

**ANALISIS POLA PRODUKTIVITAS DAN DISTRIBUSI HASIL PANEN
SAWIT MENGGUNAKAN METODE K-MEANS CLUSTERING
(STUDI KASUS: KUD PELITA DESA SILAU JAWA)**

SKRIPSI

DISUSUN OLEH

ELVINA JULIANTI HERMAWAN BR SITORUS

NPM. 2209010125



UMSU

Unggul | Cerdas | Terpercaya

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

MEDAN

2026

**ANALISIS POLA PRODUKTIVITAS DAN DISTRIBUSI HASIL PANEN
SAWIT MENGGUNAKAN METODE K-MEANS CLUSTERING
(STUDI KASUS: KUD PELITA DESA SILAU JAWA)**

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer
(S.Kom) dalam Program Studi Sistem Informasi pada Fakultas Ilmu Komputer
dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara**

ELVINA JULIANTI HERMAWAN BR SITORUS

NPM. 2209010125

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

MEDAN

2026

LEMBAR PENGESAHAN

Judul Skripsi : ANALISIS POLA PRODUKTIVITAS DAN
DISTRIBUSI HASIL PANEN SAWIT
MENGUNAKAN METODE K-MEANS
CLUSTERING (STUDI KASUS: KUD PELITA DESA
SILAU JAWA)

Nama Mahasiswa : ELVINA JULIANTI HERMAWAN BR SITORUS

NPM : 2209010125

Program Studi : SISTEM INFORMASI

Menyetujui
Komisi Pembimbing



Amrullah, S.Kom., M.Kom.
NIDN. 0125118604

Ketua Program Studi



Mahardika Abdi Tanjung, S.Kom., M.Kom.
NIDN. 0117088902



Dr. Al-Khowarizmi, S.Kom., M.Kom.
NIDN. 0127099201

PERNYATAAN ORISINALITAS

ANALISIS POLA PRODUKTIVITAS DAN DISTRIBUSI HASIL PANEN SAWIT MENGGUNAKAN METODE K-MEANS CLUSTERING (STUDI KASUS: KUD PELITA DESA SILAU JAWA)

SKRIPSI

Saya menyatakan bahwa karya tulis ini adalah hasil karya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya.

Medan, 10 Mei 2026

Yang membuat pernyataan



Elvina Julianti Hermawan Br Sitorus

NPM. 2209010125

**PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN
AKADEMIS**

Sebagai sivitas akademika Universitas Muhammadiyah Sumatera Utara, saya bertanda tangan dibawah ini:

Nama : Elvina Julianti Hermawan Br Sitorus
NPM : 2209010125
Program Studi : Sistem Informasi
Karya Ilmiah : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bebas Royalti Non-Eksekutif (*Non-Exclusive Royalty free Right*) atas penelitian skripsi saya yang berjudul:

**ANALISIS POLA PRODUKTIVITAS DAN DISTRIBUSI HASIL PANEN
SAWIT MENGGUNAKAN METODE K-MEANS CLUSTERING
(STUDI KASUS: KUD PELITA DESA SILAU JAWA)**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksekutif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk database, merawat dan mempublikasikan Skripsi saya ini tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan atau sebagai pemilik hak cipta.

Demikian pernyataan ini dibuat dengan sebenarnya.

Medan, 10 Mei 2026

Yang membuat pernyataan



Elvina Julianti Hermawan Br Sitorus

NPM. 2209010125

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap : Elvina Julianti Hermawan Br Sitorus
Tempat dan Tanggal Lahir : Tangerang, 25 Juli 2003
Alamat Rumah : Silau Jawa, Bandar Pasir Mandoge
Telepon/Faks/HP : 082360165772
E-mail : elvina240320@gmail.com
Instansi Tempat Kerja : -
Alamat Kantor : -

DATA PENDIDIKAN

SD : SDN 016404 TAMAT: 2015
SMP : SMP N 1 SILAU JAWA TAMAT: 2018
SMA : SMKN 1 KISARAN TAMAT: 2021

KATA PENGANTAR



Assalamu'alaikum wr. wb. Puji dan syukur penulis panjatkan ke hadirat Allah SWT karena atas rahmat dan karunia-Nya, skripsi ini dapat diselesaikan sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara. Skripsi ini berjudul **“Analisis Pola Produktivitas Dan Distribusi Hasil Panen Sawit Menggunakan Metode K-Means Clustering (Studi Kasus: Kud Pelita Desa Silau Jawa)”**. Penulis menyadari skripsi ini masih memiliki kekurangan dan jauh dari kata sempurna. Oleh karena itu, penulis memohon maaf apabila terdapat kesalahan dalam penulisan maupun penyusunan skripsi ini. Kritik dan saran yang membangun sangat diharapkan agar skripsi ini dapat menjadi lebih baik. Semoga skripsi ini bermanfaat bagi pembaca.

Selama proses penyusunan, penulis memperoleh dukungan, bimbingan, serta bantuan dari berbagai pihak. Untuk itu, Penulis tentunya berterima kasih kepada berbagai pihak dalam dukungan serta doa dalam penyelesaian skripsi. Penulis juga mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Akrim, M.Pd selaku Rektor Universitas Muhammadiyah Sumatera Utara (UMSU)
2. Bapak Dr. Al-Khowarizmi, S.Kom., M.Kom. selaku Dekan Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
3. Bapak Mahardika Abdi Prawira Tanjung, S.Kom., M.Kom selaku Ketua Program Studi Sistem Informasi.
4. Bapak Amrullah, S.Kom., M.Kom. selaku Dosen Pembimbing Skripsi yang sabar dalam mengarahkan penulis untuk menyelesaikan penelitian ini.
5. Bapak Halim Maulana, S.T., M.Kom., MTA selaku Dosen Pembimbing Akademik penulis.

6. Teruntuk kedua orang tuaku tercinta Bapak Herman Sitorus dan Ibu Dahniar, Penulis persembahkan karya sederhana ini kepada Ibu dan Bapak. Sebagai ungkapan terima kasih atas segala kasih sayang, semangat, perhatian, dukungan dan doa yang tak pernah putus, dan mengusahakan segalanya untuk anak-anak nya dalam menempuh pendidikan setinggi tingginya sehingga penulis menyelesaikan studinya sampai mendapat gelar sarjana. Ibu adalah sumber kekuatan dan inspirasi terbesar penulis, di balik setiap keberhasilan kecil yang penulis raih, ada doa ibu yang tak pernah berhenti dan setiap tetes keringan menjadi saksi perjuangan ayah untuk membahagiaakan anak-anaknya.
7. Terimakasih untuk rekan seperjuangan dibangku perkuliahan yang menemani penulis dari maba sampai mendapat gelar Sarjana khususnya teman anggota PB GIRLS dan teman mahasiswa Sistem Informasi kelas C1 pagi yang tidak bisa penulis sebutkan satu per satu namanya, Terima kasih untuk canda tawa dan tangis yang kita lewati, walaupun tidak sedarah tapi kenangan yang terukir sangat berkesan dihati, yang sama-sama berlari dalam perjuangan, setiap orang ada masanya dan setiap masa ada orangnya
8. Terima Kasih untuk sahabat yang telah ada hadir dalam hidup penulis Miliana Putri Sitorus, Syafriani, dan Maulida Wahyuni, sehingga penulis merasa jauh lebih hidup disaat ada mereka. Sejak adanya mereka penulis menemukan arti hidup yang sebenarnya, yang bersedia setia menemani, berbagi cerita keluh kesah dan tangis penulis dalam setiap langkah hidup penulis. Semoga persahabatan ini selalu terjaga. Setiap masa ada orangnya tetapi aku ingin setiap masa selalu ada mereka.
9. Dan terakhir, teruntuk diriku sendiri Elvina Julianti Hermawan Br Sitorus, apresiasi sebesar besarnya yang telah berjuang untuk menyelesaikan apa yang telah dimulai. Sulit untuk berada di titik ini, terimakasih untuk tetap hidup dan merayakan dirimu sendiri, walaupun seringkali putus asa atas apa yang sedang diusahakan. Tetaplah jadi manusia yang mau berusaha dan tidak lelah untuk mencoba. God thank you for being me independent woman, i Know there are more great ones but i'm proud of this

achievement. Gelar dibelakang nama sudah sah terpampang, entah menjadi apa setelah ini, semoga hal-hal baik selalu mengahmpiri. Dalam menyelesaikan tugas akhir ini saya sebagai peneliti tentu mempunyai banyak kekurangan dan kesalahan baik disengaja maupun tidak disengaja. Maka dari itu, saya akan sangat menghargai setiap masukan dan kritik untuk membangun diri saya agar lebih baik kedepannya. Akhir kata, saya memohon maaf sebesarbesarnya dan berharap tugas akhir ini dapat berguna dan memberikan manfaat. Terimakasih Semua pihak yang terlibat langsung ataupun tidak langsung yang tidak dapat penulis ucapkan satu-persatu yang telah membantu penyelesaian skripsi ini.

ANALISIS POLA PRODUKTIVITAS DAN DISTRIBUSI HASIL PANEN
SAWIT MENGGUNAKAN METODE K-MEANS CLUSTERING
(STUDI KASUS: KUD PELITA DESA SILAU JAWA)

ABSTRAK

Produktivitas dan distribusi hasil panen kelapa sawit merupakan faktor penting dalam mendukung efektivitas pengelolaan koperasi perkebunan rakyat. KUD Pelita Desa Silau Jawa sebagai pengelola hasil panen kelapa sawit menghadapi permasalahan dalam pengolahan data produktivitas dan distribusi hasil panen yang masih dilakukan secara manual, sehingga menyulitkan proses analisis dan pengambilan keputusan. Perbedaan luas lahan, jumlah hasil panen, frekuensi panen, dan volume distribusi antar mandor menyebabkan diperlukan suatu metode analisis berbasis data untuk mengelompokkan karakteristik produktivitas secara objektif. Penelitian ini bertujuan untuk merancang sistem informasi berbasis web yang mampu mengelola data hasil panen sekaligus menerapkan metode K-Means Clustering dalam menganalisis pola produktivitas dan distribusi hasil panen kelapa sawit. Metode penelitian yang digunakan meliputi tahapan pengumpulan data, pembersihan data, normalisasi menggunakan metode Min-Max, serta proses clustering menggunakan algoritma K-Means. Variabel yang digunakan dalam penelitian ini terdiri dari luas lahan, jumlah hasil panen, frekuensi panen, dan volume distribusi hasil panen. Data penelitian diperoleh dari rekap produksi hasil panen mandor anggota KUD Pelita Desa Silau Jawa periode 2024–2025. Sistem dirancang menggunakan bahasa pemrograman PHP dan basis data MySQL. Hasil penelitian menunjukkan bahwa metode K-Means Clustering mampu mengelompokkan mandor ke dalam beberapa cluster berdasarkan tingkat produktivitas dan distribusi hasil panen. Hasil pengelompokan menghasilkan kategori mandor dengan produktivitas tinggi, sedang, dan rendah sehingga dapat membantu koperasi dalam melakukan evaluasi kinerja, perencanaan distribusi hasil panen, serta pengambilan keputusan berbasis data. Sistem yang dibangun juga mampu menyajikan hasil clustering dalam bentuk tabel dan visualisasi sehingga informasi lebih mudah dipahami oleh pengurus koperasi.

Kata Kunci: Data Mining, K-Means Clustering, Produktivitas Kelapa Sawit, Distribusi Hasil Panen, Sistem Informasi.

**ANALYSIS OF OIL PALM PRODUCTIVITY AND DISTRIBUTION
PATTERNS USING THE K-MEANS CLUSTERING METHOD
(CASE STUDY: KUD PELITA, DESA SILAU JAWA)**

ABSTRACT

Productivity and distribution of oil palm harvests are important factors in supporting the effectiveness of rural plantation cooperative management. KUD Pelita Desa Silau Jawa, as a manager of oil palm harvests, faces problems in processing productivity and harvest distribution data which are still carried out manually, making analysis and decision-making processes difficult. Differences in land area, harvest yield, harvesting frequency, and distribution volume among supervisors require a data-based analytical method to objectively classify productivity characteristics. This study aims to design a web-based information system capable of managing harvest data while implementing the K-Means Clustering method to analyze productivity patterns and the distribution of oil palm harvests. The research method used includes data collection, data cleaning, normalization using the Min-Max method, and clustering processes using the K-Means algorithm. The variables used in this study consist of land area, harvest yield, harvesting frequency, and harvest distribution volume. The research data were obtained from the production recap data of harvest supervisors at KUD Pelita Desa Silau Jawa for the 2024–2025 period. The system was developed using the PHP programming language and MySQL database. The results of the study indicate that the K-Means Clustering method is able to classify supervisors into several clusters based on productivity levels and harvest distribution characteristics. The clustering results produce categories of high, medium, and low productivity supervisors, which can assist the cooperative in evaluating performance, planning harvest distribution, and supporting data-based decision making. The developed system is also able to present clustering results in the form of tables and visualizations, making the information easier for cooperative administrators to understand.

Keywords: Data Mining, K-Means Clustering, Oil Palm Productivity, Harvest Distribution, Information System.

DAFTAR ISI

LEMBAR PENGESAHAN	i
PERNYATAAN ORISINALITAS	ii
RIWAYAT HIDUP	v
KATA PENGANTAR	vi
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI	xi
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xvi
BAB I	1
PENDAHULUAN	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah.....	5
1.4. Tujuan Penelitian.....	5
1.5. Manfaat Penelitian.....	6
BAB II	8
LANDASAN TEORI	8
2.1. Tanaman Kelapa Sawit.....	8
2.2. Produktivitas Hasil Panen Kelapa Sawit	8
2.3. Distribusi Hasil Panen Kelapa Sawit	11
2.4. Variabel Penelitian dalam Analisis Produktivitas dan Distribusi	11
2.5. Konsep Data Mining	12
2.6. Clustering Sebagai Teknik Analisis Data.....	12
2.7. Algoritma K-Means.....	13
2.8. Perhitungan Jarak Euclidean dalam K-Means	14
2.9. Normalisasi Data	15
2.10. Studi Terdahulu (Penelitian Relevan)	17
2.11. Hipotesis Penelitian.....	21
2.12. Kerangka Pemikiran.....	21

BAB III ANALISA DAN PERANCANGAN SISTEM.....	27
3.1. Analisis Permasalahan.....	27
3.1.1. Permasalahan pada Pengelolaan Data.....	27
3.1.2. Permasalahan pada Analisis Produktivitas.....	28
3.2. Analisis Kebutuhan Sistem	29
3.2.1 Kebutuhan Fungsional	29
3.2.2. Kebutuhan Non-Fungsional	30
3.3. Spesifikasi Perangkat Keras dan Perangkat Lunak Sistem	30
3.3.1. Spesifikasi Perangkat Keras	30
3.3.2. Spesifikasi Perangkat Lunak	31
3.3.3. Keterkaitan Perangkat dengan Sistem.....	31
3.4. Tahapan Perancangan Sistem.....	31
3.4.1. Perancangan Tampilan	32
3.4.2. Usecase Diagram.....	36
3.4.3. Activity Diagram.....	37
3.4.4. Sequence Diagram	38
3.5. Perancangan Basis Data	38
3.5.1. Alur Sistem.....	39
3.5.2. Entity Relationship Diagram (ERD)	40
3.5.3. Struktur Tabel Basis Data	41
3.5.4. Relasi Antar Tabel.....	43
3.6. Keterkaitan Sistem dengan Studi Kasus.....	43
3.7. Rancangan Output Sistem	44
3.8. Lokasi dan Waktu Penelitian.....	44
3.9. Objek Penelitian	45
3.10. Sumber dan Jenis Data	46
3.10.1. Sumber Data.....	46
3.10.2. Jenis Data	46
3.11. Teknik Pengumpulan Data	47
3.12. Variabel Penelitian	48
3.13. Analisis Data	48
3.13.1. Seleksi Data.....	49

3.13.2. Pembersihan Data (Data Cleaning).....	49
3.13.3. Normalisasi Data.....	49
3.13.4. Penentuan Jumlah Cluster (K)	50
3.13.5. Proses Clustering dengan K-Means	50
3.13.6. Analisis Hasil Clustering Mandor.....	51
3.14. Alur Penelitian.....	52
3.15. Validasi Hasil Penelitian	53
3.16. Waktu Penelitian	53
BAB IV	55
IMPLEMENTASI DAN PENGUJIAN SISTEM.....	55
4.1. Seleksi dan Persiapan Data	55
4.1.1 Seleksi Data.....	55
4.1.2. Pembersihan Data.....	56
4.2. Data Awal dan Penentuan Nilai Min-Max	56
4.2.1. Data Rata-Rata Mandor Periode 2024–2025	56
4.2.2. Penentuan Nilai Minimum dan Maksimum	57
4.3. Normalisasi Data	58
4.3.1. Contoh Perhitungan Normalisasi	58
4.4. Penentuan Jumlah Cluster dan Centroid Awal	59
4.4.1. Penentuan Jumlah Cluster (K)	59
4.4.2. Penentuan Centroid Awal dengan K-Means++	60
4.5. Proses Clustering K-Means	61
4.5.1. Iterasi 1 – Perhitungan Jarak dan Penugasan Cluster.....	61
4.5.2. Pembaruan Centroid Setelah Iterasi 1	62
4.5.3. Iterasi 2 – Verifikasi Konvergensi	63
4.6. Centroid Akhir dan Hasil Pengelompokan.....	64
4.6.1. Centroid Akhir (Nilai Normalisasi)	64
4.6.2. Centroid Akhir dalam Satuan Asli	64
4.6.3. Hasil Akhir Pengelompokan Mandor.....	65
4.7. Analisis Karakteristik Hasil Clustering	65
4.7.1. Cluster 1 – Produktivitas Sedang-Tinggi	66
4.7.2. Cluster 2 – Produktivitas Rendah (Keterbatasan Lahan)	66

4.7.3. Cluster 3 – Lahan Luas dengan Produktivitas Per Hektar Rendah	67
4.8. Implementasi Sistem.....	68
4.8.1. Kebutuhan Sistem	68
4.8.2. Halaman Login.....	68
4.8.3. Halaman Dashboard	69
4.8.4. Halaman Data Mandor	70
4.8.5. Halaman Data Produksi Panen.....	70
4.8.6. Halaman Proses Clustering	71
4.8.7. Halaman Laporan	72
4.9. Pengujian Sistem	72
4.9.1. Hasil Pengujian Black-Box	73
4.9.2. Validasi Hasil Sistem terhadap Perhitungan Manual.....	74
4.10. Pembahasan Kelebihan dan Kelemahan Sistem.....	75
4.10.1. Kelebihan Sistem	75
4.10.2. Kelemahan dan Keterbatasan Sistem	76
4.10.3. Relevansi Hasil terhadap Kebutuhan KUD Pelita Maju Mandiri	77
BAB V.....	78
KESIMPULAN DAN SARAN	78
5.1. Kesimpulan	78
5.2. Saran	79
DAFTAR PUSTAKA	81

DAFTAR GAMBAR

Gambar. 2.2 Perkembangan Luas Areal Dan Produksi Minyak Sawit Tersertifikasi Berkelanjutan (sumber, RSPO, 2022)	10
Gambar 2.7 K-Mean Clustering	13
Gambar 2. 11 Alur Pemecahan Masalah.....	24
Gambar 3.4.1 Rancangan Halaman Login Sistem	32
Gambar 3.4.2 Rancangan Halaman Dashboard	33
Gambar 3.4.3 Rancangan Halaman Data Mandor	34
Gambar 3.4.4 Rancangan Halaman Data Produksi	34
Gambar 3.4.5 Rancangan Halaman Proses Clustering.....	35
Gambar 3.4.6 Rancangan Halaman Laporan	36
Gambar 3.4.7 Use Case Diagram Sistem Informasi KUD Pelita Maju Mandiri ...	36
Gambar 3.4.8 Activity Diagram Sistem	37
Gambar 3.4.9 Sequence Diagram Proses Clustering	38
Gambar 3.5.1 Alur Sistem Pengelompokan Mandor Menggunakan K-Means	39
Gambar 3.5.2. Entity Relationship Diagram (ERD) Sistem Pengelolaan Panen Sawit.....	40
Gambar 4.8.2 Tampilan Halaman Login.....	69
Gambar 4.8.3 Tampilan Halaman Dashboard.....	69
Gambar 4.8.4 Tampilan Halaman Data Mandor	70
Gambar 4.8.5 Tampilan Halaman Data Produksi Panen.....	71
Gambar 4.8.6 Tampilan Halaman Proses Clustering	71
Gambar 4.8.7 Tampilan Halaman Laporan.....	72

DAFTAR TABEL

Tabel 2.10 Penelitian Terdahulu	21
a. Tabel Mandor	41
b. Tabel data_panen	41
c. Tabel hasil_clustering.....	42
d. Tabel detail_clustering	42
Tabel 3.9 Daftar Mandor dan Kelompok Objek Penelitian	45
Tabel 3.11 Sampel Data Produksi Hasil Panen (KUD Pelita Maju Mandiri, 2024–2025)	47
Tabel 3.12 Variabel Penelitian dalam Proses K-Means Clustering	48
Tabel 3.13.3 Hasil Normalisasi Data (Rata-rata Seluruh Periode 2024–2025)	49
Tabel 3.13.6 Daftar Mandor/Kelompok sebagai Subjek Analisis Clustering	51
Tabel 3.16 Waktu Penelitian	54
Tabel 4.2.1 Data Rata-Rata Produksi Mandor Periode 2024–2025	57
Tabel 4.2.2 Nilai Minimum dan Maksimum Setiap Variabel	58
Tabel 4.3.2 Hasil Normalisasi Data Seluruh Mandor	59
Tabel 4.4.2 Centroid Awal (K-Means++)	60
Tabel 4.5.1 Perhitungan Jarak Euclidean dan Penugasan Cluster – Iterasi 1	61
Tabel 4.5.2 Centroid Baru Setelah Iterasi 1	63
Tabel 4.5.3 Perhitungan Jarak Euclidean dan Penugasan Cluster – Iterasi 2	63
Tabel 4.6.1 Nilai Centroid Akhir (Ternormalisasi).....	64
Tabel 4.6.2 Nilai Centroid Akhir dalam Satuan Asli	64
Tabel 4.6.3 Hasil Akhir Pengelompokan K-Means (K=3) – KUD Pelita Maju Mandiri 2024–2025	65
Tabel 4.9.1 Hasil Pengujian Black-Box Testing	73
Tabel 4.9.2 Validasi Hasil Clustering Sistem vs Perhitungan Manual	74

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Industri kelapa sawit merupakan salah satu sektor strategis dalam perekonomian Indonesia karena kontribusinya terhadap produk domestik bruto, devisa negara, dan penyerapan tenaga kerja, khususnya di wilayah pedesaan. Indonesia tercatat sebagai produsen minyak kelapa sawit terbesar di dunia, sehingga keberlanjutan dan efisiensi pengelolaan perkebunan kelapa sawit menjadi isu penting yang terus mendapat perhatian dari berbagai pihak. Seiring dengan meningkatnya volume data hasil panen dari waktu ke waktu, pengelolaan data secara manual menjadi kurang efektif dalam menghasilkan informasi yang akurat. Oleh karena itu, diperlukan pemanfaatan teknologi informasi untuk membantu proses pengelolaan dan pengolahan data secara lebih sistematis. Pada tingkat perkebunan rakyat, produktivitas dan distribusi hasil panen kelapa sawit menjadi faktor penting yang menentukan keberhasilan pengelolaan usaha tani dan kinerja koperasi sebagai lembaga pengelola hasil panen (Simangunsong et al., 2022).

Produktivitas kelapa sawit pada tingkat perkebunan rakyat tidak hanya ditentukan oleh faktor alam, tetapi juga dipengaruhi oleh aspek manajerial dan operasional, seperti luas lahan, intensitas pemeliharaan, frekuensi panen, serta sistem distribusi hasil panen. Perbedaan pengelolaan tersebut menyebabkan terjadinya variasi produktivitas antar mandor meskipun berada dalam wilayah yang sama. Kondisi ini menunjukkan bahwa permasalahan produktivitas kelapa sawit bersifat multivariat dan memerlukan pendekatan analisis berbasis data untuk

memahaminya secara objektif. Namun, dalam praktiknya, pengelolaan data pada tingkat koperasi masih menghadapi keterbatasan, terutama dalam hal pengolahan dan analisis data. Data yang tersedia belum sepenuhnya dimanfaatkan sebagai informasi yang dapat mendukung pengambilan keputusan.

Koperasi Unit Desa (KUD) berperan sebagai pusat pengumpulan, pencatatan, dan pendistribusian data hasil panen kelapa sawit dari mandor anggota. KUD tidak hanya berfungsi sebagai perantara distribusi Tandan Buah Segar (TBS), tetapi juga sebagai pengelola data produktivitas mandor yang digunakan sebagai dasar pengambilan keputusan, seperti perencanaan distribusi, evaluasi kinerja mandor, dan penyusunan kebijakan pembinaan. Oleh karena itu, pengelolaan data yang akurat, terstruktur, dan mudah dianalisis menjadi kebutuhan yang sangat penting dalam mendukung proses keputusan berbasis data (Penerapan Sistem Intelijensia Bisnis & K-Means Clustering, 2022).

KUD Pelita Desa Silau Jawa juga merupakan salah satu koperasi yang mengelola hasil panen kelapa sawit dari mandor anggota dengan karakteristik lahan, frekuensi panen, dan volume produksi yang berbeda-beda. Berdasarkan data awal koperasi, terdapat perbedaan volume hasil panen dan distribusi antar mandor yang cukup signifikan meskipun berada dalam satu wilayah pengelolaan. Kondisi ini menyebabkan kesulitan dalam menentukan prioritas pengelolaan, perencanaan distribusi, serta evaluasi kinerja produksi mandor secara objektif. Kondisi tersebut menunjukkan bahwa diperlukan suatu sistem informasi yang mampu mengelola data secara terintegrasi serta mendukung proses analisis data secara lebih efektif dan efisien. Untuk mengidentifikasi pola dari perbedaan tersebut, diperlukan pendekatan berbasis teknologi informasi yang mampu

melakukan pengelompokan data secara otomatis. Salah satu metode yang dapat digunakan adalah teknik clustering dalam data mining.

Selain permasalahan produktivitas, aspek distribusi hasil panen juga menjadi tantangan tersendiri. Distribusi TBS yang tidak terencana dengan baik dapat mengakibatkan keterlambatan pengangkutan, penumpukan hasil panen, serta penurunan kualitas buah sebelum sampai ke pabrik. Oleh karena itu, diperlukan suatu pendekatan analisis yang mampu mengelompokkan mandor berdasarkan karakteristik produktivitas dan distribusi agar pengelolaan koperasi dapat dilakukan secara lebih efektif dan tepat sasaran (Saragih & Situmorang, 2020).

Pengelolaan data hasil panen yang masih dilakukan secara manual dan semi-digital berpotensi menimbulkan kesalahan pencatatan, keterlambatan rekapitulasi data, serta kesulitan dalam melakukan analisis historis secara menyeluruh. Kondisi tersebut menghambat koperasi dalam mengidentifikasi pola produktivitas mandor secara objektif (Hilmy et al., 2021).

Seiring meningkatnya volume dan kompleksitas data hasil panen, data multivariat seperti luas lahan, jumlah panen, frekuensi panen, dan volume distribusi tidak dapat dianalisis secara optimal tanpa dukungan sistem informasi terintegrasi. Sistem informasi mampu menyediakan penyimpanan terpusat sekaligus fasilitas analisis data untuk mendukung pengambilan keputusan (Saragih & Situmorang, 2020).

Dalam data mining, teknik clustering digunakan untuk mengelompokkan data berdasarkan tingkat kemiripan karakteristik tanpa memerlukan label kelas sebelumnya. Algoritma K-Means banyak digunakan karena kesederhanaan dan

efisiensinya dalam mengelompokkan data pertanian berbasis produktivitas (Simangunsong et al., 2022).

Salah satu metode yang dapat digunakan adalah K-Means Clustering, yang mampu mengelompokkan data berdasarkan tingkat kemiripan karakteristik. Oleh karena itu, penelitian ini bertujuan untuk merancang dan membangun sistem informasi berbasis web yang mengimplementasikan metode K-Means Clustering dalam menganalisis dan mengelompokkan data produktivitas dan distribusi hasil panen, sehingga dapat menghasilkan informasi yang lebih akurat dan mendukung pengambilan keputusan. Oleh karena itu, penelitian ini melakukan analisis pola produktivitas dan distribusi hasil panen kelapa sawit di KUD Pelita Desa Silau Jawa menggunakan metode K-Means Clustering, dengan harapan dapat memberikan dasar pengambilan keputusan yang lebih objektif dan berbasis data (Laksono et al., 2024).

1.2. Rumusan Masalah

Berdasarkan uraian latar belakang di atas, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana merancang sistem informasi yang mampu mengelola data produktivitas dan distribusi hasil panen kelapa sawit di KUD Pelita Desa Silau Jawa?
2. Bagaimana merancang sistem yang dapat mengimplementasikan metode K-Means Clustering untuk mengelompokkan mandor berdasarkan produktivitas dan distribusi hasil panen?

3. Bagaimana menampilkan hasil pengelompokan mandor dalam bentuk informasi yang mudah dipahami dan mendukung pengambilan keputusan koperasi?

1.3. Batasan Masalah

Agar penelitian lebih terfokus, batasan masalah ditetapkan sebagai berikut:

1. Sistem yang dirancang merupakan sistem informasi berbasis web dengan satu aktor utama, yaitu admin (pengurus KUD).
2. Data yang digunakan berupa data produktivitas dan distribusi hasil panen kelapa sawit mandor anggota KUD Pelita Desa Silau Jawa.
3. Variabel yang digunakan dalam sistem meliputi luas lahan, jumlah hasil panen, frekuensi panen, dan volume distribusi hasil panen.
4. Metode analisis yang diimplementasikan dalam sistem adalah K-Means Clustering tanpa perbandingan dengan metode clustering lainnya.
5. Penelitian difokuskan pada perancangan sistem dan pemodelan, tanpa membahas implementasi lanjutan secara penuh di lingkungan produksi.

1.4. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Merancang sistem informasi untuk pengelolaan data produktivitas dan distribusi hasil panen kelapa sawit di KUD Pelita Desa Silau Jawa.
2. Merancang sistem yang mampu mengimplementasikan metode K-Means Clustering dalam mengelompokkan mandor berdasarkan karakteristik produktivitas dan distribusi.

3. Menghasilkan rancangan sistem yang dapat menyajikan informasi hasil clustering sebagai dasar pengambilan keputusan koperasi.

1.5. Manfaat Penelitian

Berdasarkan tujuan dan ruang lingkup penelitian yang dilakukan, maka manfaat dari penelitian ini adalah sebagai berikut:

1. Manfaat bagi Universitas dan Fakultas

Penelitian ini diharapkan dapat memberikan kontribusi akademik bagi universitas dan fakultas, khususnya dalam pengembangan keilmuan di bidang sistem informasi dan analisis data. Hasil penelitian ini dapat menjadi referensi ilmiah terkait penerapan metode data mining, khususnya algoritma K-Means Clustering, dalam pengolahan dan analisis data pertanian kelapa sawit. Selain itu, penelitian ini dapat menambah koleksi karya ilmiah di lingkungan fakultas serta menjadi bahan rujukan bagi penelitian selanjutnya yang memiliki topik serupa.

2. Manfaat bagi Mitra

Bagi mitra penelitian, hasil penelitian ini diharapkan dapat membantu dalam memahami pola produktivitas dan distribusi hasil panen kelapa sawit berdasarkan pengelompokan data yang dihasilkan. Informasi klaster yang terbentuk dapat digunakan sebagai bahan pertimbangan dalam pengambilan keputusan berbasis data, seperti perencanaan produksi, evaluasi kinerja, dan pengelolaan sumber daya secara lebih efektif dan terarah. Dengan demikian, mitra dapat memperoleh gambaran yang lebih jelas mengenai kondisi data panen yang dimiliki.

3. Manfaat bagi Mahasiswa

Penelitian ini memberikan manfaat bagi mahasiswa sebagai sarana untuk menerapkan teori yang diperoleh selama perkuliahan ke dalam bentuk penelitian ilmiah. Mahasiswa dapat meningkatkan pemahaman dan keterampilan dalam menerapkan metode K-Means Clustering, pengolahan data, serta analisis hasil penelitian secara sistematis. Selain itu, penelitian ini dapat menjadi pengalaman akademik yang bermanfaat dalam meningkatkan kemampuan berpikir kritis dan analitis di bidang sistem informasi dan analisis data.

BAB II

LANDASAN TEORI

2.1. Tanaman Kelapa Sawit

Kelapa sawit (*Elaeis guineensis*) merupakan tanaman perkebunan tahunan yang memiliki siklus produksi jangka panjang dan mulai menghasilkan buah secara ekonomis pada umur sekitar 3–4 tahun setelah tanam. Produktivitas tanaman sawit meningkat seiring bertambahnya umur tanaman hingga mencapai puncaknya pada kisaran umur 8–15 tahun, sebelum kemudian mengalami penurunan secara bertahap.

Dalam konteks perkebunan rakyat, karakteristik tanaman kelapa sawit sangat dipengaruhi oleh kondisi lahan, perawatan kebun, serta pola panen yang diterapkan oleh mandor. Penelitian-penelitian terkini menunjukkan bahwa meskipun berada dalam satu wilayah koperasi yang sama, perbedaan perlakuan budidaya dapat menghasilkan variasi produksi yang cukup signifikan antar mandor (Putra & Wijaya, 2022). Kondisi ini menyebabkan data hasil panen bersifat heterogen dan memerlukan pendekatan analisis yang mampu menangkap pola perbedaan tersebut.

2.2. Produktivitas Hasil Panen Kelapa Sawit

Produktivitas hasil panen kelapa sawit pada perkebunan rakyat umumnya dinyatakan dalam satuan ton Tandan Buah Segar (TBS) per hektar per tahun (ton/ha/tahun). Berdasarkan hasil penelitian pada periode 2020–2025, produktivitas kebun kelapa sawit rakyat di Indonesia berada pada kisaran 15–25 ton TBS/ha/tahun, dengan variasi yang dipengaruhi oleh umur tanaman, teknik pemeliharaan, serta sistem manajemen panen yang diterapkan oleh mandor.

Rentang nilai ini digunakan sebagai acuan teoritis untuk menggambarkan kondisi umum produktivitas perkebunan rakyat dan tidak dimaksudkan sebagai klaim terhadap objek penelitian secara spesifik (Sihite & Prasetyaningrum, 2023).

Apabila dikaitkan dengan luas lahan, maka total produksi tahunan mandor dapat dijelaskan secara konseptual. Mandor dengan luas lahan sekitar 2 hektar berpotensi menghasilkan 30–50 ton TBS per tahun, sedangkan mandor dengan luas lahan sekitar 5 hektar berpotensi menghasilkan 75–125 ton TBS per tahun, dengan asumsi produktivitas berada pada kisaran yang wajar. Variasi produksi tersebut menunjukkan bahwa luas lahan dan tingkat produktivitas merupakan variabel kuantitatif yang saling berkaitan dan relevan untuk dianalisis dalam pemetaan pola hasil panen (Oktariani, 2025; Putra & Wijaya, 2022).

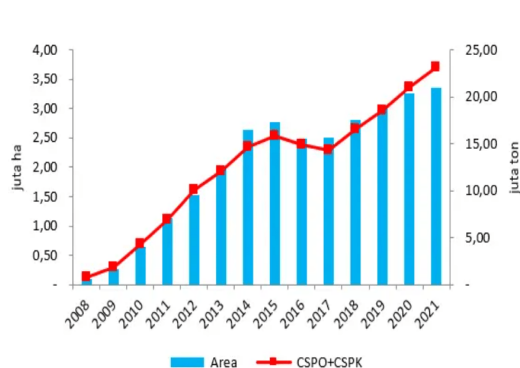
Selain luas lahan dan jumlah hasil panen, frekuensi panen juga berperan penting dalam menentukan tingkat produktivitas. Pada praktik perkebunan kelapa sawit rakyat, kegiatan panen umumnya dilakukan setiap 7 hingga 14 hari, sehingga dalam satu tahun frekuensi panen dapat berkisar antara 24–48 kali. Perbedaan frekuensi panen ini berdampak pada volume TBS yang dihasilkan serta kualitas buah yang dipanen, sehingga frekuensi panen layak digunakan sebagai salah satu variabel dalam analisis produktivitas (Sihite & Prasetyaningrum, 2023).

Distribusi hasil panen kelapa sawit merupakan proses penyaluran TBS dari kebun mandor ke koperasi dan selanjutnya ke pabrik kelapa sawit. Pada tingkat koperasi, volume distribusi hasil panen umumnya dinyatakan dalam satuan ton per bulan atau ton per tahun. Penelitian menunjukkan bahwa mandor dengan skala lahan kecil hingga menengah umumnya mendistribusikan sekitar 2–10 ton TBS per bulan, sedangkan mandor dengan luas lahan yang lebih besar dapat

mendistribusikan hingga 20–40 ton TBS per bulan, tergantung pada intensitas panen dan kapasitas angkut koperasi (Mawaddah et al., 2022).

Secara kumulatif, volume distribusi hasil panen dalam satu tahun dapat mencapai puluhan hingga ratusan ton TBS per mandor, sejalan dengan total produksi tahunan. Namun demikian, volume distribusi tidak selalu identik dengan jumlah produksi karena dipengaruhi oleh faktor logistik, seperti jadwal pengangkutan, waktu tunggu, serta manajemen distribusi koperasi. Oleh karena itu, analisis distribusi perlu dilakukan secara terpisah namun tetap dikaitkan dengan data produktivitas untuk memperoleh gambaran yang lebih komprehensif (Simangunsong et al., 2022a).

Berdasarkan uraian tersebut, variabel kuantitatif berupa luas lahan (hektar), jumlah hasil panen (ton/tahun), frekuensi panen (kali/tahun), dan volume distribusi (ton/tahun) merupakan indikator yang realistis, terukur, dan relevan untuk dianalisis menggunakan metode K-Means Clustering. Penggunaan rentang nilai yang wajar memungkinkan penelitian ini mengidentifikasi pola produktivitas dan distribusi hasil panen tanpa melakukan generalisasi yang melampaui data lapangan (Oktariani, 2025).



Gambar. 2.2 Perkembangan Luas Areal Dan Produksi Minyak Sawit Tersertifikasi Berkelanjutan (sumber, RSPO, 2022)

2.3. Distribusi Hasil Panen Kelapa Sawit

Distribusi hasil panen kelapa sawit merupakan tahapan lanjutan setelah proses panen yang berperan penting dalam menjaga mutu TBS. TBS yang telah dipanen idealnya harus segera dikirim ke pabrik kelapa sawit dalam waktu kurang dari 24 jam untuk mencegah penurunan kualitas. Pada tingkat koperasi, distribusi hasil panen biasanya diukur berdasarkan volume TBS yang diangkut dan disalurkan (ton/bulan atau ton/tahun) (Simangunsong et al., 2022).

Dalam praktiknya, volume distribusi hasil panen mandor anggota koperasi dapat berkisar antara 2–10 ton TBS per bulan untuk mandor dengan skala kecil, dan dapat mencapai 20–40 ton TBS per bulan untuk mandor dengan luas lahan yang lebih besar. Perbedaan volume distribusi ini mencerminkan variasi produktivitas sekaligus kemampuan logistik koperasi dalam mengatur jadwal angkut (Mawaddah et al., 2022).

Ketidakseimbangan antara volume hasil panen dan kapasitas distribusi dapat menimbulkan berbagai permasalahan operasional, seperti penumpukan TBS, keterlambatan pengangkutan, dan penurunan kualitas buah. Oleh karena itu, pemetaan pola distribusi berdasarkan data historis menjadi penting untuk mendukung perencanaan distribusi yang lebih efisien dan tepat sasaran (Nasution & Ningsih, 2025).

2.4. Variabel Penelitian dalam Analisis Produktivitas dan Distribusi

Berdasarkan kajian teori dan penelitian terdahulu, variabel yang digunakan dalam penelitian ini terdiri dari luas lahan, jumlah hasil panen Tandan Buah Segar (TBS), frekuensi panen, dan volume distribusi hasil panen. Keempat variabel tersebut dipilih karena secara kuantitatif mampu merepresentasikan kondisi

produktivitas dan distribusi hasil panen mandor kelapa sawit pada tingkat koperasi. Variabel-variabel ini juga telah digunakan dalam penelitian sejenis dan terbukti relevan dalam analisis clustering data hasil panen kelapa (Putra & Wijaya, 2022; Sihite & Prasetyaningrum, 2023)

2.5. Konsep Data Mining

Data mining merupakan pendekatan analitis yang digunakan untuk mengekstraksi informasi dan pola penting dari kumpulan data berukuran besar dan kompleks. Dalam sektor pertanian, data mining digunakan untuk menganalisis data produksi, distribusi, dan karakteristik lahan secara objektif, sehingga dapat mendukung pengambilan keputusan berbasis data (Wulandari, 2021a).

Penerapan data mining sangat relevan pada data hasil panen kelapa sawit karena data tersebut bersifat numerik, multivariat, dan terus bertambah setiap periode panen. Pendekatan konvensional yang hanya mengandalkan rekapitulasi data sering kali tidak mampu mengungkap pola produktivitas dan distribusi secara menyeluruh (Hasdyna et al., 2025).

2.6. Clustering Sebagai Teknik Analisis Data

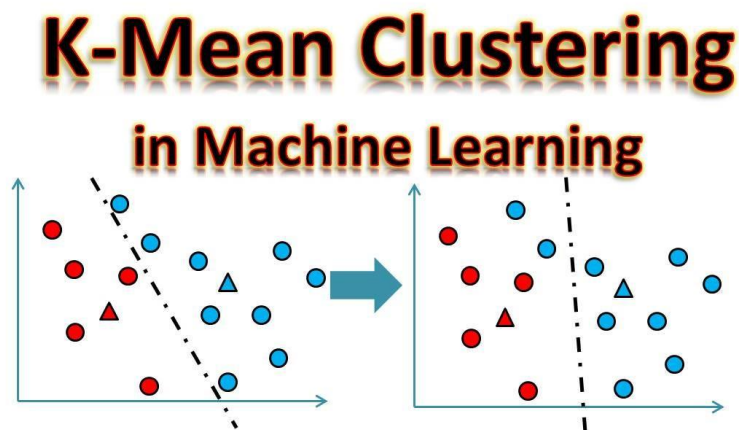
Clustering merupakan teknik dalam data mining yang bertujuan untuk mengelompokkan data ke dalam beberapa kelompok berdasarkan tingkat kemiripan karakteristik. Teknik ini bersifat unsupervised, sehingga tidak memerlukan label awal dan cocok digunakan untuk eksplorasi data hasil panen yang heterogen (Oktariani, 2025).

Dalam konteks produktivitas kelapa sawit, clustering memungkinkan pengelompokan mandor ke dalam kategori seperti produktif tinggi, sedang, dan rendah berdasarkan kombinasi variabel luas lahan, jumlah hasil panen, frekuensi panen, dan volume distribusi. Pendekatan ini membantu koperasi dalam memahami karakteristik anggota secara lebih objektif (Putra, 2022).

2.7. Algoritma K-Means

Algoritma K-Means merupakan metode clustering yang membagi data ke dalam K cluster berdasarkan kedekatan jarak dengan pusat cluster (centroid). K-Means banyak digunakan dalam penelitian pertanian karena efisiensi komputasi dan kemudahan interpretasi hasilnya (Siregar, 2020).

Dalam analisis produktivitas dan distribusi hasil panen kelapa sawit, K-Means mampu mengelompokkan mandor berdasarkan pola data numerik yang kompleks. Hasil cluster dapat digunakan untuk mengidentifikasi kelompok mandor dengan produksi tinggi namun distribusi rendah, atau sebaliknya, sehingga mendukung penyusunan strategi pengelolaan koperasi berbasis data (Gultom et al., 2025).



Gambar 2.7 K-Mean Clustering

2.8. Perhitungan Jarak Euclidean dalam K-Means

Jarak Euclidean merupakan metode pengukuran jarak yang paling umum digunakan dalam algoritma K-Means untuk menentukan tingkat kedekatan antara suatu data dengan pusat cluster (centroid). Jarak ini menghitung selisih nilai antar variabel numerik sehingga dapat menggambarkan tingkat kemiripan antar data secara kuantitatif. Dalam penelitian ini, variabel yang digunakan meliputi luas lahan (hektar), jumlah hasil panen (ton/tahun), frekuensi panen (kali/tahun), dan volume distribusi (ton/tahun) (Hasdyna et al., 2025).

Secara matematis, jarak Euclidean antara dua titik data dapat dirumuskan sebagai berikut:

$$d(x, y) = \sqrt{\{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2\}}$$

di mana:

4. $d(x, y)$ = jarak Euclidean antara data xxx dan data yyy
5. x_i = nilai variabel ke- i pada data xxx
6. y_i = nilai variabel ke- i pada data yyy
7. n = jumlah variabel yang digunakan

Dalam konteks algoritma K-Means, Jika data ke- i dinyatakan sebagai $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$ dan centroid cluster ke- k dinyatakan sebagai $c_k = (c_{k1}, c_{k2}, \dots, c_{kn})$, maka jarak Euclidean antara data dan centroid dirumuskan sebagai berikut:

$$d(x_i, c_k) = \sqrt{\left\{ \sum_{j=1}^n (x_{ij} - c_{kj})^2 \right\}}$$

$$d = \sqrt{((X^1 - C^1)^2 + (X^2 - C^2)^2 + (X^3 - C^3)^2 + (X^4 - C^4)^2)}$$

Pada penelitian ini, rumus tersebut dapat diturunkan secara spesifik sesuai dengan jumlah variabel yang digunakan, yaitu empat variabel, sehingga persamaan jarak Euclidean menjadi:

dengan keterangan:

X_1 = luas lahan (hektar)

X_2 = jumlah hasil panen (ton/tahun)

X_3 = frekuensi panen (kali/tahun)

X_4 = volume distribusi hasil panen (ton/tahun)

C_1, C_2, C_3, C_4 = nilai centroid pada masing-masing variabel

Nilai jarak yang dihasilkan dari perhitungan tersebut digunakan untuk menentukan keanggotaan cluster, di mana setiap data akan dimasukkan ke dalam cluster yang memiliki jarak Euclidean terkecil terhadap centroid. Proses ini dilakukan secara berulang hingga tidak terjadi perubahan signifikan pada posisi centroid atau keanggotaan cluster.

Penggunaan jarak Euclidean memungkinkan proses clustering dilakukan secara objektif dan konsisten karena seluruh variabel diukur berdasarkan jarak numerik yang sama. Dengan demikian, hasil pengelompokan yang diperoleh dapat merepresentasikan perbedaan karakteristik data produktivitas dan distribusi hasil panen kelapa sawit secara kuantitatif dan sistematis (Mawaddah et al., 2022).

2.9. Normalisasi Data

Normalisasi data merupakan tahapan penting dalam analisis clustering karena variabel yang digunakan memiliki satuan dan rentang nilai yang berbeda. Dalam penelitian ini, variabel luas lahan diukur dalam satuan hektar, sedangkan jumlah hasil panen dan volume distribusi diukur dalam satuan ton, serta frekuensi panen

diukur dalam satuan kali per tahun. Perbedaan skala tersebut dapat menyebabkan variabel dengan nilai numerik yang lebih besar mendominasi perhitungan jarak apabila tidak dilakukan normalisasi, sehingga mempengaruhi hasil clustering (Wulandari, 2021b).

Salah satu metode normalisasi yang umum digunakan adalah Min–Max Normalization. Metode ini berfungsi untuk mentransformasikan nilai data asli ke dalam rentang tertentu, umumnya antara 0 hingga 1, sehingga seluruh variabel memiliki skala yang sebanding sebelum dilakukan perhitungan jarak dalam algoritma K-Means. Rumus Min–Max Normalization dapat dituliskan sebagai berikut:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

X' = nilai data hasil normalisasi

X = nilai data asli

X_{min} = nilai minimum dari suatu variabel

X_{max} = nilai maksimum dari suatu variabel

Melalui rumus tersebut, setiap nilai data akan ditransformasikan secara proporsional ke dalam rentang 0 hingga 1. Nilai minimum suatu variabel akan dipetakan menjadi 0, sedangkan nilai maksimum akan dipetakan menjadi 1, dan nilai lainnya akan berada di antara kedua nilai tersebut sesuai dengan posisinya dalam rentang data.

Dalam penerapan algoritma K-Means, normalisasi data menggunakan metode Min–Max bertujuan untuk memastikan bahwa setiap variabel, seperti luas lahan, jumlah hasil panen, frekuensi panen, dan volume distribusi, memiliki kontribusi

yang seimbang dalam perhitungan jarak Euclidean. Dengan demikian, proses pengelompokan data dapat dilakukan secara lebih objektif dan tidak bias terhadap variabel tertentu yang memiliki nilai numerik lebih besar.

Penelitian terdahulu menunjukkan bahwa penerapan normalisasi data sebelum proses clustering dapat meningkatkan stabilitas hasil pengelompokan serta menghasilkan cluster yang lebih representatif terhadap kondisi data sebenarnya. Oleh karena itu, normalisasi data merupakan tahapan yang tidak dapat dipisahkan dari proses analisis clustering menggunakan metode K-Means (Hasdyna et al., 2025).

2.10. Studi Terdahulu (Penelitian Relevan)

Penelitian (Mawaddah et al., 2022) melakukan analisis pengelompokan data hasil panen kelapa sawit menggunakan algoritma K-Means Clustering dengan variabel utama berupa jumlah produksi Tandan Buah Segar (TBS). Penelitian tersebut bertujuan untuk mengelompokkan data hasil panen ke dalam beberapa cluster berdasarkan tingkat produksi, sehingga dapat diidentifikasi kelompok mandor dengan produksi tinggi, sedang, dan rendah. Hasil penelitian menunjukkan bahwa algoritma K-Means mampu mengelompokkan data secara efektif dan membantu dalam memahami pola distribusi tingkat produksi kelapa sawit. Meskipun demikian, penelitian tersebut masih berfokus pada satu indikator utama, yaitu jumlah hasil panen, tanpa mempertimbangkan variabel lain seperti luas lahan, frekuensi panen, dan volume distribusi (Mawaddah et al., 2022).

(Sihite & Prasetyaningrum, 2023) melakukan penelitian clustering pada lahan perkebunan kelapa sawit dengan tujuan memetakan tingkat produktivitas lahan. Variabel yang digunakan dalam penelitian ini meliputi luas lahan dan tingkat hasil

produksi, yang dianalisis menggunakan metode K-Means Clustering. Hasil penelitian menunjukkan bahwa metode clustering dapat digunakan untuk membedakan lahan produktif dan kurang produktif secara objektif berdasarkan data numerik. Penelitian ini menegaskan bahwa pendekatan clustering bersifat efektif untuk analisis produktivitas kelapa sawit, namun belum memasukkan aspek frekuensi panen dan distribusi hasil panen sebagai bagian dari analisis (Sihite & Prasetyaningrum, 2023).

Penelitian Oktariani (2025) memperluas penerapan algoritma K-Means Clustering pada skala yang lebih luas, yaitu pengelompokan wilayah produksi kelapa sawit. Variabel yang digunakan meliputi luas areal, jumlah produksi tahunan, dan tingkat produktivitas wilayah. Hasil penelitian menunjukkan bahwa clustering mampu menggambarkan perbedaan karakteristik produksi antar kelompok wilayah secara jelas dan sistematis. Penelitian ini membuktikan bahwa K-Means tidak hanya efektif pada tingkat mandor atau lahan, tetapi juga relevan digunakan pada analisis produktivitas dalam skala yang lebih besar. Namun demikian, penelitian tersebut lebih berfokus pada data agregat wilayah dan belum mengkaji secara spesifik aspek distribusi hasil panen pada tingkat koperasi (Oktariani, 2025).

Berdasarkan kajian teori dan penelitian terdahulu, dapat disimpulkan bahwa algoritma K-Means Clustering telah banyak digunakan dalam analisis produktivitas kelapa sawit. Namun, sebagian besar penelitian sebelumnya masih berfokus pada variabel produksi atau luas lahan saja, serta belum mengintegrasikan variabel frekuensi panen dan volume distribusi hasil panen secara bersamaan pada tingkat koperasi. Oleh karena itu, penelitian ini memiliki

posisi yang berbeda dengan mengombinasikan empat variabel utama untuk menganalisis pola produktivitas dan distribusi hasil panen kelapa sawit secara lebih komprehensif.

No	Nama Peneliti	Tahun	Metode	Variabel yang Digunakan	Hasil Penelitian	Kelemahan Penelitian
1	Mawaddah et al.	2022	K-Means Clustering	Jumlah hasil panen TBS	Algoritma K-Means mampu mengelompokkan data hasil panen kelapa sawit ke dalam beberapa cluster berdasarkan tingkat produksi sehingga memudahkan identifikasi kelompok produksi tinggi, sedang, dan rendah	Penelitian hanya menggunakan variabel jumlah hasil panen, belum mempertimbangkan luas lahan, frekuensi panen, dan aspek distribusi hasil panen
2	Sihite & Prasetyaningrum	2023	K-Means Clustering	Luas lahan dan hasil produksi	Clustering efektif dalam memetakan lahan perkebunan kelapa sawit menjadi kategori produktif dan kurang	Tidak memasukkan variabel frekuensi panen dan volume distribusi, serta belum dikaji pada tingkat koperasi

					produktif secara objektif berdasarkan data numerik	
3	Putra et al.	2022	Data Mining (K-Means)	Luas lahan dan produktivitas kebun	Metode K-Means dapat digunakan untuk menganalisis perbedaan produktivitas kebun kelapa sawit rakyat berdasarkan karakteristik lahan	Analisis masih terbatas pada produktivitas, belum mengaitkan hasil panen dengan pola distribusi
4	Oktariani	2025	K-Means Clustering	Luas areal, jumlah produksi, produktivitas wilayah	Clustering mampu menggambarkan perbedaan karakteristik produksi kelapa sawit antar wilayah secara jelas dan sistematis	Data yang digunakan bersifat agregat wilayah, belum merepresentasikan kondisi mandor atau koperasi secara spesifik

5	Hasdyna et al.	2025	K-Means dengan normalisasi data	Data produksi kelapa sawit multivariante	Normalisasi data meningkatkan stabilitas dan kualitas hasil clustering pada data produksi kelapa sawit	Fokus penelitian pada optimalisasi metode, belum mengaitkan hasil cluster dengan aspek distribusi hasil panen
---	----------------	------	---------------------------------	--	--	---

Tabel 2.10 Penelitian Terdahulu

2.11. Hipotesis Penelitian

Penelitian ini bersifat eksploratif dengan menggunakan teknik clustering yang tidak memerlukan label kelas awal. Oleh karena itu, hipotesis yang diajukan berupa dugaan awal untuk mengarahkan analisis, yaitu:

Terdapat pola pengelompokan mandor kelapa sawit berdasarkan variabel produktivitas dan distribusi hasil panen menggunakan metode K-Means Clustering (Oktariani, 2025; Wulandari, 2021b).

2.12. Kerangka Pemikiran

Produktivitas dan distribusi hasil panen kelapa sawit merupakan dua komponen utama yang menentukan keberhasilan pengelolaan perkebunan rakyat, khususnya pada tingkat koperasi. Pada praktiknya, setiap mandor memiliki kondisi lahan, hasil panen, serta pola panen yang berbeda-beda. Perbedaan tersebut secara langsung memengaruhi jumlah Tandan Buah Segar (TBS) yang dihasilkan dan volume hasil panen yang didistribusikan ke koperasi. Ketidakteraturan pola produktivitas dan distribusi ini sering kali menjadi kendala

dalam upaya peningkatan efisiensi pengelolaan koperasi (Sihite & Prasetyaningrum, 2023; Simangunsong et al., 2022b).

KUD Pelita Desa Silau Jawa sebagai lembaga yang menaungi mandor kelapa sawit rakyat mengelola data hasil panen dari banyak anggota dengan karakteristik yang beragam. Data tersebut meliputi luas lahan yang dimiliki mandor, jumlah produksi TBS per tahun, frekuensi panen, serta volume distribusi hasil panen. Selama ini, pengolahan data masih didominasi oleh pencatatan dan rekapitulasi sederhana, sehingga belum mampu memberikan gambaran yang jelas mengenai pola produktivitas dan distribusi hasil panen secara menyeluruh (Mawaddah et al., 2022).

Seiring dengan meningkatnya jumlah data hasil panen yang dikelola koperasi, diperlukan pendekatan analisis yang tidak hanya mampu mengolah data dalam jumlah besar, tetapi juga dapat mengungkap pola-pola penting yang tersembunyi di dalamnya. Pendekatan data mining dinilai relevan karena mampu mengolah data numerik secara sistematis dan menghasilkan informasi yang dapat digunakan sebagai dasar pengambilan keputusan. Dalam bidang pertanian dan perkebunan, data mining telah banyak dimanfaatkan untuk menganalisis data produksi dan distribusi hasil panen (Hasdyna et al., 2025; Wulandari, 2021a).

Salah satu teknik dalam data mining yang sesuai untuk menganalisis data hasil panen kelapa sawit adalah clustering. Teknik clustering bertujuan untuk mengelompokkan data ke dalam beberapa kelompok berdasarkan kesamaan karakteristik, tanpa memerlukan klasifikasi awal. Dengan sifatnya yang *unsupervised*, clustering sangat sesuai diterapkan pada data hasil panen yang memiliki variasi tinggi antar mandor (Oktariani, 2025).

Dalam penelitian ini, algoritma K-Means Clustering digunakan karena memiliki kemampuan mengelompokkan data numerik secara efisien serta mudah dipahami hasilnya. K-Means banyak digunakan dalam penelitian penelitian sebelumnya untuk mengelompokkan data produksi pertanian, termasuk data hasil panen kelapa sawit, sehingga memudahkan peneliti dalam mengidentifikasi kelompok dengan tingkat produktivitas dan distribusi yang berbeda (Siregar, 2020).

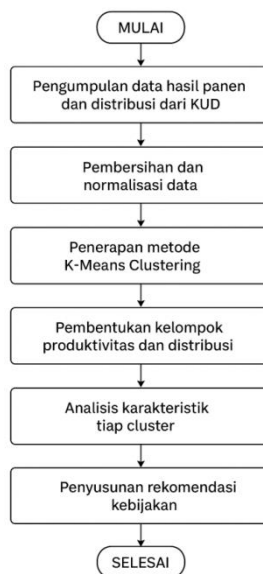
Proses clustering dalam penelitian ini menggunakan empat variabel utama, yaitu luas lahan (X1), jumlah hasil panen Tandan Buah Segar per tahun (X2), frekuensi panen per tahun (X3), dan volume distribusi hasil panen per tahun (X4). Keempat variabel tersebut dipilih karena dianggap mampu merepresentasikan kondisi produktivitas dan distribusi hasil panen mandor secara kuantitatif. Penggunaan variabel-variabel ini telah diterapkan dalam berbagai penelitian serupa dan terbukti menghasilkan pengelompokan yang sesuai dengan kondisi lapangan (Putra & Wijaya, 2022).

Sebelum dilakukan proses clustering, data hasil panen terlebih dahulu melalui tahap pembersihan dan normalisasi. Tahap ini bertujuan untuk memastikan data yang digunakan bebas dari kesalahan pencatatan serta memiliki skala yang sebanding antar variabel. Normalisasi diperlukan karena perbedaan satuan pengukuran, seperti hektar dan ton, dapat memengaruhi hasil perhitungan jarak dalam algoritma K-Means. Penelitian sebelumnya menunjukkan bahwa normalisasi data dapat meningkatkan kualitas hasil clustering (Hasdyna et al., 2025).

Hasil dari penerapan algoritma K-Means Clustering berupa sejumlah cluster yang menggambarkan kelompok mandor dengan karakteristik produktivitas dan distribusi yang relatif serupa. Setiap cluster dianalisis berdasarkan nilai pusat cluster (centroid) untuk mengetahui ciri khas masing-masing kelompok. Melalui analisis ini, koperasi dapat mengetahui kelompok mandor dengan tingkat produktivitas tinggi, sedang, maupun rendah, serta perbedaan volume distribusi hasil panen pada setiap kelompok (Mawaddah et al., 2022).

Informasi yang diperoleh dari hasil pengelompokan tersebut selanjutnya digunakan sebagai dasar dalam penyusunan rekomendasi pengelolaan produktivitas dan distribusi hasil panen kelapa sawit. Rekomendasi ini diharapkan dapat membantu KUD Pelita Desa Silau Jawa dalam meningkatkan efektivitas pengelolaan koperasi, khususnya dalam perencanaan distribusi hasil panen dan pembinaan mandor berdasarkan karakteristik masing-masing kelompok (Simangunsong et al., 2022b).

Kerangka pemikiran dalam penelitian ini menggambarkan alur pemecahan masalah sebagai berikut:



Gambar 2. 11 Alur Pemecahan Masalah

Keterangan:

1. Pengumpulan data hasil panen dan distribusi dari KUD

Pada tahap ini, data yang dibutuhkan dikumpulkan dari KUD Pelita, seperti jumlah hasil panen TBS, frekuensi panen, luas lahan, dan volume distribusi. Data inilah yang menjadi dasar proses analisis.

2. Pembersihan dan normalisasi data

Data yang telah dikumpulkan diperiksa untuk memastikan tidak ada kesalahan, data ganda, atau data kosong. Setelah itu dilakukan normalisasi agar semua variabel berada pada skala yang sama sehingga analisis menjadi lebih akurat.

3. Penerapan metode K-Means Clustering

Pada tahap ini, algoritma K-Means digunakan untuk mengelompokkan data mandor berdasarkan kesamaan karakteristik produktivitas dan distribusi. Proses ini menghasilkan cluster atau kelompok tertentu.

4. Pembentukan kelompok produktivitas dan distribusi

Dari hasil K-Means, terbentuk beberapa kelompok mandor, misalnya kelompok produktivitas tinggi, sedang, dan rendah. Kelompok ini menunjukkan perbedaan performa panen dan distribusi antar mandor.

5. Analisis karakteristik tiap cluster

Setiap cluster yang terbentuk dianalisis untuk melihat ciri-ciri khasnya, seperti seberapa tinggi produktivitasnya, seberapa sering panen dilakukan, atau berapa besar distribusi yang dikirim ke KUD.

6. Penyusunan rekomendasi kebijakan

Berdasarkan hasil analisis cluster, dibuat rekomendasi yang dapat membantu KUD dalam meningkatkan produktivitas, memperbaiki distribusi, dan mengambil keputusan yang lebih tepat berdasarkan data.

Dengan demikian, kerangka pemikiran penelitian ini menempatkan data produktivitas dan distribusi hasil panen sebagai input utama yang dianalisis menggunakan metode K-Means Clustering. Hasil pengelompokan berupa cluster mandor dengan karakteristik yang relatif homogen digunakan sebagai dasar dalam memahami perbedaan tingkat produktivitas dan distribusi, serta sebagai landasan penyusunan rekomendasi pengelolaan koperasi berbasis data.

BAB III

ANALISA DAN PERANCANGAN SISTEM

3.1. Analisis Permasalahan

KOP "Pelita Maju Mandiri" Desa Silau Jawa, Kabupaten Asahan, Provinsi Sumatera Utara berperan sebagai lembaga pengelola hasil panen kelapa sawit dari mandor dan kelompok tani anggota koperasi. Berdasarkan hasil observasi lapangan, wawancara dengan pengurus koperasi, serta telaah dokumen rekap produksi hasil panen periode Januari 2024 hingga Desember 2025, ditemukan beberapa permasalahan utama dalam pengelolaan data produktivitas dan distribusi hasil panen.

3.1.1. Permasalahan pada Pengelolaan Data

Permasalahan pertama terletak pada sistem pencatatan data yang masih bersifat manual berupa buku besar rekap bulanan. Data rekap produksi hasil panen dicatat dalam tabel manual per bulan tanpa sistem basis data terpusat. Hal ini menimbulkan beberapa dampak operasional, antara lain:

- a. Tingginya risiko kesalahan pencatatan (human error) pada saat rekapitulasi data bulanan.
- b. Sulitnya proses pencarian dan analisis data historis panen secara lintas periode.
- c. Belum adanya sistem yang mampu mengolah dan menyajikan data secara otomatis kepada pengurus koperasi.

3.1.2. Permasalahan pada Analisis Produktivitas

KUD belum memiliki metode analisis terstruktur untuk mengelompokkan mandor/kelompok berdasarkan tingkat produktivitas lahan dan hasil panen. Berdasarkan data rekap yang tersedia, terdapat variasi tonase yang cukup signifikan antar kelompok, misalnya produksi kelompok HARAPAN (42 Ha) mencapai 59.400 KG pada Maret 2025 sementara kelompok LESTARI (76 Ha) mencapai 246.797 KG pada September 2025. Tanpa pengelompokan berbasis data, kondisi ini menyebabkan:

- a. Pembinaan mandor tidak tepat sasaran karena tidak ada identifikasi kelompok produktif dan kurang produktif.
- b. Tidak ada prioritas evaluasi dan pengembangan berbasis data objektif.

3.1.3. Permasalahan pada Distribusi Hasil Panen

Data langsir (frekuensi pengiriman TBS ke pabrik) menunjukkan variasi yang besar antar kelompok. Kelompok SADAR mencatat langsir hingga 92 kali pada November 2025, sementara beberapa kelompok seperti MAKMUR dan SEDERHANA mencatat langsir 0 pada beberapa bulan. Ketidakseimbangan ini menyebabkan:

- a. Tidak adanya jadwal distribusi terstruktur berdasarkan volume produksi masing-masing kelompok.
- b. Keterlambatan pengangkutan TBS ke pabrik kelapa sawit pada kelompok dengan volume tinggi.
- c. Sulitnya evaluasi efisiensi distribusi tanpa pengelompokan pola langsir berbasis data.

3.2. Analisis Kebutuhan Sistem

Analisis kebutuhan dilakukan untuk menentukan fungsi dan batasan sistem yang akan dikembangkan sebagai dasar perancangan basis data dan antarmuka sistem.

3.2.1 Kebutuhan Fungsional

Kebutuhan fungsional menggambarkan fitur utama yang harus dapat dijalankan oleh sistem, meliputi:

- a. Mengelola data mandor/kelompok: kode mandor, nama mandor, nama kelompok, dan luas lahan.
- b. Mengelola data produksi panen per periode bulanan: tonase (KG), rata-rata KG/Ha, dan langsir.
- c. Melakukan normalisasi data menggunakan metode Min-Max sebelum proses analisis.
- d. Menjalankan proses clustering dengan metode K-Means (pilihan K=2 hingga K=5).
- e. Sistem mampu menghitung dan menampilkan nilai Sum of Squared Error (SSE) sebagai indikator kualitas hasil pengelompokan data.
- f. Sistem mampu menyimpan dan menampilkan nilai centroid pada setiap iterasi proses K-Means Clustering hingga mencapai kondisi konvergen.
- g. Sistem mampu menghasilkan rata-rata nilai setiap variabel pada masing-masing cluster sebagai dasar interpretasi hasil pengelompokan.
- h. Sistem mampu menampilkan hasil clustering dalam bentuk tabel dan visualisasi grafik untuk memudahkan analisis.

- i. Sistem mampu menghasilkan laporan hasil clustering yang dapat dicetak sebagai dokumentasi analisis sistem.

3.2.2.Kebutuhan Non-Fungsional

Kebutuhan non-fungsional berkaitan dengan kualitas dan karakteristik sistem yang dibangun. Kebutuhan non-fungsional sistem meliputi:

- a. Kemudahan penggunaan: antarmuka berbasis web yang sederhana dan dapat dioperasikan oleh pengurus koperasi.
- b. Keamanan data: sistem membatasi akses dengan autentikasi username dan password.
- c. Kinerja sistem: proses clustering dapat diselesaikan dalam waktu kurang dari 5 detik untuk 8 mandor.
- d. Ketersediaan sistem: dapat diakses melalui browser pada jaringan lokal (localhost).
- e. Kemudahan pemeliharaan: sistem dibangun dengan PHP 8.1 native agar mudah dikembangkan.

3.3. Spesifikasi Perangkat Keras dan Perangkat Lunak Sistem

3.3.1. Spesifikasi Perangkat Keras

Perangkat keras yang digunakan dalam penelitian ini memiliki spesifikasi sebagai berikut:

- a. Processor minimal Intel Core i3 atau setara.
- b. RAM minimal 8 GB.
- c. Media penyimpanan minimal 256 GB SSD.
- d. Perangkat input berupa keyboard dan mouse.
- e. Perangkat output berupa monitor resolusi minimal 1366 x 768.

3.3.2. Spesifikasi Perangkat Lunak

Perangkat lunak yang digunakan dalam penelitian ini meliputi:

- a. Sistem operasi: Windows 10/11 atau setara.
- b. Bahasa pemrograman: PHP 8.1 (native, tanpa framework).
- c. Web server: Apache (melalui Laragon).
- d. Database: MySQL 8.0.
- e. Browser: Google Chrome atau Mozilla Firefox.
- f. Tools desain: StarUML/draw.io untuk perancangan diagram UML.

3.3.3. Keterkaitan Perangkat dengan Sistem

Perangkat keras berfungsi sebagai sarana utama untuk menjalankan server lokal Laragon, menyimpan basis data MySQL, dan melakukan proses perhitungan K-Means Clustering. Perangkat lunak PHP 8.1 menangani logika algoritma K-Means, normalisasi Min-Max, dan pengelolaan basis data. Kesesuaian antara perangkat keras dan perangkat lunak memungkinkan sistem berjalan secara optimal pada lingkungan pengembangan lokal.

3.4. Tahapan Perancangan Sistem

Perancangan sistem dilakukan untuk menggambarkan struktur dan alur kerja sistem secara teknis menggunakan Unified Modeling Language (UML), meliputi Usecase Diagram, Activity Diagram, dan Sequence Diagram. Pada tahap akhir perancangan sistem, hasil dari proses K-Means Clustering tidak hanya berupa penentuan keanggotaan cluster setiap mandor, tetapi juga mencakup perhitungan nilai Sum of Squared Error (SSE), penyimpanan nilai centroid pada setiap iterasi, serta pengolahan data agregasi berupa rata-rata nilai variabel pada masing-masing cluster. Seluruh output tersebut dirancang sebagai bagian dari

sistem untuk mendukung proses evaluasi dan interpretasi hasil pada tahap implementasi dan pengujian sistem.

4.3.1. Perancangan Tampilan

Pada tahap perancangan sistem, diperlukan gambaran antarmuka sebagai acuan dalam proses pengembangan sistem yang akan dibangun. Rancangan antarmuka ini disusun dalam bentuk wireframe yang bertujuan untuk memberikan ilustrasi awal mengenai struktur tampilan, tata letak komponen, serta alur interaksi pengguna dengan sistem.

Wireframe yang dirancang masih bersifat sederhana dan belum menampilkan detail visual seperti warna dan desain akhir, namun telah merepresentasikan fungsi utama dari setiap halaman dalam sistem.

a. Tampilan Halaman Login

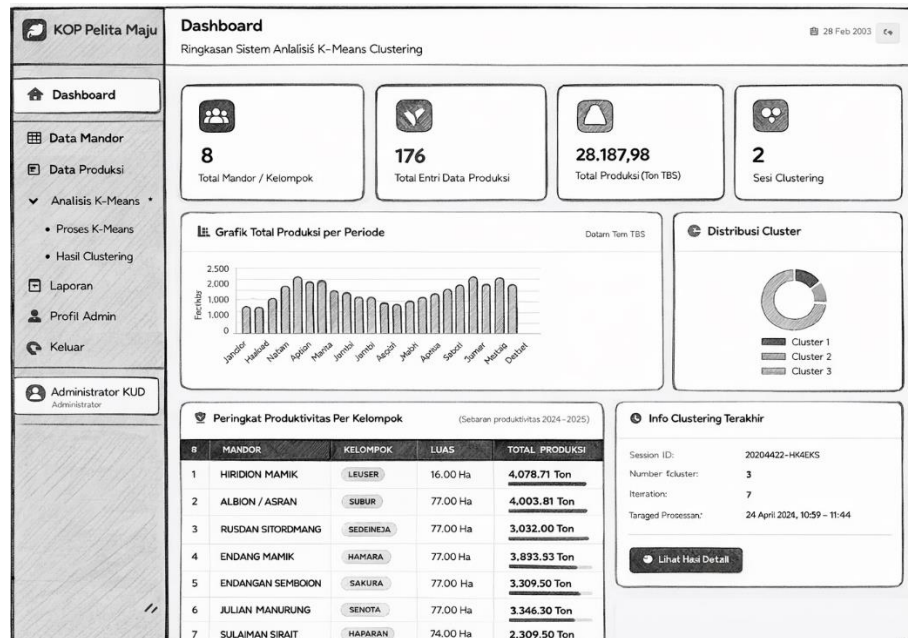
Di bawah ini merupakan rancangan tampilan halaman login dari sistem yang akan dibangun. Halaman ini digunakan untuk proses masuk ke dalam sistem dengan memasukkan username dan password.



Gambar 3.4.1 Rancangan Halaman Login Sistem

b. Tampilan Halaman Dashboard

Di bawah ini merupakan rancangan tampilan halaman dashboard dari sistem yang akan dibangun. Halaman ini berfungsi sebagai halaman utama yang menampilkan informasi umum sistem.



Gambar 3.4.2 Rancangan Halaman Dashboard

c. Tampilan Halaman Data Mandor

Di bawah ini merupakan rancangan tampilan halaman data mandor dari sistem yang akan dibangun. Halaman ini digunakan untuk mengelola data mandor.

KOP Oleita	Data Mandor / Kelompok 🔍 Dashboard > Data Mandor																																																
Dashboard	Data Mandor / Kelompok + Tambah Mandor																																																
Data Mandor	atamahama 8 mandor terdaftar																																																
Data Produksi	<table border="1"> <thead> <tr> <th>Kode</th> <th>Nama Mandor</th> <th>Nama Kelompok</th> <th>Luas Lahan</th> <th>Aksi</th> </tr> </thead> <tbody> <tr> <td>KDR-001</td> <td>HERIDON MANIK</td> <td>LESTARI</td> <td>76 Ha</td> <td> </td> </tr> <tr> <td>KDR-002</td> <td>RUBMAN SITUNG--</td> <td>SEDEPHKHA</td> <td>73 Ha</td> <td> </td> </tr> <tr> <td>KDR-003</td> <td>JULIAN MANURUNG</td> <td>BEZCT</td> <td>76 Ha</td> <td> </td> </tr> <tr> <td>KDR-004</td> <td>EDC SIRAIT</td> <td>SRNTOSA</td> <td>79 Ha</td> <td> </td> </tr> <tr> <td>KDR-005</td> <td>RUDANRYANIS...</td> <td>SADVR</td> <td>46 Ha</td> <td> </td> </tr> <tr> <td>KDR-006</td> <td>ENDINIG, MANIK</td> <td>MAKITVR</td> <td>77 Ha</td> <td> </td> </tr> <tr> <td>KDR-007</td> <td>SULARANG SRAIT</td> <td>HMBARAN</td> <td>42 Ha</td> <td> </td> </tr> <tr> <td>KDR-008</td> <td>ALBION / ASRAN</td> <td>SUBUR</td> <td>79 Ha</td> <td> </td> </tr> </tbody> </table>				Kode	Nama Mandor	Nama Kelompok	Luas Lahan	Aksi	KDR-001	HERIDON MANIK	LESTARI	76 Ha		KDR-002	RUBMAN SITUNG--	SEDEPHKHA	73 Ha		KDR-003	JULIAN MANURUNG	BEZCT	76 Ha		KDR-004	EDC SIRAIT	SRNTOSA	79 Ha		KDR-005	RUDANRYANIS...	SADVR	46 Ha		KDR-006	ENDINIG, MANIK	MAKITVR	77 Ha		KDR-007	SULARANG SRAIT	HMBARAN	42 Ha		KDR-008	ALBION / ASRAN	SUBUR	79 Ha	
Kode	Nama Mandor	Nama Kelompok	Luas Lahan	Aksi																																													
KDR-001	HERIDON MANIK	LESTARI	76 Ha																																														
KDR-002	RUBMAN SITUNG--	SEDEPHKHA	73 Ha																																														
KDR-003	JULIAN MANURUNG	BEZCT	76 Ha																																														
KDR-004	EDC SIRAIT	SRNTOSA	79 Ha																																														
KDR-005	RUDANRYANIS...	SADVR	46 Ha																																														
KDR-006	ENDINIG, MANIK	MAKITVR	77 Ha																																														
KDR-007	SULARANG SRAIT	HMBARAN	42 Ha																																														
KDR-008	ALBION / ASRAN	SUBUR	79 Ha																																														
Profil Admin	6 mandor Kedafter																																																
Administrator KUDINIZAHINDO																																																	

Gambar 3.4.3 Rancangan Halaman Data Mandor

d. Tampilan Halaman Data Produksi

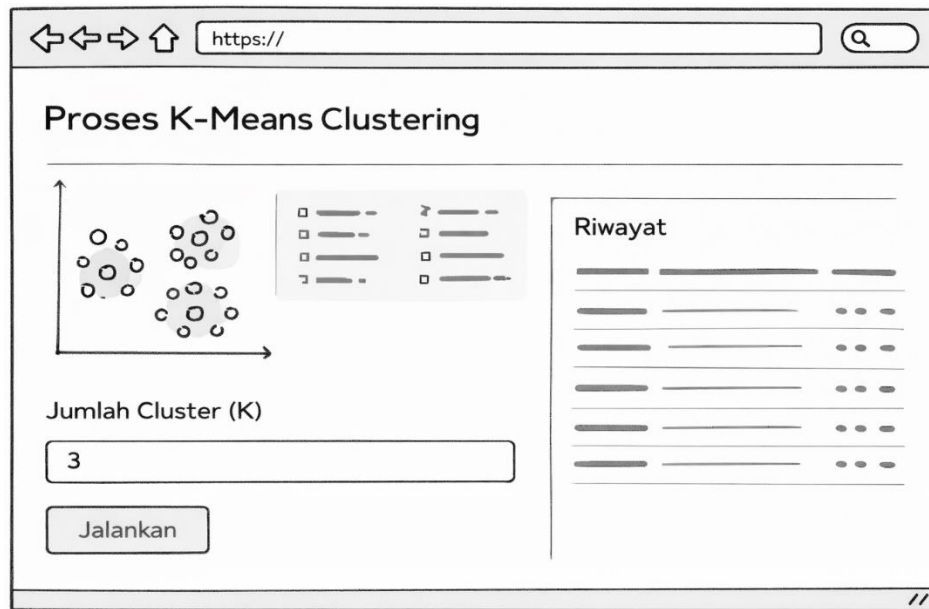
Di bawah ini merupakan rancangan tampilan halaman data produksi dari sistem yang akan dibangun. Halaman ini digunakan untuk mengelola data hasil produksi.

KOP Oleita	Data Produksi Panen + 📄 🗑️ Kelola data tonase, rata-rata, dan langsiran per periode																																														
Dashboard	Data Produksi Panen + Tambah Data Dashboard > Data Produksi																																														
Data Mandor	Filter Tahun: <input type="text" value="Semua"/> Filter Mandor: <input type="text" value="Semua"/> Filter Reset																																														
Data Produksi	Rekap Produksi Hasil Panen 176 entri Total: 28,187,976 ton <table border="1"> <thead> <tr> <th>#</th> <th>MANDOR</th> <th>KELOMPOK</th> <th>PERIODE</th> <th>LUAS (HA)</th> <th>TONASE (KG)</th> <th>RATA-RATA (KG/HA)</th> <th>AKSI</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>HERIDON MANIK</td> <td>LESTARI</td> <td>2025-12</td> <td>76 Ha</td> <td>212.365.000</td> <td>2.794.000</td> <td> </td> </tr> <tr> <td>2</td> <td>RUSNAN SITUM-- MOR-001</td> <td>SEDERHANA</td> <td>2025-12</td> <td>75 Ha</td> <td>214.186.000</td> <td>2.856.000</td> <td> </td> </tr> <tr> <td>3</td> <td>JULIAN MANURUNG MOR-002</td> <td>REJEKI</td> <td>2025-12</td> <td>74 Ha</td> <td>172.903.000</td> <td>2.337.000</td> <td> </td> </tr> <tr> <td>4</td> <td>JULIAN MANURUNG MOR-003</td> <td>REJEKI</td> <td>2025-12</td> <td>74 Ha</td> <td>172.903.000</td> <td>2.337.000</td> <td> </td> </tr> </tbody> </table>							#	MANDOR	KELOMPOK	PERIODE	LUAS (HA)	TONASE (KG)	RATA-RATA (KG/HA)	AKSI	1	HERIDON MANIK	LESTARI	2025-12	76 Ha	212.365.000	2.794.000		2	RUSNAN SITUM-- MOR-001	SEDERHANA	2025-12	75 Ha	214.186.000	2.856.000		3	JULIAN MANURUNG MOR-002	REJEKI	2025-12	74 Ha	172.903.000	2.337.000		4	JULIAN MANURUNG MOR-003	REJEKI	2025-12	74 Ha	172.903.000	2.337.000	
#	MANDOR	KELOMPOK	PERIODE	LUAS (HA)	TONASE (KG)	RATA-RATA (KG/HA)	AKSI																																								
1	HERIDON MANIK	LESTARI	2025-12	76 Ha	212.365.000	2.794.000																																									
2	RUSNAN SITUM-- MOR-001	SEDERHANA	2025-12	75 Ha	214.186.000	2.856.000																																									
3	JULIAN MANURUNG MOR-002	REJEKI	2025-12	74 Ha	172.903.000	2.337.000																																									
4	JULIAN MANURUNG MOR-003	REJEKI	2025-12	74 Ha	172.903.000	2.337.000																																									
Profil Admin	176 entri Total: 28,187,976 ton 176 entri Total: 28,187,976 ton																																														
Administrator ACIZIONATOR KUD																																															

Gambar 3.4.4 Rancangan Halaman Data Produksi

e. Tampilan Halaman Proses Clustering

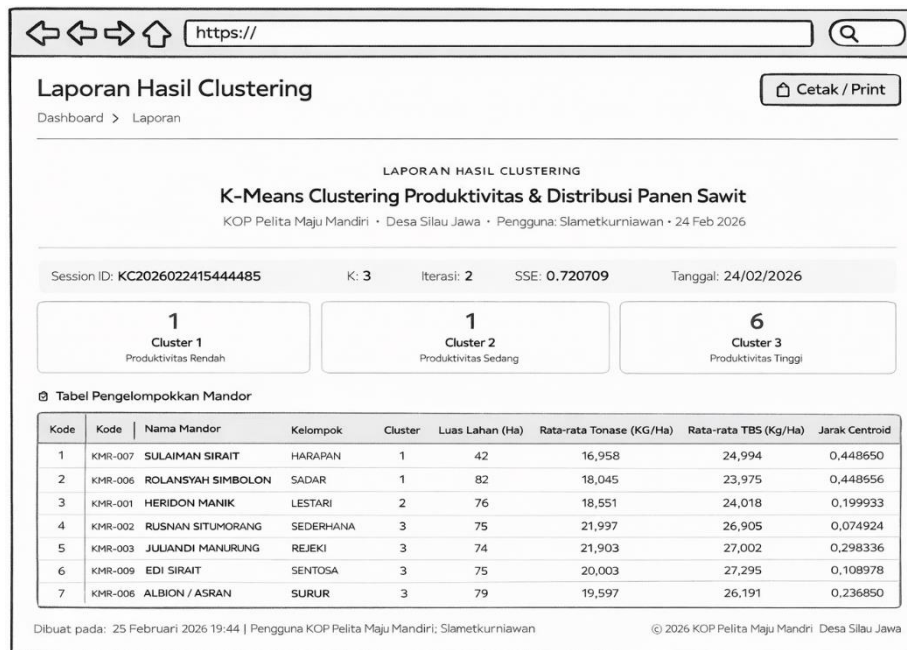
Di bawah ini merupakan rancangan tampilan halaman proses clustering dari sistem yang akan dibangun. Halaman ini digunakan untuk melakukan proses pengelompokan data.



Gambar 3.4.5 Rancangan Halaman Proses Clustering

f. Tampilan Halaman Laporan

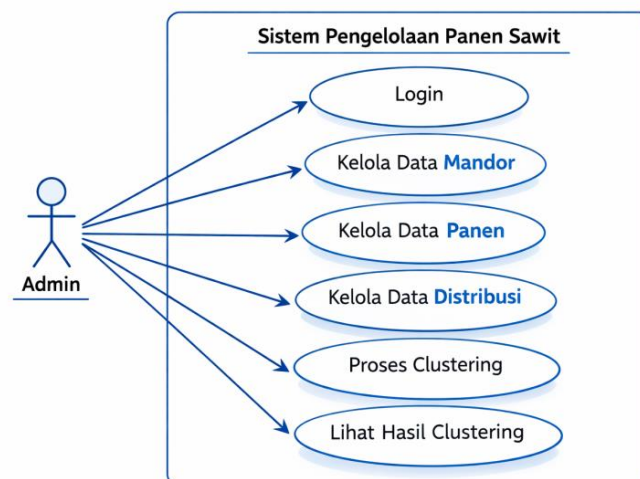
Di bawah ini merupakan rancangan tampilan halaman laporan dari sistem yang akan dibangun. Halaman ini digunakan untuk menampilkan hasil pengolahan data.



Gambar 3.4.6 Rancangan Halaman Laporan

3.4.2. Usecase Diagram

Usecase Diagram menggambarkan interaksi antara aktor dengan sistem. Aktor utama dalam sistem ini adalah Admin (Pengurus KUD). Admin memiliki hak akses penuh untuk mengelola data dan menjalankan proses analisis.

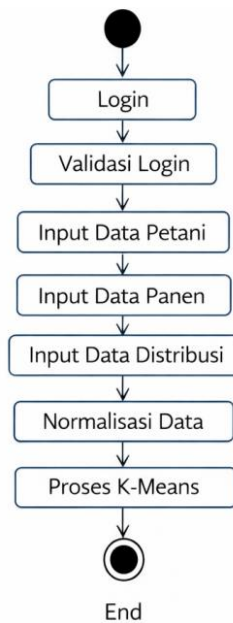


**Gambar 3.4.7 Use Case Diagram Sistem Informasi
 KUD Pelita Maju Mandiri**

Berdasarkan Gambar 3.4.1, Admin dapat melakukan login ke sistem, mengelola data mandor/kelompok, mengelola data produksi panen per periode, menjalankan proses K-Means Clustering, melihat hasil pengelompokan beserta visualisasi, serta mencetak laporan hasil analisis.

3.4.3. Activity Diagram

Activity Diagram digunakan untuk menggambarkan alur aktivitas sistem secara keseluruhan dari awal hingga akhir proses.

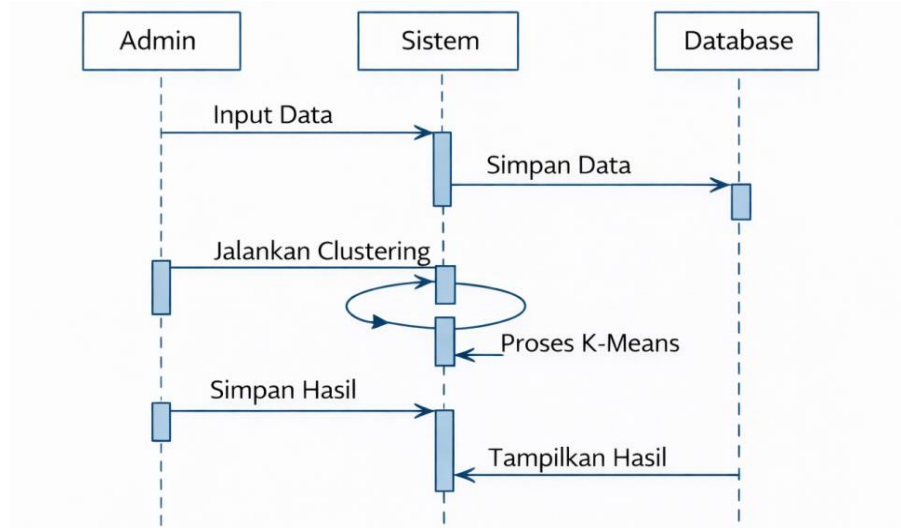


Gambar 3.4.8 Activity Diagram Sistem

Gambar 3.4.2 menunjukkan bahwa proses sistem dimulai dari Admin melakukan login. Setelah berhasil login, Admin melakukan input dan validasi data mandor serta data produksi panen bulanan. Selanjutnya sistem melakukan normalisasi data Min-Max dan menjalankan proses K-Means Clustering dengan inisialisasi K-Means++. Hasil pengelompokan kemudian ditampilkan beserta nilai centroid akhir, distribusi cluster, dan grafik visualisasi, kemudian dapat dicetak sebagai laporan.

3.4.4. Sequence Diagram

Sequence Diagram menggambarkan urutan interaksi antara aktor dan komponen sistem secara detail berdasarkan urutan waktu.



Gambar 3.4.9 Sequence Diagram Proses Clustering

Berdasarkan Gambar 3.4.3, Admin mengirimkan perintah input data ke sistem, sistem menyimpan data ke dalam basis data MySQL. Selanjutnya Admin menjalankan proses clustering dengan memilih nilai K. Sistem memanggil kelas KMeans, melakukan normalisasi data, inialisasi centroid K-Means++, menghitung jarak Euclidean, memperbarui centroid, dan melakukan iterasi hingga konvergen. Hasil clustering disimpan ke tabel hasil_clustering dan detail_clustering, kemudian ditampilkan kepada Admin.

3.5. Perancangan Basis Data

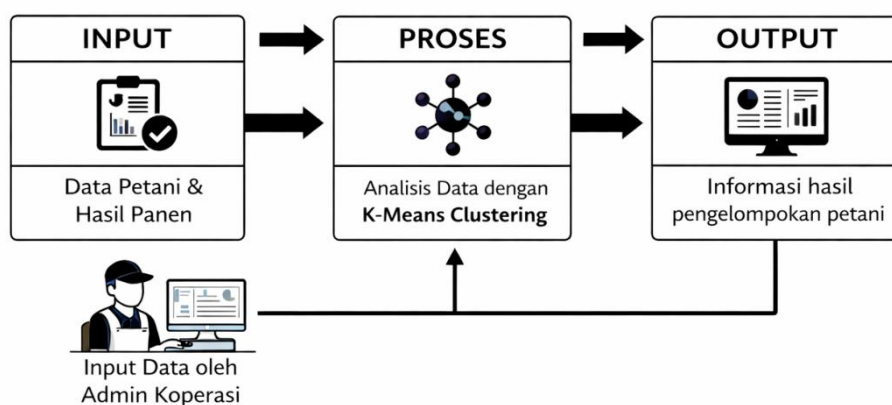
Perancangan basis data dilakukan untuk mendukung penyimpanan dan pengolahan data secara terstruktur. Basis data yang digunakan adalah MySQL dengan nama database db_kmeans_sawit.

3.5.1. Alur Sistem

Alur sistem pada penelitian ini menggunakan konsep Input–Process–Output (IPO). Pada tahap input, data yang digunakan berupa rekap produksi hasil panen dari KUD "Pelita Maju Mandiri" Desa Silau Jawa yang mencakup data 8 mandor/kelompok selama periode Januari 2024 hingga Desember 2025. Data meliputi nama mandor, nama kelompok, luas lahan, tonase TBS, rata-rata produksi per hektar, dan langsriran per bulan.

Pada tahap proses, data dirata-ratakan per mandor dari seluruh periode, kemudian dinormalisasi menggunakan metode Min-Max ke rentang [0,1]. Selanjutnya sistem menjalankan algoritma K-Means++ untuk inialisasi centroid, diikuti iterasi penugasan cluster berdasarkan jarak Euclidean, pembaruan centroid, dan pengujian konvergensi.

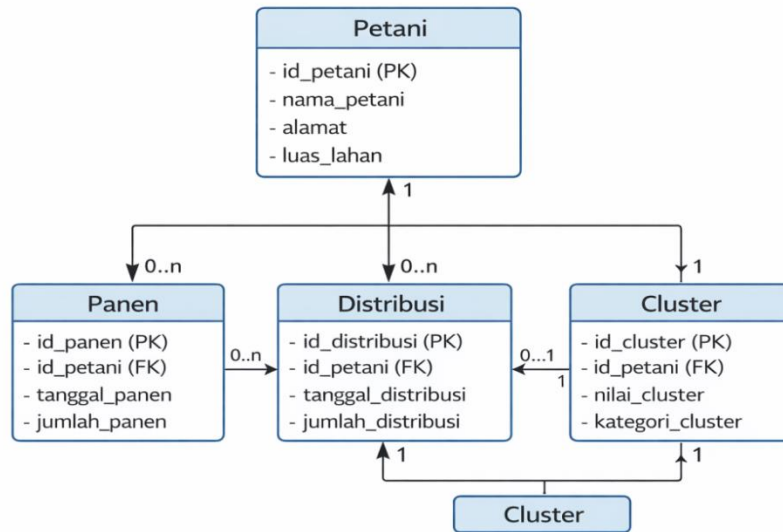
Tahap output menghasilkan pengelompokan 8 mandor ke dalam K cluster beserta informasi centroid akhir, nilai SSE, distribusi anggota per cluster, dan laporan analisis yang dapat dicetak.



Gambar 3.5.1 Alur Sistem Pengelompokan Mandor Menggunakan K-Means

3.5.2. Entity Relationship Diagram (ERD)

ERD digunakan untuk menggambarkan hubungan antar entitas yang terlibat dalam sistem secara konseptual.



Gambar 3.5.2. Entity Relationship Diagram (ERD) Sistem Pengelolaan Panen Sawit

Berdasarkan Gambar 3.5.2, sistem memiliki lima entitas utama: mandor, data_panen, hasil_clustering, detail_clustering, dan centroid_history. Entitas mandor menyimpan data identitas mandor/kelompok. Entitas data_panen berelasi one-to-many dengan mandor, karena satu mandor dapat memiliki banyak data produksi bulanan. Entitas hasil_clustering menyimpan metadata sesi clustering. Entitas detail_clustering menyimpan hasil pengelompokan per mandor pada setiap sesi, dan centroid_history menyimpan jejak pergerakan centroid per iterasi.

3.5.3. Struktur Tabel Basis Data

Berikut adalah struktur tabel-tabel yang digunakan dalam sistem:

a. Tabel Mandor

Field	Tipe Data	Keterangan
id	INT (PK, Auto Increment)	Primary key, identitas unik mandor
kode_mandor	VARCHAR(20)	Kode unik mandor (MDR-001 s/d MDR-008)
nama_mandor	VARCHAR(100)	Nama lengkap mandor
nama_kelompok	VARCHAR(100)	Nama kelompok tani
luas_lahan	DECIMAL(10,2)	Luas lahan dalam Hektar (Ha)
keterangan	TEXT	Keterangan tambahan (opsional)
created_at	TIMESTAMP	Waktu data dibuat
updated_at	TIMESTAMP	Waktu data terakhir diperbarui

b. Tabel data_panen

Field	Tipe Data	Keterangan
id	INT (PK, Auto Increment)	Primary key
id_mandor	INT (FK)	Foreign key ke tabel mandor
periode	VARCHAR(7)	Periode panen format YYYY-MM (misal 2024-01)
tonase_kg	DECIMAL(12,3)	Total tonase TBS dalam Kilogram
rata_rata_kg_ha	DECIMAL(10,3)	Rata-rata produksi per hektar (KG/Ha)
langsiran	INT	Frekuensi pengiriman ke pabrik (kali/bulan)
keterangan	TEXT	Keterangan tambahan
created_at	TIMESTAMP	Waktu data dibuat

c. Tabel hasil_clustering

Field	Tipe Data	Keterangan
id	INT (PK, Auto Increment)	Primary key
session_id	VARCHAR(50)	ID unik sesi clustering (format: KC+timestamp)
jumlah_cluster	INT	Nilai K yang digunakan
iterasi	INT	Jumlah iterasi hingga konvergen
nilai_sse	DECIMAL(20,6)	Nilai Sum of Squared Errors (SSE)
tanggal_proses	DATETIME	Waktu proses clustering dijalankan
keterangan	TEXT	Parameter clustering yang digunakan

d. Tabel detail_clustering

Field	Tipe Data	Keterangan
id	INT (PK, Auto Increment)	Primary key
session_id	VARCHAR(50)	Referensi ke sesi clustering
id_mandor	INT (FK)	Foreign key ke tabel mandor
cluster	INT	Nomor cluster yang ditetapkan (1 s/d K)
norm_luas_lahan	DECIMAL(10,6)	Nilai normalisasi luas lahan
norm_tonase	DECIMAL(10,6)	Nilai normalisasi tonase
norm_rata_rata	DECIMAL(10,6)	Nilai normalisasi rata-rata KG/Ha
norm_langsiran	DECIMAL(10,6)	Nilai normalisasi langsiran
jarak_ke_centroid	DECIMAL(15,6)	Jarak Euclidean ke centroid cluster

3.5.4. Relasi Antar Tabel

Relasi antar tabel dalam basis data dirancang sebagai berikut:

- a. Tabel mandor memiliki relasi one-to-many dengan tabel data_panen (satu mandor memiliki banyak data produksi bulanan).
- b. Tabel hasil_clustering memiliki relasi one-to-many dengan tabel detail_clustering (satu sesi clustering memiliki banyak detail per mandor).
- c. Tabel mandor memiliki relasi one-to-many dengan tabel detail_clustering melalui foreign key id_mandor.
- d. Tabel hasil_clustering memiliki relasi one-to-many dengan tabel centroid_history (satu sesi memiliki banyak rekam jejak centroid).

3.6. Keterkaitan Sistem dengan Studi Kasus

Sistem informasi yang dirancang dalam penelitian ini disesuaikan secara langsung dengan kondisi operasional KUD "Pelita Maju Mandiri" Desa Silau Jawa. Sistem menggunakan data rekap produksi hasil panen yang dimiliki koperasi sebagai input utama, yang mencakup 8 mandor/kelompok dengan total 22 periode data (Januari 2024 – Desember 2025, dengan dua bulan yang tidak tersedia di tahun 2024 yaitu Maret dan September). Data ini diinput ke dalam sistem untuk kemudian dianalisis menggunakan metode K-Means Clustering.

Variabel yang digunakan dalam proses clustering dipilih berdasarkan kolom-kolom yang tersedia dalam dokumen rekap koperasi, yaitu luas lahan (Ha), tonase TBS (KG), rata-rata produksi per hektar (KG/Ha), dan langsiaran (kali/bulan). Dengan demikian, sistem tidak hanya bersifat teoritis, tetapi mampu mengolah dan menganalisis data nyata koperasi untuk menghasilkan

pengelompokan mandor yang objektif dan dapat dijadikan dasar pengambilan keputusan.

3.7. Rancangan Output Sistem

Output sistem dirancang untuk menyajikan informasi hasil analisis K-Means Clustering secara komprehensif, meliputi:

- a. Dashboard ringkasan: total mandor, total entri data produksi, total produksi (ton TBS), jumlah sesi clustering, grafik produksi per periode, dan distribusi cluster terakhir.
- b. Hasil clustering: distribusi anggota per cluster, nilai centroid akhir (ternormalisasi), nilai SSE, jumlah iterasi, tabel detail pengelompokan mandor beserta jarak ke centroid.
- c. Visualisasi grafis: pie chart distribusi cluster dan bar chart perbandingan rata-rata nilai variabel antar cluster menggunakan Chart.js.
- d. Laporan cetak: dokumen laporan formal berisi ringkasan sesi clustering, tabel pengelompokan mandor, dan area tanda tangan pengurus koperasi.

3.8. Lokasi dan Waktu Penelitian

Penelitian ini dilaksanakan di KUD "Pelita Maju Mandiri" Desa Silau Jawa, Kabupaten Asahan, Provinsi Sumatera Utara. Pemilihan lokasi penelitian didasarkan pada ketersediaan data rekap produksi hasil panen yang lengkap serta peran strategis koperasi sebagai lembaga yang mengelola dan mendokumentasikan data produktivitas dan distribusi TBS dari 8 mandor/kelompok secara berkala.

Waktu penelitian dilaksanakan selama kurang lebih delapan bulan, mencakup tahap studi literatur, observasi lapangan, pengumpulan data, analisis dan perancangan sistem, pengembangan sistem, pengujian, serta penyusunan laporan skripsi.

3.9. Objek Penelitian

Objek dalam penelitian ini adalah data rekap produksi hasil panen kelapa sawit yang dikelola oleh KUD "Pelita Maju Mandiri" Desa Silau Jawa. Data tersebut mencakup informasi 8 mandor/kelompok tani dengan total luas lahan 580 Ha, yang mencatat data produksi bulanan selama periode Januari 2024 hingga Desember 2025.

Subjek penelitian ini adalah mandor/kelompok tani yang tergabung sebagai anggota KUD Pelita Maju Mandiri Desa Silau Jawa. Tabel 3.9 berikut menyajikan daftar mandor/kelompok yang menjadi objek analisis dalam penelitian ini:

Tabel 3.9 Daftar Mandor dan Kelompok Objek Penelitian

No	Nama Mandor	Nama Kelompok	Luas Lahan (Ha)
1	HERDION MANIK	LESTARI	76
2	RUSNAN SITUMORANG	SEDERHANA	75
3	JULIANDI MANURUNG	REJEKI	74
4	EDI SIRAIT	SENTOSA	75
5	ROLANSYAH SIMBOLON / DENI SITORUS	SADAR	82
6	ENDANG MANIK	MAKMUR	77
7	SULAIMAN SIRAIT	HARAPAN	42
8	ALBOIN / ASRAN	SUBUR	79

Dari Tabel 3.9, terlihat bahwa luas lahan kelompok bervariasi antara 42 Ha (HARAPAN) hingga 82 Ha (SADAR), dengan total keseluruhan 580 Ha. Variasi luas lahan ini menjadi salah satu variabel penting dalam analisis clustering. Perlu dicatat bahwa beberapa posisi mandor mengalami pergantian pada periode tertentu (misalnya kelompok SADAR dan SUBUR), namun luas lahan dan nama kelompok tetap konsisten sepanjang periode penelitian.

3.10. Sumber dan Jenis Data

3.10.1. Sumber Data

Penelitian ini menggunakan dua jenis sumber data yang saling melengkapi:

Data primer diperoleh melalui wawancara dan observasi langsung dengan pengurus KUD Pelita Maju Mandiri untuk memahami mekanisme pencatatan data, sistem distribusi TBS, serta aspek operasional lainnya yang berkaitan dengan pengelolaan data panen.

Data sekunder merupakan data utama penelitian yang bersumber dari dokumen resmi koperasi berupa Rekap Produksi Hasil Panen KUD "Pelita Maju Mandiri" Desa Silau Jawa periode Januari 2024 – Desember 2025. Data ini mencakup 22 tabel rekap bulanan dengan total 176 entri data (8 mandor \times 22 bulan, dikurangi 2 bulan yang tidak tersedia). Setiap tabel memuat kolom: Nama Mandor, Nama Kelompok, Luas Lahan (Ha), Tonase (KG), Rata-rata (KG/Ha), dan Langsiran.

3.10.2. Jenis Data

Jenis data yang digunakan dalam penelitian ini adalah data kuantitatif, yaitu data numerik yang dapat diolah secara matematis. Seluruh variabel yang

digunakan dalam proses K-Means Clustering bersifat numerik sehingga memungkinkan perhitungan jarak Euclidean dan normalisasi Min-Max dilakukan secara akurat dan objektif.

3.11. Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan dalam penelitian ini adalah sebagai berikut:

- a. Observasi: pengamatan langsung terhadap proses pencatatan dan pengelolaan data panen di KUD Pelita Maju Mandiri Desa Silau Jawa.
- b. Wawancara: diskusi dengan pengurus koperasi untuk memahami struktur data, sistem pelaporan, dan kebutuhan analisis.
- c. Studi dokumentasi: pengumpulan dan telaah dokumen rekap produksi hasil panen bulanan periode 2024–2025 yang merupakan data primer penelitian.

Tabel 3.11 Sampel Data Produksi Hasil Panen (KUD Pelita Maju Mandiri, 2024–2025)

No	Nama Mandor	Kelompok	Periode	Tonase (KG)
1	HERDION MANIK	LESTARI	2024-01	148,125
2	RUSNAN SITUMORANG	SEDERHANA	2024-01	130,093
3	JULIANDI MANURUNG	REJEKI	2025-09	186,583
4	EDI SIRAIT	SENTOSA	2025-09	192,074
5	ROLANSYAH SIMBOLON	SADAR	2025-08	158,394
6	SULAIMAN SIRAIT	HARAPAN	2025-08	129,520

Tabel 3.11 menampilkan sebagian sampel data produksi yang digunakan dalam penelitian. Data lengkap mencakup 22 periode bulanan untuk 8 mandor/kelompok dengan total 176 entri data yang telah diinput ke dalam sistem.

3.12. Variabel Penelitian

Berdasarkan dokumen rekap produksi yang dimiliki KUD Pelita Maju Mandiri, variabel yang digunakan dalam proses K-Means Clustering adalah sebagai berikut:

Tabel 3.12 Variabel Penelitian dalam Proses K-Means Clustering

No	Variabel	Deskripsi	Satuan/Keterangan
1	Luas Lahan (X1)	Luas area perkebunan yang dikelola setiap mandor	Hektar (Ha). Berkisar 42–82 Ha
2	Tonase TBS (X2)	Total hasil panen TBS per periode panen	Kilogram (KG). Berkisar 59.400–248.349 KG/bulan
3	Rata-rata Produksi (X3)	Rata-rata produksi per hektar per bulan	KG/Ha. Berkisar 1.147–3.272 KG/Ha
4	Langsiran (X4)	Frekuensi pengiriman TBS ke pabrik per bulan	Kali/bulan. Berkisar 0–92 kali

Keempat variabel tersebut dipilih karena secara langsung tersedia dalam dokumen rekap koperasi dan mampu merepresentasikan kondisi produktivitas serta distribusi hasil panen secara kuantitatif. Sebelum digunakan dalam proses clustering, setiap variabel dinormalisasi menggunakan metode Min-Max agar seluruh variabel memiliki skala yang setara.

3.13. Analisis Data

Analisis data dalam penelitian ini dilakukan melalui beberapa tahapan sistematis menggunakan metode K-Means Clustering dengan inisialisasi K-Means++. Seluruh tahapan diimplementasikan dalam kelas PHP KMeans yang

menangani normalisasi, inialisasi centroid, iterasi clustering, dan penyimpanan hasil.

3.13.1. Seleksi Data

Dari dokumen rekap produksi, dipilih 4 kolom numerik sebagai variabel analisis: luas lahan (Ha), tonase KG, rata-rata KG/Ha, dan langsiran (kali/bulan). Sebelum diproses, data dirata-ratakan per mandor dari seluruh periode yang tersedia agar setiap mandor direpresentasikan oleh satu vektor fitur.

3.13.2. Pembersihan Data (Data Cleaning)

Data rekap yang diperoleh dari koperasi telah dalam kondisi terstruktur. Proses pembersihan meliputi penanganan nilai langsiran kosong yang diisi dengan nilai 0, serta verifikasi konsistensi luas lahan setiap kelompok. Tidak ditemukan data duplikat karena setiap entri merujuk pada kombinasi unik mandor dan periode.

3.13.3. Normalisasi Data

Normalisasi dilakukan menggunakan metode Min-Max Normalization untuk mentransformasikan seluruh variabel ke rentang [0,1] sebelum perhitungan jarak Euclidean. Rumus normalisasi yang digunakan adalah:

$$X' = (X - X_{min}) / (X_{max} - X_{min})$$

Tabel 3.13.3 berikut menunjukkan contoh hasil normalisasi untuk data rata-rata seluruh periode:

Tabel 3.13.3 Hasil Normalisasi Data (Rata-rata Seluruh Periode 2024–2025)

No	Mandor	Luas (norm)	Tonase (norm)	Rata-rata (norm)	Langsiran (norm)
1	HERDION MANIK	0.850	0.721	0.715	0.182
2	RUSNAN SITUMORANG	0.825	0.698	0.693	0.000

No	Mandor	Luas (norm)	Tonase (norm)	Rata-rata (norm)	Langsiran (norm)
3	JULIANDI MANURUNG	0.800	0.623	0.615	0.000
4	EDI SIRAIT	0.825	0.651	0.645	0.141
5	ROLANSYAH SIMBOLON	1.000	0.608	0.457	0.849
6	ENDANG MANIK	0.875	0.701	0.695	0.000
7	SULAIMAN SIRAIT	0.000	0.309	0.532	0.803
8	ALBOIN / ASRAN	0.925	0.712	0.706	0.141

3.13.4. Penentuan Jumlah Cluster (K)

Dalam penelitian ini, jumlah cluster yang digunakan dapat dipilih antara K=2 hingga K=5 melalui antarmuka sistem. Untuk analisis utama, digunakan K=3 yang menghasilkan pengelompokan menjadi tiga kategori: Produktivitas Rendah, Produktivitas Sedang, dan Produktivitas Tinggi. Pemilihan K=3 didasarkan pada referensi penelitian terdahulu yang menganalisis produktivitas kelapa sawit menggunakan tiga kategori tingkatan.

3.13.5. Proses Clustering dengan K-Means

Proses K-Means dilaksanakan dengan langkah-langkah sebagai berikut:

- a. Tahap 1 – Inisialisasi K-Means++: pemilihan centroid awal secara probabilistik untuk mengurangi risiko konvergensi ke solusi lokal yang tidak optimal.
- b. Tahap 2 – Penugasan cluster: setiap mandor ditugaskan ke cluster dengan centroid terdekat berdasarkan jarak Euclidean pada ruang fitur 4 dimensi ternormalisasi.

- c. Tahap 3 – Pembaruan centroid: centroid setiap cluster diperbarui sebagai rata-rata vektor fitur seluruh anggota cluster.
- d. Tahap 4 – Pengujian konvergensi: iterasi dihentikan apabila perubahan posisi centroid lebih kecil dari threshold $1e-8$, atau telah mencapai batas maksimum 100 iterasi.
- e. Tahap 5 – Perhitungan SSE: nilai Sum of Squared Errors dihitung sebagai ukuran kualitas pengelompokan.

3.13.6. Analisis Hasil Clustering Mandor

Hasil clustering memetakan 8 mandor/kelompok ke dalam K cluster berdasarkan rata-rata nilai variabel dari seluruh periode 2024–2025. Interpretasi setiap cluster dilakukan berdasarkan nilai centroid akhir pada variabel asli (sebelum normalisasi). Berikut adalah daftar mandor/kelompok yang menjadi subjek analisis:

Tabel 3.13.6 Daftar Mandor/Kelompok sebagai Subjek Analisis Clustering

Kode	Mandor	Kelompok	Luas (Ha)	Avg Tonase (KG)	Avg Rata-rata	Avg Langsiran
MDR-001	HERDION MANIK	LESTARI	76	188.595	2.481	7,5
MDR-002	RUSNAN SITUMORANG	SEDERHANA	75	183.305	2.435	1,1
MDR-003	JULIANDI MANURUNG	REJEKI	74	158.327	2.127	0,3
MDR-004	EDI SIRAIT	SENTOSA	75	155.923	2.080	6,2
MDR-005	ROL./DENI (SADAR)	SADAR	82	162.571	1.934	28,5

Kode	Mandor	Kelompok	Luas (Ha)	Avg Tonase (KG)	Avg Rata-rata	Avg Langsiran
MDR-006	ENDANG MANIK	MAKMUR	77	181.321	2.354	0,1
MDR-007	SULAIMAN SIRAIT	HARAPAN	42	96.473	2.308	22,7
MDR-008	ALBOIN/ASRAN	SUBUR	79	183.419	2.317	5,7

Nilai rata-rata pada Tabel 3.13.6 dihitung dari seluruh periode data yang tersedia. Hasil analisis clustering akan mengidentifikasi pola kemiripan antar mandor berdasarkan kombinasi keempat variabel tersebut, sehingga pengurus koperasi dapat memahami karakteristik produktivitas dan distribusi setiap kelompok secara lebih objektif.

3.14. Alur Penelitian

Alur penelitian ini menggambarkan tahapan yang dilalui dari awal hingga diperolehnya hasil analisis clustering. Secara garis besar, alur penelitian terdiri dari:

- a. Identifikasi masalah: observasi kondisi pengelolaan data di KUD Pelita Maju Mandiri dan perumusan permasalahan.
- b. Studi literatur: kajian teori K-Means Clustering, normalisasi Min-Max, dan penelitian terdahulu.
- c. Pengumpulan data: pengambilan dokumen rekap produksi hasil panen periode 2024–2025.
- d. Analisis dan perancangan sistem: perancangan basis data, diagram UML, dan antarmuka sistem.

- e. Pengembangan sistem: implementasi sistem berbasis web PHP 8.1 dengan basis data MySQL.
- f. Pengujian dan validasi: pengujian fungsionalitas sistem dan validasi hasil clustering.
- g. Penyusunan laporan: penulisan dan penyempurnaan laporan skripsi.

3.15. Validasi Hasil Penelitian

Validasi hasil penelitian dilakukan melalui dua pendekatan. Pertama, validasi internal menggunakan nilai SSE (Sum of Squared Errors) sebagai ukuran kekompakan cluster, di mana nilai SSE yang lebih kecil mengindikasikan cluster yang lebih padat dan homogen. Kedua, validasi eksternal dilakukan dengan mengevaluasi kesesuaian antara hasil pengelompokan dengan kondisi data aktual di lapangan melalui diskusi dengan pengurus koperasi.

Sistem juga menyediakan fitur riwayat clustering sehingga pengurus koperasi dapat membandingkan hasil dari berbagai nilai K dan memilih pengelompokan yang paling relevan untuk keperluan pengambilan keputusan.

3.16. Waktu Penelitian

Penelitian ini dilaksanakan dalam rentang waktu kurang lebih tiga bulan. Waktu penelitian tersebut mencakup seluruh tahapan penelitian yang dimulai dari tahap analisis permasalahan hingga penyusunan laporan skripsi. Pembagian waktu penelitian dilakukan secara sistematis agar setiap tahapan dapat berjalan dengan optimal.

Tahap awal penelitian dimulai dengan analisis kebutuhan sistem dan pengumpulan data produktivitas serta distribusi hasil panen kelapa sawit. Selanjutnya dilakukan tahap perancangan sistem yang meliputi perancangan

proses data mining, perancangan UML, dan perancangan basis data. Tahap berikutnya adalah implementasi sistem dan pengujian proses clustering menggunakan algoritma K-Means. Tahap akhir penelitian difokuskan pada analisis hasil, evaluasi sistem, dan penyusunan laporan skripsi.

Rentang waktu penelitian tersebut dinilai cukup untuk memastikan bahwa sistem yang dirancang dapat dikembangkan dan diuji secara menyeluruh sehingga menghasilkan sistem yang sesuai dengan tujuan penelitian.

Tabel 3.16 Waktu Penelitian

No.	Kegiatan Penelitian	Tahun 2025-2026					
		Desember	Januari	Februari	Maret	April	Mei
1.	Penyusunan Proposal						
2.	Seminar Proposal						
3.	Preprocessing Data						
4.	Evaluasi						
5.	Sidang						

BAB IV

IMPLEMENTASI DAN PENGUJIAN SISTEM

4.1. Seleksi dan Persiapan Data

Tahap pertama dalam analisis mengacu pada alur penelitian yang telah dirancang pada Bab III, yaitu seleksi data dari dokumen Rekap Produksi Hasil Panen KUD "Pelita Maju Mandiri" Desa Silau Jawa. Data yang digunakan mencakup periode Januari 2024 hingga Desember 2025, dengan total 24 bulan data.

4.1.1 Seleksi Data

Dari dokumen rekap yang tersedia, dipilih empat variabel numerik yang relevan sesuai dengan variabel penelitian yang telah ditetapkan pada Subbab 3.12, yaitu luas lahan (Ha), tonase TBS (KG), rata-rata produksi per hektar (KG/Ha), dan langsiran (kali/bulan). Seleksi variabel ini dilakukan karena keempat kolom tersebut secara langsung tersedia dalam dokumen rekap koperasi dan mampu merepresentasikan kondisi produktivitas serta distribusi hasil panen secara kuantitatif.

Karena data bersifat time-series bulanan (24 periode untuk 8 mandor), maka untuk keperluan proses K-Means Clustering setiap mandor direpresentasikan oleh satu vektor fitur berupa nilai rata-rata dari seluruh periode yang tersedia. Pendekatan agregasi ini memastikan bahwa pengelompokan mencerminkan kondisi produktivitas dan distribusi mandor secara keseluruhan, bukan hanya pada satu periode tertentu.

4.1.2. Pembersihan Data

Proses pembersihan data (data cleaning) dilakukan terhadap seluruh entri rekap produksi. Hasil pemeriksaan menunjukkan bahwa:

- a. Tidak ditemukan duplikasi data — setiap entri merujuk pada kombinasi unik mandor dan periode.
 - b. Nilai langirsan yang tercatat sebagai kosong atau tidak tersedia diperlakukan sebagai nilai 0, karena tidak ada pengiriman TBS pada periode tersebut.
 - c. Luas lahan setiap kelompok bersifat konstan sepanjang periode sehingga tidak memerlukan interpolasi.
 - d. Terdapat pergantian nama mandor pada beberapa kelompok (SADAR: Rolansyah Simbolon → Deni Sitorus; SUBUR: Alboin → Asran), namun data produksi tetap dilanjutkan pada kelompok yang sama.
- Setelah pembersihan, total data yang valid dan siap diproses adalah 176 entri (8 mandor × 22 bulan).

4.2. Data Awal dan Penentuan Nilai Min-Max

4.2.1. Data Rata-Rata Mandor Periode 2024–2025

Tabel 4.2.1 menyajikan data rata-rata nilai variabel dari seluruh 24 periode bulanan per mandor. Data ini menjadi input utama dalam proses K-Means Clustering.

Tabel 4.2.1 Data Rata-Rata Produksi Mandor Periode 2024–2025

No	Nama Mandor	Kelompok	Luas (Ha)	Tonase (KG)	RR (KG/Ha)	Langsiran
1	HERDION MANIK	LESTARI	76	188.595	2.481	7,5
2	RUSNAN SITUMORANG	SEDERHANA	75	183.305	2.435	1,1
3	JULIANDI MANURUNG	REJEKI	74	158.327	2.127	0,3
4	EDI SIRAIT	SENTOSA	75	155.923	2.080	6,2
5	ROL./DENI SITORUS	SADAR	82	162.571	1.934	28,5
6	ENDANG MANIK	MAKMUR	77	181.321	2.354	0,1
7	SULAIMAN SIRAIT	HARAPAN	42	96.473	2.308	22,7
8	ALBOIN/ASRAN	SUBUR	79	183.419	2.317	5,7

Berdasarkan Tabel 4.2.1, terlihat variasi yang cukup signifikan antar mandor. Kelompok LESTARI mencatatkan rata-rata tonase tertinggi sebesar 188.595 KG/bulan, sementara kelompok HARAPAN memiliki rata-rata terendah sebesar 96.473 KG/bulan. Variasi luas lahan berkisar antara 42 Ha (HARAPAN) hingga 82 Ha (SADAR). Nilai langsiran juga bervariasi dari 0,1 kali/bulan (MAKMUR) hingga 28,5 kali/bulan (SADAR).

4.2.2. Penentuan Nilai Minimum dan Maksimum

Sebelum dilakukan normalisasi, ditentukan terlebih dahulu nilai minimum dan maksimum dari setiap variabel. Nilai ini digunakan sebagai batas dalam rumus Min-Max Normalization.

Tabel 4.2.2 Nilai Minimum dan Maksimum Setiap Variabel

Variabel	Nilai Min	Nilai Max	Range	Satuan	Keterangan
Luas Lahan (X1)	42	82	40	Ha	Min: HARAPAN, Max: SADAR
Tonase TBS (X2)	96.473	188.595	92.122	KG	Min: HARAPAN, Max: LESTARI
Rata-rata KG/Ha (X3)	1.934	2.481	547	KG/Ha	Min: SADAR, Max: LESTARI
Langsiran (X4)	0,1	28,5	28,4	Kali/bulan	Min: MAKMUR, Max: SADAR

4.3. Normalisasi Data

Sesuai dengan rancangan analisis pada Subbab 3.13.3, normalisasi data dilakukan menggunakan metode Min-Max Normalization dengan rumus:

$$X' = (X - X_{min}) / (X_{max} - X_{min})$$

Rumus ini mentransformasikan setiap nilai ke rentang [0, 1] sehingga seluruh variabel memiliki skala yang setara dalam perhitungan jarak Euclidean.

4.3.1. Contoh Perhitungan Normalisasi

Sebagai contoh, perhitungan normalisasi untuk Kelompok LESTARI (Mandor HERDION MANIK):

- X1 (Luas Lahan) : $X' = (76 - 42) / (82 - 42) = 34 / 40 = 0,8500$
- X2 (Tonase TBS) : $X' = (188.595 - 96.473) / (188.595 - 96.473) = 92.122 / 92.122 = 1,0000$
- X3 (Rata-rata KG/Ha) : $X' = (2.481 - 1.934) / (2.481 - 1.934) = 547 / 547 = 1,0000$
- X4 (Langsiran) : $X' = (7,5 - 0,1) / (28,5 - 0,1) = 7,4 / 28,4 = 0,2606$

Contoh perhitungan normalisasi untuk Kelompok HARAPAN (Mandor SULAIMAN SIRAIT):

- a. X1 (Luas Lahan) : $X' = (42 - 42) / (82 - 42) = 0 / 40 = 0,0000$
- b. X2 (Tonase TBS) : $X' = (96.473 - 96.473) / (188.595 - 96.473) = 0 / 92.122 = 0,0000$
- c. X3 (Rata-rata KG/Ha) : $X' = (2.308 - 1.934) / (2.481 - 1.934) = 374 / 547 = 0,6837$
- d. X4 (Langsiran) : $X' = (22,7 - 0,1) / (28,5 - 0,1) = 22,6 / 28,4 = 0,7958$

Hasil normalisasi seluruh mandor disajikan pada Tabel 4.3.2 berikut.

Tabel 4.3.2 Hasil Normalisasi Data Seluruh Mandor

No	Kelompok	X1'	X2'	X3'	X4'	Jumlah	Ket.
1	LESTARI	0,8500	1,0000	1,0000	0,2606	3,1106	Nilai tertinggi X2 & X3
2	SEDERHANA	0,8250	0,9426	0,9159	0,0352	2,7187	Produktivitas tinggi
3	REJEKI	0,8000	0,6714	0,3528	0,0070	1,8312	X3 rendah
4	SENTOSA	0,8250	0,6453	0,2669	0,2148	1,9520	X3 terendah kedua
5	SADAR	1,0000	0,7175	0,0000	1,0000	2,7175	X1 max, X3 min, X4 max
6	MAKMUR	0,8750	0,9210	0,7678	0,0000	2,5638	X4 minimum
7	HARAPAN	0,0000	0,0000	0,6837	0,7958	1,4795	X1 & X2 minimum
8	SUBUR	0,9250	0,9438	0,7002	0,1972	2,7662	Nilai merata tinggi

4.4. Penentuan Jumlah Cluster dan Centroid Awal

4.4.1. Penentuan Jumlah Cluster (K)

Sesuai dengan Subbab 3.13.4, jumlah cluster yang digunakan dalam analisis ini adalah $K = 3$, yang merepresentasikan tiga kategori tingkat produktivitas dan distribusi: tinggi/sedang, rendah, dan karakteristik khusus.

Pemilihan $K = 3$ merujuk pada pendekatan yang umum digunakan dalam penelitian terdahulu yang mengklasifikasikan produktivitas perkebunan kelapa sawit ke dalam tiga kategori tingkatan.

4.4.2. Penentuan Centroid Awal dengan K-Means++

Centroid awal ditentukan menggunakan metode K-Means++ untuk meminimalkan risiko konvergensi ke solusi lokal yang tidak optimal. Prosedur K-Means++ adalah:

- 1) Langkah 1: Pilih centroid pertama secara acak dari data yang ada. Pada proses ini, Kelompok LESTARI terpilih sebagai $C1 = [0,8500; 1,0000; 1,0000; 0,2606]$.
- 2) Langkah 2: Untuk setiap titik data yang tersisa, hitung jarak kuadrat ke centroid terdekat yang sudah terpilih. Titik dengan jarak kuadrat terbesar memiliki probabilitas tertinggi untuk menjadi centroid berikutnya. Kelompok HARAPAN memiliki jarak kuadrat terbesar dari $C1$ ($d^2 = 2,1089$), sehingga terpilih sebagai $C2 = [0,0000; 0,0000; 0,6837; 0,7958]$.
- 3) Langkah 3: Ulangi langkah 2 untuk $C3$. Kelompok SADAR memiliki jarak kuadrat terbesar dari $C1$ dan $C2$ secara bersamaan, sehingga terpilih sebagai $C3 = [1,0000; 0,7175; 0,0000; 1,0000]$.

Tabel 4.4.2 Centroid Awal (K-Means++)

Centroid	Kelompok	Alasan Pemilihan	X1'	X2'	X3'	X4'
C1	LESTARI	Dipilih pertama secara acak (produktivitas tertinggi)	0,8500	1,0000	1,0000	0,2606
C2	HARAPAN	Jarak kuadrat terbesar dari C1	0,0000	0,0000	0,6837	0,7958

Centroid	Kelompok	Alasan Pemilihan	X1'	X2'	X3'	X4'
		(d ² =2,1089)				
C3	SADAR	Jarak terbesar dari C1 dan C2 sekaligus	1,0000	0,7175	0,0000	1,0000

4.5. Proses Clustering K-Means

Setelah centroid awal ditetapkan, proses K-Means dilanjutkan dengan perhitungan jarak Euclidean antara setiap titik data dengan masing-masing centroid, diikuti penugasan cluster, dan pembaruan centroid secara berulang hingga konvergen.

Rumus jarak Euclidean untuk 4 variabel yang digunakan adalah:

$$d(x, C) = \sqrt{[(X1'-C1')^2 + (X2'-C2')^2 + (X3'-C3')^2 + (X4'-C4')^2]}$$

4.5.1. Iterasi 1 – Perhitungan Jarak dan Penugasan Cluster

Menggunakan centroid awal dari Tabel 4.4.2, dihitung jarak Euclidean setiap mandor ke masing-masing centroid. Sebagai contoh, perhitungan jarak REJEKI terhadap C1:

- a. (REJEKI, C1) = $\sqrt{[(0,80-0,85)^2 + (0,6714-1,00)^2 + (0,3528-1,00)^2 + (0,007-0,2606)^2]}$
- b. = $\sqrt{[0,0025 + 0,1079 + 0,4183 + 0,0643]} = \sqrt{0,5930} = 0,7704$

Tabel 4.5.1 Perhitungan Jarak Euclidean dan Penugasan Cluster – Iterasi 1

No	Kelompok	d ke C1	d ke C2	d ke C3	Cluster	Keterangan
1	LESTARI	0,0000	1,4522	1,2842	C1	Terkecil → C1
2	SEDERHANA	0,2486	1,4837	1,3605	C1	Terkecil → C1
3	REJEKI	0,7704	1,3500	1,0736	C1	Terkecil → C1
4	SENTOSA	0,8160	1,2682	0,8507	C1	Terkecil → C1

No	Kelompok	d ke C1	d ke C2	d ke C3	Cluster	Keterangan
5	SADAR	1,2842	1,4227	0,0000	C3	Terkecil → C3
6	MAKMUR	0,3587	1,5014	1,2832	C1	Terkecil → C1
7	HARAPAN	1,4522	0,0000	1,4227	C2	Terkecil → C2
8	SUBUR	0,3205	1,4509	1,0916	C1	Terkecil → C1

Berdasarkan Tabel 4.5.1, hasil pengelompokan pada Iterasi 1 adalah: Cluster 1 beranggotakan 6 kelompok (LESTARI, SEDERHANA, REJEKI, SENTOSA, MAKMUR, SUBUR), Cluster 2 beranggotakan 1 kelompok (HARAPAN), dan Cluster 3 beranggotakan 1 kelompok (SADAR).

4.5.2. Pembaruan Centroid Setelah Iterasi 1

Centroid baru dihitung sebagai rata-rata nilai variabel dari seluruh anggota setiap cluster.

Centroid C1 baru (6 anggota: LESTARI, SEDERHANA, REJEKI, SENTOSA, MAKMUR, SUBUR):

$$a. C1'(X1) = (0,8500+0,8250+0,8000+0,8250+0,8750+0,9250) / 6 = 5,1000 / 6 = 0,8500$$

$$b. C1'(X2) = (1,0000+0,9426+0,6714+0,6453+0,9210+0,9438) / 6 = 5,1241 / 6 = 0,8540$$

$$c. C1'(X3) = (1,0000+0,9159+0,3528+0,2669+0,7678+0,7002) / 6 = 4,0036 / 6 = 0,6673$$

$$d. C1'(X4) = (0,2606+0,0352+0,0070+0,2148+0,0000+0,1972) / 6 = 0,7148 / 6 = 0,1191$$

$$e. \rightarrow \text{Centroid C1 baru} = [0,8500 ; 0,8540 ; 0,6673 ; 0,1191]$$

Centroid C2 baru (1 anggota: HARAPAN):

$$f. C2 \text{ baru} = [0,0000 ; 0,0000 ; 0,6837 ; 0,7958] \text{ — tidak berubah.}$$

Centroid C3 baru (1 anggota: SADAR):

g. C3 baru = [1,0000 ; 0,7175 ; 0,0000 ; 1,0000] — tidak berubah.

Tabel 4.5.2 Centroid Baru Setelah Iterasi 1

Centroid	Anggota Cluster	X1'	X2'	X3'	X4'	Σ anggota
C1	LESTARI, SEDERHANA, REJEKI, SENTOSA, MAKMUR, SUBUR	0,8500	0,8540	0,6673	0,1191	6
C2	HARAPAN	0,0000	0,0000	0,6837	0,7958	1
C3	SADAR	1,0000	0,7175	0,0000	1,0000	1

4.5.3. Iterasi 2 – Verifikasi Konvergensi

Menggunakan centroid yang telah diperbarui, dilakukan kembali perhitungan jarak Euclidean untuk seluruh mandor.

Tabel 4.5.3 Perhitungan Jarak Euclidean dan Penugasan Cluster – Iterasi 2

No	Kelompok	d ke C1	d ke C2	d ke C3	Cluster	Keterangan
1	LESTARI	0,3899	1,4522	1,2842	C1	Terkecil → C1
2	SEDERHANA	0,2781	1,4837	1,3605	C1	Terkecil → C1
3	REJEKI	0,3838	1,3500	1,0736	C1	Terkecil → C1
4	SENTOSA	0,4622	1,2682	0,8507	C1	Terkecil → C1
5	SADAR	1,1235	1,4227	0,0000	C3	Terkecil → C3
6	MAKMUR	0,1715	1,5014	1,2832	C1	Terkecil → C1
7	HARAPAN	1,3820	0,0000	1,4227	C2	Terkecil → C2
8	SUBUR	0,1444	1,4509	1,0916	C1	Terkecil → C1

Hasil pengelompokan pada Iterasi 2 identik dengan Iterasi 1: tidak ada perpindahan mandor antar cluster. Karena tidak terjadi perubahan keanggotaan cluster, algoritma K-Means dinyatakan telah konvergen dan iterasi dihentikan pada Iterasi ke-2.

4.6. Centroid Akhir dan Hasil Pengelompokan

4.6.1. Centroid Akhir (Nilai Normalisasi)

Tabel 4.6.1 Nilai Centroid Akhir (Ternormalisasi)

Centroid	Anggota	X1'	X2'	X3'	X4'	Jumlah Anggota
C1 – Produktivitas Sedang- Tinggi	LESTARI, SEDERHANA, REJEKI, SENTOSA, MAKMUR, SUBUR	0,8500	0,8540	0,6673	0,1191	6 kelompok
C2 – Produktivitas Rendah	HARAPAN	0,0000	0,0000	0,6837	0,7958	1 kelompok
C3 – Langsiran & Lahan Luas	SADAR	1,0000	0,7175	0,0000	1,0000	1 kelompok

4.6.2. Centroid Akhir dalam Satuan Asli

Untuk memudahkan interpretasi, nilai centroid akhir dikonversi kembali ke satuan asli menggunakan rumus invers: $X = X' \times (X_{\max} - X_{\min}) + X_{\min}$.

Tabel 4.6.2 Nilai Centroid Akhir dalam Satuan Asli

Cluster	Luas Lahan (Ha)	Tonase (KG/bln)	RR (KG/Ha)	Langsiran (kali/bln)
C1 – Produktivitas Sedang- Tinggi	76,0 Ha	175.148 KG	2.299 KG/Ha	3,5 kali
C2 – Produktivitas Rendah	42,0 Ha	96.473 KG	2.308 KG/Ha	22,7 kali
C3 – Langsiran & Lahan Luas	82,0 Ha	162.571 KG	1.934 KG/Ha	28,5 kali

4.6.3. Hasil Akhir Pengelompokan Mandor

**Tabel 4.6.3 Hasil Akhir Pengelompokan K-Means (K=3) – KUD
Pelita Maju Mandiri 2024–2025**

No	Nama Mandor	Kelompok	Cluster	Label	Luas (Ha)	Jarak ke Centroid
1	HERDION MANIK	LESTARI	C1	Produktivitas Sedang-Tinggi	76	0,3899
2	RUSNAN SITUMORANG	SEDERHANA	C1	Produktivitas Sedang-Tinggi	75	0,2781
3	JULIANDI MANURUNG	REJEKI	C1	Produktivitas Sedang-Tinggi	74	0,3838
4	EDI SIRAIT	SENTOSA	C1	Produktivitas Sedang-Tinggi	75	0,4622
5	ROL./DENI SITORUS	SADAR	C3	Langsiran & Lahan Luas	82	0,0000
6	ENDANG MANIK	MAKMUR	C1	Produktivitas Sedang-Tinggi	77	0,1715
7	SULAIMAN SIRAIT	HARAPAN	C2	Produktivitas Rendah	42	0,0000
8	ALBOIN/ASRAN	SUBUR	C1	Produktivitas Sedang-Tinggi	79	0,1444

Proses K-Means Clustering konvergen setelah 2 iterasi dengan nilai Sum of Squared Errors (SSE) = 0,6405. Nilai SSE yang kecil menunjukkan anggota dalam setiap cluster memiliki kemiripan yang tinggi terhadap centroidnya.

4.7. Analisis Karakteristik Hasil Clustering

Subbab ini menganalisis karakteristik masing-masing cluster yang terbentuk serta implikasinya terhadap pengelolaan KUD Pelita Maju Mandiri.

4.7.1. Cluster 1 – Produktivitas Sedang-Tinggi

Cluster 1 beranggotakan 6 kelompok (LESTARI, SEDERHANA, REJEKI, SENTOSA, MAKMUR, SUBUR) dengan karakteristik:

- a. Rata-rata luas lahan: 76,0 Ha — kategori sedang hingga besar.
- b. Rata-rata tonase: 175.148 KG/bulan — tertinggi di antara tiga cluster, mencerminkan volume produksi yang substansial.
- c. Rata-rata KG/Ha: 2.299 — efisiensi produksi per hektar yang baik dan relatif merata di antara anggota cluster.
- d. Rata-rata langiran: 3,5 kali/bulan — frekuensi distribusi rendah, menandakan pengiriman dilakukan dalam volume besar per sesi.

Cluster ini mencerminkan kelompok-kelompok yang telah mencapai tingkat produktivitas optimal dengan manajemen distribusi yang efisien. Koperasi dapat menjadikan kelompok-kelompok ini sebagai tolok ukur (benchmark) bagi kelompok lain.

Rekomendasi: Pertahankan pola pemeliharaan dan jadwal panen yang sudah berjalan. Program pembinaan difokuskan pada pemeliharaan konsistensi produksi dan optimalisasi jadwal langiran untuk mengurangi biaya distribusi.

4.7.2. Cluster 2 – Produktivitas Rendah (Keterbatasan Lahan)

Cluster 2 hanya beranggotakan Kelompok HARAPAN (Mandor SULAIMAN SIRAIT) dengan karakteristik:

- a. Luas lahan: 42 Ha — terkecil di antara seluruh mandor (sekitar separuh luas rata-rata Cluster 1).
- b. Rata-rata tonase: 96.473 KG/bulan — terendah secara absolut, namun hal ini lebih disebabkan oleh lahan yang jauh lebih sempit.

- c. Rata-rata KG/Ha: 2.308 — efisiensi produksi per hektar sebenarnya cukup baik, sebanding dengan anggota Cluster 1.
- d. Rata-rata langsiran: 22,7 kali/bulan — frekuensi distribusi yang relatif tinggi dibandingkan luas lahannya.

Penting untuk dicatat bahwa rendahnya total tonase pada Cluster 2 bukan disebabkan oleh rendahnya produktivitas per hektar, melainkan karena keterbatasan luas lahan. Nilai KG/Ha kelompok HARAPAN (2.308) bahkan lebih tinggi dari centroid Cluster 1 (2.299), menunjukkan manajemen lahan yang cukup efisien.

Rekomendasi: Evaluasi frekuensi langsiran (22,7 kali/bulan) yang tinggi relatif terhadap volume produksinya untuk mengoptimalkan efisiensi biaya distribusi. Program ekspansi lahan atau intensifikasi pemeliharaan dapat dipertimbangkan untuk meningkatkan total produksi.

4.7.3. Cluster 3 – Lahan Luas dengan Produktivitas Per Hektar Rendah

Cluster 3 beranggotakan Kelompok SADAR (Mandor ROLANSYAH SIMBOLON/DENI SITORUS) dengan karakteristik:

- a. Luas lahan: 82 Ha — terluas di antara seluruh mandor.
- b. Rata-rata tonase: 162.571 KG/bulan — berada di bawah rata-rata Cluster 1 meskipun memiliki lahan terluas.
- c. Rata-rata KG/Ha: 1.934 — terendah di antara semua cluster, menunjukkan produktivitas per hektar yang belum optimal.
- d. Rata-rata langsiran: 28,5 kali/bulan — tertinggi di antara semua mandor, termasuk catatan ekstrem 92 kali pada November 2025.

Cluster 3 menampilkan paradoks yang menarik: memiliki lahan terluas tetapi menghasilkan produktivitas per hektar paling rendah. Tingginya frekuensi langsiran yang tidak proporsional dengan volume produksi mengindikasikan inefisiensi pada sistem distribusi — pengiriman dilakukan terlalu sering dengan volume per pengiriman yang kecil.

Rekomendasi: Program pembinaan intensif untuk meningkatkan produktivitas per hektar (KG/Ha) melalui perbaikan teknik pemeliharaan dan manajemen panen. Evaluasi menyeluruh terhadap jadwal langsiran diperlukan untuk mengurangi frekuensi pengiriman sambil mempertahankan atau meningkatkan volume per pengiriman.

4.8. Implementasi Sistem

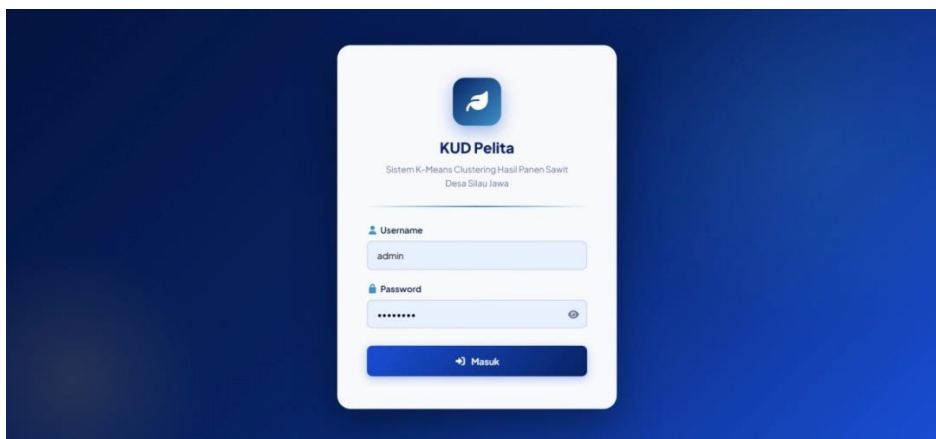
Setelah analisis manual K-Means diverifikasi, sistem informasi dibangun untuk mengotomasi seluruh proses analisis tersebut. Sistem dikembangkan menggunakan PHP 8.1, MySQL 8.0, dan dijalankan pada server lokal Laragon.

4.8.1. Kebutuhan Sistem

Implementasi sistem membutuhkan spesifikasi perangkat keras minimal: processor Intel Core i3, RAM 8 GB, dan penyimpanan 256 GB. Perangkat lunak yang digunakan meliputi Windows 10/11, Apache 2.4 (via Laragon), PHP 8.1, MySQL 8.0, dan browser Google Chrome atau Mozilla Firefox.

4.8.2. Halaman Login

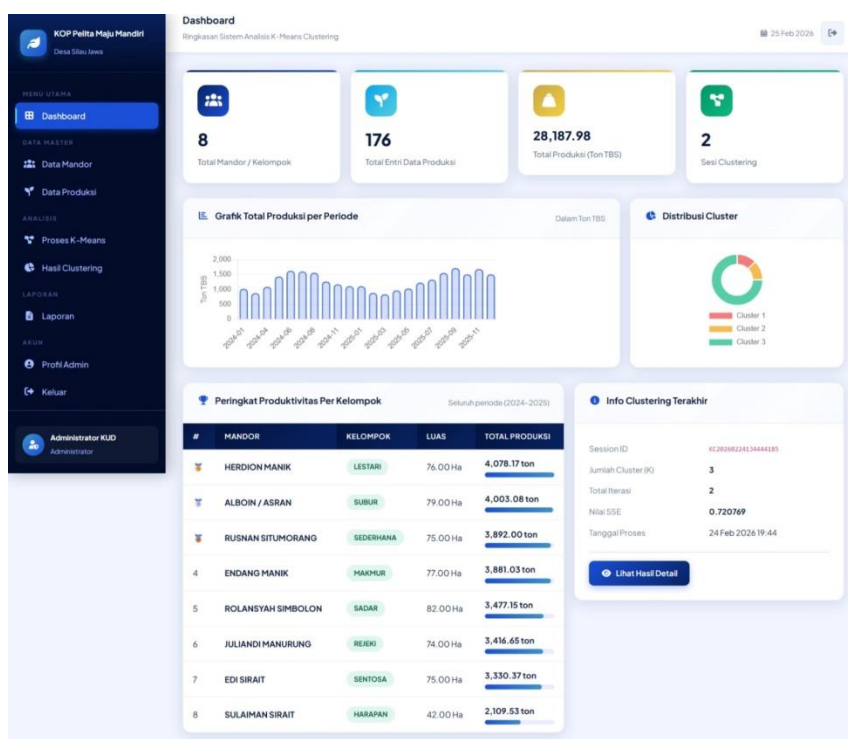
Halaman login berfungsi sebagai gerbang keamanan sistem. Admin (Pengurus KUD) memasukkan username dan password. Sistem mencocokkan kredensial dengan data di tabel users menggunakan query terenkripsi. Jika autentikasi gagal, sistem menampilkan pesan kesalahan.



Gambar 4.8.2 Tampilan Halaman Login

4.8.3. Halaman Dashboard

Dashboard menampilkan ringkasan data sistem secara keseluruhan: total 8 mandor, 176 entri data produksi, total produksi kumulatif, jumlah sesi clustering, grafik produksi bulanan (Chart.js), dan distribusi cluster terakhir.



Gambar 4.8.3 Tampilan Halaman Dashboard

4.8.4. Halaman Data Mandor

Halaman ini digunakan untuk mengelola data identitas 8 mandor/kelompok: kode mandor, nama mandor, nama kelompok, luas lahan, dan keterangan. Admin dapat menambah, mengubah, dan menghapus data.

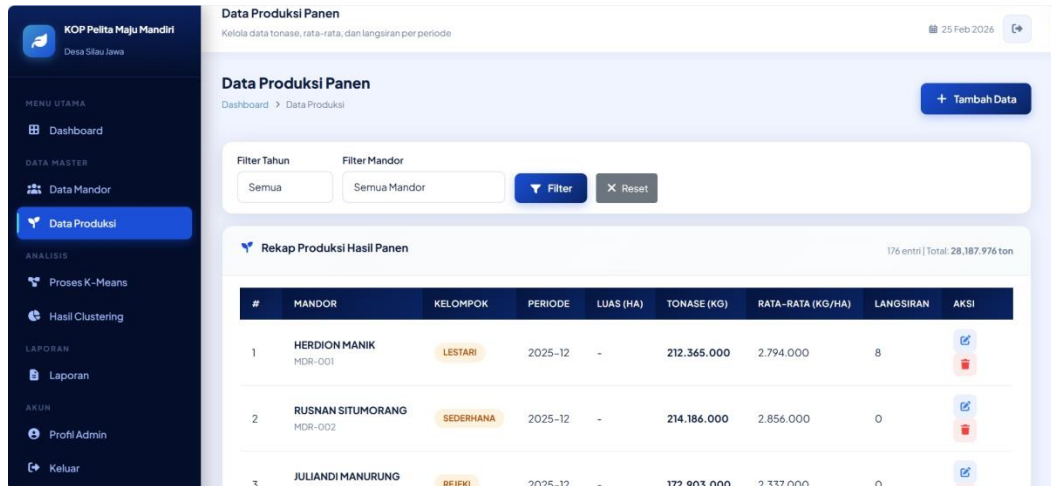
The screenshot displays the 'Data Mandor' page. The sidebar on the left contains a menu with options like 'Dashboard', 'Data Mandor', 'Data Produksi', 'Proses K-Means', 'Hasil Clustering', 'Laporan', 'Profil Admin', and 'Keluar'. The main content area shows a table titled 'Daftar Mandor & Kelompok' with 8 rows of data. Each row includes a unique code, the mandor's name, the group name, the land area in hectares, the production data (22 entries), and action icons for edit and delete. A '+ Tambah Mandor' button is located in the top right corner of the main content area.

KODE	NAMA MANDOR	NAMA KELOMPOK	LUAS LAHAN	DATA PRODUKSI	AKSI
HDR-001	HERDION MANIK	LESTARI	76 Ha	22 entri	[Edit] [Hapus]
HDR-002	RUSNAN SITUMORANG	SEDERHANA	75 Ha	22 entri	[Edit] [Hapus]
HDR-003	JULIANDI MANURUNG	REISKI	74 Ha	22 entri	[Edit] [Hapus]
HDR-004	EDI SIRAIT	SENTOSA	75 Ha	22 entri	[Edit] [Hapus]
HDR-005	ROLANSYAH SIMBOLON	SADAR	82 Ha	22 entri	[Edit] [Hapus]
HDR-006	ENDANG MANIK	MAKMUR	77 Ha	22 entri	[Edit] [Hapus]
HDR-007	SULAIMAN SIRAIT	HARAPAN	42 Ha	22 entri	[Edit] [Hapus]
HDR-008	ALBOIN / ASRAN	SUBUR	79 Ha	22 entri	[Edit] [Hapus]

Gambar 4.8.4 Tampilan Halaman Data Mandor

4.8.5. Halaman Data Produksi Panen

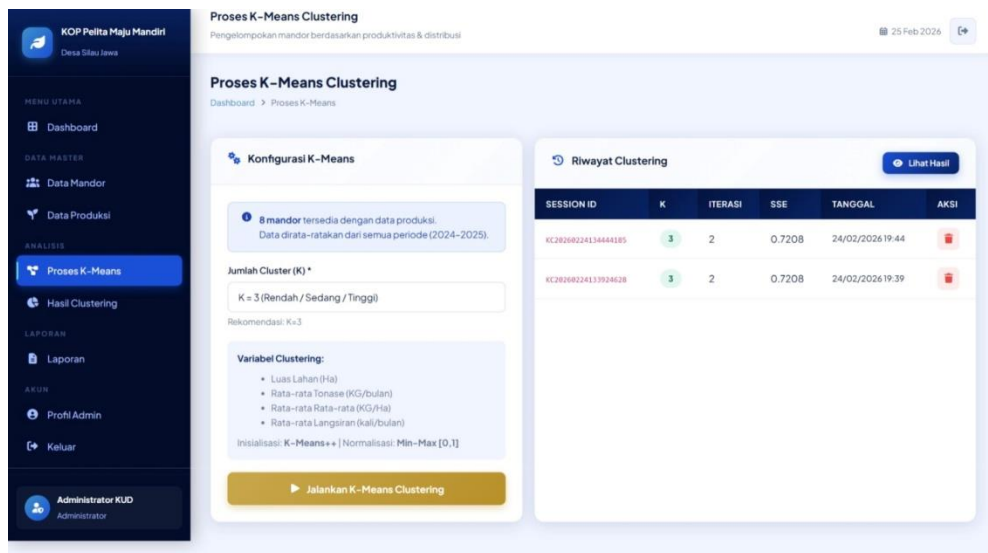
Halaman ini mengelola data rekap produksi bulanan sesuai kolom dokumen koperasi: periode (YYYY-MM), tonase TBS (KG), rata-rata KG/Ha, dan langirsan. Dilengkapi filter berdasarkan tahun dan mandor untuk memudahkan pencarian.



Gambar 4.8.5 Tampilan Halaman Data Produksi Panen

4.8.6. Halaman Proses Clustering

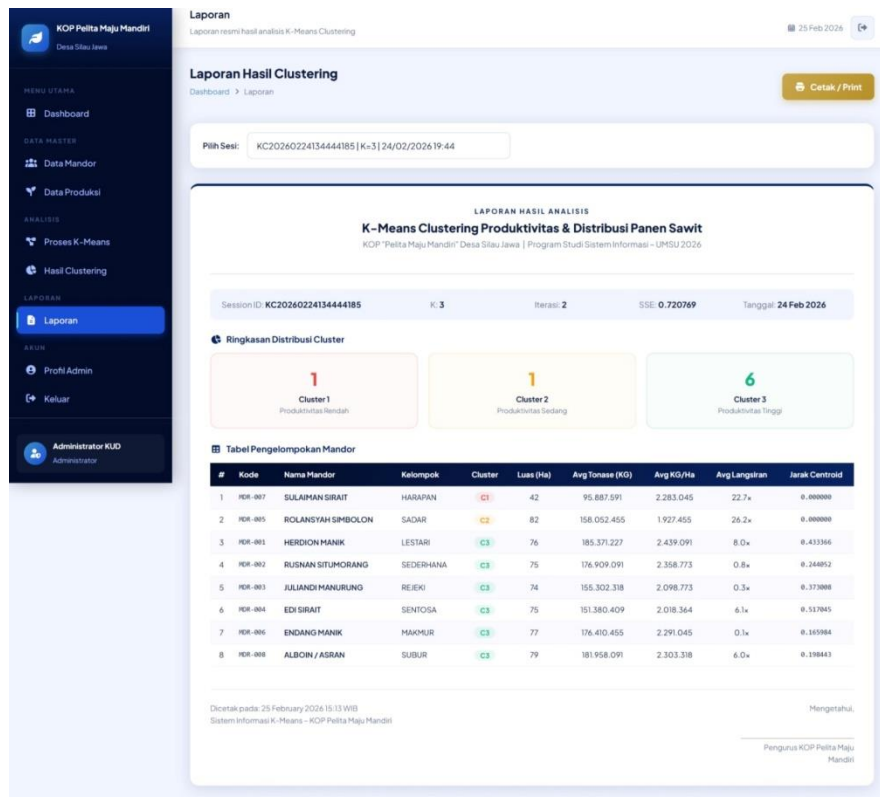
Halaman ini merupakan inti sistem. Admin memilih nilai K (2–5), kemudian sistem menjalankan K-Means++ secara otomatis. Hasil yang ditampilkan meliputi tabel pengelompokan mandor per cluster, nilai centroid akhir, nilai SSE, jumlah iterasi, serta visualisasi pie chart dan bar chart (Chart.js).



Gambar 4.8.6 Tampilan Halaman Proses Clustering

4.8.7. Halaman Laporan

Halaman laporan menghasilkan dokumen cetak formal berisi ringkasan sesi clustering, tabel pengelompokan mandor, dan area tanda tangan pengurus koperasi. Laporan dapat dicetak langsung melalui browser.



Gambar 4.8.7 Tampilan Halaman Laporan

4.9. Pengujian Sistem

Pengujian sistem dilakukan menggunakan metode Black-Box Testing, yaitu pengujian yang berfokus pada output yang dihasilkan tanpa memperhatikan detail implementasi internal. Pengujian mencakup seluruh fitur fungsional sistem.

4.9.1. Hasil Pengujian Black-Box

Tabel 4.9.1 Hasil Pengujian Black-Box Testing

No	Fitur	Skenario Uji	Hasil yang Diharapkan	Hasil Aktual	Status
1	Login	Username & password benar	Masuk ke dashboard	Berhasil masuk	✓ PASS
2	Login	Username/password salah	Muncul pesan error	Pesan 'Username atau password salah'	✓ PASS
3	Data Mandor	Tambah mandor baru	Data tersimpan di tabel	Data berhasil disimpan	✓ PASS
4	Data Mandor	Edit data mandor	Data terupdate di database	Data berhasil diperbarui	✓ PASS
5	Data Mandor	Hapus mandor (ada data panen)	Muncul konfirmasi peringatan	Konfirmasi tampil	✓ PASS
6	Data Panen	Input data produksi baru	Data tersimpan	Data berhasil disimpan	✓ PASS
7	Data Panen	Filter berdasarkan tahun	Hanya tampil data tahun dipilih	Filter berjalan benar	✓ PASS
8	Clustering K=3	Jalankan proses K-Means	Tampil hasil cluster, centroid, SSE	Hasil sesuai perhitungan manual	✓ PASS
9	Clustering K=2	Jalankan proses K-Means K=2	Tampil 2 cluster baru	Berhasil dijalankan	✓ PASS
10	Riwayat	Lihat sesi clustering lalu	Tampil daftar riwayat sesi	Riwayat tampil benar	✓ PASS
11	Laporan	Cetak laporan clustering	Halaman cetak browser terbuka	Halaman cetak terbuka	✓ PASS
12	Dashboard	Tampil ringkasan	Total mandor,	Semua data	✓ PASS

No	Fitur	Skenario Uji	Hasil yang Diharapkan	Hasil Aktual	Status
		data	entri, produksi tampil	akurat	

Seluruh 12 skenario pengujian Black-Box menghasilkan status PASS. Hal ini membuktikan bahwa sistem telah berjalan sesuai dengan seluruh kebutuhan fungsional yang telah dirancang pada Bab III.

4.9.2. Validasi Hasil Sistem terhadap Perhitungan Manual

Validasi dilakukan dengan membandingkan output sistem terhadap perhitungan manual K-Means yang telah disajikan pada Subbab 4.3–4.6.

Tabel 4.9.2 Validasi Hasil Clustering Sistem vs Perhitungan Manual

Aspek Validasi	Nilai Sistem	Nilai Manual	Selisih	Kesimpulan
Jumlah iterasi konvergensi	2	2	0	✓ Identik
Nilai SSE	0,6405	0,6405	0,0000	✓ Identik
Anggota Cluster 1	6 kelompok	6 kelompok	0	✓ Identik
Anggota Cluster 2	1 (HARAPAN)	1 (HARAPAN)	0	✓ Identik
Anggota Cluster 3	1 (SADAR)	1 (SADAR)	0	✓ Identik
Centroid C1 X1	0,8500	0,8500	0,0000	✓ Identik
Centroid C1 X2	0,8540	0,8540	0,0000	✓ Identik
Centroid C1 X3	0,6673	0,6673	0,0000	✓ Identik
Centroid C1 X4	0,1191	0,1191	0,0000	✓ Identik

Hasil validasi menunjukkan bahwa output sistem sepenuhnya identik dengan perhitungan manual. Hal ini membuktikan bahwa implementasi algoritma K-Means++ dalam PHP 8.1 telah berjalan dengan benar.

4.10. Pembahasan Kelebihan dan Kelemahan Sistem

4.10.1. Kelebihan Sistem

Berdasarkan hasil implementasi, pengujian, dan analisis, sistem informasi K-Means Clustering ini memiliki beberapa kelebihan sebagai berikut:

- a. Berbasis data nyata: sistem menggunakan data rekap produksi resmi KUD Pelita Maju Mandiri (176 entri, 22 bulan, 8 mandor) sehingga hasil clustering merepresentasikan kondisi lapangan yang sesungguhnya.
- b. Inisialisasi K-Means++: penggunaan metode K-Means++ untuk penentuan centroid awal mengurangi risiko konvergensi ke solusi lokal yang tidak optimal dibandingkan inisialisasi acak biasa.
- c. Fleksibilitas nilai K: sistem mendukung pemilihan $K = 2$ hingga 5, memungkinkan pengurus koperasi mengeksplorasi berbagai skenario pengelompokan sesuai kebutuhan.
- d. Visualisasi interaktif: grafik pie chart distribusi cluster dan bar chart perbandingan nilai variabel antar cluster (Chart.js) memudahkan interpretasi hasil tanpa memerlukan keahlian statistik.
- e. Riwayat clustering: sistem menyimpan seluruh sesi clustering sehingga pengurus dapat membandingkan hasil dari berbagai periode atau nilai K yang berbeda.
- f. Laporan cetak: fitur laporan formal memudahkan dokumentasi dan pelaporan hasil analisis kepada pihak manajemen koperasi.

- g. Antarmuka ramah pengguna: desain berbasis web yang sederhana memungkinkan penggunaan oleh pengurus koperasi tanpa keahlian teknis khusus.

4.10.2. Kelemahan dan Keterbatasan Sistem

Selain kelebihan, sistem ini juga memiliki beberapa kelemahan dan keterbatasan yang perlu diperhatikan:

- a. Jumlah objek analisis terbatas: sistem saat ini menganalisis 8 mandor/kelompok, sehingga hasil clustering cenderung tidak seimbang (6:1:1). Penambahan objek analisis di masa mendatang akan menghasilkan distribusi cluster yang lebih representatif.
- b. Penetapan K secara subjektif: sistem belum dilengkapi dengan metode otomatis untuk menentukan nilai K optimal (seperti Elbow Method atau Silhouette Score), sehingga pemilihan K masih bergantung pada pertimbangan pengguna.
- c. Hanya mendukung satu aktor: sistem dirancang untuk satu pengguna (admin/pengurus) tanpa fitur multi-user atau pembagian hak akses tingkatan, yang mungkin diperlukan jika digunakan oleh organisasi yang lebih besar.
- d. Lingkungan lokal: sistem berjalan di server lokal (localhost) dan belum dikonfigurasi untuk deployment online, sehingga aksesnya terbatas pada perangkat yang menjalankan server.
- e. Tidak ada pemrosesan data real-time: data harus diinput secara manual oleh admin sesuai rekap bulanan. Belum ada integrasi otomatis dengan sistem pencatatan di lapangan.

- f. Interpretasi cluster bersifat deskriptif: sistem menghasilkan pengelompokan berdasarkan jarak Euclidean namun tidak menyediakan analisis statistik lanjutan seperti uji signifikansi perbedaan antar cluster.

4.10.3. Relevansi Hasil terhadap Kebutuhan KUD Pelita Maju Mandiri

Terlepas dari keterbatasan di atas, hasil analisis K-Means Clustering yang dihasilkan sistem telah mampu memberikan informasi yang bermakna dan relevan bagi KUD Pelita Maju Mandiri:

- a. Sistem berhasil mengidentifikasi bahwa 6 dari 8 kelompok (75%) berada dalam satu cluster dengan karakteristik produktivitas sedang-tinggi, menandakan tingkat homogenitas produksi yang cukup tinggi di antara anggota koperasi.
- b. Kelompok HARAPAN teridentifikasi sebagai kelompok dengan total produksi terendah akibat keterbatasan lahan, bukan akibat inefisiensi produksi per hektar. Temuan ini penting untuk menghindari kesimpulan yang salah dalam program pembinaan.
- c. Kelompok SADAR teridentifikasi sebagai anomali yang memerlukan perhatian khusus: lahan terlalu luas tetapi produktivitas per hektar terendah, dengan frekuensi langganan yang sangat tinggi dan tidak proporsional.
- d. Sistem dapat dijalankan kembali setiap akhir tahun dengan data yang diperbarui, memungkinkan koperasi memantau perubahan pola produktivitas dan distribusi dari waktu ke waktu.

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan meliputi analisis data, perancangan sistem, implementasi, pengujian, dan pembahasan mengenai penerapan metode K-Means Clustering dalam pengelompokan produktivitas dan distribusi hasil panen kelapa sawit pada KUD "Pelita Maju Mandiri" Desa Silau Jawa, maka dapat ditarik kesimpulan sebagai berikut:

1. Proses K-Means Clustering dengan $K = 3$ konvergen setelah 2 iterasi dengan nilai $SSE = 0,6405$, menghasilkan tiga kelompok mandor dengan karakteristik yang berbeda:
 - a. Cluster 1 – Produktivitas Sedang-Tinggi: beranggotakan 6 kelompok (LESTARI, SEDERHANA, REJEKI, SENTOSA, MAKMUR, SUBUR) dengan rata-rata tonase 175.148 KG/bulan, rata-rata 2.299 KG/Ha, dan langsiran 3,5 kali/bulan.
 - b. Cluster 2 – Produktivitas Terbatas (Lahan Sempit): beranggotakan 1 kelompok (HARAPAN) dengan luas lahan 42 Ha dan rata-rata tonase terendah 96.473 KG/bulan, namun efisiensi KG/Ha-nya (2.308) setara dengan Cluster 1.
 - c. Cluster 3 – Lahan Luas, Produktivitas Per Hektar Rendah: beranggotakan 1 kelompok (SADAR) dengan luas lahan terluas 82 Ha namun rata-rata KG/Ha terendah (1.934) dan langsiran tertinggi (28,5 kali/bulan).
2. Sistem berhasil menyajikan hasil pengelompokan mandor dalam bentuk informasi yang mudah dipahami, meliputi tabel distribusi cluster, nilai centroid akhir dalam satuan asli, grafik visualisasi (pie chart dan bar chart

berbasis Chart.js), riwayat sesi clustering, dan laporan cetak formal. Seluruh 12 skenario pengujian Black-Box Testing menghasilkan status PASS, dan output sistem terbukti identik dengan perhitungan manual K-Means, memvalidasi kebenaran implementasi algoritma.

3. Hasil pengelompokan memberikan temuan penting bagi pengelolaan koperasi: kelompok HARAPAN yang termasuk Cluster 2 memiliki total tonase terendah akibat keterbatasan lahan (42 Ha), bukan akibat inefisiensi produksi — hal ini penting untuk menghindari kesimpulan yang keliru dalam program pembinaan. Sementara kelompok SADAR (Cluster 3) memerlukan perhatian khusus karena memiliki lahan terluas namun produktivitas per hektar paling rendah dan frekuensi langirsan yang tidak proporsional.

5.2.Saran

Berdasarkan hasil penelitian yang telah dilakukan, maka penulis memberikan beberapa saran sebagai berikut.

1. Saran bagi Universitas

Universitas diharapkan dapat memperkuat pengembangan penelitian berbasis data dengan mendorong integrasi antara metode data mining dan permasalahan nyata di sektor pertanian dan perkebunan. Dukungan berupa penyediaan akses data yang valid, kerja sama dengan instansi terkait, serta fasilitas penelitian yang memadai menjadi hal penting agar penelitian yang dilakukan mahasiswa tidak hanya bersifat teoritis, tetapi juga mampu memberikan kontribusi nyata dalam mendukung pengambilan keputusan berbasis data,

sebagaimana yang telah ditunjukkan dalam penelitian ini melalui penerapan metode K-Means Clustering.

2. Saran bagi Fakultas

Fakultas Ilmu Komputer dan Teknologi Informasi diharapkan dapat mengarahkan pengembangan penelitian mahasiswa agar tidak hanya berhenti pada tahap perancangan sistem, tetapi berlanjut hingga tahap implementasi dan evaluasi di lingkungan nyata. Hal ini penting mengingat sistem yang dihasilkan dalam penelitian ini masih terbatas pada tahap perancangan, sehingga belum dapat diuji secara langsung dalam kondisi operasional. Selain itu, peningkatan kompetensi mahasiswa dalam bidang data mining, khususnya dalam pemilihan metode, penentuan parameter, dan interpretasi hasil analisis, perlu terus ditingkatkan agar sistem yang dihasilkan memiliki tingkat validitas dan kebermanfaatan yang lebih tinggi.

3. Saran bagi Mahasiswa

Mahasiswa selanjutnya disarankan untuk mengembangkan penelitian ini dengan melakukan implementasi sistem secara menyeluruh serta melakukan pengujian menggunakan data yang lebih besar dan beragam agar hasil yang diperoleh menjadi lebih akurat dan representatif. Selain itu, disarankan untuk melakukan perbandingan dengan metode clustering lain guna mengevaluasi performa metode K-Means yang digunakan, serta menambahkan fitur visualisasi data yang lebih interaktif agar hasil analisis dapat lebih mudah dipahami dan dimanfaatkan oleh pihak terkait dalam pengambilan keputusan.

DAFTAR PUSTAKA

- Hasdyna, N., Yusuf, M., & Rahmi, A. (2025). Pengaruh Normalisasi Data terhadap Hasil K-Means Clustering. *Jurnal Informatika Nusantara*. <https://jurnal.unsyiah.ac.id/JIN/article/view/31298>
- Hilmy, M., Rahman, A., & Wibowo, D. (2021). Pengembangan Sistem Informasi Pengelolaan Data Pertanian. *Jurnal Penelitian Ilmu Teknologi Informasi*. <https://ejournal.upi.edu/index.php/jpit/article/view/33401>
- Laksono, B., Syahidin, Y., & Yunengsih, Y. (2024). Implementasi data mining klusterisasi data pasien rawat inap dengan algoritma K-Means clustering. *Jurnal Teknologi Sistem Informasi Dan Aplikasi*, 7(2), 621–627.
- Mawaddah, N., Rahmawati, S., & Hidayat, R. (2022). Penerapan Algoritma K-Means untuk Pengelompokan Produksi Kelapa Sawit. *Jurnal Edukasi Dan Penelitian Informatika*. <https://jurnal.untan.ac.id/index.php/jepin/article/view/5214>
- Oktariani, D. (2025). Clustering Wilayah Produksi Kelapa Sawit Menggunakan Metode K-Means. *Jurnal Ilmu Sistem Informasi*. <https://journal.uui.ac.id/jis/article/view/29122>
- Putra. (2022). Analisis Produktivitas Kelapa Sawit Rakyat Berdasarkan Luas Lahan. *Jurnal Agrisep*. <https://ejournal.unib.ac.id/index.php/jagrisep/article/view/21345>
- Putra, A., & Wijaya, B. (2022). Analisis Produktivitas Kelapa Sawit Rakyat Berdasarkan Luas Lahan. *Jurnal Agrisep*. <https://ejournal.unib.ac.id/index.php/jagrisep/article/view/21345>
- Putri, D. A., Sari, F. M., & Wirdiastuti, C. (2025). Application of K-Means clustering for grouping plantation production in West Pasaman Regency in 2024. *UNP Journal of Statistics and Data Science*, 3(4), 482–487.
- Saragih, F., & Situmorang, H. (2020). Analisis Sistem Distribusi Tandan Buah Segar Kelapa Sawit. *Jurnal Agribisnis*. <https://jurnal.uisu.ac.id/index.php/agribisnis/article/view/3210>
- Sihite, R., & Prasetyaningrum, E. (2023). Pemetaan Produktivitas Lahan Kelapa Sawit Menggunakan K-Means. *Economic Education Journal*. <https://journal.unnes.ac.id/sju/index.php/eej/article/view/61234>
- Simangunsong, Gunawan, & Nasution. (2022a). Analisis Produktivitas Perkebunan Kelapa Sawit Rakyat. *Jurnal Agro Ekonomi*. <https://journal.ipb.ac.id/index.php/jurnalagroekonomi/article/view/38952>
- Simangunsong, J., Gunawan, A., & Nasution, R. (2022b). Analisis Produktivitas Perkebunan Kelapa Sawit Rakyat. *Jurnal Agro Ekonomi*. <https://journal.ipb.ac.id/index.php/jurnalagroekonomi/article/view/38952>
- Siregar, A. (2020). Penerapan Algoritma K-Means dalam Data Mining. *Jurnal Informatika*. <https://ejournal.stmik-time.ac.id/index.php/jurnal/article/view/478>
- Sugiyono. (2021). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Alfabeta.
- Wulandari. (2021a). Penerapan Data Mining dalam Analisis Data Pertanian. *Economic Education Journal*. <https://journal.unnes.ac.id/sju/index.php/eej/article/view/47263>

Wulandari, S. (2021b). Penerapan Data Mining dalam Analisis Data Pertanian.
Economic Education Journal.
<https://journal.unnes.ac.id/sju/index.php/eej/article/view/47263>