



**ANALISIS BIPLLOT UNTUK MEMETAKAN
KARAKTERISTIK HASIL ANALISIS *HYBRID*
*HIERARCHICAL CLUSTERING VIA MUTUAL CLUSTER***

**(Studi pada Indikator Kemiskinan Kabupaten/Kota di Jawa
Timur Tahun 2020)**

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar Sarjana Statistika

oleh:

NAJUNDA ZUKHRUFIAH SYAHDU FIRDAUS

175090501111005



PROGRAM STUDI SARJANA STATISTIKA

JURUSAN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS BRAWIJAYA

MALANG

2021

**LEMBAR PENGESAHAN SKRIPSI****ANALISIS BILOT UNTUK MEMETAKAN
KARAKTERISTIK HASIL ANALISIS *HYBRID*
HIERARCHICAL CLUSTERING VIA *MUTUAL CLUSTER*****(Studi pada Indikator Kemiskinan Kabupaten/Kota di Jawa
Timur Tahun 2020)**

oleh:

**NAJUNDA ZUKHRUFIAH SYAHDU FIRDAUS
175090501111005****Setelah dipertahankan di depan Majelis Penguji
pada tanggal 6 Juli 2021
dan dinyatakan memenuhi syarat untuk memperoleh gelar
Sarjana Statistika****Dosen Pembimbing,****Dr. Eni Sumarminingsih, S.Si., MM.****NIP. 197705152002122009**

Mengetahui,

Ketua Jurusan Statistika

Fakultas MIPA

Universitas Brawijaya

**Rahma Fitriani, S.Si, M.Sc., Ph.D.****NIP. 197503281999032001**

**LEMBAR PERNYATAAN**

Saya yang bertanda tangan di bawah ini:

Nama : Najunda Zukhrufiah Syahdu Firdaus

NIM : 175090501111005

Jurusan : Statistika

Judul Skripsi :

**ANALISIS BIPLLOT UNTUK MEMETAKAN
KARAKTERISTIK HASIL ANALISIS *HYBRID*
HIERARCHICAL CLUSTERING VIA *MUTUAL CLUSTER***

(Studi pada Indikator Kemiskinan Kabupaten/Kota di Jawa
Timur Tahun 2020)

Dengan ini menyatakan bahwa:

1. Isi dari skripsi yang saya buat adalah benar-benar karya sendiri dan tidak menjiplak karya orang lain, selain nama-nama yang termaktub di isi dan tertulis di daftar pustaka dalam skripsi ini.
2. Apabila di kemudian hari ternyata skripsi yang saya tulis terbukti hasil jiplakan, maka saya akan bersedia menanggung segala risiko.

Demikian pernyataan ini saya buat dengan segala kesadaran.

Malang, 6 Juli 2021

Yang Menyatakan,

Najunda Zukhrufiah Syahdu Firdaus

NIM. 175090501111005



**ANALISIS BILOT UNTUK MEMETAKAN
KARAKTERISTIK HASIL ANALISIS *HYBRID
HIERARCHICAL CLUSTERING VIA MUTUAL CLUSTER*
(Studi pada Indikator Kemiskinan Kabupaten/Kota di Jawa
Timur Tahun 2020)**

ABSTRAK

Laju persentase penduduk miskin di Jawa Timur menunjukkan penurunan sejak Maret 2018 hingga September 2019, kemudian meningkat pada Maret 2020 akibat munculnya pandemi Covid-19. Analisis *hybrid hierarchical clustering via mutual cluster* digunakan untuk mengelompokkan kabupaten/kota di Jawa Timur berdasarkan indikator kemiskinan. Untuk menggali informasi mengenai karakteristik setiap klaster, digunakan analisis biplot. Data yang digunakan adalah data sekunder indikator kemiskinan kabupaten/kota di Jawa Timur tahun 2020 yang diperoleh dari Badan Pusat Statistik Jawa Timur. Kemiripan antar kabupaten/kota diukur menggunakan jarak Mahalanobis dan untuk membentuk *mutual cluster* digunakan metode *average linkage*. Hasil dari penelitian adalah Jawa Timur dibagi ke dalam 5 klaster optimal dengan validitas rata-rata *silhouette* sebesar 0.18. Pada analisis biplot, proporsi keragaman kumulatif yang dihasilkan sebesar 68.5%. Kabupaten/kota di klaster 1 cenderung memiliki karakteristik kemiskinan dengan sumber air minum yang tidak layak dan bahan bakar memasak yang kurang berkembang. Kabupaten/kota di klaster 2 dan 3 memiliki karakteristik kemiskinan multidimensi yang mirip, di mana aspek yang lebih menonjol adalah bahan bakar memasak yang kurang berkembang dan kurangnya penguasaan teknologi. Kabupaten/kota di klaster 4 cenderung memiliki karakteristik kemiskinan multidimensi dengan sanitasi yang tidak layak dan kurangnya penguasaan teknologi. Kabupaten/kota di klaster 5 cenderung memiliki karakteristik kemiskinan multidimensi dengan aspek yang menonjol adalah luas bangunan tempat tinggal $< 20m^2$.

Kata Kunci: Biplot, *Hybrid Hierarchical Clustering*, Kemiskinan, Mahalanobis, *Mutual Cluster*.



BIPLOT ANALYSIS TO MAPPING CHARACTERISTICS OF HYBRID HIERARCHICAL CLUSTERING VIA MUTUAL CLUSTER ANALYSIS RESULTS

(Study on District/City Poverty Indicators in East Java in 2020)

ABSTRACT

The percentage rate of poor people in East Java showed a decline from March 2018 to September 2019, then increased in March 2020 due to the emergence of the Covid-19 pandemic. Hybrid hierarchical clustering analysis via mutual cluster is used to group districts/cities in East Java based on poverty indicators. To obtain information about the characteristics of each cluster, biplot analysis was used. The data used is secondary data on district/city poverty indicators in East Java in 2020 obtained from the Central Statistics Agency for East Java (*BPS Jawa Timur*). Similarities between districts/cities were measured using the Mahalanobis distance and to form a mutual cluster used the average linkage method. The result of the research is that East Java is divided into 5 optimal clusters with an average silhouette validity of 0.18. In the biplot analysis, the proportion of cumulative variance produced is 68.5%. Districts/cities in cluster 1 tend to have poverty characteristics with inadequate drinking water sources and underdeveloped cooking fuels. Districts/cities in clusters 2 and 3 have similar characteristics of multidimensional poverty, where the more prominent aspects are underdeveloped cooking fuels and lack of mastery of technology. Districts/cities in cluster 4 tend to have the characteristics of multidimensional poverty with inadequate sanitation and lack of mastery of technology. Districts/cities in cluster 5 tend to have multidimensional poverty characteristics with a prominent aspect being the residential building area $< 20m^2$.

Keywords: Biplot, Hybrid Hierarchical Clustering, Poverty, Mahalanobis, Mutual Cluster.



KATA PENGANTAR

Puji dan syukur penulis haturkan ke hadirat Allah SWT yang telah melimpahkan rahmat dan karunia-Nya berupa waktu, kesehatan serta pengetahuan sehingga penulis mampu menyelesaikan skripsi yang telah disusun. Shalawat serta salam selalu tercurah kepada Rasulullah Muhammad SAW atas ajaran yang beliau sampaikan dan sebagai suri tauladan sepanjang masa. Penyusunan skripsi ini tidak lepas dari berbagai dukungan, bantuan serta doa berbagai pihak. Oleh karena itu, penulis sampaikan terima kasih kepada:

1. Ibu Dr. Eni Sumarminingsih, S.Si., MM. selaku Dosen Pembimbing atas kesempatan, ilmu dan bimbingan yang telah diberikan.
2. Bapak Darmanto, S.Si., M.Si. selaku Anggota Majelis Penguji I atas kesempatan dan arahan yang telah diberikan.
3. Ibu Dr. Suci Astutik, S.Si., M.Si. selaku Anggota Majelis Penguji II atas kesempatan dan arahan yang telah diberikan.
4. Bapak Achmad Efendi, S.Si., M.Sc., Ph.D. selaku Ketua Program Studi Sarjana Statistika Universitas Brawijaya.
5. Ibu Rahma Fitriani S.Si., M.Sc., Ph.D. selaku Ketua Jurusan Statistika Universitas Brawijaya.
6. Ayah, Ibu, Adik dan segenap keluarga yang selalu memberikan dukungan dan doa.
7. Bestari, Nawang, Aini, Diah, Nurul, Deivy, Anggun, Luluk dan teman-teman kos yang telah menemani selama proses perkuliahan, memberi bantuan dan dukungan kepada penulis.
8. Segenap keluarga besar FORKALAM dan Divisi EO Studio Statistika 2019 atas ilmu dan pengalaman yang sangat berharga.
9. Teman-teman Jurusan Statistika Universitas Brawijaya angkatan 2017 yang telah banyak memberikan masukan dan dukungan.
10. Semua pihak yang turut membantu kelancaran dalam penyelesaian skripsi.

Penulis menyadari masih terdapat kekurangan dalam naskah skripsi ini. Oleh karena itu, berbagai saran dan masukan yang membangun sangat diharapkan demi perbaikan naskah skripsi. Semoga skripsi ini bermanfaat bagi para pembaca.

Malang, 18 Juni 2021

Penulis



DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Tujuan Penelitian	5
1.4. Manfaat Penelitian	5
1.5. Batasan Masalah	5
BAB II TINJAUAN PUSTAKA	7
2.1. Statistika Deskriptif.....	7
2.2. Analisis Kluster.....	8
2.2.1. Prosedur Hierarki	8
2.2.2. Prosedur Nonhierarki.....	9
2.3. Ukuran Jarak Mahalanobis.....	10
2.4. <i>Hybrid Hierarchical Clustering via Mutual Cluster</i>	11
2.5. Validasi Kluster.....	12
2.6. Analisis Biplot	13
2.7. Indikator Kemiskinan.....	15
BAB III METODE PENELITIAN	17
3.1. Data.....	17
3.2. Variabel Penelitian.....	17
3.3. Metode Analisis Data.....	18
BAB IV HASIL DAN PEMBAHASAN	21
4.1. Analisis Deskriptif	21
4.2. <i>Hybrid Hierarchical Clustering via Mutual Cluster</i>	23
4.2.1. Pembentukan <i>Mutual Cluster</i>	23
4.2.2. <i>Hybrid Hierarchical Clustering</i>	25
4.3. Validasi Kluster.....	26
4.4. Analisis Biplot	29
4.4.1. Biplot <i>Column Metric Preserving</i>	29



4.4.2. Biplot <i>Row Metric Preserving</i>	30
4.4.3. Biplot <i>Square Root</i>	31
4.4.4. Proporsi Keragaman Kumulatif Biplot	32
4.5. Identifikasi Karakteristik Setiap Kluster	32
BAB V PENUTUP	35
5.1. Kesimpulan	35
5.2. Saran	35
DAFTAR PUSTAKA	37
LAMPIRAN	39



DAFTAR GAMBAR

Halaman

Gambar 2.1. Persentase Penduduk Miskin Jawa Timur 2018-2020, 16

Gambar 3.1. Diagram Alir..... 20

Gambar 4.1. Dendogram *Mutual Cluster*..... 24

Gambar 4.2. *Cluster* Dendogram..... 25

Gambar 4.3. Dendogram 5 Klaster Optimal 27

Gambar 4.4. Peta Klaster Kemiskinan Jawa Timur 2020..... 28

Gambar 4.5. Biplot Kemiskinan Kabupaten/Kota di Jawa Timur-
dengan $\alpha = 0$ 29

Gambar 4.6. Biplot Kemiskinan Kabupaten/Kota di Jawa Timur-
dengan $\alpha = 1$ 30

Gambar 4.7. Biplot Kemiskinan Kabupaten/Kota di Jawa Timur-
dengan $\alpha = 0.5$ 31



DAFTAR TABEL

Halaman

Tabel 1.1. Penelitian Terdahulu.....	3
Tabel 3.1. Variabel Penelitian.....	17
Tabel 4.1. Statistik Deskriptif Indikator Kemiskinan di Jawa Timur- Tahun 2020 (%).....	21
Tabel 4.2. Anggota <i>Mutual Cluster</i>	24
Tabel 4.3. Rata-Rata Indeks <i>Silhouette</i> 2 Sampai 5 Klaster.....	26
Tabel 4.4. Anggota Klaster Optimal.....	27
Tabel 4.5. Rata-rata Setiap Klaster.....	33



DAFTAR LAMPIRAN

Halaman

Lampiran 1. Data Indikator Kemiskinan Kabupaten/Kota di Jawa-Timur Tahun 2020 (%)	39
Lampiran 2. Statistik Deskriptif Indikator Kemiskinan Provinsi Jawa Timur	42
Lampiran 3. Dendogram <i>Mutual Cluster</i>	43
Lampiran 4. Dendogram <i>Hybrid Hierarchical Clustering via Mutual Cluster</i>	44
Lampiran 5. Plot Indeks <i>Silhouette</i> 2-15 Klaster	45
Lampiran 6. Dendogram dengan 5 Klaster Optimal	52
Lampiran 7. Keanggotaan Klaster	53
Lampiran 8. Peta Klaster Kemiskinan Jawa Timur Tahun 2020	54
Lampiran 9. Nilai Akar Ciri	55
Lampiran 10. Biplot ($\alpha = 1, 0$ dan 0.5)	56
Lampiran 11. <i>Source Code</i> Analisis <i>Hybrid Hierarchical Clustering via Mutual Cluster</i>	58
Lampiran 12. <i>Source Code</i> Analisis Biplot	60



BAB I PENDAHULUAN

1.1. Latar Belakang

Analisis kluster merupakan salah satu analisis multivariat untuk mengelompokkan objek berdasarkan variabel-variabel. Terdapat dua prosedur analisis kluster yaitu prosedur nonhierarki dan hierarki. Prosedur hierarki bersifat *bottom-up* dan *top-down*. Algoritma *bottom-up* memiliki kelebihan dalam segi kecepatan dan baik untuk mengidentifikasi data berukuran kecil, sedangkan *top-down* algoritma yang baik untuk data berukuran besar. Pada tahun 2006, Chipman dan Tibshirani memperkenalkan prosedur pengelompokan baru yang menggabungkan kelebihan algoritma *bottom-up* dan *top-down*. Prosedur ini dikenal dengan *hybrid hierarchical clustering*. Pada dasarnya prosedur ini juga menggabungkan algoritma *bottom-up* dan *top-down* dengan mempertahankan *mutual cluster*.

Hasil analisis kluster berupa sejumlah kluster di mana objek-objek di dalam suatu kluster bersifat homogen atau lebih mirip dibandingkan dengan objek di luar kluster, maka dari itu perlu dikaji karakteristik setiap kluster. Salah satu metode yang umum digunakan untuk mengkaji karakteristik setiap kluster, yaitu mencari rata-rata setiap kluster berdasarkan variabel. Penggunaan rata-rata cukup baik untuk mewakili karakteristik kluster, namun ada informasi lain yang kurang terjelaskan seperti keragaman dan hubungan variabel dengan objek. Untuk menggali informasi lebih mengenai karakteristik setiap kluster dapat digunakan analisis biplot. Biplot merupakan peta dua dimensi yang menggambarkan karakteristik dan hubungan dua unsur pengamatan (objek dan variabel) secara simultan. Walaupun demikian, biplot tidak menghasilkan informasi lengkap (100%) dari data karena prinsipnya yang mereduksi ruang berdimensi besar ke ruang berdimensi dua.

Kemiskinan adalah salah satu masalah pembangunan global yang pasti dialami oleh semua negara. Untuk menangani hal tersebut, sejak September 2015 sebanyak 193 negara termasuk Indonesia menyepakati agenda pembangunan bersama dalam forum Perserikatan Bangsa-Bangsa (PBB). Tujuan yang dicanangkan dalam agenda tersebut dikenal dengan *Sustainable Development Goals* (SDGs) yang berisi 17 tujuan dan 169 sasaran berlaku mulai tahun



2016 hingga 2030. Berdasarkan SDGs, “tanpa kemiskinan” merupakan tujuan pembangunan bersama yang pertama dengan mengakhiri kemiskinan dalam segala bentuk di manapun. Menurut Bappenas (2020), penghilangan kemiskinan dan kelaparan akan menjadi “tulang punggung” dari tujuan agenda pembangunan berkelanjutan. Dengan demikian Bappenas menempatkan kemiskinan dengan melihat dari berbagai aspek atau dimensi, yang dalam ilmu ekonomi pembangunan dikenal dengan istilah kemiskinan multidimensi. Hal ini juga sejalan dengan Rencana Pembangunan Jangka Panjang Nasioal (RPJPN) 2005-2025 yang melihat masalah kemiskinan ke dalam kerangka multidimensi.

Pengentasan kemiskinan di Indonesia pada praktiknya dapat di mulai dari tingkat daerah (provinsi). Kedekatan wilayah serta kemiripan karakteristik penduduk akan memudahkan pemerintah provinsi untuk mencanangkan strategi dalam rangka pengentasan kemiskinan. Jawa Timur menempati posisi persentase penduduk miskin ke-14 dari 35 provinsi di Indonesia berdasarkan Badan Pusat Statistik (BPS). Laju persentase penduduk miskin di Jawa Timur menunjukkan penurunan sejak Maret 2018 hingga September 2019 kemudian meningkat pada Maret 2020. Peningkatan persentase penduduk miskin pada periode Maret 2020 tidak hanya terjadi di Jawa Timur namun hampir semua provinsi di Indonesia. Secara nasional, BPS mencatat persentase penduduk miskin pada Maret 2020 naik 0.56 persen dibanding September 2019. Salah satu penyebabnya yakni kebijakan Pembatasan Sosial Berskala Besar (PSBB) pada awal kemunculan pandemi Covid-19 di Indonesia. Oleh karena dampak pandemi yang sangat luas termasuk segi ekonomi, pemerintah Indonesia sejak Agustus 2020 mengajak masyarakat untuk bersama bangkit dan memulihkan perekonomian nasional.

Berdasarkan uraian mengenai masalah kemiskinan, penulis ingin memetakan beberapa karakteristik kemiskinan di wilayah Jawa Timur dengan mengelompokkan kabupaten/kota. Penelitian terdahulu oleh Leleury dan Wakanubun (2015) menggunakan analisis biplot untuk memetakan karakteristik kemiskinan di Provinsi Maluku, hasil penelitian tersebut membagi Provinsi Maluku ke dalam empat kluster sesuai dengan koordinat biplot. Pada segi kemiskinan multidimensi, Rastantra (2018) mengkaji kemiskinan multidimensi di Jawa Timur tahun 2017 dengan metode Alkire-Foster dan indeks komposit. Hasil penelitian tersebut menunjukkan



indeks kemiskinan multidimensi kabupaten-kabupaten di Pulau Madura serta Kabupaten Bondowoso dan sekitarnya, memiliki tingkat kemiskinan multidimensi yang tinggi dibandingkan kabupaten/kota lainnya di Jawa Timur. Secara ringkas penelitian terdahulu mengenai karakteristik atau indikator kemiskinan tersaji pada Tabel 1.1.

Tabel 1.1. Penelitian Terdahulu Mengenai Indikator Kemiskinan

No.	Peneliti	Judul	Metode	Tahun
1.	Muhammad Irfan Afandi	Analisis Hierarki Metode <i>Linkage</i> pada Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Kemiskinan	<i>Cluster</i> dengan <i>Complete Clustering</i>	2020
2.	Volandio Ardhan Rastantra	Kemiskinan Multidimensi Provinsi dan Kabupaten/Kota di Jawa Timur	Alkire-Foster dan Indeks Komposit	2018
3.	Sri Wahyuni dan Yogo Aryo Jatmiko	Pengelompokan Kabupaten/Kota di Pulau Jawa Berdasarkan Faktor-Faktor Kemiskinan dengan Pendekatan <i>Average Linkage Hierarchical Clustering</i>	<i>Hierarchical Clustering</i>	2018
4.	Rohmah Mustafidah	Pengelompokan Kabupaten/Kota di Provinsi Jawa Tengah Berdasarkan Indikator Kemiskinan dengan <i>C-Means</i> dan <i>Fuzzy C-Means Clustering</i>	<i>C-Means</i> dan <i>Fuzzy C-Means Clustering</i>	2017

Tabel 1.1. (Lanjutan)

No.	Peneliti	Judul	Metode	Tahun
5.	Zeth A. Leleury dan Antonia E. Wakanobun	Analisis Biplot pada Pemetaan Karakteristik Kemiskinan di Provinsi Maluku	Analisis Biplot	2015

Pada segi metode pengelompokan, Wijaya (2017) membandingkan metode pautan dengan kombinasi analisis kluster hierarki dan biplot untuk memetakan wilayah di Desa Bendosari, Kecamatan Pujon, Kabupaten Malang. Sedangkan penerapan metode *hybrid hierarchical clustering via mutual cluster* dilakukan Chipman dan Tibshirani (2006) pada data *microarray*.

Perkembangan metode pengelompokan menjadi pertimbangan penulis untuk menyelesaikan masalah penelitian. Begitu juga dengan metode untuk menganalisis karakteristik kluster yang terbentuk. Oleh karena itu, penulis ingin mengombinasi *hybrid hierarchical clustering via mutual cluster* yang menggabungkan kelebihan algoritma *bottom-up* dan *top-down*, dengan analisis biplot yang dapat menjelaskan (membantu interpretasi) hasil analisis kluster. Penelitian ini diharapkan dapat memberi kontribusi dalam merancang strategi pemulihan ekonomi di Provinsi Jawa Timur akibat pandemi Covid-19, juga sebagai langkah pengentasan kemiskinan nasional yang sejalan dengan tujuan pertama SDGs.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah pada penelitian ini adalah:

1. Bagaimana hasil pengelompokan kabupaten/kota di Jawa Timur berdasarkan indikator kemiskinan tahun 2020 menggunakan analisis *hybrid hierarchical clustering via mutual cluster*?
2. Bagaimana hasil pemetaan indikator kemiskinan tahun 2020 di Jawa Timur menggunakan analisis biplot?
3. Bagaimana penerapan analisis biplot untuk eksplorasi karakteristik setiap kluster hasil analisis *hybrid hierarchical clustering via mutual cluster*?



1.3. Tujuan Penelitian

Tujuan yang ingin dicapai berdasarkan rumusan masalah adalah:

1. Menerapkan metode pengelompokan *hybrid hierarchical clustering via mutual cluster* untuk mengelompokkan kabupaten/kota di Jawa Timur berdasarkan indikator kemiskinan tahun 2020.
2. Memetakan indikator kemiskinan tahun 2020 di Jawa Timur menggunakan analisis biplot.
3. Memanfaatkan hasil analisis biplot sebagai salah satu cara eksplorasi karakteristik setiap klaster hasil analisis *hybrid hierarchical clustering via mutual cluster*.

1.4. Manfaat Penelitian

Manfaat yang akan diperoleh dari hasil penelitian ini adalah:

1. Sebagai referensi analisis lanjutan untuk eksplorasi setelah melakukan analisis pengelompokan.
2. Mengetahui bagaimana pemetaan karakteristik kemiskinan di Jawa Timur, sehingga hasil dari penelitian ini dapat menjadi bahan masukan untuk merancang strategi pengentasan kemiskinan oleh pemerintah khususnya Provinsi Jawa Timur.

1.5. Batasan Masalah

Batasan masalah pada penelitian ini adalah:

1. Data yang digunakan adalah data sekunder indikator kemiskinan kabupaten/kota di Jawa Timur tahun 2020.
2. Variabel yang digunakan berdasarkan indikator kemiskinan multidimensi yang disesuaikan dengan ketersediaan data BPS.
3. Kemiripan antar objek (kabupaten/kota) diukur menggunakan jarak Mahalanobis.



BAB II TINJAUAN PUSTAKA

2.1. Statistika Deskriptif

Statistika merupakan kumpulan teknik atau metode yang berguna untuk membuat keputusan tentang sebuah proses atau populasi berdasarkan analisis informasi yang terkandung dalam sampel (Montgomery, 2009). Metode statistika terbagi menjadi dua, yakni statistika deskriptif dan statistika inferensial. Statistika deskriptif adalah metode-metode yang berkaitan dengan penyajian data, sehingga dapat menjelaskan informasi yang terkandung dalam data. Statistika deskriptif berfungsi untuk memberikan informasi hanya mengenai data yang ada dan sama sekali tidak menarik kesimpulan apapun tentang populasi (Walpole, 1995). Statistika inferensial mencakup semua metode yang berhubungan dengan analisis sampel hingga penarikan kesimpulan mengenai populasi.

Data adalah sekumpulan fakta atau angka yang dapat dijadikan sebagai dasar dari sebuah kajian atau penelitian. Menurut sifatnya, data terbagi menjadi dua jenis yaitu kualitatif dan kuantitatif. Baik kualitatif maupun kuantitatif, data harus diolah atau dianalisis terlebih dahulu agar diperoleh informasi yang bermanfaat. Salah satu statistik dasar yang biasa digunakan untuk mendapatkan informasi dari data adalah rata-rata yang merupakan salah satu ukuran pemusatan data. Rata-rata (\bar{y}) menjelaskan nilai variabel sebagian besar objek yang diperoleh dari persamaan (2.1) berikut.

$$\bar{y} = \frac{\sum_{i=1}^n Y_i}{n} \quad (2.1)$$

Y_i : nilai variabel objek ke- i

n : ukuran sampel.

Selain ukuran pemusatan, informasi dalam data dapat diperoleh dari ukuran penyebaran, yaitu melalui statistik ragam (s^2). Apabila ingin diketahui seberapa besar keragaman dalam data, maka dapat digunakan rumus sesuai persamaan (2.2) berikut.

$$s^2 = \frac{\sum_{i=1}^n (Y_i - \bar{y})^2}{n-1} \quad (2.2)$$

Y_i : nilai variabel objek ke- i

\bar{y} : rata-rata sampel

n : ukuran sampel.

Ragam merupakan ukuran seberapa jauh objek-objek dalam data tersebar. Ragam memiliki satuan berpangkat dua, untuk



memudahkan interpretasi biasanya digunakan simpangan baku atau standar deviasi yang memiliki satuan sama dengan satuan data. Simpangan baku (s) diperoleh dari akar kuadrat ragam yang bertanda positif sesuai dengan persamaan (2.3) berikut.

$$s = +\sqrt{s^2} \quad (2.3)$$

2.2. Analisis Klaster

Analisis klaster mengelompokkan objek atau kasus menjadi kelompok-kelompok yang lebih kecil di mana setiap kelompok berisi objek yang mirip satu sama lain (Supranto, 2004). Objek yang dimaksud dalam penelitian ini adalah kabupaten/kota. Variabel-variabel digunakan sebagai pertimbangan dan dasar pengelompokan kabupaten/kota berdasarkan kemiripan. Dengan demikian, analisis klaster ini mengelompokkan kabupaten/kota sehingga terbentuk beberapa kelompok yang disebut klaster, di mana kabupaten/kota di dalam suatu klaster lebih mirip dibandingkan antar klaster. Prosedur pembentukan klaster dibagi menjadi dua yaitu hierarki dan nonhierarki.

2.2.1. Prosedur Hierarki

Hierarki merupakan prosedur yang memulai pengelompokan dengan dua atau lebih objek yang mempunyai kemiripan atau jarak paling dekat, kemudian proses dilanjutkan ke objek lain yang mempunyai kedekatan kedua. Demikian seterusnya sehingga klaster akan membentuk semacam “pohon” di mana ada hierarki (tingkatan) yang jelas antar objek, di mulai yang paling mirip hingga yang paling tidak mirip. Biasanya pengelompokan ini disajikan dalam bentuk dendogram. Dendogram adalah representasi visual langkah-langkah analisis klaster yang menunjukkan bagaimana klaster terbentuk dan nilai jarak pada setiap langkah (Usman, 2013). Prosedur hierarki menghasilkan satu set lengkap solusi klaster yang tercermin dalam dendogram. Mulai dari semua klaster beranggota objek tunggal, hingga solusi satu klaster di mana semua objek berada dalam satu klaster. Dengan demikian, prosedur hierarki menyediakan kerangka kerja yang sangat baik dan cepat untuk membandingkan beberapa solusi klaster dan membantu dalam menilai berapa banyak klaster yang harus dipertahankan.

Pada dasarnya prosedur hierarki dibedakan menjadi dua metode pengelompokan, yaitu penggabungan (*Agglomerative*) dan



pembagian (*Divisive*). Metode *Agglomerative* di mulai dengan mengasumsikan bahwa setiap objek adalah kluster. Kemudian dua objek dengan jarak terdekat digabungkan menjadi satu kluster. Selanjutnya objek ketiga akan bergabung dengan kluster baru dengan tetap memperhitungkan jarak kedekatan antar objek. Perlu diperhatikan bahwa setiap penggabungan kluster dalam metode ini selalu diikuti dengan perbaikan matriks jarak. Hair dkk. (2010) menyebutkan bahwa terdapat lima algoritma *agglomerative* yaitu pautan tunggal (*single linkage*), pautan lengkap (*complete linkage*), pautan rata-rata (*average linkage*), metode memusat (*centroid method*) dan metode Ward (*Ward's method*). Pada penelitian ini, digunakan metode *average linkage* karena cukup mewakili kluster dan sedikit terpengaruh oleh pencilan.

Average linkage adalah penggabungan yang didasarkan pada jarak rata-rata antar objek. Kriteria penggabungan yang digunakan pada prosedur *average Linkage* adalah rata-rata jarak semua objek dalam suatu kluster dengan jarak semua objek pada kluster yang lain. Algoritma pengelompokan dengan metode *average linkage* adalah sebagai berikut (Johnson dan Wichern, 2007).

- 1) Memulai dengan semua objek sebagai kluster mereka sendiri, sehingga banyak kluster sama dengan banyak objek.
- 2) Menggunakan ukuran jarak untuk menggabungkan dua objek dengan jarak terdekat, misalnya A dan B. Objek tersebut digabung untuk membentuk satu kluster baru (AB).
- 3) Jarak antara kelompok AB dengan yang lainnya ditentukan oleh:

$$d_{(AB)C} = \frac{\sum_i \sum_j d_{ij}}{N_{(AB)}N_C} \quad (2.4)$$

$d_{(AB)C}$: jarak rata-rata antara kluster (AB) dan kluster C

d_{ij} : jarak antara objek-*i* pada kluster (AB) dan objek-*j* pada kluster C

$N_{(AB)}$: banyak objek pada kluster (AB)

N_C : banyak objek pada kluster C.

- 4) Mengulangi langkah (3) sampai seluruh objek menjadi satu kluster.

2.2.2. Prosedur Nonhierarki

Prosedur nonhierarki berkebalikan dengan hierarki yang melibatkan konstruksi seperti pohon. Penetapan objek ke dalam



klaster pada prosedur nonhierarki dilakukan setelah banyak klaster ditentukan. Klaster-klaster yang terbentuk dalam prosedur nonhierarki dipengaruhi dua proses penting yang harus diperhatikan, yaitu penentuan benih klaster dan penempatan objek. Benih dalam hal ini merupakan objek pertama yang dimasukkan ke dalam suatu klaster sebagai dasar penempatan objek selanjutnya.

Menurut Hair dkk. (2010) terdapat tiga algoritma dalam prosedur nonhierarki yaitu berurutan (*sequential*), paralel (*parallel*) dan optimasi (*optimization*). Ketiga metode tersebut termasuk ke dalam grup algoritma yang dikenal sebagai algoritma *k-means*. Algoritma pengelompokan dengan metode *k-means* sebagai berikut (Johnson dan Wichern, 2007).

- 1) Membagi objek ke dalam k klaster awal
- 2) Menghitung pusat kelompok pada setiap variabel menggunakan rumus:

$$C_{(A)K} = \frac{1}{n_A} \sum_{i=1}^{n_A} x_{ki} \quad (2.5)$$

$C_{(A)K}$: pusat klaster A pada variabel k

n_A : banyaknya objek pada klaster A

x_{ki} : nilai objek ke- i pada variabel ke- k

- 3) Menghitung jarak objek terhadap pusat kelompok menggunakan ukuran jarak.
- 4) Membentuk klaster berdasarkan kedekatan jarak objek dengan pusat kelompok. Apabila objek berpindah dari klaster awal, maka titik pusat harus dihitung kembali. Perhitungan ulang pusat kelompok dilakukan ketika klaster menerima objek baru dan untuk klaster yang kehilangan objek.
- 5) Mengulangi langkah (2) hingga (4) sampai pusat kelompok tidak berubah lagi atau berhenti sesuai dengan maksimum jumlah iterasi yang ditetapkan.

2.3. Ukuran Jarak Mahalanobis

Jarak Mahalanobis (d_{ij}^2) merupakan ukuran jarak umum yang dapat menjelaskan korelasi antara variabel dengan cara memboboti setiap variabel dengan bobot yang sama (Hair dkk., 2010). Apabila terdapat multikolinieritas antar variabel pada data penelitian, maka dapat ditangani dengan penggunaan jarak Mahalanobis. Hal ini memungkinkan apabila ingin mempertahankan variabel-variabel yang digunakan.



Hair dkk. (2010), menjelaskan bahwa pendekatan jarak Mahalanobis juga mencakup pembakuan data, sehingga data asli dapat langsung digunakan. Pendekatan Mahalanobis tidak hanya melakukan pembakuan pada data dengan penskalaan menggunakan standar deviasi tapi juga menjumlahkan varian-kovarian gabungan dalam grup, yang menyesuaikan korelasi antar variabel-variabel. Feil dan Abonyi (2007) mendefinisikan rumus jarak Mahalanobis sesuai persamaan (2.6) berikut.

$$d_{ij}^2 = (\mathbf{x}_i - \mathbf{x}_j)' \mathbf{S}^{-1} (\mathbf{x}_i - \mathbf{x}_j) \quad (2.6)$$

di mana $\mathbf{x}_i = (x_1, x_2, \dots, x_p)$ dan $\mathbf{x}_j = (x_1, x_2, \dots, x_p)$

\mathbf{x}_i : vektor kolom yang berisi nilai setiap variabel objek ke- i berukuran $(p \times 1)$

\mathbf{x}_j : vektor kolom yang berisi nilai setiap variabel objek ke- j berukuran $(p \times 1)$

\mathbf{S} : matriks varian-kovarian berukuran $(p \times p)$

Jarak Mahalanobis digunakan karena terdapat variabel-variabel sosial yang memungkinkan adanya korelasi tinggi antar variabel atau bahkan terindikasi multikolinieritas.

2.4. *Hybrid Hierarchical Clustering via Mutual Cluster*

Hybrid hierarchical clustering merupakan metode gabungan dari prosedur hierarki dan nonhierarki. Metode ini menggabungkan kelebihan prosedur hierarki yang bersifat *bottom-up* dan *top-down*.

Bottom-up baik dalam segi kecepatan membentuk kluster dan mengidentifikasi data berukuran kecil, sedangkan *top-down* baik dalam mengidentifikasi data berukuran besar. *Mutual cluster* merupakan hasil suatu pengelompokan yang menggunakan jarak terbesar antara pasangan dalam kluster, yang lebih kecil dari jarak terpendek ke setiap titik di luar kluster (Chipman dan Tibshirani, 2006). Oleh karena itu, jarak maksimal antar objek dalam sebuah *mutual cluster* lebih kecil dibandingkan jarak minimal beberapa objek di luar *mutual cluster*.

Sesuai dengan tujuan *hybrid hierarchical clustering* yang menggabungkan dua prosedur, maka pengelompokan ini dilakukan dengan dua tahap. Tahap pertama secara *bottom-up* untuk menentukan *mutual cluster*, kemudian secara *top-down* dari hasil *mutual cluster* (Chipman dan Tibshirani, 2006). Algoritma *hybrid hierarchical clustering via mutual cluster* sebagai berikut:

1) Menghitung jarak antar objek menggunakan jarak Mahalanobis.



- 2) Menentukan *mutual cluster* berdasarkan algoritma *agglomerative* dengan metode *average linkage*.
- 3) Melakukan pengelompokan *hybrid hierarchical clustering* dengan algoritma *k-means* menggunakan hasil dari langkah (2). Objek akan dibagi menjadi *k* kelompok di mana *mutual cluster* yang telah terbentuk harus tetap ada dan dipertahankan.

Speed (2003) menyatakan bahwa penentuan $k = 2$ menghasilkan partisi data yang cukup masuk akal, sehingga dapat dilakukan partisi data secara rekursif menggunakan algoritma *2-means*. Tahap pertama adalah menjalankan *2-means* pada dataset lengkap, kemudian menjalankan *2-means* secara rekursif di setiap sub-kelaster, sampai setiap objek menjadi anggota kelaster tunggal. Menurut Gersho dan Gray dalam Speed (2003), istilah untuk pendekatan tersebut adalah *tree structured vector quantization* (TSVQ). Oleh karena itu *k-means* sering disebut sebagai *vector quantization*.

2.5. Validasi Klaster

Validasi klaster adalah prosedur yang mengevaluasi hasil analisis klaster secara kualitatif dan objektif. Evaluasi yang dihasilkan oleh validasi klaster berupa indeks yang dapat menunjukkan seberapa baik hasil pengelompokan yang telah dilakukan. Indeks *silhouette* adalah salah satu indeks validasi yang disebutkan oleh Chen dkk. (2002).

Kegunaan indeks *silhouette* adalah mengukur seberapa baik suatu objek dapat ditempatkan dalam suatu klaster. Indeks *silhouette* ($S_{(i)}$) dihitung sebagai derajat kepercayaan pengelompokan objek di dalam klaster yang bernilai -1 sampai 1. Indeks *silhouette* didapatkan dari rumus sesuai persamaan (2.7) berikut.

$$S_{(i)} = \frac{b_{(i)} - a_{(i)}}{\max(a_{(i)}, b_{(i)})} \quad (2.7)$$

$a_{(i)}$: rata-rata jarak antara objek ke- *i* dengan semua objek pada klaster yang sama.

$b_{(i)}$: nilai minimum dari rata-rata jarak antara objek ke-*i* dengan semua objek yang berada pada klaster lain.

Indeks *silhouette* yang semakin besar (mendekati 1) menunjukkan banyak klaster yang terbentuk semakin optimal atau terbaik (Kaufman dan Rousseeuw, 1990).



2.6. Analisis Biplot

Analisis biplot memberikan gambaran tentang objek dan variabel dari suatu matriks data ke dalam bentuk plot dua dimensi. Gambaran yang dihasilkan dapat berupa informasi tentang kedekatan antar objek, korelasi dan keragaman antar variabel serta karakteristik utama setiap objek. Perlu diperhatikan bahwa analisis biplot merupakan pereduksian dari ruang berdimensi besar ke ruang berdimensi dua. Oleh karena itu, dimungkinkan berkurangnya informasi yang akan terjelaskan dari hasil analisis biplot. Matriks data (X) yang digunakan sebagai input analisis biplot adalah

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

n : banyak objek (kabupaten/kota)

p : banyak variabel (indikator kemiskinan)

Menurut Sartono dkk.(2003), analisis biplot didasarkan pada *Singular Value Decomposition* (SVD). Bentuk umum SVD oleh Greenacre dalam Nugroho (2008) dijelaskan dalam persamaan (2.8) berikut.

$$X = ULA' \quad (2.8)$$

X : matriks data berukuran ($n \times p$)

U : matriks singular berukuran ($n \times r$)

L : matriks diagonal nilai singular berukuran ($r \times r$)

A : matriks singular berukuran ($p \times r$).

Landasan analisis biplot adalah setiap matriks berukuran ($n \times r$) yang berpangkat r dengan $r = \min(n, p)$ (Nugroho, 2008). Oleh karena biplot berdimensi dua maka pangkat $r = 2$. Didefinisikan bahwa U dan A merupakan matriks singular, sehingga $U'U = A'A = I_r$ (matriks identitas berdimensi r). Unsur-unsur diagonal L adalah akar kuadrat dari akar ciri-akar ciri $X'X$ atau XX' ($\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \dots \geq \sqrt{\lambda_n}$), sehingga unsur-unsur ini disebut nilai singular X .

Kolom-kolom A disebut vektor singular baris yang merupakan landasan ortonormal baris-baris X dalam ruang berdimensi p . Kolom-kolom U disebut vektor singular kolom yang merupakan landasan ortonormal kolom-kolom X dalam ruang berdimensi n .

Penjabaran persamaan (2.8) menjadi:

$$X = UL^\alpha L^{1-\alpha} A' \quad (2.9)$$



Misalkan $\mathbf{G} = \mathbf{U}\mathbf{L}^\alpha$ dengan \mathbf{G} adalah matriks berukuran $(n \times r)$ dan $\mathbf{H}' = \mathbf{L}^{1-\alpha}\mathbf{A}'$ dengan \mathbf{H}' adalah matriks berukuran $(p \times r)$. Sedangkan α adalah nilai faktorisasi yang besarnya $0 \leq \alpha \leq 1$, sehingga persamaan (2.9) menghasilkan koordinat berdasarkan nilai komponen utama masing-masing objek dan variabel. Dengan $r = 2$ persamaan (2.9) dapat ditulis:

$$\mathbf{X} = \mathbf{G}_2\mathbf{H}_2' \quad (2.10)$$

$$\text{di mana } \mathbf{G}_2 = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ \vdots & \vdots \\ g_{n1} & g_{n2} \end{bmatrix} \text{ dan } \mathbf{H}_2 = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \\ \vdots & \vdots \\ h_{p1} & h_{p2} \end{bmatrix}$$

Nilai α yang digunakan dapat merupakan nilai sembarang dengan $0 \leq \alpha \leq 1$. Tetapi pengambilan nilai-nilai ekstrim $\alpha = 0$ dan $\alpha = 1$ akan memudahkan dalam interpretasi. Jika $\alpha = 0$ maka $\mathbf{G} = \mathbf{U}$ dan $\mathbf{H}' = \mathbf{L}\mathbf{A}'$ disebut dengan biplot \mathbf{GH} atau *Column Metric Preserving* yang mempertahankan matriks kolom (menunjukkan variabel dalam \mathbf{X}) digunakan untuk mengetahui keragaman variabel dan korelasi antar variabel. Hal ini termasuk ke dalam salah satu tipe biplot, menurut Nugroho (2008) tipe-tipe biplot sebagai berikut.

1) Biplot *Column Metric Preserving*

Menggunakan $\alpha = 0$ yang akan mempertahankan matriks kolom (variabel) dan menampilkan hubungan antar variabel berupa matriks kovarian atau korelasi.

2) Biplot *Square Root*

Menggunakan $\alpha = 0.5$ yang akan membuat kesamaan penskalaan atau pembobot untuk baris dan kolom. Kesamaan penskalaan ini diperlukan dalam interpretasi hubungan dari dua faktor pengamatan (objek dan variabel).

3) Biplot *Row Metric Preserving*

Menggunakan $\alpha = 1$ yang akan mempertahankan matriks baris (menunjukkan objek) dan digunakan untuk mengetahui keragaman dan hubungan antar objek.

Oleh karena analisis biplot mereduksi ruang berdimensi banyak menjadi dua dimensi sehingga mengakibatkan informasi yang terkandung berkurang, maka perlu diketahui keragaman hasil analisis biplot. Mattjik dan Sumertajaya (2011) menyebutkan bahwa biplot dianggap memberikan cukup informasi jika telah memberi informasi minimal 70%. Proporsi keragaman kumulatif biplot (ρ^2) dapat diperoleh dari rumus sesuai persamaan (2.11) berikut.



$$\rho^2 = \frac{(\lambda_1 + \lambda_2)}{\sum_{k=1}^p \lambda_k} \quad (2.11)$$

- λ_1 : nilai akar ciri terbesar pertama dari X
 λ_2 : nilai akar ciri terbesar kedua dari X
 λ_k : nilai akar ciri ke- k dari X ; $k = 1, 2, \dots, p$.

Apabila ρ^2 mendekati 1 maka biplot memberikan penyajian yang semakin baik mengenai informasi data yang sebenarnya.

2.7. Indikator Kemiskinan

Kemiskinan dibagi menjadi dua definisi yaitu kemiskinan mutlak dan kemiskinan relatif. Menurut Todaro dan Smith (2012), kemiskinan mutlak atau absolut adalah kondisi tidak bisa memenuhi kebutuhan pokok makanan, pakaian dan tempat tinggal. Besarnya masalah kemiskinan absolut dicerminkan dengan jumlah penduduk yang tingkat pendapatan atau tingkat konsumsinya berada di bawah garis kemiskinan yang telah ditentukan. Garis Kemiskinan (GK) diperoleh dari penjumlahan Garis Kemiskinan Makanan (GKM) dan Garis Kemiskinan Non Makanan (GKNM). Berdasarkan konsep BPS, GKM adalah nilai pengeluaran kebutuhan minimum makanan yang disetarakan dengan 2100 kilokalori perkapita perhari. Sedangkan GKNM adalah kebutuhan minimum untuk perumahan, sandang, pendidikan dan kesehatan. Besarnya kemiskinan relatif dinyatakan dengan berapa persen dari pendapatan nasional yang diterima oleh sekelompok penduduk dengan kelas pendapatan tertentu dibandingkan dengan kelas pendapatan lainnya. BPS menyebutkan bahwa untuk mengukur kemiskinan digunakan konsep kemampuan memenuhi kebutuhan dasar. Hal ini berarti kemiskinan dipandang dari ketidakmampuan aspek ekonomi untuk memenuhi kebutuhan dasar.

Todaro dan Smith (2012) menyatakan bahwa kemiskinan tidak bisa diukur secara memadai dengan pendapatan saja karena pendapatan tidak diukur dengan sempurna. Oleh karena itu perlu adanya pengukuran kemiskinan dari berbagai aspek atau dimensi. *United Nations Development Program* (UNDP) membangun indeks kemiskinan multidimensi untuk mengukur tingkat kemiskinan dari berbagai aspek. Kemiskinan multidimensi dipandang dari tiga aspek yaitu aspek kesehatan, pendidikan dan standar kehidupan. Indikator digunakan sebagai perwakilan untuk mengukur setiap aspek. Aspek kesehatan mempunyai dua indikator:



- 1) Adanya kasus kematian anak di dalam keluarga.
- 2) Adanya anggota keluarga yang mengalami gizi buruk.

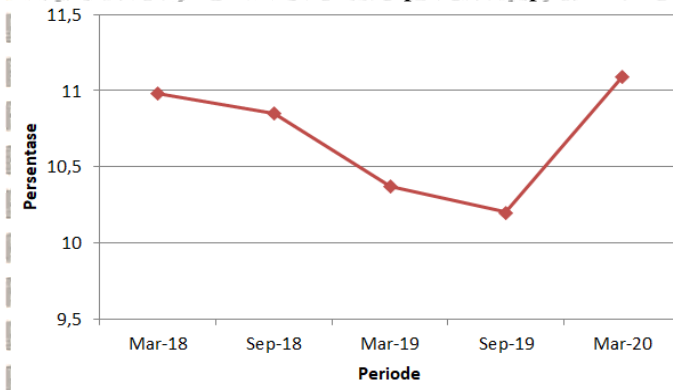
Aspek pendidikan mempunyai dua indikator:

- 1) Setidaknya ada satu anggota keluarga yang tidak menyelesaikan pendidikan selama lima tahun.
- 2) Adanya anak dengan usia sekolah (kelas satu sampai delapan) yang tidak bersekolah.

Aspek standar kehidupan diwakili oleh enam indikator:

- 1) Kekurangan listrik.
- 2) Tidak tersedia air minum yang layak.
- 3) Sanitasi yang tidak memadai.
- 4) Lantai yang tidak memadai.
- 5) Bahan bakar memasak yang tidak berkembang.
- 6) Kekurangan salah satu dari lima kepemilikan aset: telepon, radio, televisi, sepeda dan motor atau kendaraan sejenis.

Kemiskinan di Indonesia diukur oleh BPS sebanyak dua periode dalam satu tahun. Berdasarkan data pada BPS, selama dua tahun terakhir persentase penduduk miskin di Indonesia mengalami penurunan. Namun persentase penduduk miskin kembali meningkat sejak adanya pandemi Covid-19. Laju persentase kemiskinan di Indonesia selama dua tahun terakhir dapat dilihat pada Gambar 2.1.



Gambar 2.1. Persentase Penduduk Miskin Jawa Timur 2018-2020



BAB III METODE PENELITIAN

3.1. Data

Penelitian ini menggunakan data sekunder indikator kemiskinan Provinsi Jawa Timur tahun 2020. Data diperoleh dari publikasi elektronik Statistik Kesejahteraan Rakyat yang dirilis Badan Pusat Statistik Provinsi Jawa Timur pada 30 Desember 2020. Publikasi ini merupakan hasil pengumpulan data Survey Sosial Ekonomi Nasional (Susenas) Maret 2020. Objek yang digunakan yaitu 29 kabupaten dan 9 kota di Provinsi Jawa Timur. Data yang digunakan dapat dilihat pada Lampiran 1.

3.2. Variabel Penelitian

Variabel penelitian diambil dari referensi penelitian terdahulu dengan mempertimbangkan ketersediaan data yang ada di BPS Jawa Timur. Variabel-variabel yang digunakan tersaji pada tabel 3.1.

Tabel 3. 1. Variabel Penelitian

Variabel	Deskripsi	Skala
X_1	Persentase rumah tangga yang luas lantai bangunan tempat tinggalnya $< 20m^2$.	Rasio
X_2	Persentase rumah tangga yang sumber penerangan utamanya bukan listrik.	Rasio
X_3	Persentase rumah tangga yang tidak memiliki sumber air minum layak.	Rasio
X_4	Persentase rumah tangga yang tidak memiliki sanitasi layak (tidak ada fasilitas BAB sendiri/bersama).	Rasio
X_5	Persentase rumah tangga yang bahan bakar utama memasak selain gas dan listrik.	Rasio
X_6	Persentase rumah tangga yang anggota rumah tangganya tidak menguasai/memiliki telepon selular.	Rasio
X_7	Persentase penduduk yang putus sekolah jenjang SMP (usia 13-15 tahun tidak sekolah lagi).	Rasio
X_8	Persentase penduduk yang tidak sekolah usia 7-12 tahun.	Rasio



3.3. Metode Analisis Data

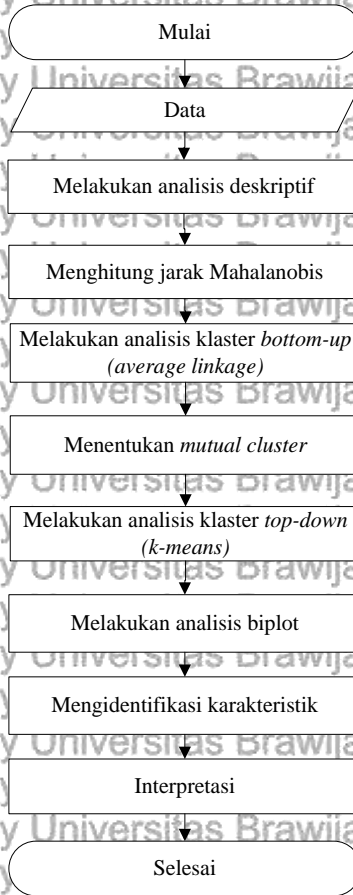
Analisis data dilakukan dengan bantuan *software* R dan melalui prosedur sebagai berikut.

- 1) Mengumpulkan data.
- 2) Melakukan analisis deskriptif untuk setiap variabel.
- 3) Menghitung jarak antar objek menggunakan jarak Mahalanobis pada persamaan (2.6).
- 4) Melakukan pengelompokan *bottom-up* berdasarkan prosedur hierarki dengan metode *average linkage* menggunakan rumus (2.4).
 - 5) Menentukan *mutual cluster* dengan tahap sebagai berikut.
 - a) Mendapatkan jarak penggabungan dari hasil *average linkage* pada langkah (5).
 - b) Memilih objek awal berdasarkan jarak penggabungan terbesar.
 - c) Menggabungkan objek (b) dengan objek terdekat ke dalam satu kluster.
 - d) Membandingkan jarak antar objek di dalam kluster dengan objek-objek lain di luar kluster.
 - Apabila jarak maksimal antar objek di dalam kluster lebih kecil dibandingkan dengan jarak minimal dengan objek di luar kluster, maka kluster pada tahap (c) menjadi *mutual cluster*.
 - Apabila jarak maksimal antar objek di dalam kluster lebih besar dibandingkan dengan jarak minimal dengan objek di luar kluster, maka kluster pada tahap (c) bukan *mutual cluster*.
 - e) Memilih objek baru yang akan dibandingkan, kemudian mengulangi tahap (c) hingga semua objek dibandingkan.
 - 6) Melakukan pengelompokan *top-down* dengan membentuk TSVQ berupa dendrogram berdasarkan prosedur *k-means* ($k = 2$) menggunakan rumus (2.5) dengan tetap mempertahankan *mutual cluster*. Tahapan pembentukan dendrogram (TSVQ) adalah sebagai berikut.
 - a) Menjalankan prosedur *2-means* pada dataset lengkap.
 - b) Menjalankan prosedur *2-means* secara rekursif di setiap sub-kluster, sampai setiap objek menjadi anggota kluster tunggal.



- 7) Menentukan banyak kluster optimal (validasi) dengan menghitung indeks *silhouette* menggunakan rumus (2.7).
- 8) Memetakan seluruh variabel dan objek dengan membentuk biplot, dengan memberi simbol dan warna berbeda pada objek untuk setiap kluster.
- 9) Menghitung proporsi keragaman kumulatif biplot dengan rumus (2.11).
- 10) Mengidentifikasi karakteristik setiap kluster berdasarkan biplot dan rata-rata setiap kluster menggunakan rumus (2.1).
- 11) Interpretasi.

Secara ringkas, prosedur penelitian digambarkan dalam diagram alir pada Gambar 3.1.



Gambar 3.1. Diagram Alir

BAB IV HASIL DAN PEMBAHASAN

4.1. Analisis Deskriptif

Gambaran awal mengenai kemiskinan di Jawa Timur dapat dilihat dari statistik deskriptif setiap variabel. Statistik yang digunakan adalah rata-rata, nilai maksimum dan minimum. Berdasarkan persamaan (2.1) statistik deskriptif setiap variabel tersaji dalam Tabel 4.1 berikut.

Tabel 4.1. Statistik Deskriptif Indikator Kemiskinan di Jawa Timur Tahun 2020 (%)

Variabel	Minimum	Maksimum	Rata-rata
X_1	0	27.41	2.82
X_2	0	0.36	0.04
X_3	0	18.55	3.79
X_4	0	38.75	7.01
X_5	0.31	63.88	14.49
X_6	2.75	24.24	10.83
X_7	0	7.63	2.33
X_8	0	1.26	0.41

Berdasarkan Tabel 4.1 setiap variabel dapat dijelaskan bahwa pada tahun 2020:

- a. Sebagian besar kabupaten/kota di Jawa Timur memiliki 2.82% rumah tangga yang luas lantai bangunan tempat tinggalnya $< 20m^2 (X_1)$. Persentase terendah dimiliki Kab. Magetan dengan 0% rumah tangga atau tidak ada rumah tangga yang memiliki luas lantai bangunan $< 20m^2$. Persentase tertinggi dimiliki Kota Surabaya yaitu sebesar 27.41% rumah tangga yang memiliki luas lantai bangunan $< 20m^2$.
- b. Pada variabel sumber penerangan rumah tangga, sekitar 0,04% rumah tangga di Jawa Timur menggunakan sumber penerangan utama bukan listrik (X_2), dengan persentase tertinggi sebesar 0.36% di Kab. Sumenep. Sebanyak 21 kabupaten dan 8 kota termasuk persentase terendah (0%), yang berarti rumah tangga di kabupaten/kota tersebut sumber penerangan utamanya telah menggunakan listrik.





- c. Terdapat sekitar 3.79% rumah tangga di Jawa Timur yang sumber air minumannya tidak layak (X_3). Beberapa kota dengan persentase rendah (0%) berarti sumber air minum setiap rumah tangganya sudah layak, di antaranya: Kota Malang, Probolinggo, Pasuruan, Mojokerto dan Madiun. Namun masih terdapat Kabupaten dengan persentase tinggi yaitu 18.55% di Kabupaten Trenggalek yang nilainya cukup jauh dengan nilai rata-rata.
- d. Kondisi kelayakan sanitasi di Jawa Timur cukup tinggi dengan setiap kabupaten/kota memiliki rata-rata 7.01% rumah tangga yang tidak memiliki sanitasi layak (X_4). Walaupun demikian, masih terdapat kabupaten dengan persentase sanitasi tidak layak tertinggi yaitu 38.75% di Kabupaten Bondowoso. Sebaliknya, beberapa kota dengan rumah tangganya yang telah memiliki sanitasi layak di antaranya: Kota Kediri, Malang dan Madiun.
- e. Sebagian besar kabupaten/kota di Jawa Timur memiliki 14.49% rumah tangga yang bahan bakar utama untuk memasak selain gas dan listrik (X_5). Persentase terendah dimiliki Kab. Gresik dengan 0.31% rumah tangga yang bahan bakar utama untuk memasak selain gas dan listrik. Persentase tertinggi dimiliki Kab. Pacitan yaitu sebesar 63.88% rumah tangga yang bahan bakar utama untuk memasak selain gas dan listrik.
- f. Pada aspek penguasaan teknologi yang dicerminkan variabel X_5 , terdapat sekitar 10.83% rumah tangga di kabupaten/kota di Jawa Timur yang tidak memiliki/menguasai telepon selular (X_6). Cukup jauh dengan nilai rata-rata, Bondowoso menjadi kabupaten dengan persentase tertinggi (24.24%) untuk rumah tangga yang tidak menguasai/memiliki telepon selular. Sebaliknya, Kota Malang memiliki persentase terendah (2.75%) yang berarti rumah tangga di Kota Malang sebagian besar telah menguasai/melek teknologi.
- g. Pada aspek pendidikan, persentase penduduk yang putus sekolah jenjang SMP (X_7) di setiap kabupaten/kota di Jawa Timur terdapat sekitar 2.33%. Kota Mojokerto menjadi kota dengan persentase putus sekolah jenjang SMP terendah (0%), yang berarti di Kota Mojokerto hampir semua penduduk menamatkan sekolah jenjang SMP atau melanjutkan ke jenjang SMA. Sebaliknya, Persentase tertinggi dimiliki Kab. Bangkalan



yaitu sebesar 7.63% penduduknya putus sekolah pada jenjang SMP.

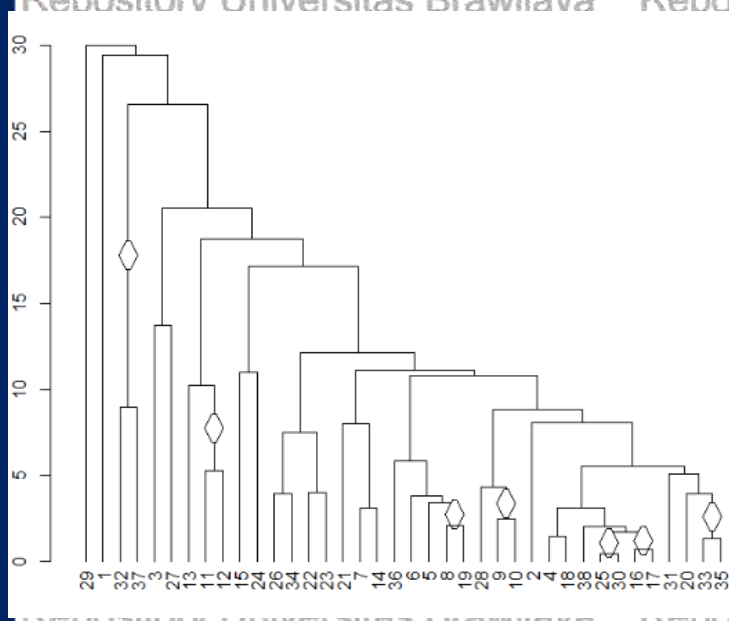
- h. Persentase penduduk yang bersekolah pada usia 7-12 tahun (X_8) juga merupakan aspek pendidikan pada kemiskinan multidimensi. Terdapat sekitar 0.41% penduduk yang tidak sekolah pada usia 7-12 tahun pada setiap kota di Jawa Timur. Persentase tertinggi dimiliki Kab. Sampang yaitu sebesar 1.26% penduduknya tidak sekolah pada usia 7-12 tahun. Beberapa kabupaten/kota dengan persentase terendah, di antaranya: Kab. Ponorogo, Kab. Tuban, Kab. Pamekasan, Kab. Sumenep dan Kota Blitar menunjukkan bahwa penduduk usia 7-12 tahun di kabupaten/kota tersebut telah bersekolah.

4.2. *Hybrid Hierarchical Clustering via Mutual Cluster*

Pengelompokan kabupaten/kota dilakukan dalam dua tahap. Tahap pertama yaitu membentuk *mutual cluster* dengan prosedur pengelompokan *bottom-up*. Tahap kedua yaitu pengelompokan berdasarkan prosedur *top-down* dengan tetap mempertahankan *mutual cluster*.

4.2.1. Pembentukan *Mutual Cluster*

Mutual cluster dibentuk berdasarkan ukuran jarak Mahalanobis dan metode yang digunakan adalah *average linkage*. *Mutual cluster* yang pertama ditentukan dengan cara memilih kabupaten/kota dengan jarak penggabungan terbesar, kemudian dicari kabupaten/kota terdekatnya untuk digabungkan dalam satu klaster. Selanjutnya, klaster yang terbentuk akan menjadi *mutual cluster* yang pertama apabila jarak maksimal antar kabupaten/kota di dalam klaster, lebih kecil dibandingkan dengan jarak minimal kabupaten/kota di luar klaster. Jika tidak terpenuhi, maka penentuan *mutual cluster* dilanjutkan sesuai jenjang dendrogram metode *average linkage*. Hasil dari pembentukan *mutual cluster* berupa dendrogram yang dapat dilihat pada Gambar 4.1 dengan kabupaten/kota yang termasuk *mutual cluster* ditandai dengan simbol heksagonal.



Gambar 4.1. Dendrogram *Mutual Cluster*

Terdapat tujuh *mutual cluster* kabupaten/kota yang terbentuk. Anggota tujuh *mutual cluster* tersebut dapat dilihat dalam Tabel 4.2 berikut.

Tabel 4.2. Anggota *Mutual Cluster*

<i>Mutual Cluster</i>	Kode Kab./Kota	Anggota
1	32, 37	Kota Malang dan Kota Surabaya.
2	11, 12	Kab. Bondowoso dan Kab. Situbondo.
3	8, 19	Kab. Lumajang dan Kab. Madiun.
4	9, 10	Kab. Jember dan Kab. Banyuwangi.
5	25, 30	Kab. Gresik dan Kota Kediri.
6	16, 17	Kab. Mojokerto dan Kab. Jombang.
7	33, 35	Kota Probolinggo dan Kota Mojokerto

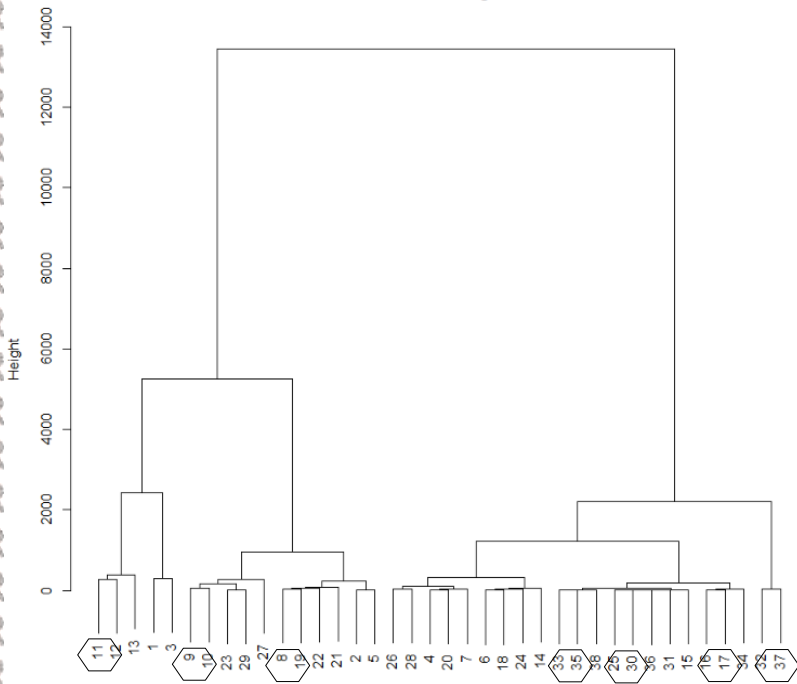
Kabupaten/kota yang termasuk ke dalam satu *mutual cluster* akan dipertahankan berada pada klaster yang sama ketika dilakukan pengelompokan secara *top-down*. Seperti pada *mutual cluster* 1, Kota Malang dan Kota Surabaya tidak akan terpisah pada saat



pengelompokan secara *top-down*, pada akhirnya kedua kota tersebut akan berada dalam kluster yang sama. Demikian juga dengan anggota *mutual cluster* 1 sampai 7.

4.2.2. Hybrid Hierarchical Clustering

Apabila *mutual cluster* telah didapatkan, kabupaten/kota dikelompokkan berdasarkan prosedur *top-down hierarchical clustering* menggunakan metode *k-means* dengan $k = 2$ atau disebut juga dengan *tree structured vector quantization* (TSVQ). Dengan mempertahankan *mutual cluster*, dibentuk struktur pohon atau dendrogram yang menampilkan semua kemungkinan kluster. Pengelompokan dimulai dengan membagi (partisi) semua kabupaten/kota di Jawa Timur ke dalam 2 kluster. Selanjutnya, partisi secara rekursif dilakukan pada setiap sub-kluster, hingga setiap kabupaten/kota menjadi anggota kluster tunggal. Dendrogram hasil *hybrid hierarchical clustering* dapat dilihat pada Gambar 4.2.



Gambar 4.2. Cluster Dendrogram



Berdasarkan Gambar 4.2 dapat dilihat bahwa anggota dari setiap *mutual cluster* saling berdekatan, sehingga dapat dipastikan *mutual cluster* tetap dipertahankan.

4.3. Validasi Klaster

Dendrogram hasil *hybrid hierarchical clustering*, selanjutnya dipangkas menjadi beberapa klaster. Pemangkasan dilakukan beberapa kali hingga didapatkan banyaknya klaster optimal. Penentuan banyaknya klaster awal berdasarkan *stopping rule*, yaitu sebanyak 2 sampai 15 klaster untuk divalidasi. Validasi 2 sampai 15 klaster dipilih karena banyaknya klaster yang lebih dari 15 memiliki rata-rata indeks *silhouette* yang semakin kecil hingga negatif. Berdasarkan persamaan (2.7) rata-rata indeks *silhouette* 2 sampai 15 klaster tersaji dalam Tabel 4.3 berikut.

Tabel 4.3. Rata-Rata Indeks *Silhouette* 2 Sampai 5 Klaster

Banyak klaster	Rata-rata indeks <i>silhouette</i>
2	0.14
3	0.06
4	0.12
5	0.18
6	0.12
7	0.15
8	0.12
9	0.10
10	0.10
11	0.07
12	0.09
13	0.05
14	-0.05
15	0

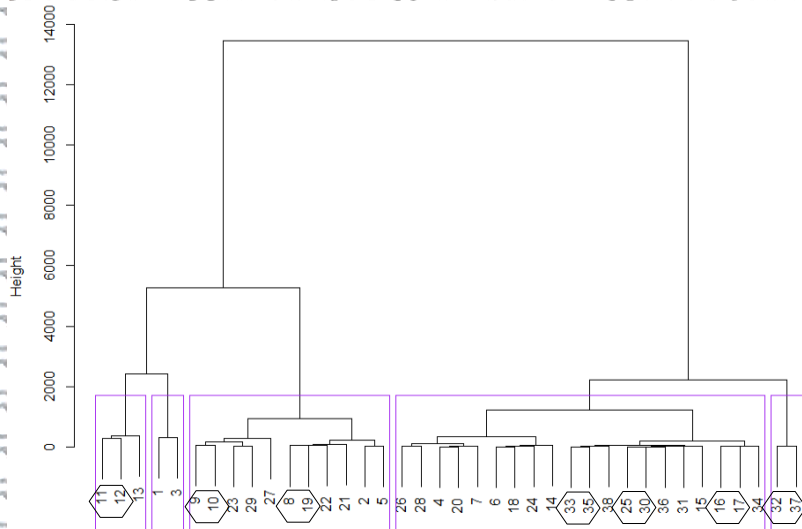
: rata-rata indeks *silhouette* terbesar.

Pada Tabel 4.3 dapat dilihat bahwa 5 klaster memiliki rata-rata indeks *silhouette* paling besar. Dengan demikian, 5 klaster merupakan banyaknya klaster optimal yang dapat dibentuk.

Apabila telah ditentukan 5 klaster sebagai banyaknya klaster optimal, kemudian diperiksa kekonsistenan *mutual cluster*-nya.



Pemangkasan dendrogram ke dalam 5 kluster dapat dilihat pada Gambar 4.3.



Gambar 4.3. Dendrogram 5 Kluster Optimal

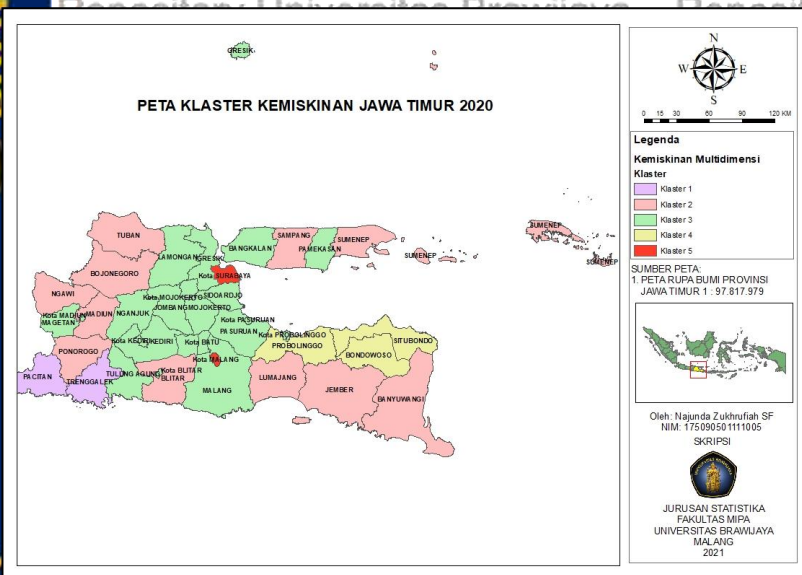
Berdasarkan Gambar 4.3 anggota setiap *mutual cluster* tidak ada yang terpisah, sehingga tetap berada di dalam kluster yang sama. Hal tersebut berarti *mutual cluster* konsisten hingga pemangkasan dendrogram ke dalam 5 kluster. Anggota setiap kluster hasil *hybrid hierarchical clustering* optimal tersaji dalam Tabel 4.4 dan peta kluster pada Gambar 4.4 berikut.

Tabel 4.4. Anggota Kluster Optimal

Kluster	Kode Kab./Kota	Anggota
1	1, 3.	Kab. Pacitan dan Kab. Trenggalek.
2	9, 10, 23, 29, 27, 8, 19, 22, 21, 2, 5.	Kab. Jember, Kab. Banyuwangi, Kab. Tuban, Kab. Sumenep, Kab. Sampang, Kab. Lumajang, Kab. Madiun, Kab. Bojonegoro, Kab. Ngawi, Kab. Ponorogo dan Kab. Blitar.
3	26, 28, 4, 20, 7, 6, 18, 24, 14, 33, 35, 38,	Kab. Bangkalan, Kab. Pamekasan, Kab. Tulungagung, Kab. Magetan, Kab. Malang, Kab. Kediri, Kab. Nganjuk, Kab. Lamongan, Kab. Pasuruan, Kota Probolinggo, Kota Mojokerto, Kota Batu,

Tabel 4.4. (Lanjutan)

Klaster	Kode Kab./Kota	Anggota
	25, 30, 36, 31, 15, 16, 17, 34.	Kab. Gresik, Kota Kediri, Kota Madiun, Kota Blitar, Kab. Sidoarjo, Kab. Mojokerto, Kab. Jombang dan Kota Pasuruan.
4	11, 12, 13.	Kab. Bondowoso, Kab. Situbondo dan Kab. Probolinggo.
5	32, 37.	Kota Malang dan Kota Surabaya.



Gambar 4.4. Peta Klaster Kemiskinan Jawa Timur 2020

Berdasarkan peta klaster pada Gambar 4.4 dapat dilihat bahwa klaster 1 mengelompok di Jawa Timur bagian barat daya dan merupakan wilayah pesisir laut selatan. Klaster 2 tersebar merata mulai dari bagian barat hingga timur. Klaster 3 dan 5 sebagian besar tersebar di tengah wilayah Jawa Timur. Klaster 4 tersebar di pesisir utara Jawa Timur bagian timur. Kedekatan wilayah dan tersebarnya kabupaten/kota anggota klaster menunjukkan kemungkinan adanya efek spasial yang disebabkan oleh kondisi geografis.

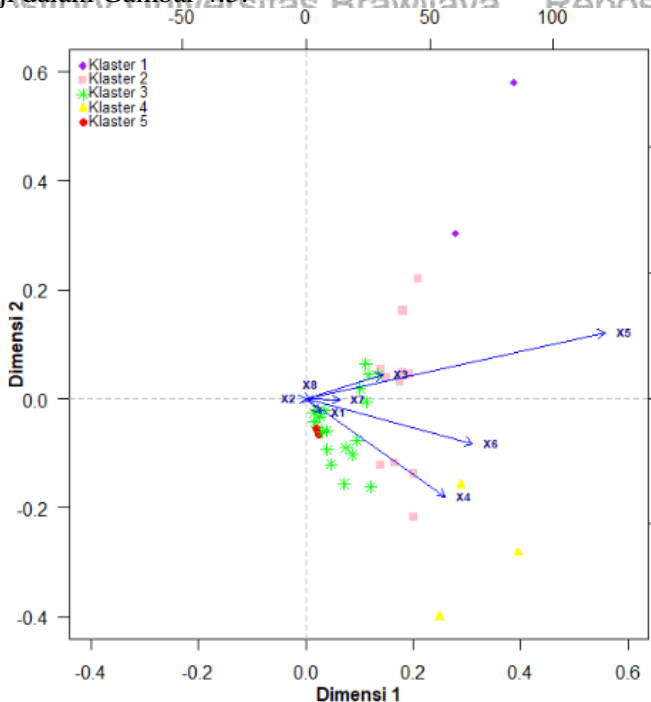


4.4. Analisis Biplot

Provinsi Jawa Timur dibagi ke dalam 5 kluster berdasarkan indikator kemiskinan, selanjutnya perlu dianalisis karakteristik dan kecenderungan antar variabel dengan kluster. Oleh karena itu, digunakan biplot untuk menggambarkan informasi tersebut. Biplot disusun berdasarkan data awal yang titik kabupaten/kotanya diberi simbol berbeda untuk setiap kluster. Kemudian indikator kemiskinan (variabel) dilambangkan sebagai vektor pada biplot. Berdasarkan kegunaannya, biplot disajikan ke dalam tiga tipe yaitu biplot *Column Metric Preserving*, *Row Metric Preserving* dan *Square Root*.

4.4.1. Biplot *Column Metric Preserving*

Biplot tipe *Column Metric Preserving* (CMP) menggunakan nilai $\alpha = 0$. Kegunaannya adalah menggambarkan kedekatan (korelasi) antar variabel pengamatan atau antar indikator kemiskinan. Hasil biplot tipe CMP kemiskinan kabupaten/kota di Jawa Timur tersaji dalam Gambar 4.5.



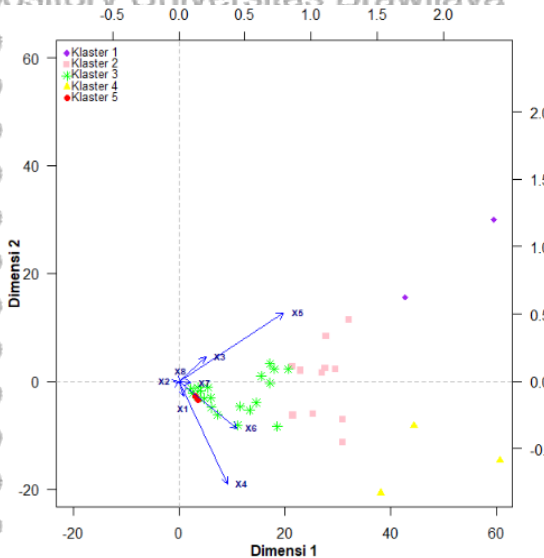
Gambar 4.5. Biplot Kemiskinan Kabupaten/Kota di Jawa Timur dengan $\alpha = 0$



Berdasarkan Gambar 4.5, antar vektor indikator kemiskinan menunjukkan sudut yang kecil ($< 90^\circ$). Dalam biplot, sudut antar vektor variabel yang semakin kecil dari 90° menunjukkan bahwa antar variabel semakin mirip atau dekat. Sedangkan vektor yang semakin panjang menunjukkan keragaman semakin besar dalam variabel tersebut. Oleh karena itu, dapat dijelaskan bahwa antar indikator saling berdekatan atau mirip, dengan kata lain antar indikator memiliki korelasi yang positif (searah). Panjang vektor pada biplot menunjukkan keragaman dalam variabel. Panjang vektor X_2 dan X_8 lebih kecil dibandingkan dengan vektor lainnya, sehingga dapat dikatakan bahwa kabupaten/kota di Jawa Timur kondisi sumber penerangan per rumah tangga dan pendidikan penduduknya hampir sama (keragamannya kecil).

4.4.2. Biplot Row Metric Preserving

Biplot tipe *Row Metric Preserving* (RMP) menggunakan $\alpha = 1$. Kegunaannya adalah memberikan tampilan biplot yang menekankan pada gambaran jarak antar kabupaten/kota, sehingga baik digunakan untuk melihat kedekatan antar kabupaten/kota. Hasil biplot RMP kemiskinan kabupaten/kota di Jawa Timur tersaji pada Gambar 4.6.



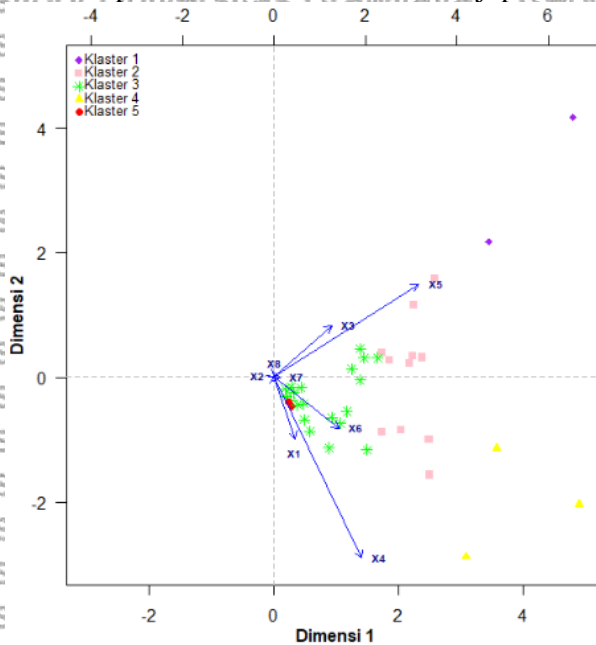
Gambar 4.6. Biplot Kemiskinan Kabupaten/Kota di Jawa Timur dengan $\alpha = 1$



Gambar 4.6 menunjukkan pembagian kabupaten/kota yang terbagi ke dalam 5 klaster sesuai dengan hasil analisis *hybrid hierarchical clustering via mutual cluster*. Dengan demikian biplot juga dapat menggambarkan kedekatan antar klaster. Klaster 5 sangat dekat dengan klaster 3, sehingga klaster 5 dapat dikatakan bertetangga dengan klaster 3. Klaster 3 dekat dengan klaster 5 dan 2, sehingga tetangga dari klaster 3 adalah klaster 5 dan 2. Anggota klaster 2 juga memiliki kedekatan dengan klaster 4 dan 1, sehingga klaster 2 juga bertetangga dengan klaster 4 dan 1.

4.4.3. Biplot *Square Root*

Biplot tipe *square root* menggunakan $\alpha = 0.5$. Kegunaannya untuk memberi penskalaan atau pembobotan untuk kedua faktor pengamatan (objek dan variabel). Oleh karena itu, biplot tipe *square root* baik untuk menggambarkan hubungan atau kecenderungan objek terhadap variabel atau sebaliknya. Hasil biplot *square root* kemiskinan kabupaten/kota di Jawa Timur tersaji pada Gambar 4.7.



Gambar 4.7. Biplot Kemiskinan Kabupaten/Kota di Jawa Timur dengan $\alpha = 0.5$



Pada Gambar 4.7 ditampilkan posisi relatif setiap objek dengan variabel, sehingga kecenderungan klaster terhadap indikator kemiskinan dapat dijelaskan. Kabupaten/kota yang ada di klaster 1 cenderung dekat dengan vektor X_3 dan X_5 . Dengan demikian, kabupaten/kota di klaster 1 memiliki karakteristik kemiskinan dengan persentase rumah tangga yang sumber air minum dan bahan bakar memasaknya tidak layak. Klaster 2 dan 3 cenderung memiliki kedekatan terhadap semua indikator, sehingga kedua klaster memiliki karakteristik kemiskinan yang hampir sama (kemiskinan multidimensi). Apabila dilihat dari biplot, posisi relatif klaster 2 lebih jauh dari titik (0,0), artinya kemiskinan multidimensi dan keragaman yang ada di dalam klaster 2 lebih tinggi dibanding klaster 3 (jauh dari rata-rata). Kabupaten/kota yang ada di klaster 4 cenderung dekat dengan vektor X_4 dan X_6 . Dengan demikian, kabupaten/kota di klaster 4 memiliki karakteristik kemiskinan multidimensi dengan persentase rumah tangga yang sanitasinya tidak layak dan kurang melek teknologi. Kabupaten/kota yang ada di klaster 5 cenderung dekat dengan ujung vektor X_1 , sehingga dapat dikatakan rumah tangga di klaster 5 memiliki tempat tinggal yang luas lantainya $< 20m^2$ dengan persentase yang cukup tinggi. Klaster 5 juga dekat dengan vektor X_4 namun jauh dari ujung vektor, sehingga dapat dikatakan bahwa rumah tangga di klaster 5 tidak memiliki sanitasi layak dengan persentase yang cukup rendah.

4.4.4. Proporsi Keragaman Kumulatif Biplot

Ketiga tipe biplot menggunakan data *input* yang sama, hanya saja berbeda nilai α -nya. Oleh karena itu, nilai akar cirinya tidak berubah (tetap). Proporsi keragaman kumulatif dari biplot kemiskinan kabupaten/kota di Jawa Timur, sebesar 68.5%. Artinya, biplot dapat menjelaskan 68.5% informasi dalam data. Nilai tersebut masih kurang dari minimal keragaman biplot 70%. Oleh sebab itu, dapat dikatakan bahwa biplot belum cukup dalam memberikan informasi dari delapan indikator kemiskinan.

4.5. Identifikasi Karakteristik Setiap Klaster

Setiap klaster kabupaten/kota diidentifikasi karakteristiknya berdasarkan rata-rata dan interpretasi biplot. Nilai rata-rata setiap klaster pada setiap variabel tersaji dalam Tabel 4.5.



Tabel 4.5. Rata-rata Setiap Klaster

Klaster	Rata-rata							
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	0.70	0.11	16.97	3.20	52.05	11.71	4.27	0.24
2	0.69	0.05	4.95	8.95	21.08	13.38	2.79	0.43
3	2.39	0.03	1.79	3.15	5.79	8.75	1.75	0.41
4	0.90	0	6.53	32.79	32.18	20.02	2.91	0.56
5	23.82	0	0.11	0.03	1.14	2.94	2.74	0.21

Berdasarkan Tabel 4.5 dan analisis biplot, karakteristik kemiskinan setiap klaster dijelaskan sebagai berikut.

a. Klaster 1 (Kemiskinan multidimensi sangat tinggi)

Menurut keragaman indikator pada biplot CMP dan nilai rata-rata, kemiskinan multidimensi klaster 1 tergolong sangat tinggi jika dibandingkan dengan klaster lain. Kabupaten/kota di klaster 1 cenderung memiliki karakteristik kemiskinan dengan sumber air minum yang tidak layak (X_3) sebesar 16.97% dan bahan bakar memasak yang kurang berkembang (X_5) sebesar 52.05%. Selain itu klaster 1 bertetangga dengan klaster 2, yang berarti setidaknya ada karakteristik kemiskinan yang mirip dari keduanya yaitu pada kondisi luas lantai bangunan $< 20m^2$.

b. Klaster 2 (Kemiskinan multidimensi sedang)

Apabila ditinjau dari keragaman indikator pada biplot CMP dan nilai rata-rata, kemiskinan multidimensi di klaster 2 tergolong sedang. Kabupaten/kota di klaster 2 memiliki karakteristik kemiskinan multidimensi yang mirip dengan klaster 3, di mana aspek yang lebih menonjol adalah bahan bakar memasak yang kurang berkembang (X_5) sebesar 21.08% dan kurangnya penguasaan teknologi (X_6) sebesar 13.38%. Selain pola kemiskinan multidimensi yang mirip, klaster 2 juga bertetangga dengan klaster 3, yang berarti setidaknya ada karakteristik kemiskinan yang mirip dari keduanya yaitu pada kondisi penduduk yang tidak sekolah pada usia 7-12 tahun.

c. Klaster 3 (Kemiskinan multidimensi rendah)

Berdasarkan keragaman indikator pada biplot CMP dan nilai rata-rata, kemiskinan multidimensi klaster 1 tergolong rendah. Kabupaten/kota di klaster 3 memiliki karakteristik kemiskinan multidimensi yang mirip dengan klaster 2, di mana aspek yang lebih menonjol adalah bahan bakar memasak yang kurang



berkembang berkembang (X_5) sebesar 5.79% dan kurangnya penguasaan teknologi (X_6) sebesar 8.75%.

d. **Klaster 4 (Kemiskinan multidimensi tinggi)**

Klaster 4 tergolong kemiskinan multidimensi tinggi apabila ditinjau dari keragaman indikator pada biplot CMP dan nilai rata-rata. Kabupaten/kota di klaster 4 cenderung memiliki karakteristik kemiskinan multidimensi dengan sanitasi yang tidak layak (X_4) sebesar 32.79% dan kurangnya penguasaan teknologi (X_6) sebesar 20.02%. Selain itu klaster 4 bertetangga dengan klaster 2, yang berarti setidaknya ada karakteristik kemiskinan yang mirip dari keduanya yaitu kondisi penduduk yang putus sekolah pada jenjang SMP.

e. **Klaster 5 (Kemiskinan multidimensi sangat rendah)**

Apabila ditinjau dari keragaman indikator pada biplot CMP dan nilai rata-rata, klaster 5 tergolong kemiskinan multidimensi yang sangat rendah. Kabupaten/kota di klaster 5 cenderung memiliki karakteristik kemiskinan multidimensi dengan aspek yang menonjol adalah luas bangunan tempat tinggal $< 20m^2$ (X_1) sebesar 23.82%. Selain itu klaster 5 bertetangga dengan klaster 3, yang berarti setidaknya ada karakteristik kemiskinan yang mirip dari keduanya, yaitu kondisi penduduk yang tidak sekolah pada usia 7-12 tahun.



BAB V PENUTUP

5.1. Kesimpulan

Berdasarkan analisis *hybrid hierarchical clustering via mutual cluster* dan biplot indikator kemiskinan kabupaten/kota di Jawa Timur pada tahun 2020, maka dapat disimpulkan bahwa:

1. Provinsi Jawa Timur dibagi ke dalam 5 klaster optimal dengan validitas rata-rata indeks *silhouette* sebesar 0.18 berdasarkan metode *hybrid hierarchical clustering via mutual cluster*.
2. Pada analisis biplot, proporsi keragaman yang dihasilkan sebesar 68.5% yang berarti biplot masih kurang dalam memberikan informasi dari delapan indikator kemiskinan. Biplot menunjukkan bahwa antar indikator kemiskinan saling berdekatan atau mirip, dengan kata lain antar indikator memiliki korelasi yang positif (searah). Selain itu, kondisi penerangan per rumah tangga dan pendidikan penduduk di Jawa Timur hampir sama (keragamannya kecil). Kemiskinan dari aspek penerangan per rumah tangga dan pendidikan penduduk cukup rendah.
3. Kemiskinan multidimensi di Jawa Timur dibagi menjadi 5 golongan yaitu kemiskinan multidimensi sangat tinggi (klaster 1), tinggi (klaster 4), sedang (klaster 2), rendah (klaster 3) dan sangat rendah (klaster 5). Di mana klaster 1 karakteristik kemiskinan yang menonjol adalah aspek sumber air minum tidak layak dan bahan bakar memasak yang kurang berkembang, Klaster 2 dan 3 karakteristik kemiskinan yang menonjol adalah aspek bahan bakar yang kurang berkembang dan kurangnya penguasaan teknologi. Klaster 4 karakteristik kemiskinan yang menonjol adalah aspek sanitasi tidak layak dan kurangnya penguasaan teknologi. Klaster 5 karakteristik kemiskinan yang menonjol adalah aspek luas bangunan tempat tinggal yang $< 20m^2$.

5.2. Saran

Berdasarkan hasil penelitian, maka saran yang dapat diberikan penulis untuk penelitian selanjutnya yaitu dalam *hybrid hierarchical clustering* dapat digunakan metode *bottom-up* yang lain seperti *single linkage*, *complete linkage*, *centroid* dan *Ward* untuk penentuan *mutual cluster*.



DAFTAR PUSTAKA

- Afandi, M. I. 2020. *Analisis Cluster Hierarki dengan Metode Complete Linkage pada Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator* (Skripsi UIN Maulana Malik Ibrahim). Diakses 26 Februari 2021 pukul 13.39 WIB melalui <http://etheses.uin-malang.ac.id/18458/>.
- Badan Pusat Statistik Jawa Timur. 2020. *Statistik Kesejahteraan Rakyat Provinsi Jawa Timur 2020*. Surabaya: BPS Provinsi Jawa Timur.
- Bappenas. 2020. *Tujuan 1 Tanpa Kemiskinan-SDGs Indonesia*. Diakses 4 Maret 2021 pukul 12.29 WIB melalui <http://sdgs.bappenas.go.id/tujuan-1/>.
- Chen, G., Jadarat, S. A., Baneerje, N., Tanaka, T. S., Ko, M. S. H. dan Zhang, M. Q. 2002. Evaluation and Comparison of Clustering Algorithms in Analyzing ES Cell Gen Expression Data. *Statistica Sinica*, 12, pp. 241-262.
- Chipman, H. dan Tibshirani, R. 2006. Hybrid Hierarchical Clustering With Applications To Microarray Data. *Biostatistics*, 7(2), pp. 286-301.
- Feil, B. dan Abonyi, J. 2007. *Cluster Analysis for Data Mining and System Identification*. Hungary: University of Pannonia.
- Gujarati, D. 2009. *Dasar-Dasar Ekonometrika*. Jakarta: Erlangga.
- Hair, J. F., Black, W.C., Babin, B. J., dan Anderson, R. E. 2010. *Multivariate Data Analysis 7th Edition*. New Jersey: Pearson Prentice Hall.
- Johnson, R. A. dan Wichern, D.W. 2007. *Applied Multivariate Statistical Analysis 6th Edition*. New Jersey: Pearson Prentice Hall.
- Kaufman, L. dan Rousseeuw, P.J. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: John Wiley & Sons.
- Leleury, Z. A. dan Wokanubun, A. E. 2015. Analisis Biplot Pada Pemetaan Karakteristik Kemiskinan di Provinsi Maluku. *Jurnal Ilmu Matematika dan Terapan*, 9(1), Hal. 21-31.
- Mattjik, A. A. dan Sumertajaya I. M. 2011. *Sidik Peubah Ganda*. Bogor: IPB Press.
- Montgomery, Douglas. 2009. *Introduction to Statistical Quality Control*. New York: John Wiley & Sons, Inc.



- Mustafidah, Rohmah. 2017. *Pengelompokan Kabupaten/Kota di Provinsi Jawa Tengah Berdasarkan Indikator Kemiskinan dengan C-Means dan Fuzzy C-Means Clustering* (Tugas Akhir, ITS). Diakses 26 Februari 2021 pukul 13.39 melalui <http://repository.its.ac.id/48528/>.
- Nugroho, S. 2008. *Statistika Multivariat Terapan*. Bengkulu: UNIB Press.
- Rastantra, V. A. 2018. *Kemiskinan Multidimensi Provinsi dan Kabupaten/Kota di Jawa Timur Tahun 2017* (Artikel, STIS). Diakses 4 Maret 2021 pukul 13.03 WIB melalui https://www.researchgate.net/publication/335501515_Kemiskinan_Multidimensi_Provinsi_dan_KabupatenKota_di_Jawa_Timur_Tahun_2017.
- Sartono, B., Affendi F. M., Syafitri U. D., Sumertajaya I. M. dan Anggraeni Y. 2003. *Analisis Peubah Ganda*. Bogor: IPB.
- Speed, T. 2003. *Statistical Analysis of Gene Expression Microarray Data*. Florida: CRC Press.
- Supranto, J. 2004. *Analisis Multivariat Arti & Interpretasi*. Jakarta: PT Rineka Cipta.
- Todaro, M. P. dan Smith, S. C. 2012. *Economic Development 11th Edition*. USA: Addison-Wesley.
- Usman, H. 2013. *Aplikasi Teknik Multivariat Untuk Riset Pemasaran*. Jakarta: PT Grefindo Persada.
- Wahyuni, S. dan Yogo A. J. 2018. *Pengelompokan Kabupaten/Kota di Pulau Jawa Berdasarkan Faktor-Faktor Kemiskinan dengan Pendekatan Average Linkage Hierarchical Clustering* (Artikel, STIS). Diakses 4 Maret 2021 pukul 15.39 WIB melalui <https://jurnal.stis.ac.id/index.php/jurnalasks/article/view/197>.
- Walpole, R. E. 1995. *Pengantar Statistika*. Edisi Ketiga. Terjemahan Bambang Sumantri. Jakarta: PT. Gramedia Pustaka Utama.
- Wijaya, S. B. 2017. *Perbandingan Metode Pautan Kombinasi Antara Analisis Biplot dengan Analisis Cluster Hierarki (Studi Kasus pada Ketahanan Sosial, ekonomi dan Ekologi Desa Bendosari Kecamatan Pujon Kabupaten Malang)* (Skripsi, Universitas Brawijaya).

LAMPIRAN

Lampiran 1. Data Indikator Kemiskinan Kabupaten/Kota di Jawa Timur Tahun 2020 (%)

Kode	Kabupaten/Kota	X1	X2	X3	X4	X5	X6	X7	X8
1	Pacitan	0.92	0.16	15.39	0.92	63.88	12.45	2.75	0.18
2	Ponorogo	0.62	0	8.85	2.06	30.74	11.9	1.07	0
3	Trenggalek	0.48	0.05	18.55	5.47	40.21	10.97	5.78	0.3
4	Tulungagung	0.74	0	2.73	0.45	14.34	11.48	1.53	0.42
5	Blitar	1.16	0.11	4.55	1.45	26.09	12.74	1.81	0.61
6	Kediri	0.83	0.08	0.93	5.56	8.83	12.19	0.53	0.97
7	Malang	1.29	0	4.4	2.76	11.62	13.55	2.04	0.04
8	Lumajang	0.42	0	5.82	8.55	24.47	13.13	2.15	0.94
9	Jember	0.61	0	9.22	22.12	18.09	14.9	2.25	0.27
10	Banyuwangi	1.2	0.09	6.85	13.27	12.98	11.21	1.41	0.26
11	Bondowoso	1.1	0	8.16	38.75	43.21	24.24	2.01	0.32
12	Situbondo	1.13	0	3.82	32.75	20.92	20.09	0.89	0.77
13	Probolinggo	0.48	0	7.6	26.88	32.42	15.72	5.83	0.58
14	Pasuruan	2.13	0	2.66	9.51	8.8	16.71	2.4	0.38
15	Sidoarjo	6.49	0.26	0.41	0.93	0.49	3.23	0.51	0.42



Lampiran 2. (Lanjutan)

Kode	Kabupaten/Kota	X1	X2	X3	X4	X5	X6	X7	X8
16	Mojokerto	2.98	0	0.44	4.77	1.61	9.01	1.03	0.3
17	Jombang	1.37	0	0.36	7.45	3.59	12.01	1.03	0.39
18	Nganjuk	0.62	0	0.31	5.14	6.72	13.75	0.6	0.38
19	Madiun	0.45	0	1.98	3.05	18.14	12.37	3	0.91
20	Magetan	0	0	0.54	2.64	12.67	10.14	0.5	0.17
21	Ngawi	0.19	0	0.51	4.25	21.95	18.19	0.72	0.38
22	Bojonegoro	0.09	0	2.34	3.68	18.53	14.26	3.51	0.15
23	Tuban	0.84	0	1.36	12.52	16.91	15.03	5.65	0
24	Lamongan	0.13	0.21	5.54	5.4	4.55	10.96	0.47	0.39
25	Gresik	5.02	0	0.9	0.63	0.31	5.83	0.53	0.34
26	Bangkalan	0.73	0	4.54	2.92	14.11	9.52	7.63	0.52
27	Sampang	0.71	0	11.24	10.83	22.4	7.35	5.49	1.26
28	Pamekasan	0.79	0	9.38	5.4	15.83	9.26	2.54	0
29	Sumenep	1.32	0.36	1.77	16.61	21.6	16.09	3.67	0
30	Kota Kediri	5.39	0	0.11	0	1.63	5.56	1.11	0.4
31	Kota Blitar	5.29	0	1.63	0.62	2.83	5.11	2.78	0
32	Kota Malang	20.23	0	0	0	0.82	2.75	4.81	0.32



Lampiran 3. (Lanjutan)

Kode	Kabupaten/Kota	X1	X2	X3	X4	X5	X6	X7	X8
33	Kota Probolinggo	2.84	0	0	3.54	3.1	4.53	1.84	0.51
34	Kota Pasuruan	2.78	0	0	3.33	1.36	7.64	4.49	0.72
35	Kota Mojokerto	2.31	0	0	0.41	0.75	3.29	0	0.44
36	Kota Madiun	4.3	0	0	0	1.68	4.23	1.9	1
37	Kota Surabaya	27.41	0	0.21	0.06	1.46	3.12	0.67	0.1
38	Kota Batu	1.75	0.07	0.98	1.52	0.92	7	1.63	0.5

Lampiran 2. Statistik Deskriptif Indikator Kemiskinan Provinsi Jawa Timur

```
> #statistik deskriptif
> summary(data1)
```

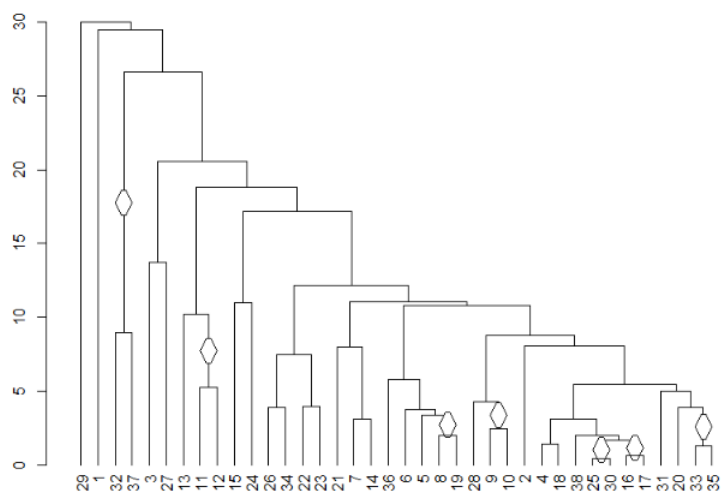
X1		X2		X3		X4	
Min.	: 0.000	Min.	:0.00000	Min.	: 0.0000	Min.	: 0.000
1st Qu.:	0.620	1st Qu.:	0.00000	1st Qu.:	0.4175	1st Qu.:	1.060
Median :	1.115	Median :	0.00000	Median :	1.8750	Median :	3.610
Mean :	2.819	Mean :	0.03658	Mean :	3.7916	Mean :	7.005
3rd Qu.:	2.663	3rd Qu.:	0.00000	3rd Qu.:	5.7500	3rd Qu.:	8.275
Max.	:27.410	Max.	:0.36000	Max.	:18.5500	Max.	:38.750

X5		X6		X7		X8	
Min.	: 0.310	Min.	: 2.750	Min.	:0.000	Min.	:0.0000
1st Qu.:	1.968	1st Qu.:	7.088	1st Qu.:	0.925	1st Qu.:	0.2000
Median :	12.825	Median :	11.345	Median :	1.870	Median :	0.3800
Mean :	14.488	Mean :	10.829	Mean :	2.331	Mean :	0.4116
3rd Qu.:	21.430	3rd Qu.:	13.700	3rd Qu.:	2.945	3rd Qu.:	0.5175
Max.	:63.880	Max.	:24.240	Max.	:7.630	Max.	:1.2600



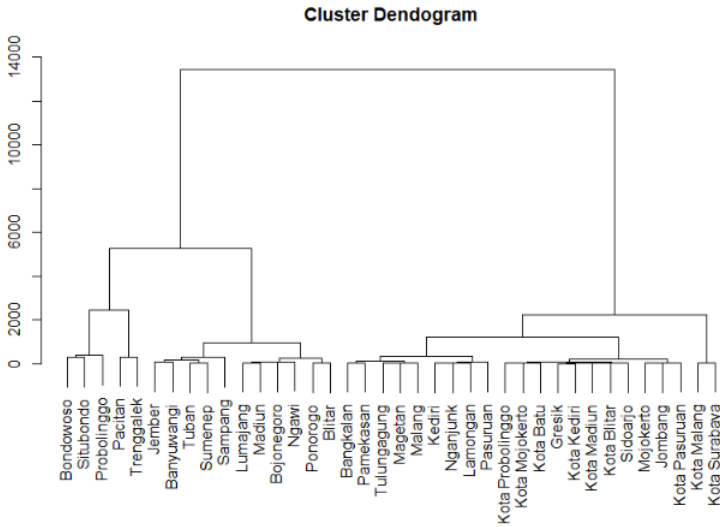


Lampiran 3. Dendogram *Mutual Cluster*



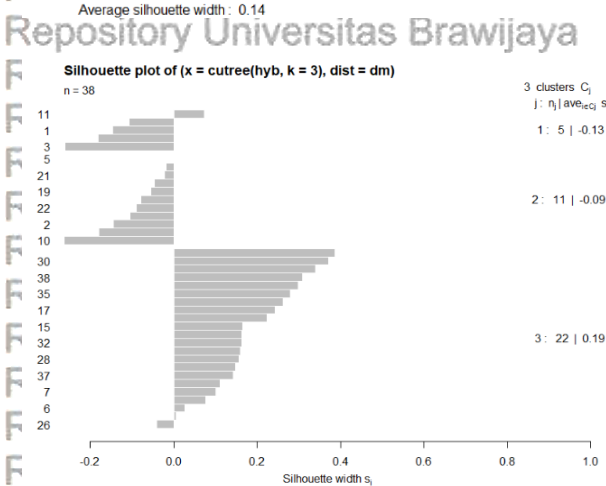
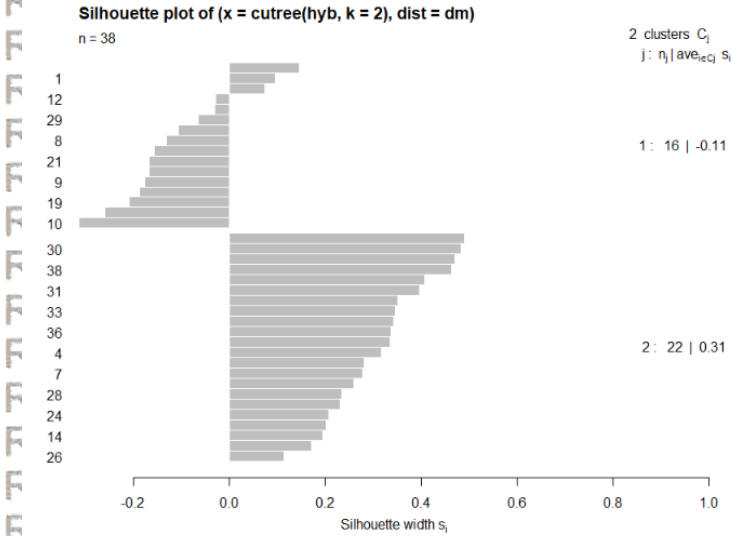


Lampiran 4. Dendrogram *Hybrid Hierarchical Clustering via Mutual Cluster*





Lampiran 5. Plot Indeks *Silhouette* 2-15 Klaster

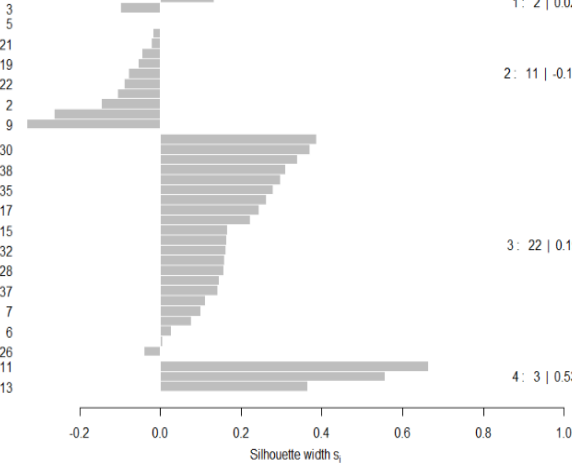




Lampiran 5. (Lanjutan)

Silhouette plot of (x = cutree(hyb, k = 4), dist = dm)

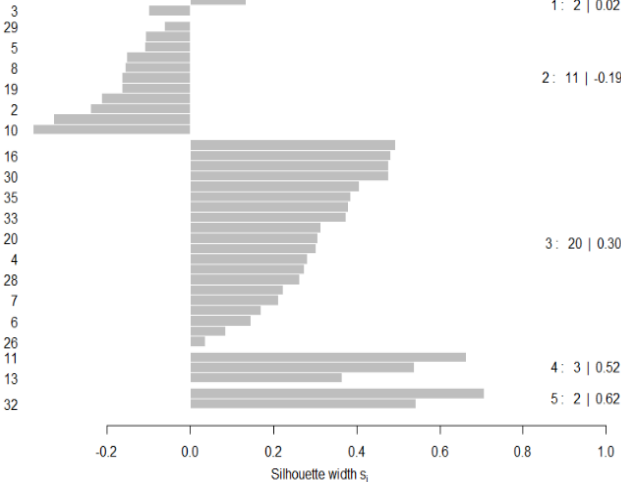
n = 38



Average silhouette width : 0.12

Silhouette plot of (x = cutree(hyb, k = 5), dist = dm)

n = 38



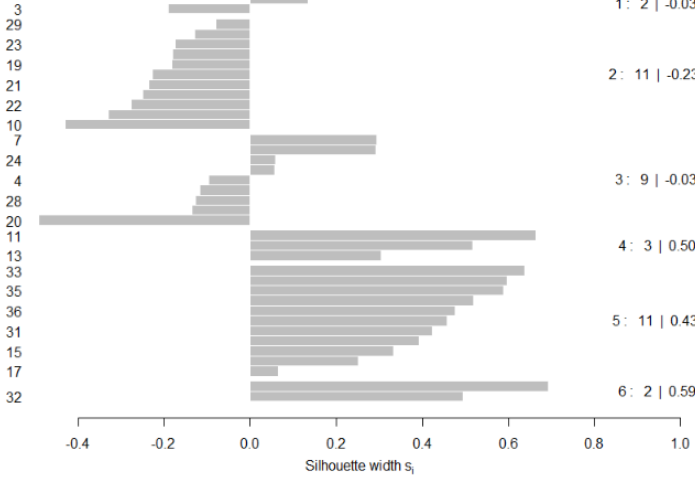
Average silhouette width : 0.18



Lampiran 5. (Lanjutan)

Silhouette plot of (x = cutree(hyb, k = 6), dist = dm)

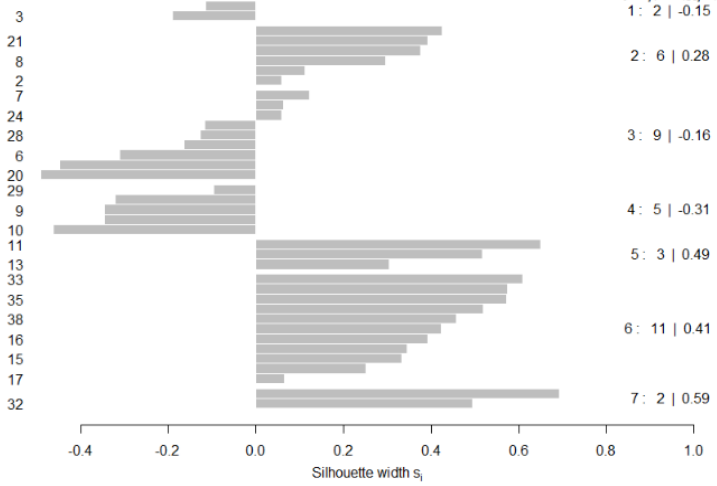
n = 38



Average silhouette width : 0.12

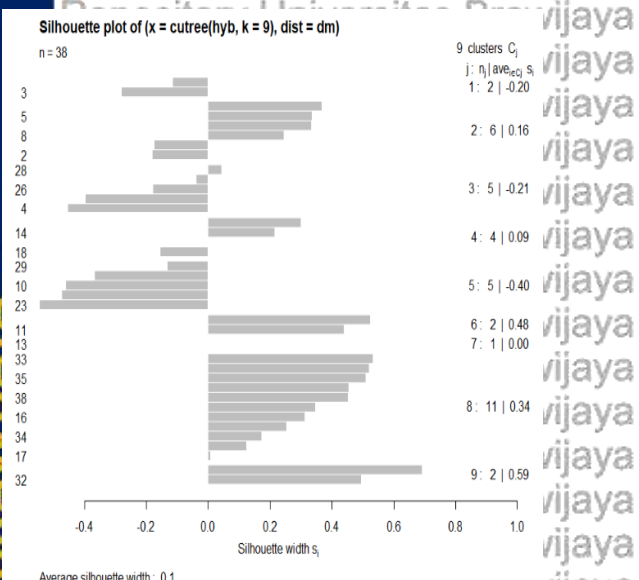
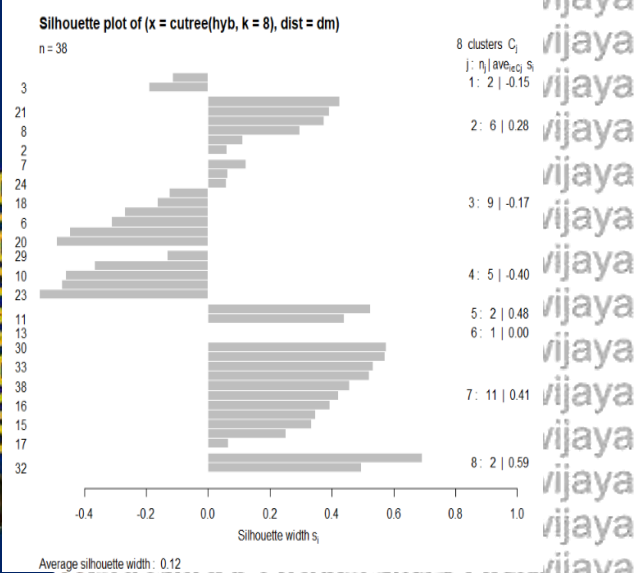
Silhouette plot of (x = cutree(hyb, k = 7), dist = dm)

n = 38



Average silhouette width : 0.15

Lampiran 5. (Lanjutan)

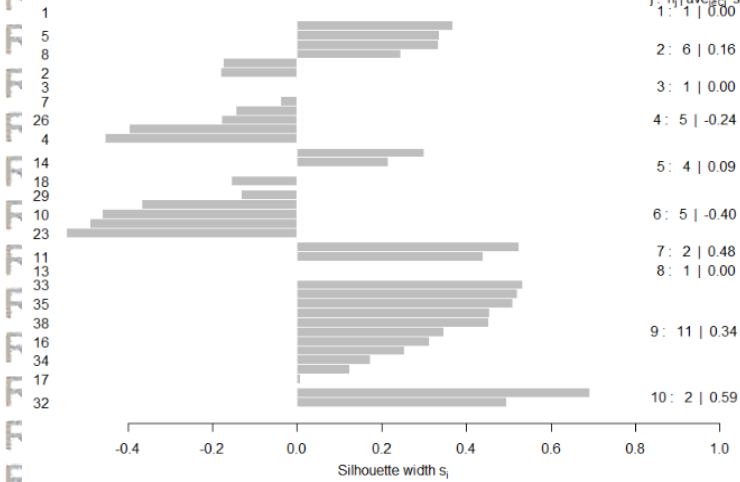




Lampiran 5. (Lanjutan)

Silhouette plot of (x = cutree(hyb, k = 10), dist = dm)

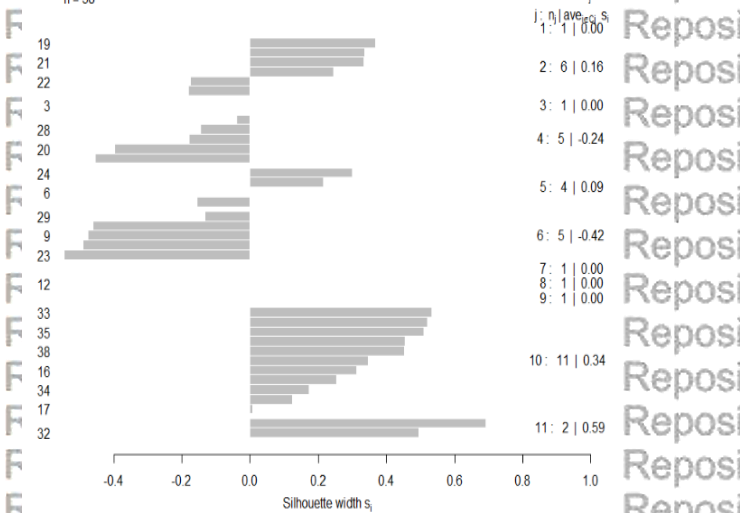
n = 38



Average silhouette width : 0.1

Silhouette plot of (x = cutree(hyb, k = 11), dist = dm)

n = 38



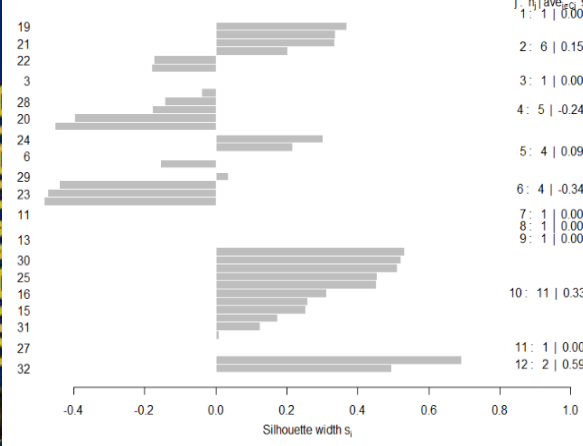
Average silhouette width : 0.07



Lampiran 5. (Lanjutan)

Silhouette plot of (x = cutree(hyb, k = 12), dist = dm)

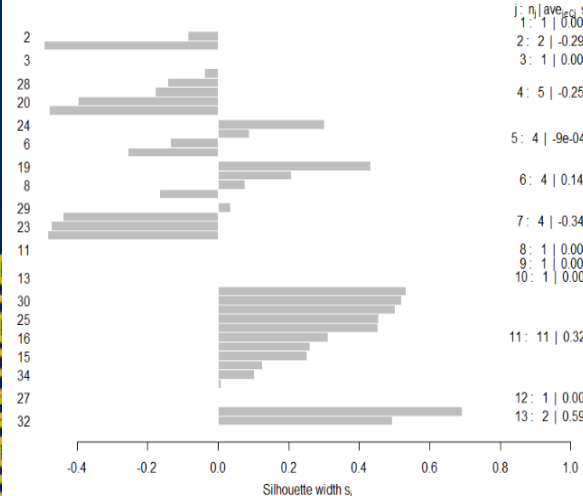
n = 38



Average silhouette width : 0.09

Silhouette plot of (x = cutree(hyb, k = 13), dist = dm)

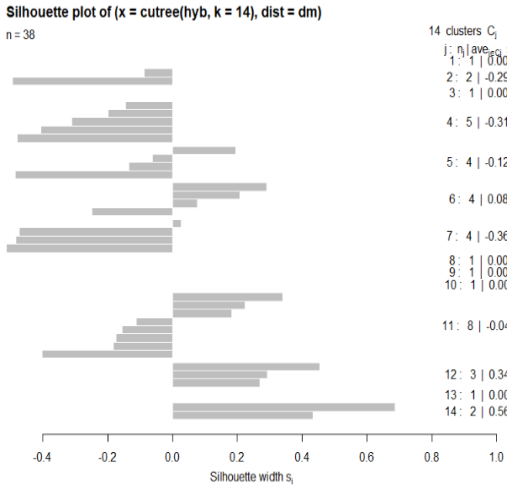
n = 38



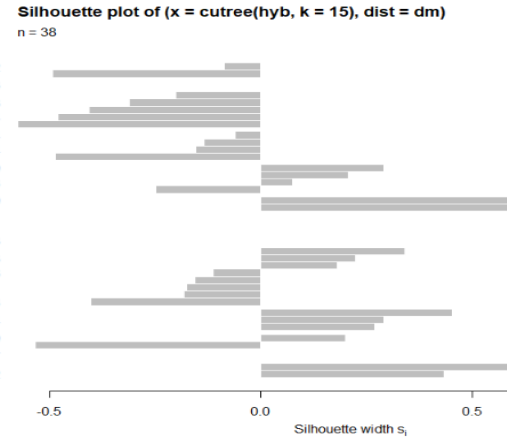
Average silhouette width : 0.05



Lampiran 5. (Lanjutan)



Average silhouette width : -0.05

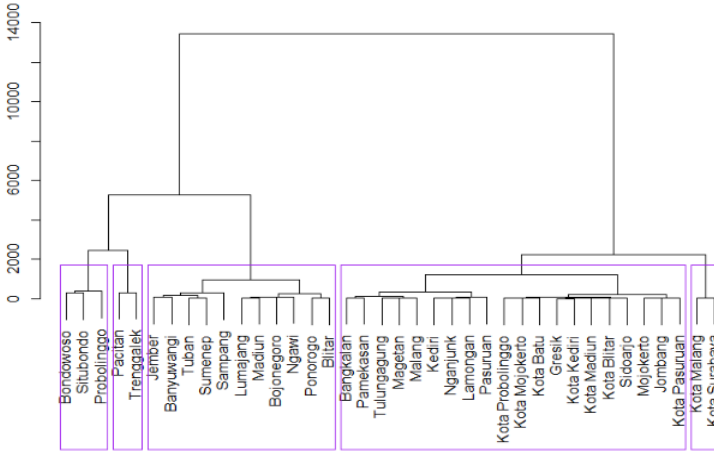


Average silhouette width : 0



Lampiran 6. Dendrogram dengan 5 Klaster Optimal

Cluster Dendrogram

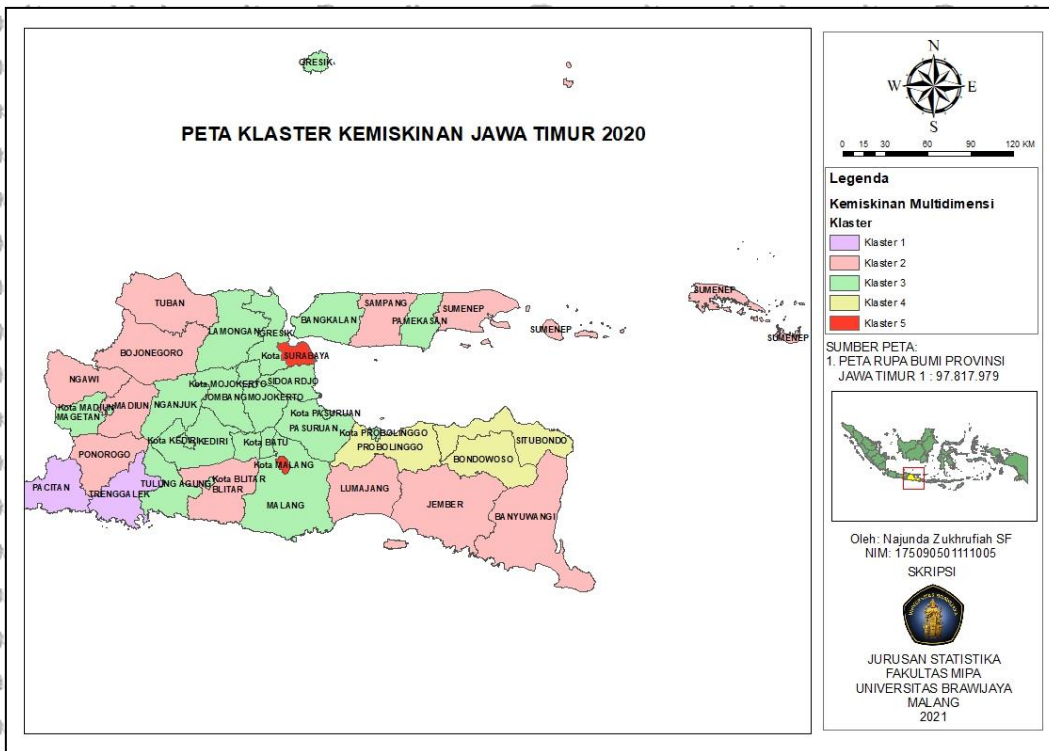


Lampiran 7. Keanggotaan Klaster

Kode	Kabupaten. Kota	X1	X2	X3	X4	X5	X6	X7	X8	klaster
1	Pacitan	0.92	0.16	15.39	0.92	63.88	12.45	2.75	0.18	1
3	Trenggalek	0.48	0.05	18.55	5.47	40.21	10.97	5.78	0.30	1
2	Ponorogo	0.62	0.00	8.85	2.06	30.74	11.90	1.07	0.00	2
5	Blitar	1.16	0.11	4.55	1.45	26.09	12.74	1.81	0.61	2
8	Lumajang	0.42	0.00	5.82	8.55	24.47	13.13	2.15	0.94	2
9	Jember	0.61	0.00	9.22	22.12	18.09	14.90	2.25	0.27	2
10	Banyuwangi	1.20	0.09	6.85	13.27	12.98	11.21	4.41	0.26	2
19	Madiun	0.45	0.00	1.98	3.05	18.14	12.37	3.00	0.91	2
21	Ngawi	0.19	0.00	0.51	4.25	21.95	18.19	0.72	0.38	2
22	Bojonegoro	0.09	0.00	2.34	3.68	18.53	14.26	3.51	0.15	2
23	Tuban	0.84	0.00	1.36	12.52	16.91	15.03	5.65	0.00	2
27	Sampang	0.71	0.00	11.24	10.83	22.40	7.35	5.49	1.26	2
29	Sumenep	1.32	0.36	1.77	16.61	21.60	16.09	3.67	0.00	2
4	Tulungagung	0.74	0.00	2.73	0.45	14.34	11.48	1.53	0.42	3
6	Kediri	0.83	0.08	0.93	5.56	8.83	12.19	0.53	0.97	3
7	Malang	1.29	0.00	4.40	2.76	11.62	13.55	2.04	0.04	3
14	Pasuruan	2.13	0.00	2.66	9.51	8.80	16.71	2.40	0.38	3
15	Sidoarjo	6.49	0.26	0.41	0.93	0.49	3.23	0.51	0.42	3
16	Mojokerto	2.98	0.00	0.44	4.77	1.61	9.01	1.03	0.30	3
17	Jombang	1.37	0.00	0.36	7.45	3.59	12.01	1.03	0.39	3
18	Nganjuk	0.62	0.00	0.31	5.14	6.72	13.75	0.60	0.38	3
20	Magetan	0.00	0.00	0.54	2.64	12.67	10.14	0.50	0.17	3
24	Lamongan	0.13	0.21	5.54	5.40	4.55	10.96	0.47	0.39	3
25	Gresik	5.02	0.00	0.90	0.63	0.31	5.83	0.53	0.34	3
26	Bangkalan	0.73	0.00	4.54	2.92	14.11	9.52	7.63	0.52	3
28	Pamekasan	0.79	0.00	9.38	5.40	15.83	9.26	2.54	0.00	3
30	Kota Kediri	5.39	0.00	0.11	0.00	1.63	5.56	1.11	0.40	3
31	Kota Blitar	5.29	0.00	1.63	0.62	2.83	5.11	1.78	0.00	3
33	Kota Probolinggo	2.84	0.00	0.00	3.54	3.10	4.53	1.84	0.51	3
34	Kota Pasuruan	2.78	0.00	0.00	3.33	1.36	7.64	4.49	0.72	3
35	Kota Mojokerto	2.31	0.00	0.00	0.41	0.75	3.29	0.00	0.44	3
36	Kota Madiun	4.30	0.00	0.00	0.00	1.68	4.23	1.90	1.00	3
38	Kota Batu	1.75	0.07	0.98	1.52	0.92	7.00	1.63	0.50	3
11	Bondowoso	1.10	0.00	8.16	38.75	43.21	24.24	2.01	0.32	4
12	Situbondo	1.13	0.00	3.82	32.75	20.92	20.09	0.89	0.77	4
13	Probolinggo	0.48	0.00	7.60	26.88	32.42	15.72	5.83	0.58	4
32	Kota Malang	20.23	0.00	0.00	0.00	0.82	2.75	4.81	0.32	5
37	Kota Surabaya	27.41	0.00	0.21	0.06	1.46	3.12	0.67	0.10	5



Lampiran 8. Peta Klaster Kemiskinan Jawa Timur Tahun 2020





Lampiran 9. Nilai Akar Ciri

F > y

5d

F [1]	153.4968067	51.7587150	38.1683541	27.1001153	16.0357541
	10.5010545	2.0725845	0.4982902		

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

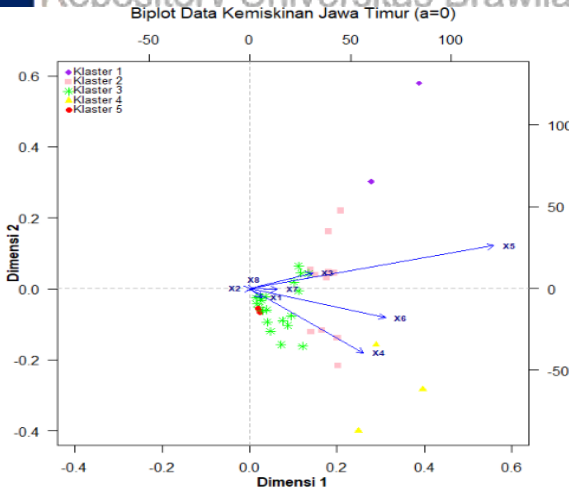
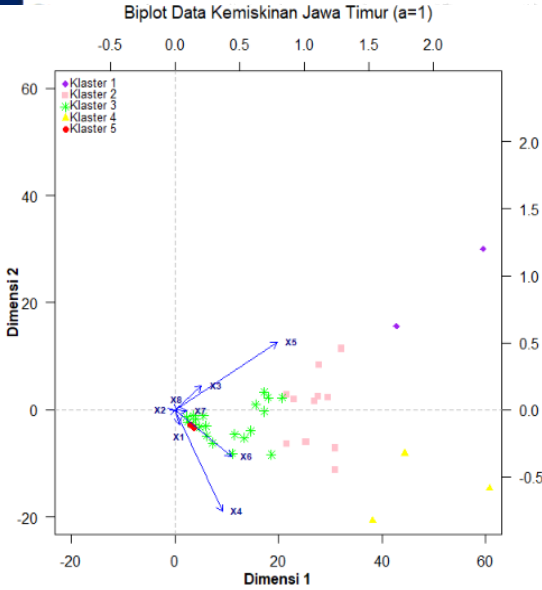
Repository Universitas Brawijaya

Repository Universitas Brawijaya

Repository Universitas Brawijaya

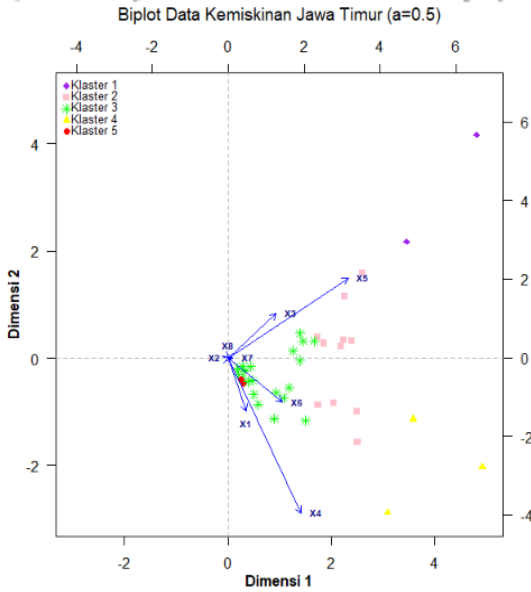


Lampiran 10. Biplot ($\alpha = 1,0$ dan $0,5$)





Lampiran 10. (Lanjutan)





Lampiran 11. *Source Code Analysis Hybrid Hierarchical Clustering via Mutual Cluster*

```

#####import data (excel)#####
library(readxl)
data0 <- as.data.frame(read_xlsx("F:/SMT-
8/SKRIPSI/dataskrip.xlsx"))
head(data0)
data1 <- data0[, -(1:2)]
head(data1)
#####statistik deskriptif#####
summary(data1)
#####Pemeriksaan nonmultikolinieritas variabel#####
library(car)
aux <- lm(data0$kode ~., data=data1)
vif(aux)
#####Jarak Mahalanobis#####
library(bitools)
dm<- D2.dist(data1,cov(data1))
#####Mutual Cluster#####
library(hybridHclust)
mc1 <- mutualCluster(distance=dm,method = "average",
plot=TRUE)
mc1
get.distances(mc1)
#####hybrid hierarchical clustering#####
hyb <- hybridHclust(data1, mc1, trace=TRUE)
plot(hyb,data0$ Kabupaten/Kota main = "Cluster Dendogram")
#####validasi#####
library(c1Valid)
si2 <- silhouette(cutree(hyb, k =2 ),dist = dm)
plot(si2)
si3 <- silhouette(cutree(hyb, k =3 ),dist = dm)
plot(si3)
si4 <- silhouette(cutree(hyb, k =4 ),dist = dm)
plot(si4)
si5 <- silhouette(cutree(hyb, k =5 ),dist = dm)
plot(si5)
si6 <- silhouette(cutree(hyb, k =6 ),dist = dm)
plot(si6)
si7 <- silhouette(cutree(hyb, k =7 ),dist = dm)
plot(si7)
si8 <- silhouette(cutree(hyb, k =8 ),dist = dm)
plot(si8)
si9 <- silhouette(cutree(hyb, k =9 ),dist = dm)

```




Lampiran 11. (Lanjutan)

```
plot(s10)
s110 <- silhouette(cutree(hyb, k =10 ),dist = dm)
plot(s110)
s111 <- silhouette(cutree(hyb, k =11 ),dist = dm)
plot(s111)
s112 <- silhouette(cutree(hyb, k =12 ),dist = dm)
plot(s112)
s113 <- silhouette(cutree(hyb, k =13 ),dist = dm)
plot(s113)
s114 <- silhouette(cutree(hyb, k =14 ),dist = dm)
plot(s114)
s115 <- silhouette(cutree(hyb, k =15 ),dist = dm)
plot(s115)
#####pembagian 5 klaster pada Dendrogram#####
plot(hyb,data0$ Kabupaten/Kota , main = "Cluster Dendrogram")
rect.hclust(hyb, k=5, border= "purple")
#####anggota klaster#####
library(plyr)
klaster <- cutree(hyb,k=5)
data2 <- arrange(data.frame(data0,klaster), klaster)
data2
```

Lampiran 12. *Source Code* Analisis Biplot

```

library(graphics)
x=data1
head(x)
y=svd(x)
y
U=y$u
L=diag(y$d)
A=y$V
#==== Biplot Alpha(θ) =====
G=U
H=A%*%L
G0 <- G[,1:2]
H0 <- H[,1:2]
#membuat plot frame baru
par()
par(pty = "s",
    cex.main = 1.2,
    cex.lab = 1,
    font.main = 1,
    font.lab = 2,
    family = "sans",
    col.main = "black",
    col.lab = "black",
    fg = "black",
    las = 1)
plot.new()
plot.window(xlim = c(-0.4, 0.6),
            ylim = c(-0.4, 0.6),
            asp = 1)
axis(side = 1,
     at = c(-0.4, -0.2, 0.0, 0.2, 0.4, 0.6),
     labels = TRUE)
axis(side = 2,
     at = c(-0.4, -0.2, 0.0, 0.2, 0.4, 0.6),
     labels = TRUE)
title(main = "Biplot Data Kemiskinan Jawa Timur (a=0)",
      line = 3,
      adj = 0.5)
#menambahkan Label X dan Y
title(xlab = "Dimensi 1",
      ylab = "Dimensi 2",
      line = 2,
      adj = 0.5)
#menambahkan point objek

```




Lampiran 12. (Lanjutan)

```

points(x = G0,
       pch =
c(18,15,18,8,15,8,8,15,15,15,17,17,17,8,8,8,8,15,
   8,15,15,15,8,8,8,15,8,15,8,8,16,8,8,8,8,16,8),
       cex = 1,
       col =
c("purple","pink","purple","green","pink","green",
  "green","pink","pink","pink","yellow","yellow",
  "yellow","green","green","green","green","green",
  "pink","green","pink","pink","pink","green","green",
  "green","pink","green","pink","green","green","red",
  "green","green","green","green","red","green"))
#menambahkan axis untuk vektor variabel
par(new = TRUE, las = 1)
plot.window(xlim = c(-87,130),
            ylim = c(-87,130),
            asp = 1)
axis(side = 3,
     at = c(-50, 0, 50, 100),
     col = "black")
axis(side = 4,
     at = c(-50, 0, 50, 100),
     col = "black")
#menambahkan kotak
box()
#menambahkan sumbu
abline(v = 0, h = 0,lty = 2, col = "grey")
#menambahkan vektor variabel
arrows(x0 = 0, x1 = H0[,1],
       y0 = 0, y1 = H0[,2],
       length = 0.08,
       col = "blue")
#menambahkan label vektor
text(x = H0[,1], y = H0[,2],labels = colnames(data1),
     font = 2,
     cex = 0.6,
     col = "navy",
     pos = c(4,2,4,4,4,4,3))
#menambahkan legenda
legend(x = "topleft",
      legend = c("Klaster 1", "Klaster 2", "Klaster 3",
                "Klaster 4", "Klaster 5"),
      pch = c(18, 15, 8, 17, 16),
      col = c("purple", "pink", "green", "yellow", "red"),

```



Lampiran 12. (Lanjutan)

```

cex = 0.8,
pt.cex = 1,
bty = "n",
x.intersp = 0.5,
y.intersp = 0.8,
xpd = FALSE,
adj = c(0, 0.25))

##### Biplot Alpha(1) #####
G=U%*L
H=A
G1 <- G[,1:2]
H1 <- H[,1:2]
#membuat plot frame baru
par()
par(pty = "s",
    cex.main = 1.2,
    cex.lab = 1,
    font.main = 1,
    font.lab = 2,
    family = "sans",
    col.main = "black",
    col.lab = "black",
    fg = "black",
    las = 1)
plot.new()
plot.window(xlim = c(-20, 60),
            ylim = c(-20, 60),
            asp = 1)
axis(side = 1,
     at = c(-20, 0, 20, 40, 60),
     labels = TRUE)
axis(side = 2,
     at = c(-20, 0, 20, 40, 60),
     labels = TRUE)
title(main = "Biplot Data Kemiskinan Jawa Timur (a=1)",
      line = 3,
      adj = 0.5)
#menambahkan Label X dan Y
title(xlab = "Dimensi 1",
      ylab = "Dimensi 2",
      line = 2,
      adj = 0.5)
#menambahkan point objek

```




Lampiran 12. (Lanjutan)

```

points(x = G1,
       pch =
c(18,15,18,8,15,8,8,15,15,15,17,17,17,8,8,8,8,8,15,
   8,15,15,15,8,8,8,15,8,15,8,8,16,8,8,8,8,16,8),
       cex = 1,
       col =
c("purple", "pink", "purple", "green", "pink", "green",
  "green", "pink", "pink", "pink", "yellow", "yellow",
  "yellow", "green", "green", "green", "green", "green",
  "pink", "green", "pink", "pink", "pink", "green", "green",
  "green", "pink", "green", "pink", "green", "green", "red",
  "green", "green", "green", "green", "red", "green"))
#menambahkan axis untuk vektor variabel
par(new = TRUE, las = 1)
plot.window(xlim = c(-0.8,2.4),
            ylim = c(-0.8,2.4),
            asp = 1)
axis(side = 3,
     at = c(-0.5, 0, 0.5,1,1.5,2),
     col = "black")
axis(side = 4,
     at = c(-0.5, 0, 0.5,1,1.5,2),
     col = "black")
#menambahkan kotak
box()
#menambahkan sumbu
abline(v = 0, h = 0, lty = 2, col = "grey")
#menambahkan vektor variabel
arrows(x0 = 0, x1 = H1[,1],
       y0 = 0, y1 = H1[,2],
       length = 0.08,
       col = "blue")
#menambahkan label vektor
text(x = H1[,1], y = H1[,2], labels = colnames(data1),
     font = 2,
     cex = 0.6,
     col = "navy",
     pos = c(1,2,4,4,4,4,3))
#menambahkan legenda
legend(x = "topleft",
      legend = c("Klaster 1", "Klaster 2", "Klaster 3",
                "Klaster 4", "Klaster 5"),
      pch = c(18, 15, 8, 17, 16),
      col = c("purple", "pink", "green", "yellow", "red"),
      cex = 0.8,

```



Lampiran 12. (Lanjutan)

```

pt.cex = 1,
  bty = "n",
  x.intersp = 0.5,
  y.intersp = 0.8,
  xpd = FALSE,
  adj = c(0,0.25))

##### Biplot Alpha(0.5) #####
G=U%*(L^0.5)
H=(L^0.5)%*A
G5 <- G[,1:2]
H5 <- H[,1:2]
#membuat plot frame baru
par()
par(mty = "s",
    cex.main = 1.2,
    cex.lab = 1,
    font.main = 1,
    font.lab = 2,
    family = "sans",
    col.main = "black",
    col.lab = "black",
    fg = "black",
    las = 1)
plot.new()
plot.window(xlim = c(-3, 5),
            ylim = c(-3, 5),
            asp = 1)
axis(side = 1,
     at = c(-2,0,2,4),
     labels = TRUE)
axis(side = 2,
     at = c(-2,0,2,4),
     labels = TRUE)
title(main = "Biplot Data Kemiskinan Jawa Timur (a=0.5)",
      line = 3,
      adj = 0.5)
#menambahkan Label X dan Y
title(xlab = "Dimensi 1",
      ylab = "Dimensi 2",
      line = 2,
      adj = 0.5)
#menambahkan point objek
points(x = G5,

```




Lampiran 12. (Lanjutan)

```

pch =
c(18,15,18,8,15,8,8,15,15,15,17,17,17,8,8,8,8,15,
8,15,15,15,8,8,8,15,8,15,8,8,16,8,8,8,16,8),
cex = 1,
col =
c("purple","pink","purple","green","pink","green",
"green","pink","pink","pink","yellow","yellow",
"yellow","green","green","green","green","green",
"pink","green","pink","pink","pink","green","green",
"green","pink","green","pink","green","green","red",
"green","green","green","green","red","green"))
#menambahkan axis untuk vektor variabel
par(new = TRUE, las = 1)
plot.window(xlim = c(-4.1,6.8),
ylim = c(-4,1,6.8),
asp = 1)
axis(side = 3,
at = c(-4,-2,0,2,4,6),
col = "black")
axis(side = 4,
at = c(-4,-2,0,2,4,6),
col = "black")
#menambahkan kotak
box()
#menambahkan sumbu
abline(v = 0, h = 0, lty = 2, col = "grey")
#menambahkan vektor variabel
arrows(x0 = 0, x1 = H5[,1],
y0 = 0, y1 = H5[,2],
length = 0.08,
col = "blue")
#menambahkan label vektor
text(x = H5[,1], y = H5[,2], labels = colnames(data1),
font = 2,
cex = 0.6,
col = "navy",
pos = c(1,2,4,4,4,4,3))
#menambahkan legenda
legend(x = "topleft",
legend = c("Klaster 1", "Klaster 2", "Klaster 3",
"Klaster 4", "Klaster 5"),
pch = c(18, 15, 8, 17, 16),
col = c("purple", "pink", "green", "yellow", "red"),
cex = 0.8,
pt.cex = 1,

```



Lampiran 12. (Lanjutan)

```
      bty = "n",
      x.intersp = 0.5,
      y.intersp = 0.8,
      xpd = FALSE,
      adj = c(0,0.25))
#==== Proporsi Keragaman kumulatif Biplot ====
eigen.v.1 <- y$d[1]
eigen.v.2 <- y$d[2]
pk=(eigen.v.1+eigen.v.2)/sum(y$d)
pk
```