

## Perbandingan Analisis *Cluster* Metode *Complete Linkage* dan Metode *Ward* dalam Pengelompokan Indeks Pembangunan Manusia di Sulawesi Selatan

Irwan<sup>1,a)</sup>, Wahidah Sanusi<sup>1,b)</sup>, Afifatun Hasanah<sup>1,c)</sup>

<sup>1</sup>Jurusan Matematika, Universitas Negeri Makassar

<sup>a)</sup>irwanthaha@unm.ac.id

<sup>b)</sup>wahidah.sanusi@unm.ac.id

<sup>c)</sup>afifatunhasanah7@gmail.com

**Abstrak.** Penelitian ini adalah penelitian terapan yang menerapkan teori-teori analisis cluster metode complete linkage dan metode ward dalam mengelompokkan Kabupaten/Kota di Provinsi Sulawesi Selatan berdasarkan Indeks Pembangunan Manusia pada tahun 2022. Analisis cluster merupakan cabang ilmu statistik analisis multivariat yang bertujuan untuk mengelompokkan objek-objek berdasarkan kemiripan karakteristik di antara objek-objek tersebut. Metode complete linkage ditentukan dari jarak terjauh antara dua objek pada cluster yang berbeda. Metode ward merupakan pengelompokan yang mampu meminimumkan Sum of Square (SSE). IPM merupakan indikator penting yang menunjukkan perkembangan dalam pembangunan sumber daya manusia dan kesejahterannya. Data penelitian bersumber dari BPS tahun 2022 dan variabel yang digunakan yaitu umur harapan hidup, harapan lama sekolah, rata-rata lama sekolah, dan pengeluaran per kapita yang disesuaikan. Hasil penelitian ini menunjukkan bahwa pengelompokan terbaik berdasarkan rasio simpangan baku terkecil dengan nilai 0,282 menggunakan metode Ward dengan 5 cluster. Cluster 1: Kepulauan Selayar, Jeneponto, Takalar, Sinjai, Pangkajene dan Kepulauan, Bone, Wajo, dan Luwu Utara. Cluster 2: Bulukumba, Gowa, Maros, Barru, Soppeng, Enrekang, dan Luwu. Cluster 3: Bantaeng, Sidenreng Rappang, dan Pinrang. Cluster 4: Tana Toraja dan Toraja Utara. Cluster 5: Makassar, Pare-pare, dan Palopo.

**Kata Kunci:** Analisis Cluster, Metode Complete Linkage, Metode Ward, IPM.

**Abstrack.** This research is applied research that apply cluster analysis theories, complete linkage method and ward method in grouping Districts/Cities in South Sulawesi Province based on the Human Development Index in 2022. Cluster analysis is a branch of statistical science, multivariate analysis which aims to group objects based on the similarity of characteristics between these objects. The complete linkage method is determined from the farthest distance between two objects in different clusters. The ward method is a grouping that is able to minimize the Sum of Square (SSE). HDI is an important indicator that shows progress in human resource development and welfare. Research data comes from BPS in 2022 and the variables used are life expectancy, expected years of schooling, average years of schooling, and adjusted per capita expenditure. The results of this research show that the best grouping is based on the smallest standard deviation ratio with a value of 0.282 using the ward method with 5 clusters. Cluster 1: Selayar Islands, Jeneponto, Takalar, Sinjai, Pangkajene and the Islands, Bone, Wajo, North Luwu. Cluster 2: Bulukumba, Gowa, Maros, Barru, Soppeng, Enrekang, Luwu. Cluster 3: Bantaeng, Sidenreng Rappang, Pinrang. Cluster 4: Tana Toraja and North Toraja. Cluster 5: Makassar, Pare-pare, Palopo.

**Key Words:** Cluster Analysis, Complete Linkage Method, Ward Method, IPM.

## PENDAHULUAN

Analisis *cluster* merupakan cabang ilmu statistik analisis multivariat yang bertujuan untuk mengelompokkan objek-objek berdasarkan kemiripan karakteristik di antara objek-objek tersebut. Analisis *cluster* memiliki dua metode yaitu, metode hirarki dan non-hirarki. Metode hirarki yang sering dipakai adalah metode agglomeratif atau penggabungan. Proses agglomeratif dimulai dengan setiap objek dalam suatu *cluster* yang terpisah kemudian objek dikelompokkan ke dalam *cluster* yang semakin banyak anggotanya (Supranto, 2004). Metode agglomeratif terdiri dari tiga jenis metode yaitu, metode *linkage*, metode *variance* dan metode *centroid*. Metode *linkage* digunakan untuk menentukan ukuran jarak antar *cluster* dengan beberapa metode yaitu *single linkage*, *complete linkage*, dan *average linkage*. Metode *variance* merupakan metode yang berusaha menghasilkan *cluster* dengan meminimumkan ragam dalam *cluster*. Metode *variance* yang sering dipergunakan adalah metode *ward*. Metode *centroid* diukur jarak antara dua *cluster* merupakan jarak antara rata-rata seluruh variabel. Seluruh metode agglomeratif memiliki prosedur berbeda-beda dalam proses *clustering*, oleh karena itu dari setiap metode memiliki kelebihan dan kekurangan masing-masing (Johnson dan Wichern, 2007; Inayatus dan Fauziyyah, 2021).

Indeks Pembangunan Manusia (IPM) merupakan indikator penting yang menunjukkan perkembangan dalam pembangunan sumber daya manusia dan kesejahterannya. Tujuan utama adanya IPM adalah menciptakan lingkungan yang memungkinkan bagi rakyatnya untuk menikmati umur panjang, sehat, dan menjalankan hidup produktif. IPM dapat mengukur keberhasilan dalam upaya membangun kualitas hidup manusia (masyarakat/penduduk) dan menentukan peringkat atau level pembangunan di suatu wilayah/negara. IPM diukur berdasarkan aspek kesehatan, pendidikan, dan ekonomi. Pada aspek kesehatan diukur melalui angka Umur Harapan Hidup (UHH), aspek pendidikan diukur melalui angka Harapan Lama Sekolah (HLS) dan Rata-Rata Lama Sekolah (RLS), sedangkan aspek ekonomi diukur melalui angka pengeluaran kapita per-tahun yang disesuaikan (A'laa, 2023).

Provinsi Sulawesi Selatan pada tahun 2022 memiliki angka IPM sebesar 72,82 yang merupakan kategori dengan IPM tinggi. Angka tersebut merupakan peningkatan dari angka IPM tahun 2021 sebesar 72,24. Hal ini menunjukkan bahwa proses pembangunan manusia yang dilakukan di Provinsi Sulawesi Selatan berhasil. Semakin tinggi angka IPM di suatu wilayah maka wilayah tersebut mengalami keberhasilan dalam proses pembangunannya. Untuk mewujudkan keberhasilan pembangunan perlu dilakukan pendataan wilayah mana saja yang membutuhkan penanganan lebih untuk memperbaiki proses pembangunannya (Tambunan, 2003).

Upaya yang dapat dilakukan untuk mewujudkan keberhasilan pembangunan di suatu wilayah adalah mengelompokkan wilayah-wilayah mana saja yang belum berhasil pada proses pembangunannya sehingga proses pembangunan yang dilakukan oleh pemerintah dapat sesuai dengan keadaan atau kondisi pembangunan di wilayah tersebut. Oleh karena itu pengelompokkan karakteristik tiap wilayah perlu dilakukan sehingga akan lebih mudah untuk mencapai keberhasilan dalam proses pembangunan wilayah dan digunakan sebagai saran untuk evaluasi kebijakan pemerintah yang telah diterapkan demi terwujudnya pemerataan pembangunan yang lebih terstruktur dan merata berdasarkan data dan karakteristik wilayah tersebut. Salah satu upaya yang dapat dilakukan dalam pengelompokkan adalah dengan menggunakan analisis *cluster*. (Wicaksono dan Yolanda, 2021). Tujuan penelitian ini adalah untuk mengelompokkan Kabupaten/Kota di Provinsi Sulawesi Selatan berdasarkan IPM menggunakan Metode *Complete Linkage* dan Metode *Ward* serta membandingkan kedua metode tersebut.

## TINJAUAN PUSTAKA

### Analisis Cluster

Analisis *cluster* merupakan teknik peubah ganda yang bertujuan untuk mengelompokkan  $n$  satuan pengamatan ke dalam  $k$  kelompok dengan ( $k < n$ ) berdasarkan  $p$  peubah (Mattjik & Sumertajaya, 2011). Tujuan utama analisis *cluster* adalah mengelompokkan objek-objek berdasarkan kesamaan karakteristik di antara objek-objek tersebut. Ciri-ciri *cluster* yang baik yaitu memiliki homogenitas (kesamaan) yang tinggi antara anggota dalam satu *cluster* (*within-cluster*) dan memiliki heterogenitas (perbedaan) yang tinggi antara *cluster* yang satu dengan *cluster* lainnya (*between cluster*). Dari dua hal tersebut dapat disimpulkan bahwa hasil analisis *cluster* yang baik adalah *cluster* yang dimana anggota-anggotanya memiliki kemiripan karekteristik yang tinggi. Sedangkan untuk *cluster* yang satu dengan yang *cluster* yang lainnya memiliki karakteristik yang sangat berbeda (Santoso, 2018).

### Uji Asumsi Analisis Cluster

1. Uji Kecukupan Sampel dilakukan untuk mengetahui apakah sampel sudah cukup mewakili populasi yang dapat diketahui dari nilai *Kaiser Meyer Olki* (KMO) yang merupakan indeks perbandingan nilai koefisien korelasi terhadap korelasi parsial. Suatu data dapat dikatakan mewakili populasi apabila  $0,5 < KMO < 1$ . Rumus KMO disajikan pada Persamaan (1) (Mahmudan, 2020; Andhika, dkk., 2021).

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2} \quad (1)$$

$$a_{ij} = \frac{-r_{ij}}{\sqrt{r_{ij}r_{ij}}} \quad (2)$$

dimana:

$p$  = banyaknya variabel

$r_{ij}$  = koefisien korelasi antara variabel  $i$  dan  $j$

$a_{ij}$  = koefisien korelasi parsial antara variabel  $i$  dan  $j$

2. Uji Multikolinieritas dilakukan untuk mengetahui hubungan linear atau korelasi yang tinggi antar variabel. Jika terjadi multikolinieritas pada salah satu variabel maka pengurangan variabel yang terjadi multikolinieritas dapat dilakukan. Suatu data dikatakan terbebas dari multikolinieritas apabila memiliki nilai *Tolerance*  $> 0,1$  dan *VIF*  $< 10$  (Sriningsih, dkk., 2018).

### Standardisasi Data

Standardisasi data dilakukan apabila terdapat perbedaan satuan yang signifikan di antara variabel yang diteliti agar dapat menghindari adanya pengaruh antara variabel satu dengan variabel lainnya. Bentuk standardisasi yang paling umum adalah mengonversi setiap variabel menjadi nilai standar atau disebut juga dengan *Z-Score* (Whendasmoro dan Joseph, 2022).

$$Z = \frac{x_i - \bar{x}}{s} \quad (3)$$

dimana:

$Z$  = nilai standardisasi data

$x_i$  = nilai data ke- $i$

$\bar{x}$  = nilai rata-rata

$s$  = standar deviasi

### Jarak Euclidean

Ukuran jarak menunjukkan kesamaan antara objek. Semakin besar jarak maka menunjukkan ketaksamaan antar objek, sebaliknya semakin kecil jarak maka menunjukkan kesamaan antar objek. Jarak euclidean didapatkan dengan cara mengkuadratkan jarak antar dua objek yang akan diukur (Johnson dan Wichern, 2007).

$$d_{ij} = \sqrt{\sum_{i=1}^p (x_{ik} - x_{jk})^2} \quad (4)$$

dimana:

$d_{ij}$  = jarak *euclidean* antara objek ke-i dengan objek ke-j

$p$  = banyak variabel *cluster*

$x_{ik}$  = nilai atau data dari objek ke-i pada variabel ke-k

$x_{jk}$  = nilai atau data dari objek ke-j pada variabel ke-k

### Metode Complete Linkage

Metode *Complete Linkage* (*farthest-neighbour method*) merupakan jarak antar *cluster* yang ditentukan oleh jarak terjauh antara dua objek dalam *cluster* yang berbeda. Dengan rumus yang disajikan pada Persamaan (5) (Johnson dan Wichern, 2007).

$$d_{(IJ)K} = \max (d_{IK}, d_{JK}) \quad (5)$$

dimana  $d_{IK}$  dan  $d_{JK}$  merupakan jarak antara anggota yang paling jauh dari *cluster I* dan *J* serta *cluster J* dan *K*.

### Metode Ward

Metode *Ward* merupakan suatu metode *cluster* hirarki yang didasari oleh hilangnya informasi akibat penggabungan objek menjadi *cluster*. Tujuan metode *ward* adalah meminimumkan ragam dalam *cluster* dan memaksimalkan ragam antar objek di *cluster* lainnya (Supranto, 2004, #). Proses pengelompokan pada metode *ward* didasarkan pada kriteria *Sum of Squares Error* (*SSE*) dengan ukuran kehomogenan antara dua objek berdasarkan jumlah kuadrat kesalahan minimal. Nilai *SSE* untuk *cluster* dengan anggota satu objek atau satu item saja adalah bernilai nol (Gundono, 2011). Jika terdapat  $N$  item yang dikelompokkan menjadi satu *cluster*, maka dapat diketahui perhitungan nilai *SSE* disajikan pada Persamaan (6) (Imasdiani, dkk., 2022).

$$SSE = \sum_{i=1}^n (X_i - \bar{X})' - (X_i - \bar{X}) \quad (6)$$

dimana:

$X_i$  = nilai data ke-i dengan  $i=1,2,3,\dots,n$

$\underline{X}$  = vektor kolom yang berisi rata-rata nilai pengamatan dalam *cluster*

$n$  = banyaknya pengamatan

### Interpretasi Cluster

Interpretasi *cluster* digunakan untuk mengetahui karakteristik tiap *cluster* dan dapat dilakukan dengan membandingkan rata-rata setiap variabel pada setiap *cluster* (*centroid*). Adapun cara menghitung *centroid* disajikan pada Persamaan (7).

$$C = \frac{\sum_{j=1}^n \sum_{k=1}^p x_{jk}}{N} \quad (7)$$

dimana:

$C$  = nilai rata-rata (*centroid*) *cluster*

$x_{jk}$  = nilai atau data dari objek  $j$  pada variabel  $k$  dalam *cluster* yang dicari nilai *centroid*

$N$  = jumlah objek dalam *cluster* yang dicari nilai *centroid*.

### Simpangan Baku

Penentuan metode terbaik pada hasil analisis *cluster* yang optimal adalah apabila objek dalam satu *cluster* memiliki sifat yang homogen dan heterogen antar *cluster*. Dalam analisis statistika yang digunakan untuk mengetahui kehomogenan antar kelompok adalah ragam atau dapat pula digunakan simpangan baku. Hasil kelompok *cluster* dikatakan baik apabila mempunyai nilai simpangan baku dalam *cluster* ( $S_w$ ) yang minimum dan nilai simpangan baku antar *cluster* ( $S_b$ ) yang maksimum. Rumus ( $S_w$ ) disajikan pada Persamaan (8) (Hardle dan Simar, 2015).

$$S_w = \frac{1}{g} \sum_{j=1}^g s_j \quad (8)$$

dimana:

$g$  = banyak *cluster* yang terbentuk

$s_j$  = simpangan baku antar *cluster*

Sedangkan rumus ( $S_b$ ) disajikan pada Persamaan (9) (Hardle dan Simar, 2015).

$$S_b = \frac{1}{g-1} \sum_{j=1}^g (\bar{X}_j - \bar{X})^2 \quad (9)$$

dimana:

$g$  = banyak *cluster* yang terbentuk

$\bar{X}_j$  = rata-rata *cluster* ke- $j$ , dimana  $j=1,2,3,\dots,g$

$\bar{X}$  = rata-rata seluruh *cluster*

Hasil perhitungan yang menghasilkan nilai rasio simpangan baku terkecil merupakan metode yang memiliki kebaikan metode yang terbaik. Rumus perhitungan nilai rasio simpangan baku disajikan pada Persamaan (10) (Hardle dan Simar, 2015).

$$\text{rasio simpangan baku} = \frac{S_b}{S_w} \quad (10)$$

### Indeks Pembangunan Manusia

Pembangunan manusia menurut *United Nations Development Programme* (UNDP) merupakan model pembangunan yang ditujukan untuk memperluas pilihan yang dapat ditumbuhkan melalui upaya pemberdayaan penduduk. Pemberdayaan penduduk ini dapat dicapai melalui upaya yang menitikberatkan pada peningkatan kemampuan dasar manusia yaitu meningkatkan derajat kesehatan, pengetahuan, dan keterampilan agar dapat digunakan untuk mempertinggi partisipasi dalam kegiatan ekonomi produktif, sosial budaya, dan politik (Zulyanto, 2016).

Indeks Pembangunan Manusia (IPM) merupakan ukuran untuk mengetahui dampak kinerja pembangunan wilayah karena menunjukkan kualitas penduduk suatu wilayah dalam hal harapan hidup, pendidikan, serta standar hidup yang layak (Latuconsina, 2017). IPM merupakan suatu angka yang mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup yang dapat mempengaruhi tingkat produktivitas yang dihasilkan oleh seseorang (Mahroji dan Nurkhasanah, 2019).

*Human Development Index* (HDI) melakukan pengukuran rata-rata pencapaian yang diraih oleh setiap warga negara yang menyangkut tiga hal dasar dari pengembangan kualitas manusia. Pengukuran ini menetapkan asumsi dasar bahwa manusia yang berkualitas adalah manusia yang

bisa hidup sehat dan panjang umur, manusia yang memiliki pendidikan yang cakap bagi kelangsungan hidupnya, dan manusia yang dapat mencapai standar yang layak dikehidupannya (Zulyanto, 2016).

Menurut UNDP, IPM digunakan untuk mengukur pembangunan manusia dengan tiga dimensi pembentuk IPM yaitu umur panjang dan hidup sehat, pengetahuan, dan standar hidup layak. Dari ketiga dimensi tersebut diturunkan empat indikator yang digunakan dalam penghitungan IPM yaitu, Umur Harapan Hidup (UHH) (Tahun), Rata-rata Lama Sekolah (RLS) (Tahun), Harapan Lama Sekolah (HLS) (Tahun), dan Pengeluaran per kapita disesuaikan (Tahun) (A'laa, 2023).

## METODE PENELITIAN

Penelitian ini merupakan penelitian terapan mengenai analisis *cluster* metode *complete linkage* dan metode *ward* dalam pengelompokkan kabupaten/kota di Sulawesi Selatan berdasarkan IPM tahun 2022 dengan objek penelitian sebanyak 24 Kab/Kota di Provinsi Sulawesi Selatan. Berikut variabel yang digunakan pada penelitian ini:

1. Umur Harapan Hidup (UHH) merupakan rata-rata perkiraan jumlah tahun yang dapat ditempuh seseorang sejak lahir.
2. Harapan Lama Sekolah (HLS) merupakan jangka waktu sekolah yang diharapkan akan dilalui oleh seorang anak pada umur tertentu di masa mendatang.
3. Rata-rata Lama Sekolah (RLS) merupakan jumlah tahun yang digunakan oleh seseorang dalam menempuh pendidikan formal.
4. Pengeluaran per kapita per tahun yang disesuaikan ditentukan dari nilai pengeluaran per kapita dan paritas daya beli.

Adapun langkah-langkah yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Mengumpulkan data IPM di Provinsi Sulawesi Selatan
2. Standardisasi data menggunakan Persamaan (3)
3. Pengukuran jarak *Euclidean* menggunakan Persamaan (4)
4. Pengelompokkan menggunakan metode *complete linkage* dan metode *ward*
  - a) Langkah pengelompokkan menggunakan metode *Complete Linkage*
    - 1) Menentukan objek bersesuaian yang memiliki jarak terdekat dalam matriks jarak  $D = \{d_{ij}\}$ ,
    - 2) Menggabungkan objek yang bersesuaian, misal objek  $I$  dan objek  $J$  yang kemudian didapatkan *cluster* ( $IJ$ ),
    - 3) Menghitung jarak maksimum antar *cluster* ( $IJ$ ) dengan objek lain katakanlah  $K$  yang belum bergabung menggunakan Persamaan (5) dan menghitung kembali matriks jarak baru dengan cara pada langkah ketiga algoritma *cluster* hirarki agglomeratif secara umum dan beri nama  $D_2$ ,
    - 4) Mengulangi langkah a).1) sampai a).3) sehingga tergabung menjadi satu *cluster*.
  - b) Langkah pengelompokkan menggunakan metode *Ward*
    - 1) Dimulai dengan memperhatikan  $N$  *cluster* yang mempunyai satu objek per *cluster* (semua responden dianggap sebagai *cluster*). Pada tahap pertama  $SSE$  bernilai nol karena setiap objek akan membentuk *cluster*,

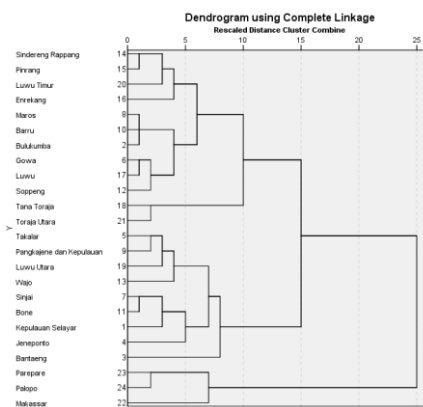
- 2) *Cluster* pertama dibentuk dengan memilih dua dari  $N$  *cluster* yang memiliki nilai *SSE* terkecil. Hal ini sejalan dengan fungsi tujuannya yaitu meminimumkan keheterogenan. Nilai *SSE* dihitung dengan Persamaan (6),
- 3)  $N-1$  kumpulan *cluster* kemudian diperhatikan lagi untuk menentukan dua dari *cluster* yang bisa meminimumkan keheterogenan. Dengan demikian  $N$  *cluster* secara sistematis dikurangi  $N-1$ ,
- 4) Mengulangi langkah b).3) dan b).4) sampai diperoleh satu *cluster* atau semua objek tergabung menjadi satu *cluster*.
5. Menentukan jumlah *cluster* yaitu 3, 4, dan 5 *cluster*
6. Interpretasi *cluster*
7. Menentukan metode terbaik dengan rasio simpangan baku dalam *cluster* ( $S_w$ ) dan simpangan baku antar *cluster* ( $S_b$ ) menggunakan Persamaan (10)

## HASIL DAN PEMBAHASAN

### Hasil Penelitian

#### *Analisis cluster metode complete linkage*

Analisis *cluster* dengan menggunakan metode *complete linkage* merupakan proses penggabungan dua objek atau lebih yang mempunyai jarak maksimum atau jarak terjauh dari anggota-anggotanya. Berikut ini merupakan hasil dendrogram dari analisis *cluster* data IPM di Provinsi Sulawesi Selatan Tahun 2022:



**GAMBAR 1.** Hasil Dendrogram Analisis *Cluster* Metode *Complete Linkage*

Penelitian ini mengelompokkan Kab/Kota di Provinsi Sulawesi Selatan yang terbagi atas tiga, empat, dan lima *cluster* dapat dilihat pada Tabel 1 sampai Tabel 3.

**TABEL 1.** Hasil dari 3 *Cluster* Metode *Complete Linkage*

<i>Cluster</i>	Jumlah Anggota	Anggota Kab/Kota
<i>Cluster 1</i>	9	Kepulauan Selayar, Bantaeng, Jeneponto, Takalar, Sinjai, Pangkajene dan Kepulauan, Bone, Wajo, dan Luwu Utara.
<i>Cluster 2</i>	12	Bulukumba, Gowa, Maros, Barru, Soppeng, Sidenreng Rappang, Pinrang, Enrekang, Luwu, Tana Toraja, Luwu Timur, dan Toraja Utara.

<i>Cluster 3</i>	3	Makassar, Pare-pare, dan Palopo.
------------------	---	----------------------------------

**TABEL 2.** Hasil dari 4 *Cluster Metode Complete Linkage*

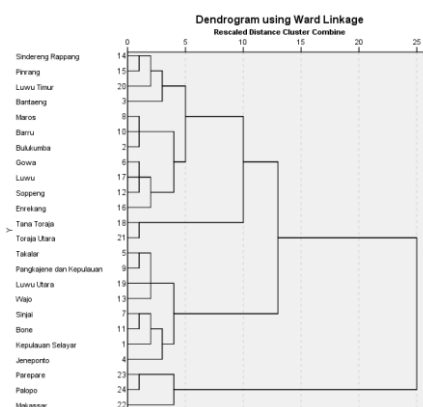
<i>Cluster</i>	<b>Jumlah Anggota</b>	<b>Anggota Kab/Kota</b>
<i>Cluster 1</i>	9	Kepulauan Selayar, Bantaeng, Jeneponto, Takalar, Sinjai, Pangkajene dan Kepulauan, Bone, Wajo, dan Luwu Utara.
<i>Cluster 2</i>	10	Bulukumba, Gowa, Maros, Barru, Soppeng, Sidenreng Rappang, Pinrang, Enrekang, Luwu, dan Luwu Timur.
<i>Cluster 3</i>	2	Tana Toraja dan Toraja Utara.
<i>Cluster 4</i>	3	Makassar, Pare-pare, dan Palopo.

**TABEL 3.** Hasil dari 5 *Cluster Metode Complete Linkage*

<i>Cluster</i>	<b>Jumlah Anggota</b>	<b>Anggota Kab/Kota</b>
<i>Cluster 1</i>	8	Kepulauan Selayar, Jeneponto, Takalar, Sinjai, Pangkajene dan Kepulauan, Bone, Wajo, dan Luwu Utara.
<i>Cluster 2</i>	10	Bulukumba, Gowa, Maros, Barru, Soppeng, Sidenreng Rappang, Pinrang, Enrekang, Luwu, dan Luwu Timur.
<i>Cluster 3</i>	1	Bantaeng.
<i>Cluster 4</i>	2	Tana Toraja dan Toraja Utara.
<i>Cluster 5</i>	3	Makassar, Pare-pare, dan Palopo.

#### *Analisis cluster metode ward*

Berikut ini merupakan hasil dendrogram dari analisis *cluster* data IPM di Provinsi Sulawesi Selatan Tahun 2022:



**GAMBAR 2.** Hasil Dendrogram Analisis *Cluster Metode Ward*

Penelitian ini mengelompokkan Kab/Kota di Provinsi Sulawesi Selatan yang terbagi atas tiga, empat, dan lima *cluster* dapat dilihat pada Tabel 4 sampai Tabel 6.

**TABEL 4.** Hasil dari 3 *Cluster Metode Ward*



<i>Cluster</i>	<b>Jumlah Anggota</b>	<b>Anggota Kab/Kota</b>
<i>Cluster 1</i>	8	Kepulauan Selayar, Jeneponto, Takalar, Sinjai, Pangkajene dan Kepulauan, Bone, Wajo, dan Luwu Utara.
<i>Cluster 2</i>	13	Bulukumba, Bantaeng, Gowa, Maros, Barru, Soppeng, Sidenreng Rappang, Pinrang, Enrekang, Luwu, Tana Toraja, Luwu Timur, dan Toraja Utara.
<i>Cluster 3</i>	3	Makassar, Pare-pare, dan Palopo.

**TABEL 5.** Hasil dari 4 *Cluster* Metode Ward

<i>Cluster</i>	<b>Jumlah Anggota</b>	<b>Anggota Kab/Kota</b>
<i>Cluster 1</i>	8	Kepulauan Selayar, Jeneponto, Takalar, Sinjai, Pangkajene dan Kepulauan, Bone, Wajo, dan Luwu Utara.
<i>Cluster 2</i>	11	Bulukumba, Bantaeng, Gowa, Maros, Barru, Soppeng, Sidenreng Rappang, Pinrang, Enrekang, Luwu, dan Luwu Timur.
<i>Cluster 3</i>	2	Tana Toraja dan Toraja Utara.
<i>Cluster 4</i>	3	Makassar, Pare-pare, dan Palopo.

**TABEL 6.** Hasil dari 5 *Cluster* Metode Ward

<i>Cluster</i>	<b>Jumlah Anggota</b>	<b>Anggota Kab/Kota</b>
<i>Cluster 1</i>	8	Kepulauan Selayar, Jeneponto, Takalar, Sinjai, Pangkajene dan Kepulauan, Bone, Wajo, dan Luwu Utara.
<i>Cluster 2</i>	7	Bulukumba, Gowa, Maros, Barru, Soppeng, Enrekang, dan Luwu
<i>Cluster 3</i>	4	Bantaeng, Sidenreng Rappang, Pinrang, dan Luwu Timur.
<i>Cluster 4</i>	2	Tana Toraja dan Toraja Utara.
<i>Cluster 5</i>	3	Makassar, Pare-pare, dan Palopo.

*Menentukan Metode Terbaik dengan Simpangan Baku*

Penentuan metode terbaik berguna untuk mengetahui seberapa baik kinerja metode *complete linkage* dan metode *ward* dalam mengelompokkan Kab/Kota di Sulawesi Selatan berdasarkan IPM. Perolehan hasil perhitungan rasio simpangan baku dari masing-masing *cluster* menggunakan metode *complete linkage* dan metode *ward* dapat dilihat pada Tabel 7.

**TABEL 7.** Rasio Simpangan Baku

<b>Banyak Cluster</b>	<b>Metode Complete Linkage</b>	<b>Metode Ward</b>
3	0,695	0,699
4	0,464	0,461
5	0,422	0,282

Berdasarkan Tabel di atas menunjukkan bahwa nilai rasio simpangan baku terkecil adalah metode *ward* dengan 5 *cluster* yang bernilai 0,282. Hal ini berarti bahwa metode tersebut merupakan metode terbaik dalam mengelompokkan Kab/Kota di Provinsi Sulawesi Selatan tahun 2022 berdasarkan IPM.

Interpretasi *cluster* digunakan untuk mengetahui karakteristik tiap *cluster* dan dapat dilakukan dengan membandingkan rata-rata setiap variabel pada setiap *cluster* (*centroid*). Adapun perolehan nilai *centroid* pada metode *ward* dengan 5 *cluster* dapat dilihat pada Tabel 8.

**TABEL 8.** *Centroid* setiap *Cluster* pada Metode *Ward 5 Cluster*

<i>Cluster</i>	<i>X1</i> (UHH)	<i>X2</i> (HLS)	<i>X3</i> (RLS)	<i>X4</i> (Pengeluaran per Kapita)
<i>Cluster 1</i>	67,81	12,77	7,59	10.658,9
<i>Cluster 2</i>	69,99	13,49	8,29	10.668,3
<i>Cluster 3</i>	70,48	12,85	7,63	12.204,7
<i>Cluster 4</i>	73,69	13,64	8,39	8.039
<i>Cluster 5</i>	71,72	15,07	11,1	14.945,7

Rata-rata setiap variabel pada indikator IPM di Provinsi Sulawesi Selatan adalah  $X1=69,89$ ,  $X2=13,36$ ,  $X3=8,36$ , dan  $X4=11.272,3$ . Berdasarkan Tabel 4.13 dapat dilihat bahwa *Cluster 1* merupakan *Cluster* dengan IPM terendah karena pada setiap variabel memiliki rata-rata dibawah rata-rata indikator IPM. Sehingga pada Kab/Kota yang berada pada *Cluster 1* perlu diperhatikan terhadap keempat variabelnya. *Cluster 2* pada variabel *X1* dan *X2* sudah mencapai lebih dari rata-rata indikator IPM, namun pada variabel *X3* dan *X4* masih dibawah rata-rata. Maka dari itu *Cluster 2* perlu diperhatikan dalam aspek pendidikan dan ekonomi. Sedangkan *Cluster 3* pada variabel *X1* dan *X4* sudah mencapai lebih dari rata-rata, namun pada variabel *X2* dan *X3* masih dibawah rata-rata indikator IPM. Maka dari itu *Cluster 3* perlu diperhatikan dalam aspek pendidikan. *Cluster 4* merupakan *Cluster* yang cukup baik karena rata-rata setiap variabel sudah mencapai lebih dari rata-rata indikator IPM terkecuali pada variabel *X4*. Maka dari itu *Cluster 4* perlu diperhatikan dalam aspek ekonomi. Sedangkan *Cluster 5* merupakan *Cluster* terbaik dan berhasil dalam pembangunan manusia karena setiap variabel sudah mencapai lebih dari rata-rata indikator IPM

## Pembahasan

Penelitian Imasdiani, dkk., (2022) mengelompokkan data Kemiskinan Kabupaten/Kota di Kalimantan Timur 2018 menggunakan metode *average linkage* dan *metode ward*, menunjukkan bahwa nilai rasio simpangan baku pada metode *ward* lebih kecil daripada metode *average linkage* yaitu sebesar 2,681.

Penelitian Hajar, dkk., (2022) mengelompokkan wilayah Desa/Kelurahan di Kabupaten Kutai Kartanegara Provinsi Kalimantan Timur berdasarkan data Potensi Desa 2018 menggunakan metode *agglomerative nesting* menunjukkan *Cluster* optimum berdasarkan nilai maksimum *Pseudo-F* yaitu 155,611 dan nilai *Icdrate* terkecil yaitu 0,596 pada metode *complete linkage* dengan jumlah *Cluster* yang terbentuk adalah dua *Cluster*.

Pada penelitian ini mengelompokkan Kabupaten/Kota di Provinsi Sulawesi Selatan dengan menggunakan metode *complete linkage* dan metode *ward* berdasarkan IPM pada tahun 2022 terbentuk *Cluster* terbaik berdasarkan perbandingan nilai rasio simpangan baku terkecil yaitu 0,282 pada metode *ward* dengan jumlah *Cluster* yang terbentuk adalah lima *Cluster*. *Cluster* pertama beranggotakan Kepulauan Selayar, Jeneponto, Takalar, Sinjai, Pangkajene dan Kepulauan, Bone, Wajo, dan Luwu Utara, *Cluster* kedua beranggotakan Bulukumba, Gowa, Maros, Barru, Soppeng, Enrekang, dan Luwu, *Cluster* ketiga beranggotakan Bantaeng, Sidenreng Rappang, dan Pinrang, *Cluster* keempat beranggotakan Tana Toraja dan Toraja Utara, dan *Cluster* kelima beranggotakan Makassar, Pare-pare, dan Palopo.

## KESIMPULAN

Adapun kesimpulan dalam penelitian pada pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan pada Tahun 2022 berdasarkan IPM dengan menggunakan metode *Complete Linkage* dan metode *Ward* adalah menghasilkan nilai rasio simpangan baku pada metode *complete linkage* dengan 3 *Cluster* = 0,695, 4 *Cluster* = 0,464, dan 5 *Cluster* = 0,422 dan nilai rasio simpangan baku pada metode *ward* dengan 3 *Cluster* = 0,699, 4 *Cluster* = 0,461, dan 5 *Cluster* = 0,282.

Hasil perbandingan analisis *Cluster* metode *complete linkage* dan metode *ward* dalam pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Tahun 2022 berdasarkan IPM menunjukkan bahwa metode terbaik berdasarkan nilai rasio simpangan baku terkecil adalah menggunakan Metode *Ward* dengan 5 *Cluster* dengan nilai 0,282. Kab/Kota *Cluster* 1 yaitu; Kepulauan Selayar, Jeneponto, Takalar, Sinjai, Pangkajene dan Kepulauan, Bone, Wajo, dan Luwu Utara. Kab/Kota *Cluster* 2 yaitu; Bulukumba, Gowa, Maros, Barru, Soppeng, Enrekang, dan Luwu. Kab/Kota *Cluster* 3 yaitu; Bantaeng, Sidenreng Rappang, dan Pinrang. Kab/Kota *Cluster* 4 yaitu; Tana Toraja dan Toraja Utara. Kab/Kota *Cluster* 5 yaitu; Makassar, Pare-pare, dan Palopo.

## DAFTAR PUSTAKA

- A'laa, R. D. (2023). *Indeks Pembangunan Manusia Provinsi Sulawesi Selatan 2022*. Badan Pusat Statistik Provinsi Sulawesi Selatan. <https://sulsel.bps.go.id/>
- Andika, L., Winner, Lazarus, & Pontoh, R. S. (2021). Application of Clustering Using the K-Means Method in Indonesian Provinces Based on Infrastructure Data in 2020. 15-15. [https://scholar.google.com/scholar?lookup=0&q=application+of+clustering+using+k-means+method+in+indonesian&hl=id&as\\_sdt=0,5#d=gs\\_qabs&t=1710890772150&u=%23p%3DAwpg\\_x7-eOAJ](https://scholar.google.com/scholar?lookup=0&q=application+of+clustering+using+k-means+method+in+indonesian&hl=id&as_sdt=0,5#d=gs_qabs&t=1710890772150&u=%23p%3DAwpg_x7-eOAJ)
- Hajar, M., Nasution, Y. N., & Prangga, S. (2022). Penerapan Metode Agglomerative Nesting (AGNES) pada Pengelompokan Wilayah Desa atau Kelurahan di Kabupaten Kutai Kartagaya. *Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya*, 2, 456-469. <https://jurnal.fmipa.unmul.ac.id/index.php/SNMSA/article/view/922>
- Hardle, W., & Simar, L. (2015). *Applied Multivariate Statistical Analysis* (4th ed.). Springer-Verlag. [https://eprints.ukh.ac.id/id/eprint/185/1/2015\\_Book\\_AppliedMultivariateStatistical.pdf](https://eprints.ukh.ac.id/id/eprint/185/1/2015_Book_AppliedMultivariateStatistical.pdf)
- Imasdiani, Purnamasari, I., & Amijaya, F. D. T. (2022). Perbandingan Analisis Cluster dengan Menggunakan Metode Average Linkage dan Metode Ward. *EKSPONENSIAL*, 13(1), 9-18. <https://doi.org/10.30872/eksponsional.v13i1.875>
- Inayatus, & Fauziyyah. (2021, Februari 22). *Introduction to Hierarchical Clustering*. <https://algotech.netlify.app/blog/introduction-to-hierarchical-clustering/>
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis* (6th ed.). Pearson Education, Inc. <https://www.webpages.uidaho.edu/~stevel/519/Applied%20Multivariate%20Statistical%20Analysis%20by%20Johnson%20and%20Wichern.pdf>
- Latuconsina, Z. M. Y. (2017). Analisis Faktor-faktor yang Mempengaruhi Indeks Pembangunan Manusia Kabupaten Malang Berbasis Pendekatan Perwilayahan dan Regresi Panel. *Jurnal Perencanaan Pembangunan Wilayah dan Pedesaan*, 1(2), 202-216. <https://doi.org/10.29244/jp2wd.2017.1.2.202-216>

- Mahmudan, A. (2020). Clustering of District or City in Central Java Based COVID-19 Case Using K-Means Clustering. *Jurnal Matematika, Statistika, dan Komputasi*, 17(1), 1-13. 10.20956/jmsk.v%vi%i.10727
- Mahroji, D., & Nurkhasanah, I. (2019). Pengaruh Indeks Pembangunan Manusia terhadap Tingkat Pengangguran di Provinsi Banten. *Jurnal Ekonomi-Qu*, 9(1), 51-72. <http://dx.doi.org/10.35448/jequ.v9i1.5436>
- Mattjik, A. A., & Sumertajaya, I. M. (2011). *Sidik Peubah Ganda dengan Menggunakan SAS*. IPB PRESS. [https://scholar.google.co.id/citations?view\\_op=view\\_citation&hl=id&user=4rfuzqoAAAJ&citation\\_for\\_view=4rfuzqoAAAJ:JQOojiI6XY0C](https://scholar.google.co.id/citations?view_op=view_citation&hl=id&user=4rfuzqoAAAJ&citation_for_view=4rfuzqoAAAJ:JQOojiI6XY0C)
- Santoso, S. (2018). *Mahir Statistik Multivariat dengan SPSS*. Elex Media Komputindo. [https://scholar.google.co.id/citations?view\\_op=view\\_citation&hl=id&user=4rfuzqoAAAJ&citation\\_for\\_view=4rfuzqoAAAJ:JQOojiI6XY0C](https://scholar.google.co.id/citations?view_op=view_citation&hl=id&user=4rfuzqoAAAJ&citation_for_view=4rfuzqoAAAJ:JQOojiI6XY0C)
- Sriningsih, M., Hatidja, D., & Prang, J. D. (2018). Penanganan Multikolinearitas dengan Menggunakan Analisis Regresi Komponen Utama pada Kasus Impor Beras di Provinsi Sulut. *Jurnal Ilmiah Sains*, 18(1), 18-24. <https://doi.org/10.35799/jis.18.1.2018.19396>
- Supranto. (2004). *Analisis Multivariat Arti dan Interpretasi*. Rineka Cipta.
- Tambunan, T. T. H. (2003). *Perekonomian Indonesia*. Jakarta Ghalia Indonesia. <http://kin.perpusnas.go.id/DisplayData.aspx?pId=108840&pRegionCode=UNTAR&pClientId=650>
- Whendasromo, R. G., & Joseph. (2022). Analisis Penerapan Normalisasi Data dengan Menggunakan Z-Score pada Kinerja Algoritma K-NN. *JURIKOM*, 9(4), 872-876. <http://dx.doi.org/10.30865/jurikom.v9i4.4526>
- Wicaksono, A. S., & Yolanda, A. M. (2021). Pengelompokan Kabupaten/Kota di Provinsi Nusa Tenggara Timur Berdasarkan Indikator Indeks Pembangunan Manusia Menggunakan K-Medoids Clustering. *Jurnal Statistika Terapan*, 1(1), 79-90. <https://doi.org/10.5300/jstar.v1i1.7>
- Zulyanto, A. (2016). Pengeluaran Pemerintah dan Indeks Pembangunan Manusia (IPM) di Provinsi Bengkulu. *Jurnal Ekonomi-Qu*, 6(2), 115-139. <http://dx.doi.org/10.35448/jequ.v6i2.4340>