

**PERBANDINGAN INDEKS VALIDITAS CALINSKI-HARABASZ,
HARTIGAN dan KRZANOWSKI-LAI pada ANALISIS *CLUSTER*
HIRARKI**

SKRIPSI

**oleh :
ARSY KAMILIN
0410950006-95**



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2008**

**'PERBANDINGAN INDEKS VALIDITAS CALINSKI-
HARABASZ, HARTIGAN dan KRZANOWSKI-LAI pada
ANALISIS *CLUSTER* HIRARKI**

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Sains dalam bidang Statistika

oleh:
ARSY KAMILIN
0410950006-95



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG**

2008

LEMBAR PENGESAHAN SKRIPSI

**PERBANDINGAN INDEKS VALIDITAS CALINSKI-
HARABASZ, HARTIGAN dan KRZANOWSKI-LAI pada
ANALISIS *CLUSTER* HIRARKI**

oleh:
ARSY KAMILIN
0410950006-95

Setelah dipertahankan di depan Majelis Penguji
pada tanggal 3 September 2008
dan dinyatakan memenuhi syarat untuk memperoleh gelar
Sarjana Sains dalam bidang Statistika

Pembimbing I

Pembimbing II

Ir. Atiek Iriany, M.S.
NIP. 131 759 544

Ir. Heni Kusdarwati, M.S.
NIP. 131 652 676

Mengetahui,
Ketua Jurusan Matematika
Fakultas MIPA Universitas Brawijaya,

Dr. Agus Suryanto, M.Sc.
NIP. 132 126 049

LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama : Arsy Kamilin
NIM : 0410950006-95
Jurusan : Matematika
Penulis skripsi berjudul : PERBANDINGAN INDEKS
VALIDITAS CALINSKI-
HARABASZ, HARTIGAN
dan KRZANOWSKI-LAI
pada ANALISIS *CLUSTER*
HIRARKI

Dengan ini menyatakan bahwa :

1. Isi dari skripsi yang saya buat adalah benar-benar karya sendiri dan tidak menjiplak karya orang lain, selain nama-nama yang termaktub di isi dan tertulis di daftar pustaka dalam skripsi ini.
2. Apabila di kemudian hari ternyata skripsi yang saya tulis terbukti hasil jiplakan, maka saya akan bersedia menanggung segala risiko.

Demikian pernyataan ini saya buat dengan segala kesadaran.

Malang, 3 September 2008
Yang menyatakan,

Arsy Kamilin
NIM. 0410950006-95

PERBANDINGAN INDEKS VALIDITAS CALINSKI-HARABASZ, HARTIGAN dan KRZANOWSKI-LAI pada ANALISIS *CLUSTER* HIRARKI

ABSTRAK

Analisis *cluster* merupakan suatu teknik yang digunakan untuk mengelompokkan beberapa objek ke dalam beberapa *cluster* sehingga objek-objek dalam satu *cluster* bersifat homogen. Salah satu masalah penting dalam analisis *cluster* adalah menentukan banyaknya *cluster* optimal. Banyaknya *cluster* optimal dapat diperoleh dari indeks validitas. Indeks validitas yang dibahas dalam penelitian ini adalah indeks Calinski-Harabasz (CH), Hartigan (H) dan Krzanowski-Lai (KL). Penelitian ini bertujuan menentukan banyaknya *cluster* optimal berdasarkan ketiga indeks dan menentukan indeks validitas terbaik berdasarkan nilai *Cluster Tightness Measure* (CTM). Indeks validitas yang menghasilkan CTM terkecil merupakan indeks validitas terbaik. Data yang digunakan dalam penelitian ini adalah data sekunder. Banyaknya *cluster* optimal berdasarkan indeks H lebih kecil atau sama dengan *cluster* optimal berdasarkan indeks KL dan lebih kecil dari *cluster* optimal berdasarkan indeks CH ($H \leq KL < CH$). Indeks CH selalu menghasilkan banyaknya *cluster* yang maksimal untuk data yang tidak menggerombol membentuk *cluster* dengan jelas. Jika nilai CTM yang dibandingkan hanya antara indeks H dan indeks KL, dapat disimpulkan bahwa indeks KL lebih baik dari indeks H.

Kata kunci: Analisis *cluster*, *cluster* optimal, indeks validitas

COMPARISON OF CALINSKI-HARABASZ INDEX, HARTIGAN INDEX and KRZANOWSKI-LAI INDEX in HIERARCHICAL CLUSTER ANALYSIS

ABSTRACT

Cluster analysis is a technique used for grouping observations into clusters such that observations in each cluster are similar to each other. A major challenge in cluster analysis is the estimation of the optimal number of clusters. The optimal number of clusters can be estimated by validity index. Validity indexes explained in this research are Calinski-Harabasz (CH) index, Hartigan (H) index and Krzanowski-Lai (KL) index. This research will estimate the optimal number of clusters with those indexes and determine the best index based on Cluster Tightness Measure (CTM). Index with the smallest CTM is the best index. Secondary data are used in this research. The optimal number of clusters for CH index is less than or equal the optimal number of clusters for KL index less than the optimal number of clusters for CH index. CH index always gives the optimal number of clusters as many as the number of objects, for the data that is not clearly making clusters. If the CTM compared just among H index and KL index, it can be concluded that KL is the best index.

Keywords: Cluster analysis, optimal cluster, validity index

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul Perbandingan Indeks Validitas Calinski-Harabasz, Hartigan dan Krzanowski-Lai pada Analisis Cluster Hirarki sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains dalam bidang Statistika.

Ketiga indeks validitas merupakan indeks yang dihitung berdasarkan jumlah kuadrat dalam *cluster*. Tiga indeks tersebut dibandingkan pada analisis *cluster* hirarki dengan *average linkage*. CTM digunakan untuk membandingkan ketiga indeks validitas tersebut.

Penulis mengucapkan terima kasih kepada:

1. Ibu Ir. Atiek Iriany, M.S., selaku Dosen Pembimbing I atas waktu, tenaga dan arahan yang telah diberikan kepada penulis selama penyusunan skripsi.
2. Ibu Ir. Heni Kusdarwati, M.S., selaku Dosen Pembimbing II yang telah menyisihkan waktu dan tenaga serta memberikan pengarahan kepada penulis selama penyusunan skripsi.
3. Ibu Suci Astutik, S.Si., M.Si., Ibu Eni Sumarminingsih, S.Si., M.M. dan Bapak Adji Achmad Rinaldo Fernandes, S.Si., M.Sc. selaku Dosen Penguji.
4. Semua pihak yang telah membantu dalam penulisan skripsi.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna karena keterbatasan kemampuan penulis. Oleh karena itu, penulis mengharap kritik dan saran. Akhir kata, penulis berharap semoga skripsi ini dapat bermanfaat bagi pembaca, baik yang ingin mempelajari analisis *cluster* hirarki dan indeks validitas.

Malang, 3 September 2008

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL	x
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN	
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah	2
1.4. Tujuan Penelitian.....	2
1.5. Manfaat Penelitian.....	2
BAB II TINJAUAN PUSTAKA	
2.1. Penyajian Data.....	3
2.2. Multikolinieritas	4
2.3. Analisis Komponen Utama.....	4
2.3.1. Pembentukan Komponen Utama	5
2.3.2. Tipe Masukan	7
2.3.3. Skor Komponen Utama	7
2.4. Analisis <i>Cluster</i>	8
2.5. Ukuran Jarak.....	9
2.6. Analisis <i>Cluster</i> Hirarki	9
2.6.1. Metode <i>Agglomerative</i>	10
2.6.2. Metode <i>Divisive</i>	11
2.7. Indeks Validitas <i>Cluster</i>	11
2.7.1. Indeks Validitas Calinski-Harabasz.....	11
2.7.2. Indeks Validitas Hartigan	12
2.7.3. Indeks Validitas Krzanowski-Lai	12
2.8. <i>Cluster Tightness Measure</i> (CTM).....	13
BAB III METODE PENELITIAN	
3.1. Data.....	15
3.2. Metode Penelitian	15

3.3	Diagram Alir.....	16
-----	-------------------	----

BAB IV HASIL DAN PEMBAHASAN

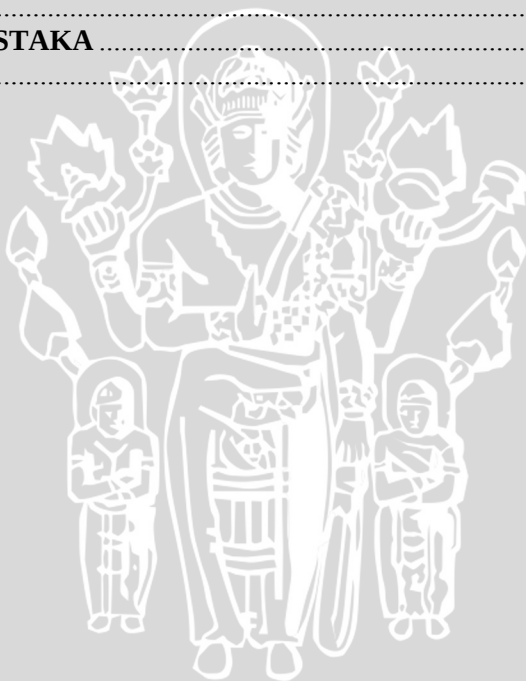
4.1.	Data Usaha Ternak Sapi di Kabupaten Lamongan.....	19
4.2.	Data Pengusaha Tani Brokoli Organik di Kecamatan Pujon.....	22
4.3.	Data Kabupaten/Kota di Propinsi Jawa Timur Berdasarkan Penduduk Usia 10 Tahun Ke Atas Menurut Kabupaten/Kota dan Status Pendidikan.....	26
4.4.	Data PDRB Beberapa Sub Sektor Kabupaten Ponorogo..	29
4.5.	Data Industri Tempe di Kota Malang.....	32

BAB V KESIMPULAN DAN SARAN

5.1.	Kesimpulan.....	39
5.2.	Saran.....	39

DAFTAR PUSTAKA	41
-----------------------------	----

LAMPIRAN	45
-----------------------	----



DAFTAR GAMBAR

	Halaman
Gambar 3.1. Diagram alir metode penelitian.....	17
Gambar 4.1. Dendrogram Data 1	19
Gambar 4.2. Diagram pencar (<i>scatterplot</i>) dan perbandingan <i>cluster</i> yang terbentuk pada Data 1	20
Gambar 4.3. Plot antara banyaknya <i>cluster</i> dan $W(k)$ Data 1	20
Gambar 4.4. Dendrogram Data 2	23
Gambar 4.5. Diagram pencar (<i>scatterplot</i>) dan perbandingan <i>cluster</i> yang terbentuk pada Data 2	23
Gambar 4.6. Plot antara banyaknya <i>cluster</i> dan $W(k)$ Data 2	24
Gambar 4.7. Dendrogram Data 3	26
Gambar 4.8. Diagram pencar (<i>scatterplot</i>) dan perbandingan <i>cluster</i> yang terbentuk pada Data 3	27
Gambar 4.9. Plot antara banyaknya <i>cluster</i> dan $W(k)$ Data 3	27
Gambar 4.10. Dendrogram Data 4	30
Gambar 4.11. Plot antara banyaknya <i>cluster</i> dan $W(k)$ Data 4	30
Gambar 4.12. Dendrogram Data 5	33
Gambar 4.13. Diagram pencar (<i>scatterplot</i>) dan perbandingan <i>cluster</i> yang terbentuk pada Data 5	33
Gambar 4.14. Plot antara banyaknya <i>cluster</i> dan $W(k)$ Data 5	34

DAFTAR TABEL

	Halaman
Tabel 4.1. Banyaknya <i>cluster</i> optimal tiga indeks validitas Data 1	21
Tabel 4.2. Nilai CTM untuk tiga indeks validitas Data 1	22
Tabel 4.3. Banyaknya <i>cluster</i> optimal tiga indeks validitas Data 2	24
Tabel 4.4. Nilai CTM untuk tiga indeks validitas Data 2	25
Tabel 4.5. Banyaknya <i>Cluster</i> optimal tiga indeks validitas Data 3	28
Tabel 4.6. Nilai CTM untuk tiga indeks validitas Data 3	29
Tabel 4.7. Banyaknya <i>cluster</i> optimal tiga indeks validitas Data 4	31
Tabel 4.8. Nilai CTM untuk tiga indeks validitas Data 4	32
Tabel 4.9. Banyaknya <i>cluster</i> optimal tiga indeks validitas Data 5	34
Tabel 4.10. Nilai CTM untuk tiga indeks validitas Data 5	35
Tabel 4.11. Banyaknya <i>cluster</i> optimal dan nilai CTM lima data	36

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data	45
Lampiran 2. Hasil Pembakuan Data	51
Lampiran 3. Pengujian Multikolinieritas	55
Lampiran 4. Skor Komponen Utama	56
Lampiran 5. Indeks Validitas	62



BAB I PENDAHULUAN

1.1. Latar Belakang

Analisis *cluster* merupakan suatu teknik yang digunakan untuk mengelompokkan beberapa objek ke dalam beberapa *cluster* sehingga objek-objek dalam satu *cluster* bersifat homogen. Setiap *cluster* memiliki karakteristik yang berbeda dengan *cluster* lain.

Saat ini analisis *cluster* telah digunakan di berbagai cabang ilmu. Misal di bidang psikologi untuk melakukan klasifikasi individu berdasarkan kepribadian, di bidang riset pemasaran untuk mengelompokkan pelanggan berdasarkan produk yang digunakan dan di bidang kimia untuk melakukan klasifikasi senyawa berdasarkan ciri-ciri kimiawi dan sebagainya.

Salah satu masalah penting dalam analisis *cluster* adalah menentukan banyaknya *cluster* optimal. *Cluster* optimal adalah *cluster* yang padat antar individu dalam *cluster* dan terisolasi dari *cluster* lain dengan baik. Cara paling sederhana untuk menentukan banyaknya *cluster* yang dibentuk adalah dengan melihat dendrogram. Kemudian muncul metode penentuan banyaknya *cluster* optimal dengan menggunakan indeks validitas berdasarkan jumlah kuadrat dalam *cluster* atau $W(k)$. Contohnya indeks validitas Hartigan, Krzanowski-Lai dan Calinski-Harabasz. Indeks validitas mengukur kecukupan dari struktur dalam mendapatkan *cluster* hasil analisis, yang berarti struktur *cluster* memberikan informasi sebenarnya mengenai data. Menurut penelitian yang telah dilakukan oleh Anggraini (2006), indeks validitas lebih sensitif dalam menentukan banyaknya *cluster* optimal. Menurut Mufidah (2004), indeks validitas sebaiknya dihitung, karena selain untuk menentukan *cluster* optimal, indeks validitas juga dapat digunakan untuk menentukan jarak dan jenis pautan yang sesuai dengan data.

Penelitian ini dilakukan untuk mengetahui indeks validitas yang paling baik, di antara tiga indeks yang dibandingkan, berdasarkan nilai *Cluster Tightness Measure* (CTM). CTM merupakan suatu nilai yang digunakan untuk mengukur kebaikan hasil analisis *cluster* berdasarkan simpangan baku tiap peubah pada tiap kelompok. Indeks validitas yang menghasilkan CTM terkecil merupakan indeks validitas terbaik.

1.2. Rumusan Masalah

1. Berapa banyaknya *cluster* optimal dalam analisis *cluster* hirarki dengan indeks validitas Calinski-Harabasz, Hartigan dan Krzanowski-Lai?
2. Indeks validitas apa yang terbaik untuk menentukan banyaknya *cluster* optimal?

1.3. Batasan Masalah

Pada penelitian ini permasalahan dibatasi pada analisis *cluster* hirarki dengan *average linkage* menggunakan jarak Euclid. Penentuan indeks validitas terbaik menggunakan nilai CTM.

1.4. Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Menentukan banyaknya *cluster* optimal dengan indeks validitas Calinski-Harabasz, Hartigan dan Krzanowski-Lai dalam analisis *cluster* hirarki.
2. Membandingkan ketiga indeks validitas berdasarkan nilai CTM.

1.5. Manfaat Penelitian

Dengan penelitian ini diharapkan dapat diketahui indeks validitas apa yang paling baik dalam menentukan banyaknya *cluster* optimal sehingga mempermudah penentuan banyaknya *cluster* optimal pada penelitian selanjutnya.

BAB II TINJAUAN PUSTAKA

2.1. Penyajian Data

Misal peubah yang diamati adalah $X_1, X_2, X_3, \dots, X_p$ dan pengamatan dilakukan terhadap setiap objek anggota sampel berukuran N . Hasil pengamatan objek-objek tersebut disajikan sebagai matriks:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \dots & x_{NP} \end{bmatrix}$$

di mana N = banyaknya objek

P = banyaknya peubah

Jarak Euclid dapat digunakan untuk menentukan kedekatan jika satuan seluruh peubah sama. Jika satuan peubah berbeda, jarak Euclid tetap dapat digunakan, namun data perlu dibakukan terlebih dahulu hingga menjadi matriks skor baku Z :

$$Z = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1p} \\ z_{21} & z_{22} & \dots & z_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ z_{N1} & z_{N2} & \dots & z_{NP} \end{bmatrix}$$

$$\text{di mana } z_{np} = \frac{x_{np} - \bar{x}_p}{s_p}$$

$$n = 1, 2, 3, \dots, N$$

$$p = 1, 2, 3, \dots, P$$

$$s_p = \text{simpangan baku peubah ke-}p$$

Menurut Hair *et al.* (1998), multikolinieritas seringkali ditemukan pada data multivariat. Jika jarak yang digunakan adalah jarak Euclid, adanya multikolinieritas pada analisis *cluster* dapat menyebabkan objek masuk dalam *cluster* yang salah sehingga hasil analisis *cluster* tidak valid. Oleh karena itu, sebelum dilakukan analisis *cluster* perlu dilakukan uji apakah terdapat multikolinieritas.

2.2. Multikolinieritas

Multikolinieritas merupakan suatu istilah yang digunakan untuk menunjukkan adanya korelasi yang tinggi di antara beberapa peubah. Menurut Bluman (1992), korelasi digunakan untuk menentukan apakah terdapat hubungan antar peubah. Koefisien korelasi menunjukkan tingkat keeratan hubungan antara peubah ke- p dan peubah ke- m di mana $p, m = 1, 2, 3, \dots, P$. Koefisien korelasi populasi antara X_p dan X_m dinotasikan dengan ρ_{pm} , sedangkan koefisien korelasi contoh dinotasikan dengan r_{pm} . Koefisien korelasi contoh dihitung melalui rumus:

$$r_{pm} = \frac{\text{cov}(X_p, X_m)}{\sqrt{\text{var}(X_p) \text{var}(X_m)}}$$

di mana:

$\text{cov}(X_p, X_m)$ = peragam peubah X_p dan X_m

$\text{var}(X_p)$ = ragam peubah ke- p

$\text{var}(X_m)$ = ragam peubah ke- m

VIF (*Variance Inflation Factor*) merupakan salah satu indikator untuk menentukan apakah terdapat multikolinieritas. Nilai VIF merupakan unsur-unsur diagonal utama invers matriks korelasi peubah-peubah X . Nilai $\text{VIF} > 10$ menunjukkan adanya multikolinieritas (Draper dan Smith, 1992). Matriks korelasi adalah sebagai berikut:

$$\mathbf{R} = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & 1 \end{bmatrix}$$

Untuk menghilangkan multikolinieritas, dapat dilakukan analisis komponen utama sebelum dilakukan analisis *cluster*.

2.3. Analisis Komponen Utama

Analisis komponen utama (AKU) merupakan teknik analisis multivariat yang dilakukan untuk mereduksi peubah dan membentuk peubah baru yang ortogonal. Peubah baru tersebut dinamakan komponen utama (Afifi dan Clark, 1990).

Menurut Gaspersz (1992), untuk menerangkan keragaman total sistem, p komponen utama dapat dihasilkan dari p peubah asal. Namun, seringkali keragaman total itu dapat diterangkan oleh sejumlah kecil komponen utama (q), di mana $q < p$.

Dalam AKU vektor peubah asal $\underline{X} = (X_1, X_2, X_3, \dots, X_p)$ ditransformasi menjadi vektor peubah baru $\underline{Y}$. Bentuk persamaan AKU dinyatakan sebagai:

$$Y_w = a_{1w}X_1 + a_{2w}X_2 + \dots + a_{pw}X_p$$

di mana $a_{1w}, a_{2w}, \dots, a_{pw}$ merupakan koefisien komponen utama ke- w , $1 \leq w \leq q$, $Y_1, Y_2, \dots, Y_q$ saling bebas dan vektor peubah baru $\underline{Y}$ menjelaskan proporsi keragaman vektor peubah asal sebesar mungkin.

2.3.1. Pembentukan Komponen Utama

Persamaan AKU dapat diuraikan menjadi:

$$Y_1 = a_{11}X_1 + a_{21}X_2 + \dots + a_{p1}X_p$$

$$Y_2 = a_{12}X_1 + a_{22}X_2 + \dots + a_{p2}X_p$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots$$

$$Y_p = a_{1p}X_1 + a_{2p}X_2 + \dots + a_{pp}X_p$$

di mana:

$$Var(Y_w) = a'_w \Sigma a_w$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_p^2 \end{bmatrix}$$

Penentuan komponen utama $Y_w = a_{1w}X_1 + a_{2w}X_2 + \dots + a_{pw}X_p$ berarti mencari vektor koefisien $a'_w = (a_{1w}, a_{2w}, a_{3w}, \dots, a_{pw})$ dihitung sedemikian hingga ragam komponen utama mencapai nilai maksimum, dengan syarat $a'_w a_w = 1$. Persoalan ini dapat diselesaikan dengan fungsi Lagrange yang memaksimumkan

$\underline{a}'_w \underline{\Sigma} \underline{a}_w$ dengan kendala $\underline{a}'_w \underline{a}_w = 1$ atau $\underline{a}'_w \underline{a}_w - 1 = 0$. Jadi, dapat dibentuk fungsi Lagrange sebagai berikut:

$$L = \underline{a}'_w \underline{\Sigma} \underline{a}_w - \lambda_w (\underline{a}'_w \underline{a}_w - 1)$$

Apabila L diturunkan terhadap $\underline{a}_w$ kemudian disamadengkan nol, maka diperoleh:

$$\begin{aligned} \frac{\partial L}{\partial \underline{a}} &= 2 \underline{a}_w \underline{\Sigma} - 2 \lambda_w \underline{a}_w = 0 \\ 2(\underline{\Sigma} - \lambda_w \underline{I}) \underline{a}_w &= 0 \\ (\underline{\Sigma} - \lambda_w \underline{I}) \underline{a}_w &= 0 \end{aligned} \quad (2.1)$$

Persamaan (2.1) akan menghasilkan solusi *non trivial* apabila matriks $(\underline{\Sigma} - \lambda_w \underline{I})$ merupakan matriks singular.

$$|\underline{\Sigma} - \lambda_w \underline{I}| = 0 \quad (2.2)$$

Penyelesaian persamaan (2.2) akan menghasilkan akar-akar ciri $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_p$ di mana $\lambda_1 > \lambda_2 > \lambda_3 > \dots > \lambda_p$. Setiap akar ciri λ_w akan menentukan vektor ciri yang dinormalkan dari $\underline{\Sigma}$, yaitu $\underline{a}_w$.

Penentuan akar ciri yang akan digunakan dalam komponen utama pertama diperoleh dari penguraian persamaan (2.1) sebagai berikut:

$$\begin{aligned} (\underline{\Sigma} - \lambda_w \underline{I}) \underline{a}_w &= 0 \\ \underline{\Sigma} \underline{a}_w - \lambda_w \underline{a}_w &= 0 \\ \underline{\Sigma} \underline{a}_w &= \lambda_w \underline{a}_w \end{aligned} \quad (2.3)$$

Jika kedua sisi persamaan (2.3) digandaawalkan dengan $\underline{a}'_w$ maka hasilnya akan menjadi:

$$\underline{a}'_w \underline{\Sigma} \underline{a}_w = \underline{a}'_w \lambda_w \underline{a}_w$$

dan dengan mengingat batasan $\underline{a}'_w \underline{a}_w = 1$, maka diperoleh:

$$\underline{a}'_w \underline{\Sigma} \underline{a}_w = \lambda_w \quad (2.4)$$

Persamaan (2.4) memperlihatkan bahwa ragam setiap komponen utama berpadanan dengan nilai setiap akar ciri yang ada. Jadi, harus dipilih akar ciri terbesar dari matriks $\underline{\Sigma}$ untuk dipergunakan dalam komponen utama pertama agar ragam komponen utama pertama maksimum.

Keragaman yang dapat dijelaskan oleh komponen utama ke- w terhadap keragaman total adalah:

$$\frac{\lambda_w}{\sum_{w=1}^p \lambda_w} \times 100\%$$

dan keragaman total yang dapat dijelaskan oleh q komponen utama pertama adalah:

$$\frac{\sum_{w=1}^q \lambda_w}{\sum_{w=1}^p \lambda_w} \times 100\%$$

di mana q adalah banyaknya komponen utama yang ditentukan (Gaspersz, 1992).

2.3.2. Tipe Masukan

Menurut Johnson dan Wichern (2002), ada dua tipe masukan pada analisis komponen utama yaitu:

1. Matriks ragam-peragam

Digunakan jika semua peubah yang diamati memiliki satuan pengukuran yang sama. Jika komponen utama ke- w adalah:

$$Y_w = a_{1w}X_1 + a_{2w}X_2 + \dots + a_{pw}X_p$$

maka

$$\text{Var}(Y_w) = a'_w \Sigma a_w = \lambda_w \quad w = 1, 2, \dots, p$$

$$\text{Cov}(Y_w, Y_s) = a'_w \Sigma a_s = 0 \quad w \neq s$$

2. Matriks korelasi

Pada data yang mempunyai satuan pengukuran berbeda, digunakan matriks korelasi sebagai masukan. Matriks korelasi populasi ρ diduga dengan matriks korelasi contoh ($\mathbf{R}$).

2.3.3. Skor Komponen Utama

Menurut Gaspersz (1992), skor komponen utama dari individu ke- n untuk komponen utama Y_w yang dihasilkan dari matriks ragam-peragam adalah:

$$SK_{wn} = \begin{bmatrix} a_{w1} & a_{w2} & \cdots & a_{wp} \end{bmatrix} \begin{bmatrix} x_{1n} - \bar{x}_1 \\ x_{2n} - \bar{x}_2 \\ \vdots \\ x_{pn} - \bar{x}_p \end{bmatrix}$$

$$SK_{wn} = \mathbf{a}'_w (x_n - \bar{\mathbf{x}})$$

di mana

SK_{wn} = skor komponen ke- w objek ke- n

Jika komponen utama dihasilkan dari matriks korelasi $\mathbf{R}$, maka matriks data objek digantikan oleh matriks data skor baku $\mathbf{Z}$. Skor komponen utama akan digunakan untuk analisis *cluster* jika terdapat multikolinieritas pada peubah.

2.4. Analisis Cluster

Analisis *cluster* adalah suatu metode analisis yang bertujuan meletakkan objek-objek ke dalam beberapa *cluster* berdasarkan nilai pengamatan sehingga objek-objek yang berada dalam satu *cluster* mirip satu sama lain. Selain itu, analisis *cluster* juga dapat digunakan untuk meringkas data (Everitt, 1980).

Misal pada n objek dilakukan pengamatan terhadap p peubah, maka analisis *cluster* akan mengelompokkan objek-objek ke dalam k *cluster*. Jarak antara objek-objek dalam satu *cluster* akan lebih kecil daripada jarak antara objek dalam suatu *cluster* terhadap objek dalam *cluster* lain.

Dua metode pembentukan *cluster* yakni analisis *cluster* hirarki dan nonhirarki. Menurut Sharma (1996), kelebihan analisis *cluster* hirarki adalah banyaknya *cluster* tidak perlu diketahui. Akan tetapi, dalam analisis *cluster* hirarki objek yang telah masuk dalam suatu *cluster* tidak dapat berpindah ke *cluster* lain.

Setelah analisis *cluster*, perlu dilakukan validasi *cluster*, yaitu suatu prosedur untuk mengevaluasi hasil analisis *cluster* secara kuantitatif sehingga diperoleh *cluster* optimal. *Cluster* optimal adalah *cluster* yang padat antar individu dalam *cluster* dan terisolasi dari *cluster* lain dengan baik serta memberikan informasi yang sebenarnya mengenai data (Jain dan Dubes, 1988). Banyaknya *cluster* optimal ditentukan melalui penerapan beberapa indeks validitas seperti Hartigan, Krzanowski-Lai dan Calinski-Harabasz.

2.5. Ukuran Jarak

Pengelompokan objek dalam analisis *cluster* didasarkan terhadap kedekatan antar objek. Format penyajian data yang digunakan untuk analisis *cluster* hirarki adalah matriks jarak. Jain dan Dubes (1988) menyajikan jarak dalam bentuk matriks setangkup:

$$\begin{bmatrix} d(1,1) & d(1,2) & \cdots & d(1,n) \\ d(2,1) & d(2,2) & \cdots & d(2,n) \\ \vdots & \vdots & \ddots & \vdots \\ d(n,1) & d(n,2) & \cdots & d(n,n) \end{bmatrix}$$

di mana $d(i, j)$ merupakan jarak antara objek ke- i dan objek ke- j .

Menurut Mardia *et al.* (1979), suatu fungsi terhadap objek i dan j , $d(i, j)$, disebut sebagai fungsi jarak jika memenuhi karakteristik:

1. $d(i, j) = d(j, i)$
2. $d(i, j) \geq 0$
3. $d(i, i) = 0$

Hartigan (1975) menyatakan cara yang biasa digunakan untuk menunjukkan kedekatan adalah dengan menggunakan kumpulan jarak antara pasangan-pasangan objek. Jarak yang biasa digunakan pada analisis *cluster* adalah jarak Euclid:

$$d(i, j) = \left\{ \sum_{v=1}^p (x_{iv} - x_{jv})^2 \right\}^{1/2}$$

2.6. Analisis Cluster Hirarki

Dillon dan Goldstein (1984) menyatakan analisis *cluster* hirarki merupakan analisis *cluster* yang melakukan penggabungan atau pemisahan objek secara berulang-ulang. Hasil dari analisis ini biasanya berupa dendrogram, yaitu diagram pohon dua dimensi yang menggambarkan penggabungan atau pemisahan objek. Menurut Afifi dan Clark (1990), analisis *cluster* hirarki dibagi menjadi metode *agglomerative* dan *divisive*.

2.6.1. Metode *Agglomerative*

Dalam metode ini, analisis dimulai dengan n *cluster* yang merupakan n objek dalam penelitian dan diakhiri dengan terbentuknya satu *cluster*. Langkah-langkah metode ini adalah:

1. Membuat matriks jarak Euclid antar objek.
2. Mencari nilai jarak terdekat antara dua objek kemudian menggabungkan kedua objek tersebut menjadi satu *cluster*.
3. Menghitung jarak antara *cluster* yang terbentuk dengan objek lain di luar *cluster* dengan menggunakan metode pautan.
4. Menyusun matriks jarak yang baru.
5. Mengulangi langkah 2, 3 dan 4 hingga terbentuk satu *cluster*.

Sharma (1996) menyajikan tiga metode pautan *agglomerative*:

1. *Single linkage* (pautan tunggal)

Metode *single linkage* membentuk *cluster* berdasarkan jarak terdekat. Dua objek yang memiliki jarak terdekat digabungkan menjadi satu *cluster*. Setelah itu dicari jarak terdekat lagi dan dilihat apakah objek yang memiliki jarak terdekat tersebut bergabung dengan dua objek pertama atau membentuk *cluster* baru dengan objek yang lain. Langkah ini dilakukan terus hingga seluruh objek bergabung membentuk satu *cluster*.

$$d_{(AB)C} = \min\{d_{AC}, d_{BC}\}$$

2. *Complete linkage* (pautan lengkap)

Prosedur metode *complete linkage* mirip dengan metode *single linkage*. Yang menjadi perbedaan adalah kriteria pembentukan *cluster* yang didasari pada jarak terjauh.

$$d_{(AB)C} = \max\{d_{AC}, d_{BC}\}$$

3. *Average linkage* (pautan rata-rata)

Langkah-langkah metode *average linkage* sama dengan metode *single linkage* dan *complete linkage*. Letak perbedaannya hanya pada kriteria pembentukan *cluster*. Kriteria yang digunakan metode ini adalah jarak rata-rata antar objek.

$$d_{(AB)C} = \text{rata - rata}\{d_{AC}, d_{BC}\}$$

2.6.2. Metode *Divisive*

Dalam metode ini, langkah pertama yang dilakukan adalah membentuk satu *cluster* yang isinya seluruh objek. Objek yang paling berbeda dikeluarkan dari *cluster* dan membentuk *cluster* baru. Langkah ini diulangi untuk objek yang paling berbeda hingga terbentuk n *cluster*.

2.7. Indeks Validitas *Cluster*

Menurut Jäkel dan Nöllenburg (2008) hasil analisis *cluster* perlu diuji dengan indeks validitas. Indeks ini dapat digunakan untuk menentukan banyaknya *cluster* optimal. Bolshakova dan Azuaje (2008) menyatakan bahwa dalam aplikasi di mana tidak ada dugaan mengenai banyaknya *cluster* optimal, diperlukan suatu metode validasi *cluster* untuk mengukur kesahihan hasil analisis *cluster*. Indeks validitas *cluster* menunjukkan kualitas hasil analisis *cluster*. Indeks ini mengukur kecukupan struktur dalam mendapatkan *cluster* hasil analisis. Kecukupan struktur *cluster* berarti struktur *cluster* memberikan informasi sebenarnya mengenai data. Indeks validitas tersebut dapat digunakan sebagai dasar penentuan banyaknya *cluster* optimal.

Tiga indeks validitas yang dibahas dalam penelitian ini didasarkan pada jumlah kuadrat dalam *cluster*, $W(k)$. Berdasarkan penelitian yang telah dilakukan oleh Tibshirani, Walther dan Hastie (2000), nilai $W(k)$ akan menurun ketika banyaknya *cluster* (k) meningkat. Akan tetapi pada suatu nilai k tertentu $W(k)$ akan semakin mendatar. Nilai k di mana nilai $W(k)$ mulai mendatar merupakan banyaknya *cluster* yang optimal. Nilai k tersebut akan diduga dengan menggunakan indeks validitas Calinski-Harabasz, (CH), Hartigan (H) dan Krzanowski-Lai (KL).

2.7.1. Indeks Calinski-Harabasz

Calinski dan Harabasz (1974) menyajikan suatu indeks yang didasari oleh *cluster* yang padat dan terisolasi dengan baik. Indeks ini dirumuskan:

$$CH(k) = \frac{B(k)/(k-1)}{W(k)/(n-k)} \quad (2.5)$$

di mana $B(k)$ adalah jumlah kuadrat antar *cluster*, $W(k)$ adalah jumlah kuadrat dalam *cluster* dan k adalah banyaknya *cluster*.

$$B(k) = \frac{1}{2} \sum_{l=1}^k \sum_{\substack{i \in C_l \\ j \notin C_l}} d^2(i, j)$$

$$W(k) = \frac{1}{2} \sum_{l=1}^k \sum_{i, j \in C_l} d^2(i, j)$$

Banyaknya *cluster* optimal diperoleh dari nilai k yang memaksimumkan $CH(k)$.

2.7.2. Indeks Hartigan

Hartigan (1975) memberikan rumus berdasarkan kepadatan *cluster* dan memilih banyaknya *cluster* seminimal mungkin:

$$H(k) = \left(\frac{W(k)}{W(k+1)} - 1 \right) (n - k - 1) \quad (2.6)$$

Indeks ini merupakan rasio yang menyatakan penurunan keragaman dalam *cluster* ketika banyaknya *cluster* berubah dari k menjadi $k+1$. Nilai $H(k)$ yang besar (>10) menunjukkan bahwa penambahan *cluster* dibutuhkan. Hal ini disebabkan ketika $H(k)$ bernilai besar maka $W(k)$ jauh lebih besar dari $W(k+1)$. Banyaknya *cluster* optimal diperoleh dari nilai k terkecil yang menghasilkan $H(k) \leq 10$.

2.7.3. Indeks Krzanowski-Lai

Krzanowski dan Lai (1985) mendefinisikan suatu indeks berdasarkan penurunan nilai jumlah kuadrat dalam *cluster*. Didefinisikan

$$DIFF(k) = \left| (k-1)^{2/p} W(k-1) - k^{2/p} W(k) \right|$$

dan pilih k yang memaksimumkan nilai:

$$KL(k) = \frac{DIFF(k)}{DIFF(k+1)} \quad (2.7)$$

Misal g merupakan banyaknya *cluster* optimal. Untuk $k \leq g$, objek yang berbeda akan membentuk *cluster* lain pada setiap langkah

pembentukan *cluster* yang berurutan dan mengakibatkan nilai $W(k)$ menurun cukup banyak. Untuk $k > g$, penambahan banyaknya *cluster* hanya akan menurunkan nilai $W(k)$ yang tidak banyak. Jadi $\text{DIFF}(k)$ diharapkan akan bernilai kecil untuk seluruh nilai k kecuali $k=g$. Hal ini berarti nilai $\text{KL}(k)$ akan maksimum untuk k optimal.

2.8. *Cluster Tightness Measure (CTM)*

Epps dan Ambikairajah (2008) telah membuat suatu rumus untuk mengukur kebaikan (*efficacy*) hasil analisis *cluster* berdasarkan simpangan baku tiap peubah dari masing-masing *cluster*:

$$CTM = \frac{1}{k} \sum_{l=1}^k \left(\frac{1}{p} \sum_{v=1}^p \frac{\sigma_v^l}{\sigma_v^G} \right) \quad (2.8)$$

di mana σ_v^l merupakan simpangan baku *cluster* ke- l untuk peubah ke- v dan σ_v^G adalah simpangan baku seluruh data untuk peubah ke- v . Jika CTM bernilai nol, maka seluruh data berada pada pusat *cluster*. Jadi, *cluster* yang terbentuk dikatakan baik jika memiliki CTM kecil.

Nilai CTM berdasarkan banyaknya *cluster* optimal yang diperoleh dari ketiga indeks validitas yang telah dijelaskan di subbab 2.7.1, 2.7.2, dan 2.7.3 akan dibandingkan. Indeks yang menghasilkan nilai CTM terkecil adalah metode terbaik.

UNIVERSITAS BRAWIJAYA



BAB III METODE PENELITIAN

3.1. Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari:

1. Data usaha ternak sapi di Kabupaten Lamongan. Data diperoleh dari skripsi Fathur Rohman tahun 2001.
2. Data pengusaha brokoli di Kecamatan Pujon. Data diperoleh dari skripsi Made Lusianawati Mendra tahun 2007.
3. Data kabupaten/kota di propinsi Jawa Timur berdasarkan data penduduk berumur 10 tahun ke atas menurut kabupaten/kota dan status pendidikan tahun 2004. Data diperoleh dari laporan PKL Dwi Hari Cahyono tahun 2005.
4. Data Pendapatan Domestik Regional Bruto (PDRB) beberapa sub sektor Kabupaten Ponorogo tahun 2004. Data diperoleh dari laporan PKL Yunaika Ferlya tahun 2005.
5. Data industri tempe di Kota Malang tahun 2005. Data diperoleh dari skripsi Tresni Wulandari tahun 2006.

Selanjutnya data disebut Data 1 hingga Data 5 dan dapat dilihat pada Lampiran 1.

3.2. Metode Penelitian

Penelitian terbagi menjadi 7 langkah utama:

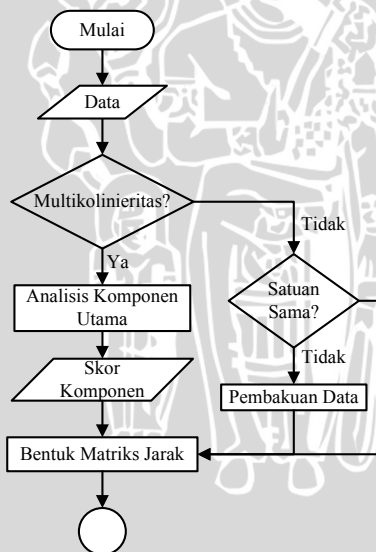
1. Melakukan uji multikolinieritas seperti yang dijelaskan pada sub bab 2.2. Jika terdapat multikolinieritas lakukan analisis komponen utama seperti yang dijelaskan pada sub bab 2.3 dan lanjutkan ke langkah 3. Jika tidak terdapat multikolinieritas maka tidak perlu melakukan analisis komponen utama dan lanjutkan ke langkah 2.
2. Melakukan pembakuan data jika satuan data tidak sama.
3. Melakukan analisis *cluster* hirarki *agglomerative* seperti yang dijelaskan pada sub bab 2.6.1 dengan *average linkage*.
4. Menghitung indeks $CH(k)$ dengan persamaan (2.5), indeks $H(k)$ dengan persamaan (2.6) dan indeks $KL(k)$ dengan persamaan (2.7).

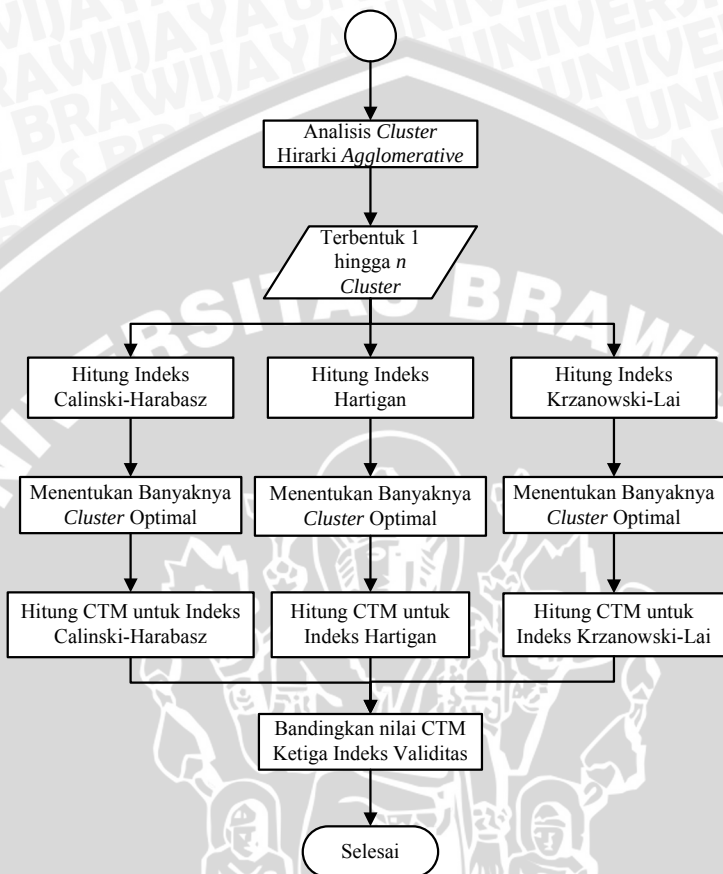
5. Menentukan banyaknya *cluster* optimal berdasarkan ketiga indeks validitas. Untuk indeks $CH(k)$ dipilih nilai k yang memaksimumkan indeks $CH(k)$, untuk indeks $H(k)$ dipilih nilai k terkecil yang menghasilkan $H(k) \leq 10$, untuk indeks $KL(k)$ dipilih nilai k yang memaksimumkan indeks $KL(k)$.
6. Menghitung CTM berdasarkan banyaknya *cluster* optimal untuk ketiga indeks validitas dengan persamaan (2.8).
7. Membandingkan nilai CTM ketiga indeks validitas. Indeks validitas yang menghasilkan CTM terkecil merupakan yang terbaik.

Untuk membantu perhitungan digunakan *software* Minitab 14 dan *package* clusterSim pada *software* R 2.7.0.

3.3. Diagram Alir

Berikut ini diagram alir penentuan banyaknya *cluster* optimal dengan indeks validitas Calinski-Harabasz, Hartigan dan Krzanowski-Lai:





Gambar 3.1. Diagram alir metode penelitian

UNIVERSITAS BRAWIJAYA



BAB IV

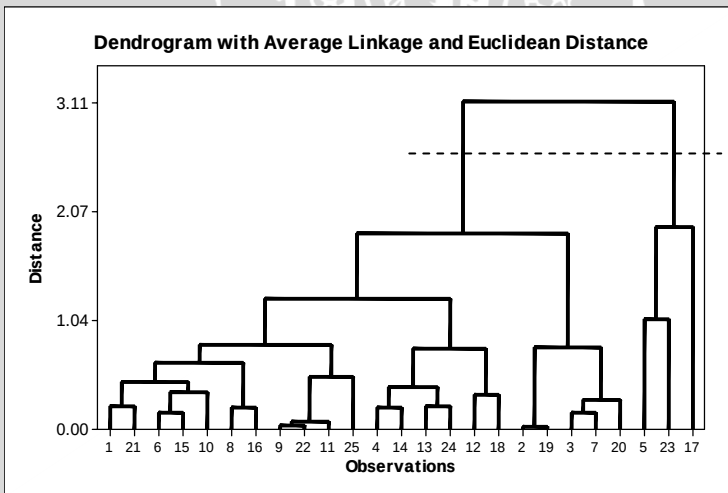
HASIL dan PEMBAHASAN

4.1. Data Usaha Ternak Sapi di Kabupaten Lamongan (Data 1)

Satuan peubah-peubah Data 1 tidak sama, sehingga perlu dilakukan pembakuan data terlebih dahulu. Hasil pembakuan data dapat dilihat pada Lampiran 2.

Pengujian multikolinieritas pada Lampiran 3 menunjukkan bahwa terdapat multikolinieritas pada peubah Data 1. Oleh karena itu perlu dilakukan analisis komponen utama. Skor komponen utama dapat dilihat pada Lampiran 4. Komponen utama pertama mampu menjelaskan keragaman sebesar 88.5% dan komponen utama kedua menjelaskan 11.5%.

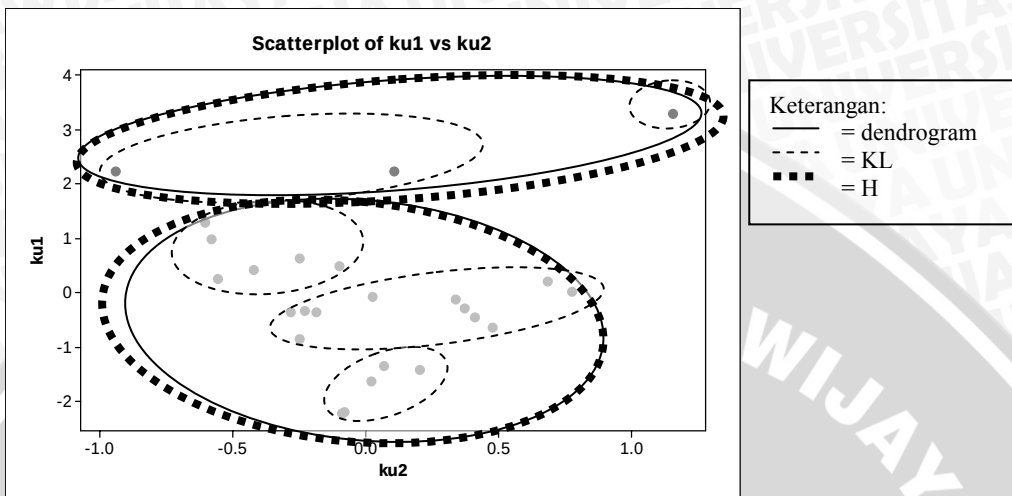
Skor komponen utama akan digunakan untuk analisis *cluster*. Hasil analisis *cluster* yang berupa dendrogram dapat dilihat pada Gambar 4.1.



Gambar 4.1. Dendrogram Data 1

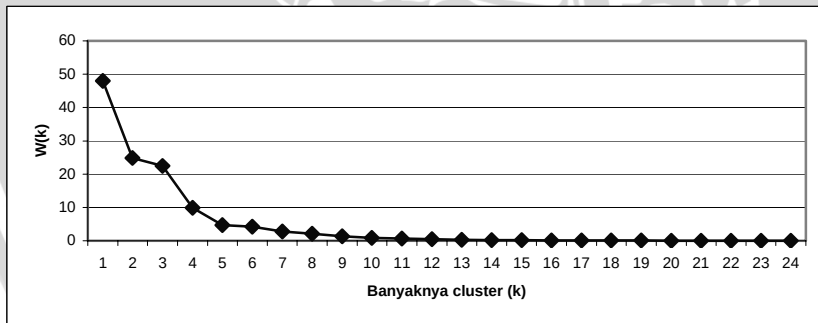
Berdasarkan dendrogram pada Gambar 4.1, banyaknya *cluster* yang terbentuk adalah dua *cluster*.

Diagram pencar serta *cluster* yang terbentuk berdasarkan dendrogram dan indeks validitas Data 1 adalah seperti pada Gambar 4.2. Terlihat bahwa dendrogram dan indeks validitas menghasilkan *cluster* optimal yang berbeda.



Gambar 4.2. Diagram pencar (*scatterplot*) dan perbandingan *cluster* yang terbentuk pada Data 1

Untuk menentukan banyaknya *cluster* optimal menurut indeks validitas CH, H dan KL, harus dihitung $W(k)$ terlebih dahulu. Plot antara banyaknya *cluster* dan $W(k)$ dapat dilihat pada Gambar 4.3 dan nilai $W(k)$ terlampir pada Lampiran 5.



Gambar 4.3. Plot antara banyaknya *cluster* dan $W(k)$ Data 1

Dari Gambar 4.3 dapat dilihat bahwa $W(k)$ mulai mendatar ketika k bernilai 5, ini berarti banyaknya *cluster* optimal kemungkinan adalah lima. Hal ini berbeda dengan penentuan banyaknya *cluster* dengan dendrogram (seperti pada Gambar 4.1) yang menghasilkan

dua *cluster*. Gambar 4.3 memperlihatkan bahwa $W(k)$ belum mendatar ketika k bernilai 2. Ini berarti penentuan banyaknya *cluster* dengan dendrogram tidak peka terhadap perubahan nilai $W(k)$.

Untuk menentukan banyaknya *cluster* optimal harus dihitung indeks validitas terlebih dahulu. Nilai seluruh indeks validitas dapat dilihat pada Lampiran 5. Berdasarkan ketiga indeks validitas diperoleh banyaknya *cluster* optimal seperti pada Tabel 4.1.

Tabel 4.1. Banyaknya *cluster* optimal tiga indeks validitas Data 1

Indeks Validitas	Banyaknya <i>Cluster</i> Optimal
CH	24
H	2
KL	5

Indeks H menghasilkan dua *cluster* optimal. Lampiran 5 memperlihatkan bahwa nilai $W(k)$ hanya mengalami sedikit penurunan ketika k bergerak dari 2 ke 3, sementara rasio antara $W(2)$ dan $W(3)$ adalah sebesar 1.106. Jadi berdasarkan indeks H, tidak perlu dilakukan penambahan *cluster* sehingga banyaknya *cluster* optimal adalah 2.

Indeks KL menghasilkan lima *cluster* optimal, hal ini sesuai dengan Gambar 4.3. Nilai $W(k)$ pada Lampiran 5 memperlihatkan bahwa selisih antara $W(4)$ dan $W(5)$ adalah 5.1371, jauh lebih besar dibandingkan selisih antara $W(5)$ dan $W(6)$ yaitu 0.5519. Hal ini berarti, ketika banyaknya *cluster* bertambah dari 4 menjadi 5, penurunan nilai $W(k)$ lebih tinggi jika dibandingkan dengan perubahan nilai $W(k)$ ketika banyaknya *cluster* bertambah dari 5 menjadi 6.

Banyaknya *cluster* optimal menurut indeks CH merupakan banyaknya *cluster* maksimal yang dapat dibentuk, yaitu 24. Hal ini disebabkan pada saat banyaknya *cluster* mencapai maksimal, maka keragaman antar *cluster* akan maksimal dan keragaman dalam *cluster* akan minimal. Berdasarkan indeks ini, analisis *cluster* seakan-akan menjadi tidak berguna. Pada Gambar 4.2 dan Gambar 4.3 dapat dilihat bahwa tidak tepat jika *cluster* yang terbentuk sebanyak 24 *cluster*. Dapat disimpulkan bahwa indeks CH kurang cocok untuk menentukan banyaknya *cluster* optimal pada Data 1.

Nilai CTM ketiga indeks validitas dapat dilihat pada Tabel 4.2.

Tabel 4.2. Nilai CTM untuk tiga indeks validitas Data 1

Indeks Validitas	CTM
CH	0.000470
H	1.045491
KL	0.285145

Berdasarkan nilai CTM pada Tabel 4.2, terlihat bahwa indeks CH menghasilkan CTM terkecil. Karena indeks CH menghasilkan banyaknya *cluster* optimal 24, yang merupakan banyaknya *cluster* maksimal yang dapat dibentuk, nilai CTM yang dihasilkan menjadi kecil. Tapi, jika banyaknya *cluster* optimal menurut indeks CH dibandingkan dengan Gambar 4.3, maka banyaknya *cluster* optimal menurut indeks CH kurang cocok digunakan.

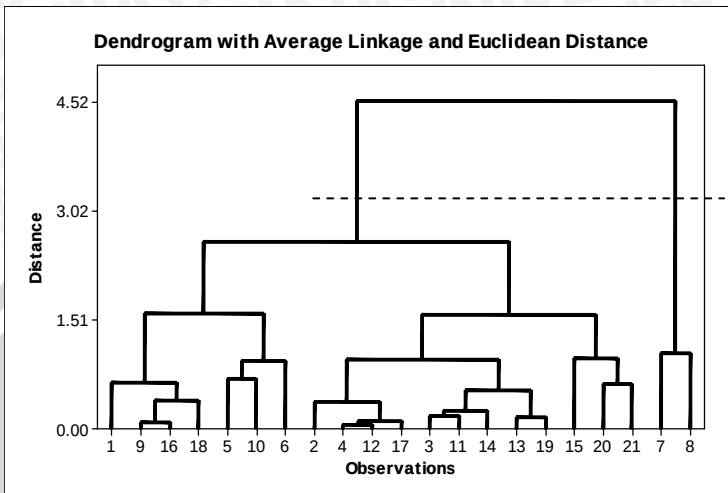
Dari hasil ini dapat diperkirakan bahwa jika nilai CTM digunakan untuk membandingkan hasil analisis *cluster* pada data dengan banyaknya *cluster* optimal yang jauh berbeda, maka kesimpulan yang dihasilkan bisa tidak tepat. Jika indeks CH dihilangkan dan nilai CTM yang dibandingkan adalah indeks H dan KL, maka terlihat bahwa indeks KL lebih baik dari indeks H.

4.2. Data Pengusaha Tani Brokoli Organik di Kecamatan Pujon (Data 2)

Peubah-peubah pada data 2 tidak memiliki satuan yang sama, oleh karena itu perlu dilakukan pembakuan data. Hasil pembakuan data dapat dilihat pada Lampiran 2.

Dari pengujian multikolinieritas pada Lampiran 3, diketahui bahwa terdapat multikolinieritas pada peubah data 2. Oleh karena itu perlu dilakukan analisis komponen utama. Hasil dari analisis komponen utama dapat dilihat pada Lampiran 4. Komponen utama pertama mampu menjelaskan keragaman sebesar 89.6% dan komponen utama kedua menjelaskan 7.9%. Jadi total keragaman yang mampu dijelaskan dua komponen utama adalah 97.5%.

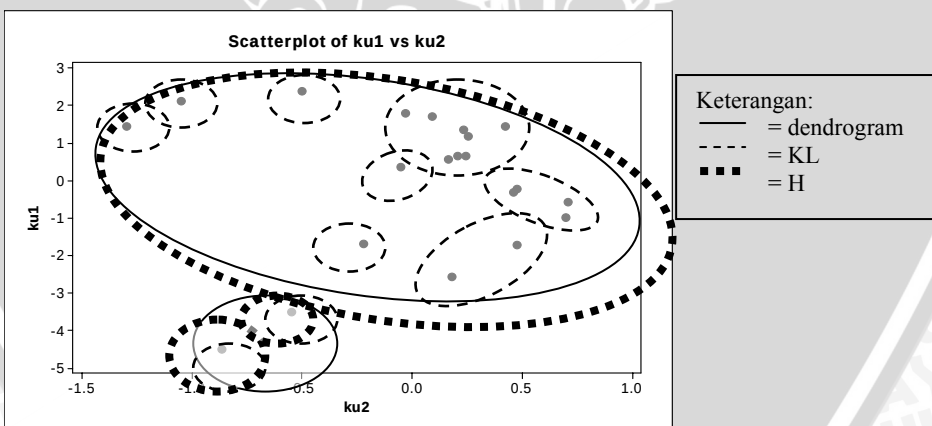
Skor komponen utama yang diperoleh akan digunakan untuk analisis *cluster* dan hasil analisis *cluster* adalah dendrogram pada Gambar 4.4.



Gambar 4.4. Dendrogram Data 2

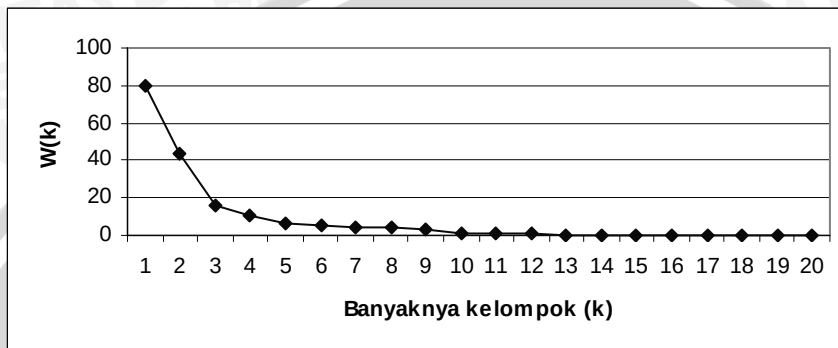
Berdasarkan dendrogram pada Gambar 4.4, banyaknya *cluster* yang terbentuk adalah dua *cluster*.

Diagram pencar serta *cluster* yang terbentuk berdasarkan dendrogram dan indeks validitas Data 2 adalah seperti pada Gambar 4.5. Terlihat bahwa dendrogram dan indeks validitas menghasilkan *cluster* optimal yang berbeda.



Gambar 4.5. Diagram pencar (*scatterplot*) dan perbandingan *cluster* yang terbentuk pada Data 2

Untuk menentukan banyaknya *cluster* optimal menurut indeks validitas CH, H dan KL, harus dihitung $W(k)$ terlebih dahulu. Plot antara banyaknya *cluster* dan $W(k)$ dapat dilihat pada Gambar 4.6.



Gambar 4.6. Plot antara banyaknya *cluster* dan $W(k)$ Data 2

Gambar 4.6 memperlihatkan bahwa $W(k)$ mulai mendatar ketika k bernilai 3, berarti banyaknya *cluster* optimal kemungkinan adalah tiga. Hal ini berbeda dengan penentuan banyaknya *cluster* dengan dendrogram yang menghasilkan dua *cluster* (Gambar 4.4). Dari Gambar 4.6 terlihat bahwa $W(k)$ belum mendatar ketika k bernilai 2. Hal ini berarti penentuan banyaknya *cluster* dengan dendrogram tidak peka terhadap perubahan nilai $W(k)$.

Untuk menentukan banyaknya *cluster* optimal harus dihitung indeks validitas terlebih dahulu. Nilai seluruh indeks validitas ditampilkan pada Lampiran 5. Banyaknya *cluster* optimal berdasarkan ketiga indeks validitas ditampilkan pada Tabel 4.3.

Tabel 4.3. Banyaknya *cluster* optimal tiga indeks validitas Data 2

Indeks Validitas	Banyaknya Cluster Optimal
CH	20
H	3
KL	10

Indeks H menghasilkan tiga *cluster* optimal. Hal ini sesuai dengan Gambar 4.6. Pada saat $W(k)$ mulai mendatar, rasio antara $W(k)$ dan $W(k+1)$ kecil sehingga indeks H yang dihasilkan juga kecil. Rasio antara $W(3)$ dan $W(4)$ adalah 1.435.

Indeks KL menghasilkan sepuluh *cluster* optimal. Hal ini disebabkan nilai $W(k)$ menurun cukup banyak ketika k bernilai 10. Dari nilai $W(k)$ pada Lampiran 5 dapat dilihat bahwa selisih antara $W(9)$ dan $W(10)$, yaitu 2.075842, jauh lebih besar dibandingkan selisih antara $W(10)$ dan $W(11)$, yaitu 0.282463. Rasio keduanya bernilai 7.349. Artinya, ketika banyaknya *cluster* bertambah dari 9 menjadi 10, terjadi penurunan nilai $W(k)$ tujuh kali lipat lebih tinggi dibandingkan dengan penurunan nilai $W(k)$ ketika banyaknya *cluster* bertambah dari 10 menjadi 11. Hal ini menghasilkan nilai indeks KL terbesar untuk Data 2. Jadi, banyaknya *cluster* optimal menurut indeks KL adalah sepuluh.

Banyaknya *cluster* optimal berdasarkan indeks CH merupakan banyaknya *cluster* maksimal yang dapat dibentuk, yaitu 20 *cluster*. Hal ini disebabkan pada saat banyaknya *cluster* mencapai maksimal, keragaman antar *cluster* akan maksimal dan keragaman dalam *cluster* akan minimal. Berdasarkan indeks ini, analisis *cluster* seakan-akan menjadi tidak berguna. Pada Gambar 4.5 dan Gambar 4.6 dapat dilihat bahwa tidak tepat jika *cluster* yang terbentuk sebanyak 20 *cluster*. Dapat disimpulkan bahwa indeks CH tidak cocok untuk menentukan banyaknya *cluster* optimal pada Data 2.

Nilai CTM ketiga indeks validitas dapat dilihat pada Tabel 4.4.

Tabel 4.4. Nilai CTM untuk tiga indeks validitas Data 2

Indeks Validitas	CTM
CH	0.000829
H	0.424790
KL	0.161172

Nilai CTM pada Tabel 4.4 memperlihatkan bahwa indeks CH menghasilkan CTM terkecil. Tapi, banyaknya *cluster* optimal menurut indeks CH tidak cocok untuk digunakan pada Data 2 karena tidak sesuai dengan Gambar 4.6.

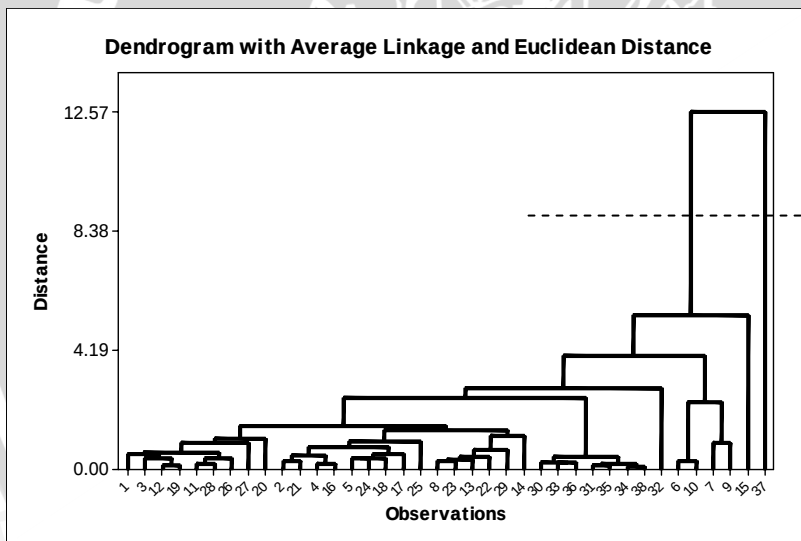
Dari hasil ini dapat diperkirakan bahwa jika nilai CTM digunakan untuk membandingkan hasil analisis *cluster* pada data dengan banyaknya *cluster* optimal yang jauh berbeda, maka kesimpulan yang dihasilkan bisa menjadi tidak tepat. Jika indeks CH dihilangkan dan nilai CTM yang dibandingkan adalah indeks H dan KL, maka dapat dilihat bahwa indeks KL lebih baik dari indeks H.

4.3. Data Kabupaten/Kota di Propinsi Jawa Timur Berdasarkan Penduduk Usia 10 Tahun Ke Atas Menurut Kabupaten/Kota dan Status Pendidikan (Data 3)

Peubah-peubah pada Data 3 tidak memiliki satuan yang sama, jadi perlu dilakukan pembakuan data. Hasil pembakuan data dapat dilihat pada Lampiran 2.

Pada Lampiran 3 terlihat bahwa terdapat multikolinieritas pada peubah Data 3. Oleh karena itu, perlu dilakukan analisis komponen utama sebelum analisis *cluster*. Hasil analisis komponen utama ditampilkan pada Lampiran 4. Komponen utama pertama mampu menjelaskan keragaman sebesar 71.9% dan komponen utama kedua menjelaskan 18.9%. Jadi, total keragaman yang mampu dijelaskan dua komponen utama adalah 90.8%.

Dendrogram hasil analisis *cluster* Data 3 dapat dilihat pada Gambar 4.7.

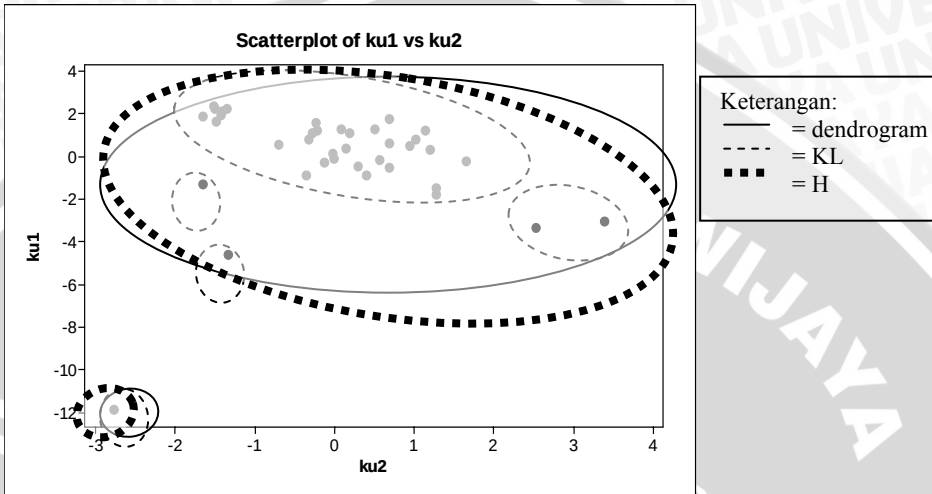


Gambar 4.7. Dendrogram Data 3

Berdasarkan dendrogram pada Gambar 4.7, terbentuk dua *cluster*.

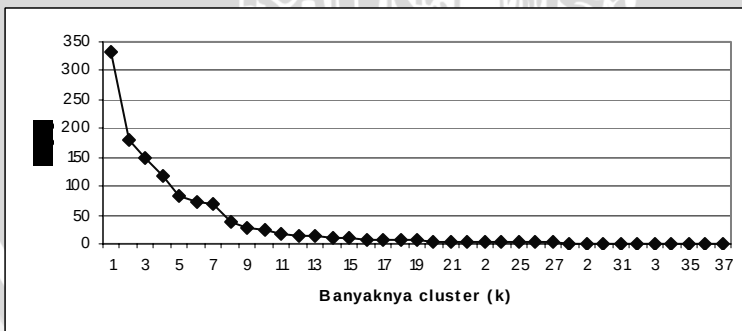
Diagram pencar serta *cluster* yang terbentuk berdasarkan dendrogram dan indeks validitas Data 3 adalah seperti pada Gambar

4.8. Terlihat bahwa dendrogram dan indeks validitas menghasilkan *cluster* optimal yang berbeda.



Gambar 4.8. Diagram pencar (*scatterplot*) dan perbandingan *cluster* yang terbentuk pada Data 3

Untuk menentukan banyaknya *cluster* optimal menurut indeks validitas CH, H dan KL, harus dihitung $W(k)$ terlebih dahulu. Nilai $W(k)$ ditampilkan pada Lampiran 5 dan plot antara banyaknya *cluster* dan $W(k)$ ditampilkan pada Gambar 4.9.



Gambar 4.9. Plot antara banyaknya *cluster* dan $W(k)$ Data 3

Dari Gambar 4.9 terlihat bahwa $W(k)$ mulai mendatar ketika k bernilai 8, berarti kemungkinan banyaknya *cluster* optimal adalah delapan *cluster*. Hal ini berbeda dengan penentuan banyaknya *cluster* dengan dendrogram seperti pada Gambar 4.7 yang menghasilkan dua *cluster*. Gambar 4.9 memperlihatkan bahwa $W(k)$ belum mendatar ketika k bernilai 2, berarti penentuan banyaknya *cluster* dengan dendrogram tidak peka terhadap perubahan nilai $W(k)$.

Untuk menentukan banyaknya *cluster* optimal, harus dihitung indeks validitas terlebih dahulu. Nilai seluruh indeks validitas ditampilkan pada Lampiran 5. Banyaknya *cluster* optimal berdasarkan ketiga indeks validitas ditampilkan pada Tabel 4.5.

Tabel 4.5. Banyaknya *cluster* optimal tiga indeks validitas Data 3

Indeks Validitas	Banyaknya <i>Cluster</i> Optimal
CH	37
H	2
KL	5

Indeks H menghasilkan dua *cluster* optimal. Penyebabnya adalah ketika k bergerak dari 1 ke 2 nilai $W(k)$ menurun cukup banyak dan ketika k bergerak dari 2 ke 3 nilai $W(k)$ hanya menurun sedikit. Jika dihitung, rasio antara $W(2)$ dan $W(3)$ adalah 1.2064. Pada saat k bernilai 2, rasio antara $W(2)$ dan $W(3)$ mengecil sehingga indeks H yang dihasilkan juga kecil. Seperti pada Data 1 dan Data 2, indeks H selalu menghasilkan sesedikit mungkin *cluster* optimal.

Indeks KL menghasilkan lima *cluster* optimal. Hal ini disebabkan ketika k bertambah dari 4 ke 5 nilai $W(k)$ menurun cukup banyak. Dari nilai $W(k)$ pada Lampiran 5 terlihat bahwa selisih antara $W(4)$ dan $W(5)$ jauh lebih besar dibandingkan selisih antara $W(5)$ dan $W(6)$, yaitu 36.9732 dan 7.471. Rasio selisih $W(4)$ dan $W(5)$ dengan $W(5)$ dan $W(6)$ adalah 4.9489 dan nilai ini menghasilkan nilai indeks KL terbesar seperti yang terlihat pada Lampiran 5. Penurunan nilai $W(k)$ menunjukkan penurunan keragaman dalam *cluster* setelah dibentuk *cluster* baru. Jadi, banyaknya *cluster* optimal menurut indeks KL adalah 5.

Pada data ini indeks CH tetap menghasilkan *cluster* optimal yang maksimal, yaitu 20 *cluster*. Hal ini tidak sesuai dengan Gambar 4.8 dan Gambar 4.9. Indeks CH menghasilkan *cluster* optimal yang menunjukkan seakan-akan analisis *cluster* menjadi tidak berguna.

Pada Gambar 4.8 dan Gambar 4.9 dapat dilihat bahwa tidak tepat jika *cluster* yang terbentuk adalah 20 *cluster*. Untuk data ini dapat disimpulkan bahwa indeks CH tidak cocok untuk menentukan banyaknya *cluster* optimal.

Nilai CTM ketiga indeks validitas dapat dilihat pada Tabel 4.6.

Tabel 4.6. Nilai CTM untuk tiga indeks validitas Data 3

Indeks Validitas	CTM
CH	0.0012368
H	0.3636676
KL	0.0200026

Nilai CTM pada Tabel 4.6 memperlihatkan bahwa indeks CH tetap menghasilkan CTM terkecil. Tapi, banyaknya *cluster* optimal menurut indeks CH tidak cocok untuk digunakan pada Data 3. Jadi nilai CTM untuk indeks CH tidak perlu diperhatikan.

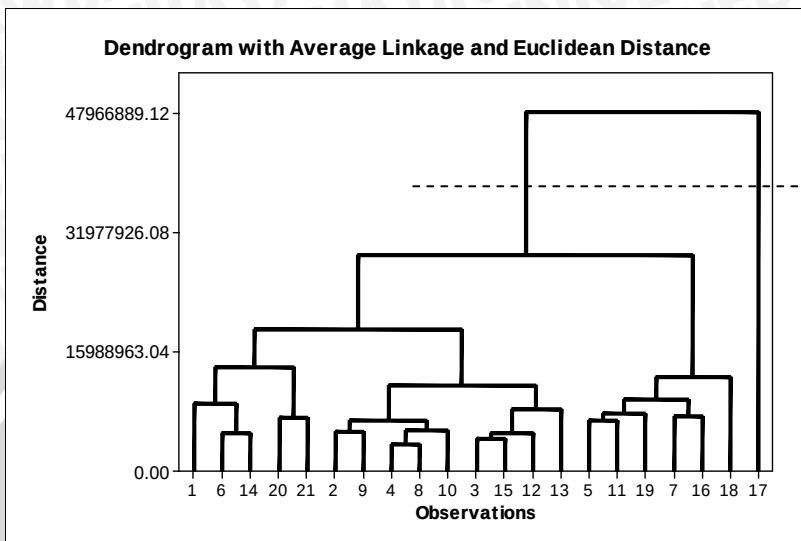
Dari nilai CTM tiga indeks validitas dapat diperkirakan bahwa jika CTM digunakan untuk membandingkan hasil analisis *cluster* pada data dengan banyaknya *cluster* optimal yang jauh berbeda maka kesimpulan yang dihasilkan bisa menjadi tidak tepat. Jika indeks CH dihilangkan dan nilai CTM yang dibandingkan adalah indeks H dan KL, maka terlihat bahwa indeks KL lebih baik dari indeks H.

4.4. Data PDRB Beberapa Sub Sektor Kabupaten Ponorogo (Data 4)

Peubah-peubah Data 4 memiliki satuan yang sama, jadi tidak perlu dilakukan pembakuan data. Data 4 ditampilkan pada Lampiran 1.

Pengujian multikolinieritas memperlihatkan bahwa tidak terdapat multikolinieritas antar peubah Data 4, sehingga data asli dapat digunakan untuk analisis *cluster*. Pengujian multikolinieritas dapat dilihat pada Lampiran 3.

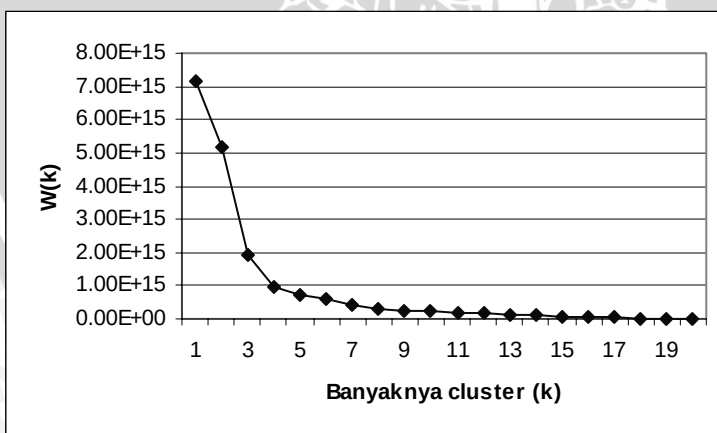
Dendrogram hasil analisis *cluster* Data 4 diperlihatkan pada Gambar 4.10.



Gambar 4.10. Dendrogram Data 4

Berdasarkan dendrogram pada Gambar 4.10, ada dua *cluster* yang terbentuk.

Untuk menentukan banyaknya *cluster* optimal menurut indeks validitas CH, H dan KL, harus dihitung $W(k)$ terlebih dahulu. Nilai $W(k)$ ditampilkan pada Lampiran 5. Plot antara banyaknya *cluster* dan $W(k)$ ditampilkan pada Gambar 4.11.



Gambar 4.11. Plot antara banyaknya *cluster* dan $W(k)$ Data 4

Gambar 4.11 memperlihatkan bahwa $W(k)$ mulai mendatar ketika k bernilai 4, berarti kemungkinan banyaknya *cluster* optimal adalah empat. Hal ini berbeda dengan penentuan banyaknya *cluster* dengan dendrogram pada Gambar 4.10. yang menghasilkan dua *cluster*. Gambar 4.11 juga memperlihatkan bahwa $W(k)$ belum mendatar ketika k bernilai 2. Ini berarti penentuan banyaknya *cluster* dengan dendrogram tidak peka terhadap perubahan nilai $W(k)$.

Untuk menentukan banyaknya *cluster* optimal harus dihitung indeks validitas terlebih dahulu. Nilai seluruh indeks validitas terlampir pada Lampiran 5. Banyaknya *cluster* optimal berdasarkan ketiga indeks validitas disajikan pada Tabel 4.7.

Tabel 4.7. Banyaknya *cluster* optimal tiga indeks validitas Data 4

Indeks Validitas	Banyaknya <i>Cluster</i> Optimal
CH	20
H	1
KL	4

Indeks H menghasilkan satu *cluster* optimal. Penyebabnya adalah ketika k bernilai 1, rasio antara $W(1)$ dan $W(2)$ kecil, yaitu 1.3829, sehingga indeks H yang dihasilkan juga kecil. Hal ini berbeda dengan Gambar 4.11 yang menunjukkan bahwa banyaknya *cluster* optimal kemungkinan adalah 4.

Indeks KL menghasilkan empat *cluster* optimal, sesuai dengan Gambar 4.11. Hal ini disebabkan ketika k bertambah dari 3 ke 4 nilai $W(k)$ menurun cukup banyak. Nilai $W(k)$ pada Lampiran 5 memperlihatkan bahwa selisih antara $W(3)$ dan $W(4)$ adalah 9.96×10^{14} . Nilai ini jauh lebih besar dibandingkan selisih antara $W(4)$ dan $W(5)$, yaitu sebesar 2.03×10^{14} . Rasio penurunan kedua selisih adalah 4.9064 dan rasio ini menghasilkan nilai KL terbesar. Jadi banyaknya *cluster* optimal menurut indeks KL adalah 4.

Pada Data 4, indeks CH tetap menghasilkan *cluster* optimal sebanyak *cluster* maksimal yang dapat dibentuk, yaitu 20. Hal ini disebabkan pada saat banyaknya *cluster* maksimal, keragaman antar *cluster* menjadi maksimal dan keragaman dalam *cluster* menjadi minimal. Akan tetapi, hal ini menunjukkan seakan-akan analisis *cluster* menjadi tidak berguna. Banyaknya *cluster* optimal menurut indeks CH tidak sesuai dengan Gambar 4.11 yang menunjukkan bahwa $W(k)$ mulai mendatar ketika k bernilai 4.

Nilai CTM ketiga indeks validitas dapat dilihat pada Tabel 4.8.

Tabel 4.8. Nilai CTM untuk tiga indeks validitas Data 4

Indeks Validitas	CTM
CH	0.01245
H	1
KL	0.56325

Tabel 4.8 memperlihatkan bahwa indeks CH menghasilkan CTM terkecil. Tapi, banyaknya *cluster* optimal menurut indeks CH tidak sesuai untuk digunakan pada Data 4.

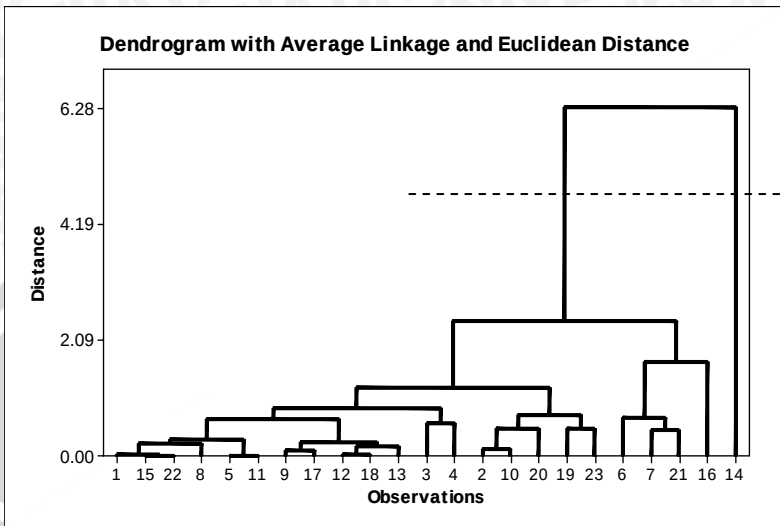
Dari hasil ini dapat diperkirakan bahwa jika nilai CTM digunakan untuk membandingkan hasil analisis *cluster* pada data dengan banyaknya *cluster* optimal yang jauh berbeda, maka kesimpulan yang dihasilkan bisa menjadi tidak tepat. Jika indeks CH dihilangkan dan nilai CTM yang dibandingkan adalah antara indeks H dan KL, maka terlihat bahwa indeks KL lebih baik dari indeks H.

4.5. Data Industri Tempe di Kota Malang (Data 5)

Peubah-peubah pada Data 5 memiliki satuan tidak sama, sehingga perlu dilakukan pembakuan data. Hasil pembakuan data dilampirkan pada Lampiran 2.

Pengujian multikolinieritas memperlihatkan bahwa terdapat multikolinieritas pada peubah Data 5. Hasil pengujian multikolinieritas dapat dilihat pada Lampiran 3. Oleh karena itu, perlu dilakukan analisis komponen utama. Komponen utama pertama mampu menjelaskan keragaman sebesar 93.8% dan komponen utama kedua menjelaskan 6.2%. Jadi total keragaman yang mampu dijelaskan dua komponen utama adalah 100%.

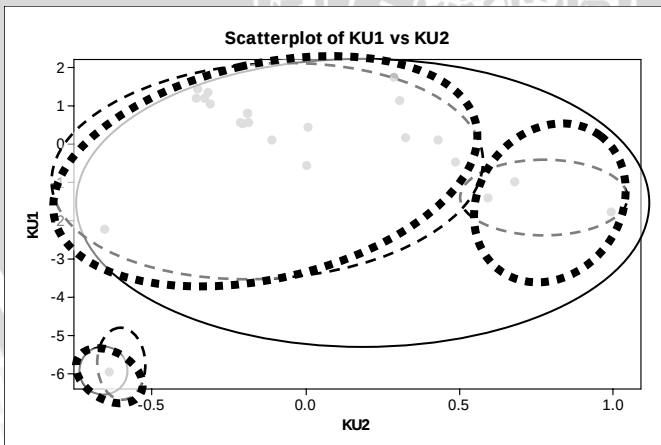
Skor komponen utama digunakan untuk analisis *cluster* dan hasil dari analisis *cluster* adalah dendrogram seperti pada Gambar 4.12.



Gambar 4.12. Dendrogram Data 5

Berdasarkan dendrogram pada Gambar 4.12, banyaknya *cluster* yang terbentuk adalah dua *cluster*.

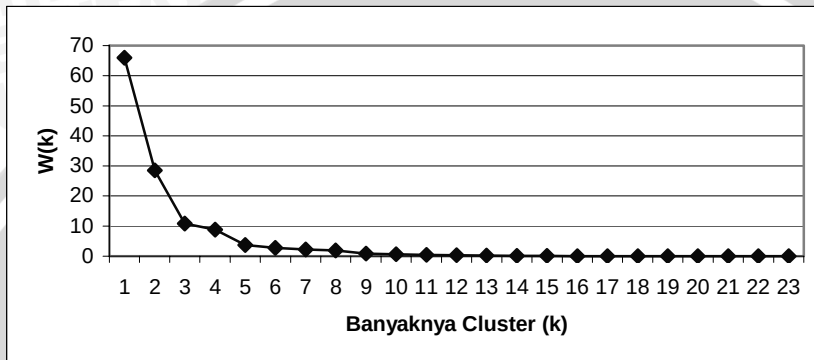
Diagram pencar serta *cluster* yang terbentuk berdasarkan dendrogram dan indeks validitas Data 4 adalah seperti pada Gambar 4.13. Terlihat bahwa dendrogram dan indeks validitas menghasilkan *cluster* optimal yang berbeda.



Keterangan:
 — = dendrogram
 - - - = KL
 . . . = H

Gambar 4.13. Diagram pencar (*scatterplot*) dan perbandingan *cluster* yang terbentuk pada Data 5

Untuk menentukan banyaknya *cluster* optimal menurut indeks validitas CH, H dan KL, harus dihitung $W(k)$ terlebih dahulu. Nilai $W(k)$ ditampilkan pada Lampiran 5. Plot antara banyaknya *cluster* dan $W(k)$ dapat dilihat pada Gambar 4.14.



Gambar 4.14. Plot antara banyaknya *cluster* dan $W(k)$ Data 5

Gambar 4.14 memperlihatkan bahwa $W(k)$ mulai mendatar ketika k bernilai 3, berarti kemungkinan banyaknya *cluster* optimal adalah 3. Hal ini berbeda dengan penentuan banyaknya *cluster* dengan dendrogram, seperti pada Gambar 4.12, yang menghasilkan dua *cluster*. Dari Gambar 4.14 dapat dilihat bahwa $W(k)$ belum mendatar ketika k bernilai 2. Hal ini berarti penentuan banyaknya *cluster* dengan dendrogram tidak peka terhadap perubahan nilai $W(k)$.

Untuk menentukan banyaknya *cluster* optimal harus dihitung indeks validitas terlebih dahulu. Nilai seluruh indeks validitas dilampirkan pada Lampiran 5. Banyaknya *cluster* optimal berdasarkan tiga indeks validitas ditampilkan pada Tabel 4.9.

Tabel 4.9. Banyaknya *cluster* optimal tiga indeks validitas Data 5

Indeks Validitas	Banyaknya Cluster Optimal
CH	22
H	3
KL	3

Indeks H menghasilkan tiga *cluster* optimal. Hal ini sesuai dengan Gambar 4.14. Ketika k bernilai 3, rasio antara $W(3)$ dan $W(4)$

kecil, yaitu 1.2326, sehingga indeks H yang dihasilkan juga kecil. Hal ini terlihat pada Lampiran 5.

Indeks KL menghasilkan tiga *cluster* optimal, sesuai dengan Gambar 4.14. Nilai $W(k)$ pada Lampiran 5 memperlihatkan bahwa selisih antara $W(2)$ dan $W(3)$ adalah 17.6727, jauh lebih besar dibandingkan selisih antara $W(3)$ dan $W(4)$, yaitu 2.04246. Hal ini berarti ketika banyaknya *cluster* bertambah dari 2 menjadi 3, penurunan nilai $W(k)$ lebih tinggi daripada penurunan nilai $W(k)$ ketika banyaknya *cluster* bertambah dari 3 menjadi 4. Rasio penurunan kedua selisih adalah 8.6526 dan rasio ini menghasilkan nilai KL terbesar. Jadi, banyaknya *cluster* optimal menurut indeks KL adalah tiga.

Pada Data 5, indeks CH menghasilkan 22 *cluster* optimal. Hal ini menunjukkan seakan-akan analisis *cluster* menjadi tidak berguna. Banyaknya *cluster* optimal menurut indeks CH tidak sesuai dengan Gambar 4.14, yang menunjukkan bahwa $W(k)$ mulai mendatar ketika k bernilai 3.

Nilai CTM tiga indeks validitas ditampilkan pada Tabel 4.10.

Tabel 4.10. Nilai CTM untuk tiga indeks validitas Data 5

Indeks Validitas	CTM
CH	0.0001777
H	0.2805400
KL	0.2805400

Berdasarkan nilai CTM terlihat bahwa indeks CH menghasilkan CTM terkecil. Tapi, jika banyaknya *cluster* optimal menurut indeks CH dibandingkan dengan Gambar 4.13, maka banyaknya *cluster* optimal menurut indeks CH kurang cocok digunakan.

Dari hasil ini dapat diperkirakan bahwa jika nilai CTM digunakan untuk membandingkan hasil analisis *cluster* pada data dengan banyaknya *cluster* optimal yang jauh berbeda, maka kesimpulan yang dihasilkan bisa tidak tepat. Jika indeks CH dihilangkan dan nilai CTM yang dibandingkan adalah indeks H dan KL, maka terlihat bahwa indeks KL sama baik dengan indeks H.

Banyaknya *cluster* optimal dan CTM yang diperoleh dari lima data yang telah dianalisis ditampilkan pada Tabel 4.11.

Tabel 4.11. Banyaknya *cluster* optimal dan nilai CTM lima data

Data	Banyaknya <i>Cluster</i> Optimal			CTM		
	CH	H	KL	CH	H	KL
1	24	2	5	0.0004700	1.0454910	0.2851450
2	20	3	10	0.0008290	0.4247900	0.1611720
3	37	2	5	0.0012368	0.3636676	0.0200026
4	20	1	4	0.0124500	1	0.5632500
5	22	3	3	0.0001777	0.2805400	0.2805400

Tabel 4.11 memperlihatkan bahwa indeks H selalu menghasilkan banyaknya *cluster* optimal yang paling sedikit. Hal ini berarti jika seorang peneliti ingin membentuk *cluster* sesedikit mungkin maka sebaiknya peneliti tersebut menggunakan indeks H.

Berdasarkan lima data yang telah diteliti, terlihat bahwa indeks KL merupakan indeks yang paling sering menghasilkan banyaknya *cluster* optimal yang mendekati banyaknya *cluster* optimal berdasarkan plot $W(k)$. Indeks KL dapat mendeteksi rasio selisih $W(k)$ yang paling besar. Jadi indeks ini sensitif terhadap perubahan nilai $W(k)$. Berdasarkan nilai CTM, indeks ini hampir selalu menghasilkan CTM yang lebih kecil daripada indeks H. Jadi dapat disimpulkan bahwa indeks KL lebih baik dari indeks H.

Pada lima data yang dianalisis, indeks CH selalu menghasilkan banyaknya *cluster* maksimal. Hal ini sesuai dengan rumus indeks CH yang merupakan rasio antara $B(k)$ dan $W(k)$, di mana nilai $B(k)$ cenderung naik (untuk data yang tidak menggerombol dengan jelas) dan nilai $W(k)$ selalu turun. Hal ini menyebabkan semakin banyak *cluster* yang terbentuk, semakin besar nilai indeks CH dan indeks CH cenderung selalu menghasilkan *cluster* yang sangat banyak. Hal ini menyebabkan analisis *cluster* seakan-akan menjadi tidak berguna. Menurut penelitian Milligan dan Cooper (1985), indeks CH hanya cocok digunakan untuk data yang menggerombol dengan jelas. Menurut penelitian lain oleh Heer dan Chi (2008) disimpulkan bahwa indeks CH tidak mampu menentukan banyaknya *cluster* yang sesuai untuk data yang tidak membentuk *cluster* dengan jelas. Karena data yang digunakan dalam penelitian ini tidak menggerombol dengan jelas, indeks CH selalu menghasilkan *cluster* yang sangat banyak.

Dari Tabel 4.11, dapat disimpulkan bahwa untuk data yang tidak menggerombol dengan jelas, banyaknya *cluster* optimal H lebih kecil atau sama dengan *cluster* optimal KL lebih kecil dari *cluster* optimal CH ($H \leq KL < CH$).

Berdasarkan nilai CTM pada Tabel 4.11, terlihat bahwa indeks CH selalu menghasilkan CTM terkecil. Hal ini disebabkan karena indeks CH selalu menghasilkan *cluster* yang maksimal sehingga dapat menurunkan keragaman *cluster-cluster* yang terbentuk. Hal ini sesuai dengan rumus CTM yaitu rasio antara simpangan baku dibagi banyaknya peubah dan banyaknya *cluster* yang terbentuk. Semakin banyak *cluster* yang terbentuk maka simpangan baku *cluster* akan semakin kecil sehingga nilai CTM juga semakin kecil. Meskipun indeks CH selalu menghasilkan CTM terkecil, banyaknya *cluster* menurut indeks CH terlalu banyak sehingga analisis *cluster* yang dilakukan menjadi tidak berguna. Ini berarti CTM kurang tepat digunakan untuk membandingkan indeks validitas pada data yang menghasilkan banyaknya *cluster* optimal yang jauh berbeda. Jika nilai CTM yang dibandingkan hanya antara indeks H dan indeks KL, dapat disimpulkan bahwa indeks KL lebih baik dari indeks H.

Dari lima data yang diteliti dapat disimpulkan bahwa penentuan banyaknya *cluster* menggunakan dendrogram tidak peka terhadap perubahan keragaman dalam *cluster*. Tidak seperti indeks validitas H dan KL yang menentukan banyaknya *cluster* optimal berdasarkan $W(k)$. Jadi, dapat disimpulkan bahwa indeks validitas lebih peka terhadap perubahan keragaman dalam *cluster* untuk menentukan banyaknya *cluster* optimal.

UNIVERSITAS BRAWIJAYA



BAB V

KESIMPULAN dan SARAN

5.1. Kesimpulan

Dari penelitian ini dapat disimpulkan bahwa:

1. Banyaknya *cluster* optimal berdasarkan indeks H lebih kecil atau sama dengan *cluster* optimal berdasarkan indeks KL dan lebih kecil dari *cluster* optimal berdasarkan indeks CH ($H \leq KL < CH$). Indeks CH selalu menghasilkan banyaknya *cluster* yang maksimal untuk data yang tidak menggerombol dengan jelas.
2. Indeks CH selalu menghasilkan CTM terkecil. Akan tetapi indeks ini kurang cocok digunakan pada data yang tidak menggerombol dengan jelas, bahkan indeks ini menghasilkan banyaknya *cluster* optimal yang lebih banyak daripada dendrogram. Jika yang dibandingkan adalah nilai CTM indeks H dan KL, maka indeks KL adalah indeks terbaik.

5.2. Saran

Untuk peneliti yang ingin memperoleh sedikit *cluster* disarankan agar menggunakan indeks H dalam menentukan banyaknya *cluster* optimal. Peneliti yang ingin memperoleh banyaknya *cluster* optimal berdasarkan rasio penurunan $W(k)$ disarankan untuk menggunakan indeks KL.

UNIVERSITAS BRAWIJAYA



DAFTAR PUSTAKA

- Afifi, A.A. dan V. Clark. 1990. *Computer-Aided Multivariate Analysis. Second Edition*. Chapman & Hall. New York.
- Anggraini, R. 2006. *Perbandingan Uji Nilai Statistik dan Uji Validitas Untuk Menentukan Kelompok Optimal Dalam Analisis Kelompok Hirarki*. Fakultas MIPA Universitas Brawijaya. Malang. Skripsi. Tidak dipublikasikan.
- Bolshakova, N. dan F. Azuaje. 2008. *Cluster Validation Techniques for Genome Expression Data*. <https://www.cs.tcd.ie/publications/tech-reports/reports.02/TCD-CS-2002-33.pdf>. Tanggal akses: 12 Mei 2008.
- Cahyono, D.H. 2005. *Laporan Praktek Kerja Lapang di Badan Pusat Statistik Kota Malang*. Fakultas MIPA Universitas Brawijaya. Malang. Tidak dipublikasikan.
- Calinski, R.B. dan J. Harabasz. 1974. *A Dendrite Method for Cluster Analysis. Communications in Statistics*. Vol. 3: 1-27.
- Dillon, W.R. dan M. Goldstein. 1984. *Multivariate Analysis: Methods & Applications*. John Wiley & Sons. New York.
- Draper, N.R. dan H. Smith. 1992. *Analisis Regresi Terapan. Edisi Kedua*. Percetakan PT Gramedia. Jakarta
- Epps, J. dan E. Ambikairajah. 2008. *Visualisation of Reduced-Dimension Microarray Data Using Gaussian Mixture Model*. <http://crpit.com/confpapers/CRPITV45Epps.pdf>. Tanggal akses: 28 April 2008.
- Everitt, B.S. 1980. *Cluster Analysis. Second Edition*. Heineman Educational Books Ltd. London.
- Ferlya, Y. 2005. *Laporan Praktek Kerja Lapang di Badan Pusat Statistik Kabupaten Ponorogo*. Fakultas MIPA Universitas Brawijaya. Malang. Tidak dipublikasikan.

- Gasperz, V. 1992. *Teknik Analisis dalam Penelitian Percobaan*. Tarsito. Bandung.
- Hair, J.F., R.E. Anderson, R.L. Tatham dan W.C. Black. 1998. *Multivariate Data Analysis with Readings. Fifth Edition*. Prentice-Hall, Inc. New Jersey.
- Hartigan, J.A. 1975. *Clustering Algorithms*. John Wiley and Sons, Inc. New York.
- Heer, J. dan Chi, E.H. 2008. *Mining the Structure of User Activity using cluster stability*. <http://www2.parc.com/istl/groups/uir/publications/items/UIR-2002-02-Heer-SIAM-ClusterStability.pdf>. Tanggal akses: 3 Maret 2008.
- Jain, A.K. dan R.C. Dubes. 1988. *Algorithms for Clustering Data*. Prentice-Hall, Inc. New Jersey.
- Jäkel, J. dan M. Nöllenburg. 2008. *Validation in the Cluster Analysis of Gene Expression Data*. <http://i11www.iti.uni-karlsruhe.de/members/noelle/pub/jn-vcage-04.pdf>. Tanggal akses: 3 Maret 2008.
- Johnson, R.A. dan D.W. Wichern. 2002. *Applied Multivariate Statistical Analysis*. Prentice-Hall, Inc. New Jersey.
- Krzanowski, W.J. dan Y.T. Lai. 1985. *A Criterion for Determining the Number of Groups in A Data Set Using Sum of Squares Clustering*. *Biometrics*. Vol. 44: 23-44.
- Mardia, K.V., J.T. Kent dan J.M. Bibby. 1979. *Multivariate Analysis*. Academic Press. New York.
- Mendra, M.L. 2007. *Analisis Faktor-Faktor Sosial Ekonomi yang Berpengaruh terhadap Produksi dan Pendapatan Usaha Tani Brokoli Organik (Kasus di Kecamatan Pujon, Kabupaten Malang)*. Fakultas Pertanian Universitas Brawijaya. Malang. Skripsi. Tidak dipublikasikan.

Milligan, G.W. dan M.C. Cooper. 1985. *An Examination of Procedures of Determining the Number of Cluster in A Data Set*. *Psychometrika*. Vol. 50: 159-179.

Mufidah, R. 2004. *Analisis Kelompok Optimal dengan Kombinasi Indeks Validitas pada Analisis Kelompok Hirarki*. Fakultas MIPA Universitas Brawijaya. Malang. Skripsi. Tidak dipublikasikan.

Rohman, F. 2001. *Manajemen Reproduksi dan Pelaksanaan IB serta Evaluasi Kualitas Semen Sapi Potong Sebelum Pelaksanaan IB di Kabupaten Lamongan*. Fakultas Peternakan Universitas Brawijaya. Malang. Skripsi. Tidak dipublikasikan.

Sharma, S. 1996. *Applied Multivariate Techniques*. John Wiley and Sons, Inc. New York.

Tibshirani, W., G. Walther dan T. Hastie. 2000. *Estimating The Number of Clusters in A Dataset via The Gap Statistics*. Technical Report 208 Departement of Statistics Stanford University. Stanford.

Wulandari, T. 2006. *Penerapan Analisis Cluster Hirarki pada Data Berskala Campuran dengan Menggunakan Analisis Komponen Utama Nonlinier (Studi Kasus di Sentra Industri Tempe Kota Malang)*. Fakultas MIPA Universitas Brawijaya. Malang. Skripsi. Tidak dipublikasikan.

UNIVERSITAS BRAWIJAYA



Lampiran 1. Data

Data 1. Data Usaha Ternak Sapi di Kabupaten Lamongan

No.	Kecamatan	X1	X2
1	Lamongan	1.438	36,26
2	Deket	75	1,87
3	Turi	844	17,33
4	Tikung	3.230	32,07
5	Kembangbuh	4.653	72,57
6	Sukodadi	1.864	40,09
7	Pucuk	619	18,22
8	Krg Geneng	1.839	50,27
9	Sekaran	2.237	28,03
10	Babat	2.247	35,97
11	Kdngpring	2.290	27,08
12	Sugio	4.045	42,83
13	Modo	3.317	42,75
14	Ngimbang	3.262	36,63
15	Bluluk	1.659	38,23
16	Sukorame	1.537	48,82
17	Sambeng	4.676	104,91
18	Mantup	4.420	47,49
19	Krgbnangun	62	1,44
20	Glagah	590	12,21
21	Kalitengah	1.143	34,16
22	Paciran	2.188	28,57
23	Selokuro	5.801	56,71
24	Brondong	2.992	42,66
25	Laren	1.718	19,96

Keterangan:

X1 : Jumlah ternak (ekor)

X2 : Kepadatan populasi sapi (ekor/km²)

Lampiran 1. (Lanjutan)

Data 2. Data Pengusaha Tani Brokoli Organik di Kecamatan Pujon

No.	X1	X2	X3	X4
1	0,200	52.575,933	772.500	6.864.000
2	0,150	42.235,35	592.000	4.921.875
3	0,050	35.107,688	560.000	1.363.250
4	0,100	41.950,125	592.000	3.505.600
5	0,300	54.165,086	737.750	12.600.000
6	0,350	68.107,29	790.000	12.871.950
7	0,500	74.342,66	860.000	24.885.000
8	0,400	68.868,19	819.500	20.960.000
9	0,150	44.939,29	712.000	5.817.000
10	0,250	60.793,37	781.250	10.251.500
11	0,050	37.933,44	565.000	2.077.500
12	0,100	42.985,39	592.000	3.347.400
13	0,025	31.671	500.000	1.112.000
14	0,010	34.977,3	580.000	1.391.875
15	0,050	25.520,33	450.000	10.156.000
16	0,150	46.284,9	712.000	6.300.000
17	0,100	43.399,84	590.000	4.181.100
18	0,150	49.703,7	750.000	6.313.650
19	0,025	31.807,24	525.000	1.100.000
20	0,010	24.318,75	425.000	1.126.000
21	0,150	20.414,22	400.000	906.528

Keterangan:

X1 : Luas lahan (Ha)

X2 : Bunga modal (Rp)

X3 : Tenaga kerja (Rp/hok)

X4 : Total penerimaan (Rp)

Lampiran 1. (Lanjutan)

Data 3. Data Kabupaten/Kota di Propinsi Jawa Timur Berdasarkan Penduduk Usia 10 Tahun Ke Atas Menurut Kabupaten/Kota dan Status Pendidikan (Dalam Satuan Orang)

Kabupaten/Kota	X1	X2	X3	X4	X5	X6	X7	X8	X9
Kab. Pacitan	168.294	188.613	69.528	11.422	13.512	1.872	841	4.667	217
Kab. Ponorogo	284.607	222.834	123.672	54.384	32.055	3.429	3.030	15.987	897
Kab. Trenggalek	162.356	235.528	93.348	33.557	21.982	2.558	3.082	12.927	574
Kab. Tulungagung	234.592	292.702	124.431	82.131	44.027	3.167	6.334	16.103	0
Kab. Blitar	335.090	310.777	154.481	65.699	38.056	3.109	5.443	17.505	0
Kab. Kediri	375.859	461.130	222.710	102.289	52.752	5.999	7.363	10.954	0
Kab. Malang	757.903	657.086	304.588	132.868	64.565	7.421	6.510	25.116	0
Kab Lumajang	325.452	317.256	106.872	62.268	16.752	2.364	1.668	11.424	348
Kab. Jember	822.183	565.813	248.461	135.358	43.444	4.073	9.096	22.401	3.937
Kab. Banyuwangi	453.647	407.687	223.991	104.617	54.937	7.556	5.748	18.310	425
Kab. Bondowoso	292.590	188.647	53.339	34.689	13.979	2.425	2.038	6.953	249
Kab. Situbondo	247.440	150.832	59.513	50.658	7.601	3.105	1.132	6.082	550
Kab. Probolinggo	405.499	286.179	92.137	43.641	20.865	3.202	636	4.792	0
Kab. Pasuruan	506.800	402.230	141.500	78.510	37.970	3.320	3.030	9.570	0
Kab. Sidoarjo	218.559	345.693	313.784	316.546	123.703	4.510	15.326	80.415	902
Kab. Mojokerto	245.314	230.498	177.314	90.129	37.633	3.595	3.356	15.851	924
Kab. Jombang	289.581	313.068	192.576	94.796	45.993	6.458	4.722	19.374	0
Kab. Nganjuk	261.639	316.673	143.617	83.921	37.224	6.123	6.316	9.913	0
Kab. Madiun	191.622	187.073	99.604	48.705	29.499	1.388	2.941	6.947	0
Kab. Magetan	138.272	198.304	93.840	45.712	37.840	2.544	6.384	13.688	184
Kab. Ngawi	266.664	240.506	110.495	43.706	33.210	4.141	4.469	7.175	0
Kab. Bojonegoro	352.085	395.462	155.871	89.293	21.604	1.579	1.155	9.703	0
Kab. Tuban	327.912	362.484	130.154	57.446	13.722	2.306	1.776	8.488	392
Kab. Lamongan	333.638	315.269	221.096	125.069	24.679	3.429	1.432	11.913	1.119

Lampiran 1. (Lanjutan)

Kabupaten/Kota	X1	X2	X3	X4	X5	X6	X7	X8	X9
Kab. Gresik	212.958	265.689	189.324	146.286	40.995	3.690	7.677	26.577	288
Kab. Bangkalan	363.590	229.637	60.959	48.217	6.970	828	3.333	7.466	0
Kab. Sampang	427.348	169.660	39.527	12.394	1.926	1.373	820	4.043	0
Kab. Pamekasan	242.112	243.952	53.642	40.779	11.321	4.710	217	5.297	217
Kab. Sumenep	433.274	272.581	98.045	59.747	5.134	1.108	0	5.154	0
Kota Kediri	30.906	60.600	48.682	38.279	19.897	1.414	2.323	10.100	505
Kota Blitar	16.983	26.826	22.236	19.023	9.588	1.377	1.224	6.171	153
Kota Malang	128.718	153.794	109.201	174.419	47.352	5.943	7.075	37.922	1.698
Kota Probolinggo	45.432	45.176	28.024	24.256	14.716	1.760	1.604	4.720	560
Kota Pasuruan	29.195	48.521	23.334	26.206	9.159	1.491	2.130	4.828	142
Kota Mojokerto	16.027	23.970	18.706	23.547	7.943	611	1.081	4.982	188
Kota Madiun	18.469	30.952	32.412	36.427	17.009	1.168	3.139	8.833	365
Kota Surabaya	342.722	492.002	445.974	549.226	192.820	16.172	29.856	160.476	4.976
Kota Batu	33.488	48.265	34.137	23.413	5.769	762	644	4.018	510

Keterangan:

- X1 : Jumlah penduduk berumur 10 tahun ke atas yang tidak berijazah
- X2 : Jumlah penduduk berumur 10 tahun ke atas yang berijazah SD
- X3 : Jumlah penduduk berumur 10 tahun ke atas yang berijazah SLTP
- X4 : Jumlah penduduk berumur 10 tahun ke atas yang berijazah SMU
- X5 : Jumlah penduduk berumur 10 tahun ke atas yang berijazah SMK
- X6 : Jumlah penduduk berumur 10 tahun ke atas yang berijazah Diploma I/II
- X7 : Jumlah penduduk berumur 10 tahun ke atas yang berijazah Diploma III/Sarmud
- X8 : Jumlah penduduk berumur 10 tahun ke atas yang berijazah Diploma IV/S1
- X9 : Jumlah penduduk berumur 10 tahun ke atas yang berijazah S2/S3

Lampiran 1. (Lanjutan)

Data 4. Data Pendapatan Domestik Regional Bruto (PDRB) Beberapa Sub Sektor Kabupaten Ponorogo (Dalam Satuan Ribuan Rupiah)

Kecamatan	X1	X2	X3	X4	X5	X6	X7
NGRAYUN	13.670.893	2.314.277	4.265.333	4.737.928	973	2.407.723	688.654
SLAHUNG	26.581.487	6.631.478	5.944.058	1.050.957	36.522	9.017.953	644.646
BUNGKAL	34.049.790	4.534.643	5.335.948	704.673	45.399	10.143.658	681.555
SAMBIT	27.099.375	2.694.506	3.297.496	417.081	220.964	15.263.838	325.201
SAWOO	49.247.728	5.141.652	6.834.810	3.092.712	136.215	15.425.188	270.824
SOOKO	18.825.058	1.529.977	2.869.250	723.748	67.115	6.001.376	806.541
PULUNG	54.965.145	1.944.120	4.642.189	2.594.195	39.206	17.303.267	1.503.689
MLARAK	24.343.820	3.536.677	5.498.682	187.448	208.747	15.503.174	277.631
SIMAN	27.133.143	1.938.616	6.440.824	36.683	129.306	11.003.486	204.387
JETIS	22.422.454	6.919.169	5.464.422	77.034	144.636	14.214.827	1.067.235
BALONG	52.476.147	5.259.819	4.607.930	714.211	335.059	10.373.943	304.760
KAUMAN	34.129.498	2.290.635	1.456.037	77.400	80.241	9.740.794	1.046.702
JAMBON	31.985.421	1.333.138	4.642.189	491.914	51.907	2.162.058	141.427
BADEGAN	21.811.299	552.355	6.706.336	439.458	68.979	6.431.324	110.271
SAMPUNG	35.712.686	1.346.468	6.355.175	820.957	742.468	7.972.822	335.071
SUKOREJO	56.027.311	581.162	616.675	268.150	110.422	12.080.602	645.212
PONOROGO	12.992.878	1.794.265	6.800.551	3.668	112.186	51.628.199	787.842
BABADAN	44.598.567	4.823.923	3.956.995	8.804	182.827	21.730.070	1.236.277
JENANGAN	46.087.012	1.588.942	5.550.072	79.234	759.124	10.091.258	567.472
NGEBEL	4.085.760	2.017.448	1.730.115	1.230.335	246.020	6.102.500	112.094
PUDAK	7.462.569	33.436	222.688	1.081.404	19.062	363.326	21.359

Keterangan:

X1 : PDRB sektor pertanian subsektor tanaman pangan

X2 : PDRB sektor pertanian subsektor tanaman perkebunan

X3 : PDRB sektor pertanian subsektor peternakan

X4 : PDRB sektor pertanian subsektor kehutanan

X5 : PDRB sektor pertanian subsektor perikanan

Lampiran 1. (Lanjutan)

X6 : PDRB sektor industri

X7 : PDRB sektor keuangan

Data 5. Data Industri Tempe di Kota Malang

No.	NAMA	X1	X2	X3
1	KASIAN	82.125.000	18.250	2.150.000
2	SAMUKI	142.350.000	31.025	1.750.000
3	SUPANGAT	65.700.000	14.600	1.000.000
4	MANI	25.185.000	5.475	500.000
5	SULKAH	67.160.000	14.600	1.950.000
6	RASEMO	293.825.000	63.875	2.649.000
7	SEKAK	226.665.000	49.275	2.371.000
8	SLAMET	98.550.000	21.900	2.000.000
9	NGATMANI	24.820.000	5.475	1.650.000
10	REBAN	131.400.000	29.200	1.825.000
11	TAWI	67.160.000	14.600	1.950.000
12	SYUKUR	32.850.000	7.300	1.775.000
13	MULYONO	40.370.000	10.950	1.900.000
14	RIDWAN	503.700.000	109.500	8.480.000
15	MUFID	82.125.000	18.250	2.200.000
16	BIONO	251.850.000	54.750	5.235.000
17	MUSTOFA	16.425.000	3.650	1.600.000
18	MUSTAQIM	32.850.000	7.300	1.825.000
19	SUMARI	183.960.000	40.150	2.190.000
20	BADRUS	116.800.000	25.550	2.450.000
21	IWAN H.	254.040.000	54.750	2.880.000
22	M. AINUR	83.950.000	18.250	2.200.000
23	SUGIANTO	167.535.000	36.500	2.894.000

Keterangan:

X1 : Nilai produksi (rupiah/tahun)

X2 : Banyaknya bahan baku (kg/tahun)

X3 : Nilai investasi (rupiah)

Lampiran 2. Hasil Pembakuan Data

1. Data 1

No.	Z1	Z2
1	-0.59756	-0.01966
2	-1.49077	-1.60975
3	-0.98682	-0.89493
4	0.57680	-0.21339
5	1.50933	1.65921
6	-0.31839	0.15743
7	-1.13427	-0.85378
8	-0.33477	0.62812
9	-0.07395	-0.40019
10	-0.06739	-0.03307
11	-0.03921	-0.44412
12	1.11089	0.28412
13	0.63381	0.28042
14	0.59777	-0.00255
15	-0.45273	0.07143
16	-0.53268	0.56108
17	1.52440	3.15451
18	1.35664	0.49958
19	-1.49929	-1.62963
20	-1.15328	-1.13166
21	-0.79088	-0.11676
22	-0.10606	-0.37522
23	2.26165	0.92589
24	0.42083	0.27626
25	-0.41406	-0.77332

Lampiran 2. (Lanjutan)

2. Data 2

No.	Z1	Z2	Z3	Z4
1	0.32681	0.55178	1.02165	0.01511
2	-0.04217	-0.14487	-0.30616	-0.27957
3	-0.78012	-0.62507	-0.54156	-0.81952
4	-0.41114	-0.16409	-0.30616	-0.49446
5	1.06475	0.65884	0.76602	0.88544
6	1.43373	1.59813	1.15038	0.92670
7	2.54065	2.01821	1.66532	2.74945
8	1.80270	1.64939	1.36739	2.15391
9	-0.04217	0.03729	0.57659	-0.14375
10	0.69578	1.10539	1.08602	0.52910
11	-0.78012	-0.43470	-0.50478	-0.71115
12	-0.41114	-0.09434	-0.30616	-0.51846
13	-0.96460	-0.85660	-0.98294	-0.85764
14	-1.07530	-0.63385	-0.39444	-0.81518
15	-0.78012	-1.27097	-1.35076	0.51461
16	-0.04217	0.12795	0.57659	-0.07046
17	-0.41114	-0.06642	-0.32087	-0.39196
18	-0.04217	0.35827	0.85613	-0.06839
19	-0.96460	-0.84742	-0.79903	-0.85946
20	-1.07530	-1.35192	-1.53466	-0.85552
21	-0.04217	-1.61498	-1.71857	-0.88882

Lampiran 2. (Lanjutan)

3. Data 3

No.	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9
1	-0.567	-0.445	-0.621	-0.760	-0.562	-0.568	-0.665	-0.461	-0.312
2	0.069	-0.217	-0.045	-0.313	-0.030	-0.017	-0.249	-0.047	0.356
3	-0.600	-0.132	-0.368	-0.530	-0.319	-0.325	-0.239	-0.159	0.039
4	-0.205	0.249	-0.037	-0.024	0.313	-0.109	0.379	-0.043	-0.525
5	0.345	0.369	0.283	-0.195	0.142	-0.130	0.209	0.008	-0.525
6	0.567	1.371	1.009	0.186	0.563	0.893	0.574	-0.231	-0.525
7	2.656	2.678	1.880	0.504	0.901	1.397	0.412	0.286	-0.525
8	0.292	0.413	-0.224	-0.231	-0.469	-0.394	-0.508	-0.214	-0.183
9	3.007	2.069	1.283	0.530	0.296	0.211	0.903	0.187	3.340
10	0.993	1.015	1.022	0.210	0.625	1.445	0.267	0.037	-0.108
11	0.112	-0.445	-0.794	-0.518	-0.548	-0.372	-0.437	-0.378	-0.280
12	-0.135	-0.697	-0.728	-0.352	-0.731	-0.131	-0.610	-0.410	0.015
13	0.729	0.205	-0.381	-0.425	-0.351	-0.097	-0.704	-0.457	-0.525
14	1.283	0.979	0.144	-0.062	0.139	-0.055	-0.249	-0.282	-0.525
15	-0.293	0.602	1.978	2.415	2.596	0.366	2.087	2.308	0.360
16	-0.146	-0.166	0.526	0.059	0.130	0.042	-0.187	-0.052	0.382
17	0.096	0.385	0.688	0.108	0.369	1.056	0.072	0.076	-0.525
18	-0.057	0.409	0.167	-0.006	0.118	0.937	0.375	-0.270	-0.525
19	-0.440	-0.455	-0.301	-0.372	-0.103	-0.739	-0.266	-0.378	-0.525
20	-0.731	-0.380	-0.363	-0.403	0.136	-0.330	0.388	-0.132	-0.344
21	-0.030	-0.099	-0.185	-0.424	0.003	0.236	0.024	-0.370	-0.525
22	0.437	0.934	0.297	0.050	-0.330	-0.672	-0.605	-0.277	-0.525
23	0.305	0.714	0.024	-0.281	-0.556	-0.414	-0.487	-0.322	-0.140
24	0.337	0.399	0.991	0.423	-0.242	-0.017	-0.553	-0.196	0.573
25	-0.323	0.069	0.653	0.643	0.226	0.076	0.634	0.340	-0.242
26	0.500	-0.172	-0.713	-0.377	-0.749	-0.938	-0.191	-0.359	-0.525
27	0.849	-0.571	-0.941	-0.750	-0.894	-0.745	-0.669	-0.484	-0.525
28	-0.164	-0.076	-0.790	-0.455	-0.624	0.437	-0.783	-0.438	-0.312
29	0.881	0.115	-0.318	-0.257	-0.802	-0.838	-0.825	-0.443	-0.525
30	-1.318	-1.298	-0.843	-0.481	-0.379	-0.730	-0.383	-0.263	-0.029
31	-1.395	-1.523	-1.125	-0.681	-0.674	-0.743	-0.592	-0.406	-0.375
32	-0.784	-0.677	-0.199	0.936	0.408	0.874	0.519	0.754	1.142
33	-1.239	-1.401	-1.063	-0.627	-0.527	-0.608	-0.520	-0.459	0.025

Lampiran 2. (Lanjutan)

No.	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9
34	-1.328	-1.379	-1.113	-0.606	-0.686	-0.703	-0.420	-0.455	-0.385
35	-1.400	-1.542	-1.162	-0.634	-0.721	-1.014	-0.619	-0.450	-0.340
36	-1.386	-1.496	-1.016	-0.500	-0.461	-0.817	-0.228	-0.309	-0.167
37	0.386	1.577	3.384	4.837	4.577	4.496	4.847	5.234	4.359
38	-1.304	-1.380	-0.998	-0.635	-0.783	-0.961	-0.702	-0.485	-0.024

4. Data 5

No.	Z1	Z2	Z3
1	-0.4159	-0.4108	-0.1646
2	0.1066	0.1019	-0.4181
3	-0.5584	-0.5573	-0.8933
4	-0.9099	-0.9235	-1.2101
5	-0.5457	-0.5573	-0.2913
6	1.4207	1.4202	0.1516
7	0.8381	0.8343	-0.0246
8	-0.2734	-0.2643	-0.2596
9	-0.9130	-0.9235	-0.4814
10	0.0116	0.0287	-0.3705
11	-0.5457	-0.5573	-0.2913
12	-0.8434	-0.8502	-0.4022
13	-0.7781	-0.7037	-0.3230
14	3.2415	3.2512	3.8465
15	-0.4159	-0.4108	-0.1329
16	1.0566	1.0540	1.7903
17	-0.9859	-0.9967	-0.5131
18	-0.8434	-0.8502	-0.3705
19	0.4676	0.4681	-0.1392
20	-0.1150	-0.1178	0.0255
21	1.0756	1.0540	0.2979
22	-0.4000	-0.4108	-0.1329
23	0.3251	0.3216	0.3069

Lampiran 3. Pengujian Multikolinieritas

1. Data 1

Correlations: x1, x2

Pearson correlation of x1 and x2 = 0.770
P-Value = 0.000

2. Data 2

Predictor	VIF
x1	9.9
x2	22.5
x3	13.7
x4	6.6

3. Data 3

Predictor	VIF
x1	7.8
x2	17.6
x3	24.8
x4	42.9
x5	47.6
x6	4.4
x7	21.8
x8	50.1
x9	3.7

4. Data 4

Predictor	VIF
x1	1.3
x2	1.3
x3	1.6
x4	1.2
x5	1.3
x6	1.6
x7	1.7

5. Data 5

Predictor	VIF
x1	3004.1
x2	3032.8
x3	3.9

Lampiran 4. Skor Komponen Utama

1. Data 1

Principal Component Analysis: x1, x2

Eigenanalysis of the Correlation Matrix

Eigenvalue	1.7705	0.2295
Proportion	0.885	0.115
Cumulative	0.885	1.000

Variable	PC1	PC2
x1	0.707	-0.707
x2	0.707	0.707

No.	KU1	KU2
1	-0.43644	0.40863
2	-2.1924	-0.08413
3	-1.3306	0.06498
4	0.25696	-0.55875
5	2.24049	0.10598
6	-0.11381	0.33645
7	-1.40576	0.19834
8	0.20743	0.68087
9	-0.33527	-0.23069
10	-0.07104	0.02427
11	-0.34177	-0.28631
12	0.98642	-0.58462
13	0.64646	-0.24989

No.	KU1	KU2
14	0.42088	-0.42449
15	-0.26962	0.37063
16	0.02008	0.7734
17	3.30849	1.15266
18	1.31255	-0.60603
19	-2.21248	-0.09217
20	-1.61569	0.01529
21	-0.6418	0.47668
22	-0.34032	-0.19033
23	2.25393	-0.94453
24	0.49291	-0.10223
25	-0.83961	-0.25403

Lampiran 4. (Lanjutan)

2. Data 2

Principal Component Analysis: x1, x2, x3, x4

Eigenanalysis of the Correlation Matrix

Eigenvalue	3.5850	0.3146	0.0722	0.0282
Proportion	0.896	0.079	0.018	0.007
Cumulative	0.896	0.975	0.993	1.000

Variable	PC1	PC2	PC3	PC4
x1	-0.506	-0.355	-0.760	0.201
x2	-0.514	0.336	-0.024	-0.789
x3	-0.492	0.622	0.190	0.579
x4	-0.488	-0.612	0.621	0.038

No.	KU1	KU2	KU3	KU4
1	-0.9585	0.69528	-0.05798	0.223054
2	0.38277	-0.05304	-0.19636	-0.08232
3	1.38217	0.2317	-0.00431	-0.0092
4	0.68421	0.20307	-0.04907	-0.1497
5	-1.68604	-0.22227	-0.12914	0.172457
6	-2.56451	0.17611	-0.33342	-0.26993
7	-4.48333	-0.87089	0.045762	-0.01016
8	-3.48328	-0.55358	0.188701	-0.06327
9	-0.2111	0.47392	0.051454	0.290654
10	-1.71223	0.47588	-0.02011	-0.08236
11	1.21335	0.25227	0.065419	-0.13391
12	0.66008	0.2412	-0.06566	-0.20564
13	1.83008	-0.03166	0.033841	-0.1209
14	1.46157	0.42244	0.250842	0.023697
15	1.46061	-1.30465	0.686283	0.082635
16	-0.29347	0.45956	0.094799	0.221951
17	0.59121	0.16405	0.009461	-0.23134
18	-0.55027	0.70952	0.143678	0.2023

Lampiran 4. (Lanjutan)

No.	KU1	KU2	KU3	KU4
19	1.73584	0.08687	0.067453	-0.02166
20	2.41081	-0.50315	0.026313	-0.07204
21	2.13003	-1.05264	-0.80796	0.235673

3. Data 3

Principal Component Analysis: x1, x2, x3, x4, x5, x6, x7, x8, x9

Eigenanalysis of the Correlation Matrix

Eigenvalue	6.4688	1.7058	0.3899	0.2098	0.0957
Proportion	0.719	0.189	0.043	0.023	0.011
Cumulative	0.719	0.908	0.951	0.975	0.986

Eigenvalue	0.0674	0.0336	0.0197	0.0093	
Proportion	0.007	0.004	0.002	0.001	
Cumulative	0.993	0.997	0.999	1.000	

Variable	PC1	PC2	PC3	PC4	PC5
x1	-0.171	0.680	0.272	-0.050	0.520
x2	-0.270	0.558	-0.108	-0.066	-0.199
x3	-0.366	0.193	-0.184	-0.209	-0.634
x4	-0.377	-0.164	-0.036	-0.230	-0.001
x5	-0.376	-0.141	-0.254	-0.136	0.077
x6	-0.353	-0.001	-0.196	0.906	0.017
x7	-0.370	-0.197	-0.071	-0.121	0.337
x8	-0.362	-0.273	-0.013	-0.179	0.318
x9	-0.296	-0.170	0.878	0.104	-0.260

Variable	PC6	PC7	PC8	PC9
x1	-0.152	0.368	-0.055	-0.015
x2	0.191	-0.700	0.178	-0.046
x3	-0.045	0.447	-0.162	0.348
x4	-0.542	-0.185	-0.442	-0.504
x5	0.224	0.311	0.609	-0.481
x6	-0.107	0.029	-0.059	0.021
x7	0.658	-0.042	-0.487	0.135
x8	-0.366	-0.202	0.344	0.607
x9	0.127	-0.024	0.116	-0.064

Lampiran 4. (Lanjutan)

No.	KU1	KU2	KU3	KU4	KU5	KU6	KU7	KU8	KU9
1	1.6482	-0.23966	0.06902	0.05604	-0.15337	0.068148	-0.09772	0.20841	0.104128
2	0.2027	-0.02574	0.40316	0.15724	-0.08631	0.013476	0.216262	0.231996	0.080063
3	0.842	-0.33649	0.13567	0.04831	-0.22261	0.239117	-0.26217	0.193892	0.169243
4	-0.0577	-0.02157	-0.62012	-0.2273	0.140408	0.373591	-0.16233	0.009141	-0.10346
5	-0.1209	0.55265	-0.4774	-0.27397	0.145127	0.224928	0.073081	0.021192	0.170181
6	-1.407	1.27405	-1.00213	0.27671	-0.30322	0.455456	-0.09814	-0.16176	-0.09613
7	-2.9873	3.38168	-0.92655	0.16741	0.11307	-0.09933	0.109205	0.108337	0.094778
8	0.6433	0.67941	0.15878	-0.15353	-0.02259	-0.17213	-0.3273	0.085616	0.045215
9	-3.3165	2.52213	3.08905	-0.32386	-0.14331	0.586411	0.077876	-0.05828	-0.07586
10	-1.7232	1.27027	-0.59077	0.79396	-0.13257	0.016441	0.293464	0.034714	-0.01709
11	1.306	0.07452	0.2455	0.13859	0.407786	-0.05377	0.013139	0.01042	-0.01095
12	1.301	-0.23048	0.45876	0.41584	0.132798	-0.28809	0.05424	-0.12332	-0.02499
13	0.8671	1.00896	-0.03466	0.20031	0.30557	-0.25556	0.054839	0.162627	-0.11128
14	-0.1959	1.65246	-0.24465	-0.18759	0.352895	-0.07188	-0.01558	0.159415	-0.22275
15	-4.5674	-1.3463	-1.19103	-1.64633	0.020122	-0.12163	0.165432	0.115435	-0.13294
16	-0.2334	-0.13205	0.18729	-0.01289	-0.54465	-0.09588	0.337208	0.061202	0.020097
17	-0.8238	0.39584	-0.91341	0.63038	-0.2317	-0.10806	0.193388	-0.00259	0.099711
18	-0.4213	0.29341	-0.78715	0.72321	-0.01464	0.286911	-0.11611	-0.26814	-0.07385
19	1.1391	-0.2902	-0.26767	-0.40903	-0.03953	0.147956	0.136794	0.075147	-0.08727
20	0.5835	-0.71462	-0.37455	-0.14617	0.109949	0.599181	-0.03961	0.03697	0.028183
21	0.4551	0.14292	-0.45215	0.36623	0.154166	0.283628	0.147634	-0.01057	-0.03001
22	0.3869	1.19883	-0.23724	-0.65218	-0.33954	-0.29422	-0.39743	0.048968	-0.04366
23	0.5464	0.94294	0.15018	-0.20004	-0.27811	-0.05951	-0.42107	0.027001	0.118802
24	-0.4844	0.67315	0.46056	-0.16828	-0.95062	-0.51981	0.183242	-0.17277	-0.00608
25	-0.8428	-0.36934	-0.57353	-0.39812	-0.19319	-0.01113	-0.01052	-0.44687	0.105074
26	1.3332	0.5001	0.23105	-0.4916	0.631067	-0.00093	-0.11218	-0.26853	0.072995
27	1.8137	0.67984	0.45908	-0.07398	0.828295	-0.23962	0.268625	-0.06712	0.127615
28	1.1306	0.18282	-0.01332	0.90536	0.068662	-0.28914	-0.32939	0.112929	-0.08283
29	1.2511	1.13076	0.26678	-0.45071	0.278882	-0.4953	0.01194	-0.11391	0.021212
30	1.7111	-1.49849	0.19797	-0.0801	-0.13743	0.085054	0.065903	0.095063	-0.01382
31	2.3106	-1.51661	0.05171	0.08733	-0.00245	-0.01856	0.0603	0.026211	0.049103

Lampiran 4. (Lanjutan)

No.	KU1	KU2	KU3	KU4	KU5	KU6	KU7	KU8	KU9
32	-1.2289	-1.66321	0.54321	0.56588	0.016197	-0.299595	-0.125207	-0.122407	-0.220473
33	1.9791	-1.42808	0.34999	0.19088	-0.06815	0.084832	0.097264	0.079978	-0.080584
34	2.1754	-1.4176	0.02365	0.07971	0.04086	0.080639	-0.024587	-0.098438	0.008751
35	2.4419	-1.52747	0.15523	-0.13874	-0.017701	-0.024275	0.032724	-0.002162	-0.001906
36	1.9086	-1.66813	0.14043	-0.11534	0.041934	-0.16823	0.083211	-0.052557	-0.014599
37	-11.8427	-2.77149	0.50642	0.42856	0.323769	-0.159815	-0.149941	0.045414	0.110008
38	2.2769	-1.35921	0.42283	-0.08222	-0.229875	-0.03576	0.013499	0.019357	0.025378

4. Data 5

Principal Component Analysis: x1, x2, x3

Eigenanalysis of the Correlation Matrix

Eigenvalue	2.8130	0.1868	0.0002
Proportion	0.938	0.062	0.000
Cumulative	0.938	1.000	1.000

Variable	PC1	PC2	PC3
x2	-0.587	0.397	0.705
x3	-0.588	0.390	-0.709
x4	-0.556	-0.831	0.004

No.	KU1	KU2	KU3
1	0.57725	-0.1885	-0.0029
2	0.11012	0.42941	0.00134
3	1.15253	0.30326	-0.0024
4	1.75049	0.28417	0.00795
5	0.81011	-0.1919	0.00889
6	-1.7535	0.99185	-0.0038
7	-0.9689	0.67844	-0.0002
8	0.46037	0.00414	-0.0065
9	1.34685	-0.3226	0.00858

Lampiran 4. (Lanjutan)

No.	KU1	KU2	KU3
10	0.18252	0.32366	-0.0136
11	0.81011	-0.1919	0.00889
12	1.21882	-0.3322	0.00612
13	1.05034	-0.3149	-0.0514
14	-5.955	-0.6413	-0.0027
15	0.55962	-0.2148	-0.0027
16	-2.2362	-0.657	0.0053
17	1.4503	-0.3537	0.00899
18	1.20118	-0.3585	0.00625
19	-0.4723	0.48386	-0.0025
20	0.12262	-0.1128	0.00246
21	-1.417	0.59041	0.01284
22	0.55032	-0.2086	0.00844
23	-0.5507	-0.0005	0.00259

Lampiran 5. Indeks Validitas

1. Data 1

k	W(k)	B(k)	H(k)	KL(k)	CH(k)
1	48.	-	23.192858	-	-
2	24.9	23.1	2.557878	0.101659	21.3374
3	22.502	25.4982	29.429514	0.631869	12.4648
4	9.8711	38.1289	23.873370	1.771826	27.0385
5	4.734	43.266	2.771160	11.115340	45.6968
6	4.1821	43.8179	9.744836	0.263048	39.8139
7	2.812	45.188	6.508516	1.847330	48.2088
8	2.0945	45.9054	9.742292	0.647002	53.2266
9	1.359	46.6409	9.442671	1.295181	68.6407
10	0.8737	47.1262	4.390287	2.922352	89.8984
11	0.6856	47.3143	7.462775	0.583890	96.6195
12	0.4578	47.5421	7.015673	1.339214	122.7285
13	0.305	47.6949	4.779733	1.814078	156.3881
14	0.223	47.7769	3.779396	1.457826	181.2930
15	0.1696	47.8303	4.582282	0.920158	201.4642
16	0.1197	47.8802	2.429818	2.259186	239.9762
17	0.0963	47.9036	2.773895	0.891063	248.6932
18	0.0736	47.9263	3.509348	0.884472	268.0565
19	0.0512	47.9488	5.208187	0.915530	312.3346
20	0.0293	47.9706	4.593226	1.620669	430.2377
21	0.0166	47.9833	11.950956	0.986221	577.4498
22	0.0049	47.9951	14.478618	2.890149	1398.6740
23	0.0011	47.9989	10.608780	4.440176	4112.1230
24	0.0002	47.9997	-	3.347249	8921.4680

Lampiran 5. (Lanjutan)

2. Data 2

k	W(k)	B(k)	H(k)	KL(k)	CH(k)
1	80	-	17.381156	-	-
2	43.8	36.2285	36.198644	0.5182441	15.72581
3	15.57742	64.4227	8.260735	6.6304746	37.22076
4	10.85704	69.1431	11.992760	0.7372006	36.08812
5	6.51579	73.4842	4.633717	3.5229531	45.11142
6	5.12018	74.8799	2.066805	3.7221912	43.87335
7	4.53444	75.4656	2.122202	0.7156505	38.83307
8	3.97243	76.0276	2.547996	0.6600375	35.54356
9	3.36077	76.6392	21.001748	0.1916296	34.20616
10	1.284928	78.7150	3.381157	8.1504191	74.87344
11	1.002465	78.9975	4.940090	0.7954464	78.80268
12	0.691783	79.3081	8.484687	0.8866736	93.79800
13	0.374243	79.6257	4.781386	2.4075956	141.84041
14	0.244401	79.7555	5.901316	1.1762176	175.71301
15	0.140651	79.8593	8.714607	1.2570582	243.33412
16	0.0626519	79.9373	3.896813	3.1290371	425.29276
17	0.0379822	79.9620	4.031343	1.3948156	526.29273
18	0.0210275	79.9789	4.839930	1.4116346	671.17239
19	0.0095149	79.9904	7.493340	1.6288042	934.05096
20	0.0027204	79.9972	-	2.4092144	1547.70747

3. Data 3

k	W(k)	B(k)	H(k)	KL(k)	CH(k)
1	333	-	32.125486	-	-
2	180.449	152.552	7.637095	6.26188807	30.43467
3	149.575	183.426	9.33685	0.66739266	21.4605
4	118.771	214.23	15.820086	0.65644991	20.44221
5	81.7978	251.203	3.417584	7.10048109	25.33587
6	74.3268	255.593	2.176916	1.94640489	22.27342
7	69.727	263.274	27.426191	0.06752783	19.50814

Lampiran 5. (Lanjutan)

k	W(k)	B(k)	H(k)	KL(k)	CH(k)
8	37.5469	295.454	12.3834	3.012377	33.72399
9	26.8294	306.172	2.610104	6.237497	41.36761
10	24.682	308.319	11.37142	0.232257	38.86282
11	17.72987	315.271	4.981228	2.699323	48.0112
12	15.0521	317.949	2.458742	2.318036	49.92768
13	13.79578	319.205	4.620191	0.525113	48.20394
14	11.71419	321.287	3.392622	1.496552	50.6348
15	10.31452	322.686	5.242456	0.707355	51.39661
16	8.46541	324.536	4.664935	1.277981	56.2277
17	7.03797	325.963	3.628734	1.429595	60.78914
18	6.04149	326.959	2.675979	1.468558	63.67045
19	5.358666	327.642	1.960637	1.445505	64.54054
20	4.880251	328.121	2.147238	0.932898	63.69701
21	4.38468	328.616	1.994046	1.117654	63.70514
22	3.947367	329.054	2.000052	1.030846	63.5129
23	3.531841	329.469	1.969003	1.054864	63.6039
24	3.144841	329.856	2.009796	1.021541	63.84534
25	2.773265	330.228	2.419595	0.884554	64.49974
26	2.364607	330.636	2.95534	0.911904	67.11793
27	1.926627	331.074	2.198376	1.46703	72.70333
28	1.355063	331.373	2.21802	1.076797	75.37369
29	1.355063	331.646	2.311523	1.058883	78.66829
30	1.100651	331.9	1.793013	1.3867	83.18634
31	0.917805	332.083	5.671203	0.463901	84.426
32	0.537085	332.464	2.827219	2.464191	119.8128
33	0.382574	332.619	2.725642	1.281118	135.8527
34	0.26307	332.738	2.528309	1.341156	153.3191

Lampiran 5. (Lanjutan)

k	W(k)	B(k)	H(k)	KL(k)	CH(k)
35	0.174715	332.826	2.473631	1.311361	168.0875
36	0.107959	332.893	6.962571	0.872709	176.2098
37	0.032512	332.969	-	2.300699	284.5209

4. Data 4

k	W(k)	B(k)	H(k)	KL(k)	CH(k)
1	7.15E+15	-	8.025817	-	-
2	5.17E+15	1.98E+15	33.426282	0.2309103	7.261453
3	1.94E+15	5.22E+15	19.940641	2.9325207	24.230174
4	9.44E+14	6.20E+15	4.931057	5.4344320	37.214567
5	7.41E+14	6.40E+15	2.978713	1.8848698	34.561255
6	6.31E+14	6.52E+15	9.586499	0.3335118	30.988422
7	3.94E+14	6.74E+15	3.591895	3.2767855	39.941063
8	3.18E+14	6.82E+15	2.730970	1.4344507	39.846917
9	2.66E+14	6.87E+15	2.649008	1.0808450	38.754771
10	2.21E+14	6.92E+15	2.001849	1.4086885	38.261650
11	1.90E+14	6.97E+15	1.772468	1.1569809	36.694129
12	1.63E+14	6.97E+15	1.834966	0.9787686	34.991931
13	1.38E+14	7.01E+15	1.829246	1.0370898	33.866120
14	1.15E+14	7.05E+15	4.035842	0.5527715	33.022413
15	7.62E+13	7.07E+15	2.324333	2.0191447	39.758680
16	5.72E+13	7.09E+15	1.920689	1.3422191	41.302144
17	4.33E+13	7.10E+15	2.484347	0.9267971	40.972711
18	2.90E+13	7.13E+15	3.113845	1.1017801	43.380009
19	1.63E+13	7.14E+15	4.366639	1.2802787	48.662209
20	6.63E+12	7.14E+15	-	1.4198220	56.678224

Lampiran 5. (Lanjutan)

5. Data 5

k	W(k)	B(k)	H(k)	KL(k)	CH(k)
1	66	-	30.2705	-	-
2	28.4961	37.5040	35.92212	0.91392031	27.63829
3	10.8234	55.1767	4.884586	58.71742489	50.97906
4	8.78094	57.2190	26.94114	0.03458962	41.26974
5	3.74126	62.2588	6.739675	6.14423362	74.88507
6	2.76164	63.2382	4.055777	2.08954893	77.85586
7	2.25381	63.7460	2.557207	2.12045239	75.42314
8	1.9591	64.0403	22.60384	0.09505245	70.0471
9	0.81199	65.1881	4.242209	7.51516949	140.49339
10	0.63298	65.3676	4.738433	0.96046872	149.16721
11	0.47292	65.5281	4.439349	1.21910091	166.27277
12	0.35253	65.6478	5.04573	1.03286767	186.21898
13	0.24818	65.7529	7.64854	0.91163503	220.78362
14	0.14639	65.8542	9.350517	1.33712852	311.43771
15	0.07565	65.9234	10.27872	1.65430897	497.95812
16	0.03532	65.9710	10.22962	1.93695284	871.64346
17	0.0155	65.9906	8.121585	2.28940291	1596.54601
18	0.00717	65.9513	26.16283	1.33827814	2705.36161
19	0.00134	66.0770	5.665722	8.09999968	10958.04618
20	0.00063	66.2691	16.02342	1.33744231	16608.79828
21	0.00013	68.4535	-	3.84564904	52656.50329
22	0	69.2547	-	-	Inf