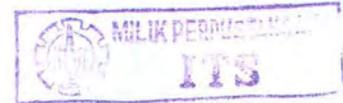


17.615 / H / 03



TUGAS AKHIR

ANALISIS HOMOGENITAS UNTUK MENGETAHUI KARAKTERISTIK RUMAH TANGGA MISKIN DI SURABAYA TAHUN 2001



RSSt
519 535
Ros
a-1

2002

Oleh:

DYAH ASTRI ROSMASARI

NRP. 1399.100.503

**JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER**

SURABAYA

2002

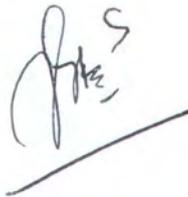
PERPUSTAKAAN ITS	
Tgl. Terima	16/09/02
Terima Dari	H
No. Agenda Frp.	21.6391

LEMBAR PENGESAHAN

**ANALISIS HOMOGENITAS UNTUK MENGETAHUI
KARAKTERISTIK RUMAH TANGGA MISKIN DI SURABAYA
TAHUN 2001**

Mengetahui / Menyetujui

Dosen Pembimbing

A handwritten signature in black ink, appearing to be 'Destri Susilaningrum', with a long horizontal line extending to the right.

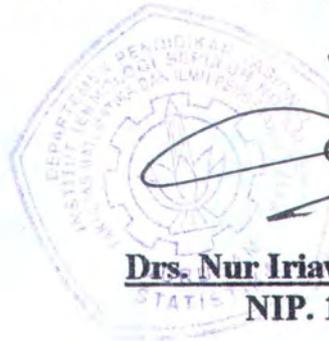
Dra. Destri Susilaningrum, MSi
NIP. 131 569 365

LEMBAR PENGESAHAN

ANALISIS HOMOGENITAS UNTUK MENGETAHUI KARAKTERISTIK RUMAH TANGGA MISKIN DI SURABAYA TAHUN 2001

Mengetahui

Ketua Jurusan Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember Surabaya



Drs. Nur Iriawan, M.Ikom, Ph.D.

NIP. 131 782 011

ABSTRAK

Analisis Homogenitas untuk mengetahui karakteristik rumah tangga miskin di Surabaya tahun 2001

Oleh : Dyah Astri Rosmasari (Nrp. 1399 100 503)

Pembimbing : Dra. Destri Susilaningrum, M.Si (Nip. 131 569 365)

Dalam beberapa kasus data multivariabel, jenis data bisa berupa data kuantitatif maupun kualitatif. Sehingga suatu data dapat mempunyai skala pengukuran nominal, ordinal, interval atau rasio. Metode multivariat selama ini banyak diterapkan pada data berskala interval atau rasio (data kuantitatif). Adapun salah satu metode yang dapat digunakan untuk menyelesaikan kasus data berskala nominal atau ordinal adalah *Homogeneity Analysis* (Homals). Metode Homals merupakan prosedur optimal scalling yang analog dengan Analisis Faktor, yaitu mampu mereduksi variabel, dimana antar variabel baru (*dimensi*) yang terbentuk saling independen satu sama lain dan tetap mampu menjelaskan keragaman data sehingga tidak banyak informasi yang hilang.

Transformasi yang digunakan adalah *Gram-Schmidt Transformation*, dengan nilai akar karakteristik (*eigen value*) yang digunakan harus lebih dari 1/jumlah variabel. Analisis terhadap 10 variabel nominal rumah tangga miskin di Surabaya tahun 2001 dapat mereduksi menjadi 6 dimensi yang mampu menjelaskan total keragaman data sebesar 88,79%. Penerapan metode *Homals* hanya sebagai analisis antara, karena masih perlu analisis lanjutan terhadap output yang dihasilkan, agar hasil analisis dapat diinterpretasikan dan lebih berguna.

Berdasarkan hasil yang didapat dari penerapan metode Homals pada data multivariabel nominal rumah tangga miskin tersebut, selanjutnya akan dilakukan *segmentasi* rumah tangga sampel dengan menggunakan metode pengelompokan nonhirarki (*k-means methods*), yaitu analisis *cluster* terhadap *object scores* pada 6 *dimensi* yang terbentuk. Menggunakan metode *k-means* dengan $k = 5$ kelompok (*segmen*) karena data penelitian merupakan data multivariat dengan sampel besar serta Kota Surabaya terbagi dalam 5 (lima) wilayah. Segmen I terdiri dari 1498 rumah tangga, segmen II terdiri 10049 rumah tangga, segmen III terdiri dari 3585 rumah tangga, segmen IV terdiri dari 1711 rumah tangga dan segmen V terdiri dari 147 rumah tangga. Kemudian dilakukan Analisis Tabulasi Silang antara *segmen* dengan variabel *geografi* dan *variabel lainnya* untuk melihat karakteristik yang signifikan dari tiap-tiap segmen.

Setelah dilakukan analisis diskriminan ternyata ada 10 variabel yang berperan dalam pembentukan kelompok. Variabel-variabel tersebut adalah Pembelian Pakaian dalam setahun (X_1), Fasilitas Air bersih (X_2), Kepemilikan rumah (X_4), Jenis dinding (X_5), Jenis Lantai (X_6), Sarana Tempat Buang Air Besar – TBAB (X_7), Sumber Penerangan (X_8), Pendidikan ART umur 6 – 15 tahun (X_9), Sumber Keuangan Rumah tangga (X_{10}), Pelayanan Kesehatan (X_{11}). Sedangkan tingkat ketepatan dalam pengklasifikasian sebesar 91,9 %.

KATA PENGANTAR

Assalamu'alaikum Wr. Wb

Dengan mengucapkan puji dan syukur kehadiran Allah SWT, yang telah memberikan taufik dan hidayah-Nya serta senantiasa melindungi, membimbing dan memberikan kemudahan - kemudahan sehingga penulis dapat menyelesaikan Tugas Akhir ini. Tugas Akhir ini merupakan syarat dan kelengkapan dalam menyelesaikan studi Strata Satu di jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember Surabaya.

Tugas Akhir ini mengambil judul "**Analisis Homogenitas untuk mengetahui Karakteristik Rumah Tangga Miskin di Surabaya Tahun 2001**" dengan menggunakan metode *Homogeneity Analysis* (Homals). Tugas akhir ini pada intinya mereduksi multivariabel Kondisi Rumah Tangga Miskin di Surabaya Tahun 2001 sebagai dasar untuk melakukan segmentasi rumah tangga sampel.

Penulis menyadari masih ada bagian yang tidak sempurna dalam Tugas Akhir ini. Untuk itu penulis sangat mengharapkan adanya masukan dan saran membangun yang berkenaan dengan tugas akhir ini. Demikian besar harapan penulis semoga Tugas Akhir ini dapat bermanfaat adanya, baik didunia pendidikan maupun untuk kepentingan baik lainnya.

Wassalamu'alaikum Wr. Wb.

Surabaya, Juli 2002

Penulis

UCAPAN TERIMA KASIH

Kupersembahkan untuk :

1. Allah SWT Sang Pencipta dan Pemilik Takdir.
2. Rasulullah Muhammad SAW sebagai suri tauladan seluruh umat manusia.
3. Bapak, Ibu di Banyuwangi yang telah mencurahkan kasih sayang, perhatian, nasehat, do'a serta dukungannya sehingga ananda bisa menyelesaikan studi ini.
4. Bapak, Ibu di Salatiga atas perhatian, kasih sayang, do'a serta dukungannya sehingga ananda bisa menyelesaikan studi ini.
5. Suamiku tercinta (Mas Wisnu) atas kesabaran menemani hari-hari berat selama proses penyelesaian tugas akhir ini, dukungan, serta kasih sayang dan do'anya sehingga Aci dapat menyelesaikan tugas akhir ini.
6. Adik-adikku (dik Zaki, dik Yusuf, dik Faiz, dik Edi, dik Lilik juga Satria) trim's atas dukungan kalian.
7. Lek Utik sekeluarga atas bantuannya selama ini.
8. Bpk Drs. Nur Iriawan , M.Ikom, Ph.D, selaku Ketua Jurusan Statistika, FMIPA, Institut Teknologi Sepuluh Nopember Surabaya.
9. Ibu Dra. Destri Susilaningrum, MSi, selaku dosen pembimbing Tugas Akhir, yang telah memberikan dukungan, masukan, pengarahan dan bimbingan dalam penyelesaian tugas Akhir ini.
10. Prof. Dra. Susanti Linuwih, Mstat, Ph.D , Ibu Dra. Lucia Aridinanti, MS dan Ibu Dra. Kartika Fitriyani, MSi selaku dosen penguji.
11. Para Dosen dan segenap Civitas Akademika di Jurusan Statistika .
12. Bapak Farikin atas kepercayaan dan bantuannya.
13. Teman-teman LJ Angkatan 99 : Mbak Puji, Mas Hadi, Mas Tulus, Mas Eki, Mas Fandi, Mas Agus, Mas Nasrudin, Mas Rahmat, Mas Maserul, Mas Yudi, Ety dan Eny atas kebersamaan dan kerjasamanya.
14. Adik-adik : Linda, Nita Nanik, Zya, Dede atas persahabatan dan perhatiannya selama ini.
15. Semua pihak yang tidak bisa disebutkan satu persatu.

DAFTAR ISI

	<i>Halaman</i>
ABSTRAK	i
KATA PENGANTAR	ii
UCAPAN TERIMA KASIH	iii
DAFTAR ISI	iv
DAFTAR TABEL	vi
DAFTAR LAMPIRAN	vii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Manfaat Penelitian	2
1.5 Batasan Masalah	3
BAB II TINJAUAN PUSTAKA	4
2.1 Tinjauan Rumah Tangga Miskin	4
2.2 Tinjauan Statistik	4
2.2.1 Analisis Komponen Utama Nonlinier	6
2.2.2 <i>Homogeneity Analysis (Homals)</i>	6
2.2.3 Analisis Kelompok (<i>Cluster Analysis</i>)	11
2.2.4 <i>Cross Tabulation</i> Dua Dimensi	17
2.2.5 Analisis Diskriminan	22

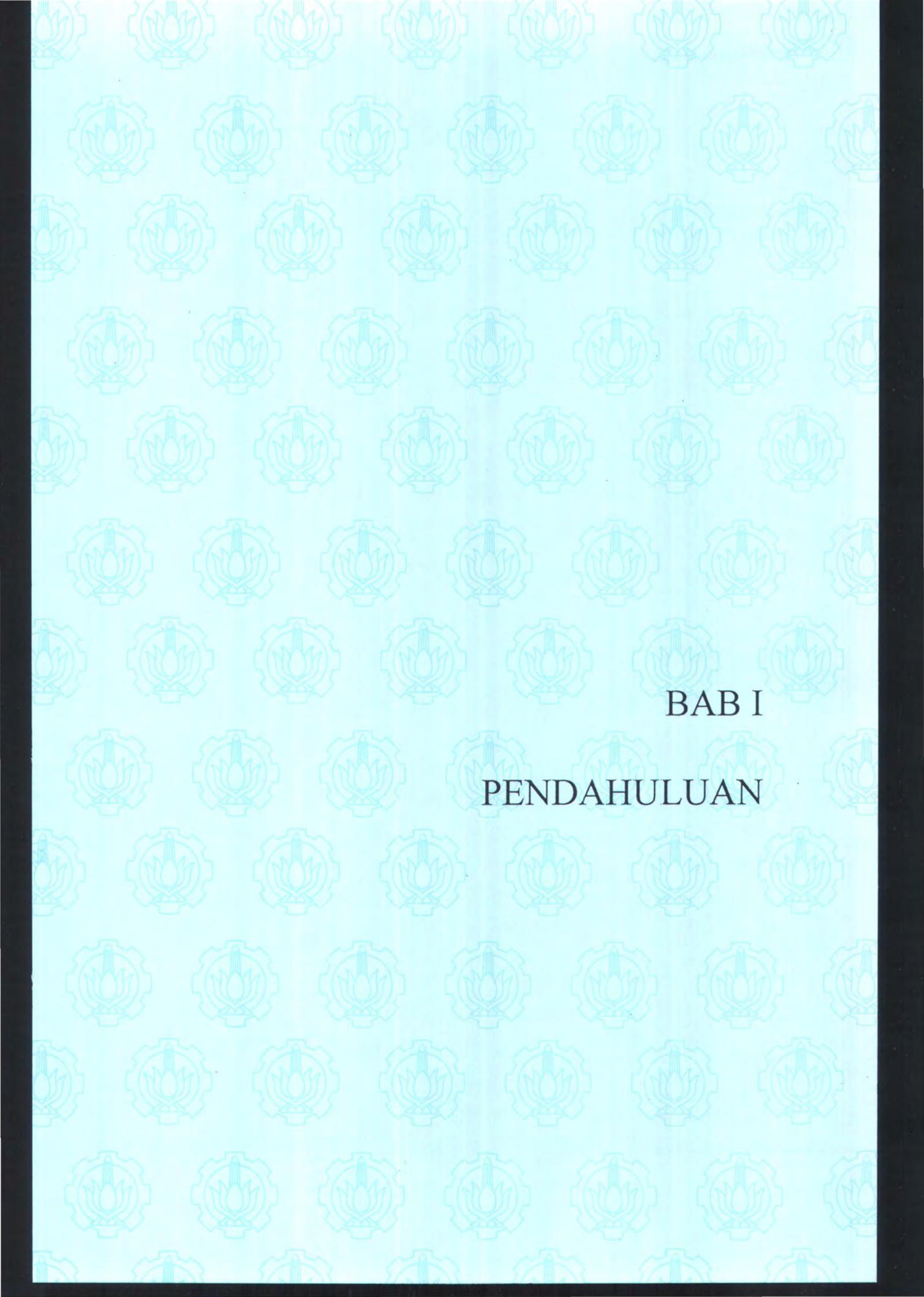
	<i>Halaman</i>
BAB III METODOLOGI	26
3.1 Bahan Penelitian	26
3.2 Variabel Penelitian	28
3.3 Metode Penelitian	34
BAB IV ANALISIS DATA DAN PEMBAHASAN	35
4.1 Analisis Homogeneity (<i>Homals</i>)	35
4.2 Analisis <i>Cluster</i>	38
4.3 Analisis Diskriminan	44
BAB V KESIMPULAN DAN SARAN	47
5.1 Kesimpulan	47
5.2 Saran	48
DAFTAR PUSTAKA	49
LAMPIRAN	50

DAFTAR TABEL

No.	JUDUL	Hal
2.1	Perbandingan antara Metode <i>Homals</i> , <i>Princals</i> dan <i>Overals</i> . . .	5
2.2	Tabel Kontingensi Dua Dimensi	18
3.1	Jumlah Rumah Tangga Sampel menurut Wilayah dan Kecamatan	27
4.1	Proporsi Keragaman yang dapat Diterangkan oleh <i>Homals</i> . .	35
4.2	Matriks <i>Discrimination Measures</i> per Variabel per Dimensi.	36
4.3	Tabel Anova untuk Uji Kesamaan Rata-Rata	40
4.4	Uji <i>Chi-Square</i> antara <i>Segmen</i> dengan Variabel	41

DAFTAR LAMPIRAN

No.	JUDUL	Halaman
A	Pengolahan Homogeneity Analysis.50
B	Pengolahan Analisis Cluster dengan Metode K-Means.59
C	Pengolahan Tabulasi Silang antara Segmen dengan Variabel .	.62
D	Pengolahan Analisis Diskriminan.73



BAB I

PENDAHULUAN

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam beberapa kasus data multivariabel, jenis data bisa kuantitatif maupun kualitatif. Sehingga suatu data dapat mempunyai berbagai skala pengukuran yaitu skala nominal, ordinal, interval atau rasio. Metode multivariate selama ini banyak diterapkan pada data yang berskala interval maupun rasio (data kuantitatif). Pada kenyataannya banyak juga data multivariabel diskrit (berskala nominal atau ordinal) yang dijumpai dan analisisnya tidak dapat menggunakan metode multivariate yang diterapkan pada data berskala interval atau rasio.

Adapun metode yang dapat digunakan untuk menyelesaikan kasus yang berskala data nominal atau ordinal adalah *Homogeneity Analysis* (HOMALS). Metode HOMALS merupakan prosedur *optimal scalling* yang analog dengan Analisis Komponen Utama yang dilanjutkan dengan Analisis Faktor.

Penelitian kali ini menggunakan data multivariabel nominal tentang penduduk miskin di Surabaya tahun 2001 yang bersumber dari Badan Pusat Statistika Surabaya, dimana masing – masing variabel mempunyai beberapa kategori.

Berdasarkan kelompok variabel tersebut akan dibentuk pengelompokan (segmentasi) rumah tangga sampel. Selanjutnya dari segmentasi rumah tangga sampel yang terbentuk akan dibandingkan dengan variabel geografi kota Surabaya dimana rumah tangga sampel dibedakan menjadi 5 kategori wilayah yaitu wilayah Surabaya Barat, Timur, Selatan, Utara dan wilayah Surabaya Pusat.

1.2 Permasalahan

Dari latar belakang tersebut, maka permasalahan yang timbul adalah :

1. Bagaimana segmentasi (pengelompokan) rumah tangga sampel berdasar hasil yang didapat dari analisis dengan metode Homals ?
2. Bagaimanakah karakteristik rumah tangga dari tiap-tiap segmen ?
3. Faktor – faktor apa saja yang membedakan kelompok rumah tangga miskin di Surabaya ?

1.3 Tujuan

Berdasarkan permasalahan yang ada, maka tujuan dari penelitian ini adalah :

1. Mengetahui segmentasi (pengelompokan) rumah tangga sampel berdasar hasil yang didapat dari analisis dengan metode Homals.
2. Mengetahui karakteristik rumah tangga dari tiap-tiap segmen .
3. Mengetahui faktor – faktor yang membedakan kelompok rumah tangga miskin di Surabaya

1.4 Manfaat

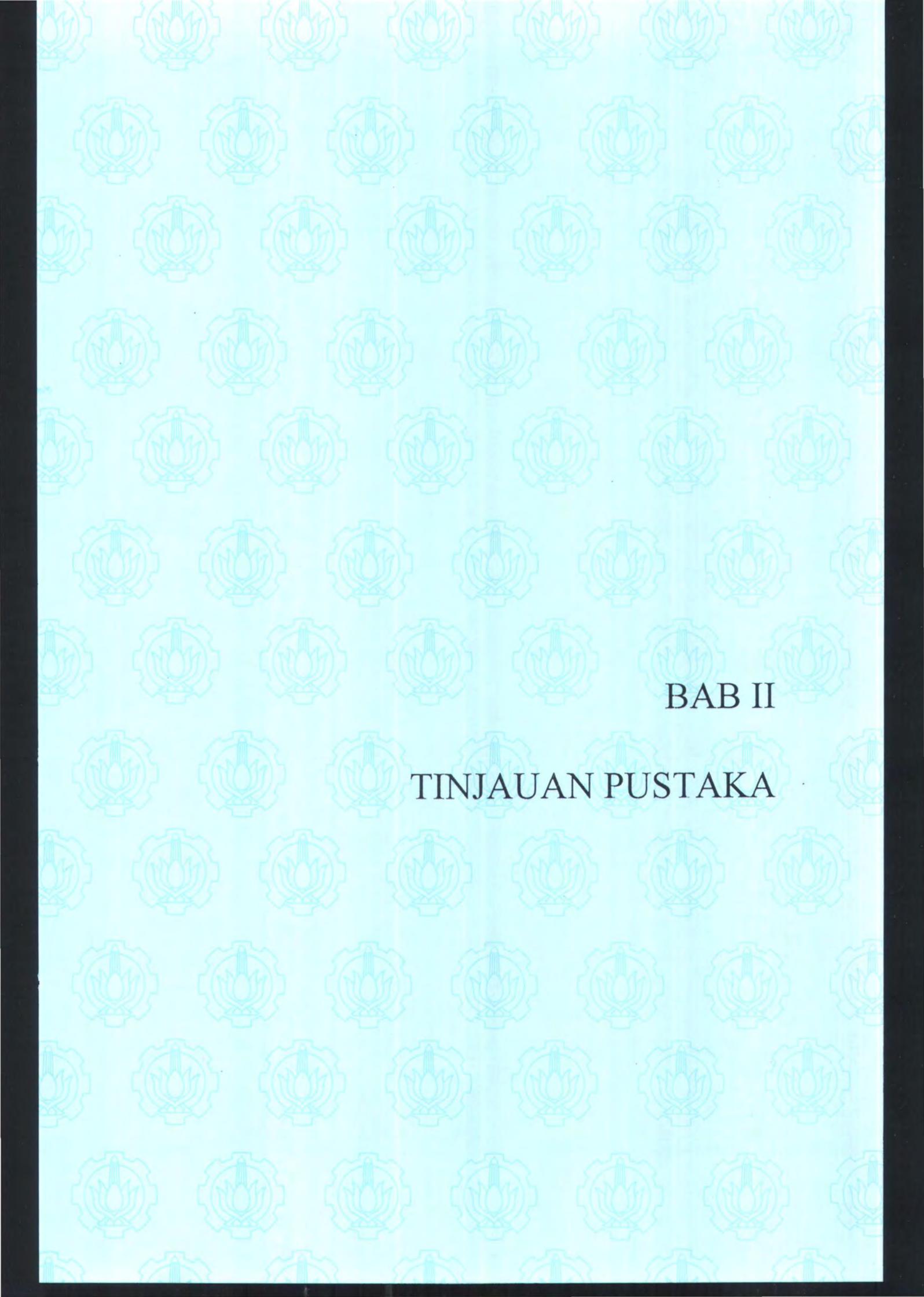
1. Dapat memahami metode Homals dengan algoritma ALS sebagai analisis korespondensi berganda yang analogi dengan Analisis komponen utama.
2. Dengan menggunakan metode Homals pada data Rumah tangga miskin di Surabaya tahun 2001, diharapkan dapat diketahui segmentasi rumah tangga sampel serta mengetahui karakteristik rumah tangga dari tiap – tiap segmen dan faktor – faktor apa saja yang membedakan kelompok rumah tangga miskin di Surabaya

1.5 Batasan Masalah

Dalam tugas akhir ini data yang digunakan adalah data sekunder dari BPS Surabaya tentang penduduk miskin di Surabaya tahun 2001 yang semuanya merupakan jenis data kategorikal yang meliputi :

1. Kelompok variabel kondisi rumah tangga miskin, yang terdiri dari 11 variabel kategorikal.
2. Kelompok variabel geografi, yaitu variabel wilayah yang menunjukkan dimana rumah tangga sampel berada.

Pengolahannya menggunakan metode Analisis Homogeneity (Homals) yang diolah menggunakan paket program *SPSS 8.0 for windows*. Pengolahan dengan metode Homals dapat mereduksi m variabel menjadi p dimensi yang menghasilkan data baru berupa *obyek score*(X). Kemudian dilakukan analisis lanjutan dengan Analisis Cluster terhadap *obyek score*. Selanjutnya dilakukan tabulasi silang antara segmen dengan variabel geografi kota Surabaya.



BAB II

TINJAUAN PUSTAKA

BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Rumah Tangga Miskin

Melalui studi yang telah dilakukan BPS menggunakan data Susenas 1998, ITS Surabaya menggunakan data Susenas 1997, dan diskusi instansi terkait (Departemen Sosial, BKKBN, Dinas Kesehatan dan BPS) dicoba untuk merumuskan sebuah indikator praktis. Pengukuran variabel lebih cenderung menyangkut aspek ekonomi dan kurang menyangkut aspek sosial. Hal ini sesuai dengan sifat bantuan yang lebih bersifat ekonomi. Pertimbangan lain yang lebih mendasari adalah aspek ekonomi lebih terukur. Penentuan skor Rumah tangga miskin didahului oleh satu penelitian yang dilakukan oleh BPS Propinsi Jawa Timur. Menggunakan metode Statistika dengan data susenas 1997 mencakup berbagai variabel kependudukan, pendidikan, kesehatan, ketenagakerjaan, perumahan, sosial budaya, komposisi pengeluaran rumah tangga dan lainnya diperoleh variabel pembeda kelompok masyarakat miskin, mendekati miskin, dan tidak miskin.

2.2 Tinjauan Statistik

Sesuai dengan tujuan penelitian, maka analisis yang dilakukan adalah metode *Homogeneity Analysis* (Homals), yang merupakan salah satu metode penskalaan optimal. Secara umum ada tiga prosedur penskalaan optimal, yaitu *Homogeneity Analysis* (Homals), *Nonlinear Principal Component Analysis* (Princals) dan

Nonlinear Canonical Correlation Analysis (Overals). Ketiga metode tersebut adalah metode analisis multivariat yang bertujuan untuk mengetahui hubungan di antara data multivariabel dengan jalan mereduksi m variabel menjadi p dimensi ($p < m$), yang hasilnya dinyatakan dalam data baru yang berupa *Object Scores (X)* dan *Category Quantification (Y_j)*. Yang membedakan penggunaan metode tersebut adalah jenis variabel random, skala pengukuran, jenis data dan banyaknya kelompok variabel yang dianalisis, seperti ditunjukkan dalam tabel 2.1.

Sementara data yang digunakan dalam penelitian ini selain multivariabel juga merupakan data multivariat. Data multivariabel adalah data dengan banyak variabel (peubah), sedangkan data multivariat adalah data yang dikumpulkan dari dua atau lebih observasi (obyek) dengan mengukur observasi tersebut dengan beberapa karakteristik.

Tabel 2.1 Perbandingan antara Metode *Homals*, *Princals* dan *Overals*

Uraian	Metode		
	<i>Homals</i>	<i>Princals</i>	<i>Overals</i>
Variabel Random	Diskrit	Diskrit Kontinu	Diskrit Kontinu
Skala Pengukuran	Nominal	Nominal Ordinal Interval Rasio	Nominal Ordinal Interval Rasio
Jenis Data	Kualitatif	Kualitatif Kuantitatif	Kualitatif Kuantitatif
Banyaknya Kelompok Variabel	1	1	> 1

2.2.1 Analisis Komponen Utama Nonlinier (*Nonlinear PCA/Princals*)

Analisis Komponen Utama Nonlinier merupakan teori pengembangan dari Analisis Komponen Utama yang dapat menerima variabel-variabel dengan jenis pengukuran yang berbeda-beda (nominal, ordinal, rasio dan interval) dan dianalisis secara simultan. Bila jenis data yang dianalisis adalah *kuantitatif*, maka hasil yang diperoleh dari Analisis Komponen Utama Nonlinier akan ekuivalen dengan hasil dari Analisis Komponen Utama yang dilanjutkan dengan Analisis Faktor. Analisis Komponen Utama Nonlinier disebut juga *PRINCALS (Principal Component Analysis by means of Alternating Least Squares)* atau Analisis Komponen Utama dengan menggunakan pendekatan *alternating least squares*.

2.2.2 *Homogeneity Analysis (HOMALS)*

Homogeneity analysis (HOMALS) yang dikenal juga sebagai *Categorical PCA*, merupakan bentuk khusus dari *Nonlinear PCA (Princals)*, yang dapat menerima variabel-variabel dengan jenis pengukuran yang sama (*homogen*), yaitu variabel *kualitatif* yang berskala nominal. Variabel-variabel tersebut dalam indikator matriksnya berupa kode kategori. Dengan menggunakan metode *HOMALS* pada data multivariabel yang berskala nominal, akan menghasilkan data baru yang dinyatakan dalam bentuk *object scores (X)* dan *category quantification (Y)*.

Untuk memudahkan pemahaman rumusan yang akan dibahas dalam metode *HOMALS*, maka akan diberikan notasi-notasi yang digunakan, yaitu :

n = banyaknya observasi atau kasus (obyek)

m = banyaknya variabel

p = ukuran dimensi

Untuk variabel j ; dimana $j = 1, 2, \dots, m$

k_j = banyaknya kategori yang valid dari variabel j

h_j = vektor $n \times 1$ dengan observasi kategorikal

G_j = matriks indikator untuk variabel j , dengan ukuran $n \times k_j$, yang elemennya:

$$g_{(j)ir} = \begin{cases} 1, & \text{jika obyek ke-}i \text{ berada dalam kategori ke-}r \text{ dari variabel } j \\ 0, & \text{jika obyek ke-}i \text{ tidak berada dalam kategori ke-}r \text{ dari variabel } j \end{cases}$$

D_j = matriks diagonal ordo $k_j \times k_j$ yang berisi marginal univariate, yaitu jumlah

$$\text{kolom dari } G_j \text{ atau } D_j = G_j^T G_j$$

M_j = binary matriks diagonal $n \times n$, dengan elemen pada diagonal utama

didefinisikan sebagai berikut :

$$m_{(j)ii} = \begin{cases} 1, & \text{jika observasi ke-}i \text{ berada dalam range } [1, k_j] \\ 0, & \text{jika observasi ke-}i \text{ tidak berada dalam range } [1, k_j] \end{cases}$$

Matriks kuantifikasinya adalah :

X = *Object Scores*, dengan ordo $n \times p$

Y_j = *Category quantification*, dengan ordo $k_j \times p$

Y = Matriks *Category quantification*, dengan ordo $\sum_j k_j \times p$

Penghitungan *object scores* (X) dan *category quantification* (Y_j) adalah dengan meminimumkan fungsi obyektif berikut :

$$\begin{aligned} \sigma(X; Y_1, Y_2, \dots, Y_m) &= \sigma(X : Y) = 1/m \sum_{j=1}^m SSQ(X - G_j Y_j) \dots \dots \dots (2.1) \\ &= 1/m \sum_{j=1}^m tr((X - G_j Y_j)' M_j (X - G_j Y_j)) \end{aligned}$$

dengan kendala :

$$1) \quad X' M_* X = mnI, \text{ dimana matrik } M_* = \sum_{j=1}^m M_j \text{ dan } I \text{ adalah matrik Identitas}$$

dengan ordo $p \times p$.

$$2) \quad u' M_* X = 0, \text{ dengan } u \text{ dinotasikan sebagai vektor-}n \text{ dengan elemen satu.}$$

Untuk mendapatkan solusi dari minimisasi fungsi obyektif dengan algoritma *ALS (Alternative Least Square)*, maka persamaan normal dari persamaan (2.1) adalah

$$1. \quad D_j Y_j = G_j' X, \quad j = 1, \dots, m; \text{ sehingga } Y_j = D_j^{-1} G_j' X$$

$$2. \quad mX = \sum_{j=1}^m G_j Y_j \text{ sehingga } X = 1/m \sum_{j=1}^m G_j Y_j$$

2.2.2.1 Iterasi untuk Optimisasi

HOMALS mengikuti algoritma *ALS* yang dilakukan dengan proses iterasi, dimana langkah-langkah iterasinya adalah :

1. Inisialisasi

Object scores X diberi inisial dengan bilangan random, yang dinormalkan

sehingga $u' M_* X = 0$ dan $X' M_* X = mnI$, didapat \tilde{X} . Maka *Category*

quantification yang pertama diperoleh sebagai berikut $\tilde{Y}_j = D_j^{-1} G_j' \tilde{X}$

2. Transformasi baku *object scores*

Misalkan *matrik auxiliary scores* yang pertama Z dihitung dari $Z \leftarrow \sum_{j=1}^m M_j G_j \tilde{Y}_j$

dan dipusatkan (*centered*) ke M_* : $\tilde{Z} \leftarrow \{M_* - (M_* u u' M_* / u' M_* u)\} Z$

3. Ortonormalisasi

Masalah ortonormalisasi adalah menemukan M_* -ortonormal X^+ yang mendekati ke \tilde{Z} pada metode *least square*. Dalam HOMALS, dilakukan setting:

$X^+ \leftarrow m^{1/2} M_*^{-1/2} \text{GRAM}(M_*^{-1/2} \tilde{Z})$, yang mana sama dengan estimasi *least square* sampai dengan rotasi.

4. Mendapatkan *category quantification*

Untuk $j = 1, \dots, m$; *category quantification* yang baru dihitung dengan rumus:

$$Y^+ \leftarrow D_j^{-1} G_j' X^+$$

5. Test Kekonvergenan

Perbedaan antara nilai *loss function* $\sigma(\tilde{X} : \tilde{Y}) - \sigma(X^+ : Y^+)$ dibandingkan dengan kriteria konvergensi sebesar $\varepsilon \approx$ bilangan positif yang kecil. Ulangi langkah 2-4 selama perbedaan dari *loss function* masih melebihi ε . Mencapai konvergen, jika perbedaan nilai *loss function* lebih kecil atau sama dengan ε .

6. Rotasi

Sebagai indikasi pada langkah 3, selama iterasi orientasi dari X dan Y masih belum mudah diinterpretasikan, maka dilakukan rotasi.

2.2.2.2 Diagnostik

1. Rank Maksimum

Maksimum rank p_{\max} mengindikasikan jumlah maksimum dimensi yang dapat dihitung dari sebarang kumpulan data.

Secara umum adalah :

$$p_{\max} = \min \left\{ (n-1), \left(\sum_{j=1}^m k_j \right) - \max(m_1, 1) \right\},$$

dimana m_1 adalah jumlah variabel tanpa *missing values*. Meskipun jumlah dimensi *nontrivial* mungkin lebih kecil dari p_{\max} ketika $m=2$, *HOMALS* mengikuti semua dimensionalities sampai dengan p_{\max} . Penentuan banyaknya dimensi biasanya ditentukan dengan memperhatikan bahwa banyaknya dimensi tidak akan mengurangi informasi dari data sebelumnya. Umumnya penentuan banyaknya dimensi memperhatikan nilai akar ciri dari dimensi yang terakhir, yaitu lebih besar dari $1/m$.

2. Frekuensi Marginal (*Marginal Frequencies*)

Tabel Frekuensi memberikan marginal univariat dan jumlah *missing values* (suatu nilai yang berada diluar range analisis) untuk setiap variabel. Frekuensi marginal dihitung dari jumlahan kolom D_j atau total jumlahan dari M_j .

3. Ukuran Pembeda (*Discrimination Measure*)

Merupakan varian dimensi dari variabel yang dikuantifikasi (*quantified variables*). Untuk variabel j dan dimensi s maka :

$$\eta_{js}^2 = y'_{(j)s} D_j y_{(j)s} / n$$

dimana $y_{(j)s}$ adalah kolom ke- s dari Y_j , berhubungan dengan $G_j y_{(j)s}$, yaitu variabel yang dikuantifikasikan ke- s

4. Akar Ciri (*Eigenvalue*)

Penghitungan akar ciri yang dilaporkan adalah setelah konvergen pada langkah ke-6. Dengan option *history*, jumlah akar ciri yang dilaporkan selama iterasi adalah dengan judul '*Total Fit*'. Ternyata jumlah nilai akar ciri akan sama dengan *trace* dari matriks asli, jumlah ini dapat dihitung sebagai

$$1/m \sum_j \sum_s \eta_{js}^2. \text{ Nilai dari } \sigma(X:Y) \text{ adalah sama dengan } p-1/m \sum_j \sum_s \eta_{js}^2$$

2.2.3 Analisis Kelompok (*Cluster Analysis*)

Untuk mengelompokkan sekumpulan data yang diukur dari beberapa variabel, analisis statistik yang digunakan adalah analisa kelompok. Dengan analisa ini data yang terbagi dalam kelompok – kelompok yang sifatnya homogen antar anggota kelompok, sedang obyek antar kelompok mempunyai sifat yang relatif berbeda (*Karson*, 1982).

Untuk mengelompokkan individu berdasarkan ukuran keserupaan yang biasanya ditunjukkan oleh ukuran jarak. Jika ada n obyek dengan p variabel, maka sebelum dilakukan pengelompokkan, terlebih dahulu ditentukan ukuran kedekatan sifat antar obyek. Ukuran yang biasa digunakan sebagai interpretasi kedekatan jarak antara dua obyek adalah jarak euclid (*euclidean distance*). Jarak euclid antara $\mathbf{X}' = [X_1, X_2, \dots, X_p]$ dan $\mathbf{Y}' = [Y_1, Y_2, \dots, Y_p]$ adalah :

$$\begin{aligned} D_{(X,Y)}^2 &= (X_1 - Y_1)^2 + (X_2 - Y_2)^2 + \dots + (X_p - Y_p)^2 \\ &= (X - Y)'(X - Y) = (Y - X)'(Y - X) \dots \dots \dots (2.2) \end{aligned}$$

Semakin besar nilai D akan semakin jauh tingkat keserupaan antara kedua individu dan sebaliknya jika nilai dari D semakin kecil maka akan semakin dekat tingkat keserupaan antar kedua individu tersebut.

Dalam membentuk kelompok, terlebih dahulu ditentukan metode apa yang akan digunakan. Bila banyaknya kelompok yang akan muncul dari pengamatan ditentukan terlebih dahulu, maka lebih baik jika digunakan metode pengelompokan non hierarki. Sebaliknya jika banyaknya kelompok yang muncul tidak diketahui atau tidak dapat ditentukan banyaknya terlebih dahulu maka digunakan metode pengelompokan berhierarki.

2.2.3.1 Metode Pengelompokan Hierarki

Metode ini digunakan apabila banyaknya kelompok yang akan dipilih tidak ditentukan sebelumnya. Metode ini dimulai dengan mengelompokkan semua obyek menjadi n kelompok, dimana masing-masing kelompok beranggotakan satu obyek atau obyek itu sendiri. Penentuan banyaknya kelompok dan anggota-anggotanya ditentukan melalui diagram pohon (*dendogram*) yang terbentuk. Penentuan banyaknya kelompok biasanya sangat bergantung pada subyektifitas peneliti atas dasar pertimbangan-pertimbangan tertentu. Ada beberapa cara atau aturan dalam menentukan jarak antara dua kelompok, yaitu :

1. Metode Pautan Tunggal (*Single Linkage*)

Metode ini menggunakan aturan jarak minimum antar kelompok. Proses penggabungan diawali dengan menemukan dua obyek yang mempunyai jarak minimum untuk selanjutnya menjadi satu kelompok. Jarak minimum antara kelompok uv dan w adalah :

$$D_{(uv)w} = \min [D_{uw}, D_{vw}] \dots \dots \dots (2.3)$$

dimana : D_{uw} = jarak antara kelompok u dan w

D_{vw} = jarak antara kelompok v dan w

2. Metode Pautan Lengkap (*Complete Linkage*)

Metode ini menggunakan aturan jarak maksimum antar kelompok. Proses penggabungan diawali dengan dua obyek yang mempunyai jarak minimum.

Jarak maksimum antara kelompok uv dengan w adalah :

$$D_{(uv)w} = \max [D_{uw}, D_{vw}] \dots \dots \dots (2.4)$$

dimana : D_{uw} = jarak antara kelompok u dan w

D_{vw} = jarak antara kelompok v dan w

3. Metode Pautan Rata-rata (*Average Linkage*)

Sebagaimana halnya dengan dua metode sebelumnya, langkah awal metode ini dimulai dengan menemukan dua obyek yang mempunyai jarak minimum menjadi satu kelompok. Langkah selanjutnya didasarkan aturan rata-rata jarak antar dua kelompok. Jarak antara kelompok uv dan w adalah :

$$D_{(uv)w} = \frac{1}{N_{uv}N_w} \sum \sum D_{ij} \dots \dots \dots (2.5)$$

dimana :

N_{uv} = Jumlah anggota dalam kelompok uv

N_w = Jumlah anggota dalam kelompok w

D_{ij} = Jarak antar obyek j dalam kelompok uv dengan obyek j dalam kelompok w

Definisi penggabungan dari metode ini adalah :

$$D_{(uv)w} = (D_{uw} + D_{vw})/2$$

$$D_{(uvw)z} = (D_{uz} + D_{vz} + D_{wz})/3 \text{ dan seterusnya} \dots \dots \dots (2.6)$$

2.2.3.2 Metode Pengelompokan Non Hierarki

Metode ini dikenal sebagai metode *K-Means*, yang bertujuan mengelompokkan seluruh obyek (observasi) ke dalam k kelompok ($k < n$), dimana banyaknya kelompok yang akan dipilih ditentukan sebelumnya. Metode pengelompokan ini cocok untuk data multivariat yang berukuran besar (banyaknya sampel/observasi besar).

Langkah-langkah metode *K-Means* (dari *Mac Queen*) :

- 1) Partisi n obyek ke dalam k kelompok secara sembarang dan hitung centroidnya (mean) untuk masing-masing kelompok.
- 2) Hitung kuadrat jarak euclid (d^2_E) masing-masing obyek terhadap centroid k kelompok yang diberikan. Kemudian kelompokkan obyek-obyek tsb ke dalam kelompok yang mempunyai d^2_E terkecil.
- 3) Hitung centroid baru dari masing-masing kelompok, lalu kembali ke langkah 2, sampai hasilnya konvergen.

Metode *K-Means* akan mengelompokkan obyek sedemikian hingga jarak tiap-tiap obyek ke pusat kelompok dalam satu kelompok adalah minimum. Adapun prosedur *K-means* secara lengkap adalah sebagai berikut :

1. Menyusun matriks $n \times p$ dari hasil pengukuran n obyek dalam p variabel yang dinotasikan dengan $A(i,j)$, dimana $i = 1, 2, \dots, n$; dan $j = 1, 2, \dots, p$.
2. Menentukan besarnya K kelompok yang dibentuk.
3. Melakukan proses penyekatan untuk mengalokasikan tiap obyek dari n obyek ke dalam salah satu kelompok dari kelompok $1, 2, \dots, K$. Penyekatan ini dinotasikan $P(n, K)$.

Adapun langkah-langkah penyekatan adalah sebagai berikut:

- Menentukan kelompok awal dengan memasukkan obyek ke-i kedalam kelompok ke- l , dengan rumus :

$$l = \frac{K[\text{sum}(i) - \min]}{\max - \min} + 1 \dots\dots\dots(2.7)$$

dimana :

K = jumlah kelompok yang dikehendaki

$\text{Sum}(i)$ = penjumlahan elemen baris ke- i dari matriks $n \times p$

$$= A(i, 1) + A(i, 2) + \dots\dots\dots + A(i, p)$$

\max = nilai max dari $\text{sum}(i)$, untuk semua $i = 1, 2, \dots, n$

\min = nilai min dari $\text{sum}(i)$, untuk semua $i = 1, 2, \dots, n$

- Menghitung $B(l, j)$, yaitu rata-rata dari variabel ke- j pada semua obyek dalam kelompok ke- l dan menyusun matriks $l \times j$.
- Menghitung jarak Euclidean antara obyek ke- i dan kelompok ke- l dengan rumus jarak :

$$D(i, l) = \left(\sum_{j=1}^p [A(i, j) - B(l, j)]^2 \right)^{1/2} \dots\dots\dots(2.8)$$

dimana :

$D(i, l)$ = jarak antara obyek ke- i dan kelompok ke- l

$A(i, j)$ = elemen baris ke- i kolom ke- j dari matriks $n \times p$

$$(i = 1, 2, \dots, n; \text{ dan } j = 1, 2, \dots, p)$$

- Menghitung kesalahan penyekatan :

$$E[P(n, K)] = \sum_{i=1}^n D[i, l(i)]^2 \dots\dots\dots(2.9)$$

dimana :

$E[P(n, K)]$ = kesalahan dari penyekatan n obyek ke dalam K kelompok

$l(i)$ = kelompok dimana obyek ke- i berada

$D[i, l(i)]$ = jarak Euclidean antara obyek ke- i dan kelompok dimana obyek ke- i berada.

➤ Memeriksa apakah perpindahan obyek ke- i dari kelompok saat ini ke kelompok baru akan menghasilkan pengurangan kesalahan penyekatan. Perubahan kesalahan penyekatan saat ini $[l(i)]$ ke kelompok yang baru dinyatakan sebagai $R_{l(i),l}$.

Perubahan kesalahan penyekatan saat ini $[l(i)]$ ke kelompok yang baru

$$R_{l(i),l} = \frac{n(l)D(i,l)^2}{n(l)+l} - \frac{n[l(i)]D[i,l(i)]^2}{n[l(i)]-1} \dots\dots\dots (2.10)$$

dimana :

$R_{l(i),l}$ = perubahan kesalahan penyekatan dari kelompok saat ini $[l(i)]$ ke kelompok baru l .

$n[l(i)]$ = jumlah obyek yang berada dalam kelompok saat ini $[l(i)]$

$n(l)$ = jumlah obyek yang berada dalam kelompok baru (l) .

Bila $R_{l(i),l} < 0$, maka perpindahan obyek ke- i dari kelompok saat ini ke kelompok baru tidak memperkecil kesalahan penyekatan, sehingga obyek tersebut tidak perlu dipindahkan dari kelompoknya saat ini.

Bila $R_{l(i),l} > 0$, maka perpindahan obyek ke- i dari kelompok saat ini ke kelompok baru akan memperkecil kesalahan penyekatan, sehingga obyek harus dipindahkan dari kelompok saat ini kekelompok yang baru.

➤ Menghitung nilai $B(l, j)$ yang baru dan menyusun matriks $l \times j$ yang baru.

Selanjutnya kembali ke langkah 3, 4, 5 demikian seterusnya sampai tidak ada lagi perpindahan obyek yang bisa menghasilkan pengurangan kesalahan penyekatan.

2.2.4 Tabel Kontingensi (*Cross Tabulation*) Dua Dimensi

Tabel kontingensi dua dimensi adalah tabel hasil tabulasi silang antara dua variabel yang bersifat kategori dimana masing-masing variabel terdiri dari beberapa kategori. Masing-masing kategori tersebut harus memenuhi syarat-syarat sebagai berikut:

i. *Homogen*

Setiap kategori terdiri atas obyek yang sama, artinya obyek dengan karakteristik sama terletak dalam satu kategori.

ii. *Mutually exclusive* (saling asing)

Mutually exclusive artinya antara kategori yang satu dengan lainnya saling lepas, sehingga setiap pengamatan hanya akan termuat dalam salah satu kategori suatu variabel.

iii. *Mutually exhaustive*

Mutually exhaustive artinya variabel terpartisi habis oleh seluruh bagian variabel (kategori), sehingga setiap obyek pasti masuk pada salah satu kategori.

Tabel 2.2 di bawah menggambarkan suatu tabel kontingensi dua dimensi. Misal variabel A dan B merupakan variabel yang bersifat kategori (level), dimana variabel A terdiri atas I kategori dan B terdiri atas J kategori, sehingga tabel kontingensi dua dimensi memiliki ukuran $I \times J$ sel.

Tabel 2.2. Tabel Kontingensi Dua Dimensi

Variabel A	Variabel B					Total	
	1	2	j		J
1	X ₁₁	X ₁₂	X _{1j}	X _{1J}	X ₁₊
2	X ₂₁	X ₂₂	X _{2j}	X _{2J}	X ₂₊
•
•
i	X _{i1}	X _{i2}	X _{ij}	X _{iJ}	X _{i+}
•
•
I	X _{I1}	X _{I2}	X _{Ij}	X _{IJ}	X _{I+}
Total	X ₊₁	X ₊₂	X _{+j}	X _{+J}	X ₊₊

dimana :

X_{ij} = frekuensi dalam kategori ke-i variabel A, kategori ke-j variabel B

$$X_{i+} = \sum_{j=1}^J X_{ij} = \text{total frekuensi baris ke-i}$$

$$X_{+j} = \sum_{i=1}^I X_{ij} = \text{total frekuensi kolom ke-j}$$

$$X_{++} = \sum_{i=1}^I \sum_{j=1}^J X_{ij} = \text{total frekuensi seluruh observasi}$$

Tabulasi silang (*cross tabulation*) tersebut merupakan gabungan frekuensi distribusi dua variabel. Gabungan dua frekuensi distribusi ini secara statistik dapat ditest dengan menggunakan Chi-square statistik, untuk menentukan apakah hubungan antara dua variabel itu signifikan atau tidak. Ukuran asosiasi ini dapat juga ditunjukkan dengan ukuran statistik berupa koefisien kontingensi, phi, tau, gamma,

dan lainnya. Yang perlu diingat bahwa signifikansi antara dua variabel tidak hanya ditentukan oleh hubungan antar sel dalam tabel, tapi juga ditentukan oleh besarnya sampel yang diobservasi. Untuk sampel yang besar meskipun hubungan antar variabel tersebut lemah, namun secara statistik dapat menunjukkan signifikansi yang tinggi.

2.2.4.1 Ukuran Statistik dari Tabel Kontingensi

1) Chi-square :

Chi-square adalah uji statistik signifikansi yang membantu untuk menentukan ada tidaknya hubungan sistematis antara dua variabel. Hal ini dilakukan dengan jalan membandingkan frekuensi yang diobservasi dengan frekuensi yang diharapkan, dengan formula sebagai berikut :

$$\chi_{hitung}^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(x_{ij} - \hat{m}_{ij})^2}{\hat{m}_{ij}} ;$$

$$\text{dimana } \hat{m}_{ij} = \frac{X_{i+} \cdot X_{+j}}{n}$$

Untuk mengetahui ada tidaknya hubungan antara variabel A dan B digunakan uji independensi.

Hipotesis yang digunakan dalam pengujian independensi ini adalah:

H_0 : tidak ada hubungan antara variabel A dan B

H_1 : ada hubungan antara variabel A dan B

Dalam bentuk peluang, hipotesis H_0 dan H_1 dapat ditulis sebagai berikut:

$$H_0 : \pi_{ij} = \pi_{i+} \cdot \pi_{+j}$$

$$H_1 : \pi_{ij} \neq \pi_{i+} \cdot \pi_{+j}$$

Jika H_0 benar, maka : $\pi_{ij} = \pi_{i+} \cdot \pi_{+j}$

Taksiran parameter untuk π_{ij} adalah :

$$\hat{\pi}_{ij} = \frac{X_{i+}}{n} \cdot \frac{X_{+j}}{n} \dots\dots\dots (2.11)$$

Nilai harapan sel frekuensi: $E(X_{ij}) = m_{ij} = n \cdot \pi_{ij}$

Taksiran nilai harapan jika H_0 benar adalah :

$$\hat{m}_{ij} = n \cdot \hat{\pi}_{ij} = n \hat{\pi}_{i+} \cdot \hat{\pi}_{+j} = n \cdot \frac{X_{i+}}{n} \cdot \frac{X_{+j}}{n}$$

$$\hat{m}_{ij} = \frac{X_{i+} \cdot X_{+j}}{n} \dots\dots\dots (2.12)$$

Statistik uji yang digunakan adalah:

$$\chi^2_{hitung} = \sum_{i=1}^I \sum_{j=1}^J \frac{(x_{ij} - \hat{m}_{ij})^2}{\hat{m}_{ij}} \dots\dots\dots (2.13)$$

Jika H_0 benar, maka statistik uji ini (χ^2_{hitung}) mendekati distribusi *Chi-Square* dengan $db=(I-1)(J-1)$.

Kriteria penolakan H_0 adalah jika $\chi^2_{hitung} \geq \chi^2_{(db, \alpha)}$ atau $p\text{-value} \leq \alpha$.

Dari pengujian ini didapatkan koefisien chi-square, besarnya derajat bebas (df), dan level signifikansinya. Level signifikansi ini dapat diplotkan dalam kurva Chi-square, apakah terletak di daerah penolakan atau di daerah penerimaan. Bila terletak dalam daerah penolakan, berarti kita menolak H_0 dan menerima H_1 , atau sebaliknya. Bila level signifikansi adalah 0,0000 berarti hubungan antara dua variabel tersebut sangat erat dengan tingkat signifikansi kurang dari 0,001. Hal ini menunjukkan tingkat signifikansi yang sangat tinggi (*very highly significant*). Didalam tabel chi-square level signifikansi selain kurang dari 0,001 adalah kurang dari 0,01; kurang dari 0,05 dan kurang dari

0,10. Chi-square hanya membantu untuk menentukan apakah dua variabel tersebut independent atau dependent. Chi-square tidak menunjukkan seberapa kuat hubungan antar variabel tersebut.

2) Phi :

Untuk tabel 2x2, Phi statistik sangat cocok untuk mengukur seberapa jauh kekuatan hubungan antara dua variabel. Phi adalah mengoreksi proporsi dengan jumlah sampel, dengan perumusan sbb :

$$\Phi = \left(\frac{\chi^2}{n} \right)^{1/2} \dots \dots \dots (2.14)$$

Nilai Phi sama dengan 0 bila dua variabel tidak ada hubungan, dan sama dengan 1 bila dua variabel mempunyai hubungan kuat.

3) Cramer's V :

Cramer's V adalah modifikasi dari Phi. Cramer's V lebih cocok bila yang dianalisis berupa tabel yang besar. Dengan perumusan sbb :

$$V = \left(\frac{\Phi^2}{\min(I - 1), (J - 1)} \right)^{1/2} \dots \dots \dots (2.15)$$

V juga bergerak dari 0 sampai +1, kalau banyak kategori dimasukkan.

4) Koefisien Kontingensi :

Ukuran tingkat hubungan antar variabel yang lain yang berdasarkan Chi-square adalah koefisien kontingensi, dengan perumusan sbb :

$$C = \left(\frac{\chi^2}{\chi^2 + n} \right)^{1/2} \dots \dots \dots (2.16)$$

Koefisien kontingensi dapat digunakan untuk menganalisis tabel berbagai ukuran. Seperti halnya Cramer's V, koefisien kontingensi mempunyai nilai

minimum sama dengan 0 namun nilai maksimumnya adalah tergantung dari ukuran tabel yang dianalisis, untuk tabel 2x2 nilai maksimumnya adalah 0,707. Dengan demikian koefisien kontingensi C lebih cocok untuk membandingkan tabel yang sama antara baris dan kolomnya.

2.2.5 Analisis Diskriminan

Analisis diskriminan digunakan untuk mengetahui apakah kelompok-kelompok yang telah terbentuk berbeda secara statistika. Analisis diskriminan atau pengklasifikasian merupakan bagian dari teknik analisis multivariat untuk memisahkan terhadap obyek-obyek kedalam alokasi obyek baru berdasarkan grup-grup yang ada.

Analisis diskriminan dikembangkan oleh *Renold Fisher* yang pada dasarnya bertujuan :

- Mendapatkan kombinasi linier saling bebas dari beberapa variabel yang dapat digunakan sebagai dasar pengelompokkan suatu obyek atau kasus kedalam suatu kelompok tertentu.
- Mendapatkan variabel pembeda untuk tiap-tiap kelompok. Jadi dengan analisis dapat diketahui apakah pengelompokkan yang dibuat secara statistik nyata atau tidak. Juga akan didapatkan variabel yang membedakan antar kelompok atau variabel pembeda.

Dalam analisis diskriminan masing-masing populasi diamati dalam beberapa variabel random (p) dan berdistribusi probabilitas bersama tertentu. Misalkan ada k populasi $\pi_1, \pi_2, \dots, \pi_k$. Dan x adalah vektor untuk p variabel random dan $f_i(x)$ adalah distribusi probabilitas dalam p variabel dari x dari populasi ke- i (π_i).

Masalah yang dihadapi dalam analisis diskriminan adalah menentukan partisi dari ruang sampel menjadi beberapa daerah R_1, R_2, \dots, R_k sehingga jika suatu obyek masuk dalam daerah R_i maka obyek tersebut ditempatkan dalam populasi π_i .

P_1, P_2, \dots, P_k adalah probabilitas suatu obyek masuk dalam $\pi_1, \pi_2, \dots, \pi_k$. P_{i+j} adalah probabilitas bersyarat sebuah obyek masuk kedalam π_j pada saat obyek sebenarnya berada dalam π_i dan anggota populasi ke- j (π_j), dinyatakan :

$$P_{ij} = P_{i+j} P_j \dots \dots \dots (2.17)$$

$E_j = \Pr$ (obyek berasal dari π_j dan salah masuk kelompok)

$$= \sum_{i=1}^k P_{i,j} \dots \dots \dots (2.18)$$

$$= P_j \sum_{i=1}^k P_{i,j} \quad j = 1, 2, \dots, k \quad i \neq j$$

$f_j(x)$ adalah fungsi distribusi dari π_j , maka :

$$P_{i,j} = \int_{R_i} f_j(x) dx \dots \dots \dots (2.19)$$

$P_{i,j}$ adalah probabilitas suatu obyek ditempatkan didaerah R_i padahal obyek tersebut masuk kedaerah R_j , sehingga :

$$E = \sum_{j=1}^k E_j$$

$$= 1 - \sum_{j=1}^k P_j \int_{R_i} f_j(x) dx \dots \dots \dots (2.20)$$

Permasalahan yang dihadapi sekarang adalah memilih R_i sedemikian hingga meminimumkan probabilitas salah masuk kelompok. Misal ada dua kelompok maka :

$$1 - E = P_1 \int_{R_1} f_1(x) dx + P_2 \int_{R_2} f_2(x) dx \dots \dots \dots (2.21)$$

dimana :

$$P_1 + P_2 = 1$$

R_1 dan R_2 partisi ruang sampel dari x

Kedua daerah tersebut dapat dinyatakan sebagai :

$$R_1 = \left[x \left| \frac{f_1(x)}{f_2(x)} \geq \frac{P_2}{P_1} \right. \right]$$

$$R_2 = \left[x \left| \frac{f_1(x)}{f_2(x)} < \frac{P_2}{P_1} \right. \right]$$

Jika ada k daerah :

$$R_j = \left[x \left| \frac{f_j(x)}{f_i(x)} > \frac{P_i}{P_j} \right. \right] \dots \dots \dots (2.22)$$

$$i=1,2,\dots,k \quad j=1,2,\dots,k \quad i \neq j$$

Jika $f_j(x)$ multivariate normal dengan mean μ_j dan matrik varian kovarian Σ , maka:

$$R_j = \left[x \left| x' \Sigma^{-1}(\mu_j - \mu_i) - (\mu_j - \mu_i)' \Sigma^{-1}(\mu_j - \mu_i) / 2 > \ln \left(\frac{P_i}{P_j} \right) \right. \right]$$

$$i = 1,2,\dots,k \quad j = 1,2,\dots,k \quad i \neq j$$

$$= \left[x \left| Y_{ij} > \ln \left(\frac{P_i}{P_j} \right) \right. \right]$$

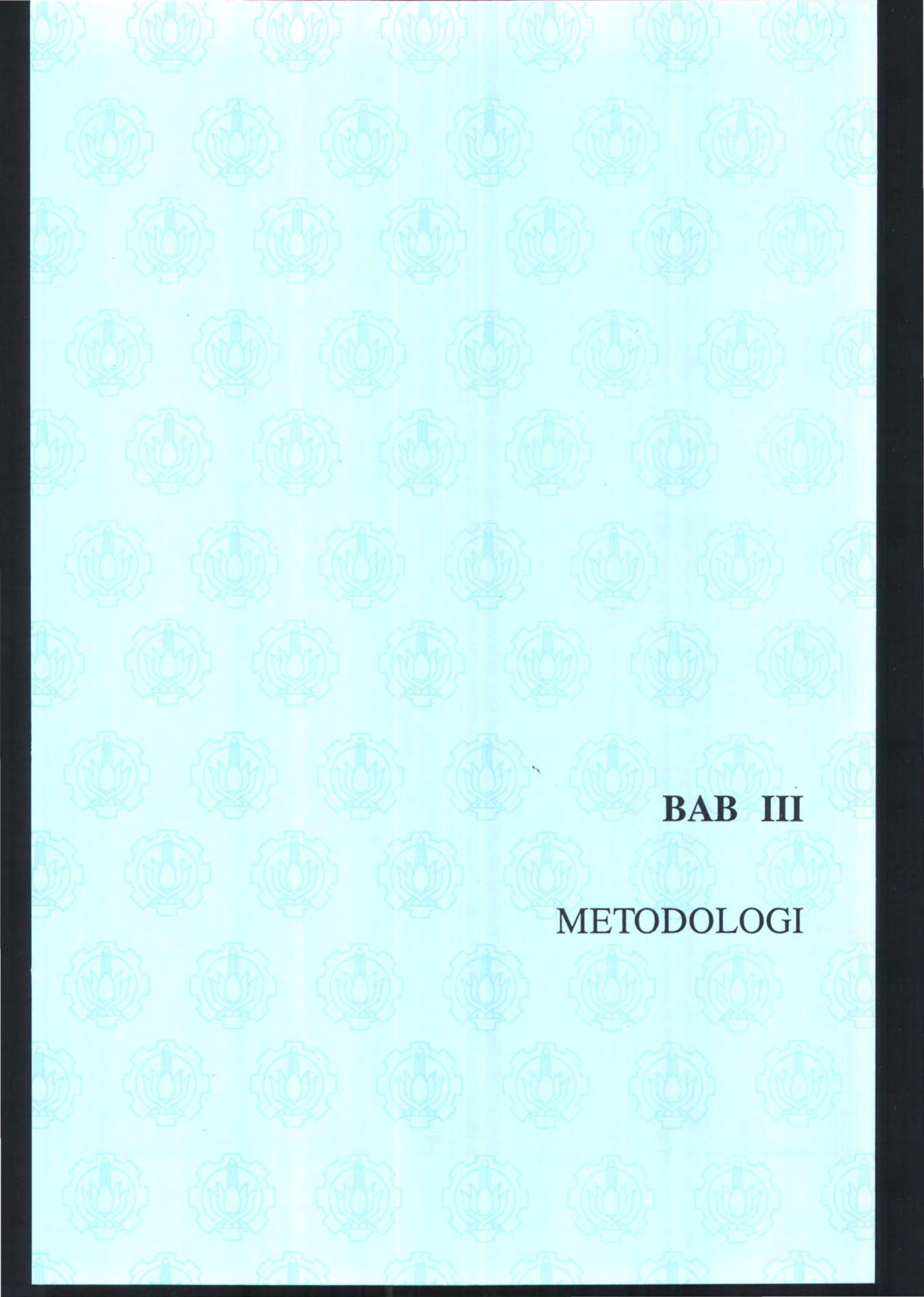
$$i = 1,2,\dots,k \quad j = 1,2,\dots,k \quad i \neq j$$

Jika obyek berasal dari π_j , maka pengamatan n_1, n_2, \dots, n_k cukup besar, maka Y_{ij} didekati dengan distribusi normal. Dalam hal ini π_j ditaksir dengan x_j dan Σ dengan s maka :

$$Y_{ij} = x' s^{-1}(x_j - x_i) - (x_j - x_i)' s^{-1}(x_j - x_i) / 2$$

Apabila penelitian melibatkan banyak variabel, maka untuk efisiensi dalam menentukan variabel mana yang berperan dalam pembentukan fungsi diskriminan dapat ditempuh melalui analisis diskriminan bertatar (*stepwise Discriminant*).

Analisis ini digunakan untuk menghilangkan informasi dari variabel yang kurang berguna dalam pembentukan fungsi diskriminan. Analisis diskriminan bertatar dimulai dengan pemilihan variabel pembeda yang paling berarti (yang mempunyai nilai F terbesar atau nilai Wilk's Lambda terkecil). Selanjutnya variabel pembeda kedua dipilih yang paling berarti berikutnya, yang mampu meningkatkan kriteria perbedaan setelah dikombinasi dengan variabel terpilih pertama. Variabel pembeda ketiga dan seterusnya juga dipilih menurut kemampuannya menyumbang perbedaan setelah dikombinasikan dengan variabel terpilih sebelumnya. Pada setiap langkah, variabel yang telah dipilih dapat dikeluarkan dari analisa jika ternyata variabel tersebut melemahkan daya perbedaan setelah dikombinasikan dengan variabel pembeda yang terpilih kemudian.



BAB III

METODOLOGI

BAB III

METODOLOGI

3.1 Bahan Penelitian

Data yang digunakan dalam penelitian kali ini adalah data sekunder yang berasal dari BPS Propinsi Jawa Timur tentang penduduk miskin tahun 2001. Secara metodologis untuk memperkirakan penduduk miskin yang telah dilakukan BPS menggunakan 2 pendekatan yaitu pendekatan wilayah dan pendekatan rumah tangga. Inti pendekatan yang pertama adalah memperkirakan penduduk miskin melalui kantong kantong kemiskinan yang berupa desa miskin (tertinggal). Sedangkan pendekatan yang kedua adalah pendekatan rumah tangga yang mengacu kepada ketidakmampuan rumah tangga dalam memenuhi kebutuhan minimum.

Dalam penelitian kali ini rumah tangga miskin yang diambil sebagai sampel adalah sebesar 16990 Rumah tangga yang tersebar di seluruh kota Surabaya yang terbagi menjadi 5 wilayah :

1. Surabaya Pusat, terdiri dari 4 kecamatan
2. Surabaya Utara, terdiri dari 4 kecamatan
3. Surabaya Timur, terdiri dari 7 kecamatan
4. Surabaya Selatan, terdiri dari 8 kecamatan
5. Surabaya Barat, terdiri dari 5 kecamatan

Banyaknya rumah tangga yang terpilih sebagai sampel di kota Surabaya adalah sebagai berikut :

Tabel 3.1 Jumlah Rumah Tangga sampel menurut wilayah & kecamatan

Wilayah	Kecamatan	Jumlah Rumah Tangga
Pusat	Tegalsari	544
	Bubutan	255
	Genteng	425
	Simokerto	935
Utara	Kenjeran	390
	Semampir	2550
	Pabean Cantikan	830
	Krembangan	383
Timur	Rungkut	816
	Tenggilis Mejoyo	221
	Gunung Anyar	272
	Sukolilo	595
	Mulyorejo	170
	Gubeng	460
	Tambaksari	1496
Selatan	Karangpilang	170
	Wiyung	340
	Dukuh Pakis	255
	Wonocolo	300
	Jambangan	340
	Gayungan	85
	Wonokromo	800
	Sawahan	1630
Barat	Lakarsantri	650
	Tandes	715
	Sukomanunggal	272
	Asemrowo	462
	Benowo	629
Total Rumah Tangga		16990

3.2 Variabel Penelitian

Variabel – variabel yang digunakan dalam penelitian ini adalah :

I. Variabel Rumah Tangga Miskin terdiri dari :

➤ Variabel 1 : Pembelian Pakaian dalam setahun

Pakaian adalah sesuatu yang layak untuk dipakai yang merupakan alat penutup badan yang biasanya terdiri dari penutup atas yang terdiri dari baju, jaket kemeja, Tshirt, jas. Serta penutup bawah misalnya celana pendek, celana panjang, sarung dan sebagainya. Pembelian yang dicatat adalah pembelian pakaian baru selama satu tahun. Semakin makmur ekonomi seseorang, kebutuhan sandang semakin diperhatikan dan sebaliknya.

Variabel ini terdiri dari 3 kategori pilihan :

- Kapan saja
- 2 – 3 kali
- Maks 1 kali

➤ Variabel 2 : Fasilitas air bersih

Pengelompokkan didasarkan atas faktor fasilitas air bersih, apakah perusahaan pribadi atau bersama. Kecenderungan orang yang berekonomi layak akan mengusahakan kebutuhan air bersihnya sendiri, baik dengan cara membeli air kemasan, atau dengan cara lain seperti berlangganan PDAM atau membeli pompa untuk mengusahakan air minum bersyarat kesehatan. Sedangkan perusahaan secara bersama seperti membeli pompa secara bersama, membuat sumur untuk beberapa keluarga, perusahaan mata air secara bersama dan lain-lain. Sedangkan air hujan dan sungai disebut pemakaian umum.

Variabel ini terdiri dari 3 kategori :

- Sendiri, apabila fasilitas air bersih digunakan khusus oleh rumah tangga tersebut saja.
- Bersama, apabila fasilitas air bersih digunakan oleh beberapa rumah tangga tertentu.
- Umum/tidak ada, apabila fasilitas air bersih dapat digunakan oleh semua orang yang tinggal disekitarnya.

➤ **Variabel 3 : Persentase Pengeluaran Makanan**

Persentase pengeluaran untuk konsumsi rumah tangga . Dilihat dari prioritas pengeluaran (pengeluaran sebagai proxy pendapatan), semakin besar pengeluaran untuk makanan, maka semakin miskin keadaan seseorang. Dasar yang melandasi adalah hukum Engel yang sudah dibuktikan dengan data Susenas 1996 dan 1999.

Variabel ini terdiri dari 3 kategori :

- < 60 %
- 60 % - 70 %
- > 70 %

➤ **Variabel 4 : Kepemilikan Rumah**

Pengelompokkan berdasarkan atas faktor kepemilikan, bahwa orang yang mampu biasanya mengusahakan kepemilikan rumah sendiri.

Variabel ini terdiri dari 3 kategori :

- Milik Sendiri, rumah / tempat tinggal yang benar-benar menjadi milik rumah tangga.
- Sewa/kontrak, tempat tinggal yang disewa secara teratur.

- Bebas sewa, tempat tinggal yang dihuni berasal dari pihak lain tanpa bayaran apapun.

➤ **Variabel 5 : Jenis Dinding**

Didasarkan kemampuan untuk mengusahakan pemenuhan kebutuhan rumah tangga yang sehat dan rapi, biasanya akan mengupayakan dinding yang kokoh, nyaman untuk ditempati.

Variabel ini terdiri dari 3 kategori :

- Tembok
- Papan kayu sederhana
- Bambu/lainnya

➤ **Variabel 6 : Jenis Lantai**

Luas lantai yang ditempati dan digunakan untuk kebutuhan sehari-hari. Bagian yang bukan termasuk luas lantai adalah lantai ruang khusus untuk usaha, lantai kandang ternak, lantai jemuran. Didasarkan atas kemampuan mengusahakan pengubinan lantai yang dipunyai supaya memenuhi syarat kesehatan dan estetika.

Variabel ini terdiri dari 3 kategori :

- Ubin
- Plester semen, pasangan bata, papan
- Tanah / pasir

➤ **Variabel 7 : Sarana Tempat Buang Air Besar**

Didasarkan atas kemampuan mengusahakan tempat buang air besar (TBAB) yang sehat dan estetis, pada rumah tangga miskin biasanya sudah tidak mampu lagi memikirkan untuk mengusahakan TBAB yang sehat dan estetis.

Variabel ini terdiri dari 3 kategori :

- Sendiri, adalah TBAB yang digunakan rumah tangga bersangkutan yang biasanya terletak didalam lingkungan rumah.
- Bersama, adalah bukan milik rumah tangga bersangkutan, dan digunakan untuk beberapa rumah tangga, biasanya terletak diluar tempat tinggal yang bersangkutan.
- Umum, bukan milik pribadi yang digunakan untuk umum, biasanya terdapat di tempat keramaian.

➤ **Variabel 8 : Sumber Penerangan**

Didasarkan atas kemampuan mengusahakan penerangan didalam rumahnya. Semakin miskin biasanya fasilitas penerangan yang mampu diusahakan juga sangat minimal.

Variabel ini terdiri dari 3 kategori :

- Listrik sendiri, sumber penerangan PLN maupun non PLN yang khusus digunakan oleh rumah tangga tersebut (pelanggan).
- Listrik bersama / petromak, apabila sumber penerangan digunakan oleh beberapa rumah tangga tertentu tanpa pencatat beban sendiri – sendiri.
- Pelta / senter

➤ **Variabel 9 : Pendidikan Anggota Rumah Tangga umur 6 – 15 tahun**

Didasarkan kemampuan menyekolahkan anggota rumah tangga pada usia sekolah, kalau rumah tangga miskin biasanya ada anggota rumah tangga usia sekolah yang tidak sekolah.

Variabel ini terdiri dari 3 kategori :

- Semua sekolah
- Ada yang tidak sekolah
- Tidak ada yang sekolah

➤ **Variabel 10 : Sumber Keuangan Rumah Tangga**

Didasarkan pada pekerjaan tetap memberikan pendapatan yang stabil sehingga mudah dalam memenuhi kebutuhan seluruh anggota rumah tangga

Variabel ini terdiri dari 3 kategori :

- Tetap
- Tidak tetap
- Tidak ada

➤ **Variabel 11 : Pelayanan Kesehatan**

Didasarkan kemampuan mengusahakan perolehan pelayanan kesehatan, semakin miskin akan sulit memperoleh pelayanan kesehatan yang baik

Variabel ini terdiri dari 3 kategori :

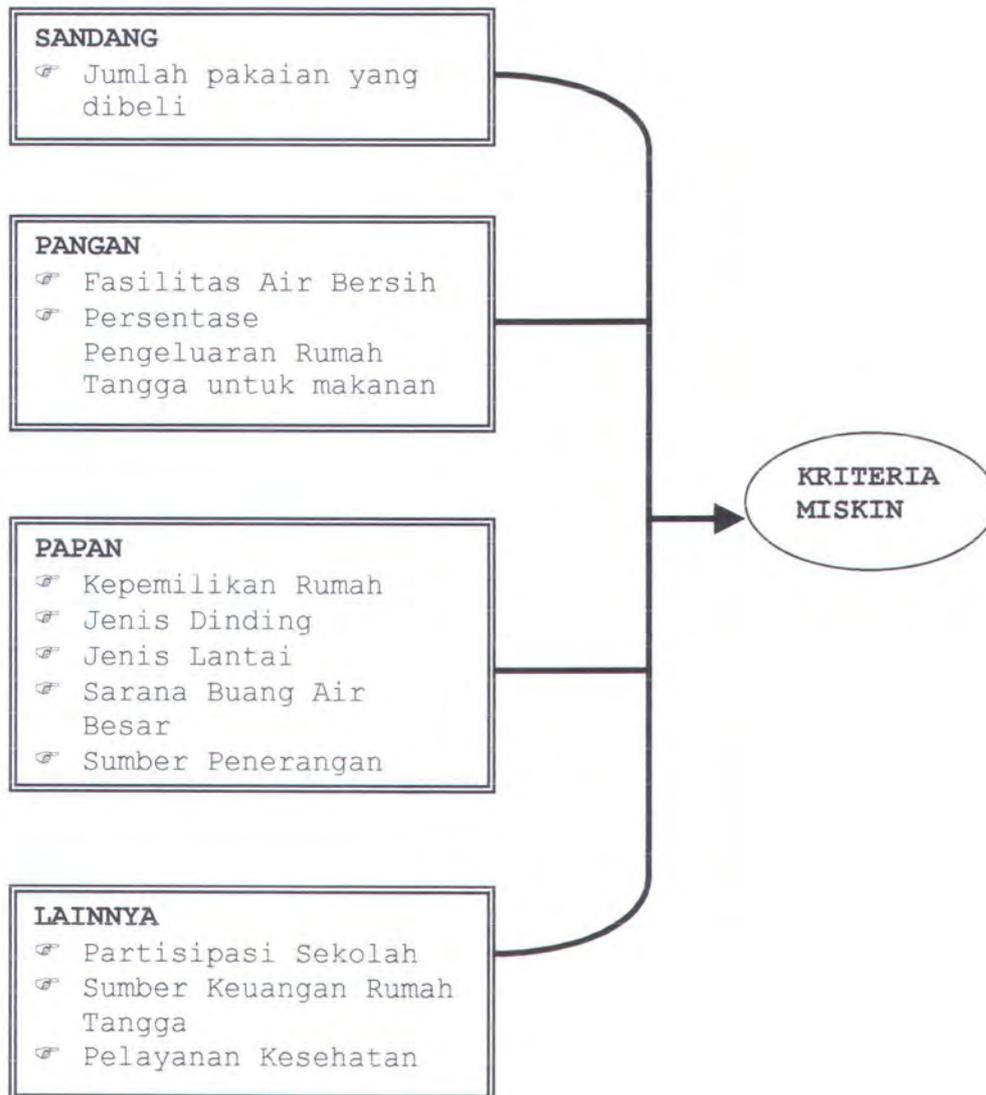
- Dokter Praktek
- RSUD, Polindes, Puskesmas, Bidan, Mantri
- Lainnya



II. Variabel Geografi

Variabel geografi yaitu wilayah yang menunjukkan rumah tangga sampel berada. Variabel geografi ini dibedakan menjadi Surabaya Pusat, Utara, Timur, Selatan dan Surabaya Barat.

Kerangka pikir pengukuran rumah tangga miskin :



Tentunya spesifikasi indikator variabel lainnya masih ada seperti variabel tempat tinggal, urban atau rural, variabel pendidikan kepala rumah tangga dan sebagainya. Walaupun demikian, didasarkan atas faktor dominan yang mempengaruhi kriteria rumah tangga miskin, serta aspek efisiensi baik dalam pendataan maupun pencacahan yang juga didukung oleh faktor pembiayaan survei, rumusan kesebelas variabel tersebut cukup mewakili dan memadai.

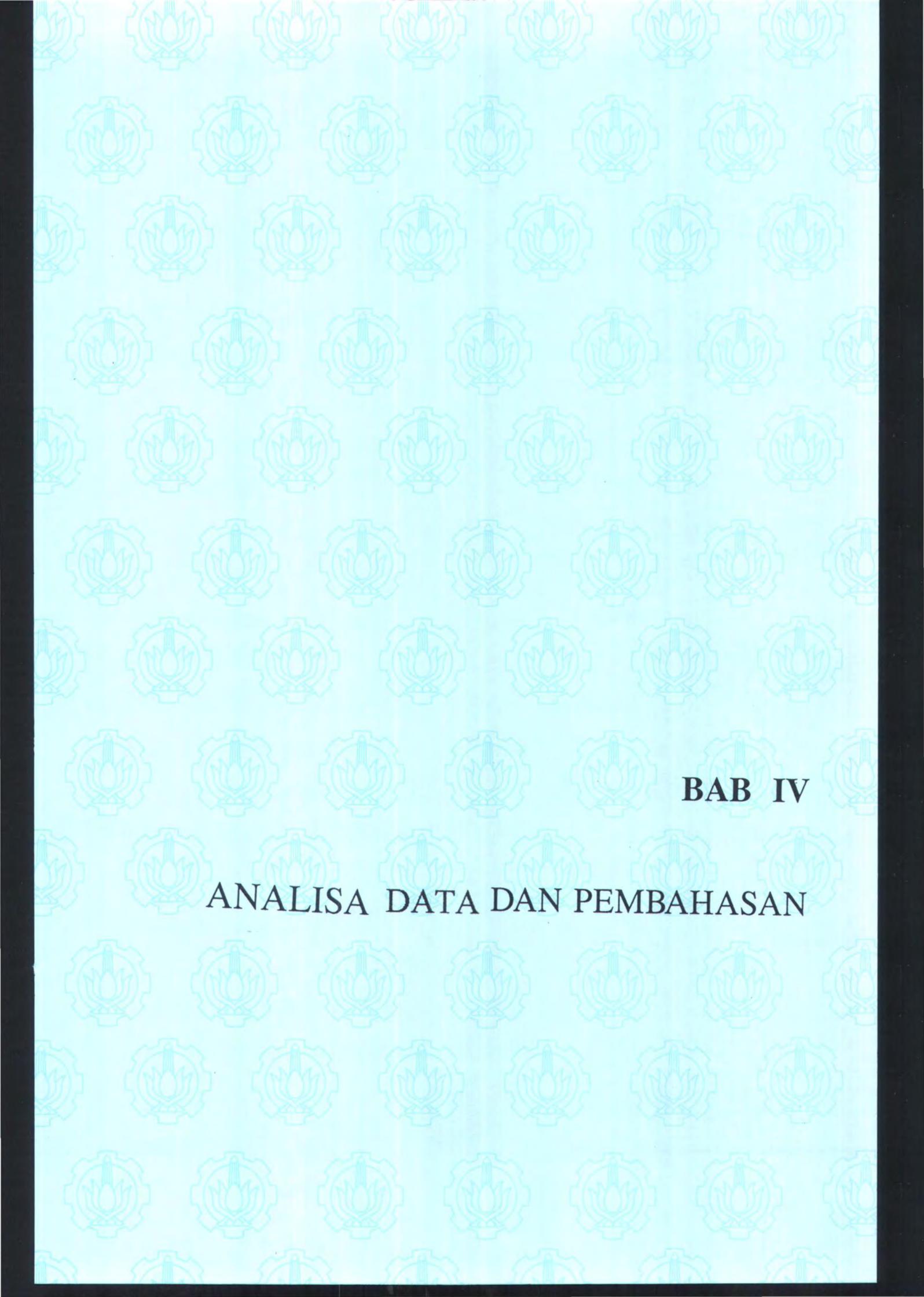
3.3 Metode Penelitian

Tujuan dari Tugas Akhir ini adalah untuk mengkaji dan menerapkan metode Analisis Homogeneity yang analog dengan *Categorical PCA* yang dilanjutkan dengan analisis faktor. Kemudian untuk mengetahui segmentasi rumah tangga sampel digunakan analisis kelompok yang dilanjutkan dengan analisis silang antara segmen dan variabel demografi rumah tangga untuk mengetahui karakteristik rumah tangga dari tiap segmen.

Pengolahan data dengan menggunakan metode Homals, dilakukan melalui paket program SPSS 8.0 *for windows* terhadap 11 variabel penduduk miskin. Setelah itu dilanjutkan dengan melakukan analisis data dan interpretasi output hasil pengolahan dengan metode Homals yang berupa data baru (obyek skor dan kategori kuantifikasi) menurut variabel baru.

Setelah itu akan dilakukan analisis lanjutan yaitu antara lain :

- ☞ Analisis Kelompok terhadap obyek skor untuk seluruh dimensi yang ada. Analisis kelompok ini bertujuan untuk mendapatkan kelompok – kelompok rumah tangga miskin yang mempunyai keserupaan (relatif homogen), dan antar kelompok mempunyai sifat yang relatif heterogen.
- ☞ Tabulasi silang antara segmen dengan variabel geografi rumah tangga maupun variabel lain untuk mengetahui karakteristik tiap segmen.
- ☞ Melakukan analisis diskriminan untuk mengetahui variabel pembeda dan untuk mengetahui apakah pengelompokan yang dibuat sudah signifikan atau belum secara statistik.



BAB IV

ANALISA DATA DAN PEMBAHASAN

BAB IV

ANALISIS DATA DAN PEMBAHASAN

Setelah dilakukan pengolahan terhadap data mengenai rumah tangga miskin di kota Surabaya tahun 2001, yang terdiri dari 11 variabel dengan menggunakan paket program SPSS 8.0 *for windows*, Ada satu variabel yaitu variabel Prosentase Pengeluaran Makanan (Var_3) semua responden ternyata menjawab >70 % dari penghasilan digunakan untuk makanan. Sehingga untuk variabel Prosentase Pengeluaran Makanan tidak diikutkan dalam analisis. Hasil yang didapat adalah sebagai berikut :

4.1 *Homogeneity Analysis*

Pengolahan terhadap 10 variabel rumah tangga miskin di Kota Surabaya tahun 2001 dengan menggunakan *Homals* diperoleh hasil 6 dimensi yang mampu menerangkan sekitar 88,79 % dari total keragaman dengan perincian sebagai berikut :

Tabel 4.1 Proporsi keragaman yang dapat diterangkan oleh Homals

Dimensi	Akar Ciri	Proporsi Keragaman (%)	Proporsi Keragaman Kumulatif (%)
1	0,2445	24,45	24,45
2	0,1647	16,47	40,92
3	0,1342	13,42	54,34
4	0,1209	12,09	66,43
5	0,1163	11,63	78,06
6	0,1073	10,73	88,79

Dari output dapat diketahui nilai akar ciri (*eigenvalue*) untuk tiap dimensi. Nilai akar ciri digunakan untuk menentukan jumlah dimensi yang digunakan dalam analisis, nilai akar ciri yang diambil adalah yang bernilai lebih besar dari pada 1/jumlah variabel

(1/10 atau 0,10). Dari hasil diatas akar ciri yang melebihi nilai 0,10 sebanyak 6 akar ciri yang akhirnya diputuskan menggunakan 6 dimensi. Dari ke 6 dimensi dapat menerangkan keragaman sebesar 88,79 % dari total keragaman, dimana dimensi pertama mampu menerangkan sebesar 24,45 % dari total keragaman, dimensi kedua mampu menerangkan sebesar 16,47 % dari total keragaman, kemudian dimensi ketiga mampu menerangkan sebesar 13,42 % dari total keragaman, dan dimensi keempat mampu menerangkan sebesar 12,09 % dari total keragaman, dimensi kelima menerangkan sebesar 11,67 % dan dimensi keenam mampu menerangkan 10,73 % dari total keragaman. Sedangkan dimensi berikutnya kemungkinan tidak akan lebih banyak menambah informasi. Jumlah dimensi maksimum yang mungkin terbentuk adalah sejumlah variabel yang ada dikurangi satu (9 dimensi).

Selain itu juga dihasilkan matriks *Discrimination Measures* (ukuran pembeda) per variabel per dimensi, yang merupakan varian dimensi dari variabel yang dikuantifikasi. *Discrimination Measures* disebut juga *Contribution of Variables* yaitu nilai yang menunjukkan kontribusi masing-masing variabel asal pada setiap dimensi yang terbentuk, analog dengan *loading factor* pada Analisis Faktor. Dari nilai *discrimination measure* ini dapat diketahui variabel-variabel yang mendukung dimensi tertentu.

Tabel 4.2 Matriks *Discrimination Measures* per Variabel per Dimensi

Variabel	Dimensi 1	Dimensi 2	Dimensi 3	Dimensi 4	Dimensi 5	Dimensi 6
1	0.020	0.046	0.018	0.020	0.318	0.178
2	0.260	0.165	0.023	0.078	0.339	0.054
4	0.556	0.012	0.280	0.037	0.028	0.074
5	0.145	0.493	0.035	0.281	0.036	0.026
6	0.107	0.435	0.144	0.188	0.124	0.037
7	0.631	0.349	0.007	0.017	0.081	0.014
8	0.692	0.085	0.093	0.030	0.008	0.103
9	0.019	0.002	0.127	0.198	0.043	0.227
10	0.006	0.049	0.275	0.224	0.105	0.068
11	0.009	0.010	0.341	0.138	0.080	0.292

Berdasarkan tabel 4.3 di atas dapat diketahui variabel-variabel yang masuk ke dalam dimensi tertentu, yaitu :

- 1) **Dimensi pertama** didukung oleh variabel Kepemilikan rumah (Var_4) dengan nilai discrimination 0,556, Sarana Tempat Buang Air Besar (Var_7) dengan nilai discrimination 0,631, Sumber penerangan (Var_8) dengan nilai discrimination 0,692.
- 2) **Dimensi kedua** didukung oleh variabel Jenis Dinding (Var_5) dengan nilai discrimination 0,493, Jenis Lantai (Var_6) dengan nilai discrimination 0,435
- 3) **Dimensi ketiga** didukung oleh variabel Sumber keuangan rumah tangga (Var_10) dengan nilai discrimination 0,275 dan variabel Pelayanan kesehatan (Var_11) dengan nilai discrimination 0,341.
- 4) **Dimensi keempat** tidak ada variabel yang mendukung. Hal ini mengindikasikan tidak ada variabel yang *dominan* mendukung dimensi keempat, tetapi tetap ada kontribusi dari tiap-tiap variabel pada dimensi ini.

- 5) **Dimensi kelima** didukung oleh variabel Pembelian pakaian dalam setahun (Var_1) dengan nilai discrimination 0,318 dan variabel Fasilitas Air Bersih (Var_2) dengan nilai discrimination 0,339.
- 6) **Dimensi keenam** didukung oleh variabel Pendidikan ART umur 6 – 15 tahun (Var_9) dengan nilai discrimination 0.227

4.2. Analisis Cluster

Setelah diketahui kontribusi tiap variabel terhadap masing masing dimensi, kemudian dilakukan analisis lanjutan, yaitu analisis *cluster*. Analisis kelompok ini bertujuan memisahkan obyek pengamatan kedalam beberapa kelompok yang memiliki sifat berbeda antar kelompok, sehingga obyek yang terletak dalam satu kelompok akan mempunyai sifat yang relatif homogen. Karena kota Surabaya terbagi dalam 5 wilayah serta data yang digunakan dalam penelitian ini berukuran cukup besar, yaitu 16990 rumah tangga sampel, maka metode yang akan digunakan adalah pengelompokan non hierarkhi (*k-means method*) dengan $k=5$. Untuk menentukan segmentasi rumah tangga sampel, dilakukan cluster terhadap obyek skor (X) pada keenam dimensi yang ada. Dari hasil pengelompokan menjadi 5 segmen ini perlu dilakukan uji kesamaan rata-rata cluster, untuk mengetahui apakah pembentukan menjadi 5 segmen tersebut sudah signifikan, dilakukan Uji kesamaan rata-rata dengan hipotesa yang digunakan adalah:

1. Uji kesamaan rata-rata cluster pada dimensi 1:

$$H_0 : \pi_1 = \pi_2 = \pi_3 = \pi_4 = \pi_5 \text{ (tidak ada perbedaan antar cluster)}$$

$$H_1 : \text{minimal ada satu } \pi_i \text{ yang tidak sama; } i=1,2,\dots,5;$$

(ada perbedaan antar cluster)

2. Uji kesamaan rata-rata cluster pada dimensi 2:

$H_0 : \pi_1 = \pi_2 = \pi_3 = \pi_4 = \pi_5$ (tidak ada perbedaan antar cluster)

$H_1 : \text{minimal ada satu } \pi_i \text{ yang tidak sama; } i=1,2,\dots,5;$

(ada perbedaan antar cluster)

3. Uji kesamaan rata-rata cluster pada dimensi 3:

$H_0 : \pi_1 = \pi_2 = \pi_3 = \pi_4 = \pi_5$ (tidak ada perbedaan antar cluster)

$H_1 : \text{minimal ada satu } \pi_i \text{ yang tidak sama; } i=1,2,\dots,5;$

(ada perbedaan antar cluster)

4. Uji kesamaan rata-rata cluster pada dimensi 4:

$H_0 : \pi_1 = \pi_2 = \pi_3 = \pi_4 = \pi_5$ (tidak ada perbedaan antar cluster)

$H_1 : \text{minimal ada satu } \pi_i \text{ yang tidak sama; } i=1,2,\dots,5;$

(ada perbedaan antar cluster)

5. Uji kesamaan rata-rata cluster pada dimensi 5:

$H_0 : \pi_1 = \pi_2 = \pi_3 = \pi_4 = \pi_5$ (tidak ada perbedaan antar cluster)

$H_1 : \text{minimal ada satu } \pi_i \text{ yang tidak sama; } i=1,2,\dots,5;$

(ada perbedaan antar cluster)

6. Uji kesamaan rata-rata cluster pada dimensi 6:

$H_0 : \pi_1 = \pi_2 = \pi_3 = \pi_4 = \pi_5$ (tidak ada perbedaan antar cluster)

$H_1 : \text{minimal ada satu } \pi_i \text{ yang tidak sama; } i=1,2,\dots,5;$

(ada perbedaan antar cluster)

Statistik uji yang digunakan adalah:

$$F_{hitung} = \frac{MS_C}{MS_E}$$

Kriteria penolakan H_0 adalah jika $F_{hitung} \geq F_{(db_1, db_2, \alpha)}$; atau $p\text{-value} \leq \alpha$,

dimana $db_1 = (k-1) = (5-1) = 4$ dan $db_2 = (n-k) = (16990-5) = 16985$

Tabel 4.3 Tabel Uji Kesamaan Rata-Rata Cluster pada 6 Dimensi

	Cluster		Error		F_{hitung}	Sig.
	Mean Square	df	Mean Square	df		
Dimension 1	286,005	4	0,933	16985	306,563	,000
Dimension 2	1358,309	4	0,680	16985	1996,310	,000
Dimension 3	2662,091	4	0,373	16985	7129,963	,000
Dimension 4	1349,682	4	0,682	16985	1977,725	,000
Dimension 5	2023,840	4	0,524	16985	3864,677	,000
Dimension 6	1486,392	4	0,650	16985	2285,891	,000

Pada Tabel 4.3 dapat dilihat hasil pengujian untuk keenam hipotesa di atas, yang menunjukkan bahwa dengan tingkat signifikansi 5 %, ternyata secara statistik H_0 ditolak (H_1 diterima), yang berarti ada perbedaan yang signifikan diantara kelima cluster (segmen) yang dibentuk. Jadi pengelompokkan rumah tangga sampel menjadi 5 (lima) segmen tersebut secara statistik sudah signifikan, dimana rumah tangga yang berada dalam satu segmen yang sama diduga mempunyai kemiripan dalam hal karakteristik, sehingga dapat membedakan dari segmen yang lain.

Dan dari hasil pengelompokan menjadi 5 segmen tersebut didapatkan :

- 1) **Segmen pertama**, terdiri dari 1498 rumah tangga;
- 2) **Segmen kedua**, terdiri dari 10049 rumah tangga;
- 3) **Segmen ketiga**, terdiri dari 3585 rumah tangga;
- 4) **Segmen keempat**, terdiri dari 1711 rumah tangga
- 5) **Segmen kelima**, terdiri dari 147 rumah tangga;

Berdasarkan hasil segmentasi tersebut dilanjutkan dengan melihat hubungan (*dependensi*) antara segmen yang terbentuk (5 segmen) dengan variabel wilayah. Untuk mengetahui ada tidaknya hubungan antara 2 (dua) variabel kategorikal (nominal) digunakan Analisis Tabulasi Silang (*Cross Tabulation*) serta dilakukan uji independensi (*Chi-Square*) untuk melihat karakteristik tiap segmen yang signifikan.

Adapun hipotesa yang digunakan dalam pengujian independensi ini adalah:

1. Uji independensi antara segmen dan variabel tertentu:

$$H_0 : \pi_{ij} = \pi_{i+} \cdot \pi_{+j}; i,j=1,2,\dots,5;$$

(tidak ada hubungan antara segmen dan variabel tertentu)

$$H_1 : \pi_{ij} \neq \pi_{i+} \cdot \pi_{+j}; i,j=1,2,\dots,5;$$

(ada hubungan antara segmen dan variabel tertentu)

Statistik uji yang digunakan adalah:

$$\chi_{hitung}^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(x_{ij} - \hat{m}_{ij})^2}{\hat{m}_{ij}}; \text{ dimana } \hat{m}_{ij} = \frac{X_{i+} \cdot X_{+j}}{n}$$

Kriteria penolakan H_0 adalah jika $\chi_{hitung}^2 \geq \chi_{(db, \alpha)}^2$ atau $p\text{-value} \leq \alpha$.

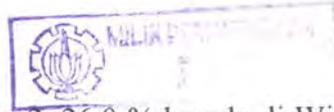
Tabel 4.4 Uji *Chi-Square* antara Segmen dengan Variabel kaarakteristik

No	Karakteristik	χ^2_{hitung}	db	Sig
1	Wilayah	1140.943	16	0.000
2	Pembelian pakaian	9298.918	8	0.000
3	Fasilitas air Bersih	1124.605	8	0.000
4	Kepemilikan rumah	1322.534	8	0.000
5	Jenis dinding terluas	2480.026	8	0.000
6	Jenis lantai terluas	8278.708	8	0.000
7	Sarana Tempat Buang Air Besar	908.047	8	0.000
8	Sumber Penerangan	11927.028	8	0.000
9	ART umur 6 – 15 tahun	3247.020	8	0.000
10	Sumber Keuangan Rumah tangga	12275.011	8	0.000
11	Pelayanan Kesehatan	8662.341	8	0.000

Tabel 4.4 menampilkan hasil pengujian untuk hipotesa di atas, yang menunjukkan bahwa dengan tingkat signifikansi 5 %, ternyata secara statistik H_0 ditolak (H_1 diterima). Ini berarti ada hubungan yang signifikan antara segmen dengan variabel variabel tersebut.

Setelah dilakukan tabulasi silang didapat karakteristik tiap segmen yaitu :

- Dari 1498 rumah tangga yang berada dalam segmen 1, 33.9% berada di Wilayah Surabaya Pusat, dengan karakteristik 98,2 % membeli pakaian max 1 kali dalam setahun, 74,7 % Fasilitas air bersih secara bersama, status kepemilikan rumah 39.1 % milik sendiri, Jenis dinding terluas 50.3 % terbuat dari tembok, jenis lantai terluas 74,1 % terbuat dari semen/plester/pasangan bata/papan, sarana tempat buang air besar 44.1 % secara bersama, sumber penerangan 65.8 % menggunakan listrik bersama/petromak, anggota rumah tangga yang berumur 6 – 15 tahun 67.2 % semua sekolah, sumber keuangan 58.9 % tidak ada, pelayanan kesehatan 52.8 % memilih lainnya.



- Dari 10049 rumah tangga yang berada dalam segmen 2, 26.0 % berada di Wilayah Surabaya Selatan, dengan karakteristik 99,9 % membeli pakaian max 1 kali dalam setahun, 74,7 % Fasilitas air bersih 76,3 % menggunakan fasilitas umum, status kepemilikan rumah 48.4 % sewa / kontrak, Jenis dinding terluas 42.8 % terbuat dari papan kayu, jenis lantai terluas 90.4 % terbuat dari semen/ plester/pasangan bata/papan, sarana tempat buang air besar 38.8 % secara bersama, sumber penerangan 68.9 % menggunakan listrik bersama/petromak, anggota rumah tangga yang berumur 6 – 15 tahun 82.2 % semua sekolah, sumber keuangan 99.2 % tidak tetap, pelayanan kesehatan semuanya memilih RSU / Puskesmas / Bidan atau mantri.
- Dari 3585 rumah tangga yang berada dalam segmen 3, sebagian besar berada di Wilayah Surabaya Timur (31.0%), dengan karakteristik 97.3 % membeli pakaian max 1 kali dalam setahun, 79.3 % menggunakan Fasilitas air bersih umum, status kepemilikan rumah 40.9 % milik sendiri, Jenis dinding terluas 84.9 % terbuat dari tembok, jenis lantai terluas 66.4 % terbuat dari ubin, sarana tempat buang air besar 46.8 % secara bersama, sumber penerangan 53.4 % menggunakan listrik bersama/petromak, anggota rumah tangga yang berumur 6 – 15 tahun 84.9% semua sekolah, sumber keuangan 69.5 % tidak tetap, pelayanan kesehatan 93.6 % memilih di RSU/ puskesmas/bidan atau mantri.
- Dari 1711 rumah tangga yang berada dalam segmen 4, 25.7% berada di Wilayah Surabaya Utara, dengan karakteristik 52.3 % membeli pakaian 1 - 2 kali dalam setahun, 48.3 % Fasilitas air bersih secara bersama, status kepemilikan rumah 56.6 % kontrak / sewa, Jenis dinding terluas 53 % terbuat dari tembok, jenis lantai terluas 83.1 % terbuat dari semen/plester/pasangan bata/papan, sarana tempat buang

air besar 65.5 % secara bersama, sumber penerangan 77.9 % menggunakan listrik bersama/petromak, anggota rumah tangga yang berumur 6 – 15 tahun 40.6 % semua sekolah, sumber keuangan 92.9 % tidak tetap, pelayanan kesehatan 97.4 % di RSUD / puskesmas/bidan atau mantri.

- Dari 147 rumah tangga yang berada dalam segmen 5, 29.3% berada di Wilayah Surabaya Timur, dengan karakteristik 84.4 % membeli pakaian max 1 kali dalam setahun, 95.9 % menggunakan Fasilitas air bersih umum , status kepemilikan rumah 40.1 % bebas sewa, Jenis dinding terluas 57.1 % terbuat dari bambu, jenis lantai terluas 62.6 % terbuat dari tanah/pasir, sarana tempat buang air besar 70.7 % menggunakan fasilitas umum, sumber penerangan 68.0 % menggunakan pelita, anggota rumah tangga yang berumur 6 – 15 tahun 44.2 % tidak ada yang sekolah, sumber keuangan 85.7 % tidak tetap, pelayanan kesehatan 82.3 % di RSUD / puskesmas / bidan atau mantri.

4.3 Analisis Diskriminan

Dari 11 variabel yang digunakan, ingin diketahui variabel apa yang secara bersama dapat membedakan kelompok satu dengan lainnya. Dengan menggunakan prosedur bertatar (*stepwise*) pemilihan variabel diseleksi berdasarkan nilai Statistik F dan Wilks Lambda (Λ).

Dari 11 variabel yang ada ternyata 10 variabel yang berperan dalam pembentukan kelompok. Variabel-variabel tersebut adalah :

1. Pembelian Pakaian dalam setahun (X_1)
2. Fasilitas Air bersih (X_2)
3. Kepemilikan rumah (X_4)

4. Jenis dinding (X_5)
5. Jenis Lantai (X_6)
6. Sarana Tempat Buang Air Besar – TBAB (X_7)
7. Sumber Penerangan (X_8)
8. Pendidikan ART umur 6 – 15 tahun (X_9)
9. Sumber Keuangan Rumah tangga (X_{10})
10. Pelayanan Kesehatan (X_{11})

Sedangkan fungsi diskriminan yang terbentuk adalah :

- a. Fungsi Diskriminan 1

$$D = -0,157 X_1 - 0,139 X_2 - 0,123 X_4 + 0,101 X_5 + 0,519 X_6 + 0,042 X_7 + 0,141 X_8 + 0,155 X_9 + 0,865 X_{10} + 0,737 X_{11}$$

- b. Fungsi Diskriminan 2

$$D = 0,944 X_1 + 0,154 X_2 + 0,046 X_4 + 0,041 X_5 + 0,004 X_6 + 0,117 X_7 - 0,130 X_8 - 0,572 X_9 + 0,166 X_{10} + 0,201 X_{11}$$

- c. Fungsi Diskriminan 3

$$D = 0,069 X_1 - 0,102 X_2 - 0,297 X_4 + 0,210 X_5 + 0,703 X_6 + 0,063 X_7 + 0,273 X_8 - 0,077 X_9 - 0,073 X_{10} - 0,501 X_{11}$$

- d. Fungsi Diskriminan 4

$$D = 0,173 X_1 + 0,291 X_2 + 0,102 X_4 + 0,227 X_5 - 0,162 X_6 - 0,191 X_7 + 0,745 X_8 + 0,360 X_9 - 0,150 X_{10} + 0,065 X_{11}$$

Untuk mengetahui apakah fungsi diskriminan yang disusun yang mengkaji hubungan antar kelompok layak atau tidak maka dilakukan pengujian terhadap vektor rata – rata. Statistik uji yang digunakan adalah V-Barlett.

Hipotesisnya:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k$$

H_1 : sedikitnya ada 2 kelompok yang berbeda rata – ratanya

Dari output didapat nilai Wilks Lambda = 0,078

Maka nilai V-Barlett sebesar :

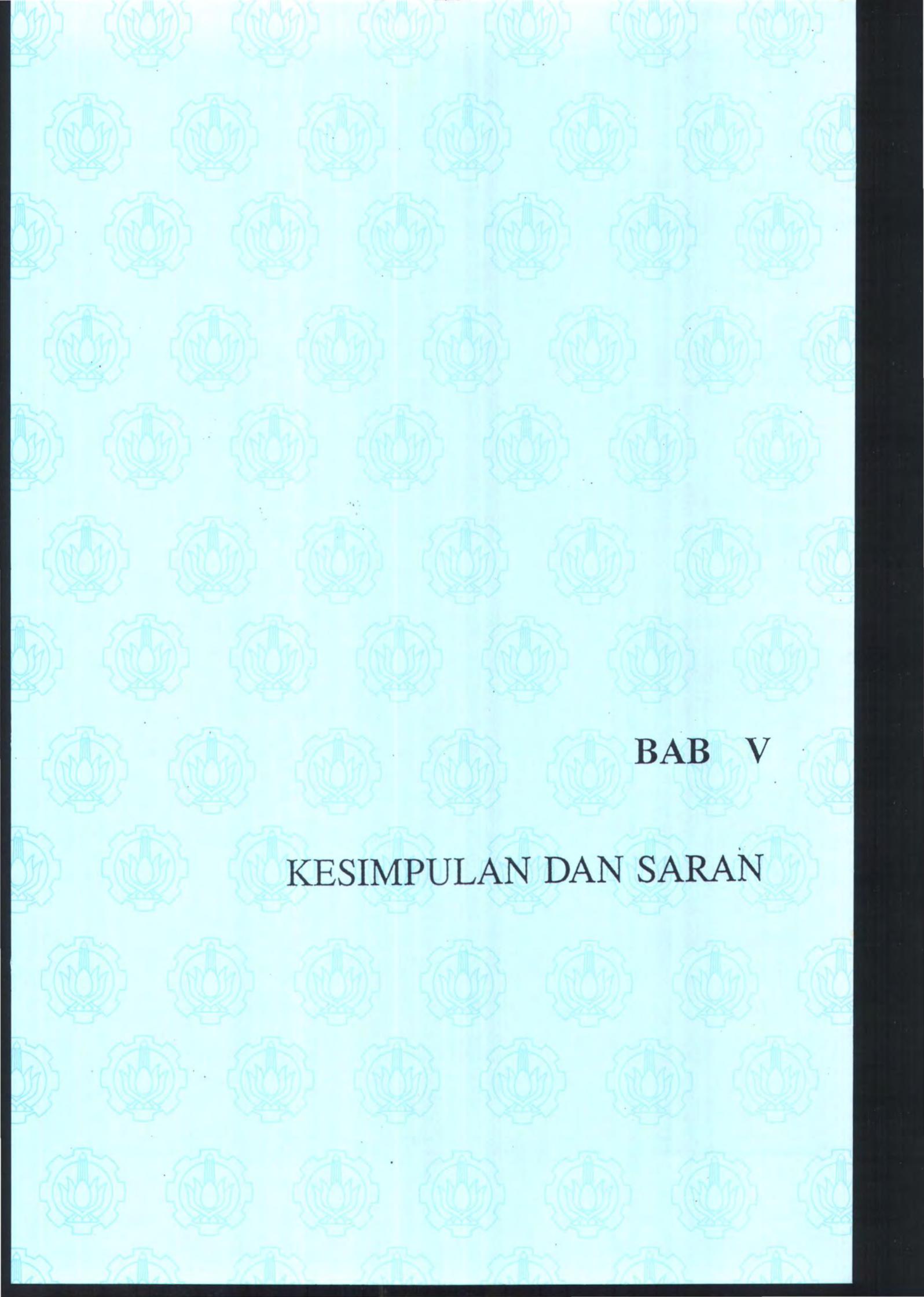
$$\text{V-Barlett} = 43392.173$$

$$\text{Statistik tabel } \chi^2_{p(k-1)\alpha} = \chi^2_{40(0,05)} = 124.3$$

Kesimpulan :

Karena statistik hitung V-Barlett lebih besar dari statistik tabel maka hipotesa awal (H_0) ditolak. Artinya berdasarkan data yang ada sedikitnya ada dua kelompok yang berbeda rata-ratanya, sehingga dengan adanya perbedaan rata-rata kedua kelompok tersebut maka fungsi diskriminan layak digunakan.

Sedangkan tingkat ketepatan dalam pengklasifikasian sebesar 91,9 %. Dari keempat fungsi diskriminan dihasilkan skor diskriminan untuk masing-masing grup. Skor diskriminan ini sangat berguna untuk mengetahui jika ada suatu obyek baru yang belum diketahui kelompoknya akan dimasukkan kedalam jajaran salah satu kelompok tersebut dengan syarat menggunakan variabel yang menjadi dasar pembentukan fungsi tersebut.



BAB V

KESIMPULAN DAN SARAN

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis data dan pembahasan, serta dengan melihat tujuan dari tugas akhir ini, maka dapat diambil kesimpulan sebagai berikut:

1. Hasil pengolahan terhadap 10 variabel rumah tangga miskin di Kota Surabaya tahun 2001 diperoleh hasil 6 dimensi yang mampu menerangkan sekitar 88,79 % dari total keragaman. Dari nilai *discrimination measure* ini dapat diketahui variabel-variabel yang mendukung dimensi tertentu yaitu :

- ☞ **Dimensi pertama** didukung oleh variabel Kepemilikan rumah (Var_4), Sarana Tempat Buang Air Besar (Var_7), Sumber penerangan (Var_8)
- ☞ **Dimensi kedua** didukung oleh variabel Jenis Dinding (Var_5) dan variabel Jenis Lantai (Var_6)
- ☞ **Dimensi ketiga** didukung oleh variabel Sumber keuangan rumah tangga (Var_10) dan variabel Pelayanan kesehatan (Var_11)
- ☞ **Dimensi keempat** tidak ada variabel yang mendukung. Hal ini mengindikasikan tidak ada variabel yang *dominan* mendukung dimensi keempat, tetapi tetap ada kontribusi dari tiap-tiap variabel pada dimensi ketiga.
- ☞ **Dimensi kelima** didukung oleh variabel Pembelian pakaian dalam setahun (Var_1) dan variabel Fasilitas Air bersih (Var_2)

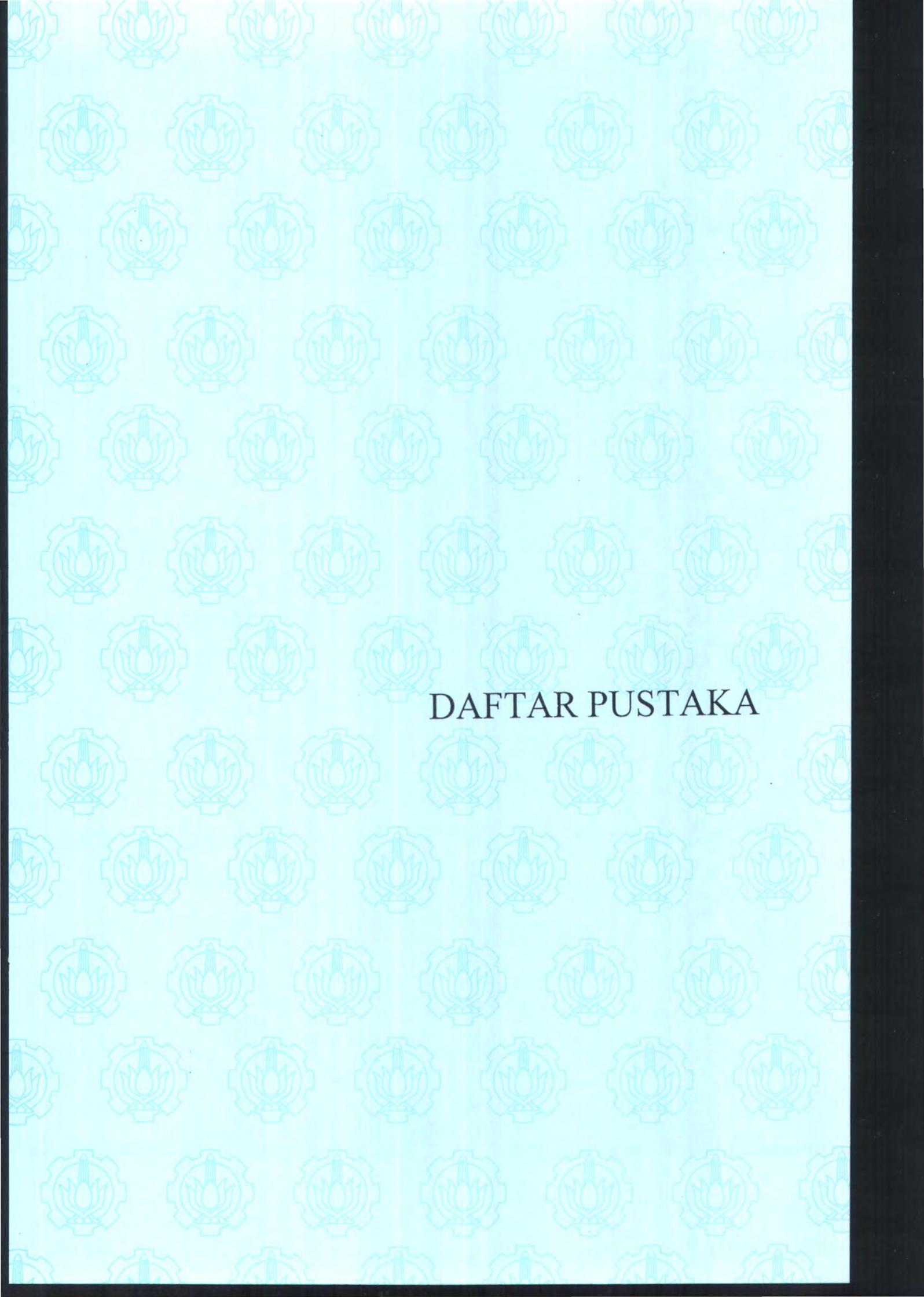
☞ **Dimensi keenam** didukung oleh variabel Pendidikan ART berumur 6 – 15 tahun (Var_9)

Dari hasil pengelompokan menjadi 5 segmen tersebut didapatkan :

- 1) **Segmen pertama**, terdiri dari 1498 rumah tangga;
 - 2) **Segmen kedua**, terdiri dari 10049 rumah tangga;
 - 3) **Segmen ketiga**, terdiri dari 3585 rumah tangga;
 - 4) **Segmen keempat**, terdiri dari 1711 rumah tangga
 - 5) **Segmen kelima**, terdiri dari 147 rumah tangga;
2. Dari hasil analisis diskriminan diperoleh 10 variabel pembeda yaitu Pembelian Pakaian dalam setahun (X_1), Fasilitas Air bersih (X_2), Kepemilikan rumah (X_4), Jenis dinding (X_5), Jenis Lantai (X_6), Sarana Tempat Buang Air Besar – TBAB (X_7), Sumber Penerangan (X_8), Pendidikan ART umur 6 – 15 tahun (X_9), Sumber Keuangan Rumah tangga (X_{10}), Pelayanan Kesehatan (X_{11}).

5.2 Saran

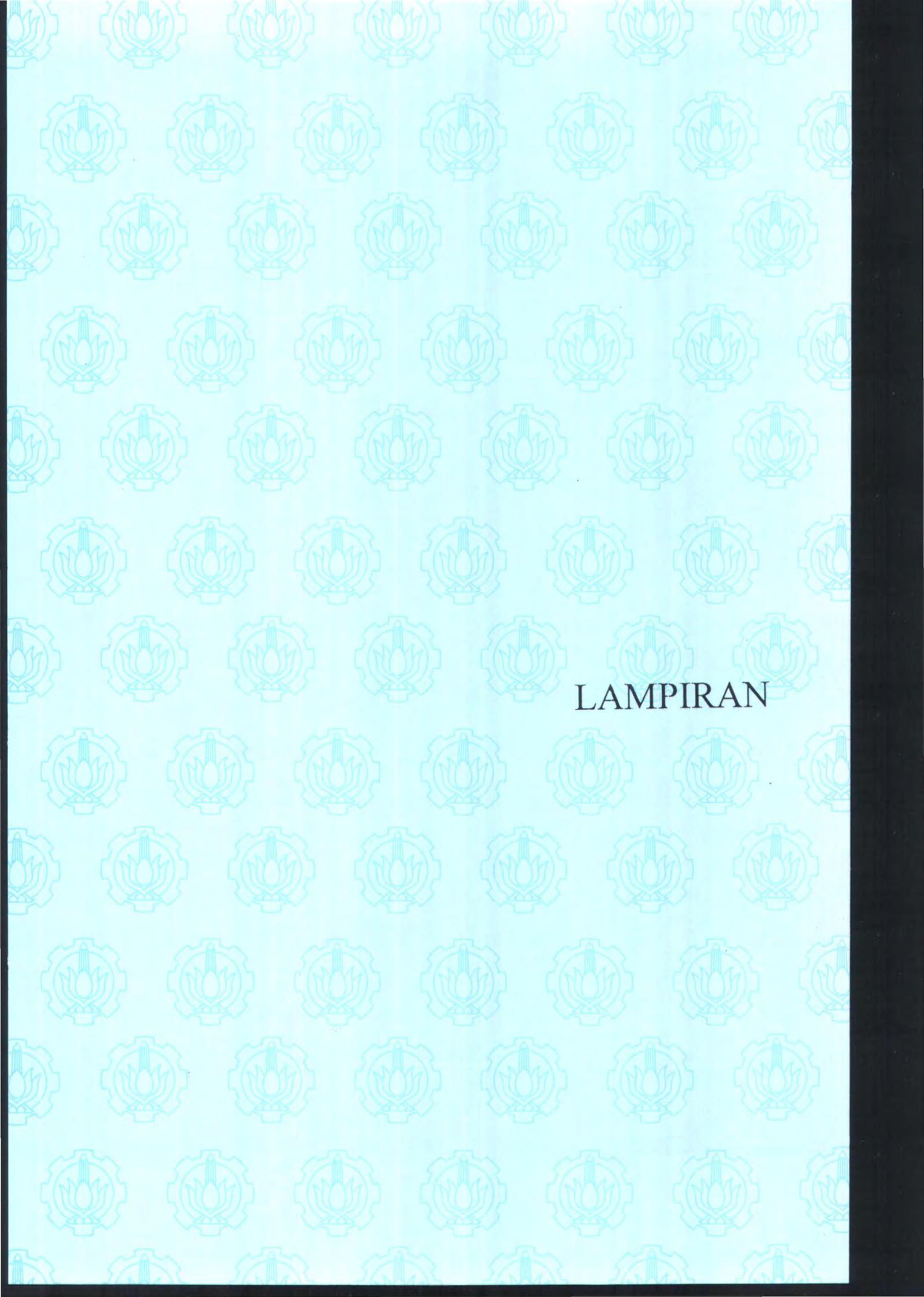
Perlu dilakukan penambahan variabel – variabel yang berhubungan dengan identitas responden (kepala rumah tangga) sehingga dapat juga diketahui karakteristik kepala rumah tangga tiap segmen. Misalnya tingkat pendidikan tertinggi, jenis kelamin kepala rumah tangga, Status perkawinan kepala rumah tangga.



DAFTAR PUSTAKA

DAFTAR PUSTAKA

- Agresti, A. (1996), *An Introduction to Categorical Data Analysis*, John Wiley & Sons. Inc, New York.
- Anderson, T.W. (1984), *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons. Inc, New York, Second Edition.
- BPS, Pendataan Kemiskinan dengan Indikator Baru di Jawa Timur, Surabaya 2001.
- Desposato, S.W., *Estimating Legislators Ideal Point Using HOMALS : Applied to Party-Switching in Brazilian Chamber of Deputies, 49th Session, 1991-1995*, Slide Persentation, University of California at Los Angeles.
- Gifi, Albert. (1990), *Nonlinear Multivariate Analysis*, Chichester: Wiley.
- Gringhuis, G and Israel, A. (1999). *Analyzing Well-Being In Relation To Characteristics Of The Population*, preprint, Statistics Nederlands.
- Johnson, R.A. dan Wichern, D.W. (1992), *Applied Multivariate Statistical Analysis*, Prentice Hall, New Jersey, Third Edition.
- Karson, Marvin J, *Multivariate Statistical Method, University of Albama, U.S.A*, 1982
- Michailidis, G and Leeuw, J.D., *Multilevel Homogeneity Analysis*, preprint, University of California at Los Angeles.



LAMPIRAN

LAMPIRAN A

The number of observations used in the analysis = 16990

List of Variables

=====

Variable	Variable label	Number of Categories
VAR_1	Pembelian pakaian dalam setahun	3
VAR_2	Fasilitas Air Minum	3
VAR_4	Kepemilikan rumah	3
VAR_5	Jenis dinding	3
VAR_6	Jenis Lantai	3
VAR_7	Sarana Tempat Buang Air Besar	3
VAR_8	Sumber Penerangan	3
VAR_9	Pendidikan ART umur 6 - 15 tahun	3
VAR_10	Sumber Keuangan Rumah Tangga	3
VAR_11	Pelayanan Kesehatan	3

Marginal Frequencies

=====

Variable	Missing	Categories		
		1	2	3
VAR_1	0	234	1029	15727
VAR_2	0	1136	3294	12560
VAR_4	0	6596	7269	3125
VAR_5	0	8849	6127	2014
VAR_6	0	3063	12824	1103
VAR_7	0	4506	7379	5105
VAR_8	0	5697	11193	100
VAR_9	0	13055	2765	1170
VAR_10	0	1286	14761	943
VAR_11	0	237	15898	855

The History of Iterations

=====

Number of Iterations	Total Fit	Difference between two Consecutive Iterations
1	,0007119	,0007119
2	,7095074	,7087956
3	,8029677	,0934603
4	,8404153	,0374476
5	,8547868	,0143715
6	,8616962	,0069094
7	,8656489	,0039528
8	,8682383	,0025894
9	,8701483	,0019100
10	,8717102	,0015618
11	,8730959	,0013857
12	,8743967	,0013007
13	,8756570	,0012604
14	,8768923	,0012353
15	,8780986	,0012063
16	,8792606	,0011621
17	,8803588	,0010982
18	,8813745	,0010156
19	,8822937	,0009192
20	,8831094	,0008156
21	,8838208	,0007114
22	,8844327	,0006119
23	,8849535	,0005208
24	,8853934	,0004399
25	,8857633	,0003699
26	,8860736	,0003103
27	,8863339	,0002603
28	,8865525	,0002186
29	,8867367	,0001841
30	,8868922	,0001556
31	,8870242	,0001320
32	,8871367	,0001125
33	,8872330	,0000963
34	,8873158	,0000828
35	,8873873	,0000715
36	,8874493	,0000620
37	,8875033	,0000540
38	,8875505	,0000472
39	,8875919	,0000414
40	,8876284	,0000364
41	,8876605	,0000322
42	,8876890	,0000285
43	,8877143	,0000253
44	,8877368	,0000225

45	,8877568	,0000200
46	,8877747	,0000179
47	,8877908	,0000160
48	,8878051	,0000144
49	,8878180	,0000129
50	,8878297	,0000116
51	,8878401	,0000105
52	,8878496	,0000094

The iterative process stops because the convergence has been reached in 52 iteration(s).

Dimension	Eigenvalue
1	,2445
2	,1647
3	,1342
4	,1209
5	,1163
6	,1073

=====
Discrimination measures per variable per dimension
 =====

Variable	Dimension					
	1	2	3	4	5	6
VAR_1	,020	,046	,018	,020	,318	,180
VAR_2	,260	,165	,023	,078	,339	,053
VAR_3	,556	,012	,280	,037	,028	,074
VAR_4	,145	,493	,035	,281	,036	,026
VAR_5	,107	,435	,144	,188	,124	,037
VAR_6	,631	,349	,007	,017	,081	,014
VAR_7	,692	,085	,093	,030	,008	,103
VAR_8	,019	,002	,127	,198	,043	,226
VAR_9	,006	,049	,275	,224	,105	,068
VAR_10	,009	,010	,341	,138	,080	,292

Marginal Frequencies and Category Quantifications

Variable: VAR_1 Pembelian pakaian dalam setahun

Category	Marginal Frequency
1	234
2	1029
3	15727

Missing: 0

Category Quantifications

Category	Dimensions					
	1	2	3	4	5	6
1	-,62	,01	-,65	,69	-2,64	2,98
2	-,47	-,85	-,42	-,46	-1,81	,89
3	,04	,06	,04	,02	,16	-,10

Variable: VAR_2 Fasilitas Air Minum

Category	Marginal Frequency
1	1136
2	3294
3	12560

Missing: 0

Category Quantifications

Category	Dimensions					
	1	2	3	4	5	6
1	1,68	-,27	,12	,63	-1,08	-,41
2	-,61	-,79	-,31	-,50	-,94	-,38
3	,01	,23	,07	,07	,35	,14

Variable: VAR_4

Kepemilikan rumah

Category	Marginal Frequency
1	6596
2	7269
3	3125

Missing: 0

Category Quantifications

Category	Dimensions					
	1	2	3	4	5	6
1	,91	,11	,11	,11	-,15	,07
2	-,72	-,02	,37	,07	-,01	,18
3	-,24	-,20	-1,09	-,40	,33	-,56

Variable: VAR_5

Jenis dinding

Category	Marginal Frequency
1	8849
2	6127
3	2014

Missing: 0

Category Quantifications

Category	Dimensions					
	1	2	3	4	5	6
1	,29	-,56	-,12	-,13	,15	,05
2	-,51	,29	,24	,57	-,07	,08
3	,25	1,58	-,23	-1,14	-,43	-,44

Variable: VAR_6

Jenis Lantai

Category	Marginal Frequency
1	3063
2	12824
3	1103

Missing: 0

Category Quantifications

Category	Dimensions					
	1	2	3	4	5	6
1	,67	-,85	-,52	-,40	,64	,39
2	-,18	,02	,21	,21	-,08	-,11
3	,23	2,17	-,98	-1,38	-,84	,16

Variable: VAR_7

Sarana Tempat Buang Air Besar

Category	Marginal Frequency
1	4506
2	7379
3	5105

Missing: 0

Category Quantifications

Category	Dimensions					
	1	2	3	4	5	6
1	1,30	-,11	,10	,13	-,25	,06
2	-,62	-,53	-,09	-,15	-,14	-,13
3	-,25	,86	,04	,09	,43	,14

Variable: VAR_8

Sumber Penerangan

Category	Marginal Frequency
1	5697
2	11193
3	100

Missing: 0

Category Quantifications

Category	Dimensions					
	1	2	3	4	5	6
1	1,17	-,04	,03	,10	-,07	,06
2	-,59	-,01	,02	-,03	,04	-,06
3	-,34	3,78	-3,95	-2,08	-,94	4,11

Variable: VAR_9

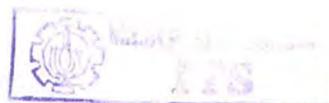
Pendidikan ART umur 6 - 15 tahun

Category	Marginal Frequency
1	13055
2	2765
3	1170

Missing: 0

Category Quantifications

Category	Dimensions					
	1	2	3	4	5	6
1	,07	-,02	,12	-,24	,11	-,24
2	-,16	,09	-,01	,89	-,25	,52
3	-,40	-,03	-1,30	,59	-,60	1,42



Variable: VAR_10

Sumber Keuangan Rumah Tangga

Category	Marginal Frequency
1	1286
2	14761
3	943

Missing: 0

Category Quantifications

Category	Dimensions					
	1	2	3	4	5	6
1	-,14	-,77	-,22	-1,08	1,08	,65
2	-,01	,06	,16	-,01	-,11	-,01
3	,28	,13	-2,12	1,56	,33	-,81

Variable: VAR_11

Pelayanan Kesehatan

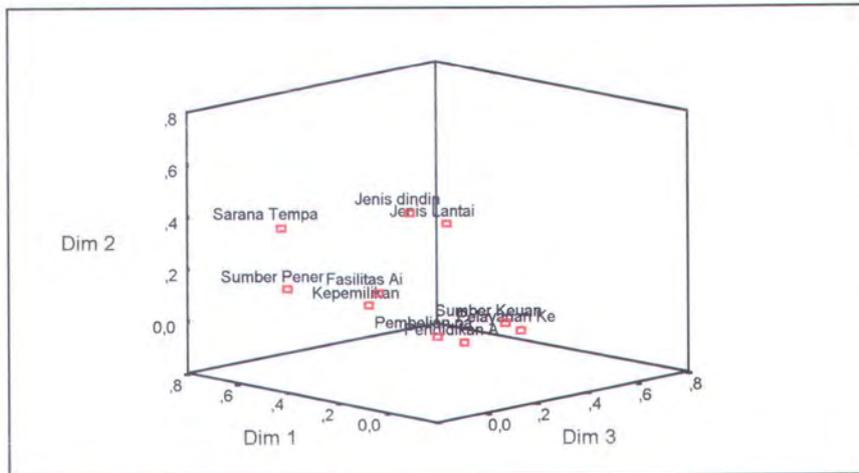
Category	Marginal Frequency
1	237
2	15898
3	855

Missing: 0

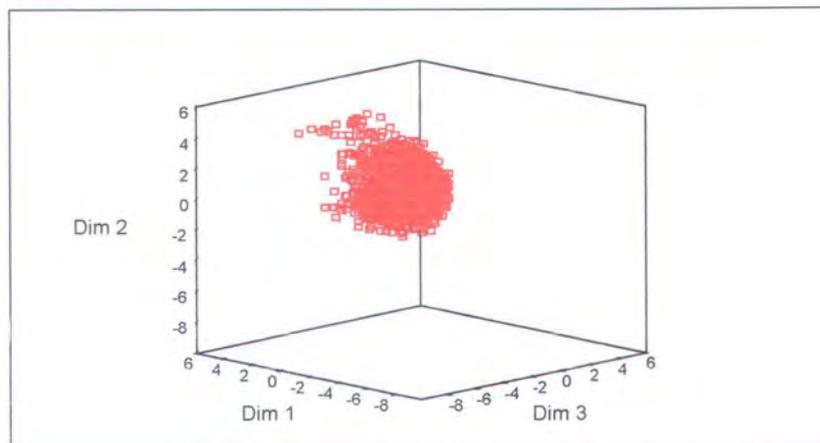
Category Quantifications

Category	Dimensions					
	1	2	3	4	5	6
1	,12	-,72	-,09	-,93	2,22	4,35
2	,02	,00	,14	-,07	-,06	-,03
3	-,41	,23	-2,53	1,55	,42	-,73

Discrimination Measures

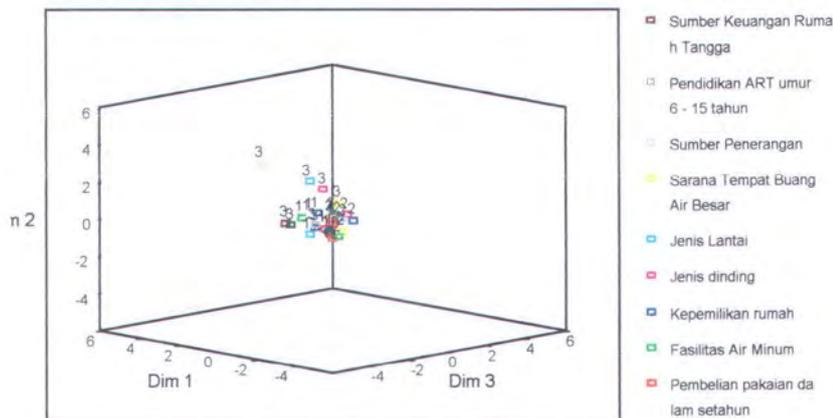


Object Scores



Cases weighted by number of objects.

Category Quantifications



LAMPIRAN B :

ANALISIS CLUSTER

Initial Cluster Centers

	Cluster				
	1	2	3	4	5
Dimension 1	-,06	,76	1,18	-1,06	-,34
Dimension 2	-,07	,11	-1,03	4,47	5,36
Dimension 3	-4,54	,32	-,40	-4,37	-8,99
Dimension 4	1,72	2,59	-1,97	-1,20	-,92
Dimension 5	2,53	-4,00	4,22	-3,86	-,70
Dimension 6	-1,67	2,79	5,13	8,49	3,12

Iteration History

Iteration	Change in Cluster Centers				
	1	2	3	4	5
1	5,015	5,078	4,574	4,844	4,244
2	,370	,192	,309	1,468	,373
3	,178	6,161E-02	,327	1,500	,150
4	,388	9,642E-02	,637	2,023	,534
5	,197	9,434E-02	,209	,963	,909
6	,192	9,395E-02	,143	,742	,309
7	,125	2,828E-02	6,761E-02	,213	,373
8	,187	3,698E-02	8,947E-02	,127	,254
9	,269	5,942E-02	3,280E-02	,146	8,673E-02
10	,410	4,492E-02	9,530E-02	,117	4,556E-02
11	,380	3,972E-02	3,842E-02	7,099E-02	9,679E-02
12	7,123E-02	2,300E-03	1,706E-02	2,838E-02	6,958E-02
13	7,807E-03	8,743E-04	1,879E-03	4,394E-03	2,727E-02
14	3,539E-03	2,356E-03	6,598E-03	3,086E-03	,000
15	,000	,000	,000	,000	,000

a Convergence achieved due to no or small distance change. The maximum distance by which any center has changed is ,000. The current iteration is 15. The minimum distance between initial centers is 7,846.

Final Cluster Centers

	Cluster				
	1	2	3	4	5
Dimension 1	-0,01	-0,05	0,42	-0,53	-0,29
Dimension 2	0,13	0,31	-0,78	-0,61	3,32
Dimension 3	-2,04	0,49	-0,24	-0,32	-3,26
Dimension 4	1,45	0,05	-0,70	0,07	-1,81
Dimension 5	0,45	-0,07	0,86	-1,70	-0,99
Dimension 6	-0,85	-0,25	0,41	1,06	3,39

Distances between Final Cluster Centers

Cluster	1	2	3	4	5
1		3,011	3,257	3,737	6,517
2	3,011		1,949	2,479	6,384
3	3,257	1,949		2,915	6,320
4	3,737	2,479	2,915		5,798
5	6,517	6,384	6,320	5,798	

ANOVA

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Dimension 1	286,005	4	,933	16985	306,563	,000
Dimension 2	1358,309	4	,680	16985	1996,310	,000
Dimension 3	2662,091	4	,373	16985	7129,963	,000
Dimension 4	1349,682	4	,682	16985	1977,725	,000
Dimension 5	2023,840	4	,524	16985	3864,677	,000
Dimension 6	1486,392	4	,650	16985	2285,891	,000

Number of Cases in each Cluster

Cluster	1	1498
	2	10049
	3	3585
	4	1711
	5	147
Valid		16990
Missing		0

LAMPIRAN C

Crosstab Segmen * Wilayah

		WILAYAH					Total
		Barat	Pusat	Selatan	Timur	Utara	
Segmen 1	Expected Count	240,5	190,4	345,6	355,3	366,2	1498
	% within Segmen	10,1%	33,9%	17,9%	27,8%	10,2%	100%
	% within WILAYAH	5,6%	23,5%	6,8%	10,3%	3,7%	8,8%
	% of Total	,9%	3,0%	1,6%	2,5%	,9%	8,8%
Segmen 2	Expected Count	1613,5	1277,0	2318,5	2383,6	2456,4	10049
	% within Segmen	17,0%	11,2%	26,0%	20,2%	25,6%	100%
	% within WILAYAH	62,5%	52,1%	66,8%	50,5%	61,9%	59,1%
	% of Total	10,0%	6,6%	15,4%	12,0%	15,1%	59,1%
Segmen 3	Expected Count	575,6	455,6	827,1	850,4	876,3	3585
	% within Segmen	12,2%	8,7%	21,0%	31,0%	27,1%	100%
	% within WILAYAH	16,1%	14,5%	19,2%	27,5%	23,4%	21,1%
	% of Total	2,6%	1,8%	4,4%	6,5%	5,7%	21,1%
Segmen 4	Expected Count	274,7	217,4	394,8	405,8	418,2	1711
	% within Segmen	23,4%	10,8%	15,3%	24,9%	25,7%	100%
	% within WILAYAH	14,7%	8,6%	6,7%	10,6%	10,6%	10,1%
	% of Total	2,4%	1,1%	1,5%	2,5%	2,6%	10,1%
Segmen 5	Expected Count	23,6	18,7	33,9	34,9	35,9	147
	% within Segmen	23,1%	19,7%	15,0%	29,3%	12,9%	100%
	% within WILAYAH	1,2%	1,3%	,6%	1,1%	,5%	,9%
	% of Total	,2%	,2%	,1%	,3%	,1%	,9%
Total	Expected Count	2728,0	2159,0	3920,0	4030,0	4153,0	16990
	% within Segmen	16,1%	12,7%	23,1%	23,7%	24,4%	100%
	% within WILAYAH	100,0%	100,0%	100,0%	100,0%	100,0%	100%
	% of Total	16,1%	12,7%	23,1%	23,7%	24,4%	100%

Chi-Square Tests

	Value	df	Sig
Pearson Chi-Square	1140,943	16	,000
Likelihood Ratio	1017,185	16	,000
N of Valid Cases	16990		

a 0 cells (,0%) have expected count less than 5. The minimum expected count is 18,68.

Crosstab Segmen * Pembelian Pakaian dalam setahun

		Pembelian Pakaian dalam setahun		Total
		1	2	
Segmen 1	Count	1471	27	1498
	Expected Count	1386,6	111,4	1498,0
	% within Segmen	98,2%	1,8%	100,0%
	% of Total	8,7%	,2%	8,8%
Segmen 2	Count	10043	6	10049
	Expected Count	9302,0	747,0	10049,0
	% within Segmen	99,9%	,1%	100,0%
	% of Total	59,1%	,0%	59,1%
Segmen 3	Count	3489	96	3585
	Expected Count	3318,5	266,5	3585,0
	% within Segmen	97,3%	2,7%	100,0%
	% of Total	20,5%	,6%	21,1%
Segmen 4	Count	600	1111	1711
	Expected Count	1583,8	127,2	1711,0
	% within Segmen	35,1%	64,9%	100,0%
	% of Total	3,5%	6,5%	10,1%
Segmen 5	Count	124	23	147
	Expected Count	136,1	10,9	147,0
	% within Segmen	84,4%	15,6%	100,0%
	% of Total	,7%	,1%	,9%
Total	Count	15727	1263	16990
	Expected Count	15727,0	1263,0	16990,0
	% within Segmen	92,6%	7,4%	100,0%
	% of Total	92,6%	7,4%	100,0%

Keterangan : 1 = Maks 1 kali
2 = Kapan saja / 2- 3 kali

Chi-Square Tests

	Value	df	Sig.
Pearson Chi-Square	9216,076	4	,000
Likelihood Ratio	5394,669	4	,000
Linear-by-Linear Association	4785,864	1	,000
N of Valid Cases	16990		

a 0 cells (,0%) have expected count less than 5. The minimum expected count is 10,93.

Crosstab Segmen * Fasilitas Air Bersih

		Fasilitas Air Minum			Total
		1	2	3	
Segmen 1	Expected Count	100,2	290,4	1107,4	1498,0
	% within Segmen	6,8%	18,5%	74,7%	100,0%
	% within Fas.Air Minum	9,0%	8,4%	8,9%	8,8%
	% of Total	,6%	1,6%	6,6%	8,8%
Segmen 2	Expected Count	671,9	1948,3	7428,8	10049,0
	% within Segmen	7,9%	15,8%	76,3%	100,0%
	% within Fas.Air Minum	70,2%	48,1%	61,1%	59,1%
	% of Total	4,7%	9,3%	45,1%	59,1%
Segmen 3	Expected Count	239,7	695,1	2650,2	3585,0
	% within Segmen	4,0%	16,8%	79,3%	100,0%
	% within Fas. Air Minum	12,5%	18,2%	22,6%	21,1%
	% of Total	,8%	3,5%	16,7%	21,1%
Segmen 4	Expected Count	114,4	331,7	1264,9	1711,0
	% within Segmen	5,5%	48,3%	46,2%	100,0%
	% within Fas. Air Minum	8,3%	25,1%	6,3%	10,1%
	% of Total	,6%	4,9%	4,6%	10,1%
Segmen 5	Expected Count	9,8	28,5	108,7	147,0
	% within Segmen	,7%	3,4%	95,9%	100,0%
	% within Fas. Air Minum	,1%	,2%	1,1%	,9%
	% of Total	,0%	,0%	,8%	,9%
Total	Expected Count	1136,0	3294,0	12560,0	16990,0
	% within Segmen	6,7%	19,4%	73,9%	100,0%
	% within Fas. Air Minum	100,0%	100,0%	100,0%	100,0%
	% of Total	6,7%	19,4%	73,9%	100,0%

Keterangan :
 1 = Sendiri
 2 = Bersama
 3 = Umum / Tidak ada

Chi-Square Tests

	Value	df	Sig
Pearson Chi-Square	1124,605	8	,000
Likelihood Ratio	958,549	8	,000
Linear-by-Linear Association	64,704	1	,000
N of Valid Cases	16990		

a 0 cells (,0%) have expected count less than 5. The minimum expected count is 9,83.

Crosstab Segmen * Kepemilikan Rumah

		Kepemilikan rumah			Total
		1	2	3	
Segmen 1	Expected Count	581,6	640,9	275,5	1498,0
	% within Segmen	39,1%	27,4%	33,5%	100,0%
	% within Kepemilikan rumah	8,9%	5,7%	16,1%	8,8%
	% of Total	3,4%	2,4%	3,0%	8,8%
Segmen 2	Expected Count	3901,3	4299,4	1848,3	10049,0
	% within Segmen	39,6%	48,4%	12,0%	100,0%
	% within Kepemilikan rumah	60,3%	66,9%	38,6%	59,1%
	% of Total	23,4%	28,6%	7,1%	59,1%
Segmen 3	Expected Count	1391,8	1533,8	659,4	3585,0
	% within Segmen	40,9%	27,1%	32,0%	100,0%
	% within Kepemilikan rumah	22,2%	13,4%	36,7%	21,1%
	% of Total	8,6%	5,7%	6,8%	21,1%
Segmen 4	Expected Count	664,3	732,0	314,7	1711,0
	% within Segmen	31,1%	56,6%	12,3%	100,0%
	% within Kepemilikan rumah	8,1%	13,3%	6,8%	10,1%
	% of Total	3,1%	5,7%	1,2%	10,1%
Segmen 5	Expected Count	57,1	62,9	27,0	147,0
	% within Segmen	23,8%	36,1%	40,1%	100,0%
	% within Kepemilikan rumah	,5%	,7%	1,9%	,9%
	% of Total	,2%	,3%	,3%	,9%
Total	Expected Count	6596,0	7269,0	3125,0	16990,0
	% within Segmen	38,8%	42,8%	18,4%	100,0%
	% within Kepemilikan rumah	100,0%	100,0%	100,0%	100,0%
	% of Total	38,8%	42,8%	18,4%	100,0%

Keterangan :
 1 = Milik Sendiri
 2 = Sewa / kontrak
 3 = Bebas sewa

Chi-Square Tests

	Value	df	Sig
Pearson Chi-Square	1322,534	8	,000
Likelihood Ratio	1282,778	8	,000
Linear-by-Linear Association	28,320	1	,000
N of Valid Cases	16990		

a 0 cells (,0%) have expected count less than 5. The minimum expected count is 27,04.

Crosstab Segmen * Jenis Dinding

		Jenis dinding			Total
		1	2	3	
Segmen 1	Expected Count	780,2	540,2	177,6	1498,0
	% within Segmen	50,3%	39,6%	10,1%	100,0%
	% within Jenis dinding	8,5%	9,7%	7,5%	8,8%
	% of Total	4,4%	3,5%	,9%	8,8%
Segmen 2	Expected Count	5233,9	3623,9	1191,2	10049,0
	% within Segmen	41,1%	42,8%	16,1%	100,0%
	% within Jenis dinding	46,7%	70,2%	80,2%	59,1%
	% of Total	24,3%	25,3%	9,5%	59,1%
Segmen 3	Expected Count	1867,2	1292,8	425,0	3585,0
	% within Segmen	84,9%	12,6%	2,6%	100,0%
	% within Jenis dinding	34,4%	7,4%	4,6%	21,1%
	% of Total	17,9%	2,7%	,5%	21,1%
Segmen 4	Expected Count	891,2	617,0	202,8	1711,0
	% within Segmen	53,0%	42,9%	4,1%	100,0%
	% within Jenis dinding	10,2%	12,0%	3,5%	10,1%
	% of Total	5,3%	4,3%	,4%	10,1%
Segmen 5	Expected Count	76,6	53,0	17,4	147,0
	% within Segmen	10,9%	32,0%	57,1%	100,0%
	% within Jenis dinding	,2%	,8%	4,2%	,9%
	% of Total	,1%	,3%	,5%	,9%
Total	Expected Count	8849,0	6127,0	2014,0	16990,0
	% within Segmen	52,1%	36,1%	11,9%	100,0%
	% within Jenis dinding	100,0%	100,0%	100,0%	100,0%
	% of Total	52,1%	36,1%	11,9%	100,0%

Keterangan :
 1 = Tembok
 2 = Papan kayu sederhana
 3 = Bambu / lainnya

Chi-Square Tests

	Value	df	Sig
Pearson Chi-Square	2480,026	8	,000
Likelihood Ratio	2593,084	8	,000
Linear-by-Linear Association	377,777	1	,000
N of Valid Cases	16990		

a 0 cells (,0%) have expected count less than 5. The minimum expected count is 17,43.

Crosstab Segmen * Jenis Lantai

		Jenis Lantai			Total
		1	2	3	
Segmen 1	Expected Count	270,1	1130,7	97,3	1498,0
	% within Segmen	19,4%	74,0%	6,5%	100,0%
	% within Jenis Lantai	9,5%	8,6%	8,9%	8,8%
	% of Total	1,7%	6,5%	,6%	8,8%
Segmen 2	Expected Count	1811,7	7585,0	652,4	10049,0
	% within Segmen	2,0%	90,4%	7,6%	100,0%
	% within Jenis Lantai	6,5%	70,8%	69,4%	59,1%
	% of Total	1,2%	53,5%	4,5%	59,1%
Segmen 3	Expected Count	646,3	2705,9	232,7	3585,0
	% within Segmen	66,4%	32,4%	1,2%	100,0%
	% within Jenis Lantai	77,8%	9,0%	3,9%	21,1%
	% of Total	14,0%	6,8%	,3%	21,1%
Segmen 4	Expected Count	308,5	1291,5	111,1	1711,0
	% within Segmen	10,8%	83,1%	6,1%	100,0%
	% within Jenis Lantai	6,0%	11,1%	9,4%	10,1%
	% of Total	1,1%	8,4%	,6%	10,1%
Segmen 5	Expected Count	26,5	111,0	9,5	147,0
	% within Segmen	4,1%	33,3%	62,6%	100,0%
	% within Jenis Lantai	,2%	,4%	8,3%	,9%
	% of Total	,0%	,3%	,5%	,9%
Total	Expected Count	3063,0	12824,0	1103,0	16990,0
	% within Segmen	18,0%	75,5%	6,5%	100,0%
	% within Jenis Lantai	100,0%	100,0%	100,0%	100,0%
	% of Total	18,0%	75,5%	6,5%	100,0%

Keterangan : 1 = Ubin
 2 = Plester/semen/pasangan bata/papan
 3 = Tanah / pasir

Chi-Square Tests

	Value	df	Sig
Pearson Chi-Square	8278,708	8	,000
Likelihood Ratio	7139,351	8	,000
Linear-by-Linear Association	718,037	1	,000
N of Valid Cases	16990		

a 0 cells (,0%) have expected count less than 5. The minimum expected count is 9,54.

Crosstab Segmen * STBAB

		Sarana Tempat Buang Air Besar (STBAB)			Total
		1	2	3	
Segmen 1	Expected Count	397,3	650,6	450,1	1498,0
	% within Segmen	23,3%	44,1%	32,6%	100,0%
	% within STBAB	7,7%	8,9%	9,6%	8,8%
	% of Total	2,1%	3,9%	2,9%	8,8%
Segmen 2	Expected Count	2665,1	4364,4	3019,4	10049,0
	% within Segmen	25,4%	38,8%	35,9%	100,0%
	% within STBAB	56,6%	52,8%	70,6%	59,1%
	% of Total	15,0%	22,9%	21,2%	59,1%
Segmen 3	Expected Count	950,8	1557,0	1077,2	3585,0
	% within Segmen	34,6%	46,8%	18,6%	100,0%
	% within STBAB	27,5%	22,7%	13,1%	21,1%
	% of Total	7,3%	9,9%	3,9%	21,1%
Segmen 4	Expected Count	453,8	743,1	514,1	1711,0
	% within Segmen	20,5%	65,5%	14,1%	100,0%
	% within STBAB	7,8%	15,2%	4,7%	10,1%
	% of Total	2,1%	6,6%	1,4%	10,1%
Segmen 5	Expected Count	39,0	63,8	44,2	147,0
	% within Segmen	10,2%	19,0%	70,7%	100,0%
	% within STBAB	,3%	,4%	2,0%	,9%
	% of Total	,1%	,2%	,6%	,9%
Total	Expected Count	4506,0	7379,0	5105,0	16990,0
	% within Segmen	26,5%	43,4%	30,0%	100,0%
	% within STBAB	100,0%	100,0%	100,0%	100,0%
	% of Total	26,5%	43,4%	30,0%	100,0%

Keterangan : 1 = Sendiri
2 = Bersama
3 = Umum / Tidak ada

Chi-Square Tests

	Value	df	Sig
Pearson Chi-Square	908,047	8	,000
Likelihood Ratio	919,620	8	,000
Linear-by-Linear Association	121,405	1	,000
N of Valid Cases	16990		

a 0 cells (,0%) have expected count less than 5. The minimum expected count is 38,99.

Crosstab Segmen * Sumber Penerangan

		Sumber Penerangan		Total
		1	2	
Segmen 1	Count	512	986	1498
	Expected Count	502,3	995,7	1498,0
	% within Segmen	34,2%	65,8%	100,0%
	% of Total	3,0%	5,8%	8,8%
Segmen 2	Count	3127	6922	10049
	Expected Count	3369,6	6679,4	10049,0
	% within Segmen	31,1%	68,9%	100,0%
	% of Total	18,4%	40,7%	59,1%
Segmen 3	Count	1669	1916	3585
	Expected Count	1202,1	2382,9	3585,0
	% within Segmen	46,6%	53,4%	100,0%
	% of Total	9,8%	11,3%	21,1%
Segmen 4	Count	378	1333	1711
	Expected Count	573,7	1137,3	1711,0
	% within Segmen	22,1%	77,9%	100,0%
	% of Total	2,2%	7,8%	10,1%
Segmen 5	Count	11	136	147
	Expected Count	49,3	97,7	147,0
	% within Segmen	7,5%	92,5%	100,0%
	% of Total	,1%	,8%	,9%
Total	Count	5697	11293	16990
	Expected Count	5697,0	11293,0	16990,0
	% within Segmen	33,5%	66,5%	100,0%
	% of Total	33,5%	66,5%	100,0%

Keterangan : 1 = Listrik Sendiri
2 = Listrik bersama / petromak / pelita

Chi-Square Tests

	Value	df	Sig.
Pearson Chi-Square	444,586	4	,000
Likelihood Ratio	451,433	4	,000
Linear-by-Linear Association	,970	1	,325
N of Valid Cases	16990		

a 0 cells (,0%) have expected count less than 5. The minimum expected count is 49,29.

Crosstab Segmen * Pendidikan ART umur 6 - 15 tahun

		Pendidikan ART umur 6 - 15 tahun			Total
		1	2	3	
Segmen 1	Expected Count	1151,1	243,8	103,2	1498,0
	% within Segmen	67,2%	18,6%	14,2%	100,0%
	% within Pendidikan	7,7%	10,1%	18,2%	8,8%
	% of Total	5,9%	1,6%	1,3%	8,8%
Segmen 2	Expected Count	7721,6	1635,4	692,0	10049,0
	% within Segmen	82,2%	16,4%	1,3%	100,0%
	% within Pendidikan	63,3%	59,7%	11,5%	59,1%
	% of Total	48,6%	9,7%	,8%	59,1%
Segmen 3	Expected Count	2754,7	583,4	246,9	3585,0
	% within Segmen	84,5%	10,5%	5,0%	100,0%
	% within Pendidikan	23,2%	13,6%	15,3%	21,1%
	% of Total	17,8%	2,2%	1,1%	21,1%
Segmen 4	Expected Count	1314,7	278,5	117,8	1711,0
	% within Segmen	40,6%	25,6%	33,8%	100,0%
	% within Pendidikan	5,3%	15,8%	49,4%	10,1%
	% of Total	4,1%	2,6%	3,4%	10,1%
Segmen 5	Expected Count	113,0	23,9	10,1	147,0
	% within Segmen	42,2%	13,6%	44,2%	100,0%
	% within Pendidikan	,5%	,7%	5,6%	,9%
	% of Total	,4%	,1%	,4%	,9%
Total	Expected Count	13055,0	2765,0	1170,0	16990,0
	% within Segmen	76,8%	16,3%	6,9%	100,0%
	% within Pendidikan	100,0%	100,0%	100,0%	100,0%
	% of Total	76,8%	16,3%	6,9%	100,0%

Keterangan :
 1 = Semua sekolah
 2 = Ada yang tidak sekolah
 3 = Tidak ada yang sekolah

Chi-Square Tests

	Value	df	Sig
Pearson Chi-Square	3247,020	8	,000
Likelihood Ratio	2462,590	8	,000
Linear-by-Linear Association	910,879	1	,000
N of Valid Cases	16990		

a 0 cells (,0%) have expected count less than 5. The minimum expected count is 10,12.

Crosstab Segmen * Sumber Keuangan Rumah Tangga

		Sumber Keuangan Rumah Tangga			Total
		1	2	3	
Segmen 1	Expected Count	113,4	1301,5	83,1	1498,0
	% within Segmen	2,4%	38,7%	58,9%	100,0%
	% within Sumber Keuangan RT	2,8%	3,9%	93,5%	8,8%
	% of Total	,2%	3,4%	5,2%	8,8%
Segmen 2	Expected Count	760,6	8730,6	557,8	10049,0
	% within Segmen	,8%	99,2%	,0%	100,0%
	% within Sumber Keuangan RT	5,9%	67,6%	,0%	59,1%
	% of Total	,4%	58,7%	,0%	59,1%
Segmen 3	Expected Count	271,4	3114,7	199,0	3585,0
	% within Segmen	30,2%	69,5%	,3%	100,0%
	% within Sumber Keuangan RT	84,2%	16,9%	1,1%	21,1%
	% of Total	6,4%	14,7%	,1%	21,1%
Segmen 4	Expected Count	129,5	1486,5	95,0	1711,0
	% within Segmen	4,9%	92,9%	2,2%	100,0%
	% within Sumber Keuangan RT	6,5%	10,8%	3,9%	10,1%
	% of Total	,5%	9,4%	,2%	10,1%
Segmen 5	Expected Count	11,1	127,7	8,2	147,0
	% within Segmen	4,8%	85,7%	9,5%	100,0%
	% within Sumber Keuangan RT	,5%	,9%	1,5%	,9%
	% of Total	,0%	,7%	,1%	,9%
Total	Expected Count	1286,0	14761,0	943,0	16990,0
	% within Segmen	7,6%	86,9%	5,6%	100,0%
	% within Sumber Keuangan RT	100,0%	100,0%	100,0%	100,0%
	% of Total	7,6%	86,9%	5,6%	100,0%

Keterangan :
 1 = Tetap
 2 = Tidak Tetap
 3 = Tidak Ada

Chi-Square Tests

	Value	df	Sig
Pearson Chi-Square	12275,011	8	,000
Likelihood Ratio	7352,536	8	,000
Linear-by-Linear Association	2399,145	1	,000
N of Valid Cases	16990		

a 0 cells (,0%) have expected count less than 5. The minimum expected count is 8,16.

Crosstab Segmen * Pelayanan Kesehatan

		Pelayanan Kesehatan		Total
		1	2	
Segmen 1	Count	791	707	1498
	Expected Count	75,4	1422,6	1498,0
	% within Segmen	52,8%	47,2%	100,0%
	% of Total	4,7%	4,2%	8,8%
Segmen 2	Count	0	10049	10049
	Expected Count	505,7	9543,3	10049,0
	% within Segmen	,0%	100,0%	100,0%
	% of Total	,0%	59,1%	59,1%
Segmen 3	Count	11	3574	3585
	Expected Count	180,4	3404,6	3585,0
	% within Segmen	,3%	99,7%	100,0%
	% of Total	,1%	21,0%	21,1%
Segmen 4	Count	35	1676	1711
	Expected Count	86,1	1624,9	1711,0
	% within Segmen	2,0%	98,0%	100,0%
	% of Total	,2%	9,9%	10,1%
Segmen 5	Count	18	129	147
	Expected Count	7,4	139,6	147,0
	% within Segmen	12,2%	87,8%	100,0%
	% of Total	,1%	,8%	,9%
Total	Count	855	16135	16990
	Expected Count	855,0	16135,0	16990,0
	% within Segmen	5,0%	95,0%	100,0%
	% of Total	5,0%	95,0%	100,0%

Keterangan :

1 = Lainnya

2 = Dokter Praktek, RSUD, Polindes, Puskesmas, Bidan, Mantri

Chi-Square Tests

	Value	df	Sig.
Pearson Chi-Square	7901,125	4	,000
Likelihood Ratio	4105,826	4	,000
Linear-by-Linear Association	1710,457	1	,000
N of Valid Cases	16990		
N of Valid Cases	16990		

a 0 cells (,0%) have expected count less than 5. The minimum expected count is 7,40.

LAMPIRAN D :**ANALISIS DISKRIMINAN****Variables Entered/Removed**

Step	Entered	Wilks' Lambda	Sig.
1	Pembelian pakaian dalam setahun (var_1)	,503	,000
2	Pelayanan Kesehatan (var_11)	,311	,000
3	Sumber Keuangan Rumah Tangga (var_10)	,184	,000
4	Jenis Lantai (var_6)	,111	,000
5	Pendidikan ART umur 6 - 15 tahun (var_9)	,089	,000
6	Sumber Penerangan (var_8)	,086	,000
7	Kepemilikan rumah (var_4)	,083	,000
8	Fasilitas Air Bersih (var_2)	,080	,000
9	Jenis dinding (var_5)	,078	,000
10	Sarana Tempat Buang Air Besar (var_7)	,078	,000

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 4	,078	43392,173	40	,000
2 through 4	,245	23911,025	27	,000
3 through 4	,618	8167,750	16	,000
4	,968	545,377	7	,000

Standardized Canonical Discriminant Function Coefficients

	Function			
	1	2	3	4
Pembelian pakaian dalam setahun	-,157	,944	,069	,173
Fasilitas Air Minum	-,139	,154	-,102	,291
Kepemilikan rumah	-,123	,046	-,297	,102
Jenis dinding	,101	,041	,210	,227
Jenis Lantai	,519	,004	,703	-,162
Sarana Tempat Buang Air Besar	,042	,117	,063	-,191
Sumber Penerangan	,141	-,130	,237	,745
Pendidikan ART umur 6 - 15 tahun	,155	-,572	-,077	,360
Sumber Keuangan Rumah Tangga	,865	,166	-,073	-,150
Pelayanan Kesehatan	,737	,201	-,501	,065

Functions at Group Centroids

Cluster Number of Case	Function			
	1	2	3	4
1	3,538	,960	-1,487	1,785E-02
2	8,498E-02	,456	,535	-3,841E-02
3	-2,132	9,533E-02	-,920	6,322E-02
4	,708	-3,598	-5,406E-02	-8,298E-02
5	1,901	-1,392	1,629	1,868

Classification Results

Original	Cluster Number of Case	Predicted Group Membership					Total
		1	2	3	4	5	
Count	1	1371	97	19	6	5	1498
	2	0	9795	247	6	1	10049
	3	0	356	3163	66	0	3585
	4	17	514	63	1099	18	1711
	5	17	54	6	21	49	147
%	1	91,5	6,5	1,3	,4	,3	100,0
	2	,0	97,5	2,5	,1	,0	100,0
	3	,0	9,9	88,2	1,8	,0	100,0
	4	1,0	30,0	3,7	64,2	1,1	100,0
	5	11,6	36,7	4,1	14,3	33,3	100,0

91,1% of original grouped cases correctly classified.