

Pengelompokan Kecamatan Berdasarkan Produksi Hortikultura Menggunakan Analisis Klaster Hierarki di Kabupaten Toraja Utara

Hasraty Paongan, Charles Eferaim Mongi*, Jantje D Prang

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

*Corresponding author: mongicharles@unsrat.ac.id

Abstrak

Tujuan dari penelitian ini adalah untuk mengetahui hasil pengelompokan kecamatan berdasarkan potensi produksi hortikultura di kabupaten Toraja Utara menggunakan analisis klaster hierarki metode *complete linkage* dan untuk mengetahui karakteristik setiap kelompok klaster yang terbentuk. Analisis klaster yaitu analisis yang mengelompokkan elemen yang mirip sebagai objek penelitian menjadi klaster yang berbeda dan klaster saling meniadakan (*mutually exclusive*). Analisis klaster terbagi atas dua metode yaitu hierarki dan non hierarki. Pada penelitian ini digunakan metode hierarki yaitu *complete linkage*. Metode *complete linkage* adalah proses clustering yang didasarkan pada jarak terjauh antar objeknya, kemudian diterapkan dalam pengelompokan kecamatan di kabupaten Toraja Utara berdasarkan produksi hortikultura dan menggunakan jarak Manhattan. Hasil pengelompokan dari penelitian ini ada 5 klaster, klaster 1 terdiri dari 7 kecamatan, klaster 2 terdiri dari 1 kecamatan, klaster 3 terdiri dari 2 kecamatan, klaster 4 terdiri dari 10 kecamatan dan klaster 5 terdiri dari 1 kecamatan dan pada klaster 1 pada produksi tomat dan buncis adalah nilai rata-rata tertinggi di antara variabel tiap klaster dan untuk produksi labu siam adalah nilai rata-rata tertinggi pada klaster 1, klaster 2 pada produksi labu siam adalah nilai rata-rata tertinggi pada klaster 2, klaster 3 pada produksi cabai rawit dan kacang Panjang adalah nilai rata-rata tertinggi di antara variabel tiap klaster dan untuk produksi labu siam merupakan nilai rata-rata tertinggi pada klaster 3, klaster 4 pada produksi terung adalah nilai rata-rata tertinggi di antara variabel tiap klaster dan untuk produksi cabai besar adalah nilai rata-rata tertinggi antar variabel dan klaster dan untuk klaster 5 pada produksi bayam dan bawang daun adalah nilai rata-rata tertinggi di antara variabel tiap klaster dan untuk produksi cabai rawit merupakan adalah nilai rata-rata tertinggi pada klaster 5.

Kata kunci: Klaster; hirarki; *complete linkage*

Clustering Districts Based on Horticultural Production Using Hierarchical Cluster Analysis in North Toraja District

Abstract

This study aimed to determine the results of sub-district grouping based on horticultural production potential in the North Toraja district using hierarchical cluster analysis of the complete linkage method and to determine the characteristics of each cluster group formed. Cluster analysis is an analysis that groups similar elements as research objects into different clusters and mutually exclusive clusters. Cluster analysis is divided into two methods, namely hierarchical and non-hierarchical. In this study, a hierarchical method was used, namely *complete linkage*. The *complete linkage* method is a clustering process based on the furthest distance between objects, then applied in grouping districts in North Toraja district based on horticultural production and using Manhattan distance. The grouping results of this study are 5 clusters, cluster 1 consists of 7 districts, cluster 2 consists of 1 district, cluster 3 consists of 2 districts, cluster 4 consists of 10 districts and cluster 5 consists of 1 district and in cluster 1 on tomato production and beans is the highest average value among the variables of each cluster and for the production of chayote is the highest average value in cluster 1, cluster 2 for chayote production is the highest average value in cluster 2, cluster 3 for cayenne pepper production and Long bean is the highest average value among the variables of each cluster and for the production of chayote is the highest average value in cluster 3, cluster 4 in eggplant production is the highest average value among the variables of each cluster and for large chili production is the highest average value between variables and clusters and for cluster 5 in the production of spinach and scallions is the highest average value among the variables of each cluster and for cayenne pepper production is the highest average value in cluster 5.

Keywords: Cluster, hierarchy, complete linkage

PENDAHULUAN

Hortikultura (horticulture) berasal dari bahasa latin hortus (tanaman kebun) dan cultura/colere (budidaya), dan dapat diartikan sebagai budidaya tanaman kebun. Kemudian hortikultura digunakan secara lebih luas bukan hanya untuk budidaya di kebun (Zulkarnain, 2009).

Clustering merupakan aktivitas (*task*) yang bertujuan mengelompokkan data yang memiliki kemiripan antara satu data dengan data lainnya ke dalam kluster atau kelompok sehingga data dalam satu kluster memiliki tingkat kemiripan (*similarity*) yang maksimum dan data antar kluster memiliki kemiripan yang minimum. Analisis kluster bertujuan menemukan kelompok objek sedemikian rupa sehingga objek-objek dalam grup akan sama (atau terkait) satu sama lain dan berbeda dari (atau tidak terkait) objek-objek dalam grup lain (Tan *et al*, 2018).

Penelitian tentang analisis kluster pernah dilakukan oleh Alfrina *et al* (2019) tentang analisis cluster terhadap prestasi mahasiswa jurusan matematika fakultas matematika dan ilmu pengetahuan alam Universitas Sam Ratulangi berdasarkan nilai akhir mata kuliah wajib tahun 2018. Pengelompokan kecamatan di kabupaten Toraja Utara berdasarkan hasil produksi pertanian menggunakan analisis gerombol (Sambaru *et al*, 2022). Perbandingan metode Single Linkage, Complete Linkage dan Average Linkage dalam pengelompokan kecamatan berdasarkan variabel jenis ternak Kabupaten Sidoarjo (Mu'afa dan Ulinuha, 2019) berdasarkan hasil penelitian perbandingan metode Single Linkage, Complete Linkage dan Average Linkage, metode Complete Linkage yang paling baik dari ketiga metode yang digunakan di lihat dari nilai rasio simpangan baku Sw terhadap Sb, dan penelitian tentang pengelompokan kecamatan di Kabupaten Minahasa berdasarkan data hasil produksi pertanian tahun 2019 dengan menggunakan analisis komponen utama (aku) dan analisis gerombol pernah dilakukan oleh Makuan *et al* (2022).

TINJAUAN PUSTAKA

Hortikultura

Hortikultura (horticulture) berasal dari bahasa latin hortus (tanaman kebun) dan cultura/colere (budidaya), dan dapat diartikan sebagai budidaya tanaman kebun. Kemudian hortikultura digunakan secara lebih luas bukan hanya untuk budidaya di kebun. Istilah hortikultura digunakan pada jenis tanaman yang dibudidayakan. Bidang kerja hortikultura meliputi pembenihan, pembibitan, kultur jaringan, produksi tanaman, hama dan penyakit, panen, pengemasan dan distribusi. Hortikultura merupakan salah satu metode budidaya pertanian modern (Zulkarnain, 2009).

Analisis Kluster dengan Metode Hierarki

Analisis kluster yaitu analisis untuk mengelompokkan elemen yang mirip sebagai objek penelitian menjadi kluster yang berbeda dan kluster saling meniadakan (*mutually exclusive*). Adapun ciri-ciri kluster adalah homogenitas (kesamaan) yang tinggi antar anggota dalam satu kluster (*within-cluster*) dan heterogenitas (perbedaan) yang tinggi antar kluster yang satu dengan kluster yang lain (*between-cluster*).

Metode Penggabungan (*Agglomerative*)

Klasterisasi adalah metode analisis data yang sering dimasukkan sebagai salah satu metode yang bertujuan untuk mengelompokkan data dengan karakteristik yang sama kedalam wilayah yang sama dan karakteristik yang beda kedalam wilayah yang lain. Klasterisasi Agglomerative merupakan bagian dari clustering berdasarkan hirarki. Klasterisasi Agglomerative dimulai dengan membuat setiap data merupakan sebuah kluster. Kluster-kluster tersebut akan bergabung berdasarkan jarak antar kluster. Penggabungan kluster akan terus terjadi sampai kluster yang dihasilkan sejumlah yang diinginkan (Aggarwal, 2014).

Mengukur Jarak Antar Objek dengan Menggunakan *Manhattan Distance*

Manhattan distance adalah metode perhitungan jarak pada ruang jarak dengan menerapkan konsep selisih mutlak (Buaton et al, 2016). Jarak *Manhattan* dirumuskan sebagai berikut (Mongi, 2015):

$$d_{i,j} = \sum_{k=1}^p |X_{ik} - X_{jk}| \quad (1)$$

Keterangan:

- $d_{i,j}$ adalah jarak antara data ke- i dengan data ke- j
- $X_{i,k}$ adalah nilai data ke- i pada variabel ke- k
- $X_{j,k}$ adalah data ke- j pada variabel ke- k
- n adalah jumlah variabel

Metode Complete Linkage (Farthest Neighbor)

Complete Linkage adalah proses clustering yang didasarkan pada jarak terjauh antar objeknya (maksimum distance). Metode ini baik untuk kasus clustering dengan normal data set distribution. Complete Linkage disebut juga pendekatan tetangga terjauh. Dasarnya adalah jarak maksimum, dimana seluruh objek dalam suatu kluster dikaitkan satu sama lain pada suatu jarak maksimum atau dengan kesamaan minimum (Tiro dan Aswi, 2010). Rumus yang digunakan adalah

$$d_{(uv)w} = \max \{d_{uw}, d_{vw}\} \quad (2)$$

Keterangan:

- d_{uw} adalah jarak terdekat dari U dan W
- d_{vw} adalah jarak terdekat dari V dan W

METODOLOGI PENELITIAN

Waktu dan Tempat

Penelitian ini dilakukan dari bulan Oktober 2022 sampai Mei 2023, di rumah dan di Perpustakaan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sam Ratulangi Manado.

Sumber Data

Data yang digunakan dalam penelitian ini yaitu data sekunder mengenai hasil produksi hortikultura yaitu data yang dipublikasikan oleh Badan Pusat Statistik (BPS) Kabupaten Toraja Utara yaitu Toraja Utara dalam angka 2022.

Objek dan Variabel Penelitian

Objek dalam penelitian ini adalah seluruh kecamatan di Kabupaten Toraja Utara yang terdiri dari 21 kecamatan, yaitu: Sopai, Kesu, Sanggalangi, Buntao, Rantebua, Nanggala, Tondon, Tallunglipu, Rantepao, Tikala, Sesean, Balusu, Sa'dan, Bangkelekila, Sesean Suloara, Kapala Pitu, Dende Piongan Napo, Awan Rante Karua, Rindingallo, Buntu Pepasan, dan Baruppu.

Variabel yang akan diamati dalam hasil produksi sektor hortikultura antara lain produksi cabai rawit (X_1), produksi kacang panjang (X_2), produksi tomat (X_3), produksi labu siam (X_4), produksi terung (X_5), produksi cabai besar (X_6), produksi buncis (X_7), produksi bayam (X_8), produksi bawang daun (X_9). Dikarenakan produksi sektor pertanian khususnya sektor pertanian hortikultura banyak data yang kosong sehingga peneliti hanya menggunakan 9 variabel yang memiliki produksi yang mencukupi.

Metode Analisis Data

Berikut analisis data:

1. Mengumpulkan data produksi Hortikultura tahun 2021.
2. Melakukan Deskripsi Data.
3. Menghitung jarak menggunakan jarak Manhattan.
4. Membentuk Kelompok Kecamatan di Kabupaten Toraja Utara dengan metode Complete Linkage.

5. Menentukan jumlah kluster dan anggota kluster yang terbentuk.
6. Menghitung rata-rata dan kemudian menentukan nilai terbesar dari rata-rata pada masing-masing variabel pada setiap kluster yang akan dijadikan sebagai karakteristik kluster tersebut.
7. Menarik kesimpulan.

HASIL DAN PEMBAHASAN

Deskripsi Data

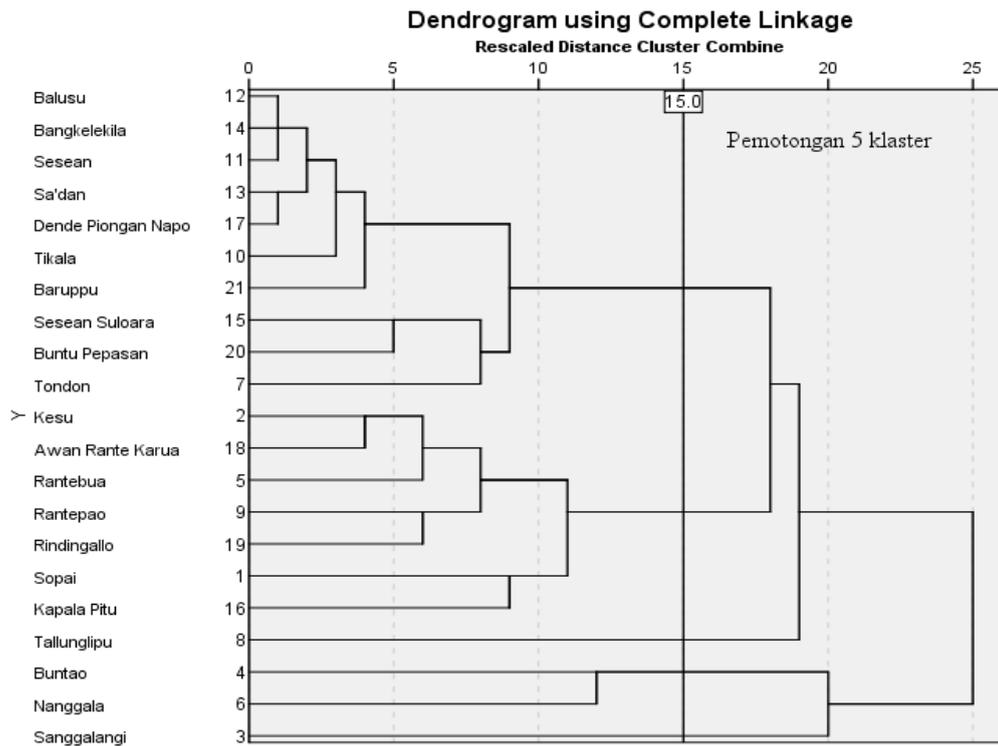
Tabel 1 merupakan hasil analisis statistik yang digunakan untuk melihat gambaran dari data sektor produksi hortikultura Kabupaten Toraja Utara. Analisis statistik deskriptif meliputi nilai minimum, nilai maksimum, rata-rata dan standar deviasi.

Tabel 1. Analisis Deskripsi Sektor Produksi Hortikultura Kabupaten Toraja Utara

	Statistik Deskriptif				
	N	Minimum	Maximum	Rata-rata	Standar Deviasi
X_1	21	8	2394	589.00	706.435
X_2	21	12	1849	504.19	498.663
X_3	21	0	1845	388.38	472.748
X_4	21	0	3740	772.76	932.555
X_5	21	0	1390	435.86	399.206
X_6	21	0	2160	411.86	556.140
X_7	21	0	970	320.24	295.750
X_8	21	0	1069	377.67	348.357
X_9	21	0	977	277.71	273.715

Pengelompokan Menggunakan Metode *Complete Linkage*

Pada proses pengelompokan ini digunakan metode complete linkage. Pada metode ini dilakukan setelah jarak sudah terbentuk, di mana pada penelitian ini digunakan Manhattan distance. Karena pengelompokan dimulai dari jarak terdekat, maka dimulai dari data (objek) kecamatan Balusu dan kecamatan Bangkelekila atau data ke 12 dan 14 dengan jarak sebesar 285 yang dapat dilihat pada lampiran 2. Dendrogram yang diperoleh pada penelitian ini disajikan dalam Gambar 1.



Gambar 1. Dendogram Metode Complete Linkage menggunakan jarak Manhattan

Pada metode complete linkage, pemotongan dendogram terjadi di klaster ke 5 dikarenakan jarak antara klaster 5 dan 4 sudah memiliki jarak yang lumayan jauh dan juga klaster 5 ke 6 memiliki jarak yang lumayan jauh. Ada 4 kelompok gerombol tersebut adalah sebagai berikut;

- Klaster 1 terdiri dari 7 kecamatan yaitu Sopai, Kesu, Rantebua, Rantepao, Kapala Pitu, Awan Rante Karua, Rindingallo.
- Klaster 2 terdiri dari 1 kecamatan yaitu Sanggalangi.
- Klaster 3 terdiri dari 2 kecamatan yaitu Buntao, Nanggala.
- Klaster 4 terdiri dari 10 kecamatan yaitu Tondon, Tikala, Sesean, Balusu, Sa'dan, Bangkelekila, Sesean Suloara, Dende Pongan Napo, Buntu Pepsan, Baruppu.
- Klaster 5 terdiri dari 1 kecamatan yaitu Tallunglipu.

Karakteristi Klaster

Proses profiling data dengan menerapkan proses rata-rata di setiap klaster. Proses profiling ini menghasilkan ciri-ciri spesifik atau karakteristik pada setiap klaster yang terbentuk.

Tabel 2. Nilai Rata-rata Produksi untuk Tiap Klaster

	Klaster				
	1	2	3	4	5
Cabe Rawit (X_1)	115	328	2394	1800	1404
Kacang Panjang (X_2)	1205	302	1849	1470	615
Tomat (X_3)	1845	209	153	780	1014
Labu siam (X_4)	2025	420	3740	2110	815
Terung (X_5)	1165	279	665	1390	734
Cabai Besar (X_6)	105	252	0	2160	1177
Buncis (X_7)	970	223	518	445	615
Bayam (X_8)	475	305	115	535	967
Bawang Daun (X_9)	410	241	230	311	516

Berdasarkan Tabel 5 untuk keterangan warna kuning menunjukkan nilai rata-rata tertinggi antar klaster tiap variabel, untuk warna hijau menunjukkan nilai rata-rata tertinggi antar variabel tiap klaster dan pada warna biru menunjukkan nilai rata-rata tertinggi antar variabel dan klaster, diketahui karakteristik setiap klaster melalui nilai rata-rata setiap klaster, kemudian dihasilkan interpretasi sebagai berikut:

1. Klaster 1, pada variabel hasil produksi tomat dan buncis dengan keterangan warna hijau adalah nilai rata-rata tertinggi di antara variabel tiap klaster dan untuk produksi labu siam dengan keterangan warna kuning adalah nilai rata-rata tertinggi di klaster 1.
2. Klaster 2, pada produksi labu siam dengan keterangan warna kuning adalah nilai rata-rata tertinggi di klaster 2.
3. Klaster 3, pada variabel hasil produksi cabai rawit dan kacang Panjang dengan keterangan warna hijau adalah nilai rata-rata tertinggi di antara variabel tiap klaster dan untuk produksi labu siam dengan keterangan warna biru merupakan nilai rata-rata tertinggi antar variabel dan klaster 3.
4. Klaster 4, pada variabel hasil produksi terung dengan keterangan warna hijau adalah nilai rata-rata tertinggi di antara variabel tiap klaster dan untuk produksi cabai besar dengan keterangan warna biru adalah nilai rata-rata tertinggi antar variabel dan klaster 4.
5. Klaster 5, pada variabel hasil produksi bayam dan bawang daun dengan keterangan warna hijau adalah nilai rata-rata tertinggi di antara variabel tiap klaster dan untuk produksi cabai rawit dengan keterangan warna kuning merupakan adalah nilai rata-rata tertinggi di klaster 5.

KESIMPULAN

Berdasarkan hasil dan pembahasan mengenai pengelompokan kecamatan menurut produksi hortikultura di kabupaten Toraja Utara, maka dapat diambil kesimpulan sebagai berikut.

1. Dihasilkan 5 kelompok klaster. Klaster 1 tujuh kecamatan yaitu kecamatan Sopai, Kesu, Rantebua, Rantepao, Kapala Pitu, Awan Rante Karua dan Rindingallo, klaster 2 satu kecamatan yaitu kecamatan Sanggalangi, klaster 3 dua kecamatan yaitu kecamatan Buntao dan Nanggala, klaster 4 sepuluh kecamatan yaitu kecamatan Tondon, Tikala, Sesean, Balusu, Sa'dan, Bangkelekila, Sesean Suloara, Dende Piongan Napo, Buntu Pepasan dan Baruppu dan klaster 5 satu kecamatan yaitu kecamatan Tallunglipu.
2. Terdapat karakteristik masing-masing klaster yaitu klaster 1 pada produksi tomat dan buncis adalah nilai rata-rata tertinggi di antara variabel tiap klaster dan untuk produksi labu siam adalah nilai rata-rata tertinggi pada klaster 1, klaster 2 pada produksi labu siam adalah nilai rata-rata tertinggi pada klaster 2, klaster 3 pada produksi cabai rawit dan kacang Panjang adalah nilai rata-rata tertinggi di antara variabel tiap klaster dan untuk produksi labu siam merupakan nilai rata-rata tertinggi pada klaster 3, klaster 4 pada produksi terung adalah nilai rata-rata tertinggi di antara variabel tiap klaster dan untuk produksi cabai besar adalah nilai rata-rata tertinggi antar variabel dan klaster dan untuk klaster 5 pada produksi bayam dan bawang daun adalah nilai rata-rata tertinggi di antara variabel tiap klaster dan untuk produksi cabai rawit merupakan adalah nilai rata-rata tertinggi pada klaster 5.

DAFTAR PUSTAKA

- Aggarwal, C. Charu. 2014. *Data Clustering Algorithms and Applications*. Taylor & Francis. Minneapolis.U.S.A
- Alfrina., D, Hatidja., J, Titaley. 2019. Analisis Cluster Terhadap Prestasi Mahasiswa Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sam Ratulangi Berdasarkan Nilai Akhir Mata Kuliah Wajib Tahun 2018. *Jurnal D'cartesiaN*. 8(1): 36-44.
- BPS Kabupaten Toraja Utara. 2022. *Toraja Utara Dalam Angka 2022*. <https://torutkab.bps.go.id/subject/55/hortikultura.html> [diakses 15 November 2022].
- Buaton, R., Sundari, Y., dan Maulita, Y. 2016. Clustering Tindak Kekerasan Pada Anak Menggunakan Algoritma K-Means Dengan Perbandingan Kedekatan Manhattan City Dan Euclidean. *Jurnal Media Informasi Analisa dan Sistem*, 1(2):47-53.

- Johnson, R.A., dan Winchern, D. W. 2007. *Applied Multivariate Statistical Analysis* (6 th edition). United State of America.
- Mongi, C. 2015. Penggunaan Data Two Step Clustering untuk Data Campuran. *Jurnal d' Cartesian*, 04(1):31.
- Mu'afa, S. F dan N. Ulinnuha. 2019. Perbandingan Metode Single Linkage, Complete Linkage dan Average Linkage dalam Pengelompokan Kecamatan Berdasarkan Variabel Jenis Ternak Kabupaten Sidoarjo. *Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, 4(2):1-5.
- Sambaru, R., H.A. H. Komalig., dan C. E. Mongi. 2022. Pengelompokan Kecamatan di Kabupaten Toraja Utara Berdasarkan Hasil Produksi Pertanian Menggunakan Analisis Gerombol. *Jurnal LPPM Bidang Sains dan Teknologi*,7(1):6-10.
- Tan P, N., M. Steinbach., A, Karpatne., V, Kumar. 2018. *Introduction to Data Mining* (2nd Edition). New York.
- Tiro, A.M dan Aswi. 2010. *Statistika Deskriptif Peubah Banyak*. Andira Publisher. Makassar.
- Zulkarnain. 2009. *Dasar-dasar Hortikultura*. Bumi Aksara. Jakarta.