



TUGAS AKHIR - KS184822

**PENGELOMPOKAN KABUPATEN/KOTA DI
PROVINSI PAPUA BERDASARKAN INDIKATOR
INDEKS PEMBANGUNAN MANUSIA
MENGUNAKAN METODE *C-MEANS* DAN *FUZZY
C-MEANS CLUSTERING***

**REZKIANA SUNJADEVA
NRP 062116 4000 0036**

**Dosen Pembimbing
Erma Oktania Permatasari, S.Si., M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS SAINS DAN ANALITIKA DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2020**



TUGAS AKHIR - KS184822

**PENGELOMPOKAN KABUPATEN/KOTA DI
PROVINSI PAPUA BERDASARKAN INDIKATOR
INDEKS PEMBANGUNAN MANUSIA
MENGUNAKAN METODE C-MEANS DAN *FUZZY
C-MEANS CLUSTERING***

**REZKIANA SUNJADEVA
NRP 062116 4000 0036**

**Dosen Pembimbing
Erma Oktania Permatasari, S.Si., M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS SAINS DAN ANALITIKA DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2020**

(Halaman ini sengaja dikosongkan)



FINAL PROJECT - KS184822

**CLASSIFICATION OF THE REGENCIES/CITIES IN
PAPUA PROVINCE BASED ON THE HUMAN
DEVELOPMENT INDEX USING *C-MEANS* AND
*FUZZY C-MEANS CLUSTERING METHODS***

**REZKIANA SUNJADEVA
SN 062116 4000 0036**

**Supervisor
Erma Oktania Permatasari, S.Si., M.Si.**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF SCIENCE AND DATA ANALYTICS
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2020**

(Halaman ini sengaja dikosongkan)

LEMBAR PENGESAHAN

PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI PAPUA BERDASARKAN INDIKATOR INDEKS PEMBANGUNAN MANUSIA MENGGUNAKAN METODE *C-MEANS* DAN *FUZZY C-MEANS CLUSTERING*


TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Statistika
pada
Program Studi Sarjana Departemen Statistika
Fakultas Sains dan Analitika Data
Institut Teknologi Sepuluh Nopember

Oleh :

Rezkianna Sunjadeva
NRP. 062116 4000 0036

Disetujui oleh Pembimbing Tugas Akhir :
Erma Oktania Permatasari, S.Si., M.Si.
NIP. 19881007 201404 2 002

()



Mengetahui,
Kepala Departemen

Dr. Dra. Kartika Fithriasari, M.Si.
NIP. 19691212 199303 2 002

SURABAYA, AGUSTUS 2020

(Halaman ini sengaja dikosongkan)

PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI PAPUA BERDASARKAN INDIKATOR INDEKS PEMBANGUNAN MANUSIA MENGGUNAKAN METODE C-MEANS DAN FUZZY C-MEANS CLUSTERING

Nama Mahasiswa : Rezkiana Sunjadeva
NRP : 062116 4000 0036
Departemen : Statistika-FMKSD-ITS
Dosen Pembimbing : Erma Oktania P., S.Si., M.Si.

Abstrak

Indonesia merupakan negara dengan jumlah penduduk terbesar keempat di dunia. Besarnya jumlah penduduk di Indonesia dapat dijadikan sebagai aset untuk meningkatkan produktivitas masyarakat sehingga dapat meningkatkan kesejahteraan. Fokus RAPBN di tahun 2020 salah satunya merupakan penguatan kualitas SDM yang sehat, cerdas, terampil, dan sejahtera. IPM merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia. IPM dibentuk melalui tiga dimensi dasar yaitu umur panjang dan hidup sehat, pengetahuan, serta standar hidup yang layak. IPM di Indonesia pada tahun 2018 sebesar 71,39, nilai tersebut dibawah target pemerintah yang mematok target sebesar 71,50. IPM terendah yaitu Provinsi Papua sebesar 60,06. Pada penelitian ini dilakukan pengelompokan kabupaten/kota di Papua berdasarkan indikator IPM dengan metode c-means dan fuzzy c-means. Hasil analisis didapatkan kesimpulan bahwa metode fuzzy c-means dengan jumlah kelompok sebanyak 5 merupakan metode terbaik dalam melakukan pengelompokan. Kelompok 3 merupakan kelompok dengan indikator IPM tinggi, kelompok 1 dengan indikator IPM cukup tinggi, kelompok 5 dengan indikator IPM sedang, kelompok 2 dengan indikator IPM cukup rendah, dan kelompok 4 dengan indikator IPM rendah.

Kata kunci: *C-Means, Fuzzy C-Means, Indeks Pembangunan Manusia, Provinsi Papua.*

(Halaman ini sengaja dikosongkan)

**CLASSIFICATION OF THE REGENCIES/CITIES IN
PAPUA PROVINCE BASED ON THE HUMAN
DEVELOPMENT INDEX USING C-MEANS AND FUZZY
C-MEANS CLUSTERING METHODS**

Name : Rezkiana Sunjadeva
Student Number : 062116 4000 0036
Department : Statistics
Supervisor : Erma Oktania P., S.Si., M.Si

Abstract

Indonesia is the fourth most populous country in the world. The large number of people in Indonesia can be used as an asset to increase the productivity of society so as to improve welfare. One of the focuses of the 2020 Draft State Budget is strengthening the quality of human resources who are healthy, smart, skilled and prosperous. HDI is an important indicator to measure success in efforts to build the quality of human life. HDI is formed through three basic dimensions, namely long life and healthy life, knowledge, and a decent standard of living. The HDI in Indonesia in 2018 was 71.39, this value is below the government's target of 71.50. The lowest HDI was Papua Province at 60.06. In this study, regencies / municipalities in Papua were grouped based on HDI indicators using the c-means and fuzzy c-means methods. The results of the analysis concluded that the fuzzy c-means method with the number of groups of 5 is the best method for grouping. Group 3 is a group with a high HDI indicator, group 1 with a fairly high HDI indicator, group 5 with a moderate HDI indicator, group 2 with a fairly low HDI indicator, and group 4 with a low HDI indicator.

Keywords: *C-Means, Fuzzy C-Means, Human Development Index, Papua Province.*

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah SWT yang telah melimpahkan segala rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan Laporan Tugas Akhir yang berjudul : **“Pengelompokan Kabupaten/Kota di Provinsi Papua Berdasarkan Indikator Indeks Pembangunan Manusia Menggunakan Metode *C-Means* dan *Fuzzy C-Means Clustering*”**. Selama proses penyusunan Laporan Tugas Akhir ini, penulis dapat menyelesaikan dengan baik dan lancar tidak lepas dari adanya bantuan berbagai pihak. Oleh karena itu, dengan penuh hormat, ketulusan, dan rendah hati, penulis ingin mengucapkan terima kasih kepada :

1. Ayah tercinta Defi Setiawan dan Ibu tersayang Iva Aquariustina atas doa, nasehat, kasih sayang, dan dukungan yang selalu diberikan kepada penulis sehingga termotivasi untuk menyelesaikan Tugas Akhir ini.
2. Ibu Dr. Dra. Kartika Fithriasari, M.Si. selaku Kepala Departemen Statistika FSAD dan Dr. Santi Wulan, S.Si., M.Si. selaku Sekretaris Departemen Bidang Akademik yang telah menyediakan fasilitas untuk menyelesaikan Tugas Akhir ini.
3. Ibu Erma Oktania Permatasari, S.Si., M.Si. selaku dosen pembimbing yang telah meluangkan waktu, mengarahkan, membimbing dan memberikan dukungan bagi penulis dalam menyelesaikan Tugas Akhir.
4. Ibu Dr. Dra. Agnes Tuti Rumiati, M.Sc. dan Ibu Dr. Vita Ratnasari, S.Si., M.Si. selaku dosen penguji yang telah memberikan koreksi dan saran-saran untuk kesempurnaan Tugas Akhir.
5. Bapak Prof. Dr. Drs. I Nyoman Budiantara, M.Si selaku dosen wali yang telah memberikan nasehat, motivasi dan bimbingan selama ini.
6. Seluruh dosen Statistika ITS yang telah memberikan ilmu dan pengetahuan, serta segenap karyawan Departemen

Statistika ITS yang selalu siap siaga membantu dalam administrasi.

7. Keluarga yang selalu mendoakan dan memberikan kasih sayang dan dukungan sehingga penulis dapat mengerjakan laporan Tugas Akhir ini dengan lancar.
8. Sahabat tercinta Marita Qori'atunnadyah, Nurul Rizqiyah, dan Herni Anggi Riski yang selalu ada saat suka dan duka sehingga penulis menjadi bersemangat dalam mengerjakan Tugas Akhir.
9. Rivi Monica Pratiwi, Kinanthi Sukma Wening, Al'Awwa Prima Nadya Putri yang selalu memberi dukungan, semangat dan menghibur serta menjadi teman berbagi cerita dalam suka maupun duka.
10. Nur Achmey Selgi, Nimas sefrida, dan M. Naufal Abdullah yang telah membantu dan memberikan saran selama mengerjakan Tugas Akhir dan teman-teman TR16GER lainnya yang mengajarkan arti kebersamaan selama 4 tahun ini.
11. Semua pihak yang telah memberikan dukungan dan membantu dalam keberhasilan penulis yang tidak dapat penulis sebutkan satu persatu.

Semoga laporan yang penulis susun dapat bermanfaat dan dapat digunakan sebagaimana mestinya. Penulis menyadari apabila pembuatan Laporan Tugas Akhir ini masih jauh dari kesempurnaan, besar harapan dari penulis untuk menerima kritik dan saran yang berguna untuk perbaikan di masa mendatang. Serta tidak lupa penulis memohon maaf apabila terdapat banyak kekurangan dalam laporan yang telah penulis susun. Atas perhatian dan dukungannya penulis sampaikan ucapan terima kasih.

Surabaya, Juni 2020

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR TABEL	xv
DAFTAR GAMBAR	xvii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	5
1.3 Tujuan	5
1.4 Manfaat.....	5
1.5 Batasan Masalah	6
BAB II TINJAUAN PUSTAKA	7
2.1 Statistika Deskriptif	7
2.2 Analisis <i>Cluster</i>	7
2.3 <i>C-Means Clustering</i>	7
2.4 <i>Fuzzy C-Means Clustering</i>	8
2.5 <i>Pseudo F-Statistic</i>	11
2.6 <i>Internal Cluster Dispersion (Icdrate)</i>	13
2.7 <i>One-Way MANOVA</i>	13
2.7.1 Uji Distribusi Normal Multivariat.....	13
2.7.2 Pengujian Homogenitas	14
2.7.3 Pengujian <i>One-way MANOVA</i>	15
2.8 Pengujian <i>One-Way ANOVA</i>	18
2.9 Indeks Pembangunan Manusia	19
BAB III METODOLOGI PENELITIAN	23
3.1 Sumber Data	23
3.2 Variabel Penelitian.....	23
3.3 Struktur Data.....	23
3.4 Langkah Analisis	24
3.5 Diagram Alir	25

BAB IV ANALISIS DAN PEMBAHASAN	27
4.1 Karakteristik Indikator Indeks Pembangunan Manusia di Provinsi Papua Tahun 2018.....	27
4.1.1 Indikator Pendidikan	28
4.1.2 Indikator Umur Panjang dan Hidup Sehat	31
4.1.3 Indikator Standar Hidup Layak	32
4.2 Pengelompokan dengan Metode <i>C-Means Clustering</i>	33
4.3 Pengelompokan dengan Metode <i>Fuzzy C-Means Clustering</i>	36
4.4 Pemilihan Metode Terbaik.....	39
4.5 Karakteristik Kelompok Metode Terbaik	40
4.5.1 Uji Perbedaan Karakteristik	41
4.5.2 Perbedaan Karakteristik Antar Kelompok	43
BAB V PENUTUP	45
5.1 Kesimpulan	45
5.2 Saran	46
DAFTAR PUSTAKA	47
LAMPIRAN	49
BIODATA PENULIS	59

DAFTAR TABEL

Tabel 2.1	<i>One-Way</i> MANOVA.....	16
Tabel 2.2	Distribusi Wilk's Lambda.....	16
Tabel 2.2	Distribusi Wilk's Lambda (lanjutan)	17
Tabel 2.3	<i>One-Way</i> ANOVA	18
Tabel 2.4	Nilai Maksimum dan Minimum setiap Komponen IPM	20
Tabel 3.1	Variabel Penelitian.....	23
Tabel 3.2	Struktur Data Penelitian	24
Tabel 4.1	Karakteristik Indikator IPM di Papua Tahun 2018	27
Tabel 4.2	Pengelompokan dengan Metode <i>C-Means</i>	34
Tabel 4.3	Nilai <i>Pseudo F-Statistic</i> Metode <i>C-Means</i>	34
Tabel 4.4	Daftar Kabupaten/Kota pada 4 <i>Cluster</i>	36
Tabel 4.5	Pengelompokan dengan Metode <i>Fuzzy C-Means</i> ..	37
Tabel 4.6	Nilai <i>Pseudo F-Statistic</i> Metode <i>Fuzzy C-Means</i> ..	37
Tabel 4.7	Daftar Kabupaten/Kota pada 5 <i>Cluster</i>	39
Tabel 4.8	Perbandingan Nilai <i>Icdrate</i>	40
Tabel 4.9	Hasil Pengujian <i>One-Way</i> ANOVA.....	42
Tabel 4.10	Perbedaan Rata-Rata Masing-Masing Variabel pada Setiap Kelompok	43
Tabel 4.11	Status Pemeringkatan Kelompok	44

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

Gambar 3.1	Diagram Alir	25
Gambar 3.1	Diagram Alir (lanjutan).....	26
Gambar 4.1	Harapan Lama Sekolah Kabupaten/Kota di Papua Tahun 2018.....	29
Gambar 4.2	Rata-Rata Lama Sekolah Kabupaten/Kota di Papua Tahun 2018.....	30
Gambar 4.3	Perbandingan Harapan Lama Sekolah dengan Rata-Rata Lama Sekolah.....	31
Gambar 4.4	Angka Harapan Hidup Kabupaten/Kota di Papua Tahun 2018.....	32
Gambar 4.5	Pengeluaran Perkapita Kabupaten/Kota di Papua Tahun 2018.....	33

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

Lampiran 1	Data Penelitian	49
Lampiran 2	Hasil Pengelompokan dengan Metode <i>C-Means</i> ..	50
Lampiran 3	Hasil Pengelompokan dengan Metode <i>Fuzzy C-Means</i>	51
Lampiran 4	<i>Syntax</i> Pengelompokan dengan Metode <i>Fuzzy C-Means</i>	52
Lampiran 4	<i>Syntax</i> Pengelompokan dengan Metode <i>Fuzzy C-Means</i> (lanjutan).....	53
Lampiran 5	<i>Syntax</i> Perhitungan Nilai <i>Pseudo F-Statistic</i> dan <i>Icdrate</i>	53
Lampiran 5	<i>Syntax</i> Perhitungan Nilai <i>Pseudo F-Statistic</i> dan <i>Icdrate</i> (lanjutan).....	54
Lampiran 6	<i>Macro Minitab</i> Pemeriksaan Asumsi Normal Multivariat.....	55
Lampiran 6	<i>Macro Minitab</i> Pemeriksaan Asumsi Normal Multivariat (lanjutan)	56
Lampiran 7	<i>Output</i> Nilai q_j dan d_j^2	56
Lampiran 7	<i>Output</i> Nilai q_j dan d_j^2 (lanjutan).....	57
Lampiran 8	<i>Output</i> Nilai Korelasi q_j dan d_j^2	57
Lampiran 9	<i>Output</i> Nilai <i>Critical Point</i> untuk Uji Normal Multivariat.....	57
Lampiran 10	<i>Output</i> Uji Homogenitas	58
Lampiran 11	<i>Output</i> Uji <i>One-Way</i> MANOVA	58
Lampiran 12	Surat Pernyataan Data Sekunder.....	59

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara dengan jumlah penduduk terbesar ke empat di dunia setelah negara China, India dan Amerika Serikat. Berdasarkan hasil sensus penduduk yang dilakukan oleh Badan Pusat Statistik (BPS) pada tahun 2010, jumlah penduduk Indonesia mencapai angka 237.641.326 jiwa. Jumlah penduduk tersebut mengalami peningkatan sebesar 15,21% jika dibandingkan dengan jumlah penduduk pada tahun 2000 yang mencapai angka 206.264.595 jiwa, sehingga laju pertumbuhan penduduk dari tahun 2000 hingga 2010 mencapai 1,52 juta jiwa per tahun (BPS, 2012). Besarnya jumlah penduduk di Indonesia dapat dijadikan sebagai aset untuk meningkatkan produktivitas masyarakat sehingga dapat meningkatkan kesejahteraan dan pertumbuhan ekonomi apabila didukung dengan kualitas Sumber Daya Manusia (SDM) yang baik. Namun, pertumbuhan penduduk yang tinggi juga dapat menimbulkan masalah apabila SDM yang ada tidak dimanfaatkan dengan baik, seperti tingginya angka pengangguran yang tidak sebanding dengan jumlah lapangan pekerjaan yang tersedia. Oleh karena itu pembangunan manusia guna memajukan kualitas Sumber Daya Manusia perlu dilakukan untuk meningkatkan pertumbuhan ekonomi dan kesejahteraan masyarakat.

Jokowi dalam pidato pembacaan Nota Keuangan dan Rancangan Anggaran Pendapatan dan Belanja Negara mengungkapkan bahwa fokus RAPBN di tahun 2020 diarahkan pada lima hal utama yang salah satunya merupakan penguatan kualitas SDM yang sehat, cerdas, terampil, dan sejahtera (Kompas, 2019). Indeks Pembangunan Manusia (IPM) merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia (masyarakat/penduduk). IPM dibentuk melalui tiga dimensi dasar yaitu umur panjang dan hidup sehat, pengetahuan, dan standar hidup yang layak. Dimensi umur

panjang dan hidup sehat digambarkan oleh angka harapan hidup saat lahir, dimensi pengetahuan digambarkan oleh rata-rata lama sekolah dan harapan lama sekolah, serta standar hidup yang layak digambarkan oleh pengeluaran perkapita disesuaikan (BPS, 2019).

Indeks Pembangunan Manusia di Indonesia pada tahun 2018 sebesar 71,39, dimana nilai tersebut belum mencapai target pemerintah yang mematok target IPM tahun 2018 sebesar 71,50 (Maharrani, 2019). Dari 38 provinsi di Indonesia tercatat bahwa IPM terendah yaitu pada Provinsi Papua sebesar 60,06 (BPS, 2018). Terdapat 29 kabupaten/kota di Provinsi Papua dimana 17 dari 29 kabupaten tersebut berstatus pembangunan manusia rendah yang memiliki nilai IPM di bawah 60 yang dilihat dari pengkategorian capaian IPM (BPS, 2018). Hal ini menandakan bahwa kualitas hidup manusia di kabupaten/kota Provinsi Papua masih banyak yang tergolong rendah. Selain itu berdasarkan data BPS, sepuluh kabupaten di Provinsi Papua merupakan penyandang IPM terendah di Indonesia pada tahun 2018. Kabupaten dengan capaian IPM terendah atau menempati posisi pertama terendah di Indonesia yaitu Kabupaten Nduga dengan IPM sebesar 29,42. Posisi kedua terendah diduduki oleh Kabupaten Puncak dengan capaian IPM sebesar 41,81. Kabupaten selanjutnya yang memiliki IPM terendah terdapat di Kabupaten Pegunungan Bintang yaitu sebesar 44,2. Kabupaten Memberamo Tengah dan Intan Jaya memiliki selisih IPM yang kecil yaitu masing-masing mencapai 46,41 dan 46,55. Posisi keenam hingga kedelapan terendah mencatatkan IPM kisaran 47 yaitu pada Kabupaten Yalimo mencapai 47,13, Kabupaten Lanny Jaya mencapai 47,34, dan Kabupaten Puncak Jaya mencapai 47,39. Kabupaten Yahukimo menduduki peringkat kesembilan terendah dengan IPM sebesar 48,51. Kabupaten kesepuluh terendah di Indonesia ditempati oleh Kabupaten Tolikara dengan IPM sebesar 48,85. IPM tertinggi di Provinsi Papua diduduki oleh Kota Jayapura dengan IPM mencapai angka 79,58 dimana capaian tersebut telah memenuhi target dari pemerintah sebesar 71,50. Namun jika dibandingkan dengan Kabupaten Nduga dengan IPM

29,42, telah terjadi ketimpangan pembangunan manusia antar wilayah di Provinsi Papua. Hal ini dikarenakan adanya perbedaan yang sangat jauh antara IPM tertinggi dan terendah pada kabupaten/kota di Provinsi Papua. Oleh karena itu perlu dilakukan pengelompokan kabupaten/kota berdasarkan indikator Indeks Pembangunan Manusia yang homogen.

Analisis *cluster* merupakan salah satu analisis multivariat yang digunakan untuk mengelompokkan objek-objek berdasarkan kemiripan karakteristik yang dimilikinya. Dimana objek-objek dalam suatu kelompok (*cluster*) sangat tinggi sedangkan tingkat kemiripan karakteristik objek antar *cluster* satu dengan yang lainnya rendah (Hair, Black, Babin, & Anderson, 2010). Terdapat dua metode *cluster* yaitu hierarki dan non-hierarki. Metode non-hierarki digunakan untuk pengelompokan objek dimana banyaknya *cluster* yang akan dibentuk dapat ditentukan terlebih dahulu sebagai bagian dari prosedur pengelompokan. Selain itu metode ini dapat diterapkan pada data yang lebih besar dibandingkan metode hierarki (Johnson & Wichern, 2007). Metode *c-means* dan *fuzzy c-means clustering* merupakan contoh metode pengelompokan non-hierarki. *C-means* merupakan salah satu metode data *clustering* non-hierarki dengan mempartisi data yang ada ke dalam bentuk satu atau lebih kelompok atau *cluster*. Metode ini mempartisi data ke dalam kelompok atau *cluster* sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu kelompok dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain. Tujuan dari metode ini adalah untuk meminimalisasikan *objective function* yang diset dalam proses *clustering* dimana pada umumnya berusaha meminimalisasikan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok (Agusta, 2007). *Fuzzy c-means* merupakan metode pengembangan dari *c-means clustering* dengan pembobotan *fuzzy*. Mingoti dan Lima pada tahun 2006 melakukan perbandingan pengelompokan metode hierarki dan non-hierarki dengan mensimulasikan data sebanyak 2530 set data. Didapatkan

hasil bahwa metode *fuzzy c-means* memiliki hasil paling baik terutama pada kasus *outlier* dan *overlapping* jika dibandingkan dengan pengelompokan hierarki (*single, average, complete linkage*), *c-means*, dan SOM *neural network* (Mingoti & Lima, 2006).

Penelitian yang berkaitan dengan metode *c-means* dan *fuzzy c-means* pernah dilakukan oleh Cendiana Aprilia Haryono (2016) dalam tugas akhirnya yang berjudul “Penerapan Metode *C-Means* dan *Fuzzy C-Means* pada Pengelompokan Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Pembangunan Ekonomi”. Berdasarkan hasil analisis tersebut didapatkan hasil bahwa metode *c-means clustering* memiliki nilai *icdrate* terkecil jika dibandingkan dengan metode *fuzzy c-means*, sehingga metode *c-means* merupakan metode terbaik dalam melakukan pengelompokan. Muhammad Fikri Masteriarsa (2019) juga pernah melakukan penelitian terkait metode *c-means* dan *fuzzy c-means* dalam tugas akhirnya yang berjudul “Pemetaan Daya Saing Indonesia di Tingkat Dunia Berdasarkan Kinerja Pembangunan dengan Metode *K Means* dan *Fuzzy C-Means*”. Pada penelitiannya tersebut menunjukkan bahwa metode *fuzzy c-means* merupakan metode terbaik untuk memetakan negara didunia berdasarkan indikator pembangunan pariwisata.

Berdasarkan dari uraian yang telah disampaikan dan beberapa penelitian sebelumnya, akan dilakukan pengelompokan kabupaten/kota di Provinsi Papua. Pengelompokan dilakukan berdasarkan indikator Indeks Pembangunan Manusia pada tahun 2018 menggunakan metode *c-means* dan *fuzzy c-means*. Pengelompokan berdasarkan indikator yang memengaruhi IPM tersebut bertujuan untuk mengelompokkan kabupaten/kota di Provinsi Papua berdasarkan kesamaan karakteristik yang dimiliki. Setelah itu dilakukan perbandingan kedua metode untuk mendapatkan hasil pengelompokan terbaik, sehingga penelitian ini dapat memberikan informasi sebagai acuan pemerintah Provinsi Papua dalam membuat kebijakan atau upaya perbaikan

bagi wilayah yang termasuk dalam kelompok kabupaten/kota dengan IPM yang rendah.

1.2 Rumusan Masalah

Permasalahan yang muncul dalam penelitian ini adalah sebagai berikut.

1. Bagaimana karakteristik kabupaten/kota di Provinsi Papua berdasarkan indikator Indeks Pembangunan Manusia pada tahun 2018?
2. Bagaimana hasil pengelompokan kabupaten/kota Provinsi Papua berdasarkan indikator Indeks Pembangunan Manusia pada tahun 2018 menggunakan metode *c-means* dan *fuzzy c-means*?
3. Bagaimana perbandingan hasil pada metode *c-means* dan *fuzzy c-means*?
4. Bagaimana karakteristik pada masing-masing kelompok yang terbentuk?

1.3 Tujuan

Berdasarkan rumusan masalah di atas, tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik kabupaten/kota di Provinsi Papua berdasarkan indikator Indeks Pembangunan Manusia pada tahun 2018.
2. Mendapatkan hasil pengelompokan kabupaten/kota Provinsi Papua berdasarkan indikator Indeks Pembangunan Manusia pada tahun 2018 menggunakan metode *c-means* dan *fuzzy c-means*.
3. Membandingkan hasil pada metode *c-means* dan *fuzzy c-means*.
4. Mendeskripsikan karakteristik pada masing-masing kelompok yang terbentuk.

1.4 Manfaat

Penelitian ini bermanfaat sebagai informasi tentang pengelompokan kabupaten/kota di Provinsi Papua berdasarkan

indikator Indeks Pembangunan Manusia menggunakan metode *c-means* dan *fuzzy c-means clustering*. Selain itu, dapat dijadikan sebagai gambaran untuk meningkatkan indeks komponen/indikator IPM yaitu peningkatan indeks harapan lama sekolah, indeks rata-rata lama sekolah, indeks angka harapan hidup, dan indeks pengeluaran perkapita di Provinsi Papua sehingga terwujud IPM di kabupaten/kota Provinsi Papua yang memenuhi target pemerintah. Manfaat bagi peneliti adalah dapat menerapkan ilmu statistik dalam menangani permasalahan khususnya dibidang sosial kependudukan.

1.5 Batasan Masalah

Dalam penelitian ini data yang digunakan adalah data indikator Indeks Pembangunan Manusia di Provinsi Papua pada tahun 2018.

BAB II

TINJAUAN PUSTAKA

Pada bagian ini diuraikan tinjauan pustaka terkait dengan metode yang digunakan untuk melakukan pengelompokan, yaitu sebagai berikut.

2.1 Statistika Deskriptif

Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Dalam statistika deskriptif terdapat dua macam penyajian yaitu visualisasi data berupa grafik/diagram dan tabel dalam satu variabel maupun dua variabel serta penyajian dalam bentuk ukuran-ukuran statistik yaitu ukuran pemusatan dan penyebaran (Walpole, 1995).

2.2 Analisis Cluster

Analisis *cluster* merupakan salah satu analisis multivariat yang bertujuan untuk mengelompokkan objek-objek berdasarkan kemiripan karakteristik yang dimilikinya. Tingkat kemiripan karakteristik yang dimaksud adalah objek-objek dalam suatu kelompok (*cluster*) sangat tinggi sedangkan tingkat kemiripan karakteristik objek antar *cluster* satu dengan yang lainnya rendah (Hair, Black, Babin, & Anderson, 2010).

2.3 C-Means Clustering

C-means merupakan metode *clustering* non-hierarki yang paling sering digunakan (Johnson & Wichern, 2007). *C-means* merupakan salah satu metode data *clustering* non-hierarki dengan mempartisi data yang ada ke dalam bentuk satu atau lebih kelompok atau *cluster*. Metode ini mempartisi data ke dalam kelompok atau *cluster* sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu kelompok dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain. Tujuan dari data metode ini adalah untuk meminimalisasikan *objective function* yang diset dalam proses

clustering dimana pada umumnya berusaha meminimalisasikan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok (Agusta, 2007). Algoritma atau langkah dari metode *c-means* adalah sebagai berikut.

1. Menentukan jumlah cluster yang akan dibentuk (besarnya c)
2. Mengalokasikan data kedalam c cluster secara random
3. Menghitung nilai *centroid* rata-rata dari masing-masing cluster. Rumus yang digunakan untuk menghitung nilai *centroid* adalah sebagai berikut.

$$v_{cj} = \frac{\sum_{i=1}^{n_c} x_{ij}}{n_c} \quad (2.1)$$

Keterangan :

v_{cj} : *centroid cluster* ke- c untuk variabel ke- j

n_c : banyak data yang menjadi anggota *cluster* ke- c

x_{ij} : data ke- i pada variabel ke- j

i : indeks objek

c : indeks *cluster*

j : indeks variabel

4. Mengelompokkan masing-masing data ke *centroid* terdekat (menggunakan jarak *euclidean*) dengan rumus :

$$d_{ci} = \sqrt{\left(\sum_{j=1}^J (x_{ij} - v_{cj})^2 \right)} \quad (2.2)$$

Keterangan :

d_{ci} : jarak data ke- i dengan pusat *cluster* ke- c

Kembali ke langkah 3 apabila masih terdapat perpindahan data dari satu *cluster* yang lain (Agusta, 2007).

2.4 Fuzzy C-Means Clustering

Metode *fuzzy c-means clustering* mengalokasikan kembali data ke dalam masing-masing *cluster* dengan memanfaatkan teori *fuzzy*. Pada metode *fuzzy c-means* digunakan variabel derajat

keanggotaan (*membership function*) yang merujuk pada seberapa besar kemungkinan suatu data bisa menjadi anggota ke dalam suatu *cluster*. Selain itu, diperkenalkan juga suatu variabel m yang merupakan *weighting exponent* dari derajat keanggotaan. m memiliki wilayah nilai lebih besar dari satu ($m > 1$). Sampai sekarang ini tidak ada ketentuan yang jelas berapa besar nilai m yang optimal dalam melakukan proses optimasi suatu permasalahan *clustering*. Nilai m yang umumnya digunakan adalah 2 (Agusta, 2007).

Pada metode *fuzzy c-means* menerapkan konsep bahwa penentuan pusat kelompok untuk pertama kali akan menandai lokasi rata-rata pada setiap kelompok. Pada konsep ini setiap data dapat menjadi anggota pada beberapa kelompok sekaligus, sehingga dapat dikatakan bahwa batas-batas dalam *fuzzy c-means* bersifat lembut (*soft*). Sehingga dengan melakukan perbaikan pusat kelompok dan nilai keanggotaan melalui perulangan algoritma *fuzzy c-means* akan didapatkan pusat kelompok yang akan menuju lokasi yang tepat. Perulangan pada algoritma ini didasarkan pada minimasi fungsi objektif pada persamaan di bawah ini.

$$J(U, c_1, \dots, c_g) = \sum_{c=1}^g J_c = \sum_{c=1}^g \sum_j^n u_{ci}^m d_{ci}^2 \quad (2.3)$$

Keterangan :

u_{ci} : keanggotaan objek ke- i dan kelompok ke- c

c_g : matriks *centroid* dari masing-masing kelompok

n : banyak data

c : banyak kelompok

m : *weight exponent*

$d_{ci} = \|c_g - x_i\|$ adalah jarak euclidean kelompok ke- c wilayah *centroid* ke- i .

Fungsi objektif tersebut menggambarkan jarak dari titik data yang diberikan ke pusat kelompok yang terbobot oleh derajat jarak keanggotaan titik data tersebut. *Membership function* untuk

suatu data kedalam kelompok tertentu dapat dihitung berdasarkan persamaan di bawah ini.

$$u_{ci} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ci}}{d_{ki}} \right)^{2/(m-1)}} \quad (2.4)$$

Keterangan :

u_{ci} : keanggotaan data ke- i kelompok ke- c

d_{ci} : nilai *centroid* data ke- i kelompok ke- c

d_{ki} : nilai *centroid* data ke- i kelompok ke- k

c : banyak kelompok

m : *weight exponent*

Membership fuction memiliki nilai jangkauan $0 \leq u_{ci} \leq 1$.

Sedangkan untuk menghitung nilai *centroid cluster* dapat menggunakan persamaan (2.5).

$$c_i = \frac{\sum_{i=1}^n u_{ci}^m x_i}{\sum_{i=1}^n u_{ci}^m} \quad (2.5)$$

dimana x_i merupakan objek atau data ke- i (Jang, Sun, & Mizutani, 1997).

Algoritma yang digunakan pada metode *fuzzy c-means* untuk menentukan pusat kelompok c_i dengan *membership matrix* U adalah sebagai berikut.

1. Menentukan banyak *cluster* atau kelompok yang akan dibentuk (c).
2. Menentukan pangkat pembobot ($m > 1$). Besarnya nilai pembobot yang sering digunakan adalah 2.
3. Menentukan batas toleransi atau kriteria perulangan (*threshold*).
4. Membentuk matriks partisi awal U (derajat keanggotaan). Matriks U merupakan bilangan random antara 0 sampai

dengan 1 dengan struktur berdasarkan persamaan di bawah ini.

$$\mathbf{U} = \begin{bmatrix} \mu_{11}(x_1) & \mu_{12}(x_2) & \cdots & \mu_{1i}(x_i) \\ \mu_{21}(x_1) & \mu_{22}(x_2) & \cdots & \mu_{2i}(x_i) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{c1}(x_1) & \mu_{c2}(x_2) & \cdots & \mu_{ci}(x_j) \end{bmatrix} \quad (2.6)$$

5. Menghitung nilai *centroid fuzzy* $c_i, i=1, \dots, c$ berdasarkan persamaan (2.5) yang kemudian akan dibentuk matriks baru nilai *centroid* untuk menghitung fungsi objektif.
6. Menghitung fungsi objektif *fuzzy c-means* berdasarkan persamaan (2.3). Nilai fungsi objektif digunakan untuk penentuan iterasi akan berlanjut atau berhenti setelah dibandingkan dengan nilai kriteria penolakan. Iterasi berhenti ketika nilai fungsi objektif dibawah batas toleransi (*threshold*).
7. Ketika nilai fungsi objektif diatas *threshold*, maka dilakukan perghitungan *membership function* atau matriks partisi \mathbf{U} yang baru berdasarkan persamaan (2.4) dan dilanjutkan dengan kembali ke langkah 5. Matriks ini digunakan untuk menentukan kelompok pada masing-masing observasi setelah iterasi berhenti.

2.5 *Pseudo F-Statistic*

Pseudo F-statistic digunakan untuk menentukan banyaknya kelompok yang optimum. *Pseudo F-statistic* tertinggi menunjukkan bahwa kelompok tersebut merupakan hasil yang optimal, dimana keragaman dalam kelompok homogen sedangkan antar kelompok heterogen. Rumus yang digunakan untuk mendapatkan *pseudo F-statistic* adalah sebagai berikut (Orpin & Kostylev, 2006).

$$\text{Pseudo } F \text{ - statistic} = \frac{\left(\frac{R^2}{c-1} \right)}{\left(\frac{1-R^2}{n-c} \right)} \quad (2.7)$$

Dimana $R^2 = \frac{(SST - SSW)}{SST}$ dengan nilai SST dan SSW berdasarkan rumus berikut.

$$\begin{aligned} SST &= \sum_{i=1}^n \sum_{c=1}^c \sum_{j=1}^J (x_{ijc} - \bar{x}_j)^2 \\ &= (x_{111} - \bar{x}_1)^2 + (x_{211} - \bar{x}_1)^2 + \dots + (x_{ijc} - \bar{x}_j)^2 \\ &= (x_{111} - \bar{x}_1)(x_{111} - \bar{x}_1) + (x_{211} - \bar{x}_1)(x_{211} - \bar{x}_1) + \dots + (x_{ijc} - \bar{x}_j)(x_{ijc} - \bar{x}_j) \end{aligned} \quad (2.8)$$

$$\begin{aligned} SSW &= \sum_{i=1}^n \sum_{c=1}^c \sum_{j=1}^J (x_{ijc} - \bar{x}_{jc})^2 \\ &= (x_{111} - \bar{x}_{11})^2 + (x_{211} - \bar{x}_{11})^2 + \dots + (x_{ijc} - \bar{x}_{jc})^2 \\ &= (x_{111} - \bar{x}_{11})(x_{111} - \bar{x}_{11}) + (x_{211} - \bar{x}_{11})(x_{211} - \bar{x}_{11}) + \dots + (x_{ijc} - \bar{x}_{jc})(x_{ijc} - \bar{x}_{jc}) \end{aligned} \quad (2.9)$$

Keterangan :

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

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

i : objek penelitian

c : banyaknya kelompok

j : variabel penelitian

x_{ijc} : sampel ke-*i* pada kelompok ke-*c* dan variabel ke-*j*

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

\bar{x}_{jc} : rata-rata sampel pada kelompok ke-*c* dan variabel ke-*j*.

2.6 *Internal Cluster Dispersion (Icdrate)*

Icdrate merupakan metode yang digunakan untuk membandingkan metode *cluster* yang terbaik dengan mengevaluasi performansi algoritma dengan menggunakan persentase rata-rata dari klasifikasi yang benar (*recovery rate*) dan nilai persebaran data-data dalam cluster (*internal cluster dispersion rate*) dari hasil akhir pengelompokan (Mingoti & Lima, 2006). Persamaan yang digunakan adalah sebagai berikut.

$$Icdrate = 1 - \frac{SSB}{SST} = 1 - \frac{SST - SSW}{SST} = 1 - R^2 \quad (2.10)$$

Keterangan :

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

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

SSB : *Sum Square Between* ($SST - SSW$)

R^2 : *recovery rate*.

2.7 *One-Way MANOVA*

Sebelum dilakukan MANOVA (*Multivariate Analyze of Varians*), perlu dilakukan pemeriksaan asumsi. Data harus memenuhi asumsi data berdistribusi normal multivariat dan matriks varians kovarians bersifat homogen.

2.7.1 Uji Distribusi Normal Multiariat

Analisis normal multivariat merupakan perluasan dari distribusi normal univariat dengan minimal dua variabel pengamatan. Dalam analisis multivariat, asumsi normal multivariat perlu dilakukan untuk memastikan bahwa data pengamatan mengikuti distribusi normal (Rencher, 2002).

Uji normal multivariat dapat dilakukan dengan menggunakan uji koefisien korelasi. Berikut merupakan uji koefisien korelasi untuk melihat apakah data memenuhi asumsi distribusi normal multivariat atau tidak.

Hipotesis :

H_0 : Data berdistribusi normal multivariat

H_1 : Data tidak berdistribusi normal multivariat

Statistik uji :

$$r_Q = \frac{\sum_{j=1}^n (x_{(j)} - \bar{x})(q_{(j)} - \bar{q})}{\sqrt{\sum_{j=1}^n (x_{(j)} - \bar{x})^2} \sqrt{\sum_{j=1}^n (q_{(j)} - \bar{q})^2}} \quad (2.11)$$

dimana $q_{(j)}$ merupakan kuantil normal standar dan dirumuskan

dengan $q_{(j)} = \frac{j - \frac{1}{2}}{n}, j = 1, 2, \dots, n$.

Dikatakan tolak H_0 jika $r_Q < r_{(\alpha, n)}$. Jika hasil dari statistik uji memiliki hasil yang kurang dari tabel normal probabilitas koefisien korelasi $r_{(\alpha, n)}$, maka dapat disimpulkan bahwa data tidak berdistribusi normal multivariat, begitu pun sebaliknya (Johnson & Wichern, 2007).

2.7.2 Pengujian Homogenitas

Mengetahui kehomogenan matriks varians kovarians dapat dilakukan dengan menggunakan Uji Box's M. Pengujian hipotesis Box's M adalah sebagai berikut.

Hipotesis :

$H_0 : \Sigma_1 = \Sigma_2 = \dots = \Sigma_g = \Sigma$

H_1 : minimal ada satu $\Sigma_g \neq \Sigma_j$ untuk $g \neq j$ dimana $c = 1, 2, \dots, g$

Statistik uji :

$$\begin{aligned} \chi_{hitung}^2 &= (1-u)M \\ &= (1-u) \left\{ \left[\sum_{c=1}^g (n_c - 1) \right] \ln |S_{pooled}| - \sum_{c=1}^g [(n_c - 1) \ln |S_c|] \right\} \quad (2.12) \end{aligned}$$

dimana :

$$Spool = \frac{1}{\sum_{c=1}^g (n_c - 1)} - \left\{ (n_1 - 1)S_1 + \dots + (n_g - 1)S_g \right\} \quad (2.13)$$

$$u = \left[\sum_{c=1}^g \frac{1}{(n_c - 1)} - \frac{1}{\sum_{c=1}^g (n_c - 1)} \right] \left[\frac{2p^2 + 3p - 1}{6(p+1)(g-1)} \right] \quad (2.14)$$

Keterangan :

- c : jumlah subgrup sebanyak g
- p : banyaknya variabel yang digunakan
- n_c : ukuran sampel untuk kelompok ke- c
- S_c : matriks varians kovarians kelompok ke- c
- S_{pool} : matriks varians kovarians gabungan

Apabila ditetapkan tingkat signifikansi α sebesar 0,05 maka gagal tolak H_0 jika nilai $\chi_{hitung}^2 \leq \chi_{\frac{1}{2}(g-1)p(p+1)}^2$ artinya matriks varians kovarians bersifat homogen (Johnson & Wichern, 2007).

2.7.3 Pengujian *One-Way* MANOVA

One-Way MANOVA digunakan untuk membandingkan rata-rata dua populasi atau lebih dengan variabel dependen lebih dari satu atau untuk mengkaji pengaruh dari suatu perlakuan terhadap respon (Johnson & Wichern, 2007). MANOVA digunakan untuk mengetahui kemiripan antar kelompok yang telah terbentuk.

Hipotesis yang digunakan dalam *one-way* MANOVA yaitu,

H_0 : $\tau_1 = \tau_2 = \dots = \tau_g = 0$ (tidak ada efek perlakuan antar kelompok)

H_1 : minimal ada satu $\tau_c \neq 0$, $c=1,2,\dots,g$ (terdapat minimal 1 efek perlakuan antar kelompok)

Berikut merupakan tabel dari pengujian *one-way* MANOVA.

Tabel 2.1 *One-Way* MANOVA

Sumber Variasi	Matriks Jumlah Kuadrat	Derajat Bebas
Perlakuan	$B = \sum_{c=1}^g n_c (\bar{X}_c - \bar{X})(\bar{X}_c - \bar{X})^T$	$g - 1$
Error	$W = \sum_{c=1}^g \sum_{i=1}^{n_c} (x_{ci} - \bar{X}_c)(x_{ci} - \bar{X}_c)^T$	$\sum_{c=1}^g n_c - g$
Total Terkoreksi	$B + W = \sum_{c=1}^g \sum_{i=1}^{n_c} (x_{ci} - \bar{X})(x_{ci} - \bar{X})^T$	$\sum_{c=1}^g n_c - 1$

Statistik uji yang digunakan pada uji *One-Way* MANOVA yaitu Wilk's Lambda. Berdasarkan Tabel 2.1, rumus persamaan dari Wilk's Lambda yang digunakan sebagai berikut.

$$\Lambda^* = \frac{|W|}{|B + W|} \quad (2.15)$$

Keterangan :

W : *matrix sum of square residuals*

B : *matrix sum of square treatment*

\bar{X}_c : rata-rata kelompok ke- c

X_{ci} : data ke- i pada kelompok ke- c

n_c : banyak anggota pada kelompok ke- c

Hasil yang didapatkan dari distribusi Wilk's Lambda dapat dilakukan pendekatan dengan distribusi F.

Tabel 2.2 Distribusi Wilk's Lambda

Jumlah Variabel	Jumlah Grup	Distribusi Sampling
$p = 1$	$g \geq 2$	$\left(\frac{\sum n_c - g}{g - 1} \right) \left(\frac{1 - \wedge^*}{\wedge^*} \right) \sim F_{g-1, \sum n_c - g}$

Tabel 2.2 Distribusi Wilk's Lambda (lanjutan)

Jumlah Variabel	Jumlah Grup	Distribusi Sampling
$p = 2$	$g \geq 2$	$\left(\frac{\sum n_c - g - 1}{g - 1} \right) \left(\frac{1 - \sqrt{\wedge^*}}{\sqrt{\wedge^*}} \right) \sim F_{2(g-1), 2(\sum n_i - g - 1)}$
$p \geq 1$	$g = 2$	$\left(\frac{\sum n_c - p - 1}{p} \right) \left(\frac{1 - \wedge^*}{\wedge^*} \right) \sim F_{p, \sum n_c - p - 1}$
$p \geq 1$	$g = 3$	$\left(\frac{\sum n_c - p - 2}{p - 2} \right) \left(\frac{1 - \wedge^*}{\wedge^*} \right) \sim F_{2p, 2(\sum n_c - p - 2)}$

Dikatakan tolak H_0 jika $F_{hitung} \geq F_{tabel}$.

Ketika asumsi homogenitas maupun normal multivariat tidak terpenuhi, maka digunakan statistik uji *Pillai's Trace* yang bersifat robust terhadap asumsi yang terlanggar. Statistik uji *Pillai's Trace* dapat dihitung berdasarkan persamaan di bawah ini.

$$F = \frac{(2N + s + 1)V^{(s)}}{(2m + s + 1)(s - V^{(s)})} \quad (2.16)$$

dengan

$$V^{(s)} = tr[\mathbf{B}(\mathbf{W} + \mathbf{B})^{-1}] \quad (2.17)$$

$$N = \frac{1}{2}(VE - p - 1) \quad (2.18)$$

$$m = \frac{1}{2}(|VH - p| - 1) \quad (2.19)$$

Keterangan :

s : $\min(VH, p)$

VH : derajat bebas hipotesis

VE : derajat bebas eror

p : banyaknya variabel

Diperoleh keputusan tolak H_0 jika nilai $F \geq F_{s(2m+s+1),s(2N+s+1),\alpha}$ (Rencher, 2002).

2.8 Pengujian *One-Way* ANOVA

One-way ANOVA digunakan untuk uji perbedaan kelompok ketika variabel terikat yang digunakan hanya satu atau uji perbedaan pada variabel-variabel antar anggota kelompok (Johnson & Wichern, 2007). Hipotesis yang digunakan adalah sebagai berikut.

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

$$H_1: \text{minimal ada 1 } \tau_c \neq 0, c = 1, 2, \dots, g$$

Berikut merupakan tabel dari pengujian *one-way* ANOVA.

Tabel 2.3 *One-Way* ANOVA

Sumber Variasi	Matriks Jumlah Kuadrat	Derajat Bebas
Perlakuan	$SS_{tr} = \sum_{c=1}^g n_c (\bar{X}_c - \bar{X})^2$	$g - 1$
Error	$SS_{res} = \sum_{c=1}^g \sum_{i=1}^{n_c} (x_{ci} - \bar{X}_c)^2$	$\sum_{c=1}^g n_c - g$
Total Terkoreksi	$SS_{total} = \sum_{c=1}^g \sum_{i=1}^{n_c} (x_{ci} - \bar{X})^2$	$\sum_{c=1}^g n_c - 1$

Statistik uji yang digunakan pada uji *One-Way* ANOVA adalah sebagai berikut.

$$F = \frac{SS_{tr} / (g - 1)}{SS_{res} / \left(\sum_{c=1}^g n_c - g \right)} \quad (2.20)$$

Dikatakan tolak H_0 ketika $F_{hitung} > F_{g-1, \sum n_c - g, (\alpha)}$.

2.9 Indeks Pembangunan Manusia

Indeks Pembangunan Manusia merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia (masyarakat/penduduk). IPM dibentuk melalui tiga dimensi dasar yaitu umur panjang dan hidup sehat, pengetahuan, dan standar hidup yang layak. Dimensi umur panjang dan hidup sehat digambarkan oleh angka harapan hidup saat lahir. Dimensi pengetahuan digambarkan oleh rata-rata lama sekolah dan harapan lama sekolah. Serta standar hidup yang layak digambarkan oleh pengeluaran perkapita yang disesuaikan (BPS, 2019). IPM dibagi menjadi empat kategori yaitu :

- Rendah, jika $IPM < 60$
- Sedang, jika $60 \leq IPM < 70$
- Tinggi, jika $70 \leq IPM < 80$
- Sangat tinggi, jika $IPM \geq 80$.

Perhitungan IPM dilakukan dengan menghitung indeks dari setiap komponen. Formula yang digunakan untuk menghitung indeks setiap komponen pembentuk IPM yaitu:

- Dimensi pendidikan

$$I_{HLS} = \frac{HLS - HLS_{min}}{HLS_{maks} - HLS_{min}} \quad (2.21)$$

$$I_{RLS} = \frac{RLS - RLS_{min}}{RLS_{maks} - RLS_{min}} \quad (2.22)$$

$$I_{pendidikan} = \frac{I_{HLS} + I_{RLS}}{2} \quad (2.23)$$

- Dimensi kesehatan

$$I_{kesehatan} = \frac{AHH - AHH_{min}}{AHH_{maks} - AHH_{min}} \quad (2.24)$$

- Dimensi standar hidup layak

$$I_{pengeluaran} = \frac{\ln(\text{pengeluaran}) - \ln(\text{pengeluaran}_{min})}{\ln(\text{pengeluaran}_{maks}) - \ln(\text{pengeluaran}_{min})} \quad (2.25)$$

Keterangan :

I_{HLS} : Indeks harapan lama sekolah

I_{RLS} : Indeks rata-rata lama sekolah

I_{AHH} : Indeks angka harapan hidup

$I_{pengeluaran}$: Indeks pengeluaran perkapita disesuaikan

Menghitung indeks masing-masing komponen IPM digunakan batas dan maksimum dan minimum seperti pada tabel di bawah ini.

Tabel 2.4 Nilai Maksimum dan Minimum Setiap Komponen IPM

Komponen IPM	Satuan	Minimum	Maksimum
Angka harapan hidup	Tahun	20	85
Harapan lama sekolah	Tahun	0	18
Rata-rata lama sekolah	Tahun	0	15
Pengeluaran perkapita	Rupiah	1007436	26572352

Berdasarkan ketiga nilai indeks komponen tersebut, didapatkan rumus perhitungan IPM sebagai berikut (BPS, 2018).

$$IPM = \sqrt[3]{I_{kesehatan} \times I_{pendidikan} \times I_{pengeluaran}} \times 100 \quad (2.26)$$

Penjelasan dari indikator pembentuk IPM adalah sebagai berikut.

a. Harapan lama sekolah

Harapan lama sekolah didefinisikan sebagai lamanya sekolah (dalam tahun) yang diharapkan akan dirasakan oleh anak pada usia tertentu di masa mendatang. Menghitung harapan lama sekolah mencakup pendidikan dari penduduk usia 7 tahun ke atas (BPS, 2018). Harapan lama sekolah dapat dirumuskan sebagai berikut.

$$HLS'_a = FK \times \sum_{i=a}^n \frac{E'_i}{P'_i} \quad (2.27)$$

Keterangan :

HLS'_a : harapan lama sekolah pada umur a di tahun t

E'_i : jumlah penduduk usia i yang bersekolah pada tahun t

- P'_i : jumlah penduduk usia i pada tahun t
 FK : faktor koreksi persantren
 i : usia ($a, a + 1, \dots, n$)

b. Rata-rata lama sekolah

Rata-rata lama sekolah merupakan rata-rata lamanya (tahun) penduduk usia 25 tahun ke atas dalam menjalani pendidikan formal (BPS, 2018). Rumus rata-rata lama sekolah sebagai berikut.

$$RLS = \frac{1}{n} \times \sum_{i=1}^n x_i \quad (2.28)$$

Keterangan :

- RLS : rata-rata lama sekolah penduduk usia 25 tahun ke atas
 x_i : lama sekolah penduduk ke- i yang berusia 25 tahun
 n : jumlah penduduk usia 25 tahun ke atas

c. Angka harapan hidup

Angka harapan hidup merupakan jumlah tahun yang diharapkan dapat ditempuh oleh bayi yang baru lahir untuk hidup dengan asumsi bahwa pola angka kematian menurut umur pada saat kelahiran sama sepanjang usia bayi (BPS, 2019).

d. Pengeluaran perkapita disesuaikan

Standar hidup layak digambarkan oleh pengeluaran perkapita disesuaikan yang ditentukan dari nilai pengeluaran perkapita dan paritas daya beli (BPS, 2019). Daya beli merupakan kemampuan masyarakat dalam membelanjakan uangnya dalam bentuk barang maupun jasa dimana daya beli antar wilayah atau daerah berbeda-beda. Perhitungan pengeluaran perkapita disesuaikan adalah sebagai berikut (BPS, 2018).

$$Y^{**} = \frac{Y^*}{PPP} \quad (2.29)$$

dengan

$$Y^* = \frac{Y}{IHK} \times 100 \quad (2.30)$$

$$PPP_j = \prod_{i=1}^m \left(\frac{p_{ij}}{p_{ik}} \right)^{\frac{1}{m}} \quad (2.31)$$

Keterangan :

- Y^{**} : pengeluaran perkapita disesuaikan
- Y^* : pengeluaran perkapita harga konstan
- Y : pengeluaran perkapita setahun
- IHK : Indeks Harga Konsumen tahun dasar 2012
- PPP_j : paritas daya beli di wilayah j
- p_{ij} : harga komoditas i di kabupaten/kota j
- p_{ik} : harga komoditi i di Jakarta Selatan
- m : jumlah komoditi

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari buku Provinsi Papua dalam Angka 2019 yang dipublikasikan oleh Badan Pusat Statistik (BPS). Unit penelitian berupa kabupaten/kota di Provinsi Papua dengan jumlah kabupaten/kota sebanyak 29 dan data yang diambil merupakan data pada tahun 2018.

3.2 Variabel Penelitian

Variabel yang digunakan untuk melakukan pengelompokan kabupaten/kota di Provinsi Papua terdiri dari data berupa indikator Indeks Pembangunan Manusia dengan jumlah variabel yang digunakan empat variabel. Terdapat tiga indikator pembentuk Indeks Pembangunan Manusia yaitu indikator pendidikan, indikator kesehatan, dan indikator ekonomi atau standar hidup layak. Di bawah ini merupakan variabel yang digunakan.

Tabel 3.1 Variabel Penelitian

Indikator	Variabel	Keterangan	Satuan
Pendidikan	X ₁	Harapan lama sekolah	Tahun
	X ₂	Rata-rata lama sekolah	Tahun
Umur panjang dan hidup sehat	X ₃	Angka harapan hidup	Tahun
Standar hidup layak	X ₄	Pengeluaran perkapita disesuaikan	Ribu Rupiah

3.3 Struktur Data

Struktur data yang digunakan dalam penelitian ini adalah sebagai berikut.

Tabel 3.2 Struktur Data Penelitian

Kab/Kota	X_1	X_2	X_3	X_4
1	$X_{1,1}$	$X_{2,1}$	$X_{3,1}$	$X_{4,1}$
2	$X_{1,2}$	$X_{2,2}$	$X_{3,2}$	$X_{4,2}$
3	$X_{1,3}$	$X_{2,3}$	$X_{3,3}$	$X_{4,3}$
4	$X_{1,4}$	$X_{2,4}$	$X_{3,4}$	$X_{4,4}$
5	$X_{1,5}$	$X_{2,5}$	$X_{3,5}$	$X_{4,5}$
\vdots	\vdots	\vdots	\vdots	\vdots
29	$X_{1,29}$	$X_{2,29}$	$X_{3,29}$	$X_{4,29}$

3.4 Langkah Analisis

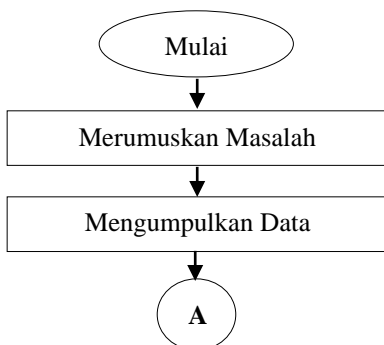
Langkah analisis digunakan untuk menggambarkan langkah-langkah penelitian yang akan dilakukan secara urut. Langkah analisis yang digunakan adalah sebagai berikut.

1. Merumuskan masalah.
2. Mengumpulkan data.
3. Melakukan eksplorasi data untuk mengetahui karakteristik dari kabupaten/kota Provinsi Papua berdasarkan indikator Indeks Pembangunan Manusia tahun 2018.
4. Melakukan pengelompokan dengan metode *c-means* dengan *software* SPSS dengan langkah sebagai berikut.
 - a. Menentukan jumlah *cluster* yang akan dibentuk
 - b. Mengalokasikan data ke masing-masing *cluster* secara *random*
 - c. Menentukan *centroid* dari data yang ada pada masing-masing *cluster*
 - d. Mengelompokkan data ke *centroid* terdekat menggunakan jarak euclidean
 - e. Mengulangi langkah 4.c ketika terdapat perpindahan data dari satu *cluster* ke *cluster* yang lainnya.
5. Melakukan pengelompokan dengan metode *fuzzy c-means* menggunakan *software* R dengan langkah sebagai berikut.
 - a. Menentukan jumlah *cluster* yang akan dibentuk
 - b. Menentukan nilai pembobot (m)
 - c. Menentukan batas toleransi atau kriteria penghentian perulangan (ε)

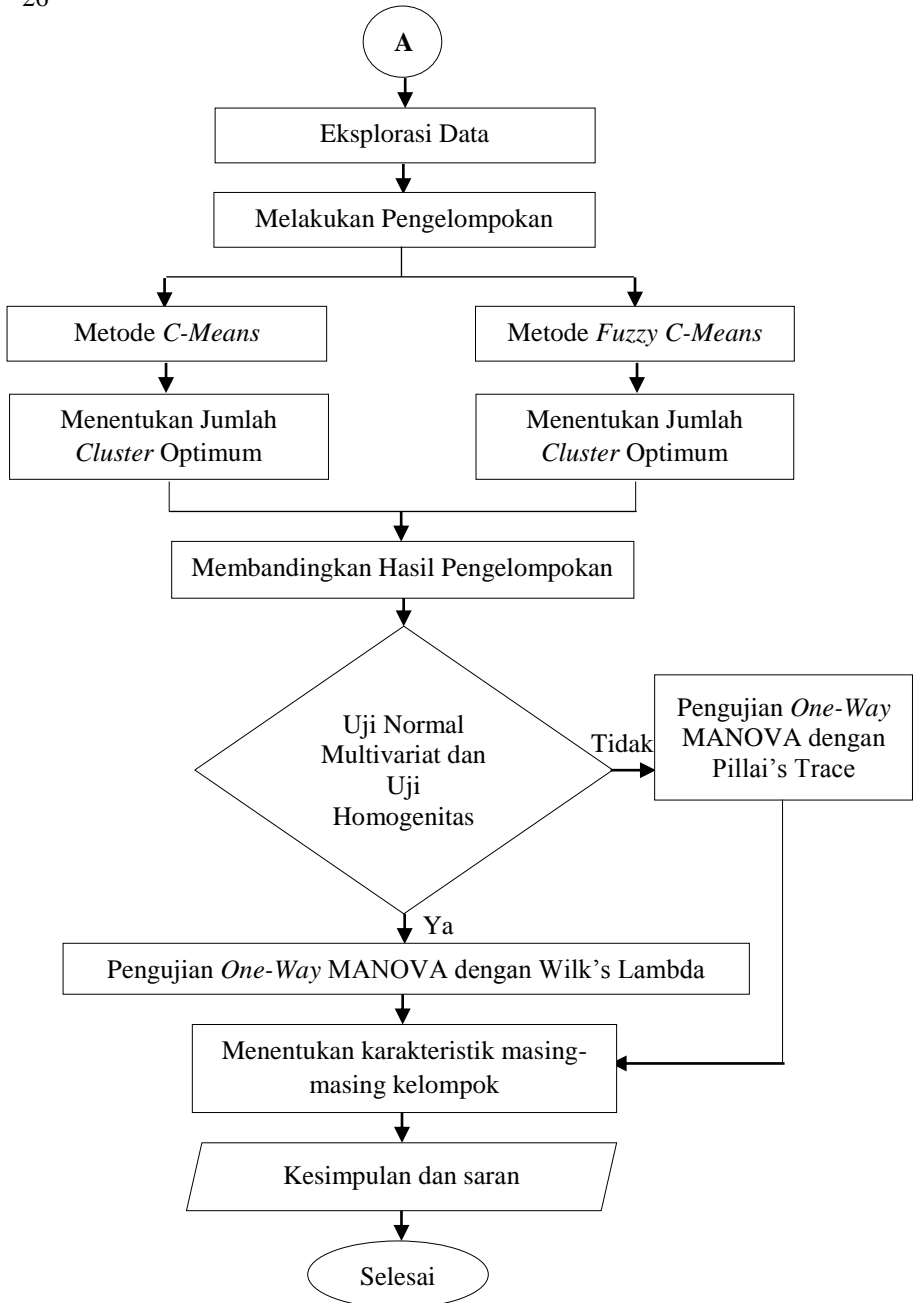
- d. Membentuk matriks partisi awal U yang merupakan bilangan *random* antara 0 sampai dengan 1
 - e. Menghitung pusat dari masing-masing *cluster*
 - f. Menghitung fungsi objektif
 - g. Apabila nilai $P < \varepsilon$ maka proses dapat dihentikan, namun jika tidak dilanjutkan dengan menghitung matriks partisi U yang baru (*membership function*) dan mengulangi langkah 5.e.
6. Menentukan jumlah *cluster* optimum pada kedua metode melalui nilai *pseudo f-statistic* dengan *software R*.
 7. Membandingkan hasil pengelompokan kedua metode dengan perhitungan nilai *icdrate* dengan *software R*.
 8. Melakukan pengujian asumsi normal multivariat menggunakan *Minitab* dan uji homogenitas dengan *software SPSS*.
 9. Melakukan pengujian perbedaan karakteristik antar kelompok dengan *one-way MANOVA* dan *one-way ANOVA* menggunakan *software SPSS*.
 10. Menentukan karakteristik pada masing-masing kelompok yang terbentuk.
 11. Menarik kesimpulan dan saran.

3.5 Diagram Alir

Diagram alir menggambarkan alur perjalanan pembuatan laporan. Diagram alir yang digunakan dalam penelitian ini adalah sebagai berikut.



Gambar 3.1 Diagram Alir



Gambar 3.1 Diagram Alir (lanjutan)

BAB IV ANALISIS DAN PEMBAHASAN

Bab ini akan membahas mengenai hasil analisis data untuk menjawab rumusan masalah pada penelitian ini. Hal-hal yang akan dibahas meliputi karakteristik data indikator Indeks Pembangunan Manusia di Provinsi Papua tahun 2018, melakukan pengelompokan kabupaten/kota di Provinsi Papua berdasarkan indikator IPM pada tahun 2018 dengan metode *c-means* dan *fuzzy c-means*, hingga didapatkan jumlah kelompok optimum pada masing-masing metode dan didapatkan pula metode terbaik untuk melakukan pengelompokan. Selanjutnya dilakukan deskripsi karakteristik pada masing-masing kelompok yang telah terbentuk.

4.1 Karakteristik Indikator Indeks Pembangunan Manusia di Provinsi Papua Tahun 2018

Indeks Pembangunan Manusia dibentuk berdasarkan beberapa indikator yaitu harapan lama sekolah, rata-rata lama sekolah, angka harapan hidup saat lahir, dan pengeluaran perkapita. Di bawah ini merupakan karakteristik dari masing-masing indikator.

Tabel 4.1 Karakteristik Indikator IPM di Papua Tahun 2018

Variabel	Mean	Maksimum	Minimum	Varians
X ₁	10,06	14,99	2,59	8,56
X ₂	5,89	11,30	0,85	9,23
X ₃	64,69	72,06	54,82	15,08
X ₄	6960	14922	4131	6614521

Berdasarkan Tabel 4.1 dapat diketahui karakteristik dari masing-masing indikator pembentuk IPM. Ukuran pemusatan data yang digunakan yaitu rata-rata, nilai maksimum, dan nilai minimum. Sedangkan ukuran penyebaran data yang digunakan yaitu varians. Diketahui rata-rata harapan lama sekolah di Provinsi Papua tahun 2018 sebesar 10,06, artinya secara rata-rata anak usia tujuh tahun yang masuk jenjang pendidikan formal di Papua pada tahun 2018 memiliki peluang untuk bersekolah

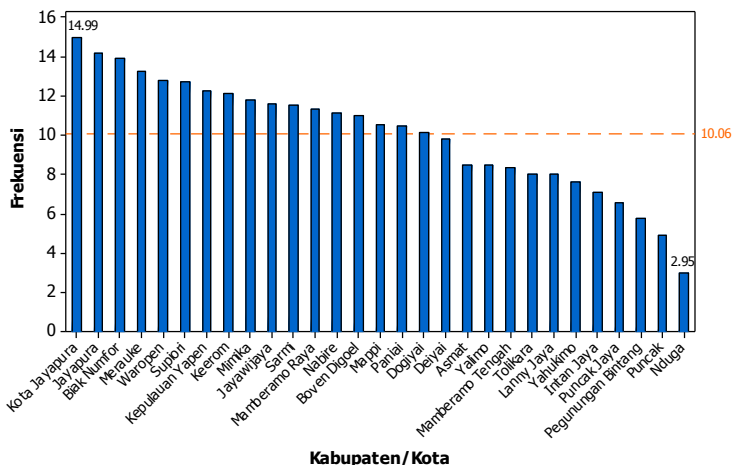
selama 10,06 tahun atau setara dengan kelas XI. Rata-rata lama sekolah memiliki *mean* sebesar 5,89 yang menandakan bahwa secara rata-rata penduduk Provinsi Papua yang berusia 25 tahun ke atas telah menempuh pendidikan selama 5,89 tahun atau hampir menamatkan kelas VI. Rata-rata harapan hidup di Papua pada tahun 2018 berdasarkan Tabel 4.1 sebesar 64,69, yang artinya secara rata-rata bayi yang baru lahir di Papua pada tahun 2018 memiliki peluang untuk bertahan hidup sampai dengan 64,69 tahun. Sedangkan pendapatan perkapita di Provinsi Papua memiliki rata-rata sebesar 6960, yang artinya secara rata-rata pengeluaran penduduk Papua selama setahun adalah Rp.6.960.000. Keragaman terendah yaitu variabel harapan lama sekolah sebesar 8,56, sedangkan keragaman tertinggi pada variabel pengeluaran perkapita sebesar 6614521 yang menandakan bahwa pendapatan perkapita di Provinsi Papua tidak merata. Deskripsi atau informasi mengenai nilai maksimum dan minimum ditampilkan pada uraian di bawah ini.

4.1.1 Indikator Pendidikan

Indikator pendidikan digambarkan oleh harapan lama sekolah dan rata-rata lama sekolah.

a. Harapan Lama Sekolah

Harapan lama sekolah merupakan angka untuk mengetahui lamanya sekolah (dalam tahun) yang dapat dirasakan oleh anak pada usia tertentu di masa mendatang. Menghitung harapan lama sekolah mencakup pendidikan dari penduduk usia 7 tahun ke atas. Provinsi Papua terdiri dari 29 kabupaten/kota. Deskripsi dari 29 kabupaten/kota di Papua berdasarkan harapan lama sekolah adalah sebagai berikut.

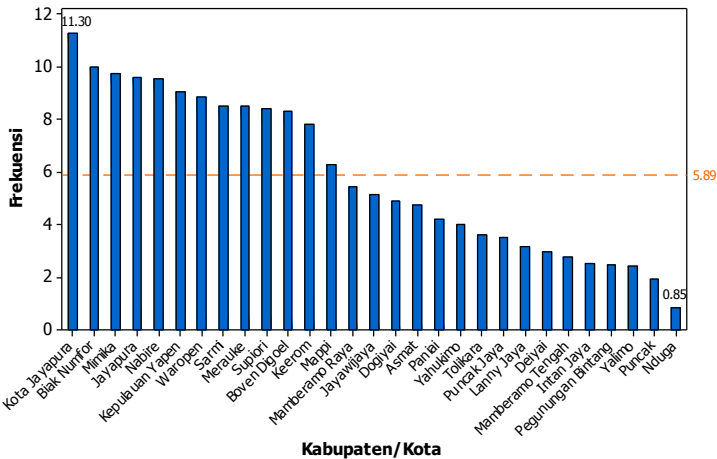


Gambar 4.1 Harapan Lama Sekolah Kabupaten/Kota di Papua Tahun 2018

Berdasarkan Gambar 4.1 diketahui bahwa terdapat 17 kabupaten/kota di Papua yang memiliki harapan lama sekolah di atas rata-rata, sedangkan 12 kabupaten lainnya memiliki harapan lama sekolah di bawah rata-rata. 12 kabupaten tersebut yaitu Kabupaten Puncak Jaya, Asmat, Yahukimo, Pegunungan Bintang, Tolikara, Nduga, Lanny Jaya, Mamberamo Tengah, Yalimo, Puncak, Intan Jaya, dan Deiyai. Kabupaten/kota dengan harapan lama sekolah tertinggi adalah Kota Jayapura, yaitu sebesar 14,99. Sedangkan kabupaten/kota dengan harapan lama sekolah terendah adalah Kabupaten Nduga yaitu sebesar 2,95.

b. Rata-Rata Lama Sekolah

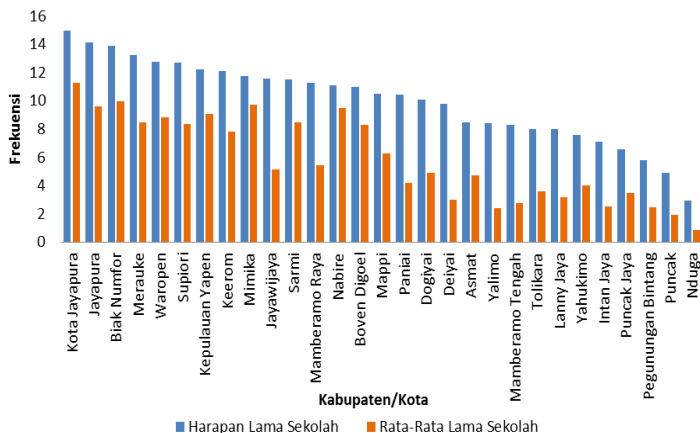
Rata-rata lama sekolah digunakan untuk mengetahui rata-rata lamanya (tahun) penduduk usia 25 tahun ke atas dalam menjalani pendidikan formal. Di bawah ini merupakan deskripsi dari 29 kabupaten/kota di Papua berdasarkan rata-rata lama sekolah.



Gambar 4.2 Rata-Rata Lama Sekolah Kabupaten/Kota di Papua Tahun 2018

Pada Gambar 4.2 menunjukkan bahwa di Papua terdapat 13 kabupaten/kota yang memiliki rata-rata lama sekolah di atas rata-rata dan 16 kabupaten/kota lainnya berada dibawah rata-rata. Banyaknya kabupaten/kota yang berada di bawah rata-rata menandakan rata-rata pendidikan yang ditempuh masih tergolong rendah. Kabupaten/kota dengan rata-rata lama sekolah tertinggi adalah Kota Jayapura, yaitu sebesar 11,30 yang artinya secara rata-rata penduduk Kota Jayapura yang berusia 25 tahun ke atas telah menempuh pendidikan selama 11,30 tahun atau setara dengan kelas XII. Sedangkan kabupaten/kota dengan rata-rata lama sekolah terendah adalah Kabupaten Nduga, yaitu sebesar 0,85 yang artinya secara rata-rata penduduk Kabupaten Nduga yang berusia 25 tahun ke atas telah menempuh pendidikan selama 0,85 tahun atau hampir menamatkan kelas I.

Selain itu akan diketahui nilai harapan lama sekolah akan sebanding atau tidak dengan nilai rata-rata lama sekolah berdasarkan gambar berikut.

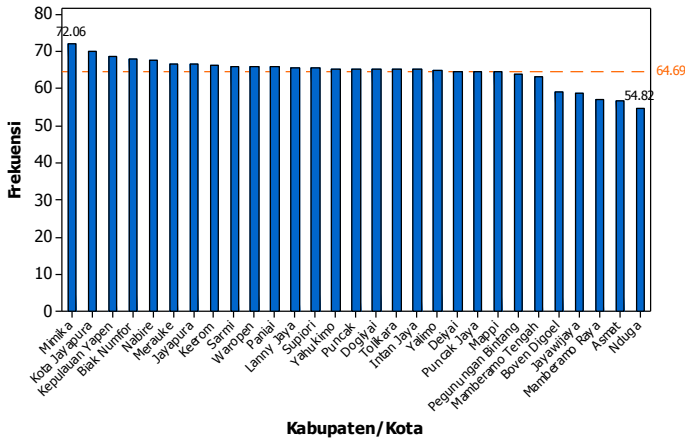


Gambar 4.3 Perbandingan Harapan Lama Sekolah dengan Rata-Rata Lama Sekolah

Berdasarkan Gambar 4.3 diketahui bahwa tingginya harapan lama sekolah tidak menjadikan nilai rata-rata lama sekolah semakin tinggi dan sebaliknya. Kabupaten Jayawijaya, Mamberamo Raya, Paniai, Deiyai, Yalimo, Mamberamo Tengah, Tolikara, Lanny Jaya, Intan Jaya, Pegunungan Bintang, Puncak, dan Nduga memiliki rata-rata lama sekolah yang sangat sedikit jika dibandingkan dengan harapan lama sekolah yang dapat ditempuh, dimana rata-rata lama sekolahnya tidak mencapai setengah dari nilai harapan lama sekolahnya.

4.1.2 Indikator Umur Panjang dan Hidup Sehat

Umur panjang dan hidup sehat digambarkan oleh angka harapan hidup. Angka harapan hidup merupakan angka yang digunakan untuk mengetahui jumlah tahun yang diharapkan dapat ditempuh oleh seseorang sejak lahir. Deskripsi dari 29 kabupaten/kota di Papua berdasarkan harapan lama sekolah adalah sebagai berikut.

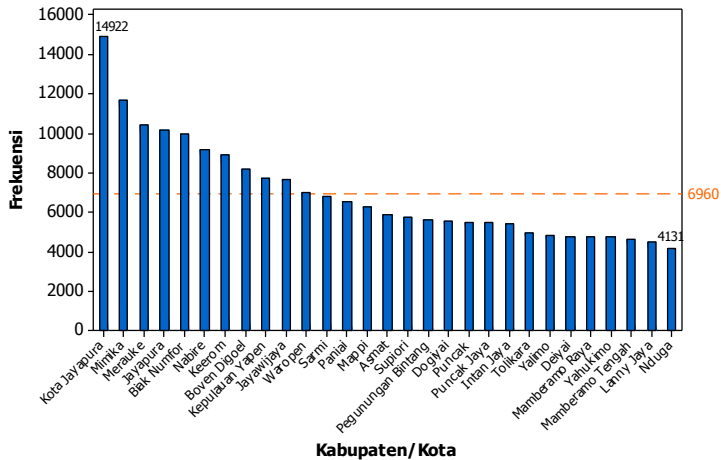


Gambar 4.4 Angka Harapan Hidup Kabupaten/Kota di Papua Tahun 2018

Berdasarkan Gambar 4.4 dapat dilihat bahwa dari 29 kabupaten/kota di Papua, terdapat 20 kabupaten/kota yang memiliki angka harapan hidup di atas rata-rata, dan 9 kabupaten/kota lainnya berada di bawah rata-rata. 9 kabupaten yang memiliki angka harapan hidup di bawah rata-rata yaitu Kabupaten Jayawijaya, Puncak Jaya, Boven Digoel, Mappi, Asmat, Pegunungan Bintang, Mamberamo Raya, Nduga, dan Mamberamo Tengah. Kabupaten Mimika sebagai kabupaten dengan angka harapan hidup tertinggi di Papua, yaitu sebesar 72,06. Sedangkan kabupaten dengan angka harapan hidup terendah sebesar 54,82 pada Kabupaten Nduga.

4.1.3 Indikator Standar Hidup Layak

Standar hidup layak digambarkan oleh pengeluaran perkapita disesuaikan. Pengeluaran perkapita disesuaikan menunjukkan biaya yang dikeluarkan untuk konsumsi yang ditentukan dari nilai pengeluaran perkapita dan paritas daya beli. Di bawah ini merupakan deskripsi dari 29 kabupaten/kota di Papua berdasarkan pengeluaran perkapita.



Gambar 4.5 Pengeluaran Perkapita Kabupaten/Kota di Papua Tahun 2018

Berdasarkan Gambar 4.5 diperoleh informasi bahwa rata-rata pengeluaran perkapita disesuaikan di Papua tahun 2018 sebesar 6960 dengan 11 kabupaten/kota memiliki pengeluaran perkapita di atas rata-rata dan 18 kabupaten/kota lainnya berada di bawah rata-rata. Kota Jayapura memiliki pengeluaran perkapita tertinggi yaitu sebesar 14922, yang artinya secara rata-rata pengeluaran penduduk Kota Jayapura selama setahun adalah Rp.14.922.000. Sedangkan kabupaten dengan pengeluaran perkapita terendah pada Kabupaten Nduga sebesar 4131, yang artinya secara rata-rata pengeluaran penduduk Kabupaten Nduga selama setahun adalah Rp.4.131.000.

4.2 Pengelompokan dengan Metode *C-Means Clustering*

Pengelompokan kabupaten/kota di Provinsi Papua berdasarkan indikator IPM yang pertama dilakukan menggunakan metode *c-means*. Pada pengelompokan dengan metode *c-means* dilakukan penentuan jumlah *cluster* yang akan dibentuk secara random terlebih dahulu, yang kemudian dilanjutkan dengan perhitungan *centroid* data ke masing-masing *cluster*. Penentuan

kelompok dilakukan dengan mengalokasikan data berdasarkan jarak terdekat antara data dengan *centroid*. Terdapat 29 kabupaten/kota di Papua yang akan dikelompokkan pada beberapa *cluster*. Pada penelitian ini dilakukan pengelompokan dengan menggunakan jumlah *cluster* sebanyak 2 hingga 5 *cluster* menggunakan *software* SPSS. Berikut merupakan hasil pengelompokan pada masing-masing *cluster* yang terbentuk.

Tabel 4.2 Pengelompokan dengan Metode *C-Means*

Kelompok ke-	Jumlah <i>Cluster</i>			
	2	3	4	5
1	7	2	6	1
2	22	19	7	6
3		8	15	15
4			1	5
5				2

Tabel 4.2 menunjukkan jumlah anggota (kabupaten/kota) tiap kelompok pada hasil pengelompokan dari 29 kabupaten/kota di Papua dengan 2 hingga 5 *cluster*. Jumlah *cluster* yang optimum dapat ditentukan menggunakan nilai *pseudo f-statistic* yang memiliki nilai paling besar diantara 2 hingga 5 *cluster* tersebut. Di bawah ini merupakan nilai *pseudo f-statistic* pada masing-masing *cluster*.

Tabel 4.3 Nilai *Pseudo F-Statistic* Metode *C-Means*

Jumlah <i>Cluster</i>	<i>Pseudo F-Statistic</i>
2	67,82766
3	81,74019
4	121,6873
5	117,9964

Berdasarkan perhitungan nilai *pseudo f-statistic* pada Tabel 4.3 dengan 2 hingga 5 *cluster* menggunakan metode *c-means* didapatkan *cluster* optimum untuk mengelompokkan kabupaten/kota di Papua berdasarkan indikator Indeks Pembangunan Manusia yaitu sebanyak 4 *cluster*. Melakukan pengelompokan dengan metode *c-means* dilakukan dengan penentuan jumlah *cluster* yang diinginkan terlebih dahulu. Contoh

pada jumlah *cluster* sebanyak 4 ($c=4$). Setelah menentukan jumlah *cluster*, dilanjutkan dengan pengalokasian data keempat *cluster* secara *random*. Kemudian menghitung *centroid* atau pusat *cluster* berdasarkan rumus pada persamaan (2.1). Selanjutnya dilanjutkan dengan pengalokasian data berupa kabupaten/kota di Provinsi Papua dengan menggunakan jarak terdekat. Jarak terdekat dihitung berdasarkan rumus pada persamaan (2.2), sehingga didapatkan anggota pada keempat *cluster*. Ketika anggota pada setiap *cluster* sudah terbentuk, maka dilihat keanggotaan yang baru terdapat perbedaan atau tidak dengan penentuan anggota masing-masing *cluster* berdasarkan hasil *random* sebelumnya. Jika terdapat perbedaan anggota, maka kembali ke langkah perhitungan *centroid* atau pusat *cluster* yang baru hingga tidak terjadi perpindahan kabupaten/kota Provinsi Papua dari satu *cluster* ke *cluster* lainnya. Perhitungan secara manual untuk jumlah *cluster* optimum sebanyak 4 *cluster* pada metode *c-means* adalah sebagai berikut.

$$SST = \sum_{i=1}^n \sum_{c=1}^C \sum_{j=1}^J (x_{ic}^j - \bar{x}^j)^2$$

$$SST = 10,11459322 + 16,8949346 + \dots + 63396189,55$$

$$SST = 185207510,4$$

$$SSW = \sum_{i=1}^n \sum_{c=1}^C \sum_{j=1}^J (x_{ic}^j - \bar{x}_c^j)^2$$

$$SSW = 0,256711111 + 2,064011111 + \dots + 110933,4044$$

$$SSW = 11870388$$

$$R^2 = \frac{(SST - SSW)}{SST} = \frac{185207510,4 - 11870388}{185207510,4} = 0,9359$$

$$\text{Pseudo } F_{(4\text{kelompok})} = \frac{\left[\frac{R^2}{(c-1)} \right]}{\left[\frac{(1-R^2)}{(n-c)} \right]} = \frac{\left[\frac{0,9359}{(4-1)} \right]}{\left[\frac{(1-0,9359)}{(29-4)} \right]} = 121,6873$$

Hasil dari pengelompokan sebanyak 4 *cluster* dapat dilihat pada Tabel 4.4.

Tabel 4.4 Daftar Kabupaten/Kota pada 4 *Cluster*

Kelompok 1	Merauke	Jayapura	Nabire
	Biak Numfor	Mimika	Keerom
Kelompok 2	Jayawijaya	Kep. Yapen	Paniai
	Boven Digoel	Mappi	Sarmi
	Waropen		
Kelompok 3	Puncak Jaya	Asmat	Yahukimo
	Peg. Bintang	Tolikara	Supiori
	Mamberamo Jaya	Nduga	Lanny Jaya
	Dogiyai	Yalimo	Puncak
	Mamberamo Tengah	Intan Jaya	Deiyai
Kelompok 4	Kota Jayapura		

4.3 Pengelompokan dengan Metode *Fuzzy C-Means Clustering*

Metode selanjutnya yang digunakan untuk melakukan pengelompokan kabupaten/kota di Provinsi Papua berdasarkan indikator IPM yaitu menggunakan metode *fuzzy c-means clustering*. Pada metode ini terdapat *membership function* yang menunjukkan seberapa besar kemungkinan suatu data dapat menjadi anggota pada suatu *cluster*. Sebelum melakukan pengelompokan dengan metode *fuzzy c-means* terlebih dahulu menentukan jumlah *cluster* yang akan dibentuk, nilai awal pembobot sebagai bobot dari *membership function*, dan toleransi perulangan. Pada analisis ini nilai awal pembobot (m) sebesar 2 dan batas toleransi atau kriteria perulangan sebesar 10^{-6} . Jumlah kabupaten/kota di Papua sebanyak 29 dan akan dikelompokkan ke dalam beberapa *cluster*. Pengelompokan dilakukan dengan menggunakan jumlah *cluster* sebanyak 2 hingga 5 *cluster* dan akan dipilih *cluster* yang paling optimum menggunakan *software R*. Hasil pengelompokan pada masing-masing *cluster* ditampilkan pada Tabel 4.5.

Tabel 4.5 Pengelompokan dengan Metode *Fuzzy C-Means Clustering*

Kelompok ke-	Jumlah <i>Cluster</i>			
	2	3	4	5
1	21	9	1	5
2	8	19	6	10
3		1	15	1
4			7	8
5				5

Jumlah anggota (kabupaten/kota) tiap kelompok pada hasil pengelompokan dari 29 kabupaten/kota di Papua dengan 2 hingga 5 *cluster* menggunakan *fuzzy c-means* dapat dilihat pada Tabel 4.5. Penentuan *cluster* optimum dapat ditentukan menggunakan nilai *pseudo f-statistic* yang memiliki nilai paling besar diantara 2 hingga 5 *cluster* tersebut. Di bawah ini merupakan nilai *pseudo f-statistic* pada masing-masing *cluster*.

Tabel 4.6 Nilai *Pseudo F-Statistic* Metode *Fuzzy C-Means*

Jumlah <i>Cluster</i>	<i>Pseudo F-Statistic</i>
2	69,08024
3	77,58484
4	121,6873
5	130,4365

Table 4.6 menunjukkan perhitungan nilai *pseudo f-statistic* dengan 2 hingga 5 *cluster* menggunakan metode *fuzzy c-means*. Jumlah *cluster* optimum yang terpilih untuk mengelompokkan kabupaten/kota di Papua berdasarkan indikator Indeks Pembangunan Manusia yaitu sebanyak 5 *cluster*. Hal ini dapat dilihat pada nilai *pseudo f-statistic* pada 5 *cluster* yang merupakan nilai tertinggi yaitu sebesar 130,4365. Melakukan pengelompokan dengan metode *fuzzy c-means* dilakukan dengan penentuan jumlah *cluster* yang diinginkan terlebih dahulu, contoh pada jumlah *cluster* sebanyak 5 ($c=5$). Selain itu juga menentukan nilai awal pembobot (m) sebesar 2 dan batas toleransi atau kriteria perulangan sebesar 10^{-6} . Kemudian dilakukan pembangkitan bilangan *random* antara 0 sampai dengan 1 sebagai elemen matriks partisi awal U dengan baris sebanyak 29 yang sesuai dengan jumlah kabupaten/kota di Provinsi Papua dan kolom

sebanyak 5 sesuai dengan banyaknya *cluster* yang akan dibentuk. Setelah itu menentukan *centroid* pada masing-masing *cluster* dan dilanjutkan dengan menghitung nilai fungsi objektif. Nilai fungsi objektif digunakan untuk penentuan iterasi akan berlanjut atau berhenti setelah dibandingkan dengan nilai kriteria penolakan. Iterasi berhenti ketika nilai fungsi objektif dibawah batas toleransi. Namun, ketika fungsi objektif memiliki nilai yang lebih besar dari batas toleransi sebesar 10^{-6} maka dilakukan perhitungan matriks partisi U yang baru (*membership function*) menggunakan rumus pada persamaan (2.4). Kemudian dilanjutkan dengan perhitungan nilai *centroid* dan fungsi objektif kembali hingga didapatkan nilai fungsi objektif kurang dari 10^{-6} . Ketika fungsi objektif kurang dari 10^{-6} maka iterasi berhenti dan dilanjutkan dengan menentukan anggota pada kelima *cluster* dengan melihat matriks derajat keanggotaan (*membership function*) pada iterasi terakhir. Misalnya pada kabupaten pertama yang memiliki derajat keanggotaan pada kelima *cluster*, kabupaten tersebut memiliki derajat keanggotaan terbesar pada *cluster* ketiga sehingga kabupaten pertama akan menjadi anggota *cluster* ketiga dan seterusnya hingga kabupaten ke-29. Perhitungan secara manual untuk jumlah *cluster* optimum pada metode *fuzzy c-means* adalah sebagai berikut.

$$SST = \sum_{i=1}^n \sum_{c=1}^C \sum_{j=1}^J (x_{ic}^j - \bar{x}^j)^2$$

$$SST = 10,11459322 + 16,8949346 + \dots + 63396189,55$$

$$SST = 185207510,4$$

$$SSW = \sum_{i=1}^n \sum_{c=1}^C \sum_{j=1}^J (x_{ic}^j - \bar{x}_c^j)^2$$

$$SSW = 0,150544 + 0,015376 + \dots + 1043257,96 + 843825,96$$

$$SSW = 8144775,9$$

$$R^2 = \frac{(SST - SSW)}{SST} = \frac{185207510,4 - 8144775,9}{185207510,4} = 0,9560$$

$$\text{Pseudo } F_{(5\text{kelompok})} = \frac{\left[\frac{R^2}{(c-1)} \right]}{\left[\frac{(1-R^2)}{(n-c)} \right]} = \frac{\left[\frac{0,9560}{(5-1)} \right]}{\left[\frac{(1-0,9560)}{(29-5)} \right]} = 130,4365$$

Hasil dari pengelompokan sebanyak 5 *cluster* dapat dilihat pada Tabel 4.7.

Tabel 4.7 Daftar Kabupaten/Kota pada 5 *Cluster*

Kelompok 1	Merauke	Jayapura	Nabire
	Biak Numfor	Mimika	
Kelompok 2	Paniai	Puncak Jaya	Mappi
	Asmat	Peg. Bintang	Sarmi
	Supiori	Puncak	Dogiyai
	Intan Jaya		
Kelompok 3	Kota Jayapura		
Kelompok 4	Yahukimo	Tolikara	Nduga
	Lanny Jaya	Mamberamo Raya	Yalimo
Kelompok 5	Deiyai	Mamberao Tengah	
	Jayawijaya	Kep. Yapen	Keerom
	Waropen	Boven Digoel	

4.4 Pemilihan Metode Terbaik

Setelah didapatkan jumlah *cluster* optimum pada kedua metode dengan melihat nilai *pseudo f-statistic* paling tinggi, tahapan selanjutnya yaitu mendapatkan metode terbaik yang digunakan untuk pengelompokan. Penentuan metode terbaik dilakukan dengan membandingkan metode *c-means* dan *fuzzy c-means* sehingga diperoleh metode yang paling sesuai untuk mengelompokkan kabupaten/kota di Papua berdasarkan indikator Indeks Pembangunan Manusia. Metode terbaik dapat dilihat dari nilai *icdrate*. *Icdrate* merupakan nilai yang menunjukkan tingkat dispersi antar anggota dalam satu kelompok. Hasil pengelompokan dikatakan baik ketika suatu kelompok memiliki nilai *icdrate* yang kecil. Hal ini dikarenakan perbedaan dalam kelompok semakin kecil atau homogen.

Tabel 4.8 Perbandingan Nilai *Icdrate*

Metode	Cluster Optimum	<i>Icdrate</i>
<i>C-Means</i>	4	0,0641
<i>Fuzzy C-Means</i>	5	0,0439

Tabel 4.8 menunjukkan perbandingan nilai *icdrate* pada kedua metode. Diketahui bahwa metode *fuzzy c-means* dengan jumlah *cluster* sebanyak 5 memiliki nilai *icdrate* terkecil yaitu sebesar 0,0439, sehingga metode terbaik yang terpilih untuk digunakan dalam mengelompokkan kabupaten/kota di Papua berdasarkan indikator Indeks Pembangunan Manusia yaitu metode *fuzzy c-means clustering*. Perhitungan manual nilai *icdrate* untuk kedua metode adalah sebagai berikut.

$$icdrate_{c-means} = 1 - R^2 = 1 - 0,9359 = 0,0641$$

$$icdrate_{fuzzy\ c-means} = 1 - R^2 = 1 - 0,9560 = 0,0439$$

4.5 Karakteristik Kelompok Metode Terbaik

Pada pengelompokan kabupaten/kota di Papua berdasarkan indikator Indeks Pembangunan Manusia menggunakan metode *c-means* dan *fuzzy c-means* telah disimpulkan bahwa penggunaan metode *fuzzy c-means* merupakan metode yang lebih baik digunakan untuk pengelompokan. Berdasarkan hal tersebut, diharapkan terdapat perbedaan karakteristik pada masing-masing kelompok terhadap seluruh indikator Indeks Pembangunan Manusia.

4.5.1 Uji Perbedaan Karakteristik

Mengetahui ada tidaknya perbedaan karakteristik pada kelompok yang terbentuk dapat dilakukan menggunakan metode *one-way* MANOVA dan *one-way* ANOVA. Sebelum dilakukan uji tersebut, terlebih dahulu dilakukan uji normal multivariat untuk mengetahui apakah data telah berdistribusi normal secara multivariat atau tidak serta uji homogenitas untuk mengetahui apakah varians antar kelompok homogen atau tidak.

a. Uji Distribusi Normal Multivariat

Pengujian distribusi normal multivariat digunakan untuk mengetahui apakah data berdistribusi normal multivariat atau tidak. Pada uji normal multivariat didapatkan nilai korelasi sebesar 0,980 yang terdapat pada Lampiran 8. Nilai tersebut akan dibandingkan dengan *critical point* dari tabel normal probabilitas koefisien korelasi (*normal probability plot correlation coefficient (PPCC) distribution*). *Critical point* yang diperoleh dengan menggunakan *alpha* 5% yaitu sebesar 0,9622 yang terdapat pada Lampiran 9, sehingga diperoleh keputusan gagal tolak H_0 yang artinya bahwa data mengikuti distribusi normal multivariat.

b. Pengujian Homogenitas

Pengujian homogenitas terhadap matriks varians kovarians dapat dilakukan dengan menggunakan Uji Box's M. Pengujian hipotesis Box's M pada penelitian ini adalah sebagai berikut.

Hipotesis :

$$H_0 : \Sigma_1 = \Sigma_2 = \dots = \Sigma_g = \Sigma$$

$$H_1 : \text{minimal ada satu } \Sigma_g \neq \Sigma_j \text{ untuk } g \neq j \text{ dimana } c = 1, 2, \dots, 5$$

Pengujian dilakukan dengan menggunakan tingkat signifikansi 5%. Hasil pengujian didapatkan nilai Box's M sebesar 68,138 yang terdapat pada Lampiran 10. Nilai $\chi^2_{\frac{1}{2}(g-1)p(p+1)}$

dengan *alpha* 5%, dan derajat bebas sebesar 40 didapatkan sebesar 55,758. Nilai Box's M yang diperoleh lebih besar dari $\chi^2_{\frac{1}{2}(g-1)p(p+1)}$. Sehingga diperoleh keputusan tolak H_0 dan disimpulkan bahwa matriks varians kovarians tidak homogen.

c. Pengujian One-Way MANOVA

Setelah dilakukan uji normal multivariat dan pengujian homogenitas didapatkan hasil bahwa data telah mengikuti distribusi normal secara multivariat dan memiliki matriks varians kovarians yang tidak homogen. Sehingga pengujian perbedaan karakteristik dengan menggunakan *One-Way MANOVA*

digunakan statistik uji *Pillai's Trace*. Pada analisis ini, faktor atau perlakuan yang diduga memberikan perbedaan pada variabel respon adalah kelompok yang terbentuk. Sedangkan variabel respon pada pengujian *One-Way* MANOVA yaitu variabel indikator Indeks Pembangunan Manusia.

Pada Lampiran 11 diketahui hasil pengujian *One-Way* MANOVA memiliki nilai statistik uji *Pillai's Trace* $F = 2,744$. Sedangkan nilai $F_{64,384,0.05}$ yaitu sebesar 1,344. Jika dibandingkan antara kedua nilai maka nilai F lebih besar daripada $F_{64,384,0.05}$ dan menghasilkan keputusan tolak H_0 yang artinya terdapat perbedaan pada kelompok yang terbentuk.

d. Pengujian *One-Way* ANOVA

One-way ANOVA digunakan untuk uji perbedaan pada variabel-variabel antar anggota kelompok. Hasil pengujian *one-way* ANOVA adalah sebagai berikut.

Tabel 4.9 Hasil Pengujian *One-Way* ANOVA

Variabel	F_{hitung}
Harapan lama sekolah	6,904
Rata-rata lama sekolah	15,176
Angka harapan hidup	3,284
Pengeluaran perkapita disesuaikan	130,444

Berdasarkan Tabel 4.9 diketahui masing-masing nilai F_{hitung} dari setiap variabel. Nilai tersebut akan dibandingkan dengan $F_{4,24,0.05}$ sebesar 2,78. Jika dibandingkan dengan $F_{4,24,0.05}$, terlihat bahwa pada keempat variabel memiliki nilai F_{hitung} yang lebih besar sehingga tolak H_0 yang artinya terdapat perbedaan karakteristik keempat variabel terhadap kelompok yang terbentuk sehingga keempat variabel memberikan pengaruh pada terbentuknya kelompok.

4.5.2 Perbedaan Karakteristik Antar Kelompok

Pengelompokan kabupaten/kota di Provinsi Papua berdasarkan indikator IPM didapatkan hasil bahwa metode *fuzzy c-means* merupakan metode terbaik yang digunakan untuk pengelompokan. Jumlah kelompok yang terbentuk sebanyak 5

cluster dengan hasil pengujian *one-way* MANOVA menunjukkan bahwa adanya perbedaan pada kelima kelompok yang terbentuk dan keempat variabel memberikan pengaruh perbedaan antar kelompok yang terbentuk berdasarkan hasil pengujian *one-way* ANOVA. Di bawah ini merupakan deskripsi dari masing-masing kelompok yang terbentuk.

Tabel 4.10 Perbedaan Rata-Rata Masing-Masing Variabel pada Setiap Kelompok

Kelompok	Kelompok				
	1	2	3	4	5
N	5	10	1	8	5
Harapan lama sekolah	12,852	8,829	14,99	8,05875	11,944
Rata-rata lama sekolah	9,476	4,751	11,3	3,16625	7,852
Angka harapan hidup	68,23	64,355	70,15	62,71	63,868
Pengeluaran perkapita	10280,4	5877,3	14922	4656,875	7896,6

Berdasarkan Tabel 4.9 didapatkan karakteristik (rata-rata) masing-masing kelompok yang terbentuk pada data indikator IPM di Provinsi Papua tahun 2018. Diketahui bahwa kelompok 3 secara rata-rata memiliki nilai tertinggi diantara kelompok lainnya pada masing-masing indikator. Hal ini menandakan bahwa kelompok 3 merupakan kelompok dengan nilai indikator IPM yang tinggi sehingga dapat dikatakan bahwa indikator pada kelompok tersebut sudah baik. Kelompok 1 merupakan kelompok dengan rata-rata tertinggi kedua pada setiap variabel setelah kelompok 3, namun jika dibandingkan dengan kelompok 3 terlihat bahwa rata-rata pada indikator pengeluaran perkapita disesuaikan memiliki selisih yang cukup besar sehingga perlu dilakukan perbaikan pada indikator tersebut. Setelah kelompok 1, indikator dengan rata-rata tertinggi yaitu pada kelompok 5, namun pada indikator angka harapan hidup memiliki rata-rata lebih kecil dibandingkan dengan kelompok 2. Selain itu indikator rata-rata lama sekolah dan pengeluaran perkapita juga memiliki nilai yang kecil sehingga pada kelompok 5 perbaikan perlu difokuskan pada

indikator angka harapan hidup, rata-rata lama sekolah dan pengeluaran perkapita disesuaikan. Kelompok 4 merupakan kelompok dengan rata-rata terendah pada keseluruhan indikator sehingga perbaikan perlu difokuskan pada seluruh indikator, dan kelompok dengan rata-rata terendah kedua yaitu pada kelompok 2 dimana indikator harapan lama sekolah, rata-rata lama sekolah dan pengeluaran perkapita memiliki rata-rata yang kecil sehingga perlu dilakukan penanganan atau perbaikan yang terfokus pada ketiga indikator tersebut.

Selain itu jika dilihat dari banyaknya kabupaten/kota pada masing-masing kelompok, kelompok 2 dan 4 memiliki jumlah anggota yang banyak namun memiliki rata-rata yang lebih rendah jika dibandingkan dengan kelompok lainnya, sehingga perlu dilakukan perbaikan yang lebih terfokus pada kedua kelompok tersebut.

Berdasarkan pencapaian rata-rata pada tiap variabel, maka dapat diberikan status pemeringkatan pada masing-masing kelompok kabupaten/kota yang terbentuk sebagai berikut.

Tabel 4.10 Status Pemeringkatan Kelompok

Kelompok ke-	Status
1	Indikator IPM cukup tinggi
2	Indikator IPM cukup rendah
3	Indikator IPM tinggi
4	Indikator IPM rendah
5	Indikator IPM sedang

BAB V KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan yang diperoleh berdasarkan analisis dan pembahasan yang diuraikan pada Bab IV adalah sebagai berikut.

1. Berdasarkan analisis karakteristik pada data indikator Indeks Pembangunan Manusia di Provinsi Papua diperoleh bahwa Kota Jayapura merupakan daerah yang memiliki indikator IPM paling tinggi dengan harapan lama sekolah, rata-rata lama sekolah, dan pendapatan perkapita tertinggi. Sedangkan Kabupaten Nduga merupakan daerah yang memiliki indikator IPM paling rendah dengan harapan lama sekolah, rata-rata lama sekolah, angka harapan hidup dan pendapatan perkapita terendah jika dibandingkan dengan daerah lain.
2. Pengelompokan kabupaten/kota di Provinsi Papua dengan metode *c-means* didapatkan hasil pengelompokan optimum sebanyak 4 *cluster*. Sedangkan pengelompokan dengan menggunakan metode *fuzzy c-means* menghasilkan 5 *cluster* sebagai hasil pengelompokan yang paling optimum yang didasarkan berdasarkan nilai *pseudo f-statistic* terbesar.
3. Metode terbaik untuk mengelompokkan kabupaten/kota di Provinsi Papua yaitu metode *fuzzy c-means* yang memiliki nilai *icdrate* lebih kecil daripada metode *c-means*.
4. Pada uji *one-way* MANOVA dengan statistik uji *Pillai's Trace* menunjukkan bahwa terdapat perbedaan pada kelompok yang terbentuk dan pada uji *one-way* ANOVA menunjukkan bahwa keempat variabel memberikan pengaruh terhadap perbedaan karakteristik antar kelompok. Kelompok 3 merupakan kelompok dengan indikator IPM tinggi. Kelompok 1 adalah kelompok dengan indikator IPM cukup tinggi namun perlu dilakukan perbaikan pada indikator pengeluaran perkapita disesuaikan. Kelompok 5

dengan indikator IPM sedang dimana perbaikan perlu difokuskan pada indikator angka harapan hidup, rata-rata lama sekolah dan pengeluaran perkapita disesuaikan. Kelompok 2 dengan indikator IPM cukup rendah dan perbaikan perlu difokuskan pada indikator harapan lama sekolah, rata-rata lama sekolah dan pengeluaran perkapita. Kelompok 4 adalah kelompok dengan indikator IPM rendah sehingga perlu adanya perbaikan pada masing-masing indikator.

5.2 Saran

Indeks Pembangunan Manusia (IPM) merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia (masyarakat/penduduk), dimana Provinsi Papua merupakan provinsi dengan IPM terendah di Indonesia. Sehingga saran untuk pemerintah yaitu perlu dilakukan perbaikan yang terfokus pada indikator dengan nilai yang rendah berdasarkan masing-masing kelompok.

DAFTAR PUSTAKA

- Agusta, Y. (2007). K-Means - Penerapan, Permasalahan Metode Terkait. *Jurnal Sistem Informasi*, 47-60.
- BPS. (2012). Dipetik Januari 20, 2020, dari Penduduk Indonesia Menurut Provinsi 1971, 1980, 1990, 2000, 2010: <https://www.bps.go.id/statictable/2009/02/20/1267/penduduk-indonesia-menurut-provinsi-1971-1980-1990-1995-2000-dan-2010.html>
- BPS. (2018). Indeks Pembangunan Manusia. Jakarta: Badan Pusat Statistika.
- BPS. (2019). Statistik Indonesia 2019. Jakarta: Badan Pusat Statistik.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis*. New Jersey: Prentice Hall.
- Jang, J.-S. R., Sun, C.-T., & Mizutani, E. (1997). *Neuro-Fuzzy and Soft Computing*. New York: Prentice Hall.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis Sixth Edition*. New Jersey: Prentice Hall.
- Kompas. (2019). *Pidato Jokowi Soal RAPBN dan Nota Keuangan 2020*. Diambil kembali dari <https://money.kompas.com/jeo/naskah-lengkap-pidato-jokowi-tentang-rapbn-dan-nota-keuangan-2020>
- Maharrani, A. (2019, April). *Pembangunan Manusia*. Dipetik January 2020, dari <https://beritagar.id/artikel/berita/indeks-pembangunan-manusia-naik-tapi-tak-capai-target>
- Mingoti, S. A., & Lima, J. O. (2006). Comparing SOM Neural Network with Fuzzy C-Means, C-Means and Traditional Hierarchical Clustering Algorithms. *European Journal of Operational Research*, 1742-1759.

- Orpin, A. R., & Kostylev, V. E. (2006). Towards a Statistically Valid Method of Textural Sea Floor Characterization of Benthic Habitats. *Marine Geology*, 209-222.
- Rencher, A. C. (2002). *Methods of Multivariate Analysis* (2 ed.). New York: John Wiley and Sons, Inc.
- Walpole, R. E. (1995). *Pengantar Statistika Edisi Ketiga*. (B. Sumantri, Penerj.) Jakarta: PT. Gramedia Pustaka Utama.

LAMPIRAN

Lampiran 1 Data Penelitian

No.	Kabupaten/Kota	X ₁	X ₂	X ₃	X ₄	No.	Kabupaten/Kota	X ₁	X ₂	X ₃	X ₄
1	Merauke	13.24	8.49	66.71	10430	16	Sarmi	11.55	8.52	66	6814
2	Jayawijaya	11.58	5.17	58.99	7637	17	Keerom	12.14	7.83	66.35	8918
3	Jayapura	14.17	9.6	66.66	10160	18	Waropen	12.77	8.87	65.99	6978
4	Nabire	11.14	9.53	67.72	9143	19	Supiori	12.72	8.39	65.53	5769
5	Kepulauan Yapen	12.24	9.07	68.85	7739	20	Mamberamo Raya	11.3	5.46	57.18	4755
6	Biak Numfor	13.94	10	68	9969	21	Nduga	2.95	0.85	54.82	4131
7	Paniai	10.47	4.2	65.94	6535	22	Lanny Jaya	8.01	3.18	65.79	4517
8	Puncak Jaya	6.59	3.51	64.65	5459	23	Mamberamo Tgh	8.33	2.78	63.14	4609
9	Mimika	11.77	9.76	72.06	11700	24	Yalimo	8.46	2.44	65.1	4799
10	Boven Digoel	10.99	8.32	59.16	8211	25	Puncak	4.93	1.95	65.33	5506
11	Mappi	10.53	6.29	64.56	6268	26	Dogiyai	10.13	4.91	65.32	5522
12	Asmat	8.47	4.74	56.88	5882	27	Intan Jaya	7.11	2.51	65.26	5440
13	Yahukimo	7.59	4.01	65.52	4737	28	Deiyai	9.79	2.99	64.83	4761
14	Peg. Bintang	5.79	2.49	64.08	5578	29	Kota Jayapura	14.99	11.3	70.15	14922
15	Tolikara	8.04	3.62	65.3	4946						

Lampiran 2 Hasil Pengelompokan dengan Metode *C-Means*

No.	Kabupaten/Kota	Jumlah <i>Cluster</i>				
		2	3	4	5	6
1	Merauke	1	3	1	5	3
2	Jayawijaya	2	3	2	2	5
3	Jayapura	1	3	1	4	3
4	Nabire	1	3	1	4	3
5	Kepulauan Yapen	2	3	2	2	5
6	Biak Numfor	1	3	1	4	3
7	Paniai	2	2	2	2	2
8	Puncak Jaya	2	2	3	3	2
9	Mimika	1	1	1	5	1
10	Boven Digoel	2	3	2	4	5
11	Mappi	2	2	2	2	2
12	Asmat	2	2	3	3	2
13	Yahukimo	2	2	3	3	6
14	Pegunungan Bintang	2	2	3	3	2
15	Tolikara	2	2	3	3	6
16	Sarmi	2	2	2	2	5
17	Keerom	1	3	1	4	3
18	Waropen	2	2	2	2	5
19	Supiori	2	2	3	3	2
20	Mamberamo Raya	2	2	3	3	6
21	Nduga	2	2	3	3	6
22	Lanny Jaya	2	2	3	3	6
23	Mamberamo Tengah	2	2	3	3	6
24	Yalimo	2	2	3	3	6
25	Puncak	2	2	3	3	2
26	Dogiyai	2	2	3	3	2
27	Intan Jaya	2	2	3	3	2
28	Deiyai	2	2	3	3	6
29	Kota Jayapura	1	1	4	1	4

Lampiran 3 Hasil Pengelompokan dengan Metode *Fuzzy C-Means*

No.	Kabupaten/Kota	Jumlah <i>Cluster</i>				
		2	3	4	5	6
1	Merauke	2	1	2	4	4
2	Jayawijaya	1	1	4	5	5
3	Jayapura	2	1	2	4	4
4	Nabire	2	1	2	4	3
5	Kepulauan Yapen	1	1	4	5	5
6	Biak Numfor	2	1	2	4	4
7	Paniai	1	2	4	1	5
8	Puncak Jaya	1	2	3	1	6
9	Mimika	2	1	2	4	4
10	Boven Digoel	2	1	4	5	3
11	Mappi	1	2	4	1	6
12	Asmat	1	2	3	1	6
13	Yahukimo	1	2	3	3	2
14	Pegunungan Bintang	1	2	3	1	6
15	Tolikara	1	2	3	3	2
16	Sarmi	1	2	4	1	5
17	Keerom	2	1	2	5	3
18	Waropen	1	2	4	5	5
19	Supiori	1	2	3	1	6
20	Mamberamo Raya	1	2	3	3	2
21	Nduga	1	2	3	3	2
22	Lanny Jaya	1	2	3	3	2
23	Mamberamo Tengah	1	2	3	3	2
24	Yalimo	1	2	3	3	2
25	Puncak	1	2	3	1	6
26	Dogiyai	1	2	3	1	6
27	Intan Jaya	1	2	3	1	6
28	Deiyai	1	2	3	3	2
29	Kota Jayapura	2	3	1	2	1

Lampiran 4 *Syntax* Pengelompokan dengan Metode *Fuzzy C-Means*

```
#Memasukkan data
Data = read.csv("E:/Papua/Dataolah.csv")
library(ppclust)
library(cluster)
library(fclust)
library(factoextra)

#Membentuk 2 Kelompok
set.seed(123)
res.fcm1 = fcm(Data, m=2, centers=2, nstart=1, iter.max=100,
con.val=1e-06)
as.data.frame(res.fcm1$u)
res.fcm1$func.val
res.fcm1$iter
res.fcm1$best.start
summary(res.fcm1)

#Membentuk 3 Kelompok
set.seed(123)
res.fcm2 = fcm(Data, m=2, centers=3, nstart=1, iter.max=100,
con.val=1e-06)
as.data.frame(res.fcm2$u)
res.fcm2$func.val
res.fcm2$iter
res.fcm2$best.start
summary(res.fcm2)

#Membentuk 4 Kelompok
set.seed(123)
res.fcm3 = fcm(Data, m=2, centers=4, nstart=1, iter.max=100,
con.val=1e-06)
as.data.frame(res.fcm3$u)
res.fcm3$func.val
```

Lampiran 4 *Syntax* Pengelompokan dengan Metode *Fuzzy C-Means* (lanjutan)

```

res.fcm3$iter
res.fcm3$best.start
summary(res.fcm3)

#Membentuk 5 Kelompok
set.seed(123)
res.fcm4 = fcm(Data, m=2, centers=5, nstart=1, iter.max=100,
con.val=1e-06)
as.data.frame(res.fcm4$u)
res.fcm4$func.val
res.fcm4$iter
res.fcm4$best.start
summary(res.fcm4)

```

Lampiran 5 *Syntax* Perhitungan Nilai *Pseudo F-Statistic* dan *Icdrate*

```

Data <- read.csv("E:/Data.csv", header=TRUE)
ps = function(Data, nc)
{
n = dim(Data)[1] #banyak data
p = dim(Data)[2] #kolom
X = Data[,1:(p-1)] #semua kolom kecuali kolom cluster
Group = Data[,p] #kolom cluster, kolom cluster di akhir
p = dim(X)[2]
Mean.X = matrix(ncol = p, nrow = (nc+1))
for (i in 1:nc)
{
for (j in 1:p)
{
Mean.X[i,j] = mean(X[which(Group==i),j])
Mean.X[(nc+1),j] = mean(X[,j])
}
}
}

```

Lampiran 5 *Syntax* Perhitungan Nilai *Pseudo F-Statistic* dan *Icdrate* (lanjutan)

```

}
SST = matrix(ncol=p, nrow=n)
for (i in 1:n)
{
for (j in 1:p)
{
SST[i,j] = (X[i,j] - Mean.X[(nc+1),j])^2
}
}
SST = sum(sum(SST))
SSE = matrix(ncol=p, nrow=n)
for (i in 1:n)
{
for (j in 1:p)
{
for (k in 1:nc)
{
if (Group[i]==k)
{
SSE[i,j] = (X[i,j] - Mean.X[k,j])^2
}
}
}
}
SSE = sum(sum(SSE))
Rsq = (SST-SSE)/SST
icdrate = 1-Rsq
Pseudo = (Rsq/(nc-1))/((1-Rsq)/(n-nc))
list(Rsq=Rsq, icdrate=icdrate, Pseudo=Pseudo)
}
ps(Data,2) #2 merupakan jumlah cluster

```

Lampiran 6 *Macro Minitab* Pemeriksaan Asumsi Normal Multivariat

```

Macro
NormalMultivariate X.1-X.p qc dj22
MConstant i j n p Prop Tengah
MColumn x.1-x.p xj Kali d dj2 qc Prob dj22
MMatrix MCova MCovaI xjxbar

#-- 1.1. Dapatkan Nilai dj2 --#
let n=count(x.1)
Covariance X.1-X.p MCova
print MCova
invers MCova MCovaI
do i=1:n
do j=1:p
let xj(j)=x.j(i)-mean(x.j)
enddo
copy xj xjxbar
mult MCovaI xjxbar Kali
let d=Kali*xj
let dj2(i)=sum(d)
enddo
print dj2

#-- 1.2. Dapatkan Nilai qc --#
do i=1:n
let Prob(i)=1-(n-i+0.5)/n
enddo
INVCDF Prob qc;
Chisquare p.

#-- 1.3 Buat Plot dj2 dengan qc --#
sort dj2 dj22
plot dj22*qc;
symbol.

```

Lampiran 6 *Macro Minitab* Pemeriksaan Asumsi Normal Multivariat (lanjutan)

```
#-- 2. Mencari Proporsi --#
INVCDF 0.5 Tengah;
  Chisquare p.
let Prop=0
do i=1:n
  if dj2(i)<=Tengah
    let Prop=Prop+1
  endif
enddo
let Prop=Prop/n
print Prop
name qc 'qc'
name dj22 'dj2'
endmacro
```

Lampiran 7 *Output* Nilai qj dan d_j^2

No.	Kabupaten/Kota	qc	d_j^2
1	Merauke	0.03478	0.1749
2	Jayapura	0.10622	0.2399
3	Nabire	0.18030	0.2962
4	Biak Numfor	0.25723	0.3000
5	Mimika	0.33725	0.3493
6	Paniai	0.42059	0.6303
7	Puncak Jaya	0.50756	0.6424
8	Mappi	0.59849	0.6593
9	Asmat	0.69374	0.6834
10	Pegunungan Bintang	0.79376	0.7162
11	Sarmi	0.89905	0.7194
12	Supiori	1.01019	0.8428
13	Puncak	1.12787	0.8489
14	Dogiyai	1.25291	0.8589

Lampiran 7 *Output* Nilai q_j dan d_j^2 (lanjutan)

No.	Kabupaten/Kota	q_c	d_j^2
15	Intan Jaya	1.38629	0.9022
16	Kota Jayapura	1.52921	1.0639
17	Yahukimo	1.68313	1.3485
18	Tolikara	1.84990	1.5830
19	Mamberamo Raya	2.03184	1.8343
20	Nduga	2.23201	1.8413
21	Lanny Jaya	2.45446	1.8437
22	Mamberamo Tengah	2.70479	2.0732
23	Yalimo	2.99099	2.7427
24	Deiyai	3.32510	3.1286
25	Jayawijaya	3.72644	3.5846
26	Kepulauan Yapen	4.22907	4.1946
27	Boven Digoel	4.90201	4.6666
28	Keerom	5.92366	6.9091
29	Waropen	8.12089	10.3219

Lampiran 8 *Output* Korelasi Nilai q_j dan d_j^2

Correlations: q_c , d_j^2

Pearson correlation of q_c and d_j^2 = 0.980
P-Value = 0.000

Lampiran 9 *Output* Nilai *Critical Point* untuk Uji Normal Multivariat

N	A	
	0,01	0,05
26	0,9393	0,9590
27	0,9413	0,9600
28	0,9428	0,9615
29	0,9441	0,9622
30	0,9462	0,9634
31	0,9476	0,9644
32	0,9490	0,9652

Lampiran 10 *Output Uji Homogenitas***Box's Test of Equality
of Covariance****Matrices^a**

Box's M	68.138
F	1.413
df1	30
df2	796.045
Sig.	.071

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + cluster

Lampiran 11 *Output Uji One-Way MANOVA***Multivariate Tests^a**

Effect		Value	F	Hypothesis df	Error df	Sig.
cluster	Pillai's Trace	1.255	2.744	16.000	96.000	.001
	Wilks' Lambda	.031	8.641	16.000	64.794	.000
	Hotelling's Trace	23.061	28.105	16.000	78.000	.000
	Roy's Largest Root	22.707	136.239 ^c	4.000	24.000	.000

Lampiran 12 Surat Pernyataan Data Sekunder**SURAT PERNYATAAN**

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FSAD ITS:

Nama : Rezkiiana Sunjadeva

NRP : 06211640000036

menyatakan bahwa data yang digunakan dalam Tugas Akhir/ ~~Thesis~~ ini merupakan data sekunder yang diambil dari ~~penelitian / buku/ Tugas Akhir/ Thesis/~~ publikasi lainnya yaitu:

Sumber : Website Badan Pusat Statistik Provinsi Papua

Keterangan : Provinsi Papua Dalam Angka 2019

Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui
Pembimbing Tugas Akhir



(Erma Oktania Permatasari, S.Si., M.Si.)
NIP. 19881007 201404 2 002

Surabaya, 12 Juni 2020



(Rezkiiana Sunjadeva)
NRP. 062116 4000 0036

*(coret yang tidak perlu)

(Halaman ini sengaja dikosongkan)

BIODATA PENULIS



Penulis dilahirkan di Lumajang, 11 Maret 1998 dengan nama lengkap Rezki Sunjadeva namun biasa dipanggil Rezk. Penulis merupakan anak pertama dari dua bersaudara oleh pasangan suami istri Defi Setiawan dan Iva Aquariustina. Pendidikan formal yang ditempuh oleh penulis antara lain SDN Rowokangkung 1 (2004-2010), SMPN 1 Yosowilangun (2010-2013), dan SMAN 2 Lumajang (2013-2016). Setelah lulus, penulis diterima sebagai mahasiswa Departemen Statistika ITS melalui jalur SNMPTN pada pilihan pertama dengan NRP 06211640000036. Selama masa perkuliahan, penulis aktif dalam kegiatan organisasi, kepanitiaan, dan pelatihan. Organisasi kampus yang pernah di ikuti oleh penulis adalah Himpunan Mahasiswa Statistika ITS (HIMASTA-ITS) sebagai staf Departemen Keilmiah periode 2017-2018 dan Sekretaris Departemen Keilmiah pada periode 2018-2019. Selain itu, penulis menjadi staf Keakhwatan FORSIS-ITS 38/39 dan Kabinet Departemen Keakhwatan FORSIS-ITS 39/40. Penulis juga aktif dalam kepanitiaan yang diadakan oleh HIMASTA-ITS yaitu Pekan Raya Statistika (PRS) 2018 sebagai Sie Konsumsi dan Kesehatan serta kepanitiaan dalam kegiatan yang diadakan oleh HIMASTA-ITS lainnya. Bagi pembaca yang ingin berdiskusi, memberikan saran, dan kritik mengenai Tugas Akhir ini untuk perbaikan kedepannya dapat disampaikan melalui *e-mail* rezkianadeva11@gmail.com.

(Halaman ini sengaja dikosongkan)