
PENGELOMPOKAN DESA ATAU KELURAHAN DI KUTAI KARTANEGARA MENGGUNAKAN ALGORITMA *DIVISIVE ANALYSIS*

Ilham Adnan Kasoqi¹, Memi Nor Hayati², Rito Goejantoro³

¹Program Studi Statistika, Jurusan Matematika, FMIPA, Universitas Mulawarman

^{2,3}Dosen Program Studi Statistika, Jurusan Matematika, FMIPA, Universitas Mulawarman

Alamat e-mail : ilhamkasoqi@gmail.com

ABSTRACT

Potential Villages (PODES) provide data on the existence, availability and development of the potential of each government administrative area. In order to make it easier for governments to make policies for a region, it is necessary to group the village and sub-districts. Cluster analysis is an analysis that aims to group objects based on the information that found in the data. One of the cluster analysis methods is the divisive analysis, which is a hierarchical grouping method with a top-down approach, where all objects are placed in one cluster and then sequentially divided into separate groups. This research aim to group villages or sub-districts in Kutai Kartanegara based on the determinants of village backwardness and obtaining the silhouette coefficient value from the optimal cluster analysis using the divisive analysis algorithm. The data used is the 2018 PODES data in Kutai Kartanegara and used 15 variables from natural and environmental factors, facilities infrastructure and access factors as well as socio-economic factors of the population. The results of the optimal cluster formed in the grouping of villages or sub-districts in Kutai Kartanegara using the divisive analysis method are 2 clusters. Cluster 1 consisting of 230 villages or sub-districts and cluster 2 consisting of 2 sub-districts. Silhouette coefficient value for data validation from clustering village or sub-districts in Kutai Kartanegara using the divisive analysis method produces 2 clusters is 0,744 which states that the cluster structure formed in this grouping is a strong structure.

Keywords: divisive analysis, village potential, silhouette coefficient

PENDAHULUAN

Pembangunan direncanakan dan dilaksanakan sesuai dengan keadaan daerah, kemampuan untuk berkembang dan kemajuan yang ingin dicapai secara daerah. Kemajuan yang ingin dicapai ini merupakan tuntutan sekaligus tantangan bagi daerah itu sendiri. Adapun keberhasilan suatu daerah dalam usaha mencapai kemajuan yang diinginkan sangat ditentukan oleh kemampuan penyelenggara daerah serta keadaan dan kedudukan daerah di antara daerah lainnya.

Pemberlakuan Undang-Undang Nomor 32 Tahun 2004 tentang Pemerintahan Daerah dan Undang-Undang Nomor 33 Tahun 2004 tentang Perimbangan Keuangan antara Pemerintah Pusat dan Pemerintah Daerah menuntut pemerintah daerah untuk melakukan desentralisasi dan memacu pertumbuhan ekonomi guna peningkatan kesejahteraan masyarakat. Kedua Undang-Undang tersebut memiliki makna yang penting bagi daerah, karena daerah dapat mengatur, menggali dan menikmati sumber-sumber potensi ekonomi

serta sumber daya alamnya tanpa ada intervensi terlalu jauh dari pemerintah pusat.

Pembangunan daerah diprioritaskan bagi daerah-daerah dengan jumlah penduduk miskin terbesar. Dalam upaya pengentasan kemiskinan, Badan Pusat Statistik (BPS) berperan dalam menyediakan data tentang kemiskinan. Kabupaten Kutai Kartanegara (Kukar) tercatat sebagai penyumbang penduduk miskin terbesar di Provinsi Kalimantan Timur (Kaltim) dengan jumlah penduduk miskin 56.560 jiwa pada Tahun 2018. Kota Samarinda dan Kutai Timur berada di posisi tertinggi kedua dan ketiga dengan jumlah penduduk miskin masing-masing sebesar 39.230 jiwa dan 33.020 jiwa [1].

Suatu wilayah mempunyai karakteristik tertentu, yang meliputi aspek fisik dan aspek non fisik. Aspek-aspek ini merupakan salah satu cara untuk menerangkan keragaman antar wilayah, salah satunya wilayah desa atau kelurahan. Suatu wilayah desa atau kelurahan memiliki kondisi yang berbeda dengan wilayah lainnya. Potensi Desa (PODES) menyediakan data tentang keberadaan, ketersediaan dan perkembangan potensi yang dimiliki setiap wilayah administrasi pemerintah. Potensi tersebut meliputi sarana dan prasarana wilayah serta potensi ekonomi, sosial, budaya dan aspek kehidupan masyarakat lainnya [2]. Agar memudahkan pemerintah daerah dalam memberikan bantuan dan kebijakan suatu wilayah, perlu dilakukan pengelompokan desa/kelurahan.

Salah satu metode yang umum digunakan untuk pengelompokan objek adalah *clustering*. Analisis *cluster* merupakan analisis yang bertujuan untuk mengelompokkan data (objek) yang didasarkan hanya pada informasi yang ditemukan dalam data yang menggambarkan objek tersebut dan hubungan di antaranya. *Clustering* menurut strukturnya dibagi menjadi dua, yaitu *non hierarchical clustering*

dan *hierarchical clustering*. Pada pengelompokan *hierarchical*, satu data tunggal bisa dianggap sebagai satu kelompok, dua kelompok atau lebih. Kelompok kecil dapat bergabung menjadi sebuah kelompok besar dan begitu seterusnya hingga seluruh data dapat bergabung menjadi sebuah kelompok dan akan membentuk semacam pohon dimana ada *hierarki* (tingkatan) yang jelas antar objek dari yang paling mirip hingga yang paling tidak mirip [3].

Strategi untuk pengelompokan *hierarchical* pada umumnya dibagi menjadi dua jenis yaitu *agglomerative* (pemusatan) dan *divisive* (penyebaran). Pengelompokan hierarki *agglomerative* merupakan metode pengelompokan hierarki dengan pendekatan *bottom-up*, yaitu pasangan objek terhubung secara berurutan untuk menghasilkan *cluster* yang lebih besar. Pengelompokan hierarki *divisive* menggunakan pendekatan *top-down*, yaitu semua objek ditempatkan dalam satu *cluster* dan kemudian secara berturut-turut dibagi menjadi kelompok-kelompok yang terpisah. Contoh metode *divisive* ialah *monothetic divisive clustering* dan *divisive analysis* [4].

Pengelompokan *divisive analysis* merupakan metode pengelompokan hierarki yang memecah belah dengan menggunakan strategi *top-down*. Dimulai dengan menempatkan semua objek dalam satu *cluster* besar, yang merupakan *root* hierarki. Kemudian membagi *root cluster* menjadi beberapa sub-*cluster* yang lebih kecil, dan secara berulang memecah *cluster* tersebut menjadi yang lebih kecil. Proses partisi berlanjut hingga setiap *cluster* di level terendah yang hanya mengandung satu objek, atau objek dalam sebuah *cluster* cukup mirip satu sama lain atau bisa disebut dengan *splinter group* [5].

Dalam rangka penyaluran bantuan, pemerintah harus menetapkan daerah tertinggal agar bantuan tersebut dapat

disalurkan dengan tepat. Penentuan desa tertinggal dapat dilakukan dengan mengelompokkan desa/kelurahan sesuai karakteristik wilayahnya. Berdasarkan latar belakang tersebut, maka penulis tertarik membuat penelitian dengan judul “Pengelompokan Desa atau Kelurahan di Kutai Kartanegara Menggunakan Algoritma *Divisive Analysis* (Studi Kasus: Data Potensi Desa Kutai Kartanegara Tahun 2018)”.

METODOLOGI PENELITIAN

Sumber Data dan Variabel Penelitian

Penelitian ini menggunakan data sekunder yang diperoleh dari BPS Provinsi Kalimantan Timur yang beralamat di Jalan Kemakmuran Samarinda. Variabel yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Kepadatan Penduduk (X_1)
2. Ketersediaan sarana pendidikan/sekolah (X_2)
3. Ketersediaan tenaga kesehatan (X_3)
4. Ketersediaan sarana kesehatan (X_4)
5. Jumlah keluarga berlangganan telepon kabel (X_5)
6. Jumlah Penginapan (X_6)
7. Jumlah bangunan pasar permanen/semi permanen (X_7)
8. Jumlah supermarket/ minimarket/pasar swalayan/toserba (X_8)
9. Jumlah keberadaan bank (X_9)
10. Jumlah penduduk memperoleh fasilitas perkreditan (X_{10})
11. Jumlah keberadaan Koperasi Non KUD lainnya (X_{11})
12. Jumlah keluarga pengguna listrik PLN (X_{12})
13. Jumlah keluarga tinggal di bantaran tepi sungai (X_{13})
14. Jumlah keluarga tinggal di pemukiman kumuh (X_{14})
15. Jumlah penderita gizi buruk (X_{15})

Pengolahan data pada penelitian ini menggunakan metode pengelompokan *divisive analysis*.

Metode Analisis

Tahapan dari pengelompokan *divisive analysis* dalam penelitian ini adalah sebagai berikut:

1. Melakukan standarisasi data.
3. Mendeteksi multikolinieritas
4. Melakukan pengelompokan data dengan menerapkan algoritma *divisive analysis* untuk setiap $k = 2, 3, 4,$ dan 5 dengan tahapan sebagai berikut

- a. Menghitung jarak Euclid

$$d(i, r) = \sqrt{(x_{i1}-x_{r1})^2 + (x_{i2}-x_{r2})^2 + \dots + (x_{iq}-x_{rq})^2} \quad (1)$$

- b. Menghitung rata-rata setiap objek berdasarkan matriks jarak Euclid

$$\bar{d}_i = \frac{1}{n-1} \sum_{r=1}^n d(i, r) \quad , r \neq i \quad (2)$$

- c. Menentukan objek yang memiliki nilai rata-rata yang terbesar dan berubah menjadi *splinter group*.
- d. Menghitung selisih nilai antara elemen matriks *splinter group* dengan nilai rata-rata setiap objek yang tersisa
- e. Mengulang langkah (b) sampai (d) sedemikian sehingga semua nilai selisih antara elemen matriks *splinter group* dengan nilai rata-rata bernilai negatif dan *cluster* terbagi menjadi dua *cluster* baru.
5. Menghitung nilai *silhouette coefficient* dengan tahapan sebagai berikut:

- a. Menghitung rata-rata jarak data ke- i dengan semua data pada *cluster* yang sama

$$a_i = \frac{1}{N_p - 1} \sum_{r=1}^{N_p-1} d_{i,r} \quad , r \neq i \quad (3)$$

- b. Menghitung rata-rata jarak data ke- i dengan semua data pada *cluster* yang berbeda

$$b_i = \frac{1}{N_p} \sum_{r=1}^{N_p} d_{i,r} \quad (4)$$

c. Menghitung nilai $SC_1(i)$

$$SC_1(i) = \frac{b_i - a_i}{\max\{a_i, b_i\}}, i = 1, 2, \dots, n \quad (5)$$

d. Menghitung rata-rata nilai $SC_2(p)$

$$SC_2(p) = \frac{1}{N_p} \sum_{x_i \in C_p} SC_1(i) \quad (6)$$

e. Menghitung nilai SC global

$$SC = \frac{\sum_{p=1}^K (N_p \times SC_2(p))}{\sum_{p=1}^K N_p} \quad (7)$$

f. Menentukan nilai k optimal berdasarkan nilai *silhouette coefficient* terbesar.

HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini ialah data Potensi Desa Kabupaten Kutai Kartanegara pada tahun 2018. Data tersebut terdiri dari 232 desa/kelurahan. Data akan dianalisis menggunakan metode *divisive* dalam *clustering*, di mana jumlah *cluster* yang akan digunakan adalah 2, 3, 4, dan 5 *cluster*. Evaluasi hasil *clustering* menggunakan metode *silhouette coefficient*. Variabel yang digunakan dalam penelitian ini sebanyak 15 variabel yang tersedia di BPS Provinsi Kaltim dengan jenis data numerik.

Statistika Deskriptif

Berikut merupakan gambaran umum mengenai data yang digunakan dalam penelitian, yaitu data Potensi Desa (PODES) Kutai Kartanegara Tahun 2018.

Tabel 1. Statistika Deskriptif

Variabel	Mean	St.Dev	Min	Max
X_1	119,859	233,740	0,11	1687,11
X_2	6,655	5,139	0	32
X_3	3,039	3,365	0	29
X_4	5,069	8,376	0	80
X_5	5,138	47,743	0	700
X_6	0,341	1,087	0	9
X_7	0,388	0,640	0	3
X_8	36,733	50,959	0	475
X_9	0,328	0,844	0	6
X_{10}	0,328	0,774	0	5
X_{11}	0,478	1,102	0	9
X_{12}	915,047	980,791	6	7261
X_{13}	62,155	159,305	0	1125
X_{14}	4,125	28,529	0	347
X_{15}	0,155	0,543	0	4

Analisis Cluster dengan *Divisive analysis*

Sebelum memasuki tahap analisis *cluster*, terlebih dahulu dilakukan standarisasi data dan selanjutnya dilakukan juga pendeteksian multikolinieritas. Salah satu metode untuk mendeteksi multikolinieritas adalah dengan menghitung nilai *Variance Inflation Factor (VIF)* dari setiap variabel. Dalam penelitian ini didapatkan hasil bahwa tidak terjadi multikolinieritas pada semua variabel.

Selanjutnya, data dianalisis menggunakan metode *divisive* dalam *clustering*. Pertama, menghitung jarak Euclid menggunakan persamaan (1). Perhitungan jarak Euclid untuk objek ke-2 dan objek ke-1 adalah sebagai berikut. Perhitungan jarak Euclid dilanjutkan hingga objek ke-232 dan objek ke-231.

$$\begin{aligned} d(2,1) &= \sqrt{(x_{2,1}-x_{1,1})^2 + \dots + (x_{2,15}-x_{1,15})^2} \\ &= \sqrt{((-0,070)-(-0,372))^2 + \dots + ((-0,286)-(-0,286))^2} \\ &= 3,123 \end{aligned}$$

Setelah itu, menghitung nilai rata-rata jarak Euclid setiap objek menggunakan persamaan (2). Perhitungan nilai rata-rata

jarak Euclid setiap objek dilanjutkan hingga objek ke-232.

$$\begin{aligned} \bar{d}_1 &= \frac{1}{232-1} (3,123 + 4,799 + \dots + 1,882) \\ &= 3,439 \end{aligned}$$

Hasil perhitungan dapat dilihat di bawah ini.

Tabel 2. Nilai Rata-rata Jarak Euclid Setiap Objek

D	1	2	...	232	Rata-rata
1	0	3,123	...	1,882	3,439
2	3,123	0	...	3,081	4,657
3	4,794	2,051	...	4,665	5,975
⋮	⋮	⋮	⋮	⋮	⋮
210	18,378	18,907	...	19,133	19,043
⋮	⋮	⋮	...	⋮	⋮
232	1,882	3,081	...	0	3,375

Selanjutnya, menentukan objek yang memiliki nilai rata-rata terbesar. Objek yang memiliki nilai rata-rata terbesar akan dipisah dan berubah menjadi *splinter group*. Berdasarkan Tabel 2, objek yang memiliki nilai rata-rata terbesar adalah objek ke-210 yaitu sebesar 19,043 sehingga objek tersebut menjadi *splinter group 1*. Kemudian menghitung ulang nilai rata-rata setiap objek yang tersisa tanpa melibatkan objek ke-210 yang telah menjadi *splinter group 1*. Setelah itu, menghitung selisih antara nilai elemen matriks *splinter group 1* dengan nilai rata-rata setiap objek yang tersisa. Hasil perhitungan dapat dilihat pada Tabel 3 berikut.

Tabel 3. Selisih antara Nilai Elemen Matriks *Splinter group 1* dengan Nilai Rata-rata Setiap Objek

D	Rata-rata	<i>Splinter 1</i> (Objek ke-210)	Selisih
1	3,374	18,378	-15,004
2	4,595	18,907	-14,313
⋮	⋮	⋮	⋮
205	14,881	13,062	1,819
⋮	⋮	⋮	⋮
232	3,307	19,133	-15,826

Selanjutnya, menentukan objek yang memiliki nilai selisih terbesar antara elemen matriks *splinter group* dengan nilai rata-rata.

Jika nilai selisih tersebut bernilai positif, maka objek yang memiliki nilai selisih terbesar bergabung dengan *splinter group*. Berdasarkan Tabel 3, objek ke-205 memiliki nilai selisih terbesar yaitu 1,819 sehingga objek tersebut bergabung dengan *splinter group 1*. Kemudian, menghitung selisih nilai antara elemen matriks *splinter group 1* dengan nilai rata-rata setiap objek yang tersisa. Hasil perhitungan dapat dilihat pada Tabel 4 di bawah ini.

Tabel 4. Selisih antara Nilai Elemen Matriks *Splinter group 1* dengan Nilai Rata-rata Setiap Objek yang Tersisa

D	Rata-rata	<i>Splinter 1</i> (Objek ke-205)	Selisih
1	3,327	14,219	-10,893
2	4,553	14,179	-9,626
⋮	⋮	⋮	⋮
214	13,803	14,402	-0,599
⋮	⋮	⋮	⋮
232	3,258	14,547	-11,289

Selanjutnya, mengulangi tahapan di atas sedemikian sehingga nilai selisih terbesar antara elemen matriks *splinter group 1* dengan nilai rata-rata yang tersisa bernilai negatif dan membentuk *splinter group* baru. Karena objek ke-214 memiliki nilai selisih terbesar yang negatif, maka objek ke-214 menjadi *splinter group 2*.

Proses iterasi untuk 2 *cluster* dihentikan sampai terbentuk *splinter group 2*. Untuk pembentukan 3 *cluster*, proses iterasi berhenti pada *splinter group 3*. Begitupun hingga pembentukan 5 *cluster*. Hasil pengelompokan *divisive analysis* untuk 2, 3, 4 dan 5 *cluster* dapat dilihat pada Tabel 5

Tabel 5. Anggota dari *Cluster* yang Terbentuk

Jumlah Cluster	Anggota Cluster
2	{(1,2,...,232),(210,205)}
3	{(1,2,...,232),(210,205),(178)}
4	{(1,2,...,232),(210,205),(178),(96)}
5	{(1,2,...,232),(210,205),(178),(96),(98,214)}

Setelah melewati tahapan dalam *divisive analysis*, dilakukan evaluasi hasil *clustering* menggunakan metode *Silhouette coefficient*. Hal ini dilakukan untuk mengetahui kualitas dari hasil pengelompokan. Berikut tahapan dalam menghitung nilai *Silhouette coefficient*.

1. Menghitung rata-rata jarak data ke-*i* dengan semua data yang berada pada *cluster* yang sama dengan menggunakan persamaan (3).

$$\begin{aligned}
 a_1 &= \frac{1}{N_1 - 1} \sum_{r=1}^{N_1 - 1} d_{1,r} \\
 &= \frac{1}{230 - 1} \sum_{r=1}^{229} d_{1,r} \\
 &= \frac{1}{229} [d(1, 2) + d(1, 3) + \dots + d(1, 232)] \\
 &= 3,327
 \end{aligned}$$

Dengan cara yang sama dihitung semua nilai a_i untuk anggota dalam *cluster* 1 sampai dengan anggota terakhir yaitu a_{232} . Begitupun pada *cluster* 2.

2. Menghitung rata-rata jarak data ke-*i* dengan semua data pada *cluster* yang berbeda dengan menggunakan persamaan (4)

$$\begin{aligned}
 b_1 &= \frac{1}{N_2} \sum_{r=1}^{N_2} d_{i,r} \\
 &= \frac{1}{2} [d(1, 205) + d(1, 210)] \\
 &= \frac{1}{2} [14, 219 + 18, 378] \\
 &= 16,299
 \end{aligned}$$

Selanjutnya dengan cara yang sama dihitung semua nilai b_i untuk anggota dalam *cluster* 1 sampai dengan anggota terakhir yaitu b_{232} . Begitupun pada *cluster* 2.

3. Menghitung nilai SC_1 data ke-*i* dengan menggunakan persamaan (5)

$$\begin{aligned}
 SC_1(1) &= \frac{b_1 - a_1}{\max\{a_1, b_1\}} \\
 &= \frac{16,299 - 3,329}{\max\{(16,299), (3,329)\}} = 0,796
 \end{aligned}$$

Demikian seterusnya dengan cara yang sama dihitung $SC_1(i)$ untuk $i = 1, 2, \dots, 232$.

4. Menghitung nilai SC_2 pada *cluster* ke-*p* dengan menggunakan persamaan (6). Perhitungan nilai SC untuk *cluster* 1 adalah sebagai berikut

$$\begin{aligned}
 SC_2(1) &= \frac{1}{N_1} \sum_{x_i \in C_1}^{N_1} SC_1(i) \\
 &= \frac{1}{230} (SC_1(1) + \dots + SC_1(232)) \\
 &= \frac{1}{230} (0,796 + \dots + 0,807) \\
 &= 0,748
 \end{aligned}$$

Perhitungan nilai SC_2 untuk *cluster* 2 adalah sebagai berikut:

$$\begin{aligned}
 SC_2(2) &= \frac{1}{N_2} \sum_{x_i \in C_2}^{N_2} SC_1(i) \\
 &= \frac{1}{2} (SC_1(205) + SC_1(210)) \\
 &= \frac{1}{2} (0,122 + 0,315) \\
 &= 0,218
 \end{aligned}$$

5. Menghitung nilai SC Global dengan menggunakan persamaan (7)

$$\begin{aligned}
 SC &= \frac{\{(n_1 \times SC_2(1)) + (n_2 \times SC_2(2))\}}{(n_1 + n_2)} \\
 &= \frac{\{(230 \times 0,748) + (2 \times 0,218)\}}{(230 + 2)} \\
 &= 0,744
 \end{aligned}$$

Hasil validasi untuk 2 cluster berdasarkan nilai *Silhouette coefficient (SC) global* adalah sebesar 0,744 artinya pengelompokan secara keseluruhan dapat dikatakan *strong structure* (struktur yang kuat), artinya kumpulan objek *i* yang bergabung dalam cluster yang sama memiliki karakteristik yang sangat mirip.

Dilakukan perhitungan *Silhouette coefficient* seperti langkah 1 sampai dengan langkah 5 untuk clustering dengan metode *divisive analysis* pada pembentukan 3 cluster, 4 cluster dan 5 cluster.

- Menentukan nilai k optimal berdasarkan nilai *silhouette coefficient* terbesar. Peneliti membandingkan nilai SC global dari hasil pembentukan 2, 3, 4 dan 5 cluster yang dapat dilihat pada Tabel 6 berikut ini.

Tabel 6. Nilai SC Global

Jumlah Cluster	SC
2	0,744
3	0,728
4	0,714
5	0,680

Berdasarkan Tabel 6 dapat diketahui bahwa nilai SC untuk evaluasi hasil pengelompokan data potensi desa berdasarkan variabel yang diduga sebagai penentu ketertinggalan desa di Kabupaten Kutai kartanegara dengan metode *divisive analysis* memiliki nilai SC yang berbeda-beda. Nilai SC yang terbesar adalah pengelompokan pada 2 cluster yaitu 0,744. Oleh karena itu, dapat diputuskan bahwa pengelompokan yang paling

optimal pada pengelompokan desa/kelurahan di Kabupaten Kutai Kartanegara dengan metode *divisive analysis* adalah 2 cluster. Cluster 1 beranggotakan 230 Desa/Kelurahan yang terdiri dari Desa Anggana, Handil Terusan, Kutai Lama, dan seterusnya hingga Desa Teluk Dalam. Sedangkan Cluster 2 beranggotakan 2 Desa/Kelurahan yang terdiri dari Kelurahan Loa Ipuh dan Melayu.

Karakteristik Hasil Analisis Cluster

Berdasarkan hasil pengelompokan untuk 2 cluster, dilakukan perhitungan nilai rata-rata variabel untuk melihat karakteristik dari masing-masing cluster.

Tabel 7. Nilai Rata-rata Variabel Setiap Cluster

Variabel	Cluster 1	Cluster 2
X_1	112,041	1.018,889
X_2	6,504	24
X_3	2,878	21,500
X_4	4,443	77
X_5	5,183	0
X_6	0,330	1,500
X_7	0,378	1,500
X_8	33,330	428
X_9	0,304	3
X_{10}	0,330	0
X_{11}	0,452	3,500
X_{12}	873,822	5.656
X_{13}	61,426	146
X_{14}	4,161	0
X_{15}	0,157	0

Berdasarkan Tabel 7 dapat disimpulkan bahwa dari segi kesehatan yang diwakili oleh variabel ketersediaan tenaga kesehatan (X_3), ketersediaan sarana kesehatan (X_4), jumlah keluarga tinggal di pemukiman kumuh (X_{14}) dan jumlah penderita gizi buruk (X_{15}), cluster 2 lebih baik dibandingkan cluster 1. Hal ini dapat dilihat dari variabel ketersediaan tenaga kesehatan (X_3) dan ketersediaan sarana kesehatan (X_4), di mana cluster 2 memiliki nilai rata-rata yang lebih tinggi dibandingkan

cluster 1 dan variabel jumlah keluarga tinggal di pemukiman kumuh (X_{14}) dan jumlah penderita gizi buruk (X_{15}) pada *cluster* 2 memiliki nilai rata-rata yang lebih rendah dibandingkan *cluster* 1.

Dari segi pendidikan yang diwakili oleh variabel ketersediaan sarana pendidikan/sekolah (X_2), *cluster* 2 lebih baik dibandingkan *cluster* 1. Hal ini dapat dilihat dari nilai rata-rata *cluster* 2 yang lebih tinggi dibandingkan *cluster* 1.

Selanjutnya dari segi ekonomi diwakili oleh variabel jumlah keluarga pengguna listrik PLN (X_{12}) dan jumlah keluarga tinggal di bantaran tepi sungai (X_{13}). Pada variabel jumlah keluarga pengguna listrik PLN (X_{12}), *cluster* 2 lebih baik dibandingkan *cluster* 1. Hal ini dapat dilihat dari nilai rata-rata *cluster* 2 yang lebih tinggi dibandingkan *cluster* 1. Pada variabel jumlah keluarga tinggal di bantaran tepi sungai (X_{13}), *cluster* 1 lebih baik dibandingkan *cluster* 2. Hal ini dapat dilihat dari nilai rata-rata *cluster* 1 yang lebih rendah dibandingkan *cluster* 2.

KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan, maka kesimpulan yang dapat diambil adalah sebagai berikut

1. *Cluster* optimal yang terbentuk pada pengelompokan desa/kelurahan di Kabupaten Kutai Kartanegara berdasarkan variabel yang diduga sebagai penentu ketertinggalan desa dengan menggunakan metode *divisive analysis* adalah sebanyak 2 *cluster* yaitu *cluster* 1 dan *cluster* 2. *Cluster* 1 beranggotakan 230 desa/kelurahan dan *cluster* 2 beranggotakan 2 desa/kelurahan.
2. Nilai *silhouette coefficient* optimal untuk validasi data hasil *clustering* desa/kelurahan di Kabupaten Kutai Kartanegara dengan menggunakan metode *divisive analysis* menghasilkan 2

cluster dengan nilai sebesar 0,744 yang menyatakan bahwa struktur *cluster* yang terbentuk pada pengelompokan ini adalah *strong structure*.

DAFTAR PUSTAKA

- [1] Badan Pusat Statistik. (2018). *Laporan Perekonomian Provinsi Kalimantan Timur 2018*. Samarinda: Badan Pusat Statistik
- [2] Fathia, A., Rahmawati, R., & Tarno. (2016). Analisis Klaster Kecamatan Di Kabupaten Semarang Berdasarkan Potensi Desa Menggunakan Metode Ward dan *Single Linkage*. *Jurnal Gaussian*, Vol 5, No 4, 801-810
- [3] Prasetyo, E. (2012). *Data Mining: Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta: Andi Offset.
- [4] Gorenescu, F. (2011). *Data Mining: Concept, Model, And Technique*. Berlin: Springer.
- [5] Han, J., Kamber, M. & Pei, J.. (2012). *Data Mining: Concepts and Techniques Third Edition*. San Fransisco: Morgan Kauffman Publisher.