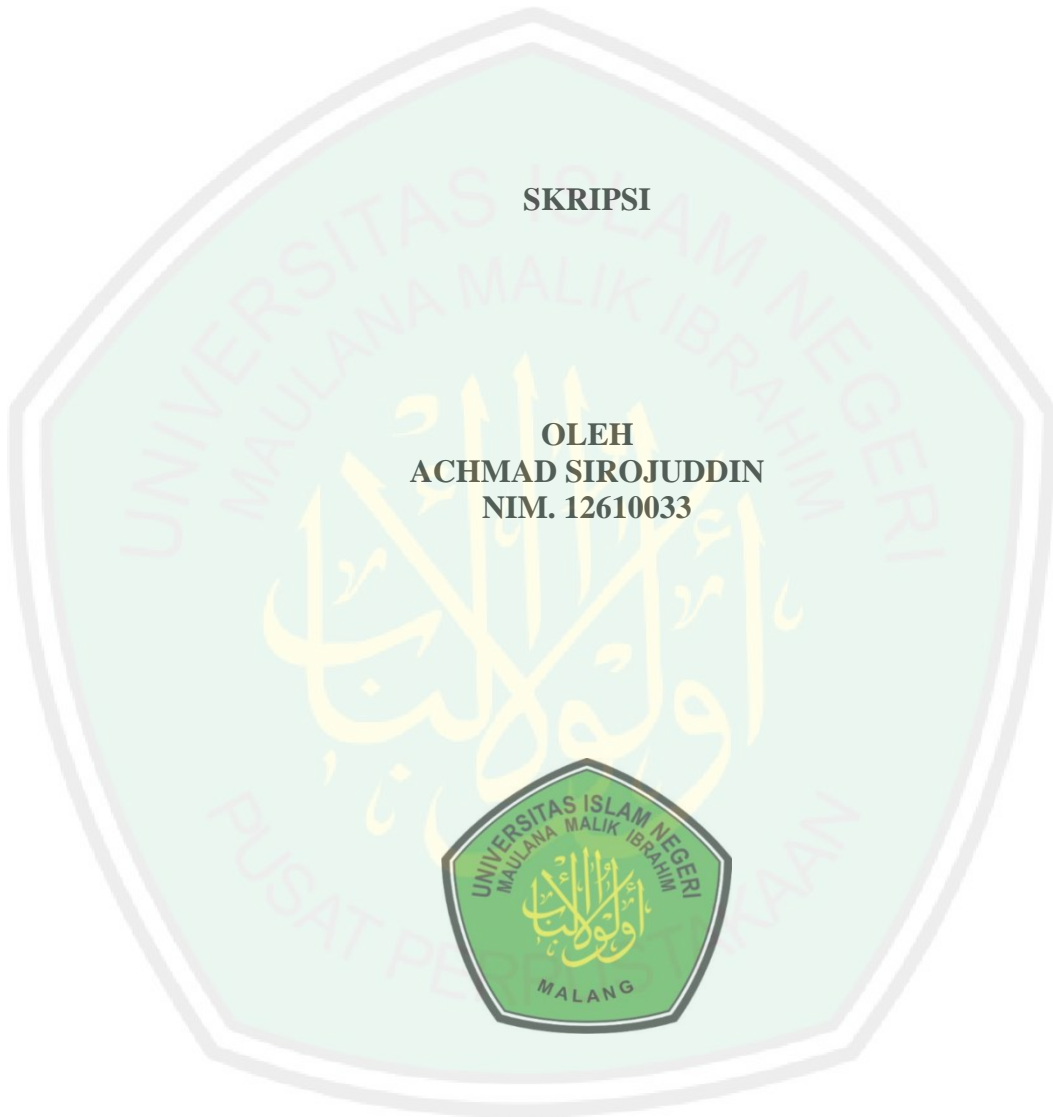


**ANALISIS *CLUSTER* PADA KABUPATEN/KOTA
DI PROVINSI JAWA TIMUR
BERDASARKAN INDIKATOR INDEKS PEMBANGUNAN MANUSIA**

SKRIPSI

**OLEH
ACHMAD SIROJUDDIN
NIM. 12610033**



**JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2016**

**ANALISIS *CLUSTER* PADA KABUPATEN/KOTA
DI PROVINSI JAWA TIMUR
BERDASARKAN INDIKATOR INDEKS PEMBANGUNAN MANUSIA**

SKRIPSI

**Diajukan Kepada
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Sains (S.Si)**

**Oleh
Achmad Sirojuddin
NIM. 12610033**

**JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2016**

ANALISIS *CLUSTER* PADA KABUPATEN/KOTA
DI PROVINSI JAWA TIMUR
BERDASARKAN INDIKATOR INDEKS PEMBANGUNAN MANUSIA


SKRIPSI

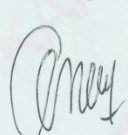
Oleh
Achmad Sirojuddin
NIM. 12610033

Telah Diperiksa dan Disetujui untuk Diuji
Tanggal 15 Desember 2016


Pembimbing I,

Pembimbing II,


Dr. Sri Harini, M.Si
NIP. 19731014 2001 12 2 002


Mohammad Jamhuri, M.Si
NIP. 19810502 200501 1 004

Mengetahui,
Ketua Jurusan Matematika


Dr. Abdussakir, M.Pd
NIP. 19751006 200312 1 001

**ANALISIS CLUSTER PADA KABUPATEN/KOTA
DI PROVINSI JAWA TIMUR
BERDASARKAN INDIKATOR INDEKS PEMBANGUNAN MANUSIA**

SKRIPSI

Oleh
Achmad Sirojuddin
NIM. 12610033

Telah Dipertahankan di Depan Dewan Penguji Skripsi dan
Dinyatakan Diterima Sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Sains (S.Si)

Tanggal 27 Desember 2016

Penguji Utama	: Dr. H. Imam Sujarwo, M.Pd
Ketua Penguji	: Ir. Nanang Widodo, M.Si
Sekretaris Penguji	: Dr. Sri Harini, M.Si
Anggota Penguji	: Mohammad Jamhuri, M.Si

Mengetahui,
Ketua Jurusan Matematika

Dr. Abdussakir, M.Pd
NIP. 19751006 200312 1 001

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Achmad Sirojuddin

NIM : 12610033

Jurusan : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Analisis *Cluster* pada Kabupaten/Kota di Provinsi Jawa Timur

berdasarkan Indikator Indeks Pembangunan Manusia

menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar rujukan. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 18 Agustus 2016
Yang membuat pernyataan,



Achmad Sirojuddin
NIM. 12610033

MOTO

“Percaya dan yakin karena Allah Swt. selalu bersamaku”



PERSEMBAHAN

Penulis persembahkan skripsi ini kepada orang yang sangat penulis kasihi dan sayangi:

Ibunda Nur Miati dan ayahanda H. Zulkifli tercinta.

Saudara-saudara dan teman-teman yang senantiasa memberikan semangat, arahan, dan juga materi, sehingga penulis dapat berjuang menyelesaikan perkuliahan di Universitas Islam Negeri Maulana Malik Ibrahim Malang.



KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji syukur penulis panjatkan ke hadirat Allah Swt., karena berkat limpahan rahmat, hidayah, dan inayah-Nya, skripsi ini dapat diselesaikan dengan baik. Shalawat beserta salam semoga tetap tercurahkan kepada nabi Muhammad Saw., yang telah membawa manusia dari alam jahiliyah menuju alam yang berilmu seperti sekarang ini.

Skripsi yang berjudul “Analisis *Cluster* pada Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Indeks Pembangunan Manusia” ini penulis susun untuk memenuhi salah satu syarat menyelesaikan pendidikan program studi strata satu (S-1) di Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Dalam proses penyusunannya tidak terlepas dari bantuan berbagai pihak yang telah mendorong dan membimbing penulis, baik tenaga, ide-ide, maupun pemikiran. Oleh karena itu, dalam kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Prof. Dr. H. Mudjia Rahardjo, M.Si, selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. drh. Hj. Bayyinatul Muchtaromah, M.Si, selaku dekan Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Abdussakir, M.Pd, selaku ketua Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Dr. Sri Harini, M.Si, selaku dosen pembimbing I yang selalu sabar membimbing penulis dengan segala ilmu yang dimiliki serta senantiasa

memberikan doa, arahan, nasihat, dan motivasi dalam melakukan penelitian kepada penulis.

5. Mohammad Jamhuri, M.Si, selaku dosen pembimbing II yang telah memberikan bimbingan, arahan, dan berbagi ilmunya kepada penulis.
6. Segenap sivitas akademika Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang terutama seluruh dosen, terima kasih karena telah memberikan ilmunya selama masa perkuliahan.
7. Ibu dan bapak yang selalu memberikan perhatian, dukungan, materi, doa, semangat, kasih sayang, serta motivasi kepada penulis sehingga dapat menyelesaikan skripsi ini.
8. Seluruh teman-teman di Jurusan Matematika angkatan 2012 yang telah memberikan dukungan dan yang selalu mendoakan agar tetap kuat dalam menjalani semuanya.
9. Semua pihak yang secara langsung atau tidak langsung telah ikut memberikan bantuan dalam menyelesaikan skripsi ini.

Terakhir penulis berharap semoga skripsi ini dapat memberikan hal yang bermanfaat dan menambah wawasan bagi pembaca dan khususnya bagi penulis juga.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Malang, Desember 2016

Penulis

DAFTAR ISI

HALAMAN JUDUL	
HALAMAN PENGAJUAN	
HALAMAN PERSETUJUAN	
HALAMAN PENGESAHAN	
HALAMAN PERNYATAAN KEASLIAN TULISAN	
HALAMAN MOTO	
HALAMAN PERSEMBAHAN	
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR SIMBOL	xii
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xv
ABSTRAK	xvi
ABSTRACT	xvii
ملخص	xviii
 BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	5
1.6 Sistematika Penulisan	6
 BAB II KAJIAN PUSTAKA	
2.1 Analisis <i>Cluster</i>	7
2.2 Analisis Faktor	11
2.2.1 Pembentukan Matriks Varians dan Kovarians	12
2.2.2 Metode <i>Principal Component</i>	15
2.2.3 Kriteria Penentuan Jumlah Faktor	20
2.2.4 Rotasi Faktor	21
2.3 Uji Asumsi	23
2.3.1 Uji Kecukupan Data	23
2.3.2 Uji <i>Bartlett</i> (Kebebasan Antar Variabel)	24
2.4 Nilai Karakteristik dan Matriks Ortogonal	25

2.5	<i>Principal Component Analysis (PCA)</i>	29
2.6	<i>Calinski-Harabasz Pseudo F-statistic</i>	31
2.7	<i>Internal Cluster Dispersion Rate (icdrate)</i>	32
2.8	<i>Multivariate Analyze of Varians (MANOVA)</i>	33
	2.8.1 Pemeriksaan Distribusi Normal Multivariat	33
	2.8.2 Pengujian MANOVA	34
2.9	Pembangunan Manusia	35
2.10	<i>Cluster</i> dalam Pandangan Islam	38

BAB III METODE PENELITIAN

3.1	Pendekatan Penelitian	39
3.2	Sumber Data	39
3.3	Variabel Penelitian	39
3.4	Analisis Data	40
	3.4.1 Pembuktian Asumsi Analisis Faktor	40
	3.4.2 Pengelompokan kabupaten/kota di Provinsi Jawa Timur	40
	3.4.3 Karakteristik Setiap Kelompok	42

BAB IV PEMBAHASAN

4.1	Pembuktian Asumsi pada Model Persamaan Analisis Faktor	43
4.2	Pengelompokan Kabupaten/Kota di Provinsi Jawa Timur	47
	4.2.1 Deskriptif Indikator IPM 2015	47
	4.2.2 Analisis Faktor	54
	4.2.3 Analisis <i>Cluster</i> Hirarki	62
4.3	Karakteristik Setiap Kelompok	72
4.4	Kajian Agama Tentang <i>Cluster</i>	72

BAB V PENUTUP

5.1	Kesimpulan	75
5.2	Saran	76

DAFTAR RUJUKAN	77
-----------------------------	----

LAMPIRAN-LAMPIRAN

RIWAYAT HIDUP

DAFTAR SIMBOL



X	: variabel asli/awal
Z	: variabel baku
S	: standar deviasi
\bar{X}	: rata-rata X
F	: <i>common factor</i>
L	: <i>loading factor</i>
μ	: rata-rata
ε	: <i>specific factor</i>
φ	: matriks diagonal
σ_{ii}	: nilai kumunalitas yang ditambahkan nilai varians spesifik
φ_i	: varians khusus ke- i
h_i^2	: kumunalitas ke- i
res_{ij}	: elemen-elemen matriks sisa selain elemen diagonal utama pada variabel ke- i dan variabel ke- j .
p	: banyaknya variabel yang diamati
$\ln R $: nilai determinan dari matriks korelasi
n	: banyaknya observasi
p	: banyaknya variabel
P	: matriks korelasi
I	: matriks identitas
$d(i, j)$: jarak antara objek i ke objek j .
\bar{x}_i	: nilai tengah pada <i>cluster</i> ke- i .

- \bar{x}_j : nilai tengah pada *cluster* ke- j .
- S^{-1} : matriks ragam peragam gabungan
- W : matriks *sum of square residuals*
- B : matriks *sum of square treatment*
- n : jumlah sampel
- l : banyaknya kelompok
- n_l : banyaknya anggota kelompok l



DAFTAR TABEL

Tabel 2.1	Klasifikasi Kecukupan Sampel	24
Tabel 4.1	Nilai <i>Measure of Sampling Adequacy</i> (MSA)	55
Tabel 4.2	Hasil KMO dan Uji <i>Bartlett</i>	56
Tabel 4.3	Tabel Total <i>Variance Explained</i>	57
Tabel 4.4	Tabel Kumunalitas	59
Tabel 4.5	Tabel Komponen Matriks	59
Tabel 4.6	Tabel Rotasi Matriks	60
Tabel 4.7	Tabel Signifikansi Loading Faktor	61
Tabel 4.8	Tabel Hasil Pengelompokan Variabel ke dalam Faktor	61
Tabel 4.9	Tabel Penamaan Faktor	62
Tabel 4.10	Nilai <i>Pseudo F</i> Masing-masing Metode <i>Cluster</i>	63
Tabel 4.11	Nilai <i>Icdrate</i> untuk Setiap Pengelompokan	67
Tabel 4.12	Uji Homogenitas Matriks Varians Kovarians	71
Tabel 4.13	Perbedaan Karakteristik Setiap Kelompok	72

DAFTAR GAMBAR

Gambar 4.1	Tingkat Partisipasi Angkatan Kerja	47
Gambar 4.2	Tingkat Pengangguran Terbuka	48
Gambar 4.3	Angka Kematian Bayi	49
Gambar 4.4	Angka Harapan Hidup	49
Gambar 4.5	Angka Keluhan Kesehatan	50
Gambar 4.6	Angka Partisipasi Sekolah Usia 16-18 tahun	50
Gambar 4.7	Angka Partisipasi Kasar Usia 16-18 tahun	51
Gambar 4.8	Angka Partisipasi Murni Usia 16-18 tahun	52
Gambar 4.9	Angka Penduduk yang Tidak Bersekolah	53
Gambar 4.10	Angka Penduduk Buta Huruf	53
Gambar 4.11	Angka Penduduk Miskin	54
Gambar 4.12	Gambar <i>Scree Plot</i>	58
Gambar 4.13	Dendogram <i>Complete Linkage</i>	64
Gambar 4.14	Dendogram <i>Average Linkage</i>	65
Gambar 4.15	Dendogram <i>Ward's Linkage</i>	66
Gambar 4.16	Pengelompokan Kabupaten/Kota <i>Complete Linkage</i>	68
Gambar 4.17	Pengelompokan Kabupaten/Kota <i>Average Linkage</i>	68
Gambar 4.18	Pengelompokan Kabupaten/Kota <i>Ward's Linkage</i>	69
Gambar 4.19	Pemeriksaan Normal Multivariat	70

ABSTRAK

Sirojuddin, Achmad. 2016. **Analisis Cluster pada Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Indeks Pembangunan Manusia**. Skripsi. Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Sri Harini, M.Si. (II) Mohammad Jamhuri, M.Si.

Kata Kunci: *cluster*, hirarki, *average linkage*, MANOVA, indikator IPM

Analisis *cluster* adalah salah satu teknik interdependensi yang dapat menggambarkan kedekatan jarak atau kemiripan antara objek dan variabel. Analisis *cluster* terbagi atas dua metode, yaitu hirarki dan nonhirarki. Dalam metode hirarki terdapat 3 metode antara lain metode *Complete Linkage*, *Average Linkage* dan *Ward's Linkage*. Metode ini diterapkan dalam pengelompokan kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator IPM. Tujuannya yaitu untuk mengetahui daerah mana saja yang memiliki kedekatan atau kemiripan. Untuk mengetahui setiap kelompok memiliki perbedaan yang signifikan, maka digunakan uji MANOVA untuk dapat membuktikan perbedaan setiap kelompok. Dari 38 kabupaten/kota di Provinsi Jawa Timur diperoleh metode terbaik yaitu metode *Average Linkage* sebanyak 5 kelompok dengan melihat nilai *icdrate* yang paling terkecil. Kelompok yang terbentuk di antaranya, kelompok 1 terdiri dari Kabupaten Pacitan, Ponorogo, Blitar, Kediri, Malang, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Nganjuk, Bojonegoro, Tuban, Pamekasan, dan Sumenep. Kelompok 2 terdiri hanya Kabupaten Trenggalek. Kelompok 3 terdiri dari Kabupaten Ponorogo, Sidoarjo, Mojokerto, Jombang, Magetan, Ngawi, Lamongan, Gresik, Kota Kediri, Blitar, Malang, Pasuruan, Probolinggo, Mojokerto, Madiun, Surabaya, dan Batu. Kelompok 4 terdiri dari Kabupaten Kediri, Malang, Pasuruan, Bangkalan, dan Sampang. Kabupaten Madiun merupakan satu-satunya anggota yang masuk dalam kelompok 5.

ABSTRACT

Sirojuddin, Achmad. 2016. **Cluster Analysis on the District/City in East Java Province Based on the Indicators of Human Development Index.** Thesis. Mathematics Department, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University Malang. Advisor: (I) Dr. Sri Harini, M.Si. (II) Mohammad Jamhuri, M.Si.

Keywords: cluster, hierarchy, *average linkage*, MANOVA, HDI indicators

Cluster analysis is one of the interdependence techniques which can portray the close distance or similarity between object and variable. Cluster analysis is divided into two methods, namely hierarchy and non-hierarchy. There are three methods in a hierarchical method including Complete Linkage, Average Linkage, and Ward's linkage. This method is implemented in the grouping of districts/cities in East Java province based on the indicators of the HDI. It aims to discover the areas which have a proximity or similarity. To determine that each group has a significant difference, MANOVA test is executed to prove a difference from each group. Out of the 38 districts/cities in East Java province, the best method obtained is Average Linkage taken from 5 groups by looking at the smallest value of icdrate. Some of the successfully formed groups are group 1 which consists of Pacitan, Ponorogo, Blitar, Kediri, Malang, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Nganjuk, Bojonegoro, Tuban, Pamekasan, and Sumenep District. Group 2 only consists of Trenggalek District. Group 3 consists of Ponorogo, Sidoarjo, Mojokerto, Jombang, Magetan, Ngawi, Lamongan, Gresik, Kediri, Blitar, Malang, Pasuruan, Probolinggo, Mojokerto, Madiun, Surabaya, and Batu. Group 4 consists of Kediri, Malang, Pasuruan, Bangkalan, and Sampang District. Madiun District is the only member of groups 5.

ملخص

أحمد سراج الدين، 2016. التحليل العنقودي في المقاطعة أو المدينة في جاوى الشرقية على أساس مؤشرات التنمية البشرية. البحث الجامعي. قسم الرياضيات، كلية العلوم والتكنولوجيا بجامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانق. المشرف الأول: د. سري هريني الماجستير. المشرف الثاني: محمد جمهوري الماجستير.

الكلمات الأساسية: العنقودي، هرمي، الربط المتوسط، مانوفا، مؤشرات التنمية البشرية.

التحليل العنقودي هو إحدى الأساليب التي تصف التقارب أو التشابه بين الأشياء والمتغيرات. وينقسم التحليل العنقودي إلى نوعين؛ الهرمي وغير الهرمي. في الأسلوب الهرمي، يشمل ثلاثة طرائق؛ الربط الكامل، الربط المتوسط وربط وارد (*ward's*). يتم تطبيق هذا الأسلوب في جميع المقاطعات أو المدن في محافظة جاوى الشرقية على أساس مؤشرات التنمية البشرية. والهدف من ذلك، هو تحديد أي مقاطعات لها وجه التقارب أو التشابه. ولتحديد كل مجموعة لديها فرق كبير فيستخدم اختبار مانوفا (*MANOVA*) لإثبات تجانسها. تم الحصول من ثمان وثلاثين مقاطعة أو مدينة محافظة جاوى الشرقية أفضل أسلوب وهو أسلوب الربط المتوسط على خمس مجموعات بالنظر إلى أقل درجة معدل تشتت العنقود الداخلي (*icdrate*). تشكل المجموعة على نحو التالي؛ المجموعة الأولى تتكون من مقاطعة فاجيتان، فونروغو، بلييتار، كادري، مالانق، بابووانجى، بوندووسو، سيتوبوندو، فروبولينجو، عانجوك، بوجونيغورو، توبان، بامكاسان، وسومنب. والمجموعة الثانية تتكون من مقاطعة ترينغاليك فقط. المجموعة الثالثة تتكون من مقاطعة فونورغو، سيدوارجو، موجوكرطو، جومبانج، ماغتان، غاوي، لمونجان، غرسيك، مدينة كادري، بلييتار، مالانق، باسوروان، فروبولينجو، موجوكرطو، ماديون، سورابايا، وباتو. وأما المجموعة الرابعة تتكون من مقاطعة كادري، مالانق، باسوروان، بنكالان، وسامفانج. ومقاطعة ماديون هي الوحيدة في المجموعة الخامسة.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Ilmu statistika merupakan ilmu yang mempelajari bagaimana merencanakan, mengumpulkan, menganalisis, menginterpretasi, dan mempresentasikan data. Dari kumpulan data, statistik dapat digunakan untuk menyimpulkan atau mendeskripsikan data. Sebagian besar konsep dasar statistik mengasumsikan teori probabilitas. Ilmu statistika juga banyak diterapkan dalam berbagai disiplin ilmu alam, sosial, pendidikan, dan bisnis. Ilmu statistika juga digunakan dalam pemerintahan untuk berbagai macam tujuan. Ilmu statistika memiliki peranan penting dalam perkembangan ilmu pengetahuan dan teknologi. Ilmu statistika juga memegang peranan penting dalam penelitian, baik dalam penyusunan model, perumusan hipotesis, pengembangan alat, instrumen pengumpulan data, penyusunan desain penelitian, penentuan sampel, dan dalam analisis data (Irianto, 2004).

Berdasarkan berbagai pengertian statistika tersebut, maka dalam pengaplikasian penelitian ini, ilmu statistika akan diaplikasikan dalam bidang Indeks Pembangunan Manusia (IPM). Menurut BPS (2015), IPM mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup. Sebagai ukuran kualitas hidup, IPM dibangun melalui pendekatan tiga dimensi dasar. Dimensi tersebut mencakup pendidikan, kesehatan, serta kehidupan yang layak. Ketiga dimensi tersebut memiliki pengertian sangat luas karena terkait banyak faktor.

IPM juga merupakan suatu tolak ukur maju atau tidaknya suatu wilayah ataupun daerah, karena dengan tingkat IPM yang tinggi suatu daerah akan dikatakan berhasil dalam program pembangunannya. Untuk membangun IPM yang berkualitas tinggi pemerintah harus membuat program-program serta fasilitas yang memadai untuk meningkatkan IPM, namun juga harus dengan dukungan masyarakat agar program-program tersebut dapat berjalan.

Pada ilmu statistika, apabila variabel yang digunakan lebih dari dua, maka akan semakin rumit untuk melakukan analisis statistik. Salah satu metode yang dapat mengatasi hal tersebut adalah analisis multivariat. Analisis multivariat merupakan metode statistik yang memungkinkan melakukan penelitian terhadap lebih dari dua variabel. Salah satu teknik analisis multivariat yang dapat digunakan untuk memahami dan menjelaskan seperangkat variabel adalah analisis *cluster*. Analisis *cluster* adalah salah satu teknik multivariat yang bertujuan mengklasifikasi suatu objek-objek ke dalam suatu kelompok-kelompok yang berbeda antara kelompok satu dengan lainnya. Objek-objek yang telah diklasifikasikan dalam satu *cluster* merupakan objek-objek yang memiliki kedekatan jarak relatif sama dengan objek lainnya (Narimawati, 2008).

Pada al-Quran juga telah dijelaskan mengenai pengelompokan yang dituangkan dalam surat asy-Syura ayat 15, yaitu:

فَلِذَلِكَ فَادَعُ^ط وَأَسْتَقِمْ^ط كَمَا أُمِرْتَ^ط وَلَا تَتَّبِعْ أَهْوَاءَهُمْ^ط وَقُلْ ءَامَنْتُ بِمَا أَنْزَلَ اللَّهُ مِنْ
كِتَابٍ وَأُمِرْتُ لِأَعْدِلَ^ط بَيْنَكُمْ^ط اللَّهُ رَبُّنَا وَرَبُّكُمْ^ط لَنَا أَعْمَلْنَا^ط وَلَكُمْ أَعْمَلُكُمْ^ط لَا حُجَّةَ
بَيْنَنَا وَبَيْنَكُمْ^ط اللَّهُ يَجْمَعُ بَيْنَنَا^ط وَإِلَيْهِ الْمَصِيرُ^ط

“Maka karena itu serulah (mereka kepada agama ini) dan tetaplah sebagaimana diperintahkan kepadamu dan janganlah mengikuti hawa nafsu mereka dan katakanlah: "Aku beriman kepada semua kitab yang diturunkan Allah Swt. dan

aku diperintahkan supaya berlaku adil di antara kamu. Allah-lah Tuhan kami dan Tuhan kamu. Bagi kami amal-amal kami dan bagi kamu amal-amal kamu. Tidak ada pertengkaran antara kami dan kamu, Allah Swt. mengumpulkan antara kita dan kepada-Nya-lah kembali (kita)" (QS. Asy-Syura/42:15).

Paparan makna dari surat asy-Syura ayat 15 telah menjelaskan bahwa Allah Swt. telah mengelompokkan manusia menjadi beberapa kelompok atau kaum, dengan pegangan kitab-kitab suci Allah Swt. yang telah diturunkan-Nya. Agar tidak terjadi pertentangan atau pertengkaran antara satu kaum dengan kaum lainnya, sehingga tercipta kedamaian di antara umat-umat beragama.

Penelitian sebelumnya yang membahas analisis *cluster* adalah penelitian Amalyah (2012) berjudul "Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Indonesia Sehat 2010", terdapat 7 kelompok dengan metode *Complete Linkage* dari 38 kabupaten/kota. Oleh karena itu, pada penelitian ini menggunakan metode Analisis *Cluster* Hirarki untuk pengelompokan dan pemetaan kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator IPM sehingga nantinya dihasilkan beberapa kelompok-kelompok yang berbeda yang dapat menunjukkan karakteristik masing-masing dari setiap kelompok tersebut. Sebelum dilakukan pengelompokan menggunakan analisis *cluster*, maka perlu dilakukan analisis faktor terlebih dahulu untuk mereduksi variabel dan dilanjutkan dengan pengelompokan kabupaten/kota di Provinsi Jawa Timur menggunakan analisis *cluster*.

Berdasarkan uraian di atas maka penulis mengangkat permasalahan dan menyusunnya dalam sebuah penelitian yang berjudul "Analisis *Cluster* pada Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Indeks Pembangunan Manusia".

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah dalam penelitian ini, yaitu:

1. Bagaimana pembuktian asumsi dari model persamaan analisis faktor?
2. Apa metode yang terbaik untuk mengelompokkan kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator IPM menggunakan hasil nilai *icdrate*?
3. Apa saja variabel-variabel yang mencirikan tiap kelompok berdasarkan indikator IPM?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan yang ingin dicapai dalam penelitian ini, yaitu:

1. Untuk mengetahui pembuktian asumsi dari model persamaan analisis faktor.
2. Untuk mengetahui metode yang terbaik untuk mengelompokkan kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator IPM menggunakan hasil nilai *icdrate*.
3. Untuk mengetahui variabel-variabel yang mencirikan tiap kelompok berdasarkan indikator IPM.

1.4 Manfaat Penelitian

a. Bagi Peneliti

1. Untuk menambah wawasan mengenai analisis *cluster*.
2. Dapat melakukan pengelompokan data menggunakan analisis *cluster*.

3. Dapat memahami metode *Complete Linkage*, *Average Linkage*, dan *Ward's Linkage*.

b. Bagi Mahasiswa

Penelitian ini dapat dijadikan sebagai bahan rujukan dan pengembangan pembelajaran statistika mengenai analisis *cluster*.

c. Bagi Instansi

1. Penelitian ini sebagai sumbangan pemikiran keilmuan matematika, khususnya dalam bidang statistika.

2. Meningkatkan peran serta Fakultas Sains dan Teknologi UIN Maulana Malik Ibrahim Malang dalam pengembangan wawasan keilmuan matematika dan statistika.

d. Bagi pihak lain

Penelitian ini sebagai sarana untuk mengetahui daerah di Jawa Timur yang memiliki indikator IPM, kesehatan maupun kemiskinan yang ada dalam proses pembangunan IPM .

1.5 Batasan Masalah

Agar mendekati sasaran yang diharapkan, maka perlu diadakan pembatasan permasalahan, antara lain:

1. Metode yang digunakan dalam penelitian ini ada tiga diantaranya, metode *Ward's Linkage*, *Complete Linkage*, dan *Average Linkage*.

2. Data dalam penelitian ini adalah data dari Badan Pusat Statistik (BPS) Provinsi Jawa Timur pada tahun 2015.

1.6 Sistematika Penulisan

Agar dapat membaca hasil penelitian ini, maka dalam penyajiannya ditulis berdasarkan suatu sistematika yang secara garis besar dibagi menjadi lima bab, yaitu:

Bab I Pendahuluan

Meliputi latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan.

Bab II Kajian Pustaka

Berisi tentang teori-teori yang berhubungan dengan pembahasan antara lain analisis *cluster*, analisis faktor, indikator IPM, serta pandangan Islam terhadap *cluster*.

Bab III Metode Penelitian

Berisi pendekatan penelitian, sumber data, variabel penelitian, dan analisis data. .

Bab IV Pembahasan

Pada bab ini berisi tentang pembahasan pengelompokan kabupaten/kota di Provinsi Jawa Timur dengan menggunakan analisis *cluster* dan penjelasan mengenai kajian agama tentang *cluster*.

Bab V Penutup

Berisi kesimpulan dan saran.

BAB II

KAJIAN PUSTAKA

2.1 Analisis *Cluster*

Analisis *cluster* adalah salah satu teknik multivariat yang bertujuan mengklasifikasi suatu objek-objek ke dalam suatu kelompok-kelompok yang berbeda antara kelompok satu dengan lainnya. Objek-objek yang telah diklasifikasikan dalam satu *cluster* merupakan objek-objek yang memiliki kedekatan jarak relatif sama dengan objek lainnya (Narimawati, 2008).

Analisis *cluster* mengelompokkan sejumlah n objek berdasarkan p varians yang secara relatif mempunyai kesamaan karakteristik di antara objek-objek tersebut, sehingga keragaman di dalam suatu kelompok tersebut lebih kecil dibandingkan keragaman antar kelompok. Objek dapat berupa barang, jasa, tumbuhan, binatang, dan orang (responden, konsumen, atau yang lainnya). Objek tersebut akan diklasifikasi ke dalam satu atau lebih *cluster* (kelompok) sehingga objek-objek yang berada dalam satu *cluster* (kelompok) akan mempunyai satu kemiripan atau kesamaan karakter.

Prosedur pembentukan *cluster* terbagi menjadi 2, yaitu hirarki dan nonhirarki. Pembentukan *cluster* hirarki mempunyai sifat sebagai pengembangan suatu hirarki atau struktur mirip pohon bercabang. Metode *Cluster* Hirarki merupakan metode pengelompokan yang mana jumlah kelompok yang akan dibuat belum diketahui. Teknik ini diproses dengan baik melalui penggabungan berurutan (*agglomerative*) atau pembagian berurutan (*divisive*).

Menurut Johnson (1967), cara kerja metode *Cluster* Hirarki yaitu, diberikan sekumpulan N item yang akan dicluster, dan sebuah matrik $N \times N$ yang menyatakan jarak antar item pada N :

1. Mulai dengan membuat *cluster* sebanyak N , masing-masing *cluster* mempunyai sebuah item. Misalnya jarak antar *cluster* sama dengan jarak antar item yang dikandungnya.
2. Cari sepasang *cluster* yang jaraknya terdekat, dan jadikan sebuah *cluster* baru. Jadi sekarang kita mempunyai $N - 1$ *cluster*.
3. Hitung jarak antara *cluster* yang baru dengan masing-masing *cluster* yang lainnya.

Menurut Johnson dan Wichern (2007), metode yang digunakan untuk menghitung jarak antar *cluster* ada 2, antara lain.

a. Jarak Euclidian

Jarak Euclidian adalah jarak yang paling umum dan paling sering digunakan dalam analisis *cluster*. Jarak Euclidian antar dua titik dapat terdefiniskan dengan jelas. Jarak yang digunakan adalah peubah kontinu. Jarak Euclidian antara *cluster* ke- i dan ke- j dari p peubah didefinisikan:

$$d_{(i,j)} = \left[\sum_{t=1}^p (\bar{x}_i - \bar{x}_j)^2 \right]^{\frac{1}{2}} \quad (2.1)$$

dengan:

- $d_{(i,j)}$: jarak antara objek i ke objek j .
- \bar{x}_i : nilai tengah pada *cluster* ke- i .
- \bar{x}_j : nilai tengah pada *cluster* ke- j .
- p : banyaknya peubah yang diamati.

b. Jarak Mahalanobis

Jarak Mahalanobis sangat berguna dalam menghilangkan atau mengurangi perbedaan skala pada masing-masing komponen. Pada permasalahan tertentu, pada saat menentukan jarak, perlu juga dipertimbangkan varians dan kovarians. Jarak Mahalanobis didefinisikan:

$$d_{(i,j)} = [(\bar{x}_i - \bar{x}_j)S^{-1}(\bar{x}_i - \bar{x}_j)]^{\frac{1}{2}} \quad (2.2)$$

dengan:

- $d_{(i,j)}$: jarak antara objek i ke objek j .
- \bar{x}_i : nilai tengah pada *cluster* ke- i .
- \bar{x}_j : nilai tengah pada *cluster* ke- j .
- S^{-1} : matriks varians dan kovarians gabungan.

4. Ulangi langkah 2 dan 3 sampai semua item menjadi sebuah *cluster* dengan N item. Tentunya tidak ada gunanya mempunyai N item yang dikelompokkan menjadi satu *cluster* besar.

Terdapat beberapa teknik pengelompokan dalam metode *Cluster* Hirarki yang dapat digunakan di antaranya adalah:

2.1.1 Metode *Ward's Linkage*

Pada metode *Ward's Linkage*, jarak antara dua kelompok yang terbentuk adalah *sum of square* di antara dua kelompok tersebut. Pengelompokan metode *Ward's Linkage* adalah meminimumkan peningkatan kriteria *Error Sum of Square* (ESS). Dua kelompok yang memiliki peningkatan ESS paling minimum, maka akan berkelompok. Jika kelompok sebanyak K maka ESS merupakan jumlahan dari $+ ESS_K$ atau dapat dituliskan $ESS = ESS_1 + ESS_2 + \dots + ESS_K$. Ketika semua kelompok bergabung menjadi satu kelompok dari N objek, maka untuk

menghitung jarak antara dua kelompok menggunakan metode *Ward's Linkage*, dapat menggunakan rumus sebagai berikut:

$$ESS = \sum_{j=1}^N (x_j - \bar{x})'((x_j - \bar{x})) \quad (2.3)$$

dengan x_j adalah ukuran asosisasi multivariat dengan j item dan \bar{x} adalah rata-rata dari seluruh item (Johnson dan Wichern, 2007).

2.1.2 Metode *Complete Linkage*

Metode *Complete Linkage* menggunakan prinsip jarak antar objek. Dasar dari metode ini adalah jarak maksimum atau jarak terjauh. Dengan demikian, metode *Complete Linkage* menjamin bahwa seluruh item pada kelompok terdapat jarak maksimum (kesamaan minimum). Algoritma aglomeratif umum dimulai dengan menemukan elemen dalam $D = \{d_{i,k}\}$ dan menggabungkan objek yang berkorespondensi, misalkan U dan V untuk membentuk kelompok (UV) . Langkah selanjutnya jarak antara (UV) dan kelompok lainnya, misalkan W dapat ditulis sebagai berikut:

$$d_{(UV)W} = \max\{d_{UW}, d_{VW}\} \quad (2.4)$$

yang mana d_{UW} dan d_{VW} merupakan jarak terjauh antara anggota kelompok U dan W serta kelompok V dan W , begitu juga sebaliknya (Johnson dan Wichern, 2007).

2.1.3 Metode *Average Linkage*

Metode *Average Linkage* menghitung jarak dua *cluster* yang disebut sebagai jarak rata-rata yang mana jarak tersebut dihitung pada masing-masing *cluster*.

$$d_{(UV)W} = \frac{\sum_i \sum_k d_{ik}}{N_{(UV)}N_W} \quad (2.5)$$

dengan d_{ik} merupakan jarak antar objek i dalam *cluster* (UV) dan objek k dalam *cluster* W . Sedangkan $N_{(UV)}$ dan N_w berturut-turut merupakan jumlah objek dalam *cluster* (UV) dan (W) (Johnson dan Wichern, 2007).

2.2 Analisis Faktor

Analisis faktor merupakan suatu teknik untuk menganalisis tentang saling ketergantungan dari beberapa variabel secara simultan dengan tujuan untuk menyederhanakan dari bentuk hubungan antara beberapa variabel yang diteliti menjadi sejumlah faktor yang lebih sedikit dari pada variabel yang diteliti. Hal ini berarti, analisis faktor dapat juga menggambarkan tentang struktur data dari suatu penelitian (Suliyanto, 2005).

Analisis faktor adalah suatu teknik interdependensi (*interdependence technique*), yang mana tidak ada pembagian variabel menjadi variabel bebas dan variabel tergantung dengan tujuan utama yaitu mendefinisikan struktur yang terletak di antara variabel-variabel dalam analisis. Analisis ini menyediakan alat-alat untuk menganalisis struktur hubungan interen atau korelasi di antara sejumlah besar variabel dengan menerangkan korelasi yang baik antar variabel, yang diasumsikan untuk merepresentasikan dimensi-dimensi dalam data (Hair, dkk, 1995).

Pada prinsipnya analisis faktor digunakan untuk mengelompokkan beberapa variabel yang memiliki kemiripan untuk dijadikan satu faktor, sehingga dimungkinkan dari beberapa atribut yang mempengaruhi satu komponen variabel dapat diringkas menjadi beberapa faktor utama yang jumlahnya lebih sedikit.

Tahap-tahap dalam analisis faktor yaitu pembentukan matriks varians dan kovarians, metode *principal component*, kriteria penentuan jumlah faktor, dan rotasi faktor.

2.2.1 Pembentukan Matriks Varians dan Kovarians

Menurut Johnson dan Wichern (2007), secara umum analisis faktor ortogonal disusun seperti model analisis regresi multivariat. Setiap variabel awal dinyatakan sebagai kombinasi linier dari faktor-faktor yang mendasari. Misalkan terdapat X variabel random dengan variabel sebanyak p , yang memiliki rata-rata μ dan matriks kovarians Σ merupakan penyusun model faktor. Variabel F_1, F_2, \dots, F_m merupakan faktor yang nilainya tidak terobservasi, $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ merupakan kesalahan (*error*) atau faktor spesifik. Secara khusus, persamaan model analisis faktor sebagai berikut:

$$\begin{aligned} X_1 - \mu_1 &= l_{11}F_1 + l_{12}F_2 + \dots + l_{1p}F_p + \varepsilon_1 \\ X_2 - \mu_2 &= l_{21}F_1 + l_{22}F_2 + \dots + l_{2p}F_p + \varepsilon_2 \\ &\vdots \\ X_p - \mu_p &= l_{p1}F_1 + l_{p2}F_2 + \dots + l_{pm}F_m + \varepsilon_p \end{aligned} \quad (2.6)$$

atau dalam notasi matriks:

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \\ X_p \end{pmatrix}_{p \times 1} - \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \vdots \\ \mu_p \end{pmatrix}_{p \times 1} = \begin{pmatrix} l_{11} & l_{12} & l_{13} & \dots & l_{1m} \\ l_{21} & l_{22} & l_{23} & \dots & l_{2m} \\ l_{31} & l_{32} & l_{33} & \dots & l_{3m} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ l_{p1} & l_{p2} & l_{p3} & \dots & l_{pm} \end{pmatrix}_{p \times m} \begin{pmatrix} F_1 \\ F_2 \\ F_3 \\ \vdots \\ F_p \end{pmatrix}_{m \times 1} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_p \end{pmatrix}_{p \times 1}$$

atau

$$(X - \mu)_{p \times 1} = L_{p \times m} F_{m \times 1} + \varepsilon_{p \times 1} \quad (2.7)$$

dengan:

X : vektor acak

F : *common factor*

L : *loading factor*

μ : rata-rata

ε : *spesific factor*

Dengan banyaknya jumlah yang tidak terobservasi, verifikasi model faktor dari observasi X_1, X_2, \dots, X_p tidak diperlukan. Tetapi penambahan asumsi tentang vektor random F dan ε pada model (2.7) menandakan hubungan kovarians.

Penambahan asumsi sebagai berikut:

$$E[F] = 0_{(m \times 1)}, Cov(F) = E[FF'] = I_{(m \times m)} \quad (2.8)$$

$$E[\varepsilon] = 0_{(p \times 1)}, Cov(\varepsilon) = E[\varepsilon\varepsilon'] = \varphi_{(p \times p)} = \begin{bmatrix} \varphi_1 & 0 & \dots & 0 \\ 0 & \varphi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \varphi_p \end{bmatrix} \quad (2.9)$$

dengan:

φ = matriks diagonal yang mana F dan ε saling bebas sehingga,

$$Cov(\varepsilon, F) = E[\varepsilon F] = 0_{(p \times m)}.$$

Model ortogonal dari analisis faktor berakibat kepada struktur kovarians untuk variabel acak X , yaitu:

$$\begin{aligned} (X - \mu)(X - \mu)' &= (LF + \varepsilon)(LF + \varepsilon)' \\ &= (LF + \varepsilon)((LF)' + \varepsilon') \\ &= LF(LF)' + \varepsilon(LF)' + \varepsilon\varepsilon' \end{aligned} \quad (2.10)$$

Berdasarkan persamaan (2.8) dan persamaan (2.9), maka matriks kovariansnya menjadi:

$$\begin{aligned}
\Sigma = Cov(X) &= E(X - \mu)(X - \mu)' \\
&= E[LF(LF)' + \varepsilon(LF)' + LF\varepsilon + \varepsilon\varepsilon'] \\
&= LL'E(FF') + L'E(\varepsilon F') + LE(F\varepsilon') + E(\varepsilon\varepsilon') \\
&= LL'.I + L'.0 + L.0 + \psi \\
&= LL' + \psi
\end{aligned} \tag{2.11}$$

Persamaan (2.11) dapat ditulis dengan:

$$Var(X_i) = l_{i1}^2 + \dots + l_{im}^2 + \psi_i \tag{2.12}$$

$$Cov(X_i, X_k) = l_{i1}l_{k1} + \dots + l_{im}l_{km} \tag{2.13}$$

Kovarians untuk variabel acak X dan faktor umum F :

$$\begin{aligned}
(X - \mu)F' &= (LF + \varepsilon)F' \\
&= LFF' + \varepsilon F'
\end{aligned} \tag{2.14}$$

Berdasarkan pada persamaan (2.7) dan (2.8), maka:

$$\begin{aligned}
Cov(X, F) &= E(X - \mu)F' \\
&= E[LFF' + \varepsilon F'] \\
&= LE(FF') + E(\varepsilon F') \\
&= LI + 0 \\
&= L
\end{aligned} \tag{2.15}$$

Secara umum dapat ditulis sebagai berikut:

$$Cov(X_i, F_j) = l_{ij} \text{ dengan } i = 1, 2, \dots, p \text{ dan } j = 1, 2, \dots, m \tag{2.16}$$

Varians dari variabel ke- i yang disumbangkan oleh m *common factor* disebut kumunalitas ke- i . Nilai $Var(X_i) = \sigma_{ii}$ merupakan nilai kumunalitas yang ditambahkan dengan nilai varians spesifik atau uniknya. Dari persamaan (2.12) dan persamaan (2.13) diperoleh:

$$\underbrace{\sigma_{ii}}_{\text{Var}(X_i)} = \underbrace{l_{i1}^2 + l_{i2}^2 + \dots + l_{im}^2}_{\text{kumunalitas}} + \underbrace{\varphi_i}_{\text{variansi spesifik}} \quad (2.17)$$

Kumunalitas ke- i dinotasikan sebagai h_i^2 , dengan menggunakan notasi di atas diperoleh persamaan:

$$h_i^2 = l_{i1}^2 + l_{i2}^2 + \dots + l_{im}^2 \quad (2.18)$$

dan

$$\sigma_{ii} = h_i^2 + \varphi_i, i = 1, 2, \dots, p \quad (2.19)$$

Nilai kumunalitas ke- i merupakan jumlah kuadrat dari *loading factor* variabel ke- i pada m faktor. Sehingga total varians dari model faktor dapat ditulis:

$$\text{tr}(\hat{O}) = \sum_{i=1}^p \sigma_{ii}$$

2.2.2 Metode *Principal Component*

Menurut Supranto (2004), metode *Principal Component* bertujuan untuk mengestimasi parameter pada analisis faktor, yaitu varians spesifik ($\psi_{(p \times p)}$), kumunalitas (h), dan matriks faktor *loading* ($L_{(p \times m)}$). Matriks varians kovarians yaitu S yang merupakan estimator (penduga) bagi matriks varians kovarians populasi yang tidak diketahui yaitu \hat{O} . Komponen utama analisis faktor pada matriks varians kovarians populasi \hat{O} memiliki pasangan nilai eigen dan vektor eigen (λ_i, e_i) dimana $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$.

Misalkan X_1, X_2, \dots, X_p merupakan sampel *random* yang teramati sebanyak p komponen. Dari data tersebut diperoleh rata-rata sample \bar{x} , matriks varians kovarians S , dan matriks korelasi R . Untuk mengestimasi nilai \hat{L} , substitusikan \hat{O} dengan S pada persamaan (2.11).

Perhatikan bahwa:

$$\hat{O} = LL' + \psi$$

substitusikan \hat{O} dengan S sehingga diperoleh:

$$S \cong \hat{L}\hat{L}' + \psi$$

Pada metode *Principal Component*, nilai $\hat{\psi}$ diabaikan sehingga S dapat difaktorkan menjadi:

$$S \cong \hat{L}\hat{L}' \quad (2.20)$$

Selanjutnya dengan spektral dekomposisi diperoleh:

$$S = E \ddot{E} E' \quad (2.21)$$

dengan E merupakan matriks ortogonal yang kolom-kolomnya merupakan vektor-vektor eigen yang telah dibakukan (matriks ortonormal) dari matriks S dan

$$\ddot{E} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \lambda_p \end{bmatrix} \quad (2.22)$$

dengan $\lambda_1, \lambda_2, \dots, \lambda_p$ adalah nilai untuk matriks S dan $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$.

Kemudian bentuk spektral dekomposisi pada persamaan (2.21) difaktorkan menjadi:

$$\ddot{E} = \ddot{E}_2^{-1} \ddot{E}_2 \quad (2.23)$$

dengan:

$$\ddot{E}_2^{-1} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \lambda_p \end{bmatrix} \quad (2.24)$$

sehingga:

$$S = E \ddot{E} E'$$

$$\begin{aligned}
&= E \ddot{E}_2^{\frac{1}{2}} \ddot{E}_2^{\frac{1}{2}} E' \\
&= (E \ddot{E}_2^{\frac{1}{2}}) (E \ddot{E}_2^{\frac{1}{2}})' \tag{2.25}
\end{aligned}$$

Berdasarkan persamaan (2.24) dan (2.25) diperoleh:

$$\hat{L} = (E \ddot{E}_2^{\frac{1}{2}}) \tag{2.26}$$

dengan $\ddot{E}_2^{\frac{1}{2}}$ merupakan matriks yang entri-entrinya merupakan nilai eigen terbesar bersesuaian dengan $\lambda_1, \lambda_2, \dots, \lambda_m$ dan E merupakan vektor eigen yang bersesuaian dengan $\ddot{E}_2^{\frac{1}{2}}$. Akan tetapi, \hat{L} tidak dapat secara langsung dinyatakan sebagai $E \ddot{E}_2^{\frac{1}{2}}$ karena $E \ddot{E}_2^{\frac{1}{2}}$ merupakan matriks berdimensi $p \times p$ sedangkan \hat{L} merupakan matriks berdimensi $p \times m$ dengan $m < p$. Karena itulah diambil \ddot{E}_1 memuat m nilai eigen terbesar dan E_1 yang memuat m vektor eigen terbesar.

Berdasarkan persamaan (2.25), dengan mengambil \ddot{E}_1 sebagai matriks yang terdiri dari nilai eigen terbesar $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ dan E_1 sebagai vektor eigen $e_1, e_2, e_3, \dots, e_m$ yang bersesuaian \ddot{E}_1 . Maka dapat diperoleh nilai dari \hat{L} sebagai berikut:

$$\hat{L}_{(p \times m)} = E_{1(p \times m)} \Lambda_{21(m \times m)}^{\frac{1}{2}}$$

atau

$$\begin{aligned}
\hat{L}_{(p \times m)} &= [e_1 \sqrt{\lambda_1} : e_1 \sqrt{\lambda_1} : \dots : e_m \sqrt{\lambda_m}]_{(p \times m)} \\
&= [\sqrt{\lambda_1} e_1 : \sqrt{\lambda_2} e_2 : \dots : \sqrt{\lambda_m} e_m]_{(p \times m)}
\end{aligned}$$

dengan $\hat{L} = [\hat{l}_{i1}, \hat{l}_{i2}, \dots, \hat{l}_{im}]$, $i = 1, 2, \dots, p$.

Persamaan (2.20) ketika diaplikasikan pada matriks S atau matriks korelasi R disebut sebagai solusi komponen utama. Pada analisis faktor, komponen utama pada matriks kovarians sampel S merupakan pasangan nilai eigen dan vektor

eigen $(\widehat{\lambda}_1, \widehat{e}_1), (\widehat{\lambda}_2, \widehat{e}_2), \dots, (\widehat{\lambda}_p, \widehat{e}_p)$ dengan $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$. Jika $m < p$ merupakan banyaknya faktor umum yang terbentuk dari p komponen atau variabel yang diamati, maka estimasi dari *loading factor* adalah sebagai berikut:

$$\widehat{L} = [\sqrt{\widehat{\lambda}_1}e_1 : \sqrt{\widehat{\lambda}_2}e_2 : \dots : \sqrt{\widehat{\lambda}_m}e_m]_{(p \times m)}$$

Estimasi dari varians khusus yang diperoleh dari elemen diagonal matriks $S - \widehat{L}\widehat{L}'$ yaitu:

$$\psi = \begin{bmatrix} \widehat{\psi}_1 & 0 & \dots & 0 \\ 0 & \widehat{\psi}_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \widehat{\psi}_p \end{bmatrix}$$

$$\widehat{\psi} = s_{ii} - \sum_{j=1}^m \widehat{l}_{ij}^2. \quad (2.27)$$

Sedangkan untuk varians khusus yang diperoleh dari elemen diagonal matriks $R - \widehat{L}\widehat{L}'$, yaitu:

$$\widehat{\psi} = 1 - \sum_{j=1}^m \widehat{l}_{ij}^2. \quad (2.28)$$

Estimasi dari kumunalitas yaitu:

$$\begin{aligned} \widehat{h}_i^2 &= \widehat{l}_{i1}^2 + \widehat{l}_{i2}^2 + \dots + \widehat{l}_{im}^2 \\ &= \sum_{j=1}^m \widehat{l}_{ij}^2 \end{aligned} \quad (2.29)$$

Idealnya kontribusi dari beberapa faktor umum awal terhadap varians sampel variabel seharusnya cukup besar. Kontribusi faktor umum pertama terhadap varians sampel s_{ii} dinyatakan \widehat{l}_{i1}^2 . Maka kontribusi dari faktor umum pertama terhadap varians total $tr(S) = s_{11} + s_{22} + \dots + s_{pp}$ didefinisikan sebagai berikut:

$$\widehat{l}_{11}^2 + \widehat{l}_{21}^2 + \dots + \widehat{l}_{p1}^2 = \sum_{i=1}^p \widehat{l}_{i1}^2 = \left(\sqrt{\widehat{\lambda}_1} \widehat{e}_{11} \right)^2$$

$$= \widehat{\lambda}_1 \sum_{i=1}^p (\widehat{e}_{i1})^2 = \lambda_1$$

dengan \widehat{e}_1 adalah vektor eigen satuan yang memiliki panjang satu. Sehingga secara umum kontribusi dari faktor umum ke- j terhadap varians total adalah:

$$\begin{aligned} \widehat{t}_{i1}^2 + \widehat{t}_{i2}^2 + \dots + \widehat{t}_{ip}^2 &= \sum_{i=1}^p \widehat{t}_{i1}^2 = \sum_{i=1}^p \left(\sqrt{\widehat{\lambda}_1} \widehat{e}_{i1} \right)^2 \\ &= \widehat{\lambda}_1 \sum_{i=1}^p (\widehat{e}_{i1})^2 \\ &= \widehat{\lambda}_1 \end{aligned} \quad (2.30)$$

Secara umum, proporsi dari varians sampel total yang berasal dari faktor umum ke- j adalah:

$$\begin{aligned} \frac{\widehat{\lambda}_j}{s_{11} + s_{21} + \dots + s_{pp}} &\text{ untuk analisis faktor pada } S \\ \frac{\widehat{\lambda}_j}{p} &\text{ untuk analisis faktor pada } R \end{aligned} \quad (2.31)$$

Kriteria pada persamaan (2.31) seringkali digunakan dalam menentukan banyaknya faktor umum. Dengan menggunakan program computer, dapat ditentukan banyaknya faktor umum berdasarkan pada banyaknya nilai eigen dari matriks korelasi R atau dari matriks varians kovarians S yang lebih dari rata-rata nilai eigen. Setelah seluruh nilai taksiran parameter didapatkan kemudian dihitung matriks sisa. Matriks sisa didefinisikan sebagai selisih dari korelasi sampel R atau dari matriks S dengan nilai-nilai estimasi yang diperoleh.

$$\text{Matriks sisa} = R - \widehat{L}\widehat{L}' + \widehat{\psi} \text{ atau } S - \widehat{L}\widehat{L}' + \widehat{\psi}$$

Tahap selanjutnya untuk mengetahui apakah nilai estimasi yang diperoleh merupakan solusi yang tepat, perlu dilakukan pengecekan dengan menghitung

nilai RMSR (*Root Mean Square Residual*) dari matriks sisa yang didapat. RMSR didefinisikan sebagai:

$$RMSR = \sqrt{\frac{\sum_{i=1}^p \sum_{j=1}^p res_{ij}^2}{\frac{p(p-1)}{2}}} \quad (2.32)$$

dengan:

res_{ij} : elemen-elemen matriks sisa selain elemen diagonal utama pada variabel ke- i dan variabel ke- j .

p : banyaknya variabel yang diamati,

Semakin kecil nilai RMSR yang diperoleh maka semakin baik model faktor yang diperoleh.

2.2.3 Kriteria Penentuan Jumlah Faktor

Analisis faktor digunakan untuk menghasilkan faktor yang jumlahnya lebih sedikit daripada jumlah variabel yang diolah. Pendekatan yang digunakan untuk menentukan berapa jumlah faktor yang diperoleh dalam penelitian ini adalah berdasarkan nilai eigen, persentase varians, dan *scree plot*.

Kriteria pertama dilakukan berdasarkan nilai eigen. Nilai eigen menunjukkan jumlah varians yang berhubungan pada suatu faktor. Faktor yang mempunyai nilai eigen lebih dari atau sama dengan 1 akan dipertahankan, dan faktor yang mempunyai nilai eigen kurang dari 1 tidak akan diikutsertakan dalam model, karena variabel yang nilainya kurang dari 1 tidak lebih baik daripada variabel aslinya (Supranto, 2004). Nilai eigen terakhir yang mempunyai nilai lebih dari atau sama dengan 1 tersebut dipilih sebagai titik penghentian ekstraksi.

Kriteria kedua adalah berdasarkan persentase varians. Jumlah faktor yang diambil ditentukan berdasarkan jumlah kumulatif varians yang telah dicapai. Jika

nilai kumulatif persentase variansnya sudah mencukupi (lebih dari setengah dari seluruh varians variabel awalnya) maka ekstraksi faktor dapat dihentikan.

Kriteria ketiga ditentukan berdasarkan *scree plot*. *Scree plot* adalah grafik yang menunjukkan relasi antara faktor dengan nilai eigennya. Penentuan kriteria ini dilakukan dengan membuat *plot* nilai eigen terhadap banyaknya faktor yang akan diekstraksi. Nilai eigen tersebut *diplotkan* pada arah vertikal, sedangkan banyaknya faktor (m) *diplotkan* pada arah horizontal. Banyaknya faktor pada kriteria ini ditentukan berdasarkan penurunan (*slope*) *plot* nilai eigen tersebut. Pada saat *scree* mulai mendatar atau merata dan nilai eigen berada pada nilai lebih dari satu dan kurang dari satu, di sinilah terdapat titik penghentian ekstraksi jumlah faktor. Titik tersebut menunjukkan banyaknya faktor yang dapat diekstraksi.

2.2.4 Rotasi Faktor

Menurut Supranto (2004), tujuan utama dari proses rotasi adalah tercapainya kesederhanaan terhadap faktor dan meningkatnya kemampuan interpretasinya. Dua metode rotasi dalam analisis faktor yang terus dikembangkan oleh banyak peneliti adalah metode Rotasi Ortogonal dan metode Rotasi *Oblique*. Rotasi ortogonal merupakan rotasi yang dilakukan dengan mempertahankan sumbu secara tegak lurus satu dengan yang lainnya. Dengan melakukan rotasi ini, maka setiap faktor *independent* terhadap faktor lain karena sumbunya saling tegak lurus. Rotasi ortogonal digunakan bila analisis bertujuan untuk mereduksi jumlah variabel tanpa mempertimbangkan seberapa berartinya faktor yang diekstraksi.

Prosedur perotasian *oblique* tidak mempertahankan sumbu tegak lurus lagi. Dengan rotasi ini maka korelasi antar faktor masih diperhitungkan, karena

sumbu faktor tidak saling tegak lurus satu dengan lainnya. Rotasi *oblique* digunakan untuk memperoleh jumlah faktor secara teoritis yang mampu menjelaskan variabel-variabel yang terkait.

Pada metode Rotasi Ortogonal dikenal beberapa pengukuran analitik, di antaranya metode *Quartimax*, *Varimax*, dan *Equimax*. Pada metode Rotasi *Quartimax*, tujuan akhir yang ingin dicapai adalah menyederhanakan baris sebuah matriks faktor. Nilai *loading factor* dirotasi sehingga sebuah variabel akan mempunyai *loading factor* tinggi pada salah satu faktor, dan faktor-faktor yang lain dibuat sekecil mungkin. Pemusatan metode rotasi ini adalah penyederhanaan struktur pada baris matriksnya. Metode ini tidak banyak dikembangkan oleh para peneliti karena tidak berhasil digunakan untuk memperoleh struktur yang sederhana. Pada akhirnya metode ini akan membuat sebuah faktor yang terlalu umum dan tujuan rotasi tidak akan dicapai.

Metode *Varimax* memfokuskan analisisnya pada penyederhanaan kolom matriks faktor. Penyederhanaan secara maksimum dapat terjadi apabila hanya ada nilai 0 dan 1 dalam sebuah kolom. Pada metode ini terjadi kecenderungan menghasilkan beberapa nilai *loading factor* yang tinggi (mendekati -1 atau +1) dan beberapa nilai *loading factor* mendekati 0 pada masing-masing kolom matriks. Logika interpretasi akan lebih mudah ketika korelasi antara faktor dan variabel bernilai +1 atau -1, karena hal ini mengindikasikan adanya asosiasi yang sempurna yang sifatnya positif atau negatif. Nilai 0 mengindikasikan adanya asosiasi yang sangat kurang. Teknik *varimax* mencoba menghasilkan nilai *loading factor* yang besar atau faktor lainnya sekecil mungkin. Struktur yang dihasilkan

ini jauh lebih sederhana jika dibandingkan dengan metode *Quartimax*. Selain itu, metode *Varimax* ini dapat membedakan faktor dengan lebih jelas.

Metode *Equimax* merupakan gabungan antara metode *Quartimax* dan *Varimax*. Fokus dari metode ini adalah dengan menyederhanakan baris atau kolom matriks faktor. Namun pada perkembangannya metode ini tidak diterima secara meluas atau jarang digunakan.

2.3 Uji Asumsi

2.3.1 Uji Kecukupan Data

Tahap pertama sebelum masuk ke tahap analisis factor, sebelumnya perlu dilakukan beberapa asumsi yaitu asumsi kecukupan data dan korelasi antar variabel. Uji kecukupan data diperlukan untuk memastikan bahwa data yang telah dikumpulkan dan disajikan dalam laporan penimbangan tersebut adalah cukup secara objektif. Uji *Kaiser Mayer Olkin* (KMO) bertujuan untuk mengetahui apakah semua data yang telah diambil cukup untuk difaktorkan. Hipotesis dari KMO adalah sebagai berikut:

H_0 : jumlah data cukup untuk analisis faktor

H_1 : jumlah data tidak cukup untuk analisis faktor

Statistik uji:

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2} \quad (2.33)$$

dengan:

$i = 1, 2, 3, \dots, p$ dan $j = 1, 2, \dots, p$ serta $i \neq j$

dengan:

r_{ij} : koefisien korelasi (hubungan antara 2 variabel) antara variabel i dan j .

a_{ij} : koefisien korelasi parsial (hubungan antara 2 variabel yang mengendalikan variabel lain) antara variabel i dan j .

Sampel akan dikatakan layak untuk dilakukan analisis faktor jika nilai KMO lebih dari 0,5 yang mana dengan kriteria ditunjukkan pada Tabel 2.1 (Hair, dkk, 1995).

Tabel 2.1 Klasifikasi Kecukupan Sampel

Nilai KMO	Keterangan
0,8-1,0	Sangat Bagus
0,7-0,8	Bagus
0,6-0,7	Cukup Bagus
0,5-0,6	Tidak Cukup Bagus
$\leq 0,5$	Tidak Layak

2.3.2 Uji *Bartlett* (Kebebasan Antar Variabel)

Uji *Bartlett* bertujuan untuk mengetahui apakah terdapat hubungan antar variabel dalam kasus multivariat. Jika variabel X_1, X_2, \dots, X_p *independent* (bersifat saling bebas), maka matriks korelasi antar variabel sama dengan matriks identitas. Sehingga untuk menguji kebebasan antar variabel ini, uji *Bartlett* menyatakan hipotesis sebagai berikut:

H_0 : $P = I$ (Tidak ada korelasi antar variabel)

H_1 : $P \neq I$ (Ada korelasi antar variabel)

Statistik uji:

$$X^2 = - \left(n - 1 - \frac{2p + 5}{6} \right) \ln |R| \quad (2.34)$$

dengan:

$\ln|R|$: nilai determinan dari matriks korelasi

- n : banyaknya observasi
 p : banyaknya variabel
 P : matriks korelasi
 I : matriks identitas.

Keputusan:

Tolak H_0 bila $X^2_{hitung} > X^2_{\alpha; 1/2p(p-1)}$ maka variabel saling berkorelasi, hal ini berarti terdapat hubungan antar variabel (Morisson, 1990).

2.4 Nilai Karakteristik dan Matriks Ortogonal

Menurut Supranto (2003), bentuk kuadrat dalam teori matriks mempunyai peranan penting di dalam aplikasi model-model statistik yang linier, misalnya teori tentang *multiple regression*, *experimental design*, dan lain sebagainya. Untuk membicarakan bentuk kuadrat tersebut diperlukan pengertian tentang nilai-nilai karakteristik (*characteristic values*).

Gudono (2011) menjelaskan ada beberapa teknik untuk menghitung nilai eigen dan vektor eigen. Salah satu cara sederhana adalah dengan menggunakan formula persamaan karakteristik (*characteristic equation*) matriks A berdimensi $n \times n$ berikut ini:

$$\det(A - \lambda I) = 0 \quad (2.35)$$

dengan memecahkan persamaan tersebut, maka akan diperoleh nilai eigen (λ) dan vektor eigen (X).

Menurut Gudono (2011), ada beberapa sifat istimewa nilai eigen dan vektor eigen, antara lain:

1. Jumlah nilai eigen sama dengan *trace* matriks yang bersangkutan.
2. Suatu matriks ataupun transposnya memiliki nilai eigen yang sama.

3. Hasil kali nilai-nilai eigen suatu matriks sama dengan determinan matriks tersebut.

Persoalan nilai karakteristik didefinisikan sebagai mencari nilai suatu skalar v dan bersama dengan vektor $X \neq 0$, memenuhi persamaan berikut:

$$AX = vX \quad (2.36)$$

dengan A suatu n matriks persegi, v dan X masing-masing disebut nilai eigen dari matriks A dan vektor eigen. Sebagai ilustrasi, misalnya untuk persamaan (2.36) diambil $n = 2$, maka:

$$\begin{aligned} AX &= vX \\ AX - vX &= (A - v)X = 0 \end{aligned} \quad (2.37)$$

atau dapat ditulis sebagai berikut:

$$\begin{bmatrix} a_{11} - v & a_{12} \\ a_{21} & a_{22} - v \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

sehingga persamaan menjadi:

$$\begin{aligned} (a_{11} - v)x_1 + a_{12}x_2 &= 0 \\ a_{21}x_1 + (a_{22} - v)x_2 &= 0 \end{aligned}$$

Artinya, persamaan (2.36) dan (2.37) sebenarnya sama. Apabila matriks $(A - vI)$ *non-singular matrix*, dapat dikalikan persamaan (2.37) dengan inversnya, dan satu-satunya pemecahan adalah $X = 0$. Jadi akan ada suatu pemecahan yang mana $X \neq 0$, hanya jika matriks $(A - vI)$ adalah singular yaitu jika:

$$\det(A - vI) = 0 \quad (2.38)$$

(Supranto, 2003).

Persamaan (2.38) akan menghasilkan suatu *polynomial* dengan variabel yang tidak diketahui v , yang kemudian dapat dipecahkan untuk v sehingga akan

diperoleh nilai eigen. Sebagai contoh, jika $n = 2$, maka persamaan (2.38) dapat ditulis sebagai berikut:

$$\begin{vmatrix} a_{11} - v & a_{12} \\ a_{21} & a_{22} - v \end{vmatrix} = 0$$

yaitu:

$$(a_{11} - v)(a_{22} - v) - a_{12}a_{21} = v^2 - (a_{11} + a_{22})v + (a_{11}a_{22} - a_{12}a_{21}) = 0$$

Persamaan tersebut merupakan bentuk *polynomial* pangkat dua. Dengan rumus kuadrat untuk memecahkan persamaan tersebut sebagai berikut:

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

dengan:

$$x = v$$

$$a = 1$$

$$b = -(a_{11} + a_{22})$$

$$c = (a_{11}a_{22} - a_{12}a_{21})$$

sehingga diperoleh:

$$v_1 = \frac{1}{2}(a_{11} + a_{22}) + \sqrt{(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}$$

$$v_2 = \frac{1}{2}(a_{11} + a_{22}) - \sqrt{(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}$$

Dalam hal yang khusus yaitu yang menyangkut matriks simetris yang mana

$a_{12} = a_{21}$, sehingga:

$$v_1 = \frac{1}{2}(a_{11} + a_{22}) + \sqrt{(a_{11} + a_{22})^2 - 4a_{11}a_{22} + 4a_{12}^2}$$

$$v_2 = \frac{1}{2}(a_{11} + a_{22}) - \sqrt{(a_{11} + a_{22})^2 - 4a_{11}a_{22} + 4a_{12}^2}$$

(2.39)

(Supranto, 2003).

Misalkan matriks simetris A , $n \times n$, mempunyai nilai eigen v_1, v_2, \dots, v_n dan ada kemungkinan bahwa beberapa di antaranya mempunyai nilai yang sama. Bersesuaian dengan nilai eigen ini adalah himpunan vektor eigen. Vektor tersebut menjadi *standard* dengan jalan membuat suatu jumlah kuadrat elemen-elemennya mempunyai nilai satu, sehingga ortogonal (artinya masing-masing nilai eigen akan memberikan vektor eigen), misalkan C_1, C_2, \dots, C_n sedemikian sehingga:

$$C_i' C_j = 0, i \neq j; i, j = 1, 2, \dots, n$$

Tanpa menghilangkan sifat umum, vektor-vektor tersebut dapat dibuat *standard* (normal) sedemikian sehingga $C_i' C_i = 1$ untuk semua i . Suatu himpunan vektor-vektor ortogonal yang telah dibuat normal, disebut himpunan ortogonal (Supranto, 2003).

Misalkan C merupakan matriks $n \times n$, dengan kolom-kolomnya terdiri dari vektor-vektor C_i dan kemudian ditulis dua syarat sebagai berikut:

$$C_i' C_j = d_{ij}$$

dengan:

$$d_{ij} = 0, \text{ jika } i \neq j$$

$$d_{ij} = 1, \text{ jika } i = j$$

maka akibatnya:

$$C' C = I_n$$

sehingga:

$$C' = C^{-1}$$

Yaitu, bahwa transpos dari C sama dengan inversnya. Matriks yang mempunyai sifat demikian dinamakan matriks ortogonal. Jadi suatu matriks C yang dibentuk

dari vektor eigen yang *standard* dari suatu matriks akan merupakan matriks ortogonal, yaitu transpos sama dengan invers:

$$C' = C^{-1} \text{ atau } C'C = CC' = I \quad (2.40)$$

(Supranto, 2003).

2.5 Principal Component Analysis (PCA)

Menurut Supranto (2004), PCA merupakan suatu teknik mereduksi data multivariat (banyak data) yang mencari untuk mengubah (mentransformasi) suatu matriks data awal/asli menjadi satu himpunan kombinasi linier yang lebih sedikit, akan tetapi menyerap sebagian besar jumlah varians dari data awal. Tujuan utama dari PCA adalah menjelaskan sebanyak mungkin jumlah varians data asli dengan sedikit mungkin komponen utama yang disebut faktor. Banyaknya faktor (komponen) yang dapat diekstrak dari data awal/asli adalah sebanyak variabel yang ada. Misalkan ada m komponen (faktor) yang dapat diekstrak dari p variabel asli, maka paling banyak $m = p$, artinya banyaknya faktor sama dengan banyaknya variabel. Hal ini tidak diinginkan, sebab menjadi tidak hemat, maka m harus kurang dari p , artinya banyaknya komponen atau faktor yang harus dipertahankan harus sedikit mungkin, akan tetapi sudah mencakup sebagian besar informasi yang terkandung di dalam data asli.

Yamin (2011), juga menjelaskan PCA pada dasarnya teknik statistik yang bertujuan untuk menyederhanakan variabel yang diamati dengan cara mereduksi dimensinya (disebut juga sebagai teknik pereduksian data). Prinsip utama dalam PCA adalah terdapatnya korelasi di antara variabel. Sehingga jika hal ini terjadi, maka ada estimasi peneliti bahwa sesungguhnya beberapa variabel tersebut dapat

direduksi. Praktisnya, teknik PCA ini berusaha untuk mereduksi dengan mengelompokkan variabel yang saling berkorelasi dan melakukan proses transformasi variabel tersebut ke dalam suatu komponen/dimensi baru yang mana tidak berkorelasi sama sekali. Setelah beberapa komponen hasil PCA yang bebas multikolinieritas diperoleh, komponen-komponen tersebut menjadi variabel estimasi baru yang akan diregresikan atau dianalisis pengaruhnya terhadap variabel respon (Y) dengan menggunakan analisis regresi.

Johnson dan Wichern (1998) menjelaskan cara pembentukan komponen utama ada dua cara, yaitu pembentukan komponen utama berdasarkan matriks kovarians dan pembentukan komponen utama berdasarkan matriks korelasi. Penggunaan matriks kovarians dapat dilakukan jika variabel yang diamati mempunyai satuan pengukuran yang sama, sedangkan matriks korelasi digunakan jika satuan dari variabel yang diamati berbeda. Secara umum tahapan menentukan komponen utama untuk data dengan skala pengukuran tidak sama, sehingga dapat dituliskan sebagai berikut:

1. Matriks Z yang merupakan matriks yang berisi data dari variabel prediktor X yang distandarisasi atau dibakukan.
2. $Z'Z$ adalah matriks korelasi dari matriks Z . Cara mereduksi komponen utama dimulai dari prosedur seleksi nilai eigen $\lambda_1, \lambda_2, \dots, \lambda_p$ yang mana jumlah dari nilai eigen ini akan sama dengan *trace* matriks korelasi atau jumlah diagonal matriks korelasi, yaitu:

$$\sum_{j=1}^n \lambda_j = tr(Z'Z)$$

Jika nilai eigen diurutkan dari nilai terbesar sampai nilai terkecil, maka pengaruh komponen utama K_j berbeda dengan pengaruh λ_j . Ini berarti bahwa komponen-komponen tersebut menerangkan proporsi keragaman terhadap variabel respon Y yang semakin lama semakin kecil. Komponen utama K_j saling ortogonal sesamanya dan melalui suatu hubungan:

$$K_j = g_{1j}z_1 + g_{2j}z_2 + \dots + g_{pj}z_p \quad (2.41)$$

Vektor eigen g_j diperoleh dari setiap nilai eigen yang memenuhi suatu sistem persamaan:

$$(Z'Z - \lambda_j I)g_j = 0$$

2.6 Calinski-Harabasz *Pseudo F-statistic*

Salah satu metode alternatif yang digunakan untuk menentukan banyaknya kelompok optimum adalah *Pseudo F-statistic* yang dirumuskan oleh Calinski dan Harabasz (1974). Penelitian oleh Miligan dan Cooper (1985) menunjukkan bahwa *Pseudo F-statistic* yang selanjutnya disebut *Pseudo F*, memberikan hasil terbaik di antara 3 metode yang merupakan metode yang dapat digunakan secara global. *Pseudo F* tertinggi pada beberapa simulasi menunjukkan bahwa kelompok tersebut mampu memberikan hasil yang optimum, yang mana keragaman dalam kelompok sangat homogen sedangkan antar kelompok sangat heterogen. Rumus *Pseudo F* tertulis pada persamaan (2.42) (Orpin dan Kostylev, 2006).

$$Pseudo F = \frac{\left(\frac{R^2}{c-1}\right)}{\left(\frac{1-R^2}{n-c}\right)} \quad (2.42)$$

$$R^2 = \frac{(SST - SSW)}{SST} \quad (2.43)$$

$$SST = \sum_{i=1}^c \sum_{j=1}^{n_c} \sum_{k=1}^p (x_{ijk} - \bar{x}_k)^2 \quad (2.44)$$

$$SSW = \sum_{i=1}^c \sum_{j=1}^{n_c} \sum_{k=1}^p (x_{ijk} - \bar{x}_{ik})^2 \quad (2.45)$$

dengan:

SST : total jumlah dari kuadrat jarak terhadap rata-rata keseluruhan

SSW : total jumlah dari kuadrat jarak sampel terhadap rata-rata kelompoknya

n : banyaknya sampel

c : banyaknya kelompok

n_c : banyaknya data pada kelompok n_c

p : banyaknya variabel

x_{ijk} : kelompok ke- i pada sampel ke- j dan variabel ke- k

\bar{x}_k : rata-rata seluruh sampel pada variabel- k

\bar{x}_{jk} : rata-rata kelompok ke- i pada variabel ke- k

2.7 Internal Cluster Dispersion Rate (icdrate)

Ada beberapa kriteria dalam menilai kebaikan pengelompokan yang pada intinya untuk menilai homogenitas dalam kelompok dan heterogenitas antar kelompok. Merujuk pada Eviritt dalam Mingoti dan Lima (2006), perbandingan metode pengelompokan dapat diukur dengan menghitung rata-rata persebaran internal kelompok terhadap partisi secara keseluruhan. Metode ini seringkali

digunakan dalam mengestimasi akurasi dari algoritma pengelompokan. Perhitungan *icdrate* ditunjukkan oleh persamaan berikut:

$$icdrate = 1 - \frac{SSB}{SST} = 1 - SST - \frac{SSW}{SST} = 1 - R^2 \quad (2.46)$$

$$SSB = \sum_{j=1}^c \sum_{k=1}^p (\bar{x}_{jk} - \bar{x}_k)^2 \quad (2.47)$$

dengan:

SSB : *Sum of Squared Between-Groups*

SST : *Total Sum of Squared Partition*

R^2 : *Recovery Rate = SSB/SST*

c : banyaknya kelompok

n : banyaknya sampel

2.8 Multivariate Analyze of Varians (MANOVA)

Analisis statistika multivariat digunakan untuk membandingkan rata-rata dua populasi atau lebih. Pengertian tersebut menjelaskan bahwa metode analisis varians multivariat digunakan untuk mengkaji pengaruh dari satu atau lebih perlakuan terhadap respon. Adapun asumsi yang harus dipenuhi sebelum melakukan pengujian dengan MANOVA yaitu matriks varians kovarians antar perlakuan identik/*homogeneity*, sampel acak dari populasi yang berbeda adalah *independent* dan setiap populasi memiliki distribusi multivariat normal (Johnson dan Wichern, 2007).

2.8.1 Pemeriksaan Distribusi Normal Multivariat

Menurut Johnson dan Wichern (2007), pemeriksaan data multivariat normal dengan menggunakan *plot* phi-kuadrat yang berdasarkan pada perhitungan

nilai jarak kuadrat (d_j^2). Pemeriksaan multivariat normal dilakukan dengan cara membuat *plot* khi-kuadrat dari d_j^2 dan q_j . Tahap-tahapan dalam pembuatan *plot* khi-kuadrat adalah sebagai berikut:

- a. Menghitung d_j^2 yaitu jarak umum yang dikuadratkan dengan perhitungan sebagai berikut:

$$d_j^2 = (X_j - \bar{X})'S^{-1}(X_j - \bar{X}) \quad (2.48)$$

dengan S^{-1} merupakan invers matriks varians kovarians yang berukuran $k \times k$ dengan elemen matriks adalah sebagai berikut,

$$S_{jk} = \frac{\sum_{i=1}^n (X_{ij} - \bar{X}_j)(X_{ik} - \bar{X}_k)}{n - 1} \quad (2.49)$$

- b. Mengurutkan nilai d_j^2 dari nilai yang terkecil sampai yang terbesar,

$$d_{(1)}^2 \leq d_{(2)}^2 \leq \dots \leq d_{(n)}^2 \quad (2.50)$$

- c. Membuat *plot* d_j^2 dengan titik koordinat $X^2 \left(p \frac{j-0,5}{n} \right)$.

2.8.2 Pengujian MANOVA

MANOVA memiliki model sebagai berikut:

$$X_{ij} = \mu + \tau_i + e_{ij}, j = 1, 2, \dots, n \text{ dan } i = 1, 2, \dots, g$$

Hipotesis :

$$H_0 : \tau_1 = \tau_2 = \dots = \tau_g = 0$$

$$H_1 : \text{Minimal ada satu } \tau_g \neq 0$$

Statistik uji yang digunakan adalah *Wilk's Lambda* dengan perhitungan melalui tabel MANOVA. Dengan nilai *Wilk's Lambda* sebagai berikut:

$$\Lambda^* = \frac{|W|}{|B + W|} \quad (2.51)$$

Nilai *Wilk's Lambda* dapat diketahui dengan statistik uji *F* yaitu tolak H_0 jika nilai

$$\Lambda^* > F_{n_l-1, n-n_l^*}$$

dengan:

W : matriks *sum of square residuals*

B : matriks *sum of square treatment*

n : jumlah sampel

l : banyaknya kelompok

n_l : banyaknya anggota pada kelompok l

2.9 Pembangunan Manusia

Manusia adalah kekayaan bangsa yang sesungguhnya. Tujuan utama dari pembangunan adalah menciptakan lingkungan yang memungkinkan bagi rakyatnya untuk menikmati umur panjang, sehat, dan menjalankan kehidupan produktif. Hal ini tampaknya merupakan suatu kekayaan yang sederhana. Tetapi hal ini seringkali terlupakan oleh berbagai kesibukan jangka pendek untuk mengumpulkan harta dan uang. Pada tahun 1990 *United Nations Development Program* (UNDP) dalam laporannya “*Global Human Development Report*” memperkenalkan konsep “Pembangunan Manusia (*Human Development*)” sebagai paradigma baru model pembangunan. Menurut UNDP, pembangunan manusia dirumuskan sebagai perluasan pilihan bagi penduduk (*enlarging the choices of people*), yang dapat dilihat sebagai proses upaya ke arah “perluasan pilihan” dan sekaligus sebagai taraf yang dicapai dari upaya tersebut. Pada saat yang sama pembangunan manusia dapat dilihat juga sebagai pembangunan (*formation*) kemampuan manusia melalui perbaikan taraf kesehatan,

pengetahuan, dan keterampilan, sekaligus sebagai pemanfaatan (*utilization*) kemampuan/ketrampilan mereka. Konsep pembangunan tersebut jauh lebih luas pengertiannya dibandingkan konsep pembangunan ekonomi yang menekankan pada pertumbuhan (*economic growth*), kebutuhan dasar, kesejahteraan masyarakat, atau pengembangan sumber daya manusia.

Menurut UNDP (1995), untuk memperluas pilihan-pilihan manusia, konsep pembangunan manusia harus dibangun dari empat dimensi yang tidak terpisahkan. Berdasarkan konsep di atas maka untuk menjamin tercapainya tujuan pembangunan manusia, ada empat unsur pokok yang perlu diperhatikan yaitu:

a. Produktivitas (*Productivity*)

Masyarakat harus mampu untuk meningkatkan produktifitas mereka dan berpartisipasi penuh dalam proses mencari penghasilan dan lapangan pekerjaan. Oleh karena itu, pembangunan ekonomi merupakan bagian dari model pembangunan manusia.

b. Pemerataan (*equality*)

Masyarakat harus mempunyai akses untuk memperoleh kesempatan yang adil. Semua hambatan terhadap peluang ekonomi dan politik harus dihapuskan sehingga masyarakat dapat berpartisipasi di dalam dan memperoleh manfaat dari peluang-peluang yang ada.

c. Kestinambungan (*Sustainability*)

Akses untuk memperoleh kesempatan harus dipastikan bahwa tidak hanya untuk generasi sekarang tetapi juga untuk generasi yang akan datang. Semua jenis pemodalan baik itu fisik, manusia, dan lingkungan hidup harus dilengkapi.

d. Pemberdayaan (*Empowerment*)

Pembangunan harus dilakukan oleh masyarakat, dan bukan hanya untuk mereka. Masyarakat harus berpartisipasi penuh dalam mengambil keputusan dan proses-proses yang mempengaruhi kehidupan mereka.

2.9.1 Indeks Pembangunan Manusia (IPM)

IPM merupakan indeks komposit yang dihitung sebagai rata-rata sederhana dari tiga indeks dasar yaitu indeks harapan hidup, indeks pendidikan, dan indeks standar hidup layak. Menurut UNDP, IPM mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup. Menurut BPS (2015), IPM dibangun melalui pendekatan tiga dimensi dasar yaitu:

- a. Dimensi umur panjang dan kesehatan
- b. Dimensi pengetahuan
- c. Dimensi kehidupan yang layak

Ketiga dimensi tersebut memiliki pengertian sangat luas karena terkait banyak faktor. Untuk mengukur dimensi umur panjang dan kesehatan, digunakan angka harapan hidup waktu lahir. Selanjutnya untuk mengukur dimensi pengetahuan digunakan gabungan indikator angka buta huruf, rata-rata lama sekolah, dan angka partisipasi sekolah. Adapun untuk mengukur dimensi kehidupan yang layak digunakan indikator kemampuan daya beli (*Purchasing Power Parity*). Kemampuan daya beli masyarakat terhadap sejumlah kebutuhan pokok yang dilihat dari rata-rata besarnya pengeluaran perkapita sebagai pendekatan pendapatan yang mewakili capaian pembangunan untuk hidup layak.

2.10 Cluster dalam Pandangan Islam

Istilah *cluster* memiliki arti yang sama dengan kelompok-kelompok. Dalam Islam dijelaskan pula hal-hal yang berkaitan dengan kelompok atau pengelompokan. Salah satu pengelompokan yang telah diajarkan Islam adalah tentang ciri-ciri seseorang yang termasuk golongan orang yang diberi petunjuk oleh Allah Swt. ataupun yang tidak. Seperti pada firman Allah Swt. berikut ini:

مَنْ يَهْدِ اللَّهُ فَهُوَ الْمُهْتَدِيٌّ وَمَنْ يُضِلِّ فَأُولَئِكَ هُمُ الْخَاسِرُونَ ﴿١٧٨﴾

“Barang siapa yang diberi petunjuk oleh Allah Swt., maka dialah yang memperoleh petunjuk; dan barang siapa yang disesatkan Allah Swt., maka merekalah orang-orang yang merugi”(QS. Al-A’raf/7:178).

Pada penjelasan Ibnu Katsir, Allah Swt. berfirman bahwa barang siapa yang diberi petunjuk oleh Allah Swt., maka tidak ada yang dapat menyesatkannya dan barang siapa yang disesatkan oleh-Nya, maka sesungguhnya dia telah merugi, kecewa, dan sesat tanpa dapat dielakkan lagi. Karena sesungguhnya sesuatu yang dikehendaki oleh Allah Swt. pasti terjadi, dan sesuatu tidak dikehendaki-Nya pasti tidak terjadi (Abdullah, 2007).

Surat al-A’raf ayat 178 menjelaskan tentang golongan orang-orang yang diberikan petunjuk oleh Allah Swt. dengan golongan orang-orang yang merugi. Jika seseorang memiliki satu ciri-ciri yang sama seperti yang dipaparkan ayat di atas maka dia akan dikelompokkan ke dalam kelompok yang memiliki sifat atau ciri-ciri yang sama sepertinya. Begitu pula pada analisis *cluster* ini, jika suatu objek memiliki kesamaan atau kedekatan jarak dengan suatu variabel maka akan terletak pada *cluster* yang sama.

BAB III

METODE PENELITIAN

3.1 Pendekatan Penelitian

Pendekatan penelitian yang digunakan pada penelitian ini adalah pendekatan studi literatur deskriptif kuantitatif. Pada studi literatur yaitu dengan mengumpulkan bahan-bahan pustaka yang dibutuhkan oleh peneliti sebagai acuan dalam menyelesaikan penelitian. Sedangkan pendekatan deskriptif kuantitatif yaitu dengan menganalisis data dan menyusun data yang sudah ada sesuai dengan kebutuhan peneliti.

3.2 Sumber Data

Pada penelitian ini data yang digunakan adalah indikator IPM di kabupaten/kota Jawa Timur, yang bersumber dari Badan Pusat Statistik (BPS) Provinsi Jawa Timur. Unit observasi penelitian ini adalah 29 kabupaten dan 9 kota di Provinsi Jawa Timur.

3.3 Variabel Penelitian

Adapun variabel yang digunakan dalam penelitian ini, antara lain:

X_1 : persentase tingkat partisipasi angkatan kerja.

X_2 : persentase tingkat pengangguran terbuka.

X_3 : persentase angka kematian bayi.

X_4 : persentase angka harapan hidup.

X_5 : persentase angka keluhan kesehatan.

X_6 : persentase Angka Partisipasi Sekolah (APS) kelompok usia 16-18 tahun.

X_7 : persentase Angka Partisipasi Kasar (APK) kelompok usia 16-18 tahun.

X_8 : persentase Angka Partisipasi Murni (APM) kelompok usia 16-18 tahun.

X_9 : persentase angka penduduk 10 tahun ke atas yang tidak pernah bersekolah.

X_{10} : persentase angka buta huruf penduduk 10 tahun ke atas.

X_{11} : persentase angka penduduk miskin.

3.4 Analisis Data

3.4.1 Pembuktian Asumsi Analisis Faktor

Langkah-langkah pembuktian asumsi analisis faktor sebagai berikut:

1. Membuktikan bahwa asumsi $Cov(F)$ merupakan identitas.
2. Membuktikan bahwa asumsi $Cov(\varepsilon)$ merupakan matriks diagonal.
3. Membuktikan bahwa asumsi $Cov(\varepsilon, F)$ saling bebas.
4. Menarik kesimpulan

3.4.2 Pengelompokan Kabupaten/Kota di Provinsi Jawa Timur

Adapun langkah-langkah analisis data untuk melakukan penelitian tentang indikator IPM di kabupaten/kota Provinsi Jawa Timur pada tahun 2015 antara lain:

1. Membuktikan asumsi-asumsi pada persamaan analisis faktor.
2. Melakukan analisis statistika deskriptif untuk mengkaji karakteristik kabupaten/kota di Provinsi Jawa Timur dengan menggunakan penyajian diagram.

3. Mengelompokkan kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator-indikator IPM tahun 2015 dengan langkah-langkah analisis sebagai berikut:
 - a. Melakukan penyelidikan apakah terdapat korelasi yang signifikan antar variabel dengan menggunakan uji *Barlett* dan KMO untuk kelayakan suatu data.
 - b. Melakukan analisis faktor untuk menganalisis lebih lanjut variabel-variabel yang dapat menggambarkan kelompok kabupaten/kota.
 - c. Memperoleh hasil kelompok kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator IPM tahun 2015 dengan metode *Cluster* Hirarki. Metode *Cluster* Hirariki yang digunakan adalah metode *Complete Linkage*, *Average Linkage*, dan *Ward's Linkage* dengan menggunakan acuan banyak kelompok optimum berdasarkan nilai *Pseudo F*.
4. Menentukan metode terbaik di antara 3 metode tersebut menggunakan nilai *icdrate*.
5. Membuat pemetaan dari hasil pengelompokan kabupaten/kota di Provinsi Jawa Timur yang hasil kelompoknya sudah *optimum* dengan peta tematik menggunakan *software Geoda*.
6. Mengetahui perbandingan antar kelompok yang terbentuk menggunakan pengujian MANOVA dengan langkah-langkah sebagai berikut:
 - a. Melakukan pengujian asumsi normal multivariat.
 - b. Melakukan pengujian independensi untuk mengetahui apakah matriks varians kovarians bersifat homogen atau tidak.

- c. Melakukan pengujian vektor rata-rata untuk mengetahui adanya perbedaan yang signifikan antar kelompok yang telah terbentuk di Provinsi Jawa Timur.

7. Menarik kesimpulan.

3.4.3 Karakteristik Setiap Kelompok

Langkah-langkah dalam menentukan karakteristik setiap kelompok adalah sebagai berikut:

1. Menghitung nilai rata-rata dari setiap variabel yang terdapat dalam masing-masing kelompok yang terbentuk.
2. Menentukan nilai rata-rata yang terbesar dari setiap variabel dalam masing-masing kelompok yang akan menjadi karakteristik kelompok tersebut.
3. Menarik kesimpulan.

BAB IV PEMBAHASAN

4.1 Pembuktian Asumsi pada Model Persamaan Analisis Faktor

Secara umum analisis faktor disusun seperti model dalam analisis regresi multivariat. Setiap variabel awal dinyatakan sebagai kombinasi linier dari faktor-faktor yang mendasari. Misalkan vektor acak X , dengan banyak komponen p dan mempunyai mean μ dan matriks kovarians Σ merupakan penyusunan model faktor. Secara matematis model analisis faktor ditulis seperti pada persamaan (2.7).

Asumsi-asumsi yang harus dipenuhi oleh variabel random F dan ε dalam analisis faktor adalah:

Asumsi 1. Untuk $j = 1, 2, 3, \dots, m$, berlaku $E(F_j) = 0$, $Var(F_j) = 1$, dan $Cov(F_j, F_k) = 0$ untuk $j \neq k$ maka diperoleh $Cov(F) = E(FF') = I_{m \times m}$.

Bukti:

Berdasarkan asumsi 1 diperoleh bahwa F_j, F_k saling bebas,

$$Cov(F_j, F_k) = 0, E(F_j, F_k) = E(F_j)E(F_k)$$

Misal $F = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix}, F' = [F_1, F_2, \dots, F_m]$.

$$\begin{aligned} Var(F_j) &= E(F_j - E(F_j))^2 = 1 \\ &= E(F_j^2) - (E(F_j))^2 = 1, \text{ karena } E(F_j) = 0 \\ &= E(F_j^2) - (0)^2 = 1 \end{aligned}$$

$$= E(F_j^2) = 1.$$

$$\text{Cov}(F) = \text{Cov}(F_i, F_j) = \text{Cov} \begin{bmatrix} F_1 F_1 & F_1 F_2 & \cdots & F_1 F_m \\ F_2 F_1 & F_2 F_2 & \cdots & F_2 F_m \\ \vdots & \vdots & \ddots & \vdots \\ F_m F_2 & F_m F_2 & \cdots & F_m F_m \end{bmatrix}$$

$$= \text{Cov} \begin{bmatrix} F_1^2 & F_1 F_2 & \cdots & F_1 F_m \\ F_2 F_1 & F_2^2 & \cdots & F_2 F_m \\ \vdots & \vdots & \ddots & \vdots \\ F_m F_2 & F_m F_2 & \cdots & F_m^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$E(F, F') = I_{m \times m}$$

$$E(F, F') = E \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix} [F_1, F_2, \dots, F_m]$$

$$= E \begin{bmatrix} F_1 F_1 & F_1 F_2 & \cdots & F_1 F_m \\ F_2 F_1 & F_2 F_2 & \cdots & F_2 F_m \\ \vdots & \vdots & \ddots & \vdots \\ F_m F_2 & F_m F_2 & \cdots & F_m F_m \end{bmatrix} = E \begin{bmatrix} F_1^2 & F_1 F_2 & \cdots & F_1 F_m \\ F_2 F_1 & F_2^2 & \cdots & F_2 F_m \\ \vdots & \vdots & \ddots & \vdots \\ F_m F_2 & F_m F_2 & \cdots & F_m^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I_{m \times m}$$

Diperoleh $E(F_j) = 0$, untuk $j = 1, 2, 3, \dots, m$, dan $\text{Cov}(F) = E(F, F') = I_{m \times m}$.

Asumsi 2. Untuk $i = 1, 2, 3, \dots, p$ berlaku $E(\varepsilon_i) = 0$, $\text{Var}(\varepsilon_i) = \psi$, dan

$\text{Cov}(\varepsilon_i, \varepsilon_k) = 0$ untuk $i \neq k$, maka diperoleh $\text{Cov}(\varepsilon) = E(\varepsilon \varepsilon') = \Psi_{m \times m} =$

$$\begin{bmatrix} \varphi_1 & 0 & \cdots & 0 \\ 0 & \varphi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \varphi_p \end{bmatrix} \text{ yang mana } \Psi_{m \times m} \text{ adalah matriks diagonal.}$$

Bukti:

Berdasarkan asumsi 2 diperoleh bahwa $\varepsilon_i, \varepsilon_k$ saling bebas, $Cov(\varepsilon_i, \varepsilon_k) = 0$,

$$E(\varepsilon_i, \varepsilon_k) = E(\varepsilon_i)E(\varepsilon_k).$$

$$\text{Misal } \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{bmatrix}, \varepsilon' = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p].$$

$$\begin{aligned} Var(\varepsilon_i) &= E(\varepsilon_i - E(\varepsilon_i))^2 = \psi \\ &= E(\varepsilon_i^2) - (E(\varepsilon_i))^2 = \psi, \text{ karena } E(\varepsilon_i) = 0 \\ &= E(\varepsilon_i^2) - (0)^2 = \psi \\ &= E(\varepsilon_i^2) = \psi \end{aligned}$$

$$\begin{aligned} Cov(\varepsilon) &= Cov(\varepsilon_i, \varepsilon_k) = Cov \begin{bmatrix} \varepsilon_1 \varepsilon_1 & \varepsilon_1 \varepsilon_2 & \dots & \varepsilon_1 \varepsilon_p \\ \varepsilon_2 \varepsilon_1 & \varepsilon_2 \varepsilon_2 & \dots & \varepsilon_2 \varepsilon_p \\ \vdots & \vdots & & \vdots \\ \varepsilon_p \varepsilon_1 & \varepsilon_p \varepsilon_2 & \dots & \varepsilon_p \varepsilon_p \end{bmatrix} \\ &= Cov \begin{bmatrix} \varepsilon_1^2 & \varepsilon_1 \varepsilon_2 & \dots & \varepsilon_1 \varepsilon_p \\ \varepsilon_2 \varepsilon_1 & \varepsilon_2^2 & \dots & \varepsilon_2 \varepsilon_p \\ \vdots & \vdots & & \vdots \\ \varepsilon_p \varepsilon_1 & \varepsilon_p \varepsilon_2 & \dots & \varepsilon_p^2 \end{bmatrix} \\ &= \begin{bmatrix} \psi_1 & 0 & \dots & 0 \\ 0 & \psi_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & \varepsilon_p \varepsilon_2 & \dots & \psi_p \end{bmatrix} = \Phi_{p \times p}. \end{aligned}$$

$$E(\varepsilon, \varepsilon^2) = \Phi_{p \times p}$$

$$\begin{aligned} E(\varepsilon, \varepsilon^2) &= E \left[\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{bmatrix} [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p] \right] \\ &= E \begin{bmatrix} \varepsilon_1 \varepsilon_1 & \varepsilon_1 \varepsilon_2 & \dots & \varepsilon_1 \varepsilon_p \\ \varepsilon_2 \varepsilon_1 & \varepsilon_2 \varepsilon_2 & \dots & \varepsilon_2 \varepsilon_p \\ \vdots & \vdots & & \vdots \\ \varepsilon_p \varepsilon_1 & \varepsilon_p \varepsilon_2 & \dots & \varepsilon_p \varepsilon_p \end{bmatrix} = E \begin{bmatrix} \varepsilon_1^2 & \varepsilon_1 \varepsilon_2 & \dots & \varepsilon_1 \varepsilon_p \\ \varepsilon_2 \varepsilon_1 & \varepsilon_2^2 & \dots & \varepsilon_2 \varepsilon_p \\ \vdots & \vdots & & \vdots \\ \varepsilon_p \varepsilon_1 & \varepsilon_p \varepsilon_2 & \dots & \varepsilon_p^2 \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} \psi_1 & 0 & \cdots & 0 \\ 0 & \psi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \varepsilon_p \varepsilon_2 & \cdots & \psi_p \end{bmatrix} = \Phi_{p \times p}.$$

Sehingga diperoleh $E(\varepsilon_i) = 0$, untuk $i = 1, 2, 3, \dots, m$, dan $Cov(\varepsilon) = E(\varepsilon, \varepsilon') = \psi_{p \times p}$

Asumsi 3. Untuk F dan ε saling bebas, sehingga $Cov(\varepsilon, F) = E(\varepsilon, F) = 0_{p \times m}$.

Bukti:

$$E(\varepsilon, F') = E \left[\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{bmatrix} [F_1, F_2, \dots, F_m] \right]$$

$$= \begin{bmatrix} \varepsilon_1 \varepsilon_1 & \varepsilon_1 \varepsilon_2 & \cdots & \varepsilon_1 \varepsilon_p \\ \varepsilon_2 \varepsilon_1 & \varepsilon_2 \varepsilon_2 & \cdots & \varepsilon_2 \varepsilon_p \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_p \varepsilon_1 & \varepsilon_p \varepsilon_2 & \cdots & \varepsilon_p \varepsilon_p \end{bmatrix}$$

$$\text{karena } F \text{ dan } \varepsilon \text{ saling bebas, } E(\varepsilon, F') = E(\varepsilon)E(F') = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = 0_{p \times m}$$

$$\begin{aligned} Cov(\varepsilon, F) &= E(\varepsilon - E(\varepsilon))(F - E(F)) \\ &= E[\varepsilon F - \varepsilon E(F) - E(\varepsilon)F + E(\varepsilon)E(F)] \\ &= E(\varepsilon F) - E(\varepsilon)E(F) - E(\varepsilon)E(F) + E(\varepsilon)E(F) \\ &= E(\varepsilon F) - E(\varepsilon)E(F) = E(\varepsilon)E(F) - E(\varepsilon)E(F) = 0 \end{aligned}$$

$$Cov(\varepsilon, F) = Cov \begin{bmatrix} \varepsilon_1 F_1 & \varepsilon_1 F_2 & \cdots & \varepsilon_1 F_m \\ \varepsilon_2 F_1 & \varepsilon_2 F_2 & \cdots & \varepsilon_2 F_m \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_p F_1 & \varepsilon_p F_2 & \cdots & \varepsilon_p F_m \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = 0_{p \times m}$$

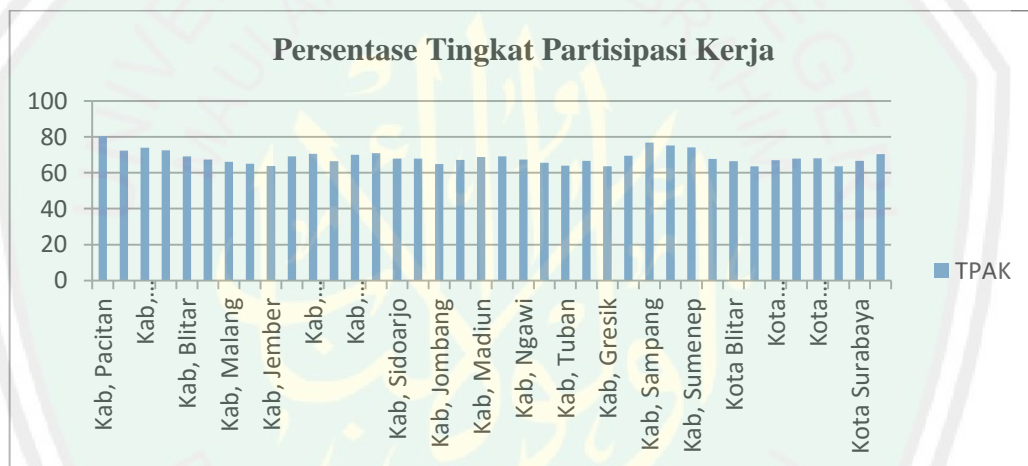
Sehingga untuk F dan ε yang saling bebas diperoleh:

$$Cov(\varepsilon, F) = E(\varepsilon, F') = 0_{p \times m}$$

4.2 Pengelompokan Kabupaten/Kota di Provinsi Jawa Timur

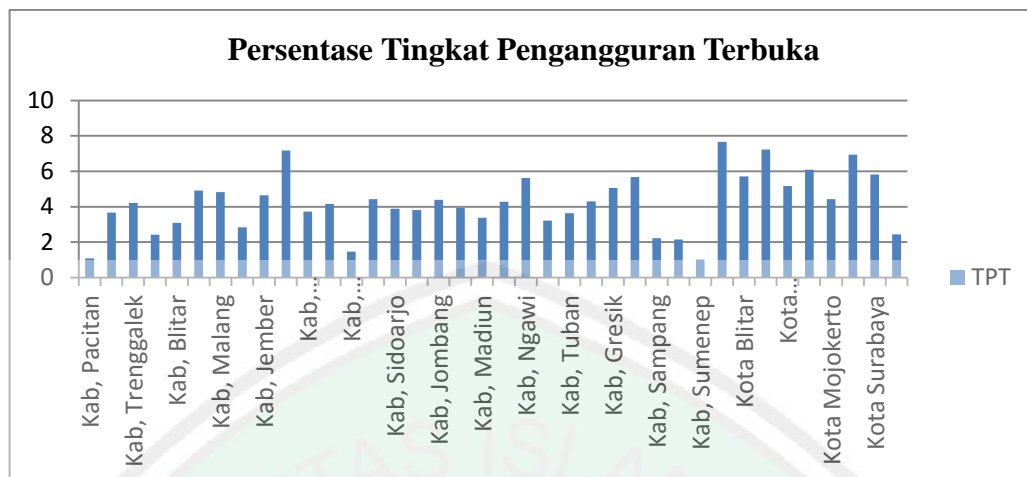
4.2.1 Deskriptif Indikator IPM 2015

Statistika deskriptif dilakukan untuk mengetahui gambaran secara umum karakteristik untuk masing-masing indikator yang digunakan. Berikut ini adalah data deskriptif dengan menggunakan diagram.



Gambar 4.1 Tingkat Partisipasi Angkatan Kerja

Berdasarkan Gambar 4.1 dapat diketahui bahwa tingkat partisipasi angkatan kerja di Provinsi Jawa timur, pada Kabupaten Pacitan mencapai 80,64% yang merupakan tingkat partisipasi angkatan kerja tertinggi. Sedangkan tingkat partisipasi angkatan kerja di Provinsi Jawa Timur terendah terjadi pada Kota Malang mencapai 60,56% dan terendah kedua terjadi pada Kota Probolinggo mencapai 63.61%. Rata-rata tingkat partisipasi angkatan kerja di Provinsi Jawa Timur sebesar 68,27%.



Gambar 4.2 Tingkat Pengangguran Terbuka

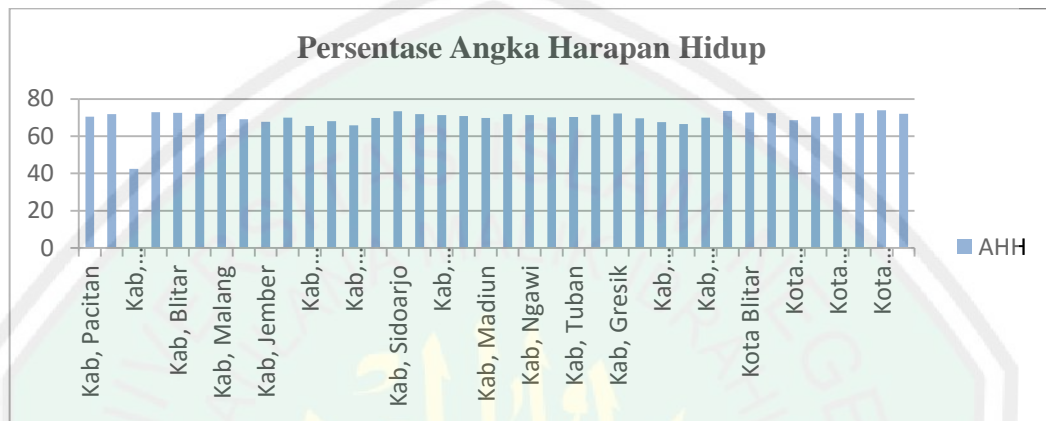
Berdasarkan Gambar 4.2 dapat diketahui bahwa tingkat pengangguran terbuka di Provinsi Jawa Timur yang tertinggi terjadi pada Kota Kediri mencapai 8,46%. Sedangkan tingkat pengangguran terbuka di Provinsi Jawa Timur terendah terjadi pada Kabupaten Pacitan mencapai 0,97% dan tingkat pengangguran terbuka terendah kedua terjadi pada Kabupaten Bondowoso mencapai 1,75%. Rata-rata tingkat pengangguran terbuka di Provinsi Jawa Timur sebesar 4,36%.



Gambar 4.3 Angka Kematian Bayi

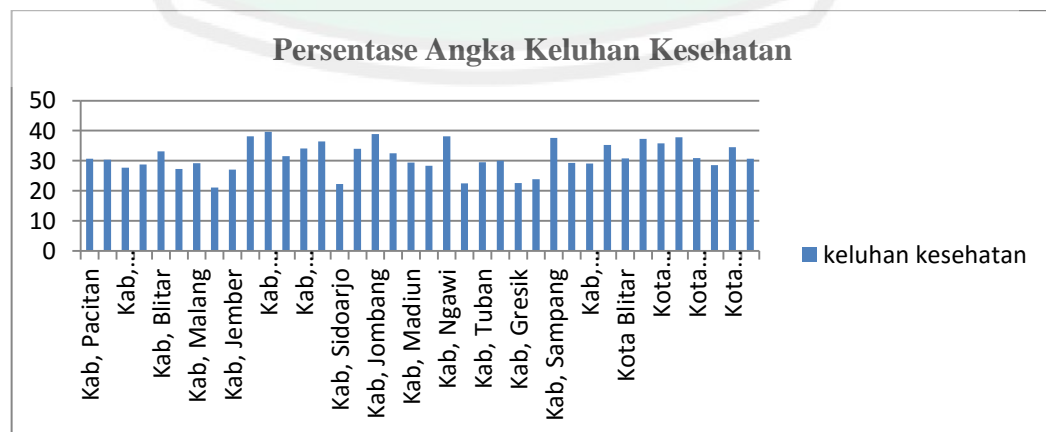
Berdasarkan Gambar 4.3 dapat diketahui bahwa angka kematian bayi di Provinsi Jawa Timur pada tahun 2015 yang tertinggi terjadi pada Kabupaten Probolinggo mencapai 61,48 % dan tertinggi kedua terjadi pada Kabupaten Jember mencapai 54,72%. Sedangkan untuk angka kematian bayi terendah terjadi

pada Kota Blitar mencapai 17,99 % dan untuk angka kematian bayi yang terendah kedua terjadi pada Kabupaten Trenggalek yang mencapai 20,23%. Rata-rata angka kematian bayi untuk seluruh kabupaten/kota di Provinsi Jawa Timur yaitu 31,91%.



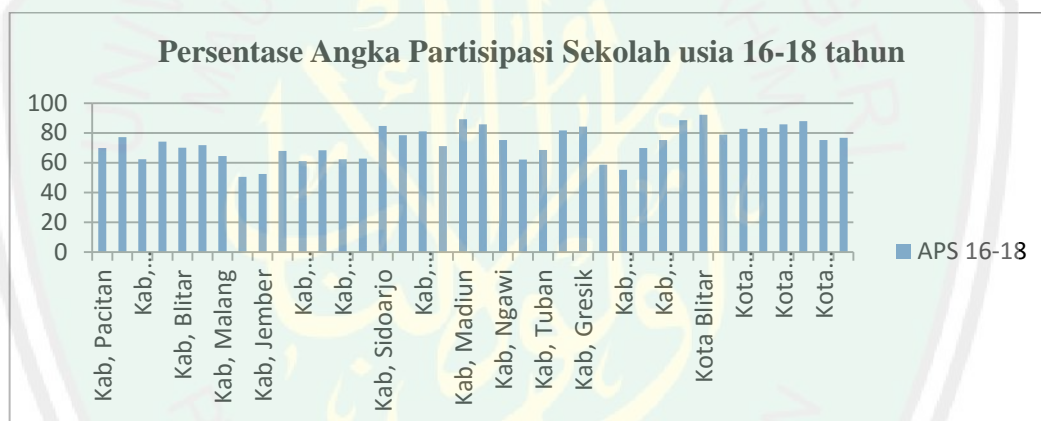
Gambar 4.4 Angka Harapan Hidup

Berdasarkan Gambar 4.4 dapat diketahui bahwa angka harapan hidup di Provinsi Jawa Timur pada tahun 2015 yang tertinggi terjadi pada Kota Surabaya mencapai 73,85% dan disusul oleh Kota Kediri mencapai 73,52% yang merupakan angka harapan hidup tertinggi kedua. Sedangkan untuk angka harapan hidup terendah terjadi pada Kabupaten Trenggalek mencapai 42,51%. Rata-rata angka harapan hidup untuk seluruh kabupaten/kota di Provinsi Jawa Timur yaitu 69,90%.



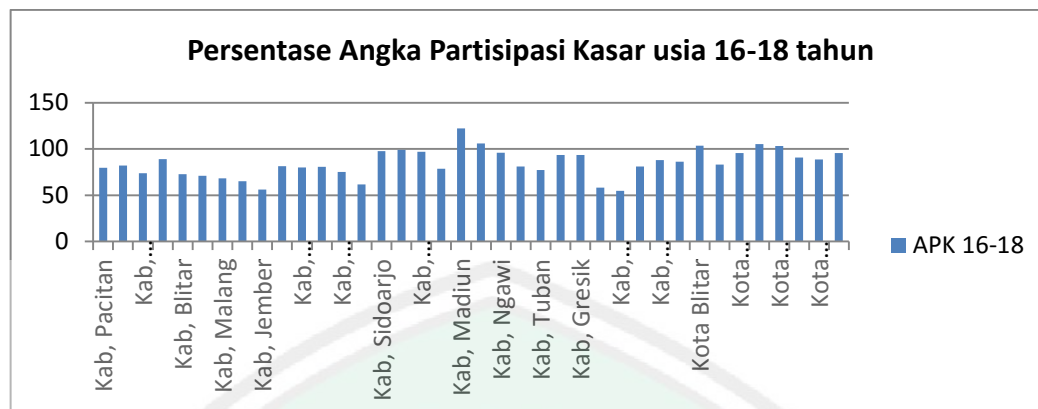
Gambar 4.5 Angka Keluhan Kesehatan

Berdasarkan Gambar 4.5 dapat diketahui bahwa yang mempunyai keluhan kesehatan di Provinsi Jawa Timur pada tahun 2015 yang tertinggi terjadi pada Kabupaten Bondowoso mencapai 39,63% dan disusul oleh Kabupaten Jombang mencapai 38,91% yang mempunyai keluhan kesehatan tertinggi kedua. Sedangkan yang mempunyai keluhan kesehatan terendah terjadi pada Kabupaten Lumajang mencapai 21,12%, selanjutnya yang mempunyai keluhan kesehatan terendah kedua Kabupaten Sidoarjo mencapai 22,31%. Rata-rata yang mempunyai keluhan kesehatan untuk seluruh kabupaten/kota di Provinsi Jawa timur yaitu 31,18%.



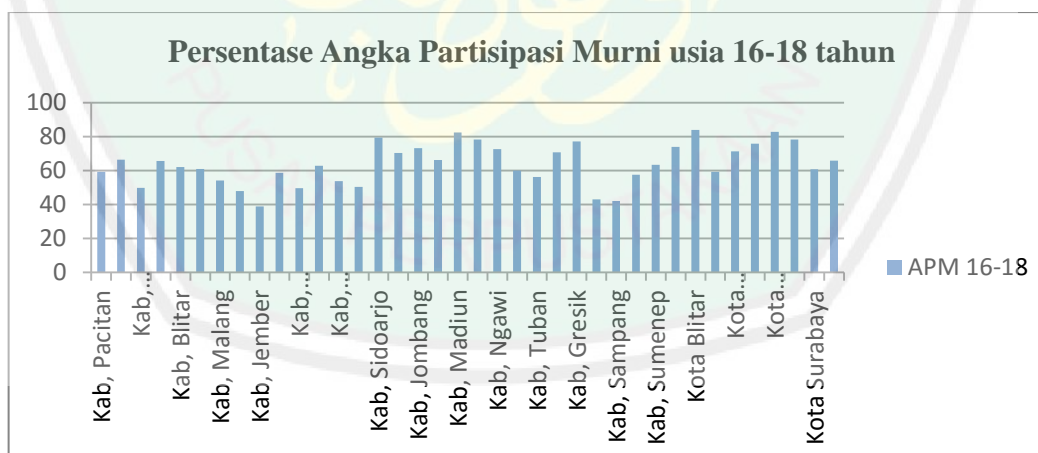
Gambar 4.6 Angka Partisipasi Sekolah usia 16-18 tahun

Berdasarkan Gambar 4.6 dapat diketahui bahwa angka partisipasi sekolah untuk umur 16-18 di Provinsi Jawa Timur pada tahun 2015 yang tertinggi terjadi pada Kota Blitar mencapai 92,17% dan disusul oleh Kabupaten Madiun mencapai 89,22% yang merupakan angka partisipasi sekolah untuk umur 16-18 tertinggi kedua. Sedangkan angka partisipasi sekolah untuk umur 16-18 terendah terjadi pada Kabupaten Lumajang mencapai 50,61%, selanjutnya angka partisipasi sekolah untuk umur 16-18 terendah kedua Kabupaten Jember mencapai 52,52%. Rata-rata angka partisipasi sekolah untuk umur 16-18 untuk seluruh kabupaten/kota di Provinsi Jawa Timur yaitu 73,34%.



Gambar 4.7 Angka Partisipasi Kasar Usia 16-18 tahun

Berdasarkan Gambar 4.7 dapat diketahui bahwa angka partisipasi kasar untuk umur 16-18 di Provinsi Jawa Timur pada tahun 2015 yang tertinggi terjadi pada Kabupaten Madiun mencapai 122,09%. Sedangkan angka partisipasi kasar untuk umur 16-18 terendah terjadi pada Kabupaten Sampang mencapai 54,63%, selanjutnya angka partisipasi kasar untuk umur 16-18 terendah kedua Kabupaten Jember mencapai 55,96%. Rata-rata angka partisipasi kasar untuk umur 16-18 untuk seluruh kabupaten/kota di Provinsi Jawa Timur yaitu 84,52%.



Gambar 4.8 Angka Partisipasi Murni Usia 16-18 tahun

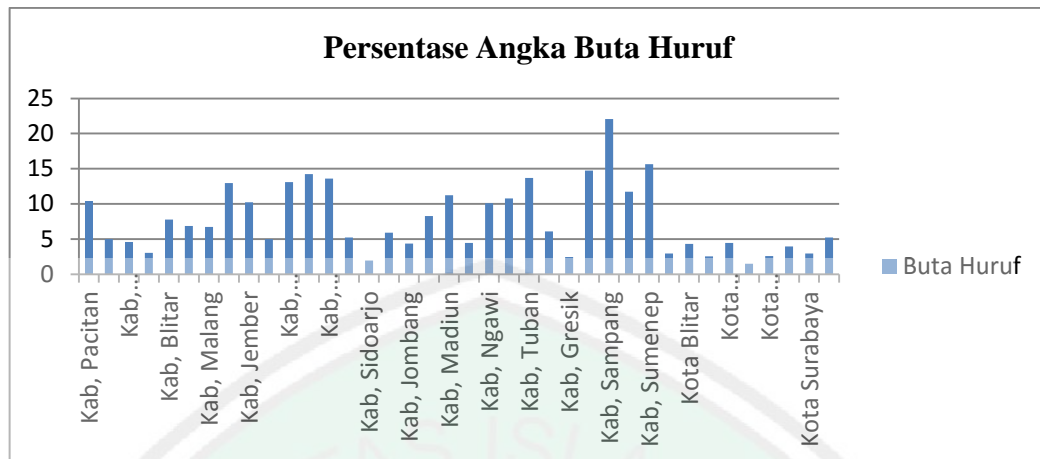
Berdasarkan Gambar 4.8 dapat diketahui bahwa angka partisipasi murni untuk umur 16-18 di Provinsi Jawa Timur pada tahun 2015 yang tertinggi terjadi pada Kota Blitar mencapai 83,83% dan disusul oleh Kota Mojokerto mencapai 82,79% yang merupakan angka partisipasi murni untuk umur 16-18 tertinggi

kedua, sedangkan angka partisipasi murni untuk umur 16-18 terendah terjadi pada Kabupaten Jember mencapai 38,87%, selanjutnya angka partisipasi murni untuk umur 16-18 terendah kedua yaitu Kabupaten Sampang mencapai 42,01%. Rata-rata angka partisipasi murni untuk umur 16-18 untuk seluruh kabupaten/kota di Provinsi Jawa Timur yaitu 63,74%.



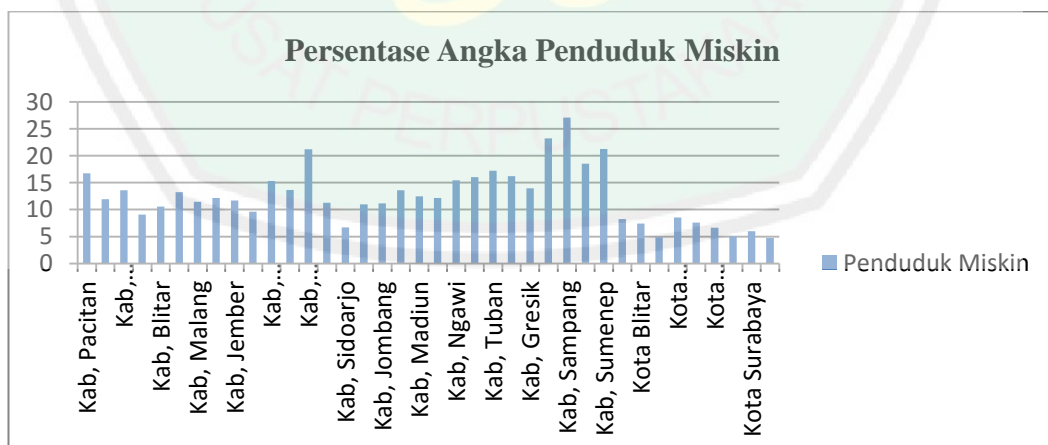
Gambar 4.9 Angka Penduduk yang Tidak Bersekolah

Berdasarkan Gambar 4.9 dapat diketahui bahwa penduduk yang tidak bersekolah di Provinsi Jawa Timur tahun 2015 yang tertinggi terjadi pada Kabupaten Sampang dengan populasi mencapai 26.27% dan penduduk yang tidak bersekolah yang tertinggi kedua terjadi pada Kabupaten Sumenep mencapai 23,14%. Sedangkan penduduk yang tidak bersekolah di Provinsi Jawa Timur terendah terjadi pada Kota Madiun mencapai 1,11% dan penduduk yang tidak bersekolah terendah kedua terjadi pada Kabupaten Sidoarjo mencapai 1,22%. Rata-rata penduduk yang tidak bersekolah di Provinsi Jawa Timur sebesar 8,00%.



Gambar 4.10 Angka Penduduk Buta Huruf

Berdasarkan Gambar 4.10 dapat diketahui bahwa penduduk yang buta huruf di Provinsi Jawa Timur tahun 2015 yang tertinggi terjadi pada Kabupaten Sampang dengan populasi mencapai 22,07% dan penduduk yang buta huruf yang tertinggi kedua terjadi pada Kabupaten Sumenep mencapai 15,63%. Sedangkan penduduk yang buta huruf di Provinsi Jawa Timur terendah terjadi pada Kota Pasuruan mencapai 1,48% dan penduduk yang buta huruf terendah kedua terjadi pada Kabupaten Sidoarjo mencapai 1,95 %. Rata-rata penduduk yang buta huruf di Provinsi Jawa Timur sebesar 7,69%.



Gambar 4.11 Angka Penduduk Miskin

Berdasarkan Gambar 4.11 dapat diketahui bahwa penduduk miskin di Provinsi Jawa Timur tahun 2015 yang tertinggi terjadi pada pada Kabupaten

Sampang dengan populasi mencapai 27,08% dan penduduk miskin yang tertinggi kedua terjadi pada Kabupaten Sumenep mencapai 21,22 %. Sedangkan penduduk miskin di Provinsi Jawa Timur terendah terjadi pada Kota Batu mencapai 4,77% dan penduduk miskin terendah kedua terjadi pada Kota Malang mencapai 4,87%. Rata-rata penduduk miskin di Provinsi Jawa Timur sebesar 12,53%.

4.2.2 Analisis Faktor

Sebelum melakukan pengelompokan, maka perlu dilakukan reduksi variabel, karena terdapat varians yang tinggi. Hal ini dilakukan untuk mengatasi adanya korelasi antar variabel yang dapat mengganggu proses pembentukan kelompok. Adapun tahapan-tahapannya sebagai berikut:

1. Pemilihan Variabel

Tahap pertama sebelum dilakukannya analisis *cluster* pada kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator IPM, maka perlu dilakukan analisis faktor terlebih dahulu. Analisis faktor bertujuan untuk mereduksi dimensi data dengan cara menyatakan variabel asal sebagai kombinasi linier sejumlah faktor.

a. *Measure of Sampling Adequacy* (MSA)

MSA digunakan untuk mengetahui apakah variabel memadai untuk dianalisis lebih lanjut. Dapat diketahui melalui nilai *anti-image correlation matriks* sebagai berikut:

Tabel 4.1 Nilai *Measure of Sampling Adequacy* (MSA)

Variabel	<i>Measure of Sampling Adequacy</i> (MSA)	Keterangan
X1	0,516	Memadai
X2	0,587	Memadai
X3	0,861	Memadai
X4	0,638	Memadai
X5	0,172	Tidak Memadai
X6	0,813	Memadai
X7	0,789	Memadai
X8	0,792	Memadai
X9	0,779	Memadai
X10	0,823	Memadai
X11	0,893	Memadai

Tabel 4.1 dapat diketahui nilai MSA dari setiap variabel. Jika nilai MSA lebih dari 0,5 maka variabel tersebut sudah memadai untuk dianalisis lebih lanjut dan jika nilai MSA kurang dari 0,5 maka variabel tersebut harus dikeluarkan satu persatu dari analisis, dan tidak digunakan lagi dalam analisis selanjutnya. Dari hasil perhitungan, diperoleh bahwa variabel X_5 memiliki nilai MSA kurang dari 0,5 sehingga variabel keluhan kesehatan tidak dimasukkan ke dalam analisis selanjutnya. Sehingga dapat disimpulkan bahwa terdapat 10 variabel yang dapat digunakan untuk analisis faktor.

b. Uji KMO dan Uji *Bartlett*

Penggunaan analisis faktor ini digunakan untuk mereduksi variabel karena adanya korelasi antar variabel. Untuk mengetahui adanya korelasi atau tidak maka dilakukan pengujian independensi menggunakan uji *Bartlett* dengan menggunakan hipotesis sebagai berikut:

Hipotesis:

$H_0: P = I$ (Tidak ada korelasi antar variabel)

$H_1: P \neq I$ (Ada korelasi antar variabel)

Tabel 4.2 Hasil KMO dan Uji Bartlett

<i>Kaiser-Meyer-Olkin Measure of Sampling Adequacy.</i>	0,764
<i>Approx. Chi-Square</i>	364,567
<i>Bartlett's Test of Sphericit df</i>	55
<i>Sig.</i>	0,000

Tabel 4.2 menunjukkan bahwa terdapat korelasi antar variabel karena nilai $p\text{-value} < \alpha = 5\%$. Selanjutnya, melakukan pengujian untuk mengetahui apakah jumlah data cukup untuk analisis faktor menggunakan uji KMO yang mana hasil KMO sebesar 0,764 sehingga nilai KMO $> 0,5$ dan artinya analisis faktor ini cukup bagus dilakukan karena nilai KMO lebih dari 0,5, sehingga dapat disimpulkan bahwa analisis faktor tepat digunakan untuk menyederhanakan kumpulan 10 variabel tersebut.

2. Pembentukan Faktor

Tahap berikutnya setelah diperoleh variabel yang memenuhi asumsi analisis faktor, maka langkah selanjutnya adalah membentuk faktor untuk menentukan struktur yang mendasari hubungan antar variabel awal. Adapun metode yang digunakan untuk pembentukan faktor ini adalah metode *Principal Component* dengan analisis matriks varians dan kovarians dan dilakukan dengan rotasi *varimax* untuk memudahkan dalam interpretasi. Adapun langkahnya adalah sebagai berikut:

a. Penentuan Jumlah Faktor

Jumlah faktor yang akan dibentuk ditentukan dengan melakukan kombinasi beberapa kriteria sehingga diperoleh jumlah faktor yang paling sesuai dengan data. Kriteria pertama yang digunakan untuk penentuan jumlah faktor adalah nilai eigen. Faktor yang memiliki nilai eigen lebih dari 1 akan dipertahankan dan yang nilai eigennya kurang dari 1 tidak akan diikutsertakan

dalam model. Dari Tabel 4.3 didapatkan nilai eigen yang lebih dari 1 berada pada 9 faktor, yaitu faktor 1, faktor 2, faktor 3, faktor 4, faktor 5, faktor 6, faktor 7, faktor 8, dan faktor 9. Dari kriteria ini dapat diperoleh jumlah faktor yang digunakan adalah sebanyak 9 faktor.

Tabel 4.3 Total Variance Explained

Comp	Initial Eigen values			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cum %	Total	% of Variance	Cum %
1	571,429	61,004	61,004	5,002	45,474	45,474
2	106,500	12,184	73,188	1,341	12,187	57,660
3	33,669	8,733	81,921			
4	28,200	4,825	86,746			
5	24,092	3,522	90,268			
6	14,268	2,988	93,256			
7	8,829	1,983	95,239			
8	3,443	1,310	96,548			
9	2,063	0,830	97,378			
10	0,701	0,706	98,085			

Kriteria yang kedua adalah berdasarkan nilai persentase varians total yang dapat dijelaskan dengan banyaknya faktor yang akan dibentuk. Dari Tabel 4.3 diperoleh jika variabel-variabel diringkas menjadi beberapa faktor, maka nilai total varians dapat dijelaskan sebagai berikut:

Jika 10 variabel diekstraksi menjadi 1 faktor, maka diperoleh varians total sebesar

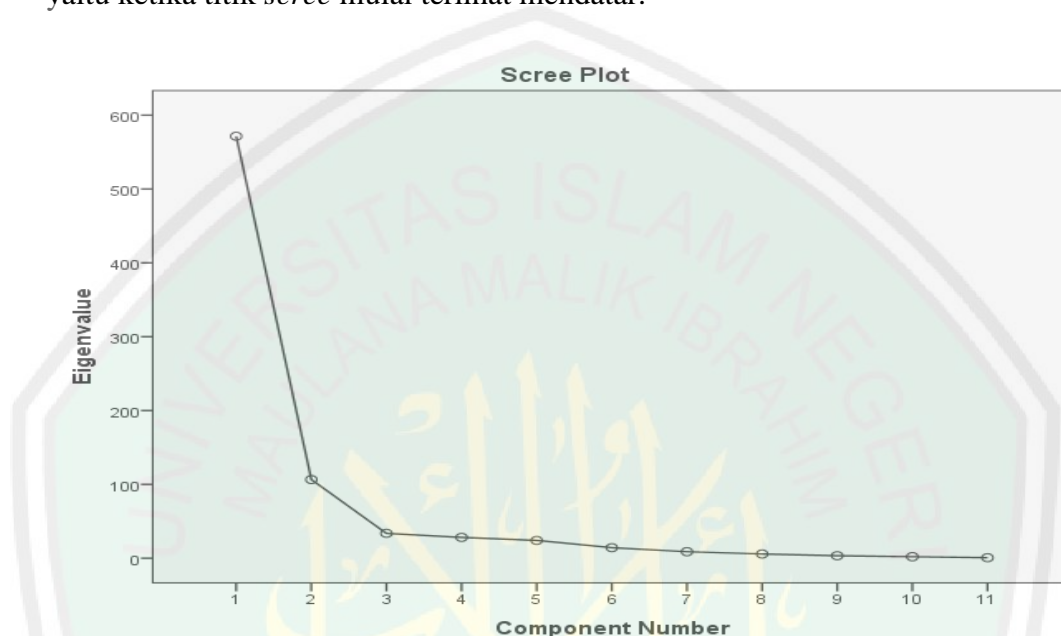
$$\frac{5,002}{10} \times 100\% = 50,02\%.$$

Jika 10 variabel diekstraksi menjadi 2 faktor, maka diperoleh varians total sebesar

$$\frac{1,341}{10} \times 100\% = 13,41\%.$$

Tahap selanjutnya dengan mengekstraksi 10 variabel menjadi 2 faktor, maka dihasilkan varians total kumulatif yang cukup besar yakni 63,43%, artinya dari 2 faktor yang telah terbentuk tersebut sudah mewakili 10 variabel indikator IPM yang menjelaskan tentang kondisi pembangunan manusia di Provinsi Jawa Timur tahun 2015 sebesar 63,43% .

Kriteria ketiga adalah penentuan berdasarkan *scree plot*. *Scree plot* merupakan *plot* nilai eigen terhadap jumlah faktor yang diekstraksi. Titik pada tempat di mana *scree* mulai terjadi menunjukkan banyaknya faktor yang tepat, yaitu ketika titik *scree* mulai terlihat mendatar.



Gambar 4.12 *Scree Plot*

Berdasarkan Gambar 4.12 diketahui bahwa titik *scree* mulai mendatar pada ekstraksi variabel awal pada titik ke-3. Sehingga dari kombinasi ketiga kriteria, yaitu nilai eigen, persentase varians total, dan *scree plot* dapat disimpulkan bahwa ekstraksi faktor yang paling tepat adalah 2 faktor.

b. Kumunalitas

Kumunalitas pada dasarnya adalah jumlah varians dari suatu variabel yang dapat dijelaskan oleh faktor yang ada. Lebih lengkapnya pada Tabel 4.4.

Tabel 4.4 Tabel Kumunalitas

Variabel	Initial	Extraction
X1	1,000	0,083
X2	1,000	0,195
X3	1,000	0,944
X4	1,000	0,140
X6	1,000	0,917
X7	1,000	0,960
X8	1,000	0,951
X9	1,000	0,813
X10	1,000	0,714
X11	1,000	0,601

Berdasarkan Tabel 4.4 dapat diketahui bahwa variabel Angka Partisipasi Kasar (APK) usia 16-18 tahun merupakan variabel dengan varians terbesar yaitu 0,960 dan yang memiliki varians terendah adalah variabel tingkat partisipasi kerja yaitu 0,083.

c. Komponen Matriks

Tahap selanjutnya setelah diperoleh 2 faktor adalah jumlah faktor paling optimum, maka Tabel 4.5 menunjukkan distribusi ke-10 variabel pada 2 faktor yang terbentuk sedangkan angka yang ada pada Tabel 4.5 adalah *loading factor*, yang menunjukkan besar kasar korelasi antara variabel dengan kedua faktor. Proses penentuan variabel mana yang akan dimasukkan ke faktor 1, ataupun faktor 2 dilakukan dengan melakukan perbandingan besar korelasi. Lebih lengkapnya dapat dilihat pada Tabel 4.5.

Tabel 4.5 Tabel Komponen Matriks

Variabel	Loading Factor	
	1	2
X1	-0,256	0,133
X2	0,316	-0,310
X3	0,763	0,601
X4	0,369	0,057
X5	0,059	0,143
X6	0,952	0,104
X8	0,926	0,320
X9	0,962	0,156
X10	-0,719	0,544
X11	-0,705	0,466
X12	-0,640	0,438

d. Rotasi Matriks

Proses rotasi dilakukan untuk mendapatkan faktor dengan *loading factor* yang cukup jelas untuk interpretasi. Komponen matriks rotasi merupakan matriks korelasi yang memperlihatkan distribusi variabel yang lebih jelas dan nyata jika dibandingkan dengan komponen matriks. Dari Tabel 4.6 dapat diketahui bahwa nilai antar *loading factor* telah cukup dibedakan dan siap dilakukan interpretasi. Seluruh variabel mempunyai *loading factor* yang tinggi pada salah satu faktor dan mempunyai *loading factor* yang cukup kecil untuk faktor lainnya.

Tabel 4.6 Tabel Rotasi Matrik

Variabel	<i>Loading Factor</i>	
	1	2
X1	0,265*	-0,116
X2	-0,439	0,050*
X3	0,947*	-0,215
X4	-0,188	0,323*
X6	-0,519	0,805*
X7	-0,334	0,921*
X8	-0,485	0,846*
X9	0,876*	-0,216
X10	0,806*	-0,255
X11	0,743*	-0,222

keterangan:

Bold dan *) memiliki nilai *loading* (korelasi) cukup tinggi

e. Interpretasi Jumlah Faktor

Penentuan signifikansi nilai *loading factor* untuk menentukan pengelompokan variabel ke dalam faktor yang sesuai. Nilai *loading factor* akan dianggap signifikan jika nilainya lebih dari 0.

Tabel 4.7 Tabel Signifikansi *Loading Factor*

Variabel	<i>Loading Factor</i>			
	1	Keterangan	2	Keterangan
X1	0,265	Signifikan	-0,116	-
X2	-0,439	-	0,050	Signifikan
X3	0,947	Signifikan	-0,215	-
X4	-0,188	-	0,323	Signifikan
X6	-0,519	-	0,805	Signifikan
X7	-0,334	-	0,921	Signifikan
X8	-0,485	-	0,846	Signifikan
X9	0,876	Signifikan	-0,216	-
X10	0,806	Signifikan	-0,255	-
X11	0,743	Signifikan	-0,222	-

Berdasarkan Tabel 4.7 diperoleh nilai *loading factor* dari setiap variabel, dari hasil tersebut variabel-variabel yang memiliki nilai signifikan atau yang nilai *loading factornya* lebih dari 0 akan dikelompokkan ke dalam faktor sebagai berikut:

Tabel 4.8 Tabel Hasil Pengelompokan Variabel ke dalam Faktor

Faktor	Variabel
1	X1, X3, X9, X10, X11
2	X2, X4, X6, X7, X8

Setelah terbentuk 2 faktor yang masing-masing beranggotakan 5 variabel, maka dilakukan penamaan faktor berdasarkan karakteristik yang sesuai dengan anggotanya.

Faktor 1. Berdasarkan Tabel 4.8 dapat diketahui bahwa anggota dari faktor 1 terdiri dari variabel X1 (persentase tingkat partisipasi angkatan kerja), X3 (persentase angka kematian bayi), X9 (persentase angka penduduk yang tidak bersekolah), X10 (persentase angka buta huruf), dan X11 (persentase angka penduduk miskin). Secara umum anggota dari faktor 1 menggambarkan tentang kemiskinan, sehingga faktor 1 dinamakan dengan faktor kemiskinan.

Faktor 2. Berdasarkan Tabel 4.8 dapat diketahui bahwa anggota dari faktor 2 terdiri dari variabel X2 (persentase tingkat pengangguran terbuka), X4 (persentase angka harapan hidup), X6 (persentase APS 16-18), X7 (persentase APK 16-18),

dan X8 (persentase APM 16-18). Secara umum anggota dari faktor 2 menggambarkan tentang pendidikan, sehingga faktor 2 dinamakan faktor pendidikan.

Tabel 4.9 Tabel Penamaan Faktor

Faktor	Variabel	Nama
1	X1, X3, X9, X10, X11	Kemiskinan
2	X2, X4, X6, X7, X8	Pendidikan

4.2.3 Analisis *Cluster* Hirarki

Analisis *cluster* hirarki merupakan suatu metode yang tidak membutuhkan suatu asumsi yang dibuat dalam jumlah kelompok atau struktur kelompok. Analisis *cluster* hirarki sendiri merupakan suatu metode pengelompokan yang jumlah kelompok yang akan dibuat belum diketahui.

Analisis ini menggunakan 3 metode *Cluster* Hirarki yaitu metode *Complete Linkage*, *Average Linkage*, dan *Ward's Linkage*. Selain itu, jumlah pengelompokan dalam metode Hirarki ini menggunakan perkiraan jumlah kelompok 2 sampai 5 dan penentuan banyaknya kelompok yang paling optimum dari masing-masing metode akan dilihat berdasarkan nilai *Pseudo F* yang terbesar dan hasilnya dapat dilihat pada Tabel 4.10 berikut ini:

Tabel 4.10 Nilai *Pseudo F* Masing-masing Metode *Cluster*

Banyaknya Kelompok	<i>Pseudo F</i>		
	<i>Complete</i>	<i>Average</i>	<i>Ward's</i>
2	0,94334449	0,970661502	1,455197947
3	1,074106593	1,448009903	1,074106593
4	1,348922854*	1,666869092	1,348922854
5	1,094693611	2,039498108*	1,49987804*

keterangan:

Bold dan *) memiliki nilai *Pseudo F* tinggi

Tabel 4.10 di atas menunjukkan nilai *Pseudo F* pada masing-masing metode. Metode *Complete Linkage* memiliki nilai optimum 4 kelompok dan untuk *Average Linkage* serta *Ward's Linkage* masing-masing memiliki nilai optimum 5 kelompok. Sedangkan hasil dendogram masing-masing metode pada

pengelompokan kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator IPM tahun 2015 sebagai berikut:

a. Metode *Complete Linkage*

Metode ini digunakan untuk melihat jarak atau kesamaan antara pasangan-pasangan objek dan kelompok yang dibentuk dari kesatuan individu dengan menggabungkan tetangga terjauhnya. Banyaknya kelompok pada metode *Complete Linkage* ini berdasarkan Tabel 4.10 dilihat dari nilai *Pseudo F* yang memiliki jumlah kelompok optimum sebanyak 4 kelompok. Untuk kelompok 1 terdapat 14 kabupaten, kelompok 2 terdapat 1 kabupaten, kelompok 3 terdapat 18 kabupaten dan kota, dan kelompok 4 terdapat 5 kabupaten. Berikut merupakan rincian anggota kabupaten/kota dari setiap kelompok.

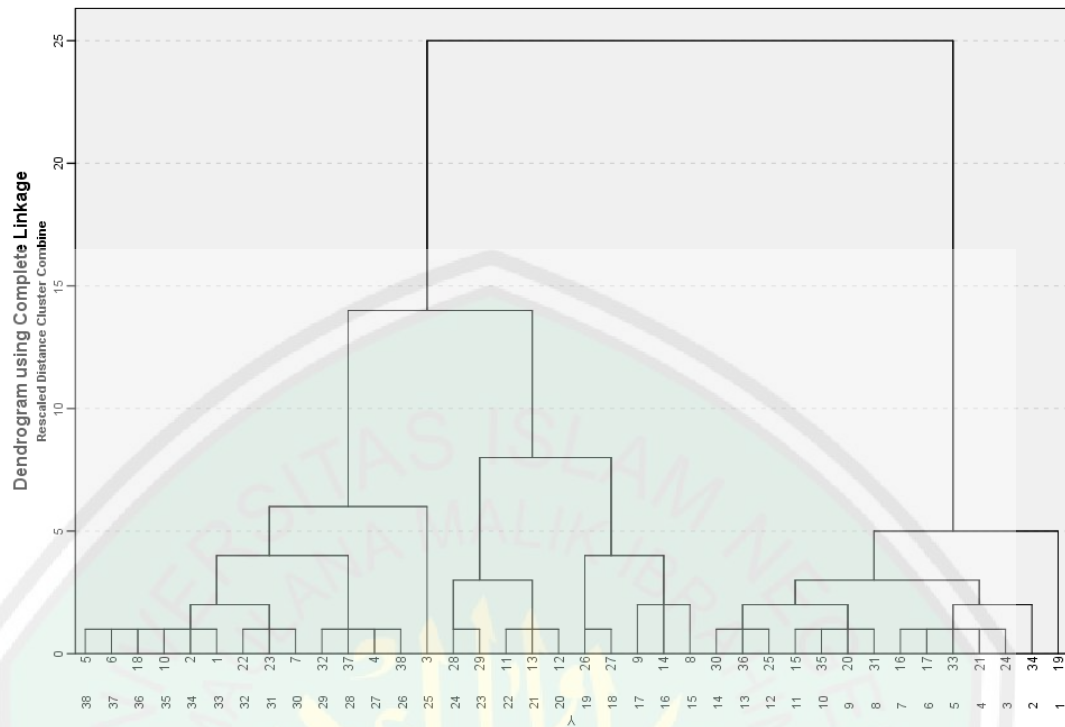
Kelompok 1: Kabupaten Pacitan, Ponorogo, Trenggalek, Blitar, Kediri, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Nganjuk, Bojonegoro, Tuban, Pamekasan, dan Sumenep.

Kelompok 2: Kabupaten Trenggalek.

Kelompok 3: Kabupaten Tulungagung, Sidoarjo, Mojokerto, Jombang, Madiun, Magetan, Ngawi, Lamongan, Gresik, Kota Kediri, Blitar, Malang, Pasuruan, Probolinggo, Mojokerto, Madiun, Surabaya, dan Batu.

Kelompok 4: Kabupaten Lumajang, Jember, Pasuruan, Bangkalan, dan Sampang.

Untuk lebih jelasnya dapat dilihat dari hasil dendogram untuk data indikator IPM provinsi Jawa Timur berikut:



Gambar 4.13 Dendrogram *Complete Linkage*

b. *Average Linkage*

Average linkage merupakan salah satu metode *Cluster* Hirarki yang didasarkan pada rata-rata jarak jauh dari seluruh objek pada suatu *cluster* dengan seluruh objek pada *cluster* lain. Hasil pada Gambar 4.14 di bawah ini menunjukkan bahwa hasil pengelompokan dengan menggunakan metode *Average Linkage* diperoleh 5 kelompok dari hasil *cluster*, dengan kelompok 1 terdapat 14 kabupaten, kelompok 2 terdapat 1 kabupaten, kelompok 3 terdapat 17 kabupaten dan kota, kelompok 4 terdapat 5 kabupaten dan kelompok 5 terdapat 1 kabupaten. Berikut merupakan rincian anggota kabupaten/kota dari setiap kelompok:

Kelompok 1: Kabupaten Pacitan, Ponorogo, Blitar, Kediri, Malang, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Nganjuk, Bojonegoro, Tuban, Pamekasan, dan Sumenep .

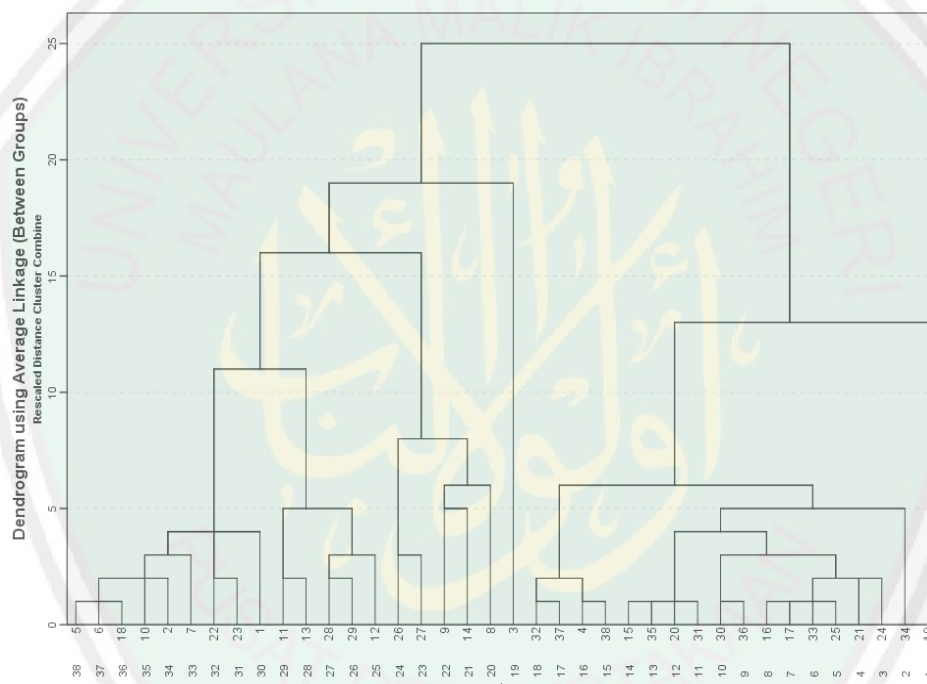
Kelompok 2: Kabupaten Trenggalek.

Kelompok 3: Kabupaten Ponorogo, Sidoarjo, Mojokerto, Jombang, Magetan, Ngawi, Lamongan, Gresik, Kota Kediri, Blitar, Malang, Pasuruan, Probolinggo, Mojokerto, Madiun, Surabaya, dan Batu.

Kelompok 4: Kabupaten Kediri, Malang, Pasuruan, Bangkalan, dan Sampang.

Kelompok 5: Kabupaten Madiun

Untuk lebih jelasnya dapat dilihat dari hasil dendrogram untuk data indikator IPM 2015 sebagai berikut:



Gambar 4.14 Dendrogram Average Linkage

c. Metode *Ward's Linkage*

Metode ini menghitung jarak antara dua kelompok yang terbentuk dengan meminimumkan peningkatan kriteria *Error Sum of Square* (ESS). Jumlah kelompok optimum yang terbentuk sebanyak 5 kelompok, dengan kelompok 1 terdapat 14 kabupaten, kelompok 2 terdapat 1 kabupaten, kelompok 3 terdapat 5 kabupaten, kelompok 4 terdapat 5 kabupaten dan kelompok 5 terdapat 14

kabupaten dan kota. Berikut merupakan rincian anggota kabupaten/kota dari setiap kelompok:

Kelompok 1: Kabupaten Pacitan, Ponorogo, Tulungagung, Blitar, Kediri, Malang, Banyuwangi, Nganjuk, Bojonegoro, Tuban, Kota Malang, Surabaya, dan Batu.

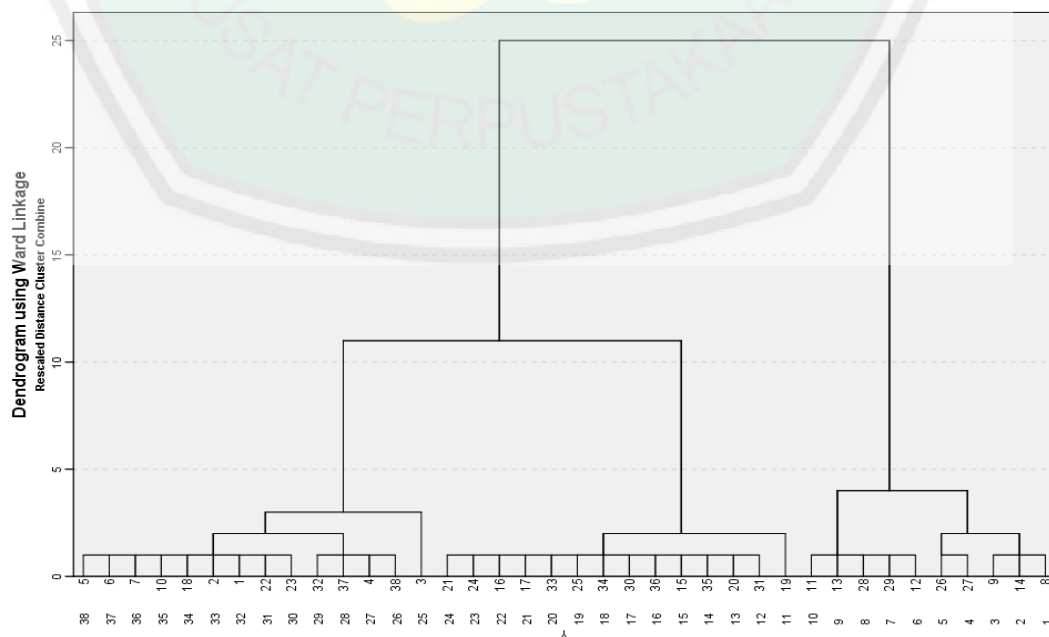
Kelompok 2: Kabupaten Trenggalek.

Kelompok 3: Kabupaten Lumajang, Jember, Pasuruan, Bangkalan, dan Sampang.

Kelompok 4: Kabupaten Jember, Banyuwangi, Bondowoso, Pamekasan, dan Sumenep.

Kelompok 5: Kabupaten Sidoarjo, Mojokerto, Jombang, Madiun, Magetan, Ngawi, Lamongan, Gresik, Kota Kediri, Blitar, Probolinggo, Pasuruan, Mojokerto, dan Madiun.

Untuk lebih jelasnya dapat dilihat dari hasil dendrogram untuk data indikator IPM 2015 sebagai berikut:



Gambar 4.15 Dendrogram *Ward's Linkage*

d. Pemilihan Metode Terbaik

Menurut Mingoti dan Lima (2006), kebaikan hasil pengelompokan dapat dilihat dari penyebaran internal dalam kelompok atau disebut dengan *internal cluster dispersion rate (icdrate)*. Semakin kecil nilai *icdratenya*, maka semakin baik hasil pengelompokan dari masing-masing metode. Hasil pemilihan metode terbaik dapat dilihat pada Tabel 4.11.

Tabel 4.11 Nilai *icdrate* untuk Setiap Pengelompokan

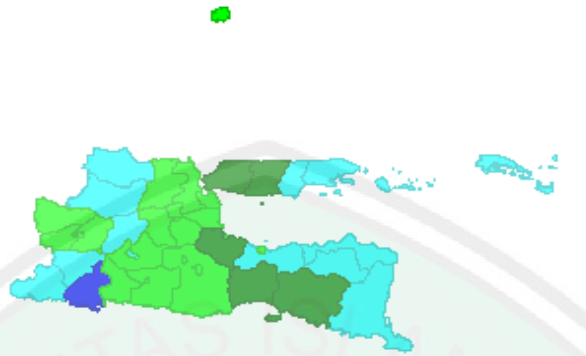
Metode	Banyak Kelompok Optimum	Icdrate
<i>Complete Linkage</i>	4	0,893636997
<i>Average Linkage</i>	5	0,801788378*
<i>Ward's Linkage</i>	5	0,84616443

Keterangan:

Bold dan *) memiliki nilai *icdrate* yang paling kecil

Berdasarkan Tabel 4.11 dapat diketahui bahwa pada hasil banyaknya kelompok yang optimum untuk masing-masing metode *Cluster Hirarki* memberikan keragaman yang cukup berbeda dengan metode lainnya. Hal ini disebabkan karena hasil kelompok yang optimum memiliki jumlah yang berbeda yaitu *complete linkage* dengan 4 kelompok dan *average linkage* serta *ward's linkage* yang masing-masing 5 kelompok. Pada Tabel 4.11 juga menunjukkan penilaian *icdrate* dari ketiga metode yang digunakan dalam pembentukan anggota pengelompokan. Metode yang memiliki nilai *icdrate* paling kecil yaitu metode *Average Linkage* dalam pengelompokan kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator IPM tahun 2015 dengan pembagian optimum kelompoknya sebanyak 5 kelompok. Untuk lebih jelasnya dapat dilihat pada masing-masing pemetaannya sebagai berikut:

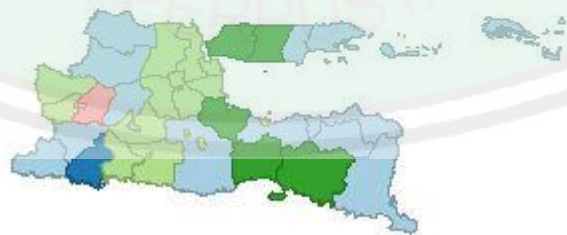
a. *Complete Linkage*



Gambar 4.16 Pengelompokan Kabupaten/Kota *Complete Linkage*

Gambar 4.16 menunjukkan bahwa warna biru muda masuk ke dalam *cluster* 1 sedangkan warna biru tua masuk ke dalam *cluster* 2, untuk warna hijau muda masuk ke dalam *cluster* ke 3 dan *cluster* 4 ditandai dengan warna hijau tua. *Cluster* 3 memiliki anggota paling banyak di antara ke 3 *cluster* lainnya dengan 18 kabupaten dan kota.

b. *Average Linkage*



Gambar 4.17 Pengelompokan Kabupaten/Kota *Average Linkage*

Gambar 4.17 menunjukkan bahwa terdapat 5 kelompok yang dikategorikan. *Cluster* 1 ditandai dengan warna biru muda, untuk *cluster* 2 ditandai dengan warna biru tua, *cluster* 3 ditandai dengan warna hijau muda,

cluster 4 ditandai dengan warna hijau tua, dan *cluster* 5 ditandai dengan warna merah muda. *Cluster* 3 merupakan *cluster* yang memiliki anggota yang paling banyak di antara 4 *cluster* lainnya.

c. *Ward's Linkage*



Gambar 4.18 Pengelompokan Kabupaten/Kota *Ward's Linkage*

Gambar 4.18 menunjukkan bahwa terdapat 5 kelompok yang dikategorikan. *Cluster* 1 ditandai dengan warna biru muda, untuk *cluster* 2 ditandai dengan warna biru tua, *cluster* 3 ditandai dengan warna hijau muda, *cluster* 4 ditandai dengan warna hijau tua, dan *cluster* 5 ditandai dengan warna merah muda. *Cluster* 1 dan *cluster* 5 merupakan *cluster* yang memiliki anggota yang paling banyak di antara 3 *cluster* lainnya.

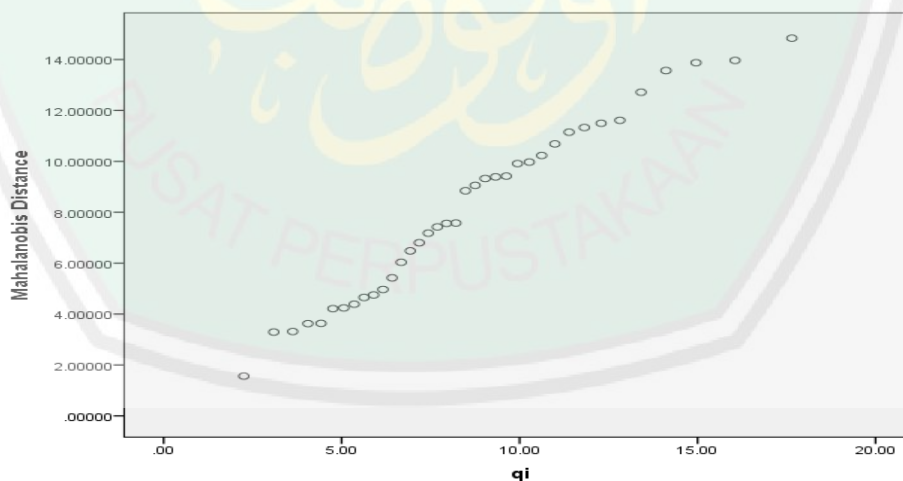
e. Evaluasi Hasil Pengelompokan

Berdasarkan analisis pengelompokan di atas maka diperoleh metode yang terbaik yaitu metode *Average Linkage*, selanjutnya dilakukan evaluasi pada hasil pengelompokan tersebut. Evaluasi ini digunakan untuk menguji apakah terdapat perbedaan yang cukup signifikan minimum antar dua kelompok yang terbentuk.

Evaluasi hasil pengelompokan akan dilakukan dengan menggunakan MANOVA, yang mana asumsi yang harus dipenuhi sebelum dilakukan analisis

MANOVA adalah pemeriksaan variabel dependen berdistribusi multivariat normal dan matriks varians kovarians bersifat homogen. Dalam penelitian ini, yang akan dianalisis MANOVA adalah 2 faktor yang telah terbentuk dari variabel asli yaitu indikator IPM. Selanjutnya, akan dilakukan pemeriksaan distribusi normal.

Pemeriksaan distribusi normal pada data penelitian ini dapat diketahui dengan menggunakan *plot* khi-kuadrat. *Plot* ini didasarkan perhitungan nilai jarak kuadrat (d_j^2) pada setiap pengamatan dan *plot* khi-kuadrat berupa garis lurus yang berada di tengah grafik. Berdasarkan hasil perhitungan normal multivariat diperoleh nilai jarak kuadrat sebesar 0,547 yang artinya bahwa data mengikuti sebaran distribusi normal karena lebih dari 50%. Selain itu juga terlihat secara visual dari Gambar 4.19 di bawah ini yang menunjukkan pola data mengikuti sebaran distribusi normal multivariat.



Gambar 4.19 Pemeriksaan Normal Multivariat

Asumsi homogenitas matriks varians kovarians yang harus dipenuhi sebelum dilakukan uji perbedaan dengan menggunakan MANOVA. Pengujian homogenitas matriks varians kovarians diperoleh dengan menggunakan *Box's M*

dan hipotesis yang digunakan dalam pengujian matriks varians kovarians adalah sebagai berikut:

$$H_0 : \Sigma_1 = \Sigma_2$$

$$H_1 : \Sigma_1 \neq \Sigma_2$$

Keputusan gagal tolak H_0 jika $p\text{-value} > \alpha$ ($\alpha = 5\%$)

Tabel 4.12 Uji Homogenitas Matriks Varians Kovarians

Keterangan	Nilai
Box's M	114,163
Uji F	1,152
P-value	0,209

Tabel 4.12 di atas menunjukkan bahwa diperoleh nilai $p\text{-value}$ pada pengujian homogenitas varians kovarians $> \alpha = 5\%$ yang artinya adalah gagal tolak H_0 maka dapat disimpulkan bahwa data tentang indikator IPM sudah homogen dan memenuhi asumsi untuk dilakukan pengujian perbedaan dengan MANOVA. Diketahui bahwa kedua asumsi sebelum melakukan uji perbedaan kelompok dengan MANOVA, baik distribusi normal multivariat dan homogenitas matriks varians kovarians telah terpenuhi secara statistik. Adapun hipotesis yang diberikan dalam pengujian perbedaan kelompok dengan MANOVA adalah sebagai berikut:

Hipotesis:

$$H_0: \tau_1 = \tau_2 = \dots = \tau_g = 0$$

$$H_1: \text{minimal ada satu } \tau_g \neq 0$$

Berdasarkan hasil pengujian menggunakan MANOVA, diperoleh nilai *Wilk's Lamda* sebesar 0,000 sehingga nilai $p\text{-value}$ menunjukkan $< \alpha = 5\%$. Hal ini dapat dikatakan bahwa terdapat perbedaan rata-rata antar kelompok kabupaten/kota di Jawa Timur yang terbentuk.

4.3 Karakteristik Setiap Kelompok

Pengujian MANOVA yang telah dilakukan menunjukkan terdapat perbedaan antara 5 kelompok yang terbentuk oleh metode *Average Linkage*. Setelah itu akan diketahui masing-masing kelompok berdasarkan variabel-variabel indikator IPM. Perbedaan untuk setiap kelompok dapat dilihat pada Tabel 4.13 berikut:

Tabel 4.13 Perbedaan Karakteristik Setiap Kelompok

Kelompok 1	Kelompok 2	Kelompok 3	Kelompok 4	Kelompok 5
-Persentase Penduduk Miskin	-Persentase Tingkat Partisipasi Kerja	-Persentase Angka Harapan Hidup -Persentase Tingkat Pengangguran Terbuka	-Persentase Angka Kematian Bayi -Persentase Angka Buta Huruf	-Persentase Angka Partisipasi Sekolah (16-18 tahun) -Persentase Angka Partisipasi Kasar (16-18 tahun) -Persentase Angka Partisipasi Murni (16-18 tahun) -Persentase Penduduk yang Tidak Pernah Bersekolah

4.4 Kajian Agama Tentang Cluster

Pokok utama dari analisis *cluster* yaitu pengelompokan suatu objek yang memiliki kemiripan nilai dan jarak antar objek lainnya. Jika kedua objek tersebut memiliki kemiripan nilai dan jarak, maka dua objek tersebut akan berada dalam satu *cluster*.

Banyak ayat al-Quran yang menjelaskan ciri-ciri seseorang pada suatu kelompok tertentu. Salah satunya yaitu Allah Swt. telah menurunkan kitab suci al-Quran kepada kelompok hamba-hamba yang terpilih. Seperti pada firman Allah Swt. dalam al-Quran surat al-Fathir ayat 32:

ثُمَّ أَوْرَثْنَا الْكِتَابَ الَّذِينَ اصْطَفَيْنَا مِنْ عِبَادِنَا فَمِنْهُمْ ظَالِمٌ لِنَفْسِهِ ۖ وَمِنْهُمْ مُقْتَصِدٌ وَمِنْهُمْ سَابِقٌ بِالْخَيْرَاتِ بإِذْنِ اللَّهِ ۗ ذَٰلِكَ هُوَ الْفَضْلُ الْكَبِيرُ ﴿٣٢﴾

“Kemudian kitab itu Kami wariskan kepada orang-orang yang Kami pilih di antara hamba-hamba Kami, lalu di antara mereka ada yang menganiaya diri mereka sendiri dan di antara mereka ada yang pertengahan dan di antara mereka ada (pula) yang lebih dahulu berbuat kebaikan dengan izin Allah Swt.. Yang demikian itu adalah karunia yang amat besar”(QS. Al-Fathir/35:32).

Pada ayat ini, diterangkan bahwa Allah Swt. menurunkan al-Quran sebagai satu pedoman hidup untuk manusia baik kehidupan di dunia maupun di akhirat. Untuk mencapai kedua hal tersebut maka manusia dituntut untuk mampu memahami, membaca, dan mengamalkan apa yang terkandung dalam kitab Allah Swt. tersebut. Namun banyak di antara kelompok manusia yang menyia-nyiakan hal tersebut dan tidak mau menggunakan kitab al-Quran sebagai pedoman hidupnya.

Allah Swt. berfirman, *“Kemudian Kami jadikan orang-orang yang mengamalkan kitab yang besar yang membenarkan kitab-kitab sebelumnya adalah orang-orang yang Kami pilih di antara hamba-hamba Kami”*, mereka adalah ummat nabi Muhammad Saw. Kemudian mereka terbagi menjadi 3 golongan. Kelompok pertama pada penjelasan Ibnu Katsir terdapat pada kalimat *“lalu di antara mereka ada yang menganiaya diri mereka sendiri”*, yang dimaksudkan di sini adalah orang yang melalaikan sebagian dari pekerjaan yang diwajibkan atasnya dan mengerjakan sebagian dari hal-hal yang diharamkan. Kelompok kedua terdapat pada kalimat *“dan di antara mereka ada yang pertengahan”*, yang dimaksudkan di sini adalah orang yang menunaikan hal-hal yang diwajibkan atas dirinya dan meninggalkan hal-hal yang diharamkan, tetapi ada kalanya dia meninggalkan sebagian dari hal-hal yang disunnahkan dan

mengerjakan sebagian dari hal-hal yang dimakruhkan. Kelompok ketiga terdapat pada kalimat “*dan di antara mereka ada (pula) yang lebih dahulu berbuat kebaikan dengan izin Allah Swt.*”, yang dimaksudkan adalah orang yang mengerjakan semua kewajiban dan hal-hal yang disunnahkan, juga meninggalkan semua hal-hal yang diharamkan, yang dimakruhkan, dan sebagian hal yang diperbolehkan (Abdullah, 2007).

Ayat ini menjadi petunjuk bagi manusia untuk mengetahui dirinya masuk dalam kelompok yang mana dan memilih al-Quran sebagai pedoman hidupnya. Jika manusia itu tergolong pada kelompok yang pertama dan yang kedua, mereka dapat bertaubat dan mendekatkan dirinya kembali dengan Allah Swt. dengan cara menjadikan al-Quran sebagai pedoman hidupnya di dunia maupun di akhirat.

BAB V

PENUTUP

5.1 Kesimpulan

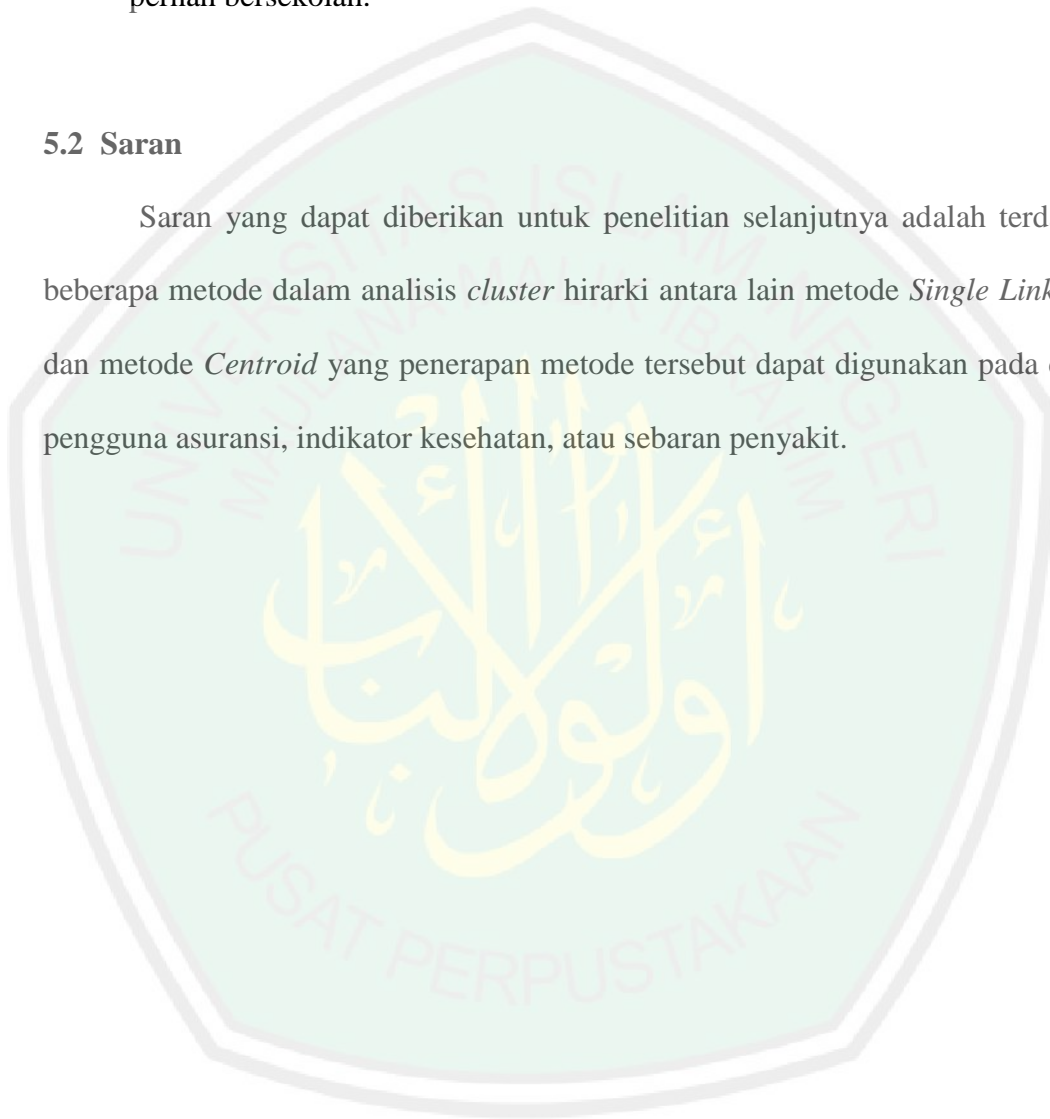
Berdasarkan pembahasan pada penelitian ini, maka dapat diambil kesimpulan sebagai berikut:

1. Pembuktian terhadap asumsi pada model persamaan analisis faktor yaitu:
 - a. Diperoleh $E(F_j) = 0$, untuk $j = 1, 2, 3, \dots, m$, dan $Cov(F) = E(F, F') = I_{m \times m}$.
 - b. Diperoleh $E(\varepsilon_i) = 0$, untuk $i = 1, 2, 3, \dots, m$, dan $Cov(\varepsilon) = E(\varepsilon, \varepsilon') = \psi_{p \times p}$.
 - c. Untuk F dan ε yang saling bebas diperoleh $Cov(\varepsilon, F) = E(\varepsilon, F') = 0_{p \times m}$.
2. Hasil pengelompokan terbaik dibentuk oleh *average linkage* sebanyak 5 kelompok. Hal ini dilihat dari nilai *icdrate* yang menunjukkan bahwa nilai terendah terdapat pada metode *Average Linkage* yakni sebesar 0,801788378 sehingga metode *Average Linkage* merupakan metode yang terbaik di antara metode *Complete Linkage* ataupun *Ward's Linkage*.
3. Variabel-variabel yang mencirikan setiap kelompok sebagai berikut:
 - a. Kelompok 1 dominan terhadap persentase penduduk miskin.
 - b. Kelompok 2 dominan terhadap persentase tingkat partisipasi kerja.
 - c. Kelompok 3 dominan terhadap persentase angka harapan hidup dan persentase tingkat pengangguran terbuka.
 - d. Kelompok 4 dominan terhadap persentase angka kematian bayi dan persentase angka buta huruf.

- e. Kelompok 5 dominan terhadap persentase angka partisipasi sekolah (16-18 tahun), persentase angka partisipasi kasar (16-18 tahun), persentase angka partisipasi murni (16-18 tahun), dan persentase penduduk yang tidak pernah bersekolah.

5.2 Saran

Saran yang dapat diberikan untuk penelitian selanjutnya adalah terdapat beberapa metode dalam analisis *cluster* hirarki antara lain metode *Single Linkage* dan metode *Centroid* yang penerapan metode tersebut dapat digunakan pada data pengguna asuransi, indikator kesehatan, atau sebaran penyakit.



DAFTAR RUJUKAN

- Abdullah. 2007. *Tafsir Ibnu Katsir jilid 4*. Jakarta: Pustaka Imam Asy-Syafi'i.
- Abdullah. 2007. *Tafsir Ibnu Katsir jilid 6*. Jakarta: Pustaka Imam Asy-Syafi'i.
- Amalyah, F. 2012. *Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan indikator Indonesia Sehat 2010*. Skripsi tidak dipublikasikan. Surabaya: ITS Surabaya.
- Astuti, A.D. 2014. *Partial Least Square (PLS) dan Principal Component Regression (PCR) untuk Regresi Linear dengan Multikolinearitas pada Kasus Indeks Pembangunan Manusia di Kabupaten Gunung Kidul*. Skripsi tidak dipublikasikan. Yogyakarta: Universitas Negeri Yogyakarta
- BPS. 2015. *Indeks Pembangunan Manusia 2015*. Surabaya: Badan Pusat Statistik.
- Calinski, T., dan Harabasz, J. 1974. *A Dendrit Method for Cluster Analysis*. Poznan: Taylor & Francis, Inc.
- Gudono. 2011. *Analisis Data Multivariat*. Yogyakarta: BPFE.
- Hair, J.F., Anderson R.E., Tatham, R.L., dan Black, W.C. 1995. *Multivariate Data Analysis with Readings*. New Jersey: Pearson New International Edition.
- Irianto, A. 2004. *Statistik Konsep Dasar & Aplikasinya*. Jakarta: Prenada Media Group.
- Johnson dan Wichern. 1998. *Applied Multivariate Statistical Analysis 4nd Edition*. New Jersey: Prentice-Hall, Inc.
- Johnson dan Wichern. 2007. *Applied Multivariate Statistical Analysis 6nd Edition*. New Jersey: Prentice-Hall International, Inc.
- Johnson, S.C. 1967. *Hierarchical Clustering Schemes*. New Jersey: Prentice-Hall, Inc.
- Milligan, G.W., dan Cooper, M.C. 1985. *An Examination of Procedures for Determining the Number of Clusters in Data Set*. New Jersey: Prentice-Hall, Inc.
- Mingoti, S.A., dan Lima, J.O. 2006. Comparing SOM Neural Network with Fuzzy C-Means, K-Means and Traditional Hierarchical Clustering Algorithms. *European journal of Operational Research*, (Online), 173 (3): 169-175, (<http://www.sciencedirect.com>), diakses 27 September 2016.

- Morrison, D.F. 1990. *Multivariate Statistical Method 3rd Edition*. New York: Mc Graw-Hill Book Company.
- Narimawati, U. 2008. *Metodologi Penelitian Kualitatif dan Kuantitatif, Teori dan Aplikasi*. Bandung: Agung Media.
- Orpin, A.R., dan Kostylev, V.E. 2006. *Toward's a Statistically Valid Method of Textural Sea Floor Characterization of Benthic Habitats*. New Zealand: *Marine Geology*.
- Suliyanto. 2005. *Analisis Data Dalam Aplikasi Pemasaran*. Bogor: Ghalia Indonesia Anggota IKAPI.
- Supranto, J. 2003. *Pengantar Matriks* . Jakarta: PT. Rineka Cipta.
- Supranto, J. 2004. *Analisis Multivariat Arti dan Interpretasi*. Jakarta: Rineka Cipta.
- Sutanto, H. 2009. *Cluster Analysis*. Yogyakarta: Jurnal Pendidikan Matematika: FMIPA Yogyakarta: UNY.
- UNDP. 1995. *1995 Human Development Report*. New York: UNDP.
- Yamin, S.R. 2011. *Regresi dan Korelasi dalam Genggaman Anda*. Jakarta: Salemba Empat.

LAMPIRAN - LAMPIRAN

Lampiran 1. Variabel Penelitian

Kabupaten/Kota	X₁	X₂	X₃	X₄	X₅	X₆
Pacitan	80,28	1,08	21,66	70,5	30,69	69,79
Ponorogo	72,31	3,66	24,86	71,88	30,4	77,22
Trenggalek	74	4,2	20,23	42,51	27,75	62,33
Tulungagung	72,57	2,42	20,87	72,88	28,76	74,05
Blitar	69,12	3,08	22,68	72,5	33,1	70,14
Kediri	67,28	4,91	25,79	72,04	27,28	71,8
Malang	66,04	4,83	28,63	71,78	29,18	64,44
Lumajang	65,09	2,83	36,03	69,07	21,12	50,61
Jember	63,74	4,64	54,72	67,8	27,08	52,52
Banyuwangi	69,15	7,17	30,82	69,93	38,15	67,92
Bondowoso	70,55	3,72	50,93	65,43	39,63	61,13
Situbondo	66,47	4,15	53,06	68,08	31,56	68,28
Probolinggo	69,92	1,47	61,48	65,75	34,13	62,25
Pasuruan	70,91	4,43	48,61	69,83	36,45	62,77
Sidoarjo	67,94	3,88	22,78	73,43	22,31	84,72
Mojokerto	67,8	3,81	22,82	71,76	33,97	78,51
Jombang	64,82	4,39	26,8	71,37	38,91	81,11
Nganjuk	67,17	3,93	29,88	70,87	32,54	71,17
Madiun	68,73	3,38	30,2	69,76	29,4	89,22
Magetan	69,14	4,28	21,77	71,91	28,39	85,76
Ngawi	67,29	5,61	24,81	71,33	38,12	75,17
Bojonegoro	65,49	3,21	37,87	70,11	22,51	62,2
Tuban	64	3,63	31,59	70,25	29,49	68,49
Lamongan	66,64	4,3	32,82	71,47	30,01	81,55
Gresik	63,66	5,06	22,13	72,2	22,58	84,26
Bangkalan	69,44	5,68	53,12	69,62	23,84	58,73
Sampang	76,85	2,22	49,5	67,48	37,61	55,34
Pamekasan	75,08	2,14	47,48	66,56	29,3	69,81
Sumenep	74,1	1,01	46,77	70,02	29,12	75,13
Kota Kediri	67,77	7,66	22,08	73,52	35,27	88,43
Kota Blitar	66,46	5,71	17,99	72,7	30,78	92,17
Kota Malang	63,66	7,22	21,28	72,3	37,24	78,91
Kota Probolinggo	66,94	5,16	21,52	68,52	35,78	82,78
Kota Pasuruan	67,78	6,09	37,12	70,54	37,82	83,12
Kota Mojokerto	68,07	4,42	20,92	72,39	30,92	85,78
Kota Madiun	63,54	6,93	22,11	72,41	28,59	87,77
Kota Surabaya	66,56	5,82	21,91	73,85	34,49	75,19
Kota Batu	70,38	2,43	27,08	72,06	30,69	76,67

Lanjutan Variabel Penelitian

Kabupaten/Kota	X₇	X₈	X₉	X₁₀	X₁₁
Pacitan	79,62	59,26	6,45	10,43	16,73
Ponorogo	81,91	66,31	10,18	4,98	11,92
Trenggalek	73,68	49,85	4,5	4,6	13,56
Tulungagung	89,1	65,51	3,27	3,03	9,07
Blitar	72,56	62,07	4,97	7,79	10,57
Kediri	71,17	60,8	4,84	6,84	13,23
Malang	68,29	54	5,36	6,73	11,48
Lumajang	65,08	47,78	11,84	12,97	12,14
Jember	55,96	38,87	13,25	10,23	11,68
Banyuwangi	81,51	58,63	8,64	5,01	9,61
Bondowoso	79,95	49,56	13,81	13,09	15,29
Situbondo	80,75	62,75	14,96	14,23	13,65
Probolinggo	75,11	53,75	13,68	13,59	21,21
Pasuruan	61,55	50,25	8,01	5,2	11,26
Sidoarjo	97,66	79,32	1,22	1,95	6,72
Mojokerto	99,13	70,37	3,25	5,91	10,99
Jombang	97,1	73,09	3,96	4,34	11,17
Nganjuk	78,47	66,18	5,54	8,29	13,6
Madiun	122,09	82,29	9,18	11,21	12,45
Magetan	106,04	78,22	4,36	4,43	12,19
Ngawi	95,86	72,57	11,1	10,11	15,45
Bojonegoro	81,2	60,01	8,64	10,76	16,02
Tuban	77,18	56,26	13,26	13,7	17,23
Lamongan	93,62	70,61	8,8	6,08	16,18
Gresik	93,45	77,16	2,85	2,46	13,94
Bangkalan	58,3	42,91	19,36	14,73	23,23
Sampang	54,63	42,01	26,27	22,07	27,08
Pamekasan	81,12	57,45	17,88	11,73	18,53
Sumenep	87,94	63,28	23,14	15,63	21,22
Kota Kediri	86,35	73,87	1,47	2,96	8,23
Kota Blitar	103,66	83,83	1,67	4,33	7,42
Kota Malang	83,15	59,23	2,01	2,55	4,87
Kota Probolinggo	95,47	71,2	5,98	4,43	8,55
Kota Pasuruan	105,26	75,7	4,19	1,48	7,6
Kota Mojokerto	103,11	82,79	1,38	2,56	6,65
Kota Madiun	90,83	78,31	1,11	3,95	5,02
Kota Surabaya	88,65	60,69	1,79	2,96	6
Kota Batu	95,4	65,7	1,89	5,21	4,77

Keterangan:

- X_1 : Tingkat Partisipasi Angkatan Kerja
- X_2 : Tingkat Pengangguran Terbuka
- X_3 : Angka Kematian Bayi
- X_4 : Angka Harapan Hidup
- X_5 : Keluhan Kesehatan
- X_6 : Angka Partisipasi Sekolah (APS) usia 16-18 tahun
- X_7 : Angka Partisipasi Kasar (APK) usia 16-18 tahun
- X_8 : Angka Partisipasi Murni (APM) usia 16-18 tahun
- X_9 : Tidak Bersekolah
- X_{10} : Buta Huruf
- X_{11} : Penduduk Miskin



Lampiran 2. Anti-image Correlationmatriks

Anti-image Matrices

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	
Anti-image Covariance	X1	,364	,221	,094	,132	-,197	-,040	,033	,021	-,051	,056	-,035
	X2	,221	,301	,046	,096	-,187	-,048	,064	,013	-,042	,058	,021
	X3	,094	,046	,279	-,008	-,081	,019	-,008	,005	-,081	,030	,005
	X4	,132	,096	-,008	,620	-,094	-,041	,104	-,037	-,023	-,025	,079
	X5	-,197	-,187	-,081	-,094	,692	-,015	-,074	,038	,024	-,044	,037
	X6	-,040	-,048	,019	-,041	-,015	,066	-,018	-,040	-,012	,020	-,006
	X7	,033	,064	-,008	,104	-,074	-,018	,125	-,043	-,014	-,003	,039
	X8	,021	,013	,005	-,037	,038	-,040	-,043	,056	,014	-,007	-,017
	X9	-,051	-,042	-,081	-,023	,024	-,012	-,014	,014	,091	-,063	-,048
	X10	,056	,058	,030	-,025	-,044	,020	-,003	-,007	-,063	,115	-,041
	X11	-,035	,021	,005	,079	,037	-,006	,039	-,017	-,048	-,041	,193
Anti-image Correlation	X1	,516 ^a	,668	,296	,278	-,392	-,255	,154	,151	-,278	,273	-,132
	X2	,668	,587 ^a	,159	,223	-,410	-,337	,329	,101	-,253	,311	,087
	X3	,296	,159	,861 ^a	-,019	-,184	,140	-,046	,039	-,510	,169	,021
	X4	,278	,223	-,019	,638 ^a	-,143	-,203	,374	-,198	-,098	-,094	,230
	X5	-,392	-,410	-,184	-,143	,172 ^a	-,072	-,253	,193	,096	-,155	,102
	X6	-,255	-,337	,140	-,203	-,072	,813 ^a	-,200	-,664	-,160	,234	-,051
	X7	,154	,329	-,046	,374	-,253	-,200	,789 ^a	-,516	-,130	-,028	,251
	X8	,151	,101	,039	-,198	,193	-,664	-,516	,792 ^a	,203	-,089	-,160
	X9	-,278	-,253	-,510	-,098	,096	-,160	-,130	,203	,779 ^a	-,618	-,361
	X10	,273	,311	,169	-,094	-,155	,234	-,028	-,089	-,618	,823 ^a	-,275
	X11	-,132	,087	,021	,230	,102	-,051	,251	-,160	-,361	-,275	,893 ^a

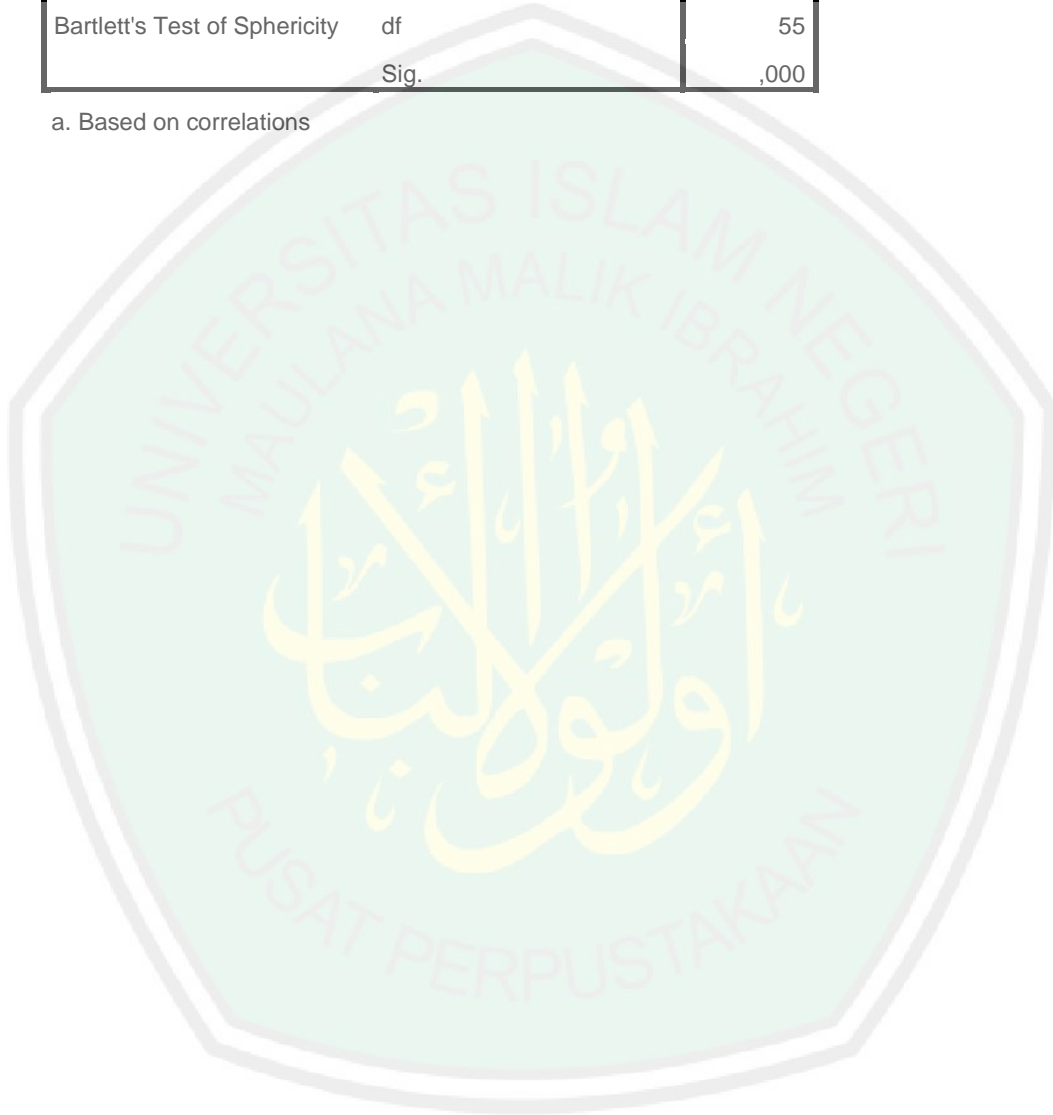
a. Measures of Sampling Adequacy(MSA)

Lampiran 3. Nilai KMO dan Bartlet

KMO and Bartlett's Test^a

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,764
Approx. Chi-Square		364,567
Bartlett's Test of Sphericity	df	55
	Sig.	,000

a. Based on correlations



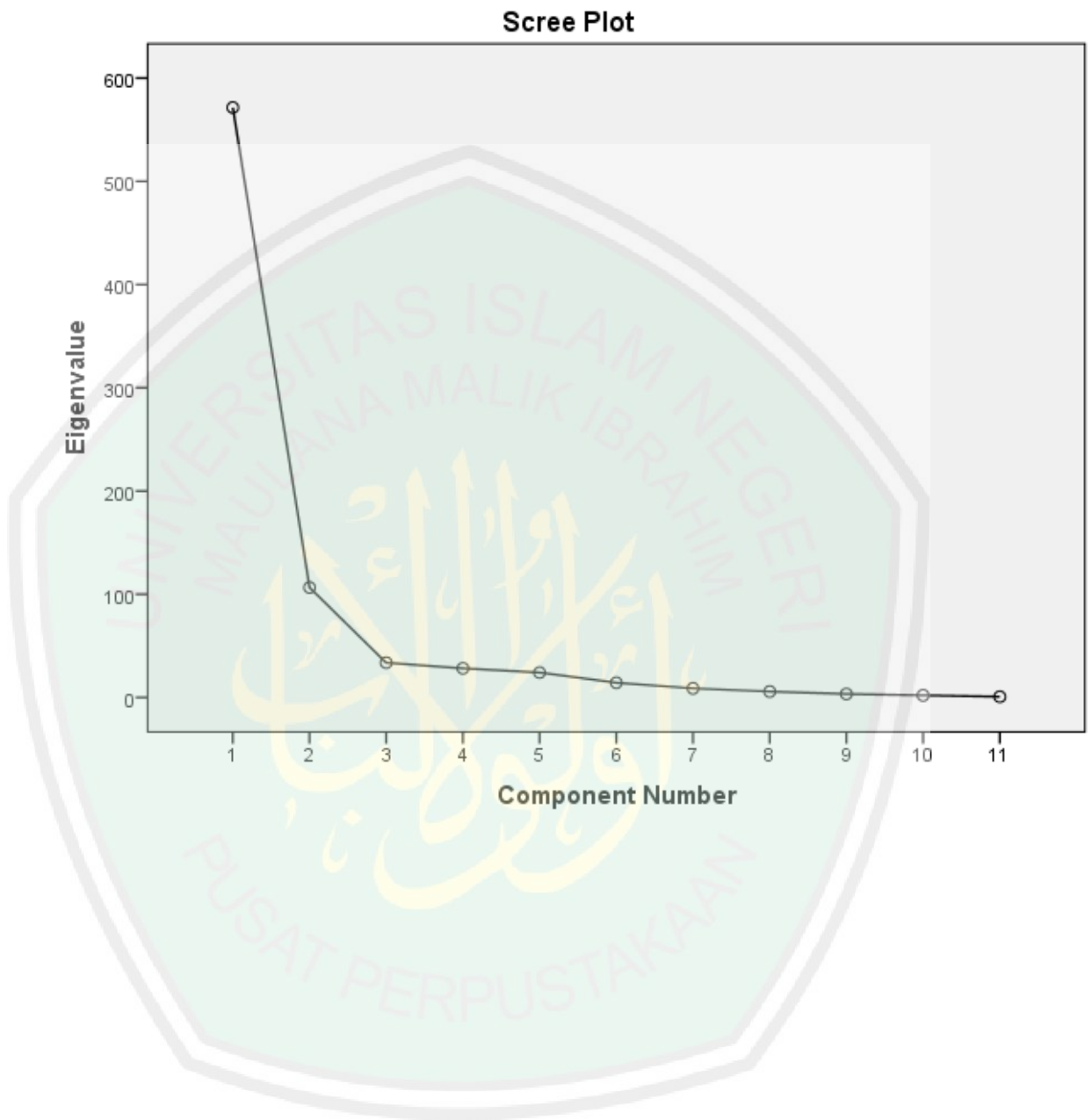
Lampiran 4. Total Variance Explained

Total Variance Explained										
	Component	Initial Eigenvalues ^a			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
		Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
Raw	1	571,429	71,526	71,526	571,429	71,526	71,526	290,841	36,405	36,405
	2	106,500	13,331	84,857	106,500	13,331	84,857	387,088	48,452	84,857
	3	33,669	4,214	89,072						
	4	28,200	3,530	92,601						
	5	24,092	3,016	95,617						
	6	14,268	1,786	97,403						
	7	8,829	1,105	98,508						
	8	5,711	,715	99,223						
	9	3,443	,431	99,654						
	10	2,063	,258	99,912						
	11	,701	,088	100,000						
Rescaled	1	571,429	71,526	71,526	5,002	45,474	45,474	3,785	34,408	34,408
	2	106,500	13,331	84,857	1,341	12,187	57,660	2,558	23,252	57,660
	3	33,669	4,214	89,072						
	4	28,200	3,530	92,601						
	5	24,092	3,016	95,617						
	6	14,268	1,786	97,403						
	7	8,829	1,105	98,508						
	8	5,711	,715	99,223						
	9	3,443	,431	99,654						
	10	2,063	,258	99,912						
	11	,701	,088	100,000						

Extraction Method: Principal Component Analysis.

a. When analyzing a covariance matrix, the initial eigenvalues are the same across the raw and rescaled solution.

Lampiran 5. Scree Plot



Lampiran 6. Tabel Kumunalitas

Communalities

	Raw		Rescaled	
	Initial	Extraction	Initial	Extraction
X1	14,626	1,219	1,000	,083
X2	2,722	,532	1,000	,195
X3	152,241	143,657	1,000	,944
X4	25,292	3,529	1,000	,140
X5	24,766	,597	1,000	,024
X6	116,869	107,188	1,000	,917
X7	228,815	219,720	1,000	,960
X8	142,720	135,675	1,000	,951
X9	40,258	32,750	1,000	,813
X10	23,462	16,759	1,000	,714
X11	27,136	16,303	1,000	,601

Extraction Method: Principal Component Analysis.

Lampiran 7. Tabel Komponen Matriks

Component Matrix^a

	Raw		Rescaled	
	Component		Component	
	1	2	1	2
X1	-,980	,508	-,256	,133
X2	,521	-,511	,316	-,310
X3	-9,417	7,414	-,763	,601
X4	1,857	,287	,369	,057
X5	,295	,714	,059	,143
X6	10,292	1,124	,952	,104
X7	14,009	4,844	,926	,320
X8	11,497	1,867	,962	,156
X9	-4,564	3,453	-,719	,544
X10	-3,416	2,255	-,705	,466
X11	-3,333	2,279	-,640	,438

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

RIWAYAT HIDUP



Achmad Sirojuddin dilahirkan di Makassar pada tanggal 27 September 1994, anak kedua dari empat bersaudara, dari pasangan Bapak H. Zulkifli dan Ibu Nur Miati. Pendidikan dasar ditempuh di kampung halamannya di SDN Mandaran Rejo 1 yang ditamatkan pada tahun 2006.

Pada tahun yang sama dia melanjutkan pendidikan menengah pertama di SMP Negeri 3 Pasuruan. Pada tahun 2009 dia menamatkan pendidikannya, kemudian melanjutkan pendidikan menengah atas SMA Negeri 3 Pasuruan dan menamatkan pendidikan tersebut pada tahun 2012. Pendidikan berikutnya dia tempuh di Universitas Islam Negeri Maulana Malik Ibrahim Malang melalui jalur SMPTN Tulis dengan mengambil Jurusan Matematika Fakultas Sains dan Teknologi.



KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI
MAULANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Gajayana No. 50 Dinoyo Malang Telp./Fax.(0341)558933

BUKTI KONSULTASI SKRIPSI

Nama : Achmad Sirojuddin
NIM : 12610033
Fakultas/Jurusan : Sains dan Teknologi/Matematika
Judul Skripsi : Analisis Cluster pada Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Indeks Pembangunan Manusia
Pembimbing I : Dr. Sri Harini, M.Si
Pembimbing II : Mohammad Jamhuri, M.Si

No.	Tanggal	Hal	Tanda Tangan
1.	29 Agustus 2016	Konsultasi Bab I & Bab II	1.
2.	14 September 2016	Konsultasi Bab III	2.
3.	28 September 2016	Konsultasi Agama Bab I	3.
4.	30 September 2016	Konsultasi Agama Bab II	4.
5.	5 Desember 2016	Revisi Bab I, Bab II, Bab III & Bab IV	5.
6.	6 Desember 2016	Konsultasi Bab IV	6.
7.	7 Desember 2016	Konsultasi Agama Bab IV	7.
8.	6 Desember 2016	Konsultasi Seluruh Bab	8.
9.	6 Desember 2016	Revisi Agama Bab IV	9.
10.	13 Desember 2016	ACC Keseluruhan	10.
11.	14 Desember 2016	ACC Agama Keseluruhan	11.

Malang, 14 Desember 2016

Mengetahui,

Ketua Jurusan Matematika

Dr. Abdussakir, M.Pd

NIP. 19751006 200312 1 001