

**ANALISIS CLUSTER KOMPONEN UTAMA NONLINIER DAN
ANALISIS LATENT CLASS CLUSTER PADA DATA BERSKALA
CAMPURAN**

SKRIPSI

oleh :

IRFAN HADI KURNIAWAN

115090500111055



PROGRAM STUDI STATISTIKA

JURUSAN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS BRAWIJAYA

MALANG

2015

REPOSITORY.UB.AC.ID

UNIVERSITAS
BRAWIJAYA



REPOSITORY.UB.AC.ID

UNIVERSITAS
BRAWIJAYA



REPOSITORY.UB.AC.ID

UNIVERSITAS
BRAWIJAYA





ANALISIS CLUSTER KOMPONEN UTAMA NONLINIER DAN ANALISIS LATENT CLASS CLUSTER PADA DATA BERSKALA CAMPURAN

SKRIPSI

oleh:

IRFAN HADI KURNIAWAN

115090500111055



PROGRAM STUDI STATISTIKA

JURUSAN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS BRAWIJAYA

MALANG

2015



LEMBAR PENGESAHAN SKRIPSI
ANALISIS CLUSTER KOMPONEN UTAMA NONLINIER DAN
ANALISIS LATENT CLASS CLUSTER PADA DATA BERSEKALA
CAMPURAN

oleh :

IRFAN HADI KURNIAWAN

115090500111055

Setelah dipertahankan di depan Majelis Penguji
pada tanggal 28 Oktober 2015
dan dinyatakan memenuhi syarat untuk memperoleh gelar
Sarjana Sains dalam bidang Statistika

Mengetahui,
Dosen Pembimbing

Dr. Ir. Atiek Iriany, MS

NIP. 196308091988022001

Ketua Jurusan Matematika
Fakultas MIPA Universitas Brawijaya

Ratno Bagus Edy Wibowo, S.Si, M.Si, Ph.D

NIP. 197509082000031003



LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : IRFAN HADI KURNIAWAN
NIM : 115090500111055
Program Studi : STATISTIKA
Penulis Skripsi Berjudul :

ANALISIS *CLUSTER* KOMPONEN UTAMA *NONLINIER* DAN ANALISIS *LATENT CLASS CLUSTER* PADA DATA BERSEKALA CAMPURAN

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 dalam daftar pustaka dalam Skripsi ini.
2. Apabila dikemudian hari ternyata yang saya tulis terbukti hasil jiplakan, maka saya akan bersedia menanggung resiko yang akan saya terima.

Demikian pernyataan ini saya buat dengan segala kesadaran.

Malang, 28 Oktober 2015

Yang menyatakan

Irfan Hadi Kurniawan
NIM. 115090500111055

ANALISIS CLUSTER KOMPONEN UTAMA NONLINIER DAN ANALISIS LATENT CLASS CLUSTER PADA DATA BERSKALA CAMPURAN

ABSTRAK

Pada suatu penelitian, seringkali menemui berbagai macam skala yang digunakan pada data, antara lain data kategori dan data kontinu. Dalam suatu kasus terdapat suatu penelitian yang melibatkan dua jenis skala data, kategori dan kontinu yang disebut data campuran. Pada metode pengelompokan data berskala campuran terdapat analisis yang mampu diterapkan, pada kali ini menggunakan Analisis *Cluster* Komponen Utama *Nonlinier* dan Analisis *Latent Class Cluster*. Tujuan dari penelitian ini adalah membandingkan metode Analisis *Cluster* Komponen Utama *Nonlinier* dan Analisis *Latent Class Cluster* berdasarkan nilai Rasio Keragaman dalam pengelompokan pada kasus data berskala campuran menggunakan data Industri Sentra Tempe di Kota Malang Tahun 2005. Nilai rasio keragaman yang lebih kecil diindikasikan sebagai metode yang menghasilkan pengelompokan yang lebih baik. Hasil penelitian menunjukkan nilai rasio keragaman Analisis *Cluster* Komponen Utama *Nonlinier* sebesar 1.2 sedangkan Analisis *Latent Class Cluster* sebesar 1.609 sehingga dapat disimpulkan bahwa Analisis *Cluster* Komponen Utama *Nonlinier* lebih baik dibandingkan dengan Analisis *Latent Class Cluster* pada data Industri Sentra Tempe di Kota Malang Tahun 2005.

Kata Kunci : Analisis *Cluster* Komponen Utama *Nonlinier*, *Latent Class Cluster*, *Skala Campuran*

CLUSTER NONLINIER PRICIPAL COMPONENT ANALYSIS AND LATENT CLASS CLUSTER ANALYSIS ON THE MIXTURED SCALE DATA

ABSTRACT

In some research, often meet various scales used in data, such as categories and continuous data. Some cases there are a research involving two types of data scale, categories and continuous so-called mixture data. In the grouping method for mixture scale data there is an analysis which is able to be applied, at this time using Cluster Nonlinear Principal Component Analysis and Latent Class Cluster Analysis. The purpose of this research was to compare between Cluster Nonlinier Principal Component Analysis and Latent Class Cluster Analysis based on the ratio of diversity in the grouping the case of mixture scale data using Industrial Data Centers Tempe in Malang on 2005. The value ratio of diversity smaller is indicated as a method that produces better grouping. The results showed The value ratio of diversity for Cluster Nonlinier Principal Component Analysis is 1.2 while Latent Class Cluster Analysis is 1.609, so can be concluded that Cluster Nonlinier Principal Component Analysis better than Latent Class Cluster Analysis for Industrial Data Centers Tempe in Malang on 2005.

Keywords: Cluster Nonlinier Principal Component Analysis, Latent Class Cluster Analysis, Mixture Scale



KATA PENGANTAR

Puji syukur kehadiran Allah SWT karena atas segala limpahan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Analisis *Cluster* Komponen Utama *Nonlinier* dan Analisis *Latent Class Cluster* Pada Data Berskala Campuran” sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains dalam bidang Statistika.

Dalam pelaksanaan dan penyusunan skripsi ini, penulis telah banyak dibantu oleh berbagai pihak, baik berupa doa, bimbingan, saran, dorongan, dan semangat. Oleh karena itu, dalam kesempatan ini penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Dr. Ir. Atiek Iriany, MS selaku dosen pembimbing atas bimbingan, ilmu, waktu, kesabaran dan nasihat yang telah diberikan.
2. Dr. Suci Astutik, S.Si., M.Si selaku dosen penguji I atas waktu, ilmu serta saran yang telah diberikan.
3. Ir. Heni Kusdarwati, MS selaku dosen penguji II atas waktu, ilmu serta saran yang telah diberikan.
4. Dr. Ratno Bagus EW, M.Si., selaku Ketua Jurusan Matematika Fakultas MIPA Universitas Brawijaya.
5. Bapak dan Ibu dosen Jurusan Matematika atas ilmu dan bimbingan yang diberikan
6. Seluruh staff dan karyawan Jurusan Matematika Fakultas MIPA Universitas Brawijaya.
7. Bapak, Ibuk, yang senantiasa memberikan doa, semangat, kasih sayang, kesabaran, dan segalanya, serta mas Arif dan dek Fahmi yang selalu memberikan dukungan, semangat dan doa dalam setiap langkah untuk menuntun ilmu dan menyelesaikan tugas akhir.
8. Deani, Rizka, dan Nindiya atas waktu, bantuan, dan doanya bersama-sama melewati semhas dan ujian skripsi.
9. Rahman, Isal, Thomas, Irsyad, Putra, Qoid, dan Adit atas waktu dan semangatnya, cepet lulus.



10. Lita, Chandra, Niar, Ganis, Ika, Yanuriska, Susi, Erul, Angga, Aldo, Vita, Vida, Niluh, Manda, Fauzi, Dian, Dita, Linda, Ardi, Zaman, Nandi, Suci, Choir, Bagus, dan seluruh teman-teman Statistika 2011 atas semangat, doa, dan segala bantuan yang telah diberikan.
11. Direksi Studio Statistika 2014 (Uli, Puput, Ernia, Umi, Putri, Izza, Umar, Arin, Faik, dan Ayu) atas segala pengalaman, kekeluargaan, dan persahabatan.
12. Seluruh Pengurus dan Warga HIMAMASTA serta Seluruh Direksi dan Pengurus Studio Statistika atas semangat, doa, dukungan, dan pengalamannya.
13. Umi Faida atas semangat, doa, semangat, kesabaran, perhatian, dan dukungan yang telah diberikan.
14. Meta, Tyas, Dion, Fandi, Ahmad, Bowo, Suci, Novita, dan Meiga atas doa, waktu dan dukungannya. Malang – Ponorogo serasa dekat.
15. Kontrakkan Manja dan Cengger Ayam Dalam 76 sebagai tempat keluh kesah yang telah diberikan.
16. Semua pihak yang telah membantu dan berdoa dalam penyusunan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan mengingat keterbatasan kemampuan penulis. Untuk itu, dengan segala kerendahan hati penulis mengharap kritik dan saran.

Semoga skripsi ini bermanfaat bagi semua pihak yang membutuhkan.

Malang, Oktober 2015

Penulis



DAFTAR ISI

Halaman

HALAMAN JUDUL i

LEMBAR PENGESAHAN ii

LEMBAR PERNYATAAN iii

ABSTRAK vi

ABSTRACT v

KATA PENGANTAR vi

DAFTAR ISI viii

DAFTAR GAMBAR x

DAFTAR TABEL xi

DAFTAR LAMPIRAN xii

BAB I PENDAHULUAN 1

 1.1 Latar Belakang 1

 1.2 Rumusan Masalah 2

 1.3 Tujuan 2

 1.4 Manfaat 2

 1.5 Batasan Masalah 2

BAB II TINJAUAN PUSTAKA 3

 2.1 Skala Pengukuran Data 3

 2.2 Analisis *Cluster* Klasik 4

 2.3 Analisis Komponen Utama *Nonlinier* 5

 2.3.1 Analisis *Cluster* Komponen Utama *Nonlinier* 8

 2.3.2 Analisis *Cluster* Hirarki 9

 2.3.3 Validitas *Cluster* 10

 2.4 Analisis *Latent Class Cluster* 11

 2.4.1 Pendugaan Parameter LCC 14

 2.4.2 Penentuan Kelompok 17

 2.5 Keragaman Kelompok 18

BAB III METODE PENELITIAN 21

 3.1 Sumber Data 21

 3.2 Metode Analisis 21

BAB IV HASIL DAN PEMBAHASAN 25

4.1 Statistika Deskriptif 25

4.2 Analisis *Cluster* Komponen Utama *Nonlinier* 26

4.2.1 Analisis Komponen Utama *Nonlinier* 26

4.2.2 Analisis *Cluster* Hirarki dan Validitas

Cluster 26

4.3 Hasil Analisis *Latent Class Cluster* 28

4.3.1 Penentuan Model Terbaik 28

4.3.2 Pendugaan Parameter 30

4.4 Rasio Keragaman 30

4.5 Interpretasi *Cluster* 31

BAB V KESIMPULAN DAN SARAN 35

5.1 Kesimpulan 35

5.2 Saran 35

DAFTAR PUSTAKA 37

LAMPIRAN 41

DAFTAR GAMBAR

Gambar 3.1 Diagram Alir Proses ACKUN dan LCC 23



DAFTAR TABEL

Halaman

Tabel 3.1 Peubah dan Skala Data	21
Tabel 4.1 Statistik Deskriptif Variabel Jumlah Tenaga Kerja, Nilai Produksi, Bahan Baku, dan Nilai Investasi.	25
Tabel 4.2 Tabel Frekuensi Variabel Teknologi dan Aneka Olahan	25
Tabel 4.3 Hasil Indeks <i>Dunn</i> , DB, dan RMSD	27
Tabel 4.4 Hasil Peringkat Indeks Gabungan	28
Tabel 4.5 Keanggotaan 4 <i>Cluster</i>	28
Tabel 4.6 Model <i>Latent Class Cluster</i>	29
Tabel 4.7 Keanggotaan 5 <i>Cluster</i>	29
Tabel 4.8 Pendugaan Parameter	30
Tabel 4.9 Rasio Keragaman	31
Tabel 4.10 Hasil Analisis Deskriptif Setiap <i>Cluster</i>	31



DAFTAR LAMPIRAN

Halaman

- Lampiran 1. Data Industri Tempe di Kota Malang Tahun 2005 41
- Lampiran 2. Histogram Variabel Tenaga Kerja, Nilai Produksi,
Bahan Baku, dan Nilai Investasi 45
- Lampiran 3. *Pie Chart* Variabel Teknologi dan Aneka Olahan 47
- Lampiran 4. Kategori Peubah Selang Kelas 48
- Lampiran 5. *Syntax* SPSS untuk Komponen Utama *Nonlinier* 52
- Lampiran 6. Skor Komponen Objek 53
- Lampiran 7. Akar Ciri Komponen Utama *Nonlinier* dan
Komponen *Loading* 57
- Lampiran 8. Hasil *Output* Indeks Validitas 58
- Lampiran 9. Hasil Keanggotaan ACKUN 59
- Lampiran 10. Hasil *Output Latent Class Cluster* 63
- Lampiran 11. Hasil Keanggotaan LCC 65

BAB I PENDAHULUAN

1.1 Latar Belakang

Suatu populasi terdapat berbagai macam karakteristik yang mampu dijelaskan oleh data. Untuk mengetahui seberapa besar dan seberapa kuat karakteristik dalam data yang berfungsi menjelaskan keragaman populasi, umumnya melibatkan lebih dari satu peubah sehingga lebih rinci dalam mengklasifikasikan suatu karakteristik tertentu.

Menurut Hair dkk. (1998), analisis *cluster* klasik merupakan metode pengelompokan multivariat, tujuan utamanya mengelompokkan objek berdasarkan karakteristik yang dimiliki dan analisis ini akan menghasilkan kelompok-kelompok dengan homogenitas tinggi pada setiap kelompok dan heterogenitas tinggi antar kelompok, sehingga unit-unit pengamatan dalam satu kelompok mempunyai ciri-ciri lebih homogen dibandingkan unit pengamatan dalam kelompok lain (Mattjik dkk, 2002).

Pada kenyataannya seringkali dalam proses analisis, terdapat permasalahan dasar yang harus diatasi, salah satunya yaitu perbedaan jenis ukuran data. Menurut Siegel (1997) secara umum data memiliki 4 tingkat yaitu data berskala nominal, ordinal, interval, dan rasio. Dalam suatu penelitian seringkali tidak hanya melibatkan peubah yang memiliki jenis data kontinu (interval dan rasio), namun sering dijumpai berbagai macam peubah yang diamati memiliki jenis data kategori (nominal dan ordinal). Analisis *cluster* klasik yang merupakan metode pengelompokan data hanya bisa digunakan untuk permasalahan dengan menggunakan jenis data kontinu (interval dan rasio) karena dalam prosesnya menggunakan perhitungan matriks.

Terdapat metode yang sering digunakan untuk mengelompokkan data berskala campuran (kontinu dan kategori), salah satunya adalah Analisis *Cluster Hirarki* dengan transformasi Komponen Utama *Nonlinier* (Wulandari, 2006). Dalam penerapannya sama seperti menggunakan Analisis *Cluster Hirarki* namun dilakukan transformasi pada data menggunakan Komponen Utama *Nonlinier* yang akan menghasilkan skor komponen objek yang memiliki skala rasio sehingga dapat dianalisis lebih lanjut menggunakan analisis *cluster Hierarki* (Gifi, 1990).



Selain dengan Analisis *Cluster* Komponen Utama *Nonlinier*, terdapat metode lain yang dapat digunakan untuk mengelompokkan data pada data campuran yaitu analisis *Latent Class Cluster*, dalam proses pengelompokan dilihat berdasarkan peluang dari setiap anggota secara langsung, sehingga setiap objek diasumsikan masuk dalam setiap kelompok tertentu (Vermunt dan Magidson, 2004).

Pada penelitian ini akan membandingkan antara kedua metode yang mampu digunakan untuk pengelompokan data berskala campuran, yaitu menggunakan metode Analisis Komponen Utama *Nonlinier* dan *Latent Class Cluster*.

1.2 Rumusan Masalah

Apakah terdapat perbedaan hasil pada pengelompokan menggunakan metode Analisis *Cluster* Komponen Utama *Nonlinier* dan Analisis *Latent Class Cluster*?

1.3 Tujuan

Membandingkan metode Analisis *Cluster* Komponen Utama *Nonlinier* dan Analisis *Latent Class Cluster* berdasarkan nilai Rasio Keragaman dalam pengelompokan pada kasus data berskala campuran.

1.4 Batasan Masalah

Penentuan model terbaik dilihat berdasarkan keragaman kelompok, baik keragaman dalam kelompok maupun antar kelompok.

1.5 Manfaat

Manfaat lebih lanjut dari penelitian ini adalah memberikan informasi lebih lanjut mengenai metode Analisis *Cluster* Komponen Utama *Nonlinier* dan Analisis *Latent Class Cluster*. Selain itu untuk memberikan pandangan metode mana yang lebih baik digunakan untuk mengatasi kasus pengelompokan data pada skala campuran.



BAB II TINJAUAN PUSTAKA

2.1 Skala Pengukuran Data

Skala merupakan hasil dari suatu pengukuran pada beberapa jenis skala yang bervariasi, sedangkan pengukuran adalah pemberian angka terhadap objek atau fenomena menurut aturan tertentu. Teori pengukuran terdiri dari teori yang terpisah dan berbeda, masing-masing memiliki tingkat pengukuran yang berlainan. Menurut Siegel (1997), terdapat 4 tingkatan dalam skala pengukuran data, yaitu skala nominal, skala ordinal, skala interval, dan skala rasio.

Skala nominal merupakan skala yang digunakan hanya untuk pengklasifikasian suatu objek atau sesuatu yang diamati, berupa lambang-lambang atau angka-angka yang digunakan untuk mengidentifikasi kelompok-kelompok sebagai induk dari suatu objek (Siegel, 1997). Salah satu contoh yaitu pemberian label di suatu supermarket, misal untuk produk kaos diberi skor 1, kemudian produk pasta gigi diberi skor 2, dan seterusnya. Selain itu misal pada suatu biodata mengenai jenis kelamin, misal untuk laki-laki diberi skor 1, dan untuk perempuan diberi skor 2.

Skala ordinal merupakan skala pada suatu objek dalam kelompok dan memiliki suatu hubungan tertentu antara objek-objek yang memungkinkan memiliki tingkatan atau ranking antar objek. Tingkatan disini bisa diartikan jenis kelas-kelas, misalnya lebih tinggi, lebih disukai, lebih sulit, dan seterusnya atau bisa juga dengan tanda $<$, $>$, dan lain-lain yang mengartikan bahwa sesuatu lebih besar atau lebih kecil daripada, lebih suka atau lebih tidak suka daripada, dan seterusnya (Siegel, 1997). Salah satu contoh yaitu misal tingkatan pada sabuk bela diri, mulai dari rendah yaitu putih $>$ kuning $>$ hijau $>$ orange $>$ biru $>$ coklat. Tingkatan ini juga mampu diberikan skor sebagai berikut :

- Sabuk Putih, diberi skor 1
- Sabuk Kuning, diberi skor 2
- Sabuk Hijau, diberi skor 3 dan seterusnya

Skala interval merupakan skala yang tidak hanya menunjukkan sifat dari skala ordinal (urutan) saja, namun juga diketahui ukuran dari jarak antara kedua angka pada skala tersebut. Dalam skala interval, titik nol dan unit pengukurannya adalah sembarang (Siegel, 1997). Salah satu contoh apabila kita mengukur suhu dengan skala yang berbeda



(Celcius dan Fahrenheit), unit pengukuran dan titik nol dalam mengukur suhu adalah sembarang, keduanya mengandung informasi yang sama banyaknya dan sama jenisnya, hal ini dikarenakan keduanya berhubungan secara *linier*. Misal pada Celcius titik beku pada 0 derajat, tapi pada Fahrenheit titik beku berada pada 32 derajat.

Suatu skala memiliki semua ciri suatu skala interval dan disamping itu memiliki suatu titik nol sejati sebagai titik asalnya, maka skala itu dinamakan skala rasio. Salah satu contoh apabila kita mengukur massa atau berat, skala ons dan pon memiliki titik nol sejati. Rasio antara dua berat adalah independen terhadap unit pengukurannya sehingga dua berat dalam pon memiliki rasio yang sama, begitu juga untuk ons dan lainnya (Siegel, 1997).

2.2 Analisis Cluster Klasik

Menurut Mattjik, dkk. (2002), Analisis *Cluster* Klasik merupakan suatu metode dalam analisis peubah ganda yang bertujuan untuk mengelompokkan n satuan pengamatan ke dalam k kelompok dengan ($k < n$) berdasarkan p peubah, sehingga unit-unit pengamatan dalam satu kelompok mempunyai ciri-ciri yang lebih homogen dibandingkan unit pengamatan dalam kelompok lain. Pengelompokan tersebut dilakukan berdasarkan ukuran kedekatan antar masing-masing individu. Ukuran kedekatan sering disebut dengan jarak. Masalah yang mendasar dalam analisis ini adalah ukuran kedekatan yang digunakan serta penentuan cara pengelompokanya.

Adapun tujuan dilakukannya Analisis *Cluster* Klasik adalah untuk (Sitepu dkk, 2011):

1. Menggali data / eksplorasi data
2. Mereduksi data menjadi kelompok data baru dengan jumlah lebih kecil atau dinyatakan dengan klasifikasi data
3. Menggeneralisasi suatu populasi untuk memperoleh suatu hipotesis
4. Menduga karakteristik dalam data

Dalam aplikasinya, Analisis *Cluster* banyak digunakan dalam berbagai bidang, antara lain psikologi (pengklasifikasian individu ke dalam sifat kepribadian), analisis wilayah (pengklasifikasian kota ke dalam klasifikasi jenis penempatan berdasarkan peubah demografi dan fiskal), penelitian pemasaran (pengklasifikasian pelanggan kedalam segmentasi alasan faktor pembelian dan penggunaan produk), dan ilmu kimia (pengklasifikasian alasan senyawa performa suatu sifat), dan lain

lain. Banyak sekali masalah-masalah di masyarakat yang menggunakan Analisis *Cluster* Klasik, karena hasil dan interpretasinya bermanfaat dan luas.

Dalam Analisis *Cluster* Klasik terdapat 2 teknik pengelompokan yang sering digunakan, yaitu Analisis *Cluster Hirarki* dan Analisis *Cluster Non Hirarki*. Perbedaan secara mendasar dari dua teknik tersebut adalah pada *Cluster Hirarki* untuk proses pengelompokan terjadi berdasarkan tingkatan atau perhitungan dari analisis yang digunakan, dengan kata lain kita belum tahu berapa kelompok yang akan dibentuk, sedangkan *Cluster Non Hirarki* membentuk pengelompokan karakteristik berdasarkan kriteria sudah diketahui kelompok yang akan dibentuk.

2.3 Analisis Komponen Utama *Nonlinier*

Analisis Komponen Utama (*Principal Component Analysis* / PCA) adalah teknik yang digunakan untuk menyederhanakan suatu data, dengan cara mentransformasi *linier* sehingga terbentuk sistem koordinat baru dengan ragam maksimum. PCA dapat digunakan untuk mereduksi dimensi suatu data tanpa mengurangi karakteristik data tersebut secara signifikan. Pereduksian data dilakukan dengan cara mentransformasi *linier* sehingga terbentuk sistem koordinat baru dengan ragam maksimum tanpa mengurangi karakteristik data tersebut secara signifikan. Fungsi dari Analisis komponen utama sebagai berikut (Johnson dan Wichern, 2002):

1. Identifikasi peubah baru yang mendasari data peubah ganda.
2. Mereduksi jumlah himpunan peubah yang tidak berkorelasi dengan mempertahankan sebanyak mungkin keragaman data.
3. Menghilangkan peubah-peubah asal yang tidak memberi informasi yang penting.

Perbedaan mendasar antara analisis komponen utama dengan analisis komponen utama *nonlinier* adalah mengenai kegunaan yang mana pada analisis komponen utama tidak memungkinkan untuk dianalisis data dengan hubungan antar peubah *nonlinier*, atau memiliki skala pengukuran yang berbeda, sedangkan pada analisis komponen utama *nonlinier* memungkinkan untuk diaplikasikan pada berbagai tingkat skala pengukuran peubah tidak hanya pada numerik, tapi juga bisa pada ordinal dan nominal (Konig, 2002).



Analisis komponen utama *nonlinier* merupakan pengembangan dari analisis komponen utama dengan menggunakan pendekatan *Alternating Least Square* atau sering disebut PRINCALS (*Principal Componen Analysis by Alternating Least Square*) (Gifi, 1989). Secara umum konsep dari analisis komponen utama *nonlinier* adalah membentuk suatu peubah baru dengan tetap mampu menjelaskan populasi dari peubah awal, untuk peubah baru dengan skala kategori diberikan suatu nilai numerik sehingga mampu dihitung dalam transformasi, hal tersebut disebut dengan kuantifikasi (Kroonenberg, dkk. 1997).

Kategori peubah sebaiknya sebagian besar memiliki frekuensi yang lebih besar dari pada 4 (Gifi, 1990) dan banyaknya kategori peubah dengan dengan skala matriks adalah 7-10 kategori (Kroonenberg, dkk. 1997). Pada penelitian sebelumnya (Wulandari, 2006) menjelaskan bahwa banyaknya kategori perubah dapat ditentukan berdasarkan pada peneliti sendiri karena diperoleh kesimpulan bahwa semakin bertambahnya kategori untuk peubah tidak menjamin bahwa proporsi keragaman akan semakin besar, sehingga lebih baik tetap berpedoman bahwa banyaknya kategori antara 4 sampai 10 kategori.

Setelah menentukan jumlah peubah kategori dalam kelas, selang kelas ditentukan dalam dengan rumus sebagai berikut (Yitnosumarto, 1994) :

$$I = \frac{R}{k} \tag{2.1}$$

$$R = X_t - X_r$$

$$k = 1 + 3.22 \log n$$

di mana :

- I : selang dalam kelas
- R : kisaran atau *range*
- X_t : nilai maksimal / pengamatan tertinggi
- X_r : nilai minimal / pengamatan tertinggi
- k : banyaknya kategori peubah
- n : banyaknya pengamatan

Gifi (1990) menjelaskan bahwa bila terdapat suatu data terbentuk dalam matriks **H** berukuran $n \times m$, maka digunakan notasi sebagai berikut :



n : banyaknya pengamatan (objek)

m : banyaknya peubah

p : banyaknya dimensi (untuk peubah $j, j=1, \dots, m$)

h_j : Vektor n kategori pengamatan

k_j : Banyaknya kategori pada peubah ke- j

G_i : Matriks indikasi untuk peubah ke- j pada ukuran $n \times k_j$
 Elemen dari G_i diartikan sebagai berikut ($i = 1, \dots, n; k = \dots, K_j$)

$$G_j = (g_{jtk}) = \begin{pmatrix} g_{j1k_1} & \dots & g_{j1k_{K_j}} \\ \vdots & & \vdots \\ g_{jnk_1} & \dots & g_{jnk_{K_j}} \end{pmatrix} = (g_{j1}, g_{j2}, \dots, g_{jK_j}) \quad (2.2)$$

$g_{ijk} = \begin{cases} 1, & \text{jika objek ke-}i \text{ berada pada kategori ke-}k \text{ dalam peubah } j \\ 0, & \text{jika objek ke-}i \text{ tidak berada pada kategori ke-}k \text{ dalam peubah } j \end{cases}$

matriks G_j merupakan matriks yang berisi frekuensi setiap kategori peubah, dengan persamaan sebagai berikut :

$$G_j u = M_j u \quad (2.3)$$

$$M^* = \sum_{j=1}^m M_j = mI \quad (2.4)$$

$$D_j = G_j' G_j \quad (2.5)$$

u : Vektor kolom dengan seluruh elemennya adalah 1

M_j : Matriks diagonal yang berupa matriks I .

Didefinisikan sebagai berikut

$$m(j)_{ii} = \begin{cases} 1, & \text{jika pengamatan ke-}i \text{ berada pada selang } [1, k_j] \\ 0, & \text{jika pengamatan ke-}i \text{ berada di luar selang } [1, k_j] \end{cases}$$

D_j : Matriks diagonal berisikan marginal univariat, contohnya jumlah kolom dari G_j



Model Analisis Komponen Utama *Nonlinier* adalah sebagai berikut :

$$Z = XY \tag{2.6}$$

di mana X merupakan matriks skor komponen objek dengan ordo $n \times p$ di mana p merupakan banyaknya dimensi dan Y merupakan *multiple* kategori kuantifikasi. Tujuan utama dari PRINCALS adalah mencari nilai komponen X dan kuantifikasi kategori Y_j (untuk $j=1, \dots, m$), dengan didasari oleh teori *meet loss* yang bertujuan untuk meminimumkan fungsi *homogeneity loss* (σ_M) sebagai berikut (Gifi, 1989) :

$$\sigma(X; Y) = \frac{1}{m} \sum_j \text{tr} (X - G_j Y_j)' M_j (X - G_j Y_j) \tag{2.7}$$

Fungsi tersebut minimal dalam batasan normalisasi $X' M * X = mn I$, di mana $M * = \sum_j M_j$, dan I merupakan Matriks identitas $p \times p$ dalam pencantuman M_j ke dalam $\sigma(X; Y)$ menjamin bahwa tidak ada pengaruh dari nilai di luar selang $[1, k_j]$ yang mungkin dapat menyebabkan kasus data hilang. Skor komponen akan terpusat memenuhi $u' M * X = 0$.

2.3.1 Analisis *Cluster* Komponen Utama *Nonlinier*

Menurut Hair, dkk. (1998), dalam Analisis *Cluster* Klasik terdapat beberapa asumsi, pertama yaitu kecukupan sampel dalam menjelaskan suatu populasi atau dengan kata lain mampu mewakili populasi dan asumsi kedua adalah pengaruh multikolinieritas. Dalam kenyataannya asumsi multikolinieritas sering kali diatasi menggunakan analisis komponen utama, yaitu dengan mengubah peubah dalam data menjadi peubah baru yang saling bebas dengan tetap mempunyai karakteristik untuk mewakili populasi (Jolliffe, 2002).

Konig (2002) menjelaskan bahwa analisis komponen utama hanya dapat digunakan untuk data yang mempunyai skala kontinu dengan asumsi antar peubah *linier*. Hal ini tidak memungkinkan apabila hubungan antar peubah memiliki skala pengukuran yang berbeda, mengindikasikan bahwa hubungan antar peubah *Nonlinier*.



Kemampuan analisis komponen utama dalam mengatasi multikolinieritas dalam data memiliki batasan hanya pada antar peubah dengan memiliki satuan yang berbeda, hal ini mampu diatasi dengan pembakuan pada data, akan tetapi apabila skala pengukuran peubah berbeda perlu dilakukan analisis lain yang mampu mengatasi permasalahan pada peubah dengan hubungan *Nonlinier* atau memiliki perbedaan skala pengukuran (Wulandari, 2006).

Dalam analisis komponen utama *Nonlinier*, nilai yang dihasilkan sama dengan analisis komponen utama, yaitu berupa nilai skor komponen suatu objek yang pada akhirnya digunakan untuk analisis *cluster*. Wulandari (2006) menjelaskan bahwa skor komponen suatu objek memiliki skala pengukuran rasio atau berjenis data matriks sehingga dapat dianalisis lebih lanjut dengan Analisis *Cluster* Klasik.

2.3.2 Analisis *Cluster* Hirarki

Analisis *cluster* memiliki dua macam metode, yaitu menggunakan analisis *cluster hirarki* dan non *hirarki*. Pada penelitian kali ini menggunakan metode *hirarki* karena belum diketahuinya *cluster* yang akan dibentuk. Analisis *cluster hirarki* adalah analisis yang mana melakukan pengelompokkan data dengan cara mengukur jarak pada setiap obyek dengan cara mengelompokkan dua atau lebih objek yang mempunyai kesamaan paling dekat, kemudian proses berlanjut ke objek berikutnya yang memiliki kedekatan kedua, dan demikian seterusnya. Kemudian dari hasil proses tersebut membentuk sebuah diagram yang disebut dendogram. Dendogram merupakan diagram pohon untuk memisalkan atau mengilustrasikan hasil dari *cluster* oleh pendekatan *Cluster* Hirarki.

Pada dendogram menunjukkan penggabungan (*agglomerative*) atau pembagian (*divisive*). Metode penggabungan (*agglomerative*) pada penelitian kali ini menggunakan ukuran kedekatan / jarak *Euclidean* dengan persamaan sebagai berikut:

$$d(i, j) = \left[\left(\sum_{k=1}^m (x_{ik} - x_{jk})^2 \right) \right]^{1/2} \quad (2.8)$$

dengan menggunakan ukuran kemiripan sifat yang sering digunakan adalah Metode Pautan Rata-rata (*Average Linkage*) sebagai berikut :

$$d_{KL} = \frac{1}{n_K n_L} \sum_{i \in C_K} \sum_{j \in C_L} d(x_i, x_j) \quad (2.9)$$



2.3.3 Validitas *Cluster*

Menurut Hair dkk. (1998), validitas disertakan oleh peneliti dengan maksud meyakinkan bahwa hasil kelompok *cluster* yang terbentuk mampu mewakili atau menjelaskan populasi secara umum maupun spesifik. Oleh karena itu uji validitas *cluster* perlu dilakukan setelah melakukan berbagai jenis macam analisis *cluster* untuk mengevaluasi hasil dari suatu analisis *cluster* yang sesuai (Indriani, 2013). Validitas *cluster* juga diperlukan untuk pengukuran kebaikan dari suatu proses dalam analisis *cluster*.

Terdapat beberapa indeks yang digunakan dalam validitas dalam pengelompokan, antara lain Indeks *Dunn*, Indeks *Davies Bouldin*, dan Indeks *RMSSDT*.

A. Indeks *Dunn*

Indeks *Dunn* (*D*) dituliskan dalam rumus sebagai berikut :

$$D = \min_{i=1..n_c} \left\{ \min_{j=i+1..n_c} \left(\frac{d(c_i, c_j)}{\max_{k=1..n_c} (\text{diam}(c_k))} \right) \right\} \quad (2.10)$$

di mana :

$$d(c_i, c_j) = \min_{x \in c_i, y \in c_j} \{d(x, y)\}$$

$$\text{diam}(c_k) = \max_{x, y \in c_i} \{d(x, y)\}$$

Dalam penggunaan indeks *Dunn* ini dilihat berdasarkan pada gagasan indentifikasi sebagai kelompok yang padat dan terpisah dengan baik. Jika data teridentifikasi sebagai kelompok yang baik, dilihat dari jarak antar kelompok biasanya lebih besar dan diameter dari *cluster* diharapkan lebih kecil, sehingga semakin besar nilai Indeks *Dunn* diindikasikan terdapat konfigurasi *cluster* yang lebih baik (Kovacs dkk., 2006).

B. Indeks *Davies Bouldin*

Indeks ini dilihat berdasarkan ukuran kesamaan *cluster* (R_{ij})

$$R_{ij} = \frac{s_i + s_j}{d_{ij}} \quad (2.11)$$

$$d_{ij} = d(v_i, v_j), \quad s_i = \frac{1}{\|c_{i1}\|} \sum_{x \in c_i} d(x, v_i)$$

kemudian Indeks *Davies Bouldin* didefinisikan sebagai berikut :

$$DB = \frac{1}{n_c} \sum_{i=1}^{n_c} R_i, \text{ di mana}$$

$$R_i = \max_{j=1, \dots, n_c, i \neq j} (R_{ij}), i = 1, \dots, n_c$$

di mana :

R_{ij} : Ukuran kesamaan *cluster*

s_i : Jarak d dalam kelompok

d_{ij} : Jarak antar kelompok

n_c : Jumlah kelompok yang terbentuk

Indeks *Davies Bouldin* mengindikasikan keeratan dalam *cluster*, semakin kecil nilai Indeks *Davies Bouldin* semakin lebih baik kelompok yang terbentuk (Kovacs, dkk. 2006).

C. Indeks RMSSDT

RMSSDT (*root mean square standart deviation*) merupakan salah satu indeks yang sering kali digunakan untuk mengevaluasi hasil dari proses analisis *Cluster*. Indeks RMSSDT dituliskan dalam rumus sebagai berikut :

$$RMSSDT = \sqrt{\frac{\sum_{j=1, \dots, d} \sum_{k=1}^{n_{ij}} (x_k - \bar{x}_j)^2}{\sum_{j=1, \dots, d} (n_{ij} - 1)}} \quad (2.12)$$

di mana

\bar{x}_j : rata-rata kelompok ke-j

n_{ij} : banyaknya objek ke-i pada kelompok ke-j

Semakin kecil nilai RMSSDT diindikasikan bahwa kelompok yang terbentuk akan semakin baik.

2.4 Analisis *Latent Class Cluster*

Latent class model sering digunakan pada ilmu sosial yang bertujuan mengklasifikasikan individu atau objek pada perbedaan dalam sebuah grup atau kelompok yang didasari oleh beberapa respon pada sekumpulan objek yang diamati *latent class* model peubah campuran untuk mengakomodasikan semua jenis data (termasuk ordinal dan nominal). Model yang dianjurkan diduga menggunakan *Maximum*



Likelihood dengan algoritma EM (*Expectation Maximum*) (Moustaki dan Papageorgiou, 2004).

Kegunaan utama dari analisis *Latent Class Cluster* yaitu untuk mengidentifikasi dari beberapa objek ke dalam suatu kelompok (*cluster*) yang mempunyai kesamaan sifat dari segi kesukaan, nilai, karakteristik, perilaku, dan lain sebagainya yang di dalamnya termasuk K -kategori peubah laten, yang di mana setiap kategori diwakilkan dalam setiap *cluster* (Vermunt dan Magidson, 2004).

Salah satu kelebihan dari analisis *Latent Class Cluster* adalah kemampuan model yang mampu ditempatkan dan digunakan pada berbagai jenis data yang memberikan kemungkinan untuk menguji kelayakan model (Moustaki dan Papageorgiou, 2004).

Berikut adalah fungsi sebaran peluang bersama dari peubah dengan data berskala campuran yang diamati sebagai berikut :

$$f(x_h) = \sum_{j=1}^K \eta_j g(x_h|j) \quad (2.13)$$

di mana

η_j : peluang awal (prior) kelas j ($\eta_j = \frac{n_j}{N}$)

$g(x_h|j)$: fungsi distribusi dari peubah yang berupa nominal, ordinal, atau kontinyu (interval atau rasio).

Pada *Latent Class Cluster*, diasumsikan faktor jarak konsisten untuk K *cluster* dan setiap kelas ke- j terdapat peluang keanggotaannya. Hal tersebut ditengahi sebagai peluang *prior* karena memberikan peluang untuk masuk kedalam suatu kelompok untuk melihat suatu data.

Setiap peubah dalam analisis *Latent Class Cluster* memiliki sebaran bersyarat sesuai dengan jenis skala pada data yang digunakan. Pada peubah berskala kategori mengikuti distribusi multinomial, sedangkan untuk peubah berskala kontinyu mengikuti distribusi normal.

1. Model *Latent Class Cluster* untuk Skala Nominal

Pada peubah berskala nominal distribusi yang diasumsikan pada model adalah distribusi multinomial sebagai berikut:

$$g_i(x_i|j) = \prod_{s=1}^{c_i} (\pi_{ij(s)})^{x_i(s)} \quad (2.14)$$



di mana

$\pi_{ij(s)}$: Peluang suatu objek akan berada dikelas j untuk kategori s dan peubah i .

j : Banyaknya kelas yang terbentuk

s : Banyaknya kategori ($s = 1, 2, \dots, c_i$)

$x_{i(s)} \begin{cases} \text{bernilai 1 jika respon jatuh pada kategori} \\ s (s = 1, 2, \dots, c_i) \\ \text{bernilai 0 untuk selainnya} \end{cases}$

2. Model *Latent Class Cluster* untuk Skala Ordinal

Peubah berskala ordinal dengan indikator x_i didefinisikan seperti peubah berskala nominal dengan distribusi yang diasumsikan pada model adalah distribusi multinomial dengan penambahan pada order atau tingkatan dalam data, sebagai berikut :

$$g_i(x_i|I) = \prod_{s=1}^{m_i} \pi_{ij(s)}^{x_{i(s)}} \quad (2.15)$$

$$= \prod_{s=1}^{m_i} (\gamma_{ijs(s)} - \gamma_{ijs(s-1)})^{x_{i(s)}}$$

di mana

$\pi_{ij(s)}$: Peluang suatu objek akan berada di kelas j untuk kategori s dan peubah i

j : Banyaknya kelas yang terbentuk

s : Banyaknya kategori ($s = 1, 2, \dots, c_i$)

$\gamma_{ijs(s)}$: Peluang kumulatif dari respon (dengan $\gamma_{ijs(s)} =$

$\pi_{ij(1)} + \pi_{ij(2)} + \dots + \pi_{ij(s)}$)
 $x_{i(s)} \begin{cases} 1 \text{ untuk } i \text{ dengan kategori jawaban } s \\ 0 \text{ untuk selainnya} \end{cases}$

3. Model *Latent Class Cluster* untuk Skala Interval dan Rasio (Kontinyu)

Pada data berskala interval dan rasio diasumsikan model mengikuti distribusi normal, dengan bentuk sebagai berikut :

$$g(x_i | \mu_{ij}, \sigma_i^2) = (2\pi)^{-1/2} \sigma_i^{-1/2} \exp \left[-\frac{1}{2\sigma_i^2} (x_i - \mu_{ij})^2 \right] \quad (2.16)$$

di mana

μ_{ij} : rata-rata peubah ke- i pada kelas j

σ_i^2 : ragam dari peubah ke- i

π : 3.14

\exp : 2.718

Sehingga diperoleh fungsi sebaran peluang bersama (persamaan 2.15) pada data nominal, ordinal, dan kontinyu pada persamaan 2.14, 2.15, dan 2.16 diperoleh sebagai berikut

$$\begin{aligned} f(x_h) &= \sum_{j=1}^K \eta_j g(x_h | j) \\ &= \sum_{j=1}^K \eta_j \left[\prod_{i=1}^N \left(\prod_{s=1}^{c_i} (\pi_{ij(s)})^{x_{i(s)}} \right) \right] \left[\prod_{i=1}^0 \left(\prod_{s=1}^{m_i} (\gamma_{ij(s)})^{\gamma_{ij(s-1)}} \right) \right] \left[\prod_{i=1}^Q \left((2\pi)^{-1/2} \sigma_i^{-1/2} \exp \left[-\frac{1}{2\sigma_i^2} (x_i - \mu_{ij})^2 \right] \right) \right] \end{aligned}$$

2.4.1 Pendugaan Parameter *Latent Class Cluster*

Pada *Latent Class Cluster* pendugaan parameter yang digunakan adalah metode *Maximum Likelihood*. Metode *Maximum Likelihood* merupakan salah satu cara untuk mengestimasi parameter yang tidak diketahui. Prosedur pendugaan ini menguji apakah pendugaan maksimum yang tidak diketahui dari fungsi *likelihood* suatu sampel nilainya sudah memaksimalkan fungsi *likelihood* (Haeruddin, 2013).

Pada metode *Maximum Likelihood*, misalkan $x_1, x_2, x_3, \dots, x_n$ merupakan peubah yang saling bebas dan mempunyai fungsi kepadatan peluang $f(x, \theta)$ di mana θ merupakan parameter, maka fungsi *likelihood* dapat didefinisikan sebagai berikut :

$$L(x_1, x_2, x_3, \dots, x_n | \theta) = f(x_1, \theta) f(x_2, \theta) f(x_3, \theta) \dots f(x_n, \theta) \quad (2.17)$$

Jika fungsi *likelihood* terdiferensialkan pada θ , maka estimasi *Maximum Likelihood* dapat diperoleh melalui persamaan berikut :

$$(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_n) = \frac{\partial L(x_1, x_2, \dots, x_n; \theta)}{\partial \theta_i} = 0 \quad (2.18)$$

dengan $i = 1, 2, \dots, n$

Langkah-langkah untuk menentukan estimasi maksimum *likelihood* dari θ_i adalah :

1. Menentukan fungsi *likelihood* pada persamaan (2.18)
2. Membentuk logaritma natural *likelihood*
3. Membentuk persamaan *likelihood* dan menyelesaikan

$$\frac{\partial L(x_1, x_2, \dots, x_n; \theta)}{\partial \hat{\theta}_i} = 0$$

4. Mendapatkan estimasi *maksimum likelihood* dari θ_i , yaitu $\hat{\theta}_i$

Pada *Latent Class Cluster*, setiap skala pengukuran data memiliki parameter yang berbeda-beda. Parameter dengan skala nominal menggunakan x , untuk parameter skala ordinal menggunakan γ , sedangkan untuk parameter skala kontinu bisa menggunakan μ dan σ .

Fungsi *log likelihood* untuk digunakan sebagai sampel acak berukuran n adalah sebagai berikut :

$$L = \sum_{h=1}^n \log f(x_h) \text{ dan} \quad (2.19)$$

$$L(\theta|x) = \sum_{h=1}^n \log \sum_{j=1}^K \eta_j g(x_h|j) \quad (2.20)$$

$$L(\theta|x) = \sum_{h=1}^n \log \sum_{j=1}^K \eta_j \left[\prod_{i=1}^{c_i} \left(\prod_{s=1}^{c_i} (\pi_{ij(s)})^{x_i(s)} \right) \right] \left[\prod_{i=1}^0 \left(\prod_{s=1}^{m_i} (\gamma_{ij(s)})^{y_{ij(s-1)}} \right) \right] \left[\prod_{i=1}^0 \left((2\pi)^{-1/2} \sigma_i^{-1/2} \exp \left[-\frac{1}{2\sigma_i^2} (x_i - \mu_{ij})^2 \right] \right) \right] \quad (2.21)$$

Algoritma EM (*Expectation Maximum*) merupakan sebuah metode iteratif untuk estimasi *Maximum Likelihood* (MLE) yang berguna dalam permasalahan data yang tidak lengkap (*missing data*) (Haeruddin, 2013). Secara spesifikasi iterasi menggunakan algoritma



EM berfungsi untuk memaksimalkan fungsi *likelihood*. Dalam setiap iterasi pada Algoritma EM ini terdapat 2 tahap, yaitu tahap Ekspektasi atau tahap *E-step* dan tahap Maksimisasi atau tahap *M-step*.

Menurut Haeruddin (2013), *E-step* bertujuan untuk menemukan ekspektasi bersyarat dari *missing* data dengan syarat data yang diketahui nilainya dan penduga parameternya, kemudian mensubstitusikan nilai ekspektasi yang diperoleh terhadap *missing* data. Sedangkan *M-step* bertujuan untuk memaksimalkan fungsi loglikelihood dengan cara mencari turunan parsial dari fungsi *log likelihood* tersebut.

Moustaki dan Papageorgiou (2004) menjelaskan bahwa langkah-langkah algoritma EM sebagai berikut :

1. Memilih nilai inisial untuk pendugaan awal peluang $h(j|x_h)$.
2. Memperoleh pendugaan awal untuk model parameter.
3. Memperoleh pendugaan baru untuk $h(j|x_h)$
4. Kembali ke langkah 2 dan seterusnya sampai diperoleh konvergensi.

Peluang awal kelas diestimasi sebagai berikut :

$$\hat{\eta}_j = \sum_{h=1}^n \hat{h}(j|x_h) / n \tag{2.22}$$

Peluang bersyarat diestimasi bahwa $x_i = s$ diberikan kelas ke- j untuk variabel nominal adalah sebagai berikut :

$$\hat{\pi}_{ij(s)} = \sum_{h=1}^n x_{ih(s)} \hat{h}(j|x_h) / (n\hat{\eta}_j) \tag{2.23}$$

Peluang bersyarat kumulatif diestimasi bahwa $x_i \leq s$ diberikan kelas ke- j untuk variabel ordinal adalah sebagai berikut :

$$\hat{\gamma}_{ij(s)} = \hat{\gamma}_{ij(s-1)} + \sum_{h=1}^n x_{ih(s)} \hat{h}(j|x_h) / (n\hat{\eta}_j) \tag{2.24}$$

Estimasi parameter untuk variabel kontinu adalah sebagai berikut:

$$\hat{\mu}_{ij} = \sum_{h=1}^n x_{ih(s)} \hat{h}(j|x_h) / (n\hat{\eta}_j) \tag{2.25}$$

dan ragam (diasumsikan konstan) setiap kelas adalah sebagai berikut :

$$\hat{\sigma}_i^2 = \frac{\sum_{h=1}^n \sum_{j=1}^K (x_{ih} - \hat{\mu}_{ij})^2 \hat{h}(j|x_h)}{\sum_{h=1}^n \sum_{j=1}^K \hat{h}(j|x_h)} \quad (2.26)$$

dan juga peluang posterior diestimasi sebagai berikut :

$$\hat{h}(j|x_h) = \hat{\eta}_j \hat{g}(x_h|j) / \hat{f}(x_h) \quad (2.27)$$

Setelah didapatkan hasil yang optimal menggunakan iterasi dengan EM, kemudian prosedur ini akan menghasilkan model yang konvergen sehingga didapatkan estimasi parameter yang stabil. Oleh karena itu pendugaan parameter dalam analisis *Latent Class Cluster* dilakukan menggunakan *Maximum Likelihood* dengan iterasi algoritma EM.

2.4.2 Penentuan Kelompok

Asumsi kebebasan lokal merupakan asumsi yang mendasari analisis *Latent Class Cluster*. Ketidakcocokan model dalam *Latent Class Cluster* disebabkan oleh pelanggaran pada asumsi tersebut. Sehingga apabila terdapat pelanggaran asumsi, maka dapat diatasi dengan menambah jumlah kelompok. Uji ketepatan model salah satu yang digunakan adalah BIC (*Bayesian Information Criterion*) (Vermunt dan Magidson, 2004).

BIC mencerminkan peningkatan nilai jumlah kuadrat residual dan jumlah parameter dari model yang digunakan, sehingga model dengan nilai BIC yang lebih kecil mengindikasikan model tersebut model yang terbaik karena model yang dihasilkan lebih bisa menjelaskan ragam dalam data. Menurut Vermunt dan Magidson (2004) nilai BIC dapat diperoleh menggunakan rumus sebagai berikut :

$$BIC = -2LL + \ln(N)M \quad (2.28)$$

di mana :

N : Banyaknya objek pengamatan

M : Banyaknya parameter

LL : Nilai maksimum fungsi loglikelihood dari suatu model



2.5 Keragaman Kelompok

Dalam proses pengelompokan, terdapat 2 jenis keragaman, yaitu keragaman dalam kelompok (*Sum of Square Within Cluster* / SSW) dan keragaman antar kelompok (*Sum of Square Between Cluster* / SSB).

1. Keragaman untuk data kontinyu

Dijelaskan oleh Bunker, dkk (1996) pada Umah (2014), bahwa persamaan untuk menghitung SSW dan SSB dengan data kontinyu adalah sebagai berikut :

$$SSW = \frac{1}{c} \sum_{c=1}^c S_c \tag{2.29}$$

$$SSB = \left[\frac{1}{c-1} \sum_{c=1}^c (\bar{x}_c - \bar{x})^2 \right]^{1/2} \tag{2.30}$$

di mana

SSW : Keragaman dalam kelompok

SSB : Keragaman antar kelompok

S_c : Simpangan baku kelompok ke -c

\bar{x}_c : Rata-rata kelompok ke-c

\bar{x} : Rata-rata total

C : Banyaknya kelompok

2. Keragaman untuk data kategori

Menurut Light dan Margolin (1997) pada Umah (2014) dijelaskan bahwa persamaan rumus untuk data kategori adalah sebagai berikut :

$$SSW = \left[\frac{1}{n-c} \left(\frac{n}{2} - \frac{1}{2} \sum_{c=1}^c \frac{1}{n_c} \sum_{k=1}^K n_{kc} \right) \right]^{1/2} \tag{2.31}$$

$$SSB = \left[\frac{1}{c-1} \left(\frac{1}{2} \left(\sum_{c=1}^c \frac{1}{n_c} \sum_{k=1}^K n_{kc}^2 \right) - \frac{1}{2n} \sum_{k=1}^K n_{k.}^2 \right) \right]^{1/2} \tag{2.32}$$

di mana

n_k : banyaknya objek pada kategori ke-k, $k=1,2,...,K$

n_c : banyaknya objek pada kelompok ke-c, $c=1,2,...,C$

n_{kc} : banyaknya objek pada kategori ke-k, kelompok ke-c

Hasil pengelompokan dikatakan baik apabila keragaman dalam kelompok (SSW) memiliki keragaman yang rendah sedangkan keragaman antar kelompok (SSB) memiliki keragaman yang tinggi.

Repository Universitas Brawijaya

REPOSITORY.UB.AC.ID

UNIVERSITAS
BRAWIJAYA



REPOSITORY.UB.AC.ID

UNIVERSITAS
BRAWIJAYA



REPOSITORY.UB.AC.ID

UNIVERSITAS
BRAWIJAYA



BAB III

METODE PENELITIAN

3.1 Sumber Data

Data tentang Industri Tempe Sentra di Kota Malang tahun 2005 merupakan data sekunder yang diperoleh dari penelitian Wulandari (2006) dengan peubah dan skala data sebagai berikut :

- Jumlah Objek : 102
- Peubah dan Skala data disajikan pada Tabel 3.1

Tabel 3.1. Peubah dan Skala Data

Peubah	Skala Data
X1 : Jumlah Tenaga Kerja (orang)	Rasio
X2 : Nilai Produksi (rupiah/tahun)	Rasio
X3 : Jumlah Bahan Baku (kg/tahun)	Rasio
X4 : Nilai Investasi (rupiah/tahun)	Rasio
X5 : Teknologi yang Digunakan (1: Tradisional, 2: Menggunakan alat modern)	Nominal
X6 : Aneka Olahan (1: Hanya Produksi Tempe, 2: Juga Memproduksi Kripik Tempe)	Nominal

Selengkapnya data disajikan pada Lampiran 1.

3.2 Metode Analisis

- Langkah-langkah dalam analisis *cluster* adalah sebagai berikut :
 1. Pengkategorian data, pada data kontinyu diubah menjadi data kategori dengan bantuan selang kelas pada persamaan 2.1
 2. Melakukan transformasi menggunakan Analisis Komponen Utama *Nonlinier* pada peubah dengan skala nominal dan ordinal untuk mendapatkan skor komponen pada objek dengan skala interval atau rasio menggunakan persamaan 2.2 sampai dengan 2.7.
 3. Dilakukan analisis *cluster hirarki* terhadap semua data menggunakan jarak *Euclidean* pada persamaan 2.8 dan metode penggabungan menggunakan *average linkage* menggunakan persamaan 2.9.



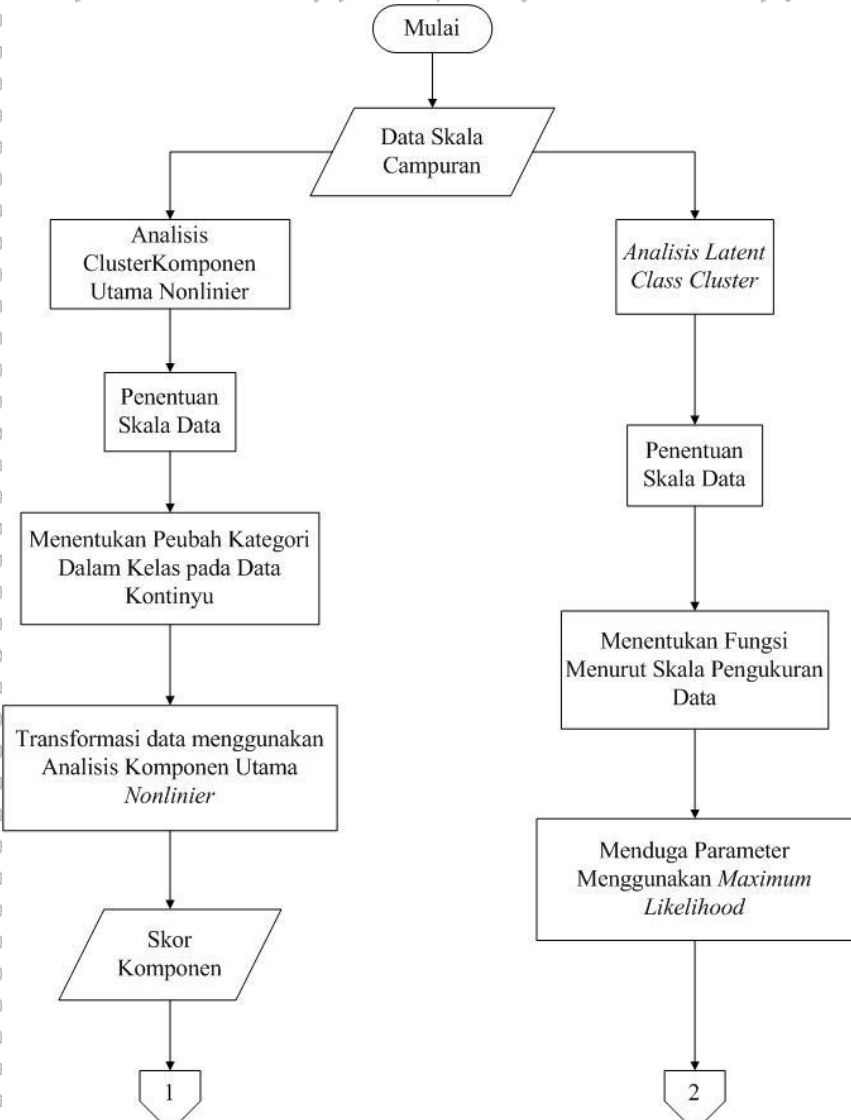
4. Melakukan uji validitas menggunakan Indeks *Dunn* pada persamaan 2.10, Indeks *DB* pada persamaan 2.11, dan Indeks *RMSSDT* pada persamaan 2.13.
5. Menentukan kelompok optimal pada setiap metode menggunakan Indeks Gabungan.
6. Menentukan *SSW* (*sum-of-square within cluster*) dengan persamaan 2.29 dan 2.31 dan *SSB* (*sum-of-square between cluster*) dengan persamaan 2.30 dan 2.32 pada semua data kontinu dan kategori.
7. Menentukan rasio keragaman dengan cara membagi nilai *SSW/SSB*.

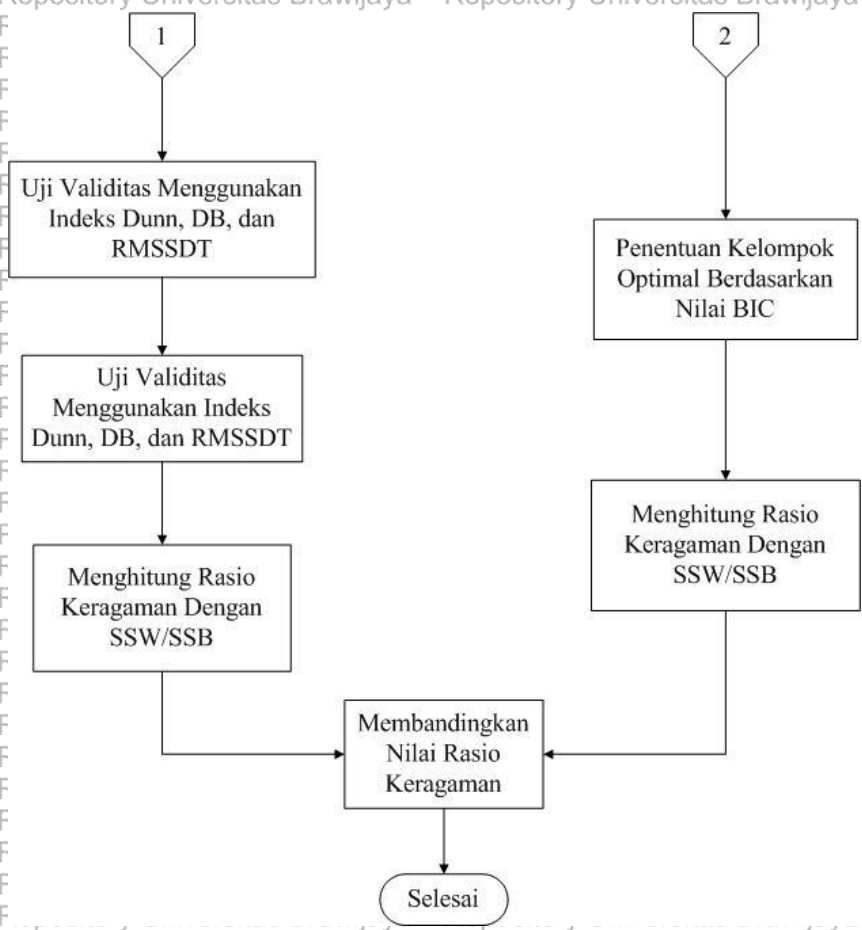
➤ Langkah-langkah dalam *Latent Class Cluster* adalah sebagai berikut :

1. Menentukan model *Latent Class Cluster* sesuai dengan skala dari data (nominal, ordinal, dan kontinu) menggunakan persamaan 2.14 sampai dengan 2.16.
2. Melakukan pendugaan parameter dari model dengan menggunakan metode *Maximum Likelihood* pada persamaan 2.17 sampai dengan 2.21 dengan menggunakan iterasi algoritma *EM* pada persamaan 2.22 sampai dengan 2.27.
3. Menentukan banyaknya kelompok yang terbentuk menggunakan nilai dari *BIC* dengan persamaan 2.28.
4. Menentukan *SSW* (*sum-of-square within cluster*) dengan persamaan 2.29 dan 2.31 dan *SSB* (*sum-of-square between cluster*) dengan persamaan 2.30 dan 2.32 pada semua data kontinu dan kategori.
5. Menentukan rasio keragaman dengan cara membagi nilai *SSW/SSB*.

➤ Membandingkan nilai rasio keragaman antara Analisis *Cluster* Komponen Utama *Nonlinier* dan *Latent Class Cluster*

Dari metode analisis di atas, tahapan analisis disajikan dalam diagram alir pada Gambar 3.1:





Gambar 3.1. Diagram Alir Proses Analisis *Cluster* Komponen Utama *Nonlinier* dan *Latent Class Cluster*



BAB IV
HASIL DAN PEMBAHASAN

4.1 Statistik Deskriptif

Hasil Statistik Deskriptif dari data Industri Tempe di Kota Malang tahun 2005 disajikan pada Tabel 4.1 dan 4.2 :

Tabel 4.1. Statistik Deskriptif Variabel Jumlah Tenaga Kerja, Nilai Produksi, Bahan Baku, dan Nilai Investasi.

Variabel	Total Data	Rata Rata	Ragam	Nilai Minimal	Nilai Maksimal
Jumlah Tenaga Kerja	102	2.598	0.6586	1	6
Nilai Produksi	102	230413456	5.32E+17	1168E+4	71175E+5
Bahan Baku	102	30960	6.05E+08	3650	155125
Nilai Investasi	102	3716304	3.33E+13	500000	53000000

Dari Tabel 4.1 dijelaskan total data, rata-rata, ragam, nilai minimal, dan nilai maksimal untuk setiap peubah kontinyu, hasil histogram bisa dilihat pada Lampiran 2. Diketahui bahwa setiap variabel kontinyu memiliki kisaran yang cukup besar, diindikasikan pada data tersebut memiliki pencilan. Pada Tabel 4.1 juga menjelaskan bahwa tidak adanya *missing* data sehingga jumlah objek setiap variabel tepat 102 dan kemudian mampu dianalisis menggunakan Analisis *Cluster* Komponen Utama *Nonlinier* maupun Analisis *Latent Class Cluster*.

Tabel 4.2. Tabel Frekuensi Variabel Teknologi dan Aneka Olahan

Variabel	Kategori	
	1	2
Teknologi	3	99
Aneka Olahan	70	32



Dari Tabel 4.2 diketahui bahwa pada data kategori pada variabel Teknologi, terdapat 3 Industri atau sekitar 2.9 % dari Industri Tempe di Kota Malang masih menggunakan teknologi tradisional untuk mengolah tempe, sedangkan terdapat 99 Industri atau sekitar 97.1% dari Industri Tempe di Kota Malang sudah menggunakan teknologi modern. Dijelaskan juga pada Tabel 4.2 untuk variabel Aneka Olahan, terdapat 70 Industri atau sekitar 68.6% dalam produksinya hanya memproduksi tempe, sedangkan terdapat 32 Industri atau 31.4% dalam produksinya selain memproduksi tempe juga memproduksi kripik tempe. Diagram kue bisa dilihat pada lampiran 3.

4.2 Analisis Cluster Komponen Utama Nonlinier

4.2.1 Analisis Komponen Utama Nonlinier

Pada Analisis Komponen Utama *Nonlinier*, data yang digunakan dalam analisis berupa skala kategori sehingga data dengan skala ordinal ataupun rasio harus dikategorikan terlebih dahulu. Pada data yang digunakan, terdapat 4 peubah yang memiliki skala ordinal atau rasio (X_1 , X_2 , X_3 , X_4) yang kemudian perlu dikategorikan, sehingga diperoleh pengkategorian menurut selang kelas, untuk hasil selengkapnya dapat dilihat pada Lampiran 4.

Pada penelitian kali ini, untuk menghitung Analisis Komponen Utama *Nonlinier* menggunakan *software* SPSS dengan menggunakan *syntax* SPSS pada Lampiran 5. Dijelaskan pada tinjauan pustaka bahwa nilai yang dihasilkan dari Analisis Komponen Utama *Nonlinier* berupa skor komponen objek yang memiliki skala rasio. Hasil Analisis Komponen Utama *Nonlinier* berupa skor komponen objek dapat dilihat pada Lampiran 6.

4.2.2 Analisis Cluster Hirarki dan Validitas Cluster

Pada Analisis *Cluster* Hirarki, data yang digunakan adalah hasil dari Analisis Komponen Utama *Nonlinier* berupa skor komponen objek. Jarak yang digunakan pada Analisis *Cluster* Hirarki menggunakan jarak *Eucliden* dengan menerapkan metode pautan rata-rata (*Average Linkage*).

Cluster yang dihasilkan akan dicari yang paling optimal pada setiap *cluster*. Kriteria *cluster* yang optimal menggunakan 3 Indeks Validitas, yaitu Indeks *Dunn*, Indeks *Davies Bouldin*, dan Indeks *RMSSDT* dengan indikasi yang berbeda pada setiap Indeks.

Berikut merupakan hasil untuk penentuan jumlah *cluster* yang optimal berdasarkan indeks validitas dapat dilihat pada Tabel 4.3:

Tabel 4.3. Hasil Indeks *Dunn*, DB, dan RMSSDT

Jumlah Cluster	RMSSTD	DB	Dunn
10	63.626	0.483	0.251
9	89.604	0.466	0.216
8	3.564	0.521	0.216
7	93.58	0.433	0.387
6	94.576	0.424	0.387
5	96.57	0.312	0.719
4	97.575	0.302	0.723
3	100.517	0.228	0.679
2	101.522	0.236	0.608

Pada Indeks *Dunn* kriteria yang digunakan adalah semakin besar nilai Indeks diindikasikan terdapat hasil *cluster* yang lebih baik, diperoleh hasil pada jumlah *cluster* 4. Pada Indeks DB kriteria yang digunakan adalah semakin kecil nilai Indeks semakin baik kelompok yang terbentuk, diperoleh hasil pada jumlah *cluster* 3. Pada Indeks RMSSDT kriteria yang digunakan adalah semakin kecil nilai Indeks semakin baik kelompok yang dihasilkan, diperoleh hasil pada jumlah *cluster* 8.

Pada Tabel 4.3, antara Indeks *Dunn*, Indeks DB, dan Indeks RMSSDT menghasilkan jumlah *cluster* optimal yang berbeda. Oleh karena itu diperlukan nilai gabungan yang digunakan untuk memutuskan jumlah *cluster* yang optimal. Digunakanlah Indeks gabungan, setiap Indeks diperingkat sesuai dengan kriteria masing-masing Indeks dengan urutan peringkat mulai dari 1, 2, dan seterusnya. Pada Indeks gabungan diperoleh dari jumlah peringkat setiap *cluster* pada 3 Indeks yang digunakan, sehingga semakin kecil Indeks gabungan maka diperoleh jumlah *cluster* yang optimal.

Hasil peringkat Indeks gabungan dapat dilihat pada Tabel 4.4 berikut :

Tabel 4.4. Hasil peringkat Indeks Gabungan

Jumlah Cluster	RMSSTD	Rank	DB	Rank	Dunn	Rank	Indeks Gabungan
10	63.626	2	0.483	8	0.251	7	17
9	89.604	3	0.466	7	0.216	8	18
8	3.564	1	0.521	9	0.216	8	18
7	93.58	4	0.433	6	0.387	5	15
6	94.576	5	0.424	5	0.387	5	15
5	96.57	6	0.312	4	0.719	2	12
4	97.575	7	0.302	3	0.723	1	11
3	100.517	8	0.228	1	0.679	3	12
2	101.522	9	0.236	2	0.608	4	15

Pada Tabel 4.4 menunjukkan bahwa jumlah *cluster* yang optimal berada pada 4 *cluster* berdasarkan Indeks gabungan paling minimal. Keanggotaan 4 *cluster* disajikan pada Tabel 4.5 :

Tabel 4.5. Keanggotaan 4 Cluster

Kelompok	1	2	3	4
Nomor Objek	66	67	14,50,63,74,90	1,2,3,4,...,99,100,101,102 (Kecuali objek pada kelompok 1,2, dan 3)

Hasil keanggotaan kelompok bisa dilihat pada Lampiran 9.

4.3 Hasil Analisis *Latent Class Cluster*

4.3.1 Penentuan Model Terbaik

Pada Analisis *Latent Class Cluster* dilakukan dengan menggunakan *software* XL Stat. Model *Latent Class Cluster* yang terbentuk dari data Industri tempe di Kota Malang tahun 2005 disajikan pada Tabel 4.6. Tujuan analisis *Latent Class Cluster* untuk mengidentifikasi banyaknya *cluster* yang dibutuhkan untuk menjelaskan hubungan diantara variabel. Banyaknya *cluster* yang terbentuk dilihat dari nilai BIC model, semakin kecil nilai BIC diindikasikan model tersebut menghasilkan kelompok yang paling optimal.

Tabel 4.6. Model *Latent Class Cluster*

Jumlah Cluster	LL	BIC(LL)	Jumlah Parameter
2	-4936.855	9970.835	21
3	-4831.565	9811.129	32
4	-4835.644	9870.161	43
5	-4694.194	9638.136	54
6	-4670.993	9642.609	65
7	-4650.241	9651.979	76
8	-4632.482	9667.337	87
9	-4622.708	9698.663	98
10	-4604.406	9712.934	109

Berdasarkan Tabel 4.6 dapat dilihat bahwa nilai BIC paling kecil berada pada *cluster* 5, sehingga diperoleh untuk Analisis *Latent Class Cluster* menghasilkan kelompok yang optimal sebanyak 5 *cluster*. Hasil keanggotaan selengkapnya bisa dilihat pada Lampiran 11. Keanggotaan 5 *cluster* disajikan pada Tabel 4.7 :

Tabel 4.7. Keanggotaan 5 *Cluster*

Kelompok	1	2	3	4	5
Nomor Objek	8,9,12,...., 98,100,101	1,5,6,10, 11,13,15,18, 19,21,23,27, 29,33,35,37, 42,46,51,54, 55,76,77,79, 82,87,88, 94,95,99,102	7,14,16,36, 44,50,57,63, 64,69,75, 80,90	2,3,4, 43,74	66,67, 72,73
	(Kecuali objek pada kelompok 2,3,4, dan 5)				

4.3.2 Pendugaan Parameter

Pendugaan parameter yang digunakan pada Analisis *Latent Class Cluster* adalah *Maximum Likelihood Estimation* (MLE). Penduga MLE

diperoleh melalui fungsi *log-likelihood* yang dimaksimumkan menggunakan metode iteratif yaitu algoritma *Expectation-Maximization* (EM).

Tabel 4.8. Pendugaan Paramater

Iterasi EM	15
Log-posterior	-4825.665
L^2	9390.359
Nilai Konvergen	0.010

Besarnya nilai L^2 menunjukkan bahwa jumlah hubungan ketidakbebasan yang ada pada data yang mampu dijelaskan dengan menggunakan 5 *cluster*. Berdasarkan Tabel 4.8 pendugaan menggunakan algoritma EM dilakukan sebanyak 15 iterasi kemudian diperoleh nilai L^2 sebesar 9390.359 dengan nilai konvergen 0.010.

4.4 Rasio Keragaman

Suatu *Cluster* dikatakan baik apabila memiliki kehomogenan yang tinggi dalam suatu *cluster* dan kehomogenan yang rendah antar *cluster*. Homogenitas dalam *cluster* dapat dihitung dengan SSW (*Sum of Square Within Cluster*) dan homogenitas antar *cluster* dapat dihitung menggunakan SSB (*Sum of Square Between Cluster*).

Dalam penelitian kali ini terdapat peubah kontinyu dan peubah kategori yang kemudian dalam penghitungan keragaman juga akan berbeda. Nilai SSW dan SSB untuk peubah kontinyu akan ditunjukkan menggunakan persamaan 2.31 dan 2.32, sedangkan nilai SSW dan SSB untuk peubah kategori dihitung menggunakan persamaan 2.33 dan 2.34.

Hasil perhitungan SSW dan SSB pada peubah kontinyu dan kategori bisa dilihat pada Tabel 4.9 berikut :



Tabel 4.9 Rasio Keragaman

Objek	Jumlah Variabel		Metode			
	Kontinyu	Kategori	Analisis <i>Cluster</i> Komponen Utama <i>Nonlinier</i>		Latent Class Cluster	
			Jumlah Kelompok	Rata-Rata Nilai SSW/SSB	Jumlah Kelompok	Rata-Rata Nilai SSW/SSB
102	4	2	4	1.20009	5	1.60915

Pada Tabel 4.9 diketahui rasio keragaman antara kedua metode, Analisis *Cluster* Komponen Utama *Nonlinier* dan Metode *Latent Class Cluster*. Dari hasil Tabel 4.9 diperoleh rasio keragaman Analisis *Cluster* Komponen Utama *Nonlinier* sebesar 1,20009 dan metode *Latent Class Cluster* memiliki rasio keragaman sebesar 1,60915 sehingga diketahui bahwa Analisis *Cluster* Komponen Utama *Nonlinier* memiliki nilai rasio keragaman yang lebih minimum dan diindikasikan merupakan metode terbaik dalam menangani data campuran dibandingkan dengan metode *Latent Class Cluster*.

4.5 Interpretasi

Diperoleh metode terbaik yaitu Analisis *Cluster* Komponen Utama *Nonlinier* dengan anggota secara lengkap terdapat pada Lampiran 9. Analisis deskriptif untuk 4 *cluster* disajikan pada Tabel 4.10 :

Tabel 4.10 Hasil Analisis Deskriptif Setiap *Cluster*

Kelompok	X1	X2	X3	X4	X5	X6
<i>Cluster 1</i>	5	7117500000	155125	12000000	2	1
<i>Cluster 2</i>	3	2263000000	49275	4800000	2	1
<i>Cluster 3</i>	4.2	503555000	79935	6732000	2	1,4
<i>Cluster 4</i>	2.48	122146289.5	26882.21	3458979	1.97	1.32

Tabel 4.10 menunjukkan rata-rata dari setiap variabel pada masing-masing *cluster*. Pada kelompok 1 yang beranggotakan Bapak Suwaji memiliki nilai yang sangat tinggi dan tertinggi pada setiap



variabel, baik dari jumlah tenaga kerja, nilai produksi, jumlah bahan baku dan bahkan nilai investasi memiliki perbedaan yang sangat tinggi dengan kelompok yang lain, pada industri Pak Suwaji ini sudah menggunakan alat modern dalam mengolah namun dalam produksinya hanya memproduksi tempe saja, dengan dibantu sekitar 5 tenaga kerja, industri Pak Suwaji ini menghasilkan hasil yang sangat tinggi dibandingkan dengan industri yang lainnya.

Pada kelompok 2 yang beranggotakan Bapak Sutrisno memiliki nilai yang tinggi dibandingkan dengan kelompok yang lain namun masih lebih rendah dibandingkan dengan kelompok 1 dengan Bapak Suwaji. Pada industri milik Bapak Sutrisno memiliki hasil yang cukup tinggi pada nilai produksi namun memiliki nilai yang cukup rendah pada jumlah bahan baku dan nilai investasi. Bapak Sutrisno dalam industrinya sudah menggunakan alat modern namun dalam pengolahannya hanya memproduksi tempe saja dengan hanya dibantu 3 tenaga kerja menghasilkan nilai produksi yang tinggi dan jumlah bahan baku dan nilai investasi yang rendah.

Kelompok 3 memiliki jumlah anggota sebanyak 5 industri tempe (anggota bisa dilihat pada Lampiran 9) dengan nilai sekitar rata-rata karena masih berada dibawah dari kelompok 1 namun beberapa nilai lebih besar daripada kelompok 2. Industri pada kelompok ini memiliki nilai produksi yang cukup rendah dibandingkan dengan kelompok yang lain namun memiliki jumlah bahan baku dan nilai investasi yang cukup tinggi, hanya lebih rendah dibandingkan dengan kelompok 1. Pada kelompok 3 ini merupakan kelompok yang cukup ideal karena dengan nilai produksi yang cukup rendah menghasilkan jumlah bahan baku dan nilai investasi yang cukup tinggi dengan dibantu sekitar 4 tenaga kerja. Industri pada kelompok 3 ini sudah menggunakan alat produksi yang modern dengan produk yang dihasilkan terdapat beberapa industri yang hanya produksi tempe saja dan juga beberapa industri yang memproduksi kripik tempe.

Kelompok 5 beranggotakan 95 industri yang tersebar di Kota Malang dengan nilai yang rendah pada seluruh variabel dibandingkan dengan kelompok yang lain, baik dari nilai produksi, jumlah bahan baku, dan nilai investasi memiliki perbedaan yang cukup signifikan dibandingkan dengan kelompok lain. Pada kelompok 5 ini dalam proses produksinya sudah menggunakan alat modern walaupun terdapat beberapa industri yang masih menggunakan teknologi tradisional dalam

proses produksinya, selain itu produk yang dihasilkan pada kelompok 5 ini terdapat beberapa industri yang hanya produksi tempe saja dan juga beberapa industri yang memproduksi kripik tempe. Secara umum kelompok inilah industri tempe yang paling banyak berada di Kota Malang namun memiliki nilai yang rendah dibandingkan dengan kelompok yang lain.



BAB V KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil dan pembahasan, diperoleh kesimpulan sebagai berikut :

Pada Analisis *Cluster* Komponen Utama *Nonlinier* terbentuk sebanyak 4 kelompok, sedangkan menggunakan *Latent Class Cluster* terbentuk sebanyak 5 kelompok. Dalam proses perbandingannya, Analisis *Cluster* Komponen Utama *Nonlinier* memiliki nilai rasio keragaman yang lebih kecil (1.20009) dibandingkan dengan nilai keragaman dari *Latent Class Cluster* (1.60915), sehingga dapat disimpulkan bahwa Analisis *Cluster* Komponen Utama *Nonlinier* mampu menghasilkan proses pengelompokan yang lebih baik dibandingkan dengan *Latent Class Cluster* pada data berskala campuran dalam kasus Industri Tempe di Kota Malang Tahun 2005.

5.2 Saran

Pengelompokan pada data berskala campuran (kategori dan kontinyu) didapatkan hasil yang lebih optimal dan disarankan menggunakan Analisis *Cluster* Komponen Utama *Nonlinier*. Untuk penelitian selanjutnya lebih mengkaji dalam topik membandingkan berbagai macam Indeks validitas yang digunakan untuk menentukan jumlah *cluster* karena masih banyak indeks validitas yang lain dan ditentukan indeks mana yang paling baik untuk menghasilkan kelompok yang paling optimal.

DAFTAR PUSTAKA

- Bolshakova, N. dan Azuaje, F. dan. 2001. *Improving Expression Data Mining through Cluster Validity*. Departement of Computer Science. Trinity College Dublin. Ireland. <https://www.cs.tcd.ie/publications/tech-reports/reports.02/TCD-CS-2002-34.pdf>. Diakses tanggal 26 Februari 2015.
- Gifi, A. 1989. *Algorithm Descriptions for Anchor, Homals, Princals, and Overal*.
- Gifi, A. 1990. *Nonlinier Multivariat Analysis*. John Wiley & Sons, New York.
- Haerudiin, 2013. *Latent Class Regression Analysis Untuk Data kategorik Dengan Satu Kovariat*. (<http://repository.unej.ac.id/bitstream/handle/123456789/6277/Haeruddin%20-%20081810101052.pdf?sequence=1>). Diakses tanggal 12 Maret 2015.
- Hair, J.F.Jr., Anderson, R.E., Tatham, R.I. dan Black, W.C.. 1998. *Multivariate Data Analysis*. Fifth Ed., Prentice Hall International, Inc. New Jersey.
- Indriani, Y.N. 2013. *Perbandingan Jumlah Kelompok Optimal Pada Metode Single Linkage Dan Complete Linkage Dengan Indeks Validitas Silhouette*. <http://jurnal-online.um.ac.id/data/artikel/artikelAC484AF441FECB1CF2BA7AE839AF053B.pdf>. Diakses tanggal 12 Maret 2015.
- Jhonson, R.A dan Wichren, D.W. 2007. *Applied Multivariate Statistical Analysis Fifth Ed*. Prentice Hall International, Inc. New Jersey.
- Jolliffe, I.T. 2002. *Principal Component Analysis*. Springer-Verlag, Inc. New York.



Konig, R. 2002. *On The Rotation of Non-Linear Principal Component Analysis (PRINCALS) Solution : Description of a Procedur.* http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_50.pdf. Diakses tanggal 24 Februari 2015.

Kovacs, F., Legany, C., dan Babos, A. 2005. *Cluster Validity Measurement Techniques.* *Jurnal 6th International symposium of hungarian researchers on computational intelligence.* <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.2.848&rep=rep1&type=pdf>. Diakses tanggal 26 Februari 2015.

Kroonenberg, P.M., Harch, B.D., Basford, K.E., dan Cruickshank, A. 1997. *Combine Analysis Of Categorical And Numerical Description Of Australian Groundnut Accessions Using Nonlinear Principal Component Analysis.* https://openaccess.leidenuniv.nl/bitstream/handle/1887/11597/7_702_078.pdf?sequence=1. Diakses tanggal 26 Februari 2015.

Vermunt, J.K. dan Magidson, J. 2004. *Latent Class Model.* The Sage Handbook Of Quantitative Methodology For The Social Sciences. Thousands Oakes. Sage.

Mattjik, A., Sumertajaya, I.M., Wijayanto, H., Indahwati, Kurnia, A., dan Sartono, B. 2002. *Aplikasi Analisis Peubah Ganda.* Depdiknas. Bogor.

Moustaki, I. dan Papageorgiou, I. 2004. *Latent Class Model For Mixed Variabel With Applications In Archaeometry.* *Journal of Computational Statistics and Data Analysis.*

Sitepu, R., Irmeilyana, dan B. Gultom. 2011. " *Analisis Cluster Terhadap Tingkat Pencemaran Udara pada Sektor Industri di Sumatera Selatan,* " *Jurnal* dari Universitas Sriwijaya. <http://jpsmpaunsri.files.wordpress.com/2011/11/v14-no3-a-3-sitepu-11-17.pdf>. Diakses tanggal 25 Februari 2015.



Siegel, S. 1997. *Statistika-Nonparametrik*. Gramedia Pustaka Utama. Jakarta.

Umah, F. 2014. *Analisis Two Step Cluster (TSC) Dan Analisis Latent Class Cluster (LCC) Pada Pengelompokan Data Berekala Campuran Kategorik Dan Kontinu*. Skripsi. Jurusan Matematika Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Brawijaya, Malang.

Wulandari, T. 2006. *Penerapan Analisis Cluster Hirarki Pada Data Berekala Campuran dengan Menggunakan Analisis Komponen Utama Nonlinier*. Skripsi. Jurusan Matematika Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Brawijaya, Malang.

LAMPIRAN

Lampiran I. Data Industri Tempe di Kota Malang Tahun 2005

No	Nama	X1	X2	X3	X4	X5	X6
1	Kasian	3	82125000	18250	2150000	2	2
2	Samuki	4	142350000	31025	1750000	1	1
3	Supangat	4	65700000	14600	1000000	1	1
4	Mani	3	25185000	5475	500000	1	1
5	Sulkah	3	67160000	14600	1950000	2	1
6	Rasemo	3	293825000	63875	2649000	2	1
7	Sekak	4	226665000	49275	2371000	2	1
8	Slamet	2	98550000	21900	2000000	2	1
9	Ngatmani	2	24820000	5475	1650000	2	1
10	Reban	3	131400000	29200	1825000	2	1
11	Tawi	3	67160000	14600	1950000	2	1
12	Syukur	2	32850000	7300	1775000	2	1
13	Mulyono	3	50370000	10950	1900000	2	2
14	Ridwan	6	503700000	109500	8480000	2	1
15	Mufid	3	82125000	18250	2200000	2	1
16	Biono	4	251850000	54750	5235000	2	1
17	Mustofa	2	16425000	3650	1600000	2	1
18	Mustaqim	3	32850000	7300	1825000	2	1
19	Sumari	3	183960000	40150	2190000	2	2
20	Badrus	2	11680000	25550	2450000	2	2
21	Iwan H.	3	254040000	54750	2880000	2	1
22	M. Ainur	2	83950000	18250	2200000	2	2
23	Sugianto	3	167535000	36500	2894000	2	1
24	Syamsul A.	2	118260000	25550	2500000	2	2
25	Machmud	2	41062500	9125	1925000	2	1
26	Juwanto	2	118260000	25550	2400000	2	1
27	Suwandi	3	83950000	18250	2000000	2	2
28	Rochmad N.	2	100375000	21900	2320000	2	1
29	A. Soleh	3	83950000	18250	1800000	2	1
30	Umar Fauzi	2	117530000	25550	2400000	2	1
31	Sumiati	2	65700000	14600	2776000	2	2
32	Senu	2	41062500	9125	1338000	2	1
33	Chamdani	3	117895000	25550	3340000	2	1
34	Moh. Sani	2	65700000	14600	1850000	2	1

Lampiran 1. (Lanjutan)

No	Nama	X1	X2	X3	X4	X5	X6
35	Ainul M.	3	205860000	45625	2258000	2	2
36	Kukuh K.	3	335800000	73000	7690000	2	2
37	Sunyoto	3	167535000	36500	2652000	2	2
38	Wasir	2	167900000	36500	2396000	2	2
39	Ivan K.	2	74460000	16425	1840000	2	2
40	Matali	2	67160000	14600	1930000	2	2
41	Junaidi	2	100740000	21900	2240000	2	2
42	Haryono	3	83950000	18250	2300000	2	2
43	Djari	1	33215000	7300	1338000	2	2
44	Ketang G.	4	335800000	73000	6950000	2	2
45	Sanusi	2	50370000	10950	2225000	2	2
46	Winanik	3	186150000	40150	3360000	2	2
47	Syafi'i	2	67160000	14600	1925000	2	2
48	Syafi I A.	2	50187500	10950	2250000	2	2
49	Churoti	2	111690000	24455	2868000	2	2
50	Yusuf	4	419750000	91250	7555000	2	2
51	Nanto	3	209875000	45625	4750000	2	2
52	Abdul M.	2	75555000	16425	2276000	2	2
53	Agus H.	2	142350000	31025	2300000	2	2
54	Rifkik	3	201480000	43800	2802000	2	2
55	Achadun	3	109500000	23725	2868000	2	2
56	H. Asy'Ari	2	117530000	25550	2675000	2	2
57	Chamim	3	337260000	73000	7100000	2	2
58	Anas M.	2	124830000	27375	2960000	2	2
59	M. Jalal	2	58400000	12775	2230000	2	2
60	M. Anshori	2	58765000	12775	2190000	2	2
61	Asmu'i	2	83950000	18250	2322000	2	2
62	R. Djumadi	2	142350000	32025	2780000	2	2
63	Achmad S.	2	419750000	91250	7100000	2	2
64	Imam C.	4	284700000	62050	5600000	2	2
65	Suyanto	2	33215000	7300	2568000	2	2
66	Suwaji	5	7117500000	155125	12000000	2	2
67	Sutrisno	3	2263000000	49275	4800000	2	2
68	Didik Karim	2	50370000	10950	1800000	2	2
69	Buadi	2	271560000	58400	7010000	2	2
70	Anshori	2	41975000	9125	2000000	2	2
71	Erfin E.	2	58765000	12775	2200000	2	2

Lampiran 1. (Lanjutan)

No	Nama	X1	X2	X3	X4	X5	X6
72	Dianur	2	251850000	54750	53000000	2	2
73	Usman	3	124830000	27375	27000000	2	1
73	Usman	3	124830000	27375	27000000	2	1
74	Syafi'i I.	4	755555000	16425	2275000	2	1
75	Mustakin	2	250390000	54750	5650000	2	2
76	Kasiono	3	167900000	36500	2450000	2	1
77	Soni	3	32850000	7300	1850000	2	1
78	Yoyok	2	41975000	9125	2150000	2	2
79	Suruli	3	98550000	21900	2300000	2	1
80	Nanda	2	201480000	43800	4500000	2	1
81	Isman	2	49275000	10950	2275000	2	1
82	Abdul K.	3	49640000	10950	2345000	2	1
83	Ridho I	2	50370000	10950	1900000	2	1
84	Mariono	2	142350000	31025	2875000	2	2
85	Royan	2	168630000	36500	2001000	2	1
86	Jasmani	2	149650000	32850	2105000	2	1
87	Sholeh	3	134320000	29200	2820000	2	1
88	Hadi S.	3	167900000	36500	2500000	2	1
89	Siti Aisyiyah	2	67160000	14600	2653000	2	1
90	A. Suedi	5	419020000	91250	8250000	2	2
91	Agus Salim	2	92345000	20075	2430000	2	1
92	Romeli	2	210240000	45625	2588000	2	1
93	Kaseri	2	144540000	31390	2920000	2	2
94	Dachroni	3	208780000	45625	4750000	2	2
95	Supardi	3	226300000	49275	3170000	2	1
96	Thoyib	2	83950000	18250	2500000	2	1
97	Subakri	2	127020000	27740	2700000	2	2
98	Satik	2	167900000	36500	2450000	2	1
99	Sujiono	3	100740000	21900	2625000	2	1
100	Gimin	2	98550000	21900	2250000	2	1
101	Soli	2	125560000	27375	2825000	2	1
102	Solikin	3	82125000	18250	2775000	2	1

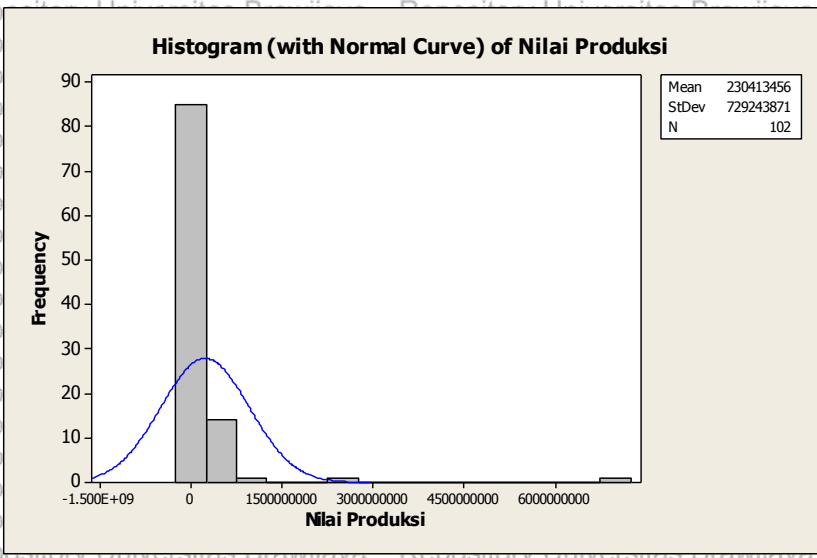
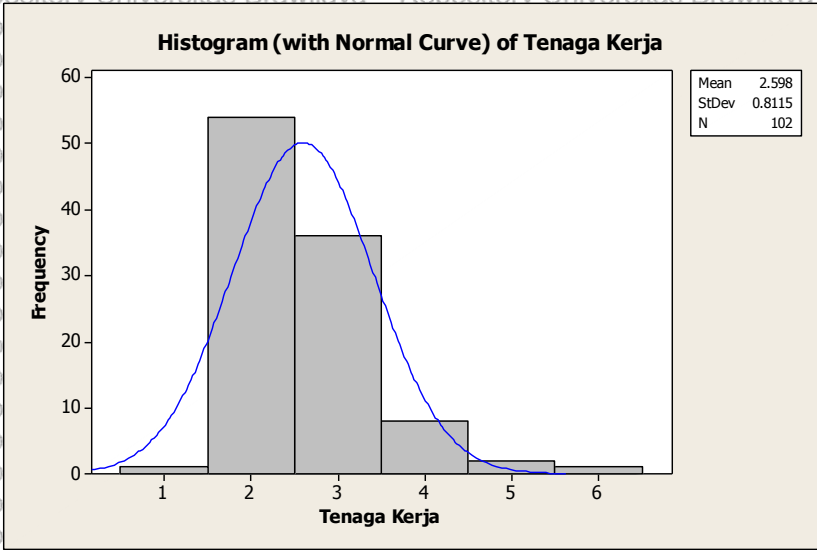


Keterangan :

- X1 : Jumlah Tenaga Kerja (orang)
- X2 : Nilai Produksi (rupiah/tahun)
- X3 : Jumlah Bahan Baku (kg/tahun)
- X4 : Nilai Investasi (rupiah/tahun)
- X5 : Teknologi yang Digunakan (1: Tradisional ; 2: Alat Modern)
- X6 : Aneka Olahan (1: Hanya Produksi Tempe ; 2: Juga Memproduksi Kripik Tempe)

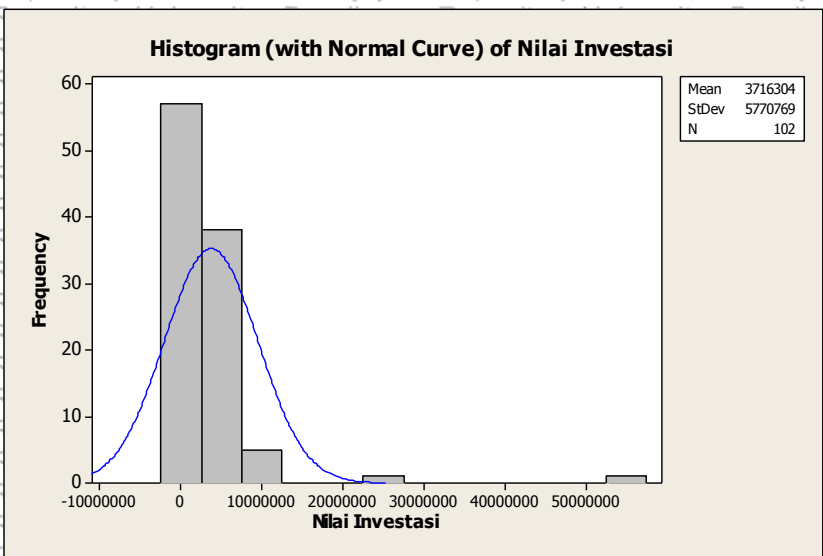
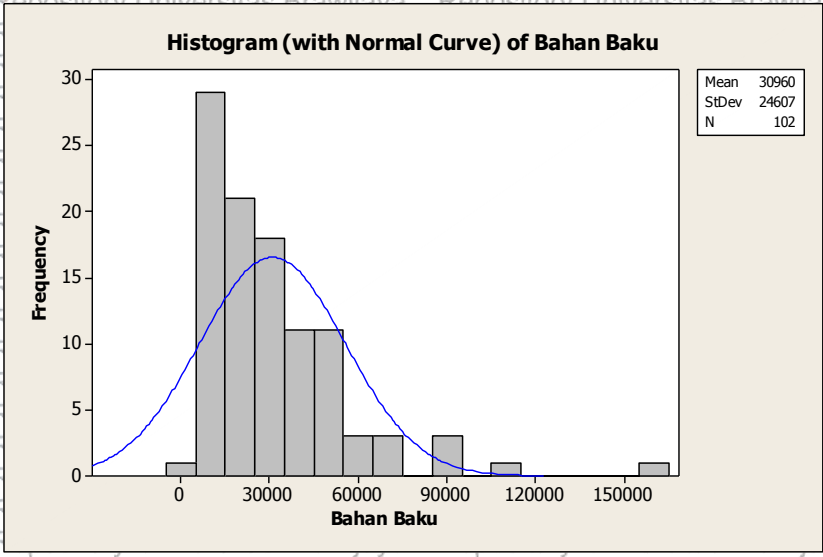


Lampiran 2. Histogram Variabel Tenaga Kerja, Nilai Produksi, Bahan Baku, dan Nilai Investasi.



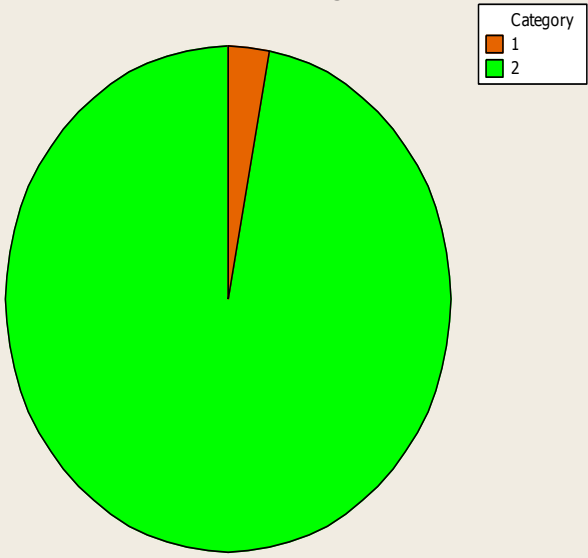


Lampiran 2. (Lanjutan)

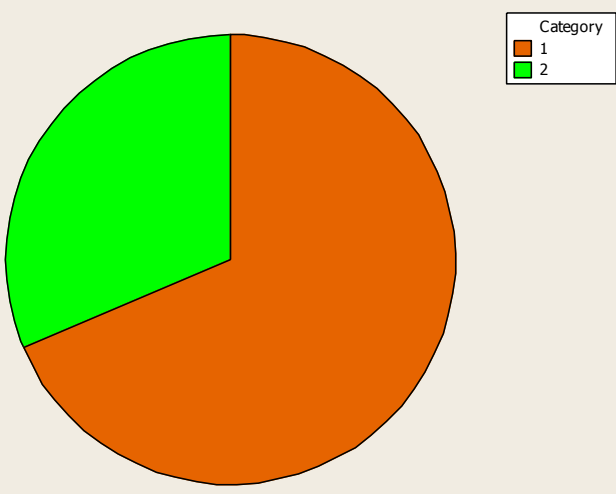


Lampiran 3. *Pie Chart* Variabel Teknologi dan Aneka Olahraga

Pie Chart of Teknologi



Pie Chart of Aneka Olahraga



Lampiran 4. Kategori Peubah Selang Kelas

No	Nama	X1	X2	X3	X4	X5	X6
1	Kasian	2	1	1	1	2	2
2	Samuki	2	1	2	1	1	1
3	Supangat	2	1	1	1	1	1
4	Mani	2	1	1	1	1	1
5	Sulkah	2	1	1	1	2	1
6	Rasemo	2	1	3	1	2	1
7	Sekak	2	1	3	1	2	1
8	Slamet	1	1	1	1	2	1
9	Ngatmani	1	1	1	1	2	1
10	Reban	2	1	2	1	2	1
11	Tawi	2	1	1	1	2	1
12	Syukur	1	1	1	1	2	1
13	Mulyono	2	1	1	1	2	2
14	Ridwan	4	1	5	2	2	1
15	Mufid	2	1	1	1	2	1
16	Biono	2	1	3	1	2	1
17	Mustofa	1	1	1	1	2	1
18	Mustaqim	2	1	1	1	2	1
19	Sumari	2	1	2	1	2	2
20	Badrus	1	1	2	1	2	2
21	Iwan H.	2	1	3	1	2	1
22	M. Ainur	1	1	1	1	2	2
23	Sugianto	2	1	2	1	2	1
24	Syamsul A.	1	1	2	1	2	2
25	Machmud	1	1	1	1	2	1
26	Juwanto	1	1	2	1	2	1
27	Suwandi	2	1	1	1	2	2
28	Rochmad N.	1	1	1	1	2	1
29	A. Soleh	2	1	1	1	2	1
30	Umar Fauzi	1	1	2	1	2	1
31	Sumiati	1	1	1	1	2	2

Lampiran 4. (Lanjutan)

No	Nama	X1	X2	X3	X4	X5	X6
32	Senu	1	1	1	1	2	1
33	Chamdani	2	1	2	1	2	1
34	Moh. Sani	1	1	1	1	2	1
35	Ainul M.	2	1	2	1	2	2
36	Kukuh K.	2	1	4	1	2	1
37	Sunyoto	2	1	2	1	2	1
38	Wasir	1	1	2	1	2	1
39	Ivan K.	1	1	1	1	2	2
40	Matali	1	1	1	1	2	1
41	Junaidi	1	1	1	1	2	2
42	Haryono	2	1	1	1	2	2
43	Djari	1	1	1	1	2	1
44	Ketang G.	2	1	4	1	2	1
45	Sanusi	1	1	1	1	2	2
46	Winanik	2	1	2	1	2	1
47	Syafi'i	1	1	1	1	2	1
48	Syafi' A.	1	1	1	1	2	2
49	Churoti	1	1	1	1	2	1
50	Yusuf	2	1	5	1	2	1
51	Nanto	2	1	2	1	2	2
52	Abdul M.	1	1	1	1	2	1
53	Agus H.	1	1	2	1	2	1
54	Rifkik	2	1	2	1	2	2
55	Achadun	2	1	1	1	2	1
56	H. Asy'Ari	1	1	2	1	2	2
57	Chamim	2	1	4	1	2	2
58	Anas M.	1	1	2	1	2	1
59	M. Jalal	1	1	1	1	2	1
60	M. Anshori	1	1	1	1	2	1
61	Asmu'i	1	1	1	1	2	2
62	R. Djumadi	1	1	2	1	2	1
63	Achmad S.	1	1	5	1	2	2
64	Imam C	2	1	3	1	2	1

Lampiran 4. (Lanjutan)

No	Nama	X1	X2	X3	X4	X5	X6
65	Suyanto	1	1	1	1	2	1
66	Suwaji	3	7	7	2	2	1
67	Sutrisno	2	3	3	1	2	1
68	Didik Karim	1	1	1	1	2	2
69	Buadi	1	1	3	1	2	2
70	Anshori	1	1	1	1	2	2
71	Erfm E.	1	1	1	1	2	2
72	Dianur	1	1	3	7	2	2
73	Usman	2	1	2	4	2	1
74	Syafi' I.	2	1	1	1	2	1
75	Mustakin	1	1	3	1	2	2
76	Kasiono	2	1	2	1	2	1
77	Soni	2	1	1	1	2	1
78	Yoyok	1	1	1	1	2	2
79	Suruli	2	1	1	1	2	1
80	Nanda	1	1	2	1	2	1
81	Isman	1	1	1	1	2	1
82	Abdul K.	2	1	1	1	2	1
83	Ridho I	1	1	1	1	2	1
84	Mariono	1	1	2	1	2	2
85	Royan	1	1	2	1	2	1
86	Jasmani	1	1	2	1	2	1
87	Sholeh	2	1	2	1	2	1
88	Hadi S.	2	1	2	1	2	1
89	Siti Aisyiyah	1	1	1	1	2	1
90	A. Suedi	3	1	5	2	2	2
91	Agus Salim	1	1	1	1	2	1
92	Romeli	1	1	2	1	2	1
93	Kaseri	1	1	2	1	2	2
94	Dachroni	2	1	2	1	2	2
95	Supardi	2	1	3	1	2	1
96	Thoyib	1	1	1	1	2	1
97	Subakri	1	1	2	1	2	2

Lampiran 4. (Lanjutan)

No	Nama	X1	X2	X3	X4	X5	X6
98	Satik	1	1	2	1	2	1
99	Sujiono	2	1	1	1	2	1
100	Gimin	1	1	1	1	2	1
101	Soli	1	1	2	1	2	1
102	Solikin	2	1	1	1	2	1

Lampiran 5. *Syntax* SPSS Komponen Utama *Nonlinier*

PRINCALS VARIABLES = X1(4) X2 X3 X4 (7) X5 X6 (2)

/ANALYSIS= X5 X6 (SNOM)

X1 X2 X3 X4 (NUME)

/DIMENSION=6

/MAXITER=100

/CONVERGENCE=.00001

/PRINT=FREQ EIGEN LOADINGS QUANT HISTORY OBJECT

/PRINT=EIGEN LOADINGS HISTORY OBJECT

/PLOT=NONE

/SAVE=DIM(6)

Lampiran 6. Skor Komponen Objek

No	Z1	Z2	Z3	Z4	Z5	Z6
1	-0.18	-0.57	0.49	1.23	-0.6	-1.81
2	0.44	3.65	3.83	1.18	1.52	1.3
3	0.05	3.77	3.89	1.03	1.63	0.27
4	0.05	3.77	3.89	1.03	1.63	0.27
5	-0.05	0.67	-0.3	-0.34	-0.74	-1.28
6	0.74	0.43	-0.42	-0.04	-0.97	0.79
7	0.74	0.43	-0.42	-0.04	-0.97	0.79
8	-0.69	0.25	-0.46	-0.73	0.48	0.02
9	-0.69	0.25	-0.46	-0.73	0.48	0.02
10	0.34	0.55	-0.36	-0.19	-0.86	-0.25
11	-0.05	0.67	-0.3	-0.34	-0.74	-1.28
12	-0.69	0.25	-0.46	-0.73	0.48	0.02
13	-0.18	-0.57	0.49	1.23	-0.6	-1.81
14	3.11	0.49	0.63	-0.06	-3.61	-0.02
15	-0.05	0.67	-0.3	-0.34	-0.74	-1.28
16	0.74	0.43	-0.42	-0.04	-0.97	0.79
17	-0.69	0.25	-0.46	-0.73	0.48	0.02
18	-0.05	0.67	-0.3	-0.34	-0.74	-1.28
19	0.22	-0.69	0.43	1.38	-0.72	-0.77
20	-0.43	-1.11	0.26	0.99	0.5	0.53
21	0.74	0.43	-0.42	-0.04	-0.97	0.79
22	-0.82	-0.99	0.32	0.84	0.62	-0.5
23	0.34	0.55	-0.36	-0.19	-0.86	-0.25
24	-0.43	-1.11	0.26	0.99	0.5	0.53
25	-0.69	0.25	-0.46	-0.73	0.48	0.02
26	-0.3	0.13	-0.52	-0.59	0.36	1.06
27	-0.18	-0.57	0.49	1.23	-0.6	-1.81
28	-0.69	0.25	-0.46	-0.73	0.48	0.02
29	-0.05	0.67	-0.3	-0.34	-0.74	-1.28
30	-0.3	0.13	-0.52	-0.59	0.36	1.06
31	-0.82	-0.99	0.32	0.84	0.62	-0.5
32	-0.69	0.25	-0.46	-0.73	0.48	0.02



Lampiran 6. (Lanjutan)

No	Z1	Z2	Z3	Z4	Z5	Z6
33	0.34	0.55	-0.36	-0.19	-0.86	-0.25
34	-0.69	0.25	-0.46	-0.73	0.48	0.02
35	0.22	-0.69	0.43	1.38	-0.72	-0.77
36	1.13	0.31	-0.48	0.11	-1.09	1.83
37	0.34	0.55	-0.36	-0.19	-0.86	-0.25
38	-0.3	0.13	-0.52	-0.59	0.36	1.06
39	-0.82	-0.99	0.32	0.84	0.62	-0.5
40	-0.69	0.25	-0.46	-0.73	0.48	0.02
41	-0.82	-0.99	0.32	0.84	0.62	-0.5
42	-0.18	-0.57	0.49	1.23	-0.6	-1.81
43	-0.69	0.25	-0.46	-0.73	0.48	0.02
44	1.13	0.31	-0.48	0.11	-1.09	1.83
45	-0.82	-0.99	0.32	0.84	0.62	-0.5
46	0.34	0.55	-0.36	-0.19	-0.86	-0.25
47	-0.69	0.25	-0.46	-0.73	0.48	0.02
48	-0.82	-0.99	0.32	0.84	0.62	-0.5
49	-0.69	0.25	-0.46	-0.73	0.48	0.02
50	1.52	0.18	-0.54	0.26	-1.2	2.87
51	0.22	-0.69	0.43	1.38	-0.72	-0.77
52	-0.69	0.25	-0.46	-0.73	0.48	0.02
53	-0.3	0.13	-0.52	-0.59	0.36	1.06
54	0.22	-0.69	0.43	1.38	-0.72	-0.77
55	-0.05	-0.67	-0.3	-0.34	-0.74	-1.28
56	-0.43	1.11	0.26	0.99	0.5	0.53
57	1	-0.93	0.31	1.68	-0.95	1.31
58	-0.3	0.13	-0.52	-0.59	0.36	1.06
59	-0.69	0.25	-0.46	-0.73	0.48	0.02
60	-0.69	0.25	-0.46	-0.73	0.48	0.02
61	-0.82	-0.99	0.32	0.84	0.62	-0.5
62	-0.3	0.13	-0.52	-0.59	0.36	1.06
63	0.75	-1.47	0.08	1.44	0.15	3.65
64	0.74	0.43	-0.42	-0.04	-0.97	0.79
65	-0.69	0.25	-0.46	-0.73	0.48	0.02

Lampiran 6. (Lanjutan)

No	Z1	Z2	Z3	Z4	Z5	Z6
66	6.87	-0.25	-2.54	0.6	5.7	-2.08
67	1.94	0.4	-1.38	0.21	1.8	-1.02
68	-0.82	-0.99	0.32	0.84	-0.62	-0.5
69	-0.03	-1.23	0.2	1.14	-0.39	1.57
70	-0.82	-0.99	0.32	0.84	-0.62	-0.5
71	-0.82	-0.99	0.32	0.84	-0.62	-0.5
72	1.78	4.36	5.18	-5.48	-0.58	-0.12
73	1.25	-1.02	2.13	-3.5	-0.76	-1.09
74	-0.05	0.67	-0.3	-0.34	-0.74	-1.28
75	-0.03	-1.23	0.2	1.14	-0.39	1.57
76	0.34	0.55	-0.36	-0.19	-0.86	-0.25
77	-0.05	0.67	-0.3	-0.34	-0.74	-1.28
78	-0.82	-0.99	0.32	0.84	-0.62	-0.5
79	-0.05	0.67	-0.3	-0.34	-0.74	-1.28
80	-0.3	0.13	-0.52	-0.59	-0.36	1.06
81	-0.69	0.25	-0.46	-0.73	-0.48	0.02
82	-0.05	0.67	-0.3	-0.34	-0.74	-1.28
83	-0.69	0.25	-0.46	-0.73	-0.48	0.02
84	-0.43	-1.11	0.26	0.99	0.5	0.53
85	-0.3	0.13	-0.52	-0.59	-0.36	1.06
86	-0.3	0.13	-0.52	-0.59	-0.36	1.06
87	0.34	0.55	-0.36	-0.19	-0.86	-0.25
88	0.34	0.55	-0.36	-0.19	-0.86	-0.25
89	-0.69	0.25	-0.46	-0.73	-0.48	0.02
90	2.34	-1.16	1.25	1.12	-2.25	0.76
91	-0.69	0.25	-0.46	-0.73	-0.48	0.02
92	-0.3	0.13	-0.52	-0.59	-0.36	1.06
93	-0.43	-1.11	0.26	0.99	0.5	0.53
94	0.22	-0.69	0.43	1.38	-0.72	-0.77
95	0.74	0.43	-0.42	-0.04	-0.97	0.79
96	-0.69	0.25	-0.46	-0.73	-0.48	0.02
97	-0.43	-1.11	0.26	0.99	0.5	0.53
98	-0.3	0.13	-0.52	-0.59	-0.36	1.06



Lampiran 6. (Lanjutan)

No	Z1	Z2	Z3	Z4	Z5	Z6
99	-0.05	0.67	-0.3	-0.34	-0.74	-1.28
100	-0.69	0.25	-0.46	-0.73	0.48	0.02
101	-0.3	0.13	-0.52	-0.59	0.36	1.06
102	-0.05	0.67	-0.3	-0.34	-0.74	-1.28

Lampiran 7. Akar Ciri Komponen Utama *Nonlinier* Dan Komponen Loading

Dimension Eigenvalue

1	.3164
2	.2063
3	.1587
4	.1353
5	.1154
6	.0678

Component Loadings

Variable	Dimension					
	1	2	3	4	5	6
X5	.032	-.650	-.673	-.187	.278	-.107
X6	-.114	-.709	.348	.593	.044	-.099
X1	.718	.304	.094	.189	-.499	.313
X2	.712	-.011	-.284	.064	.597	-.229
X3	.843	-.169	-.065	.137	-.091	.479
X4	.388	-.438	.534	.605	.015	-.078

Lampiran 8. Hasil *Output* Indeks Validitas

Joining Cluster	Distance	Member of Cluster	RMSSTD	DB	Dunn
Case 102	0.984	63	63.626	0.483	0.251
Case 94	1.121	89	89.604	0.466	0.216
Case 14	1.122	2	3.564	0.521	0.216
Case 95	1.377	93	93.58	0.433	0.387
Case 67	1.5	94	94.576	0.424	0.387
Case 63	1.939	96	96.57	0.312	0.719
Case 73	2.058	97	97.575	0.302	0.723
Case 90	2.54	100	100.517	0.228	0.679
Case 72	3.767	101	101.522	0.236	0.608

Lampiran 9. Hasil Keanggotaan Kelompok Analisis *Cluster* Komponen
Utama *Nonlinier*

Kelompok 1	X1	X2	X3	X4	X5	X6
Suwaji	5	7117500000	155125	12000000	2	1

Kelompok 2	X1	X2	X3	X4	X5	X6
Sutrismo	3	2263000000	49275	4800000	2	1

Kelompok 3	X1	X2	X3	X4	X5	X6
Ridwan	6	503700000	109500	8480000	2	1
Yusuf	4	419750000	91250	7555000	2	1
Achmad S.	2	419750000	91250	7100000	2	2
Syafi'I I.	4	755555000	16425	2275000	2	1
A. Suedi	5	419020000	91250	8250000	2	2

Kelompok 4	X1	X2	X3	X4	X5	X6
Kasian	3	82125000	18250	2150000	2	2
Samuki	4	142350000	31025	1750000	1	1
Supangat	4	65700000	14600	1000000	1	1
Mani	3	25185000	5475	500000	1	1
Sulkah	3	67160000	14600	1950000	2	1
Rasemo	3	293825000	63875	2649000	2	1
Sekak	4	226665000	49275	2371000	2	1
Slamet	2	98550000	21900	2000000	2	1
Ngatmani	2	24820000	5475	1650000	2	1
Reban	3	131400000	29200	1825000	2	1
Tawi	3	67160000	14600	1950000	2	1
Syukur	2	32850000	7300	1775000	2	1
Mulyono	3	50370000	10950	1900000	2	2
Mufid	3	82125000	18250	2200000	2	1

Lampiran 9. (Lanjutan)

Kelompok 4	X1	X2	X3	X4	X5	X6
Biono	4	251850000	54750	5235000	2	1
Mustofa	2	16425000	3650	1600000	2	1
Mustaqim	3	32850000	7300	1825000	2	1
Sumari	3	183960000	40150	2190000	2	2
Badrus	2	11680000	25550	2450000	2	2
Iwan H.	3	254040000	54750	2880000	2	1
M. Ainur	2	83950000	18250	2200000	2	2
Sugianto	3	167535000	36500	2894000	2	1
Syamsul A.	2	118260000	25550	2500000	2	2
Machmud	2	41062500	9125	1925000	2	1
Juwanto	2	118260000	25550	2400000	2	1
Suwandi	3	83950000	18250	2000000	2	2
Rochmad N.	2	100375000	21900	2320000	2	1
A. Soleh	3	83950000	18250	1800000	2	1
Umar Fauzi	2	117530000	25550	2400000	2	1
Sumiati	2	65700000	14600	2776000	2	2
Senu	2	41062500	9125	1338000	2	1
Chamdani	3	117895000	25550	3340000	2	1
Moh. Sani	2	65700000	14600	1850000	2	1
Ainul M.	3	205860000	45625	2258000	2	2
Kukuh K.	3	335800000	73000	7690000	2	1
Sunyoto	3	167535000	36500	2652000	2	1
Wasir	2	167900000	36500	2396000	2	1
Ivan K.	2	74460000	16425	1840000	2	2
Matali	2	67160000	14600	1930000	2	1
Junaidi	2	100740000	21900	2240000	2	2
Haryono	3	83950000	18250	2300000	2	2
Djari	1	33215000	7300	1338000	2	1
Ketang G.	4	335800000	73000	6950000	2	1
Sanusi	2	50370000	10950	2225000	2	2

Lampiran 9. (Lanjutan)

Kelompok 4	X1	X2	X3	X4	X5	X6
Winanik	3	186150000	40150	3360000	2	1
Syafi'i	2	67160000	14600	1925000	2	1
Syafi'i A.	2	50187500	10950	2250000	2	2
Churoti	2	111690000	24455	2868000	2	1
Nanto	3	209875000	45625	4750000	2	2
Abdul M.	2	75555000	16425	2276000	2	1
Agus H.	2	142350000	31025	2300000	2	1
Rifkik	3	201480000	43800	2802000	2	2
Achadun	3	109500000	23725	2868000	2	1
H. Asy'Ari	2	117530000	25550	2675000	2	2
Chamim	3	337260000	73000	7100000	2	2
Anas M.	2	124830000	27375	2960000	2	1
M. Jalal	2	58400000	12775	2230000	2	1
M. Anshori	2	58765000	12775	2190000	2	1
Asmu'i	2	83950000	18250	2322000	2	2
R. Djumadi	2	142350000	32025	2780000	2	1
Imam C.	4	284700000	62050	5600000	2	1
Suyanto	2	33215000	7300	2568000	2	1
Didik Karim	2	50370000	10950	1800000	2	2
Buadi	2	271560000	58400	7010000	2	2
Anshori	2	41975000	9125	2000000	2	2
Erfin E.	2	58765000	12775	2200000	2	2
Dianur	2	251850000	54750	53000000	2	2
Usman	3	124830000	27375	27000000	2	1
Mustakin	2	250390000	54750	5650000	2	2
Kasiono	3	167900000	36500	2450000	2	1
Soni	3	32850000	7300	1850000	2	1
Yoyok	2	41975000	9125	2150000	2	2
Suruli	3	98550000	21900	2300000	2	1

Lampiran 9. (Lanjutan)

Kelompok 4	X1	X2	X3	X4	X5	X6
Nanda	2	201480000	43800	4500000	2	1
Isman	2	492750000	10950	2275000	2	1
Abdul K.	3	496400000	10950	2345000	2	1
Ridho I	2	503700000	10950	1900000	2	1
Mariono	2	142350000	31025	2875000	2	2
Royan	2	168630000	36500	2001000	2	1
Jasmani	2	149650000	32850	2105000	2	1
Sholeh	3	134320000	29200	2820000	2	1
Hadi S.	3	167900000	36500	2500000	2	1
Siti Aisyiyah	2	67160000	14600	2653000	2	1
Agus Salim	2	92345000	20075	2430000	2	1
Romeli	2	210240000	45625	2588000	2	1
Kaseri	2	144540000	31390	2920000	2	2
Dachroni	3	208780000	45625	4750000	2	2
Supardi	3	226300000	49275	3170000	2	1
Thoyib	2	83950000	18250	2500000	2	1
Subakri	2	127020000	27740	2700000	2	2
Satik	2	167900000	36500	2450000	2	1
Sujiono	3	100740000	21900	2625000	2	1
Gimin	2	98550000	21900	2250000	2	1
Soli	2	125560000	27375	2825000	2	1
Solikin	3	82125000	18250	2775000	2	1

Lampiran 10. Hasil *Output Latent Class Cluster*

Nbr Clusters	LL	BIC(LL)	AIC(LL)	AIC3(LL)	Number of parameters
2	4936.855	9970.835	9915.710	9936.710	21
3	4831.565	9811.129	9727.130	9759.130	32
4	4835.644	9870.161	9757.287	9800.287	43
5	4694.194	9638.136	9496.388	9550.388	54
6	4670.993	9642.609	9471.985	9536.985	65
7	4650.241	9651.979	9452.481	9528.481	76
8	4632.482	9667.337	9438.964	9525.964	87
9	4622.708	9698.663	9441.415	9539.415	98
10	4604.406	9712.934	9426.812	9535.812	109





Lampiran 10. (Lanjutan)

Estimation summary (5 Clusters):

EM iterations	15
Log-posterior	-4825.665
L^2	9390.359
Final convergence value	0.010

Log-likelihood statistics (5 Clusters):

Log-likelihood(LL)	-4694.194
Log-prior	-130.431
Log-posterior	-4824.624
BIC(LL)	9638.136
AIC(LL)	9496.388
AIC3(LL)	9550.388
CAIC(LL)	9692.136
SABIC(LL)	9467.570

Lampiran 11. Hasil Keanggotaan Kelompok *Latent Class Cluster*

Observation	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster
Obs1	0.000	1.000	0.000	0.000	0.000	2
Obs2	0.000	0.000	0.000	1.000	0.000	4
Obs3	0.000	0.000	0.000	1.000	0.000	4
Obs4	0.000	0.002	0.000	0.998	0.000	4
Obs5	0.000	0.997	0.000	0.003	0.000	2
Obs6	0.000	0.961	-0.039	0.000	0.000	2
Obs7	0.000	0.000	-0.964	0.029	0.007	3
Obs8	1.000	0.000	0.000	0.000	0.000	1
Obs9	0.999	0.000	0.000	0.001	0.000	1
Obs10	0.000	0.999	0.000	0.001	0.000	2
Obs11	0.000	0.997	0.000	0.003	0.000	2
Obs12	1.000	0.000	0.000	0.000	0.000	1
Obs13	0.000	1.000	0.000	0.000	0.000	2
Obs14	0.000	0.000	-0.995	0.000	0.005	3
Obs15	0.000	0.998	-0.000	0.002	0.000	2
Obs16	0.000	0.000	-0.999	0.000	0.001	3
Obs17	0.999	0.000	0.000	0.001	0.000	1
Obs18	0.000	0.992	0.000	0.008	0.000	2
Obs19	0.000	1.000	0.000	0.000	0.000	2
Obs20	1.000	0.000	0.000	0.000	0.000	1
Obs21	0.000	0.995	0.005	0.000	0.000	2
Obs22	1.000	0.000	0.000	0.000	0.000	1
Obs23	0.000	1.000	0.000	0.000	0.000	2
Obs24	1.000	0.000	0.000	0.000	0.000	1
Obs25	1.000	0.000	0.000	0.000	0.000	1
Obs26	1.000	0.000	0.000	0.000	0.000	1
Obs27	0.000	1.000	0.000	0.000	0.000	2
Obs28	1.000	0.000	0.000	0.000	0.000	1
Obs29	0.000	0.997	0.000	0.003	0.000	2
Obs30	1.000	0.000	0.000	0.000	0.000	1
Obs31	1.000	0.000	0.000	0.000	0.000	1
Obs32	0.999	0.000	0.000	0.001	0.000	1
Obs33	0.000	1.000	0.000	0.000	0.000	2

Lampiran 11. (Lanjutan)



Observation	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster
Obs34	1.000	0.000	0.000	0.000	0.000	1
Obs35	0.000	1.000	0.000	0.000	0.000	2
Obs36	0.000	0.000	0.999	0.000	0.001	3
Obs37	0.000	1.000	0.000	0.000	0.000	2
Obs38	1.000	0.000	0.000	0.000	0.000	1
Obs39	1.000	0.000	0.000	0.000	0.000	1
Obs40	1.000	0.000	0.000	0.000	0.000	1
Obs41	1.000	0.000	0.000	0.000	0.000	1
Obs42	0.000	1.000	0.000	0.000	0.000	2
Obs43	0.000	0.000	0.000	1.000	0.000	4
Obs44	0.000	0.000	1.000	0.000	0.000	3
Obs45	1.000	0.000	0.000	0.000	0.000	1
Obs46	0.000	1.000	0.000	0.000	0.000	2
Obs47	1.000	0.000	0.000	0.000	0.000	1
Obs48	1.000	0.000	0.000	0.000	0.000	1
Obs49	1.000	0.000	0.000	0.000	0.000	1
Obs50	0.000	0.000	0.999	0.000	0.001	3
Obs51	0.000	0.900	0.100	0.000	0.000	2
Obs52	1.000	0.000	0.000	0.000	0.000	1
Obs53	1.000	0.000	0.000	0.000	0.000	1
Obs54	0.000	0.999	0.001	0.000	0.000	2
Obs55	0.000	0.999	0.000	0.001	0.000	2
Obs56	1.000	0.000	0.000	0.000	0.000	1
Obs57	0.000	0.000	1.000	0.000	0.000	3
Obs58	1.000	0.000	0.000	0.000	0.000	1
Obs59	1.000	0.000	0.000	0.000	0.000	1
Obs60	1.000	0.000	0.000	0.000	0.000	1
Obs61	1.000	0.000	0.000	0.000	0.000	1
Obs62	1.000	0.000	0.000	0.000	0.000	1
Obs63	0.000	0.000	1.000	0.000	0.000	3
Obs64	0.000	0.000	1.000	0.000	0.000	3
Obs65	1.000	0.000	0.000	0.000	0.000	1
Obs66	0.000	0.000	0.000	0.000	1.000	5

Lampiran 11. (Lanjutan)

Observation	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster
Obs67	0.000	0.000	0.000	0.000	1.000	5
Obs68	1.000	0.000	0.000	0.000	0.000	1
Obs69	0.000	0.000	1.000	0.000	0.000	3
Obs70	1.000	0.000	0.000	0.000	0.000	1
Obs71	1.000	0.000	0.000	0.000	0.000	1
Obs72	0.000	0.000	0.000	0.000	1.000	5
Obs73	0.000	0.000	0.000	0.000	1.000	5
Obs74	0.000	0.000	0.000	0.996	0.003	4
Obs75	0.000	0.000	1.000	0.000	0.000	3
Obs76	0.000	1.000	0.000	0.000	0.000	2
Obs77	0.000	0.993	0.000	0.007	0.000	2
Obs78	1.000	0.000	0.000	0.000	0.000	1
Obs79	0.000	0.999	0.000	0.001	0.000	2
Obs80	0.024	0.000	0.971	0.003	0.002	3
Obs81	1.000	0.000	0.000	0.000	0.000	1
Obs82	0.000	0.997	0.000	0.003	0.000	2
Obs83	1.000	0.000	0.000	0.000	0.000	1
Obs84	1.000	0.000	0.000	0.000	0.000	1
Obs85	1.000	0.000	0.000	0.000	0.000	1
Obs86	1.000	0.000	0.000	0.000	0.000	1
Obs87	0.000	1.000	0.000	0.000	0.000	2
Obs88	0.000	1.000	0.000	0.000	0.000	2
Obs89	1.000	0.000	0.000	0.000	0.000	1
Obs90	0.000	0.000	1.000	0.000	0.000	3
Obs91	1.000	0.000	0.000	0.000	0.000	1
Obs92	0.997	0.000	0.003	0.000	0.000	1
Obs93	1.000	0.000	0.000	0.000	0.000	1
Obs94	0.000	0.901	0.099	0.000	0.000	2
Obs95	0.000	0.998	0.002	0.000	0.000	2
Obs96	1.000	0.000	0.000	0.000	0.000	1
Obs97	1.000	0.000	0.000	0.000	0.000	1
Obs98	1.000	0.000	0.000	0.000	0.000	1

Lampiran 11. (Lanjutan)

Observation	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster
Obs99	0.000	0.999	0.000	0.001	0.000	2
Obs100	1.000	0.000	0.000	0.000	0.000	1
Obs101	1.000	0.000	0.000	0.000	0.000	1
Obs102	0.000	0.999	0.000	0.001	0.000	2