Received: 06/12/2025 | Revised: 21/03/2026 | Accepted: 12/04/2026
DOI: 10.69816/jgd.v3i1.48775

ORIGINAL ARTICLE

TAMALANREA
JOURNAL OF GOVERNMENT AND DEVELOPMENT

Regional development in West Java Province: Clustering population density, human development index, and life expectancy

Weky Agesty[1], **Nayla Bunga Safa Felisa**[2], **Aulia Sri Barokah**[3] and **Hilma Mutiara Winata**[4]

Affiliation
[1,2,3,4] Department of Public Administration, Faculty of Social and Political Sciences,
Sunan Gunung Djati State Islamic University,
Bandung, Indonesia, 40614

Correspondence
wekyagesty2288@gmail.com

Funding Information
The authors declare that no specific grant was received from funding agencies in the public, commercial, or not-for-profit sectors for the completion of this study.

Abstract
This study investigates regional development patterns in West Java Province by examining the relationships among population density, the Human Development Index (HDI), and life expectancy, while also identifying regional characteristics through K-means clustering analysis. The study aims to provide a comprehensive overview of these indicators, evaluate their linear relationships, and classify districts and municipalities according to their development profiles. A quantitative research approach was employed, utilizing descriptive statistics, Pearson's correlation analysis, and K-means clustering. Secondary data were obtained from official government publications. The findings reveal a strong positive correlation between population density and HDI, as well as between population density and life expectancy. The clustering analysis identified two distinct groups: highly urbanized and densely populated areas characterized by higher HDI scores and longer life expectancy, and less densely populated districts exhibiting comparatively lower levels of human development. These results underscore persistent disparities in regional development between urban and non-urban areas in West Java, which are associated with unequal access to education, healthcare services, and infrastructure. The study provides empirical evidence that can support regional governments in designing more targeted and equitable development policies tailored to the specific characteristics and needs of each cluster.

Keywords
Regional Development, Population Density, Human Development Index, Life Expectancy, Clustering Analysis, Quality of Life.

1 | INTRODUCTION

Regional development is a crucial aspect for improving public welfare and achieving an equitable quality of life. However, West Java Province, as the most populous region in Indonesia (with more than 49 million residents), faces significant challenges related to development disparities across its districts and cities (BPS West Java Province, 2023). While major urban areas have experienced rapid infrastructure growth, many regions, particularly inland and rural areas, still face limited access to basic services (Surtiadi & Sunsun, 2017). To map these dynamics, macro indicators such as population density, the Human Development Index (HDI), and life expectancy serve as key variables for a comprehensive evaluation of development quality. Research on regional development dynamics has been widely conducted, particularly on clustering regions based on economic and social indicators (Rocha et al., 2020).

Previous studies in West Java have focused on regional clustering using single development indicators, particularly the Human Development Index (HDI), or village-level development measures (Irsyifa Mayzela Afnan et al., 2024). These studies have successfully identified regional disparities and provided useful classifications for policymaking. However, most existing studies emphasize classification outcomes without examining the statistical relationships among demographic and human development indicators. In addition, previous research generally relies on a single development dimension, limiting a comprehensive understanding of regional development dynamics (Fitrianigsih & Dwi Kartikasari, 2024). Consequently, there remains a lack of empirical evidence regarding how population density relates to human development achievements and health outcomes across districts and cities in West Java. Addressing this gap is important because effective regional development policies require not only regional categorization but also an understanding of the interactions among key development indicators (Fratesi & Perucca, 2019).

Although previous studies have successfully identified regional disparities and classified districts and cities based on development indicators, several limitations remain (Sawicki & Flynn, 1996). Most studies focus primarily on a single development dimension, particularly the Human Development Index (HDI), while limited attention has been given to examining the interrelationships between demographic characteristics and human development outcomes. Furthermore, existing studies generally emphasize regional classification results without first investigating the statistical relationships among key development indicators. Consequently, there is still limited empirical evidence regarding how population density relates to HDI and life expectancy across districts and cities in West Java (Hendajany & Riyadi, 2022).

Addressing this gap is important because understanding both the relationships among development indicators and regional typologies can provide a stronger evidence base for designing more targeted and effective development policies (Avdiushchenko & Zając, 2019). Therefore, this study contributes to the literature by integrating Pearson correlation analysis and K-Means clustering to simultaneously examine development relationships and regional development patterns in West Java. In addition, no studies have explicitly examined the linear relationships between population density and HDI and life expectancy before clustering, even though population density plays an important role in determining the efficiency of public service delivery and the pressure on infrastructure (Maulu et al., 2021)

Prior to clustering, all variables were standardized using Z-score normalization to eliminate scale differences among variables (Moeller, 2025). The optimal number of clusters was determined using the Elbow Method by evaluating the Within-Cluster Sum of Squares (WCSS) across several cluster solutions. The results indicated that k = 2 provided the most interpretable grouping structure and represented the point where additional clusters produced only marginal reductions in WCSS. Therefore, two clusters were selected for the final K-Means analysis.

The urgency of this research lies in the need for an integrated understanding of the relational patterns among development indicators (population density, HDI, and life expectancy) and the simultaneous mapping of regional characteristics (Man et al., 2021). Formulating effective development policies requires a data foundation that not only reveals the status of development (clustering) but also identifies the driving and inhibiting factors (correlation).

The novelty of this study lies in the integration of Pearson correlation analysis and K-Means clustering within a single analytical framework. Unlike previous studies that primarily focused on regional classification using a single development indicator, this research simultaneously examines the relationships among population density, HDI, and life expectancy before identifying regional development typologies (Long et al., 2020). By combining relationship analysis and regional classification, this study provides a more comprehensive understanding of both the determinants and patterns of regional development disparities in West Java. The findings offer practical insights for local governments in designing more targeted and evidence-based development policies (Arnold et al., 2024)

2 | LITERATURE REVIEW

2.1 | Regional Development and Growth Pole Theory

Regional development is inherently a dynamic and interactive process aimed at enhancing societal welfare sustainably through the strategic integration of economic, social, and demographic dimensions (Adamowicz, 2023). Theoretically, development should progress equitably across all geographical spaces. In empirical reality, however, economic growth and capital accumulation exhibit highly uneven distributions, thereby creating sharp spatial polarizations between advanced urban centers and relatively lagging rural peripheries. This structural disparity is frequently triggered by variations in natural resource endowments, disparate levels of accessibility, and the skewed concentration of infrastructural investments.

Developmental polarization can be profoundly analyzed through the lens of Growth Pole Theory, originally conceptualized by François Perroux (Torre, 2025). The theory posits that development does not appear everywhere simultaneously; rather, it concentrates at specific points or "growth poles" possessing dominant economic agglomeration forces. These growth poles are characterized by the presence of large-scale industries, technological innovations, and expansive labor markets capable of drawing production factors from surrounding regions. In the context of West Java, major urban centers like Bandung City and Bekasi City act as growth poles that heavily absorb migration inflows and capital.

Furthermore, the impact of a growth pole on its surrounding hinterlands is elucidated by the Core-Periphery Theory (Liang, 2024). This framework partitions geographical space into an advanced "core" that dominates economically, and a dependent "periphery." Ideally, the core generates positive spread effects (or spillover effects) through technology transfer, demand for raw materials, and employment opportunities in the periphery. However, if the centrifugal pull of the center is overly dominant, backwash effects occur instead—whereby capital and skilled labor are drained from the periphery to the core, subsequently widening the developmental gap.

In West Java Province, this core-periphery dichotomy is visibly manifested in the spatial division between the Northern-Western and Southern-Eastern regions. The northern and western corridors, directly contiguous with the *Jabodetabek* megapolitan area, have rapidly developed as industrial cores and primary economic hubs. Conversely, the southern and eastern regencies, characterized by rugged topography, limited transportation infrastructure, and dispersed settlements, tend to remain isolated and occupy a peripheral status. Consequently, spatial disparities in developmental achievements at the district and city levels become an inevitable geographical outcome (Priatama et al., 2022).

2.2 | The Interrelationship Between Population Density, HDI, and Life Expectancy

Population density serves as a crucial demographic indicator depicting human concentration within a specific spatial unit. Within regional planning discourse, population density is not merely perceived as a social burden but also as an asset that catalyzes efficiency in public service delivery. When population is concentrated in specific areas (urbanization), governments can minimize per capita costs required to construct essential public infrastructure. A dense population base establishes a robust market that incentivizes both public and private sectors to provide high-quality educational and healthcare facilities (Fabre & Straub, 2023).

The nexus between population density and the quality of human capital is directly reflected in the achievements of the Human Development Index (HDI). The HDI is a composite indicator encompassing the dimensions of a long and healthy life, knowledge, and a decent standard of living. In densely populated urban areas, access to reputable schools, higher education institutions, vocational training centers, and modern economic opportunities is far more abundant than in sparsely populated regions. This resource allocation efficiency in high-density areas triggers improvements in school enrollment rates and per capita income, which consequently elevates the overall HDI score significantly (Sari & Tiwari, 2024).

Beyond the educational and economic dimensions, a well-managed population density also exhibits a positive linear correlation with Life Expectancy (LE). Life expectancy reflects the public health status, which is heavily influenced by environmental quality, nutritional adequacy, and the accessibility of medical services. Regions with high population concentrations typically possess vastly superior healthcare infrastructure, such as Type-A referral hospitals, medical specialists, and structured sanitation programs. Rapid and close access to emergency medical facilities in dense urban settings acts as a primary determinant for higher life expectancy among local populations.

Meticulous urban planning and sufficient environmental carrying capacity are mandatory to sustain this positive relationship, as it is not entirely absolute. If population growth in dense areas is unmatched by infrastructure expansion, these regions risk facing threats of congestion, pollution, slums, and severe strain on basic service networks a phenomenon known as urban diseconomies of agglomeration. Therefore, a robust positive correlation among population density, HDI, and life expectancy indicates that a region is still operating in a phase where agglomeration economies significantly outweigh the corresponding congestion costs (Sun & Abdullah, 2025).

In the macro context of West Java, the interaction of these three variables produces a unique pattern of disparity. Highly dense cities demonstrate absolute dominance in HDI components and life expectancy due to mature urban ecosystems. Meanwhile, expansive regencies with low population densities face formidable challenges characterized by a high-cost economy and high-cost delivery of public services. Dispersed settlements in rural and remote areas render the distribution of educational facilities and health clinics inefficient, thereby hindering the acceleration of overall quality of life.

2.3 | Application of K-Means Clustering in Regional Disparity Analysis

Objective mapping and classification of regional characteristics represent the foundational steps required for policymakers to address regional disparities effectively. A highly prominent and effective multivariate statistical tool utilized for this purpose is the non-hierarchical K-Means Clustering technique. This method operates by partitioning a set of observations into a predetermined number of clusters (k), where each data point is assigned to the cluster with the nearest center (*centroid*) (Karthikeyan, 2020).

The operational mechanism of the K-Means algorithm centers on minimizing within-cluster variance and maximizing between-cluster variance. Through an iterative approach, the algorithm computes the Euclidean distance from each data object to the centroids, updates the centroid positions based on the mean of the newly assigned members, and repeats this process until convergence is reached. The primary advantage of this method lies in its computational efficiency when handling multivariate datasets, enabling it to detect latent similarity patterns across regions rapidly and accurately (Guo et al., 2022).

Prior to executing the K-Means algorithm, a critical data preprocessing stage must be undertaken, namely data normalization or standardization. Given that regional development indicators such as population density, HDI, and life expectancy possess highly divergent measurement units and scale ranges, direct cluster analysis without transformation would yield biased groupings. The application of Z-score standardization serves to transform all variables into a uniform scale with a mean of zero and a standard deviation of one. This ensures that every variable contributes an equitable and balanced weight during the inter-object distance calculation.

The primary challenge in applying K-Means lies in determining the most appropriate and representative number of clusters (*k*). Empirical testing is conducted using the Elbow Method to mitigate researcher subjectivity. This method evaluates the Within-Cluster Sum of Squares (WCSS) representing the total squared internal distances—across various values of *k*. As the number of clusters increases, the WCSS value naturally declines. The optimal *k* point is identified at the segment of the graph that forms a sharp angle or inflection resembling an "elbow," where adding more clusters beyond this point no longer yields a significant reduction in WCSS (Oti et al., 2021).

In regional development studies, the implementation of K-Means clustering functions not merely as a technical mathematical apparatus but also carries profound policy interpretation value. The clustering results simplify the complexity of spatial data into easily interpretable regional typologies, thereby allowing governments to identify priority areas for intervention. Integrating correlation analysis to scrutinize inter-variable relationships with K-Means clustering to examine actual spatial groupings establishes a rigorous, evidence-based policy framework for formulating targeted regional development strategies.

3 | METHODS

3.1 | Research Design

This study employs a quantitative research approach to examine the relationship between population density, human development, and health outcomes across districts and cities in West Java Province. Quantitative methods are particularly suitable for identifying statistical relationships among variables and generating objective evidence that can support regional development planning and policy formulation (Kusumastuti et al., 2020).

To achieve the research objectives, this study integrates two analytical techniques: Pearson correlation analysis and K-Means clustering. Pearson correlation is used to assess the strength and direction of the linear relationships between Population Density (PD), Human Development Index (HDI), and Life Expectancy (LE). This method enables the identification of whether regions with higher population concentrations tend to exhibit better human development outcomes and longer life expectancy.

In addition, K-Means clustering is employed to classify districts and cities into several groups based on similarities in their demographic and human development characteristics. As an unsupervised machine learning technique, K-Means facilitates the identification of natural groupings within the data without requiring predefined categories. The resulting clusters provide a clearer understanding of regional disparities and enable the identification of areas that share similar development profiles (Guo et al., 2022).

3.2 | Data Sources and Variables

This study utilises secondary data obtained from the West Java Provincial Government Work Plan (RKPD Jawa Barat) 2025 and official statistical publications issued by the Indonesian Central Bureau of Statistics (Badan Pusat Statistik/BPS). The dataset comprises observations from all districts and cities in West Java Province, making the study population identical to the sample through the application of a total sampling technique (BPS Provinsi Jawa Barat, 2025).

Three variables were selected to represent demographic conditions and human development outcomes. Population Density (PD) was used to measure the concentration of population within a specific area and was expressed as the number of inhabitants per square kilometre. The Human Development Index (HDI) was employed as a composite indicator reflecting achievements in education, health, and living standards. Life Expectancy (LE) was used to represent the average expected lifespan of the population and served as an indicator of regional health conditions.

The selection of these variables was based on theoretical and empirical considerations suggesting that population concentration may influence both human development achievements and public health outcomes. By integrating demographic and development indicators, this study seeks to provide a comprehensive assessment of regional disparities across districts and cities in West Java (Singh, 2021).

3.3 | Pearson Correlation Analysis

Pearson correlation analysis was employed to examine the strength and direction of the linear relationships among the selected variables. This statistical technique is widely used to measure the degree of association between continuous variables and is represented by the correlation coefficient (r), which ranges from −1 to +1. Positive values indicate a direct relationship, whereas negative values indicate an inverse relationship. Values closer to zero suggest weaker associations between variables (Kusumastuti et al., 2020).

In this study, the correlation analysis focused on two primary relationships: Population Density and Human Development Index, as well as Population Density and Life Expectancy. These relationships were examined to determine whether districts and cities with higher population concentrations tend to exhibit better human development outcomes and higher levels of public health.

The statistical significance of the correlation coefficients was evaluated to ensure that the observed relationships were not attributable to random variation. The results of the correlation analysis provide preliminary insights into the extent to which demographic concentration is associated with development and health indicators across West Java (Saccenti et al., 2020).

3.4 | K-Means Clustering Analysis

Following the correlation analysis, K-Means clustering was applied to classify districts and cities into groups based on similarities in Population Density, Human Development Index, and Life Expectancy. K-Means is a partition-based clustering algorithm that aims to group observations into clusters while minimising variation within clusters and maximising differences between clusters. Prior to clustering, all variables were standardised using z-score transformation to eliminate differences in measurement scales and ensure that each variable contributed equally to the clustering process. Standardisation is essential because Population Density, HDI, and Life Expectancy are measured using different units and numerical ranges (Guo et al., 2022).

The clustering procedure involved determining the optimal number of clusters, assigning observations to the nearest centroid based on Euclidean distance, and iteratively recalculating centroid positions until stable cluster memberships were obtained. The resulting clusters were then interpreted according to their centroid values and overall characteristics to identify patterns of regional development and demographic concentration. The combination of K-Means clustering and Pearson correlation enables a comprehensive analysis by simultaneously examining statistical relationships among variables and identifying groups of districts and cities with similar development profiles. All statistical analyses and clustering procedures were conducted using Google Colab (Python) and IBM SPSS Statistics (Kenger et al., 2023).

4 | RESULTS AND DISCUSSION

4.1 | Descriptive Statistics

The data used in this study consist of three main variables: population density, the Human Development Index (HDI), and life expectancy across 27 districts and cities in West Java Province. Descriptive statistics provide an initial overview of each variable's distribution before proceeding with the correlation and clustering analyses.

Table 1. Descriptive Statistics Results.

Variable	N	Minimum	Maximum	Mean	Std. Deviation
Population Density (People/km^2)	27	392	15421	3856.96	4575.144
Human Development Index (HDI)	27	68.18	83.29	74.1119	4.34493
Life Expectancy (years)	27	70.19	75.79	73.0022	1.40876
Valid N (listwise)	27				

Table 1 presents the descriptive statistics, the population density variable ranges from 392 to 15,421 people/km^2, with a mean (average) of 3,856.96 people/km^2. This wide range indicates a substantial imbalance in population distribution across West Java, with specific areas, particularly urban regions, exhibiting extremely high population concentrations, while others show significantly lower densities. The standard deviation of 4,575.144 further underscores the considerable variation in population distribution across regions.

The Human Development Index (HDI) ranges from 68.18 to 83.29, with an average of 74.11. The standard deviation of 4.34493 suggests a noticeable disparity in human development achievements across regions, though the variation is smaller than that observed in population density. Overall, the average HDI indicates that most regions in West Java fall into the high human development category according to the BPS classification (Badan Pusat Statistik, 2024).

For the life expectancy variable, the minimum recorded value is 70.19 years, the maximum is 75.79 years, and the average is 73.00 years. The standard deviation of 1.40876 indicates that differences in health levels across regions are relatively moderate compared to those of the other variables. This implies that access to healthcare services is relatively evenly distributed, although minor disparities remain among districts and cities (Xu et al., 2022). The descriptive statistics reveal substantial variation in population density across districts and cities in West Java, while HDI and life expectancy show relatively smaller variations. These differences indicate the presence of regional development disparities and provide a basis for further correlation and clustering analyses (Majumder et al., 2022).

4.2 | Pearson Correlation Test Results

The Pearson correlation analysis between population density and the Human Development Index (HDI) shows a strong, statistically significant relationship with a correlation coefficient of 0.904. This positive correlation indicates that higher population density is associated with higher HDI values. The significance level of $p < 0.001$ confirms that the relationship is statistically significant at the 99% confidence level. With 27 observations, these findings suggest that highly populated, typically urban areas tend to have better access to education, healthcare, and economic opportunities, thereby contributing to higher HDI scores in West Java (Arya Pramaditya et al., 2024).

Table 2. Pearson correlation results for population density, HDI, and life expectancy.

Description	Pearson Correlation	Sig. (2-tailed)	N
Relationship Between Population Density and HDI	0.904	<0.001	27
Relationship Between Population Density and Life Expectancy	0.714	<0.001	27

Meanwhile, the Pearson correlation analysis between population density and life expectancy shows a strong positive relationship (r = 0.714). This means that regions with higher population density tend to have higher life expectancy. The p-value of <0.001 also indicates statistical significance at the 99% confidence level (Table 2). Based on 27 observations, this finding suggests that densely populated and urbanized areas, which generally have better access to healthcare services, hospitals, and public facilities, tend to experience higher life expectancy than less populated regions (Galvani-Townsend et al., 2022).

The strong positive correlations between population density and HDI ($r = 0.904$, $p < 0.001$) and between population density and life expectancy ($r = 0.714$, $p < 0.001$) across 27 regions in West Java align with real-world conditions. Urban areas such as Bandung City, with population densities exceeding 15,000 people/km^2, benefit from greater access to healthcare, education, and economic infrastructure, which supports improvements in both HDI and life expectancy. Data from BPS reports that West Java's HDI reached 74.24 in 2023, with an average life expectancy of 73.8 years (BPS Jawa Barat, 2024). In highly populated areas such as Bekasi City, life expectancy is among the highest (75.79 years). Other studies also confirm the link between urbanization and improvements in HDI and life expectancy, driven by more efficient allocation of public resources in densely populated regions (Mohamad Amin et al., 2024; Nguea, 2023; Shao & Kong, 2024). The correlation value of 0.904 is consistent with existing literature and the province's urbanization trend, which has increased by 1.26% per year.

4.3 | K-Means Clustering Results

The clustering analysis using the three variables, population density, HDI, and life expectancy, shows that the regions in West Java form two main groups with notably different characteristics (Table 3). The first cluster (Cluster 0) is dominated by regions with relatively low population density, averaging approximately 1,352 people per square kilometer. Areas in this cluster also exhibit lower human development outcomes, as reflected in the average HDI value of 71.97. In addition, their life expectancy falls within the moderate range, with an average of 72.48 years.

Table 3. Cluster Statistics Results

Cluster	Population density			HDI			Life Expectancy (years)			Cluster		
	mean	median	std	mean	median	std	mean	median	std	mean	median	std
0	1352.0	1026.0	883.02	71.97	71.62	2.23	72.48	72.51	1.12	0.0	0.0	0.0
1	11014.0	10123.0	2749.45	80.22	79.69	2.70	74.50	74.50	1.04	1.0	1.0	0.0

On the other hand, the second cluster (Cluster 1) represents regions with very high population density, averaging 11,014 people per square kilometer. Interestingly, regions in this cluster demonstrate better human development performance. This is indicated by their higher HDI of 80.22 and a longer life expectancy of 74.50 years.

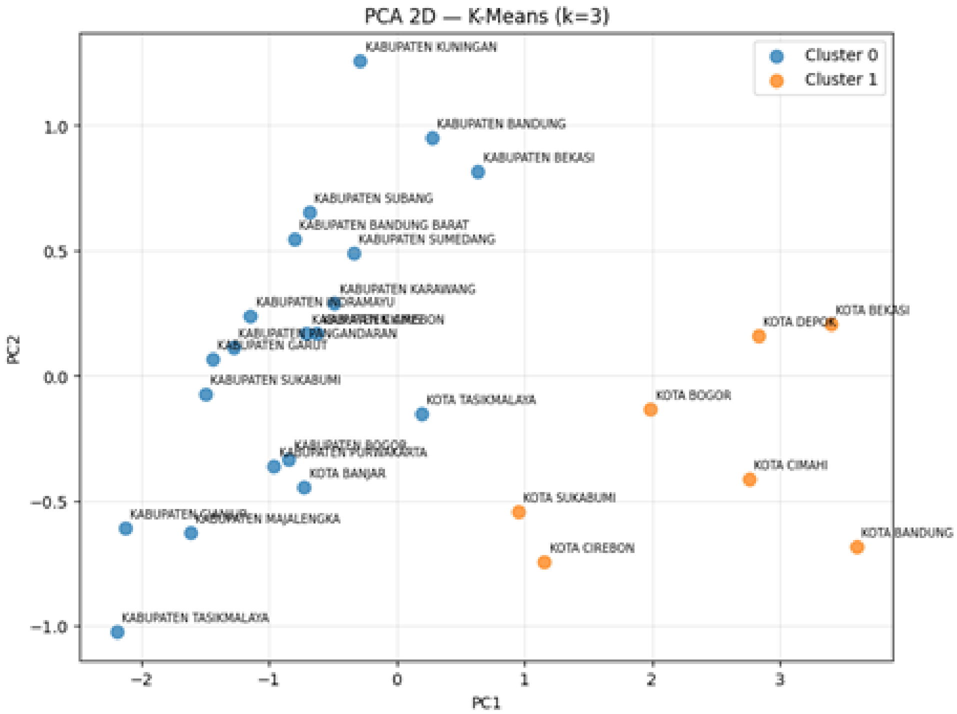

Fig. 1. PCA plot results.

The differences in characteristics between the two clusters become more apparent through the 2D PCA visualization in Fig. 1 In the plot, the regions in Cluster 0 appear tightly grouped on one side of the diagram, indicating substantial similarity among the areas within this cluster. In contrast, the regions in Cluster 1 are more widely dispersed on the opposite side, reflecting significant differences in population density, HDI, and life expectancy.

The PCA visualization further confirms the effectiveness of the clustering process by showing a clear separation between the two clusters within the reduced-dimensional space. This separation indicates that the selected variables population density, HDI, and life expectancy are capable of distinguishing regions with different development characteristics. The limited overlap between clusters suggests that the K-Means algorithm successfully identified distinct regional development patterns in West Java. Furthermore, the distribution of observations in the PCA plot demonstrates that variations in population density are closely associated with differences in human development outcomes, supporting the results of the correlation analysis.

This pattern directly supports the correlation findings, which show strong relationships between population density and both HDI and life expectancy. In other words, regions with higher population density tend to exhibit better human development outcomes (Bille et al., 2023). The clustering results naturally group regions by these characteristics, reinforcing the interpretation that population density is a key variable in explaining variation in human development across West Java (Pravitasari et al., 2021).

These findings are also consistent with field conditions. Highly populated cities typically have better access to educational services, healthcare, infrastructure, and economic opportunities (Fitrah Maharani et al., 2025). This contributes to improved quality of life, which is reflected in higher HDI and life expectancy values. Meanwhile, larger, less densely populated districts face challenges in ensuring equitable public service delivery, resulting in slower human development progress (Dick-Sagoe, 2020). This further strengthens the conclusion that population density is one of the main determinants of human development disparities across regions in West Java.

Thus, regional grouping through clustering methods serves not only as a technical analytical result but also reinforces and validates the correlation findings that higher population density is associated with better human development outcomes (Pravitasari et al., 2021). This demonstrates that population density plays a crucial role in illustrating both the inequalities and the development potential of regions in West Java (Wedi & Fathurrahman, 2025).

4.4 | Pearson Correlation and K-Means Clustering for Regional Development

The results of the correlation test and K-Means analysis provide a comprehensive overview of how population distribution dynamics influence the quality of human development in West Java. The strong correlations between population density and both the Human Development Index (HDI) and life expectancy (LE) indicate that areas with higher population density tend to have greater access to education, healthcare, and socio-economic infrastructure (Chen et al., 2023). When these correlations are linked with the K-Means output, patterns of development inequality become increasingly evident: major urban cities such as Bandung, Bekasi, Depok, and Cimahi are grouped within clusters characterized by higher human development achievements, whereas regencies such as Cianjur, Garut, Kuningan, and Sukabumi fall into clusters with relatively lower performance.

This pattern reflects on-the-ground conditions, showing that development in West Java remains spatially centralized(Saksono, 18 C.E.). Rapid growth occurs in urban and metropolitan areas, while larger, less densely populated regions experience slower development. (Dick-Sagoe, 2020) highlights that urban regions benefit from better public service infrastructure due to concentrated economic activity and service provision. Meanwhile, Naisy (2025) emphasizes that many regencies face challenges in delivering education and health services due to dispersed settlements, complex geographical conditions, and limited fiscal capacity (Madubun, 2024). These empirical findings reinforce the clustering results, suggesting that spatial isolation and low population density contribute to the lower levels of human development in regency areas.

The correlation and clustering results are also relevant to West Java's strategic development, especially the disparities between the North–South and West–East regions. The northern and western parts, closely connected to industrial zones and the Greater Jakarta area, show significantly higher human development outcomes. In contrast, the southern region, characterized by hilly topography and distant settlements, experiences slower development. This structural disparity is reflected in the clustering results, where southern and eastern regions tend to be grouped into clusters with lower HDI–LE values.

These findings can also be interpreted through the perspective of regional development theories, particularly Growth Pole Theory and Core–Periphery Theory (Madubun, 2024). Growth Pole Theory suggests that development tends to concentrate in specific growth centers where economic activities, infrastructure, and investment are highly concentrated. In the context of West Java, major urban areas such as Bandung, Bekasi, and Depok function as growth poles that attract resources and generate development spillover effects. Meanwhile, Core–Periphery Theory explains that regions located closer to economic centers tend to experience more rapid development than peripheral regions. This perspective helps explain why districts located in southern and eastern West Java, which generally have lower accessibility and weaker economic integration, are grouped into clusters with lower HDI and life expectancy values. Therefore, the observed disparities are not merely demographic differences but also reflect uneven spatial distribution of development opportunities and public resources.

These cluster categories provide an important insight for policymakers: regional disparities are not merely statistical differences but reflect challenges in accessibility, local economic structure, public service quality, spatial connectivity, and fiscal capacity. Therefore, the K-Means grouping can serve as a basis for formulating more precise development interventions, such as improving healthcare facilities, equalizing education services, accelerating transportation and road infrastructure, and strengthening local economic sectors based on regional potential (Gharehbaghi et al., 2020).

Overall, the integration of correlation and clustering results confirms that population density is not merely a demographic figure but a key indicator of the pace of regional development. Leveraging these findings enables the West Java Provincial Government to establish more targeted development priorities, reduce city–regency disparities, and promote more inclusive and equitable development across the province.

5 | CONCLUSION

The results of the correlation analysis and K-Means clustering demonstrate that regional development in West Java Province continues to face persistent spatial disparities, primarily characterized by a structural divide between advanced urban centers and lagging regencies. Highly urbanized regions exhibit a strong, mutually reinforcing nexus of high population density, superior Human Development Index (HDI) scores, and longer Life Expectancy (LE), whereas less-developed areas form a distinct peripheral cluster marked by lower socio-demographic outcomes. This concentration of development in specific urban agglomerations confirms the prevailing core-periphery dynamics within the province, where the northern and western economic corridors heavily outperform the southern and eastern rural zones. Consequently, utilizing correlation results to understand the interdependencies among these key macro indicators, combined with clustering typologies to map regional profiles, establishes a rigorous empirical foundation for designing more precise and targeted regional equalization policies.

Addressing these deep-rooted interregional disparities requires local governments to transition from uniform approaches toward highly differentiated, evidence-based development strategies that prioritize the underserved peripheral cluster. A fundamental policy instrument involves strengthening fiscal equalization mechanisms to bridge the fiscal capacity gap among local governments, thereby ensuring that less-developed regencies can provide standardized public services comparable to advanced urban areas. In the education sector, strategic interventions must focus on rectifying the uneven distribution of human capital by implementing robust teacher redistribution policies, enhancing welfare incentives for educators in remote areas, and expanding rural educational infrastructure. Concurrently, accelerating investments in primary healthcare facilities and critical transportation networks is imperative to overcome geographic isolation, improve access to vital services, and foster long-term inclusive economic integration across all regions in West Java.

Future research could substantially enrich these empirical findings by incorporating a broader spectrum of socio-economic variables to capture the multidimensional nature of regional disparities. Integrating macro indicators such as absolute poverty rates, gross regional domestic product (GRDP) growth, direct healthcare accessibility indexes, and digital infrastructure quality would provide a more comprehensive diagnostic of regional performance. Furthermore, subsequent studies should explore the application of alternative non-hierarchical clustering techniques, hierarchical algorithms, or advanced spatial econometrics—such as exploratory spatial data analysis (ESDA) and spatial autoregressive models to uncover deeper spatial patterns, localized spillover effects, and temporal dynamics in regional development.

Acknowledgments

The authors would like to express their sincere gratitude to all individuals and parties who contributed to the completion of this research, either directly or indirectly, through valuable support, insights, suggestions, and constructive feedback.

Disclosure Statement

The authors declare that there are no conflicts of interest regarding the publication of this article. The research was conducted independently, and the authors received no financial or commercial support that could have influenced the study's design, analysis, interpretation, or publication.

Data Availability Statement

The data used in this study were obtained from publicly available secondary sources published by the Central Statistics Agency (BPS) and the West Java Provincial Statistics Agency. The datasets analyzed during the current study are available from the corresponding author upon reasonable request.

Reference

Adamowicz, M. (2023). Local development model as an element of regional sustainable strategy. In *Sustainable Regional Planning*. IntechOpen. https://doi.org/10.5772/intechopen.109534

Arnold, L., Bimczok, S., Clemens, T., Brand, H., & Starke, D. (2024). Implementing evidence ecosystems in the public health service: Development of a framework for designing tailored training programs. *PLOS ONE*, *19*(4), e0292192. https://doi.org/10.1371/journal.pone.0292192

Arya Pramaditya, Y., Bintang Karisma, D., & Muhammad Naufal, V. (2024). Assessing the role of inequality and human development indices in regional economic growth: West Java. *Creative Research Journal*, *10*(02), 133–146. https://doi.org/10.34147/crj.v10i02.389

Avdiushchenko, A., & Zając, P. (2019). Circular Economy Indicators as a Supporting Tool for European Regional Development Policies. *Sustainability*, *11*(11), 3025. https://doi.org/10.3390/su11113025

Badan Pusat Statistik. (2024). *Indeks Pembangunan Manusia 2024*.

BPS Provinsi Jawa Barat. (2025). *Jawa Barat Dalam Angka 2025*. Badan Pusat Statistik Provinsi Jawa Barat.

Bille, R. A., Jensen, K. E., & Buitenwerf, R. (2023). Global patterns in urban green space are strongly linked to human development and population density. *Urban Forestry & Urban Greening*, *86*, 127980. https://doi.org/10.1016/j.ufug.2023.127980

BPS Jawa Barat. (2024). *Indeks Pembangunan Manusia Provinsi Jawa Barat 2023*.

Chen, L., Chen, T., Lan, T., Chen, C., & Pan, J. (2023). The Contributions of Population Distribution, Healthcare Resourcing, and Transportation Infrastructure to Spatial Accessibility of Health Care. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, *60*. https://doi.org/10.1177/00469580221146041

Dick-Sagoe, C. (2020). Decentralization for improving the provision of public services in developing countries: A critical review. *Cogent Economics & Finance*, *8*(1), 1804036. https://doi.org/10.1080/23322039.2020.1804036

Fabre, A., & Straub, S. (2023). The Impact of Public–Private Partnerships (PPPs) in Infrastructure, Health, and Education. *Journal of Economic Literature*, *61*(2), 655–715. https://doi.org/10.1257/jel.20211607

Fitrah Maharani, Fionasari, D., & Putri, A. M. (2025). The Influence of Trust in Government, Tax Fairness, Tax Education on Tax Morale. *Balance : Jurnal Akuntansi Dan Manajemen*, *4*(2), 1276–1288. https://doi.org/10.59086/jam.v4i2.1007

Fitrianigsih, D., & Dwi Kartikasari, M. (2024). Application of K-Means Clustering with the Elbow Method to Group Districts/Cities Based on Factors Affecting the Human Development Index in West Java Province. *Emerging Statistics and Data Science Journal*, *2*(2), 250–257. https://doi.org/10.20885/esds.vol2.iss.2.art18

Fratesi, U., & Perucca, G. (2019). EU regional development policy and territorial capital: A systemic approach. *Papers in Regional Science*, *98*(1), 265–282. https://doi.org/10.1111/pirs.12360

Galvani-Townsend, S., Martinez, I., & Pandey, A. (2022). Is life expectancy higher in countries and territories with publicly funded health care? Global analysis of health care access and the social determinants of health. *Journal of Global Health*, *12*, 04091. https://doi.org/10.7189/jogh.12.04091

Gharehbaghi, K., Clarkson, I., Hurst, N., & Rahmani, F. (2020). Transportation development for regional infrastructure: Implications for Australian rural areas. *Transportation Research Procedia*, *48*, 4003–4011. https://doi.org/10.1016/j.trpro.2021.04.002

Guo, Q., Yin, Z., & Wang, P. (2022). An Improved Three-Way K-Means Algorithm by Optimizing Cluster Centers. *Symmetry*, *14*(9), 1821. https://doi.org/10.3390/sym14091821

Hendajany, N., & Riyadi, D. R. (2022). Determinants of Regency/City Typology Based on HDI Indicators: Case from West Java, Indonesia. *Journal of Regional and Rural Development Planning*, *6*(3), 249–261. https://doi.org/10.29244/jp2wd.2022.6.3.249-261

Irsyifa Mayzela Afnan, Siti Hasanah, Anwar Fitrianto, Erfiani, & Alfa Nugraha. (2024). Village Clustering in West Java Province Based on the 2021 Village Development Index (IPD) Using the K-Prototypes Algorithm. *Jurnal Statistika Dan Aplikasinya*, *7*(2), 174–183. https://doi.org/10.21009/JSA.07206

Karthikeyan, B., George, D.J., Manikandan, G., & Thomas, T., (2020). A Comparative Study on K-Means Clustering and Agglomerative Hierarchical Clustering. *International Journal of Emerging Trends in Engineering Research*, *8*(5), 1600–1604. https://doi.org/10.30534/ijeter/2020/20852020

Kenger, Öm. N., Kenger, Z. D., Özceylan, E., & Mrugalska, B. (2023). Clustering of Cities Based on Their Smart Performances: A Comparative Approach of Fuzzy C-Means, K-Means, and K-Medoids. *IEEE Access*, *11*, 134446–134459. https://doi.org/10.1109/ACCESS.2023.3333753

Kusumastuti, A., Khoiron, A. M., & Achmadi. (2020). *Quantitative Research Methods*. Yogyakarta: CV Budi Utama.

Liang, S. (2024). Border Effects and the Growth Mechanism of Peripheral Economic Growth Centers in Integration. In *International Regional Economic Integration and the Development of China's Borderland Economies* (pp. 53–128). Springer Nature Singapore. https://doi.org/10.1007/978-981-97-3044-5_2

Long, X., Yu, H., Sun, M., Wang, X.-C., Klemeš, J. J., Xie, W., Wang, C., Li, W., & Wang, Y. (2020). Sustainability evaluation based on the Three-dimensional Ecological Footprint and Human Development Index: A case study on the four island regions in China. *Journal of Environmental Management*, *265*, 110509. https://doi.org/10.1016/j.jenvman.2020.110509

Madubun, J. (2024). Public services in island sub-districts: Towards geography-based governance. *Australian Journal of Public Administration*, *83*(3), 308–327. https://doi.org/10.1111/1467-8500.12586

Majumder, S., Kayal, P., Chowdhury, I. R., & Das, S. (2022). Regional disparities and development in India: evidence from Wroclow Taxonomy and K-means clustering. *GeoJournal*, *88*(3), 3249–3282. https://doi.org/10.1007/s10708-022-10805-2

Man, W., Wang, S., & Yang, H. (2021). Exploring the spatial-temporal distribution and evolution of population aging and social-economic indicators in China. *BMC Public Health*, *21*(1), 966. https://doi.org/10.1186/s12889-021-11032-z

Maulu, S., Hasimuna, O. J., Mutale, B., Mphande, J., & Siankwilimba, E. (2021). Enhancing the role of rural agricultural extension programs in poverty alleviation: A review. *Cogent Food & Agriculture*, *7*(1). https://doi.org/10.1080/23311932.2021.1886663

Moeller, J. (2025). Why and when you should avoid using z-scores in graphs displaying profile or group differences. *Journal for Person-Oriented Research*, *11*(2), 58–78. https://doi.org/10.17505/jpor.2025.28091

Mohamad Amin, N. A., Shaari, M. S. S., Sulong, A., & Masnan, F. (2024). The impacts of urbanization and economic growth on life expectancy in the Asean-5 countries. *Asian People Journal (APJ)*, *7*(1), 126–137. https://doi.org/10.37231/apj.2024.7.1.605

Nguea, S. M. (2023). Improving human development through urbanization, demographic dividend and biomass energy consumption. *Sustainable Development*, *31*(4), 2517–2535. https://doi.org/10.1002/sd.2528

Oti, E. U., Olusola, M. O., Eze, F. C., & Enogwe, S. U. (2021). Comprehensive Review of K-Means Clustering Algorithms. *International Journal of Advances in Scientific Research and Engineering*, *07*(08), 64–69. https://doi.org/10.31695/IJASRE.2021.34050

Pravitasari, A. E., Rustiadi, E., Priatama, R. A., Murtadho, A., Kurnia, A. A., Mulya, S. P., Saizen, I., Widodo, C. E., & Wulandari, S. (2021). Spatiotemporal Distribution Patterns and Local Driving Factors of Regional Development in Java. *ISPRS International Journal of Geo-Information*, *10*(12), 812. https://doi.org/10.3390/ijgi10120812

Priatama, R. A., Rustiadi, E., Widiatmaka, W., & Pravitasari, A. E. (2022). Physical Geographical Factors Leading to the Disparity of Regional Development: The Case Study of Java Island. *Indonesian Journal of Geography*, *54*(2). https://doi.org/10.22146/ijg.66729

Rocha, H., Kunc, M., & Audretsch, D. B. (2020). Clusters, economic performance, and social cohesion: a system dynamics approach. *Regional Studies*, *54*(8), 1098–1111. https://doi.org/10.1080/00343404.2019.1668550

Saccenti, E., Hendriks, M. H. W. B., & Smilde, A. K. (2020). Corruption of the Pearson correlation coefficient by measurement error and its estimation, bias, and correction under different error models. *Scientific Reports*, *10*(1), 438. https://doi.org/10.1038/s41598-019-57247-4

Saksono, A. (18 C.E.). *Changing impacts of state-led rural development policies in Indonesia from 1967 to 2018*. https://doi.org/10.26190/unsworks/22803

Sari, V. A., & Tiwari, S. (2024). The Geography of Human Capital: Insights from the Subnational Human Capital Index in Indonesia. *Social Indicators Research*, *172*(2), 673–702. https://doi.org/10.1007/s11205-024-03322-x

Sawicki, D. S., & Flynn, P. (1996). Neighborhood Indicators: A Review of the Literature and an Assessment of Conceptual and Methodological Issues. *Journal of the American Planning Association*, *62*(2), 165–183. https://doi.org/10.1080/01944369608975683

Shao, J., & Kong, X. (2024). Sustainable urban expansion and human development in <scp>China</scp> : An analysis using an environmentally improved human development index. *Sustainable Development*, *32*(5), 4397–4412. https://doi.org/10.1002/sd.2909

Singh, A. K. (2021). Population Growth and Economic Development: Theoretical Arguments and Empirical Findings— A Survey of Literature. *Indian Journal of Human Development*, *15*(3), 486–502. https://doi.org/10.1177/09737030211062105

Sun, T., & Abdullah, M. A. bin. (2025). Impact of Industrial Agglomeration on the Upgrading of China's Automobile Industry: The Threshold Effect of Human Capital and Moderating Effect of Government. *Sustainability*, *17*(7), 3090. https://doi.org/10.3390/su17073090

Surtiadi, E., & Sunsun, S. (2017). *Regional planning and development* (Andrea Emmas Prawitasari, Ed.). Yayasan Pustaka Obor Indonesia.

Torre, A. (2025). *François Perroux (1903–1987): Father of French Regional Science and Growth Pole Theory* (pp. 55–70). https://doi.org/10.1007/978-3-031-90625-1_4

Wedi, A., & Fathurrahman, R. (2025). An Analysis of the Determinant Factors of Poverty and Inequality in Indonesia and Their Implications for Village Development Planning. *Journal La Sociale*, *6*(5), 1477–1494. https://doi.org/10.37899/journal-la-sociale.v6i5.2389

Xu, R., Yue, W., Wei, F., Yang, G., Chen, Y., & Pan, K. (2022). Inequality of public facilities between urban and rural areas and its driving factors in ten cities of China. *Scientific Reports*, *12*(1), 13244. https://doi.org/10.1038/s41598-022-17569-2