

SKRIPSI

***CLUSTERING POTENSI KOPI MENGGUNAKAN METODE K-
MEANS BERBASIS GEOGRAPHIC INFORMATION SYSTEM
(GIS)***

***CLUSTERING COFFEE POTENTIAL USING THE K-MEANS
METHOD BASED ON GEOGRAPHIC INFORMATION SYSTEM
(GIS)***

Diajukan untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer



ABDULLAH AZZAM

D0221304

PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS SULAWESI BARAT

MAJENE

2025

HALAMAN PERSETUJUAN

SKRIPSI

CLUSTERING POTENSI KOPI MENGGUNAKAN METODE K-MEANS BERBASIS GEOGRAPHIC INFORMATION SYSTEM (GIS)

Telah disiapkan dan disusun oleh

**ABDULLAH AZZAM
D0221304**

Telah dipertahankan di depan Tim Penguji

Pada tanggal 18 September 2025

Susunan Tim Penguji

Pembimbing I



Arnita Irianti, S.Si., M.Si
NIP: 198708062018032001

Penguji I



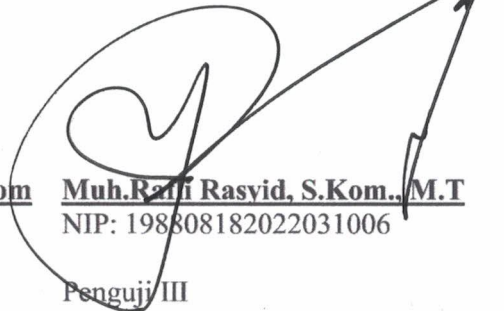
Diny Anggriani Adnas, S.ST., M.T
NIP: 199301232025062007

Pembimbing II



Nuralamsah Zulkarnaim, S.Kom., M.Kom
NIP: 198910142019031013

Penguji II



Muh. Rafi Rasyid, S.Kom., M.T
NIP: 198808182022031006

Penguji III



Muh. Fahmi Rustan, S.Kom., M.T
NIP: 199112272019031010

PERNYATAAN ORISINALITAS

Saya dengan ini menyatakan bahwa sepanjang pengetahuan saya, di dalam naskah skripsi ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu perguruan tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis disitasi dalam naskah ini dan disebutkan dalam daftar referensi.

Apabila ternyata dalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur plagiasi, saya bersedia skripsi ini digugurkan dan gelar akademik yang telah saya peroleh (sarjana) dibatalkan, serta diproses sesuai dengan perundang-undangan yang berlaku (UU No.20 Tahun 2003, Pasal 25 ayat 2 Pasal 70)

Majene, 10 September 2025



Abdullah Azzam
NIM.D0221304

ABSTRAK

Abdullah Azzam. *Clustering* Potensi Kopi Menggunakan Metode *K-Means* Berbasis *Geographic Information System* (Gis). (Dibimbing Oleh **Arnita Irianti** dan **Nuralamsah Zulkarnaim**).

Kabupaten Mamasa di Sulawesi Barat memiliki potensi besar dalam pengembangan komoditas kopi, namun masih minim sistem pendukung berbasis data spasial untuk analisis potensi lahan. Penelitian ini bertujuan membangun sistem yang mengintegrasikan algoritma *K-Means* manual dengan visualisasi peta interaktif berbasis *Geographic Information System* (GIS) untuk mengelompokkan potensi kopi secara spasial. Data dianalisis berdasarkan atribut numerik seperti luas lahan, jumlah petani, produksi, dan produktivitas, dengan proses normalisasi menggunakan MinMaxScaler. Evaluasi kluster menggunakan *Davies-Bouldin Index* (DBI) menunjukkan bahwa hasil *clustering* memiliki tingkat kompaksi dan separasi yang baik. Sistem berhasil membentuk tiga kluster deskriptif (“Rendah”, “Sedang”, “Tinggi”) yang divisualisasikan dalam peta interaktif berbasis Streamlit. Uji coba menunjukkan semua fitur, mulai dari unggah data, *clustering*, hingga visualisasi GIS, berjalan sesuai ekspektasi. Hasil penelitian ini memberikan kontribusi nyata dalam pengambilan keputusan strategis di sektor perkebunan kopi.

Kata kunci: *K-Means*, Kopi Mamasa, GIS, *Clustering*

ABSTRACT

Abdullah Azzam. *Clustering Coffee Potential Using The K-Means Method Based On Geographic Information System (Gis). (Supervised by Arnita Irianti and Nuralamsah Zulkarnaim).*

Mamasa Regency in West Sulawesi holds significant potential for coffee commodity development, yet lacks spatial data-based support systems for land potential analysis. This study aims to develop a system that integrates a manual K-Means algorithm with interactive map visualization based on Geographic Information System (GIS) to spatially cluster coffee potential. The data were analyzed using numerical attributes such as land area, number of farmers, production, and productivity, with normalization performed using MinMaxScaler. Cluster evaluation using the Davies-Bouldin Index (DBI) indicates that the clustering results demonstrate good compactness and separation. The system successfully formed three descriptive clusters (“Low”, “Medium”, “High”), which are visualized through an interactive map built with Streamlit. Testing showed that all features from data upload, clustering, to GIS visualization functioned as expected. The findings of this research provide tangible contributions to strategic decision-making in the coffee plantation sector.

Keywords: *K-Means, Mamasa Coffee, GIS, Clustering*

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kopi merupakan komoditas strategis dalam sektor perkebunan Indonesia, dengan Sumatera Selatan, Aceh, dan Lampung sebagai provinsi penghasil utama. Sulawesi Barat, khususnya Kabupaten Mamasa, juga memiliki potensi besar melalui Kopi Mamasa yang dikenal akan kualitas dan cita rasanya yang khas. Daerah ini memiliki peluang ekspor yang menjanjikan dan peran penting dalam memperkuat industri kopi nasional. Komoditas ini merupakan salah satu andalan Kabupaten Mamasa dengan potensi ekspor yang besar. Dengan akses yang lebih baik terhadap teknologi dan informasi terkini, petani kopi di Mamasa dapat meningkatkan produksi serta memaksimalkan nilai ekonomi dari komoditas bernilai tinggi ini. Selain itu, permintaan terhadap komoditas kopi terus meningkat setiap tahun, menjadikannya salah satu bahan baku perkebunan yang paling penting di pasar internasional (Chairuddin & Abdullah, 2023).

Keterbatasan investasi di sektor pertanian menjadi salah satu hambatan utama dalam mendorong pertumbuhan ekonomi dan pengembangan bidang perkebunan (Adha & Andy, 2022). Investasi terhambat karena kurangnya data yang komprehensif terkait kualitas dan proyeksi hasil lahan. Penata Kelola Penanaman Modal Ahli Madya DPMPTSP Sulbar, Satriawan Hasan Sulur, menekankan pentingnya penyusunan peta potensi yang akurat dan komprehensif. Dalam pernyataannya di situs resmi DPMPTSP Sulawesi Barat, beliau mengatakan “dengan peta potensi yang baik, kita dapat memberikan gambaran yang jelas kepada para investor tentang berbagai peluang yang ada di sektor perkebunan kopi. Ini adalah langkah strategis untuk mengoptimalkan potensi sumber daya alam kita dalam mendorong pertumbuhan ekonomi” (DPMPTSP Sulbar, 2024).

Kemajuan teknologi informasi telah menyediakan berbagai solusi untuk mengatasi masalah ini. Salah satu algoritma yang dapat diterapkan untuk mengelompokkan lahan kopi adalah algoritma *K-Means*, yang bekerja dengan menetapkan jumlah kelompok awal serta menentukan nilai pusat kluster (*centroid*) awalnya (Alif et al., 2022). Algoritma *K-Means* adalah salah satu algoritma

clustering dalam *data mining* yang menggunakan pendekatan *unsupervised learning*. Algoritma ini bekerja dengan mempelajari data secara mandiri untuk menentukan kelas atau klaster masing-masing.

Penelitian yang dilakukan oleh Kaligis dan Yulianto (2022), mengelompokkan kinerja pegawai dengan metode *Clustering K-Means*, K-Medoids, dan X-Means. Berdasarkan hasil analisis menggunakan *davies bouldin index* (DBI), nilai yang diperoleh untuk masing-masing algoritma adalah: *K-Means* sebesar -0.377, K-Medoids sebesar -0.930, dan X-Means sebesar -0.497. Oleh karena itu, algoritma terbaik untuk pengelompokan data kinerja pegawai dalam penelitian tersebut adalah *K-Means*, karena memiliki nilai DBI terkecil. Penelitian yang berjudul “Perbandingan Algoritma *K-Means* dan K-Medoids untuk Pemetaan Hasil Produksi Buah-Buahan” telah dilakukan oleh Prasetyaningrum dan Susanti (2023), Studi ini membandingkan metode *K-Means* dan K-Medoids. Hasil pengujian menunjukkan bahwa algoritma *K-Means* menghasilkan nilai DBI yang lebih kecil, seperti terlihat pada pengujian keempat dengan $k=6$ yang mencapai nilai 0,296. Sementara itu, algoritma K-Medoids menghasilkan klaster optimal pada pengujian pertama dengan $k=3$, dengan nilai DBI sebesar 0,507.

Integrasi *K-Means* dengan Sistem Informasi Geografis (GIS) sangat diperlukan agar hasil analisis menjadi lebih optimal. GIS berperan penting dalam mengelola data spasial yang rumit, terutama mengingat bahwa peta dan data statistik seringkali cepat usang, sehingga sulit menemukan layanan yang menyediakan data dan informasi yang benar-benar akurat (Falah et al., 2023). Selain itu, penggunaan GIS yang terhubung dengan platform web mendukung proses pengambilan keputusan berbasis data menjadi lebih cepat dan lebih tepat.

Penelitian dengan judul "***Clustering Potensi Kopi Menggunakan Metode K-Means Berbasis Geographic Information System (GIS)***" dipilih oleh penulis berdasarkan permasalahan yang diuraikan sebelumnya. Tujuan penelitian ini untuk mengimplementasikan algoritma *K-Means* dalam pengelompokan lahan di setiap kecamatan kabupaten Mamasa, yang akan divisualisasikan dengan *geographic information system* (GIS).

1.2 Rumusan Masalah

Berdasarkan latar belakang yang ada, masalah yang akan dibahas dalam penelitian ini adalah:

1. Bagaimana cara mengkluster lahan kopi di Mamasa menggunakan metode *K-Means*?
2. Bagaimana cara memvisualisasikan metode *K-Means* ke dalam sistem berbasis GIS untuk pemetaan lahan kopi di kabupaten tersebut?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Mengembangkan sistem pemetaan lahan kopi di Kabupaten Mamasa dengan metode *K-Means* berbasis GIS.
2. Menyediakan informasi berbasis data untuk mendukung pengambilan keputusan dan meningkatkan investasi di sektor perkebunan kopi.

1.4 Manfaat Penelitian

Penelitian ini diharapkan memberikan beberapa manfaat, antara lain:

1. Mengembangkan metode analisis spasial berbasis GIS dengan algoritma *K-Means* untuk sektor perkebunan kopi, khususnya di Kabupaten Mamasa. Hasil penelitian ini dapat memperkaya literatur tentang penerapan *Clustering* dalam pemetaan potensi lahan dan menjadi rujukan bagi studi analisis spasial dan optimalisasi lahan berbasis data.
2. Membantu pemerintah daerah, petani, dan investor mengidentifikasi lahan potensial untuk budidaya kopi secara akurat dan berbasis data. Sistem GIS yang dikembangkan mendukung pengambilan keputusan terkait perencanaan produksi dan investasi, sehingga pengembangan perkebunan kopi dapat lebih efektif dan efisien.

1.5 Batasan Masalah

Agar tetap menjaga fokus penelitian, sejumlah batasan masalah ditentukan dalam penelitian ini adalah sebagai berikut:

1. Penelitian hanya mencakup penerapan metode *K-Means* dalam pengklasteran lahan kopi dengan integrasi ke dalam sistem GIS berbasis web tanpa fitur analisis tambahan.
2. Penelitian menggunakan algoritma *K-Means* standar tanpa modifikasi atau pengembangan lebih lanjut,

BAB II

KAJIAN PUSTAKA

2.1 Landasan Teori

2.1.1 Potensi Kopi Mamasa

Kabupaten Mamasa adalah salah satu daerah di Sulawesi Barat yang memiliki potensi besar untuk pengembangan industri kopi. Mamasa yang terletak di daerah pegunungan memiliki iklim dan kondisi tanah yang ideal untuk pertumbuhan tanaman kopi, khususnya jenis Arabika, yang menjadi primadona di pasar kopi internasional. Petani kopi di Kabupaten Mamasa sangat menyadari bahwa kopi dapat dijadikan tanaman konservasi sekaligus memiliki nilai ekonomis yang tinggi (Pratama et al., 2024).

Karakteristik geografis Kabupaten Mamasa sangat membantu pertumbuhan kopi, terutama dengan variasi ketinggian yang ideal untuk Arabika dan Robusta. Pengembangan kopi di Sulawesi barat mencakup 23.419 hektar, dengan sentra utama di Kabupaten Mamasa seluas 19.117 hektar, menurut data dari Litbang Pemprov Sulbar. Produksi kopi berkualitas tinggi yang menarik perhatian pembeli domestik dan internasional dimungkinkan oleh ketinggian lahan di atas 1.000 mdpl dan kondisi iklim yang mendukung (Badan Pusat Statistik Provinsi Sulawesi Barat, 2024). Petani lokal menghadapi dua tantangan utama yang menghambat produktivitas, yaitu: pengelolaan lahan yang tidak efektif dan kurangnya informasi mengenai potensi lahan. Diharapkan, dengan upaya pemerintah dalam meningkatkan dukungan teknis dan penyebaran informasi, masalah ini dapat teratasi dan produktivitas kopi Mamasa dapat meningkat secara signifikan.

2.1.2 Normalisasi

Untuk mendapatkan hasil yang optimal, tahapan pra proses *data mining* sangat penting. Preprocessing meningkatkan efisiensi penerapan algoritma. Normalisasi adalah proses preprocessing data. Proses normalisasi dilakukan untuk menstandarkan nilai-nilai dalam dataset agar berada dalam skala yang konsisten (Muhima et al., 2023). Nilai fitur atau variabel dalam kumpulan data dapat diubah dengan menggunakan skala Min-Max, yang biasanya berkisar antara 0 dan 1, tetapi juga dapat disesuaikan sesuai kebutuhan.

Rumus Normalisasi Min Max sebagai berikut:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2.1)$$

Keterangan:

- X' = nilai yang dinormalisasi
- X = nilai asli yang ingin dinormalisasi
- X_{min} = nilai minimum dari dataset
- X_{max} = nilai maksimum dari dataset

2.1.3 K-Means

K-Means adalah salah satu teknik pengelompokan data yang paling umum dan mudah digunakan. Algoritma ini membagi data ke dalam sejumlah kluster yang ditetapkan pada awal (k). Setiap data dihubungkan ke pusat kluster, berdasarkan jarak terdekat, yang biasanya merupakan jarak geometris.

Algoritma *K-Means* adalah sebagai berikut:

1. Tentukan berapa banyak kluster (k) yang ingin Anda miliki.
2. Pilih pusat kluster (*centroid*) awal yang berbeda.
3. Data dimasukkan ke kluster dengan jarak terdekat ke *centroid* berdasarkan hasil perhitungan *Euclidean distance*.

$$d = \sqrt{\sum_{i=1}^n (x_i - c_i)^2} \quad (2.2)$$

Keterangan:

- d = jarak titik data dengan *centroid*
 - x_i = nilai titik data
 - c_i = nilai titik *centroid*
4. Posisi *centroid* yang diperbarui didasarkan pada posisi data rata-rata dalam masing-masing kluster.

$$\mu_k = \frac{1}{N_k} \sum_{j=1}^{N_k} x_j \quad (2.3)$$

Keterangan:

- μ_k = *centroid* baru untuk kelompok ke- k
- N_k = jumlah data dalam kelompok ke- k
- x_j = data ke- j dalam kelompok ke- k

5. Sampai posisi *centroid* tidak berubah atau ada perubahan kecil, ulangi langkah 3 dan 4.

Metode *K-Means* dikenal karena kemudahan implementasinya dan kecepatan dalam memproses data, terutama pada dataset besar, seperti yang diungkapkan dalam penelitian Al Masykur et al., (2023) yang menunjukkan bahwa *K-Means* dapat memberikan hasil *Clustering* yang baik ketika diterapkan pada pemetaan pengelompokan lahan produksi tandan buah segar. Dengan penggunaan *K-Means* di Mamasa, diharapkan dapat memberikan gambaran yang lebih jelas tentang wilayah potensi produksi kopi.

2.1.4 *Clustering*

Clustering adalah salah satu teknik dalam *data mining* yang digunakan untuk mengelompokkan data berdasarkan karakteristik atau pola tertentu. Data dengan karakteristik yang mirip akan dikelompokkan ke dalam satu klaster, sedangkan data yang berbeda akan dimasukkan ke klaster lain. Metode *Clustering* banyak diterapkan di berbagai bidang karena kemampuannya dalam mengidentifikasi pola data yang besar dan kompleks. Pada penelitian ini, *Clustering* diterapkan untuk mengelompokkan lahan berdasarkan beberapa parameter guna menemukan lahan dengan target investasi paling layak untuk budidaya kopi.

Clustering biasanya digunakan dalam perkebunan untuk memetakan lahan berdasarkan produktivitas, tipe tanah, dan faktor iklim. Oleh karena itu, perkembangan produktivitas hasil bumi dari sektor perkebunan, dalam hal ini perkebunan perlu dipantau supaya bisa menjadi perhatian instansi terkait dalam mengoptimalkan daerah penghasil tersebut (Pratiwi et al., 2022). Pengelompokan menjadi metode yang relevan untuk pengelompokan lahan kopi Mamasa, yang bervariasi berdasarkan produktivitas dan luas lahan.

2.1.5 *Davies Bouldin Index (DBI)*

Davies-Bouldin Index (DBI) adalah salah satu metrik evaluasi yang digunakan untuk menilai kualitas hasil *Clustering*. Metrik ini pertama kali diperkenalkan oleh David L. Davies dan Donald W. Bouldin pada tahun 1979 sebagai cara untuk mengevaluasi hasil pengelompokan data secara *unsupervised*. DBI bekerja dengan

mengukur tingkat kompaksi dan separasi antar klaster yang terbentuk dari algoritma *Clustering*. Pengukuran menggunakan *Davies-Bouldin Index* (DBI) menghasilkan nilai yang bersifat absolut, di mana nilai DBI terkecil menunjukkan kinerja yang lebih baik dalam hasil klasterisasi (Soepriyono & Triayudi, 2023). DBI dapat dirumuskan sebagai berikut:

$$DBI = \frac{1}{k} \sum_{i=1}^k \left(\frac{s_i + s_j}{d_{ij}} \right) \quad (2.4)$$

Keterangan:

k = jumlah klaster

s_i = rata-rata jarak antara titik data dalam klaster i dengan *centroid* i (kompaksi).

s_j = rata-rata jarak antara titik data dalam klaster j dengan *centroid* klaster j (kompaksi).

d_{ij} = Jarak antara *centroid* klaster i dan *centroid* klaster j (separasi).

Davies-Bouldin Index memiliki beberapa keunggulan yang membuatnya menjadi matrik evaluasi yang efektif. Salah satu kelebihan adalah tidak membutuhkan label data, sehingga sangat cocok untuk algoritma unsupervised seperti *K-Means*. Selain itu, DBI cukup mudah diterapkan karena perhitungannya langsung menggunakan hasil *Clustering*. Metrik ini juga fleksibel dan dapat digunakan untuk berbagai jenis dataset, termasuk data yang memiliki banyak dimensi, sehingga cocok untuk beragam kebutuhan analisis *Clustering*.

2.1.6 Geographic Information System (GIS)

Geographic Information System (GIS) atau lebih dikenal sistem informasi geografis (SIG) merupakan teknologi yang memungkinkan pengumpulan, penyimpanan, pengelolaan, dan analisis data geografis atau spasial. *Geographic Information System* (GIS) merupakan teknologi berbasis komputer yang dimanfaatkan untuk pemetaan dan analisis data spasial, dengan tujuan mengungkap kondisi serta kejadian di permukaan bumi (Meilani et al., 2025). SIG telah diimplementasikan di berbagai sektor, termasuk perencanaan tata ruang, pengelolaan sumber daya alam, dan pemetaan perkebunan.

Di sektor perkebunan SIG digunakan untuk memvisualisasikan sebaran lahan, menganalisis kesesuaian lahan untuk berbagai jenis komoditas, serta mengoptimalkan pengelolaan lahan agar lebih efisien. Pemanfaatan *Web GIS* bersama dengan algoritma *K-Means* memungkinkan pemetaan dan pengelompokan data spasial, yang berguna dalam mengenali wilayah yang memiliki potensi investasi serta memberikan dukungan bagi pemerintah dalam menyusun kebijakan yang lebih efektif (Nurdin et al., 2023).

2.1.7 Python

Python, yang diciptakan oleh Guido van Rossum pada tahun 1991, merupakan bahasa pemrograman tingkat tinggi yang dikenal luas karena fleksibilitas dan kemudahan penggunaannya. Sintaksisnya yang sederhana dan mudah dipahami memungkinkan pengembang menulis kode yang lebih sederhana dan mudah dipahami. Filosofi desain Python dikenal dengan sebutan “*The Zen of Python*”, yang mengutamakan kejelasan, kesederhanaan, dan konsistensi dalam penulisan kode (Adawiyah Ritonga & Yahfizham Yahfizham, 2023). Python mendukung berbagai paradigma pemrograman, seperti berorientasi objek, fungsional, dan imperatif, sehingga cocok untuk berbagai aplikasi, termasuk analisis data, komputasi ilmiah, dan pengembangan aplikasi web interaktif.

Keunggulan utama Python terletak pada ekosistem pustakanya yang kaya, yang mendukung berbagai kebutuhan pengembangan. Pustaka seperti NumPy memungkinkan komputasi numerik yang efisien, Pandas menyediakan alat untuk analisis dan pengelolaan data, dan Streamlit mempermudah pembuatan aplikasi web interaktif dengan antarmuka yang intuitif. Dalam konteks penelitian ini, Streamlit digunakan untuk mengembangkan sistem berbasis web yang mengintegrasikan proses *Clustering K-Means* dengan visualisasi Sistem Informasi Geografis (GIS).

2.1.8 Website

Website adalah media informasi yang tersedia di internet. *Website* terdiri dari kumpulan halaman yang biasanya terorganisir dalam satu domain atau subdomain dan berada dalam jaringan *World Wide Web* (WWW) di internet. Situs web adalah

kumpulan halaman yang dapat diakses melalui peramban internet. Halaman-halaman ini umumnya ditulis dalam bahasa HTML (*HyperText Markup Language*) dan dapat menyertakan elemen multimedia seperti gambar, video, dan animasi. Beberapa jenis situs web termasuk situs statis, dinamis, e-commerce, blog, dan portal berita.

Dengan kemajuan teknologi, website telah menjadi alat penting dalam berbagai aspek kehidupan, termasuk pendidikan, bisnis, dan hiburan. Penyebaran informasi melalui *website* sangat cepat dan mencakup area yang luas tanpa batasan jarak dan waktu. *Website* menjadi alat penting untuk menyampaikan informasi dan mempromosikan produk agar dikenal oleh masyarakat luas (Nurlailah & Nova Wardani, 2023). Perkembangan teknologi web telah memungkinkan desain responsif, sehingga *website* kini dapat diakses dengan mudah melalui berbagai perangkat seperti *smartphone*, tablet, maupun komputer desktop. Hal ini tentunya meningkatkan aksesibilitas informasi bagi pengguna di seluruh dunia.

2.1.9 Metode *Black Box*

Pengujian perangkat lunak merupakan salah satu aspek dalam metode perancangan perangkat lunak. Pengujian perangkat lunak melibatkan evaluasi sistem atau komponennya dengan tujuan untuk menemukan kesalahan dan memastikan bahwa perangkat lunak berfungsi sesuai dengan spesifikasi yang ditentukan (Rahman Abdillah et al., 2024). Pengujian *black box* merupakan teknik evaluasi perangkat lunak yang menitikberatkan pada fungsionalitas sistem tanpa memperhatikan struktur internal atau kode sumbernya. Pendekatan ini digunakan untuk memastikan bahwa perangkat lunak berjalan sesuai dengan spesifikasi yang telah ditetapkan dengan mengamati hubungan antara masukan yang diberikan dan keluaran yang dihasilkan. Fokus utama dalam metode ini adalah menilai validitas data masukan, kesesuaian hasil keluaran, serta kinerja fitur-fitur utama yang terdapat dalam sistem yang diuji.

2.2 Penelitian Sebelumnya

Terkait dengan topik penelitian, berikut beberapa penelitian sebelumnya yang digunakan:

Tabel 2. 1 Penelitian Terdahulu

NO	Judul dan Peneliti	Hasil	Keterikatan
1.	Penerapan Metode <i>K-Means Clustering</i> Pada Pemetaan Lahan Kopi Di Kabupaten Malang. Dimas Alif Fajar Fadhillah, Ahmad Faisol, Nurlaily Vendyansyah (2022)	Penelitian menunjukkan bahwa sebagian besar wilayah di Kabupaten Malang memiliki produksi tanaman kopi yang rendah. Dengan metode <i>K-Means Clustering</i> , data terbagi menjadi tiga klaster, yaitu: klaster rendah mendominasi dengan jumlah data tertinggi setiap tahun, klaster sedang memiliki jumlah data yang lebih sedikit, dan klaster tinggi mencakup paling sedikit wilayah, menunjukkan distribusi produksi kopi yang tidak merata.	Menggunakan algoritma K Means sebagai metode <i>Clustering</i> dan visualisasi hasil pemetaan menggunakan GIS.
2.	Pengaruh Tenaga Kerja dan Investasi di Sektor Pertanian terhadap Pertumbuhan Ekonomi Sektor Pertanian di Indonesia. Ajeng Afrillia Adha, Puti Andiny (2022)	Penelitian menggunakan data time series 2000–2020 dengan regresi linier berganda menunjukkan bahwa tenaga kerja dan investasi pertanian memiliki pengaruh negatif dan tidak signifikan terhadap pertumbuhan ekonomi sektor pertanian di Indonesia pada tingkat signifikansi 5%.	Menunjukkan bahwa sektor pertanian memerlukan pendekatan baru seperti pemetaan spasial dan analisis <i>Clustering</i> untuk mendukung kebijakan yang lebih tepat sasaran, seperti dalam

			penelitian ini.
3.	Perbandingan Algoritma <i>K-Means</i> Dan K-Medoids Untuk Pemetaan Hasil Produksi Buah-Buahan. Eka Prasetyaningrum, Puji Susanti (2023).	Penelitian menunjukkan bahwa algoritma <i>K-Means</i> lebih optimal dibandingkan K-Medoids dalam mengelompokkan hasil produksi buah-buahan di kabupaten Kotawaringin Timur, dengan nilai DBI terbaik sebesar 0,296 untuk 6 klaster.	Menggunakan metode Algoritma <i>K-Means</i> untuk analisis produktivitas buah.
4.	Analisa Perbandingan Algoritma <i>K-Means</i> , K Medoids, Dan X-Means Untuk Pengelompokkan Kinerja Pegawai. Gideon Bartolomeus Kaligis, Sri Yulianto (2022).	Berdasarkan hasil uji Karena algoritma <i>K-Means</i> memiliki nilai DBI terkecil maka menjadi yang terbaik dalam menentukan data kinerja pegawai ke dalam kriteria klaster yang dibutuhkan. Klaster yang dihasilkan oleh metode <i>K-Means</i> memiliki kriteria nilai tinggi (C1) dan kriteria nilai rendah (C0), masing-masing dengan 3 pegawai.	Menggunakan algoritma <i>K-Means</i> untuk mengelompokkan kinerja pegawai.
5.	Sistem Informasi Geografis Berbasis web Tentang Pemetaan Tempat Penjualan Oleh-Oleh Di Pekanbaru.	Penelitian ini menghasilkan sebuah sistem informasi geografis berbasis web yang dapat memudahkan wisatawan dalam mencari dan menemukan lokasi	Sistem Informasi geografis (GIS) sebagai visualisasi data.

	Ghazi Falah, Dicko Andrian, Teguh Alzikri, Debi Setiawan (2023)	penjualan oleh-oleh di Pekanbaru secara praktis dan efisien.	
6.	Mapping System Model and <i>Clustering</i> of Fishery Products using <i>K-Means</i> Algorithm with Web GIS Approach. Nurdin, Taufiq, Fajriana, Muhammad Zia Ulhaq (2023)	Sistem informasi berbasis WebGIS berhasil dibangun untuk mengelompokkan produk perikanan menjadi dua kluster: ikan unggulan dan biasa menggunakan algoritma <i>K-Means</i> .	Menggunakan algoritma <i>K-Means</i> dan Web GIS untuk klasifikasi spasial berbasis data atribut. Relevan dengan pendekatan pemetaan potensi kopi.
7.	An Improved <i>Clustering</i> Based on <i>K-Means</i> for Hotspots Data. Rani Rotul Muhima, Muchamad Kurniawan, Septiyawan Rosetya Wardhana, Anton Yudhana, Sunardi, Mitra Adhimukti (2023)	Pengembangan GAP <i>K-Means</i> untuk meningkatkan efisiensi dan kualitas klustering. Evaluasi dilakukan menggunakan DBI, Silhouette, SSE, dan F-measure.	Mendukung pentingnya normalisasi dan evaluasi <i>Clustering</i> menggunakan DBI, yang juga diterapkan dalam penelitian ini.
8.	Analisis Potensi Titik Rawan Kemacetan di Kecamatan Jebres	Penelitian ini menemukan 7 titik rawan kemacetan pada 5 ruas jalan di Kecamatan Jebres,	Relevan dalam penggunaan GIS sebagai alat visualisasi spasial

	Berbasis Sistem Informasi Geografis. Farah Meilani, Vidda Arlysia, Syahita Martha De Wela, Abel Surya Mahendra (2025)	dikategorikan dalam dua tingkat kerawanan: kuning (rawan) dan merah (sangat rawan). Penelitian menggunakan pendekatan deskriptif berbasis SIG untuk mengidentifikasi titik-titik tersebut dan mengevaluasi dampak kemacetan seperti pemborosan BBM, polusi udara, dan stress.	untuk pengambilan keputusan. Keterkaitan utamanya terletak pada pendekatan spasial dalam menganalisis fenomena wilayah tertentu, seperti juga dalam pemetaan potensi kopi.
9.	Mengidentifikasi Kelompok Sektor Perkebunan di Indonesia Berdasarkan Produktivitas Hasil Bumi. Hesti Pratiwi, Ari Purno Wahyu Wibowo (2022)	Hasil klasifikasi di dapat tiga klaster sektor perkebunan yang diantaranya adalah "great" yang berarti di atas target, "good" yang berarti belum mencapai target namun masih di atas rata-rata, dan "underperformed" yang berarti sektor perkebunan dalam klaster ini perlu perhatian khusus karena produktivitasnya di bawah performa.	Menggunakan algoritma <i>K-Means</i> untuk mengelompokkan data percobaan.
10.	Analisis Skalogram: Studi Kasus Pengembangan Kopi Mamasa Berbasis Kesesuaian Lahan dan Sarana Fasilitas Penunjang di	Hasil penelitian menunjukkan, lahan kopi di Kabupaten Mamasa umumnya tergolong marginal (S3) dengan kendala utama berupa temperatur, curah hujan, dan kemiringan lereng. Produksi dapat	Melakukan analisis untuk kepentingan investasi pada sektor perkebunan kopi Mamasa.

	Kabupaten Mamasa. Zulkarnain Chairuddin, Nuryahya Abdullah (2023).	ditingkatkan melalui intensifikasi lahan dengan teknologi dan agroforestri. Dukungan pemerintah sangat diperlukan untuk mendukung produksi dan kepastian pasar. Fasilitas pendukung kopi di daerah ini masih terbatas dan membutuhkan perhatian lebih.	
11.	Perbandingan Kinerja Algoritma <i>Clustering Data mining</i> Untuk Prediksi Harga Saham Pada Reksadana dengan <i>Davies Bouldin Index</i> . Gatot Soepriyono, Agung Triayudi (2023)	Penelitian ini menyimpulkan bahwa <i>K-Means</i> memiliki kinerja lebih baik dibandingkan K-Medoids, dengan nilai DBI yang lebih kecil, yang menunjukkan klaster yang lebih kompak dan terpisah.	Penelitian ini membuktikan bahwa algoritma <i>K-Means</i> efektif untuk mengelompokkan data lahan tanpa label kelas (unsupervised), seperti potensi lahan kopi, dan menggunakan Davies-Bouldin Index (DBI) untuk mengevaluasi kualitas hasil <i>Clustering</i> secara empiris.
12.	Studi Literatur Perbandingan Bahasa Pemrograman C ++ Dan	Penelitian ini membahas perbandingan antara bahasa pemrograman C++ dan Python, yang	Menggunakan bahasa pemrograman python untuk melakukan

	Bahasa Pemrograman Python Pada Algoritma Pemrograman. Adawiyah Ritonga, Yahfizham Yahfizham (2023).	masing-masing memiliki kelebihan. C++ lebih cocok untuk pengembangan perangkat lunak desktop namun lebih kompleks, sedangkan Python lebih mudah dipahami dan umum digunakan untuk pembuatan situs web. Tujuan utamanya adalah memahami perbedaan, fungsi, dan keunggulan kedua bahasa tersebut.	perhitungan dan olah data.
13.	Perancangan Website Sebagai Media Informasi Dan Promosi Oleh-Oleh Khas Kota Pagaram. Ela Nurlailah, Kiki Rizky Nova Wardani (2023)	Dari penelitian ini didapatkan bahwa sistem berbasis web online yang dibangun efektif dalam menyampaikan informasi kepada pengunjung mengenai toko oleh-oleh khas Pagaram, membantu penjual dalam mempromosikan toko agar dikenal luas, serta mendukung Dinas Pariwisata kota Pagaram dalam memperkenalkan aset-aset kota secara lebih optimal.	Menggunakan website untuk menampilkan data hasil analisis.
14.	Pengujian Perangkat Lunak Sistem Informasi Inventori pada Usaha Jasa	Penelitian ini menggunakan teknik pengujian metode black box testing dengan teknik pengujian	Pengujian perangkat lunak dengan metode <i>black box</i> yang fokus pada

	Pengiriman Paket. Rahman Abdillah, Rudi Hermawan, Wawan Hermawansyah, Ibnu Adkha, Heri Arifin (2024)	Equivalence Partitioning, yang berarti hanya perbagian fitur yang diuji secara random dan perlu diperbaiki atau dikembangkan lebih lanjut oleh pengembang (programmer).	validasi fungsi sistem berdasarkan masukan dan keluaran yang dihasilkan.
15.	Penerapan Metode <i>K-Means Clustering</i> untuk Pemetaan Pengelompokan Lahan Produksi Tandan Buah Segar. Abdussalam Al Masykur, Siska Kurnia Gusti, Suwanto Sanjaya, Febi Yanto, Fadhilah Syafria (2023)	Penelitian menunjukkan bahwa metode <i>K-Means Clustering</i> efektif untuk memetakan lahan produksi tandan buah segar dengan tingkat kecocokan 99,54% menggunakan RapidMiner. Dengan nilai <i>Davies Bouldin Index</i> 0,921, tiga klaster diidentifikasi: C1 (Produktivitas Sedang, 96 blok), C2 (Produktivitas Rendah, 41 blok), dan C3 (Produktivitas Tinggi, 79 blok). Metode ini berpotensi memberikan informasi berharga bagi perusahaan.	Penggunaan metode <i>K-Means</i> yang sama untuk analisis <i>Clustering</i>. Selain itu, pendekatan berbasis GIS dalam penelitian ini sejalan dengan pemanfaatan data geospasial yang dilakukan dalam penelitian sebelumnya.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini berhasil mengembangkan sistem analisis potensi kopi berbasis web yang mengintegrasikan algoritma *K-Means* dengan visualisasi *Geographic Information System* (GIS) menggunakan platform Streamlit. Sistem ini mengelompokkan potensi kopi di 17 kecamatan Kabupaten Mamasa berdasarkan atribut numerik (luas lahan, jumlah petani, produktivitas, jumlah produksi) dan memvisualisasikan hasilnya pada peta interaktif. Berdasarkan hasil implementasi, pengujian, dan analisis, berikut adalah kesimpulan:

1. Keberhasilan Pengembangan Sistem:

Sistem pemetaan potensi kopi berbasis web berhasil dikembangkan dengan mengintegrasikan algoritma *K-Means* manual dengan *Geographic Information System* (GIS). Arsitektur sistem mencakup empat komponen utama, yaitu antarmuka pengguna, modul pemrosesan data, modul visualisasi GIS, serta basis data. Integrasi ini memungkinkan proses unggah data, pra-pemrosesan menggunakan normalisasi Min-Max, klusterisasi dengan algoritma *K-Means*, penyimpanan hasil dalam basis data, hingga visualisasi spasial berbasis peta interaktif berjalan secara terpadu.

2. Efektivitas *Clustering K-Means*:

Algoritma *K-Means* mampu mengelompokkan potensi kopi menjadi tiga klaster, yaitu Rendah, Sedang, dan Tinggi, berdasarkan variabel luas lahan, jumlah petani, produksi, dan produktivitas. Pemilihan nilai $k = 3$ dipertimbangkan untuk memudahkan interpretasi serta mendukung kebutuhan analisis kebijakan investasi. Evaluasi menggunakan *Davies Bouldin Index* (DBI) menunjukkan hasil klaster yang cukup kompak dan terpisah, sehingga dapat dikategorikan optimal.

3. Kontribusi Visualisasi GIS:

Integrasi sistem dengan GIS memberikan kemudahan dalam memahami distribusi spasial potensi kopi di Kabupaten Mamasa. Peta interaktif yang menampilkan hasil klusterisasi memudahkan pengguna, baik pemerintah daerah, petani, maupun investor, dalam mengidentifikasi wilayah potensial.

Hal ini menunjukkan bahwa GIS mampu meningkatkan keterbacaan data dan mendukung pengambilan keputusan berbasis lokasi.

4. Manfaat Praktis Penelitian:

Sistem yang dibangun tidak hanya berfungsi sebagai alat analisis akademik, tetapi juga memiliki manfaat praktis dalam mendukung perencanaan pembangunan perkebunan kopi di Sulawesi Barat. Pemerintah daerah dapat menggunakan hasil klasterisasi untuk menentukan prioritas pengembangan wilayah, petani dapat memahami posisi lahannya dalam konteks produktivitas regional, dan investor memperoleh gambaran awal tentang daerah yang layak untuk investasi.

Sistem ini memenuhi tujuan penelitian, menyediakan alat analisis berbasis data untuk transformasi digital di sektor perkebunan kopi.

5.2 Saran

Berdasarkan hasil dan kekurangan sistem, berikut adalah rekomendasi untuk pengembangan lebih lanjut:

1. Pengembangan Algoritma: Membandingkan K-Means dengan algoritma lain seperti K-Medoids, DBSCAN, atau Fuzzy C-Means guna memperoleh hasil klasterisasi yang lebih akurat.
2. Penambahan Variabel Analisis: Memperluas indikator dengan memasukkan faktor iklim, infrastruktur, dan akses pasar agar hasil analisis lebih komprehensif.
3. Integrasi Sistem: Mengembangkan sistem ke arah Decision Support System (DSS) dengan fitur prediksi produktivitas, simulasi investasi, serta rekomendasi wilayah prioritas.

Saran-saran ini diharapkan meningkatkan performa, akurasi, dan kegunaan sistem untuk analisis potensi perkebunan kopi di wilayah lain.

DAFTAR PUSTAKA

- Adawiyah Ritonga, & Yahfizham Yahfizham. (2023). Studi Literatur Perbandingan Bahasa Pemrograman C++ dan Bahasa Pemrograman Python pada Algoritma Pemrograman. *Jurnal Teknik Informatika Dan Teknologi Informasi*, 3(3), 56–63. <https://doi.org/10.55606/jutiti.v3i3.2863>
- Adha, A. A., & Andy, P. (2022). PENGARUH TENAGA KERJA DAN INVESTASI DI SEKTOR PERTANIAN TERHADAP PERTUMBUHAN EKONOMI SEKTOR PERTANIAN DI INDONESIA. *SAMUKA (Jurnal Samudra Ekonomika)*, 6, 40–49.
- Al Masykur, A., Gusti, S. K., Sanjaya, S., Yanto, F., & Syafria, F. (2023). Penerapan Metode K-Means Clustering untuk Pemetaan Pengelompokan Lahan Produksi Tandan Buah Segar. *Jurnal Informatika*, 10(1). <https://doi.org/10.31294/inf.v10i1.15621>
- Alif, D., Fadhilah, F., Faisol, A., & Vendyansyah, N. (2022). PENERAPAN METODE K-MEANS CLUSTERING PADA PEMETAAN LAHAN KOPI DI KABUPATEN MALANG. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 6, Issue 1).
- Chairuddin, Z., & Abdullah, N. (2023). Analisis Skalogram: Studi Kasus Pengembangan Kopi Mamasa Berbasis Kesesuaian Lahan dan Sarana Fasilitas Penunjang di Kabupaten Mamasa. *Jurnal Ecosolum*, 12(1), 86–104. <https://doi.org/10.20956/ecosolum.v12i1.25663>
- Falah, G., Andrean, D., Alzikri, T., Setiawan, D., & Riau, U. M. (2023). *Sistem Informasi Geografis Berbasis web Tentang Pemetaan Tempat Penjualan Oleh-Oleh Di Pekanbaru* (Vol. 3, Issue 2). <http://maps.google.com>.
- Kaligis, G. B., & Yulianto, S. (2022). ANALISA PERBANDINGAN ALGORITMA K-MEANS, K-MEDOIDS, DAN X-MEANS UNTUK PENGELOMPOKAN KINERJA PEGAWAI. *Jurnal Penerapan Teknologi Informasi Dan Komunikasi*.
- Meilani, F., Arlysia, V., De Wela, S. M., & Mahendra, A. S. (2025). Analisis Potensi Titik Rawan Kemacetan di Kecamatan Jebres Berbasis Sistem Informasi Geografis. *ENVIRO: Journal of Tropical Environmental Research*, 27(1), 32–39. <https://doi.org/10.20961/enviro.v27i1.101206>
- Muhima, R. R., Kurniawan, M., Wardhana, S. R., Yudhana, A., Sunardi, & Adhimukti, M. (2023). An improved clustering based on K-means for hotspots data. *Indonesian Journal of Electrical Engineering and Computer Science*, 31(2), 1109–1117. <https://doi.org/10.11591/ijeecs.v31.i2.pp1109-1117>
- Nurdin, Taufiq, Fajriana, & Ulhaq, M. Z. (2023). Mapping System Model and Clustering of Fishery Products using K-Means Algorithm with Web GIS Approach. In *Original Research Paper International Journal of Intelligent Systems and Applications in Engineering IJISAE* (Vol. 2023, Issue 3). www.ijisae.org
- Nurlailah, E., & Nova Wardani, K. R. (2023). PERANCANGAN WEBSITE SEBAGAI MEDIA INFORMASI DAN PROMOSI OLEH-OLEH KHAS KOTA PAGARALAM. *JIPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 8(4), 1175–1185. <https://doi.org/10.29100/jipi.v8i4.4006>
- Parlika, R., Afifudin, M., Pradana, I. A., Dimas, Y., Wiratama, W., & Holis, M. N. (2022). STUDI LITERATUR EFISIENSI MODEL RAPID APPLICATION DEVELOPMENT DALAM PENGEMBANGAN PERANGKAT LUNAK (2014-

- 2022). *Positif : Jurnal Sistem Dan Teknologi Informasi*.
<http://mcastud.com/student-project-development-go/>
- Prasetyaningrum, E., & Susanti, P. (2023). Perbandingan Algoritma K-Means Dan K-Medoids Untuk Pemetaan Hasil Produksi Buah-Buahan. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 7(4), 1775.
<https://doi.org/10.30865/mib.v7i4.6477>
- Pratama, I., Idrus, M., & Sahade. (2024). ANALISIS KELAYAKAN FINANSIALUSAHA TANI KOPI ARABIKA DI DESA MASOSO KECAMATAN BAMBANG KABUPATEN MAMASA. *JURNAL REVENUE*, 5.
<https://doi.org/10.46306/rev.v5i1>
- Pratiwi, H., Purno, A., & Wibowo, W. (2022). Mengidentifikasi Kelompok Sektor Perkebunan di Indonesia Berdasarkan Produktivitas Hasil Bumi. *DIFFUSION JOURNAL OF SYSTEM AND INFORMATION TECHNOLOGY*, 54(1).
- Rahman Abdillah, Rudi Hermawan, Wawan Hermawansyah, Ibnu Adkha, & Heri Arifin. (2024). Pengujian Perangkat Lunak Sistem Informasi Inventori pada Usaha Jasa Pengiriman Paket. *Polygon : Jurnal Ilmu Komputer Dan Ilmu Pengetahuan Alam*, 2(4), 166–175. <https://doi.org/10.62383/polygon.v2i4.199>
- Rahmani, Z., Farkhatin, N., & Ningsih, R. (2022). *PERANCANGAN SISTEM INFORMASI AKADEMIK PADA PAUD GODWILLING MAMPANG DEPOK BERBASIS JAVA MENGGUNAKAN METODE R&D*.
- Soepriyono, G., & Triayudi, A. (2023). Perbandingan Kinerja Algoritma Clustering Data Mining Untuk Prediksi Harga Saham Pada Reksadana dengan Davies Bouldin Index. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 7(4), 2061.
<https://doi.org/10.30865/mib.v7i4.6623>
- Wajdi, Muh. F., & Tandililing, M. (2022). *IMPLEMENTASI METODE RESEARCH AND DEVELOPMENT (R & D) PADA APLIKASI PENGELOLAAN ARSIP*.
<https://doi.org/https://doi.org/10.56708/progres.v14i1.320>
- Yulian Pamuji, F., Rofiqul Muslikh, A., Muhammad Arief, R., & Muti, D. (2024). KOMPARASI METODE MEAN DAN KNN IMPUTATION DALAM MENGATASI MISSING VALUE PADA DATASET KECIL. *JIP (Jurnal Informatika Polinema)*, 10, 257–264.
<https://doi.org/https://doi.org/10.33795/jip.v10i2.5031>