

METODE IMPUTASI UNTUK MENGATASI DATA HILANG  
PADA ANALISIS KOMPONEN UTAMA NONLINIER  
DATA ORDINAL

SKRIPSI

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

oleh :

NUR ISTIQOMAH  
0910953007-95



PROGRAM STUDI STATISTIKA  
JURUSAN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS BRAWIJAYA  
MALANG

2013

LEMBAR PENGESAHAN SKRIPSI

METODE IMPUTASI UNTUK MENGATASI DATA HILANG  
PADA ANALISIS KOMPONEN UTAMA NONLINIER  
DATA ORDINAL

oleh :

NUR ISTIQOMAH

0910953007-95

Setelah dipertahankan di depan Majelis Pengaji  
Pada tanggal 11 Desember 2013  
dan dinyatakan memenuhi syarat untuk memperoleh gelar  
Sarjana Sains dalam bidang Statistika

Pembimbing I

Dr. Ir. Solimun, MS  
NIP. 196112151987031002

Pembimbing II

Rahma Fitriani, S.Si., M.Sc, Ph.D  
NIP. 19760328 199903 2 001

Mengetahui,  
Ketua Jurusan Matematika  
Fakultas MIPA Universitas Brawijaya

Dr. Abdul Rouf Alghofari, M.Sc  
NIP. 19670907 199203 1 001

## LEMBAR PERNYATAAN

Saya yang bertandatangan di bawah ini :

Nama : Nur Istiqomah  
NIM : 0910953007-95  
Jurusan : Matematika  
Program Studi : Statistika  
Penulis Skripsi Berjudul : Metode Imputasi untuk Mengatasi Data Hilang pada Analisis Komponen Utama Nonlinier Data Ordinal

Dengan ini menyatakan bahwa :

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

Demikian pernyataan ini dibuat dengan segala kesadaran.

Malang, 11 Desember 2013  
Yang menyatakan,

(Nur Istiqomah)  
NIM. 0910953007-95

# METODE IMPUTASI UNTUK MENGATASI DATA HILANG PADA ANALISIS KOMPONEN UTAMA NONLINIER

## DATA ORDINAL

### ABSTRAK

Analisis Komponen Utama merupakan analisis peubah ganda untuk mentransformasi p peubah asal menjadi k peubah baru, di mana antar vektor dari peubah baru tersebut adalah saling bebas. Jenis data yang digunakan dalam penelitian bisa nominal, ordinal, interval, rasio, atau campuran. Jika jenis data yang digunakan adalah nominal atau ordinal, maka analisis yang digunakan merupakan pengembangan Analisis Komponen Utama yang disebut Analisis Komponen Utama Nonlinier. Salah satu permasalahan yang sering dihadapi dalam survei dengan cara memberikan kuesioner adalah terdapat data hilang atau data tidak lengkap. Apabila persentase data hilang cukup besar dan pola data hilang adalah acak atau MCAR (*Missing Completely at Random*), pendugaan terhadap nilai data hilang (metode imputasi) menjadi alternatif pilihan yang layak untuk dilakukan agar data tersebut menjadi data lengkap dan dapat diproses lebih lanjut. Dalam penelitian ini, dibandingkan tiga jenis metode imputasi untuk mengatasi data hilang pada data ordinal. Metode tersebut adalah *Median Imputation* (MEI), *Mode Imputation* (MOI), *Forward Imputation* (FOI). Untuk mengetahui metode terbaik, maka ketiganya dibandingkan menggunakan nilai kosinus (korelasi) dan RMSE (*Root Mean Square Error*). Metode FOI lebih baik dari MEI dan MOI dalam mengatasi data hilang pada data ordinal karena menghasilkan nilai kosinus paling besar, dan menghasilkan nilai RMSE paling kecil. Pada penelitian ini, belum dapat dibuktikan bahwa semakin besar persentase data hilang maka nilai Bobot Peubah dan Skor Komponen Objek yang dihasilkan dari metode Imputasi juga semakin berbeda dari nilai bobot peubah dan skor komponen objek pada data lengkap.

Kata Kunci : Analisis Komponen Utama Nonlinier, *Forward Imputation*, *Mode Imputation*, *Median Imputation*.

# **IMPUTATION METHOD FOR HANDLING MISSING DATA ON NONLINEAR PRINCIPAL COMPONENT ANALYSIS ORDINAL DATA**

## **ABSTRACT**

*Principal Component Analysis is an analysis of multiple variables to transform the  $p$  original variables into  $k$  new variables. Vectors of the new variables are independent. Data used was in the research can be nominal, ordinal, interval, ratio, or a mixture. If the type of data is nominal or ordinal, the analysis was used Nonlinear Principal Component Analysis. One of the problems often encountered in the survey is missing data or incomplete data . If percentage of missing data is large and the pattern is random or MCAR (Missing Completely at Random), estimate the value of missing data ( imputation method ) being a viable alternative so that the data is complete and the data can be processed further. In this research, compared three types of imputation methods to handling missing data on ordinal data . The methods are Median Imputation (MEI), Mode Imputation (MOI), Forward Imputation (FOI). To determine the best method, then all three methods are compared using cosine values (correlation) and RMSE (Root Mean Square Error). The FOI method better than the MEI and MOI method in handling missing data on ordinal data, because FOI method has the greatest value of the cosine, and has the smallest RMSE value. In this research, has not been proven that the more large percentage of missing data make variable loadings and object score resulting from imputation methods are also increasingly different from the variable loadings and object scores on the full data.*

**Keywords :** Nonlinear Principal Component Analysis, Forward Imputation, Mode Imputation, Median Imputation

## KATA PENGANTAR

Puji Syukur kepada Tuhan Yang Maha Esa, sehingga penulis dapat menyelesaikan Skripsi yang berjudul “Metode Imputasi untuk Mengatasi Data Hilang pada Analisis Komponen Utama Nonlinier Data Ordinal”. Dalam penyusunan Skripsi ini, banyak pihak yang telah memberikan bantuan kepada penulis. Oleh kerena itu, pada kesempatan ini penulis ingin menyampaikan terima kasih kepada :

1. Bapak Dr. Ir. Solimun, MS, selaku Dosen Pembimbing I atas segala motivasi, bimbingan, masukan, dan nasehat selama proses penyelesaian Skripsi ini hingga dapat terselesaikan dengan baik.
2. Ibu Rahma Fitriani, S.Si., M.Sc., Ph.D, selaku Dosen Pembimbing II atas segala bimbingan dan nasehat selama proses penyelesaian Skripsi ini hingga dapat terselesaikan dengan baik.
3. Ibu Dr. Atiek Iriany, MS., selaku Dosen Pengaji atas saran dan masukan yang telah diberikan.
4. Bapak Dr. Abdul Rouf Alghofari, M.Sc, selaku Ketua Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang.
5. Orang tua, keluarga, dan teman-teman untuk segala kasih sayang, semangat, dan doa yang selalu diberikan kepada penulis selama ini.
6. Teman-teman Program Studi Statistika 2009 atas bantuan, dukungan, dan kerjasamanya.
7. Seluruh staf pengajaran Jurusan Matematika atas bantuan dan kerjasamanya.
8. Semua pihak yang telah membantu dalam penyusunan Skripsi ini yang tidak dapat penulis sebutkan seluruhnya.

Penulis menyadari bahwa Skripsi ini masih kurang sempurna. Oleh karena itu, saran dan kritik yang membangun sangat penulis harapkan untuk penulisan yang lebih baik. Semoga Skripsi ini dapat memberikan manfaat bagi pembaca.

Malang, Desember 2013

Penulis

## DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL.....</b>	<b>i</b>
<b>LEMBAR PENGESAHAN SKRIPSI .....</b>	<b>ii</b>
<b>LEMBAR PERNYATAAN .....</b>	<b>iii</b>
<b>ABSTRAK .....</b>	<b>iv</b>
<b>ABSTRACT .....</b>	<b>v</b>
<b>KATA PENGANTAR .....</b>	<b>vi</b>
<b>DAFTAR ISI .....</b>	<b>vii</b>
<b>DAFTAR GAMBAR .....</b>	<b>ix</b>
<b>DAFTAR TABEL .....</b>	<b>x</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xi</b>
 <b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah .....	3
1.4 Tujuan Penelitian .....	3
1.5 Manfaat Penelitian .....	4
 <b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Jenis Data .....	5
2.2 Analisis Komponen Utama .....	6
2.2.1 Input menggunakan Matriks Ragam Peragam .....	6
2.2.2 Input Menggunakan Matriks Korelasi .....	9
2.3 Analisis Komponen Utama Nonlinier .....	11
2.4 Data Hilang ( <i>Missing Data</i> ) .....	16
2.4.1 Persentase Data Hilang .....	17
2.4.2 Jenis-Jenis Data Hilang berdasarkan Tingkat Keacakan .....	17
2.4.3 Diagnosa Keacakan dari Data Hilang .....	18
2.5 Data Hilang pada Analisis Komponen Utama Nonlinier .....	19
2.6 Metode Imputasi.....	20
2.6.1 Pendekatan <i>Nearest Neighbor</i> (NN) .....	24
2.7 Metode Imputasi Terbaik.....	25
 <b>BAB III METODE PENELITIAN</b>	
3.1 Data .....	27

3.2 Pembentukan Data Hilang .....	29
3.3 Metode Analisis.....	31
3.3.1 Metode Analisis Penelitian .....	31
3.3.2 Analisis Komponen Utama Nonlinier .....	33
3.3.3 <i>Median Imputation</i> .....	34
3.3.4 <i>Mode Imputation</i> .....	37
3.3.5 <i>Forward Imputation</i> .....	39
<b>BAB IV HASIL DAN PEMBAHASAN</b>	
4.1 Data Lengkap .....	43
4.1.1 Analisis Komponen Utama Nonlinier Data Lengkap .....	43
4.1.1.1 Penentuan Skala Pengukuran .....	43
4.1.1.2 Penentuan Banyaknya Dimensi .....	43
4.1.1.3 Penentuan Skor Komponen Objek dan Kuantifikasi Kategori .....	45
4.1.1.4 Penentuan Bobot Peubah .....	45
4.1.1.5 Tes Konvergensi .....	45
4.2 Data Lengkap Hasil Imputasi .....	46
4.2.1 Analisis Komponen Utama Nonlinier Data Lengkap Hasil Imputasi .....	47
4.2.1.1 Penentuan Skala Pengukuran .....	47
4.2.1.2 Penentuan Banyaknya Dimensi .....	47
4.2.1.3 Penentuan Skor Komponen Objek dan Kuantifikasi Kategori .....	50
4.2.1.4 Penentuan Bobot Peubah .....	50
4.2.1.5 Tes Konvergensi .....	50
4.3 Penentuan Metode Imputasi Terbaik .....	50
4.4 Pengaruh Semakin Besar Persentase Data Hilang Terhadap Bobot Peubah dan Skor Komponen Objek yang Dihasilkan dari Metode Imputasi .....	55
<b>BAB V KESIMPULAN DAN SARAN</b>	
5.1 Kesimpulan .....	61
5.2 Saran .....	61
<b>DAFTAR PUSTAKA .....</b>	63
<b>LAMPIRAN .....</b>	67

## DAFTAR GAMBAR

	Halaman
Gambar 2.1 <i>Forward Imputation</i> : langkah awal dari Prosedur iteratif dengan menguraikan data matriks H .....	22
Gambar 2.2 Grafis yang mempresentasikan prosedur <i>Forward Imputation</i> .....	23
Gambar 3.1 Diagram Alir Metode Penelitian .....	32
Gambar 3.2 Diagram Alir Analisis Komponen Utama Nonlinier .....	34
Gambar 3.3 Diagram Alir Metode <i>Median Imputation</i> .....	36
Gambar 3.4 Diagram Alir Metode <i>Mode Imputation</i> .....	38
Gambar 3.5 Diagram Alir Metode <i>Forward Imputation</i> ....	41



## DAFTAR TABEL

	Halaman
Tabel 2.1 Data matriks H .....	19
Tabel 2.2 Matriks Indikator G dari Matriks H .....	20
Tabel 3.1 Daftar data dengan Banyak Nilai Pengamatan yang Hilang .....	30
Tabel 4.1 Nilai Akar Ciri Seluruh Dimensi pada Data Lengkap (Data 1) .....	44
Tabel 4.2 Akar Ciri, Proporsi Keragaman dan Proporsi Keragaman Kumulatif Ketiga Dimensi pada Data 1 .....	44
Tabel 4.3 Tes Konvergensi Data 1 .....	45
Tabel 4.4 Nilai Akar Ciri Seluruh Dimensi pada Data Lengkap Hasil Imputasi (Data 1) .....	48
Tabel 4.5 Metode Imputasi Terbaik Berdasarkan Nilai $\cos(\underline{a}_t, \underline{a}_\theta)$ Terbesar .....	52
Tabel 4.6 Metode Imputasi Terbaik Berdasarkan nilai RMSE Terkecil .....	55
Tabel 4.7 Pengaruh Semakin Besar Persentase Data Hilang dalam Menghasilkan Nilai Kosinus .....	57
Tabel 4.8 Pengaruh Semakin Besar Persentase Data Hilang dalam Menghasilkan Nilai RMSE .....	59

## DAFTAR LAMPIRAN

	Halaman
Lampiran 1	Data Lengkap .....
Lampiran 2	Data Hilang pada Setiap Kondisi .....
Lampiran 3	Hasil Pendugaan Nilai Data Hilang .....
Lampiran 4	Nilai Akar Ciri Seluruh Dimensi pada Data lengkap .....
Lampiran 5	Penentuan Banyak Dimensi yang Digunakan 80
Lampiran 6	Akar Ciri, Proporsi Keragaman danproporsi Keragaman Kumulatif padaData lengkap ... 81
Lampiran 7	Nilai Skor Komponen Objek pada Data Lengkap .....
Lampiran 8	Nilai Kuantifikasi pada Data lengkap .....
Lampiran 9	Nilai Bobot Peubah pada data lengkap .....
Lampiran 10	Tes konvergensi pada Data Lengkap .....
Lampiran 11	Nilai Akar Ciri Seluruh Dimensi padaData lengkap Hasil Imputasi .....
Lampiran 12	Akar Ciri, Proporsi Keragaman dan Proporsi Keragaman Kumulatif pada Data Lengkap Hasil Imputasi .....
Lampiran 13	Nilai Skor Komponen Objek pada DataLengkap Hasil Imputasi .....
Lampiran 14	Nilai Kuantifikasi pada Data Lengkap Hasil Imputasi .....
Lampiran 15	Nilai Bobot Peubah pada data Lengkap Hasil Imputasi .....
Lampiran 16	Tes konvergensi pada Data Lengkap Hasil Imputasi .....
Lampiran 17	Hasil Perhitungan $\cos(\underline{a}_t, \underline{a}_\theta)$ .....
Lampiran 18	Hasil Perhitungan RMSE <sub>t</sub> .....
Lampiran 19	Diagram Garis Metode Imputasi Paling Baik Berdasarkan Nilai $\cos(\underline{a}_t, \underline{a}_\theta)$ Terbesar .....
Lampiran 20	Diagram Garis Metode Imputasi Paling Baik Berdasarkan Nilai RMSE <sub>t</sub> Terkecil .. 99
Lampiran 21	Diagram Garis Pengaruh Semakin Besar

Lampiran 22

Persentase Data Hilang dalam Menghasilkan Nilai $\cos(\underline{a}_t, \underline{a}_\theta)$ .....	101
Diagram Garis Pengaruh Semakin Besar Persentase Data Hilang dalam Menghasilkan Nilai RMSE <sub>t</sub> .....	103



## **BAB I**

### **PENDAHULUAN**

#### **1.1 Latar Belakang**

Dalam pengumpulan data, selain pengamatan dengan satu peubah (peubah tunggal) terdapat pula pengamatan dengan lebih satu peubah (peubah ganda). Analisis peubah ganda yang bertujuan untuk mentransformasi p peubah asal menjadi k peubah baru di mana vektor antar peubah baru yang terbentuk adalah saling bebas menurut Johnson dan Wichern (2002) disebut dengan Analisis Komponen Utama (AKU). Masing-masing k peubah baru yang terbentuk juga merupakan kombinasi linier dari p peubah lama. Pada awalnya Analisis Komponen Utama (AKU) digunakan untuk kasus dengan data berskala interval dan rasio. Sedangkan untuk data berskala nominal dan ordinal menurut Gifi (1983) suatu analisis yang digunakan merupakan pengembangan Analisis Komponen Utama (AKU) yang disebut Analisis Komponen Utama Nonlinier.

Sama halnya dengan Analisis Komponen Utama (AKU), Analisis Komponen Utama Nonlinier juga tetap mempertahankan prinsip utama bahwa komponen utama yang terbentuk sebagai peubah baru merupakan kombinasi linier dari peubah asal, di mana peubah baru tersebut mampu menjelaskan sebanyak mungkin keragaman data sehingga nantinya tidak banyak informasi yang hilang.

Terdapat beberapa jenis data yang digunakan dalam penelitian bidang sosial, yaitu nominal, ordinal, interval, rasio, atau campuran. Data tersebut diperoleh melalui berbagai macam teknik pengumpulan data, salah satunya yaitu dengan cara membagikan kuesioner kepada responden. Pada penelitian ini, difokuskan pada data dengan skala ordinal. Oleh karena itu analisis yang digunakan adalah Analisis Komponen Utama Nonlinier.

Salah satu permasalahan yang sering dihadapi dalam survei adalah data tidak lengkap/data hilang (*missing data*). Hal tersebut disebabkan oleh beberapa hal di antaranya adalah responden tidak memberikan jawaban pada kuesioner, kelalaian saat melakukan perekaman data, data tidak dicatat dengan benar atau oleh hal lain yang tidak diketahui sebabnya. Jika data hilang dengan persentase yang besar dibiarkan begitu saja maka dapat mempengaruhi analisis statistik, terutama ketika beberapa peubah secara sistematis hilang

untuk beberapa objek. Cara termudah untuk menangani masalah tersebut adalah membuang objek-objek yang mempunyai pengamatan hilang dan hanya menganalisa objek-objek yang tidak mempunyai pengamatan hilang. Tetapi prosedur tersebut membuang informasi dan dapat mengakibatkan hasil yang buruk. Selain itu jika banyak responden yang disurvei sudah sesuai dengan banyak responden yang dibutuhkan, maka pengurangan objek-objek akan membuat data yang digunakan menjadi tidak cukup, sehingga perlu dilakukan survei tambahan untuk mendapatkan data pengganti. Oleh karena itu, apabila persentase data hilang cukup besar, pendugaan terhadap nilai data hilang menjadi alternatif pilihan yang layak untuk dilakukan agar data tersebut menjadi data lengkap dan dapat diproses lebih lanjut.

Dalam survei, banyak faktor yang menyebabkan data hilang/ data tidak lengkap tidak dapat dilakukan penggantian sampel. Beberapa faktor tersebut diantaranya adalah keterbatasan biaya, tingkat kesulitan yang tinggi serta keterbatasan jumlah sampel yang mungkin untuk diobservasi.

Sebelum memberikan perlakuan (*treatment*) untuk menangani data hilang, terlebih dahulu harus diketahui tingkat keacakan atau pola data hilang. Jika data hilang memiliki pola acak atau MCAR (*Missing Completely at Random*) maka untuk menangani permasalahan data hilang, salah satu metode umum yang digunakan adalah mengganti data hilang tersebut menggunakan metode imputasi. Menurut Acuna dan Rodriguez (2003) metode imputasi adalah proses pengisian atau penggantian nilai-nilai yang hilang pada sekumpulan data (*dataset*) dengan nilai-nilai yang mungkin berdasarkan informasi yang didapatkan pada data tersebut. Penggantian data hilang dilakukan agar diperoleh data lengkap (tidak terdapat data hilang) dengan menggunakan metode imputasi, sehingga analisa lebih lanjut untuk data lengkap bisa dilakukan.

Dalam penelitian ini, tiga jenis metode imputasi digunakan untuk mengatasi data hilang pada data ordinal. Dua di antaranya adalah metode yang umum untuk digunakan, yaitu metode *Median Imputation* (MEI) dan *Mode Imputation* (MOI). Kedua metode tersebut akan dibandingkan dengan metode imputasi baru yang diusulkan oleh Ferrari *et al.* (2011) untuk mengatasi data hilang pada data ordinal yaitu metode *Forward Imputation* (FOI).

Dalam penelitian Ferrari *et al.* (2011), diperoleh bahwa metode *Forward Imputation* adalah metode yang lebih baik dalam

mengatasi data hilang untuk peubah kategori yang disajikan dalam konteks Analisis Komponen Utama Nonlinier. Dalam penelitian ini, ingin mengetahui apakah metode baru yang diusulkan tersebut adalah metode yang lebih baik dalam mengatasi data hilang dari pada metode *Median Imputation* dan *Mode Imputation*. Oleh karena itu, dalam penelitian ini digunakan sepuluh data yang disajikan dalam konteks Analisis Komponen Utama Nonlinier data ordinal. Sepuluh data tersebut dilakukan pembentukan data hilang dan kemudian diatasi menggunakan metode *Median Imputation*, *Mode Imputation*, dan *Forward Imputation*. Setelah dilakukan metode imputasi maka data hilang (data tidak lengkap) tersebut menjadi data lengkap hasil imputasi. Data lengkap hasil imputasi yang terbentuk dari masing-masing metode dilakukan Analisis Komponen Utama Nonlinier untuk mendapatkan nilai bobot peubah dan skor komponen objek. Kemudian dilakukan perbandingan nilai bobot peubah dan skor komponen objek dari ketiga metode metode imputasi. Berdasarkan hasil perbandingan tersebut dapat diketahui metode mana yang lebih baik dalam mengatasi data hilang.

## 1.2 Rumusan Masalah

1. Metode apa yang lebih baik dalam mengatasi data hilang pada Analisis Komponen Utama Nonlinier data ordinal?
2. Bagaimanakah pengaruh semakin besar persentase data hilang terhadap nilai bobot peubah (*variable loadings*) yang dihasilkan dari metode imputasi?
3. Bagaimanakah pengaruh semakin besar persentase data hilang terhadap skor komponen objek (*object scores*) yang dihasilkan dari metode imputasi?

## 1.3 Batasan Masalah

1. Data yang digunakan adalah data ordinal
2. Pola data hilang yang digunakan adalah acak atau MCAR
3. Metode yang digunakan untuk mengatasi data hilang adalah metode imputasi : *Median imputation*, *Mode imputation*, dan *Forward imputation*

## 1.4 Tujuan Penelitian

1. Mengkaji apakah metode *Forward imputation* lebih baik dalam mengatasi data hilang daripada metode *Median imputation* dan *Mode imputation*

2. Mempelajari pengaruh semakin besar persentase data hilang terhadap nilai bobot peubah (*variable loadings*) yang dihasilkan dari metode imputasi
3. Mempelajari pengaruh semakin besar persentase data hilang terhadap skor komponen objek (*object scores*) yang dihasilkan dari metode imputasi

## 1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah memperoleh informasi apakah metode *Forward Imputation* terbukti adalah metode yang lebih baik penanganannya dalam mengatasi data hilang pada Analisis Komponen Utama Nonlinier data ordinal.



## BAB II

### TINJAUAN PUSTAKA

#### 2.1 Jenis Data

Data adalah kumpulan informasi yang diperoleh dari suatu pengamatan, dapat berupa angka, lambang atau sifat. Berdasarkan skala ukurnya, data dapat dibedakan menjadi empat, yaitu data nominal, ordinal, interval dan rasio. Skala pengukuran diperlukan untuk mengklasifikasikan peubah yang akan diukur supaya tidak terjadi kesalahan dalam menentukan analisis data dan langkah penelitian selanjutnya. Berikut adalah empat jenis data berdasarkan skala pengukuran (Clark dan Schkade,1983):

##### 1. Data Nominal

Data nominal adalah tingkatan data paling rendah menurut tingkat pengukurannya. Data nominal membedakan benda atau peristiwa yang satu dengan yang lainnya berdasarkan nama (mengandung unsur identifikasi atau penamaan saja). Jadi, jika angka-angka atau lambang-lambang digunakan untuk mengidentifikasi suatu benda atau peristiwa, maka angka-angka tersebut adalah data nominal. Contoh data nominal antara lain : jenis kelamin (laki-laki ; perempuan) dan nama kota (Malang ; Surabaya; Blitar). Setiap jenis kelamin dan kota diberi skor (dikuantitatifkan) sebelum dilakukan analisis, misalnya : Jenis kelamin (laki-laki = 1 ; perempuan = 2) dan nama kota (Kota Malang = 1, Kota Blitar = 2 dan Kota Surabaya =3 ).

##### 2. Data Ordinal

Data ordinal adalah data yang selain mengandung unsur identifikasi atau penamaan juga mengandung unsur tingkatan atau urutan (data disusun menurut peringkat masing-masing). Contoh dari data ordinal adalah kepribadian tenaga penjualan, misalnya, bisa diperingkat dari yang “paling buruk” sampai yang “paling baik”. Jika ingin memeringkatkan  $n$  buah peristiwa berdasarkan suatu ciri tertentu, kita boleh menetapkan nomor 1 untuk peristiwa yang ciri tertentunya paling kurang, nomor 2 untuk peristiwa yang ciri tertentunya kedua paling kurang, dan seterusnya hingga nomor  $n$  untuk peristiwa dengan kadar ciri tertentu yang paling tinggi. Pada contoh sebelumnya : paling buruk diberi skor 1, buruk diberi skor 2, cukup baik diberi skor 3, baik diberi skor 4, dan paling baik diberi skor 5.

### 3. Data Interval

Data interval mempunyai tingkatan lebih tinggi dari data ordinal. Data interval dapat dibeda-bedakan antara yang satu dan lainnya, kemudian diurutkan, perbedaan antara peringkat yang satu dan yang lainnya mempunyai arti, yakni apabila satuan pengukurannya tetap. Data interval dapat diurutkan dan diukur dengan pasti dengan unit yang sama tetapi tidak memiliki nilai nol mutlak. Sebagai contoh adalah suhu udara yang diukur dalam satuan Celcius, jika dalam suatu ruangan bersuhu nol maka bukan berarti ruangan tersebut tidak mempunyai suhu atau temperatur udara.

### 4. Data Rasio

Apabila pengukuran yang dilakukan memiliki sifat-sifat yang terdapat pada ketiga jenis data sebelumnya serta sifat tambahan bahwa ratio antara masing-masing pengukuran mempunyai arti dan memiliki nilai nol mutlak, maka disebut dengan data rasio. Contoh data rasio adalah pengukuran jarak yang diukur dalam satuan centimeter. Jarak antara satu titik dengan dirinya sendiri adalah nol. Semua operator metematika dapat dilakukan terhadap data rasio.

## 2.2 Analisis Komponen Utama

Analisis Komponen Utama pada dasarnya bertujuan untuk mentransformasi  $p$  peubah asal menjadi  $k$  peubah baru yang saling orthogonal. Masing-masing  $k$  peubah baru tersebut merupakan kombinasi linier dari  $p$  peubah asal. Keragaman informasi yang dimiliki oleh  $p$  peubah asal sebagian besar dapat diterangkan oleh  $k$  peubah baru (Johnson & Wichern, 2002).

Ada 2 macam data masukan (input) pada Analisis Komponen Utama yaitu menggunakan matriks ragam peragam atau menggunakan matriks korelasi. Jika peubah asal memiliki satuan yang sama maka Analisis Komponen Utama menggunakan matriks ragam peragam, sedangkan jika peubah asal memiliki satuan yang berbeda, maka digunakan matriks korelasi .

### 2.2.1 Input Menggunakan Matriks Ragam Peragam

Menurut Johnson dan Wichern (2002) komponen utama didefinisikan sebagai kombinasi linier dari  $p$  peubah yang dinyatakan dalam bentuk matriks sebagai berikut :

$$\underline{KU} = \underline{A}' \underline{X};$$

$$\begin{bmatrix} \underline{KU}_1 \\ \underline{KU}_2 \\ \vdots \\ \underline{KU}_k \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pk} \end{bmatrix}' \begin{bmatrix} \underline{X}_1 \\ \underline{X}_2 \\ \vdots \\ \underline{X}_p \end{bmatrix}$$

Atau dalam bentuk kombinasi linier sebagai berikut :

$$\begin{aligned} \underline{KU}_1 &= a_{11}\underline{X}_1 + a_{21}\underline{X}_2 + \dots + a_{p1}\underline{X}_p = \underline{a}_1' \underline{X} \\ \underline{KU}_2 &= a_{12}\underline{X}_1 + a_{22}\underline{X}_2 + \dots + a_{p2}\underline{X}_p = \underline{a}_2' \underline{X} \\ &\vdots \\ \underline{KU}_k &= a_{1k}\underline{X}_1 + a_{2k}\underline{X}_2 + \dots + a_{pk}\underline{X}_p = \underline{a}_k' \underline{X} \end{aligned} \quad (2.1)$$

dengan

$\underline{KU}_1, \underline{KU}_2, \dots, \underline{KU}_k$  = komponen utama dari  $\underline{X}$

$\underline{X}_1, \underline{X}_2, \dots, \underline{X}_p$  = peubah asal

$\underline{a}_k$  = vektor koefisien.

Pemilihan vektor koefisien komponen utama  $\underline{a}_k = (a_{1k}, a_{2k}, \dots, a_{pk})$  adalah sedemikian rupa sehingga ragam dari komponen utama  $Var(\underline{KU}_k)$  maksimum dengan batasan  $\underline{a}_k' \underline{a}_k = 1$ . Ragam dari komponen utama pertama dapat dirumuskan sebagai berikut (Timm, 1975):

$$Var(\underline{KU}_1) = Var(\underline{a}_1' \underline{X}) = \underline{a}_1' \Sigma \underline{a}_1 \quad (2.2)$$

Agar  $Var(\underline{KU}_1)$  maksimum dengan batasan  $\underline{a}_1' \underline{a}_1 = 1$  maka digunakan metode penggandaan Lagrange yaitu:

$$f(\underline{a}_1, \lambda_1) = \underline{a}_1' \Sigma \underline{a}_1 - \lambda_1 (\underline{a}_1' \underline{a}_1 - 1) \quad (2.3)$$

Fungsi tersebut akan mencapai maksimum jika turunan parsial pertama  $f(\underline{a}_1, \lambda_1)$  terhadap  $\underline{a}_1$  sama dengan nol.

$$\frac{\partial f(\underline{a}_1, \lambda_1)}{\partial \underline{a}_1} = 2 \sum \underline{a}_1 - 2\lambda_1 \underline{a}_1 = 0$$

atau

$$\sum \underline{a}_1 = \lambda_1 \underline{a}_1 \quad (2.4)$$

Persamaan (2.4) dipenuhi oleh  $\lambda_1$  dan  $\underline{a}_1$  yang merupakan pasangan akar ciri dan vektor ciri matriks ragam peragam  $\Sigma$ . Akibatnya  $\underline{a}_1' \Sigma \underline{a}_1 = \underline{a}_1' \lambda_1 \underline{a}_1 = \lambda_1 \underline{a}_1' \underline{a}_1 = \lambda_1$ . Oleh karena itu  $Var(\underline{KU}_1) = \underline{a}_1' \Sigma \underline{a}_1 = \lambda_1$ , di mana  $\lambda_1$  adalah akar ciri yang terbesar dari matriks ragam peragam  $\Sigma$  dan  $\underline{a}_1$  adalah vektor ciri yang bersesuaian dengan  $\lambda_1$ .

Selanjutnya ragam dari komponen utama kedua dirumuskan sebagai berikut :

$$Var(\underline{KU}_2) = Var(\underline{a}'_2 \underline{X}) = \underline{a}'_2 \Sigma \underline{a}_2 \quad (2.5)$$

Vektor pembobot  $\underline{a}'_2$  dipilih sedemikian hingga keragaman komponen utama kedua maksimum, serta ortogonal terhadap vektor pembobot  $\underline{a}'_1$  dari komponen utama pertama. Agar ragam dari komponen utama kedua maksimum, serta antara komponen utama kedua tidak berkorelasi dengan komponen utama pertama, maka vektor pembobot  $\underline{a}_2$  dipilih sedemikian sehingga  $\underline{KU}_2 = \underline{a}'_2 \underline{X}$  tidak berkorelasi dengan  $\underline{KU}_1 = \underline{a}'_1 \underline{X}$

Ragam dari komponen utama kedua akan maksimum dengan kendala  $\underline{a}'_2 \underline{a}_2 = 1$  dan  $cov(\underline{KU}_1, \underline{KU}_2) = cov(\underline{a}'_1 \underline{X}, \underline{a}'_2 \underline{X}) = \underline{a}'_1 \Sigma \underline{a}_2 = 0$ . Karena  $\underline{a}_1$  adalah vektor ciri dari  $\Sigma$  dan  $\Sigma$  adalah matriks simetrik, maka:

$$\underline{a}'_1 \Sigma = \underline{a}'_1 \Sigma' = (\Sigma \underline{a}_1)' = (\lambda_1 \underline{a}_1)' = \lambda_1 \underline{a}'_1$$

Kendala  $\underline{a}'_1 \Sigma \underline{a}_2 = \lambda_1 \underline{a}'_1 \underline{a}_2 = 0$  dapat dituliskan sebagai  $\underline{a}'_1 \underline{a}_2 = 0$ .

Jadi fungsi Lagrange yang dimaksimumkan adalah :

$$f(\underline{a}_2, \lambda_2, \lambda) = \underline{a}'_2 \Sigma \underline{a}_2 - \lambda_2 (\underline{a}'_2 \underline{a}_2 - 1) - \lambda (\underline{a}'_1 \underline{a}_2 - 0) \quad (2.6)$$

Fungsi ini mencapai maksimum jika turunan parsial pertama  $f(\underline{a}_2, \lambda_2, \lambda)$  terhadap  $\underline{a}_2$  sama dengan nol, sehingga diperoleh

$$\frac{\partial f(\underline{a}_2, \lambda_2, \lambda)}{\partial \underline{a}_2} = 2\Sigma \underline{a}_2 - 2\lambda_2 \underline{a}_2 - \lambda \Sigma \underline{a}_1 = 0 \quad (2.7)$$

Jika persamaan (2.7) dikalikan dengan  $\underline{a}'_1$  maka diperoleh :

$$2\underline{a}'_1 \Sigma \underline{a}_2 - 2\lambda_2 \underline{a}'_1 \underline{a}_2 - \lambda \underline{a}'_1 \Sigma \underline{a}_1 = 0 \quad \text{karena } \underline{a}'_1 \Sigma = \lambda_1 \underline{a}'_1 \text{ maka}$$

$$2\underline{a}'_1 \Sigma \underline{a}_2 - 2\lambda_2 \underline{a}'_1 \underline{a}_2 - \lambda \lambda_1 \underline{a}'_1 \underline{a}_1 = 0$$

$$2\underline{a}'_1 \Sigma \underline{a}_2 - 0 - \lambda \lambda_1 = 0$$

Oleh karena  $2\underline{a}'_1 \Sigma \underline{a}_2 = 0$  maka  $\lambda = 0$ . Dengan demikian persamaan (2.7) setelah diturunkan terhadap  $\underline{a}_2$  menjadi :

$$\begin{aligned} \frac{\partial f(\underline{a}_2, \lambda_2, \lambda)}{\partial \underline{a}_2} &= 2\Sigma \underline{a}_2 - 2\lambda_2 \underline{a}_2 = 0 \\ \Sigma \underline{a}_2 - \lambda_2 \underline{a}_2 &= 0 \end{aligned} \quad (2.8)$$

Jadi  $\lambda_2$  dan  $\underline{a}_2$  merupakan pasangan akar ciri dan vektor ciri dari matriks ragam peragam  $\Sigma$ .  $\underline{a}_2$  adalah vektor ciri yang bersesuaian dengan akar ciri terbesar kedua dari matriks ragam peragam  $\Sigma$ .

Secara umum komponen utama ke- $k$  dapat dituliskan sebagai :

$$\underline{KU}_k = a_{1k}\underline{X}_1 + a_{2k}\underline{X}_2 + \dots + a_{pk}\underline{X}_p$$

$$\underline{KU}_k = \underline{a}_k \underline{X}$$

dengan :

$$\underline{a}'_k = (a_{1k}, a_{2k}, \dots, a_{pk})' \text{ dan } \underline{a}'_k \underline{a}_k = 1$$

Vektor pembobot  $\underline{a}'_k$  diperoleh dengan memaksimumkan keragaman komponen utama ke- $k$ , yaitu :

$$Var \underline{KU}_k = \underline{a}'_k \Sigma \underline{a}_k$$

dengan kendala:

$$\underline{a}'_k \underline{a}_k = 1 \text{ untuk } k = k \quad \text{serta}$$

$$\underline{a}'_k \underline{a}_k = 0 \text{ untuk } k \neq k$$

dengan kendala tersebut, maka sesama komponen utama tidak berkorelasi.

Total ragam peubah asal akan sama dengan total ragam komponen utama.

$$\begin{aligned} \sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} &= \sum_{i=1}^p Var(X_{ii}) = \lambda_1 + \lambda_2 + \dots + \lambda_p \\ &= \sum_{k=1}^p Var(\underline{KU}_k) \end{aligned}$$

Keragaman yang dapat diterangkan oleh komponen utama ke- $k$  terhadap keragaman total adalah :

$$\frac{\lambda_k}{\sum_{k=1}^p \lambda_k} \times 100\% \quad (2.9)$$

Sedangkan secara kumulatif, keragaman total yang dapat dijelaskan oleh  $m$  komponen utama adalah :

$$\frac{\sum_{k=1}^m \lambda_k}{\sum_{k=1}^p \lambda_k} \times 100\% \quad (2.10)$$

di mana  $m < p$ ;  $m$  = banyak komponen utama yang digunakan.

## 2.2.2 Input Menggunakan Matriks Korelasi

Bila  $p$  peubah asal diukur dengan satuan pengukuran berbeda maka peubah tersebut perlu ditransformasi ke skor baku. Pembakuan peubah asal  $X$  ke dalam peubah  $Z$  dapat dilakukan dengan menggunakan rumus (Johnson & Wichern, 2002) :

$$Z_{ij} = \frac{X_{ij} - \mu_j}{\sqrt{\sigma_{jj}}} \quad (2.11)$$

di mana :

i = objek pengamatan (1,2,...,n)

j = peubah (1,2,...,p)

$\mu_j$  = rata-rata peubah asal ke-j

$\sqrt{\sigma_{jj}}$  = simpangan baku peubah asal

Persamaan transformasi Z bila dinyatakan dalam bentuk matriks :

$$\underline{Z} = (\underline{V}^{1/2})^{-1} (\underline{X} - \underline{\mu}) \quad (2.12)$$

Dengan  $\underline{V}^{1/2}$  adalah matriks diagonal simpangan baku peubah asal, sedangkan elemen yang lain adalah nol, dan dalam bentuk matriks  $(\underline{V}^{1/2})^{-1}$  dapat ditulis sebagai berikut :

$$(\underline{V}^{1/2})^{-1} = \begin{bmatrix} \frac{1}{\sqrt{\sigma_{11}}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{\sigma_{22}}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{\sigma_{pp}}} \end{bmatrix}$$

Hubungan antara matriks ragam peragam  $\underline{\Sigma}$  dengan matriks korelasi peubah asal  $\underline{R}$  dapat dinyatakan dengan :

$$\underline{R} = (\underline{V}^{1/2})^{-1} \underline{\Sigma} (\underline{V}^{1/2})^{-1} \quad (2.13)$$

Dengan demikian komponen utama dari Z dapat ditentukan dari vektor ciri yang diperoleh melalui matriks korelasi peubah asal  $\underline{R}$ , sehingga komponen utama ke-k adalah :

$$\begin{aligned} \underline{KU}_k &= \underline{a}'_k \underline{Z} \\ &= a_{1k} Z_1 + a_{2k} Z_2 + \cdots + a_{pk} Z_p \end{aligned} \quad (2.14)$$

Pada pembentukan komponen utama dari matriks korelasi terdapat bahwa jumlah dari semua akar ciri adalah p, yaitu sama dengan banyak peubah  $X_p$ .

$$\begin{aligned} \sum_{k=1}^p \text{Var}(\underline{KU}_k) &= \sum_{k=1}^p \text{Var}(\underline{Z}_k) \\ &= \lambda_1 + \lambda_2 + \cdots + \lambda_p = p \end{aligned} \quad (2.15)$$

Keragaman yang dapat diterangkan oleh komponen utama ke-k adalah :

$$\frac{\lambda_k}{p} \times 100\% , \quad k = 1, 2, \dots, p \quad (2.16)$$

## 2.3 Analisis Komponen Utama Nonlinier

Pada Analisis Komponen Utama, data yang digunakan adalah data berskala rasio dan interval. Untuk data berskala nominal dan ordinal, Gifi (1983) menyatakan bahwa analisis yang digunakan bukan Analisis Komponen Utama, melainkan Analisis Komponen Utama Nonlinier. Analisis Komponen Utama Nonlinier merupakan pengembangan dari Analisis Komponen Utama dan biasa disebut dengan PRINCALS (*Principal Component Analysis by Alternating Least Squares*) atau Analisis Komponen Utama dengan menggunakan pendugaan *alternating least squares*.

Sama halnya dengan Analisis Komponen Utama, Analisis Komponen Utama Nonlinier juga tetap mempertahankan prinsip utama bahwa komponen utama yang terbentuk sebagai peubah baru merupakan kombinasi linier dari peubah asal, di mana peubah baru tersebut harus mampu menjelaskan sebanyak mungkin keragaman informasi yang terkandung dalam data. Aspek baru dalam Analisis Komponen Utama Nonlinier adalah bahwa korelasi antar kategori peubah dengan dimensi yang terbentuk harus sudah ditentukan. Hal tersebut dapat dicapai dengan memberikan nilai numerik pada kategori peubah. Pemberian nilai numerik pada kategori suatu peubah disebut kuantifikasi (*quantification*). Kuantifikasi yang sama untuk setiap dimensi pada suatu kategori peubah disebut kuantifikasi tunggal. Sedangkan kuantifikasi terpisah untuk setiap dimensi disebut kuantifikasi berganda (Kroonenberg *et al.*, 1997).

Metode statistik multivariat yang digunakan untuk mencapai ukuran satu dimensi adalah Analisis Komponen Utama Nonlinier yang merupakan teknik pengurangan dimensi untuk data kategori. Analisis Komponen Utama Nonlinier memperoleh tiga kelompok unsur, yaitu bobot peubah (*variable loadings*), kategori kuantifikasi (*category quantifications*), dan skor komponen objek (*object scores*). Bobot peubah mengukur pentingnya peubah dalam menentukan indikator. Kategori kuantifikasi menetapkan skala kuantitatif untuk masing-masing peubah dan menjaga urutan kategori. Skor komponen objek merangkum semua peubah ke dalam nilai tunggal untuk setiap objek (Ferrari *et al.*, 2011).

Pada Analisis Komponen Utama Nonlinier, kategori semua peubah dengan skala selain numerik diberi kuantifikasi kategori dengan skala numerik yang sesuai. Di mana Analisis Komponen Utama Linier bertujuan untuk mengoptimalkan atau mencari rata-rata kuadrat korelasi antara peubah asal dengan komponen yang

terbentuk, sedangkan Analisis Komponen Utama Nonlinier bertujuan untuk mengoptimalkan atau mencari rata-rata kuadrat korelasi yang optimal antara peubah yang diberi kuantifikasi kategori dengan komponen utama yang terbentuk. Rata-rata kuadrat korelasi tersebut menunjukkan besarnya total keragaman yang dapat diterangkan oleh komponen utama. Dalam pencarian nilai optimal tersebut, bobot peubah dan kategori kuantifikasi akan memberikan nilai yang bervariasi sampai nilai optimum ditemukan. Sedangkan dalam Analisis Komponen Utama Linier, hanya bobot peubah yang bervariasi (Konig, 2002).

Apabila terdapat suatu data yang dibentuk ke dalam matriks  $\mathbf{H}$  yang berukuran  $n \times m$ , maka untuk memudahkan perhitungan Analisis Komponen Utama Nonlinier dipakai notasi (Gifi, 1990):

$n$  = banyak pengamatan (objek) ;  $i = 1, 2, \dots, n$

$m$  = banyak peubah ;  $j = 1, 2, \dots, m$

$k_j$  = banyak kategori pada peubah ke- $j$

$\underline{h}_j$  = vektor kolom ke- $j$  dari matriks  $H$  berukuran  $n \times 1$

$\mathbf{G}_j$  = matriks indikator dari  $\underline{h}_j$  berukuran  $n \times k_j$

di mana :

$$\mathbf{G}_{(j)i} = \begin{cases} 1, & \text{jika objek ke-}i \text{ berada dalam kategori ke-}r \text{ dari} \\ & \text{peubah } j \\ 0, & \text{jika objek ke-}i \text{ tidak berada dalam kategori ke-}r \text{ dari} \\ & \text{peubah } j \end{cases}$$

$i = 1, 2, \dots, n$  ;  $r = 1, 2, \dots, k_j$

$\mathbf{G} = (\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_j, \dots, \mathbf{G}_m)$  berukuran  $n \times \sum_{j=1}^m k_j$

Pada kasus data lengkap (tanpa *missing data*),  $\mathbf{G}_j$  dikatakan lengkap apabila setiap baris pada  $\mathbf{G}_j$  hanya mempunyai satu unsur bernilai satu dan yang lainnya bernilai nol. Dari matriks  $\mathbf{G}_j$  akan diperoleh frekuensi tiap kategori pada setiap peubah. Persamaan berikut akan berlaku jika mempunyai data lengkap :

$$\mathbf{G}_j \underline{u} = \mathbf{M}_j \underline{u} \quad (2.17)$$

$$\mathbf{D}_j = \mathbf{G}'_j \mathbf{G}_j \quad (2.18)$$

$$\mathbf{M}_* = \sum_{j=1}^m \mathbf{M}_j = m \mathbf{I}, \text{ karena } \mathbf{M}_j = \mathbf{I} \quad (2.19)$$

Isi diagonal utama matriks  $\mathbf{M}_j$  adalah :

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

di mana :

$\underline{\mathbf{u}}$  = vektor kolom yang semua elemennya adalah 1

$\mathbf{M}_j$  = matriks bujur sangkar yang berupa matriks  $I$

$\mathbf{D}_j$  = matriks diagonal yang merupakan jumlah kolom dari  $\mathbf{G}_j$

Analisis Komponen Utama Nonlinier didasarkan pada teori *meet loss* yang bertujuan untuk meminimumkan fungsi *homogeneity loss* ( $\sigma_M$ ). Fungsi *homogeneity loss* adalah sebagai berikut :

$$\sigma_M(\mathbf{X}, \mathbf{Y}) = m^{-1} \sum_{j=1}^m (\mathbf{X} - \mathbf{G}_j \mathbf{Y}_j)' (\mathbf{X} - \mathbf{G}_j \mathbf{Y}_j) \quad (2.20)$$

dengan normalisasi  $AVE(x_s) = 0$  untuk dimensi  $s = 1, 2, \dots, p$  dan  $\mathbf{X}' \mathbf{X} = I$ .

Fungsi  $\sigma_M$  diminimumkan menggunakan metode *alternating least square* untuk mendapatkan pendugaan nilai bobot peubah  $a$ . Algoritma untuk menghitung  $p$  dimensi pertama secara simultan dan meminimumkan  $\sigma_M$  dengan menggunakan *alternating least square* adalah (Gifi, 1990) :

$$\mathbf{Y}_j = \mathbf{D}_j^{-1} \underline{\mathbf{G}}_j' \mathbf{X} \quad (2.21)$$

$$\mathbf{Z} = \sum_{j=1}^m \underline{\mathbf{G}}_j \mathbf{Y}_j \quad (2.22)$$

$$\mathbf{X} = GRAM(\mathbf{Z}) \quad (2.23)$$

di mana :

$\mathbf{X}$  = matriks skor komponen objek (*object scores*) berukuran  $n \times p$  di mana  $p$  adalah banyak dimensi ( $p \leq m$ )

$\mathbf{Y}_j$  = matriks *multiple category quantification* berukuran  $k_j \times p$

$\underline{\mathbf{y}}_j$  = *single category quantification* untuk peubah  $h_j$

$AVE(x_s)$  = vektor yang merupakan rata-rata kolom dari elemen matriks  $\mathbf{X}_s$ .

$GRAM(\mathbf{Z})$  = matriks ortogonal dari ortogonalisasi Gram-Schmidt dari matriks  $\mathbf{Z}$

$a_j$  = bobot peubah (*variable loadings*)

Sasaran dari metode ini adalah untuk mencari skor komponen objek  $\mathbf{X}$  dan kategori kuantifikasi  $\mathbf{Y}_j$  sedemikian hingga nilai  $\sigma_M$  dapat minimal. Berikut adalah langkah-langkahnya (Gifi, 1983) :

1. Inisialisasi

- $\mathbf{X}$  diberi inisial (permulaan) dengan bilangan random yang dinormalkan sehingga  $\underline{u}' \mathbf{M}_* \mathbf{X} = 0$  dan  $\mathbf{X}' \mathbf{M}_* \mathbf{X} = mn I$ , kemudian didapatkan  $\tilde{\mathbf{X}}$
- $\tilde{\mathbf{Y}}_j$  dilakukan permulaan nilai dengan persamaan :  

$$\tilde{\mathbf{Y}}_j = \mathbf{D}_j^{-1} \mathbf{G}_j' \tilde{\mathbf{X}} \quad (2.24)$$
- $\tilde{\mathbf{y}}_j$  dilakukan permulaan yang digambarkan sebagai  $k_j$  bilangan bulat yang berurutan dan dinormalkan sehingga  $\underline{u}' \mathbf{D}_j \tilde{\mathbf{y}}_j = 0$  dan  $\tilde{\mathbf{y}}_j' \mathbf{D}_j \tilde{\mathbf{y}}_j = n$
- Bobot peubah juga dilakukan permulaan dengan persamaan :  

$$\tilde{a}_{jp} = \tilde{\mathbf{X}}' \mathbf{G}_j \tilde{\mathbf{y}}_j \quad (2.25)$$

2. Update Skor Komponen Objek  $\mathbf{X}$

Nilai  $\mathbf{X}$  baru (*update object scores*) dengan nilai  $\mathbf{Y}$  yang sudah diketahui dapat dihitung dengan langkah-langkah sebagai berikut:

$$\mathbf{Z} = \sum_{j=1}^m \mathbf{M}_j \mathbf{G}_j \mathbf{Y}_j \quad (2.26)$$

$$\tilde{\mathbf{Z}} = \left\{ \mathbf{M}_* - \left( \frac{\mathbf{M}_* \underline{u} \underline{u}' \mathbf{M}_*}{\underline{u}' \mathbf{M}_* \underline{u}} \right) \right\} \mathbf{Z} \quad (2.27)$$

3. Ortonormalisasi

Ortonormalisasi dilakukan untuk mencari  $\mathbf{M}_*$  yang ortonormal terhadap  $\mathbf{X}^+$  dan terdekat dengan  $\tilde{\mathbf{Z}}$ . Hal tersebut dilakukan dengan menentukan  $\mathbf{X}^+$  melalui persamaan :

$$\mathbf{X} = m^{1/2} \mathbf{M}_*^{-1/2} GRAM \left( \mathbf{M}_*^{-1/2} \tilde{\mathbf{Z}} \right) \quad (2.28)$$

4. Update Kategori Kuantifikasi

Setelah diperoleh nilai  $\mathbf{X}^+$  maka langkah berikutnya adalah memperbarui nilai kategori kuantifikasi dengan persamaan sebagai berikut :

$$\tilde{\mathbf{Y}}_j = \mathbf{D}_j^{-1} \mathbf{G}_j' \mathbf{X} \quad (2.29)$$

$$\underline{a}_{jp} = (\tilde{\mathbf{Y}}_j' \mathbf{D}_j \tilde{\mathbf{y}}_j) / (\tilde{\mathbf{y}}_j' \mathbf{D}_j \tilde{\mathbf{y}}_j) \quad (2.30)$$

$$\mathbf{y}_j = \tilde{\mathbf{Y}}_j \underline{a}_{jp} / \underline{a}'_{jp} \underline{a}_{jp} \quad (2.31)$$

$$Y_j = \underline{y}_j \underline{a}'_{jp} \quad (2.32)$$

5. Uji Konvergensi

Keragaman maksimum diperoleh dengan melalukan proses iterasi untuk mencapai kriteria konvergensi. Kriteria konvergensi ditentukan dengan selisih Total Fit atau peubah iterasi mendekati  $\varepsilon$  (bilangan positif yang kecil), sehingga *Total Fit* yang terakhir diperoleh merupakan keragaman kumulatif yang mampu menerangkan total keragaman.

$$TFIT = p - \sigma_M(\mathbf{X}, \mathbf{Y}) \quad (2.33)$$

di mana  $p$  adalah banyak dimensi yang terbentuk. Pada tahap iterasi, langkah kedua sampai keempat akan diulang sampai didapatkan nilai yang konvergen.

Jika data peubah ganda tanpa ada data hilang (*missing data*), Gifi (1989) menyatakan bahwa akar ciri dapat dicari dari  $m^{-1}\underline{R}(\mathbf{Q})$ .  $\underline{R}(\mathbf{Q})$  adalah matriks korelasi antara skala peubah yang optimal pada semua himpunan matriks  $\mathbf{Q}$ , di mana  $\underline{Q}_j = \mathbf{G}_j Y_j$ .

Banyak dimensi maksimum yang mungkin terbentuk adalah sama dengan  $m$ , di mana  $m$  adalah banyak peubah. Untuk menentukan berapa banyak dimensi yang digunakan dalam melakuan PRINCALS, maka terlebih dahulu harus mencari akar ciri dari masing-masing dimensi. Misalkan banyak peubah adalah 14, maka banyak dimensi maksimum yang mungkin terbentuk adalah 14. Selanjutnya, dari ke-14 dimensi tersebut, dicari akar ciri dari masing-masing dimensi. Setelah didapatkan akar ciri dari masing-masing dimensi, maka dapat ditentukan berapa banyak dimensi yang digunakan. Gifi (1983) menyatakan bahwa dalam PRINCALS banyak dimensi yang akan digunakan ditetapkan dengan kriteria akar ciri yang lebih besar dari  $1/m$ , di mana  $m$  adalah banyak peubah. Misalkan pada dimensi pertama sampai kelima memiliki akar ciri yang lebih besar dari  $1/m$ , maka banyak dimensi yang digunakan dalam melakuan PRINCALS adalah 5. Banyak dimensi yang terbentuk menunjukkan banyak komponen utama yang terbentuk. Secara umum, dengan melihat hubungan dari hasil yang diperoleh, maka komponen utama ke-  $p$  dapat dituliskan sebagai berikut :

$$\begin{aligned}
 KU_1 &= (a_{11}y_1 + a_{21}y_2 + \dots + a_{j1}y_j)/(a_{11}^2 + a_{21}^2 + \dots + a_{j1}^2) \\
 &= (\underline{a}_{j1}' \underline{y}_j)/(\underline{a}_{j1}^2)' \\
 &= \frac{\underline{a}_{j1}}{(\underline{a}_{j1}^2)'} \underline{y}_j \\
 &\vdots \\
 KU_p &= (a_{1p}y_1 + a_{2p}y_2 + \dots + a_{jp}y_j)/(a_{1p}^2 + a_{2p}^2 + \dots + a_{jp}^2) \\
 &= (\underline{a}_{jp}' \underline{y}_j)/(\underline{a}_{jp}^2)' \\
 &= \frac{\underline{a}_{jp}}{(\underline{a}_{jp}^2)'} \underline{y}_j
 \end{aligned} \tag{2.34}$$

## 2.4 Data Hilang (*Missing Data*)

Data hilang adalah adanya sel-sel kosong pada satu atau beberapa peubah. Data hilang dapat terjadi karena informasi yang dibutuhkan untuk sesuatu pada satu atau beberapa peubah tidak diberikan, sulit dicari atau memang informasi tersebut tidak tersedia (Nova & Mukid, 2011).

Penyebab data hilang ada dua, yaitu dari faktor eksternal (dari peneliti) dan faktor internal (dari objek). Contoh dari faktor eksternal yaitu data tidak dicatat dengan benar, kelalaian saat melakukan perekaman, sehingga melewatkannya bagian yang harus dicatat. Sedangkan contoh dari faktor internal yaitu responden dengan tidak sengaja melewati bagian dari kuesioner, responden dengan sengaja tidak mau memberikan beberapa jawaban pada kuesioner dikarenakan alasan tertentu, atau oleh hal lain yang tidak diketahui sebabnya (Nova & Mukid, 2011).

Dampak yang terjadi dari data hilang jika persentasenya cukup besar dan tidak diberikan perlakuan (*treatment*) untuk menanganinya adalah (Hair *et al.*, 2010) :

- Terjadi pengurangan sampel yang tersedia untuk analisis. Jika perlakuan (*treatment*) untuk data hilang tidak diterapkan, maka setiap observasi dengan data hilang pada salah satu peubah akan dikeluarkan dari analisis.
- Hasil statistik dapat menjadi bias jika pola data hilang tidak acak. Bias terjadi ketika pola data hilang menyebabkan data tertentu hilang dan data hilang mengakibatkan hasil yang salah.

Setiap analisis multivariat dimulai dengan pemeriksaan data hilang. Jika terdapat data hilang, maka dilakukan identifikasi

data hilang, yaitu dengan cara memeriksa tingkat data hilang (persentase data hilang) dan jenis data hilang yang terbentuk.

#### **2.4.1 Persentase Data Hilang**

Penentuan banyak atau persentase data hilang yaitu menentukan jumlah data hilang untuk setiap peubah, setiap objek, dan bahkan keseluruhan. Cara untuk menilai seberapa banyak data hilang adalah dengan tabulasi : persentase peubah dengan data hilang untuk setiap objek dan persentase objek dengan data hilang untuk setiap peubah.

Masalah utama dalam penentuan besarnya persentase data hilang adalah untuk menentukan apakah persentase data hilang cukup rendah untuk tidak mempengaruhi hasil. Menurut Hair *et al* (2010), jika persentase data hilang cukup rendah dan terdapat pola acak, maka dapat digunakan salah satu teknik imputasi untuk mengatasi data hilang. Jika persentase data yang hilang terlalu tinggi, maka peneliti harus mempertimbangkan pendekatan khusus untuk mendiagnosa keacakan pola data hilang sebelum melanjutkan untuk menerapkan perlakuan.

Perlakuan sederhana yang harus dipertimbangkan sebelum melanjutkan untuk mendiagnosa keacakan yaitu menghapus objek atau peubah dengan persentase data hilang yang besar. Ketika dilakukan penghapusan peubah maka harus didasarkan pada alasan toritis, seperti harus memastikan bahwa peubah alternatif sangat berkorelasi dengan peubah asli sehingga dapat mewakili maksud dari peubah asli (Munro, 2005)

#### **2.4.2 Jenis-Jenis Data Hilang Berdasarkan Tingkat Keacakan**

Sebelum memberikan perlakuan (*treatment*) yang sesuai untuk mengatasi data hilang, terlebih dahulu harus diketahui tingkat keacakan data hilang. Beberapa jenis data hilang berdasarkan tingkat keacakan data hilang adalah sebagai berikut :

a) *MNAR (Missing Not at Random)*

Misalkan terdapat dua peubah ( $X_1$  dan  $X_2$ ),  $X_1$  tidak memiliki data hilang tetapi  $X_2$  memiliki data hilang. Jika data hilang pada  $X_2$  berkorelasi dengan  $X_1$ , maka data hilang bersifat tidak acak (Hair *et al.*, 2010).

b) *MAR (Missing at Random)*

Menurut Croninger dan Douglas (2005) data hilang disebut MAR jika nilai-nilai yang hilang dari peubah, misalkan  $X_1$ , tergantung pada peubah  $X_2$  atau peubah lainnya, tetapi tidak tergantung pada  $X_1$ . Sebagai contoh :

Diasumsikan bahwa jenis kelamin responden (peubah  $X_1$ ) diketahui, dan ditanyakan tentang pendapatan rumah tangga (peubah  $X_2$ ). Misalkan data hilang yang terdapat pada pendapatan rumah tangga  $X_2$  adalah acak untuk laki-laki dan perempuan, tetapi terjadi pada frekuensi yang lebih tinggi untuk laki-laki daripada perempuan. Meskipun data hilang beroperasi secara acak, tetapi setiap perlakuan yang diterapkan pada data hilang masih akan mencerminkan pola yang sama seperti ketika terdapat data hilang. Hal tersebut dikarenakan jenis kelamin mempengaruhi distribusi akhir dari nilai pendapatan rumah tangga.

c) MCAR(*Missing Completely at Random*)

Menurut Croninger dan Douglas (2005) MCAR terjadi ketika kemungkinan terjadinya data hilang untuk peubah, misalkan  $X_1$ , tidak berhubungan dengan nilai  $X_1$  itu sendiri atau dengan peubah lainnya (terjadinya data hilang tidak berhubungan dengan nilai semua peubah).

Dengan mengambil contoh seperti pada MAR, maka data hilang disebut MCAR jika data hilang yang terdapat pada pendapatan rumah tangga ( $X_2$ ) secara acak hilang dalam proporsi yang sama bagi laki-laki dan perempuan. Jika hal tersebut terjadi, maka salah satu perlakuan dapat diterapkan untuk mengatasi data hilang.

#### 2.4.3 Diagnosa Keacakan dari Data Hilang

Hal yang harus dilakukan ketika terdapat data hilang adalah memastikan apakah data hilang terjadi dengan cara benar-benar acak. Ketika jumlah data sedikit, peneliti mungkin dapat secara langsung melihat pola dari data hilang. Tetapi jika ukuran sampel dan jumlah peubah meningkat, maka dibutuhkan tes diagnostik empiris. Beberapa program statistik menambahkan teknik khusus yang dirancang untuk analisis data hilang, yang umumnya mencakup satu atau dua tes diagnostik.

- 1) Diagnostik pertama adalah menilai pola data hilang dari peubah  $X$  tunggal dengan membentuk dua kelompok : pengamatan peubah  $X$  dengan data hilang dan pengamatan peubah  $X$  dengan tidak ada data hilang. Uji statistik kemudian dilakukan untuk menentukan apakah ada perbedaan yang signifikan antara kedua kelompok. Perbedaan signifikan mengidentifikasi kemungkinan pola data hilang adalah tidak acak.

- Contoh : diambil dari contoh sebelumnya, yaitu pendapatan rumah tangga dan jenis kelamin. Dibentuk dua kelompok responden, yaitu kelompok dengan data hilang pada pertanyaan pendapatan rumah tangga dan kelompok yang menjawab pertanyaan tersebut (tidak terdapat data hilang). Kemudian akan dibandingkan persentase jenis kelamin untuk setiap kelompok. Jika salah satu jenis kelamin, misalnya laki-laki, ditemukan dalam proporsi yang lebih besar pada kelompok data hilang, maka diduga pola data hilang tersebut adalah tidak acak.
- 2) Diagnostik kedua adalah uji keseluruhan keacakan yang menentukan apakah data hilang dapat diklasifikasikan sebagai MCAR. Tes ini menganalisis pola data hilang pada semua peubah dan membandingkannya dengan pola yang diharapkan untuk data hilang (pola acak). Jika tidak terdapat perbedaan yang signifikan, data hilang dapat diklasifikasikan sebagai MCAR (Hair *et al.*, 2010).

## 2.5 Data Hilang pada Analisis Komponen Utama Nonlinier

Pada kasus data lengkap (tanpa data hilang),  $G_j$  dikatakan lengkap apabila setiap baris pada  $G_j$  hanya mempunyai satu unsur bernilai satu dan yang lainnya bernilai nol. Sedangkan pada kasus data tidak lengkap (terdapat data hilang),  $G_j$  dikatakan tidak lengkap jika ada baris yang semua nilai elemennya bernilai nol. Matriks indikator tidak lengkap bisa dikuantifikasi dengan prinsip yang sama seperti yang diuraikan dalam bagian 2.3 untuk kasus data lengkap (Gifi, 1990).

Berikut adalah contoh matriks H yang mempunyai data hilang. Matriks H memiliki  $n=10$ ,  $m= 3$ ,  $k_j = 3$  ( $j = 1,2,3$ ). Elemen dari H adalah ‘*category labels*’ : peubah pertama mempunyai kategori ‘a’, ‘b’, ‘c’; yang kedua ‘p’, ‘q’, ‘r’; yang ketiga ‘u’, ‘v’, ‘w’ (dengan nol frekuensi untuk w).

Tabel 2.1 Data matiks H

Objek	Peubah		
	1	2	3
1	-	P	U
2	b	Q	V
3	-	R	V
4	a	P	U
5	b	P	V
6	c	P	V
7	a	P	U
8	a	P	V
9	c	-	V
10	a	P	V

dengan (-) adalah nilai yang hilang.

Berikut adalah tabel matriks indikator G yang berasal dari Tabel 2.1

Tabel 2.2 Matriks Indikator G dari data matriks H

Objek	G <sub>1</sub>			G <sub>2</sub>			G <sub>3</sub>		
	a	b	c	p	q	r	u	v	w
1	0	0	0	1	0	0	1	0	0
2	0	1	0	0	1	0	0	1	0
3	0	0	0	0	0	1	0	1	0
4	1	0	0	1	0	0	1	0	0
5	0	1	0	1	0	0	0	1	0
6	0	0	1	1	0	0	0	1	0
7	1	0	0	1	0	0	1	0	0
8	1	0	0	0	1	0	0	1	0
9	0	0	1	0	0	0	0	1	0
10	1	0	0	1	0	0	0	1	0

Pada Tabel 2.2 terlihat bahwa matriks indikator G<sub>1</sub> dan G<sub>2</sub> mempunyai baris yang semua elemennya bernilai nol. Hal tersebut dikarenakan matriks H adalah matriks data tidak lengkap (terdapat data hilang).

Menurut Gifi (1990), program PRINCALS meminimumkan  $\sigma_M$  dan menyediakan pilihan *missing data passive* apabila terdapat data hilang, yaitu menghapus nilai data yang hilang. Hal tersebut berarti fungsi *loss* yang mengandung matriks R(Q) yang menghasilkan akar ciri, hasilnya tidak dapat diinterpretasikan bila terdapat data hilang.

Data hilang dapat diatasi dengan metode imputasi. Dengan metode imputasi, maka data tidak lengkap akan menjadi data lengkap. Setelah menjadi data lengkap, maka analisis lebih lanjut dapat diselesaikan.

## 2.6 Metode Imputasi

Menurut Acuna & Rodriguez (2003) metode imputasi adalah proses pengisian atau penggantian nilai-nilai yang hilang pada data dengan nilai-nilai yang mungkin yang berdasarkan informasi yang didapatkan pada data. Terdapat beberapa metode yang termasuk teknik imputasi, diantaranya adalah *median imputation*, *mode imputation*, dan *forward imputation*.

### a) Median Imputation

*Median imputation* adalah metode di mana data hilang untuk peubah digantikan oleh median dari semua nilai yang terdapat pada peubah yang didasarkan pada semua nilai yang diketahui/valid. Metode imputasi ini diterapkan secara terpisah pada masing-masing peubah yang mengandung data hilang (Acuna & Rodriguez, 2003).

b) *Mode Imputation*

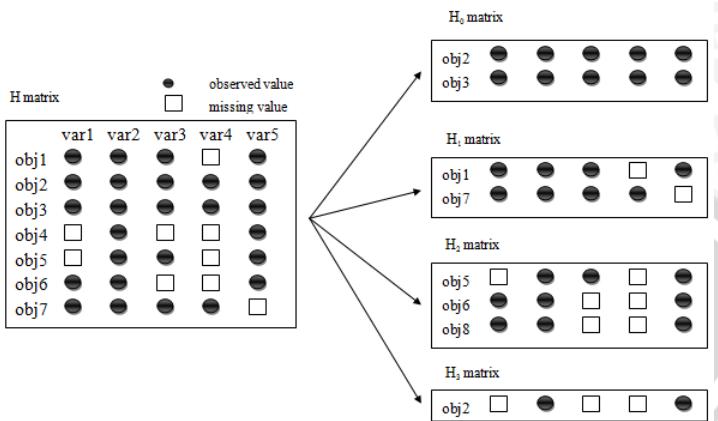
*Mode imputation* adalah metode di mana nilai-nilai yang hilang untuk peubah digantikan oleh modus dari semua nilai yang terdapat pada peubah yang didasarkan pada semua nilai yang diketahui/valid. Metode imputasi ini diterapkan secara terpisah pada masing-masing peubah yang mengandung nilai-nilai yang hilang.

Permasalahan data hilang dalam atribut bertipe kategori dapat digunakan metode *Mode Imputation*. Dalam *Mode Imputation* struktur hubungan antara data tidak dianggap. Menurut Acuna & Rodriguez (2003), keberadaan relasi yang tinggi antara atribut yang memiliki data hilang dengan suatu atribut yang lain dapat menyebabkan teknik *Mode Imputation* tidak berguna atau bahkan berbahaya.

c) *Forward Imputation*

Ferrari *et al.* (2011) mengusulkan alternatif penanganan data hilang dengan tujuan mengubah suatu dataset tidak lengkap menjadi data lengkap dengan mempertimbangkan pola hubungan diseluruh peubah. Usulan ini didasarkan pada algoritma iteratif yang menggunakan hasil Analisis Komponen Utama Nonlinier pada data tanpa ada data hilang dan imputasi yang dilakukan berurutan dengan metode *nearest neighbor*. Proses yang berurutan ini dimulai dari objek yang memiliki data hilang dengan jumlah terendah dan berakhir pada objek yang memiliki data hilang dengan jumlah tertinggi. Metode yang diusulkan tersebut bernama *Forward Imputation*.

Dalam Gambar 2.1 pemisahan data matriks mulai digambarkan untuk kasus sederhana dari lima peubah dan 8 objek.



(Ferrari *et al.*, 2011)

Gambar 2.1 *Forward imputation* : langkah awal dari prosedur iteratif dengan menguraikan data matriks H

Langkah awal dari prosedur iteratif ini adalah dengan cara menguraikan data matriks H menjadi  $K+1$  sub matriks ( $H_k$ ), di mana  $k$  adalah banyak pengamatan yang hilang dalam satu objek ( $k = 0, 1, \dots, K < m$ ) dan  $m$  adalah banyak peubah.

Metode *Forward Imputation* menggunakan seluruh informasi yang terkandung dalam data, memperhitungkan nilai yang diamati pada peubah lain (bertentangan dengan metode *median imputation* dan *mode imputation*). Selain itu, metode tersebut juga mempertahankan karakteristik Analisis Komponen Utama NonLinier, memastikan bahwa peran setiap objek dalam menafsirkan hasil terkait dengan banyak peubah yang diamati, dan memperhitungkan peran peubah-peubah dengan bobot (*loading*) dari peubah tersebut. Dalam memperhitungkan bobot (*loading*) dari peubah, maka digunakan jarak *weighted Minkowski* seperti berikut :

$$\min_z d_r(u_i^{(k)}; u_z^0) = \left( \sum_j a_j^{(k-1)} |G_j(i)\underline{y}_j^{(k-1)} - G_j(z)\underline{y}_j^{(k-1)}|^r \right)^{\frac{1}{r}} \quad (2.35)$$

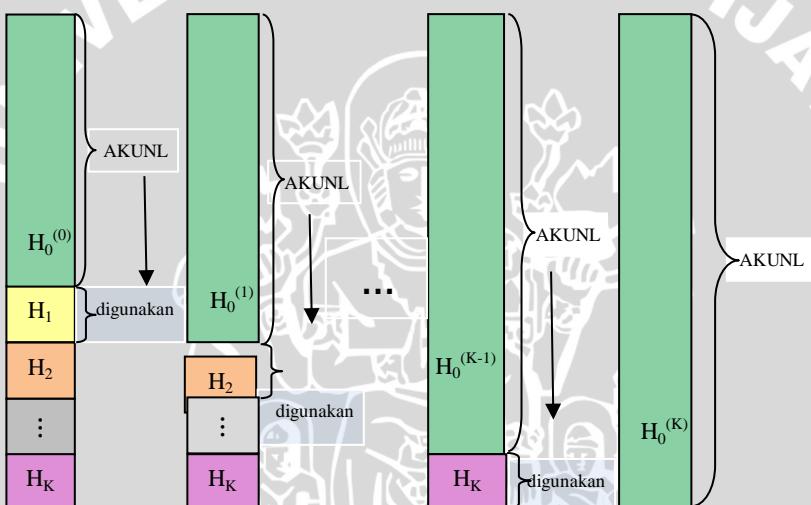
Untuk  $i = 1, \dots, n_k$  (ini adalah siklus pada objek) :

pilih objek (baris)  $u_i^{(k)}$  pada matriks tidak lengkap  $H_k$  dan mengidentifikasi objek pada matriks lengkap  $H_0^{(k-1)}$  yang paling dekat dalam jarak *weighted Minkowski*. Hasil yang terbentuk tidak terpengaruh oleh objek  $u_i^{(k)}$  mana yang terlebih dahulu

masuk ke dalam proses imputasi. Imputasi yang dilakukan dari setiap  $u_i^{(k)}$  hanya berdasarkan pada matriks  $H_0^{(k-1)}$ , dan akan diperbarui menjadi  $H_0^{(k)}$  ketika semua  $u_i^{(k)}$  telah lengkap.

Dengan  $r > 0$ ,  $u_z^{(k-1)} \in H_0^{(k-1)}$  dan hanya peubah yang diamati pada kedua objek  $u_i^{(k)}$  dan  $u_z^0$ . Jarak  $d_r(u_i^{(k)}, u_z^0)$  antara objek  $u_i^{(k)}$  dengan data hilang dan setiap objek  $u_z^0$  pada  $H_0^{(k-1)}$  adalah *weighted Minkowski*.

Pada Gambar 2.2 menunjukkan grafis yang merepresentasikan prosedur *forward Imputation* secara keseluruhan.



Keterangan :

AKUNL = Analisis Komponen Utama Nonlinier

(Ingrasia et al., 2011)

Gambar 2.2 Grafis yang merepresentasikan prosedur *Forward Imputation*

Secara umum, grafis tersebut menjelaskan proses *Forward Imputation* mulai dari penghitungan Analisis Komponen Utama Nonlinier menggunakan matriks lengkap  $H_0^{(0)}$  kemudian dari hasil tersebut digunakan untuk proses imputasi secara iteratif ke matriks tidak lengkap  $H_k$  ( $k= 1, 2, \dots, K$ ). Proses imputasi yang dilakukan menerapkan metode *nearest neighbor*. Setelah semua proses imputasi selesai, maka matriks tersebut menjadi matriks lengkap  $H_0^{(K)}$ . Dengan

terbentuknya matriks data lengkap, maka bisa dilakukan NLPKA pada matriks tersebut.

### 2.6.1 Pendekatan *Nearest Neighbor* (NN)

Sebuah data yang mengandung satu atau lebih data hilang yang harus diperhitungkan disebut sebagai objek target. Dalam pendekatan *Nearest Neighbor*, dilakukan pengukuran jarak antara objek target dan setiap objek yang lain, kemudian dipilih K objek terdekat dengan objek target (Wasito & Mirkin, 2005).

Untuk menerapkan pendekatan *Nearest Neighbor* (NN), berikut adalah dua hal yang harus dilakukan :

- 1) Mengukur jarak

Ukuran kedekatan masing-masing individu disebut dengan jarak. Ukuran jarak digunakan ketika terdapat dua objek yang berada pada titik yang berbeda. Jarak antar objek juga disebut dengan ukuran kemiripan atau *similarity*. Brimberg (1989) menyatakan bahwa individu p dan q adalah fungsi jarak jika sifat-sifat dasar berikut terpenuhi :

1.  $d(p,q) \geq 0 \rightarrow$  positif
2.  $d(p,q) = 0$  jika  $p = q$
3.  $d(p,r) \leq d(p,q) + d(q,r) \rightarrow$  ketidaksamaan segitiga (*triangle inequality*)
4.  $d(p,q) = d(q,p) \rightarrow$  simetri

Salah satu ukuran kedekatan adalah jarak *Minkowski*. Li *et al.* (2003) menyatakan bahwa jarak *Minkowski* secara umum dapat digunakan untuk mengukur kemiripan antar objek. Misalkan dua objek ke-i dan ke-j yang berada pada p dimensi ditulis dengan dua vektor yaitu  $(x_{i1}, x_{i2}, \dots, x_{ik})$  dan  $(x_{j1}, x_{j2}, \dots, x_{jk})$ . Jarak ini dihitung sebagai berikut :

$$d(i,j) = \left[ \sum_{k=1}^p |x_{ik} - x_{jk}|^r \right]^{1/r} \quad (2.36)$$

Secara umum, untuk  $r = 2$  jarak *Minkowski* lebih dikenal dengan jarak *Euclidean*, sedangkan untuk  $r = 1$  jarak *Minkowski* menjadi jarak *Manhattan* atau *City Block*.

Dalam konteks Analisis Komponen Utama Nonlinier, untuk memperhitungkan peran peubah maka perlu dilakukan penghitungan bobot (*loading*) dari peubah tersebut. Oleh karena itu, fungsi jarak yang digunakan adalah *Weighted Minkowski*. Fungsi jarak *Weighted Minkowski* didefinisikan seperti pada Persamaan (2.35)

## 2) Pemilihan *neighborhood*

Prinsip pemilihan objek terdekat yaitu pada semua objek, hanya objek yang tidak memiliki nilai yang hilang dalam peubah yang sama pada objek target yang memiliki nilai yang hilang (Wasito & Mirkin, 2005)

Wasito & Mirkin (2005) mengatakan bahwa secara singkat teknik *Nearest Neighbor* adalah : mengambil baris pertama yang memiliki nilai yang hilang sebagai objek target  $X_i$ , menemukan  $K$  *nearest neighbor* dan membentuk matriks  $X$  yang terdiri dari objek target dan *neighbors*. Kemudian menerapkan algoritma imputasi ke matriks  $X$  dengan hanya memasukkan pada nilai yang hilang yang terdapat pada objek target. Ulangi hal tersebut sampai semua nilai yang hilang diisi, kemudian menjadi matriks data lengkap.

Algoritma *Nearest neighbor* berbasis imputasi adalah sebagai berikut :

- 1) ambil baris pertama yang memiliki nilai yang hilang, dan baris tersebut adalah sebagai objek target  $X_i$
- 2) cari  $K$  *neighbors* dari  $X_i$
- 3) membuat data matriks  $X$  yang terdiri dari  $X_i$  dan  $K$  *neighbors* yang dipilih
- 4) terapkan algoritma imputasi ke matriks  $X$  dan masukkan ke dalam nilai yang hilang di  $X_i$
- 5) jika tidak ada nilai yang hilang dalam data, maka proses berhenti, jika masih terdapat nilai yang hilang maka kembali ke Langkah 1.

## 2.7 Metode Imputasi Terbaik

Penentuan metode imputasi yang terbaik dalam mengatasi data hilang menggunakan hasil dari Analisis Komponen Utama Nonlinier setelah terbentuk data lengkap. Analisis Komponen Utama Nonlinier dihitung pada dataset yang lengkap (data lengkap awal sebelum dilakukan pembentukan data hilang) dan pada dataset dengan data hilang buatan yang ditangani dengan ketiga metode imputasi (*Median Imputation (MEI)*, *Mode Imputation (MOI)* dan *Forward Imputation (FOI)*). Dengan metode imputasi tersebut, data tidak lengkap menjadi data lengkap hasil imputasi. Kemudian hasil dari perhitungan Analisis Komponen Utama Nonlinier pada data lengkap dan data lengkap hasil imputasi dievaluasi dengan cara perbandingan.

Ferrari *et al.* (2011) mengatakan bahwa perbandingan didasarkan pada vektor bobot peubah (*variable loadings*)  $a$  dan vektor dari skor komponen objek (*object scores*) X.

- a) Berdasarkan vektor bobot peubah (*variable loadings*)

Untuk membandingkan bobot peubah, kosinus antara vektor dari bobot peubah  $a_t$  untuk setiap  $t$  metode perlakuan data hilang dan vektor bobot peubah dari data lengkap (data lengkap sebelum dilakukan imputasi)  $a_\theta$  dihitung sebagai berikut:

$$\cos(a_t, a_\theta) = \frac{a_t' a_\theta}{\|a_t\| \|a_\theta\|} \quad (2.38)$$

Semakin kecil sudut kosinus maka nilainya akan menuju 1, dan hal tersebut berarti semakin dekat kedua vektor. Jika kedua vektor semakin dekat, maka dapat dikatakan bahwa korelasi bobot peubah dari data lengkap hasil metode imputasi dengan bobot peubah dari data lengkap adalah semakin kuat. Jika korelasinya mendekati 1, maka kemampuan metode  $t$  tersebut untuk mengatasi data hilang adalah lebih baik.

- b) Berdasarkan vektor skor komponen objek (*object scores*)

Untuk membandingkan skor komponen objek, dapat dihitung dengan persamaan *Root Mean Square Error* (RMSE) (Ferrari et al, 2011):

$$RMSE_t = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_{t,i} - X_{\theta,i})^2} \quad (2.39)$$

di mana :

$X_{t,i}$  = skor komponen objek (*object scores*) untuk objek  $i$  yang dihitung dengan metode  $t$  pada matriks data lengkap hasil imputasi

$X_{\theta,i}$  = skor komponen objek (*object scores*) untuk objek  $i$  pada matriks data lengkap

$n$  = banyak objek.

Semakin kecil nilai RMSE maka kesalahan hasil imputasi semakin kecil, begitu juga sebaliknya. Semakin kecil kesalahan imputasi maka menunjukkan bahwa metode tersebut adalah metode terbaik.

Analisis ini dilakukan untuk 3 kondisi tingkat data hilang (kondisi 1 = 0.5% ; kondisi 2 = 1% ; kondisi 3 = 1.5%) dan pola data hilang yang dihasilkan di bawah kondisi yang sama.

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Data**

Data yang digunakan dalam penelitian ini adalah 10 data sekunder dengan jenis skala ordinal. Data tersebut adalah data lengkap. Sepuluh data tersebut akan digunakan untuk menganalisis penentuan metode imputasi terbaik dalam mengatasi data hilang. Dalam penelitian ini, masing-masing data dibagi menjadi dua bagian, yaitu data lengkap dan data tidak lengkap (terdapat nilai-nilai yang hilang). Data yang terdapat nilai hilang merupakan data lengkap yang kemudian dibuat menjadi data hilang dengan pola acak atau *Missing Completely at Random* (MCAR). Berikut adalah 10 data yang akan digunakan dalam penelitian :

❖ **Data 1**

Data yang pertama diambil dari hasil survei Badan Pusat Statistik Kabupaten Lamongan (2012) yang berjudul “Pendapat Responden Tentang Pelayanan Publik pada Instansi A”. Data lengkap terdiri dari 14 peubah dan 100 objek sehingga total pengamatan adalah 1400 pengamatan. Data ke-1 beserta keterangan peubah-peubahnya dapat dilihat pada Lampiran 1.

❖ **Data 2**

Data yang kedua diambil dari skripsi Putri (2008) mahasiswa Jurusan Teknologi industri Pertanian Fakultas Teknologi Pertanian Universitas Brawijaya yang berjudul “Analisis Kualitas Pelayanan terhadap Tingkat Kepuasan Konsumen (studi kasus di restoran Boyong Kalegan Yogyakarta)”. Data lengkap terdiri dari 23 peubah dan 72 objek sehingga total pengamatan adalah 1656 pengamatan. Data ke-2 beserta keterangan peubah-peubahnya dapat dilihat pada Lampiran 1.

❖ **Data 3**

Data yang ketiga diambil dari skripsi Hapsari (2010) mahasiswa Jurusan Sosial Ekonomi pertanian Fakultas Pertanian Universitas Brawijaya yang berjudul “Analisis Persepsi Konsumen terhadap Jus Buah Dalam Kemasan (studi kasus di Giant-Hypermarket Mall Olympic Garden Malang)”. Data lengkap terdiri dari 9 peubah dan 49 objek sehingga total pengamatan adalah 441 pengamatan. Merk jus buah dalam kemasan yang digunakan dalam skripsi tersebut adalah Buavita dan ABC Juice, tapi dalam penelitian ini peneliti hanya menggunakan data merk Buavita. Data

ke-3 beserta keterangan peubah-peubahnya dapat dilihat pada Lampiran 1.

❖ Data 4

Data yang keempat diambil dari skripsi Irawati (2002) mahasiswa Jurusan Manajemen Fakultas Ekonomi Universitas Brawijaya yang berjudul “Persepsi Konsumen terhadap Bauran Pemasaran Rokok Starmild 16 pada Perusahaan Rokok PT. Bentoel Prima Malang ”. Data lengkap terdiri dari 14 peubah dan 56 objek sehingga total pengamatan adalah 784 pengamatan. Data ke-4 beserta keterangan peubah-peubahnya dapat dilihat pada Lampiran 1.

❖ Data 5

Data yang kelima diambil dari skripsi Tohari (2005) mahasiswa Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Brawijaya yang berjudul “Penerapan Multivariate dalam Menentukan Segmentasi Konsumen Plasa Araya dan *Positioning* terhadap Pesaingnya”. Data lengkap terdiri dari 17 peubah dan 103 objek sehingga total pengamatan adalah 1751 pengamatan. Data ke-5 beserta keterangan peubah-peubahnya dapat dilihat pada Lampiran 1.

❖ Data 6

Data yang keenam diambil dari skripsi Nugraha (2011) mahasiswa Jurusan Manajemen Fakultas Ekonomi Universitas Brawijaya yang berjudul “Pengukuran Kualitas Pelayanan Berdasarkan persepsi Konsumen terhadap Hotel Ollino Garden Malang”. Data lengkap terdiri dari 20 peubah dan 100 objek sehingga total pengamatan adalah 2000 pengamatan. Data ke-6 beserta keterangan peubah-peubahnya dapat dilihat pada Lampiran 1.

❖ Data 7

Data yang ketujuh diambil dari kripsi Maharani (2007) mahasiswa Jurusan Teknologi industri Pertanian Fakultas Teknologi Pertanian Universitas Brawijaya yang berjudul “Persepsi Konsumen terhadap Kualitas Bakpao Telo dengan Metode *Importance-performance Analysis*”. Data lengkap terdiri dari 19 peubah dan 75 objek sehingga total pengamatan adalah 1425 pengamatan. Data ke-7 beserta keterangan peubah-peubahnya dapat dilihat pada Lampiran 1.

❖ Data 8

Data yang kedelapan diambil dari skripsi Wahyuni (2008) mahasiswa Jurusan Sosial Ekonomi Pertanian Fakultas Pertanian Universitas Brawijaya yang berjudul “Persepsi Konsumen terhadap Apel Lokal dan Apel Impor”. Data lengkap terdiri dari 11 peubah

dan 82 objek sehingga total pengamatan adalah 902 pengamatan. Terdapat beberapa jenis apel yang digunakan dalam skripsi tersebut, tetapi dalam penelitian ini peneliti hanya menggunakan data apel manalagi. Data ke-8 beserta keterangan peubah-peubahnya dapat dilihat pada Lampiran 1.

❖ Data 9

Data yang kesembilan diambil dari skripsi Fitrida (2009) mahasiswa Jurusan Sosial Ekonomi Pertanian Fakultas Pertanian Universitas Brawijaya yang berjudul “Analisis Tingkat Kepuasan dan Persepsi Konsumen terhadap Produk Teh *Cup Tong Tji* (studi kasus di Malang Town Square dan Mal Olympic Garden Malang)”. Data lengkap terdiri dari 14 peubah dan 70 objek sehingga total pengamatan adalah 980 pengamatan. Data ke-9 beserta keterangan peubah-peubahnya dapat dilihat pada Lampiran 1.

❖ Data 10

Data yang kesepuluh diambil dari skripsi Wahyuningsih (2012) mahasiswa Jurusan Teknologi industri Pertanian Fakultas Teknologi Pertanian Universitas Brawijaya yang berjudul “Analisis Persepsi Konsumen terhadap *Puree Mangga Podang* dengan Metode *Importance Performance Analysis* (studi kasus pada UKM Sari Buah dan Dodol di Kota Batu)”. Data lengkap terdiri dari 15 peubah dan 31 objek sehingga total pengamatan adalah 465 pengamatan. Data ke-10 beserta keterangan peubah-peubahnya dapat dilihat pada Lampiran 1.

### 3.2 Pembentukan Data Hilang

Pada penelitian ini, data yang digunakan adalah data lengkap (tanpa data hilang). Agar tujuan penelitian tercapai, maka dilakukan penghapusan/penghilangan beberapa nilai pengamatan. Penghapusan nilai-nilai pengamatan dilakukan secara acak dengan bantuan *software R-2.15.2*. Banyak pengamatan yang dihapus sesuai dengan persentase data hilang yang diinginkan (0,5%, 1% dan 1,5%). Berikut adalah daftar dari 10 data beserta banyak nilai pengamatan yang dihilangkan.

Tabel 3.1 Daftar Data dengan Banyak Nilai Pengamatan yang Hilang

Data	Banyak Pengamatan	Data Hilang (%)	Banyak Pengamatan yang Hilang
1	1400	0.5	7
		1	14
		1.5	21
2	1656	0.5	9
		1	17
		1.5	25
3	441	0.5	3
		1	5
		1.5	7
4	784	0.5	4
		1	8
		1.5	12
5	1751	0.5	9
		1	18
		1.5	27
6	2000	0.5	10
		1	20
		1.5	30
7	1425	0.5	8
		1	15
		1.5	22
8	902	0.5	5
		1	10
		1.5	14
9	980	0.5	5
		1	10
		1.5	15
10	465	0.5	3
		1	5
		1.5	7

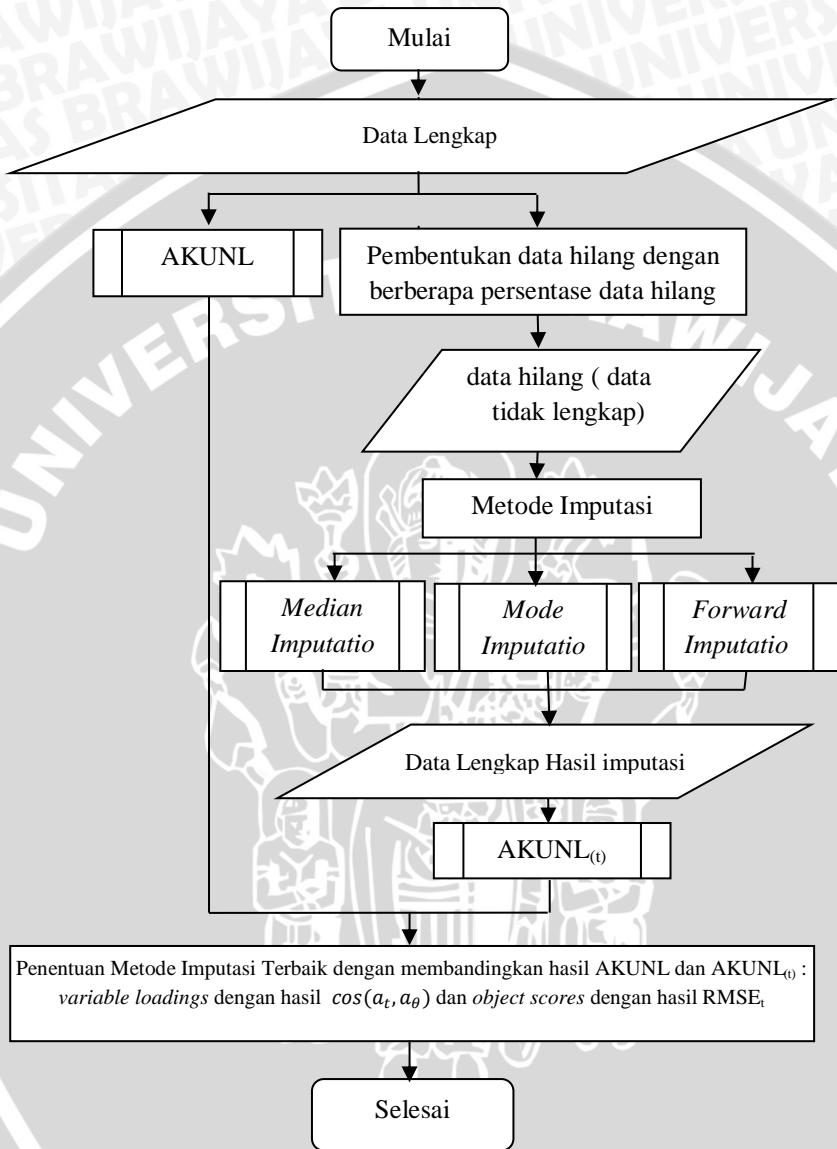
### **3.3 Metode Analisis**

Metode analisis yang digunakan untuk mengatasi data hilang pada penelitian ini adalah metode *Median Imputation*, *Mode Imputation*, dan *Forward Imputation* dalam konteks Analisis Komponen Utama Nonlinier (AKUNL) data ordinal. Langkah-langkah analisis penelitian ini adalah :

#### **3.3.1 Metode Analisis Penelitian**

1. Mempersiapkan data lengkap
2. Melakukan Analisis Komponen Utama Nonlinier
3. Pembentukan data tidak lengkap, yaitu membuat data hilang sebanyak :  
kondisi 1 = 0.5%, kondisi 2 = 1%, dan kondisi 3 = 1.5%, dari banyak pengamatan pada data lengkap (membuat data hilang berbagai persentase dengan pola acak)
4. Melakukan penanganan data hilang untuk masing-masing kondisi dengan menggunakan metode imputasi : *Median Imputation* (MEI), *Mode Imputation* (MOI), dan *Forward Imputation* (FOI)
5. Melakukan AKUNL<sub>(t)</sub> pada data lengkap yang terbentuk dari setiap metode imputasi (t = MEI, MOI, dan FOI)
6. Membandingkan hasil AKUNL dan AKUNL<sub>(t)</sub> berdasarkan:
  - Bobot peubah (*Variable loadings*) → dengan menggunakan Persamaan (2.37)
  - Skor Komponen Objek (*Object scores*) → dengan menggunakan Persamaan (2.38)

Analisis metode imputasi dilakukan dengan bantuan *Software R-2.15.2* sedangkan Analisis Komponen Utama Nonlinier (AKUNL) menggunakan bantuan *software SPSS 17.0*. Secara sistematis langkah-langkah analisis dalam penelitian ini dapat dilihat pada Gambar 3.1.



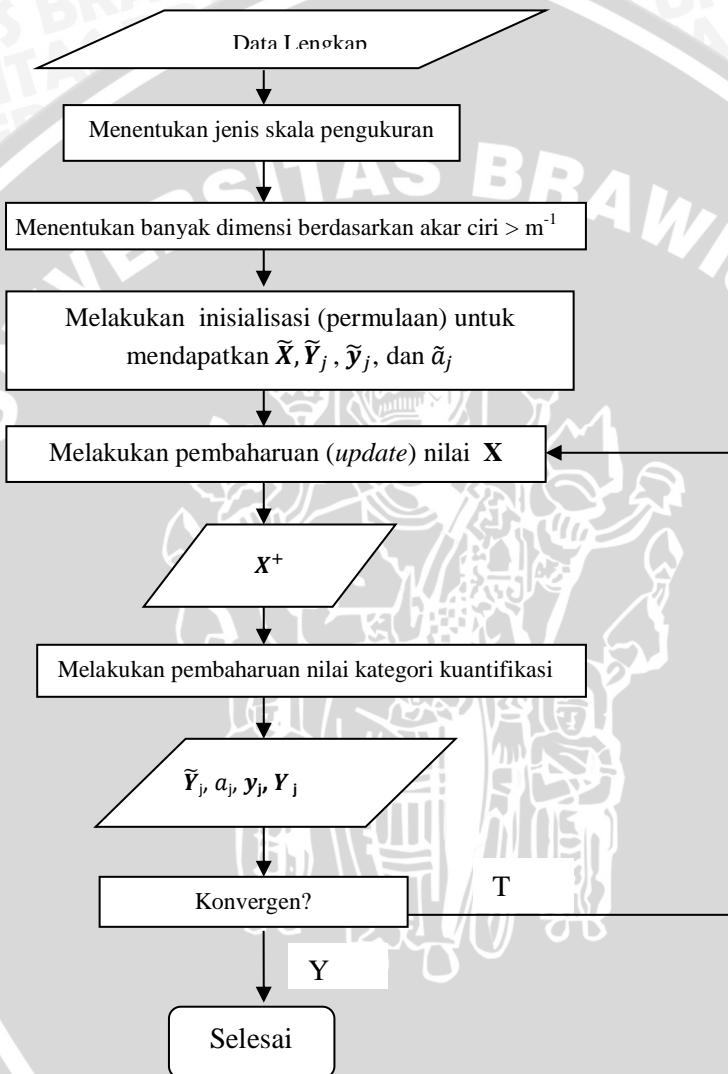
Gambar 3.1. Diagram Alir Metode Penelitian

### 3.3.2 Analisis Komponen Utama Nonlinier (AKUNL)

Untuk masing-masing kondisi, langkah-langkah Analisis Komponen Utama Nonlinier dalam penelitian ini dapat dijelaskan sebagai berikut :

1. Menyiapkan data lengkap
2. Menentukan skala pengukuran peubah untuk masing-masing peubah
3. Menentukan banyak dimensi berdasarkan kriteria akar ciri  $> m^{-1}$
4. Melakukan inisialisasi (permulaan) untuk mendapatkan  $\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}_j, \tilde{\mathbf{y}}_j$ , dan  $\tilde{a}_j$  seperti pada persamaan (2.24) dan (2.25)
5. Melakukan pembaharuan (*Update*) nilai  $\mathbf{X}$  seperti pada persamaan (2.28)
6. Melakukan pembaharuan (*update*) nilai kategori kuantifikasi untuk mendapatkan  $\tilde{\mathbf{Y}}_j, \underline{a}_j, \mathbf{y}_j, \mathbf{Y}_j$  seperti pada persamaan (2.29), (2.30), (2.31) dan (2.32)
7. Melakukan tes konvergensi dengan menghitung nilai *total fit* pada persamaan (2.33) kemudian dibandingkan dengan kriteria konvergensi  $\varepsilon$ . Apabila selisih *total fit* masih jauh dari  $\varepsilon$  maka kembali melakukan iterasi (kembali kelahkah 3) sampai nilai selisih *total fit* mendekati  $\varepsilon$ , agar diperoleh total keragaman maksimum

Langkah-langkah Analisis Komponen Utama Nonlinier AKUNL atau AKUNL<sub>(t)</sub> secara sistematis dapat dilihat pada Gambar 3.2



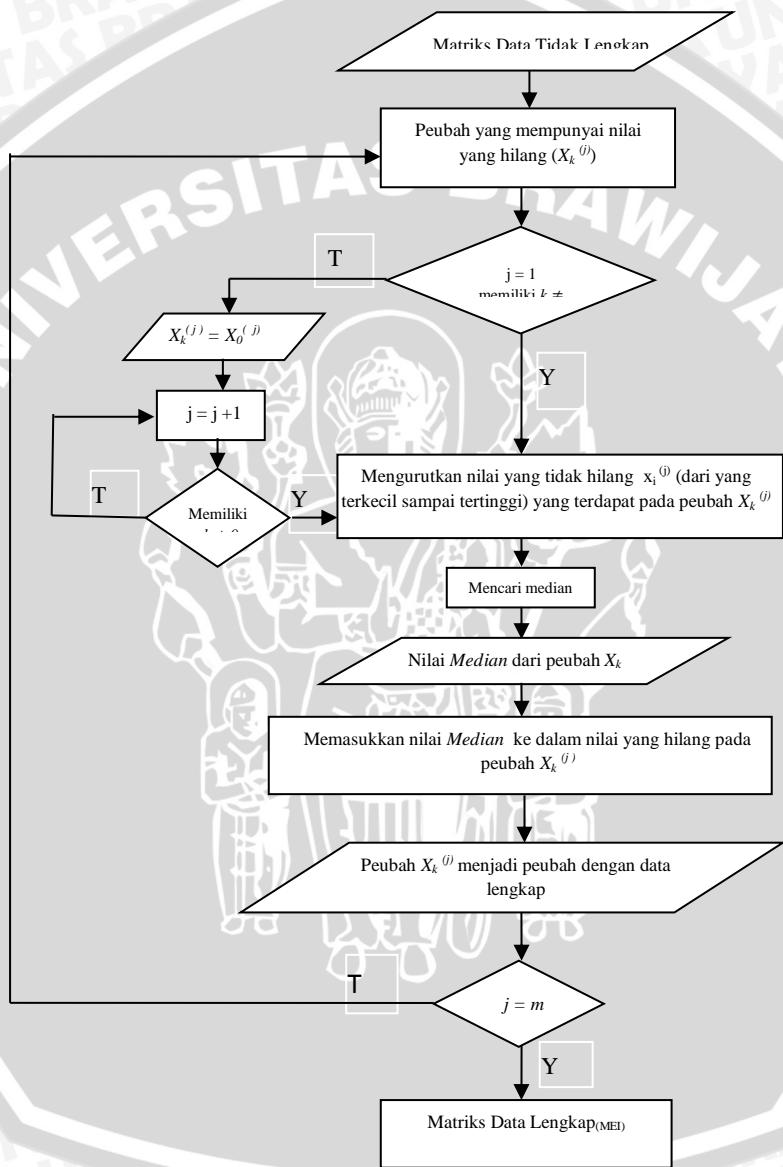
Gambar 3.2. Diagram Alir Analisis Komponen Utama Nonlinier (AKUNL)

### **3.3.3 Median Imputation**

Untuk masing-masing kondisi, langkah-langkah *Median Imputation* dalam penelitian ini dapat dijelaskan sebagai berikut :

1. menyiapkan matriks data tidak lengkap
2. pengumpulan informasi tentang peubah mana yang mempunyai nilai yang hilang ( $X_k^{(j)}$ ), dimana  $j = 1, 2, \dots, m$  ;  $k \neq 0$  (k adalah banyak pengamatan yang hilang)
3.  $j = 1$  memiliki  $k \neq 0$   
jika  $j = 1$  tidak memiliki  $k \neq 0$  maka menjadi  $X_k^{(j)} = X_0^{(j)}$
4.  $j = j + 1$
5. jika tidak ada peubah yang memiliki  $k \neq 0$ , yaitu  $X_k^{(j)} = X_0^{(j)}$ ,  
maka kembali ke Langkah 4
6. mengumpulkan nilai-nilai pada peubah  $X_k^{(j)}$  yang tidak hilang,  
kemudian mengurutkan nilai tersebut dari terkecil sampai  
terbesar
7. mencari median, maka diperoleh nilai median dari peubah  $X_k^{(j)}$
8. memasukkan nilai median dari peubah  $X_k^{(j)}$  ke dalam nilai yang  
hilang pada peubah  $X_k^{(j)}$ , maka peubah  $X_k^{(j)}$  menjadi peubah  
dengan data lengkap
9. Ulangi Langkah 4-8 sampai  $j = m$
10. Pada langkah terakhir peubah  $X_0^{(j)} + X_k^{(j)}$  lengkap dijadikan satu  
sehingga membentuk matriks data lengkap dari metode *Median  
Imputation*  $H_{nxm(MEI)}$

Langkah-langkah *Median Imputation* (MEI) secara sistematis dapat dilihat pada Gambar 3.3



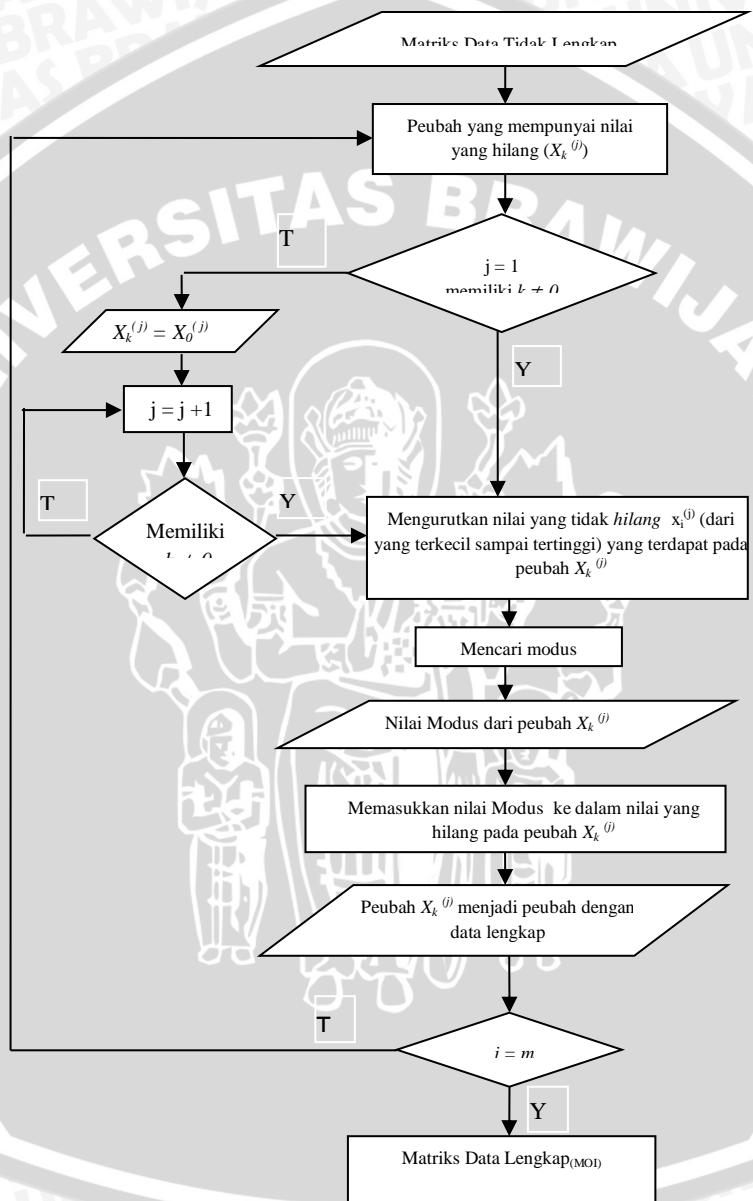
Gambar 3.3. Diagram Alir Metode *Median Imputation*

### **3.3.4 Mode Imputation**

Untuk masing-masing kondisi, langkah-langkah *Mode Imputation* dalam penelitian ini dapat dijelaskan sebagai berikut :

1. menyiapkan matriks data tidak lengkap
2. pengumpulan informasi tentang peubah mana yang mempunyai nilai yang hilang ( $X_k^{(j)}$ ), dimana  $j = 1, 2, \dots, m$  ;  $k \neq 0$  (k adalah banyak pengamatan yang hilang)
3.  $j = 1$  memiliki  $k \neq 0$   
jika  $j = 1$  tidak memiliki  $k \neq 0$  maka menjadi  $X_k^{(j)} = X_0^{(j)}$
4.  $j = j + 1$
5. jika tidak ada peubah yang memiliki  $k \neq 0$ , yaitu  $X_k^{(j)} = X_0^{(j)}$ ,  
maka kembali ke Langkah 4
6. mengumpulkan nilai-nilai pada peubah  $X_k^{(j)}$  yang tidak hilang,  
kemudian mengurutkan nilai tersebut dari terkecil sampai  
terbesar
7. mencari modus, maka diperoleh nilai modus dari peubah  $X_k^{(j)}$
8. memasukkan nilai modus dari peubah  $X_k^{(j)}$  ke dalam nilai yang  
hilang pada peubah  $X_k^{(j)}$ , maka peubah  $X_k^{(j)}$  menjadi peubah  
dengan data lengkap
9. Ulangi Langkah 4-8 sampai  $j = m$
10. Pada langkah terakhir peubah  $X_0^{(j)} + X_k^{(j)}$  lengkap dijadikan satu  
sehingga membentuk matriks data lengkap dari metode *Median  
Imputation*  $H_{nxm(MOI)}$

Langkah-langkah *Mode Imputation* (MOI) secara sistematis dapat dilihat pada Gambar 3.4



Gambar 3.4. Diagram Alir Metode *Mode Imputation*

### 3.3.5. Forward Imputation

Untuk masing-masing kondisi, langkah-langkah *Forward Imputation* dalam penelitian ini adalah di bawah ini.

Misalkan  $H$  adalah matriks data awal dengan dimensi  $n \times m$ , dipengaruhi oleh data hilang dengan pola acak atau MCAR. Prosedur terdiri dari langkah-langkah berikut :

1. Susun kembali matriks  $H$  dalam rangka untuk membagi/menguraikan menjadi  $K+1$  sub-matrik ( $H_k$ ) yang terpisah-pisah. Sub-matriks  $H_k$  mencakup pada baris, hanya objek dengan tepat  $k$  pengamatan yang hilang,  $k = 0, 1, \dots, K < m$ , dimana  $K$  adalah banyak pengatan hilang yang paling besar dalam tiap baris. Beberapa matriks mungkin tidak memiliki elemen yang hilang. Dalam hal ini, tentukan  $H_k = H_{0xm}$ . Submatriks ini dari semua matriks  $H$  yang memiliki jumlah kolom  $m$  yang sama (berasal dari semua kolom yang sama), tetapi berbeda jumlah baris (berasal dari baris yang berbeda) dan sub matriks tersebut mengikuti aturan sebagai berikut :
  - 1a. Matriks  $H_0^{(0)}$  memiliki dimensi  $n_0^{(0)} \times m$  dan semua berisi  $n_0^{(0)}$  objek yang tidak dipengaruhi oleh pengamatan *missing*. Maka matriks ini hanya matriks lengkap di langkah awal yang diperbarui pada setiap langkah.
  - 1b. matriks  $H_k$  : memiliki dimensi  $n_k \times m$  dan mencakup semua objek  $n_k$  dengan  $k$  pengamatan hilang
$$n_0^{(0)} + \sum_{i=1}^K n_k = n$$
2. Tentukan  $k = 0$
3. Lakukan Analisis Komponen Utama Nonlinier pada matriks lengkap  $H_0^{(k)}$  untuk menghitung peubah *loadings*  $a_j^{(k)}$  dan *category quantification*  $G_j(i)y_j^{(k)}$   $\forall j = 1, \dots, m$  dan  $\forall i = 1, \dots, n_0^{(k)}$
4.  $k = k + 1$
5. Jika tidak ada objek yang tepat memiliki  $k$  nilai yang hilang, yaitu  $H_k = H_{0xm}$ , maka kembali ke langkah 4
6. Penerapan metode imputasi *nearest neighbour*. Untuk setiap  $k$  dan setiap baris di  $H_k$ , metode imputasi *nearest neighbor* diimplementasikan dengan penggunaan *loadings* dan *quantification* dari AKUNL yang dilakukan pada matriks lengkap  $H_0^{(k-1)}$ .

Metode *Forward Imputation* memperhitungkan peran peubah yang lain yaitu dengan cara menghitung bobot (*loading*)

tiap peubah. Oleh karena itu, digunakan jarak *weighted Minkowski* seperti pada Persamaan (2.36).

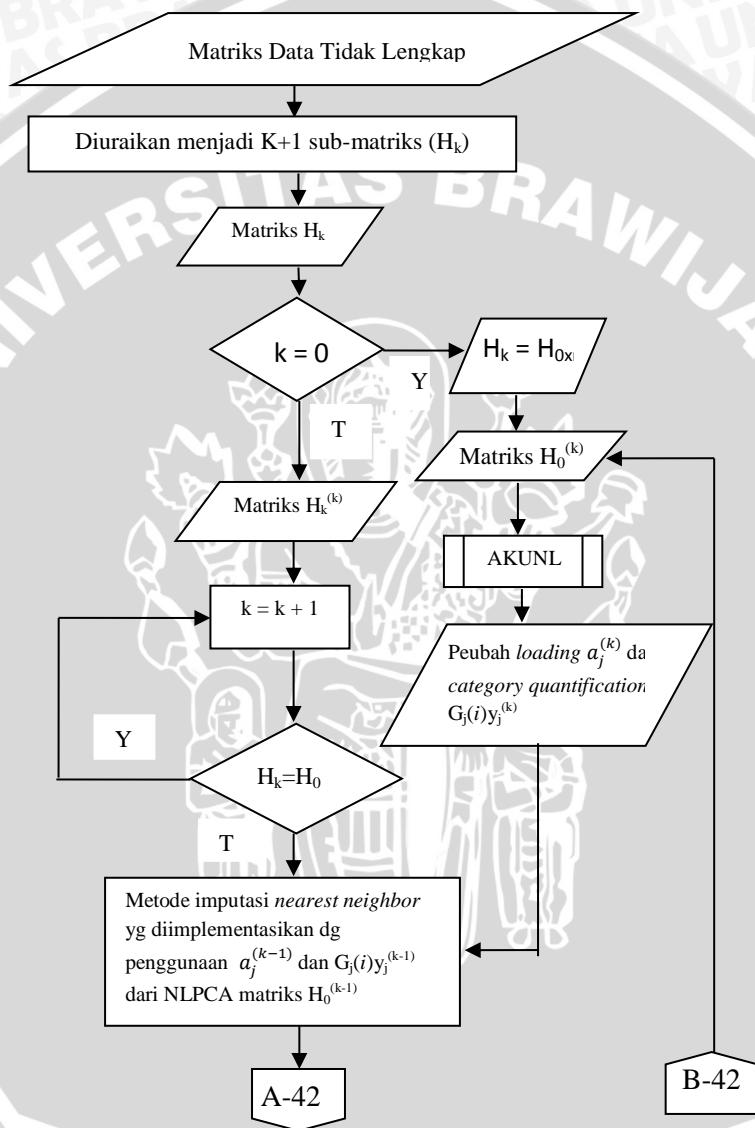
Objek yang memiliki matriks lengkap  $H_0^{(k-1)}$  yang meminimalkan  $d$  disebut  $u_{i^*}^0$ . Kategori yang sesuai diamati pada  $u_{i^*}^0$  dimasukkan pada kategori yang hilang pada  $u_i^{(k)}$ .

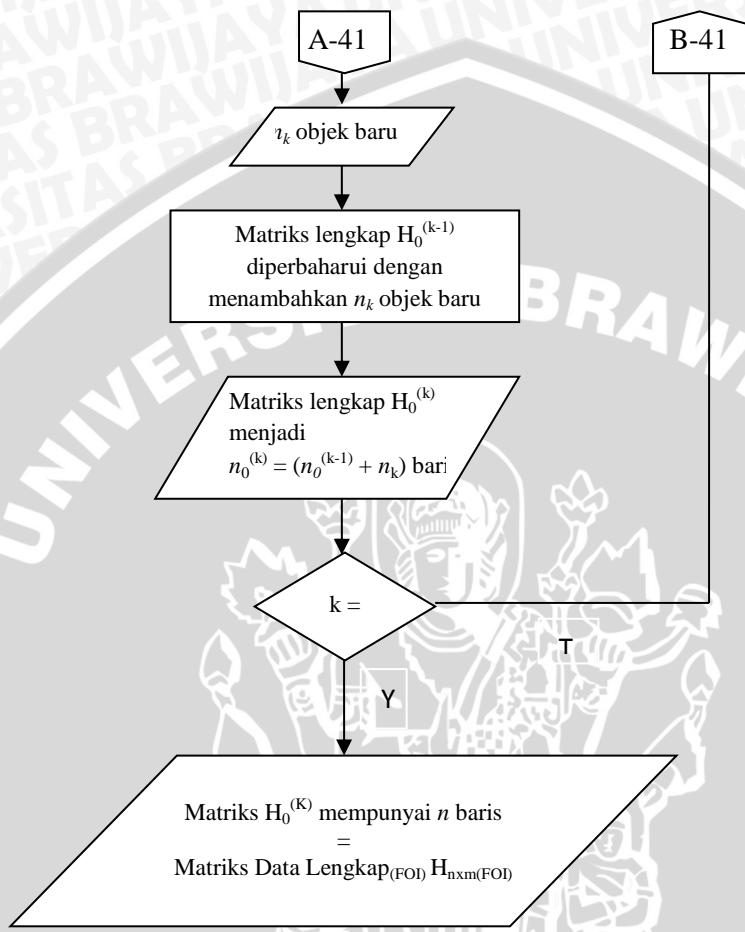
Pada akhir langkah ini, didapatkan  $n_k$  objek baru dengan  $k$  nilai yang telah dihitungkan.

7. Perbarui matriks lengkap  $H_0^{(k-1)}$  dengan menambahkan objek-objek baru  $n_k$  dengan  $k$  nilai yang telah diperhitungkan; matriks  $H_0^{(k)}$  sekarang menjadi  $n_0^{(k)} = (n_0^{(k-1)} + n_k)$  baris
8. Ulangi Langkah 3-7 sampai  $k = K$ . Pada langkah terakhir matriks  $H_0^{(K)}$  mempunyai  $n$  baris
9. Pada langkah terakhir, matriks  $H_0^{(K)}$  mempunyai  $n$  baris, sehingga matriks  $H_0^{(K)}$  menjadi matriks data lengkap dari metode *Forward Imputation*  $H_{nxm(FOI)}$



Langkah-langkah *Forward Imputation* (FOI) secara sistematis dapat dilihat pada Gambar 3.5





Gambar 3.5. Diagram Alir Metode *Forward Imputation*

## **BAB IV**

### **HASIL DAN PEMBAHASAN**

#### **4.1 Data Lengkap**

Data lengkap adalah data awal sebelum dilakukan pembentukan data hilang. Data lengkap yang digunakan dalam penelitian ini sebanyak 10 data sekunder. Sepuluh data tersebut dapat dilihat pada Lampiran 1. Data lengkap tersebut dilakukan Analisis Komponen Utama Nonlinier untuk mendapatkan bobot peubah (*variable loadings*) dan skor komponen objek (*object scores*).

##### **4.1.1 Analisi Komponen Utama Nonlinier pada Data Lengkap**

###### **4.1.1.1 Penentuan Skala Pengukuran**

Langkah awal dalam Analisis Komponen Utama Nonlinier adalah menentukan skala pengukuran untuk setiap peubah. Sepuluh data yang digunakan dalam penelitian ini menggunakan data berskala ordinal. Karena pada seluruh data sudah memiliki kategori, maka tidak perlu dilakukan penentuan kategori. Penentuan kategori hanya dilakukan jika data berskala rasio atau interval.

###### **4.1.1.2 Penentuan Banyak Dimensi**

Pada seluruh data, jumlah peubah untuk setiap data berbeda-beda, jadi jumlah dimensi maksimum untuk setiap data juga berbeda-beda. Jumlah dimensi maksimum yang mungkin terbentuk adalah sama dengan  $m$ , di mana  $m$  adalah jumlah peubah. Untuk menentukan berapa jumlah dimensi yang digunakan, maka terlebih dahulu harus mencari akar ciri dari masing-masing dimensi. Pada data 1 sampai 10, besarnya nilai akar ciri untuk masing-masing dimensi dapat dilihat pada Lampiran 4. Setelah didapatkan akar ciri dari masing-masing dimensi, maka dapat ditentukan berapa jumlah dimensi yang akan digunakan. Jumlah dimensi yang digunakan ditetapkan dengan kriteria nilai akar ciri yang lebih besar dari  $1/m$  di mana  $m$  adalah jumlah peubah. Penentuan jumlah dimensi yang digunakan untuk data 1 sampai 10 dapat dilihat pada Lampiran 5. Setelah ditentukan jumlah dimensi yang digunakan, maka dilakukan iterasi untuk memaksimumkan keragaman kumulatif yang dapat menjelaskan keragaman data. Besarnya akar ciri, proporsi keragaman dan proporsi keragaman kumulatif untuk data 1 sampai 10 dapat dilihat pada Lampiran 6.

Sebagai contoh pada data 1, jumlah peubah adalah 14, jadi jumlah dimensi maksimum yang mungkin terbentuk adalah 14.

Selanjutnya, dari 14 dimensi tersebut, dicari akar ciri dari masing-masing dimensi.

Tabel 4.1 Nilai Akar Ciri Seluruh Dimensi pada Data lengkap (Data 1)

Dimensi	Akar ciri
1	0,5762
2	0,0978
3	0,0719
4	0,0405
5	0,0374
6	0,0338
7	0,0324
8	0,0289
9	0,0235
10	0,0152
11	0,0142
12	0,0126
13	0,0083
14	0,0073

Berdasarkan nilai akar ciri pada Tabel 4.1, maka dapat ditentukan berapa jumlah dimensi yang digunakan untuk melakukan PRINCALS. Jumlah dimensi yang digunakan untuk data 1 yaitu berdasarkan kriteria nilai akar ciri yang lebih besar dari  $1/m = 1/14 = 0,0714$ . Maka pada data 1 diputuskan untuk menggunakan jumlah dimensi sebesar 3 karena nilai akar ciri pada dimensi keempat dan berikutnya kurang dari 0,0714. Setelah ditentukan jumlah dimensi yang digunakan, maka dilakukan iterasi untuk memaksimumkan keragaman kumulatif yang dapat menjelaskan keragaman data. Akar ciri, proporsi keragaman dan proporsi keragaman kumulatif ketiga dimensi untuk data 1 adalah sebagai berikut :

Tabel 4.2 Akar Ciri, Proporsi Keragaman dan Proporsi Keragaman Kumulatif Ketiga Dimensi pada Data 1

Dimensi	Akar Ciri	Proporsi Keragaman (%)	Proporsi Keragaman Kumulatif (%)
1	0,7332	73,32	73,32
2	0,0923	9,23	82,55
3	0,0772	7,72	90,27

Besarnya keragaman kumulatif yang dapat dijelaskan oleh data 1 sebesar 90,27% dari total keragaman. Pada dimensi 1 memiliki akar ciri sebesar 0,7332 yang dapat menerangkan 73,32% dari total keragaman data, pada dimensi ke-2 memiliki akar ciri sebesar 0,0923 yang dapat menerangkan 9,23% dari total keragaman data, pada dimensi ke-3 memiliki akar ciri sebesar 0,0772 yang dapat menerangkan 7,72% dari total keragaman.

#### **4.1.1.3 Penentuan Skor Komponen Objek dan Kuantifikasi Kategori**

Skor komponen objek (*object scores*) digunakan sebagai nilai skor komponen yang dapat digunakan untuk analisis lebih lanjut. Skor komponen objek yang terbentuk berupa skala rasio. Nilai skor komponen objek untuk data 1 sampai 10 dapat dilihat pada Lampiran 7.

Pemberian kuantifikasi atau hitungan kategori setiap peubah dibuat sedemikian hingga peubah dengan kuantifikasi yang baru memiliki korelasi dengan dimensi pertama setinggi mungkin. Nilai kuantifikasi selengkapnya untuk data 1 sampai 10 dapat dilihat pada Lampiran 8.

#### **4.1.1.4 Penentuan Bobot peubah**

Masuknya peubah ke dalam suatu dimensi tergantung pada besarnya bobot peubah yang merupakan nilai korelasi antara peubah yang telah dikuantifikasi dengan dimensi. Peubah dengan nilai bobot peubah terbesar memiliki peran utama dalam dimensi tersebut. Besarnya bobot peubah setiap peubah untuk data 1 sampai 10 dapat dilihat pada Lampiran 9.

#### **4.1.1.5 Tes Konvergensi**

Pada data 1 sampai 10, iterasi maksimum yang digunakan sebanyak 100 iterasi dengan kriteria konvergensi sebesar  $1 \times 10^{-5}$ . Banyak iterasi yang dibutuhkan untuk mencapai nilai yang konvergen pada data 1 sampai 10 dapat dilihat pada Lampiran 10.

Sebagai contoh pada data 1, berikut adalah hasil tes konvergensinya:

Tabel 4.3 Tes Konvergensi Data 1

Iterasi	Total Fit	Total Loss	Multiple loss	Single Loss	Perubahan Iterasi
1	0,7939033	2,2060967	2,0856777	0,1204190	0,0479316
2	0,8393167	2,1606833	1,9992694	0,1614139	0,0454134
3	0,8610940	2,1389060	1,9658401	0,1730659	0,0217773
:	:	:	:	:	:
18	0,9026847	2,0973153	2,0236304	0,0736849	0,0000177
19	0,9026947	2,0973053	,0730435	0,0730435	0,0000100
20	0,9027004	2,0972996	2,0247573	0,0725423	0,0000057

Berdasarkan Tabel 4.3, dapat dilihat bahwa banyak iterasi yang dibutuhkan untuk mencapai nilai yang konvergen adalah 20 iterasi

dengan *total fit* sebesar 0,9027004. *Total fit* sebesar 0,9027004 menunjukkan *fit* yang bagus dan dapat dikatakan bahwa komponen utama yang baru dapat menerangkan 90,27% dari total keragaman data asal.

## 4.2 Data Lengkap Hasil Imputasi

Setelah diperoleh skor komponen objek (*object scores*) dan bobot peubah (*variable loadings*) dari data lengkap, kemudian data tersebut dilakukan pembentukan data hilang pola acak dengan 3 kondisi. Kondisi 1 adalah pembentukan data hilang dengan persentase data hilang 0,5%, kondisi 2 adalah pembentukan data hilang dengan persentase data hilang 1%, dan kondisi 3 adalah pembentukan data hilang dengan persentase 1,5%. Banyaknya pengamatan yang hilang pada setiap kondisi dapat dilihat pada Tabel 3.1, dan Tabel letak dari data hilang yang dihasilkan dapat dilihat pada Lampiran 2.

Setelah terbentuk data hilang, maka dilakukan metode imputasi untuk menduga nilai dari data hilang. Metode imputasi adalah proses pengisian atau penggantian nilai-nilai yang hilang pada dataset dengan nilai-nilai yang mungkin berdasarkan informasi yang didapatkan pada dataset. Metode Imputasi yang digunakan adalah *Median Imputation* (MEI), *Mode Imputation* (MOI) dan *Forward Imputation* (FOI). Penggantian data hilang dengan metode MEI dilakukan dengan cara mengganti nilai-nilai yang hilang dengan nilai median dari data yang tidak hilang. Metode MOI dilakukan dengan cara mengganti nilai-nilai yang hilang dengan nilai modus dari data yang tidak hilang. Metode MEI dan MOI dilakukan secara terpisah untuk setiap peubah. Jika dalam satu peubah terdapat lebih dari 1 nilai yang hilang, maka nilai hasil imputasi pada peubah tersebut akan sama. Sedangkan pada metode FOI tidak dilakukan secara terpisah untuk setiap peubah. Metode FOI mempertimbangkan hubungan pada seluruh peubah.

Setelah dilakukan metode imputasi, maka data tidak lengkap menjadi data lengkap hasil metode imputasi. Hasil pendugaan nilai data hilang tersebut dapat dilihat pada Lampiran 3. Setelah terbentuk data lengkap dari hasil imputasi, kemudian dilakukan Analisis Komponen Utama Nonlinier untuk mendapatkan nilai bobot peubah dan skor komponen objek.

## **4.2.1 Analisis Komponen Utama Nonlinier Data Lengkap Hasil Imputasi**

### **4.2.1.1 Penentuan Skala Pengukuran**

Langkah awal dalam Analisis Komponen Utama Nonlinier adalah menentukan skala pengukuran untuk setiap peubah. Karena jenis data lengkap hasil imputasi ini adalah sama seperti data lengkap yang awal, maka data ini adalah data ordinal.

### **4.2.1.2 Penentuan Banyaknya Dimensi**

Pada ke-90 data lengkap hasil imputasi yang terbentuk, jumlah peubah untuk setiap data berbeda-beda, jadi jumlah dimensi maksimum untuk setiap data juga berbeda-beda. Seperti yang dijelaskan sebelumnya, jumlah dimensi maksimum yang mungkin terbentuk adalah sama dengan jumlah peubah  $m$ . Untuk menentukan berapa jumlah dimensi yang digunakan, maka terlebih dahulu harus mencari akar ciri dari masing-masing dimensi. Pada seluruh data lengkap hasil imputasi, besarnya nilai akar ciri untuk masing-masing dimensi dapat dilihat pada Lampiran 11. Setelah didapatkan akar ciri dari masing-masing dimensi, maka dapat ditentukan berapa jumlah dimensi yang digunakan. Jumlah dimensi yang digunakan ditetapkan dengan kriteria nilai akar ciri yang lebih besar dari  $1/m$  di mana  $m$  adalah jumlah peubah. Penentuan jumlah dimensi yang digunakan untuk setiap data hasil imputasi dapat dilihat pada Lampiran 5. Setelah ditentukan jumlah dimensi yang digunakan, maka dilakukan iterasi untuk memaksimumkan keragaman kumulatif yang dapat menjelaskan keragaman data. Besarnya akar ciri, proporsi keragaman dan proporsi keragaman kumulatif untuk setiap data hasil imputasi dapat dilihat pada Lampiran 12.

Sebagai contoh pada data 1, jumlah peubah adalah 14, jadi jumlah dimensi maksimum yang mungkin terbentuk adalah 14. Selanjutnya, dari 14 dimensi tersebut, dicari akar ciri dari masing-masing dimensi.

Tabel 4.4 Nilai Akar Ciri Seluruh Dimensi pada Data Lengkap Hasil Imputasi (Data 1)

Kondisi 1 (persentase data hilang 0,5%)

Forward Imputation		Mode Imputation		Median Imputation	
Dimensi	Akar ciri	Dimensi	Akar ciri	Dimensi	Akar ciri
1	0,5757	1	0,5757	1	0,5757
2	0,0981	2	0,0981	2	0,0981
3	0,0724	3	0,0724	3	0,0724
4	0,0411	4	0,0411	4	0,0411
5	0,0368	5	0,0368	5	0,0368
6	0,0345	6	0,0345	6	0,0345
7	0,0321	7	0,0321	7	0,0321
8	0,0286	8	0,0286	8	0,0286
9	0,0235	9	0,0235	9	0,0235
10	0,0152	10	0,0152	10	0,0152
11	0,0139	11	0,0139	11	0,0139
12	0,0125	12	0,0125	12	0,0125
13	0,0084	13	0,0084	13	0,0084
14	0,0072	14	0,0072	14	0,0072

Kondisi 2 (persentase data hilang 1%)

Forward Imputation		Mode Imputation		Median Imputation	
Dimensi	Akar ciri	Dimensi	Akar ciri	Dimensi	Akar ciri
1	0,5748	1	0,5602	1	0,5602
2	0,0992	2	0,0991	2	0,0991
3	0,0723	3	0,0699	3	0,0699
4	0,0399	4	0,0467	4	0,0467
5	0,0373	5	0,0407	5	0,0407
6	0,0336	6	0,0364	6	0,0364
7	0,0331	7	0,0334	7	0,0334
8	0,0285	8	0,0287	8	0,0287
9	0,0235	9	0,0231	9	0,0231
10	0,0151	10	0,0166	10	0,0166
11	0,0144	11	0,015	11	0,015
12	0,0132	12	0,0137	12	0,0137
13	0,0087	13	0,0088	13	0,0088
14	0,0066	14	0,0077	14	0,0077

Kondisi 3 (persentase data hilang 1,5%)

Forward Imputation		Mode Imputation		Median Imputation	
Dimensi	Akar ciri	Dimensi	Akar ciri	Dimensi	Akar ciri
1	0,5781	1	0,5774	1	0,5774
2	0,0995	2	0,099	2	0,099
3	0,0701	3	0,07	3	0,07
4	0,0396	4	0,0393	4	0,0393
5	0,0378	5	0,0381	5	0,0381
6	0,0346	6	0,0341	6	0,0341
7	0,0317	7	0,0321	7	0,0321
8	0,028	8	0,0274	8	0,0274
9	0,0238	9	0,0238	9	0,0238
10	0,017	10	0,0172	10	0,0172
11	0,0123	11	0,0134	11	0,0134
12	0,0114	12	0,0114	12	0,0114
13	0,0087	13	0,0087	13	0,0087
14	0,0075	14	0,0081	14	0,0081

Berdasarkan nilai akar ciri pada tabel 4.4, maka dapat ditentukan berapa jumlah dimensi yang digunakan untuk melakukan PRINCALS. Jumlah dimensi yang digunakan untuk data lengkap

hasil imputasi pada data 1 yaitu berdasarkan kriteria nilai akar ciri yang lebih besar dari  $1/m = 1/14 = 0,0714$ .

Pada persentase data hilang sebesar 0,5%, ketika metode yang digunakan untuk mengatasi data hilang adalah FOI, MOI dan MEI diputuskan untuk menggunakan jumlah dimensi sebesar 3 karena nilai akar ciri pada dimensi keempat dan berikutnya kurang dari 0,0714.

Pada persentase data hilang sebesar 1%, ketika metode yang digunakan untuk mengatasi data hilang adalah FOI diputuskan untuk menggunakan jumlah dimensi sebesar 3 karena nilai akar ciri pada dimensi keempat dan berikutnya kurang dari 0,0714. Sedangkan ketika metode yang digunakan adalah MOI dan MEI diputuskan untuk menggunakan jumlah dimensi sebesar 2 karena nilai akar ciri pada dimensi ketiga dan berikutnya kurang dari 0,0714.

Pada persentase data hilang sebesar 1,5%, ketika metode yang digunakan untuk mengatasi data hilang adalah FOI, MOI dan MEI diputuskan untuk menggunakan jumlah dimensi sebesar 2 karena nilai akar ciri pada dimensi ketiga dan berikutnya kurang dari 0,0714.

Pada data lengkap hasil imputasi yang berasal dari data 1, semakin besar persentase data hilang mengakibatkan banyak dimensi yang terbentuk juga semakin berbeda dari banyak dimensi yang terbentuk pada data lengkap. Hal tersebut menunjukkan bahwa perbedaan persentase data hilang akan mempengaruhi hasil analisis.

Setelah ditentukan banyak dimensi yang digunakan, maka dilakukan iterasi untuk memaksimumkan keragaman kumulatif yang dapat menjelaskan keragaman data. Akar ciri, proporsi keragaman dan proporsi keragaman kumulatif masing-masing dimensi yang terbentuk untuk data lengkap hasil imputasi pada data 1 dapat dilihat pada Lampiran 12. Berdasarkan pada lampiran 12, pada persentase data hilang sebesar 0,5%, besarnya keragaman kumulatif yang dapat dijelaskan oleh data lengkap hasil imputasi metode FOI, MOI dan MEI adalah sama, yaitu sebesar 90,28% dari total keragaman. Pada persentase data hilang sebesar 1%, besarnya keragaman kumulatif yang dapat dijelaskan oleh data lengkap hasil imputasi metode FOI adalah sebesar 90,28% dari total keragaman, sedangkan metode MOI dan MEI adalah sebesar 80,18% dari total keragaman. Dan pada persentase data hilang sebesar 1,5%, besarnya keragaman kumulatif yang dapat dijelaskan oleh data lengkap hasil imputasi metode FOI,

MOI dan MEI adalah sama, yaitu sebesar 83,4% dari total keragaman.

#### **4.2.1.3 Penentuan Skor Komponen Objek dan Kuantifikasi Kategori**

Skor komponen objek digunakan sebagai nilai skor komponen yang dapat digunakan untuk analisis lebih lanjut. Nilai skor komponen objek untuk setiap data hasil imputasi dapat dilihat pada Lampiran 13.

Pemberian kuantifikasi atau hitungan kategori setiap peubah dibuat sedemikian hingga peubah dengan kuantifikasi yang baru memiliki korelasi dengan dimensi pertama setinggi mungkin. Nilai kuantifikasi selengkapnya untuk setiap data hasil imputasi dapat dilihat pada Lampiran 14

#### **4.2.1.4 Penentuan Bobot Peubah**

Masuknya peubah ke dalam suatu dimensi tergantung pada besarnya bobot peubah yang merupakan nilai korelasi antara peubah yang telah dikuantifikasi dengan dimensi. Peubah dengan nilai bobot peubah terbesar memiliki peran utama dalam dimensi tersebut. Besarnya bobot peubah setiap peubah dapat dilihat pada Lampiran 15.

#### **4.2.1.5 Tes Konvergensi**

Iterasi maksimum yang digunakan untuk data ke-7 kondisi 2 metode *Forward Imputation* sebanyak 1000, sedangkan untuk data hasil imputasi yang lainnya iterasi maksimum yang digunakan sebanyak 100 iterasi dengan kriteria konvergensi sebesar  $1 \times 10^{-5}$ . Pada data ke-7 kondisi 2 metode *Forward Imputation* digunakan sebanyak 1000 iterasi karena banyak iterasi yang dibutuhkan lebih dari 100. Banyaknya iterasi yang dibutuhkan untuk mencapai nilai yang konvergen dapat dilihat pada Lampiran 16.

### **4.3 Penentuan Metode Imputasi Terbaik**

Setelah diperoleh nilai bobot peubah (*variable loadings*) dan skor komponen objek (*object scores*) dari data lengkap dan data lengkap hasil imputasi, maka dapat dilakukan penentuan metode imputasi yang terbaik dalam mengatasi data hilang.

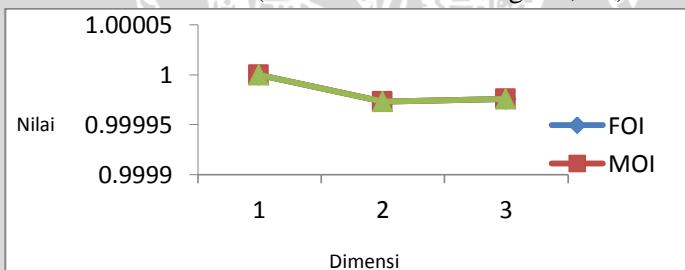
Penentuan metode imputasi terbaik didasarkan pada perbandingan vektor bobot peubah (*variable loadings*)  $\underline{a}$  dan vektor dari skor komponen objek (*object scores*) X.

a) vektor bobot peubah (*variable loadings*)

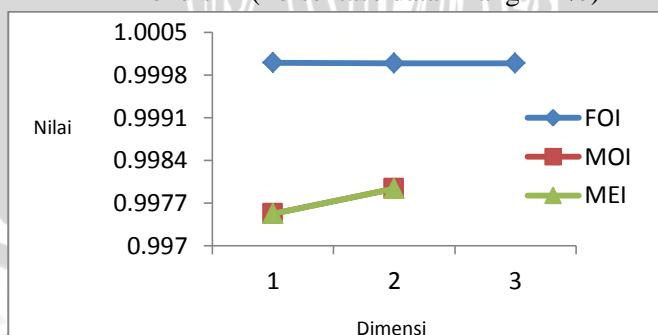
Penentuan metode imputasi terbaik didasarkan pada perbandingan vektor bobot peubah (*variable loadings*), yaitu kosinus antara vektor dari bobot peubah  $\underline{a}_t$  untuk setiap  $t$  metode imputasi dan vektor bobot peubah dari data lengkap yang awal  $\underline{a}_\theta$  dihitung menggunakan Persamaan (2.38). Hasil perhitungan  $\cos(\underline{a}_t, \underline{a}_\theta)$  dari setiap metode imputasi pada setiap kondisi dapat dilihat pada Lampiran 17.

Semakin kecil sudut kosinus, maka nilai  $\cos(\underline{a}_t, \underline{a}_\theta)$  akan mendekati 1, dan hal tersebut berarti semakin dekat kedua vektor. Jika kedua vektor semakin dekat, maka dapat dikatakan bahwa korelasi bobot peubah dari data lengkap hasil metode imputasi dengan bobot peubah dari data lengkap adalah semakin kuat. Jika korelasinya mendekati 1, maka kemampuan metode  $t$  tersebut untuk mengatasi data hilang adalah lebih baik. Untuk melihat lebih jelas metode yang memiliki korelasi paling mendekati 1, maka nilai  $\cos(\underline{a}_t, \underline{a}_\theta)$  pada Lampiran 17 disajikan dalam bentuk diagram garis seperti pada Lampiran 19. Berikut adalah contoh diagram garis untuk data 1 :

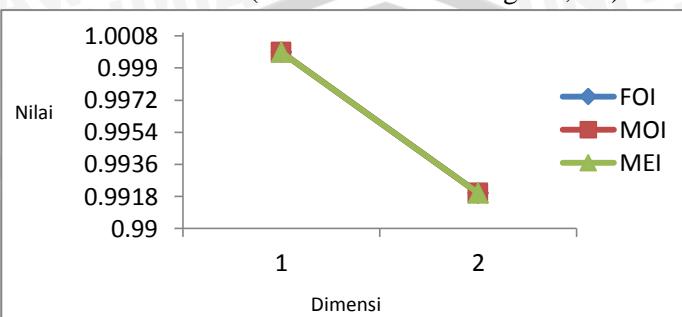
Kondisi 1 (Persentase data hilang = 0,5%)



Kondisi 2 (Persentase data hilang = 1%)



Kondisi 3 (Persentase data hilang = 1,5%)



Gambar 4.1 Diagram Garis Metode Imputasi Paling Baik Berdasarkan Nilai  $\cos(\underline{a}_t, \underline{a}_\theta)$  Terbesar

Berdasarkan gambar 4.1, dapat dilihat bahwa untuk kondisi 1 dan 3, ketiga metode imputasi adalah metode yang lebih baik dalam mengatasi data hilang. Hal tersebut dapat dilihat dari garis yang saling berhimpit. Pada kondisi 2, metode FOI adalah metode yang lebih baik dalam mengatasi data hilang. Hal tersebut dapat dilihat dari garis metode FOI yang paling mendekati nilai 1.

Hasil diagram garis untuk data 2 sampai 10 dapat dilihat pada Lampiran 19. Berdasarkan Lampiran 19, berikut rangkuman hasil metode imputasi yang terbaik untuk setiap kondisi pada setiap data yang disajikan dalam Tabel 4.5.

Tabel 4.5 Metode Imputasi Terbaik Berdasarkan Nilai  $\cos(\underline{a}_t, \underline{a}_\theta)$  Terbesar

Data	Metode Imputasi yang memiliki Nilai Kosinus Paling Mendekati 1		
	Kondisi 1	Kondisi 2	Kondisi 3
1	FOI, MOI, MEI	FOI	FOI, MOI, MEI
2	FOI	FOI	FOI
3	FOI	FOI	FOI
4	FOI	FOI	FOI
5	FOI	FOI	FOI, MOI, MEI
6	FOI	FOI	FOI
7	FOI	FOI	FOI
8	FOI	FOI	FOI
9	FOI	FOI	FOI
10	FOI	FOI, MOI, MEI	FOI

Keterangan :

FOI = Forward Imputation

MOI = Mode Imputation

MEI = Median Imputation

Kondisi 1 = data hilang dengan persentase data hilang 0,5%

Kondisi 2 = data hilang dengan persentase data hilang 1%

Kondisi 3 = data hilang dengan persentase data hilang 1,5%

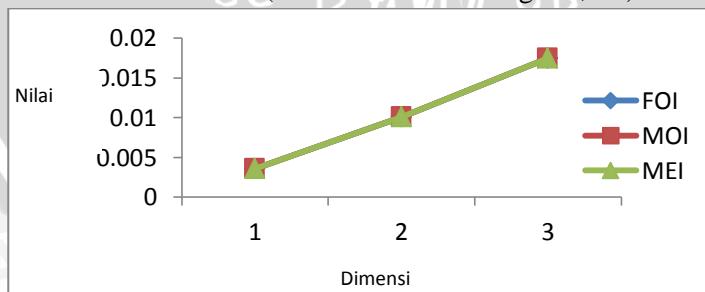
Berdasarkan Tabel 4.5, dapat dilihat bahwa pada kondisi 1, 9 dari 10 data menyatakan bahwa metode *Forward Imputation* (FOI) adalah metode imputasi yang lebih baik dalam mengatasi data hilang. Pada kondisi 2, 9 dari 10 data juga menyatakan bahwa metode *Forward Imputation* (FOI) adalah metode imputasi terbaik dalam mengatasi data hilang. Dan pada kondisi 3, 8 dari 10 data menyatakan bahwa metode *Forward Imputation* (FOI) adalah metode imputasi yang lebih baik dalam mengatasi data hilang. Jadi pada penelitian ini, jika dilihat dari nilai bobot peubah (*variable loadings*) yang dihasilkan, berapapun besarnya persentase data hilang, metode imputasi yang lebih baik dalam mengatasi data hilang adalah metode *Forward Imputation*.

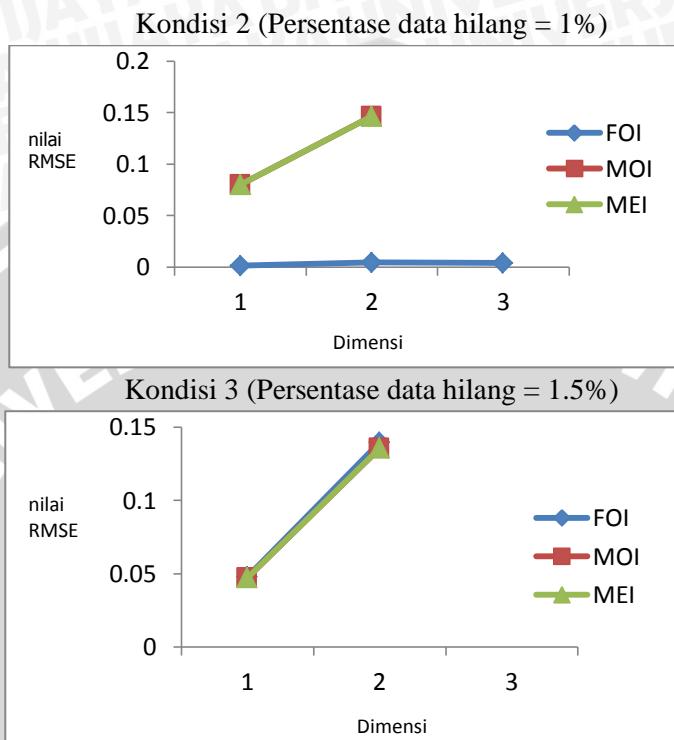
b) Skor Komponen Objek (*object scores*)

Penentuan metode imputasi terbaik juga didasarkan pada perbandingan skor komponen objek. Perbandingan skor komponen objek dapat dihitung dengan persamaan *Root Mean Square Error* (RMSE) pada persamaan (2.39). Hasil perhitungan RMSE<sub>t</sub> dari setiap metode imputasi pada setiap kondisi dapat dilihat pada Lampiran 18.

Semakin kecil nilai RMSE<sub>t</sub> maka kesalahan hasil imputasi semakin kecil, begitu juga sebaliknya. Semakin kecil kesalahan imputasi maka menunjukkan bahwa metode tersebut adalah metode terbaik. Untuk melihat lebih jelas metode mana yang memiliki nilai RMSE<sub>t</sub> paling kecil, maka pada Lampiran 18 disajikan dalam bentuk diagram garis seperti pada Lampiran 20. Berikut adalah contoh diagram garis untuk data 1 :

Kondisi 1 (Persentase data hilang = 0,5%)





Gambar 4.2 Diagram Garis Metode Imputasi Paling Baik Berdasarkan Nilai RMSE<sub>t</sub> paling kecil

Berdasarkan gambar 4.2, dapat dilihat bahwa untuk kondisi 1 dan 3, ketiga metode imputasi adalah metode paling dalam mengatasi data hilang. Hal tersebut dapat dilihat dari garis yang saling berimpit. Pada kondisi 2, metode FOI adalah metode yang lebih baik dalam mengatasi data hilang. Hal tersebut dapat dilihat dari garis metode FOI yang paling mendekati nilai 0.

Hasil diagram garis untuk data 2 sampai 10 dapat dilihat pada Lampiran 20. Berdasarkan Lampiran 20, berikut rangkuman hasil metode imputasi terbaik untuk setiap kondisi pada setiap data yang disajikan dalam Tabel 4.6.

Tabel 4.6. Metode Imputasi Terbaik Berdasarkan Nilai RMSE Terkecil

Data	Metode Imputasi yang memiliki Nilai RMSE paling Kecil		
	Kondisi 1	Kondisi 2	Kondisi 3
1	FOI, MOI, MEI	FOI	FOI, MOI, MEI
2	FOI	FOI	FOI
3	FOI	FOI	FOI
4	FOI	FOI	FOI
5	FOI	FOI	FOI, MOI, MEI
6	FOI	FOI	FOI
7	FOI	FOI	FOI
8	FOI	FOI	FOI
9	FOI	FOI	FOI
10	FOI	FOI, MOI, MEI	FOI

Keterangan :

FOI = *Forward Imputation*

MOI = *Mode Imputation*

MEI = *Median Imputation*

Kondisi 1 = data hilang dengan persentase data hilang 0,5%

Kondisi 2 = data hilang dengan persentase data hilang 1%

Kondisi 3 = data hilang dengan persentase data hilang 1,5%

Berdasarkan Tabel 4.6, dapat dilihat bahwa pada kondisi 1 dan 2, 9 dari 10 data menyatakan bahwa metode *Forward Imputation* (FOI) adalah metode imputasi yang lebih baik dalam mengatasi data hilang. Dan pada kondisi 3, 8 dari 10 data menyatakan bahwa metode *Forward Imputation* (FOI) adalah metode imputasi yang lebih baik dalam mengatasi data hilang. Jadi pada penelitian ini, jika dilihat dari nilai skor komponen objek (*object scores*) yang dihasilkan, berapapun besarnya persentase data hilang, metode imputasi yang lebih baik dalam mengatasi data hilang adalah metode *Forward Imputation*.

#### 4.4 Pengaruh Semakin Besar Persentase Data Hilang Terhadap Bobot Peubah dan Skor Komponen Objek yang Dihasilkan Dari Metode Imputasi

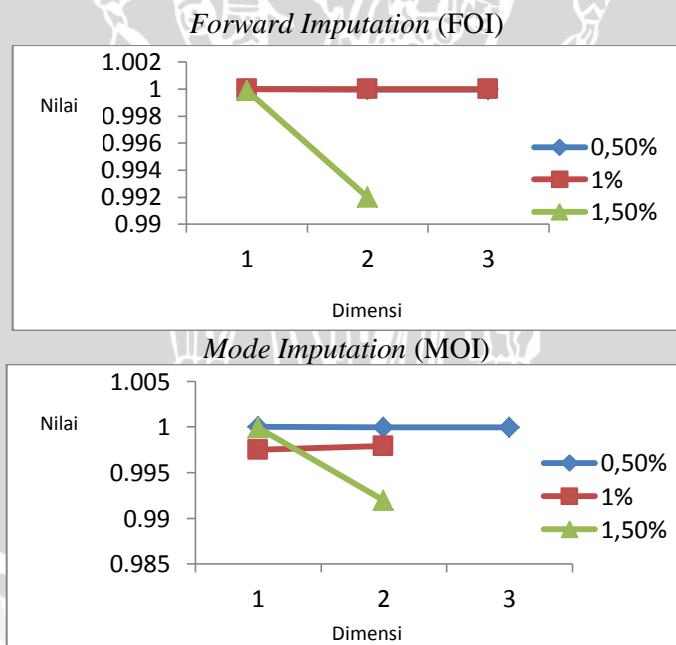
Semakin besar persentase data hilang, maka semakin besar juga kemungkinan kesalahan hasil imputasi. Jika terjadi kesalahan hasil imputasi maka nilai bobot peubah dan skor komponen objek yang dihasilkan akan berbeda dari nilai bobot peubah dan skor komponen objek pada data lengkap. Untuk mengetahui apakah nilai bobot peubah dan skor komponen objek yang dihasilkan dari metode imputasi akan berbeda dengan nilai bobot peubah dan skor

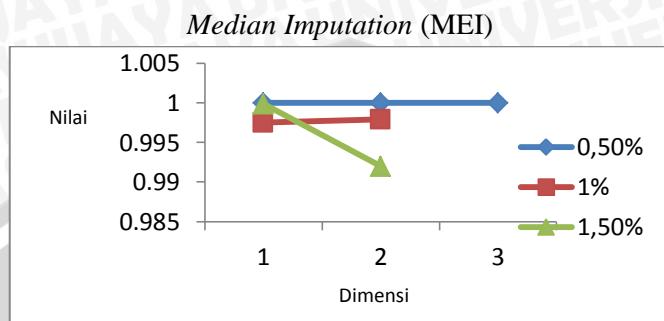
komponen objek yang dihasilkan dari data lengkap, maka dilakukan perhitungan nilai kosinus (korelasi) dan nilai RMSE<sub>t</sub> seperti pada pembahasan 4.3.

Hasil perhitungan nilai kosinus (korelasi) dan nilai RMSE<sub>t</sub> dapat dilihat pada Lampiran 17 dan 18. Jika semakin besar persentase data hilang menghasilkan nilai korelasi yang semakin kecil (menjauhi nilai 1), maka dapat dikatakan bahwa semakin besar persentase data hilang berpengaruh terhadap nilai bobot peubah yang dihasilkan dari metode imputasi. Dan jika semakin besar persentase data hilang menghasilkan nilai RMSE<sub>t</sub> yang semakin besar (menjauhi nilai 0), maka dapat dikatakan bahwa semakin besar persentase data hilang berpengaruh terhadap skor komponen objek yang dihasilkan dari metode imputasi. Jadi, untuk melihat apakah semakin besar persentase data hilang berpengaruh terhadap nilai bobot peubah dan skor komponen objek yang dihasilkan dari metode imputasi, maka perhitungan nilai kosinus dan nilai RMSE<sub>t</sub> pada Lampiran 17 dan 18 disajikan dalam bentuk diagram garis sebagai berikut :

- Berdasarkan nilai Kosinus

DATA 1 :





Gambar 4.3 Diagram Garis Pengaruh Semakin Besar Persentase Data Hilang dalam Menghasilkan Nilai Kosinus

Berdasarkan Gambar 4.3, dapat dilihat bahwa:

1. Diagram garis pada metode FOI menunjukkan bahwa untuk banyak data hilang 0,5% dan 1% garisnya berhimpit. Tetapi jika dilihat berdasarkan Lampiran 17, nilai kosinus (korelasi) untuk banyak data hilang 1% lebih mendekati nilai 1 dari pada untuk data hilang 0,5%. Hal tersebut berarti semakin besar persentase data hilang maka tidak menghasilkan nilai kosinus (korelasi) yang semakin kecil (menjauhi nilai 1)
2. Diagram garis pada metode MOI dan MEI menunjukkan bahwa semakin besar persentase data hilang, garis yang dihasilkan semakin menjauhi nilai 1. Hal tersebut berarti semakin besar persentase data hilang maka menghasilkan nilai kosinus yang semakin kecil (menjauhi nilai 1).

Hasil diagram garis untuk data 2 sampai 10 dapat dilihat pada Lampiran 21. Berikut rangkuman hasil yang disajikan dalam Tabel 4.7 berdasarkan Lampiran 21:

Tabel 4.7 Pengaruh Semakin Besar Persentase Data Hilang dalam Menghasilkan Nilai Kosinus

Data	Pengaruh Perbedaan Tingkat Data Hilang		
	FOI	MOI	MEI
1	Tidak Berpengaruh	Berpengaruh	Berpengaruh
2	Berpengaruh	Berpengaruh	Berpengaruh
3	Tidak Berpengaruh	Tidak Berpengaruh	Tidak Berpengaruh
4	Tidak Berpengaruh	Tidak Berpengaruh	Tidak Berpengaruh
5	Berpengaruh	Berpengaruh	Berpengaruh
6	Berpengaruh	Berpengaruh	Berpengaruh
7	Berpengaruh	Tidak Berpengaruh	Tidak Berpengaruh
8	Tidak Berpengaruh	Tidak Berpengaruh	Tidak Berpengaruh
9	Berpengaruh	Berpengaruh	Berpengaruh
10	Tidak Berpengaruh	Tidak Berpengaruh	Tidak Berpengaruh

Keterangan :

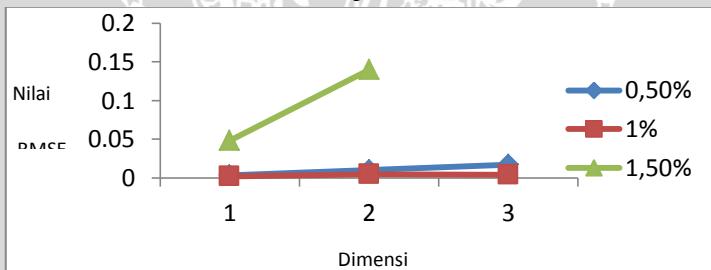
FOI	= <i>Forward Imputation</i>
MOI	= <i>Mode Imputation</i>
MEI	= <i>Median Imputation</i>
Tidak Berpengaruh	Semakin besar persentase data hilang, maka nilai kosinus yang dihasilkan tidak semakin kecil
Berpengaruh	Semakin besar persentase data hilang, maka nilai kosinus yang dihasilkan semakin kecil

Berdasarkan Tabel 4.7, dapat dilihat bahwa untuk metode *Forward Imputation* (FOI), *Mode Imputation* (MOI) dan *Median Imputation* (MEI) hanya 5 dari 10 data yang menyatakan bahwa semakin besar persentase data hilang maka nilai kosinus yang dihasilkan semakin kecil (menjauhi 1). Hal tersebut berarti, semakin besar persentase data hilang maka nilai bobot peubah yang dihasilkan dari metode imputasi semakin berbeda dari nilai bobot peubah pada data lengkap.

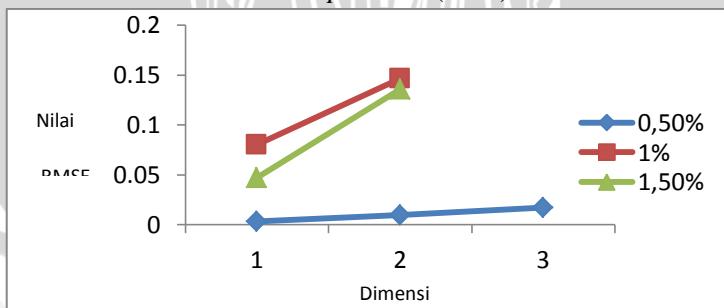
- b) Berdasarkan nilai RMSE<sub>t</sub>

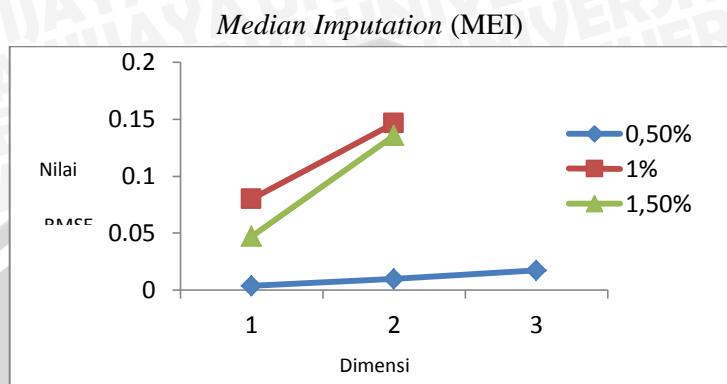
DATA 1 :

*Forward Imputation (FOI)*



*Mode Imputation (MOI)*





Gambar 4.4 Diagram Garis Pengaruh Semakin Besar Persentase Data Hilang dalam Menghasilkan Nilai RMSE<sub>t</sub>

Berdasarkan Gambar 4.4, dapat dilihat bahwa:

1. Diagram garis pada metode FOI menunjukkan bahwa garis pada banyak data hilang 1% lebih mendekati nilai 0 dari pada garis pada banyak data hilang 0,5%. Hal tersebut berarti semakin besar persentase data hilang maka tidak menghasilkan nilai RMSE<sub>t</sub> yang semakin besar (menjauhi nilai 0)
2. Diagram garis pada metode MOI dan MEI menunjukkan bahwa garis pada banyak data hilang 1,5% lebih mendekati nilai 0 dari pada garis pada banyak data hilang 1%. Hal tersebut berarti semakin besar persentase data hilang maka tidak menghasilkan nilai RMSE<sub>t</sub> yang semakin besar (menjauhi nilai 0).

Hasil diagram garis untuk data 2 sampai 10 dapat dilihat pada Lampiran 22. Berikut rangkuman hasil yang disajikan dalam Tabel 4.8 berdasarkan Lampiran 22 :

Tabel 4.8 Pengaruh Semakin Besar Persentase Data Hilang dalam Menghasilkan nilai RMSE<sub>t</sub>

Data	Pengaruh Perbedaan Tingkat Data Hilang		
	FOI	MOI	MEI
1	Tidak Berpengaruh	Tidak Berpengaruh	Tidak Berpengaruh
2	Berpengaruh	Berpengaruh	Berpengaruh
3	Tidak Berpengaruh	Tidak Berpengaruh	Tidak Berpengaruh
4	Tidak Berpengaruh	Tidak Berpengaruh	Tidak Berpengaruh
5	Berpengaruh	Berpengaruh	Berpengaruh
6	Berpengaruh	Berpengaruh	Berpengaruh
7	Berpengaruh	Tidak Berpengaruh	Tidak Berpengaruh
8	Tidak Berpengaruh	Tidak Berpengaruh	Tidak Berpengaruh
9	Berpengaruh	Berpengaruh	Berpengaruh
10	Tidak Berpengaruh	Tidak Berpengaruh	Tidak Berpengaruh

Keterangan :

FOI = *Forward Imputation*

MOI = *Mode Imputation*

MEI = *Median Imputation*

Tidak Berpengaruh	Semakin besar persentase data hilang, maka nilai $\text{RMSE}_t$ yang dihasilkan tidak semakin besar
Berpengaruh	Semakin besar persentase data hilang, maka nilai $\text{RMSE}_t$ yang dihasilkan semakin besar

Berdasarkan Tabel 4.8 dapat dilihat bahwa untuk metode *Forward Imputation* (FOI) 5 dari 10 data menyatakan bahwa semakin besar persentase data hilang maka nilai RMSE yang dihasilkan juga semakin besar. Sedangkan untuk metode *Mode Imputation* (MOI) dan *Median Imputation* (MEI) 4 dari 10 data menyatakan bahwa semakin besar persentase data hilang maka nilai RMSE yang dihasilkan juga semakin besar. Hal tersebut berarti, semakin besar persentase data hilang maka skor komponen objek (*object scores*) yang dihasilkan dari metode imputasi semakin berbeda dari skor komponen objek (*object scores*) pada data lengkap.

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

1. Berdasarkan nilai bobot peubah (*variable loadings*) dan skor komponen objek (*object scores*) yang dihasilkan dari metode imputasi, berapapun besarnya persentase data hilang, metode imputasi yang lebih baik dalam mengatasi data hilang pada Analisis Komponen Utama Nonlinier data ordinal adalah *Forward Imputation* (FOI).
2. Berdasarkan hasil penelitian ini, belum dapat ditunjukkan bahwa semakin besar persentase data hilang maka nilai bobot peubah (*variable loadings*) yang dihasilkan dari metode imputasi juga semakin berbeda dari nilai bobot peubah (*variable loadings*) pada data lengkap. Hal tersebut dikarenakan hanya 5 dari 10 data yang mendukung pernyataan tersebut.
3. Berdasarkan hasil penelitian ini, belum dapat ditunjukkan bahwa semakin besar persentase data hilang maka skor komponen objek (*object scores*) yang dihasilkan dari metode imputasi juga semakin berbeda dari skor komponen objek (*object scores*) pada data lengkap. Hal tersebut dikarenakan hanya 4 dari 10 data yang mendukung pernyataan tersebut.

#### 5.2 Saran

Penelitian selanjutnya disarankan untuk :

1. Memperbesar selang persentase data hilang yang dibentuk untuk melihat apakah semakin besar persentase data hilang, maka nilai bobot peubah dan skor komponen objek yang dihasilkan dari metode imputasi juga semakin berbeda dari nilai bobot peubah dan skor komponen objek pada data lengkap.
2. Mencari tahu apakah metode imputasi menghasilkan bobot peubah dan skor komponen objek yang sama pada data lengkap ketika jenis data hilang yang diberikan adalah tidak acak

UNIVERSITAS BRAWIJAYA



## DAFTAR PUSTAKA

- Acuna, E. and Rodriguez, C. 2003. *The Treatment of Missing Values and its Effect in the Classifier Accuracy.* <http://academic.uprm.edu/eacuna/IFCS04r.pdf>. Diakses tanggal 20 Januari 2013.
- Badan Pusat Statistik Kabupaten Lamongan. 2012. *Pendapat Responden Tentang Pelayanan Publik pada Instansi A.* BPS Kabupaten Lamongan. Lamongan.
- Brimberg, J. 1989. *Properties of Distance Function and Minisum Location Models.* Open Access Dissertations and Theses. <http://digitalcommons.mcmaster.ca/cgi/viewcontent.cgi?article=4637&context=opendissertations>. Diakses tanggal 10 April 2013.
- Clark, C.T. and Schkade, L.L. 1983. *Statistical Analysis for Administrative Decisions.* South-Western Publishing Co. Cincinnati, Ohio.
- Croninger, R.G., and Douglas, K.M. 2005. *Missing Data and institutional Research.* New directions for institutional research. San Francisco. Chapter 3, pp.33-50
- Daniel, W. W. 1989. *Statistika Nonparametrik Terapan.* PT Gramedia, Jakarta.
- Ferrari, P. A., Annoni, P., Barbiero, A., and Manzi, G. 2011. *An Imputation Method for Categorical Variables With Application to Nonlinear Principal Component Analysis.* *Computational Statistics & Data Analysis.* Computational Statistic and Data Analysis 55, hal. 2410-2420.
- Fitrida, M.R. 2009. *Analisis Tingkat Kepuasan dan Persepsi Konsumen terhadap Produk Teh Cup Tong Tji (studi kasus di Malang Town Square dan Mal Olympic Garden Malang).*

Skripsi Fakultas Pertanian Universitas brawijaya. Malang (tidak dipublikasikan).

- Gifi, A. 1983. *PRINCALS User's Guide*. Departement of Datatheory, Faculty of Social Sciences, University of Leiden, the Netherlands.
- \_\_\_\_\_. 1989. *Algorithm Descriptions for ANACOR, HOMALS, PRINCALS, and OVERALS*. [http://www.datatheory.nl/pdfs/89/89\\_01.pdf](http://www.datatheory.nl/pdfs/89/89_01.pdf). Diakses tanggal 10 April 2013
- \_\_\_\_\_. 1990. *Nonlinear Multivariate Analysis*. John Wiley & Sons Inc. New York.
- Gumus, M., Hamarat, B. 2004. *Knowledge Management Perceptions of Managers. Journal of Knowledge Management Pratice*. <http://www.tlanc.com/articl64.html>. Diakses tanggal 12 Februari 2013.
- Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. 2010. *Multivariate Data Analysis*. Pearson Education, Inc., Upper Saddle River, New Jersey.
- Hapsari, S.K. 2010. *Analisis Persepsi Konsumen terhadap Jus Buah Dalam Kemasan (studi kasus di Giant-Hypermarket Mall Olympic Garden Malang)*. Skripsi Fakultas Pertanian Universitas brawijaya. Malang (tidak dipublikasikan).
- Ingrasia, S., Rocci,R., and Vichi, M. 2011. *New Perspectives in Statistical Modelling and Data Analysis*. Springer. New York.
- Konig, R. 2002. *On the Rotation of Non-Linear Principal Components Analysis (PRINCALS) Solution : Description of A Procedure*. ZUMA-Nachrichten 50, Hal. 114-120.
- Irawati, E.. 2002. *Persepsi Konsumen terhadap Bauran Pemasaran Rokok Starmild 16 pada Perusahaan Rokok PT. Bentoel Prima Malang*. Skripsi Fakultas Ekonomi Universitas brawijaya. Malang (tidak dipublikasikan).

- Johnson, R.A. and Wichern, D. W. 2002. *Applied Multivariate Statistical Analysis fifth edition*. Prentice-hall, Inc. New Jersey.
- Kroonenberg, P. M., Harch, B. D., Basford, K. E., and Cruickshank, A. 1997. *Combined Analysis of Categorical and numerical Descriptors of Australian Groundnut Accession Using Nonlinear Principle Component Analysis*. *Journal of Agricultural, Biological, and Environmental Statistic*, Vol. 2, Hal. 294-312.
- Li, B., Chang, E., and Wu, Y. 2003. *Discovery of A Perceptual Distance Function for Measuring Image Similarity*. <http://infolab.stanford.edu/~echang/dpf-ext.pdf>. Diakses tanggal 4 April 2013.
- Maharani, S. 2007. *Persepsi Konsumen terhadap Kualitas Bakpao Telo dengan Metode Importance-performance Analysis*. Skripsi Fakultas Teknologi Pertanian Universitas brawijaya. Malang (tidak dipublikasikan).
- Munro, B.H., 2005. *Statistical Methods for Health Care Research, Fifth Edition*. Lippincott Williams & Wilkins, America
- Nova, M. dan Mukid, M.A. 2011. *Pendugaan Data Hilang dengan Menggunakan Data Augmentation*. Media Statistika, Vol. 4, No. 2, Hal. 73-87.
- Nugraha, R.S. 2011. *Pengukuran Kualitas Pelayanan Berdasarkan persepsi Konsumen terhadap Hotel Ollino Garden Malang*. Skripsi Fakultas Ekonomi Universitas brawijaya. Malang (tidak dipublikasikan).
- Putri, H.Y. 2008. *Analisis Kualitas Pelayanan terhadap Tingkat Kepuasan Konsumen (studi kasus di restoran Boyong Kalegan Yogyakarta)*. Skripsi Fakultas Teknologi pertanian Universitas brawijaya. Malang (tidak dipublikasikan).
- Sanders, D. H. And Smidt, R. K. 2000. *Statistical a First Course*. Edisi ke-6. McGraw-Hill, Boston.

- Sudjana. 2002. *Metoda Statistika*, edisi 6. PT Tarsito. Bandung.
- Timm, N.H. 1975. *Multivariate Analysis with Application in Education and Psychology*. Brooks/Cole Publishing Company. California.
- Tohari, A. 2005. *Penerapan Multivariate dalam Menentukan Segmentasi Konsumen Plasa Araya dan Positioning terhadap Pesaingnya*. Skripsi Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas brawijaya. Malang (tidak dipublikasikan).
- Wahyuni, A.D. 2008. *Persepsi Konsumen terhadap Apel Lokal dan Apel Impor*. Skripsi Fakultas Pertanian Universitas brawijaya. Malang (tidak dipublikasikan).
- Wahyuningsih, I. 2012. *Analisis Persepsi Konsumen terhadap Puree Mangga Podang dengan Metode Importance Performance Analysis (studi kasus pada UKM Sari Buah dan Dodol di Kota Batu)*. Skripsi Fakultas Teknologi Pertanian Universitas brawijaya. Malang (tidak dipublikasikan).
- Wasito, I. And Mirkin, B. 2005. *Nearest Neighbour Approach in the Least-Squares data Imputation Algorithms*. Information Sciences 169, Hal. 1-25.

## LAMPIRAN

### Lampiran 1. Data Lengkap

#### Data 1

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X 10	X 11	X 12	X 13	X 14
1	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	3	3	3	3	3	3	3	3	3	3	3	3	3	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
97	3	3	3	3	3	3	3	3	3	3	3	3	3	3
98	3	3	3	3	3	3	2	2	2	2	2	2	3	3
99	3	3	3	3	3	3	3	3	3	3	3	3	3	3
100	3	2	2	2	2	2	3	3	3	3	3	3	3	3

Keterangan :

- X<sub>1</sub> : Kemudahan prosedur pelayanan
- X<sub>2</sub> : Kesesuaian persyaratan pelayanan dengan jenis pelayanannya
- X<sub>3</sub> : Kejelasan dan kepastian petugas
- X<sub>4</sub> : Kedisiplinan petugas dalam memberikan pelayanan
- X<sub>5</sub> : Tanggungjawab petugas dalam memberikan pelayanan
- X<sub>6</sub> : Kemampuan petugas dalam memberikan pelayanan
- X<sub>7</sub> : Kecepatan pelayanan
- X<sub>8</sub> : Keadilan untuk mendapatkan pelayanan
- X<sub>9</sub> : Kesopanan dan keramahan petugas
- X<sub>10</sub> : Kewajaran biaya untuk mendapatkan pelayanan
- X<sub>11</sub> : Kesesuaian biaya yang dibayar dengan biaya yang ditetapkan
- X<sub>12</sub> : Ketetapan pelaksanaan pelayanan terhadap jadwal waktu pelayanan yang telah ditetapkan
- X<sub>13</sub> : Kenyamanan di lingkungan unit pelayanan
- X<sub>14</sub> : Keamanan pelayanan di lingkungan unit pelayanan

## Lampiran 1 (Lanjutan)

Kriteria penilaian :

Peubah	Kriteria Penilaian			
	1	2	3	4
X <sub>1</sub>	Tidak mudah	Kurang mudah	Mudah	Sangat Mudah
X <sub>2</sub>	Tidak sesuai	Kurang sesuai	Sesuai	Sangat sesuai
X <sub>3</sub>	Tidak jelas	Kurang jelas	Jelas	Sangat jelas
X <sub>4</sub>	Tidak disiplin	Kurang disiplin	Disiplin	Sangat disiplin
X <sub>5</sub>	Tidak bertanggung jawab	Kurang bertanggung jawab	Bertanggung jawab	Sangat bertanggung jawab
X <sub>6</sub>	Tidak mampu	Kurang mampu	Mampu	Sangat mampu
X <sub>7</sub>	Tidak cepat	Kurang cepat	Cepat	Sangat cepat
X <sub>8</sub>	Tidak adil	Kurang adil	Adil	Sangat adil
X <sub>9</sub>	Tidak sopan & tidak ramah	Kurang sopan & kurang ramah	Sopan & ramah	Sangat sopan dan ramah
X <sub>10</sub>	Tidak wajar	Kurang wajar	Wajar	Sangat wajar
X <sub>11</sub>	Selalu tidak sesuai	Kadang-kadang sesuai	Lebih sering sesuai	Selalu sesuai
X <sub>12</sub>	Selalu tidak tepat	Kadang-kadang tepat	Lebih sering tepat	Selalu tepat
X <sub>13</sub>	Tidak nyaman	Kurang nyaman	Nyaman	Sangat nyaman
X <sub>14</sub>	Tidak aman	Kurang aman	Aman	Sangat aman

## Data 2

No	X 1	X 2	X 3	X 4	X 5	X 6	X 7	X 8	X 9	X 10	X 11	X 12	X 13	X 14	X 15	X 16	X 17	X 18	X 19	X 20	X 21	X 22	X 23
1	5	5	4	4	4	3	3	3	3	4	4	4	3	3	5	5	4	4	3	4	4	4	4
2	5	5	4	4	4	4	5	4	3	3	4	4	4	4	5	5	4	4	3	4	4	4	4
3	5	5	5	4	4	4	4	5	4	5	4	4	4	4	4	5	5	5	3	3	3	3	3
4	5	5	4	4	4	4	3	3	4	4	4	4	4	4	4	5	5	4	4	4	4	4	4
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
59	5	4	4	4	4	3	3	4	3	3	3	3	3	3	3	5	5	3	3	3	3	3	3
70	5	5	5	5	4	4	4	3	3	3	3	3	3	4	5	5	4	4	3	3	3	4	4
71	5	5	5	4	5	4	4	4	3	3	3	3	5	5	5	5	3	3	3	3	3	3	3
72	5	4	4	5	3	4	4	4	3	5	4	3	3	3	5	5	3	3	3	3	3	3	3

Keterangan :

- X<sub>1</sub> : keragaman variasi menu
- X<sub>2</sub> : penampilan hidangan makanan dan minuman
- X<sub>3</sub> : ketersediaan hiburan musik live, rakit dan sarana bermain
- X<sub>4</sub> : ruangan dan bangunan serta area parkir yang bagus
- X<sub>5</sub> : kebersihan dan kerapian penampilan karyawan
- X<sub>6</sub> : kecepatan dalam pelayanan dan penyajian
- X<sub>7</sub> : ketepatan dalam melayani pesanan dan penyajian
- X<sub>8</sub> : ketepatan penghitungan nota
- X<sub>9</sub> : kemudahan pembayaran
- X<sub>10</sub> : ketanggungan saat pelanggan membutuhkan bantuan

## Lampiran 1 (Lanjutan)

- X<sub>11</sub> : karyawan selalu menjaga kelengkapan tissue dan tusuk gigi di meja
- X<sub>12</sub> : kemampuan merespon dengan cepat keluhan serta saran dari pelanggan
- X<sub>13</sub> : kemampuan dan tanggapan dalam membantu memilih pesan
- X<sub>14</sub> : karyawan segera meminta maaf apabila terjadi kesalahan
- X<sub>15</sub> : citra rasa produk
- X<sub>16</sub> : kesegaran produk
- X<sub>17</sub> : pengetahuan karyawan memberikan informasi tentang menu yang tersedia
- X<sub>18</sub> : kesabaran dan ketelatenan karyawan dalam melayani pelanggan
- X<sub>19</sub> : lingkungan yang aman dan nyaman
- X<sub>20</sub> : keramahan karyawan
- X<sub>21</sub> : intonasi dan emosi karyawan baik dalam berkomunikasi
- X<sub>22</sub> : perlakuan yang sama dalam melayani pelanggan
- X<sub>23</sub> : perhatian terhadap keluhan pelanggan

Kriteria penilaian :

- 1 = sangat tidak setuju  
2 = tidak setuju  
3 = cukup setuju  
4 = setuju  
5 = sangat setuju

## Data 3

No	X1	X2	X3	X4	X5	X6	X7	X8	X9
1	5	4	5	5	5	5	4	5	4
2	4	3	2	4	3	5	4	3	4
3	4	4	4	4	3	4	4	4	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
46	5	4	4	5	4	5	4	4	4
47	4	4	5	5	2	5	4	4	3
48	4	4	3	3	4	5	4	5	4
49	4	4	5	5	5	5	5	5	3

Keterangan :

X<sub>1</sub> : Rasa

X<sub>2</sub> : Aroma

X<sub>3</sub> : Keterkenalan merk “Buavita”

X<sub>4</sub> : Label halal

## Lampiran 1 (Lanjutan)

- X<sub>5</sub> : komposisi  
X<sub>6</sub> : Tanggal kadaluarsa  
X<sub>7</sub> : Kemasan  
X<sub>8</sub> : Pilihan Rasa Buah  
X<sub>9</sub> : Manfaat/khasiat

Kriteria penilaian :  
1 = tidak baik  
2 = kurang baik  
3 = cukup baik  
4 = baik  
5 = sangat baik

## Data 4

N o	X1	X2	X3	X4	X5	X6	X7	X8	X9	X 10	X 11	X 12	X 13	X 14
1	2	5	5	4	4	4	5	3	5	5	5	5	5	3
2	3	5	5	2	5	3	5	2	2	5	3	4	3	3
3	1	4	3	1	3	4	3	4	1	5	4	3	5	3
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
54	4	2	5	1	3	4	5	3	1	5	1	2	4	4
55	4	4	4	1	2	1	5	4	2	5	2	2	1	1
56	3	3	4	4	5	4	4	3	2	5	2	3	4	4

Keterangan :

X<sub>1</sub> : Perhatian utama terhadap merk

X<sub>2</sub> : Daya tarik kemasan

X<sub>3</sub> : Kualitas

X<sub>4</sub> : Tingkat kecocokan Rasa

X<sub>5</sub> : Perbandingan harga

X<sub>6</sub> : Perubahan harga

X<sub>7</sub> : Kesesuaian harga

X<sub>8</sub> : Persepsi harga

X<sub>9</sub> : Daya tarik iklan

X<sub>10</sub> : Personal selling

X<sub>11</sub> : Sponsorship

X<sub>12</sub> : Sales promotion

X<sub>13</sub> : Intensitas pengecer

X<sub>14</sub> : Kemudahan tempat

## Lampiran 1 (Lanjutan)

Kriteria Penilaian :

Peubah	Kriteria Penilaian				
	1	2	3	4	5
X <sub>1</sub>	Tidak menjadi perhatian	Kurang menjadi perhatian	Cukup menjadi perhatian	Menjadi perhatian	Sangat menjadi perhatian
X <sub>2</sub>	Tidak menarik	Kurang menarik	Cukup menarik	Menarik	Sangat menarik
X <sub>3</sub>	Tidak dipengaruhi	Kurang dipengaruhi	Cukup dipengaruhi	Dipengaruhi	Sangat dipengaruhi
X <sub>4</sub>	Tidak cocok	Kurang cocok	Agak cocok	Cocok	Sangat cocok
X <sub>5</sub>	Tidak penting	Kurang penting	Cukup penting	Penting	Sangat penting
X <sub>6</sub>	Tidak terpengaruh	Kurang terpengaruh	Cukup terpengaruh	Terpengaruh	Sangat terpengaruh
X <sub>7</sub>	Tidak memperhatikan	Kurang memperhatikan	Cukup memperhatikan	Memperhatikan	Sangat memperhatikan
X <sub>8</sub>	Tidak setuju	Kurang setuju	Cukup setuju	Setuju	Sangat setuju
X <sub>9</sub>	Tidak mempengaruhi	Kurang mempengaruhi	Cukup mempengaruhi	Mempengaruhi	Sangat mempengaruhi
X <sub>10</sub>	Tidak mempengaruhi	Kurang mempengaruhi	Cukup mempengaruhi	Mempengaruhi	Sangat mempengaruhi
X <sub>11</sub>	Tidak tertarik	Kurang tertarik	Cukup tertarik	Tertarik	Sangat tertarik
X <sub>12</sub>	Tidak tertarik	Kurang tertarik	Cukup tertarik	Tertarik	Sangat tertarik
X <sub>13</sub>	Tidak mempengaruhi	Kurang mempengaruhi	Cukup mempengaruhi	Mempengaruhi	Sangat mempengaruhi
X <sub>14</sub>	Tidak mempengaruhi	Kurang mempengaruhi	Cukup mempengaruhi	Mempengaruhi	Sangat mempengaruhi

## Data 5

N o	X 1	X 2	X 3	X 4	X 5	X 6	X 7	X 8	X 9	X 10	X 11	X 12	X 13	X 14	X 15	X 16	X 17
1	3	3	3	2	3	3	3	3	3	2	4	4	3	2	3	4	4
2	4	3	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5
3	4	3	3	4	4	4	3	4	3	3	3	4	3	3	2	3	2
4	3	4	3	3	3	4	3	3	3	3	2	3	3	2	4	3	1
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
100	3	3	3	5	3	3	4	2	4	3	2	3	2	2	1	4	4
101	4	3	3	4	3	5	5	5	5	4	3	3	3	2	4	4	3
102	3	5	4	4	4	3	4	4	5	4	3	4	4	2	2	4	2
103	5	5	5	5	5	1	1	1	1	1	1	1	1	1	1	1	1

Keterangan :

X<sub>1</sub> : kesesuaian harga dengan merk dan kualitas

X<sub>2</sub> : kelengkapan barang

X<sub>3</sub> : kualitas barang

X<sub>4</sub> : harga barang

X<sub>5</sub> : jaminan keamanan/garansi barang rusak

X<sub>6</sub> : penampilan fisik karyawan

X<sub>7</sub> : sikap karyawan terhadap konsumen

X<sub>8</sub> : cara menyampaikan informasi kepada konsumen

Kriteria penilaian :

1 = sangat kurang

2 = kurang

3 = cukup

4 = baik

5 = sangat baik

## Lampiran 1 (Lanjutan)

- X<sub>9</sub> : kebersihan lantai dan barang
- X<sub>10</sub> : layanan spesial
- X<sub>11</sub> : lokasi yang strategis
- X<sub>12</sub> : AC, musik serta penerangan
- X<sub>13</sub> : jumlah troli dan keranjang
- X<sub>14</sub> : luas ruangan, tatanan dan susunan barang
- X<sub>15</sub> : kecukupan informasi letak barang
- X<sub>16</sub> : sarana pengaduan/ komplain/ keluhan bagi konsumen
- X<sub>17</sub> : parkir kendaraan

## Data 6

N o	X 1	X 2	X 3	X 4	X 5	X 6	X 7	X 8	X 9	X 10	X 11	X 12	X 13	X 14	X 15	X 16	X 17	X 18	X 19	X 20
1	4	4	3	4	5	5	5	5	3	4	4	3	4	5	3	5	4	5	4	
2	3	3	4	5	5	4	4	5	4	4	4	3	3	5	5	5	4	5	3	
3	4	4	3	4	3	3	4	5	4	3	4	3	3	3	3	5	3	3	3	
4	5	5	5	5	4	4	5	4	4	4	4	5	4	5	4	3	5	5	3	
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
97	4	3	4	4	3	4	5	3	3	3	5	4	3	4	4	4	4	5	4	
98	3	3	3	3	3	3	4	3	3	3	4	4	3	5	3	4	3	4	3	
99	5	4	5	4	5	3	4	5	5	4	4	4	5	4	4	5	5	5	5	
100	5	2	4	4	4	3	4	3	3	3	4	3	3	4	4	3	3	5	5	

Keterangan :

- X<sub>1</sub> : Fasilitas fisik
- X<sub>2</sub> : Kenyamanan ruangan
- X<sub>3</sub> : Penampilan karyawan
- X<sub>4</sub> : Keadaan interior
- X<sub>5</sub> : Pelayanan tepat
- X<sub>6</sub> : Prosedur pelayanan
- X<sub>7</sub> : Keahlian kerja karyawan
- X<sub>8</sub> : Kemampuan karyawan dalam mengerjakan tugasnya
- X<sub>9</sub> : Pelayanan dengan segera
- X<sub>10</sub> : Menerima respon keluhan konsumen
- X<sub>11</sub> : Kesediaan karyawan dalam memberikan informasi yang dibutuhkan konsumen
- X<sub>12</sub> : Karyawan menggunakan tutur bahasa yang sopan dan mudah dimengerti
- X<sub>13</sub> : Jaminan memberikan hasil yang maksimal dari pelayanan yang diberikan

Kriteria penilaian :

- 1 = tidak baik
- 2 = kurang baik
- 3 = cukup baik
- 4 = baik
- 5 = sangat baik

## Lampiran 1 (Lanjutan)

X<sub>14</sub> : Memberikan rasa aman kpd konsumen selama menginap

X<sub>15</sub> : Karyawan bersikap ramah

X<sub>16</sub> : Pelayanan yg diberikan sesuai dg harga yang telah dibayar

X<sub>17</sub> : Karyawan memberikan perhatian individual pada setiap konsumen

X<sub>18</sub> : Karyawan melayani konsumen dengan cara simpatik

X<sub>19</sub> : Pelayanan yang diberikan tidak membeda-bedakan antara pelanggan satu dengan lainnya

X<sub>20</sub> : Karyawan mampu memahami kebutuhan tamu

## Data 7

N o	X 1	X 2	X 3	X 4	X 5	X 6	X 7	X 8	X 9	X 10	X 11	X 12	X 13	X 14	X 15	X 16	X 17	X 18	X 19
1	4	5	3	3	3	5	5	4	4	3	4	4	4	3	4	4	4	4	4
2	5	4	4	4	4	4	4	4	4	4	4	4	4	5	4	4	4	5	4
3	3	3	3	4	4	3	3	4	4	4	3	3	3	3	4	3	4	4	4
4	3	3	4	5	5	5	3	4	3	3	3	3	3	4	3	4	5	5	4
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
73	3	3	3	4	4	4	3	4	2	2	3	3	4	4	3	4	4	4	4
74	4	5	3	4	3	3	3	4	5	3	4	4	5	5	3	4	5	5	5
75	3	3	3	3	3	3	3	4	4	3	4	4	3	3	4	4	3	4	3

Keterangan :

X<sub>1</sub> : bahan baku utama pembuatan bakpao

X<sub>2</sub> : rasa bakpao

X<sub>3</sub> : Harga

X<sub>4</sub> : keragaman bakpao dalam rasa

X<sub>5</sub> : tekstur bakpao/ kelembutan bakpao

X<sub>6</sub> : kandungan gizi

X<sub>7</sub> : kecepatan pelayanan pembelian

X<sub>8</sub> : iklan atau promosi

X<sub>9</sub> : bentuk bakpao

X<sub>10</sub>: warna

X<sub>11</sub> : aroma bakpao

X<sub>12</sub> : bentuk

X<sub>13</sub>: warna kemasan

X<sub>14</sub>: nama merek

X<sub>15</sub>: informasi produk  
pada kemasan

X<sub>16</sub>: masa kadaluarsa/  
umur produk

X<sub>17</sub>: jaminan halal

X<sub>18</sub>: terdaftar pada  
departemen  
kesehatan

X<sub>19</sub>: jaminan keamanan  
kemasan  
pangan (HACCP)

Kriteria penilaian :

1 = sangat buruk

4 = Baik

2 = Buruk

5 = Sangat baik

3 = Cukup buruk

## Lampiran 1 (Lanjutan)

### Data 8

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
1	4	4	4	3	5	3	2	4	4	2	3
2	5	5	5	4	5	5	4	5	5	4	5
3	3	3	3	3	3	3	3	3	3	3	3
4	4	5	5	4	5	5	1	5	4	1	5
:	:	:	:	:	:	:	:	:	:	:	:
79	4	5	5	4	4	3	2	5	4	1	3
80	4	4	4	4	4	3	4	4	4	4	3
81	3	3	3	3	4	3	3	4	3	3	3
82	5	3	3	2	5	5	2	3	3	5	5

Keterangan :

- X<sub>1</sub> : Tingkat pertimbangan dalam jenis
- X<sub>2</sub> : Tingkat pertimbangan dalam warna
- X<sub>3</sub> : Tingkat pertimbangan dalam rasa
- X<sub>4</sub> : Tingkat pertimbangan dalam ukuran
- X<sub>5</sub> : Tingkat pertimbangan dalam manfaat
- X<sub>6</sub> : Tingkat pertimbangan dalam harga
- X<sub>7</sub> : Tingkat pertimbangan dalam label
- X<sub>8</sub> : Tingkat pertimbangan dalam ketersediaan
- X<sub>9</sub> : Tingkat pertimbangan dalam kesegaran
- X<sub>10</sub> : Tingkat pertimbangan dalam kerenyahan
- X<sub>11</sub> : Tingkat pertimbangan dalam bebas pestisida

Kriteria penilaian :

- |                                  |                            |
|----------------------------------|----------------------------|
| 1 = sangat tidak dipertimbangkan | 4 = dipertimbangkan        |
| 2 = tidak dipertimbangkan        | 5 = sangat dipertimbangkan |
| 3 = cukup dipertimbangkan        |                            |

### Data 9

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14
1	5	2	5	5	5	5	5	4	5	4	4	4	5	4
2	5	4	5	3	4	5	5	3	5	3	5	4	5	5
3	4	4	4	3	4	3	4	3	3	4	5	3	2	2
4	4	4	4	5	5	4	4	4	4	4	4	4	5	4
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
67	3	3	3	4	4	4	4	4	3	3	3	3	4	4
68	3	3	3	4	4	5	4	4	5	4	4	3	3	3
69	4	4	4	3	4	4	4	4	3	3	4	3	3	3
70	4	4	4	4	4	4	3	4	3	3	4	3	4	4

Keterangan :

- X<sub>1</sub> : *brand image* yang dimiliki teh Tong Tji
- X<sub>2</sub> : gaya hidup akibat mengkonsumsi teh *cup* Tong Tji
- X<sub>3</sub> : manfaat teh Tong Tji

## Lampiran 1 (Lanjutan)

- X<sub>4</sub> : kemurnian rasa teh Tong Tji  
X<sub>5</sub> : aroma Teh Tong Tji  
X<sub>6</sub> : kepraktisan untuk dikonsumsi  
X<sub>7</sub> : kesegaran ketika dikonsumsi  
X<sub>8</sub> : harga yang lebih murah  
X<sub>9</sub> : variasi rasa produk  
X<sub>10</sub> : produk tidak cepat basi  
X<sub>11</sub> : produk higienis  
X<sub>12</sub> : homogenitas kualitas produk  
X<sub>13</sub> : gambar yang menarik pada kemasan *cup*  
X<sub>14</sub> : pemberian warna gambar yang menarik pada kemasan *cup*
- Kriteria penilaian :  
1 = tidak baik  
2 = kurang baik  
3 = cukup baik  
4 = baik  
5 = sangat baik

## Data 10

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X <sub>10</sub>	X <sub>11</sub>	X <sub>12</sub>	X <sub>13</sub>	X <sub>14</sub>	X <sub>15</sub>
1	3	3	3	3	3	4	4	3	3	4	4	2	2	2	2
2	2	2	4	3	4	2	2	2	4	4	4	2	2	2	2
3	2	2	2	3	3	3	3	3	3	4	2	2	2	2	1
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
29	2	4	2	2	2	2	2	2	4	3	4	1	2	2	2
30	2	2	4	3	4	2	2	2	4	3	4	3	3	1	1
31	2	2	4	4	4	2	2	2	2	2	4	2	2	2	4

Keterangan :

- X<sub>1</sub> : Rasa  
X<sub>2</sub> : Aroma  
X<sub>3</sub> : Warna  
X<sub>4</sub> : Tekstur  
X<sub>5</sub> : Tingkat kekentalan  
X<sub>6</sub> : Tingkat kelarutan  
X<sub>7</sub> : kegunaan *puree* mangga  
Podang sebagai bahan baku  
X<sub>8</sub> : kegunaan *puree* mangga Podang  
sebagai bahan tambahan  
pembuatan produk olahan lai

- X<sub>9</sub>: kandungan gizi  
X<sub>10</sub>: manfaat *puree* untuk kesehatan  
X<sub>11</sub>: masa kadaluarsa  
X<sub>12</sub>: bahan kemasan  
X<sub>13</sub>: bentuk kemasan

X<sub>14</sub>: informasi produk

X<sub>15</sub>: jaminan halal

Kriteria penilaian :

- 1 = Sangat jelek  
2 = Jelek  
3 = Cukup  
4 = Baik  
5 = Sangat baik

## Lampiran 2. Data Hilang pada Setiap Kondisi

### DATA 1 :

KONDISI 1							
No	X1	X2	X3	X6	X8	X13	
13	3	3	-	3	3	3	3
64	3	-	3	3	3	3	3
65	-	3	3	3	3	3	3
79	3	3	3	3	-	3	3
83	-	3	3	3	3	3	3
94	3	3	2	3	3	-	3
96	3	3	3	-	3	3	3

KONDISI 2										
No	X2	X3	XS	X6	X9	X10	X11	X12	X13	X14
6	3	3	3	-	3	3	3	3	3	3
7	3	2	1	1	1	4	-	3	1	1
8	3	3	-	3	3	3	3	3	3	3
36	-	3	3	3	-	3	3	3	3	3
45	3	3	3	3	3	-	3	3	3	3
47	3	3	3	3	3	3	3	1	-	3
50	-	3	3	3	3	3	3	3	1	3
61	3	2	3	3	2	3	2	2	3	-
65	-	3	3	3	3	3	3	3	3	3
70	2	2	2	2	2	-	2	2	3	3
73	3	3	3	3	-	3	3	3	3	3
74	3	-	3	3	3	2	3	3	3	3
97	3	3	3	3	3	3	3	-	3	3
98	3	-	3	3	2	2	2	2	3	3

KONDISI 3										
No	X1	X3	X4	XS	X6	X7	X9	X10	X11	X12
1	3	3	-	3	3	3	3	3	3	3
5	3	3	3	3	-	2	3	3	3	3
10	3	3	3	-	3	3	3	3	3	3
19	3	3	3	3	3	3	-	3	4	3
21	3	-	3	3	3	3	3	4	3	3
32	3	-	3	3	3	3	3	3	3	3
36	3	3	-	3	3	3	3	3	3	3
46	3	3	3	-	3	3	3	3	3	3
47	3	3	3	3	3	3	3	3	-	1
51	3	3	-	3	3	3	3	3	3	3
54	3	3	-	3	3	3	3	3	3	3
57	3	4	3	4	3	3	3	3	3	-
59	3	-	3	3	3	3	3	2	2	2
66	-	3	3	4	4	4	4	4	4	4
73	2	3	3	3	3	2	-	3	3	3
76	3	3	2	2	2	2	-	3	3	3
78	3	4	3	3	3	3	3	-	4	4
83	3	3	-	3	3	3	3	3	3	3
88	1	1	-	1	1	1	1	1	1	1
89	3	3	3	3	3	3	-	3	3	3
90	2	-	3	3	3	2	2	3	2	2

### DATA 10 :

KONDISI 1			
No	X2	X6	X7
1	3	-	4
13	2	4	-
12	-	2	2

KONDISI 2				
No	X7	X9	X12	X14
9	2	4	2	2
14	3	-	5	4
26	-	2	1	2
29	2	4	-	2

KONDISI 3						
No	X1	X2	X7	XS	X10	X11
1	3	3	4	3	4	2
2	2	2	2	2	-	4
4	4	4	-	4	4	2
20	2	2	2	-	3	4
21	4	4	3	4	2	4
26	-	2	2	2	4	2
29	2	2	2	2	3	2

### Keterangan :

Kondisi 1 = persentase data hilang 0,5%

Kondisi 2 = persentase data hilang 1%

Kondisi 3 = persentase data hilang 1,5%

= nilai yang hilang

### Lampiran 3. Hasil Pendugaan Nilai Data Hilang

#### DATA 1

No	Forward Imputation					Mode Imputation					Median Imputation							
	X1	X2	X3	X6	X8	X13	X1	X2	X3	X6	X8	X13	X1	X2	X3	X6	X8	X13
13	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
64	3	3	3	3	3	3	3	3	3	3	3	3	64	3	3	3	3	3
65	3	3	3	3	3	3	65	3	3	3	3	3	65	3	3	3	3	3
79	3	3	3	3	3	3	79	3	3	3	3	3	79	3	3	3	3	3
83	3	3	3	3	3	3	83	3	3	3	3	3	83	3	3	3	3	3
94	3	3	2	3	3	3	94	3	3	2	3	3	94	3	3	2	3	3
96	3	3	3	3	3	3	96	3	3	3	3	3	96	3	3	3	3	3

No	KONDISI 2																			
	Forward Imputation					Mode Imputation					X2	X3	X5	X6	X9	X10	X11	X12	X13	X14
X2	X3	X5	X6	X9	X10	X11	X12	X13	X14	No	X2	X3	X5	X6	X9	X10	X11	X12	X13	X14
6	3	3	3	3	3	3	3	3	6	3	3	3	3	3	3	3	3	3	3	3
7	3	2	1	1	1	4	1	3	1	1	7	3	2	1	1	1	4	3	1	1
8	3	3	3	3	3	3	3	3	8	3	3	3	3	3	3	3	3	3	3	3
36	3	3	3	3	3	3	3	3	36	3	3	3	3	3	3	3	3	3	3	3
45	3	3	3	3	3	3	3	3	45	3	3	3	3	3	3	3	3	3	3	3
47	3	3	3	3	3	3	1	3	47	3	3	3	3	3	3	3	3	1	3	3
50	3	3	3	3	3	3	3	3	1	50	3	3	3	3	3	3	3	3	1	3
61	3	2	3	3	2	3	2	2	3	61	3	2	3	3	2	3	2	2	3	3
65	3	3	3	3	3	3	3	3	65	3	3	3	3	3	3	3	3	3	3	3
70	2	2	2	2	2	3	2	2	2	70	2	2	2	2	2	3	2	2	3	3
73	3	3	3	3	3	3	3	3	73	3	3	3	3	3	3	3	3	3	3	3
74	3	3	3	3	3	2	3	3	74	3	3	3	3	3	2	3	3	3	3	3
97	3	3	3	3	3	3	3	3	97	3	3	3	3	3	3	3	3	3	3	3
98	3	3	3	3	2	2	2	1	2	3	98	3	3	3	3	2	2	2	2	3

No	KONDISI 2																			
	Median Imputation																			
X2	X3	X5	X6	X9	X10	X11	X12	X13	X14											
6	3	3	3	3	3	3	3	3	3											
7	3	2	1	1	1	4	1	3	1											
8	3	3	3	3	3	3	3	3	3											
36	3	3	3	3	3	3	3	3	3											
45	3	3	3	3	3	3	3	3	3											
47	3	3	3	3	3	3	3	1	3											
50	3	3	3	3	3	3	3	3	3											
61	3	2	3	3	2	3	2	2	3											
65	3	3	3	3	3	3	3	3	3											
70	2	2	2	2	2	3	2	2	3											
73	3	3	3	3	3	3	3	3	3											
74	3	3	3	3	3	2	3	3	3											
97	3	3	3	3	3	3	3	3	3											
98	3	3	3	3	2	2	2	2	3											

No	KONDISI 3																			
	Forward Imputation					Mode Imputation														
X1	X3	X4	X5	X6	X7	X9	X10	X11	X12	No	X1	X3	X4	X5	X6	X7	X9	X10	X11	X12
1	3	3	3	3	3	3	3	3	1	1	3	3	3	3	3	3	3	3	3	3
5	3	3	3	3	3	3	2	3	3	3	5	3	3	3	3	2	3	3	3	3
10	3	3	3	3	3	3	3	3	3	10	3	3	3	3	3	3	3	3	3	3
19	3	3	3	3	3	3	3	4	3	19	3	3	3	3	3	3	4	3	3	3
21	3	3	3	3	3	3	4	3	3	21	3	3	3	3	3	3	3	4	3	3
32	3	2	3	3	3	3	3	3	3	32	3	3	3	3	3	3	3	3	3	3
36	3	3	3	3	3	3	3	3	36	3	3	3	3	3	3	3	3	3	3	3
46	3	3	3	3	3	3	3	3	3	46	3	3	3	3	3	3	3	3	3	3
47	3	3	3	3	3	3	3	3	2	47	3	3	3	3	3	3	3	3	3	1
51	3	3	3	3	3	3	3	3	3	51	3	3	3	3	3	3	3	3	3	3
54	3	3	3	3	3	3	3	3	3	54	3	3	3	3	3	3	3	3	3	3
57	3	4	3	4	3	3	3	3	3	57	3	4	3	4	3	3	3	3	3	3
59	3	2	3	3	3	2	2	2	2	59	3	3	3	3	3	3	2	2	2	2
66	3	3	3	4	4	4	4	4	4	66	3	3	3	4	4	4	4	4	4	4
73	2	3	3	3	3	2	3	3	3	73	2	3	3	3	2	2	3	3	3	3
76	3	3	2	2	2	2	3	3	3	76	3	2	2	2	2	2	3	3	3	3
78	3	4	3	3	3	3	3	4	4	78	3	4	3	3	3	3	3	3	4	4
83	3	3	3	3	3	3	3	3	3	83	3	3	3	3	3	3	3	3	3	3
88	1	1	1	1	1	1	1	1	1	88	1	1	1	1	1	1	1	1	1	1
89	3	3	3	3	3	3	3	3	3	89	3	3	3	3	3	3	3	3	3	3
90	2	3	3	3	3	2	2	3	2	90	2	3	3	3	3	2	2	3	2	2

No	KONDISI 3 Median Imputation											
	X1	X3	X4	X5	X6	X7	X9	X10	X11	X12		
1	3	3	3	3	3	3	3	3	3	3		
5	3	3	3	3	3	2	3	3	3	3		
10	3	3	3	3	3	3	3	3	3	3		
19	3	3	3	3	3	3	3	4	3	3		
21	3	3	3	3	3	3	3	4	3	3		
32	3	3	3	3	3	3	3	3	3	3		
36	3	3	3	3	3	3	3	3	3	3		
46	3	3	3	3	3	3	3	3	3	3		
47	3	3	3	3	3	3	3	3	3	1		
51	3	3	3	3	3	3	3	3	3	3		
54	3	3	3	3	3	3	3	3	3	3		
57	3	4	3	4	3	3	3	3	3	3		
59	3	3	3	3	3	3	3	2	2	2		
66	3	3	3	4	4	4	4	4	4	4		
73	2	3	3	3	3	2	3	3	3	3		
76	3	3	2	2	2	2	3	3	3	3		
78	3	4	3	3	3	3	3	3	4	4		
83	3	3	3	3	3	3	3	3	3	3		
88	1	1	3	1	1	1	1	1	1	1		
89	3	3	3	3	3	3	3	3	3	3		
90	2	3	3	3	3	2	2	3	2	2		

## DATA 10

No	KONDISI 1										
	Forward Imputation			Mode Imputation			Median Imputation				
No	X2	X6	X7	No	X2	X6	X7	No	X2	X6	X7
1	3	3	4	1	3	2	4	1	3	3	4
13	2	4	4	13	2	4	2	15	2	4	2
31	2	2	2	31	2	2	2	31	3	2	2

No	KONDISI 2																
	Forward Imputation					Mode Imputation			Median Imputation								
No	X7	X9	X12	X14	X15	No	X7	X9	X12	X14	X15	No	X7	X9	X12	X14	X15
9	2	4	2	2	3	9	2	4	2	2	2	9	2	4	2	2	2
14	3	4	5	4	4	14	3	4	5	4	4	14	3	4	5	4	4
26	2	2	1	2	2	26	2	2	1	2	2	26	2	2	1	2	2
29	2	4	2	2	2	29	2	4	2	2	2	29	2	4	2	2	2

No	KONDISI 3																				
	Forward Imputation					Mode Imputation			Median Imputation												
No	X1	X3	X7	X8	X10	X11	X12	X8	X10	X11	X14	No	X1	X3	X7	X8	X10	X11	X14		
1	3	3	4	3	4	4	2	1	3	4	4	2	1	3	3	4	3	4	2		
2	2	2	2	2	2	4	2	2	2	2	2	9	2	2	2	2	2	3	4	2	
4	4	4	4	4	4	4	2	4	4	4	4	4	2	4	4	4	4	4	4	2	
20	2	2	2	2	3	4	2	20	2	2	2	2	3	4	2	20	2	2	3	4	2
21	4	4	3	4	2	4	2	21	4	4	3	4	2	4	2	21	4	4	3	4	2
26	3	3	2	2	4	2	2	26	2	3	2	2	4	2	2	26	2	3	2	2	2
29	2	2	2	2	3	3	2	29	2	2	2	2	3	4	2	29	2	2	2	3	4

Keterangan :

Kondisi 1 = persentase data hilang 0,5%

Kondisi 2 = persentase data hilang 1%

Kondisi 3 = persentase data hilang 1,5%

## Lampiran 4. Nilai Akar Ciri Seluruh Dimensi pada Data Lengkap

		Nilai Akar Ciri Seluruh Dimensi Data Sebenarnya																	
DATA 1		DATA 2		DATA 3		DATA 4		DATA 5		DATA 6		DATA 7		DATA 8		DATA 9		DATA 10	
Dimensi	Eigen value	Dimensi	Eigen value	Dimensi	Eigen value	Dimensi	Eigen value	Dimensi	Eigen value	Dimensi	Eigen value	Dimensi	Eigen value	Dimensi	Eigen value	Dimensi	Eigen value	Dimensi	Eigen value
1	0,5762	1	0,1554	1	0,4222	1	0,2233	1	0,3058	1	0,2855	1	0,2833	1	0,359	1	0,2295	1	0,3281
2	0,0978	2	0,1294	2	0,1565	2	0,1622	2	0,1734	2	0,1024	2	0,0973	2	0,1504	2	0,1386	2	0,1221
3	0,0719	3	0,1087	3	0,1361	3	0,1115	3	0,0981	3	0,0796	3	0,091	3	0,1181	3	0,1197	3	0,1082
4	0,0405	4	0,0932	4	0,0763	4	0,0967	4	0,0741	4	0,0694	4	0,0717	4	0,083	4	0,0883	4	0,0933
5	0,0374	5	0,0724	5	0,0616	5	0,0897	5	0,0552	5	0,0571	5	0,0616	5	0,0682	5	0,0766	5	0,0761
6	0,0338	6	0,0593	6	0,0514	6	0,0577	6	0,0483	6	0,0517	6	0,0567	6	0,0609	6	0,0634	6	0,0648
7	0,0324	7	0,0521	7	0,0392	7	0,0507	7	0,0415	7	0,0482	7	0,0492	7	0,0508	7	0,0577	7	0,0541
8	0,0289	8	0,0464	8	0,0328	8	0,0427	8	0,0338	8	0,0399	8	0,0428	8	0,0378	8	0,0524	8	0,0344
9	0,0235	9	0,0448	9	0,024	9	0,0372	9	0,0306	9	0,0357	9	0,0395	9	0,0335	9	0,048	9	0,0298
10	0,0152	10	0,0382			10	0,0345	10	0,0303	10	0,0303	10	0,0367	10	0,0218	10	0,0407	10	0,0229
11	0,0142	11	0,0305			11	0,0311	11	0,0234	11	0,0297	11	0,0346	11	0,0164	11	0,0313	11	0,0198
12	0,0126	12	0,0271			12	0,0247	12	0,0196	12	0,0276	12	0,0271			12	0,0256	12	0,0138
13	0,0083	13	0,0251			13	0,0225	13	0,0181	13	0,0246	13	0,0218			13	0,0156	13	0,0132
14	0,0073	14	0,0229			14	0,0155	14	0,0161	14	0,023	14	0,0203			14	0,0126	14	0,0127
		15	0,0193				15	0,0117	15	0,0208	15	0,0178						15	0,0067
		16	0,0169				16	0,011	16	0,0198	16	0,0164							
		17	0,0137				17	0,0091	17	0,0183	17	0,0122							
		18	0,0117						18	0,0144	18	0,0113							
		19	0,0094						19	0,0116	19	0,0086							
		20	0,0072						20	0,0106									
		21	0,0066																
		22	0,0059																
		23	0,0037																

## Lampiran 5. Penentuan Banyak Dimensi yang Digunakan

### ➤ Data Lengkap

Penentuan Banyaknya Dimensi			
DATA	Banyaknya dimensi awal	Kriteria akar ciri ( $1/m$ )	Banyaknya dimensi yang digunakan
1	14	0,0714	3
2	23	0,0435	9
3	9	0,1111	3
4	14	0,0714	5
5	17	0,0588	4
6	20	0,0500	6
7	19	0,0526	6
8	11	0,0909	3
9	14	0,0714	5
10	15	0,0667	5

### ➤ Data Lengkap Hasil Imputasi

Penentuan Banyaknya Dimensi							
Data	Percentase Data Hilang	Metode Imputasi	Banyaknya dimensi awal	Kriteria akar ciri ( $1/m$ )	Banyaknya dimensi yang digunakan		
1	Kondisi 1	Forward Imputation	14	0,0714	3		
		Mode Imputation			3		
		Median Imputation			3		
	Kondisi 2	Forward Imputation			3		
		Mode Imputation			2		
		Median Imputation			2		
	Kondisi 3	Forward Imputation			2		
		Mode Imputation			2		
		Median Imputation			2		
.	.	.	14	0,0714	.		
		.			.		
		.			.		
		.			.		
		.			.		
	10	Kondisi 1	15	0,0667	5		
					5		
					5		
		Kondisi 2			5		
					5		
	.	Kondisi 3			5		
					5		
					5		
		.			5		
					5		

# Kondisi 1 = persentase data hilang 0,5%

Kondisi 2 = persentase data hilang 1%

Kondisi 3 = persentase data hilang 1,5%

## Lampiran 6. Akar Ciri, Proporsi Keragaman dan Proporsi Keragaman Kumulatif pada Data Lengkap

Data	Dimensi	Akar Ciri	Proporsi Kumulatif (%)	Proporsi Keragaman Kumulatif (%)
1	1	0,7332	73,32	73,32
	2	0,0923	9,23	82,55
	3	0,0772	7,72	90,27
2	1	0,157	15,7	15,7
	2	0,1285	12,85	28,55
	3	0,1065	10,65	39,2
	4	0,0967	9,67	48,87
	5	0,0799	7,99	56,86
	6	0,07	7	63,86
	7	0,0567	5,67	69,53
	8	0,0519	5,19	74,72
	9	0,0437	4,37	79,09
	:	:	:	:
9	1	0,3065	30,65	30,65
	2	0,1733	17,33	47,98
	3	0,131	13,1	61,08
	4	0,0939	9,39	70,47
	5	0,0644	6,44	76,91
10	1	0,3297	32,97	32,97
	2	0,1565	15,65	48,62
	3	0,1202	12,02	60,64
	4	0,1076	10,76	71,4
	5	0,0995	9,95	81,35

## Lampiran 7. Nilai Skor Komponen Objek pada Data Lengkap

➤ Data 1 :

Objek	Dimensi		
	1	2	3
1	-0,24	0,03	-0,14
2	-0,24	0,03	-0,14
3	-0,24	0,03	-0,14
4	-0,24	0,03	-0,14
⋮	⋮	⋮	⋮
97	-0,24	0,03	-0,14
98	-0,18	0,04	-0,06
99	-0,24	0,03	-0,14
100	-0,17	-0,13	-0,37

➤ Data 2 :

Objek	Dimensi								
	1	2	3	4	5	6	7	8	9
1	-1,09	-0,3	-0,66	0,32	0,97	0,82	-0,45	0,78	0,39
2	-0,57	1,74	0,64	-0,04	2,11	0,98	1,31	-0,49	2,47
3	1,15	1,49	0,77	-0,24	-0,96	-0,77	0,02	1,38	0,34
4	-0,92	0,74	-0,85	0,15	0,19	0,55	-0,12	0,93	-0,4
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
69	1,09	-1,7	-0,73	-0,02	0,37	-0,11	0,84	-0,47	0,3
70	-0,26	-0,86	1,07	-0,88	1,08	1,04	-0,36	-0,28	1,51
71	1,18	-1	1,4	-0,71	0,42	0,06	0,63	-0,18	1,44
72	1,09	-0,66	-0,15	-0,14	-0,3	-0,25	1,48	-0,28	-0,58

.

.

.

.

➤ Data 9 :

Objek	Dimensi				
	1	2	3	4	5
1	-0,68	1,19	0,94	0,26	-1,09
2	-0,57	1,08	0	-0,94	0,84
3	0,64	-0,66	-0,45	-0,69	0,49
4	0,39	-0,35	0,46	0,23	-1,14
⋮	⋮	⋮	⋮	⋮	⋮
67	0,8	-0,85	-0,08	0,5	-0,27
68	0,25	-0,18	0,72	0,54	1,15
69	0,76	-0,81	-0,11	0,51	-0,32
70	0,83	-0,89	-0,15	0,54	-0,16

➤ Data 10 :

Objek	Dimensi				
	1	2	3	4	5
1	-0,56	0,78	1,33	-0,76	0,39
2	0,71	-0,23	-0,17	1,06	-0,2
3	0,26	-0,01	1,12	-0,01	-1,15
4	-1	-0,07	1,06	-0,92	0,45
⋮	⋮	⋮	⋮	⋮	⋮
28	1,05	-0,14	0,18	0,23	-1,34
29	0,97	0,54	0,35	-0,33	-0,42
30	0,63	-0,85	-1,05	-0,32	-1,06
31	0,74	-0,84	-0,13	0,7	-0,37

## Lampiran 8. Nilai Kuantifikasi pada Data Lengkap

DATA 1 :

Variable: X1		
Type: Ordinal	Missing:	0
Category:	Marginal Frequency	Quantification
1	4	-4,90
2	8	,13
3	88	,21
4	0	,00

Single Category Coordinates			
Category	Dimension		
	1	2	3
1	4,04	1,59	-,81
2	-,11	-,04	,02
3	-,17	-,07	,04
4	,00	,00	,00

Multiple Category Coordinates			
Category	Dimension		
	1	2	3
1	4,04	1,59	-,82
2	-,10	,07	,25
3	-,17	-,08	,01
4	,00	,00	,00

::	::	::
----	----	----

Variable: X14		
Type: Ordinal	Missing:	0
Category:	Marginal Frequency	Quantification
1	4	-4,89
2	4	,12
3	89	,21
4	3	,37

Single Category Coordinates			
Category	Dimension		
	1	2	3
1	4,50	,77	1,35
2	-,11	,02	,03
3	-,20	-,03	-,06
4	-,34	-,06	-,10

Multiple Category Coordinates			
Category	Dimension		
	1	2	3
1	4,50	,84	1,34
2	,26	-,1,03	,13
3	-,20	,01	-,06
4	-,29	,01	-,31

DATA 10 :

Variable: X1		
Type: Ordinal	Missing:	0
Category:	Marginal Frequency	Quantification
1	1	-1,49
2	16	-,87
3	8	,10
4	6	,10
5	0	,00

Single Category Coordinates			
Category	Dimension		
	1	2	3
1	,80	-,69	,08
2	,46	-,40	,05
3	-,58	,50	-,06
4	-,58	,50	-,06
5	,00	,00	,00

Multiple Category Coordinates			
Category	Dimension		
	1	2	3
1	1,05	-,14	,18
2	,44	-,44	,04
3	-,60	,86	,23
4	-,55	,05	-,44
5	,00	,00	,00

::	::	::
----	----	----

Variable: X15		
Type: Ordinal	Missing:	0
Category:	Marginal Frequency	Quantification
1	4	-,74
2	14	-,74
3	2	-,74
4	9	,135
5	2	,135

Single Category Coordinates			
Category	Dimension		
	1	2	3
1	,40	,45	-,06
2	,40	,45	-,06
3	,40	,45	-,06
4	-,72	-,82	,11
5	-,72	-,82	,11

Multiple Category Coordinates			
Category	Dimension		
	1	2	3
1	,44	-,31	,26
2	,24	,67	,01
3	1,39	,43	-,12
4	-,83	-,70	,11
5	-,24	-,32	,11

## Lampiran 9. Nilai Bobot Peubah pada Data Lengkap

DATA 1

Peubah	Dimensi		
	1	2	3
X1	-0,824	-0,325	0,166
X2	-0,694	0,51	0,238
X3	-0,912	0,162	0,306
X4	-0,908	0,165	0,308
X5	-0,933	-0,291	0,137
X6	-0,932	-0,292	0,138
X7	-0,833	-0,182	-0,303
X8	-0,92	-0,155	-0,284
X9	-0,935	-0,288	0,131
X10	-0,898	0,174	0,322
X11	-0,797	0,491	-0,229
X12	-0,733	0,511	-0,343
X13	-0,682	-0,081	-0,479
X14	-0,922	-0,157	-0,276

•  
•  
•

DATA 10

Variable	Dimension				
	1	2	3	4	5
X1	-0,534	0,46	-0,053	-0,443	0,431
X2	-0,462	0,408	-0,083	-0,544	0,47
X3	-0,359	0,744	-0,098	0,214	-0,282
X4	-0,535	0,282	-0,629	0,098	-0,248
X5	-0,063	-0,38	-0,39	0,393	0,522
X6	-0,655	-0,133	0,545	-0,163	0,015
X7	-0,839	0,083	0,349	-0,083	-0,04
X8	-0,787	0,043	0,295	0,09	0,019
X9	-0,603	0,414	-0,336	0,461	-0,19
X10	-0,674	0,066	0,07	0,489	0,298
X11	-0,534	-0,203	0,39	0,015	-0,582
X12	-0,617	-0,42	-0,374	-0,347	-0,094
X13	-0,338	-0,402	-0,534	-0,488	-0,39
X14	-0,627	-0,497	-0,164	0,184	0,173
X15	-0,534	-0,605	0,08	0,166	0,126

## Lampiran 10. Tes Konvergensi pada Data Lengkap

### DATA 1 :

* The History of Iterations *						
Iteration	Total Fit	Total Loss	Multiple Loss	Single Loss	Iteration Change	
1	,7939033	2,2060967	2,0856777	,1204190	,0479316	
2	,8393167	2,1606833	1,9992694	,1614139	,0454134	
3	,8610940	2,1389060	1,9658401	,1730659	,0217773	
4	,8728491	2,1271508	1,9600803	,1670706	,0117551	
5	,8810141	2,1189859	1,9665519	,1524340	,0081650	
6	,8875201	2,1124799	1,9766947	,1378752	,0065060	
7	,8926250	2,1073750	1,9862871	,1210879	,0051049	
8	,8963527	2,1036473	1,9942630	,1093843	,0037276	
9	,8988811	2,1011189	2,0009028	,1002161	,0025284	
10	,9004781	2,0995219	2,0063284	,0931934	,0015971	
11	,9014347	2,0985653	2,0106674	,0878979	,0009565	
12	,9019886	2,0980114	2,0140853	,0839261	,0005540	
13	,9023034	2,0976966	2,0167630	,0809336	,0003148	
14	,9024806	2,0975194	2,0188616	,0786579	,0001772	
15	,9025799	2,0974201	2,0205095	,0769105	,0000994	
16	,9026357	2,0973643	2,0218058	,0755585	,0000558	
17	,9026670	2,0973330	2,0228268	,0745061	,0000313	
18	,9026847	2,0973153	2,0236304	,0736849	,00000177	
19	,9026947	2,0973053	2,0242618	,0730435	,0000100	
20	,9027004	2,0972996	2,0247573	,0725423	,00000057	

The iterative process stops because the convergence test value is reached.

### DATA 10 :

* The History of Iterations *						
Iteration	Total Fit	Total Loss	Multiple Loss	Single Loss	Iteration Change	
1	,7569057	4,2430943	3,8853070	,3577873	,0291105	
2	,7702138	4,2297862	3,8763481	,3534381	,0133081	
3	,7771715	4,2228285	3,8672366	,3555919	,0069578	
4	,7820951	4,2179049	3,8602256	,3576793	,0049236	
5	,7858001	4,2141999	3,8555066	,3586933	,0037049	
6	,7889858	4,2110142	3,8528081	,3582060	,0031857	
7	,7918352	4,2081648	3,8501794	,3579854	,0028493	
8	,7944593	4,2055407	3,8469007	,3586400	,0026241	
9	,7969193	4,2030807	3,8431420	,3599387	,0024600	
10	,7992510	4,2007490	3,8391693	,3615797	,0023317	
.	.	.	.	.	.	
.	.	.	.	.	.	
.	.	.	.	.	.	
44	,8133750	4,1866250	3,8198222	,3668028	,00000149	
45	,8133895	4,1866105	3,8197273	,3668832	,00000145	
46	,8134035	4,1865965	3,8196330	,3669635	,00000140	
47	,8134172	4,1865828	3,8195392	,3670436	,00000137	
48	,8134305	4,1865695	3,8194460	,3671236	,00000133	
49	,8134434	4,1865566	3,8193531	,3672034	,00000130	
50	,8134561	4,1865439	3,8192607	,3672832	,00000127	
51	,8134685	4,1865315	3,8191686	,3673629	,00000124	
52	,8134806	4,1865194	3,8190769	,3674425	,00000121	
53	,8134920	4,1865080	3,8189856	,3675224	,00000115	
54	,8135023	4,1864977	3,8188755	,3676222	,00000103	
55	,8135118	4,1864882	3,8187515	,3677366	,00000095	

The iterative process stops because the convergence test value is reached.

## Lampiran 11. Nilai Akar Ciri Seluruh Dimensi pada Data Lengkap Hasil Imputasi

### DATA 1

Kondisi 1						Kondisi 2						Kondisi 3					
FOI		MOI		MEI		FOI		MOI		MEI		FOI		MOI		MEI	
D	AC	D	AC	D	AC	D	AC	D	AC	D	AC	D	AC	D	AC	D	AC
1	0,5757	1	0,5757	1	0,5757	1	0,5748	1	0,5602	1	0,5602	1	0,5781	1	0,5774	1	0,5774
2	0,0981	2	0,0981	2	0,0981	2	0,0992	2	0,0991	2	0,0991	2	0,0995	2	0,099	2	0,099
3	0,0724	3	0,0724	3	0,0724	3	0,0723	3	0,0699	3	0,0699	3	0,0701	3	0,07	3	0,07
4	0,0411	4	0,0411	4	0,0411	4	0,0399	4	0,0467	4	0,0467	4	0,0396	4	0,0393	4	0,0393
5	0,0368	5	0,0368	5	0,0368	5	0,0373	5	0,0407	5	0,0407	5	0,0378	5	0,0381	5	0,0381
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
14	0,0072	14	0,0072	14	0,0072	14	0,0066	14	0,0077	14	0,0077	14	0,0075	14	0,0081	14	0,0081
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

### DATA 10

Kondisi 1						Kondisi 2						Kondisi 3					
FOI		MOI		MEI		FOI		MOI		MEI		FOI		MOI		MEI	
D	AC	D	AC	D	AC	D	AC	D	AC	D	AC	D	AC	D	AC	D	AC
1	0,329	1	0,3277	1	0,3276	1	0,3243	1	0,3254	1	0,3254	1	0,3381	1	0,3332	1	0,3333
2	0,1246	2	0,1218	2	0,1242	2	0,1234	2	0,1246	2	0,1246	2	0,1211	2	0,1187	2	0,1203
3	0,1059	3	0,1091	3	0,109	3	0,1124	3	0,1087	3	0,1087	3	0,1066	3	0,1082	3	0,1061
4	0,093	4	0,0924	4	0,089	4	0,0922	4	0,0929	4	0,0929	4	0,0966	4	0,0925	4	0,0926
5	0,0772	5	0,0755	5	0,0756	5	0,0782	5	0,0782	5	0,0782	5	0,0738	5	0,0752	5	0,0756
6	0,0633	6	0,0652	6	0,0654	6	0,0615	6	0,0619	6	0,0619	6	0,0606	6	0,0574	6	0,0573
7	0,0539	7	0,054	7	0,0549	7	0,0546	7	0,054	7	0,054	7	0,0536	7	0,0549	7	0,0526
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
15	0,0067	15	0,0068	15	0,0067	15	0,0065	15	0,0063	15	0,0063	15	0,0067	15	0,0053	15	0,0054

Keterangan :

Kondisi 1 = banyak data hilang 0,5%

Kondisi 2 = banyak data hilang 1%

Kondisi 3 = banyak data hilang 1,5%

FOI = Forward Imputation

MOI = Mode Imputation

D = Dimensi

AC = Akar ciri

## Lampiran 12. Akar Ciri, Proporsi Keragaman dan Proporsi Keragaman Kumulatif pada Data Lengkap Hasil Imputasi

### ➤ DATA 1

Data	Kondisi	Metode Imputasi	Dimensi	AC	PK (%)	PKK (%)
1	1	FOI	1	0,7333	73,33	73,33
			2	0,0923	9,23	82,56
			3	0,0772	7,72	90,28
		MOI	1	0,7333	73,33	73,33
			2	0,0923	9,23	82,56
	2	MEI	1	0,7333	73,33	73,33
			2	0,0923	9,23	82,56
			3	0,0772	7,72	90,28
		FOI	1	0,7331	73,33	73,33
			2	0,0924	9,23	82,56
			3	0,0772	7,72	90,28
	3	MOI	1	0,7109	71,09	71,09
			2	0,0909	9,09	80,18
			1	0,7109	71,09	71,09
		MEI	2	0,0909	9,09	80,18
			1	0,7424	74,24	74,24
	10	FOI	2	0,0916	9,16	83,4
			1	0,7425	74,25	74,25
		MOI	2	0,0915	9,15	83,4
			1	0,7425	74,25	74,25
		MEI	2	0,0915	9,15	83,4
			1	0,3398	33,98	33,98
		FOI	2	0,146	14,6	48,58
			3	0,1253	12,53	61,11
			4	0,1057	10,57	71,68
			5	0,0993	9,93	81,61
		MOI	1	0,3048	30,48	30,48
			2	0,2301	23,01	53,49
			3	0,1188	11,88	65,37
			4	0,1049	10,49	75,86
			5	0,0925	9,25	85,11
		MEI	1	0,3057	30,57	30,57
			2	0,2291	22,91	53,48
			3	0,1196	11,96	65,44
			4	0,103	10,3	75,74
			5	0,0887	8,87	84,61
	2	FOI	1	0,3049	30,49	30,49
			2	0,2106	21,06	51,55
			3	0,1189	11,89	63,44
			4	0,1063	10,63	74,07
		MOI	5	0,0908	9,08	83,15
			1	0,3068	30,68	30,68
			2	0,2116	21,16	51,84
			3	0,1182	11,82	63,66
			4	0,1057	10,57	74,23
		MEI	5	0,0894	8,94	83,17
			1	0,3068	30,68	30,68
			2	0,2116	21,16	51,84
			3	0,1182	11,82	63,66
			4	0,1057	10,57	74,23
	3	FOI	5	0,0894	8,94	83,17
			1	0,3374	33,74	33,74
			2	0,1587	15,87	49,61
			3	0,1167	11,67	61,28
			4	0,106	10,6	71,88
			5	0,1017	10,17	82,05

		MOI	1	0,3258	32,58	32,58
			2	0,1632	16,32	48,9
			3	0,1193	11,93	60,83
			4	0,11	11	71,83
			5	0,0916	9,16	80,99
		MEI	1	0,3316	33,16	33,16
			2	0,1548	15,48	48,64
			3	0,1184	11,84	60,48
			4	0,1137	11,37	71,85
			5	0,0921	9,21	81,06

Keterangan :

- Kondisi 1 = persentase data hilang 0,5%  
 Kondisi 2 = persentase data hilang 1%  
 Kondisi 3 = persentase data hilang 1,5%  
 AC = Akar Ciri  
 PK = Proporsi Keragaman  
 PKK = Proporsi Keragaman Kumulatif



## Lampiran 13. Nilai Skor Komponen Objek pada Data Lengkap Hasil Imputasi

Data 1

Kondisi	Objek	FOI			MOI			MOI		
		Dimensi			Dimensi			Dimensi		
		1	2	3	1	2	3	1	2	3
1	1	-0,24	-0,24	-0,24	0,04	-0,14	0,04	-0,24	0,04	-0,14
	2	-0,24	-0,24	-0,24	0,04	-0,14	0,04	-0,24	0,04	-0,14
	3	-0,24	-0,24	-0,24	0,04	-0,14	0,04	-0,24	0,04	-0,14
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	98	-0,18	-0,18	-0,18	0,04	-0,05	0,04	-0,18	0,04	-0,05
	99	-0,24	-0,24	-0,24	0,04	-0,14	0,04	-0,24	0,04	-0,14
	100	-0,17	-0,17	-0,17	-0,13	-0,38	-0,13	-0,17	-0,13	-0,38
2	1	-0,24	0,03	-0,14	-0,28	0,08	•	-0,28	0,08	•
	2	-0,24	0,03	-0,14	-0,28	0,08	•	-0,28	0,08	•
	3	-0,24	0,03	-0,14	-0,28	0,08	•	-0,28	0,08	•
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	98	-0,18	0,04	-0,06	-0,11	-0,15	•	-0,11	-0,15	•
	99	-0,24	0,03	-0,14	-0,28	0,08	•	-0,28	0,08	•
	100	-0,17	-0,13	-0,37	-0,05	-0,08	•	-0,05	-0,08	•
3	1	-0,27	0,06	•	-0,27	0,06	•	-0,27	0,06	•
	2	-0,27	0,06	•	-0,27	0,06	•	-0,27	0,06	•
	3	-0,27	0,06	•	-0,27	0,06	•	-0,27	0,06	•
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	98	-0,13	-0,14	•	-0,13	-0,13	•	-0,13	-0,13	•
	99	-0,27	0,06	•	-0,27	0,06	•	-0,27	0,06	•
	100	-0,05	-0,07	•	-0,05	-0,07	•	-0,05	-0,07	•

•

•

•

Data 10

Kondisi	Objek	FOI					MOI					MOI				
		Dimensi					Dimensi					Dimensi				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	1	-0,44	0,35	1,37	-0,63	-0,73	0,12	-0,68	-0,24	1,45	0,78	0,08	0,65	0,12	1,42	0,9
	2	0,74	-0,05	-0,06	1,09	0,24	0,7	0,6	0,35	-0,48	-0,69	-0,69	-0,63	-0,34	-0,51	-0,75
	3	0,28	-0,6	1,15	0,03	0,76	0,3	0,2	0,06	1,22	-0,77	-0,31	-0,23	-0,25	1,26	-0,73
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	29	1,01	0,23	0,4	-0,41	0,25	0,73	0,16	0,23	0,41	0,42	-0,73	-0,21	-0,24	0,27	0,43
	30	0,63	-0,25	-1,24	-0,34	1,34	-0,84	0,56	0,6	-0,36	-0,78	-0,84	-0,6	-0,62	-0,41	-0,83
	31	0,75	-0,61	-0,39	0,73	0,62	-0,43	0,6	-0,25	-0,84	-1,2	-0,31	-0,23	0,6	-0,91	-0,42
2	1	0,78	-0,11	-0,15	-1,61	-0,37	-0,77	-0,15	-0,25	-1,6	0,32	-0,77	-0,15	-0,25	-1,6	0,32
	2	-0,68	-0,02	0,11	0,74	0,77	0,68	-0,01	0,11	0,78	-0,73	0,68	-0,01	0,11	0,78	-0,73
	3	-0,24	-0,43	-0,66	-0,71	1,06	0,26	-0,44	-0,7	-0,57	-1,05	0,26	-0,44	-0,7	-0,57	-1,05
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	29	-0,78	0,13	-0,7	-0,44	-0,23	0,78	0,13	-0,7	-0,4	0,22	0,78	0,13	-0,7	-0,4	0,22
	30	-1	-0,2	-0,41	0,6	0,43	0,99	-0,13	-0,35	0,63	-0,44	0,99	-0,13	-0,35	0,63	-0,44
	31	0,82	-0,43	0,1	0,81	0,61	0,83	-0,42	0,12	0,82	-0,56	0,83	-0,42	0,12	0,82	-0,56
3	1	-0,64	-0,57	1,46	-0,4	0,68	0,68	0,52	-1,17	0,89	-0,72	0,65	0,53	1,22	0,9	-0,83
	2	0,94	0,2	-0,31	-0,04	-0,8	-0,93	-0,19	0,22	-0,21	0,8	-0,95	-0,09	-0,25	-0,16	0,86
	3	0,24	0,38	0,98	-0,53	-0,9	-0,2	-0,14	-0,77	0,78	1,13	-0,27	-0,24	0,7	0,84	1,03
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	29	1,22	-0,68	0,16	0,03	0,43	-0,99	0,62	-0,56	0,3	0,1	-1,03	0,45	0,57	0,22	0
	30	0,69	0,84	-1,04	-0,84	-0,46	-0,72	-0,86	-0,11	-1,06	0,82	-0,67	-0,76	0,05	-1,28	1
	31	0,7	0,67	-0,36	0,36	-0,87	-0,71	-0,66	0,56	0,03	0,82	-0,71	-0,57	-0,56	0,01	0,89

Keterangan :

FOI = Forward Imputation

MOI = Mode Imputation

MEI = Median Imputation

Kondisi 1 = persentase data hilang 0,5%

Kondisi 2 = persentase data hilang 1%

Kondisi 3 = persentase data hilang 1,5%

## Lampiran 14. Nilai Kuantifikasi pada Data Lengkap Hasil Imputasi

### ➤ DATA 1

Kondisi 1 :

*Forward Imputation*

Variable: X1			
Type: Ordinal	Missing: 0		
Category:	Marginal Frequency	Quantification	
1	4	-4,90	
2	8	,13	
3	88	,21	
4	0	,00	

Single Category Coordinates		Multiple Category Coordinates					
Category	Dimension	Category	Dimension				
1	2	3	1	2	3		
1	4,03	1,60	-,80	1	4,03	1,59	-,81
2	-,11	-,04	,02	2	-,10	,07	,26
3	-,17	-,07	,03	3	-,17	-,08	,01
4	,00	,00	,00	4	,00	,00	,00

Variable: X14			
Type: Ordinal	Missing: 0		
Category:	Marginal Frequency	Quantification	
1	4	-4,89	
2	4	-,12	
3	89	,21	
4	3	,37	

Single Category Coordinates		Multiple Category Coordinates					
Category	Dimension	Category	Dimension				
1	2	3	1	2	3		
1	4,50	,76	1,35	1	4,49	,83	1,34
2	-,11	,02	,03	2	,26	-1,03	,14
3	-,20	-,03	-,06	3	-,20	,01	-,06
4	-,34	-,06	-,10	4	-,29	,00	-,31

Mode Imputation			
Variable: X1			
Type: Ordinal	Missing: 0		
Category:	Marginal Frequency	Quantification	
1	4	-4,90	
2	8	,13	
3	88	,21	
4	0	,00	

Single Category Coordinate		Multiple Category Coordinates					
Category	Dimension	Category	Dimension				
1	2	3	1	2	3		
1	4,03	1,60	-,80	1	4,03	1,59	-,81
2	-,11	-,04	,02	2	-,10	,07	,26
3	-,17	-,07	,03	3	-,17	-,08	,01
4	,00	,00	,00	4	,00	,00	,00

Variable: X14			
Type: Ordinal	Missing: 0		
Category:	Marginal Frequency	Quantification	
1	4	-4,89	
2	4	-,12	
3	89	,21	
4	3	,37	

Single Category Coordinates		Multiple Category Coordinates					
Category	Dimension	Category	Dimension				
1	2	3	1	2	3		
1	4,50	,76	1,35	1	4,49	,83	1,34
2	-,11	,02	,03	2	,26	-1,03	,14
3	-,20	-,03	-,06	3	-,20	,01	-,06
4	-,34	-,06	-,10	4	-,29	,00	-,31

## Lampiran 14 (Lanjutan)

### *Median Imputation*

Variable: X1		
Type: Ordinal	Missing:	0
Category	Marginal Frequency	Quantification
1	4	-4,90
2	8	,13
3	88	,21
4	0	,00

Single Category Coordinates			Multiple Category Coordinates				
Category	Dimension		Category	Dimension			
1	2	3	1	2	3		
1	4,03	1,60	-,80	1	4,03	1,59	-,81
2	-,11	-,04	,02	2	-,10	,07	,26
3	-,17	-,07	,03	3	-,17	-,08	,01
4	,00	,00	,00	4	,00	,00	,00

⋮	⋮	⋮
---	---	---

Variable: X14		
Type: Ordinal	Missing:	0
Category	Marginal Frequency	Quantification
1	4	-4,89
2	4	-,12
3	89	,21
4	3	,37

Single Category Coordinates			Multiple Category Coordinates				
Category	Dimension		Category	Dimension			
1	2	3	1	2	3		
1	.4,50	.76	1,35	1	4,49	,83	1,34
2	,11	,02	,03	2	,26	-,103	,14
3	-,20	-,03	-,06	3	-,20	,01	-,06
4	-,34	-,06	-,10	4	-,29	,00	-,31

dst.

## Lampiran 15. Nilai Bobot Peubah pada Data Lengkap Hasil Imputasi

**DATA 1**

Kondisi	Peubah	FOI			MOI			MEI		
		Dimensi			Dimensi			Dimensi		
		1	2	3	1	2	3	1	2	3
1	X1	-0,824	-0,326	0,164	-0,824	-0,326	0,164	-0,824	-0,326	0,164
	X2	-0,694	0,508	0,241	-0,694	0,508	0,241	-0,694	0,508	0,241
	X3	-0,912	0,16	0,307	-0,912	0,16	0,307	-0,912	0,16	0,307
	:	:	:	:	:	:	:	:	:	:
	X12	-0,733	0,513	-0,34	-0,733	0,513	-0,34	-0,733	0,513	-0,34
	X13	-0,682	-0,075	-0,481	-0,682	-0,075	-0,481	-0,682	-0,075	-0,481
	X14	-0,922	-0,156	-0,277	-0,922	-0,156	-0,277	-0,922	-0,156	-0,277
	:	:	:	:	:	:	:	:	:	:
2	X1	-0,824	-0,325	0,165	-0,819	-0,326	0,165	-0,819	-0,326	0,165
	X2	-0,694	0,509	0,24	-0,689	0,501	0,24	-0,689	0,501	0,24
	X3	-0,912	0,161	0,306	-0,685	0,183	0,306	-0,685	0,183	0,306
	:	:	:	:	:	:	:	:	:	:
	X12	-0,733	0,512	-0,342	-0,742	0,499	-0,342	-0,742	0,499	-0,342
	X13	-0,682	-0,079	-0,48	-0,695	-0,105	-0,48	-0,695	-0,105	-0,48
	X14	-0,922	-0,156	-0,277	-0,92	-0,156	-0,277	-0,92	-0,156	-0,277
	:	:	:	:	:	:	:	:	:	:
3	X1	-0,82	-0,313	0,16	-0,82	-0,313	0,16	-0,82	-0,313	0,16
	X2	-0,693	0,522	0,24	-0,693	0,524	0,24	-0,693	0,524	0,24
	X3	-0,915	0,183	0,306	-0,915	0,183	0,306	-0,915	0,183	0,306
	:	:	:	:	:	:	:	:	:	:
	X12	-0,741	0,47	0,16	-0,741	0,469	0,16	-0,741	0,469	0,16
	X13	-0,732	-0,156	0,306	-0,732	-0,155	0,306	-0,732	-0,155	0,306
	X14	-0,926	-0,191	0,277	-0,926	-0,191	0,277	-0,926	-0,191	0,277
	:	:	:	:	:	:	:	:	:	:

**DATA 10**

K o n d i s i	P e u b a h	FOI					MOI					MEI				
		Dimensi					Dimensi					Dimensi				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	X1	-0,539	0,472	0,047	-0,431	-0,429	0,61	0,349	0,038	-0,298	-0,547	0,32	0,663	0,139	0,084	0,545
	X2	-0,465	0,43	-0,017	-0,546	-0,447	0,528	0,311	0,035	-0,324	-0,656	0,238	0,655	0,272	-0,107	0,543
	X3	-0,416	0,684	0,26	0,16	0,342	0,483	0,609	-0,403	0,034	0,309	0,391	0,489	-0,626	-0,001	0,004
	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
	X13	-0,375	-0,032	-0,644	-0,435	0,426	-0,1	-0,785	-0,481	0,133	-0,316	0,422	0,468	-0,026	-0,549	-0,215
	X14	-0,63	-0,449	-0,344	0,131	-0,103	0,509	-0,453	0,341	0,047	-0,066	0,681	-0,066	0,527	-0,172	0,098
	X15	-0,539	-0,607	-0,163	0,11	-0,055	0,443	-0,568	0,35	0,198	0,052	0,485	-0,142	0,637	-0,088	-0,23
	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
2	X1	0,61	0,349	0,036	-0,298	-0,547	-0,617	0,34	0,02	-0,339	0,521	-0,617	0,34	0,02	-0,339	0,521
	X2	0,528	0,311	0,035	-0,324	-0,656	-0,534	0,307	0,025	-0,37	0,63	-0,534	0,307	0,025	-0,37	0,63
	X3	0,483	0,609	-0,403	0,034	0,309	-0,498	0,592	-0,41	0,062	-0,313	-0,498	0,592	-0,41	0,062	-0,313
	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
	X13	-0,1	-0,785	-0,481	0,133	-0,316	0,143	-0,789	-0,463	0,119	0,323	0,143	-0,789	-0,463	0,119	0,323
	X14	0,504	-0,453	0,341	0,347	-0,066	-0,499	-0,46	0,358	0,326	0,091	-0,499	-0,46	0,358	0,326	0,091
	X15	0,433	0,568	0,35	0,198	0,052	-0,42	-0,577	0,359	0,181	0,038	-0,42	-0,577	0,359	0,181	-0,038
	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
3	X1	-0,55	-0,456	0,063	-0,132	0,586	0,612	0,271	-0,4	-0,199	-0,477	0,607	0,304	0,403	-0,235	-0,442
	X2	-0,461	-0,419	0,039	-0,136	0,711	0,458	0,339	-0,244	-0,12	-0,71	0,453	0,335	0,256	-0,195	-0,693
	X3	-0,402	-0,676	-0,049	-0,29	-0,397	0,381	0,754	-0,074	-0,143	0,321	0,41	0,749	0,091	-0,054	0,328
	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
	X13	-0,286	0,514	-0,529	-0,538	0,164	0,236	-0,639	-0,318	-0,608	0,082	0,295	-0,514	0,242	-0,709	0,155
	X14	-0,607	0,441	-0,204	0,346	0,043	0,615	-0,433	0,4	0,013	-0,065	0,614	-0,454	-0,367	-0,015	-0,077
	X15	-0,534	0,558	0,021	0,317	-0,036	0,521	-0,547	0,291	0,195	0,01	0,537	-0,57	-0,242	0,181	0
	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:

Keterangan :

FOI = Forward Imputation

MOI = Mode Imputation

MEI = Median Imputation

Kondisi 1 = persentase data hilang 0,5%

Kondisi 2 = persentase data hilang 1%

Kondisi 3 = persentase data hilang 1,5%

## Lampiran 16. Tes Konvergensi pada Data Lengkap Hasil Imputasi

### DATA 1

#### Kondisi 1:

##### *Forward Imputation*

* The History of Iterations *					
Iteration	Total Fit	Total Loss	Multiple Loss	Single Loss	Iteration Change
1	,7937503	2,2062497	2,0869163	,1193334	,0474961
2	,8387273	2,1612727	2,0027608	,1585119	,0449770
3	,8605987	2,1394013	1,9697277	,1696736	,0218715
.	.	.	.	.	.
18	,9027316	2,0972684	2,0238827	,0733857	,0000156
19	,9027402	2,0972598	2,0244118	,0728480	,0000086

The iterative process stops because the convergence test value is reached.

#### Kondisi 2 :

##### *Forward Imputation*

* The History of Iterations *					
Iteration	Total Fit	Total Loss	Multiple Loss	Single Loss	Iteration Change
1	,7944973	2,2055027	2,0831197	,1223830	,0482271
2	,8398310	2,1601690	1,9979554	,1622136	,0453337
3	,8613862	2,1386138	1,9657601	,1728537	,0215552
.	.	.	.	.	.
19	,9026917	2,0973083	2,0225947	,0747136	,0000101
20	,9026975	2,0973025	2,0230722	,0742303	,0000058

The iterative process stops because the convergence test value is reached.

##### *Mode Imputation*

* The History of Iterations *					
Iteration	Total Fit	Total Loss	Multiple Loss	Single Loss	Iteration Change
1	,7937503	2,2062497	2,0869163	,1193334	,0474961
2	,8387273	2,1612727	2,0027608	,1585119	,0449770
3	,8605987	2,1394013	1,9697277	,1696736	,0218715
.	.	.	.	.	.
18	,9027316	2,0972684	2,0238827	,0733857	,0000156
19	,9027402	2,0972598	2,0244118	,0728480	,0000086

The iterative process stops because the convergence test value is reached.

##### *Mode Imputation*

* The History of Iterations *					
Iteration	Total Fit	Total Loss	Multiple Loss	Single Loss	Iteration Change
1	,7067749	1,2932251	1,2037483	,0894768	,0474233
2	,7539071	1,2460929	1,1256463	,1204466	,0471322
3	,7754975	1,2245025	1,0939793	,1305232	,0215905
.	.	.	.	.	.
18	,8017639	1,1982361	1,1583773	,0398588	,0000134
19	,8017728	1,1982272	1,1589421	,0392851	,0000089

The iterative process stops because the convergence test value is reached.

##### *Median Imputation*

* The History of Iterations *					
Iteration	Total Fit	Total Loss	Multiple Loss	Single Loss	Iteration Change
1	,7937503	2,2062497	2,0869163	,1193334	,0474961
2	,8387273	2,1612727	2,0027608	,1585119	,0449770
3	,8605987	2,1394013	1,9697277	,1696736	,0218715
.	.	.	.	.	.
18	,9027316	2,0972684	2,0238827	,0733857	,0000156
19	,9027402	2,0972598	2,0244118	,0728480	,0000086

The iterative process stops because the convergence test value is reached.

##### *Median Imputation*

* The History of Iterations *					
Iteration	Total Fit	Total Loss	Multiple Loss	Single Loss	Iteration Change
1	,7067749	1,2932251	1,2037483	,0894768	,0474233
2	,7539071	1,2460929	1,1256463	,1204466	,0471322
3	,7754975	1,2245025	1,0939793	,1305232	,0215905
.	.	.	.	.	.
18	,8017639	1,1982361	1,1583773	,0398588	,0000134
19	,8017728	1,1982272	1,1589421	,0392851	,0000089

The iterative process stops because the convergence test value is reached.

## Lampiran 16 (Lanjutan)

### Kondisi 3 :

#### *Forward Imputation*

#### *Mode Imputation*

* The History of Iterations *						
Iteration	Total Fit	Total Loss	Multiple Loss	Single Loss	Iteration Change	Iteration
1	,7301788	1,2698212	1,1839891	,0858320	,0526510	
2	,7824593	1,2175407	1,1038408	,1136999	,0522805	
3	,8061593	1,1938407	1,0748246	,1190161	,0237000	
.	.	.	.	.	.	
14	,8340130	1,1659870	1,1325629	,0334241	,0000171	
15	,8340212	1,1659788	1,1332043	,0327745	,0000082	

The iterative process stops because the convergence test value is reached.

* The History of Iterations *						
Iteration	Total Fit	Total Loss	Multiple Loss	Single Loss	Iteration Change	Iteration
1	,7294112	1,2705888	1,1874572	,0831316	,0530746	
2	,7821359	1,2178641	1,1077855	,1100786	,0527248	
3	,8059588	1,1940412	1,0789420	,1150992	,0238229	
.	.	.	.	.	.	
14	,8339749	1,1660251	1,1335719	,0324531	,0000134	
15	,8339812	1,1660188	1,1341083	,0319105	,0000063	

The iterative process stops because the convergence test value is reached.

#### *Median Imputation*

* The History of Iterations *						
Iteration	Total Fit	Total Loss	Multiple Loss	Single Loss	Iteration Change	Iteration
1	,7294112	1,2705888	1,1874572	,0831316	,0530746	
2	,7821359	1,2178641	1,1077855	,1100786	,0527248	
3	,8059588	1,1940412	1,0789420	,1150992	,0238229	
4	,8176322	1,1823678	1,0766288	,1057389	,0116734	
.	.	.	.	.	.	
12	,8339326	1,1660674	1,1315619	,0345055	,0000635	
13	,8339616	1,1660384	1,1327694	,0332690	,0000290	
14	,8339749	1,1660251	1,1335719	,0324531	,0000134	
15	,8339812	1,1660188	1,1341083	,0319105	,0000063	

The iterative process stops because the convergence test value is reached.

dst.

Keterangan :

Kondisi 1 = persentase data hilang 0,5%

Kondisi 2 = persentase data hilang 1%

Kondisi 3 = persentase data hilang 1,5%

## Lampiran 17. Hasil Perhitungan $\cos(\underline{a}_t, \underline{a}_\theta)$

Data 1

Dimensi	$\cos(a_t, a_\theta)$								
	Kondisi 1 (0.5%)			Kondisi 2 (1%)			Kndisi 3 (1.5%)		
	FOI	MOI	MEI	FOI	MOI	MEI	FOI	MOI	MEI
1	1	1	1	0,999999955	0,997523292	0,997523292	0,999879642	0,999879403	0,999879403
2	0,999973	0,999973	0,999973	0,999993427	0,997934471	0,997934471	0,991980998	0,991978183	0,991978183
3	0,999976	0,999976	0,999976	0,999994103					

Data 10

Dimensi	$\cos(a_t, a_\theta)$								
	Kondisi 1 (0.5%)			Kondisi 2 (1%)			Kndisi 3 (1.5%)		
	FOI	MOI	MEI	FOI	MOI	MEI	FOI	MOI	MEI
1	0,998816118	0,983859493	0,98074677	0,96140521	0,967064803	0,967064803	0,997895641	0,996938616	0,998400037
2	0,909201809	0,725238496	0,718662564	0,88357908	0,878626525	0,878626525	0,953855736	0,958284064	0,966951501
3	0,934616364	0,62363539	0,569711252	0,814322625	0,801847575	0,801847575	0,992374302	0,725830357	0,742867057
4	0,984729123	0,547560971	0,565924661	0,732266803	0,747557711	0,747557711	0,785944899	0,690947572	0,713940093
5	0,989290345	0,852038869	0,862725899	0,894395034	0,903677791	0,903677791	0,83789786	0,889283485	0,911645591

## Lampiran 18. Hasil Perhitungan RMSE<sub>t</sub>

Data 1

Dimensi	RMSE (t)								
	Kondisi 1 (0.5%)			Kondisi 2 (1%)			Kndisi 3 (1.5%)		
	FOI	MOI	MEI	FOI	MOI	MEI	FOI	MOI	MEI
1	0,003606	0,003606	0,003606	0,001732	0,080293	0,080293	0,047958	0,047202	0,047202
2	0,01005	0,01005	0,01005	0,004583	0,146404	0,146404	0,139682	0,135761	0,135761
3	0,017464	0,017464	0,017464	0,003873					

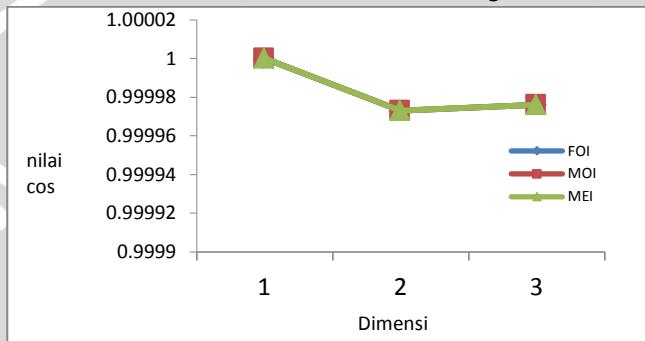
Data 10

Dimensi	RMSE (t)								
	Kondisi 1 (0.5%)			Kondisi 2 (1%)			Kondisi 3 (1.5%)		
	FOI	MOI	MEI	FOI	MOI	MEI	FOI	MOI	MEI
1	0,08128	0,456526	0,48159	0,274508	0,294059	0,294059	0,090892	0,093464	0,110117
2	0,374299	0,796853	0,802036	0,499965	0,516271	0,516271	0,235379	0,224564	0,192253
3	0,443265	0,928134	0,941493	0,998085	1,013534	1,013534	0,213753	0,791039	0,778996
4	0,192781	0,786893	0,795128	0,809932	0,798121	0,798121	0,762515	0,804517	0,72587
5	0,225868	0,788156	0,856286	0,951171	0,957456	0,957456	0,704552	0,4761	0,412158

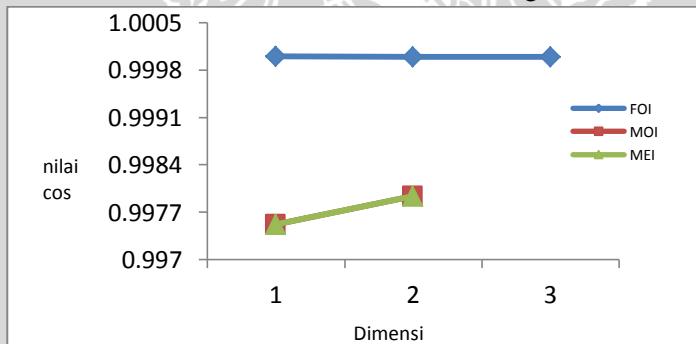
**Lampiran19. Diagram Garis Metode Imputasi Paling Baik Berdasarkan Nilai  $\cos(\underline{a}_t, \underline{a}_\theta)$  Terbesar**

**DATA 1**

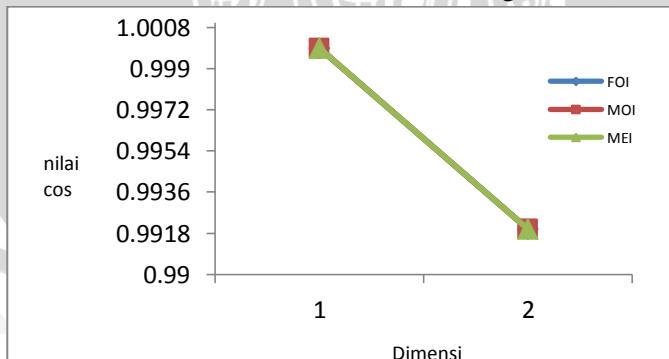
Kondisi 1 (Percentase data hilang = 0,5%)



Kondisi 2 (Percentase data hilang = 1%)

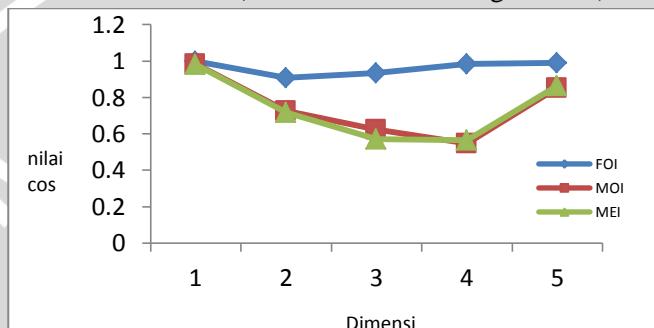


Kondisi 3 (Percentase data hilang = 1,5%)

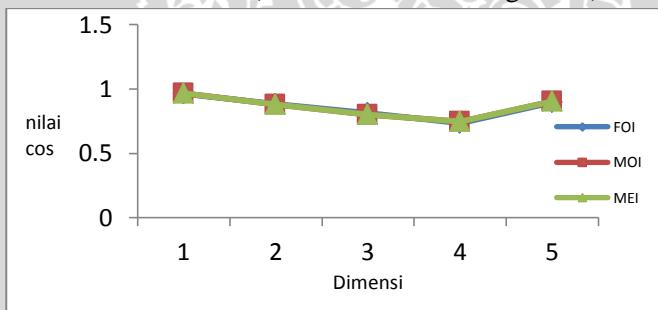


## DATA 10

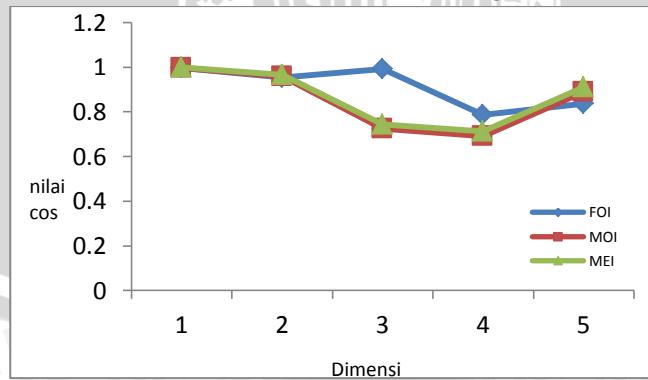
Kondisi 1 (Persentase data hilang = 0,5%)



Kondisi 2 (Persentase data hilang = 1%)



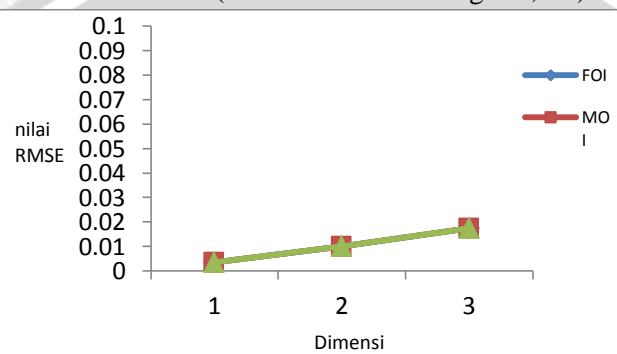
Kondisi 3 (Persentase data hilang = 1,5%)



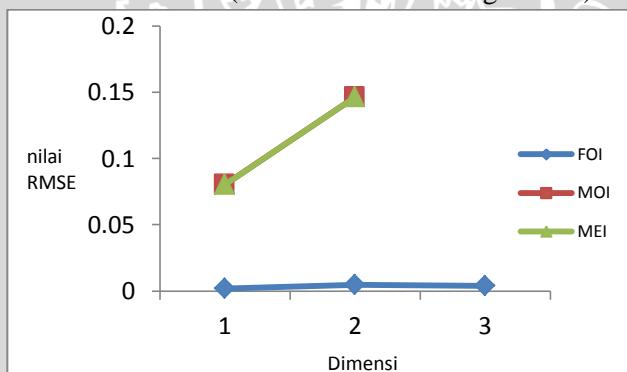
## Lampiran 20. Diagram Garis Metode Imputasi Paling Baik Berdasarkan Nilai RMSE<sub>t</sub> Terkecil

DATA 1 :

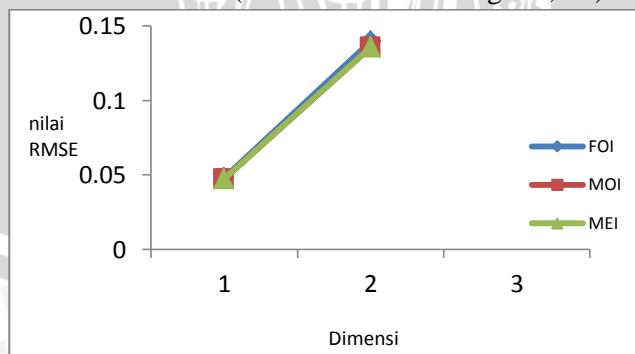
Kondisi 1 (Persentase data hilang = 0,5%)



Kondisi 2 (Persentase data hilang = 01%)



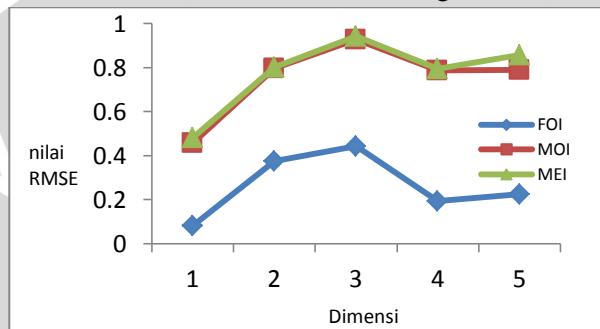
Kondisi 3 (Persentase data hilang = 1,5%)



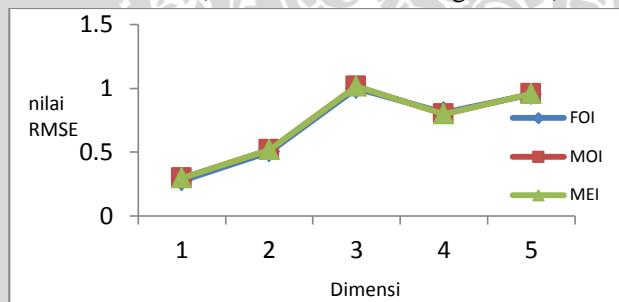
## Lampiran 20 (Lanjutan)

DATA 10 :

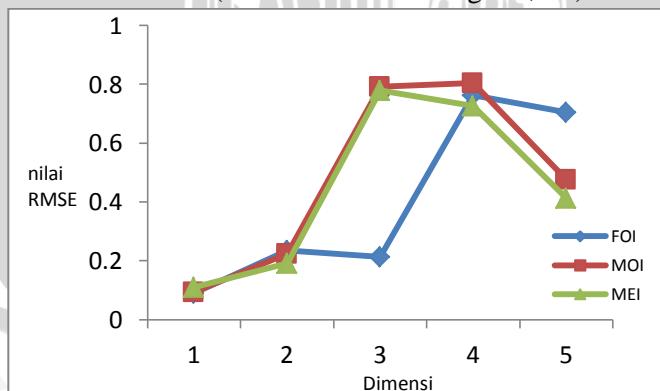
Kondisi 1 (Persentase data hilang = 0,5%)



Kondisi 2 (Persentase data hilang = 01%)

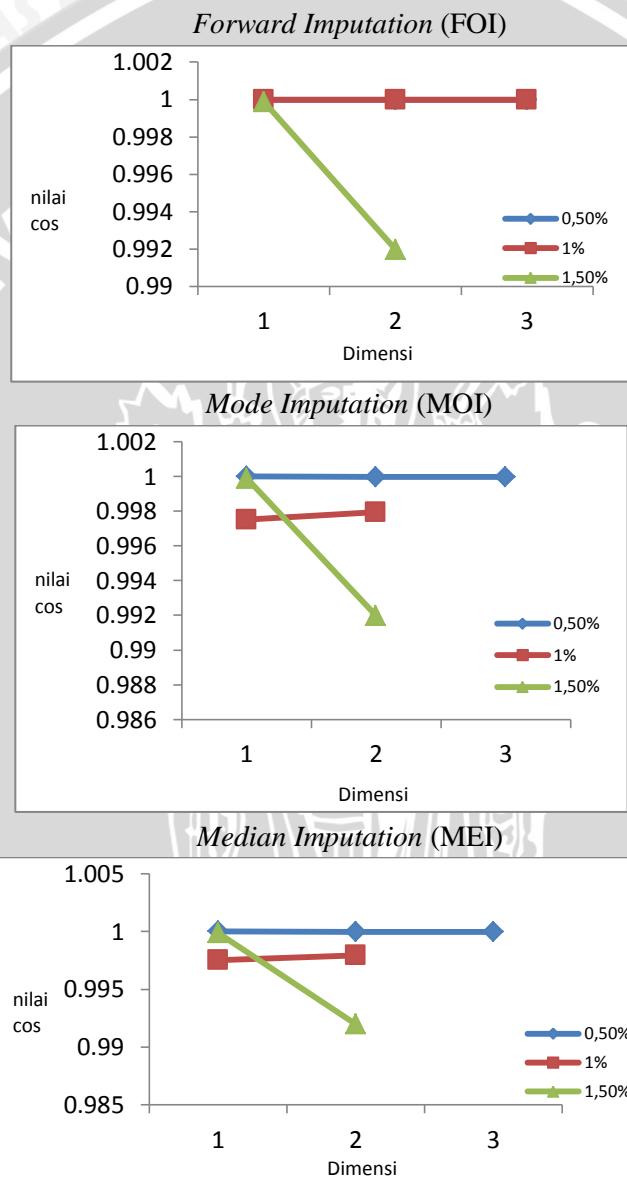


Kondisi 3 (Persentase data hilang = 1,5%)



**Lampiran 21. Diagram Garis Pengaruh Semakin Besar Persentase Data Hilang dalam Menghasilkan Nilai  $\cos(\underline{a}_t, \underline{a}_\theta)$**

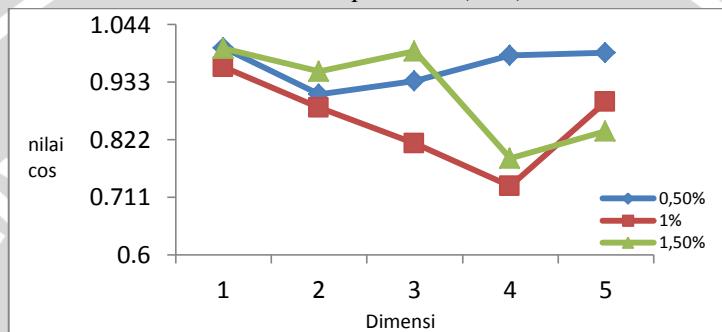
DATA 1 :



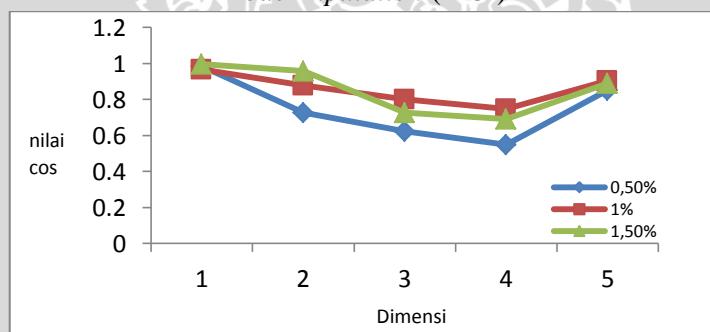
## Lampiran 21 (Lanjutan)

DATA 10 :

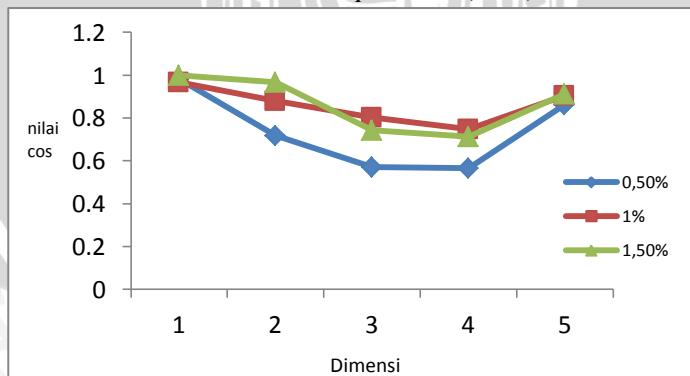
*Forward Imputation (FOI)*



*Mode Imputation (MOI)*

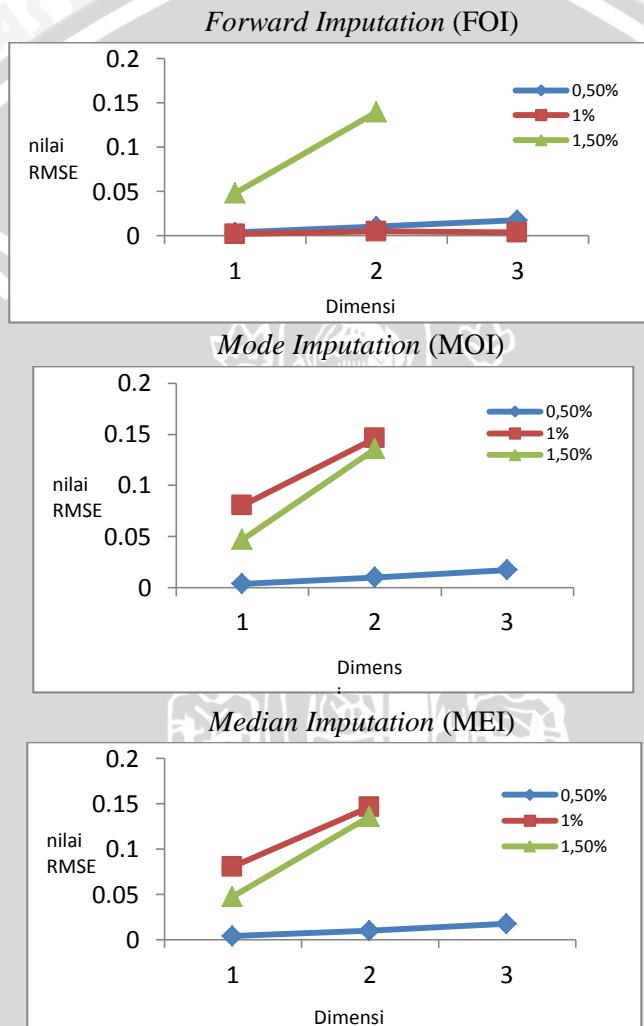


*Median Imputation (MEI)*



**Lampiran 22. Diagram Garis Pengaruh Semakin Besar Persentase Data Hilang dalam Menghasilkan Nilai RMSE<sub>t</sub>**

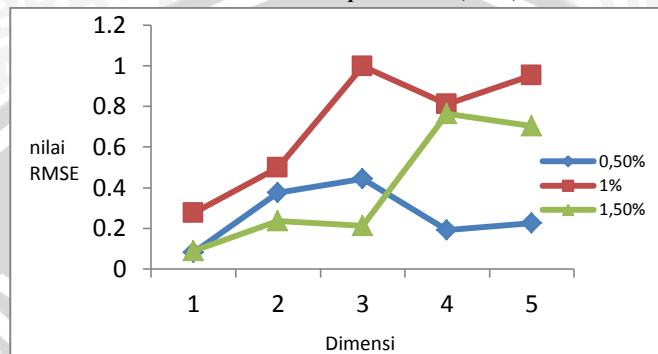
DATA 1 :



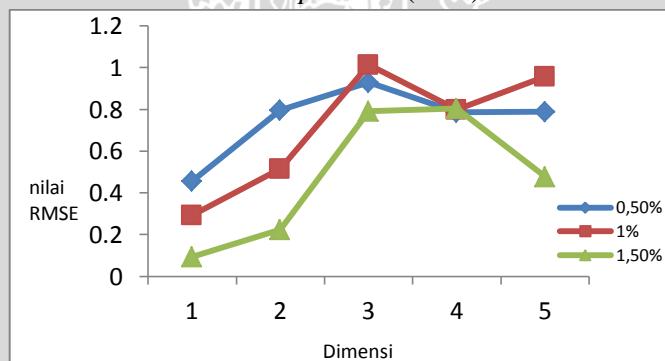
## Lampiran 22 (Lanjutan)

### DATA 10 :

*Forward Imputation (FOI)*



*Mode Imputation (MOI)*



*Median Imputation (MEI)*

