

Pengelompokkan Kabupaten/Kota Di Sulawesi Berdasarkan Indikator Pendidikan Menggunakan Analisis Klaster Average Linkage Dan Median Linkage

(District/City Clustering In Sulawesi Based On Education Indicators Using Average Linkage Cluster And Median Linkage Analysis)

Eka Oktavianty^{1*}, Junaidi¹ dan Lilies Handayani¹

¹ Program Studi Statistika, Jurusan Matematika, Fakultas MIPA, Universitas Tadulako, Jl. Soekarno Hatta Km 9 Tondo Palu Sulawesi Tengah 94118.

Keywords: *Educational Indicators, Cluster Analysis, Average Linkage, Median Linkage, Standard Deviation Ratio.*

Keywords: *Indikator Pendidikan, Analisis Klaster, Average Linkage, Median Linkage, Rasio Simpangan Baku.*

* Corresponding Author :
ekaoktavianty97@gmail.com

Abstract

Education is important to improve the quality of human resources, but the important issue that often arises is the problem of the low quality of education and equitable distribution of education. Measuring instruments used to see how well the quality of education is an indicator of education. To find out the ranking of education or the characteristics of the level of education, the grouping is done based on educational indicators using cluster analysis. Cluster analysis is a multivariate technique that classifies objects into different groups between one group and another group. In this study, the methods will be used average linkage and median linkage in classifying Regencies /Cities in Sulawesi based on education indicators. The results of the analysis showed that the median linkage obtained a standard deviation ratio value of 0.061 smaller than the standard deviation ratio average linkage value of 0.078. The method that has the smallest ratio is the method with the best performance. So that grouping City Districts in Sulawesi based on education indicators in 2017 is better to use the median linkage.

Abstrak

Pendidikan merupakan hal yang penting untuk meningkatkan kualitas sumber daya manusia, akan tetapi permasalahan yang sering muncul adalah isu tentang rendahnya mutu pendidikan dan pemerataan pendidikan. Alat ukur yang digunakan untuk melihat seberapa baik kualitas pendidikan adalah indikator pendidikan. Untuk mengetahui pemerataan pendidikan atau karakteristik tingkat pendidikan maka dilakukan pengelompokkan berdasarkan indikator pendidikan menggunakan analisis klaster. Analisis klaster adalah salah satu teknik multivariat yang mengklasifikasikan suatu objek-objek ke dalam kelompok yang berbeda antara kelompok satu dengan kelompok lainnya. Pada penelitian ini akan digunakan metode average linkage dan median linkage dalam mengelompokkan Kabupaten/Kota di Sulawesi berdasarkan indikator pendidikan. Hasil dari analisis menunjukkan bahwa median linkage memperoleh nilai rasio simpangan baku sebesar 0,061 lebih kecil dibandingkan dengan nilai rasio simpangan baku average linkage yakni 0,078. Metode yang mempunyai rasio terkecil merupakan metode dengan kinerja terbaik. Sehingga pengelompokkan Kabupaten/Kota di Sulawesi berdasarkan indikator pendidikan tahun 2017 lebih baik menggunakan median linkage.

Latar Belakang

Menurut Supranto (2004) analisis multivariat merupakan analisis yang digunakan untuk memahami struktur data yang melibatkan lebih dari satu variabel atau teknik analisis yang melibatkan banyak variabel. Multivariat terbagi dua, yaitu analisis dependensi dan analisis interdependensi. Analisis kluster termasuk dalam analisis multivariat interdependensi atau saling ketergantungan (Gudono, 2015).

Analisis kluster adalah salah satu teknik multivariat yang mengklasifikasi suatu objek-objek ke dalam kelompok yang berbeda antara kelompok satu dengan kelompok lainnya (Goreti dkk, 2016). Analisis kluster dibagi dua yaitu metode hierarki dan nonhierarki. Kluster hierarki memulai pengelompokan dengan dua atau lebih objek yang mempunyai kesamaan paling dekat dan kemudian diteruskan ke objek yang mempunyai kedekatan kedua, demikian seterusnya hingga kluster membentuk semacam pohon dimana ada tingkatan yang jelas antar objek. Alat yang membantu memperjelas proses hierarki disebut dendrogram, dendrogram merupakan representasi visual dari langkah analisis kluster yang menunjukkan bagaimana kluster terbentuk dan nilai koefisien jarak pada setiap langkah (Sitepu dkk, 2011).

Metode kluster hierarki diantaranya yaitu single linkage, complete linkage, average linkage, centroid linkage, ward linkage, dan median linkage (Goreti dkk, 2016). Keuntungan penggunaan metode hierarki dalam analisis kluster adalah mempercepat pengolahan dan menghemat waktu karena data yang diinputkan akan membentuk hierarki atau membentuk tingkatan sendiri sehingga mempermudah dalam penafsiran, namun kelemahan dalam metode ini adalah seringnya terdapat kesalahan pada data outlier, perbedaan ukuran jarak yang digunakan, dan terdapatnya variabel yang tidak relevan (Hardius dan Nurdin 2013).

Penelitian terkait dengan kluster hierarki dilakukan oleh (Fadliana dan Rozi, 2015) menggunakan agglomerative hierarchical clustering pada data kualitas pelayanan keluarga berencana. Pada penelitiannya, berdasarkan hasil uji validitas kluster metode average linkage memberikan solusi kluster yang lebih baik dibandingkan dengan metode single linkage, complete linkage dan ward. Penelitian selanjutnya oleh (Ningsih dkk, 2016) menggunakan metode complete linkage dan average linkage pada data produksi palawija menghasilkan metode average linkage merupakan metode terbaik berdasarkan hasil perhitungan rasio simpangan baku.

Pendidikan merupakan hal yang penting untuk meningkatkan kualitas sumber daya manusia, pentingnya pendidikan bagi manusia diantaranya yaitu mampu memberikan ilmu pengetahuan yang luas bagi manusia dan pendidikan merupakan salah satu pilar utama dalam proses peningkatan kemampuan dan daya saing suatu bangsa, sehingga pembangunan pendidikan sangat diperlukan. Pembangunan pendidikan merupakan investasi dalam meningkatkan kualitas sumber daya manusia yang berperan penting dalam meningkatkan pertumbuhan ekonomi dan menurunkan tingkat kemiskinan serta pengangguran (BPS Indonesia, 2018).

Pentingnya peran pendidikan terhadap kemajuan bangsa, pengukuran dan penghitungan indikator-indikator pendidikan perlu dilakukan untuk melihat sejauh mana pemerataan pendidikan. Agar dapat mengetahui pemerataan pendidikan atau karakteristik tingkat pendidikan maka dilakukan pengelompokan menggunakan analisis kluster. Berdasarkan uraian di atas, pada penelitian ini akan dilakukan pengelompokan Kabupaten/Kota di Sulawesi berdasarkan indikator pendidikan menggunakan metode kluster *average linkage* dan *median linkage*.

Bahan dan Metode

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) provinsi-provinsi di Sulawesi yaitu data indikator pendidikan Kabupaten/Kota di Sulawesi tahun 2017. Jumlah Provinsi yang digunakan sebanyak 6 Provinsi dan 81 Kabupaten/Kota dengan 11 variabel. Berikut ini merupakan variabel-variabel yang digunakan dalam penelitian, yaitu:

1. Angka Partisipasi Sekolah Umur 7-12
2. Angka Partisipasi Sekolah Umur 13-15
3. Angka Partisipasi Sekolah Umur 16-18
4. Angka Partisipasi Kasar Umur 7-12
5. Angka Partisipasi Kasar Umur 13-15
6. Angka Partisipasi Kasar Umur 16-18
7. Angka Partisipasi Murni Umur 7-12
8. Angka Partisipasi Murni Umur 13-15
9. Angka Partisipasi Murni Umur 16-18
10. Belum/Tidak Pernah Sekolah
11. Rata-rata Lama Sekolah

Metode yang digunakan pada penelitian ini adalah analisis kluster metode *average linkage* dan *median linkage*. Langkah-langkah analisis yang akan digunakan pada penelitian ini adalah sebagai berikut:

1. Mengumpulkan data, yang diambil dari web resmi BPS Provinsi di Sulawesi pada tahun 2017
2. Melakukan uji kelayakan data, yaitu sampel harus mewakili populasi dengan dengan melihat nilai KMO atau MSA dan melihat hubungan antar variabel dengan uji Bartlett, untuk mengetahui apakah data terdapat multikolinearitas dengan melihat nilai VIF.
3. Menghitung nilai kemiripan antar objek menggunakan rumus *euclidean*.
4. Memilih prosedur analisis klaster.
5. Menginterpretasi klaster.
6. Menentukan metode terbaik.
7. Menarik kesimpulan

Hasil dan Pembahasan

Tujuan dalam penelitian ini untuk memperoleh pengelompokan dan membandingkan hasil analisis menggunakan klaster *average linkage* dan *median linkage* Kabupaten/Kota di Sulawesi berdasarkan indikator pendidikan dan penelitian ini menggunakan *software R*. Berikut merupakan proses

pengelompokan dengan menggunakan *average linkage* dan *median linkage* Kabupaten/Kota di Sulawesi Tengah berdasarkan indikator pendidikan.

Uji Multikolinearitas

Berdasarkan Tabel 1 terlihat bahwa pada variabel X₄ sebesar 12.816 dan X₇ sebesar 15.658, variabel yang memiliki nilai VIF lebih dari 10 artinya terjadi multikolinearitas, sehingga variabel yang memiliki nilai VIF paling tinggi dihilangkan. Variabel X₇ dihilangkan maka dilakukan kembali pengujian multikolinearitas. Berdasarkan Tabel 2 terlihat nilai VIF dari semua variabel kurang dari 10, sehingga dapat disimpulkan tidak terjadi multikolinearitas pada 10 variabel.

Uji Kaiser Meyer Olkin

Berdasarkan Tabel 3 menunjukkan nilai KMO sebesar 0.622 > 0.5, sehingga dapat disimpulkan sampel yang digunakan sudah cukup untuk dianalisis lebih lanjut. Tabel 4 menunjukkan nilai MSA dari semua variabel melebihi 0.5, artinya setiap variabel layak digunakan.

Tabel 1. Nilai VIF 11 Variabel

Variabel	VIF
X1	3.157
X2	2.410
X3	5.498
X4	12.816
X5	5.366
X6	2.125
X7	15.658
X8	9.278
X9	4.464
X10	1.774
X11	1.981

Tabel 2. Nilai VIF 10 Variabel

Variabel	VIF
X1	3.157
X2	2.375
X3	5.472
X4	2.426
X5	5.223
X6	2.075
X8	7.919
X9	4.458
X10	1.675
X11	1.974

Tabel 3. Nilai KMO

Kaiser Meyer Olkin	0.622
--------------------	-------

Tabel 4. Nilai MSA

Variabel	MSA
X1	0.507
X2	0.735
X3	0.567
X4	0.723
X5	0.699
X6	0.550
X8	0.577
X9	0.643
X10	0.545
X11	0.635

Tabel 5. Nilai Bartlett

Nilai Bartlett	484.7
Df	45

Uji Bartlett

Uji Bartlett menyatakan hipotesis sebagai berikut (Suliyanto, 2005):

$H_0 : \rho = I$ (ada korelasi antar variabel)

$H_1 : \rho \neq I$ (tidak ada korelasi antar variabel)

Keputusan: Tolak H_0 bila $X_{hitung}^2 > X_{\alpha;1/2p(p-1)}^2$

Berdasarkan Tabel 5 terlihat bahwa nilai Bartlett sebesar 484.7, Derajat bebas sebesar 45, dan chi-square sebesar 61.65, sehingga disimpulkan bahwa tidak terjadi korelasi antar variabel karena $484.7 > 61.65$.

Menghitung Nilai Kemiripan

Kemiripan antar objek dapat diukur dengan menggunakan ukuran jarak. Menentukan nilai kemiripan antar objek menggunakan jarak *euclidean* karena data indikator pendidikan Kabupaten/Kota di Sulawesi memiliki satuan data yang sama. Perhitungan jarak *euclidean* menggunakan rumus sebagai berikut:

$$d_{ij} = \sqrt{\sum_{k=1}^p (X_{ik} - X_{jk})^2}$$

Sehingga diperoleh nilai kemiripan antar objek berukuran 81x81 sebagai berikut:

$$D = \begin{bmatrix} 0.000 & 14.64 & 17.40 & \dots & 19.98 \\ 14.64 & 0.000 & 11.13 & \dots & 20.11 \\ 17.40 & 11.13 & 0.000 & \dots & 20.22 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 27.15 & 17.93 & 16.20 & \dots & 31.98 \\ 20.72 & 16.35 & 14.67 & \dots & 29.58 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 19.60 & 20.42 & 14.00 & \dots & 18.09 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 80.16 & 80.06 & 79.35 & \dots & 82.75 \\ 71.61 & 71.39 & 71.00 & \dots & 74.29 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 19.98 & 20.11 & 20.22 & \dots & 0.000 \end{bmatrix}$$

Tabel 6. Rata-rata Variabel Metode Average Linkage

Klaster	Rata-rata Variabel
1	69.124
2	76.309
3	64.46
4	74.45
5	29.884

Tabel 7. Rata-rata Variabel Metode Median Linkage

Klaster	Rata-rata Variabel
1	70.18
2	76.309
3	64.46
4	75.168
5	29.884

Tabel 8. Rasio Metode Average Linkage dan Median Linkage

Rasio	Average Linkage	Median Linkage
$\frac{S_w}{S_B}$	0.078	0.061

Proses Analisis Klaster

Proses pengelompokan dilakukan setelah jarak antar variabel diukur dengan menggunakan jarak *euclidean*. Hasil analisis mengindikasikan adanya kedekatan antar objek yang bisa dijadikan panduan objek mana yang memiliki kemiripan karakteristik yang sama. Dua objek dengan karakteristik yang sama digambarkan sebagai dua titik yang posisinya berdekatan, semakin dekat posisi dua buah objek maka semakin mirip dan semakin jauh posisi dua buah titik objek maka semakin berbeda.

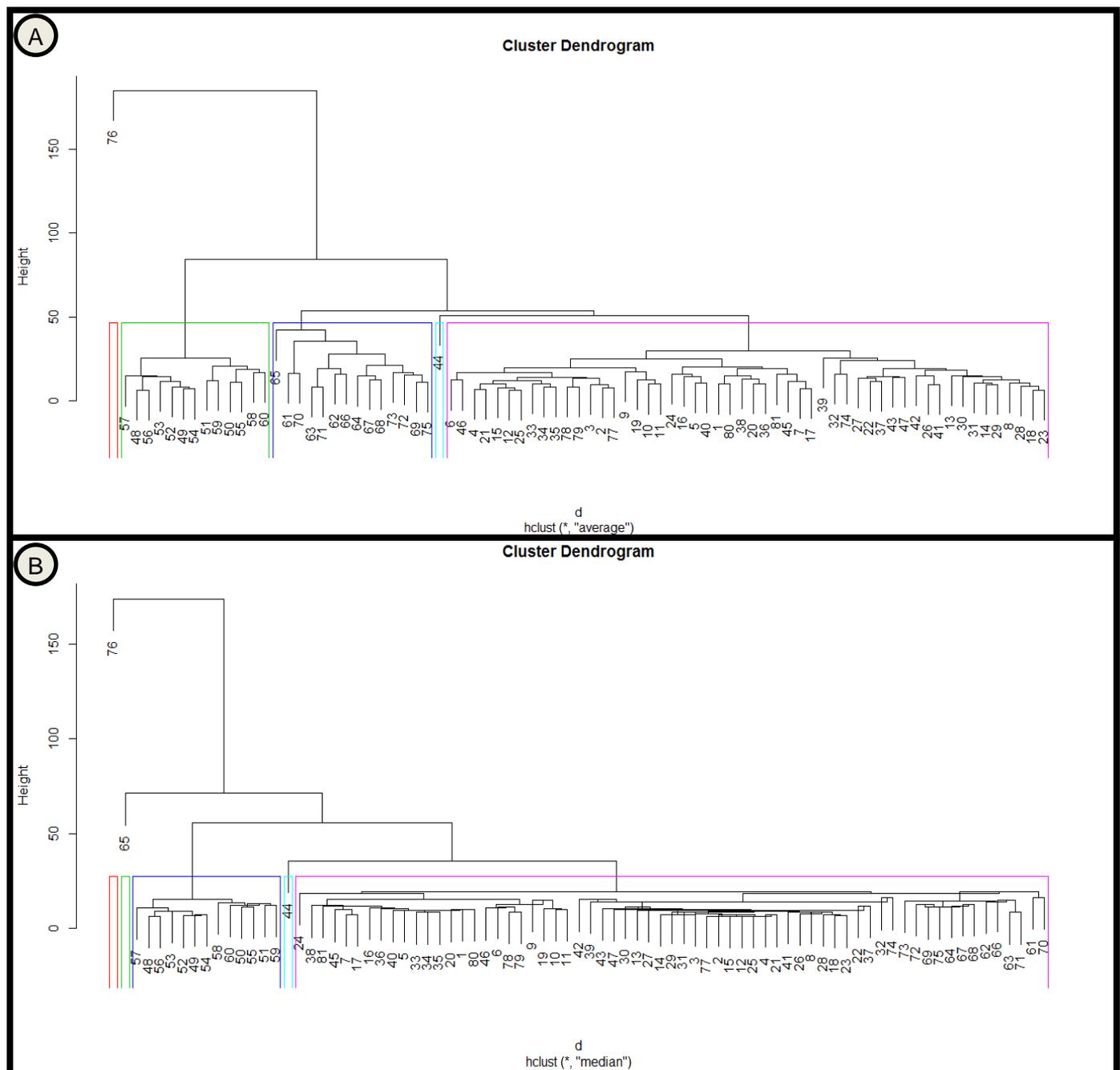
Berdasarkan Gambar 1A terlihat bahwa metode *average linkage* menunjukkan anggota klaster 1 terdiri dari 52 Kabupaten/ Kota, klaster 2 terdiri dari 1 Kabupaten/ Kota, klaster 3 terdiri dari 13 Kabupaten/ Kota, klaster 4 terdiri dari 14 kabupaten/Kota dan klaster 5 terdiri dari 1 Kabupaten/Kota. Berdasarkan Gambar 1B terlihat bahwa metode *median linkage* menunjukkan anggota klaster 1 terdiri dari 65 Kabupaten/ Kota, klaster 2 terdiri dari 1 Kabupaten/ Kota, klaster 3 terdiri dari 13 Kabupaten/ Kota, klaster 4 terdiri dari 1 kabupaten/Kota dan klaster 5 terdiri dari 1 Kabupaten/Kota.

Interpretasi Klaster *Average linkage* dan *median linkage*

Berdasarkan Tabel 6, Tabel 7, Gambar 1A dan Gambar 1B diketahui bahwa pengelompokan Kabupaten/Kota di Sulawesi menjadi 5 klaster yang memiliki jumlah anggota yang berbeda pada masing-masing klaster. Hasil pengklasteran yang telah terbentuk diperoleh kelompok Kabupaten/Kota di Sulawesi berdasarkan indikator pendidikan dengan tingkat atau kualitas pendidikan yang sangat baik hingga sangat kurang baik berturut-turut adalah

sebagai berikut:

1. Klaster 5 yang beranggotakan 1 Kabupaten/Kota dengan tingkat atau kualitas pendidikan yang sangat baik.
2. Klaster 3 yang beranggotakan 13 Kabupaten/Kota dengan tingkat atau kualitas pendidikan yang baik.
3. Klaster 1 yang beranggotakan 65 Kabupaten/Kota dengan tingkat atau kualitas pendidikan yang cukup baik.
4. Klaster 4 yang beranggotakan 1 Kabupaten/Kota dengan tingkat atau kualitas pendidikan yang



Gambar 1. Dendrogram Analisis Klaster: *Average linkage* (A), *Median linkage* (B)

kurang baik.

5. Kluster 2 yang beranggotakan 1 Kabupaten/Kota dengan tingkat atau kualitas pendidikan yang sangat kurang baik.

Penentuan Metode Terbaik

Setelah terbentuk 5 kluster dari metode *average linkage* dan *median linkage*, selanjutnya penentuan metode yang menghasilkan kluster terbaik. Pengecekan dilakukan menggunakan simpangan baku dalam kluster dan simpangan baku antar kluster.

1. Simpangan Baku Dalam Kluster

Perhitungan nilai simpangan baku dalam kluster metode *average linkage*, sebagai berikut:

$$S_W = K^{-1} \sum_{k=1}^K S_k$$

$$S_W = 5^{-1}(s_1 + s_2 + s_3 + s_4 + s_5)$$

$$S_W = \frac{(s_1 + s_2 + s_3 + s_4 + s_5)}{5}$$

$$S_W = \frac{(3.049+0+2.289+2.154+0)}{5}$$

$$S_W = \frac{7.492}{5} = 1.49$$

Perhitungan nilai simpangan baku dalam kluster metode *median linkage*, sebagai berikut:

$$S_W = K^{-1} \sum_{k=1}^K S_k$$

$$S_W = 5^{-1}(s_1 + s_2 + s_3 + s_4 + s_5)$$

$$S_W = \frac{(s_1 + s_2 + s_3 + s_4 + s_5)}{5}$$

$$S_W = \frac{(3.589+0+2.289+0+0)}{5}$$

$$S_W = \frac{5.878}{5} = 1.18$$

2. Simpangan Baku Antar Kluster

$$S_B = \left[(K - 1)^{-1} \sum_{k=1}^K (\bar{X}_k - \bar{X})^2 \right]^{1/2}$$

Perhitungan nilai simpangan baku antar kluster metode *average linkage*, sebagai berikut:

$$S_B = \left(\frac{(69.12 - 62.838)^2 + \dots + (29.85 - 62.838)^2}{5 - 1} \right)^{1/2}$$

$$S_B = \left(\frac{1446.64}{4} \right)^{1/2} = (361.66)^{1/2} = 19.017$$

Perhitungan nilai simpangan baku antar kluster metode *median linkage*, sebagai berikut:

$$S_B = \left(\frac{(70.18 - 63.19)^2 + \dots + (29.85 - 63.19)^2}{5 - 1} \right)^{1/2}$$

$$S_B = \left(\frac{1477.68}{4} \right)^{1/2} = (369.42)^{1/2} = 19.220$$

Untuk mengetahui kinerja kedua metode atau melihat kualitas pengelompokkan yang terbaik, dilihat berdasarkan nilai rasio terkecil. Berdasarkan Tabel 8 metode *median linkage* memperoleh nilai rasio sebesar 0.061 lebih kecil dibandingkan dengan rasio *average linkage* sebesar 0.078. Sehingga dapat disimpulkan pengelompokkan Kabupaten/Kota di Sulawesi berdasarkan indikator pendidikan tahun 2017 lebih baik dengan menggunakan metode *median linkage*.

Ucapan Terimakasih

Penulis mengucapkan terima kasih kepada pihak Badan Pusat Statistik Provinsi-Provinsi di Sulawesi yang telah memperkenankan penulis untuk mengambil data mengenai indikator pendidikan.

Daftar Pustaka

Badan Pusat Statistik Provinsi Sulawesi Barat [BPS Sulawesi Barat]. (2017). *Statistik Kesejahteraan Rakyat*. Mamuju: BPS Sulawesi Barat.

Badan Pusat Statistik Provinsi Gorontalo [BPS Gorontalo]. (2017). *Indikator Pendidikan*. Gorontalo: BPS Gorontalo.

Badan Pusat Statistik Indonesia [BPS Indonesia]. (2018). *Potret Pendidikan Indonesia*. Jakarta: BPS Indonesia.

Badan Pusat Statistik Provinsi Sulawesi Selatan [BPS Sulawesi Selatan]. (2017). *Statistik Pendidikan*. Makassar: BPS Sulawesi Selatan.

Badan Pusat Statistik Provinsi Sulawesi Tengah [BPS Sulawesi Tengah]. (2017). *Statistik Kesejahteraan Rakyat*. Palu: BPS Sulawesi Tengah.

Badan Pusat Statistik Provinsi Sulawesi Tenggara [BPS Sulawesi Tenggara]. (2017). *Indikator Statistik Kesejahteraan Rakyat*. Kendari: BPS Sulawesi Tenggara.

Badan Pusat Statistik Provinsi Sulawesi Utara [BPS Sulawesi Utara]. (2017). *Statistik Kesejahteraan Rakyat*. Manado: BPS Sulawesi Utara.

Fadliana, A., dan Rozi, F. (2015). *Penerapan Agglomerative Hierarchical Clustering untuk Klasifikasi Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Kualitas Pelayanan Keluarga Berencana*. *Jurnal Cauchy*, 4(1): 25-40.

Goreti, M., Novia, Y., dan Wahyuningsih, S. (2016). *Perbandingan Hasil Analisis Cluster dengan Menggunakan Metode Single Linkage dan Metode C-Means*. *Jurnal Eksponensial*, 7(1): 9-16.

- Gudono.(2015). *Anlisis Data Multivariat Edisi Empat*.Yogyakarta: BPFE.
- Hardius, U., dan Nurdin, S.(2013). *Aplikasi Teknik Multivariate Riset Pemasaran*. Jakarta: Raja Grafindo Persada.
- Ningsih, S., Wahyuningsih, S., dan Nasution, Yuki, N. (2016). *Perbandingan Kinerja Metode Complete Linkage dan Average linkage dalam Menentukan Hasil Analisis Cluster*. Jurnal Sains dan Teknologi, 1(1): 46-50.
- Sitepu, R., Irmeilyana., dan Gultom, B. (2011). *Analisis Klaster Terhadap Tingkat Pencemaran Udara pada Sektor Industri Di Sumatera Selatan*.Jurnal Penelitian Sains, 14(3): 11-17.
- Suliyanto. (2005). *Analisis Data dalam Aplikasi Pemasaran*. Bogor: Penerbit Ghalia Indonesia.
- Supranto, J. (2004). *Analisis Multivariat Arti dan Interpretasi Edisi Pertama*. Jakarta: PT Rineka Cipta.
-