

Klasterisasi Sektor Pertanian Produk Domestik Regional Bruto (PDRB) Menurut Lapangan Usaha untuk Perencanaan Wilayah Berkelanjutan di Indonesia

Clustering of the Agricultural Sectors of Gross Regional Domestic Product (GRDP) by Industry for Sustainable Regional Planning in Indonesia

Ayub Sahala Gultom¹, Nicea Roona Paranoan², Caecilia Bintang Girik Allo³, Lolita Tuhumena⁴

Diterima: 26 Agustus 2025

Disetujui: 28 November 2025

Abstrak: Pembangunan wilayah berkelanjutan memerlukan pemetaan struktur ekonomi regional yang komprehensif, terutama pada sektor primer di bidang pertanian yang menjadi basis utama perekonomian di banyak provinsi Indonesia. Penelitian ini bertujuan mengklasterisasi sektor pertanian Produk Domestik Regional Bruto (PDRB) menurut lapangan usaha guna mengidentifikasi pola dan karakteristik pertumbuhan antar wilayah. Data diambil dari publikasi Statistik Nilai Tukar Petani BPS, mencakup subsektor tanaman pangan, hortikultura, perkebunan, peternakan, dan perikanan, dianalisis menggunakan metode K-Means dan Hierarchical. Evaluasi dengan Davies-Bouldin Index (DBI) menunjukkan K-Means lebih unggul (DBI 0,17) dibandingkan Hierarchical (DBI 1,36), menandakan klaster yang lebih kompak dan terpisah jelas. Hasil analisis menghasilkan empat klaster utama: dua klaster berorientasi ekspor melalui perkebunan dan perikanan, serta dua lainnya fokus pada subsektor pangan dan hortikultura dengan diversifikasi seimbang. Empat klaster ini dinilai paling bermakna secara konseptual dan kebijakan karena mencerminkan tipologi struktural utama ekonomi primer Indonesia. Pembagian tersebut konsisten dengan teori regional differentiation dan place-based development, yang menekankan pentingnya keunggulan komparatif lokal dalam perencanaan pembangunan. Setiap klaster merepresentasikan arah kebijakan berbeda penguatan ekspor, ketahanan pangan, diversifikasi pedesaan, dan keberlanjutan lingkungan. Hasil penelitian ini yaitu berupa klasterisasi dapat digunakan sebagai place based policy untuk mendorong pertumbuhan inklusif, memperkuat ketahanan pangan, mengembangkan ekspor berbasis sumber daya lokal, serta mendukung pengelolaan sumber daya alam secara berkelanjutan. Pendekatan ini memperkuat dasar empiris bagi pengambilan keputusan pembangunan wilayah, sehingga strategi pembangunan dapat disesuaikan dengan karakteristik spesifik tiap provinsi, sekaligus mendorong tercapainya target SDGs terkait ekonomi, pengentasan kemiskinan, dan pembangunan berkelanjutan

Kata kunci: klasterisasi, K-Means, Hierarchical Clustering, place-based, policy

Abstract: Sustainable regional development requires a comprehensive mapping of regional economic structures, particularly within the primary sector, which serves as the main economic foundation in many Indonesian provinces. This study aims to cluster the primary sector of the Gross Regional Domestic Product (GRDP) by industry classification to identify interregional

¹ Program Studi Planologi, Universitas Cenderawasih

^{2,3} Program Studi Statistika, Universitas Cenderawasih

⁴ Program Studi Ilmu Perikanan, Universitas Cenderawasih

Korespondensi: gultomayub@gmail.com

growth patterns and characteristics. Data were obtained from the Statistics of Farmers' Terms of Trade (BPS), covering the subsectors of food crops, horticulture, plantations, livestock, and fisheries, and analyzed using the K-Means and Hierarchical clustering methods. Evaluation using the Davies-Bouldin Index (DBI) showed that K-Means performed better (DBI = 0.17) compared to Hierarchical (DBI = 1.36), indicating more compact and clearly separated clusters. The analysis resulted in four main clusters: two clusters oriented toward exports through plantation and fisheries subsectors, and two others focusing on food crop and horticultural subsectors with balanced diversification. These four clusters are considered the most conceptually and policy-relevant, as they reflect the structural typologies of Indonesia's primary economy. The classification aligns with the theories of regional differentiation and place-based development, emphasizing the importance of local comparative advantages in development planning. Each cluster represents a distinct policy direction export strengthening, food security, rural diversification, and environmental sustainability. The resulting clusterization can be utilized as a place based policy framework to promote inclusive growth, enhance food resilience, develop resource-based exports, and support sustainable natural resource management. This approach strengthens the empirical foundation for regional development decision-making, allowing strategies to be tailored to the specific characteristics of each province while advancing the achievement of SDG targets related to economic growth, poverty alleviation, and sustainable development.

Keywords: clustering, K-Means, Hierarchical Clustering, place based policy.

PENDAHULUAN

Pembangunan wilayah yang berkelanjutan merupakan salah satu agenda strategis dalam kebijakan pembangunan nasional Indonesia. Konsep ini tidak hanya menekankan pada pertumbuhan ekonomi semata, melainkan juga menitikberatkan pada pemerataan, ketahanan pangan, serta keberlanjutan lingkungan (Todaro & Smith, 2021; United Nations, 2015; Bappenas, 2020; OECD, 2020). Dalam konteks tersebut, pemetaan struktur ekonomi regional menjadi aspek yang krusial, terutama untuk sektor primer yang masih menjadi basis utama perekonomian di banyak provinsi di Indonesia. Sektor pertanian, yang meliputi subsektor tanaman pangan, hortikultura, perkebunan, peternakan, dan perikanan, memiliki kontribusi signifikan dalam penyediaan lapangan kerja, pemenuhan kebutuhan domestik, serta penciptaan nilai tambah melalui orientasi ekspor (Badan Pusat Statistik, 2024; FAO, 2021; Kementerian Pertanian RI, 2023).

Namun, kontribusi sektor primer antar provinsi tidak seragam. Terdapat variasi struktur subsektor yang terdapat pada sektor primer mencerminkan perbedaan orientasi pembangunan daerah. Beberapa wilayah lebih berorientasi ekspor melalui komoditas perkebunan dan perikanan, sementara wilayah lain lebih menitikberatkan pada diversifikasi pangan dan hortikultura untuk memperkuat ketahanan pangan lokal. Perbedaan struktur ini mengindikasikan perlunya analisis yang lebih dalam mengenai pola pertumbuhan sektor primer antar wilayah. Salah satu pendekatan yang dapat digunakan adalah metode analisis kluster yang mampu mengelompokkan wilayah berdasarkan kesamaan karakteristik ekonomi (Hair et al., 2019; Everitt et al., 2011; Jain, 2010; Xu & Wunsch, 2009).

Analisis kluster memungkinkan pemetaan variasi antar wilayah secara lebih sistematis dan objektif. Dengan menggunakan algoritma seperti K-Means dan *Hierarchical Clustering*, pola distribusi sektor dalam PDRB dapat diidentifikasi secara lebih jelas (Johnson & Wichern, 2018; Jain, 2010; Zhang et al., 1996). Evaluasi kinerja model klusterisasi dapat dilakukan dengan indeks validasi, salah satunya *Davies-Bouldin Index* (DBI) yang mengukur tingkat kompak dan terpisahnya suatu kluster (Davies & Bouldin, 1979; Rousseeuw, 1987; Hastie et al., 2021). Hasil penelitian sebelumnya menunjukkan bahwa K-Means sering memberikan kluster yang lebih baik dalam konteks ekonomi

dibandingkan metode Hierarchical, karena mampu menghasilkan struktur kluster yang lebih homogen dan jelas (IDSS, 2025; Widiyanto et al., 2024).

Dalam konteks pembangunan wilayah, klasterisasi subsektor pertanian tidak hanya bermanfaat untuk tujuan analitis, tetapi juga memiliki implikasi strategis dalam perumusan kebijakan. Konsep pembangunan berbasis tempat (*place-based development*) menekankan pentingnya memahami keunggulan spesifik suatu wilayah untuk merancang kebijakan yang sesuai dengan potensi dan kebutuhan lokal (Rodríguez-Pose et al., 2012; CLPorter, 1998; OECD, 2020). Hasil klasterisasi sektor primer PDRB dapat memberikan dasar empiris yang kuat bagi pemerintah pusat maupun daerah dalam merumuskan strategi pembangunan yang lebih tepat sasaran, baik dalam mendorong ekspor komoditas unggulan maupun memperkuat ketahanan pangan nasional. Dengan memahami struktur kluster sektor pertanian, kebijakan pembangunan dapat diarahkan tidak hanya pada peningkatan PDRB, tetapi juga pada efisiensi sumber daya, pengurangan emisi, serta peningkatan kesejahteraan masyarakat secara berkelanjutan.

Dengan demikian, penelitian mengenai klasterisasi sektor primer PDRB menjadi penting untuk dilakukan. Selain memberikan gambaran empiris tentang variasi struktur subsektor yang menjadi bagian dari sektor primer PDRB antar provinsi di Indonesia, penelitian ini juga berkontribusi terhadap penguatan literatur dan praktik pembangunan wilayah berbasis data. Secara ilmiah, penelitian ini memiliki tiga kontribusi utama, yaitu menyediakan bukti empirik terkini berbasis provinsi terkait struktur ekonomi sektor primer yang belum banyak dieksplorasi dalam konteks pembangunan berkelanjutan di Indonesia, mengintegrasikan validasi kluster multi-metrik (dengan menggunakan *Davies-Bouldin Index* dan *Silhouette Coefficient Index* untuk memastikan ketepatan dan reliabilitas hasil klasterisasi secara komprehensif, serta menawarkan lintasan kebijakan (*policy pathways*) berbasis hasil kluster untuk mendukung formulasi strategi pembangunan wilayah yang adaptif terhadap potensi sektoral dan kondisi geografis masing-masing provinsi. Dengan pendekatan tersebut, penelitian ini tidak hanya memperkuat basis teoretis mengenai analisis kluster ekonomi regional, tetapi juga memberikan arah praktis dalam implementasi *evidence based regional development policy* yang sejalan dengan pencapaian *Sustainable Development Goals* (SDGs).

METODE

Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan metode analisis kluster untuk mengelompokkan wilayah berdasarkan kemiripan kontribusi sektor pertanian terhadap Produk Domestik Regional Bruto (PDRB) di Indonesia. Pendekatan ini dipilih untuk memberikan gambaran objektif mengenai struktur ekonomi regional serta mengidentifikasi pola pengelompokan wilayah yang relevan dalam perencanaan wilayah berkelanjutan. Pengelompokan dalam penelitian ini sebanyak 34 provinsi di Indonesia sebelum pemekaran. DKI Jakarta dimasukkan sebagai wilayah analisis meskipun tidak memiliki sektor pertanian PDRB.

Sumber Data dan Teknik Analisis Data

Data yang digunakan dalam penelitian ini merupakan data sekunder, yaitu data kontribusi sektor pertanian PDRB terhadap total PDRB menurut lapangan usaha di masing-masing provinsi yang meliputi subsektor: tanaman pangan, hortikultura, perkebunan, peternakan, dan perikanan. Sumber data yang digunakan berasal dari publikasi *publikasi Statistik Nilai Tukar Petani* Badan Pusat Statistik Indonesia tahun 2024. Penelitian ini menggunakan pendekatan deskriptif-analitis yang terdiri dari empat tahap dasar: pengumpulan data, pengelompokan data, penyajian data, dan penarikan kesimpulan (Miles et al., 2014). Data yang telah dikumpulkan akan dianalisis

menggunakan teknik analisis kluster yaitu dengan metode Hirarkhi (*Hierarchical Clustering*) dan Non-Hirarkhi (*K-Means Clustering*).

K-Means memiliki keterbatasan yaitu sensitivitas terhadap *outlier*. Algoritma K-Means sangat sensitif terhadap *outlier* karena menggunakan mean (rata-rata) sebagai pusat kluster. *Outlier* dengan nilai ekstrem dapat menarik posisi centroid menjauh dari pusat sebenarnya, sehingga mengubah struktur kluster secara signifikan (Han et al. 2012). Misalnya, satu nilai ekstrem dapat menyebabkan anggota kluster lain dikelompokkan secara keliru. Berikut ini cara mengatasi sensitivitas terhadap *outlier* (Kaufman & Rousseeuw, 2009).

1. Pra-pemrosesan data: Lakukan deteksi dan pembersihan outlier menggunakan metode statistik (Z-score, IQR) atau algoritma seperti Isolation Forest dan DBSCAN.
2. Gunakan varian yang lebih robust: Terapkan K-Medoids (PAM) atau K-Medians, yang menggunakan median atau medoid (titik data sebenarnya) sebagai pusat kluster sehingga lebih tahan terhadap outlier.
3. Normalisasi data: Lakukan standardisasi atau scaling agar perbedaan skala antar variabel tidak memperbesar pengaruh outlier.

Selain itu, K-Means memiliki sensitivitas terhadap parameter k (jumlah kluster). Pemilihan jumlah kluster (k) sangat memengaruhi hasil. Tidak ada kriteria internal dalam K-Means untuk menentukan k optimal secara otomatis. Jika k terlalu kecil maka informasi heterogenitas data hilang. Sedangkan jika terlalu besar, maka muncul *overfitting*. Adapun cara mengatasi sensitivitas terhadap parameter k yaitu (Rousseeuw, 1987):

1. Gunakan metode evaluasi seperti Elbow Method, Silhouette Coefficient, Gap Statistic, atau Davies-Bouldin Index untuk menentukan K yang optimal.
2. Terapkan algoritma adaptif seperti X-Means atau G-Means, yang memperluas K-Means dengan mekanisme otomatis menentukan jumlah kluster berdasar distribusi statistik.

Metode *Hierarchical Clustering* memiliki keterbatasan yaitu sensitivitas terhadap *outlier*. Hierarchical clustering (terutama versi *agglomerative*) juga rentan terhadap outlier karena proses penggabungan dilakukan secara bertahap tanpa mekanisme penghapusan atau penyesuaian data ekstrem. Satu *outlier* dapat menjadi kluster tersendiri atau menyebabkan struktur dendrogram menyimpang. Cara mengatasi permasalahan tersebut diantaranya (Ward, 1963; Hair et al., 2019):

1. Pra-pemrosesan data untuk mendeteksi dan menghapus outlier sebelum clustering.
2. Gunakan *robust linkage methods*, seperti *Ward's method* yang meminimalkan variansi intra kluster, atau median linkage, yang relatif lebih stabil terhadap nilai ekstrem.
3. Terapkan *trimmed hierarchical clustering*, di mana sejumlah kecil data dengan jarak terbesar diabaikan selama proses penggabungan.

Selain itu metode *Hierarchical Clustering* memiliki keterbatasan yaitu memiliki sensitivitas terhadap parameter dan skala. *Hierarchical clustering* memerlukan pemilihan metrik jarak (misalnya Euclidean, Manhattan, Mahalanobis) dan metode *linkage* (*single, complete, average, Ward*). Kombinasi yang tidak sesuai dapat menghasilkan kluster yang tidak representatif. Selain itu, skala variabel yang berbeda dapat mendistorsi jarak antar objek. Cara mengatasi permasalahan ini diantaranya yaitu:

1. Normalisasi data sebelum analisis agar semua variabel berkontribusi seimbang.
2. Eksperimen dengan beberapa kombinasi *linkage* dan metrik kemudian validasi hasil menggunakan *cophenetic correlation coefficient* atau *Silhouette Score*.
3. Gunakan *hybrid approach*, misalnya melakukan K-Means setelah *Hierarchical* (atau sebaliknya) untuk memperbaiki hasil akhir.

Berikut ini definisi seluruh variabel yang digunakan dalam penelitian ini.

Tabel 1. Definisi Operasional Variabel Penelitian

No.	Subsektor	Definisi Variabel
1.	Tanaman Pangan	Nilai tambah bruto yang dihasilkan dari kegiatan budidaya tanaman pangan seperti padi, jagung, ubi kayu, ubi jalar, dan kacang-kacangan. Menggambarkan kontribusi subsektor pertanian pangan terhadap total PDRB daerah.
2.	Hortikultura	Nilai tambah bruto dari produksi buah-buahan, sayuran, tanaman hias, dan tanaman obat. Mencerminkan tingkat diversifikasi produk pertanian bernilai tinggi dan potensi ekspor nontradisional.
3.	Perkebunan	Nilai tambah bruto yang berasal dari budidaya tanaman perkebunan seperti kelapa sawit, kakao, kopi, karet, cengkeh, dan kelapa. Menunjukkan kontribusi subsektor berorientasi ekspor terhadap struktur ekonomi regional.
4.	Peternakan	Nilai tambah bruto yang dihasilkan dari produksi ternak besar (sapi, kerbau), ternak kecil (kambing, domba, babi), unggas, serta hasil olahan seperti susu, telur, dan daging. Mengindikasikan peran subsektor protein hewani dalam ketahanan pangan dan pendapatan rumah tangga.
5.	Perikanan	Nilai tambah bruto dari kegiatan penangkapan dan budidaya ikan laut dan perairan darat (termasuk udang, rumput laut, dan ikan konsumsi). Mewakili potensi ekonomi berbasis sumber daya pesisir dan laut.

Tahap-tahap untuk menganalisis data adalah sebagai berikut:

1. Menginput data
2. Melakukan analisis kluster menggunakan Non-Hirarkhi (K-Means Clustering) dan metode Hirarkhi
3. Melakukan uji perbandingan nilai Davies-Bouldin Index (DBI) kedua metode
4. Interpretasi hasil analisis kluster
5. Menarik kesimpulan dan saran.

Preprocessing Data

Tahap awal yang dilakukan dalam penelitian ini adalah *preprocessing* data, yaitu serangkaian langkah untuk memastikan bahwa data yang digunakan telah bersih, konsisten, dan siap diproses lebih lanjut dalam algoritma *clustering*. Dari hasil pemeriksaan, ditemukan adanya sejumlah nilai yang hilang (*missing values*) dalam hal ini adalah wilayah DKI Jakarta tidak memiliki sektor primer pertanian PDRB pada seluruh variabel sektor primer pertanian PDRB. Apabila tidak ditangani dengan baik, nilai yang hilang ini dapat menimbulkan bias dalam hasil analisis. Untuk mengatasi masalah tersebut digunakan teknik imputasi rata-rata (*mean imputation*), di mana nilai yang kosong digantikan dengan nilai rata-rata dari variabel yang bersangkutan.

Setelah proses imputasi, langkah berikutnya adalah melakukan normalisasi data dengan metode *Z-Score* menggunakan bantuan *software RStudio*. Normalisasi sangat penting karena setiap variabel dalam dataset memiliki satuan dan skala yang berbeda, sehingga tanpa normalisasi proses *clustering* akan bias terhadap variabel yang memiliki rentang nilai lebih besar. Dengan penerapan *Z-Score*, setiap variabel ditransformasikan sehingga memiliki rata-rata nol dan standar deviasi satu. Hal ini menjamin bahwa kontribusi tiap variabel dalam proses pengelompokan memiliki bobot yang sama, sehingga hasil *clustering* lebih representatif dan tidak dipengaruhi oleh perbedaan skala antar variabel.

HASIL DAN PEMBAHASAN

1. Analisis Statistika Deskriptif

Dalam penelitian ini, analisis statistika deskriptif dilakukan untuk mengetahui karakteristik umum masing-masing provinsi di Indonesia tahun 2024 menurut sektor primer pertanian PDRB.

Tabel 2. Hasil Analisis Statistika Deskriptif Variabel Penelitian

Variabel Penelitian	Nilai Minimum	Nilai Maksimum	Rata-rata	Standar Deviasi
Subsektor Tanaman Pangan (X_1)	92,74	117,78	102,06	6,11
Subsektor Tanaman Hortikultura (X_2)	88,36	137,60	107,91	10,77
Subsektor Tanaman Perkebunan (X_3)	91,31	170,64	119,23	22,96
Subsektor Peternakan (X_4)	93,28	110,49	102,21	4,63
Subsektor Perikanan (X_5)	91,45	114,44	104,42	5,65

Hasil analisis statistika deskriptif pada Tabel 2 dapat diuraikan sebagai berikut:

Subsektor Tanaman Pangan (X_1)

- Rata-rata (mean) sebesar 102,06 menunjukkan kinerja subsektor tanaman pangan relatif stabil dan moderat dibanding subsektor lainnya.
- Standar deviasi sebesar 6,11 tergolong rendah, menandakan variasi antarwilayah (atau antar periode) tidak terlalu besar.
- Berdasarkan hasil analisis statistika deskriptif terlihat bahwa subsektor tanaman pangan relatif konsisten, mencerminkan ketahanan pangan yang cukup baik dan distribusi produktivitas yang merata antar wilayah.

Subsektor Tanaman Hortikultura (X_2)

- Rata-rata sebesar 107,91 menunjukkan kontribusi moderat.
- Standar deviasi sebesar 10,77 lebih tinggi dibanding subsektor pangan, artinya variasi *output* hortikultura lebih besar.
- Berdasarkan hasil analisis statistika deskriptif terlihat bahwa terdapat ketimpangan produktivitas antar wilayah yang bisa disebabkan oleh perbedaan akses infrastruktur, teknologi, dan kondisi agroklimat yang bervariasi.

Subsektor Tanaman Perkebunan (X_3)

- Rata-rata sebesar 119,23 adalah yang tertinggi di antara semua subsektor.
- Standar deviasi sebesar 22,96 juga tertinggi, menunjukkan fluktuasi yang sangat besar antar wilayah atau waktu.
- Berdasarkan hasil analisis statistika deskriptif terlihat bahwa subsektor perkebunan merupakan motor utama pertumbuhan ekonomi primer, namun tidak merata pada beberapa daerah memiliki basis perkebunan unggul (seperti sawit, kakao, kopi), sementara lainnya masih tertinggal.

Subsektor Peternakan (X_4)

- Rata-rata sebesar 102,21 menunjukkan kinerja stabil.
- Standar deviasi sebesar 4,63 merupakan yang terendah, menandakan distribusi paling seragam di antara subsektor lainnya.
- Berdasarkan hasil analisis statistika deskriptif terlihat bahwa subsektor peternakan tidak terlalu fluktuatif, cenderung tumbuh secara inkremental dan merata di seluruh wilayah.

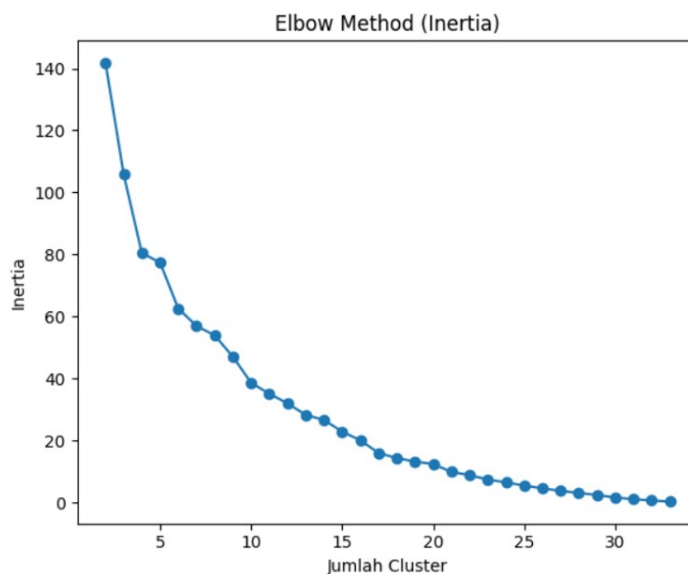
Subsektor Perikanan (X_5)

- Rata-rata sebesar 104,42 menunjukkan kontribusi cukup signifikan.
- Standar deviasi sebesar 5,65 menandakan variasi moderat.

- c. Berdasarkan hasil analisis statistika deskriptif terlihat bahwa subsektor perikanan memiliki potensi pertumbuhan yang stabil, namun masih terdapat disparitas antar wilayah pesisir tergantung infrastruktur, teknologi tangkap, dan akses pasar.

2. Hasil *Clustering* dengan K-Means

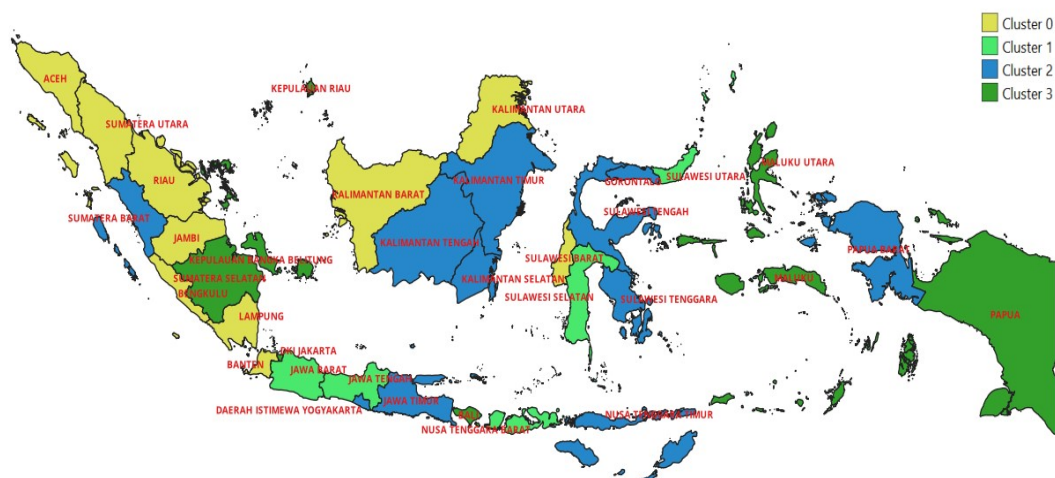
Untuk menentukan jumlah kluster yang paling sesuai, digunakan metode elbow yang didasarkan pada analisis nilai inerti atau jumlah kuadrat jarak antara data dan pusat kluster.



Gambar 1. Grafik Elbow

Hasil analisis *elbow method* (Gambar 1) menunjukkan pola penurunan nilai *inertia* yang signifikan dari kluster ke-2 hingga kluster ke-4. Penurunan tajam ini mengindikasikan bahwa penambahan kluster dalam rentang tersebut secara substansial meningkatkan kompaktitas dan homogenitas dalam masing-masing kluster, sehingga setiap tambahan kluster memberikan kontribusi nyata dalam mengurangi variasi internal. Setelah mencapai kluster ke-4, penurunan nilai *inertia* menjadi lebih lambat dan tidak terlalu signifikan, menunjukkan bahwa penambahan kluster selanjutnya hanya memberikan perbaikan marginal dalam struktur kluster. Pola inilah yang menandai adanya titik tekukan (*elbow point*) pada $k = 4$. Secara interpretatif, titik tekukan ini menjadi indikator jumlah kluster yang paling ideal, karena pada titik tersebut keseimbangan antara jumlah kluster dan efisiensi pemisahan data tercapai. Dengan kata lain, memilih $k = 4$ memungkinkan pembagian data yang optimal, di mana kluster yang terbentuk cukup kompak secara internal namun tetap berbeda secara jelas antara kluster satu dengan lainnya. Oleh karena itu, $k = 4$ dipilih sebagai jumlah kluster final untuk analisis berikutnya, memastikan hasil klusterisasi mampu merepresentasikan pola dan karakteristik utama dalam distribusi data secara akurat tanpa menimbulkan kompleksitas yang tidak perlu. Selain menggunakan metode elbow, digunakan pula nilai *Silhouette Coefficient Index* (SC) untuk mendukung pembentukan jumlah kelompok. Dalam penelitian ini diperoleh nilai SC tertinggi sebesar 0,19 untuk metode *K-Means* dengan terbentuk 4 kelompok.

Berdasarkan hasil analisis metode elbow dan nilai SC diperoleh bahwa terdapat 4 kluster yang terbentuk, sehingga berikut ini adalah pemetaan wilayah masing-masing provinsi di Indonesia dengan menggunakan metode *K-Means*.



Gambar 2. Pemetaan Wilayah Berdasarkan Metode K-Means

Adapun hasil karakteristik masing-masing klaster adalah sebagai berikut:

- a. Cluster 0 – Dominasi perkebunan: Provinsi yang termasuk dalam *cluster* ini adalah Provinsi Aceh, Sumatera Utara, Riau, Jambi, Bengkulu, Lampung, Banten, Kalimantan Barat, Kalimantan Utara, dan Sulawesi Barat. *Cluster* ini ditandai oleh dominasi subsektor perkebunan, yang berarti struktur ekonomi pertanian wilayahnya sangat bertumpu pada komoditas perkebunan. Contoh komoditas utama di wilayah ini antara lain kelapa sawit, karet, kopi, dan kakao. Tingginya kontribusi subsektor ini mencerminkan orientasi ekspor dan daya dukung lahan yang luas, namun juga menunjukkan kerentanan terhadap fluktuasi harga global. Subsektor lain seperti hortikultura, pangan, peternakan, dan perikanan relatif moderat sehingga tidak menjadi penopang utama. Pernyataan ini didukung oleh data subsektor primer per provinsi, terlihat dengan jelas bahwa provinsi-provinsi yang tergabung dalam klaster ini menunjukkan nilai subsektor tanaman perkebunan yang secara konsisten jauh lebih tinggi dibandingkan subsektor lainnya, seperti tanaman pangan, hortikultura, peternakan, dan perikanan. Fenomena ini menegaskan bahwa struktur ekonomi pertanian di wilayah-wilayah tersebut sangat bertumpu pada komoditas perkebunan, terutama kelapa sawit, karet, kopi, dan kakao.
- b. Cluster 1 – Kuat di hortikultura dan pangan: Provinsi yang termasuk dalam *cluster* 1 adalah Provinsi Jawa Barat, Jawa Tengah, Nusa Tenggara Barat, Sulawesi Utara, dan Sulawesi Selatan. *Cluster* ini memiliki ciri khas kekuatan subsektor hortikultura (sayuran, buah-buahan, dan tanaman florikultura) serta tanaman pangan (padi, jagung, kedelai). Daerah-daerah dalam kelompok ini umumnya dikenal sebagai lumbung pangan nasional, terutama Jawa Barat dan Jawa Tengah yang berperan besar dalam produksi beras. Diversifikasi pertanian relatif baik karena subsektor lainnya tetap berkontribusi meskipun tidak dominan. Hal ini menunjukkan bahwa wilayah cluster ini berorientasi pada ketahanan pangan domestik. Pernyataan ini didukung dari data empiris yang menunjukkan bahwa *Cluster* ini memperoleh nilai indeks tertinggi di antara subsektor primer lainnya. Hal ini merefleksikan struktur ekonomi yang berorientasi pada ketahanan pangan nasional, diversifikasi pertanian, serta produktivitas tinggi sektor pangan dan hortikultura.
- c. Cluster 2 – Seimbang/merata semua subsektor: Provinsi yang termasuk dalam *cluster* 2 adalah Provinsi Sumatera Barat, D.I. Yogyakarta, Jawa Timur, Nusa

Tenggara Timur, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, dan Papua Barat. Cluster ini menunjukkan struktur subsektor yang merata tanpa ada subsektor yang terlalu menonjol. Distribusi kekuatan antar subsektor (pangan, hortikultura, perkebunan, peternakan, perikanan) cenderung setara. Kondisi ini menandakan adanya diversifikasi ekonomi pertanian yang relatif baik, yang memungkinkan wilayah ini lebih tahan terhadap guncangan pada subsektor tertentu. Dengan demikian, *cluster* ini bisa dikategorikan sebagai wilayah dengan sistem pertanian yang resilien dan beragam. Hal ini didukung dengan data PDRB menurut lapangan usaha pada sektor primer pertanian, provinsi-provinsi dalam *Cluster 2* menunjukkan distribusi kekuatan subsektor yang relatif merata, baik pada subsektor tanaman pangan, hortikultura, perkebunan, peternakan, maupun perikanan. Nilai pada masing-masing subsektor bergerak dalam rentang 90–120, tanpa ada subsektor yang mendominasi secara ekstrem.

- d. Cluster 3 – Lemah di pangan dan hortikultura, kuat di perikanan dan perkebunan: Provinsi yang termasuk dalam *cluster* ini adalah Provinsi Sumatera Selatan, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, Bali, Maluku, Maluku Utara, dan Papua. Cluster ini ditandai oleh rendahnya kontribusi subsektor tanaman pangan dan hortikultura, namun relatif kuat di subsektor perikanan dan perkebunan. Wilayah ini umumnya memiliki potensi laut yang besar (misalnya Maluku, Maluku Utara, Papua) serta sektor perkebunan (misalnya kelapa sawit di Sumatera Selatan dan Kep. Riau). Struktur ini menunjukkan adanya peluang pengembangan berbasis sumber daya perikanan tangkap maupun budidaya, serta perkebunan untuk pasar domestik maupun ekspor. Namun, kelemahan pada subsektor pangan dapat menimbulkan tantangan dalam mencapai swasembada pangan di wilayah ini. Hal ini didukung dengan data PDRB sektor pertanian menurut lapangan usaha, provinsi-provinsi dalam *Cluster* ini menunjukkan pola ketimpangan subsektor yang khas di mana kontribusi subsektor tanaman pangan dan hortikultura relatif rendah (sekitar 90–107), sementara subsektor perkebunan dan perikanan memiliki nilai tinggi (sekitar 105–120). Pola ini menunjukkan pergeseran struktur ekonomi dari basis agraris ke arah ekonomi berbasis sumber daya laut dan perkebunan komersial.

3. Hasil *Clustering* dengan *Hierarchical*

Selain metode K-Means, analisis juga dilakukan dengan metode *Hierarchical Clustering* menggunakan *linkage* tertentu. Hasil dendrogram pada Gambar 3 menunjukkan adanya pemotongan pada jarak sekitar 5, yang menghasilkan pembentukan empat klaster. Pemotongan pada jarak tersebut dianggap ideal karena pada titik ini struktur klaster masih jelas terbentuk, dengan setiap klaster memiliki jumlah anggota yang relatif seimbang. Pemilihan ini juga mencegah data terpecah menjadi terlalu banyak bagian kecil, yang justru dapat mengaburkan interpretasi.

Selain itu, garis penggabungan yang lebih panjang dalam dendrogram menandakan adanya perbedaan yang signifikan antar klaster. Oleh karena itu, pemotongan pada jarak 5 dipandang tepat, karena memberikan keseimbangan antara tingkat ketelitian (*accuracy*) dan kesederhanaan model (*simplicity*). Dengan demikian, metode *hierarchical* juga mendukung hasil dari metode K-Means, bahwa jumlah klaster ideal untuk penelitian ini adalah empat. Selain dilihat dari hasil dendrogram metode *hierarchical clustering* untuk memperoleh jumlah kelompok yang terbentuk, dapat didukung dengan melihat nilai SC yang diperoleh. Nilai SC tertinggi yang diperoleh dengan menggunakan metode *hierarchical clustering* sebesar 0,23 dengan terbentuknya 4 kelompok / klaster.

serupa terlihat pada Maluku dan Maluku Utara, di mana subsektor perikanan dan perkebunan sedikit menonjol, tetapi tidak menggeser subsektor lainnya secara ekstrem. Keseimbangan ini mencerminkan daya tahan ekonomi (*resilience*) sektor pertanian terhadap guncangan subsektor tertentu, karena ketergantungan pada satu komoditas rendah. Wilayah dalam *cluster* ini cenderung memiliki kombinasi antara potensi darat dan laut, serta struktur ekonomi yang mulai bergeser menuju pola agribisnis multikomoditas

- b. Cluster 1 – Relatif kuat di hortikultura dan pangan, lemah di perikanan: Provinsi yang termasuk dalam *cluster* ini diantaranya Provinsi Sumatera Barat, Bengkulu, D.I. Yogyakarta, Jawa Timur, Nusa Tenggara Timur, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Sulawesi Tenggara, Gorontalo, dan Papua Barat. Cluster ini memperlihatkan dominasi subsektor hortikultura dan tanaman pangan dengan nilai berkisar 110–125, sementara subsektor perikanan relatif rendah, berkisar 95–100. Misalnya, Provinsi D.I. Yogyakarta memiliki nilai hortikultura sebesar 126,04 dan pangan sebesar 102,69, sedangkan perikanan hanya sebesar 91,45. Pola serupa terlihat pada Jawa Timur dengan nilai pangan sebesar 111,94 dan hortikultura sebesar 117,79, serta perikanan sebesar 99,14. Kondisi ini menunjukkan bahwa wilayah-wilayah tersebut lebih berorientasi pada pertanian darat, baik dalam bentuk hortikultura sayuran dan buah-buahan, maupun tanaman pangan seperti padi dan jagung. Faktor penentu utamanya adalah ketersediaan lahan subur, infrastruktur irigasi, dan tradisi pertanian intensif. Sebaliknya, kontribusi subsektor perikanan rendah karena minimnya garis pantai aktif atau rendahnya potensi perikanan tangkap, terutama di provinsi daratan seperti Yogyakarta dan Kalimantan Tengah. Dengan demikian, *cluster* ini menggambarkan wilayah agraris yang padat aktivitas pertanian darat namun relatif lemah di subsektor kelautan.
- c. Cluster 2 – Sangat kuat di pangan dan hortikultura, lemah di Perkebunan: Provinsi yang termasuk dalam *cluster* ini yaitu Provinsi Jawa Barat, Jawa Tengah, Nusa Tenggara Barat, dan Sulawesi Utara, Cluster ini memperlihatkan dominan subsektor tanaman pangan dan hortikultura, dengan nilai tertinggi di antara semua cluster (pangan dan hortikultura >115), namun subsektor perkebunan relatif lemah (<105). Misalnya, Jawa Tengah mencatat nilai tanaman pangan sebesar 114,63 dan hortikultura sebesar 115,13, sementara perkebunan hanya sebesar 105,41. Demikian pula, Nusa Tenggara Barat (NTB) memiliki nilai hortikultura sangat tinggi yaitu sebesar 137,6 dan pangan sebesar 117,78, tetapi perkebunan hanya sebesar 93,51. Hal ini menegaskan bahwa daerah-daerah dalam cluster ini mengandalkan tanaman pangan dan hortikultura sebagai sektor unggulan. Faktor penunjangnya meliputi: iklim dan kesuburan tanah yang ideal untuk hortikultura (sayuran dan buah tropis); kedekatan dengan pasar besar (jawa barat dan jawa tengah) sehingga produk hortikultura memiliki nilai ekonomi tinggi; dan orientasi produksi untuk konsumsi domestik dan distribusi antar wilayah. Namun, kelemahan di subsektor perkebunan menandakan keterbatasan lahan untuk tanaman tahunan atau komoditas ekspor (sawit, karet, kopi). Cluster ini menggambarkan wilayah dengan basis pertanian intensif dan berorientasi konsumsi dalam negeri, bukan ekspor bahan mentah.
- d. Cluster 3 – Dominasi perkebunan: Provinsi yang termasuk dalam *cluster* ini yaitu Provinsi Aceh, Sumatera Utara, Riau, Jambi, Kalimantan Barat, Kalimantan Utara, dan Sulawesi Barat. Cluster ini sangat jelas ditandai oleh tingginya nilai subsektor perkebunan yang berada pada rentang 133–170, jauh di atas subsektor lain. Misalnya, Riau memiliki nilai subsektor perkebunan sebesar 164,79 dan Kalimantan Utara bahkan mencapai 170,64. Sebaliknya, subsektor hortikultura dan peternakan berada pada kisaran 95–105, yang menunjukkan ketimpangan

subsektor. Kondisi ini sesuai dengan karakter wilayah yang memiliki lahan luas dan berorientasi ekspor komoditas perkebunan, seperti kelapa sawit, karet, kopi, dan kakao. Subsektor perikanan dan pangan berfungsi sebagai pendukung, bukan penggerak utama. Ciri utama cluster ini adalah: ketergantungan tinggi terhadap komoditas perkebunan ekspor, khususnya sawit dan karet; kontribusi besar terhadap PDRB sektor pertanian, namun disertai risiko tinggi terhadap fluktuasi harga global; ketimpangan subsektor yang menyebabkan kerentanan ekonomi apabila harga komoditas utama menurun. Dengan demikian, cluster ini menggambarkan wilayah agraris ekspor yang padat lahan, namun kurang terdiversifikasi.

4. Perbandingan dan Pembahasan

Untuk membandingkan performa kedua metode, digunakan ukuran evaluasi *Davies-Bouldin Index* (DBI), di mana nilai yang lebih rendah menunjukkan kualitas kluster yang lebih baik.

Tabel 3. Hasil Perbandingan K-Means dan Hierarchical

Metode	Jumlah Kluster	Davies-Bouldin Index
K-Means	4	0.17
Hierarchical	4	1.36

Hasil evaluasi performa algoritma menunjukkan bahwa metode *K-Means Clustering* memiliki kemampuan yang lebih baik dibandingkan dengan *Hierarchical Clustering* dalam membentuk struktur kluster subsektor pertanian di tingkat provinsi. Hal ini dibuktikan melalui perbandingan nilai *Davies-Bouldin Index* (DBI) yang digunakan sebagai ukuran kualitas kluster, di mana nilai yang lebih rendah mencerminkan kualitas kluster yang lebih baik. Berdasarkan hasil perhitungan, metode K-Means menghasilkan nilai DBI sebesar 0,17, sedangkan metode Hierarchical Clustering memiliki nilai 1,36. Nilai DBI yang rendah pada K-Means menunjukkan bahwa kluster yang terbentuk memiliki tingkat kompak (*intra cluster similarity*) yang tinggi serta pemisahan antar kluster (*inter cluster dissimilarity*) yang lebih jelas, sehingga secara teknis dapat dinyatakan bahwa struktur kluster yang dihasilkan oleh K-Means lebih optimal.

Keunggulan performa K-Means ini didukung oleh karakteristik algoritmanya yang mampu meminimalkan jarak antar anggota kluster terhadap centroid, sehingga distribusi data antar kluster menjadi lebih homogen. Dengan empat kluster yang terbentuk, hasil analisis memperlihatkan perbedaan yang jelas dalam struktur subsektor pertanian antar wilayah di Indonesia. Cluster 0 dan Cluster 3 cenderung berorientasi pada komoditas ekspor, khususnya subsektor perkebunan dan perikanan, sedangkan Cluster 1 dan Cluster 2 menunjukkan orientasi berbeda yang lebih menekankan pada ketahanan pangan domestik dan diversifikasi usaha tani. Temuan ini memperlihatkan adanya variasi spasial dalam pola pembangunan pertanian, di mana sebagian wilayah lebih menonjol pada subsektor berorientasi ekspor, sementara wilayah lain menguatkan sektor pangan dan hortikultura untuk kebutuhan dalam negeri.

Secara substantif, hasil klusterisasi ini memiliki nilai strategis dalam konteks perumusan kebijakan pembangunan pertanian berbasis wilayah dan penguatan daya saing regional berbasis kluster ekonomi (Porter, 1998; OECD, 2020; World Bank, 2021). Dengan mengidentifikasi provinsi-provinsi yang memiliki kesamaan karakteristik subsektor, pemerintah dapat mengarahkan kebijakan yang lebih spesifik dan berbasis potensi lokal. Misalnya, kluster dengan dominasi perkebunan dan perikanan dapat diarahkan pada penguatan rantai nilai ekspor melalui hilirisasi dan diversifikasi produk,

sedangkan klaster dengan dominasi tanaman pangan dan hortikultura dapat difokuskan pada peningkatan produktivitas, ketahanan pangan, dan sistem pertanian berkelanjutan.

Namun demikian, meskipun hasil DBI memberikan indikasi kuat tentang kualitas internal klaster yang terbentuk, terdapat keterbatasan penting terkait validitas dan reliabilitas hasil yang perlu diperhatikan secara kritis. Keterbatasan ini terletak pada minimnya analisis tambahan mengenai kestabilan atau konsistensi hasil clustering, misalnya melalui validasi silang (*cross validation*), analisis sensitivitas terhadap parameter awal, atau perbandingan antar algoritma dengan variasi jumlah klaster (k). Tanpa langkah validasi tersebut, terdapat risiko bahwa hasil klaster yang diinterpretasikan sebagai representasi struktur ekonomi pertanian regional justru merupakan artefak statistik yang muncul akibat konfigurasi data tertentu, bukan struktur alami yang stabil. Sebagai contoh, dalam metode K-Means, hasil pengelompokan sangat sensitif terhadap penentuan titik awal (*centroid*) dan jumlah klaster (k). Jika inisialisasi *centroid* berbeda atau data diubah sebagian (misalnya melalui *resampling*), hasil klasterisasi dapat berubah signifikan. Oleh karena itu, penggunaan satu metrik evaluasi tunggal seperti DBI, meskipun secara teknis valid, belum cukup menjamin reliabilitas empiris dari hasil pengelompokan. Analisis sensitivitas dan uji stabilitas diperlukan untuk memastikan bahwa struktur klaster tetap konsisten di bawah berbagai skenario atau variasi parameter.

Metode K-Means memiliki sensitivitas tinggi terhadap penentuan inisialisasi pusat klaster, sehingga pemilihan titik awal yang kurang tepat berpotensi menghasilkan solusi lokal yang bukan merupakan pengelompokan optimal. Selain itu, metode ini juga sangat dipengaruhi oleh proses normalisasi data; perbedaan skala antar variabel dapat mendominasi pembentukan klaster apabila standardisasi tidak dilakukan secara konsisten. K-Means juga mengasumsikan bahwa klaster bersifat sferis dan memiliki ukuran relatif serupa, sehingga kurang ideal ketika data memiliki bentuk klaster yang tidak beraturan atau mengandung outlier. Sementara itu, metode Hierarchical, meskipun lebih stabil terhadap inisialisasi, tetap memiliki kelemahan seperti sensitivitas terhadap noise dan kompleksitas komputasi yang lebih tinggi. Menyadari batasan ini memberikan pemahaman yang lebih seimbang mengenai keandalan hasil, serta menegaskan perlunya pengecekan sensitivitas dan validasi silang untuk memastikan stabilitas struktur klaster.

Berdasarkan hasil analisis terdapat implikasi kebijakan per klaster sebagai berikut:

Cluster 0 – Dominasi Perkebunan

Karakteristik: Bergantung pada subsektor perkebunan (sawit, karet, kopi) dengan kontribusi subsektor lain sedang.

Isu Lingkungan & Sosial: Risiko deforestasi tinggi, ketimpangan kepemilikan lahan, dan emisi karbon dari ekspansi perkebunan.

Arah Kebijakan:

1. Diversifikasi nilai tambah perkebunan melalui penguatan industri hilir (bioenergi, oleokimia, pupuk organik berbasis limbah perkebunan).
2. Investasi ramah lingkungan (*green investment*) dengan sertifikasi *sustainable agriculture* (RSPO, ISPO).
3. Program rehabilitasi lahan kritis dan penerapan *agroforestry* untuk menekan deforestasi.
4. Pemberdayaan petani kecil dan koperasi Perkebunan guna mengurangi ketimpangan distribusi pendapatan dan meningkatkan SDG 8 dan 12.

Cluster 1 – Kuat di Hortikultura dan Tanaman Pangan

Karakteristik: Didominasi oleh subsektor pangan dan hortikultura; menjadi pusat produksi pangan nasional.

Isu Lingkungan & Sosial: Degradasi tanah dan air akibat intensifikasi, kerentanan terhadap perubahan iklim.

Arah Kebijakan:

1. Program ketahanan pangan berkelanjutan melalui inovasi precision agriculture dan efisiensi penggunaan pupuk/pestisida.
2. Pembangunan infrastruktur cold-chain nasional untuk mengurangi food loss (mendukung SDG 12).
3. Manajemen air terpadu (integrated water resources management) untuk menghadapi risiko kekeringan dan banjir (SDG 13).
4. Pelibatan masyarakat lokal dan perempuan dalam koperasi pangan guna memperkuat ketahanan sosial ekonomi (SDG 2 dan 8).

Cluster 2 – Seimbang/Merata di Semua Subsektor

Karakteristik: Struktur subsektor pertanian relatif merata, menunjukkan keseimbangan sektor primer yang potensial.

Isu Lingkungan & Sosial: Kemiskinan pedesaan dan keterbatasan akses pasar bagi UMKM agroindustri.

Arah Kebijakan:

1. Model pembangunan agointegratif (pertanian, perikanan, peternakan) berbasis ekoregion untuk mendukung SDG 9 dan 12.
2. Pengembangan agroindustri desa dan UMKM pengolahan hasil pertanian untuk meningkatkan nilai tambah lokal.
3. Program penguatan pasar digital dan rantai pasok inklusif agar wilayah ini menjadi buffer zone pangan nasional.
4. Integrasi indikator sosial (kemiskinan pedesaan, ketimpangan wilayah) dalam perencanaan klaster agar tercapai pertumbuhan yang inklusif (SDG 8).

Cluster 3 – Lemah di Pangan dan Hortikultura, Kuat di Perikanan dan Perkebunan

Karakteristik: Suboptimal di pangan-hortikultura, namun kuat di subsektor perikanan (tangkap/budidaya) dan perkebunan.

Isu Lingkungan & Sosial: Tekanan ekosistem pesisir, polusi perairan, serta keterbatasan lahan produktif di perkotaan.

Arah Kebijakan:

1. Investasi cold-chain dan logistik maritim untuk efisiensi rantai pasok hasil laut dan pengurangan kerusakan produk (SDG 9 dan 12).
2. Integrasi perikanan dengan industri pengolahan hasil laut (surimi, fillet, rumput laut olahan) untuk memperkuat marine-based economy.
3. Program ketahanan pangan berbasis lahan terbatas (urban farming, hidroponik, aquaponik) di kawasan padat penduduk.
4. Rehabilitasi ekosistem pesisir dan pengelolaan perikanan berkelanjutan guna menekan emisi dan menjaga stok ikan (SDG 13).

KESIMPULAN DAN SARAN

Hasil analisis *clustering* menunjukkan bahwa struktur subsektor pertanian antar wilayah di Indonesia sangat beragam. Metode K-Means terbukti lebih optimal dibandingkan metode Hierarchical, berdasarkan nilai *Davies-Bouldin Index* yang lebih rendah (0.17). Keempat klaster yang terbentuk mencerminkan diferensiasi wilayah berdasarkan subsektor dominan maupun kelemahannya.

Temuan ini mendukung pentingnya perencanaan pembangunan berbasis klaster wilayah yang adaptif, inklusif, dan selaras dengan Tujuan Pembangunan Berkelanjutan (SDGs), terutama ketahanan pangan dan pertanian berkelanjutan (SDG 2), pertumbuhan ekonomi inklusif dan produktivitas tenaga kerja (SDG 8), pembangunan infrastruktur dan

industrialisasi berkelanjutan (SDG 9), konsumsi dan produksi berkelanjutan (SDG 12), serta aksi terhadap perubahan iklim (SDG 13).

Perspektif lanjutan dan penelitian ke depan untuk memperkuat relevansi kebijakan dan keakuratan model:

1. Menggunakan data panel multi tahun agar dapat memantau dinamika temporal struktur ekonomi sektor primer dan transisi antar kluster.
2. Mengintegrasikan analisis spasial berbasis GIS dengan peta kluster berwarna dan legenda yang jelas, untuk mendukung perencanaan wilayah dan evidence-based policymaking.
3. Memasukkan indikator lingkungan dan sosial (emisi karbon, deforestasi, kemiskinan, ketimpangan) dalam model klusterisasi berikutnya, agar analisis lebih berkelanjutan secara multidimensi.
4. Mengembangkan model hibrid (CLR, PCA, K-Means) untuk meningkatkan robustness hasil kluster terhadap outlier dan multikolinearitas variabel subsektor.

DAFTAR PUSTAKA

- Badan Pusat Statistik. (2024). Statistik nilai tukar petani 2024. *Badan Pusat Statistik*.
- Bappenas. (2020). Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2020–2024. *Kementerian Perencanaan Pembangunan Nasional*.
- Davies, D. L., & Bouldin, D. W. (1979). A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-1*(2), 224–227. <https://doi.org/10.1109/TPAMI.1979.4766909>
- Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). Wiley.
- FAO. (2021). The state of food and agriculture 2021. *Food and Agriculture Organization of the United Nations*.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.
- Hastie, T., Tibshirani, R., & Friedman, J. (2021). *The Elements of Statistical Learning* (2nd ed.). Springer.
- IDSS. (2025). Cluster Analysis and Validation Using K-Means and Hierarchical Methods. *Institute for Digital Statistical Studies*.
- 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>
- Johnson, R. A., & Wichern, D. W. (2018). *Applied Multivariate Statistical Analysis* (7th ed.). Pearson.
- Kaufman, L., & Rousseeuw, P. J. (2009). *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley.
- Kementerian Pertanian Republik Indonesia. (2023). Outlook sektor pertanian Indonesia 2023. *Kementerian Pertanian RI*.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook* (3rd ed.). Sage Publications.
- OECD. (2020). Regional Development Policies in OECD Countries. *OECD Publishing*.
- Porter, M. E. (1998). Clusters and the New Economics of Competition. *Harvard Business Review Press*.
- Rodríguez-Pose, A., Crescenzi, R., & Di Cataldo, M. (2012). Social and Economic Regional Disparities in the European Union. *Environment and Planning A: Economy and Space, 44*(6), 1273–1291. <https://doi.org/10.1068/a444496>
- 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)
- Todaro, M. P., & Smith, S. C. (2021). *Economic Development* (13th ed.). Pearson.
- United Nations. (2015). Transforming Our World: The 2030 Agenda for Sustainable Development. *United Nations*.
- Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association, 58*(301), 236–244. <https://doi.org/10.1080/01621459.1963.10500845>
- Widiyanto, A., Nugroho, S., & Prasetyo, E. (2024). Comparative Analysis of K-Means and Hierarchical Clustering Methods in Regional Economic Classification. *Journal of Data Science and Regional Analysis, 8*(1), 45–57.
- World Bank. (2021). Indonesia economic prospects: Boosting the recovery. *World Bank*.

Xu, R., & Wunsch, D. (2009). Clustering. *Wiley-IEEE Press*.

Zhang, T., Ramakrishnan, R., & Livny, M. (1996). BIRCH: An Efficient Data Clustering Method for Very Large Databases. *ACM SIGMOD Record*, 25(2), 103-114. <https://doi.org/10.1145/235968.233324>