

# PENGELOMPOKAN KABUPATEN/KOTA DI SULAWESI SELATAN BERDASARKAN INDIKATOR KESEHATAN DENGAN MENGGUNAKAN METODE *AVERAGE LINKAGE*

Wahidah Alwi<sup>1</sup>, Adnan Sauddin<sup>1</sup>, Vivi Feromida<sup>1</sup>

<sup>1</sup>Jurusan Matematika UIN Alauddin Makassar

E-mail: wahidah.alwi@uin-alauddin.ac.id

**Abstrak:** Penelitian ini membahas tentang pengelompokan kabupaten/kota di Sulawesi Selatan berdasarkan indikator kesehatan dengan menggunakan metode *Average Linkage*. Kondisi kesehatan di Provinsi Sulawesi Selatan saat ini, masih terbatasnya jangkauan dan akses pelayanan kesehatan. Tujuan dari penelitian ini adalah untuk mengetahui daerah yang memiliki kemiripan karakteristik berdasarkan indikator kesehatan di Sulawesi Selatan. Terdapat beberapa metode yang dapat digunakan dalam pengelompokan, salah satunya analisis *cluster Average Linkage*. Dimana analisis *cluster* adalah salah satu teknik multivariat yang tujuan utamanya mengelompokkan objek berdasarkan karakteristik yang mereka miliki. Dari hasil penelitian, diperoleh bahwa daerah yang memiliki kemiripan karakteristik dari segi kesehatan di Provinsi Sulawesi Selatan berdasarkan variabel yang digunakan dibentuk 4 *cluster*. Kabupaten/Kota pada *cluster* 4 yaitu Kota Palopo, *cluster* 3 terdiri dari 2 Kabupaten/Kota yaitu Kabupaten Tana Toraja dan Toraja Utara, dan *cluster* 2 hanya terdiri dari 1 Kabupaten/Kota yaitu Kabupaten Bulukumba. Adapun *cluster* 1 terdiri dari 20 Kabupaten/Kota yaitu semua Kabupaten/Kota kecuali Kabupaten Bulukumba, Tana Toraja, Toraja Utara dan Kota Palopo.

**Kata Kunci:** Analisis Cluster , Average Linkage, Indikator Kesehatan

## PENDAHULUAN

Kondisi kesehatan di Provinsi Sulawesi Selatan saat ini, masih terbatasnya jangkauan dan akses pelayanan kesehatan terutama pada daerah-daerah terpencil, perbatasan dan kepulauan karena belum didukung oleh sarana transportasi yang memadai. Selain itu, belum meratanya penyebaran beberapa tenaga kesehatan serta adanya tugas rangkap bagi petugas di tingkat Puskesmas menyebabkan tidak optimalnya pelayanan kesehatan. Kondisi umum pembangunan kesehatan di Provinsi Sulawesi Selatan dipengaruhi oleh berbagai faktor lingkungan, perilaku dan pelayanan kesehatan.

Pada Susenas 2014, penduduk Sulawesi Selatan yang mengalami keluhan kesehatan sekitar 24,78% dan 50,60% yang mengakibatkan terganggunya kegiatan sehari-hari. Gangguan kesehatan dapat mengakibatkan terganggunya aktivitas sehari-hari, baik dalam melakukan pekerjaan, bersekolah, mengurus rumah tangga maupun aktivitas lainnya. Gangguan kesehatan yang biasanya terjadi pada masyarakat yaitu

panas, batuk, pilek, diare, sakit kepala, sakit gigi, dll. Semakin banyak penduduk yang mengalami gangguan kesehatan, maka tingkat kesehatan di daerah tersebut rendah.

Dalam rangka memudahkan proses penanganan masalah kesehatan, maka perlu dilakukan pengelompokan berdasarkan indikator kesehatan tiap daerah agar dapat diberi bantuan dan ditingkatkan pelayanan kesehatannya. Permasalahan seperti gangguan kesehatan, pemberian imunisasi dan indikator lain dapat diatasi dengan tepat.

## TINJAUAN PUSTAKA

Analisis *cluster* adalah salah satu teknik multivariat yang tujuan utamanya adalah mengelompokkan objek berdasarkan karakteristik yang mereka miliki (Hair dkk, 2010). Pengelompokan dilakukan atas dasar kesamaan atau jarak. Analisis dilakukan dengan menempatkan pengamatan secara simultan untuk kemudian dihitung sebagai a profiling referensi karakteristik masing-masing *cluster* untuk menjelaskan perbedaan yang terjadi pada setiap indikator (Rahmat, 2017).

Berbeda dengan teknik multivariat lainnya, analisis ini tidak mengestimasi set variabel secara empiris sebaliknya menggunakan set variabel yang ditentukan oleh peneliti itu sendiri. Fokus dari analisis *cluster* adalah membandingkan objek berdasarkan set variabel, hal inilah yang menyebabkan para ahli mendefinisikan set variabel sebagai tahap kritis dalam analisis *cluster*. Set variabel *cluster* adalah suatu set variabel yang mempresentasikan karakteristik yang dipakai objek-objek.

Proses dari analisis *cluster* merupakan pengelompokan objek yang dapat dilakukan dengan menggunakan dua macam metode diantaranya metode non hirarki dan hirarki. Adapun metode non hirarki adalah metode pengelompokan yang telah ditentukan jumlah kelompok terlebih dahulu. Sedangkan metode hirarki jumlah kelompok belum diketahui dan metode ini merupakan metode pengelompokan yang terstruktur dan bertahap.

Metode hirarki merupakan metode pengelompokan yang hasilnya disajikan secara bertingkat atau berjenjang dari  $n$ ,  $(n-1)$ , ... , 1 kelompok. Tipe dasar dalam metode ini adalah aglomerasi (pemusatan) dan pemecahan. Dalam metode aglomerasi tiap observasi pada mulanya dianggap sebagai *cluster* tersendiri sehingga terdapat *cluster* sebanyak jumlah observasi. Kemudian dua *cluster* yang terdekat kesamaannya digabung menjadi suatu *cluster* baru, sehingga jumlah *cluster* berkurang satu pada tiap tahap. Sebaliknya pada metode pemecahan dimulai dari satu *cluster* yang besar yang mengandung seluruh observasi, selanjutnya observasi-observasi yang paling tidak sama dipisah dan dibentuk *cluster-cluster* yang lebih kecil. Proses ini dilakukan hingga tiap observasi menjadi *cluster* sendiri-sendiri (Ermawati, 2012).

Beberapa metode *agglomerative* pada analisis *cluster*, diantaranya sebagai berikut:

### 1) *Single linkage Method* (Pautan Tunggal)

Metode ini didasarkan pada jarak minimum. Dimulai dengan dua objek yang dipisahkan dengan jarak paling pendek maka keduanya akan ditempatkan pada *cluster* pertama, dan seterusnya. Metode ini dikenal pula dengan nama pendekatan tetangga terdekat.

Untuk menghitung jarak cluster (UV) dengan cluster-cluster yang lain dapat dirumuskan dengan:

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

Dimana:

Nilai  $d_{UW}$  dan  $d_{VW}$  menggambarkan jarak terdekat antara *cluster* U dengan W serta V dengan W.

2) *Complete linkage Method* (Pautan Lengkap)

Metode ini disebut juga pendekatan tetangga terjauh. Dasarnya adalah jarak maksimum. Dalam metode ini seluruh objek dalam suatu *cluster* dikaitkan satu sama lain pada suatu jarak maksimum atau kesamaan minimum.

Untuk menghitung jarak *cluster* (UV) dengan *cluster-cluster* yang lain dapat dirumuskan dengan:

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

Catatan:

Nilai  $d_{UW}$  dan  $d_{VW}$  menggambarkan jarak terdekat antara *cluster* U dengan W serta V dengan W (Johnson dan Wichern, 2007).

3) *Average linkage Method* (Pautan Rata-Rata)

Metode ini menghitung jarak antara dua *cluster* yang disebut sebagai jarak rata-rata dimana jarak tersebut dihitung pada masing-masing *cluster*. Pada berbagai keadaan, metode ini dianggap lebih stabil dibandingkan dengan kedua metode sebelumnya.

Adapun untuk menghitung jarak *cluster* (UV) dengan *cluster-cluster* yang lain dapat dirumuskan dengan (Sumertajaya dan Matjik):

$$d_{(UV)W} = \text{rata - rata } (d_{UW}, d_{VW})$$

Atau

$$d_{(UV)W} = \frac{\sum_i \sum_k d_{ik}}{N_{(UV)} N_w}$$

Catatan:

-Nilai  $d_{UW}$  dan  $d_{VW}$  menggambarkan jarak terdekat antara *cluster* U dengan W serta V dengan W.

- $d_{ik}$  merupakan jarak antara obyek i dalam *cluster* (UV) dan obyek k dalam *cluster* W.

4) *Ward's Method*

Dalam metode ini jarak antara dua *cluster* adalah jumlah kuadrat antara dua *cluster* untuk seluruh variabel. Metode ini cenderung digunakan untuk mengkombinasikan *cluster-cluster* dengan jumlah kecil.

5) *Centroid Method*

Jarak antara dua *cluster* adalah jarak antar centroid *cluster* tersebut. Centroid *cluster* adalah nilai tengah observasi pada variabel dalam suatu set variabel *cluster* (Ermawati, 2012).

### Ukuran Kemiripan Objek

Pengelompokan data dari struktur yang kompleks ke struktur yang lebih sederhana didasarkan pada “kedekatan” atau kesamaan”.

Adapun jarak yang biasa digunakan dalam analisis pengklateran diantaranya adalah (Yulianto dan Hidayatullah, 2014):

#### 1) Jarak Euclidian

Jarak Euclidian adalah jarak yang paling umum dan paling sering digunakan dalam analisis cluster. Jarak Euclidian antara dua titik dapat terdefiniskan dengan jelas. Jarak digunakan adalah peubah kontinu.

Jarak Euclidian antara *cluster* ke-*i* dan ke-*j* dari *p* peubah didefinisikan:

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2}$$

Atau

$$d(i, j) = \left[ \sum_{j=1}^m \sum_{i=1}^n (x_j - y_i)^2 \right]^{1/2}$$

Dimana:

$d(i, j)$  = jarak antara objek *i* ke objek *j*

$x_j$  = nilai tengah pada cluster ke-*j*

$y_i$  = nilai tengah pada cluster ke-*i*

$p$  = banyaknya peubah yang diamati

Pada jarak *Euclidean* mempunyai asumsi bahwa antar peubah tidak saling berkorelasi, memiliki satuan pengukuran yang sama dan pengukuran pembakuan mempunyai rata-rata nol dan standar deviasi satu.

#### 2) Jarak *Manhattan (City-Block)*

Jarak ini memudahkan jarak rata-rata dimensi-dimensi secara menyilang. Dalam banyak kasus, ukuran jarak ini mendapatkan hasil yang mirip terhadap *Euclidean* sederhana.

Jarak *City-Block* ditulis sebagai berikut:

$$d(i, j) = \sum_i |x_i - y_i|$$

Dimana:

$d(i, j)$  = jarak antara objek *i* ke objek *j*

#### 3) *The Chebyshev Distance*

*The Chebyshev Distance* antar dua obyek ialah nilai perbedaan mutlak yang maksimum pada tiap variabel. Ukuran jarak ini cocok untuk menentukan 2 objek, sebagai berbeda jika mereka berbeda terhadap segala sesuatu pada dimensi. Jarak *Chebyshev* dihitung sebagai berikut (Sumertajaya dan Matjik):

$$d(i, j) = \text{maksimum } |x_i - y_i|$$

### ***Menentukan Jumlah Cluster***

Nilai Eigen merupakan nilai karakteristik dari matriks bujursangkar. Untuk mencari nilai eigen, menggunakan beberapa perhitungan matriks dasar seperti penjumlahan matriks dan perkalian matriks.

Definisi 2.1

Untuk skalar sebarang  $\lambda$ , skalar  $\lambda$  disebut nilai eigen (*eigenvalue*). Nilai eigen dapat diperoleh dengan menentukan persamaan karakteristik terlebih dahulu, yaitu  $\det(\lambda I - A) = 0$ .

Nilai eigen dapat menjelaskan besarnya kontribusi keragaman masing-masing komponen utama dalam menjelaskan keragaman data asal. Besarnya kontribusi keragaman relatif yang mampu dijelaskan oleh komponen utama ke- $i$  adalah sebesar (Johnson dan Wichern, 2007):

$$\text{Total Varian} = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \dots + \lambda_p}$$

Atau dapat juga dinyatakan dalam besaran persen sebagai berikut:

$$\text{Total Varian} = \frac{\lambda_i}{\sum_{i=1}^p \lambda_i} \times 100\%$$

Sedangkan besarnya keragaman kumulatif untuk  $q$  buah komponen utama dapat

$$\text{dituliskan sebagai berikut : Total Kumulatif Varian} = \frac{\sum_{j=1}^q \lambda_j}{\sum_{i=1}^p \lambda_i} \times 100\%$$

Faktor yang dipertimbangkan bermakna apabila Keragaman kumulatif minimal 75%.

### ***Menginterpretasikan Cluster***

Tahap interpretasi meliputi pengujian tiap cluster untuk menamai dan menandai dengan suatu label yang secara akurat dapat menjelaskan kealamian *cluster*. Selain itu, tahapan ini dilakukan untuk menjelaskan karakteristik dari setiap kelompok berdasarkan dengan tujuan memberi label pada masing-masing kelompok tersebut. Proses ini dimulai dengan suatu ukuran yang sering digunakan yaitu *centroid cluster*.

*Centroid cluster* adalah rata-rata nilai objek yang terdapat dalam *cluster* pada tiap variabel dengan persamaan berikut:

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

Dimana:

$C$  : Nilai *centroid* tiap *cluster*

$x_{jk}$  : Nilai dari objek  $j$  pada variabel  $k$  dalam *cluster* yang dicari nilai *centroid*

$N$  : Banyaknya objek dalam *cluster* yang dicari nilai *centroid*

Membuat profil dan interpretasi *cluster* tidak hanya untuk memperoleh suatu gambaran saja melainkan pertama, menyediakan suatu rata-rata untuk menilai suatu korespondensi pada *cluster* yang terbentuk. Kedua, profil *cluster* tersebut memberikan arah bagi penilaian terhadap signifikansi praktis.

## METODE PENELITIAN

Data dari Dinas Kesehatan Provinsi Sulawesi Selatan tentang indikator kesehatan. Variabel yang digunakan yaitu, akses air minum layak, akses sanitasi layak, cakupan persalinan oleh tenaga kesehatan, cakupan pemeriksaan neonatal-1, cakupan imunisasi dasar lengkap, cakupan penimbangan balita, ratio dokter/puskesmas dan ratio bidan/puskesmas.

## PROSEDUR PENELITIAN

Metode yang digunakan adalah metode hirarki *Average Linkage* dan ukuran jarak *euclidean*.

Adapun prosedur yang akan dilakukan dalam penelitian ini adalah sebagai berikut:

1. Merumuskan masalah
2. Menentukan ukuran kemiripan objek
3. Menemukan jarak terdekat antar objek
4. Menggabungkan objek-objek yang sesuai
5. Menentukan jumlah *cluster*
6. Menginterpretasi *cluster*

## HASIL DAN PEMBAHASAN

### Profile Data

Nilai statistika deskriptif dari masing-masing variabel.

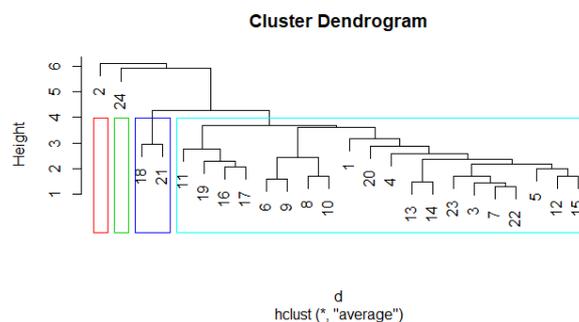
Tabel 1. Statistik Deskriptif

Variabel	Min	Mean	Max	Std.
X <sub>1</sub>	7,81	73,27	100	20,5102
X <sub>2</sub>	27,44	67,55	92,27	17,3083
X <sub>3</sub>	87,30	93,87	99,30	3,50742
X <sub>4</sub>	3,98	95,17	100	19,5073
X <sub>5</sub>	87,66	99,56	112,29	5,62688
X <sub>6</sub>	56,09	76,29	88,34	8,39788
X <sub>7</sub>	4,14	8,493	19,53	3,0319
X <sub>8</sub>	13,86	43,31	103,93	21,2449

Dari Tabel 4.1 tersebut di atas, tampak bahwa nilai maksimum terbesar indikator kesehatan berada pada variabel cakupan imunisasi dasar lengkap yaitu sebesar 112,29 dan nilai minimum terbesar berada pada variabel cakupan imunisasi dasar lengkap juga yaitu sebanyak 87,66 serta standar deviasi tertinggi sebesar 21,2449 pada variabel ratio bidan/puskesmas.

### ANALISIS CLUSTER

Pengelompokan objek tersebut dapat ditunjukkan dengan jelas pada gambar dendrogram pengelompokan kabupaten/kota berdasarkan indikator kesehatan di Provinsi Sulawesi Selatan. Berdasarkan hasil proses analisis *cluster* yang menggunakan dendrogram dengan metode *average linkage*, maka diperoleh 4 kelompok kabupaten/kota seperti terlihat pada gambar berikut.



Gambar 4.1 Dendrogram anggota *cluster*

Dari hasil penelitian, diperoleh bahwa daerah yang memiliki kemiripan karakteristik berdasarkan indikator kesehatan di Provinsi Sulawesi Selatan berdasarkan variabel yang digunakan dibentuk 4 *cluster*. Kabupaten/Kota pada *cluster* 4 yaitu Kota Palopo, *cluster* 3 terdiri dari 2 Kabupaten/Kota yaitu Kabupaten Tana Toraja dan Toraja Utara, dan *cluster* 2 hanya terdiri dari 1 Kabupaten/Kota yaitu Kabupaten Bulukumba. Adapun *cluster* 1 terdiri dari 20 Kabupaten/Kota yaitu semua Kabupaten/Kota kecuali Kabupaten Bulukumba, Tana Toraja, Toraja Utara dan Kota Palopo.

### INTERPRETASI

Interpretasi masing-masing *cluster* sebagai berikut.

1. *Cluster* 1 mempunyai nilai rata-rata yang tinggi untuk variabel Cakupan Imunisasi Dasar Lengkap ( $X_5$ ). Adapun variabel Ratio Dokter/puskesmas paling rendah dibandingkan dengan *cluster* lain. Sedangkan untuk semua variabel memiliki nilai rata-rata sedang kecuali variabel Ratio Dokter/puskesmas ( $X_7$ ).
2. *Cluster* 2 mempunyai nilai rata-rata yang lebih tinggi untuk variabel Akses Sanitasi Layak ( $X_2$ ). Adapun variabel Cakupan Persalinan oleh Tenaga Kesehatan ( $X_3$ ), Cakupan Pemeriksaan Neonatal-1 ( $X_4$ ), Cakupan Imunisasi Dasar Lengkap ( $X_5$ ), dan variabel Ratio bidan/puskesmas ( $X_8$ ) memiliki nilai paling rendah daripada *cluster* lainnya.
3. *Cluster* 3 mempunyai nilai rata-rata paling tinggi untuk variabel Akses Air Minum Layak ( $X_1$ ), Cakupan Persalinan oleh Tenaga Kesehatan ( $X_3$ ), dan variabel Ratio bidan/puskesmas ( $X_8$ ) dibandingkan *cluster* yang lain. Adapun variabel Akses Sanitasi Layak ( $X_2$ ) dan cakupan Penimbangan Balita ( $X_6$ ) memiliki nilai paling rendah daripada *cluster* yang lain.

4. *Cluster* 4 mempunyai nilai rata-rata yang paling tinggi untuk Cakupan Pemeriksaan Neonatal-1 ( $X_4$ ), Cakupan Imunisasi Dasar Lengkap ( $X_5$ ), dan Cakupan Penimbangan Balita ( $X_6$ ), dan Ratio dokter/puskesmas dibandingkan *cluster* lain.

## KESIMPULAN

Berdasarkan hasil analisis yang dilakukan, maka didapatkan bahwa daerah yang memiliki kemiripan karakteristik berdasarkan indikator kesehatan di Sulawesi Selatan yaitu:

1. *Cluster* 1 terdiri dari 20 Kabupaten/Kota yaitu semua Kabupaten/Kota kecuali Kabupaten Bulukumba, Tana Toraja, Toraja Utara dan Kota Palopo.
2. *Cluster* 2 terdiri dari 1 Kabupaten/Kota yaitu Kabupaten Bulukumba.
3. *Cluster* 3 terdiri dari 2 Kabupaten/Kota yaitu Tana Toraja dan Toraja Utara.
4. *Cluster* 4 terdiri dari 1 Kabupaten/Kota yaitu Kota Palopo.

## DAFTAR PUSTAKA

- Ermawati. *Statistika Multivariat Terapan*, (Makassar: Alauddin University Press, 2012)
- Hair, Joseph F., dkk. *Multivariate Data Analysis A Global Perspective*. (New Jersey: Pearson Education, 2010)
- Johnson, Richard A. dan Wichern, Dean W. *Applied Multivariate Statistical Analysis Fifth Edition*, (Amerika: Prentice-Hall, 2007)
- Rahmat, Abdul. *Clustering in Education*. European Research Studies Journal, 2017. Vol.XX edisi 3A.
- Sumertajaya, I Made dan Matjik, Ahmad Ansori. *Sidik Peubah Ganda dengan Menggunakan SAS*
- Yulianto. Safa'at dan Hidayatullah, Kishera. *Analisis Cluster untuk Pengelompokan Kabupaten/Kota di Provinsi Jawa Tengah Berdasarkan Indikator Kesejahteraan Rakyat*, (Semarang: Akd. Statistika Muhammadiyah Semarang, 2014)