

## PENGELOMPOKAN DESA/KELURAHAN DI KOTA DENPASAR MENURUT INDIKATOR PENDIDIKAN

Ni Wayan Aris Aprilia A.P.<sup>§1</sup>, I Gusti Ayu Made Srinadi<sup>2</sup>, Kartika Sari<sup>3</sup>

<sup>1</sup>Jurusan Matematika, Fakultas MIPA - Universitas Udayana [Email: arisapriliasari27@gmail.com]

<sup>2</sup>Jurusan Matematika, Fakultas MIPA - Universitas Udayana [Email: srinadi@unud.ac.id]

<sup>3</sup>Jurusan Matematika, Fakultas MIPA - Universitas Udayana [Email: sari\_kaartika@yahoo.co.id]

<sup>§</sup>Corresponding Author

### ABSTRACT

*Cluster analysis is one of data analysis used to classify objects in clusters which has objects with the same characteristics, whereas the other cluster has different characteristics. One part of the method of analysis cluster is hierarchy method. In a hierarchical method there are methods of linkage in the form of incorporation. Generally, methods of linkage is divided into 5 methods: single linkage, complete linkage, average linkage, Ward and centroid. The purpose of this study was to determine the best method of linkage among the method of single linkage, complete linkage, average linkage, and Ward, using Euclidean and Pearson proximity distance. Base on the smallest value of CTM (Cluster Tightness Measure), the best method of linkage as a result of this research was average linkage in Pearson distance.*

*Keywords: Euclidean distance, Pearson distance, Linkage methods, CTM*

### 1. PENDAHULUAN

Metode pautan merupakan bagian dari metode hirarki dalam analisis gerombol (*cluster analysis*), yaitu suatu proses penggabungan. Analisis gerombol adalah salah satu teknik analisis data yang digunakan untuk mengelompokkan obyek-obyek, sehingga dalam satu gerombol tergabung obyek dengan karakteristik yang sama, dan memiliki karakteristik yang berbeda dengan gerombol lain (Hair, et al [1]). Metode-metode yang merupakan metode pautan meliputi metode *single linkage*, *complete linkage*, *average linkage*, *Ward* dan *centroid*.

Pembentukan matriks jarak baru dalam metode pautan dengan metode *single linkage* menggunakan kriteria jarak minimum, sering disebut pendekatan tetangga terdekat (*nearest-neighbor*). Jarak minimum antara (*ij*) dengan kelompok lain misalkan kelompok *w* dituliskan sebagai:

$D_{(ij)w} = \min\{d_{iw}, d_{jw}\}$  dengan  $d_{iw}$  dan  $d_{jw}$  secara berturut-turut adalah jarak dari gerombol

*i* ke gerombol *w* dan dari gerombol *j* ke gerombol *w*,  $D_{(ij)w}$  merupakan jarak terdekat antara gerombol *i* dan *w* serta gerombol *j* dan *w*. Misalkan dimiliki matriks jarak objek 1 sampai dengan objek 5 seperti matriks berikut.

$$D = d_{ij} = \begin{matrix} & \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 0 & & & & \\ 9 & 0 & & & \\ 3 & 7 & 0 & & \\ 6 & 5 & 9 & 0 & \\ 11 & 10 & 2 & 8 & 0 \end{bmatrix} \end{matrix}$$

Objek dengan jarak terdekat digabung menjadi satu yaitu objek 3 dan 5 menjadi (35), selanjutnya dilakukan pembentukan matriks jarak dengan metode pautan *single linkage* dengan perhitungan sebagai berikut:

$$d_{(35)1} = \min\{d_{31}, d_{51}\} = \min\{3, 11\} = 3$$

$$d_{(35)2} = \min\{d_{32}, d_{52}\} = \min\{7, 10\} = 7$$

$$d_{(35)4} = \min\{d_{34}, d_{54}\} = \min\{9, 8\} = 8$$

Matriks jarak baru yang terbentuk adalah:

$$d_{ij} = \begin{matrix} & \begin{matrix} (35) \\ 1 \\ 2 \\ 4 \end{matrix} \\ \begin{matrix} (35) \\ 1 \\ 2 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 3 & 7 & 8 \\ 3 & 0 & 9 & 6 \\ 7 & 9 & 0 & 5 \\ 8 & 6 & 5 & 0 \end{bmatrix} \end{matrix}$$

Jarak terdekat adalah objek (35) disimbulkan objek (3\*) dengan objek 1, selanjutnya objek (3\*) digabung dengan objek 1 membentuk (3\*1). Prosedur perhitungan matriks jarak baru identik dengan yang dilakukan sebelumnya, dengan perhitungan:

$$d_{(3*1)2} = \min\{d_{3*2}, d_{12}\} = \min\{7,9\} = 7$$

$$d_{(3*1)4} = \min\{d_{3*4}, d_{14}\} = \min\{8,6\} = 6$$

Matriks jarak baru menjadi:

$$d_{ij} = \begin{matrix} & (3*1) & 2 & 4 \\ \begin{matrix} (3*1) \\ 2 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 7 & 6 \\ 7 & 0 & 5 \\ 6 & 5 & 0 \end{bmatrix} \end{matrix}$$

Prosedur ini akan dilakukan sampai semua pasangan objek dengan jarak minimum diperoleh dan bergabung menjadi satu gerombol. Hasil perhitungan digambarkan dalam bentuk dendogram.

Pada metode *complete linkage*, pembentukan matriks jarak baru berdasarkan jarak maksimum antara dua objek atau gerombol. *Average linkage* menggunakan jarak rata-rata antara dua gerombol. Metode *Ward* ditentukan dari jumlah kuadrat jarak dari dua gerombol berdasarkan semua variabel. Jarak dalam metode *centroid* merupakan jarak antara dua gerombol yang didefinisikan sebagai jarak *Euclidean* antara vektor mean (Rencher [2]).

Pendidikan merupakan salah satu unsur penting dalam pembangunan. Untuk mengetahui karakteristik tingkat pendidikan desa/kelurahan di Kota Denpasar dilakukan pengelompokan berdasarkan data indikator pendidikan. Indikator-indikator pendidikan dalam penelitian ini meliputi banyak sekolah TK, SD, SMP, SMA/SMK, banyak siswa TK, SD, SMP, SMA/SMK, dan banyak guru TK, SD, SMP, SMA/SMK. Penelitian ini bertujuan untuk mengetahui metode pautan terbaik di antara metode *single linkage*, *complete linkage*, *average linkage*, dan *Ward* dalam mengelompokkan desa/kelurahan di Kota Denpasar menurut indikator pendidikan.

Kriteria yang digunakan dalam penentuan metode terbaik adalah *Cluster Tightness Measure* (CTM). CTM adalah ukuran perkiraan efektivitas dalam kelompok, rancangan ukuran

antar kelompok, dan ukuran pemisahan dalam suatu kelompok. CTM diukur berdasarkan simpangan baku pada masing-masing kelompok seperti ditulis dalam persamaan berikut (Sutanto[3]):

$$CTM = \frac{1}{r} \sum_{t=1}^r \left( \frac{1}{p} \sum_{m=1}^p \frac{S_{tm}}{S_m} \right)$$

dengan

r : banyaknya kelompok

p : banyaknya variabel

S<sub>tm</sub> : simpangan baku pada kelompok ke-t untuk variabel ke-m

S<sub>m</sub> : simpangan baku variabel ke-m

Pemilihan metode pautan terbaik dilihat berdasarkan nilai CTM terkecil.

## 2. METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder yang berasal dari publikasi Badan Pusat Statistik (BPS) yaitu data Dinas Pendidikan Pemuda dan Olahraga, yaitu data indikator pendidikan pada 43 desa/kelurahan di Kota Denpasar tahun 2013. Indikator pendidikan yang menjadi variabel penelitian sebanyak 12 variabel yaitu banyak sekolah TK, SD, SMP, SMA/SMK, banyak siswa TK,SD SMP,SMA/SMK, dan banyak guru TK, SD, SMP, SMA/SMK.

Langkah-langkah analisis data sebagai berikut:

1. Melakukan standarisasi data.
2. Melakukan pncecekan korelasi.
3. Jika ada korelasi yang signifikan antara variabel dilakukan analisis komponen utama untuk membentuk komponen-komponen yang saling bebas, selanjutnya skor komponen yang diperoleh dipandang sebagai variabel untuk analisis lebih lanjut, jika tidak ada korelasi yang signifikan maka dilanjutkan ke langkah 4.
4. Mengukur kesamaan obyek (*similarity*) menggunakan jarak *Euclidean* [2] dengan rumus:

$$d_{ij}^E = \sqrt{\sum_{k=1}^p (x_{ki} - x_{kj})^2} \tag{1}$$

dan jarak Pearson yang menggunakan rumus (Sutanto [3]):

$$d_{ij}^p = \sqrt{\sum_{k=1}^p \frac{(x_{ki} - x_{kj})^2}{var(x_k)}} \quad (2)$$

- Menjelaskan setiap anggota kelompok yang terbentuk pada masing-masing metode pautan yaitu *single linkage* (Johnson & Wichern [4]) dengan rumus:

$$D_{(ij)w} = \min\{d_{iw}, d_{jw}\} \quad (3)$$

*complete linkage* [4] dengan rumus:

$$D_{(ij)w} = \max\{d_{iw}, d_{jw}\} \quad (4)$$

*average linkage* [4] dengan rumus:

$$D_{(ij)w} = \frac{\sum_n \sum_k d_{nk}}{N_{(ij)} N_w} \quad (5)$$

serta metode *ward* (Rencher [2]) menggunakan rumus:

$$SSE_{ij} = \sum_{k=1}^{n_{ij}} (\mathbf{y}_k - \bar{\mathbf{y}}_{ij})' (\mathbf{y}_k - \bar{\mathbf{y}}_{ij}) \quad (6)$$

- Melakukan pemilihan metode terbaik berdasarkan nilai CTM terkecil.
- Interpretasi hasil.

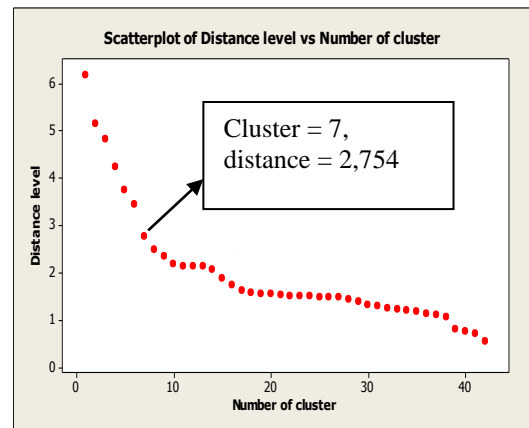
### 3. HASIL DAN PEMBAHASAN

#### 3.1 Deskripsi Data

Desa/kelurahan di Kota Denpasar ada sebanyak 43 dengan beberapa indikator pendidikan saling berkorelasi secara signifikan. Terlebih dahulu dilakukan analisis komponen utama untuk memperoleh skor komponen yang saling bebas. Skor komponen-skor komponen yang selanjutnya dianalisis dengan analisis gerombol untuk mengelompokkan desa/kelurahan di Kota Denpasar.

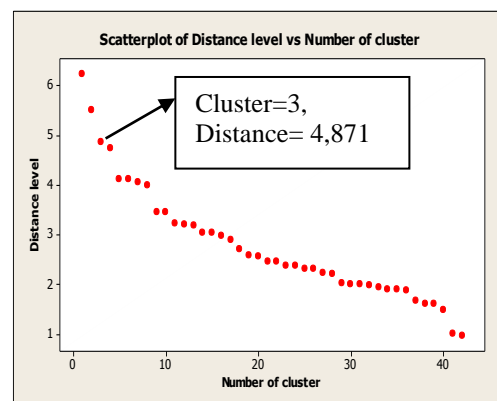
#### 3.2 Penentuan Gerombol dengan Metode *Single Linkage*

Pengelompokan dilakukan berdasarkan kedua jarak kedekatan (jarak Euclidean dan Pearson). Pada jarak *Euclidean*, plot diagram pencar memperlihatkan bahwa banyak gerombol yang terbentuk adalah 7, seperti pada Gambar 1.



Gambar 1. Diagram Pencar Metode *Single Linkage* pada Jarak *Euclidean*

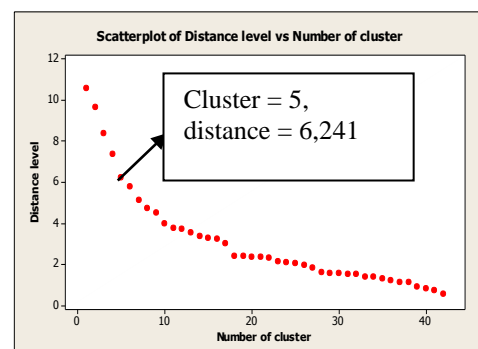
Pada jarak *Pearson*, banyak gerombol yang terbentuk adalah 3 seperti pada Gambar 2.



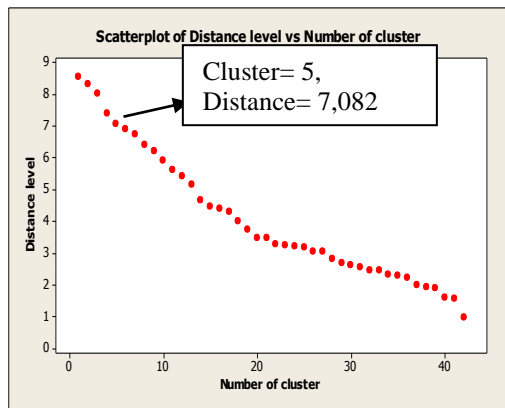
Gambar 2. Diagram Pencar Metode *Single Linkage* pada Jarak *Pearson*

#### 3.3 Penentuan Banyak Gerombol dengan Metode *Complete Linkage*

Proses penggerombolan dan pembentukan matriks jarak baru menggunakan persamaan (4). Sedangkan plot diagram pencar digunakan untuk menentukan banyaknya gerombol seperti pada Gambar 3 dan Gambar 4.



Gambar 3. Plot Diagram Pencar Metode *Complete Linkage* pada Jarak *Euclidean*

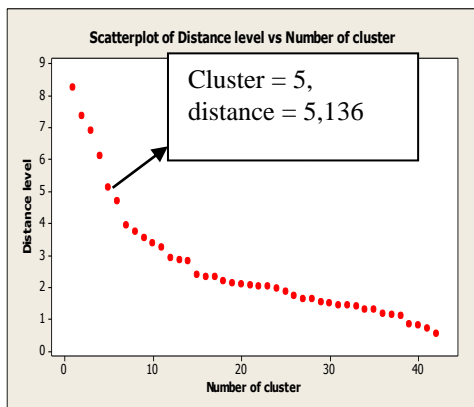


Gambar 4. Plot Diagram Pencar Metode *Complete Linkage* pada Jarak *Pearson*

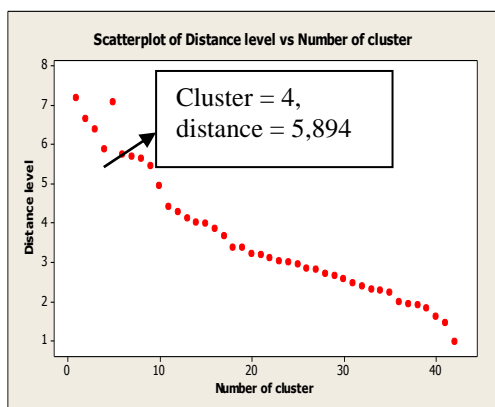
Terbentuk 5 gerombol dengan metode *Complete Linkage* pada jarak *Euclidean* dan jarak *Pearson*.

### 3.4 Penentuan Banyak Gerombol dengan Metode *Average Linkage*

Banyaknya gerombol yang terbentuk pada metode *average linkage* dilihat berdasarkan plot diagram pencar antara jarak *Euclidean* dan *Pearson*, dilihat pada Gambar 5 dan Gambar 6.



Gambar 5. Plot Diagram Metode *Average Linkage* pada Jarak *Euclidean*

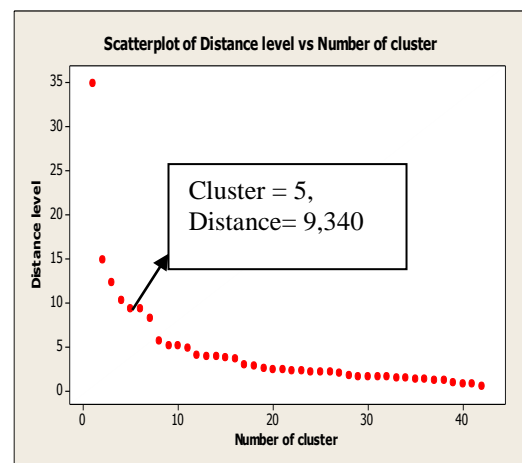


Gambar 6. Plot Diagram Pencar Metode *Average Linkage* pada Jarak *Pearson*

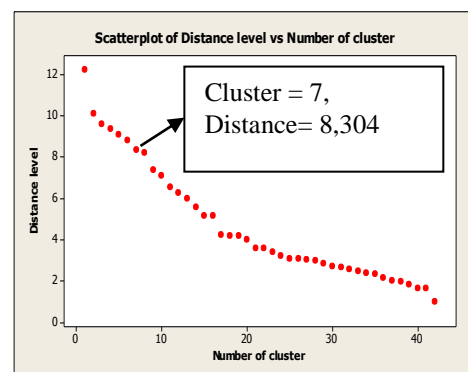
Terbentuk 5 gerombol pada jarak *Euclidean* dan 4 gerombol pada jarak *Pearson* dengan metode *Average Linkage*.

### 3.5 Penentuan Banyak Gerombol dengan Metode *Ward*

Pada metode *Ward*, dengan menggunakan plot diagram pencar untuk jarak *Euclidean* banyak gerombol yang terbentuk dapat dilihat pada Gambar 7. Gambar 8 menunjukkan banyak gerombol yang terbentuk pada jarak *Pearson*.



Gambar 7. Plot Diagram Pencar Metode *Ward* pada Jarak *Euclidean*



Gambar 8. Plot Diagram Pencar Metode *Ward* pada Jarak *Pearson*

Dengan metode *Ward*, pada jarak *Euclidean* terbentuk 5 gerombol dan 7 gerombol terbentuk pada jarak *Pearson*.

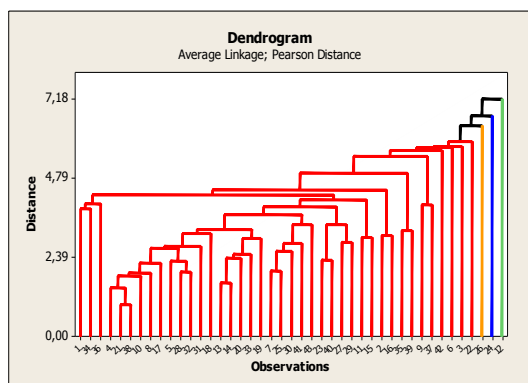
### 3.6 Penentuan Metode Terbaik

Pemilihan metode terbaik dengan menggunakan kriteria *Cluster Tightness Measure (CTM)* terkecil. Hasil perhitungan *CTM* pada masing-masing metode pautan dapat dilihat pada Tabel 1

Tabel 1. Metode Terbaik Menggunakan CTM

Metode	Nilai CTM dari Jarak	
	<i>Euclidean</i>	<i>Pearson</i>
<i>Single Linkage</i>	0,0909	0,0196
<i>Complete Linkage</i>	0,0409	0,0411
<i>Average Linkage</i>	0,0399	0,0183
<i>Ward</i>	0,0663	0,0535

Berdasarkan Tabel 1 dapat dilihat bahwa metode pautan terbaik dalam mengelompokkan desa/kelurahan di kota Denpasar berdasarkan indikator pendidikan metode *average linkage* pada jarak kedekatan *Pearson*. Dendrogram hasil pengelompokan metode *average linkage* pada jarak *Pearson* dapat dilihat pada Gambar 9.



Gambar 9. Dendrogram Hasil Pengelompokan *Average Linkage* pada Jarak *Pearson*

Dengan jarak kedekatan *Pearson*, metode pautan *Average Linkage*, 43 desa/kelurahan di kota Denpasar dikelompokkan dalam 4 kelompok. Satu kelompok beranggotakan 40 desa/kelurahan dan 3 desa/kelurahan lainnya masing-masing merupakan kelompok tersendiri, diuraikan pada Tabel 2.

Hasil pengelompokan ini menunjukkan bahwa Desa Sumerta Klod, Desa Dangin Puri Kauh, dan Desa Dangin Puri Kangen memiliki indikator pendidikan dengan karakteristik yang berbeda dari ke-40 desa/kelurahan lainnya di kota Denpasar yang tergabung dalam kelompok 1. Pada kelompok 1, variabel yang dominan menentukan karakteristik pendidikan pada kelompok ini adalah variabel  $X_1$  (banyak sekolah TK).

Tabel 2. Pengelompokan Desa/Kelurahan dengan Metode *Average Linkage* pada Jarak *Pearson*

Kel	Anggota Kelompok	
I	1. Desa Pemogan	25. Kelurahan Sumerta
	2. Kel. Pedungan	26. Desa Sumerta Kaja
	3. Kel.Sesetan	27. Desa Sumerta Kauh
	4. Desa Serangan	28. Kelurahan Dangin Puri
	5. Desa Sidakarya	29. Kelurahan Penatih
	6. Kelurahan Panjer	30. Desa Penatih Dangin Puri
	7. Kelurahan Renon	31. Desa Pemecutan Kaja
	8. Desa Sanur Kauh	32. Desa Dauh Puri Kauh
	9. Kelurahan Sanur	33. Desa Dauh Puri Klod
	10. Desa Sanur Kaja	34. Kelurahan Dauh Puri
	11. Desa.Dangin Puri Klod	35. Desa Dauh Puri Kangan
	12. Desa Dauh Puri Kaja	36. Kelurahan Pemecutan
	13. Desa Dangin Puri Kaja	37. Desa Tegal Arum
	14. Kel. Tonja	38. Desa Tegal Kerta
	15. Kel. Peguyangan	39. Kelurahan Padangsambian
	16. Kelurahan Ubung	40. Desa Padangsambian
	17. Desa Ubung Kaja	
	18. Desa Peguyangan Kaja	
	19. Desa Peguyangan Kangan	
	20. Desa Padangsambian Klod	
	21. Desa Pemecutan Klod	
	22. Kelurahan Kesiman	
	23. Desa Kesiman Petilan	
	24. Desa Kesiman Kertalangu	
II	Desa Sumerta Klod	
III	Desa Dangin Puri Kauh	
IV	Desa Dangin Puri Kangan	

#### 4. KESIMPULAN

Berdasarkan kriteria nilai CTM terkecil diperoleh bahwa metode pautan terbaik yang digunakan dalam pengelompokan desa/kelurahan di kota Denpasar berdasarkan indikator pendidikan adalah metode *average linkage*.

**DAFTAR PUSTAKA**

- [1] Hair, Joseph.F. et.al. 1995. *Multivariate Data Analysis with Readings Fourth Edition*. United State of America: Prentice-Hall International, Inc.
- [2] Rencher, A.C. 2002. *Methods of Multivariate Analysis Second Edition*. New York: John Wiley and Son, Inc.
- [3] Sutanto, H.T. 2009. *Cluster Analysis*. <http://core.ac.uk/download/pdf/11064649.pdf>. Diakses 14 April 2015
- [4] Johnson, R.A. & Dean.W.Wichern. 2007. *Applied Multivariate Statistical Analysis Sixth Edition*. United State of America: Pearson Education, Inc.