

IMPLEMENTASI METODE CLUSTERING UNTUK PEMETAAN WILAYAH PRODUKSI DAN EKSPOR KOPI DI INDONESIA

Arya Dwi Saputra¹; Jefri Jaya²; Teny Handhayani^{3*}; Manatap Sitorus Dolok Lauro⁴

Fakultas Teknologi Informasi^{1,2,3,4}

Universitas Tarumanagara^{1,2,3,4}

<https://untar.ac.id/>^{1,2,3,4}

arya.535210049@stu.untar.ac.id¹, jefri.535210048@stu.untar.ac.id², tenyh@fti.untar.ac.id^{3*},
manataps@fti.untar.ac.id⁴

(*) Corresponding Author



Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi-Non Komersial 4.0 Internasional.

Abstract— *Coffee is one of the main agricultural commodities in Indonesia, but the distribution of production and export contribution is still uneven. This study aims to map the patterns of coffee production and export in Indonesia using clustering methods, namely K-Means and Hierarchical Agglomerative Clustering (AHC). The data used includes coffee production by province and regency (2015–2022), as well as coffee export data by destination country (2016–2023), obtained from BDSP and BPS. The system is developed in the form of an interactive website that allows users to upload datasets, select clustering methods, and view analysis results in the form of tables, graphs, and interactive maps. Clustering quality is evaluated using the Silhouette Score and Davies-Bouldin Index (DBI). The testing results show that the optimal number of clusters is two for all datasets, with the highest Silhouette score reaching 0.85 and the lowest DBI of 0.21, indicating good clustering quality. AHC is more effective in analyzing export and provincial-level production data, while K-Means performs better for regency-level data. This system is expected to provide insights into the distribution patterns of coffee production and exports and support decision-making in the agricultural sector, particularly for coffee commodities.*

Keywords: *agglomerative, coffee, clustering, export, k-means.*

Abstrak— Kopi merupakan salah satu komoditas pertanian unggulan di Indonesia, namun distribusi produksi dan kontribusi eksportnya masih belum merata. Ketimpangan ini mempengaruhi efektivitas pengembangan sektor kopi nasional. Penelitian ini bertujuan untuk memetakan pola produksi dan ekspor kopi di Indonesia menggunakan metode clustering K-Means dan Hierarchical Agglomerative Clustering (AHC). Penelitian ini menggunakan data produksi kopi per provinsi dan kabupaten (2015–2022) serta data ekspor kopi berdasarkan negara tujuan (2016–2023), yang diperoleh dari BDSP dan BPS. Sistem dikembangkan dalam bentuk website interaktif yang memungkinkan pengguna mengunggah dataset, memilih metode clustering, dan melihat hasil analisis dalam bentuk tabel, grafik, dan peta interaktif. Evaluasi kualitas clustering dilakukan menggunakan Silhouette Score dan Davies-Bouldin Index (DBI). Hasil pengujian menunjukkan bahwa jumlah cluster optimal adalah dua untuk semua dataset, dengan nilai Silhouette tertinggi mencapai 0.85 dan DBI terendah sebesar 0.21, yang menunjukkan kualitas pengelompokan yang baik. Metode AHC unggul dalam menganalisis data ekspor dan produksi tingkat provinsi, sedangkan metode K-Means lebih efektif untuk data produksi tingkat kabupaten. Sistem ini diharapkan dapat memberikan wawasan terhadap pola distribusi produksi dan ekspor kopi serta mendukung pengambilan keputusan dalam sektor pertanian, khususnya komoditas kopi.

Kata kunci: *agglomerative, kopi, klasterisasi, ekspor, k-means.*

PENDAHULUAN

Kopi merupakan tanaman perkebunan yang tumbuh di daerah tropik dan banyak ditanam di Indonesia. Tanaman ini bisa tumbuh di berbagai lokasi, selama tanahnya tidak tandus dan suhunya tidak terlalu tinggi, sehingga tidak cocok untuk tumbuh di daerah yang memenuhi kondisi tersebut (Harun, 2022). Kopi adalah biji yang diambil dari pohon yang disebut coffee, dan satu pohon bisa menghasilkan sekitar satu kilogram kopi setiap tahunnya. Kopi telah menjadi komoditas pertanian unggulan di Brasil, Vietnam, Kolombia, dan Indonesia (Freitas et al., 2024). Pemerintah Indonesia mendorong peningkatan produksi dan kualitas kopi varietas Arabika dan Robusta (Ashardiono & Trihartono, 2024). Peningkatan kualitas sensor biji kopi dapat dilakukan melalui pembentukan cita rasa dan prekursor cita rasa, optimalisasi jalur fermentasi biji kopi dan mikroorganisme, serta pembatasan pembentukan cita rasa yang tidak enak (Febrianto & Zhu, 2023). Kopi luwak merupakan salah satu kopi premium yang dihasilkan Indonesia (Farag et al., 2023).

Hasil perkebunan menjadi komoditas ekspor Indonesia. Tujuan dari ekspor ini adalah untuk mendapatkan keuntungan, karena harga barang di luar negeri biasanya lebih tinggi dibandingkan saat dijual di pasar dalam negeri. Indonesia mengekspor berbagai produk. Ini mencakup produk migas seperti minyak tanah, solar, dan bensin, serta produk perkebunan seperti karet, kopi, kakao, dan sebagainya.

Penelitian menunjukkan bahwa rata-rata pasar ekspor biji kopi Indonesia di pasar internasional sekitar 6%, sehingga Indonesia berada di posisi ketiga setelah Brazil dan Kolombia. (Adetya et al., 2025). Penelitian tersebut belum membahas tentang distribusi negara tujuan ekspor kopi Indonesia.

Sebuah studi mengkaji lokasi produksi Kopi Arabika di Kabupaten Temanggung (Fihry et al., 2023). Studi ini membahas lingkungan yang paling cocok untuk pertumbuhan Kopi Arabika. Clustering adalah cara mengelompokkan data atau objek yang memiliki karakteristik serupa. Biasanya digunakan untuk mencari pusat *cluster* secara iteratif dengan menghitung jarak terpendek antara setiap data dan pusat setiap klaster (Apriyani et al., 2023). Clustering negara tujuan ekspor kopi dilakukan menggunakan metode K-Means tetapi hasil clustering belum divisualisasikan dalam bentuk pemetaan (Putra, 2023). penelitian sebelumnya masih terbatas pada analisis produksi atau ekspor secara parsial, dan belum menggabungkan keduanya dalam satu sistem yang terintegrasi. Selain itu, belum banyak pendekatan visual

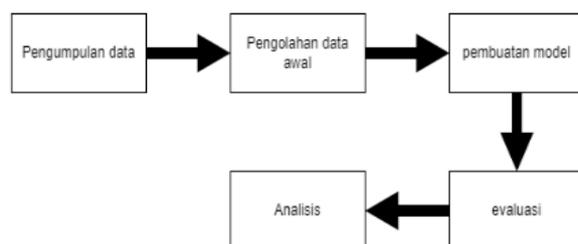
interaktif berbasis web yang dapat membantu pemangku kebijakan dalam memahami persebaran spasial produksi dan ekspor kopi secara komprehensif.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk memetakan pola distribusi wilayah produksi dan ekspor kopi di Indonesia menggunakan metode clustering, serta menyajikannya dalam bentuk sistem interaktif berbasis web yang dapat mendukung pengambilan keputusan dalam sektor pertanian, khususnya komoditas kopi.

BAHAN DAN METODE

Alur kerja penelitian ditampilkan pada Gambar 1 yang terdiri atas pengumpulan data, pengolahan data awal, pembuatan model, evaluasi dan analisis. Tahap pertama yaitu pengumpulan data. Dataset yang digunakan yaitu data produksi dan ekspor kopi Indonesia. Data produksi kopi dikumpulkan dari <https://bdsp2.pertanian.go.id/bdsp/id/lokasi>. Sedangkan data ekspor kopi dikumpulkan dari <https://www.bps.go.id/id/publication?keyword=kopi&sort=latest>. Dataset produksi kopi terdiri atas variabel lokasi kota/kabupaten di Indonesia, luas lahan (hektar), jumlah produksi (ton), dan produktivitas dari tahun 2015 – 2022. Dataset ekspor terdiri atas variable nama negara tujuan, volume ekspor (ton), dan nilai ekspor (\$) dari tahun 2016 - 2023. Tahap kedua yaitu pengolahan data. Dataset diunduh dari website kemudian dilakukan pra-pemrosesan data yaitu memilih sampel yang tidak ada nilai yang hilang. Dataset produksi dan ekspor kopi kemudian dijadikan inputan untuk algoritma clustering. Tahap ketiga yaitu pembuatan model. Penelitian ini menggunakan library scikit-learn <https://scikit-learn.org/stable/>. Seluruh proses pengolahan data dan pemodelan dilakukan menggunakan bahasa pemrograman Python, dengan library utama meliputi:

1. scikit-learn untuk pemodelan clustering dan evaluasi
2. pandas untuk manipulasi data
3. matplotlib dan plotly untuk visualisasi grafik
4. serta folium untuk visualisasi spasial.



Sumber : (Andrian et.al.,2024)

Gambar 1. Alur penelitian

Penelitian ini menggunakan dua metode clustering: K-Means dan Hierarchical Agglomerative clustering. K-Means Clustering bekerja dengan cara membagi data menjadi dua atau lebih kelompok. (Abiodun M. Ikotun and Absalom E. Ezugwu and Laith Abualigah and Belal Abuhaija and Jia Heming, 2023). Data dibagi ke dalam kelompok dengan cara ini. Ini berarti bahwa kelompok data terdiri dari kelompok yang memiliki ciri yang mirip, dan kelompok lain terdiri dari kelompok yang memiliki ciri yang berbeda. Tujuan dari pengelompokan data ini adalah untuk mendapatkan kelompok data yang memiliki kesamaan tinggi dengan variasi yang rendah (Mirafatabzadeh, Seyed Mahdi and Colombo, Cristian Giovanni and Longo, Michela and Foadelli, 2023). Rumus untuk menghitung *centroid* dari suatu *cluster* adalah :

$$C_j = \frac{1}{N} \sum_{i=1}^N X_i. \quad (1)$$

Dimana :

C_j adalah centroid dari cluster j

N adalah jumlah data dalam cluster j

X_i adalah nilai data ke- i dalam cluster

Jarak antara data ke centroid dihitung dengan rumus perhitungan jarak Euclidean:

$$d(X, C) = \sqrt{\sum_{i=1}^n (X_i - C_i)^2} \quad (2)$$

Dimana :

X adalah titik data

C adalah centroid

n adalah jumlah dimensi dalam dataset

Kelebihan dari metode K-Means clustering adalah kecepatan dalam pengolahan data yang besar, namun kelemahannya adalah memerlukan penentuan jumlah cluster yang tepat serta rentan terhadap outlier. Metode K-Means clustering berhasil digunakan untuk pengelompokan wilayah panen padi di Indonesia (Ishak et al., 2024) dan pengelompokan kota di pulau Jawa berdasarkan kondisi meteorologi (Handhayani & Rusdi, 2023).

Metode kedua yang digunakan adalah Agglomerative Hierarchical Clustering, teknik ini adalah teknik clustering membentuk hirarki sehingga terbentuk struktur (Chhabra, Anshuman and Mohapatra, 2022; Monath et al., 2021). Terdapat dua jenis algoritma Clustering Hierarkis, yaitu Agglomerative dan Divisive. (Pratama Simanjuntak & Khaira, 2021). Agglomerative clustering berhasil digunakan untuk melakukan clustering wilayah di Sumatra berdasarkan data meteorologi (Handhayani & Lewenus, 2024).

Cara kerja agglomerative dimulai dengan mengumpulkan setiap titik data secara terpisah, di mana pada awalnya masing-masing titik data berada dalam kelompok sendiri dan jumlah kelompok sama dengan jumlah seluruh titik data. (Emmendorfer, Leonardo Ramos; Canuto, 2021). Langkah selanjutnya, data yang paling dekat dibentuk menjadi satu kelompok baru. Jarak antar kelompok tersebut dihitung lagi, kemudian kelompok yang paling dekat digabungkan bersama. Beberapa metode yang digunakan pada Hierarchical clustering yaitu single linkage, average linkage, centroid linkage dan complete linkage (Li et al., 2022).

Single linkage menggunakan jarak terdekat antara dua cluster, metode ini cocok untuk data yang terpisah jauh dan memiliki cluster berbentuk konsentris. Namun, metode ini memiliki kelemahan, yaitu jika data mengandung noise atau pola tidak teratur, maka akan terbentuk cluster yang tidak terstruktur dengan baik. (Fahrudin et al., 2021).

Tahap keempat adalah evaluasi hasil pengelompokan. Untuk menilai kualitas pengelompokan digunakan metrik Silhouette Score. Metrik ini mengukur seberapa baik setiap data dalam kelompoknya dibandingkan dengan kelompok lainnya. Nilai Silhouette Score berkisar dari -1 sampai 1, di mana nilai yang lebih tinggi menunjukkan pengelompokan yang lebih baik. Jika nilainya mendekati 1, berarti data tersebut berada di dalam cluster yang tepat. Jika nilainya mendekati 0, berarti data berada di antara dua cluster. Dan jika nilainya negatif, artinya data lebih cocok berada di cluster yang berbeda. (Hasan, 2024; Suraya et al., 2023).

Silhouette Score dapat dihitung dengan rumus :

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (3)$$

Dimana:

$a(i)$ adalah jarak rata-rata suatu data dengan anggota cluster yang sama

$b(i)$ adalah jarak rata-rata suatu data dengan anggota cluster terdekat lainnya.

Davies-Bouldin Index dapat digunakan untuk memberikan gambaran tentang seberapa baik setiap kelompok terpisah satu sama lain dan seberapa dekat kelompok-kelompok tersebut (Umagapi et al., 2023). Nilai indeks yang tinggi menunjukkan ketidakjelasan atau tumpang tindih dalam pengelompokan, sedangkan nilai indeks yang lebih rendah menunjukkan pengelompokan yang lebih dan terpisah dengan baik (Suraya et al., 2023). Rumus Davies-Bouldin Index adalah:

$$DB = \frac{1}{n} \sum_{i=1}^n \max_{j \neq i} \left(\frac{\sigma_i + \sigma_j}{d_{ij}} \right) \quad (4)$$

Dimana:

σ_i adalah rata-rata jarak antara data dalam cluster i ke centroidnya.

d_{ij} adalah jarak centroid ke cluster i dan j .

Tahapan ke lima yaitu analisis. Tahapan ini menganalisis hasil clustering. Analisis dilengkapi dengan visualisasi data.

HASIL DAN PEMBAHASAN

Pengujian dilakukan terhadap masing-masing data dengan algoritma K-Means Clustering dan Agglomerative Hierarchical Clustering. Eksperimen dilakukan menggunakan jumlah cluster 2, 3, 4, 5, 6, dan 7 untuk mencari jumlah cluster dengan nilai Silhouette tertinggi:

Tabel 1. Pengujian Dataset Ekspor Agglo

Algo	Cluster	SS	DBI
Agglomerative	2	0.79	0.55
Agglomerative	3	0.76	0.21
Agglomerative	4	0.70	0.36
Agglomerative	5	0.66	0.34
Agglomerative	6	0.64	0.41
Agglomerative	7	0.63	0.45

Sumber : (Hasil Penelitian, 2025)

Tabel 2. Pengujian dataset ekspor K-Means

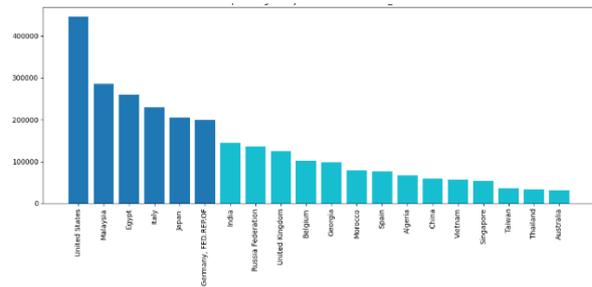
Algo	Cluster	SS	DBI
K-Means	2	0.79	0.55
K-Means	3	0.74	0.35
K-Means	4	0.70	0.36
K-Means	5	0.67	0.39
K-Means	6	0.64	0.48
K-Means	7	0.62	0.49

Sumber : (Hasil Penelitian, 2025)



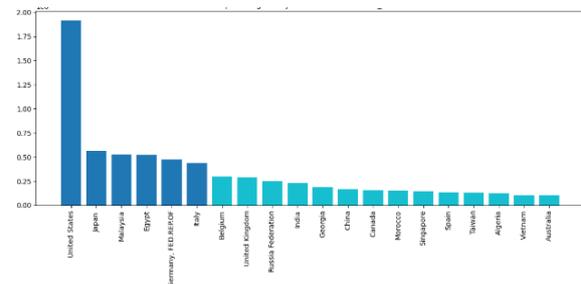
Sumber : (Hasil Penelitian, 2025)

Gambar 1. Pemetaan clustering wilayah ekspor



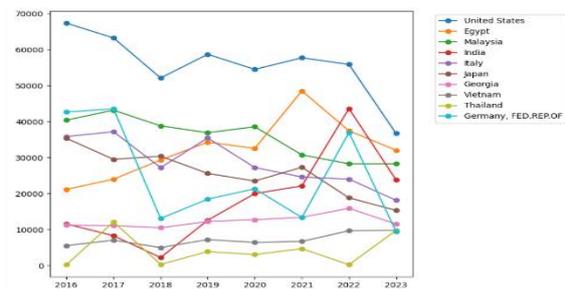
Sumber: (Hasil Penelitian, 2025)

Gambar 2. Grafik total ekspor



Sumber: (Hasil Penelitian, 2025)

Gambar 3. Grafik volume ekspor



Sumber: (Penelitian, 2025)

Gambar 4. Grafik Tren Total Ekspor

Metode Agglomerative Hierarchical Clustering dengan jumlah cluster 2 menunjukkan hasil terbaik pada pengujian pertama terhadap data ekspor kopi ke negara tujuan. Nilai Silhouette Score sebesar 0.79 dan nilai DBI sebesar 0.55 menunjukkan hasil pengelompokan yang sangat baik dan kualitas clustering yang cukup baik, Hasil eksperimen ditampilkan pada Tabel 1 dan Tabel 2.

Negara-negara tujuan ekspor dibagi menjadi dua kelompok besar menurut pola hasil clustering. Kelompok pertama terdiri dari negara-negara dengan volume dan nilai ekspor tinggi, seperti Amerika Serikat, Jepang, dan Jerman, yang merupakan pasar utama kopi di Indonesia. Kelompok kedua terdiri dari negara-negara dengan kontribusi ekspor lebih rendah, seperti Mesir, Aljazair, dan Malaysia. Grafik total ekspor ditampilkan pada Gambar 2. Sedangkan Gambar 3 menjelaskan grafik volume ekspor.

Negara seperti Amerika Serikat, Jepang, dan Jerman membentuk cluster utama dengan volume ekspor tinggi, sesuai dengan studi tentang negara tujuan utama kopi Indonesia (Suryadi Muzahidi Aziz & Nur Azizah Komara Rifai, 2022). Sementara itu, negara seperti Mesir dan Malaysia masuk cluster kedua dengan kontribusi lebih rendah karena faktor permintaan dan kebijakan impor. Gambar 2 dan 3 menunjukkan distribusi yang mencerminkan segmentasi pasar ekspor kopi: dominasi oleh beberapa negara besar, dan sisanya tersebar dalam volume kecil.

Tabel 3. Pengujian dataset kabupaten agglomerative

Algo	Cluster	SS	DBI
Agglomerative	2	0.83	0.54
Agglomerative	3	0.50	0.70
Agglomerative	4	0.40	0.78
Agglomerative	5	0.41	0.70
Agglomerative	6	0.38	0.66
Agglomerative	7	0.44	0.66

Sumber : (Hasil Penelitian, 2025)

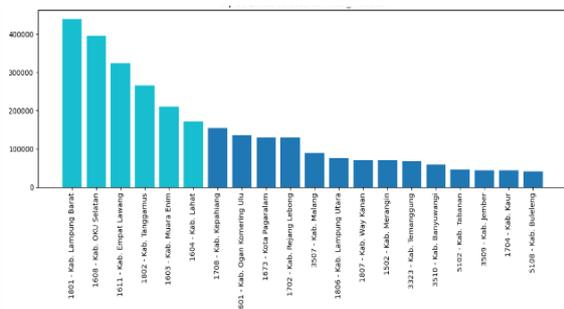
Tabel 4. Hasil pengujian analisis dataset kabupaten metode K-means

Algo	Cluster	SS	DBI
K-Means	2	0.85	0.41
K-Means	3	0.75	0.60
K-Means	4	0.45	0.70
K-Means	5	0.46	0.70
K-Means	6	0.43	0.70
K-Means	7	0.43	0.77

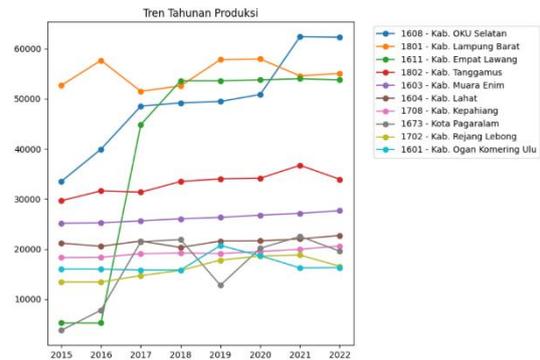
Sumber : (Hasil Penelitian, 2025)



Sumber: (Hasil Penelitian, 2025)
 Gambar 5. Pemetaan clustering wilayah kabupaten



Sumber: (Hasil Penelitian, 2025)
 Gambar 6. Grafik batang produksi



Sumber: (Hasil Penelitian, 2025)
 Gambar 7. Grafik tren produksi

Selain itu, pengujian model untuk data kopi tingkat kabupaten menampilkan bahwa metode pengelompokan K-Means pada saat jumlah cluster 2 menghasilkan nilai silhouette yaitu 0.85 dan nilai DBI sebesar 0.41, yang menunjukkan bahwa pengelompokan data kabupaten ke dalam dua kelompok besar dan memberikan pemisahan cluster yang sangat baik. Tabel 3 menampilkan hasil eksperimen dataset kabupaten menggunakan Agglomerative clustering. Tabel 4 menampilkan hasil eksperimen menggunakan algoritma K-Means.

Hasil clustering pada data produksi tingkat kabupaten membuktikan adanya perbedaan luas lahan dan produksi kopi dari setiap daerah. Kabupaten-kabupaten seperti Gayo Lues, Tapanuli Utara, dan Bener Meriah termasuk dalam cluster produksi tinggi, sementara banyak kabupaten di wilayah timur Indonesia termasuk dalam cluster produksi rendah. Ini menunjukkan bahwa ada perbedaan produksi di berbagai lokasi geografis, yang dapat dipertimbangkan saat mengatur kebijakan pengembangan komoditas kopi. Gambar dapat dilihat pada Gambar 5.

Kabupaten Gayo Lues, Tapanuli Utara, dan Bener Meriah tergabung dalam cluster produksi tinggi, didukung oleh kondisi alam dan sistem pertanian yang intensif. Kabupaten di wilayah timur cenderung masuk cluster rendah karena keterbatasan infrastruktur dan produktivitas. Gambar 5 hingga 7 menunjukkan visualisasi spasial dan tren yang membedakan jelas dua kelompok ini.

Tabel 5. Hasil pengujian analisis dataset provinsi metode Agglomerative

Algo	Cluster	SS	DBI
Agglomerative	2	0.72	0.41
Agglomerative	3	0.50	0.64
Agglomerative	4	0.51	0.58
Agglomerative	5	0.52	0.50
Agglomerative	6	0.51	0.38
Agglomerative	7	0.48	0.40

Sumber : (Hasil Penelitian, 2025)

Tabel 6. Hasil pengujian analisis dataset Provinsi metode K-Means

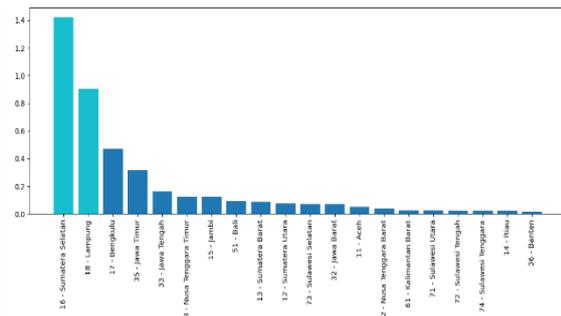
Algo	Cluster	SS	DBI
K-Means	2	0.72	0.41
K-Means	3	0.50	0.64
K-Means	4	0.43	0.56
K-Means	5	0.48	0.51
K-Means	6	0.51	0.38
K-Means	7	0.48	0.40

Sumber : (Hasil Penelitian, 2025)



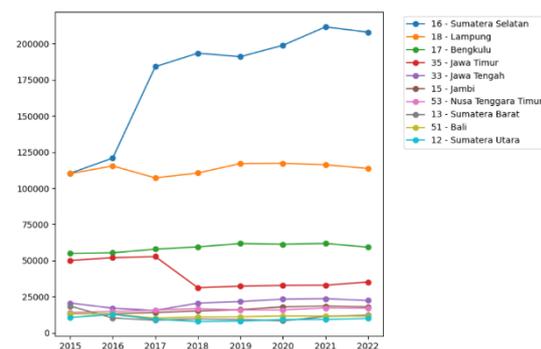
Sumber : (Hasil Penelitian, 2025)

Gambar 8. Pemetaan clustering wilayah provinsi



Sumber : (Hasil Penelitian, 2025)

Gambar 9. Grafik batang produksi



Sumber : (Hasil Penelitian, 2025)

Gambar 10. Grafik tren produksi

Pada pengujian terakhir pada data produksi tingkat provinsi, metode Agglomerative Hierarchical Clustering dengan jumlah cluster 2 menghasilkan hasil terbaik; model ini menghasilkan nilai Silhouette Score sebesar 0.72 dan nilai DBI sebesar 0.41, yang menunjukkan bahwa pemisahan antar cluster sangat baik dan berkualitas. Hasil

eksperimen clustering dataset provinsi menggunakan Agglomerative clustering ditampilkan pada Tabel 5. Tabel 6 menunjukkan hasil eksperimen menggunakan algoritma K-Means.

Pada data produksi tingkat provinsi, hasil clustering juga menunjukkan pemisahan yang jelas antara provinsi dengan kontribusi besar terhadap produksi nasional (seperti Aceh, Sumatera Utara, dan Lampung) dengan provinsi lainnya. Cluster produksi tinggi didominasi provinsi di wilayah barat Indonesia, sementara provinsi di wilayah tengah dan timur umumnya tergolong dalam cluster produksi rendah. Pola ini mencerminkan ketergantungan produksi kopi Indonesia pada wilayah tertentu. Penamaan kategori untuk masing-masing cluster, seperti “tinggi”, “sedang”, dan “rendah”, ditentukan berdasarkan nilai total produksi (atau volume ekspor) dari setiap cluster. Cluster dengan nilai tertinggi diberi label “tinggi”, diikuti oleh “sedang”, dan “rendah”. Penamaan ini bertujuan untuk memudahkan interpretasi visual dalam peta dan grafik tren tahunan. Grafik tren produksi kopi tingkat kabupaten disediakan oleh Gambar 7. Peta hasil pemetaan wilayah provinsi ditampilkan oleh Gambar 8. Gambar 9 menampilkan grafik produksi kopi dari setiap daerah. Grafik tren produksi kopi tingkat provinsi dapat dilihat pada Gambar 10. Provinsi Aceh, Lampung, dan Sumatera Utara mendominasi produksi dan tergolong dalam cluster tinggi, sesuai dengan data BPS. Wilayah tengah dan timur seperti NTB dan Papua Barat masuk dalam cluster rendah karena kontribusi produksinya minim. Gambar 8–10 menggambarkan pola ketimpangan produksi nasional secara spasial dan tahunan.

Berdasarkan hasil pengujian yang telah dilakukan terhadap tiga jenis dataset yaitu dataset ekspor kopi berdasarkan negara tujuan, dataset produksi kopi tingkat kabupaten, dan dataset produksi kopi tingkat provinsi, diperoleh sejumlah temuan penting dari penerapan metode clustering K-Means dan Agglomerative Hierarchical Clustering (AHC) dengan variasi jumlah cluster dari 2 hingga 7.

Pada dataset ekspor, nilai silhouette score dan Davies-Bouldin Index (DBI) menunjukkan bahwa metode Agglomerative Clustering dan K-Means Clustering memiliki nilai evaluasi yang sama dengan jumlah cluster 2. Nilai silhouette score mencapai 0.79 dan DBI sebesar 0.55, yang mengindikasikan bahwa hasil pengelompokan cukup baik dengan pemisahan antar cluster yang jelas dan internal cluster yang kompak.

Pada dataset produksi kopi tingkat kabupaten, metode K-Means dengan 2 cluster memberikan hasil terbaik dengan silhouette score 0.85 dan DBI sebesar 0.41. Ini menunjukkan bahwa pengelompokan data kabupaten menjadi dua

kelompok besar berhasil menangkap perbedaan signifikan antara wilayah, seperti kabupaten dengan produksi tinggi dan rendah. Nilai silhouette yang tinggi menegaskan bahwa data dalam cluster memiliki kedekatan tinggi dan jaraknya cukup jauh dari cluster lain. Untuk dataset produksi kopi tingkat provinsi, hasil terbaik juga diperoleh pada jumlah cluster 2, dan diperoleh oleh kedua metode yang memberikan nilai yang sama, yaitu Silhouette score 0.72 dan DBI 0.41.

KESIMPULAN

Evaluasi kualitas clustering menggunakan dua metrik, yaitu Silhouette Score dan Davies-Bouldin Index (DBI), menunjukkan bahwa jumlah cluster optimal untuk seluruh dataset adalah dua. Jumlah tersebut memberikan nilai evaluasi terbaik: Silhouette Score tertinggi mencapai 0.85 dan DBI terendah sebesar 0.21, yang mencerminkan kualitas pengelompokan yang baik dan pemisahan antar cluster yang jelas. 3.

Berdasarkan hasil pengelompokan, data ekspor menghasilkan dua pola utama, cluster pertama memiliki mitra dagang besar seperti Amerika Serikat, Jepang, dan Jerman. sedangkan cluster kedua mencakup negara dengan volume ekspor lebih rendah seperti Mesir dan Malaysia. Data produksi kabupaten menunjukkan dominasi produksi oleh kabupaten-kabupaten di Sumatera seperti Gayo Lues dan Tapanuli Utara. Data produksi provinsi mengindikasikan bahwa produksi kopi nasional terkonsentrasi di wilayah barat Indonesia seperti Aceh, Lampung, dan Sumatera Utara, sementara provinsi di wilayah tengah dan timur menunjukkan tingkat produksi yang relatif rendah.

Penelitian selanjutnya dapat mengembangkan sistem ini dengan menambahkan fitur prediksi produksi atau ekspor menggunakan model machine learning, serta memperluas cakupan data spasial hingga level desa. Selain itu, pendekatan clustering dapat diuji pada komoditas pertanian lainnya untuk membandingkan pola distribusi antar komoditas.

REFERENSI

- Abiodun M. Ikotun and Absalom E. Ezugwu and Laith Abualigah and Belal Abuhaija and Jia Heming. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, 622, 178–210. <https://doi.org/https://doi.org/10.1016/j.ins.2022.11.139>
- Adetya, A., Rifin, A., & Nurmalina, R. (2025). Posisi Pasar Ekspor Biji Kopi Indonesia Di Pasar Internasional: Pendekatan Almost Ideal Demand System (AIDS). *Forum Agribisnis*, 15(1), 114–123. <https://doi.org/10.29244/fagb.15.1.114-123>
- Apriyani, P., Dikananda, A. R., & Ali, I. (2023). Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi. *Hello World Jurnal Ilmu Komputer*, 2(1), 20–33. <https://doi.org/10.56211/helloworld.v2i1.230>
- Ashardiono, F., & Trihartono, A. (2024). Optimizing the potential of Indonesian coffee: a dual market approach. *Cogent Social Sciences*, 10(1). <https://doi.org/10.1080/23311886.2024.2340206>
- Chhabra, Anshuman and Mohapatra, P. (2022). Fair Algorithms for Hierarchical Agglomerative Clustering. *International Conference on Machine Learning and Applications (ICMLA)*, 206–211. <https://doi.org/10.1109/ICMLA55696.2022.00036>
- Emmendorfer, Leonardo Ramos; Canuto, A. M. de P. (2021). A generalized average linkage criterion for Hierarchical Agglomerative Clustering. *Applied Soft Computing*, 100, 106990. <https://doi.org/https://doi.org/10.1016/j.asoc.2020.106990>
- Fahrudin, T. M., Riyantoko, P. A., Hindrayani, K. M., & Swari, M. H. P. (2021). Cluster Analysis of Hospital Inpatient Service Efficiency Based on BOR, BTO, TOI, AvLOS Indicators using Agglomerative Hierarchical Clustering. *Telematika*, 18(2), 194. <https://doi.org/10.31315/telematika.v18i2.4786>
- Farag, M. A., Mohamed, T. A., El-Hawary, E. A., & Abdelwareth, A. (2023). Metabolite Profiling of Premium Civet Luwak Bio-Transformed Coffee Compared with Conventional Coffee Types, as Analyzed Using Chemometric Tools. *Metabolites*, 13(2), 173. <https://doi.org/10.3390/metabo13020173>
- Febrianto, N. A., & Zhu, F. (2023). Coffee bean processing: Emerging methods and their effects on chemical, biological and sensory properties. *Food Chemistry*, 412, 135489. <https://doi.org/10.1016/j.foodchem.2023.135489>
- Fiqhry, A. T., Santoso, T. N. B., & Ardiani, F. (2023). Kajian Produksi Kopi Arabika (Coffea

- arabica) pada Berbagai Ketinggian Tempat di Kabupaten Temanggung. *Jurnal Agro Industri Perkebunan*, 81–90. <https://doi.org/10.25181/jaip.v12i2.3497>
- Freitas, V. V., Borges, L. L. R., Vidigal, M. C. T. R., dos Santos, M. H., & Stringheta, P. C. (2024). Coffee: A comprehensive overview of origin, market, and the quality process. *Trends in Food Science & Technology*, 146, 104411. <https://doi.org/10.1016/j.tifs.2024.104411>
- Handhayani, T., & Lewenusa, I. (2024). An Analysis of Meteorological Data in Sumatra and Nearby using Agglomerative Clustering. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 8(2), 234–241. <https://doi.org/10.29207/resti.v8i2.5663>
- Handhayani, T., & Rusdi, Z. (2023). K-Means Using Dynamic Time Warping For Clustering Cities in Java Island According to Meteorological Conditions. *2023 Eighth International Conference on Informatics and Computing (ICIC)*, 1–6. <https://doi.org/10.1109/ICIC60109.2023.10381899>
- Harun, S. (2022). Analisis Produksi Kopi di Indonesia Tahun 2015-2020 Menggunakan Metode Cobb-Douglass. *Jurnal Ilmiah Ekonomi Pembangunan*, 1(2), 102–109.
- Hasan, Y. (2024). Pengukuran Silhouette Score dan Davies-Bouldin Index pada Hasil Cluster. *KAKIFIKOM (Kumpulan Artikel Karya Ilmiah Fakultas Ilmu Komputer)*, 06(01), 60–74.
- Ishak, Y. A., Handhayani, T., Sitorus, M. D. L., William, W., Pragantha, J., & Lewenusa, I. (2024). Advanced Clustering Approach for Mapping Regions of Paddy Productivity in Indonesia Using Intelligent K-Means. *IEEE Asia-Pacific Conference on Geoscience, Electronics and Remote Sensing Technology (AGERS)*, 269–274.
- Li, T., Rezaeipanah, A., & Tag El Din, E. M. (2022). An ensemble agglomerative hierarchical clustering algorithm based on clusters clustering technique and the novel similarity measurement. *Journal of King Saud University - Computer and Information Sciences*, 34(6), 3828–3842. <https://doi.org/10.1016/j.jksuci.2022.04.010>
- Miraftebzadeh, Seyed Mahdi and Colombo, Cristian Giovanni and Longo, Michela and Foadelli, F. (2023). K-Means and Alternative Clustering Methods in Modern Power Systems. *IEEE Access*, 11. <https://doi.org/10.1109/ACCESS.2023.3327640>
- Monath, N., Dubey, K. A., Guruganesh, G., Zaheer, M., Ahmed, A., McCallum, A., Mergen, G., Najork, M., Terzihan, M., Tjanaka, B., Wang, Y., & Wu, Y. (2021). Scalable Hierarchical Agglomerative Clustering. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1245–1255. <https://doi.org/10.1145/3447548.3467404>
- Pratama Simanjuntak, K., & Khaira, U. (2021). Hotspot Clustering in Jambi Province Using Agglomerative Hierarchical Clustering Algorithm. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1(1), 7–16.
- Putra, I. K. N. (2023). Implementasi Metode K-Means Clustering Pada Ekspor Kopi Berdasarkan Negara Tujuan. *Jurnal Ilmu Komputer Dan Bisnis*, 14(1), 45–55. <https://doi.org/10.47927/jikb.v14i1.332>
- Suraya, S., Sholeh, M., & Lestari, U. (2023). Evaluation of Data Clustering Accuracy using K-Means Algorithm. *International Journal of Multidisciplinary Approach Research and Science*, 2(01), 385–396. <https://doi.org/10.59653/ijmars.v2i01.504>
- Suryadi Muzahidi Aziz, & Nur Azizah Komara Rifai. (2022). Pengelompokan Ekspor Kopi Menurut Negara Tujuan Menggunakan Metode K-Means Clustering dengan Silhouette Coefficient. *Bandung Conference Series: Statistics*, 2(2), 416–424. <https://doi.org/10.29313/bcss.v2i2.4536>
- Umagapi, I. T., Umaternate, B., Komputer, S., Pasca Sarjana Universitas Handayani, P., Kepegawaian Daerah Kabupaten Pulau Morotai, B., & Riset dan Inovasi, B. (2023). Uji Kinerja K-Means Clustering Menggunakan Davies-Bouldin Index Pada Pengelompokan Data Prestasi Siswa. *Seminar Nasional Sistem Inform asi Dan Teknologi (SISFOTEK)*, 7(1), 303–308.