

ANALISIS KLASTER UNTUK PENGELOMPOKAN KABUPATEN/KOTA DI PROPINSI SULAWESI SELATAN BERDASARKAN INDIKATOR KESEJAHTERAAN RAKYAT

Wahidah Alwiⁱ, Muh. Hasrulⁱⁱ

ⁱ Prodi Matematika FST, UINAM, Wahidah.alwi@uin-alauddin.ac.id

ⁱⁱ Mahasiswa Program Studi Matematika-FST, UINAM

ABSTRAK, Tujuan utama dalam pembangunan suatu Negara adalah meningkatkan kesejahteraan rakyat. Salah satu penyebab dari permasalahan kesejahteraan rakyat adalah pembangunan yang dilakukan oleh pemerintah tidak terlaksana secara merata atau dengan kata lain pemerintah dalam melaksanakan pembangunan di suatu daerah tidak tepat sasaran. salah satu solusi yang dapat diterapkan adalah pengidentifikasian karakteristik berdasarkan tingkat kesejahteraan rakyat tiap daerah sehingga pemerintah dapat mengambil atau memutuskan kebijakan dan strategi yang baik/tepat sasaran dalam pembangunan. Dalam penelitian ini akan dibahas mengenai pengelompokan Kabupaten/Kota di Propinsi Sulawesi Selatan berdasarkan indikator kesejahteraan rakyat dengan analisis *cluster*. Dimana analisis *cluster* merupakan teknik pengelompokan objek-objek berdasarkan kemiripan karakteristik yang dimiliki. Tujuan dari penelitian ini adalah untuk mengelompokkan Kabupaten/Kota di Propinsi Sulawesi Selatan berdasarkan beberapa indikator kesejahteraan rakyat. Salah satu metode dalam analisis *cluster* untuk mengelompokkan adalah metode *Average Linkage* yaitu metode yang ditentukan dari rata-rata jarak seluruh objek pada *cluster* lainnya. Dari hasil analisis diperoleh bahwa pengelompokan 24 Kabupaten/Kota di Propinsi Sulawesi Selatan dapat dibentuk tiga kelompok (*cluster*), yaitu *Cluster 1* terdiri 21 Kabupaten/Kota dimana *cluster* ini sangat dipengaruhi oleh variabel Kepemilikan Rumah Sendiri (X_{11}). Selain dari variabel PDRB (X_1) dan Kepemilikan Rumah Sendiri (X_{11}) untuk ketujuh variabel lainnya memiliki rata-rata (*centroid*) yang paling rendah diantara *cluster* lainnya., *Cluster 2* terdiri 2 Kabupaten/Kota dimana pada *cluster* ini dominan terhadap variabel Rata-Rata Lama Sekolah. *Cluster 2*, untuk variabel PDRB (X_1) memiliki rata-rata yang paling rendah. Sedangkan untuk kedelapan variabel lainnya memiliki rata-rata (*centroid*) yang cukup tinggi (sedang)., dan *Cluster 3* terdiri 1 Kabupaten/Kota yaitu Kota Makassar dengan variabel yang mempengaruhi yaitu variabel Kepadatan Penduduk (X_2). Pada *Cluster* ini untuk kedelapan variabel lainnya memiliki rata-rata (*centroid*) yang paling tinggi diantara *cluster* lainnya. Sedangkan untuk variabel Kepemilikan Rumah Sendiri (X_{11}) merupakan yang paling rendah.

Kata Kunci: Analisis Cluster, Kesejahteraan Rakyat, Average Linkage

1. PENDAHULUAN

Kesejahteraan rakyat selalu menjadi topik yang menarik untuk dibahas. Sebagaimana di setiap negara yang menjadi tujuan utama dalam pembangunan adalah peningkatan kesejahteraan rakyat. Sepertihalnya di Indonesia, kesejahteraan rakyat adalah salah satu tujuan negara yang tertuang dalam Pembukaan UUD 1945 alinea IV.

Kesejahteraan rakyat pada dasarnya merupakan suatu kondisi yang bentuknya dinamis atau dengan kata lain nilai kuantitatifnya tidak akan pernah berhenti karena akan terus berubah seiring dengan perkembangan kebutuhan hidup manusia.

Jumlah penduduk propinsi Sulawesi Selatan Tahun 2015 sebesar 8.520.304 jiwa dengan kepadatan penduduk sekitar 186,18 jiwa per kilometer persegi. Jumlah penduduk terus mengalami kenaikan dari tahun ke tahun sebesar 1.8% atau 10,1 juta. Ketika jumlah penduduk terus bertambah, berarti pemerintah juga harus terus menambah jumlah fasilitas hidup layak bagi masyarakatnya. Selain masalah tentang kependudukan, angka pengangguran di Propinsi Sulawesi Selatan juga masih tinggi. Sesuai data Tahun 2015, angka pengangguran terbuka mencapai 218.311 jiwa atau sekitar 5,95%. Dibandingkan tahun sebelumnya dengan tingkat pengangguran terbuka berjumlah 212.857 jiwa atau sekitar 5,08%. Implikasinya, tingkat pengangguran terbuka akan semakin meningkat, jika tidak ada perubahan strategi dalam penciptaan lapangan kerja. Maka dari itu salah satu solusi guna mengatasi hal tersebut adalah dengan pengidentifikasian karakteristik berdasarkan tingkat kesejahteraan rakyat tiap daerah sehingga pemerintah dapat mengambil

atau memutuskan kebijakan dan strategi yang baik/tepat sasaran dalam pembangunan

Analisis multivariat merupakan salah satu solusi dalam hal mengelompokkan objek-objek yang mempunyai kesamaan karakteristik. Salah satu dari analisis multivariat yang dapat digunakan yaitu analisis klaster. Analisis klaster bertujuan untuk mengelompokkan objek-objek pengamatan berdasarkan karakteristik yang dimiliki. Ukuran kesamaan yang digunakan adalah ukuran jarak antar objek. Kedua objek yang memiliki jarak paling dekat akan bergabung menjadi satu klaster. Kedekatan jarak yang dimiliki menunjukkan bahwa kedua objek tersebut memiliki tingkat kesamaan karakteristik. Objek penelitian ini adalah tingkat kesejahteraan rakyat dengan 11 indikator yaitu PDRB Tiap Kabupaten/Kota (X_1), Kepadatan Penduduk (X_2), Jumlah Penduduk Miskin (X_3), Daya Beli (X_4), Jumlah Angkatan Kerja (X_5), Angka Harapan Hidup (X_6), Angka Melek Huruf (X_7), Rata-Rata Lama Sekolah (X_8), Angka Harapan Lama Sekolah (X_9), Tingkat Pengangguran Terbuka (X_{10}), Kepemilikan Rumah Sendiri (X_{11}).

2. TINJAUAN PUSTAKA

Analisis Multivariat

Analisis multivariat yaitu analisis yang digunakan terhadap data yang terdiri dari banyak variabel, dimana variabelnya saling berkorelasi satu sama lain, dengan $i = 1, 2, \dots, n$ dan $j = 1, 2, \dots, p$. Data multivariat dengan n pengamatan pada p variabel ditampilkan sebagai berikut

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2k} & \dots & x_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{j1} & x_{j2} & \dots & x_{jk} & \dots & x_{jp} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nk} & \dots & x_{np} \end{bmatrix} \text{ Matriks } \mathbf{X}$$

berisi data yang terdiri dari semua pengamatan terhadap semua variabel dimana n baris dan p kolom.

Analisis Cluster

Analisis *cluster* adalah salah satu teknik dari analisis multivariat yang mempunyai tujuan utama yaitu untuk mengelompokkan objek-objek

yang berdasarkan kemiripan karakteristik yang dimilikinya. Analisis *cluster* akan membagi sejumlah data pada satu atau beberapa *cluster* tertentu. Sebuah *cluster* yang baik adalah *cluster* yang memiliki :

- Homogenitas (kesamaan) yang tinggi antara anggota dalam satu *cluster* (*within-cluster*).
- Heterogenitas (perbedaan) yang tinggi antara *cluster* yang satu dengan *cluster* lainnya (*between cluster*).

Prosedur Analisis Cluster

Rumusan Masalah

Dalam hal ini dilakukan pemilihan variabel-variabel yang akan digunakan untuk pengelompokan (pembentukan *cluster*). Tahap ini merupakan tahap yang penting karena memilih satu atau dua variabel yang tidak sesuai dengan masalah pengelompokan dan akan menyebabkan penyimpangan pada hasil pengelompokan.

Data Outliers

Outliers adalah data pengamatan dengan nilainya “ekstrem” yang muncul karena memiliki karakteristik yang unik atau terlihat sangat jauh berbeda dari observasi-observasi lainnya. Deteksi data yang *outlier* dapat dilakukan dengan melakukan standarisasi data (*Z-Score*) setelah itu melakukan pengecekan dimana dikatakan sebuah data *outlier* apabila nilainya tidak terletak dalam selang $(-2,5; +2,5)$.

Standarisasi Data

Perlunya dilakukan proses standarisasi data dengan transformasi pada data asli sebelum dianalisis lebih lanjut apabila variabel yang diteliti memiliki perbedaan ukuran satuan yang besar. Standarisasi dilakukan terhadap variabel yang relevan ke dalam bentuk *z-score*.

Asumsi-Asumsi

Terdapat dua asumsi yang harus dipenuhi dalam analisis *cluster* yaitu sampel mewakili populasi dan tidak terjadi multikolinearitas. Untuk mengetahui apakah sampel yang diambil benar-benar dapat mewakili populasi yang ada dibutuhkan nilai *Kaiser-Meyer Olkin (KMO)*. Dimana nilai *KMO* kurang dari 0.5 menandakan bahwa sampel yang diambil tidak dapat mewakili populasi yang ada, (Machfudhoh, 2013). Sedangkan untuk mendeteksi ada atau tidaknya multikolinearitas dapat dilakukan dengan

menghitung koefisien korelasi sederhana (korelasi *Pearson*) antar variabel bebas, jika terdapat nilai yang mencapai atau melebihi 0,8 maka terjadi multikolinearitas. Dalam mengatasi masalah multikolinearitas pada data adalah dengan *Metode Principal Component Analysis* (PCA) atau analisis komponen utama

Ukuran Kemiripan Objek

Dalam menjelaskan atau mengukur kemiripan/kesamaan antar objek digunakan pendekatan ukuran jarak (*distance*). Semakin besar nilai jarak antara objek berarti semakin besar pula perbedaan karakteristik antara objek. Sedangkan semakin kecil nilai jarak antara objek berarti semakin tinggi pula kemiripan/kesamaan karakteristik antara objek. Terdapat dua pengamatan dalam ruang p -dimensi $\mathbf{x}' = (x_1, x_2, \dots, x_p)$ dan $\mathbf{y}' = (y_1, y_2, \dots, y_p)$. Jarak dari dua pengamatan tersebut dapat dihitung dalam berbagai cara, seperti jarak *Euclid*. Jika jarak dua titik \mathbf{x} dan \mathbf{y} ditulis dengan $d(\mathbf{x}, \mathbf{y})$, rumus perhitungannya dapat dihitung sebagai berikut :

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2}$$

$$= \sqrt{(\mathbf{x} - \mathbf{y})'(\mathbf{x} - \mathbf{y})}$$

Metode Pengelompokan

Terdapat dua macam metode dalam proses analisis *cluster* yaitu metode hirarki dan metode non hirarki. Adapun metode hierarki adalah metode pengelompokan yang terstruktur dan bertahap serta jumlah kelompok atau *cluster* belum diketahui. Sedangkan pada metode non hirarki, telah ditentukan jumlah kelompok terlebih dahulu.

Pada metode hirarki terdapat dua cara dalam analisis cluster yaitu dengan cara penggabungan (*agglomerative*) dan pemisahan (*divisive*). Metode hirarki dengan cara penggabungan dilakukan dengan menggabungkan objek secara bertahap, sehingga nantinya hanya diperoleh satu kelompok saja. Sebaliknya cara pemisahan pada metode hirarki dimulai dengan membentuk satu kelompok/*cluster* besar yang anggotanya seluruh objek pengamatan. Kemudian dipisah menjadi kelompok yang lebih kecil, hingga akhirnya satu

kelompok hanya beranggotakan satu objek pengamatan saja.

Dalam penelitian ini menggunakan *Average Linkage* dalam pengelompokannya yaitu menentukan rata-rata jarak seluruh objek suatu *cluster* terhadap seluruh objek pada *cluster* lainnya. Adapun langkah-langkahnya sebagai berikut :

- i.) Menentukan objek bersesuaian yang memiliki jarak terdekat dalam matriks jarak $\mathbf{D} = \{d_{ik}\}$,
- ii.) Menggabungkan objek yang bersesuaian tersebut, katakanlah objek U dan objek V , yang kemudian didapatkan *cluster*(UV).
- iii.) Mengitung jarak antara *cluster*(UV) dengan objek lain katakanlah W yang belum bergabung, dengan rumus

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

dimana :

d_{ik} : jarak antara obyek i pada *cluster* UV dan obyek k pada *cluster*

$N_{(UV)}$: Jumlah item pada *cluster*

N_W : Jumlah item pada *cluster* UV dan W

dan menghitung kembali matriks jarak baru dengan cara pada langkah ketiga algoritama *cluster* hirarki *agglomerative* secara umum, dan beri nama \mathbf{D}_2 .

- iv.) Mengulangi Langkah 2 sampai bergabung menjadi satu *cluster*.

Interpretasi Cluster

Tahapan interpretasi *cluster* adalah untuk mencari karakter setiap kelompok yang khas, salah satunya dapat dilakukan dengan membandingkan *mean* atau *centroid* pada setiap kelompok. Adapun cara menghitung *mean* (*centroid*) yaitu :

$$C = \frac{\sum_{j=1}^n \sum_{k=1}^p x_{jk}}{N}$$

dimana :

C : Nilai rata-rata (*centroid*) cluster

x_{jk} : Nilai atau data dari objek j pada variabel k dalam *cluster* yang dicari nilai *centroid*

N :Jumlah objek dalam *cluster* yang dicari nilai *centroid*

3. METODOLOGI PENELITIAN

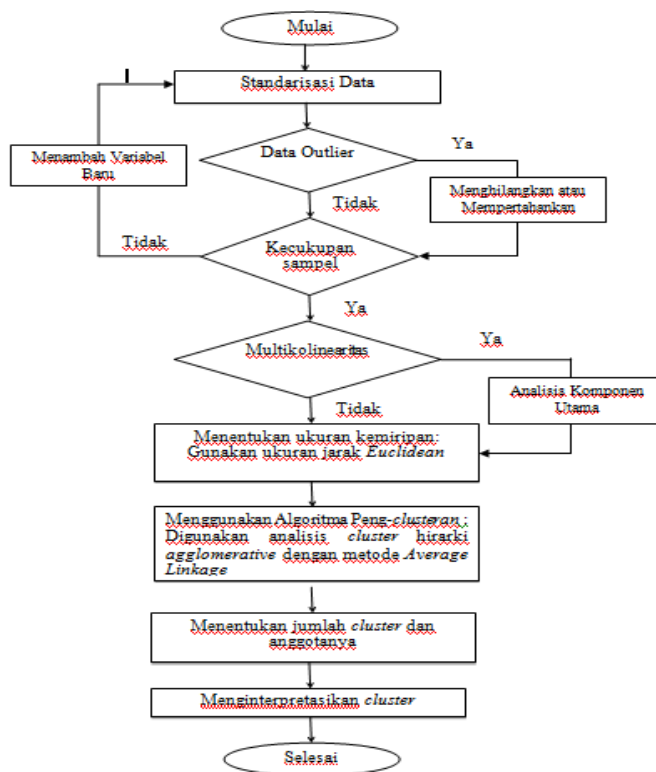
Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari BPS Provinsi Sulawesi Selatan yaitu data indikator kesejahteraan rakyat Tahun 2015. Dengan teknik pengumpulan data yaitu (1). Dokumentasi, (2) Observasi, dan (3) wawancara.

Adapun variabel yang digunakan yaitu PDRB(X_1) yaitu jumlah nilai barang dan jasa akhir dalam satuan Rupiah yang dihasilkan oleh seluruh unit ekonomi pada suatu daerah, kepadatan penduduk (X_2) yaitu perbandingan jumlah penduduk dengan luas wilayahnya dalam satuan Jiwa/ Km^2 , jumlah penduduk miskin (X_3) yaitu jumlah penduduk yang memiliki rata-rata pengeluaran dibawah garis kemiskinan dalam satuan Jiwa, daya beli (X_4) yaitu kemampuan masyarakat dalam membelanjakan uangnya dalam bentuk barang maupun jasa dengan satuan Rupiah, jumlah angkatan kerja (X_5) yaitu jumlah penduduk usia kerja (15 tahun dan lebih) yang bekerja, atau punya pekerjaan namun sementara tidak bekerja dan pengangguran dalam satuan Jiwa, angka harapan hidup (X_6) yaitu rata-rata tahun hidup dalam satuan Tahun yang masih akan dijalani oleh seseorang yang telah berhasil mencapai umur x , pada suatu tahun tertentu, dalam situasi mortalitas yang berlaku di lingkungan masyarakatnya, angka melek huruf (X_7) yaitu persentase penduduk usia 10 tahun ke atas yang dapat membaca dan menulis dalam huruf latin atau lainnya, dan rata-rata lama sekolah (X_8) yaitu jumlah tahun belajar penduduk usia 15 tahun ke atas yang telah diselesaikan dalam pendidikan formal (tidak termasuk tahun yang mengulang) dengan satuan Tahun, angka harapan lama sekolah (X_9) yaitu lamanya sekolah yang diharapkan akan dirasakan oleh anak pada umur tertentu di masa mendatang dalam satuan Tahun, tingkat pengangguran terbuka (X_{10}) yaitu Persentase pengangguran terhadap jumlah angkatan kerja, dan kepemilikan rumah sendiri (X_{11}) yaitu Persentase jumlah rumah tangga dengan status kepemilikan rumah sendiri. Dimana rumah sendiri merupakan jika tempat tinggal tersebut pada waktu pendataan

betul-betul sudah milik kepala rumah tangga atau salah satu seorang anggota rumah tangga.

Dalam penelitian ini digunakan analisis *cluster* hirarki *agglomerative* dengan metode *Average Linkage*, dengan langkah-langkah sebagai berikut: (1) Menentukan objek bersesuaian yang memiliki jarak terdekat dalam matriks jarak $D = \{d_{ik}\}$, (2) Menggabungkan objek yang bersesuaian tersebut, katakanlah objek U dan objek V , yang kemudian didapatkan *cluster* (UV), (3) Menghitung jarak antara *cluster* (UV) dengan objek lain katakanlah W yang belum bergabung dan menghitung kembali matriks jarak baru, (4) Mengulangi langkah 2 sampai bergabung menjadi satu *cluster*. Hasil dari *Average Linkage* ini dapat ditampilkan dalam bentuk dendogram atau diagram pohon

Adapun prosedur dari analisis cluster dalam penelitian ini disajikan dalam diagram alur sebagai berikut:



4. HASIL PENELITIAN

Statistika Deskriptif

Sebelum dilakukan analisis *cluster* terlebih dahulu dianalisis secara deskriptif.

Tabel 4.1 Statistika Deskriptif Indikator Kesejahteraan Rakyat Tahun 2015

Statistik Deskriptif				
Var	Min	Mak	Mean	Std Deviasi
(X ₁)	2724	88740	10393	17105,97
(X ₂)	40	8246	645,4	1648,47
(X ₃)	8,41	75,01	33,24	17,74
(X ₄)	6,273	16,669	9,907	2,007
(X ₅)	13542	348501	121311	77104,28
(X ₆)	65,49	72,8	68,6	2,1
(X ₇)	83,5	98,22	91,82	3,33
(X ₈)	5,64	10,77	7,54	1,23
(X ₉)	11,61	15,02	12,82	0,91
(X ₁₀)	0,9	12,07	5,23	2,96
(X ₁₁)	68,22	98,84	88,22	8,59

Berdasarkan Tabel 4.1 diperoleh nilai minimum, maksimum, mean dan standar variasi dari sebelas variabel. Nilai tersebut menggambarkan perbedaan karakteristik dari setiap variabel. Sebelum di lakukan uji asumsi terlebih dahulu dilakukan standarisasi data yang dilakukan apabila terdapat perbedaan satuan yang signifikan diantara variabel-variabel yang diteliti. Karena dalam penelitian ini, data yang digunakan satuan variabelnya berbeda maka dilakukan standarisasi data terlebih dahulu. Selanjutnya, berdasarkan data penelitian yang telah distandarisi, jika ada data yang nilainya tidak berada diantara ± 2,5 berarti data tersebut outlier. Hasil pengamatan terhadap data yang outlier ditampilkan dalam Tabel 4.15 sebagai berikut :

Tabel 4.2. Deteksi *Outlier*

Kab/Kota	Variabel	Nilai
Bone	ZX ₅	2,94652
Makassar	ZX ₁	4,5801
	Zx ₂	4,61067
	ZX ₄	2,87021
	ZX ₈	2,61633

Berdasarkan Tabel 4.2 diperoleh data outlier, langkah berikutnya adalah tetap mempertahankan data outlier tersebut karena merupakan representasi dari populasi yang diteliti. Data yang outlier ini memberikan gambaran bahwa Kabupaten Bone dan Kota Makassar pada variabel tertentu sangat berbeda karakteristiknya dibanding dengan kabupaten/kota lainnya yaitu: jumlah angkatan kerja (ZX₅) di Kab. Bone sangat berbeda dengan kabupaten yang lain, begitupula PDRB (ZX₁) Kota Makassar sangat berbeda dengan PDRB di

kabupaten yang lainnya, kepadatan penduduk (ZX₂) , daya beli (ZX₄) dan rata-rata lama sekolah (ZX₈) sangat berbeda dengan kabupaten yang lainnya.

Uji Asumsi

Uji Kecukupan Sampel

Uji kecukupan sampel bertujuan untuk menguji apakah sampel yang digunakan cukup untuk dianalisis lebih lanjut, dapat dilihat dari nilai *Kaiser Meyer Olkin (KMO)*.

Hipotesis :

H₀: Sampel belum cukup untuk dianalisis lebih lanjut

H₁: Sampel sudah cukup untuk dianalisis lebih lanjut

$$\text{Statistik uji: } KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p \rho_{ij}^2}$$

Pengambilan keputusan : nilai *KMO* > 0,5 maka tolak H₀.

Tabel 4.3 Uji KMO

Uji		Nilai
<i>KMO</i>		0,694
<i>Bartlett's Test Of Sphericity</i>	<i>Approx. Chi-Square</i>	237,71
	<i>Df</i>	55
	<i>Sig</i>	2.2e-16

Berdasarkan Tabel 4.3 menunjukkan bahwa nilai *KMO* sebesar 0,694. Karena 0,694 > 0,5 berarti tolak H₀ sehingga dapat disimpulkan sampel digunakan sudah cukup untuk dianalisis lebih lanjut.

Uji Multikolinearitas

Dalam mengetahui ada tidaknya multikolinearitas dapat dilihat dari nilai koefisien korelasi sederhana (*korelasi Pearson*). Jika terdapat nilai yang mencapai atau melebihi ± 0,8 maka terjadi multikolinearitas. Setelah dilakukan perhitungan menggunakan program R diperoleh hasil bahwa variabel X₁ dengan X₂, X₇ dengan X₈, X₈ dengan X₉, X₈ dengan X₁₁ serta X₉ dengan X₁₁ mengalami multikolinearitas karena nilai korelasi lebih dari ±0,80. Maka dari itu dilakukan analisis komponen utama atau *principal component analysis* untuk mengatasi masalah tersebut.

IV.3 Principal Component Analysis

Dalam analisis komponen utama ada beberapa asumsi yang harus terpenuhi yaitu :

a. Uji Kecukupan Sampel

Uji ini telah dilakukan sebelumnya dan diperoleh nilai *KMO* sebesar 0,694 (Tabel 4.2) dan telah memenuhi asumsi.

b. Uji Bartlett

Uji ini bertujuan untuk mengetahui ada tidaknya korelasi antar variabel

Hipotesis :

$H_0 : \mathbf{R} = \mathbf{I}$ (Tidak ada korelasi)

$H_1 : \mathbf{R} \neq \mathbf{I}$ (Terdapat korelasi)

Statistik uji :

$$\chi^2_{hitung} = -\left\{n-1 - \frac{2p+5}{6}\right\} \ln|\mathbf{R}|$$

Pengambilan keputusan : Tolak H_0 jika $\chi^2_{hitung} > \chi^2_{\frac{\alpha}{2}, p(p-1)}$ atau $\text{sig.} < \alpha=0,05$

Adapun hasil pengujian *Bartlett* dapat di lihat pada Tabel 4.3 sebagai berikut :

Tabel 4.4 Uji Bartlett

Uji		Nilai
Bartlett's Test	Approx. Chi-Square	237,71
	df	55
	sig.	2.2e-16

Berdasarkan Tabel 4.4 menunjukkan bahwa nilai *p-value* sebesar 2.2e-16. Karena *p-value* < 0,05 berarti tolak H_0 sehingga dapat disimpulkan terdapat korelasi antar variabel.

c. Nilai MSA

Tabel 4.5 Nilai *MSA*

Var	<i>MSA</i>
X1	0,609
X2	0,645
X3	0,341
X4	0,768
X5	0,424
X6	0,568
X7	0,682
X8	0,783
X9	0,802
X10	0,813
X11	0,78

Nilai *MSA* digunakan untuk mengetahui variabel-variabel mana saja yang layak dilakukan *PCA*. Nilai *MSA* yang kurang dari 0,5 menunjukkan bahwa variabel tersebut tidak layak untuk dianalisis *PCA*. Adapun hasil

perhitungan nilai *MSA* dari kesebelas variabel yang ditampilkan dalam Tabel 4.5.

Berdasarkan Tabel 4.5 dari sebelas variabel yang ada, variabel X_3 dan X_5 masing-masing memiliki nilai *MSA* kurang dari 0,5 sehingga hanya kesembilan variabel yang layak untuk dilakukan *PCA*.

Selanjutnya melakukan analisis komponen utama dengan langkah-langkah sebagai berikut :

1. Membuat matrks \mathbf{Z} yang berisi data variabel yang distandarisasi
2. Membuat matriks korelasi \mathbf{R} yaitu dengan rumus :

$$\mathbf{R} = \frac{1}{n-1} \mathbf{Z}'\mathbf{Z}$$

3. Mencari nilai eigen $\lambda_1, \lambda_2, \dots, \lambda_p$ yang memenuhi persamaan

$$|\mathbf{R} - \lambda \mathbf{I}| = 0$$

4. Mencari eigen vector (\hat{e}_i) yang bersesuaian dengan nilai eigen dan memenuhi persamaan

$$(\mathbf{R} - \lambda \mathbf{I})\hat{e}_i = 0$$

5. Menentukan banyaknya komponen yang dipilih berdasarkan nilai eigen yang menjelaskan besarnya kontrobusi keragaman atau varian masing-masing komponen utama. Nilai eigen selalu diurutkan dari yang terbesar ke terkecil.

6. Bentuk persamaan komponen (*PC*) seperti berikut :

$$PC_i = \hat{e}_i \mathbf{z} = \hat{e}_{i1} z_1 + \hat{e}_{i2} z_2 + \dots + \hat{e}_{ip} z_p$$

Setelah dilakukan langkah-langkah seperti di atas diperoleh bahwa Komponen 1 dan Komponen 2 secara bersama-sama telah dapat menjelaskan keragaman atau variansi dari kesembilan variabel sebesar 79,89 %.Selanjutnya menentukan *eigen vector* atau koefisien dari *principal component* yang akan digunakan untuk membentuk persamaan *PC* sebagai berikut :

$$PC_1 = 0,287z_1 + 0,313z_2 + 0,313z_4 + 0,246z_6 + 0,345z_7 + 0,394z_8 + 0,372z_9 + 0,329z_{10} - 0,376z_{11}$$

$$PC_2 = -0,510z_1 - 0,450z_2 - 0,341z_4 + 0,467z_6 + 0,224z_7 + 0,171z_8 + 0,217z_9 - 0,121z_{10} - 0,249z_{11}$$

Selanjutnya adalah menghitung *PCscore* dari kedua persaman *PC*. Nilai *PC score* inilah yang digunakan untuk analisis lebih lanjut.

Ukuran Kemiripin Objek

Menghitung kemiripan antar objek dengan menggunakan jarak *Euclidean*. Dihitung jarak antara Kab. Kepulauan Selayar dan Kab. Bulukumba (obyek 1 dan 2).

$$d(1,2) = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2}$$

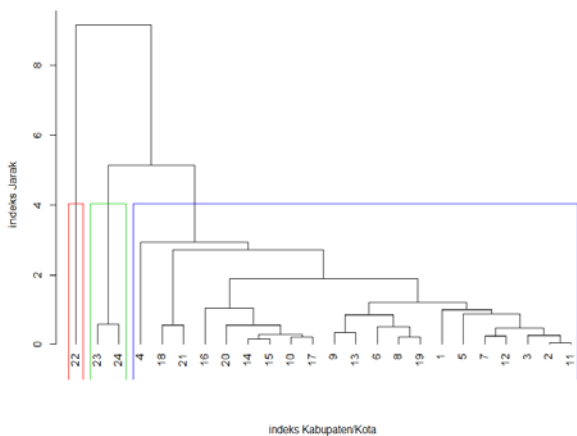
$$= \sqrt{((-1,63688) - (-1,52448))^2 + (0,41790 - (-0,64439))^2}$$

$$= 1,068$$

selanjutnya adalah menghitung jarak antara seluruh objek sehingga nantinya diperoleh matriks jarak antara objek.

Analisis Cluster Hirarki Metode Average Linkage

Dengan menggunakan metode *Average Linkage* bahwa objek nomor 2 (Kab. Bulukumba) dan nomor 11 (Kab. Bone) bergabung menjadi satu *cluster*, karena memiliki jarak paling kecil yaitu 0,044. Hal ini menunjukkan bahwa jarak antara kedua kabupaten tersebut merupakan jarak yang paling dekat dari banyaknya kombinasi jarak 24 kabupaten/kota. Selanjutnya pada tahap kedua jarak yang terdekat adalah objek nomor 14 (Sindereng Rappang) bergabung dengan nomor 15 (Pinrang) dengan jarak yaitu 0,147. Demikian seterusnya, sehingga semua objek bergabung menjadi satu cluster. Hasil dari proses ini ditampilkan dalam Gambar 4.1 sebagai berikut :



Gambar 4.1 Dendrogram Analisis Cluster Metode Average Linkage

Pada Gambar 4.1 menunjukkan bahwa terdapat 3 kelompok/*cluster* yaitu *cluster* 1 terdiri 21 Kabupaten/Kota, *cluster* 2 terdiri dua Kabupaten/Kota, dan *cluster* 3 terdiri satu

Kabupaten/Kota. Dalam hal ini dapat dilihat dalam Tabel 4.6 sebagai berikut:

Tabel 4.6 Cluster dan Anggotanya

Cluster	Kabupaten/Kota
Cluster 1	Kepulauan Selayar, Gowa, Barru, Sidenreng Rappang, Pinrang, Luwu, Luwu Timur, Bulukumba, Bantaeng, Sinjai, Maros, Pangkajenne Kepulauan, Bone, Soppeng, Wajo, Luwu Utara, Takalar, Tana Toraja, Toraja Utara, Enrekang, dan Jeneponto
Cluster 2	Palopo, dan Pare-Pare
Cluster 3	Makassar

Interpretasi Cluster

Dalam hal menginterpretasikan hasil *cluster* digunakan nilai-nilai rata-rata (*centroid*) tiap *cluster*. Nilai *centroid* tiap *cluster* diperoleh sebagaimana dalam Tabel 4.6 berikut ini :

Tabel 4.6 Centroid Cluster

Variabel	Cluster 1	Cluster 2	Cluster 3
X ₁	-0,18246	-0,37419	4,58011
X ₂	-0,2423	0,23878	4,61066
X ₄	-0,25548	1,24739	2,87021
X ₆	-0,14603	0,8518	1,36312
X ₇	-0,25127	1,69986	1,87695
X ₈	-0,32433	2,09727	2,61634
X ₉	-0,29993	2,08859	2,12137
X ₁₀	-0,27147	1,70375	2,29325
X ₁₁	0,29938	-1,97968	-2,32766

5. KESIMPULAN DAN SARAN

Berdasarkan analisis dan pembahasan, maka dapat disimpulkan hasil pengelompokan Kabupaten/Kota di provinsi Sulawesi Selatan berdasarkan indikator kesejahteraan rakyat dengan menggunakan metode *Average Linkage* yaitu :

1. *Cluster* 1 terdiri 21 kabupaten/kota, dengan indikator oleh variabel Kepemilikan Rumah Sendiri (X₁₁). Selain dari variabel PDRB (X₁) dan Kepemilikan Rumah Sendiri (X₁₁) untuk ketujuh variabel lainnya memiliki rata-rata nilai centroid yang paling rendah di antara *cluster* lainnya.
2. *Cluster* 2 terdiri dua kabupaten/kota. Dengan indikator terhadap variabel Rata-Rata Lama Sekolah. *Cluster* 2, untuk variabel PDRB (X₁) memiliki rata-rata yang paling rendah.

Sedangkan untuk kedelapan variabel lainnya memiliki rata-rata nilai *centroid* yang cukup tinggi (sedang).

3. *Cluster* 3 terdiri satu kabupaten/kota yaitu, Kota Makassar dengan variabel yang mempengaruhi yaitu variabel Kepadatan Penduduk (X_2). Pada *Cluster* ini untuk kedelapan variabel lainnya memiliki rata-rata nilai *centroid* yang paling tinggi diantara cluster lainnya. Sedangkan untuk variabel Kepemilikan Rumah Sendiri (X_{11}) memiliki rata-rata nilai *centroid* paling rendah.

Untuk penelitian selanjutnya, peneliti menyarankan untuk penggunaan metode yang lain dalam analisis cluster hirarki sehingga dapat menjadi bahan perbandingan hasil analisis.

6. DAFTAR PUSTAKA

- [1] Badan Pusat Statistik Provinsi Sulawesi Selatan. 2015. "*Indikator Kesejahteraan Rakyat Sulawesi Selatan Tahun 2015*". Makassar : BPS Provinsi Sulawesi Selatan.
- [2] Ermawati. 2012. "*Statistika Multivariat Terapan*". Makassar : Alauddin University Press.
- [3] Johnson, Richard & Dean. 2007. "*Applied Multivariate Statistical Analysis Sixth Edition*". New York: Prentice-Hall International. Mattjik, Ahmad Ansori & I Made Sumertajaya, "*Sidik Peubah Ganda dengan Menggunakan SAS*". Bogor : IPB Press.
- [4] Laraswati, Tri Febriana. 2014. *Perbandingan Kinerja Metode Complete, Linkage Metode Average Linkage, dan Metode K-means Dalam Menentukan Hasil Analisis Cluster*. Yogyakarta : Universitas Negeri Yogyakarta.
- [5] Machfudhoh, Siti dan Nuri Wahyuningsih. "Analisis Cluster Kabupaten/Kota Berdasarkan Pertumbuhan Ekonomi Jawa Timur". *Jurnal Sains Dan Seni Pomits*, Vol. 2, No.1 (2013).http://digilib.its.ac.id/public/ITS-paper-3759_7-1_210100028-paper.pdf.