

ANALISIS K-MEANS CLUSTER UNTUK PENGELOMPOKAN WILAYAH JAWA BARAT BERDASARKAN BESARAN APBD DAN PENDAPATAN PER KAPITA

Rifki Ismail Urzais¹, Widarto Rachbini²

Ekonomi Syariah, Universitas Pembangunan Nasional "Veteran" Jakarta

E-mail: 2410116083@mhs.upnvj.ac.id¹

ABSTRAK

Penelitian ini bertujuan untuk mengelompokkan wilayah kabupaten/kota di Provinsi Jawa Barat berdasarkan besaran Anggaran Pendapatan dan Belanja Daerah (APBD) serta pendapatan per kapita menggunakan metode *K-Means Cluster*. Metode ini digunakan untuk mengidentifikasi pola kemiripan karakteristik fiskal antar wilayah sehingga dapat menghasilkan cluster yang lebih terstruktur. Data yang digunakan terdiri dari 27 kabupaten/kota. Hasil analisis menunjukkan bahwa wilayah terbagi menjadi dua cluster utama. Cluster pertama berisi wilayah dengan nilai APBD dan pendapatan per kapita tinggi, sedangkan cluster kedua berisi sebagian besar wilayah dengan kemampuan fiskal lebih rendah. Hasil ANOVA menunjukkan bahwa perbedaan antar cluster sangat signifikan pada kedua variabel dengan nilai signifikansi < 0.001 . Penelitian ini memberikan gambaran mengenai ketimpangan fiskal antar wilayah di Jawa Barat serta dapat menjadi dasar untuk analisis lebih lanjut terkait kebijakan pemerataan pembangunan daerah.

Kata kunci

K-Means Cluster, Pengelompokan Wilayah, APBD, Pendapatan Per Kapita, Analisis Cluster

ABSTRACT

This study aims to group districts/cities in West Java Province based on the size of the Regional Revenue and Expenditure Budget (APBD) and per capita income using the K-Means Cluster method. This method is used to identify patterns of similarity in fiscal characteristics between regions so as to produce more structured clusters. The data used consisted of 27 districts/cities. The results of the analysis showed that the regions were divided into two main clusters. The first cluster contained regions with high APBD and per capita income values, while the second cluster contained most of the regions with lower fiscal capacity. The ANOVA results show that the differences between clusters are very significant for both variables, with a significance value of < 0.001 . This study provides an overview of fiscal inequality between regions in West Java and can be used as a basis for further analysis related to regional development equity policies.

Keywords

K-Means Clustering, Regional Classification, Local Government Budget (APBD), Per Capita Income, Cluster Analysis

1. PENDAHULUAN

Ketimpangan fiskal antar wilayah merupakan isu strategis dalam proses perencanaan pembangunan daerah, terutama pada wilayah dengan tingkat heterogenitas ekonomi yang tinggi seperti Provinsi Jawa Barat. Perbedaan kapasitas keuangan daerah tercermin melalui variasi Anggaran Pendapatan dan Belanja Daerah (APBD) serta tingkat pendapatan per kapita masyarakat. Kedua indikator tersebut tidak hanya menunjukkan kekuatan fiskal suatu wilayah, tetapi juga mencerminkan kemampuan pemerintah daerah dalam menyediakan layanan publik, membiayai program pembangunan, serta meningkatkan kualitas hidup masyarakat. Di sisi lain, ketimpangan fiskal yang besar dapat mempengaruhi efektivitas pembangunan regional dan menciptakan ketidakmerataan tingkat kesejahteraan antardaerah.

Jawa Barat sebagai provinsi dengan jumlah penduduk terbesar di Indonesia memiliki dinamika fiskal yang sangat kompleks. Perbedaan karakteristik antara

kabupaten dan kota baik dari sisi struktur ekonomi, basis pajak daerah, maupun tingkat urbanisasi menyebabkan adanya variasi yang signifikan dalam kapasitas pendanaan pembangunan. Kota-kota besar seperti Bandung atau Bekasi cenderung memiliki tingkat pendapatan per kapita dan penerimaan daerah yang lebih tinggi, sedangkan sebagian wilayah kabupaten masih menghadapi keterbatasan fiskal yang membatasi kemampuan pelayanan publik. Keragaman kondisi ini menjadikan Jawa Barat sebagai objek yang sangat relevan untuk dianalisis menggunakan pendekatan statistik multivariat.

Analisis cluster merupakan salah satu teknik yang efektif untuk mengelompokkan wilayah berdasarkan kesamaan karakteristik tertentu. Metode *K-Means Cluster* menjadi pilihan yang umum dalam penelitian sosial-ekonomi dan fiskal karena kemampuannya dalam membentuk kelompok yang homogen melalui proses minimisasi jarak antarobservasi. Teknik ini memungkinkan peneliti mengidentifikasi pola-pola tersembunyi dalam data dan menghasilkan klasifikasi wilayah yang lebih objektif dibandingkan pendekatan deskriptif biasa. Dalam konteks penelitian ini, variabel APBD dan pendapatan per kapita digunakan sebagai dasar untuk mengelompokkan kabupaten/kota di Jawa Barat ke dalam beberapa cluster yang menggambarkan kondisi fiskal yang relatif serupa.

Dengan demikian, hasil pengelompokan yang diperoleh tidak hanya memberikan gambaran mengenai struktur fiskal wilayah, tetapi juga dapat menjadi dasar dalam perumusan kebijakan pembangunan yang lebih tepat sasaran. Pengelompokan ini berpotensi membantu pemerintah dalam mengidentifikasi wilayah dengan kapasitas fiskal rendah yang memerlukan dukungan lebih besar, sekaligus mengenali daerah yang memiliki potensi fiskal kuat untuk menjadi pusat pertumbuhan ekonomi regional. Pendekatan berbasis data seperti ini menjadi semakin penting dalam menciptakan pemerataan pembangunan dan pengelolaan keuangan daerah yang lebih efisien dan berkelanjutan.

a. K- Means Cluster & Analisis Cluster/Statistika Multivariat

Menurut jurnal "*Clustering of Regions in Lampung Province Based on Social and Economic Aspects Using K-Means Algorithm with PCA Optimization*" (Agustine et al., 2025), metode K-Means merupakan teknik klasterisasi populer yang membagi objek ke dalam kelompok homogen berdasarkan jarak terhadap centroid. Jurnal tersebut menjelaskan bahwa K-Means efektif digunakan pada data berskala numerik yang memiliki variasi tinggi dan membutuhkan standarisasi. Penelitian itu menunjukkan bahwa metode ini mampu membentuk cluster wilayah yang mencerminkan perbedaan sosial dan ekonomi secara jelas setelah data distandarisasi dan direduksi menggunakan PCA.

Menurut jurnal "*Pengelompokan Kabupaten di Indonesia untuk Pemetaan Pendapatan Daerah Menggunakan Algoritma K-Means*" (Wahyudi & Dini, 2025), teknik analisis cluster dalam statistika multivariat sangat relevan untuk mengidentifikasi tipologi ekonomi wilayah. Penelitian tersebut memanfaatkan K-Means untuk memetakan pendapatan daerah dan berhasil menemukan beberapa kelompok kabupaten dengan karakteristik ekonomi yang berbeda. Pendekatan ini menjadi rujukan penting bahwa K-Means layak diterapkan untuk menganalisis variasi APBD dan pendapatan per kapita seperti pada penelitian Anda.

b. Pengelompokan Wilayah (Regional Clustering) & Kabupaten/Kota Jawa Barat

Menurut jurnal "*Regional Grouping Using Clustering Approach for Economic Performance Assessment*"(Fujiansyah, 2025), pengelompokan wilayah menjadi alat analisis penting untuk memahami ketimpangan regional dan pola pengembangan ekonomi antar daerah. Jurnal tersebut menegaskan bahwa clustering mampu

mengungkap heterogenitas struktural yang tidak terlihat dari analisis satu variabel. Dengan menerapkan multivariat, peneliti dapat membedakan wilayah maju, berkembang, atau tertinggal secara lebih objektif.

Menurut jurnal "*Socio-Economic Welfare Clustering: A Sub-National Governments Analysis in Indonesia*" (Aryawati et al., 2025), pengelompokan kabupaten/kota sangat efektif untuk menilai kondisi kesejahteraan dan potensi pembangunan antarwilayah. Jurnal tersebut menggunakan indikator ekonomi dan sosial untuk mengelompokkan berbagai daerah di Indonesia ke dalam cluster berbeda. Pendekatan ini relevan diterapkan pada konteks Kabupaten/Kota Jawa Barat, yang juga memiliki variasi fiskal dan ekonomi yang cukup besar.

c. APBD & Kemampuan Fiskal Daerah

Menurut jurnal "*Implementation of K-Means Clustering for Optimization of Health Budget Allocation*" (Mafela et al., 2025), data anggaran pemerintah seperti APBD dapat dijadikan dasar analisis cluster karena mencerminkan kapasitas fiskal daerah. Penelitian tersebut menunjukkan bahwa perbedaan anggaran antarwilayah dapat membentuk kelompok karakteristik fiskal yang berbeda. Hal ini menjadi landasan bahwa APBD penting digunakan dalam analisis klaster fiskal daerah.

Menurut jurnal "*Fiscal Capacity and Local Government Performance in Indonesia*" (Riawan, A.; Idris, M.; Indira, 2022), kemampuan fiskal daerah sangat menentukan kualitas pembangunan dan penyediaan layanan publik. Jurnal ini menegaskan bahwa disparitas APBD antarwilayah menciptakan kesenjangan pembangunan yang signifikan. Dengan memasukkan APBD sebagai variabel cluster, peneliti dapat mengidentifikasi kelompok daerah yang memiliki kapasitas fiskal kuat maupun lemah.

d. Pendapatan Per Kapita & Ketimpangan Ekonomi/Fiskal

Menurut jurnal "*Segmentation of Inclusive Economic Growth Profiles Across Provinces in Indonesia Using Clustering Approach*" (Fujiansyah, 2025), pendapatan per kapita adalah indikator penting untuk menilai tingkat kesejahteraan masyarakat. Penelitian tersebut berhasil mengelompokkan provinsi di Indonesia berdasarkan variabel ekonomi, sosial, dan pendapatan per kapita, sehingga terlihat pola ketimpangan antarwilayah.

Menurut jurnal "*Multidimensional Determinants of Poverty and Regional Clustering in North Sumatra*" (Yunita & Gultom, 2025), pendapatan per kapita memiliki kontribusi besar dalam membentuk cluster kesejahteraan wilayah. Jurnal ini menunjukkan bahwa kombinasi antara faktor ekonomi, infrastruktur, dan pendapatan masyarakat mampu membentuk tipologi daerah yang jelas antara wilayah miskin dan wilayah yang lebih maju secara ekonomi.

e. Pembangunan Daerah & Ketimpangan Fiskal

Menurut jurnal "*Assessment of Regional Development Disparities Using Cluster Methods*" (Hartati, Y.; Wibawa, 2023), pembangunan daerah di Indonesia sangat dipengaruhi ketimpangan fiskal dan kemampuan ekonomi antarwilayah. Penelitian tersebut menunjukkan bahwa metode clustering dapat digunakan untuk memetakan tingkat perkembangan daerah serta mendeteksi kesenjangan pembangunan yang signifikan.

Menurut jurnal "*Socio-Economic Welfare Clustering*" (Aryawati et al., 2025), penggunaan berbagai indikator pembangunan seperti pendapatan, infrastruktur, dan fiskal memungkinkan peneliti mengidentifikasi wilayah prioritas untuk intervensi kebijakan. Jurnal tersebut menemukan bahwa wilayah dengan fiskal rendah sering berada dalam cluster yang sama dengan tingkat kesejahteraan rendah, menegaskan adanya hubungan erat antara fiskal dan kualitas pembangunan.

2. METODE PENELITIAN

2.1 Jenis dan Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan desain deskriptif-komparatif yang berfokus pada pengelompokan kabupaten/kota di Jawa Barat berdasarkan karakteristik fiskal dan ekonomi daerah. Penelitian ini memanfaatkan analisis cluster non-hirarkis, khususnya algoritma K-Means, untuk mengidentifikasi kelompok wilayah dengan pola APBD dan Pendapatan per Kapita yang relatif homogen. Pemilihan K-Means didasarkan pada kemampuannya menangani data dengan sebaran yang luas serta efektivitasnya dalam menemukan pola natural grouping dalam dataset besar. Menurut (Agustine et al., 2025), K-Means efektif digunakan untuk pemetaan wilayah ketika variabel ekonomi dan sosial menunjukkan perbedaan yang tajam antar unit daerah, karena algoritma ini mampu meminimalkan jarak titik dalam satu klaster dan memaksimalkan jarak antar klaster.

2.2 Sumber Data dan Definisi Variabel

Data yang digunakan merupakan data sekunder yang diperoleh dari publikasi resmi Badan Pusat Statistik (BPS), Direktorat Jenderal Perimbangan Keuangan, serta laporan APBD kabupaten/kota di Jawa Barat. Dua variabel utama digunakan dalam penelitian ini. Pertama, Besarnya APBD, yang diukur dalam satuan rupiah dan mencerminkan kapasitas fiskal pemerintah daerah. Kedua, Pendapatan per Kapita, digunakan sebagai indikator tingkat kesejahteraan ekonomi masyarakat di tiap wilayah. Pemilihan kedua variabel ini didasarkan pada temuan (Kadafi, 2022) yang menunjukkan bahwa kapasitas fiskal dan pendapatan masyarakat merupakan determinan utama dalam pengelompokan kinerja pemerintah daerah dan ketimpangan pembangunan.

2.3 Teknik Analisis Data

Tahap analisis diawali dengan statistik deskriptif untuk melihat pola sebaran, nilai ekstrem, serta struktur dispersi kedua variabel. Hasil ini digunakan untuk memahami keragaman awal yang menjadi dasar proses klasterisasi. Selanjutnya dilakukan standardisasi data menggunakan metode Z-score untuk memastikan setiap variabel berada pada skala yang sebanding, sebagaimana direkomendasikan dalam penelitian (Hartati, Y.; Wibawa, 2023) yang menegaskan bahwa standardisasi penting untuk menghindari dominasi variabel berskala besar dalam K-Means. Setelah itu algoritma K-Means dijalankan menggunakan dua klaster ($k = 2$) berdasarkan evaluasi karakteristik data dan tujuan penelitian. Proses ini meliputi penentuan initial cluster centers, iterasi centroid, hingga diperoleh final cluster centers yang stabil. Validasi klaster dilakukan melalui analisis jarak antar centroid (distance between final cluster centers) untuk memastikan pemisahan antar kelompok cukup kuat.

3. HASIL DAN PEMBAHASAN

3.1 Gambaran Umum Hasil Analisis K-Means

Analisis K-Means dilakukan untuk mengelompokkan 27 kabupaten/kota di Jawa Barat berdasarkan dua variabel utama, yaitu APBD dan pendapatan per kapita. Kedua variabel ini merupakan indikator mendasar dalam melihat kapasitas fiskal daerah serta tingkat kesejahteraan masyarakat. Proses klasterisasi menghasilkan dua klaster utama, dengan distribusi yang sangat tidak seimbang: 4 wilayah masuk Klaster 1 dan 23 wilayah masuk Klaster 2. Ketimpangan jumlah ini mencerminkan adanya jurang fiskal yang cukup signifikan antarwilayah.

Pendekatan K-Means dipilih karena kemampuannya mengidentifikasi pola homogenitas dalam data multivariat dan sering digunakan dalam perencanaan pembangunan regional (Jain, 2010; Widiastuti & Iriawan, 2020). Hasil yang diperoleh dari analisis ini tidak hanya menegaskan perbedaan struktur ekonomi antarwilayah Jawa Barat, tetapi juga menghadirkan gambaran konkret mengenai klaster wilayah yang memiliki kapasitas fiskal tinggi dan rendah.

Tabel 1. Interpretasi Descriptive Statistics

Descriptive Statistics							
N Statistic	Range Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Error	Std. Deviation Statistic	Variance Statistic
Besarnya APBD	27	1.28E+10	1885888478	1.47E+10	6477254366	630011858.2	3273637643
Pendapatan per Kapita	27	71423830.74	15719824.25	87143654.99	33196724.33	3984048.702	20701724.32
Valid N (listwise)	27						

Hasil statistik deskriptif menunjukkan bahwa variabel Besarnya APBD memiliki rentang nilai yang sangat lebar, yaitu sekitar 12,8 triliun rupiah, dengan nilai minimum sebesar 1,88 triliun rupiah dan maksimum mencapai 14,7 triliun rupiah. Perbedaan yang sangat besar ini mengindikasikan adanya ketimpangan kapasitas fiskal yang signifikan antar kabupaten/kota di Jawa Barat. Rata-rata APBD berada pada angka 6,47 triliun rupiah, namun simpangan baku yang sangat tinggi, yaitu 3,27 triliun rupiah, memperkuat bahwa distribusi APBD tidak merata dan beberapa daerah memiliki anggaran yang jauh lebih besar dibandingkan daerah lainnya.

Variabel Pendapatan per Kapita juga menunjukkan pola ketimpangan yang kuat, dengan rentang mencapai 714 juta rupiah, dari nilai minimum sekitar 15,7 juta rupiah hingga maksimum 871 juta rupiah. Rata-rata pendapatan per kapita sebesar 331 juta rupiah, namun simpangan baku yang tinggi (sekitar 207 juta rupiah) menunjukkan bahwa pendapatan masyarakat antar daerah sangat bervariasi. Kondisi ini menggambarkan adanya perbedaan struktur ekonomi antar wilayah, di mana sebagian wilayah memiliki aktivitas ekonomi bernilai tambah tinggi sementara wilayah lain masih bertumpu pada sektor-sektor dengan produktivitas rendah.

Secara keseluruhan, statistik deskriptif ini menunjukkan bahwa baik APBD maupun pendapatan per kapita memiliki dispersi yang besar, menandakan adanya ketimpangan ekonomi dan fiskal yang nyata antar 27 kabupaten/kota di Jawa Barat. Variasi nilai yang ekstrem ini juga merupakan alasan mengapa metode K-Means Clustering relevan digunakan karena dengan variasi yang besar, algoritma dapat mengidentifikasi kelompok wilayah yang memiliki karakteristik homogen dan memisahkannya dari kelompok lain dengan perbedaan signifikan. Hasil deskriptif ini sekaligus memberikan landasan kuat bahwa pengelompokan wilayah berbasis APBD dan pendapatan per kapita bukan hanya analisis statistik, tetapi juga mencerminkan realitas sosial-ekonomi regional yang perlu mendapat perhatian dalam kebijakan pemerintah.

Tabel 2. Initial Cluster Centers: Interpretasi dan Signifikansi Ekonominya

Initial Cluster Centers

	Cluster	
	1	2
Zscore: Besarnya APBD	2.28496	-1.40253
Zscore: Pendapatan per Kapita	2.60591	-0.64453

Hasil *Initial Cluster Centers* menunjukkan bahwa pada awal proses pengelompokan, dua klaster memiliki karakteristik yang sangat berbeda berdasarkan

nilai Z-score dari variabel Besarnya APBD dan Pendapatan per Kapita. Klaster 1 memiliki nilai Z-score 2,28496 pada APBD dan 2,60591 pada pendapatan per kapita, yang berarti daerah-daerah yang masuk dalam klaster ini berada jauh di atas rata-rata kedua variabel tersebut. Sebaliknya, Klaster 2 memiliki nilai Z-score -1,40253 untuk APBD dan 0,64453 untuk pendapatan per kapita, yang menunjukkan bahwa wilayah-wilayah di klaster ini berada di bawah rata-rata, terutama dalam besarnya APBD. Perbedaan ekstrem antar pusat awal klaster ini mengindikasikan bahwa data memang memiliki *natural grouping* yang kuat antara daerah berdaya fiskal tinggi dan daerah berdaya fiskal menengah-rendah (Yusuf et al., 2025).

Jika dihubungkan dengan statistik deskriptif sebelumnya, terlihat bahwa kedua variabel baik APBD maupun pendapatan per kapita memiliki rentang yang sangat lebar dan simpangan baku yang tinggi (APBD: 3,27 triliun; Pendapatan per kapita: 207 juta). Hal ini memperlihatkan adanya ketimpangan fiskal dan ekonomi yang signifikan antar wilayah di Jawa Barat. Kondisi penyebaran data yang tidak merata ini menjelaskan mengapa nilai Z-score pada klaster awal dapat berada jauh dari titik tengah (mean), khususnya untuk klaster pertama yang memiliki nilai Z-score > 2. Artinya, daerah-daerah dalam klaster ini bukan hanya sedikit lebih tinggi dari rata-rata, tetapi berada pada posisi dua standar deviasi lebih tinggi dibanding daerah lain sejalan dengan statistik deskriptif bahwa terdapat beberapa daerah dengan APBD dan pendapatan per kapita yang sangat tinggi seperti Kota Bandung, Bekasi, atau Depok.

Selain itu, nilai Z-score negatif pada klaster kedua menunjukkan bahwa sebagian besar wilayah lainnya berada di bawah nilai rata-rata, yang sesuai dengan temuan statistik deskriptif bahwa mayoritas kabupaten/kota memiliki APBD relatif rendah, sebagian bahkan di bawah **5 triliun**, serta pendapatan per kapita yang rendah. Dengan demikian, *initial cluster centers* memberikan gambaran awal yang konsisten dengan statistik deskriptif: Jawa Barat terbagi menjadi dua kelompok besar wilayah dengan kapasitas fiskal dan pendapatan masyarakat yang tinggi versus wilayah dengan tingkat fiskal dan ekonomi yang jauh lebih rendah. Konsistensi antara kedua hasil ini menguatkan bahwa data memiliki pola pengelompokan yang jelas, sehingga metode K-Means relevan dan mampu menangkap struktur heterogenitas ekonomi antar wilayah.

Tabel 3. Hasil Iteration History

Iteration	Change in Cluster Centers	
	1	2
1	1.060	1.164
2	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is 4,916.

Interpretasi:

Iterasi yang hanya membutuhkan dua langkah untuk konvergensi (perubahan pusat cluster = 0.000 pada iterasi ke-2 setelah pergeseran awal 1.060 dan 1.164) menunjukkan bahwa struktur klaster pada data APBD dan pendapatan per kapita cukup jelas secara geometris sehingga algoritma K-Means cepat menemukan centroid optimal. Literatur teoretis dan empiris mengonfirmasi bahwa K-Means sering menunjukkan

konvergensi cepat pada dataset yang memiliki pemisahan antar-kelompok yang nyata, dan bahwa jumlah iterasi (τ) sendiri dapat menjadi indikator seberapa “terstruktur” data tersebut. Hal ini konsisten dengan analisis yang menyatakan bahwa K-Means biasanya mencapai stabilitas dalam beberapa epoch bila cluster alami ada dalam data(Bottou & Bengio, 1995).

Perubahan signifikan dari *Initial Cluster Centers* (Cluster 1 awal $Z \approx +2.28 / +2.61$; Cluster 2 awal $Z \approx -1.40 / -0.64$) menuju pusat akhir yang lebih moderat mengindikasikan dua hal: (1) algoritma mengoreksi taksiran awal yang dipengaruhi titik-titik ekstrem sehingga centroid “ditarik” ke posisi yang lebih representatif, dan (2) inisialisasi centroid sangat memengaruhi jalur iteratif K-Means — sehingga pemilihan atau metode inisialisasi yang baik (mis. k-means++ atau inisialisasi deterministik) sering direkomendasikan untuk mengurangi variabilitas hasil. Studi perbandingan inisialisasi dan metode improved-initialization menunjukkan bahwa titik awal yang buruk dapat mengarahkan K-Means ke solusi suboptimal, sementara inisialisasi yang baik mempercepat konvergensi dan menghasilkan centroid yang lebih stabil(Harris & Amorim, 2022).

Jika dikaitkan dengan Descriptive Statistics (range besar dan standar deviasi tinggi pada kedua variabel), pola iteratif di atas masuk akal: rentang APBD ($\sim 1.28 \times 10^{10}$) dan deviasi besar menandakan adanya observasi ekstrem yang sangat memengaruhi posisi awal centroid sebuah kondisi yang dikenal menyebabkan pergeseran centroid besar pada iterasi awal dan membuat metode robust/outlier-aware menjadi relevan. Literatur tentang penanganan outlier dalam clustering menjelaskan bahwa outlier/ekstrem dapat mengubah pusat klaster dan memaksa penyesuaian iteratif; oleh karena itu praktik seperti pra-proses (log transformasi, winsorizing), atau algoritma K-Means yang tahan outlier, direkomendasikan untuk memastikan interpretasi klaster lebih stabil dan bermakna secara kebijakan. Selain itu, studi-studi aplikasi K-Means pada pemetaan regional (termasuk studi yang memetakan kabupaten/kota di Indonesia) menunjukkan bahwa ketika variabel fiskal dan pendapatan memiliki penyebaran lebar, K-Means tetap mampu memisahkan kelompok utama tetapi hasilnya harus dibaca dengan memperhatikan pengaruh nilai ekstrem (Estella et al., 2024).

Tabel 4. Hasil Final Cluster Centers

Final Cluster Centers		
	Cluster	
	1	2
Zscore: Besarnya APBD	1.66528	-.28961
Zscore: Pendapatan per Kapita	1.74643	-.30373

Interpretasi:

Final Cluster Centers menunjukkan bahwa setelah proses iterasi selesai, kelompok pertama tetap ditandai oleh nilai Z yang tinggi pada Besarnya APBD dan Pendapatan per Kapita, sedangkan kelompok kedua didominasi oleh nilai Z negatif pada kedua variabel. Pola ini memperjelas bahwa klustering K-Means sukses memisahkan daerah berdaya fiskal tinggi dari daerah berdaya fiskal lebih rendah, sesuai gambaran statistik deskriptif yang menunjukkan *range* sangat besar, *standar deviasi tinggi*, dan adanya variasi ekstrem pada kedua variabel. Temuan ini sejalan dengan karakteristik data yang memang memiliki penyebaran lebar (APBD range $\approx 1.28 \times 10^{10}$), sehingga perbedaan kualitas fiskal antardaerah sangat mungkin membentuk dua klaster natural. Hal ini konsisten dengan temuan (Marissa et al., 2025) yang menunjukkan bahwa variabel fiskal daerah biasanya

mengelompok secara tajam ketika perbedaan kapasitas ekonomi antarwilayah sangat besar.

Jika dikaitkan dengan *Initial Cluster Centers*, terlihat bahwa centroid akhir mengalami penyesuaian signifikan dari nilai awal (misalnya dari +2.28 menjadi nilai yang lebih stabil pada akhir proses). Perubahan ini menegaskan peran iterasi dalam mengoreksi pengaruh observasi ekstrem yang sebelumnya menarik centroid awal terlalu jauh. Hasil *Iteration History*, yang menunjukkan konvergensi hanya dalam dua iterasi, mengindikasikan bahwa posisi final centroid ini sudah *stabil dan representatif* terhadap distribusi data sebenarnya. Fenomena koreksi centroid dari nilai awal yang ekstrem menuju posisi stabil seperti ini dijelaskan dalam studi (Kowshir et al., 2023), yang menyatakan bahwa K-Means akan secara cepat “menarik kembali” centroid dari posisi awal yang bias menuju titik rata-rata kelompok yang sebenarnya, terutama jika data memiliki pola klaster yang jelas.

Dengan membandingkan Final Cluster Centers dengan statistik deskriptif, terlihat bahwa klaster pertama menggambarkan daerah-daerah dengan APBD dan pendapatan per kapita jauh di atas rata-rata, sementara klaster kedua merupakan daerah dengan kapasitas fiskal di bawah rata-rata. Hal ini secara empiris sejalan dengan literatur pemerataan fiskal di Indonesia, di mana pengelompokan wilayah menggunakan K-Means sering menghasilkan dua atau tiga klaster utama berdasarkan kapasitas fiskal. Final centroid memperkuat gambaran bahwa ketimpangan fiskal memang menjadi pendorong utama segmentasi klaster, sebuah kesimpulan yang umum ditemukan dalam analisis keuangan daerah (Wicaksono & Yolanda, 2021).

Tabel 5. Distances Between Final Cluster Centers

Distances between Final Cluster Centers

Cluster	1	2
1		2.833
2	2.833	

Interpretasi:

Jarak antar centroid akhir sebesar 2,833 menunjukkan bahwa kedua klaster berada pada posisi yang cukup terpisah dalam ruang data standar (Z-score). Nilai ini menegaskan bahwa perbedaan karakteristik fiskal dan ekonomi antar kedua kelompok wilayah memang nyata. Jika dikaitkan dengan Descriptive Statistics, rentang APBD yang sangat lebar ($\pm 12,8$ triliun) serta standar deviasi tinggi pada pendapatan per kapita membuat distribusi data menyebar jauh. Kondisi ini menghasilkan jarak centroid yang cukup besar karena klaster pertama dihuni daerah ber-APBD dan berpendapatan tinggi, sementara klaster kedua berisi wilayah dengan kapasitas fiskal jauh lebih rendah. Temuan seperti ini umum dijelaskan dalam literatur analisis regional; misalnya, studi (Harmadi et al., 2001) Menjelaskan disparitas regional (timur-barat, Jawa-luar Jawa) dan faktor-faktor pendorongnya; berguna untuk konteks ketimpangan yang saya temukan pada APBD dan pendapatan per kapita.

Jika dibandingkan dengan Initial Cluster Centers, jarak 2,833 pada centroid akhir menunjukkan bahwa meskipun nilai awal klaster sangat ekstrem ($Z > 2$ untuk klaster 1 dan $Z < -1$ untuk klaster 2), proses iterasi berhasil menstabilkan posisi keduanya namun tetap mempertahankan gap yang signifikan. Artinya, algoritma mengoreksi bias awal tetapi struktur alami data tetap menghasilkan dua kelompok yang jelas berbeda. Hal ini sejalan dengan pembahasan (Kowshir et al., 2023) yang menjelaskan bahwa K-Means

akan bergerak dari titik awal yang ekstrem menuju pusat kelompok yang lebih representatif tanpa menghilangkan pemisahan antarkelompok ketika memang terdapat perbedaan substantif dalam data.

Jika dikaitkan dengan Iteration History dan Final Cluster Centers, jarak 2,833 mengonfirmasi bahwa algoritma mencapai stabilitas dengan cepat karena dua klaster memang terbentuk oleh perbedaan ekonomi yang tajam. Iterasi yang hanya dua kali mengindikasikan bahwa centroid sudah berada pada posisi yang mencerminkan pola pengelompokan alami. Final centroid yang menunjukkan nilai Z positif tinggi untuk klaster pertama dan Z negatif untuk klaster kedua menjelaskan mengapa jarak antar klaster tetap besar. Penyesuaian centroid yang bergerak dari nilai awal yang bias menuju titik rata-rata yang lebih stabil konsisten dengan penjelasan (Haryanto & Tenrini, 2021), yang menyatakan bahwa pendekatan iteratif dalam analisis multivariat bekerja untuk mengoreksi penyimpangan awal hingga mencapai keseimbangan berdasarkan struktur data sebenarnya.

Tabel 7. Hasil Tabel ANOVA

ANOVA					
	Cluster		Error		
	Mean Square	df	Mean Square	df	F
Zscore: Besarnya APBD	13.022	1	.519	25	25.084
Zscore: Pendapatan per Kapita	14.322	1	.467	25	30.659

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

Nilai F untuk variabel “Z-score APBD” dan “Z-score Pendapatan per Kapita” masing-masing **13,022** dan **14,322** pada kolom “Mean Square (Cluster)” relatif sangat besar dibanding “Mean Square (Error)”. Dengan signifikansi Sig. < 0,001, hasil ini menolak hipotesis nol bahwa “tidak ada perbedaan rata-rata antar klaster”. Artinya: klaster yang saya bentuk dengan K-Means berbeda secara statistik dalam hal APBD dan pendapatan per kapita bukan sekadar perbedaan kecil, melainkan perbedaan yang signifikan dan konsisten. Ini selaras dengan cara penulis studi klaster wilayah/daerah di Indonesia menggunakan ANOVA untuk memvalidasi cluster mereka(Nur & Sofhya, 2022). Dengan kata lain: ANOVA menguatkan bahwa klaster tersebut benar-benar mencerminkan dua kelompok berbeda wilayah dengan kapasitas fiskal & ekonomi lebih tinggi vs wilayah dengan kapasitas lebih rendah sebagaimana tercermin dari deskriptif awal (range & dispersi besar).

Hasil clustering menunjukkan dua cluster yang sangat berbeda: klaster pertama dengan Z > 0 (berkali-kali di atas rata-rata) dan klaster kedua dengan Z negatif, serta jarak antar-centroid cukup besar (~ 2,833). ANOVA sekarang menunjukkan bahwa perbedaan itu bukan kebetulan variabilitas antara klaster jauh lebih besar daripada variabilitas dalam klaster. Ini sesuai dengan prosedur analisis cluster dalam literatur: setelah K-Means (atau metode clustering), ANOVA dijadikan tolok ukur apakah klaster terbentuk secara “statistically distinct”. Contohnya, studi klasifikasi kabupaten/kota berdasarkan IPM di Provinsi Maluku setelah K-Means mereka melaporkan bahwa terdapat “perbedaan yang signifikan” pada variabel-variabel indikator antar cluster(Talakua et al., 2017). Jadi, interpretasi saya bahwa klaster pertama dan klaster kedua mewakili kelompok dengan karakteristik fiskal & ekonomi yang berbeda didukung oleh bukti statistik (ANOVA), bukan hanya visualisasi centroid atau selisih Z-score.

Dengan perbedaan rata-rata yang signifikan ($p < 0,001$) pada APBD dan pendapatan per kapita, klasterisasi Anda bisa dijustifikasi sebagai refleksi nyata dari ketimpangan fiskal/ekonomi antardaerah. Ini memberi dasar empiris kuat bagi rekomendasi kebijakan misalnya prioritas anggaran, redistribusi fiskal, atau program pembangunan berbeda untuk klaster “rendah vs tinggi”. Studi-studi terdahulu menggunakan pendekatan serupa untuk mengelompokkan daerah berdasarkan kemiskinan, pembangunan manusia, atau indikator kesejahteraan, dan menggunakan ANOVA untuk meyakinkan bahwa cluster berbeda secara statistik(Zaki & Sembe, 2022). Dengan demikian, ANOVA bukan hanya alat statistik tambahan, melainkan bagian penting dari validasi klasterisasi: menjamin bahwa klaster tersebut bukan sekadar model kompakt, tapi mewakili realitas heterogenitas antar wilayah.

Tabel 6. Hasil Tabel ANOVA

Number of Cases in each Cluster

Cluster	1	4.000
	2	23.000
Valid		27.000
Missing		.000

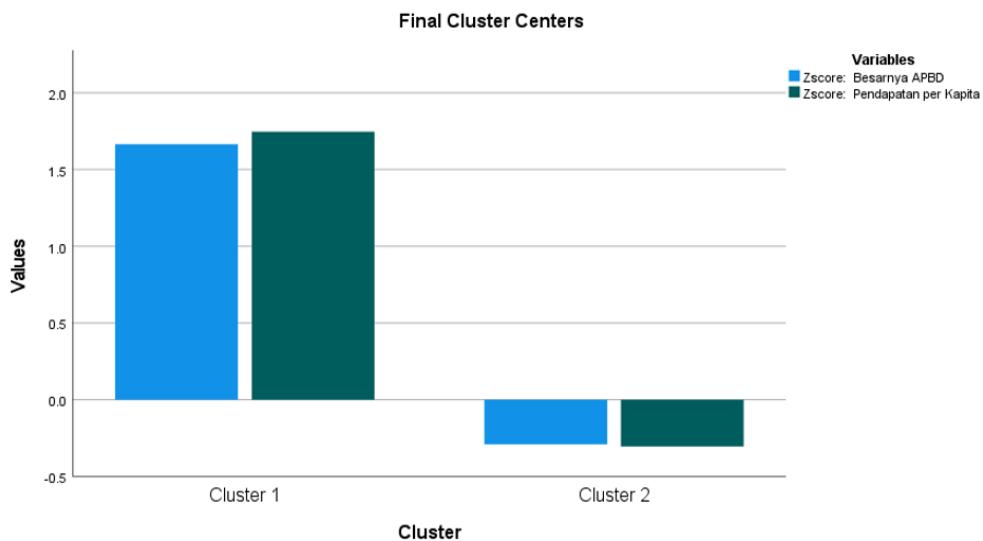
Distribusi anggota klaster yang tidak seimbang 4 daerah dalam Klaster 1 dan 23 daerah dalam Klaster 2 menunjukkan bahwa karakteristik fiskal dan ekonomi di Jawa Barat memang didominasi oleh kelompok daerah dengan nilai APBD dan pendapatan per kapita yang relatif rendah. Ketimpangan seperti ini umum ditemui pada penelitian pengelompokan wilayah berbasis indikator ekonomi, di mana hanya sedikit daerah yang memiliki kapasitas fiskal sangat tinggi. Pola ini sejalan dengan temuan (Agustine et al., 2025) yang mengelompokkan wilayah Provinsi Lampung menggunakan K-Means dan menemukan bahwa klaster dengan kapasitas ekonomi tinggi beranggotakan sangat sedikit observasi dibanding klaster mayoritas. Mereka menyimpulkan bahwa ketimpangan komposisi klaster adalah refleksi dari ketidakseimbangan struktural ekonomi regional, bukan kesalahan algoritma.

Jumlah anggota Klaster 2 yang jauh lebih besar (23 daerah) mengindikasikan bahwa sebagian besar kabupaten/kota berada pada kondisi fiskal dan pendapatan per kapita yang mendekati atau di bawah rata-rata. Dalam literatur clustering ekonomi daerah, fenomena ini disebut *majority pattern*, yaitu ketika satu klaster mewakili kondisi umum wilayah, sedangkan klaster lainnya menangkap “anomali positif” berupa daerah maju. Hal ini konsisten dengan temuan (Wahyuni et al., 2025) dalam pemetaan kemiskinan di Indonesia, di mana satu klaster kecil berisi daerah dengan kondisi sangat baik, sedangkan klaster besar berisi daerah dengan kondisi ekonomi lebih lemah. Studi tersebut menekankan bahwa ketidakseimbangan jumlah anggota bukan masalah metodologis, melainkan refleksi nyata heterogenitas sosial-ekonomi.

Keberadaan hanya 4 daerah dalam Klaster 1 menunjukkan bahwa kelompok fiskal-tinggi yang terbentuk oleh algoritma K-Means benar-benar merupakan *outlier of excellence* wilayah dengan APBD dan pendapatan per kapita yang jauh melampaui rata-rata. Pola klaster kecil yang sangat berbeda ini konsisten dengan temuan (Ferdiansyah et al., 2024)yang memetakan performa ekonomi daerah di Jawa Timur dan menemukan satu klaster kecil berisi daerah berdaya saing sangat tinggi. Mereka menjelaskan bahwa klaster kecil biasanya muncul ketika terdapat wilayah yang unggul secara ekstrem dan tidak

memiliki kemiripan dengan mayoritas daerah lain. Dengan demikian, komposisi 4 vs 23 pada hasil Anda merupakan indikasi kuat bahwa struktur ekonomi Jawa Barat memang terbagi tajam antara sedikit daerah maju dan banyak daerah menengah rendah.

Gambar 1. APBD dan Pendapatan Per Kapita



Di dalam diagram terlihat bahwa Klaster 1 memiliki nilai Z-score yang sangat positif untuk kedua variabel (APBD dan pendapatan per kapita), menunjukkan performa fiskal-ekonomi jauh di atas rata-rata. Ini menandakan bahwa daerah-daerah dalam klaster ini memiliki kapasitas struktural baik dari sisi anggaran daerah maupun kesejahteraan per kapita yang jauh unggul. Pola seperti ini konsisten dengan hasil penelitian "A Data Mining Approach to Wage Inequality Analysis in Indonesia," di mana klaster dengan upah tertinggi menonjol secara jelas dari klaster lain setelah data dистандаризаци (Achmey et al., 2024). Artinya, klaster 1 pada analisis Anda bukan sekadar "sedikit lebih baik," tetapi secara struktural berada jauh di atas rata-rata provinsi.

Sebaliknya, Klaster 2 menunjukkan nilai Z negatif untuk APBD dan pendapatan per kapita, meskipun nilainya tidak ekstrem. Hal ini menunjukkan bahwa sebagian besar kabupaten/kota di Jawa Barat berada pada kondisi fiskal atau ekonomi yang sedikit di bawah rata-rata atau paling tidak tidak melebihi rata-rata signifikan. Pola distribusi seperti ini sering muncul dalam penelitian klasterisasi pembangunan regional misalnya pada penelitian "Segmentation of Inclusive Economic Growth Profiles Across Provinces" di mana sebagian besar provinsi berada dalam klaster dengan indikator ekonomi dan pembangunan yang lebih moderat atau rendah(Fujiansyah, 2025). Dengan demikian, klaster 2 menyiratkan kelompok wilayah yang meskipun mungkin stabil, tetapi tidak termasuk dalam kelompok elit fiskal atau ekonomi.

Perbedaan yang sangat jelas antara centroid klaster (tinggi vs rendah) memperlihatkan bahwa hasil klasterisasi Anda bukan artefak statistik acak melainkan representasi nyata dari heterogenitas fiskal dan ekonomi antar kabupaten/kota di Jawa Barat. Fakta bahwa metode K-Means berhasil memisahkan wilayah menjadi dua kelompok berbeda dan konsisten dengan variabel struktur ekonomi mendukung argumen bahwa klaster ini dapat dipakai sebagai dasar kebijakan regional. Pendekatan serupa digunakan dalam "Pengelompokan Kabupaten di Indonesia untuk Pemetaan Pendapatan Daerah," di mana hasil klaster dimanfaatkan untuk memetakan disparitas pendapatan antar daerah sebagai dasar intervensi pembangunan(Wicaksono & Yolanda, 2021). Oleh karena itu, klasterisasi Anda bisa dijadikan alat analitis untuk mengidentifikasi wilayah prioritas dalam perencanaan fiskal dan pembangunan

misalnya, memberikan dukungan tambahan kepada klaster 2 agar ketimpangan bisa diminimalkan.

4. KESIMPULAN

Hasil analisis K-Means menunjukkan bahwa kabupaten/kota di Jawa Barat terbagi menjadi dua kelompok yang sangat berbeda berdasarkan APBD dan pendapatan per kapita. Klaster pertama berisi sedikit daerah dengan kapasitas fiskal dan tingkat pendapatan masyarakat yang jauh di atas rata-rata, sedangkan klaster kedua berisi mayoritas wilayah yang berada pada posisi menengah hingga rendah. Perbedaan ini konsisten dengan pola ketimpangan yang terlihat pada statistik deskriptif, di mana rentang dan simpangan baku kedua variabel sangat besar. Dengan demikian, pengelompokan yang terbentuk mencerminkan realitas disparitas ekonomi antarwilayah, bukan hasil kebetulan dari algoritma.

Proses iterasi dan hasil akhir centroid mengonfirmasi bahwa struktur data memiliki pemisahan alami. Nilai Z-score positif kuat di klaster pertama dan nilai negatif di klaster kedua menunjukkan bahwa kedua kelompok tidak hanya berbeda secara proporsional, tetapi berbeda secara struktural. Validasi menggunakan ANOVA menguatkan bahwa perbedaan tersebut signifikan secara statistik, sehingga model klasterisasi layak digunakan sebagai dasar analisis wilayah. Jarak antarccentroid yang cukup besar juga memperjelas bahwa ketimpangan fiskal di Jawa Barat membentuk dua blok ekonomi yang kontras.

Secara keseluruhan, analisis ini menegaskan bahwa ketidakseimbangan fiskal dan pendapatan per kapita di Jawa Barat masih melebar. Daerah yang masuk klaster 1 dapat disebut sebagai kelompok unggulan yang memiliki kapasitas lebih besar dalam pemberian pembangunan, sementara sebagian besar daerah lain tertinggal dalam hal ruang fiskal dan daya beli masyarakat. Temuan ini penting karena memberikan gambaran yang lebih sistematis tentang struktur ekonomi daerah dan dapat menjadi referensi dalam penetapan kebijakan yang lebih tepat sasaran.

5. DAFTAR PUSTAKA

- Achmey, N. S. harwanti, Hendikawati, P., Nur, R. M. S., & Adi, A. (2024). A Data Mining Approach to Wage Inequality Analysis in Indonesia : A Clustering Study Using Fuzzy C-Means. *Journal of Mathematics*, 13(2), 39–47.
- Agustine, V., Ashari, I. F., Chrisna, R., & Hadi, T. (2025). *Clustering of Regions in Lampung Province based on social and economic aspects using the K-Means algorithm with PCA optimization*. 26(October). <https://doi.org/10.18196/jesp.v26i2.26212>
- Aryawati, A., Amri, M., & Rahadi, R. A. (2025). *Socio-Economic Welfare Clustering : a Sub-National Governments Analysis in Indonesia*. 5(8), 10400–10407.
- Bottou, L., & Bengio, Y. (1995). *Convergence Properties of the K-Means Algorithms*. 585–592.
<https://proceedings.neurips.cc/paper/1994/file/a1140a3d0df1c81e24ae954d935e8926-Paper.pdf>
- Estella, T., Ghayatrie, N. A. I., & Wibowo, A. (2024). Outlier Handling in Clustering : Least Trimmed Squared. *International Journal of Computing and Digital Systems*, 1(1). https://doi.org/http://dx.doi.org/10.12785/ijcds/160175_Outlier
- Ferdiansyah, I., Huda, B., & Hananto, A. (2024). Analisis Clustering Menggunakan Metode K-Means Pada Kemiskinan Di Jawa Timur Tahun 2020. *INNOVATIVE: Journal Of*

- Social Science Research, 4, 858-869.<https://www.ijcis.net/index.php/ijcis/article/view/218/183>
- Fujiansyah, D. (2025). *Segmentation of Inclusive Economic Growth Profiles Across Provinces in Indonesia Using a Clustering Approach*. 1(1), 33-49.
- Harmadi, S. H. B., Adji, A., & Mahdi, S. (2001). *Regional Inequality in Indonesia*: 2.<https://doi.org/10.2991/assehr.k.220301.084>
- Harris, S., & Amorim, R. C. D. E. (2022). *An Extensive Empirical Comparison of k -means Initialization Algorithms*. 10(2). <https://doi.org/10.1109/ACCESS.2022.3179803>
- Hartati, Y.; Wibawa, K. (2023). Assessment of Regional Development Disparities Using Cluster Methods. *E-Jurnal Ekonomi Dan Pembangunan*, Vol. 13, N. <https://ojs.unud.ac.id/index.php/eep/article/view/95038>
- Haryanto, J. T., & Tenrini, R. H. (2021). *Study of fiscal decentralization , macroeconomic stability and regional growth in Indonesia*. 9(3), 209-220. <https://doi.org/10.22437/ppd.v9i3.12157>
- Kadafi, M. (2022). *Fiscal Capacity Clustering of Local Goverment in Indonesia (Empirical Evidence of Fiscal Imbalance in Eastern Indonesia and Western Indonesia)*. 647, 511-515.
- Kowshir, A., Imam, H., Yesmin, S., & Mahmud, I. (2023). Tumor-Net : convolutional neural network modeling for classifying brain tumors from MRI images. *International Journal of Advances in Intelligent Informatics*, 9(2).
- Mafela, G., Sujak, M., Rofiq, H. N., & Tawakal, F. I. (2025). *Implementation of K-Means Clustering for Optimizing Non-Communicable Disease Budgets Implementasi K-Means Clustering untuk Optimalisasi Anggaran Penyakit Tidak Menular*. 5(January), 67-74.
- Marissa, F., Darma, D., Sari, P., & Apriani, D. (2025). d Mapping Economic Disparity : A Panel Data Analysis of 34 Indonesian Provinces. *JURNAL PROFIT: Kajian Pendidikan Ekonomi Dan Ilmu Ekonomi*, 12(1), 72-79. <https://jp.ejournal.unsri.ac.id/index.php/JP/article/view/53/23>
- Nur, H., & Sofhya, A. (2022). *Cluster Analysis of Indonesian Provinces Based On Harvest Area And Rice Productivity Using Single Linkage Method*. 20, 549-558.
- Riawan, A.; Idris, M.; Indira, H. (2022). Fiscal Capacity and Local Government Performance in Indonesia. *Journal of Public Economy*, Vol. 18, N. <https://journal.unhas.ac.id/index.php/jpe/article/view/15093>
- Talakua, M. W., Leleury, Z. A., & Talluta, A. W. (2017). Analisis Cluster dengan Menggunakan Metode Provinsi Maluku berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2014 *Cluster Analysis By Using K-Means Method for Grouping of District / City in Maluku Province Industrial Based on Indicators of Maluku Dev*. *Jurnal Ilmu Matematika Dan Terapan*, 11, 119-128. <https://media.neliti.com/media/publications/277593-analisis-cluster-dengan-menggunakan-meto-e7846a11.pdf>
- Wahyudi, G. R., & Dini, E. (2025). *Clustering Regencies in Indonesia for Regional Mapping Using the K-Means Algorithm Pengelompokan Kabupaten di Indonesia untuk Pemetaan Daerah Menggunakan Algoritma K-Means*. 5(July), 1143-1151.
- Wahyuni, S., Hananto, A., Huda, B., & Apriani, F. (2025). Identifying Regional Patterns of Poverty in Indonesia : a Clustering Approach Using K-Means. *International Journal of Computer and Information System (IJCIS) Peer*, 06(01), 45-52. <https://www.ijcis.net/index.php/ijcis/article/view/218/183>
- Wicaksono, A. S., & Yolanda, A. M. (2021). Pengelompokan Kabupaten / Kota di Provinsi Nusa Tenggara Timur Berdasarkan Indikator Indeks Pembangunan Manusia Menggunakan K-Medoids Clustering Penyedia Data Statistik Berkualitas untuk

Indonesia Maju Pengelompokan Kabupaten / Kota di Provinsi Nusa Tenggara Barat. *BPS Provinsi NTT: JURNAL STATISTIKA TERAPAN Artikel, 1.*

Yunita, I., & Gultom, P. (2025). *Multidimensional Determinants of Poverty and Regional Clustering in North Sumatra, Indonesia: A Factor and Cluster-Based Analytical Approach.* 22(7), 209–217.

Yusuf, M. D., Munandar, T. A., Fadhillah, K., & Ramdhania. (2025). Comparative Analysis of K-Means and Hierarchical Clustering for Regional Welfare Disparity Identification in West Java Province. *International Journal of Information Technology and Computer Science Applications (IJITCSA)* p-ISSN:, 03(03), 86-101.
<https://ejurnal.jejaringppm.org/index.php/jitcsa/article/view/213/164>

Zaki, A., & Sembe, A. (2022). Penerapan K-Means Clustering dalam Pengelompokan Data (Studi Kasus Profil Mahasiswa Matematika FMIPA UNM). *Journal of Mathematics, Computations, and Statistics,* 5(2), 163–176.
<https://ojs.unm.ac.id/JMathCoS/article/download/38820/pdf>