

Analisis Spasial dan Klasterisasi Ketimpangan Aset Koperasi di Tingkat Kabupaten/Kota Provinsi Jawa Barat

Ery Supriyadi Rustidja¹, Udin Hidayat²

Universitas Koperasi Indonesia^{1,2}

erysr@ikopin.ac.id²

ABSTRAK

Koperasi memegang peranan krusial dalam menopang stabilitas ekonomi akar rumput di tingkat daerah. Namun, sebaran modal yang tercermin dari nilai total aset koperasi antar-wilayah seringkali mengalami ketimpangan yang lebar, sehingga memerlukan sistem pemetaan yang objektif untuk mendukung pengambilan kebijakan. Penelitian ini bertujuan untuk mengelompokkan 27 kabupaten/kota di Provinsi Jawa Barat berdasarkan karakteristik kemiripan total aset koperasi menggunakan pendekatan sains data. Metode penelitian yang diterapkan adalah kuantitatif berbasis data mining dengan membandingkan dua arsitektur algoritma clustering, yaitu K-Means Clustering sebagai perwakilan metode non-hierarki dan Agglomerative Hierarchical Clustering dengan Ward's Method sebagai metode hirarki. Data sekunder berupa total aset koperasi berskala satu dimensi diperoleh dari portal resmi Open Data Jabar. Penentuan jumlah kelompok optimal ditentukan melalui Elbow Method, sedangkan pengujian keandalan struktur kluster dievaluasi menggunakan metrik validasi internal Silhouette Coefficient. Hasil analisis grafik Elbow secara konsisten menunjukkan bahwa jumlah partisi paling efisien untuk dataset ini adalah tiga kluster ($SK=3\$$). Melalui tahapan evaluasi model, algoritma Hierarchical Clustering (Metode Ward) terbukti menghasilkan performa terbaik dengan rata-rata skor Silhouette sebesar 0,665, mengungguli algoritma K-Means yang mencatatkan skor 0,587. Metode Ward dinilai lebih andal dalam memecah data satu dimensi yang memiliki varians timpang tanpa terdistorsi penuh oleh pencilan ekstrem. Hasil pemetaan terbaik mengisolasi Kota Bandung sebagai kluster mandiri tunggal dengan aset super tinggi (Rp 5,15 Triliun), diikuti oleh empat wilayah penyangga industri pada kluster aset tinggi (Rp 1,61 T – Rp 2,94 T), serta 22 kabupaten/kota lainnya yang mendominasi kelompok aset sedang-rendah (di bawah Rp 1,30 Triliun). Temuan ini mengonfirmasi adanya polarisasi ekonomi yang nyata pada ekosistem koperasi di Jawa Barat. Implikasi dari riset ini merekomendasikan pemerintah daerah untuk menerapkan kebijakan pembinaan terarah berbasis kluster (data-driven policy), di mana intervensi bantuan modal stimulan dan pendampingan manajemen diutamakan secara masif pada 22 wilayah berkembang di kluster terendah.

Kata Kunci: Koperasi, Total Aset, Clustering, K-Means, Ward's Method

I. PENDAHULUAN

Koperasi merupakan salah satu pilar utama dalam sistem perekonomian Indonesia yang berfungsi untuk meningkatkan kesejahteraan anggota pada khususnya dan masyarakat pada umumnya (Undang-Undang Nomor 25 Tahun 1992 Tentang Perkoperasian). Di Provinsi Jawa Barat, perkembangan koperasi menunjukkan dinamika yang sangat masif, baik dari segi kuantitas institusi maupun akumulasi modal usaha. Salah satu indikator finansial paling krusial yang mencerminkan skala ekonomi, kapasitas operasional, dan tingkat kepercayaan anggota terhadap suatu koperasi adalah total aset. Informasi mengenai total aset koperasi di tingkat kabupaten/kota menjadi modal data yang sangat berharga bagi pemangku kebijakan untuk memetakan kekuatan ekonomi daerah.

Namun demikian, sebaran total aset koperasi antar-wilayah di Jawa Barat menunjukkan adanya indikasi ketimpangan yang cukup signifikan. Beberapa wilayah metropolitan atau pusat industri memiliki akumulasi aset koperasi yang menembus angka triliunan rupiah, sementara beberapa kabupaten/kota lainnya masih berada di skala miliaran rupiah. Fenomena

kesenjangan ekonomi regional ini jika tidak diidentifikasi dengan tepat akan menghambat efektivitas program pembinaan dan penyaluran insentif dari pemerintah, karena kebijakan yang bersifat *one-size-fits-all* menjadi tidak relevan (Sugiastuti & Pratama, 2022). Oleh karena itu, diperlukan sebuah mekanisme pengelompokan (*clustering*) wilayah yang objektif dan terukur berdasarkan karakteristik total aset tersebut.

Dalam domain sains data, pengelompokan data satu dimensi (*1-Dimensional data*) seperti total aset per wilayah umumnya diselesaikan menggunakan algoritma *Unsupervised Learning*. Algoritma *K-Means* merupakan pendekatan yang paling populer digunakan karena efisiensinya dalam mengelompokkan data berdasarkan jarak terdekat ke pusat massa (*centroid*) (Jain, 2010). Meski demikian, data finansial regional seringkali memiliki karakteristik sebaran yang tidak merata dan rentan terhadap nilai ekstrem (*outlier*) (Wu et al., 2008). Penggunaan *K-Means* secara tunggal tanpa adanya komparasi seringkali menghasilkan partisi kelompok yang kurang optimal atau bias terhadap wilayah beraset sangat besar.

Penelitian terdahulu mengenai pemetaan koperasi mayoritas masih menggunakan pendekatan statistik deskriptif atau pengelompokan manual berdasarkan ambang batas (*threshold*) tertentu yang bersifat subjektif (Everitt et al., 2011). Keterbatasan tersebut membuka *research gap* mengenai bagaimana menguji keandalan berbagai algoritma sains data dalam menangani data komparatif wilayah berskala satu dimensi. Untuk mengatasi keterbatasan *K-Means* dalam menangani varians data yang timpang (Aggarwal & Reddy, 2018), pendekatan *Hierarchical Clustering* dengan *Ward's Method* dapat diterapkan sebagai pembanding guna melihat struktur kedekatan wilayah secara bertahap melalui representasi *dendrogram* (Murtagh & Legendre, 2014).

II. TINJAUAN PUSTAKA

A. Koperasi

Koperasi merupakan badan usaha yang beranggotakan orang-orang atau badan hukum koperasi dengan melandaskan kegiatannya berdasarkan prinsip koperasi sekaligus sebagai gerakan ekonomi rakyat yang berdasar atas asas kekeluargaan. Dalam perspektif ekonomi regional, entitas koperasi tidak hanya berfungsi sebagai lembaga keuangan mikro (Tambunan, 2017), melainkan juga sebagai indikator stabilitas ekonomi akar rumput di suatu daerah (Birchall, 2013; Spear & Bidet, 2005).

Total aset merupakan indikator agregat finansial yang mencerminkan keseluruhan sumber daya ekonomi yang dimiliki atau dikendalikan oleh koperasi, yang mencakup aset lancar, aset tetap, serta investasi jangka Panjang (Subramanyam & Wild, 2014). Nilai total aset mencerminkan skala kapitalisasi, kapasitas ekspansi usaha, tingkat likuiditas, serta kapasitas fungsional lembaga keuangan tersebut dalam menyerap risiko ekonomi di suatu wilayah. Pemetaan total aset secara spasial dan makro penting dilakukan untuk mengidentifikasi kutub-kutub pertumbuhan (*growth poles*), konsentrasi modal, serta wilayah-wilayah yang mengalami stagnasi atau keteringgalan pertumbuhan ekonomi (disparitas regional) (Krugman, 1991).

B. *Unsupervised Learning* dan Karakteristik Data 1-Dimensi

Clustering atau klasterisasi merupakan salah satu implementasi utama dari paradigma *Unsupervised Learning* dalam domain sains data. Berbeda dengan *Supervised Learning* yang

mengandalkan *label target (ground truth)* untuk proses pelatihan model, *Unsupervised Learning* mengeksplorasi struktur intrinsik dari kumpulan data mentah untuk menemukan pola, kemiripan, atau kelompok tersembunyi secara mandiri berdasarkan karakteristik kedekatan fitur yang dimiliki (Bishop, 2009; Hastie et al., 2017).

Karakteristik data satu dimensi (1D), seperti data total aset wilayah tanpa adanya variabel pendukung lainnya, memiliki keunikan dan tantangan tersendiri dalam komputasi. Pada data multi-dimensi, algoritma mencari pola geometris di dalam ruang hiper-dimensi (*hyperspace*) melalui kombinasi berbagai fitur. Sebaliknya, pada data 1-Dimensi, ruang pencarian bersifat *linear* dan terbatas pada kerapatan objek di sepanjang garis bilangan numerik tunggal. Tantangan utama klasterisasi 1D adalah menetapkan titik potong yang optimal tanpa memicu terjadinya *trivially sorted groups* (sekadar mengurutkan data lalu membaginya sama rata). Klasterisasi 1D berbasis sains data bertujuan untuk meminimalkan varians internal di dalam kelompok (*within-cluster variance*) dan memaksimalkan jarak pemisah antar-kelompok (*between-cluster variance*). Pendekatan ini secara ilmiah menggantikan pembagian klasifikasi konvensional yang bersifat subjektif, seperti penentuan ambang batas (*threshold*) manual yang rawan bias menjadi sistem pembagian yang berbasis karakteristik data aktual (*data-driven decision*) (Cao, 2010).

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. 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.

$$d = \sqrt{\sum (x_i - c_i)^2} \quad (1)$$

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*.

E. *Hierarchical Clustering (Ward's Method)*

Berbeda dengan *K-Means* yang membagi data secara langsung ke dalam sejumlah *K* kelompok yang kaku, *Hierarchical Clustering* membangun struktur kelompok secara bertahap dan berpola menyerupai pohon (*tree-like structure*). Penelitian ini menggunakan pendekatan *Agglomerative Hierarchical Clustering*, yaitu metode *bottom-up* di mana proses komputasi dimulai dengan menganggap setiap objek data tunggal sebagai satu klaster mandiri yang terpisah. Secara berurutan dan berulang, algoritma akan mencari pasangan klaster yang memiliki tingkat kemiripan paling tinggi untuk digabungkan menjadi satu klaster baru yang lebih besar, hingga pada tahap akhir seluruh objek menyatu ke dalam satu klaster tunggal raksasa.

Untuk menentukan klaster mana yang paling layak digabungkan pada setiap iterasi, penelitian ini menerapkan *Ward's Method* (Kriteria Varian Minimum). Metode Ward memiliki karakteristik yang berbeda dengan metode tautan tradisional yang mengukur jarak geometris antar-titik terluar (Murtagh & Legendre, 2014). Metode Ward didasarkan pada pendekatan *Analysis of Variance* (ANOVA) yang bertujuan untuk meminimalkan kenaikan total jumlah kuadrat kesalahan (*Sum of Squares Error / SSE*) di dalam internal klaster pada setiap tahapan penggabungan. Secara matematis, rumus kenaikan varians atau jarak antar-klaster saat menggabungkan dua klaster *A* dan *B* dinyatakan sebagai berikut:

$$\Delta SSE_{AB} = \frac{n_A n_B}{n_A + n_B} ||c_A - c_B||^2 \quad (2)$$

Di mana n_A dan n_B adalah jumlah objek data yang terdapat di dalam klaster *A* dan *B*, sedangkan c_A dan c_B berturut-turut adalah nilai titik pusat (*centroid*) dari masing-masing klaster tersebut.

Kelebihan utama dari penggunaan Metode Ward dalam data ekonomi 1-Dimensi adalah kemampuannya menghasilkan struktur kelompok dengan ukuran yang cenderung seimbang dan memiliki varians internal yang sangat rapat (Murtagh & Contreras, 2012). Hasil akhir dari urutan penggabungan ini dipetakan secara visual ke dalam bentuk diagram pohon hirarki yang disebut *dendrogram*. Melalui *dendrogram*, peneliti dapat menganalisis tingkatan hubungan kekerabatan antar-wilayah secara fleksibel dan menentukan titik potong (*cutting point*) klaster terbaik secara objektif (Legendre & Legendre, 2012; Yim & Ramdeen, 2015).

F. *Metrik Validasi Klaster*

Karena algoritma klasterisasi termasuk dalam kategori *Unsupervised Learning*, data yang digunakan tidak memiliki label asli (*ground truth*) sebagai acuan pembandingan. Oleh karena itu, pengujian performa model tidak bisa menggunakan metrik akurasi konvensional. Sebagai gantinya, diperlukan metrik validasi internal yang mengevaluasi kualitas klaster murni berdasarkan struktur matematis data hasil pengelompokan.

1. *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 2.

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

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

2. *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 3.

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

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* dan *Hierarchical Clustering* (Metode Ward). Pendekatan kuantitatif digunakan karena data yang dianalisis berupa data numerik tunggal (*one-dimensional data*) yang terdiri dari variabel total aset koperasi.

Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Provinsi Jawa Barat berdasarkan kemiripan karakteristik total aset, sehingga dapat diketahui kelompok wilayah dengan aset koperasi berkategori tinggi, sedang, dan rendah secara objektif berbasis data (*data-driven*).

Jenis data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh secara legal dari portal resmi Pemerintah Provinsi Jawa Barat, yaitu Open Data Jabar (Dinas Koperasi dan Usaha Kecil Provinsi Jawa Barat, 2026). Data yang digunakan berupa data total aset koperasi (Rp) yang diakumulasikan per wilayah administratif kabupaten/kota di Provinsi Jawa Barat. Data yang dianalisis difokuskan pada periode pencatatan terbaru untuk mencerminkan kondisi terkini sebaran modal koperasi di Provinsi Jawa Barat.

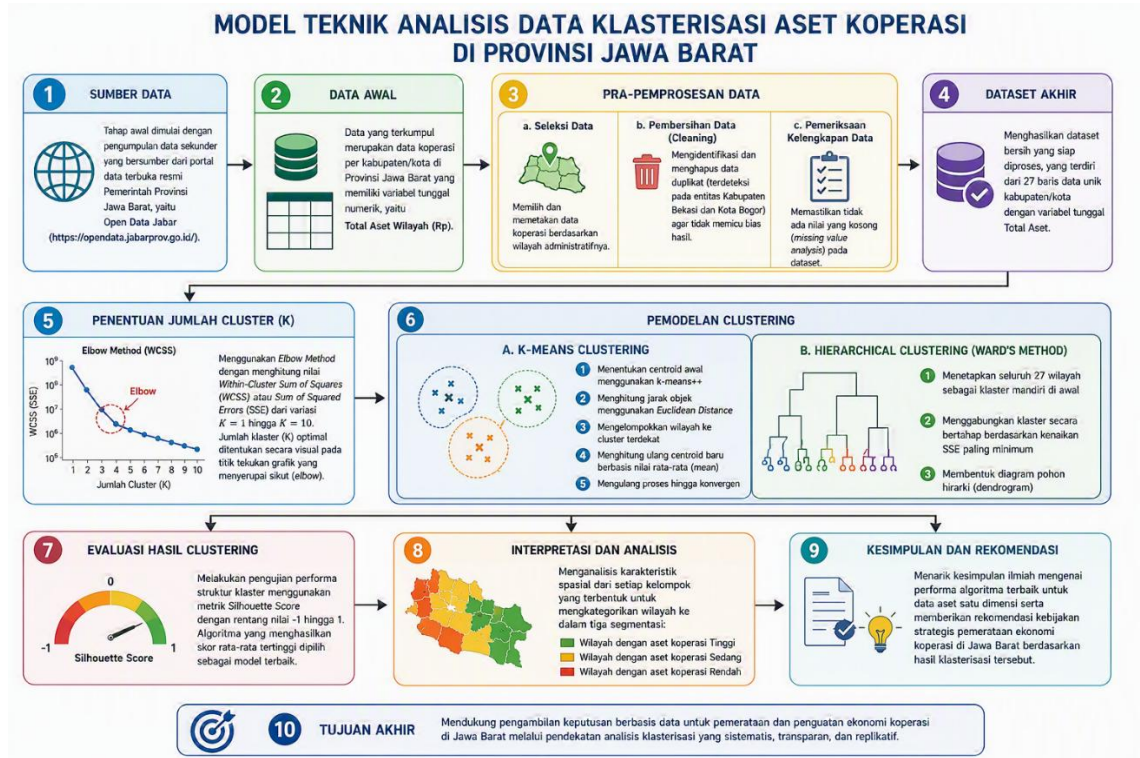
Objek dalam penelitian ini adalah kabupaten/kota di Provinsi Jawa Barat, yang terdiri dari 27 wilayah administratif administratif (18 kabupaten dan 9 kota), di antaranya Kabupaten Bandung, Kabupaten Bandung Barat, Kabupaten Bekasi, Kabupaten Bogor, Kabupaten Ciamis, Kabupaten Cianjur, Kabupaten Cirebon, Kabupaten Garut, Kabupaten Indramayu, Kabupaten Karawang, Kabupaten Kuningan, Kabupaten Majalengka, Kabupaten Pangandaran, Kabupaten Purwakarta, Kabupaten Subang, Kabupaten Sukabumi, Kabupaten Sumedang, Kabupaten Tasikmalaya, Kota Bandung, Kota Banjar, Kota Bekasi, Kota Bogor, Kota Cimahi, Kota Cirebon, Kota Depok, Kota Sukabumi, dan Kota Tasikmalaya. Setiap kabupaten/kota diperlakukan sebagai satu unit data yang dianalisis berdasarkan variabel tunggal total aset.

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

Tabel 1. Operasionalisasi Variabel Penelitian

Variabel	Keterangan	Skala Data
Total Aset	Akumulasi nilai seluruh sumber daya ekonomi/modal yang dikelola oleh koperasi di suatu kabupaten/kota (Rp)	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 dan *Preprocessing Data*

Dataset yang diperoleh secara resmi dari portal Open Data Jabar memuat data total aset koperasi di tingkat kabupaten/kota se-Provinsi Jawa Barat yang mencakup 27 wilayah administratif unik (18 kabupaten dan 9 kota). Karena data yang diperoleh telah terverifikasi bersih dan bebas dari kendala *missing values* maupun redundansi input, tahapan pra-pemrosesan difokuskan langsung pada analisis karakteristik sebaran data menggunakan statistik deskriptif.

Total aset koperasi di Provinsi Jawa Barat menunjukkan tingkat ketimpangan atau varians yang sangat ekstrem pada skala linear. Nilai tertinggi dicatatkan oleh Kota Bandung dengan total aset mencapai Rp 5.151.322.872.310, sedangkan nilai terendah berada di Kabupaten Pangandaran dengan total aset sebesar Rp 66.020.484.030. Jarak rentang yang masif dan mencapai triliunan rupiah ini menjadi indikator awal adanya polarisasi ekonomi koperasi yang nyata di Jawa Barat, sehingga memperkuat urgensi dilakukannya pengelompokan berbasis algoritma sains data. Lebih jelasnya, total aset koperasi masing-masing kota/kabupaten dapat dilihat pada Tabel 2.

Tabel 2. Total Aset Koperasi Masing-masing Kota/Kabupaten di Jawa Barat

Kota/Kabupaten	Total Aset (Rp.)
Kabupaten Bandung	1,235,275,371,362
Kabupaten Bandung Barat	556,994,646,849
Kabupaten Bekasi	1,612,614,866,613
Kabupaten Bogor	2,155,545,387,137
Kabupaten Ciamis	440,224,256,983
Kabupaten Cianjur	2,632,919,889,302
Kabupaten Cirebon	768,841,521,890
Kabupaten Garut	518,631,359,367
Kabupaten Indramayu	684,297,379,922
Kabupaten Karawang	916,741,225,449
Kabupaten Kuningan	426,880,150,185
Kabupaten Majalengka	217,390,966,768
Kabupaten Pangandaran	66,020,484,030
Kabupaten Purwakarta	566,108,375,203
Kabupaten Subang	357,639,233,016
Kabupaten Sukabumi	314,679,960,635
Kabupaten Sumedang	1,066,150,308,455
Kabupaten Tasikmalaya	311,353,073,966
Kota Bandung	5,151,322,872,310
Kota Banjar	91,026,401,303
Kota Bekasi	2,947,101,161,777
Kota Bogor	557,004,879,103
Kota Cimahi	382,392,668,532
Kota Cirebon	1,077,058,623,092
Kota Depok	1,308,085,319,197
Kota Sukabumi	121,787,697,841
Kota Tasikmalaya	1,064,276,237,740

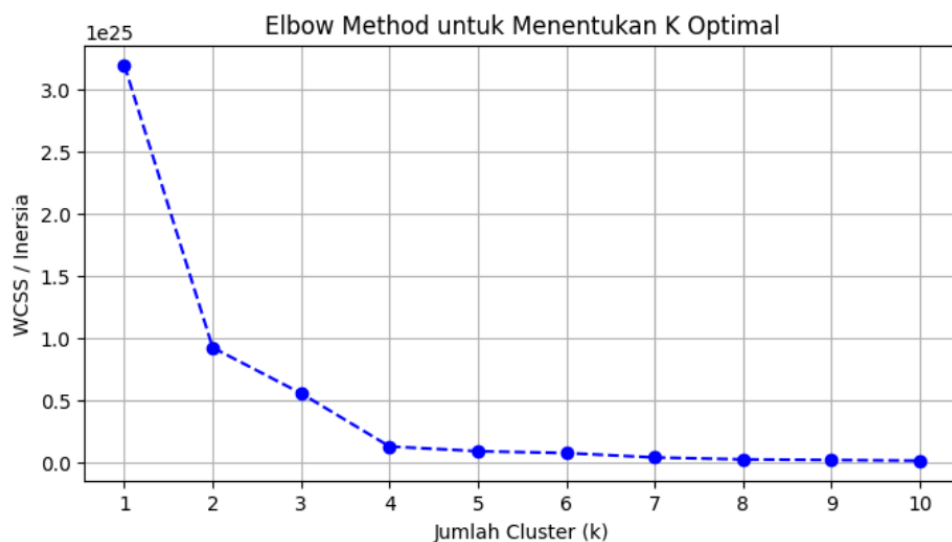
Sumber: Open Data Jawa Barat, data diolah (2026)

B. *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*.

C. 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 3. Hasil *K-Means Clustering*

Cluster	Wilayah
0	Kota Bandung, Kota Bekasi, Kabupaten Cianjur
1	Kabupaten Bogor, Kabupaten Bekasi, Kota Depok, Kabupaten Bandung, Kota Cirebon, Kabupaten Sumedang, Kota Tasikmalaya, Kabupaten Karawang
2	Kabupaten Cirebon, Kabupaten Indramayu, Kabupaten Purwakarta, Kota Bogor, Kabupaten Bandung Barat, Kabupaten Garut, Kabupaten Ciamis, Kabupaten Kuningan, Kota Cimahi, Kabupaten Subang, Kabupaten Sukabumi, Kabupaten Tasikmalaya, Kabupaten Majalengka, Kota Sukabumi, Kota Banjar, Kabupaten Pangandaran

Sumber: Data diolah (2026)

Analisis komputasi *K-Means* menunjukkan bahwa algoritma berbasis *centroid* rata-rata ini cenderung membagi data berdasarkan rentang jarak global yang proporsional. Batas bawah

untuk Cluster 0 ditarik pada angka Rp 2,63 Triliun, sementara batas bawah Cluster 1 ditahan oleh Kabupaten Karawang di angka Rp 916 Miliar.

D. Hasil Klasterisasi *Hierarchical Clustering (Ward's Method)*

Sebagai pembanding bagi algoritma berbasis partisi langsung, dilakukan pemodelan menggunakan *Agglomerative Hierarchical Clustering* dengan *Ward's Linkage*. Keunggulan metode Ward yang meminimalkan peningkatan varians internal (*internal variance*) menghasilkan pola partisi kelompok yang secara matematis berbeda dengan *K-Means*. Hasil *clustering* dipaparkan pada Tabel 4.

Tabel 4. Hasil *Hierarchical Clustering*

Cluster	Wilayah
0	Kota Bandung
1	Kota Bekasi, Kabupaten Cianjur, Kabupaten Bogor, Kabupaten Bekasi
2	Kota Depok, Kabupaten Bandung, Kota Cirebon, Kabupaten Sumedang, Kota Tasikmalaya, Kabupaten Karawang, Kabupaten Cirebon, Kabupaten Indramayu, Kabupaten Purwakarta, Kota Bogor, Kabupaten Bandung Barat, Kabupaten Garut, Kabupaten Ciamis, Kabupaten Kuningan, Kota Cimahi, Kabupaten Subang, Kabupaten Sukabumi, Kabupaten Tasikmalaya, Kabupaten Majalengka, Kota Sukabumi, Kota Banjar, Kabupaten Pangandaran

Sumber: Data diolah (2026)

E. Analisis Komparasi dan Evaluasi Performa Model

Berdasarkan visualisasi hasil akhir, terdapat perbedaan pola pengelompokan yang kontras antara *K-Means* (Tabel 3) dan Metode Ward (Tabel 4). Pada algoritma *K-Means*, Kota Bandung digabungkan bersama dengan Kota Bekasi dan Kabupaten Cianjur ke dalam satu kelompok (Cluster Tinggi). Namun, pada Metode Ward, Kota Bandung diisolasi sendirian di Cluster 0, sedangkan Kota Bekasi dan Kabupaten Cianjur ditarik ke bawah untuk bergabung dengan Kabupaten Bogor dan Kabupaten Bekasi di Cluster 1.

Fenomena ini membuktikan secara teoritis sifat Metode Ward yang sangat sensitif dalam menjaga kerapatan varians (*Sum of Squares Error*). Karena jarak antara Kota Bandung (Rp 5,15 T) dan Kota Bekasi (Rp 2,94 T) terlampaui lebar dengan *gap* mencapai Rp 2,21 Triliun, algoritma Ward menilai bahwa memaksakan penggabungan kedua wilayah tersebut akan merusak homogenitas varians internal kelompok. Oleh karena itu, Metode Ward memilih mengisolasi Kota Bandung sebagai klaster mandiri tunggal.

Untuk menentukan secara objektif algoritma mana yang menghasilkan kualitas partisi terbaik pada data finansial satu dimensi ini, dilakukan pengujian metrik *Silhouette Coefficient*. Hasil komparasi metrik validasi internal ditunjukkan pada Tabel 5.

Tabel 5. Perbandingan Nilai *Silhouette Score Model*

Algoritma Klasterisasi	Rata-rata <i>Silhouette Score</i>	Keterangan Struktur Klaster
K-Means Clustering	0,587	Cukup Baik, namun rentan bias terhadap batas pengelompokan linear akibat pencilan
Hierarchical Clustering (Ward)	0,665	Baik, memiliki struktur kelompok yang lebih padat (<i>higher cohesion</i>) dan batas partisi yang lebih alami

Sumber: data diolah, 2026

Berdasarkan hasil pengujian pada Tabel 5, algoritma *Hierarchical Clustering* dengan *Ward's Method* menghasilkan skor *Silhouette* tertinggi sebesar 0,665. Perolehan skor ini membuktikan secara empiris bahwa metode hirarki lebih unggul dan kokoh untuk melakukan klasterisasi data finansial satu dimensi yang memiliki tingkat ketimpangan tinggi dan mengandung pencilan ekstrem.

Secara komputasi, nilai 0,665 menunjukkan bahwa struktur kelompok yang dibentuk oleh metode Ward memiliki pemisahan antar-klaster (*separation*) yang lebih baik dan konsisten pada garis bilangan tunggal. *K-Means* dengan skor 0,587 cenderung memaksakan pembagian berbasis titik tengah linear (*centroid*), yang menyebabkannya kurang optimal ketika harus menarik batas kelompok pada data yang memiliki sebaran tidak merata (*skewed data*).

Dari sisi interpretasi substantif, hasil dari *Hierarchical Clustering* (Tabel 4) direkomendasikan sebagai model utama yang paling representatif untuk digunakan oleh instansi pemerintah sebagai dasar pengambilan keputusan berbasis data (*data-driven policy*) dalam program intervensi modal dan pembinaan ekosistem koperasi di Provinsi Jawa Barat.

V. KESIMPULAN

Berdasarkan hasil analisis dan eksperimen komputasi terhadap data total aset koperasi di 27 kabupaten/kota Provinsi Jawa Barat, penelitian ini berhasil merumuskan beberapa kesimpulan penting. Pengujian menggunakan *Elbow Method* secara konsisten menunjukkan bahwa jumlah kelompok paling optimal dan efisien untuk mempartisi data numerik satu dimensi ini adalah tiga klaster. Melalui pengujian metrik validasi internal *Silhouette Coefficient*, algoritma *Hierarchical Clustering* dengan *Ward's Method* terbukti menghasilkan performa model terbaik dengan skor sebesar 0,665, yang secara signifikan mengungguli algoritma *K-Means Clustering* dengan skor 0,587. Metode Ward terbukti jauh lebih andal dan rigid dalam menangani karakteristik data finansial regional yang memiliki tingkat ketimpangan tinggi serta mengandung nilai pencilan ekstrem.

Hasil pemetaan terbaik dari metode hirarki ini berhasil membagi wilayah Jawa Barat secara alami ke dalam tiga tingkatan ekonomi koperasi yang berbeda. Cluster 0 dengan kategori aset super tinggi diisi oleh Kota Bandung secara tunggal dengan total aset mencapai Rp 5,15 Triliun. Selanjutnya, Cluster 1 dengan kategori aset tinggi diisi oleh empat wilayah industri dan penyangga utama, yaitu Kota Bekasi, Kabupaten Cianjur, Kabupaten Bogor, dan Kabupaten Bekasi dengan rentang aset antara Rp 1,61 Triliun hingga Rp 2,94 Triliun. Sementara itu, Cluster 2 menjadi kelompok mayoritas yang menampung 22 kabupaten/kota lainnya dengan skala aset di bawah Rp 1,30 Triliun. Dominasi jumlah wilayah di Cluster 2 ini

mengonfirmasi adanya pola polarisasi ekonomi dan kesenjangan ruang yang nyata pada ekosistem koperasi di Provinsi Jawa Barat.

Bertitik tolak dari temuan objektif tersebut, beberapa rekomendasi strategis dapat diajukan untuk meningkatkan efisiensi tata kelola ekonomi daerah. Dinas Koperasi dan UMKM atau instansi terkait disarankan untuk mulai menerapkan kebijakan pembinaan yang berbasis karakteristik klaster (*data-driven policy*) guna menggantikan sistem umum yang kurang merata. Wilayah yang berada di Cluster 0 dan Cluster 1 dapat diarahkan pada program modernisasi, digitalisasi masif, serta ekspansi pasar berskala nasional. Sebaliknya, 22 wilayah di Cluster 2 harus dijadikan prioritas utama untuk menerima intervensi bantuan modal stimulan, restrukturisasi organisasi, serta pendampingan manajemen bisnis dasar. Bagi penelitian selanjutnya, keterbatasan variabel tunggal dalam riset ini dapat dikembangkan dengan menambahkan dimensi fitur baru seperti jumlah anggota aktif, total omset usaha, dan Sisa Hasil Usaha (SHU) agar hasil pengelompokan koperasi di masa mendatang menjadi lebih komprehensif.

DAFTAR PUSTAKA

- Aggarwal, C. C., & Reddy, C. K. (2018). *Data Clustering* (C. C. Aggarwal & C. K. Reddy, Eds.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315373515>
- Birchall, Johnston. (2013). *Resilience in a downturn : the power of financial cooperatives*. International Labour Office.
- Bishop, C. M. . (2009). *Pattern recognition and machine learning*. Springer Science + Business Media.
- Cao, Longbing. (2010). *Domain driven data mining*. Springer.
- Dinas Koperasi dan Usaha Kecil Provinsi Jawa Barat. (2026). *Jumlah Asset Koperasi Berdasarkan Kabupaten/Kota di Jawa Barat*. <https://opendata.jabarprov.go.id>.
- Everitt, Brian., Landau, Sabine., Leese, Morven., & Stahl, Daniel. (2011). *Cluster Analysis, 5th Edition*. John Wiley & Sons.
- Han, Jiawei., Kamber, Micheline., & Pei, Jian. (2012). *Data mining : concepts and techniques*. Elsevier/Morgan Kaufmann.
- Hastie, Trevor., Tibshirani, Robert., & Friedman, J. H. . (2017). *The elements of statistical learning : data mining, inference, and prediction*. Springer.
- 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>
- Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy*, 99(3), 483–499. <http://www.jstor.org/stable/2937739>
- Larose, D. T. ., & Larose, C. D. . (2014). *Discovering knowledge in data : an introduction to data mining*. Wiley.
- Legendre, Pierre., & Legendre, Louis. (2012). *Numerical ecology*. Elsevier.

- Murtagh, F., & Contreras, P. (2012). Algorithms for hierarchical clustering: an overview. *WIREs Data Mining and Knowledge Discovery*, 2(1), 86–97. <https://doi.org/10.1002/widm.53>
- Murtagh, F., & Legendre, P. (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *Journal of Classification*, 31(3), 274–295. <https://doi.org/10.1007/s00357-014-9161-z>
- 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)
- Spear, R., & Bidet, E. (2005). Social enterprise for work integration in 12 european countries: a descriptive analysis*. *Annals of Public and Cooperative Economics*, 76(2), 195–231. <https://doi.org/10.1111/j.1370-4788.2005.00276.x>
- Subramanyam, K. R. ., & Wild, John. (2014). *Financial statement analysis K.R. Subramanyam, John Wild*. McGraw-Hill.
- Sugiasuti, R. H., & Pratama, M. R. (2022). Dampak Buruk Pembangunan Tanpa Pemerataan : Kesenjangan Ekonomi Antar Wilayah di Indonesia. *Profit : Jurnal Administrasi Bisnis*, 16(1), 79–90.
- Tambunan, T. T. H. (2017). Women Entrepreneurs in MSEs in Indonesia: Their Motivations and Main Constraints. *International Journal of Gender and Women's Studies*, 5(1). <https://doi.org/10.15640/ijgws.v5n1a9>
- Tan, P.-Ning., Steinbach, Michael., & Kumar, Vipin. (2006). *Introduction to data mining*. Pearson Addison Wesley.
- Undang-Undang Nomor 25 Tahun 1992 Tentang Perkoperasian (1992).
- Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z.-H., Steinbach, M., Hand, D. J., & Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1), 1–37. <https://doi.org/10.1007/s10115-007-0114-2>
- Yim, O., & Ramdeen, K. T. (2015). Hierarchical Cluster Analysis: Comparison of Three Linkage Measures and Application to Psychological Data. *The Quantitative Methods for Psychology*, 11(1), 8–21. <https://doi.org/10.20982/tqmp.11.1.p008>