

PEMETAAN KETIMPANGAN PENDIDIKAN MENENGAH ATAS DI INDONESIA BERDASARKAN KLASTER PROVINSI MENGUNAKAN METODE *K-MEANS*

Dinda Aska Amalia, Rista Ayu Restari, Syifa Nurul Azmi & Hilma Mutiara Winata
UIN Sunan Gunung Djati Bandung, Indonesia
Email: dndaaskaamalia@gmail.com*

Riwayat Artikel

Diterima: 30 Oktober 2025
Disetujui: 30 November 2025
Diterbitkan: 13 Desember 2025

Abstract

Inequality in upper secondary education in Indonesia remains a challenge in terms of equalizing the quality of human resources. This study aims to map the level of educational inequality between provinces using the K-Means Clustering algorithm based on data from the Central Statistics Agency (BPS) for 2024. The analysis was conducted on seven indicators, namely School Participation Rate (APS), Literacy Rate (AMH), Average Length of Schooling (RLS), Expected Length of Schooling (HLS), number of schools, number of teachers, and number of high school students through the stages of data cleaning, feature scaling, and determining the optimal number of clusters using the Elbow Method, Silhouette Score, and Davies–Bouldin Index. The results showed three optimal clusters ($K=3$) with a Silhouette Score of 0.587 and a Davies–Bouldin Index of 0.5369. Cluster 0 represents provinces with high educational achievement and infrastructure, cluster 1 describes disadvantaged areas with limited access to education, and cluster 2 contains moderately developed regions. This mapping confirms spatial and structural educational inequality, especially in eastern Indonesia, and is expected to form the basis for region-based affirmative policies to promote equitable senior secondary education in Indonesia.

Keywords : Educational Inequality, Senior Secondary Education, Clustering, *K-Means*.

Abstrak

Ketimpangan pendidikan menengah atas di Indonesia masih menjadi tantangan dalam pemerataan kualitas sumber daya manusia. Penelitian ini bertujuan memetakan tingkat ketimpangan pendidikan antar provinsi menggunakan algoritma *K-Means Clustering* berbasis data Badan Pusat Statistik (BPS) tahun 2024. Analisis dilakukan terhadap tujuh indikator yaitu Angka Partisipasi Sekolah (APS), Angka Melek Huruf (AMH), Rata-rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), jumlah sekolah, jumlah guru, dan jumlah murid SMA melalui tahapan *data cleaning*, *feature scaling*, serta penentuan jumlah kluster optimal dengan *Elbow Method*, *Silhouette Score*, dan *Davies–Bouldin Index*. Hasil menunjukkan tiga kluster optimal ($K=3$) dengan *Silhouette Score* 0,587 dan *Davies–Bouldin Index* 0,5369. *Cluster 0* merepresentasikan provinsi dengan capaian pendidikan dan infrastruktur tinggi, *cluster 1* menggambarkan daerah tertinggal dengan akses pendidikan terbatas, dan *cluster 2* berisi wilayah berkemajuan sedang. Pemetaan ini mengkonfirmasi ketimpangan pendidikan yang bersifat spasial dan struktural terutama di wilayah timur Indonesia, serta diharapkan menjadi dasar kebijakan afirmatif berbasis wilayah guna mendorong pemerataan pendidikan menengah atas di Indonesia.

Kata Kunci: Ketimpangan Pendidikan, Pendidikan Menengah Atas, Klasterisasi, *K-Means*.

A. PENDAHULUAN

Pendidikan merupakan proses sadar dan terencana yang membantu seseorang mengembangkan potensi, kepribadian, kecerdasan, akhlak, dan keterampilan yang dibutuhkan bagi kehidupannya (Makkawaru, 2019). Pendidikan juga berperan penting dalam mewujudkan pemerataan kesempatan belajar dan mengurangi ketimpangan khususnya pada jenjang menengah atas di Indonesia. Melalui pendidikan yang merata dan berkualitas, suatu negara dapat meningkatkan produktivitas, kesejahteraan sosial, serta daya saing global. Namun, di Indonesia pemerataan pendidikan masih menjadi tantangan besar yang berimplikasi pada ketimpangan kualitas sumber daya manusia antarwilayah (Vega et al., 2024).

Berbagai faktor berkontribusi terhadap masih tingginya ketimpangan pendidikan di Indonesia terutama pada jenjang menengah atas. Perbedaan tingkat kesejahteraan antarwilayah, disparitas infrastruktur pendidikan, serta ketimpangan akses terhadap sumber daya seperti guru berkualitas dan sarana belajar menjadi penyebab utama (Melyana R Pugu, Elmiwati, 2025). Wilayah perkotaan umumnya memiliki fasilitas pendidikan yang lebih lengkap dan akses yang lebih mudah terhadap teknologi pembelajaran. Sementara daerah pedesaan dan terpencil kerap menghadapi keterbatasan baik dari segi jumlah sekolah, tenaga pendidik, maupun dukungan pemerintah daerah (Risna Fadillah, Respy Desmaryani, 2025). Kondisi ini tidak hanya menciptakan kesenjangan dalam hal akses pendidikan, tetapi juga mempengaruhi mutu pembelajaran dan hasil belajar peserta didik di berbagai provinsi di Indonesia.

Ketimpangan pendidikan di Indonesia tidak hanya disebabkan oleh perbedaan ekonomi antarwilayah, tetapi juga oleh distribusi sumber daya pendidikan yang belum merata seperti ketersediaan guru dan fasilitas sekolah (Indriaty et al., 2025). Studi yang dilakukan oleh Arfan dan Pekei (2025) menunjukkan bahwa persebaran guru dan sekolah menengah atas di Indonesia masih terkonsentrasi di wilayah perkotaan sehingga daerah terpencil mengalami keterbatasan akses dan kualitas pendidikan (Arfan & Pekei, 2025). Hal ini sejalan dengan penelitian dari Antika dan Ismasari yang memetakan pola distribusi sekolah menggunakan teknik klasterisasi dan menemukan bahwa ketimpangan antar provinsi berdampak pada rendahnya pemerataan kualitas pendidikan (Kamalia & Nawangsih, 2025). Oleh karena itu, analisis berbasis data spasial dengan metode seperti *K-Means Clustering* diperlukan untuk mengidentifikasi klaster wilayah yang memiliki karakteristik serupa dalam hal pemerataan pendidikan menengah atas sehingga kebijakan pendidikan dapat disusun secara lebih tepat sasaran.

Penelitian terdahulu oleh Mashitawai et al. (2025) mengelompokkan kabupaten/kota di Provinsi Nusa Tenggara Timur berdasarkan enam indikator kualitas pendidikan APS, APK, RLS, rasio guru dan murid SMA, AMH, dan jumlah SMA menggunakan algoritma *K-Means* dan *K-Medoids*. Hasilnya menunjukkan bahwa *K-Means* lebih optimal dengan *Silhouette Score* sebesar 0,459 dan membentuk dua klaster utama yaitu wilayah dengan kualitas pendidikan tinggi dan rendah (Mashitawai et al., 2025). Penelitian serupa oleh Lenama et al. (2023) juga menunjukkan efektivitas metode klasterisasi dalam memetakan disparitas pendidikan di berbagai daerah Indonesia. Namun, sebagian besar studi tersebut masih berfokus pada wilayah tertentu dan belum menyoroti pemerataan pendidikan menengah atas antar provinsi secara nasional (Nony Lenama, Maria A. Kleden, 2023).

Selain itu, penelitian oleh Wardhana, Kharisma, dan Fauzy (2023) menunjukkan bahwa ketimpangan pendidikan di Indonesia dipengaruhi oleh berbagai faktor ekonomi dan sosial seperti ketimpangan pendapatan, tingkat putus sekolah, serta pengeluaran pemerintah untuk pendidikan. Hasil tersebut memperkuat bahwa disparitas pendidikan antar provinsi tidak hanya disebabkan oleh aspek internal lembaga pendidikan, tetapi juga oleh kondisi sosial

ekonomi daerah. Namun, kajian tersebut belum menggambarkan pola spasial ketimpangan pendidikan antar provinsi secara visual dan berkelompok sehingga diperlukan pendekatan analisis klaster untuk melihat peta distribusinya secara lebih komprehensif (Wardhana et al., 2023).

Berdasarkan tinjauan penelitian terdahulu, masih terdapat kesenjangan dalam kajian pemerataan pendidikan menengah atas pada tingkat nasional. Sebagian besar studi sebelumnya hanya berfokus pada wilayah atau provinsi tertentu, sehingga belum mampu menggambarkan pola ketimpangan pendidikan antar provinsi secara menyeluruh di Indonesia. Oleh karena itu, penelitian ini difokuskan untuk menjawab pertanyaan utama: *"Bagaimana pola ketimpangan pendidikan menengah atas antar provinsi di Indonesia, serta bagaimana karakteristik klaster provinsi yang terbentuk berdasarkan hasil analisis menggunakan metode K-Means Clustering?"*

Sejalan dengan permasalahan tersebut, penelitian ini bertujuan untuk memetakan tingkat ketimpangan pendidikan menengah atas antar provinsi di Indonesia berdasarkan indikator pendidikan antara lain Angka Partisipasi Sekolah (APS) 16-18, Angka Melek Aksara 15-24, Rata-Rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), Jumlah Sekolah SMA, Jumlah Guru SMA, dan Jumlah Murid SMA. Melalui metode *K-Means Clustering*, penelitian ini diharapkan mampu mengelompokkan provinsi-provinsi di Indonesia ke dalam klaster yang memiliki karakteristik serupa dalam hal pemerataan dan kualitas pendidikan.

Hasil penelitian ini diharapkan dapat memberikan kontribusi nyata bagi pemerintah, khususnya dalam perumusan kebijakan pemerataan pendidikan menengah atas berbasis data dan spasial. Dengan adanya pemetaan klaster pendidikan ini, diharapkan pengambil kebijakan dapat menetapkan prioritas pembangunan pendidikan secara lebih efektif dan terarah sesuai dengan kondisi dan kebutuhan tiap wilayah. Selain itu, penelitian ini juga diharapkan dapat memperkaya kajian empiris mengenai analisis ketimpangan pendidikan di Indonesia yang masih terbatas dilakukan pada tingkat nasional.

B. KAJIAN PUSTAKA

Pendidikan

Menurut Undang-Undang Nomor 20 Tahun 2003 Pasal 1 Ayat (1), pendidikan merupakan suatu usaha yang dilakukan secara sadar dan terencana untuk menciptakan lingkungan belajar serta proses pembelajaran yang memungkinkan peserta didik aktif dalam mengembangkan potensi dirinya. Dalam proses tersebut, berbagai kemampuan yang dimiliki anak perlu diasah mencakup aspek kognitif, intuitif, serta berbagai keterampilan lainnya agar dapat menjadi bekal dalam menghadapi persaingan di masa depan (Pemerintah Republik Indonesia, 2003). Pendidikan juga dapat dipahami sebagai suatu proses yang dilaksanakan melalui berbagai cara untuk mendorong peserta didik agar mampu mengembangkan dirinya secara aktif. Proses ini meliputi pengembangan kemampuan berpikir logis, peningkatan pengetahuan, pembentukan mental dan spiritual yang kuat, pengendalian diri, serta penguasaan keterampilan yang berguna bagi kehidupan (Dihe & Wangdra, 2023). Pada penelitian ini, jenjang pendidikan menengah atas (SMA/SMK) dipandang sebagai fase strategis pembentukan kualitas sumber daya manusia.

Ketimpangan Pendidikan

Ketimpangan pendidikan di Indonesia masih menjadi permasalahan mendasar dalam pembangunan sumber daya manusia. Menurut Anwar (2022), perbedaan kondisi geografis, ekonomi, dan sosial menyebabkan ketimpangan akses pendidikan antarwilayah. Daerah perkotaan dengan infrastruktur lebih baik menunjukkan partisipasi dan mutu pendidikan yang lebih tinggi dibandingkan wilayah pedesaan (Anwar, 2022). Hasil studi Aditomo dan Felicia (2018) menegaskan bahwa status sosial ekonomi keluarga berpengaruh nyata

terhadap capaian belajar siswa (Aditomo et al., 2018). Ketimpangan tersebut juga tampak pada jenjang pendidikan menengah atas khususnya pada angka partisipasi kasar (APK) SMA yang menunjukkan kesenjangan mencolok antarprovinsi. Hal ini menegaskan bahwa pemerataan akses pendidikan SMA menjadi indikator penting dalam menilai kualitas pembangunan manusia antar provinsi di Indonesia. Pemetaan ketimpangan ini diperlukan untuk mengidentifikasi wilayah prioritas yang membutuhkan intervensi kebijakan pendidikan sekaligus sebagai dasar penentuan alokasi dan strategi pemerataan oleh pemerintah.

Clustering

Clustering merupakan salah satu metode dalam analisis data yang berfungsi untuk mengelompokkan sekumpulan data ke dalam beberapa kelompok berdasarkan kemiripan karakteristik antar data. Proses pengelompokan ini dilakukan secara otomatis menggunakan model statistik, sehingga setiap kelompok (*cluster*) berisi data dengan ciri-ciri yang relatif serupa. Secara umum, teknik *clustering* digunakan untuk menganalisis data yang belum memiliki label atau kategori tertentu dengan tujuan mengidentifikasi pola dan hubungan tersembunyi dalam kumpulan data tersebut. Melalui pendekatan ini, data dapat dikelompokkan ke dalam sejumlah cluster yang memiliki tingkat kesamaan tinggi di dalam kelompok dan perbedaan yang jelas antar kelompok (Rohma et al., 2025). Dalam penelitian ini, *clustering* diterapkan untuk mengelompokkan provinsi berdasarkan indikator pendidikan menengah atas guna memetakan tingkat ketimpangan antarwilayah.

K-Means

K-Means merupakan salah satu algoritma *clustering* yang paling banyak digunakan dalam analisis data. Metode ini termasuk dalam kategori *unsupervised learning* yaitu pendekatan dalam *machine learning* yang digunakan untuk mengelompokkan data tanpa adanya label atau kategori yang telah ditentukan sebelumnya (Sinaga, K. P., & Yang, 2020). Proses kerja *K-Means* dilakukan dengan cara mengelompokkan data berdasarkan rata-rata posisi atau pusat titik (*centroid*) dari setiap *cluster* yang terbentuk. Metode ini digunakan untuk mengidentifikasi pola ketimpangan pendidikan SMA antar provinsi secara kuantitatif dan objektif.

Teori Modal Manusia (*Human Capital Theory*)

Penelitian ini didasarkan pada Teori Modal Manusia (*Human Capital Theory*) yang menjelaskan bahwa pendidikan merupakan investasi penting dalam meningkatkan kualitas sumber daya manusia (Becker, 1964). Semakin tinggi tingkat dan pemerataan pendidikan di suatu wilayah, semakin besar pula kontribusi masyarakat terhadap produktivitas dan pertumbuhan ekonomi daerah. Ketimpangan pendidikan menengah atas di Indonesia dapat dipahami sebagai bentuk ketimpangan dalam distribusi modal manusia antarprovinsi. Provinsi dengan akses pendidikan lebih baik akan menghasilkan tenaga kerja dengan kompetensi tinggi, sedangkan wilayah dengan keterbatasan fasilitas pendidikan cenderung tertinggal dalam pembangunan ekonomi.

Teori Ketimpangan Regional (*Regional Disparities Theory*)

Teori Ketimpangan Regional (*Regional Disparities Theory*) menjelaskan bahwa perbedaan pembangunan antarwilayah terjadi karena ketimpangan distribusi sumber daya, infrastruktur, dan kebijakan publik (Todaro, M. P., & Smith, 2012). Dalam pendidikan, teori ini menegaskan bahwa wilayah dengan infrastruktur pendidikan yang baik cenderung memiliki capaian pendidikan yang lebih tinggi sehingga menciptakan kesenjangan spasial dalam kualitas sumber daya manusia. Untuk memahami dan mengukur fenomena tersebut secara empiris, penelitian ini menggunakan metode *K-Means Clustering* sebagai alat analisis kuantitatif yang mampu mengelompokkan provinsi berdasarkan kemiripan indikator pendidikan menengah atas. Melalui pendekatan ini, teori-teori tersebut diterjemahkan ke

dalam bentuk pemetaan data (*data-driven mapping*) yang menggambarkan pola nyata ketimpangan pendidikan antar provinsi sekaligus menjadi dasar empiris bagi perumusan kebijakan pemerataan pendidikan nasional.

Teori Pembangunan Manusia (*Human Development Theory*)

Teori ini diperkenalkan oleh *United Nations Development Programme* (UNDP) melalui *Human Development Report* 1990. Menurut UNDP (1990), pembangunan manusia diukur dari tiga dimensi utama yaitu pendidikan, kesehatan, dan pendapatan (Masouras, A., Daskou, S., Pistikou, V., & Dimitriou, 2023). Pendidikan berperan penting dalam memperluas kemampuan dan pilihan individu, serta menjadi ukuran penting dalam indeks pembangunan manusia (*Human Development Index/HDI*). Dalam penelitian ini, teori ini digunakan untuk menjelaskan hubungan antara pendidikan dan kualitas hidup masyarakat. Pada penelitian ini, kinerja pendidikan menengah atas memengaruhi komponen pendidikan dalam IPM melalui indikator rata-rata dan harapan lama sekolah, sehingga mencerminkan ketimpangan pembangunan manusia antar provinsi di Indonesia.

C. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode deskriptif-eksploratif untuk memetakan ketimpangan pendidikan menengah atas antar provinsi di Indonesia (Siroj et al., 2024). Pendekatan kuantitatif digunakan karena fokus penelitian ini adalah pada pengolahan data numerik dari indikator pendidikan yang bersumber dari Badan Pusat Statistik (BPS) tahun 2024 meliputi Angka Partisipasi Sekolah (APS) usia 16–18 tahun, Angka Melek Aksara usia 15–24 tahun, Rata-Rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), serta Jumlah Sekolah, Guru, dan Murid SMA (Badan Pusat Statistik, 2024d, 2024c, 2024b, 2024a, 2024e). Metode deskriptif digunakan untuk menggambarkan kondisi pendidikan di tiap provinsi. Sedangkan metode eksploratif diterapkan untuk menemukan pola dan kelompok provinsi yang memiliki kemiripan karakteristik pendidikan menggunakan algoritma *K-Means Clustering*. Analisis data dilakukan menggunakan Google Colab dengan bahasa pemrograman Python yang mendukung proses komputasi statistik dan *machine learning*. Titik elbow juga dianalisis untuk melihat penurunan inersia yang paling menonjol.

Tahapan analisis mencakup enam langkah utama, yaitu: (1) *Data Importing* dari sumber BPS dalam format CSV; (2) *Data Cleaning* untuk menghapus nilai hilang dan anomali agar dataset bersih dan siap digunakan; (3) *Feature Scaling* menggunakan metode *Min-Max Scaler* agar setiap indikator pendidikan memiliki skala yang sebanding; (4) penerapan algoritma *K-Means Clustering* untuk mengelompokkan provinsi berdasarkan kemiripan indikator pendidikan; (5) *Model Validation* menggunakan metode *Elbow Method*, *Silhouette Score*, dan *Davies–Bouldin Index* guna menentukan jumlah cluster optimal; serta (6) *Visualization* hasil klaster dalam bentuk *scatter plot* dan peta sebaran provinsi untuk memudahkan interpretasi spasial. Pendekatan ini memungkinkan identifikasi pola dan klasifikasi wilayah secara objektif, sekaligus memberikan gambaran empiris mengenai tingkat ketimpangan pendidikan menengah atas antar provinsi di Indonesia.

D. HASIL DAN PEMBAHASAN

Dinamika Awal Ketimpangan Pendidikan antar Provinsi

Ketimpangan dalam pendidikan menengah atas di Indonesia tidak hanya bersumber dari perbedaan ekonomi antar wilayah tetapi juga mencerminkan variasi akses, capaian, serta kapasitas layanan yang disediakan oleh pemerintah daerah. Untuk memetakan dinamika tersebut, penelitian ini menggunakan tujuh indikator pendidikan dari Badan Pusat Statistik (2024) yang mencakup aspek partisipasi (APS 16–18), angka melek aksara (AMA 15–24), *outcome* jangka pendek dan panjang (RLS dan HLS), serta kapasitas layanan (jumlah

sekolah, guru, dan murid SMA). Ketujuh variabel ini diolah menggunakan normalisasi *z-score* untuk memastikan skala yang seragam dalam proses klasterisasi. Setelah dataset berhasil ditampilkan, dilakukan analisis statistik deskriptif untuk memberikan gambaran awal karakteristik setiap indikator yang digunakan.

Tabel 1. Statistik Deskriptif Indikator Pendidikan Menengah Atas

Variabel	N	Min	Max	Mean	Std. Dev
APS 16–18	38	47,65	90,36	75,07052632	7,525172153
Angka Melek Aksara 15–24	38	90,46	99,97	99,32105263	2,127599573
Rata-rata Lama Sekolah	38	4,21	11,49	8,841315789	1,227810997
Harapan Lama Sekolah	38	9,63	15,7	13,20973684	1,021693049
Jumlah Sekolah SMA	38	40	1820	386,1842105	391,577396
Jumlah Guru SMA	38	887	44558	9473,947368	9652,679597
Jumlah Murid SMA	38	13821	842760	142109,6579	170420,4

Sumber: BPS, data diolah oleh penulis

Tabel 1 menunjukkan variasi yang cukup jauh antar provinsi dalam hal akses, capaian, dan kapasitas layanan pendidikan menengah atas. Dapat terlihat bahwasanya jumlah murid SMA di provinsi dengan populasi terendah tercatat sebanyak 13,821, sedangkan provinsi tertinggi mencapai 842,760 murid. Perbedaan ini mengindikasikan adanya kesenjangan yang sangat terlihat jelas dalam beban layanan pendidikan. Ketimpangan juga terlihat pada indikator partisipasi dan kualitas pendidikan dimana APS berkisar antara 47,65% hingga 90,36%. Sementara rata-rata lama sekolah bervariasi dari 4,21 tahun hingga 11,49 tahun yang mencerminkan perbedaan nyata dalam pencapaian pendidikan dasar hingga menengah atas. Fakta ini menggambarkan bahwa kualitas pendidikan bukan hanya soal kurikulum, tetapi juga soal siapa yang punya akses terhadapnya dan siapa yang tertinggal.

Temuan deskriptif ini menunjukkan bahwa pendekatan satu kebijakan untuk semua wilayah cenderung tidak efektif mengingat karakteristik pendidikan di tiap provinsi berbeda secara substansial. Oleh karena itu, diperlukan metode pengelompokan wilayah yang berbasis data untuk memahami pola ketimpangan secara lebih objektif dan terarah. Gagasan ini selaras dengan pandangan Becker (1964) yang menekankan peran pendidikan sebagai investasi strategis dalam peningkatan kualitas sumber daya manusia. Serta sejalan dengan pemikiran Todaro, M. P., & Smith (2012) mengenai dampak distribusi layanan publik terhadap kesenjangan pembangunan. Dengan dasar tersebut, pendekatan klasterisasi menjadi langkah relevan untuk membahas lebih lanjut struktur ketimpangan pendidikan antar wilayah.

Penentuan Nilai K Optimal untuk Klasterisasi

Seluruh indikator pendidikan dinormalisasi menggunakan *z-score* untuk memastikan skala yang sebanding antarprovinsi. Tahap berikutnya adalah menentukan jumlah klaster (K) yang paling tepat untuk memetakan ketimpangan pendidikan berdasarkan kemiripan karakteristik antarwilayah. Nilai K yang diuji dalam penelitian ini adalah 2 hingga 5, dengan evaluasi menggunakan tiga metrik utama yaitu *inertia*, *silhouette score*, dan *Davies–Bouldin Index*. Masing-masing metrik memberikan gambaran tentang efisiensi pengelompokan dan jarak antar klaster. Hasil data pengelompokan yang akan diuji dapat dilihat pada gambar di bawah ini

	K	Inertia	Silhouette	Davies-Bouldin
0	2	184.879902	0.502639	0.722233
1	3	81.357236	0.587116	0.536983
2	4	64.512531	0.285364	1.017014
3	5	44.856705	0.320618	0.778834

Gambar 1. Data Uji K
Sumber: Diolah Peneliti, 2025

Hasil evaluasi menunjukkan bahwa $K = 3$ memberikan hasil paling optimal. Dikarenakan nilai *silhouette score* tertinggi (0,587) menunjukkan pemisahan kluster yang jelas, *Davies-Bouldin Index* terendah (0,5369) menunjukkan dispersi internal yang kecil antar kluster, dan penurunan *inertia* yang penting dari $K = 2$ ke $K = 3$ menguatkan pemilihan ini.

Dengan mempertimbangkan ketiga indikator tersebut, jumlah kluster optimal ditetapkan sebanyak tiga sehingga analisis ketimpangan pendidikan SMA akan dilakukan berdasarkan tiga kelompok provinsi. Pengelompokan ini memberikan dasar untuk membedakan provinsi berdasarkan kesamaan karakteristik pendidikan yang kemudian dapat dimanfaatkan sebagai kerangka kerja dalam menyusun kebijakan pendidikan yang lebih baik. Hasil klasterisasi ini tidak hanya menggambarkan variasi antarwilayah, tetapi juga membuka ruang bagi pemerintah untuk merancang kebijakan pendidikan yang lebih sesuai dengan kebutuhan *riil* masing-masing kelompok provinsi. Daripada menggunakan pendekatan seragam, segmentasi berbasis data seperti ini dapat menjadi dasar perumusan strategi afirmatif yang lebih terarah dan terukur.

Hasil Pengelompokan Wilayah Pendidikan Menggunakan *K-Means*

Pada bagian ini, digunakan algoritma *K-Means Clustering* dengan bantuan pustaka *Scikit-Learn* untuk mengelompokkan data indikator pendidikan menengah atas antarprovinsi. Jumlah kluster yang digunakan ditentukan sebanyak tiga *cluster* ($K=3$) dengan parameter *random_state* sebesar 42 agar hasil pengelompokan tetap konsisten setiap kali dijalankan.

	APS 16-18	ANGKA MELEK AKSARA 15-24	RATA-RATA LAMA SEKOLAH	HARAPAN LAMA SEKOLAH	JUMLAH SEKOLAH SMA	JUMLAH GURU SMA	JUMLAH MURID SMA
0	-0.084240	0.297200	-0.054736	-0.064213	2.416119	2.425378	2.514488
1	-3.124441	-4.192139	-3.034396	-3.382138	-0.833830	-0.890930	-0.748388
2	0.205808	0.224859	0.196492	0.219410	-0.249900	-0.247489	-0.267537

Gambar 2. Algoritma *K-Means*
Sumber: Diolah Peneliti, 2025

Hasil analisis *cluster* menggunakan *K-Means* ($K=3$) menunjukkan adanya tiga kelompok provinsi yang berbeda secara signifikan dalam hal ketimpangan pendidikan SMA.

- *Cluster 0* mencakup provinsi dengan kapasitas pendidikan tertinggi ditunjukkan oleh jumlah sekolah, guru, dan murid SMA yang jauh di atas rata-rata nasional. Kondisi ini mencerminkan kapasitas penyediaan pendidikan yang lebih baik sehingga ketimpangan antar wilayah relatif kecil.
- *Cluster 1* merupakan kelompok dengan keterbatasan paling besar terutama pada indikator APS, angka melek aksara, rata-rata dan harapan lama sekolah yang jauh di bawah nasional, serta ketersediaan sekolah dan guru yang minim. *Cluster* ini banyak berasal dari wilayah timur Indonesia menunjukkan kebutuhan intervensi yang lebih mendesak dalam penguatan akses dan infrastruktur pendidikan.

- *Cluster 2* mencakup provinsi dengan capaian pendidikan yang cukup baik meskipun infrastruktur tidak sebesar *cluster 0* berisi provinsi dengan capaian pendidikan yang cukup baik, meskipun tidak sekuat *cluster 0* dalam hal infrastruktur. *Cluster* ini mencerminkan wilayah dengan kondisi pendidikan yang stabil, namun masih menghadapi keterbatasan dalam hal kapasitas layanan.

Pengelompokan wilayah ini memberikan dasar penting untuk memahami ketimpangan pendidikan antar provinsi secara geografis. Hasil klasterisasi dapat dimanfaatkan sebagai acuan dalam perumusan kebijakan pendidikan yang lebih relevan dengan menyesuaikan intervensi terhadap kebutuhan dan karakteristik masing-masing wilayah. Hasil klasterisasi menggunakan algoritma *K-Means* membagi seluruh provinsi di Indonesia ke dalam tiga kelompok berdasarkan tujuh indikator pendidikan menengah atas yang mencerminkan aspek akses, capaian, dan kapasitas layanan. Pengelompokan ini membantu mengidentifikasi ketimpangan pendidikan secara spasial di tingkat nasional yang belum banyak dibahas. Berikut ini hasilnya:

Tabel 2. Hasil *Clustering* Provinsi

PROVINSI	CLUSTER	PROVINSI	CLUSTER
Aceh	2	Kalimantan Barat	2
Sumatera Utara	0	Kalimantan Tengah	2
Sumatera Barat	2	Kalimantan Selatan	2
Riau	2	Kalimantan Timur	2
Jambi	2	Kalimantan Utara	2
Sumatera Selatan	2	Sulawesi Utara	2
Bengkulu	2	Sulawesi Tengah	2
Lampung	2	Sulawesi Selatan	2
Kep. Bangka Belitung	2	Sulawesi Tenggara	2
Kep. Riau	2	Gorontalo	2
Dki Jakarta	2	Sulawesi Barat	2
Jawa Barat	0	Maluku	2
Jawa Tengah	0	Maluku Utara	2
Di Yogyakarta	2	Papua Barat	2
Jawa Timur	0	Papua Barat Daya	2
Banten	2	Papua	2
Bali	2	Papua Selatan	2
Nusa Tenggara Barat	2	Papua Tengah	1
Nusa Tenggara Timur	2	Papua Pegunungan	1

Sumber: Data diolah penulis hasil clustering

Berdasarkan Tabel 2, *cluster* 0 mencakup provinsi-provinsi dengan dukungan infrastruktur pendidikan yang relatif tinggi baik dari sisi jumlah sekolah, guru, maupun murid SMA. Wilayah ini umumnya berada di Indonesia bagian barat, dan mencerminkan sebaran layanan pendidikan yang lebih merata sebagaimana juga dicatat oleh Arfan & Pekei (2025). Sebaliknya, *cluster* 1 terdiri dari provinsi dengan kondisi paling tertinggal ditandai oleh rendahnya angka partisipasi sekolah, keterbatasan guru, serta akses fasilitas pendidikan yang sangat terbatas. Wilayah-wilayah ini sebagian besar berada di kawasan timur dan pedalaman selaras dengan temuan Risna Fadillah & Respy Desmaryani (2025) serta Melyana R Pugu & Elmiwati (2025) yang menyoroti disparitas akses antara kota dan daerah terpencil. *Cluster* 2 merupakan kelompok menengah yang relatif stabil mencerminkan daerah dengan kondisi pendidikan yang moderat namun belum sepenuhnya optimal seperti juga dicatat dalam studi Lenama et al., (2023).

Temuan ini memperkuat pandangan bahwa ketimpangan pendidikan di Indonesia bersifat struktural dan berlangsung dalam dimensi spasial yang luas. Klasterisasi dalam penelitian ini tidak hanya memvisualkan disparitas, tetapi juga berkontribusi dalam pengembangan pendekatan berbasis wilayah untuk menyusun kebijakan pendidikan yang lebih responsif. Studi terdahulu seperti Mashitawai et al., (2025) dan Kamalia & Nawangsih (2025) telah membuktikan efektivitas teknik *cluster* dalam lingkup lokal, namun belum menyentuh tataran nasional secara komprehensif. Melalui pendekatan ini, ketimpangan antar wilayah juga dapat dipahami dalam kerangka *Human Development Theory* (Masouras et al., 2023), di mana pendidikan menjadi salah satu dimensi utama pembangunan manusia. Ketimpangan pada indikator seperti rata-rata lama sekolah dan harapan lama sekolah dalam hasil *cluster* ini merefleksikan kesenjangan kualitas hidup antar wilayah yang lebih luas. Oleh karena itu, hasil klasterisasi ini dapat menjadi dasar perumusan kebijakan pendidikan yang tidak seragam melainkan adaptif terhadap kebutuhan dan kondisi spesifik setiap klaster wilayah.

Pemetaan Ketimpangan Pendidikan SMA di Indonesia

Peta berikut menggambarkan hasil klasterisasi ketimpangan pendidikan SMA di seluruh provinsi Indonesia. Setiap provinsi diwarnai sesuai dengan *cluster* yang mereka masuki yang menunjukkan tingkat ketimpangan pendidikan berdasarkan berbagai indikator seperti Angka Partisipasi Sekolah (APS), Jumlah Sekolah, dan Jumlah Guru SMA.



Gambar 3. Pemetaan Klaster Ketimpangan Pendidikan SMA di Indonesia
Sumber: Diolah Peneliti, 2025

Visualisasi peta pada Gambar 3 memberikan gambaran nyata mengenai pola ketimpangan pendidikan menengah atas di Indonesia secara geografis. Peta ini memperjelas bagaimana provinsi-provinsi terdistribusi ke dalam tiga klaster berdasarkan indikator akses, kualitas, dan kapasitas layanan pendidikan. Visualisasi seperti ini sangat penting untuk melihat sejauh mana ketimpangan terjadi antarwilayah, serta sebagai dasar untuk menyusun kebijakan pendidikan yang lebih sesuai kondisi daerah dan adaptif. Melalui peta ini, distribusi klaster antarprovinsi dapat diamati dengan jelas, memberikan visual terhadap kompleksitas dan sebaran ketimpangan yang tidak selalu tampak dari angka statistik semata.

Terlihat bahwa *cluster* 1 (warna hijau terang) mencakup provinsi-provinsi di wilayah timur Indonesia seperti Papua Pegunungan dan Papua Selatan yang menunjukkan kondisi pendidikan paling tertinggal. Provinsi dalam *cluster* ini memiliki nilai terendah dalam hampir seluruh indikator dari angka partisipasi sekolah, melek aksara, hingga ketersediaan sekolah dan guru. Temuan ini menegaskan bahwa pemerintah harus memberikan prioritas tinggi pada daerah yang tergolong dalam *cluster* tersebut melalui intervensi afirmatif yang berbasis kebutuhan nyata. Upaya yang dapat ditempuh mencakup percepatan pembangunan infrastruktur pendidikan, penataan distribusi guru yang lebih proporsional, serta penyediaan skema insentif yang terukur bagi tenaga pendidik di wilayah terpencil guna mengurangi disparitas layanan pendidikan.

Cluster 2 (warna kuning) merupakan kelompok dengan jumlah provinsi terbesar dan tersebar di berbagai wilayah Indonesia. *Cluster* ini menggambarkan daerah dengan kinerja pendidikan yang relatif konsisten tetapi masih berada di bawah potensi maksimalnya. Untuk meningkatkan performa *cluster* ini, intervensi yang berorientasi pada penguatan kualitas seperti pengembangan kompetensi guru secara berkelanjutan, modernisasi fasilitas pembelajaran, serta pemanfaatan teknologi pendidikan yang terukur dapat menjadi strategi untuk mendorong peningkatan capaian secara sesuai dengan daerahnya.

Sementara itu, *Cluster* 0 (warna hijau kebiruan) terdiri atas sedikit provinsi yang menempati posisi tertinggi dalam capaian indikator pendidikan termasuk Sumatera Utara, Jawa Barat, Jawa Tengah, dan Jawa Timur. Provinsi-provinsi ini menunjukkan kapasitas kelembagaan dan sumber daya pendidikan yang sangat kuat. Kendati demikian, konsistensi performa perlu terus dijaga melalui kebijakan yang fokus pada pemeliharaan kualitas, antisipasi stagnasi, dan eksplorasi inovasi baru. Selain itu, *cluster* ini memiliki potensi untuk berperan sebagai rujukan pengembangan dengan catatan bahwa adopsi praktik terbaik perlu disesuaikan dengan kapasitas masing-masing daerah.

Faktor Penyebab Ketimpangan Pendidikan pada Wilayah Tertinggal

Berdasarkan hasil klasterisasi dalam penelitian ini, provinsi-provinsi dalam *Cluster* 1 menunjukkan capaian pendidikan yang paling rendah dibanding kelompok lainnya. Wilayah seperti Papua Pegunungan dan Papua Selatan secara konsisten menempati posisi terbawah dalam hampir seluruh indikator yang dianalisis mulai dari partisipasi sekolah usia 16–18 tahun, angka melek aksara, rata-rata dan harapan lama sekolah, hingga jumlah sekolah dan guru SMA. Ketimpangan ini tidak terbatas pada satu aspek layanan, tetapi bersifat menyeluruh dan saling berkaitan yang mencerminkan hambatan struktural yang telah berlangsung lama. Salah satu faktor utama yang berkontribusi terhadap kondisi ini adalah tantangan geografis dan keterisolasian wilayah di mana provinsi-provinsi tersebut umumnya berada pada kawasan dengan topografi rumit, infrastruktur terbatas, dan jarak tempuh antar permukiman yang jauh. Hal ini berdampak pada minimnya pembangunan sarana pendidikan dan ketimpangan distribusi guru terutama mereka yang memiliki kompetensi dan kualifikasi yang memadai. Di sisi lain, kapasitas fiskal pemerintah daerah yang rendah juga memperlemah upaya perluasan akses dan peningkatan mutu pendidikan sehingga layanan yang tersedia tidak cukup menjangkau kebutuhan riil masyarakat di wilayah tersebut.

Ketimpangan tersebut dapat dijelaskan melalui teori ketimpangan regional yang dikemukakan oleh Todaro, M. P., & Smith (2012) yang menekankan bahwa perbedaan kemajuan antarwilayah terjadi karena distribusi sumber daya, infrastruktur, dan layanan publik yang tidak merata. Dalam pendidikan, ketika perhatian pembangunan lebih banyak diarahkan ke wilayah-wilayah yang secara ekonomi lebih maju dan mudah dijangkau, maka daerah-daerah dengan tantangan geografis akan terus mengalami ketertinggalan. Ketidakseimbangan ini tidak hanya berdampak pada rendahnya angka partisipasi atau kualitas layanan pendidikan, tetapi juga memperlebar kesenjangan sosial dan ekonomi antarwilayah dalam jangka panjang. Dengan kata lain, pendidikan yang tidak merata berpotensi melanggengkan ketimpangan pembangunan manusia antarprovinsi.

Temuan ini diperkuat oleh hasil penelitian Wardhana et al., (2023) yang menyatakan bahwa ketimpangan pendidikan di Indonesia erat kaitannya dengan ketimpangan ekonomi, tingkat putus sekolah, dan rendahnya belanja pemerintah daerah untuk sektor pendidikan. Penelitian tersebut menjelaskan bahwa daerah-daerah dengan pendapatan rendah cenderung memiliki kemampuan terbatas untuk meningkatkan mutu pendidikan di wilayahnya. Kondisi ini selaras dengan temuan klusterisasi pada penelitian ini yang memperlihatkan bahwa provinsi-provinsi dengan tingkat pendidikan terendah juga merupakan wilayah dengan kapasitas fiskal dan pembangunan yang lemah. Dengan demikian, faktor ekonomi dan kebijakan publik sangat mempengaruhi peluang suatu daerah dalam meningkatkan akses dan kualitas pendidikan secara merata.

Dalam situasi seperti ini, pendekatan satu kebijakan untuk seluruh wilayah menjadi tidak relevan. Daerah-daerah yang masuk dalam *cluster* 1 memerlukan intervensi yang dirancang secara khusus berbasis data dan sesuai dengan kondisi lokal. Pemerintah pusat dapat mempertimbangkan pemberian insentif khusus untuk guru di daerah terpencil, memperkuat pembangunan infrastruktur sekolah yang sesuai dengan kondisi geografis, serta memperluas dukungan dana afirmatif bagi daerah dengan capaian pendidikan yang masih tertinggal. Selain itu, pelibatan pemerintah daerah dalam perencanaan dan pelaksanaan program pendidikan menjadi penting agar intervensi yang diberikan tepat sasaran. Pendekatan seperti ini akan lebih efektif daripada kebijakan seragam, karena mempertimbangkan karakteristik dan tantangan nyata di wilayah yang paling tertinggal.

Rekomendasi Kebijakan Strategis untuk Pemerataan Pendidikan Menengah Atas

Perbedaan karakteristik antar klaster yang ditunjukkan melalui hasil analisis klusterisasi menegaskan bahwa kebijakan pendidikan tidak dapat dirancang secara seragam untuk seluruh wilayah. Setiap klaster memiliki kombinasi tantangan dan potensi yang berbeda baik dari sisi akses, kualitas, maupun kapasitas penyelenggaraan pendidikan. Oleh karena itu, diperlukan formulasi rekomendasi kebijakan yang dirancang secara terarah sesuai dengan kebutuhan dan kondisi masing-masing kelompok provinsi yang tergabung dalam klaster tertentu.

Cluster 0: Program Konsolidasi Infrastruktur dan Ekosistem Sekolah Menengah

Kebijakan yang direkomendasikan untuk klaster ini adalah Program Konsolidasi Infrastruktur dan Ekosistem Sekolah Menengah yaitu kebijakan yang berfokus pada pembangunan ulang fondasi layanan pendidikan di tingkat SMA. Kebijakan ini mencakup pembangunan sarana fisik sekolah, modernisasi ruang belajar, dan penyediaan laboratorium sains dasar, dengan prioritas pemerataan kualitas fasilitas antar sekolah. Model kebijakan ini mengacu pada pendekatan "*Basic Infrastructure Guarantee*" yang diterapkan Tiongkok pada wilayah pedesaan tertinggal, di mana pemerintah menetapkan standar minimum nasional bagi sarana pendidikan menengah (Daniel B.S. Ang, 1999). Dalam jangka pendek, kebijakan ini memastikan seluruh sekolah mencapai standar minimum kelayakan belajar. Dalam jangka panjang, strategi ini menurunkan kesenjangan mutu antardaerah dan memungkinkan siswa mengikuti proses belajar dalam lingkungan yang setara.

Cluster 1: Teacher Mobility & Professional Rotation Program

Kebijakan yang direkomendasikan untuk Klaster 1 adalah *Teacher Mobility & Professional Rotation Program*, yaitu sistem rotasi guru berkinerja tinggi yang diterapkan lintas kabupaten/kota maupun antar provinsi untuk memperkuat pemerataan kualitas pengajaran di SMA. Program ini mengadaptasi praktik Korea Selatan yang mewajibkan guru berpindah sekolah setiap beberapa tahun untuk menghindari konsentrasi mutu guru di wilayah tertentu (Wang, 2022), serta pendekatan Singapura yang menempatkan *master teachers* di sekolah-sekolah yang memerlukan peningkatan kapasitas pedagogis secara sistematis (Guo & Li, 2024).

Pelaksanaan program ini dimulai dari penyusunan indikator performa guru berbasis data untuk mengidentifikasi pendidik dengan kinerja unggul. Guru-guru terpilih kemudian ditempatkan secara berkala pada SMA yang membutuhkan penguatan kualitas pengajaran, dengan jaminan bahwa sekolah asal tidak akan mengalami kekurangan tenaga pendidik. Untuk menjaga keberlanjutan partisipasi, guru yang mengikuti rotasi diberikan insentif berupa percepatan karier, akses prioritas ke pelatihan profesional, serta tambahan tunjangan sesuai lokasi penugasan. Selama masa penugasan, guru tidak hanya mengajar tetapi juga bertindak sebagai mentor melalui kegiatan *lesson study*, pembinaan rekan sejawat, dan pengembangan metode pembelajaran, yang memungkinkan peningkatan kapasitas sekolah lokal berlangsung meskipun rotasi telah selesai.

Keberhasilan kebijakan ini dievaluasi melalui peningkatan capaian belajar siswa, penguatan kualitas praktik pembelajaran, serta perubahan indikator ketimpangan seperti APS SMA atau disparitas mutu antarsekolah. Dalam jangka panjang, skema rotasi guru ini diharapkan mampu mengurangi ketimpangan kompetensi pedagogis antarwilayah, mempercepat penyebaran praktik pembelajaran efektif, dan mewujudkan ekosistem pendidikan menengah yang lebih setara di seluruh Indonesia.

Cluster 2: Program Akselerasi Keadilan Akses SMA berbasis Digital dan Bantuan Individu

Kebijakan untuk klaster ini adalah Program Akselerasi Keadilan Akses SMA Berbasis Digital dan Bantuan Individu, yang berfokus pada pengurangan hambatan ekonomi dan geografis dalam mengikuti pendidikan menengah. Elemen kebijakan meliputi bantuan biaya pendidikan model “*Merit-Needs Hybrid Scholarship*”, kelas jarak jauh sinkron, serta platform pembelajaran nasional yang menyerupai sistem *digital learning* Korea Selatan dan Tiongkok. Kebijakan ini didasarkan pada hasil implementasi “*Smart Education Korea*” dan “*China Smart Classroom Reform*” (Gopinathan & Lee, 2018). Strategi jangka pendeknya memastikan siswa tidak terhambat biaya maupun jarak. Dampak jangka panjangnya adalah peningkatan partisipasi SMA secara merata, khususnya pada provinsi terpencil atau berbiaya hidup rendah.

E. KESIMPULAN

Penelitian ini menunjukkan bahwa ketimpangan pendidikan menengah atas di Indonesia merupakan permasalahan yang bersifat struktural dan spasial, dipengaruhi oleh ketidakseimbangan ekonomi, distribusi tenaga pendidik, serta ketersediaan infrastruktur pendidikan antarwilayah. Melalui penerapan algoritma *K-Means Clustering* terhadap tujuh indikator pendidikan dari data BPS tahun 2024, diperoleh tiga kelompok provinsi dengan karakteristik yang berbeda. Provinsi di wilayah barat umumnya memiliki capaian pendidikan tinggi dengan dukungan fasilitas dan sumber daya yang memadai, sementara wilayah tengah menunjukkan kondisi menengah yang stabil, dan wilayah timur cenderung tertinggal dengan akses serta partisipasi pendidikan yang rendah. Pola ini menegaskan adanya ketimpangan

regional dalam pemerataan mutu pendidikan menengah atas yang membutuhkan kebijakan pembangunan berbasis data dan diferensiasi wilayah.

Hasil penelitian ini memberikan implikasi praktis bagi perumusan kebijakan pendidikan nasional. Pemerintah perlu mengimplementasikan strategi pemerataan pendidikan melalui optimalisasi distribusi tenaga pendidik, peningkatan sarana dan prasarana pendidikan di daerah tertinggal, serta pemanfaatan teknologi pembelajaran digital untuk memperluas akses pendidikan berkualitas. Selain itu, sinergi antara pemerintah pusat dan daerah menjadi faktor penting dalam memperkuat implementasi kebijakan berbasis klaster wilayah. Penelitian selanjutnya disarankan untuk mengintegrasikan metode klasterisasi lain seperti *K-Medoids* atau *Hierarchical Clustering* serta menambahkan variabel sosial ekonomi, alokasi anggaran pendidikan, dan indeks pembangunan manusia guna memperluas cakupan analisis dan menguatkan dasar temuan secara empiris. Dengan demikian, hasil riset ini diharapkan dapat menjadi landasan strategis bagi kebijakan pemerataan pendidikan menengah atas yang berkeadilan, terukur, dan berkelanjutan di seluruh Indonesia.

REFERENSI

- Aditomo, A., Fakultas, P. D., Universitas, P., Felicia, N., Fakultas, P. D., & Universitas, P. (2018). Ketimpangan Mutu Dan Akses Pendidikan Di Indonesia : Potret Berdasarkan Survei Pisa 2015 Metode Analisis. *Kilas Pendidikan*, 75(3), 417–453. https://www.researchgate.net/publication/327261461_Ketimpangan_Mutu_Dan_Akses_Pendidikan_Di_Indonesia_Potret_Berdasarkan_Survei_Pisa_2015
- Anwar, M. S. (2022). Ketimpangan Aksesibilitas Pendidikan Multikultural Pendidikan Dalam Perspektif. *Foundasia*, 13(1), 1–15. <https://doi.org/10.21831/Foundasia.V13i1.47444>
- Arfan, U., & Pekei, Y. (2025). Evaluasi Distribusi Guru-Siswa Dan Ketersediaan Sekolah Untuk Mendukung Pembangunan Pendidikan Menggunakan K-Means Clustering. *Malcom: Indonesian Journal Of Machine Learning And Computer Science*, 5(4), 1508–1516. <https://doi.org/10.57152/Malcom.V5i4.207>
- Badan Pusat Statistik. (2024a). *[Metode Baru] Harapan Lama Sekolah (Tahun)*, 2024. Badan Pusat Statistik. <https://www.bps.go.id/id/statistics-table/2/Nde3izi=-Metode-Baru--Harapan-Lama-Sekolah--Tahun-.html>
- Badan Pusat Statistik. (2024b). *[Metode Baru] Rata-Rata Lama Sekolah (Tahun)*, 2024. Badan Pusat Statistik. <https://www.bps.go.id/id/statistics-table/2/Nde1izi=-Metode-Baru-Rata-Rata-Lama-Sekolah.html>
- Badan Pusat Statistik. (2024c). *Angka Melek Aksara Penduduk 15-24 Tahun Menurut Provinsi*, 2024. Badan Pusat Statistik. <https://www.bps.go.id/id/statistics-table/2/Mtq2mimy/Angka-Melek-Huruf-Penduduk-Umur-15-24-Tahun-Menurut-Provinsi.html>
- Badan Pusat Statistik. (2024d). *Angka Partisipasi Sekolah (Aps) Menurut Provinsi Dan Kelompok Umur*, 2024. Badan Pusat Statistik. <https://www.bps.go.id/id/statistics-table/2/Mjixmsmy/Angka-Partisipasi-Sekolah--Aps--Menurut-Provinsi-Dan-Kelompok-Umur.html>
- Badan Pusat Statistik. (2024e). *Jumlah Sekolah, Guru, Dan Murid Sekolah Menengah Atas (Sma) Di Bawah Kementerian Pendidikan, Kebudayaan, Riset, Dan Teknologi Menurut Provinsi*, 2024/2025. Badan Pusat Statistik. <https://www.bps.go.id/id/statistics-table/3/Ytfsrmnubehowe5ztuzsdwvhohhmmfppwm5vmfp6mdkjmwm==/Jumlah-Sekolah--Guru--Dan-Murid-Sekolah-Menengah-Atas--Sma--Di-Bawah-Kementerian-Pendidikan--Kebudayaan--Riset--Dan-Teknologi-Menurut-Pr>

- ovinsi--2024.Html?Year=2024
- Becker, G. S. (1964). Human Capital: A Theoretical And Empirical Analysis, With Special Reference To Education. *University Of Chicago Press*.
<https://doi.org/10.7208/Chicago/9780226041223.001.0001>
- Daniel B.S. Ang, A.-Y. C. And C.-K. L. (1999). Virtual Project-Based Learning : A Case Study Of Hong Kong And Singapore Students' Collaboration. *Mera-Era Joint Conference, Malacca, Malaysia*, 1462–1472.
<https://repository.nie.edu.sg/server/api/core/bitstreams/48edc94a-b5c7-409a-a4e3-17007640bba8/content>
- Dihe, L., & Wangdra, Y. (2023). Pendidikan Adalah Faktor Penentu Daya Saing Bangsa. *Prosiding Seminar Nasional Ilmu Sosial Dan Teknologi*, 12(2), 84–90.
- Gopinathan, S., & Lee, M. H. (2018). Excellence And Equity In High-Performing Education Systems : Policy Lessons From Singapore And Hong Kong / Excelencia Y Equidad En Sistemas Educativos De Alto Rendimiento : Lecciones De Las Políticas Educativas En Singapur Y Hong Kong. *Infancia Y Aprendizaje Journal For The Study Of Education And Development*, 41(2), 203–247. <https://doi.org/10.1080/02103702.2018.1434043>
- Guo, Y., & Li, X. (2024). Regional Inequality In China ' S Educational Development : An Urban-Rural Comparison. *Heliyon*, 10(4), 1–14.
<https://doi.org/10.1016/j.heliyon.2024.E26249>
- Indriaty, D., Astriani, D., Sabrifha, E., & Aqilla, L. (2025). Strategi Kebijakan Pendidikan Dalam Menghadapi Ketimpangan Sosial Dan Budaya. *Journal Home Page*, 5(2), 551–565. <https://jurnal.permapendis-sumut.org/index.php/pema>
- Kamalia, A. Z., & Nawangsih, I. (2025). Identifikasi Pola Tingkat Kesenjangan Ketuntasan Pendidikan Di Indonesia Dengan Menggunakan Metode K-Medoids Cluster Ing Identification Of Patterns In Educational Completion Disparities In Indonesia Using The K-Medoids Clustering Method. *Jurnal Teknologi Informasi Dan Ilmu Komputer (Jtiik)*, 12(2), 321–330. <https://doi.org/10.25126/jtiik.2025129219>
- Makkawaru, M. (2019). Pentingnya Pendidikan Bagi Kehidupan Dan Pendidikan Karakter Dalam Dunia Pendidikan. *Jurnal Konsepsi*, 8(3), 116–119.
<https://www.p3i.my.id/index.php/konsepsi/article/view/87>
- Mashitawai, R. K., Nugrahaeni, W., Putri, B., Cahyoko, F. D., Maret, U. S., Surakarta, K., & Tengah, P. J. (2025). Penerapan Algoritma K-Means Dan K-Medoids Untuk Klasterisasi Wilayah Berdasarkan Indikator Kualitas Pendidikan Di Provinsi Nusa Tenggara Timur Tahun 2023. *Prosiding Seminar Nasional Sains Dan Teknologi Seri Iv*, 2(2), 48–57.
<https://conference.ut.ac.id/index.php/saintek/article/view/6448>
- Masouras, A., Daskou, S., Pistikou, V., & Dimitriou, D. (2023). Global Developments In Nation Branding And Promotion: Theoretical And Practical. *Igi Global*.
- Melyana R Pugu, Elmiwati, R. (2025). Ketimpangan Pendidikan Di Indonesia: Studi Kebijakan Dan Evaluasi Inovasi Untuk Menjembatani Perbedaan Kelas Sosial. *Jurnal Pembelajaran Dan Pengembangan Diri*, 5(3). <https://doi.org/10.47353/bj.v5i3.610>
- Nony Lenama, Maria A. Kleden, I. G. P. (2023). K-Means Clustering Analysis Pada Pengelompokan Kabupaten/Kota Di Provinsi Nusa Tenggara Timur Berdasarkan Indikator Pendidikan. *Jurnal Cakrawala Ilmiah*, 2(9), 3365–3376.
<http://bajangjournal.com/index.php/jci>
- Pemerintah Republik Indonesia. (2003). *Undang-Undang Nomor 20 Tahun 2003 Tentang Sistem Pendidikan Nasional*.
- Risna Fadillah, Respy Desmaryani, A. L. (2025). Analisis Ketimpangan Sarana Dan Prasarana Pendidikan Di Daerah Pedesaan. *Jurnal Multidisiplin*, 03(02), 217–225.
<https://e-journal.naurendigiton.com/index.php/jam/article/view/1772>

- Rohma, G., Asyafiiyah, U., Widyastuti, A., & Andriani, F. (2025). Analisis Wilayah Prioritas Pembangunan Di Provinsi Jawa Timur Berdasarkan Indikator Sosial Menggunakan Metode K-Means Clustering. *Aritech: Journal Of Computer Science And Artificial Intelligence*, 5(1), 32–48. <https://doi.org/10.29240/Arcitech.V5i1.13488>
- Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-Means Clustering Algorithm. *Ieee Access*, 8(1), 80716–80727. <https://doi.org/10.1109/Access.2020.2988796>
- Siroj, R. A., Afgani, W., Septaria, D., Zahira, G., Kuantitatif, P., Ilmiah, P., & Data, A. (2024). Metode Penelitian Kuantitatif Pendekatan Ilmiah Untuk Analisis Data. *Jurnal Review Pendidikan Dan Pengajaran*, 7(3), 11279–11289. <http://journal.universitaspahlawan.ac.id/index.php/jrpp>
- Todaro, M. P., & Smith, S. C. (2012). Economic Development (12th Ed.). *Addison-Wesley*.
- Vega, A., P, J. A., & A, M. R. (2024). Kesetaraan Akses Pendidikan : Analisis Pengimplementasian Nilai Pancasila Dalam Pemerataan Akses Pendidikan Di Indonesia. *Jurnal Lentera Ilmu (Jli)*, 1(2), 44–57. <https://doi.org/10.59971/Li.V1i2.51>
- Wang, R. (2022). The Implementation Dilemma And Solution Of Teacher Exchange Rotation In Compulsory Education. *International Journal Of Education And Humanities*, 4(3), 219–222. <https://doi.org/10.54097/Ijeh.V4i3.1811>
- Wardhana, A., Kharisma, B., & Fauzy, M. Z. (2023). Determinan Ketimpangan Pendidikan Provinsi Di Indonesia. *Jurnal Ekonomi & Manajemen Universitas Bina Sarana Informatika*, 21(1), 101–111. <https://doi.org/10.31294/Jp.V2i1i1>