

**ANALISIS TWO STEP CLUSTER (TSC) DAN ANALISIS
LATENT CLASS CLUSTER (LCC) PADA PENGELOMPOKAN
DATA BERSKALA CAMPURAN KATEGORIK DAN
KONTINU**

SKRIPSI

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

oleh:

FACHRIYATUL UMAH
0910950038-95



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2014**

**LEMBAR PENGESAHAN SKRIPSI
PERBANDINGAN METODE PEMULUSAN EKSPONENSIAL
GANDA (LINIER SATU PARAMETER) DAN METODE
PEMULUSAN ADAPTIF SATU PARAMETER DARI BROWN**

oleh:
FACHRIYATUL UMAH
0910950038-95

**Setelah dipertahankan di depan Majelis Penguji
pada tanggal 23 Januari 2014
dan dinyatakan memenuhi syarat untuk memperoleh gelar
Sarjana Sains dalam bidang Statistika**

Pembimbing I

Pembimbing II

Dr. Ir. Solimun, MS
NIP. 196112151987031002

Samingun Handoyo, S.Si., M.Cs
NIP. 197304151998021002

Mengetahui,
Ketua Jurusan Matematika
Fakultas MIPA Universitas Brawijaya

Drs. Abdul Rouf Alghofari, M.Sc. Ph.D
NIP. 196709071992031001

LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : **Fachriyatul Umah**
NIM : **0910950038**
Jurusan : **Matematika**
Penulis Skripsi Berjudul : **Analisis *Two Step Cluster (TSC)* dan *Analisis Latent Class Cluster (LCC)* pada Pengelompokan Data Berskala Campuran**

Dengan ini menyatakan bahwa:

- 1. Skripsi ini adalah benar-benar karya saya sendiri dan bukan hasil plagiat dari karya orang lain. Karya-karya yang tercantum dalam Daftar Pustaka Skripsi ini, semata-mata digunakan sebagai acuan/referensi.**
- 2. Apabila dikemudian hari diketahui bahwa isi Skripsi saya merupakan hasil plagiat, maka saya bersedia menanggung resiko yang akan saya terima.**

Demikian pernyataan ini dibuat dengan segala kesadaran.

Malang, 23 Januari 2014
Yang menyatakan,

Fachriyatul Umah
NIM. 0910950038

ABSTRAK

Analisis *Two Step Cluster (TSC)* dan Analisis *Latent Class Cluster (LCC)* pada Pengelompokan Data Berskala Campuran Kategorik dan Kontinu

Analisis *cluster* (kelompok) merupakan analisis yang digunakan untuk mengelompokkan objek-objek berdasarkan ukuran kemiripan. Data yang digunakan pada analisis *cluster* dapat terdiri dari data kategorik saja, data kontinu saja, maupun campuran antara data kategorik dan data kontinu. Terkait dengan permasalahan pengelompokan pada data berskala campuran, terdapat beberapa metode yang dapat digunakan, diantaranya adalah metode *Two Step Cluster (TSC)* dan *Latent Class Cluster (LCC)*. Penelitian ini bertujuan untuk mengetahui perbedaan hasil pengelompokan pada data berskala campuran dengan menggunakan analisis *TSC* dan *LCC*, serta mengetahui metode manakah yang lebih baik dalam pengelompokannya. Kriteria perbandingan yang digunakan yaitu rasio keragaman antara *SSW (Sum of Square Within Cluster)* dengan *SSB (Sum of Square Between Cluster)*. Hasil pengelompokan yang baik adalah jika objek dalam kelompok yang sama memiliki keragaman yang kecil sedangkan objek antar kelompok memiliki keragaman yang besar. Data yang digunakan pada penelitian ini adalah 20 data sekunder yang memiliki variabel berskala campuran kategorik dan kontinu. Hasil penelitian menunjukkan dari 20 data yang digunakan, terdapat 4 data yang menghasilkan banyaknya kelompok yang sama dan 16 data lainnya menghasilkan banyaknya kelompok yang berbeda antara analisis *TSC* dan *LCC*. Nilai rata-rata dari 20 rata-rata rasio keragaman (*SSW/SSB*) untuk analisis *TSC* sebesar 1,3006 dan untuk analisis *LCC* sebesar 0,6227, sehingga dapat disimpulkan bahwa analisis *Latent Class Cluster* menghasilkan pengelompokan yang lebih baik dari analisis *Two Step Cluster*.

Kata kunci: *Cluster*, Analisis *Two Step Cluster (TSC)*, Analisis *Latent Class Cluster (LCC)*

ABSTRACT

Two Step Cluster Analysis (TSC) and Latent Class Cluster Analysis (LCC) in Clustering Mixed Scale Data Categories and Continuous

Clusters analysis is an analysis that is used to classify objects based on similarity measure. The data used in the cluster analysis can consist of categories data only, continuous data only, or a mix between categories and continuous data. Related to the issue clustering of mixed scale data, there are several methods that can be used, including the method of Two Step Cluster analysis (TSC) and Latent Class Cluster analysis (LCC). This research aims to know the differences results in clustering mixed scale data by using analysis of TSC and LCC, and find out which method is better in its clustering. Criteria of comparison that used in this research is ratio of the diversity between SSW (Sum of Square Within a Cluster) and SSB (Sum of Square Between Clusters). A good clustering results is if the object in the same group have small diversity whereas objects between groups has a great diversity. The data used in this research is 20 secondary data that has a variable mixture of categories and continuous scale. The results showed that from 20 data used in this research, there were only 4 data that produces same number of cluster and 16 other data yield different number of cluster in analysis of TSC and LCC. Average value of 20 average ratio of diversity (SSW/SSB) for TSC is 1,3006 and for LCC is 0,6227, so the conclusion is Latent Class Cluster (LCC) analysis produces a better clustering than Two Step Cluster (TSC) in clustering mixed scale data categories and continuous.

Keyword : Cluster, Two Step Cluster Analysis (TSC), Latent Class Cluster Analysis (LCC)

KATA PENGANTAR

Alhamdulillah puji syukur kehadirat Allah SWT atas rahmat dan hidayah-Nya penulis dapat menyelesaikan skripsi yang berjudul “Analisis *Two Step Cluster (TSC)* dan Analisis *Latent Class Cluster (LCC)* pada Pengelompokan Data Berskala Campuran Kategorik dan Kontinu”. Dalam penyusunan skripsi ini, banyak pihak yang telah memberikan bantuan kepada penulis. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan terimakasih kepada:

1. Bapak Dr. Ir. Solimun, MS., selaku dosen pembimbing I, Bapak Samingun Handoyo, S.Si., M.Cs, selaku dosen pembimbing II, dan Ibu Dr. Ir. Atiek Iriany, MS., selaku ketua penguji yang telah memberikan motivasi, bimbingan, masukan dan pengarahan hingga skripsi ini terselesaikan dengan baik.
2. Bapak Drs. Abdul Rouf Alghofari, M.Sc. Ph.D, selaku Ketua Jurusan Matematika FMIPA Universitas Brawijaya Malang.
3. Orang Tua, Kakak, Adik dan keluarga besar atas dukungan dan doa yang diberikan kepada penulis.
4. Mbak Fitriyani Dewi Maulidya atas bantuan dan semangat yang diberikan kepada penulis.
5. Teman-teman Statistika Universitas Brawijaya angkatan 2008, 2009, dan 2010 atas bantuan, dukungan dan kerjasamanya.
6. Teman-teman satu bimbingan atas bantuan, dukungan dan kerjasamanya.
7. Seluruh pihak yang telah membantu penyusunan skripsi ini.

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

Malang, 23 Januari 2014

Penulis



DAFTAR ISI

HALAMAN JUDUL i

LEMBAR PENGESAHAN SKRIPSI ii

LEMBAR PERNYATAAN iii

ABSTRAK iv

ABSTRACT v

KATA PENGANTAR vi

DAFTAR ISI vii

DAFTAR GAMBAR ix

DAFTAR TABEL x

DAFTAR LAMPIRAN xi

BAB I. PENDAHULUAN

1.1. Latar Belakang 1

1.2. Rumusan Masalah 2

1.3. Tujuan Penelitian 2

1.4. Batasan Masalah 3

1.5. Manfaat Penelitian 3

BAB II. TINJAUAN PUSTAKA

2.1. Analisis *Cluster* 5

2.2. Analisis *Two Step Cluster* 5

2.2.1. Ukuran Jarak 5

2.2.2. Tahapan Pengelompokan 7

2.2.3. Penentuan Banyaknya Kelompok 8

2.3. Analisis *Latent Class Cluster* 12

 A. Variabel Berskala Nominal 12

 B. Variabel Berskala Ordinal 12

 C. Variabel Berskala Interval dan Rasio (Kontinu) 13

2.3.1. Pendugaan Parameter 15

2.3.2. Penentuan Banyaknya Kelompok 16

2.4. Keragaman Kelompok 17

 A. Keragaman untuk Data Kontinu 17

 B. Keragaman untuk Data Kategorik 18

2.5. Uji Hipotesis Dua Nilai Tengah 19



BAB III. DATA DAN METODE PENELITIAN

3.1. Sumber Data..... 21

3.2. Metode Analisis..... 30

3.2.1. Analisis *Two Step Cluster*..... 30

3.2.2. Analisis *Latent Class Cluster*..... 30

BAB IV. HASIL DAN PEMBAHASAN

4.1. Hasil Analisis *Two Step Cluster*..... 35

4.2. Hasil Analisis *Latent Class Cluster*..... 36

4.3. Pembahasan..... 38

4.3.1. Uji Nilai Tengah Rata-Rata Rasio Keragaman..... 40

BAB V. KESIMPULAN DAN SARAN

5.1. Kesimpulan..... 47

5.2. Saran..... 47

DAFTAR PUSTAKA..... 49

LAMPIRAN..... 53

DAFTAR GAMBAR

3.1.	Diagram Alir Metode Penelitian	32
3.2.	Diagram Alir Analisis <i>Two Step Cluster</i>	33
3.3.	Diagram Alir Analisis <i>Latent Class Cluster</i>	33
4.1.	Grafik Rasio Keragaman pada Analisis <i>Two Step Cluster</i> dan Analisis <i>Latent Class Cluster</i>	40
4.2.	Grafik Banyaknya Kelompok pada Analisis <i>Two Step Cluster</i> dan Analisis <i>Latent Class Cluster</i>	42
4.3.	Grafik antara Rasio Keragaman dan Banyaknya Objek pada Analisis <i>Two Step Cluster</i> dan Analisis <i>Latent Class Cluster</i>	44



DAFTAR TABEL

3.1. Sumber Data.....	21
4.1. Hasil Analisis <i>Two Step Cluster</i>	35
Hasil Analisis <i>Latent Class Cluster</i>	37
4.3. Hasil Analisis <i>Two Step Cluster</i> dan Analisis <i>Latent Class Cluster</i>	38
4.4. Hasil Uji Nilai Tengah pada Rata-Rata Rasio Keragaman <i>Two Step Cluster</i> dan <i>Latent Class Cluster</i>	41



DAFTAR LAMPIRAN

Lampiran 1. Data Sekunder	53
Lampiran 2. <i>Output Analisis Two Step Cluster</i>	66
Lampiran 3. Rasio Keragaman Analisis <i>Two Step Cluster</i>	76
Lampiran 4. <i>Output Analisis Latent Class Cluster</i>	80
Lampiran 5. Rasio Keragaman Analisis <i>Latent Class Cluster</i>	87
Lampiran 6. <i>Output Uji Ragam</i>	91
Lampiran 7. <i>Output Uji Nilai Tengah Rata-Rata Rasio Keragaman</i>	92
Lampiran 8. Langkah-langkah Analisis <i>Two Step Cluster</i>	93
Lampiran 9. Langkah-langkah Analisis <i>Latent Class Cluster</i>	96

BAB I PENDAHULUAN

1.1 Latar Belakang

Perkembangan metode untuk menganalisis maupun interpretasi data terus berkembang seiring dengan berkembangnya pengetahuan dan masalah-masalah baru yang muncul. Karakteristik suatu populasi terkadang tidak cukup dipahami hanya dengan mengamati satu variabel saja. Pada umumnya, terdapat permasalahan di mana banyaknya variabel yang diamati dalam suatu penelitian terdiri lebih dari satu ataupun dua variabel. Analisis *multivariate* merupakan metode statistika yang dalam analisisnya melibatkan banyak variabel. Salah satu penerapan analisis *multivariate* ialah analisis kelompok (*Cluster Analysis*). Analisis *cluster* digunakan untuk mengelompokkan objek-objek menjadi beberapa kelompok berdasarkan ukuran kemiripan yang diamati, sehingga diperoleh kemiripan objek dalam kelompok yang sama dibandingkan antar objek dari kelompok yang berbeda.

Data peubah ganda pada analisis *cluster* dapat terdiri dari data kategorik saja (skala nominal dan ordinal), data kontinu saja (skala interval dan rasio), maupun campuran antara data kategorik dan kontinu. Terkait dengan permasalahan pengelompokan pada data berskala campuran, terdapat beberapa metode yang dapat digunakan. Metode yang biasa digunakan dalam analisis *cluster* data campuran yaitu analisis *Two Step Cluster (TSC)* atau analisis kelompok dua tahap. Disebut dengan analisis kelompok dua tahap karena pada prosedur pengelompokan objeknya dilakukan melalui dua tahapan, yaitu tahap pembentukan kelompok awal (*preclustering*) dan tahap pembentukan kelompok akhir. Saat variabel yang digunakan merupakan kombinasi antara variabel berskala kontinu dan kategorik maka ukuran jarak yang digunakan adalah jarak *Log-likelihood* (Chan, 2005).

Selain dengan menggunakan analisis *Two Step Cluster (TSC)*, terdapat alternatif metode lain yang dapat digunakan untuk mengatasi masalah pengelompokan pada data berskala campuran yaitu analisis *Latent Class Cluster*. Analisis *Latent Class Cluster (LCC)* memungkinkan pengelompokan objek pada variabel berskala

campuran. Dalam analisis *Latent Class Cluster* (LCC) pengelompokan kasus didasarkan pada peluang, jadi kasus dikelompokkan ke dalam kelompok-kelompok berdasarkan peluang keanggotaannya yang dihitung secara langsung dari model, dengan pendugaan parameter menggunakan metode maksimum *likelihood*.

Secara garis besar, antara analisis *Two Step Cluster* dan analisis *Latent Class Cluster* memiliki perbedaan pada dasar pengelompokan yang digunakan. Pada analisis *Two Step Cluster*, dasar pengelompokan yang digunakan adalah dengan berdasarkan kedekatan atau jarak, sedangkan pada analisis *Latent Class Cluster* pengelompokan didasarkan pada model dan bersifat probabilistik. Dengan perbedaan dasar pengelompokan dalam hal *clustering* data berskala campuran pada kedua analisis tersebut, maka pada penelitian ini akan dilakukan perbandingan dari kedua metode tersebut untuk mengetahui perbedaan hasil pengelompokan pada data berskala campuran kategorik dan kontinu.

1.2 Rumusan Masalah

Berdasarkan latar belakang, dapat dirumuskan masalah yang mendasari penelitian ini sebagai berikut:

1. Apakah terdapat perbedaan hasil pengelompokan menggunakan analisis *Two Step Cluster* dan analisis *Latent Class Cluster* pada data berskala campuran?
2. Metode pengelompokan manakah yang lebih baik antara analisis *Two Step Cluster* dan analisis *Latent Class Cluster* dalam pengelompokan data berskala campuran?

1.3 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah:

1. Mengetahui perbedaan hasil pengelompokan pada data berskala campuran dengan menggunakan analisis *Two Step Cluster* dan analisis *Latent Class Cluster*.
2. Membandingkan analisis *Two Step Cluster* dan Analisis *Latent Class Cluster* dalam pengelompokan pada data berskala campuran.

1.4 Batasan Masalah

Batasan masalah dalam penelitian ini antara lain:

1. Data yang digunakan merupakan 20 data sekunder yang menggunakan variabel berskala campuran antara kategorik dan kontinu.
2. Perbedaan yang digunakan sebagai acuan perbandingan adalah perbedaan banyaknya pengelompokan yang terbentuk.
3. Penentuan metode yang terbaik dari kedua analisis dengan menggunakan keragaman kelompok, baik keragaman antar kelompok maupun keragaman dalam kelompok.

1.5 Manfaat

Manfaat dari penelitian ini diharapkan dapat memberikan informasi mengenai penggunaan analisis *Two Step Cluster* dan analisis *Latent Class Cluster* pada pengelompokan data berskala campuran. Selain itu, dapat pula diketahui perbedaan dari kedua metode tersebut baik secara teori maupun hasil yang diperoleh sehingga akan didapatkan informasi metode mana yang lebih baik dalam mengatasi masalah pengelompokan pada data berskala campuran.



BAB II TINJAUAN PUSTAKA

2.1. Analisis Cluster

Analisis *cluster* merupakan suatu analisis *multivariate* yang digunakan untuk mengelompokkan objek pengamatan menjadi beberapa kelompok berdasarkan ukuran kemiripan antar objek, sehingga objek-objek yang berada dalam satu kelompok memiliki kemiripan yang lebih homogen dibandingkan objek dari kelompok yang berbeda (Johnson dan Wichern, 2007).

Secara umum terdapat dua metode yang digunakan dalam analisis *cluster*, yaitu metode hirarki dan metode nonhirarki. Metode nonhirarki dimulai dengan menentukan terlebih dahulu banyaknya kelompok yang diinginkan sehingga sifat pengelompokkannya tidaklah alamiah karena dikondisikan untuk banyaknya kelompok tertentu. Metode hirarki digunakan ketika banyaknya kelompok yang terbentuk ditentukan berdasarkan hasil dari analisis.

Pengelompokan metode hirarki dibagi dalam dua algoritma, yaitu *agglomerative* dan *divisive*. Pada algoritma *agglomerative*, proses pengelompokan berlangsung dengan penggabungan obyek dalam kelompok-kelompok dan hasil akhirnya semua obyek bergabung menjadi satu kelompok besar. Sedangkan pada algoritma *divisive* bersifat sebaliknya, dimulai dari satu kelompok besar yang mencakup semua kelompok, selanjutnya kelompok yang memiliki ketidakmiripan cukup besar akan dipisahkan sehingga membentuk kelompok yang lebih kecil (Johnson dan Wichern, 2007).

2.2 *Two Step Cluster Analysis* (Analisis Kelompok Dua Tahap)

Analisis *Two Step Cluster* merupakan suatu analisis yang memberi solusi pada masalah kombinasi tipe variabel kontinu dan kategorik. Metode ini juga dapat menentukan banyaknya kelompok optimal dengan sendirinya (Chan, 2005).

2.2.1 Ukuran Jarak

Chan (2005) mengatakan bahwa saat variabel yang digunakan merupakan kombinasi antara variabel berskala kontinu dan kategorik maka ukuran jarak yang digunakan adalah jarak *Log-likelihood*. Jarak antara kelompok j dan s didefinisikan sebagai berikut:



$$d_{j,s} = \xi_j + \xi_s - \xi_{(j,s)} \quad (2.1)$$

dengan:

$$\xi_j = -N_j \sum_{k=1}^{K^A} \frac{1}{2} \log(\sigma_k^2 + \sigma_{jk}^2) + E_{j,k}^{K^A}$$

$$\xi_s = -N_s \sum_{k=1}^{K^B} \frac{1}{2} \log(\sigma_k^2 + \sigma_{sk}^2) + E_{s,k}^{K^B}$$

$$\xi_{(j,s)} = -N_{(j,s)} \sum_{k=1}^{K^A} \frac{1}{2} \log(\sigma_k^2 + \sigma_{(j,s)k}^2) + E_{(j,s)k}^{K^A}$$

$$E_{j,k} = - \sum_{l=1}^{L_k} \frac{N_{jkl}}{N_j} \log \frac{N_{jkl}}{N_j}$$

$$E_{s,k} = - \sum_{l=1}^{L_k} \frac{N_{skl}}{N_s} \log \frac{N_{skl}}{N_s}$$

$$E_{(j,s)k} = - \sum_{l=1}^{L_k} \frac{N_{(j,s)kl}}{N_{(j,s)}} \log \frac{N_{(j,s)kl}}{N_{(j,s)}}$$

di mana:

$d_{(j,s)}$ = jarak antara kelompok j dan s

ξ_j = log likelihood pada kelompok j

ξ_s = log likelihood pada kelompok s

$\xi_{(j,s)}$ = log likelihood pada kelompok gabungan j dan s

N_j = banyaknya objek pada kelompok j

N_s = banyaknya objek pada kelompok s

$N_{(j,s)}$ = banyaknya objek pada kelompok gabungan j dan s

N_{jkl} = banyaknya objek di kelompok j untuk variabel

kategorik ke-k dengan kategori ke-l

N_{skl} = banyaknya objek di kelompok s untuk variabel

kategorik ke-k dengan kategori ke-l

$N_{(j,s)kl}$ = banyaknya objek pada kelompok gabungan j dan s

untuk variabel kategorik ke-k dengan kategori ke-l

- L_k = banyaknya kategori untuk variabel kategorik ke-k
- σ_k^2 = penduga ragam variabel kontinu ke-k
- σ_{jk}^2 = penduga ragam kelompok j pada variabel kontinu ke-k
- σ_{sk}^2 = penduga ragam kelompok s pada variabel kontinu ke-k
- $\sigma_{(j,s)k}^2$ = penduga ragam kelompok gabungan j dan s pada variabel kontinu ke-k
- E_{jk} = nilai harapan kelompok j pada variabel kategorik ke-k
- E_{sk} = nilai harapan kelompok s pada variabel kategorik ke-k
- $E_{(j,s)k}$ = nilai harapan kelompok j dan s pada variabel kategorik ke-k
- K^A = banyaknya variabel kontinu
- K^B = banyaknya variabel kategorik

2.2.2 Tahapan Pengelompokan

Prosedur pengelompokan objek dalam analisis *Two Step Cluster* dilakukan melalui dua tahapan, yaitu tahap pembentukan kelompok awal (*Preclustering*) dan tahap pembentukan kelompok akhir.

A. Tahap Pengelompokan Awal (*Preclustering*)

Tujuan pengelompokan awal (*Preclustering*) adalah untuk memasukkan data matriks baru dengan objek yang lebih sedikit pada langkah selanjutnya. Pada tahap ini data yang ada dimasukkan satu per satu, kemudian ditentukan data tersebut harus masuk pada kelompok yang telah terbentuk sebelumnya atau membentuk kelompok yang baru. Data yang masuk dihitung jaraknya dengan menggunakan ukuran jarak yang telah ditentukan. Apabila jarak tersebut kurang dari kriteria ukuran penerimaan (*Threshold Distance*) maka data tersebut tergabung dengan *subcluster* pertama, tetapi jika sebaliknya maka data membentuk suatu *subcluster* yang baru. Batas penerimaan (*Threshold Distance*) merupakan suatu nilai yang dimulai dari nol dan akan berubah mengikuti ukuran jarak terkecil di antara jarak *subcluster-subcluster* yang telah terbentuk pada satu tingkat pengelompokan sebelumnya.

Metode pengelompokan pada tahap *preclustering* analisis kelompok dua tahap dapat dijelaskan sebagai berikut:

1. Inisialisasikan obyek pertama sebagai *subcluster* pertama yang dibentuk.
 2. Mencari nilai jarak antara *subcluster* pertama dengan obyek kedua.
 3. Apabila jarak antara *subcluster* pertama dengan obyek kedua lebih besar dari *threshold* (θ), maka dibentuk suatu *subcluster* baru. *Subcluster* baru tersebut merupakan obyek kedua itu sendiri.
 4. Sebaliknya, apabila syarat pada bagian (3) tidak terpenuhi, maka terjadi penggabungan obyek kedua dengan *subcluster* pertama.
 5. Mencari nilai jarak antara obyek ketiga dengan *subcluster* pertama maupun dengan *subcluster* kedua.
 6. Memilih nilai jarak terbesar diantara keduanya.
 7. Apabila nilai jarak terbesar lebih besar dari nilai *threshold* (θ), maka dibentuk suatu *subcluster* baru yang merupakan obyek ketiga itu sendiri.
 8. Sebaliknya, apabila syarat pada bagian (7) tidak terpenuhi, maka terjadi penggabungan obyek ketiga dengan *subcluster* yang memiliki nilai jarak terbesar.
 9. Hal itu berlangsung secara terus-menerus sampai obyek yang terakhir sehingga terbentuk *subcluster-subcluster* sebanyak r *subcluster* yang akan dikelompokkan lagi pada tahap kedua.
- (Theodoris dan Koutroumbas, 2003).

B. Tahap Pembentukan Kelompok Akhir

Pada tahap pembentukan kelompok akhir, hasil dari pengelompokan tahap pertama (*preclustering*) yaitu *subcluster* dikelompokkan menggunakan metode hirarki dengan prosedur penggabungan (*agglomerative*). Tiap-tiap *subcluster* yang terbentuk pada tahap pertama akan digabungkan satu persatu sesuai dengan ukuran jarak yang telah ditetapkan. Prosedur ini berakhir sampai seluruh *subcluster* menjadi suatu kelompok.

2.2.3 Penentuan Banyaknya Kelompok yang Terbentuk

Pada SPSS *Two Step Cluster*, banyaknya kelompok dapat diperoleh secara otomatis. Untuk menentukan banyaknya kelompok yang terbentuk, didasarkan pada nilai BIC, solusi kelompok yang terbaik memiliki nilai BIC terkecil. Rumus BIC untuk kelompok J adalah sebagai berikut:

$$BIC J = -2 \sum_{j=1}^J \xi_j + m_j \log(N) \quad (2.2)$$

di mana:

$$\xi_j = -N \sum_{k=1}^{K^A} \frac{1}{2} \log(\sigma_k^2 + \sigma_{jk}^2) + E_{jk}$$

$$E_{jk} = - \sum_{l=1}^{L_k} \frac{N_{jkl}}{N_j} \log \frac{N_{jkl}}{N_j}$$

$$m_j = J - 2K^A + (L_k - 1) \quad k=1$$

- N = banyaknya objek keseluruhan
- N_j = banyaknya objek di dalam kelompok j
- N_{jk} = banyaknya objek di dalam kelompok j untuk variabel kategorik ke-k dengan kategori ke-l
- σ_k^2 = penduga ragam variabel kontinu ke-k
- σ_{jk}^2 = penduga ragam kelompok j pada variabel kontinu ke-k
- K^A = banyaknya variabel kontinu
- K^B = banyaknya variabel kategorik
- L_k = banyaknya kategori untuk variabel kategorik ke-k

Misalkan terdapat sebuah data dengan 3 objek yang telah dilakukan analisis *Two Step Cluster*, pada tahap pengelompokan akhir dengan metode hirarki diperoleh hasil *cluster membership*:

- 1 cluster = ABC
- 2 cluster = AB dan C
- 3 cluster = A, B, dan C

data yang digunakan terdiri dari 3 variabel dengan 2 variabel berskala kategorik dan 1 variabel berskala kontinu:

	X1	X2	X3
A	3	1	1,1
B	2	1	0,5
C	1	2	0,7

kemudian pada tahap penentuan banyaknya kelompok, dihitung nilai BIC dengan persamaan (2.2)

1. Untuk 1 kelompok (ABC)

$$E_{jk} = - \sum_{l=1}^{L_k} \frac{N_{jkl}}{N_j} \log \frac{N_{jkl}}{N_j}$$

$$E_{jk} = - \frac{1}{3} \log \frac{1}{3} + \frac{1}{3} \log \frac{1}{3} + \frac{1}{3} \log \frac{1}{3} + \dots - \frac{2}{3} \log \frac{2}{3} + \frac{1}{3} \log \frac{1}{3}$$

$$= -0,159 - 0,159 - 0,159 + (-0,117 - 0,154)$$

$$= (0,477) + (0,276)$$

$$= 0,753$$

$$\xi_j = -N_j \sum_{k=1}^{K^A} \frac{1}{2} \log(\sigma_k^2 + \sigma_{jk}^2) + E_{jk}$$

$$\xi_1 = -3 \left[\frac{1}{2} \log(0,093 + 0,093) + 0,753 \right]$$

$$= -3(-0,3653 + 0,753)$$

$$= -1,164$$

$$m_j \log(N) = \sum_{k=1}^{K^B} 2K^A + (L_k - 1) \log(N)$$

$$m_j \log(N) = 2 \times 1 + (3 - 1) + (2 - 1) \log 3$$

$$= 2 + 3 \log 3$$

$$= 2,385$$

$$BIC_j = -2 \sum_{j=1}^J \xi_j + m_j \log(N)$$

$$BIC_1 = -2(-1,164) + 2,385 = 4,713$$

2. Untuk 2 kelompok (AB dan C)

$$E_{1k} = 0,3 \quad E_{2k} = 0$$

$$\xi_1 = -0,057 \quad \xi_2 = 1,545$$

$$m_j \log(N) = 4,77$$

$$BIC_1 = -2 \sum_{j=1}^k \xi_j + m_j \log(N)$$

$$BIC_2 = (-2 \cdot -0,057 + 1,545) + 4,77$$

$$= 1,794$$

3. Untuk 3 kelompok (A, B, dan C)

$$E_{1k} = 0$$

$$E_{2k} = 0$$

$$E_{3k} = 0$$

$$\xi_1 = 1,547$$

$$\xi_2 = 1,547$$

$$\xi_3 = 1,547$$

$$m_j \log(N) = 7,155$$

$$BIC_3 = -2 \sum_{j=1}^k \xi_j + m_j \log(N)$$

$$BIC_3 = -2 \cdot 1,547 + 1,547 + 1,547 + 7,155$$

$$= -2 \cdot 4,64 + 7,155$$

$$= -9,28 + 7,155$$

$$= -2,125$$

Banyaknya Kelompok	Anggota Kelompok	BIC
1	ABC	4,713
2	AB dan C	1,794
3	A, B, C	-2,125

Berdasarkan perhitungan nilai BIC, diperoleh BIC terkecil adalah pada saat terbentuk 3 kelompok, sehingga dapat diambil kesimpulan bahwa banyaknya kelompok yang terbentuk pada contoh kasus ini sebanyak 3 kelompok dengan masing-masing anggota kelompok merupakan objek itu sendiri, yaitu A, B, dan C.

2.3 Analisis Latent Class Cluster (LCC)

Latent Class Cluster digunakan untuk mengklasifikasikan individu atau objek ke dalam kelompok yang berbeda berdasarkan pada respon untuk sekumpulan indikator yang diamati. Model yang diusulkan dihitung menggunakan *maximum likelihood* dengan algoritma *Expectation Maximum* (EM). *Latent Class Cluster* untuk variabel campuran biner dan metrik diperluas untuk mengakomodasi semua jenis data (termasuk ordinal dan nominal). (Moustaki dan Papageorgiou, 2004). Vermunt dan Magidson (2005) mengatakan bahwa pengelompokan dengan analisis *Latent Class Cluster* didasarkan pada peluang, jadi kasus dikelompokkan ke dalam kelompok-kelompok berdasarkan peluang keanggotaannya yang dihitung secara langsung dari model. Hal ini jugalah yang membedakan analisis *Latent Class Cluster* (LCC) dengan analisis *cluster* yang biasa digunakan untuk mengelompokkan objek atau kasus berdasarkan jarak atau kedekatan.

Masing-masing variabel pengamatan dalam analisis LCC memiliki sebaran bersyarat sesuai dengan jenis skala data pengamatan. Untuk variabel kategorik (skala data nominal dan ordinal) mengikuti distribusi bersyarat multinomial, sedangkan untuk variabel dengan data berskala kontinu mengikuti distribusi normal.

A. Variabel Berskala Nominal

Pada variabel dengan data berskala nominal, model mengikuti distribusi Multinomial (Moustaki dan Papageorgiou, 2004):

$$g_i x_i j = \prod_{s=1}^{c_i} (\pi_{ij(s)})^{x_{i(s)}} \quad (2.3)$$

di mana:

$\pi_{ij(s)}$ = peluang objek akan berada di kelas j untuk kategori s dan

variabel ke- i ($\pi_{ij(s)} = \frac{n_{is}}{n_j}$)

j = banyaknya kelas yang terbentuk

s = banyaknya kategori jawaban ($s = 1, 2, \dots, c_i$)

i untuk i dengan kategori jawaban s ($s = 1, 2, \dots, c_i$)

$x_{i(s)}$ = 0 untuk selainnya

B. Variabel Berskala Ordinal

Untuk variabel berskala ordinal, x_i didefinisikan seperti pada variabel berskala nominal yaitu berdistribusi multinomial namun disertai urutan (order) dengan bentuk sebagai berikut (Moustaki dan Papageorgiou, 2004):

$$g_i x_i J = \prod_{s=1}^{m_i} \pi_{ij s}^{x_i(s)} \prod_{s=1}^{m_i} (\gamma_{ij s} - \gamma_{ij(s-1)})^{x_i(s)} \quad (2.4)$$

di mana:

$\pi_{ij(s)}$ = peluang objek akan berada di kelas j untuk kategori s dan variabel ke- i ($\pi_{ij(s)} = \frac{m_{is}}{n_j}$)

j = banyaknya kelas yang terbentuk

s = banyaknya kategori jawaban ($s = 1, 2, \dots, m_i$)

$\gamma_{ij s}$ = peluang kumulatif dari respon

dengan $\gamma_{ij s} = \pi_{ij 1} + \pi_{ij 2} + \dots + \pi_{ij s}$

dimana $j = 1, \dots, k$ dan $s = 1, \dots, m_i$

$x_i(s) = 1$ untuk i dengan kategori jawaban s

$x_i(s) = 0$ untuk selainnya

C. Variabel Berskala Interval dan Rasio (Kontinu)

Pada kasus di mana variabel dengan data berskala Interval dan Rasio (Kontinu), diasumsikan model mengikuti distribusi Normal dengan bentuk sebagai berikut (Moustaki dan Papageorgiou, 2004):

$$g x_i \mu_{ij} \sigma_i^2 = (2\pi)^{-\frac{1}{2}} \sigma_i^{-1} \exp \left\{ -\frac{1}{2\sigma_i^2} (x_i - \mu_{ij})^2 \right\} \quad (2.5)$$

di mana:

μ_{ij} = rata-rata variabel ke- i di kelas j

σ_i^2 = varians dari variabel ke- i

$\pi = 3,14$

$\exp = 2,718$

Model LCC memiliki fungsi sebaran peluang bersama yang merupakan peluang dari variabel-variabel dengan data berskala campuran yang diamati, sebagaimana berikut (Moustaki dan Papageorgiou, 2004):

$$f(x_h) = \prod_{j=1}^K \eta_j g_{x_h j} \quad (2.6)$$

di mana:

η_j = peluang awal (*prior*) kelas j ($\eta_j = \frac{n_j}{N}$)
 $g_{x_h j}$ = fungsi distribusi dari variabel yang dapat berupa nominal, ordinal, maupun kontinu
 dengan $g_{x_h j}$ dapat berupa variabel dengan data bertipe nominal, ordinal, dan kontinu seperti pada persamaan (2.4) (2.5) dan (2.6)
 maka diperoleh persamaan $g_{x j}$ sebagai berikut:

$$g_{x j} = \prod_{i=1}^N c_i \prod_{s=1}^{m_i} \pi_{ij s}^{x_{i s}} (Y_{ij s} - Y_{ij s-1})^{x_{i s}} \quad j = 1, \dots, k \quad (2.7)$$

$$(2\pi)^{-\frac{1}{2}} \sigma_i^{-1} \exp -\frac{1}{2\sigma_i^2} (x_i - \mu_{ij})^2$$

dengan N adalah banyaknya variabel dengan data berskala nominal, O adalah banyaknya variabel dengan data berskala ordinal, dan Q adalah banyaknya variabel dengan data berskala kontinu. Sehingga fungsi peluang bersama dari variabel-variabel pengamatan sesuai dengan bentuk persamaan (2.7) sebagaimana berikut:

$$f(x_h) = \prod_{j=1}^K \eta_j g_{x_h j} = \prod_{j=1}^K \prod_{i=1}^N c_i \prod_{s=1}^{m_i} \pi_{ij s}^{x_{i s}} (Y_{ij s} - Y_{ij s-1})^{x_{i s}} \quad (2.8)$$

$$(2\pi)^{-\frac{1}{2}} \sigma_i^{-1} \exp -\frac{1}{2\sigma_i^2} (x_i - \mu_{ij})^2$$

2.3.1 Pendugaan Parameter

Pendugaan parameter pada *latent class cluster* untuk variabel dengan tipe campuran menggunakan metode *maximum likelihood*. Metode *maximum likelihood* didasarkan oleh ide untuk mendapatkan estimator yang memaksimumkan *likelihood*-nya. Pendugaan parameter dilakukan pada fungsi peluang bersama yang diperoleh pada persamaan (2.8) di mana π adalah parameter yang diduga untuk variabel dengan data berskala nominal, γ adalah parameter yang diduga untuk variabel dengan data berskala ordinal, μ dan σ adalah parameter yang diduga untuk variabel dengan data berskala kontinu. Fungsi *log likelihood* untuk suatu sampel acak berukuran n adalah (Moustaki dan Papageorgiou, 2004):

$$\begin{aligned} L(\pi, \gamma, \mu, \sigma) &= \sum_{h=1}^n \log f(x_h) \\ L(\pi, \gamma, \mu, \sigma) &= \sum_{h=1}^n \log \left[\sum_{j=1}^K \eta_j g(x_h; \mu_j, \sigma_j) \right] \\ &= \sum_{h=1}^n \log \left[\sum_{j=1}^K \eta_j \prod_{s=1}^{C_s} \pi_{ij|s}^{x_{i|s}} \prod_{i=1}^0 2\pi^{-\frac{1}{2}} \sigma_i^{-1} \exp \left\{ -\frac{1}{2\sigma_i^2} x_i^2 \right\} \right] \end{aligned} \quad (2.10)$$

Untuk memaksimumkan fungsi *likelihood* dilakukan dengan proses iterasi menggunakan metode *Expectation Maximum* dan *Newton Raphson*. Proses iterasi menggunakan metode EM terdiri dari dua tahap yang dilakukan sampai diperoleh hasil yang konvergen, yaitu:

1. *Expectation* (E), dalam proses ini dilakukan identifikasi karakteristik dari masing-masing kelompok yang terbentuk
2. *Maximization* (M), proses identifikasi komposisi kelompok dan data agar memaksimumkan nilai *likelihood* data terhadap model kelompok yang dihasilkan.

Setelah dilakukan proses iterasi dengan EM dan diperoleh hasil yang optimal, maka proses iterasi dialihkan pada metode *Newton Raphson*. Metode iterasi *Newton Raphson* memiliki sifat yang lebih cepat berkerja dalam melakukan pendugaan parameter ketika nilainya mendekati optimal. Oleh karena itu, pendugaan parameter dalam analisis LC *Cluster* dilakukan menggunakan metode MLE dengan iterasi EM dan *Newton Raphson* (Vermunt dan Magidson, 2005).

2.3.2 Penentuan Banyaknya Kelompok yang Terbentuk

Berdasarkan model yang dihasilkan dari analisis, maka akan dipilih model terbaik yang mampu menjelaskan data. *Bayesian Information Criteria* (BIC) adalah sebuah kriteria statistik untuk memilih model. Model dengan nilai BIC yang lebih kecil dipilih sebagai model yang terbaik, karena nilai BIC yang lebih kecil menunjukkan bahwa model yang dihasilkan lebih bisa menjelaskan variasi dari data. Nilai untuk BIC dapat diperoleh menggunakan rumus (Vermunt dan Magidson, 2005):

$$BIC = -2LL + \ln N M \quad (2.11)$$

di mana:

N = Banyaknya objek pengamatan

M = Banyaknya parameter

LL = Nilai maksimum fungsi *log-likelihood* dari suatu model yang diestimasi

Misalkan pada suatu kasus pengelompokan yang terdiri dari 52 objek pengamatan diperoleh hasil:

	LL	M
1-Cluster	-1151,32	8
2-Cluster	-950,819	17
3-Cluster	-929,146	26
4-Cluster	-908,637	35
5-Cluster	-891,071	44

Maka perhitungan BIC untuk setiap kelompoknya adalah sebagai berikut:

$$\begin{aligned} BIC (1 \text{ kelompok}) &= -2LL + \ln N M \\ &= (-2 \times -1151,32) + (\ln (52) \times 8) \\ &= 2302,64 + 31,61 \\ &= 2334,249 \end{aligned}$$

$$\begin{aligned} BIC (2 \text{ kelompok}) &= -2LL + \ln N M \\ &= (-2 \times -950,81) + (\ln (52) \times 17) \\ &= 1901,63 + 67,17 \\ &= 1968,801 \end{aligned}$$

$$\begin{aligned} \text{BIC (3 kelompok)} &= -2LL + \ln N \cdot M \\ &= (-2 \times -929,14) + (\ln (52) \times 26) \\ &= 1858,28 + 102,73 \\ &= 1961,012 \end{aligned}$$

$$\begin{aligned} \text{BIC (4 kelompok)} &= -2LL + \ln N \cdot M \\ &= (-2 \times -908,637) + (\ln (52) \times 35) \\ &= 1817,27 + 138,29 \\ &= 1955,563 \end{aligned}$$

$$\begin{aligned} \text{BIC (5 kelompok)} &= -2LL + \ln N \cdot M \\ &= (-2 \times -891,071) + (\ln (52) \times 44) \\ &= 1782,14 + 173,85 \\ &= 1955,994 \end{aligned}$$

Dari perhitungan tersebut, untuk menentukan banyaknya kelompok yang terbentuk dipilih berdasarkan nilai BIC terkecil yang diperoleh. Nilai BIC terkecil didapatkan pada saat banyaknya kelompok yang terbentuk sebanyak 4, dengan nilai BIC sebesar 1955,563. Sehingga, disimpulkan bahwa banyaknya kelompok yang terbentuk dari kasus ini adalah sebanyak 4 kelompok.

2.4 Keragaman Kelompok

Pada dasarnya dalam pengelompokan terdapat dua keragaman, yaitu keragaman dalam kelompok (*Sum of Square Within Cluster / SSW*) dan keragaman antar kelompok (*Sum of Square Between Cluster / SSB*).

A. Keragaman untuk data kontinu

Bunkers *et al* (1996) dalam Dewi *et al* (2012) menerangkan persamaan rumus sebagai berikut untuk menghitung SSW dan SSB data dengan variabel kontinu

$$SSW = \frac{1}{C} \sum_{c=1}^C S_c^2 \quad (2.12)$$

$$SSB = \frac{1}{C-1} \sum_{c=1}^C (x_c - \bar{x})^2 \quad (2.13)$$

di mana:

SSW = Keragaman dalam kelompok

SSB = Keragaman antar kelompok

S_c = Simpangan baku kelompok ke-c

$$S_c = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

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

\bar{x} = Rata-rata total

C = Banyaknya kelompok

B. Keragaman untuk data kategorik

Light dan Margolin (1997) dalam Dewi *et al* (2012) menerangkan persamaan rumus sebagai berikut untuk menghitung SSW dan SSB data dengan variabel kategorik

$$SSW = \frac{1}{n-C} \sum_{c=1}^C \frac{1}{2} \sum_{k=1}^K \frac{1}{n_c} n_{kc}^2 \quad (2.14)$$

$$SSB = \frac{1}{C-1} \sum_{c=1}^C \frac{1}{2} \sum_{k=1}^K \frac{1}{n_c} n_{kc}^2 - \frac{1}{2n} \sum_{k=1}^K n_k^2 \quad (2.15)$$

di mana:

SSW = Keragaman dalam kelompok

SSB = Keragaman antar kelompok

n_k = banyaknya objek pada kategori ke-k, $k=1,2,\dots,K$

n_c = banyaknya objek pada kelompok ke-c, $c=1,2,\dots,C$

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

Hasil pengelompokan yang baik adalah jika objek dalam kelompok yang sama memiliki keragaman yang rendah sedangkan objek antar kelompok memiliki tingkat keragaman yang tinggi. Dengan kata lain, objek dalam satu kelompok memiliki tingkat kemiripan yang tinggi dan objek berbeda kelompok memiliki tingkat kemiripan yang rendah (Serban dan Grigoreta, 2006).

Rasio keragaman *cluster* dapat dihitung dengan membagi nilai SSW dan SSB. Nilai rasio keragaman yang kecil mengindikasikan *cluster* yang baik.

2.5 Uji Hipotesis Dua Nilai Tengah

Dalam praktik seringkali ingin diketahui apakah terdapat perbedaan yang berarti (*significant*) dari dua rata-rata. Pada penelitian ini ingin diketahui apakah terdapat perbedaan hasil pengelompokan berdasarkan nilai rata-rata rasio keragaman antara analisis *Two Step Cluster* dan analisis *Latent Class Cluster*, dengan menggunakan nilai rata-rata yang diperoleh dari rata-rata rasio keragaman untuk 20 data pada masing-masing metode analisis yang digunakan dalam penelitian ini akan dilakukan uji nilai tengah untuk mengetahui perbedaan antara metode analisis *Two Step Cluster* dan metode analisis *Latent Class Cluster* tersebut.

Secara umum, situasi yang melibatkan uji dua nilai tengah terkadang belum diketahui apakah ragam dari kedua populasi yang digunakan sama atau berbeda, sehingga perlu dilakukan uji untuk mengetahuinya, dengan menggunakan hipotesis:

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

statistik uji yang digunakan:

$$f = \frac{S_1^2}{S_2^2} \quad (2.16)$$

di mana s_1^2 dan s_2^2 adalah ragam dari masing-masing sampel, dengan derajat bebas yang digunakan $v_1 = n_1 - 1$ dan $v_2 = n_2 - 1$, untuk n adalah banyaknya ukuran sampel yang digunakan.

Setelah diketahui apakah terdapat perbedaan nilai ragam atau tidak dari kedua sampel yang digunakan, kemudian bisa dilakukan uji nilai tengah dengan menggunakan hipotesis:

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 > \mu_2$$

Jika pada uji ragam diperoleh hasil H_0 diterima, yang menunjukkan bahwa ragam dari kedua populasi sama, maka statistik uji yang digunakan adalah:



$$t_{hit} = \frac{(X_1 - X_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (2.17)$$

di mana:

$$s_p^2 = \frac{s_1^2 n_1 - 1 + s_2^2 (n_2 - 1)}{n_1 + n_2 - 2}$$

dengan derajat bebas yang digunakan $v = n_1 + n_2 - 2$.

Sedangkan jika pada uji ragam diperoleh hasil H_0 ditolak, yang menunjukkan bahwa ragam dari kedua populasi tidak sama, maka statistik uji yang digunakan adalah:

$$t_{hit} = \frac{(X_1 - X_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (2.18)$$

dengan derajat bebas yang digunakan :

$$v = \frac{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}}$$

(Walpole *et al.*, 2011)

BAB III

METODE PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan 20 data sekunder dengan variabel yang digunakan berskala campuran. Data yang digunakan dapat dilihat pada Tabel 3.1 berikut:

Tabel 3.1 Data Sekunder Penelitian

Data	Sumber	n	Variabel
1.	Skripsi Hermayudi, W.A (2004) Judul: Pengelompokan Industri Tempe Berdasarkan Klasifikasi Industri dengan <i>Cluster Analysis</i> (studi kasus di sentra industri tempe kota Malang) Fakultas: Teknik Pertanian Universitas Brawijaya	102	X_1 = Jumlah tenaga kerja (orang) X_2 = Nilai produksi (rupiah/tahun) X_3 = Jumlah bahan baku (kg/tahun) X_4 = Nilai investasi (Rp) X_5 = Teknologi yang digunakan 1 = alat modern 2 = alat tradisional X_6 = Aneka olahan 1 = Juga memproduksi selain tempe (keripik) 2 = Hanya memproduksi tempe
2.	Skripsi Nini Suryani (2011) Judul: Penerapan Model Regresi ZIGP pada Banyaknya Siswa Gagal Ujian Nasional di SMA/SMK Kab. Malang Fakultas: MIPA Universitas Brawijaya	64	X_1 = Akreditasi sekolah 1 = A 2 = B 3 = C 4 = Belum terakreditasi X_2 = Waktu sekolah 1 = Pagi 2 = Siang 3 = Pagi dan Siang

Data	Sumber	n	Variabel
			X_3 = Rasio penerimaan dan pendaftaran X_4 = Rasio guru dan murid X_5 = Rasio murid dan kelas
3.	Skripsi Ainun Nazhiirah (2010) Judul: Perbandingan penduga ML dan Bayesian MCMC pada Model Regresi Logistik Multilevel Fakultas: MIPA Universitas Brawijaya	79	X_1 = Asal TK 1 = Al Hidayah I 2 = Indria Tama 3 = Al Hidayah II X_2 = Skor pemilihan warna X_3 = Skor kerapian X_4 = Skor kebersihan X_5 = Jenis kelamin 0 = Perempuan 1 = Laki-Laki
4.	Dinas Pendidikan Provinsi Jawa Timur dalam Skripsi Ainun Nazhiirah (2010) Judul: Perbandingan penduga ML dan Bayesian MCMC pada Model Regresi Logistik Multilevel Fakultas: MIPA Universitas Brawijaya	104	X_1 = Status sekolah 1 = Negeri 0 = Swasta X_2 = Lama sekolah berdiri (tahun) X_3 = Banyaknya siswa terakhir X_4 = Banyaknya guru terakhir X_5 = Banyaknya alumni yang diterima di dunia usaha dan industri pada setahun terakhir X_6 = Status tanah bangunan 1 = Milik sendiri 0 = Menyewa X_7 = Nilai rata-rata ujian nasional sekolah X_8 = IPM tiap Kabupaten/Kota asal sekolah

Data	Sumber	n	Variabel
5.	Skripsi Nadhifa Isnayni (2011) Judul: Analisis Kualitas Angsuran Kredit Peserta Program Kredit Usaha Rakyat (KUR) Fakultas: Ekonomi Universitas Brawijaya	90	X_1 = Tingkat Pendidikan $1 = SD$ $2 = SLTP$ $3 = SLTA$ $4 = PT$ X_2 = Jumlah tanggungan keluarga X_3 = Jumlah pinjaman (Juta rupiah) X_4 = Jangka waktu (bulan) X_5 = Omzet (Juta rupiah) X_6 = Laba (Juta rupiah) X_7 = Pinjaman lain $1 = Ada$ $0 = Tidak ada$
6.	Skripsi Lely Marina (2009) Judul: Pengaruh Atribut Komite Audit dan <i>Financial Distress</i> terhadap Kualitas Pelaporan Keuangan (Studi Empiris pada Perusahaan yang Terdaftar di BEI tahun 2009) Fakultas: Ekonomi Universitas Brawijaya	58	X_1 = Jumlah anggota X_2 = Independensi komite(%) X_3 = Kompetensi komite audit $1 = Terdapat anggota komite audit yang berlatar belakang pendidikan dan berpengalaman di bidang keuangan$ $0 = Tidak terdapat anggota komite audit yang berlatar belakang pendidikan dan berpengalaman di bidang keuangan$ X_4 = Frekuensi pertemuan dalam 1 tahun X_5 = Tingkat kondisi (kesulitan) keuangan X_6 = Ukuran perusahaan (diukur dengan natural logaritma total aset)

Data	Sumber	n	Variabel
7.	Skripsi Dieta F (2012) Judul: Analisis Faktor- Faktor yang Mempengaruhi Keputusan Bank dalam Pemberian Kredit Usaha Mikro (MKM) (Studi kasus pada PT. Bank X Tbk. Kantor Cabang Probolinggo) Fakultas: Ekonomi Universitas Brawijaya	98	X_1 = Usia (tahun) X_2 = Laba pertahun (omzet) (Juta rupiah) X_3 = Taksasi jaminan (rupiah) X_4 = Status Agunan 1 = Marketable 0 = Unmarketable X_5 = Tingkat Pendidikan 1 = SD 2 = SELTP 3 = SLTA 4 = PT X_6 = Jenis Usaha 1 = Perdagangan 0 = Non perdagangan X_7 = Lama usaha (Tahun)
8.	Skripsi Rizka Niko (2011) Judul: Analisis Karakteristik Debitur Kredit UMKM sebagai Upaya Mengurangi Resiko Munculnya Kredit Macet Fakultas: Ekonomi Universitas Brawijaya	70	X_1 = Jenis kelamin 0 = Laki-laki 1 = Perempuan X_2 = Tingkat pendidikan 1 = SD 2 = SLTP 3 = SLTA 4 = PT X_3 = Status nasabah 1 = Nasabah lama 0 = Nasabah baru X_4 = Jumlah tanggungan dalam keluarga X_5 = Lama usaha yang digeluti debitur (tahun) X_6 = Besar pinjaman (rupiah) X_7 = Jangka waktu pengembalian kredit yang disepakati (bulan) X_8 = Omzet per bulan (juta rupiah) X_9 = Laba (rupiah) X_{10} = Jaminan kredit (rupiah)

Data	Sumber	n	Variabel
9.	Skripsi Eveline (2002) Judul: Hubungan Antara Faktor Resiko dengan Derajat Klinis Skabies di RSUD dr. Saiful Anwar Kota Malang Fakultas: Kedokteran Universitas Brawijaya	66	$X_1 =$ Usia (tahun) $X_2 =$ Tingkat pendidikan 1 = Belum Sekolah 2 = TK 3 = SD 4 = SMP 5 = SMA 6 = PT $X_3 =$ Tingkat sosial ekonomi: 1 = Rendah (< Rp. 500.000) 2 = Menengah (antara Rp. 500.000 s.d Rp. 1.000.000) 3 = Tinggi (> Rp. 1.000.000) $X_4 =$ Jumlah teman sekamar $X_5 =$ Higiene pribadi 0 = Buruk 1 = Baik
10.	Skripsi Iswenda .N (2007) Judul: Faktor-Faktor Penentu Tingkat Gangguan Jiwa Pasien Rumah Sakit Jiwa Daerah Dr. Rm. Soedjarwadi Di Kabupaten Klaten, Jawa Tengah, Tahun 2006 Sekolah Tinggi Ilmu Statistik (STIS)	100	$X_1 =$ Umur (tahun) $X_2 =$ Jenis kelamin 0=Laki-laki 1=Wanita $X_3 =$ Status tempat tinggal 0 = Desa 1 = Kota $X_4 =$ Tingkat pendidikan 1 = Tidak tamat SD 2 = SD-SMP 3 = > SMA $X_5 =$ Status kerja 1 = Bekerja 0 = Tidak $X_6 =$ Frekuensi muncul gejala 0 = Jarang 1 = Sering $X_7 =$ Frekuensi dirawat $X_8 =$ Kondisi ekonomi 0 = Buruk 1 = Baik $X_9 =$ Penyakit keturunan 0 = Tidak ada 1 = Ada

Data	Sumber	n	Variabel
11.	Skripsi M. Sahri (2011) Judul: Variabel-Variabel yang Mempengaruhi Pendapatan Petani Apel Fakultas: Ekonomi Universitas Brawijaya	50	$X_1 =$ Modal (rupiah) $X_2 =$ Luas lahan/m ² $X_3 =$ Jumlah tenaga kerja $X_4 =$ Tingkat pendidikan 1 = SD 2 = SLTP 3 = SLTA 4 = PT $X_5 =$ Umur $X_6 =$ Jumlah Anggota Keluarga $X_7 =$ Pengeluaran per Musim Panen (rupiah)
12.	Skripsi Perhatiningsih, S.F (2002) Judul: Perbandingan Profil Tekanan Darah Selama Serangan Infark Serebral Akut Dan Infark Miokard Akut Dari Penderita Hipertensi Di Rsud Dr. Saiful Anwar Malang Fakultas: Kedokteran Universitas Brawijaya	30	$X_1 =$ Usia (tahun) $X_2 =$ Jenis Kelamin 0 = Perempuan 1 = Laki-laki $X_3 =$ Hipertensi 0 = Tidak 1 = Ya $X_4 =$ Diabetes Melitus 0 = Ya 1 = Tidak
13.	Skripsi Catur Oktavina (2013) Judul: Analisis Produktivitas Wanita Pekerja Sektor Informal (Studi Kasus: Pedagang di Kota Malang) Fakultas: Ekonomi Universitas Brawijaya	100	$X_1 =$ Umur (tahun) $X_2 =$ Status pernikahan 1 = Tidak menikah 2 = Menikah $X_3 =$ Pendidikan (tahun) $X_4 =$ Pengalaman (tahun) $X_5 =$ Daerah 0 = Luar Malang 1 = Malang

Data	Sumber	n	Variabel
			X_6 = Lokasi 0 = Kurang strategis 1 = Strategis X_7 = Jarak (Km) X_8 = Kepemilikan usaha 0 = Milik orang lain 1 = Milik sendiri X_9 = Status pekerjaan 0 = Sampingan 1 = Utama X_{10} = Sistem upah 0 = Bukan bulanan 1 = Bulanan X_{11} = Jenis dagangan 0 = Non makanan 1 = Makanan X_{12} = Hub. dengan pemilik 0 = Tidak ada 1 = Ada X_{13} = Tanggungan keluarga (orang) X_{14} = Keamanan 0 = Kurang aman 1 = Aman X_{15} = Akses kredit 0 = Tidak ada 1 = Ada
14.	Skripsi Irawan Wahyu Trio (2004) Judul: Hubungan Tingkat Pengetahuan Keluarga tentang Pencegahan Penularan Tuberculosis Paru dengan Jumlah Anggota Keluarga Yang Tertular TBC Paru di Rumah Sakir Paru Batu Fakultas Kedokteran Universitas Brawijaya	47	X_1 = Umur (tahun) X_2 = Tingkat pendidikan 1 = SD 2 = SMP 3 = SMA 4 = PT X_3 = Pekerjaan 1 = PNS 2 = Swasta 3 = Tidak Bekerja 4 = Lain-lain X_4 = Pernah Mendapat Informasi tentang TBC 1 = Pernah 2 = Tidak X_5 = Nilai pengetahuan (%) X_6 = Nilai penerapan (%)

Data	Sumber	n	Variabel
15.	S. Weisberg, (1985). <i>Applied Linier Regression</i> . Second Edition. New York: John Wiley and Sons	52	X_1 = Jenis kelamin $1 = Lk$ $0 = Pr$ X_2 = Jabatan $1 =$ Assistant Proffesor $2 =$ Associate Proffesor $3 =$ Full Proffesor X_3 = Lama mengajar (tahun) X_4 = Gelar $1 =$ Doktor $0 =$ Master X_5 = Lama memperoleh gelar (tahun) X_6 = Gaji (dolar)
16.	Utama. M.P (2012) Judul : analisis faktor- faktor yang mempengaruhi keputusan kredit usaha rakyat (KUR). Fakultas Ekonomi dan Bisnis Universitas Brawijaya Malang	89	X_1 = Tingkat Pendidikan $1 =$ SD $2 =$ SMP $3 =$ SMA $4 =$ Diploma $5 =$ S1 X_2 = Jumlah tanggungan keluarga X_3 = Lama usaha X_4 = Laba usaha X_5 = Jumlah pinjaman X_6 = Jangka waktu pinjaman
17.	Hosmer, D. W., Lemeshow, S. And Sturdiant, R. X. (2013). <i>Applied Logistic Regression</i> : Third Edition	200	X_1 = Status setelah keluar dari rumah sakit $0 =$ hidup $1 =$ mati X_2 = Umur (Tahun) X_3 = Jenis kelamin $0 =$ Pr $1 =$ Lk X_4 = Ras kulit $0 =$ hitam $1 =$ putih X_5 = Presentase luka bakar (%) X_6 = Adanya gangguan pernafasan akibat kebakaran $0 =$ Tidak $1 =$ Ya

Data	Sumber	n	Variabel
20.	Skripsi Darmawi (2011). Judul: Analisis Faktor- Faktor yang Mempengaruhi Pengambilan Keputusan Petani dalam Pemanfaatan Kredit LKM Prima Tani Fakultas Pertanian Universitas Brawijaya	71	$X_1 =$ Umur $X_2 =$ Tingkat pendidikan 1 = SD 2 = SMP 3 = SMA 4 = SL $X_3 =$ Jumlah anggota keluarga $X_4 =$ Luas kepemilikan lahan $X_5 =$ Pekerjaan sampingan 1 = Memiliki 0 = Tidak

3.2 Metode Analisis

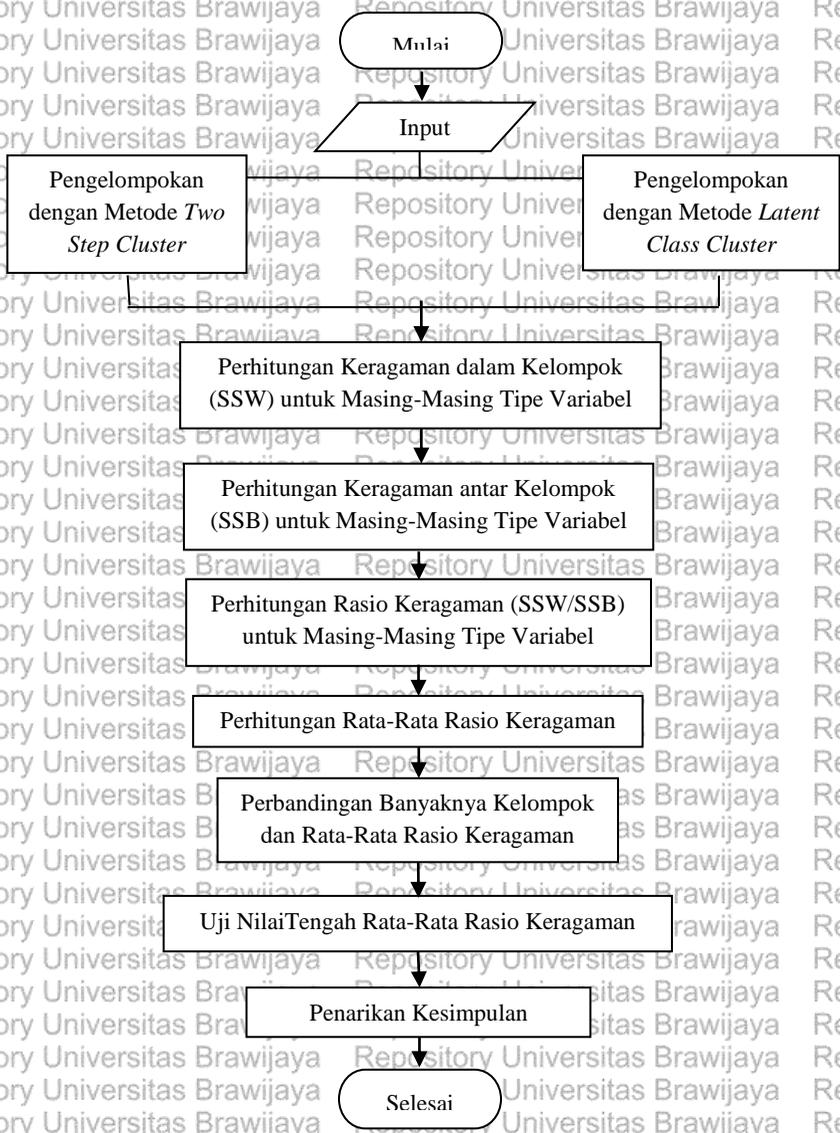
3.2.1 Analisis *Two Step Cluster*

1. Pengelompokan tahap awal (*preclustering*).
2. Pembentukan kelompok akhir menggunakan metode analisis berhirarki dengan prosedur penggabungan (*agglomerative*).
3. Menetapkan banyaknya kelompok yang terbentuk berdasarkan nilai BIC yang terkecil.
4. Melakukan identifikasi keanggotaan kelompok (*cluster membership*).
5. Menghitung keragaman dalam kelompok (SSW) dan antar kelompok (SSB) untuk masing-masing tipe variabel.
6. Menghitung rasio keragaman (SSW/SSB) untuk masing-masing tipe variabel.
7. Menghitung rata-rata dari rasio keragaman data variabel kontinu dan kategorik
8. Membandingkan nilai rata-rata rasio keragaman dan banyaknya jumlah kelompok yang terbentuk dengan analisis *Latent Class Cluster*.
9. Melakukan uji nilai tengah untuk rata-rata rasio keragaman yang diperoleh dengan hasil dari analisis *Latent Class Cluster*.

3.2.2 Analisis *Latent Class Cluster*

1. Membentuk model LC *Cluster*
2. Melakukan pendugaan parameter menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan metode iterasi *Expectation Maximum* (EM) dan *Newton Raphson*.

3. Menetapkan banyaknya kelompok yang terbentuk berdasarkan nilai BIC yang terkecil.
 4. Melakukan identifikasi keanggotaan kelompok (*cluster membership*) berdasarkan model LC *Cluster* yang terbentuk.
 5. Menghitung keragaman dalam kelompok (SSW) dan antar kelompok (SSB) untuk masing-masing tipe variabel.
 6. Menghitung rasio keragaman (SSW/SSB) untuk masing-masing tipe variabel.
 7. Menghitung rata-rata dari rasio keragaman data variabel kontinu dan kategorik.
 8. Membandingkan nilai rata-rata rasio keragaman dan banyaknya jumlah kelompok yang terbentuk dengan analisis *Two Step Cluster*.
 9. Melakukan uji nilai tengah untuk rata-rata rasio keragaman yang diperoleh dengan hasil dari analisis *Two Step Cluster*.
- Tahap analisis data tersebut dapat digambarkan dalam diagram alir sebagaimana tertera dalam Gambar 3.1



Gambar 3.1 Diagram Alir Perbandingan Analisis *Two Step Cluster* dan Analisis *Latent Class Cluster*

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil Analisis *Two Step Cluster*

Pengelompokan pada analisis *Two Step Cluster* dilakukan dengan menggunakan *software* SPSS v18. Banyaknya kelompok yang terbentuk dengan menggunakan analisis *Two Step Cluster* disajikan pada Tabel 4.1 dan dapat dilihat pada Lampiran 2

Tabel 4.1 Hasil Analisis *Two Step Cluster*

Data	Banyaknya Kelompok yang Terbentuk	Rata-Rata Rasio Keragaman
1	3	0,484
2	2	2,209
3	3	0,613
4	2	0,542
5	2	1,242
6	2	2,277
7	3	0,664
8	2	0,873
9	3	0,827
10	2	3,965
11	2	1,414
12	2	0,845
13	4	2,151
14	3	0,765
15	2	0,344
16	2	0,061
17	4	0,722
18	3	0,528
19	3	3,506
20	3	1,977

dari Tabel 4.1 dapat dilihat bahwa banyaknya kelompok yang terbentuk pada 20 data yang dianalisis dengan menggunakan analisis *Two Step Cluster* hanya berkisar antara 2-4 kelompok saja.

Berdasarkan dari banyaknya kelompok yang terbentuk tersebut, pada setiap datanya kemudian dilakukan perhitungan keragaman, baik keragaman dalam kelompok ($SSW / \text{Sum of Square Within Cluster}$) maupun keragaman antar kelompok ($SSB / \text{Sum of Square Between Cluster}$) dengan menggunakan persamaan (2.12), sampai persamaan (2.15). Nilai rasio keragaman diperoleh dengan cara membagi nilai SSW dengan SSB untuk masing-masing skala variabel (kategorik dan kontinu). Nilai rasio yang didapatkan untuk masing-masing skala variabel tersebut kemudian dirata-rata sehingga diperoleh nilai rata-rata rasio keragaman seperti pada Tabel 4.1. Nilai keragaman untuk masing-masing tipe variabel pada analisis *Two Step Cluster* dapat dilihat pada Lampiran 3.

Nilai rata-rata rasio keragaman yang kecil mengindikasikan *cluster* yang baik. Dari Tabel 4.1 dapat dilihat nilai rata-rata rasio keragaman paling kecil sebesar 0,061 yang merupakan rata-rata rasio keragaman pada data ke-16, sedangkan rata-rata rasio keragaman paling besar yaitu 3,965 merupakan rata-rata rasio keragaman pada data ke-10. Sehingga dapat diketahui bahwa dari 20 data yang pada penelitian ini, pengelompokan yang paling baik dengan menggunakan analisis *Two Step Cluster* adalah pada data ke-16.

4.2 Hasil Analisis *Latent Class Cluster*

Pengelompokan pada analisis *Latent Class Cluster* dilakukan dengan menggunakan *software* Latent GOLD v4.0. Penentuan banyaknya kelompok yang terbentuk dapat ditentukan dengan memilih nilai BIC terkecil, banyaknya kelompok yang terbentuk pada analisis *Latent Class Cluster* disajikan pada Tabel 4.2 dan dapat dilihat pada Lampiran 4, dari Tabel 4.2 dapat dilihat bahwa banyaknya kelompok yang terbentuk pada 20 data yang dianalisis dengan menggunakan analisis *Latent Class Cluster* lebih beragam jika dibandingkan dengan menggunakan analisis *Two Step cluster*, di mana banyaknya kelompok yang terbentuk berkisar antara 2-9 kelompok.

Berdasarkan banyaknya kelompok yang terbentuk pada *Latent Class Cluster*, pada setiap datanya kemudian dilakukan perhitungan keragaman, baik keragaman dalam kelompok (SSW) maupun keragaman antar kelompok (SSB). Nilai keragaman yang diperoleh tersebut kemudian dihitung rasionya dengan cara membagi nilai SSW

dengan *SSB* untuk masing-masing tipe variabel. Hasil rasio untuk masing-masing tipe variabel tersebut kemudian dirata-rata sehingga diperoleh nilai rata-rata rasio keragaman seperti pada Tabel 4.2. Nilai keragaman masing-masing tipe variabel dengan menggunakan analisis *Latent Class Cluster* dapat dilihat pada Lampiran 5.

Tabel 4.2 Hasil analisis *Latent Class Cluster*

Data	Banyaknya Kelompok yang Terbentuk	Rata-Rata Rasio Keragaman
1	6	0,195
2	3	0,398
3	3	0,694
4	9	0,174
5	4	0,273
6	3	0,228
7	2	0,819
8	4	1,293
9	4	0,561
10	3	0,422
11	4	0,377
12	2	0,391
13	4	2,874
14	2	0,767
15	4	0,364
16	8	0,019
17	7	0,289
18	4	0,201
19	2	1,937
20	3	0,178

Nilai rata-rata rasio keragaman yang kecil mengindikasikan *cluster* yang baik. Dari Tabel 4.2 dapat dilihat nilai rata-rata rasio keragaman paling kecil sebesar 0,0197 yang merupakan rata-rata rasio keragaman data ke-16, sedangkan rata-rata rasio keragaman paling besar yaitu 2,874 merupakan rata-rata rasio keragaman data ke-13. Sehingga dapat diketahui bahwa dari 20 data yang digunakan pada penelitian ini, pengelompokan yang paling baik dengan menggunakan analisis *Latent Class Cluster* adalah pada data ke-16.

4.3 Pembahasan

Hasil pengelompokan dan nilai rata-rata rasio keragaman dengan menggunakan analisis *Two Step Cluster* dan *Latent Class Cluster* secara keseluruhan dapat dilihat pada Tabel 4.3

Tabel 4.3 Hasil Analisis *Two Step Cluster* dan Analisis *Latent Class Cluster*

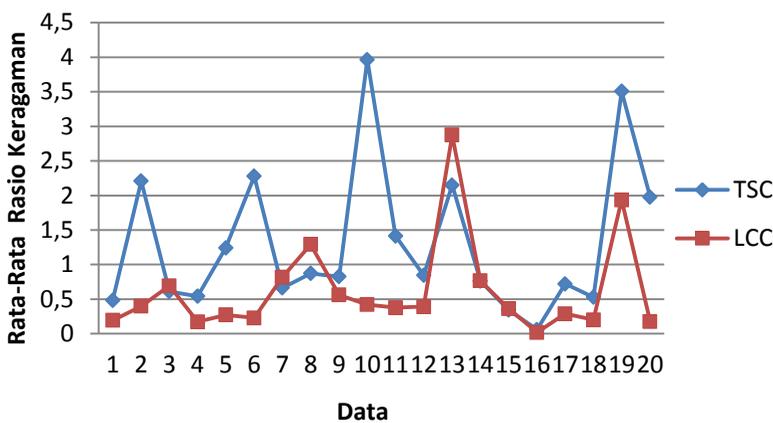
Data	n	Banyak Variabel		Banyak Kelompok		Rata-Rata Rasio Keragaman (SSW/SSB)	
		Kategorik	Kontinu	TSC	LCC	TSC	LCC
1	102	2	4	3	6	0,484	0,195
2	64	2	3	2	3	2,209	0,398
3	79	2	3	3	3	0,613	0,694
4	104	2	6	2	9	0,542	0,174
5	90	2	5	2	4	1,242	0,273
6	58	1	5	2	3	2,277	0,228
7	98	3	4	3	2	0,664	0,819
8	70	3	7	2	4	0,873	1,293
9	66	3	2	3	4	0,827	0,561
10	100	7	2	2	3	3,965	0,422
11	50	1	6	2	4	1,414	0,377
12	30	3	1	2	2	0,848	0,391
13	100	10	5	4	4	2,151	2,874
14	47	3	3	3	2	0,765	0,767
15	52	3	3	2	4	0,344	0,364
16	89	1	5	2	8	0,061	0,019
17	200	4	2	4	7	0,722	0,289
18	150	6	1	3	4	0,528	0,201
19	48	2	3	3	2	3,506	1,937
20	71	2	3	3	3	1,977	0,178

Berdasarkan Tabel 4.3 dapat dilihat bahwa pada data ke-1 nilai rata-rata rasio keragaman (SSW/SSB) dengan menggunakan metode analisis *Two Step Cluster* sebesar 0,484, sedangkan nilai rata-rata rasio keragaman (SSW/SSB) dengan menggunakan metode analisis *Latent Class Cluster* sebesar 0,195. Nilai rata-rata rasio keragaman yang didapatkan pada analisis *Latent Class Cluster* lebih kecil dari analisis *Two Step Cluster*, sehingga dapat disimpulkan bahwa hasil pengelompokan pada data ke-1 lebih baik dengan menggunakan analisis *Latent Class Cluster* dibandingkan ketika menggunakan analisis *Two Step Cluster*.

Nilai rata-rata rasio keragaman (SSW/SSB) pada data ke-3 dengan menggunakan metode analisis *Two Step Cluster* sebesar 0,613, sedangkan nilai rata-rata rasio keragaman (SSW/SSB) dengan menggunakan metode analisis *Latent Class Cluster* sebesar 0,694. Berdasarkan nilai tersebut diketahui bahwa nilai rata-rata rasio keragaman pada analisis *Two Step Cluster* lebih kecil dari nilai rata-rata rasio keragaman yang diperoleh pada analisis *Latent Class Cluster*, sehingga dapat disimpulkan bahwa hasil pengelompokan pada data ke-3 lebih baik ketika menggunakan analisis *Two Step Cluster* dibandingkan ketika menggunakan analisis *Latent Class Cluster*.

Secara keseluruhan, dari 20 data yang digunakan terdapat 14 data pada analisis *Latent Class Cluster* yang memiliki nilai rata-rata rasio keragaman lebih kecil dibandingkan dengan analisis *Two Step Cluster*, untuk mengetahui perbandingan rata-rata rasio keragaman dari masing-masing metode analisis dapat dilihat berdasarkan grafik pada Gambar 4.1

Berdasarkan Grafik pada Gambar 4.1 dapat diketahui bahwa dari 20 data yang dianalisis, sebagian besar memiliki nilai rata-rata rasio keragaman yang lebih kecil ketika menggunakan analisis *Latent Class Cluster (LCC)* dibandingkan dengan analisis *Two Step Cluster (TSC)*. Pada analisis *Two Step Cluster* hanya terdapat 6 data yang nilai rata-rata rasio keragamannya lebih kecil dibandingkan dengan analisis *Latent Class Cluster*, yaitu pada data ke-3, 7, 8, 13, 14, dan 15.



Gambar 4.1 Grafik rata-rata rasio keragaman pada analisis *Two Step Cluster (TSC)* dan *Latent Class Cluster (LCC)*

4.3.1 Uji Nilai Tengah Rata-Rata Rasio Keragaman

Dari hasil yang telah diperoleh dan didapatkan nilai rata-rata rasio keragaman dari 20 data yang digunakan untuk masing-masing metode analisis pengelompokan pada penelitian ini, diketahui terdapat 14 data yang nilai rata-rata rasio keragamannya lebih kecil (lebih baik) ketika menggunakan analisis *Latent Class Cluster*, dan hanya 6 data yang nilai rata-rata rasio keragamannya lebih kecil (lebih baik) ketika menggunakan analisis *Two Step Cluster*.

Untuk mengetahui apakah memang secara signifikan terdapat perbedaan dari rata-rata rasio keragaman yang dihasilkan pada pengelompokan data berskala campuran dengan menggunakan analisis *Two Step Cluster* dan analisis *Latent Class Cluster* maka dilakukan uji dengan menggunakan nilai rata-rata rasio keragaman yang telah diperoleh tersebut. Nilai rata-rata rasio keragaman yang diperoleh pada analisis *Two Step Cluster* dan analisis *Latent Class Cluster* diuji dengan menggunakan uji dua nilai tengah, sehingga dari hasil pengujian tersebut akan dapat diketahui apakah secara signifikan terdapat perbedaan berdasarkan nilai rata-rata rasio keragamannya.

Sebelum dilakukan uji nilai tengah, terlebih dahulu perlu diuji apakah ragam untuk rata-rata rasio keragaman pada analisis *Two Step*

Cluster sama dengan analisis *Latent Class Cluster*, untuk itu digunakan hipotesis seperti berikut:

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

menggunakan *software* Minitab dengan digunakan α sebesar 0,05 diperoleh nilai *p-value* sebesar 0,071. Nilai *p-value* tersebut lebih besar dari α yang digunakan, sehingga dapat diambil keputusan untuk menerima H_0 dan disimpulkan bahwa ragam dari rata-rata rasio keragaman untuk *Two Step Cluster* dan *Latent Class Cluster* pada 20 data yang digunakan dalam penelitian ini adalah sama. *Output* dari hasil analisis ragam ini dapat dilihat pada Lampiran 6.

Dengan diperoleh kesimpulan bahwa ragam rata-rata rasio keragaman dari kedua metode tersebut sama, maka untuk uji nilai tengah pada rata-rata rasio keragaman analisis *Two Step Cluster* dan analisis *Latent Class Cluster* ini dapat digunakan rumus (2.17) dalam perhitungannya, dengan menggunakan *software* Minitab dan digunakan α sebesar 0,05. Untuk mengetahui apakah memang dari 20 data yang digunakan pada penelitian ini secara signifikan terbukti bahwa rata-rata rasio keragaman yang dihasilkan pada analisis *Two Step Cluster* tidak memberikan hasil yang lebih baik dari analisis *Latent Class Cluster*, maka digunakan hipotesis seperti berikut:

$$H_0 : \mu_{TSC} = \mu_{LCC}$$

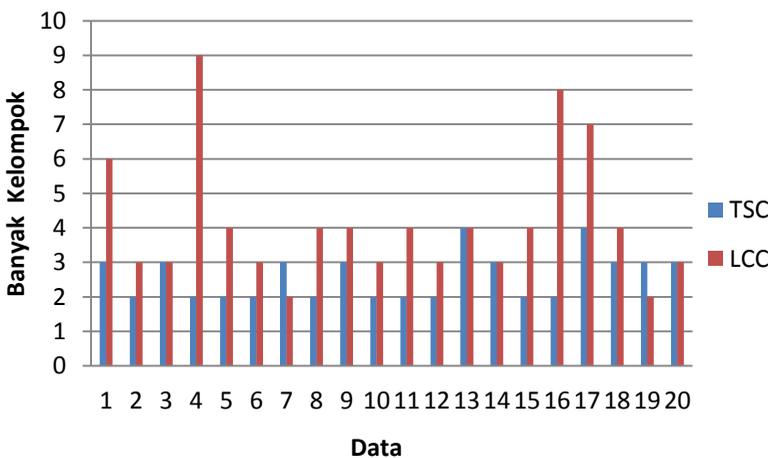
$$H_1 : \mu_{TSC} > \mu_{LCC}$$

Tabel 4.4 Hasil Uji Nilai Tengah pada Rata-Rata Rasio Keragaman *Two Step Cluster (TSC)* dan *Latent Class Cluster (LCC)*

Analisis	Rata-Rata dari 20 Rata-Rata Rasio Keragaman (SSW/SSB)	Nilai Statistik Uji	P-Value	Keputusan
LCC	0,6227	2,40	0,011	Terima H_1
TSC	1,3006			

Berdasarkan hasil analisis seperti yang ditampilkan pada Tabel 4.4 diperoleh nilai *p-value* sebesar 0,011, nilai *p-value* tersebut lebih kecil dari α yang digunakan, sehingga dapat diambil keputusan untuk menerima H_1 . *Output* dari hasil analisis uji nilai tengah pada 20 rata-

rata rasio keragaman untuk analisis *Latent Class Cluster* dan Analisis *Two Step Cluster* dapat dilihat pada Lampiran 7. dengan keputusan untuk menerima H_1 , maka disimpulkan bahwa nilai rata-rata rasio keragaman yang diperoleh pada analisis *Two Step Cluster* lebih besar dari analisis *Latent Class Cluster*, sehingga dapat diketahui bahwa analisis *Two Step Cluster* tidak memberikan hasil pengelompokan yang lebih baik dari analisis *Latent Class Cluster* atau dapat disimpulkan bahwa dari 20 data yang digunakan pada penelitian ini didapatkan hasil pengelompokan untuk data berskala campuran pada analisis *Latent Class Cluster* lebih baik dibandingkan dengan menggunakan analisis *Two Step Cluster*. Hal tersebut dimungkinkan karena pada analisis *Latent Class Cluster* pengelompokan objek didasarkan pada peluang keanggotaannya yang dihitung secara langsung dari model, sehingga sifat pengelompokannya lebih teliti dan lebih akurat dibandingkan dengan analisis *Two Step Cluster* yang dalam pengelompokannya hanya didasarkan pada perhitungan jarak.



Gambar 4.2 Grafik banyaknya kelompok pada analisis *Two Step Cluster* dan *Latent Class Cluster*

Banyaknya kelompok yang diperoleh dari kedua analisis tersebut juga berbeda, pada analisis *Two Step Cluster* banyaknya kelompok yang diperoleh hanya berkisar pada 2 hingga 4 kelompok saja, sedangkan pada analisis *Latent Class Cluster* banyaknya

kelompok yang diperoleh lebih beragam yaitu antara 2 hingga 9 kelompok. Terdapat 4 data yang memiliki banyak kelompok sama ketika menggunakan analisis *Latent Class Cluster* maupun dengan analisis *Two Step Cluster*, yaitu pada data ke-3 sebanyak 3 cluster, data ke-13 sebanyak 4 cluster, data ke-14 sebanyak 3 cluster, dan data ke-20 sebanyak 3 cluster.

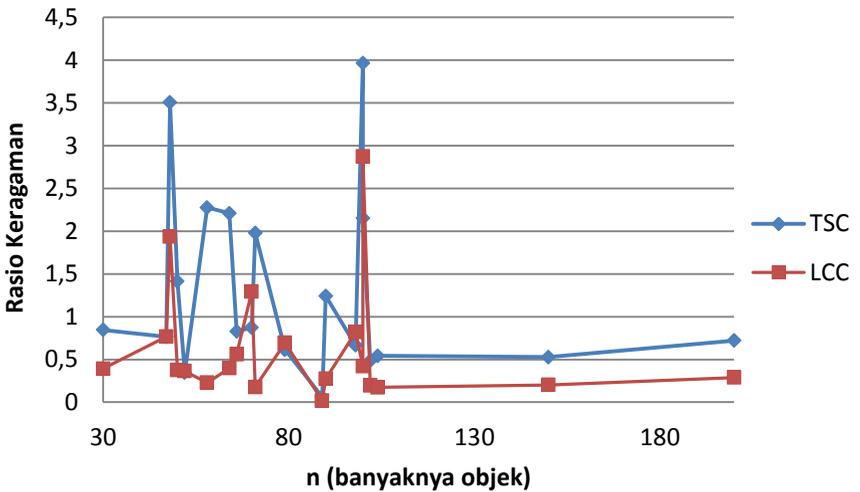
Pada Tabel 4.3 terdapat keterangan mengenai banyaknya variabel kategorik dan variabel kontinu yang digunakan. Berdasarkan dari Tabel tersebut terlihat bahwa tidak terdapat pola yang terbentuk antara perbandingan banyaknya variabel yang digunakan (kontinu dan kategorik) dengan nilai rasio keragaman yang diperoleh, sehingga dapat disimpulkan bahwa dari 20 data yang digunakan dalam penelitian ini perbandingan banyaknya variabel kategorik dan variabel kontinu ketika menggunakan analisis *Two Step Cluster* dan analisis *Latent Class Cluster* tidak berpengaruh terhadap kebaikan suatu pengelompokan atau nilai rasio keragaman yang diperoleh.

Hal tersebut dimungkinkan karena pada perhitungan keragaman untuk mengetahui kebaikan pengelompokan yang digunakan dalam penelitian ini, jumlah variabel kontinu dan kategorik yang digunakan tidak diperhitungkan dalam rumus keragamannya yaitu pada persamaan (2.12) – (2.15), sehingga berdasarkan rumus keragaman tersebut terlihat bahwa perbandingan jumlah variabel kontinu dan kategorik yang digunakan tidak berpengaruh terhadap nilai rasio keragaman yang dihasilkan baik ketika menggunakan analisis *Two Step Cluster* maupun analisis *Latent Class Cluster*.

Grafik pada Gambar 4.3 merupakan grafik antara banyaknya objek yang digunakan dengan rasio keragaman yang diperoleh. Dari grafik tersebut ingin diketahui pola yang terbentuk dari banyaknya objek yang digunakan dengan nilai rata-rata rasio keragaman, agar dapat diketahui apakah terdapat hubungan atau pengaruh dari banyaknya objek yang digunakan terhadap kebaikan suatu pengelompokan yang dilihat berdasarkan nilai rata-rata rasio keragaman.

Pada gambar tersebut terlihat tidak terdapat pola yang terbentuk, sehingga dapat diketahui bahwa banyaknya data yang digunakan tidak berpengaruh terhadap nilai rasio keragaman atau

baik tidaknya suatu pengelompokan pada analisis *Two Step Cluster* dan analisis *Latent Class Cluster*.



Gambar 4.3 Grafik antara rasio keragaman dan banyaknya objek pada analisis *Two Step Cluster* dan analisis *Latent Class Cluster*

Secara keseluruhan, dari 20 data yang digunakan pada penelitian ini memiliki tujuan yang berbeda-beda untuk pengelompokannya. Sebagian besar data yang digunakan bertujuan untuk mengetahui pemerataan objek dari kasus yang diteliti. Seperti pada data ke-1, dilakukan pengelompokan agar dapat diketahui peta industri tempe di Kota Malang dengan kelebihan dan kekurangan untuk masing-masing kelompok yang terbentuk. Berdasarkan kelompok-kelompok yang terbentuk tersebut dapat ditentukan cara atau strategi yang akan digunakan untuk meningkatkan kualitas produksi industri-industri tempe, sehingga dari hasil yang diperoleh dapat dimanfaatkan untuk peningkatan tingkat perekonomian.

Selain data ke-1, fungsi pengelompokan untuk pemerataan objek juga digunakan pada data ke-2, data ke-4, dan data ke-11. Pada data ke-2 dan data ke-4 tujuan pemerataan objek digunakan untuk kasus pendidikan. Diharapkan dari pengelompokan yang terbentuk dapat diketahui kelebihan dan kekurangan dari masing-masing kelompok, sehingga berdasarkan informasi tersebut dapat ditentukan

BAB V KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan dari hasil dan pembahasan, dapat diperoleh kesimpulan sebagai berikut :

1. Terdapat perbedaan banyaknya kelompok yang terbentuk pada analisis *Two Step Cluster* dan analisis *Latent Class Cluster*. Dari 20 data yang digunakan, hanya terdapat 4 data yang menghasilkan banyaknya kelompok yang sama dengan menggunakan analisis *Two Step Cluster* dan *Latent Class Cluster*, sedangkan 16 data lainnya memiliki hasil pengelompokan yang berbeda.

2. Nilai rata-rata dari 20 rata-rata rasio keragaman (SSW/SSB) untuk analisis *Two Step Cluster* sebesar 1,3006 dan untuk analisis *Latent Class Cluster* sebesar 0,6227, sehingga dapat disimpulkan bahwa analisis *Latent Class Cluster* menghasilkan pengelompokan yang lebih baik dari *Two Step Cluster*.

5.2 Saran

Pada pengelompokan data berskala campuran (kategorik dan kontinu) untuk mendapatkan hasil yang lebih baik disarankan dengan menggunakan analisis *Latent Class Cluster*. Pada penelitian selanjutnya disarankan untuk membandingkan kedua metode tersebut menggunakan data simulasi dengan berbagai macam karakteristik data dan variabel yang digunakan sehingga akan lebih banyak informasi yang bisa diperoleh.

DAFTAR PUSTAKA

- Chan, Y.H. 2005. *Cluster Analysis*. Singapore. Med J. Singapore. <http://www.sma.org.sg/smj/4604/4604bs1.pdf>. Diakses pada tanggal 2 Agustus 2013
- Collet. 1994. *Modelling Survival Data In Medical Research, Second Edition*. Chapman & Hall
- Darmawi. 2011. *Analisis Faktor-Faktor yang Mempengaruhi Pengambilan Keputusan Petani dalam Pemanfaatan Kredit Lembaga Keuangan Mikro (LKM) Prima Tani*. Tugas Akhir. Fakultas Pertanian. Universitas Brawijaya
- Dewi, A., Sutikno dan Kuswanto, H. 2012. *Metode Cluster Ensembl untuk Pengelompokan Desa Perdesaan di Provinsi Riau*. Tugas Akhir. Institut Teknologi Sepuluh November (ITS)
- Dieta. F. 2012. *Analisis Faktor-Faktor yang Mempengaruhi Keputusan Bank dalam Pemberian Kredit Usaha Mikro (MKM) (Studi Kasus pada PT. Bank X Tbk. Kantor Cabang Probolinggo)*. Tugas Akhir. Fakultas Ekonomi. Universitas Brawijaya Malang
- Eveline. 2002. *Hubungan Antara Faktor Resiko dengan Derajat Klinis Skabies di RSUD Dr. Saiful Anwar Kota Malang*. Tugas Akhir. Fakultas Kedokteran. Universitas Brawijaya Malang
- Hermayudi. W.A. 2004. *Pengelompokan Industri Tempe Berdasarkan Klasifikasi Industri dengan Cluster Analysis (Studi Kasus di Sentra Industri Tempe Kota Malang)*. Tugas Akhir. Fakultas Teknik Pertanian. Universitas Brawijaya Malang
- Hosmer, D. W., Lemeshow, S. And Sturdiart, R. X. 2013. *Applied Logistic Regression, 3rd Edition*. John Wiley & Son, Inc

Isnayni, N. 2011. *Analisis Kualitas Angsuran Kredit Peserta Program Kredit Usaha Rakyat (KUR)*. Tugas Akhir. Fakultas Ekonomi. Universitas Brawijaya Malang

Iswenda, N. 2007. *Faktor-Faktor Penentu Tingkat Gangguan Jiwa Pasien Rumah Sakit Jiwa Daerah Dr. Rm. Soedjarwadi di Kabupaten Klaten, Jawa Tengah*. Tugas Akhir. Sekolah Tinggi Ilmu Statistika

Johnson, R.A. and D.W. Wincern. 2007. *Applied Multivariate Statistical analysis*. Prentice Hall Inc: New Jersey.

Marina, L. 2009. *Pengaruh Atribut Komite Audit dan Financial Distress Terhadap Kualitas Pelaporan Keuangan (Studi Empiris pada Perusahaan yang Terdaftar di BEI Tahun 2009)*. Tugas Akhir. Fakultas Ekonomi. Universitas Brawijaya Malang

Moustaki, I. and Papageorgiou, I. 2004. *Latent class models for mixed outcomes with applications in Archaeometry*. Computational Statistics and Data Analysis

Nazhiirah, A. 2010. *Perbandingan Penduga ML dan Bayessian MCMC pada Model Regresi Logistik Multilevel*. Tugas Akhir. Fakultas MIPA. Universitas Brawijaya Malang

Niko, R. 2011. *Analisis Karakteristik Debitur Kredit UMKM Sebagai Upaya Mengurangi Resiko Munculnya Kredit Macet*. Tugas Akhir. Fakultas Ekonomi. Universitas Brawijaya Malang

Oktavina, C. 2013. *Analisis Produktivitas Wanita Pekerja Sektor Informal (Studi Kasus: Pedagang di Kota Malang)*. Tugas Akhir. Fakultas Ekonomi. Universitas Brawijaya Malang

Prehatiningsih, 2002. *Perbandingan Profil Tekanan Darah Serangan Infark Serebral Akut dan Infark Miokard Akut dari Penderita Hipertensi*. Tugas Akhir. Fakultas Kedokteran. Universitas Brawijaya Malang

Sahri, M. 2011. *Variabel-Variabel yang Mempengaruhi Pendapatan Petani Apel (Studi kasus: Petani Apel di Desa Bumiaji)*. Tugas Akhir. Fakultas Ekonomi. Universitas Brawijaya Malang

Serban, G, and Grigoreta SM. 2006. *A Comparison of Clustering Techniques In Aspect Mining*. Studia Univ. Babeş-Bolyai, Informatica, Volume LI

Suryani, N. 2011. *Penerapan Model Regresi ZIGP pada Banyaknya Siswa Gagal Ujian Nasional di SMA/SMK Kabupaten Malang*. Tugas Akhir. Fakultas MIPA. Universitas Brawijaya Malang

Theodoris, S and K. Koutroumbas. 2003. *Pattern Recognition Second Edition*. Academic Press: San Diego.

Utama, M.P. 2012. *Analisis Faktor-Faktor yang Mempengaruhi Keputusan Kredit Usaha Rakyat (KUR)*. Tugas Akhir. Fakultas Ekonomi dan Bisnis Universitas Brawijaya Malang

Vermunt J.K and Magidson J. 2005. *Latent Gold's User's Guide*. Statistical Innovations, Inc: Boston.

Wahyu, B. 2007. *Analisis Keterkaitan Antara Tingkat Kepuasan, Motivasi, Loyalitas, dan Kinerja Guru SMA (Studi Kasus pada Guru SMA Negeri dan Swasta di Kota Mojokerto Tahun Ajaran 2006/2007)*. Tugas Akhir Sekolah Tinggi Ilmu Statistik (STIS)

Wahyu, I. 2004. *Hubungan Tingkat Pengetahuan Keluarga tentang Pencegahan Penularan Tuberculosis Paru dengan Jumlah Anggota Keluarga Yang Tertular TBC Paru di Rumah Sakit Paru Bat. Tugas Akhir.* Fakultas Kedokteran, Universitas Brawijaya

Walpole, R., Myers, R., Myers, S., Ye, K. 2011. *Probability & Statistics For Engineers & Scientists, Ninth Edition.* Pretince Hall Inc: New Jersey.

Weisberg, S. 1985. *Applied Linier Regression, Second Edition.* John Wiley and Sons: New York

Lampiran 1. Data Sekunder yang digunakan dalam Analisis

1. Data 1

No	X1	X2	X3	X4	X5	X6
1	3	82125000	18250	2150000	2	2
2	4	142350000	31025	1750000	1	1
3	4	65700000	14600	1000000	1	1
4	3	25185000	5475	500000	1	1
5	3	67160000	14600	1950000	2	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
99	3	100740000	21900	2625000	2	1
100	2	98550000	21900	2625000	2	1

X_1 = Jumlah tenaga kerja (orang)
(rupiah/tahun)

X_2 = Nilai produksi

X_3 = Jumlah bahan baku (kg/thn)

X_4 = Nilai investasi (Rp)

X_5 = Teknologi yang digunakan

1 = alat modern 2 = alat tradisional

X_6 = Aneka olahan

1 = Juga memproduksi selain tempe (keripik)

2 = Hanya memproduksi tempe

2. Data 2

No	X1	X2	X3	X4	X5
1	2	1	1,23	0,15	23,6
2	2	1	1,04	0,21	11,83
3	1	1	1,03	0,13	29,89
⋮	⋮	⋮	⋮	⋮	⋮
63	4	1	0	0,58	8,67
64	4	1	0	0,36	15,67

X_1 = Akreditasi sekolah

1 = A 2 = B

3 = C

4 = Belum terakreditasi

X_2 = Waktu penyelenggaraan sekolah

1 = Pagi 2 = Siang

3 = Campuran

Lampiran 1. (Lanjutan)

X_3 = Rasio penerimaan dan pendaftaran murid

X_4 = Rasio guru dan

X_5 = Rasio murid dan kelas

3. Data 3

No	X1	X2	X3	X4	X5
1	1	255	230	200	1
2	1	250	215	195	0
3	1	245	215	195	1
4	1	255	195	190	0
5	1	245	215	190	1
⋮	⋮	⋮	⋮	⋮	⋮
78	3	225	170	140	1
79	3	200	120	130	0

X_1 = Asal TK

1 = Al Hidayah I

2 = Indria Tama

3 = Al Hidayah II

X_2 = Skor pemilihan warna

X_3 = Skor kerapian

X_4 = Skor kebersihan

X_5 = Jenis kelamin

0 = Perempuan

1 = Laki-Laki

4. Data 4

No	X1	X2	X3	X4	X5	X6	X7	X8
1	0	18	1647	50	62	1	22,56	67,5
2	0	30	114	30	20	1	21,34	66
3	0	37	488	25	35	1	20,54	66
4	1	6	232	16	42	1	19,58	69,9
5	0	18	127	19	61	1	20,55	70,2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
103	0	22	75	21	25	0	22,81	69,9
104	0	24	184	29	40	1	21,67	70,2

Lampiran 1. (Lanjutan)

X_1 = Status sekolah

1 = Negeri

0 = Swasta

X_2 = Lama sekolah berdiri (tahun)

X_3 = Banyaknya siswa terakhir

X_4 = Banyaknya guru terakhir

X_5 = Banyaknya alumni yang diterima di dunia usaha dan industri pada setahun terakhir

X_6 = Status tanah bangunan

1 = Milik sendiri

0 = Menyewa

X_7 = Nilai rata-rata ujian nasional sekolah

X_8 = IPM tiap Kabupaten/Kota asal sekolah

5. Data 5

No	X1	X2	X3	X4	X5	X6	X7
1	3	2	5000000	24	17500000	2400000	1
2	1	1	3000000	24	6500000	1500000	0
3	1	2	5000000	24	7500000	1800000	0
4	1	1	4000000	24	7500000	1500000	0
5	1	2	5000000	24	18750000	2550000	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
89	2	3	10000000	24	18750000	2400000	0
90	2	3	5000000	36	12000000	1200000	0

X_1 = Tingkat Pendidikan

1 = SD

2 = SLTP

3 = SLTA

4 = PT

X_2 = Jumlah tanggungan keluarga

X_3 = Jumlah pinjaman (Juta rupiah)

X_4 = Jangka waktu (bulan)

X_5 = Omzet (Juta rupiah)

X_6 = Laba (Juta rupiah)

X_7 = Pinjaman lain

1 = Ada

0 = Tidak ada

Lampiran 1. (Lanjutan)

6. Data 6

No	X1	X2	X3	X4	X5	X6
1	3	100	1	14	0,978	16,265
2	3	100	1	1	0,684	15,111
3	3	100	1	1	0,988	14,634
4	3	100	1	2	0,703	12,242
:	:	:	:	:	:	:
57	3	100	1	12	0,437	13,811
58	4	100	1	5	0,403	15,828

X_1 = Jumlah anggota di dalam komite audit

X_2 = Independensi komite audit (%)

X_3 = Kompetensi komite audit

1 = Terdapat anggota komite audit yang berlatar belakang pendidikan dan berpengalaman di bidang keuangan

0 = Tidak terdapat anggota komite audit yang berlatar belakang pendidikan dan berpengalaman di bidang keuangan

X_4 = Frekuensi pertemuan komite audit dalam 1 tahun

X_5 = Tingkat kondisi (kesulitan) keuangan

X_6 = Ukuran perusahaan (diukur dengan natural logaritma total aset)

7. Data 7

No	X1	X2	X3	X4	X5	X6	X7
1	35	45113452	25000000	1	3	1	10
2	24	112312785	47000000	0	4	1	2
3	43	14740000	8000000	1	3	1	15
4	40	211940700	175000000	1	2	1	12
5	26	74589000	50000000	1	2	1	3
:	:	:	:	:	:	:	:
97	40	211940700	175000000	1	2	1	12
98	26	36958421	20000000	1	4	1	2

Lampiran 1. (Lanjutan)

X_1 = Usia (tahun)

X_2 = Laba pertahun (omzet) (Juta rupiah)

X_3 = Taksasi jaminan (rupiah)

X_4 = Status Agunan

1 = Marketable

0 = Unmarketable

X_5 = Tingkat Pendidikan

1 = SD 2 = SLTP

3 = SLTA 4 = PT

X_6 = Jenis Usaha

1 = Perdagangan

0 = Non perdagangan

X_7 = Lama usaha (Tahun)

8. Data 8

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
1	0	1	1	3	6	15000000	24	18000000	3200000	27500000
2	0	3	0	4	6	10000000	24	11000000	2200000	22500000
3	1	2	0	4	6	15000000	24	14000000	2200000	32000000
4	0	1	1	3	7	7000000	24	15000000	2400000	15000000
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
69	1	2	1	2	7	15000000	36	70000000	3600000	40000000
70	0	2	0	3	7	10000000	36	12000000	2400000	28500000

X_1 = Jenis kelamin

0 = Laki-laki

1 = Perempuan

X_2 = Tingkat pendidikan

1 = SD 2 = SLTP

3 = SLTA 4 = PT

X_3 = Status nasabah

1 = Nasabah lama

0 = Nasabah baru

X_4 = Jumlah tanggungan dalam keluarga

X_5 = Lama usaha yang digeluti debitur (tahun)

X_6 = Besar pinjaman (rupiah)

X_7 = Jangka waktu pengembalian kredit yang disepakati (bulan)

X_8 = Omzet per bulan (juta rupiah)

X_9 = Laba (rupiah)

X_{10} = Jaminan kredit (rupiah)

Lampiran 1. (Lanjutan)

9. Data 9

No	X1	X2	X3	X4	X5
1	14	4	2	18	0
2	19	5	1	3	0
3	16	4	1	20	0
4	7	3	2	2	0
5	3	1	1	3	0
⋮	⋮	⋮	⋮	⋮	⋮
65	29	5	2	2	1
66	33	6	3	1	1

X_1 = Usia (tahun)

X_2 = Tingkat pendidikan

1 = Belum Sekolah

2 = TK

3 = SD

4 = SMP

5 = SMA

6 = PT

X_3 = Tingkat sosial ekonomi dengan kategori:

1 = Rendah (< Rp. 500.000)

2 = Menengah (antara Rp. 500.000 s.d Rp. 1.000.000)

3 = Tinggi (> Rp. 1.000.000)

X_4 = Jumlah teman sekamar

X_5 = Higieni pribadi

0 = Buruk

1 = Baik

10. Data 10

No	X1	X2	X3	X4	X5	X6	X7	X8	X9
1	58	1	0	1	1	1	1	0	0
2	41	1	0	1	0	1	1	0	0
3	52	1	0	1	1	1	1	0	0
4	44	1	0	2	1	1	1	0	0
5	38	1	0	1	0	1	3	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
99	52	1	0	1	0	1	7	1	0
100	47	0	0	2	1	1	6	1	0

Lampiran 1. (Lanjutan)

X_1 = Umur (tahun)

X_2 = Jenis kelamin

0 = Laki-laki

1 = Perempuan

X_3 = Status tempat tinggal

0 = Desa

1 = Kota

X_4 = Tingkat pendidikan

1 = Tidak tamat SD

2 = SD-SMP

3 = \geq SMA

X_5 = Status kerja

0 = Tidak bekerja

1 = Bekerja

X_6 = Frekuensi muncul gejala

0 = Jarang

1 = Sering

X_7 = Frekuensi dirawat

X_8 = Kondisi ekonomi

0 = Buruk

1 = Baik

X_9 = Penyakit keturunan (gen)

0 = Tidak ada

1 = Ada

11. Data 11

No	X1	X2	X3	X4	X5	X6	X7
1	30500000	9000	30	1	53	5	13800000
2	10550000	900	2	1	59	5	3600000
3	76250000	20000	60	1	64	4	18000000
4	27125000	4800	40	2	37	4	9600000
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
49	41500000	8000	58	1	50	4	12000000
50	25000000	4500	25	2	47	4	4200000

X_1 = Modal (rupiah)

X_2 = Luas lahan/m²

X_3 = Jumlah tenaga kerja

X_4 = Tingkat pendidikan

1 = SD

2 = SLTP

3 = SLTA

4 = PT

X_5 = Umur

X_6 = Jumlah Anggota Keluarga

X_7 = Pengeluaran per Musim Panen (rupiah)

Lampiran 1. (Lanjutan)

12. Data 12

No	X1	X2	X3	X4
1	63	1	0	0
2	63	1	0	1
3	54	1	0	0
4	49	1	0	0
:	:	:	:	:
29	41	1	0	0
30	43	1	0	0

X₁ = Usia (tahun)

X₂ = Jenis Kelamin

0 = Perempuan

1 = Laki-laki

X₃ = Hipertensi

0 = Tidak

1 = Ya

X₄ = Diabetes Melitus

0 = Ya

1 = Tidak

13. Data 13

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15
1	50	1	6	12	1	1	1	0	1	0	1	1	1	1	1
2	35	0	12	2	0	0	4	0	0	1	1	6	0	0	1
3	37	0	12	2	1	1	3	0	0	0	0	2	0	0	0
4	45	1	6	15	1	1	1	0	1	0	0	4	1	0	0
5	34	0	12	5	1	1	3	0	0	0	1	2	1	1	1
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
99	55	1	6	13	0	1	0,5	1	1	0	1	0	1	0	1
100	59	1	6	10	1	1	3	1	0	0	1	0	2	0	1

X₁ = Umur (tahun)

X₂ = Status pernikahan

1 = Tidak menikah

2 = Menikah

X₃ = Pendidikan (tahun)

Lampiran 1. (Lanjutan)

X_4 = Pengalaman (tahun)

X_5 = Daerah

0 = Luar Malang

1 = Malang

X_6 = Lokasi

0 = Kurang strategis

1 = Strategis

X_7 = Jarak (Km)

X_8 = Kepemilikan usaha

0 = Milik orang lain

1 = Milik sendiri

X_9 = Status pekerjaan

0 = Luar Malang

1 = Malang

X_{10} = Sistem upah

0 = Bukan bulanan

1 = Bulanan

X_{11} = Jenis dagangan

0 = Non makanan

1 = Makanan

X_{12} = Hub. dengan pemilik

0 = Tidak ada

1 = Ada hubungan

X_{13} = Tanggungan keluarga (orang)

X_{14} = Keamanan

0 = Kurang aman

1 = Aman

X_{15} = Akses kredit

0 = Tidak ada akses

1 = Ada akses

14. Data 14

No	X1	X2	X3	X4	X5	X6
1	32	2	1	1	61	72
2	36	2	2	1	61	67
3	42	1	4	2	72	39
4	38	1	4	2	56	50
:	:	:	:	:	:	:
46	41	2	1	1	72	61
47	23	1	4	1	83	44

X_1 = Umur (tahun)

X_2 = Tingkat pendidikan

1 = SD

2 = SMP

3 = SMA

4 = PT

Lampiran 1. (Lanjutan)

- X_3 = Pekerjaan
 1 = PNS 2 = Swasta 3 = Tidak Bekerja 4 = Lain-lain
 X_4 = Pernah Mendapat Informasi Pencegahan Penularan TBC
 1 = Pernah 2 = Tidak Pernah
 X_5 = Nilai pengetahuan (%) X_6 = Nilai penerapan (%)

15.Data 15

No	X1	X2	X3	X4	X5	X6
1	1	3	25	1	35	36350
2	1	3	13	1	22	35350
3	1	3	10	1	23	28200
4	0	3	7	1	27	26775
⋮	⋮	⋮	⋮	⋮	⋮	⋮
51	0	1	1	1	1	15000
52	0	1	0	1	2	20300

X_1 = Jenis kelamin

1 = Laki-laki 0 = Perempuan

X_2 = Jabatan

1 = Assistant Prof 2 = Associate Prof 3 = Full Prof

X_3 = Lama mengajar (tahun)

X_4 = Gelar

1 = Doktor 0 = Master

X_5 = Lama memperoleh gelar (tahun)

X_6 = Gaji (dolar)

16.Data 16

No	X1	X2	X3	X4	X5	X6
1	3	3	3	750000	5000000	12
2	1	1	5	750000	5000000	12
3	1	0	10	450000	5000000	12
⋮	⋮	⋮	⋮	⋮	⋮	⋮
88	3	5	4	250000	5000000	18
89	1	7	1	700000	5000000	12

Lampiran 1. (Lanjutan)

X_1 = Tingkat Pendidikan

1 = SD

2 = SMP

3 = SMA

4 = diploma

5 = S1

X_2 = Jumlah tanggungan keluarga

X_5 = Jumlah pinjaman

X_3 = Lama usaha

X_6 = Waktu pinjaman

X_4 = Laba usaha

17. Data 17

No	X1	X2	X3	X4	X5	X6
1	0	26	1	1	25,3	0
2	0	2	0	0	5	0
3	0	22	0	0	2	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
199	1	63	1	1	36	0
200	1	60	0	0	31	1

X_1 = Status setelah keluar dari rumah sakit

0 = hidup

1 = mati

X_2 = Umur (Tahun)

X_3 = Jenis kelamin

0 = perempuan

1 = laki-laki

X_4 = Ras kulit

0 = hitam

1 = putih

X_5 = Presentase luka bakar (%)

X_6 = Adanya gangguan pernafasan akibat kebakaran

0 = Tidak

1 = Ya

18. Data 18

No	X1	X2	X3	X4	X5	X6	X7
1	1	23	2	1	1	2	2
2	1	23	2	1	2	1	3
3	1	24	2	1	2	1	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
149	2	59	3	3	1	2	1
150	1	74	1	3	2	1	3

Lampiran1. (Lanjutan)
20. Data 20

No	X1	X2	X3	X4	X5
1	46	1	4	0,25	1
2	52	1	7	0,5	1
3	48	1	6	1	1
4	43	1	4	1	1
:	:	:	:	:	:
70	23	1	2	0,25	0
71	31	1	3	0,06	0

$X_1 =$

$X_2 =$

$X_3 =$

$X_4 =$

$X_5 =$

Umur

Tingkat pendidikan

1 = SD 2 = SMP 3 = SMA 4 = SL

Jumlah anggota keluarga

Luas kepemilikan lahan

Pekerjaan sampingan

1 = Memiliki 0 = Tidak memiliki

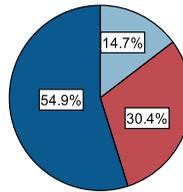
Lampiran 2. Output Analisis *Two Step Cluster*

1. Data 1

Model Summary

Algorithm	TwoStep
Input Features	6
Clusters	3

Cluster Sizes



Cluster
1
2
3

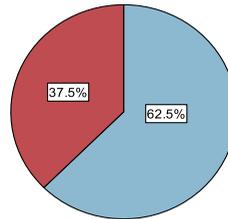
Size of Smallest Cluster	15 (14.7%)
Size of Largest Cluster	56 (54.9%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	3.73

2. Data 2

Model Summary

Algorithm	TwoStep
Input Features	5
Clusters	2

Cluster Sizes



Cluster
1
2

Size of Smallest Cluster	24 (37.5%)
Size of Largest Cluster	40 (62.5%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.67

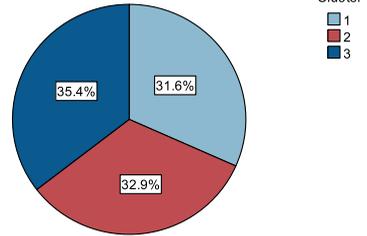
Lampiran 2. (Lanjutan)

3. Data 3

Model Summary

Algorithm	TwoStep
Input Features	5
Clusters	3

Cluster Sizes



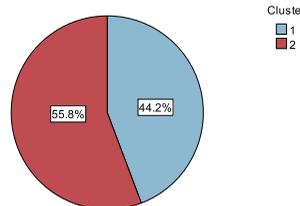
Size of Smallest Cluster	25 (31.6%)
Size of Largest Cluster	28 (35.4%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.12

4. Data 4

Model Summary

Algorithm	TwoStep
Input Features	8
Clusters	2

Cluster Sizes



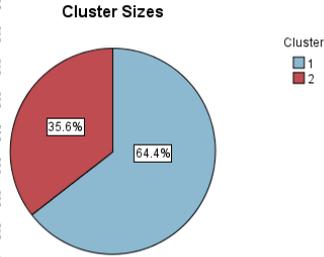
Size of Smallest Cluster	46 (44.2%)
Size of Largest Cluster	58 (55.8%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.26

Lampiran 2. (Lanjutan)

5. Data 5

Model Summary

Algorithm	TwoStep
Input Features	7
Clusters	2

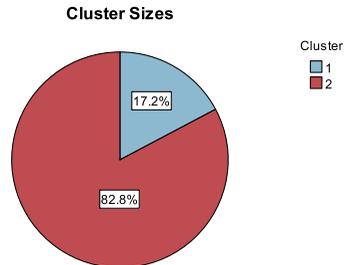


Size of Smallest Cluster	32 (35.6%)
Size of Largest Cluster	58 (64.4%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.81

6. Data 6

Model Summary

Algorithm	TwoStep
Input Features	6
Clusters	2



Size of Smallest Cluster	10 (17.2%)
Size of Largest Cluster	48 (82.8%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	4.80

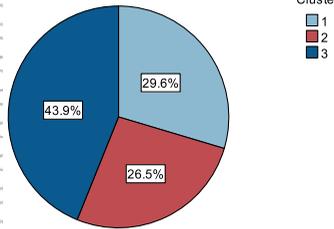
Lampiran 2. (Lanjutan)

7. Data 7

Model Summary

Algorithm	TwoStep
Input Features	7
Clusters	3

Cluster Sizes



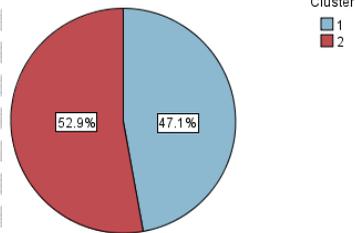
Size of Smallest Cluster	26 (26.5%)
Size of Largest Cluster	43 (43.9%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.65

8. Data 8

Model Summary

Algorithm	TwoStep
Input Features	10
Clusters	2

Cluster Sizes



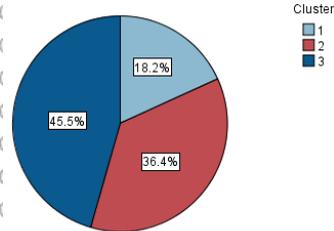
Size of Smallest Cluster	33 (47.1%)
Size of Largest Cluster	37 (52.9%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.12

Lampiran 2. (Lanjutan) 9. Data 9

Model Summary

Algorithm	TwoStep
Input Features	5
Clusters	3

Cluster Sizes



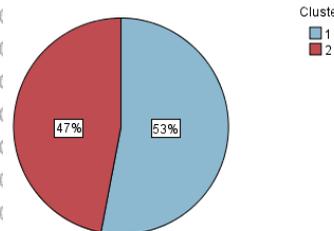
Size of Smallest Cluster	12 (18.2%)
Size of Largest Cluster	30 (45.5%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	2.50

10. Data 10

Model Summary

Algorithm	TwoStep
Input Features	9
Clusters	2

Cluster Sizes



Size of Smallest Cluster	47 (47%)
Size of Largest Cluster	53 (53%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.13

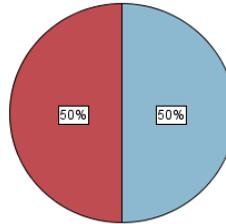
Lampiran 2. (Lanjutan)

11. Data 11

Model Summary

Algorithm	TwoStep
Input Features	7
Clusters	2

Cluster Sizes



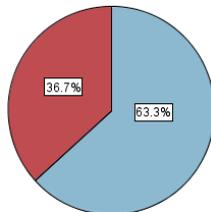
Size of Smallest Cluster	25 (50%)
Size of Largest Cluster	25 (50%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.00

12. Data 12

Model Summary

Algorithm	TwoStep
Input Features	4
Clusters	2

Cluster Sizes



Size of Smallest Cluster	11 (36.7%)
Size of Largest Cluster	19 (63.3%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.73

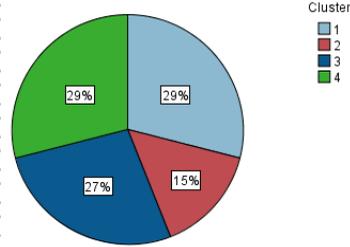
Lampiran 2. (Lanjutan)

13. Data 13

Model Summary

Algorithm	TwoStep
Input Features	15
Clusters	4

Cluster Sizes



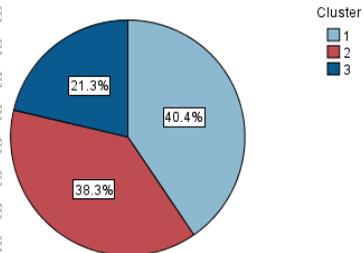
Size of Smallest Cluster	15 (15%)
Size of Largest Cluster	29 (29%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.93

14. Data 14

Model Summary

Algorithm	TwoStep
Input Features	6
Clusters	3

Cluster Sizes



Size of Smallest Cluster	10 (21.3%)
Size of Largest Cluster	19 (40.4%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.90

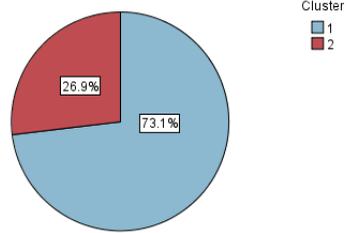
Lampiran 2. (Lanjutan)

15. Data 15

Model Summary

Algorithm	TwoStep
Input Features	6
Clusters	2

Cluster Sizes



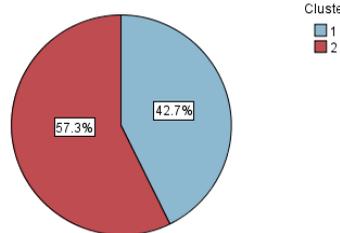
Size of Smallest Cluster	14 (26.9%)
Size of Largest Cluster	38 (73.1%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	2.71

16. Data 16

Model Summary

Algorithm	TwoStep
Input Features	6
Clusters	2

Cluster Sizes



Size of Smallest Cluster	38 (42.7%)
Size of Largest Cluster	51 (57.3%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.34

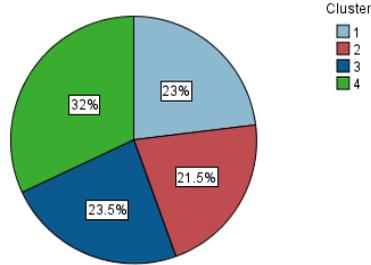
Lampiran 2. (Lanjutan)

17. Data 17

Model Summary

Algorithm	TwoStep
Input Features	6
Clusters	4

Cluster Sizes



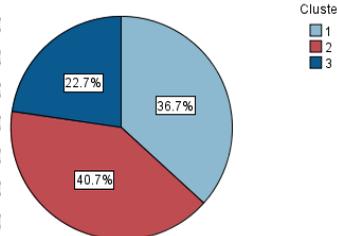
Size of Smallest Cluster	43 (21.5%)
Size of Largest Cluster	64 (32%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.49

18. Data 18

Model Summary

Algorithm	TwoStep
Input Features	7
Clusters	3

Cluster Sizes



Size of Smallest Cluster	34 (22.7%)
Size of Largest Cluster	61 (40.7%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.79

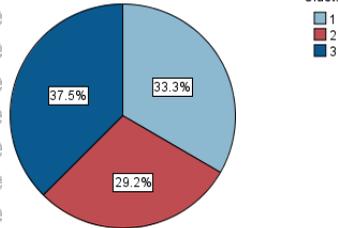
Lampiran 2. (Lanjutan)

19. Data 19

Model Summary

Algorithm	TwoStep
Input Features	7
Clusters	3

Cluster Sizes



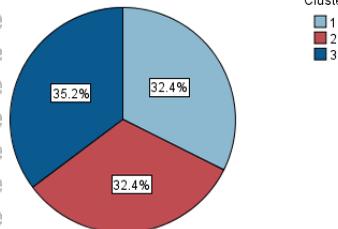
Size of Smallest Cluster	14 (29.2%)
Size of Largest Cluster	18 (37.5%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.29

20. Data 20

Model Summary

Algorithm	TwoStep
Input Features	5
Clusters	3

Cluster Sizes



Size of Smallest Cluster	23 (32.4%)
Size of Largest Cluster	25 (35.2%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.09

Lampiran 3. Rasio Keragaman Analisis *Two Step Cluster*

DATA1	KATEGORIK	KONTINU
SSW	0,714	26988682,8
SSB	15,793	29206617,91
RASIO	0,045	0,924
KERAGAMAN	0,484	

DATA2	KATEGORIK	KONTINU
SSW	0,709	4,881
SSB	10,178	1,122
RASIO	0,069	4,349
KERAGAMAN	2,209	

DATA3	KATEGORIK	KONTINU
SSW	0,714	23,731
SSB	9,389	20,628
RASIO	0,076	1,151
KERAGAMAN	0,613	

DATA4	KATEGORIK	KONTINU
SSW	0,712	41,297
SSB	13,593	39,985
RASIO	0,052	1,032
KERAGAMAN	0,542	

DATA5	KATEGORIK	KONTINU
SSW	0,711	1664071,247
SSB	12,293	685642,7487
RASIO	0,057	2,427
KERAGAMAN	1,242	

Lampiran 3. (Lanjutan)

DATA6	KATEGORIK	KONTINU
SSW	0,716	3,987
SSB	12,048	0,886
RASIO	0,059	4,496
KERAGAMAN	2,277	

DATA7	KATEGORIK	KONTINU
SSW	0,709	27204455,98
SSB	13,611	21293451,73
RASIO	0,052	1,277
KERAGAMAN	0,664	

DATA8	KATEGORIK	KONTINU
SSW	0,711	4263832,913
SSB	9,904	2548577,516
RASIO	0,071	1,673
KERAGAMAN	0,872	

DATA9	KATEGORIK	KONTINU
SSW	0,679	11,373
SSB	1,175	10,572
RASIO	0,578	1,075
KERAGAMAN	0,827	

DATA10	KATEGORIK	KONTINU
SSW	0,709	8,118
SSB	19,926	1,028
RASIO	0,035	7,896
KERAGAMAN	3,965	

Lampiran 3. (Lanjutan)

DATA11	KATEGORIK	KONTINU
SSW	0,712	2824094,613
SSB	5,621	1045097,463
RASIO	0,126	2,702
KERAGAMAN	1,414	

DATA12	KATEGORIK	KONTINU
SSW	0,711	10,862
SSB	10,883	6,6563
RASIO	0,065	1,631
KERAGAMAN	0,848	

DATA13	KATEGORIK	KONTINU
SSW	0,696	4,691
SSB	23,384	1,098
RASIO	0,029	4,271
KERAGAMAN	2,151	

DATA14	KATEGORIK	KONTINU
SSW	0,671	9,689
SSB	9,451	6,639
RASIO	0,071	1,459
KERAGAMAN	0,765	

DATA15	KATEGORIK	KONTINU
SSW	0,537	1405,855
SSB	11,772	2183,725
RASIO	0,045	0,643
KERAGAMAN	0,344	

Lampiran 3. (Lanjutan)

DATA16	KATEGORIK	KONTINU
SSW	0,715	5795914,043
SSB	6,536	440403278,3
RASIO	0,109	0,013
KERAGAMAN	0,061	

DATA17	KATEGORIK	KONTINU
SSW	0,711	17,931
SSB	22,359	12,695
RASIO	0,031	1,412
KERAGAMAN	0,722	

DATA18	KATEGORIK	KONTINU
SSW	0,706	7,435
SSB	22,967	7,251
RASIO	0,031	1,025
KERAGAMAN	0,528	

DATA19	KATEGORIK	KONTINU
SSW	0,421	14,827
SSB	8,175	2,132
RASIO	0,051	6,962
KERAGAMAN	3,506	

DATA20	KATEGORIK	KONTINU
SSW	0,542	4,736
SSB	9,586	1,215
RASIO	0,056	3,897
KERAGAMAN	1,977	

Lampiran 4. Output Analisis Latent Class Cluster

1. Data 1

	LL	BIC(LL)	Class.Err.
1-Cluster	-5027,46	10101,17	0
2-Cluster	-4812,2	9721,531	0,0159
3-Cluster	-4712,69	9573,377	0,0166
4-Cluster	-4660,28	9519,433	0,0127
5-Cluster	-4597,49	9444,731	0,0128
6-Cluster	-4561,17	9422,956	0,0212
7-Cluster	-4573,48	9498,463	0,0245
8-Cluster	-4549,59	9501,56	0,0179
9-Cluster	-4527,66	9508,564	0,0209
10-Cluster	-4501,19	9506,495	0,0255

2. Data 2

	LL	BIC(LL)	Class.Err.
1-Cluster	-425,473	896,6929	0
2-Cluster	-344,555	776,4464	0,0251
3-Cluster	-316,683	762,2918	0,0796
4-Cluster	-300,587	771,6891	0,0775
5-Cluster	-290,598	793,2991	0,1173

3. Data 3

	LL	BIC(LL)	Class.Err.
1-Cluster	-1594,35	3214,925	0
2-Cluster	-1277,05	2606,542	0,0055
3-Cluster	-1198,21	2475,066	0,0449
4-Cluster	-1188,32	2481,512	0,0594
5-Cluster	-1182,37	2495,832	0,0845

Lampiran 4. (Lanjutan)

4. Data 4

	LL	BIC(LL)	Class.Err.
1-Cluster	-12490,9	25028,17	0
2-Cluster	-6367,16	12831,85	0
3-Cluster	-4792,67	9733,963	0
4-Cluster	-4157,09	8513,896	0,0006
5-Cluster	-3717,64	7686,086	0,0015
6-Cluster	-3538,97	7379,822	0,0002
7-Cluster	-3332,33	7017,629	0,0041
8-Cluster	-3201,74	6807,541	0,0031
9-Cluster	-3123,15	6701,451	0,0041
10-Cluster	-3124,57	6755,369	0,0033

5. Data 5

	LL	BIC(LL)	Class.Err.
1-Cluster	-4873,1	9800,19	0
2-Cluster	-4773,89	9651,266	0,0188
3-Cluster	-4663,75	9480,491	0,0038
4-Cluster	-4621,94	9446,376	0,0022
5-Cluster	-4606,37	9464,732	0,029

6. Data 6

	LL	BIC(LL)	Class.Err.
1-Cluster	-672,76	1378,004	0
2-Cluster	-524,864	1118,756	0,0024
3-Cluster	-419,379	944,329	0,0001
4-Cluster	-402,492	947,0996	0,0001
5-Cluster	-392,114	962,8873	0,0001

Lampiran 4. (Lanjutan)

7. Data 7

	LL	BIC(LL)	Class.Err.
1-Cluster	-3146,92	6495,588	0
2-Cluster	-2994,21	6387,319	0,0287
3-Cluster	-2925,71	6447,466	0,0442
4-Cluster	-2886,64	6566,481	0,0382
5-Cluster	-2868,79	6727,925	0,0471

8. Data 8

	LL	BIC(LL)	Class.Err.
1-Cluster	-6,4E+07	128946233,9	0
2-Cluster	-1,9E+07	38533305,41	0
3-Cluster	-6200570	12401322,35	0
4-Cluster	-2046042	4092326,347	0
5-Cluster	-2046042	4092386,153	0

9. Data 9

	LL	BIC(LL)	Class.Err.
1-Cluster	-785,85	1617,786	0
2-Cluster	-630,169	1335,753	0,0003
3-Cluster	-591,054	1286,849	0,0217
4-Cluster	-572,595	1279,258	0,0434
5-Cluster	-559,373	1282,143	0,0405

10. Data 10

	LL	BIC(LL)	Class.Err.
1-Cluster	-1108,32	2262,698	0
2-Cluster	-986,714	2065,532	0,0339
3-Cluster	-956,14	2050,435	0,0586
4-Cluster	-934,073	2052,352	0,045
5-Cluster	-921,97	2074,199	0,0535

Lampiran 4. (Lanjutan)

11. Data 11

	LL	BIC(LL)	Class.Err.
1-Cluster	-2910,57	5868,081	0
2-Cluster	-2740,48	5574,845	0,0096
3-Cluster	-2659,45	5459,735	0,0151
4-Cluster	-2626,31	5440,393	0,0073
5-Cluster	-2605,91	5446,545	0,0133

12. Data 12

	LL	BIC(LL)	Class.Err.
1-Cluster	-164,914	343,4329	0
2-Cluster	-154,028	338,6676	0,0972
3-Cluster	-150,452	348,5212	0,1119
4-Cluster	-148,483	361,5886	0,1099
5-Cluster	-146,972	375,5732	0,1559

13. Data 13

	LL	BIC(LL)	Class.Err.
1-Cluster	-2082,82	4239,313	0
2-Cluster	-1970,5	4092,962	0,0302
3-Cluster	-1898,77	4027,796	0,0384
4-Cluster	-1857,54	4023,635	0,0351
5-Cluster	-1832,05	4050,925	0,0288

14. Data 14

	LL	BIC(LL)	Class.Err.
1-Cluster	-683,072	1412,345	0
2-Cluster	-644,396	1385,046	0,0034
3-Cluster	-623,559	1393,424	0,0085
4-Cluster	-615,19	1426,738	0,0166
5-Cluster	-608,528	1463,465	0,0344

Lampiran 4. (Lanjutan)

15. Data 15

	LL	BIC(LL)	Class.Err
1-Cluster	-1151,32	2334,245	0
2-Cluster	-950,819	1968,808	0,0005
3-Cluster	-929,146	1961,024	0,0367
4-Cluster	-908,637	1955,568	0,0312
5-Cluster	-891,071	1955,996	0,0424

16. Data 16

	LL	BIC(LL)	Class.Err
1-Cluster	-2,5E+08	4,99E+08	0
2-Cluster	-9,6E+07	1,91E+08	0
3-Cluster	-5,6E+07	1,12E+08	0
4-Cluster	-1,6E+07	31748776	0
5-Cluster	-8815837	17631863	0
6-Cluster	-8512376	17024976	0
7-Cluster	-1872449	3745159	0
8-Cluster	-1867533	3735363	0
9-Cluster	-1867533	3735399	0
10-Cluster	-8507453	17015273	0,0002

Lampiran 4. (Lanjutan)

17. Data 17

	LL	BIC(LL)	Class.Err.
1-Cluster	-3942,55	7922,188	0
2-Cluster	-2442,06	4963,59	0,0027
3-Cluster	-2154,75	4431,358	0,0152
4-Cluster	-2075,34	4314,92	0,0397
5-Cluster	-2027,48	4261,589	0,0558
6-Cluster	-1979,54	4208,108	0,0583
7-Cluster	-1950,44	4192,283	0,0809
8-Cluster	-1932,6	4199	0,0863
9-Cluster	-1913,44	4203,056	0,0873
10-Cluster	-1901,06	4220,691	0,0964

18. Data 18

	LL	BIC(LL)	Class.Err.
1-Cluster	-1236,69	2533,505	0
2-Cluster	-1094,2	2293,625	0,01
3-Cluster	-1067,85	2286,025	0,0368
4-Cluster	-1042,55	2280,522	0,0129
5-Cluster	-1023,79	2288,091	0,0495

19. Data 19

	LL	BIC(LL)	Class.Err.
1-Cluster	-896,285	1835,154	0
2-Cluster	-861,043	1811,123	0,0003
3-Cluster	-849,518	1834,528	0,0773
4-Cluster	-837,192	1856,331	0,0767
5-Cluster	-821,815	1872,032	0,0874

Lampiran 4. (Lanjutan)
20. Data 20

	LL	BIC(LL)	Class.Err.
1-Cluster	-625,996	1281,83	0
2-Cluster	-554,082	1167,842	0,0379
3-Cluster	-532,272	1154,06	0,0393
4-Cluster	-519,436	1158,226	0,0531
5-Cluster	-505,953	1161,1	0,0444

Lampiran 5. Rasio Keragaman Analisis *Latent Class Cluster*

DATA1	KATEGORIK	KONTINU
SSW	0,712	12845526,59
SSB	13,994	37719312,27
RASIO	0,0501	0,341
KERAGAMAN	0,195	

DATA2	KATEGORIK	KONTINU
SSW	0,689	3,479
SSB	11,551	4,713
RASIO	0,059	0,738
KERAGAMAN	0,398	

DATA3	KATEGORIK	KONTINU
SSW	0,713	26,349
SSB	9,021	20,106
RASIO	0,079	1,311
KERAGAMAN	0,694	

DATA4	KATEGORIK	KONTINU
SSW	0,698	22,477
SSB	15,671	73,901
RASIO	0,044	0,304
KERAGAMAN	0,174	

DATA5	KATEGORIK	KONTINU
SSW	0,446	1228171,36
SSB	19,461	2343463,629
RASIO	0,022	0,524
KERAGAMAN	0,273	

Lampiran 5. (Lanjutan)

DATA6	KATEGORIK	KONTINU
SSW	0,712	2,403
SSB	18,721	5,732
RASIO	0,038	0,419
KERAGAMAN	0,228	

DATA7	KATEGORIK	KONTINU
SSW	0,711	32888190,77
SSB	18,721	20546325,22
RASIO	0,037	1,601
KERAGAMAN	0,819	

DATA8	KATEGORIK	KONTINU
SSW	0,695	4116663,524
SSB	9,007	1640458,155
RASIO	0,077	2,509
KERAGAMAN	1,293	

DATA9	KATEGORIK	KONTINU
SSW	0,673	7,399
SSB	7,627	7,148
RASIO	0,088	1,034
KERAGAMAN	0,561	

DATA10	KATEGORIK	KONTINU
SSW	0,703	5,078
SSB	23,287	6,234
RASIO	0,031	0,814
KERAGAMAN	0,422	

Lampiran 5. (Lanjutan)

DATA11	KATEGORIK	KONTINU
SSW	0,671	2979939,003
SSB	5,331	4737177,601
RASIO	0,125	0,629
KERAGAMAN	0,377	

DATA12	KATEGORIK	KONTINU
SSW	0,712	8,082
SSB	7,266	11,832
RASIO	0,097	0,683
KERAGAMAN	0,391	

DATA13	KATEGORIK	KONTINU
SSW	0,692	4,539
SSB	24,916	0,793
RASIO	0,027	5,721
KERAGAMAN	2,874	

DATA14	KATEGORIK	KONTINU
SSW	0,696	10,248
SSB	9,146	7,028
RASIO	0,076	1,458
KERAGAMAN	0,767	

DATA15	KATEGORIK	KONTINU
SSW	0,518	1289,654
SSB	9,564	1909,917
RASIO	0,054	0,675
KERAGAMAN	0,364	

Lampiran 5. (Lanjutan)

DATA16	KATEGORIK	KONTINU
SSW	0,182	108458,08
SSB	10,269	4971388,74
RASIO	0,017	0,021
KERAGAMAN	0,019	

DATA17	KATEGORIK	KONTINU
SSW	0,711	9,274
SSB	21,184	17,033
RASIO	0,033	0,544
KERAGAMAN	0,289	

DATA18	KATEGORIK	KONTINU
SSW	0,663	6,111
SSB	53,273	15,742
RASIO	0,012	0,388
KERAGAMAN	0,201	

DATA19	KATEGORIK	KONTINU
SSW	0,705	14,155
SSB	4,003	3,827
RASIO	0,176	3,698
KERAGAMAN	1,937	

DATA20	KATEGORIK	KONTINU
SSW	0,542	3,474
SSB	13,371	10,994
RASIO	0,041	0,315
KERAGAMAN	0,178	

Lampiran 6. *Output Uji Ragam*

Test for Equal Variances: TSC; LCC

95% Bonferroni confidence intervals for standard deviations

	N	Lower	StDev	Upper
TSC	20	0,776451	1,05982	1,64197
LCC	20	0,506989	0,69201	1,07213

F-Test

Test statistic = 2,35; p-value = 0,071

Lampiran 7. Output Uji Dua Nilai Tengah Rata-Rata Rasio Keragaman

Two-Sample T-Test and CI: TSC; LCC

	N	Mean	StDev	SE Mean
TSC	20	1,30	1,06	0,24
LCC	20	0,623	0,692	0,15

Difference = mu (TSC) - mu (LCC)

Estimate for difference: 0,677900

95% lower bound for difference: 0,200729

T-Test of difference = 0 (vs >): T-Value = 2,40 P-Value = 0,011 DF = 38

Lampiran 8. Langkah-langkah analisis *Two Step Cluster (TSC)*

1. Masukkan data pada *worksheet* SPSS v18

	id	x1	x2	x3	x4	x5	x6	x7	var
1	1	1	23	2	1	1	2	2	
2	2	1	23	2	1	2	1	3	
3	3	1	24	2	1	2	1	3	
4	4	1	24	2	1	2	1	3	
5	5	2	24	2	1	2	1	1	
6	6	2	24	2	1	2	1	1	
7	7	2	24	2	1	2	1	1	
8	8	1	25	2	1	2	1	1	
9	9	1	25	2	1	2	1	1	
10	10	1	25	2	1	1	2	3	
11	11	2	25	2	1	2	1	3	
12	12	2	26	2	1	2	1	3	
13	13	2	27	2	1	2	1	3	
14	14	1	27	2	1	1	2	2	
15	15	2	27	2	1	1	1	3	
16	16	2	27	2	1	1	1	3	
17	17	2	27	2	1	2	1	3	
18	18	2	27	2	1	1	2	3	
19	19	1	28	2	1	1	2	2	
20	20	2	28	2	1	1	2	1	
21	21	2	29	2	1	1	1	1	
22	22	2	29	2	1	1	2	1	
23	23	1	29	3	1	1	3	3	

2. Lakukan penyesuaian skala data untuk setiap variabel pada *“variable view”*

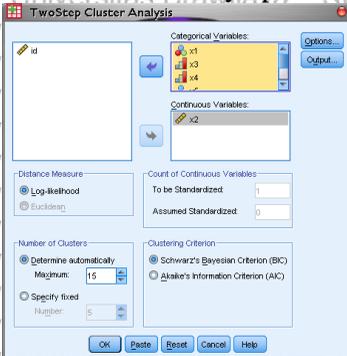
Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
x1	Numeric	8	0		None	None	8	Right	Scale	Input
x2	Numeric	8	0		None	None	8	Right	Nominal	Input
x3	Numeric	8	0		None	None	8	Right	Scale	Input
x4	Numeric	8	0		None	None	8	Right	Ordinal	Input
x5	Numeric	8	0		None	None	8	Right	Ordinal	Input
x6	Numeric	8	0		None	None	8	Right	Nominal	Input
x7	Numeric	8	0		None	None	8	Right	Ordinal	Input

3. Kembali pada *worksheet* data, klik *Analyze* → *classify* → *twostep cluster*

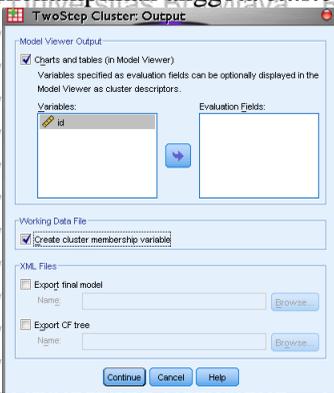
The screenshot shows the SPSS menu structure: **Analyze** → **Classify** → **Two-Step Cluster...**. The **Two-Step Cluster...** dialog box is open, showing options for **Display** (Cluster Centers, Cluster Membership, etc.) and **Save** (Cluster Membership, etc.).

Lampiran 8. (Lanjutan)

4. Masukkan variabel kategorik pada kolom *categorical variables* dan variabel kontinyu pada kolom *continuous variables*



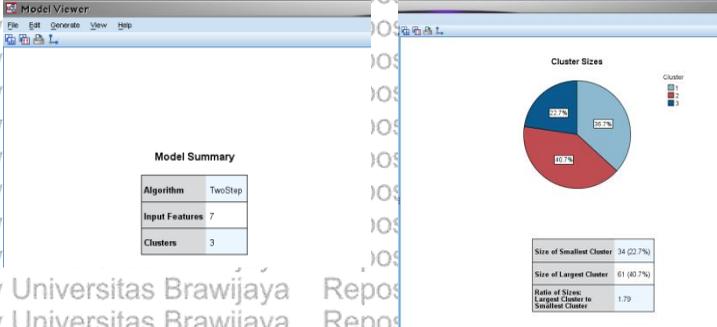
5. Klik *output* sehingga muncul tampilan seperti di bawah ini. Beri centang pada “*create cluster membership variable*” untuk menampilkan anggota *cluster* bagi setiap objek



6. Klik *continue*, kemudian klik *OK* untuk mengetahui hasil analisis *Two Step Cluster*.

Lampiran 8. (Lanjutan)

7. Hasil *Output* yang diperoleh dari analisis *Two Step Cluster* ditampilkan seperti di bawah ini

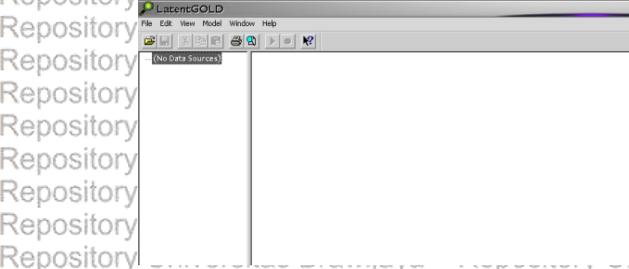


Cluster Membership untuk setiap objek:

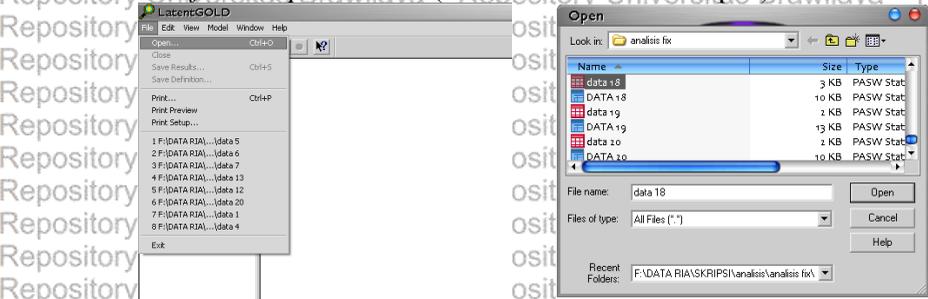
	id	x1	x2	x3	x4	x5	x6	x7	TSC_B9	var
1	1	1	23	2	1	1	2	2	1	
2	2	1	23	2	1	2	1	3	3	
3	3	1	24	2	1	2	1	3	3	
4	4	1	24	2	1	2	1	3	3	
5	5	2	24	2	1	2	1	1	3	
6	6	2	24	2	1	2	1	1	3	
7	7	2	24	2	1	2	1	3	3	
8	8	1	25	2	1	2	1	1	3	
9	9	1	25	2	1	2	1	1	3	
10	10	1	25	2	1	1	2	3	1	
11	11	2	25	2	1	2	1	3	3	
12	12	2	26	2	1	2	1	3	3	
13	13	2	27	2	1	2	1	3	3	
14	14	1	27	2	1	1	2	2	1	
15	15	2	27	2	1	1	1	3	3	
16	16	2	27	2	1	1	1	3	3	
17	17	2	27	2	1	2	1	3	3	
18	18	2	27	2	1	1	2	3	1	
19	19	1	28	2	1	1	2	2	1	
20	20	2	28	2	1	1	2	1	1	
21	21	2	29	2	1	2	1	1	3	
22	22	2	29	2	1	1	2	1	1	
23	23	1	29	3	1	1	3	3	1	

Lampiran 9. Langkah-langkah analisis *Latent Class Cluster (LCC)*

1. Buka *software* Latent Gold 4.0

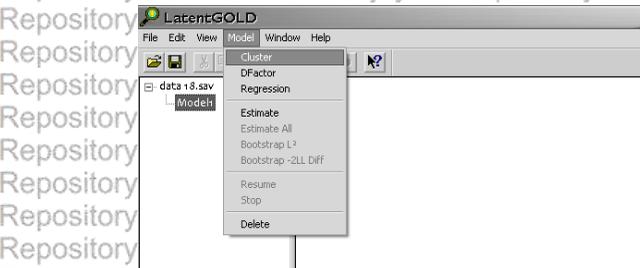


2. Klik *file* → *open* → *data* (folder lokasi data disimpan)



File yang digunakan dalam *software* Latent Gold v4.0 bertipe *.sav* sehingga data yang akan digunakan perlu disimpan dalam bentuk file data SPSS terlebih dahulu.

3. Klik *model* → *cluster*

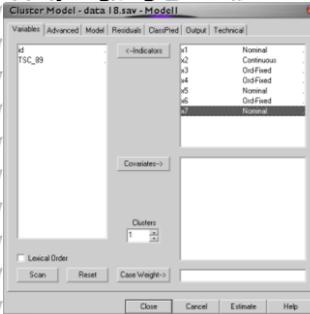


Lampiran 9. (Lanjutan)

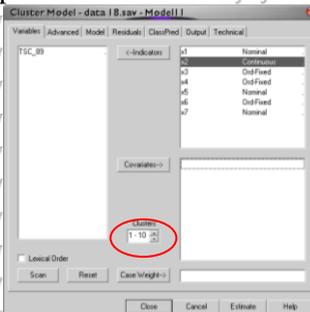
4. Pindahkan semua indikator atau variabel ke dalam kolom “*indicators*” sehingga muncul tampilan seperti di bawah ini



5. Klik kanan pada tiap variabel untuk menentukan setiap jenis skala data yang digunakan



6. Pada kolom “*clusters*” diisikan angka 1 sampai banyaknya cluster yang ingin dicoba untuk mendapatkan model terbaik dalam pembentukan model LCC



Lampiran 9. (Lanjutan)

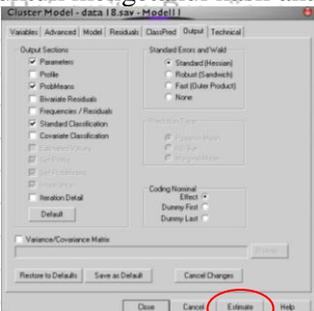
7. Pada menu *ClassPred*, masukkan id (banyaknya objek yang digunakan) pada kolom “← ID”



8. Pada menu *output* centang point-point yang ingin ditampilkan pada hasil analisis



9. Jika semua prosedur telah selesai dilakukan maka klik *estimate* untuk mengetahui hasil analisis *Latent Class Cluster*



Lampiran 9. (Lanjutan)

10. Hasil *output* yang diperoleh dari analisis *Latent Class Cluster* ditampilkan sebagai berikut:

LatentGOLD

File Edit View Model Window Help

data: i8.sav

File name: F:\DATA RIA\SKRIPSI\analisis\analisis fix\data
File size: 2408 bytes
File date: 2013-Okt-13 21:03:40

	LL	BIC(LL)	Class.Err.	
Model1	1-Cluster	-1236,6885	2533,5046	0,0000
Model2	2-Cluster	-1094,2010	2293,6253	0,0100
Model3	3-Cluster	-1067,8531	2286,0252	0,0368
Model4	4-Cluster	-1042,5835	2280,5218	0,0129
Model5	5-Cluster	-1023,7902	2288,0909	0,0495
Model6	6-Cluster	-1014,8036	2315,2135	0,0458
Model7	7-Cluster	-1000,7270	2332,1559	0,0632
Model8	8-Cluster	-995,7970	2367,3916	0,0498
Model9	9-Cluster	-991,6060	2404,1054	0,0407
Model10	10-Cluster	-986,4734	2438,9358	0,0484
Model11				

LatentGOLD

File Edit View Model Window Help

data: i8.sav

	x1	x2	x3	x4	x5	x6	x7	ObsFreq	Modal	Cluster1	Cluster2	Cluster3	Cluster4
Model1 - L2 = 1072,5979	1	23	2	1	1	2	2	1,0000	2	0,0000	0,9993	0,0000	0,0007
Model2 - L2 = 787,6219	1	23	2	1	2	1	3	1,0000	4	0,0000	0,0002	0,0000	0,9998
Model3 - L2 = 734,9271	1	24	2	1	2	1	3	2,0000	4	0,0000	0,0002	0,0000	0,9998
Model4 - L2 = 734,9271	1	25	2	1	1	2	3	1,0000	2	0,0015	0,9988	0,0000	0,0018
Model4 - L2 = 684,0204	1	25	2	1	2	1	1	2,0000	4	0,0000	0,0001	0,0000	0,9999
Parameters	1	27	2	1	1	2	2	1,0000	2	0,0001	0,9994	0,0000	0,0005
ProbMeans	1	28	2	1	1	2	2	1,0000	2	0,0002	0,9994	0,0000	0,0004
Standard Classificat	1	29	3	1	1	3	3	1,0000	2	0,0072	0,9928	0,0000	0,0000
Model5 - L2 = 652,0186	1	30	2	1	1	1	3	1,0000	2	0,0038	0,9963	0,0000	0,0017
Model5 - L2 = 619,8078	1	30	2	1	2	1	2	1,0000	4	0,0000	0,0011	0,0000	0,9989
Model8 - L2 = 605,9018	1	31	2	1	1	2	2	2,0000	2	0,0004	0,9993	0,0000	0,0003
Model9 - L2 = 581,5693	1	34	2	1	2	1	3	1,0000	4	0,0000	0,0005	0,0000	0,9994
Model10 - L2 = 578,5322	1	35	2	2	1	4	1	1,0000	1	0,9960	0,0032	0,0009	0,0000
Parameters	1	36	2	1	2	1	2	1,0000	4	0,0000	0,0021	0,0000	0,9979
ProbMeans	1	36	2	1	2	3	1,0000	1	0,9951	0,0042	0,0007	0,0000	
Standard Classificat	1	38	2	1	2	1	1	1,0000	1	0,9934	0,0028	0,0039	0,0000
Model11	1	39	2	2	1	1	3	1,0000	1	0,9958	0,0031	0,0007	0,0003
Parameters	1	39	2	2	1	3	1	1,0000	1	0,9915	0,0011	0,0073	0,0000
ProbMeans	1	39	2	2	1	4	1	1,0000	1	0,9870	0,0008	0,0115	0,0000
Standard Classificat	1	39	3	2	1	2	3	1,0000	1	0,9979	0,0021	0,0000	0,0000
Parameters	1	40	2	2	1	2	1	3,0000	1	0,9929	0,0015	0,0056	0,0000
ProbMeans	1	40	2	2	1	3	1	1,0000	1	0,9904	0,0009	0,0088	0,0000
Standard Classificat	1	40	2	2	1	3	3	2,0000	1	0,9971	0,0007	0,0021	0,0000
Parameters	1	41	2	2	1	2	1	1,0000	1	0,9922	0,0011	0,0087	0,0000
ProbMeans	1	41	2	2	1	3	3	1,0000	1	0,9969	0,0008	0,0025	0,0000
Standard Classificat	1	42	2	2	1	2	3	1,0000	1	0,9973	0,0007	0,0019	0,0000
Parameters	1	42	2	2	1	3	3	2,0000	1	0,9965	0,0004	0,0030	0,0000
ProbMeans	1	42	2	3	1	4	1	1,0000	3	0,1119	0,0000	0,8881	0,0000
Standard Classificat	1	44	2	2	1	2	3	1,0000	1	0,9968	0,0004	0,0028	0,0000
Parameters	1	44	2	3	1	1	3	1,0000	1	0,9883	0,0000	0,4137	0,0000
ProbMeans	1	44	2	3	1	3	1	2,0000	3	0,1211	0,0000	0,8789	0,0000
Standard Classificat	1	45	2	2	1	3	3	1,0000	1	0,9946	0,0002	0,0052	0,0000
Parameters	1	45	2	3	1	1	3	1,0000	1	0,5416	0,0000	0,4584	0,0000
ProbMeans	1	45	2	3	1	2	3	2,0000	3	0,4291	0,0000	0,5709	0,0000
Standard Classificat	1	46	2	2	1	2	3	1,0000	1	0,9958	0,0002	0,0040	0,0000
Parameters	1	47	2	2	2	1	1	1,0000	1	0,9801	0,0002	0,0197	0,0000
ProbMeans	1	47	2	2	2	2	2	1,0000	4	0,0000	0,0000	0,0000	0,0000

