

## ANALISIS PENGELOMPOKKAN KEBUN KELAPA SAWIT PTPN IV REGIONAL 4 BERDASARKAN PRODUKTIVITAS TANDAN BUAH SEGAR MENGUNAKAN METODE *K-MEANS*

Aghnina Aulia Rahma<sup>1,\*</sup>, Gusmanely. Z<sup>2</sup>

<sup>1,2</sup>Program Studi Matematika, FST, Universitas Jambi

\*Corresponding author: [aghninaauliarahma@gmail.com](mailto:aghninaauliarahma@gmail.com)

### ABSTRACT

*This study aims to classify the oil palm plantations of PT Perkebunan Nusantara IV Regional 4 based on the productivity of Fresh Fruit Bunches using the K-Means Clustering method. The data used consist of 12 oil palm plantations during the period of January to June 2025, with five main variables: land area, number of productive trees, number of bunches, average bunch weight, and total FFB production. Before the analysis was conducted, the data were standardized using the Z-Score method, and the optimal number of clusters was determined using the Elbow method. The results show that the plantations can be grouped into three clusters based on FFB productivity, namely high, medium, and low productivity clusters. This clustering provides useful information for company management in evaluating the performance of each plantation and formulating strategies to improve productivity in the future*

**Keywords:** Cluster, Fresh Fruit Bunches, Oil Palm, Productivity, K-Means

### ABSTRAK

Penelitian ini bertujuan untuk mengelompokkan kebun kelapa sawit PT Perkebunan Nusantara IV Regional 4 berdasarkan produktivitas Tandan Buah Segar menggunakan metode *K-Means Clustering*. Data yang digunakan adalah 12 kebun kelapa sawit selama periode Januari hingga Juni 2025 dengan lima variabel utama yaitu, luas lahan, jumlah pohon produksi, jumlah tandan, rata-rata berat tandan dan total produksi Tandan Buah Segar. Sebelum dilakukan analisis, data distandarisasi menggunakan metode *Z-Score* dan jumlah *cluster* ditentukan menggunakan metode *Elbow*. Hasil analisis menunjukkan bahwa kebun-kebun dapat dikelompokkan menjadi tiga *cluster* berdasarkan produktivitas Tandan Buah Segar, yaitu *cluster* dengan produktivitas tinggi, sedang, dan rendah. Pengelompokkan ini memberikan informasi yang berguna bagi manajemen perusahaan dalam mengevaluasi kinerja setiap kebun serta merumuskan strategi peningkatan produktivitas di masa mendatang.

**Kata Kunci:** Klaster, Tandan Buah Segar, Kelapa Sawit, Produktivitas, *K-Means*

---

#### ARTICLE INFO

Submission received: 10 November 2025

Accepted: 29 December 2025

Revised: 15 December 2025

Published: 31 December 2025

Available on: <https://doi.org/10.32493/sm.v7i3.54618>

---

**StatMat: Jurnal Statistika dan Matematika** is licenced under a Creative Commons Attribution-ShareAlike 4.0 International License.

---

## 1. PENDAHULUAN

Kelapa sawit merupakan salah satu komoditas unggulan dan utama di Indonesia yang berperan penting dalam sektor perkebunan dan perekonomian nasional. Produk utama yang dihasilkan berupa minyak sawit mentah atau CPO (*Crude Palm Oil*) dan minyak inti sawit

atau PKO (*Palm Kernel Oil*) yang memiliki nilai ekonomi yang tinggi dan menjadi salah satu penyumbang devisa negara terbesar dibandingkan dengan komoditas perkebunan yang lain. Salah satu komponen utama dalam proses produksi minyak sawit adalah Tandan Buah Segar yang mencerminkan efisiensi dan keberhasilan manajemen budidaya (Lubis & Lubis, 2018). Produktivitas Tandan Buah Segar dipengaruhi oleh beberapa faktor, yang secara umum dapat dibagi menjadi tiga kategori yaitu faktor genetik, faktor lingkungan, dan faktor teknik budidaya. Faktor genetik mencakup jenis bibit unggul dan umur tanaman yang menjadi kunci untuk menentukan potensi hasil dasar. Faktor lingkungan, seperti keadaan tanah, luas lahan, bentuk lahan, dan curah hujan, menciptakan kondisi pertumbuhan yang optimal bagi tanaman. Di sisi lain, faktor teknik budidaya termasuk pemupukan, pengendalian hama dan gulma, serta perawatan lainnya memegang peranan penting dalam mengoptimalkan potensi yang sudah ada (Ginting *et al.*, 2022).

PT Perkebunan Nusantara IV Regional 4 merupakan bagian dari Holding perusahaan perkebunan negara (PalmCo) yang mengelola 12 kebun kelapa sawit di wilayah Jambi dan Sumatra Barat. Dalam operasionalnya, ditemui kondisi dimana sebagian kebun memiliki jumlah pohon yang banyak namun menghasilkan tandan dengan berat rata-rata yang rendah, dan terdapat juga kebun dengan rata-rata berat tandan yang tinggi meski jumlah tandan relatif sedikit. Variasi ini pada akhirnya berdampak pada tingkat produktivitas Tandan Buah Segar. Penelitian yang dilakukan oleh Suhartono (2023) menunjukkan bahwa kepadatan tanaman dapat mempengaruhi produktivitas Tandan Buah Segar serta jumlah tandan dan rata-rata berat tandan adalah faktor penting dalam menentukan hasil panen kelapa sawit. Selain itu, luas lahan tidak selalu berbanding lurus dengan tingkat produktivitas karena efisiensi pengelolaan kebun juga ikut menentukan hasil (Yosephine *et al.*, 2025). Keadaan yang tidak konsisten ini menghasilkan variasi dalam tingkat produktivitas Tandan Buah Segar, sehingga menyulitkan manajemen dalam menilai kinerja dan merencanakan strategi yang tepat. Untuk memahami variasi tersebut secara lebih terstruktur, diperlukan analisis yang mampu mengelompokkan kebun kelapa sawit berdasarkan kesamaan karakteristik produktivitasnya, sehingga manajemen perusahaan dapat lebih mudah dalam menyusun strategi peningkatan produktivitas.

Untuk memahami variasi tersebut, diperlukan analisis yang mampu mengelompokkan kebun kelapa sawit berdasarkan kemiripan karakteristik produktivitas Tandan Buah Segar. Salah satu metode yang dapat digunakan untuk mengelompokkan data kuantitatif adalah

metode K-Means (Fauziah, & Basir, C., 2024). Metode ini merupakan metode pengelompokan data (*clustering*) yang bekerja dengan cara mengevaluasi setiap data dan mengelompokkan kedalam salah satu cluster berdasarkan jarak terdekat terhadap pusat *centroidi* (Permata Sari, Y., & Basir, C., 2025). Metode K-Means dianggap paling tepat untuk penelitian ini karena dinilai sederhana, efisien dan mampu memberikan hasil pengelompokkan yang mudah di pahami. K-Means dapat mengelompokkan kebun dengan jelas ke dalam kategori produktivitas tertentu. Selain itu, metode K-Means sering dipakai dalam penelitian di sektor pertanian, sehingga relevan untuk mengidentifikasi pola produktivitas kebun sawit berdasarkan faktor-faktor yang dianalisis (Asroni & Adrian, 2015).

## 2. METODE PENELITIAN

### 2.1 Pengumpulan Data

Data yang diperoleh dalam penelitian ini berasal dari data Internal perusahaan PT Perkebunan Nusantara IV Regional 4. Data memuat informasi mengenai produksi Tandan Buah Segar dari seluruh kebun kelapa sawit di PT Perkebunan Nasional IV Regional 4 selama periode Januari hingga Juni 2025. PT Perkebunan Nusantara IV Regional 4 terdiri atas 12 kebun. Variabel yang digunakan pada penelitian ini didefinisikan pada tabel 1.

**Tabel 1. Definisi Variabel**

Notasi Variabel	Definisi Variabel	Satuan
$x_1$	Luas Lahan	Hektare
$x_2$	Jumlah Pohon Produksi	Pohon
$x_3$	Jumlah Tandan	Tandan
$x_4$	Rata-Rata Berat Tandan	Kilogram
$x_5$	Hasil Produksi TBS	Kilogram

### 2.2 Analisis Data

Langkah-langkah dalam menjalankan metode *K-Means Clustering* adalah sebagai berikut:

#### 1. Standarisasi Data

Proses standarisasi perlu dilakukan jika terdapat perbedaan signifikan dalam ukuran satuan di antara variabel-variabel yang sedang diteliti. Perbedaan satuan yang signifikan dapat membuat perhitungan dalam analisis *cluster* menjadi tidak akurat. Standarisasi dilakukan dengan menggunakan metode *Z-Score*, dengan tujuan menyamakan skala antar

variabel sehingga seluruh variabel memiliki rata-rata nol dan standar deviasi satu. (Suparwito, 2024).

## 2. Pemeriksaan Asumsi

Sebelum melakukan analisis *cluster* dengan menggunakan metode *K-Means* terdapat 2 asumsi yang perlu dipenuhi terlebih dahulu yaitu:

### a. Sampel Representatif

Sebelum melakukan analisis pengelompokan, dilakukan pengujian kelayakan sampel terlebih dahulu. Kelayakan sampel dilakukan dengan melihat nilai *Kaiser Meyer Olkin* (KMO) (Nikmah & Mar'ah, 2025). Berikut Persamaan *Kaiser Meyer Olkin* (KMO):

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r^2_{ij}}{\sum_{i=1}^p \sum_{j=1}^p r^2_{ij} + \sum_{i=1}^p \sum_{j=1}^p a^2_{ij}} \quad i = 1, 2, 3, \dots, p; \quad j = 1, 2, 3, \dots, p \quad (1)$$

Keterangan:

$r^2_{ij}$  : Korelasi antara variabel ke- $i$  dan ke- $j$

$a^2_{ij}$  : Besarnya koefisien parsial antara variabel ke- $i$  dan ke- $j$

Sampel dinyatakan layak atau representatif jika nilai KMO > 0,5.

### b. Uji Multikolinieritas

Uji multikolinearitas ini diperlukan untuk mengetahui adanya hubungan korelasi yang kuat antar variabel. Beberapa hal yang dapat dilakukan apabila data mengandung multikolinieritas yaitu dapat mengurangi variabel dengan jumlah yang sama pada setiap set atau dengan menggunakan analisis komponen utama (*Principal Component Analysis*) yang mereduksi variabel menjadi beberapa faktor sehingga tidak mengandung multikolinieritas.

## 3. Metode *Elbow*

Metode *elbow* adalah teknik yang digunakan untuk menentukan jumlah *cluster* optimal secara empiris dan sederhana. Metode ini didasarkan pada prinsip bahwa peningkatan jumlah *cluster* akan mengurangi variansi dalam setiap *cluster*. Dalam metode *elbow*, penentuan jumlah *cluster* didasarkan pada nilai *Within sum of square* (WSS). Jumlah *cluster* yang optimal ditandai dengan penurunan *Within sum of square* (WSS) yang paling signifikan pada grafik, yang membentuk pola seperti siku (Jollyta *et al.*, 2021).

## 4. *K-Means Clustering*

Langkah-langkah menjalankan metode *K-Means* secara sistematis adalah sebagai berikut:

1. Menentukan jumlah *cluster* ( $k$ ) yang diinginkan sesuai dengan tujuan analisis.

2. Menentukan pusat *centroid* awal, biasanya dipilih secara acak dari data yang tersedia.
3. Menentukan jarak antara setiap data ke setiap *centroid* yang telah ditentukan menggunakan rumus *Euclidean Distance*. Rumus *Euclidean Distance*:

$$d_{ij} = \sqrt{\sum_{k=1}^m (V_{ik} - c_{jk})^2} \quad (2)$$

dengan:

$d_{ij}$  : Jarak antara data ke- $i$  ke *centroid* ke- $j$

$m$  : Banyaknya variabel yang diamati

$V_{ik}$  : Nilai data ke  $i$  pada variable ke- $k$

$c_{jk}$  : nilai *centroid* ke- $j$  pada variabel ke- $k$

4. Menghitung ulang *centroid* untuk setiap *cluster* berdasarkan rata-rata seluruh data yang berada dalam *centroid* tersebut. Rumusnya adalah:

$$c = \sum_{i=1}^n \frac{x_i}{n} \quad (3)$$

dengan:

$c$  : *Centroid* pada *cluster*

$x_i$  : data ke- $i$

$n$  : Banyaknya objek data dalam suatu *cluster* yang terbentuk

5. Mengulangi proses mulai dari langkah ke-3 hingga posisi *centroid* tidak lagi berubah secara signifikan, atau tidak terjadi perpindahan anggota antar *cluster*. Kondisi ini disebut konvergen, yang menandakan bahwa proses pengelompokan telah stabil.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Analisis Deskriptif

Analisis deskriptif dilakukan untuk memberikan gambaran umum mengenai karakteristik data dari 12 kebun kelapa sawit PT Perkebunan Nusantara IV Regional 4 yang digunakan dalam penelitian. Hasil statistik deskriptif dari kelima variabel penelitian yang dianalisis disajikan pada tabel 2.

**Tabel 2.** Analisis Deskriptif Statistika

Variabel	N	Minimum	Maksimum	Mean	Std. Deviation
Luas Lahan	12	547,55	4.416,90	2.314,73	1.182,02
Jumlah Pohon Kelapa Sawit	12	100.637	498.792	309.458,75	132.487,77
Jumlah Tandan	12	168.063	1.914.891	1.016.009,67	540.145
Rata-Rata Berat Tandan	12	8,45	77,88	22,13	18,45
Produksi TBS	12	2.514.222	35.468.090	18.719.533,33	9.713.138,38

Hasil analisis deskriptif menunjukkan adanya variasi yang cukup besar antar kebun. Misalnya, kebun dengan luas lahan terbesar adalah Bukit Kausar (4.416,9 ha), sedangkan kebun dengan luas lahan terkecil adalah Rimbo Dua (547,55 ha). Dari sisi jumlah pohon produksi, kebun Bunut memiliki jumlah pohon terbanyak, yaitu 498.792 pohon, sedangkan yang paling sedikit adalah Bukit Cermin sebanyak 100.637 pohon. Selain itu, rata-rata berat tandan tertinggi terdapat pada kebun Bunut dengan nilai 77,88 kg per tandan, sedangkan yang terendah terdapat pada kebun Lagan dengan 8,45 kg per tandan. Hal ini menunjukkan adanya perbedaan kondisi produksi antar kebun yang cukup signifikan, baik dari sisi luas area, jumlah pohon, maupun hasil produksi.

### 3.2 Standarisasi Data

Berdasarkan data yang telah dikumpulkan selanjutnya dilakukan proses standardisasi menggunakan metode *Z-Score* untuk menyamakan skala antarvariabel. Standarisasi data terhadap variabel dapat dihitung dengan rumus *Z-Score* sebagai berikut:

$$Zx_{ik} = \frac{x_{ik} - \bar{x}_k}{\sigma_k} \quad i = 1, 2, 3, \dots, n; \quad k = 1, 2, 3, \dots, n \quad (4)$$

Keterangan:

$Z$  : Nilai Standarisasi

$x_{ik}$  : Nilai objek ke- $i$  pada variabel ke- $k$

$\bar{x}_k$  : Rata-rata nilai variabel ke- $k$

$\sigma_k$  : Standar deviasi variabel ke- $k$

Hasil Standarisasi data yang digunakan dalam proses pengelompokkan disajikan dalam tabel 3 berikut:

**Tabel 3.** Hasil Standarisasi

Nama Kebun	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
Ophir	0,7693	0,4967	0,7682	-0,1129	1,0267
Bunut	1,4293	1,4291	-1,0378	3,0220	1,7243
Tanjung Lebar	-0,4844	-0,6957	0,9182	-0,6219	-0,2678
Rimbo Satu	0,4552	-0,9523	1,1321	-0,1380	1,3547
Batanghari	-0,7527	-0,2847	-0,3589	-0,1398	-0,2721
Durian Luncuk	-0,4461	-0,3679	-0,0858	-0,3449	-0,3529
P 50 Kota	-0,6265	-0,9976	-0,8501	-0,3606	-1,0397
Solok Selatan	-0,6986	0,1220	-0,6917	0,4449	0,0793
Bukit Cermin	-0,6244	-1,5762	-1,5699	-0,3889	-1,6684
Bukit Kausar	1,7785	1,4001	1,6641	-0,5214	0,5400
Rimbo Dua	-1,4951	-1,1150	-0,5058	-0,0970	-0,3715
Lagan	0,6686	0,6369	0,6173	-0,7416	-0,7528



### 3.3 Asumsi Analisis *Cluster*

#### a. Sampel Representatif

**Tabel 4.** Hasil Nilai KMO

Uji	Nilai
<i>Kiser Meyer Olkin Measure of Sampling Adequacy</i>	0,530

Berdasarkan tabel 4, diperoleh nilai KMO sebesar  $0,530 > 0,5$ . Hal ini berarti sampel yang diambil dari populasi yang digunakan telah mampu untuk mewakili keseluruhan populasi yang ada. Sehingga dapat disimpulkan bahwa data layak di analisis lebih lanjut karena telah memenuhi asumsi.

#### b. Uji Multikolinieritas

Uji multikolinieritas dilakukan untuk mengetahui ada atau tidaknya hubungan yang kuat antarvariabel independen. Pengujian dilakukan menggunakan Uji Korelasi *Rank Spearman*, Nilai korelasi  $> 0,5$  menunjukkan adanya indikasi multikolinieritas. Hasil uji korelasi *Rank Spearman* dapat dilihat pada matriks korelasi berikut:

$$\rho = \begin{bmatrix} 1 & 0,7622378 & 0,4475524 & -0,1748252 & 0,5454545 \\ 0,7622378 & 1 & 0,4335664 & 0,1888112 & 0,7902098 \\ 0,4475524 & 0,4335664 & 1 & -0,3706294 & 0,4195804 \\ -0,1748252 & 0,188112 & -0,37066294 & 1 & 0,5104895 \\ 0,5454545 & 0,7902098 & 0,4195804 & -0,5104895 & 1 \end{bmatrix}$$

Berdasarkan hasil matriks korelasi, terdapat beberapa variabel yang saling berkorelasi dengan nilai  $\rho > 0,5$ . Hal ini menunjukkan bahwa terdapat korelasi yang kuat antar variabel. Adapun cara mengatasi korelasi yang kuat antar variabel tersebut dapat dengan menggunakan metode *Principal Component Analysis* (PCA).

### 3.4 *Principal Component Analysis*

*Principal Component Analysis* digunakan karena dapat mengatasi masalah multikolinieritas di antara variabel tanpa perlu mengeluarkan variabel dari analisis. Variabel dapat dikelompokkan menjadi suatu faktor, yaitu nilai eigen yang memenuhi kriteria  $\lambda > 1$ , sedangkan nilai eigen dengan  $\lambda < 1$  tidak digunakan dalam penentuan jumlah faktor yang terbentuk sehingga proses faktorisasi harus berhenti pada suatu faktor.

**Tabel 5.** Komponen Data Hasil Analisis PCA

PC (Komponen Utama)	Nilai Eigen	Total Varian (%)	Total Kumulatif (%)
$PC_1$	3,094	61,890	61,890
$PC_2$	1,449	28,985	90,874
$PC_3$	0,353	7,053	97,927
$PC_4$	0,071	1,423	99,350
$PC_5$	0,033	0,650	100,000

Berdasarkan tabel 5, nilai eigen menunjukkan besarnya variansi yang dapat dijelaskan oleh masing-masing komponen utama. Hanya terdapat dua komponen yang memiliki nilai eigen lebih besar dari 1, yaitu  $PC_1$  dan  $PC_2$ . Langkah selanjutnya adalah menghitung nilai komponen utama untuk setiap kebun. Nilai ini diperoleh dengan mensubstitusikan data yang telah distandarisi ke dalam persamaan komponen utama berdasarkan nilai loading faktor dari masing-masing variabel. Hasil perhitungan skor komponen utama untuk setiap kebun disajikan pada tabel 6.

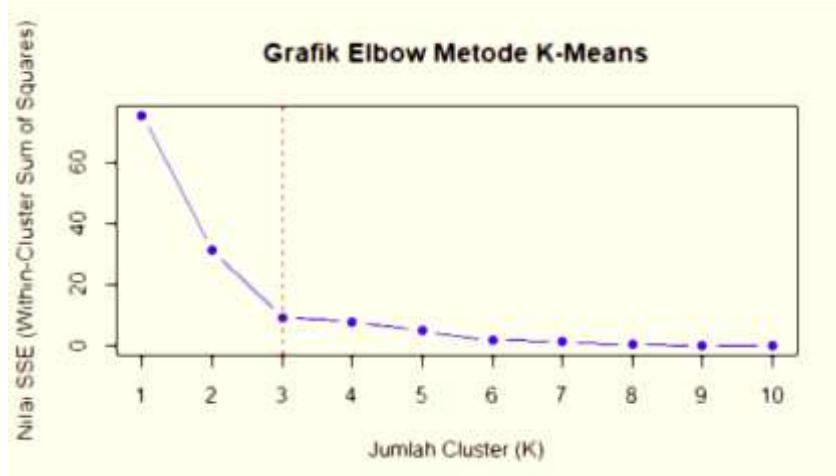
**Tabel 6.** Hasil Komponen Utama

Nama Kebun	$V_1$	$V_2$
Ophir	-1,4142	-0,5306
Bunut	-2,0989	2,9486
Tanjung Lebar	0,6519	-1,0337
Rimbo Satu	-1,7672	-0,7274
Batanghari	0,8304	0,1786
Durian Luncuk	0,7329	-0,1899
P 50 Kota	1,7675	0,2458
Solok Selatan	0,3410	0,8561
Bukit Cermin	2,6410	0,6326
Bukit Kausar	-2,3310	-1,6238
Rimbo Dua	1,7513	0,3917
Lagan	-0,2945	-1,1479

### 3.5 Metode *Elbow*

Penentuan jumlah cluster dilakukan menggunakan metode *elbow*, yaitu dengan melihat grafik hubungan antara jumlah *cluster* ( $k$ ) dan nilai *Within Sum of Squares* (WSS).





**Gambar 1.** Grafik *Elbow K-Means*

Berdasarkan hasil nilai *Elbow* yang ditampilkan pada gambar 1, penurunan nilai *Within Sum of Squares* (WSS) terlihat cukup tajam dari  $k = 1$  sampai  $k = 3$ . Namun setelah  $k = 3$ , penurunan nilai WSS mulai melandai, sehingga penambahan jumlah *cluster* tidak lagi memberikan pengurangan yang signifikan terhadap variasi data. Hal ini menunjukkan bahwa tiga *cluster* merupakan jumlah yang optimal karena sudah mampu menggambarkan perbedaan antar kebun dengan jelas tanpa membuat analisis menjadi terlalu kompleks,

### 3.6 *K-Means Clustering*

*Centroid* awal dipilih secara acak dari data yang telah dinormalisasi, dengan masing-masing mewakili satu kebun yang berbeda. Adapun titik awal *centroid* tersebut adalah:

$c_1$ : Kebun Ophir

$c_2$ : Kebun Rimbo Satu

$c_3$ : Kebun Bukit Kausar

Pemilihan secara acak ini mengikuti pendekatan dasar algoritma *K-Means*, yang selanjutnya akan diperbarui secara iteratif hingga posisi *centroid* tidak berubah signifikan. Selanjutnya menghitung jarak antara *centroid* dan objek data menggunakan ukuran jarak *Euclidean*.

**Tabel 7.** Hasil Perhitungan *Euclidean* Iterasi 1

Nama Kebun	$c_1$	$c_2$	$c_3$	Jarak Terdekat	Cluster
Ophir	0,0000	3,5459	2,1265	0	1
Bunut	3,5459	0,0000	4,8399	0	2
Tanjung Lebar	2,1265	4,8399	0,0000	0	3
Rimbo Satu	0,4041	3,6909	2,4384	0,4041	1
Batanghari	2,3540	4,0316	1,2253	1,2253	3
Durian Luncuk	2,1740	4,2271	0,8477	0,8477	3
P 50 Kota	3,2751	4,7175	1,6975	1,6975	3
Solok Selatan	2,2369	3,2143	1,9152	1,9152	3
Bukit Cermin	4,2187	5,2755	2,5948	2,5948	3
Bukit Kausar	1,4267	4,5783	3,0407	1,4267	1
Rimbo Dua	3,2971	4,6218	1,8001	1,8001	3
Lagan	1,2786	4,4762	0,9533	0,9533	3

Setelah jarak setiap kebun terhadap *centroid* telah diperoleh, selanjutnya adalah mengevaluasi apakah terjadi perubahan pada keanggotaan *cluster* yang telah terbentuk. Tahap berikutnya yaitu menentukan *centroid* baru dengan menghitung rata-rata nilai dari seluruh kebun yang termasuk dalam masing-masing *cluster*. Hasil *centroid* baru untuk iterasi ke 2 di sajikan dalam tabel 8.

**Tabel 8.** *Centroid* Baru Iterasi 2

Centroid	$V_1$	$V_2$
$c_1$	-1,8375	-0,9606
$c_2$	-2,0989	2,9486
$c_3$	1,0527	-0,0083

Hasil perhitungan jarak serta pengelompokkan kebun berdasarkan jarak terdekat terhadap masing-masing *centroid* iterasi ke-2 disajikan pada tabel 9.

**Tabel 9.** Hasil Perhitungan *Euclidean* iterasi 2

Nama Kebun	$c_1$	$c_2$	$c_3$	Jarak Terdekat	Cluster
Ophir	0,6033	3,5459	2,5216	0,6033	1
Bunut	3,9179	0,0000	4,3215	0	2
Tanjung Lebar	2,4904	4,8399	1,1009	1,1009	3
Rimbo Satu	0,2436	3,6909	2,9101	0,2436	1
Batanghari	2,9009	4,0316	0,2904	0,2904	3
Durian Luncuk	2,6834	4,2271	0,3677	0,3677	3
P 50 Kota	3,8015	4,7175	0,7587	0,7587	3
Solok Selatan	2,8366	3,2143	1,1197	1,1197	3
Bukit Cermin	4,7534	5,2755	1,7127	1,7127	3
Bukit Kausar	0,8267	4,5783	3,7495	0,8267	1
Rimbo Dua	3,8350	4,6218	0,8050	0,8050	3
Lagan	1,5542	4,4762	1,7645	1,5542	1

Karena posisi *cluster* yang diperoleh pada iterasi ke-2 masih mengalami perubahan, maka proses pengelompokan dilanjutkan hingga mencapai kondisi konvergen. Kondisi konvergen dicapai pada iterasi ke-3 yang ditandai dengan posisi *cluster* sudah tidak mengalami perpindahan. Hasil akhir pengelompokkan disajikan kedalam tabel 10 berikut:

**Tabel 10.** Hasil Akhir *Cluster*

<i>Cluster</i>	Jumlah	Nama Kebun Kelapa Sawit
$c_1$	4	Kebun Ophir, Rimbo Satu, Bukit Kausar dan Lagan
$c_2$	1	Kebun Bunut
$c_3$	7	Kebun Tanjung Lebar, Batanghari, Durian Luncuk, Pangkalan 50 Kota, Bukit Cermin Solok Selatan, dan Rimbo Dua

Karakteristik masing-masing *cluster* dapat diketahui melalui perhitungan nilai rata-rata kedua komponen yang terbentuk. Hasil perhitungan rata-rata disajikan pada tabel 11.

**Tabel 11.** Nilai Rata-Rata Tiap Variabel

	Luas Lahan (Ha)	Jumlah Pohon (Pohon)	Jumlah Tandan (Tandan)	Rata-Rata Berat Tandan (Kg)	Produksi TBS (Kg)
$c_1$	3.408	424.925	1.580.698	15	23.985.705
$c_2$	4.004	498.792	455.431	78	35.468.090
$c_3$	1.449	216.413	773.413	18	13.317.641

Berdasarkan hasil pengelompokan menggunakan metode *K-Means Clustering*, diperoleh tiga *cluster* kebun kelapa sawit dengan tingkat produktivitas yang berbeda. *Cluster* 1 merupakan *cluster* dengan kategori produktivitas sedang karena memiliki rata-rata jumlah tandan yang tinggi yaitu 1.580.698 tandan, namun rata-rata berat tandan hanya 15 kg dengan total produksi 23.985.705 kg. Hal ini menunjukkan bahwa meskipun hasil panen banyak, ukuran tandan yang kecil membuat produktivitas belum optimal. *Cluster* 2 merupakan *cluster* dengan kategori produktivitas tinggi karena memiliki luas lahan besar yaitu 4.004 ha dan rata-rata berat tandan tertinggi yaitu 78 kg, meskipun jumlah tandannya hanya 455.431. Hal ini menunjukkan bahwa kebun dalam *cluster* ini mampu menghasilkan produksi TBS yang tinggi karena ukuran tandan yang besar dan efisiensi produksi yang baik. *Cluster* 3 merupakan *cluster* dengan kategori produktivitas rendah karena memiliki rata-rata luas lahan dan jumlah pohon yang kecil, masing-masing 1.449 ha dan 216.431 pohon, dengan total produksi hanya 13.317.641 kg. Kondisi ini menunjukkan bahwa kebun dalam *cluster* ini memiliki hasil produksi yang masih rendah dan memerlukan peningkatan dalam pengelolaan serta pemeliharaan tanaman.

#### 4. KESIMPULAN

Berdasarkan analisis yang telah dilakukan diperoleh kesimpulan bahwa terbentuk 3 *cluster* kebun kelapa sawit PT Perkebunan Nusantara IV Regional 4 berdasarkan produktivitas Tandan Buah Segar. Dimana *Cluster* 1 terdiri dari Kebun Ophir, Rimbo Satu, Bukit Kausar dan Lagan. *Cluster* ini ditandai dengan produktivitas sedang, karena keseimbangan antara jumlah tandan dan berat tandan belum sepenuhnya optimal. *Cluster* 2 hanya terdiri dari Kebun Bunut. *Cluster* ini menunjukkan produktivitas tinggi dengan karakteristik ukuran tandan yang sangat besar, sehingga mampu menghasilkan produktivitas tinggi meskipun jumlah tandan relatif sedikit. *Cluster* 3 mencakup Kebun Tanjung Lebar, Batanghari, Durian Luncuk, Pangkalan 50 Kota, Bukit Cermin Solok Selatan, dan Rimbo Dua. *Cluster* ini ditandai dengan jumlah pohon dan luas lahan yang relatif kecil sehingga berdampak pada rendahnya produksi Tandan Buah Segar.

#### 5. DAFTAR PUSTAKA

- Asroni, & Adrian, R. (2015). Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang (*Implementation Method for K-Means Clustering Based Student Value with Weka Interface a Case Stud*). *Jurnal Ilmiah Semesta Teknika*, 18(1), 76–82.
- Fauziah, & Basir, C. (2024). Analisis Klaster dengan Metode K-Means Berdasarkan Usia Warga yang Divaksin Covid-19 di Kelurahan Grogol Selatan. *Jurnal Bayesian: Jurnal Ilmiah Statistika dan Ekonometrika*, 4(1), 53–62. <https://bayesian.lppmbinabangsa.id/index.php/home/article/view/76>
- Permata Sari, Y., & Basir, C. (2025). Analisis Klaster dengan Metode K-Means pada Persebaran Kasus Covid-19 Berdasarkan Desa di Kecamatan Kemang-Bogor. *Pelita : Jurnal Penelitian Dan Karya Ilmiah*, 24(2), 33–45. <https://doi.org/10.33592/pelita.v24i2.5504>
- Putri Br Ginting, I, Jamil, M., & Supristiwendi, S. (2022). Faktor-Faktor Yang Mempengaruhi Produksi Kelapa Sawit (*Elaeis Guineensis Jacq.*) Di PT. Perkebunan Nusantara II Kebun Sawit Seberang Kabupaten Langkat. *Jurnal Penelitian Agrisamudra*, 9(1).
- Jollyta, D., Siddik, M., Mawengkang, H., & Efendi, S. (2021). Teknik Evaluasi Cluster Solusi Menggunakan Phyton dan Rapidminer. *Deepublish*.
- Lubis, M. F., & Lubis, I. (2018). Analisis Produksi Kelapa Sawit (*Elaeis Guineensis Jacq.*) Di Kebun Buatan, Kabupaten Pelalawan, Riau. *Buletin Agrohorti*, 6(2), 281–286.
- Mustofani, D. (2022). Penerapan Uji Korelasi Rank Spearman Untuk Mengetahui Hubungan Tingkat Pengetahuan Ibu Terhadap Tindakan Swamedikasi Dalam Penanganan Demam Pada Anak. *Jurnal UJMC*, 9, 9–13.
- Nikmah, A., Mar'ah, I.S. (2025). Pengelompokan Provinsi di Indonesia Berdasarkan Kesehatan Jiwa Menggunakan Algoritma KMeans. *Journal of Sciencetech Research and Development*, 7(2), 1–11.
- Suhartono, M.J., Setyawan, H., & Aji, W.A. (2023). Identifying Plant Age to Determine



Production Trend of Oil Palm Fresh Fruit Bunches. *Agricultural Journal*, 6(2), 378-384.

Suparwito, H., Gunawan, R., Binanto, I., Wisyastuti, W., & Kumalasanti, R.A. (2024). Pengantar Manajemen Data. Sanata Dharma University Press.

Yosephine, I. O., Effendi, Z., & Fitriana, D. (2025). Study of the Productivity of Oil Palm Plants (*Elaeis guineensis* Jacq.) on Different Topographies. *Jurnal Teknik Pertanian Lampung (Journal of Agricultural Engineering)*, 14(4), 1349–1358.