

Klasterisasi Pulau Penghasil Kelapa Sawit di Indonesia Berdasarkan Luas Areal, Produksi dan Tenaga Kerja Menggunakan Metode K-Means Clustering

Gulfi Oktariani

Manajemen, Universitas Bakrie
e-mail: gulfioktariani19@gmail.com

Abstrak

Kelapa sawit merupakan salah satu komoditas utama di Indonesia yang berkontribusi besar terhadap perekonomian nasional. Penelitian ini bertujuan untuk melakukan *clustering* Pulau penghasil kelapa sawit berdasarkan luas areal, produksi, dan tenaga kerja menggunakan metode K-Means untuk menganalisis data dan menggunakan bahasa pemrograman python. Hasil analisis menunjukkan bahwa terdapat jumlah *cluster* yang optimal sebanyak $k=2$, *cluster* 1 terdiri dari wilayah Sumatera dan Kalimantan yang memiliki luas areal terbesar, produksi tinggi, serta jumlah tenaga kerja yang tinggi. Sementara itu, cluster 2 mencakup wilayah dengan luas lahan yang lebih rendah seperti Jawa, Sulawesi, dan Maluku & Papua, cenderung memiliki produksi dan tenaga kerja lebih sedikit. Hasil penelitian ini, memberikan rekomendasi untuk wilayah perkebunan besar diperlukan peningkatan efisiensi produksi, infrastruktur, serta penerapan teknologi pertanian modern guna meningkatkan daya saing di pasar global. Sementara, untuk wilayah dengan luas lahan yang rendah perlu didorong untuk melakukan diversifikasi produk dan meningkatkan akses pasar bagi petani lokal.

Kata kunci: Kelapa Sawit, Luas Areal, Produksi, Tenaga Kerja, K-Means Clustering

Abstract

Palm oil is one of Indonesia's main commodities, contributing significantly to the national economy. This study aims to cluster palm oil-producing islands based on plantation area, production, and workforce using the K-Means method for data analysis and the Python programming language. The analysis results indicate that the optimal number of clusters is $k=2$. Cluster 1 consists of Sumatra and Kalimantan, which have the largest plantation areas, high production, and a large workforce. Meanwhile, Cluster 2 includes regions with smaller plantation areas, such as Java, Sulawesi, and Maluku & Papua, which tend to have lower production and fewer workers. The findings of this study provide recommendations for large plantation areas, where improving production efficiency, infrastructure, and the adoption of modern agricultural technology are necessary to enhance competitiveness in the global market. On the other hand, regions with smaller plantation areas should be encouraged to diversify their products and improve market access for local farmers.

Keywords : Palm Oil, Plantation Area, Production, Workforce, K-Means Clustering

PENDAHULUAN

Komoditas kelapa sawit telah memberikan kontribusi signifikan bagi perekonomian Indonesia. Menurut data dari (Badan Pusat Statistik, 2025), pertumbuhan ekonomi Indonesia pada tahun 2024 mencapai 5,02%, sedikit melambat dibandingkan 5,05% pada tahun 2023. Terdiri dari sektor Pertanian, Perikanan, dan Kehutanan, untuk sektor Pertanian yang mencakup produksi kelapa sawit memberikan kontribusi penting bagi pertumbuhan ekonomi Indonesia yang tumbuh positif sebesar 1,69%. Nilai ekonomi sektor perkelapasawitan dari hulu hingga hilir pada tahun 2023 mencapai lebih dari Rp750 Triliun, yang berkontribusi sekitar 3,5% terhadap PDB Nasional (Jayadi, 2024).

Luas areal kelapa sawit Indonesia tahun 2024 berdasarkan data dari Kementerian Pertanian sebesar 16.833.985 Ha, yang terdiri dari areal perkebunan sawit rakyat (PSR) sebesar 6.290.966 Ha (37,37%), perkebunan besar negara (PBN) sebesar 545.709 Ha (3,24%) dan perkebunan besar swasta (PBS) sebesar 8.598.982 Ha (51,08%), selain itu terdapat luasan yang akan dikonfirmasi (LAD) sebesar 1.398.328 Ha (8,31%). Sedangkan jumlah produksi kelapa sawit Indonesia tahun 2024 sebesar 47.474.604 Ton dalam bentuk minyak sawit (*crude palm oil*/CPO), dengan rincian produksi dari PSR sebesar 16.478.307 Ton (34,70%), produksi dari PBN sebesar 2.159.991 Ton (4,54%) dan produksi dari PBS sebesar 28.836.306 Ton (60,74%). Berdasarkan data produksi dan luas areal kelapa sawit, dapat dihitung bahwa produktivitas kelapa sawit Indonesia pada tahun 2024 sebesar 3,82 Ton/Ha, dengan rincian produktivitas untuk PSR sebesar 3,27 Ton/Ha, produktivitas untuk PBN sebesar 4,40 Ton/Ha dan produktivitas untuk PBS sebesar 3,80 Ton/Ha (Direktorat Jenderal Perkebunan, 2024).

Perkebunan Kelapa Sawit di Indonesia tersebar di 29 provinsi, pada tahun 2024 provinsi penghasil kelapa sawit terbesar yaitu provinsi Riau yang mencapai luas areal sebesar 3,49 Juta Ha atau sekitar 20,76% dari total luas areal perkebunan kelapa sawit di Indonesia. Sementara Pulau dengan luas areal perkebunan kelapa sawit terbesar yaitu Pulau Sumatera mencapai 10,20 Juta Ha dan Pulau Kalimantan mencapai 5,98 Juta Ha (Direktorat Jenderal Perkebunan, 2024). Dari tujuh wilayah pulau-kepulauan yang ada di Indonesia hanya Pulau Bali dan Nusa Tenggara yang tidak memiliki perkebunan sawit eksisting. Hal ini disebabkan oleh karakteristik ekosistem di Kepulauan Bali dan Nusa Tenggara yang didominasi oleh pegunungan denudasional bermaterial campuran batuan beku luar dan batuan piroklastik vegetasi padang rumput monsun pamah, ekosistem ini memiliki lapisan tanah yang sangat tipis sehingga sawit tidak dapat tumbuh di pulau ini (Madani Berkelanjutan, 2024).

Kelapa sawit merupakan salah satu komoditas pertanian yang berkontribusi besar dalam penerimaan devisa negara, tahun 2023 menyumbang sebesar USD 25,61 miliar dengan volume ekspor sebesar 38,23 juta ton, neraca volume perdagangan mengalami peningkatan sebesar 4,68% (Pusat Data dan Sistem Informasi Pertanian Sekretariat Jenderal Kementerian Pertanian, 2024). Berdasarkan data (Badan Pusat Statistik, 2024) dari sisi penyerapan tenaga kerja di sektor pertanian, kehutanan dan perikanan dari total tenaga kerja di Indonesia sebesar 28,64%. Komoditas kelapa sawit sangat penting bagi pembangunan nasional perkebunan kelapa sawit yang memiliki pengaruh bagi perekonomian masyarakat dan dapat menyerap lapangan kerja yang lebih besar.

Pertumbuhan industri kelapa sawit dipengaruhi oleh berbagai faktor, di antaranya luas areal perkebunan, produksi, dan tenaga kerja yang terlibat dalam proses produksi. Meskipun demikian, distribusi produksi kelapa sawit di setiap Pulau tidak merata, sehingga perlu dilakukan analisis lebih lanjut untuk memahami pola penyebarannya. Dalam upaya memahami perbedaan karakteristik Pulau penghasil kelapa sawit, digunakan metode *clustering* untuk mengelompokkan Pulau berdasarkan faktor-faktor utama yang mempengaruhi industri ini. Metode K-Means *Clustering* merupakan salah satu metode yang dapat diimplementasikan untuk mengelompokkan berdasarkan Pulau yang memiliki karakteristik atau kesamaan dalam luas areal, produksi, dan tenaga kerja. Dengan adanya klasterisasi ini, pemerintah dan pelaku industri dapat mengambil keputusan yang lebih tepat dalam pengelolaan dan pengembangan perkebunan kelapa sawit di Indonesia.

Dari latar belakang yang telah dijabarkan diatas, tujuan dari penelitian ini yaitu melakukan *clustering* Pulau penghasil kelapa sawit di Indonesia berdasarkan luas areal, produksi, dan tenaga kerja menggunakan metode K-Means *Clustering*, menentukan jumlah *cluster* optimal yang dapat merepresentasikan perbedaan karakteristik dari luas areal, produksi dan tenaga kerja perkebunan kelapa sawit. Dengan adanya pengelompokkan ini diharapkan dapat memberikan informasi dan rekomendasi yang dapat digunakan oleh pemerintah atau pelaku industri dalam pengelolaan perkebunan kelapa sawit. Penelitian ini diharapkan dapat digunakan untuk membantu membuat kebijakan yang lebih efektif untuk mendukung pertumbuhan komoditas kelapa sawit yang berkelanjutan di seluruh wilayah Indonesia.

METODE

Data yang digunakan dalam penelitian ini dikumpulkan dari Buku Statistik Perkebunan dari tahun 2013-2024, Kementerian Pertanian tentang Perkebunan Kelapa Sawit yang mencakup data berdasarkan Pulau di Indonesia yang terdiri dari Sumatera, Jawa, Kalimantan, Sulawesi dan Maluku & Papua, luas areal, produksi dan tenaga kerja yang terlibat dalam perkebunan kelapa sawit. Periode data yang digunakan selama 11 tahun untuk mendapatkan pola pertumbuhan dari seluruh Pulau di Indonesia. Dalam melakukan analisis data terdapat beberapa proses atau tahapan yang dilakukan agar dalam proses pengolahan data penelitian dapat dimengerti dan mudah dipahami. Adapun rancangan blok diagram penelitian dapat dilihat pada gambar 1.



Gambar 1. Rancangan blok diagram penelitian

Pada tahapan pertama yang akan dilakukan yaitu tahap pengumpulan data, data yang digunakan dalam penelitian ini menggunakan data dari Buku Statistik Perkebunan Kementerian Pertanian terkait perkebunan kelapa sawit di Indonesia dari tahun 2013-2024 yang memiliki jumlah data sebanyak 60 data.

Tahap kedua melakukan *data preprocessing* untuk mempermudah proses pengolahan data seperti *data cleaning* dan *data normalization*. Pada proses *data cleaning* akan melakukan pembersihan data yang memiliki nilai *null* atau kosong, dan melakukan pemeriksaan terhadap *outliers* pada data yang digunakan, kemudian melakukan penghapusan *outliers* yang dapat memengaruhi keseimbangan data. Selanjutnya tahap *data normalization* akan melakukan proses *standardscaler* yang digunakan untuk melakukan standarisasi terhadap nilai dari variabel luas areal, produksi dan tenaga kerja untuk memastikan setiap variabel yang digunakan memiliki skala yang sebanding. Tahap visualisasi data menggunakan *Exploratory Data Analysis (EDA)*, metode EDA ini merupakan bagian dari proses sebelum melakukan pemodelan, tujuannya untuk melakukan visualisasi data pada variabel luas areal, produksi dan tenaga kerja.

Pada tahap pemodelan atau pemilihan metode, peneliti menggunakan metode *K-Means clustering*, untuk menentukan jumlah *cluster* yang optimal menggunakan metode *Elbow* untuk melihat titik optimal dari jumlah *cluster* berdasarkan *Within-Cluster Sum of Squares (WCSS)*. Dalam pengelompokan Pulau berdasarkan kemiripan data dalam luas areal, produksi dan tenaga kerja yang akan dilakukan proses *clustering* untuk menemukan hubungan antara ketiga variabel yang digunakan. Pada tahapan ini akan menginterpretasikan hasil *clustering* untuk memahami karakteristik masing-masing kelompok Pulau penghasil sawit di Indonesia.

Metode K-Means merupakan algoritma yang mengklasterisasikan berdasarkan partisi dan melakukan *clustering* melalui proses iterasi berkelanjutan sampai dengan bertemu kondisi akhir, proses iterasi berhenti dan hasil *clustering* adalah *output* (Dwiki Aulia Fakhri, 2021). Metode ini membagi data kedalam kelompok sehingga data yang berkarakteristik sama dimasukkan ke dalam satu kelompok sementara data yang berkarakteristik berbeda dimasukkan dalam kelompok yang lain (Nayla Nur Afidaha, 2023). Algoritma ini bekerja dengan cara menentukan pusat *cluster* secara acak, menghitung jarak data terhadap pusat *cluster*, dan memperbarui posisi pusat *cluster* hingga tidak ada perubahan signifikan dalam iterasi. K-Means sering digunakan dalam analisis eksploratif untuk mengidentifikasi pola dalam data yang besar.

Menurut (Wakhidah, 2017) langkah-langkah melakukan *clustering* dengan algoritma K-Means adalah sebagai berikut :

- 1) Menentukan jumlah *cluster* (*k*);
- 2) Menentukan nilai *centroid*;

Dalam menentukan nilai *centroid* untuk awal iterasi, nilai awal *centroid* dilakukan secara acak. Sedangkan jika menentukan nilai *centroid* yang merupakan tahap dari iterasi, maka digunakan rumus sebagai berikut.

$$v_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} x_{kj}$$

dimana :

v_{ij} : centroid rata-rata *cluster* ke- i untuk variabel ke- j

N_i : jumlah data yang menjadi anggota *cluster* ke- i

i, k : indeks dari *cluster*

j : indeks dari variabel

x_{kj} : nilai data ke- k yang ada di dalam *cluster* tersebut untuk variable ke- j

- 3) Menghitung jarak antara titik *centroid* dengan titik tiap objek, untuk menghitung jarak tersebut dapat menggunakan *Euclidean Distance*, yaitu:

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

dimana :

D_e : *Euclidean Distance*

i : banyaknya objek

(x, y) : koordinat objek

(s, t) : koordinat *centroid*

- 4) Pengelompokkan objek dengan memperhitungkan jarak minimum objek.
5) Kembali ke tahap 2, lakukan perulangan hingga nilai *centroid* yang dihasilkan tetap dan anggota *cluster* tidak berpindah ke *cluster* lain.

K-Means *Clustering* termasuk dalam teknik non-hierarki yang membagi data menjadi satu atau lebih *cluster*, dengan data yang sama dikelompokkan bersama dalam satu *cluster* dan data yang beragam dibagi menjadi beberapa *cluster*. Karena mudah digunakan, algoritma ini adalah yang paling banyak digunakan. Kekurangan dari algoritma ini adalah sangat bergantung pada inialisasi *cluster* (Tifani Amalina, 2022).

HASIL DAN PEMBAHASAN

Pengumpulan data dilakukan dengan menghimpun informasi melalui Buku Statistik Perkebunan dari Direktorat Jenderal Perkebunan, Kementerian Pertanian yang terdiri dari data tahun, pulau, luas areal, produksi dan tenaga kerja dari tahun 2013-2024, data ini terdiri dari Pulau Sumatera, Jawa, Bali & Nusa Tenggara, Kalimantan, Sulawesi, dan Maluku & Papua, akan tetapi untuk Pulau Bali & Nusa Tenggara tidak memiliki perkebunan kelapa sawit sehingga hanya terdapat data dari 6 pulau di Indonesia yaitu dengan total dataset sebanyak 60 data/baris.

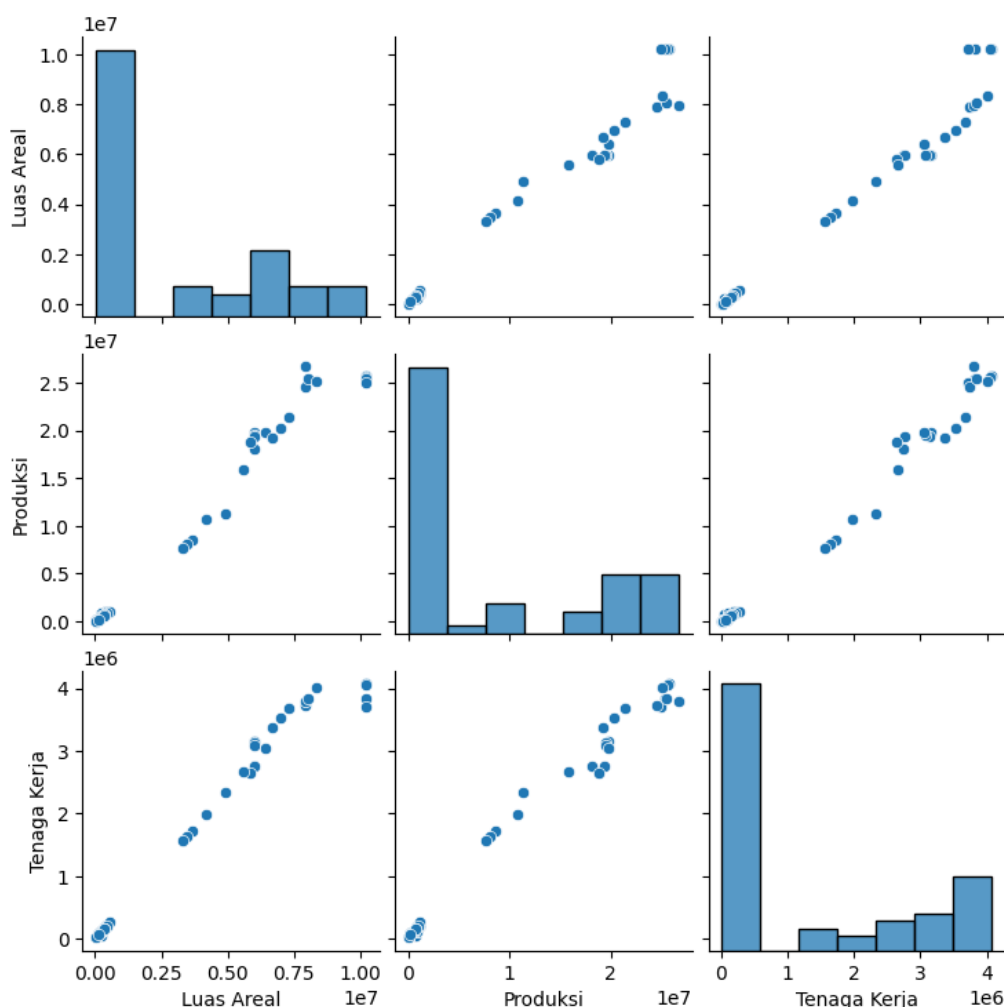
Proses pengolahan data dilakukan dengan tahap *data preprocessing* yaitu proses data cleaning untuk melakukan pembersihan data yang memiliki nilai *null* atau kosong pada dataset ini tidak ada data yang memiliki nilai *null* dan tidak ada data yang *outliers*. Selanjutnya tahap *data normalization* menggunakan *standardscaler* untuk melakukan standarisasi terhadap nilai dari variabel luas areal, produksi dan tenaga kerja untuk memastikan setiap variabel yang digunakan memiliki skala yang sebanding.

Tabel 1. Hasil tahap standardscaler

Indeks	Luas Areal (Ha)	Produksi (Ton)	Tenaga Kerja
0	2.142458	1.788141	1.801455
1	-0.806546	-0.797654	-0.829189
2	0.919001	1.190592	1.201145

Indeks	Luas Areal (Ha)	Produksi (Ton)	Tenaga Kerja
....			
59	-0.779967	-0.789227	-0.796103

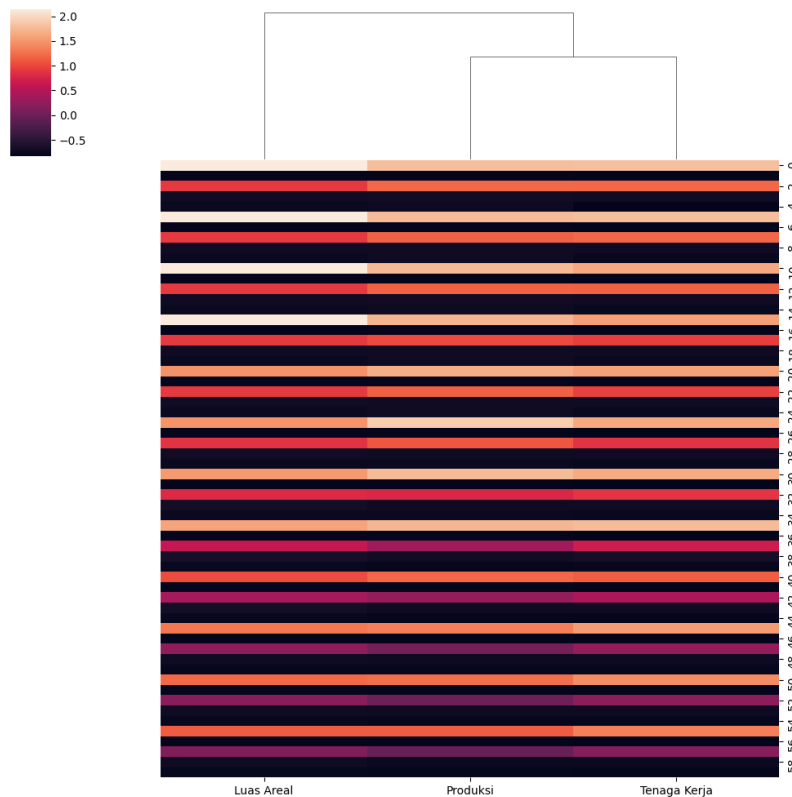
Tahap visualisasi data menggunakan *Exploratory Data Analysis* (EDA). Metode EDA ini merupakan bagian dari proses sebelum melakukan pemodelan. EDA dapat dikatakan sebagai proses menganalisis dataset untuk mengenali dan menyimpulkan karakteristik dari data yang digunakan. Pendekatan yang digunakan pada proses ini yaitu menggunakan *bivariate analysis* untuk mengetahui bagaimana korelasi antara variabel dan menemukan tren serta pola dalam data. Selain itu, berbagai teknik visualisasi seperti heatmap, scatter plot, dan bar chart digunakan untuk memberikan pemahaman yang lebih mendalam mengenai distribusi data serta hubungan antarvariabel, sehingga dapat membantu dalam pengambilan keputusan dan perumusan strategi berbasis data.



Gambar 2. Pair Plot hubungan antara 3 variabel

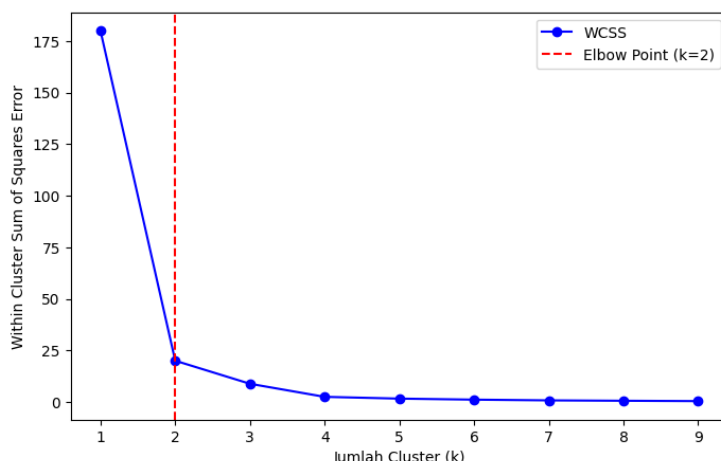
Gambar 2 merupakan hasil dari pair plot yang menggambarkan hubungan antar variabel. Dapat dilihat bahwa semakin besar luas areal perkebunan kelapa sawit, maka produksi yang dihasilkan juga semakin tinggi serta membutuhkan tenaga kerja yang lebih banyak. Sementara itu, peningkatan hasil produksi kelapa sawit juga berbanding lurus dengan tenaga kerja yang

dibutuhkan untuk mendukung proses budidaya, panen, dan pengolahan hasil. Hal ini menunjukkan adanya hubungan positif yang kuat antara luas areal, produksi, dan tenaga kerja dalam industri kelapa sawit, yang mengindikasikan bahwa peningkatan skala perkebunan dapat berdampak langsung terhadap peningkatan produksi dan penyerapan tenaga kerja.



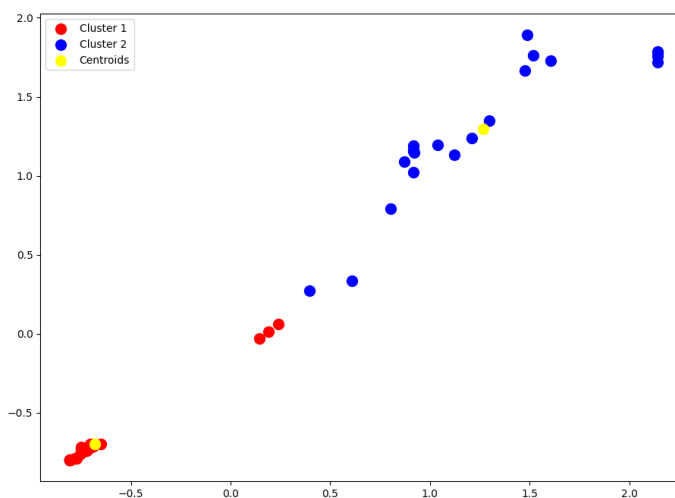
Gambar 3. Heatmap korelasi antara 3 variabel

Berdasarkan heatmap yang dihasilkan pada Gambar 3, terdapat beberapa pola menarik yang dapat dianalisis terkait hubungan antara luas areal, produksi, dan tenaga kerja di sektor kelapa sawit di Indonesia. Diketahui bahwa warna yang mendekati angka 2 berwarna merah muda berarti memiliki hubungan yang kuat, sedangkan yang mendekati -0,5 berwarna ungu tua berarti memiliki hubungan lemah atau negatif. Secara umum, Pulau dengan luas areal yang lebih besar cenderung memiliki produksi yang tinggi, yang menunjukkan korelasi positif antara kedua variabel ini. Semakin luas perkebunan, semakin besar pula hasil produksi yang diperoleh. Selain itu, produksi kelapa sawit juga berkaitan dengan tenaga kerja, di mana Pulau dengan produksi tinggi umumnya membutuhkan lebih banyak tenaga kerja. Namun, ada beberapa Pulau yang memiliki produksi besar tetapi tenaga kerja yang tidak terlalu banyak, yang kemungkinan disebabkan oleh mekanisasi dalam proses produksi.



Gambar 4. Jumlah cluster metode elbow

Setelah memahami korelasi antara variabel, proses selanjutnya menentukan jumlah *cluster* menggunakan metode elbow berdasarkan nilai SSE (*Sum of Square Error*). Semakin besar nilai SEE maka semakin rendah kualitas *cluster*, sementara semakin kecil nilai SEE maka kualitas *cluster* semakin baik. Berdasarkan hasil uji *cluster* pada gambar 4 menunjukkan bahwa jumlah *cluster* $k=1$ memiliki nilai SEE yang paling tinggi, ketika jumlah *cluster* $k=2$ nilai SEE mengalami penurunan yang signifikan, pada *cluster* $k=3$ dan seterusnya hingga *cluster* $k=9$ kembali mengalami penurunan. Jumlah *cluster* yang membentuk siku terlihat jelas pada jumlah *cluster* $k=2$ merupakan jumlah *cluster* optimal untuk digunakan pada data ini.



Gambar 5. Hasil metode K-Means menggunakan jumlah cluster k=2

Gambar 5 merupakan hasil dari *cluster* $k=2$ menggunakan metode K-Means *clustering*, untuk persebaran objek atau titik-titik berwarna merah merupakan *cluster* $k=1$, sedangkan yang berwarna biru merupakan hasil *cluster* $k=2$. Terdapat titik berwarna kuning merupakan nilai *centroid* yang merupakan titik pusat data dalam proses *clustering*.

Tabel 1. Hasil Cluster k=1

Tahun	Pulau	Luas Areal (Ha)	Produksi(Ton)	Tenaga Kerja
2024	Sumatera	10.201.658	25.751.247	4.079.166
2024	Kalimantan	5.982.568	19.818.500	3.152.639
2023	Sumatera	10.201.658	25.688.752	4.058.905
2023	Kalimantan	5.982.568	19.468.048	3.141.172

Tahun	Pulau	Luas Areal (Ha)	Produksi(Ton)	Tenaga Kerja
2022	Sumatera	10.201.658	25.453.682	3.832.077
2022	Kalimantan	5.982.568	19.516.365	3.079.620
2021	Sumatera	10.201.658	25.039.330	3.713.502
2021	Kalimantan	5.982.568	18.141.629	2.752.621
2020	Sumatera	7.907.810	24.555.698	3.731.436
2020	Kalimantan	5.990.789	19.377.282	2.764.260
2019	Sumatera	7.944.520	26.765.541	3.802.812
2019	Kalimantan	5.820.406	18.835.810	2.647.545
2018	Sumatera	8.047.920	25.467.966	3.835.419
2018	Kalimantan	5.588.075	15.872.812	2.665.377
2017	Sumatera	8.355.034	25.152.448	4.014.548
2017	Kalimantan	4.919.027	11.309.484	2.335.615
2016	Sumatera	6.393.364	19.856.284	3.046.528
2015	Sumatera	7.285.853	21.387.654	3.677.583
2014	Sumatera	6.985.723	20.297.100	3.526.145
2013	Sumatera	6.682.228	19.232.296	3.372.959

Hasil dari persebaran masing-masing *cluster* kemudian akan dilihat objek atau data yang termasuk kedalam *cluster* $k=1$. Tabel 1 merupakan hasil dari *cluster* $k=1$, dari hasil menunjukkan bahwa yang tergabung dalam *cluster* ini adalah Pulau Sumatera dan Kalimantan yang memiliki jumlah luas areal terbesar di antara 6 Pulau penghasil sawit di Indonesia, dengan jumlah produksi paling rendah 11,30 juta ton dan paling tinggi 26,76 juta ton, dan memiliki rata-rata tenaga kerja sebanyak 3,36 juta orang.

Tabel 2. Hasil Cluster $k=2$

Tahun	Pulau	Luas Areal (Ha)	Produksi (Ton)	Tenaga Kerja
2024	Jawa	32.020	78.266	18.990
2024	Sulawesi	390.008	1.065.911	204.032
2024	Maluku & Papua	227.731	760.681	47.266
2023	Jawa	32.020	79.701	18.922
2023	Sulawesi	390.008	1.011.075	209.224
2023	Maluku & Papua	227.731	836.723	118.491
2022	Jawa	32.020	62.507	19.337
2022	Sulawesi	390.008	1.029.638	204.105
2022	Maluku & Papua	227.731	757.481	118.500
2021	Jawa	32.020	59.445	18.703
2021	Sulawesi	390.008	1.012.997	202.492
2021	Maluku & Papua	227.731	868.080	131.291
2020	Jawa	32.982	56.243	18.216
2020	Sulawesi	416.626	921.041	194.599
2020	Maluku & Papua	238.391	831.581	121.504
2019	Jawa	32.958	63.487	18.587
2019	Sulawesi	418.846	896.614	199.627
2019	Maluku & Papua	239.880	558.795	120.566
2018	Jawa	35.042	84.430	19.435
2018	Sulawesi	436.061	991.590	198.639
2018	Maluku & Papua	219.252	466.833	110.188
2017	Jawa	32.976	58.675	18.669
2017	Sulawesi	569.535	1.077.878	271.089
2017	Maluku & Papua	154.002	214.143	91.902
2016	Jawa	36.742	60.294	20.530

Tahun	Pulau	Luas Areal (Ha)	Produksi (Ton)	Tenaga Kerja
2016	Kalimantan	4.177.368	10.729.133	1.978.347
2016	Sulawesi	439.224	921.359	209.086
2016	Maluku & Papua	154.767	163.891	87.382
2015	Jawa	36.163	63.617	22.216
2015	Kalimantan	3.638.024	8.590.871	1.722.931
2015	Sulawesi	349.464	727.309	171.121
2015	Maluku & Papua	135.303	179.481	76.592
2014	Jawa	34.935	61.671	21.438
2014	Kalimantan	3.471.843	8.127.398	1.643.651
2014	Sulawesi	334.235	687.611	163.652
2014	Maluku & Papua	129.494	170.699	73.334
2013	Jawa	33.712	59.720	20.654
2013	Kalimantan	3.306.523	7.680.148	1.564.956
2013	Sulawesi	318.880	647.907	156.123
2013	Maluku & Papua	123.677	161.932	70.055

Hasil dari persebaran *cluster* $k=2$ pada tabel 2, menunjukkan bahwa *cluster* ini terdiri dari Pulau Jawa, Sulawesi, dan Maluku & Papua yang memiliki luas lahan rata-rata hanya 553 ribu ha, dengan jumlah produksi paling rendah 56 ribu ton dan paling tinggi 1,07 juta ton, dan hanya memiliki rata-rata tenaga kerja sebanyak 105 ribu orang. Sementara terdapat 4 data dari Pulau Kalimantan pada tahun 2013-2016 masuk kedalam *cluster* ini karena memiliki luas areal yang hanya 3,30 - 4,17 juta ha, dengan jumlah produksi 7,68 - 10,7 juta ton dan hanya memiliki jumlah tenaga kerja paling tinggi sebanyak 1,97 juta orang.

SIMPULAN

Berdasarkan dari hasil analisis data perkebunan kelapa sawit dari tahun 2013-2024 menggunakan metode K-Means *clustering* menunjukkan bahwa jumlah *cluster* yang optimal yaitu menggunakan $k=2$ berdasarkan luas areal, produksi dan tenaga kerja, dengan hasil dari *cluster* 1 sebanyak 20 data yang terdiri dari Pulau Sumatera dan Kalimantan yang merupakan pusat utama industri kelapa sawit dengan luas perkebunan yang sangat besar, produksi tinggi, serta penyerapan tenaga kerja yang mencapai jutaan orang. Untuk wilayah ini, diperlukan peningkatan infrastruktur dan efisiensi produksi guna mempercepat distribusi hasil perkebunan serta menekan biaya operasional. Selain itu, penerapan teknologi pertanian modern dan sistem irigasi yang lebih baik perlu didorong untuk meningkatkan produktivitas tanpa harus memperluas lahan. Standarisasi dan sertifikasi internasional seperti ISPO harus diperkuat guna meningkatkan daya saing di pasar global, sementara kesejahteraan tenaga kerja perlu ditingkatkan melalui pelatihan dan fasilitas pendukung.

Cluster 2 terdiri dari 40 data yang mencakup Pulau Jawa, Sulawesi, Maluku & Papua memiliki skala perkebunan yang lebih kecil, produksi yang lebih rendah, serta tenaga kerja yang lebih sedikit, namun tetap menunjukkan potensi pertumbuhan, terutama di Sulawesi dan Maluku & Papua. Pada wilayah ini, strategi yang dapat diterapkan adalah meningkatkan skala usaha serta mendorong diversifikasi produk berbasis kelapa sawit. Produktivitas lahan juga harus ditingkatkan melalui penerapan teknik pertanian presisi dan penggunaan benih unggul. Selain itu, pendampingan dan pelatihan bagi petani lokal menjadi langkah penting untuk memastikan mereka memiliki keterampilan dalam mengelola perkebunan secara berkelanjutan. Infrastruktur dan akses pasar juga perlu diperkuat agar petani dan pelaku usaha kecil dapat lebih mudah memasarkan hasil produksinya. Pada *cluster* ini juga mencakup Pulau Kalimantan dari tahun 2013-2016 yang memiliki luas areal, produksi, dan tenaga kerja yang tidak terlalu tinggi di bandingkan tahun 2017-2024, sehingga data ini tergolong dalam *cluster* 2.

Hasil analisis *clustering* ini dapat menjadi dasar dalam perencanaan kebijakan berbasis data yang lebih tepat sasaran. Pemerintah perlu memperkuat regulasi dan pengawasan terkait tata kelola lahan, lingkungan, serta tenaga kerja guna mencegah eksploitasi sumber daya alam yang

berlebihan. Secara keseluruhan, industri kelapa sawit di Indonesia masih sangat bergantung pada Sumatera dan Kalimantan, yang menjadi pusat produksi terbesar, sementara wilayah lain berperan dalam skala lebih kecil namun tetap berkontribusi terhadap pertumbuhan industri. Dengan tren pertumbuhan yang ada, masih terdapat peluang untuk pengembangan industri kelapa sawit di wilayah Sulawesi dan Maluku & Papua di masa mendatang.

DAFTAR PUSTAKA

- Tifani Amalina, D. B. (2022). Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food. *Jurnal Ilmiah Wahana Pendidikan*, 574-583.
- Dwiki Aulia Fakhri, S. D. (2021). Optimalisasi Pelayanan Perpustakaan terhadap Minat Baca Menggunakan Metode K-Means Clustering. *Jurnal Informasi dan Teknologi*, 160-166.
- Nayla Nur Afidaha, M. (2023). Penerapan Metode Clustering dengan Algoritma K-Means untuk Pengelompokan Data Migrasi Penduduk Tiap Kecamatan di Kabupaten Rembang. *PRISMA, Prosiding Seminar Nasional Matematika*, 729-738.
- Wakhidah, N. (2017). Clustering Menggunakan K-Means Algorithm (K-Means Algorithm Clustering). *Fakultas Teknologi Informasi dan Komunikasi Universitas Semarang*.
- Direktorat Jenderal Perkebunan. (2024). *Buku Statistik Perkebunan 2023-2025*. Jakarta: Direktorat Jenderal Perkebunan.
- Badan Pusat Statistik. (2025, 2 5). *Ekonomi Indonesia Tahun 2024 Tumbuh 5,03 Persen (C-to-C). Ekonomi Indonesia Triwulan IV-2024 Tumbuh 5,02 Persen (Y-on-Y). Ekonomi Indonesia Triwulan IV-2024 Tumbuh 0,53 Persen (Q-to-Q)*. Retrieved from Badan Pusat Statistik.
- Badan Pusat Statistik. (2024). *Keadaan Angkatan Kerja Di Indonesia Februari 2024*.
- Jayadi, F. (2024, 9 12). *Nilai Ekonomi Sektor Perkelapasawitan Capai Rp750 Triliun di 2023*. Retrieved from emiten news.com.
- Madani Berkelanjutan. (2024). *Kertas Kebijakan Urgensi Perbaikan Tata Kelola Sawit melalui Kebijakan Penghentian Pemberian Izin dalam Perspektif Ekonomi dan Daya Dukung Daya Tampung Lingkungan Hidup (D3TLH)*.
- Pusat Data dan Sistem Informasi Pertanian Sekretariat Jenderal Kementerian Pertanian. (2024). *Analisis Kinerja Perdagangan Komoditas Kelapa Sawit 2024*. Jakarta: Pusat Data dan Sistem Informasi Pertanian Sekretariat Jenderal Kementerian Pertanian.