

Pengelompokan Kecamatan di Kabupaten Kepulauan Sangihe Berdasarkan Luas Panen dan Produksi Hortikultura Menggunakan Metode *K-Means*

Jeliana Hamel, Hanny Komalig*, Deiby Salaki

Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sam Ratulangi, Manado, 95115

*Corresponding author: komaligh@unsrat.ac.id

Abstrak

Penelitian ini bertujuan untuk mengelompokkan kecamatan-kecamatan di Kabupaten Kepulauan Sangihe berdasarkan luas panen dan produksi hortikultura, serta mengetahui variabel penelitian apa saja yang mempengaruhi dimensi komponen utama pembentuk gerombol. Data yang digunakan dalam penelitian ini merupakan data sekunder tahun 2020 yaitu data luas panen dan produksi hortikultura yang diperoleh dari publikasi Badan Pusat Statistik (BPS) Kabupaten Kepulauan Sangihe. Analisis yang digunakan dalam penelitian ini adalah Analisis Komponen Utama dan Analisis Gerombol dengan metode *k-means*. Hasil Analisis Komponen Utama yaitu KU_1 dan KU_2 yang mewakili 86,1% variabilitasnya dan terbentuk 2 kelompok dengan menggunakan Analisis Gerombol. Gerombol pertama terdiri dari 4 Kecamatan dan gerombol kedua terdiri dari 11 kecamatan. KU_1 dipengaruhi oleh variabel luas panen cabai rawit, luas panen tomat, luas panen ketimun, produksi cabai rawit, produksi tomat, dan produksi ketimun. Sedangkan, KU_2 dipengaruhi oleh variabel luas panen cabai besar, luas panen kangkung, produksi cabai besar, dan produksi kangkung.

Kata kunci : Analisis Komponen Utama, Analisis Gerombol, *K-Means*

Grouping of Districts in Sangihe Islands Regency Based on Harvested Area and Horticultural Production Using the *K-Means* Method

Abstract

*This study aims to classify sub-districts in Sangihe Islands Regency based on harvested area and horticultural production, and to find out which research variables influence the dimensions of the main cluster-forming components. The data used in this study is secondary data for 2020, namely data on harvested area and horticultural production obtained from publications from the Central Statistics Agency for Sangihe Islands Regency. The analysis used in this study is Principal Component Analysis and Cluster Analysis with the *k-means* method. The results of the Main Component Analysis, namely KU_1 and KU_2 , represented 86.1% of the variability and 2 groups were formed using Cluster Analysis. The first cluster consists of 4 districts and the second cluster consists of 11 districts. KU_1 is influenced by the variables of cayenne pepper harvested area, tomato harvested area, cucumber harvested area, cayenne pepper production, tomato production, and cucumber production. Meanwhile, KU_2 is influenced by the variables of large chili harvested area, kale harvested area, large chili production, and kale production.*

Keywords: Principal Component Analysis, Cluster Analysis, K-Means

PENDAHULUAN

Kabupaten Kepulauan Sangihe sebagian penduduknya bekerja sebagai petani. Berdasarkan publikasi Profil Ketenagakerjaan Kabupaten Kepulauan Sangihe 2021, sebesar 25,95% penduduk bekerja di sektor pertanian. Salah satu subsektor pertanian adalah pertanian hortikultura.

Hortikultura merupakan gabungan bahasa Latin, *hortus* yang mengandung arti kebun dan *culture* yang berarti bercocok tanam. Hortikultura bisa didefinisikan sebagai cara budidaya tanaman yang dilakukan di kebun dan halaman rumah (Fakultas Pertanian Universitas Medan Area, 2020).

Suatu wilayah mempunyai karakteristik tertentu, sehingga ada kemungkinan bahwa suatu wilayah memiliki kondisi atau karakteristik yang berbeda dengan wilayah lainnya. Di Kabupaten Kepulauan Sangihe, setiap Kecamatan menghasilkan produksi hortikultura dengan jumlah yang berbeda-beda, serta dengan luas panen yang berbeda pula. Maka, berdasarkan perbedaan kedua variabel tersebut, sudah menunjukkan bahwa kecamatan-kecamatan yang ada di Kabupaten Kepulauan Sangihe memiliki karakteristik yang berbeda-beda khususnya di sektor pertanian yang dalam hal ini yaitu pertanian hortikultura. Oleh karena perbedaan luas panen dan jumlah produksi hortikultura di setiap Kecamatan inilah, maka perlu dilakukan pengelompokan untuk mengetahui kecamatan-kecamatan yang memiliki kemiripan berdasarkan kedua variabel tersebut.

Ada dua Analisis Multivariat yang digunakan dalam penelitian ini, yaitu Analisis Komponen Utama dan Analisis Gerombol.

Penelitian sebelumnya pernah dilakukan oleh Pangkey *et al.* (2018) tentang Aplikasi Analisis Komponen Utama dan Analisis Gerombol pada Varietas Tanaman Hias Krisan (*Chrysanthemum morifolium* R.) di Kota Tomohon. Penelitian juga pernah dilakukan oleh Takaredase *et al.* (2019) tentang Pengelompokan Desa di Kabupaten Kepulauan Sangihe Berdasarkan Indikator Sosial Ekonomi dengan Menggunakan Analisis Komponen Utama dan Analisis Gerombol. Penelitian yang serupa juga pernah dilakukan oleh Mukuan *et al.* (2022) tentang Pengelompokan Kecamatan di Kabupaten Minahasa Berdasarkan Data Hasil Produksi Pertanian Tahun 2019 dengan Menggunakan Analisis Komponen Utama dan Analisis Gerombol.

Dalam penelitian ini digunakan variabel luas panen dan produksi hortikultura, yang terdiri dari 5 jenis tanaman hortikultura yaitu cabai besar, cabai rawit, tomat, ketimun, dan kangkung.

Berdasarkan uraian di paragraf-paragraf sebelumnya, maka akan dilakukan suatu penelitian untuk mengelompokkan kecamatan-kecamatan di Kabupaten Kepulauan Sangihe berdasarkan luas panen dan produksi hortikultura, serta mengetahui variabel penelitian apa saja yang mempengaruhi dimensi komponen utama pembentuk gerombol, menggunakan Analisis Komponen Utama dan Analisis Gerombol dengan metode *k-means*.

METODOLOGI PENELITIAN

Waktu Penelitian

Penelitian ini dilaksanakan dari bulan Oktober 2022 sampai bulan Februari 2023, mulai dari studi tentang topik yang diteliti, penyusunan makalah, pengambilan data, sampai pengolahan data.

Prosedur Penelitian

Pada penelitian ini digunakan Analisis Komponen Utama dan Analisis Gerombol. Metode Penggerombolan yang akan digunakan dalam penelitian ini adalah metode tak berhirarki yaitu metode *k-means*. Data akan diolah dengan bantuan *Microsoft Office Excel dan Minitab 19*.

Prosedur 1

Melakukan standarisasi data karena terdapat perbedaan skala pengukuran, dengan rumus :

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (1)$$

Prosedur 2

Melakukan analisis korelasi untuk mengetahui hubungan antar variabel penelitian. Koefisien korelasi (r) adalah nilai yang menunjukkan kuat atau lemah hubungan antar x dan y (Walpole dan Raymond, 1995). Nilai r terletak antara -1 dan 1 ($-1 \leq r \leq 1$), nilai r dapat dihitung dengan persamaan :

$$r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}} \quad (2)$$

Dimana : r_{xy} = koefisien korelasi antara x dan y
 n = banyaknya data
 x_i = variabel x
 y_i = variabel y

Prosedur 3

Melakukan analisis komponen utama. Analisis Komponen Utama (AKU) secara teknis merupakan suatu teknik mereduksi data multivariat yang mengubah variabel asal (suatu matriks data/asli) menjadi variabel baru yang merupakan kombinasi linier dari variabel asal (suatu set kombinasi linier yang lebih sedikit), akan tetapi menyerap sebagian besar jumlah varian dari data awal. Tujuan utamanya menjelaskan sebanyak mungkin jumlah varian data asli dengan sedikit mungkin komponen utama (Supranto, 2004).

Menurut Mattjik dan Sumertajaya (2011), kombinasi linier dari p peubah asal, dapat ditulis:

$$Y = AX \quad (3)$$

dimana :

$$Y = \begin{bmatrix} Y_1 \\ \dots \\ Y_p \end{bmatrix}, A = \begin{bmatrix} a_{11} & \dots & a_{1p} \\ \dots & \dots & \dots \\ a_{p1} & \dots & a_{pp} \end{bmatrix}, X = \begin{bmatrix} X_1 \\ \dots \\ X_p \end{bmatrix}$$

Sehingga komponen utama dapat ditulis sebagai :

$$Y_1 = \alpha_{11}X_1 + \alpha_{12}X_2 + \dots + \alpha_{1p}X_p \quad (4)$$

Prosedur 4

Membuat plot skor komponen utama untuk penentuan nilai k yang akan digunakan dalam melakukan analisis gerombol.

Prosedur 5

Melakukan analisis gerombol dengan metode *K-Means* untuk mengelompokkan objek penelitian. Analisis gerombol merupakan teknik peubah ganda yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan kemiripan karakteristik yang dimilikinya (Mattjik dan Sumertajaya, 2011).

Menurut Handoyo *et al.* (2014), algoritma *K-Means* adalah metode klastering berbasis jarak yang membagi data ke dalam sejumlah klaster dan hanya bekerja pada atribut numerik. Secara umum metode analisis klaster *K-Means* menggunakan algoritma sebagai berikut :

1. Tentukan k sebagai jumlah klaster yang dibentuk.
2. Bangkitkan k titik pusat klaster (*centroid*) awal secara random. Penentuan *centroid* awal dilakukan secara random/acak dari objek-objek yang tersedia sebanyak k klaster, kemudian untuk menghitung *centroid* klaster ke- i berikutnya digunakan rumus sebagai berikut :

$$v = \frac{\sum_{i=1}^n x_i}{n}; i = 1, 2, 3, \dots, n \quad (5)$$

dimana : v = *centroid* pada klaster
 x_i = objek ke- i
 n = banyaknya data

3. Hitung jarak setiap objek ke masing-masing *centroid* dari masing-masing klaster. Untuk menghitung jarak antara objek dengan *centroid* digunakan jarak *Euclidian*, dengan rumus:

$$d_{x,y} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (6)$$

Dengan : $d_{x,y}$ = jarak antara objek x dengan objek y
 x_i = nilai tengah pada gerombol x hingga i
 y_i = nilai tengah pada gerombol y hingga i
 n = banyaknya peubah yang diamati

4. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan centroidnya.
5. Tentukan posisi *centroid* baru (C_k) dengan cara menghitung nilai rata-rata dari data yang ada pada *centroid* yang sama, dengan rumus :

$$C_k = \left(\frac{1}{n_k}\right) \sum d_i \quad (7)$$

Dimana n_k adalah jumlah dokumen dalam kluster k dan d_i adalah dokumen dalam kluster k . Kembali ke langkah 3 jika posisi *centroid* baru dengan *centroid* lama, tidak sama.

HASIL DAN PEMBAHASAN

Hasil 1

Hasil standarisasi data dapat dilihat dalam Tabel 1.

Tabel 1. Hasil Standarisasi Data

Kecamatan	X_1	X_2	X_3	X_4	X_5
Manganitu Selatan	0.03807	-0.30174	-0.33809	-0.34693	-0.42846
Tatoareng	-0.53297	-0.71320	-0.70559	-0.76887	-0.42846
Tamako	-0.53297	1.59098	1.05838	2.32536	1.46179
Tabukan Selatan	-0.53297	0.68577	0.10290	-0.06564	0.07561
Tabukan Selatan Tengah	-0.53297	-0.63091	-0.63209	-0.62822	-0.55447
Tabukan Selatan Tenggara	-0.53297	-0.63091	-0.48509	-0.76887	-0.68049
Tabukan Tengah	0.03807	-0.71320	-0.48509	-0.62822	-0.68049
Manganitu	-0.53297	-0.38403	-0.19110	0.35630	0.07561
Tahuna	2.32223	0.43889	0.02940	0.49695	2.97399
Tahuna Timur	-0.53297	-0.79549	-0.70559	-0.62822	-0.55447
Tahuna Barat	0.60911	2.57849	3.18985	2.04406	-0.17642
Tabukan Utara	2.32223	0.43889	0.10290	0.49695	0.45366
Nusa Tabukan	-0.53297	-0.71320	-0.48509	-0.76887	-0.68049
Kepulauan Marore	-0.53297	-0.87778	-0.63209	-0.76887	-0.30244
Kendahe	-0.53297	0.02743	0.17640	-0.34693	-0.55447
Manganitu Selatan	0.24330	-0.00084	-0.18903	-0.22720	-0.33432
Tatoareng	-0.48661	-0.42075	-0.29065	-0.33847	-0.56621
Tamako	-0.48661	1.18353	-0.24908	0.51546	0.53529
Tabukan Selatan	-0.48661	-0.16055	-0.24908	-0.29448	0.07150
Tabukan Selatan Tengah	-0.48661	-0.44228	-0.29007	-0.33329	-0.62419
Tabukan Selatan Tenggara	-0.48661	-0.42972	-0.28776	-0.33847	-0.68216
Tabukan Tengah	-0.24330	-0.42793	-0.27853	-0.33329	-0.68216
Manganitu	-0.48661	-0.39922	-0.26871	-0.31259	-0.21837
Tahuna	1.45982	-0.30231	-0.27448	-0.29706	3.14413
Tahuna Timur	-0.48661	-0.44946	-0.29065	-0.31259	0.76719
Tahuna Barat	0.36496	3.29207	3.61240	3.53009	-0.45027
Tabukan Utara	3.04129	-0.26104	-0.24677	-0.26342	0.59327
Nusa Tabukan	-0.48661	-0.42434	-0.17460	-0.33847	-0.68216
Kepulauan Marore	-0.48661	-0.44049	-0.28950	-0.33847	-0.30533
Kendahe	-0.48661	-0.31667	-0.23349	-0.31776	-0.56621

Hasil 2

Hasil analisis korelasi dapat dilihat dalam Tabel 2.

Tabel 2. *Pairwise Pearson Correlations*

Sample 1	Sample 2	Correlation	95% CI for ρ	P-Value
X ₂	X ₁	0.348	(-0.200, 0.730)	0.204
X ₃	X ₁	0.254	(-0.297, 0.678)	0.362
X ₄	X ₁	0.330	(-0.220, 0.720)	0.230
X ₅	X ₁	0.639	(0.189, 0.868)	0.010
X ₆	X ₁	0.948	(0.846, 0.983)	0.000
X ₇	X ₁	0.136	(-0.404, 0.606)	0.628
X ₈	X ₁	0.169	(-0.376, 0.627)	0.546
X ₉	X ₁	0.151	(-0.392, 0.616)	0.592
X ₁₀	X ₁	0.684	(0.264, 0.886)	0.005
X ₃	X ₂	0.945	(0.840, 0.982)	0.000
X ₄	X ₂	0.923	(0.779, 0.974)	0.000
X ₅	X ₂	0.448	(-0.083, 0.781)	0.094
X ₆	X ₂	0.300	(-0.250, 0.704)	0.277
X ₇	X ₂	0.883	(0.677, 0.961)	0.000
X ₈	X ₂	0.718	(0.326, 0.899)	0.003
X ₉	X ₂	0.810	(0.510, 0.935)	0.000
X ₁₀	X ₂	0.243	(-0.307, 0.672)	0.382
X ₄	X ₃	0.853	(0.606, 0.950)	0.000
X ₅	X ₃	0.242	(-0.309, 0.671)	0.385
X ₆	X ₃	0.198	(-0.350, 0.645)	0.479
X ₇	X ₃	0.965	(0.897, 0.989)	0.000
X ₈	X ₃	0.886	(0.685, 0.962)	0.000
X ₉	X ₃	0.942	(0.832, 0.981)	0.000
X ₁₀	X ₃	0.056	(-0.470, 0.552)	0.843
X ₅	X ₄	0.560	(0.067, 0.833)	0.030
X ₆	X ₄	0.290	(-0.261, 0.698)	0.295
X ₇	X ₄	0.809	(0.507, 0.934)	0.000
X ₈	X ₄	0.568	(0.079, 0.837)	0.027
X ₉	X ₄	0.706	(0.304, 0.895)	0.003
X ₁₀	X ₄	0.319	(-0.231, 0.715)	0.246
X ₆	X ₅	0.483	(-0.039, 0.798)	0.068
X ₇	X ₅	0.139	(-0.402, 0.608)	0.621
X ₈	X ₅	-0.051	(-0.549, 0.473)	0.856
X ₉	X ₅	0.046	(-0.477, 0.546)	0.870
X ₁₀	X ₅	0.898	(0.715, 0.966)	0.000
X ₇	X ₆	0.085	(-0.447, 0.572)	0.763
X ₈	X ₆	0.105	(-0.431, 0.586)	0.710
X ₉	X ₆	0.089	(-0.443, 0.575)	0.751
X ₁₀	X ₆	0.530	(0.024, 0.820)	0.042
X ₈	X ₇	0.914	(0.755, 0.971)	0.000
X ₉	X ₇	0.977	(0.930, 0.992)	0.000
X ₁₀	X ₇	-0.037	(-0.539, 0.484)	0.895
X ₉	X ₈	0.977	(0.930, 0.993)	0.000
X ₁₀	X ₈	-0.130	(-0.602, 0.410)	0.645
X ₁₀	X ₉	-0.082	(-0.570, 0.449)	0.771

Hasil yang diperoleh dari analisis korelasi antar variabel menunjukkan bahwa semua variabel memiliki korelasi dimana nilai r terletak antara -1 dan 1 ($-1 \leq r \leq 1$) sehingga dapat dilanjutkan dengan Analisis Komponen Utama.

Hasil 3

Untuk menentukan komponen utama maka perlu mencari nilai eigen dan vektor eigen. Hasil proporsi dari nilai eigen dituangkan dalam Tabel 3.

Tabel 3. Proporsi Nilai Eigen

KU	Nilai Eigen	Proporsi (%)	Kumulatif (%)
1	5.6195	0.562	0.562
2	2.9921	0.299	0.861
3	0.9114	0.091	0.952
4	0.3209	0.032	0.984
5	0.0674	0.007	0.991
6	0.0550	0.006	0.997
7	0.0195	0.002	0.999
8	0.0091	0.001	0.999
9	0.0041	0.000	1.000
10	0.0010	0.000	1.000

Komponen utama yang pertama (KU_1) proporsi nilai eigennya hanya 56,2% sehingga dianggap belum dapat menggambarkan keragaman total. Tetapi bila dijumlahkan proporsi dari 2 komponen utama yaitu KU_1 dan KU_2 menghasilkan proporsi 86,1% yang artinya kedua komponen utama tersebut dapat mewakili total variabilitasnya.

Tabel 4 merupakan vektor eigen dari 2 komponen utama dengan masing-masing skor dituangkan dalam Tabel 5.

Tabel 4. Matriks Komponen Utama

Variabel	KU_1	KU_2
X_1	0.192	0.450
X_2	0.403	-0.011
X_3	0.407	-0.135
X_4	0.379	0.045
X_5	0.184	0.450
X_6	0.162	0.415
X_7	0.391	-0.205
X_8	0.350	-0.242
X_9	0.378	-0.229
X_{10}	0.115	0.499

Tabel 5. Skor Antar 2 Komponen Utama

	Skor 1	Skor 2
	-0.61381	-0.10990
	-1.58554	-0.61387
	2.67343	0.12532
	-0.12655	-0.23594
	-1.50526	-0.70102
	-1.52488	-0.81471
	-1.34974	-0.45343
	-0.65810	-0.25205
	1.64374	4.76810
	-1.43603	0.00182
	7.08808	-2.60658
	1.20971	2.98488
	-1.51636	-0.84227
	-1.57599	-0.43124
	-0.72269	-0.81910

Dari matriks komponen utama pada Tabel 4, jika diturunkan dalam bentuk persamaan matematis adalah sebagai berikut :

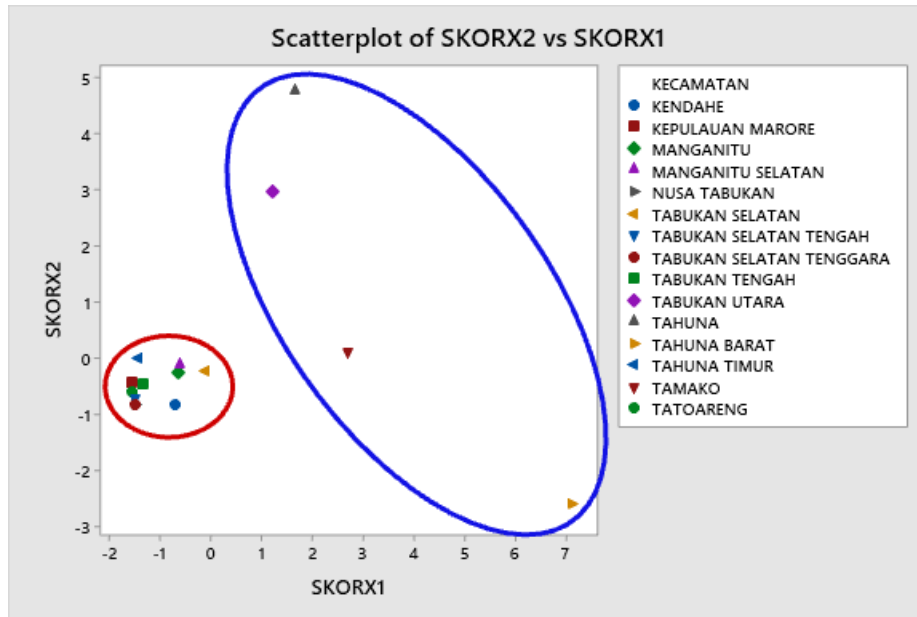
$$KU_1 = 0,192X_1 + 0,403X_2 + 0,407X_3 + 0,379X_4 + 0,184X_5 + 0,162X_6 + 0,391X_7 + 0,350X_8 + 0,378X_9 + 0,115X_{10}$$

$$KU_2 = 0,450X_1 - 0,011X_2 - 0,135X_3 + 0,045X_4 + 0,450X_5 + 0,415X_6 - 0,205X_7 - 0,242X_8 - 0,229X_9 + 0,499X_{10}$$

Variabel luas panen cabai rawit (X_2), luas panen tomat (X_3), luas panen ketimun (X_4), produksi cabai rawit (X_7), produksi tomat (X_8), dan produksi ketimun (X_9) memberikan korelasi yang besar pada KU_1 , yang artinya variabel-variabel tersebut adalah variabel yang lebih mempengaruhi dimensi komponen utama pertama. Sedangkan, variabel luas panen cabai besar (X_1), luas panen kangkung (X_5), produksi cabai besar (X_6), dan produksi kangkung (X_{10}) memberikan korelasi yang besar pada KU_2 , yang artinya variabel-variabel tersebut adalah variabel yang lebih mempengaruhi dimensi komponen utama kedua.

Hasil 4

Plot skor dua komponen utama pada Tabel 5, disajikan dalam Gambar 1.



Gambar 1. Plot Skor 2 Komponen Utama

Gambar 1 menunjukkan bahwa 2 komponen utama yaitu KU_1 dan KU_2 membentuk 2 kelompok. Jumlah kelompok ini yang akan menjadi acuan pengelompokan kecamatan dalam analisis gerombol ($k = 2$).

Hasil 5

Dalam melakukan analisis gerombol digunakan skor komponen-komponen utama yang ada di Tabel 5 sebagai variabel indikatornya. Hasil analisis gerombol dapat dilihat pada Tabel 6, Tabel 7, Tabel 8, dan Tabel 9.

Tabel 6. Hasil Akhir Analisis Gerombol

	<i>Number of Observations</i>	<i>Within Cluster Sum of Squares</i>	<i>Average Distance from Centroid</i>	<i>Maximum Distance from Centroid</i>
Gerombol 1	4	14.404	1.722	2.811
Gerombol 2	11	0.776	0.251	0.453

Tabel 6 menunjukkan jumlah observasi dari 2 gerombol yang terbentuk dimana gerombol pertama memuat 4 Kecamatan dan gerombol kedua memuat 11 Kecamatan. Jumlah kuadrat dari gerombol (*within cluster sum of squares*) dan jarak rata-rata dari pusat gerombol (*average distance from centroid*) bernilai rendah maka gerombol tersebut memiliki tingkat kemiripan yang besar. Ini berarti gerombol 2 lebih mirip dari gerombol 1.

Tabel 7. Nilai Rata-rata dari Setiap Gerombol

Variabel	Gerombol 1	Gerombol 2	<i>Grand Centroid</i>
Skor 1	1.3304	-0.4838	0.0000
Skor 2	0.7619	-0.2771	0.0000

Tabel 8. Jarak Rata-rata Antar Gerombol

	Gerombol 1	Gerombol 2
Gerombol 1	0.0000	2.0906
Gerombol 2	2.0906	0.0000

Tabel 9. Penempatan Kecamatan ke dalam Gerombol

Gerombol 1	Gerombol 2
Tamako	Manganitu Selatan
Tahuna	Tatoareng
Tahuna Barat	Tabukan Selatan
Tabukan Utara	Tabukan Selatan Tengah
	Tabukan Selatan Tenggara
	Tabukan Tengah Manganitu
	Tahuna Timur
	Nusa Tabukan
	Kepulauan Marore
	Kendahe

KESIMPULAN

1. Analisis Komponen Utama dari 15 Kecamatan dengan 10 variabel penelitian membentuk 2 komponen utama yaitu KU_1 dan KU_2 yang mewakili 86,1% variabilitasnya. Dari plot skor 2 komponen utama terbentuk 2 gerombol sebagai dasar penentuan nilai k untuk proses penggerombolan menggunakan metode tak berhirarki k -means. Analisis gerombol k -means dengan nilai $k = 2$ menghasilkan 2 gerombol. Gerombol pertama terdiri dari 4 Kecamatan dan gerombol kedua terdiri dari 11 Kecamatan.
2. KU_1 dipengaruhi oleh variabel luas panen cabai rawit (X_2), luas panen tomat (X_3), luas panen ketimun (X_4), produksi cabai rawit (X_7), produksi tomat (X_8), dan produksi ketimun (X_9). Sedangkan, KU_2 dipengaruhi oleh variabel luas panen cabai besar (X_1), luas panen kangkung (X_5), produksi cabai besar (X_6), dan produksi kangkung (X_{10}).

DAFTAR PUSTAKA

- Badan Pusat Statistik Kabupaten Kepulauan Sangihe. 2021. Kabupaten Kepulauan Sangihe Dalam Angka 2021. Katalog/*Catalog*: 1102001.7103. BPS Kabupaten Kepulauan Sangihe, Tahuna.
- Badan Pusat Statistik Kabupaten Kepulauan Sangihe. 2022. Profil Ketenagakerjaan Kabupaten Kepulauan Sangihe 2021. Katalog : 2303003.7103. BPS Kabupaten Kepulauan Sangihe, Tahuna.
- Fakultas Pertanian Universitas Medan Area. 2020. Tanaman Hortikultura. FAPERTA UMA, Medan.
- Handoyo, R., M.R. Rumani., dan S.M. Nasution. 2014. Perbandingan Metode *Clustering* Menggunakan Metode Single Lingkage dan *K-Means* pada Pengelompokkan Dokumen. *JSM STMIK Mikroskil*. **15(2)**: 72-82.
- Mattjik, A.A., dan I.M. Sumertajaya. 2011. Sidik Peubah Ganda dengan Menggunakan SAS. Edisi ke-1. IPB Press, Bogor.
- Mukuan, C.V., F.D. Pongoh., dan H.A.H. Komalig. 2022. Pengelompokan Kecamatan di Kabupaten Minahasa Berdasarkan Data Hasil Produksi Pertanian Tahun 2019 dengan Menggunakan Analisis Komponen Utama dan Analisis Gerombol. *d'Cartesian : Jurnal Matematika dan Aplikasi*. **11(1)**: 12-17.
- Pangkey, R.A., Y.A.R. Langi., dan H.A.H. Komalig. 2018. Aplikasi Analisis Komponen Utama dan Analisis Gerombol pada Varietas Tanaman Hias Krisan (*Chrysanthemum morifolium* R.) di Kota Tomohon. *d'Cartesian : Jurnal Matematika dan Aplikasi*. **7(2)**: 73-77.
- Supranto, J. 2004. Analisis Multivariat : Arti dan Interpretasi. PT. Rineka Cipta. Jakarta.
- Takaredase, S.Y., H. Komalig., dan J.S. Kekenusa. 2019. Pengelompokan Desa di Kabupaten Kepulauan Sangihe Berdasarkan Indikator Sosial Ekonomi dengan Menggunakan Analisis Komponen Utama dan Analisis Gerombol. *d'Cartesian : Jurnal Matematika dan Aplikasi*. **8(1)**: 45-48.
- Walpole, R., dan M. Raymond. 1995. Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan. ITB. Bandung.