



TUGAS AKHIR - SS141501

**PENGELOMPOKAN PROVINSI
DI INDONESIA BERDASARKAN INDIKATOR
KESEJAHTERAAN RAKYAT MENGGUNAKAN METODE
C-MEANS DAN *FUZZY C-MEANS CLUSTERING***

ANNISA SAJIDAH
NRP 1311 100 135

Dosen Pembimbing
Dr. Dra. Ismaini Zain, M.Si

Program Studi S1 Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember
Surabaya 2016



TUGAS AKHIR - SS141501

**PENGELOMPOKAN PROVINSI
DI INDONESIA BERDASARKAN INDIKATOR
KESEJAHTERAAN RAKYAT MENGGUNAKAN METODE
C-MEANS DAN *FUZZY C-MEANS CLUSTERING***

**ANNISA SAJIDAH
NRP 1311 100 135**

**Dosen Pembimbing
Dr. Dra. Ismaini Zain, M.Si**

**Program Studi S1 Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember
Surabaya 2016**



FINAL PROJECT - SS141501

**GROUPING PROVINCES IN INDONESIA
BASED ON WELFARE INDICATORS
USING *C*-MEANS AND *FUZZY C*-MEANS**

**ANNISA SAJIDAH
NRP 1311 100 135**

**Supervisor
Dr. Dra. Ismaini Zain, M.Si**

**Undergraduate Program of Statistics
Faculty of Mathematics and Natural Sciences
Institut Teknologi Sepuluh Nopember
Surabaya 2016**

LEMBAR PENGESAHAN

**PENGELOMPOKAN PROVINSI DI INDONESIA
BERDASARKAN INDIKATOR KESEJAHTERAAN
DENGAN MENGGUNAKAN C-MEANS DAN FUZZY C-
MEANS**

TUGAS AKHIR

**Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains pada
Program Studi S-1 Jurusan Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember**

Oleh :

**ANNISA SAJIDAH
NRP. 1311 100 135**

Disetujui oleh Pembimbing Tugas Akhir :

**Dr. Dra. Ismaini Zain, M.Si.
NIP. 19600525 198803 2 001**

()

Mengetahui
Ketua Jurusan Statistika FMIPA-ITS


**Dr. Suhartono
NIP. 19710929 199512 1 001**

SURABAYA, JANUARI 2016

Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Kesejahteraan Rakyat dengan Metode *c-means* dan *fuzzy c-means clustering*

Nama Mahasiswa : Annisa Sajidah
NRP : 1311 100 135
Jurusan : Statistika FMIPA-ITS
Pembimbing : Dr. Dra. Ismaini Zain, M.Si

ABSTRAK

*Permasalahan sosial yang terjadi di Indonesia memberikan dampak yang cukup serius kepada kesejahteraan rakyat di suatu daerah. Kesejahteraan rakyat berkaitan dengan tujuan pembangunan nasional namun permasalahan yang ada adalah tidak mempunya pemerintah dalam memilah permasalahan di setiap daerah sehingga pembangunan yang dilakukan dapat merata. Penelitian ini bertujuan memberikan informasi kepada pihak terkait dalam pembangunan nasional agar memberikan penanganan yang sesuai kepada provinsi yang dianggap kurang sejahtera berdasarkan tingkat homogenitasnya dengan metode pengelompokan *c-means* dan *fuzzy c-means*. Pengelompokan tersebut memberikan hasil bahwa dengan *c-means* memberikan hasil lebih baik dibandingkan dengan *fuzzy c-means clustering* berdasarkan nilai *icd rate* terendahnya dimana *cluster optimum* yang terbentuk ada sebanyak 2 *cluster* yang dilihat dari nilai *pseudo f-statistics* tertinggi. Hasil *c-means* memberikan hasil karakteristik provinsi papua membutuhkan perhatian lebih dibandingkan provinsi lainnya dari segala dimensi. Sedangkan dengan *fuzzy c-means* kedua kelompok mampu meberikan karakteristik tersendiri sesuai dengan komponen utama yang membentuknya.*

Kata Kunci – *C-Means Clustering, Fuzzy C-Means Clusering, Icd rate, Indikator Kesejahteraan Rakyat. Pseudo F-Statistics.*

Grouping Provinces in Indonesia Based On Welfare Indicators Using *c-means* and *fuzzy c-means clustering*

Name of Student : Annisa Sajidah
NRP : 1311 100 135
Department : Statistics
Advisor : Dr. Dra. Ismaini Zain, M.Si

ABSTRACT

*Social problems that occur in Indonesia provide a fairly serious impact to the welfare of people in that area. Welfare indicators concerned with national development goals but the problem is the inability of the government for sorting out the problems in each area so that development can be done evenly. This study aims to provide information to stakeholders in national development in order to provide appropriate treatment to provinces that are considered less prosperous based on the level of homogeneity by the method of grouping *c-means* and *fuzzy c-means*. Grouping by *fuzzy c-means* providing better results than with *c-means* clustering based on the value *icd-rate* cause it is the lowest rate in which the optimum cluster is formed there were two clusters are seen from the pseudo highest *f-statistics*. The result by *c-means* provide the characteristics of Papua province needs more attention compared the other provinces from all dimensions. While the *fuzzy c-means* two groups that gave the distinct characteristics in accordance with the principal components of dimensions that make it up.*

Kata Kunci – C-Means Clustering, Fuzzy C-Means Clustering, Icd rate, Pseudo F-Statistics, Welfare Indicators.

“Halaman ini sengaja dikosongkan”

“Halaman ini sengaja dikosongkan”

KATA PENGANTAR

Puji syukur kehadiran Allah SWT, atas limpahan rahmat yang tidak pernah berhenti sehingga terselesaikannya Tugas Akhir yang berjudul

“Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Kesejahteraan Rakyat Dengan Metode *C-Means* dan *Fuzzy C-Means Clustering*”

dengan baik. Penyusunan Tugas Akhir ini tidak terlepas dari bantuan dan dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini ucapan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT yang telah melimpahkan segala rahmat dan kasih sayangNya sehingga terselesaikan Tugas Akhir ini dengan baik.
2. Bapak Dr. Suhartono M.Sc selaku Ketua Jurusan Statistika yang telah memberikan banyak fasilitas untuk kelancaran penyelesaian Tugas Akhir ini.
3. Ibu Dr. Dra Ismaini Zain M.Si selaku dosen pembimbing yang telah sabar dalam memberikan bimbingan dan saran atas bantuan dan semua informasi yang diberikan. Meluangkan segala kesempatan dan waktu yang ada untuk memberikan bimbingan terhadap Tugas Akhir ini.
4. Ibu Dr. Vita Ratnasari, S.Si, M.Si dan Dra. Ibu Madu Ratna, M.Si selaku dosen penguji atas saran dan kritiknya yang sangat membangun.
5. Ibu Santi Putri selaku dosen wali atas dukungan, semangat yang diberikan, nasehat setiap semester, dan saran-saran yang berguna pada saat perwalian.
6. Bapak, Ibu, dan keluarga dirumah atas segala doa, kasih sayang, dan dukungan. Keluarga terbaik di dunia yang telah dianugerahkan Allah SWT sert penyemangat disaat semangat mulai surut dan disaat menemui kendala.

7. Indi Yasinta dan Ummu Habibah saudara selama masa kuliah. Terimakasih banyak buat dukungan yang tak berhenti mengalir dan bersedia menyediakan tempat untuk berkeluh kesah saat kesulitan datang saat menyelesaikan Tugas Akhir ini.
8. Fiscy Aprilia, Aprilia Tri Wahyu, Anggraeni Rahma Dewi, Fatkhiyatur Rizki dan Ilafi Andalita telah menjadi teman senasib sepenanggungan dan memberikan dukungan banyak hal dan akhirnya Tugas Akhir ini terselesaikan dengan baik.
9. Teman-teman seperjuangan statistika great 2011 sigma 22 yang selalu bersama dalam dekapan hangatnya sebuah keluarga.
10. Teman-teman AIESEC Surabaya yang selalu membantu selama berada disini dan bersama untuk *giving impact to the society based on Global Issues*.
11. Keluarga BEM FMIPA ITS 12-13 dan BEM FMIPA ITS 13-14 yang selalu menyemangati sehingga Tugas Akhir ini terselesaikan dengan baik.
12. Keluarga Karate ITS yang selalu menjadi keluarga dan menjadi teman seperjuangan mengharumkan nama ITS dengan menjadi atlet pertandingan untuk ITS.
13. Keluarga Gudang Ilmu Indonesia yang menjadi keluarga dalam berbagi ilmu kepada adik-adik yang membutuhkan.
14. Serta semua pihak yang telah mendukung dan tidak dapat penulis sebutkan satu persatu.

Harapannya Tugas Akhir ini dapat bermanfaat bagi kita semua serta saran dan kritik yang bersifat membangun guna perbaikan di masa mendatang.

Surabaya, Januari 2016

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	
ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	3
1.3 Tujuan	3
1.4 Manfaat Penelitian	3
1.5 Batasan Masalah	4
BAB II TINJAUAN PUSTAKA	
2.1 Distribusi Normal Multivariat	5
2.2 Uji KMO.....	5
2.3 Uji Barlett.....	6
2.4 Analisis <i>Cluster</i>	6
2.5 <i>C-means Clustering</i>	7
2.6 <i>Fuzzy C-means Clustering</i>	9
2.7 Fungsi Keanggotaan.....	11
2.7.1 Representasi Linear	11
2.7.2 Representasi Segitiga.....	13
2.8 <i>Principal Component Anaysis (PCA)</i>	14
2.9 <i>Pseudo f-statistics</i>	15
2.10 <i>Internal Cluster Dispertion Rate (icd-rate)</i>	16
2.11 <i>One-way ANOVA</i>	17
BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data.....	19
3.2 Variabel Penelitian	19
3.3 Struktur Data	20

3.4 Definisi Operasional.....	20
3.5 Langkah-langkah Penelitian	22
3.6 Diagram Alir.....	24
BAB IV ANALISIS DAN PEMBAHASAN	
4.1 Deskripsi Indikator Kesejahteraan Rakyat	25
4.2 Pengelompokan Provinsi di Indonesia dengan <i>C-Means</i> dan <i>Fuzzy C-Means Clustering</i>	32
4.3 Perbandingan Hasil Pengelompokan dengan <i>C-Means</i> dan <i>Fuzzy C-Means Clustering</i>	41
BAB V KESIMPULAN DAN SARAN	
5.1 Kesimpulan	43
5.2 Saran	44
DAFTAR PUSTAKA	45
LAMPIRAN	47

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Representasi Linear Naik.....	12
Gambar 2.2 Representasi Linear Turun.....	13
Gambar 2.3 Representasi Kurva Segitiga.....	14
Gambar 3.1 Diagram Alir Penelitian.....	24
Gambar 4.1 Indeks Gini Provinsi di Indonesia.....	28
Gambar 4.2 Persentase Penduduk Miskin di Indonesia.....	29
Gambar 4.3 Tingkat Pengangguran Terbuka.....	30
Gambar 4.4 Plot Distribusi Normal Multivariat.....	31
Gambar 4.5 <i>Scree plot</i>	33
Gambar 4.6 Peta Pengelompokan dengan <i>C-means clustering</i> ..	36
Gambar 4.7 Peta Pengelompokan dengan <i>Fuzzy C-means clustering</i>	40

“Halaman ini sengaja dikosongkan”

DAFTAR TABEL

	Halaman
Tabel 2.1 Uji <i>One-way</i> ANOVA	18
Tabel 3.1 Variabel Indikator Kesehatan.....	27
Tabel 3.2 Variabel Indikator Pendidikan	19
Tabel 3.3 Variabel Indikator Ekonomi.....	20
Tabel 3.4 Struktur Data.....	20
Tabel 4.1 Statistika Deskriptif Berdasarkan Indikator Kesehatan.....	25
Tabel 4.2 Statistika Deskriptif Berdasarkan Indikator Pendidikan	29
Tabel 4.3 Uji Barlett dan Uji KMO Indikator Kesejahteraan Rakyat Indonesia	32
Tabel 4.4 <i>Eigenvalues dan percentage of variance</i>	33
Tabel 4.5 Nilai Pseudo f-statistics Metode <i>C-means</i>	35
Tabel 4.6 Anggota Cluster Dengan <i>C-means</i>	35
Tabel 4.7 Nilai Pseudo f-statistics Metode <i>Fuzzy c-means</i>	38
Tabel 4.8 Pengelompokan Optimum Dengan <i>Fuzzy C-Means Clustering</i> Fungsi Keanggotaan Linear Naik	40
Tabel 4.9 Pengelompokan Optimum Dengan <i>Fuzzy C-Means Clustering</i> Fungsi Keanggotaan Linear Turun dan Segitiga	40
Tabel 4.10 <i>Icd Rate C-Means dan Fuzzy C-means Clustering</i>	42

“Halaman ini sengaja dikosongkan”

DAFTAR LAMPIRAN

	Halaman
Lampiran A	Data Penelitian..... 47
Lampiran B	Data Penelitian Standarisasi..... 48
Lampiran C	Principal Component Analysis..... 49
Lampiran D	Principal Component Analysis & Derajat Keanggotaan Linear Naik 50
Lampiran E	Principal Component Analysis & Derajat Keanggotaan Linear Turun 51
Lampiran F	Principal Component Analysis & Derajat Keanggotaan Segitiga 52
Lampiran G	Matriks U Linear Naik 53
Lampiran H	Matriks U Linear Turun 54
Lampiran I	Matriks U Segitiga..... 55
Lampiran J	Macro Minitab Normal Multivariat 56
Lampiran K	Syntax Matab Fuzzy C-Means 57
Lampiran L	Syntax Matlab Pseudo F Statistics 59
Lampiran M	Hasil Cluster Dengan C-Means..... 60
Lampiran N	Hasil Cluster Dengan Fuzzy C-Means Linear naik..... 61
Lampiran O	Hasil Cluster Dengan Fuzzy C-Means Linear Turun 62
Lampiran P	Hasil Cluster Dengan Fuzzy C-Means Segitiga..... 63

“Halaman ini sengaja dikosongkan”

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara berkembang dengan tingkat pertumbuhan penduduk setiap tahunnya cukup tinggi dimana persebaran penduduk serta lahan untuk tinggal di Indonesia tidak merata. Hal ini mengakibatkan banyak masalah sosial terjadi di Indonesia sehingga perlu dilakukan perencanaan pembangunan agar terlepas dari permasalahan sosial. Pembangunan adalah rangkaian upaya mewujudkan tujuan nasional bangsa Indonesia yang maju, mandiri, sejahtera dan berkeadilan sesuai dengan UUD 1945 alinea keempat yang menyebutkan bahwa tujuan pembangunan nasional adalah mencerdaskan kehidupan bangsa dan memajukan kesejahteraan umum yang mana hal ini mengandung makna pemberdayaan penduduk.

Sejalan dengan adanya pembangunan nasional, *United Nation* atau Perserikatan Bangsa-Bangsa pada tahun 2000 mendeklarasikan kesepakatan Tujuan Pembangunan Milenium (*Millenium Development Goal/ MDG's*). Terdapat delapan poin tujuan MDG's yaitu menanggulangi kemiskinan dan kelaparan, upaya mencapai pendidikan dasar untuk semua, mendorong kesetaraan gender dan pemberdayaan perempuan, menurunkan angka kematian anak dan balita, meningkatkan angka kesehatan ibu, memerangi HIV/AIDS, malaria dan penyakit menular lainnya, memastikan kelestarian lingkungan hidup dan mengembangkan kemitraan global untuk pembangunan. Sehingga pada intinya pembangunan adalah menciptakan lingkungan yang memungkinkan masyarakat menikmati umur panjang, sehat, dan menjalankan kehidupan produktif yang mana pada intinya adalah untuk meningkatkan kesejahteraan rakyat.

Dalam mengukur tercapai tidaknya pembangunan di suatu wilayah dibutuhkan indikator-indikator yang mampu mengukur kesejahteraan rakyat dan dijadikan landasan ukuran keberhasilan. Indikator kesejahteraan rakyat cukup luas atau multidimensional

dan juga kompleks sehingga suatu taraf kesejahteraan rakyat hanya dapat dinilai melalui indikator-indikator terukur dari berbagai aspek pembangunan..

Indikator kesejahteraan rakyat terdiri dari indikator pendidikan, indikator ketenagakerjaan, indikator demografi, indikator kesehatan, dan indikator sosial lainnya. (Badan Pusat Statistik,2014). Penelitian ini mengambil indikator ekonomi, pendidikan, dan kesehatan. Ketiga indikator ini diambil berdasarkan pemilihan poin pada tujuan MDG's dari *United Nation* dikarenakan indikator ini merupakan dimensi yang cukup dekat dengan kesejahteraan masyarakat dimana metode yang digunakan adalah pengelompokan dengan *c-means* dan *fuzzy c-means clustering*.

Analisis kelompok (*cluster analysis*) merupakan salah satu teknik dalam analisis multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek sehingga objek paling dekat kesamaannya dengan objek yang paling dekat kesamaannya berada dalam kelompok yang sama dan yang memiliki karakteristik berbeda berada pada kelompok lainnya. (Johnson, 2002).Metode *c-means clustering* adalah suatu metode pengelompokan non hirarki yang mencoba mempartisi data menjadi suatu kelompok sehingga data yang memiliki karakteristik yang sama akan berada dalam satu kelompok yang sama dan memiliki karakteristik berbeda dengan data yang berada pada kelompok lain sedangkan metode *fuzzy c-means* adalah suatu metode pengelompokan objek yang memanfaatkan dasar pembobotan dengan teori himpunan *fuzzy*. Pengelompokan dengan *fuzzy* sering digunakan karena memberikan hasil yang cukup baik dalam hal meningkatkan homogenitas tiap kelompok yang dihasilkan. (Shihab, 2000).

Penelitian terdahulu pada Ririn, 2015 dalam menentukan kelompok kota dan kabupaten di Jawa Timur berdasarkan indikator kemiskinan menggunakan metode *c-means clustering* dan *fuzzy c-means clustering* memberikan hasil bahwa menggunakan *c-means clustering* lebih baik dibandingkan dengan metode *fuzzy c-means clustering*.

Penelitian ini diharapkan mampu memberikan saran kepada BAPPENAS mengenai aspek aspek pembangunan yang harus difokuskan di tiap-tiap provinsi agar mencapai tujuan pembangunan nasional yang diharapkan.

1.2 Rumusan Masalah

Pembangunan nasional yang dilakukan oleh pemerintah kurang efisien dikarenakan kurang mampunya pemerintah dalam mengidentifikasi permasalahan yang terjadi di Indonesia maka rumusan masalah dari penelitian tersebut adalah sebagai berikut.

1. Bagaimana gambaran umum indikator kesejahteraan rakyat di Indonesia?
2. Bagaimana pengelompokan provinsi di Indonesia berdasarkan indikator kesejahteraan rakyat dengan analisa kluster *c-means* dan *fuzzy c-means* dengan *cluster* optimum?
3. Manakah metode yang lebih baik diantara keduanya berdasarkan nilai *icd rate* terkecil dan *SSW* terbesar?

1.3 Tujuan

Berdasarkan rumusan masalah yang telah dijabarkan, maka tujuan yang ingin dicapai pada penelitian ini adalah sebagai berikut.

1. Mendeskripsikan gambaran kesejahteraan rakyat tiap-tiap provinsi di Indonesia.
2. Mengelompokkan provinsi di Indonesia berdasarkan indikator kesejahteraan rakyat dengan metode *c-means* dan *fuzzy c-means*.
3. Membandingkan hasil pengelompokkan provinsi di Indonesia dengan metode *c-mean* dan *fuzzy c-means* berdasarkan nilai *icd rate* terkecil dan *SSW* terbesar.

1.4 Manfaat Penelitian

Berdasarkan latar belakang alinea-7 yang telah dipaparkan, rumusan masalah dan tujuan penelitian maka penelitian ini diharapkan memberikan manfaat agar mampu memberikan informasi kepada instansi terkait mengenai pengelompokan provinsi di Indonesia berdasarkan indikator kesejahteraan rakyat terutama pada indikator kesehatan, pendidikan dan ekonomi sehingga dapat menunjang keberhasilan pembangunan nasional.

1.5 Batasan Masalah

Batasan masalah yang digunakan dalam penelitian ini adalah menggunakan indikator kesejahteraan rakyat pada dimensi indikator pendidikan, kesehatan dan ekonomi pada tahun 2014 yang terdiri dari 11 variabel. Nilai pembobot yang digunakan dalam metode *fuzzy c-means* bernilai 2. Syntax *matlab* yang digunakan dalam penelitian ini berdasarkan Tugas Akhir yang digunakan Anggraeni Rahma Dewi yang berjudul “Pengelompokan Kabupaten/Kota di Jawa Timur berdasarkan Indikator Kemiskinan Menggunakan Metode *C-means* dan *Fuzzy C-Means*” dengan iterasi sebanyak 1000.

BAB II TINJAUAN PUSTAKA

2.1 Distribusi Normal Multivariat

Untuk mengetahui apakah data mengikuti distribusi normal, maka dilakukan uji distribusi normal multivariat. Pengujian distribusi normal multivariat dilakukan untuk memperkuat dugaan bahwa data telah berdistribusi normal multivariat dan sebagai asumsi dasar yang harus dipenuhi sebelum pengujian lainnya (Johnson & Wichern, 2007). Untuk melakukan pengujian distribusi normal multivariat maka hipotesis yang diberikan adalah sebagai berikut.

H_0 : Data berdistribusi normal multivariat

H_1 : Data tidak berdistribusi normal multivariat.

Pemeriksaan distribusi normal multivariat dapat dilakukan dengan melihat *q-q plot* atau *scatter plot* dari nilai *square distance*.

$$d_j^2 = (x_j - \bar{x})' S^{-1} (x_j - \bar{x}) ; j = 1, 2, 3 \dots n \quad (2.1)$$

dimana :

d_j^2 : Nilai jarak data ke- j

x_j : Variabel ke- j

\bar{x} : Nilai vektor rata-rata

S^{-1} : Nilai invers matriks varians-kovarians

Jika *plot* cenderung membentuk garis lurus lebih dari 0,5 dan nilai $d_j^2 \leq \chi_{p;0.5}^2$ maka data disimpulkan tidak berdistribusi normal multivariat.

2.2 Uji KMO

Uji KMO adalah suatu pengujian yang digunakan untuk mengukur kecukupan sampling secara menyeluruh dan mengukur kecukupan sampling untuk setiap indikator dengan menggunakan nilai korelasi antar variabel. Hipotesis yang diberikan untuk pengujian ini adalah sebagai berikut.

H_0 : Jumlah data cukup untuk difaktorkan

H_1 : Jumlah data tidak cukup untuk difaktorkan

Statistik Uji yang digunakan dalam pengujian ini adalah sebagai berikut.

$$KMO = \frac{\sum_{i=1}^P \sum_{j=1}^P r_{ij}^2}{\sum_{i=1}^P \sum_{j=1}^P r_{ij}^2 + \sum_{i=1}^P \sum_{j=1}^P a_{ij}^2} \quad (2.2)$$

dimana :

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

r_{ij} : Korelasi antara variabel i dan j

a_{ij} : Korelasi parsial antara variabel i dan j

Berdasarkan Hair, *et.al* (2010), sampel akan dikatakan layak untuk dilakukan analisis faktor apabila nilai $KMO > 0.5$.

2.3 Uji Barlett

Uji Barlett merupakan pengujian statistik untuk menguji persamaan korelasi antara dua atau lebih variabel independen untuk dilihat nilai korelasinya (Johnson & Wichern, 2007). Hipotesis yang diberikan dalam pengujian ini adalah sebagai berikut

$H_0: \boldsymbol{\rho} = \mathbf{I}$

$H_0: \boldsymbol{\rho} \neq \mathbf{I}$

Statistik Uji yang digunakan dalam pengujian ini adalah sebagai berikut.

$$\chi_{hitung}^2 = - \left[(n - 1) - \frac{(2p + 5)}{6} \right] \ln |R| \quad (2.3)$$

dimana :

n : Jumlah observasi

ρ : Jumlah variabel

$|R|$: Determinan dari matriks korelasi

Keputusan tolak H_0 jika nilai $\chi_{hitung}^2 > \chi_{\alpha, \frac{p(p-1)}{2}}^2$

2.4 Analisis Cluster

Metode pengelompokan (*clustering methods*) merupakan suatu teknik statistika multivariate yang mempunyai tujuan utama dalam mengelompokkan objek-objek berdasarkan kesamaan karakteristik yang dimiliki oleh objek tersebut. Dalam analisis

cluster ini, objek akan dikelompokkan sehingga setiap objek yang paling kesamaannya dengan objek lain yang berada dalam kelompok yang sama.

Metode pengelompokan dibedakan menjadi dua yaitu metode hierarki (*hierarchical clustering*) dan metode non-hierarki (*non-hierarchical clustering*) (Johnson & Wichern, 2007) dimana metode hierarki dengan berbagai pendekatan *single-linkage, complete linkage* dan *minimum-varians*. Metode non hierarki atau disebut juga metode partisi juga telah dikembangkan dengan berbagai pendekatan algoritma. Perbedaan yang terlihat jelas adalah pada permulaan prosedurnya dimana metode hierarki mengelompokkan suatu pengamatan secara bertahap, sedangkan pada metode non hierarki dilakukan dengan melakukan partisi pada ruang sampel.

2.5 C-means Cluster

C-means merupakan salah satu metode data *clustering* non hierarki yang berusaha mempartisi data yang ada kedalam bentuk satu *cluster*/kelompok sehingga data yang memiliki karakteristik yang sama akan dikelompokkan pada satu kelompok dan yang memiliki karakteristik yang berbeda akan dikelompokkan dalam kelompok yang lainnya. Dalam Mac Queen (1967) dan Johnson (2002) algoritma dasar untuk *c-means cluster* adalah sebagai berikut.

1. Menentukan jumlah *cluster*/kelompok initial, misal C.
2. Mengalokasikan data kedalam *cluster*/kelompok secara random.
3. Menghitung *centroid* rata-rata dari data yang ada di masing masing *cluster*/kelompok. Rumus untuk menghitung nilai *centroid* adalah sebagai berikut.

$$v_{ij} = \frac{\sum_{k=1}^{N_i} x_{kj}}{N_i} \quad (2.4)$$

dimana :

- v_{ij} : *centroid*/rata rata *cluster* ke-*i* untuk variabel ke-*j*
 N_i : Jumlah data yang menjadi anggota *cluster*

- i, k : indeks *cluster*
 j : indeks dari variabel
 x_{kj} : nilai data ke- k yang ada di dalam *cluster* tersebut untuk variabel ke- j

4. Mengalokasikan masing-masing data ke *centroid*/rata-rata terdekat (jarak yang digunakan adalah jarak *Euclidean*) dengan rumus sebagai berikut.

$$D_{ik} = \sqrt{\left(\sum_{j=1}^j (x_{kj} - v_{ij})^2 \right)} \quad (2.5)$$

dimana :

- D_{ik} : jarak *Euclidean* ke- i objek ke- k
 k : indeks objek
 j : indeks variabel variabel
 x : Koordinat objek
 v_{ij} : *centroid*/rata rata *cluster* ke- i untuk variabel ke- j
5. Kembali ke step-3, apabila masih ada data yang berpindah *cluster* atau apabila perubahan nilai *centroid* ada yang diatas nilai *threshold* yang ditentukan atau apabila perubahan nilai pada *objective function* yang digunakan diatas nilai *threshold* yang ditentukan. *Objective function* yang digunakan dalam *c-means cluster* adalah sebagai berikut.

$$J(X, U, V) = \sum_{k=1}^N \sum_{i=1}^C (u_{ik}) D(x_k, v_i)^2 \quad (2.6)$$

dimana :

- N : Banyaknya data
 C : Banyaknya *cluster*
 u_{ik} : Keanggotaan data ke- k kedalam *cluster* ke- i
 v_i : Nilai *centroid* (rata-rata) *cluster* ke- i
 $D(x_k, v_i)^2$: *square Euclidean distance space* antara data ke- k terhadap *centroid cluster* ke- i

2.6 Fuzzy c-means clustering

Fuzzy c-means merupakan suatu teknik pengelompokan yang dikembangkan dari *c-means* dengan mempertimbangkan himpunan keanggotaan *fuzzy* (samar) sebagai dasar pembobotan yang ditemukan pertama kali oleh Jim Bezdek pada tahun 1981 dan digunakan dalam bidang pengenalan pola (*patern recognition*), dan masih terus dikembangkan sampai saat ini. Dalam metode *FCM Cluster* ini dipergunakan variabel *membership function* *uik*, yang merujuk pada seberapa besar probabilitas suatu data bisa menjadi anggota ke dalam suatu kelompok (*cluster*) dan variabel *m* yang merupakan *weighing exponent* dari *membership function uik*. Metode ini merupakan pengembangan dari *C-Means Cluster* dengan pembobotan *fuzzy*. Konsep dasar *fuzzy c-means* adalah pertama kali menentukan pusat *cluster* yang akan menandai lokasi rata-rata untuk setiap *cluster*. Dalam konsep *fuzzy c-means clustering*, setiap objek bisa menjadi anggota dari beberapa *cluster*. Oleh karena itu batas-batas dalam *c-means clustering* bersifat tegas (*hard*) dan *fuzzy c-means clustering* bersifat lembut (*soft*) (Santosa, 2007).

Metode *fuzzy c-means clustering* pertama kali dengan menentukan pusat *cluster* yang akan menandai lokasi rata-rata untuk setiap *cluster*. Pada konsisi awal, pusat *cluster* ini dikatakan belum akurat. Tiap data memiliki derajat keanggotaan untuk tiap *cluster*. Dengan cara memperbaiki pusat *cluster* dan nilai keanggotaan data secara berulang, maka dapat dilihat bahwa pusat *cluster* akan menuju lokasi yang tepat. Perulangan ini didasarkan pada minimasi fungsi objektif yang menggambarkan jarak dari titik data yang diberikan ke pusat *cluster* yang terbobot oleh derajat jarak keanggotaan titik data tersebut. Berikut ini merupakan langkah-langkah dalam *fuzzy c-means clustering*.

- 1 Menentukan
 - a. Matriks x berukuran $I \times J$
 - I = Jumlah objek
 - J = Jumlah variabel
 - b. Jumlah cluster yang akan dibentuk ($C \geq 2$)

- c. Pangkat (pembobot eksponen/ $w > 1$). Akan tetapi yang sering digunakan sebanyak 2.
- d. Maksimum iterasi yang umum digunakan sebanyak 100
- e. Kriteria penghentian (threshold/ $\varepsilon =$ nilai positif yang sangat kecil). Akan tetapi yang sering digunakan sebanyak 10^{-6} .
2. Membentuk matriks partisi awal U (derajat keanggotaan dalam *cluster*) yang merupakan matriks partisi awal.

$$U = \begin{bmatrix} \mu_{11}(x_1) & \mu_{12}(x_2) & \dots & \mu_{1l}(x_l) \\ \mu_{21}(x_1) & \mu_{22}(x_2) & \dots & \mu_{2l}(x_l) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{c1}(x_1) & \mu_{c2}(x_2) & \dots & \mu_{cl}(x_l) \end{bmatrix}$$

3. Menghitung pusat *cluster* V untuk setiap *cluster*

$$V_{ij} = \frac{\sum_{i=1}^{I_i} (U_{ik})^w x_{kj}}{\sum_{i=1}^{I_i} (U_{ik})^w} \quad (2.7)$$

dimana:

V_{ij} : titik pusat centroid (rata-rata) *cluster* ke- i variabel ke- j

U_{ik} : derajat keanggotaan *cluster* ke- i objek ke- k

x_{kj} : nilai objek ke- k yang ada dalam *cluster* tersebut untuk variabel ke- j

I_i : jumlah objek yang menjadi anggota *cluster* ke- c

i : indeks cluster

j : indeks variabel

k : indeks objek

w : pembobot eksponen

4. Perbaiki derajat keanggotaan setiap data pada setiap *cluster* (perbaiki matriks partisi).

$$U_{ik} = \left[\sum_{j=1}^i \left(\frac{D_{ik}}{D_{jk}} \right)^{2/(w-1)} \right]^{-1} \quad (2.8)$$

dengan

$$D_{ik} = \sqrt{\left(\sum_{i=1}^j (x_{kj} - v_{ij})^2 \right)} \quad (2.9)$$

- D_{ik} : jarak *euclidean cluster* ke- i objek ke- k
 D_{jk} : jarak *euclidean* variabel ke- j objek ke- k
 j : indeks variabel
 i : indeks cluster
 k : indeks objek
 x_{ij} : nilai objek ke- k yang ada dalam *cluster* tersebut untuk variabel ke- j
 v_{ij} : titik pusat (*centroid/rata-rata*) cluster ke- i variabel ke- j

5 Menentukan kriteria penghentian iterasi, yaitu perubahan matriks partisi pada iterasi sekarang dan iterasi sebelumnya.

$$\Delta = |U^l - U^{l-1}| \quad (2.10)$$

dimana

l : iterasi ke- t

U : derajat keanggotaan

Apabila $\Delta < \varepsilon$ maka iterasi dihentikan. Tetapi, jika tidak maka kembali ke langkah-3 (Bezdek, Ehrlich, & Full, 1984).

2.7 Fungsi Keanggotaan

Fungsi keanggotaan adalah suatu kurva yang menunjukkan pemetaan titik-titik input data kedalam nilai keanggotaannya. Salah satu cara yang dapat digunakan untuk mendapatkan nilai keanggotaan adalah dengan melalui pendekatan fungsi (Kusumadewi & Hartati, 2006)

2.6.1 Representasi Linear

Bentuk representasi linear merupakan bentuk paling sederhana dan menjadi pilihan yang baik untuk mendekati konsep yang kurang jelas.

1. Representasi Linear Naik

Terdapat 2 keadaan himpunan *fuzzy* yang linear yaitu kenaikan himpunan dimulai pada nilai domain yang memiliki derajat keanggotaan nol (0) bergerak ke kanan menuju domain yang memiliki derajat keanggotaan lebih tinggi. Rumus fungsi keanggotaan linear naik dipaparkan pada rumus 2.7.

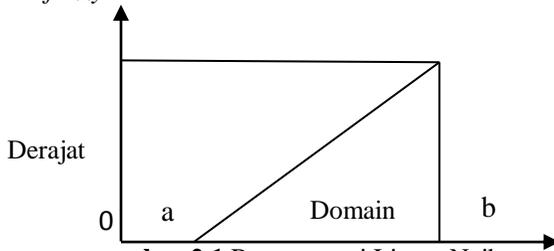
$$\mu(x) = \begin{cases} 0 & ; x \leq a \\ \frac{(x - a)}{(b - a)} & ; a \leq x \leq b \\ 1 & ; x \geq b \end{cases} \quad (2.11)$$

dimana:

a : nilai minimum dari domain

b : nilai maksimum dari domain

Gambar 2.1 merupakan gambar representasi fungsi keanggotaan *fuzzy* kurva linear naik.



Gambar 2.1 Representasi Linear Naik

2. Representasi Linear Turun

Fungsi keanggotaan linear yang kedua merupakan kebalikan dari yang pertama, yakni representasi linear turun. Garis lurus dimulai dari nilai domain dengan derajat keanggotaan tertinggi pada sisi kiri, kemudian bergerak menurun pada domain yang lebih rendah. Rumus fungsi keanggotaan linear turun dipaparkan pada rumus 2.8.

Derajat Keanggotaan :

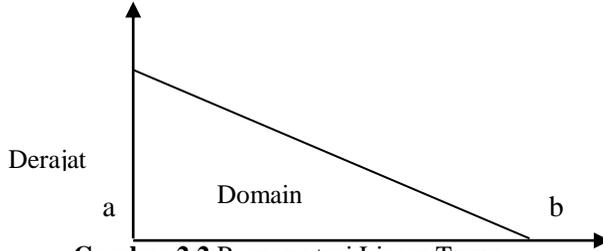
$$\mu(x) = \begin{cases} \frac{(b - x)}{(b - a)} & ; a \leq x \leq b \\ 0 & ; x \geq b \end{cases} \quad (2.12)$$

dimana:

a : nilai minimum dari domain

b : nilai maksimum dari domain

Gambar 2.2 merupakan gambar representasi fungsi keanggotaan *fuzzy* kurva linear turun.



Gambar 2.2 Representasi Linear Turun

2.6.2 Representasi Kurva Segitiga

Fungsi keanggotaan segitiga pada dasarnya seperti bentuk segitiga hanya saja ada beberapa titik yang memiliki nilai keanggotaan 1. Rumus fungsi keanggotaan linear naik dipaparkan pada rumus 2.9.

Derajat Keanggotaan :

$$\mu_{(x)} = \begin{cases} 0 ; x \leq a \text{ atau } x \geq c \\ \frac{(x - a)}{(b - a)} ; a \leq x \leq b \\ \frac{(c - x)}{(c - b)} ; b \leq x \leq a \end{cases} \quad (2.13)$$

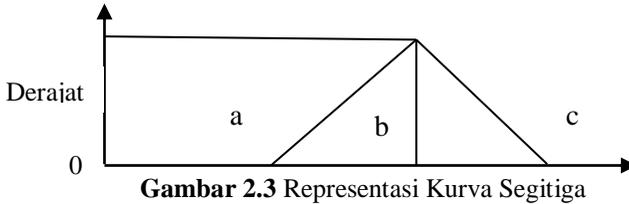
dimana:

a : nilai minimum dari domain

b : nilai median dari domain

c : nilai maksimum dari suatu domain

Gambar 2.3 merupakan gambar representasi fungsi keanggotaan *fuzzy* kurva segitiga.



Gambar 2.3 Representasi Kurva Segitiga

2.8 Principle Component Analysis (PCA)

Principle component analysis (PCA/analisis komponen utama) digunakan untuk menjelaskan struktur matriks varians kovarians dari suatu set variabel melalui kombinasi linier dari variabel-variabel. Secara umum, komponen utama berguna untuk mereduksi dan menginterpretasi variabel-variabel. Misalkan, terdapat J buah variabel yang terdiri atas I buah objek dan J buah variabel tersebut dibuat sebanyak K buah komponen utama (dengan $K < J$) yang merupakan kombinasi linier atas J buah variabel. K komponen utama tersebut dapat menggantikan J buah variabel yang membentuk tanpa kehilangan banyak informasi mengenai keseluruhan variabel. Secara umum analisis komponen utama merupakan analisis *intermediate* yang berarti hasil komponen utama dapat digunakan untuk analisis lebih lanjut.

Komponen utama tergantung kepada matrik varian kovarian Σ dan matrik korelasi r dari seluruh variabel $x_1, x_2, x_3, \dots, x_j$ dimana pada analisis tidak memerlukan asumsi populasi harus berdistribusi *Multivariate Normal*. Apabila komponen utama diturunkan dari populasi *Multivariate Normal* interpretasi dan inferensi dapat dibuat dari komponen sampel. Penyusutan variabel dimensi dengan cara mengambil sejumlah kecil komponen yang mampu menerangkan bagian terbesar keragaman data. Apabila komponen utama yang diambil sebanyak K buah, dimana $K < J$, maka proporsi keragaman yang bisa diterangkan sebagai berikut.

$$\text{keragaman total} = \frac{\lambda_k}{\lambda_1 + \lambda_2 + \dots + \lambda_K} \quad (2.14)$$

dimana $k = 1, 2, \dots, K$

Sehingga nilai proporsi dari varian total populasi dapat diterangkan oleh komponen pertama, kedua atau sampai sejumlah K komponen utama secara bersama-sama adalah semaksimal mungkin. Tidak ada ketentuan berapa besar proporsi keragaman data yang dianggap cukup mewakili keragaman total. Meskipun jumlah komponen utama berkurang dari variabel asal tetapi ini merupakan gabungan dari variabel-variabel asal sehingga informasi yang diberikan tidak berubah.

Pemilihan komponen utama yang digunakan didasarkan pada *eigen value* (λ) yang bernilai lebih besar dari 1 ($\lambda_k > 1$). Hal ini dikarenakan *eigen value* berasal dari matrik varian kovarian Σ dan matrik korelasi r yang merupakan *standardized* dari matriks varian kovarian Σ dengan rata-rata sebesar 1 (Rencher, 2002). Secara Ideal, banyak komponen utama yang secara kumulatif telah dapat menerangkan sekitar 60% atau lebih variasi dalam data. Nilai λ adalah akan mengandung nilai *eigen vector* (e'_i). Masing-masing nilai dari *eigen vector* (e'_i) dikalikan dengan masing-masing variabel asli. Sehingga menghasilkan persamaan analisis komponen utama sesuai dengan nilai k dengan keragaman total telah terpenuhi (Johnson & Wichern, 2007).

2.9 Pseudo F-Statistics

Penentuan jumlah *cluster* optimum yang pembentukan *cluster* ditentukan oleh jarak *Euclidean* sesuai metode yang digunakan, maka menggunakan *pseudo-statistic* (Timm, Applied Multivariate Analysis, 2002). Nilai *pseudo f-statistics* tertinggi menunjukkan bahwa jumlah kelompok telah optimal, dimana keseragaman dalam kelompok sangat homogen sedangkan antar kelompok sangat heterogen. Menurut Orpin & Kostlev, 2006 rumus yang digunakan dalam menghitung nilai *pseudo f-statistics* adalah sebagai berikut.

$$\text{Pseudo } F\text{-statistics} = \frac{\left(\frac{R^2}{i-1}\right)}{\left(\frac{1-R^2}{I-i}\right)} \quad (2.15)$$

dengan

$$R^2 = \frac{(SST - SSW)}{SST} \quad (2.16)$$

$$SST = \sum_{n=1}^N \sum_{i=1}^I \sum_{j=1}^J (x_{ni}^j - \bar{x}^j)^2 \quad (2.17)$$

$$SSW = \sum_{n=1}^N \sum_{i=1}^I \sum_{j=1}^J (x_{ni}^j - \bar{x}_i^j)^2 \quad (2.18)$$

dimana

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

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

N : banyak objek

I : banyak *cluster*

j : banyak variabel

x_{ni}^j : sampel ke- n kelompok ke- i variabel ke- j

\bar{x}^j : rata-rata seluruh sampel pada variabel ke- j

\bar{x}_i^j : rata-rata sampel pada kelompok ke- i variabel ke- j

2.10 Internal Cluster Dispersion Rate (Icdrate)

Internal Cluster Dispersion Rate (Icd rate) merupakan metode yang digunakan untuk membandingkan metode kluster yang terbaik dengan mengevaluasi performansi algoritma dengan menggunakan prosentase rata-rata dari klasifikasi yang benar (*recovery rate*). Nilai persebaran data-data dalam *cluster (Icd rate)* dari hasil akhir pengelompokan.

$$\begin{aligned} \text{Icdrate} &= 1 - \frac{SSB}{SST} = 1 - \frac{SST_SSW}{SST} \\ &= 1 - R^2 \end{aligned} \quad (2.19)$$

dengan

$$R^2 = \frac{(SST - SSW)}{SST} \quad (2.20)$$

$$SSB = SST - SSW \quad (2.21)$$

SST:total jumlah dari kuadrat jarak terhadap rata-rata keseluruhan
SSW:total jumlah dari kuadrat jarak objek terhadap rata-rata kelompoknya.

R^2 :*Recovery Rate*

2.11 One-way ANOVA

Menentukan respon mana yang dipengaruhi oleh perlakuan yang dalam hal ini adalah hasil *cluster* dapat diperoleh melalui pengujian *One-way ANOVA (Analysis of Variance)*. Berikut adalah hipotesis yang digunakan dalam pengujian *One-way ANOVA*

$$H_0: \tau_1 = 0$$

$$H_1: \tau_1 \neq 0$$

Statistik uji:

$$F_{hitung} = \frac{\sum_{t=1}^T n_t (\bar{X}_t - \bar{X})^2 / (t-1)}{\sum_{j=1}^J \sum_{t=1}^T (\bar{X}_{ij} - \bar{X})^2 / (\sum_{t=1}^T n_t - t)} \quad (2.22)$$

Keterangan:

Treat = nilai *Sum of Square* residual

Treat + Error = nilai *Sum of Square* total

Tolak H_0 Fhitung lebih besar dari F-tabel. Berikut merupakan tabel uji *One-way ANOVA*:

Tabel 2.1 Uji *One-way* ANOVA

Sumber Variasi	Jumlah Kuadrat	Derajat Bebas (db)
Perlakuan	$Treat = \sum_{t=1}^T n_t (\bar{X}_t - \bar{X})^2$	$T - 1$
Residual (Error)	$Error = \sum_{j=1}^J \sum_{t=1}^T (\bar{X}_{tj} - \bar{X}_j)^2$	$\sum_{t=1}^T n_t - T$
Total	$Treat + Error = \sum_{j=1}^J \sum_{t=1}^T (\bar{X}_{tj} - \bar{X})^2$	$\sum_{t=1}^T n_t - 1$

(Johnson & Wichern, 2007).

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian Tugas Akhir ini adalah data indikator kesehatan, pendidikan dan ekonomi pada tahun 2014 yang diambil pada data Survey Ekonomi Nasional (SUSENAS) , Survey Angkatan Kerja (SAKERNAS), Profil Kesehatan Indonesia dari Badan Pusat Statistik (BPS) Indonesia. Unit penelitian yang digunakan adalah sebanyak 33 provinsi yang ada di Indonesia.

3.2 Variabel Penelitian

Berdasarkan tinjauan pustaka yang telah dijabarkan, variabel yang digunakan dalam penelitian ini adalah

1. Indikator Kesehatan

Indikator kesehatan merupakan salah satu indikator kesejahteraan rakyat yang mengukur kesejahteraan rakyat dari derajat kesehatan.

Tabel 3.1 Variabel Indikator Kesehatan

Variabel	Keterangan	Skala
X1	ASI Eksklusif	Rasio
X2	Angka Kesakitan malaria	Rasio
X3	Imunisasi Anak Sekolah	Rasio
X4	Rumah Tangga Air Minum Layak	Rasio
X5	Rumah Tangga Sanitasi Layak	Rasio

2. Indikator Pendidikan

Indikator pendidikan merupakan salah satu indikator kesejahteraan rakyat yang mengukur kesejahteraan rakyat dari mutu pendidikan.

Tabel 3.2 Variabel Indikator Pendidikan

Variabel	Keterangan	Skala
X6	Angka Melek Huruf	Rasio
X7	Angka Partisipasi Sekolah	Rasio
X8	Angka Lama Sekolah	Rasio

3. Indikator Ekonomi

Indikator ekonomi dalam penelitian ini merupakan salah satu indikator kesejahteraan rakyat yang mengukur kesejahteraan rakyat dari mutu demografi dan ketenagakerjaan.

Tabel 3.3 Variabel Indikator Ekonomi

Variabel	Keterangan	Skala
X9	Indeks Gini	Rasio
X10	Penduduk Miskin	Rasio
X11	Tingkat Pengangguran Terbuka	Rasio

3.3 Struktur Data

Struktur data yang digunakan dalam penelitian ini adalah sebagai berikut.

Tabel 3.4 Struktur Data

Provinsi	X_1	X_2	...	X_{11}
(1)	(2)	(3)	...	(1)
1	X_{11}	X_{21}	...	$X_{15;1}$
2	X_{12}	X_{22}	...	$X_{15;2}$
3	X_{13}	X_{23}	...	$X_{15;3}$
⋮	⋮	⋮	⋮	⋮
33	$X_{1;33}$	$X_{2;33}$...	$X_{15;33}$

3.4 Definisi Operasional

Berikut adalah definisi operasional masing-masing indikator untuk masing-masing variabel indikator kesehatan, pendidikan, dan ekonomi.

1. Konsep Indikator Kesehatan

Konsep kesehatan merupakan variabel yang mampu mengukur indikator kesejahteraan rakyat berdasarkan tingkat mutu kesehatan/ pembangunan kesehatan masing-masing provinsi di Indonesia.

- a. ASI Eksklusif adalah persentase bayi usia 0-6 bulan yang yang diberi air susu ibu tanpa tambahan dan minuman lain kepada bayi sejak lahir sampai usia 6 bulan.
- b. Angka Kesakitan Malaria adalah persentase penduduk yang mengalami keluhan malara hingga mengganggu aktivitas sehari-hari. Dapat diukur dengan perbandingan

jumlah penderita malaria dengan jumlah penduduk pada tempat dan waktu yang sama per 100.000 penduduk.

- c. Imunisasi Anak Sekolah persentase anak sekolah yang diimunisasi setiap provinsi. Dalam penelitian ini variabel imunisasi anak sekolah diambil untuk anak sekolah di kelas 2 dan 3 SD.
- d. Rumah Tangga Dengan Akses Air Bersih persentase rumah tinggal dengan bersanitasi air bersih. Air bersih meliputi air yang dibeli, PAM/PDAM, mata air dan sumur terlindungi.
- e. Rumah Tangga Dengan Sanitasi Layak adalah persentase rumah tinggal dengan ketersediaan fasilitas buang air, ketersediaan tangki pembuangan tinja dan sanitasi layak lainnya.

2. Konsep Indikator Pendidikan

Konsep pendidikan adalah variabel yang menjelaskan indikator kesejahteraan rakyat berdasarkan tingkat mutu pendidikan masing-masing provinsi di Indonesia.

- a. Angka Melek Huruf adalah persentase penduduk usia 15 tahun keatas yang bisa membaca dan menulis serta mengerti kalimat sederhana di kehidupan sehari-hari
- b. Angka Partisipasi Sekolah adalah proporsi dari semua anak yang masih sekolah pada suatu kelompok umur tertentu terhadap jumlah penduduk dengan kelompok umur sesuai. Penelitian ini khusus mengambil proporsi anak berusia 16-18 tahun keatas yang merupakan usia diluar wajib belajar 9 tahun yang dicanangkan pemerintah yang dinyatakan dalam persen.
- c. Angka Lama Sekolah adalah rata-rata jumlah tahun yang dihabiskan oleh penduduk usia 15 tahun keatas untuk menempuh semua jenis pendidikan formal yang pernah dijalani.

3. Konsep Indikator Ekonomi

Konsep ekonomi merupakan variabel yang menjelaskan kesejahteraan rakyat yang dikaitkan dengan tingkat pengeluaran dan pendapatan.

- a. Indeks Gini adalah ukuran pemerataan pendapatan yang dihitung berdasarkan kelas pendapatan. Nilai koefisien Gini terletak antara nol yang mencerminkan pemerataan sempurna dan satu yang menggambarkan ketidakmerataan sempurna.
- b. Penduduk Miskin adalah persentase penduduk yang memiliki rata-rata pengeluaran perkapita dibawah garis kemiskinan.
- c. Tingkat Pengangguran Terbuka adalah indikasi tentang penduduk usia kerja yang termasuk dalam kelompok pengangguran dan dapat diukur sebagai persentase jumlah pengangguran terhadap jumlah angkatan kerja.

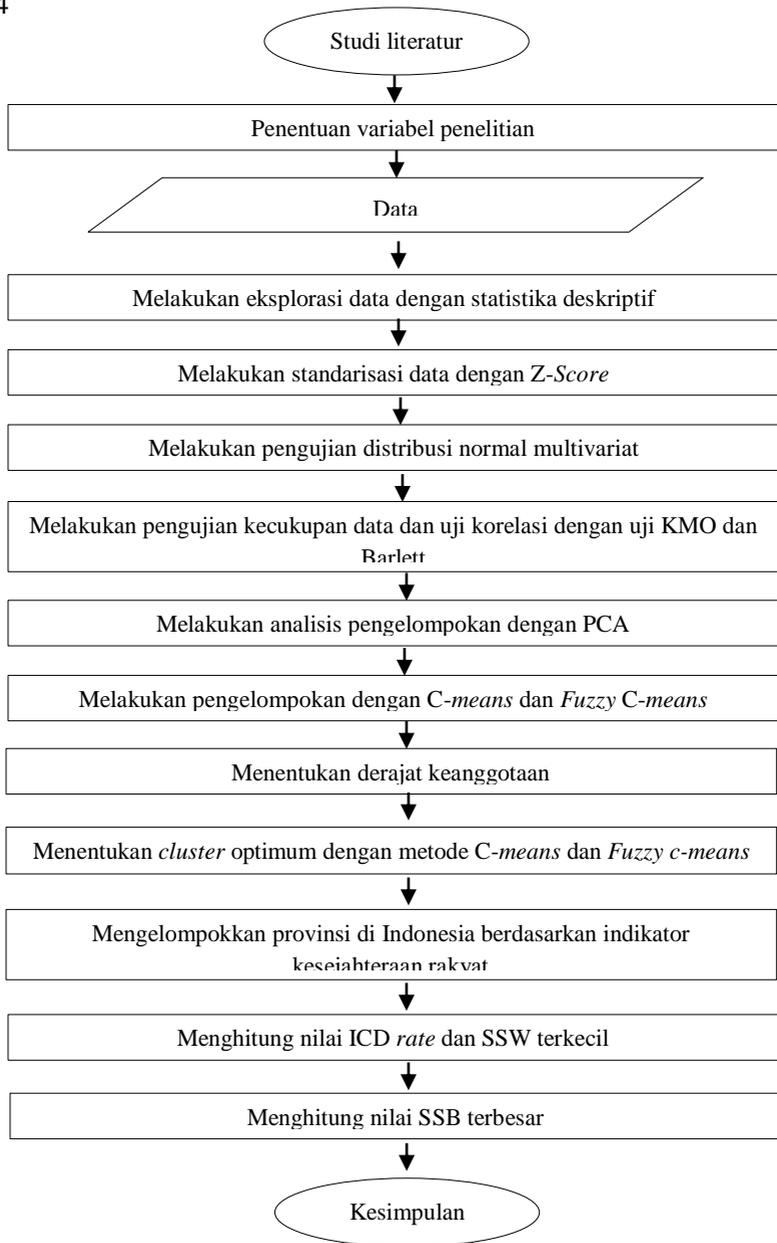
3.5 Langkah-langkah Penelitian

Berdasarkan sumber data dan variabel penelitian yang telah dipaparkan sebelumnya, langkah analisis yang akan dilakukan dalam penelitian Tugas akhir ini adalah sebagai berikut.

1. Menggambarkan indikator kesejahteraan rakyat Indonesia di tiap-tiap provinsi. Langkah langkah yang dilakukan adalah sebagai berikut.
 - a. Mengumpulkan data yang berisi variabel penelitian Tugas akhir
 - b. Analisa Statistika Deskriptif
2. Pengelompokan provinsi di Indonesia berdasarkan indikator kesejahteraan rakyat menggunakan *c-means* dan *fuzzy c-means* dengan *cluster* optimum. Langkah langkah yang dilakukan adalah sebagai berikut.
 - a. Standarisasi data dengan Z-score
 - b. Melakukan pengujian distribusi normal multivariat.

- c. Melakukan pengujian kecukupan data dan uji korelasi antar variabel penelitian dengan uji KMO dan uji barlett.
 - d. Melakukan Analisis Pengelompokan dengan *Principal Component Analysis*.
 - e. Melakukan Pengelompokan dengan *C-means* dan *Fuzzy c-means*.
 - f. Menentukan derajat keanggotaan melalui fungsi keanggotaan linier naik, linier turun, dan segitiga sebagai inputasi dari metode *fuzzy c-means clustering* dengan pereduksian variabel menggunakan *principle component analysis*.
 - g. Menentukan jumlah *cluster* optimum dengan metode *c-means* dan *fuzzy c-means clustering* melalui nilai *pseudo-f statistics* terbesar. Dalam *fuzzy c-means clustering* digunakan fungsi keanggotaan linier naik, linier turun dan segitiga.
 - h. Mengelompokkan provinsi di Indonesia berdasarkan indikator kesejahteraan rakyat sesuai dengan jumlah *cluster* optimum dan menentukan karakteristik masing-masing *cluster*.
3. Membandingkan metode *c-means clustering* dan *fuzzy c-means clustering* dilakukan dengan langkah-langkah sebagai berikut.
 - a. Menghitung nilai *icd rate* dan SSW terkecil
 - b. Menghitung nilai SSB terbesar.
 - c. Menentukan perbedaan karakteristik

Berdasarkan langkah analisis yang telah dipaparkan maka diagram alir penelitian ini adalah sebagai berikut.



Gambar 3.1 Diagram Alir Penelitian

BAB IV

ANALISIS DAN PEMBAHASAN

Pada bab ini akan dilakukan analisis dan pembahasan tentang pengelompokan indikator kesejahteraan rakyat di Indonesia di tahun 2014 dengan menggunakan metode *c-means clustering* dan *fuzzy c-means clustering* dimana data yang digunakan adalah data sekunder yang diperoleh dari hasil Survey Badan Pusat Statistika Nasional.

4.1 Deskripsi Indikator Kesejahteraan Rakyat

Deskripsi kondisi indikator kesejahteraan rakyat tiap provinsi dapat digambarkan dengan analisis statistika deskriptif. Deskripsi mengenai kondisi kesejahteraan rakyat dapat diketahui dengan ukuran pemusatan dan penyebaran. Ukuran pemusatan menggunakan nilai rata-rata (*mean*), minimum dan maksimum. Ukuran penyebaran menggunakan nilai standard deviasi dan varians.

4.1.1 Deskripsi Bidang Kesehatan

Sub-bab ini merupakan deskripsi mengenai kondisi kesejahteraan rakyat bidang kesehatan yang dilihat dari beberapa variabel yaitu Pemberian ASI Eksklusif, Angka Kesakitan Malaria, Imunisasi Anak Sekolah, Rumah Tangga Akses Air Minum Layak dan Rumah Tangga Akses Sanitasi Layak. Analisa statistika deskriptif untuk indikator kesejahteraan rakyat bidang kesehatan yang mengacu pada data Lampiran A dapat digambarkan dengan Tabel 4.1.

Tabel 4.1 Statistika Deskriptif Berdasarkan Indikator Kesehatan

Variabel	Tertinggi		Terendah		Mean
	Provinsi	%	Provinsi	%	%
ASI EKSLUSIF	NTB	84,7	Jawa Barat	21,8	59,3
ANGKA MALARIA	Papua	29,5	Bali	0	2,31
IMUNISASI ANAK	Jawa Tengah	99,5	Sulawesi Utara	43,2	91,6
AKSES AIR MINUM	Bali	93,2	Bengkulu	35,17	65,3
AKSES SANITASI	DKI Jakarta	86,8	NTT	12,77	55,9

Berdasarkan Tabel 4.1 didapatkan bahwa ada beberapa provinsi menduduki peringkat tertinggi dan terendah untuk beberapa indikator kesehatan. Nusa Tenggara Barat (NTB) merupakan provinsi yang memiliki persentase ASI Eksklusif tertinggi nasional dan Jawa Barat merupakan yang terendah dan ternyata NTB merupakan provinsi yang mengeluarkan peraturan pemerintah daerah tentang pemberian air susu eksklusif dimana dalam peraturan tersebut mengatur mengenai tugas dan tanggung jawab pemerintah dalam pengembangan program ASI Eksklusif. *Means* sebesar 59,3% dan terdapat 13 provinsi berada dibawah rata-rata ASI Eksklusif nasional.

Peringkat tertinggi Angka Kesakitan Malaria adalah Papua dan terendah berada di Bali. Papua memiliki daerah pesisir pantai dan dataran rendah dan wilayah pegunungan tengah merupakan wilayah endemik malaria sehingga hal ini menyebabkan angka malaria di Papua masih tinggi. Terdapat 8 provinsi berada diatas rata-rata Angka Malaria nasional dan butuh perhatian lebih dari pemangku kepentingan masyarakat. Persentase Imunisasi Anak juga merupakan indikator kesehatan rakyat yang perlu diperhatikan karena terdapat 9 provinsi yang memiliki persentase dibawah rata-rata sebesar 91,67% terutama Sulawesi utara karena hal ini menunjukkan kurangnya kesadaran masyarakat untuk memberikan vaksin kepada anak agar terhindar dari penyakit.

Persentase Rumah Tangga Akses Air Minum dan Sanitasi layak merupakan indikator kesehatan yang perlu diperhatikan karena ternyata Indonesia merupakan negara dengan 5 negara akses sanitasi dan air minum bersih yang rendah di dunia. Terdapat 9 provinsi yang berada dibawah rata-rata nasional dan NTT merupakan provinsi dengan sanitasi terendah di Indonesia dan ternyata NTT juga merupakan provinsi dengan angka kesakitan diare tertinggi karena hal ini berbanding lurus dengan sanitasi yang buruk. Pemerintah sebaiknya memberikan perhatian lebih kepada daerah yang memiliki akses sanitasi dan air minum yang kurang layak demi kesejahteraan rakyat yang lebih baik sehingga pembangunan nasional tercapai dengan baik.

4.1.2 Deskripsi Bidang Pendidikan

Indikator Kesejahteraan Rakyat yang mampu mengindikasikan rakyat tersebut sejahtera adalah dengan melihat mutu pendidikan rakyat nya. Terdapat beberapa variabel yang merupakan indikator mutu pendidikan suatu daerah yaitu Angka Melek Huruf (AMH), Angka Partisipasi Sekolah (APS) dan Angka Lama Sekolah (ALS). Berikut merupakan statistika deskriptif yang berupa peringkat di masing masing variable yang mengacu pada data Lampiran A dipaparkan dalam Tabel 4.2.

Tabel 4.2 Statistika Deskriptif Berdasarkan Indikator Pendidikan

Variabel	Tertinggi		Terendah		Means
	Provinsi	Nilai	Provinsi	Nilai	
AMH	Sulawesi utara	99,6	Papua	70,78	95,4
APS	DI Yogyakarta	86,44	Papua	61,63	72,7
ALS	DKI Jakarta	10,98	Papua	6,32	8,59

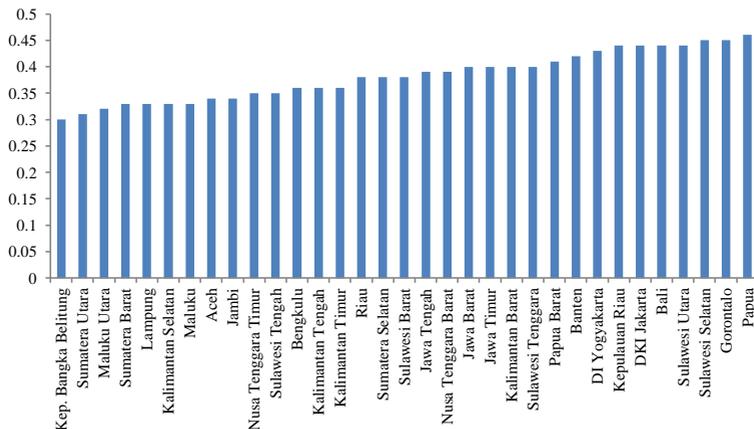
Tabel 4.2 memberikan informasi bahwa provinsi Papua merupakan provinsi dengan mutu pendidikan yang dari segala aspek cukup tertinggal baik dari kemampuan membaca yang diindikasikan dengan AMH, kesadaran akan pentingnya pendidikan formal yang dapat dilihat dari APS dan ALS. Ternyata hal ini dikarenakan kurangnya tenaga pengajar dan infrastruktur yang kurang memadai. Maka, berdasarkan fakta ini sebaiknya pemerintah memberikan perhatian lebih di bidang pendidikan kepada provinsi Papua agar memiliki mutu pendidikan yang lebih baik.

Peringkat tertinggi di nasional Angka Melek Huruf adalah Sulawesi Utara hal ini dikarenakan di daerah ini pemerintah daerah memang sedang menggalakkan wajib membaca dan menggalakkan berbagai terobosan untuk meningkatkan masyarakat agar mendapatkan pendidikan yang berkualitas dengan adanya rintisan sekolah nasional (SSN). Angka Partisipasi Sekolah dalam penelitian ini adalah untuk SMA keatas karena biasanya tertinggi adalah provinsi DI Yogyakarta dan terdapat 17

provinsi yang memiliki Angka Partisipasi Sekolah SMA yang berada di bawah rata-rata dan membutuhkan perhatian dari pemerintah setempat. Angka Lama Sekolah terlama adalah DKI Jakarta dimana provinsi ini merupakan provinsi terbesar di Indonesia dan merupakan ibukota negara Indonesia sehingga pemerintah daerah provinsi ini memang serius dalam melakukan pembangunan karna merupakan *role model* untuk provinsi lainnya.

4.1.3 Deskripsi Bidang Ekonomi

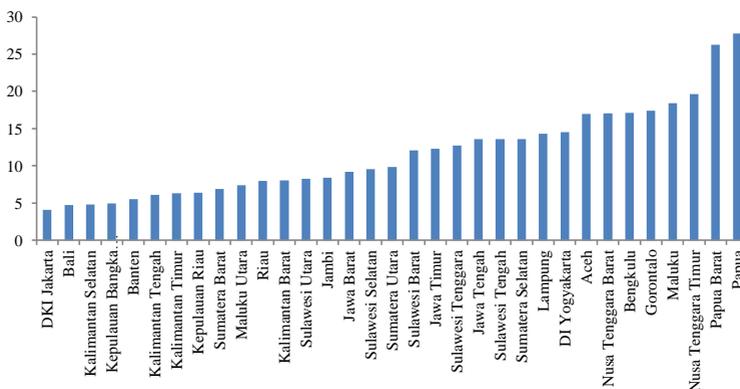
Indikator ekonomi memberikan gambaran ketimpangan pendapatan, penduduk miskin dan pengangguran terbuka yang ada pada suatu daerah dimana indikator tersebut mampu merepresentasikan keadaan kesejahteraan rakyat yang ada pada daerah tersebut (BPS,2012)



Gambar 4.1 Indeks Gini Provinsi di Indonesia

Tabel 4.1 mengacu pada data Lampiran A merupakan tabel yang menunjukkan Indeks Gini provinsi di Indonesia. Indeks Gini merupakan indikasi ketimpangan pendapatan di suatu daerah dan digagas oleh Bank Dunia. Variabel yang diambil adalah gabungan pendapatan seluruh golongan yaitu golongan 40 persen

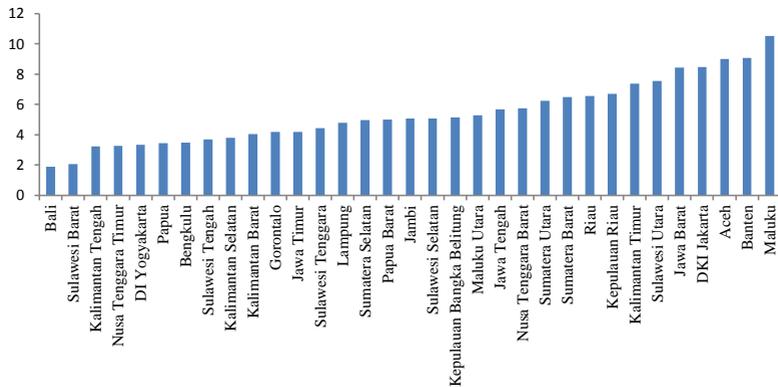
berpendapatan rendah, 40 persen pendapatan sedang, 20 persen golongan tinggi. Provinsi Kepulauan Bangka Belitung merupakan provinsi yang memiliki ketimpangan pendapatan terendah di Indonesia dan ini mengindikasikan bahwa tingkat ketimpangan antar kelompok cukup rendah. Indikator lain yang mampu menggambarkan kesejahteraan rakyat adalah banyaknya penduduk miskin yang ada pada suatu daerah. Berikut merupakan grafik persentase penduduk miskin setiap provinsi di Indonesia yang mengacu pada data Lampiran A.



Gambar 4.2 Persentase Penduduk Miskin di Indonesia

Gambar 4.2 merupakan grafik persentase penduduk miskin yang ada di Indonesia dan didapatkan bahwa Papua merupakan provinsi dengan persentase penduduk miskin tertinggi. Menurut BPS dari tahun 2010 hingga 2015 Indonesia mengalami tren menurun selama periode tersebut. Jika diukur dari tempat tinggal penduduk miskin lebih banyak berada di daerah pedesaan dibandingkan dengan daerah perkotaan dan ini dapat dilihat dari provinsi Papua yang merupakan daerah pedalaman di Indonesia yang memiliki persentase penduduk miskin sebesar 27,8 %.

Indikator lain yang menggambarkan kesejahteraan rakyat suatu daerah adalah banyaknya penduduk usia kerja yang tidak memiliki pekerjaan yang digambarkan dengan tingkat pengangguran terbuka yang mengacu pada data Lampiran A.



Gambar 4.3 Tingkat Pengangguran terbuka di Indonesia

Gambar 4.3 menunjukkan gambar grafik Tingkat Pengangguran Terbuka yang merupakan salah satu indikator ketenagakerjaan yang pada penelitian ini masuk dalam indikator ekonomi yang mampu mengindikasikan capaian hasil pembangunan nasional. Maluku merupakan provinsi di Indonesia yang memiliki tingkat pengangguran terbuka tertinggi di Indonesia dan ini mencerminkan besarnya jumlah penduduk di Maluku yang masuk dalam kategori usia kerja dan pengangguran.

4.2 Pengelompokan Provinsi di Indonesia Dengan *C-mean Clustering* dan *Fuzzy C-means Clustering*

Pemaparan dengan statistika deskriptif memberikan gambaran kondisi kesejahteraan rakyat di tiap provinsi Indonesia maka selanjutnya perlu dilakukan pengelompokan dengan metode *c-means* dan *fuzzy c-means* berdasarkan indikator kesejahteraan rakyat agar tujuan dalam penelitian dapat dicapai. Pengelompokan memerlukan asumsi yang harus dipenuhi yaitu normal multivariate

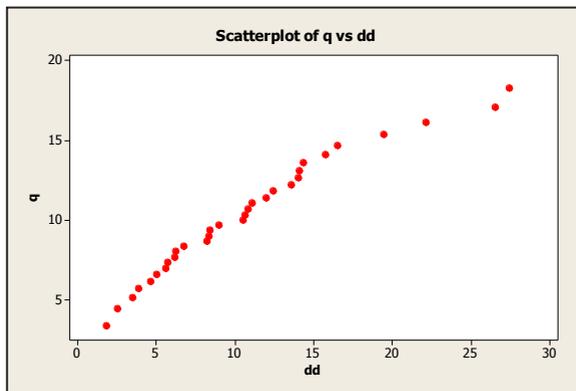
dan multikolinearitas sehingga perlu dilakukan pengujian asumsi terlebih dahulu.

4.2.1 Uji Asumsi Multivariat

Pengujian ini dilakukan secara serentak seluruh variabel indikator kesejahteraan rakyat di Indonesia tahun 2014. Pengujian ini diperlukan untuk memenuhi asumsi dalam analisis multivariat.

4.2.1.1 Asumsi Multivariat Normal

Dalam melakukan analisis data multivariat, perlu dilakukan pengujian asumsi distribusi normal untuk mengetahui apakah data telah berdistribusi normal multivariat sehingga layak untuk dilakukan analisis multivariat selanjutnya. Hipotesis dan statistik uji mengacu pada persamaan 2.1. Berikut adalah hasil dari pengujian normal multivariat untuk semua indikator kesejahteraan rakyat yang mengacu pada Lampiran J merupakan *output* dari *software* macro *minitab*.



Gambar 4.4 Plot Distribusi Normal Multivariat

Gambar 4.4 menunjukkan bahwa jika dilihat dari visual plot QQ cenderung membentuk garis lurus dan ini diperkuat dengan nilai t yang bernilai 0,5152 yang berarti lebih dari 0,5 dan memberikan keputusan gagal tolak H_0 . Sehingga dapat disimpulkan bahwa data indikator kesejahteraan rakyat di

Indonesia pada tahun 2014 memenuhi asumsi distribusi normal multivariat.

4.2.1.2 Asumsi Korelasi dan Kecukupan Data

Pengujian ini dilakukan untuk mengetahui apakah data indikator kesejahteraan rakyat di Indonesia memiliki korelasi dan untuk menguji apakah data telah cukup untuk dilanjutkan ke analisis selanjutnya. Uji korelasi dilakukan dengan *Barlett-test* dan uji kecukupan data dilakukan dengan *KMO-test*. Hipotesis dan statistik uji mengacu pada persamaan 2.2 dan 2.3. Berikut merupakan hasil dari Uji Barlett dan Uji KMO.

Tabel 4.3 Uji Barlett dan Uji KMO Indikator Kesejahteraan Rakyat Indonesia

<i>Approx. Chi Square</i>	157,14
Df	55
<i>Sig.</i>	0,000
<i>Kaiser Meyer Olkin Measure of Sampling Adequacy (KMO)</i>	0,583

Tabel 4.3 menunjukkan bahwa nilai *chi-square* untuk indikator kesejahteraan rakyat di Indonesia tahun 2014 sebesar 157,14 dan *p-value* sebesar sebesar 0,000 yang mana nilai ini memberikan keputusan tolak H_0 karena *p-value* kurang dari α (0,05) sehingga dapat disimpulkan bahwa antar variabel memiliki korelasi yang signifikan. nilai KMO adalah sebesar 0,583 dan nilai ini lebih besar dari α (0,05) sehingga memberikan keputusan tolak H_0 yang artinya data Indikator Kesejahteraan Rakyat Indonesia tahun 2014 telah cukup sehingga dapat dilanjutkan dengan analisis multivariat selanjutnya.

4.2.2 *Principal Component Analysis (PCA)*

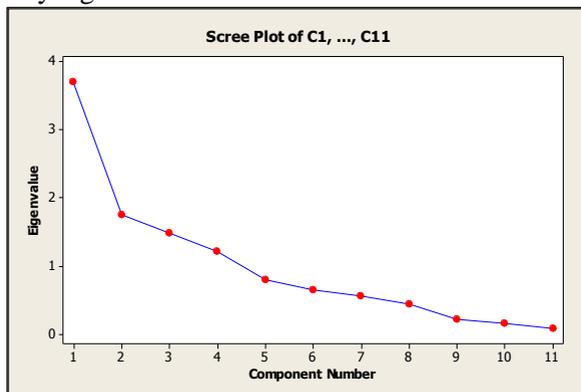
Penelitian ini menggunakan 11 variabel dan 33 obyek (provinsi) di Indonesia sehingga perlu dilakukan pereduksian dengan menggunakan *Principle component Analysis (PCA)* agar tiap komponen yang terbentuk dapat mewakili seluruh variabel. Untuk mentukan berapa banyak komponen utama yang terbentuk

dapat diketahui melalui *output total variance explained* pada Tabel 4.4 berikut ini.

Tabel 4.4 *Eigenvalues dan Percentage of Variance*

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative%
1	3,704	33,672	33,672
2	1,742	15,834	49,506
3	1,473	13,387	62,893
4	1,217	11,062	73,955
5	0,791	7,186	81,141
6	0,645	5,863	87,005
7	0,555	5,045	92,049
8	0,439	3,994	96,043
9	0,217	1,969	98,013
10	0,146	1,328	99,340
11	0,073	0,660	100,000

Berdasarkan Tabel 4.4 tersebut didapat bahwa terdapat 4 nilai *eigen* yang bernilai lebih dari 1 yaitu pada komponen 1 hingga 4 yang mana ini diartikan bahwa reduksi dari 11 variabel tersebut dapat dibentuk 4 komponen utama dan total nilai keragaman yang dapat dijelaskan sebesar 73,955%. Berikut ini merupakan gambar *scree plot* yang mampu menjelaskan komponen yang terbentuk secara visual.



Gambar 4.5 *Scree Plot*

Gambar 4.5 menunjukkan bahwa pada komponen 1 sampai 2 grafik turun secara drastis. Hal ini dikarenakan *range* terbesar antara komponen 1 dan 2, namun dari titik 2 ke 3 dan titik 3 ke 4 juga mengalami penurunan namun tidak cukup drastis sehingga komponen utama yang diduga terbentuk ada sebanyak 4 komponen utama yang dapat mewakili variabel-variabel pada data Indikator Kesejahteraan Rakyat Indonesia tahun 2014. Berikut ini merupakan persamaan berdasarkan 4 komponen utama yang mengacu pad Lampiran C telah terbentuk.

$$PC_1 = -0,072 x_1 - 0,382 x_2 - 0,101 x_3 + 0,355 x_4 + 0,330 x_5 + 0,406 x_6 + 0,239 x_7 + 0,439 x_8 + 0,016 x_9 - 0,332 x_{10} + 0,280 x_{11}$$

$$PC_2 = 0,490 x_1 - 0,358 x_2 + 0,398 x_3 - 0,073 x_4 - 0,118 x_5 + 0,133 x_6 - 0,054 x_7 - 0,135 x_8 - 0,426 x_9 - 0,360 x_{10} - 0,323 x_{11}$$

$$PC_3 = 0,403 x_1 + 0,095 x_2 + 0,3 x_3 + 0,401 x_4 + 0,355 x_5 - 0,362 x_6 + 0,305 x_7 + 0,019 x_8 + 0,402 x_9 - 0,126 x_{10} - 0,224 x_{11}$$

$$PC_4 = 0,189 x_1 + 0,155 x_2 + 0,038 x_3 - 0,180 x_4 - 0,358 x_5 + 0,126 x_6 + 0,675 x_7 + 0,340 x_8 - 0,248 x_9 + 0,353 x_{10} + 0,064 x_{11}$$

Berdasarkan persamaan tersebut, selanjutnya akan diperoleh nilai *score component* yang pada pengelompokan menggunakan metode ini, nilai tersebut digunakan sebagai faktor yang mengelompokkan provinsi di Indonesia dan dilampirkan pada Lampiran C.

4.2.3 Pengelompokan Provinsi dengan *C-means Clustering*

Analisis yang dilakukan selanjutnya adalah melakukan pengelompokan dengan metode *c-means clustering*. Berikut ini merupakan pemaparan hasil dari analisis pengelompokan dengan 4 *principal component* yang telah terbentuk berdasarkan variabel indikator kesejahteraan rakyat (Lampiran C) Sehingga pengelompokan yang didapatkan berdasarkan kombinasi linear dari variabel masing-masing dimensi.

4.2.3.1 Penentuan *Cluster Optimum*

Penentuan *cluster optimum* dilakukan dengan menghitung nilai *pseudo f-statistics* tertinggi. Penentuan *cluster optimum*

dilakukan dengan membentuk *cluster* yang berjumlah 2-4 *cluster* terlebih dahulu kemudian menentukan nilai *pseudo f-statistics* pada masing masing *cluster* untuk melihat *cluster* optimum yang terbentuk dan dijelaskan pada Tabel 4.5.

Tabel 4.5 Nilai *Pseudo F-statistics* metode *C-Means*

Jumlah Cluster	<i>Pseudo F-statistics</i>
2	7,560
3	7,018
4	5,734

Tabel 4.5 diatas merujuk pada Lampiran M yang menunjukkan bahwa nilai *pseudo f-statistics* untuk metode *c-means clustering* dengan jumlah 2 sampai 4 *cluster*. Nilai *pseudo f-statistics* terbesar merupakan jumlah pengelompokan yang optimum, sehingga memberikan kesimpulan bahwa jumlah *cluster* optimum untuk metode *c-means clustering* adalah 2 *cluster* karna nilai *pseudo f-statistics* nya paling tinggi dibandingkan dengan *cluster* lain yaitu sebesar 7,560.

4.2.3.2 Pengelompokan Dengan *Cluster* Optimum

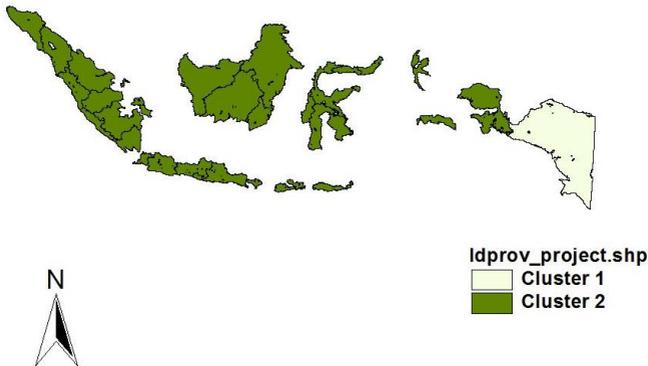
Pada pemaparan sebelumnya telah diketahui bahwa *cluster* yang terbentuk adalah sebanyak 2 *cluster* dan berikut ini merupakan hasil dari Tabel 4.5 merupakan hasil dari analisis pengklasteran dengan *c-means clustering*.

Tabel 4.6 Anggota *Cluster* Dengan *C-Means*

	Klaster 1	Klaster 2
Aceh	Gorontalo	Papua
Sumatera Utara	Maluku	
Sumatera Barat	Maluku Utara	
Riau	Papua Barat	
Jambi	Sulawesi Barat	
Sumatera Selatan	Sulawesi Tengah	
Bengkulu	Sulawesi Selatan	
Lampung	Sulawesi Tenggara	
Kep. Bangka Belitung	Nusa Tenggara Timur	
Kep. Riau	Sulawesi Utara	
DKI Jakarta	Kalimantan Timur	
DI Yogyakarta	Kalimantan Tengah	

Banten	Kalimantan Selatan
Bali	Kalimantan Barat
Nusa Tenggara Barat	Jawa Barat
Jawa Timur	Jawa Tengah

Tabel 4.6 diatas merupakan hasil analisis pengelompokan dengan *c-means clustering* dengan jumlah *cluster* optimum sebanyak 2 *cluster*. Anggota pada *cluster* ada sebanyak 32 provinsi dan hanya 1 provinsi yang berada pada *cluster* 2. Berikut hasil pengelompokan dengan *c-means clustering* secara visual.



Gambar 4.6 Peta Pengelompokan Dengan *C-Means Clustering*

Gambar 4.6 diatas diketahui bahwa Papua merupakan provinsi yang berada dalam *cluster* 2 sedangkan 32 provinsi lainnya berada pada *cluster* 1. 32 provinsi selain Papua merupakan anggota *cluster* 1 yang merupakan provinsi dengan karakteristik kombinasi linear yang hampir berdekatan atau homogen. Dan ini memberikan kesimpulan bahwa pengelompokan dengan kombinasi linear dari tiga dimensi dan membentuk 4 komponen utama kesejahteraan rakyat tersebut, Papua merupakan provinsi yang memiliki karakteristik berbeda dari 32 provinsi lainnya. Provinsi ini merupakan salah satu provinsi dengan *title* daerah tertinggal sehingga pemerintah sebaiknya memberikan perhatian khusus kepada provinsi ini dan melakukan pembangunan di provinsi ini untuk kesejahteraan rakyat yang ada di Papua.

4.2.4 Pengelompokan Provinsi dengan *Fuzzy C-means Clustering*

Dalam keseharian kondisi ketidakpastian lebih sering ditemui dalam mengklasifikasikan suatu kondisi. Misalnya usia tua atau muda adalah sesuatu yang tidak dapat dinyatakan dengan pasti, seperti usia berapa yang memisahkan antara tua dan muda. Lain halnya dengan jenis kelamin laki-laki atau perempuan yang sudah pasti. Demikian juga dengan kondisi temperatur dingin, normal atau panas, tidak ada batasan suhu yang pasti untuk membedakan kondisi tersebut. Dalam kasus-kasus dimana mendefinisikan keanggotaan suatu kelompok tidak jelas batasannya atau menjadi tidak pasti sehingga dalam menggunakan metode *c-means clustering* metode ini menjadi kurang sesuai .

Variabel yang digunakan dalam penelitian ini adalah indikator kesejahteraan rakyat dari berbagai dimensi namun masing-masing variabel tidak memiliki batasan yang jelas sehingga perlu dilakukan pengelompokan dengan *fuzzy c-means* yang memiliki sifat lebih fleksibel dan diharapkan memberikan hasil pengelompokan yang lebih homogen untuk masing-masing anggota dalam satu kelompok.

Sebelum dilakukan analisis pengelompokan, variabel direduksi menggunakan komponen utama atau *principal component analysis* yang mana hasil analisis komponen utama menghasilkan 4 komponen dan score component yang dihasilkan dilampirkan pada lampiran C. Pengelompokan dengan *fuzzy c-means* dengan analisis komponen utama akan dileburkan dengan merata-ratakan *score component* yang dihasilkan (Vijayarajan & Muttan, 2014) . Hasil rata-rata dari *score component* yang dihasilkan dilampirkan pada Lampiran C.

Rata-rata *score component* digunakan untuk mendapatkan probabilitas objek provinsi sesuai *cluster* nya atau derajat keanggotaannya dan dilampirkan dalam Lampiran D- Lampiran F. Derajat kenggotaan digunakan untuk mendapatkan matriks U yang merupakan matriks fungsi keanggotaan sesuai dengan fungsi

keanggotaan yang digunakan dan dilampirkan pada Lampiran G-Lampiran I.

4.2.3.1 Penentuan *Cluster Optimum*

Pengelompokan Indikator Kesejahteraan Rakyat di Indonesia tahun 2014 dengan metode *fuzzy c-means clustering* adalah dengan membentuk *cluster* yang berjumlah 2-4 *cluster* lalu kemudian menentukan nilai *pseudo f-statistics* pada masing masing *cluster* untuk melihat *cluster optimum* namun pengelompokan menggunakan metode *fuzzy c-means clustering*, memerlukan derajat keanggotaan sesuai dengan fungsi keanggotaan yang telah dipilih. Sehingga diperlukan fungsi keanggotaan (*membership function*) yaitu suatu kurva yang menunjukkan pemetaan titik-titik input data ke dalam nilai keanggotaan dimana penelitian ini menggunakan fungsi keanggotaan linier naik, linier turun dan segitiga. Fungsi keanggotaan ini akan membentuk matriks derajat keanggotaan yang berukuran *cluster x obyek* sebagai input saat melakukan pengelompokan menggunakan *fuzzy c-means clustering*.

Pengelompokan dengan metode FCM *clustering* akan dibentuk berdasarkan 3 fungsi keanggotaan, yaitu representasi linier naik, linier turun dan kurva segitiga dengan kelompok awal sama yaitu 2 hingga 4 *cluster*. Berikut ini merupakan nilai *pseudo f-statistics* untuk masing-masing fungsi keanggotaan untuk mendapatkan *cluster optimum* yang dipaparkan dalam Tabel 4.7

Tabel 4.7 Nilai *Pseudo F-statistics* metode *Fuzzy C-Means*

Jumlah <i>Cluster</i>	Fungsi Keanggotaan		
	Linear Naik	Linear Turun	Segitiga
2	5,493	5,493	5,493
3	3,616	3,616	3,616
4	2,372	2,372	2,372

Berdasarkan Tabel 4.7 tersebut diketahui bahwa nilai *pseudo f-statistics* tertinggi ada pada *cluster* 2 untuk masing masing fungsi keanggotaan dan ini memberikan informasi bahwa *cluster optimum* yang hanya mampu terbentuk adalah 2 *cluster*. Maka, hal ini menyimpulkan bahwa walaupun walaupun fungsi keanggotaan

yang digunakan berbeda, *cluster* optimum yang terbentuk menghasilkan hasil yang sama yaitu 2 *cluster*. Jadi, kesimpulannya dengan menggunakan metode *fuzzy c-means clustering* kelompok optimum yang digunakan adalah 2 *cluster*.

4.2.3.2 Pengelompokan Dengan *Cluster* Optimum

Berdasarkan nilai *pseudo f-statistics* untuk metode *fuzzy c-means clustering*, *cluster* optimum yang dihasilkan adalah sama dengan metode *c-means clustering* yaitu sebesar 2 *cluster*. Hasil pengelompokan provinsi di Indonesia berdasarkan indikator kesejahteraan rakyat dengan jumlah *cluster* optimum dengan masing-masing fungsi keanggotaan dipaparkan pada Tabel 4.8 berikut.

Tabel 4.8 Pengelompokan Optimum Dengan *Fuzzy C-Means Clustering* Fungsi Keanggotaan Linear Naik

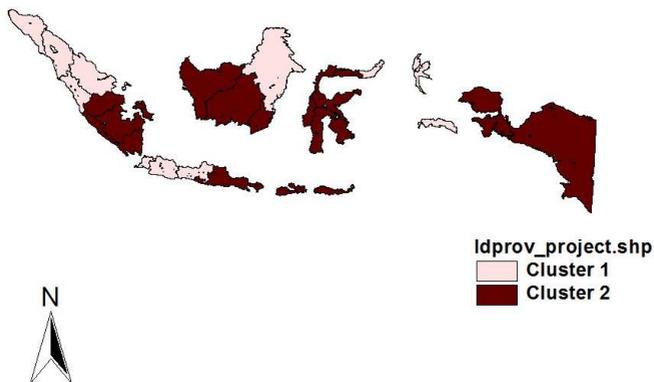
Klaster 1	Klaster 2	
Aceh	Jambi	Sulawesi Tengah
Sumatera Utara	Sumatera Selatan	Sulawesi Selatan
Sumatera Barat	Bengkulu	Sulawesi Tenggara
Riau	Lampung	Gorontalo
Kepulauan Riau	Kep. Bangka Belitung	Sulawesi Barat
DKI Jakarta	DI Yogyakarta	Papua Barat
Jawa Barat	Jawa Timur	Papua
Jawa Tengah	Bali	
Banten	Nusa Tenggara Barat	
Kalimantan Timur	Nusa Tenggara Timur	
Sulawesi Utara	Kalimantan Barat	
Maluku	Kalimantan Tengah	
Maluku Utara	Kalimantan Selatan	

Berdasarkan Tabel 4.8 diatas, didapatkan bahwa dengan representatif linear naik jumlah anggota *cluster* 1 adalah sebanyak 13 provinsi dan *cluster* 2 ada sebanyak 20 provinsi. Berikut ini merupakan hasil pengelompokan dengan metode *fuzzy c-means clustering* representatif linear turun dan segitiga.

Tabel 4.9 Pengelompokan Optimum Dengan *Fuzzy C-Means Clustering* Fungsi Keanggotaan Linear Turun dan Segitiga

	Klaster 1	Klaster 2
Jambi	Sulawesi Tengah	Aceh
Sumatera Selatan	Sulawesi Selatan	Sumatera Utara
Bengkulu	Sulawesi Tenggara	Sumatera Barat
Lampung	Gorontalo	Riau
Kep. Bangka Belitung	Sulawesi Barat	Kepulauan Riau
DI Yogyakarta	Papua Barat	DKI Jakarta
Jawa Timur	Papua	Jawa Barat
Bali		Jawa Tengah
Nusa Tenggara Barat		Banten
Nusa Tenggara Timur		Kalimantan Timur
Kalimantan Barat		Sulawesi Utara
Kalimantan Tengah		Maluku
Kalimantan Selatan		Maluku Utara

Berdasarkan pemaparan tersebut, didapatkan bahwa dengan representatif segitiga, linear naik dan linear turun menghasilkan hasil yang sama. pengelompokan dengan *fuzzy c-means clustering* dengan visualisasi gambar digambarkan pada gambar berikut ini.



Gambar 4.7 Peta Pengelompokan Dengan *C-Means Clustering*

Berdasarkan hasil yang telah dipaparkan tersebut maka didapatkan informasi mengenai karakteristik masing masing provinsi di masing masing *cluster*. Setelah diketahui anggota kelompok masing-masing *cluster* makadilanjutkan dengan

menganalisis karakteristik masing-masing *cluster* dan diketahui bahwa *cluster* 1 adalah *cluster* dengan provinsi yang memiliki karakteristik kesejahteraan rakyat yang kurang layak pada indikator kesehatan Imunisasi Anak, Sanitasi Layak dan Air Bersih dan Angka Kesakitan Malaria. Sedangkan *cluster* 2 memiliki karakteristik indikator kesejahteraan rakyat yang kurang layak pada indikator kesehatan di Pemberian ASI Eksklusif. Pada Indikator kesejahteraan rakyat Indikator pendidikan di seluruh dimensi baik AMH, APS dan TPT, anggota provinsi di *cluster* 1 memiliki kesejahteraan yang kurang layak namun pada *cluster* 2, derajat mutu pendidikan tersebut dapat dikatakan cukup baik dan memenuhi target pemerintah dalam tujuan pembangunan nasional. Pada indikator kesejahteraan rakyat dimensi ekonomi, anggota di *cluster* 2 merupakan anggota provinsi yang membutuhkan perhatian pada tingkat pengangguran terbuka karena cukup tinggi pada provinsi tersebut sedangkan pada *cluster* 1, anggota provinsi tersebut merupakan provinsi dengan indikator kesejahteraan rakyat dimenasi ekonomi yang bermasalah pada jumlah penduduk miskin yang cukup tinggi.

4.3 Perbandingan Hasil Pengelompokan dengan *C-means Clustering* dan *Fuzzy C-means Clustering*

Metode yang digunakan dalam penelitian ini adalah menggunakan metode pengelompokan dengan *c-means* dan *fuzzy c-means*. Dalam menentukan *cluster* optimum, digunakan nilai *pseudo f-statistics*. Pengelompokan berdasarkan masing-masing metode telah dipaparkan di sub bab sebelumnya. Tahapan selanjutnya adalah membandingkan kedua metode manakah yang lebih baik dengan menggunakan nilai *icd-rate* masing masing metode. *Icdrate* merupakan nilai yang menunjukkan tingkat dispersi antar anggota dalam satu *cluster*. Hasil pengelompokan dikatakan baik jika suatu kelompok memiliki nilai *icdrate* yang semakin kecil. Hal ini dikarenakan perbedaan dalam kelompok tersebut semakin kecil atau sering dinamakan semakin homogen.

Selain itu perlu menentukan nilai dari SSW (*Sum Square Within*) yang merupakan nilai jarak total antar anggota dalam satu *cluster* dengan pusat *cluster* dan SSB (*Sum Square Between*) yang merupakan nilai jarak total antar pusat *cluster*. Berikut merupakan nilai *icdrate* dengan pengelompokan sebanyak 2 *cluster* untuk metode CM dan FCM *clustering* menggunakan 3 fungsi keanggotaan yakni representatif linear naik, linear turun, dan segitiga. Namun, berdasarkan pamaran diatas, nilai pseudo *f-statistics* dan anggota *cluster* yang dihasilkan ketiga kurva adalah sama sehingga dapat dipastikan nilai *icd-rate* yang akan bernilai sama untuk metode *fuzzy c-means clustering*.

Tabel 4.10 *Icd Rate C-Means dan Fuzzy C-means Clustering*

Nilai	<i>C-Means</i>	<i>Fuzzy C-means</i>
SSW	282,988	299,014
SSB	69,011	52,985
<i>Icd-Rate</i>	0,803	0,849

Tabel diatas menunjukkan bahwa metode dengan *Fuzzy c-means clustering* memiliki nilai *icdrate* yang paling rendah walau tidak berbeda dengan nilai *icd-rate fuzzy c-means* dan ini menunjukkan bahwa dispersi yang terjadi dengan metode pengelompokan *c-means clustering* cukup kecil dan metode ini merupakan metode yang lebih baik dibandingkan dengan metode *fuzzy c-means*.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Hasil dari analisis dan pembahasan mengenai pengelompokan provinsi di di Indonesia berdasarkan indikator kesejahteraan rakyat dapat dirangkum dalam kesimpulan berikut ini.

1. Gambaran umum kesejahteraan rakyat Indonesia pada masing masing provinsi di Indonesia melalui analisis statistika deskriptif didapatkan hasil untuk dimensi kesehatan terdapat 13 provinsi yang berada dibawah rata-rata nasional untuk ASI Eksklusif dan NTB merupakan provinsi dengan ASI Eksklusif tertinggi. 8 provinsi diatas rata-rata untuk Angka Kesakitan Malaria dan Papua adalah provinsi dengan Angka Kesakitan Malaria tertinggi di Indonesia. Sulawesi Utara adalah provinsi dengan Angka Imunisasi Anak tertinggi dan terdapat 9 provinsi dibawah rata-rata nasional untuk Angka Imunisasi Anak. Variabel Persentase Rumah Tangga Akses Air Minum Layak dan Sanitasi Layak terdapat 9 provinsi yang berada dibawah rata-rata nasional dan NTT merupakan akses sanitasi terendah di Indonesia. Dimensi pendidikan memberikan gambaran bahwa Papua merupakan provinsi dengan indikator pensisikan terburuk di Indonesia dikarenakan kurangnya tenaga pengajar di provinsi tersebut. Dimensi bidang ekonomi memberikan gambaran Bangka Belitung adalah provinsi dengan tingkat ketimpangan terbesar di Indonesia, Penduduk miskin di Indonesia terbesar berada di Papua dan Maluku merupakan provinsi dengan Tingkat Pengangguran Tertinggi di Indonesia.
2. Pengelompokan dengan *c-means* dan *fuzzy c-means* memenuhi asumsi normal multivariat, kecukupan data dan korelasi antar variabel yang signifikan. Variabel yang cukup banyak dan memiliki korelasi yang cukup kuat maka dilakukan reduksi variabel dengan PCA dan terbentuk 4 PC. Cluster optimum yang dihasilkan kedua metode adalah 2

cluster yang dilihat dari nilai *pseudo f-statistics* sebesar 7,560 dengan 32 provinsi berada pada *cluster* 1 dan hanya Papua yang berada di *cluster* 2 serta karakteristik yang terbentuk adalah provinsi Papua merupakan provinsi yang kesejahteraan rakyat nya kurang layak dari segala dimensi. Dengan *fuzzy c-means clustering* diketahui bahwa nilai *pseudo f-statistics* untuk semua kurva representatif adalah sama yaitu sebesar 5,493. Anggota *cluster* yang terbentuk dari pengklasteran tersebut terdapat 13 provinsi yang berada pada *cluster* 1 dan 20 lainnya pada *cluster* 2.

3. Perbandingan kedua metode *c-means* dan *fuzzy c-means* memberikan hasil bahwa *c-means* memberikan hasil lebih baik dibandingkan dengan *c-means* yang dilihat dari nilai *icd rate* sebesar 0,803 untuk metode *c-means* yang menunjukkan tingkat dispersi yang rendah.

5.2 Saran

Saran yang dapat diberikan untuk penelitian selajutnya adalah memperbanyak variabel yang berkaitan dengan kesejahteraan rakyat terkait dengan pembangunan nasional dan MDG's dan juga menambahkan fungsi *fuzzy* untuk mendapatkan hasil yang lebih baik. Saran yang dapat diberikan untuk BAPPENAS adalah sebaiknya lebih fokus dalam pembangunann nasional sesuai dengan karakteristik permasalahan sosial yang berkaitan dengan kesejahteraan rakyat. Sehingga pembangunan dapat berjalan lebih baik dan dana yang dikeluarkan untuk pembangunan nasional jadi lebih efisien dimana dalam penelitian ini memberikan informasi bahwa papua adalah provinsi yang membutuhkan perhatian dari pemangku kepentingan di negara ini.

DAFTAR PUSTAKA

- Ahnaf, A., Purnama, H., Syahbudin, I., & dkk. (2012). *Pemantauan Perkembangan Kesejahteraan Rakyat*. Jakarta: Badan Pusat Statistik.
- Anwar, A. (2013). *Masalah Jumlah Penduduk Indonesia*. Padang, Sumbar: Mahasiswa Fakultas Sastra, Universitas Andalas.
- Badan Pusat Statistik. (2014). *Indeks Pembangunan Manusia Metode Baru*. Jakarta: Badan Pusat Statistik.
- Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The Fuzzy C-Means Clustering Algorithm. *Computers and Geoscience* printed in USA, vol 10 no.2-3 pp 191-203.
- Johnson, R. A., & Wichern, D. W. (2007). *Nonhierarchical Clustering Methods*. In *Applied Multivariate Statistical Analysis*. 7th Edition (p. 696). USA: Prentice Hall.
- Hasbullah, J. (2012). *Tangguh Dengan Statistika*. Bandung: Nuansa Cendekia.
- Karti, H. S. (2013). *Pengelompokan Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Pendidikan SMA/SMK/MA dengan Metode C-Means dan Fuzzy C-Means*. SAINS DAN SENI POMITS, Vol. 2, No.2.
- Kusumadewi, S., & Hartati, S. (2006). *Neuro Fuzzy: Integrasi Sistem Fuzzy & Jaringan Syaraf*. Yogyakarta: Graha Ilmu.
- Lailiyah. (2011). *pengelompokan kabupaten/kota di Jawa Timur berdasarkan kesamaan nilai faktor-faktor yang mempengaruhi tingkat pengangguran terbuka dengan metode hierarchi dan nonhierarchi*. In T. Akhir. Surabaya: Jurusan Statistika FMIPA ITS.
- Mardianto, S. (2012). *Kemiskinan di Indonesia*. Aceh: Universitas Syiah Kuala.

- Mingoti, S. A., & Lima, J. O. (2006). Comparing SOM Neural Network with Fuzzy C-Means,. *European Journal of Operational Research*, 1742–1759.
- Orpin, & Kostlev. (2006). Toward a statistically valid method of textural sea floor characterization of benthic habitats. *Marine Geology*.
- Rencher, A. C. (2002). *Method of Multivariate Analysis*. In S. Edition, *Wiley Interscience* (pp. 410-412). Canada: A John Wiley & Sons, Inc. Publication.
- Santosa, B. (2007). *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Sukim. (2011). Studi tentang Metode C-Means Cluster dan Fuzzy C-Means Cluster serta Aplikasinya pada Kasus Pengelompokan Desa/Kelurahan berdasarkan status Ketertinggalan. In T. Akhir. Surabaya: Jurusan Statistika FMIPA ITS.
- Timm, N. H. (2002). *Applied Multivariate Analysis*. United State of America: Springer.
- Vijayarajan,R.,&Muttan.S. (2014). Fuzzy C-means clustering based principal component averaging fusion. *International Journal of Fuzzy System*

LAMPIRAN

LAMPIRAN A DATA PENELITIAN

PROVINSI	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
Aceh	55.4	0.16	87.8	61.1	29.54	97.42	80.89	9.45	0.34	16.98	9.02
Sumatera Utara	37.6	0.69	96.8	67.13	67.64	98.57	75.78	9.44	0.31	9.85	6.23
Sumatera Barat	73.6	0.18	92.3	61.2	40.07	98.44	81.97	8.94	0.33	6.89	6.5
Riau	55.7	0.13	94.5	73.8	45.73	98.75	75.3	9	0.38	7.99	6.56
Jambi	64.3	0.84	97.2	60.78	55.25	97.77	70.41	8.5	0.34	8.39	5.08
Sumsel	64.5	0.31	94.3	59.14	58.24	98.14	67.84	8.31	0.38	13.62	4.96
Bengkulu	78.5	2.17	95.9	35.17	39.85	97.52	77.92	8.85	0.36	17.09	3.47
Lampung	63.7	0.55	95.9	51.48	34.67	96.54	68.75	8.05	0.33	14.31	4.79
Kep. Belitung	54.9	0.86	98.5	62.53	80.13	97.6	65.78	7.87	0.3	4.97	5.14
Kep. Riau	50.3	0.41	90.4	83.27	59.68	98.71	81.57	9.92	0.44	6.4	6.69
DKI Jakarta	67.1	0	93.6	91.23	86.81	99.54	70.23	10.98	0.44	4.09	8.47
Jawa Barat	21.8	0.01	87.3	63.92	63.25	97.96	65.48	8.39	0.4	9.18	8.45
Jawa Tengah	60	0.05	99.5	71.41	69.07	92.98	67.54	7.62	0.39	13.58	5.68
DI Yogyakarta	70.8	0.012	98.3	77.7	82.54	94.44	86.44	9.67	0.43	14.55	3.33
Jawa Timur	74	0.01	92.5	74.82	65.13	91.36	70.25	7.77	0.4	12.28	4.19
Banten	65	0	95.8	67.76	68.09	97.24	66.25	8.8	0.42	5.51	9.07
Bali	72.2	0	98.9	93.22	74.35	92.56	81.59	8.88	0.44	4.76	1.9
NTB	84.7	0.78	88.5	63.94	59.35	86.96	75.68	7.58	0.39	17.05	5.75
NTT	77.4	12.81	92.5	52.65	12.77	91.18	73.96	7.45	0.35	19.6	3.26

LAMPIRAN A
DATA PENELITIAN((LANJUTAN)

PROVINSI	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
Kalimantan Barat	49.5	0.17	95	60.91	44.97	92.3	66.48	7.47	0.4	8.07	4.04
Kaltmantan Tengah	40.7	1.32	93.9	58.73	29.48	98.82	65.84	8.48	0.36	6.07	3.24
Kalimantan Selatan	67.2	1.35	95.1	57.67	19.36	98.19	67.18	8.21	0,33	4.81	3.8
Kalimantan Timur	67.8	0.32	91.2	75.11	66.95	98.59	80.5	9.62	0.36	6.31	7.38
Sulawesi Utara	38.5	0.94	43.2	70.16	68.17	99.6	71.98	9.26	0.44	8.26	7.54
Sulawesi Tengah	56.4	0.8	89.3	58.26	54.51	97.08	73.64	8.45	0.35	13.61	3.68
Sulawesi Selatan	69.3	0.1	94.3	68.68	72.97	91.26	69.38	8.32	0.45	9.54	5.08
Sulawesi Tenggara	65.5	0.46	91.3	73.74	66.5	94.03	72.25	8.83	0.4	12.7	4.43
Gorontalo	58.8	0.84	94.3	66.18	59.83	97.9	68.69	7.69	0.45	17.41	4.18
Sulawesi Barat	65	0.25	83.9	50.88	59.48	92.27	66.97	7.55	0.38	12.05	2.08
Maluku	45.3	6	94	63.01	61.7	98.77	77.48	9.62	0.33	18.44	10.51
Maluku Utara	62.2	3.32	89.7	61.98	58.97	98.36	74.83	8.93	0.32	7.41	5.29
Papua Barat	27.3	10.85	83.5	68.8	65.8	96.75	79.87	9.56	0.41	26.26	5.02
Papua	52.2	29.57	96.2	49.42	24.78	70.78	61.63	6.32	0.46	27.8	3.44

LAMPIRAN B
DATA PENELITIAN STANDARISASI Z-Score

PROVINSI	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
Aceh	-0.27	-0.38	-0.41	-0.36	-1.44	0.37	1.32	0.91	-0.19	0.88	1.72
Sumatera Utara	-1.50	-0.28	0.53	0.15	0.64	0.58	0.49	0.90	-0.48	-0.31	0.39
Sumatera Barat	0.99	-0.37	0.06	-0.35	-0.86	0.56	1.50	0.37	-0.29	-0.80	0.52
Riau	-0.25	-0.38	0.29	0.73	-0.55	0.62	0.42	0.43	0.19	-0.62	0.55
Jambi	0.34	-0.26	0.58	-0.39	-0.04	0.44	-0.38	-0.11	-0.19	-0.55	-0.15

LAMPIRAN B**DATA PENELITIAN STANDARISASI Z-Score (LANJUTAN)**

Sumsel	0.36	-0.35	0.27	-0.53	0.13	0.50	-0.79	-0.31	0.19	0.32	-0.21
Bengkulu	1.33	-0.02	0.44	-2.58	-0.87	0.39	0.84	0.27	0.00	0.90	-0.92
Lampung	0.30	-0.31	0.44	-1.19	-1.16	0.21	-0.65	-0.59	-0.29	0.44	-0.29
Kep. Belitung	-0.30	-0.25	0.71	-0.24	1.32	0.41	-1.13	-0.78	-0.58	-1.12	-0.12
Kep. Riau	-0.62	-0.33	-0.13	1.54	0.20	0.61	1.43	1.42	0.77	-0.88	0.61
DKI Jakarta	0.54	-0.41	0.20	2.22	1.68	0.76	-0.41	2.56	0.77	-1.26	1.46
Jawa Barat	-2.59	-0.40	-0.46	-0.12	0.40	0.47	-1.18	-0.23	0.39	-0.42	1.45
Jawa Tengah	0.05	-0.40	0.82	0.52	0.72	-0.45	-0.84	-1.05	0.29	0.31	0.13
DI Yogyakarta	0.79	-0.40	0.69	1.06	1.45	-0.18	2.22	1.15	0.68	0.48	-0.98
Jawa Timur	1.02	-0.40	0.09	0.81	0.50	-0.75	-0.40	-0.89	0.39	0.10	-0.58
Banten	0.39	-0.41	0.43	0.21	0.66	0.34	-1.05	0.22	0.58	-1.03	1.74
Bali	0.89	-0.41	0.75	2.39	1.00	-0.53	1.43	0.30	0.77	-1.15	-1.66
NTB	1.75	-0.27	-0.33	-0.12	0.19	-1.56	0.48	-1.10	0.29	0.89	0.17
NTT	1.25	1.84	0.09	-1.09	-2.35	-0.78	0.20	-1.24	-0.10	1.32	-1.02
Kalimantan Barat	-0.68	-0.38	0.35	-0.38	-0.60	-0.57	-1.01	-1.21	0.39	-0.60	-0.65
Kaltantan Tengah	-1.29	-0.17	0.23	-0.56	-1.44	0.63	-1.12	-0.13	0.00	-0.94	-1.03
Kalimantan Selatan	0.55	-0.17	0.36	-0.66	-1.99	0.51	-0.90	-0.42	-3.49	-1.14	-0.76
Kalimantan Timur	0.59	-0.35	-0.05	0.84	0.60	0.59	1.26	1.10	0.00	-0.90	0.94
Sulawesi Utara	-1.44	-0.24	-5.06	0.41	0.67	0.77	-0.12	0.71	0.77	-0.57	1.02
Sulawesi Tengah	-0.20	-0.26	-0.25	-0.61	-0.08	0.31	0.15	-0.16	-0.10	0.32	-0.82
Sulawesi Selatan	0.69	-0.39	0.27	0.29	0.93	-0.77	-0.54	-0.30	0.87	-0.36	-0.15
Sulawesi Tenggara	0.43	-0.32	-0.04	0.72	0.58	-0.25	-0.08	0.25	-3.49	0.17	-0.46

LAMPIRAN B

DATA PENELITIAN STANDARISASI Z-Score (LANJUTAN)

Gorontalo	-0.04	-0.26	0.27	0.07	0.21	0.46	-0.66	-0.98	0.87	0.95	-0.58
Sulawesi Barat	0.39	-0.36	-0.81	-1.24	0.19	-0.58	-0.93	-1.13	0.19	0.06	-1.58
Maluku	-0.97	0.65	0.24	-0.20	0.31	0.62	0.77	1.10	-0.29	1.12	2.43
Maluku Utara	0.20	0.18	-0.21	-0.29	0.17	0.55	0.34	0.36	-0.39	-0.71	-0.05
Papua Barat	-2.21	1.50	-0.85	0.30	0.54	0.25	1.16	1.03	0.48	2.42	-0.18
Papua	-0.49	4.78	0.47	-1.36	-1.69	-4.54	-1.80	-2.45	0.97	2.68	-0.93

Keterangan :

x_1 : Pemberian ASI Eksklusif (%)

x_2 : Angka Kesakitan Malaria

x_3 : Imunisasi Anak Sekolah (%)

x_4 : Rumah Tangga Air Minum Layak (%)

x_5 : Rumah Tangga Sanitasi Layak (%)

x_6 : Angka Melek Huruf

x_7 : Angka Partisipasi Sekolah

x_8 : Angka Lama Sekolah

x_9 : Indeks Gini

x_{10} : Penduduk Miskin (%)

x_{11} : Tingkat Pengangguran Terbuka

LAMPIRAN C
PRINCIPLE COMPONENT ANALYSIS

No	Provinsi	PC1	PC2	PC3	PC4	Mean PC
1	Aceh	0.6572	-0.8986	-0.9871	2.1735	0.2362
2	Sumatera Utara	1.3840	-0.3887	-0.5491	0.1872	0.1583
3	Sumatera Barat	0.8077	0.9567	-0.1401	1.5325	0.7892
4	Riau	1.1092	0.1063	-0.1664	0.2468	0.3240
5	Jambi	0.0415	0.9457	-0.2698	-0.261	0.1140
6	Sumatera Selatan	-0.3497	0.4560	-0.2395	-0.452	-0.146
7	Bengkulu	-1.4129	1.0696	-0.2432	2.0092	0.3556
8	Lampung	-1.3107	0.8614	-0.9419	0.2513	-0.285
9	Bangka Belitung	0.2742	0.9983	-0.4141	-1.734	-0.219
10	Kepulauan Riau	2.4877	-0.7755	0.6715	0.5371	0.7302
11	DKI Jakarta	3.6148	-0.4415	1.1993	-0.810	0.8905
12	Jawa Barat	0.8367	-1.6722	-1.8823	-1.653	-1.092
13	Jawa Tengah	-0.4222	0.2168	0.7012	-1.307	-0.202
14	DI Yogyakarta	1.4228	0.1193	2.7437	1.2093	1.3738
15	Jawa Timur	-0.4529	0.5865	1.3278	-0.960	0.1251
16	Banten	1.1944	0.0464	-0.1574	-1.237	-0.038
17	Bali	1.3843	1.0232	3.0312	-0.356	1.270
18	NTB	-1.2162	0.2277	1.4819	0.2388	0.1830
19	NTT	-3.4967	0.2939	0.1102	1.6014	-0.372
20	Kalimantan Barat	-1.1583	0.4421	-0.4693	-1.411	-0.649
21	Kalimantan Tengah	-0.5855	0.5656	-1.6629	-0.758	-0.610
22	Kalimantan Selatan	-0.9788	3.0698	-2.5023	0.6485	0.0593
23	Kalimantan Timur	2.1735	0.1398	0.6056	0.7288	0.9119
24	Sulawesi Utara	2.1565	-3.1823	-2.0095	-0.887	-0.980
25	Sulawesi Tengah	-0.3461	0.1968	-0.3369	0.2162	-0.067
26	Sulawesi Selatan	-0.0047	0.2307	1.2124	-1.221	0.0541
27	Sulawesi Tenggara	0.2904	1.7021	-0.5811	0.5876	0.4998
28	Gorontalo	-0.6945	-0.1436	0.3633	-0.759	-0.308
29	Sulawesi Barat	-1.5966	0.5993	-0.2054	-1.073	-0.569
30	Maluku	1.0524	-1.8044	-0.7117	1.443	-0.005
31	Maluku Utara	0.5661	0.3981	-0.3416	0.3091	0.2329
32	Papua Barat	-0.0594	-3.2326	0.1278	1.4230	-0.435
33	Papua	-7.3680	-2.7130	1.2358	-0.456	-2.325

LAMPIRAN D
PRINCIPLE COMPONENT ANALYSIS & DERAJAT
KEANGGOTAAN LINEAR NAIK

No	Provinsi	μPC	2C	3C	4C
1	Aceh	0.692	1	1	2
2	Sumatera Utara	0.671	1	1	2
3	Sumatera Barat	0.842	1	1	1
4	Riau	0.716	1	1	2
5	Jambi	0.659	1	2	2
6	Sumatera Selatan	0.589	1	2	2
7	Bengkulu	0.725	1	1	2
8	Lampung	0.552	1	2	2
9	Bangka Belitung	0.569	1	2	2
10	Kepulauan Riau	0.826	1	1	1
11	DKI Jakarta	0.869	1	1	1
12	Jawa Barat	0.333	2	2	3
13	Jawa Tengah	0.574	1	2	2
14	DI Yogyakarta	1.000	1	1	1
15	Jawa Timur	0.662	1	1	2
16	Banten	0.618	1	2	2
17	Bali	0.972	1	1	1
18	NTB	0.678	1	1	2
19	NTT	0.528	1	2	2
20	Kalimantan Barat	0.453	2	2	3
21	Kalimantan Tengah	0.464	2	2	3
22	Kalimantan Selatan	0.645	1	2	2
23	Kalimantan Timur	0.875	1	1	1
24	Sulawesi Utara	0.364	2	2	3
25	Sulawesi Tengah	0.610	1	2	2
26	Sulawesi Selatan	0.643	1	2	2
27	Sulawesi Tenggara	0.764	1	1	1
28	Gorontalo	0.545	1	2	2
29	Sulawesi Barat	0.475	2	2	3
30	Maluku	0.627	1	2	2
31	Maluku Utara	0.692	1	1	2
32	Papua Barat	0.511	1	2	2
33	Papua	0.000	2	3	4

LAMPIRAN E
PRINCIPLE COMPONENT ANALYSIS & DERAJAT
KEANGGOTAAN FUNGSI KEANGGOTAAN LINEAR
TURUN

No	Provinsi	μPC	2C	3C	4C
1	Aceh	0.3075	2	3	3
2	Sumatera Utara	0.3286	2	3	3
3	Sumatera Barat	0.1580	2	3	4
4	Riau	0.2838	2	3	3
5	Jambi	0.3406	2	2	3
6	Sumatera Selatan	0.4109	2	2	3
7	Bengkulu	0.2752	2	3	3
8	Lampung	0.4484	2	2	3
9	Bangka Belitung	0.4306	2	2	3
10	Kepulauan Riau	0.1740	2	3	4
11	DKI Jakarta	0.1306	2	3	4
12	Jawa Barat	0.6668	1	1	2
13	Jawa Tengah	0.4262	2	2	3
14	DI Yogyakarta	0.0000	2	3	4
15	Jawa Timur	0.3376	2	2	3
16	Banten	0.3818	2	2	3
17	Bali	0.0279	2	3	4
18	NTB	0.3219	2	3	3
19	NTT	0.4722	2	2	3
20	Kalimantan Barat	0.5469	1	2	2
21	Kalimantan Tengah	0.5364	1	2	2
22	Kalimantan Selatan	0.3554	2	2	3
23	Kalimantan Timur	0.1249	2	3	4
24	Sulawesi Utara	0.6365	1	2	2
25	Sulawesi Tengah	0.3896	2	2	3
26	Sulawesi Selatan	0.3567	2	2	3
27	Sulawesi Tenggara	0.2363	2	3	4
28	Gorontalo	0.4548	2	2	3
29	Sulawesi Barat	0.5252	1	2	2
30	Maluku	0.3728	2	2	3
31	Maluku Utara	0.3084	2	3	3
32	Papua Barat	0.4891	2	2	3
33	Papua	1.0000	1	1	1

LAMPIRAN F
DATA PRINCIPLE COMPONENT ANALYSIS & DERAJAT
KEANGGOTAAN SEGITIGA

No	Provinsi	μPC	2C	3C	4C
1	Aceh	0.8620	1	1	1
2	Sumatera Utara	0.9210	1	1	1
3	Sumatera Barat	0.4430	2	2	3
4	Riau	0.7955	1	1	1
5	Jambi	0.9546	1	1	1
6	Sumatera Selatan	0.9158	1	1	1
7	Bengkulu	0.7715	1	1	1
8	Lampung	0.8575	1	1	1
9	Bangka Belitung	0.8852	1	1	1
10	Kepulauan Riau	0.4877	2	2	3
11	DKI Jakarta	0.3662	2	2	3
12	Jawa Barat	0.5180	1	2	2
13	Jawa Tengah	0.8920	1	1	1
14	DI Yogyakarta	0.0000	2	3	4
15	Jawa Timur	0.9462	1	1	1
16	Banten	0.9610	1	1	1
17	Bali	0.0782	2	3	4
18	NTB	0.9023	1	1	1
19	NTT	0.8206	1	1	1
20	Kalimantan Barat	0.7044	1	1	2
21	Kalimantan Tengah	0.7207	1	1	2
22	Kalimantan Selatan	0.9961	1	1	1
23	Kalimantan Timur	0.3500	2	2	3
24	Sulawesi Utara	0.5651	1	2	2
25	Sulawesi Tengah	0.9489	1	1	1
26	Sulawesi Selatan	1.0000	1	1	1
27	Sulawesi Tenggara	0.6623	1	1	2
28	Gorontalo	0.8475	1	1	1
29	Sulawesi Barat	0.7381	1	1	2
30	Maluku	0.9751	1	1	1
31	Maluku Utara	0.8645	1	1	1
32	Papua Barat	0.7943	1	1	1
33	Papua	0.0000	2	3	4

LAMPIRAN G
MATRIKS U LINEAR NAIK

2 Cluster		3 Cluster			4 Cluster			
1	0	1	0	0	0	1	0	0
1	0	1	0	0	0	1	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	0	1	0	0
1	0	0	1	0	0	1	0	0
1	0	0	1	0	0	1	0	0
1	0	1	0	0	0	1	0	0
1	0	0	1	0	0	1	0	0
1	0	0	1	0	0	1	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	1	0	0	0
0	1	0	1	0	0	0	1	0
1	0	0	1	0	0	1	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	0	1	0	0
1	0	0	1	0	0	1	0	0
1	0	0	1	0	0	1	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	0	1	0	0
1	0	0	1	0	0	1	0	0
0	1	0	1	0	0	0	1	0
0	1	0	1	0	0	0	1	0
1	0	0	1	0	0	1	0	0
1	0	1	0	0	1	0	0	0
0	1	0	1	0	0	0	1	0
1	0	0	1	0	0	1	0	0
1	0	0	1	0	0	1	0	0
1	0	1	0	0	1	0	0	0
0	1	0	1	0	0	0	1	0
1	0	0	1	0	0	1	0	0
1	0	0	1	0	0	1	0	0
1	0	0	1	0	0	1	0	0
0	1	0	0	1	0	0	0	1

LAMPIRAN H
MATRIKS U LINEAR TURUN

2 Cluster		3 Cluster			4 Cluster			
0	1	0	0	1	0	0	1	0
0	1	0	0	1	0	0	1	0
0	1	0	0	1	0	0	0	1
0	1	0	0	1	0	0	1	0
0	1	0	1	0	0	0	1	0
0	1	0	1	0	0	0	1	0
0	1	0	0	1	0	0	1	0
0	1	0	1	0	0	0	1	0
0	1	0	1	0	0	0	1	0
0	1	0	0	1	0	0	0	1
0	1	0	0	1	0	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	1	0	0	0	1	0
0	1	0	0	1	0	0	0	1
0	1	0	1	0	0	0	1	0
0	1	0	1	0	0	0	1	0
0	1	0	0	1	0	0	0	1
0	1	0	0	1	0	0	1	0
0	1	0	0	1	0	0	0	1
0	1	0	0	1	0	0	0	1
0	1	0	1	0	0	0	1	0
0	1	0	1	0	0	0	1	0
0	1	0	0	1	0	0	0	1
0	1	0	1	0	0	0	1	0
0	1	0	1	0	0	0	1	0
0	1	0	0	1	0	0	0	1
0	1	0	1	0	0	0	1	0
0	1	0	0	1	0	0	0	1
0	1	0	1	0	0	0	1	0
0	1	0	0	1	0	0	1	0
0	1	0	1	0	0	0	1	0
0	1	0	1	0	0	0	1	0
1	0	1	0	0	1	0	0	0

LAMPIRAN I
MATRIKS U SEGITIGA

2 Cluster		3 Cluster			4 Cluster			
1	0	1	0	0	1	0	0	0
1	0	1	0	0	1	0	0	0
0	1	0	1	0	0	0	1	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	1	0	0	0
0	1	0	1	0	0	0	1	0
0	1	0	1	0	0	0	1	0
1	0	0	1	0	0	1	0	0
1	0	1	0	0	1	0	0	0
0	1	0	0	1	0	0	0	1
1	0	1	0	0	1	0	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	0	1	0	0
1	0	1	0	0	0	1	0	0
1	0	1	0	0	1	0	0	0
0	1	0	1	0	0	0	1	0
1	0	0	1	0	0	1	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	0	1	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	1	0	0	0
1	0	1	0	0	1	0	0	0
0	1	0	0	1	0	0	0	1

LAMPIRAN J

MACRO MINITAB NORMAL MULTIVARIAT

```

macro
qq x.1-x.p
mconstant i n p t chis
mcolumn d x.1-x.p dd pi q ss tt
mmatrix s sinv ma mb mc md
let n=count(x.1)
cova x.1-x.p s
invert s sinv
do i=1:p
  let x.i=x.i-mean(x.i)
enddo
do i=1:n
  copy x.1-x.p ma;
  use i.
  transpose ma mb
  multiply ma sinv mc
  multiply mc mb md
  copy md tt
  let t=tt(1)
  let d(i)=t
enddo
set pi
  1:n
end
let pi=(pi-0.5)/n
sort d dd
invcdf pi q;
chis p.
plot q*dd
invcdf 0.5 chis;
chis p.
let ss=dd<chis
let t=sum(ss)/n
print t
endmacro

```

LAMPIRAN K

SYNTAX MATLAB FUZZY C-MEANS

```

clear;clc;
%program K-means data asli
X=load('E:\data.txt');
klaster=4;
w=2;
P(1)=1;
delta=1;
jarak=1;
t=1;
MaxItr=100;
eps=10^(-6);
[n,m]=size(X);
x=transpose(X);
u=load('E:\matriku.txt');
U=transpose(u);

%Normalisasi matrix partisi awal
Q=sum(U);
for k=1:n;
U(:,k)=U(:,k)/Q(k);
end;

%proses iterasi
while (t<MaxItr)&&(delta>eps);

%menghitung pusat klaster
for i=1:klaster
Tk=0;
    for k=1:n;
Tk=Tk+U(i,k).^w;
    end;
    for j=1:m;
Tj=0;
        for k=1:n;
Tj=Tj+(U(i,k)^w)*x(j,k);
        end;

```

```

                                v(j,i)=Tj/Tk;
                                end;
end
%jarak antara matrix yang akan dicluster dengan pusat cluster

for i=1:klaster;
    for k=1:n;
        d(i,k)=0;
        for j=1:m;
            d(i,k)=(d(i,k)+((x(j,k)-v(j,i)).^2)).^0.5;
        end
    end;
end;
end;
%menghitung fungsi obyektif
P(t+1)=0;
for k=1:n;
    for i=1:klaster;
        P(t+1)=P(t+1)+(d(i,k)^2)*(U(i,k)^w);
    end;
end;

%mengupdate matriks partisi
U_lama=U;
U=zeros(klaster,n);
for k=1:n;
    for i=1:klaster;
        if d(i,k)~=0;
            for j=1:klaster;
                U(i,k)=U(i,k)+((d(i,k)^2)/(d(j,k)^2))^(1/(w-1));
            end;
            U(i,k)=1/U(i,k);
            else
                U(i,k)=0;
            end;
        end;
    end;
end;
%hitung eror
delta=max(max(abs(U(i,k)-U_lama)));

%naikan iterasi
t=t+1;
end;

```

```

maxU=max(U)
index1=find(U(1,:)==maxU)
index2=find(U(2,:)==maxU)
index3=find(U(3,:)==maxU)
index4=find(U(4,:)==maxU)

```

LAMPIRAN L

SYNTAX MATLAB *PSEUDO F STATISTICS*

```

%menghitung pseudo f statistics
p=load('C:\KULIAH\Mata Kuliah\Semester 8\TA\Olahan Revisi\Tcluster.txt');
n=length(p);
x=p(:,1:8);
for j=1:4
k=max(p(:,j+8));
m=mean(x);
rm=repmat(m,n,1);
dm=(x-rm).^2;
jum=sum(dm);
sst=sum(jum);
ssw=0;
for i=1:k
anggota=find([p(1:n,j+8)]==i);
dataC=x(anggota,:);
na=size(dataC,1);
rata=mean(dataC);
kurang=dataC-repmat(rata,na,1);
total=sum(sum(kurang.^2,2));
ssw=ssw+total;
end
ssb=(sst-ssw);
rsq=ssb/sst;
msb=rsq/(k-1);
msw=(1-rsq)/(n-k);
pf(j)=(msb/msw);
icdrate(j)=(1-rsq);
filename='pf.xlsx';
xlswrite(filename,pf,'Sheet1','a2:a5');
filename='icdrate.xlsx';
xlswrite(filename,icdrate,'Sheet1','a2:a5');
end

```

LAMPIRAN M
HASIL *CLUSTER* DENGAN C-MEANS

No	Provinsi	2C
1	Aceh	2
2	Sumatera Utara	2
3	Sumatera Barat	2
4	Riau	2
5	Jambi	2
6	Sumatera Selatan	2
7	Bengkulu	2
8	Lampung	2
9	Bangka Belitung	2
10	Kepulauan Riau	2
11	DKI Jakarta	2
12	Jawa Barat	2
13	Jawa Tengah	2
14	DI Yogyakarta	2
15	Jawa Timur	2
16	Banten	2
17	Bali	2
18	NTB	2
19	NTT	2
20	Kalimantan Barat	2
21	Kalimantan Tengah	2
22	Kalimantan Selatan	2
23	Kalimantan Timur	2
24	Sulawesi Utara	2
25	Sulawesi Tengah	2
26	Sulawesi Selatan	2
27	Sulawesi Tenggara	2
28	Gorontalo	2
29	Sulawesi Barat	2
30	Maluku	2
31	Maluku Utara	2
32	Papua Barat	2
33	Papua	1

LAMPIRAN N
HASIL *CLUSTER* DENGAN FUZZY C-MEANS LINEAR
NAIK

No	Provinsi	2C	3C	4C
1	Aceh	1	1	1
2	Sumatera Utara	1	1	1
3	Sumatera Barat	1	1	1
4	Riau	1	1	1
5	Jambi	2	2	2
6	Sumatera Selatan	2	2	2
7	Bengkulu	2	3	4
8	Lampung	2	2	4
9	Bangka Belitung	2	2	2
10	Kepulauan Riau	1	1	1
11	DKI Jakarta	1	1	1
12	Jawa Barat	1	1	1
13	Jawa Tengah	1	2	2
14	DI Yogyakarta	2	3	4
15	Jawa Timur	2	3	4
16	Banten	1	1	1
17	Bali	2	3	4
18	NTB	2	2	2
19	NTT	2	3	4
20	Kalimantan Barat	2	3	4
21	Kalimantan Tengah	2	3	4
22	Kalimantan Selatan	2	3	4
23	Kalimantan Timur	1	1	1
24	Sulawesi Utara	1	1	1
25	Sulawesi Tengah	2	3	4
26	Sulawesi Selatan	2	2	2
27	Sulawesi Tenggara	2	3	4
28	Gorontalo	2	3	4
29	Sulawesi Barat	2	3	4
30	Maluku	1	1	1
31	Maluku Utara	1	2	2
32	Papua Barat	2	3	4
33	Papua	2	3	4

LAMPIRAN O
HASIL *CLUSTER* DENGAN FUZZY C-MEANS LINEAR
TURUN

No	Provinsi	2C	3C	4C
1	Aceh	2	3	4
2	Sumatera Utara	2	3	4
3	Sumatera Barat	2	3	4
4	Riau	2	3	4
5	Jambi	1	1	3
6	Sumatera Selatan	1	1	3
7	Bengkulu	1	2	1
8	Lampung	1	1	1
9	Bangka Belitung	1	1	3
10	Kepulauan Riau	2	3	4
11	DKI Jakarta	2	3	4
12	Jawa Barat	2	3	4
13	Jawa Tengah	2	1	3
14	DI Yogyakarta	1	2	1
15	Jawa Timur	1	2	1
16	Banten	2	3	4
17	Bali	1	2	1
18	NTB	1	1	3
19	NTT	1	2	1
20	Kalimantan Barat	1	2	1
21	Kalimantan Tengah	1	2	1
22	Kalimantan Selatan	1	2	1
23	Kalimantan Timur	2	3	4
24	Sulawesi Utara	2	3	4
25	Sulawesi Tengah	1	2	1
26	Sulawesi Selatan	1	1	3
27	Sulawesi Tenggara	1	2	1
28	Gorontalo	1	2	1
29	Sulawesi Barat	1	2	1
30	Maluku	2	3	4
31	Maluku Utara	2	1	3
32	Papua Barat	1	2	1
33	Papua	1	2	1

LAMPIRAN P
HASIL *CLUSTER* DENGAN FUZZY C-MEANS SEGITIGA

No	Provinsi	2C	3C	4C
1	Aceh	2	2	3
2	Sumatera Utara	2	2	3
3	Sumatera Barat	2	2	3
4	Riau	2	2	3
5	Jambi	1	1	1
6	Sumatera Selatan	1	1	1
7	Bengkulu	1	3	4
8	Lampung	1	1	4
9	Bangka Belitung	1	1	1
10	Kepulauan Riau	2	2	3
11	DKI Jakarta	2	2	3
12	Jawa Barat	2	2	3
13	Jawa Tengah	2	1	1
14	DI Yogyakarta	1	3	4
15	Jawa Timur	1	3	4
16	Banten	2	2	3
17	Bali	1	3	4
18	NTB	1	1	1
19	NTT	1	3	4
20	Kalimantan Barat	1	3	4
21	Kalimantan Tengah	1	3	4
22	Kalimantan Selatan	1	3	4
23	Kalimantan Timur	2	2	3
24	Sulawesi Utara	2	2	3
25	Sulawesi Tengah	1	3	4
26	Sulawesi Selatan	1	1	1
27	Sulawesi Tenggara	1	3	4
28	Gorontalo	1	3	4
29	Sulawesi Barat	1	3	4
30	Maluku	2	2	3
31	Maluku Utara	2	1	1
32	Papua Barat	1	3	4
33	Papua	1	3	4

BIODATA PENULIS



Annisa Sajidah Lahir di Medan, 27 Februari 1994. Penulis yang memiliki hobi membaca, menulis, bernang dan karate merupakan putri pertama Bapak Sumali dan Ibu Cahyoginarti serta memiliki dua adik bernama Fathan Abdallah dan Syahid Hasan Yasin. Sebelumnya, penulis menempuh pendidikan formal di SD Pertiwi Medan, SMPN 7 Medan, dan

SMAN 4 Medan. Setelah lulus SMA, penulis melanjutkan pendidikan ke jenjang perguruan tinggi di Jurusan Statistika ITS pada tahun 2011. Pada tahun kedua sebagai mahasiswa, penulis aktif berorganisasi sebagai staff Saintek BEM FMIPA ITS 12-13 dan berlanjut pada tahun ketiga sebagai Sekretaris Departemen DAGRI BEM FMIPA ITS 13-14. Di Himpunan Mahasiswa Jurusan Penulis aktif sebagai staff HUMAS SCC HIMASTA ITS 12-13 dan berlanjut sebagai sekretaris SCC HIMASTA ITS 13-14. Penulis juga aktif di UKM KARATE ITS dari tahun pertama hingga tahun terakhir perkuliahan dan aktif sebagai atlit yang mengikuti pertandingan karate mahasiswa di berbagai tingkat. Selain mengikuti organisasi di dalam kampus, penulis juga senang mengikuti organisasi luar kampus yaitu AIESEC LC Surabaya dan *take role* di berbagai kegiatan serta pernah menjadi Team Leader di AIESEC sebagai Organizing Committee President of International Career Day 2015 yang memberikan *impact to the society*. Penulis juga dalam kepanitiaan dalam kampus dan luar kampus serta aktif sebagai *freelance* surveyor dan administrator pertandingan.

Untuk berdiskusi lebih lanjut mengenai tugas akhir, hubungi penulis melalui email: annisasajidah@gmail.com