SHM PUBLISHER
SURYA HIJAU MANFAAT

Journal of Information System Exploration and Research

https://joiser.shmpublisher.com

p-ISSN 2964-1160 | e-ISSN 2963-6361

A Multi-Level Clustering Framework for Provincial Educational Facility Equalization in Indonesia

Antika Zahrotul Kamalia[1*], Zaenur Rozikin[2], Hemdani Rahendra Herlianto[3], Hendra Arya Syaputra[4], Asep Arwan Sulaeman[5], Choiriyatun Nisa Latansa[6]

[1,2,3,4,5,6]*Department of Informatics Engineering, Universitas Pelita Bangsa, Indonesia*

DOI: https://doi.org/10.52465/joiser.v4i2.13
Received 28 May 2026; Accepted 01 July 2026; Available online 02 July 2026

Article Info

Keywords:
Decision support;
Educational facility;
K-Means;
Multi-level clustering;
Provincial prioritization

Abstract

Equitable distribution of educational facilities is crucial for development planning, as regional disparities in facility availability can constrain access to education. This study identifies priority areas for school-facility equalization in Indonesia based on 2024 data covering 38 provinces and village/urban-ward-based facility availability by education level. Unlike single-stage clustering studies, this study combines macro-prioritization and micro-level need profiling to identify which provinces should be prioritized but also which education levels require attention. The analysis includes log1p transformation, standardization, optimal cluster selection using Elbow and Silhouette criteria, and the application of Level-1 and Level-2 K-Means clustering. The Level-1 results produce three priority groups: High Priority, Medium Priority, and Low Priority, with the optimal structure at K=3. The Level-2 analysis within the high-priority group is most stable at K2=2, distinguishing provinces dominated by primary and lower-secondary facility shares from those with a more balanced composition and relatively higher tertiary share. The Silhouette values indicate that the selected clusters provide reasonably separated groupings. The proposed framework provides a data-driven priority map and level-specific need profiles. The results can support staged infrastructure planning and differentiated interventions across provincial priority groups to strengthen educational facility equalization in Indonesia.

1. Introduction

Equitable access to education remains a central development priority in Indonesia, yet the distribution of educational services is still uneven across regions [1], [2]. Differences in geography, settlement patterns, and local development capacity can translate into unequal access to school

* ***Corresponding Author:***

Antika Zahrotul Kamalia,
Department of Informatics Engineering,
Universitas Pelita Bangsa, Indonesia,
Jl. Inspeksi Kalimalang No. 9, Cibatu, South Cikarang, Bekasi Regency, West Java 17530, Indonesia.
Email: antika.kamalia@pelitabangsa.ac.id

facilities, particularly in rural and remote areas [3]. Such inequality may not be fully captured by aggregate education indicators alone, because the practical barrier often lies in whether basic education services are physically available near where people live namely, whether villages/urban wards have access to primary, secondary, and post-secondary educational facilities.

This study is also positioned as a supply-side complement to prior outcome-oriented evidence. In our earlier provincial study on educational completion disparities using K-Medoids clustering (2018–2023), we found pronounced completion gaps in central and eastern Indonesia (e.g., East Nusa Tenggara and Papua), which reflect limited access to and availability of educational facilities and highlight the need for data-driven policy and infrastructure improvement [4]. These findings motivate the present work to directly measure and prioritize facility availability as a concrete basis for planning equalization interventions.

Clustering techniques have been widely used to profile educational conditions by grouping schools or regions with similar characteristics [5]. Prior studies have applied methods such as K-Means, K-Medoids, and Fuzzy C-Means to segment provinces or institutions based on education-related indicators, demonstrating that clustering supports targeted policy design by enabling differentiated interventions across groups rather than applying uniform strategies nationwide [6], [7], [8]. However, many existing studies still rely mainly on aggregated outcome indicators (e.g., attainment or participation) rather than facility-availability measures that more directly capture physical service coverage. In addition, most works adopt a single-stage clustering design that produces segmentation but provides limited diagnostic insight into which education levels (e.g., primary, secondary, tertiary) are most underserved within the highest-priority areas.

Based on the above literature, three gaps can be identified. First, research that uses village/urban-ward-based facility availability as a supply-side proxy for educational equity across multiple education levels remains limited. Second, clustering outputs are not always translated into operational priority categories (e.g., high/medium/low priority) that are directly usable for planning and budgeting. Third, the predominance of single-stage clustering means that many studies do not conduct a focused second-stage analysis within the highest-priority group to uncover distinct, level-specific shortage patterns.

To address these gaps, this study proposes a multi-level clustering framework that offers three main contributions. First, it adopts a supply-side educational equity perspective by using village/urban-ward-based facility availability indicators across education levels, rather than relying solely on outcome-based indicators such as participation, completion, or dropout rates. Second, it employs a two-stage clustering design in which Level-1 clustering maps provincial priority groups, while Level-2 clustering further diagnoses education-level-specific facility needs within the high-priority group using proportional-share features. Third, the framework produces both an operational priority classification and a level-specific need profile, enabling policymakers to identify not only which provinces should be prioritized but also which education levels require strengthening. Therefore, the proposed approach provides a more actionable basis for data-driven planning for educational facility equalization.

2. Literature Review

Educational inequality in Indonesia is territorially uneven and closely associated with differences in access to educational services across regions. Pronounced educational completion disparities in central and eastern Indonesia (e.g., East Nusa Tenggara and Papua), reflecting limited access to and availability of educational facilities and highlighting the need for data-driven interventions [4]. This motivates a complementary supply-side perspective: beyond outcome indicators (e.g., completion), equity planning should quantify the availability of educational facilities near where people live as a basis for prioritizing equalization policies.

Clustering has been widely applied in educational analytics to group regions or schools based on similar characteristics, using methods such as K-means, K-medoids, and fuzzy C-means. In general, clustering supports policy targeting by enabling differentiated strategies across clusters rather than uniform interventions [9], [10], [11]. However, many studies still rely on aggregated outcome-based indicators and use single-stage clustering, which often yields segmentation without a detailed diagnosis of level-specific facility needs (e.g., primary, secondary, or tertiary).

Table 1. Comparison of previous studies and research gaps

Author/Year	Focus	Method	Limitation	Gap Addressed by This Study
Kamalia and Nawangsih (2025) [4]	Educational completion disparity across Indonesian provinces	K-Medoids clustering	Focuses on outcome-based indicators, particularly educational completion, rather than direct facility availability	This study complements the outcome perspective by using village/urban-ward-based facility availability as a supply-side indicator
Hasnataeni et al. (2025) [10]	School accreditation clustering in West Java	K-Means, K-Medoids, and Fuzzy C-Means	Focuses on school accreditation and does not address educational facility equalization across provinces	This study applies clustering to support national-level facility equalization planning
Nurkhofifah et al. (2025) [11]	Junior high school education clustering based on density and dropout ratios	Quartile and K-Means methods	Limited to junior high school indicators and does not examine multi-level educational facilities	This study covers multiple education levels from primary to higher education
Vidyananta and Dermawan (2025) [8]	Clustering Indonesian provinces based on elementary school indicators	K-Means and K-Medoids	Focuses mainly on elementary school conditions and uses a single-stage clustering approach	This study uses a two-stage clustering framework to identify macro priorities and micro-level needs
Rahmawati and Fauzan (2024) [9]	Provincial clustering based on education indicators	K-Medoids and outlier handling	Provides regional segmentation but does not specifically diagnose facility needs by education level	This study provides priority classification and level-specific facility need profiling
Jiang et al. (2022) [24]	Spatial pattern of basic education resources in rural areas	Spatial analysis of education resources	Emphasizes spatial distribution but is not designed as a multi-stage provincial priority classification framework for Indonesia	This study integrates facility-availability indicators with policy-oriented priority mapping for Indonesian provinces

As summarized in Table 1. previous studies have contributed to educational clustering, regional segmentation, and spatial analysis of educational resources. However, three gaps remain evident. First, limited work uses village/urban-ward-based facility availability indicators as a supply-side proxy for multi-level educational equity. Second, clustering outputs are not consistently translated into operational priority categories suitable for planning, such as high-, medium-, and low-priority groups. Third, most studies rely on single-stage clustering and rarely perform a focused second-stage analysis within the highest-priority group to uncover distinct education-level shortage patterns. To address these gaps, this study proposes a multi-level clustering framework in which Level-1 identifies provincial priority groups, while Level-2 profiles education-level needs among high-priority provinces. This framework produces more actionable recommendations for educational facility equalization planning by identifying both which provinces should be prioritized and which education levels require strengthening.

3. Method

This section describes the research design, data sources, preprocessing procedures, feature engineering, clustering stages, and evaluation metrics used in this study. The methodological workflow consists of data collection, data quality checking, transformation and standardization, Level-1 clustering for provincial priority mapping, Level-2 clustering for high-priority profiling, and internal validation to assess cluster quality and interpretability.

4.1. Research design and stages

This study employs a quantitative design, secondary data, and an unsupervised learning approach. The main objective is to group Indonesian provinces based on the similarity of village/urban ward-level school facility availability profiles to derive operational priority categories for facility equalization. We implement a multi-level (multi-stage) clustering framework with two stages: (1) Level-1 macro clustering, which maps all provinces into priority groups (high/medium/low) using the counts of villages/urban wards with facilities at each education level (primary, lower secondary, upper secondary/general, vocational, and higher education). Before clustering, the data are cleaned and processed through log1p transformation and standardization; the optimal number of clusters (K) is selected using the Elbow (inertia) and Silhouette criteria. (2) Level-2 micro clustering is then applied only to high-priority provinces to reveal distinct, level-specific shortage patterns using proportion/share features across education levels; the number of subclusters (K2) is chosen based on the Silhouette Score. Clustering quality is assessed using internal validation metrics (Silhouette, DBI, and CHI) and supported by visual diagnostics (e.g., PCA 2D, cluster profiles, and share-composition plots) to ensure that the results are interpretable and actionable for equalization planning. The overall research workflow is presented in Figure 1.

Figure 1. Research stages

4.2. Data sources and research variables

The data used is secondary data from 2024, including the number of villages/sub-districts with school facilities by province and education level. The unit of analysis is the province (all provinces in Indonesia). The main variable is the number of villages with facilities at the following levels: elementary, middle, high school, vocational school, and university Table 2.

Table 2. Provinces that have school facilities

Province	Number of Villages with School Facilities by Province and Education Level				
	Elementary School (SD)	**Middle School (SMP)**	**High School(SMA)**	**Vocational School (SMK)**	**College/University**
	2024	**2024**	**2024**	**2024**	**2024**
Aceh	3382	1421	735	205	119
Sumatera Utara	5003	2319	1147	712	202
Sumatera Barat	1256	808	415	180	102
Riau	1811	1210	631	259	80
Jambi	1484	810	393	169	42
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
Maluku Utara	1090	616	294	142	25
Papua Tengah	430	135	50	36	16
Papua Pegunungan	530	169	48	15	8

The data were read from an Excel file and then pre-processed through: (1) harmonizing province names, (2) converting education-level columns into numeric values, and (3) checking for missing values and duplicates. If missing values were found, they were handled conservatively. To represent the overall capacity of multi-level facility availability, the total number of villages/urban wards with educational facilities was computed for each province as follows Formulas 1-4 [12], [13]:

4.2.1. Total facilities per province

$$T_i = \sum_{j=1}^{5} X_{ij} \tag{1}$$

where T_i is the total facilities in province i, and X_{ij} is the number of villages/urban wards with facilities at education level j (with $j = 1, \ldots, 5$ representing SD, SMP, SMA/SMU, SMK, and Higher Education). For Level-2 analysis (to capture level-specific shortage patterns among high-priority provinces), the proportional contribution of each education level to the total was calculated as:

4.2.2. Proportion (share) by education level

$$S_{ij} = \frac{X_{ij}}{T_i} \tag{2}$$

where S_{ij} is the share of level j in province i, X_{ij} is the facility count at level j, and T_i is the total facility count across all levels in province i.

Because the variables are count-based and may be skewed, a log transformation was applied:

4.2.3. log1p transformation for count data

$$X'_{ij} = \ln(1 + X_{ij}) \tag{3}$$

where X'_{ij} is the transformed feature value, $\ln(\cdot)$ is the natural logarithm, and the constant 1 ensures that zero values remain valid.

Finally, features were standardized using Z-score normalization to ensure comparability across different scales:

4.2.4. Z-score standardization

$$Z_{ij} = \frac{X'_{ij} - \mu_j}{\sigma_j} \tag{4}$$

where Z_{ij} is the standardized value for province i and feature/level j, X'_{ij} is the log1p-transformed value, μ_j is the mean of feature j, and σ_j is the standard deviation of feature j.

4.3. Level-1 clustering (macro clustering) using K-Means

K-Means partitions the data into K clusters by minimizing the within-cluster sum of squares (WCSS), also referred to as inertia. The objective function is formula 5 [14]:

4.3.1. K-Means objective function

$$\min \sum_{k=1}^{K} \sum_{x_i \in C_k} \| x_i - \mu_k \|^2 \tag{5}$$

where C_k denotes the set of data points assigned to cluster k, and μ_k is the centroid (mean vector) of cluster k. In Level-1, the feature vector for each province is $[SD, SMP, SMA/SMU, SMK, PT]$, representing the counts of villages/urban wards with educational facilities at each level. These count features are processed using log1p transformation and Z-score standardization, so that the clustering reflects similarity in multi-level facility availability profiles and is not dominated by variables with larger scales. The number of clusters K is selected using two complementary criteria: the Elbow method (based on inertia/WCSS) and the Silhouette score.

K-Means was selected as the main clustering algorithm because the objective of this study is to produce interpretable, centroid-based priority groups from a small provincial tabular dataset. The method is computationally simple, transparent, and suitable for policy-oriented segmentation, where each cluster can be interpreted through its centroid and average facility profile. Compared with more complex approaches such as Gaussian Mixture Models, Spectral Clustering, or DBSCAN, K-Means provides more direct and practical interpretation for classifying provinces into high-, medium-, and low-priority groups. Hierarchical clustering with Ward linkage is used as a supporting diagnostic tool through a dendrogram, while density-based methods such as DBSCAN are less suitable because the dataset is small and the objective is not to detect arbitrary-shape dense regions or outliers, but to generate clear priority categories for educational facility equalization planning.

4.4. Data preprocessing and quality check

Before clustering, several preprocessing and data quality checking steps were conducted to ensure that the dataset was suitable for analysis show Table 3. The initial dataset consisted of 38 provinces and five education-level variables: SD, SMP, SMA, SMK, and University. Province names were harmonized by removing unnecessary spaces, and all education-level columns were converted into numeric format. The data quality check showed that there were no missing values and no duplicate province records; therefore, no province was removed from the dataset. The final dataset remained 38 provinces.

After the quality check, feature engineering was performed by calculating the total number of villages/urban wards with school facilities across all education levels for each province. In addition, proportional share variables were computed to represent the contribution of each education level to the province's total facility availability. For Level-1 clustering, the count variables were transformed using log1p to reduce skewness and then standardized using Z-score normalization before applying K-Means. For Level-2 clustering, the share variables were standardized and used to identify education-level-specific patterns within the high-priority group.

Table 3. Data quality check summary

Check	Result	Action
Initial observations	38 provinces	Retained
Final observations	38 provinces	Used for clustering
Missing values	0	No imputation required
Duplicate province records	0	No record removed
Non-numeric education-level columns	Checked	Converted to numeric
Province names	Checked	Harmonized/trimmed
Feature engineering	Total and share variables	Computed
Transformation	log1p and standardization	Applied before clustering

4.5. Equity priority labeling (level-1 results)

The Level-1 K-Means clustering produces numeric cluster labels, which are then converted into High, Medium, and Low Priority categories based on the average total facility availability within each cluster. Specifically, for each province i, the total number of villages/urban wards with educational facilities across levels is denoted as T_i. Let $\bar{T}_k$ represent the mean of T_i among provinces assigned to cluster k. The cluster with the smallest $\bar{T}_k$ is labeled as High Priority (indicating lower overall facility availability), followed by Medium Priority, and the cluster with the largest $\bar{T}_k$ is labeled as Low Priority. As an optional policy-support measure, a standardized priority score (0–100) can be computed to facilitate ranking is formula 6 [15]:

4.5.1. Standardized priority score (0–100)

$$P_i = 100\left(1 - \frac{T_i - \min(T)}{\max(T) - \min(T)}\right) \tag{6}$$

where P_i is the priority score for province i (higher values indicate higher priority), T_i is the total facility count for province i, and $\min(T)$ and $\max(T)$are the minimum and maximum T_i values across all provinces.

4.6. Level-2 clustering (micro clustering) on high priority groups

Level-2 clustering is conducted to further differentiate the High-Priority provinces identified in Level-1 and to reveal education-level-specific shortage patterns within this group. First, the dataset is filtered to include only provinces classified as High Priority based on the Level-1 results. Unlike Level-1, which uses transformed and standardized facility counts, Level-2 uses proportion (share) features to capture the internal composition of facility availability across education levels.

Specifically, for each high-priority province i, a share vector is constructed as $[Share_{SD}, Share_{SMP}, Share_{SMA}, Share_{SMK}, Share_{HE}]$, where each element represents the proportion of villages/urban wards with facilities at the corresponding education level relative to the province's total facility count. These share features are then standardized to ensure comparable contributions across dimensions. The optimal number of subclusters (K_2) is determined using the Silhouette Score over a candidate range (e.g., $K_2 = 2$to 5). After selecting K_2, K-Means is applied to assign each high-priority province to a Level-2 subcluster.

The resulting subclusters are interpreted using their mean share profiles to identify which education levels tend to be relatively underrepresented within each subcluster. This micro-level profiling enables more targeted recommendations, for example, distinguishing high-priority provinces that are dominated by primary/lower-secondary facility shares from those with a more balanced composition or relatively higher tertiary shares.

4.7. Level-2 clustering (micro clustering) on high priority groups

Cluster quality assessment in this study relies on internal validation metrics, since the approach is unsupervised and no ground-truth labels are available for direct comparison. The clustering results are evaluated using three main indicators: the Silhouette Score, which measures within-cluster cohesion and between-cluster separation; the Davies–Bouldin Index (DBI), which assesses the ratio of within-cluster dispersion to between-cluster distances (lower values indicate better clustering); and the Calinski–Harabasz Index (CHI), which compares between-cluster dispersion to within-cluster dispersion (higher values indicate clearer cluster structure). In addition to these numerical metrics, evaluation is supported by PCA-based 2D projections to visually inspect cluster separability and by cluster profiling (mean feature patterns) to ensure that each cluster has interpretable characteristics in the context of educational facility equalization planning.

4. Results and Discussion

This section presents the results and discussion of the proposed multi-level clustering framework for mapping educational facility equalization priorities in Indonesia. The discussion begins with the data characteristics and the determination of the optimal number of clusters, followed by the Level-1 clustering analysis that classifies provinces into priority groups. Furthermore, Level-2 clustering is applied to provinces categorized as high priority to identify education-level-specific facility need patterns. The results are discussed through tables, visualizations, cluster profiles, and policy-oriented interpretations so that the findings can provide practical contributions to more targeted educational facility equalization planning.

4.1. Data description

The data used is the number of villages with school facilities by province and level of education in 2024 (elementary, junior high, senior high, vocational, and tertiary education) for 38 provinces. This variable represents the availability of village-based educational facilities (not the number of schools), making it relevant for assessing spatial equality in educational services across regions [16]. Descriptively, the total number of villages with school facilities per province shows a wide distribution (minimum 577 and maximum 18,089) and is skewed. Therefore, in data processing, log1p transformation and standardization were performed before clustering so that the Euclidean-based distance calculation in K-Means is not dominated by provinces with very large scales [17].

Table 4. Statistik total desa berfasilitas sekolah (T) per provinsi

Statistik	Nilai
n (provinsi)	38
Rata-rata T	3.782,5
Median T	2.760,5
Minimum T	577
Maksimum T	18.089

The summary in Table 4. shows a gap between provinces: some provinces have a very high total number of villages with school facilities, while others have numbers ranging from hundreds to thousands, indicating unequal potential access to village-based education services.

4.2. Determining the number of clusters (Level-1)

The number of clusters for Level-1 was selected by comparing the Elbow method (inertia/WCSS) and the Silhouette Score over $K = 2–6$. The results show that the Silhouette Score reaches its highest value at $K = 3$(≈ 0.493), therefore $K = 3$was chosen as the Level-1 macro-clustering structure because it provides the most balanced trade-off between within-cluster cohesion and between-cluster separation [18].

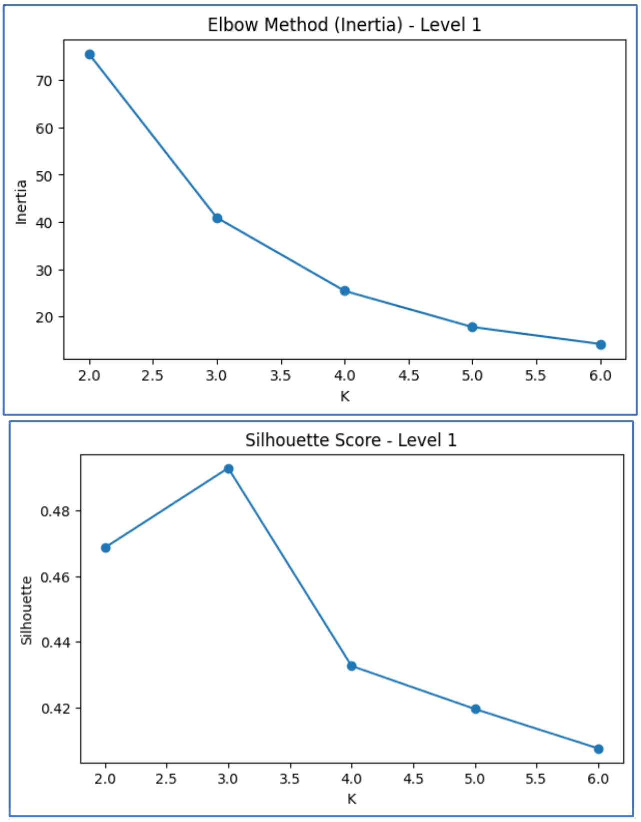

Figure 2. Elbow and silhouette score (level-1)

As shown in Figure 2. the Elbow plot indicates the steepest decrease in inertia from $K = 2$ to $K = 3$, followed by a more gradual decline for larger K, suggesting diminishing returns when additional clusters are introduced. This pattern is consistent with the Silhouette curve, which peaks at $K = 3$. Selecting $K = 3$ also supports clearer interpretation in subsequent analysis, as the resulting clusters can be operationalized into three priority categories (high/medium/low) for facility equalization planning. For additional validation, internal indices such as the Davies–Bouldin Index (Davies & Bouldin, 1979) and the Calinski–Harabasz Index (Caliński & Harabasz, 1974) may be reported in the evaluation subsection to further corroborate the selected clustering structure.

4.3. Robustness analysis of preprocessing and clustering scenarios

To examine the sensitivity of the clustering results to preprocessing choices, a robustness analysis was conducted by comparing several preprocessing and clustering scenarios. The evaluated scenarios included raw count data with Standard Scaler and K-Means, log1p-transformed data with Standard Scaler and K-Means, raw count data with Robust Scaler and K-Means, and log1p-transformed data with Standard Scaler and Agglomerative Clustering using Ward linkage. Each scenario was evaluated using three internal validation metrics: Silhouette Score, Davies–Bouldin Index (DBI), and Calinski–Harabasz Index (CHI).

Table 5. Robustness analysis of preprocessing and clustering scenarios

Scenario	Silhouette	Davies–Bouldin	Calinski–Harabasz
Raw + Standard Scaler + KMeans	0.572282	0.552142	133.727070
Raw + Robust Scaler + KMeans	0.565066	0.580505	141.497743
Log1p + Standard Scaler + KMeans	0.492766	0.601250	63.736114
Log1p + Standard Scaler + Agglomerative (Ward)	0.481363	0.543311	58.627409

The robustness results show Table 5. that the raw-count scenarios produce higher internal validation scores, particularly in terms of Silhouette and Calinski–Harabasz values. However, because the dataset consists of count-based facility indicators with a highly skewed distribution, relying directly on raw counts may cause provinces with very large administrative scales to dominate the clustering structure. Therefore, the log1p transformation with Standard Scaler was retained as the main preprocessing strategy. Although its internal validation score is slightly lower than the raw-count scenarios, it reduces the dominance of extreme values and produces more balanced, interpretable, and policy-relevant priority groups. The Ward-based agglomerative result was used as a supporting comparison and showed a broadly comparable cluster structure, indicating that the proposed priority grouping is not solely dependent on one clustering procedure.

4.4. Level-1 clustering results: equalization priority

Level-1 clustering produced three priority groups based on multi-level facility profiles (SD–HE) after log transformation and standardization. The clusters were then labeled using the mean total facility count $\boldsymbol{T}$ within each cluster: High Priority (lowest mean $\boldsymbol{T}$), Medium Priority, and Low Priority (highest mean $\boldsymbol{T}$). In this study, “priority” indicates that provinces with smaller total numbers of villages/urban wards having school facilities require greater attention for facility equalization planning.

Table 6. Summary of total facility count (T) by level-1 priority group

Level-1 Priority	Number of Provinces	Mean T	Min T	Max T
High Priority	15	999.4	577	1,989
Medium Priority	19	3,752.4	1,995	6,098
Low Priority	4	14,362.0	9,383	18,089

Table 6. shows a clear separation in the scale of total facility coverage. The High Priority group lies far below the overall provincial average, while the Low Priority group is dominated by provinces with very large Tvalues, suggesting broader village-based service coverage in aggregate (BPS, 2024).

Table 7. Comparative characteristics of level-1 clusters

Prioritas_L1	n_provinces	mean_total	mean_SD	mean_SMP	mean_SMA	mean_SMK	mean_HE
High Priority	15	999.4	541.3	248.2	109.3	70.1	30.6
Medium Priority	19	3752.4	1922.9	1046.4	485.5	225.4	72.2
Low Priority	4	14362	6954	3807.5	1817	1419	364.5

Table 7. shows that the High Priority group has the lowest average facility availability across all education levels, indicating that these provinces require greater attention in educational facility equalization planning. The Medium Priority group shows moderate facility availability, suggesting the need for targeted improvement in underserved areas, particularly at the SMP, SMA, and SMK levels. Meanwhile, the Low Priority group has the highest aggregate facility availability; therefore, policy

attention in this group should focus more on intra-provincial equity, service quality, and capacity improvement rather than merely increasing the number of facilities.

Table 8. Priority results (Level-1) per province (based on T)

Provinsi	Total desa berfasilitas (T)	Prioritas (Level-1)	Provinsi	Total desa berfasilitas (T)	Prioritas (Level-1)
Kalimantan Utara	577	High Priority	Sumatera Barat	2761	Medium Priority
Papua Tengah	667	High Priority	Kalimantan Tenggara	2897	Medium Priority
Papua Selatan	721	High Priority	Sulawesi Utara	3034	Medium Priority
Papua Barat	743	High Priority	Nusa Tenggara Barat	3072	Medium Priority
Papua Barat Daya	752	High Priority	Jambi	2898	Medium Priority
Kep. Bangka Belitung	758	High Priority	Sulawesi Tengah	3449	Medium Priority
Papua Pegunungan	770	High Priority	Sulawesi Tenggara	3395	Medium Priority
Kep. Riau	884	High Priority	Kalimantan Barat	3916	Medium Priority
Papua	1072	High Priority	Riau	3991	Medium Priority
DKI Jakarta	1074	High Priority	Banten	4179	Medium Priority
DI Yogyakarta	1100	High Priority	Lampung	5170	Medium Priority
Gorontalo	1201	High Priority	Sumatera Selatan	5477	Medium Priority
Sulawesi Barat	1340	High Priority	Aceh	5862	Medium Priority
Bali	1343	High Priority	Sulawesi Selatan	6073	Medium Priority
Bengkulu	1989	High Priority	Nusa Tenggara Timur	6098	Medium Priority
Kalimantan Timur	1995	Medium Priority	Sumatera Utara	9383	Low Priority
Maluku	2149	Medium Priority	Jawa Tengah	14976	Low Priority
Maluku Utara	2167	Medium Priority	Jawa Barat	15000	Low Priority
Kalimantan Tengah	2760	Medium Priority	Jawa Timur	18089	Low Priority

Table 8. Level-1 Priority Results by Province (based on T) lists provinces and their assigned Level-1 priority categories. The High Priority group contains many provinces with very low T values (including several provinces in the Papua region and provinces with smaller village bases), whereas the Low Priority group is dominated by provinces with very high T values (notably Java and North Sumatra). From a policy perspective, these results provide an initial screening tool to refine priority areas for equalizing the availability of village-based educational facilities.

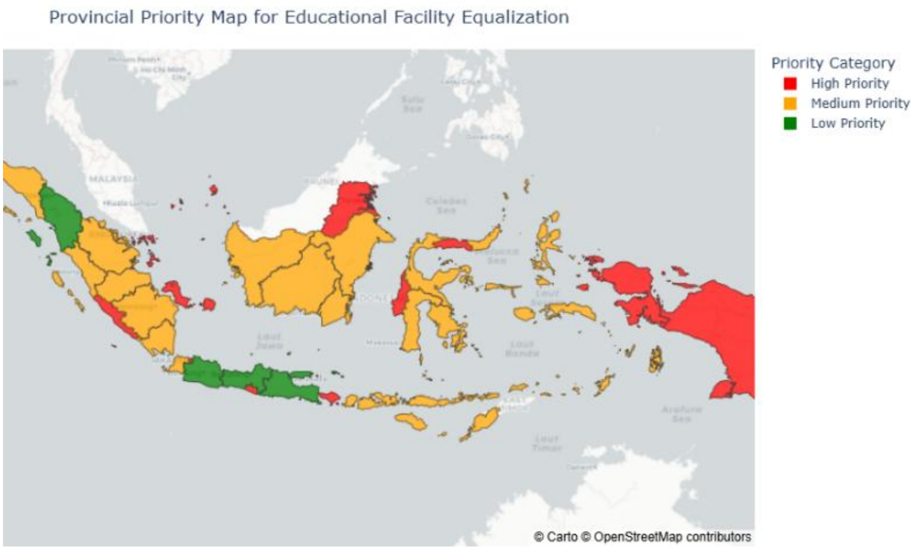

Figure 3. Provincial priority map for educational facility equalization based on level-1 clustering.

Figure 3. presents the spatial distribution of Level-1 priority categories across Indonesian provinces. The map shows that high-priority provinces are concentrated in several eastern and peripheral regions, including the Papua area and selected island provinces, while low-priority provinces are dominated by provinces with large aggregate facility availability such as Java and North Sumatra. This spatial visualization strengthens the interpretation of the clustering results by showing how educational facility equalization priorities vary across regions. Therefore, the map can serve as an initial decision-support tool for identifying provinces that require greater attention in educational facility planning.

However, the Level-1 priority categories should be interpreted with caution because the indicator used in this study is based on the number of villages/urban wards with facilities. This measure is sensitive to the total number of administrative units in each province. Highly urban provinces, such as DKI Jakarta and DI Yogyakarta, may show relatively low total facility counts because they have fewer village/urban-ward units, not necessarily because they experience severe facility shortages. Therefore, the priority categories should ideally be complemented with coverage rates, population density, school-age population, accessibility indicators, and facility quality measures to support more accurate policy interpretation.

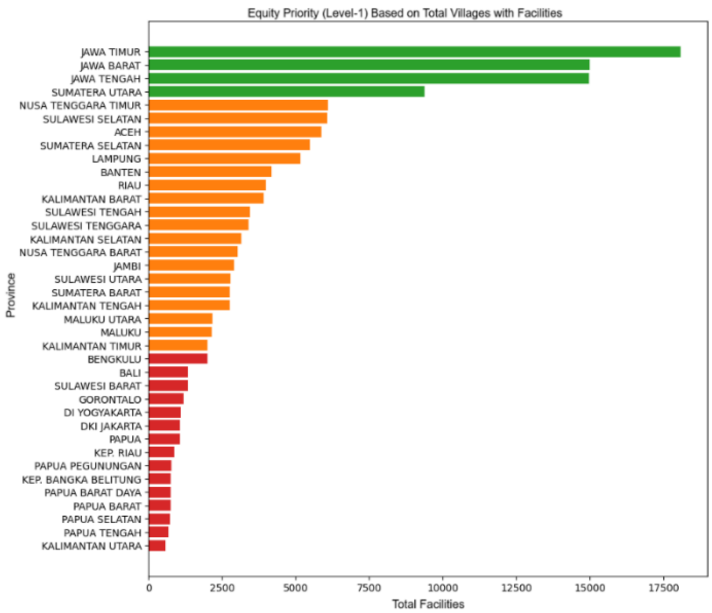

Figure 4. Level-1 clustering results

Figure 4. presents a provincial ranking based on the total number of villages/urban wards that have educational facilities across levels (SD, SMP, SMA/SMU, SMK, and higher education), with bar colors indicating the Level-1 clustering priority categories (high, medium, and low priority). Provinces with the highest totals are concentrated in the low-priority group (e.g., East Java, West Java, Central Java, and North Sumatra), suggesting broader aggregate coverage of village-based educational facilities compared with other provinces. In contrast, the high-priority group is dominated by provinces with relatively small totals (located at the lower part of the chart), making them key candidates for facility equalization efforts under a village-based availability perspective. This visualization strengthens the interpretation of the Level-1 K-Means results by making the scale differences across priority groups clearly visible and easy to communicate for planning purposes [19].

Nevertheless, because the indicator used is a count of villages/urban wards with facilities (rather than a ratio normalized by the total number of villages/urban wards, population size, or measures of quality/capacity), some highly urban provinces or provinces with fewer administrative village units may appear as high priority even though their infrastructure context differs [20]. Therefore, Figure 2 should be interpreted as an initial screening map based on facility availability and ideally complemented by additional indicators in future work or policy discussion, such as coverage rates (per total villages/urban wards), school-age population density, and accessibility measures.

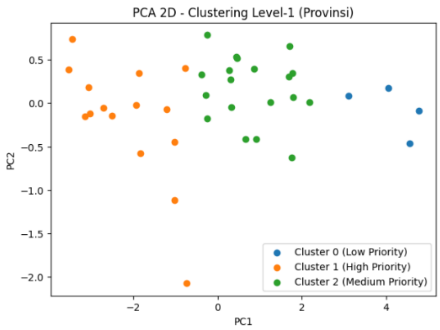

Figure 5. Level-1 clustering results

Figure 5. presents a 2D PCA projection (PC1–PC2) of provinces to support the interpretation of the Level-1 clustering results. The plot shows a reasonably clear separation pattern: Cluster 0 (Low Priority) is concentrated on the positive side of PC1 and appears relatively separated from the other groups, whereas Cluster 1 (High Priority) tends to lie on the negative side of PC1. Cluster 2 (Medium Priority) is located in the central region and partially overlaps with both the high- and low-priority clusters. This structure suggests that the clustering output is meaningful: low-priority provinces exhibit more distinct feature characteristics, while the medium-priority group acts as a transitional segment with similarities to both extremes. It should be emphasized that PCA is used here only for visualization and diagnostic validation of cluster separability; the clusters are still formed using the original features after log1p transformation and standardization, not by PCA itself [21].

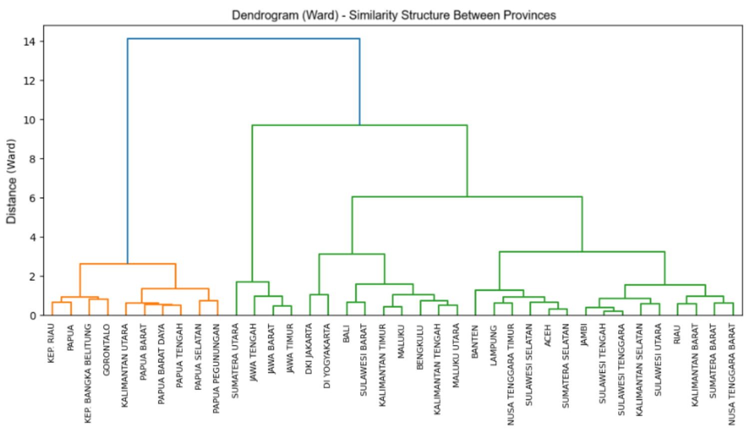

Figure 6. Dendogram (ward) of the similarity structure between provinces

The Ward dendrogram in Figure 6. illustrates the hierarchical similarity structure among provinces. Provinces with more similar school-facility profiles (SD–HE) are merged at lower linkage distances (heights). In contrast, merges occurring at higher distances indicate groups that are substantially different from one another. The presence of several major branches that only join at relatively high linkage distances suggests that the provincial data contain a meaningful, non-random grouping structure, supporting the use of clustering for provincial segmentation.

In addition, the dendrogram provides a useful sanity check for the Level-1 K-Means solution. Cutting the tree at an appropriate height typically yields two to four broad groups, consistent with selecting K=3 based on the Elbow and Silhouette criteria. Ward linkage is chosen because it merges clusters by minimizing the increase in within-cluster variance, aligning well with the variance-minimization principle that underlies centroid-based clustering approaches [22].

4.5. Determining the number of clusters (Level-2)

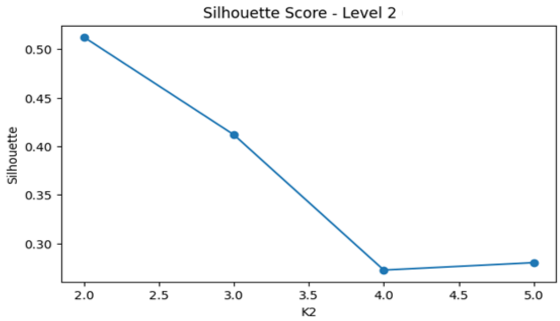

Figure 7. Determination of the number of Level-2 (K2) subclusters

The Level-2 Silhouette plot in Figure 7. shows that the highest Silhouette value is achieved at K2 = 2 (slightly above 0.50), and then decreases sharply at K2 = 3 and remains low for K2 = 4–5. This pattern indicates that provinces within the High-Priority group are most stably separated into two subgroups with distinct education-level composition profiles (share SD–HE). Increasing the number of subclusters beyond two tends to increase overlap between groups and reduce separation quality. Therefore, K2 = 2 is selected as the best configuration for Level-2 micro clustering, enabling clearer interpretation and more targeted level-specific recommendations for high-priority provinces.

4.6. Level-2 clustering results: deepening on high priority

Level-2 clustering was applied exclusively to High Priority provinces to differentiate education-level shortage patterns using proportion (share) features across levels (SD–HE). The selection of $\boldsymbol{K_2}$in Level-2 produced the best Silhouette value at $\boldsymbol{K_2 = 2}$(≈ 0.512), indicating that two distinct facility-composition patterns exist within the high-priority group.

Table 9. Level-2 subclusters at high priority

Province (High Priority)	**Total (T)**	**Subcluster (Level-2)**
Kalimantan Utara	577	Type A (primary school-dominant; high school/vocational school/university low)
Papua Tengah	667	Type A (primary school-dominant; high school/vocational school/university low)
Papua Selatan	721	Type A (primary school-dominant; high school/vocational school/university low)
Papua Barat	743	Type A (primary school-dominant; high school/vocational school/university low)
Papua Barat Daya	752	Type A (primary school-dominant; high school/vocational school/university low)

Province (High Priority)	Total (T)	Subcluster (Level-2)
Kep. Bangka Belitung	758	Type A (primary school-dominant; high school/vocational school/university low)
Papua Pegunungan	770	Type A (primary school-dominant; high school/vocational school/university low)
Kep. Riau	884	Type A (primary school-dominant; high school/vocational school/university low)
Papua	1072	Type A (primary school-dominant; high school/vocational school/university low)
Gorontalo	1201	Type A (primary school-dominant; high school/vocational school/university low)
Sulawesi Barat	1340	Type A (primary school-dominant; high school/vocational school/university low)
Bali	1343	Type A (primary school-dominant; high school/vocational school/university low)
Bengkulu	1989	Type A (primary school-dominant; high school/vocational school/university low)
DKI Jakarta	1074	Type B (primary school-dominant; high school/vocational school/university low)
DI Yogyakarta	1100	Type B (primary school-dominant; high school/vocational school/university low)

Based on Table 9. Type A (13 provinces) exhibits a strongly skewed composition toward the basic levels on average, SD ≈ 58.8% and SMP ≈ 24.1%—while the shares of SMA, SMK, and Higher Education are relatively small (SMA ≈ 9.7%, SMK ≈ 5.2%, HE ≈ 2.1%). This pattern suggests that facility availability is concentrated at the primary and lower-secondary stages, with comparatively limited coverage at higher levels, potentially increasing barriers to progression into upper-secondary and tertiary education in high-priority areas [23].

In contrast, Type B (2 provinces: DKI Jakarta and DI Yogyakarta) shows a more balanced composition and a notably higher share of Higher Education (≈ 9.0%). Therefore, the low total $\boldsymbol{T}$ in these two provinces is more likely influenced by their urban context and smaller village/urban-ward administrative bases rather than a broad shortage of facilities across levels. Consequently, policy recommendations for Type B should be guided by more urban-relevant indicators (e.g., service availability per sub-district/kelurahan, school capacity, and quality), whereas Type A is more directly aligned with interventions targeting the physical equalization of upper-secondary and tertiary facility coverage.

Table 10. Level-2 subcluster profile

Type_L2	n_provinces	Share_SD	Share_SMP	Share_SMA	Share_SMK	Share_HE
Type A	13	0.588324	0.241442	0.096785	0.052336	0.021113
Type B	2	0.321541	0.259624	0.169228	0.159551	0.090056

Table 10. presents the Level-2 subcluster profile within the High Priority group. Type A consists of 13 provinces and is strongly dominated by primary and lower-secondary facility shares, with average shares of SD = 0.588 and SMP = 0.241, while the shares of SMA, SMK, and Higher Education remain relatively low. This indicates that Type A provinces require stronger intervention at upper-secondary, vocational, and tertiary education levels. In contrast, Type B consists of 2 provinces and shows a more balanced facility composition, including a relatively higher share of Higher Education (0.090). Therefore, Type B provinces should not be interpreted solely as areas with broad facility shortages; instead, policy attention should focus more on urban-relevant indicators such as service capacity, quality, and intra-area distribution. These Level-2 profiles show that high-priority provinces are not homogeneous and require differentiated intervention strategies.

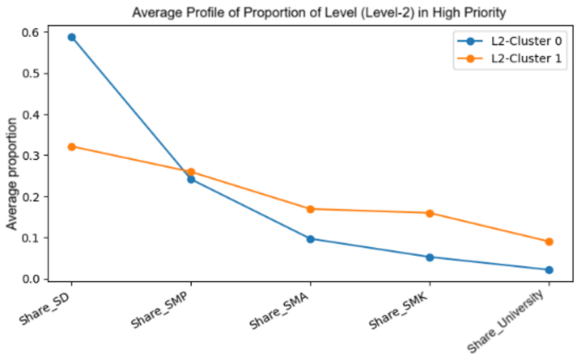

Figure 8. Average profile of level proportion (share) resulting from Level-2 clustering

Figure 8. and the accompanying Level-2 profile table show clear differences between the two subclusters within the High-Priority group. L2-Cluster 0 is strongly dominated by the primary (SD) and lower-secondary (SMP) levels, with average shares of approximately 0.588 and 0.241, respectively. In contrast, the shares of upper-secondary (SMA), vocational (SMK), and especially Higher Education (HE) are relatively small (approximately 0.097, 0.052, and 0.021). This pattern indicates that facility availability is concentrated at the basic education levels and becomes weaker at more advanced levels [24].

By contrast, L2-Cluster 1 shows a more balanced composition, with relatively higher shares at the upper-secondary and vocational levels (SMA ≈ 0.169; SMK ≈ 0.160) as well as a substantially larger share of Higher Education (≈ 0.090). This suggests that the intervention needs of this subgroup differ from those of L2-Cluster 0. Therefore, the Level-2 results do not merely identify "high-priority provinces," but also classify them according to their education-level-specific shortage patterns, allowing equalization policies to be directed more precisely for example, by prioritizing the expansion of SMA, SMK, and HE facilities in subclusters that remain heavily concentrated at the SD–SMP levels.

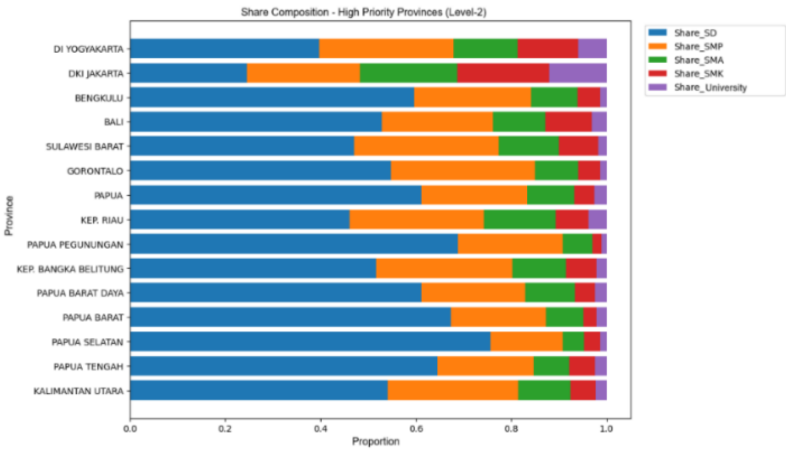

Figure 9. Composition of the proportion of levels (share) in high-priority provinces based on Level-2 clustering results (2024)

The stacked bar chart in Figure 9. illustrates the education-level composition (from Share_SD to Share_Higher_Education) for each province classified as High Priority, thereby helping explain the internal variation that underlies the Level-2 subcluster separation. Overall, many high-priority provinces exhibit dominant shares in primary (SD) and lower-secondary (SMP) facilities, while the shares for upper-secondary (SMA) and vocational (SMK) are smaller, and the higher education share is the smallest. This pattern suggests that facility availability is concentrated at the basic levels and becomes weaker at higher levels [25]. Beyond the numerical clustering results, further interpretation is needed to understand why certain provinces appear as high priority and how the priority map can be translated into policy action.

4.7. Regional interpretation and policy use of the priority map

The clustering results should be interpreted as reflections of regional development characteristics, not merely as numerical groupings. Several eastern provinces, particularly in the Papua region, appear in the High Priority group due to relatively limited village/urban-ward-based facility availability, which may be related to geographical barriers, dispersed settlements, uneven infrastructure development, and administrative reorganization. In contrast, urban provinces such as DKI Jakarta and DI Yogyakarta require more careful interpretation because their lower facility counts are influenced by smaller administrative bases and urban characteristics; therefore, policy attention should focus more on capacity, quality, accessibility, and intra-area equity. From a policy perspective, Level-1 clustering can support the identification of provinces requiring greater planning and budgeting attention, while Level-2 clustering helps determine which education levels should be strengthened. Thus, the proposed priority map can serve as a decision-support tool for differentiated educational facility equalization strategies rather than uniform interventions across all provinces.

4.8. Limitations and future research

This study has several limitations. The indicator used only measures the quantity of villages/urban wards with educational facilities and does not capture facility quality, capacity, teacher availability, accessibility, or transportation constraints. In addition, the facility counts have not been normalized by the total number of villages/urban wards, school-age population, or regional population density, and demand-side variables such as enrollment demand, poverty, and socioeconomic conditions are not yet included. Since clustering is an exploratory unsupervised learning method, the results should be interpreted as a decision-support tool rather than causal evidence. Future studies should integrate GIS-based accessibility analysis, normalized coverage indicators, facility quality measures, socioeconomic variables, spatial clustering, and multi-year data to produce more precise and policy-relevant recommendations for educational facility equalization.

5. Conclusion

This study concludes that the proposed multi-level clustering framework provides an operational approach for mapping educational facility equalization priorities across Indonesian provinces using 2024 village/urban-ward-based facility availability data. The Level-1 results classify provinces into High, Medium, and Low Priority groups, while the Level-2 analysis within the high-priority group identifies education-level-specific needs, particularly differences between provinces dominated by basic education facilities and those with more balanced facility composition. Scientifically, this study contributes a two-stage clustering approach for supply-side educational facility equity analysis, and practically, it offers a priority map and level-specific profiles to support staged infrastructure planning and differentiated policy interventions. However, the findings should be interpreted as exploratory decision-support results because the analysis is based on facility counts and does not yet include coverage ratios, facility quality, accessibility, capacity, or demographic demand. Future research should integrate GIS-based accessibility analysis, normalized coverage indicators, socioeconomic and demand-side variables, and spatial clustering methods to produce more precise recommendations for educational facility equalization.

Credit Authorship Contribution Statement

Antika Zahrotul Kamalia : Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Visualization, Writing – original draft, Project administration. **Zaenur Rozikin :** Methodology, Validation, Formal analysis, Writing – review & editing. **Hemdani Rahendra Herlianto :** Software, Data curation, Investigation, Visualization. **Hendra Arya Syaputra :** Validation, Resources, Writing – review & editing. **Asep Arwan Sulaeman :** Supervision, Methodology, Validation, Writing – review & editing. **Choiriyatun Nisa Latansa :** Supervision, Project administration, Writing – review & editing.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The data that support the findings of this study are publicly available from the Central Statistics Agency (Badan Pusat Statistik/BPS) of Indonesia. The processed datasets and analysis codes used during the current study are available from the corresponding author upon reasonable request.

Use of Artificial Intelligence

The authors used artificial intelligence tools solely to assist with language refinement, grammar checking, and improving the readability of the manuscript. All scientific content, study design, data collection, data analysis, interpretation of the results, and final manuscript preparation were performed, verified, and approved by the authors, who take full responsibility for the content of this article.

References

[1] M. P. Hababil, M. K. Firdaus, N. Nazhmi, M. D. Hamdani, M. R. Alghifary, and A. Fadilla, "Analisis Pengaruh Pemerataan Ekonomi Dalam Upaya Menghapus Ketimpangan Sosial-Ekonomi Antar Masyarakat," *J. Macroecon. Soc.*, vol. 1, no. 4, pp. 1–9, 2024, doi: 10.47134/jmsd.v1i4.276.

[2] I. Widiastuti, "Assessing the Impact of Education Policies in Indonesia: Challenges, Achievement, and Future Direction," *AL-ISHLAH J. Pendidik.*, vol. 17, no. 2, pp. 1955–1964, 2025, doi: 10.35445/alishlah.v17i2.6803.

[3] F. Apriliansyah, M. Ahmad, and H. Rochimah, "Analisis Kesenjangan Pembiayaan Pendidikan Di Indonesia : Studi Literatur Terhadap Perbedaan Akses Dan Kualitas Antara Sekolah Negeri dan Swasta di Daerah Perkotaan dan Pedesaan," *J. Manaj. Pendidik.*, vol. 10, no. 2, pp. 753–758, 2025.

[4] A. Z. Kamalia and I. Nawangsih, "Identifikasi Pola Tingkat Kesenjangan Ketuntasan Pendidikan Di Indonesia Dengan Menggunakan Metode K-Medoids Clustering," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 12, no. 2, pp. 321–330, 2025, doi: 10.25126/jtiik.2025129219.

[5] A. F. Mohamed Nafuri, N. S. Sani, N. F. A. Zainudin, A. H. A. Rahman, and M. Aliff, "Clustering Analysis for Classifying Student Academic Performance in Higher Education," *Appl. Sci.*, vol. 12, no. 19, 2022, doi: 10.3390/app12199467.

[6] M. Ardan, "Perbandingan performa algoritma k-means dan fuzzy c-means dalam klasterisasi data populasi negara," *J. Komput. dan Inform.*, vol. 19, no. 1, pp. 73–82, 2024.

[7] G. F. Anwar, U. Khaira, and P. E. P. Utomo, "Clustering Wilayah Di Indonesia Berdasarkan Kualitas Pendidikan Menggunakan Algoritma Fuzzy C-Means," *Inf. Syst. J.*, vol. 8, no. 2, pp. 167–178, 2025, doi: https://doi.org/10.24076/infosjournal.2025v8i02.2442.

[8] I. P. A. Vidyananta and K. T. Dermawan, "Perbandingan Algoritma K-Means Dan K- Medoids Dalam Pengelompokan Provinsi Di Indonesia Berdasarkan Indikator Keadaan Sekolah Dasar," vol. 13, no. 3, 2025, doi: http://dx.doi.org/10.23960/jitet.v13i3S1.8145.

[9] O. Rahmawati and A. Fauzan, "Provincial Clustering Based on Education Indicators: K-Medoids Application and K-Medoids Outlier Handling," *Barekeng*, vol. 18, no. 2, pp. 1167–1178, 2024, doi: 10.30598/barekengvol18iss2pp1167-1178.

[10] Y. Hasnataeni, M. R. Nurhambali, R. Ardhani, S. Hafsah, and A. M. Soleh, "Comparison of clustering analysis of K-means, K-medoids, and fuzzy C-means methods: case study of school accreditation in west java," *J. Soft Comput. Explor.*, vol. 6, no. 2, pp. 79–88, 2025, doi: 10.52465/joscex.v6i2.575.

[11] E. Nurkhofifah, D. Athina, A. Ristiyanti Tarida, and F. Amelia Pratiwi, "Clustering of Junior High School Education in West Java Based on Density and Dropout Ratios Using Quartile and KMeans Methods," *Proc. Int. Conf. Data Sci. Off. Stat.*, vol. 2025, no. 1, pp. 483–511, 2025, doi: 10.34123/icdsos.v2025i1.662.

[12] Y. Wang, D. Zhai, W. Xie, and S. Huang, "Spatial optimization of hierarchical healthcare facilities driven by multi-source data: a case study of Shenyang, China," *Front. Public Heal.*, vol. 13, no. July, pp. 1–22, 2025, doi: 10.3389/fpubh.2025.1640070.

[13] C. Xu and Y. Sun, "Building-Scale Accessibility Assessment of Sports Facilities: A Spatial Equity Perspective," *Land*, vol. 15, no. 3, pp. 1–25, 2026, doi: 10.3390/land15030522.
[14] A. Rykov, R. C. De Amorim, V. Makarenkov, and B. Mirkin, "Inertia-Based Indices to Determine the Number of Clusters in K-Means: An Experimental Evaluation," *IEEE Access*, vol. 12, no. December 2023, pp. 11761–11773, 2024, doi: 10.1109/ACCESS.2024.3350791.
[15] S. Yildirim, "XGBoost-Driven Evaluation of Clustering Methods for MOBA Player Segmentation," *J. Emerg. Comput. Technol.*, vol. 5, no. 1, pp. 84–95, 2025, doi: 10.57020/ject.1770052.
[16] Badan Pusat Statistik, "Jumlah Desa yang Memiliki Fasilitas Sekolah Menurut Provinsi dan Tingkat Pendidikan," BPS. [Online]. Available: https://www.bps.go.id/id/statistics-table/2/MjA1IzI=/jumlah-desa-yang-memiliki-fasilitas-sekolah-menurut-provinsi-dan-tingkat-pendidikan.html
[17] A. H. Eyeleko, T. Feng, and Y. Yan, "A Unified Clustering-Based Anonymization for Privacy-Preserving Data Publishing with Multidimensional Privacy Quantification," *Inf.*, vol. 17, no. 3, pp. 1–40, 2026, doi: 10.3390/info17030302.
[18] M. Karagiannidou, C. Vasilakos, E. Kokinou, and N. Gerarchakis, "High-Resolution Eutrophication Mapping Using Multispectral UAV Imagery and Unsupervised Classification: Assessment in the Almyros Stream (Crete, Greece)," *Remote Sens.*, vol. 18, no. 3, pp. 1–31, 2026, doi: 10.3390/rs18030501.
[19] N. Alzahrani, M. Meccawy, H. Samra, and H. A. El-Sabagh, "Identifying Weekly Student Engagement Patterns in E-Learning via K-Means Clustering and Label-Based Validation," *Electron.*, vol. 14, no. 15, pp. 1–27, 2025, doi: 10.3390/electronics14153018.
[20] X. Yang and M. Wang, "Diversification and Spatial Differentiation of Villages' Functional Types in the New Period of China: Results from Hierarchical Urban-Rural Spatial Relations and Townships Size," *Land*, vol. 11, no. 2, 2022, doi: 10.3390/land11020171.
[21] M. Briscik, M. A. Dillies, and S. Déjean, "Improvement of variables interpretability in kernel PCA," *BMC Bioinformatics*, vol. 24, no. 1, pp. 1–21, 2023, doi: 10.1186/s12859-023-05404-y.
[22] G. U. Nugraha, H. Bakti, R. F. Lubis, A. Mulyono, Y. Ulfa, and Y. Sudrajat, "Data driven resistivity zonation integrating inversion kriging and clustering for subsurface characterization in groundwater exploration," *Discov. Water*, vol. 5, no. 1, 2025, doi: 10.1007/s43832-025-00261-7.
[23] A. Rizqi, K. Ainur, A. Qonitatun, and S. N. Luwistiana, "Kesenjangan Kualitas Sekolah : Tantangan Pemerataan Pendidikan Dasar di Pesisiran Jawa dan Implikasinya di ranah Keberagaman islam," *Al Madjid J. Pendidik. Islam*, vol. 2, no. 1, pp. 1–29, 2026.
[24] L. Jiang, J. Chen, Y. Tian, and J. Luo, "Spatial Pattern and Influencing Factors of Basic Education Resources in Rural Areas around Metropolises—A Case Study of Wuhan City's New Urban Districts," *ISPRS Int. J. Geo-Information*, vol. 11, no. 11, 2022, doi: 10.3390/ijgi11110576.
[25] M. N. Huda, D. A. Pratiwi, A. Chairil, Solihan, S. R. R. Pramono, and M. Y. Hidayatullah, "Analisis Spasial Persebaran Fasilitas Kesehatan," *JPIG (Jurnal Pendidik. dan Ilmu Geogr.*, vol. 11, no. 1, pp. 93–105, 2026, doi: https://doi.org/10.21067/jpig.v11i1.13157 93.