

Analisis Klasifikasi Produktivitas Tanaman Biofarmaka di Kabupaten Sumedang Menggunakan Metode K-Means Clustering

Ery Supriyadi Rustidja¹
Universitas Koperasi Indonesia¹
erysr@ikopin.ac.id¹

ABSTRAK

Penelitian ini bertujuan untuk mengelompokkan komoditas tanaman biofarmaka di Kabupaten Sumedang berdasarkan tingkat produksi, luas panen, dan produktivitas menggunakan metode *K-Means Clustering*. Data yang digunakan merupakan data sekunder tahun 2025 yang diperoleh dari Badan Pusat Statistik Kabupaten Sumedang. Tahapan analisis meliputi *preprocessing data* dengan normalisasi menggunakan metode *Min-Max*, penentuan jumlah *cluster* menggunakan metode *Elbow*, serta evaluasi hasil *clustering* menggunakan *Silhouette Score*.

Hasil penelitian menunjukkan bahwa jumlah cluster optimal adalah tiga cluster. *Cluster 0* terdiri dari komoditas kencur dan lengkuas yang memiliki tingkat produksi dan luas panen tinggi namun produktivitas relatif rendah. *Cluster 1* terdiri dari jahe dan kunyit yang memiliki produktivitas tinggi meskipun produksi dan luas panen rendah. Sementara itu, cluster 2 terdiri dari kapulaga, jeruk nipis, dan serai yang memiliki tingkat produksi, luas panen, dan produktivitas relatif rendah.

Hasil evaluasi menggunakan *Silhouette Score* sebesar 0,500 menunjukkan bahwa kualitas *clustering* berada pada kategori cukup baik, sehingga metode *K-Means* mampu memberikan pengelompokan yang cukup representatif. Penelitian ini diharapkan dapat memberikan informasi dalam menentukan prioritas pengembangan komoditas biofarmaka di Kabupaten Sumedang.

Kata Kunci: *K-Means Clustering, Tanaman Biofarmaka, Produktivitas, Clustering, Sumedang*

I. PENDAHULUAN

Sektor pertanian memiliki peran strategis dalam pembangunan ekonomi nasional, terutama dalam mendukung ketahanan pangan, penyediaan lapangan kerja, serta peningkatan kesejahteraan masyarakat. Selain komoditas pangan utama, subsektor hortikultura dan biofarmaka juga menjadi bagian penting yang berkontribusi terhadap diversifikasi produk pertanian. Tanaman biofarmaka seperti jahe, kunyit, kencur, dan kapulaga memiliki nilai ekonomi yang tinggi serta semakin dibutuhkan dalam industri kesehatan dan herbal (Almira et al., 2023).

Kabupaten Sumedang merupakan salah satu daerah yang memiliki potensi dalam pengembangan tanaman biofarmaka. Berdasarkan data produksi tanaman biofarmaka tahun 2025, beberapa komoditas menunjukkan tingkat produksi yang cukup tinggi, seperti kencur dengan produksi mencapai sekitar 11.859.688 kg, kunyit sebesar 4.132.334 kg, lengkuas sebesar 4.437.387 kg, dan jahe sebesar 1.945.451 kg. Selain itu, luas panen yang cukup besar pada beberapa komoditas menunjukkan potensi pengembangan yang signifikan, namun belum diketahui secara pasti komoditas mana yang memiliki kinerja produksi terbaik.

Meskipun demikian, tidak seluruh komoditas biofarmaka memiliki tingkat produksi dan produktivitas yang sama. Perbedaan tersebut terlihat dari variasi nilai produksi, luas panen, dan produktivitas antar komoditas yang dapat dilihat pada Tabel 1. Kapulaga memiliki produktivitas yang relatif rendah dibandingkan komoditas lain, sedangkan lengkuas memiliki tingkat produktivitas yang lebih tinggi. Variasi tersebut menunjukkan adanya perbedaan

performa antar komoditas, namun hingga saat ini belum terdapat klasifikasi yang jelas mengenai komoditas biofarmaka berdasarkan tingkat kinerja produksinya di Kabupaten Sumedang.

Tabel 1. Tingkat Produksi dan Produktivitas Tanaman Biofarmaka di Kabupaten Sumedang

No	Komoditas	Tingkat Produksi (Kg)	Produktivitas (Kg/m ²)
1	Jahe	1.945.451	3,39
2	Kapulaga	390.507	1,26
3	Kencur	11.859.688	2,93
4	Kunyit	4.132.334	3,64
5	Lengkuas	4.437.387	4,58
6	Jeruk Nipis	93.395	22,05
7	Serai	3.500	4,38

Sumber: Badan Pusat Statistik Kabupaten Sumedang, data diolah (2026)

Selama ini, analisis dalam sektor pertanian sering difokuskan pada perbandingan antar wilayah atau kecamatan. Namun, dalam konteks data yang tersedia, unit analisis yang lebih relevan adalah komoditas tanaman itu sendiri, bukan wilayah. Hal ini dikarenakan data yang digunakan berbasis jenis tanaman biofarmaka, sehingga pendekatan klasifikasi yang tepat adalah mengelompokkan komoditas berdasarkan karakteristik produksinya, bukan berdasarkan lokasi geografis.

Pengelompokan komoditas biofarmaka menjadi penting untuk mengidentifikasi komoditas unggulan, komoditas yang berkembang sedang, serta komoditas yang masih tertinggal. Dengan adanya klasifikasi tersebut, pemangku kebijakan dapat lebih mudah menentukan prioritas pengembangan, baik dalam hal peningkatan produksi, efisiensi lahan, maupun optimalisasi produktivitas.

Oleh karena itu, diperlukan suatu metode analisis yang mampu mengelompokkan komoditas secara objektif berdasarkan karakteristik numerik yang dimiliki, salah satunya adalah metode *K-Means Clustering*. Metode ini merupakan teknik *data mining* yang digunakan untuk mengelompokkan data berdasarkan kemiripan karakteristik tertentu. Dalam penelitian ini, *K-Means* digunakan untuk mengelompokkan komoditas biofarmaka berdasarkan variabel produksi, luas panen, dan produktivitas, sehingga dapat menghasilkan *cluster* yang merepresentasikan tingkat kinerja masing-masing komoditas.

Penelitian terdahulu yang dilakukan oleh Kholifah & Bahri (2025) serta Anwar et al. (2024) menunjukkan bahwa metode *K-Means Clustering* efektif dalam mengelompokkan data pertanian, khususnya untuk mengidentifikasi ketimpangan berbasis wilayah. Namun, kedua penelitian tersebut lebih berfokus pada pengelompokan wilayah atau kecamatan berdasarkan indikator pertanian, sehingga belum membahas klasifikasi komoditas pertanian secara langsung. Oleh karena itu, penelitian ini berfokus pada klasifikasi komoditas biofarmaka untuk memberikan perspektif yang berbeda dibandingkan penelitian sebelumnya.

Beberapa penelitian terdahulu menunjukkan bahwa metode *K-Means* efektif dalam melakukan pengelompokan data, khususnya dalam sektor pertanian (Nikhar & Thakre, 2020). Penelitian yang menggunakan *K-Means* dalam analisis komoditas pertanian menunjukkan bahwa metode ini mampu mengidentifikasi kelompok komoditas unggulan dan menentukan

prioritas pengembangan. Selain itu, penelitian lain juga menunjukkan bahwa penerapan *K-Means* pada data produktivitas pertanian dapat memberikan hasil klasifikasi yang cukup akurat dan mudah diinterpretasikan (Septianto et al., 2025; Trisnawati & Purnamasari, 2023).

Berdasarkan uraian tersebut, maka penelitian ini bertujuan untuk mengelompokkan komoditas tanaman biofarmaka di Kabupaten Sumedang menggunakan metode *K-Means Clustering*. Dengan adanya penelitian ini, diharapkan dapat diperoleh informasi mengenai klasifikasi komoditas biofarmaka yang dapat digunakan sebagai dasar dalam pengambilan kebijakan serta pengembangan sektor pertanian yang lebih efektif dan terarah.

II. TINJAUAN PUSTAKA

A. Pertanian dan Tanaman Biofarmaka

Sektor pertanian merupakan salah satu pilar utama pembangunan ekonomi, khususnya dalam mendukung ketahanan pangan, penyediaan bahan baku industri, serta peningkatan kesejahteraan masyarakat. Dalam perkembangannya, sektor pertanian tidak hanya berfokus pada komoditas pangan utama, tetapi juga mencakup subsektor hortikultura dan tanaman biofarmaka.

Tanaman biofarmaka adalah tanaman yang memiliki kandungan zat aktif yang bermanfaat untuk kesehatan dan sering digunakan sebagai bahan baku obat tradisional maupun herbal (Almira et al., 2023). Komoditas seperti jahe, kunyit, kencur, dan kapulaga merupakan contoh tanaman biofarmaka yang memiliki nilai ekonomi tinggi serta permintaan pasar yang terus meningkat.

Pengembangan tanaman biofarmaka menjadi penting karena selain memiliki potensi ekonomi, komoditas ini juga mendukung industri kesehatan dan pengobatan alami. Oleh karena itu, diperlukan analisis yang tepat untuk mengidentifikasi potensi dan kinerja masing-masing komoditas biofarmaka (Satria et al., 2023).

B. Produksi, Luas Panen, dan Produktivitas

Produksi pertanian merupakan jumlah hasil panen yang diperoleh dari suatu kegiatan budidaya dalam periode tertentu. Produksi dipengaruhi oleh berbagai faktor seperti luas lahan, teknologi budidaya, serta kondisi lingkungan (Soekartawi, 2005).

Luas panen menunjukkan luas area yang digunakan untuk menghasilkan suatu komoditas. Variabel ini penting karena menggambarkan skala produksi suatu tanaman.

Sementara itu, produktivitas merupakan ukuran efisiensi produksi yang diperoleh dari perbandingan antara hasil produksi dengan luas lahan yang digunakan. Produktivitas digunakan untuk menilai tingkat keberhasilan suatu kegiatan pertanian dalam memanfaatkan sumber daya yang tersedia (Nicholson & Snyder, 2012). Secara matematis, produktivitas dapat dirumuskan sebagai:

$$\text{Produktivitas} = \frac{\text{Produksi}}{\text{Luas Panen}} \quad (1)$$

Ketiga variabel tersebut saling berkaitan dan digunakan sebagai indikator utama dalam menilai kinerja suatu komoditas pertanian. Oleh karena itu, variabel produksi, luas panen, dan produktivitas sering digunakan dalam analisis klasifikasi dan pengelompokan data pertanian.

C. *Data Mining*

Data mining adalah proses pola atau informasi yang bermakna dari sekumpulan data dalam jumlah besar dengan menggunakan teknik tertentu (Han et al., 2012). *Data mining* merupakan bagian dari proses *Knowledge Discovery in Database* (KDD) yang bertujuan untuk mengolah data menjadi informasi yang berguna (Larose & Larose, 2014).

Teknik dalam *data mining* meliputi berbagai metode, seperti klasifikasi, regresi, asosiasi, dan *clustering*. Dalam penelitian ini, metode yang digunakan adalah *clustering*, yaitu teknik pengelompokan data berdasarkan kemiripan karakteristik tanpa menggunakan label.

D. *Clustering*

Clustering merupakan metode dalam *data mining* yang bertujuan untuk mengelompokkan objek ke dalam beberapa kelompok (*cluster*) sehingga objek dalam satu *cluster* memiliki tingkat kemiripan yang tinggi, sedangkan antar *cluster* memiliki perbedaan yang signifikan (Jain, 2010).

Clustering termasuk ke dalam metode *unsupervised learning*, di mana data tidak memiliki label sebelumnya. Hasil dari *clustering* berupa kelompok-kelompok data yang dapat digunakan untuk mengetahui pola atau struktur dalam data.

Dalam konteks penelitian ini, *clustering* digunakan untuk mengelompokkan komoditas biofarmaka berdasarkan karakteristik produksi, luas panen, dan produktivitas.

E. *K-Means Clustering*

K-Means Clustering merupakan salah satu metode *clustering* yang paling banyak digunakan dalam analisis data (Jain, 2010). Metode ini bekerja dengan membagi data ke dalam sejumlah cluster (k) berdasarkan kedekatan jarak terhadap titik pusat *cluster* (*centroid*). Langkah-langkah *K-Means*, diantaranya: (1) menentukan jumlah klaster (k); (2) menentukan *centroid* awal; (3) menghitung jarak setiap data ke *centroid*; (4) mengelompokkan data ke *cluster* terdekat; (5) menghitung *centroid* baru; (6) mengulangi proses hingga konvergen.

Perhitungan jarak menggunakan *euclidean distance* (d) yang dapat dilihat pada rumus 2.

$$d = \sqrt{\sum (x_i - c_i)^2} \quad (2)$$

Keterangan untuk rumus 2: d = *euclidean distance*; x_i = data ke- i ; c_i = *centroid* ke- i .

K-Means memiliki kelebihan yaitu mudah diimplementasikan dan efisien dalam mengelompokkan data dalam jumlah besar. Namun, metode ini juga memiliki kekurangan, seperti sensitivitas terhadap pemilihan *centroid* awal dan jumlah *cluster*.

F. *Elbow Method*

Elbow method digunakan untuk menentukan jumlah *cluster* optimal dalam *K-Means*. Metode ini dilakukan dengan menghitung nilai *Sum of Squared Error* (SSE) untuk berbagai jumlah *cluster*, kemudian memilih jumlah *cluster* yang menghasilkan penurunan nilai SSE yang mulai melambat (Tan et al., 2006). Rumus SSE dapat dilihat pada rumus 3.

$$SSE = \sum (\text{jarak data ke centroid})^2 \quad (3)$$

Jumlah *cluster* optimal ditentukan berdasarkan grafik SSE terhadap k , dimana titik *elbow* menunjukkan bahwa penurunan *error* mulai melambat.

G. *Silhouette Score*

Silhouette score digunakan untuk mengukur kualitas hasil *clustering*. Nilai ini menunjukkan seberapa baik suatu data berada dalam *cluster*-nya dibandingkan dengan *cluster* lain (Rousseeuw, 1987). Rumus untuk ini dapat dilihat pada rumus 4.

$$S = \frac{b - a}{\max(a, b)} \quad (4)$$

Keterangan untuk rumus 4: S = *silhouette score*; a = jarak rata-rata ke *cluster sendiri*; b = jarak rata-rata ke *cluster* lain. Adapun hasil dari *silhouette score* memiliki makna: (1) jika S mendekati 1, maka *cluster* sangat baik; (2) jika S mendekati 0, maka *cluster* kurang jelas; dan (3) jika S mendekati -1, maka terjadi salah pengelompokan.

III. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis *data mining*, khususnya teknik *clustering* menggunakan algoritma *K-Means*. Pendekatan kuantitatif digunakan karena data yang dianalisis berupa data numerik yang terdiri dari variabel produksi, luas panen, dan produktivitas tanaman biofarmaka.

Penelitian ini bertujuan untuk mengelompokkan komoditas tanaman biofarmaka berdasarkan kemiripan karakteristik produksi, sehingga dapat diketahui kelompok komoditas unggulan, sedang, dan kurang berkembang.

Jenis data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari instansi resmi, yaitu Badan Pusat Statistik (BPS) Kabupaten Sumedang. Data yang digunakan berupa: (1) produksi tanaman biofarmaka (kg); (2) luas panen (m^2); dan (3) produktivitas (kg/m^2). Data yang dianalisis difokuskan pada tahun 2025 untuk mencerminkan kondisi terkini sektor biofarmaka di Kabupaten Sumedang.

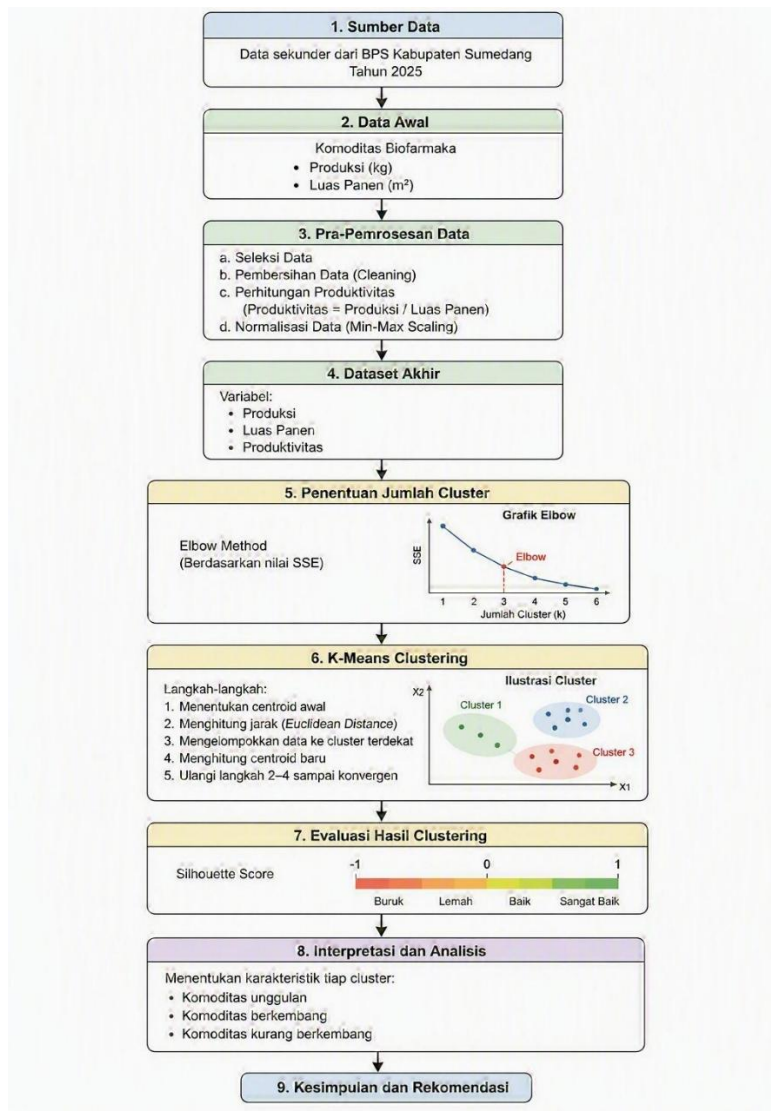
Objek dalam penelitian ini adalah komoditas tanaman biofarmaka di Kabupaten Sumedang, diantaranya (1) jahe; (2) kapulaga; (3) kencur; (4) kunyit; (5) lengkuas; (6) jeruk nipis; dan (7) serai. Setiap komoditas diperlakukan sebagai satu unit data yang dianalisis berdasarkan variabel produksi, luas panen, dan produktivitas.

Berdasarkan paparan sebelumnya, maka variabel yang digunakan dalam penelitian ini adalah:

Tabel 2. Operasionalisasi Variabel Penelitian

Variabel	Keterangan	Skala Data
Produksi	Jumlah hasil panen (kg)	Rasio
Luas Panen	Luas lahan yang dipanen (m^2)	Rasio
Produktivitas	Hasil produksi per satuan luas (kg/m^2)	Rasio

Selanjutnya terkait teknik analisis data dalam penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Teknik Analisis Data

IV. HASIL PENELITIAN DAN PEMBAHASAN

A. Deskripsi Data Penelitian

Data yang digunakan dalam penelitian ini terdiri dari 7 komoditas tanaman biofarmaka di Kabupaten Sumedang tahun 2025. Variabel yang digunakan meliputi produksi, luas panen, dan produktivitas. Data diperoleh dari Badan Pusat Statistik Kabupaten Sumedang dan telah melalui tahap pembersihan serta normalisasi sebelum dilakukan analisis *clustering*.

Tabel 3. Tingkat Produksi, Luas Panen, dan Produktivitas Tanaman Biofarmaka di Kabupaten Sumedang Tahun 2025

Jenis Tanaman	Produksi (kg)	Luas Panen (m ²)	Produktivitas (kg/m ²)
Jahe	1.945.451	574.563	3,39
Kapulaga	390.507	309.140	1,26
Kencur	11.859.688	4.048.520	2,93
Kunyit	4.132.334	1.134.579	3,64
Lengkuas	4.437.387	969.267	4,58
Jeruk Nipis	93.395	4.236	22,05
Serai	3.500	800	4,38

Sumber: Badan Pusat Statistik Kabupaten Sumedang, data diolah (2026)

Berdasarkan tabel 3, terlihat bahwa tanaman kencur memiliki nilai produksi tertinggi, sedangkan serai memiliki nilai produksi terendah. Sedangkan dilihat dari produktivitas, jeruk nipis menunjukkan nilai yang lebih tinggi dibandingkan komoditas lainnya.

B. Hasil *Preprocessing* Data

Tahap *preprocessing* data dilakukan sebelum proses *clustering* untuk memastikan kualitas data yang digunakan. Proses ini meliputi seleksi data, pengecekan kelengkapan data, serta normalisasi variabel menggunakan metode *Min-Max*. Normalisasi dilakukan untuk menghindari perbedaan skala antar variabel produksi, luas panen, dan produktivitas. Berdasarkan hasil normalisasi pada Tabel 4, terlihat bahwa setiap variabel telah berada pada rentang nilai 0 hingga 1, sehingga data telah siap digunakan pada proses *clustering* menggunakan metode *K-Means*. Selain itu, terlihat bahwa komoditas kencur memiliki nilai maksimum pada variabel produksi dan luas panen, sedangkan jeruk nipis memiliki nilai maksimum pada variabel produktivitas.

Tabel 4. Hasil Normalisasi Data Tingkat Produksi, Luas Panen, dan Produktivitas Tanaman Biofarmaka di Kabupaten Sumedang Tahun 2025

Jenis Tanaman	Produksi	Luas Panen	Produktivitas
Jahe	0,163792	0,141750	0,102130
Kapulaga	0,032642	0,076176	0,000000
Kencur	1,000000	1,000000	0,080164
Kunyit	0,348243	0,280103	0,114458
Lengkuas	0,373972	0,239262	0,159486
Jeruk Nipis	0,007582	0,000849	1,000000
Serai	0,000000	0,000000	0,149716

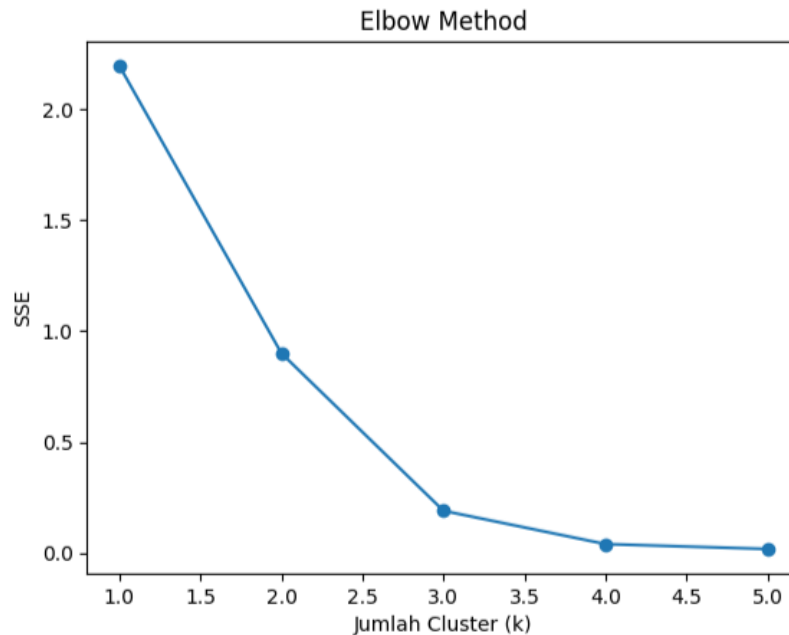
Sumber: Data diolah (2026)

C. *Elbow Method*

Penentuan jumlah *cluster* optimal dalam penelitian ini dilakukan menggunakan metode *Elbow*, yaitu dengan menganalisis nilai *Sum of Squared Error* (SSE) pada beberapa jumlah *cluster*. Metode ini bertujuan untuk menentukan jumlah *cluster* yang paling tepat sehingga hasil *clustering* dapat merepresentasikan struktur data secara optimal.

Nilai SSE dihitung berdasarkan jumlah kuadrat jarak antara data dengan *centroid* pada masing-masing *cluster*. Semakin kecil nilai SSE menunjukkan bahwa jarak antar data dalam satu *cluster* semakin dekat. Namun demikian, penambahan jumlah *cluster* secara terus-menerus akan menurunkan nilai SSE, sehingga diperlukan suatu titik optimal yang disebut sebagai *elbow*.

Hasil pengujian metode *Elbow* pada data tanaman biofarmaka ditunjukkan pada Gambar 2.



Gambar 2. Grafik *Elbow Method*

Berdasarkan Gambar 2, terlihat bahwa penurunan nilai SSE terjadi secara signifikan dari *cluster* 1 ke *cluster* 3. Setelah jumlah *cluster* mencapai 3, penurunan nilai SSE cenderung melambat dan tidak terlalu signifikan. Hal ini menunjukkan bahwa titik *elbow* berada pada nilai $k = 3$.

Dengan demikian, jumlah *cluster* yang digunakan dalam penelitian ini adalah 3 *cluster*, yang selanjutnya digunakan dalam proses *clustering* menggunakan metode *K-Means*. Pemilihan jumlah *cluster* sebanyak 3 juga mempertimbangkan keseimbangan antara kompleksitas model dan interpretasi hasil *clustering*.

D. Hasil *Clustering K-Means*

Setelah dilakukan penentuan jumlah *cluster* menggunakan metode *Elbow*, diperoleh jumlah *cluster* optimal sebanyak $k = 3$. Selanjutnya, proses *clustering* dilakukan menggunakan metode *K-Means* untuk mengelompokkan komoditas tanaman biofarmaka di Kabupaten Sumedang berdasarkan variabel produksi, luas panen, dan produktivitas.

Hasil pengolahan data menggunakan metode *K-Means* menghasilkan pengelompokan komoditas ke dalam tiga *cluster*, sebagaimana ditunjukkan pada Tabel 5.

Tabel 5. Hasil *Clustering* Komoditas Tanaman Biofarmaka

Cluster	Komoditas
0	Kencur
	Lengkuas
1	Jahe
	Kunyit
2	Kapulaga
	Jeruk Nipis
	Serai

Sumber: Data diolah (2026)

Selain itu, untuk mengetahui karakteristik masing-masing *cluster*, dilakukan perhitungan nilai rata-rata (*centroid*) dari setiap variabel pada tiap *cluster*. Kriteria nilai *centroid* dalam

penelitian ini ditentukan berdasarkan rentang nilai hasil normalisasi, yaitu (1) 0,0 – 0,33 dikategorikan rendah; (2) 0,34 - 0,66 dikategorikan sedang; dan (3) 0,67 – 1,00 dikategorikan tinggi. Hasil perhitungan *centroid* dapat dilihat pada Tabel 6.

Tabel 6. Nilai *Centroid* Tiap *Cluster*

Cluster	Produksi		Luas Panen		Produktivitas	
	Skor	Kriteria	Skor	Kriteria	Skor	Kriteria
0	1,000	Tinggi	1,000	Tinggi	0,080	Rendah
1	0,007	Rendah	0,000	Rendah	1,000	Tinggi
2	0,183	Rendah	0,148	Rendah	0,105	Rendah

Sumber: Data diolah (2026)

Berdasarkan hasil *clustering* tersebut, dapat dilihat bahwa setiap *cluster* memiliki karakteristik yang berbeda. *Cluster* 0 terdiri dari komoditas dengan nilai produksi dan luas panen yang tinggi, yaitu kencur dan lengkuas. *Cluster* ini dapat dikategorikan sebagai kelompok komoditas unggulan.

Cluster 1 terdiri dari komoditas dengan nilai produksi dan luas panen pada tingkat rendah, namun memiliki produktivitas yang tinggi. *Cluster* ini menunjukkan komoditas yang cukup stabil dan memiliki potensi untuk dikembangkan lebih lanjut.

Sementara itu, *cluster* 2 terdiri dari komoditas dengan nilai produksi relatif rendah, yaitu kapulaga, jeruk nipis, dan serai. Namun demikian, jeruk nipis memiliki nilai produktivitas yang sangat tinggi dibandingkan komoditas lainnya, sehingga menunjukkan adanya potensi efisiensi lahan yang tinggi meskipun dari sisi produksi masih rendah.

E. Evaluasi Hasil *Clustering* dengan *Silhouette Score*

Evaluasi kualitas hasil *clustering* dalam penelitian ini dilakukan menggunakan metode *Silhouette Score*. Metode ini digunakan untuk mengukur seberapa baik setiap data berada dalam *cluster* yang telah terbentuk, serta untuk melihat tingkat pemisahan antar *cluster*.

Nilai *Silhouette Score* dihitung berdasarkan perbandingan antara jarak rata-rata suatu data terhadap anggota *cluster* yang sama (*cohesion*) dan jarak rata-rata data tersebut terhadap *cluster* terdekat lainnya (*separation*). Nilai *Silhouette Score* berada pada rentang -1 hingga 1, di mana nilai yang mendekati 1 menunjukkan bahwa data telah terkelompok dengan baik, sedangkan nilai yang mendekati 0 menunjukkan bahwa pengelompokan kurang jelas, dan nilai negatif menunjukkan adanya kesalahan dalam pengelompokan.

Berdasarkan hasil perhitungan menggunakan metode *K-Means*, diperoleh nilai *Silhouette Score* sebesar 0,500. Nilai tersebut menunjukkan bahwa hasil *clustering* berada pada kategori cukup baik, di mana data dalam setiap *cluster* memiliki tingkat kedekatan yang relatif baik, meskipun masih terdapat beberapa data yang memiliki kedekatan dengan *cluster* lain.

Hal ini mengindikasikan bahwa struktur *cluster* yang terbentuk sudah mampu merepresentasikan pola data, namun pemisahan antar *cluster* belum sepenuhnya optimal. Kondisi ini dapat disebabkan oleh adanya perbedaan skala atau karakteristik yang cukup ekstrem pada beberapa komoditas, seperti nilai produktivitas jeruk nipis yang jauh lebih tinggi dibandingkan komoditas lainnya.

Dengan demikian, metode *K-Means* yang digunakan dalam penelitian ini telah mampu menghasilkan pengelompokan yang cukup representatif, meskipun masih terdapat ruang untuk peningkatan kualitas clustering pada penelitian selanjutnya.

V. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa metode *K-Means Clustering* pada komoditas tanaman biofarmaka di Kabupaten Sumedang menghasilkan 3 *cluster*, yaitu *cluster* 0 yang terdiri dari kencur dan lengkuas, *cluster* 1 yang terdiri dari jahe dan kunyit, serta *cluster* 2 yang terdiri dari kapulaga, jeruk nipis, dan serai.

Hasil pengelompokan menunjukkan bahwa setiap *cluster* memiliki karakteristik yang berbeda, di mana *cluster* 0 merupakan komoditas dengan tingkat produksi dan luas panen tinggi, *cluster* 1 menunjukkan komoditas dengan produktivitas tinggi namun produksi dan luas panen rendah, sedangkan *cluster* 2 merupakan komoditas dengan tingkat produksi, luas panen, dan produktivitas yang relatif rendah.

Hasil evaluasi *clustering* menunjukkan nilai *silhouette score* sebesar 0,500 yang mengindikasikan bahwa hasil *clustering* berada pada kategori cukup baik, sehingga metode *K-Means* mampu menghasilkan pengelompokan yang cukup representatif, meskipun masih terdapat peluang untuk peningkatan kualitas *clustering*.

DAFTAR PUSTAKA

- Almira, Z., Yundari, Y., & Im'roah, N. (2023). Implementasi Metode Polinomial Newton Gregory untuk Mengestimasi Produksi Tanaman Biofarmaka di Kalimantan Barat. *Jurnal EurekaMatika*, 11(1), 69–78. <https://doi.org/10.17509/jem.v11i1.57451>
- Anwar, A. M., Rinaldi, A. R., & Mulyawan, M. (2024). Perbandingan Algoritma K-Means dan K-Medoids Dalam Pengelompokan Kabupaten dan Kota Berdasarkan Tanaman Biofarmaka. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), 3388–3393. <https://doi.org/10.36040/jati.v7i6.8188>
- Han, Jiawei., Kamber, Micheline., & Pei, Jian. (2012). *Data mining : concepts and techniques*. Elsevier/Morgan Kaufmann.
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651–666. <https://doi.org/10.1016/j.patrec.2009.09.011>
- Kholifah, F. N., & Bahri, S. (2025). Pengelompokan Daerah Produksi Tanaman Biofarmaka Menurut Jenis Tanaman dengan Metode K – Means Clustering. *ESTIMASI: Journal of Statistics and Its Application*, 6(1), 61–69. <https://doi.org/https://doi.org/10.20956/ejsa.v6i1.27200>
- Larose, D. T. ., & Larose, C. D. . (2014). *Discovering knowledge in data : an introduction to data mining*. Wiley.
- Nicholson, Walter., & Snyder, Christopher. (2012). *Microeconomic theory : basic principles and extensions*. South-Western/Cengage Learning.

- Nikhar, M., & Thakre, L. (2020). Smart Agricultural Farm Enhancement with K-Means Learning. *International Journal of Innovative Technology and Exploring Engineering*, 9(8), 166–170. <https://doi.org/10.35940/ijitee.H6222.069820>
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Satria, I., Manurung, A. A., & Hanafiah, M. A. (2023). Analisis Perkembangan Produksi Tanaman Biofarmaka (Obat) di Indonesia Menggunakan Algoritma Resilient. *BRAHMANA: Jurnal Penerapan Kecerdasan Buatan*, 5(1), 122–131.
- Septianto, M. A., Faqih, A., & Rinaldi, A. R. (2025). Klasterisasi Data Produksi Pertanian di Kabupaten Cirebon dengan Algoritma K-Means. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(2). <https://doi.org/10.23960/jitet.v13i2.6174>
- Soekartawi. (2005). *Agribisnis Teori dan Aplikasinya*. Raja Grafindo Persada.
- Tan, P.-Ning., Steinbach, Michael., & Kumar, Vipin. (2006). *Introduction to data mining*. Pearson Addison Wesley.
- Trisnawati, P., & Purnamasari, A. I. (2023). Penerapan Pengelompokan Produktivitas Hasil Pertanian Menggunakan Algoritma K-Means. *Infotek: Jurnal Informatika Dan Teknologi*, 6(2), 249–257. <https://doi.org/10.29408/jit.v6i2.10198>

