

# Pengelompokan Potensi Produksi Bawang Merah di Jawa Tengah Menggunakan Analisis *Linkage Hierarchical Clustering* dan *K-Means Clustering*

Auryn Clara Jangkobus<sup>1</sup>, Rahmatina Hidayati<sup>2</sup>

<sup>1,2</sup>*Sistem Informasi, Universitas Merdeka Malang. 082290559700*  
*e-mail: aurnyclara27@gmail.com<sup>1</sup>, rahmatina.hidayati@unmer.ac.id<sup>2</sup>*

---

## ABSTRAK

### **Kata Kunci:**

Bawang Merah  
*K-Means Clustering*  
*Linkage Hierarchical Clustering*

Bawang merah adalah tanaman hortikultura unggulan di Jawa Tengah yang signifikan bagi perekonomian regional. Pada 2021, produksi bawang merah mencapai 556.057 ton, turun dari 500.992 ton pada tahun sebelumnya. Ketidakefisienan pasokan sering disebabkan oleh perbedaan iklim dan cuaca. Untuk menjaga keseimbangan pasokan, permintaan, dan stabilitas harga, distribusi merata diperlukan dari daerah produksi tinggi ke rendah. Penelitian ini menggunakan metode pengelompokan (*Clustering*) untuk mengidentifikasi daerah potensial produksi bawang merah di Jawa Tengah, termasuk *Agglomerative Hierarchical Clustering* dan *K-means Clustering*. Pada 2020, *K-Means* mengelompokkan 15 Kota/Kabupaten dalam cluster 2, meningkat menjadi 25 pada 2021. *Linkage Hierarchical Clustering* mengelompokkan 28 Kabupaten/Kota dalam cluster 3. Suhu dan curah hujan signifikan mempengaruhi produksi bawang merah. Penelitian ini memberikan rekomendasi kebijakan untuk meningkatkan produksi dan kesejahteraan petani, sehingga distribusi dan produksi bawang merah dapat lebih efektif dan efisien, membantu menjaga stabilitas ekonomi regional.

## ABSTRACT

### **Keyword:**

Red Onion  
*K-Means Clustering*  
*Linkage Hierarchical Clustering*

*Red garlic is a leading horticultural crop in Central Java that is significant for the regional economy. By 2021, red garlic production had reached 556,057 metric tons, down from 500,992 metric tons the previous year. Supply inefficiency is often caused by climate and weather differences. To maintain the balance of supply, demand, and price stability, even distribution is required from high to low production areas. The research uses clustering methods to identify potential areas of red garlic production in Central Java, including agglomerative hierarchical clustering and K-means clusters. By 2020, K-Means had grouped 15 cities and districts in cluster 2, increasing to 25 by 2021. Linkage hierarchical clustering groups 28 districts or cities in cluster 3. Temperatures and rainfall have significantly affected garlic production. This study provides policy recommendations to improve farmers' production and well-being, so that the distribution and production of garlic can be more effective and efficient, helping to maintain regional economic stability.*

## I. PENDAHULUAN

Bawang merah merupakan tanaman hortikultura penting di Jawa Tengah yang berfungsi sebagai komoditas sayuran unggulan dan memiliki peran signifikan dalam perekonomian regional, serta digunakan dalam masakan sehari-hari dan obat tradisional [1]. Tanaman ini telah lama dibudidayakan oleh petani di wilayah ini dan berkontribusi pada stabilisasi harga pangan, sebanding dengan cabai dan bawang putih. Iklim di Jawa Tengah yang sangat mendukung pertumbuhan bawang merah menjadikannya sebagai salah satu produsen utama bawang merah di Indonesia. Namun, pada periode 2020-2021, harga bawang merah menurun drastis di tingkat petani akibat curah hujan tinggi yang berdampak pada kualitas bawang merah. Sebagai respons, Kepala Dinas Ketahanan Pangan Jawa Tengah meluncurkan Gerakan Peduli Petani Bawang Merah untuk membantu petani menghadapi penurunan harga dan mencegah kerugian lebih lanjut, dengan memanfaatkan teknologi informasi untuk memprediksi perubahan iklim dan meningkatkan kesejahteraan petani [2]. Data Badan Pusat Statistik tahun 2020 menunjukkan bahwa Jawa Tengah adalah pusat produksi bawang merah terbesar di Indonesia, dengan produksi mencapai 556.057 ton pada tahun 2021 dari luas tanam 53.645 hektar. Meski demikian, ada penurunan produksi dibandingkan tahun sebelumnya yang mencapai 500.992 ton, disebabkan oleh ketidakseimbangan antara produksi yang tinggi namun tidak sepenuhnya terjual serta variasi iklim di setiap wilayah. Daerah dengan iklim yang ideal untuk bawang merah umumnya memproduksi lebih banyak, sedangkan daerah dengan cuaca buruk mengalami penurunan produksi, mengakibatkan ketidakefisienan dalam distribusi. Untuk mengatasi masalah ini, penelitian ini menggunakan metode pengelompokan atau clustering untuk mengelompokkan data berdasarkan kesamaan karakteristik, dengan membandingkan dua metode utama: *Agglomerative Hierarchical Clustering* dan *K-means Clustering* [3].

Penelitian ini akan menerapkan kedua metode tersebut untuk mengelompokkan daerah di Jawa Tengah berdasarkan potensi produksi bawang merah, dengan memanfaatkan data tentang produksi, luas panen, dan kondisi iklim seperti suhu, curah hujan, dan durasi penyinaran matahari. Rumusan masalah penelitian mencakup karakteristik potensi produksi bawang merah di berbagai daerah di Jawa Tengah, hasil pengelompokan menggunakan metode *Linkage Hierarchical Clustering* dan *K-means Clustering*, serta faktor-faktor signifikan yang mempengaruhi potensi produksi. Tujuan utama penelitian ini adalah untuk menganalisis dan membandingkan potensi produksi bawang merah di berbagai daerah, serta merumuskan rekomendasi kebijakan yang dapat meningkatkan potensi produksi bawang merah di Jawa Tengah [4].

Penelitian ini didukung oleh penelitian terdahulu tentang pengelompokan telah banyak dilakukan, seperti pada penelitian [5] menggunakan *K-means Clustering* untuk mengidentifikasi wilayah Kabupaten Bima yang berpotensi area produksi dan luas panen bawang merah tinggi, sedang dan rendah yang sesuai dengan geografis dan iklim wilayahnya. Pada penelitian [6] hasil perbandingan *Single Linkage*, *Complete Linkage*, dan *Average Linkage* pada Kesejahteraan

Masyarakat di Jawa Timur dalam perhitungan analisis kesejahteraan masyarakat, dengan 3 *Cluster* merupakan yang terbaik dengan nilai *indeks silhouette* sebesar 0,6054. Serta pada penelitian [7] pengelompokan provinsi Indonesia membentuk lima *Cluster* dengan karakteristik yang berbeda masing-masing dikategorikan baik dan sangat baik, cukup dan sangat buruk dalam cakupan pelayanan kesehatan ibu.

Metode *Linkage Hierarchical Clustering* dan *K-means Clustering* digunakan dalam penelitian ini untuk mengelompokkan daerah-daerah tersebut dan mengidentifikasi faktor-faktor berpengaruh secara signifikan terhadap potensi produksi bawang merah. Tujuan akhirnya adalah merumuskan rekomendasi kebijakan yang tepat untuk meningkatkan potensi produksi tanaman bawang merah. Variabel yang menjadi fokus dalam penelitian ini adalah suhu, curah hujan, jumlah hari hujan, tekanan udara, durasi penyinaran matahari.

Oleh karena itu, penulis akan melakukan penelitian pengelompokan Potensi Produksi Bawang Merah di Jawa Tengah dengan metode *Linkage Hierarchical Clustering* dan *K-means Clustering*.

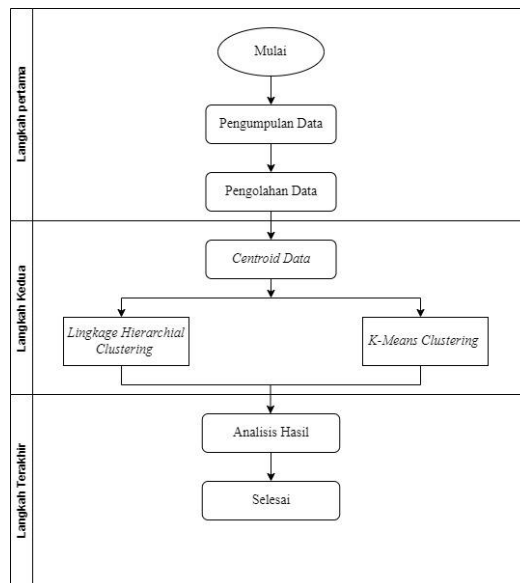
## II. METODE

Dalam penelitian ini, penulis menggunakan metode deskriptif kuantitatif untuk memahami data numerik dan menganalisis hubungan antara variabel secara objektif. Penelitian dilakukan secara virtual tanpa lokasi fisik khusus, menggunakan perangkat lunak dan perangkat keras yang tersedia. Data diperoleh dari situs web Badan Pusat Statistik (BPS) Jawa Tengah, mencakup 29 Kabupaten dan 6 Kota di Jawa Tengah.

**Tabel 1.** Karakteristik Iklim Jawa Tengah

No	Variabel	Ciri-Ciri
1.	Iklim	Tropis
2.	Suhu	22,1°C sampai 35,5°C
3.	Curah Hujan	2000 mm <sup>3</sup> per tahun
4.	Penyinaran Matahari	67%
5.	Tekanan Udara	46%-95%

Variabel yang digunakan merupakan data yang mencakup jumlah produksi bawang merah, luas panen, suhu rata-rata, durasi penyinaran matahari, curah hujan, dan hari hujan dari tahun 2020 hingga 2021 yang menjadi variabel dalam penelitian ini[8]. Proses penelitian dimulai dengan pengumpulan dan pengolahan data, diikuti dengan langkah *centroid* dan penerapan algoritma *Linkage Hierarchical Clustering* serta *K-means Clustering*, dan diakhiri dengan analisis hasil pengelompokan. Alur penelitian ditunjukkan pada gambar 1 di bawah ini.:



**Gambar 1.** Alur Penelitian

### Populasi dan Sampel

Populasi dan sampel yang digunakan dalam penelitian ini adalah keseluruhan data produksi bawang merah dari 35 Kabupaten dan Kota di Jawa Tengah tahun 2020 sampai dengan tahun 2021.

### Teknik Pengumpulan Data

Pada tahap ini, data dan informasi dikumpulkan dari berbagai sumber dengan menggunakan analisis kuantitatif dan data sekunder dari website Badan Pusat Statistik (BPS) Jawa Tengah. Data yang dikumpulkan mencakup produksi bawang merah, luas panen, dan kondisi iklim di Kabupaten dan Kota di Jawa Tengah untuk tahun 2020-2021.

### Teknik Analisis Data

Setelah data dikelompokkan, pengujian dilakukan dengan metode *silhouette coefficient* untuk menilai seberapa baik setiap objek ditempatkan dalam kluster atau kelompok. Data mining adalah proses mendapatkan informasi menarik dari data berskala besar (*big data*) dan menemukan pola tersembunyi untuk menghasilkan informasi yang berguna[9]. Salah satu alat dalam data mining adalah Analisis *Cluster*, yang mengelompokkan objek-objek berdasarkan kesamaan karakteristik. Tujuan analisis ini adalah mengumpulkan data dari berbagai sumber untuk mengelompokkan objek-objek sesuai tingkat kemiripannya, menggunakan konsep pengukuran jarak dan kesamaan antar objek[10].

- **Linkage Hierarchical Clustering**

*Linkage Hierarchical Clustering* adalah proses di mana setiap titik dianggap sebagai *cluster* tunggal dan secara bertahap menggabungkan *cluster* terdekat berdasarkan metrik jarak tertentu hingga hanya satu *cluster* yang tersisa. *Agglomerative Clustering* mengelompokkan data dari bawah ke atas, dimulai dari kelompok kecil dan menggabungkannya menjadi kelompok yang lebih besar [11]. Peneliti menggunakan metode ini untuk mengevaluasi keoptimalan hasil *cluster*, termasuk *Single Linkage*, *Average Linkage*, *Complete Linkage*, dan *Ward Method*.

Algoritma *Agglomerative Hierarchical Clustering* mencakup beberapa metode pengukuran jarak antar cluster[12]:

1. *Single Linkage* Mengukur jarak terkecil antara objek dan *cluster*

$$d_{(UV)W} = \min(d_{UW}, d_{VW}) \quad (1)$$

2. *Complete Linkage*: Mengukur jarak terjauh antara objek antar *cluster*

$$d_{(UV)W} = \max(d_{UW}, d_{VW}) \quad (2)$$

3. *Average Linkage*: Mengukur rata-rata jarak antara semua objek di satu *cluster* dengan semua objek di cluster lain.

$$d_{(UV)W} = \frac{d_{(UW)} + d_{(VW)}}{n_{(UV)}nW} \quad (3)$$

4. *Ward Linkage*: Menggunakan jarak kuadrat dalam setiap cluster untuk memastikan *cluster* yang terbentuk memiliki variasi data yang kecil dan berbeda jauh dari cluster lain

$$d_{(UV)W} = \frac{[(n_W + n_U)d_{(UW)} + (n_W + n_V)d_{(VW)}] - n_W d_{(UV)}}{n_W + n_{(UV)}} \quad (4)$$

- ***K-Means Clustering***

Tujuan *K-means Clustering* adalah membagi data menjadi satu atau lebih kelompok berdasarkan kesamaan atribut. Data dengan atribut serupa dikelompokkan bersama, sedangkan data dengan atribut berbeda ditempatkan di kelompok lain [13]. Metode ini bertujuan untuk meminimalkan kesamaan antar cluster dan memaksimalkan kesamaan di dalam *cluster*. *K-means Clustering* dapat mengelompokkan banyak data dengan cepat, tetapi hasilnya sangat dipengaruhi oleh pemilihan pusat awal setiap *cluster* [14].

Langkah-langkah *K-Means Clustering* adalah sebagai berikut:

1. Pilih jumlah *cluster* (*k*), misalnya dengan metode *silhouette*.
2. Hitung jarak antara data dan centroid menggunakan rumus jarak *Euclidean*:

$$d(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + \dots + (X_{ki} - X_{kj})^2} \quad (5)$$

Keterangan:

$d(i, j)$ : jarak data ke-*i* ke pusat Cluster *j*

$x_{ki}$  : data ke-*i* yang berkaitan dengan atribut data ke-*k*

$x_{kj}$  : titik pusat ke-*j* pada atribut data ke-*k*

3. Tempatkan data di cluster terdekat berdasarkan jarak *centroid*.
4. Hitung ulang *centroid* setiap *cluster*.
5. Proses selesai jika nilai centroid tidak berubah[3].

## Analisis Hasil

Pada tahap ini, hasil clustering dari dataset produksi bawang merah dianalisis untuk mengevaluasi efektivitas masing-masing metode. Pengelompokan terbaik dari metode *Hierarchical Clustering* dan *K-means Clustering* akan digunakan untuk mengidentifikasi wilayah dengan potensi produksi bawang merah.

### III. HASIL DAN PEMBAHASAN

#### Hasil

*Dataset* yang digunakan dalam pengumpulan data ini adalah data mentah produksi bawang merah. Data tersebut diolah oleh peneliti dan dilakukan proses clustering. Parameter dalam dataset meliputi produksi bawang merah tahun 2020-2021, luas lahan panen, suhu, curah hujan, tekanan udara, hari hujan, dan penyinaran matahari. Pengolahan data dilakukan menggunakan *Microsoft Excel* dan aplikasi *Orange* untuk mempermudah prosesnya.

No	Kabupaten/ Kota	Produksi (kuintal) 2020	Luas Panen (ha) 2020	Suhu Udara (°C) 2020	Curah Hujan (mm · ) 2020	Hari Hujan (Hari) 2020	Tekanan Udara (mb)	Penyinaran Matahari (h)
01	Cilacap	8303	68	26	4863	230	1010,7	66,70
02	Bayungas	1177	20	28	4353,5	210	0	0
03	Purbalingga	107	7	0	0	0	0	0
04	Banjarnegara	662	13	24,52	20566	222	0	3,80
05	Kebumaha	1033	16	24,63	2822	136	1004,8	0
06	Purworejo	233	3	0	32038	135	0	0
07	Wonorebo	464	3	28,8	4304	0	0	0
08	Magelang	3664	103	26	3162	188	0	0
09	Bongohli	100120	3151	27,3	2326	123	1011	0
10	Klaten	2375	36	0	2202	0	0	0
11	Sukoharjo	4860	54	0	0	0	0	0
12	Wonogiri	24783	231	0	2305	120	0	0
13	Karanganyar	36834	474	0	2363	136	0	0
14	Sragen	20575	205	26,3	2551	151	0	0
15	Grobogan	204407	2385	0	0	0	0	0

No	Kabupaten/ Kota	Produksi (kuintal) 2021	Luas Panen (ha) 2021	Suhu Udara (°C) 2021	Curah Hujan (mm · ) 2021	Hari Hujan (Hari) 2021	Tekanan Udara (mb)	Penyinaran Matahari (h)
01	Cilacap	7684	34	27,44	3050	233	1010	55
02	Bayungas	723	37	28,00	2696	118	0	0
03	Purbalingga	565	13	0,00	2731	178	0	0
04	Banjarnegara	322	7	23,34	3751	198	0	100
05	Kebumaha	2104	38	27,00	4928	191	0	42
06	Purworejo	157	3	0,00	76265	0	0	0
07	Wonorebo	1530	16	0,00	3218	111	0	0
08	Magelang	13520	160	27,00	3121	162	0	0
09	Bongohli	88600	1204	32,80	2303	160	0	0
10	Klaten	3736	55	0,00	2080	138	0	0
11	Sukoharjo	7070	79	0,00	2746	221	0	0
12	Wonogiri	17895	263	28,00	3044	141	0	0
13	Karanganyar	56804	654	0,00	1185	0	0	0
14	Sragen	22363	202	0,00	2579	152	0	0
15	Grobogan	26114	2353	0,00	4679	93	0	0

(a) Data Produksi Tahun 2020

(b) Data Produksi Tahun 2021

Gambar 2. Pengumpulan data mentah di Ms.Excel

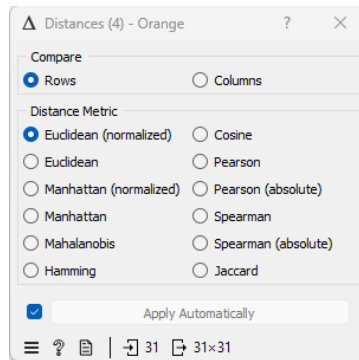
#### Preprocessing Data

Pada tahap ini, data yang dikumpulkan mengalami preprocessing untuk mempersiapkannya sebelum pemodelan menggunakan *Microsoft Excel*. Langkah-langkahnya meliputi penanganan missing value. Missing value terjadi saat nilai atau informasi pada variabel tertentu hilang. Data yang tidak digunakan mencakup produksi dan luas lahan dari kota-kota yang tidak memproduksi bawang merah. Dari 35 baris data awal, penyesuaian missing value mengurangi menjadi 31 baris. Dari 7 variabel, hanya 5 variabel yang digunakan karena tekanan udara dan penyinaran matahari memiliki banyak nilai nol dan tidak tersedia dalam data.

#### Transformasi Data

Pada tahap ini, data dinormalisasi agar nilainya berada dalam rentang tertentu, seperti -1 hingga 1 atau 0 hingga 1. Ini memastikan semua variabel memiliki skala yang sama, mencegah dominasi satu variabel, dan meningkatkan akurasi algoritma pemrosesan data. Setelah normalisasi, dilakukan pengelompokan menggunakan algoritma *Agglomerative Hierarchical Clustering* dan *K-means Clustering* dengan atribut jarak *Euclidean* (dinormalisasi) pada aplikasi *Orange*. Normalisasi bertujuan untuk menyamakan kontribusi fitur dalam perhitungan jarak, mempermudah klusterisasi

dengan menilai kesamaan data, dan mendeteksi anomali dengan mengukur jarak titik data dari pusat kluster atau rata-rata.

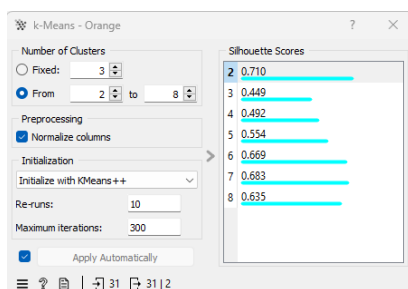


Gambar 3. Tampilan Transformasi Data

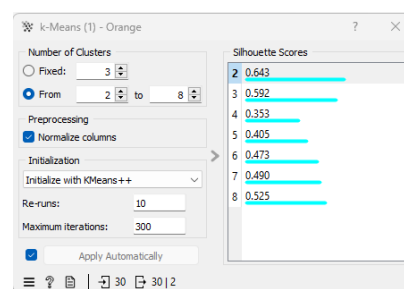
## Proses Clustering

### 1. Clustering K-Means Tahun 2020

- Penelitian ini menerapkan metode *K-Means*, *Silhouette Plot* menampilkan koefisien *silhouette* untuk setiap data cluster dan nilai rata-rata untuk seluruh *cluster*.
- Pada tahap ini, *cluster* akan ditentukan berdasarkan potensi tertinggi, sedang, dan rendah. Dalam pengaturan *K-means* di *Orange*, ditetapkan Fixed 3 untuk melihat hasil dari tiap cluster tersebut. Penentuan cluster didasarkan pada nilai *silhouette scores*, yang mengukur kualitas dan kekuatan cluster; semakin tinggi nilai *silhouette scores*, semakin baik hasilnya. Berdasarkan nilai *silhouette scores* 0,710 pada data *outlier*[15]. Jika menghapus data *outlier* (Kabupaten Brebes) maka *silhouette scores* terbaik yaitu 0,643. Berikut adalah gambar hasil penentuan *cluster*.



(a) *Silhouette score* data *outlier K-Means* 2020



(b) *Silhouette score* tanpa data *outlier K-Means* 2020

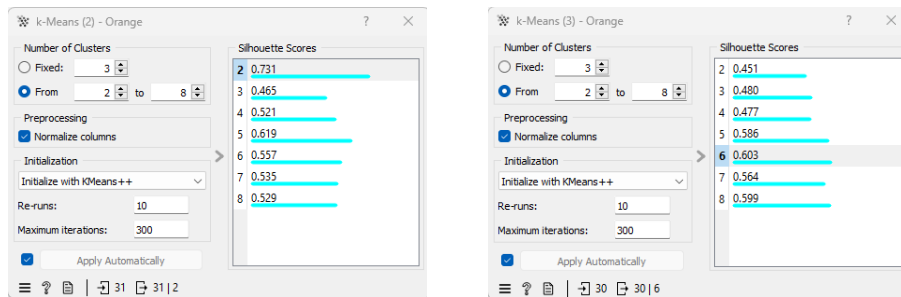
Gambar 4. Penentuan *Cluster* data *K-Means* 2020

- Clustering* pada C1 mencakup 29 Kabupaten/Kota, sedangkan C2 hanya memiliki 1 Kabupaten. Kabupaten Brebes di C2 merupakan data outlier karena produksi di Brebes jauh lebih tinggi dibandingkan Kabupaten/Kota lain. Setelah menghapus data outlier (Kabupaten Brebes), menghasilkan 11 Kabupaten/Kota tetap berada di C1, 15 Kabupaten/Kota di C2, dan 1 Kabupaten berada di C3.
- Langkah berikutnya adalah menentukan karakteristik setiap *cluster* yang telah dibentuk. Karakteristik iklim masing-masing *Cluster* tanpa data outlier pada *K-Means* 2020.C1,

memiliki curah hujan sedikit banyak dengan rata-rata 51,50 sampai 128 hari hujan dan C2, memiliki suhu udara 25,45°C-27,65°C, dengan curah hujan lebih banyak dan rata-rata 64,50 sampai 191,50 hari hujan.

## 2. Clustering K-Means Tahun 2021

- Berdasarkan nilai *silhouette scores*, Cluster yang optimal yaitu 2 Cluster dengan nilai *silhouette scores* 0,731. Jika menghapus data Kabupaten Brebes maka *silhouette scores* terbaik yaitu 0,603.



(a) Silhouette score data outlier K-Means 2021

(b) Silhouette score tanpa data outlier K-Means 2021

**Gambar 5.** Tampilan hasil clustering pada Scatterplot K-Means 2021

- Clustering pada C1 terdapat 20 Kabupaten, pada C2 terdapat 1 Kabupaten, dan pada C3 terdapat 11 Kabupaten. Cluster yang dihasilkan setelah menghapus data Kabupaten Brebes dalam Data Tabel silhouette score pada C1 (tinggi) terdapat 1 Kabupaten, pada C2 (sedang) terdapat 25 Kabupaten, dan pada C3 (rendah) terdiri dari 4 Kabupaten.
- Karakteristik masing-masing Cluster tanpa data outlier K-Means 2021

C1, memiliki curah hujan terbanyak dengan C2, memiliki suhu udara 27,10 °C, 132 sampai 171 hari hujan dan C3, memiliki curah hujan yang sedikit lebih banyak, dengan jumlah 98,50 sampai 145,50 hari hujan.

## 3. Clustering Linkage Hierarchical Clustering 2020

Selanjutnya, data produksi bawang merah dianalisis menggunakan metode *Linkage Hierarchical Clustering* dengan aplikasi khusus. *Distances* menggunakan *Dendrogram* untuk memberikan gambaran visual memudahkan penentuan jumlah cluster yang optimal. *Box Plot* membantu mengidentifikasi nilai-nilai dalam setiap cluster pada level tertentu.

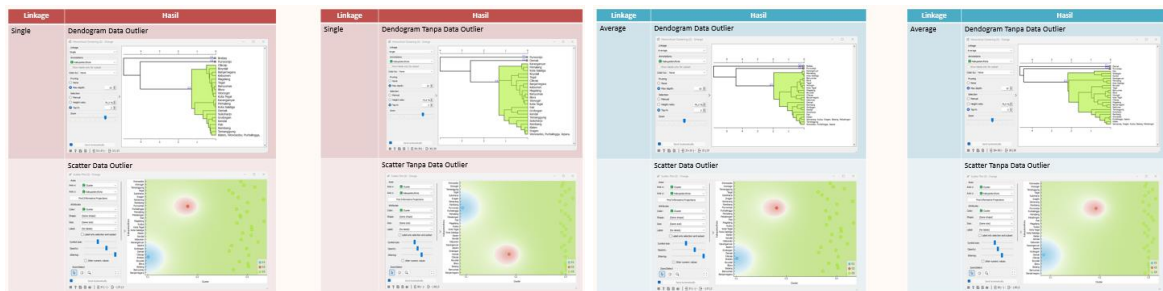
Berdasarkan hasil clustering pada Gambar 7, tahun 2020 dengan metode *Single Linkage*, *Average Linkage*, *Complete Linkage*, dan *Ward Linkage*, metode *Ward Linkage* memberikan hasil terbaik. Tanpa data outlier, *Ward Linkage* menghasilkan cluster C1 dengan 1 kabupaten, C2 dengan 15 kabupaten/kota, dan C3 dengan 15 kabupaten/kota. Dengan data outlier, hasilnya adalah C1 dengan 1 kabupaten, C2 dengan 14 kabupaten/kota, dan C3 dengan 15 kabupaten/kota. Metode *Ward Linkage* menunjukkan pembagian yang lebih seimbang antara cluster sedang dan rendah, serta lebih sensitif terhadap perubahan data, memberikan hasil yang lebih optimal dalam pengelompokan data dibandingkan metode lainnya.



## Karakteristik Iklim 2020

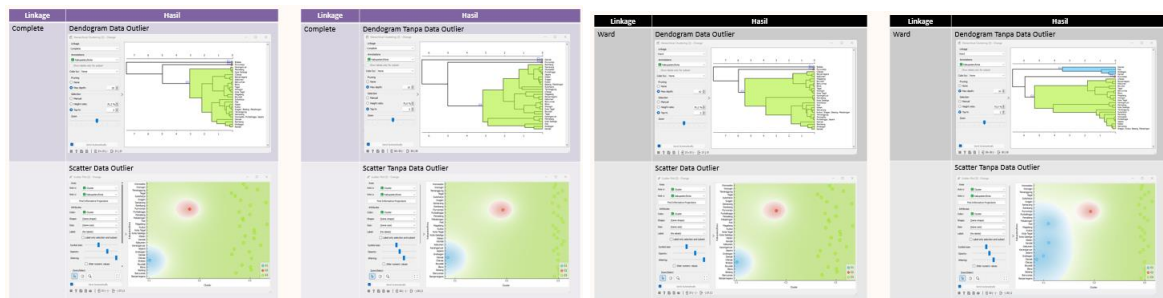
Dari hasil *clustering*, karakteristik iklim dengan metode *Single Linkage*, *Average Linkage*, *Complete Linkage*, dan *Ward Linkage*, metode *Complete Linkage* memberikan hasil terbaik dalam pengelompokan berdasarkan curah hujan dan suhu udara. *Cluster C2* dalam *Complete Linkage* memiliki curah hujan terbanyak dengan 195 hari hujan, sedangkan *cluster C3* memiliki suhu antara 12,26-26,5°C dan curah hujan yang sedikit lebih banyak dengan rata-rata 124,50 sampai 155 hari hujan. Metode ini menunjukkan variasi yang jelas dan detail dalam pengelompokan berdasarkan parameter yang dianalisis.

### 4. Clustering Linkage Clustering 2021



a) *Single Linkage* 2021

b) *Average Linkage* 2021



c) *Complete Linkage* 2021

d) *Ward Linkage* 2021

**Gambar 6.** Tampilan Hasil *Clustering Dendrogram* dan *Scatterplot Hierarchical Clustering* Tahun 2021

Berdasarkan hasil clustering pada gambar 9, dengan metode *Single Linkage*, *Average Linkage*, *Complete Linkage*, dan *Ward Linkage* untuk tahun 2021, metode yang memberikan hasil terbaik adalah *Ward Linkage*. Metode ini menunjukkan variasi yang signifikan ketika data outlier dihapus dan dimasukkan. Hasil *clustering* tanpa *outlier* menunjukkan 1 kabupaten di C1, 1 kabupaten di C2, dan 28 kabupaten/kota di C3. Dengan data outlier, hasilnya adalah 4 kabupaten di C1, 1 kabupaten di C2, dan 25 kabupaten/kota di C3. Hal ini menunjukkan bahwa *Ward Linkage* lebih sensitif dan optimal dalam mengelompokkan data.

## Karakteristik Iklim 2021

Berdasarkan hasil *clustering* menggunakan metode *Single Linkage*, *Average Linkage*, *Complete Linkage*, dan *Ward Linkage*, metode *Complete Linkage* memberikan hasil terbaik. Dalam *Complete Linkage*, *Cluster C1* memiliki curah hujan yang sedikit lebih banyak dengan 195 hari hujan, *Cluster C2* memiliki curah hujan paling banyak, dan *Cluster C3* memiliki suhu antara 12,26-26,5°C dengan

curah hujan yang lebih sedikit, rata-rata antara 115,50 hingga 170,50 hari hujan. Metode ini menunjukkan pengelompokan yang lebih detail dan bervariasi berdasarkan curah hujan dan suhu udara, sehingga memberikan hasil optimal dalam analisis karakteristik iklim.

## **Pembahasan**

Berdasarkan hasil penelitian pada tabel 4.5, metode *K-Means* menunjukkan variasi yang signifikan dalam komposisi *cluster* antara tahun 2020 dan 2021. Metode *Linkage Hierarchical Clustering (Single, Average, dan Complete Linkage)* menunjukkan pola yang konsisten tetapi terpusat di C3, sementara *Ward Linkage* memberikan hasil yang lebih seimbang dan dianggap terbaik karena memisahkan cluster berdasarkan jumlah hari hujan dan suhu udara, baik pada tahun 2020 maupun 2021. *Complete Linkage* juga konsisten, sedangkan *Single* dan *Average Linkage* kurang jelas dalam memisahkan *cluster*. Pada tahun 2020, *Ward Linkage* menunjukkan hasil terbaik dengan C1 memiliki curah hujan sedikit lebih banyak, C2 dengan curah hujan lebih banyak, dan C3 dengan suhu lebih rendah serta curah hujan lebih sedikit. Pada tahun 2021, *Complete Linkage* menunjukkan hasil baik dengan C1 memiliki curah hujan sedikit lebih banyak, C2 dengan curah hujan paling banyak, dan C3 dengan suhu lebih rendah serta curah hujan sedikit. *Ward Linkage* pada tahun 2021 juga seimbang dengan C1 dan C2 memiliki curah hujan lebih banyak dan C3 dengan suhu lebih rendah serta curah hujan lebih sedikit. Berdasarkan hasil penelitian diketahui bahwa hasil *clustering* menggunakan *K-Means* pada tahun 2020 ada 31 Kabupaten/Kota dengan 1 Kabupaten yang *outlier* (Kabupaten Brebes). Jika data *outlier* dihapus, terdapat 30 Kabupaten/Kota dengan 28 Kabupaten di C1 (*cluster* rendah), 1 Kabupaten di C2 (*cluster* tertinggi), dan 1 Kabupaten di C3. Pada tahun 2021, *K-Means* menunjukkan 31 Kabupaten/Kota dengan 1 Kabupaten *outlier*. Setelah menghapus *outlier*, terdapat 30 Kabupaten/Kota dengan 1 Kabupaten di C1, 26 Kabupaten di C2, dan 3 Kabupaten di C3. Hasil *clustering* menggunakan *Linkage Hierarchical Clustering (Single, Average, Complete, dan Ward Linkage)* pada tahun 2020 dan 2021 menunjukkan bahwa metode *Ward Linkage* memberikan hasil terbaik dan lebih seimbang dibandingkan metode lainnya.

Kelompok terbanyak pada pengelompokan 3 kategori berada di *cluster* 2 (sedang) dengan metode *K-Means* tahun 2020, menunjukkan iklim di 15 kota/kabupaten di Jawa Tengah dengan suhu udara 25-27°C dan curah hujan lebih banyak (64-191 hari hujan). Pada tahun 2021, 25 kota/kabupaten masuk kategori sedang (*cluster* 2). Metode *Linkage Hierarchical Clustering (Single, Average, dan Complete Linkage)* tahun 2020 dan 2021 menunjukkan pengelompokan terbanyak pada *cluster* 3 dengan karakteristik iklim di 28 kabupaten/kota dan *Ward* di 25 kabupaten/kota, dengan suhu udara rata-rata 26-27,5°C dan curah hujan lebih sedikit, rata-rata 132-171 hari hujan.

## **IV. SIMPULAN**

Berdasarkan penelitian yang dilakukan, dapat disimpulkan bahwa metode *K-Means* menghasilkan variasi signifikan dalam komposisi *cluster* antara tahun 2020 dan 2021. Metode

*Linkage Hierarchical Clustering (Single, Average, dan Complete Linkage)* menunjukkan pola konsisten tetapi terpusat di C3, sementara *Ward Linkage* memberikan hasil terbaik dengan pemisahan *cluster* berdasarkan karakteristik jumlah hari hujan dan suhu udara. Pada tahun 2020, *Ward Linkage* memberikan hasil terbaik dengan C1 memiliki curah hujan sedikit lebih banyak, C2 curah hujan lebih banyak, dan C3 suhu lebih rendah serta curah hujan lebih sedikit. Pada tahun 2021, *Complete Linkage* juga memberikan hasil baik dengan pola serupa, dan *Ward Linkage* tetap seimbang, menjadikannya metode yang paling optimal. Faktor-faktor yang signifikan mempengaruhi potensi produksi bawang merah di Jawa Tengah adalah suhu udara, curah hujan, dan jumlah hari hujan. *Cluster 1* memiliki suhu udara, curah hujan, dan jumlah hari hujan yang sedikit, di mana pengairan yang sedikit juga tidak terlalu baik bagi pertumbuhan bawang merah. *Cluster 2*, dengan suhu udara rata-rata 25-27°C dan volume curah hujan yang sedikit lebih banyak, merupakan iklim yang sesuai bagi pertumbuhan bawang merah. *Cluster 3* adalah daerah dengan suhu dan curah hujan berlebihan, sehingga potensi pertumbuhannya rendah.

#### DAFTAR RUJUKAN

- [1] R. Despita *et al.*, “Produksi Bawang Merah pada Musim Hujan dengan Aplikasi Rhizobakteria Pemacu Tumbuh Tanaman Shallot Production in The Rainy Season with Plant Growth Promoting Rhizobacteria Application,” 2021.
- [2] L. T. W. Astuti, A. Daryanto, Y. Syaikat, and H. K. Daryanto, “Analisis Resiko Produksi Usahatani Bawang Merah pada Musim Kering dan Musim Hujan di Kabupaten Brebes,” *Jurnal Ekonomi Pertanian dan Agribisnis*, vol. 3, no. 4, pp. 840–852, Jan. 2019, doi: 10.21776/ub.jepa.2019.003.04.19.
- [3] Asep Muhidin, “ANALISA METODE HIERARCHICAL CLUSTERING DAN K-MEAN DENGAN MODEL LRFMP PADA SEGMENTASI PELANGGAN,” 2017.
- [4] A. Azzahra and A. W. Wijayanto, “SISTEMASI: Jurnal Sistem Informasi Perbandingan Agglomerative Hierarchical dan K-Means dalam Pengelompokan Provinsi Berdasarkan Pelayanan Kesehatan Maternal Comparison of Agglomerative Hierarchical and K-Means in Grouping Provinces Based on Maternal Health Services.”
- [5] A. A. Maulana, A. W. Rafii, Y. A. Anjelina, and E. Widodo, “Pengelompokan Kecamatan di Kabupaten Bima Berdasarkan Jumlah Produksi dan Luas Panen Bawang Merah Tahun 2021 Menggunakan K-Means Clustering,” 2023.
- [6] Y. Reinaldi, N. Ulinnuha, and Moh. Hafiyusholeh, “Comparison of Single Linkage, Complete Linkage, and Average Linkage Methods on Community Welfare Analysis in Cities and Regencies in East Java,” *Jurnal Matematika, Statistika dan Komputasi*, vol. 18, no. 1, pp. 130–140, Sep. 2021, doi: 10.20956/j.v18i1.14228.

- [7] Fithriyah Azzahrah, S. Annas, and Z. Rais, "Hybrid Hierarchical Clustering dalam Pengelompokan Daerah Rawan Bencana Tanah Longsor di Sulawesi Selatan," *VARIANSI: Journal of Statistics and Its application on Teaching and Research*, vol. 4, no. 3, pp. 153–161, Dec. 2022, doi: 10.35580/variansiunm38.
- [8] P. Kesesuaian, L. Dan, and C. Hujan, "MENINGKATKAN PRODUKSI BAWANG MERAH MELALUI."
- [9] A. Tri, R. Dani, S. Wahyuningsih, and N. A. Rizki, "Penerapan Hierarchical Clustering Metode Agglomerative pada Data Runtun Waktu," *Jambura Journal of Mathematics*, vol. 1, 2019, [Online]. Available: <http://ejournal.ung.ac.id/index.php/jjom,P->
- [10] A. Alfira, F. Hermin, and E. D. Wiraningsih, "Analisis Hybrid Mutual Clustering menggunakan Jarak Square Euclidean."
- [11] L. Rahmawati, J. Informatika, S. W. Sihwi, and E. Suryani, "ANALISA CLUSTERING MENGGUNAKAN METODE K-MEANS DAN HIERARCHICAL CLUSTERING (STUDI KASUS : DOKUMEN SKRIPSI JURUSAN KIMIA, FMIPA, UNIVERSITAS SEBELAS MARET)."
- [12] A. Septianingsih, S. A. Pertama, D. Kependudukan, P. Sipil, and K. Tangerang, "PEMETAAN KABUPATEN KOTA DI PROVINSI JAWA TIMUR BERDASARKAN TINGKAT KASUS PENYAKIT MENGGUNAKAN PENDEKATAN AGGLOMERATIF HIERARCHICAL CLUSTERING," vol. 3, no. 2, 2022, doi: 10.46306/lb.v3i2.
- [13] U. Muhammadiyah Mataram Mataram, L. Sulistina, M. Ibrahim, and H. Ratu Perwira Negara, "Seminar Nasional LPPM UMMAT Perbedaan Tingkat Akurasi Metode K-Means dan Hierarchical Clustering di Bidang Peramalan dan Klasifikasi," vol. 1, 2022, [Online]. Available: <https://www.scopus.com/>
- [14] A. Azzahra and A. W. Wijayanto, "SISTEMASI: Jurnal Sistem Informasi Perbandingan Agglomerative Hierarchical dan K-Means dalam Pengelompokan Provinsi Berdasarkan Pelayanan Kesehatan Maternal Comparison of Agglomerative Hierarchical and K-Means in Grouping Provinces Based on Maternal Health Services."
- [15] R. F. Nasution, "IJM: Indonesian Journal of Multidisciplinary Analisis Perbandingan Clustering Based Dengan Density Based Dalam Mendeteksi Outlier," *IJM: Indonesian Journal of Multidisciplinary*, vol. 2, 2024, [Online]. Available: <https://journal.csspublishing/index.php/ijm>