

# DATA MINING *K-MEDOIDS* DAN *K-MEANS* UNTUK PENGELOMPOKAN POTENSI PRODUKSI KELAPA SAWIT DI INDONESIA

Faiq Husain Pratama<sup>1)</sup>, Agung Triayudi <sup>\*2)</sup> dan Eri Mardiani <sup>3)</sup>

<sup>1,2,3)</sup> Fakultas Teknologi Komunikasi dan Informatika, Universitas Nasional

Jl. Sawo Manila, Pasar Minggu, Jakarta Selatan, Daerah Khusus Ibukota Jakarta, 12520

e-mail: [faiqhusain3@gmail.com](mailto:faiqhusain3@gmail.com) <sup>1)</sup>, [agungtriayudi@civitas.unas.ac.id](mailto:agungtriayudi@civitas.unas.ac.id) <sup>2)</sup>, [Erimardiani1@gmail.com](mailto:Erimardiani1@gmail.com) <sup>3)</sup>

\*Penulis Korespondensi

## ABSTRAK

*Kelapa sawit merupakan tanaman golongan palma yang memiliki periode produksi setiap tahunnya. Penyebaran terbesar kelapa sawit berada di Indonesia. Indonesia memiliki luas perkebunan mencapai 17,32 juta hektar. Detailnya hasil produksi 26,57 ton dengan luas kebun 8,51 juta hektar. Menurut data USDA, pada tahun 2022 menurun karena berbagai faktor. Untuk itu perlu dilakukan klasifikasi potensi produksi kelapa sawit dan identifikasi peluang keberhasilan produksi disetiap lokasi perkebunan kelapa sawit. Dengan ini dilakukan penelitian pembuatan sistem clustering untuk melihat potensi produksi kelapa sawit dengan memakai kombinasi 2 metode yaitu K-Medoids dan K-Means. K-Medoids berfungsi untuk penentuan cluster sesuai dengan data variable yang paling rendah(Cluster 1) 18 wilayah, sedang (Cluster 2) 5 wilayah, dan tinggi (Cluster 3)/(Centroid) 2 wilayah pada potensi hasil Luas areal, produksi, dan produktivitas kelapa sawit. Algoritma K-Means berfungsi untuk mengelompokkan rata rata luas tanah 514.885,72 Ha, produksi 1.931.882,84 Ton dan produktifitas 3.227,08 Kg/Ha, dengan pembagian potensi rendah, sedang dan tinggi. Kombinasi dari kedua algoritma berfungsi sangat baik karena masing masing memiliki peran tersendiri yang sesuai dengan kebutuhan penelitian. Dari penggabungan 2 metode K-Medoids dan K-Means mendapatkan hasil ketiga klaster bahwa hasil produksi kelapa sawit yang memiliki potensi rendah 72% sedang 20%, tinggi 2%. Segmentasi ini disebabkanoleh kesamaan karakteristik perkebunan berdasarkan kesamaan dari luas, produksi, dan produktivitas. Yang memiliki potensi produksi tertinggi kelapa sawit ada 2 provinsi yaitu Kalimantan Barat dan Riau.*

**Kata Kunci:** Clustering, Kelapa Sawit, K-Medoids, K-Means

## ABSTRACT

*Oil palm is a type of palm plant that has a production period every year. The largest distribution of oil palm is in Indonesia. Indonesia has a plantation area of 17.32 million hectares. In detail, the production yield is 26.57 tons with a garden area of 8.51 million hectares. According to USDA data, in 2022 it will decline due to various factors. For this reason, it is necessary to classify the potential for oil palm production and the chances of successful production at each location of oil palm plantations. With this, a research was conducted on making a clustering system to see the potential for oil palm production by using a combination of 2 methods, namely K-Medoids and K-Means. K-Medoids function to enable clusters according to the data variables with the lowest (Cluster 1) 18 regions, medium (Cluster 2) 5 regions, and high (Cluster 3)/(Centroid) 2 regions on potential yield Area, production, and palm oil productivity. The K-Means algorithm is used to classify the average land area of 514,885.72 Ha, production of 1,931,882,84 tons and productivity of 3,227,08 Kg/Ha, with the distribution of low, medium and high potential. The combination of the two works very well because each has its own role according to the research needs. From the combination of the 2 methods K-Medoids and K-Means, the results of the three clusters show that the yield of palm oil which has low potential is 72%, 20%, high is 2%. This segmentation is caused by the similarity of plantation characteristics based on the similarity of area, production, and productivity. There are 2 provinces that have the highest production potential of palm oil, namely West Kalimantan and Riau.*

**Keywords:** Clustering, Palm Oil, K-Medoids, K-Means

## I. PENDAHULUAN

**D**I Indonesia tersebar jutaan hektar perkebunan kelapa sawit. Tanaman dari kelompok palma ini memiliki waktu produksi tahunan yang memungkinkan mereka tumbuh di daerah tropis. Tanaman kelapa sawit tumbuh secara alami di Afrika Barat, atau lebih tepatnya di daratan Nigeria. Dengan ditanamnya 4 bibit kelapa sawit di Kebun Raya Bogor pada tahun 1848, menjadikan titik awal dari budidaya tanaman tersebut di Nusantara. Karena banyaknya manfaat yang diperoleh, tanaman ini fokus dibudidayakan sejak tahun 1910 di wilayah Sumatera hingga Jawa, tepatnya di Jawa Barat. Tujuan dari dibudidayakannya kelapa sawit tidak lain untuk terus menjaga dan mengembangkan perkebunan karena perkebunan inilah yang menjadi benang merah

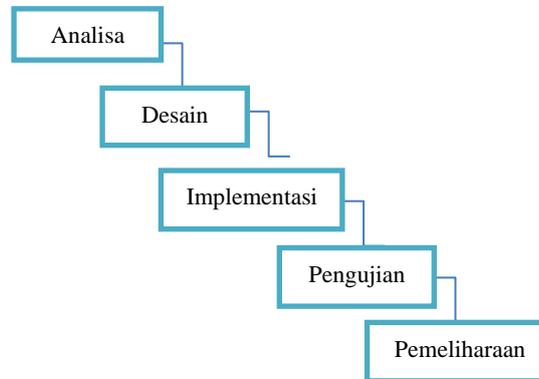
dari kesediaan minyak sawit dunia. Pada tahun 2022 mengalami penurunan, masalah penurunan produksi kelapa sawit di wilayah Jawa Barat dan sebagian provinsi di Indonesia merupakan masalah yang sebelumnya tidak dapat diprediksi karena dipengaruhi oleh faktor alam. Dan saat itu survey ini dilakukan, Indonesia sedang mengalami penurunan persediaan minyak sawit. Menurut GAPKI (Gabungan Pengusaha Kelapa Sawit Indonesia), faktor cuaca menjadi penyebab turunnya produksi kelapa sawit pada tahun 2021.

Beberapa penelitian juga telah dilakukan untuk mengatasi masalah penurunan produksi kelapa sawit diantaranya yaitu adanya kajian dengan judul “Analisis Pengelompokan Tanaman Kelapa Sawit Pada Tingkat Potensi Tertentu Berdasarkan Luas Tanaman Menggunakan Algoritma K-Means Badan Pusat Statistik Sumatera Bagian Utara” dan “Penerapan Metode K-Means untuk Clustering Produk Toko Online dalam Menentukan Tingkat Persediaan”. Kedua penelitian tersebut sedikitnya memiliki tujuan yang sama yaitu untuk memahami potensi kelapa sawit serta untuk mengklasifikasikan tingkat produksi pada perkebunan kelapa sawit [1]. Metode yang tepat untuk digunakan untuk melihat dan menangkap potensi produksi perkebunan kelapa sawit yang mempunyai persamaan karakteristik yang didasarkan pada luas area, hasil/produksi, dan produktivitas adalah metode Algoritma K-Medoids dan K-Means. K-Medoids termasuk metode clustering partisi yang mengelompokkan satu set objek ke dalam satu set kluster. Penelitian ini menggunakan kombinasi kedua algoritma untuk mengklasifikasikan potensi lokasi buah kelapa sawit Indonesia. Algoritma K-Means ialah langkah-langkah pengklasifikasian data yang non-hierarki dengan membagi data menjadi dua atau lebih kelompok [2]. Kelemahan dari penelitian tersebut adalah belum mampu melihat potensi produksi kelapa sawit secara mengerucut. Berdasarkan permasalahan diatas, telah dikembangkan sebuah sistem yang mengimplementasikan proses clustering untuk produksi kelapa sawit menggunakan kombinasi algoritma K-Medoids dan K-Means. Sehingga sistem tersebut dapat digunakan untuk mengidentifikasi produsen kelapa sawit yang potensial dan profitable sebab sebagai negara agraris dengan iklim tropis, Indonesia didukung oleh sub-sektor pertanian [3]. Beberapa penelitian telah dilakukan di beberapa wilayah di Indonesia untuk mendukung produksi kelapa sawit di Indonesia. Titik fokus dari penelitian ini terletak pada bagian produksi kelapa sawit. Trisna Yuniarti dan Dahliyah Hayati juga mengelompokkan potensi produksi kelapa sawit mereka sendiri pada tahun 2021. Tujuan dari penelitian tersebut untuk mengkluster perkebunan kelapa sawit di Nusantara berdasar kepada produktivitas areal.

Hampir disetiap belahan wilayah Indonesia terdapat perkebunan kelapa sawit. Namun dari segi produktivitas, luas, dan total produksi di tiap wilayah, belum jelas potensi perkebunan seperti apa. Untuk mengetahui lebih jelas, teknik data mining clustering K-Means dapat digunakan sebagai metode untuk melihat potensi paling subur. Beberapa bab penelitian telah dilakukan menggunakan metode K-Means guna mengelompokkan data penjualan [4, 5]. Ada studi lain yang juga menggunakan metode ini untuk mengelompokkan distribusi produk [6, 7]. Teknik itupun juga dapat disandingkan bersama metode lain seperti pada penelitian yang ditujukan untuk memperkirakan profil penggunaan beban panas skala besar di sektor residensial terfokus dalam industri manufaktur [8]. Dengan teknik yang sama, dapat dilakukan studi clustering yang difokuskan pada perkebunan kelapa sawit untuk memetakan hasil/produksi buah kelapa sawit, menganalisis blok perkebunan kelapa sawit yang paling produktif dan yang tidak. Menentukan pengelompokan jenis kelapa sawit yang produktif untuk ditingkatkan dan mengupayakan kelapa sawit tidak produktif menjadi produktif [9]. Clustering/kluster merupakan suatu teknik pengelompokan data sesuai dengan jenis data yang sama [10]. Sehingga peneliti dapat menggunakan data ini lebih dalam untuk mendapatkan wawasan dan informasi baru secara lengkap dan terperinci. K-Means adalah pengelompokan berdasarkan kesamaan antara objek dan statis yang memiliki waktu linier rendah. K-Means biasanya digunakan dalam data mining [11]. Berdasarkan penelitian-penelitian sebelumnya, peneliti mengembangkan sistem klasifikasi potensi produksi kelapa sawit dengan menggabungkan dua algoritma yaitu algoritma K-Medoids dan K-Means dalam mengelompokkan potensi produksi kelapa sawit di Indonesia. Sistem yang diusulkan adalah sistem berbasis web, sehingga perusahaan dapat mengaksesnya kapan saja dan dari platform mana saja. Sistem aplikasi yang dikembangkan juga berbasis web, sehingga memberikan kemudahan bagi user/pengguna smartphone. Pengembangan sistem aplikasi menggunakan PHP, Python, Visual Code, Waterfall, dan Mysql. Suatu ciptaan halaman web yang dinamis menggunakan PHP karena memiliki keunggulan melakukan koneksi berbagai macam software sistem manajemen (DBMS). Konektivitas beberapa DBMS juga sangat baik, antara lain Sybase, Oracle, Mysql, MsSql, Solid, SQL Server, Adabas, Postgres SQL, Velocis, Filepro, Unix Dbm, dBase, dan lain lain [12]. Tujuan dari sebuah penelitian ini adalah untuk menciptakan sistem pengelompokan potensial produksi kelapa sawit dengan mengklasifikasikan perkebunan kelapa sawit di Indonesia berdasarkan luas area, hasil produksi dan produktivitas. Diharapkan sistem yang dibuat ini mampu memberikan kedetailan dalam menyampaikan informasi dan memiliki fungsi dalam menyusun strategi peningkatan produksi kelapa sawit di Indonesia dengan menggunakan metode/teknik kombinasi Clustering K-Means dan K-Medoids.

## II. METODE PENELITIAN

Penelitian tentang pengelompokan potensi produksi kelapa sawit di Indonesia menggunakan metode *waterfall*. Metode tersebut terlihat Gambar 1 di bawah:

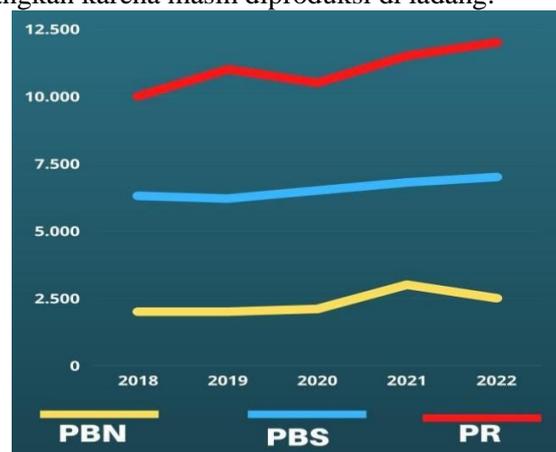


Gambar 1. Metode waterfall

Pada tahap pengembangan sebuah sistem, metode ini bersifat sistematis, serta rangkaian dalam sistem perancangan perangkat lunak dilakukan secara urut. Metode *waterfall* atau disebut juga metode air terjun merupakan metode yang sering dipakai untuk melakukan pembaharuan pada sistem aplikasi yang sedang berjalan. Terdapat beberapa tahapan pada metode ini, yaitu analisis aplikasi, desain aplikasi, implementasi aplikasi, pengujian aplikasi, dan pemeliharaan [13].

### A. Analisa

Pada fase ini, mengumpulkan persyaratan untuk sistem, memahami perangkat lunak yang dibuat, dan mendokumentasikan hasil analisis dalam spesifikasi persyaratan. Dalam penelitian ini, kebutuhan perangkat lunak diambil dari tinjauan literatur dan perbandingan dengan aplikasi yang ada, serta data yang diperoleh dari Direktorat Perkebunan Kementerian Pertanian Republik Indonesia tahun 2021 melalui *website* <https://ditjenbun.pertanian.go.id/>. Dengan berdasarkan pada data tersebut, variable diberikan pada setiap ukuran data yang digunakan. X1 (Data Luas Area), X2 (data hasil/produksi), dan X3 (Data Produktivitas). Untuk data lahan pertanian dihitung total Tanaman Unggul (TM) dan Tanaman Rusak (TR). Tanaman yang rusak sudah tidak mencukupi lagi, tetapi dipertimbangkan karena masih diproduksi di ladang.



Gambar 2. Produktivitas kelapa sawit Indonesia tahun 2018 – 2022

Gambar di atas merupakan data produktivitas kelapa sawit di Indonesia berdasarkan provinsi dan keadaan tanaman tahun 2022 yang penulis peroleh dari publikasi Kementerian Pertanian RI yang terdiri dari luas areal dan produksi kelapa sawit pada PR (Perkebunan Rakyat) berwarna merah, PBN (Perkebunan Besar Negara) berwarna kuning dan PBS (Perkebunan Besar Swasta) berwarna biru.

Pada Tabel I membahas mengenai data *mining* untuk mengklasifikasikan luas area, hasil/produksi, dan produktivitas dari sebuah perkebunan kelapa sawit sesuai dengan kategori yang telah ditetapkan.

TABEL I  
PENENTUAN KATEGORI

Kategori Penelitian	Satuan
Luas Area TMB	Ha.
Luas Area (TM)	Ha.
Luas Area (TR)	Ha.
Produksi	Ton.
Produktivitas	Kg/Ha.

Tahapan dari proses segmentasi data perkebunan adalah tahap analisis. Setelah mengolah menggunakan teknik Clustering *K-Means*, selanjutnya dilakukan analisis yang lebih detail untuk merinci bagian-bagian perkebunan kelapa sawit di berbagai wilayah di Indonesia berdasarkan pada persamaan karakteristik luas area/wilayah, data hasil produksi, dan produktivitas.

TABEL II  
DATA LUAS AREA DAN HASIL PRODUKSI TAHUN 2022

PROVINSI	LUAS AREA TBM (Ha)	LUAS AREA TM (Ha)	LUAS AREA TR (Ha)	PRODUKSI (Ton)	PRODUKTIVITAS (Kg/Ha)
ACEH	97.261Ha	360.302Ha	30.440Ha	1.134.606 Ton	3.149 KG/HA
SUMUT	119.461Ha	1.1176.501Ha	29.117 Ha	5.776.781 Ton	4.910 KG/HA
SUMBAR	41.451Ha	337.604Ha	14.254 Ha	1.312.253 Ton	3.887 KG/HA
RIAU	330.605Ha	2.441.069Ha	82.268 Ha	9.984.315 Ton	4.090 KG/HA
KEPRI	2.079Ha	5.140Ha	178 Ha	20.020 Ton	3.895 KG/HA
JAMBI	192.179Ha	840.481Ha	41.939 Ha	3.022.565 Ton	3.596 KG/HA
SUMSEL	141.012Ha	1.022.994Ha	33.958 Ha	4.267.023 Ton	4.171 KG/HA
KEP. BABEL	34.190Ha	204.379Ha	1.244 Ha	843.047 Ton	4.125 KG/HA
BENGKULU	52.991Ha	268.190Ha	4.069 Ha	1.063.404 Ton	3.965 KG/HA
LAMPUNG	18.601Ha	172.710Ha	5.001 Ha	384.948 Ton	2.229 KG/HA
DKI JAKARTA	0	0	0	0	0
JAWA BARAT	599	12.865Ha	0	33.093 Ton	2.572 KG/HA
BANTEN	1.587	14.990Ha	2.666 Ha	27.423 Ton	1.829 KG/HA
D.I.YOGYAKARTA	0	0	0	0	0
JAWA TIMUR	0	0	0	0	0
BALI	0	0	0	0	0
NTB	0	0	0	0	0
NTT	0	0	0	0	0
KALBAR	316.531Ha	1.666.521Ha	56.152 Ha	5.471.407 Ton	3.283 KG/HA
KALTENG	216.123Ha	1.754.028Ha	48.509 Ha	7.685.407 Ton	4.382 KG/HA
KALSEL	67.255Ha	410.288Ha	19.917 Ha	1.561.147 Ton	3.805 KG/HA
KALTIM	205.300Ha	1.088.044Ha	20.262 Ha	3.823.221 Ton	3.514 KG/HA
KALTARA	22.897Ha	130.395Ha	4.385 Ha	301.607 Ton	2.513 KG/HA
SULUT	0	0	0	0	0
GORONTALO	1.781Ha	11.517Ha	0	4.975 Ton	432 KG/HA
SULTENG	26.255Ha	113.890Ha	5.727 Ha	371.717 Ton	3.264 KG/HA
SULSEL	9.559Ha	26.830Ha	8.347 Ha	100.317 Ton	3.739 KG/HA
SULBAR	38.225Ha	100.732Ha	17.223 Ha	348.015 Ton	3.455 KG/HA
SULTRA	11.030Ha	93.695Ha	5.576 Ha	76.300 Ton	814 KG/HA
GORONTALO	1.781Ha	11.517Ha	0	4.975 Ton	432 KG/HA
SULTENG	26.255Ha	113.890Ha	5.727 Ha	371.717 Ton	3.264 KG/HA
SULSEL	9.559Ha	26.830Ha	8.347 Ha	100.317 Ton	3.739 KG/HA
SULBAR	38.225Ha	100.732Ha	17.223 Ha	348.015 Ton	3.455 KG/HA
SULTRA	11.030Ha	93.695Ha	5.576 Ha	76.300 Ton	814 KG/HA
MALUKU	794Ha	9.799Ha	259 Ha	19.145 Ton	1.954 KG/HA
MALUKU UTARA	5.541Ha	0	0	0	0

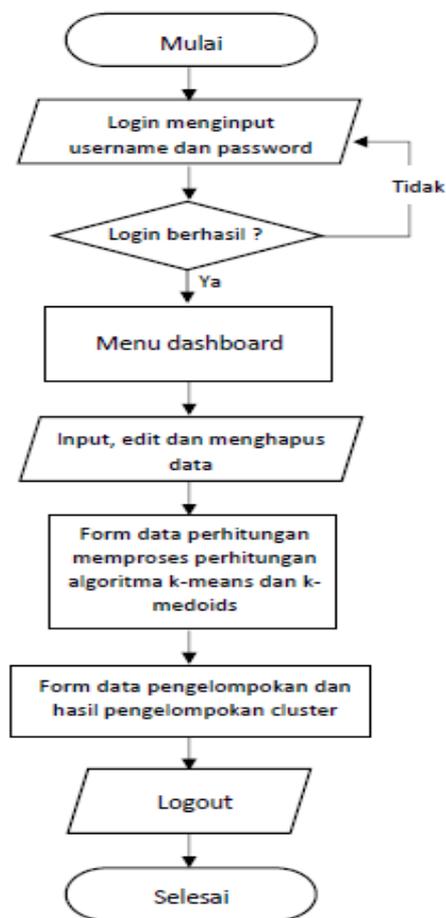
PAPUA	28.674Ha	111.855Ha	19.191 Ha	557.559 Ton	4.985 KG/HA
PAPUA BARAT	5.126Ha	45.892Ha	0	106.413Ton	2.319 KG/HA

### B. Desain Sistem

Pada tahap ini melakukan desain sistem pengelompokan potensi produksi kelapa sawit berbasis web. Padatahap ini membuat desain spesifikasi kebutuhan dari analisis dan desain sistem. Di tahap ini akan menghasilkan Flowchart, Usecase dan diagram Activity yang merupakan pemodelan data dari tahap analisa. Tahap ini juga menghasilkan desain tampilan sistem, struktur data, dan arsitektur lunak.

#### 1) Flowchart

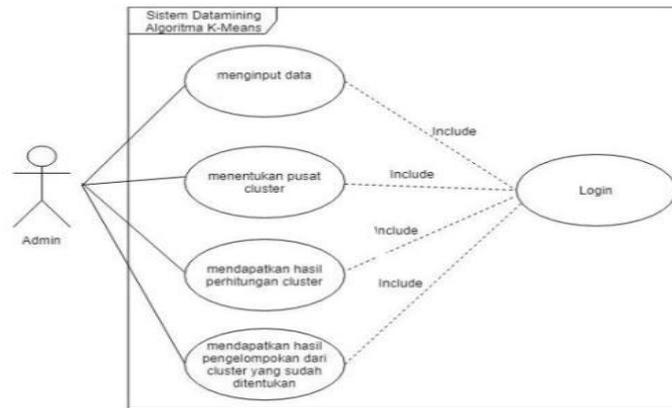
Flowchart merupakan gambaran umum dari sebuah pemodelan aplikasi yang akan dibuat. Berikut *Flowchart* yang dibuat.



Gambar 3. Flowchart

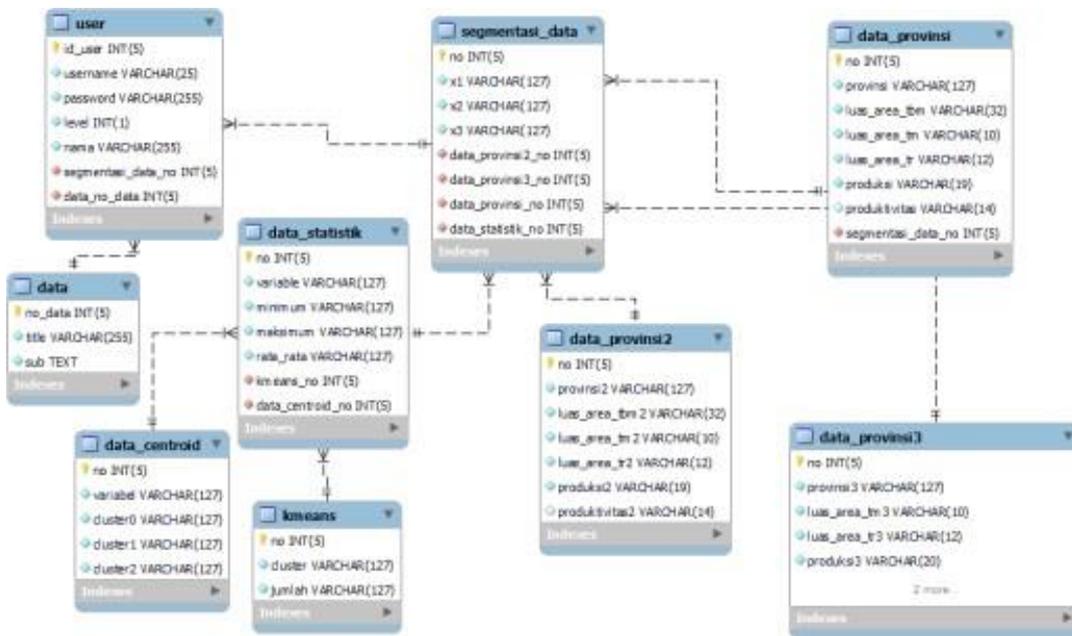
#### 2) Use Case

Merupakan pemodelan fungsional pada aplikasi yang tujuannya untuk mendeskripsikan kebutuhan sistem yang sudah dianalisis sebelumnya. Terlihat *Use Case* di Gambar 3.



Gambar 4. Use case diagram

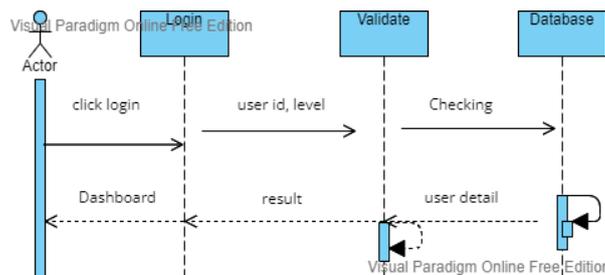
Pada gambar di atas menjelaskan yang dapat mengakses fungsi tersebut ialah admin, fungsi tersebut seperti yang dipanah untuk Include masuk ke dalam sistem yaitu login. setelah login admin dapat akses fungsi lainnya.



Gambar 5. Database desain

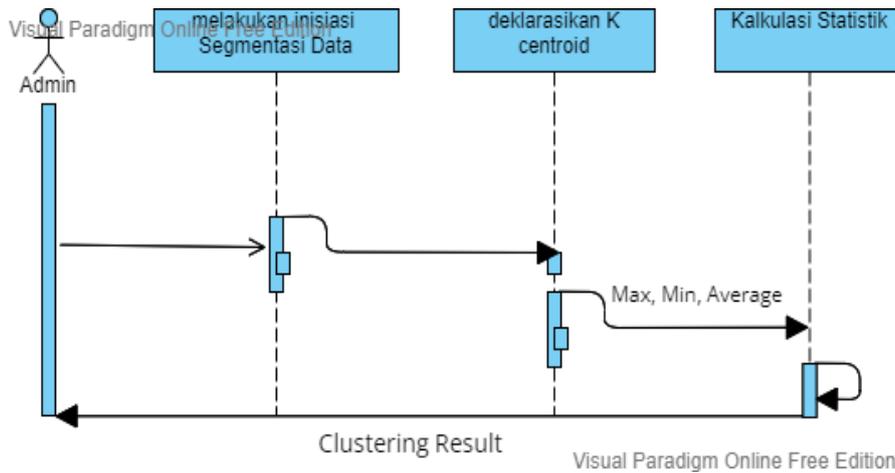
Gambar di atas adalah gambar desain database yang digunakan dalam sistem yang dibuat. Selanjutnya Gambar 5 ditampilkan diagram *database* untuk pembuatan sistem. *Database* merupakan kumpulan berbagai data yang dihimpun menjadi satu kesatuan untuk mempermudah proses pemodelan dalam suatu sistem.

3) *Diagram Activity:*

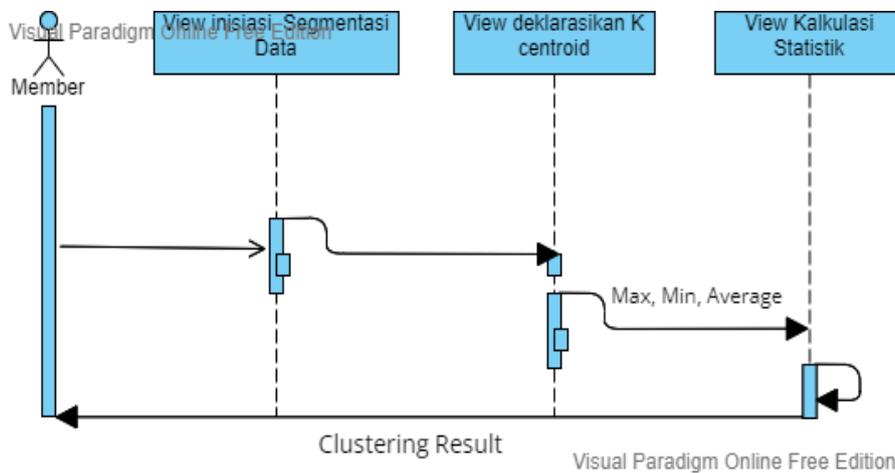


Gambar 6. Sequence diagram login

Selanjutnya Gambar 6 yaitu *sequence diagram login*. Gambar tersebut menunjukkan proses dari *login* hingga perhimpunan data yang masuk sampai pada menu sebelumnya.

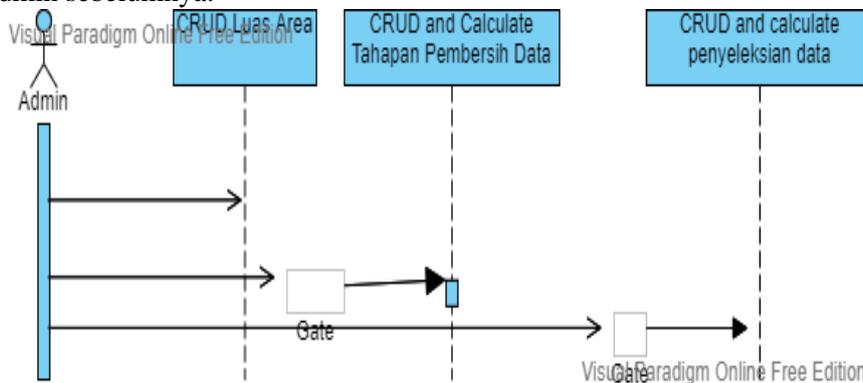


Gambar 7. Diagram admin clustering



Gambar 8. Diagram member clustering

Setelah *database* terbentuk, maka diketahui alur admin dan *member* ketika mengakses sistem yang dibuat oleh peneliti, admin melakukan inisiasi segmentasi data untuk merubah data-data lama ke yang baru. Dilanjut dengan Gambar 7, yang mana aktivitas dari *member* yang mengakses sistem akan diperlihatkan dengan beberapa data yang dimasukkan oleh admin sebelumnya.



Gambar 9. Penghimpunan data

Pada tahap penghimpunan data inilah admin dapat mengelompokkan data/mengimpun data.

**C. Pengkodean**

3) *Algoritma K-Means*: Salah satu algoritma pengklasteran yang masuk dalam kelompok *Unsupervised Learning*

adalah *K-Means* yang dalam penggunaannya untuk pembagian data menjadi beberapa sekte sistem partisi. Disetiap kluster terdapat Centroid (titik Pusat) yang menggambarkan kluster tersebut [14]. Berikut langkah-langkahnya [15]:

- Menentukan jumlah  $k$  (*cluster*) pada *dataset*
- Menentukan (*centroid*) secara *random*
- rumus *Euclidean Distance* untuk hitung jarak yang dekat dengan centroid berikut:

$$d = \sqrt{[x_i - s_i]^2 + [y_i - t_i]^2}$$

Keterangan:

$d$  = *Euclidean Distance*

$i$  = Banyak objek

$x, y$  = Titik koordinat objek  $s, t$  =

Titik koordinat *centroid*

- Rata-rata dalam sebuah kluster merupakan pusat kluster. Menghitung pusat kluster yang baru dengan anggota kluster yang sekarang. Dihitung menggunakan rumus:

$$V_{ij} = \frac{\sum_{k=1}^{N} X_{kj} N_{ik}}{N_{ij}}$$

Ket:

$V_{ij}$  = *Centroid (titik pusat)* rata-rata kluster ke- $i$  untuk variable ke- $j$

$N_{ij}$  = Jumlah anggota kluster ke- $i$ ,  $k$  =

Indeks dari kluster

$j$  = Indeks dari variable  $X$

$kj$  = Nilai data ke- $k$  variable ke- $j$  untuk *cluster* tersebut

- Kembali melakukan langkah ke-3 dan langkah ke-4 dan literasi mencapai *centroid* yang nilainya optimal.
- 4) *Algoritma K-Medoids*: Algoritma ini merupakan satu dari sekian banyak metode partisi kluster pembatas, karena memakai objek yang paling terpusat dikluster menjadi pusat kluster dari nilai rata-rata objek dalam sebuah kluster [16]. Langkah-langkahnya sebagai berikut di bawah ini:
- Inisialisasi *centroid cluster* sebanyak jumlah *cluster*.
  - Pada setiap data dialokasikan ke kluster yang paling dekat dengan menggunakan persamaan ukuran jarak *Euclidean Distance* dengan persamaan (1).
  - Memilih acak sebuah objek pada masing-masing kluster sebagai calon *Medoids* baru.
  - Menghitung jarak pada setiap objek yang berada pada masing-masing kluster dengan calon *Medoids* baru.
  - Menghitung total simpangan ( $s$ ) dengan nilai total *distance* baru hingga total *distance* lama. Jika  $S < 0$ , maka tukar objek dengan data cluster untuk membentuk sekumpulan  $k$  objek baru sebagai *Medoids*.
  - Mengulangi step by step dari 3-5 hingga tidak terjadi perubahan *Medoids*, sehingga didapat kluster beserta masing-masing anggota [17].

### III. HASIL DAN PEMBAHASAN

#### A. Pengolahan Data & Hasil Perhitungan

*Data set* tiap provinsi yang ditanami tumbuhan kelapa sawit terdapat pada Tabel I. Data tersebut selanjutnya dilakukan pembersihan data. Pembersihan data berfungsi agar lebih efisien dan efektif. Data yang meliputi provinsi tidak memiliki nilai dibersihkan dan menampilkan data terbaru yang ditujukan pada Tabel III. Data tersebut mengalami perubahan dari 34 data provinsi menjadi 26 *dataset*.

TABEL III  
TAHAPAN PEMBERSIHAN DATA

PROVINSI	LUAS AREAL TBM (HA)	LUAS AREAL TM (HA)	LUAS AREAL TR (HA)	PRODUKSI (TON)	PRODUKTIVITAS (KG/HA)
ACEH	97.261Ha	360.302Ha	30.440Ha	1.134.606 Ton	3.149 KG/HA
SUMUT	119.461Ha	1.1176.501Ha	29.117 Ha	5.776.781 Ton	4.910 KG/HA
SUMBAR	41.451Ha	337.604Ha	14.254 Ha	1.312.253 Ton	3.887 KG/HA
RIAU	330.605Ha	2.441.069Ha	82.268 Ha	9.984.315 Ton	4.090 KG/HA
KEPRI	2.079Ha	5.140Ha	178 Ha	20.020 Ton	3.895 KG/HA

JAMBI	192.179Ha	840.481Ha	41.939 Ha	3.022.565 Ton	3.596 KG/HA
SUMSEL	141.012Ha	1.022.994Ha	33.958 Ha	4.267.023 Ton	4.171 KG/HA
KEP. BABEL	34.190Ha	204.379Ha	1.244 Ha	843.047 Ton	4.125 KG/HA
BENGKULU	52.991Ha	268.190Ha	4.069 Ha	1.063.404 Ton	3.965 KG/HA
LAMPUNG	18.601Ha	172.710Ha	5.001 Ha	384.948 Ton	2.229 KG/HA
JAWA BARAT	599	12.865Ha	0	33.093 Ton	2.572 KG/HA
BANTEN	1.587	14.990Ha	2.666 Ha	27.423 Ton	1.829 KG/HA
KALBAR	316.531Ha	1.666.521Ha	56.152 Ha	5.471.407 Ton	3.283 KG/HA
KALTENG	216.123Ha	1.754.028Ha	48.509 Ha	7.685.407 Ton	4.382 KG/HA
KALSEL	67.255Ha	410.288Ha	19.917 Ha	1.561.147 Ton	3.805 KG/HA
KALTIM	205.300Ha	1.088.044Ha	20.262 Ha	3.823.221 Ton	3.514 KG/HA
KALTARA	22.897Ha	130.395Ha	4.385 Ha	301.607 Ton	2.513 KG/HA
GORONTALO	1.781Ha	11.517Ha	0	4.975 Ton	432 KG/HA
SULTENG	26.255Ha	113.890Ha	5.727 Ha	371.717 Ton	3.264 KG/HA
SULSEL	9.559Ha	26.830Ha	8.347 Ha	100.317 Ton	3.739 KG/HA
SULBAR	38.225Ha	100.732Ha	17.223 Ha	348.015 Ton	3.455 KG/HA
SULTRA	11.030Ha	93.695Ha	5.576 Ha	76.300 Ton	814 KG/HA
MALUKU	794Ha	9.799Ha	259 Ha	19.145 Ton	1.954 KG/HA
MALUKU UTARA	5.541Ha	0	0	0	0
PAPUA	28.674Ha	111.855Ha	19.191 Ha	557.559 Ton	4.985 KG/HA
PAPUA BARAT	5.126Ha	45.892Ha	0	106.413Ton	2.319 KG/HA

Untuk proses segmentasi maka data perkebunan dipilih. Hasil pemilihan data yang terdapat pada Tabel III menunjukkan bahwa data luas tanaman dewasa (TM), luas tanaman rusak (TR), produktivitas, dan produktivitas digunakan dalam proses transformasi data. Wilayah yang belum ditumbuhi (TBM) kelapa sawit tidak digunakan. Sedangkan data TR (tanaman rusak) akan tetap digunakan. Karena tetap menghasilkan walau tidak optimal.

Bukti bahwa ada perkebunan kelapa sawit di Maluku yang tidak produktif disingkirkan dan dari 26 dataset menjadi 25 dataset. Perpindahan data dengan memilih atribut. Di ketahui pada Tabel III data luas tanaman dewasa (TM) dan luas tanaman rusak (TR) masih dalam format terpisah. Keduanya digabungkan nilai data untuk menghasilkan data baru. Hasil dari perpindahan data berupa kumpulan 25 *dataset* yang diolah menggunakan algoritma *K-Means*. Untuk menyederhanakan proses setiap variabel diberikan atribut khusus.

TABEL IV  
PENYELEKSIAN DATA

PROVINSI	LUAS AREAL TM (HA)	LUAS AREAL TR (HA)	PRODUKSI (TON)	PRODUKTIVITAS (KG/HA)
ACEH	360.302Ha	30.440Ha	1.134.606 Ton	3.149 KG/HA
SUMUT	1.1176.501Ha	29.117 Ha	5.776.781 Ton	4.910 KG/HA
SUMBAR	337.604Ha	14.254 Ha	1.312.253 Ton	3.887 KG/HA
RIAU	2.441.069Ha	82.268 Ha	9.984.315 Ton	4.090 KG/HA
KEPRI	5.140Ha	178 Ha	20.020 Ton	3.895 KG/HA
JAMBI	840.481Ha	41.939 Ha	3.022.565 Ton	3.596 KG/HA
SUMSEL	1.022.994Ha	33.958 Ha	4.267.023 Ton	4.171 KG/HA
KEP. BABEL	204.379Ha	1.244 Ha	843.047 Ton	4.125 KG/HA
BENGKULU	268.190Ha	4.069 Ha	1.063.404 Ton	3.965 KG/HA
LAMPUNG	172.710Ha	5.001 Ha	384.948 Ton	2.229 KG/HA
JAWA BARAT	12.865Ha	*	33.093 Ton	2.572 KG/HA
BANTEN	14.990Ha	2.666 Ha	27.423 Ton	1.829 KG/HA
KALBAR	1.666.521Ha	56.152 Ha	5.471.407 Ton	3.283 KG/HA
KALTENG	1.754.028Ha	48.509 Ha	7.685.407 Ton	4.382 KG/HA
KALSEL	410.288Ha	19.917 Ha	1.561.147 Ton	3.805 KG/HA
KALTIM	1.088.044Ha	20.262 Ha	3.823.221 Ton	3.514 KG/HA
KALTARA	130.395Ha	4.385 Ha	301.607 Ton	2.513 KG/HA
GORONTALO	11.517Ha	*	4.975 Ton	432 KG/HA
SULTENG	113.890Ha	5.727 Ha	371.717 Ton	3.264 KG/HA

SULSEL	26.830Ha	8.347 Ha	100.317 Ton	3.739 KG/HA
SULBAR	100.732Ha	17.223 Ha	348.015 Ton	3.455 KG/HA
SULTRA	93.695Ha	5.576 Ha	76.300 Ton	814 KG/HA
MALUKU	9.799Ha	259 Ha	19.145 Ton	1.954 KG/HA
PAPUA	111.855Ha	19.191 Ha	557.559 Ton	4.985 KG/HA
PAPUA BARAT	45.892Ha	*	106.413Ton	2.319 KG/HA

Untuk melakukan segmentasi data hasil transformasi diproses dengan data *mining* dengan metode *K- Means* yang membagi kebun kelapa sawit menjadi beberapa segmen berdasarkan tiga variable yaitu X1= Luas, X2= Produksi, X3= Produktivitas. Segmentasi dilakukan dengan menggunakan *software Rapid Miner*. Hal pertama yang dapat dilakukan adalah menjalankan desain *K-Means* dengan memilih algoritma *K-Means* dan *power import* data yang terdapat pada Tabel IV.

TABEL V  
SEGMENTASI DATA

Id	(X1)	(X2)	(X3)
01	390.742Ha	1.134.606 Ton	3.149 KG/HA
02	1.205.618Ha	5.776.781 Ton	4.910 KG/HA
03	351.858Ha	1.312.253 Ton	3.887 KG/HA
04	2.523.337Ha	9.984.315 Ton	4.090 KG/HA
05	5.318Ha	20.020 Ton	3.895 KG/HA
06	882.420Ha	3.022.565 Ton	3.596 KG/HA
07	1.056.952Ha	4.267.023 Ton	4.171 KG/HA
08	205.623Ha	843.047 Ton	4.125 KG/HA
09	272.259Ha	1.063.404 Ton	3.965 KG/HA
010	177.711Ha	384.948 Ton	2.229 KG/HA
011	12.865Ha	33.093 Ton	2.572 KG/HA
012	17.656Ha	27.423 Ton	1.829 KG/HA
013	1.722.673Ha	5.471.407 Ton	3.283 KG/HA
014	1.802.537Ha	7.685.407 Ton	4.382 KG/HA
015	430.205Ha	1.561.147 Ton	3.805 KG/HA
016	1.108.306Ha	3.823.221 Ton	3.514 KG/HA
017	134.780Ha	301.607 Ton	2.513 KG/HA
018	11.517Ha	4.975 Ton	432 KG/HA
019	119.671Ha	371.717 Ton	3.264 KG/HA
020	35.117Ha	100.317 Ton	3.739 KG/HA
021	117.955Ha	348.015 Ton	3.455 KG/HA
022	99.271Ha	76.300 Ton	814 KG/HA
023	10.058Ha	19.145 Ton	1.954 KG/HA
024	131.046Ha	557.559 Ton	4.985 KG/HA
025	45.892Ha	106.413Ton	2.319 KG/HA

Nilai titik berat pusat ditentukan secara acak dan otomatis menggunakan jenis pengukuran yang digunakan berupa pengukuran numerik jarak *Euclidean*. Nilai tersebut berdasarkan luas lantai X1, X2, X3 yang di tujukan pada Tabel VI.

TABEL VI  
DATA CENTROID PADA CENTROID TABEL

Variabel	Cluster 1	Cluster 2	Cluster 3 (Centroid)
X1	142.753,7	1.195.193,8	2.162.937
X2	459.221,6	4.472.199,4	8.835.042,5
X3	2.929,5	3.894,8	4.236,0

Hasil variable X1, X2, X3 dengan statistik minimum, maksimum dan rata-rata ditunjukkan pada Tabel VII.

TABEL VII  
STATISTIK NILAI SETIAP VARIABEL

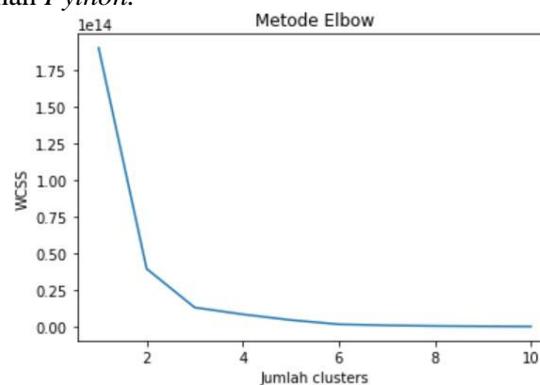
Variabel	Minimum	Maksimum	Rata-rata
X1	5.318	2.523.337	514.855,72
X2	4.975	9.984.315	1.931.882,84
X3	432	4.985	3.227,08

1) *Clustering K-Means*: Representasi *K-Means Clustering* dapat dilihat berdasarkan hasil penyaringan data menggunakan *K-Means* yang ditinjau dari ketiga variable X1, X2, X3 untuk perkebunan kelapa sawit mendapat 3 kluster yaitu *Cluster 1*, *Cluster 2*, *Centroid* . Di bagian Tabel VIII merupakan hasil segmentasi yang telah menumakan hasil dari pengolahan data.

TABEL VIII HASIL  
SEGMENTASI

PROVINSI	Cluster1	Cluster2	Centroid
ACEH	*		
SUMUT		*	
SUMBAR	*		
RIAU			*
KEPRI	*		
JAMBI		*	
SUMSEL		*	
KEP. BABEL	*		
BENGKULU	*		
LAMPUNG	*		
JAWA BARAT	*		
BANTEN	*		
KALBAR		*	
KALTENG			*
KALSEL	*		
KALTIM		*	
KALTARA	*		
GORONTALO	*		
SULTENG	*		
SULSEL	*		
SULBAR	*		
SULTRA	*		
MALUKU	*		
PAPUA	*		
PAPUA BARAT	*		

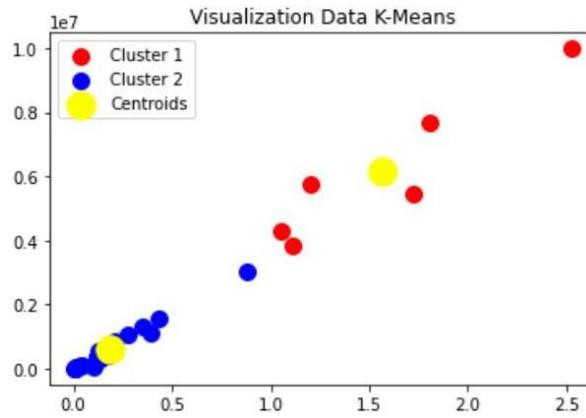
*Dataset* yang diambil dari variabel X1, X2, dan X3 menjadi tolak ukur dalam memvisualisasi data menggunakan bahasa pemrograman *Python*.



Gambar 10. Metode elbow

Metode Elbow merupakan metode yang berada didalam Python untuk membentuk struktur. dalam menentukan angka cluster K yang tepat bisa menggunakan metode tersebut Terkait jumlah cluster dapat disesuaikan oleh

seberapa banyak dan jenis data.



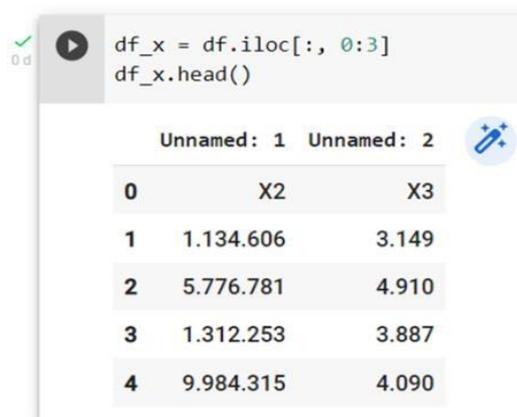
Gambar 11. Data visualization cluster k-means

Gambar 10 menunjukkan hasil dari data *visualization* menggunakan *python* setelah variabel telah ditentukan, kemudian dapat diproyeksikan dengan model algoritma *K-Means*. Dapat dilihat bahwa cluster 1 merupakan provinsi terendah dalam potensi produksi kelapa sawit, cluster 2 sedang, centroid tertinggi dengan 2 provinsi berwarna kuning.

2) *Clustering K-Medoids*: Hasilnya tiga model segmentasi terlihat di tabel 9. Hasil yang pertama (*cluster1*) 18 item/wilayah, (*cluster 2*) terdiri dari 5 item/wilayah, (*centroid*) terdiri dari 2 item/wilayah.

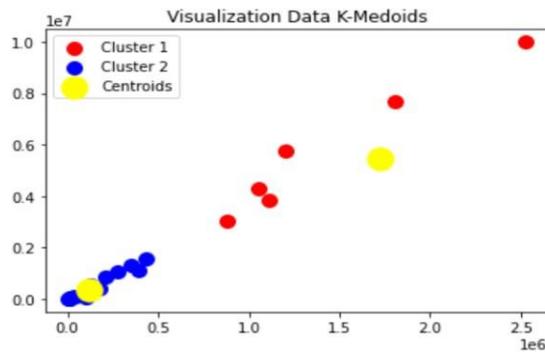
TABEL IX  
HASIL SEGMENTASI DENGAN K-MEDOIDS CLUSTERING

Cluster	Jumlah
Cluster1	18 Item
Cluster 2	5 Item
Centroid	2 Item



Gambar 12. Menentukan variabel yang diklasterkan

Gambar 12 merupakan Tahap penentuan variabel yang diklasterkan. Tahapan dalam visualisasi data ialah menyiapkan lingkungan pengembangan aplikasi antara lain meng-*install* berbagai *libraries* seperti *pandas*, *numpy*, *matplotlib*, dan *scikit-learn*, hingga mengonfigurasi *dataframe* menjadi model algoritma *K-Means* & *K-Medoids*.



Gambar 13. Data visualization cluster k-medoids

Gambar 13 merupakan grafik data visualization menggunakan *python* dengan menggunakan algoritma *K-Medoids*. Dari data-data yang telah di klusterkan dengan penggabungan kedua metode dapat diasumsikan sebagai berikut :

*Asumsi 1*TABEL X  
ASUMSI 1 (CLUSTER 1)

<b>Id</b>	<b>Rentang Luas</b>	<b>Rentang Produksi</b>	<b>Rentang Produktivitas</b>
1, 3, 5, 8, 9, 10, 11, 12, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25	5.318 Ha – 430.205 Ha	4.975 Ton – 1.561.147 Ton	432 Kg/Ha – 4.985 Kg/Ha

*Cluster 1* (Rendah) terdapat 18 wilayah yang memiliki potensi rendah terdiri 18 provinsi Pada *Cluster 0* merupakan luas area terletak pada rentang 5.318 Ha sampai 430.205 Ha. Produksi pada rentang antara 4.975Ton sampai 1.561.147 Ton dan produktivitas terletak pada rentang 432 Kg/Ha hingga 4.985 Kg/Ha.

*Asumsi 2*TABEL XI  
ASUMSI 2 (CLUSTER 2)

<b>Id</b>	<b>Rentang Luas</b>	<b>Rentang Produksi</b>	<b>Rentang Produktivitas</b>
2, 6, 7, 13, 16	882.420 Ha – 1.772.673 Ha	3.002.565 Ton – 5.776.781 Ton	3.283 Kg/Ha – 4.910 Kg/Ha

*Cluster2* (sedang) terdapat 5 wilayah yang memiliki potensi sedang terdiri dari 5 provinsi. Pada *Cluster 2* diketahui luas area terletak pada rentang 882.420 Ha sampai 1.772.673 Ha. Sedangkan produksi pada rentang antara 3.002.565 Ton sampai 5.776.781 Ton dan produktivitas teletak pada rentang 3.283 Kg/Hahingga 4.910 Kg/Ha.

*Asumsi 3*TABEL XII  
ASUMSI 3 (CENTROID)

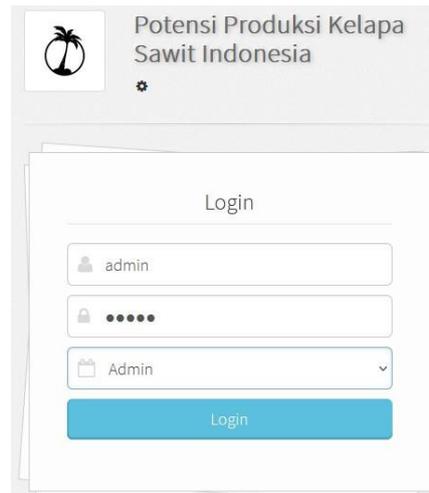
<b>Id</b>	<b>Rentang Luas</b>	<b>Rentang Produksi</b>	<b>Rentang Produktivitas</b>
4, 14	1.802.537 Ha – 2.523.337 Ha	7.802.537 Ton – 9.984.315 Ton	4.090 Kg/Ha – 4.382 Kg/Ha

*Centroid* (tinggi) terdapat 2 wilayah yang memiliki potensi tinggi terdiri dari 2 provinsi. Pada *Centroid* diketahui luas area terletak pada rentang1.802.537 Ha sampai 2.523.337 Ha. Sedangkan produksi pada rentang antara

7.802.537 Ton sampai 9.984.315 Ton dan produktivitas terletak pada rentang 4.090 Kg/Ha hingga 4.382 Kg/Ha.

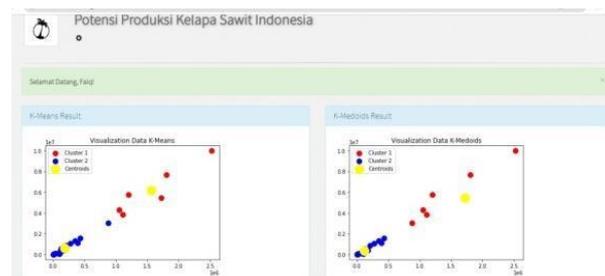
### B. Implementasi

Tahap-tahap dalam mengimplementasikan sesuatu tahap perancangan sistem berdasarkan metode *waterfall*, pengimplementasian disusun dalam bentuk sebuah program aplikasi. Sistem tersebut terlihat pada beberapa gambar dibawah ini. Tampilan tersebut berjalan dengan sempurna, akan tetapi masih memerlukan penelitian lebih lanjut. Hasil dari data yang telah diolah maka diaplikasikan kedalam sistem yang telah dibuat. Tampilan tersebut terletak pada Gambar 13.



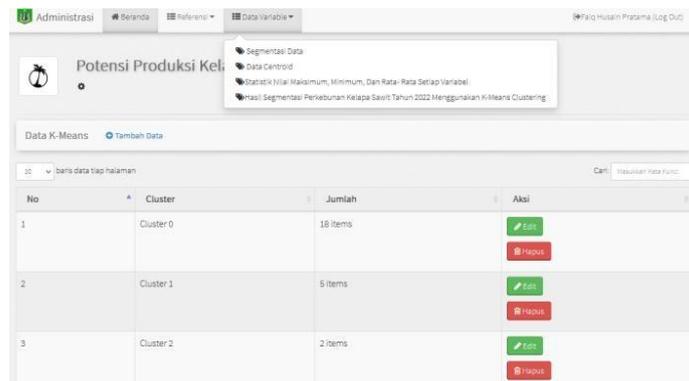
Gambar 14. Tampilan login

Di dalam Gambar 14. Terdapat tampilan menu *login* pada sistem yang telah dibuat, *login* tersebut berfungsi sebagai tahap awal ketika *user* atau admin ingin mengakses sistem tersebut, pada tahap awal dilakukan registrasi terlebih dahulu dengan memasukkan *Username* dan *Password* setelah itu juga terdapat pilihan *user* atau admin. Setelah melakukan registrasi, *user* atau admin dapat mengakses sistem tersebut yang di tampilkan seperti Gambar 15 dibawah ini. Tampilan tersebut adalah tampilan menu admin.

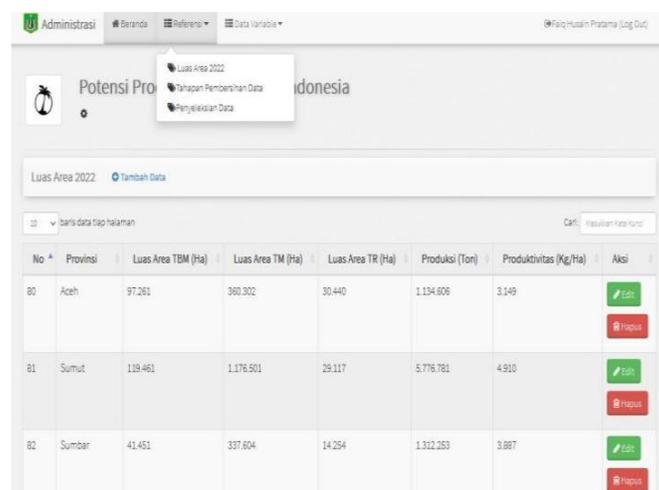


Gambar 15. Tampilan dashboard

Tampilan dashboard pada sistem yang dibuat berisi tentang informasi grafik kluster dengan metode K-Means dan K-Medoids dimana menu admin dimana ada menu beranda, referensi dan data variabel. Pada setiap tools nya memiliki fungsi masing-masing yaitu untuk memasukkan dan mengeluarkan data yang telah diolah sehingga secara otomatis dengan perhitungan algoritma, sistem yang diakses oleh user sesuai dengan hasil yang diinginkan oleh admin penginput data.



Gambar 16. Menu segmentasi data



Gambar 17. Menu luas area tanaman

Selain terdapat menu tampilan utama pada admin. Juga terdapat beberapa menu yaitu tampilan login admin, referensi, dan data variable. Yang terekam pada tampilan Gambar 15. Gambar tersebut ditampilkan menu segmentasidata. Data pada menu diatas dapat dirubah dengan perhitungan algoritma. Untuk mengganti mengetahui menambah dan mengurangi data luas area tanaman terdapat pada menu referensi yaitu dibagian Luas area, luas area ini menunjukkan luas area yang ditanami oleh kelapa sawit, seperti data tabel 2 diatas. Tidak jauh berbeda dengan ketika *loginUser*. Ketika *login user* tampilannya samanamun tidak *user* tidak mampu untuk mengubah, mengolah, menambahatau mengurangi data yang telah dimasukkan oleh admin. Pengujian sistem aplikasi potensi produksi kelapa sawit ini menggunakan 10.000 *user* pada *database* dan tidak mengalami kendala sedikitpun.

Pengujian sistem selanjutnya menggunakan GTMetrix untuk mengukur kinerja situs dan Loadimpact untuk mengukur ketahanan sistem yang dibuat.



Gambar 18. Pengujian GTMetrix

Sesuai dengan tampilan Gambar 19 terlihat bahwa sistem yang dibuat memiliki performa 98% dan structur 94% dengan grade A.



Gambar 19. Pengujian loadimpact

Jika dengan pengujian Loadimpact sesuai dengan Gambar 19 di atas bahwa sistem yang telah dibuat memiliki hasil rata-rata item yang telah diuji 912ms dan 1.137 request dengan error 0 dan 16 per second.

K-Medoid berfungsi untuk penentuan cluster sesuai dengan data variable yang paling tinggi, sedang dan rendah pada Luas areal kelapa sawit. Algoritma K-means berfungsi untuk mengelompokkan rata rata luas tanah, produksi dan produktifitas, dengan pembagian potensi rendah, sedang dan tinggi.

#### IV. KESIMPULAN

Dalam tahap akhir sebuah penelitian dan pengimplementasian sebuah sistem pengelompokan potensi produksi kelapa sawit dengan kombinasi 2 algoritma yaitu K-Medoids dan K-Means guna mengelompokkan kelapa sawit berdasar kepada luas area, Produksi dan produktivitas telah mendapatkan hasil dengan data variable yang paling rendah (Cluster 1) 18 wilayah, sedang (Cluster 2) 5 Wilayah, dan tinggi (Cluster 3) / Centroids 2 wilayah pada potensi hasil Luas areal, produksi, dan produktivitas kelapa sawit. Algoritma K-means berfungsi untuk mengelompokkan rata rata luas tanah 514.885,72 Ha, produksi 1.931.882,84 Ton dan produktifitas 3.227,08 Kg/Ha, dengan pembagian potensi rendah, sedang dan tinggi. Kombinasi dari kedua algoritma berfungsi sangat baik karena masing masing memiliki peran tersendiri yang sesuai dengan kebutuhan penelitian. Dari penggabungan 2 metode K-Medoid dan K-Means mendapatkan hasil ketiga klaster bahwa hasil produksi kelapa sawit yang memiliki potensi rendah 72% sedang 20%, tinggi 2%. Segmentasi ini disebabkan oleh kesamaan karakteristik perkebunan berdasarkan kesamaan dari luas, produksi, dan produktivitas. Yang memiliki potensi tertinggi kelapa sawit ada 2 provinsi yaitu Kalimantan Barat dan Riau.

Hal ini menjadikan kedua algoritma cocok untuk pengklasteran potensi produksi kelapa sawit karena keakurasiannya sesuai dengan perhitungan manual yang telah dijabarkan pada bagian pengolahan data.

#### DAFTAR PUSTAKA

- [1] K.Vajen, M.Jesper, F.Pag, and U.Jordan, "Annual Industrial and Commercial Heat Load Profiles: Modeling Based On K-Means Clustering and Regression Analysis," 2021.
- [2] Siti Hajar, A.Windarto, A.Wanto, E.Irawan, A.A.Novany, "Penerapan K-Means Clustering Pada Ekspor Minyak Kelapa Sawit Menurut Negara Tujuan," SAINTEK, P.314, 2020
- [3] Trisna Yuniarti And Dahliyah Hayati, "Segmentasi Perkebunan Kelapa Sawit Dengan Data Mining Teknik K-Means Clustering Berdasarkan Luas Areal, Produksi Dan Produktivitas," Vol. 2, No. 2, P56-64, 2021.
- [4] E.T.E.Handayani, Handoko, dan Fauziah, "Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Dengan Menggunakan Metode K-Means Clustering," J.Ilm. Tekno Dan Rekayasa, Vol.25, No.1, Pp. 76-88, 2020.
- [5] F.T.Kesuma, Feryanti dan S.P.Tamba, "Penerapan Data Mining Untuk Menentukan Penjualan Sparepart Toyota Dengan Metode K-Means Clustering," (Jusikom Prima), Vol. 2, 2021.
- [6] U.B.Mulia, "Jumlah Ritel," Vol. XI, no. 1, pp.32-44
- [7] F.Pag, M.Jesper, K.Vajen, And U.Jordan, "Annual Industrial And Comercial Head Loadprofiles: Modeling Based On K-Means Clustering And Regression Analysis," Energy Convers Manag. X, Vol. 10, No. March, 2021,
- [8] H.S.Tambunan, I.S.Danamik, D.F.Pasaribu, E.Irawan, And K. Kunci, "Memanfaatkan Algoritma K-Means Dalam Memetakan Potensi Hasil Produksi Kelapa Sawit Ptpn Iv Marihat," Bios Teknologi, Informasi Dan Rekayasa Komputer, 2021.
- [9] P.Deepak And A.Jurek-Loughrey, "Multi View Clustering," Pp. 27-53.
- [10] Y. Liu, X. Du, And S. Ma. :Innovative Study On Clustering Center And Distance Measurement Of K-Means Algoritma Based On User Data Of Jd Mall," No. 01. Springer Us, 2021.
- [11] Rizaldi, "Penerapan Waterfall Dalam Membangun Sistem Informasi Pengolahan Data," Vol. No. 1, Pp. 71-80, 2021
- [12] Eri Mardiani, Nur Ramansyah, Wahyudi, Et.Al, "Kumpulan Latihan PHP," Jakarta: PT Elex Media Komputindo, 2021.
- [13] Yuniarti And Hayati "Inventory: Industrial Vocational E-Journal On Agroindustry," Volume. 2, 2021.

- [14] L.Y.Hutabarat, H. D.Tampubolon, D.Gultom, F.R.Iلمي H Zer, And D. Hartama, “Penerapan Algoritma K-Means Untuk Mengetahui Tingkat Tindak Kejahatan Daerah Pematangsiantar,” J. Teknol. Inf, Volume. 4, 2020.
- [15] Ai Rohmah, Falentino Sembiring, Adhitia Erfina, “Implementasi Algoritma K-Means Clustering Analysis Untuk Menentukan Hambatan Pembelajaran Daring,” SISMATIK Universitas Nusa Putra, 2021.
- [16] Minami, et.al, “Klasterisasi Penyakit Menggunakan Algoritma K-Medoids Pada Dinas Kesehatan Kabupaten Agam Provinsi Sumatera Barat,” Jurnal Nasional Pendidikan Teknik Informatika Volume 10, Nomor 3, Desember 2021.
- [17] Nur Arief1, Irfan Sudahri Damanik, Eka Irawan, “Penerapan Algoritma K- Medoids Dalam Mengelompokkan Tingkat Kasus Kejahatan di Setiap Provinsi,” STIKOM Tunas Bangsa, Vol 2, No 3, pp. 111-116, Desember 2021. <https://djournal.com/klik>
- [18] A Triayudi, I Fitri, WO Widyarto – Proceedings of the 2nd International Conference on Electronics, Biomedical Engineering, and Health Informatics. 2022
- [19] Triayudi A, Sumiati S, Dwiyatno S, Karyaningsih D, Susilawati (2021) Measure the effectiveness of information systems with the naïve bayes classifier method. IAES Int J Artif Intell 10.