

Identifikasi Karakteristik Desa di Provinsi Bengkulu Tahun 2018 Berdasarkan *Latent Class Cluster* (LCC) (*Identification of Village Characteristics Based on Latent Class Cluster (LCC)*)

Debora Chrisinta

Universitas Timor

Kefamenanu, Nusa Tenggara Timur

E-mail: deborachrisinta@unimor.ac.id

ABSTRAK

Permasalahan sosial ekonomi di Indonesia yang paling mendasar pada saat ini adalah masalah ketidakmerataan pembangunan ekonomi. Provinsi Bengkulu termasuk dalam kategori 10 Provinsi termiskin di Indonesia. Oleh karena itu perlu dilakukan analisis pengelompokan karakteristik masing-masing desa/kelurahan yang ada di Provinsi Bengkulu, agar sasaran kebijakan pemerintah lebih terarah. Salah satu teknik statistika yang dapat digunakan untuk mengidentifikasi karakteristik desa adalah analisis penggerombolan, salah satu metodenya adalah *Latent Class Clustering* (LCC). Metode ini berbasis sebaran data dan dapat digunakan ketika data yang ditemui berasal dari peubah campuran yaitu, kategorik dan numerik. Data yang digunakan adalah data Potensi Desa (PODES) tahun 2018. Peubah yang digunakan berasal dari dimensi Indeks Pembangunan Desa (IPD). Gerombol optimal yang diperoleh adalah 3 gerombol, gerombol terbaik ditunjukkan pada gerombol 2 yang memberikan IPD yang relatif lebih tinggi daripada gerombol 1 dan 3. Karakteristik yang ditunjukkan dalam gerombol 2 menunjukkan desa yang memiliki sarana dan prasarana yang sudah memadai dilihat berdasarkan dimensi IPD pada peubah numerik dan kategorik yang digunakan. Sedangkan gerombol 1 dan 3 memiliki karakteristik IPD yang cenderung tidak memiliki beberapa fasilitas ataupun sarana dan prasarana yang memadai.

Kata kunci: Analisis penggerombolan, *Latent Class Clustering* (LCC), PODES

ABSTRACT

The most basic socio-economic problem in Indonesia at this time is the problem of inequality in economic development. Bengkulu Province is included in the category of 10 poorest provinces in Indonesia. Therefore, it is necessary to analyze the grouping characteristics of each village in Bengkulu Province, so that the government's policy targets are more focused. One of the statistical techniques that can be used to identify village characteristics is cluster analysis, one of which is Latent Class Clustering (LCC). This method is based on data distribution and can be used when the data encountered comes from mixed variables, namely, categorical and numeric. The data used is the Village Potential data in 2018. The variables used are derived from the dimensions of the Village Development Index (VDI). The optimal clusters obtained are 3 clusters, the best clusters are shown in cluster 2 which gives a relatively higher VDI than clusters 1 and 3. The characteristics shown in cluster 2 indicate that villages that have adequate facilities and infrastructure are assessed based on the dimensions of IPD on numerical variables and the categories used. Meanwhile, clusters 1 and 3 have VDI characteristics which tend not to have several facilities or adequate infrastructure.

Keywords: Clustering analysis, *Latent Class Clustering* (LCC), Village Potential Data

PENDAHULUAN

Permasalahan sosial ekonomi di Indonesia yang paling mendasar pada saat ini adalah masalah ketidakmerataannya pembangunan ekonomi dan pendapatan warga negara di seluruh wilayah Indonesia, khususnya desa/kelurahan. Hingga saat ini, desa belum dipandang secara maksimal sebagai komponen yang potensial bagi pembangunan ekonomi. Mengingat pentingnya peranan desa dalam hal pemasokan pangan bangsa selama ini, dikarenakan Indonesia sebagai negara agraris terutama sektor pertanian yang menjadikan sebagai pusat perekonomian masyarakat secara menyeluruh. Pemerintah menjadikan pembangunan pertanian menjadi prioritas yang utama bagi pembangunan nasional. Adapun dalam meningkatkan pembangunan, aspek paling utama yang perlu diperhatikan adalah pemerataan pembangunan desa yang dapat juga diperhatikan dari berbagai aspek dan tidak hanya pada sektor pembangunan pertanian. Oleh karena itu, perhatian pemerintah sangat diperlukan dalam menangani permasalahan ini. Salah satu solusinya adalah dengan mengeluarkan kebijakan sesuai karakteristik dan potensi masing-masing wilayah khususnya desa/kelurahan. Oleh karena itu, diperlukan pengelompokan terhadap desa/kelurahan, agar sasaran kebijakan pemerintah lebih terarah sesuai dengan karakteristik masing-masing desa/kelurahan. Teknik penggerombolan menjadi sangat penting guna

memperoleh ketepatan pengelompokkan. Teknik penggerombolan akan diterapkan untuk mengelompokkan desa yang ada di Provinsi Bengkulu berdasarkan data PODES Tahun 2018. Provinsi Bengkulu termasuk dalam 10 Provinsi termiskin di Indonesia per September 2021 berdasarkan data BPS. Oleh karena itu, perlu diketahui karakteristik desa yang terdapat di Provinsi Bengkulu untuk mengetahui gambaran umum dalam upaya meningkatkan pembangunan desa.

Beberapa metode yang telah dikembangkan dalam mengatasi permasalahan penggerombolan pada peubah campuran adalah *Latent Class Cluster (LCC)*, metode *Gower*, algoritma *K-Prototypes*, *Two Step Cluster Analysis (TSCA)*, *K-Means Method* dan *cluster ensemble*. Munthe (2019), melakukan perbandingan terhadap TSCA dan *K-Prototypes*, yang menunjukkan bahwa TSCA memberikan solusi yang terbaik. Berbeda dengan Munthe (2019), sebelumnya Fachriyatul (2014) telah melakukan penelitian terhadap perbandingan metode TSCA dan LCC pada pengelompokkan data berskala campuran yang menunjukkan bahwa LCC menghasilkan pengelompokkan yang lebih baik berdasarkan nilai rata-rata rasio keragaman. Hermawati (2018) dalam penelitiannya juga mengkaji perbandingan TSCA, metode *Gower* dan *cluster ensemble* dengan kinerja yang paling baik diberikan oleh *cluster ensemble*.

Visualisasi hasil penggerombolan objek-objek yang berasal dari beberapa peubah numerik dapat dilakukan dengan menggunakan *Chernoff face*. Namun, untuk visualisasi hasil penggerombolan objek yang berasal dari peubah kategorik belum tersedia dan penyajian hasil penggerombolan masih menggunakan tabel.

Berdasarkan beberapa penelitian sebelumnya menunjukkan bahwa LCC menunjukkan metode yang cenderung menghasilkan kualitas gerombol lebih baik. Oleh karena itu pada penelitian ini melakukan penggerombolan LCC pada data PODES Provinsi Bengkulu Tahun 2018 dengan objek pengamatan adalah desa-desa yang ada di Provinsi Bengkulu.

METODE

Analisis Gerombol

Analisis gerombol (*cluster analysis*) merupakan salah satu teknik peubah ganda yang digunakan untuk mengelompokkan sekumpulan objek amatan berdasarkan karakteristik objek. Karakteristik objek dalam suatu gerombol haruslah memiliki tingkat kemiripan yang tinggi, sedangkan objek antar gerombol memiliki tingkat kemiripan yang rendah/ketidakmiripan. Oleh karena itu, keragaman dalam suatu gerombol minimum dan sebaliknya untuk keragaman antar gerombol maksimum (Mattjik dan Sumertajaya, 2011).

Analisis gerombol termasuk dalam metode klasifikasi tidak terbimbing (*unsupervised learning*). Metode *unsupervised learning* merupakan suatu metode analisis data dengan tujuan mengelompokkan objek yang hampir sama dalam suatu area tertentu. Bermula dari data yang ada, kemudian mengelompokkan data tersebut menjadi 2 atau lebih kelompok. Beberapa metode yang dapat digunakan pada metode klasifikasi tidak terbimbing adalah *Hierarchical Clustering*, *K-Means*, dan *Fuzzy C-Means*. *Hierarchical Clustering* atau dikenal dengan metode hirarki dapat digunakan jika kelompok yang akan dibentuk belum diketahui. Metode ini memiliki dua jenis teknik yaitu teknik pemusatan (*agglomerative*) dan penyebaran (*divisive*). Umumnya, teknik pemusatan yang sering digunakan dan terbagi lagi menjadi 5 metode yaitu *Single Linkage*, *Complete Linkage*, *Average Linkage*, *Ward's Method*, dan *Centroid*. Berbeda dengan hirarki, *K-Means*, dan *Fuzzy C-Means* merupakan kebalikan dari metode hirarki yaitu dalam melakukan pengelompokkan objek jumlah gerombol yang diinginkan telah ditentukan dan sering disebut sebagai metode tak berhirarki (*non-hierarchical clustering*).

Kedua metode tersebut, baik hirarki dan tak berhirarki cocok diterapkan pada peubah numerik. Namun, tidak menutup kemungkinan bahwa data yang diamati tidak hanya berasal dari peubah numerik, bisa saja berasal dari peubah campuran. Oleh karena itu, pada penelitian ini akan mengkaji metode penggerombolan yang dapat diterapkan pada peubah campuran. Metode yang terpilih adalah hasil dari kajian perbandingan metode terbaik dari penelitian sebelumnya yaitu LCC.

Metode Latent Class Clustering (LCC)

Metode *Latent Class Clustering (LCC)* merupakan salah satu analisis penggerombolan yang menggunakan peluang posterior dalam melakukan pengelompokkan objek (Vermunt dan Magidson, 2002). Misalkan terdapat vektor p peubah yang merupakan respon objek ke- h dan jika ditulis $\mathbf{x}'_h = [x_{1h}, x_{2h}, \dots, x_{ph}]$ dengan masing-masing peubah memiliki sebaran bersyarat seperti multinomial dan normal. Secara umum, sebaran peluang model LCC yang berasal dari peubah campuran dapat dituliskan sebagai berikut:

$$f(\mathbf{x}_h) = \sum_{j=1}^K \eta_j [g(\mathbf{x}_h|j)] \dots \dots \dots (1)$$

dengan, $g(\mathbf{x}_h|j)$ adalah fungsi campuran peubah numerik dan kategorik, serta η_j adalah prior kelompok ke- j .

Sebaran bersyarat bahwa objek terdapat dalam kelas j jika diketahui \mathbf{x} disebut sebagai peluang posterior dan dapat ditulis sebagai:

$$h(j|\mathbf{x}_h) = \frac{\eta_j [g(\mathbf{x}_h|j)]}{f(\mathbf{x}_h)} \dots \dots \dots (2)$$

Peubah nominal diasumsikan mengikuti sebaran multinomial, dapat dituliskan sebagai berikut:

$$g_i(\mathbf{x}_i|j) = \prod_{s=1}^{c_i} (\pi_{ij(s)})^{x_{i(s)}} \dots \dots \dots (3)$$

dengan c_i merupakan jumlah kategori pada peubah ke- i dengan $\sum_{s=1}^{c_i} x_{i(s)} = 1$, serta $\pi_{ij(s)}$ adalah peluang suatu objek kategori s pada peubah i yang terdapat pada kelas j . Selanjutnya, untuk peubah numerik yang diasumsikan menyebar normal peluangnya dapat dituliskan sebagai berikut:

$$g_i(\mathbf{x}_i|\mu_{ij}, \sigma_i^2) = (2\pi)^{-1/2} \sigma_i^{-1/2} \exp \left[-\frac{1}{2\sigma_i^2} (x_i - \mu_{ij})^2 \right] \dots \dots \dots (4)$$

dengan $-\infty < x_i < \infty$, μ_{ij} adalah parameter lokasi dari peubah kontinu x_i pada kelas ke- j dan σ_i^2 adalah ragam pada peubah ke- i .

Parameter pada masing-masing peubah diduga dengan membentuk terlebih dahulu fungsi *log-likelihood* sebagai berikut:

$$L = \sum_{h=1}^n \log f(\mathbf{x}_h) = \sum_{h=1}^n \log \sum_{j=1}^K \eta_j g(\mathbf{x}_h|j), \quad j = 1, \dots, K \dots \dots \dots (5)$$

Fungsi diatas dapat dimaksimumkan dengan menggunakan algoritma *Expectation Maximization* (EM), dengan $\sum_{j=1}^K \eta_j = 1$. Persamaan (6) memberikan hasil pendugaan untuk peluang posterior berikut:

$$\hat{h}(j|\mathbf{x}_h) = \hat{\eta}_j \hat{g}(\mathbf{x}_h|j) / \hat{f}(\mathbf{x}_h) \dots \dots \dots (6)$$

Rasio Simpangan Baku (Sw/Sb)

Evaluasi hasil penggerombolan dapat dilihat berdasarkan rasio antara nilai simpangan baku di dalam kelompok (S_w) dan simpangan baku antar kelompok (S_B). Selain itu pada proses simulasi akan dilihat pula nilai akurasi yang dihasilkan metode dalam menjelaskan ketepatan metode pada saat melakukan penggerombolan. Kehomogenan gerombol yang maksimum dan hetegogenitas yang maksimum antar gerombol dapat dilihat dalam bentuk nilai rasio S_w dan S_B yang kecil (Bunkers *et al.* 1996). Rumus perhitungan S_w dan S_B untuk peubah dengan tipe numerik adalah sebagai berikut:

$$S_w = \frac{1}{C} \sum_{c=1}^C S_c \dots \dots \dots (7)$$

$$S_B = \left[\frac{1}{C-1} \sum_{c=1}^C (\bar{x}_c - \bar{x})^2 \right]^{\frac{1}{2}} \dots \dots \dots (8)$$

dengan, S_w adalah nilai simpangan baku di dalam kelompok, S_B adalah nilai simpangan baku antar kelompok, S_c adalah nilai simpangan baku kelompok ke- c , \bar{x}_c adalah rata-rata kelompok ke- c , \bar{x} adalah rata-rata keseluruhan kelompok dan C adalah jumlah kelompok yang terbentuk. Sedangkan S_w dan S_B untuk peubah dengan tipe kategorik adalah sebagai berikut:

$$S_w = [MSW]^{\frac{1}{2}} \dots \dots \dots (9)$$

$$S_B = [MSB]^{\frac{1}{2}} \dots \dots \dots (10)$$

$$MSW = \frac{SSW}{n-C} = \frac{1}{n-C} \left[\sum_{c=1}^C \left(\frac{n_c}{2} - \frac{1}{2n_c} \sum_{k=1}^K n_{kc}^2 \right) \right] \dots \dots \dots (11)$$

$$MSB = \frac{SSB}{C-1} = \frac{1}{C-1} \left[\frac{1}{2} \left(\sum_{c=1}^C \frac{1}{n_c} \sum_{k=1}^K n_{kc}^2 \right) - \frac{1}{2n} \sum_{k=1}^K n_k^2 \right] \dots \dots \dots (12)$$

$$n = \sum_{c=1}^C n_c = \sum_{k=1}^K n_k = \sum_{k=1}^K \sum_{c=1}^C n_{kc} \dots \dots \dots (13)$$

dengan S_w adalah nilai simpangan baku di dalam kelompok data kategorik, S_B adalah nilai simpangan baku antar kelompok data kategorik, MSW adalah *Mean of squares within*, MSB adalah *Mean of squares between c*, SSW adalah *Sum of squares within*, SSB adalah *Sum of squares between*, C adalah jumlah kelompok yang terbentuk, n adalah banyaknya pengamatan, n_k adalah jumlah pengamatan dengan kategori ke- k , n_c adalah jumlah pengamatan pada kelompok ke- c , dan n_{kc} adalah jumlah pengamatan dengan kategori ke- k dan kelompok ke- c .

Data dan Sumber Data

Pada penulisan ini menggunakan data PODES yang diperoleh dari Badan Pusat Statistik (BPS) dengan satuan pengamatan adalah seluruh desa dan kelurahan di Provinsi Bengkulu pada tahun 2018 dengan total sebanyak 1514 desa/kelurahan. Daftar peubah disajikan pada Tabel 1 berikut.

Tabel 1. Daftar peubah penelitian berdasarkan pembagian dimensi IPD

Dimensi	Peubah	Keterangan
a. Pelayanan dasar	X1 : Jumlah keseluruhan sarana pendidikan	Numerik
	X2 : Jumlah keseluruhan sarana kesehatan	Numerik
	X3 : Jumlah keseluruhan tenaga kesehatan	Numerik
b. Kondisi infrastruktur	X4 : Jumlah sarana dan prasarana ekonomi	Numerik
	X13 : Bahan bakar memasak sebagian besar keluarga	Kategorik (6 kategori)*
	X5 : Jumlah keluarga pengguna listrik	Numerik
	X14 : Sumber air minum sebagian besar keluarga	Kategorik (10 kategori)*
	X15 : Sumber air mandi/cuci sebagian besar keluarga	Kategorik (8 kategori)*
	X16 : Kantor pos/pos pembantu/rumah pos	Kategorik (4 kategori)*
	X17 : Perusahaan/agen jasa ekspedisi swasta	Kategorik (4 kategori)*
	X18 : Fasilitas internet di kantor kepala desa	Kategorik (4 kategori)*
	X19 : Keberadaan warga yang menggunakan telepon seluler	Kategorik (3 kategori)*
c. Aksesibilitas/Transportasi	X20 : Keberadaan angkutan umum	Kategorik (3 kategori)*
	X21 : Jenis permukaan jalan darat antar desa yang terluas	Kategorik (4 kategori)*
	X6 : Waktu tempuh dari kantor kepala desa ke kantor camat	Numerik
	X7 : Biaya transportasi dari kantor kepala desa ke kantor camat	Numerik
	X8 : Waktu tempuh dari kantor kepala desa ke kantor bupati/walikota	Numerik
	X9 : Biaya transportasi dari kantor kepala desa ke kantor bupati/walikota	Numerik
d. Pelayanan Umum	X22 : Fasilitas/lapangan olahraga Sepak bola	Kategorik (4 kategori)*
	X23 : Fasilitas/lapangan olahraga bola voli	Kategorik (4 kategori)*
	X24 : Fasilitas/lapangan olahraga bulu tangkis	Kategorik (4 kategori)*
e. Penyelenggara Pemerintahan	X10 : Umur Kepala desa	Numerik
	X11 : Umur Sekretaris desa	Numerik
	X12 : Jumlah aparatur pemerintahan desa	Numerik

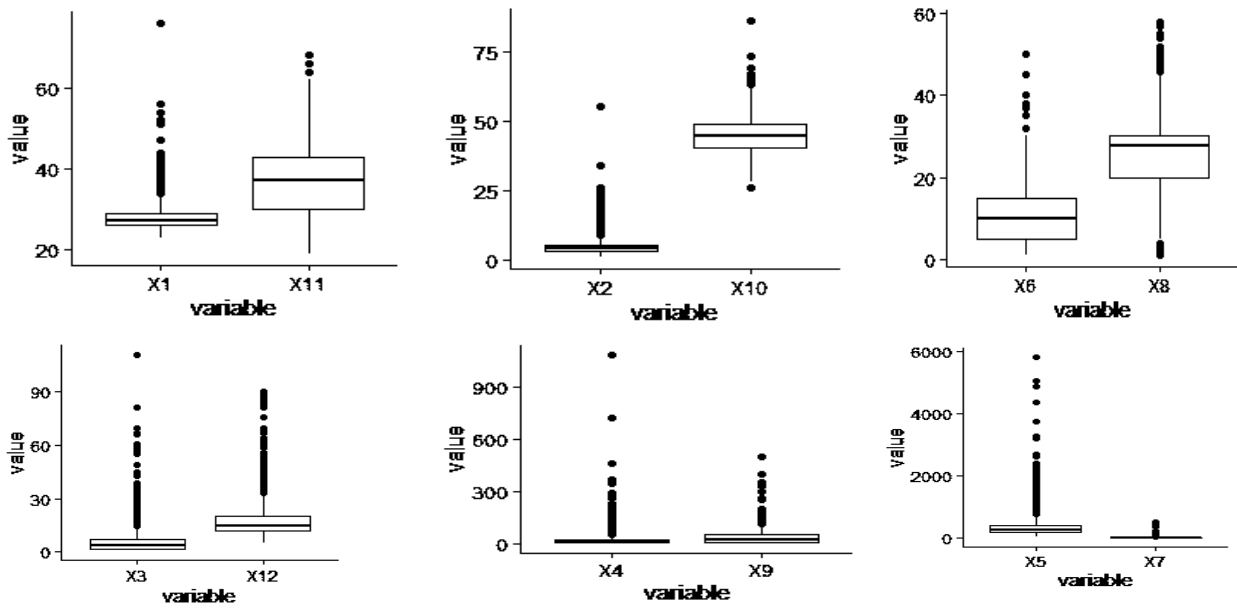
*Masing-masing kategori dapat dilihat pada Buku IPD 2018 yang diterbitkan oleh BPS

Prosedur analisis data terbagi menjadi dua yaitu memisahkan data menjadi dua sesuai dengan jenisnya kemudian diterapkan pada metode LCC. Berikut adalah langkah penggerombolan menggunakan metode LCC:

- a. Eksplorasi data
- b. Melakukan evaluasi kinerja metode
- c. Penggerombolan LCC dilakukan dengan tahapan:
 - 1) Definisikan banyaknya gerombol yang terbentuk.
 - 2) Menduga parameter *probabilitas posterior* dengan menggunakan metode EM.
 - 3) Melakukan pengelompokan objek berdasarkan nilai *probabilitas posterior* yang tertinggi.
 - 4) Mencari gerombol optimal
- d. Visualisasi gerombol optimal yang dihasilkan metode terbaik

HASIL DAN PEMBAHASAN

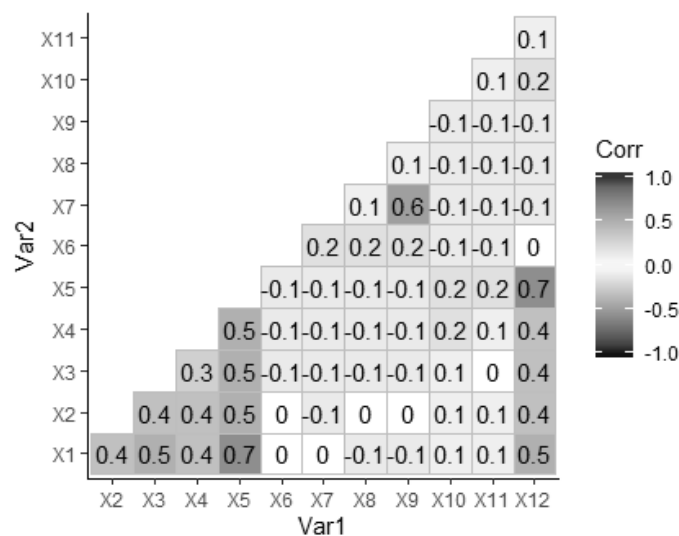
Peubah yang digunakan disesuaikan dengan IPD yang terdiri dari 12 peubah numerik dan 12 peubah kategorik. Dua jenis peubah ini diasumsikan mengikuti sebaran normal dan multinomial. Hal ini dikarenakan data yang ada memiliki karakteristik parameter yang dimiliki oleh kedua sebaran tersebut. Karakteristik pada peubah numerik disajikan dengan menggunakan diagram kotak garis (*boxplot*). Sedangkan untuk peubah kategorik disajikan dalam persentase masing-masing kategori yang muncul, serta ada beberapa kategori dalam peubah yang tidak muncul diakibatkan adanya desa yang tidak memuat kategori tersebut. Visualisasi sebaran nilai 12 peubah numerik disajikan pada Gambar 2 berikut.



Gambar 1. Karakteristik peubah numerik data PODES

Penyajian Gambar 2 dilakukan dengan menampilkan 2 peubah yang memiliki sebaran nilai yang hampir sama agar jelas terlihat pencilan yang ada dalam masing-masing peubah. Sebaran nilai tertinggi ada pada peubah X5 yang merupakan jumlah keluarga pengguna listrik di desa Provinsi Bengkulu. Semua peubah numerik memiliki pencilan atas. Pencilan atas ini merupakan nilai yang jauh lebih besar daripada nilai lainnya. Namun pada peubah X8 dan X10 ditemukan pencilan bawah artinya terdapat nilai yang jauh lebih kecil daripada nilai lainnya. Peubah X8 merupakan waktu tempuh dari kantor kepala desa ke kantor Bupati/Walikota dan X10 merupakan umur kepala desa. Berdasarkan pencilan bawah yang ada pada kedua peubah tersebut artinya terdapat kepala desa yang relatif lebih muda daripada kepala desa lainnya dan terdapat beberapa desa yang memiliki waktu tempuh lebih cepat menuju kantor Bupati/Walikota.

Korelasi antar peubah numerik pada data PODES disajikan pada Gambar 3 berikut. Nilai korelasi terbesar antar peubah numerik ditunjukkan oleh nilai 0.7. Pada data PODES ditemukan adanya peubah yang saling bebas ditandai dengan nilai korelasi 0 dan memiliki hubungan yang cenderung kecil antar peubah ditandai dengan nilai korelasi -0.1 sampai 0.4 . Berdasarkan identifikasi adanya korelasi di beberapa peubah sehingga diperlukan tindakan transformasi menggunakan analisis komponen utama agar pada saat diterapkan metode LCC tidak menyebabkan adanya objek pengamatan yang saling tumpang tindih ketika dilakukan penggerombolan.

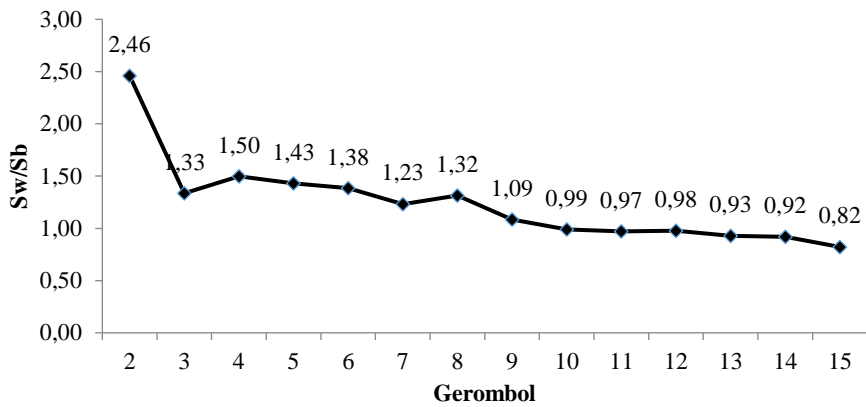


Gambar 2. Nilai korelasi peubah numerik data PODES

Karakteristik peubah kategorik diperoleh bahwa lebih dari 90% desa menggunakan LPG 3kg sebagai bahan bakar memasak dan sisanya menggunakan kayu bakar. Penggunaan air untuk minum, mandi atau mencuci terdapat lebih dari 50% desa menggunakan air sumur. Namun, melihat kondisi infrastruktur terkait keberadaan kantor pos dan agen jasa pengiriman swasta terlihat bahwa lebih dari 90% desa yang tidak memiliki infrastruktur tersebut.

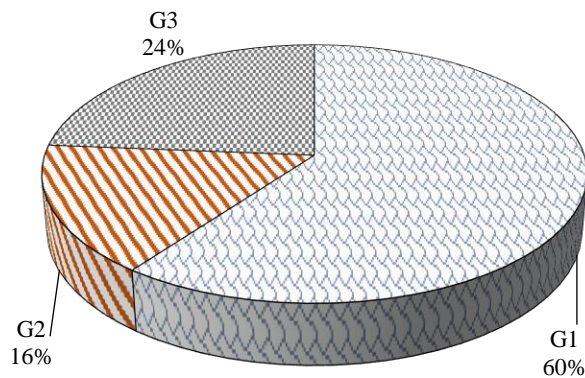
Hal ini berbanding terbalik dengan kondisi jalan yang sudah lebih dari 80% berbentuk aspal/beton. Ditemukan pula bahwa terdapat sekitar 60% tidak terdapat fasilitas internet di kantor desa/lurah. Lain halnya dengan kepemilikan telepon seluler bahwa terdapat lebih dari 90% desa sebagaimana besar warganya telah memiliki telepon seluler. Selanjutnya, dilihat dari aspek transportasi hanya terdapat sekitar 42% desa yang memiliki angkutan umum tanpa trayek tetap. Terkait dengan keberadaan pelayanan umum dari aspek fasilitas olahraga sepak bola, bola voli dan bulu tangkis terdapat lebih dari 50% desa memiliki fasilitas tersebut.

Penerapan penggerombolan desa berdasarkan data PODES menggunakan metode LCC diawali dengan melakukan transformasi skor komponen utama, mengingat bahwa satuan pada peubah numerik berbeda-beda. Selain itu juga dilakukan uji kenormalan data pada peubah numerik. Namun, karena pada uji tidak menunjukkan data menyebar normal maka pada penelitian ini menggunakan pendekatan grafik sebaran normal. Pembentukan objek-objek menjadi sebuah gerombol ditentukan berdasarkan nilai probabilitas posterior terbesar. Pemilihan gerombol optimal berdasarkan kestabilan nilai Sw/Sb yang dihasilkan banyaknya gerombol yang ditentukan yaitu 2 gerombol sampai 15 gerombol. Gambar 4 menampilkan hasil perhitungan nilai rata-ran Sw/Sb dari 12 peubah numerik dan 12 peubah kategorik.



Gambar 4 Nilai Sw/Sb penggerombolan LCC

Berdasarkan Gambar 4 diperoleh bahwa hasil gerombol optimal metode LCC adalah 3 gerombol. Persentase desa yang tersebar dalam 3 gerombol optimal yang dihasilkan LCC disajikan dalam Gambar 5. Desa yang termasuk dalam gerombol 1 (G1) ada sebanyak 60% atau ada sebanyak 911 desa yang ada dalam gerombol ini. Gerombol 2 (G2) ada sebanyak 244 desa atau 16%. Sedangkan gerombol 3 (G3), memuat sebanyak 359 desa atau sebesar 24%. Banyaknya desa pada masing-masing gerombol memberikan karakteristik yang berbeda sesuai dengan peubah yang digunakan.



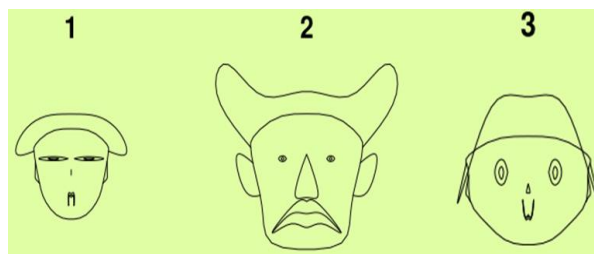
Gambar 5. Persentase banyaknya desa dalam masing-masing gerombol

Tabel 2. Deskripsi ciri wajah pada peubah numerik

Peubah	Ciri Wajah	Keterangan
X1	<i>height of face/ width of nose</i>	Semakin tinggi bentuk wajah/semakin lebar bentuk hidung maka semakin banyak sarana pendidikan
X2	<i>width of face / width of ear</i>	Semakin lebar bentuk wajah/semakin lebar bentuk telinga maka semakin banyak sarana kesehatan
X3	<i>structure of face/ height of ear</i>	Semakin tidak teratur bentuk wajah/semakin tinggi bentuk telinga maka semakin banyak tenaga kesehatan
X4	<i>height of mouth</i>	Semakin tinggi bentuk mulut maka semakin banyak sarana dan prasarana ekonomi
X5	<i>width of mouth</i>	semakin lebar bentuk mulut semakin banyak pengguna listrik
X6	<i>smiling</i>	semakin melengkung keatas bentuk mulut semakin besar waktu tempuh dari kantor kepala desa ke kantor camat
X7	<i>height of eyes</i>	semakin tinggi bentuk mata maka semakin besar biaya transportasi ke kantor camat
X8	<i>width of eyes</i>	semakin lebar bentuk mata maka semakin tinggi waktu tempuh ke kantor bupati/walikota
X9	<i>height of hair</i>	semakin tinggi bentuk rambut maka semakin besar biaya transportasi ke kantor bupati/walikota
X10	<i>width of hair</i>	semakin lebar bentuk rambut maka semakin tinggi umur kepala desa
X11	<i>style of hair</i>	semakin melengkung keatas bentuk rambut maka semakin tinggi umur kades
X12	<i>height of nose</i>	semakin tinggi bentuk hidung maka semakin banyak aparatur pemerintahan desa

Visualisasi peubah numerik untuk masing-masing gerombol disajikan menggunakan *Chernoff face* dengan terlebih dahulu menghitung nilai rata-rata pada masing-masing peubah dalam gerombol. Pembuatan *Chernoff face* menggunakan package *aplpack* pada program R dengan menggunakan fungsi *faces*. Deskripsi *Chernoff face* masing-masing peubah diberikan pada Tabel 3.

Gambar 6 menampilkan karakteristik gerombol optimal yang dihasilkan metode LCC berdasarkan struktur wajah. Pembentukan *Chernoff face* berdasarkan rata-rata peubah numerik masing-masing gerombol dengan objek pengamatan desa yang masuk dalam ketiga gerombol yang berbeda-beda. Berdasarkan yang diperoleh Gerombol 2 menunjukkan nilai rata-rata yang jauh lebih besar dari pada gerombol 1 dan 3. Kecuali pada bentuk mata dan bentuk senyum, yang menunjukkan waktu dan biaya tempuh menuju kantor camat dan bupati/walikota. Artinya pada gerombol 2 menunjukkan karakteristik desa yang lebih unggul sesuai dengan peubah penelitian yang digunakan. Pada gerombol 1 dan 3 yang sangat membedakan adalah gerombol 3 menunjukkan waktu tempuh dan biaya dari kantor kepala desa ke kantor camat, serta biaya transportasi ke kantor bupati/walikota dan yang paling besar. Selain itu, struktur wajah antara gerombol 3 cenderung lebih besar daripada gerombol 1. Hal ini menandakan bahwa karakteristik pada gerombol 3 cenderung lebih baik daripada gerombol satu pada beberapa peubah. Namun, mengingat bahwa biaya transportasi menuju camat dan walikota lebih besar dari gerombol 1, maka antara gerombol 1 dan 3 tidak dapat dikatakan saling mengungguli.



Gambar 6. *Chernoff face* gerombol optimal metode LCC

Visualisasi peubah kategorik tidak dapat dilakukan, karena belum tersedia penyajian gerombol yang berasal dari peubah kategorik. Oleh karena itu, penyajian karakteristik peubah kategorik dilakukan dengan menggunakan tabel.

Karakteristik gerombol optimal untuk peubah kategorik disajikan pada Tabel 3 berdasarkan persentase kategori yang muncul dalam gerombol. Gerombol 1 memuat persentase kemunculan kategori yang cenderung lebih besar daripada gerombol 2 dan 3. Keberadaan angkutan umum, kantor pos, fasilitas internet di kantor

desa dan ketersediaan beberapa fasilitas olahraga pada gerombol 1 persentase yang paling besar muncul pada kategori tidak ada. Hal ini menunjukkan bahwa, desa yang berada dalam gerombol 2 dan 3 memuat desa dengan kategori yang lebih baik daripada gerombol 1. Namun, antara kedua gerombol tersebut, gerombol 2 menunjukkan persentase kategori tidak ada fasilitas infrastruktur lebih kecil daripada gerombol 3.

Tabel 3. Persentase desa di Provinsi Bengkulu menurut peubah kategorik pada gerombol optimal

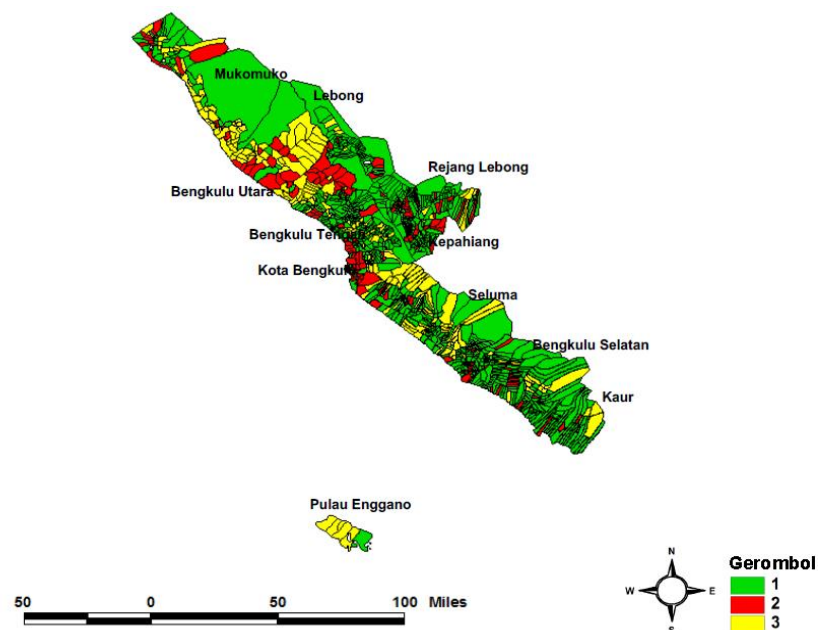
Peubah Kategorik	Kategori	Persentase Gerombol		
		1	2	3
X13: Bahan Bakar Memasak	1 Gas Kota	55.68	15.92	20.94
	2 LPG 3kg			
	3 LPG >3kg			
	4 Minyak tanah			
	5 Kayu bakar			
	6 Lainnya			
X14 : Sumber air minum sebagian besar keluarga	1 Air Kemasan bermerk	2.71	2.91	1.78
	2 Air isi ulang	8.19	3.43	0.46
	3 Ledeng dengan meteran (PAM/PDAM)			
	4 Ledeng tanpa meteran	0.13		
	5 Sumur bor tanpa pompa	5.15	2.64	2.25
	6 Sumur	36.53	6.54	16.71
	7 Mata air	5.48	0.46	1.72
	8 Sungai/danau/kolam/waduk /situ/embung/bendungan	1.12		0.53
	9 Air hujan	0.99		0.26
	10 Lainnya			
X15 : Sumber air mandi/cuci sebagian besar keluarga	1 Ledeng dengan meteran (PAM/PDAM)	7.79	4.76	0.59
	2 Ledeng tanpa meteran	0.99	0.26	0.20
	3 Sumur bor tanpa pompa	5.42	2.84	2.25
	4 Sumur	35.87	7.40	16.25
	5 Mata air	5.42	0.53	1.85
	6 Sungai/danau/kolam/waduk /situ/embung/bendungan	4.56	0.26	2.44
	7 Air hujan	0.13	0.07	0.13
	8 Lainnya			
X16 : Kantor pos/pos pembantu/rumah pos	1 Beroperasi	2.44	3.24	0.13
	2 Jarang beroperasi	0.26		0.13
	3 Tidak beroperasi	0.20	0.13	0.20
	4 Tidak ada	57.27	12.75	23.25
X17 : Perusahaan/agen jasa ekspedisi swasta	1 Beroperasi	2.18	4.16	0.20
	2 Jarang beroperasi	0.40	0.13	0.46
	3 Tidak beroperasi	0.13	0.07	0.13
	4 Tidak ada	57.46	11.76	22.92
X18 : Fasilitas internet di kantor kepala desa/lurah	1 Berfungsi	17.50	7.27	9.31
	2 Jarang berfungsi	1.59	0.33	1.06
	3 Tidak berfungsi	1.19	0.73	0.92
	4 Tidak ada	39.89	7.79	12.42
X19 : Keberadaan warga pengguna handphone	1 Sebagian besar warga	57.46	15.59	21.60
	2 Sebagian kecil warga	1.78	0.33	1.85
	3 Tidak ada	0.92	0.20	0.26
X20 : Keberadaan angkutan umum	1 Ada, dengan trayek tetap	13.08	6.74	2.31
	2 Ada, tanpa trayek tetap	30.45	7.46	4.29
	3 Tidak ada angkutan umum	16.64	1.92	17.11
X21 : Jenis permukaan jalan antar desa/kel. yang terluas	1 Aspal/beton	55.61	14.73	11.16
	2 diperkeras (kerikil,batu,dll)	4.23	1.39	9.91
	3 Tanah	0.20		2.11
	4 lainnya	0.13		0.53
X22 : Fasilitas/lapangan olahraga Sepak bola	1 Ada, baik	13.94	6.01	9.11
	2 Ada, rusak parah	4.82	1.12	2.58
	3 Ada, rusak sedang	1.19	0.07	0.92
	4 Tidak ada	40.22	8.92	11.10

Peubah Kategorik	Kategori	Persentase Gerombol			
		1	2	3	
X23 :	1	Ada, baik	28.67	9.71	16.45
Fasilitas/lapangan	2	Ada, rusak parah	7.86	1.65	3.17
olahraga bola voli	3	Ada, rusak sedang	1.98	0.20	0.86
	4	Tidak ada	21.66	4.56	3.24
X24 :	1	Ada, baik	10.70	6.01	6.61
Fasilitas/lapangan	2	Ada, rusak parah	4.29	0.86	1.72
olahraga bulu	3	Ada, rusak sedang	0.59		0.40
tangkis	4	Tidak ada	44.58	9.25	14.99

Tampilan sebaran desa pada gerombol optimal yang diperoleh disajikan dalam bentuk peta tematik (Gambar 7). Area pada peta didominasi oleh warna hijau yang menunjukkan gerombol 1 artinya jika ditinjau dari karakteristik peubah kategorik maka menunjukkan bahwa kebanyakan desa di Provinsi Bengkulu yang tidak memiliki fasilitas infrastruktur. Gerombol 2 yang di definisikan sebagai warna merah paling banyak berada di kabupaten Bengkulu Utara, serta terlihat pula di hampir semua wilayah Kota Bengkulu. Hal ini menunjukkan bahwa, berdasarkan dimensi IPD sesuai peubah penelitian yang digunakan termasuk pada desa yang memiliki kondisi yang lebih baik daripada desa lainnya yang berada pada gerombol 1 dan 3. Adapun gerombol 3 yang memiliki karakteristik yang hampir mirip dengan gerombol 1 juga memiliki sebaran di semua kabupaten. Sebaran desa yang masuk dalam gerombol optimal disajikan dalam Tabel 4 berikut.

Tabel 3. Persentase desa di Provinsi Bengkulu menurut peubah kategorik pada gerombol optimal

No	Kabupaten	Banyaknya desa dalam gerombol		
		1	2	3
1	Bengkulu	10	56	1
2	Bengkulu Selatan	129	24	5
3	Bengkulu Tengah	87	10	46
4	Bengkulu Utara	96	44	80
5	Kaur	146	22	27
6	Kepahiang	95	14	8
7	Lebong	90	7	7
8	Mukomuko	46	21	85
9	Rejang Lebong	109	37	10
10	Seluma	103	9	90



Gambar 7. Peta tematik level Desa di Provinsi Bengkulu berdasarkan gerombol optimal

Berdasarkan Gambar 7 diperoleh bahwa gerombol 1 yang ditandai dengan *legend* warna hijau menunjukkan sebaran desa yang memiliki karakteristik persentase ketersediaan sarana dan prasarana yang kurang memadai jika dilihat berdasarkan peubah kategorik dibandingkan gerombol 2 dan 3, namun jika dilihat berdasarkan peubah numerik desa yang terdapat dalam gerombol 2 yang ditunjukkan dengan *legend* warna merah memiliki pembangunan yang lebih baik dari segi pendidikan, sarana kesehatan, jumlah aparatur desa dan sebagainya dibandingkan desa yang ada di gerombol 1 dan 3. Gerombol 3 yang ditunjukkan dengan *legend* warna kuning termasuk dalam desa cukup memadai dengan ketersediaan sarana dan prasarana yang dibutuhkan desa jika dilihat berdasarkan peubah numerik dan kategorik.

KESIMPULAN

Hasil penggerombolan menggunakan metode LCC diperoleh gerombol optimal sebanyak 3 gerombol, dengan masing-masing karakteristik yang berbeda-beda tiap gerombol. Gerombol terbaik ditunjukkan pada gerombol 2 yang memberikan IPD yang relatif lebih tinggi daripada gerombol 1 dan 3. Gerombol 1 memuat persentase kemunculan kategori yang cenderung lebih besar daripada gerombol 2 dan 3. Keberadaan angkutan umum, kantor pos, fasilitas internet di kantor desa dan ketersediaan beberapa fasilitas olahraga pada gerombol 1 persentase yang paling besar muncul pada kategori tidak ada. Hal ini menunjukkan bahwa, desa yang berada dalam gerombol 2 dan 3 memuat desa dengan kategori yang lebih baik daripada gerombol 1. Namun, antara kedua gerombol tersebut, gerombol 2 menunjukkan persentase kategori tidak ada fasilitas infrastruktur lebih kecil daripada gerombol 3. Apabila dilihat berdasarkan peubah numerik gerombol 2 menunjukkan karakteristik desa yang lebih unggul sesuai dengan peubah penelitian yang digunakan. Pada gerombol 1 dan 3 yang sangat membedakan adalah gerombol 3 menunjukkan waktu tempuh dan biaya dari kantor kepala desa ke kantor camat, serta biaya transportasi ke kantor bupati/walikota dan yang paling besar. Selain itu, struktur wajah antara gerombol 3 cenderung lebih besar daripada gerombol 1. Hal ini menandakan bahwa karakteristik pada gerombol 3 cenderung lebih baik daripada gerombol satu pada beberapa peubah. Namun, mengingat bahwa biaya transportasi menuju camat dan walikota lebih besar dari gerombol 1, maka antara gerombol 1 dan 3 tidak dapat dikatakan saling mengungguli.

DAFTAR PUSTAKA

- Badan Pusat Statistik.(2018).Podes Provinsi Bengkulu. Jakarta (ID): BPS
- Badan Pusat Statistik.(2019).Indeks Pembangunan Desa 2018. Jakarta (ID): BPS
- Bunkers MJ, Miller JR, DeGaetano AT. (1996). Definition of climate regions in the northern plains using an objective cluster modification technique. *Journal of Climate* 9(1):130–146.
- Fachriyatul. (2014). Analisis Two Step Cluster (TSC) dan Analisis Latent Class Cluster (LCC) pada Pengelompokan Data Berskala Campuran Kategorik dan Kontinu. Malang: Universitas Brawijaya
- Hermawati F. (2018). Perbandingan *Cluster ensemble*, *Two Step Cluster* Dan Metode Gower Pada Penggerombolan Data Bertipe Campuran [Tesis]. Bogor (Id): Institut Pertanian Bogor
- Mattjik AA, Sumertajaya IM. 2011. *Sidik Peubah Ganda dengan Menggunakan SAS*. Bogor (ID): IPB Press
- Munthe AD. 2019. Penggerombolan Desa/Kelurahan Di Provinsi Nusa Tenggara Timur Menggunakan Algoritma *Two Step Cluster* (TSC) Dan Algoritma *K-Prototypes*[Tesis].Bogor (ID): Institut Pertanian Bogor
- Vermunt, J.K., dan Magidson, J. (2002). *Latent Class Cluster Analysis*. Dalam Hagenars, J.A., dan McCutcheon, A.L.,(edited), *Applied Latent Class Analysis*, 89-106. Cambridge University Press.