

**PENENTUAN JUMLAH KELOMPOK PADA ANALISIS
KELOMPOK METODE *SINGLE LINKAGE* MENGGUNAKAN
BEBERAPA INDEKS VALIDITAS**

SKRIPSI

Oleh:
FITRIATI SHOLIAH
0610950022-95



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2011**

UNIVERSITAS BRAWIJAYA



**PENENTUAN JUMLAH KELOMPOK PADA ANALISIS
KELOMPOK METODE *SINGLE LINKAGE* MENGGUNAKAN
BEBERAPA INDEKS VALIDITAS**

SKRIPSI

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

Oleh:
FITRIATI SHOLIAH
0610950022-95



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2011**

LEMBAR PENGESAHAN SKRIPSI

**PENENTUAN JUMLAH KELOMPOK PADA ANALISIS
KELOMPOK METODE *SINGLE LINKAGE* MENGGUNAKAN
BEBERAPA INDEKS VALIDITAS**

**Oleh:
FITRIATI SHOLIAH
NIM. 0610950022**

**Setelah dipertahankan di depan Majelis Penguji
pada tanggal 11 Februari 2011
dan dinyatakan memenuhi syarat untuk memperoleh gelar Sarjana
Sains dalam bidang Statistika**

Pembimbing I,

Pembimbing II,

**Dr. Ir. Solimun, MS
NIP. 19611215 198709 1 002**

**Dr. Ir. Ni Wayan Surya W, MS
NIP. 19551102 198103 2 001**

**Mengetahui,
Ketua Jurusan Matematika
Fakultas MIPA
Universitas Brawijaya**

**Dr. Abdul Rouf Al-Ghofari, M.Sc
NIP. 19670907 199203 1 001**

LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : FITRIATI SHOLIAH
NIM : 0610950022 - 95
Program Studi : STATISTIKA
Penulisan Skripsi berjudul :

**Penentuan Jumlah Kelompok Pada Analisis Kelompok Metode
Single Linkage Menggunakan Beberapa Indeks Validitas**

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.
2. Apabila di kemudian hari ternyata Skripsi yang saya tulis terbukti hasil jiplakan, maka saya akan bersedia menanggung segala risiko yang akan saya terima.

Demikian pernyataan ini dibuat dengan segala kesadaran.

Malang, 11 Februari 2011
Yang menyatakan,

(FITRIATI SHOLIAH)
NIM. 0610950022

PENENTUAN JUMLAH KELOMPOK PADA ANALISIS KELOMPOK METODE *SINGLE LINKAGE* MENGGUNAKAN BEBERAPA INDEKS VALIDITAS

ABSTRAK

Analisis kelompok merupakan salah satu metode pengelompokan dengan cara membagi atau memecah objek-objek pada kelompok yang besar menjadi kelompok-kelompok yang lebih kecil dimana objek dalam kelompok relatif sama dan objek pada kelompok lain relatif tidak sama. Secara garis besar metode pengelompokan pada analisis kelompok dibedakan menjadi dua yaitu metode pengelompokan hirarki dan nonhirarki. Kesulitan utama pada analisis kelompok adalah menentukan jumlah kelompok optimal yang memperkecil ukuran kesalahan dalam kelompok. Beberapa penelitian telah menghasilkan metode penentuan jumlah kelompok optimal dengan menggunakan indeks validitas. Pada penelitian ini akan digunakan indeks validitas *gap statistic*, Hartigan, Krazanowski-Lai, Silhouette, dan Calinski-Harabasz dalam menentukan jumlah kelompok optimal. Tujuan dari penelitian ini adalah membandingkan kelima indeks validitas berdasarkan nilai *Cluster Tighness Measure* (CTM). Data yang digunakan yaitu 50 data sekunder dengan skala ukur rasio dengan jumlah objek dan variabel yang berbeda-beda. Berdasarkan hasil penelitian diperoleh kesimpulan bahwa indeks validitas Calinski-Harabasz merupakan indeks validitas terbaik. Indeks validitas Calinski-Harabasz memberikan nilai CTM terkecil pada 54% data penelitian. Sementara itu, indeks validitas Krazanowski-Lai, Silhouette, Hartigan, dan *gap statistic* berturut-turut memberikan nilai CTM terkecil pada 48%, 30%, 14%, dan 8% data penelitian.

Kata kunci : Analisis kelompok, indeks validitas, *gap statistic*, Hartigan, Krazanowski-Lai, Silhouette, Calinski-Harabasz, *Cluster Tighness Measure* (CTM).

ESTIMATED THE NUMBER OF CLUSTERS IN CLUSTER ANALYSIS OF SINGLE LINKAGE METHOD USING SOME OF VALIDITY INDEX

ABSTRACT

Cluster analysis is method of grouping by dividing objects in the large group into groups of smaller where the objects are relatively similar in the groups and objects in other groups are relatively dissimilar. Clustering methods can be divided into two types namely hierarchical clustering method and nonhierarchical clustering method. The main difficulty in the cluster analysis is to determine the optimal number of cluster that decrease the size of the error in the cluster. Some research has produced a method of determining the optimal number of cluster using the validity index. In this research was used the gap statistic, Hartigan, Krazanowski-Lai, Silhouette, and Calinski-Harabasz in the determining of optimal number of cluster. The purpose of this study was to compare the validity of these index based on the value of Cluster Tighness Measure (CTM). The data which is used 50 secondary data with a scale measuring the ratio, the number of objects and variables that vary. Based on research results it is concluded that the validity index Calinski-Harabasz is best validity index. Validity index Calinski-Harabasz gives the smallest CTM value in 54% of research data. Meanwhile, the validity index Krazanowski-Lai, Silhouette, Hartigan, and the gap statistic in a row gives the smallest value of CTM at 48%, 30%, 14%, and 8% of research data.

Keywords : Cluster analysis, validity index, *gap statistic*, Hartigan, Krazanowski-Lai, Silhouette, Calinski-Harabasz, *Cluster Tighness Measure* (CTM).

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah melimpahkan rahmat, karunia, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Penentuan Jumlah Kelompok pada Analisis Kelompok Metode *Single Linkage* Menggunakan Beberapa Indeks Validitas”. Pada skripsi ini penentuan jumlah kelompok menggunakan 5 indeks validitas. Tujuan dari penelitian ini adalah membandingkan kelima indeks validitas untuk mengetahui indeks validitas terbaik berdasarkan nilai CTM. Dalam pelaksanaan dan penyusunan skripsi ini, banyak pihak yang telah membantu baik berupa bimbingan, dukungan, dan saran. Oleh karena itu, penulis mengucapkan terimakasih kepada :

1. Bapak Dr. Ir. Solimun, MS selaku dosen pembimbing I dan Ibu Dr. Ir. Ni Wayan Surya Wardhani, MS selaku dosen pembimbing II yang telah memberikan bimbingan, dukungan, serta saran.
2. Ibu Dra. Ani Budi Astuti, M.Si, Ibu Eni Sumarminingsih, S.Si., MM, dan Bapak Adji Achmad Rinaldo Fernandes, S.Si., M.Sc selaku dosen penguji yang telah memberikan saran dan kritik.
3. Bapak Dr. Abdul Rouf A., M.Sc selaku Ketua Jurusan Matematika Fakultas MIPA Universitas Brawijaya.
4. Ibu, Bapak, Kakak, dan Adek atas doa, kasih sayang, serta dukungan moril dan materil yang tidak pernah putus
5. Teman-teman Statistika 2006 atas dukungan, semangat, dan persahabatan yang tidak akan terlupakan
6. Teman-teman kos SS 259B atas motivasi, perhatian, dan kebersamaan yang tidak akan terlupakan
7. Karyawan tata usaha dan semua pihak yang telah membantu dalam proses penyelesaian skripsi ini

Penulis menyadari bahwa skripsi ini masih terdapat kekurangan. Oleh karena itu, penulis menerima saran dan kritik yang membangun demi kesempurnaan skripsi ini. Semoga skripsi ini dapat bermanfaat bagi banyak pihak.

Malang, 11 Februari 2011

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	xi
 BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan	3
1.5 Manfaat	3
 BAB II TINJAUAN PUSTAKA	
2.1 Skala Pengukuran Data	5
2.2 Korelasi	6
2.3 Pembakuan Data	7
2.4 Analisis Komponen Utama	8
2.4.1 Matriks Masukan	10
2.4.2 Skor Komponen	11
2.4.3 Proporsi Keragaman	11
2.4.4 Penentuan Komponen Utama yang Digunakan	12
2.5 Analisis Kelompok	12
2.6 Analisis Kelompok Hirarki	13
2.6.1 Metode <i>Agglomerative</i>	14
2.6.2 Metode <i>Divisive</i>	16
2.6.3 Ukuran Kedekatan Jarak	16
2.7 Estimasi Jumlah Kelompok	17
2.7.1 <i>Gap Statistic</i>	17
2.7.2 Hartigan	19
2.7.3 Krzanowski-Lai	19

	Halaman
2.7.4 Silhouette	20
2.7.5 Calinski-Harabasz	20
2.8 Cluster Tightness Measure (CTM)	20

BAB III METODE PENELITIAN

3.1 Data Penelitian.....	23
3.2 Metode Penelitian	30

BAB IV HASIL PEMBAHASAN

4.1 Pengujian Koefisien Korelasi	33
4.2 Analisis Komponen Utama	34
4.3 Analisis Kelompok Hirarki	36

BAB V KESIMPULAN DAN SARAN

5.1 Kesimpulan	41
5.2 Saran	41

DAFTAR PUSTAKA	43
-----------------------------	----

LAMPIRAN	47
-----------------------	----

DAFTAR TABEL

	Halaman
Tabel 3.1 Data Penelitian	23
Tabel 4.1 Pengujian Koefisien Korelasi.....	33
Tabel 4.2 Hasil Analisis Komponen Utama Data 4	35
Tabel 4.3 Jumalak Kelompok dan nilai CTM yang dihasilkan oleh Indeks Validitas <i>Gap statistic</i> (Gap), Hartigan (H), Krzanowski-Lai (KL), Silhouette (S), dan Calinski-Harabasz (CH)	36



DAFTAR GAMBAR

	Halaman
Gambar 3.1 Diagram Alir metode Penelitian	31

UNIVERSITAS BRAWIJAYA



DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data Penelitian.....	47
Lampiran 2. Uji Koefisien Korelasi	97
Lampiran 3. Hasil Analisis Komponen Utama Data 4.....	112
Lampiran 4. Nilai dan hasil pengelompokan indeks validitas (Gap statistic, Hartigan, Krazanowski-Lai, Silhouette dan Calinski-Harabasz) menggunakan <i>package</i> ClusterSim pada <i>software</i> R.2.7.2	113



UNIVERSITAS BRAWIJAYA



BAB I

PENDAHULUAN

1.1. Latar Belakang

Sejak zaman dahulu manusia sudah mengenal dan melakukan klasifikasi atau pengelompokan objek-objek yang ada disekelilingnya. Pengelompokan pada masa itu dilakukan dengan sangat sederhana yaitu berdasarkan ciri fisik dan manfaat objek tersebut bagi manusia. Contoh pengelompokan yang dilakukan pada waktu itu antara lain pengelompokan hewan berdasarkan jumlah kaki yaitu hewan berkaki dua dan hewan berkaki empat, pengelompokan tumbuhan berdasarkan manfaat bagi manusia yaitu tumbuhan obat-obatan dan penghasil pangan. Timbul keterbatasan ketika harus mengelompokkan banyak objek, terlebih jika objek tersebut memiliki karakteristik yang bermacam-macam dan bentuk yang tidak jelas. Oleh karena itu, diperlukan metode khusus untuk membagi objek-objek tersebut ke dalam kelompok yang lebih kecil sehingga bisa diidentifikasi.

Kelompok adalah sekumpulan objek yang memiliki kesamaan diantara anggotanya dan memiliki ketidaksamaan dengan objek lain pada kelompok lainnya. Analisis kelompok adalah salah satu teknik yang digunakan untuk menggabungkan objek-objek ke dalam kelompok-kelompok, di mana setiap objek dalam suatu kelompok memiliki karakteristik yang homogen dan setiap kelompok berbeda dengan kelompok lainnya (Sharma, 1996). Perkembangan selanjutnya, analisis kelompok banyak digunakan di berbagai disiplin ilmu antara lain pengetahuan alam, kedokteran, ekonomi, pertanian, peternakan, pendidikan dan disiplin ilmu lainnya. Contoh penggunaan analisis kelompok di bidang pengetahuan alam yaitu pengelompokan air sungai berdasarkan kandungan zat radioaktif, di bidang pertanian yaitu pengelompokan kacang tanah berdasarkan berat kering, jumlah daun, panjang tanaman, dan jumlah polong isi, di bidang pendidikan yaitu pengelompokan provinsi di Indonesia berdasarkan tingkat kelulusan.

Metode pengelompokan pada analisis kelompok pada dasarnya ada dua, yaitu metode pengelompokan hirarki dan nonhirarki. Metode pengelompokan hirarki digunakan apabila belum ada informasi jumlah kelompok yang akan dipilih. Sedangkan metode pengelompokan nonhirarki bertujuan untuk mengelompokkan n objek ke dalam k kelompok ($k < n$), di mana nilai k sudah ditentukan sebelumnya (Prayitno, 2007).

Kesulitan utama dalam analisis kelompok adalah menentukan jumlah kelompok optimal yang memperkecil ukuran kesalahan dalam kelompok. Kelompok optimal merupakan kelompok yang padat antar objek dalam kelompok dan terisolasi dari kelompok lain dengan baik yang berarti struktur kelompok yang terbentuk memberikan informasi yang sebenarnya mengenai data. Secara umum, menaikkan jumlah kelompok akan memperkecil ukuran kesalahan dalam kelompok. Ditinjau dari segi statistik, jumlah kelompok di mana ukuran kesalahan mulai turun mengindikasikan kelayakan banyaknya kelompok yang akan digunakan (Kariyam dan Subanar, 2002).

Beberapa penelitian telah menghasilkan metode penentuan jumlah kelompok optimal dengan menggunakan indeks validitas, antara lain dari Calinski dan Harabasz (1974), Hartigan (1975), Krazanowski dan Lai (1985), Kaufman dan Rousseeuw (1990) mengusulkan statistik *silhouette*. Tibshirani, Walther, dan Hatie (2001) mengusulkan metode *gap statistic*. Penelitian sebelumnya (Kariyam dan Subanar, 2002), membandingkan 4 indeks validitas didapatkan kesimpulan bahwa indeks validitas *Gap statistic* memberikan ketepatan jumlah prediksi yang tinggi pada data yang mengelompok dengan baik dan memberikan hasil yang tidak jauh berbeda dengan indeks validitas lain. Pada penelitian Kamilin (2008) menggunakan 3 indeks validitas, kesimpulan yang dihasilkan adalah indeks validitas Calinski-Harabasz selalu menghasilkan jumlah kelompok maksimum jika dibandingkan dengan indeks validitas Hartigan, dan Krazanowski-Lai. Mufidah (2004) meneliti tentang indeks validitas gabungan. Hasil penelitian ini memberikan kesimpulan bahwa analisis kelompok dengan menggunakan metode penggabungan berbeda memberikan nilai kelompok optimal yang sama tetapi dapat juga memberikan nilai kelompok optimal yang berbeda.

Berdasarkan uraian tersebut, pada penelitian ini akan digunakan indeks validitas *Gap statistic*, Hartigan, Krzanowski-Lai, *Silhouette*, Calinski-Harabasz untuk menentukan jumlah kelompok optimal. Data yang digunakan adalah 50 data sekunder yang diambil dari skripsi, penelitian, dan buku. Hasil pengelompokan akan dibandingkan berdasarkan nilai *Cluster Tightness Measure* (CTM).

CTM merupakan suatu nilai yang digunakan untuk mengukur kebaikan hasil analisis kelompok berdasarkan simpangan baku setiap variabel pada setiap kelompok. Indeks yang menghasilkan CTM terkecil merupakan indeks validitas terbaik.

1.2. Rumusan Masalah

Indeks validitas (*Gap statistic*, Hartigan, Krzanowski-Lai, Silhouette, Calinski-Harabasz) apa yang terbaik untuk menentukan kelompok optimal pada analisis kelompok hirarki metode *single linkage*.

1.3. Batasan Masalah

Pada penelitian ini analisis kelompok yang digunakan adalah analisis kelompok hirarki. Ukuran kedekatan yang digunakan adalah ukuran kedekatan jarak Euclid. Penentuan indeks validitas terbaik berdasarkan pada nilai CTM.

1.4. Tujuan Penelitian

Membandingkan indeks validitas (*Gap statistic*, Hartigan, Krzanowski-Lai, Silhouette, Calinski-Harabasz) berdasarkan nilai CTM.

1.5. Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah dapat diperoleh informasi indeks validitas yang paling baik berdasarkan nilai CTM dalam menentukan kelompok optimal pada analisis kelompok metode *single linkage* sehingga mempermudah penentuan indeks validitas yang akan digunakan pada penelitian selanjutnya.

UNIVERSITAS BRAWIJAYA



BAB II TINJAUAN PUSTAKA

2.1. Skala Pengukuran Data

Dalam menentukan statistik uji yang tepat untuk himpunan data, sangat penting untuk mengetahui tipe data tersebut. Beberapa percobaan telah dilakukan untuk merumuskan pengelompokan tipe data. Pengelompokan tipe data yang diterima adalah tipe data yang diusulkan oleh Stevens (1951) yaitu data nominal, ordinal, interval, dan rasio (Afifi dan Clark, 1990).

1. Data Nominal

Data nominal disebut juga data kategori karena angka pada data ini hanya sebagai label atau simbol dari objek. Contoh data nominal adalah variabel jenis kelamin yaitu laki-laki dan perempuan. Digunakan dua symbol untuk melambangkan variabel tersebut yaitu laki-laki = 1 dan perempuan = 2. Contoh data nominal yang terdiri dari lebih dari dua variabel yaitu data ras, agama, pekerjaan.

2. Data Ordinal

Data ordinal merupakan tingkatan data yang lebih tinggi. Angka pada data ordinal tidak hanya sebagai label tetapi juga menunjukkan peringkat. Setiap kelompok pada data ini bisa dibandingkan dengan hubungan lebih dari atau kurang dari. Contoh data ordinal adalah kepuasan konsumen terhadap suatu barang. Sangat puas = 1, puas = 2, kurang puas = 3, dan tidak puas = 4.

3. Data Interval

Data interval memiliki jarak yang sama antar datanya tetapi tidak memiliki nilai nol mutlak artinya pada saat nilai variabel nol bukan berarti tidak ada karakteristik yang diukur pada saat tersebut. Contoh data interval adalah variabel suhu dalam derajat Fahrenheit. Jarak antara 12° dan 13° sama dengan jarak antara 13° dan 14° .

4. Data Rasio

Data rasio merupakan tingkatan data yang paling tinggi. Data rasio memiliki sifat data nominal, ordinal, interval dan memiliki nilai absolute nol. Pada data rasio semua operasi matematika bisa dilakukan. Contoh data rasio adalah data tinggi badan, berat badan, umur, dll

2.2. Korelasi

Korelasi merupakan istilah statistika yang digunakan untuk mengukur kekuatan hubungan antar variabel yang digambarkan oleh besarnya koefisien korelasi. Koefisien korelasi berupa konstanta yang menggambarkan tingkat keeratan hubungan antar dua variabel. Besaran dari koefisien korelasi tidak menunjukkan hubungan sebab akibat antara dua variabel, tetapi hanya menggambarkan keterkaitan linier antar variabel (Walpole, 1992).

Ukuran hubungan linier antara dua variabel X dan Y dirumuskan sebagai berikut:

$$r_{x,y} = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{\sqrt{\left(n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right) \left(n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right)}} \quad (2.1)$$

di mana:

$r_{x,y}$: koefisien korelasi variabel X dan Y

n : banyaknya objek

Nilai koefisien korelasi ($r_{x,y}$) berkisar antara -1 sampai 1. Hubungan linier sempurna terjadi jika nilai $r_{x,y} = 1$ atau $r_{x,y} = -1$. Bila $r_{x,y}$ mendekati 1 atau -1 hubungan antara kedua peubah kuat dan dikatakan terdapat korelasi yang tinggi antar keduanya. Akan tetapi bila $r_{x,y}$ mendekati nol, hubungan linier antar X dan Y sangat lemah. Untuk $r_{x,y} = 0$, diartikan tidak terdapat hubungan linier antara X dan Y (Walpole, 1992).

Uji koefisien korelasi dilakukan untuk mengetahui adanya korelasi antar dua peubah dengan hipotesis sebagai berikut:

$H_0 : \rho = 0$ lawan

$H_1 : \rho \neq 0$

Dengan statistik uji sebagai berikut:

$$t_{hit} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

di mana:

r : koefisien korelasi

n : banyaknya observasi

Pengujian korelasi dilakukan menggunakan sebaran t dengan derajat bebas ($n - 2$) dan taraf nyata (α) tertentu dalam menentukan daerah kritik yang menjadi dasar pengambilan keputusan. Ketentuan dalam uji

koefisien korelasi ini adalah apabila $|t_{hit}| \leq t_{tabel}$ maka terima H_0 , artinya kedua variabel saling bebas, jika $|t_{hit}| > t_{tabel}$ maka tolak H_0 , artinya ada hubungan antara dua variabel (Walpole, 1992).

Berbagai *software* statistika khususnya untuk perhitungan nilai statistik uji biasanya disertai dengan suatu indikator pengambilan keputusan yaitu *p-value*. *P-value* dinyatakan sebagai taraf (keberartian) terkecil sehingga nilai uji statistik yang diamati masih berarti. Besarnya nilai α yang digunakan sebagai pembanding *p-value* tergantung pada peneliti mengenai seberapa resiko salah yang masih bisa ditolerir. Semakin kecil *p-value*, semakin kecil peluang kesalahan untuk menolak H_0 (Walpole dan Myers, 1995).

2.3. Pembakuan Data

Variabel-variabel yang memiliki satuan yang berbeda pada analisis kelompok akan berpengaruh pada hasil pengelompokan. Mengubah satuan variabel mengakibatkan perubahan matriks jarak pada analisis kelompok. Sebagai contoh terdapat 3 objek dengan 2 variabel (Y_1, Y_2): (2, 5), (4, 2), (7, 9), dengan menggunakan jarak *Euclid* didapatkan matriks $D = (d_{ij})$

$$D_1 = \begin{bmatrix} 0 & 3,6 & 6,4 \\ 3,6 & 0 & 7,6 \\ 6,4 & 7,6 & 0 \end{bmatrix}$$

jika satuan Y_1 diubah dari meter menjadi sentimeter maka matriks jarak menjadi:

$$D_2 = \begin{bmatrix} 0 & 200 & 500 \\ 200 & 0 & 300 \\ 500 & 300 & 0 \end{bmatrix}$$

Pada matriks jarak pertama jarak terbesar adalah jarak antara objek kedua dan ketiga (d_{23}), sedangkan pada matriks jarak kedua jarak terbesar adalah jarak antara objek pertama dan ketiga (d_{13}). Untuk mengatasi masalah tersebut, setiap variabel harus dibakukan terlebih dahulu sebelum membentuk matriks jarak (Rencher, 2002).

Pembakuan data yang umum adalah mengubah nilai setiap variabel ke nilai baku (dikenal dengan Z score) dengan cara mengurangkan nilai setiap variabel dengan nilai rata-ratanya dibagi dengan simpangan baku variabel tersebut. Proses ini mengubah nilai awal menjadi nilai baku dengan rata-rata 0 dan ragam 1. Transformasi

ini menghilangkan pengaruh perbedaan satuan variabel yang akan dianalisis (Hair *et al.*, 1998).

Rumus untuk menghitung nilai baku adalah sebagai berikut:

$$Z_i = \frac{x_i - \bar{x}}{s} \quad (2.2)$$

di mana: x_i : objek ke- i

$\bar{x}$: rata-rata

s : simpangan baku

Z_i : nilai baku objek ke- i

(Walpole, 1992)

2.4. Analisis Komponen Utama

Analisis komponen utama bertujuan menjelaskan struktur ragam peragam dari sekumpulan variabel melalui beberapa kombinasi linier variabel-variabel tersebut. Secara umum sasaran hasil dari analisis komponen utama adalah reduksi data untuk kebutuhan interpretasi (Johnson dan Wichern, 2022).

Analisis komponen utama merupakan suatu metode statistika yang digunakan untuk mentransformasi sekumpulan variabel menjadi variabel baru di mana variabel yang dihasilkan tidak saling berkorelasi dan merupakan kombinasi linier variabel asal. Variabel yang dihasilkan dinamakan komponen utama (Afifi dan Clark, 1990)

Pada analisis komponen utama walaupun terbentuk p komponen utama yang menjelaskan keragaman total, tetapi biasanya sebagian besar keragaman total bisa dijelaskan oleh k komponen utama di mana $k \leq p$. Analisis komponen utama bisa digunakan sebagai analisis awal untuk analisis yang lain. Hasil analisis komponen utama bisa digunakan sebagai masukan pada analisis regresi berganda dan analisis kelompok (Johnson dan Wichern, 2022).

Pada analisis komponen utama, vektor peubah asal yaitu $\mathbf{X} = (X_1, X_2, \dots, X_p)$ ditransformasi menjadi vektor peubah baru yaitu $\mathbf{Y} = (Y_1, Y_2, \dots, Y_p)$. Bentuk persamaan pada analisis komponen utama adalah sebagai berikut:

$$\begin{aligned} Y_1 &= \mathbf{a}'_1 \mathbf{X} = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \\ Y_2 &= \mathbf{a}'_2 \mathbf{X} = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p \\ &\vdots \\ Y_p &= \mathbf{a}'_p \mathbf{X} = a_{p1}X_1 + a_{p2}X_2 + \dots + a_{pp}X_p \end{aligned}$$

di mana $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p$ adalah vektor koefisien untuk setiap komponen utama yang bersesuaian. Rumus ragam dan peragam untuk peubah Y adalah:

$$\text{Var}(Y_i) = \mathbf{a}_i' \boldsymbol{\Sigma} \mathbf{a}_i \quad i = 1, 2, \dots, p$$

$$\text{Cov}(Y_i) = \mathbf{a}_i' \boldsymbol{\Sigma} \mathbf{a}_k \quad i, k = 1, 2, \dots, p$$

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

di mana $\boldsymbol{\Sigma}$ adalah matriks ragam peragam peubah asal. Mencari komponen utama $Y_i = \mathbf{a}_i' \mathbf{X}$ berarti mencari vektor koefisien $\mathbf{a}_i$ sedemikian sehingga $\mathbf{a}_i' \boldsymbol{\Sigma} \mathbf{a}_i$ adalah ragam komponen utama yang akan mencapai nilai maksimum dengan syarat bahwa $\mathbf{a}_i' \mathbf{a}_i = 1$. Permasalahan ini dapat diselesaikan menggunakan fungsi Lagrange.

$$\text{Maksimumkan} \quad : \quad \mathbf{a}_i' \boldsymbol{\Sigma} \mathbf{a}_i$$

$$\text{Batasan} \quad : \quad \mathbf{a}_i' \mathbf{a}_i = 1 \text{ atau } \mathbf{a}_i' \mathbf{a}_i - 1 = 0$$

Didapatkan fungsi Lagrange sebagai berikut :

$$L = \mathbf{a}_i' \boldsymbol{\Sigma} \mathbf{a}_i - \lambda_i (\mathbf{a}_i' \mathbf{a}_i - 1) \quad (2.3)$$

Apabila L diturunkan terhadap $\mathbf{a}_i$ kemudian disamakan dengan nol maka diperoleh:

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{a}} &= 2\mathbf{a}_i \boldsymbol{\Sigma} - 2\lambda_i \mathbf{a}_i = 0 \\ 2(\boldsymbol{\Sigma} - \lambda_i \mathbf{I}) \mathbf{a}_i &= 0 \\ (\boldsymbol{\Sigma} - \lambda_i \mathbf{I}) \mathbf{a}_i &= 0 \end{aligned} \quad (2.4)$$

Persamaan (2.4) akan menghasilkan jawaban nontrivial yang berarti hanya ada satu jawaban yang bersifat unik atau khas apabila matriks $(\boldsymbol{\Sigma} - \lambda_i \mathbf{I})$ merupakan matriks singular, yaitu determinan dari matriks tersebut sama dengan nol

$$|(\boldsymbol{\Sigma} - \lambda_i \mathbf{I})| = 0 \quad (2.5)$$

Persamaan (2.5) akan menghasilkan akar-akar karakteristik (nilai Eigen) yaitu $\lambda_1, \lambda_2, \dots, \lambda_p$ dimana $\lambda_1 > \lambda_2 > \dots > \lambda_p > 0$. Setiap akar karakteristik λ_i akan menentukan vektor karakteristik (vektor Eigen) $\mathbf{a}_i$.

Penentuan akar karakteristik yang akan digunakan pada komponen utama pertama yaitu:

$$\begin{aligned}(\Sigma - \lambda_i I) \mathbf{a}_i &= 0 \\ \Sigma \mathbf{a}_i - \lambda_i I \mathbf{a}_i &= 0 \\ \Sigma \mathbf{a}_i &= \lambda_i I \mathbf{a}_i\end{aligned}\quad (2.6)$$

Jika kedua sisi persamaan (2.6) digandaawalkan dengan $\mathbf{a}_i'$ maka diperoleh persamaan :

$$\mathbf{a}_i' \Sigma \mathbf{a}_i = \lambda_i \quad (2.7)$$

Dengan demikian dapat diketahui bahwa ragam setiap komponen utama bersesuaian dengan nilai setiap akar karakteristik. Persamaan (2.7) menunjukkan bahwa ragam komponen utama maksimum adalah akar karakteristik terbesar dari matriks Σ .

2.4.1. Matriks Masukan

Ada dua tipe masukan pada analisis komponen utama yaitu:

1. Matriks ragam-peragam

Bilamana unit satuan seluruh variabel yang akan dianalisis memiliki unit satuan sama maka matriks masukan yang digunakan adalah matriks ragam peragam. Misal Σ merupakan matriks kovariansi dari vektor acak $X' = [X_1 + X_2 + \dots X_p]$ dengan pasangan nilai eigen dan vektor eigen yang saling ortonormal adalah $(\lambda_1, \mathbf{a}_1), (\lambda_2, \mathbf{a}_2), \dots, (\lambda_p, \mathbf{a}_p)$ dimana $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ maka komponen utama ke- i didefinisikan sebagai berikut:

$$Y_i = \mathbf{a}_i' X = a_{i1} X_1 + a_{i2} X_2 + \dots + a_{ip} X_p, \quad i = 1, 2, \dots, p$$

2. Matriks korelasi

Bilaman variabel yang akan dianalisis memiliki perbedaan range yang sangat besar (mempunyai ukuran satuan variabel yang berbeda) maka matriks masukan yang digunakan adalah matriks korelasi. Dengan matriks masukan berupa matriks korelasi, maka data dari seluruh variabel ditransformasi ke dalam bentuk normal baku sehingga semua variabel akan memiliki unit satuan sama.

Komponen utama ke- i ; Y_i yang dibentuk berdasarkan variabel-variabel yang telah dibakukan $Z' = Z_1, Z_2, \dots, Z_p$ dengan $Cov(Z) = \rho$ didefinisikan sebagai berikut :

$$Y_i = a_i'Z = a_{i1}Z_1 + a_{i2}Z_2 + + a_{ip}Z_p, \quad i = 1, 2, \dots, p$$

2.4.2. Skor Komponen

Apabila komponen utama sudah terbentuk, tahap selanjutnya adalah menghitung skor komponen dari setiap individu yang akan digunakan untuk analisis lebih lanjut. Jika vektor skor dari individu ke-i adalah $(X_{i1}, X_{i2}, \dots, X_{ip})$ maka skor komponen dari individu ke-i pada komponen utama Y_j yang dihasilkan dari matriks ragam permang adalah:

$$SK_{ij} = (a_{1j}, a_{2j}, \dots, a_{pj}) \begin{bmatrix} X_{i1} - \bar{X}_1 \\ X_{i2} - \bar{X}_2 \\ \vdots \\ X_{ip} - \bar{X}_p \end{bmatrix}$$

atau $SK_{ij} = a_j^t (X_i - \bar{X})$ (2.8)

di mana:

SK_{ij} : skor komponen ke-j dari individu ke-i

a_j^t : vektor koefisien komponen utama ke-j

X_i : vektor data individu ke-i

$\bar{X}$: vector rata-rata variabel asal

Jika komponen utama dihasilkan dari matriks korelasi, maka rumus skor komponen utama dari individu ke-I adalah:

$$SK_{ij} = a_j^t D^{1/2} (X_i - \bar{X}) \quad (2.9)$$

di mana:

$$D^{1/2} : \text{diag} \left(\frac{1}{\sqrt{S_{11}}}, \dots, \frac{1}{\sqrt{S_{pp}}} \right)$$

S_{jj} menyatakan elemen diagonal ke-j dari S , $j = 1, 2, \dots, p$

2.4.3. Proporsi Keragaman

Setiap komponen utama memiliki nilai kepentingan relatif yaitu nilai yang menunjukkan seberapa besar pentingnya komponen utama ke-j diukur dari keragaman total yang dapat diterangkan oleh komponen

utama tersebut. Semakin besar nilai ini maka komponen utama tersebut dianggap semakin penting atau bermakna.

Besarnya kepentingan relatif adalah rasio antara ragam komponen utama ke- j dengan ragam total. Proporsi keragaman total yang dijelaskan oleh komponen utama ke- j adalah:

$$\frac{\lambda_j}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \times 100\%$$

(Johnson dan Wichern, 2002)

2.4.4. Penentuan Komponen Utama yang Digunakan

Interpretasi lebih lanjut didasarkan pada komponen pokok yang bermakna. Komponen-komponen pokok yang dianggap bermakna bilamana memenuhi kriteria tertentu. Marpaung dalam Astuti (2007) melakukan kompilasi terhadap beberapa pustaka berkenaan dengan penentuan komponen pokok yang dianggap bermakna, hasil tinjauan diberikan sebagai berikut :

1. Mengambil nilai eigen yang lebih besar dari 1 ($\lambda_j \geq 1$)
2. Memilih k buah komponen utama sebagai penyumbang terbesar terhadap keragaman data.

$$\frac{\sum_{j=1}^k \lambda_j}{\sum_{j=1}^p \lambda_j} > 0.75 \text{ dalam hal ini } p \text{ adalah jumlah variabel awal atau}$$

jumlah semua komponen utama yang dihasilkan.

Kriteria mana yang paling tepat digunakan tidak ada petunjuk jelas secara teoritis, sehingga dapat dipilih salah satu bilamana antar keduanya.

2.5. Analisis Kelompok

Analisis kelompok merupakan salah satu metode pengelompokan dengan cara membagi atau memecah objek-objek pada kelompok yang besar menjadi kelompok-kelompok yang lebih kecil di mana objek dalam kelompok relatif sama dan objek pada kelompok lain relatif tidak sama (Lattin *et al.*, 2006).

Secara garis besar metode pengelompokan pada analisis kelompok dibedakan menjadi dua yaitu metode pengelompokan hirarki dan nonhirarki. Metode pengelompokan hirarki digunakan apabila belum ada informasi jumlah kelompok. Sedangkan metode pengelompokan nonhirarki jumlah kelompok yang akan terbentuk sudah ditetapkan terlebih dahulu (Dillon dan Goldstein, 1984).

Pada metode pengelompokan hirarki terdapat dua metode pengelompokan yaitu metode *agglomerative* dan *divisive*. Algoritma metode *agglomerative* dimulai dengan setiap objek menjadi kelompok sehingga terdapat n kelompok, kemudian objek yang paling mirip membentuk kelompok, demikian seterusnya sampai terbentuk satu kelompok. Algoritma metode *divisive* merupakan kebalikan dari metode *agglomerative* yaitu pada awalnya semua objek membentuk satu kelompok, kemudian objek yang paling berbeda memisahkan diri dari kelompok sebelumnya, demikian seterusnya sampai terbentuk n kelompok (Timm, 2002).

Pada metode pengelompokan nonhirarki dimulai dengan memilih sejumlah nilai kelompok awal sesuai dengan jumlah yang diinginkan dan kemudian objek digabungkan ke dalam kelompok-kelompok tersebut. Pada metode pengelompokan nonhirarki terdapat tiga metode pengelompokan yaitu metode *sequential threshold*, *parallel threshold* dan *optimizing partitioning*. Metode *Sequential threshold* melakukan pengelompokan dengan terlebih dahulu memilih satu objek dasar yang akan dijadikan nilai awal kelompok, kemudian semua objek yang ada pada jarak terdekat dengan kelompok ini akan bergabung, kemudian dipilih kelompok kedua dan semua objek yang memiliki jarak paling dekat bergabung dengan kelompok ini, demikian seterusnya sampai terbentuk beberapa kelompok dengan semua objek didalamnya. Metode *parallel threshold* secara prinsip sama dengan metode *sequential threshold*, tetapi pemilihan objek awal kelompok dilakukan sekaligus, kemudian melakukan penggabungan objek ke dalam kelompok secara bersamaan. Metode *optimizing partitioning* adalah penggabungan dari metode *sequential threshold* dan *parallel threshold* dengan melakukan optimasi pada penempatan objek yang ditukar untuk kelompok lainnya dengan pertimbangan kriteria optimasi (Laboratorium Data Mining Jurusan Teknik Industri Universitas Islam Indonesia, 2009).

Hair *et al.* (1998), menyatakan konsep kesamaan adalah hal yang penting dalam analisis kelompok. Kesamaan antar objek adalah ukuran korespondensi antar objek. Ada tiga metode yang dapat diterapkan,

yaitu ukuran korelasi, ukuran jarak dan ukuran asosiasi. Ukuran korelasi dan jarak digunakan untuk data metrik (data interval dan rasio), sedangkan ukuran asosiasi digunakan untuk data nonmetrik (data nominal dan ordinal).

2.6. Analisis Kelompok Hirarki

Metode pengelompokan hirarki melibatkan penggabungan atau pemecahan secara berulang-ulang. Salah satu hal penting atau keistimewaan yang membedakan metode pengelompokan hirarki dari metode pengelompokan lainnya adalah penempatan objek dalam kelompok dilakukan secara pasti, artinya objek yang sudah berada pada suatu kelompok tidak akan pindah dan digabungkan dengan objek lain yang sudah berada pada kelompok lainnya. Metode pengelompokan hirarki dibedakan menjadi dua yaitu metode *agglomerative* dan *divisive* (Dillon dan Goldstein, 1984).

2.6.1. Metode *Agglomerative*

Menurut Rencher (2002), metode *agglomerative* dimulai dengan n kelompok, setiap kelompok terdiri dari satu objek. Pada tahap selanjutnya objek yang paling dekat membentuk kelompok kecil, proses tersebut berlanjut sampai semua objek berada pada satu kelompok. Berikut adalah algoritma pengelompokan metode *agglomerative*:

1. Mulai dengan n kelompok, setiap kelompok terdiri dari satu objek kemudian membentuk matriks jarak berukuran $n \times n$.
2. Cari matriks jarak untuk pasangan kelompok yang terdekat. Misalkan jarak terdekat yaitu jarak kelompok A dan B.
3. Gabungkan kelompok A dan B berdasarkan pautan yang telah ditentukan, label kelompok baru yaitu (AB).
4. Bentuk matriks jarak yang baru.
5. Ulangi langkah 2, 3, dan 4 sampai terbentuk satu kelompok.

Rencher (2002) memberikan lima metode pengelompokan *agglomerative* yaitu :

1. *Single Linkage*

Metode *single linkage* atau pautan tunggal didasarkan pada jarak minimum. Dimulai dengan dua objek yang memiliki jarak paling dekat keduanya akan ditempatkan pada kelompok pertama, langkah ini dilakukan terus sehingga seluruh objek membentuk satu kelompok.

$$D(A, B) = \min\{d(y_i, y_j)\} \quad (2.10)$$

2. *Complete Linkage*

Metode *complete linkage* atau peutan lengkap didasarkan pada jarak maksimum. Algoritma pengelompokan sama seperti metode pautan tunggal, tetapi didasarkan pada jarak terjauh.

$$D(A, B) = \max\{d(y_i, y_j)\} \quad (2.11)$$

3. *Average Linkage*

Metode *average linkage* atau pautan rata-rata didasarkan pada jarak rata-rata antar objek. Pengelompokan dimulai dari pasangan objek dengan jarak paling mendekati jarak rata-rata.

$$D(A, B) = \frac{1}{n_A n_B} \sum_{i=1}^{n_A} \sum_{j=1}^{n_B} (y_i, y_j) \quad (2.12)$$

4. *Metode Centroid*

Jarak antara dua kelompok adalah jarak terdekat antar centroid kelompok tersebut. Centroid kelompok adalah nilai tengah objek pada variabel dalam suatu set variabel kelompok.

$$D(A, B) = d(\bar{y}_A, \bar{y}_B) \quad (2.13)$$

$$\bar{y}_A = \sum_{i=1}^{n_A} \frac{y_i}{n_A}$$

$$\bar{y}_B = \sum_{i=1}^{n_B} \frac{y_i}{n_B}$$

Dimana $\bar{y}_A$ dan $\bar{y}_B$ adalah vektor rata-rata untuk vektor observasi pada A dan B.

5. *Metode Ward*

Metode ward menggunakan jumlah kuadrat jarak dalam kelompok dan antar kelompok. Jika AB adalah kombinasi dari kelompok A dan B, maka jumlah kuadrat jarak dalam kelompok adalah :

$$SSE_A = \sum_{i=1}^{n_A} (y_i - \bar{y}_A)(y_i - \bar{y}_A) \quad (2.14)$$

$$SSE_B = \sum_{i=1}^{n_B} (y_i - \bar{y}_B)(y_i - \bar{y}_B)$$

$$SSE_{AB} = \sum_{i=1}^{n_{AB}} (y_i - \bar{y}_{AB})(y_i - \bar{y}_{AB})$$

dimana $\bar{y}_{AB} = \frac{(n_A \bar{y}_A + n_B \bar{y}_B)}{(n_A + n_B)}$ dan $n_A, n_B, n_{AB} = n_A + n_B$ masing-masing merupakan jumlah objek pada kelompok A, B dan AB. Pada metode ini kelompok yang digabungkan yaitu kelompok yang memperkecil kenaikan jumlah kuadrat, didefinisikan sebagai berikut:

$$I_{AB} = SSE_{AB} - (SSE_A + SSE_B)$$

2.6.2. Metode *Divisive*

Dillon dan Goldstein (1984) menyatakan metode ini merupakan kebalikan dari metode *agglomerative*. Algoritma metode *divisive* dimulai dengan semua objek membentuk satu kelompok, pada langkah berikutnya objek yang paling berbeda memisahkan diri dengan kelompok tersebut, begitu seterusnya sampai terbentuk n kelompok.

2.6.3. Ukuran Kedekatan Jarak

Konsep dasar pengukuran pada analisis kelompok adalah konsep pengukuran jarak (*distance*) dan kesamaan (*similarity*). *Distance* adalah ukuran jarak pisah antar objek sedangkan *similarity* adalah ukuran kedekatan. Pengukuran jarak (*distance type measure*) digunakan untuk data yang bersifat metrik (Dillon dan Goldstein, 1984).

Giudici (2003) menjelaskan bahwa jika x dan y adalah baris dari matriks data maka $d(x, y)$ merupakan jarak antara dua objek jika dan hanya jika memenuhi syarat sebagai berikut :

1. Nonnegatif : $d(x, y) \geq 0$ untuk semua x dan y
2. Identitas : $d(x, y) = 0 \leftrightarrow x = y$ untuk semua x dan y
3. Simetris : $d(x, y) = d(y, x)$ untuk semua x dan y
4. Ketidaksamaan segitiga : $d(x, y) \leq d(x, z) + d(y, z)$ untuk semua x, y dan z

Jarak antar objek bisanya disajikan dalam bentuk matriks jarak seperti berikut:

$$D = \begin{bmatrix} 0 & d_{12} & \dots & d_{1n} \\ d_{21} & 0 & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \dots & 0 \end{bmatrix}$$

Ukuran kedekatan jarak merupakan ukuran yang paling sering digunakan, diterapkan untuk data berskala metrik. Ukuran kedekatan jarak yang sering digunakan pada analisis kelompok adalah ukuran jarak *Euclid*. Jarak *Euclid* adalah metrika yang paling sering digunakan untuk menghitung kesamaan dua vektor. Jarak *Euclid* menghitung akar dari kuadrat perbedaan dua vektor. Jika data dinyatakan dalam bentuk matriks X yang anggotanya X_{ij} , maka jarak *Euclid* didefinisikan sebagai berikut:

$$d(i, j) = \sqrt{\sum_{k=1}^p (X_{ik} - X_{jk})^2}, i \neq j \quad (2.15)$$

di mana:

$d(i, j)$: jarak *Euclid* objek data ke- i dan objek data ke- j pada peubah ke- k

p : banyaknya peubah

X_{ik} : data ke- i pada peubah ke- k

X_{jk} : data ke- j pada peubah ke- k

2.7. Estimasi Jumlah Kelompok

Estimasi adalah perkiraan mengenai nilai (*value*), jumlah (*amount*), ukuran (*size*), atau berat (*weight*) dari sesuatu. Dalam konteks analisis kelompok, estimasi jumlah kelompok adalah perkiraan mengenai jumlah kelompok optimal yang akan terbentuk dari suatu himpunan data. Pada analisis kelompok, jika tidak ada dugaan mengenai jumlah kelompok optimal, diperlukan suatu metode validitas kelompok untuk mengukur kebaikan hasil analisis kelompok. Indeks validitas kelompok menunjukkan kualitas hasil analisis kelompok, indeks ini mengukur kecukupan struktur dalam mendapatkan kelompok hasil analisis. Kecukupan struktur kelompok berarti struktur kelompok memberikan informasi sebenarnya mengenai data (Bolshakova dan Azuaje, 2008).

2.7.1. Gap statistic

Tibshirani *et al.* (2001), menjelaskan prosedur *gap statistic* sebagai salah satu metode estimasi jumlah kelompok optimal. Misalkan X_{ij} adalah pengamatan pada objek ke- i dan variabel ke- j . Objek data dikelompokkan menjadi k kelompok yaitu $C_1, C_2, \dots, C_k$ dengan C_r

merupakan pengamatan pada kelompok ke- r dan n_r adalah banyaknya objek pada kelompok ke- r . Jika $d_{ii'}$ merupakan jarak antara pengamatan ke- i dan ke- i' maka didefinisikan :

$$D_r = \sum_{i, i' \in C_r} d_{ii'}$$

Merupakan jumlah jarak semua titik dalam kelompok r

$$W_k = \sum_{r=1}^k \frac{1}{2n_r} D_r \quad (2.16)$$

Merupakan jumlah kuadrat gabungan dalam k kelompok.

Menurut Tibshirani *et al.* (2001), algoritma *gap statistic* adalah sebagai berikut :

1. Mengelompokan data dan mengubah-ubah banyaknya kelompok mulai dari $k = 1, 2, \dots, K$ yang masing-masing memberikan ukuran dispersi dalam kelompok W_k
2. Membangkitkan B rujukan data (resampling) menggunakan distribusi referensi variabel secara uniform pada daerah nilai hasil observasi variabel. Setiap himpunan data yang dibangkitkan kemudian dikelompokan dan setiap kelompok akan memberikan ukuran dispersi dalam kelompok W_{kb}^* dimana $b = 1, 2, \dots, B$ dan $k = 1, 2, \dots, K$. Perhitungan estimasi *gap statistic* adalah:

$$Gap(k) = \left(\frac{1}{b} \right) \sum_b \log(W_{kb}^*) - \log(W_k) \quad (2.17)$$

3. Misalkan $l = \left(\frac{1}{b} \right) \sum_b \log(W_{kb}^*)$, maka dapat dihitung standar deviasinya yaitu:

$$sd_k = \sqrt{\left(\frac{1}{b} \right) \sum_b \{ \log(W_{kb}^*) - l \}^2}$$

$$\text{Dan didefinisikan } s_k = sd_k \sqrt{\left(1 + \frac{1}{B} \right)}$$

Jumlah kelompok (k) optimum diestimasi melalui nilai k terkecil sedemikian sehingga:

$$Gap(k) \geq Gap(k+1) - s_{k+1}$$

2.7.2. Hartigan

Hartigan dalam Tibshirani *et al.* (2001) mengusulkan indeks untuk menentukan jumlah kelompok optimal yang dilambangkan $H(k)$ yang dirumuskan sebagai berikut:

$$H(k) = \left(\frac{\text{tr}(W_k)}{\text{tr}(W_{k+1})} - 1 \right) (n - k - 1) \quad (2.18)$$

$$W_k = \sum_{r=1}^k \sum_{i=1}^{n_r} (x_{r_i} - \bar{x}_r)(x_{r_i} - \bar{x}_r)^T$$

di mana:

- $\text{tr}(W_k)$: trace matriks jumlah kuadrat dalam kelompok
- k : banyaknya kelompok
- x_{r_i} : objek ke- i pada kelompok ke- r
- $\bar{x}_r$: rata-rata dalam kelompok ke- r
- n_r : banyaknya objek pada kelompok ke- r

Hartigan menyarankan bahwa kelompok ditambah jika $H(k) > 10$. Jumlah kelompok (k) optimal diestimasi melalui nilai $k \geq 1$ terkecil sedemikian sehingga besarnya $H(k) \leq 10$.

2.7.3. Krzanowski-Lai

Krzanowski dan Lai dalam Tibshirani *et al.* (2001) mengusulkan indeks untuk menentukan jumlah kelompok optimal yang dilambangkan $KL(k)$.

$$DIFF(k) = (k-1)^{\frac{2}{p}} \text{tr}(W_{k+1}) - k^{\frac{2}{p}} \text{tr}(W_k)$$

di mana:

- $\text{tr}(W_k)$: trace matriks jumlah kuadrat dalam kelompok
- k : banyaknya kelompok

dipilih k yang memaksimumkan nilai:

$$KL = \left| \frac{DIFF(k)}{DIFF(k+1)} \right| \quad (2.19)$$

2.7.4. Silhouette

Kauffman dan Rousseeuw dalam Tibshirani *et al.* (2001) mengusulkan statistik *silhouette* yang dilambangkan $s(i)$ untuk menentukan jumlah kelompok optimal yang dirumuskan sebagai berikut:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2.20)$$

di mana:

$a(i)$: rata-rata jarak objek ke- i ke objek lain dalam kelompok

$b(i)$: minimum dari rata-rata jarak objek ke- i dengan semua anggota dari kelompok yang lain.

Jumlah kelompok (k) optimal adalah k yang memaksimumkan rata-rata $s(i)$

2.7.5. Calinski-Harabasz

Calinski dan Harabasz dalam Tibshirani *et al.* (2001) mengusulkan indeks untuk menentukan jumlah kelompok optimal yang dilambangkan $CH(k)$ yang dirumuskan sebagai berikut:

$$CH(k) = \frac{\text{tr}(B_k) / (k-1)}{\text{tr}(W_k) / (n-k)} \quad (2.21)$$

$$B_k = \sum_{r=1}^k n_r (\bar{x}_r - \bar{x})(\bar{x}_r - \bar{x})^T$$

di mana:

$\text{tr}(W_k)$: trace matriks jumlah kuadrat dalam kelompok

$\text{tr}(B_k)$: trace matriks jumlah kuadrat antar kelompok

k : banyaknya kelompok

$\bar{x}$: rata-rata total

n : banyaknya objek

Jumlah kelompok (k) optimal adalah k yang memaksimumkan $CH(k)$.

2.8. Cluster Tightness Measure (CTM)

Epps dan Ambikairajah (2008) merumuskan *Cluster Tightness Measure* (CTM) sebagai ukuran kebaikan dari hasil pengelompokan berdasarkan simpangan baku setiap peubah pada masing-masing kelompok, yang dirumuskan sebagai berikut:

$$CTM = \frac{1}{k} \sum_{k=1}^K \left(\frac{1}{p} \sum_{j=1}^p \frac{{}_k\sigma_j}{\sigma_j} \right) \quad (2.22)$$

di mana:

${}_k\sigma_j$: simpangan baku pada kelompok ke- k untuk peubah ke- j

σ_j : simpangan baku seluruh data untuk peubah ke- j

p : banyaknya peubah

K : banyaknya kelompok

Kelompok yang terbentuk dikatakan baik jika memiliki nilai CTM kecil.



UNIVERSITAS BRAWIJAYA



BAB III METODE PENELITIAN

3.1. Data Penelitian

Data yang digunakan pada penelitian ini adalah 50 data sekunder yang diambil dari beberapa penelitian. Data penelitian secara lengkap dapat dilihat pada Lampiran 1 dengan klasifikasi pada Tabel 3.1.

Tabel 3.1. Data Penelitian

Data	Deskripsi	Jumlah Objek	Jumlah Variabel
1	Pengelompokan stasiun pengambilan sampel di Sungai Code Yogyakarta berdasarkan konsentrasi gamma radionuklida (Taftazani, 2006, Pusat Teknologi Akselerator dan Proses Bahan (PTAPB)-BATAN: Lampiran 1; Data 1)	11	9
2	Data kadar nutrisi plasma nutfah jagung (Yasin <i>et al.</i> , 2005, Balai Penelitian Tanaman Serealia Maros: Lampiran 1; Data 2)	12	7
3	Data beban pencemaran udara pada sektor industri di Jawa Tengah tahun 2005 (dalam ton) (Purwaningsih, 2007, Skripsi, Universitas Negeri Semarang: Lampiran 1; Data 3)	14	5
4	Data rata-rata angka criminal kota per 100000 populasi (Rancher, 2002, John Willey and Sons. New York: Lampiran 1; Data 4)	16	7
5	Data variabel unsur indikator kemiskinan di Kecamatan Kras Kabupaten Kediri bulan Juli 2006 (Fitrianingsih, 2003, Laporan PKL, Universitas Brawijaya Malang: Lampiran 1; Data 5)	16	4
6	Data karakteristik beberapa varietas padi (Rubiyo, 2005, Balai Pengkajian Teknologi Pertanian Bali: Lampiran 1; Data 6)	16	4

Tabel 3.1. (Lanjutan)

Data	Deskripsi	Jumlah Objek	Jumlah Variabel
7	Data tinggi tanaman, jumlah anakan maksimum, dan jumlah anakan produktif beberapa varietas padi (Rubiyo, 2005, Balai Pengkajian Teknologi Pertanian Bali: Lampiran 1; Data 7)	16	3
8	Data panjang malai, bobot kering panen, dan bobot 1000 butir gabah beberapa varietas padi (Rubiyo, 2005, Balai Pengkajian Teknologi Pertanian Bali: Lampiran 1; Data 8)	16	3
9	Data kandungan protein, lisin, dan triptofan beberapa varietas jagung QPM (Yasin <i>et al.</i> , 2005, Balai Penelitian Tanaman Serealia Maros: Lampiran 1; Data 9)	16	3
10	Respon beberapa genotip kedelai (<i>Glycinemax</i> (L.) Merrill) terhadap genangan (Anwar, 2008, Skripsi, Universitas Brawijaya Malang: Lampiran 1; data 10)	17	6
11	Data indikator kemiskinan di Kabupaten Bangkalan (Krismayanti, 2003, Skripsi, Universitas Brawijaya Malang: Lampiran 1; Data 11)	18	4
12	Data karakter agronomi pada kondisi tercekam (perlakuan) dan bobot kering akar 18 varietas padi sawah (Munir <i>et al.</i> , 2009, BPTP Sukarami Sumatera Barat: Lampiran 1; Data 12)	18	3
13	Data karakter agronomi 18 varietas padi pada sawah lahan rawa (Munir <i>et al.</i> , 2009, BPTP Sukarami Sumatera Barat: Lampiran 1; Data 13)	18	3

Tabel 3.1. (Lanjutan)

Data	Deskripsi	Jumlah Objek	Jumlah Variabel
14	Data persentase tenaga kerja di bidang pertanian, industri dan perusahaan jasa di beberapa negara (http://lib.stat.cmu.edu/DASL/Datafiles/LaborForce.html : Lampiran 1; Data 14)	19	3
15	Data tingkat produksi dan penggunaan faktor produksi usaha tani jagung pada petani yang mengikuti program PHBM (Andress, 2003, Skripsi, Universitas Brawijaya Malang: Lampiran ; Data 15)	20	2
16	Daftar analisis bahan makanan lauk pauk nabati (per 100 gram bahan) (Yonathan, 1999, Skripsi, Universitas Brawijaya Malang: Lampiran 1; Data 16)	20	3
17	Data pengeluaran untuk periklanan di beberapa perusahaan (http://lib.stat.cmu.edu/DASL/Datafiles/tvadsdat.html : Lampiran 1; Data 17)	21	2
18	Data tingkat kepekaan kuman <i>Escherichia Coli</i> terhadap beberapa antimikroba di RSUD DR Saiful Anwar Malang (Wicaksono, 1999, Skripsi, Universitas Brawijaya Malang: Lampiran; Data 18)	24	3
19	Data konsumsi protein di beberapa negara eropa (http://lib.stat.cmu.edu/DASL/Datafiles/Protein.html : Lampiran 1; Data 19)	25	9
20	Data rata-rata tinggi tanaman dan jumlah polong per tanaman dari hasil penelitian interaksi genotip x lingkungan galur-galur kacang tanah (<i>Arachis hypogala L.</i>) (Wibowo, 2003, Skripsi, Universitas Brawijaya Malang: Lampiran 1; Data 20)	25	3

Tabel 3.1. (Lanjutan)

Data	Deskripsi	Jumlah Objek	Jumlah Variabel
21	Data persentase tenaga kerja industri di beberapa negara eropa (http://lib.stat.cmu.edu/DASL/Datafiles/EuropeanJobs.html : Lampiran 1; Data 21)	26	9
22	Data karakteristik industri kerajinan rumah tangga di 27 provinsi di Indonesia tahun 1996 (Purnanto, 2003, Skripsi, Universitas Brawijaya Malang: Lampiran 1; Data 22)	27	5
23	Data penduduk Surabaya berdasarkan tingkat sosial budaya di 28 kecamatan (Andayani, 2004, Skripsi, Institut Teknologi Sepuluh November Surabaya: Lampiran 1; Data 23)	28	8
24	Data gaji dan total penerimaan tim sepak bola (http://lib.stat.cmu.edu/DASL/Datafiles/qbacksalarydat.html : Lampiran ; Data 24)	28	2
25	Data varietas jagung inbrida (Wijaya, 2003, Skripsi, Universitas Bina Nusantara: Lampiran 1; Data 25)	30	5
26	Nilai rata- rata hasil ujian nasional 2008 Madrasah Aliyah Negeri program IPS (Sjafrudin, 2008, http://pendis.depag.go.id/file/dokumen/analisis200803.pdf : Lampiran 1; Data 26)	32	6
27	Data rata-rata konsumsi protein perkapita sehari (gram) penduduk Kabupaten Malang menurut kelompok makanan tahun 2007 (Sulistiyowati, 2009, Laporan PKL, Universitas Brawijaya Malang: Lampiran 1; Data 27)	33	8
28	Nilai rata- rata hasil ujian nasional 2008 Madrasah Tsanawiyah (Sjafrudin, 2008, http://pendis.depag.go.id/file/dokumen/analisis200803.pdf : Lampiran 1; Data 28)	33	4

Tabel 3.1. (Lanjutan)

Data	Deskripsi	Jumlah Objek	Jumlah Variabel
29	Data penduduk berumur 10 tahun ke atas menurut kabupaten/ kota (Cahyono, 2005, Laporan PKL, Universitas Brawijaya Malang: Lampiran 1; Data 29)	35	5
30	Data curah hujan harian dari 36 pos pengamatan hujan selama 12 periode di wilayah Kabupaten Indramayu (Haryoko, 2007, Badan Meteorologi dan Geofisika: Lampiran 1; Data 30)	36	3
31	Data karakteristik berbagai merk mobil (http://lib.stat.cmu.edu/DASL/Datafiles/Cars.html ; Lampiran 1; Data 31)	38	5
32	Data indikator kekuatan tekanan semen (Suarto, 2004, Skripsi, Universitas Brawijaya Malang: Lampiran 1; Data 32)	39	6
33	Data luas daerah, ukuran biji dan tinggi tanaman beberapa spesies pohon oak (http://lib.stat.cmu.edu/DASL/Datafiles/Acorn.html ; Lampiran 1; Data 33)	39	3
34	Data luas wilayah, jumlah penduduk, kepadatan penduduk dan rumah tangga miskin di Kabupaten Bogor tahun 2008 (Triana, 2009, Skripsi, Institut Pertanian Bogor: Lampiran 1; Data 34)	40	4
35	Data polusi udara 41 kota di Amerika Serikat (http://lib.stat.cmu.edu/DASL/Datafiles/Airpollution.html ; Lampiran 1; Data 35)	41	6
36	Data angka kematian perseribu populasi di beberapa negara (http://lib.stat.cmu.edu/DASL/Datafiles/cigcancerdat.html : Lampiran 1 ; Data 36)	44	4

Tabel 3.1. (Lanjutan)

Data	Deskripsi	Jumlah Objek	Jumlah Variabel
37	Data rata-rata konsumsi alkohol tiap daerah di Jepang dengan 5 jenis minuman beralkohol (Wijayanto <i>et al.</i> , 1998, Forum Statistika dan Komputasi: Lampiran 1: Data 37)	46	5
38	Data pendapatan perkapita, belanja umum dan karakteristik ekonomi beberapa negara (http://lib.stat.cmu.edu/DASL/Datafiles/pubexpendat.htm ; Lampiran 1; Data 38)	48	4
39	Data rata-rata gaji profesor 50 universitas di Amerika Serikat (http://lib.stat.cmu.edu/DASL/Datafiles/facultysalaries.html ; Lampiran 1; Data 39)	49	3
40	Data plasma nutfah kapas (Sa'adah, 2009, Laporan PKL, Universitas Brawijaya Malang; Lampiran 1; Data 40)	50	6
41	Data pelanggan PLN di kecamatan Sleman (Hayati, 2008, Skripsi, Universitas Islam Indonesia Yogyakarta; Lampiran 1; Data 41)	50	4
42	Data rata-rata intensitas serangan, tinggi tanaman, rasio BK/BB dari hasil penelitian uji ketahanan 50 varietas kacang hijau terhadap Black Gram Mottle Virus (Utami, 2008, Skripsi, Universitas Brawijaya Malang; Lampiran 1; Data 42)	50	3
43	Data jumlah polong, tinggi tanaman dan hasil genotip terpilih beberapa varietas kacang tanah (Azzahra, 2002, Balai Penelitian Pertanian Lahan Rawa; Lampiran 1; Data 43)	51	3

Tabel 3.1. (Lanjutan)

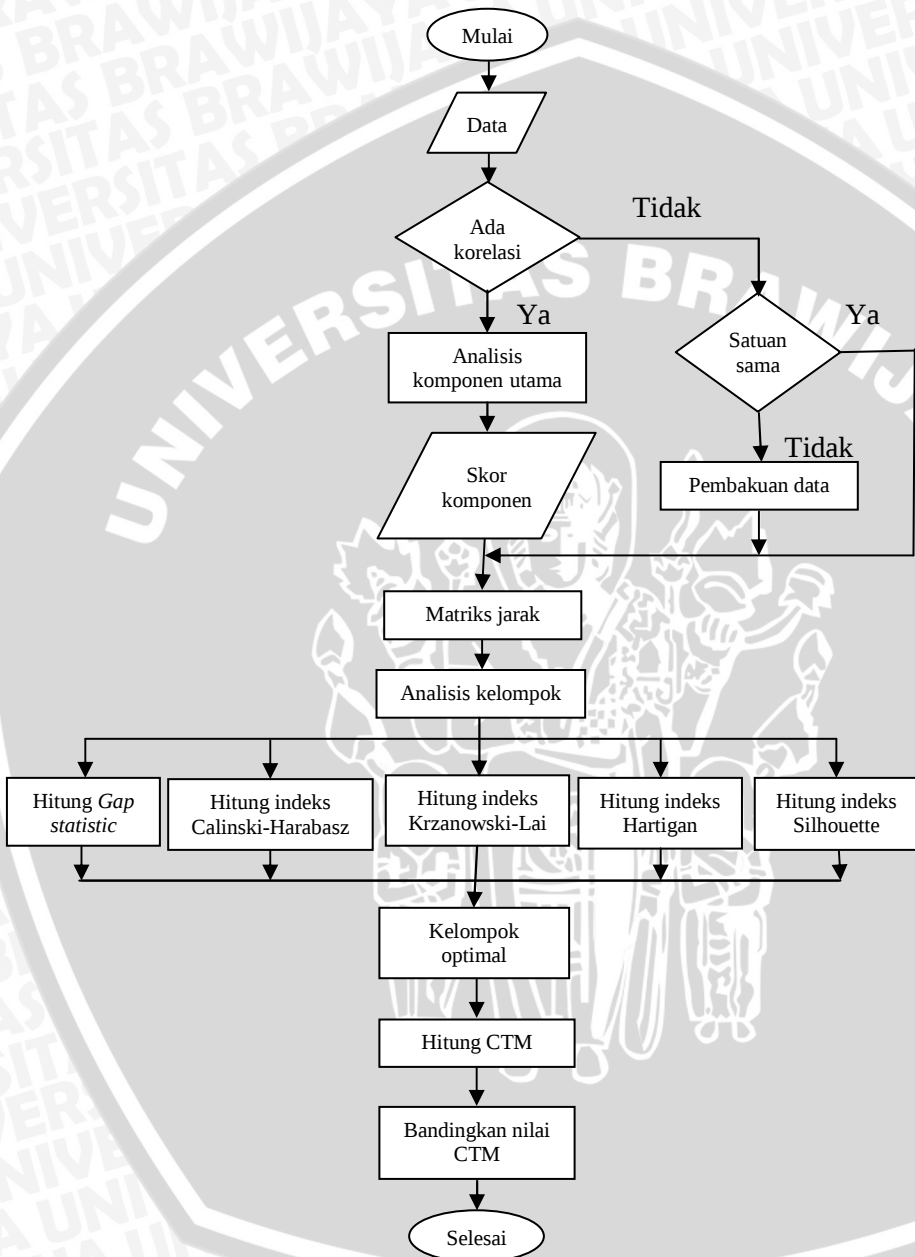
Data	Deskripsi	Jumlah Objek	Jumlah Variabel
44	Data rata-rata gaji pengajar dan pengeluaran persiswa di 50 negara bagian dan daerah Kolombia (http://lib.stat.cmu.edu/DASL/Datafiles/EducationalSpending.html : Lampiran 1; Data 44)	51	2
45	Data rata-rata gaji guru dan pengeluaran setiap siswa di 50 negara bagian dan daerah Kolombia (http://lib.stat.cmu.edu/DASL/Datafiles/teacherpaydat.html : Lampiran 1; Data 44)	51	2
46	Data karakteristik 53 danau di Florida (http://lib.stat.cmu.edu/DASL/Datafiles/mercuryinbass.html : Lampiran 1; Data 46)	53	5
47	Data karakteristik beberapa varietas kacang tanah (Lestari, 2009, PKL, Universitas Brawijaya Malang: Lampiran 1; Data 47)	58	6
48	Data karakteristik beberapa merek sereal (http://lib.stat.cmu.edu/DASL/Datafiles/cereals.html : Lampiran 1; Data 48)	69	8
49	Data industri tempe di kota Malang (Hermayudi, 2004, Skripsi, Universitas Brawijaya Malang: Lampiran 1; Data 49)	74	3
50	Data karakteristik beberapa perusahaan (http://lib.stat.cmu.edu/DASL/Datafiles/companies.html : Lampiran 1 ; Data 50)	79	6

3.2. Metode Penelitian

Tahapan-tahapan analisis pada penelitian ini adalah sebagai berikut:

1. Melakukan uji korelasi seperti yang dijelaskan pada sub bab 2.2. jika ada korelasi lakukan analisis komponen utama seperti yang dijelaskan pada sub bab 2.4, jika tidak terdapat korelasi langsung kelangkah 2.
2. Melakukan pembakuan data jika satuan data tidak sama menggunakan persamaan (2.2).
3. Membentuk matriks jarak menggunakan persamaan (2.15)
4. Melakukan analisis kelompok hirarki metode *single linkage* menggunakan persamaan (2.10)
5. Menghitung *gap statistic* menggunakan persamaan (2.17), indeks Hartigan menggunakan persamaan (2.18), indeks Krzanowski-Lai menggunakan persamaan (2.19), indeks Silhouette menggunakan persamaan (2.20), dan indeks Calinski-Harabasz menggunakan persamaan (2.21).
6. Menentukan jumlah kelompok optimal berdasarkan kelima kreteria di atas.
7. Menghitung nilai CTM kelima indeks validitas menggunakan persamaan (2.22).
8. Membandingkan nilai CTM untuk kelima indek validitas.

Perhitungan pada analisis ini menggunakan *software* SPSS versi 15, Minitab 13, R. 2.7. 2, dan *Microsoft Excel 2007*. *Software* SPSS versi 15 digunakan untuk uji koefisien korelasi, Minitab 13 digunakan untuk analisis komponen utama, R. 2.7. 2 digunakan untuk analisis kelompok dan menghitung indeks validitas, dan *Microsoft Excel 2007* digunakan untuk menghitung nilai CTM. Tahapan-tahapan pada penelitian ini dapat dilihat pada gambar berikut ini:



Gambar 3.1 Diagram Alir Metode Penelitian

UNIVERSITAS BRAWIJAYA



BAB IV

HASIL DAN PEMBAHASAN

4.1. Pengujian Koefisien Korelasi

Seperti pada analisis yang lain, pada analisis kelompok juga harus memenuhi asumsi yaitu tidak terdapat korelasi antar peubah. Ada atau tidaknya korelasi sangat berpengaruh pada hasil pengelompokan sehingga masalah korelasi sangat diperhatikan pada analisis kelompok. Sebelum analisis kelompok dilakukan terlebih dahulu harus dilakukan pengujian koefisien korelasi. Hasil pengujian koefisien korelasi untuk semua data penelitian disajikan pada Tabel 4.1.

Tabel 4.1. Pengujian Koefisien Korelasi

Data	Uji koefisien korelasi	Data	Uji koefisien korelasi
1	Ada korelasi	26	Ada korelasi
2	Tidak ada korelasi	27	Ada korelasi
3	Ada korelasi	28	Ada korelasi
4	Ada korelasi	29	Ada korelasi
5	Ada korelasi	30	Ada korelasi
6	Ada korelasi	31	Ada korelasi
7	Ada korelasi	32	Ada korelasi
8	Tidak ada korelasi	33	Tidak ada korelasi
9	Ada korelasi	34	Ada korelasi
10	Ada korelasi	35	Ada korelasi
11	Ada korelasi	36	Ada korelasi
12	Ada korelasi	37	Ada korelasi
13	Ada korelasi	38	Ada korelasi
14	Ada korelasi	39	Ada korelasi
15	Tidak ada korelasi	40	Ada korelasi
16	Tidak ada korelasi	41	Ada korelasi
17	Ada korelasi	42	Tidak ada korelasi
18	Ada korelasi	43	Ada korelasi
19	Ada korelasi	44	Ada korelasi
20	Ada korelasi	45	Ada korelasi

Tabel 4.1. (Lanjutan)

Data	Uji koefisien korelasi	Data	Uji koefisien korelasi
21	Ada korelasi	46	Ada korelasi
22	Ada korelasi	47	Ada korelasi
23	Ada korelasi	48	Ada korelasi
24	Ada korelasi	49	Ada korelasi
25	Ada korelasi	50	Ada korelasi

Tabel 4.1 merupakan ringkasan dari hasil uji koefisien korelasi semua data penelitian. Uji koefisien korelasi secara lengkap dapat dilihat pada Lampiran 2.

Pada Tabel 4.1 dapat dilihat bahwa 88% (44 dari 50 data penelitian) terdapat korelasi antar peubah. Sebagian besar data penelitian yang digunakan berkorelasi, sehingga perlu dilakukan analisis komponen utama untuk mereduksi data sehingga didapatkan skor komponen sebagai masukan pada analisis kelompok. Pada data penelitian ini terdapat enam data yang tidak berkorelasi yaitu data 2, 8, 15, 16, 33, dan 42. Pada data yang tidak berkorelasi, analisis kelompok bisa dilakukan secara langsung dengan memperhatikan satuan data. Jika satuan data sama maka langsung membentuk matriks jarak, jika satuan data tidak sama terlebih dahulu dilakukan pembakuan data. Pada data penelitian ini data 2, 15, dan 16 tidak berkorelasi dan memiliki satuan yang sama sehingga bisa langsung dilakukan analisis kelompok, sedangkan pada data 7, 33, dan 42 tidak berkorelasi dan satuan data tidak sama sehingga perlu dilakukan pembakuan data terlebih dahulu.

4.2. Analisis Komponen Utama

Masalah korelasi pada analisis kelompok dapat diatasi dengan analisis komponen utama. Analisis komponen utama bertujuan menjelaskan struktur ragam peragam dari sekumpulan peubah melalui beberapa peubah baru dimana peubah baru saling bebas dan merupakan kombinasi linier dari peubah asal. Sebagai contoh, pada data 4 yaitu data rata-rata angka kriminal kota per 100000 populasi. Hasil analisis komponen utama untuk data 4 disajikan pada Tabel 4.2.

Tabel 4.2. Hasil Analisis Komponen Utama Data 4

Peubah	Komponen Utama						
	K ₁	K ₂	K ₃	K ₄	K ₅	K ₆	K ₇
X ₁	-0.284	0.601	0.291	0.318	0.443	0.343	0.242
X ₂	-0.434	0.058	0.281	-0.607	0.009	0.240	-0.551
X ₃	-0.387	0.193	-0.454	0.506	-0.390	0.071	-0.439
X ₄	-0.459	0.265	0.090	-0.236	-0.455	-0.459	0.487
X ₅	-0.388	-0.404	0.294	0.336	0.391	-0.547	-0.188
X ₆	-0.346	-0.597	0.126	0.139	-0.222	0.557	0.360
X ₇	-0.316	-0.092	-0.721	-0.296	0.492	0.004	0.207
Nilai eigen	3.462	1.330	0.958	0.611	0.361	0.170	0.108
Proporsi keragaman (%)	0.495	0.190	0.137	0.087	0.052	0.024	0.015
Keragaman Kumulatif (%)	0.495	0.685	0.821	0.909	0.960	0.985	1.000

Tabel 4.2 merupakan ringkasan hasil analisis komponen utama data 4. Hasil analisis komponen utama secara lengkap dapat dilihat pada Lampiran 3.

Pada Tabel 4.2 dapat dilihat bahwa 82.1% keragaman total dapat dijelaskan oleh tiga komponen utama yaitu K₁, K₂, dan K₃. Persamaan komponen utama pertama adalah sebagai berikut :

$$K_1 = -0.284X_1 - 0.434X_2 - 0.387X_3 - 0.459X_4 - 0.388X_5 - 0.346X_6 - 0.316X_7$$

$$K_2 = 0.601X_1 + 0.058X_2 + 0.193X_3 + 0.265X_4 - 0.404X_5 - 0.597X_6 - 0.092X_7$$

$$K_3 = 0.291X_1 + 0.281X_2 - 0.454X_3 + 0.090X_4 + 0.294X_5 + 0.126X_6 - 0.721X_7$$

Skor komponen pada analisis komponen utama merupakan masukan pada analisis kelompok. Pada data penelitian ini, terdapat 44 data penelitian yang menggunakan skor komponen sebagai masukan pada analisis kelompok.

4.3. Analisis Kelompok Hirarki

Data penelitian yang sudah tidak berkorelasi dan memiliki satuan yang sama, selanjutnya dilakukan analisis kelompok hirarki metode *single linkage* menggunakan jarak *Euclid*. Pada penelitian ini penentuan jumlah kelompok diestimasi menggunakan lima indeks validitas dan penentuan indeks validitas terbaik berdasarkan nilai CTM. Jumlah kelompok dan nilai CTM yang dihasilkan oleh kelima indeks validitas tersebut disajikan pada Tabel 4.3.

Tabel 4.3. Jumlah Kelompok dan nilai CTM yang dihasilkan oleh Indeks Validitas *Gap statistic* (Gap), Hartigan (H), Krzanowski-Lai (KL), Silhouette (S), dan Calinski-Harabasz (CH).

Data	Gap		H		KL		S		CH	
	k	Nilai CTM	k	Nilai CTM	k	Nilai CTM	k	Nilai CTM	k	Nilai CTM
1	2	0.3829	2	0.3829	5	0.0649	5	0.0949	5	0.0649
2	4	0.3153	2	0.7352	2	0.7352	2	0.7352	2	0.7352
3	2	0.4812	3	0.1814	2	0.4812	4	0.0213	6	0.0029
4	2	0.4747	2	0.4747	5	0.2006	6	0.1824	5	0.2006
5	2	0.3895	2	0.3895	5	0.1248	2	0.3895	3	0.2109
6	2	0.4217	2	0.4217	3	0.0409	2	0.4217	2	0.4217
7	2	0.4217	2	0.4217	6	0.0409	6	0.0409	6	0.0409
8	2	0.3978	2	0.3978	4	0.1428	3	0.2022	3	0.2022
9	2	0.5025	3	0.5058	5	0.5295	6	0.1755	5	0.5295
10	3	0.2519	3	0.2519	3	0.2519	2	0.3153	7	0.0414
11	2	0.4139	2	0.4139	3	0.3175	2	0.4139	3	0.3175
12	2	0.7848	2	0.7848	4	0.5299	2	0.7848	6	0.2979
13	2	0.3139	3	0.2341	3	0.2341	2	0.3139	5	0.1184
14	2	0.2279	4	0.1042	4	0.1042	2	0.2279	4	0.1042
15	2	0.4409	3	0.3684	4	0.1455	7	0.0429	7	0.0429
16	2	0.5066	2	0.5066	2	0.5066	7	0.0869	7	0.0869
17	2	0.483	3	0.1939	8	0.0452	7	0.0024	8	0.0452

Tabel 4.3. (Lanjutan)

Data	Gap		H		KL		S		CH	
	k	Nilai CTM	k	Nilai CTM	k	Nilai CTM	k	Nilai CTM	k	Nilai CTM
18	2	0.4004	2	0.4004	4	0.1723	6	0.0298	8	0.0556
19	2	0.4958	2	0.4958	2	0.4958	10	0.1463	6	0.2866
20	2	0.3341	2	0.3341	4	0.1472	2	0.3341	6	0.1115
21	2	0.416	3	0.6515	3	0.6515	2	0.416	6	0.1796
22	2	0.4473	2	0.4473	6	0.1928	8	0.0057	6	0.1928
23	3	0.2325	3	0.2325	3	0.2325	3	0.2335	3	0.2325
24	2	0.3668	2	0.3668	8	0.0634	2	0.3668	10	0.0367
25	2	0.4929	2	0.4929	9	0.0717	10	0.0605	9	0.0717
26	2	0.4846	2	0.4846	5	1.9474	2	0.4846	5	1.9474
27	2	0.4944	2	0.4944	5	0.1818	3	0.3278	7	0.1225
28	2	0.4537	2	0.4537	3	0.2938	2	0.4537	2	0.4537
29	2	0.7759	3	0.5244	3	0.5244	2	0.7759	7	0.0516
30	2	0.4308	2	0.4308	4	0.2276	2	0.4308	5	0.164
31	2	0.3129	3	0.223	3	0.223	2	0.3129	3	0.223
32	2	0.4029	2	0.4029	8	0.1163	2	0.4029	8	0.1163
33	2	0.4361	2	0.4361	10	0.0732	2	0.4361	3	0.2818
34	2	0.6498	2	0.6498	5	0.1654	2	0.6498	2	0.6498
35	2	0.3782	2	0.3782	3	0.2336	2	0.3782	2	0.3782
36	3	0.4288	3	0.4288	3	0.4288	3	0.4288	3	0.4288
37	3	0.6494	3	0.6494	9	0.1305	2	0.4728	3	0.6494
38	2	0.8973	2	0.8973	8	0.2912	2	0.8973	13	0.1923
39	2	0.4838	2	0.4838	8	0.0774	6	0.1096	11	0.0438
40	2	0.2808	2	0.4474	8	0.0592	2	0.4474	6	0.0661
41	3	0.7019	2	0.7019	11	0.152	2	0.7019	12	0.1349
42	2	0.4779	2	0.4779	8	0.1009	2	0.4779	4	0.2758
43	2	0.45	2	0.45	6	0.2641	2	0.45	7	0.1253

Tabel 4.3. (Lanjutan)

Data	Gap		H		KL		S		CH	
	k	Nilai CTM	k	Nilai CTM	k	Nilai CTM	k	Nilai CTM	k	Nilai CTM
44	2	0.3242	2	0.3242	9	0.0535	2	0.3242	9	0.0535
45	2	0.3216	2	0.3216	12	0.0225	2	0.3216	12	0.0225
46	3	0.2842	2	0.4645	8	0.1132	2	0.4645	8	0.1132
47	2	0.6321	2	0.6321	4	0.2062	2	0.6321	2	0.6321
48	3	0.3904	3	0.3904	11	0.1051	3	0.3904	3	0.3904
49	3	0.1393	4	0.0991	3	0.1393	2	0.4742	3	0.1393
50	2	0.1859	4	0.0493	15	0.0052	2	0.1859	2	0.1859

keterangan :

k : jumlah kelompok

Tabel 4.3 merupakan ringkasan hasil estimasi jumlah kelompok dan nilai CTM kelima indeks validitas. Hasil analisis kelompok secara lengkap dapat dilihat pada Lampiran 4.

Pada Tabel 4.3 dapat dilihat bahwa hasil pengelompokan kelima indeks validitas tidak sama. Pada tabel tersebut juga dapat dilihat bahwa 96% (48 dari 50 data penelitian) menghasilkan jumlah kelompok yang tidak sama untuk kelima indeks validitas. Kelima indeks validitas tersebut menghasilkan jumlah kelompok yang sama hanya pada data 23. dan 36.

Pada Tabel 4.3 menunjukkan bahwa indeks validitas *Gap statistic*, *Hartigan*, dan *Silhouette* cenderung menghasilkan jumlah kelompok yang sama, yaitu 48% (24 dari 50 data penelitian) yang digunakan menghasilkan jumlah kelompok yang sama untuk ketiga indeks validitas tersebut. Indeks validitas *Krzanowski-Lai* cenderung menghasilkan kelompok yang sama dengan indeks validitas *Calinski-Harabasz*, yaitu 38% (19 dari 50 data penelitian) menghasilkan jumlah kelompok yang sama untuk kedua indeks validitas tersebut.

Jumlah kelompok yang terbentuk menggunakan kelima indeks validitas *Gap statistic*, *Hartigan*, *Krzanowski-Lai*, *Silhouette*, dan *Calinski-Harabasz* memberikan hasil yang berbeda-beda sehingga diperlukan suatu ukuran kebaikan hasil pengelompokan. *Cluster Tightness Measure* (CTM) merupakan ukuran kebaikan dari hasil pengelompokan berdasarkan simpangan baku setiap peubah pada masing-masing kelompok.

Pada Tabel 4.3 dapat dilihat nilai CTM untuk kelima indeks validitas. Tabel tersebut menunjukkan bahwa jumlah kelompok yang paling besar selalu menghasilkan nilai CTM yang paling kecil kecuali pada data 17, 18, dan 26. Hal ini terjadi karena rumus CTM merupakan rasio antara ragam setiap kelompok dengan ragam total dibagi jumlah kelompok yang terbentuk, semakin banyak jumlah kelompok biasanya akan menghasilkan nilai CTM yang semakin kecil.

Pada Tabel 4.3 dapat dilihat bahwa nilai CTM untuk kelima indeks validitas berbeda-beda karena pendefinisian rumus CTM berdasarkan jumlah kelompok yang terbentuk. Tabel 4.4 menunjukkan indeks validitas Krzanowski-Lai, dan Calinski-Harabasz cenderung menghasilkan nilai CTM yang lebih kecil dibandingkan indeks validitas yang lain karena kedua indeks validitas tersebut biasanya menghasilkan jumlah kelompok yang lebih besar daripada indeks validitas yang lain.

Nilai CTM yang berwarna merah merupakan nilai CTM yang terkecil untuk kelima indeks validitas untuk setiap data penelitian. Sebagai contoh pada data 2, nilai CTM untuk indeks validitas *Gap statistic*, Hartigan, Krzanowski-Lai, Silhouette, dan Calinski-Harabasz berturut-turut adalah 0.3153, 0.7352, 0.7352, 0.7352, dan 0.7352. Indeks validitas *Gap statistic* memberikan nilai CTM terkecil yaitu 0.3153, berdasarkan nilai CTM tersebut dapat disimpulkan bahwa indeks validitas *Gap statistic* merupakan indeks validitas terbaik untuk data 2.

Tabel 4.3 menunjukkan indeks validitas Calinski-Harabasz memberikan nilai CTM terkecil paling banyak yaitu pada 54% (27 dari 50 data penelitian). Selanjutnya diikuti indeks validitas Krzanowski-Lai sebanyak 48% (24 dari 50 data penelitian) memberikan nilai CTM terkecil, sedangkan indeks validitas Silhouette, Hartigan, dan *Gap statistic* berturut-turut sebanyak 30%, 14%, dan 8% dari data penelitian menghasilkan nilai CTM terkecil. Berdasarkan uraian sebelumnya dapat dikatakan indeks validitas Krzanowski-Lai merupakan indeks validitas yang paling banyak menghasilkan nilai CTM terkecil dibandingkan 4 indeks validitas yang lain, oleh karena itu dapat disimpulkan indeks validitas Calinski-Harabasz merupakan indeks validitas terbaik untuk mengestimasi jumlah kelompok pada analisis kelompok metode *single linkage*.

Pada bab sebelumnya disebutkan bahwa permasalahan utama pada analisis kelompok adalah berapa jumlah kelompok optimal dari suatu himpunan data yang bisa memperkecil ukuran kesalahan dalam kelompok yang terbentuk. Penentuan jumlah kelompok merupakan hal

yang sulit dilakukan karena tidak ada penjelasan yang tepat tentang kelompok yang terbentuk, artinya setiap orang memiliki anggapan yang berbeda-beda mengenai jumlah kelompok yang akan terbentuk. pada penelitian ini penentuan jumlah kelompok optimal menggunakan lima indeks validitas yang memiliki formula yang berbeda-beda. Indeks validitas *Gap statistic* membangkitkan data dengan menggunakan distribusi referensi, ada dua distribusi referensi yang digunakan pada *Gap statistic* yaitu distribusi referensi uniform dan pendekatan komponen utama dari himpunan data (Tibshirani *et al.*, 2001). Indeks validitas Hartigan, Krzanowski-Lai, dan Calinski-Harabasz menggunakan jumlah kuadrat dalam kelompok dan jumlah kuadrat antar kelompok dalam perhitungannya, sedangkan indeks validitas *Silhouette* menggunakan rata-rata jarak setiap titik ke titik lain dalam kelompok dan di luar kelompok.

Berdasarkan Tabel 4.3, indeks validitas Calinski-Harabasz biasanya menghasilkan nilai CTM yang paling kecil, karena indeks validitas Calinski-Harabasz biasanya menghasilkan jumlah kelompok yang lebih banyak dibandingkan indeks validitas yang lain. Hal ini sesuai dengan rumus CTM yaitu rasio antar simpangan baku kelompok dengan simpangan baku himpunan data penelitian dibagi jumlah variabel dan jumlah kelompok yang terbentuk. Sedangkan indeks validitas *Gap statistic* selalu menghasilkan nilai CTM paling besar karena jumlah kelompok yang terbentuk selalu lebih sedikit dibandingkan indeks validitas yang lain. Menurut (Tibshirani *et al.*, 2001) estimasi indeks validitas *Gap statistic* akan baik pada data yang terpisah dengan baik, ketika himpunan data tidak terpisah dengan baik maka estimasi jumlah kelompok tidak akan bisa dijelaskan secara tepat.

Berdasarkan nilai CTM dapat disimpulkan indeks validitas Calinski-Harabasz merupakan indeks validitas terbaik untuk mengestimasi jumlah kelompok pada analisis kelompok metode *single linkage*.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan yaitu membandingkan indeks validitas dapat diketahui indeks validitas terbaik untuk mengestimasi jumlah kelompok berdasarkan nilai CTM sebagai ukuran kebaikan pengelompokan. Indeks validitas Calinski-Harabasz memberikan nilai CTM terkecil paling banyak yaitu 54% (27 dari 50 data penelitian), indeks validitas Krzanowski-Lai, Silhouett, Hartigan, dan *Gap statistic* berturut-turut memberikan nilai CTM terkecil pada 48%, 30%, 14%, dan 8%. Oleh karena itu, dapat disimpulkan indeks validitas Calinski-Harabasz merupakan indeks validitas terbaik untuk mengestimasi jumlah kelompok pada analisis kelompok metode *single linkage*.

5.2 Saran

Saran dari penelitian ini adalah:

1. Dalam menentukan indeks validitas yang akan digunakan pada analisis kelompok metode *single linkage* disarankan menggunakan indeks validitas Calinski-Harabasz
2. Untuk penelitian selanjutnya disarankan menggunakan model simulasi untuk membandingkan tingkat akurasi hasil pengelompokan indeks validitas.

UNIVERSITAS BRAWIJAYA



DAFTAR PUSTAKA

- Afifi, A. A dan V. Clark. 1990. *Computer Aided Multivariate Analysis*. Chapman & Hall, New York.
- Andayani, N. 2004. *Analisis Komponen Utama dan Analisis Kelompok untuk Mengelompokan Kecamatan di Kotamadya Surabaya Berdasarkan Tingkat Sosial Budaya*. Skripsi Fakultas MIPA Institut Teknologi Sepuluh November Surabaya (tidak dipublikasikan)
- Andress, S. 2003. *Analisis Pendapatan Usaha Tani Jagung dalam Program Pengelolaan Sumberdaya Hutan Bersama Masyarakat (Studi Kasus di Dusun Pondok Asem Desa Kedungsari Kecamatan Tegaldimo Kabupaten Banyuwangi)*. Skripsi Fakultas Pertanian Universitas Brawijaya Malang (tidak dipublikasikan)
- Anwar, N. 2008. *Respon Beberapa Genotipe Kedelai (Glycinemax (L.) Merrill) Terhadap Genangan*. Skripsi Fakultas Pertanian Universitas Brawijaya Malang (tidak dipublikasikan)
- Astutik, S. 2007. *Diktat Analisis Multivariat*. Jurusan Matematika, Fakultas MIPA, Universitas Brawijaya Malang.
- Bolshakova, N. dan F. Azuaje. 2008. *Cluster Validation Techniques for Genome Expression Data*. <http://www.cs.tcd.ie/publications/techreports.02/TCD-CS-2002-33.pdf>. Tanggal Akses 20 Maret 2010.
- Dillon, W. R dan M Goldstein. 1984. *Multivariate Analysis: Methods & Applications*. John Wiley and Sons. New York.
- Epps, J dan E. Mbikairajah. 2008. *Visualisation of Reduced Dimension Microarray Data Using Gaussian Mixture Model*. <http://crpit.com/confpapers/CRPITV45Epps.pdf>. Tanggal Akses 15 Maret 2010.

- Giudici, P. 2003. *Applied Data Mining, Statistical Methods for Business and Industry*. John Wiley and Sons. New York.
- Hair, J. F., R. E. Anderson, R. L. Thatham dan W. c. Black. 1998. *Multivariate Data Analysis Fifth Edition*. Prentice-Hall, Inc. New Jersey.
- Hayati, F. 2008. *Analisis Faktor-Faktor yang Mempengaruhi Permintaan Konsumen Terhadap Listrik pada Rumah tangga (Studi Kasus Dusun Nambongan Desa Caturhar Kecamatan Sleman Kabupaten Sleman Daerah Istimewa Yogyakarta)*. Skripsi Fakultas Ekonomi Universitas Islam Indonesia Yogyakarta (tidak dipublikasikan)
- Hermayudi, W. A. 2004. *Pengelompokkan Industri Tempe Di Wilayah Kota Malang*. Skripsi Fakultas Teknologi Pertanian Universitas Brawijaya Malang (tidak dipublikasikan)
- Johnson, R. A. dan D. W. Wichern. 2002. *Applied Multivariate Statistical Analysis Fifth Edition*. Prentice- Hall, Inc. New Jersey.
- Kamilin, A. 2008. *Perbandingan Indeks Validitas Calinski-Harabasz, Hartigan, dan Krzanowski-Lai Pada Analisis Kelompok Hirarki*. . Skripsi Program Studi Statistika Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang (Tidak dipublikasikan).
- Kariyam dan Subanar. 2002. *Perbandingan Beberapa Indeks Kriteria Penentuan Jumlah Optimal Kelompok*. Program Studi Matematika. Program Pascasarjana Universitas Gadjah Mada Yogyakarta.
- Krismayanti, E. T. 2003. *Kajian Pendugaan Parameter Pada Analisis Faktor Dengan Menggunakan Metode Komponen Utama dan Metode Kemungkinan Maksimum*. Skripsi Program Studi Statistika Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang (Tidak dipublikasikan).
- Laboratorium Data Mining Jurusan Teknik Industri Fakultas Teknologi Industri Universitas Islam Indonesia. 2009. *Clustering*.

http://www.adadata.id.or.id/dl_jump.php?id=49&Login=31c097c3fb5534f53825c99e05a4cbe5. Tanggal Akses 20 April 2010.

- Lattin, J. M., J. D. Carroll, dan P. E. Green. 2006. *Analyzing Multivariate Data*. Duxbury Press. South Western.
- Mufidah, R. 2004. *Analisis Kelompok Optimal dengan Kombinasi Indeks Validitas Pada Analisis Kelompok Hirarki*. . Skripsi Program Studi Statistika Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang (Tidak dipublikasikan).
- Purnanto, K. 2002. *Metode Grafis Chernoff Faces dalam Analisis Kelompok*. Skripsi Program Studi Statistika Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang (Tidak dipublikasikan).
- Purwaningsih, D. W. 2007. *Analisis Cluster Terhadap Tingkat Pencemaran udara Pada Sektor Industri Jawa Tengah*. Skripsi Fakultas MIPA Universitas Negeri Semarang (Tidak dipublikasikan)
- Rencher, A. C. 2002. *Methods of Multivariate Analysis Second Edition*. John Wiley and Sons. New York.
- Sharma, S. 1996. *Applied Multivariate Techniques*. John Wiley and Sons, Inc. New York.
- Suarto, E. A. 2004. *Kajian Regresi Komponen Utama dan Regresi New Stepwise dalam Pemilihan Model Regresi Linier Berganda*. Skripsi Program Studi Statistika Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang (tidak dipublikasikan)
- Timm, N. H. 2002. *Applied Multivariate Analysis*. University of Pittsburgh. USA.
- Tibshirani, W., G. Walther dan T. Hastie. 2001. *Estimating The Number of Clusters in A Dataset via The Gap Statistics*. Technical Report 208 Departement of Statistics Stanford University. Stanford.

Triana, R. L. 2009. *Analisis Dampak Pemekaran Wilayah Terhadap Pengelompokan Kecamatan Berdasarkan Beberapa Peubah Sosial Ekonomi di Kabupaten Bogor Tahun 2008*. Skripsi Fakultas Ekonomi dan Manajemen Institut Pertanian Bogor (tidak dipublikasikan)

Utami, R. S. 2008. *Uji Ketahanan 50 Varietas/ Galur Kacang Hijau Terhadap Black Gram Mottle Virus dan Pengaruh Umur Tanaman Waktu Inkubasi BGMV Terhadap Pertumbuhan dan Hasil Kacang Hijau (Vigna Radiatu L. Wilczek)*. Skripsi Fakultas Pertanian Universitas Brawijaya Malang (tidak dipublikasikan)

Walpole, R. E. 1992. *Pengantar Statistika, Edisi ke-3*. Terjemahan Sumantri, B. PT. GramediaPustaka Utama. Jakarta.

Walpole, R. E. dan R. H. Myers. 1995. *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuan, Edisi Keempat*. Terjemahan R. K. Sembiring. Penerbit ITB. Bandung.

Wibowo, P. 2003. *Interaksi Genotip X Lingkungan Galur-Galur Harapan Kacang Tanah (Arachis hypogala L.)*. Skripsi Fakultas Pertanian Universitas Brawijaya Malang (tidak dipublikasikan)

Wicaksono, Y. I. 2003. *Pola Kepekaan Kuman-Kuman yang Potensial Menyebabkan Infeksi Nosokomial Terhadap Beberapa Antimikroba di RSUD Dr. Saiful Anwar Malang Pada Tahun 2001*. Skripsi . Skripsi Fakultas Kedokteran Universitas Brawijaya Malang (tidak dipublikasikan)

Wijaya, A. 2007. *Pengelompokan jagung inbrida melalui analisis cluster dengan pendekatan metode agglomerative berbasis computer*. Skripsi Fakultas MIPA Universitas Bina Nusantara Jakarta (tidak dipublikasikan)

Yonathan, L. A. 1999. *Pemilihan Makanan Murah Bergizi bagi Ibu Hamil Sebagai Salah Satu Upaya Pencegahan Lahirnya Bayi dengan Berat Lahir Rendah*. Skripsi Fakultas Kedokteran Universitas Brawijaya Malang (tidak dipublikasikan)

Lampiran 1. Data Penelitian

Data 1. Data hasil pengukuran aktivitas konsentrasi Gamma (γ) radionuklida terhadap sampel air sungai Code Yogyakarta di 11 stasiun,

No	Stasiun	X1	X2	X3	X4	X5		X9
1	Turgo	0.0292	0.09	0.056	0.113	0.167		0.099
2	Boyong bawah	0.0701	0.09	0.058	0.116	0.167		0.128
3	Ngentak	0.0974	0.094	0.058	0.116	0.169		0.227
4	Ring road utara	0.1316	0.094	0.058	0.12	0.169		0.269
5	Sardjito	0.6167	0.094	0.058	0.125	0.169		0.354
6	Retjobuntung	0.6312	0.094	0.058	0.127	0.169		0.524
7	Tungkak	0.6075	0.097	0.058	0.127	0.169		0.85
8	Karang Kajen	0.7755	0.097	0.064	0.13	0.172		0.879
9	Ring road Selatan	0.5948	0.097	0.064	0.13	0.172		1.247
10	Abang	0.1658	0.1	0.064	0.132	0.178		1.261
11	Pacar	0.0311	0.12	0.065	0.132	0.178		1.672

Keterangan :

X1 = Debit air (m^3/detik)

X2 = Aktivitas Gamma ^{210}Pb (Bq/Liter)

X3 = Aktivitas Gamma ^{226}Ra (Bq/Liter)

X4 = Aktivitas Gamma ^{212}Pb (Bq/Liter)

X5 = Aktivitas Gamma ^{214}Pb (Bq/Liter)

X6 = Aktivitas Gamma ^{208}Ti (Bq/Liter)

X7 = Aktivitas Gamma ^{214}Bi (Bq/Liter)

X8 = Aktivitas Gamma ^{228}Ac (Bq/Liter)

X9 = Aktivitas Gamma ^{40}K (Bq/Liter)

Sumber : Taftazani, A. 2006. *Usulan Kegiatan 2005-2006. Program Kali Bersih (Prokasih) Sungai Code Yogyakarta*. Yogyakarta: Pusat Teknologi Akselerator dan Proses Bahan (PTAPB)-BATAN.

Lampiran 1. (Lanjutan)

Data 2. Data kadar nutrisi plasma nutfah jagung

No	Varietas	X1	X2	X3	X4	X5	X6	X7
1	Lokal Blora	10.27	1.98	7.38	4.69	2.36	75.58	21.3
2	Lokal Bantul	10.68	1.79	8.1	5.13	2.08	74.3	20.45
3	Lokal Gorontalo	9.94	2.04	8.88	4.95	2.26	74.12	18.1
4	Pakelo	9.83	2.26	7.79	4.84	2.86	75.14	19.14
5	Lokal Jeneponto	10.47	2.09	8.63	4.92	2.16	73.81	19.25
6	Koasa Takalar	10.32	2.17	8.34	4.95	2.73	74.14	18.87
7	Tafenpah	10.22	2.08	9.38	4.64	2.31	73.59	17.84
8	Pen Kikis	10.65	1.59	7.69	5.25	2.45	74.75	22.73
9	Pulut Cilla	10.87	2.1	8.81	4.96	2.01	73.18	7.24
10	Pulut Takalar	10.32	2.15	8.06	5.06	2.15	74.35	6.98
11	Pulut Gorontalo	9.94	2.04	8.98	4.92	2.24	74.06	8.25
12	Pulut Bulukumba	10.71	2.21	8.09	4.94	2.12	73.91	6.72

Keterangan :

X1 = Kadar air (%)

X2 = Kadar abu (%)

X3 = Kadar protein(%)

X4 = Kadar lemak (%)

X5 = Kadar serat kasar (%)

X6 = Kadar karbohidrat (%)

X7 = Kadar amilosa (%)

Sumber : Yasin, M., S. Singgih., dan M. Hamdani. 2005. *Keragaman Hayati Plasma Nutfah Jagung*. Balai Penelitian Tanaman Serealia, Maros.

Lampiran 1. (Lanjutan)

Data 3. Data beban pencemaran udara pada sektor industri di Jawa Tengah tahun 2005

No	Industri	X1	X2	X3	X4	X5
1	Makanan	16941.83	157562.87	73638.52	3168.79	393.13
2	Minuman	14608.25	135861.10	62496.05	2732.06	338.94
3	Tekstil	92737.12	862476.34	403086.58	17343.74	2151.66
4	Kayu	37078.59	344839.36	161163.97	18876.55	860.28
5	Olahan kayu	63789.95	575981.54	269190.49	11582.55	1436.93
6	Kertas	4916.75	483238.54	21370.91	919.53	131.40
7	Kimia dasar	6031.05	56090.16	27387.22	1127.93	139.93
8	Non logam	5226.60	48608.60	22717.69	977.48	121.27
9	Semen	509.72	4740.50	2215.52	95.33	11.83
10	Kapur&gips	9945.85	19350.26	10694.06	1860.05	230.75
11	Logam dasar	3830.46	35624.19	21977.07	1002.93	88.87
12	Hasil olahan logam	7291.27	67810.42	31691.85	1363.61	168.34
13	Rumah sakit	63.37	58.93	284.02	11.85	1.47
14	Perhotelan	98.15	91.29	426.64	18.36	2.28

Keterangan :

X1 = Debu (ton)

X2 = SO₂ (ton)

X3 = NO₂ (ton)

X4= HC (ton)

X5 = CO (ton)

Sumber : Purwaningsih, D. W. 2007. *Analisis Cluster Terhadap Tingkat Pencemaran udara Pada Sektor Industri Jawa Tengah*. Skripsi Fakultas MIPA Universitas Negeri Semarang

Lampiran 1. (Lanjutan)

Data 4. Data rata-rata angka kriminal kota per 100000 populasi

No	City	X1	X2	X3	X4	X5	X6	X7
1	Atlanta	16.5	24.8	106	147	1112	905	494
2	Boston	4.2	13.3	122	90	982	669	954
3	Chicago	11.6	24.7	340	242	808	609	645
4	Dallas	18.1	34.2	184	293	1668	901	602
5	Denver	6.9	41.5	173	191	1534	1368	780
6	Detroit	13.0	35.7	477	220	1566	1183	788
7	Hartford	2.5	8.8	68	103	1017	724	468
8	Honolulu	3.6	12.7	42	28	1457	1102	637
9	Houston	16.8	26.6	289	186	1509	787	697
10	Kansas City	10.8	43.2	255	226	1494	955	765
11	Los Angeles	9.7	51.8	286	355	1902	1386	862
12	New Orleans	10.3	39.7	266	283	1056	1036	776
13	New York	9.4	19.4	522	267	1674	1392	848
14	Portland	5.0	23.0	157	144	1530	1281	488
15	Tucson	5.1	22.9	85	148	1206	756	483
16	Washington	12.5	27.6	524	217	1496	1003	793

Keterangan :

X1 = Rata-rata banyaknya kejadian kriminal pembunuhan (kejadian)

X2 = Rata-rata banyaknya kejadian kriminal pemerkosaan (kejadian)

X3 = Rata-rata banyaknya kejadian kriminal perampokan (kejadian)

X4 = Rata-rata banyaknya kejadian kriminal penyerangan (kejadian)

X5 = Rata-rata banyaknya kejadian kriminal pembongkaran dan pencurian (kejadian)

X6 = Rata-rata banyaknya kejadian kriminal pencurian (kejadian)

X7 = Rata-rata banyaknya kejadian kriminal pencurian mobil (kejadian)

Sumber : Rencher, A. C. 2002. *Methods of Multivariate Analysis Second Edition*. John Willey and Sons. New York.

Lampiran 1. (Lanjutan)

Data 5. Data variabel unsur indikator kemiskinan di Kecamatan Kras Kabupaten Kediri bulan Juli 2006

No	Desa	X1	X2	X3	X4
1	Kras	1869.8	0.400	0.145	99.1
2	Jambean	1613.69	0.403	0.180	97.9
3	Banjaranyar	1397.59	0.403	0.182	97.7
4	Purwodadi	1794.88	0.397	0.142	98.5
5	Nyawangan	1738.73	0.403	0.162	98.6
6	Pelas	1094.44	0.403	0.213	97.9
7	Setonorejo	1175.92	0.403	0.150	97.8
8	Bleber	1247.79	0.403	0.169	98.3
9	Mojosari	1256.76	0.403	0.202	98.7
10	Jabang	1192.66	0.403	0.241	99.2
11	Butuh	1321.6	0.403	0.220	98.9
12	Rejomulyo	1140.45	0.403	0.196	98.7
13	Karangtalun	1003.46	0.403	0.139	97.8
14	Bendosari	1302.08	0.403	0.205	98.7
15	Kanigoro	1054.46	0.403	0.242	99.1
16	Krandang	793.24	0.403	0.183	98.6

Keterangan :

X1 = Kepadatan penduduk (jiwa/km²)

X2 = Proporsi usia tidak produktif

X3 = Proporsi penerima BLT

X4 = Presentase tamat SD (%)

Sumber : Fitrianingsih. 2003. Laporan Praktek Kerja Lapangan. Universitas Brawijaya Malang.

Lampiran 1. (Lanjutan)

Data 6. Data kadar air, gabah isi/malai, gabah hampa/malai beberapa varietas padi

No	Varietas	X1	X2	X3	X4
1	IR71606-02-ML-002	13.87	23.27	96.09	19.29
2	S3393-02-ML-0006	14.20	23.12	88.98	15.91
3	S3382-02-ML-0007	14.37	24.93	109.30	24.25
4	S4325D-02-ML-0008	13.70	22.54	96.40	11.07
5	IR73885-02-ML-0010	14.13	22.42	101.29	25.21
6	IR69734-02-ML-0011	13.73	23.53	96.35	25.05
7	BP50F-02-ML-0013	13.70	22.08	93.56	69.76
8	B10299B-02-ML-0018	13.33	21.71	80.95	21.89
9	B10386E-02-ML-0019	13.27	21.82	78.87	35.94
10	BP138E-02-ML-0021	14.90	20.63	174.76	58.81
11	BP138E-02-ML-0024	14.03	21.79	133.71	28.02
12	BP342B-02-ML-0025	13.80	24.29	116.26	60.73
13	IR 64	13.60	22.59	86.02	12.98
14	Memberamo	13.43	23.30	93.19	30.14
15	Ciherang	13.53	23.41	110.44	16.65
16	Tukad Balian	13.13	20.63	76.12	13.07

Keterangan :

X1 = Bobot 1000 butir (gram)

X2 = Kadar air gabah (%)

X3 = Gabah isi/malai (butir)

X4 = Gabah isi/malai (butir)

Sumber : Rubiyo, Suprpto, dan A. Darajat. 2005. *Evaluasi Beberapa Galur Harapan Padi Sawah di Bali*. Bali: Balai Pengkajian Teknologi Pertanian

Lampiran 1. (Lanjutan)

Data 7. Data tinggi tanaman, jumlah anakan maksimum, dan jumlah anakan produktif beberapa varietas padi

No	Varietas	X1	X2	X3
1	IR71606-02-ML-0002	85.80	21.07	18.80
2	S3393-02-ML-0006	85.13	21.00	18.80
3	S3382-02-ML-0007	97.87	23.80	21.67
4	S4325D-02-ML-0008	86.80	22.20	20.20
5	IR73885-02-ML-0010	91.87	21.07	18.90
6	IR69734-02-ML-0011	84.87	19.00	17.07
7	BP50F-02-ML-0013	93.80	16.07	15.60
8	B10299B-02-ML-0018	78.87	20.87	18.73
9	B10386E-02-ML-0019	99.20	19.40	17.20
10	BP138E-02-ML-0021	97.87	11.67	9.53
11	BP138E-02-ML-0024	97.60	17.53	15.51
12	BP342B-02-ML-0025	103.73	13.87	12.20
13	IR64	85.67	25.00	22.87
14	Memberamo	95.67	18.53	16.67
15	Ciherang	85.53	20.13	18.53
16	Tukad Balian	83.20	20.33	18.20

Keterangan :

X1 = Tinggi tanaman (cm)

X2 = Jumlah anakan maksimum (per rumpun)

X3 = Jumlah anakan produktif (per rumpun)

Sumber : Rubiyo, Suprpto, dan A. Darajat. 2005. *Evaluasi Beberapa Galur Harapan Padi Sawah di Bali*. Bali: Balai Pengkajian Teknologi Pertanian

Lampiran 1. (Lanjutan)

Data 8. Data panjang malai, bobot kering panen, dan bobot 1000 butir gabah beberapa varietas padi

No	Varietas	X1	X2	X3
1	IR71606-02-ML-002	21.92	8.39	26.27
2	S3393-02-ML-0006	21.54	7.68	28.13
3	S3382-02-ML-0007	23.11	8.37	23.57
4	S4325D-02-ML-0008	22.15	7.73	28.07
5	IR73885-02-ML-0010	22.56	7.58	24.17
6	IR69734-02-ML-0011	21.80	8.38	24.60
7	BP50F-02-ML-0013	21.93	7.87	28.30
8	B10299B-02-ML-0018	22.01	6.12	21.53
9	B10386E-02-ML-0019	22.22	9.50	30.30
10	BP138E-02-ML-0021	24.23	8.02	27.60
11	BP138E-02-ML-0024	19.55	8.60	26.30
12	BP342B-02-ML-0025	25.68	5.83	27.53
13	IR64	21.38	8.45	27.17
14	Memberamo	22.15	9.00	28.10
15	Ciherang	21.46	7.59	27.13
16	Tukad Balian	20.16	6.80	27.63

Keterangan :

X1 = Panjang malai (cm)

X2 = Bobot kering panen (ton/Ha)

X3 = Bobot 1000 butir gabah (gram)

Sumber : Rubiyo, Suprpto, dan A. Darajat. 2005. *Evaluasi Beberapa Galur Harapan Padi Sawah di Bali*. Bali: Balai Pengkajian Teknologi Pertanian

Lampiran 1. (Lanjutan)

Data 9. Data kandungan protein, lisin, dan triptofan beberapa varietas jagung QPM

No	Varietas	X1	X2	X3
1	S99TLYQ GH “A”	10.00	0.48	0.091
2	S99TLYQ GH “AB”	8.56	0.421	0.081
3	S99TLYQ-A	9.90	0.468	0.091
4	S99TLYQ-AB	10.38	0.477	0.093
5	S00TLYQ-B	9.85	0.48	0.092
6	S00TLYQ-AB	9.25	0.456	0.086
7	POZA RICA 8365	9.39	0.441	0.085
8	IBOPERENDRA 8565	9.03	0.452	0.086
9	TOCUMEN 8565	9.45	0.447	0.084
10	POZA RICA 8765	8.49	0.378	0.078
11	ACROSS 8765	10.22	0.432	0.096
12	ACROSS 8666	10.53	0.411	0.093
13	POZA RICA 8666	9.46	0.425	0.084
14	POP.61C1QPM TEYF	8.95	0.411	0.081
15	B i s m a	8.82	0.275	0.065
16	L a m u r u	8.49	0.278	0.064

Keterangan :

X1 = Kandungan protein (%)

X2 = Kandungan lisin (%)

X3 = Kandungan triptofan (%)

Sumber : Yasin, M., S. Singgih., dan M. Hamdani. 2005. *Keragaman Hayati Plasma Nutfah Jagung*. Balai Penelitian Tanaman Serealia, Maros.

Lampiran 1. (Lanjutan)

Data 10. Data varietas kedelai

No	Genotip	X1	X2	X3	X4	X5	X6
1	MLGG 0133	0.9	0.97	1.21	1.3	1.32	0.81
2	MLGG 0144	1.4	0.93	1.51	1.3	1.38	1.33
3	MLGG 0231	1.11	1.49	0.85	0.58	0.83	0.28
4	MLGG 0236	1.11	0.46	1.34	1.55	1.59	1.04
5	MLGG 0258	0.98	0.68	0.3	0.24	0.26	0.36
6	MLGG 0470	0.72	0.54	0.59	0.45	0.52	0.97
7	MLGG 0532	1.58	1.86	3.33	3.23	3.14	2.08
8	MLGG 0537	1.8	1.17	2	4.5	4.89	6.67
9	MLGG 0553	0.43	0.75	0.39	0.7	0.7	0.31
10	MLGG 0583	0.84	0.69	0.33	0.44	0.41	0.41
11	MLGG 0602	1.45	2.04	0.64	0.52	0.5	0.48
12	MLGG 0714	0.85	0.41	0.66	0.39	0.35	0.33
13	MLGG 0720	0.66	1.39	0.58	0.42	0.31	0.34
14	MLGG 0781	0.89	0.88	0.52	0.49	0.67	0.81
15	MLGG 0817	0.66	1.24	1.24	1.16	0.91	0.69
16	Panderman	0.66	0.27	0.72	0.54	0.66	0.38
17	Lawit	1.47	2.17	2.38	1.91	1.34	1.53

Keterangan :

X1 = Tinggi tanaman

X2 = Jumlah cabang

X3 = Jumlah buku subur

X4 = Jumlah biji

X5 = Hasil biji

Sumber : Anwar, N. 2008. *Respon Beberapa Genotipe Kedelai (Glycinemax (L.) Merrill) Terhadap Genangan*. Tugas Akhir Jurusan Budidaya Pertanian Fakultas Pertanian Universitas Brawijaya Malang

Lampiran 1. (Lanjutan)

Data 11. Data indikator kemiskinan di Kabupaten Bangkalan

No	Kecamatan	X1	X2	X3	X4
1	Kamal	1.89	3.21	30.10	18.80
2	Labang	2.08	3.52	28.34	7.06
3	Kwanyar	3.07	4.22	27.18	14.56
4	Modung	3.30	5.05	22.51	12.90
5	Blega	5.18	6.21	23.23	17.69
6	Konang	3.75	4.45	21.80	10.98
7	Galis	4.11	4.92	22.38	14.43
8	Tanah Merah	3.27	4.90	28.32	16.33
9	Tragah	2.24	3.78	24.22	13.25
10	Socah	2.47	4.42	27.98	11.94
11	Bangkalan	1.47	3.25	39.57	18.13
12	Burneh	2.61	5.15	25.12	24.73
13	Arosbaya	2.45	4.47	28.70	17.38
14	Geger	4.10	7.12	20.75	16.43
15	Kokop	3.08	5.68	18.95	10.61
16	Tanjungbumi	2.46	4.28	25.37	18.09
17	Sepulu	3.28	4.92	21.52	14.38
18	Klampis	2.74	4.52	25.76	13.19

Keterangan :

X1 = Luas lahan pertanian (Ha)

X2 = Produksi pertanian (ton)

X3 = Kepadatan penduduk (per Km²)

X4 = Pendapatan asli daerah (juta rupiah)

Sumber : Krismayanti, E. T. 2003. *Kajian Pendugaan Parameter Pada Analisis Faktor Dengan Menggunakan Metode Komponen Utama dan Metode Kemungkinan Maksimum*. Skripsi Program Studi Statistika Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang

Lampiran 1. (Lanjutan)

Data 12. Data karakter agronomi pada kondisi tercekam (perlakuan) dan bobot kering akar 18 varietas padi sawah

No	Varietas	X1	X2	X3
1	IR42	6.90	0.79	0.60
2	IR64	17.90	0.76	0.72
3	IR66	8.10	0.47	0.26
4	IR74	18.80	0.77	0.49
5	Widas	22.80	1.05	1.02
6	Memberamo	0.66	0.31	0.87
7	Digul	9.40	1.12	0.76
8	Cisadane	2.15	1.05	0.46
9	Punggur	2.08	0.95	0.90
10	Cirata	22.70	1.95	0.88
11	Celebes	2.05	0.63	0.85
12	Cisantana	2.06	0.88	0.87
13	Kalimas	0.58	0.34	0.37
14	Banyuasin	1.77	0.66	0.60
15	Batang Piaman	1.48	0.68	0.97
16	Batang Lembang	3.37	1.88	1.03
17	Ketan Lokal	2.11	1.00	1.03
18	Randah Kuning	2.14	1.09	0.87

Keterangan :

X1 = Total bobot kering (gram)

X2 = Bobot kering tajuk (gram)

X3 = Bobot kering akar (gram)

Sumber : Munir, R., Sunadi., dan W. haryoko. 2009. *Penapisan Varietas Padi Toleran Salinitas pada Lahan Rawa di Kabupaten Pesisir Selatan*. BPTP Sukarami Sumatera Barat

Lampiran 1. (Lanjutan)

Data 13. Data karakter agronomi 18 varietas padi pada sawah lahan rawa

No	Varietas	X1	X2	X3
1	IR42	63.62	1.68	0.51
2	IR64	17.50	0.26	0.66
3	IR66	26.04	0.53	0.60
4	IR74	29.51	0.83	1.02
5	Widas	24.57	0.43	0.80
6	Memberamo	19.72	0.65	0.87
7	Digul	37.09	1.01	0.85
8	Cisadane	21.08	0.50	0.46
9	Punggur	24.30	0.62	0.90
10	Cirata	8.62	0.65	0.31
11	Celebes	22.86	0.35	0.85
12	Cisantana	39.35	0.77	0.87
13	Kalimas	15.47	0.69	0.37
14	Banyuasin	19.67	0.59	0.60
15	Btg. piaman	42.21	0.66	0.97
16	Btg. lembang	22.91	0.74	1.03
17	Ketan	14.00	0.94	1.03
18	Randah kuning	17.93	0.76	0.87

Keterangan :

X1 = Total bobot kering (gram)

X2 = Bobot kering tajuk (gram)

X3 = Bobot kering akar (gram)

Sumber : Munir, R., Sunadi., dan W. haryoko. 2009. *Penapisan Varietas Padi Toleran Salinitas pada Lahan Rawa di Kabupaten Pesisir Selatan*. BPTP Sukarami Sumatera Barat

Lampiran 1. (Lanjutan)

Data 14. Data persentase tenaga kerja di bidang pertanian, industri dan perusahaan jasa di beberapa negara

No	Negara	X1	X2	X3
1	CANADA	13	43	45
2	SWEEDEN	14	53	33
3	SWITZERLAND	11	56	33
4	LUXEMBOURG	15	51	34
5	U. KINGDOM	4	56	40
6	DENMARK	18	45	37
7	W. GERMANY	15	60	25
8	FRANCE	20	44	36
9	BELGUIM	6	52	42
10	NORWAY	20	49	32
11	ICELAND	25	47	29
12	NETHERLANDS	11	49	40
13	AUSTRIA	23	47	30
14	IRELAND	36	30	34
15	ITALY	27	46	28
16	JAPAN	33	35	32
17	GREECE	56	37	21
18	PORTUGAL	44	33	23
19	TURKEY	79	12	9

Keterangan :

X1 = Tenaga kerja di bidang pertanian (%)

X2 = Tenaga kerja di bidang industry (%)

X3 = Tenaga kerja di perusahaan jasa (%)

Sumber : OECD Economic Development

(<http://lib.stat.cmu.edu/DASL/Datafiles/LaborForce.html>)

Lampiran 1. (Lanjutan)

Data 15. Data tingkat produksi dan penggunaan faktor produksi usaha tani jagung pada petani yang mengikuti program PHBM

Petani	X1	X2
1	64.00	115.00
2	82.00	100.00
3	70.67	106.67
4	60.00	106.67
5	68.00	110.00
6	64.00	120.00
7	68.00	105.00
8	66.67	100.00
9	68.00	106.67
10	72.00	120.00
11	74.00	120.00
12	74.00	130.00
13	62.67	113.33
14	78.00	140.00
15	80.00	140.00
16	64.00	106.67
17	68.00	140.00
18	74.00	120.00
19	72.00	130.00
20	62.00	130.00

Keterangan

X1 = Benih jagung Hibrida C7 yang digunakan petani (kg)

X2 = Banyaknya pupuk KCl yang digunakan petani (kg)

Sumber : Andress, S. 2003. *Analisis Pendapatan Usaha Tani Jagung dalam Program Pengelolaan Sumberdaya Hutan Bersama Masyarakat (Studi Kasus di Dusun Pondok Asem Desa Kedungsari Kecamatan Tegaldimo Kabupaten Banyuwangi)*. Skripsi Fakultas Pertanian Universitas Brawijaya Malang

Lampiran 1. (Lanjutan)

Data 16. Daftar analisis bahan makanan lauk pauk nabati (per 100 gram bahan)

No	Bahan Makanan	X1	X2	X3
1	Air kelapa muda	0.2	0.1	3.8
2	Daging kelapa muda	1.0	0.9	14.0
3	Daging kelapa setengah tua	4.0	15.0	10.0
4	Daging kelapa tua	3.4	34.7	14.0
5	Kenari	15.0	66.0	13.0
6	Biji kecipir	32.8	17.0	36.0
7	Kluwak	10.0	24.0	13.5
8	Biji koro benguk	0.0	3.0	55.0
9	Biji koro krupuk	24.0	0.7	22.1
10	Biji koro weduk	8.3	1.5	61.0
11	Kwaci	22.2	42.1	13.8
12	Biji muda lamtoro	30.6	0.3	15.4
13	Biji tua lamtoro	5.7	0.5	26.2
14	Biji segar nangka	10.6	0.1	36.7
15	Oncom	13.0	6.0	22.6
16	Biji pala	7.5	36.4	40.1
17	Pati kacang hijau	4.5	1.0	83.5
18	Pete segar	0.4	2.0	22.0
19	Susu kedelai	3.5	2.5	5.0
20	Tahu	7.8	4.6	1.6

Keterangan :

X1 = Protein (%)

X2 = Lemak (%)

X3 = Karbohidrat (%)

Sumber : Yonathan, L. A. 1999. *Pemilihan Makanan Murah Bergizi bagi Ibu Hamil Sebagai Salah Satu Upaya Pencegahan Lahirnya Bayi dengan Berat Lahir Rendah*. Skripsi Fakultas Kedokteran Universitas Brawijaya Malang

Lampiran 1. (Lanjutan)

Data 17. Data pengeluaran untuk periklanan di beberapa perusahaan

No	Perusahaan	X1	X2
1	MILLER LITE	50.1	32.1
2	PEPSI	74.1	99.6
3	STROH'S	19.3	11.7
4	FED'L EXPRESS	22.9	21.9
5	BURGER KING	82.4	60.8
6	COCO-COLA	40.1	78.6
7	MC DONALD'S	185.9	92.4
8	MCI	26.9	50.7
9	DIET COLA	20.4	21.4
10	FORD	166.2	40.1
11	LEVI'S	27	40.8
12	BUD LITE	45.6	10.4
13	ATT/BELL	154.9	88.9
14	CALVIN KLEIN	5	12
15	WENDY'S	49.7	29.2
16	POLAROID	26.9	38
17	SHASTA	5.7	10
18	MEOW MIX	7.6	12.3
19	OSCAR MEYER	9.2	23.4
20	CREST	32.4	71.1
21	KIBBLES 'N BITS	6.1	4.4

Keterangan

X1 = Biaya periklanan (juta dolar)

X2 = Laba (juta dolar)

Sumber : TV Ad Yields

(<http://lib.stat.cmu.edu/DASL/Datafiles/Protein.html>)

Lampiran 1. (Lanjutan)

Data 18. Data tingkat kepekaan kuman *Escherichia Coli* terhadap beberapa antimikroba di RSUD DR Saiful Anwar Malang

No	Jenis Antibiotik	X1	X2	X3
1	Ampicillin	73.41	2.18	24.40
2	Amoxicillin	72.22	2.40	27.78
3	Amox+Clav Acid	13.64	18.58	67.79
4	Cefadroxil	30.75	26.19	43.06
5	Cefotaxim	6.24	14.69	79.07
6	Ceftriaxone	8.51	12.87	78.61
7	Cefuroxim	18.81	56.44	24.74
8	Cephalotin	53.16	17.11	29.74
9	Cephadrine	31.78	27.13	41.09
10	Kanamychine	26.98	4.56	68.45
11	Gentamycine	15.81	0.79	83.40
12	Netilmycine	6.14	1.58	92.28
13	Amikacine	5.77	1.19	93.04
14	Ciprofoxacin	17.59	3.36	79.05
15	Ofloxacin	17.84	1.45	80.71
16	Norfloxacin	19.25	1.26	79.50
17	Tetracycline	72.17	3.58	24.25
18	Chloramphenicol	49.10	2.78	47.02
19	Nalidixic Acid	31.16	5.07	63.77
20	Nitrofurantoin	11.68	8.03	80.29
21	Centrimoxazole	61.45	0.48	38.07
22	Sulfonamide	73.81	0.99	25.20
23	Doxycycline	66.14	1.58	32.28
24	Fosfomycine	12.37	1.03	86.60

Keterangan :

X1 = Resisten

X2 = Intermediate

X3 = Sensitif

Sumber : Wicaksono, Y. I. 2003. *Pola Kepekaan Kuman-Kuman yang Potensial Menyebabkan Infeksi Nosokomial Terhadap Beberapa Antimikroba di RSUD Dr. Saiful Anwar Malang Pada Tahun 2001*. Skripsi . Skripsi Fakultas Kedokteran Universitas Brawijaya Malang

Lampiran 1. (Lanjutan)

Data 19. Data konsumsi protein di beberapa negara Eropa

No	Country	X1	X2	X3	X4	X5	X6	X7	X8	X9
1	Albania	10.1	1.4	0.5	8.9	0.2	42.3	0.6	5.5	1.7
2	Austria	8.9	14	4.3	19.9	2.1	28	3.6	1.3	4.3
3	Belgium	13.5	9.3	4.1	17.5	4.5	26.6	5.7	2.1	4
4	Bulgaria	7.8	6	1.6	8.3	1.2	56.7	1.1	3.7	4.2
5	Czechoslovakia	9.7	11.4	2.8	12.5	2	34.3	5	1.1	4
6	Denmark	10.6	10.8	3.7	25	9.9	21.9	4.8	0.7	2.4
7	E Germany	8.4	11.6	3.7	11.1	5.4	24.6	6.5	0.8	3.6
8	Finland	9.5	4.9	2.7	33.7	5.8	26.3	5.1	1	1.4
9	France	18	9.9	3.3	19.5	5.7	28.1	4.8	2.4	6.5
10	Greece	10.2	3	2.8	17.6	5.9	41.7	2.2	7.8	6.5
11	Hungary	5.3	12.4	2.9	9.7	0.3	40.1	4	5.4	4.2
12	Ireland	13.9	10	4.7	25.8	2.2	24	6.2	1.6	2.9
13	Italy	9	5.1	2.9	13.7	3.4	36.8	2.1	4.3	6.7
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
25	Yugoslavia	4.4	5	1.2	9.5	0.6	55.9	3	5.7	3.2

Keterangan :

X1 = Konsumsi daging merah

X2 = Konsumsi daging putih

X3 = Konsumsi telur

X4 = Konsumsi susu

X5 = Konsumsi ikan

X6 = Konsumsi sereal

X7 = Konsumsi strachy

X8 = Konsumsi kacang-kacangan

X9 = Konsumsi sayur dan buah

Sumber : Protein

(<http://lib.stat.cmu.edu/DASL/Datafiles/Protein.html>)

Lampiran 1. (Lanjutan)

Data 20. Data rata-rata tinggi tanaman dan jumlah polong per tanaman dari hasil penelitian interaksi genotip x lingkungan galur-galur kacang tanah (*Arachis hypogala L.*)

No	Varietas	X1	X2	X3
1	Mn-21	1.32	35.20	11.97
2	IP99-52	1.37	37.38	11.83
3	P-9409	1.30	33.63	12.93
4	P-9407	1.40	39.87	11.83
5	I-11	1.30	30.05	14.17
6	CF3-11	1.57	27.60	16.15
7	Lm/LT-93-BI-131	1.37	30.92	14.00
8	Lm/LT-93-B2-14	1.34	35.45	11.63
9	87123/86680-93-B75	1.40	36.28	13.27
10	Lm/LT-93-B2-20	1.48	33.08	16.83
11	Lm/LT-93-B2-65	1.35	36.15	11.93
12	Lm/LT-93-B1-169	1.50	36.62	10.07
13	GH 7904	1.33	35.20	9.60
14	GH 7920	1.20	36.87	9.23
15	ICGV 87055	1.37	36.65	11.53
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
25	MHS/91278-99-C-180-5	1.12	37.72	11.83

Keterangan :

X1 = Berat polong kering (ton)

X2 = Tinggi tanaman (cm)

X3 = Jumlah polong/ tanaman

Sumber : Wibowo, P. 2003. *Interaksi Genotip X Lingkungan Galur-Galur Harapan Kacang Tanah (Arachis hypogala L.)*. Skripsi Fakultas Pertanian Universitas Brawijaya Malang

Lampiran 1. (Lanjutan)

Data 21. Data persentase tenaga kerja industri di beberapa negara eropa

No	Negara	X1	X2	X3	X4	X5	...	X9
1	Belgium	3.3	0.9	27.6	0.9	8.2	...	7.2
2	Denmark	9.2	0.1	21.8	0.6	8.3	...	7.1
3	France	10.8	0.8	27.5	0.9	8.9	...	5.7
4	W. Germany	6.7	1.3	35.8	0.9	7.3	...	6.1
5	Ireland	23.2	1.0	20.7	1.3	7.5	...	6.1
6	Italy	15.9	0.6	27.6	0.5	10.0	...	5.7
7	Luxembourg	7.7	3.1	30.8	0.8	9.2	...	6.2
8	Netherlands	6.3	0.1	22.5	1.0	9.9	...	6.8
9	United Kingdom	2.7	1.4	30.2	1.4	6.9	...	6.4
10	Austria	12.7	1.1	30.2	1.4	9.0	...	7.0
11	Finland	13.0	0.4	25.9	1.3	7.4	...	7.6
12	Greece	41.4	0.6	17.6	0.6	8.1	...	6.7
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
25	USSR	23.7	1.4	25.8	0.6	9.2	...	9.3

Keterangan :

X1 = Persentase tenaga kerja di bidang pertanian

X2 = Persentase tenaga kerja di bidang mining

X3 = Persentase tenaga kerja di bidang industri

X4 = Persentase tenaga kerja di bidang penyediaan bahan industri

X5 = Persentase tenaga kerja di bidang bangunan

X6 = Persentase tenaga kerja di bidang perusahaan jasa

X7 = Persentase tenaga kerja di bidang keuangan

X8 = Persentase tenaga kerja di bidang sosial

X9 = Persentase tenaga kerja di bidang transportasi dan komunikasi

Sumber : European Jobs

(<http://lib.stat.cmu.edu/DASL/Datafiles/EuropeanJobs.html>)

Lampiran 1. (Lanjutan)

Data 22. Data karakteristik industri kerajinan rumah tangga di 27 provinsi di Indonesia tahun 1996

No	Provinsi	X1	X2	X3	X4	X5
1	Aceh	89065	1.59	2.82	2688.25	5105.24
2	Sumut	56656	1.83	6.95	3203.27	6393.693
3	Sumbar	58120	1.72	7.14	4025.46	6579.098
4	Riau	20681	1.96	4.11	3603.2	7251.149
5	Jambi	13777	2.21	8.94	6031.6	9824.699
6	Sumsel	48673	1.94	8.67	3704.89	7064.58
7	Bengkulu	9883	2.14	15.59	3642.94	7787.24
8	Lampung	63335	2.05	6.93	3333.15	5797.711
9	Jakarta	20186	2.1	11.43	3595.01	6366.689
10	Jabar	346219	1.98	6.84	4354.01	7611.288
11	Jateng	661812	1.76	1.07	2156.86	3495.126
12	Yogya	80309	1.69	1.43	2291.1	3903.284
13	Jatim	561142	1.8	1.55	2335.65	3987.187
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
27	Irja	8792	2.1	4.77	3768.58	8910.319

Keterangan :

X1 = Jumlah total usaha di tiap propinsi (unit)

X2 = Rata-rata perbandingan jumlah pekerja dengan banyaknya usaha dalam setiap propinsi

X3 = Prosentase jumlah usaha yang memiliki ijin usaha

X4 = Rata-rata total input antara (ribu rupiah)

X5 = Rata-rata total nilai output (ribu rupiah)

Sumber : Purnanto, K. 2002. *Metode Grafis Chernoff Faces dalam Analisis Kelompok*. Skripsi Program Studi Statistika Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang

Lampiran 1. (Lanjutan)

Data 23. Data penduduk Surabaya berdasarkan tingkat sosial budaya di 28 kecamatan

No	Kecamatan	X1	X2	X3	X4	X5	...	X8
1	Tegalsari	49	150	112	1035	14	378	208
2	Genteng	36	115	89	808	19	516	845
3	Bubutan	50	112	101	956	14	330	98
4	Simokerto	42	90	76	703	8	165	101
5	Pabean	21	46	45	460	9	252	138
6	Semampir	53	126	133	1254	15	350	150
7	Krembang	50	131	103	974	18	573	165
8	Kenjeran	37	97	104	822	14	524	224
9	Tam. Sari	85	220	140	1536	23	591	266
10	Gubeng	70	178	115	1320	23	512	424
11	Rungkut	34	119	45	591	13	333	144
12	Tenggilis	24	64	33	440	7	146	130
13	Gun. Anyar	14	42	43	300	6	99	0
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
28	Lakarsantri	23	72	74	727	13	337	118

Keterangan :

X1 = Jumlah sekolah TK

X2 = Banyaknya guru TK

X3 = Jumlah sekolah SD

X4 = Banyaknya guru SD

X5 = Jumlah sekolah SLTP

X6 = Banyaknya guru SLTP

X7 = Jumlah sekolah SMU

X8 = Banyaknya guru SMU

Sumber : Andayani, N. 2004. *Analisis Komponen Utama dan Analisis Kelompok untuk Mengelompokkan Kecamatan di Kotamadya Surabaya Berdasarkan Tingkat Sosial Budaya*. Skripsi Fakultas MIPA ITS Surabaya

Lampiran 1. (Lanjutan)

Data 24. Data gaji dan total penerimaan tim sepak bola

No	TEAM	X1	X2
1	BILLS	650	24249
2	BENGALS	1050	20666
3	BROWNS	967	19413
4	BRONCOS	500	21992
5	OILERS	1700	21399
6	COLT	2000	22022
7	CHIEFS	1100	25859
8	RAIDERS	1300	20390
9	DOLPHINS	1400	23738
10	PATRIOTS	2250	23294
11	JETS	800	19063
12	STEELERS	3500	30131
13	CHARGERS	1200	18698
14	SEAHAWKS	1250	25348
15	FALCONS	2250	25642
.	.	.	.
.	.	.	.
.	.	.	.
28	REDSKINS	1450	20780

Keterangan :

X1 = Gaji pemain gelandang (ribu dolar)

X2 = Total gaji tim (ribu dolar)

Sumber : Q-back and Team Salaries

(<http://lib.stat.cmu.edu/DASL/Datafiles/qbacksalarydat.html>)

Lampiran 1. (Lanjutan)

Data 25. Data varietas jagung inbrida

No	Varietas	X1	X2	X3	X4	X5
1	G1	3862.42	61	63	99	167.4
2	G2	3529.167	61	63	101	177.4
3	G3	3062.225	52	55	88	148.4
4	G4	1660.98	54	55	87	165.2
5	G5	3913.412	53	58	98	177.2
6	G6	4173.449	56	56	98	183.4
7	G7	2433.489	49	49	82	147.8
8	G8	3673.932	56	51	87	187
9	G9	3504.941	55	56	87	154.2
10	G10	4073.824	56	57	88	162.6
11	G11	3487.647	54	57	90	167.6
12	G12	3602.729	58	61	88	176.6
13	G13	2352.634	49	50	82	145
14	G14	3776.042	52	59	95	175.2
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
30	G30	870.557	52	55	88	118.4

Keterangan :

X1 = Berat biji (gram)

X2 = Umur bunga jantan (hari)

X3 = Umur bunga jantan (hari)

X4 = Umur panen (hari)

X5 = Tinggi tanaman (cm)

Sumber : Wijaya, A. 2007. *Pengelompokan jagung inbrida melalui analisis cluster dengan pendekatan metode agglomerative berbasis computer*. Skripsi Fakultas MIPA Universitas Bina Nusantara Jakarta

Lampiran 1. (Lanjutan)

Data 26. Nilai rata- rata hasil ujian nasional 2008 Madrasah Aliyah Negeri program IPS

No	PROVINSI	X1	X2	X3	X4	X5	X6
1	DKI	7.27	6.64	6.98	8.00	7.96	6.83
2	JABAR	7.56	7.09	7.89	8.17	8.32	6.84
3	JATENG	7.07	6.34	6.14	7.60	7.27	6.14
4	DIY	7.62	5.87	6.08	7.46	7.63	6.14
5	JATIM	7.65	6.86	7.47	7.88	8.31	6.16
6	NAD	5.72	5.84	6.19	6.68	6.74	5.55
7	SUMUT	6.95	6.98	7.53	7.96	7.46	7.03
8	SUMBAR	6.64	6.14	6.45	7.31	7.66	5.88
9	RIAU	6.65	6.09	6.41	7.66	6.93	6.37
10	JAMBI	5.92	6.36	6.55	7.68	8.08	7.02
11	SUMSEL	7.13	7.10	7.40	7.80	7.37	7.10
12	LAMPUNG	6.73	7.02	6.89	8.24	7.67	7.33
13	KALBAR	6.54	5.03	6.00	6.59	6.77	5.56
14	KALTENG	7.00	6.42	6.37	7.19	6.59	6.62
15	KALSEL	6.60	6.92	6.63	7.53	7.29	6.73
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
32	PABAR	7.08	6.53	6.87	8.17	6.72	8.69

Keterangan :

X1 = Rata-rata nilai bahasa indonesia

X2 = Rata-rata nilai bahasa inggris

X3 = Rata-rata nilai matematika

X4 = Rata-rata nilai ekonomi

X5 = Rata-rata nilai sosiologi

X6 = Rata-rata nilai geografi

Sumber : Sjafrudin. 2008. *Analisis Hasil Ujian Nasional Madrasah Aliyah Negeri Tahun 2008* <http://pendis.depag.go.id/file/dokumen/analisis200803.pdf>

Lampiran 1. (Lanjutan)

Data 27. Data rata-rata konsumsi protein per-kapita sehari penduduk Kabupaten Malang menurut kelompok makanan tahun 2007

No	Kecamatan	X1	X2	X3	X4	...	X8
1	Sumbermanjing	21.90	0.30	16.40	2.40	...	0.70
2	Pagelaran	20.90	0.40	13.30	2.80	...	0.60
3	Pakis	21.70	0.60	8.40	3.50	...	0.60
4	Tirtoyudo	18.60	0.50	11.30	4.60	...	0.60
5	Tajinan	18.90	0.50	10.80	3.50	...	0.60
6	Kromengan	19.20	0.30	10.10	3.60	...	1.00
7	Gondanglegi	22.30	0.40	9.20	3.00	...	0.70
8	Pakisaji	18.80	0.30	7.60	3.30	...	1.00
9	Dampit	20.80	0.30	15.40	3.50	...	0.80
10	Pujon	15.70	0.30	10.50	4.20	...	0.50
11	Dau	17.50	0.30	6.70	5.00	...	0.50
12	Lawang	20.80	0.30	6.10	4.00	...	0.50
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
33	Pagak	22.20	0.80	20.30	2.40	...	0.60

Keterangan :

X1 = Konsumsi padi-padian (gram)

X2 = Konsumsi umbi-umbian (gram)

X3 = Konsumsi ikan (gram)

X4 = Konsumsi daging (gram)

X5 = Konsumsi telur dan susu (gram)

X6 = Konsumsi sayuran (gram)

X7 = Konsumsi kacang-kacangan (gram)

X8 = Konsumsi buah-buahan (gram)

Sumber : Sulistiyowati, Y. 2009. Laporan Praktek Kerja Lapang. Universitas Brawijaya Malang.

Lampiran 1. (Lanjutan)

Data 28. Nilai rata- rata hasil ujian nasional 2008 Madrasah Tsanawiyah

No	PROVINSI	X1	X2	X3	X4
1	JABAR	7.33	7.46	7.43	7.59
2	BALI	7.19	7.33	7.36	7.89
3	JATIM	7.24	7.35	7.13	7.53
4	BANTEN	6.87	7.18	7.29	7.50
5	GTLO	6.39	7.82	7.57	7.04
6	SUMUT	6.75	7.34	7.09	7.21
7	DKI	7.63	6.87	6.68	7.04
8	SULUT	6.62	6.71	7.03	7.27
9	NAD	6.25	7.22	7.10	7.03
10	SULTRA	6.32	7.17	6.97	7.00
11	JAMBI	6.46	7.02	6.89	7.02
12	MALUKU	6.41	6.85	6.83	7.06
13	SUMSEL	6.60	7.03	6.90	6.61
14	KALTIM	6.85	6.75	6.75	6.75
15	NTB	6.15	7.11	6.60	6.84
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
33	NTT	6.10	4.77	4.71	5.22

Keterangan :

X1 = Rata-rata nilai bahasa indonesia

X2 = Rata-rata nilai bahasa inggris

X3 = Rata-rata nilai matematika

X4 = Rata-rata nilai IPA

Sumber : Sjafrudin. 2008. *Analisis Hasil Ujian Nasional Madrasah Tsanawiyah Negeri Tahun 2008*. <http://pendis.depag.go.id/file/dokumen/analisis200803.pdf>

Lampiran 1. (Lanjutan)

Data 29. Data penduduk berumur 10 tahun ke atas menurut kabupaten/kota

No	Kabupaten/Kota	X1	X2	X3	X4	X5
1	Pacitan	168294	188613	69528	11422	13512
2	ponorogo	284607	222834	123672	54384	32055
3	trenggalek	162356	235528	93348	33557	21982
4	tulungagung	234592	292702	124431	82131	44027
5	blitar	335090	310777	154481	65699	38056
6	kediri	375859	461130	222710	102289	52752
7	malang	757903	657086	304588	132868	64565
8	lumajang	325452	317256	106872	62268	16752
9	jember	822183	565813	248461	135358	43444
10	banyuwangi	453647	407687	223991	104617	54937
11	bondowoso	292590	188647	53339	34689	13979
12	situbondo	247440	150832	59513	50658	7601
13	probolinggo	405499	286179	92137	43641	20865
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
35	kota Mojokerto	16027	23970	18706	23547	7943
34	kota Pasuruan	29195	48521	23334	26206	9159
35	kota Mojokerto	16027	23970	18706	23547	7943

Keterangan :

X1 = Penduduk berumur 10 tahun ke atas yang tidak berijazah

X2 = Penduduk berumur 10 tahun ke atas yang berijazah SD

X3 = Penduduk berumur 10 tahun ke atas yang berijazah SLTP

X4 = Penduduk berumur 10 tahun ke atas yang berijazah SMU

X5 = Penduduk berumur 10 tahun ke atas yang berijazah SM kejuruan

Sumber : Cahyono. 2009. Laporan Praktek Kerja Lapang. Universitas Brawijaya Malang.

Lampiran 1. (Lanjutan)

Data 30. Data curah hujan harian dari 36 pos pengamatan hujan selama 3 periode di wilayah Kabupaten Indramayu

No	Stasiun	X1	X2	X3
1	Anjatan	43	39	30
2	Bondan	73	89	37
3	Bugel	36	19	24
4	Bugis	66	52	31
5	Bangkir	72	65	69
6	Bulak/kandanghaur	50	33	26
7	Bantarhuni	84	79	53
8	Cidempet	59	45	41
9	Cigugur	44	31	21
10	Cikedung	61	73	38
11	Cipancuh	77	77	43
12	Gabuswetan	59	52	25
13	Gantar	75	60	42
14	Indramayu	108	48	68
15	Jatibarang	58	58	37
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
36	Wanguk	64	34	28

Keterangan :

X1 = Pengukuran pada periode pertama

X2 = Pengukuran pada periode kedua

X3 = Pengukuran pada periode ketiga

Sumber : Haryoko, U. 2007. Pewilayahan hujan untuk menentukan pola hujan (contoh kasus kabupaten indramayu). Badan Meteorologi dan Geofisika

Lampiran 1. (Lanjutan)

Data 31. Data karakteristik berbagai merk mobil

No	Mobil	MPG	Dri	Hor	Dis	Cyl
1	Buick Estate Wagon	16.9	2.73	155	350	8
2	Ford Country Squire Wagon	15.5	2.26	142	351	8
3	Chevy Malibu Wagon	19.2	2.56	125	267	8
4	Chrysler LeBaron Wagon	18.5	2.45	150	360	8
5	Chevette	30	3.7	68	98	4
6	Toyota Corona	27.5	3.05	95	134	4
7	Datsun 510	27.2	3.54	97	119	4
8	Dodge Omni	30.9	3.37	75	105	4
9	Audi 5000	20.3	3.9	103	131	5
10	Volvo 240 GL	17	3.5	125	163	6
11	Saab 99 GLE	21.6	3.77	115	121	4
12	Peugeot 694 SL	16.2	3.58	133	163	6
13	Buick Century Special	20.6	2.73	105	231	6
14	Mercury Zephyr	20.8	3.08	85	200	6
15	Dodge Aspen	18.6	2.71	110	225	6
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
38	VW Rabbit	31.9	3.78	71	89	4

Keterangan :

X1 = Konsumsi bahan bakar (mil/galaon)

X2 = Konsumsi umbi-umbian (gram)

X3 = Konsumsi ikan (gram)

X4 = Konsumsi daging (gram)

X5 = Konsumsi telur dan susu (gram)

Sumber : Cars

(<http://lib.stat.cmu.edu/DASL/Datafiles/Cars.html>)

Lampiran 1. (Lanjutan)

Data 32. Data indikator kekuatan tekanan semen

Jenis adonan semen	X1	X2	X3	X4	X5
1	5.98	21.6	65.6	21.99	52.94
2	6	21.8	65.07	24.73	48.56
3	5.9	21.38	65.51	20.35	54.27
4	5.88	21.86	65.28	24.63	49.79
5	5.8	21.2	65.61	18.42	56.15
6	5.82	21.9	65.06	27.83	46.33
7	5.86	21.6	65.45	22.01	52.91
8	5.84	21.32	65.72	18.78	56.13
9	5.86	21.46	65.72	20.04	54.99
10	5.96	21.68	65.39	23.31	51.5
11	6.42	21.86	65.15	28.38	45.46
12	5.96	21.14	65.8	17.62	56.99
13	5.92	21.18	65.85	25.5	57.3
14	6	21.08	65.69	17.77	56.56
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
39	5.86	22.04	56.13	24	48.29

Keterangan :

X1 = Kandungan oksida alumina (Al_2O_3) dalam satuan persen

X2 = Kandungan oksida silica (SiO_2) dalam satuan persen

X3 = Kandungan oksidakapur (CaO) dalam satuan persen

X4 = Kandungan dikalsium silikat (2CaOSiO_2) dalam satuan persen

X5 = Kandungan trikalsium silikat (3CaOSiO_2) dalam satuan persen

Sumber : Suarto, E. A. 2004. *Kajian Regresi Komponen Utama dan Regresi New Stepwise dalam Pemilihan Model Regresi Linier Berganda*. Skripsi Program Studi Statistika Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang

Lampiran 1. (Lanjutan)

Data 33. Data luas daerah, ukuran biji dan tinggi tanaman beberapa spesies pohon oak

No	Spesies	X1	X2	X3
1	<i>Quercus alba</i> L.	24196	1.4	27
2	<i>Quercus bicolor</i> Willd.	7900	3.4	21
3	<i>Quercus macrocarpa</i> Michx.	23038	9.1	25
4	<i>Quercus prinoides</i> Willd.	17042	1.6	3
5	<i>Quercus Prinus</i> L.	7646	10.5	24
6	<i>Quercus stellata</i> Wang.	19938	2.5	17
7	<i>Quercus virginiana</i> Mill	7985	0.9	15
8	<i>Quercus Michauxii</i> Nutt.	8897	6.8	0.3
9	<i>Quercus lyrata</i> Walt.	8982	1.8	24
10	<i>Quercus Laceyi</i> Small.	233	0.3	11
11	<i>Quercus Chapmanii</i> Sarg.	1598	0.9	15
12	<i>Quercus Durandii</i> Buckl.	1745	0.8	23
13	<i>Quercus Muehlenbergii</i> Engelm	17042	2	24
14	<i>Quercus ilicifolia</i> Wang.	4082	1.1	3
15	<i>Quercus incana</i> Bartr.	3775	0.6	13
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
39	<i>Quercus tomentella</i> Engelm	13	7.1	18

Keterangan :

X1 = Luas daerah yang mencakup spesies pohon oak (km² x 100)

X2 = Ukuran biji buah (cm³)

X3 = Tinggi tanaman oak (meter)

Sumber : Acorns

(<http://lib.stat.cmu.edu/DASL/Datafiles/Acorn.html>)

Lampiran 1. (Lanjutan)

Data 34. Data luas wilayah, jumlah penduduk, kepadatan penduduk dan rumah tangga miskin di Kabupaten Bogor Tahun 2008

No	Kecamatan	X1	X2	X3	X4
1	Rumpin	111.01	124626	1123	7899
2	Cigudeg	158.90	11331	713	8542
3	Sukajaya	76.28	62924	825	5632
4	Jasinga	208.07	95223	458	9321
5	Tenjo	64.45	63935	992	5154
6	Parung Panjang	62.59	93558	1495	5645
7	Nanggung	135.25	88139	652	6135
8	Leuwiliang	61.77	111164	18	11429
9	Leuwisadeng	32.83	73048	2225	6598
10	Pamijahan	80.88	136006	1682	10399
11	Cibungbulang	32.66	123007	3766	9062
12	Ciampea	51.06	139478	2732	7786
13	Tenjolaya	23.68	50883	2149	4696
14	Citeureup	67.19	170123	2532	8539
15	Cibinong	43.37	251562	58	7217
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
40	Gunung Putri	56.29	22578	4011	2773

Keterangan :

X1 = Luas wilayah (km^2)

X2 = Jumlah penduduk (jiwa)

X3 = Kepadatan penduduk (jiwa/km^2)

X4 = Jumlah rumah tangga miskin (rumahtangga)

Sumber : Triana, R. L. 2009. *Analisis Dampak Pemekaran Wilayah Terhadap Pengelompokan Kecamatan Berdasarkan Beberapa Peubah Sosial Ekonomi di Kabupaten Bogor Tahun 2008*. Skripsi Fakultas Ekonomi dan Manajemen Institut Pertanian Bogor

Lampiran 1. (Lanjutan)

Data 35. Data polusi udara 41 kota di Amerika Serikat

No	Kota	X1	X2	X3	X4	X5	X6
1	Phoenix	10	70.3	213	582	6	7.05
2	Little Rock	13	61	91	132	8.2	48.52
3	San Francisco	12	56.7	453	716	8.7	20.66
4	Denver	17	51.9	454	515	9	12.95
5	Hartford	56	49.1	412	158	9	43.37
6	Wilmington	36	54	80	80	9	40.25
7	Washington	29	57.3	434	757	9.3	38.89
8	Jacksonville	14	68.4	136	529	8.8	54.47
9	Miami	10	75.5	207	335	9	59.8
10	Atlanta	24	61.5	368	497	9.1	48.34
11	Chicago	110	50.6	3344	3369	10.4	34.44
12	Indianapolis	28	52.3	361	746	9.7	38.74
13	Des Moines	17	49	104	201	11.2	30.85
14	Wichita	8	56.6	125	277	12.7	30.58
15	Louisville	30	55.6	291	593	8.3	43.11
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
41	Milwaukee	16	45.7	569	717	11.8	29.07

Keterangan :

X1 = Kandungan SO₂ udara (microgram/m³)

X2 = Rata-rata suhu setiap tahun (F)

X3 = Jumlah perusahaan yang menggunakan 20 atau lebih pekerja

X4 = Jumlah populasi (ribu jiwa)

X5 = Rata-rata kecepatan angin setiap tahun (mil/jam)

X6 = Rata-rata munculnya salju setiap tahun (inci)

Sumber : Air Pollution

(<http://lib.stat.cmu.edu/DASL/Datafiles/Airpollution.html>)

Lampiran 1. (Lanjutan)

Data 36. Data angka kematian per seribu populasi di beberapa negara

No	Negara	X1	X2	X3	X4
1	AL	2.90	17.05	1.59	6.15
2	AZ	3.52	19.80	2.75	6.61
3	AR	2.99	15.98	2.02	6.94
4	CA	4.46	22.07	2.66	7.06
5	CT	5.11	22.83	3.35	7.20
6	DE	4.78	24.55	3.36	6.45
7	DC	5.60	27.27	3.13	7.08
8	FL	4.46	23.57	2.41	6.07
9	ID	3.08	13.58	2.46	6.62
10	IL	4.75	22.80	2.95	7.27
11	IN	4.09	20.30	2.81	7.00
12	IO	4.23	16.59	2.90	7.69
13	KS	2.91	16.84	2.88	7.42
14	KY	2.86	17.71	2.13	6.41
15	LA	4.65	25.45	2.30	6.71
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
44	AK	3.46	25.88	4.32	4.90

Keterangan :

X1 = Kematian per 100K populasi karena kangker kandung kemih

X2 = Kematian per 100K populasi karena kangker paru-paru

X3 = Kematian per 100K populasi karena kangker

X4 = Kematian per 100K populasi karena leukimia

Sumber : Smoking and Cancer

(<http://lib.stat.cmu.edu/DASL/Datafiles/cigcancerdat.html>)

Lampiran 1. (Lanjutan)

Data 37. Data rata-rata konsumsi alkohol tiap daerah di Jepang dengan 5 jenis minuman beralkohol

No	Prefecture	X1	X2	X3	X4	X5
1	Aichi	8.63	2.67	56.89	1.08	1.11
2	Akita	20.31	5.25	57.09	1.04	1.77
3	Aomori	11.71	7.80	56.36	1.20	2.12
4	Chiba	7.15	5.56	40.52	1.10	1.02
5	Ehime	10.99	4.08	52.73	0.67	0.94
6	Fukui	13.31	1.57	56.51	0.62	1.07
7	Fukuoka	8.55	7.10	52.87	1.05	1.21
8	Fukushima	14.58	5.25	47.07	0.89	1.79
9	Gifu	11.65	2.09	49.13	0.73	0.87
10	Gunma	9.63	7.80	44.02	1.07	0.92
11	Hiroshima	10.97	4.68	59.14	1.09	1.25
12	Hokkaido	8.18	8.91	55.48	1.85	1.72
13	Hyogo	9.23	2.35	51.86	0.99	1.21
14	Ibaragi	9.88	5.06	42.50	0.81	1.01
15	Ishikawa	15.62	1.74	58.01	0.89	1.19
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
46	Yamanashi	9.73	7.5	42.97	5.08	0.93

Keterangan :

X1 = Sake (yen)

X2 = Shochu (yen)

X3 = Bear (yen)

X4 = Wine (yen)

X5 = Whisky (yen)

Sumber : Wijayanto, H. 1998. *Analisis Pengembangan Konsep Produk Menggunakan Metode Biplot*. Forum statistika dan Komputai

Lampiran 1. (Lanjutan)

Data 38. Data pendapatan perkapita, belanja umum dan karakteristik ekonomi beberapa negara

No	Negara	X1	X2	X3	X4
1	ME	19.7	6.9	29.6	11
2	NH	17.7	14.7	26.4	11.2
3	VT	0	3.7	28.5	11.2
4	MA	85.2	10.2	25.1	11.1
5	RI	86.2	1	25.3	10.4
6	CT	77.6	25.4	25.2	9.6
7	NY	85.5	12.9	24	10.1
8	NJ	78.9	25.5	24.8	9.2
9	PA	77.9	7.8	25.7	10
10	DE	68.8	39.9	26.4	8
11	MD	78.2	31.1	27.5	7.3
12	VA	50.9	21.9	28.8	7.3
13	MI	73.1	22.2	28	8.2
14	OH	69.5	21.8	26.9	9.2
15	IN	48.1	18.3	27.5	9.6
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
48	CA	86.5	48.5	26.2	8.8

Keterangan :

X1 = Persentase populasi yang tinggal di area metropolitan

X2 = Persentase perubahan populasi pada tahun 1950-1960

X3 = Persentase populasi yang berumur 5-19 tahun

X4 = Persentase populasi diatas 65 tahun

Sumber : Public Expenditure

(<http://lib.stat.cmu.edu/DASL/Datafiles/pubexpendat.htm>)

Lampiran 1. (Lanjutan)

Data 39. Data rata-rata gaji profesor 50 universitas di Amerika Serikat

No	Universitas	X1	X2	X3
1	DUKE	83.0	57.5	46.1
2	VANDERBILT	78.9	49.7	42.5
3	WASHINGTON UNIV.	75.4	51.4	43.8
4	TULANE	70.2	50.8	41.5
5	CAL TECH	93.3	70.0	56.4
6	CARNEGIE MELLON	79.3	55.0	49.4
7	CORNELL	71.0	52.4	44.9
8	VIRGINIA	71.6	47.8	39.5
9	TEXAS	70.3	45.7	40.8
10	ROCHESTER	75.5	51.7	43.9
11	NEBRASKA	63.5	46.0	40.3
12	UNIV. OF IOWA	69.5	50.4	42.8
13	STANFORD	91.2	64.4	50.0
14	COLORADO	64.9	49.6	42.4
15	UNIV. OF PENN.	90.5	64.1	56.2
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
49	COLUMBIA	83.4	55.8	43.0

Keterangan :

X1 = Rata-rata gaji profesor pada tahun 1992

X2 = Rata-rata perkumpulan profesor pada tahun 1992

X3 = Rata-rata gaji asisten profesor pada tahun 1992

Sumber : Faculty Salaries

(<http://lib.stat.cmu.edu/DASL/Datafiles/facultysalaries.html>)

Lampiran 1. (Lanjutan)

Data 40. Data plasma nutfah kapas

No	Plasma Nutfah	X1	X2	X3	X4	X5	X6
1	AC 134	3.90	5.90	10.30	13.40	26.00	25.00
2	Acala Messila Valley	1.80	2.50	3.70	8.10	10.05	17.70
3	Stoneville 7	1.90	4.40	7.00	4.90	9.00	19.70
4	SR 1-F4-71	1.90	3.20	4.90	6.00	11.70	14.00
5	L4 x Rex/1	1.40	2.80	2.70	4.70	9.30	20.80
6	Delta Queen	2.00	2.50	4.20	5.80	11.90	15.35
7	Compinas 81/4	2.20	5.80	6.70	9.30	21.30	23.50
8	HL 1	1.50	3.10	6.40	5.20	9.80	9.80
9	Kapas Mesir	3.10	6.90	9.10	3.10	6.10	13.70
10	Pima Cotton	3.70	5.80	7.20	4.50	4.20	4.70
11	76077-8301-2	3.20	3.80	4.40	4.80	11.90	16.00
12	H6	2.60	3.40	5.60	7.30	15.40	13.60
13	ISA 205 B	3.70	6.50	9.60	10.20	25.00	26.00
14	SU 28	2.60	3.10	5.10	4.90	7.50	8.70
15	Kanesia 2	2.10	4.80	5.40	8.70	18.90	21.20
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
50	Kanesia 9	3.40	3.70	7.00	10.35	23.50	26.00

Keterangan :

X1 = Jumlah buah segmen tengah

X2 = Jumlah buah segmen bawah

X3 = Jumlah buah segmen atas

X4 = Produksi segmen tengah

X5 = Produksi segmen tengah

X6 = Produksi segmen atas

Sumber : Sa'adah, Z. 2009. Laporan Praktek Kerja Lapang. Universitas Brawijaya Malang

Lampiran 1. (Lanjutan)

Data 41. Data pelanggan PLN di kecamatan Sleman

Objek	X1	X2	X3	X4
1	700000	60000	40	157
2	300000	185000	40	34
3	650000	76000	96	53
4	2650000	75000	250	235
5	400000	150000	48	103
6	300000	75000	40	82
7	1963000	45000	220	222
8	950000	22500	27	108
9	300000	63000	40	63
10	700000	225000	104	64
11	1200000	50000	180	151
12	300000	30000	120	82
13	1586000	150000	200	82
14	1775000	325000	150	119
15	700000	50000	150	226
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
50	600000	35000	56	56

Keterangan :

X1 = Pendapatan total keluarga perbulan (Rupiah)

X2 = Pengeluaran bahan bakar (minyak tanah, kayu bakar, gas, dan premium/solar) perbulan (Rupiah)

X3 = Luas bangunan (m^2)

X4 = Permintaan listrik pada rumah tangga (Kwh)

Sumber : Hayati, F. 2008. *Analisis Faktor-Faktor yang Mempengaruhi Permintaan Konsumen Terhadap Listrik pada Rumah tangga (Studi Kasus Dusun Nambongan Desa Caturhar Kecamatan Sleman Kabupaten Sleman Daerah Istimewa Yogyakarta)*. Skripsi UII Yogyakarta

Lampiran 1. (Lanjutan)

Data 42. Data rata-rata intensitas serangan, tinggi tanaman, rasio BK/BB dari hasil penelitian uji ketahanan 50 varietas kacang hijau terhadap Black Gram Mottle Virus

No	Varietas	X1	X2	X3
1	Betet	77.40	15.20	0.233
2	Bhakti	78.82	15.00	0.190
3	Gelatik	77.86	14.40	0.179
4	Kenari	80.52	14.60	0.216
5	Lokal wongsorejo	72.16	15.80	0.218
6	Manyar	68.82	18.00	0.189
7	Merak	67.57	10.20	0.224
8	Merpati	77.01	13.40	0.209
9	Parkit	74.94	16.80	0.199
10	Sriti	73.78	14.00	0.198
11	Walet	74.15	14.80	0.189
12	EVO 947	73.30	16.00	0.193
13	IPBM 79-8-82	77.10	14.20	0.206
14	MLG 1	72.49	22.00	0.217
15	MLG 2	74.30	21.96	0.236
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
50	VC 61488-16	74.89	19.60	0.185

Keterangan :

X1 = Intensitas serangan (%)

X2 = Tinggi tanaman (cm)

X3 = Rasio BK/BB

Sumber : Utami, R. S. 2008. *Uji Ketahanan 50 Varietas/ Galur Kacang Hijau Terhadap Black Gram Mottle Virus dan Pengaruh Umur Tanaman Waktu Inkubasi BGMV Terhadap Pertumbuhan dan Hasil Kacang Hijau (Vigna Radiatu L. Wilczek)*. Skripsi Universitas Brawijaya Malang

Lampiran 1. (Lanjutan)

Data 43. Data jumlah polong, tinggi tanaman dan hasil genotipe terpilih beberapa varietas kacang tanah

No	Genotip	X1	X2	X3
1	Mlg 9096	2.75	7.25	32.90
2	Mlg 9098	3.00	7.75	31.72
3	Mlg 9094	4.38	11.50	32.05
4	Mlg 9089	3.47	10.75	35.47
5	Mlg 9055	2.25	11.50	51.97
6	Mlg 9090	2.34	9.25	39.63
7	Mlg 9070	3.30	13.50	45.73
8	Mlg 9007	3.02	14.00	38.88
9	Mlg 9080	3.23	13.50	40.40
10	Mlg 9020	3.47	13.50	44.45
11	Mlg 9093	3.84	8.75	29.90
12	Mlg 9048	2.38	12.00	36.97
13	Mlg 9076	2.47	8.50	35.13
14	Mlg 9049	2.35	12.50	35.47
15	Mlg 9056	2.61	11.75	55.90
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
51	Lokal	3.41	11.25	37.27

Keterangan :

X1 = Hasil polong (gram/tanaman)

X2 = Jumlah polong

X3 = Tinggi tanaman (cm)

Sumber : Azzahra, F. dan Koesrini. 2002. *Penampilan Genotip-genotip Kacang Tanah di Lahan Lebak Dangkal*. Balai Penelitian Pertanian Lahan Rawa

Lampiran 1. (Lanjutan)

Data 44. Data rata-rata gaji pengajar dan pengeluaran per siswa di 50 negara bagian dan daerah Kolombia

No	Negara	X1	X2
1	Maine	19.6	3.35
2	Vt.	20.3	3.55
3	R.I.	29.5	4.67
4	N.Y.	30.7	5.71
5	Pa.	25.9	4.17
6	Ind.	24.3	3.16
7	Mich.	30.2	3.78
8	Minn.	27.4	3.98
9	Mo.	22	3.16
10	S.Dak	18.1	2.97
11	Kan.	27.6	3.91
12	Md.	27.2	4.35
13	Va.	23.4	3.59
14	N.C.	22.8	3.37
15	Ga.	22.1	2.98
.	.	.	.
.	.	.	.
.	.	.	.
51	Alaska	41.5	8.35

Keterangan :

X1 = Jumlah pembayaran (ribu)

X2 = Rata-rata jumlah pengeluaran per siswa (ribu)

Sumber : Educational Spending

(<http://lib.stat.cmu.edu/DASL/Datafiles/EducationalSpending.html>)

Lampiran 1. (Lanjutan)

Data 45. Data rata-rata gaji guru dan pengeluaran setiap siswa di 50 negara bagian dan daerah Kolombia

No	Negara	X1	X2
1	ME	19583	3346
2	NH	20263	3114
3	VT	20325	3554
4	MA	26800	4642
5	RI	29470	4669
6	CT	26610	4888
7	NY	30678	5710
8	NJ	27170	5536
9	PA	25853	4168
10	OH	24500	3547
11	IN	24274	3159
12	IL	27170	3621
13	MI	30168	3782
14	WI	26525	4247
15	MN	27360	3982
.	.	.	.
.	.	.	.
.	.	.	.
51	HA	25845	3766

Keterangan :

X1 = Rata-rata gaji guru sekolah umum setiap tahun (\$)

X2 = Pengeluaran setiap siswa di sekolah umum (\$)

Sumber : Teacher Pay by States

(<http://lib.stat.cmu.edu/DASI/Datafiles/teacherpaydat.html>)

Lampiran 1. (Lanjutan)

Data 46. Data karakteristik 53 danau di Florida

No	Danau	X1	X2	X3	X4	X5
1	Alligator	5.90	6.10	3.00	0.70	1.23
2	Annie	3.50	5.10	1.90	3.20	1.33
3	Apopka	116.00	9.10	44.10	128.30	0.04
4	Blue Cypress	39.40	6.90	16.40	3.50	0.44
5	Brick	2.50	4.60	2.90	1.80	1.20
6	Bryant	19.60	7.30	4.50	44.10	0.27
7	Cherry	5.20	5.40	2.80	3.40	0.48
8	Crescent	71.40	8.10	55.20	33.70	0.19
9	Deer Point	26.40	5.80	9.20	1.60	0.83
10	Dias	4.80	6.40	4.60	22.50	0.81
11	Dorr	6.60	5.40	2.70	14.90	0.71
12	Down	16.50	7.20	13.80	4.00	0.50
13	Eaton	25.40	7.20	25.20	11.60	0.49
14	East Tohopeka	7.10	5.80	5.20	5.80	1.16
15	Farm-13	128.00	7.60	86.50	71.10	0.05
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
53	Yale	71.80	7.90	20.50	8.80	0.27

Keterangan :

X1 = Kalsium karbonat (mg/L)

X2 = pH

X3 = Kalsium (mg/L)

X4 = Kloropil (mg/L)

X5 = Rata-rata konsentrasi merkuri (parts/juta)

Sumber : Mercury in Bass

(<http://lib.stat.cmu.edu/DASL/Datafiles/mercuryinbass.html>)

Lampiran 1. (Lanjutan)

Data 47. Data karakteristik beberapa varietas kacang tanah

No	Plasma nutfah	X1	X2	X3	X4	X5	X6
1	LOCAL	30	22.08	45.6	24.6	52.8	33.8
2	Balit Bogor	31	20.82	54.6	36.4	48.8	30.4
3	Balit Bogor	33	19.18	48.8	28.4	48.6	38
4	Balit Bogor	31	18.2	66.8	38.6	68	37
5	Balit Bogor	31	17.46	50.8	27	50	42.8
6	I-10	32	7.04	56.2	29.8	56	24.8
7	Purworejo	42	5.44	47	25.8	42.6	30.2
8	IV-18	31	17.4	55	36	58	23.6
9	V-5	32	27.76	54	31.4	54	34.8
10	V-26	31	25.24	61.6	35.6	70	34.4
11	V-35	31	19	59	27.4	59.8	44.4
12	V-39	31	15.62	50.8	29.2	49.4	38.4
13	VI-2	31	14.26	64.2	35.4	62.2	37.8
14	VI-12	31	17.74	51.2	28.4	50.4	28.6
15	VI-24	31	8.34	57	34.4	57.2	33.4
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
58	ICGV 94217	32	22.04	53.6	27	46.2	46.2

Keterangan :

X1 = Umur berbunga (hari)

X2 = Berat polong (gr)

X3 = Panjang daun (mm)

X4 = Lebar daun (mm)

X5 = Panjang petiol (mm)

X6 = Tinggi tanaman (cm)

Sumber : Lestari, T. K. 2009. Laporan Praktek Kerja Lapang. Universitas Brawijaya Malang

Lampiran 1. (Lanjutan)

Data 48. Data karakteristik beberapa merek sereal

No	Nama	X1	X2	X3	X4	...	X8
1	100%_Bran	70	4	1	130	...	280
2	100%_Natural_Bran	120	3	5	15	...	135
3	All-Bran	70	4	1	260	...	320
4	All-Bran_with_Extra_	50	4	0	140	...	330
5	Apple_Cinnamon	110	2	2	180	...	70
6	Apple_Jacks	110	2	0	125	...	30
7	Basic_4	130	3	2	210	...	100
8	Bran_Chex	90	2	1	200	...	125
9	Bran_Flakes	90	3	0	210	...	190
10	Cap'n'Crunch	120	1	2	220	...	35
11	Cheerios	110	6	2	290	...	105
12	Cinnamon_Toast	120	1	3	210	...	45
13	Clusters	110	3	2	140	...	105
14	Cocoa_Puffs	110	1	1	180	...	55
15	Corn_Chex	110	2	0	280	...	25
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	35
69	Total_Corn_Flakes	110	2	1	200	...	20

Keterangan :

X1 = Kalori per bungkus

X2 = Protein (gram)

X3 = Lemak (gram)

X4 = Sodium (milligram)

X5 = Serat (gram)

X6 = Karbohidrat (gram)

X7 = Gula (gram)

X8 = Potasium (milligram)

Sumber : Cereals

(<http://lib.stat.cmu.edu/DASL/Datafiles/cereals.html>)

Lampiran 1. (Lanjutan)

Data 49. Data industri tempe di kota Malang

No	Nama/ Perusahaan	X1	X2	X3
1	Kasian	89425	18250	2150
2	Samuki	143080	29200	1750
3	Supangat	71540	14600	1000
4	Mani	26827	5475	500
5	Sulkah	71540	14600	1950
6	Rasemo	89425	18250	2200
7	Sekak	89425	18250	2250
8	Slamet	35770	7300	1800
9	Ngatmani	17885	3650	1650
10	Reban	44713	9125	1825
11	Tawi	71540	14600	1950
12	Syukur	35770	7300	1775
13	Mulyono	53655	10950	1900
14	Ridwan	357700	73000	6165
15	Mufid	89425	18250	2200
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
74	Solikin	89425	18250	2775

Keterangan :

X1 = Nilai produksi/tahun (Rp. 1000)

X2 = Jumlah bahan baku/tahun (Kg)

X3 = Nilai investasi (Rp. 1000)

Sumber : Hermayudi, W. A. 2004. *Pengelompokkan Industri Tempe Di Wilayah Kota Malang*. Skripsi Fakultas Teknologi Pertanian Universitas Brawijaya Malang

Lampiran 1. (Lanjutan)

Data 50. Data karakteristik beberapa perusahaan

No	Perusahaan	X1	X2	X3		X6
1	Air Products	2687	1870	1890		18.2
2	Allied Signal	13271	9115	8190		143.8
3	American Electric Power	13621	4848	4572		23.4
4	American Savings Bank FSB	3614	367	90		1.1
5	AMR	6425	6131	2448		49.5
6	Apple Computer	1022	1754	1370		4.8
7	Armstrong World Industries	1093	1679	1070		20.8
8	Bally Manufacturing	1529	1295	444		19.4
9	Bank South	2788	271	304		2.1
10	Bell Atlantic	19788	9084	10636		79.4
11	H&R Block	327	542	959		2.8
12	Brooklyn Union Gas	1117	1038	478		3.8
13	California First Bank	5401	550	376		4.1
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
79	FW Woolworth	2535	5958	1921	177	118.1

Keterangan :

X1 = Jumlah kekayaan (juta)

X2 = Jumlah penjualan (juta)

X3 = Nilai pasar perusahaan (juta)

X4 = Keuntungan (juta)

X5 = Cash flow (juta)

X6 = Jumlah tenaga kerja (ribu)

Sumber : Companies

(<http://lib.stat.cmu.edu/DASL/Datafiles/companies.html>)

Lampiran 2. Uji Koefisien Korelasi

Data 1

		Correlations								
		X1	X2	X3	X4	X5	X6	X7	X8	X9
X1	Pearson Correlation	1	-.160	.153	.512	-.098	.229	.147	.347	.162
	Sig. (2-tailed)		.637	.653	.107	.774	.498	.666	.296	.633
	N	11	11	11	11	11	11	11	11	11
X2	Pearson Correlation	-.160	1	.701*	.641*	.821**	.758**	.836**	.540	.833**
	Sig. (2-tailed)	.637		.016	.034	.002	.007	.001	.086	.001
	N	11	11	11	11	11	11	11	11	11
X3	Pearson Correlation	.153	.701*	1	.812**	.880**	.912**	.895**	.678*	.898**
	Sig. (2-tailed)	.653	.016		.002	.000	.000	.000	.022	.000
	N	11	11	11	11	11	11	11	11	11
X4	Pearson Correlation	.512	.641*	.812**	1	.779**	.924**	.908**	.804**	.887**
	Sig. (2-tailed)	.107	.034	.002		.005	.000	.000	.003	.000
	N	11	11	11	11	11	11	11	11	11
X5	Pearson Correlation	-.098	.821**	.880**	.779**	1	.930**	.923**	.640*	.897**
	Sig. (2-tailed)	.774	.002	.000	.005		.000	.000	.034	.000
	N	11	11	11	11	11	11	11	11	11
X6	Pearson Correlation	.229	.758**	.912**	.924**	.930**	1	.962**	.747**	.952**
	Sig. (2-tailed)	.498	.007	.000	.000	.000		.000	.008	.000
	N	11	11	11	11	11	11	11	11	11
X7	Pearson Correlation	.147	.836**	.895**	.908**	.923**	.962**	1	.812**	.963**
	Sig. (2-tailed)	.666	.001	.000	.000	.000	.000		.002	.000
	N	11	11	11	11	11	11	11	11	11
X8	Pearson Correlation	.347	.540	.678*	.804**	.640*	.747**	.812**	1	.684*
	Sig. (2-tailed)	.296	.086	.022	.003	.034	.008	.002		.020
	N	11	11	11	11	11	11	11	11	11
X9	Pearson Correlation	.162	.833**	.898**	.887**	.897**	.952**	.963**	.684*	1
	Sig. (2-tailed)	.633	.001	.000	.000	.000	.000	.000	.020	
	N	11	11	11	11	11	11	11	11	11

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Data 2 (Tidak ada korelasi)

Data 3

		Correlations				
		X1	X2	X3	X4	X5
X1	Pearson Correlation	1	.895**	.997**	.876**	1.000**
	Sig. (2-tailed)		.000	.000	.000	.000
	N	14	14	14	14	14
X2	Pearson Correlation	.895**	1	.900**	.776**	.898**
	Sig. (2-tailed)	.000		.000	.001	.000
	N	14	14	14	14	14
X3	Pearson Correlation	.997**	.900**	1	.877**	.997**
	Sig. (2-tailed)	.000	.000		.000	.000
	N	14	14	14	14	14
X4	Pearson Correlation	.876**	.776**	.877**	1	.877**
	Sig. (2-tailed)	.000	.001	.000		.000
	N	14	14	14	14	14
X5	Pearson Correlation	1.000**	.898**	.997**	.877**	1
	Sig. (2-tailed)	.000	.000	.000	.000	
	N	14	14	14	14	14

**. Correlation is significant at the 0.01 level (2-tailed).

Lampiran 2. (Lanjutan)

Data 4

		Correlations						
		X1	X2	X3	X4	X5	X6	X7
X1	Pearson Correlation	1	.435	.437	.556*	.232	-.068	.063
	Sig. (2-tailed)		.092	.090	.025	.388	.802	.817
	N	16	16	16	16	16	16	16
X2	Pearson Correlation	.435	1	.315	.772**	.497	.457	.372
	Sig. (2-tailed)	.092		.235	.000	.050	.075	.155
	N	16	16	16	16	16	16	16
X3	Pearson Correlation	.437	.315	1	.606*	.340	.320	.543*
	Sig. (2-tailed)	.090	.235		.013	.197	.227	.030
	N	16	16	16	16	16	16	16
X4	Pearson Correlation	.556*	.772**	.606*	1	.420	.343	.379
	Sig. (2-tailed)	.025	.000	.013		.105	.194	.148
	N	16	16	16	16	16	16	16
X5	Pearson Correlation	.232	.497	.340	.420	1	.759**	.275
	Sig. (2-tailed)	.388	.050	.197	.105		.001	.303
	N	16	16	16	16	16	16	16
X6	Pearson Correlation	-.068	.457	.320	.343	.759**	1	.309
	Sig. (2-tailed)	.802	.075	.227	.194	.001		.245
	N	16	16	16	16	16	16	16
X7	Pearson Correlation	.063	.372	.543*	.379	.275	.309	1
	Sig. (2-tailed)	.817	.155	.030	.148	.303	.245	
	N	16	16	16	16	16	16	16

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Data 5

		Correlations			
		X1	X2	X3	X4
X1	Pearson Correlation	1	-.617*	-.420	.108
	Sig. (2-tailed)		.011	.105	.690
	N	16	16	16	16
X2	Pearson Correlation	-.617*	1	.469	-.167
	Sig. (2-tailed)	.011		.067	.536
	N	16	16	16	16
X3	Pearson Correlation	-.420	.469	1	.468
	Sig. (2-tailed)	.105	.067		.068
	N	16	16	16	16
X4	Pearson Correlation	.108	-.167	.468	1
	Sig. (2-tailed)	.690	.536	.068	
	N	16	16	16	16

*. Correlation is significant at the 0.05 level (2-tailed).

Data 6

		Correlations			
		X1	X2	X3	X4
X1	Pearson Correlation	1	.104	.789**	.300
	Sig. (2-tailed)		.701	.000	.258
	N	16	16	16	16
X2	Pearson Correlation	.104	1	-.116	-.089
	Sig. (2-tailed)	.701		.669	.744
	N	16	16	16	16
X3	Pearson Correlation	.789**	-.116	1	.467
	Sig. (2-tailed)	.000	.669		.068
	N	16	16	16	16
X4	Pearson Correlation	.300	-.089	.467	1
	Sig. (2-tailed)	.258	.744	.068	
	N	16	16	16	16

**. Correlation is significant at the 0.01 level (2-tailed).

Lampiran 2. (Lanjutan)

Data 7

Correlations

		X1	X2	X3
X1	Pearson Correlation	1	-.550*	-.548*
	Sig. (2-tailed)		.027	.028
	N	16	16	16
X2	Pearson Correlation	-.550*	1	.992**
	Sig. (2-tailed)	.027		.000
	N	16	16	16
X3	Pearson Correlation	-.548*	.992**	1
	Sig. (2-tailed)	.028	.000	
	N	16	16	16

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Data 8 (Tidak ada korelasi)

Data 9

Correlations

		X1	X2	X3
X1	Pearson Correlation	1	.600*	.829**
	Sig. (2-tailed)		.014	.000
	N	16	16	16
X2	Pearson Correlation	.600*	1	.894**
	Sig. (2-tailed)	.014		.000
	N	16	16	16
X3	Pearson Correlation	.829**	.894**	1
	Sig. (2-tailed)	.000	.000	
	N	16	16	16

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Data 10

Correlations

		X1	X2	X3	X4	X5	X6
X1	Pearson Correlation	1	.579*	.713**	.735**	.721**	.678**
	Sig. (2-tailed)		.015	.001	.001	.001	.003
	N	17	17	17	17	17	17
X2	Pearson Correlation	.579*	1	.548*	.352	.264	.186
	Sig. (2-tailed)	.015		.023	.166	.306	.476
	N	17	17	17	17	17	17
X3	Pearson Correlation	.713**	.548*	1	.828**	.751**	.556*
	Sig. (2-tailed)	.001	.023		.000	.001	.020
	N	17	17	17	17	17	17
X4	Pearson Correlation	.735**	.352	.828**	1	.985**	.905**
	Sig. (2-tailed)	.001	.166	.000		.000	.000
	N	17	17	17	17	17	17
X5	Pearson Correlation	.721**	.264	.751**	.985**	1	.933**
	Sig. (2-tailed)	.001	.306	.001	.000		.000
	N	17	17	17	17	17	17
X6	Pearson Correlation	.678**	.186	.556*	.905**	.933**	1
	Sig. (2-tailed)	.003	.476	.020	.000	.000	
	N	17	17	17	17	17	17

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Lampiran 2. (Lanjutan)

Data 11

		Correlations			
		X1	X2	X3	X4
X1	Pearson Correlation	1	.800**	-.677**	-.034
	Sig. (2-tailed)		.000	.002	.894
	N	18	18	18	18
X2	Pearson Correlation	.800**	1	-.685**	.124
	Sig. (2-tailed)	.000		.002	.625
	N	18	18	18	18
X3	Pearson Correlation	-.677**	-.685**	1	.257
	Sig. (2-tailed)	.002	.002		.304
	N	18	18	18	18
X4	Pearson Correlation	-.034	.124	.257	1
	Sig. (2-tailed)	.894	.625	.304	
	N	18	18	18	18

** . Correlation is significant at the 0.01 level (2-tailed).

Data 12

		Correlations		
		X1	X2	X3
X1	Pearson Correlation	1	.359	.009
	Sig. (2-tailed)		.143	.973
	N	18	18	18
X2	Pearson Correlation	.359	1	.472*
	Sig. (2-tailed)	.143		.048
	N	18	18	18
X3	Pearson Correlation	.009	.472*	1
	Sig. (2-tailed)	.973	.048	
	N	18	18	18

* . Correlation is significant at the 0.05 level (2-tailed).

Data 13

		Correlations		
		X1	X2	X3
X1	Pearson Correlation	1	.679**	.127
	Sig. (2-tailed)		.002	.615
	N	18	18	18
X2	Pearson Correlation	.679**	1	.012
	Sig. (2-tailed)	.002		.962
	N	18	18	18
X3	Pearson Correlation	.127	.012	1
	Sig. (2-tailed)	.615	.962	
	N	18	18	18

** . Correlation is significant at the 0.01 level (2-tailed).

Data 14

		Correlations		
		X1	X2	X3
X1	Pearson Correlation	1	-.898**	-.859**
	Sig. (2-tailed)		.000	.000
	N	19	19	19
X2	Pearson Correlation	-.898**	1	.588**
	Sig. (2-tailed)	.000		.008
	N	19	19	19
X3	Pearson Correlation	-.859**	.588**	1
	Sig. (2-tailed)	.000	.008	
	N	19	19	19

** . Correlation is significant at the 0.01 level (2-tailed).

Lampiran 2. (Lanjutan)
Data 15 (Tidak ada korelasi)
Data 16 (Tidak ada korelasi)

Data 17

Correlations			
		X1	X2
X1	Pearson Correlation	1	.651**
	Sig. (2-tailed)		.001
	N	21	21
X2	Pearson Correlation	.651**	1
	Sig. (2-tailed)	.001	
	N	21	21

** Correlation is significant at the 0.01 level

Data 18

Correlations				
		X1	X2	X3
X1	Pearson Correlation	1	-.214	-.866**
	Sig. (2-tailed)		.315	.000
	N	24	24	24
X2	Pearson Correlation	-.214	1	-.303
	Sig. (2-tailed)	.315		.150
	N	24	24	24
X3	Pearson Correlation	-.866**	-.303	1
	Sig. (2-tailed)	.000	.150	
	N	24	24	24

** Correlation is significant at the 0.01 level (2-tailed).

Data 19

Correlations										
		X1	X2	X3	X4	X5	X6	X7	X8	X9
X1	Pearson Correlation	1	.153	.586**	.503*	.061	-.500*	.135	-.349	-.074
	Sig. (2-tailed)		.465	.002	.010	.772	.011	.519	.087	.724
	N	25	25	25	25	25	25	25	25	25
X2	Pearson Correlation	.153	1	.620**	.281	-.234	-.414*	.314	-.635**	-.061
	Sig. (2-tailed)	.465		.001	.173	.260	.040	.127	.001	.771
	N	25	25	25	25	25	25	25	25	25
X3	Pearson Correlation	.586**	.620**	1	.576**	.066	-.712**	.452*	-.560**	-.046
	Sig. (2-tailed)	.002	.001		.003	.755	.000	.023	.004	.829
	N	25	25	25	25	25	25	25	25	25
X4	Pearson Correlation	.503*	.281	.576**	1	.138	-.593**	.222	-.621**	-.408*
	Sig. (2-tailed)	.010	.173	.003		.511	.002	.285	.001	.043
	N	25	25	25	25	25	25	25	25	25
X5	Pearson Correlation	.061	-.234	.066	.138	1	-.524**	.404*	-.147	.266
	Sig. (2-tailed)	.772	.260	.755	.511		.007	.045	.483	.198
	N	25	25	25	25	25	25	25	25	25
X6	Pearson Correlation	-.500*	-.414*	-.712**	-.593**	-.524**	1	-.533**	.651**	.047
	Sig. (2-tailed)	.011	.040	.000	.002	.007		.006	.000	.825
	N	25	25	25	25	25	25	25	25	25
X7	Pearson Correlation	.135	.314	.452*	.222	.404*	-.533**	1	-.474*	.084
	Sig. (2-tailed)	.519	.127	.023	.285	.045	.006		.017	.688
	N	25	25	25	25	25	25	25	25	25
X8	Pearson Correlation	-.349	-.635**	-.560**	-.621**	-.147	.651**	-.474*	1	.375
	Sig. (2-tailed)	.087	.001	.004	.001	.483	.000	.017		.065
	N	25	25	25	25	25	25	25	25	25
X9	Pearson Correlation	-.074	-.061	-.046	-.408*	.266	.047	.084	.375	1
	Sig. (2-tailed)	.724	.771	.829	.043	.198	.825	.688	.065	
	N	25	25	25	25	25	25	25	25	25

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Data 20

Correlations			
		X1	X2
X1	Pearson Correlation	1	.335
	Sig. (2-tailed)		.102
	N	25	25
X2	Pearson Correlation	.335	1
	Sig. (2-tailed)	.102	
	N	25	25
X3	Pearson Correlation	.135	-.537**
	Sig. (2-tailed)	.519	.006
	N	25	25

** Correlation is significant at the 0.01 level (2-tailed).

Lampiran 2. (Lanjutan)

Data 21

Correlations										
		X1	X2	X3	X4	X5	X6	X7	X8	X9
X1	Pearson Correlation	1	.036	-.671**	-.400*	-.538**	-.737**	-.220	-.747**	-.565**
	Sig. (2-tailed)		.862	.000	.043	.005	.000	.281	.000	.003
	N	26	26	26	26	26	26	26	26	26
X2	Pearson Correlation	.036	1	.445*	.405*	-.026	-.397*	-.443*	-.281	.157
	Sig. (2-tailed)	.862		.023	.040	.901	.045	.024	.164	.445
	N	26	26	26	26	26	26	26	26	26
X3	Pearson Correlation	-.671**	.445*	1	.385	.494*	.204	-.156	.154	.351
	Sig. (2-tailed)	.000	.023		.052	.010	.318	.447	.452	.079
	N	26	26	26	26	26	26	26	26	26
X4	Pearson Correlation	-.400*	.405*	.385	1	.060	.202	.110	.132	.375
	Sig. (2-tailed)	.043	.040	.052		.771	.323	.593	.519	.059
	N	26	26	26	26	26	26	26	26	26
X5	Pearson Correlation	-.538**	-.026	.494*	.060	1	.356	.016	.158	.388
	Sig. (2-tailed)	.005	.901	.010	.771		.074	.937	.440	.050
	N	26	26	26	26	26	26	26	26	26
X6	Pearson Correlation	-.737**	-.397*	.204	.202	.356	1	.366	.572**	.188
	Sig. (2-tailed)	.000	.045	.318	.323	.074		.066	.002	.359
	N	26	26	26	26	26	26	26	26	26
X7	Pearson Correlation	-.220	-.443*	-.156	.110	.016	.366	1	.108	-.246
	Sig. (2-tailed)	.281	.024	.447	.593	.937	.066		.601	.226
	N	26	26	26	26	26	26	26	26	26
X8	Pearson Correlation	-.747**	-.281	.154	.132	.158	.572**	.108	1	.568**
	Sig. (2-tailed)	.000	.164	.452	.519	.440	.002	.601	.002	
	N	26	26	26	26	26	26	26	26	26
X9	Pearson Correlation	-.565**	.157	.351	.375	.388	.188	-.246	.568**	1
	Sig. (2-tailed)	.003	.445	.079	.059	.050	.359	.226	.002	
	N	26	26	26	26	26	26	26	26	26

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Data 22

Correlations						
		X1	X2	X3	X4	X5
X1	Pearson Correlation	1	-.252	-.318	-.245	-.363
	Sig. (2-tailed)		.205	.106	.219	.063
	N	27	27	27	27	27
X2	Pearson Correlation	-.252	1	.500**	.554**	.666**
	Sig. (2-tailed)	.205		.008	.003	.000
	N	27	27	27	27	27
X3	Pearson Correlation	-.318	.500**	1	.503**	.525**
	Sig. (2-tailed)	.106	.008		.008	.005
	N	27	27	27	27	27
X4	Pearson Correlation	-.245	.554**	.503**	1	.950**
	Sig. (2-tailed)	.219	.003	.008		.000
	N	27	27	27	27	27
X5	Pearson Correlation	-.363	.666**	.525**	.950**	1
	Sig. (2-tailed)	.063	.000	.005	.000	
	N	27	27	27	27	27

** . Correlation is significant at the 0.01 level (2-tailed).

Lampiran 2. (Lanjutan)

Data 23

Correlations									
		X1	X2	X3	X4	X5	X6	X7	X8
X1	Pearson Correlation	1	.957**	.916**	.963**	.822**	.773**	.584**	.395*
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.001	.037
	N	28	28	28	28	28	28	28	28
X2	Pearson Correlation	.957**	1	.848**	.933**	.829**	.784**	.636**	.466*
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.012
	N	28	28	28	28	28	28	28	28
X3	Pearson Correlation	.916**	.848**	1	.955**	.782**	.778**	.542**	.404*
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.003	.033
	N	28	28	28	28	28	28	28	28
X4	Pearson Correlation	.963**	.933**	.955**	1	.831**	.803**	.611**	.438*
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.001	.020
	N	28	28	28	28	28	28	28	28
X5	Pearson Correlation	.822**	.829**	.782**	.831**	1	.869**	.684**	.529**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.004
	N	28	28	28	28	28	28	28	28
X6	Pearson Correlation	.773**	.784**	.778**	.803**	.869**	1	.734**	.615**
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000
	N	28	28	28	28	28	28	28	28
X7	Pearson Correlation	.584**	.636**	.542**	.611**	.684**	.734**	1	.894**
	Sig. (2-tailed)	.001	.000	.003	.001	.000	.000		.000
	N	28	28	28	28	28	28	28	28
X8	Pearson Correlation	.395*	.466*	.404*	.438*	.529**	.615**	.894**	1
	Sig. (2-tailed)	.037	.012	.033	.020	.004	.000	.000	
	N	28	28	28	28	28	28	28	28

** . Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Data 24

Correlations		
	X1	X2
X1	Pearson Correlation	1
	Sig. (2-tailed)	.582**
	N	.001
		28
X2	Pearson Correlation	.582**
	Sig. (2-tailed)	.001
	N	28
		28

** . Correlation is significant at the 0.01 level

Data 25

Correlations					
	X1	X2	X3	X4	X5
X1	Pearson Correlation	1	-.008	-.064	.496**
	Sig. (2-tailed)		.965	.737	.005
	N	30	30	30	30
X2	Pearson Correlation	-.008	1	.862**	.457*
	Sig. (2-tailed)	.965		.000	.011
	N	30	30	30	30
X3	Pearson Correlation	-.064	.862**	1	.478**
	Sig. (2-tailed)	.737	.000		.008
	N	30	30	30	30
X4	Pearson Correlation	.496**	.457*	.478**	1
	Sig. (2-tailed)	.005	.011	.008	
	N	30	30	30	30
X5	Pearson Correlation	.683**	.008	-.035	.300
	Sig. (2-tailed)	.000	.966	.853	.108
	N	30	30	30	30

** . Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Lampiran 2. (Lanjutan)

Data 26

Correlations						
	X1	X2	X3	X4	X5	X6
X1 Pearson Correlation	1	.602**	.455**	.454**	.568**	.336
Sig. (2-tailed)		.000	.009	.009	.001	.060
N	32	32	32	32	32	32
X2 Pearson Correlation	.602**	1	.818**	.766**	.624**	.725**
Sig. (2-tailed)	.000		.000	.000	.000	.000
N	32	32	32	32	32	32
X3 Pearson Correlation	.455**	.818**	1	.682**	.494**	.618**
Sig. (2-tailed)	.009	.000		.000	.004	.000
N	32	32	32	32	32	32
X4 Pearson Correlation	.454**	.766**	.682**	1	.618**	.738**
Sig. (2-tailed)	.009	.000	.000		.000	.000
N	32	32	32	32	32	32
X5 Pearson Correlation	.568**	.624**	.494**	.618**	1	.281
Sig. (2-tailed)	.001	.000	.004	.000		.119
N	32	32	32	32	32	32
X6 Pearson Correlation	.336	.725**	.618**	.738**	.281	1
Sig. (2-tailed)	.060	.000	.000	.000	.119	
N	32	32	32	32	32	32

** . Correlation is significant at the 0.01 level (2-tailed).

Data 27

Correlations						
	X1	X2	X3	X4	X5	X6
X1 Pearson Correlation	1	-.101	.333	-.480**	-.591**	.298
Sig. (2-tailed)		.577	.058	.005	.000	.092
N	33	33	33	33	33	33
X2 Pearson Correlation	-.101	1	.490**	-.127	-.051	.292
Sig. (2-tailed)	.577		.004	.482	.779	.099
N	33	33	33	33	33	33
X3 Pearson Correlation	.333	.490**	1	-.429*	-.180	.024
Sig. (2-tailed)	.058	.004		.013	.316	.896
N	33	33	33	33	33	33
X4 Pearson Correlation	-.480**	-.127	-.429*	1	.710**	.009
Sig. (2-tailed)	.005	.482	.013		.000	.959
N	33	33	33	33	33	33
X5 Pearson Correlation	-.591**	-.051	-.180	.710**	1	-.054
Sig. (2-tailed)	.000	.779	.316	.000		.764
N	33	33	33	33	33	33
X6 Pearson Correlation	.298	.292	.024	.009	-.054	1
Sig. (2-tailed)	.092	.099	.896	.959	.764	
N	33	33	33	33	33	33

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Lampiran 2. (Lanjutan)

Data 28

Correlations

		X1	X2	X3	X4
X1	Pearson Correlation	1	.228	.239	.352*
	Sig. (2-tailed)		.202	.181	.045
	N	33	33	33	33
X2	Pearson Correlation	.228	1	.926**	.831**
	Sig. (2-tailed)	.202		.000	.000
	N	33	33	33	33
X3	Pearson Correlation	.239	.926**	1	.888**
	Sig. (2-tailed)	.181	.000		.000
	N	33	33	33	33
X4	Pearson Correlation	.352*	.831**	.888**	1
	Sig. (2-tailed)	.045	.000	.000	
	N	33	33	33	33

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Data 29

Correlations

		X1	X2	X3	X4	X5
X1	Pearson Correlation	1	.875**	.613**	.278	.245
	Sig. (2-tailed)		.000	.000	.106	.156
	N	35	35	35	35	35
X2	Pearson Correlation	.875**	1	.837**	.502**	.526**
	Sig. (2-tailed)	.000		.000	.002	.001
	N	35	35	35	35	35
X3	Pearson Correlation	.613**	.837**	1	.811**	.824**
	Sig. (2-tailed)	.000	.000		.000	.000
	N	35	35	35	35	35
X4	Pearson Correlation	.278	.502**	.811**	1	.877**
	Sig. (2-tailed)	.106	.002	.000		.000
	N	35	35	35	35	35
X5	Pearson Correlation	.245	.526**	.824**	.877**	1
	Sig. (2-tailed)	.156	.001	.000	.000	
	N	35	35	35	35	35

**. Correlation is significant at the 0.01 level (2-tailed).

Data 30

Correlations

		X1	X2	X3
X1	Pearson Correlation	1	.523**	.524**
	Sig. (2-tailed)		.001	.001
	N	36	36	36
X2	Pearson Correlation	.523**	1	.466**
	Sig. (2-tailed)	.001		.004
	N	36	36	36
X3	Pearson Correlation	.524**	.466**	1
	Sig. (2-tailed)	.001	.004	
	N	36	36	36

**. Correlation is significant at the 0.01 level (2-tailed).

Lampiran 2. (Lanjutan)

Data 31

Correlations						
		X1	X2	X3	X4	X5
X1	Pearson Correlation	1	.417**	-.871**	-.786**	-.806**
	Sig. (2-tailed)		.009	.000	.000	.000
	N	38	38	38	38	38
X2	Pearson Correlation	.417**	1	-.589**	-.798**	-.692**
	Sig. (2-tailed)	.009		.000	.000	.000
	N	38	38	38	38	38
X3	Pearson Correlation	-.871**	-.589**	1	.872**	.864**
	Sig. (2-tailed)	.000	.000		.000	.000
	N	38	38	38	38	38
X4	Pearson Correlation	-.786**	-.798**	.872**	1	.940**
	Sig. (2-tailed)	.000	.000	.000		.000
	N	38	38	38	38	38
X5	Pearson Correlation	-.806**	-.692**	.864**	.940**	1
	Sig. (2-tailed)	.000	.000	.000	.000	
	N	38	38	38	38	38

** Correlation is significant at the 0.01 level (2-tailed).

Data 32

Correlations							
		X1	X2	X3	X4	X5	X6
X1	Pearson Correlation	1	-.010	-.083	-.240	.264	-.345*
	Sig. (2-tailed)		.954	.616	.141	.105	.031
	N	39	39	39	39	39	39
X2	Pearson Correlation	-.010	1	-.379*	-.402*	.786**	-.786**
	Sig. (2-tailed)	.954		.017	.011	.000	.000
	N	39	39	39	39	39	39
X3	Pearson Correlation	-.083	-.379*	1	.228	-.324*	.438**
	Sig. (2-tailed)	.616	.017		.163	.044	.005
	N	39	39	39	39	39	39
X4	Pearson Correlation	-.240	-.402*	.228	1	-.530**	.493**
	Sig. (2-tailed)	.141	.011	.163		.001	.001
	N	39	39	39	39	39	39
X5	Pearson Correlation	.264	.786**	-.324*	-.530**	1	-.931**
	Sig. (2-tailed)	.105	.000	.044	.001		.000
	N	39	39	39	39	39	39
X6	Pearson Correlation	-.345*	-.786**	.438**	.493**	-.931**	1
	Sig. (2-tailed)	.031	.000	.005	.001	.000	
	N	39	39	39	39	39	39

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

Data 33 (Tidak ada korelasi)

Data 34

Correlations					
		X1	X2	X3	X4
X1	Pearson Correlation	1	-.166	-.624**	.213
	Sig. (2-tailed)		.306	.000	.186
	N	40	40	40	40
X2	Pearson Correlation	-.166	1	.206	.322*
	Sig. (2-tailed)	.306		.201	.042
	N	40	40	40	40
X3	Pearson Correlation	-.624**	.206	1	-.111
	Sig. (2-tailed)	.000	.201		.495
	N	40	40	40	40
X4	Pearson Correlation	.213	.322*	-.111	1
	Sig. (2-tailed)	.186	.042	.495	
	N	40	40	40	40

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Lampiran 2. (Lanjutan)

Data 35

Correlations							
		X1	X2	X3	X4	X5	X6
X1	Pearson Correlation	1	-.434**	.645**	.494**	.095	.054
	Sig. (2-tailed)		.005	.000	.001	.556	.736
	N	41	41	41	41	41	41
X2	Pearson Correlation	-.434**	1	-.190	-.063	-.350*	.386*
	Sig. (2-tailed)	.005		.234	.697	.025	.013
	N	41	41	41	41	41	41
X3	Pearson Correlation	.645**	-.190	1	.955**	.238	-.032
	Sig. (2-tailed)	.000	.234		.000	.134	.841
	N	41	41	41	41	41	41
X4	Pearson Correlation	.494**	-.063	.955**	1	.213	-.026
	Sig. (2-tailed)	.001	.697	.000		.182	.871
	N	41	41	41	41	41	41
X5	Pearson Correlation	.095	-.350*	.238	.213	1	-.013
	Sig. (2-tailed)	.556	.025	.134	.182		.936
	N	41	41	41	41	41	41
X6	Pearson Correlation	.054	.386*	-.032	-.026	-.013	1
	Sig. (2-tailed)	.736	.013	.841	.871	.936	
	N	41	41	41	41	41	41

** . Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Data 36

Correlations					
		X1	X2	X3	X4
X1	Pearson Correlation	1	.659**	.359*	.162
	Sig. (2-tailed)		.000	.017	.293
	N	44	44	44	44
X2	Pearson Correlation	.659**	1	.283	-.152
	Sig. (2-tailed)	.000		.063	.326
	N	44	44	44	44
X3	Pearson Correlation	.359*	.283	1	.189
	Sig. (2-tailed)	.017	.063		.220
	N	44	44	44	44
X4	Pearson Correlation	.162	-.152	.189	1
	Sig. (2-tailed)	.293	.326	.220	
	N	44	44	44	44

** . Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Data 37

Correlations						
		X1	X2	X3	X4	X5
X1	Pearson Correlation	1	-.573**	.338*	.014	.376*
	Sig. (2-tailed)		.000	.022	.925	.010
	N	46	46	46	46	46
X2	Pearson Correlation	-.573**	1	-.184	.092	-.133
	Sig. (2-tailed)	.000		.221	.545	.380
	N	46	46	46	46	46
X3	Pearson Correlation	.338*	-.184	1	.134	.480**
	Sig. (2-tailed)	.022	.221		.376	.001
	N	46	46	46	46	46
X4	Pearson Correlation	.014	.092	.134	1	.315*
	Sig. (2-tailed)	.925	.545	.376		.033
	N	46	46	46	46	46
X5	Pearson Correlation	.376*	-.133	.480**	.315*	1
	Sig. (2-tailed)	.010	.380	.001	.033	
	N	46	46	46	46	46

** . Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Lampiran 2. (Lanjutan)

Data 38

Correlations					
		X1	X2	X3	X4
X1	Pearson Correlation	1	.404**	-.626**	-.041
	Sig. (2-tailed)		.004	.000	.782
	N	48	48	48	48
X2	Pearson Correlation	.404**	1	-.204	-.413**
	Sig. (2-tailed)	.004	.163	.004	.004
	N	48	48	48	48
X3	Pearson Correlation	-.626**	-.204	1	-.525**
	Sig. (2-tailed)	.000	.163	.000	.000
	N	48	48	48	48
X4	Pearson Correlation	-.041	-.413**	-.525**	1
	Sig. (2-tailed)	.782	.004	.000	
	N	48	48	48	48

** . Correlation is significant at the 0.01 level (2-tailed).

Data 39

Correlations				
		X1	X2	X3
X1	Pearson Correlation	1	.875**	.834**
	Sig. (2-tailed)		.000	.000
	N	49	49	49
X2	Pearson Correlation	.875**	1	.926**
	Sig. (2-tailed)	.000		.000
	N	49	49	49
X3	Pearson Correlation	.834**	.926**	1
	Sig. (2-tailed)	.000	.000	
	N	49	49	49

** . Correlation is significant at the 0.01 level (2-tailed).

Data 40

Correlations							
		X1	X2	X3	X4	X5	X6
X1	Pearson Correlation	1	.627**	.485**	.312*	.259	.296*
	Sig. (2-tailed)		.000	.000	.027	.069	.037
	N	50	50	50	50	50	50
X2	Pearson Correlation	.627**	1	.769**	.247	.349*	.327*
	Sig. (2-tailed)	.000		.000	.084	.013	.020
	N	50	50	50	50	50	50
X3	Pearson Correlation	.485**	.769**	1	.291*	.356*	.301*
	Sig. (2-tailed)	.000	.000		.040	.011	.034
	N	50	50	50	50	50	50
X4	Pearson Correlation	.312*	.247	.291*	1	.642**	.566**
	Sig. (2-tailed)	.027	.084	.040		.000	.000
	N	50	50	50	50	50	50
X5	Pearson Correlation	.259	.349*	.356*	.642**	1	.703**
	Sig. (2-tailed)	.069	.013	.011	.000		.000
	N	50	50	50	50	50	50
X6	Pearson Correlation	.296*	.327*	.301*	.566**	.703**	1
	Sig. (2-tailed)	.037	.020	.034	.000	.000	
	N	50	50	50	50	50	50

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Lampiran 2. (Lanjutan)

Data 41

Correlations					
		X1	X2	X3	X4
X1	Pearson Correlation	1	.186	.358*	.514**
	Sig. (2-tailed)		.196	.011	.000
	N	50	50	50	50
X2	Pearson Correlation	.186	1	.153	-.042
	Sig. (2-tailed)	.196		.290	.771
	N	50	50	50	50
X3	Pearson Correlation	.358*	.153	1	.556**
	Sig. (2-tailed)	.011	.290		.000
	N	50	50	50	50
X4	Pearson Correlation	.514**	-.042	.556**	1
	Sig. (2-tailed)	.000	.771	.000	
	N	50	50	50	50

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Data 42 (Tidak ada korelasi)

Data 43

Correlations				
		X1	X2	X3
X1	Pearson Correlation	1	.056	-.025
	Sig. (2-tailed)		.698	.859
	N	51	51	51
X2	Pearson Correlation	.056	1	.327*
	Sig. (2-tailed)	.698		.019
	N	51	51	51
X3	Pearson Correlation	-.025	.327*	1
	Sig. (2-tailed)	.859	.019	
	N	51	51	51

*. Correlation is significant at the 0.05 level (2-tailed).

Data 44

Correlations			
		X1	X2
X1	Pearson Correlation	1	.827**
	Sig. (2-tailed)		.000
	N	51	51
X2	Pearson Correlation	.827**	1
	Sig. (2-tailed)	.000	
	N	51	51

**. Correlation is significant at the 0.01 level

Data 45

Correlations			
		X1	X2
X1	Pearson Correlation	1	.835**
	Sig. (2-tailed)		.000
	N	51	51
X2	Pearson Correlation	.835**	1
	Sig. (2-tailed)	.000	
	N	51	51

**. Correlation is significant at the 0.01 level

Lampiran 2. (Lanjutan)

Data 46

Correlations

		X1	X2	X3	X4	X5
X1	Pearson Correlation	1	.719**	.833**	.478**	-.594**
	Sig. (2-tailed)		.000	.000	.000	.000
	N	53	53	53	53	53
X2	Pearson Correlation	.719**	1	.577**	.608**	-.575**
	Sig. (2-tailed)	.000		.000	.000	.000
	N	53	53	53	53	53
X3	Pearson Correlation	.833**	.577**	1	.410**	-.401**
	Sig. (2-tailed)	.000	.000		.002	.003
	N	53	53	53	53	53
X4	Pearson Correlation	.478**	.608**	.410**	1	-.491**
	Sig. (2-tailed)	.000	.000	.002		.000
	N	53	53	53	53	53
X5	Pearson Correlation	-.594**	-.575**	-.401**	-.491**	1
	Sig. (2-tailed)	.000	.000	.003	.000	
	N	53	53	53	53	53

**. Correlation is significant at the 0.01 level (2-tailed).

Data 47

Correlations

		X1	X2	X3	X4	X5	X6
X1	Pearson Correlation	1	-.511**	-.334*	-.203	-.182	-.285*
	Sig. (2-tailed)		.000	.010	.127	.171	.030
	N	58	58	58	58	58	58
X2	Pearson Correlation	-.511**	1	.243	.017	-.138	.425**
	Sig. (2-tailed)	.000		.066	.901	.300	.001
	N	58	58	58	58	58	58
X3	Pearson Correlation	-.334*	.243	1	.768**	.581**	.355**
	Sig. (2-tailed)	.010	.066		.000	.000	.006
	N	58	58	58	58	58	58
X4	Pearson Correlation	-.203	.017	.768**	1	.588**	.030
	Sig. (2-tailed)	.127	.901	.000		.000	.821
	N	58	58	58	58	58	58
X5	Pearson Correlation	-.182	-.138	.581**	.588**	1	.035
	Sig. (2-tailed)	.171	.300	.000	.000		.793
	N	58	58	58	58	58	58
X6	Pearson Correlation	-.285*	.425**	.355**	.030	.035	1
	Sig. (2-tailed)	.030	.001	.006	.821	.793	
	N	58	58	58	58	58	58

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Lampiran 2. (Lanjutan)

Data 48

Correlations									
		X1	X2	X3	X4	X5	X6	X7	X8
X1	Pearson Correlation	1	.023	.511**	.310**	-.313**	.264*	.551**	-.106
	Sig. (2-tailed)		.851	.000	.010	.009	.029	.000	.388
	N	69	69	69	69	69	69	69	69
X2	Pearson Correlation	.023	1	.221	-.059	.482**	-.141	-.328**	.539**
	Sig. (2-tailed)	.851		.067	.629	.000	.248	.006	.000
	N	69	69	69	69	69	69	69	69
X3	Pearson Correlation	.511**	.221	1	-.012	.024	-.326**	.283*	.223
	Sig. (2-tailed)	.000	.067		.922	.847	.006	.019	.065
	N	69	69	69	69	69	69	69	69
X4	Pearson Correlation	.310**	-.059	-.012	1	-.069	.332**	.137	-.033
	Sig. (2-tailed)	.010	.629	.922		.574	.005	.261	.790
	N	69	69	69	69	69	69	69	69
X5	Pearson Correlation	-.313**	.482**	.024	-.069	1	-.364**	-.149	.912**
	Sig. (2-tailed)	.009	.000	.847	.574		.002	.222	.000
	N	69	69	69	69	69	69	69	69
X6	Pearson Correlation	.264*	-.141	-.326**	.332**	-.364**	1	-.316**	-.377**
	Sig. (2-tailed)	.029	.248	.006	.005	.002		.008	.001
	N	69	69	69	69	69	69	69	69
X7	Pearson Correlation	.551**	-.328**	.283*	.137	-.149	-.316**	1	.005
	Sig. (2-tailed)	.000	.006	.019	.261	.222	.008		.964
	N	69	69	69	69	69	69	69	69
X8	Pearson Correlation	-.106	.539**	.223	-.033	.912**	-.377**	.005	1
	Sig. (2-tailed)	.388	.000	.065	.790	.000	.001	.964	
	N	69	69	69	69	69	69	69	69

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Data 49

		Correlations		
		X1	X2	X3
X1	Pearson Correlation	1	1.000**	.941**
	Sig. (2-tailed)		.000	.000
	N	74	74	74
X2	Pearson Correlation	1.000**	1	.942**
	Sig. (2-tailed)	.000		.000
	N	74	74	74
X3	Pearson Correlation	.941**	.942**	1
	Sig. (2-tailed)	.000	.000	
	N	74	74	74

** . Correlation is significant at the 0.01 level (2-tailed).

Data 50

		Correlations					
		X1	X2	X3	X4	X5	X6
X1	Pearson Correlation	1	.746**	.682**	.602**	.641**	.594**
	Sig. (2-tailed)		.000	.000	.000	.000	.000
	N	79	79	79	79	79	79
X2	Pearson Correlation	.746**	1	.879**	.814**	.855**	.924**
	Sig. (2-tailed)	.000		.000	.000	.000	.000
	N	79	79	79	79	79	79
X3	Pearson Correlation	.682**	.879**	1	.968**	.970**	.818**
	Sig. (2-tailed)	.000	.000		.000	.000	.000
	N	79	79	79	79	79	79
X4	Pearson Correlation	.602**	.814**	.968**	1	.989**	.762**
	Sig. (2-tailed)	.000	.000	.000		.000	.000
	N	79	79	79	79	79	79
X5	Pearson Correlation	.641**	.855**	.970**	.989**	1	.787**
	Sig. (2-tailed)	.000	.000	.000	.000		.000
	N	79	79	79	79	79	79
X6	Pearson Correlation	.594**	.924**	.818**	.762**	.787**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	
	N	79	79	79	79	79	79

** . Correlation is significant at the 0.01 level (2-tailed).

Lampiran 3. Hasil Analisis Komponen Utama Data 4

Principal Component Analysis: C1, C2, C3, C4, C5, C6, C7

Eigenanalysis of the Correlation Matrix

Eigenvalue	3.4622	1.3297	0.9582	0.6109	0.3608	0.1700
0.1082						
Proportion	0.495	0.190	0.137	0.087	0.052	0.024
0.015						
Cumulative	0.495	0.685	0.821	0.909	0.960	0.985
1.000						

Variable	PC1	PC2	PC3	PC4	PC5	PC6
PC7						
C1	-0.284	0.601	0.291	0.318	0.443	0.343
0.242						
C2	-0.434	0.058	0.281	-0.607	0.009	0.240
0.551						
C3	-0.387	0.193	-0.454	0.506	-0.390	0.071
0.439						
C4	-0.459	0.265	0.090	-0.236	-0.455	-0.459
0.487						
C5	-0.388	-0.404	0.294	0.336	0.391	-0.547
0.188						
C6	-0.346	-0.597	0.126	0.139	-0.222	0.557
0.360						
C7	-0.316	-0.092	-0.721	-0.296	0.492	0.004
0.207						

Lampiran 4. Nilai dan hasil pengelompokan indeks validitas (*Gap statistic*, Hartigan, Krazanowski-Lai, Silhouette dan Calinski-Harabasz) menggunakan *package* ClusterSim pada *software* R.2.7.2

1. Data 1

```
minimal nc where diffu>=0 is 2 for diffu= 0.3749  
1 1 1 1 1 1 1 1 1 1 2
```

```
"minimal nc for H<=10 equals 2 for H=  
4.44905762204298"  
"clustering for minimal nc where H<=10"  
1 1 1 1 1 1 1 1 1 1 2
```

```
"max KL for 5 clusters= 1.67323933457639"  
print("clustering for max KL")  
"clustering for max KL"  
1 2 2 2 2 2 2 3 3 4 5
```

```
"max S for 5 clusters= 0.537159748937072"  
print("clustering for max S")  
"clustering for max S"  
1 2 2 2 2 2 2 3 3 4 5
```

```
"max G1 for 5 clusters= 9.7632896508224"  
print("clustering for max G1")  
"clustering for max G1"  
1 2 2 2 2 2 2 3 3 4 5
```

2. Data 2

```
minimal nc where diffu>=0 is 4 for diffu= 0.3184  
1 2 2 2 2 2 2 3 4 4 4 4
```

```
"minimal nc for H<=10 equals 2 for H=  
6.38846900338157"  
"clustering for minimal nc where H<=10"  
1 1 1 1 1 1 1 1 2 2 2 2
```

```
"max KL for 2 clusters= 37.0177342876834"  
print("clustering for max KL")
```

Lampiran 4. (Lanjutan)

```
"clustering for max KL"
```

```
1 1 1 1 1 1 1 1 2 2 2 2
```

```
"max S for 2 clusters= 0.826075798485365"
```

```
print("clustering for max S")
```

```
"clustering for max S"
```

```
1 1 1 1 1 1 1 1 2 2 2 2
```

```
"max G1 for 2 clusters= 132.916650204149"
```

```
print("clustering for max G1")
```

```
"clustering for max G1"
```

```
[1] 1 1 1 1 1 1 1 1 2 2 2 2
```

```
\
```

3. Data 3

```
minimal nc where diffu>=0 is 2 for diffu= 0.4057
```

```
1 1 2 2 2 1 1 1 1 1 1 1 1
```

```
"minimal nc for H<=10 equals 3 for H=
```

```
6.47571052908557"
```

```
"clustering for minimal nc where H<=10"
```

```
1 1 2 3 3 1 1 1 1 1 1 1 1
```

```
"max KL for 2 clusters= 6.80497974021983"
```

```
print("clustering for max KL")
```

```
"clustering for max KL"
```

```
1 1 2 2 2 1 1 1 1 1 1 1 1
```

```
"max S for 4 clusters= 0.829381399971163"
```

```
print("clustering for max S")
```

```
"clustering for max S"
```

```
1 1 2 3 4 1 1 1 1 1 1 1 1
```

```
"max G1 for 6 clusters= 206.930818001661"
```

```
print("clustering for max G1")
```

```
"clustering for max G1"
```

```
1 1 2 3 4 5 6 6 6 6 6 6 6
```

Lampiran 4. (Lanjutan)

4. Data 4

```
minimal nc where diffu>=0 is 2 for diffu= 0.2669
1 2 1 1 1 1 1 1 1 1 1 1 1 1 1

"minimal nc for H<=10 equals 2 for H=
2.61795156729177"
"clustering for minimal nc where H<=10"
1 2 1 1 1 1 1 1 1 1 1 1 1 1 1

"max KL for 5 clusters= 8.17970118773476"
print("clustering for max KL")
"clustering for max KL"
1 2 3 1 1 1 4 4 1 1 5 1 1 4 4 1

"max S for 6 clusters= 0.291656376194138"
print("clustering for max S")
"clustering for max S"
1 2 3 1 1 1 4 5 1 1 6 1 1 5 4 1

"max G1 for 5 clusters= 4.17244813086001"
print("clustering for max G1")
"clustering for max G1"
1 2 3 1 1 1 4 4 1 1 5 1 1 4 4 1
```

5. Data 5

```
minimal nc where diffu>=0 is 2 for diffu= 0.2615
1 1 1 2 1 1 1 1 1 1 1 1 1 1 1

"minimal nc for H<=10 equals 2 for H=
4.88027550119303"
"clustering for minimal nc where H<=10"
1 1 1 2 1 1 1 1 1 1 1 1 1 1 1

"max KL for 5 clusters= 8.9280369928921"
print("clustering for max KL")
"clustering for max KL"
1 2 2 3 4 5 2 2 2 2 2 2 2 2 2 2
```

Lampiran 4. (Lanjutan)

```
"max S for 2 clusters= 0.494037359097497"  
print("clustering for max S")  
"clustering for max S"  
1 1 1 2 1 1 1 1 1 1 1 1 1 1 1  
  
"max G1 for 3 clusters= 5.34992651351256"  
print("clustering for max G1")  
"clustering for max G1"  
1 2 2 3 2 2 2 2 2 2 2 2 2 2 2
```

6. Data 6

```
minimal nc where diffu>=0 is 2 for diffu= 0.2954  
1 1 1 1 1 1 1 1 1 2 1 1 1 1 1  
  
"minimal nc for H<=10 equals 2 for H=  
5.70151262315194"  
"clustering for minimal nc where H<=10"  
1 1 1 1 1 1 1 1 1 2 1 1 1 1 1  
  
"max KL for 3 clusters= 2.16362250204781"  
print("clustering for max KL")  
"clustering for max KL"  
1 1 1 1 1 1 2 1 1 3 1 2 1 1 1  
  
"max S for 2 clusters= 0.583227520317909"  
print("clustering for max S")  
"clustering for max S"  
1 1 1 1 1 1 1 1 1 2 1 1 1 1 1  
  
"max G1 for 2 clusters= 7.89036424831558"  
print("clustering for max G1")  
"clustering for max G1"  
1 1 1 1 1 1 1 1 1 2 1 1 1 1 1
```

Lampiran 4. (Lanjutan)

7. Data 7

```
minimal nc where diffu>=0 is 2 for diffu= 0.8096
1 1 1 1 1 1 1 1 1 2 1 2 1 1 1 1

"minimal nc for H<=10 equals 2 for H=
2.79585881187831"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 2 1 2 1 1 1 1

"max KL for 6 clusters= 12.2073436955627"
print("clustering for max KL")
"clustering for max KL"
1 1 2 1 1 1 3 1 3 4 3 5 6 3 1 1

"max S for 6 clusters= 0.642560019082863"
print("clustering for max S")
"clustering for max S"
1 1 2 1 1 1 3 1 3 4 3 5 6 3 1 1

"max G1 for 6 clusters= 21.7729209778362"
print("clustering for max G1")
"clustering for max G1"
1 1 2 1 1 1 3 1 3 4 3 5 6 3 1 1
```

8. Data 8

```
minimal nc where diffu>=0 is 2 for diffu= 0.0724
1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1

"minimal nc for H<=10 equals 2 for H=
6.19320292184862"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1

"max KL for 4 clusters= 11.7100118898749"
> print("clustering for max KL")
"clustering for max KL"
```

Lampiran 4. (Lanjutan)

```
1 1 1 1 1 1 2 1 3 1 4 1 1 1 1
"max S for 3 clusters= 0.497601270675284"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 2 1 1 1 3 1 1 1 1
"max G1 for 3 clusters= 5.86476039627609"
print("clustering for max G1")
"clustering for max G1"
1 1 1 1 1 1 2 1 1 1 3 1 1 1 1
```

9. Data 9

```
minimal nc where diffu>=0 is 2 for diffu= 1.1961
1 2 1 1 1 2 2 2 2 2 1 1 2 2 2
"minimal nc for H<=10 equals 3 for H=
8.85016094789255"
"clustering for minimal nc where H<=10"
1 2 1 1 1 3 3 3 3 2 1 1 3 3 3 2
"max KL for 5 clusters= 13.3618331374577"
print("clustering for max KL")
"clustering for max KL"
1 2 1 3 1 4 4 5 4 2 3 3 4 5 5 2
"max S for 6 clusters= 0.659093953749091"
print("clustering for max S")
"clustering for max S"
1 2 1 3 1 4 4 5 4 2 3 3 4 5 6 2
"max G1 for 5 clusters= 127.194751056249"
print("clustering for max G1")
"clustering for max G1"
1 2 1 3 1 4 4 5 4 2 3 3 4 5 5 2
```

Lampiran 4. (Lanjutan)

10. Data 10

```
minimal nc where diffu>=0 is 3 for diffu= 0.4113
1 1 1 1 1 1 2 3 1 1 1 1 1 1 1 2

"minimal nc for H<=10 equals 3 for H=
1.7384847726728"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 2 3 1 1 1 1 1 1 1 2

"max KL for 3 clusters= 65.0325238259364"
print("clustering for max KL")
"clustering for max KL"
1 1 1 1 1 1 2 3 1 1 1 1 1 1 1 2

"max S for 2 clusters= 0.638276260825718"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1

"max G1 for 7 clusters= 18.6090433794093"
print("clustering for max G1")
"clustering for max G1"
1 1 2 1 3 3 4 5 3 3 6 3 3 3 1 3 7
```

11. Data 11

```
minimal nc where diffu>=0 is 2 for diffu= 0.147
1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1

"minimal nc for H<=10 equals 2 for H=
7.84884756056417"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1

"max KL for 3 clusters= 2.18575231461896"
print("clustering for max KL")
"clustering for max KL"
1 1 1 1 2 1 1 1 1 1 3 1 1 2 1 1 1 1

"max S for 2 clusters= 0.440624100868396"
print("clustering for max S")
```

Lampiran 4. (Lanjutan)

```
"clustering for max S"
1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1
"max G1 for 3 clusters= 6.56871750230275"
print("clustering for max G1")
"clustering for max G1"
1 1 1 1 2 1 1 1 1 1 3 1 1 2 1 1 1 1
```

12. Data 12

```
minimal nc where diffu>=0 is 2 for diffu= 1.2202
1 2 1 2 2 1 1 1 1 2 1 1 1 1 1 1 1
"minimal nc for H<=10 equals 2 for H=
3.22555737980861"
"clustering for minimal nc where H<=10"
1 2 1 2 2 1 1 1 1 2 1 1 1 1 1 1 1
"max KL for 4 clusters= 70.8928448901966"
print("clustering for max KL")
"clustering for max KL"
1 2 1 2 3 4 1 4 4 3 4 4 4 4 4 4 4
"max S for 2 clusters= 0.806191944889086"
print("clustering for max S")
"clustering for max S"
1 2 1 2 2 1 1 1 1 2 1 1 1 1 1 1 1
"max G1 for 6 clusters= 409.445523720011"
print("clustering for max G1")
"clustering for max G1"
1 2 1 2 3 4 5 4 4 3 4 4 4 4 4 6 4 4
```

13. Data 13

```
minimal nc where diffu>=0 is 2 for diffu= 0.3657
1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
"minimal nc for H<=10 equals 3 for H=
9.5282764948482"
```

Lampiran 4. (Lanjutan)

```
"clustering for minimal nc where H<=10"
1 2 2 2 2 2 3 2 2 2 2 3 2 2 3 2 2 2

"max KL for 3 clusters= 6.38315964890405"
print("clustering for max KL")
"clustering for max KL"
1 2 2 2 2 2 3 2 2 2 2 3 2 2 3 2 2 2

"max S for 2 clusters= 0.72250325897663"
print("clustering for max S")
"clustering for max S"
1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

"max G1 for 5 clusters= 49.966410207287"
print("clustering for max G1")
"clustering for max G1"
1 2 2 3 2 2 4 2 2 5 2 4 2 2 4 2 2 2
```

14. Data 14

```
minimal nc where diffu>=0 is 2 for diffu= 0.1277
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2

"minimal nc for H<=10 equals 4 for H=
0.95357952704687"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 2 3 3 4

"max KL for 4 clusters= 4.94962655413827"
print("clustering for max KL")
"clustering for max KL"
1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 2 3 3 4

"max S for 2 clusters= 0.687648376358214"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2

"max G1 for 4 clusters= 28.3438064046438"
print("clustering for max G1")
```

Lampiran 4. (Lanjutan)

```
"clustering for max G1"
```

```
1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 2 3 3 4
```

15. Data 15

```
minimal nc where diffu>=0 is 2 for diffu= 0.211
```

```
1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

```
"minimal nc for H<=10 equals 3 for H=
```

```
0.838503630502"
```

```
"clustering for minimal nc where H<=10"
```

```
1 2 1 1 1 1 1 1 1 1 1 1 1 3 3 1 3 1 1 1
```

```
"max KL for 4 clusters= 4.19167441922413"
```

```
print("clustering for max KL")
```

```
"clustering for max KL"
```

```
1 2 1 1 1 1 1 1 1 1 1 1 1 3 3 1 4 1 1 1
```

```
"max S for 7 clusters= 0.622388730205058"
```

```
print("clustering for max S")
```

```
"clustering for max S"
```

```
1 2 1 1 1 1 1 1 1 3 3 4 1 5 5 1 6 3 4 7
```

```
"max G1 for 7 clusters= 20.0528291231203"
```

```
print("clustering for max G1")
```

```
"clustering for max G1"
```

```
1 2 1 1 1 1 1 1 1 3 3 4 1 5 5 1 6 3 4 7
```

16. Data 16

```
minimal nc where diffu>=0 is 2 for diffu= 0.263
```

```
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1
```

```
"minimal nc for H<=10 equals 2 for H=
```

```
4.88799123373444"
```

```
"clustering for minimal nc where H<=10"
```

```
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1
```

```
"max KL for 2 clusters= 8.00350973666977"
```

```
print("clustering for max KL")
```

Lampiran 4. (Lanjutan)

```
"clustering for max KL"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1

"max S for 7 clusters= 0.455904028169805"
print("clustering for max S")
"clustering for max S"
1 1 1 1 2 3 1 4 1 4 5 1 1 1 1 6 7 1 1 1

"max G1 for 7 clusters= 7.73447326974316"
print("clustering for max G1")
"clustering for max G1"
1 1 1 1 2 3 1 4 1 4 5 1 1 1 1 6 7 1 1 1
```

17. Data 17

```
minimal nc where diffu>=0 is 2 for diffu= 0.0109
1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1

"minimal nc for H<=10 equals 3 for H=
2.17196517557042"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 2 1 1 3 1 1 2 1 1 1 1 1 1 1

"max KL for 8 clusters= 8.76984156925483"
print("clustering for max KL")
"clustering for max KL"
1 2 3 3 4 5 6 1 3 7 1 3 8 3 1 1 3 3 3 5 3

"max S for 7 clusters= 0.659560370623321"
print("clustering for max S")
"clustering for max S"
1 2 1 1 3 4 5 1 1 6 1 1 7 1 1 1 1 1 1 4 1

"max G1 for 8 clusters= 50.6678507047483"
print("clustering for max G1")
"clustering for max G1"
1 2 3 3 4 5 6 1 3 7 1 3 8 3 1 1 3 3 3 5 3
```

Lampiran 4. (Lanjutan)

18. Data 18

```
minimal nc where diffu>=0 is 2 for diffu= 0.3659
1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"minimal nc for H<=10 equals 2 for H=
1.27230707449057"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"max KL for 4 clusters= 160.836740902810"
print("clustering for max KL")
"clustering for max KL"
1 1 2 3 2 2 4 1 3 2 2 2 2 2 2 2 1 1 2 2 1 1 1 2

"max S for 6 clusters= 0.662272505318657"
print("clustering for max S")
"clustering for max S"
1 1 2 3 2 2 4 5 3 2 2 2 2 2 2 2 2 1 6 2 2 1 1 1 2
```

19. Data 19

```
minimal nc where diffu>=0 is 2 for diffu= 0.3244
1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"minimal nc for H<=10 equals 2 for H=
2.37901584464236"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"max KL for 2 clusters= 18.5708409244503"
print("clustering for max KL")
"clustering for max KL"
1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"max S for 10 clusters= 0.414360538553394"
print("clustering for max S")
"clustering for max S"
1 2 2 3 4 2 5 6 2 7 8 2 7 2 2 4 9
3 10 2 2 2 7 2 3
```

Lampiran 4. (Lanjutan)

```
"max G1 for 6 clusters= 15.9356475812094"  
print("clustering for max G1")  
"clustering for max G1"  
1 2 2 3 4 2 2 5 2 4 4 2 4 2 4 6 3 6 2 2 2 4 2 3
```

20. Data 20

```
minimal nc where diffu>=0 is 2 for diffu= 0.1304  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1  
  
"minimal nc for H<=10 equals 2 for H=  
6.21129938448229"  
"clustering for minimal nc where H<=10"  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1  
  
"max KL for 4 clusters= 10.7915965435812"  
print("clustering for max KL")  
"clustering for max KL"  
1 1 1 1 2 3 2 1 1 2 1 1 1 1 2 1 4 1 1 1 1 1 1 1  
  
"max S for 2 clusters= 0.681537256913628"  
print("clustering for max S")  
"clustering for max S"  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1  
  
"max G1 for 6 clusters= 20.3538984334327"  
print("clustering for max G1")  
"clustering for max G1"  
1 1 1 1 2 3 2 1 1 2 1 1 4 4 1 2 1 5 6 6 1 6 1 1 1
```

21. Data 21

```
minimal nc where diffu>=0 is 2 for diffu= 0.1185  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1  
1  
  
"minimal nc for H<=10 equals 3 for H=  
0.638530740269873"  
"clustering for minimal nc where H<=10"
```

Lampiran 4. (Lanjutan)

```
1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 3 1 1 1 1 1 1 1
2

"max KL for 3 clusters= 11.3601347934079"
print("clustering for max KL")
"clustering for max KL"
1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 3 1 1 1 1 1 1 1
2

"max S for 2 clusters= 0.604631559058441"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1
1

"max G1 for 6 clusters= 17.4158177136508"
print("clustering for max G1")
"clustering for max G1"
1 1 1 1 2 1 1 1 1 1 1 3 1 2 4 1 1 5 2 2 1 2 2 2 2
6
```

22. Data 22

```
minimal nc where diffu>=0 is 2 for diffu= 0.1289
1 1 1 1 1 1 1 1 1 1 1 2 1 2 1 1 1 1 1 1 1 1 1 1 1
1 1

"minimal nc for H<=10 equals 2 for H=
4.15239713718134"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 2 1 2 1 1 1 1 1 1 1 1 1 1 1
1 1

"max KL for 6 clusters= 22.9458756089252"
print("clustering for max KL")
"clustering for max KL"
1 2 2 2 3 2 2 2 2 4 5 1 5 1 1 1 2 2 3 1 6 2 2 2 2
2 2

"max S for 8 clusters= 0.467157838486911"
```

Lampiran 4. (Lanjutan)

```
print("clustering for max S")
"clustering for max S"
1 2 2 2 3 2 4 2 2 5 6 1 6 1 1 1 2 2 7 1 8 2 2 2 2
2 2

"max G1 for 6 clusters= 12.1178558646012"
print("clustering for max G1")
"clustering for max G1"
1 2 2 2 3 2 2 2 2 4 5 1 5 1 1 1 2 2 3 1 6 2 2 2 2
2 2
```

23. Data 23

```
minimal nc where diffu>=0 is 3 for diffu= 0.2902
1 2 1 1 1 1 1 1 3 3 1 1 1 1 1 3 3 1 1 1 1 1 1 1
1 1 1

"minimal nc for H<=10 equals 3 for H=
4.41713294110264"
"clustering for minimal nc where H<=10"
1 2 1 1 1 1 1 1 3 3 1 1 1 1 1 3 3 1 1 1 1 1 1 1
1 1 1

"max KL for 3 clusters= 11.2514743296161"
print("clustering for max KL")
"clustering for max KL"
1 2 1 1 1 1 1 1 3 3 1 1 1 1 1 3 3 1 1 1 1 1 1 1
1 1 1

"max S for 3 clusters= 0.544352137192304"
print("clustering for max S")
"clustering for max S"
1 2 1 1 1 1 1 1 3 3 1 1 1 1 1 3 3 1 1 1 1 1 1 1
1 1 1

"max G1 for 3 clusters= 18.6213452465702"
print("clustering for max G1")
"clustering for max G1"
1 2 1 1 1 1 1 1 3 3 1 1 1 1 1 3 3 1 1 1 1 1 1 1
1 1 1
```

Lampiran 4. (Lanjutan)

24. Data 24

```
minimal nc where diffu>=0 is 2 for diffu= 0.6387
1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1
1 1 1

"minimal nc for H<=10 equals 2 for H=
3.71354781567784"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1
1 1 1

"max KL for 8 clusters= 173.037153009678"
print("clustering for max KL")
"clustering for max KL"
1 2 2 3 2 2 4 2 4 2 2 5 2 4 6 7 8 4 4 4 4 4 2 2
2 2 2

"max S for 2 clusters= 0.613422204324115"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1
1 1 1

"max G1 for 10 clusters= 13.4368503697465"
print("clustering for max G1")
"clustering for max G1"
1 2 2 3 2 2 4 2 4 5 2 6 2 4 7 8 9
4 4 4 4 4 4 2 2 10 2 2
```

25. Data 25

```
minimal nc where diffu>=0 is 2 for diffu= 0.1166
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 2 1 1

"minimal nc for H<=10 equals 2 for H=
2.65628324053302"
"clustering for minimal nc where H<=10"
```

Lampiran 4. (Lanjutan)

```

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 2 1 1
"max KL for 9 clusters= 32.5912888016813"
print("clustering for max KL")
"clustering for max KL"
1 1 2 3 2 2 4 2 2 2 2 2 4 2 2 2 4 2 2 5 6 2 6 4 7
6 6 8 6 9
"max S for 10 clusters= 0.422648602024157"
print("clustering for max S")
"clustering for max S"
1 2 3 2 2 4 2 2 2 2 2 4 2 2 2 5 2
2 6 7 2 7 4 8 7 7 9 7 10
"max G1 for 9 clusters= 10.3056811496416"
print("clustering for max G1")
"clustering for max G1"
1 1 2 3 2 2 4 2 2 2 2 2 4 2 2 2 5 6 2 6 4 7
6 6 8 6 9

```

26. Data 26

```
minimal nc where diffu>=0 is 2 for diffu= 0.4357
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 2

"minimal nc for H<=10 equals 2 for H=
4.85564921330404"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 2

"max KL for 5 clusters= 40.1872669458123"
print("clustering for max KL")
"clustering for max KL"
1 1 1 1 1 2 1 1 1 1 1 2 1 1 1 3 2 1 1 1 1 2 4 1
1 2 1 1 2 1 5
```

Lampiran 4. (Lanjutan)

```
"max S for 2 clusters= 0.257687630929094"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 2

"max G1 for 5 clusters= 9.18685721378267"
print("clustering for max G1")
"clustering for max G1"
1 1 1 1 1 2 1 1 1 1 1 1 2 1 1 1 3 2 1 1 1 1 2 4 1
1 2 1 1 2 1 5
```

27. Data 27

```
minimal nc where diffu>=0 is 2 for diffu= 0.3123
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1

"minimal nc for H<=10 equals 2 for H=
1.66151072527676"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1

"max KL for 5 clusters= 14.7228334698149"
print("clustering for max KL")
"clustering for max KL"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 3 4 1 1 1 1 1 1
1 1 1 1 1 1 1 5

"max S for 3 clusters= 0.179732349293626"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 3 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1

"max G1 for 7 clusters= 2.39996432246904"
print("clustering for max G1")
"clustering for max G1"
```

Lampiran 4. (Lanjutan)

```
1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 3 4 5 1 1 1 1 1 1
1 1 1 1 1 1 6 7
```

28. Data 28

```
minimal nc where diffu>=0 is 2 for diffu= 0.1951
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 2 2 2

"minimal nc for H<=10 equals 2 for H=
1.31104247725541"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 2 2 2

"max KL for 3 clusters= 16.2588359451737"
print("clustering for max KL")
"clustering for max KL"
1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 3 3 3

"max S for 2 clusters= 0.518636526380975"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 2 2 2

"max G1 for 2 clusters= 21.0520283188227"
print("clustering for max G1")
"clustering for max G1"
print(clusters[which.max(res[,2]),])
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 2 2 2
```

29. Data 29

```
minimal nc where diffu>=0 is 2 for diffu= 0.1504
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1
```

Lampiran 4. (Lanjutan)

```
"minimal nc for H<=10 equals 3 for H=
2.31608814591110"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 2 1 2 1 1 1 1 1 3 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1

"max KL for 3 clusters= 6.08034099266247"
print("clustering for max KL")
"clustering for max KL"
1 1 1 1 1 1 2 1 2 1 1 1 1 1 3 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1

"max S for 2 clusters= 0.634902043637417"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1

"max G1 for 7 clusters= 17.4465033659946"
print("clustering for max G1")
"clustering for max G1"
1 1 1 1 1 2 3 1 4 2 1 1 1 1 5 1 1 1 1 1 1 1 1
1 1 1 1 6 6 7 6 6 6
```

30. Data 30

```
minimal nc where diffu>=0 is 2 for diffu= 0.3777
1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1

"minimal nc for H<=10 equals 2 for H=
2.69409447723849"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1

"max KL for 5 clusters= 7.67011461336868"
print("clustering for max KL")
"clustering for max KL"
```

Lampiran 4. (Lanjutan)

```
1 2 1 1 3 1 2 1 1 1 2 1 1 4 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1
```

```
"max S for 2 clusters= 0.496700745058341"
```

```
print("clustering for max S")
```

```
"clustering for max S"
```

```
1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1
```

```
"max G1 for 5 clusters= 7.20009180256058"
```

```
print("clustering for max G1")
```

```
"clustering for max G1"
```

```
1 2 1 1 3 1 4 1 1 1 2 1 1 5 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1
```

31. Data 31

```
minimal nc where diffu>=0 is 2 for diffu= 0.0543
```

```
1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 1 1 1 1 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2 2 2 2
```

```
"minimal nc for H<=10 equals 3 for H=
4.71232466110942"
```

```
"clustering for minimal nc where H<=10"
```

```
1 1 1 1 2 2 2 2 2 3 2 3 3 3 3 3 1 1 1 1 2 3 2 2 2
2 2 2 3 3 2 2 2 2 2 2 2 2
```

```
"max KL for 3 clusters= 64.4790983483082"
```

```
print("clustering for max KL")
```

```
"clustering for max KL"
```

```
1 1 1 1 2 2 2 2 2 3 2 3 3 3 3 3 1 1 1 1 2 3 2 2 2
2 2 2 3 3 2 2 2 2 2 2 2 2
```

```
"max S for 2 clusters= 0.510478900975877"
```

```
print("clustering for max S")
```

```
"clustering for max S"
```

```
1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 1 1 1 1 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2 2 2 2
```

```
"max G1 for 3 clusters= 53.2488818287062"
```

Lampiran 4. (Lanjutan)

```
print("clustering for max G1")
"clustering for max G1"
1 1 1 1 2 2 2 2 2 3 2 3 3 3 3 3 1 1 1 1 2 3 2 2 2
2 2 2 3 3 2 2 2 2 2 2 2 2
```

32. Data 32

```
minimal nc where diffu>=0 is 2 for diffu= 0.3019
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 2
```

```
"minimal nc for H<=10 equals 2 for H=
0.989167896820248"
```

```
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 2
```

```
"max KL for 8 clusters= 108.635093200396"
```

```
print("clustering for max KL")
"clustering for max KL"
1 2 1 2 1 3 1 1 1 2 3 1 4 1 1 1 5 5 5 5 5 1 5 5 5
1 5 1 1 1 1 6 1 7 1 1 2 3 8
```

```
"max S for 2 clusters= 0.425406477650864"
```

```
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 2
```

```
"max G1 for 8 clusters= 45.5966097438688"
```

```
print("clustering for max G1")
"clustering for max G1"
1 2 1 2 1 3 1 1 1 2 3 1 4 1 1 1 5 5 5 5 5 1 5 5 5
1 5 1 1 1 1 6 1 7 1 1 2 3 8
```

33. Data 33

```
minimal nc where diffu>=0 is 2 for diffu= 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 2 1 1
```

Lampiran 4. (Lanjutan)

```
"minimal nc for H<=10 equals 2 for H=
8.20541788399979"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 2 1 1

"max KL for 10 clusters= 5.92868907992734"
print("clustering for max KL")
"clustering for max KL"
2 3 4 5 6 2 7 2 2 2 2 6 2 2 2 8 2
9 2 2 2 3 6 6 2 2 2 2 2 2 2 2 2 2
2 10 2 2

"max S for 2 clusters= 0.52269973072453"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 2 1 1

"max G1 for 3 clusters= 7.7735145695555"
print("clustering for max G1")
"clustering for max G1"
1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1
1 1 1 1 1 1 1 1 1 1 1 3 1 1
```

34. Data 34

```
minimal nc where diffu>=0 is 2 for diffu= 0.1278
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"minimal nc for H<=10 equals 2 for H=
4.59542443615143"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"max KL for 5 clusters= 15.1676639271032"
print("clustering for max KL")
"clustering for max KL"
```

Lampiran 4. (Lanjutan)

```
1 1 1 2 1 1 1 1 1 1 1 1 1 3 4 1 1 1 1 1 1 5 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

```
"max S for 2 clusters= 0.406437695804803"
```

```
print("clustering for max S")
```

```
"clustering for max S"
```

```
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

```
"max G1 for 2 clusters= 7.69648572617424"
```

```
print("clustering for max G1")
```

```
"clustering for max G1"
```

```
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

35. Data 35

```
minimal nc where diffu>=0 is 2 for diffu= 0.1111
```

```
1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

```
"minimal nc for H<=10 equals 2 for H=
4.31371534104042"
```

```
"clustering for minimal nc where H<=10"
```

```
1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

```
"max KL for 3 clusters= 134.670704699423"
```

```
print("clustering for max KL")
```

```
"clustering for max KL"
```

```
1 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
```

```
"max S for 2 clusters= 0.658387568096074"
```

```
print("clustering for max S")
```

```
"clustering for max S"
```

```
1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

Lampiran 4. (Lanjutan)

```
"max G1 for 2 clusters= 13.8803623764797"  
print("clustering for max G1")  
"clustering for max G1"  
1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

36. Data 36

```
minimal nc where diffu>=0 is 3 for diffu= 1.0245  
1 2 1 2 2 2 2 1 2 2 1 1 1 2 2 2 2 1 1 2 2 1 2  
2 1 2 1 2 2 1 2 1 1 1 2 1 2 2 2 1 1 3  
  
"minimal nc for H<=10 equals 3 for H=  
1.61687460550354"  
"clustering for minimal nc where H<=10"  
1 2 1 2 2 2 2 1 2 2 1 1 1 2 2 2 2 1 1 2 2 1 2  
2 1 2 1 2 2 1 2 1 1 1 2 1 2 2 2 1 1 3  
  
"max KL for 3 clusters= 16.1877735738445"  
print("clustering for max KL")  
"clustering for max KL"  
1 2 1 2 2 2 2 1 2 2 1 1 1 2 2 2 2 1 1 2 2 1 2  
2 1 2 1 2 2 1 2 1 1 1 2 1 2 2 2 1 1 3  
  
"max S for 3 clusters= 0.427484697830619"  
print("clustering for max S")  
"clustering for max S"  
1 2 1 2 2 2 2 1 2 2 1 1 1 2 2 2 2 1 1 2 2 1 2  
2 1 2 1 2 2 1 2 1 1 1 2 1 2 2 2 1 1 3  
  
"max G1 for 3 clusters= 54.2152688629824"  
print("clustering for max G1")  
"clustering for max G1"  
1 2 1 2 2 2 2 1 2 2 1 1 1 2 2 2 2 1 1 2 2 1 2  
2 1 2 1 2 2 1 2 1 1 1 2 1 2 2 2 1 1 3
```

Lampiran 4. (Lanjutan)

37. Data 37

```
minimal nc where diffu>=0 is 3 for diffu= 0.4318
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 3 1 1 1 1 2
1 1 1 1 1 1 3 1 1 1 1 1 1 1 3 1 1 1 1 1 1

"minimal nc for H<=10 equals 3 for H=
0.57112709186776"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 3 1 1 1 1 2
1 1 1 1 1 1 3 1 1 1 1 1 1 1 3 1 1 1 1 1 1

"max KL for 9 clusters= 4.63976136476626"
print("clustering for max KL")
"clustering for max KL"
1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3 1 4 5 1 1 1 6
1 1 1 7 5 1 8 1 1 1 1 1 1 1 9 1 1 1 1 1 1

"max S for 2 clusters= 0.442298692751413"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

38. Data 38

```
minimal nc where diffu>=0 is 2 for diffu= 0.5414
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1
1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1

"minimal nc for H<=10 equals 2 for H=
1.16550684245894"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1
1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 2 1

"max KL for 8 clusters= 21.6181390184402"
print("clustering for max KL")
"clustering for max KL"
1 1 2 3 3 3 3 3 4 3 1 3 3 1 3 1 1 1 1 1 1 5 1
1 1 1 1 1 1 1 1 2 6 1 3 7 5 1 2 2 2 3 3 1 5 8
```

Lampiran 4. (Lanjutan)

```
"max S for 2 clusters= 0.484047400949713"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1
1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1

"max G1 for 13 clusters= 37.0495170760817"
print("clustering for max G1")
"clustering for max G1"
1 1 2 3 4 3 3 3 3 5 3 6 3 3 6 3 6
6 6 6 1 1 6 7 6 1 6 6 6 1 1 6 6 8
9 6 3 10 7 1 11 11 12 3 3 6 7 13
```

39. Data 39

```
minimal nc where diffu>=0 is 2 for diffu= 0.903
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"minimal nc for H<=10 equals 2 for H=
5.09360285271142"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"max KL for 8 clusters= 15.8853462029018"
print("clustering for max KL")
"clustering for max KL"
1 2 2 2 3 2 2 2 2 2 2 2 4 2 5 2 6 2 2 1 6 2 2 2 2
2 2 7 1 8 2 2 2 2 1 2 2 2 2 2 8 2 2 8 2 2 2 1 1

"max S for 6 clusters= 0.428069053622982"
print("clustering for max S")
"clustering for max S"
1 1 1 1 2 1 1 1 1 1 1 1 3 1 4 1 5 1 1 1 5 1 1 1 1
1 1 6 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"max G1 for 11 clusters= 25.7741283328156"
print("clustering for max G1")
"clustering for max G1"
```

Lampiran 4. (Lanjutan)

1	2	2	2	3	4	2	2	2	2	2	2	5	2	6	2	7
2	2	8	7	2	2	4	2	2	2	9	1	10	2	2	2	2
11	2	2	2	2	4	10	2	2	10	2	2	2	8	1		

40. Data 40

```
minimal nc where diffu>=0 is 2 for diffu= 0.1201
1 1 1 1 1 1 1 1 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"minimal nc for H<=10 equals 2 for H=
3.67733140112177"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"max KL for 8 clusters= 8.6395092957638"
print("clustering for max KL")
"clustering for max KL"
1 2 2 2 2 2 2 2 3 4 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 5
2 2 2 2 2 2 2 2 2 2 2 6 2 2 2 2 2 2 2 7 2 7 2 2 2 8

"max S for 2 clusters= 0.321106304986872"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 1 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"max G1 for 6 clusters= 5.50693201997446"
print("clustering for max G1")
"clustering for max G1"
1 2 2 2 2 2 2 2 3 4 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 5
2 2 2 2 2 2 2 2 2 2 2 6 2 2 2 2 2 2 2 2 2 2 2 2 2 2
```

41. Data 41

```
minimal nc where diffu>=0 is 3 for diffu= 0.1217
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 2
1 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

Lampiran 4. (Lanjutan)

```
"minimal nc for H<=10 equals 2 for H=
8.52772744641507"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"max KL for 12 clusters= 28.2218678172648"
print("clustering for max KL")
"clustering for max KL"
1 1 2 1 1 2 1 1 3 4 1 5 6 4 7 8 1
1 2 9 1 1 1 7 1 1 10 11 1 1 4 1 1 12
4 1 1 1 3 1 3 1 1 1 4 4 1 1 1

"max S for 2 clusters= 0.553039601016572"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"max G1 for 12 clusters= 20.9378810832184"
print("clustering for max G1")
"clustering for max G1"
1 1 1 2 1 1 2 1 1 3 4 1 5 6 4 7 8
1 1 2 9 1 1 1 7 1 1 10 1 1 1 4 1 1
11 12 1 1 1 3 1 3 1 1 1 4 4 1 1 1
```

42. Data 42

```
minimal nc where diffu>=0 is 2 for diffu= 0.1787
1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"minimal nc for H<=10 equals 2 for H=
7.05583447212115"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"max KL for 8 clusters= 4.62228808170707"
print("clustering for max KL")
```

Lampiran 4. (Lanjutan)

```
"clustering for max KL"
1 1 1 1 1 1 2 1 1 1 1 1 1 1 3 1 1 1 1 1 4 1 1 1 1
1 1 1 1 1 1 1 1 5 1 6 1 1 1 1 7 1 7 1 1 1 8 1 1 1

"max S for 2 clusters= 0.337461886995175"
print("clustering for max S")
"clustering for max S"
1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

"max G1 for 4 clusters= 5.636386775216"
print("clustering for max G1")
"clustering for max G1"
1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 3 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 4 1 4 1 1 1 3 1 1 1
```

43. Data 43

```
minimal nc where diffu>=0 is 2 for diffu= 0.1174
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1
1

"minimal nc for H<=10 equals 2 for H=
6.37680064251027"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1
1

"max KL for 7 clusters= 5.5775063979772"
print("clustering for max KL")
"clustering for max KL"
1 1 1 1 2 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 3 1 1 4
1 1 2 1 1 1 5 1 1 1 1 1 1 1 1 1 5 6 1 1 5 7 1 1
1

"max S for 2 clusters= 0.53195501832711"
print("clustering for max S")
"clustering for max S"
```

Lampiran 4. (Lanjutan)

1
1 2 1 1
1

```
"max G1 for 6 clusters= 13.5300803226621"
```

```
print("clustering for max G1")
```

```
"clustering for max G1"
```

1 1 1 1 2 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 3 1 1 2

1 1 2 1 1 1 4 1 1 1 1 1 1 1 1 1 4 5 1 1 4 6 1 1

1

44. Data 44

minimal nc where $\text{diffu} \geq 0$ is 2 for $\text{diffu} = 1.106$

1 1

1 1

2

"minimal nc for $H \leq 10$ equals 2 for $H =$

9.6126901624562"

```
"clustering for minimal nc where  $H \leq 10$ "
```

1 1

1 1

2

```
"max KL for 9 clusters= 46.3444207810641"
```

```
print("clustering for max KL")
```

"clustering for max KL"

1 1 2 3 4 4 5 4 1 6 4 4 1 1 1 1 1 1 1 4 1 1 4 7

4 1 4 4 4 4 4 4 1 1 1 4 8 1 1 1 1 6 1 4 1 4 4 4 4

9

```
"max S for 2 clusters= 0.753016753840711"
```

```
print("clustering for max S")
```

```
"clustering for max S"
```

1 1

1 1

2

Lampiran 4. (Lanjutan)

```
"max G1 for 9 clusters= 78.949619918067"  
print("clustering for max G1")  
"clustering for max G1"  
1 1 2 3 4 4 5 4 1 6 4 4 1 1 1 1 1 1 1 4 1 1 4 7  
4 1 4 4 4 4 4 4 1 1 1 4 8 1 1 1 1 6 1 4 1 4 4 4  
9
```

45. Data 45

```
minimal nc where diffu>=0 is 2 for diffu= 1.0873  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2  
1
```

```
"minimal nc for H<=10 equals 2 for H=  
6.91077753642758"  
"clustering for minimal nc where H<=10"  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2  
1
```

```
"max KL for 12 clusters= 21.856524147132"  
print("clustering for max KL")  
"clustering for max KL"  
1 1 2 3 2 4 5 2 1 6 2 7 2 2 1 1 1  
8 1 1 9 2 10 1 1 1 1 1 1 1 1 11 1  
1 1 1 1 1 5 2 1 6 1 6 2 2 7 12 2
```

```
"max S for 2 clusters= 0.765742750427772"  
print("clustering for max S")  
"clustering for max S"  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2  
1
```

```
"max G1 for 12 clusters= 36.7470387082630"  
print("clustering for max G1")  
"clustering for max G1"  
1 1 1 2 3 2 4 5 2 1 6 2 7 2 2 1 1  
1 8 1 1 9 2 10 1 1 1 1 1 1 1 1 11  
1 1 1 1 1 1 5 2 1 6 1 6 2 2 7 12 2
```

Lampiran 4. (Lanjutan)

46. Data 46

```
minimal nc where diffu>=0 is 3 for diffu= 0.1537
1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1
1 1 1

"minimal nc for H<=10 equals 2 for H=
5.79042440075691"
"clustering for minimal nc where H<=10"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1
1 1 1

"max KL for 8 clusters= 25.3082061263765"
print("clustering for max KL")
"clustering for max KL"
1 1 2 1 1 1 1 3 1 1 1 1 1 1 4 3 3 5 1 1 1 1 1 1
1 1 1 1 1 3 1 1 1 6 1 6 7 1 8 6 6 1 3 1 1 1 1 1
1 1 1

"max S for 3 clusters= 0.458687572547572"
print("clustering for max S")
"clustering for max S"
1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1
1 1 1

"max G1 for 8 clusters= 12.6124975545819"
print("clustering for max G1")
"clustering for max G1"
1 1 2 1 1 1 1 3 1 1 1 1 1 1 4 3 3 5 1 1 1 1 1 1
1 1 1 1 1 3 1 1 1 6 1 6 7 1 8 6 6 1 3 1 1 1 1 1
1 1 1
```

Lampiran 4. (Lanjutan)

47. Data 47

```
minimal nc where diffu>=0 is 2 for diffu= 0.0545
1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1
```

```
"minimal nc for H<=10 equals 2 for H=
3.52291662648021"
```

```
"clustering for minimal nc where H<=10"
```

```
1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1
```

```
"max KL for 4 clusters= 85.8485295433525"
```

```
print("clustering for max KL")
```

```
"clustering for max KL"
```

```
1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 3 1 4 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1
```

```
"max S for 2 clusters= 0.468131658814911"
```

```
print("clustering for max S")
```

```
"clustering for max S"
```

```
1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1
```

```
"max G1 for 2 clusters= 10.8799130726714"
```

```
print("clustering for max G1")
```

```
"clustering for max G1"
```

```
1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1
```

48. Data 48

```
minimal nc where diffu>=0 is 3 for diffu= 0.0966
1 2 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2
```

Lampiran 4. (Lanjutan)

```
"minimal nc for H<=10 equals 3 for H=
0.808316183152458"
"clustering for minimal nc where H<=10"
1 2 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

"max KL for 11 clusters= 9.42766493178323"
print("clustering for max KL")
"clustering for max KL"
2 1 3 4 4 4 4 4 4 5 4 4 4 4 4 4 4 4
4 6 4 4 4 4 4 4 7 4 4 4 4 4 4 4 4 4
4 4 4 4 4 4 4 4 4 8 8 8 4 4 4 4 4 4
4 9 9 4 10 4 4 7 4 4 7 7 7 4 11 7 4 4

"max S for 3 clusters= 0.391168397456279"
print("clustering for max S")
"clustering for max S"
1 2 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

"max G1 for 3 clusters= 9.81031728382731"
print("clustering for max G1")
"clustering for max G1"
1 2 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
```

49. Data 49

```
minimal nc where diffu>=0 is 3 for diffu= 0.7003
1 1 1 1 1 1 1 1 1 1 1 1 2 1 2 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 1 1 1
2 1 1 1 1 1 2 1 1 1 1 1 2 2 1 3 2 1 2 1 1 1 1

"minimal nc for H<=10 equals 4 for H=
1.36410775964019"
"clustering for minimal nc where H<=10"
```

Lampiran 4. (Lanjutan)

```
1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 2 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 3 1 1 2 1 1 1 1 1 1 2 1 1 1 1 1 2
2 1 1 1 1 1 2 1 1 1 1 1 2 2 1 4 2 1 2 1 1 1 1 1
```

```
"max KL for 3 clusters= 39.5020216206475"
```

```
print("clustering for max KL")
```

```
"clustering for max KL"
```

```
1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 2 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 1 1 1 2
2 1 1 1 1 1 2 1 1 1 1 1 2 2 1 3 2 1 2 1 1 1 1 1
```

```
"max S for 2 clusters= 0.79374539389082"
```

```
print("clustering for max S")
```

```
"clustering for max S"
```

```
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1
```

```
"max G1 for 3 clusters= 192.777506501787"
```

```
print("clustering for max G1")
```

```
"clustering for max G1"
```

```
1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 2 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 1 1 2
2 1 1 1 1 1 2 1 1 1 1 1 2 2 1 3 2 1 2 1 1 1 1
```

50. Data 50

```
minimal nc where diffu>=0 is 2 for diffu= 0.2515
```

```
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1
```

```
"minimal nc for H<=10 equals 4 for H=
```

```
9.77547042198963"
```

```
"clustering for minimal nc where H<=10"
```

```
1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 3 1 1 1 1 1 4 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1
```

Lampiran 4. (Lanjutan)

```
"max KL for 15 clusters= 27.0617034982362"
print("clustering for max KL")
"clustering for max KL"
1 2 3 1 1 1 1 1 1 4 1 1 1 1 1 5 1
1 1 1 1 6 7 1 1 1 1 1 1 1 1 1 8 1
1 1 7 1 1 9 1 1 10 1 11 1 1 12 1 1
1 1 1 13 1 1 1 14 1 1 1 1 1 1 1 1
1 1 1 1 1 1 15 1 1 1 1 6

"max S for 2 clusters= 0.911990324275285"
print("clustering for max S")
"clustering for max S"

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1

"max G1 for 2 clusters= 132.704705449779"
print("clustering for max G1")
"clustering for max G1"
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1
```

UNIVERSITAS BRAWIJAYA

