

TESIS
PENERAPAN ALGORITMA PARTITIONING AROUND
MEDOIDS (PAM) DAN HIERARCHICAL CLUSTERING
DALAM PENGELOMPOKKAN WILAYAH PRODUKSI
KOMODITAS TANAMAN PANGAN DI INDONESIA



disusun oleh

HERY PRIANDOKO

22.55.1205

Konsentrasi : Digital Transformation Intelligent

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026

TESIS
PENERAPAN ALGORITMA PARTITIONING AROUND
MEDIODS (PAM) DAN HIERARCHICAL CLUSTERING
DALAM PENGELOMPOKAN WILAYAH PRODUKSI
KOMODITAS TANAMAN PANGAN DI INDONESIA

APPLICATION OF PARTITIONING AROUND MEDIODS
(PAM) AND HIERARCHICAL CLUSTERING ALGORITHM
IN GROUPING FOOD CROP COMMODITY PRODUCTION
AREAS IN INDONESIA

Diajukan untuk memenuhi salah satu syarat mencapai derajat Pascasarjana
Program Studi S2 PJJ Informatika



disusun oleh

HERY PRIANDOKO

22.55.1204

Konsentrasi : Digital Transformation Intelligent

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA

2026

HALAMAN PERSETUJUAN

**PENERAPAN ALGORITMA PARTITIONING AROUND MEDOIDS
(PAM) DAN HIERARCHICAL CLUSTERING DALAM
PENGELOMPOKKAN WILAYAH PRODUKSI KOMODITAS TANAMAN
PANGAN DI INDONESIA**

**APPLICATION OF PARTITIONING AROUND MEDIOIDS (PAM) AND
HIERARCHICAL CLUSTERING ALGORITHM IN GROUPING FOOD
CROP COMMODITY PRODUCTION AREAS IN INDONESIA**

yang disusun dan diajukan oleh

Hery Priandoko

22.55.1204

telah disetujui oleh Dosen Pembimbing Tesis
pada tanggal 10 Juni 2025

Dosen Pembimbing Utama,



Alva Hendi Muhammad, S.T., M.Eng., Ph.D.
NIK. 190302493

HALAMAN PENGESAHAN

**PENERAPAN ALGORITMA PARTITIONING AROUND MEDOIDS
(PAM) DAN HIERARCHICAL CLUSTERING DALAM
PENGELOMPOKAN WILAYAH PRODUKSI KOMODITAS TANAMAN
PANGAN DI INDONESIA**

**APPLICATION OF PARTITIONING AROUND MEDIOIDS (PAM) AND
HIERARCHICAL CLUSTERING ALGORITHM IN GROUPING FOOD
CROP COMMODITY PRODUCTION AREAS IN INDONESIA**

yang disusun dan diajukan oleh

Hery Priandoko

22.55.1204

Telah dipertahankan di depan Dewan Penguji
pada tanggal 10 Juni 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Dr. Kumara Ari Yuana, S.T., M.T.
NIK. 190302575

Dr. Andi Sunyoto, M.Kom.
NIK. 190302052

Alva Hendi Muhammad, S.T., M.Eng., Ph.D.
NIK. 190302493



Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer
Tanggal 10 Juni 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : Hery Priandoko

NIM : 22.55.1204

Menyatakan bahwa Tesis dengan judul berikut:

PENERAPAN ALGORITMA PARTITIONING AROUND MEDOIDS (PAM) DAN HIERARCHICAL CLUSTERING DALAM PENGELOMPOKKAN WILAYAH PRODUKSI KOMODITAS TANAMAN PANGAN DI INDONESIA

Dosen Pembimbing Utama : Alva Hendi Muhammad, S.T., M.Eng., Ph.D.

Dosen Pembimbing Pendamping : Anggit Dwi Hartanto, M.Kom.

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian **SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini **SAYA** buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 10 Juni 2025

Yang Menyatakan,



Hery Priandoko

HALAMAN PERSEMBAHAN

Alhamdulillah segala puji dan syukur kehadiran Allah SWT yang telah memberikan kemudahan dan kelancaran kepada saya sehingga laporan tesis ini dapat terselesaikan. Laporan tesis ini saya persembahkan kepada individu-individu yang berjasa dalam perjalanan penelitian ini:

1. Keluarga dan Orang Tersayang: Tesis ini dipersembahkan kepada Orang Tua, Istri tercinta dan Saudara keluarga saya yang selalu memberikan dukungan, semangat serta motivasi tanpa henti-hentinya.
2. Pembimbing : Tesis ini dipersembahkan kepada Bapak Alva Hendi Muhammad, S.T., M.Eng., Ph.D dan Bapak Anggit Dwi Hartanto, M.Kom. sebagai penghormatan atas bimbingan, arahan, dan pengetahuan yang berharga yang telah diberikan selama penyusunan tesis ini. Terima kasih atas dedikasi dan kesabaran yang tiada duanya.
3. Institusi : Tesis ini dipersembahkan kepada Universitas AMIKOM Yogyakarta sebagai ungkapan terima kasih atas fasilitas, akses, dan sumber daya yang telah disediakan, yang telah berperan penting dalam kemajuan penelitian ini.

KATA PENGANTAR

Puji syukur kami panjatkan kehadiran Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya, sehingga tesis ini dapat diselesaikan dengan baik. Shalawat dan salam semoga selalu tercurahkan atas junjungan kita Nabi Muhammad SAW, beserta keluarga, sahabat-sahabat serta para pengikutnya yang istiqomah sampai akhir zaman.

Penyusunan tesis yang berjudul "Penerapan Algoritma Partitioning Around Medoids (PAM) dan Hierarchical Clustering dalam Pengelompokan Wilayah Produksi Komoditas Tanaman Pangan di Indonesia", merupakan salah satu syarat untuk memperoleh gelar Magister Komputer di Universitas Amikom Yogyakarta.

Dengan terselesainya laporan tesis ini, penulis berkesempatan untuk mengucapkan terima kasih kepada :

1. Orangtua, istri dan keluarga besar penulis yang selalu memberikan motivasi, arahan, dan do'a.
2. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta.
3. Ibu Prof. Dr. Kusriani, M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
4. Bapak Alva Hendi Muhammad, S.T., M.Eng., Ph.D., selaku Dosen Pembimbing I, atas bimbingan serta waktu yang telah diberikan.
5. Bapak Anggit Dwi Hartanto, M.Kom. selaku Dosen Pembimbing II, atas bimbingan serta waktu yang telah diberikan.

6. Teman-teman kuliah angkatan 7 PJJ informatika Amikom Yogyakarta atas kerjasamanya selama perkuliahan.

Penulis sangat mengharapkan kritik dan saran dari pembaca demi kesempurnaan tesis ini, karena penyusun menyadari masih banyak kekurangan. Akhir kata penulis berharap semoga laporan tesis yang sederhana ini dapat bermanfaat bagi yang berkepentingan.

Yogyakarta, 10 Juni 2025

Hery Priandoko



DAFTAR ISI

| | |
|---|-----------|
| HALAMAN JUDUL | i |
| HALAMAN PERSETUJUAN | ii |
| HALAMAN PENGESAHAN | iii |
| HALAMAN PERNYATAAN KEASLIAN TESIS | iv |
| HALAMAN PERSEMBAHAN | v |
| KATA PENGANTAR | vi |
| DAFTAR ISI | viii |
| DAFTAR TABEL | x |
| DAFTAR GAMBAR | xi |
| INTISARI | xii |
| ABSTRACT | xiii |
| BAB 1 PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah | 5 |
| 1.3 Batasan Masalah | 6 |
| 1.4 Tujuan Penelitian | 6 |
| 1.5 Manfaat Penelitian | 7 |
| BAB 2 TINJAUAN PUSTAKA | 8 |
| 2.1 Tinjauan Pustaka | 8 |
| 2.2 Keaslian Penelitian | 11 |
| 2.3 Landasan Teori | 17 |
| 2.3.1 Analisa Cluster | 17 |
| 2.3.2 Metode Hierarcical Clustering | 18 |
| 2.3.3 Metode Non Hierarcical | 18 |
| 2.3.4 Algoritma K-Means | 19 |
| 2.3.5 Algoritma Partitioning Around Medoids (PAM) | 19 |
| 2.3.6 Metode agglomerative | 21 |
| 2.3.7 Metode devisive | 22 |
| 2.3.8 Metode Single Linkage | 23 |
| 2.3.9 Metode Complete Linkage | 24 |
| 2.3.10 Metode Average Linkage | 24 |
| 2.3.11 Pengukuran Jarak Kemiripan | 25 |
| 2.3.12 Standarisasi Data | 25 |
| 2.3.13 Silhouette Coefficient | 26 |
| 2.3.14 Principal Componen Analisys (PCA) | 27 |
| BAB 3 METODE PENELITIAN | 30 |

| | |
|---|-----------|
| 3.1 Jenis, Sifat dan Pendekatan Penelitian | 30 |
| 3.2 Metode Pengumpulan Data | 32 |
| 3.2.1 Preprocessing Data | 33 |
| 3.3 Deskriptif Dataset | 34 |
| 3.4 Metode Analisis Data | 35 |
| 3.4.1 Preprocessing | 35 |
| 3.4.2 Metode Hierarchical Clustering | 37 |
| 3.4.3 Metode Partitioning Around Medoids (PAM) | 37 |
| 3.4.4 Pembuatan Dendogram dan Peta Cluster | 38 |
| 3.4.5 Analisa Cluster | 38 |
| 3.5 Alur Penelitian | 39 |
| BAB 4 HASIL PENELITIAN DAN PEMBAHASAN | 43 |
| 4.1 Pengumpulan Data | 43 |
| 4.2 Deskriptif Produksi Komoditas Tanaman | 44 |
| 4.3 Partitioning Around Medoids (PAM) | 68 |
| 4.4 Hierarchical Clustering dengan Metode Agglomerative | 77 |
| 4.5 Analisa Cluster | 91 |
| 4.6 Diskusi Teoritik | 92 |
| BAB 5 PENUTUP | 94 |
| 5.1 Kesimpulan | 94 |
| 5.2 Saran | 95 |
| DAFTAR PUSTAKA | 98 |

DAFTAR TABEL

| | |
|---|----|
| Tabel 2.1 Matriks literatur review dan posisi penelitian | 11 |
| Tabel 4.1 Hasil analisis statistik deskriptif | 45 |
| Tabel 4.2 Hasil analisis statistik deskriptif luas panen | 47 |
| Tabel 4.3 Hasil Analisis Statistik Deskriptif Produksi Tanaman Pangan | 55 |
| Tabel 4.4 Hasil Analisis Statistik Deskriptif Jumlah Bencana Alam | 62 |
| Tabel 4.5 Hasil Analisis Statistik Deskriptif Rata-rata Pendapatan Bersih ... | 64 |
| Tabel 4.6 Hasil Analisis Statistik Deskriptif Perubahan Data Nol | 65 |
| Tabel 4.7 Hasil Silhouette Score Metode PAM | 70 |
| Tabel 4.8 Nilai Principal Components (PC) | 71 |
| Tabel 4.9 Loading Komponen Utama Pertama (PC1) | 73 |
| Tabel 4.10 Loading Komponen Utama Kedua (PC2) | 73 |
| Tabel 4.11 Silhouette Score Penentuan K Optimal | 79 |
| Tabel 4.12 Profiling Hierarchical Clustering (Average Linkage, K=2) | 83 |



DAFTAR GAMBAR

| | |
|--|----|
| Gambar 2.1 Bagan Analisis Cluster | 17 |
| Gambar 3.1 Alur Penelitian | 39 |
| Gambar 4.1 Grafik Silhouette Score Metode PAM | 69 |
| Gambar 4.2 Visualisasi Clustering dengan Metode PAM | 72 |
| Gambar 4.3 Peta Cluster Metode PAM | 76 |
| Gambar 4.4 Dedrogram Cluster Metode Single Linkage | 85 |
| Gambar 4.5 Dedrogram Cluster Metode Average Linkage | 86 |
| Gambar 4.6 Dedrogram Cluster Metode Complete Linkage | 87 |
| Gambar 4.7 Silhouette Score Metode Hierarchical Clustering | 79 |
| Gambar 4.8 Dedrogram Cluster Metode Single Linkage, K=2 | 88 |
| Gambar 4.9 Dedrogram Cluster Metode Average Linkage, K=2 | 89 |
| Gambar 4.10 Dedrogram Cluster Metode Complete Linkage, K=2 | 90 |
| Gambar 4.11 Peta Cluster Metode Hierarchical Clustering | 82 |



INTISARI

Indonesia terdiri dari beberapa wilayah yang memiliki potensi dalam pemenuhan kebutuhan pangan. Salah satu sektor utama dalam pemenuhan kebutuhan pangan adalah sektor pertanian. Sektor pertanian merupakan sektor yang perlu mendapatkan perhatian yang cukup besar dari pemerintah pusat dan daerah dalam pemenuhan kebutuhan pangan nasional. Kebutuhan pangan pada saat ini sering mengalami kelangkaan sehingga masyarakat menjadi sulit untuk mendapatkan kebutuhan pangan tersebut. Masalah ketergantungan akan kebutuhan pangan dapat membahayakan ketersediaan pasokan pangan negara. Di Indonesia terdapat wilayah yang memiliki produksi komoditas tanaman pangan sehingga dapat membantu dalam ketersediaan kebutuhan pangan. Dari permasalahan yang ada, peneliti mencoba melakukan penelitian dengan melakukan pengelompokkan terhadap wilayah yang ada di Indonesia khususnya di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat untuk mengetahui wilayah mana yang memiliki produksi komoditas tanaman pangan.

Penelitian ini, pengelompokkan wilayah di Indonesia khususnya di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat yang memiliki produksi komoditas tanaman pangan akan menggunakan metode PAM dan HC. Metode HC yang akan digunakan adalah metode agglomerative cluster dengan metode linkage, seperti single, complete dan average linkage. Hasil pengelompokkan pada penelitian ini kemudian akan dianalisa untuk menentukan metode linkage yang terbaik dan berapa cluster yang sebaiknya dibentuk menggunakan nilai silhouette score.

Berdasarkan hasil penelitian, dengan menggunakan metode PAM dan metode HC diperoleh nilai silhouette yang sama sebesar 0.9083 dan jumlah cluster (k) optimal $k = 2$. Baik metode PAM dan HC memiliki jumlah anggota cluster yang sama. Cluster 1 terdiri dari 98 kabupaten / kota, cluster 2 terdiri dari 2 kabupaten / kota.

Kata Kunci: Medoids, Cluster, Hierarchical, PAM, Pangan

ABSTRACT

Indonesia consists of several regions with potential for meeting food needs. One of the main sectors contributing to this is agriculture. This sector requires significant attention from the central and regional governments to meet national food needs. Food shortages are currently common, making it difficult for people to obtain these essential foods. This dependency on food can jeopardize the country's food supply. Indonesia has several regions that produce food crops, which can contribute to food security. To address this issue, researchers conducted a study by grouping regions in Indonesia, specifically in the provinces of East Java, Central Java, and West Java, to determine which regions produce food crops.

This study, which categorizes regions in Indonesia, specifically in the provinces of East Java, Central Java, and West Java, that produce food crops, will use the PAM and HC methods. The HC method used is the agglomerative cluster method with linkage methods, such as single, complete, and average linkage. The clustering results in this study will then be analyzed to determine the best linkage method and the optimal number of clusters to form using the silhouette score.

Based on the research results, the PAM and HC methods obtained the same silhouette value of 0.9083, and the optimal number of clusters (k) was $k = 2$. Both the PAM and HC methods had the same number of cluster members. Cluster 1 consisted of 98 districts/cities, and Cluster 2 consisted of 2 districts/cities.

Keywords: *Medoids, Cluster, Hierarchical, PAM, Food*

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara agraris, terdiri dari beberapa wilayah yang memiliki potensi dalam pemenuhan kebutuhan pangan. Salah satu sektor utama dalam pemenuhan kebutuhan pangan adalah sektor pertanian. Sektor pertanian merupakan sektor yang perlu mendapatkan perhatian yang cukup besar dari pemerintah pusat dan daerah karena memiliki peran yang penting dalam rangka pembangunan ekonomi bangsa. Sektor pertanian merupakan sektor yang memiliki peranan penting bagi perekonomian nasional. Hal ini didukung oleh iklim tropis yang dimiliki negara Indonesia serta ditunjang dengan struktur tanah yang baik untuk digunakan bercocok tanam [1]. Penyediaan akan pemenuhan kebutuhan pangan dapat dilakukan melalui produksi sendiri, dengan cara memanfaatkan alokasi sumberdaya alam dan juga dapat melakukan impor dari negara lain. Konsumsi kebutuhan pangan yang besar di Indonesia harus diimbangi dengan produksi komoditas pertanian sehingga dapat mencukupi kebutuhan nasional. Oleh karena itu, negara harus memberikan perhatian penuh agar tidak menyebabkan krisis pangan. Indonesia sendiri memiliki beberapa komoditi di sektor pertanian yang menjadi prioritas utama dalam pemenuhan kebutuhan pangan, diantaranya tanaman padi, tanaman jagung, dan tanaman umbi-umbian.

Tingkat ketergantungan penduduk Indonesia terhadap kebutuhan pangan cukup tinggi, yang akan menimbulkan masalah apabila terjadi kelangkaan.

Kebutuhan pangan pada saat ini sering mengalami kelangkaan sehingga masyarakat menjadi sulit untuk mendapatkan kebutuhan pangan tersebut. Sebagai negara agraris, negara harus tetap menjaga ketersediaan kebutuhan pangan tersebut. Salah satu cara yang biasa dilakukan untuk menjamin ketersediaan kebutuhan pangan tersebut adalah dengan melakukan kebijakan import. Alasan suatu negara melakukan kebijakan impor disebabkan karena adanya kegagalan negara tersebut dalam pemenuhan kebutuhan dalam negeri, baik dalam bentuk kualitas maupun kuantitas [2]. Kebijakan impor akan dilakukan apabila produksi yang dihasilkan di dalam negeri sudah sangat minim dan tidak sesuai dengan kebutuhan dalam negeri.

Masalah ketergantungan akan kebutuhan pangan dapat membahayakan ketersediaan pasokan pangan negara. Impor akan komoditas tanaman pangan menjadi salah satu solusi untuk menjaga ketersediaan pangan di Indonesia. Impor komoditas tanaman pangan yang dilakukan oleh Indonesia menunjukkan bahwa jumlah ketersediaan komoditas pangan tidak dapat mencukupi kebutuhan pangan nasional. Hal ini diperkuat dengan adanya pasal 14 ayat 2 undang-undang nomor 18 tahun 2012 tentang pemerintah dan pemerintah daerah bertanggung jawab untuk mengembangkan sistem distribusi pangan guna mewujudkan stabilitas pasokan dan harga pangan di tingkat konsumen dan produsen. Dalam pasal ini dapat menjelaskan bahwa pemerintah pusat dan pemerintah daerah memiliki tanggung jawab untuk memastikan kestabilan pasokan pangan.

Berdasarkan data Badan Pusat Statistik (BPS) Indonesia, impor kebutuhan komoditas pangan di Indonesia dalam 5 tahun terakhir mengalami fluktuatif dan

cenderung mengalami peningkatan. Menurut lembaga pertanian dan pangan dunia (FAO) 1984 ketersediaan kebutuhan komoditas pangan mencakup beberapa aspek penting yang harus diperhatikan oleh negara-negara di seluruh dunia. FAO menekankan beberapa poin penting berkaitan dengan ketersediaan pangan, diantaranya produksi yang memadai, distribusi yang efektif, stabilisasi harga, cadangan pangan, keragaman pangan, dan keamanan pangan. Dibutuhkan langkah-langkah yang konkrit dalam pencapaian swasembada pangan diantaranya adalah dengan perencanaan dan implementasi program secara baik. Perencanaan yang baik akan menghasilkan ketersediaan data yang akurat [3]. Pemerintah harus dapat menentukan kebijakan yang tepat untuk meningkatkan jumlah produksi komoditas pangan didalam negeri. Berdasarkan data dari badan pusat statistik jumlah produksi komoditas tanaman pangan pada setiap provinsi di Indonesia berbeda-beda. Indonesia terdapat daerah-daerah yang memiliki produksi komoditas tanaman pangan sehingga dapat membantu dalam ketersediaan kebutuhan pangan negara. Untuk mendapatkan dataset mengenai daerah-daerah di Indonesia yang memiliki produksi komoditas tanaman pangan peneliti menggunakan dataset publik yang terdapat di website badan pusat statistik (BPS).

Produksi tanaman pangan dan hortikultura di berbagai wilayah di Indonesia sering menunjukkan ketidakmerataan yang signifikan, baik secara kuantitas maupun pola produksinya. Penelitian sebelumnya, menunjukkan bahwa wilayah yang dikategorikan berdasarkan produksi tanaman pangan seperti padi, jagung, kedelai, dan hortikultura mengelompokkan kabupaten / kota ke dalam kelompok produksi tinggi dan rendah berdasarkan output komoditasnya. Namun,

sebagian besar dari penelitian hanya menggunakan satu metode clustering yang sederhana, seperti K-Means atau K-Medoids tanpa membandingkan kestabilan hasil antar metode yang digunakan atau tanpa mempertimbangkan karakteristik dataset yang digunakan, seperti keberadaan nilai nol, pencilan, dan distribusi tidak normal.

Dari permasalahan yang ada, peneliti mencoba melakukan penelitian dengan melakukan pengelompokan terhadap daerah-daerah atau wilayah yang ada di Indonesia khususnya di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat untuk mengetahui daerah-daerah mana yang memiliki produksi komoditas tanaman pangan. Dalam penelitian ini, metode pengelompokan daerah-daerah atau wilayah yang ada di Indonesia khususnya di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat akan menggunakan metode partitioning around medoids (PAM) dan hierarchical clustering. Metode hierarchical clustering yang akan digunakan adalah metode agglomerative. Metode agglomerative sendiri terdapat 3 macam metode linkage, yaitu dengan metode single linkage (jarak terdekat), average linkage (jarak rata-rata) dan complete linkage (jarak terjauh).

Analisis hierarchical clustering merupakan analisis yang digunakan untuk melakukan pengelompokan data dengan cara mengukur jarak kedekatan setiap objek kemudian akan disajikan dalam bentuk dendrogram. Beberapa jenis analisis cluster yang menggunakan metode hierarchical clustering, yaitu Agglomerative (bottom-up) dan Devisive (top-down) [4]. Metode hierarchical clustering dimulai dengan mengelompokkan dua atau lebih objek yang paling mirip. Cluster tersebut kemudian berlanjut dengan objek lain dan seterusnya, membentuk semacam

pohon yang di dalamnya terdapat tingkatan (hierarki) yang jelas antar objek, dari yang paling mirip hingga yang paling tidak mirip. Hasil dari analisis hierarchical clustering dapat ditunjukkan dalam bentuk dendrogram atau dalam diagram pohon [5]. Dendrogram terbentuk dari matriks jarak yang berukuran $n \times n$. Tahap pertama dalam membentuk dendrogram adalah dengan melakukan perhitungan jarak dan menyusun setiap objek secara hirarkis sehingga objek dengan kemiripan terbesar terletak berdekatan berdasarkan jaraknya. Kemudian grup atau cluster yang terbentuk akan bergabung dengan grup atau cluster lainnya, dimana mereka bersifat cukup mirip berdasarkan jaraknya, dan seterusnya [6]. Tujuan dari analisis cluster ini nantinya akan digunakan untuk mengelompokkan daerah-daerah atau wilayah di Indonesia khususnya di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat yang memiliki produksi komoditas tanaman pangan.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang diatas, maka rumusan masalah dalam penelitian ini adalah sebagai berikut :

- a. Bagaimana cara untuk mengelompokkan daerah-daerah atau wilayah di Indonesia khususnya di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat yang memiliki produksi komoditas tanaman pangan dengan memanfaatkan metode partitioning around medoids (PAM) dan hierarchical clustering.
- b. Berapa cluster yang sebaiknya dibentuk dalam penelitian tersebut?

1.3 Batasan Masalah

Batasan masalah yang akan dilakukan dalam penelitian ini adalah sebagai berikut :

- a. Data yang digunakan untuk melakukan penelitian ini menggunakan dataset publik yang diperoleh dari website Badan Pusat Statistik (BPS)
- b. Metode cluster yang akan digunakan dalam penelitian ini adalah metode *partitioning around medoids* (PAM) dan *hierarchical clustering*.

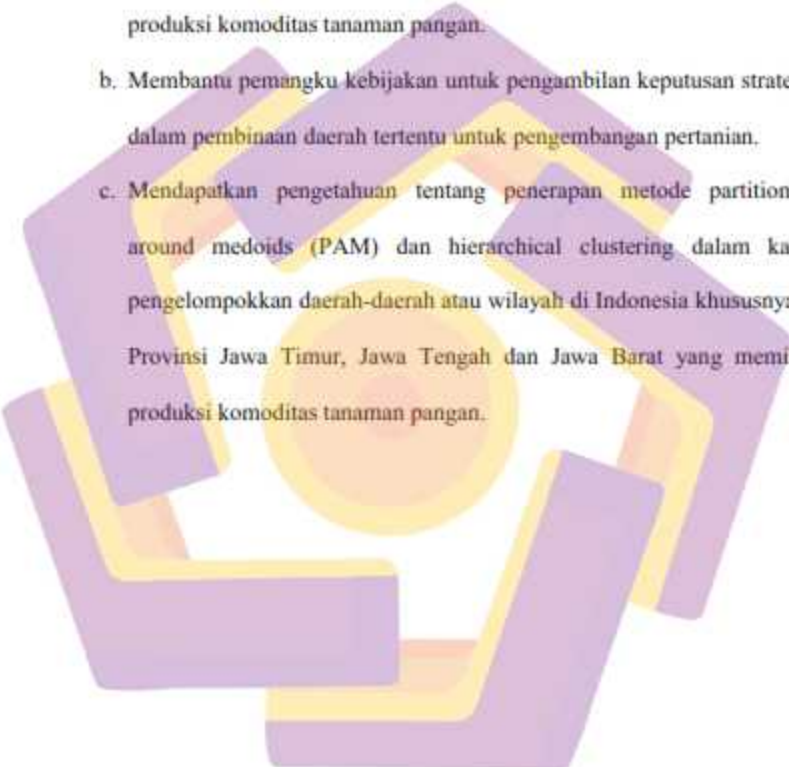
1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah :

- a. Melakukan pengelompokan daerah-daerah atau wilayah di Indonesia khususnya di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat ke dalam cluster yang homogen berdasarkan kesamaan dalam aspek produksi pangan, kondisi lingkungan (banjir, cuaca ekstrem, karhutla), dan variabel pendukung lainnya seperti upah tenaga.
- b. Mengetahui penerapan metode *partitioning around medoids* (PAM) dan *hierarchical clustering* dalam pengelompokan daerah-daerah atau wilayah di Indonesia khususnya di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat yang memiliki produksi komoditas tanaman pangan.
- c. Melakukan analisa hasil pengelompokkan agar dapat digunakan oleh pemangku kebijakan untuk pengambilan keputusan strategis dalam pengembangan pertanian.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

- a. Mengetahui daerah-daerah atau wilayah di Indonesia khususnya di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat yang memiliki produksi komoditas tanaman pangan.
 - b. Membantu pemangku kebijakan untuk pengambilan keputusan strategis dalam pembinaan daerah tertentu untuk pengembangan pertanian.
 - c. Mendapatkan pengetahuan tentang penerapan metode partitioning around medoids (PAM) dan hierarchical clustering dalam kasus pengelompokan daerah-daerah atau wilayah di Indonesia khususnya di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat yang memiliki produksi komoditas tanaman pangan.
- 

BAB 2

TINJAUAN PUSTAKA

2.1 Tujuan Pustaka

Penelitian ini dilakukan untuk melakukan pengelompokan terhadap daerah-daerah atau wilayah di Indonesia khususnya di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat yang memiliki produksi komoditas tanaman pangan yang tinggi, sehingga dalam penelitian ini diharapkan bisa menjadi pengetahuan untuk dijadikan fokus pembinaan terhadap daerah-daerah atau wilayah dengan produksi komoditas tanaman pangannya yang rendah. Penelitian ini menggunakan dataset publik yang berasal dari website badan pusat statistik (BPS). Dalam penelitian ini, peneliti akan menggunakan metode *partitioning around medoids (PAM)* dan *hierarchical clustering*. Metode *hierarchical clustering* yang akan digunakan adalah metode *agglomerative*. Metode *agglomerative* sendiri terdapat 3 macam metode cluster yaitu dengan metode *single linkage* (jarak terdekat), *average linkage* (jarak rata-rata) dan *complete linkage* (jarak terjauh).

Berikut beberapa penelitian yang sudah pernah dilakukan oleh peneliti lain dengan menggunakan metode *Partitioning Around Medoids (PAM)* dan *hierarchical clustering*. Penelitian pertama berjudul *Mapping the Diversity of Regional Characteristics Towards Sustainable Economic Strategic Area Development: A Case Study of West-East Corridor of West Sumatra Province* [7]. Penelitian ini bertujuan untuk mengetahui keberagaman karakteristik dan tipologi kawasan strategis Koridor Ekonomi Timur-Barat.

Penelitian selanjutnya berjudul Indonesian Territory Clustering Based On Harvested Area and Rice Productivity Using Clustering Algorithm [8]. Penelitian ini bertujuan untuk mengklasifikasi kota/kabupaten di Indonesia berdasarkan data produksi padi tahun 2021.

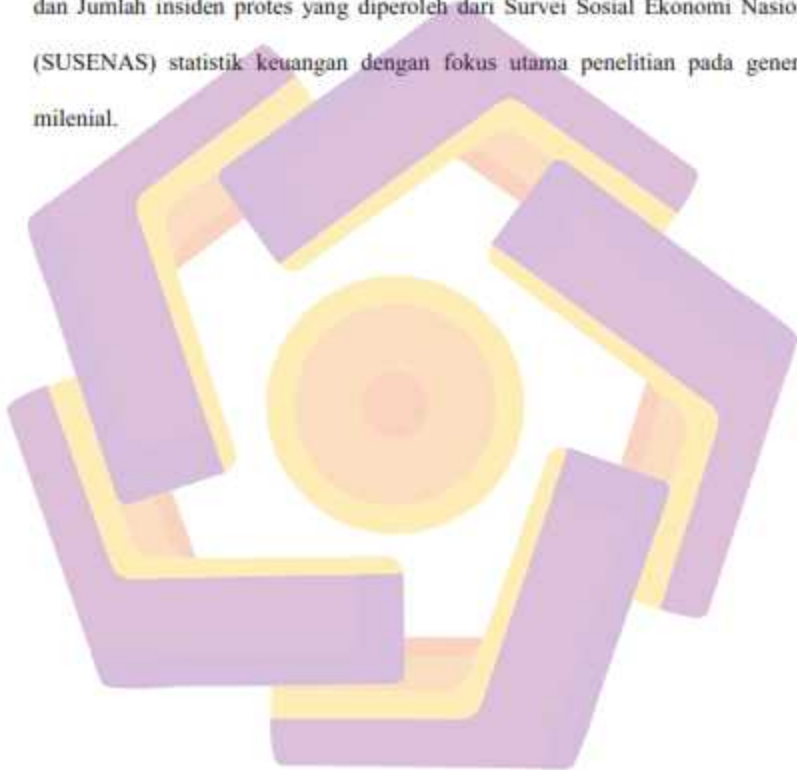
Penelitian selanjutnya berjudul Clustering and Mapping of Agricultural Production Based on Geographic Information System Using KMedoids Algorithm [9]. Penelitian ini bertujuan untuk mengembangkan sistem pemetaan dan pengelompokan produksi pertanian dengan menggunakan algoritma K-Medoids.

Penelitian selanjutnya berjudul Credit Customer Segmentation With Hierarchical Clustering At Various [10]. Penelitian ini bertujuan untuk melihat pengelompokan dan perbandingan hasil analisis kluster menggunakan metode ward linkage dengan berbagai jarak (Euclidean, Manhattan dan Mahalanobis) dengan penilaian variabel 5C (Character, Capacity, Capital, Collateral, Condition of Economy) pada nasabah KPR Bank X.

Penelitian selanjutnya berjudul Analisis Hard dan Soft Clustering Untuk Pengelompokan Indikator Ketahanan Pangan Indonesia 2021 [11]. Penelitian ini bertujuan untuk Penelitian ini bertujuan untuk mengelompokkan provinsi-provinsi di Indonesia berdasarkan karakteristik indikator ketahanan pangan dan membandingkan beberapa metode untuk mendapatkan metode pengelompokan yang terbaik.

Penelitian selanjutnya berjudul Cluster Around Latent Variable for Vulnerability Towards Natural Hazards, Non-Natural Hazards, Social Hazards in West Papua [12]. Penelitian ini bertujuan untuk menghitung kerentanan sosial

berbasis lokasi di Papua Barat yang melibatkan komponen Informasi, Teknologi, dan Komunikasi, Akses Pangan, Bencana Alam, Pernyataan Perlindungan Sosial, Akses ke Layanan Keuangan, Deskripsi sumber pendapatan rumah tangga, Jumlah kejadian banjir, jumlah bencana gempa bumi, kasus kematian COVID-19, dan Jumlah insiden protes yang diperoleh dari Survei Sosial Ekonomi Nasional (SUSENAS) statistik keuangan dengan fokus utama penelitian pada generasi milenial.



2.2 Keaslian Penelitian

Tabel 2.1 Matriks literatur review dan posisi penelitian

PENERAPAN ALGORITMA PARTITIONING AROUND MEDOIDS (PAM) DAN HIERARCHICAL CLUSTERING DALAM PENGELOMPOKKAN WILAYAH PRODUKSI KOMODITAS TANAMAN PANGAN DI INDONESIA

| No | Judul Penelitian | Nama Peneliti, Tahun, Index | Metode Penelitian | Hasil | Keunggulan dan Kelemahan | Perbandingan |
|----|--|---|---|---|--|--|
| 1 | Mapping the Diversity of Regional Characteristics Towards Sustainable Economic Strategic Area Development: A Case Study of West-East Corridor of West Sumatra Province | Siska Amelia ^{1,2*} , Eman Rustiadi ³ , Baba Barus ³ , Bambang Juanda ⁴ , International Journal of Sustainable Development and Planning, 2022 | Penelitian ini bertujuan untuk mengetahui keberagaman karakteristik dan tipologi kawasan strategis Koridor Ekonomi Timur-Barat. | Hasil analisis kluster spasial menghasilkan 3 kluster yaitu kluster perkotaan, kluster desa-kota dan kluster perdesaan sedangkan analisis overlay menghasilkan 10 karakteristik dan tipologi wilayah yang digunakan sebagai dasar penyusunan strategi dan kebijakan pengembangan wilayah serta peningkatan peluang investasi di kawasan strategis Koridor Barat-Timur dan | Dapat dilakukan penelitian dengan menggunakan metode selain metode PCA | Pada penelitian ini menggunakan objek penelitian dan metode cluster yang berbeda |

| | | | | | | |
|---|---|--|---|--|--|--|
| | | | | Provinsi Sumatera Barat pada umumnya. | | |
| 2 | INDONESIAN TERRITORY CLUSTERING BASED ON HARVESTED AREA AND RICE PRODUCTIVITY USING CLUSTERING ALGORITHM | Imelda Putri Kurniawati*, Hasih Pratiwi, Sugiyanto, Journal of Social Science, 2022 | Tujuan penelitian ini adalah untuk mengklasifikasi kota/kabupaten di Indonesia berdasarkan data produksi padi tahun 2021. | Metode yang diperoleh terbaik adalah metode FCM dengan dua kluster dan nilai siluet sebesar 0,828. Hasil klasterisasi dengan metode terbaik tersebut digunakan sebagai acuan dalam pembuatan peta klasterisasi. Areal yang masih tergolong rendah diharapkan dapat meningkatkan produktivitas padi. Algoritma PAM menghasilkan dua kluster dengan nilai koefisien siluet sebesar 0,58. Algoritma CLARA dengan 100 sampel menghasilkan tiga kluster dengan koefisien siluet sebesar 0,59. | Dapat dilakukan penelitian dengan menggunakan metode klasterisasi yang lain. | Pada penelitian ini menggunakan objek penelitian dan metode yang berbeda. Pada penelitian ini akan menggunakan metode cluster PAM dan cluster hierarki |
| 3 | Clustering and | Nurdin | Penelitian ini | Dengan | Dapat dilakukan | Pada penelitian ini |

| | | | | | |
|---|--|---|---|--|---|
| <p>Mapping of Agricultural Production Based on Geographic Information System Using KMedoids Algorithm</p> | <p>Nurdin1, Fajriana2, Rini Meiyanti3, Adelia4, Maya Maulita4, Journal of Artificial Intelligence and Technology, 2025</p> | <p>mengembangkan sistem pemetaan dan pengelompokan produksi pertanian dengan menggunakan algoritma K-Medoids.</p> | <p>menggunakan 25 titik data uji dari 27 kecamatan pada tahun 2023.</p> <p>kajian ini mengidentifikasi kluster produksi tinggi (C1) untuk komoditas seperti beras, kedelai, mangga, belimbing, jambu biji, jambu air, jeruk siam, jeruk besar, manggis, nanas, sawi, sirsak, sukun, melinjo, kacang panjang, dan duku. Kluster produksi sedang (C2) meliputi jagung, durian, pepaya, pisang, rambutan, nangka, dan jengkol, sedangkan kluster produksi rendah (C3) meliputi alpukat</p> | <p>penelitian dengan menggunakan metode klusterisasi yang berbeda atau dengan membandingkan dengan metode cluster yang lain.</p> | <p>menggunakan objek penelitian dan metode yang berbeda, pada penelitian ini menggunakan algoritma K-Medoids.</p> |
|---|--|---|---|--|---|

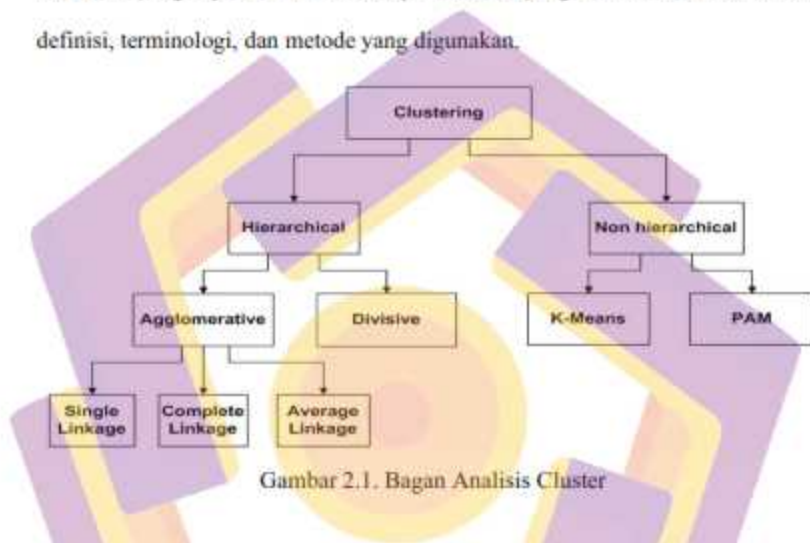
| | | | | | | |
|---|---|--|--|---|--|---|
| | | | | dan salah. | | |
| 4 | CREDIT CUSTOMER SEGMENTATION WITH HIERARCHICAL CLUSTERING AT VARIOUS DISTANCES | HENIDA RATNA AYU PUTRI ¹ , ADJI ACHMAD RINALDO FERNANDES ² , ATIEK IRIANY ³ , NURJANNAH ⁴ , SOLIMUN ⁵ , <i>Journal of Theoretical and Applied Information Technology</i> , 2023. | Tujuan penelitian ini adalah untuk melihat pengelompokan dan perbandingan hasil analisis kluster menggunakan metode ward linkage dengan berbagai jarak (Euclidean, Manhattan dan Mahalanobis) dengan penilaian variabel 5C (Character, Capacity, Capital, Collateral, Condition of Economy) pada nasabah KPR Bank X. | Hasil analisis kluster menggunakan metode Ward dengan jarak Mahalanobis adalah kluster satu terdiri dari 26 nasabah dengan aspek kepatuhan sebagai aspek tertinggi, kluster dua terdiri dari 30 nasabah dengan aspek kondisi sebagai aspek tertinggi, sedangkan kluster tiga terdiri dari 44 nasabah dengan aspek permodalan sebagai aspek tertinggi. | Penelitian dapat dilakukan dengan menggunakan metode klusterisasi yang berbeda atau dengan membandingkan dengan metode klusterisasi yang lain. | Pada penelitian ini menggunakan objek penelitian dan metode yang berbeda. Pada penelitian ini akan menggunakan metode cluster PAM dan cluster hierarki |
| 5 | Analisis Hard dan Soft Clustering Untuk Pengelompokan Indikator Ketahanan Pangan Indonesia 2021 | Winar Wahyu Prastanika#1, Arie Wahyu Wijayanto#2, <i>Jurnal Sistem dan Teknologi Informasi</i> , | Penelitian ini bertujuan untuk mengelompokkan provinsi-provinsi di Indonesia berdasarkan karakteristik indikator | Hasil penelitian ini menemukan bahwa metode K-Means merupakan metode terbaik dalam pengelompokan | Dapat dilakukan penelitian dengan menggunakan metode klusterisasi yang berbeda atau dengan menggunakan dataset yang lebih banyak. | Pada penelitian ini menggunakan objek penelitian dan metode cluster yang berbeda, pada penelitian ini menggunakan Metode clustering K- Means, K- Medoids, dan FCM |

| | | | | | |
|---|---|---|--|--|--|
| | | 2023 | ketahanan pangan dan membandingkan beberapa metode untuk mendapatkan metode pengelompokan yang terbaik. | provinsi-provinsi di Indonesia menurut tingkat ketahanan pangannya. Metode ini membentuk dua klaster. Klaster pertama terdiri dari 8 provinsi dengan tingkat kerawanan pangan, sedangkan klaster kedua terdiri dari 26 provinsi dengan tingkat ketahanan pangan. | |
| 6 | Cluster Around Latent Variable for Vulnerability Towards Natural Hazards, Non-Natural Hazards, Social Hazards in West Papua | REZZY EKO CARAKA ^{1,2,3*} YOUNGJO LEE ^{1,2} , RUNG CHING CHEN ³ , TONI TOHARUDIN ⁴ , PRANA UGIANA GIO ⁵ , ROBERT KURNIAWA N ⁶ , AND | Tujuan dari penelitian ini adalah untuk menghitung kerentanan sosial berbasis lokasi di Papua Barat yang melibatkan komponen Informasi, Teknologi, dan Komunikasi, Akses Pangan, Bencana Alam, Pernyataan Perlindungan | Setelah dilakukan pengelompokan variabel di sekitar variabel laten dengan nilai konektivitas 3.9400794, Dunn 0.9373, dan Silhouette 0.6333. Setiap faktor memberikan tanda yang menunjukkan pengaruh positif atau negatif terhadap kerentanan sosial | Dapat dilakukan penelitian dengan menggunakan metode klasterisasi yang lain. Pada penelitian ini menggunakan objek penelitian yang berbeda. |

| | | | | | | |
|--|--|---|---|---|--|--|
| | | BENS PARDAMEA N ^o , IEEE Access, 2021 | Sosial, Akses ke Layanan Keuangan, Deskripsi sumber pendapatan rumah tangga, Jumlah kejadian banjir, jumlah bencana gempa bumi, kasus kematian COVID-19, dan Jumlah insiden protes yang diperoleh dari Survei Sosial Ekonomi Nasional (SUSENAS) statistik keuangan dengan fokus utama penelitian pada generasi milenial. | dan akhirnya akan terbentuk klaster lokasi berdasarkan indeks yang diperoleh. | | |
|--|--|---|---|---|--|--|

2.3 Landasan Teori

Suatu penelitian membutuhkan teori-teori yang sudah terbukti kebenarannya dalam dunia ilmiah. Oleh karena itu, penelitian ini juga didukung oleh teori yang diperoleh dari beberapa referensi yang bermanfaat baik dari segi definisi, terminologi, dan metode yang digunakan.



Gambar 2.1. Bagan Analisis Cluster

2.3.1 Analisis Cluster

Analisis cluster adalah istilah umum untuk berbagai metode numerik dengan tujuan untuk menemukan kelompok atau kelompok pengamatan yang homogen dan terpisah dari kelompok lain. Analisis cluster merupakan teknik yang lebih primitif karena tidak ada asumsi yang dibuat mengenai jumlah kelompok. Pengelompokan dilakukan atas dasar persamaan atau jarak. Input yang diperlukan adalah ukuran kesamaan data dimana kesamaan dapat dihitung [13].

Secara umum metode cluster dapat dikelompokkan menjadi dua yaitu metode hierarchial dan metode non hierarchial. Metode hierarchial adalah metode yang membuat sebuah dekomposisi berhierarki dari kumpulan data berdasarkan kesamaan karakteristik objeknya. Sedangkan metode non hierarchial digunakan untuk pengelompokan objek, dimana jumlah klaster yang akan dibentuk dapat ditentukan sebelumnya. Beberapa jenis analisis cluster yang menggunakan metode hierarchial, diantaranya single linkage, complete linkage, average linkage. Sedangkan dalam analisis cluster non hierarchial, metode yang paling umum digunakan adalah metode K-Means [14].

2.3.2 Metode Hierarchical Clustering

Metode hierarchical clustering merupakan metode analisis cluster yang proses pengelompokannya dilakukan secara bertahap atau bertingkat, sehingga membentuk suatu tingkatan tertentu, seperti struktur pohon. Hasil clustering dengan menggunakan metode hierarchical clustering dapat ditampilkan dalam bentuk dendrogram. Dendrogram adalah grafik yang mengilustrasikan bagaimana pengelompokan terjadi [15]. Metode-metode yang bisa digunakan dalam metode hierarchical clustering adalah metode agglomerative (penyatuan) dan metode devisive (pembagian).

2.3.3 Metode Non Hierarchical

Metode non hierarchial disebut juga dengan metode k-means. Metode ini berbeda dengan metode hierarchial, pada metode non

hierarchical dimulai dengan menentukan terlebih dahulu beberapa cluster awal yang diinginkan, kemudian hasil dari objek pengamatan tersebut bergabung dan membentuk cluster. Pada metode non hierarchical terdiri dari beberapa algoritma diantaranya adalah Algoritma K-Means dan Algoritma K-Medoid.

2.3.4 Algoritma K-Means

Algoritma K-Means merupakan salah satu metode pengelompokan data non-hierarki yang dapat membagi data ke dalam bentuk dua atau lebih kelompok [16]. Data-data yang mempunyai karakteristik yang berbeda, dikelompokkan dengan kelompok atau cluster yang lain sehingga data yang berada dalam satu kelompok atau cluster memiliki tingkat variasi yang kecil.

2.3.5 Algoritma Partitioning Around Medoids (PAM)

Partitioning Around Medoids (PAM) merupakan algoritma clustering yang hampir mirip dengan algoritma clustering K-Means. Perbedaan antara kedua perhitungan ini adalah pada perhitungan Partitioning Around Medoids (PAM) memakai objek sebagai perwakilan (medoid) sebagai centroid untuk setiap cluster, sedangkan pada perhitungan algoritman K-Means memakai rata-rata sebagai centroid [17].

Algoritma Partitioning Around Medoids (PAM) memiliki keunggulan dalam mengatasi kekurangan yang ada pada algoritma K-Means, yang sensitif terhadap permasalahan data dimana terdapat data yang mempunyai perbedaan data yang sangat besar dengan data-data lain yang terdapat pada kumpulan

data. Hal ini dapat mengakibatkan terjadi persebaran data yang tidak seimbang [18].

Tahapan Algoritma Partitioning Around Medoids (PAM) sebagai berikut [19] :

1. Secara acak memilih k objek pada sekumpulan n objek untuk digunakan sebagai medoids.
2. Menggunakan Manhattan Distance untuk meletakkan objek non-medoids dalam cluster yang terdekat dengan medoids.
3. Secara acak menentukan o_{random} : objek non-medoids.
4. Menghitung S (selisih), dari total cost pertukaran medoids O_j dengan O_{random} .
5. Bila $S < 0$ ganti O_j dengan O_{random} untuk menghasilkan objek baru untuk digunakan sebagai medoid.
6. Ulang tahap 3 sampai 5 hingga medoids tidak mengalami perubahan, hingga didapatkan cluster serta anggota pada tiap-tiap cluster.
7. Mengulangi tahap 3 sampai 5 hingga tidak ada pergantian medoids, kemudian diperoleh cluster serta tiap-tiap anggota cluster. Pernyataan nilai dengan persamaan :

$$\text{Total Cost} = \sum \sqrt{\sum_{k=1}^n (x_k - y_k)^2} \quad \dots\dots\dots (1)$$

Dengan :

- n - jumlah sebuah data
- k - indikator data
- x_k - nilai atribut ke- k dari x

y_k - nilai atribut ke-k dari y

Nilai S dinyatakan dalam persamaan :

$$S = \text{Total Cost Baru} - \text{Total Cost Lama} \dots\dots\dots (2)$$

Dengan:

S - Selisih
 Total Cost Baru - Cost non-medoids
 Total Cost Lama - Cost medoids

2.3.6 Metode agglomerative

Metode agglomerative dimulai dari mengasumsikan bahwa semua objek adalah cluster. Dua objek yang jaraknya paling dekat digabungkan menjadi sebuah cluster. Proses ini berlanjut hingga terbentuk cluster yang terdiri dari seluruh objek. Agglomerative adalah metode pengelompokan hierarchial di mana setiap objek dimulai dalam cluster terpisah. Cluster dibentuk dengan cara mengelompokkan objek-objek menjadi cluster-cluster yang berisi objek lebih. Proses ini berlanjut hingga semua objek menjadi anggota dari satu cluster.

Langkah-langkah dalam menggunakan metode hierarchial agglomerasi untuk mengelompokkan N obyek [20] :

1. Mulai dengan N cluster, setiap cluster mengandung entiti tunggal dan sebuah matriks simetrik dari jarak (similarities) D - $\{d_{ik}\}$ dengan tipe $N \times N$.

2. Cari matriks jarak untuk pasangan cluster yang terdekat. Misalkan jarak antara cluster U dan V yang paling dekat adalah d_{uv} .
3. Gabungkan cluster U dan V. Label cluster yang baru dibentuk dengan (UV).

Perbarui entries pada matrik jarak dengan cara :

- a. Hapus baris dan kolom yang bersesuaian dengan cluster U dan V
 - b. Tambahkan baris dan kolom yang memberikan jarak-jarak antara cluster (UV) dan cluster - cluster yang tersisa.
4. Ulangi langkah 2 dan 3 sebanyak (N-1) kali (semua objek akan berada dalam cluster tunggal setelah algoritma berakhir). Catat identitas dari cluster yang digabungkan dan tingkat-tingkat (jarak atau similarities) dimana penggabungan terjadi.

Metode agglomerative ada beberapa macam, yaitu: metode single linkage (jarak terdekat), metode complete linkage (jarak terjauh) dan metode average linkage (jarak rata-rata).

2.3.7 Metode divisive

Metode divisive merupakan kebalikan dari metode agglomerative. Metode ini dimulai dari satu cluster yang mencakup semua objek pengamatan. Kemudian objek yang memiliki ketidakmiripan cukup besar akan dipisahkan menjadi kelompok baru, demikian seterusnya sampai

terbentuk cluster yang jumlahnya sama dengan jumlah objek yang dikelompokkan [13].

Pada setiap langkahnya dalam metode divisive terjadi penambahan kelompok ke dalam nilai dua nilai terkecil sampai akhirnya semua elemen bergabung. Metode divisive merupakan proses pengklasteran yang didasarkan pada persamaan nilai rata-rata antar objek. Ini berarti bahwa cluster hierarchical dibangun dalam $n-1$ langkah ketika data mengandung n objek [21].

2.3.8 Metode Single Linkage

Metode ini didasarkan pada jarak minimum. Dimulai dengan dua objek yang dipisahkan dengan jarak paling pendek maka keduanya akan ditempatkan pada cluster pertama, dan seterusnya. Metode ini dikenal pula dengan nama pendekatan tetangga terdekat. Pada awalnya, harus menemukan jarak terpendek dalam $D = \{d_{ij}\}$ dan menggabungkan objek-objek yang bersesuaian misalnya, U dan V , untuk mendapat klaster (UV) . Untuk langkah (3) dari algoritma di atas jarak-jarak antara (UV) dan cluster W yang lain dihitung dengan cara :

$$d_{UVW} = \min_{i \in \{U, V\}} \{d_{iW}\} \dots \dots \dots (3)$$

Besaran-besaran d_{UV} dan d_{VW} berturut-turut adalah jarak terpendek antara cluster-cluster U dan W dan juga cluster-cluster V dan W [22].

2.3.9 Metode Complete Linkage

Disebut juga pendekatan tetangga terjauh. Dasarnya adalah jarak maksimum. Dalam metode ini seluruh objek dalam suatu kluster dikaitkan satu sama lain pada suatu jarak maksimum atau dengan kesamaan minimum. Pada awalnya harus ditemukan jarak terpendek dalam $D = \{d_{ik}\}$ dan menggabungkan objek-objek yang bersesuaian misalnya, U dan V, untuk mendapat cluster (UV). Untuk langkah (3) dari algoritma di atas jarak-jarak antara (UV) dan cluster W yang lain dihitung dengan cara :

$$d_{(UV)W} = \max\{d_{UW}, d_{VW}\} \dots\dots\dots (4)$$

Besaran-besaran d_{UW} dan d_{VW} berturut-turut adalah jarak terjauh antara cluster-cluster U dan W dan juga cluster-cluster V dan W [13].

2.3.10 Metode Average Linkage

Dasarnya adalah jarak rata-rata antar observasi. Pengelompokan dimulai dari tetangga atau pasangan observasi dengan jarak paling mendekati jarak rata-rata. Pada awalnya harus ditemukan jarak terpendek dalam $D = \{d_{ik}\}$ dan menggabungkan objek-objek yang bersesuaian misalnya, U dan V, untuk mendapat cluster (UV). Untuk langkah (3) dari algoritma di atas jarak-jarak antara (UV) dan cluster W yang lain dihitung dengan cara :

$$d_{(UV)W} = \frac{\sum_i \sum_k d_{ik}}{N(UV)NW} \dots\dots\dots (5)$$

Dimana d_{ik} adalah jarak antara obyek i dalam cluster (UV) dan objek k dalam cluster W , dan N_{uv} dan N_w berturut-turut adalah banyaknya item-item dalam cluster (UV) dan W [23].

2.3.11 Pengukuran Jarak Kemiripan

Pengukuran jarak kemiripan yang akan digunakan pada penelitian ini adalah menggunakan metode manhattan distance.

Manhattan distance adalah metode pengukuran jarak yang digunakan untuk menghitung jumlah selisih mutlak antara dua titik suatu objek [24].

Proses perhitungan jarak menggunakan metode manhattan distance dapat dilakukan dengan menggunakan persamaan berikut :

$$d = \sum_{i=1}^n |x_i - y_i| \quad \dots\dots\dots (6)$$

Keterangan :

d : Jarak Manhattan antara dua titik / data

x_i : Nilai fitur ke- i dari data pertama

y_i : Nilai fitur ke- i dari data kedua

2.3.12 Standarisasi Data

Dalam data mining, data outlier dapat merujuk pada observasi atau titik data yang secara signifikan berbeda dari keseluruhan data didalam sebuah dataset. Data outlier juga sering kali merupakan data yang tidak biasa, dan data tersebut bisa menjadi penting karena data tersebut dapat mengindikasikan kesalahan dalam pengumpulan data, atau data tersebut bisa

menjadi informasi berharga yang dapat mengungkapkan pola atau anomali yang tidak terduga dalam keseluruhan data [25]. Salah satu teknik untuk mengatasi outlier pada data adalah dengan menggunakan metode robust scaler. Metode Robust Scaler merupakan salah satu teknik pre-processing dalam analisis data yang dapat digunakan untuk menormalkan data dengan mengurangi pengaruh outlier. Perhitungan dengan Metode Robust Scaler dilakukan dengan menggunakan persamaan berikut.

$$X_{Scaled} = \frac{X - Q1(X)}{Q3(X) - Q1(X)} \dots\dots\dots (7)$$

Keterangan:

X_{Scaled} : Nilai yang telah dinormalisasi

X : Nilai asli pada dataset

$Q1(X)$: Quartil pertama dari dataset X

$Q3(X)$: Quartil ketiga dari dataset X

2.3.13 Silhouette Coefficient

Silhouette Coefficient digunakan untuk melihat kualitas dari kluster yang terbentuk, seberapa baik suatu objek tersebut ditempatkan kedalam suatu cluster. Metode ini merupakan gabungan dari metode kohesi dan pemisahan. Kohesi digunakan untuk mengukur kedekatan objek yang berada pada satu cluster, sedangkan pemisahan digunakan untuk mengukur kedekatan antar cluster yang terbentuk. Nilai Koeffisien Silhouette suatu objek i sebagai berikut [26] :

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \dots\dots\dots (8)$$

Keterangan :

$a(i)$ – jarak rata-rata antara objek i dengan seluruh objek yang berada dalam kluster yang sama

$b(i)$ – jarak rata-rata minimum antara objek i dengan seluruh objek yang berada pada kluster terdekat

Koeffisien Silhouette memiliki rentang nilai antara -1 sampai 1. Nilai Koeffisien Silhouette yang mendekati 1 adalah yang lebih baik. Nilai Koeffisien Silhouette suatu kluster yang terdiri dari N objek dapat dihitung dengan rata-rata dari seluruh nilai Koeffisien Silhouette (SLH) setiap objek i sebagai berikut [27] :

$$SLH = \frac{1}{N} \sum_{i=1}^N S(i) \quad \dots\dots\dots (9)$$

Hasil pengklasteran terbaik dicapai ketika SLH maksimum, hal ini berarti meminimumkan jarak antar objek dalam kluster, sekaligus memaksimumkan jarak antar kluster [28].

2.3.14 Principal Componen Analisis (PCA)

Principal Componen Analisis (PCA) adalah suatu teknik statistik multivariat yang secara linear digunakan untuk mengubah bentuk sekumpulan data variabel asli menjadi kumpulan data variabel yang lebih kecil yang tidak berkorelasi yang dapat mewakili informasi dari kumpulan data variabel asli. Tujuan metode ini adalah untuk menyederhanakan data variabel yang diamati dengan menyusut, atau mengurangi, dimensinya. Ini dicapai dengan mengubah data variabel bebas awal menjadi data variabel baru yang tidak memiliki korelasi sama sekali [29].

Principal Component Analysis (PCA) merupakan salah satu metode yang paling populer yang dapat digunakan untuk pengurangan dimensi dan merupakan teknik statistik standar yang dapat digunakan untuk menyederhanakan dimensi suatu kumpulan data [30].

Principal Component Analysis (PCA) menghasilkan komponen utama yang diperoleh dari dekomposisi eigen value dan eigen vector dari matriks kovarians. Berikut merupakan langkah-langkah persamaan yang diperlukan dalam Principal Component Analysis (PCA) [31]. Sebagai berikut :

1. Menghitung mean (X) dari data tiap dimensi menggunakan persamaan berikut :

$$X = \frac{1}{n} \sum_{i=1}^n X_i \quad \dots\dots\dots (10)$$

Dengan :

n - Jumlah data sampel

X_i - Data Sampel

2. Menghitung covariance matrix (Cx) persamaan berikut :

$$Cx = \frac{1}{n-1} \sum_{i=1}^n (X_i - X)(X_i - X)^T \quad \dots\dots\dots (11)$$

Dengan :

n - Jumlah data sampel

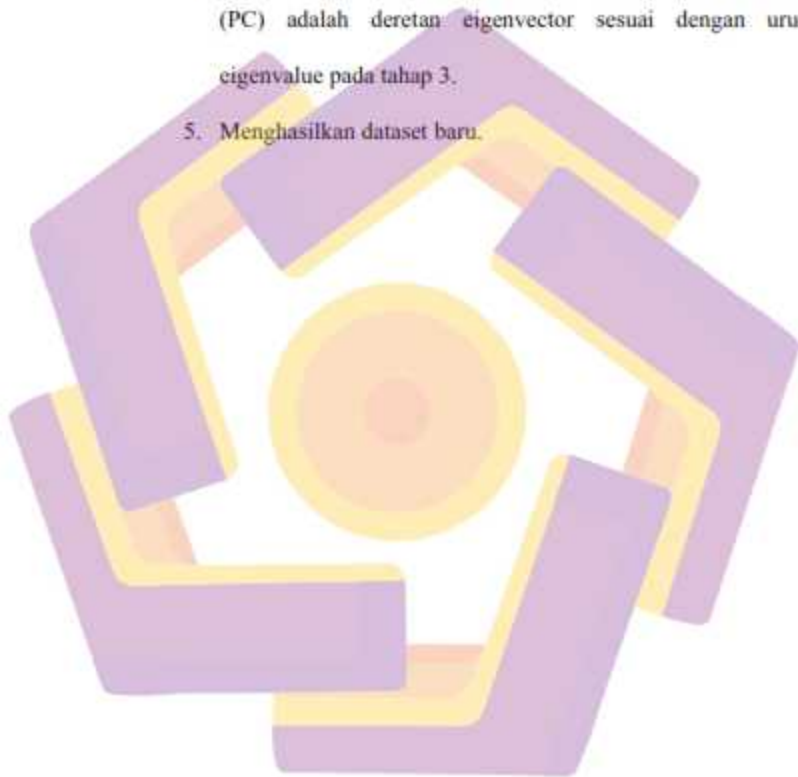
X_i - Data Sampel

X - Mean

3. Menghitung eigenvector (V_m) dan Eigenvalue (λ_m) dari covariance matrix menggunakan persamaan berikut :

$$C_x v_m = \lambda_m v_m \quad \dots\dots\dots (12)$$

4. Urutkan eigenvalue secara descending. Principal Component (PC) adalah deretan eigenvector sesuai dengan urutan eigenvalue pada tahap 3.
5. Menghasilkan dataset baru.



BAB 3

METODE PENELITIAN

3.1 Jenis, Sifat, dan Pendekatan Penelitian

Pada tahap ini akan dilakukan tahapan-tahapan penelitian untuk mengelompokkan daerah-daerah atau wilayah di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat yang memiliki potensi produksi komoditas tanaman pangan, tahapan-tahapan penelitian yang akan dilakukan adalah sebagai berikut :

- a. Menentukan rumusan masalah dan menentukan tujuan penelitian yang akan dilakukan,
- b. Mencari dataset yang diperlukan dalam penelitian. Pada penelitian ini, peneliti menggunakan dataset publik yang berasal dari website badan pusat statistik (BPS). Dataset yang akan digunakan dalam penelitian ini adalah data “Luas Panen Padi Menurut Kabupaten_Kota, 2023”, “Produksi Padi (GKG) , 2021-2023”, “Luas Panen, Produksi, dan Produktivitas Padi Menurut Kabupaten_Kota di Provinsi Jawa Tengah, 2021-2023”, “Luas Panen Tanaman Padi Kabupaten_Kota menurut Bulan di Provinsi Jawa Barat (Ha), 2023”, “Produksi Tanaman Padi Kabupaten_Kota menurut Bulan di Provinsi Jawa Barat (Ton-GKG), 2023”, “Produksi Tanaman Sayuran Menurut Kabupaten_Kota dan Jenis Tanaman di Provinsi Jawa Timur, 2023”, “Luas Panen Tanaman Sayuran dan Buah-Buahan Menurut Kabupaten_Kota dan Jenis Tanaman di Provinsi Jawa Timur, 2023”, “Produksi Tanaman Sayuran

Semusim Menurut Jenis Tanaman dan Kabupaten_Kota di Provinsi Jawa Tengah, 2023”, “Luas Panen Tanaman Sayuran Semusim Menurut Jenis Tanaman dan Kabupaten_Kota di Provinsi Jawa Tengah, 2023”, “Produksi Tanaman Sayuran Menurut Kabupaten_Kota dan Jenis Tanaman di Provinsi Jawa Barat, 2023”, “Luas Panen Tanaman Sayuran dan Buah-Buahan Menurut Kabupaten_Kota dan Jenis Tanaman di Provinsi Jawa Barat, 2023”, “Jumlah Kejadian Bencana Alam Menurut Kabupaten_Kota di Provinsi Jawa Timur, 2023”, “Jumlah Kejadian Bencana Alam Menurut Kabupaten_Kota di Provinsi Jawa Tengah, 2023”, “Jumlah Kejadian Bencana Alam Menurut Kabupaten_Kota di Provinsi Jawa Barat, 2023”, “Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten_Kota dan Lapangan Pekerjaan Utama di Provinsi Jawa Timur (rupiah) 2023”, “Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten_Kota dan Lapangan Pekerjaan Utama di Provinsi Jawa Tengah rupiah 2023”, “Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten_Kota dan Lapangan Pekerjaan Utama di Provinsi Jawa Barat (rupiah) 2023”.

- c. Melakukan seleksi dataset, dataset yang tidak memiliki nilai (N/A) akan diganti atau diisikan dengan nilai statistik dan selanjutnya akan dilakukan standardisasi dataset publik dengan menggunakan robust scale agar tidak memiliki rentang nilai yang terlalu lebar .

- d. Melakukan perhitungan jarak dataset yang telah distandarisasi menggunakan metode manhattan distance.
- e. Dari hasil perhitungan jarak data, kemudian akan dilakukan perhitungan silhouette score untuk mengetahui nilai k optimal yang akan dilakukan perhitungan pada pembuatan cluster. Pada penelitian ini cluster yang akan digunakan adalah Partitioning Around Medoids (PAM) dan hierarchical clustering dengan menggunakan metode agglomerative. Metode agglomerative yang digunakan menggunakan metode single linkage, complete linkage dan average linkage
- f. Langkah selanjutnya dari hasil perhitungan jarak data, kemudian hasil cluster akan divisualisasikan dalam bentuk dendrogram dan peta cluster untuk memudahkan dalam pembacaan daerah-daerah atau wilayah mana saja yang memiliki potensi produksi komoditas tanaman pangan. Visualisasi tersebut akan memudahkan dalam menganalisa pengelompokkan daerah-daerah atau wilayah mana saja yang memiliki potensi produksi komoditas tanaman pangan.

3.2 Metode Pengumpulan Data

Dalam tahap pengumpulan data, peneliti menggunakan dataset publik yang tersedia di website badan pusat statistik (BPS).

Dataset yang akan digunakan dalam penelitian ini adalah data “Luas Panen Padi Menurut Kabupaten_Kota, 2023”, “Produksi Padi (GKG) , 2021-2023”, “Luas Panen, Produksi, dan Produktivitas Padi Menurut Kabupaten_Kota di Provinsi Jawa Tengah, 2021-2023”, “Luas Panen Tanaman Padi Kabupaten_Kota

menurut Bulan di Provinsi Jawa Barat (Ha), 2023”, “Produksi Tanaman Padi Kabupaten_Kota menurut Bulan di Provinsi Jawa Barat (Ton-GKG), 2023”, “Produksi Tanaman Padi Kabupaten_Kota menurut Bulan di Provinsi Jawa Barat (Ton-GKG), 2023”, “Luas Panen Tanaman Sayuran dan Buah-Buahan Menurut Kabupaten_Kota dan Jenis Tanaman di Provinsi Jawa Timur, 2023”, “Produksi Tanaman Sayuran Semusim Menurut Jenis Tanaman dan Kabupaten_Kota di Provinsi Jawa Tengah, 2023”, “Luas Panen Tanaman Sayuran Semusim Menurut Jenis Tanaman dan Kabupaten_Kota di Provinsi Jawa Tengah, 2023”, “Produksi Tanaman Sayuran Menurut Kabupaten_Kota dan Jenis Tanaman di Provinsi Jawa Barat, 2023”, “Luas Panen Tanaman Sayuran dan Buah-Buahan Menurut Kabupaten_Kota dan Jenis Tanaman di Provinsi Jawa Barat, 2023”, “Jumlah Kejadian Bencana Alam Menurut Kabupaten_Kota di Provinsi Jawa Timur, 2023”, “Jumlah Kejadian Bencana Alam Menurut Kabupaten_Kota di Provinsi Jawa Tengah, 2023”, “Jumlah Kejadian Bencana Alam Menurut Kabupaten_Kota di Provinsi Jawa Barat, 2023”, “Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten_Kota dan Lapangan Pekerjaan Utama di Provinsi Jawa Timur (rupiah) 2023”, “Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten_Kota dan Lapangan Pekerjaan Utama di Provinsi Jawa Tengah rupiah 2023”, “Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten_Kota dan Lapangan Pekerjaan Utama di Provinsi Jawa Barat (rupiah) 2023”.

3.2.1. Preprocessing Data

Pada tahap preprocessing ini terdiri dari :

a. Pemilihan Data (Data Selection)

Pada tahap ini, peneliti akan menggunakan dataset yang diunduh dari website badan pusat statistik (BPS). Dataset yang akan digunakan pada penelitian ini difokuskan pada data tahun 2023.

b. Pembersihan data (Data Cleaning)

Pada tahap ini, akan mendeteksi data-data yang tidak konsisten dan data-data yang tidak memiliki nilai.

c. Transformasi Data (Data Transformation)

Pada tahap ini, dataset akan diubah kedalam bentuk tertentu untuk memenuhi kriteria data dalam metode Partitioning Around Medoids (PAM) dan hierarchical clustering menggunakan metode agglomerative.

3.3 Deskriptif Dataset

Dataset yang akan digunakan dalam penelitian ini diperoleh dari website Badan Pusat Statistik (BPS). Dataset yang diperoleh akan digunakan untuk mengelompokkan kabupaten / kota berdasarkan karakteristik dataset tersebut, seperti dari sektor pertanian dan produksi pangan, sektor bencana dan kerentanan iklim, dan sektor sosial-ekonomi. Sektor pertanian dan produksi pangan meliputi luas panen padi, produksi padi, luas panen cabai rawit, produksi cabai rawit, luas panen bawang daun, produksi bawang daun, luas panen cabai besar, produksi cabai besar, luas panen bayam, produksi bayam, luas panen kacang panjang, produksi kacang panjang, luas panen kangkung, produksi kangkung, luas panen

labu siam, produksi labu siam, luas panen sawi, produksi sawi, luas panen ketimun, produksi ketimun, luas panen terung, produksi terung, luas panen tomat, produksi tomat, luas panen jamur tiram, produksi jamur tiram. Pada sektor bencana dan kerentanan iklim, meliputi jumlah bencana alam – banjir, jumlah bencana alam – kebakaran hutan dan lahan, jumlah bencana alam - cuaca ekstrem. Pada sektor sosial-ekonomi, meliputi rata-rata upah / gaji bersih pekerja informal berdasarkan lapangan pekerjaan utama pertanian. Dataset yang diperoleh berupa data unsupervised, yaitu data yang tidak memiliki label atau target output. Dataset yang diperoleh tidak ada strukturnya, sangat random dan didalamnya memiliki potensi kelompok data, sehingga dataset tersebut sangat cocok untuk digunakan dalam proses clustering. Dataset yang diperoleh masih terdapat banyak nilai yang kosong sehingga dapat menyebabkan outlier pada saat proses clustering. Dataset yang masih terdapat nilai yang kosong tersebut diperlukan perlakuan khusus, seperti imputasi atau pengisian nilai kosong. Pada penelitian ini, pengisian data kosong adalah dengan mengisi dengan nilai nol apabila pada dataset tersebut tidak ada kejadian pada daerah tersebut. Selain mengisi dengan nilai nol, pengisian data kosong dilakukan juga dengan mengganti dengan nilai median.

3.4 Metode Analisis Data

Pada tahap ini, data dalam penelitian ini akan dianalisis dengan langkah-langkah sebagai berikut :

3.4.1. Preprocessing

Tahapan dimana untuk menyiapkan data agar dapat memenuhi standar data untuk proses Partitioning Around Medoids

(PAM) dan hierarchical clustering dengan menggunakan metode agglomerative. Pada tahapan ini meliputi:

a. Pengumpulan Data

Pada tahap ini peneliti menggunakan dataset publik yang tersedia diwebsite badan pusat statistik (BPS). Dataset publik yang didapat dari website badan pusat statistik (BPS) berupa file excel. Dataset publik tersebut memiliki beberapa atribut, antara lain : kabupaten / kota, luas panen padi, produksi padi, luas panen cabai rawit, produksi cabai rawit, luas panen bawang daun, produksi bawang daun, luas panen cabai besar, produksi cabai besar, luas panen bayam, produksi bayam, luas panen kacang panjang, produksi kacang panjang, luas panen kangkung, produksi kangkung, luas panen labu siam, produksi labu siam, luas panen sawi, produksi sawi, luas panen ketimun, produksi ketimun, luas panen terung, produksi terung, luas panen tomat, produksi tomat, luas panen jamur tiram, produksi jamur tiram, jumlah bencana alam – banjir, jumlah bencana alam – kebakaran hutan dan lahan, jumlah bencana alam - cuaca ekstrem, dan rata-rata upah/gaji bersih pekerja informal berdasarkan lapangan pekerjaan utama pertanian. Dataset yang akan digunakan pada penelitian ini difokuskan pada data tahun 2023.

b. Pemilihan Data (Data Selection)

Pada tahap pemilihan data digunakan untuk mengurangi kompleksitas atribut yang akan dikelola pada tahap transformasi data, sehingga hanya atribut yang akan di analisis saja yang akan dipilih.

c. Pembersihan data (Data Cleaning)

Pada tahap ini untuk mendeteksi data-data yang memiliki ciri-ciri sebagai berikut, data yang kurang lengkap, data yang kosong, data yang tidak konsisten.

d. Transformasi Data (Data Transformation)

Pada tahap ini adalah untuk proses perubahan data menjadi data yang sesuai kategori data dalam algoritma clustering agar proses algoritma dapat berjalan optimal.

3.4.2. Metode Hierarchical Clustering

Pada tahap ini peneliti menggunakan metode hierarchial clustering dengan menggunakan metode agglomerative. Metode agglomerative yang digunakan antara lain metode single linkage, metode complete linkage dan metode average linkage. Tiap metode akan dilakukan perhitungan jarak tiap data menggunakan metode manhattan distance. Dalam perhitungan jarak menggunakan metode manhattan distance akan menghasilkan matrik jarak yang akan dijadikan acuan dalam pembentukan cluster.

3.4.3. Metode Partitioning Around Medoids (PAM)

Pada tahap ini, selain menggunakan metode hierarchial clustering dengan menggunakan metode agglomerative. Peneliti juga menggunakan

metode partitioning around medoids (PAM). Pada metode partitioning around medoids (PAM) ini perhitungan jarak akan dilakukan dengan menggunakan metode manhattan distance. Pada perhitungan jarak akan menghasilkan matrik jarak, selanjutnya akan dilakukan perhitungan untuk pembentukan cluster dengan menggunakan metode partitioning around medoids (PAM). Langkah selanjutnya adalah penentuan nilai k optimal yang sebaiknya akan digunakan. Penentuan nilai k optimal dilakukan dengan perhitungan silhouette score. Setelah menentukan nilai k optimal, langkah selanjutnya adalah melakukan visualisasi hasil cluster menggunakan k optimal yang telah dihasilkan dari silhouette score.

3.4.4. Pembuatan Dendrogram dan Peta Cluster

Pada tahap ini, setelah dilakukan perhitungan jarak menggunakan metode manhattan distance, selanjutnya melakukan perhitungan menggunakan metode partitioning around medoids (PAM) dan hierarchical clustering dengan menggunakan metode agglomerative. Selanjutnya akan divisualisasikan hasilnya dalam bentuk dendrogram dan peta cluster untuk mengetahui persebaran daerah-daerah atau wilayah dengan potensi produksi komoditas tanaman pangan khususnya di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat.

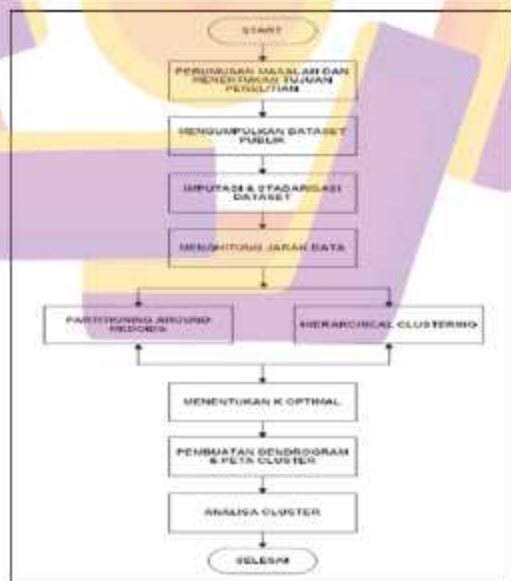
3.4.5. Analisa Cluster

Pada tahap ini, setelah perhitungan jarak data selesai selanjutnya akan divisualisasikan hasilnya dalam bentuk dendrogram dan peta cluster. Dendrogram dan peta cluster yang terbentuk dapat digunakan untuk

menunjukkan persebaran daerah-daerah atau wilayah di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat yang memiliki potensi produksi komoditas tanaman pangan. Pada tahap analisa cluster ini, jumlah cluster (k) optimal dapat ditunjukkan dengan menggunakan nilai silouette yang tertinggi. Penentuan cluster yang sudah divisualisasikan dengan menggunakan dendrogram dan peta cluster ini selanjutnya dapat dilakukan analisa hasil untuk mengetahui daerah-daerah atau wilayah khususnya di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat yang memiliki potensi produksi komoditas tanaman pangan.

3.5 Alur Penelitian

Pada penelitian ini, peneliti akan menggunakan alur penelitian sebagai berikut :



Gambar 3.1 Alur Penelitian

Pada alur penelitian ini, peneliti akan melakukan tahapan-tahapan penelitian dalam mengelompokkan daerah-daerah atau wilayah di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat yang memiliki potensi produksi komoditas tanaman pangan, sebagai berikut :

- a. Menentukan rumusan masalah dan menentukan tujuan penelitian yang akan dilakukan.
- b. Mencari dataset yang diperlukan untuk penelitian. Pada penelitian ini, peneliti menggunakan dataset publik yang berasal dari website badan pusat statistik (BPS). Dataset yang akan digunakan dalam penelitian ini adalah data “Luas Panen Padi Menurut Kabupaten_Kota, 2023”, “Produksi Padi (GKG) , 2021-2023”, “Luas Panen, Produksi, dan Produktivitas Padi Menurut Kabupaten_Kota di Provinsi Jawa Tengah, 2021-2023”, “Luas Panen Tanaman Padi Kabupaten_Kota menurut Bulan di Provinsi Jawa Barat (Ha), 2023”, “Produksi Tanaman Padi Kabupaten_Kota menurut Bulan di Provinsi Jawa Barat (Ton-GKG), 2023”, “Produksi Tanaman Padi Kabupaten_Kota menurut Bulan di Provinsi Jawa Barat (Ton-GKG), 2023”, “Luas Panen Tanaman Sayuran dan Buah-Buahan Menurut Kabupaten_Kota dan Jenis Tanaman di Provinsi Jawa Timur, 2023”, “Produksi Tanaman Sayuran Semusim Menurut Jenis Tanaman dan Kabupaten_Kota di Provinsi Jawa Tengah, 2023”, “Luas Panen Tanaman Sayuran Semusim Menurut Jenis Tanaman dan Kabupaten_Kota di Provinsi Jawa Tengah, 2023”, “Produksi Tanaman Sayuran Menurut Kabupaten_Kota dan

Jenis Tanaman di Provinsi Jawa Barat, 2023”, “Luas Panen Tanaman Sayuran dan Buah-Buahan Menurut Kabupaten_Kota dan Jenis Tanaman di Provinsi Jawa Barat, 2023”, “Jumlah Kejadian Bencana Alam Menurut Kabupaten_Kota di Provinsi Jawa Timur, 2023”, “Jumlah Kejadian Bencana Alam Menurut Kabupaten_Kota di Provinsi Jawa Tengah, 2023”, “Jumlah Kejadian Bencana Alam Menurut Kabupaten_Kota di Provinsi Jawa Barat, 2023”, “Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten_Kota dan Lapangan Pekerjaan Utama di Provinsi Jawa Timur (rupiah) 2023”, “Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten_Kota dan Lapangan Pekerjaan Utama di Provinsi Jawa Tengah rupiah 2023”, “Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten_Kota dan Lapangan Pekerjaan Utama di Provinsi Jawa Barat (rupiah) 2023”.

- c. Melakukan imputasi dataset yang memiliki nilai nol dan yang tidak memiliki nilai (N/A), dan melakukan standarisasi dataset agar memenuhi kriteria dalam analisis cluster. Perhitungan cluster yang akan digunakan pada penelitian ini menggunakan metode partitioning around medoids (PAM) dan hierarchical clustering. Model metode hierarchical clustering yang akan digunakan adalah metode agglomerative, dan metode agglomerative yang akan digunakan adalah dengan metode single linkage, complete linkage dan average linkage.

- d. Melakukan perhitungan jarak data menggunakan metode manhattan distance. Metode cluster yang akan digunakan, adalah metode partitioning around medoids (PAM) dan hierarchical clustering, hierarchical clustering yang digunakan menggunakan metode agglomerative. Metode agglomerative yang akan digunakan adalah dengan metode single linkage, complete linkage dan average linkage.
- e. Melakukan perhitungan cluster, dengan menggunakan metode partitioning around medoids (PAM) dan hierarchical clustering
- f. Menentukan nilai k optimal dengan menggunakan silhouette score
- g. Memvisualisasikan cluster hasil dari perhitungan k optimal menggunakan silhouette score. Visualisasi dalam bentuk dendrogram dan selanjutnya memvisualisasikan dalam bentuk peta cluster map untuk mengetahui persebaran kabupaten / kota di Provinsi Jawa, khususnya Jawa Timur, Jawa Tengah dan Jawa Barat.
- h. Hasil pembentukan dendrogram dan peta cluster selanjutnya dapat digunakan untuk analisa cluster dalam mengelompokkan daerah-daerah atau wilayah di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat yang memiliki potensi produksi komoditas tanaman pangan.

BAB 4

HASIL PENELITIAN DAN PEMBAHASAN

4.1 Pengumpulan Data

Dalam tahap pengumpulan data, peneliti menggunakan dataset publik yang tersedia di website badan pusat statistik (BPS). Dataset yang akan digunakan dalam penelitian ini adalah data “Luas Panen Padi Menurut Kabupaten_Kota, 2023”, “Produksi Padi (GKG) , 2021-2023”, “Luas Panen, Produksi, dan Produktivitas Padi Menurut Kabupaten_Kota di Provinsi Jawa Tengah, 2021-2023”, “Luas Panen Tanaman Padi Kabupaten_Kota menurut Bulan di Provinsi Jawa Barat (Ha), 2023”, “Produksi Tanaman Padi Kabupaten_Kota menurut Bulan di Provinsi Jawa Barat (Ton-GKG), 2023”, “Produksi Tanaman Sayuran Menurut Kabupaten_Kota dan Jenis Tanaman di Provinsi Jawa Timur, 2023”, 2023”, “Luas Panen Tanaman Sayuran dan Buah-Buahan Menurut Kabupaten_Kota dan Jenis Tanaman di Provinsi Jawa Timur, 2023”, “Produksi Tanaman Sayuran Semusim Menurut Jenis Tanaman dan Kabupaten_Kota di Provinsi Jawa Tengah, 2023”, “Luas Panen Tanaman Sayuran Semusim Menurut Jenis Tanaman dan Kabupaten_Kota di Provinsi Jawa Tengah, 2023”, “Produksi Tanaman Sayuran Menurut Kabupaten_Kota dan Jenis Tanaman di Provinsi Jawa Barat, 2023”, “Luas Panen Tanaman Sayuran dan Buah-Buahan Menurut Kabupaten_Kota dan Jenis Tanaman di Provinsi Jawa Barat, 2023”, “Jumlah Kejadian Bencana Alam Menurut Kabupaten_Kota di Provinsi Jawa Timur, 2023”, “Jumlah Kejadian Bencana Alam Menurut Kabupaten_Kota di Provinsi Jawa

Tengah, 2023”, “Jumlah Kejadian Bencana Alam Menurut Kabupaten_Kota di Provinsi Jawa Barat, 2023”, “Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten_Kota dan Lapangan Pekerjaan Utama di Provinsi Jawa Timur (rupiah) 2023”, “Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten_Kota dan Lapangan Pekerjaan Utama di Provinsi Jawa Tengah rupiah 2023”, “Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten_Kota dan Lapangan Pekerjaan Utama di Provinsi Jawa Barat (rupiah) 2023”. Dataset publik tersebut memiliki beberapa atribut, antara lain : kabupaten / kota, luas panen padi, produksi padi, luas panen cabai rawit, produksi cabai rawit, luas panen bawang daun, produksi bawang daun, luas panen cabai besar, produksi cabai besar, luas panen bayam, produksi bayam, luas panen kacang panjang, produksi kacang panjang, luas panen kangkung, produksi kangkung, luas panen labu siam, produksi labu siam, luas panen sawi, produksi sawi, luas panen ketimun, produksi ketimun, luas panen terung, produksi terung, luas panen tomat, produksi tomat, luas panen jamur tiram, produksi jamur tiram, jumlah bencana alam – banjir, jumlah bencana alam – kebakaran hutan dan lahan, jumlah bencana alam - cuaca ekstrem, dan rata-rata upah/gaji bersih pekerja informal berdasarkan lapangan pekerjaan utama pertanian. Dataset yang akan digunakan pada penelitian ini adalah dataset pada tahun 2023.

4.2 Deskriptif Produksi Komoditas Tanaman

Hasil analisis statistik deskriptif untuk menggambarkan produksi komoditas tanaman pangan di beberapa provinsi di Indonesia khususnya Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat diperoleh dari dataset publik pada

website Badan Pusat Statistik (BPS). Hasil analisis statistik deskriptif dapat disajikan pada Tabel 4.1. Pada penelitian ini, dalam menentukan daerah-daerah yang memiliki potensi produksi komoditas tanaman pangan difokuskan pada kabupaten / kota di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat, luas panen padi, produksi padi, luas panen cabai rawit, produksi cabai rawit, luas panen bawang daun, produksi bawang daun, luas panen cabai besar, produksi cabai besar, luas panen bayam, produksi bayam, luas panen kacang panjang, produksi kacang panjang, luas panen kangkung, produksi kangkung, luas panen labu siam, produksi labu siam, luas panen sawi, produksi sawi, luas panen ketimun, produksi ketimun, luas panen terung, produksi terung, luas panen tomat, produksi tomat, luas panen jamur tiram, produksi jamur tiram, jumlah bencana alam – banjir, jumlah bencana alam – kebakaran hutan dan lahan, jumlah bencana alam - cuaca ekstrem, dan rata-rata upah / gaji bersih pekerja informal berdasarkan lapangan pekerjaan utama pertanian.

Tabel 4.1. Hasil analisis statistik deskriptif

| | N | Minimum | Maximum | Mean | Std. Deviation |
|------------------------|-----|---------|------------|-------------|----------------|
| Luas_Panen_Padi | 100 | 5.82 | 230456.96 | 49245.0082 | 45342.56230 |
| Padi | 100 | 30.51 | 1424303.10 | 279348.0873 | 265543.16031 |
| Luas_Panen_Cabai_Rawit | 96 | .37 | 9881.00 | 1240.7671 | 2169.47675 |
| Cabai_Rawit | 96 | .40 | 116031.78 | 10132.4574 | 20024.54707 |
| Luas_Panen_Bawang_Daun | 59 | .21 | 3628.00 | 563.8727 | 852.14764 |
| Bawang_Daun | 59 | .36 | 57195.70 | 6946.7187 | 12269.82743 |

| | | | | | |
|---------------------------|----|------|-----------|------------|-------------|
| Luas_Panen_Cabai_Besar | 89 | .02 | 3383.00 | 267.2820 | 465.31233 |
| Cabai_Besar | 89 | .01 | 56999.00 | 3064.6967 | 7194.61102 |
| Luas_Panen_Bayam | 78 | .03 | 1850.90 | 129.3191 | 279.30744 |
| Bayam | 78 | .19 | 13742.95 | 749.1540 | 1945.65288 |
| Luas_Panen_Kacang_Panjang | 86 | .01 | 959.59 | 123.3043 | 184.99628 |
| Kacang_Panjang | 86 | .06 | 13097.50 | 1382.7361 | 2695.66322 |
| Luas_Panen_Kangkung | 91 | .03 | 2053.30 | 160.4431 | 305.97957 |
| Kangkung | 91 | .03 | 21782.37 | 1363.5798 | 2912.33046 |
| Luas_Panen_Labu_Siam | 54 | .01 | 649.50 | 64.1774 | 115.17742 |
| Labu_Siam | 54 | .08 | 86907.50 | 5068.3883 | 15158.51157 |
| Luas_Panen_Sawi | 86 | .44 | 2327.00 | 311.6143 | 519.62473 |
| Sawi | 86 | .20 | 48796.50 | 3941.2937 | 8284.77599 |
| Luas_Panen_Ketimun | 79 | .41 | 1151.74 | 153.0229 | 235.44427 |
| Ketimun | 79 | 1.60 | 23685.55 | 2541.1529 | 4884.81519 |
| Luas_Panen_Terung | 89 | .11 | 1097.00 | 139.3060 | 205.97777 |
| Terung | 89 | .96 | 50912.27 | 2891.6407 | 6609.04167 |
| Luas_Panen_Tomat | 87 | .08 | 3689.00 | 219.1968 | 495.66964 |
| Tomat | 87 | .46 | 104839.60 | 5387.3298 | 14853.32536 |
| Luas_Panen_Jamur_tiram | 72 | 3.00 | 478629.00 | 24233.6443 | 60160.17490 |
| Jamur_Tiram | 72 | .04 | 20963.79 | 657.4476 | 2528.20476 |
| Banjir | 71 | 1.00 | 24.00 | 3.4225 | 3.45858 |
| Kerhutia | 62 | 1.00 | 36.00 | 6.1935 | 8.09339 |
| Cuaca_Ekstrem | 78 | 1.00 | 145.00 | 7.1282 | 18.78617 |

| | | | | | |
|--------------------|----|-----------|------------|--------------|--------------|
| Upah | 91 | 599972.00 | 3006492.00 | 1299089.6813 | 381380.03517 |
| Valid N (listwise) | 18 | | | | |

Berdasarkan Tabel 4.1. hasil analisis statistik deskriptif. Luas panen, produksi komoditas tanaman pangan, jumlah bencana alam dan rata-rata upah / gaji bersih pekerja informal berdasarkan lapangan pekerjaan utama pertanian di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat setiap daerah memiliki pola data yang berbeda-beda, dataset awal yang diperoleh di website Badan Pusat Statistik (BPS) masih terdapat dataset yang tidak memiliki nilai (N/A). Dataset yang tidak memiliki nilai (N/A) tersebut dapat disebabkan karena belum terisinya dataset tersebut atau dapat juga disebabkan luas panen, jumlah produksi komoditas tanaman pangan, jumlah bencana alam dan rata-rata upah / gaji bersih pekerja informal berdasarkan lapangan pekerjaan utama pertanian yang tidak terdapat di daerah tersebut.

Tabel 4.2. Hasil analisis statistik deskriptif luas panen

| | N | Minimum | Maximum | Mean | Std. Deviation |
|---------------------------|-----|---------|-----------|------------|----------------|
| Luas_Panen_Padi | 100 | 5.82 | 230456.96 | 49245.0082 | 45342.56230 |
| Luas_Panen_Cabai_Rawit | 96 | .37 | 9681.00 | 1240.7671 | 2169.47675 |
| Luas_Panen_Bawang_Daun | 59 | .21 | 3628.00 | 563.8727 | 852.14764 |
| Luas_Panen_Cabai_Besar | 89 | .02 | 3383.00 | 267.2820 | 465.31233 |
| Luas_Panen_Bayam | 78 | .03 | 1850.90 | 129.3191 | 279.30744 |
| Luas_Panen_Kacang_Panjang | 86 | .01 | 959.59 | 123.3043 | 184.99628 |
| Luas_Panen_Kangkung | 91 | .03 | 2053.30 | 160.4431 | 305.97957 |

| | | | | | |
|------------------------|----|------|-----------|------------|-------------|
| Luas_Panen_Labu_Siam | 54 | .01 | 649.50 | 64.1774 | 115.17742 |
| Luas_Panen_Sawi | 86 | .44 | 2327.00 | 311.6143 | 519.62473 |
| Luas_Panen_Ketimun | 79 | .41 | 1151.74 | 153.0229 | 235.44427 |
| Luas_Panen_Terung | 89 | .11 | 1097.00 | 139.3060 | 205.97777 |
| Luas_Panen_Tomat | 87 | .08 | 3689.00 | 219.1968 | 495.66964 |
| Luas_Panen_Jamur_tiram | 72 | 3.00 | 478629.00 | 24233.6443 | 60160.17490 |
| Valid N (listwise) | 32 | | | | |

Pada tabel 4.2. Hasil analisis statistik deskriptif luas panen, terdapat dataset yang tidak memiliki nilai (N/A), diantaranya sebagai berikut :

1. Luas Panen Cabai Rawit

Pada luas panen cabai rawit, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Tegal (Jawa Tengah), Kota Cirebon (Jawa Barat).

2. Luas Panen Bawang Daun

Pada luas panen bawang daun, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Pacitan (Jawa Timur), Kabupaten Jember (Jawa Timur), Kabupaten Situbondo (Jawa Timur), Kabupaten Sidoarjo (Jawa Timur), Kabupaten Bojonegoro (Jawa Timur), Kabupaten Lamongan (Jawa Timur), Kabupaten Gresik (Jawa Timur), Kabupaten Bangkalan (Jawa Timur), Kabupaten Sampang (Jawa Timur), Kabupaten Sumenep (Jawa Timur), Kota Kediri (Kota), Kota Blitar (Jawa Timur), Kota Malang (Jawa

Timur), Kota Probolinggo (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kota Surabaya (Jawa Timur), Kabupaten Kebumen (Jawa Tengah), Kabupaten Purworejo (Jawa Tengah), Kabupaten Sukoharjo (Jawa Tengah), Kabupaten Sragen (Jawa Tengah), Kabupaten Grobongan (Jawa Tengah), Kabupaten Blora (Jawa Tengah), Kabupaten Rembang (Jawa Tengah), Kabupaten Pati (Jawa Tengah), Kabupaten Kudus (Jawa Tengah), Kabupaten Jepara (Jawa Tengah), Kota Magelang (Jawa Tengah), Kota Surakarta (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Cirebon (Jawa Barat), Karawang (Jawa Barat), Kota Bogor (Jawa Barat), Kota Sukabumi (Jawa Barat), Kota Cirebon (Jawa Barat), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat), Kota Cimahi (Jawa Barat).

3. Luas Panen Cabai Besar

Pada luas panen cabai besar, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Pacitan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kabupaten Grobongan (Jawa Tengah), Kabupaten Jepara (Jawa Tengah), Kota Magelang (Jawa Tengah), Kota Semarang (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat).

4. Luas Panen Bayam

Pada luas panen bayam, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Bondowoso (Jawa Timur), Kabupaten Situbondo (Jawa Timur), Kabupaten Probolinggo (Jawa Timur), Kabupaten Mojokerto (Jawa Timur), Kabupaten Magetan (Jawa Timur), Kabupaten Sampang (Jawa Timur), Kota Kediri (Jawa Timur), Kota Blitar (Jawa Timur), Kota Malang (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kabupaten Banjarnegara (Jawa Tengah), Kabupaten Wonosobo (Jawa Tengah), Kabupaten Sukoharjo (Jawa Tengah), Kabupaten Sragen (Jawa Tengah), Kabupaten Kudus (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kuningan (Jawa Barat), Majalengka (Jawa Barat), Subang (Jawa Barat), Kota Cirebon (Jawa Barat).

5. Luas Panen Kacang Panjang

Pada luas panen kacang panjang, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Situbondo (Jawa Timur), Kabupaten Sidoarjo (Jawa Timur), Kabupaten Mojokerto (Jawa Timur), Kota Malang (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kota Surabaya (Jawa Timur), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Kota Cirebon (Jawa Barat), Kota Depok (Jawa Barat), Kota Cimahi (Jawa Barat).

6. Luas Panen Kangkung

Pada luas panen kangkung, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Bondowoso (Jawa Timur), Kabupaten Situbondo (Jawa Timur), Kabupaten Probolinggo (Jawa Timur), Kota Kediri (Jawa Timur), Kota Pasuruan (Jawa Timur), Kabupaten Sukoharjo (Jawa Tengah), Kabupaten Karanganyar (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Kota Cirebon (Jawa Barat), Kota Depok (Jawa Barat), Kota Cimahi (Jawa Barat).

7. Luas Panen Labu Siam

Pada luas panen labu siam, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Ponorogo (Jawa Timur), Kabupaten Lumajang (Jawa Timur), Kabupaten Situbondo (Jawa Timur), Kabupaten Sidoarjo (Jawa Timur), Kabupaten Mojokerto (Jawa Timur), Kabupaten Madiun (Jawa Timur), Kabupaten Bojonegoro (Jawa Timur), Kabupaten Tuban (Jawa Timur), Kabupaten Lamongan (Jawa Timur), Kabupaten Gresik (Jawa Timur), Kabupaten Sampang (Jawa Timur), Kabupaten Sumenep (Jawa Timur), Kota Kediri (Jawa Timur), Kota Blitar (Jawa Timur), Kota Malang (Jawa Timur), Kota Probolinggo (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kota Surabaya (Jawa Timur), Kabupaten Banyumas (Jawa Tengah),

Kabupaten Kebumen (Jawa Tengah), Kabupaten Purworejo (Jawa Tengah), Kabupaten Sukoharjo (Jawa Tengah), Kabupaten Sragen (Jawa Tengah), Kabupaten Blora (Jawa Tengah), Kabupaten Pati (Jawa Tengah), Kabupaten Jepara (Jawa Tengah), Kabupaten Demak (Jawa Tengah), Kota Magelang (Jawa Tengah), Kota Surakarta (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Kuningan (Jawa Barat), Cirebon (Jawa Barat), Karawang (Jawa Barat), Bekasi (Jawa Barat), Pangandaran (Jawa Barat), Kota Sukabumi (Jawa Barat), Kota Bandung (Jawa Barat), Kota Cirebon (Jawa Barat), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat), Kota Tasikmalaya (Jawa Barat), Kota Banjar (Jawa Barat).

8. Luas Panen Sawi

Pada luas panen labu siam, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Madiun (Jawa Timur), Kabupaten Bangkalan (Jawa Timur), Kabupaten Sampang (Jawa Timur), Kota Kediri (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kabupaten Purworejo (Jawa Tengah), Kabupaten Sukoharjo (Jawa Tengah), Kabupaten Sragen (Jawa Tengah), Kabupaten Kudus (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Cirebon (Jawa Barat).

9. Luas Panen Ketimun

Pada luas panen ketimun, terdapat daerah kabupaten / kota yang tidak

memiliki nilai (N/A), diantara adalah daerah Kabupaten Situbondo (Jawa Timur), Kabupaten Probolinggo (Jawa Timur), Kabupaten Mojokerto (Jawa Timur), Kabupaten Nganjuk (Jawa Timur), Kabupaten Gresik (Jawa Timur), Kota Kediri (Jawa Timur), Kota Blitar (Jawa Timur), Kota Malang (Jawa Timur), Kota Probolinggo (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kota Surabaya (Jawa Timur), Kabupaten Sukoharjo (Jawa Tengah), Kota Surakarta (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Tegal (Jawa Tengah), Kota Cirebon (Jawa Barat), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat), Kota Cimahi (Jawa Barat).

10. Luas Panen Terung

Pada luas panen terung, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Nganjuk (Jawa Timur), Kota Malang (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Kuningan (Jawa Barat), Kota Bogor (Jawa Barat), Kota Depok (Jawa Barat).

11. Luas Panen Tomat

Pada luas panen ketimun, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Nganjuk (Jawa Timur), Kota Kediri (Jawa Timur), Kota Malang (Jawa Timur), Kota

Pasuruan (Jawa Timur), Kota Madiun (Jawa Timur), Kabupaten Sragen (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat), Kota Cimahi (Jawa Barat).

12. Luas Panen Jamur Tiram

Pada luas panen jamur tiram, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Lumajang (Jawa Timur), Kabupaten Bondowoso (Jawa Timur), Kabupaten Jombang (Jawa Timur), Kabupaten Lamongan (Jawa Timur), Kabupaten Bangkalan (Jawa Timur), Kabupaten Pamekasan (Jawa Timur), Kabupaten Sumenep (Jawa Timur), Kota Kediri (Jawa Timur), Kota Probolinggo (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kabupaten Sukoharjo (Jawa Tengah), Kabupaten Grobongan (Jawa Tengah), Kabupaten Kudus (Jawa Tengah), Kabupaten Pemalang (Jawa Tengah), Kabupaten Surakarta (Jawa Tengah), Kabupaten Salatiga (Jawa Tengah), Garut (Jawa Barat), Pangandaran (Jawa Barat), Kota Bogor (Jawa Barat), Kota Sukabumi (Jawa Barat), Kota Bandung (Jawa Barat), Kota Cirebon (Jawa Barat), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat), Kota Cimahi (Jawa Barat).

Tabel 4.3. Hasil Analisis Statistik Deskriptif Produksi Tanaman Pangan

| | N | Minimum | Maximum | Mean | Std. Deviation |
|--------------------|-----|---------|------------|-------------|----------------|
| Padi | 100 | 30.51 | 1424303.10 | 279348.0873 | 265543.16031 |
| Csbai_Rawit | 96 | .40 | 116031.78 | 10132.4574 | 20024.54707 |
| Bawang_Daun | 59 | .36 | 57195.70 | 6946.7187 | 12269.82743 |
| Cabai_Besar | 89 | .01 | 56999.00 | 3064.6967 | 7194.61102 |
| Bayam | 78 | .19 | 13742.95 | 749.1540 | 1945.65288 |
| Kacang_Panjang | 86 | .06 | 13097.50 | 1382.7361 | 2695.66322 |
| Kangkung | 91 | .03 | 21782.37 | 1363.5798 | 2912.33046 |
| Labu_Siam | 54 | .08 | 86907.50 | 5068.3883 | 15158.51157 |
| Sawi | 86 | .20 | 48796.50 | 3941.2937 | 8284.77599 |
| Kelimun | 79 | 1.60 | 23685.55 | 2541.1529 | 4884.81519 |
| Tehung | 89 | .96 | 50912.27 | 2891.6407 | 6609.04167 |
| Tomat | 87 | .46 | 104839.60 | 5387.3298 | 14853.32536 |
| Jamur_Tiram | 72 | .04 | 20963.79 | 657.4476 | 2528.20476 |
| Valid N (listwise) | 32 | | | | |

Pada tabel 4.3. Hasil analisis statistik deskriptif produksi tanaman pangan, terdapat dataset yang tidak memiliki nilai (N/A), diantaranya sebagai berikut :

1. Tanaman Cabai Rawit

Pada produksi tanaman cabai rawit, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Tegal (Jawa Tengah), Kota Cirebon (Jawa Barat).

2. Tanaman Bawang Daun

Pada produksi tanaman bawang daun, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Pacitan (Jawa Timur), Kabupaten Jember (Jawa Timur), Kabupaten Situbondo (Jawa Timur), Kabupaten Sidoarjo (Jawa Timur), Kabupaten Bojonegoro (Jawa Timur), Kabupaten Lamongan (Jawa Timur), Kabupaten Gresik (Jawa Timur), Kabupaten Bangkalan (Jawa Timur), Kabupaten Sampang (Jawa Timur), Kabupaten Sumenep (Jawa Timur), Kota Kediri (Kota), Kota Blitar (Jawa Timur), Kota Malang (Jawa Timur), Kota Probolinggo (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kota Surabaya (Jawa Timur), Kabupaten Kebumen (Jawa Tengah), Kabupaten Purworejo (Jawa Tengah), Kabupaten Sukoharjo (Jawa Tengah), Kabupaten Sragen (Jawa Tengah), Kabupaten Grobongan (Jawa Tengah), Kabupaten Blora (Jawa Tengah), Kabupaten Rembang (Jawa Tengah), Kabupaten Pati (Jawa Tengah), Kabupaten Kudus (Jawa Tengah), Kabupaten Jepara (Jawa Tengah), Kota Magelang (Jawa Tengah), Kota Surakarta (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Cirebon (Jawa Barat), Karawang (Jawa Barat), Kota Bogor (Jawa Barat), Kota Sukabumi (Jawa Barat), Kota Cirebon (Jawa Barat), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat), Kota Cimahi (Jawa Barat).

3. Tanaman Cabai Besar

Pada produksi tanaman cabai besar, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Pacitan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kabupaten Grobongan (Jawa Tengah), Kabupaten Jepara (Jawa Tengah), Kota Magelang (Jawa Tengah), Kota Semarang (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat).

4. Tanaman Bayam

Pada produksi tanaman bayam, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Bondowoso (Jawa Timur), Kabupaten Situbondo (Jawa Timur), Kabupaten Probolinggo (Jawa Timur), Kabupaten Mojokerto (Jawa Timur), Kabupaten Magetan (Jawa Timur), Kabupaten Sampang (Jawa Timur), Kota Kediri (Jawa Timur), Kota Blitar (Jawa Timur), Kota Malang (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kabupaten Banjarnegara (Jawa Tengah), Kabupaten Wonosobo (Jawa Tengah), Kabupaten Sukoharjo (Jawa Tengah), Kabupaten Sragen (Jawa Tengah), Kabupaten Kudus (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kuningan (Jawa Barat), Majalengka (Jawa Barat), Subang (Jawa Barat), Kota Cirebon (Jawa Barat).

5. Tanaman Kacang Panjang

Pada produksi tanaman kacang panjang, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Situbondo (Jawa Timur), Kabupaten Sidoarjo (Jawa Timur), Kabupaten Mojokerto (Jawa Timur), Kota Malang (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kota Surabaya (Jawa Timur), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Kota Cirebon (Jawa Barat), Kota Depok (Jawa Barat), Kota Cimahi (Jawa Barat).

6. Tanaman Kangkung

Pada produksi tanaman kangkung, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Bondowoso (Jawa Timur), Kabupaten Situbondo (Jawa Timur), Kabupaten Probolinggo (Jawa Timur), Kota Kediri (Jawa Timur), Kota Pasuruan (Jawa Timur), Kabupaten Sukoharjo (Jawa Tengah), Kabupaten Karanganyar (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Kota Cirebon (Jawa Barat), Kota Depok (Jawa Barat), Kota Cimahi (Jawa Barat).

7. Tanaman Labu Siam

Pada produksi tanaman labu siam, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Ponorogo (Jawa Timur), Kabupaten Lumajang (Jawa Timur),

Kabupaten Situbondo (Jawa Timur), Kabupaten Sidoarjo (Jawa Timur), Kabupaten Mojokerto (Jawa Timur), Kabupaten Madiun (Jawa Timur), Kabupaten Bojonegoro (Jawa Timur), Kabupaten Tuban (Jawa Timur), Kabupaten Lamongan (Jawa Timur), Kabupaten Gresik (Jawa Timur), Kabupaten Sampang (Jawa Timur), Kabupaten Sumenep (Jawa Timur), Kota Kediri (Jawa Timur), Kota Blitar (Jawa Timur), Kota Malang (Jawa Timur), Kota Probolinggo (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kota Surabaya (Jawa Timur), Kabupaten Banyumas (Jawa Tengah), Kabupaten Kebumen (Jawa Tengah), Kabupaten Purworejo (Jawa Tengah), Kabupaten Sukoharjo (Jawa Tengah), Kabupaten Sragen (Jawa Tengah), Kabupaten Blora (Jawa Tengah), Kabupaten Pati (Jawa Tengah), Kabupaten Jepara (Jawa Tengah), Kabupaten Demak (Jawa Tengah), Kota Magelang (Jawa Tengah), Kota Surakarta (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Kuningan (Jawa Barat), Cirebon (Jawa Barat), Karawang (Jawa Barat), Bekasi (Jawa Barat), Pangandaran (Jawa Barat), Kota Sukabumi (Jawa Barat), Kota Bandung (Jawa Barat), Kota Cirebon (Jawa Barat), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat), Kota Tasikmalaya (Jawa Barat), Kota Banjar (Jawa Barat).

8. Tanaman Sawi

Pada produksi tanaman sawi, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Madiun

(Jawa Timur), Kabupaten Bangkalan (Jawa Timur), Kabupaten Sampang (Jawa Timur), Kota Kediri (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kabupaten Purworejo (Jawa Tengah), Kabupaten Sukoharjo (Jawa Tengah), Kabupaten Sragen (Jawa Tengah), Kabupaten Kudus (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Cirebon (Jawa Barat).

9. Tanaman Ketimun

Pada produksi tanaman ketimun, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Situbondo (Jawa Timur), Kabupaten Probolinggo (Jawa Timur), Kabupaten Mojokerto (Jawa Timur), Kabupaten Nganjuk (Jawa Timur), Kabupaten Gresik (Jawa Timur), Kota Kediri (Jawa Timur), Kota Blitar (Jawa Timur), Kota Malang (Jawa Timur), Kota Probolinggo (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kota Surabaya (Jawa Timur), Kabupaten Sukoharjo (Jawa Tengah), Kota Surakarta (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Tegal (Jawa Tengah), Kota Cirebon (Jawa Barat), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat), Kota Cimahi (Jawa Barat).

10. Tanaman Terung

Pada produksi tanaman terung, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Nganjuk

(Jawa Timur), Kota Malang (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Kuningan (Jawa Barat), Kota Bogor (Jawa Barat), Kota Depok (Jawa Barat).

11. Tanaman Tomat

Pada produksi tanaman tomat, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Nganjuk (Jawa Timur), Kota Kediri (Jawa Timur), Kota Malang (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Madiun (Jawa Timur), Kabupaten Sragen (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Kota Tegal (Jawa Tengah), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat), Kota Cimahi (Jawa Barat).

12. Tanaman Jamur Tiram

Pada produksi tanaman jamur tiram, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah daerah Kabupaten Lumajang (Jawa Timur), Kabupaten Bondowoso (Jawa Timur), Kabupaten Jombang (Jawa Timur), Kabupaten Lamongan (Jawa Timur), Kabupaten Bangkalan (Jawa Timur), Kabupaten Pemekasan (Jawa Timur), Kabupaten Sumenep (Jawa Timur), Kota Kediri (Jawa Timur), Kota Probolinggo (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kabupaten Sukoharjo (Jawa Tengah), Kabupaten Grobongan (Jawa Tengah),

Kabupaten Kudus (Jawa Tengah), Kabupaten Pemalang (Jawa Tengah), Kabupaten Surakarta (Jawa Tengah), Kabupaten Salatiga (Jawa Tengah), Garut (Jawa Barat), Pangandaran (Jawa Barat), Kota Bogor (Jawa Barat), Kota Sukabumi (Jawa Barat), Kota Bandung (Jawa Barat), Kota Cirebon (Jawa Barat), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat), Kota Cimahi (Jawa Barat).

Tabel 4.4. Hasil Analisis Statistik Deskriptif Jumlah Bencana Alam

| | N | Minimum | Maximum | Mean | Std. Deviation |
|--------------------|----|---------|---------|--------|----------------|
| Banjir | 71 | 1.00 | 24.00 | 3.4225 | 3.45858 |
| Karhutla | 62 | 1.00 | 36.00 | 6.1935 | 8.09339 |
| Cuaca Ekstrem | 78 | 1.00 | 145.00 | 7.1282 | 16.78617 |
| Valid N (listwise) | 48 | | | | |

Pada tabel 4.4. Hasil analisis statistik deskriptif jumlah bencana alam, terdapat dataset yang tidak memiliki nilai (N/A), diantaranya sebagai berikut :

1. Jumlah Bencana Alam - Banjir

Pada jumlah bencana alam - banjir, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah Kabupaten Pacitan (Jawa Timur), Kabupaten Ponorogo (Jawa Timur), Kabupaten Trenggalek (Jawa Timur), Kabupaten Tulungagung (Jawa Timur), Kabupaten Blitar (Jawa Timur), Kabupaten Kediri (Jawa Timur), Kabupaten Nganjuk (Jawa Timur), Kabupaten Madiun (Jawa Timur), Kota Probolinggo (Jawa Timur), Kabupaten Magetan (Jawa Timur), Kabupaten Ngawi (Jawa Timur), Kabupaten Tuban (Jawa Timur),

Kabupaten Pamekasan (Jawa Timur), Kabupaten Sumenep (Jawa Timur), Kota Kediri (Jawa Timur), Kota Blitar (Jawa Timur), Kota Malang (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kota Surabaya (Jawa Timur), Kota Batu (Jawa Timur), Kabupaten Kendal (Jawa Tengah), Kota Magelang (Jawa Tengah), Kota Salatiga (Jawa Tengah), Purwakarta (Jawa Barat), Kota Sukabumi (Jawa Barat), Kota Bandung (Jawa Barat), Kota Banjar (Jawa Barat).

2. Jumlah Bencana Alam – Kebakaran Hutan dan Lahan

Pada jumlah bencana alam – kebakaran hutan dan lahan, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah Kabupaten Pacitan (Jawa Timur), Kabupaten Ponorogo (Jawa Timur), Kabupaten Trenggalek (Jawa Timur), Kabupaten Tulungagung (Jawa Timur), Kabupaten Blitar (Jawa Timur), Kabupaten Banyuwangi (Jawa Timur), Kabupaten Nganjuk (Jawa Timur), Kabupaten Madiun (Jawa Timur), Kabupaten Tuban (Jawa Timur), Kabupaten Lamongan (Jawa Timur), Kabupaten Gresik (Jawa Timur), Kabupaten Bangkalan (Jawa Timur), Kabupaten Sampang (Jawa Timur), Kabupaten Pamekasan (Jawa Timur), Kabupaten Sumenep (Jawa Timur), Kota Kediri (Jawa Timur), Kota Blitar (Jawa Timur), Kota Malang (Jawa Timur), Kota Probolinggo (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kota Surabaya (Jawa Timur), Kabupaten Purbalingga (Jawa Tengah), Kabupaten Banjarnegara (Jawa Tengah), Kabupaten Demak (Jawa

Tengah), Kabupaten Brebes (Jawa Tengah), Kota Pekalongan (Jawa Tengah), Tasikmalaya (Jawa Barat), Indramayu (Jawa Barat), Bekasi (Jawa Barat), Kota Bogor (Jawa Barat), Kota Bandung (Jawa Barat), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat), Kota Tasikmalaya (Jawa Barat), Kota Banjar (Jawa Barat).

3. Jumlah Bencana Alam – Cuaca Ekstrem

Pada jumlah bencana alam – kebakaran hutan dan lahan, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah Kabupaten Trenggalek (Jawa Timur), Kabupaten Kediri (Jawa Timur), Kabupaten Lumajang (Jawa Timur), Kabupaten Tuban (Jawa Timur), Kabupaten Bangkalan (Jawa Timur), Kabupaten Sumenep (Jawa Timur), Kota Kediri (Jawa Timur), Kota Blitar (Jawa Timur), Kota Malang (Jawa Timur), Kota Probolinggo (Jawa Timur), Kota Pasuruan (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Madiun (Jawa Timur), Kota Surabaya (Jawa Timur), Kota Batu (Jawa Timur), Kota Magelang (Jawa Tengah), Kota Salatiga (Jawa Tengah), Kota Tegal (Jawa Tengah), Kota Depok (Jawa Barat), Kota Banjar (Jawa Barat).

Tabel 4.5. Hasil Analisis Statistik Deskriptif Rata-rata Pendapatan Bersih

| | N | Minimum | Maximum | Mean | Std. Deviation |
|--------------------|----|-----------|------------|--------------|----------------|
| Upah | 91 | 599972.00 | 3006492.00 | 1299089.6813 | 381380.03517 |
| Valid N (listwise) | 91 | | | | |

Pada tabel 4.5. Hasil analisis statistik deskriptif rata-rata pendapatan bersih

sebulan pekerja informal menurut kabupaten kota dan lapangan pekerjaan utama 2023, terdapat dataset yang tidak memiliki nilai (N/A), diantaranya sebagai berikut :

1. Rata-rata Pendapatan Bersih

Pada rata-rata pendapatan bersih, terdapat daerah kabupaten / kota yang tidak memiliki nilai (N/A), diantara adalah Kota Malang (Jawa Timur), Kota Mojokerto (Jawa Timur), Kota Surabaya (Jawa Timur), Kota Surakarta (Jawa Tengah), Kota Bandung (Jawa Barat), Kota Cirebon (Jawa Barat), Kota Bekasi (Jawa Barat), Kota Depok (Jawa Barat), Kota Cimahi (Jawa Barat).

Tabel 4.6. Hasil Analisis Statistik Deskriptif Perubahan Data Nol

| | N | Minimum | Maximum | Mean | Std. Deviation |
|------------------------|-----|---------|------------|-------------|----------------|
| Luas_Panen_Padi | 100 | 5.82 | 230456.96 | 49245.0082 | 45342.56230 |
| Padi | 100 | 30.51 | 1424303.10 | 279348.0873 | 265543.16031 |
| Luas_Panen_Cabai_Rawit | 100 | .00 | 9881.00 | 1191.1364 | 2139.19998 |
| Cabai_Rawit | 100 | .00 | 116031.78 | 9727.1591 | 19717.08387 |
| Luas_Panen_Bawang_Daun | 100 | .00 | 3628.00 | 332.6849 | 709.30535 |
| Bawang_Daun | 100 | .00 | 57195.70 | 4098.5640 | 9999.57423 |
| Luas_Panen_Cabai_Besar | 100 | .00 | 3383.00 | 237.8810 | 446.67983 |
| Cabai_Besar | 100 | .00 | 56999.00 | 2727.5801 | 6851.26612 |
| Luas_Panen_Bayam | 100 | .00 | 1850.90 | 100.8689 | 252.14129 |
| Bayam | 100 | .00 | 13742.95 | 584.3401 | 1744.02081 |

| | | | | | |
|---------------------------|-----|-----------|------------|--------------|--------------|
| Luas_Panen_Kacang_Panjang | 100 | .00 | 959.59 | 106.0417 | 176.72850 |
| Kacang_Panjang | 100 | .00 | 13097.50 | 1189.1531 | 2543.91928 |
| Luas_Panen_Kangkung | 100 | .00 | 2053.30 | 146.0032 | 295.36729 |
| Kangkung | 100 | .00 | 21782.37 | 1240.8576 | 2804.35854 |
| Luas_Panen_Labu_Siam | 100 | .00 | 649.50 | 34.6558 | 90.19613 |
| Labu_Siam | 100 | .00 | 86907.50 | 2736.9297 | 11378.01749 |
| Luas_Panen_Sawi | 100 | .00 | 2327.00 | 267.9883 | 493.59497 |
| Sawi | 100 | .00 | 48796.50 | 3389.6126 | 7798.74085 |
| Luas_Panen_Ketimun | 100 | .00 | 1151.74 | 120.8881 | 218.17248 |
| Ketimun | 100 | .00 | 23685.55 | 2007.5108 | 4458.92585 |
| Luas_Panen_Terung | 100 | .00 | 1097.00 | 123.9823 | 199.07737 |
| Terung | 100 | .00 | 50912.27 | 2573.5603 | 6297.06516 |
| Luas_Panen_Tomat | 100 | .00 | 3689.00 | 190.7012 | 467.88387 |
| Tomat | 100 | .00 | 104839.60 | 4686.9770 | 13963.04045 |
| Luas_Panen_Jamur_tiram | 100 | .00 | 478629.00 | 17448.2239 | 52107.68426 |
| Jamur_Tiram | 100 | .00 | 20963.79 | 473.3623 | 2161.49283 |
| Banjir | 100 | 1.00 | 24.00 | 3.0100 | 2.97971 |
| Kerhutla | 100 | 1.00 | 36.00 | 4.6000 | 6.67424 |
| Cuaca_Ekstrem | 100 | 1.00 | 145.00 | 6.4400 | 14.86119 |
| Upah | 100 | 599972.00 | 3006492.00 | 1295488.5400 | 363813.66804 |
| Valid N (listwise) | 100 | | | | |

Berdasarkan Tabel 4.6. hasil analisis statistik deskriptif perubahan data nol. Luas panen, produksi komoditas tanaman pangan, jumlah bencana alam dan rata-

rata upah / gaji bersih pekerja informal berdasarkan lapangan pekerjaan utama pertanian di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat untuk dataset yang tidak memiliki nilai (N/A) telah diisi dengan data yang baru, diantaranya dengan menggantikan dengan nilai nol pada luas panen dan produksi dan menggantikan dengan nilai median untuk jumlah bencana alam dan rata-rata upah / gaji bersih. Pada dataset produksi komoditas tanaman pangan yang diganti dengan menggunakan nilai nol dikarenakan pada luas panen suatu komoditas tertentu juga tidak terdapat nilai (N/A) sehingga peneliti memutuskan untuk mengganti data yang tidak memiliki nilai (N/A) tersebut dengan nilai nol, dimana apabila pada luas panen tidak memiliki nilai (N/A) maka untuk produksi komoditas tanaman pangan tersebut apabila juga tidak memiliki nilai (N/A) data tersebut juga akan diisikan dengan nilai nol. Sedangkan pada dataset jumlah bencana alam dan rata-rata upah / gaji bersih, data yang tidak memiliki nilai (N/A) peneliti memutuskan untuk menggantikan dengan nilai median. Peneliti berasumsi bahwa dataset tersebut belum tentu tidak ada tetapi bisa berarti dataset tersebut tidak tersedia atau dataset tersebut belum tercatat. Langkah selanjutnya setelah dataset yang diperoleh dari website Badan Pusat Statistik (BPS) dilakukan pengisian dan penggantian data yang tidak memiliki nilai (N/A) dengan nilai nol dan nilai median, kemudian dilakukan proses standarisasi data menggunakan metode Robust Scaler. Standarisasi data pada penelitian ini menggunakan metode Robust Scaler, peneliti menggunakan metode Robust scaler karena dataset yang diperoleh memiliki perbedaan skala yang sangat luas disetiap variabelnya. Standarisasi data dengan menggunakan metode Robust Scaler dilakukan agar

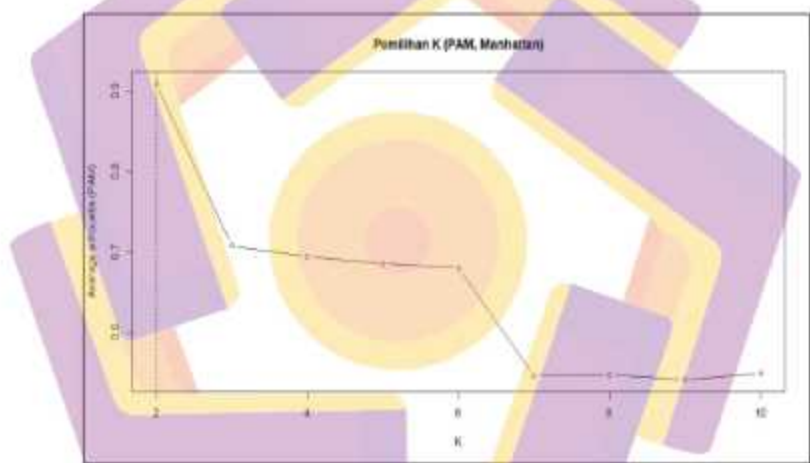
perbedaan skala tersebut tidak mengganggu pada proses cluster.

4.3 Partitioning Around Medoids (PAM)

Pada proses pembentukan cluster menggunakan algoritma Partitioning Around Medoids (PAM), setelah dataset dilakukan pengisian data yang tidak memiliki nilai (N/A). Langkah selanjutnya adalah dengan melakukan proses standarisasi data dengan menggunakan metode robust scaler. Proses standarisasi data dengan menggunakan metode robust scaler ini mengubah skala setiap variabel sehingga sebagian besar nilai berada dalam rentang sekitar -1 hingga 1, serta mengurangi pengaruh nilai ekstrem (outlier) yang dapat mendistorsi hasil analisis. Standarisasi ini diperlukan karena variabel dalam dataset memiliki rentang nilai yang sangat berbeda, sehingga penting untuk memastikan bahwa tidak ada variabel yang mendominasi perhitungan jarak pada proses clustering.

Dataset yang telah dilakukan standarisasi data menggunakan metode robust scaler, selanjutnya dilakukan tahap perhitungan jarak menggunakan metode manhattan distance. Metode perhitungan jarak menggunakan manhattan distance dipilih karena dataset yang dimiliki terdapat banyak nilai nol dan berpotensi outlier. Perhitungan jarak menggunakan metode manhattan distance akan menghasilkan matrik jarak yang nantinya akan digunakan sebagai acuan untuk proses cluster menggunakan algoritma Partitioning Around Medoids (PAM). Setelah melakukan perhitungan jarak dan clustering menggunakan metode algoritma Partitioning Around Medoids (PAM), langkah selanjutnya adalah dengan melakukan validasi hasil clustering menggunakan Silhouette Score. Perhitungan silhouette score untuk menentukan jumlah cluster (K) yang optimal.

Silhouette score yang lebih tinggi dan yang mendekati 1 menunjukkan bahwa objek berada lebih dekat dengan cluster yang menjadi anggotanya daripada dengan cluster lainnya, sehingga struktur cluster semakin baik dan pemisahan antar cluster menjadi semakin jelas. Oleh karena itu, jumlah cluster yang dipilih berdasarkan Silhouette Score yang tertinggi. Pada gambar 4.1. Grafik Silhouette Score metode PAM, diperoleh nilai $k = 2$, jumlah cluster optimal yang akan digunakan pada penelitian ini adalah $k = 2$.



Gambar 4.1. Grafik Silhouette Score Metode PAM

Pada Tabel 4.7. Hasil Silhouette Score Metode PAM, menunjukkan nilai silhouette score yang dilakukan untuk menentukan nilai k optimal. Percobaan yang dilakukan menggunakan nilai $k = 2$ sampai dengan $k = 10$. Pada percobaan tersebut didapatkan nilai silhouette tertinggi adalah $k = 2$ dengan silhouette score 0.9083. Dengan demikian, pada penelitian ini $k = 2$ ditetapkan sebagai jumlah cluster optimal karena memiliki kualitas pemisahan cluster yang paling baik

dibandingkan dengan nilai k lainnya.

Tabel 4.7. Hasil Silhouette Score Metode PAM

| No | Jumlah K | Silhouette Score |
|----|----------|------------------|
| 1 | 2 | 0.9083 |
| 2 | 3 | 0.7075 |
| 3 | 4 | 0.6943 |
| 4 | 5 | 0.6855 |
| 5 | 6 | 0.6799 |
| 6 | 7 | 0.5459 |
| 7 | 8 | 0.5475 |
| 8 | 9 | 0.5413 |
| 9 | 10 | 0.5494 |

Langkah selanjutnya adalah melakukan visualisasi cluster berdasarkan jumlah cluster (k) optimal yang diperoleh dari analisis Silhouette Score. Pada penelitian ini, nilai silhouette tertinggi sebesar 0,9083 dan diperoleh pada $k = 2$, sehingga jumlah cluster (k) optimal yang akan digunakan adalah dua cluster. Visualisasi hasil clustering menggunakan Principal Component Analysis (PCA). Principal Component Analysis (PCA) merupakan metode yang digunakan untuk mereduksi dimensi dengan mengubah sekumpulan variabel yang saling berkorelasi menjadi beberapa komponen utama atau principal components (PC) yang tidak berkorelasi, dengan masih mempertahankan sebagian besar informasi dalam data. Pada penelitian ini, hasil PCA menunjukkan bahwa komponen utama

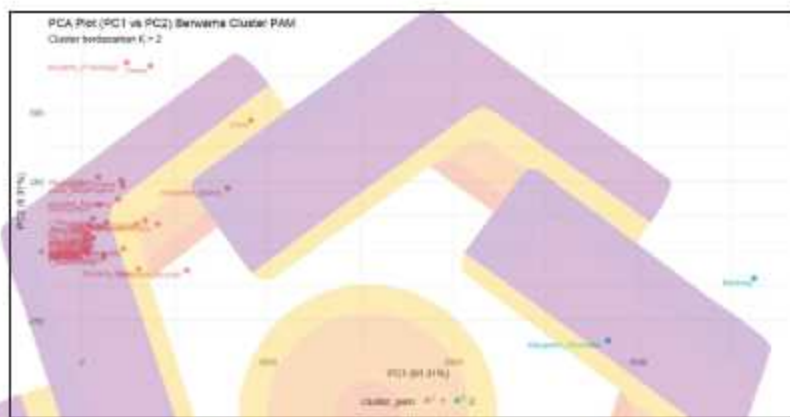
pertama (PC1) dan komponen utama kedua (PC2) mampu menjelaskan sekitar 97% variasi data, sehingga dua komponen utama tersebut sudah cukup untuk mewakili struktur data seluruh variabel (sekitar 30 variabel produksi hortikultura). Komponen utama berikutnya (PC3, PC4, dan seterusnya) tidak ditampilkan pada penelitian ini, karena kontribusi variansnya sangat kecil (sekitar kurang dari 2%) dan tidak memberikan informasi yang signifikan terhadap pola cluster.

Tabel 4.8. Nilai Principal Components (PC)

| No | Komponen | Proporsi Varians |
|----|------------|------------------|
| 1 | PC1 | 91.31 % |
| 2 | PC2 | 6.31 % |
| 3 | PC3 – PC29 | < 2% per PC |

Pada Tabel 4.8. Nilai Principal Components (PC), menunjukkan analisis komponen utama (PCA) yang diterapkan untuk mereduksi 29 variabel hortikultura menjadi beberapa komponen utama yang lebih ringkas. Hasil analisa PCA menunjukkan bahwa komponen utama pertama (PC1) menjelaskan 91,31% varians, sedangkan PC2 menjelaskan 6,31% varians. Dengan demikian, komponen utama pertama (PC1) dan komponen utama kedua (PC2) secara kumulatif menjelaskan 97,62 % variasi total data, sehingga kedua komponen ini sudah mampu mewakili informasi utama dalam dataset. Komponen utama berikutnya (PC3-PC29) masing-masing hanya menjelaskan kurang dari 2 % variasi, sehingga tidak memberikan informasi tambahan yang signifikan dan tidak digunakan dalam visualisasi lebih lanjut.

Langkah selanjutnya adalah membuat visualisasi cluster berdasarkan dari hasil analisa silhouette score. Pada hasil analisa silhouette score, jumlah cluster (k) optimal yang dihasilkan berdasarkan silhouette score tertinggi. Pada penelitian ini k optimal yang akan digunakan adalah $k = 2$ dengan silhouette score $= 0.9083$.



Gambar 4.2. Visualisasi Clustering Metode PAM

Pada Gambar 4.2. Visualisasi clustering metode PAM, menunjukkan visualisasi hasil clustering menggunakan algoritma Partitioning Around Medoids (PAM). Pada visualisasi hasil clustering tersebut, cluster 1 ditunjukkan dengan warna merah, sedangkan cluster 2 ditunjukkan dengan warna biru. Titik-titik data diplot menggunakan dua komponen utama pertama dan kedua (PC1 dan PC2) dari Principal Component Analysis (PCA), yang bersama-sama mampu menjelaskan proporsi variasi terbesar pada dataset. Berdasarkan gambar 4.2., terlihat bahwa kedua cluster terpisah dengan jelas, sehingga struktur cluster yang terbentuk dapat dikatakan baik dan mendukung hasil evaluasi silhouette score sebelumnya.

Pada Tabel 4.8. Nilai PC1 dan PC2, menunjukkan nilai proportion of

variance dari Principal Component Analysis (PCA). Komponen utama pertama (PC1) mampu menjelaskan sebesar 91,31% variasi total data, sedangkan komponen utama kedua (PC2) menjelaskan sebesar 6,31%. Dengan demikian, kedua komponen utama tersebut secara kumulatif menjelaskan 97,62 % variasi dalam data. Sehingga, komponen utama (PC1) dan komponen utama kedua (PC2) sudah dapat mewakili pola dominan pada dataset hortikultura dengan sangat baik.

Tabel 4.9. Loading Komponen Utama Pertama (PC1)

| No | Komponen | Loading PC1 | Keterangan |
|----|------------------------|---------------|------------------|
| 1 | Labu siam | 0.9487 | dominan |
| 2 | Daun bawang | 0.2967 | kontribusi kuat |
| 3 | Luas panen bawang daun | 0.0716 | kontribusi kecil |
| 4 | Luas panen labu siam | 0.0588 | kontribusi kecil |
| 5 | Tomat | 0.0349 | kecil |

Pada tabel 4.9. Loading Komponen Utama Pertama (PC1), menunjukkan tingkat intensitas produksi hortikultura, terutama pada tanaman labu siam dan tanaman bawang daun. Kabupaten / kota dengan nilai PC1 tinggi merupakan wilayah dengan skala produksi hortikultura besar.

Tabel 4.10. Loading Komponen Utama Kedua (PC2)

| No | Komponen | Loading PC2 | Keterangan |
|----|------------------------|----------------|-------------------------|
| 1 | Bawang daun | 0.9087 | sangat dominan |
| 2 | Labu siam | -0.3134 | Dominan berlawanan arah |
| 3 | Luas panen bawang daun | 0.2173 | kontribusi sedang |

| | | | |
|---|----------------------|--------|------------------|
| 4 | Tomat | 0.1073 | kontribusi kecil |
| 5 | Luas panen labu siam | 0.0554 | kontribusi kecil |

Pada tabel 4.10. Loading Komponen Utama Kedua (PC2), menunjukkan perbedaan dominasi komoditas. Produksi tanaman bawang daun mendominasi arah positif, sedangkan produksi tanaman labu siam mendominasi arah negatif. Yang artinya, daerah dengan nilai PC2 tinggi adalah daerah yang lebih fokus pada bawang daun, sedangkan nilai PC2 rendah cenderung berdominasi labu siam

Berdasarkan Tabel 4.9. dan 4.10., Principal Component Analysis (PCA) menghasilkan dua komponen utama dengan kontribusi kumulatif sebesar 97,62%, sehingga komponen utama pertama (PC1) dan komponen utama kedua (PC2) sudah mewakili sebagian besar informasi dalam dataset hortikultura.

Komponen utama pertama (PC1) memiliki nilai loading terbesar pada variabel labu siam, bawang daun, tomat, jamur tiram, dan cabai besar. Seluruh nilai loading PC1 bernilai positif, sehingga komponen utama pertama (PC1) dapat diinterpretasikan sebagai dimensi tingkat intensitas produksi hortikultura. Kabupaten / kota dengan nilai PC1 tinggi memiliki skala produksi sayuran yang besar, sedangkan nilai PC1 rendah mengindikasikan produksi yang relatif rendah.

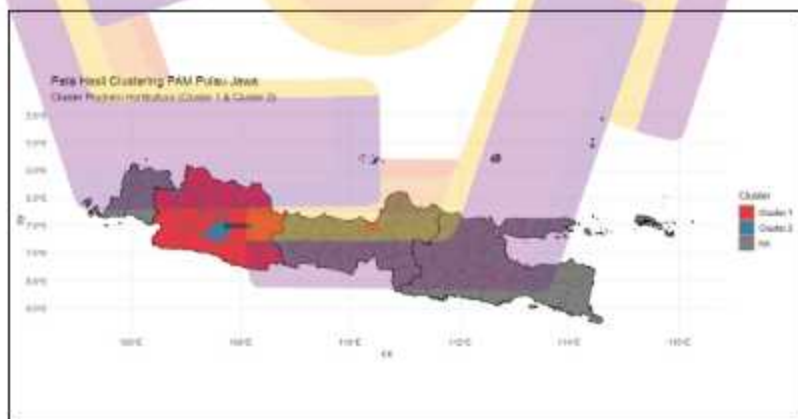
Komponen utama kedua (PC2) menekankan pada variabel yang menunjukkan perbedaan pola dominasi komoditas. Variabel bawang daun memiliki nilai loading positif sangat besar, sedangkan pada labu siam bernilai negatif cukup kuat. Dengan demikian, Komponen utama kedua (PC2) dapat mengindikasikan karakteristik spesifik komoditas. Nilai PC2 tinggi

mengindikasikan wilayah yang dominan bawang daun dan nilai PC2 rendah mengindikasikan wilayah yang dominan labu siam.

Pada cluster 1 terdiri dari 98 kabupaten / kota, sedangkan pada cluster 2 terdiri dari 2 kabupaten / kota. Cluster 1 terdiri dari daerah, Kabupaten Pacitan, Kabupaten Ponorogo, Kabupaten Trenggalek, Kabupaten Tulungagung, Kabupaten Blitar, Kabupaten Kediri, Kabupaten Malang, Kabupaten Lumajang, Kabupaten Jember, Kabupaten Banyuwangi, Kabupaten Bondowoso, Kabupaten Situbondo, Kabupaten Probolinggo, Kabupaten Pasuruan, Kabupaten Sidoarjo, Kabupaten Mojokerto, Kabupaten Jombang, Kabupaten Nganjuk, Kabupaten Madiun, Kabupaten Magetan, Kabupaten Ngawi, Kabupaten Bojonegoro, Kabupaten Tuban, Kabupaten Lamongan, Kabupaten Gresik, Kabupaten Bangkalan, Kabupaten Sampang, Kabupaten Pamekasan, Kabupaten Sumenep, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, Kota Batu, Kabupaten Cilacap, Kabupaten Banyumas, Kabupaten Purbalingga, Kabupaten Banjarnegara, Kabupaten Kebumen, Kabupaten Purworejo, Kabupaten Magelang, Kabupaten Boyolali, Kabupaten Klaten, Kabupaten Sukoharjo, Kabupaten Wonogiri, Kabupaten Karanganyar, Kabupaten Sragen, Kabupaten Grobogan, Kabupaten Blora, Kabupaten Rembang, Kabupaten Pati, Kabupaten Kudus, Kabupaten Jepara, Kabupaten Demak, Kabupaten Semarang, Kabupaten Temanggung, Kabupaten Kendal, Kabupaten Batang, Kabupaten Pekalongan, Kabupaten Pemalang, Kabupaten Tegal, Kabupaten Brebes, Kota Magelang, Kota Surakarta, Kota Salatiga, Kota Semarang, Kota Pekalongan, Kota Tegal, Bogor, Sukabumi,

Cianjur, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Subang, Purwakarta, Karawang, Bekasi, Bandung Barat, Pangandaran, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, dan Kota Banjar. Pada Cluster 2 terdiri dari 2 kabupaten / kota, Kabupaten Wonosobo dan Bandung.

Langkah selanjutnya adalah dengan menampilkan peta cluster menggunakan metode Partitioning Around Medoids (PAM). Pada Gambar 4.3. menunjukkan peta sebaran hasil clustering menggunakan metode Partitioning Around Medoids (PAM) pada wilayah Pulau Jawa. Wilayah yang termasuk cluster 1 ditampilkan dengan warna merah, sedangkan wilayah yang termasuk cluster 2 ditampilkan dengan warna biru. Daerah lain yang berwarna abu-abu merupakan kabupaten / kota yang tidak termasuk dalam dataset penelitian sehingga tidak diklaster.



Gambar 4.3. Peta Cluster metode Partitioning Around Medoids (PAM)

Pada peta cluster dengan menggunakan metode partitioning around

medoids (PAM) terlihat bahwa sebagian besar wilayah Pulau Jawa berada pada Cluster 1 (ditunjukkan dengan warna merah), sedangkan hanya Kabupaten Wonosobo dan Kota Bandung yang termasuk dalam Cluster 2 (ditunjukkan dengan warna biru) sebagai kelompok wilayah dengan karakteristik produksi hortikultura yang berbeda secara signifikan dibandingkan wilayah lainnya.

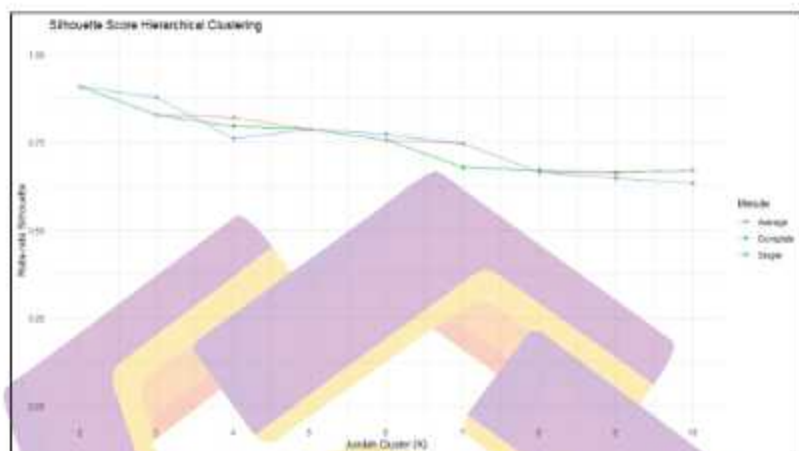
4.4 Hierarchical Clustering dengan Metode Agglomerative

Pada penelitian ini, selain menggunakan metode partitioning around medoids (PAM), peneliti juga akan menggunakan pembentukan cluster dengan menggunakan Cluster Hierarchical dengan Metode Agglomerative. Pada metode agglomerative ini akan menggunakan 3 metode linkage, antara lain : metode single linkage, average linkage dan complete linkage. Penelitian ini akan membandingkan juga 3 metode single linkage, average linkage dan complete linkage yang menghasilkan cluster terbaik. Setelah dataset dilakukan pengisian data yang tidak memiliki nilai (N/A) dengan nilai nol dan nilai median. Langkah selanjutnya adalah dengan melakukan proses standarisasi data dengan menggunakan metode robust scaler. Proses standarisasi data dengan menggunakan metode robust scaler membuat sebagian besar nilai berada diantara -1 dan 1 dan nilai-nilai yang ekstrem menjadi tidak terlalu besar. Proses standarisasi ini dilakukan agar dataset yang diperoleh tidak memiliki rentang jarak yang sangat lebar.

Langkah selanjutnya adalah melakukan perhitungan jarak, perhitungan jarak dilakukan dengan menggunakan metode manhattan distance. Metode manhattan distance ini dipilih karena dataset yang dimiliki sebelumnya banyak

memiliki data nilai nol atau outlier. Perhitungan jarak menggunakan metode manhattan distance akan menghasilkan matrik jarak yang akan digunakan sebagai acuan perhitungan pada metode linkage untuk mengelompokkan kabupaten / kota. Pada penelitian ini, metode linkage yang digunakan meliputi, metode single linkage, average linkage dan complete linkage. Hasil perhitungan dengan menggunakan metode linkage (single linkage, average linkage dan complete linkage) akan divisualisasikan dengan dendrogram. Dendrogram yang terbentuk menggunakan metode single linkage ditunjukkan pada gambar 4.4., metode single linkage menggunakan jarak minimum antar titik dari dua cluster. Dendrogram yang terbentuk menggunakan metode average linkage ditunjukkan pada gambar 4.5., metode average linkage menghitung jarak antar cluster sebagai rata-rata seluruh jarak antar titik kedua cluster. Dendrogram yang terbentuk menggunakan metode complete linkage ditunjukkan pada gambar 4.6., metode complete linkage menggunakan jarak maksimum antar titik dari dua cluster.

Langkah selanjutnya setelah membuat pengelompokkan kabupaten / kota menggunakan metode linkage (single linkage, average linkage dan complete linkage) adalah menentukan jumlah k optimal. Pada penelitian ini penentuan jumlah k optimal akan dilakukan dengan menggunakan Silhouette Score.



Gambar 4.7. Silhouette Score Metode Hierarchical Clustering

Pada Gambar 4.7. Silhouette Score Metode Hierarchical Clustering, menunjukkan nilai k optimal setiap metode linkage (single linkage, average linkage dan complete linkage). Metode linkage yang digunakan pada penelitian ini adalah metode single linkage, metode average linkage dan metode complete linkage.

Tabel 4.11. Silhouette Score Penentuan K Optimal

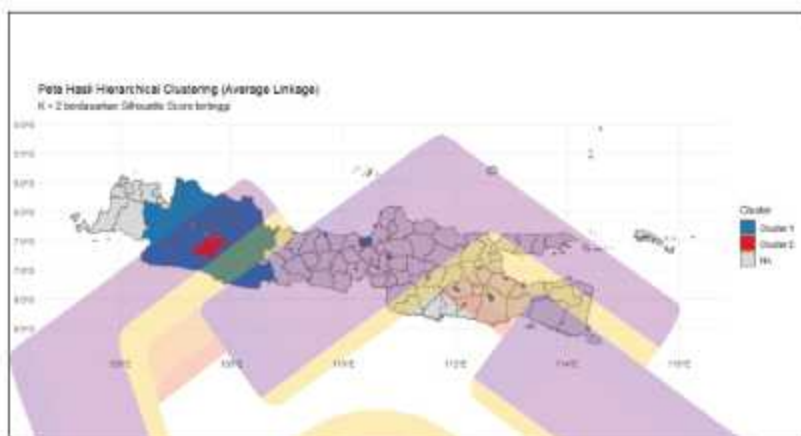
| No | K | Single Linkage | Average Linkage | Complete Linkage |
|----|---|------------------|------------------|------------------|
| 1 | 2 | 0.9083445 | 0.9083445 | 0.9083445 |
| 2 | 3 | 0.8779102 | 0.8281155 | 0.8281155 |
| 3 | 4 | 0.7610469 | 0.7953271 | 0.8193402 |
| 4 | 5 | 0.7865518 | 0.7865518 | 0.7865518 |
| 5 | 6 | 0.7721252 | 0.7559051 | 0.7559051 |

| | | | | |
|---|----|-----------|-----------|-----------|
| 6 | 7 | 0.7472882 | 0.6785319 | 0.7472882 |
| 7 | 8 | 0.6635205 | 0.6704027 | 0.6635205 |
| 8 | 9 | 0.6484637 | 0.6640308 | 0.6631242 |
| 9 | 10 | 0.6344382 | 0.6701489 | 0.6701489 |

Pada Tabel 4.11. Silhouette Score Penentuan K Optimal, menunjukkan nilai silhouette tiap metode linkage (single linkage, average linkage dan complete linkage). Pada metode single linkage menghasilkan silhouette score = 0.9083445 (mendekati nilai 1) pada nilai $k = 2$. Metode average linkage menghasilkan nilai silhouette score = 0.9083445 (mendekati nilai 1) pada nilai $k = 2$ dan Metode complete linkage menghasilkan nilai silhouette score = 0.9083445 (mendekati nilai 1) pada nilai $k = 2$. Penentuan jumlah cluster (k) optimal dilakukan menggunakan silhouette score untuk $K = 2$ hingga $K = 10$. Pada Gambar 4.7 menunjukkan bahwa seluruh metode linkage (single linkage, complete linkage dan average linkage) menghasilkan nilai Silhouette tertinggi pada $K = 2$ (0,9083). Setelah $K > 2$, nilai Silhouette cenderung menurun yang mengindikasikan penurunan kualitas pemisahan cluster. Berdasarkan evaluasi silhouette score, ketiga metode linkage (single linkage, complete linkage, dan average linkage) menghasilkan nilai yang sama pada jumlah cluster optimal $K = 2$. Pemilihan metode linkage selanjutnya dilakukan dengan mempertimbangkan karakteristik metode. Metode single linkage cenderung menghasilkan efek rantai (chaining), sementara metode complete linkage terlalu konservatif dalam memisahkan cluster. Oleh karena itu, metode average linkage dipilih sebagai metode hierarchical clustering yang

digunakan dalam penelitian ini karena memberikan kompromi terbaik antara stabilitas dan kedekatan antar objek dalam cluster. Selanjutnya hasil perhitungan metode linkage (single linkage, average linkage dan complete linkage) menggunakan silhouette score divisualisasikan menggunakan dendrogram. Pada gambar 4.8. menunjukkan visualisasi dalam bentuk dendrogram menggunakan metode single linkage dengan menggunakan hasil perhitungan silhouette score atau dengan k optimal $k = 2$. Pada gambar 4.9. menunjukkan visualisasi dalam bentuk dendrogram menggunakan metode average linkage dengan menggunakan hasil perhitungan silhouette score atau dengan k optimal $k = 2$. Pada gambar 4.10. menunjukkan visualisasi dalam bentuk dendrogram menggunakan metode complete linkage dengan menggunakan hasil perhitungan silhouette score atau dengan k optimal $k = 2$.

Langkah selanjutnya adalah dengan menampilkan peta cluster menggunakan metode hierarchical clustering. Pada Gambar 4.11. menunjukkan peta sebaran kabupaten / kota hasil clustering menggunakan metode hierarchical clustering pada wilayah Pulau Jawa. Wilayah yang termasuk cluster 1 ditampilkan dengan warna biru, sedangkan wilayah yang termasuk cluster 2 ditampilkan dengan warna merah. Daerah lain yang berwarna abu-abu merupakan kabupaten / kota yang tidak termasuk dalam dataset penelitian sehingga tidak dikluster.



Gambar 4.11. Peta Cluster metode Hierarchical Clustering

Pada cluster 1 terdiri dari 98 kabupaten / kota, sedangkan pada cluster 2 terdiri dari 2 kabupaten / kota. Cluster 1 terdiri dari daerah, Kabupaten Pacitan, Kabupaten Ponorogo, Kabupaten Trenggalek, Kabupaten Tulungagung, Kabupaten Blitar, Kabupaten Kediri, Kabupaten Malang, Kabupaten Lumajang, Kabupaten Jember, Kabupaten Banyuwangi, Kabupaten Bondowoso, Kabupaten Situbondo, Kabupaten Probolinggo, Kabupaten Pasuruan, Kabupaten Sidoarjo, Kabupaten Mojokerto, Kabupaten Jombang, Kabupaten Nganjuk, Kabupaten Madiun, Kabupaten Magetan, Kabupaten Ngawi, Kabupaten Bojonegoro, Kabupaten Tuban, Kabupaten Lamongan, Kabupaten Gresik, Kabupaten Bangkalan, Kabupaten Sampang, Kabupaten Pamekasan, Kabupaten Sumenep, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, Kota Batu, Kabupaten Cilacap,

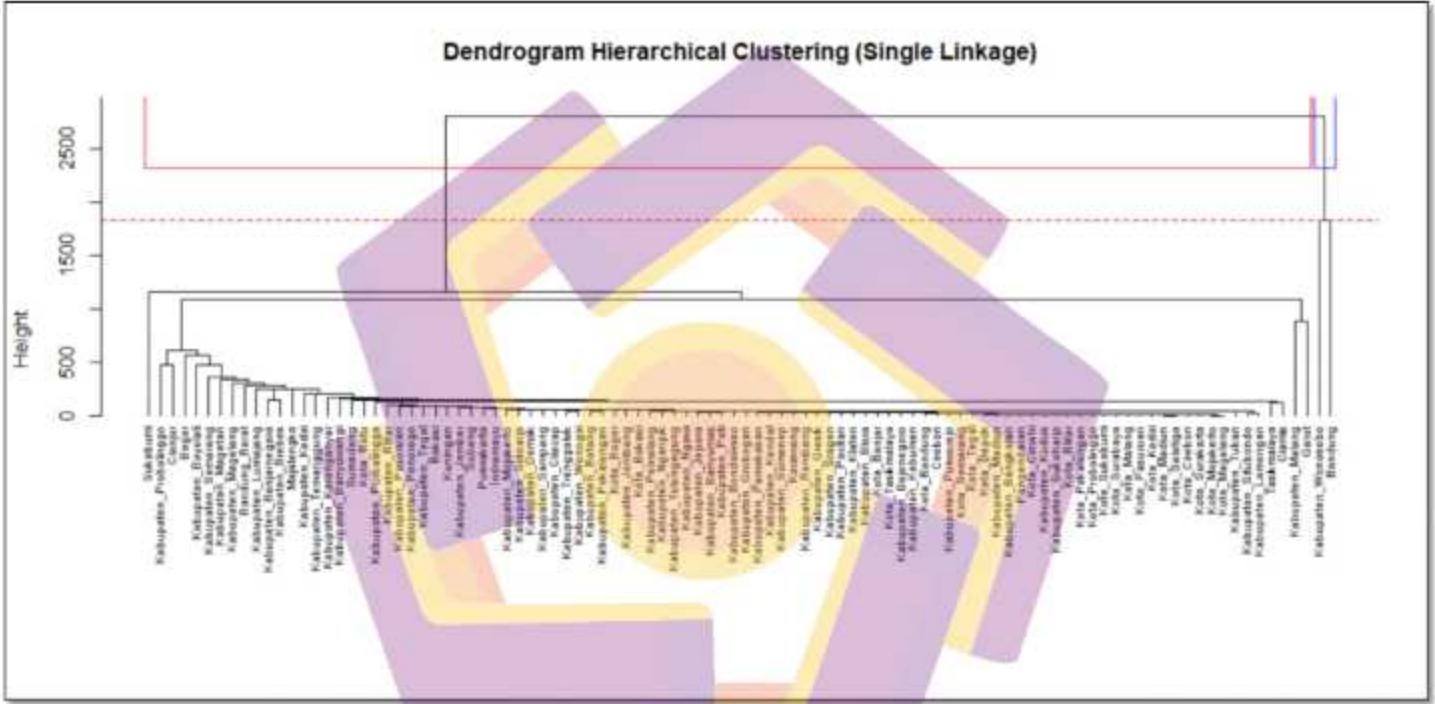
Kabupaten Banyumas, Kabupaten Purbalingga, Kabupaten Banjarnegara, Kabupaten Kebumen, Kabupaten Purworejo, Kabupaten Magelang, Kabupaten Boyolali, Kabupaten Klaten, Kabupaten Sukoharjo, Kabupaten Wonogiri, Kabupaten Karanganyar, Kabupaten Sragen, Kabupaten Grobogan, Kabupaten Blora, Kabupaten Rembang, Kabupaten Pati, Kabupaten Kudus, Kabupaten Jepara, Kabupaten Demak, Kabupaten Semarang, Kabupaten Temanggung, Kabupaten Kendal, Kabupaten Batang, Kabupaten Pekalongan, Kabupaten Pemalang, Kabupaten Tegal, Kabupaten Brebes, Kota Magelang, Kota Surakarta, Kota Salatiga, Kota Semarang, Kota Pekalongan, Kota Tegal, Bogor, Sukabumi, Cianjur, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Subang, Purwakarta, Karawang, Bekasi, Bandung Barat, Pangandaran, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, dan Kota Banjar. Pada Cluster 2 terdiri dari 2 kabupaten / kota, yaitu Kabupaten Wonosobo dan Bandung

Tabel 4.12. Profiling Hierarchical Clustering (Average Linkage, K – 2)

| Variabel | Cluster 1 | Cluster 2 | Keterangan |
|------------------------|-----------|-------------|-----------------------------|
| Labu siam | 11.4 | 78.436 | Cluster 2 sangat dominan |
| Bawang daun | 25.5 | 44.765 | Cluster 2 sangat dominan |
| Tomat | 281 | 39.102 | Cluster 2 jauh lebih tinggi |
| Luas panen labu siam | rendah | sangat luas | Intensitas lahan Cluster 2 |
| Luas panen bawang daun | rendah | sangat luas | Spesialisasi Cluster 2 |

Pada tabel 4.12. Profiling Hierarchical Clustering (Average Linkage, K –

2), menunjukkan profiling cluster yang dilakukan menggunakan nilai median. Pemilihan nilai median didasarkan pada karakteristik dataset yang bersifat tidak simetris dan memiliki banyak nilai nol, serta mengandung nilai ekstrem pada beberapa wilayah sentra produksi. Penggunaan nilai median dinilai lebih mewakili dalam menggambarkan kondisi tipikal anggota cluster dibandingkan dengan nilai rata-rata yang sensitif terhadap outlier. Dengan demikian, nilai median memberikan gambaran yang lebih stabil terhadap variasi nilai yang ekstrem antar wilayah. Berdasarkan hasil hierarchical clustering dengan metode linkage (average linkage) dan jumlah cluster (k) optimal $K = 2$, diperoleh karakteristik cluster sebagai berikut : Pada cluster 1, memiliki ciri dengan menggunakan nilai median produksi dan luas panen yang relatif rendah pada seluruh variabel utama, termasuk produksi labu siam, bawang daun, dan tomat. ciri ini dapat menunjukkan bahwa wilayah dalam cluster 1 ini belum berperan sebagai wilayah sentra utama produksi hortikultura, melainkan lebih bersifat pendukung dengan skala produksi yang terbatas. Pada cluster 1 dapat juga disebut sebagai wilayah produksi hortikultura skala rendah-menengah. Pada Cluster 2 menunjukkan bahwa nilai median yang dihasilkan jauh lebih tinggi pada variabel produksi labu siam, bawang daun, dan tomat, serta luas panen labu siam dan bawang daun. Oleh karena itu, ciri tersebut dapat mengindikasikan bahwa wilayah dalam cluster ini merupakan sentra produksi hortikultura dengan intensitas produksi dan pemanfaatan lahan yang tinggi. Sehingga pada cluster 2 ini merupakan wilayah dengan sentra produksi hortikultura skala besar.



Gambar 4.8. Dendrogram Cluster Metode Single Linkage, K-2

4.5 Analisa Cluster

Pada penelitian ini, penerapan metode partitioning around medoids (PAM) dan hierarchical clustering untuk mengelompokkan wilayah kabupaten / kota di Provinsi Jawa Timur, Jawa Tengah dan Jawa Barat. Pengelompokkan kabupaten / kota di Pulau Jawa tersebut berdasarkan karakteristik produksi hortikultura. Kedua metode (partitioning around medoids (PAM) dan hierarchical clustering) menggunakan dataset yang sama, jarak Manhattan, serta jumlah cluster (k) optimal yang ditentukan menggunakan silhouette score. Hasil evaluasi menggunakan silhouette score menunjukkan bahwa metode partitioning around medoids (PAM) maupun hierarchical clustering menghasilkan nilai silhouette tertinggi pada $K = 2$. Hal ini menunjukkan bahwa struktur dataset secara alami membentuk dua kelompok utama yang memiliki tingkat kesamaan cluster tinggi dan memiliki perbedaan antar-cluster yang jelas. Baik pada metode partitioning around medoids (PAM) maupun pada hierarchical clustering, dua cluster yang dihasilkan memiliki karakteristik yang sama, yaitu : Pada cluster 1, Merupakan wilayah yang memiliki skala produksi hortikultura yang relatif rendah hingga menengah, ditandai dengan nilai median produksi dan luas panen yang lebih kecil pada sebagian besar komoditas utama, seperti labu siam, bawang daun, dan tomat. Pada cluster 1 ini umumnya bukan merupakan sentra produksi utama, melainkan wilayah dengan produksi hortikultura yang lebih terbatas. Pada Cluster 2, Merupakan wilayah dengan tingkat produksi hortikultura yang tinggi, ditunjukkan dengan nilai median produksi dan luas panen yang jauh lebih besar, khususnya pada komoditas labu siam dan bawang daun. Cluster ini mewakili wilayah sentra

produksi hortikultura utama dengan spesifik komoditas dan pemanfaatan lahan yang tinggi. Kesamaan karakteristik kedua metode ini (metode partitioning around medoids (PAM) maupun hierarchical clustering), ini diperkuat oleh peta cluster yang dihasilkan, di mana pola sebaran cluster pada metode partitioning around medoids (PAM) dan hierarchical clustering menunjukkan distribusi wilayah yang sama.

4.6 Diskusi Teoritik

Penelitian ini menggunakan metode Partitioning Around Medoids (PAM) dan hierarchical clustering, yang secara umum sejalan dengan metode yang digunakan pada penelitian terdahulu, namun dengan beberapa perbedaan pendekatan. Penelitian Mapping the Diversity of Regional Characteristics dan Clustering and Mapping of Agricultural Production Based on GIS Using K-Medoids Algorithm, sama-sama menggunakan pendekatan clustering berbasis wilayah untuk memetakan tipologi daerah. Kesamaan dengan penelitian ini terletak pada penggunaan algoritma berbasis medoid (PAM/K-Medoids) yang relatif lebih robust terhadap outlier dibandingkan K-Means. Pada penelitian ini, menggabungkan PAM dan hierarchical clustering secara simultan, sehingga memungkinkan evaluasi konsistensi struktur cluster antar metode.

Sementara itu, pada penelitian Credit Customer Segmentation with Hierarchical Clustering menitikberatkan pada hierarchical clustering dengan berbagai jarak dan linkage, mirip dengan penelitian ini dari sisi eksplorasi jarak Manhattan. Namun, objek dan domain penelitiannya berbeda (nasabah perbankan), sedangkan penelitian ini berfokus pada wilayah produksi komoditas tanaman

pangan/hortikultura.

Berdasarkan penelitian-penelitian terdahulu, clustering terbukti efektif untuk membentuk tipologi wilayah pada konteks produksi pertanian, ketahanan pangan, maupun kerentanan berbasis lokasi. Studi pengelompokan produksi padi menegaskan bahwa indikator produksi mampu membedakan karakteristik kabupaten / kota. Pada metode clustering menegaskan pentingnya membandingkan metode untuk memperoleh struktur cluster yang paling mewakili, sedangkan pada clustering yang berbasis variabel menekankan bahwa data wilayah yang multidimensi sering memiliki korelasi tinggi sehingga memerlukan peringkasan dimensi untuk interpretasi yang lebih jelas. Oleh karena itu, penelitian ini memadukan metode *partitioning around medoids* (PAM) dan *hierarchical clustering* dengan validasi *silhouette* serta PCA untuk memudahkan interpretasi pola dominan, sehingga hasil pengelompokan wilayah dapat digunakan sebagai dasar pengambilan keputusan strategis dalam pengembangan pertanian berbasis wilayah.

Secara umum, hasil penelitian ini sejalan dengan penelitian-penelitian terdahulu yang menyatakan bahwa metode clustering efektif untuk mengidentifikasi struktur dan tipologi wilayah berdasarkan karakteristik produksi dan indikator multidimensi.

BAB 5

PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian tentang penerapan algoritma partitioning around medoids (PAM) dan hierarchical clustering dalam pengelompokan daerah-daerah atau wilayah di Indonesia yang memiliki produksi komoditas tanaman pangan, dapat diambil kesimpulan sebagai berikut :

1. Pada penelitian ini tentang penerapan algoritma partitioning around medoids (PAM) dan hierarchical clustering dalam pengelompokan daerah-daerah atau wilayah di Provinsi Jawa yang memiliki produksi komoditas tanaman pangan. Baik metode partitioning around medoids (PAM) dan hierarchical clustering sama-sama menghasilkan silhouette score sebesar 0.9083 dan jumlah cluster (k) optimal $k = 2$. Pada hierarchical clustering, metode linkage yang dipilih adalah metode average.
2. Pada penelitian menggunakan partitioning around medoids (PAM), hasil analisis Principal Component Analysis (PCA) menunjukkan di mana PC1 mewakili dimensi utama tingkat produksi hortikultura, sedangkan PC2 menggambarkan variasi dominasi jenis komoditas yang dihasilkan antar wilayah. Pada visualisasi PCA menunjukkan pemisahan antar cluster yang jelas, sehingga dapat menandakan

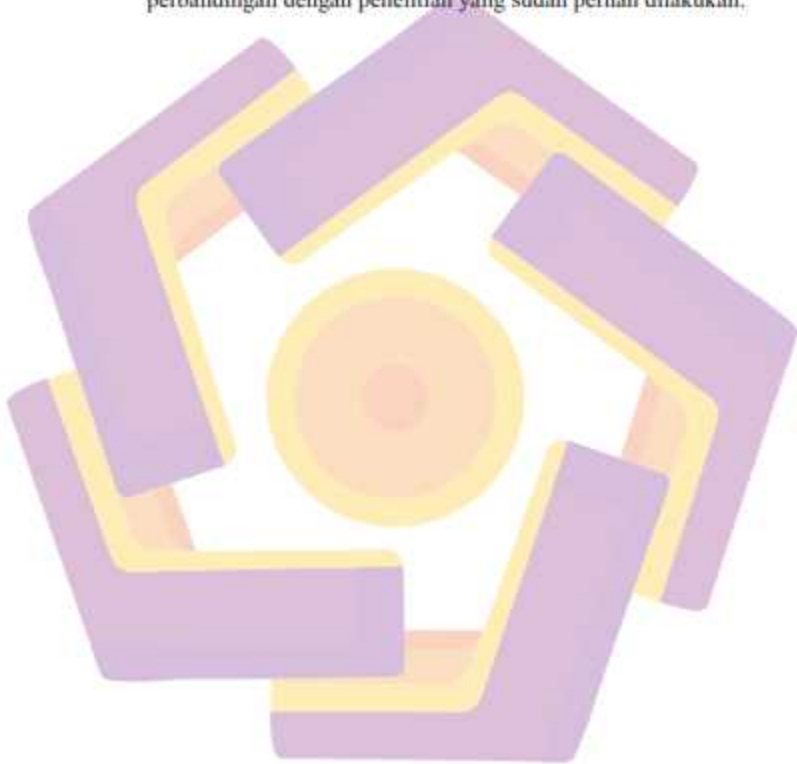
bahwa perbedaan antar cluster bukan bersifat acak, melainkan didorong oleh kombinasi variabel produksi yang dominan.

3. Hasil pengelompokan wilayah menggunakan partitioning around medoids (PAM) dan hierarchical clustering menunjukkan adanya dua cluster utama dengan karakteristik produksi hortikultura yang berbeda secara signifikan. Pada cluster 1 mewakili wilayah dengan tingkat produksi yang relatif rendah, sedangkan pada cluster 2 merupakan sentra produksi hortikultura. Perbedaan karakteristik ini dapat memberikan dasar bagi pemangku kebijakan untuk merumuskan strategi pengembangan pertanian yang lebih terarah. Kesimpulan merupakan pernyataan singkat, jelas, dan tepat tentang apa yang diperoleh, memuat keunggulan dan kelemahan, dapat dibuktikan, serta terkait langsung dengan Rumusan Masalah dan Tujuan Penelitian. Uraian pada bagian ini harus merupakan pernyataan yang pernah dianalisis/dibahas pada bagian sebelumnya, bukan pernyataan yang sama sekali baru dan tidak pernah dibahas pada bagian sebelumnya, serta merupakan jawaban atas permasalahan yang dirumuskan. Bagian ini tidak perlu ada uraian penjelasan lagi.

5.2 Saran

Beberapa saran yang dapat digunakan untuk pengembangan penelitian ini adalah :

1. Pada penelitian selanjutnya dapat menggunakan dataset dengan beberapa variabel dataset yang lebih bervariasi dan dengan jumlah dataset yang lebih besar.
2. Dapat menggunakan pendekatan dengan algoritma lain untuk perbandingan dengan penelitian yang sudah pernah dilakukan.



DAFTAR PUSTAKA

- [1] R. Zaeroni and S. Rustariyuni, "Pengaruh Produksi Beras, Konsumsi Beras Dan Cadangan Devisa Terhadap Impor Beras Di Indonesia," *E-Jurnal Ekon. Pembang. Univ. Udayana*, vol. 5, no. 9, pp. 993–1010, 2016.
- [2] D. Yulianti and Hedwigis, "The Influence of Macroeconomics Indicators to Import Rice in Indonesia. Sustainable Competitive Advantage (SCA)," 2013.
- [3] S. Konyep, "Upaya Pencapaian Swasembada Pangan Melalui Membumikan Padi Amfibi Balitbangtan di Provinsi Papua Barat," *J. Trit.*, vol. 11, no. 2, pp. 32–41, 2020, doi: 10.47687/jt.v11i2.115.
- [4] K. P. Simanjuntak and U. Khaira, "Hotspot Clustering in Jambi Province Using Agglomerative Hierarchical Clustering Algorithm Pengelompokan Titik Api di Provinsi Jambi dengan Algoritma Agglomerative Hierarchical Clustering," vol. 1, no. April, pp. 7–16, 2021.
- [5] G. A. Mauser and J. A. Hartigan, "Clustering Algorithms," *J. Mark. Res.*, vol. 14, no. 1, p. 124, 1977, doi: 10.2307/3151073.
- [6] M. H. Asnawi, P. F. Rahmah, A. Nugraha, and T. Purwandari, "Pemetaan Kabupaten / Kota di Jawa Barat Berdasarkan Jenis Usaha Pertanian Menggunakan Analisis Korespondensi," *Semin. Nas. Stat. X*, no. June, 2021.
- [7] S. Amelia, E. Rustiadi, B. Barus, and B. Juanda, "Mapping the Diversity of Regional Characteristics Towards Sustainable Economic Strategic Area Development: A Case Study of West-East Corridor of West Sumatra Province," *Int. J. Sustain. Dev. Plan.*, vol. 17, no. 1, pp. 185–193, 2022, doi: 10.18280/ijstdp.170118.
- [8] I. P. Kurniawati, H. Pratiwi, and S. Sugiyanto, "Indonesian Territory Clustering Based On Harvested Area and Rice Productivity Using Clustering Algorithm," *J. Soc. Sci.*, vol. 4, no. 1, pp. 100–110, 2023, doi: 10.46799/jss.v4i1.510.
- [9] N. Nurdin, R. Meiyanti, and M. Maulita, "Clustering and Mapping of Agricultural Production Based on Geographic Information System Using K-Medoids Algorithm," 2025.

- [10] H. Ratna, A. Putri, A. Achmad, R. Fernandes, and A. Iriany, "Credit Customer Segmentation With Hierarchical Clustering At Various Distances," *J. Theor. Appl. Inf. Technol.*, vol. 31, no. 2, 2023, [Online]. Available: www.jatit.org
- [11] W. W. Prastanika and A. W. Wijayanto, "Analisis Hard dan Soft Clustering Untuk Pengelompokan Indikator Ketahanan Pangan Indonesia 2021 Hard and Soft Clustering Analysis for Grouping Indonesian Food Security Indicators in 2021," *J. Sist. dan Teknol. Inf.*, vol. 11, no. 4, pp. 596–604, 2023, doi: 10.26418/justin.v11i4.68400.
- [12] R. E. Caraka *et al.*, "Cluster around Latent Variable for Vulnerability towards Natural Hazards, Non-Natural Hazards, Social Hazards in West Papua," *IEEE Access*, vol. 9, pp. 1972–1986, 2021, doi: 10.1109/ACCESS.2020.3038883.
- [13] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis: Pearson Prentice Hall*. 2007.
- [14] D. Widyadhana, R. B. Hastuti, I. Kharisudin, and F. Fauzi, "Perbandingan Analisis Klaster K-Means dan Average Linkage untuk Pengklasteran Kemiskinan di Provinsi Jawa Tengah," *Prism. Pros. Semin. Nas. Mat.*, vol. 4, pp. 584–594, 2021.
- [15] R. J. Freun, W. J. Wilso, and D. L. Moh, *Statistical Methods, Third Edition*. 2010. doi: 10.1016/C2009-0-20216-9.
- [16] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," *J. Tekno Kompak*, vol. 15, no. 2, p. 25, 2021, doi: 10.33365/jtk.v15i2.1162.
- [17] N. K. Kaur, U. Kaur, and D. Singh, "K-Medoid Clustering Algorithm- A Review," *Int. J. Comput. Appl. Technol.*, vol. 1, no. 1, pp. 42–45, 2014.
- [18] S. Velamparambil, S. Mackinnon-Cormier, J. Perry, R. Lemos, M. Okoniewski, and J. Leon, "GPU accelerated Krylov subspace methods for computational electromagnetics," *Proc. 38th Eur. Microw. Conf. EuMC 2008*, vol. 1, no. 9, pp. 1312–1314, 2008, doi: 10.1109/EUMC.2008.4751704.
- [19] J. Ha, M. Kambe, and J. Pe, *Data Mining: Concepts and Techniques*. 2011.

doi: 10.1016/C2009-0-61819-5.

- [20] E. Hartini, "Classification of Missing Values Handling Method During Data Mining: Review," *Sigma Epsil.*, vol. 21, no. 2, pp. 49–60, 2017.
- [21] D. Rachmatin and K. Sawitri, "Perbandingan antara metode agglomeratif, metode divisif dan metode k-means dalam analisis kluster," *Semin. Nas. Mat. UNPAR*, vol. 1, pp. 9–17, 2019, [Online]. Available: <http://eprints.itenas.ac.id/157/>
- [22] J. Supranto, "Analisis multivariat: arti dan interpretasi." PT. Rineka Cipta, 2004.
- [23] A. Sholiha, "Perbandingan Analisis Kluster Menggunakan Metode Single Linkage, Complete Linkage, Average Linkage Dan K-Means Untuk Pengelompokan," *Semarang Univ. Negeri Semarang*, vol. 1, no. 2, p. 3, 2015.
- [24] M. Nishom, "Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square," *J. Inform. J. Pengemb. IT*, vol. 4, no. 1, pp. 20–24, 2019, doi: 10.30591/jpit.v4i1.1253.
- [25] Riska Chairunisa, Adiwijaya, and Widi Astuti, "Perbandingan CART dan Random Forest untuk Deteksi Kanker berbasis Klasifikasi Data Microarray," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 5, pp. 805–812, 2020, doi: 10.29207/resti.v4i5.2083.
- [26] N. : Moh'd Alodat, "Instructor Information Schedule of Classes: Week Topics to be covered 1 Multivariate Data: Summary Statistics 2 Multivariate Data: Plots 3 The Multivariate Normal Distribution 4 Inference about multivariate means 5 Multivariate analysis of variance 6 Exam," *Prentice Hall*, p. 9, 2012, [Online]. Available: www.statsci.org
- [27] A. Afrimayani, H. Yozza, and D. Devianto, "Pengelompokan Negara Di Dunia Berdasarkan Data Runtun Waktu Realisasi Penanaman Modal Asing Di Indonesia Menggunakan Analisis Cluster," *J. Mat. UNAND*, vol. 8, no. 2, p. 157, 2019, doi: 10.25077/jmu.8.2.157-164.2019.
- [28] L. Vendramin, P. A. Jaskowiak, and R. J. G. B. Campello, "On the combination of relative clustering validity criteria," *ACM Int. Conf.*

Proceeding Ser., pp. 733–744, 2013, doi: 10.1145/2484838.2484844.

- [29] M. Wangge, “Penerapan Metode Principal Component Analysis (PCA) Terhadap Faktor-faktor yang Mempengaruhi Lamanya Penyelesaian Skripsi Mahasiswa Program Studi Pendidikan Matematika FKIP UNDANA,” *J. Cendekia J. Pendidik. Mat.*, vol. 5, no. 2, pp. 974–988, 2021, doi: 10.31004/cendekia.v5i2.465.
- [30] A. S. Ritonga and I. Muhandhis, “Teknik Data Mining Untuk Mengklasifikasikan Data Ulasan Destinasi Wisata Menggunakan Reduksi Data Principal Component Analysis (Pca),” *Edutic - Sci. J. Informatics Educ.*, vol. 7, no. 2, 2021, doi: 10.21107/edutic.v7i2.9247.
- [31] P. Di, K. Bojonegoro, D. Hedyati, and I. M. Suartana, “Penerapan Principal Component Analysis (PCA) Untuk Reduksi Dimensi Pada Proses Clustering Data Produksi,” vol. 05, pp. 49–54, 2021.

