



Pengelompokan Desa di Kabupaten Kepulauan Sangihe Berdasarkan Indikator Sosial Ekonomi dengan Menggunakan Analisis Komponen Utama dan Analisis Gerombol

Sonya Yulia Takaredase¹, Hanny Komalig¹, John S. Kekenusa^{1*}

¹Jurusan Matematika–Fakultas Matematika dan Ilmu Pengetahuan Alam–Universitas Sam Ratulangi Manado, Indonesia

*Corresponding author : johnskekenusa@unsrat.ac.id

ABSTRAK

Penelitian ini bertujuan untuk mengetahui desa yang mempunyai kemiripan berdasarkan variabel yang ada menggunakan *plot multivariate* dan untuk mengelompokan desa di Kabupaten Kepulauan Sangihe dengan menggunakan Analisis Gerombol. Data yang digunakan merupakan data sekunder tahun 2018 yang diperoleh dari Badan Pusat Statistik (BPS) Kabupaten Kepulauan Sangihe. Analisis yang digunakan adalah Analisis Komponen Utama dan Analisis Gerombol. Hasil Analisis Komponen Utama yaitu KU_1, KU_2, KU_3, KU_4 dan KU_5 yang mewakili 67,383% variabilitasnya dan terbentuk 4 kelompok dengan menggunakan Analisis Gerombol.

INFO ARTIKEL

Diterima : 11 Maret 2019

Diterima setelah revisi : 24 Maret 2019

Tersedia online : 31 Maret 2019

Kata Kunci:

Pengelompokan desa,
Indikator Sosial Ekonomi,
Analisis Komponen Utama,
Analisis Gerombol

1. PENDAHULUAN

Menurut Kamus Besar Bahasa Indonesia, kata sosial berarti berkenaan dengan masyarakat (KBBI, 2002:1454). Menurut Departemen Sosial, kata sosial adalah segala sesuatu yang dipakai sebagai acuan dalam berinteraksi antar manusia dalam konteks masyarakat atau komunitas, sebagai acuan berarti sosial bersifat abstrak yang berisi simbol-simbol berkaitan dengan pemahaman terhadap lingkungan, dan berfungsi untuk mengatur tindakan-tindakan yang dimunculkan oleh individu-individu sebagai anggota suatu masyarakat, sehingga dengan demikian sosial haruslah mencakup lebih dari seorang individu yang terikat pada suatu kesatuan interaksi karena lebih dari seorang individu berarti terdapat hak dan kewajiban dari masing-masing individu yang saling berfungsi satu dengan yang lainnya.

Sumber daya manusia merupakan potensi yang terkandung dalam diri manusia untuk mewujudkan perannya sebagai makhluk sosial yang adaptif dan transformatif yang mampu mengelolah dirinya sendiri serta seluruh potensi yang terkandung di alam menuju tercapainya kesejahteraan kehidupan dalam tatanan yang seimbang dan berkelanjutan [1].

Suatu wilayah mempunyai karakter tertentu yang meliputi aspek fisik, dan aspek nonfisik. Aspek fisik diantaranya sarana, prasarana dan akses. Aspek nonfisik diantaranya sosial budaya, ekonomi, dan penduduk. Aspek-aspek ini merupakan salah satu cara untuk menerangkan keberagaman antar wilayah, salah satunya wilayah kecamatan. Terdapat kemungkinan bahwa suatu wilayah memiliki kondisi aspek-aspek yang berbeda dengan wilayah lainnya. PODES (Potensi Desa) 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].

Tujuan khusus analisis komponen utama adalah meringkaskan bentuk hubungan di antara variabel-variabel teramati, mereduksi jumlah variabel yang banyak menjadi lebih sedikit, atau dari dimensi besar menjadi dimensi kecil, sehingga mudah diinterpretasikan. Analisis Gerombol merupakan teknik multivariat yang mempunyai tujuan utama untuk mengelompokan objek-objek berdasarkan karakteristik yang dimilikinya. Analisis gerombol mengklasifikasi objek sehingga setiap objek yang paling dekat kesamaanya dengan objek lain berada dalam gerombol yang sama [4]. Penelitian ini bertujuan yang pertama untuk mengetahui desa yang mempunyai kemiripan berdasarkan variabel yang ada menggunakan *plot multivariate* dan yang kedua bertujuan untuk mengelompokan desa di Kabupaten Kepulauan Sangihe dengan menggunakan Analisis Gerombol.

Analisis Multivariat

Analisis multivariat atau Analisis Peubah Ganda adalah Teknik-teknik analisis statistika yang memperlakukan sekelompok variabel yang saling berkorelasi sebagai satu sistem, dengan memperhitungkan korelasi antar variabel-variabel tersebut. Dengan maksud lain, metode analisis multivariat adalah suatu metode statistika yang tujuan digunakannya adalah untuk menganalisis data yang terdiri dari banyak variabel serta diduga antar variabel tersebut saling berhubungan satu sama lain [4].

Analisis Korelasi

Korelasi bertujuan untuk melihat hubungan antara dua variabel yang tidak menunjukkan hubungan fungsional (berhubungan bukan disebabkan). Pengujian korelasi tidak membedakan jenis variabel apakah variabel dependen maupun independen.

Analisis Komponen Utama

Analisis Komponen Utama (AKU) dipelopori oleh Karl Pearson pada tahun 1901 untuk peubah nonstokastik, AKU adalah teknik statistik yang dapat digunakan untuk menjelaskan struktur variansi-kovariansi dari sekumpulan variabel baru di mana variabel baru ini saling bebas, dan merupakan kombinasi linier dari variabel asal.

Dalam bentuk matematis, katakan saja bahwa Y merupakan kombinasi linier dari variabel-variabel X_1, X_2, \dots, X_p yang dapat dinyatakan sebagai [5]:

$$Y = W_1 X_1 + W_2 X_2 + \dots + W_p X_p \quad (1)$$

Dengan :

W_i adalah bobot atau koefisien untuk variabel ke i

X_i adalah variabel ke i

Y adalah kombinasi linier dari variabel X

Analisis Gerombol

Gerombol dapat diartikan sebagai kelompok dengan demikian pada dasarnya analisis gerombol akan menghasilkan sejumlah gerombol atau kelompok [6]. Analisis Gerombol (*Cluster*) adalah teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya.

Sosial Ekonomi Dalam Pembangunan

Pembangunan sosial sebagai pendekatan pembangunan yang bertujuan meningkatkan kualitas kehidupan manusia secara paripurna, yakni memenuhi kebutuhan manusia yang terentang mulai dari kebutuhan fisik sampai sosial. Secara kontekstual pembangunan sosial lebih berorientasi pada prinsip keadilan sosial ketimbang pertumbuhan ekonomi [3].

2. METODE PENELITIAN

2.1. Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan selama bulan Desember 2018 sampai dengan Maret 2019. Data yang di peroleh merupakan data sekunder yang di ambil di Badan Pusat Statistik (BPS) Kabupaten Kepulauan Sangihe. Pengolahan data dilakukan di laboratorium Statistika Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam.

2.2. Tahapan Penelitian

Langkah-langkah dalam tahapan penelitian adalah :

1. Pengambilan data
2. Membuat Analisis Korelasi untuk melihat besar kecilnya hubungan dari tiap variabel.
3. Menganalisa data menggunakan metode Analisis Komponen Utama (AKU) untuk mengetahui kombinasi linier dari indikator yang ada.
4. Membuat Plot *Multivariate* untuk mengetahui desa mana saja yang mempunyai kemiripan dengan desa yang lain berdasarkan indikator-indikator yang digunakan.
5. Membuat gerombol berdasarkan hasil score dari Komponen Utama

2.3. Variabel Penelitian

Data yang digunakan merupakan data sekunder dari Badan Pusat Statistik (BPS) Kabupaten Kepulauan Sangihe. Data tersebut merupakan data terbaru yaitu data tahun 2018.

- X_1 : Luas Area Desa (Km^2)
- X_2 : Jumlah Wilayah
- X_3 : Jumlah Lembaga Adat
- X_4 : Jumlah Penduduk Tahun 2017
- X_5 : Laju Pertumbuhan Penduduk Per Tahun (%)
- X_6 : Jumlah Total Penduduk
- X_7 : Jumlah Rumah Tangga
- X_8 : Jumlah TK
- X_9 : Jumlah SD
- X_{10} : Jumlah SMP
- X_{11} : Jumlah Puskesmas
- X_{12} : Jumlah Tempat Ibadah
- X_{13} : Jumlah Industri Makanan
- X_{14} : Jumlah Industri Lainnya
- X_{15} : Jumlah PBB Bruto
- X_{16} : Jumlah PBB Neto

2.4. Tahapan Analisis Data

Teknik analisis yang digunakan dalam pengaplikasian model Analisis Komponen Utama dan Analisis Gerombol pada penelitian ini yaitu menggunakan perangkat lunak computer yaitu *Microsoft Office Excel*, *Minitab 17* dan *IBM SPSS*, dengan langkah-langkah sebagai berikut :

1. Analisis Korelasi
2. Analisis Komponen Utama
3. Analisis Gerombol

3. HASIL DAN PEMBAHASAN

a. Analisis Korelasi

Dari 145 desa di Kabupaten Kepulauan Sangihe, terdiri dari 16 Variabel ($X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}, X_{12}, X_{13}, X_{14}, X_{15}, X_{16}$) dilakukan analisis korelasi Pearson antar variabel untuk mengetahui keeratan hubungan antar variabel.

b. Analisis Komponen Utama

Setelah dilakukan analisis korelasi, dilanjutkan dengan analisis Komponen Utama untuk menentukan jumlah kelompok yang akan digunakan dalam analisis gerombol. Pertama menentukan nilai eigen dan vektor

Pengelompokan Desa Berdasarkan Indikator Sosial Ekonomi di Kabupaten Kepulauan Sangihe Dengan Menggunakan Analisis Komponen Utama dan Analisis Gerombol

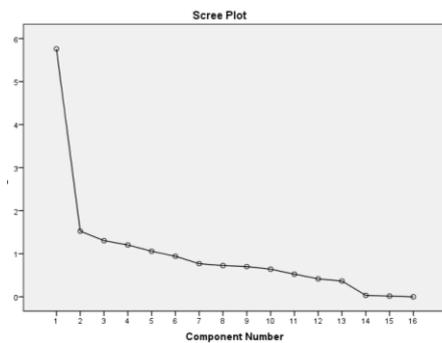
d'Cartesian : Jurnal Matematika dan Aplikasi, Vol. 8, No. 1 (Maret 2019) : 45-48

eigen dari matriks korelasi. Dalam hal ini nilai eigen merupakan varians setiap Komponen Utamanya dan vektor eigen merupakan koefisien-koefisien komponen utamanya. Hasil proporsi dari nilai eigen dituangkan dalam tabel.

Tabel 1. Proporsi nilai eigen

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	5.760	36.003	36.003
2	1.526	9.536	45.539
3	1.305	8.157	53.696
4	1.204	7.528	61.224
5	1.058	6.614	67.838
6	.942	5.887	73.725
7	.770	4.814	78.539
8	.726	4.538	83.077
9	.703	4.396	87.473
10	.643	4.017	91.489
11	.526	3.289	94.778
12	.417	2.607	97.385
13	.366	2.289	99.674
14	.034	.215	99.889
15	.018	.111	100.000
16	2.228E-8	1.393E-7	100.000

Extraction Method: Principal Component Analysis.



Gambar 1. Grafik Scree Plot

Pada tabel 1 dan gambar 1 dapat dilihat nilai eigen dari Komponen Utama pertama (KU_1) sampai dengan nilai eigen Komponen Utama ke enam belas (KU_{16}) yang semakin kecil yaitu dari kiri atas ke kanan bawah. Scree plot adalah grafik yang menunjukkan relasi antar Komponen Utama. Kemudian pada nilai vector eigen yang merupakan koefisien dari Komponen Utama dapat dilihat pada tabel 1, bahwa ada 16 Komponen Utama yang terbentuk dari data variabel asal dengan proporsi yang berbeda-beda tetapi hanya 13 Komponen Utama yang dapat menjelaskan data variabel asal, karena hasil dari analisis menunjukkan bahwa KU_{14} , KU_{15} , KU_{16} memiliki nilai eigen dibawah 0,1%.

c. Analisis Gerombol

Selanjutnya dilakukan analisis gerombol dengan menggunakan skor komponen-komponen utama yang ada pada tabel 3 sebagai variabel indikatornya, karena data awal memiliki satuan skala yang berbeda-beda sehingga digunakan skor komponen utama.

Tabel 2. Hasil Akhir Analisis Gerombol

	Number of	Whitin cluster	Average Distance	Maximum distance
	Observations	Sum of square	from centroid	from centroid
Gerombol 1	23	284.774	3.325	5.714
Gerombol 2	33	145.199	1.838	5.627
Gerombol 3	43	198.950	1.944	5.342
Gerombol 4	46	123.147	1.421	4.828

Tabel 2 merupakan hasil akhir dari analisis gerombol dengan memperlihatkan jumlah observasi dari 4 gerombol yang terbentuk dimana gerombol pertama memuat sebanyak 23 desa, gerombol kedua memuat sebanyak 33 desa, gerombol ketiga memuat sebanyak 43 desa dan gerombol keempat memuat sebanyak 46 desa. Jumlah kuadrat dari gerombol (*within cluster sum of square*) dan jarak rata-rata dari pusat gerombol (*average distane from centroid*) bernilai rendah maka gerombol tersebut memiliki tingkat kemiripan yang besar begitupun sebaliknya. Ini berarti gerombol 4 lebih mirip dari gerombol 1,2 dan 3, gerombol 2 lebih mirip dari gerombol 1 dan 3, dan gerombol 2 lebih mirip dari gerombol 1. Dalam tabel 4 nilai jarak maksimum dari centroid yang lebih tinggi yaitu pada gerombol 1 menunjukkan observasi di gerombol 1 terletak lebih jauh dari pusat gerombol.

Tabel 3. Nilai rata-rata dari setiap gerombol

Variabel	Gerombol 1	Gerombol 2	Gerombol 3	Gerombol 4	Grand Centroid
skor 1	4.2482	-0.5372	0.6743	-2.3690	0
skor 2	0.5080	0.6833	-0.7898	-0.0059	0
skor 3	0.0157	-0.8202	0.3950	0.2113	0
Skor 4	0.1913	0.1671	-0.2977	0.0628	0
skor 5	-0.0772	-0.7060	0.3062	0.2588	0

Tabel 3 menunjukkan besarnya nilai rata-rata dari skor yang merupakan variabel pada setiap gerombol. Terlihat bahwa, pada gerombol 1 memiliki nilai yang paling tinggi dari ketiga gerombol lainnya. *Grand centroid* merupakan rata-rata dari masing-masing skor.

Tabel 4. Jarak rata-rata antar gerombol

	Gerombol 1	Gerombol 2	Gerombol 3	Gerombol 4
Gerombol 1	0	4.9016	3.8714	6.6497
Gerombol 2	4.9016	0	2.5210	2.4158
Gerombol 3	3.8714	2.5210	0	2.1689
Gerombol 4	6.6497	2.4158	3.1689	0

Tabel 4 memperlihatkan perbandingan jarak rata-rata antar gerombol untuk melihat bagaimana perbedaan antar gerombol.

4. PENUTUP

4.1. Kesimpulan

1. Analisis Komponen Utama dari 145 Desa dengan 16 variabel membentuk 5 Komponen Utama yaitu KU_1, KU_2, KU_3, KU_4 dan KU_5 yang mewakili 67,383% variabilitasnya. Dari plot skor 5 komponen utama yaitu terbentuk 4 gerombol sebagai dasar penentuan ilai k untuk proses penggerombolan menggunakan metode non-hirarki k -means.
2. Analisis Gerombol k -means dengan nilai $k=4$ menghasilkan 4 gerombol. Gerombol pertama beranggotakan 23 desa, gerombol kedua beranggotakan 33 desa, gerombol ketiga beranggotakan 43 desa dan gerombol keempat beranggotakan 46 desa.

4.2. Saran

Pada penelitian ini menggunakan Analisis Komponen Utama dan Analisis Gerombol (*Cluster*), pada analisis gerombol peneliti hanya membatasi menggunakan satu metode. Disarankan bagi peneliti yang ingin membuat penelitian seperti ini untuk menggunakan metode analisi gerombol lainnya agar dapat menjadi pembanding untuk pilihan penggunaan metode terbaik.

REFERENSI

- [1] Horton, Paul B. dan Hunt, Chester L. *Sosiologi Jilid 2 Edisi Keenam*. Jakarta : PT Gelora Aksara Pratama, 1984
- [2] Deliarnov. (2012). *Perkembangan Pemikiran Ekonomi*. Jakarta. Raja Grafindo Persada.
- [3] Rochmi, Arinda. 2011. "Pengelompokan kabupaten/kota di Jawa Timur Berdasarkan Kesamaan Nilai Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka dengan Metode Hirarkhi dan Non Hirarkhi". FMIPA ITS.
- [4] Suryanto. 1988. *Metode Statistika Multivariat*. Jakarta: Departemen Pendidikan dan Kebudayaan.
- [5] Johnson, R. A. dan D. W. Wichern. 1988. *Applied Multivariate Statistical Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey.
- [6] Mattjik, A dan I. Sumertajaya. 2011. Sidik Peubah Ganda Dengan Menggunakan SAS. Bogor, Bogor.

Sonya Yulia Takaredase (sonyatakaredase@yahoo.com)



Lahir di Belae , Sulawesi Utara pada tanggal 22 Juli 1997.. Menempuh pendidikan tinggi Jurusan Matematika, FMIPA, Universitas Sam Ratulangi Manado. Tahun 2019 adalah tahun terakhir ia menempuh studi. Makalah ini merupakan hasil penelitian skripsinya yang dipublikasikan.

John S. Kekenusa (johnskekenua@unsrat.ac.id)



Lahir di Tahuna, Sulawesi Utara pada tanggal 24 Agustus 1958. Pada tahun 1982 mendapatkan gelar Sarjana yang diperoleh dari Fakultas Perikanan, Universitas Sam Ratulangi Manado. Pada tahun 1988 mendapatkan gelar Magister Statistika Terapan di Institut Pertanian Bogor, dan pada tahun 2006 mendapat gelar Doktor yang diperoleh dari MIPA, UNAIR Surabaya. Jabatan Akademik Profesor (Guru Besar, Statistika) sejak tahun 2007.

Hanny A. H.Komalig (hanoy07@yahoo.com)



Pada tahun 1990, memperoleh gelar Insinyur di Fakultas Pertanian Universitas Sam Ratulangi Manado. Kemudian Ia mengikuti *Basic Science Bridging Program* bidang Matematika di Insitut Teknologi Bandung pada tahun 1992 selama 2 tahun. Selanjutnya memperoleh gelar Magister Sains bidang Statistika di Institut Pertanian Bogor tahun 1998. Pada tahun 2008 memperoleh gelar Doktor di Universitas Airlangga Surabaya Untuk bidang Matematika Modeling, menjadi dosen di departemen Matematika, FMIPA USRAT pada tahun 2000 sampai sekarang. Bidang yang ditekuni diantaranya Statistika Multivariat Non Linear dan Reduksi Dimensi