

SKRIPSI

**PENERAPAN ANALISIS *LINKAGE HIERARCHICAL*
CLUSTERING DAN *K-MEANS CLUSTERING* DALAM
PENGELOMPOKKAN POTENSI PRODUKSI BAWANG
MERAH DI JAWA TENGAH**



Oleh:

Auryn Clara Jangkobus

20083000037

**PROGRAM S1 SISTEM INFORMASI
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS MERDEKA
MALANG
2024**

**PENERAPAN ANALISIS *LINKAGE HIERARCHICAL CLUSTERING* DAN *K-MEANS CLUSTERING* PADA
PENGELOMPOKAN POTENSI PRODUKSI BAWANG
MERAH DI JAWA TENGAH**



SKRIPSI

Diajukan kepada

Program S1 Sistem Informasi Universitas Merdeka Malang

Untuk Memenuhi Salah Satu Persyaratan

Dalam Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh:

Auryn Clara Jangkobus

NIM: 20083000037

**PROGRAM S1 SISTEM INFORMASI
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS MERDEKA
MALANG**

2024

MOTTO

“Untuk segala sesuatu ada masanya,
untuk apa pun di bawah langit ada waktunya”

Pengkotbah 3 : 1

Proses Tuhan memang tidak cepat, tapi proses Tuhan itu Tepat

Do you'r best and God do the rest

(Lakukan yang terbaik dan Tuhan akan melakukan sisanya)

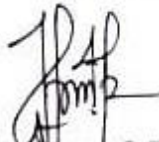
LEMBAR PERSETUJUAN SKRIPSI

Nama : Aurn Clara Jangkobus
NIM : 20083000037
Program Studi : S1 Sistem Informasi
Bidang Kajian : Analisis Data
Judul Tugas Akhir : Penerapan Analisis Linkage Hierarchical Clustering Dan
K-Means Clustering Dalam Pengelompokan Potensi
Produksi Bawang Merah Di Jawa Tengah

Malang, 31 Juli 2024

DISETUJUI DAN DITERIMA

Ketua Program Studi S1 Sistem
Informasi



Luthfi Indana, S.Pd., M.Pd.
NIDN. 0711059203

Dosen Pembimbing



Rahmatina Hidayati, S.Kom., MT.
NIDN. 0720028902

Dekan
Fakultas Teknologi Informasi



Dr. Mardiana Andarwati, S.E., M.Si
NIDN. 0716037601

LEMBAR PENGESAHAN

Dipersiapkan dan disusun oleh

Nama : Aurn Clara Jangkobus
NIM : 20083000037
Judul : Penerapan Analisis Linkage Hierarchical Clustering
Dan K-Means Clustering Pada Pengelompokan
Potensi Produksi Bawang Merah Di Jawa Tengah

Telah dipertahankan di depan Dosen Penguji pada,
Hari : Jumat
Tanggal : 12 Juli 2024
Tempat : Fakultas Teknologi Informasi
Universitas Merdeka Malang

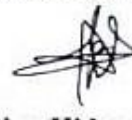
Susunan Dewan Penguji

Ketua Penguji



Firnanda Al. A. Putra, S.Kom., M.Kom.
NIDN. 0706129401

Sekretaris Penguji



Rahmatina Hidayati, S.Kom., M.T.
NIDN. 0720028902

Anggota Penguji



Bambang Nurdewanto, S.Kom., M.Kom
NIDN. 0005027501

Skripsi ini Telah Diterima Sebagai Salah Satu Persyaratan untuk memperoleh
Gelar Sarjana/Ahli Madya Komputer

Malang, 31 Juli 2024
Dekan Fakultas Teknologi Informasi
Universitas Merdeka Malang



Dr. Mardiana Andarwati, S.E., M.Si.
NIDN. 0716037601

SURAT PERNYATAAN ORISINALITAS SKRIPSI

Saya yang bertandatangan dibawah ini:

Nama : Aurny Clara Jangkobus
NPM : 20083000037
Program Studi : SI Sistem Informasi
Judul : Penerapan Analisis Linkage Hierarchical Clustering Dan K-Means Clustering Dalam Pengelompokkan Potensi Produksi Bawang Merah Di Jawa Tengah
Fakultas : Teknologi Informasi
Perguruan Tinggi : Universitas Merdeka Malang

Dengan ini menyatakan bahwa saya dengan sebenar-benarnya melakukan penelitian dan penulisan skripsi di atas benar-benar karya saya dan apabila ternyata di dalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur plagiasi, maka saya bersedia mendapatkan sanksi akademik berupa pencabutan gelar akademik sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini di buat untuk dipergunakan sebagaimana mestinya.

Malang, 31 Juli 2024



Aurny Clara Jangkobus

ABSTRAK

Bawang merah adalah tanaman hortikultura yang menjadi salah satu komoditas sayuran unggulan di Jawa Tengah, dengan dampak signifikan terhadap perekonomian regional. Produksi bawang merah di Jawa Tengah mencapai 556.057 ton pada tahun 2021, menurun dari 500.992 ton tahun sebelumnya. Ketidakefisienan dalam pasokan disebabkan oleh perbedaan iklim dan cuaca di berbagai wilayah. Oleh karena itu, diperlukan distribusi yang merata dari daerah dengan produksi tinggi ke daerah dengan produksi sedang dan rendah untuk menjaga keseimbangan pasokan dan permintaan serta stabilitas harga.

Penelitian ini menggunakan metode pengelompokan (*Clustering*) untuk mengidentifikasi daerah potensial produksi bawang merah di Jawa Tengah. Metode yang digunakan meliputi *Agglomerative Hierarchical Clustering* dan *K-means Clustering*. Hasil penelitian menunjukkan bahwa pada 2020, metode *K-Means Clustering* menunjukkan 15 Kota/Kabupaten Jawa Tengah dalam cluster 2 (suhu 25-27°C, curah hujan 64-191 hari), dan pada 2021, 25 Kota/Kabupaten. *Linkage Hierarchical Clustering* menunjukkan cluster 3 terbanyak (28 Kabupaten/Kota), *Ward Linkage* (25 Kabupaten/Kota). *Ward Linkage* (2020) dan *Complete Linkage* (2021) optimal. Suhu dan curah hujan signifikan mempengaruhi produksi bawang merah. *Ward Linkage* dan *Complete Linkage* memberikan hasil yang lebih optimal dalam pengelompokan daerah berdasarkan iklim. Faktor signifikan yang mempengaruhi produksi bawang merah adalah suhu udara dan curah hujan. Penelitian ini diharapkan dapat memberikan rekomendasi kebijakan yang tepat untuk meningkatkan produksi bawang merah dan kesejahteraan petani.

Kata kunci: Bawang merah, *K-Means Clustering*, *Linkage Hierarchical Clustering*

ABSTRACT

Shallots are a horticultural crop that is one of the leading vegetable commodities in Central Java, with a significant impact on the regional economy. Shallot production in Central Java reached 556,057 tons in 2021, a decrease from 500,992 tons the previous year. Inefficiencies in supply are caused by variations in climate and weather across different regions. Therefore, it is necessary to evenly distribute shallots from high-production areas to medium- and low-production areas to maintain supply and demand balance and price stability.

This study uses clustering methods to identify potential shallot production areas in Central Java. The methods used include Agglomerative Hierarchical Clustering and K-means Clustering. The results of the study show that in 2020, the K-Means method identified 15 cities/regencies in Central Java in cluster 2 (temperature 25-27°C, rainfall 64-191 days), and in 2021, 25 cities/regencies. Linkage Hierarchical Clustering showed the highest number in cluster 3 (28 regencies/cities), with Ward Linkage (25 regencies/cities). Ward Linkage (2020) and Complete Linkage (2021) were optimal. Temperature and rainfall significantly affect shallot production. Ward Linkage and Complete Linkage provide more optimal results in clustering regions based on climate. Significant factors influencing shallot production are air temperature and rainfall. This research is expected to provide policy recommendations to improve shallot production and the welfare of farmers.

Keywords: Shallots, K-Means Clustering, Linkage Hierarchical Clustering

LEMBAR PERSEMBAHAN

Saya mengungkapkan rasa syukur dan terima kasih yang tak terhingga kepada Tuhan Yesus Kristus atas berkat dan kasih karunia-Nya yang selalu menyertai saya hingga skripsi ini dapat terselesaikan.

Skripsi ini saya persembahkan untuk orangtua tercinta Mama Meiske, Papa Yupiter dan Pa'Albert dan keluarga saya. Terima kasih atas dukungan, kasih sayang, dan doa yang tiada hentinya diberikan kepada saya. Skripsi ini juga saya berikan kepada Adik saya tercinta Alm.Fendriko Inzaghi Takasanakeng. Serta tidak lupa untuk semua sahabat dekat yang ikut membantu dalam memberi dukungan dan motivasi kepada saya. Tak luput pula, karya ini saya persembahkan untuk diri saya sendiri yang sudah bertahan atas segala perjuangan melalui proses pendidikan dan telah kooperatif dalam mengerjakan skripsi ini hingga akhir. Terima kasih sudah melibatkan Tuhan Yesus Kristus dalam setiap perjalananmu.

KATA PENGANTAR

Puji dan Syukur kepada Tuhan Yang Maha Esa atas segala penyertaan dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul “Penerapan Analisis Linkage Hierarchical Clustering Dan K-Means Clustering Dalam Pengelompokkan Potensi Produksi Bawang Merah Di Jawa Tengah”.

Skripsi ini merupakan tugas akhir yang diajukan untuk memenuhi syarat dalam memperoleh gelar Sarjana Sistem Informasi Pada Fakultas Teknologi Informasi Universitas Merdeka Malang.

Penulis menyadari bahwa dalam penyusunan skripsi ini tidak akan terwujud tanpa ada bantuan dan dorongan berbagai pihak. Oleh karena itu pada kesempatan ini, penulis ingin menyampaikan terima kasih kepada yang terhormat:

1. Ibu Rahmatina Hidayati, S.Kom., MT., selaku Dosen Pembimbing yang telah memberikan bimbingan dan motivasi sehingga penulis bisa menyelesaikan skripsi ini dengan baik
2. Bapak Firnanda Al-Islama Achyunda Putra, S.Kom., M.Kom selaku Dosen Penguji I, dan Bapak Bambang Nurdewanto, S.Kom., M.Kom selaku Dosen Penguji II yang telah memberikan saran dan masukkan dalam penyempurnaan skripsi
3. Ibu Luthfi Indana S.Pd., M.Pd., selaku Ketua Program Studi Sistem Informasi
4. Ibu Dr. Mardiana Andarwati, SE., M.Si., selaku Dekan Fakultas Teknologi Informasi
5. Bapak Prof.Drs. Anwar Sanusi, MPA, Ph.D., selaku Rektor Universitas Merdeka Malang
6. Bapak Drs. Anis Zubair, M.Kom selaku Dosen Analisis Data yang telah memberikan ilmu pengetahuan selama penulis mengikuti perkuliahan
7. Bapak dan Ibu Dosen Program Studi S1 Sistem Informasi Fakultas Teknologi Informasi Universitas Merdeka Malang
8. Staf dan Bagian Akademik yang sudah membantu penulis selama perkuliahan
9. Kepada Orangtua penulis Mama Meiske Parera, A.Md.KP., Papa Yupiter S.H Jangkobus dan Pa’Albert Takasanakeng, A.Md.Par., yang telah memberikan segala dukungan, semangat, perhatian, doa serta telah mendidik dan membesarkan saya dalam limpahan kasih sayang

10. Teruntuk Tante Dra.Erlisiawati Parera, MM., Conny Parera, Om Revly.S Parera, A.Md.Pel., yang telah membantu, menasehati, memberikan doa, dukungan, semangat yang tidak didapatkan dimanapun, memberikan saran saat saya mengalami kesulitan dan membantu material untuk memenuhi kebutuhan saya dalam menyelesaikan skripsi
11. Teruntuk saudari tersayang penulis Adenia, yang telah berjuang bersama selama perkuliahan dalam suka dan duka serta terima kasih telah memberikan bantuan moril dan material, motivasi, semangat dan kebaikannya. Serta saudara-saudari Michiko, Amelia dan Adventus dan keluarga penulis yang telah memberikan doa, dukungan dan semangat
12. Teruntuk sahabat seperjuangan Maria Febriana Palin Werang dan Teman-teman mahasiswa Program Studi S1 Sistem Informasi Fakultas Teknologi Informasi Universitas Merdeka Malang
13. Teruntuk sahabat penulis Adel dan Ibu Kos Bu'Mina yang selalu memberikan dukungan, doa dan semangat
14. Seluruh pihak yang berperan secara langsung dan tidak langsung yang tidak dapat disebutkan satu per satu.

Akhirnya, dengan segala kerendahan hati peneliti menyadari keterbatasan penelitian ini karena keterbatasan peneliti. Demi kesempurnaan skripsi ini, saran dan kritik yang bersifat membangun sangat penulis harapkan. Semoga karya yang sederhana dapat memberi manfaat bagi semua pihak yang membutuhkan.

Malang, 31 Juli 2024

Auryn Clara Jangkobus

DAFTAR ISI

MOTTO	ii
LEMBAR PERSETUJUAN SKRIPSI	iii
LEMBAR PENGESAHAN	iv
SURAT PERNYATAAN ORISINALITAS SKRIPSI	v
ABSTRAK	vi
ABSTRACT	vii
LEMBAR PERSEMBAHAN	viii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN	xv
BIODATA PENELITI.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah	6
BAB II KAJIAN PUSTAKA	7
2.1 Landasan Teori	7
2.1.1 Tanaman Bawang Merah.....	7
2.1.2 Karakteristik Iklim di Jawa Tengah.....	7
2.1.3 <i>Data Mining</i>	8
2.1.4 <i>Linkage Hierarchical Clustering</i>	8
2.1.5 <i>K-means Clustering</i>	10
2.2 Penelitian Terdahulu	11
BAB III METODE PENELITIAN	15
3.1 Desain Penelitian	15
3.2 Lokasi Penelitian	16
3.3 Populasi dan Sampel atau Objek Penelitian	16

3.4 Teknik Pengumpulan Data	16
3.5 Teknis Analisis Data.....	18
3.5.1 Alur Algoritma <i>Hierarchical Clustering</i>	18
BAB IV HASIL DAN PEMBAHASAN	24
4.1 Hasil Penelitian.....	24
4.1.1 Pengumpulan Data.....	25
4.1.2 Preprocessing <i>Data</i>	26
4.1.3 Transformasi.....	28
4.1.4 Proses <i>Clustering</i>	29
4.2 Analisis Hasil.....	60
4.3 Pembahasan	65
BAB V PENUTUP.....	69
5.1 Kesimpulan.....	69
5.2 Saran	69
DAFTAR PUSTAKA	71
LAMPIRAN.....	74

DAFTAR TABEL

Tabel 2. 1 Karakteristik Iklim Jawa Tengah.....	7
Tabel 2. 2 Penelitian Terdahulu	11
Tabel 4. 1 Hasil <i>Clustering</i> Tahun 2020	38
Tabel 4. 2 Tabel Hasil Box Plot 2020	45
Tabel 4. 3 Hasil <i>Clustering</i> Tahun 2021	48
Tabel 4. 4 Hasil Box Plot 2021	55
Tabel 4. 5 Analisis Hasil dari Metode <i>K-Means</i> dan <i>Linkage Hierarchical Clustering</i>	60

DAFTAR GAMBAR

Gambar 3. 1	Flowchart Desain penelitian.....	15
Gambar 3. 2	Data Produksi Bawang Merah 2020 dan 2021.....	17
Gambar 3. 3	Alur Penyelesaian <i>Hierarchical Clustering</i> dan <i>K- means Clustering</i> di Ms. Excel.....	22
Gambar 3. 4	Flowchart Alur <i>Orange</i> Data Mining.....	23
Gambar 4. 1	Data Produksi dan Luas Panen Bawang Merah Tahun 2020 dan 2021.....	24
Gambar 4. 2	Pengumpulan data mentah di Ms. Excel.....	25
Gambar 4. 3	Data Produksi 2021 Sebelum di-cleaning.....	27
Gambar 4. 4	Data Produksi 2020 dan 2021 sesudah di cleaning.....	28
Gambar 4. 5	Penentuan Cluster <i>K-Means</i> Tahun 2020.....	30
Gambar 4. 6	Tampilan hasil <i>Clustering</i> pada <i>Scatterplot K-Means</i> 2020.....	31
Gambar 4. 7	Penentuan Cluster tanpa data Kabupaten Brebes.....	32
Gambar 4. 8	Tampilan <i>scatterplot</i> hasil penghapusan data <i>Outlier K-Means</i> 2020.....	33
Gambar 4. 9	Penentuan Cluster Tahun 2021.....	34
Gambar 4. 10	Penentuan Cluster Tanpa Kabupaten Brebes.....	34
Gambar 4. 11	Tampilan hasil penghapusan data <i>Outlier scatterplot</i> pada <i>K-Means</i> 2021.....	35
Gambar 4. 12	Cara penyimpanan hasil pengelolaan di aplikasi <i>Orange</i>	36
Gambar 4. 13	Tampilan lokasi penyimpanan di aplikasi <i>Orange</i>	36
Gambar 4. 14	Susunan <i>Clustering Hierarchical Clustering</i>	37

DAFTAR LAMPIRAN

Lampiran 1 Hasil Cek Plagiasi.....	74
Lampiran 2 Lembar Persetujuan Tugas Akhir.....	77
Lampiran 3 Data Tabel Outlier K-Means 2020 Pada Aplikasi Orange	78
Lampiran 4 Data Tabel Data Tanpa Outlier Pada K-Means 2020	79
Lampiran 5 Data Tabel Outlier K-Means 2021 Pada Orange.....	80
Lampiran 6 Data Tabel Tanpa Outlier K-Means 2021	81
Lampiran 7 Data Tabel Produksi Outlier Hierarchical Linkage (Single Linkage) 2020 ..	82
Lampiran 8 Data Tabel Produksi Tanpa Outlier Hierarchical Linkage (Single Linkage) 2020.....	83
Lampiran 9 Data Tabel Outlier Hierarchical Clustering (Average Linkage) 2020	84
Lampiran 10 Data Tabel Tanpa Outlier Hierarchical Clustering (Average Linkage) 2020	85
Lampiran 11 Data Tabel Outlier Hierarchical Clustering (Complete Linkage) 2020	86
Lampiran 12 Data Tabel Tanpa Outlier Hierarchical Clustering (Complete Linkage) 2020.....	87
Lampiran 13 Data Tabel Outlier Hierarchical Clustering (Ward Linkage) 2020.....	88
Lampiran 14 Data Tabel Tanpa Outlier Hierarchical Clustering (Ward Linkage) 2020 ..	89
Lampiran 15 Data Tabel Outlier Hierarchical Clustering (Single Linkage) 2021	90
Lampiran 16 Data Tabel Tanpa Outlier Hierarchical Clustering (Single Linkage) 2021	91
Lampiran 17 Data Tabel Outlier Hierarchical Clustering (Average Linkage) 2021	92
Lampiran 18 Data Tabel Tanpa Outlier Hierarchical Clustering (Average Linkage) 2021.....	93
Lampiran 19 Data Tabel Outlier Hierarchical Clustering (Complete Linkage) 2021	94
Lampiran 20 Data Tabel Tanpa Outlier Hierarchical Clustering (Complete Linkage) 2021.....	95
Lampiran 21 Data Tabel Outlier Hierarchical Clustering (Ward Linkage) 2021.....	96
Lampiran 22 Data Tabel Tanpa Outlier Hierarchical Clustering (Ward Linkage) 2021 ..	97