



Pengelompokan Tanaman Perkebunan Berdasarkan Produktivitas dan Luas Lahan dengan K- Means Clustering

Ethaniel Williano Adhi Putra, Yunus Widjaja*

Teknologi dan Desain, Sistem Informasi, Universitas Pembangunan Jaya, Tangerang Selatan
Blok B7/P, Jl. Boulevard UPJ No.65, RT.4/RW.3, Sawah Baru, Kec. Ciputat, Tangerang Selatan, Indonesia

Email: ¹ethaniel.willianoadhi@student.upj.ac.id, ^{2,*}yunus.widjaja@upj.ac.id

Email Penulis Korespondensi: yunus.widjaja@upj.ac.id

Submitted: 10/03/2026; Accepted: 06/04/2026; Published: 07/04/2026

Abstrak—Data perkebunan di Jawa Barat dikelompokkan berdasarkan luas lahan dan produktivitas tanaman menggunakan metode K-Means. Data ini diperoleh dari Open Data Jabar dari tahun 2022 hingga 2024 dan dianalisis dengan pendekatan kuantitatif. Tiga kelompok dapat diidentifikasi berdasarkan hasil clustering: satu kelompok memiliki produktivitas tinggi tetapi memiliki luas lahan yang relatif terbatas, yang lain memiliki luas lahan yang luas tetapi produktivitasnya belum optimal, dan yang terakhir memiliki produktivitas dan luas lahan yang sama-sama rendah. Hasilnya menunjukkan bahwa luas lahan tidak selalu sebanding dengan produktivitas. Studi ini menekankan betapa pentingnya memilih variabel yang relevan dan menggunakan metode secara konsisten untuk menghasilkan analisis yang lebih akurat dan mudah dipahami.

Kata Kunci: K-Means Clustering; Produktivitas Tanaman; Luas Lahan; Perkebunan Jawa Barat; Analisis Cluster.

Abstract—Plantation data in West Java was grouped based on land area and crop productivity using the K-Means method. This data was obtained from Open Data Jabar from 2022 to 2024 and analyzed using a quantitative approach. Three groups can be identified based on the clustering results: one group has high productivity but relatively limited land area, another has large land area but suboptimal productivity, and the last group has equally low productivity and land area. The results indicate that land area does not always correlate with productivity. This study emphasizes the importance of selecting relevant variables and using methods consistently to produce more accurate and understandable analyses.

Keywords: K-Means Clustering; Crop Productivity; Land Area; West Java Plantations; Cluster Analysis.

1. PENDAHULUAN

Selain berkontribusi terhadap ketahanan pangan nasional, sektor perkebunan Jawa Barat memainkan peran penting dalam mendorong perekonomian lokal. Berbagai produk seperti teh, kopi, kelapa, karet, dan kakao tersebar di berbagai wilayah dengan iklim yang berbeda. Pada akhirnya, tingkat produktivitas yang dihasilkan dipengaruhi oleh perbedaan karakteristik daerah ini. Dalam praktiknya, masih ada variasi dalam tingkat produktivitas antara wilayah. Beberapa wilayah memiliki tingkat produksi yang tinggi, sementara wilayah lain memiliki tingkat produksi yang rendah atau tidak stabil. Kondisi ini dipengaruhi oleh banyak faktor alam, seperti iklim, kesuburan tanah, dan jumlah air yang tersedia. Pengelolaan juga memainkan peran, seperti penerapan teknologi dan metode budidaya yang diterapkan di masing-masing daerah.

Selain hal-hal yang disebutkan di atas, pemanfaatan data untuk membantu pengambilan keputusan masih belum optimal. Data statistik perkebunan sebenarnya banyak tersedia, tetapi biasanya hanya digunakan sebagai laporan deskriptif tanpa diproses lebih lanjut untuk menemukan pola atau kecenderungan. Akibatnya, kebijakan yang dibuat, seperti menetapkan prioritas pengembangan atau memberikan bantuan, belum sepenuhnya bergantung pada analisis data. Meskipun demikian, dengan menggunakan pendekatan yang lebih sistematis, data dapat memberikan gambaran yang lebih jelas tentang kondisi produktivitas di setiap area dan membantu dalam menentukan tindakan yang lebih tepat. Dalam data mining, teknik clustering adalah salah satu metode yang dapat digunakan untuk memahami pola dalam data. Metode ini memungkinkan untuk mengelompokkan daerah berdasarkan ciri-ciri tertentu yang mirip, sehingga daerah dengan kondisi yang hampir identik dapat dianalisis dalam satu kelompok. Dalam penelitian ini, kelompokan dibuat berdasarkan luas lahan dan tingkat produktivitas tanaman. Kedua faktor ini dianggap dapat mewakili kondisi utama perkebunan. Diharapkan dengan pengelompokan ini dapat diperoleh gambaran yang lebih terstruktur tentang variasi produktivitas di seluruh Jawa Barat.

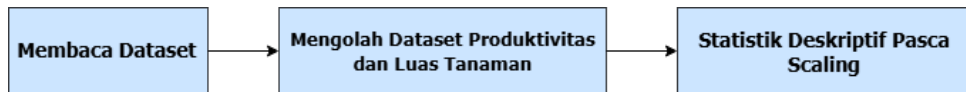
Studi sebelumnya telah menunjukkan bahwa teknik clustering, terutama K-Means, cukup efektif dalam mengelompokkan data di bidang pertanian dan perkebunan. Penelitian yang dilakukan oleh Rosa Rosiana dan Willy Prihartono (2025) menjelaskan Algoritma K-Means dipilih karena kemampuannya yang unggul dalam mengelompokkan data berukuran besar dan kemampuannya mengidentifikasi kelompok dengan karakteristik yang serupa [1]. Meskipun demikian, Pratama dan Sari (2021) menggunakan teknik yang sama untuk membagi wilayah produksi padi menjadi kelompok dan membuat saran untuk pengembangan yang paling penting [2]. Penelitian tambahan oleh Faiq dan Agung (2022) menggabungkan K-Means dan K-Medoids untuk menentukan kemungkinan produksi kelapa sawit; penelitian ini menunjukkan bahwa teknik clustering dapat menunjukkan kesamaan karakteristik wilayah [3]. Selain itu, penelitian oleh Ade Aisyah dan Nur Irhamni (2025) menunjukkan bahwa variasi dalam jumlah cluster dapat memengaruhi tingkat detail hasil pengelompokan [4]. Oleh karena itu, penentuan jumlah cluster sangat penting untuk analisis.

Namun, penelitian biasanya lebih berfokus pada penerapan teknik daripada membahas secara menyeluruh bagaimana hasil pengelompokan dapat digunakan dalam konteks yang lebih luas. Selain itu, untuk menentukan jumlah produktivitas tanaman serta luas tanaman itu sendiri, sebagian besar penelitian menggunakan satu pendekatan evaluasi. Akibatnya, hasil yang dihasilkan belum tentu menggambarkan struktur data secara optimal. Sebaliknya, penelitian ini terlalu terbatas pada satu komoditas, sehingga tidak memberikan gambaran yang cukup luas tentang keadaan di daerah tersebut. Ini menunjukkan bahwa penelitian yang mempertimbangkan hasil analisis sebagai bagian dari pemahaman kondisi produktivitas daripada hanya sebagai produk dari proses metodologis masih diperlukan.

Penelitian ini diharapkan dapat memberikan kontribusi yang lebih nyata dari sudut pandang akademis dan praktis. Dari sudut pandang akademis, penelitian ini akan meningkatkan pemanfaatan teknik clustering dalam analisis data pertanian, khususnya pada skala wilayah yang lebih luas. Dari sudut pandang praktis, hasil penelitian dapat digunakan oleh pemerintah daerah untuk membuat kebijakan yang lebih tepat sasaran. Pengembangan sektor perkebunan diharapkan dapat dilakukan secara lebih efisien dan berkelanjutan dengan pemahaman yang lebih baik terhadap pola produktivitas.

2. METODOLOGI PENELITIAN

Pada penelitian yang dilakukan dengan menggunakan pendekatan kuantitatif dengan metode data mining, khususnya dengan menggunakan metode algoritma K-Means Clustering, untuk mengelompokkan tanaman perkebunan di Provinsi Jawa Barat mulai dari jumlah produktivitas tanaman serta luas tanaman perkebunan dengan menggunakan K-Means [5]. Untuk datanya sendiri diambil di Open Data Jawa Barat. Pendekatan kuantitatif yang saya pilih karena penelitian ini bertujuan pada pengolahan data berupa angka, pengujian model secara statistik, serta pengukuran hasil berdasarkan nilai evaluasi dengan terukur [6]. Pada penelitian ini disusun dalam empat tahapan yaitu membaca dataset, data training, featurer scaling, dan clustering [7].



Gambar 1. Alur penelitian k-means

2.1 Membaca Dataset

Data dalam penelitian ini berasal dari Open Data Jabar dalam format CSV yang berisi informasi hasil perkebunan di Provinsi Jawa Barat pada periode 2022 hingga 2024 [8]. Variabel yang terdapat dalam dataset meliputi id, kode kabupaten kota, produktivitas tanaman, luas tanaman, dan tahun pengamatan [9]. Data diolah menggunakan bahasa pemrograman Python dengan bantuan library pandas sehingga dapat diubah menjadi dataframe untuk memudahkan analisis [10]. Hasil pembacaan menunjukkan bahwa dataset memiliki 207 baris observasi, dan seluruh kolom utama juga memiliki jumlah data 207, sehingga tidak ditemukan nilai kosong pada data yang digunakan [11].

2.2 Mengolah Dataset Produktivitas dan Luas Tanaman

Tabel 1. Statistik deskriptif dataset produktivitas perkebunan

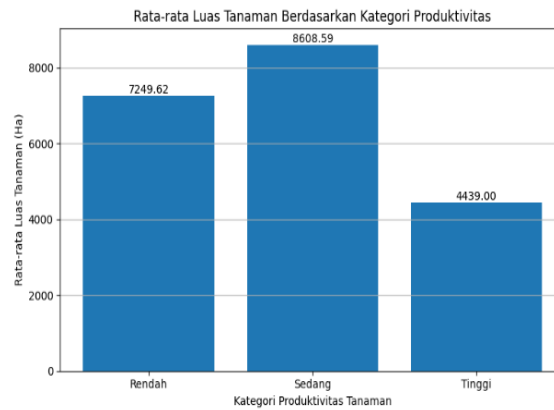
	id	kode kabupaten kota	produktivitas tanaman	luas tanaman	tahun
count	207.0	207.0	207.0	207.0	207.0
mean	104.0	3232.608695652174	2137.9468599033817	7347.652173913043	2022.6521739130435
std	59.89991652748775	31.330163538027435	3545.4832967935463	6882.811546966187	0.7000572852396709
min	1.0	3201.0	0.0	167.0	2022.0
25%	52.5	3208.0	301.0	1607.0	2022.0
50%	104.0	3215.0	637.0	4439.0	2023.0
75%	155.5	3273.0	1287.5	12119.0	2023.0
max	207.0	3279.0	26000.0	19342.0	2024.0

Selanjutnya data diolah dan dilakukan analisis statistik deskriptif untuk melihat gambaran awal data. Rata-rata produktivitas tanaman sebesar 2137,95, sedangkan rata-rata luas tanaman 7347,65 hektar, dengan rata-rata tahun 2022,65 yang menunjukkan data berasal dari periode 2022 hingga 2024 [12]. Nilai standar deviasi produktivitas 3545,48 dan luas tanaman 6882,81 menunjukkan adanya variasi data antar wilayah [13]. Nilai minimum produktivitas 0 dan luas tanaman 167 hektar, sedangkan nilai maksimum produktivitas 26000 dan luas tanaman 19342 hektar, dengan tahun maksimum 2024 [14].

Dataset yang telah dibaca kemudian dianalisis dengan statistik deskriptif untuk mendapatkan gambaran umum tentang data [15]. Hasilnya menunjukkan rata-rata produktivitas tanaman sebesar 2137,95 dan rata-rata luas tanaman 7347,65 hektar [16]. Nilai rata-rata tahun 2022,65 menandakan bahwa data berasal dari periode 2022

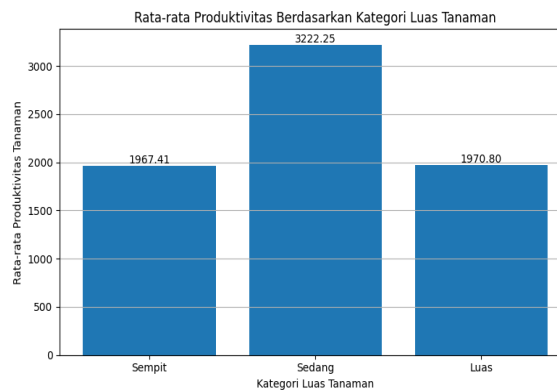
hingga 2024. Selain itu, standar deviasi produktivitas 3545,48 dan standar deviasi luas tanaman 6882,81 menunjukkan adanya perbedaan yang cukup besar antar wilayah, baik dari sisi produksi maupun luas lahan [17].

Nilai minimum produktivitas tercatat 0, sedangkan luas tanaman minimum 167 hektar dengan tahun minimum 2022 [18]. Di sisi lain, nilai maksimum produktivitas mencapai 26000 dan luas tanaman 19342 hektar, dengan tahun maksimum 2024. Median produktivitas 637 yang kurang dari nilai rata-rata menunjukkan bahwa pembagian data tidak merata karena terdapat beberapa nilai yang jauh lebih tinggi dibandingkan data lainnya.



Gambar 2. Grafik Rata – rata Luas Tanaman Berdasarkan Kategori Produktivitas

Hubungan antara produktivitas tanaman dan luas lahan tidak selalu sejalan, seperti yang ditunjukkan pada grafik. Luas lahan rata-rata untuk kategori produktivitas rendah adalah 7249,62 ha, dengan contoh serehwangi di Kabupaten Kuningan (produktivitas 258 kg luas 10.172 ha), nilam di Kabupaten Cianjur (produktivitas 0 luas 17.933 ha), dan kapok di Kabupaten Bandung (produktivitas 258 kg luas 1.084 ha), yang menunjukkan bahwa luas lahan tidak menjamin hasil tinggi. Meskipun demikian, luas lahan yang paling kecil digunakan dalam kategori lahan yang memiliki produktivitas tinggi adalah 4439,00 ha, dengan contoh produktivitas pinang di Kabupaten Tasikmalaya (13.748 kg dengan luas 12.119 ha), serehwangi di Kabupaten Kuningan (7.179 ha dengan luas 1.084 ha), dan pinang di Kabupaten Ciamis (9.611 kg dengan luas 9.844 ha). Hal ini menunjukkan bahwa tidak hanya luas lahan yang mempengaruhi produktivitas tanaman, tetapi juga faktor lain, seperti metode pengelolaan, kualitas tanah, dan penggunaan teknologi pertanian [19].



Gambar 3. Grafik Hasil Perkebunan Berdasarkan Produktivitas Tanaman Perkebunan

Grafik menunjukkan bahwa kategori lahan sedang (3222,25) memiliki produktivitas rata-rata tertinggi, sementara kategori lahan sempit (1967,41) dan luas (1970,80) masing-masing hampir sama dan lebih rendah [20]. Beberapa komoditas menunjukkan pola ini, menurut data CSV. Pada lahan sedang, tanaman seperti kopi arabika dan kelapa hibrida menghasilkan hasil yang relatif stabil dan optimal, sementara pada lahan luas, tanaman seperti nilam dan kelapa tidak selalu menghasilkan hasil yang tinggi meskipun area mereka besar, sementara lahan sempit cenderung menghasilkan hasil yang lebih sedikit karena skala produksi yang lebih kecil. Ini menunjukkan bahwa pengelolaan, jenis tanaman, dan efisiensi budidaya lebih penting daripada ukuran lahan.

2.3 Statistik Deskriptif Pasca Scaling

Dengan menggunakan variabel luas lahan dan produktivitas sebagai dasar analisis, data penelitian ini telah dinormalisasi untuk berada dalam rentang 0 hingga 1. Semua data menunjukkan kabupaten dan kota Jawa Barat [21]. Nilai yang lebih rendah dari 0 menunjukkan kondisi yang lebih rendah, sedangkan nilai yang lebih tinggi menunjukkan kondisi yang lebih tinggi. Ada perbedaan antara daerah yang memiliki banyak lahan tetapi memiliki produktivitas yang rendah dan sebaliknya. Untuk menghindari dominasi variabel K-Means, prosedur ini bertujuan

Cluster 0 berisi data dengan tingkat produktivitas rendah hingga sedang. Sebagian besar nilainya berada di bawah atau mendekati rata-rata dataset sekitar 2.137, bahkan ada yang mendekati 0, sehingga cluster ini menggambarkan wilayah dengan produksi yang relatif rendah.

b. Cluster 1

Cluster 1 menunjukkan produktivitas pada tingkat menengah hingga cukup tinggi. Nilainya lebih baik dibanding cluster 0, namun belum mencapai nilai maksimum dalam dataset. Wilayah pada cluster ini dapat dikatakan memiliki produksi yang cukup stabil .

c. Cluster 2

Cluster 2 merupakan kelompok dengan produktivitas paling tinggi. Data dengan nilai besar, termasuk yang mendekati 26.000 merupakan tanaman pala, umumnya berada pada cluster ini. Hal ini menunjukkan wilayah dengan hasil produksi yang paling unggul.

Secara umum, jumlah data pada cluster 0 biasanya paling banyak, cluster 1 berada di tingkat menengah, sedangkan cluster 2 cenderung lebih sedikit karena hanya mencakup wilayah dengan produktivitas sangat tinggi.

3.1.3 Nilai Centroid Cluster 0, 1, dan 2

Tabel 4. Hasil Nilai Centroid Cluster 0, 1, dan 2

Keterangan	Nilai Centroid Cluster 0, 1, dan 2
array	$([-0.39512485, 1.23322981], [-0.42269614, -0.70801286], [1.96108865, 0.07373321])$

a. Cluster 0

Pada cluster 0 ini Kelompok ini memiliki karakteristik produktivitas tinggi (1,23) dan luas lahan relatif rendah (-0,39), yang menunjukkan bahwa wilayah-wilayah ini mampu menghasilkan produksi yang optimal meskipun memiliki keterbatasan lahan. Kondisi ini menunjukkan adanya efisiensi dalam pengelolaan perkebunan, di mana hasil yang diperoleh tetap tinggi tanpa bergantung pada luas lahan yang besar. Cluster ini biasanya terdiri dari daerah yang menghasilkan komoditas bernilai ekonomi tinggi seperti teh, kopi, dan kakao, yang dikenal memiliki tingkat produktivitas tinggi. Tempat-tempat yang memiliki ciri-ciri ini biasanya terletak di wilayah dataran tinggi seperti Bandung, Garut, dan Tasikmalaya, yang memiliki kondisi lingkungan yang ideal untuk pertumbuhan dan hasil perkebunan.

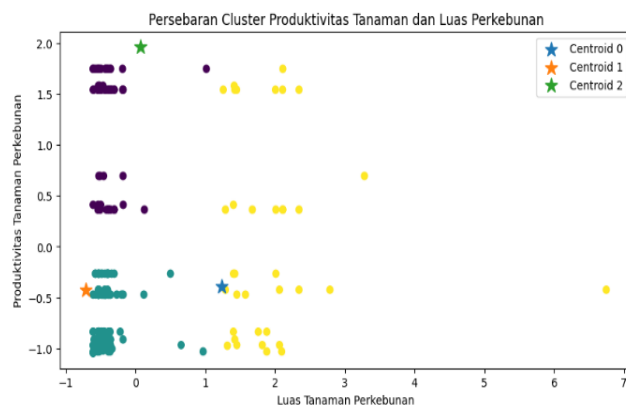
a. Cluster 1

Pada cluster 1 memiliki karakteristik produktivitas yang rendah (-0,70) dan luas lahan yang relatif rendah (-0,42), menunjukkan kondisi perkebunan yang buruk di daerah tersebut. Kondisi ini menunjukkan bahwa pemanfaatan lahan dan hasil produksi masih kurang optimal. Cluster ini menunjukkan daerah dengan pengelolaan perkebunan yang masih terbatas, di mana komoditas utama adalah kelapa dan karet dengan sistem budidaya tradisional. Wilayah dengan karakteristik seperti ini biasanya berada di daerah pesisir atau dataran rendah, seperti Cirebon dan Indramayu, yang memiliki kesulitan tersendiri dalam meningkatkan produktivitas perkebunan [23].

b. Cluster 2

Pada cluster 2 memiliki karakteristik luas lahan yang sangat tinggi (1,96) dan tingkat produktivitas sedang (0,07), menunjukkan area yang memiliki banyak potensi lahan tetapi belum dimanfaatkan dengan baik. Kondisi ini menunjukkan bahwa, meskipun ada banyak lahan yang tersedia, hasil produksi yang diperoleh belum sebanding dengan potensi yang ada. Cluster ini biasanya menggambarkan area perkebunan berskala besar di mana tanaman kelapa dan karet adalah tanaman utama. Karakteristik seperti ini umumnya ditemukan di wilayah seperti Subang, Sukabumi, dan Bogor, yang memiliki potensi lahan yang besar tetapi masih membutuhkan peningkatan dalam pengelolaan untuk memaksimalkan produktivitasnya.

3.2 Visualisasi Hasil Clustering



Gambar 4. Persebaran Cluster Produktivitas Tanaman dan Luas Perkebunan

Hasil visualisasi menunjukkan bahwa data dibagi menjadi tiga cluster, masing-masing diwakili oleh centroid, yang merupakan titik pusat dari setiap cluster dan menunjukkan nilai rata-rata dari semua anggota kelompok. Untuk memahami lebih lanjut tentang karakteristik masing-masing cluster, kita dapat melihat posisi centroid terhadap sumbu X (luas lahan) dan sumbu Y (produktivitas).

a. Centroid 0

Cluster pertama, dengan centroid 0, berada pada posisi dengan nilai luas lahan yang cenderung sedang dan produktivitas yang relatif rendah. Hasil normalisasi menunjukkan bahwa centroid ini memiliki nilai luas lahan positif (sekitar 1) dan produktivitas negatif (sekitar -0,4), menunjukkan bahwa area dalam cluster ini memiliki luas lahan yang cukup besar dibandingkan rata-rata, tetapi tingkat produktivitasnya masih di bawah rata-rata. Kondisi ini menunjukkan bahwa lahan memiliki potensi yang belum dimanfaatkan dengan baik. Dalam data perkebunan di Jawa Barat, situasi seperti ini dapat terjadi di daerah yang memiliki lahan yang luas tetapi tidak didukung oleh teknik pertanian yang efektif, penggunaan pupuk yang optimal, atau sistem pengelolaan yang baik.

b. Centroid 1

Centroid 1, cluster kedua, menunjukkan karakteristik dengan luas lahan yang kecil dan produktivitas yang rendah. Nilai centroid untuk masing-masing variabel berada pada posisi negatif, kira-kira -0,8 untuk luas lahan dan -0,4 untuk produktivitas. Cluster ini menunjukkan beberapa daerah yang memiliki keterbatasan sumber daya lahan dan hasil produksi. Di bagian kiri bawah grafik, titik-titik dalam kelompok ini terkonsentrasi, menunjukkan bahwa daerah-daerah tersebut memiliki kondisi perkebunan yang kurang menguntungkan. Kondisi ini sering terjadi di daerah dengan keterbatasan lahan, kurangnya akses teknologi, atau kondisi geografis yang tidak mendukung pengembangan perkebunan.

c. Centroid 2

Centroid 2, cluster ketiga, memiliki ciri-ciri yang paling menonjol, terutama dalam hal produktivitas. Cluster ini memiliki produktivitas yang tinggi (sekitar 2 dalam hasil normalisasi), sementara luas lahannya sedang (sekitar 0,2), menunjukkan bahwa wilayahnya mampu menghasilkan output yang tinggi meskipun luas lahannya tidak sebesar cluster lainnya. Dengan kata lain, kelompok ini menunjukkan efisiensi penggunaan lahan.

Kawasan yang telah menerapkan teknik budidaya yang lebih canggih, penggunaan bibit unggul, dan pengelolaan yang lebih terencana biasanya mengalami kondisi ini. Selain itu, kondisi lingkungan seperti kondisi iklim yang mendukung dan kesuburan tanah juga dapat berkontribusi terhadap tingginya produktivitas. Di Jawa Barat, kelompok ini mungkin menunjukkan daerah yang telah berkembang dalam industri perkebunan dengan sistem pengelolaan yang lebih baik. Dikenal bahwa industri perkebunan Jawa Barat memiliki berbagai produk unggulan seperti karet, kopi, dan teh jika dibandingkan dengan situasi nyata di lapangan. Namun, dalam proses clustering, jenis komoditas tidak digunakan sebagai variabel dalam penelitian ini. Akibatnya, penyebutan komoditas hanya digunakan sebagai ilustrasi untuk membantu memahami konteks hasil analisis, bukan sebagai bagian dari variabel yang dianalisis. Hal ini dilakukan untuk menjaga konsistensi antara metodologi dan diskusi, dan untuk mencegah penambahan asumsi yang tidak didukung oleh data [24].

Selain itu, hasil analisis ini menunjukkan bahwa hubungan antara luas lahan dan produktivitas tidak selalu bersifat linear. Ada situasi di mana luas lahan yang besar tidak diikuti oleh produktivitas yang tinggi, dan sebaliknya, wilayah dengan luas lahan yang kecil mampu menghasilkan produktivitas yang tinggi. Ini menunjukkan bahwa produktivitas tidak hanya ditentukan oleh luas lahan tetapi juga oleh faktor lain, seperti kualitas pengelolaan. Penelitian ini menekankan pemilihan variabel yang tepat untuk proses clustering. Hasil pengelompokan menjadi lebih mudah dipahami dan memiliki makna yang jelas dengan hanya menggunakan variabel yang relevan, yaitu produktivitas dan luas lahan. Hal ini juga menghindari kesalahan umum dalam analisis kelompok, seperti memasukkan variabel seperti kode wilayah, yang tidak terkait dengan tujuan penelitian. Selain itu, konsistensi dalam penggunaan metode normalisasi juga menjadi hal yang penting dalam penelitian ini. Menggunakan StandardScaler secara teratur membuat hasil analisis lebih stabil dan lebih mudah diinterpretasikan. Selain itu, ini menjawab kritik tentang penggunaan berbagai metode normalisasi dalam satu penelitian, yang dapat menyebabkan hasil yang tidak konsisten.

Secara keseluruhan, hasil klasifikasi ini memberikan gambaran yang cukup jelas tentang keadaan industri perkebunan di Jawa Barat. Cluster dengan tingkat produktivitas tinggi dapat digunakan sebagai contoh atau benchmark bagi wilayah lain untuk meningkatkan kinerja produksi. Di sisi lain, cluster dengan tingkat produktivitas rendah dapat menjadi fokus pengembangan melalui pengembangan teknologi, pelatihan SDM, dan kebijakan yang mendukung industri perkebunan. Oleh karena itu, penelitian ini tidak hanya mengumpulkan data tetapi juga memberikan pengetahuan untuk membantu pengambilan keputusan. Untuk meningkatkan produktivitas secara keseluruhan di Provinsi Jawa Barat, pemerintah daerah dan pihak terkait dapat menggunakan hasil ini untuk membuat strategi pengembangan sektor perkebunan yang lebih efisien, efisien, dan tepat sasaran. Hasil pengelompokan juga dapat digunakan sebagai dasar untuk evaluasi kebijakan di industri perkebunan, khususnya untuk menentukan prioritas pengembangan wilayah. Karena setiap kelompok memiliki fitur unik, metode yang digunakan untuk meningkatkan produktivitas tidak dapat disamaratakan. Misalnya, fokus utama dapat diarahkan pada peningkatan efisiensi pengelolaan lahan, seperti penerapan teknologi pertanian modern, penggunaan pupuk yang tepat, dan peningkatan kualitas sumber daya manusia melalui pelatihan. Sebaliknya, cluster dengan luas lahan



kecil dan produktivitas rendah memerlukan pendekatan yang lebih mendasar, seperti memberikan dukungan untuk akses lahan, modal, dan pendampingan dalam pelatihan.

Sebaliknya, kelompok dengan produktivitas tinggi dapat berfungsi sebagai contoh praktik terbaik yang dapat diterapkan oleh daerah lain. Dengan mempelajari pola pengelolaan yang diterapkan pada kelompok-kelompok ini, diharapkan daerah lain dapat meningkatkan produktivitasnya tanpa bergantung pada perluasan lahan, hal ini penting mengingat keterbatasan lahan yang semakin meningkat. Diharapkan hasil clustering ini akan memungkinkan pengambilan keputusan yang lebih terarah dan berbasis data tentang pengembangan industri perkebunan. Metode seperti ini tidak hanya memberikan gambaran tentang kondisi saat ini tetapi juga memberikan gambaran strategi yang lebih efisien untuk meningkatkan produktivitas di masa depan.

Selain itu, temuan analisis ini dapat digunakan sebagai dasar untuk perencanaan yang akan datang untuk mengembangkan sektor perkebunan secara berkelanjutan dalam jangka panjang. Pemerintah daerah dan pemangku kepentingan dapat membuat strategi yang lebih khusus untuk masing-masing wilayah dengan mengetahui pola distribusi produktivitas dan luas lahan melalui pendekatan clustering. Metode berbasis data ini memiliki kemampuan untuk menyesuaikan dengan kebutuhan dan kemungkinan yang ada di setiap wilayah, yang membuatnya lebih efektif dibandingkan dengan kebijakan umum. Selain itu, dapat dikembangkan lebih lanjut pemanfaatan hasil clustering dengan memasukkan variabel tambahan ke dalam penelitian berikutnya. Variabel seperti faktor iklim, jenis tanah, atau tingkat penggunaan teknologi pertanian dapat termasuk dalam kategori variabel ini. Tujuannya adalah untuk mendapatkan pemahaman yang lebih komprehensif tentang komponen-komponen yang memengaruhi produktivitas tanaman perkebunan. Namun, agar proses pengelompokan tidak bias, penambahan variabel harus tetap mempertimbangkan relevansinya terhadap tujuan penelitian. Oleh karena itu, penelitian ini tidak hanya membantu dalam analisis data tetapi juga menawarkan peluang untuk pengembangan studi yang lebih mendalam. Hasil yang diperoleh diharapkan dapat berfungsi sebagai pedoman untuk meningkatkan efisiensi dan produktivitas industri perkebunan, khususnya di wilayah Jawa Barat, sehingga dapat berdampak positif pada perekonomian daerah secara keseluruhan. Hasil clustering dapat digunakan sebagai alat untuk memantau dan mengevaluasi kinerja sektor perkebunan dari waktu ke waktu selain dari tujuan perencanaan dan kebijakan. Pergeseran posisi suatu area dari satu cluster ke cluster lain dapat dilihat dengan lebih jelas dengan melakukan analisis secara berkala dengan teknik yang sama. Misalnya, pergeseran dari cluster produktivitas rendah ke cluster yang lebih produktif dapat menunjukkan bahwa program intervensi telah berhasil. Sebaliknya, penurunan cluster di daerah tertentu dapat menunjukkan bahwa ada masalah yang perlu ditangani segera.

Di sisi lain, hasil ini dapat membantu pemilik perkebunan dalam membuat keputusan tentang jenis tanaman, pengelolaan lahan, dan alokasi sumber daya. Dengan memahami bagaimana wilayah mereka berada dalam suatu kelompok, pemilik perkebunan dapat membuat strategi yang lebih tepat dan realistis sesuai dengan situasi yang mereka hadapi. Ini pasti akan meningkatkan efisiensi usaha dan mengurangi risiko kerugian. Pada akhirnya, metode berbasis data seperti yang digunakan dalam penelitian ini menjadi semakin penting di tengah tuntutan untuk meningkatkan efisiensi dan produktivitas sektor pertanian. Teknik analisis seperti K-Means dapat disederhanakan menjadi pola-pola yang lebih mudah dianalisis dan digunakan sebagai dasar pengambilan keputusan.

4. KESIMPULAN

Studi ini menunjukkan bahwa metode K-Means dapat mengelompokkan informasi perkebunan di Jawa Barat berdasarkan dua faktor utama: luas lahan dan produktivitas tanaman. Hasil clustering menghasilkan tiga kelompok yang berbeda. Ada kelompok yang memiliki produktivitas tinggi tetapi memiliki lahan yang relatif terbatas, kelompok yang memiliki lahan yang luas tetapi produktivitas belum optimal, dan kelompok yang memiliki kedua keterbatasan tersebut. Hasil menunjukkan bahwa tidak hanya luas lahan yang mempengaruhi produktivitas, tetapi juga pengelolaan dan efisiensi pemanfaatan lahan. Dari segi metodologi, penelitian ini menekankan penggunaan teknik normalisasi yang konsisten dan pemilihan variabel yang tepat. Dalam proses clustering, variabel seperti kode wilayah, yang tidak memiliki makna analitis, tidak digunakan karena dapat menyebabkan bias. Selain itu, analisis ini tidak memasukkan jenis komoditas sebagai variabel. Akibatnya, interpretasi hasilnya berpusat pada hubungan antara produktivitas dan luas lahan. Kedepannya, penelitian dapat dilakukan dengan menambahkan variabel yang lebih relevan, seperti faktor lingkungan atau teknologi pertanian. Untuk membuat hasil clustering lebih konsisten dan mudah dipahami, penelitian dapat dilakukan dengan menggunakan visualisasi evaluasi seperti Metode Elbow agar hasil clustering lebih dapat diandalkan dan lebih mudah dipahami. Hasil analisis diharapkan dapat memberikan dasar yang lebih kuat untuk pengambilan kebijakan berbasis data berkat penggunaan pendekatan yang lebih konsisten serta data yang dihasilkan lebih akurat.

REFERENCES

- [1] R. Rosiana and W. Prihartono, "Jurnal Informatika Terpadu Implementasi Algoritma K-Means Untuk Pengelompokan Kecamatan Berdasarkan Produktivitas Tanaman Padi Di Kabupaten Cirebon," *Jurnal Informatika Terpadu*, vol. 11, no. 1, pp. 12–19, 2025



- [2] H. Pratiwi, A. Purno, and W. Wibowo, “Mengidentifikasi Kelompok Sektor Perkebunan di Indonesia Berdasarkan Produktivitas Hasil Bumi,” *Diffusion: Journal of Systems and Information Technology*, vol. 2, no. 2, pp. 54 – 61, 2022, doi: <https://doi.org/10.37031/diffusion.v2i2.14999>
- [3] F. H. Pratama *et al.*, “Data Mining K-Medoids Dan K-Means Untuk Pengelompokan Potensi Produksi Kelapa Sawit Di Indonesia,” *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran)*, vol 07, no. 04, pp:1294-1310, 2022
- [4] A. Aisyah, A. Putri, N. I. Sabrina, and O. Meila, “K-Means Klastering Tanaman Biofarma Zingiber officinale Indonesia Tahun 2023,” *Industrial & System Engineering Journals*, vol. 4, no.1, pp. 367-372, 2025, doi: <https://doi.org/10.37477/isejou.v4i1>
- [5] E. Febrianty, L. Awalina, and W. I. Rahayu, “Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail Optimizing Marketing Strategies with Customer Segmentation Using K-Means Clustering on Online Retail Transactions,” *Jurnal Teknologi dan Informasi (JATI)*, vol. 13, 2023, doi: 10.34010/jati.v13i2.
- [6] Y. Pitaloka Angriani *et al.*, “Implementasi Algoritma K-Means Clustering Dalam Menentukan Blok Tanaman Sawit Produktif Pada Pt Arta Prigel,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol.8, no.2, pp. 1820 -1825 2024, doi :<https://doi.org/10.36040/jati.v8i2>.
- [7] A. B. Krisna and D. M. Putri, “Analisis Segmentasi Klien Menggunakan K-Means Clustering Pada Agensi Kreatif Anoorapro,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 2, pp. 1931-1938, 2025, doi: <https://doi.org/10.36040/jati.v9i2.12822>
- [8] F. Ilmawan and A. Ramadhanu, “Implementasi Metode K-Means Untuk Klasterisasi Varietas Parpika Dengan Menggunakan Teknik Pengolahan Citra Digital,” *Jurnal Informatika Teknologi dan Sains*, vol. 7, no. 1, pp. 249–254, 2025, doi: <https://elibrary.ru/item.asp?id=81595670>
- [9] Mochamad Wahyudi Masitha Risna Saragih Solikhun, “Data Mining : Penerapan Algoritma K-Means Clustering dan K-Medoids Clustering,” Penerbit Yayasan Kita Menulis, ISBN: 978-623-7645-81-8 (print) E-ISBN: 978-623-7645-82-5 (online), pp. v-63, 2020, doi:<https://repository.bsi.ac.id/repo/files/346193/download/Salinan-Data-Mining---Mochamad-Wahyudi.pdf>.
- [10] F. Izzatul Muttaqin, Z. Rizky Ramadani, and R. Yotenka, “Implementasi Algoritma K-Means dalam Analisis Cluster Usaha Pertanian Perorangan di Provinsi D. I Yogyakarta Tahun 2023,” *Emerging Statistics and Data Science Journal*, vol. 3, no. 1, pp : 472-481, 2025 doi: <https://doi.org/10.20885/esds.vol3.iss.1.art4>.
- [11] M. A. Septianto, A. Faqih, and A. R. Rinaldi, “Klasterisasi Data Produksi Pertanian Di Kabupaten Cirebon Dengan Algoritma K-Means,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 2, Apr. 2025, doi: 10.23960/jitet.v13i2.6174.
- [12] T. Maahara, A. Saptari, and M. Ula, “Analisis Clustering Lahan Perkebunan Kopi Jenis Tertentu Menggunakan Sistem Informasi Geografis Di Kecamatan Pintu Rime Gayo Kabupaten Bener Meriah,” *JATI : Jurnal Mahasiswa Teknik Informatika* vol. 9, no. 5, pp. 8524-8529 Oktober 2025 doi: <https://doi.org/10.36040/jati.v9i5.15096>
- [13] F. Kurniawan and R. Herdiansyah, “Implementasi Algoritma K-Means Clustering dalam Pemetaan Wilayah Ideal untuk Berbagai Tanaman Perkebunan Berbasis Web,” *Ranah Research : Journal Of Multidisciplinary Research and Development*, vol. 8, no. 1, pp. 165-175, 2025, doi: 10.38035/rrj.v8i1
- [14] Mustika, Yunita Ardilla, Abraham Manuhutu, Nazaruddin Ahmad, “Data_Mining_Cetak,” Widina Bhakti Persada Bandung, ISBN:978-623-5811-14-7,2021, pp.iii-252, doi:<https://www.academia.edu/9608514>
- [15] A. Subayu, “Penerapan Metode K-Means Untuk Analisis Stunting Gizi Pada Balita: Systematic Review,” *Jurnal Sains, Nalar, dan Aplikasi Teknologi Informasi*, vol. 2, no. 1, pp. 42-50, Sep. 2022, doi: 10.20885/snati.v2i1.18.
- [16] Sekar Setyaningtyas, B. Indarmawan Nugroho, and Z. Arif, “Tinjauan Pustaka Sistematis: Penerapan Data Mining Teknik Clustering Algoritma K-Means,” *Jurnal Teknoif Teknik Informatika Institut Teknologi Padang*, vol. 10, no. 2, pp. 52–61, Oct. 2022, doi: 10.21063/jtif.2022.v10.2.52-61.
- [17] D. Selvia and A. Ramadhanu, “Implementasi Metode K-Means Clustering Dengan Teknik Pengolahan Citra Untuk Mengidentifikasi Jenis Sepatu,” *Jurnal Informatika Teknologi dan Sains*, vol. 7, no. 1, pp. 361–366, 2025, doi: <https://elibrary.ru/item.asp?id=81836722>
- [18] C. Selvi, D. Sembiring, L. Hanum, and S. Parsaoran Tamba, “Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Menentukan Judul Skripsi Dan Jurnal Penelitian (Studi Kasus Ftik Unpri),” *Jurnal Sistem Informasi dan Ilmu Komputer Prima*, vol. 5, no. 2, pp. 80-85 2022
- [19] N. A. Maori, “Metode Elbow Dalam Optimasi Jumlah Cluster Pada K-Means Clustering,” *Jurnal SIMETRIS*, vol. 14, no.2, pp. 277-287, 2023, doi: <https://pdfs.semanticscholar.org/818a/672dd55d9a0e9159eab09b02a5137edcb40.pdf>
- [20] N. N. Hasanah and A. S. Purnomo, “Implementasi Data Mining Untuk Pengelompokan Buku Menggunakan Algoritma K-Means Clustering (Studi Kasus : Perpustakaan Politeknik LPP Yogyakarta),” *Jurnal Teknologi Dan Sistem Informasi Bisnis*, vol. 4, no. 2, pp. 300–311, Jul. 2022, doi: 10.47233/jteksis.v4i2.499.
- [21] Sitti Rahmah, Muhjamil, Andi Aslindah, Aldi bas Fawait, and Yudhi Fajar Saputra, “Pengelompokan Hasil Pembelajaran Mahasiswa dengan Algoritma K-Means Clustering,” *Jurnal Pengabdian Masyarakat dan Riset Pendidikan*, vol. 3, no. 4, pp. 4215–4221, Jun. 2025, doi: 10.31004/jerkin.v3i4.1217.
- [22] T. Amalina, D. Bima, A. Pramana, and B. N. Sari, “Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food,” *Jurnal Ilmiah Wahana Pendidikan*, vol. 8, no. 15, pp. 574–583, 2022, doi: 10.5281/zenodo.7052276.
- [23] W. R. Nurputra, W. Witanti, and A. Komarudin, “Principal Component Analysis (PCA) Untuk Meningkatkan Hasil Klasterisasi Penjualan Video Game Menggunakan Algoritma K-Means” *Jurnal Locus: Penelitian dan Pengabdian*, vol. 4, no. 8, pp. 4683-4703, 2025, doi: <https://doi.org/10.58344/locus.v4i8.4151>
- [24] H. Almiatus Soleha, W. Pura Nurmawanti, U. Hidayaturrohman, R. Haiban Hirzi, A. Septiani, and J. Statistika, “Penerapan Clustering Time Series pada Pengelompokan Provinsi di Indonesia (Studi Kasus : Nilai Ekspor Non Migas di Indonesia Tahun 2016-2020),” vol. 15, no. 2, pp. 286-291, 2022. [Online]. doi: <https://doi.org/10.36456/jstat.a5550>