untuk analisis kluster menunjukkan bahwa data memenuhi asumsi distribusi normal multivariat dan dependensi. Melalui analisis kluster hirarki (complete linkage) serta non-hirarki (K-means), terbentuk dua kluster utama yang membagi kabupaten/kota di Jawa Timur berdasarkan karakteristik kesejahteraan. Dengan metode complete linkage, 27 wilayah masuk ke dalam kluster 1 dan 11 wilayah ke kluster 2, sedangkan K-means mengelompokkan 26 wilayah dalam kluster 1 dan 12 wilayah dalam kluster 2. Dari enam variabel yang digunakan, satu variabel terbukti tidak signifikan dalam memengaruhi hasil kluster. Berdasarkan hasil pemetaan sudah sesuai kriteria yang cenderung sama dimana wilayah kota mayoritas tergabung dalam satu kluster, dan kluster lainnya di dominasi oleh wilayah kabupaten.

Kata kunci: *Analisis Kluster, Complete Linkage, K-means, Kesejahteraan Sosial dan Ekonomi.*

1. LATAR BELAKANG

Received: December, 2024; Revised: December, 2024; Accepted: December, 2024; Online Available: December 11, 2024; Published: December 13, 2024;

* Oktaviana Nur Rohmatulillah, oktaviana12002@gmail.com

Kesejahteraan sosial dan ekonomi merupakan aspek penting yang mencerminkan kualitas hidup masyarakat di suatu wilayah. Kesejahteraan sendiri merupakan titik ukur bagi suatu masyarakat bahwa telah berada pada kondisi sejahtera atau disebut juga persamaan hidup yang setingkat lebih tinggi dari kehidupan sebelumnya (Farisi & dkk, 2022) Kesejahteraan ini dipengaruhi oleh berbagai faktor yang berkaitan dengan kondisi sosial, ekonomi dan lingkungan setempat. Berdasarkan data Badan Pusat Statistik Nasional (BPS), Jawa Timur merupakan provinsi dengan jumlah penduduk terbesar kedua di Indonesia, dimana jumlah penduduk besar tidak akan jauh dari tingginya angka kemiskinan, sehingga dapat menjadi hambatan dalam upaya peningkatan kesejahteraan masyarakatnya. Perbedaan karakteristik wilayah seperti ketenagakerjaan, pendidikan serta kependudukan menciptakan variasi yang signifikan dalam tingkat kesejahteraan antar daerah. Beragamnya faktor-faktor ini menunjukkan perlunya pendekatan yang sistematis untuk memahami pola keterkaitan antar indikator kesejahteraan sosial dan ekonomi guna membantu pemerintah daerah dalam merancang kebijakan yang lebih tepat sasaran. Oleh karena itu, pada penelitian ini dilakukan analisis untuk mengelompokkan faktor-faktor yang memengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023 berdasarkan karakteristik yang serupa menggunakan analisis kluster.

Analisis kluster sendiri merupakan kumpulan dari beberapa teknik pengolahan data multivariat yang memiliki tujuan utama mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya, dimana hasil dari kluster yang terbentuk harus menunjukkan homogenitas internal yang tinggi dalam satu kluster dan heterogenitas yang tinggi antar kluster (Badrutaman, Sudarno, & Maruddani, 2020). Analisis kluster terbagi menjadi dua jenis yaitu metode hirarki (*agglomerative* dan *divisive*) dan non hirarki. Pada penelitian ini dilakukan analisis kluster hirarki dan analisis kluster non hirarki untuk mengetahui metode analisis kluster dan jumlah kluster yang paling optimum.

Pada penelitian ini untuk melakukan analisis kluster dilakukan analisis karakteristik, pengujian asumsi analisis kluster yaitu pengujian distribusi normal multivariat dan pengujian korelasi independensi, analisis kluster beserta pemetannya, analisis karakteristik kluster pada faktor-faktor yang memengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023. Adapun tujuan penelitian ini yaitu mengetahui pengelompokan daerah-daerah di Provinsi Jawa Timur yang memiliki karakteristik serupa pada kesejahteraan sosial dan ekonomi. Penelitian diharapkan dapat memberi manfaat kepada penulis untuk lebih memahami terkait analisis kluster dan penerapan analisis kluster dalam kehidupan sehari-hari, bagi pembaca dapat mengetahui penerapan analisis kluster dan pengelompokan daerah-daerah di Provinsi Jawa Timur berdasarkan kesejahteraan sosial dan ekonomi, serta bagi pemerintah dapat dijadikan informasi mengenai penyebaran kondisi kesejahteraan sosial dan ekonomi untuk mewujudkan kesejahteraan bersama. Batasan masalah yang digunakan pada penelitian ini yaitu analisis kluster pada metode hirarki hanya menggunakan *average linkage* dan

pada metode non hirarki hanya menggunakan *K-means*.

2. KAJIAN TEORITIS

2.1 Presentase Penduduk Miskin

Presentase penduduk miskin merupakan banyaknya penduduk yang berada di bawah garis kemiskinan nasional dibagi dengan jumlah penduduk pada periode waktu yang sama dinyatakan dalam satuan persen (%). Kemiskinan dipandang sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Metode Perhitungan melalui Jumlah penduduk yang berada di bawah garis kemiskinan pada waktu tertentu dibagi dengan jumlah penduduk seluruhnya pada periode waktu yang sama dinyatakan dalam satuan persen (%) (Badan Pusat Statistik, 2024).

2.2 Rata-rata Pengeluaran Makanan Perkapita Sebulan

Pengeluaran per kapita adalah biaya yang dikeluarkan untuk konsumsi semua anggota rumah tangga selama sebulan dibagi dengan banyaknya anggota rumah tangga. Pengeluaran rata-rata per kapita merupakan biaya yang dikeluarkan untuk konsumsi semua anggota rumah tangga selama sebulan baik yang berasal dari pembelian, pemberian maupun produksi sendiri. Rata-rata Pengeluaran Makanan Perkapita Sebulan mengetahui pola konsumsi rumah tangga secara umum dan menilai tingkat kesejahteraan ekonomi penduduk (Pemerintah Kabupaten Pematang, 2017).

2.3 Garis Kemiskinan

Garis kemiskinan merupakan tingkat minimum pendapatan yang dianggap perlu dipenuhi untuk memperoleh standar hidup yang mencukupi. Garis Kemiskinan (GK) mencerminkan cerminan dari nilai rupiah pengeluaran minimum yang diperlukan seseorang untuk memenuhi kebutuhan pokok hidupnya selama sebulan, baik kebutuhan makanan maupun non-makanan. GK terdiri dari Garis Kemiskinan Makanan (GKM) dan Garis Kemiskinan Non-Makanan (GKNM) (Badan Pusat Statistik (BPS) Jawa Timur, 2024).

2.4 Tingkat Pengangguran Terbuka (TPT)

TPT (Tingkat Pengangguran Terbuka) adalah persentase jumlah pengangguran terhadap jumlah angkatan kerja. Indikator tersebut juga bermanfaat untuk menghitung jumlah angkatan kerja serta mengetahui tingkat pengangguran yang terjadi pada suatu wilayah. TPT yang memiliki nilai tinggi menunjukkan banyak angkatan kerja yang tidak terserap pada pasar. TPT dapat mencerminkan kinerja pasar kerja dan berlangsungnya kondisi ekonomi tertentu (Indra Kertati, 2023).

2.5 Tingkat Partisipasi Angkatan Kerja (TPAK)

Tingkat Partisipasi Angkatan Kerja (TPAK) merupakan perbandingan jumlah angkatan kerja terhadap jumlah penduduk usia kerja. TPAK mngindikasikan besarnya penduduk usia kerja yang aktif secara ekonomi di suatu wilayah. emakin tinggi TPAK kota, menunjukkan bahwa semakin tinggi pula pasokan tenaga kerja (labour supply) yang tersedia untuk memproduksi barang dan jasa suatu perekonomian di kota tersebut (Badan Pusat Statistik, 2024).

2.6 Rata-Rata Pendapatan Bersih Sebulan Pekerja Informal

Rata-Rata Pendapatan Bersih Sebulan Pekerja Informal merupakan rata-rata pendapatan dari Pekerja yang berstatus berusaha sendiri dan pekerja bebas di sektor pertanian dan non pertanian/ Employment status are self employed, casual agricultural worker, and casual non-agricultural worker. Pekerja informal berisiko terkena dampak dari ketidakstabilan ekonomi, contohnya pada masa pandemi COVID-19, yang berpotensi menyebabkan penurunan penawaran dan pendapatan (Salsabila, Fitriani, Umaidah, & Heryana, 2023).

2.7 Statistika Deskriptif

Statistika deskriptif merupakan metode analisis berupa mengumpulkan, mengolah, dan menyajikan satu data penelitian dalam bentuk tabel, grafik, maupun histogram. Salah satu bentuk analisis deskriptif yang digunakan adalah *mean*, varians, median. Statistika deskriptif berfungsi untuk menyampaikan informasi yang dapat dipahami dengan baik (Luh Titi Handayani, 2023).

2.7.1 Mean

Mean adalah rata-rata dari sekumpulan data tertentu. *Mean* merupakan salah satu ukuran pemusatan data dan ukuran paling penting dari kecenderungan sentral data terdistribusi. *Mean* juga salah satu metode paling dasar untuk analisis data, berikut adalah rumus dari *mean* (uryana, 2024).

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (1)$$

Keterangan :

- \bar{x} : Rata-rata sampel
- x_i : Data ke-i pada sampel
- n : Banyaknya data sampel

2.7.2 Standar deviasi

Standar deviasi merupakan ukuran pemusatan data yang digunakan untuk mengetahui homogenitas dari data serta dapat menjadi proksi dari sebuah ukuran risiko atau stabilitas data. Standar deviasi merupakan akar dari varians, Berikut adalah rumus dari standar deviasi (I Wayan Terimajaya, 2024).

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (2)$$

Keterangan:

- s^2 : Varians

- s : Standar deviasi
 \bar{x} : Rata-rata sampel
 x_i : Data ke-i pada sampel
 n : Banyaknya data sampel

2.7.3 Nilai Maksimum dan Minimum

Nilai maksimal adalah nilai terbesar dari kumpulan suatu data. Sedangkan nilai minimum adalah nilai terkecil dari suatu data. Nilai maksimum dan minimum juga dapat digunakan untuk menghitung *range*, yaitu dengan cara nilai maksimum dikurangi nilai minimum (Walpole, 2017).

2.8 Normalitas Multivariat Uji T-Proporsi

Uji T merupakan pengujian yang fleksibel. Pengujian ini dapat digunakan dalam kondisi jumlah data sampel sedikit maupun banyak. Uji t yang didekati dengan distribusi t untuk jumlah sampel besar nilainya akan cenderung hampir sama dengan distribusi normal. Berikut adalah langkah-langkah uji t sebagai berikut (Septin Puji Astuti, 2023).

- .Merumuskan Hipotesis:
 - H_0 : Data berdistribusi normal multivariat
 - H_1 : Data tidak berdistribusi normal multivariat
 - Menentukan taraf signifikan
 - Merumuskan daerah penolakan, Tolak H_0 jika nilai t proporsi $\leq 45\%$ atau t proporsi $> 55\%$ dan P value $< \alpha$
 - Menghitung statistik uji
- Statistik Uji :

$$d_j^2 = (x_j - \bar{x})S^{-1}(x_j - \bar{x}) \quad (3)$$

Keterangan :

- j : Banyak pengamatan, $j=1,2,\dots,n$
 x_j : Objek pengamatan ke- j
 \bar{x} : Rata-rata pengamatan ke- j
 d_j^2 : Nilai kuadrat ke- j
 S^{-1} : Invers matriks varian kovarian

2.9 Uji Korelasi Independensi (Uji Bartlett)

Uji korelasi independensi adalah salah satu teknik statistik yang digunakan untuk menguji apakah terdapat hubungan antara dua variabel dalam sebuah populasi. Uji *bartlett* test merupakan uji yang mengikuti distribusi chi square dengan derajat bebas 1 (Saputra & Natalini, 2021). Asumsi yang harus dipenuhi untuk uji MANOVA diantaranya kehomogenan ragam. Berikut adalah langkah-langkah uji *bartlett*.

Merumuskan Hipotesis:

- H_0 : $\rho = I$ (Matriks korelasi identik dengan matriks identitas atau antar variabel independen)

$H_1 : \rho \neq I$ (Matriks korelasi tidak identik dengan matriks identitas atau antar variabel independen)

Menggunakan taraf signifikan sebesar α dan daerah penolakan Tolak H_0 jika nilai $\chi^2_{hitung} > \chi^2_{(\alpha,df)}$, maka statistik uji

$$\chi^2 = (ln.n)\{B - \sum dklogs_i^2\} \quad (4)$$

Keterangan :

n = Jumlah data

B = $\sum dklogs_i^2$; yang mana $s^2 = \frac{\sum dks_i^2}{\sum dk}$

s_i^2 = Varians data untuk setiap kelompok ke i

dk = Derajat bebas

2.10 Standarisasi Data

Standarisasi data dilakukan ketika saat mengelompokan objek yang memiliki skala pengukuran yang berbeda dan cukup besar, karena analisis *klaster* mengelompokan objek yang memiliki kemiripan sama dalam *klaster* dan kemiripan diukur dari objek dengan jarak dekat (Hasmira, Alwi, & Nurfadhilah, 2023). Standarisasi data dapat berfungsi memperkecil perbedaan tiap *klaster* (Ismail, et al., 2024). Berikut adalah rumus dari standarisasi.

$$Z_{score} = \frac{X - \bar{X}}{s} \quad (5)$$

Keterangan:

X : Data amatan

μ : Nilai rata-rata

σ : Nilai standar deviasi

2.11 Analisis Cluster

Analisis *klaster* merupakan teknik analisis multivariat yang memiliki tujuan utama untuk melakukan proses pengelompokan objek-objek pengamatan yang dibagi menjadi beberapa kelompok berdasarkan karakteristik yang dimilikinya. Prinsip dasar dalam analisis *klaster* adalah melakukan proses pengelompokan objek (observasi) pada suatu *klaster* yang memiliki kemiripan yang besar dengan objek yang lainnya dalam *klaster* yang sama, tetapi sangat tidak mirip dengan anggota objek yang lain pada *klaster* yang berbeda (Hasmira, Alwi, & Nurfadhilah, 2023). Analisis *klaster* terdiri dari *klaster* hierarki dan non-hierarki.

1). Metode Hierarki

Analisis *klaster* menggunakan metode hierarki merupakan pengelompokan objek secara hierarki berdasarkan kemiripan sifatnya dimana belum diketahui jumlah *klaster* yang terbentuk. Hasil dari analisis ini ditunjukkan melalui diagram pohon atau disebut dengan dendogram (Apriliana & Widodo, 2023). Metode hierarki di analisis melalui

penggabungan berurutan (*agglomerative*) atau pembagian berurutan (*divisive*). Berikut adalah beberapa bagian dari metode hierarki yang akan disajikan sebagai berikut.

- **Complete Linkage**

Analisis hierarki menggunakan *complete linkage* merupakan kebalikan dari metode *single linkage*. Prinsip jarak pada *complete linkage* menggunakan jarak terjauh antar objek. Pengelompokan dilakukan dengan mencari nilai minimum $D = \{d_{(XY)}\}$ dan menggabungkan objek-objek saling berdekatan, berikut adalah rumus yang digunakan pada *complete linkage* (Asiska, Satyahadewi, & Perdana, 2019).

$$d_{(UV)W} = \max\{d_{UW}, d_{VW}\} \quad (7)$$

Keterangan:

$d_{(UV)W}$: Jarak antara kluster (UV) dan W

d_{UW} : Jarak antara kluster U dan W

d_{VW} : Jarak antara kluster V dan W

2). **Metode Non-Hierarki**

Metode non hierarki merupakan pengelompokan objek dimana banyaknya *klaster* yang akan dibentuk dapat ditentukan terlebih dahulu. Metode ini dapat diterapkan pada data yang lebih besar dibandingkan metode hierarki (Suhaeni, Kurnia, & Ristiyan, 2018). Metode non-hierarki yang umum digunakan adalah *K-means*.

- **K Means**

Metode *K-means* merupakan metode untuk analisis *klustering* yang bertujuan untuk menentukan cara terbaik untuk membagi entitas ke-n menjadi kelompok yang disebut *klaster*. Langkah-langkah melakukan *k-means* sebagai berikut.

1. Menentukan banyaknya *klaster*.
2. Menentukan pusat *klaster*, dan dilanjutkan dengan menghitung jarak setiap objek setiap pusat *klaster*. Jarak antara tiap titik data yang ada dan tiap centroid akan dihitung menggunakan analisis Euclidean hingga ditentukan jarak terpendek antara setiap titik data dan centroid Euclidean ditentukan jarak terpendek antara setiap titik data dan centroid.
3. Pengelompokan objek berdasarkan jarak minimum terhadap pusat *klaster*. (Saputra & Natalini, 2021).

2.12 Pemilihan Jumlah Kluster Optimum

Pemilihan jumlah kluster optimum merupakan salah satu prosedur dalam analisis kluster pada data. Jumlah kluster optimum dapat diketahui dengan kriteria nilai pseudo F-Statistics dan *icredate*. Dilihat dari nilai pseudo F-Statistics dan *icredate*, jumlah kluster dipilih apabila memiliki nilai pseudo F-Statistics lebih besar dan nilai *icredate* (Khoirunnisa, Ruchjana, Irianingsih, & Suhadi, 2020). Berikut merupakan persamaan pemilihan jumlah kluster.

- Pseudo F- *statistics*:

$$Pseudo - F = \frac{\left(\frac{R^2}{k-1}\right)}{\left(\frac{1-R^2}{n-k}\right)} \quad (9)$$

$$R^2 = \frac{(SST-SSW)}{SST} \quad (10)$$

$$SST = \sum_{i=1}^n \sum_{j=1}^c \sum_{k=1}^p (X_{ijk} - \bar{X}_j)^2 \quad (11)$$

$$SSW = \sum_{i=1}^n \sum_{j=1}^c \sum_{k=1}^p (X_{ijk} - \bar{X}_{jk})^2 \quad (12)$$

Keterangan:

SST (Sum Square Total) : Total jumlah dari kuadrat jarak sampel terhadap rata-rata keseluruhan

SSW (Sum Square Within) : Total jumlah dari kuadrat jarak sampel terhadap rata-rata kelompoknya

n : Banyaknya sampel

c : Banyaknya variabel

p : Banyaknya kelompok

X_{ijk} : Sampel ke-i pada variabel ke-j kelompok ke-k

\bar{X}_j : Rata-rata seluruh sampel pada variabel ke-j

\bar{X}_{jk} : Rata-rata sampel pada variabel ke-j dan kelompok ke-k

- *Increate*

Icdrate (internal klaster dispersion) menggambarkan tingkat dispersi atau perbedaan dalam klaster. Semakin kecil nilai *icdrate* maka semakin baik hasil pengelompokannya.

$$Icdrate = 1 - \frac{SSB}{SST} = 1 - \frac{SST-SSW}{SST} = 1 - R^2 \quad (13)$$

Keterangan :

SSB : *Sum Square Between* (SST-SSW)

R² : *Recovery Rate* (SSB/SST)

3. METODE PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari *website* Badan Pusat Statistik Jawa Timur. Data yang digunakan adalah data terkait faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur pada tahun 2023.

3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini adalah sebagai berikut.

Tabel 1. Variabel Penelitian

Variabel	Keterangan	Kategori	Satuan
X ₁	Persentase Penduduk Miskin	Rasio	Persen
X ₂	Rata-Rata Pengeluaran Makanan Perkapita Sebulan	Rasio	Rupiah
X ₃	Garis Kemiskinan	Rasio	Rupiah
X ₄	Tingkat Pengangguran Terbuka (TPT)	Rasio	-
X ₅	Tingkat Partisipasi Angkatan Kerja (TPAK)	Rasio	-
X ₆	Rata-Rata Pendapatan Bersih Sebulan Pekerja Informal	Rasio	Rupiah

3.3 Struktur Data

Struktur data yang digunakan pada penelitian ini ditunjukkan pada tabel berikut.

Tabel 2. Struktur Data

Observasi ke-	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆
1	X _{1 1}	X _{2 1}	X _{3 1}	X _{4 1}	X _{5 1}	X _{6 1}
2	X _{1 2}	X _{2 2}	X _{3 2}	X _{4 2}	X _{5 2}	X _{6 2}
3	X _{1 3}	X _{2 3}	X _{3 3}	X _{4 3}	X _{5 3}	X _{6 3}
4	X _{1 4}	X _{2 4}	X _{3 4}	X _{4 4}	X _{5 4}	X _{6 4}
5	X _{1 5}	X _{2 5}	X _{3 5}	X _{4 5}	X _{5 5}	X _{6 5}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
38	X _{1 38}	X _{2 38}	X _{3 38}	X _{4 38}	X _{5 38}	X _{6 38}

3.4 Langkah Analisis

Langkah analisis yang dilakukan dalam penelitian ini adalah sebagai berikut.

1. Mengumpulkan data faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Jawa Timur tahun 2023.
2. Mendeskripsikan karakteristik data faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Jawa Timur tahun 2023.
3. Melakukan pengujian asumsi analisis kluster yang terdiri dari uji asumsi distribusi normal multivariat dan uji independensi antar variabel pada data faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Jawa Timur tahun 2023.
4. Melakukan standarisasi pada data faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Jawa Timur tahun 2023.
5. Melakukan analisis kluster hierarki menggunakan metode *complete linkage*.
 - a. Menentukan jumlah kluster optimum berdasarkan nilai *Pseudo-F* dan *Icdrate*.
 - b. Menghitung jarak euclidean yang divisualisasikan melalui dendrogram.
 - c. Melakukan pengelompokan pada hasil kluster optimum yang terbentuk.

- d. Melakukan analisis karakteristik pada masing-masing klaster yang terbentuk.
- e. Melakukan pemetaan pada hasil klaster optimum yang terbentuk.
6. Melakukan analisis klaster non-hierarki menggunakan metode *K-Means*.
 - a. Menentukan jumlah klaster optimum berdasarkan nilai *Pseudo-F* dan *Icdrate*.
 - c. Melakukan analisis perbedaan variabel pada klaster yang terbentuk menggunakan ANOVA.
 - d. Melakukan pengelompokan pada hasil klaster optimum yang terbentuk.
 - e. Melakukan analisis karakteristik pada masing-masing klaster yang terbentuk.
 - f. Melakukan pemetaan pada hasil klaster optimum yang terbentuk.
7. Menginterpretasi hasil analisis.
8. Menarik kesimpulan dan saran.

4. HASIL DAN PEMBAHASAN

4.1 Karakteristik Data

Karakteristik data dari faktor-faktor yang memengaruhi kesejahteraan sosial dan ekonomi Jawa Timur tahun 2023 dianalisis menggunakan rata-rata, standar deviasi, nilai minimum, dan nilai maksimum yang disajikan sebagai berikut.

Tabel 3. Karakteristik Data

Var	Mean	St. Dev	Min	Prov	Max	Prov
X ₁	10,293	4,321	3,310	Kota Batu	21,760	Sampang
X ₂	643855	135429	456234	Sampang	1052321	Sumenep
X ₃	514364	91406	370643	Pacitan	742678	Surabaya
X ₄	4,663	1,429	1,710	Sumenep	8,050	Sidoarjo
X ₅	73,159	3,767	66,890	Nganjuk	81,640	Pacitan
X ₆	1639442	354774	1111614	Bondowoso	2860798	Surabaya

Berdasarkan tabel 3 diketahui bahwa pada keseluruhan variabel memiliki nilai rata-rata yang lebih besar dari pada standar deviasinya, dimana hal tersebut menunjukkan keragaman data antar provinsi cenderung kecil. Pada presentase penduduk miskin diketahui Sampang dan terendah pada Kota Batu, Rata-Rata Pengeluaran Makanan Perkapita Sebulan tertinggi ada di Sumenep dan terendah di Sampang, wilayah dengan garis kemiskinan tertinggi ada di Kota Surabaya dan terendah di Pacitan, Tingkat Pengangguran Terbuka (TPT) tertinggi di Sidoarjo dan terendah di Sumenep, Tingkat Partisipasi Angkatan Kerja (TPAK) tertinggi di Pacitan dan terendah di Nganjuk, sedangkan Rata-Rata Pendapatan Bersih Sebulan Pekerja Informal tertinggi di Kota Surabaya dan terendah di Bonowoso.

4.2 Pengujian Asumsi

Dalam analisis klaster terhadap faktor-faktor yang memengaruhi kesejahteraan sosial dan ekonomi Jawa Timur tahun 2023, terdapat asumsi yang harus dipenuhi, yaitu bahwa

data memiliki distribusi normal multivariat dan terdapat keterkaitan antar variabel dependen. Penjelasan mengenai asumsi-asumsi ini akan dijelaskan sebagai berikut.

4.2.1 Uji Asumsi Distribusi Normal Multivariat

Pengujian asumsi distribusi normal multivariat pada data faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Jawa Timur tahun 2023 dilakukan dengan menggunakan T-Proporsi dengan hipotesis sebagai berikut.

H_0 : Data faktor kesejahteraan sosial dan ekonomi Jawa Timur tahun 2023 berdistribusi normal multivariat.

H_1 : Data faktor kesejahteraan sosial dan ekonomi Jawa Timur tahun 2023 tidak berdistribusi normal multivariat.

Ditetapkan taraf signifikan (α) sebesar 0,05 dan diperoleh daerah penolakan, yaitu tolak H_0 jika T-proporsi lebih kecil dari 45% atau T-proporsi lebih besar dari 55%. Hasil statistik uji ditunjukkan sebagai berikut.

Tabel 3. Uji Distirbusi Normal Multivariat

T-Proporsi
52,63%

Berdasarkan tabel 3 diketahui bahwa nilai T-Proporsi sebesar 52,63% yang bernilai lebih besar dari 45% dan lebih kecil dari 55%, sehingga diputuskan gagal tolak H_0 yang artinya data faktor kesejahteraan sosial dan ekonomi Jawa Timur tahun 2023 telah berdistribusi normal multivariat.

4.2.2 Uji Asumsi Dependensi

Pengujian asumsi dependensi pada data faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Jawa Timur tahun 2023 dilakukan dengan menggunakan uji *bartlett* dengan hipotesis sebagai berikut.

H_0 : $\rho = \mathbf{I}$ (Matriks korelasi identik dengan matriks identitas atau antar variabel independen)

H_1 : $\rho \neq \mathbf{I}$ (Matriks korelasi tidak identik dengan matriks identitas atau antar variabel dependen)

Dengan taraf signifikan (α) sebesar 0,05 diperoleh daerah penolakan, yaitu tolak H_0 jika χ^2_{hitung} lebih besar dari $\chi^2_{(\alpha;df)}$ atau P-value lebih kecil dari α . Statistik uji ditunjukkan pada tabel berikut.

Tabel 4. Uji Bartlett

χ^2_{hitung}	$\chi^2_{(0,05;15)}$	P-value
80,796	24,996	0,000

Berdasarkan tabel 4 dapat diketahui bahwa χ^2_{hitung} sebesar 80,796 dimana bernilai lebih besar dari $\chi^2_{(0,05;15)}$ sebesar 24,996 dan diperkuat dengan P-value sebesar 0,000 yang bernilai lebih kecil dari α sebesar 0,05, diputuskan tolak H_0 yang artinya matriks korelasi tidak identik dengan matriks identitas atau antar variabel dependen, sehingga memenuhi asumsi dependensi.

4.3 Penentuan Jumlah Kluster Optimum

Dalam analisis kluster, baik hirarki maupun non-hirarki, jumlah kluster optimum yang menggambarkan faktor-faktor kesejahteraan sosial dan ekonomi di Jawa Timur tahun 2023 perlu ditentukan. Hasil penentuan jumlah kluster yang optimal adalah sebagai berikut.

Tabel 5. Penentuan Jumlah Kluster Optimum

Metode	Jumlah Kluster	Pseudo-F	Icdrate
<i>Average Linkage</i>	2	1,2263	0,9671
	3	14,9899	0,5386
	4	10,2799	0,5244
	5	12,6751	0,3943
<i>K-means</i>	2	17,2588	0,6759
	3	8,3897	0,6759
	4	9,5465	0,5428
	5	3,9551	0,6759

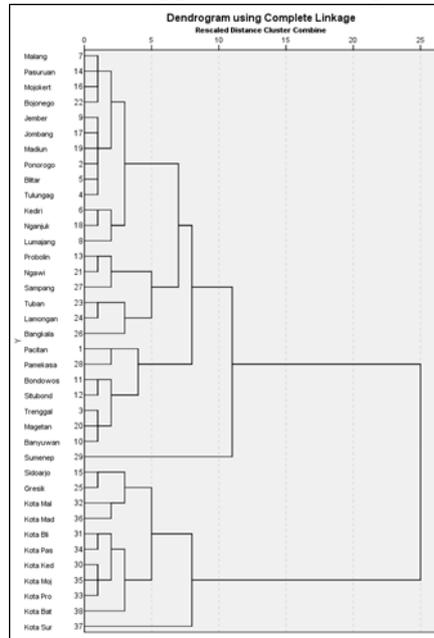
Berdasarkan Tabel 5 dapat diketahui bahwa pada analisis kluster hirarki dengan metode *average linkage* dengan 3 kluster menghasilkan nilai *Pseudo-F* tertinggi yaitu 14,9899. Sedangkan untuk kluster non-hirarki dengan metode *K-means* dengan 2 kluster menghasilkan nilai tertinggi yaitu 17,2588. Namun, untuk memudahkan analisis perbandingan dan interpretasi maka dalam penelitian ini akan digunakan analisis kluster *complete linkage* serta *K-means*, dimana kedua metode memiliki kluster yang terbentuk sebanyak 2 kluster dengan nilai *Pseudo-F* tinggi dan *Icdrate* kecil.

4.4 Pengelompokan Kluster

Pengelompokan kluster pada faktor-faktor yang memengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023 dilakukan menggunakan metode terbaik yaitu metode *complete linkage* dan metode *K-means* karena memiliki nilai *Pseudo-F* tertinggi dan *icdrate* paling rendah dengan jumlah kluster optimum sebanyak 2 kluster.

4.4.1 Pengelompokan Kluster Hirarki dengan Metode *Complete Linkage*

Hasil pemilihan kluster optimum pada metode *complete linkage* pada faktor-faktor yang memengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023 didapatkan 2 kluster. Adapun daerah-daerah yang terbentuk ditentukan berdasarkan nilai *euclidean* yang digambarkan menggunakan dendrogram sebagai berikut.



Gambar 1 Dendrogram Analisis Kluster *Complete Linkage*

Berdasarkan Gambar 1 diketahui bahwa dengan metode *complete linkage* yang terbentuk sebanyak 2 kluster. Hasil analisis kluster berdasarkan metode *complete linkage* faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023 disajikan pada tabel berikut.

Tabel 6. Hasil Analisis Kluster *Complete Linkage*

Kluster	Provinsi
1	Pacitan, Ponorogo, Trenggalek, Tulungagung, Blitar, Kediri, Malang, Lumajang, Jember, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Pasuruan, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Ngawi, Bojonegoro, Tuban, Lamongan, Bangkalan, Sampang, Pamekasan, dan Sumenep
2	Sidoarjo, Gresik, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, dan Kota Batu

4.4.2 Pengelompokan Klaster Non Hirarki dengan Metode *K-means*

Hasil pemilihan klaster optimum pada metode *K-means* pada faktor-faktor yang memengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023 didapatkan 2 klaster. Hasil analisis klaster berdasarkan metode *K-Menas* faktor-faktor yang memengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023 disajikan pada tabel berikut.

Tabel 8. Hasil Analisis Klaster *K-means*

Klaster	Provinsi
1	Pacitan, Ponorogo, Trenggalek, Tulungagung, Blitar, Kediri, Malang, Lumajang, Jember, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Pasuruan, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Ngawi, Bojonegoro, Tuban, Bangkalan, Sampang, Pamekasan, dan Sumenep
2	Sidoarjo, Lamongan, Gresik, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, dan Kota Batu

Pada metode *K-means* digunakan analisis ANOVA untuk melihat perbedaan variabel pada klaster yang terbentuk. Adapun hasil ANOVA sebagai berikut.

Hipotesis:

H_0 : Tidak terdapat perbedaan signifikan pada semua kelompok pada klaster

H_1 : Minimal ada satu kelompok pada klaster yang berbeda signifikan

Dengan taraf signifikan sebesar 5% dan daerah kritis tolak jika $P-Value < \alpha$. Sehingga diperoleh hasil statistik uji sebagai berikut.

Tabel 7. ANOVA

Variabel	F-hitung	<i>P-Value</i>
X1	19,808	0,000
X2	19,261	0,000
X3	80,664	0,000
X4	12,138	0,001
X5	2,169	0,150
X6	15,144	0,000

Berdasarkan Tabel 4.5 diketahui bahwa variabel X1, X2, X3, X4, dan X6 memiliki nilai *P-Value* yang kurang dari 0,05. Hal ini menunjukkan rata-rata pengeluaran makanan, garis kemiskinan, tingkat pengangguran terbuka, dan rata-rata pendapatan bersih sebulan pekerja informal memberikan pengaruh yang signifikan terhadap pengelompokan daerah-daerah berdasarkan faktor-faktor yang memengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023. Sedangkan variabel X5 memiliki nilai *P-Value* lebih dari 0,05 yang menunjukkan tingkat partisipasi angkatan kerja tidak memberikan pengaruh signifikan terhadap pengelompokan daerah-daerah berdasarkan faktor-faktor yang memengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023.

4.5 Karakteristik Klaster

Karakteristik klaster faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023 sebagai berikut.

1) Karakteristik Klaster Hirarki

Karakteristik klaster Hirarki menggunakan metode *complete linkage* berdasarkan faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023 sebagai berikut.

Tabel 9. Karakteristik Klaster Hirarki

Variabel	Klaster	Mean	Min	Maks
X1	1	12,033	6,530	21,760
	2	6,020	3,310	10,960
X2	1	595750,444	456234	1052321
	2	761928,182	576394	948701
X3	1	465679,926	370643	547017
	2	633861,364	554195	742678
X4	1	4,229	1,710	6,180
	2	5,727	4,060	8,050
X5	1	73,810	66,890	81,640
	2	71,562	67,580	78,990
X6	1	1531319,963	1111614	2154102
	2	1904833	1472768	2860798

Berdasarkan Tabel 9 diketahui bahwa klaster 1 memiliki nilai rata-rata faktor lebih tinggi pada variabel X1 dan X5, sedangkan klaster 2 memiliki nilai rata-rata faktor lebih tinggi pada variabel X2, X3, X4, dan X6. Hal ini menunjukkan bahwa daerah-daerah yang termasuk klaster 1 memiliki persentase penduduk miskin dan tingkat partisipasi angkatan kerja yang lebih tinggi daripada daerah pada klaster 2, sedangkan daerah-daerah yang termasuk klaster 2 memiliki rata-rata pengeluaran makanan, garis kemiskinan, tingkat pengangguran terbuka, dan rata-rata pendapatan bersih sebulan pekerja informal yang lebih tinggi daripada daerah pada klaster 1.

2) Karakteristik Klaster Non Hirarki

Karakteristik klaster Non Hirarki menggunakan metode *K-means* berdasarkan faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023 sebagai berikut.

Tabel 10. Karakteristik Klaster Non Hirarki

Variabel	Klaster	Mean	Min	Maks
X1	1	12,018	6,530	21,760

ANALISIS KLASTER PADA FAKTOR-FAKTOR YANG MEMPENGARUHI INDIKATOR KESEJAHTERAAN SOSIAL DAN EKONOMI DI PROVINSI JAWA TIMUR TAHUN 2023

	2	6,553	3,310	12,420
X2	1	590255,923	456234	1052321
	2	759984,833	576394	948701
X3	1	463412,385	370643	547017
	2	624759,250	524636	742678
X4	1	4,182	1,710	6,180
	2	5,705	4,060	8,050
X5	1	73,761	66,890	81,640
	2	71,855	67,580	78,990
X6	1	1510027,077	1111614	2154102
	2	1919841,500	1472768	2860798

Berdasarkan Tabel 10 diketahui bahwa kluster 1 memiliki nilai rata-rata faktor lebih tinggi pada variabel X1 dan X5, sedangkan kluster 2 memiliki nilai rata-rata faktor lebih tinggi pada variabel X2, X3, X4, dan X6. Hal ini menunjukkan bahwa daerah-daerah yang termasuk kluster 1 memiliki persentase penduduk miskin dan tingkat partisipasi angkatan kerja yang lebih tinggi daripada daerah pada kluster 2, sedangkan daerah-daerah yang termasuk kluster 2 memiliki rata-rata pengeluaran makanan, garis kemiskinan, tingkat pengangguran terbuka, dan rata-rata pendapatan bersih sebulan pekerja informal yang lebih tinggi daripada daerah pada kluster 1.

4.6 Pemetaan faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023

Pemetaan hasil faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023 berdasarkan kluster disajikan sebagai berikut.



Gambar 2 Pemetaan Kluster

Berdasarkan Gambar 2 dapat diketahui bahwa warna merah muda merupakan kluster 1 dan warna ungu merupakan kluster 2. Mapping tersebut menunjukkan bahwa kluster 1 mayoritas ada pada kabupaten di Jawa Timur, daerah ini lebih bergantung pada sektor pertanian dan memiliki akses yang lebih terbatas terhadap pusat industri dan

infrastruktur sehingga tingkat kesejahteraan sosial dan ekonomi masyarakat cenderung lebih rendah dibandingkan dengan klaster 2. Sedangkan klaster 2 berisikan wilayah kota yang ada di Jawa Timur yang merepakan pusat ekonomi, pendidikan dan industri di Jawa Timur, akses yang lebih baik terhadap lapangan pekerjaan, fasilitas pendidikan dan infrastruktur modern di klaster 2 mendukung tingkat kesejahteraan sosial dan ekonomi yang lebih tinggi. Pola ini menunjukkan bahwa perbedaan geografis antara wilayah perkotaan dan kabupaten berkontribusi signifikan terhadap ketimpangan kesejahteraan sosial dan ekonomi di Jawa Timur.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil yang dijelaskan pada analisis dan pembahasan didapatkan kesimpulan dari penelitian sebagai berikut.

1. Berdasarkan karakteristik diketahui bahwa keragaman antar provinsi cenderung kecil atau homogen karena nilai rata-rata lebih besar daripada nilai standar deviasinya.
2. Berdasarkan pengujian asumsi untuk analisis klaster berdasarkan faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi di Provinsi Jawa Timur tahun 2023 adalah asumsi distribusi normal multivariat dan asumsi dependensi.
3. Berdasarkan analisis klaster hirarki dengan nilai *Pseudo-F* paling tinggi menggunakan metode *complete linkage* terbentuk 3 klaster didapatkan 27 kabupaten/kota masuk ke dalam klaster 1 dan 11 kabupaten/kota masuk ke dalam klaster 2 .
4. Berdasarkan analisis klaster non hirarki dengan metode *K-means* didapatkan 26 kabupaten/kota masuk ke dalam klaster 1 dan 12 kabupaten/kota masuk ke dalam klaster 2 dan hasil pemetaan sudah sesuai kriteria yang cenderung sama dimana wilayah kota mayoritas tergabung dalam satu klaster, dan klaster lainnya di dominasi oleh wilayah kabupaten.

5.2 Saran

Saran yang didapat berdasarkan penelitian ini yaitu bagi pemerintah Provinsi Jawa Timur untuk lebih memperhatikan keadaan kesejahteraan sosial dan ekonomi daerahnya khususnya di daerah kabupaten, sehingga dapat mewujudkan kesetaraan antar daerah. Adapun saran bagi pembaca yaitu mengetahui pengelompokan klaster di Provinsi Jawa Timur tahun 2023 berdasarkan faktor-faktor yang mempengaruhi kesejahteraan sosial dan ekonomi.

DAFTAR REFERENSI

- Apriliana, T., & Widodo, E. (2023). Analisis Cluster Hierarki untuk Pengelompokan Provinsi di Indonesia berdasarkan Jumlah Base Transceiver Station dan Kekuatan Sinyal. *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*.

- Asiska, N., Satyahadewi, N., & Perdana, H. (2019). PENCARIAN CLUSTER OPTIMUM PADA SINGLE LINKAGE, COMPLETE LINKAGE DAN AVERAGE LINKAGE. *Buletin Ilmiah Math, Stat, dan Terapannya (Bimaster)*.
- Badan Pusat Statistik (BPS) Jawa Timur. (2024). *Garis Kemiskinan Menurut Kabupaten/Kota di Jawa Timur (Rupiah/Bulan/Kapita), 2022-2024*.
- Badan Pusat Statistik. (2024). Persentase Penduduk Miskin (Persen), 2024.
- Badruttaman, A., Sudarno, & Maruddani, D. I. (2020). Penerapan Analisis Kluster K Modes dengan Validasi Davies Bouldin Index dalam Menentukan Karakteristik Kanal YouTube di Indonesia vol 9; halaman 263-272. *JURNAL GAUSSIAN, Volume 9, Nomor 3, Tahun 2020, Halaman 263-272*.
- Farisi, S. A., & dkk. (2022). Peran UMKM (Usaha Mikro Kecil Menengah) dalam Meningkatkan Kesejahteraan Masyarakat) vol 9; halaman 79. *Jurnal Dinamika Ekonomi Syariah*.
- Hasmira, Alwi, W., & Nurfadhilah, K. (2023). Kemiskinan, Penentuan Cluster Hirarki Optimum dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan dengan Indikator. *Jurnal Matematika dan Statistika serta Aplikasinya*.
- I Wayan Terimajaya, N. L. (2024). *Dasar-Dasar Statistika : Konsep dan Metode Analisis*.
- Indra Kertati, H. S. (2023). *Implementasi Kebijakan Publik : Dari Hulu ke Hilir*.
- Ismail, I. H., Bahriani, M., Wahyudi, E., Nurrahman, A., Negara, H. R., Syafii, A., . . . M, M. A. (2024). *ANALISIS DATA KUANTITATIF DENGAN PROGRAM R*.
- Khoirunnisa, H., Ruchjana, B. N., Irianingsih, I., & Suhadi, B. (2020). Perbandingan Penerapan Metode Agglomerative dengan Metode K-Means pada Data Curah Hujan di Wilayah Bogor. *KUBIK: Jurnal Publikasi Ilmiah Matematika* .
- Luh Titi Handayani, A. (2023). *STATISTIK DESKRIPTIF*.
- Pemerintah Kabupaten Pematang. (2017). *Pengeluaran Penduduk*. Pematang.
- Salsabila, F., Fitriani, I., Umidah, Y., & Heryana, N. (2023). PENERAPAN METODE CRISP-DM UNTUK ANALISA PENDAPATAN BERSIH BULANAN PEKERJA INFORMAL DI PROVINSI JAWA BARAT DENGAN ALGORITMA K-MEANS. *DINAMIK*.
- Saputra, E. A., & Natalini, Y. (2021). Analisis Pengelompokan Data Nilai Siswa untuk Menentukan Siswa Berprestasi Menggunakan Metode Clustering K-Means. *Journal of Information Systems and Informatics*.
- Septin Puji Astuti, E. S. (2023). *Pengantar dan Analisis Desain Eksperimen Menggunakan MINITAB*.
- Suhaeni, C., Kurnia, A., & Ristiyani. (2018). Perbandingan Hasil Pengelompokan menggunakan Analisis Cluster Berhirarki, K-Means Cluster, dan Cluster Ensemble (Studi Kasus Data Indikator Pelayanan Kesehatan Ibu Hamil). *Jurnal Media Infotama*.
- uryana, D. (2024). *Matematika: Matematika Jilid 2* .
- Walpole, R. E. (2017). *Pengantar Statistika. In Standard Methods for the Examination of Water and Wastewater (Vol. 1)*.

