



## Analisis Gerombol untuk Pengelompokan Kabupaten/Kota di Sulawesi Utara Berdasarkan Produksi Hortikultura

Eva Damayanti<sup>1</sup>, Jantje D. Prang<sup>1</sup>, Charles E. Mongi<sup>\*</sup>

<sup>1</sup>Jurusan Matematika–Fakultas Matematika dan Ilmu Pengetahuan Alam–Universitas Sam Ratulangi Manado, Indonesia

<sup>\*</sup>Corresponding author : [charlesmongi@unsrat.ac.id](mailto:charlesmongi@unsrat.ac.id)

### ABSTRAK

Pertanian adalah suatu kegiatan manusia yang termasuk di dalamnya yaitu bercocok tanam, peternakan, perikanan dan juga kehutanan. Produksi pertanian sangatlah penting dalam kehidupan sehari-hari. Dalam penelitian ini menggunakan Analisis Gerombol yang merupakan teknik peubah ganda yang mempunyai tujuan utama untuk mengelompokkan Kabupaten/Kota dan membandingkan pengelompokan analisis penggerombolan hirarki menggunakan Metode Pautan Tunggal, Pautan Lengkap, Pautan Rata-rata, Metode Centroid dan Metode Ward berdasarkan produksi pertanian di Sulawesi Utara. Dengan menggunakan Jarak Euclidian, penggerombolan pada Metode Pautan Tunggal, Pautan Lengkap, Pautan Rata-rata dan Metode Centroid terdapat 5 gerombol, sedangkan pada Metode Ward terdapat 6 gerombol. Berdasarkan gerombol akan dilihat perbandingan menggunakan rumus CTM (*Cluster Tightness Measure*) pada Metode Pautan Tunggal, Pautan Lengkap, Pautan Rata-rata, Metode Centroid menghasilkan nilai yang sama yaitu 0,054448 sedangkan metode Ward menghasilkan nilai 0,046549. Dapat di simpulkan bahwa pemilihan metode terbaik dalam penggerombolan adalah Metode Ward.

### ABSTRACT

Agriculture is a human activity which includes farming, livestock, fisheries and forestry. Agricultural production is very important in everyday life. In this study using cluster analysis which is a multivariate technique that has the main objective to group district/city and compare groupings of hierarchical clustering analysis using the single linkage method, complete linkage, average linkage, centroid method and ward method based on agricultural production in North Sulawesi. By using Euclidian Distance, the clusters on the Single Linkage Method, Complete Linkage, Average Linkage and Centroid Method there are 5 clusters, while in the Ward Method there are 6 Clusters. Based on the cluster, that will see a comparison using the CTM (*Cluster Tightness Measure*) formula on the Single Linkage Method, Complete Linkage, Average Linkage, the Centroid Method produces the same value which is 0.054448 while the Ward Method produces a value which is 0.046549. So, it can be concluded that the selection of the best method in clustering is the Ward Method.

### Informasi Artikel

Diterima : 20 Juni 2019

Diterima setelah revisi : 23 Juli 2019

Tersedia online : 25 Juli 2019

### Kata Kunci:

Produksi Hortikultura, Jarak Euclidian, Pautan Tunggal, Metode Ward, CTM (*Cluster Tightness Measure*).

### Article Information

Received : 20 June 2019

Received after revision : 23 July 2019

Available online : 25 July 2019

### Keywords:

Horticulture Production, Euclidian Distance, Single Linkage, Ward Method, CTM (*Cluster Tightness Measure*).

## 1. PENDAHULUAN

Indonesia adalah negara kepulauan yang terbentang dari Sabang sampai Marauke. Dan Indonesia merupakan negara agraris, karena mayoritas penduduk Indonesia bermata pencarian sebagai petani. Pertanian di Indonesia merupakan sektor yang menunjang terhadap kesejahteraan dan kemakmuran rakyatnya. Apabila di sektor pertanian ini berkembang dan maju, maka rakyat Indonesia akan terjamin hidupnya dengan kemakmuran [1].

Secara umum pengertian dari pertanian adalah suatu kegiatan manusia yang termasuk di dalamnya yaitu bercocok tanam, peternakan, perikanan dan juga kehutanan. Sebagian besar dari mata pencarian masyarakat di Indonesia adalah sebagai petani, sehingga sektor pertanian sangat penting untuk dikembangkan di negara kita [2].

Dalam penelitian ini menggunakan analisis gerombol yang merupakan teknik peubah ganda yang

mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan kemiripan karakteristik yang dimilikinya. Karakteristik objek-objek dalam suatu gerombol memiliki tingkat kemiripan yang tinggi, sedangkan karakteristik antar objek pada suatu gerombol dengan gerombol lain memiliki tingkat kemiripan yang rendah. Dengan kata lain, keragaman antar gerombol maksimum. Dengan menggunakan analisis gerombol peneliti bertujuan untuk dapat mengelompokkan kabupaten/kota di Sulawesi Utara berdasarkan produksi hortikultura [3].

Penelitian sebelumnya menggunakan data BPS tentang kemiskinan dengan menggunakan metode pautan tunggal, metode pautan lengkap, metode pautan rata-rata dan metode *k-means*. Berdasarkan kesimpulan penggerombolan kecamatan di Kabupaten Malang berdasarkan peubah status kemiskinan dengan metode pautan tunggal menghasilkan 4 gerombol. Penggerombolan dengan metode pautan lengkap, metode pautan rata-rata dan metode *k-means* ( $k=2$ )

menghasilkan 2 gerombol yang sama. Hasil penggerombolan kecamatan di Kabupaten Malang berdasarkan peubah status kemiskinan yang dianggap paling mewakili adalah 2 gerombol. Hal ini ditunjang oleh 3 metode yaitu metode pautan lengkap, metode pautan rata-rata, dan metode *k-means* ( $k=2$ ). Ketiga metode tersebut menghasilkan gerombol sama dengan 2. Gerombol I tidak memiliki ciri yang menonjol karena tidak ada peubah yang memiliki rata-rata lebih tinggi dibandingkan gerombol II semua peubahnya cenderung berada di sekitar rata-rata total [4].

Dalam Analisis Gerombol dengan Metode Hirarki yaitu Analisis yang penggerombolan datanya dilakukan dengan cara mengukur jarak kedekatan pada setiap objek yang kemudian membentuk sebuah dendogram. Pada penelitian ini menggunakan ke-5 Metode Analisis penggerombolan berhirarki (Metode Pautan tunggal, Pautan Lengkap, Pautan Rata-rata, Metode Centroid, dan Metode Ward) dengan menggunakan jarak Euclidian. Untuk dapat mengelompokkan Kabupaten/Kota dan membandingkan ke-5 Metode Hirarki berdasarkan kemiripan karakteristik dari sudut pandang pertanian dengan Analisis Gerombol.

**2. Analisis Gerombol**

Analisis Gerombol merupakan suatu metode peubah ganda yang bertujuan untuk mengelompokkan  $n$  objek pengamatan ke dalam  $m$  gerombol ( $m < n$ ) berdasarkan  $p$  peubah, sehingga setiap pengamatan yang terletak dalam satu gerombol mempunyai kemiripan sifat yang lebih besar dibandingkan dengan pengamatan yang terletak pada gerombol lain [5].

Analisis gerombol berhubungan dengan data berbentuk tabel dimana baris menyatakan objek, organisme atau individu dan kolom menyatakan peubah atau karakteristik. Dasar penggerombolan  $n$  objek ke dalam  $m$  Gerombol adalah jarak, yang merupakan ukuran kedekatan antara masing-masing objek. Dalam menentukan jarak harus diperhatikan satuan peubah-peubahnya, jika satuan peubah tidak sama maka perlu dilakukan pembakuan atau standarisasi peubah [6].

Adapun jarak yang biasa digunakan dalam analisis penggerombolan diantaranya adalah jarak Euclidian, jarak mahalnobis, jarak manhattan, jarak log likelihood [7].

**2.1. Pemilihan jarak**

**2.1.1. Jarak Euclidian**

Jarak *Euclidian* adalah jarak yang paling umum dan paling sering digunakan dalam analisis gerombol. Jarak *Euclidian* antara dua titik dapat terdefiniskan dengan jelas. Jarak yang digunakan kontinu. Jarak *Euclidian* antara gerombol ke-  $i$  dan ke-  $j$  dari  $p$  peubah dirumuskan sebagai berikut [8]:

$$d(i, k) = \sqrt{\sum_{j=1}^p (x_{ij} - x_{kj})^2} \tag{1}$$

Dengan:

- $d(i, k)$  = Jarak antara objek  $i$  ke objek  $k$
- $x_{ij}$  = Nilai data ke- $i$  pada variabel ke-  $j$
- $x_{kj}$  = Nilai data ke- $k$  pada variabel ke-  $j$
- $p$  = Banyaknya peubah yang diamati.

**2.1.2. Jarak Mahalanobis**

Jarak *Mahalanobis* sangat berguna dalam menghilangkan atau mengurangi perbedaan skala pada masing-masing komponen. Pada permasalahan tertentu, pada saat menentukan jarak, perlu juga dipertimbangkan ragam dan peragaman. Jarak *Mahalanobis* dirumuskan sebagai berikut [9]:

$$d(i, j) = [(\bar{x}_i - \bar{x}_j)S^{-1}(\bar{x}_i - \bar{x}_j)]^{1/2} \tag{2}$$

Dengan:

- $d(i, j)$  = Jarak antara objek  $i$  ke objek  $j$
- $\bar{x}_j$  = Nilai tengah pada gerombol ke-  $j$
- $\bar{x}_i$  = Nilai tengah pada gerombol ke-  $i$
- $S^{-1}$  = Matriks ragam peragam gabungan antara.

**2.1.3. Jarak Manhattan**

Jarak *Manhattan* antara dua objek merupakan jumlah nilai perbedaan mutlak untuk tiap variabel. Jarak ini juga disebut jarak *Minkowski*. Ukuran ini merupakan bentuk umum dari jarak *Euclidian*, fungsi jaraknya dirumuskan sebagai berikut [9]:

Jika  $\underline{x}' = (x_1, x_2, \dots, x_p)$ ;  $p$  adalah variabel, maka  $x' = (x_{i1}, x_{i2}, \dots, x_{ip})$  adalah kumpulan variabel pada objek ke-  $i$

$$d(i, j) = [\sum_{k=1}^p |x_{ik} - y_{jk}|^r]^{1/r} \tag{3}$$

Dengan:

- $d(i, j)$  = Jarak antara objek  $i$  ke objek  $j$
- $\bar{x}_j$  = Nilai tengah pada Gerombol ke-  $j$
- $\bar{x}_i$  = Nilai tengah pada Gerombol ke-  $i$
- $p$  = Banyaknya peubah yang diamati
- $d_y$  = jarak antara dua objek ke-  $i$  dan objek ke-  $j$

**2.1.4. Jarak Log Likelihood**

Jarak *Log Likelihood* dapat ditetapkan untuk peubah kontinu maupun kategorik. Jarak antara gerombol  $j$  dengan gerombol  $s$  dirumuskan pada persamaan 4 [9]:

$$d(j, s) = \xi_j + \xi_s - \xi_{(j,s)} \tag{4}$$

**2.2. Metode Penggerombolan Berhirarki**

Metode penggerombolan berhirarki digunakan apabila banyak gerombol yang akan dibentuk belum diketahui sebelumnya. Pada metode berhirarki terdapat 2 cara pembentukan gerombol yaitu *agglomerative* dan *divisive*. Metode hirarki *agglomerative* dimulai dengan masing-masing objek adalah satu gerombol kemudian digabungkan berdasarkan kesamaan objek sampai menjadi hanya satu gerombol saja yang berisi semua objek. Sedangkan metode hirarki *divisive* adalah kebalikannya dimulai dengan satu gerombol yang anggotanya adalah seluruh objek, kemudian objek-objek yang paling jauh dipisah dan membentuk gerombol lain. Proses ini berlanjut sampai semua objek masing-masing membentuk satu gerombol [10].

Metode-metode yang bisa digunakan dalam metode hirarki adalah metode Agglomeratif (*Agglomerative Method*) dan metode Defisif (*Devisive Method*) [11].

**2.2.1. Metode Agglomeratif (Agglomerative Method)**

Metode Agglomeratif dimulai dengan menganggap bahwa setiap objek adalah sebuah gerombol. Kemudian dua objek dengan jarak terdekat digabungkan menjadi satu gerombol. Selanjutnya objek ketiga akan bergabung dengan gerombol yang ada atau bersama objek lain dan membentuk gerombol baru dengan tetap memperhitungkan jarak kedekatan antar objek. Proses akan berlanjut hingga akhirnya terbentuk satu gerombol yang terdiri dari keseluruhan objek.

Metode Hirarki Agglomeratif (*Agglomerative Method*) masih ada lima metode dalam pembentukan gerombol yaitu [12]:

1. Pautan Tunggal (Single Linkage)
2. Pautan Lengkap (Complete Linkage)

3. Pautan Rata-rata (Average Linkage)
4. Metode Centroid (Pusat)
5. Metode Ward (Ward's Method).

Berikut penjelasan mengenai beberapa pautan dalam metode penggerombolan hirarki [5]:

### 2.2.1.1. Pautan Tunggal

Jarak dua gerombol diukur dengan jarak terdekat antara sebuah objek dalam gerombol yang satu dengan sebuah objek dalam gerombol lain [5].

$$d_{mj} = \min(d_{kj}, d_{lj}) \quad (5)$$

Dengan:

$d_{mj}$  = jarak antara gerombol  $m$  dan gerombol  $j$ .

$d_{kj}$  = jarak antara tetangga terdekat gerombol  $k$  dan  $j$ ,

$d_{lj}$  = jarak antara tetangga terdekat gerombol  $l$  dan  $j$ .

### 2.2.1.2. Pautan Lengkap

Jarak dua gerombol diukur dengan jarak terjauh antara sebuah objek dalam gerombol yang satu dengan sebuah objek dalam gerombol yang lain [5].

$$d_{mj} = \max(d_{kj}, d_{lj}) \quad (6)$$

### 2.2.1.3. Pautan Rata-rata

Jarak antara dua gerombol diukur dengan jarak rata-rata antara sebuah objek dalam gerombol yang satu dengan sebuah objek dalam gerombol yang lain [5].

$$d_{mj} = \frac{N_k d_{kj} + N_l d_{lj}}{N_m} \quad (7)$$

Dengan:

$N_k, N_l$  dan  $N_m$  = jumlah objek dalam gerombol ke- $k, l, m$ .

### 2.2.1.4. Metode Centroid

Jarak antara dua buah gerombol diukur sebagai jarak *Euclidian* antara kedua rata-rata (*centroid*) gerombol. Metode Centroid adalah nilai tengah observasi pada variabel dalam suatu *set variabel cluster*. Keuntungannya adalah outlier hanya sedikit berpengaruh jika dibandingkan dengan metode lain, yang dihitung dengan formulasi [13].

$$d_{mj} = \frac{N_k d_{kj} + N_l d_{lj}}{N_m} - \frac{N_k N_l d_{kl}}{N_m^2} \quad (8)$$

Dengan:

$N_k, N_l, N_m$  = jumlah objek dalam gerombol ke- $k, l, m$ .

### 2.2.1.5. Metode Ward

Jarak antara dua gerombol adalah jumlah kuadrat penyimpangan dari metode centroid. Tujuan dari metode ward adalah untuk meminimalkan dalam *cluster sum of squares*.

$$d_{mj} = \frac{(N_j + N_k) d_{kj} + (N_j + N_l) d_{lj} - N_j d_{kl}}{N_j + N_m} \quad (9)$$

Metode ini berbeda dengan metode lainnya karena menggunakan pendekatan analisis varians untuk menghitung jarak antar gerombol atau metode ini meminimumkan jumlah kuadrat (*SSE*). *SSE* hanya dapat dihitung jika Gerombol memiliki elemen lebih dari satu objek. *SSE* Gerombol yang hanya memiliki satu objek adalah nol. Metode Ward ini dihitung berdasarkan persamaan berikut [14]:

$$SSE = \sum_{i=1}^N (X_i - \bar{X})' (X_i - \bar{X}) \quad (10)$$

Dengan:

*SSE* = Sum of Square Error

$X_i$  = Vektor kolom yang entrinya nilai objek ke- $i$  dengan  $i = 1, 2, 3, \dots, N$

$\bar{X}$  = Vektor kolom yang entrinya rata-rata nilai objek dalam Gerombol

$N$  = Banyaknya objek.

### 2.2.2. Metode Defisif (*Devisive Method*)

Proses pengelompokan dengan pendekatan metode pemecahan (Top to Down) dimulai dengan  $n$  objek yang dikelompokkan menjadi satu gerombol, kemudian gerombol tersebut dipartisi ke dalam dua gerombol pada setiap langkah diperoleh  $n$  gerombol dengan setiap gerombol memiliki satu objek. Dasar pengelompokan juga didasarkan pada jarak. Akan tetapi, teknik ini tidak banyak digunakan, sehingga tidak banyak prosedur yang dikembangkan [15].

### 2.3. Pemilihan Metode Terbaik

Pemilihan metode terbaik dalam penggerombolan dapat dilihat dari keragaman yang terbentuk, semakin kecil nilai keragamannya, maka hasil pengelompokan semakin baik. *Cluster Tightness Measure* (CTM) merupakan ukuran kebaikan dari hasil pengelompokan berdasarkan simpangan baku setiap variabel pada masing-masing kelompok. Jika CTM bernilai nol, artinya semua data berada pada pusat kelompok. Metode pengelompokan terbaik dengan menggunakan nilai CTM dihitung dengan persamaan [16]:

$$CTM = \frac{1}{K} \sum_{t=1}^K \left( \frac{1}{P} \sum_{m=1}^P \frac{S_{tm}}{S_m} \right) \quad (11)$$

Dengan:

$S_{tm}$  = Simpangan baku pada kelompok ke- $t$  untuk variabel ke- $m$

$S_m$  = Simpangan baku seluruh data untuk variabel ke- $m$

$K$  = Banyaknya kelompok

$P$  = Banyaknya variabel.

## 3. METODE PENELITIAN

### 3.1. Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan dari bulan Desember 2018 sampai Februari 2019, mulai dari penyusunan proposal, pengambilan data serta pengolahan data. Pengambilan data dilakukan di Badan Pusat Statistik (BPS) Provinsi Sulawesi Utara dan pengolahan data dilakukan di Laboratorium Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sam Ratulangi Manado.

### 3.2. Jenis dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder, dari Badan Pusat Statistik (BPS) Sulawesi Utara data produksi hortikultura 2017 [17].

### 3.3. Variabel dan Objek Penelitian

Variabel dari penelitian ini adalah produksi hortikultura Sulawesi Utara. Pada produksi hortikultura terdapat 26 variabel tapi disini peneliti hanya menggunakan 11 variabel karena pada variabel lainnya terdapat banyak 0 sehingga hanya diambil 11 variabel. 11 variabel tersebut disajikan pada Tabel 1:

Objek dari penelitian ini adalah Kabupaten/Kota di Sulawesi Utara dan objek-objek yang digunakan disajikan pada Tabel 2:

**Tabel 1.** Variabel Penelitian

Produksi Hortikultura	
Variabel	Keterangan
X1	Tomat
X2	Cabai Rawit
X3	Kangkung
X4	Cabai Besar
X5	Kacang Panjang
X6	Terung
X7	Ketimun
X8	Labu Siam
X9	Bayam
X10	Kacang Merah
X11	Bawang Daun

**Tabel 2.** Objek Penelitian

Kabupaten/Kota	
Objek	Keterangan
Y1	Bolaang Mongondow
Y2	Minahasa
Y3	Kepulauan Sangihe
Y4	Kepulauan Talaud
Y5	Minahasa Selatan
Y6	Minahasa Utara
Y7	Bolaang Mongondow Utara
Y8	Kepulauan Sitaro
Y9	Minahasa Tenggara
Y10	Bolaang Mongondow Selatan
Y11	Bolaang Mongondow Timur
Y12	Kota Manado
Y13	Kota Bitung
Y14	Kota Tomohon
Y15	Kota Kotamobagu

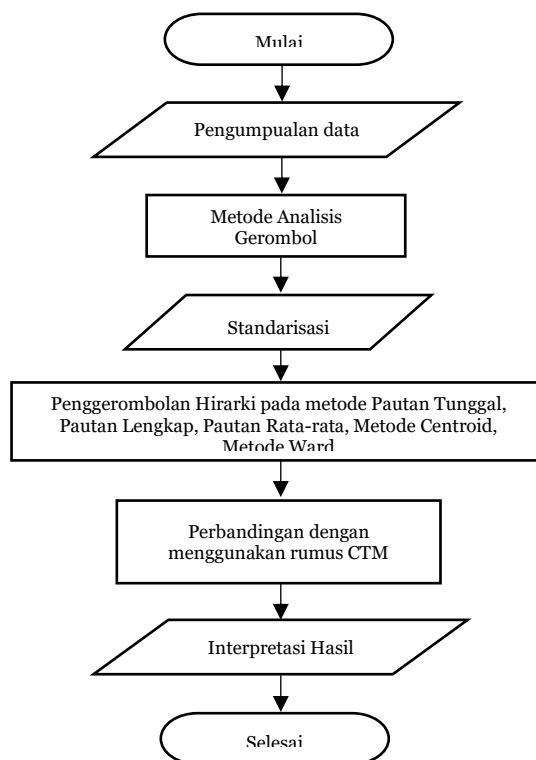
### 3.4. Metode Analisis

Adapun langkah-langkah Analisis data untuk penelitian tentang produksi hortikultura di Kabupaten/Kota Provinsi Sulawesi Utara pada tahun 2017 antara lain:

1. Melakukan Analisis statistika deskriptif untuk mengkaji karakteristik kabupaten/kota di provinsi Sulawesi utara yang di sajikan berupa tabel dengan menggunakan Aplikasi Minitab 17.
2. Melakukan hasil Analisis Penggerombolan pada metode Pautan Tunggal, Pautan Lengkap, Pautan Rata-rata, Metode Centroid, dan Metode Ward yang di sajikan dalam bentuk dendogram, setelah itu dilakukan cara pemotongan dendogram dan pengelompokan produksi hortikultura.
3. Menentukan metode terbaik diantara metode Pautan Tunggal, Pautan Lengkap, Pautan Rata-rata, Metode Centroid, dan Metode Ward pada analisis gerombol hirarki dengan menggunakan rumus CTM.
4. Menginterpretasikan hasil Analisis Gerombol Kabupaten/Kota pada Produksi hortikultura.

### 3.5. Diagram Alir (Flowchart) Penelitian

Selanjutnya, dibuat diagram alir digunakan untuk membantu analisis data yang disajikan pada Gambar 1.



**Gambar 1.** Diagram Alir Penelitian.

## 4. HASIL DAN PEMBAHASAN

### 4.1. Statistik Deskriptif Karakteristik Produksi Hortikultura di Sulawesi Utara

Sebelum dilaksanakan analisis gerombol terlebih dahulu di analisis secara deskriptif. Dalam menentukan deskriptif statistik di sajikan dengan menggunakan tabel. Dari 15 Kabupaten/Kota di Sulawesi Utara, di hasilkan distribusi produksi hortikultura seperti pada Tabel 3 data produksi hortikultura hasil statistik deskriptif tiap variabel seluruh Sulawesi utara 2017 dalam satuan kuintal.

Berdasarkan Tabel 3, terdapat 11 variabel dan 15 Kabupaten/Kota, dimana 11 variabel tersebut terdapat produksi Tomat, Cabai Rawit, Kangkung, Cabai Besar, Kacang Panjang, Terung, Ketimun, Labu Siam, Bayam, Kacang Merah dan Bawang Daun. Sedangkan 15 Kabupaten/Kota terdapat Kabupaten Bolaang Mongondow, Kabupaten Minahasa, Kabupaten Kepulauan Sangihe, Kabupaten Kepulauan Talaud, Kabupaten Minahasa Selatan, kabupaten Minahasa Utara, Kabupaten Bolaang Mongondow Utara, Kabupaten Kepulauan Sitaro, Kabupaten Minahasa Tenggara, Kabupaten Bolaang Mongondow Selatan, Kabupaten Bolaang Mongondow Timur, Kabupaten Kota Manado, Kabupaten Kota Bitung, Kabupaten Kota Tomohon, dan Kabupaten Kota Kotamobagu.

**Tabel 3.** Hasil Statistik Deskriptif tiap Variabel (Produksi Hortikultura) seluruh Sulawesi Utara 2017 dalam satuan kuintal.

Variabel	Rata-rata	Simpangan Baku	Minimum	Maksimum
Tomat	20.184	36.951	13	135.337
Cabai Rawit	8.870	15.352	279	55.783
Kangkung	2.665	3.956	0	15.291
Cabai Besar	3.251	6.036	0	22.038
Kacang Panjang	1.763	2.357	10	8.772
Terung	4.132	6.883	9	20.49
Ketimun	2.555	3.519	0	10.869
Labu Siam	4.273	10.549	0	38.455
Bayam	670	867	0	3.101
Kacang Merah	552	1.318	0	5.164
Bawang Daun	24.924	59.829	0	230.200

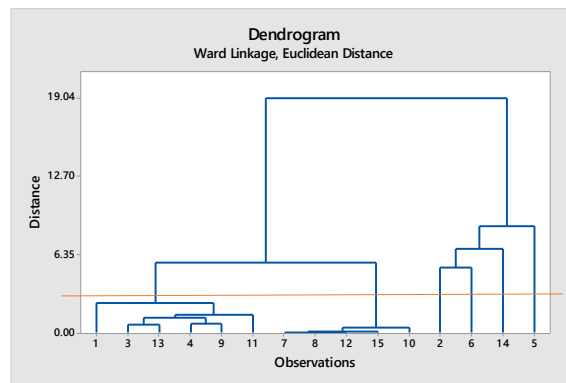
Cara pengambilan rata-rata dilakukan dengan cara dijumlahkan semua produksi tomat pada 15 kabupaten/kota kemudian dibagi 15, dilakukan hal yang sama pada produksi cabai rawit, produksi kangkung, produksi cabai besar, produksi kacang panjang, produksi terung, produksi ketimun, produksi labu siam, produksi bayam, produksi kacang merah, dan produksi bawang daun. Sehingga Produksi Bawang Daun secara rata-rata merupakan produksi tertinggi di Sulawesi Utara, kemudian diikuti dengan Produksi Tomat secara rata-rata merupakan produksi tertinggi kedua sedangkan Produksi Cabai Rawit merupakan produksi tertinggi ketiga. Untuk produksi terendah secara rata-rata ialah Produksi Kacang Merah. Simpangan Baku atau Standar deviasi bisa juga didefinisikan sebagai, rata-rata jarak penyimpangan titik-titik data diukur dari nilai rata-rata data. Simpangan baku tertinggi terdapat pada Produksi Bawang Daun, tertinggi kedua pada Produksi Tomat dan standar deviasi terendah pada produksi Bayam.  $Simpangan\ Baku = \sqrt{Var}$ , dimana varians atau ragam suatu peubah acak (Distribusi Probabilitas) adalah ukuran seberapa jauh sebuah kumpulan bilangan tersebar. Minimum atau nilai terkecil, terdapat beberapa nilai nol yaitu pada Produksi Kangkung, Produksi Cabai Besar, Produksi Ketimun, Produksi Labu Siam, Produksi Bayam, Produksi Kacang Merah dan Produksi Bawang Daun. Sedangkan Maximum atau nilai terbesar, terdapat pada Produksi Bawang Daun, Produksi Tomat dan Produksi Cabai Rawit.

**4.2. Hasil Analisis Gerombol**

Pada hasil penelitian Analisis Gerombol ini dilakukan dengan menggunakan jarak Euclidean dan Metode Pautan Tunggal, Pautan Lengkap, Pautan Rata-rata, Metode Centroid, dan Metode Ward. Pengelompokan produksi hortikultura dilakukan dengan menggunakan aplikasi Minitab 17.

**4.2.1. Metode Ward**

Hasil pengerombolan Metode Ward berdasarkan distance level 2,4545 step ke-9 terdapat 6 gerombol yang disajikan pada Gambar 2.



**Gambar 2.** Dendrogram Hasil Analisis Gerombol dengan menggunakan Metode Ward.

Pemotongan dendrogram pada metode pautan tunggal, pautan lengkap, pautan rata-rata dan metode centroid sama seperti metode ward dengan melihat distance level terjauh. Berdasarkan jarak terjauh metode pautan tunggal terdapat pada step ke-10 dengan distance level 1,60697 menghasilkan 5 gerombol, pautan lengkap dengan distance level 2,79180, pautan rata-rata dengan distance level 2,33341 dan metode ward dengan distance level 1,85525 menghasilkan 5 gerombol.

**4.3. Perbandingan Antara Pengelompokan Analisis Gerombol Hirarki**

Untuk mengetahui metode terbaik dalam analisis pengerombolan hirarki, digunakan CTM (*Cluster Tightness Measure*), dimana semakin kecil nilai keragamannya, maka hasil pengelompokkan semakin baik. Jika CTM bernilai nol, artinya semua data berada pada pusat kelompok.

Dari proses pengerombolan menggunakan ke-5 metode hirarki yang telah dijelaskan sebelumnya. Langkah selanjutnya dihitung nilai CTM, Sebelum menghitung nilai CTM terlebih dahulu mencari nilai simpangan baku pada kelompok ke-*t* untuk variabel ke-*m* ( $s_{tm}$ ) dan mencari nilai simpangan baku variabel ke-*m* ( $s_m$ ).

**4.4. Pemilihan Metode Terbaik Menggunakan Rumus CTM (Cluster Tightness Measure)**

Sebelum menentukan pemilihan metode terbaik, terlebih dahulu dilakukan pembentukan Gerombol seperti yang sudah di jelaskan pembentukan gerombol dilakukan dengan pemotongan dendrogram yang dilihat berdasarkan distance level terjauh antar simpul yang satu dengan simpul yang lain. Hasil Pengerombolan disajikan pada Tabel 4.

**Tabel 4.** Pengelompokan Produksi Hortikultura

METODE	GEROMBOL					
	I	II	III	IV	V	VI
Pautan Tunggal	1,3,13,4,9,11,7,8,12,15,10	6	2	5	14	
Pautan lengkap	1,3,13,4,9,11,7,8,12,15,10	6	5	2	14	
Pautan rata-rata	1,3,13,4,9,11,7,8,12,15,10	6	5	2	14	
Metode Centroid	1,3,13,4,9,11,7,8,12,15,10	6	5	2	14	
Metode Ward	1,3,13,4,9,11	7,8,12,15,10	2	6	14	5

4.4.1. Metode Ward

Berdasarkan jarak terjauh Metode Ward menghasilkan 6 Gerombol, Gerombol pertama terdapat 6 kelompok dan Gerombol kedua terdapat 5 kelompok sedangkan Gerombol 3, Gerombol 4, Gerombol 5 dan Gerombol 6 hanya menghasilkan 1 kelompok. Perhitungan Pemilihan Metode terbaik untuk Metode Ward sama dengan Metode Pautan Tunggal.

➤ Simpangan baku pada kelompok ke-*t* untuk variabel ke- *m* (*S<sub>tm</sub>*)

• Gerombol 1

$$S_1 = \sqrt{\frac{(X_1 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n-1}}$$

$$S_1 = \sqrt{\frac{(17.939-9.278,833)^2 + \dots + (558-9.278,833)^2}{6-1}}$$

$$S_1 = \sqrt{\frac{435.751.678,8}{5}}$$

$$S_1 = 9.335,434418$$

Untuk hasil keseluruhan nilai *S<sub>1</sub>*, ..., *S<sub>11</sub>*. Terdapat pada Tabel 5.

Tabel 5. Simpangan Baku seluruh variabel pada Gerombol 1 dengan menggunakan Metode Ward.

Simpangan Baku	Nilai
S1	9.335,434418
S2	2.855,469308
S3	658,292868
S4	1.832,009762
S5	342,8308135
S6	1.100,285493
S7	735,8868572
S8	598,9836114
S9	512,5346492
S10	414,1982215
S11	27.626,5122

• Gerombol 2

$$S_1 = \sqrt{\frac{(X_1 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n-1}}$$

$$S_1 = \sqrt{\frac{(67-62,8)^2 + \dots + (36-62,8)^2}{5-1}}$$

$$S_1 = \sqrt{\frac{13.278,8}{4}}$$

$$S_1 = 57,617$$

Untuk hasil keseluruhan nilai *S<sub>1</sub>*, ..., *S<sub>11</sub>*, terdapat pada Tabel 6.

Perhitungan simpangan baku seluruh variabel (*S<sub>tm</sub>*) metode ward gerombol 1 dapat dilihat pada tabel 7 dan gerombol 2 pada tabel 8, sedangkan gerombol 3, gerombol 4, gerombol 5 dan gerombol 6 pada metode ward hanya terdapat 1 kelompok yang mengakibatkan

nilai simpangan baku seluruh variabel (*S<sub>tm</sub>*) menghasilkan 0.

Tabel 6. Simpangan Baku seluruh variabel pada Gerombol 2 dengan menggunakan Metode Ward.

Simpangan Baku	Nilai
S1	57,61683782
S2	2.120,046462
S3	189,8768022
S4	39,7893253
S5	20,90932806
S6	34,7807418
S7	735,8868572
S8	7,155417528
S9	10,91787525
S10	0
S11	5,196152423

➤ Perhitungan  $\frac{1}{P} \sum_{m=1}^P \frac{S_{tm}}{S_m}$  pada Gerombol I, Gerombol II, Gerombol III, Gerombol IV, Gerombol V dan Gerombol VI.

• Gerombol 1

$$\frac{1}{P} \sum_{m=1}^P \frac{S_{tm}}{S_m} = \frac{1}{11} \left( \frac{9.335,434}{36.950,860} + \dots + \frac{27.626,5122}{59.829,465} \right)$$

$$= \frac{1}{11} ((0,253 + \dots + 0,462))$$

$$= \frac{1}{11} ((2,847))$$

$$= 0,259$$

• Gerombol 2

$$\frac{1}{P} \sum_{m=1}^P \frac{S_{tm}}{S_m} = \frac{1}{11} \left( \frac{57,617}{36.950,860} + \dots + \frac{5,196}{59.829,465} \right)$$

$$= \frac{1}{11} ((0,002 + \dots + 8,685))$$

$$= \frac{1}{11} ((0,226))$$

$$= 0,021$$

Pada Gerombol 3, Gerombol 4, Gerombol 5 dan Gerombol 6 dengan memasukan nilai kedalam rumus  $\frac{1}{P} \sum_{m=1}^P \frac{S_{tm}}{S_m}$  menghasilkan nilai 0 karena nilai *S<sub>tm</sub>* pada Gerombol 2, Gerombol 3, Gerombol 4, Gerombol 5 dan gerombol 6 menghasilkan 0 yang jika dibagikan angka berapapun hasilnya pasti 0.

➤ Perhitungan nilai CTM

$$CTM = \frac{1}{K} \sum_{t=1}^K \left( \frac{1}{P} \sum_{m=1}^P \frac{S_{tm}}{S_m} \right)$$

$$CTM = \frac{1}{6} (0,252645 + 0,001559 + 0 + 0 + 0 + 0)$$

$$CTM = 0,046549$$

Jadi perhitungan nilai CTM dengan menggunakan metode pautan tunggal adalah 0,054448 dan metode ward adalah 0,046549 sedangkan metode pautan lengkap, pautan rata-rata dan metode centroid menghasilkan nilai yang sama dengan metode pautan tunggal karena terdapat pengelompokan yang juga sama. Hasil perbandingan disajikan pada Tabel 7.

**Tabel 7.** Nilai CTM dengan menggunakan Metode Pautan Tunggal, Pautan Lengkap, Pautan Rata-rata, Metode Centroid dan Metode Ward.

Metode	Nilai CTM
Pautan Tunggal	0,054448
Pautan Lengkap	0,054448
Pautan Rata-rata	0,054448
Metode Centroid	0,054448
Metode Ward	0,046549

Berdasarkan tabel 9, nilai CTM menunjukkan bahwa tidak besar pengaruh nilai CTM antar metode, sehingga dapat dikatakan metode pautan tunggal, pautan lengkap, pautan rata-rata, metode centroid dan metode ward memiliki kinerja yang sama baiknya.

#### 4.5. Interpretasi Hasil Analisis Gerombol Kabupaten/Kota pada Produksi Hortikultura

Setelah melakukan perbandingan pengelompokan analisis gerombol hirarki, maka selanjutnya Menginterpretasikan dengan menggunakan metode ward pada produksi hortikultura. Pada metode ward terdapat 6 gerombol bisa dilihat pada tabel 8.

**Tabel 8.** Penggerombolan Produksi hortikultura dengan menggunakan Metode Ward.

	Metode Ward
Gerombol 1	Kabupaten Bolaang Mongondow, Kabupaten Kepulauan Sangihe, Kota Bitung, Kabupaten Talaud, Kabupaten Minahasa Tenggara, Kabupaten Bolaang Mongondow Timur
Gerombol 2	Kabupaten Bolaang Mongondow Utara, Kabupaten Kepulauan Sitaro, Kota Manado, Kota Kotamobagu dan Kabupaten Bolaang Mongondow Selatan
Gerombol 3	Kabupaten Minahasa
Gerombol 4	Kabupaten Minahasa Utara
Gerombol 5	Kota Tomohon
Gerombol 6	Kabupaten Minahasa Selatan

Pada Tabel 10 diatas dapat dilihat bahwa, gerombol pertama terdapat 6 kabupaten/kota yaitu, Kabupaten Bolaang Mongondow, Kabupaten Kepulauan Sangihe, Kota Bitung, Kabupaten Kepulauan Talaud, Kabupaten Minahasa Tenggara, Kabupaten Bolaang Mongondow Timur. Gerombol kedua terdapat 5 kabupaten/kota yaitu, Kabupaten Bolaang Mongondow Utara, Kabupaten kepulauan Sitaro, Kota Manado, Kota Kotamobagu dan Kabupaten Bolaang Mongondow Selatan, gerombol ketiga terdapat 1 kabupaten/kota yaitu, Kabupaten Minahasa, gerombol keempat terdapat

1 kabupaten/kota yaitu Kabupaten Minahasa Utara, gerombol kelima terdapat 1 kabupaten/kota yaitu Kota Tomohon dan gerombol keenam juga terdapat 1 kabupaten/kota yaitu Kabupaten Minahasa Selatan.

Gerombol pertama secara rata-rata produksi tertinggi adalah produksi bawang baun dengan 17.749,5 kg dan produksi terendah secara rata-rata adalah produksi kacang merah dengan 293,167 kg. Gerombol kedua secara rata-rata produksi tertinggi adalah produksi cabai rawit dengan 1.298 kg, produksi terendah secara rata-rata adalah produksi bawang daun dengan 3 kg dan ada juga yang tidak memproduksi pada gerombol kedua yaitu produksi kacang merah. Gerombol ketiga produksi terbanyak ialah produksi tomat dengan 69.470 kg dan produksi cabai rawit dengan 30.589 kg, sedangkan produksi tersedikit ialah produksi bayam dengan 739 kg. Gerombol keempat produksi terbanyak ialah produksi cabai rawit dengan 16.557 kg dan produksi terung dengan 13.296 kg, sedangkan produksi tersedikit ialah produksi bayam dengan 159 kg dan produksi kacang merah dengan 330 kg, adapun produksi bawang daun yang tidak memproduksi pada gerombol 4 Kabupaten Minahasa Utara. Gerombol kelima produksi terbanyak ialah produksi labu siam dengan 38.455 kg dan produksi tomat dengan 32.186 kg, sedangkan produksi tersedikit ialah produksi kacang merah dengan 180 kg dan produksi cabai rawit dengan 309 kg. Gerombol keenam produksi terbanyak ialah produksi bawang daun dengan 230.200 kg dan produksi tomat dengan 135.337 kg, sedangkan produksi tersedikit ialah produksi bayam dengan 147 kg.

## 5. Penutup

### 5.1. Kesimpulan

- Analisis penggerombolan hirarki terdapat 5 gerombol pada metode pautan tunggal, pautan lengkap, pautan rata-rata dan metode centroid sedangkan pada metode ward terdapat 6 gerombol. Metode ward gerombol 1 terdapat Kabupaten Bolaang Mongondow, Kabupaten Kepulauan Sangihe, Kota Bitung, Kabupaten Kepulauan Talaud, Kabupaten Minahasa Tenggara, Kabupaten Bolaang Mongondow Timur, gerombol 2 terdapat Kabupaten Bolaang Mongondow Utara, Kabupaten Kepulauan Sitaro, Kota Manado, Kota Kotamobagu, Kabupaten Bolaang Mongondow Selatan, gerombol 3 terdapat Kabupaten Minahasa, gerombol 4 terdapat Kabupaten Minahasa Utara, gerombol 5 terdapat Kota Tomohon, dan gerombol 6 terdapat Kabupaten Minahasan Selatan.
- Berdasarkan perbandingan penggerombolan hirarki dengan menggunakan rumus CTM (*Cluster Tightness Measure*) pada metode pautan tunggal, pautan lengkap, pautan rata-rata dan metode centroid menghasilkan 0,054448 sedangkan pada metode ward menghasilkan 0,046549. Jadi dapat di simpulkan metode ward adalah metode yang terbaik karena menghasilkan nilai yang lebih kecil.

### 5.2. Saran

Pemerintah harus menyediakan lahan lebih pada Kabupaten Bolaang Mongondow dan Kabupaten Bolaang Mongondow Timur agar produksi hortikultura pada kabupaten tersebut semakin meningkat.

REFERENSI

[1] Mangunwidjaja, D, Sailah, I. 2009. Pengantar Teknologi Pertanian. Penebar Swadaya. Bogor.

[2] Bukhori, M. 2014. Sektor Pertanian Terhadap Pembangunan di Indonesia. [Skripsi]. Fakultas Pertanian Universitas Pembangunan Nasional "Veteran". Surabaya.

[3] Mattjik. A.A., dan I. M. Sumertajaya. 2011. Sidik Peubah Ganda. Departemen Statistik FMIPA-IPB. Bogor.

[4] Suharleni, F. 2003. Analisis Gerombol (Cluster Analysis) untuk Penggerombolan Kecamatan di Kabupaten Malang Berdasarkan Status Kemiskinan. [Skripsi]. Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

[5] Johnson, Richard A., dan Dean, Wichern, W. 2002. *Applied Multivariate Statistical Analysis 6<sup>th</sup>* edition. Prentice-Hall. London.

[6] Roux, M. 1991. Basic Procedures in Hierarchical Cluster Analysis, Devillers, J., and W. Karcher. Editor, *Applied multivariate Analysis in SAR & Environmental Studies*. Kluwer Academic Publisher. Dordrecht.

[7] Mongi, C., 2015, "Penggunaan Analisis Two Step Clustering Untuk Data Campuran", Jurnal deCartesian, 4(1) : 9-19.

[8] Simamora. 2005. *Analisis Multivariat Pemasaran Edisi Pertama*. PT. Gramedia Pustaka utama: Jakarta.

[9] Anderberg MR. 1973. *Cluster Analysis for Application*. Academic Press, New York.

[10] Gudono, Ph. D. 2011. Analisis Data Multivariat Edisi Pertama. BPFE. Yogyakarta.

[11] Handoyo, R. 2014. Perbandingan Metode clustering Menggunakan Metode Single Linkage Dan K-Means pada Pengelompokan Dokumen. Jurnal SIFO Mikroskil. 15(2):73-76.

[12] Hardle, W. 2007. *Applied Multivariate Statistical Analysis*. Berlin Springer. New York.

[13] Singgih, S. 2002. Buku Latihan SPSS Statistik Multivariat. PT Elex Komputindo. Jakarta.

[14] Sarfia. 2016. Penerapan Metode Ward pada Analisis Gerombol Hirarki (Studi kasus: Komoditas peternakan di Sulawesi Tenggara). [Skripsi]. Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Halu Oleo. Kendari.

[15] Hardius, U., dan Nurdin, S. 2013. *Aplikasi Teknik Multivariate Untuk Riset Pemasaran*. PT Grafindo Persada. Jakarta.

[16] Novita, Y. 2013. *Perbandingan jumlah kelompok optimal Single Linkage dan Complete Linkage dengan indeks Validitas silhouette (studi kasus data pembangunan manusia jawa timur)*.

[Skripsi]. Universitas Negeri Malang. Malang.

[17] BPS. 2017. Statistik Pertanian. Badan Pusat Statistik. Sulawesi Utara.

**Eva Damayanti** ([evhadamayanti31@gmail.com](mailto:evhadamayanti31@gmail.com))



Lahir di Lamongan, Jawa Timur pada tanggal 23 Oktober 1997. Menempuh pendidikan tinggi di Jurusan Matematika FMIPA, Universitas Sam Ratulangi Manado. Tahun 2019 adalah tahun terakhir ia menempuh studi. Makalah ini merupakan hasil penelitian skripsinya yang dipublikasikan.

**Charles E. Mongi** ([charlesmongi@unsrat.ac.id](mailto:charlesmongi@unsrat.ac.id))



Lahir di Tondano, Minahasa Pada tanggal 4 Januari 1984. Pada tahun 2006 memperoleh gelar sarjana pada program studi Matematika, Universitas Sam Ratulangi Manado, gelar Magister Sains diperoleh dari Institut Pertanian Bogor pada tahun 2014. Dan menjadi dosen di Jurusan Matematika FMIPA Universitas Sam Ratulangi Manado Pada tahun 2008 sampai sekarang.

**Jantje D. Prang** ([jantjedprang@yahoo.com](mailto:jantjedprang@yahoo.com))



Lahir pada tanggal 20 Desember 1958. Pada tahun 1983 mendapatkan gelar Sarjana S1 (Drs) yang diperoleh dari IKIP Negeri Manado. Gelar Magister Sains (M.Si) diperoleh dari Institut Pertanian Bogor pada tahun 2006. Ia bekerja di UNSRAT di Program Studi Matematika sebagai pengajar akademik tetap UNSRAT.