

Artikel: *Akses terbuka/Open access*

**Identifikasi Karakteristik Desa/Kelurahan Di Provinsi Nusa Tenggara Timur
Menggunakan Analisis Gerombol**

Sitasi: Munthe, 2023, JSTAR 3(1), 50-62

Kronologi naskah.

Submit : 17 April 2023

Revisi : 30 Mei 2023

Diterima : 30 Mei 2023



Penyedia Data Statistik Berkualitas
untuk Indonesia Maju

Reformasi Birokrasi



IDENTIFIKASI KARAKTERISTIK DESA/KELURAHAN DI PROVINSI NUSA TENGGARA TIMUR MENGGUNAKAN ANALISIS GEROMBOL

Andrew Donda Munthe¹

¹Statistisi Ahli Muda, Badan Pusat Statistik Provinsi NTT, Indonesia

#korespondensi *author*: andrew_donda@bps.go.id

Abstract

Grouping of villages by clustering analysis can result in the identification to the appropriate policy programs to each cluster formed. The clustering algorithm that can be used on mixed variables of village grouping is the Gower Method (Gower's distance). The purpose of this study is to apply the Gower Method to village clustering in Nusa Tenggara Timur Province based on 13 mixed variables. The data used in this study was sourced from Village Potential Data Collection (PODES) in 2021. The results showed that the optimal cluster that formed from the application of the Gower Method was 3 clusters. The first cluster consisted of 1.063 villages, the second cluster consisted of 870 villages and the third cluster consisted of 1.517 villages. Based on the visualization of the results of the cluster, the area with the highest tendency to development achievement is the village in the third cluster, while the lowest is in the villages of the first cluster members.

Keyword: *clustering, mixed variables, Gower Method, villages*

1. Pendahuluan

Isu kemiskinan merupakan persoalan multidimensi yang hingga saat ini masih dihadapi oleh berbagai negara dunia. Pertemuan 193 negara anggota Perserikatan Bangsa-Bangsa (PBB) di New York pada tahun 2015 bahkan menghasilkan kesepakatan bersama untuk menjalankan pembangunan berkelanjutan hingga tahun 2030. Indonesia dan negara-negara anggota PBB lainnya pada saat itu menyepakati 17 tujuan utama pembangunan berkelanjutan atau sering disebut dengan *Sustainable Development Goals (SDGs)*.

Pengentasan Kemiskinan menjadi tujuan pertama dalam SDGs yaitu “*End poverty in all its forms everywhere*” atau mengakhiri kemiskinan dalam segala bentuk di mana pun. Pemerintah Indonesia juga telah menuangkan permasalahan multidimensi kemiskinan ke dalam Rencana Pembangunan Jangka Panjang Nasional (RPJPN) 2005-2025. Kerangka multidimensi kemiskinan bukan hanya menyangkut ukuran pendapatan tetapi juga berkaitan dengan kerentanan seseorang menjadi miskin serta ada/tidaknya pemenuhan hak dasar dalam menjalani kehidupan yang bermartabat.

Provinsi Nusa Tenggara Timur

menjadi salah satu provinsi di Indonesia dengan tingkat kemiskinan yang tinggi di Indonesia. Persentase penduduk miskin di Provinsi Nusa Tenggara Timur pada bulan September 2022 mencapai 20,23 persen atau sekitar seperlima dari total penduduk (BPS, 2023). Tingkat kemiskinan di Provinsi NTT pada September 2022 berada pada peringkat tertinggi ketiga di Indonesia, dibawah Provinsi Papua (26,80 persen) dan Provinsi Papua Barat (21,43 persen).

Penduduk miskin yang ada di Provinsi Nusa Tenggara Timur lebih dominan berada di daerah perdesaan. Penduduk miskin yang berada di daerah perdesaan mencapai 24,11 persen sedangkan penduduk miskin di daerah perkotaan hanya sekitar 9,0 persen.

Data diatas menunjukkan ketimpangan pembangunan wilayah perdesaan dan perkotaan. Oleh karena itu, perencanaan berbagai program pembangunan dalam rangka menekan angka kemiskinan perlu memperhatikan aspek wilayah. Fokusnya terutama pada wilayah administrasi di tingkat desa/kelurahan.

Pengelompokkan desa/kelurahan yang dilakukan melalui analisis statistik dapat menghasilkan identifikasi berbagai karakteristik wilayah. Hasil pengelompokkan dapat digunakan untuk merencanakan berbagai kebutuhan program kebijakan yang tepat dari kelompok atau gerombol yang terbentuk.

Analisis statistik yang biasa digunakan untuk mengelompokkan

objek adalah analisis gerombol (*cluster analysis*). Objek-objek yang diteliti dikelompokkan berdasarkan atas kemiripan ataupun ketidakmiripan antar objeknya. Terdapat 2 metode yang digunakan dalam analisis ini yaitu metode hierarki dan metode tidak berhierarki. Metode hierarki (*hierarchial clustering*) dilakukan apabila di awal penerapan analisis, jumlah gerombol yang akan terbentuk belum diketahui. Perbedaannya dengan metode tidak berhierarki (*non- hierarchial clustering*), jumlah gerombol yang akan terbentuk sudah ditetapkan terlebih dahulu sebelum proses pengelompokkan dilakukan lebih lanjut.

Secara umum, analisis gerombol dilakukan pada objek dengan amatan tidak terlalu besar dan menggunakan 1 jenis peubah saja (peubah numerik saja atau peubah kategorik saja). Permasalahan lain pada analisis gerombol adalah keberadaan pencilan (*outlier*) yang sangat mempengaruhi kualitas dari hasil akhir kelompok yang terbentuk.

Terdapat beberapa metode yang dapat digunakan untuk melakukan pengelompokkan pada data dengan peubah campuran. Metode tersebut adalah algoritma *Two Step Cluster (TSC)*, *Cluster Ensemble*, *K-Prototypes*, dan juga metode Gower.

Kelebihan metode Gower dibandingkan dengan metode lain adalah pada proses pengolahannya. Metode Gower proses pengolahannya tidak terlalu kompleks sehingga lebih

mudah untuk diterapkan (Gower, 1971).

Berdasarkan uraian diatas maka tujuan dari penelitian ini adalah untuk mengaplikasikan Analisis gerombol dengan Metode Gower untuk identifikasi karakteristik desa/kelurahan di Provinsi Nusa Tenggara Timur berdasarkan data peubah campuran.

2. Metodologi

Analisis gerombol dengan penerapan Metode Gower pada penelitian ini menggunakan peubah campuran (kombinasi peubah numerik dan peubah kategorik) yang berkaitan dengan indikator kemiskinan. Total peubah yang digunakan adalah sebanyak 13 peubah (7 peubah numerik dan 6 peubah kategorik).

Bahan dan Sumber Data

Data yang digunakan dalam penelitian ini berasal dari hasil pendataan Potensi Desa (Podes) Tahun 2021 yang bersumber dari Badan Pusat Statistik (BPS). Objek penelitian adalah seluruh desa/kelurahan di Provinsi Nusa Tenggara Timur (kondisi tahun 2021) dengan total 3.450 desa/kelurahan.

Peubah yang digunakan dalam penelitian ini adalah sebanyak 13 peubah yang didasarkan pada penelitian sebelumnya yang dilakukan oleh Munthe (2018) dan juga mempertimbangkan ketersediaan data hasil pendataan Potensi Desa (Podes) Tahun 2021.

Peubah-peubah yang digunakan pada penelitian ini terbagi menjadi 4 dimensi yaitu:

1. Dimensi Pelayanan dasar (terdiri dari 3 peubah numerik);
2. Dimensi pelayanan umum (terdiri dari 2 peubah numerik);
3. Dimensi aksesibilitas dan transportasi (terdiri dari 1 peubah numerik dan 1 peubah kategorik);
4. Dimensi kondisi infrastruktur energi, komunikasi, air bersih dan sanitasi (terdiri dari 1 peubah numerik dan 5 peubah kategorik).

Tabel 1. Peubah Penelitian

Peubah	Nama Peubah	Tipe Peubah
Dimensi 1 (Pelayanan dasar)		
X1	Rasio sarana kesehatan per 1.000 penduduk	Numerik
X2	Rasio tenaga kesehatan per 1.000 penduduk	Numerik
X3	Rasio sarana pendidikan per 1.000 penduduk	Numerik
Dimensi 2 (Pelayanan Umum)		
X4	Jumlah surat miskin atau Surat Keterangan Tidak Mampu (SKTM) yang dikeluarkan desa/kelurahan	Numerik
X5	Jarak kantor desa/kelurahan ke kantor kecamatan	Numerik
Dimensi 3 (Aksesibilitas dan Transportasi)		
X6	Jarak kantor desa/kelurahan ke kantor bupati/walikota	Numerik

Tabel 1. Peubah Penelitian (Lanjutan)

Peubah	Nama Peubah	Tipe Peubah
X7	Keberadaan angkutan umum yang melewati desa/kelurahan	Kategorik
Dimensi 4 (infrastruktur energi, komunikasi, air bersih dan sanitasi)		
X8	Persentase keluarga pengguna listrik	Numerik
X9	topografi wilayah desa/kelurahan	Kategorik
X10	Fasilitas tempat buang air besar sebagian besar keluarga	Kategorik
X11	sumber air minum sebagian besar keluarga	Kategorik
X12	Bahan Bakar untuk memasak sebagian besar keluarga	Kategorik
X13	Ketersediaan dan kualitas sinyal komunikasi seluler	Kategorik

Metode Analisis Data

Metode analisis data yang digunakan adalah analisis gerombol dengan menggunakan ukuran jarak Gower (*Gower's Distance*). Pengolahan data menggunakan software statistik Rstudio versi 2022.07.1+554.

Langkah-langkah analisis data yang dilakukan pada penelitian ini, yaitu:

1. Melakukan eksplorasi data untuk mengetahui gambaran data serta keberadaan pencilan (*outlier*) terutama pada peubah numerik.
2. Melakukan perhitungan ukuran

kemiripan Gower (*Gower's Distance*) dengan rumus:

$$d_{ij} = \frac{\sum_{k=1}^p \delta_{ijk} d_{ijk}}{\sum_{k=1}^p \delta_{ijp}} \quad (1)$$

dimana:

d_{ij} = ukuran jarak antara objek ke-i dan objek ke-j

k = banyaknya peubah dalam set data ($k=1,2,\dots,p$)

δ_{ijp} = 0, jika x_{ik} dan x_{jk} pada peubah ke-k memiliki data hilang atau peubah biner asimetris pada kedua objek adalah nol ($x_{ik} = x_{jk} = 0$)

δ_{ijp} = 1, jika lainnya

d_{ijk} = jarak antara objek ke-i dan objek ke-j pada peubah ke-k

Nilai jarak antara objek ke-i dan objek ke-j pada peubah ke-k (d_{ijk}) untuk setiap tipe data peubah dapat didefinisikan sebagai berikut:

- a. Jika peubah k merupakan peubah dengan tipe data biner atau nominal, maka:
 - $d_{ijk}=1$, jika $x_{ik} \neq x_{jk}$
 - $d_{ijk}=0$, jika $x_{ik} = x_{jk}$
- b. Jika peubah k merupakan peubah dengan tipe data berskala interval, maka:

$$d_{ijk} = \frac{|x_{ik} - x_{jk}|}{R_k} \quad (2)$$

dengan $R_k = \max(x_{lk}) - \min(x_{lk})$ dan l adalah semua objek pada peubah ke-k.

- c. Jika peubah k merupakan

peubah dengan tipe data berskala ordinal, maka d_{ijk} dihitung menggunakan Persamaan (1), dimana x_{ik} dan x_{jk} merupakan nilai peringkat pada data.

d. Jika peubah k merupakan peubah dengan tipe data rasio maka d_{ijk} dihitung juga dengan menggunakan Persamaan (1).

3. Selanjutnya menentukan banyak gerombol optimal berdasarkan matriks jarak (*Gower's Distance*) pada Poin 2. Metode yang digunakan adalah dengan menggunakan perhitungan koefisien *Silhouette* (*Silhouette Method*). Rumus perhitungannya adalah sebagai berikut:

$$s(i) = \frac{b(i) - a(i)}{\max \{a(i), b(i)\}}$$

dimana:

$s(i)$ = Koefisien *Silhouette*

$a(i)$ = Rata-rata jarak antar anggota dalam cluster

$b(i)$ = Jarak terkecil antara anggota cluster terdekatnya (*nearest neighbor*)

Nilai koefisien *Silhouette* adalah antara 0 – 1. Kriteria ketepatan dan kualitas hasil gerombol optimal semakin baik apabila nilainya mendekati 1 (Kauffman dan Rousseuw, 1990)

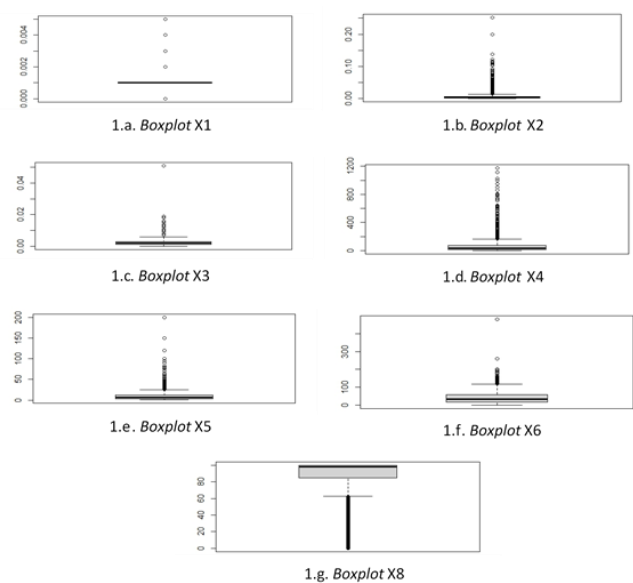
4. Setelah diperoleh gerombol optimal, langkah terakhir adalah melakukan visualisasi dan interpretasi hasil dari analisis gerombol. Untuk peubah

numerik visualisasi dilakukan dengan menggunakan *Chernoff Face* sedangkan peubah kategorik menggunakan diagram batang.

3. Hasil dan Pembahasan

Karakteristik dari peubah-peubah numerik pada penelitian ini dapat diketahui melalui diagram kotak garis (*boxplot*) seperti ditunjukkan pada Gambar 1.

Secara visual dapat terlihat bahwa X1, X2, X3, X4, X5, dan X6 nilainya menyebar pada nilai-nilai yang rendah. Sedangkan 1 peubah lainnya yaitu X8, nilai-nilainya menyebar pada nilai yang tinggi (Gambar 1.g).



Gambar 1. *Boxplot* Peubah-peubah Numerik

Semua peubah juga memiliki pencilan (*outlier*) yaitu nilai yang memiliki perbedaan besar dibandingkan dengan nilai lainnya. Keberadaan

pencilan pada masing-masing peubah disebut dengan *univariate outlier*. Keberadaan pencilan pada setiap peubah yang diteliti memberikan dampak negatif terhadap kualitas pengelompokan yang dihasilkan melalui analisis gerombol.

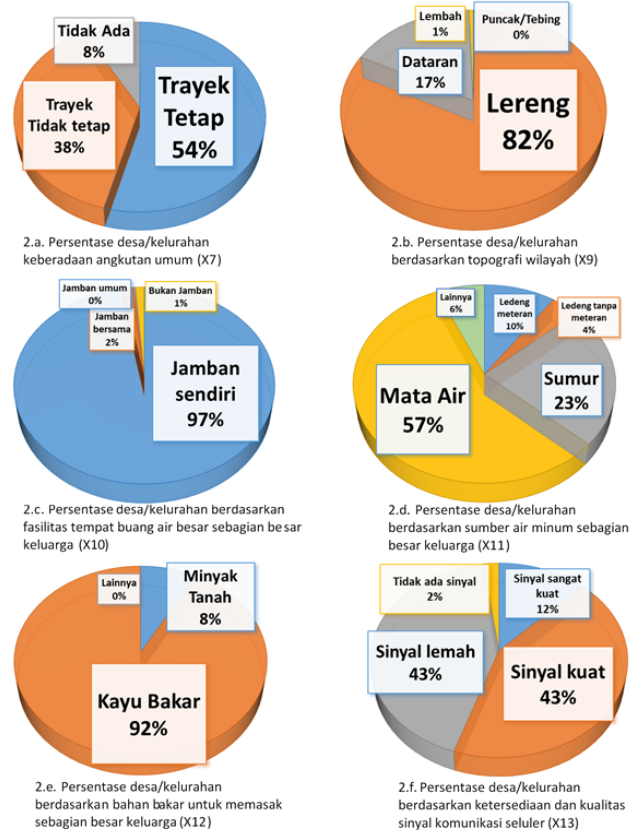
Gambaran data peubah kategorik pada penelitian seperti ditunjukkan pada Gambar 2. Persentase keberadaan angkutan umum yang melewati desa/kelurahan di Provinsi Nusa Tenggara Timur sudah relatif tinggi. Sebesar 54 persen desa/kelurahan dari total 3.450 desa/kelurahan di NTT telah memiliki sarana angkutan umum dengan trayek tetap. Akan tetapi, perlu dicermati bahwa masih ada desa di Provinsi NTT yang tidak ada angkutan umum sama sekali dengan persentase mencapai 8 persen (Gambar 2.a)

Topografi di sebagian besar wilayah desa/kelurahan di Provinsi NTT adalah lereng dengan persentase mencapai 82 persen (Gambar 2.b). Penggunaan jamban sendiri sebagai tempat buang air besar sebagian besar keluarga di Provinsi NTT telah mencapai 97 persen (Gambar 2.c). Untuk pemenuhan kebutuhan air minum, sebagian besar warga desa/kelurahan di Provinsi NTT memanfaatkan sumber mata air dengan persentase mencapai 57 persen desa/kelurahan (Gambar 2.d).

Kayu bakar merupakan bahan bakar yang digunakan untuk memasak bagi sebagian besar masyarakat desa/kelurahan di Provinsi NTT dengan persentase mencapai 92 persen

(Gambar 2.e).

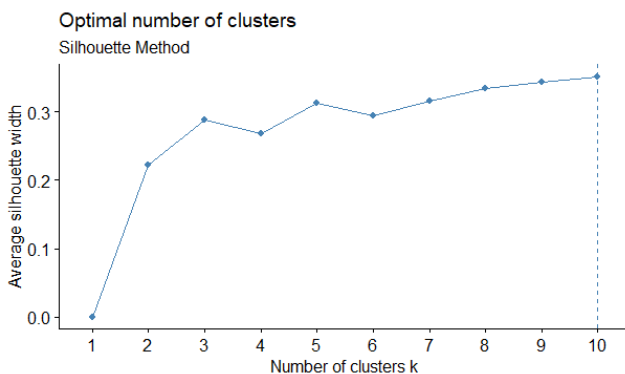
Peubah kategorik terakhir dalam penelitian ini adalah berkaitan dengan ketersediaan dan kualitas sinyal komunikasi seluler di desa/kelurahan. Meskipun secara umum sebagian besar wilayah sudah tersedia kualitas sinyal yang sangat kuat (12 persen) maupun kuat (43 persen), tetapi yang sinyalnya lemah juga masih relatif tinggi dengan persentase sebesar 43 persen desa/kelurahan. Bahkan masih ada wilayah-wilayah desa yang tidak memperoleh sinyal sama sekali dengan persentase sebesar 2 persen (Gambar 2.f).



Gambar 2. *Pie Chart* Peubah-peubah kategorik

Seluruh peubah dalam penelitian ini selanjutnya akan diolah menggunakan matriks jarak Gower (*Gower's Distance*). Peubah-peubah numerik memiliki skala data yang berbeda-beda sehingga perlu dilakukan standarisasi ke bentuk *Z score* sebelum pengolahan data dapat dilanjutkan.

Hasil pengolahan data berdasarkan perhitungan ukuran jarak Gower selanjutnya digunakan untuk memperoleh gerombol optimal. Penentuan gerombol optimal berdasarkan pada perhitungan koefisien *Silhouette* (*Silhouette Method*) seperti ditunjukkan pada Gambar 3.

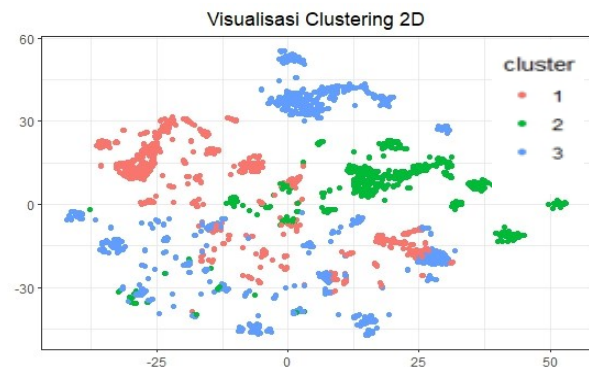


Gambar 3. Koefisien *Silhouette* pembentukan gerombol optimal

Gambar diatas menunjukkan bahwa berdasarkan nilai koefisien *Silhouette*, gerombol optimal yang dihasilkan analisis gerombol metode Gower ada pada jumlah 10 gerombol. Akan tetapi, apabila dicermati perubahan kenaikan koefisien *Silhouette* yang signifikan ada pada jumlah 3 gerombol. Selanjutnya pada jumlah 4 gerombol koefisien *Silhouette* mengalami penurunan. Selanjutnya

koefisien *Silhouette* terus mengalami peningkatan hingga jumlah 10 gerombol. Akan tetapi peningkatan nilai koefisien *Silhouette* pada jumlah 10 gerombol tidak berbeda jauh dengan jumlah 3 gerombol. Oleh karena itu, peneliti menetapkan gerombol optimal pada penelitian ini adalah pada jumlah 3 gerombol.

Visualisasi dalam bentuk 2 dimensi pengelompokan desa/kelurahan di Provinsi NTT berdasarkan 3 gerombol optimal seperti ditunjukkan pada Gambar 4. Gerombol 1 ditunjukkan dengan warna coklat, gerombol 2 warna hijau dan gerombol 3 warna biru.



Gambar 4. Visualisasi 3 gerombol optimal dalam bentuk 2 dimensi

Masing-masing objek penelitian yaitu desa/kelurahan di Provinsi NTT kemudian dikelompokkan kedalam 3 gerombol optimal tersebut. Sebanyak 1.063 desa/kelurahan merupakan anggota pada gerombol 1 atau mencapai 30,81 persen dari total desa/kelurahan yang ada di NTT hasil pendataan Potensi Desa (Podes) Tahun 2021.

Gerombol 2 merupakan gerombol dengan ukuran yang terkecil dari 2

gerombol lainnya. Gerombol 2 hanya terdiri dari 870 desa/kelurahan atau hanya 25,22 persen dari total desa/kelurahan.

Gerombol 3 yang terdiri dari 1.517 desa/kelurahan atau dengan persentase sebesar 43,97 persen merupakan gerombol dengan ukuran terbesar dibandingkan dengan gerombol lain. Sebaran desa/kelurahan pada setiap gerombol per Kabupaten/Kota seperti ditunjukkan pada Tabel 2.

Tabel 2. Distribusi Desa/Kelurahan per Kab/ Kota berdasarkan 3 gerombol optimal

No.	Nama Kabupaten/Kota	Kluster			Total
		1	2	3	
1	Sumba Barat	46	7	21	74
2	Sumba Timur	54	96	50	200
3	Kupang	47	58	72	177
4	TTS	145	39	94	278
5	TTU	46	26	121	193
6	Belu	28	5	48	81
7	Alor	79	24	72	175
8	Lembata	28	58	65	151
9	Flores Timur	67	55	128	250
10	Sikka	29	40	91	160
11	Ende	61	98	119	278
12	Ngada	15	50	86	151
13	Manggarai	20	71	132	223
14	Rote Ndao	68	4	47	119
15	Mang. Barat	39	77	53	169
16	Sumba Tengah	34	15	16	65
17	SBD	107	11	57	175
18	Nagekeo	25	31	57	113
19	Mang. Timur	55	59	63	177
20	Sabu Raijua	8	35	20	63
21	Malaka	56	9	62	127
22	Kota Kupang	6	2	43	51
NTT		1,063	870	1,517	3,450

Visualisasi hasil analisis gerombol pada peubah-peubah numerik menggunakan *Chernoff Face* yang berdasarkan perhitungan nilai rata-rata dari semua peubah numerik pada masing-masing gerombol. Visualisasi *Chernoff Face* adalah bentuk wajah (kepala, rambut, mulut, hidung, telinga) yang menggambarkan karakteristik masing-masing gerombol untuk semua peubah numerik.

Deskripsi visualisasi peubah numerik pada 3 gerombol optimal seperti ditunjukkan pada Tabel 3.

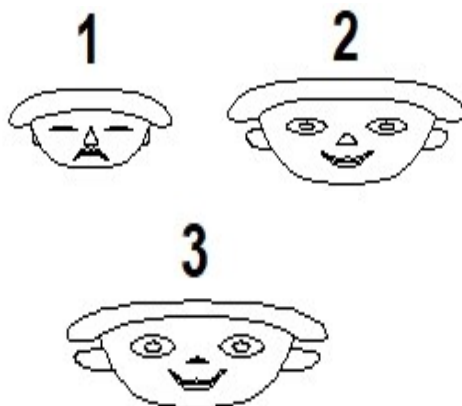
Tabel 3. Deskripsi peubah numerik sesuai dengan tampilan *Chernoff Face*

Peubah	Ciri Wajah	Keterangan
X1	<i>height of face</i>	Semakin tinggi bentuk wajah maka semakin tinggi rasio sarana kesehatan per 1000 penduduk
X2	<i>width of face</i>	Semakin lebar bentuk wajah maka semakin tinggi rasio tenaga kesehatan per 1000 penduduk
X3	<i>Smiling</i>	Semakin tersenyum maka semakin tinggi rasio sarana pendidikan per 1000 penduduk
X4	<i>height of eyes</i>	Semakin tinggi bentuk mata maka semakin banyak jumlah surat miskin atau Surat Keterangan Tidak Mampu (SKTM) yang dikeluarkan desa/kelurahan

Tabel 3. Deskripsi peubah numerik sesuai dengan tampilan *Chernoff Face*(lanjutan)

Peubah	Ciri Wajah	Keterangan
X5	<i>height of hair</i>	Semakin tinggi bentuk rambut maka semakin jauh jarak kantor desa/kelurahan ke kantor kecamatan
X6	<i>height of nose</i>	Semakin tinggi bentuk hidung maka semakin jauh jarak kantor desa/kelurahan ke kantor bupati/walikota
X8	<i>width of ear</i>	Semakin lebar bentuk telinga maka semakin tinggi persentase keluarga pengguna listrik

Chernoff Face untuk 3 gerombol optimal seperti ditunjukkan pada Gambar 5.



Gambar 5. *Chernoff Face* untuk 3 gerombol optimal

Berdasarkan gambar tersebut, Gerombol 2 dan gerombol 3 seperti memiliki banyak kesamaan. Akan tetapi bentuk beberapa ciri wajah pada gerombol 3 sedikit berbeda daripada

gerombol 2 terutama pada bentuk wajah dan senyum yang lebih lebar, bentuk rambut dan hidung yang lebih rendah serta telinga yang lebih panjang.

Hal tersebut menunjukkan bahwa wilayah desa/kelurahan di gerombol 3 secara rata-rata memiliki kondisi yang lebih baik dibandingkan dengan gerombol lainnya. Desa/Kelurahan di gerombol 3 secara rata-rata memiliki rasio sarana kesehatan, rasio tenaga kesehatan serta rasio pendidikan yang lebih baik dibandingkan gerombol yang lainnya.

Selanjutnya, untuk jarak kantor desa/kelurahan ke kantor camat (peubah X5) dan juga jarak ke kantor bupati/walikota (peubah X6) yang dicirikan melalui rambut dan hidung pada Gerombol 3 bentuknya lebih pendek daripada gerombol lainnya. Ini berarti dimensi pelayanan umum dan dimensi aksesibilitas dan transportasi pada desa/kelurahan di Gerombol 3 merupakan yang terbaik dibandingkan dengan 2 gerombol lainnya.

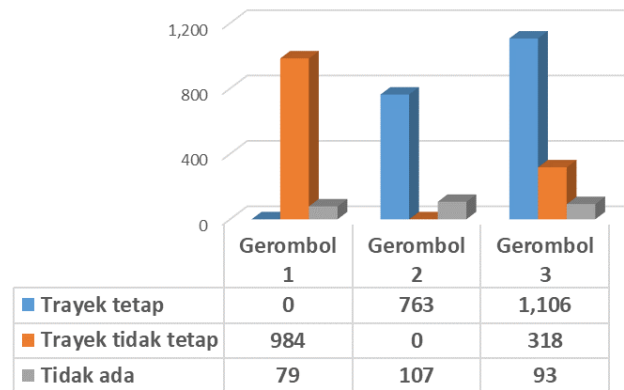
Hal yang sama ditunjukkan pada bentuk telinga (peubah X8) yang pada gerombol 3 bentuknya lebih panjang dibandingkan dengan gerombol lainnya. Artinya secara rata-rata desa/kelurahan pada gerombol 3 memiliki persentase keluarga pengguna listrik tertinggi dibandingkan dengan gerombol lainnya.

Berkebalikan dengan Gerombol 3, Gerombol 1 menjadi titik perhatian utama karena memiliki bentuk wajah penciri peubah numerik yang cenderung

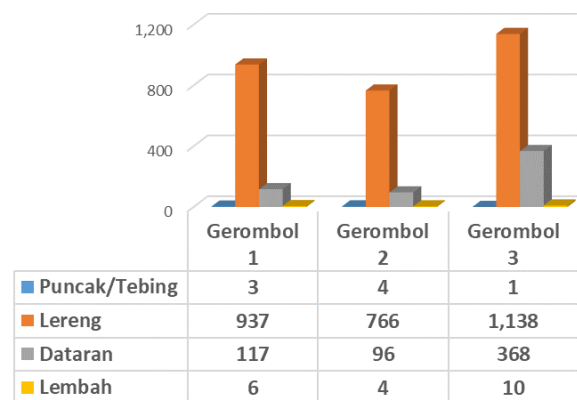
kurang baik dibandingkan dengan gerombol lainnya. Visualisasi *Chernoff Face* menunjukkan rasio sarana kesehatan, rasio tenaga kesehatan, rasio pendidikan, jumlah surat miskin yang dikeluarkan desa/kelurahan, serta persentase keluarga pengguna listrik pada gerombol ini merupakan yang paling rendah dibandingkan dengan gerombol lain.

Visualisasi hasil 3 gerombol optimal untuk peubah kategorik dilakukan melalui diagram batang (*bar chart*). Untuk peubah X7 (Keberadaan angkutan umum yang melewati desa/kelurahan) pada gerombol 1 didominasi oleh wilayah yang keberadaan angkutan umumnya ada dengan trayek tidak tetap dengan jumlah 984 desa/kelurahan. Untuk gerombol 2 dan gerombol 3 sebagian besar wilayah desa/kelurahan telah dilalui angkutan umum dengan trayek tetap (Gambar 6).

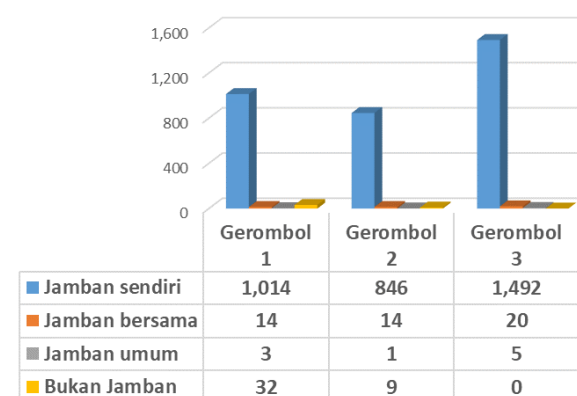
Untuk peubah X9 (topografi wilayah desa/kelurahan), semua gerombol didominasi oleh wilayah desa/kelurahan dengan topografi lereng (Gambar 7). Hal serupa juga terjadi pada peubah X10 (Fasilitas tempat buang air besar sebagian besar keluarga) yang anggota desa/keluruhannya pada masing-masing gerombol terdistribusi cukup merata (Gambar 8).



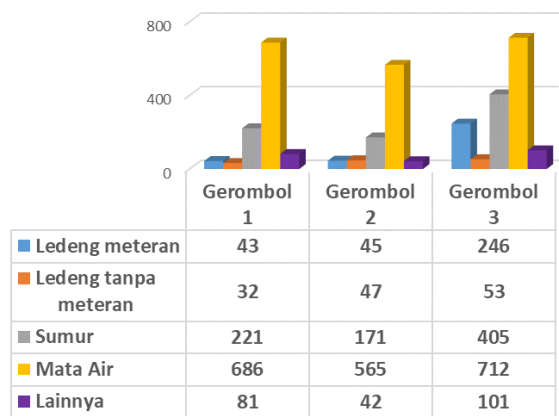
Gambar 6. Jumlah desa/kelurahan berdasarkan keberadaan angkutan umum yang melewati desa/kelurahan



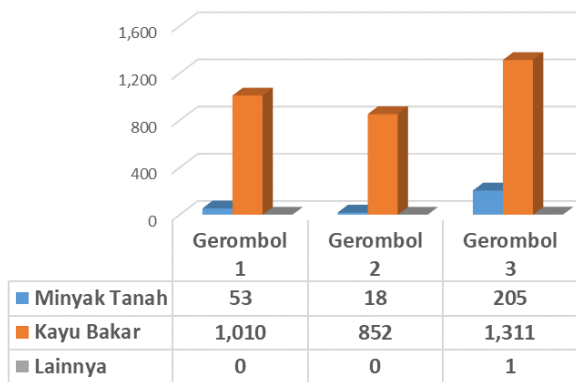
Gambar 7. Jumlah desa/kelurahan berdasarkan topografi wilayah



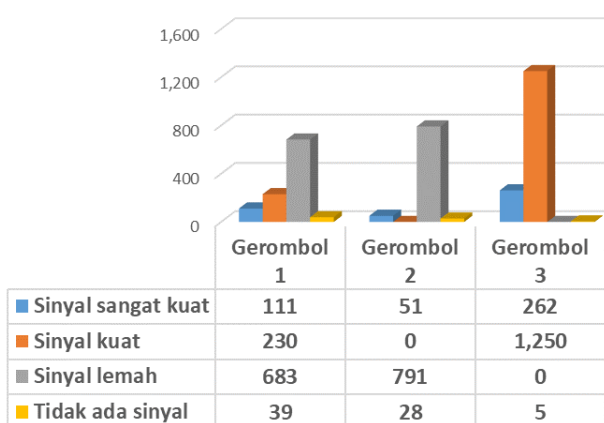
Gambar 8. Jumlah desa/kelurahan berdasarkan fasilitas tempat buang air besar sebagian besar keluarga



Gambar 9. Jumlah desa/kelurahan berdasarkan sumber air minum sebagian besar keluarga



Gambar 10. Jumlah desa/kelurahan berdasarkan bahan bakar untuk memasak sebagian besar keluarga



Gambar 11. Jumlah desa/kelurahan berdasarkan ketersediaan dan kualitas sinyal komunikasi seluler sebagian besar keluarga

Peubah kategorik X11 (sumber air minum sebagian besar keluarga) terlihat penggunaan sumber air minum yang berasal dari mata air cukup merata untuk semua gerombol. Akan tetapi gerombol 3 cenderung lebih baik karena desa/kelurahan yang penggunaan sumber airnya melalui ledeng meteran, ledeng tanpa meteran maupun sumur jumlahnya pada gerombol ini lebih banyak dibandingkan dengan gerombol lain (Gambar 9).

Peubah kategorik X12 (Bahan Bakar untuk memasak sebagian besar keluarga), desa/kelurahan yang sebagian besar keluarganya menggunakan kayu bakar sebagai bahan bakar untuk memasak terdistribusi cukup merata ke semua gerombol. Akan tetapi gerombol 3 bisa dikatakan lebih baik karena pada gerombol ini desa/kelurahan yang menggunakan minyak tanah relatif cukup banyak yaitu sebanyak 205 desa/kelurahan (Gambar 10).

Peubah kategorik X13 (Ketersediaan dan kualitas sinyal komunikasi seluler) yang paling baik ada pada gerombol 3 karena desa/kelurahan dengan sinyal kuat dan sinyal sangat kuat sebagian besar ada pada gerombol ini. Sedangkan desa/kelurahan dengan sinyal lemah dan tidak ada sinyal terdistribusi pada gerombol 1 dan gerombol 2 (Gambar 11).

Berdasarkan hasil pembahasan peubah-peubah kategorik untuk setiap gerombol maka desa/kelurahan yang kondisi pembangunan wilayahnya paling baik secara berturut turut adalah

pada gerombol 3, gerombol 2 dan terakhir gerombol 1. Hal ini sejalan dengan pembahasan gerombol pada peubah-peubah numerik.

4. Kesimpulan dan Saran

Kesimpulan

- a. Analisis gerombol yang dilakukan pada 13 peubah campuran (7 peubah numerik dan 6 peubah kategorik) menggunakan metode Gower (*Gower's Distance*) menghasilkan 3 gerombol optimal. Gerombol 1 terdiri dari 1.063 desa/kelurahan, gerombol 2 terdiri dari 870 desa/kelurahan dan gerombol 3 terdiri dari 1.517 desa/kelurahan.
- b. Berdasarkan visualisasi hasil gerombol pada peubah numerik maupun juga peubah kategorik maka wilayah yang kecenderungan capaian pembangunannya paling tinggi adalah desa/kelurahan pada gerombol 3, sedangkan yang paling rendah adalah pada wilayah desa/kelurahan anggota gerombol 1.

Saran

- a. Pemerintah perlu mengoptimalkan program pembangunan secara komprehensif dan berkelanjutan terutama pada desa/kelurahan pada Gerombol 1 yang capaian pembangunannya paling rendah dibandingkan dengan gerombol lainnya.
- b. Keterbatasan penelitian ini yaitu hanya menampilkan nilai pencilan pada masing-masing peubah numerik

(*univariate outlier*). Penelitian selanjutnya dapat melakukan perhitungan untuk nilai-nilai yang memiliki perbedaan besar dibandingkan dengan nilai lainnya pada dua atau lebih peubah numerik (*multivariate outlier*). Perlakuan atau penanganan pencilan akan meningkatkan kualitas gerombol yang terbentuk.

- c. Selain Analisis Gerombol menggunakan metode Gower, ada beberapa metode lain yang dapat digunakan untuk melakukan pengelompokan pada data dengan peubah campuran (numerik dan kategorik). Alternatif metode lain yang dapat digunakan yaitu Algoritma *Two Step Cluster*, *Cluster Ensemble*, atau juga *K-Prototypes*. Penelitian selanjutnya dapat menggunakan dan membandingkan antara metode Gower dengan metode lainnya dalam menghasilkan gerombol optimal.

Daftar Pustaka

- Aditya, A., Sari, B. N., & Padilah, T. N. (2021). Perbandingan pengukuran jarak Euclidean dan Gower pada kluster k-medoids. *Jurnal Teknologi dan Sistem Komputer*, 9(1), 1-7.
- Akay, Ö., & Yüksel, G. (2018). Clustering the mixed panel dataset using Gower's distance and k-prototypes algorithms. *Communications in Statistics-Simulation and Computation*, 47(10), 3031-3041.

- Badan Pusat Statistik. (2023). Berita Resmi Statistik No. 07/01/Th. XXVI, 16 Januari 2023. *Profil Kemiskinan di Indonesia September 2022*, Jakarta (ID): BPS.
- Chernoff H. (1973). The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association*. 68(342):361– 368.
- Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, 857-871.
- Kaufman, L., Rousseuw, P.J. (1990). *Finding Group in Data : An Introduction to Cluster Analysis*. New Jersey: John Wiley & Sons, Inc.
- Munthe, A. D. (2019). Penerapan clustering time series untuk menggerombolkan provinsi di Indonesia berdasarkan nilai produksi padi. *Jurnal Litbang Sukowati: Media Penelitian dan Pengembangan*, 2(2), 11-11.
- Munthe, A. D., Sumertajaya, I. M., & Syafitri, U. D. (2018). Penggerombolan desa/kelurahan berdasarkan indikator kemiskinan dengan menerapkan algoritma TSC dan K-Prototypes. *Indonesian Journal of Statistics and Its Applications*, 2(2), 63-76.
- Van den Hoven, J. (2015). Clustering with optimised weights for Gower's metric. *University of Amsterdam*, 14-17.