

APLIKASI ANALISIS CLUSTER DALAM MENGELOMPOKKAN KECAMATAN BERDASARKAN FAKTOR PENYEBAB PENDUDUK BUTA AKSARA DI KABUPATEN JENEPONTO

Irwan

Dosen Jurusan Matematika, Fak. Sains dan Teknologi.

UIN Alauddin Makassar

e-mail:iwan.uin@gmail.com

Abstract: Tulisan ini membuat tentang pengelompokan kecamatan berdasarkan faktor penyebab penduduk buta aksara di kabupaten Jeneponto dengan aplikasi analisis cluster. Jenis faktor yang mempengaruhi penduduk buta aksara yakni faktor ekonomi, budaya dan geografis.

Pengumpulan data dilakukan dengan cara survey melalui penggunaan kuesioner sebagai instrument penelitian. Sampel yang dilibatkan sebanyak 200 orang dengan teknik pengambilan sampelnya Proporsional Random Sampling. Data dianalisis dengan menggunakan teknik Hierarchical. Hasil penelitian menunjukkan bahwa faktor yang menyebabkan penduduk buta aksara yang paling banyak di Kabupaten Jeneponto yaitu faktor ekonomi di antara sebelas kecamatan. Adapun kecamatan yang paling banyak penduduk buta aksaranya yang disebabkan dari faktor ekonomi yakni kecamatan Binamu.

Key words: *Analisis cluster, hierarchical, buta aksara.*

I. PENDAHULUAN

Tingkat buta aksara dapat dilihat berdasarkan beberapa faktor yang menjadi aspek penilaian diantara faktor ekonomi, budaya, dan geografis. Faktor tersebut dapat menjadi variabel dalam tingkat penduduk buta aksara. Dengan adanya faktor-faktor tersebut, maka sangat mempengaruhi kehidupan masyarakat khususnya dalam hal pendidikan. Maka, pemerintah perlu mengetahui daerah-daerah mana saja yang paling banyak penduduk yang buta aksara. Agar nantinya dapat dipertimbangkan selanjutnya, dengan jalan didirikannya sekolah-sekolah bagi mereka baik dari kalangan anak-anak remaja maupun dewasa, oleh karena itu, perlu melakukan pengelompokan desa berdasarkan faktor-faktor penyebab penduduk buta aksara.

Untuk melihat kesamaan atau kemiripan suatu daerah berdasarkan faktor penyebab penduduk buta aksara. Penulis menggunakan salah satu metode dalam analisis multivariat, yaitu Cluster Analysis (gerombol). Analisis gerombol merupakan suatu metode dalam analisis peubah ganda yang bertujuan untuk mengelompokkan n satuan pengamatan ke dalam k gerombol dengan ($k < n$) berdasarkan p peubah, sehingga unit-unit pengamatan dalam satu kelompok mempunyai ciri-ciri yang lebih homogen dibandingkan unit pengamatan dalam kelompok lain. Pada analisis ini terdapat dua metode diantaranya metode *hierarki* dan *nonhierarki*.

Analisis multivariat sering digunakan untuk memecahkan permasalahan dalam penelitian yang sifatnya sangat kompleks. Analisis ini dikelompokkan dalam dua bentuk yaitu:

1. Analisis Dependensi yakni metode ketergantungan digunakan untuk menjelaskan dan memprediksi satu atau lebih variabel dependen yang didasarkan pada variabel-variabel independen yang digunakan pada analisis dependensi terdiri dari analisis diskriminan, analisis regresi berganda, multivariate analysis of variance (manova), dan korelasi kanonikal.

2. Analisis interdependensi yakni metode saling ketergantungan digunakan untuk mengetahui struktur data sekelompok objek dimensi atau variabel. Metode statistik yang sering digunakan adalah analisis komponen utama (*principal component analysis*), analisis faktor (*factor analysis*), dan analisis kelompok (*cluster analysis*), dan analisis log-linear.

3. Analisis *Cluster* merupakan metode pengelompokan, di mana data yang akan dikelompokkan belum membentuk kelompok sehingga pengelompokkan yang akan dilakukan bertujuan agar data yang terdapat di dalam kelompok yang sama relatif lebih homogen daripada data yang berada pada kelompok yang berbeda. Diharapkan dengan terbentuknya kelompok tersebut akan lebih mudah menganalisa dan lebih tepat pengambilan keputusan sehubungan dengan masalah tersebut.

Solusi Cluster Analysis bersifat tidak unik, anggota cluster untuk tiap penyelesaian/solusi tergantung pada beberapa elemen prosedur dan beberapa solusi yang berbeda dapat diperoleh dengan mengubah satu elemen atau lebih. Solusi cluster secara keseluruhan bergantung pada variabel-variabel yang digunakan sebagai dasar untuk menilai kesamaan. Penambahan atau pengurangan variabel-variabel yang relevan dapat mempengaruhi substansi hasil cluster

analysis. Secara umum cluster analysis, bisa dikatakan sebagai proses menganalisa baik tidaknya suatu proses pembentukan *cluster*. Analisa cluster bisa diperoleh dari kepadatan cluster yang dibentuk (*cluster density*). Kepadatan suatu cluster bisa ditentukan dengan *variance within cluster* (V_w) dan *variance between cluster* (V_b).

Varian tiap tahap pembentukan cluster bisa dihitung dengan rumus

$$V_c^2 = \frac{1}{n_c - 1} \sum_{i=1}^n (y_i - \bar{y}_c)^2$$

Selanjutnya dari nilai varian di atas, kita bisa menghitung nilai *variance within cluster* (V_w) dengan rumus :

$$V_w = \frac{1}{N - c} \sum_{i=1}^c (n_i - 1) \cdot V_i^2$$

Nilai *variance between cluster* (V_b) dengan rumus :

$$V_b = \frac{1}{c - 1} \sum_{i=1}^c n_i (\bar{y}_i - \bar{y})^2$$

Salah satu metode yang digunakan untuk menentukan cluster yang ideal adalah batasan *variance*, yaitu dengan menghitung kepadatan *cluster* berupa *variance within cluster* (V_w) dan *variance between cluster* (V_b). Cluster yang ideal mempunyai V_w minimum yang merepresentasikan *internal homogenitas* dan maksimum V_b yang menyatakan *external homogenitas*.

$$V = \frac{V_w}{V_b}$$

Meskipun minimum V_w menunjukkan nilai cluster yang ideal, tetapi pada beberapa kasus kita tidak bisa menggunakannya secara langsung untuk mencapai *global optimum*. Jika kita paksakan, maka solusi yang dihasilkan akan jatuh pada *local optima*. Adapun ciri Cluster Analysis yang baik adalah cluster yang mempunyai:

1. Homogenitas (kesamaan) yang tinggi antar anggota dalam satu gerombol (*variance within-cluster*).
2. Heterogenitas (perbedaan) yang tinggi antar gerombol yang satu dengan gerombol lainnya (*variance between-cluster*).

Teknik yang digunakan dalam cluster ini yakni teknik hierarchical. Metode ini memulai pengclusteran data dengan dua atau lebih objek yang mempunyai kesamaan paling dekat kemudian dilanjutkan ke objek lain yang mempunyai kedekatan kedua dan seterusnya sampai *cluster* akan membentuk semacam pohon

sehingga ada tingkatan yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip sehingga pada akhirnya hanya akan terbentuk sebuah *cluster*.

Tipe dasar dalam metode ini adalah agglomerasi dan pemecahan. Dalam metode agglomerasi 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 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. Hal penting dalam metode *hierarchical* adalah bahwa hasil pada tahap sebelumnya selalu bersarang di dalam hasil pada tahap berikutnya, membentuk sebuah pohon. Atau dengan metode ini, data tidak langsung dikelompokkan ke dalam beberapa cluster dalam 1 tahap, tetapi dimulai dari 1 cluster yang mempunyai kesamaan, dan berjalan seterusnya selama beberapa iterasi, hingga terbentuk beberapa cluster tertentu. Arah *hierarchical clustering* dibagi 2, yaitu *divisive* dan *agglomerative*

Penghitungan jarak antar obyek, maupun antar *clusternya* dilakukan dengan *Euclidian distance*, khususnya untuk data numerik. Untuk data 2 dimensi, digunakan persamaan sebagai berikut :

$$d(x,y) = \sqrt{\sum(x_i - y_i)^2}$$

Algoritma hierarchical clustering banyak diaplikasikan pada metode peng-clusteran berikut :

a. Pautan Tunggal (*Single Linkage Hierarchical Method (SLHM)*)

Bila ada dua cluster hendak digabungkan atau tidak perlu diperhatikan cukup satu anggota dari masing - masing cluster yang keduanya berhubungan atau berjarak paling dekat. Kalau sebuah cluster digabungkan hanya karena paling dekat dengan salah satu anggota maka metode ini bisa menyebabkan interpretasi cluster yang tidak jelas. Dua cluster digabungkan pada setiap tahap yang berdasarkan jarak paling dekat di antara keduanya sehingga proses ini berlanjut terus sampai diperoleh satu cluster sehingga dirumuskan dengan $D=\{d_{ik}\}$. Kemudian menggabungkan objek-objek yang sesuai, misalkan objek tersebut dilambangkan dengan gerombol U dan gerombol V untuk mendapatkan gerombol

(UV). Untuk menghitung jarak gerombol (LJV) dengan gerombol – gerombol yang lain dapat dirumuskan dengan:

$$d_{(uv)w} = \min \{d_{uw}, d_{vw}\}$$

besaran dan menggambarkan jarak terdekat antara gerombol U dan W serta V dan W. Hasil dari pautan tunggal ini dapat ditampilkan dalam bentuk dendrogram atau diagram pohon, dimana dahan atau cabang dan diagram pohon tersebut merupakan gerombolnya.

b. Pautan Lengkap (*Complete Linkage Hierarchical Method*)

Dalam metode pautan lengkap, ditentukan dan jarak terjauh antara dua objek pada gerombol yang berbeda (*Furthest Neighbor*). Metode ini dapat pula digunakan dengan baik untuk kasus dimana objek – objek yang ada berasal dan kelompok yang benar – benar berbeda. Pada metode ini tahap pertama yang harus dilakukan adalah menemukan jarak terdekat antara gerombol–gerombol tersebut, dirumuskan dengan $D=\{d_{ik}\}$. Kemudian menggabungkan objek-objek yang sesuai, misalkan objek tersebut dilambangkan dengan gerombol U dan gerombol V untuk mendapatkan gerombol gabungan (UV). Untuk menghitung jarak gerombol (UV), dengan gerombol-gerombol yang lain dapat dirumuskan dengan:

$$d_{(uv)w} = \max \{d_{uw}, d_{vw}\}$$

besaran d_{uw} dan d_{vw} menggabungkan jarak terjauh antara U dengan W serta V dengan W.

c. Pautan Rata-rata (*Average Linkage Hierarchical Method*)

Metode ini bertujuan meminimumkan rata-rata jarak semua pasangan pengamatan dan dua gerombol yang digabungkan serta cenderung membentuk gerombol dengan ragam kecil. Pada berbagai keadaan, metode ini dianggap lebih stabil dibandingkan dengan kedua metode sebelumnya. Pada metode ini tahap pertama yang harus dilakukan adalah sama seperti metode-metode sebelumnya yaitu menemukan jarak terkecil dari matriks jarak, kemudian menggabungkan objek-objek yang sesuai, misalkan objek tersebut dilambangkan dengan gerombol U dilambangkan dan gerombol V untuk mendapatkan gerombol gabungan (UV). Untuk menghitung jarak gerombol (UV) dengan gerombol-gerombol yang lain dapat dirumuskan dengan:

$$d_{(uv)w} = \text{rata-rata} \{d_{uw}, d_{vw}\}$$

besarnya d_{uw} dan d_{vw} menggambarkan jarak terjauh antara U dengan W serta V dengan W.

d. Metode Ward (*Ward's Method*)

Metode ini cenderung digunakan untuk mengkombinasi cluster-cluster dengan jumlah kecil. Dihubungkan pada centroid methods yang bahwa itu juga mengarah pada sebuah perwakilan geometric yang mana pada cluster centroids memerankan sebuah aturan yang penting. Untuk membentuk cluster-cluster, metode tersebut meminimalisir sebuah fungsi objektif yang dalam masalah ini, persamaan ukuran “squared error” seperti yang digunakan dalam *MANOVA*.

e. Centroid Linkage Hierarchical Method

Centroid Linkage adalah proses pengclusteran yang didasarkan pada jarak antar centroidnya. Metode ini bagus untuk memperkecil *variance within cluster* karena melibatkan centroid pada saat penggabungan antar cluster. Metode ini juga baik untuk data yang mengandung *outlier*.

Asumsi-asumsi dalam Cluster Analysis

Seperti hal teknik analisis lain, Cluster Analisis juga menetapkan adanya suatu asumsi. Ada dua asumsi dalam Cluster Analysis, yaitu:

a. Kecukupan Sampel untuk Merepresentasikan/Mewakili Populasi

Biasanya suatu penelitian dilakukan terhadap populasi diwakili oleh sekelompok sampel. Sampel yang digunakan dalam Cluster Analisis harus dapat mewakili populasi yang ingin dijelaskan, karena analisis ini baik jika sampel representatif. Jumlah sampel yang diambil tergantung penelitinya, seorang peneliti harus yakin bahwa sampel yang diambil representatif terhadap populasi.

b. Pengaruh *Multikolinieritas*

Ada atau tidaknya multikolinieritas antar variabel sangat diperhatikan dalam Cluster Analysis karena hal itu berpengaruh, sehingga variabel-variabel yang bersifat multikolinieritas secara eksplisit dipertimbangkan dengan lebih seksama. Alternatif lain adalah dengan mengkombinasikan kedua metode ini. Pertama, gunakan metode kemudian *hierarchical* dilanjutkan dengan metode *no hierarchical*.

II. METODE PENELITIAN

Langkah-langkah yang dilakukan untuk pengelompokan kecamatan berdasarkan faktor penyebab penduduk buta aksara di kabupaten Jeneponto adalah sebagai berikut:

- a. Menentukan desa berdasarkan faktor penyebab penduduk buta aksara. Dalam penentuan desa yang berdasarkan atas faktor penyebab penduduk buta aksara, maka langkah-langkah yang harus dilakukan, yaitu:

Langkah 1: Melakukan pendekatan cluster terhadap Kecamatan yang akan diteliti.

Langkah 2: Mengukur kesamaan (*Homogenitas*) antar cluster yang satu dengan cluster yang lain.

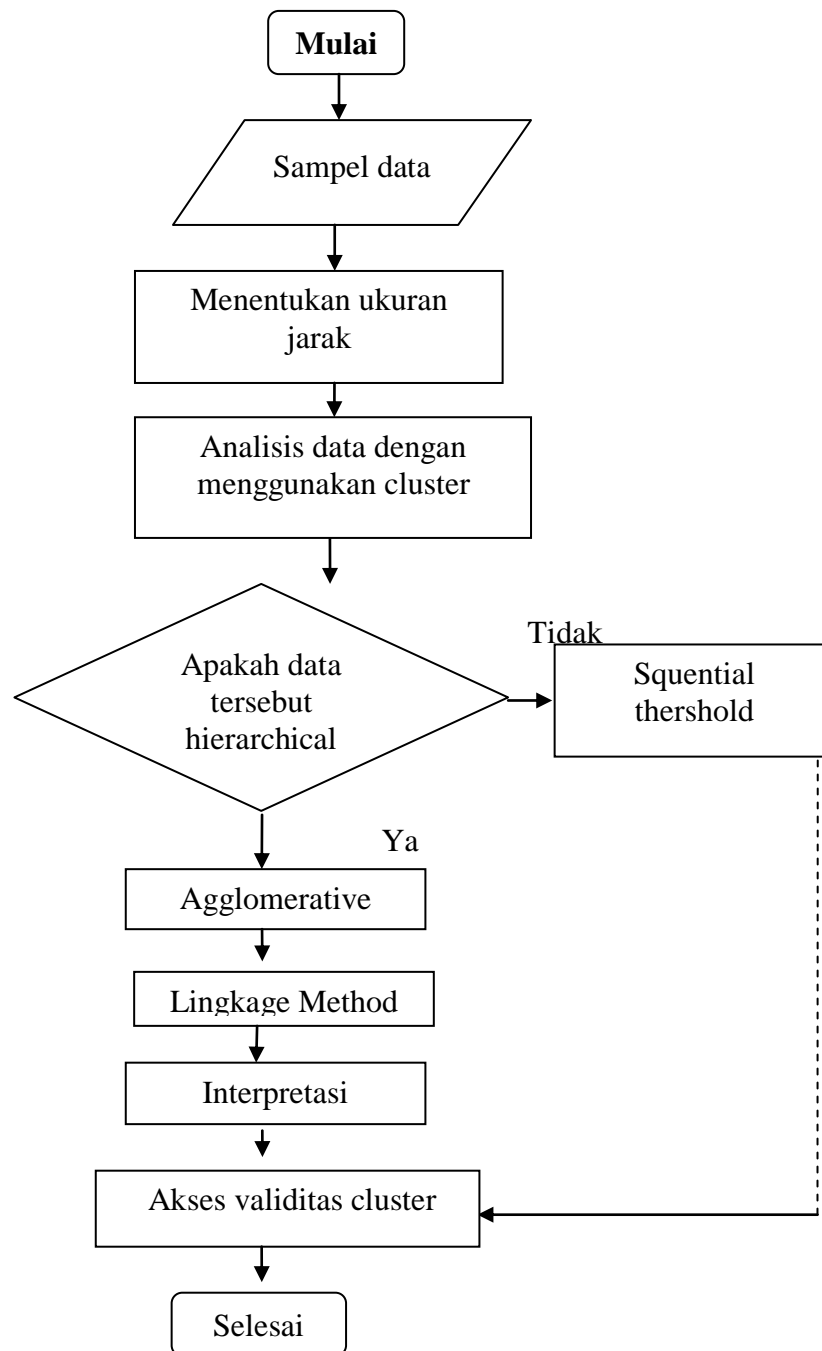
Langkah 3: Membentuk cluster, untuk membentuk cluster yang baik dicapai dengan tiga tujuan yaitu:

- 1) Deskripsi Klasifikasi Data
- 2) Penyederhanaan Data
- 3) Identifikasi Hubungan

- b. Menentukan faktor-faktor yang mempengaruhi penduduk buta aksara, maka yang harus dilakukan yaitu:

- 1) Membentuk variabel dari data yang terpilih
- 2) Bila variabel terbentuk dengan syarat sudah terpenuhi maka dianalisis dengan menggunakan metode yang terpilih.

Secara umum langkah-langkah dalam metodologi di atas, digambarkan dalam *flowchar* di bawah ini:



Gambar 1.

Flowcart analisis cluster pada pengelompokan kecamatan berdasarkan faktor penyebab penduduk buta aksara di kabupaten Jeneponto

III. PEMBAHASAN

Adapun hasil pengelompokan buta aksara yang diolah dengan menggunakan metode hierarki yaitu:

Tabel 1. Rekapitulasi Pengelohan Kasus (*Cases Processing Summary(a)*)

Case Processing Summary^a

Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
11	100.0	0	.0	11	100.0

a. Single Linkage (Between Groups)

Tabel 2. Daftar Agglomerasi (Agglomeration Schedule)

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	10	5.000	0	0	4
2	9	11	6.000	0	0	3
3	6	9	11.000	0	2	6
4	2	7	12.500	1	0	6
5	1	3	14.000	0	0	9
6	2	6	21.333	4	3	7
7	2	8	40.333	6	0	10
8	4	5	56.000	0	0	9
9	1	4	69.000	5	8	10
10	1	2	99.429	9	7	0

Tabel 3. Matriks kedekatan (proximity matrix)

Proximity Matrix

Case	Squared Euclidean Distance										
	1	2	3	4	5	6	7	8	9	10	11
1	.000	89.000	14.000	89.000	93.000	49.000	118.000	114.000	105.000	74.000	89.000
2	89.000	.000	35.000	20.000	100.000	22.000	17.000	85.000	26.000	5.000	46.000
3	14.000	35.000	.000	35.000	59.000	25.000	66.000	98.000	65.000	30.000	65.000
4	89.000	20.000	35.000	.000	56.000	54.000	53.000	137.000	82.000	25.000	102.000
5	93.000	100.000	59.000	56.000	.000	146.000	189.000	297.000	206.000	121.000	230.000
6	49.000	22.000	25.000	54.000	146.000	.000	17.000	29.000	12.000	9.000	10.000
7	118.000	17.000	66.000	53.000	189.000	17.000	.000	36.000	9.000	8.000	19.000
8	114.000	85.000	98.000	137.000	297.000	29.000	36.000	.000	29.000	52.000	11.000
9	105.000	26.000	65.000	82.000	206.000	12.000	9.000	29.000	.000	17.000	6.000
10	74.000	5.000	30.000	25.000	121.000	9.000	8.000	52.000	17.000	.000	27.000
11	89.000	46.000	65.000	102.000	230.000	10.000	19.000	11.000	6.000	27.000	.000

Tabel 3. di atas adalah hasil proses cluster (*gerombol*) metode pautan tunggal (*single group linkage*) setelah jarak antara variabel diukur dengan *jarak kuadrat eukledian (Squared Euclidean Distance)*, maka dilakukan pengelompokan variabel secara *hierarki*. Dimana cara ini merupakan pengelompokan dilakukan pelan-pelan jumlah gerombol berkurang sehingga akhirnya semua menjadi satu gerombol saja. Cara pembuatan gerombol yang dimulai dari satu atau lebih variabel yang paling mirip membentuk satu gerombol, kemudian gerombol memasukkan lagi satu variabel yang paling mirip, dinamakan *agglomerasi*, yaitu:

- a. Pada tahap atau stage 1 (dilihat baris 1) terbentuk satu gerombol dengan anggota variabel 2 (Desa Bangkala Barat) dengan variabel 10 (Desa Kelara), untuk kolom koefisien yang berisi angka 5.000 yang menyatakan jarak antara variabel dari desa Bangkala Barat dan desa Kelara seperti yang terlihat pada matrix promity sebelumnya. Karena proses matrix promity memperlihatkan jarak kedekatan antara anggota cluster sehingga bisa dilihat anggota yang mana yang paling kecil jumlah koefisiennya berarti menyatakan bahwa anggota tersebut anggota yang terdekat diantara anggota lain. Jadi desa Bangkala Barat dan desa Kelara memperlihatkan jarak antara yang terdekat diantara desa lain. Karena proses agglomerasi dimulai dengan dua variabel yang terdekat dari sekian banyak kombinasi jarak dari 11 variabel yang ada. Kemudian jika dilihat pada kolom terakhir untuk baris 1 tersebut (next stage) angka 4. Hal ini

berarti langkah gerombol selanjutnya dilakukan dengan melihat stage 4 atau baris ke 4.

- b. Pada baris ke-2 terlihat variabel 9 (Desa Arungkeke) membentuk gerombol dengan variabel 11 (desa Tarowang). Dengan demikian, sekarang gerombol dari 3 variabel yakni desa Arungkeke, desa Kelara, dan desa Tarowang. Angka pada kolom coefficient jarak antara rata-rata yang terjadi antara variabel terakhir yang bergabung dengan desa Bontoramba dengan dua variabel sebelumnya, yakni variabel 2 (desa Bangkala Barat) dan variabel 10 (desa Kelara).

Demikian selanjutnya dari stage 3 proses dilanjutkan ke stage 4, dari stage 4 proses dilanjutkan ke stage 5 dan selanjutnya sampai ke stage terakhir yakni stage 10.

Pada proses *agglomerasi* tentu bersifat kompleks khususnya perhitungan koefisien yang melibatkan sekian banyak variabel dan terus bertambah. Semakin kecil angka koefisien, semakin anggota gerombol tersebut mempunyai kemiripan satu dengan yang lain demikian pula sebaliknya makin besar koefisien, makin tidak mirip satu dengan yang lain. Proses *agglomerasi* pada akhirnya akan menyatukan semua variabel menjadi satu gerombol.

Berdasarkan hasil yang diperoleh tersebut, maka dapat dilakukan pekelompokan dengan melalui *Dendogram*. *Dendogram* merupakan *visualisasi* proses gerombol yang terjadi, seperti hasil output berikut ini;

Dendogram

* * H I E R A R C H I C A L C L U S T E R A N A L Y S I S * * *
 Dendrogram using Average Linkage (Between Groups)

		Rescaled Distance Cluster Combine					
C A S E		0	5	10	15	20	25
Label	Num	+-----+-----+-----+-----+-----					
B. Barat	2	-+-----+					
Kelara	10	--+ +-----+					
Batang	7	-----+ +-----+-----+					
Arungkek	9	-+-----+					
Tarowang	11	--+ +-----+ +-----+-----+-----					
Turatea	6	-----+					
Rumbia	8	-----+-----+					
Bangkala	1	-----+-----+-----+-----+					
Tamalate	3	-----+ +-----					
Bontoram	4	-----+-----+-----+-----+ +					
Binamu	5	-----+-----+-----+-----+ +					

Abbreviated	Extended
Name	Name
Arungkek	Arungkeke
Bontoram	Bontoramba
Tamalate	Tamalatea

Analisis Dendogram

Perhatikan bahwa skala yang digunakan bukanlah koefisien yang ada pada tabel *agglomeration schedule*, namun telah dilakukan proses skala ulang (*rescale*), dengan batasan 0 sampai 25. Pengskalaan ulang dilakukan dapat diketahui jarak antara kelompok cluster yang terdekat dapat diperoleh jarak yang paling dekat diantara cluster yang terbentuk.

Proses *agglomerasi* dimulai pada skala 0, dengan ketentuan jika sebuah garis dekat angka 0, maka variabel-variabel yang terwakali dengan garis tersebut semakin mungkin membentuk sebuah gerombol.

Pada hasil dendogram di atas, variabel nomor 2 dan 10 membentuk sebuah cluster tersendiri, karena mereka mempunyai panjang garis yang sama dan bergabung menjadi satu kesatuan, demikian pula variabel nomor 9 dan 11 membentuk cluster tersendiri.

Sebaliknya variabel 7, 6, 8, 1, 3, 4 dan 5 tidak bergabung dengan variabel lain. Karena mempunyai garis yang lebih panjang dari variabel-variabel yang disebut terdahulu. Dengan demikian, pada proses pertama telah terbentuk sembilan cluster, yakni:

1. 2 cluster yang mempunyai anggota lebih dari satu variabel yaitu, 2,10,9 dan 11 (Bangkala Barat, Kelara, Arungkeke, dan tarowang).
2. 7 cluster yang berdiri sendiri.

Kemudian proses selanjutnya yaitu terlihat variabel nomor 2 dan 10 yang sebelumnya telah bergabung, sekarang bergabung lagi dengan variabel 7 yang sebelumnya juga membentuk satu cluster tersendiri.

Demikian seterusnya, proses *dendogram* berjalan kearah kanan dengan menggunakan petunjuk panjang garis yang semakin ke kanan, hingga pada akhirnya, semua variabel akan bergabung menjadi satu cluster.

Dendogram berguna untuk menunjukkan anggota cluster yang ada jika akan ditentukan berapa cluster yang harusnya terbentuk. Sebagai contoh penulis menginginkan 4 cluster, maka dari *dendogram* terlihat cluster 1 beranggotakan

kecamatan Bontoramba (variabel 4), dan kecamatan Binamu (variabel 5), cluster 2 beraggotakan Kecamatan Bangkala (variabel 1) dan Kecamatan Tamalatea (variabel 3), cluster 3 beraggotakan Kecamatan Rumbia (variabel 8), kecamatan Turatea (variabel 6), kecamatan Batang (variabel 7), kecamatan Kelara (variabel 10) dan kecamatan Baangkala Barat (variabel 2) dan cluster 4 beraggotakan kecamatan Tarawang (variabel 11), dan kecamatan Arungkeke (variabel 9).

Dilihat dari besarnya jumlah penduduk buta aksara disetiap Kecamatan dalam 4 cluster yaitu:

1) Cluster 1 jumlah penduduknya sebesar 50 orang antara lain berdasarkan faktor ekonominya jumlah penduduk buta aksara di kecamatan Bontoramba dan kecamatan Binamu sebesar 32 orang, berdasarkan faktor budayanya jumlah penduduk buta aksara sebesar 16 orang dan berdasarkan faktor geografisnya jumlah penduduk buta aksara di kecamatan Bontoramba 2 orang sedangkan kecamatan Binamu tidak ada. Dari jumlah penduduk yang ada maka dapat disimpulkan bahwa daerah-daerah dalam cluster 1 merupakan daerah yang sedang penduduk buta aksaranya.

2) Cluster 2 jumlah penduduknya sebesar 52 orang, antara lain berdasarkan faktor ekonominya jumlah penduduk buta aksara di kecamatan Bangkala dan kecamatan Tamalatea sebesar 21 orang, berdasarkan faktor budayanya jumlah penduduk buta aksara di kecamatan Bangkala dan kecamatan Tamalatea sebesar 23 orang dan berdasarkan faktor geografisnya jumlah penduduk buta aksara di kecamatan Bangkala dan kecamatan Tamalatea sebesar 8 orang. Dari jumlah penduduk yang ada maka dapat disimpulkan bahwa daerah-daerah dalam cluster 2 merupakan daerah yang sedang penduduk buta aksaranya.

3) Cluster 3 jumlah penduduk sebesar 74 orang, antara lain berdasarkan faktor ekonominya jumlah penduduk buta aksara di kecamatan Rumbia, kecamatan Turatea, kecamatan Batang, kecamatan Kelara, dan kecamatan Bangkala Barat sebesar 36 orang, berdasarkan faktor budayanya jumlah penduduk buta aksara di kecamatan Rumbia, kecamatan Turatea, kecamatan Batang, kecamatan Kelara, dan kecamatan B.Barat sebesar 25 orang dan berdasarkan faktor geografisnya jumlah penduduk buta aksara di kecamatan Rumbia, kecamatan Turatea, kecamatan Batang, kecamatan Kelara, dan kecamatan sebesar 13 orang sedangkan B.Barat tidak ada. Dari jumlah penduduk yang ada maka dapat disimpulkan bahwa daerah-daerah dalam cluster 3 merupakan daerah yang paling banyak penduduk buta aksaranya.

4) Cluster 4 jumlah penduduknya sebesar 24 orang, antara lain berdasarkan faktor ekonominya jumlah penduduk buta aksara di kecamatan Tarowang dan kecamatan Arungkeke sebesar 9 orang, berdasarkan faktor budayanya jumlah penduduk buta aksara di kecamatan Tarowang dan kecamatan Arungkeke sebesar 11 orang dan berdasarkan faktor geografisnya jumlah penduduk buta aksara di kecamatan Tarowang dan kecamatan Arungkeke sebesar 4 orang. Dari jumlah penduduk yang ada maka dapat disimpulkan bahwa daerah-daerah dalam cluster 4 merupakan daerah yang paling sedikit penduduk buta aksaranya.

Dilihat dari jenis faktor yang mempengaruhi penduduk buta aksara disetiap kecamatan dalam 4 cluster yaitu:

1) Jenis faktor yang mempengaruhi penduduk buta aksara disetiap Kecamatan dalam cluster 1 (kecamatan Bontoramba dan kecamatan binamu) yaitu faktor ekonomi, budaya dan geografis. meskipun faktor tersebut masuk dalam cluster satu tapi jumlah penduduk buta aksaranya sedang. Hal ini disebabkan karena daerah-daerah sekitarnya termasuk mata pencahariannya sangat besar dan sudah ada sekolah-sekolah didirikan di sekitar daerah tersebut. Sehingga dapat disimpulkan bahwa daerah tersebut sudah dikategorikan berpendidikan.

2) Jenis faktor yang mempengaruhi penduduk buta aksara disetiap Kecamatan dalam cluster 2 (kecamatan Bangkala dan kecamatan Tamalatea) yaitu faktor ekonomi, budaya dan geografis. Meskipun faktor tersebut masuk dalam cluster dua tapi jumlah penduduk buta aksaranya sedang. Hal ini disebabkan karena daerah-daerah sekitarnya termasuk mata pencahariannya sangat besar dan sudah ada sekolah-sekolah didirikan di sekitar daerah tersebut. Sehingga dapat disimpulkan bahwa daerah tersebut sudah dikategorikan berpendidikan, dan mempunyai lembaga latihan kerja yang dipersiapkan bagi anak yang putus sekolah.

3) Jenis faktor yang mempengaruhi penduduk buta aksara disetiap Kecamatan dalam cluster 3 (kecamatan Rumbia, kecamatan Turatea, kecamatan Batang, kecamatan Kelara, dan kecamatan Bangkala Barat) yaitu faktor ekonomi, budaya dan geografis. Daerah ini merupakan penduduknya masih mengikuti tradisi-tradisi keluarga dan pencahariannya bagus akan tetapi penduduknya masih terlalu dipengaruhi tradisi tempo dulu. Sehingga bias disimpulkan bahwa daerah tersebut pendidikannya masih tergolong rendah dibandingkan dengan daerah-daerah lain.

4) Jenis faktor yang mempengaruhi penduduk buta aksara disetiap Kecamatan dalam cluster 4 (kecamatan Tarowang dan kecamatan Arungkeke) yaitu faktor

ekonomi, budaya dan geografis. Daerah tersebut sebagian besar sudah banyak yang mengerti pendidikan dan mata pencaharian sudah sangat besar apalagi sudah banyak sekolah-sekolah yang didirikan. Sehingga disimpulkan bahwa daerah tersebut merupakan daerah yang berpendidikan.

Hasil pengclusteran yang diperoleh dari analisis biasa menjadi bahan pertimbangan bagi pemerintah kabupaten Jeneponto dalam mengambil kebijakan. Kecamatan atau daerah yang jumlah penduduk buta aksaranya paling besar dapat menjadi kebijakan pemerintah agar nantinya dapat didirikan sekolah-sekolah bagi mereka. Demikian pula dengan melihat jenis faktor-faktor yang paling mempengaruhi yaitu kurangnya mata pencaharian sehingga banyak penduduk yang keluar kota untuk mencari nafkah untuk menghidupi keluarganya, jadi penduduk kurang menyekolahkan anak-anak mereka disebabkan karena persoalan biaya. Dan banyak penduduk yang secara turun-temurun atau tradisi yang tidak bisa dihilangkan dan sampai sekarang masih terjadi yang demikian yaitu jika melihat anaknya sudah besar terutama anak perempuan maka akan segera dipingit oleh keluarganya sendiri.

IV. KESIMPULAN

Pada analisis cluster diaplikasikan kepada pengelompokan kecamatan berdasarkan faktor penyebab penduduk buta aksara di Kabupaten Jeneponto. Maka didapat dari hasil penelitian bahwa faktor yang menyebabkan penduduk buta aksara yang paling banyak di Kabupaten Jeneponto yaitu faktor ekonomi di antara sebelas kecamatan. Adapun kecamatan yang paling banyak penduduk buta aksaranya yang disebabkan dari faktor ekonomi yakni Kecamatan Binamu.

DAFTAR PUSTAKA

- Arnold F. Steven.1944. *The Theory Of Linear Models And Multivariate Analysis*. The United States Of America.
- Bado, Alimi dan Anisa. 2006. *Analisis Gerombol (Pelatihan Analisis Data Multivariat menggunakan SPSS, MINITAB, dan SAS bagi dosen PTN kawasan Timur Indonesia)*. Departemen Pendidikan Nasional Direktorat Jenderal Pendidikan Tinggi dengan Fakultas MIPA UNHAS. Makassar.

- , 2006. *Konsep Dasar Statistik Multivariat (Pelatihan Analisis Data Multivariat menggunakan SSPSS, MINITAB, dan SAS bagi Dosen PTN kawasan Timur Indonesia)*. Departemen Pendidikan Nasional Direktorat Jenderal Pendidikan Tinggi dengan Fakultas MIPA UNHAS. Makassar.
- Bhuono, Agung. 2005. *Strategi Jitu Memilih Metode Statistik penelitian dengan SPSS*. PT. Andi. Yogyakarta.
- Departemen Statistik FMIPA ITB dengan Rektorat Jenderal Pendidikan Tinggi. *Departemen Pendidikan Nasional Aplikasi*. 2004. Modul teori. *Pelatihan Analisis Multivariat*. Departemen Pendidikan Nasional, Jakarta.
- Djojonegoro Wardiman. 2002. *Pendidikan Nasional Sebagai Proses Transformasi Budaya*. Balai Pustaka Jakarta.
- Ilmu Analisis Cluster*. 2009. (on line). [Http:// www. Youngstatistician.Com](http://www.Youngstatistician.Com). Diakses 28/12/2009.
- Jalal, Fasli. 2003. *Sulit pemberantasan Buta Aksara (on line)*. [Http: // www. Kompas Com/ kompas_Cetak / dikbud. Html](http://www.Kompas.Com/kompas_Cetak/dikbud.Html). Diakses 28/12/2009.
- J. A. Shaw Peter. 1987. *Multivariat Statistics for the Enviromental Sciences*. Oxford University Press Inc.. New York.
- Johnson Richard dan Wichern Dean. 2002. *Applied Multivariate Statiscal analysis*. United States of America.
- Mattjik Ansori Ahmad. 2002. *Analisis Peubah Ganda*. PKSDM-DIKTI DEPDIKNAS Jurusan Statistik FAK MIPA IPB. Bogor.
- Morrison F. Donald. 1967. *Multivariat Statistical Methods*. R. R. Donnelley & Sons Comppany. The United State of America.
- Rosyada, Dede. 2004. *Paradigma Pendidikan Demokratis (sebuah model perlibatan masyarakat dalam penyelenggaraan pendidikan)*. Kencana Jakarta.
- Santoso Singgih. 2002. *SPSS Statistik Multivariat*. PT. Elex media komputindo, Jakarta.
- Sekolah Tinggi Ilmu Statistik. 2006. *Penelitian meggunakan Statistik (on loine)*. [Http://www. Youngstascian.com](http://www.Youngstascian.com)). diakses 12/3/2010.

Simamora, Bilson. 2005. *Analisis Multivariat Pemasaran*. PT Gramedia Pustaka Utama. Jakarta.

Supranto, J. 2004. *Analisis Multivariat Arti dan Interpretasi*. Rineka Cipta. Jakarta.

Uci. 2010. *Cluster Analysis*. [Http://www.pc.Statistic/Statis](http://www.pc.Statistic/Statis) off.Pdf.Online.