p-ISSN 2338 – 3216
e-ISSN 2528 – 1070

JURNAL STATISTIKA
UNIVERSITAS MUHAMMADIYAH SEMARANG
14(1) 2026: 43-60

# CLUSTERING OF REGENCIES IN WEST KALIMANTAN BASED ON FINANCIAL RATIOS USING THE AVERAGE LINKAGE METHOD

**Hazwani Dhiya' Atiq Viatmaja[1*], Annisa Auliarahmi[2], Gabriella Simarmata[3]**
[1,2] Statistics Study Program, Tanjungpura University, Indonesia
[3] The Audit Board of Republic Indonesia Regional Office in West Kalimantan Province

***e-mail**: *h1091221043@student.untan.ac.id*

**Article Info:**
Received: December 26, 2025
Accepted: May 25, 2026
Available Online: June 2, 20206

**Keywords:**
Average Linkage; Budget Ratios; Cluster Analysis; Fiscal Independence; West Kalimantan

**Abstract:** Regional financial management serves as a crucial framework for assessing fiscal viability and policy impact on development. In West Kalimantan, identifying budget performance patterns is essential to support targeted financial policy decisions. This study aims to cluster regencies in West Kalimantan based on budget ratios derived from the 2024 Audited Examination Result Reports and to evaluate cluster quality. Hierarchical Cluster Analysis with the Average Linkage method and Manhattan distance was applied to four financial ratios across twelve regencies, with cluster validity assessed using Silhouette, Davies-Bouldin, and Dunn indices. The results indicate that three clusters represent the optimal solution. Cluster 1 comprises eight characterized by low fiscal independence and structural dependence on central government transfers, despite sufficiently effective locally generated revenue realization. Cluster 2 consists of three regencies demonstrating superior fiscal performance across all dimensions, supported by stronger local economic bases. Cluster 3 consists solely of Kubu Raya Regency with unique fiscal profile combining the highest Financial Independence Ratio and the lowest capital expenditure proportion in the sample, best interpreted as a fiscally distinct outlier. These findings indicate multidimensional fiscal disparities across West Kalimantan regencies that cannot be addressed through uniform policy approaches. Differentiated interventions are needed, including revenue extensification strategies and expenditure reorientation toward productive capital investment. Future research is recommended to incorporate time series data and micro fiscal variables for more comprehensive results.

## 1. INTRODUCTION

Regional financial management serves as the main framework to assessing fiscal viability, evaluating the effectiveness of budget allocation, and understanding the impact of policies on regional development. Reliable and comparable financial data are therefore essential for conducting meaningful inter regional analysis. In Indonesia, audited Local Government Financial Statements (LKPD) and Audit Reports (LHP) issued by the Audit Board (BPK) provide a credible basis for measuring fiscal performance at the regency and city levels, enabling structured comparisons across regions and supporting evidence-based policymaking [1].


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

Maintaining accountability in budget execution remains increasingly important in the context of regional fiscal disparities and the growing demand for transparent financial governance [2]. Differences in regional revenue capacity and expenditure allocation further emphasize the need to systematically map budget performance patterns in order to support more targeted and effective regional financial policies. In this regard, the use of audited LKPD data is particularly valuable, as it ensures the validity, reliability, and comparability of fiscal indicators across local governments [3].

Previous studies show that fiscal decentralization in West Kalimantan is accompanied by inefficiencies, reliance on intergovernmental transfers, and governance challenges, which reflect structural limitations in regional fiscal capacity and flexibility [4]. These conditions highlight the importance of analytical approaches capable of identifying patterns of fiscal variation across regencies. Such analysis can provide a more structured understanding of regional disparities and serve as a practical tool for improving financial policy design at the local level.

Literature shows that fiscal decentralization in West Kalimantan is accompanied by inefficiencies, reliance on transfer funds, and governance challenges, which indicate structural limitations in regional fiscal capacity and flexibility [4]. These conditions highlight the importance of analytical approaches capable of identifying patterns of fiscal variation across regencies. Such analysis can provide a more structured understanding of regional disparities and serve as a practical tool for improving financial policy design at the local level.

Clustering methods have been widely applied in regional studies to uncover patterns across areas based on socio economic and fiscal indicators. Empirical evidence shows that clustering can reveal meaningful group structures that are useful for policy formulation. For example, studies grouping Indonesian provinces based on poverty levels using hierarchical clustering methods, such as Average Linkage, demonstrate that regional variation can be organized into informative clusters that support targeted policy interventions. More broadly, hierarchical clustering approaches have been recognized as effective techniques for identifying natural groupings in data, particularly when observations exhibit hierarchical similarity structures that may not be adequately captured by partition based methods such as K-Means [5].

At the national level, clustering of regencies and cities has been conducted using various indicators, including education, human development, and poverty metrics, demonstrating the flexibility of clustering approaches across different domains [6]. However, in the context of West Kalimantan, previous studies have primarily applied clustering methods such as K-Means using non-fiscal indicators, and have not incorporated audited financial data as the basis for analysis [7]. This indicates a clear research gap, particularly in the application of hierarchical clustering methods to fiscal indicators derived from audited LKPD data [7].

To address this gap, this study employs Hierarchical Cluster Analysis using the Average Linkage method to classify regencies in West Kalimantan based on financial ratios. The Average Linkage method is particularly suitable for fiscal ratio data, which often exhibit variability and potential outliers. Unlike Single Linkage, which is highly sensitive to extreme values, and Complete Linkage, which tends to produce overly compact clusters, Average Linkage determines inter cluster distance based on the mean of all pairwise distances between cluster members. This characteristic makes it more robust in handling heterogeneous fiscal data and more appropriate for capturing balanced cluster structures in the presence of outlier observations [8].


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

This study contributes to the literature by integrating audited financial data with hierarchical clustering techniques to produce a more reliable and policy relevant classification of regional fiscal performance. By focusing on fiscal indicators derived from audited LKPD data, this research also enhances the reproducibility and credibility of clustering based four budget ratios, comprising the Financial Independence Ratio, PAD Effectiveness Ratio, Regional Tax Effectiveness Ratio, and Harmony Ratio, calculated from audited LKPD data. It also evaluates the quality of the resulting clusters using multiple validation indices. The findings are expected to provide a structured basis for identifying inter regional fiscal disparities and to support the formulation of more targeted and evidence based regional financial policies.

## 2. LITERATURE REVIEW

### 2.1. Data Standardization

Cluster analysis is one of the multivariate techniques that aims to group a number of objects into several smaller groups, where the objects in each group have a high degree of similarity to one another. Prior to distance calculation, standardization is necessary to ensure that variables measured on different scales contribute equally to the analysis [9]. This is particularly important in fiscal clustering studies where variables such as financial independence ratios and harmony ratios may have substantially different ranges and standard deviations, as observed in this dataset. The z score standardization formula is as follows

$$z = \frac{x - \bar{x}}{s} \tag{1}$$

Let $x$ represent each observed datum, then $\bar{x}$ describes the average (mean) of the data. Meanwhile $s$ is used to indicate the standard deviation.

### 2.2. Measuring Similarity Between Objects

Based on the basic principle of cluster analysis which aims to group objects with a certain level of similarity, a measure is needed to assess the extent of similarity between those objects [9]. In cluster analysis, there are various distance measures that can be used, one of which is the Manhattan distance. Manhattan distance is a method for calculating the distance between two points by summing the absolute values of the differences of each component or dimension between the two points. The Manhattan distance equation can be stated as follows :

$$d_{(XY)} = \sum_{j=1}^{p} |x_j - y_j| \tag{2}$$

with $d_{(XY)}$ denoting the Manhattan distance between object $X$ and object $Y$. The value of $p$ indicates the number of variables observed in both objects and for each $j$-th variable, $x_j$ represents the value of $j$ on object $X$, while $y_j$ is the value of $j$ on object $Y$.

### 2.3. Selecting a Clustering Procedure

The clustering procedure in cluster analysis can be categorized into two types: hierarchical and non-hierarchical methods [9]. In general, the steps in a hierarchical agglomerative clustering method for forming groups from $N$ objects are as follows:

a) Begin with the formation of $n$ initial clusters, where each object is treated as a separate cluster. If there is an $n$ x $n$ distance matrix $D = \{d_{(XY)}\}$.


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

b) Identify the pair of clusters with the smallest distance. Suppose the most similar clusters are $U$ and $V$, then $D = \{d_{(UV)}\}$, and the pair $U$ and $V$ is selected for merging.
c) Merge clusters $U$ and $V$ into a new cluster $(UV)$, and update the distance matrix as follows:
1) Remove the rows and columns corresponding to clusters $U$ and $V$
2) Add a new row and column representing the distance between the new cluster $(UV)$ and the remaining clusters.
d) Repeat steps (b) and (c) for $(n-1)$ iterations until all objects are merged into a single cluster at the end of the algorithm.

The hierarchical agglomerative clustering method used in this study is the Average Linkage method. The Average Linkage method follows a similar procedure to the two other hierarchical methods, namely Single Linkage and Complete Linkage. The main principle of this method is to use the average distance between all possible pairs of objects from two clusters. The clustering process begins by identifying the minimum distance $D = \{d_{(XY)}\}$, and merging the nearest objects, such as $U$ and $V$ to form a new cluster $(UV)$. Next, at step (c) of the previously described algorithm, the distance between the new cluster $(UV)$ and another object $W$, is calculated as follows:

$$d_{(UV)W} = \frac{1}{n_U \cdot n_W} \sum_{u \epsilon U} \sum_{w \epsilon W} d(u, w) \tag{3}$$

where $n_U$ and $n_W$ are the number of members in clusters $U$ and $W$ respectively, $u$ denotes each member of cluster $U$, and $w$ denotes each member of cluster $W$. This general formula considers all possible pairwise combinations between members of the two clusters, making it more representative than methods that rely only on the minimum or maximum distance. It should be noted that the simplified expression $d_{(UV)W} = average\ (d_{UW}, d_{VW})$ is only valid as a special case when each cluster contains exactly one member. The Average Linkage method is less sensitive to outliers than Single Linkage and less prone to chain formation than Complete Linkage, making it particularly appropriate for fiscal datasets where extreme values may be present [10].

## 2.4. Optimum Cluster Analysis

There is no definite rule in determining the ideal number of clusters in cluster analysis. The cluster considered optimum is obtained when the clustering result reaches a global optimum condition [11].

1. Cluster Density Analysis
The density of a cluster can be measured based on the minimum within-cluster variance and maximum between-cluster variance [12]. Suppose there are several clusters $c_1, c_2, \ldots, c_n$ each containing members $x_i$, where $i = 1, 2, \ldots, n$ and $n$ denotes the number of members in each cluster, while $\bar{x}_p$ represents the mean of cluster $p$, The variance of each cluster $p\left(\delta_p{}^2\right)$ can be calculated as follows:

$$\delta_p{}^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_i - \bar{x}_p\right)^2 \tag{4}$$

If $N$ represents the total number of members across all clusters, the within-cluster variance $\left(V_w{}^2\right)$ can be calculated using:


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

$$V_w{}^2 = \frac{1}{n-k}\sum_{i=1}^{k} (n_i - 1)\delta_i{}^2 \quad (5)$$

Furthermore, if each cluster $i$ has its own mean $(\bar{x})$**,** the between-cluster variance $\left(V_b{}^2\right)$ can be determined as follows:

$$V_b{}^2 = \frac{1}{k-1}\sum_{i=1}^{k} n_i(\bar{x}_i - \bar{x})^2 \quad (6)$$

Since an ideal cluster has minimum $V_w{}^2$ and maximum $V_b{}^2$, the optimal variance ratio $V$ can be expressed as:

$$V = \frac{V_w{}^2}{V_b{}^2} \times 100\% \quad (7)$$

2. Determining the Optimum Cluster
   The determination of the optimum number of clusters is carried out by analyzing the pattern of variance changes. Identification of variance movement is used to obtain clusters that reach the global optimum condition [11]. The potential position of the global optimum can be identified through implementation silhouette, elbow and gap statistics method.

a. Silhouette Method
   The silhouette method is used to determine the most optimal number of clusters by utilizing ratio-scaled data. In its application, the silhouette algorithm calculates the average distance between an object and all other objects located in the same cluster, as well as comparing it with the average distance to objects located in other clusters. A silhouette coefficient value approaching 1 indicates that the cluster formation is already optimal [12].
   One evaluation method that can be used to assess the level of optimality of clustering results is the silhouette coefficient. This evaluation functions to assess the extent to which the data has been grouped appropriately. The calculation process of the silhouette coefficient is carried out through the following steps [13]:

1) For each $i$-th object, calculate the average distance of that object to all other objects in the same cluster.
2) Next, for each $i$-th object, calculate the average distance to all objects located in different clusters, then take the smallest average distance value.
3) The silhouette coefficient value is then calculated using the following formula:

$$S_i = \frac{(b_i - a_i)}{max(a_i, b_i)} \quad (8)$$

with $S_i$ is silhouette coefficient value for the $i$-th object, $a_i$ is average distance of the $i$-th object to all objects in the same cluster, and $b_i$ is average distance of the $i$-th object to all objects in other clusters.

The silhouette coefficient value is in the range of -1 to 1. If $S_i = 1$, it means the object is in the appropriate cluster. If $S_i = 0$, the object is located between two clusters and it is unclear which cluster it belongs to, and if $S_i = -1$, then the formed cluster structure is poor, and the object should be in another cluster [13]. According to Kauffman and


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

Rousseeuw (1990), the quality of clustering results can be interpreted based on the silhouette coefficient value as shown in Table 1 below [12], [13]:

**Table 1.** Silhouette Coefficient Value

| Silhouette's Values | Interpretation |
|---|---|
| 0.71 – 1.00 | Strong Cluster |
| 0.51 – 0.70 | Good Cluster |
| 0.26 – 0.50 | Weak Cluster |
| 0.00 – 0.25 | Bad Cluster |

b. Elbow Method

The Elbow Method is used to determine the optimal number of clusters by observing the elbow point formed on the Sum of Squared Error (SSE) graph. This method is based on the largest decrease in SSE value between the number of clusters $(k-1)$ and $k$, thus that $k$ value is considered as the most appropriate number of clusters [12].

This method utilizes the SSE value from every possible number of clusters. In general, the more clusters used, the smaller the SSE value will be. Therefore, the optimal number of clusters is determined at the point where the decrease in SSE begins to slow down after a sharp decrease occurred previously. The formula for calculating SSE can be written as follows:

$$SSE = \sum_{k=1}^{K} \sum_{i \in X_k} |x_i - c_k|^2 \quad (9)$$

where $K$ represents the number of clusters, $x_i$ is the $i$-th data, and $c_k$ is the center of the $k$-th cluster.

c. Davies-Bouldin Index

The Davies-Bouldin Index (DBI) measures the average similarity between each cluster and its most similar neighboring cluster, where similarity is defined as the ratio of within-cluster scatter to between-cluster separation. A lower DBI value indicates better cluster quality, reflecting clusters that are more compact internally and more separated from one another [14]. The DBI is calculated as follows:

$$DBI = \frac{1}{k} \sum_{i=1}^{k} \max_{j \neq i} \left( \frac{s_i + s_j}{d_{ij}} \right) \quad (10)$$

where $k$ is the number of clusters, $s_i$ is the average distance between each member of cluster $i$, and the centroid of cluster $i$, $s_j$ is the average distance between each member of cluster $j$ and the centroid of cluster $j$, and $d_{ij}$ is the distance between the centroids of clusters $i$ and $j$. The optimal number of clusters corresponds to the solution with the minimum DBI value.

d. Dunn Index

The Dunn Index measures the ratio of the minimum inter cluster distance to the maximum intra cluster diameter. A higher Dunn Index value indicates better clustering quality,


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

reflecting clusters that are both compact and well separated from one another [15]. The Dunn Index is calculated as follows

$$DI = \frac{min_{1 \le i < j \le k}\, d(C_i, C_j)}{max_{1 \le m \le k}\, diam(C_m)} \tag{11}$$

where $k$ is the number of clusters, $d(C_i, C_j)$ is the distance between cluster $C_i$ and cluster $C_j$, representing the minimum separation between any two distinct clusters, and $diam(C_m)$ is the diameter of cluster $C_m$, defined as the maximum distance between any two members within the same cluster. The optimal number of clusters corresponds to the solution that maximizes the Dunn Index value, as a larger value reflects greater inter cluster separation relative to intra cluster spread. In this study, the Dunn Index is used alongside the Silhouette coefficient, Elbow, and Davies-Bouldin Index to confirm the optimal cluster solution, as convergence across multiple indices provides stronger evidence than relying on a single measure alone.

## 3. METHODOLOGY

### 3.1 Type and Source of Data

This study employs a quantitative descriptive approach. The data used are secondary data obtained from the Audit Reports (LHP) issued by the Audit Board of the Republic of Indonesia Regional Office in West Kalimantan Province, on the Local Government Financial Statements (LKPD) for all regencies in the Province of West Kalimantan. The dataset covers the 2024 fiscal year, or the most recent audited and fully published year available. The use of audited LKPD data ensures the validity, reliability, and comparability of regional financial data [16]. Moreover, regional literature indicates that the use of audited financial data enhances the quality of fiscal comparisons across regions and supports the formation of more informative fiscal clusters [12].

### 3.2 Analytical Variables

The analytical units in this study consist of 12 local government entities in West Kalimantan Province, representing 12 regencies. The scope of these analytical variables aligns with prior regional financial clustering studies that utilized audited financial data to compare fiscal performance among local government entities.

### 3.3 Research Variables

The variables used for clustering are budget ratios calculated from the Budget Realization Report (APBD) contained in the audited LKPD data. The variables are as follows:

a) Financial Independence Ratio ($x_1$)
Measures the extent to which a local government can finance its operations from Locally Generated Revenue (PAD). A higher ratio indicates greater fiscal autonomy and lower dependence on external funding [13]. The ratio is calculated as follows:

$$x_1 = \frac{Realized\ PAD}{Total\ Transfer\ Revenue} \times 100\% \tag{12}$$


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

where Total Transfer Revenue is intergovernmental transfer. The classification of financial independence is presented in Table 2

**Table 2.** Classification of Financial Independence Ratio

| Financial Independence (%) | Financial Capability | Relationship Pattern |
|---|---|---|
| $0 \leq x < 25$ | Very low | Instructive |
| $25 \leq x < 50$ | Low | Consultative |
| $50 \leq x < 75$ | Moderate | Participatory |
| $75 \leq x < 100$ | High | Delegative |

These categories indicate the degree of fiscal relationship between the local and central governments, where higher independence reflects stronger regional autonomy and reduced reliance on external funding [17].

b) PAD Effectiveness Ratio $(x_2)$
Measures the regional government's ability to realize the budgeted PAD by looking at the targets that have been set based on actual conditions. The smaller value of this ratio, the more efficient the performance of the PAD in collecting local revenue [18]. The PAD Effectiveness can be calculated using the following formula.

$$x_2 = \frac{Realized\ PAD}{Budgeted\ PAD} \times 100\% \tag{13}$$

The classification is presented in Table 3 [19].

**Table 3.** Classification of PAD Effectiveness

| Ratio Effectiveness (%) | PAD Effectiveness |
|---|---|
| $x > 100$ | Very effective |
| $90 \leq x < 100$ | Effective |
| $80 \leq x < 90$ | Sufficiently |
| $60 \leq x < 800$ | Less effective |
| $x < 60$ | Ineffective |

c) Regional Tax Effectiveness Ratio $(x_3)$
Measures the ability of the regional government to realize the budgeted PAD compared to the revenue targets that have been allocated based on the actual conditions of the region [20], The regional tax effectiveness ratio of PAD can be calculated using the following formula


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

$$x_3 = \frac{Realized\ Local\ Tax}{Budgeted\ Local\ Tax} \times 100\% \quad (14)$$

The classification is presented in Table 4 [21].

**Table 4.** Classification of Regional Tax Effectiveness Ratio

| Regional Tax Effectiveness (%) | Category |
|---|---|
| $x > 100$ | Very effective |
| $90 \leq x < 100$ | Effective |
| $80 \leq x < 90$ | Sufficiently |
| $60 \leq x < 800$ | Less effective |
| $x < 60$ | Ineffective |

d) Harmony Ratio ($x_4$)

The Harmony Ratio reflects the extent to which regional governments optimally allocate their budget between operational expenditure and capital expenditure. According to Tim Litbang Depdagri Fisipol UGM [22], the harmony ratio consists of two components, the operational expenditure ratio and the capital expenditure ratio. The operational expenditure ratio measures the proportion of operational spending relative to total regional expenditure, while the capital expenditure ratio measures the proportion of capital spending relative to total regional expenditure. The two formulas are as follows

$$Operational\ Expenditure\ Ratio = \frac{Realized\ Operational\ Expenditure}{Total\ Regional\ Expenditure} \times 100\% \quad (15)$$

$$Capital\ Expenditure\ Ratio = \frac{Realized\ Capital\ Expenditure}{Total\ Regional\ Expenditure} \times 100\% \quad (16)$$

In this study, the capital expenditure component is adopted as the clustering variable ($x_5$), as it more directly captures the orientation of regional governments toward productive development spending and reflects the quality of budget allocation in a more substantive manner [23]. The ratio is therefore calculated as follows

$$x_4 = \frac{Realized\ Capital\ Expenditure}{Budgeted\ Regional\ Expenditure} \times 100\% \quad (17)$$

The assessment of expenditure harmony follows the qualitative criteria proposed by Tim Litbang Depdagri Fisipol UGM, which are based on the relative comparison between operational and capital expenditure [22]. These criteria are presented in Table 5.

**Table 5.** Criteria for Expenditure Harmony

| Comparison | Criteria |
|---|---|
| Operational Expenditure > Capital Expenditure | Poor |


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

| Comparison | Criteria |
|---|---|
| Operational Expenditure = Capital Expenditure | Sufficient |
| Operational Expenditure < Capital Expenditure | Good |

However, this classification is used solely for descriptive interpretation and is not directly incorporated into the clustering model. The clustering analysis utilizes only the capital expenditure ratio ($x_4$) as a proxy for expenditure harmony, ensuring consistency with the selected set of variables while maintaining interpretability.

A higher value of $x_5$indicates a greater proportion of capital expenditure, suggesting that the regional government places stronger emphasis on development-oriented spending. Conversely, a lower value indicates that the budget is predominantly allocated to operational expenditure, reflecting limited allocation for long-term investment and infrastructure development.

### 3.4 Data Analysis Stages

Cluster analysis was carried out through several stages. First, data from the 2024 Audited LKPD for 12 regencies in West Kalimantan Province were collected using four predetermined financial ratios. Second, a variable redundancy check was conducted using the Pearson correlation matrix to identify any pairs of variables with correlations exceeding 0.80, which could potentially cause dominance or distortion in distance calculations. In the context of cluster analysis, the concern is not multicollinearity in the regression sense, but rather whether any variable pair is sufficiently redundant to bias inter object distances [7].

Third, all four ratios were standardized using z scores to ensure that variables with different scales and ranges contributed equally to the distance calculations. Fourth, the distance between objects was calculated using the Manhattan distance method. Fifth, the optimal number of clusters was determined using multiple complementary approaches, including the Elbow method based on within cluster dispersion (WSS), the Silhouette coefficient.

Sixth, the clustering process was performed using Hierarchical Cluster Analysis with the Average Linkage method, which calculates the distance between clusters based on the mean of all pairwise distances between their respective members. The cluster groupings were then visualized using a dendrogram. Finally, the quality of the resulting clusters was evaluated using the Silhouette coefficient, Davies-Bouldin Index, and Dunn Index to assess cluster compactness and separation, ensuring the robustness of the clustering results. All data analyses were conducted using R Studio software.

## 4. RESULTS AND DISCUSSION

### 4.1 Statistics Descriptive

The financial ratio values for each variable across the 12 regencies in West Kalimantan Province, based on the 2024 audited LKPD, are presented in Table 6.


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

**Table 6.** Budget Ratio Values for 12 Regencies in West Kalimantan 2024

| Regency | $x_1$(%) | $x_2$(%) | $x_3$(%) | $x_4$(%) |
|---|---|---|---|---|
| Sambas | 12.24 | 73.36 | 95.12 | 14.50 |
| Bengkayang | 7.61 | 71.78 | 88.99 | 17.11 |
| Landak | 4.99 | 83.97 | 59.56 | 12.46 |
| Mempawah | 14.29 | 110.97 | 103.98 | 23.05 |
| Sanggau | 10.80 | 113.56 | 122.00 | 15.10 |
| Ketapang | 11.32 | 102.43 | 89.44 | 23.46 |
| Sintang | 9.43 | 105.16 | 60.68 | 13.24 |
| Kapuas Hulu | 5.35 | 101.42 | 92.99 | 15.17 |
| Melawi | 5.18 | 83.50 | 83.76 | 10.07 |
| Kayong Utara | 6.10 | 79.31 | 95.40 | 11.38 |
| Sekadau | 6.80 | 68.66 | 57.77 | 14.39 |
| Kubu Raya | 16.47 | 95.20 | 100.31 | 10.42 |

Accordingly, the descriptive statistics are summarized in Table 7.

**Table 7.** Statistics Descriptive

| Variable | Description | Unit | Min | Mean | Max | SD |
|---|---|---|---|---|---|---|
| $x_1$ | Financial Independence Ratio | % | 4.99 | 9.21 | 16.47 | 3.83 |
| $x_2$ | PAD Effectiveness Ratio | % | 68.66 | 90.78 | 113.56 | 15.92 |
| $x_3$ | Regional Tax Effectiveness Ratio | % | 57.77 | 87.50 | 122.00 | 19.50 |
| $x_4$ | Harmony Ratio | % | 10.07 | 15.03 | 23.46 | 4.36 |

The Financial Independence Ratio ($x_1$) has an average value of 9.21% with a standard deviation of 3.83%, indicating that all regencies in West Kalimantan remain highly dependent on central government transfer funds to finance governmental and development activities. The relatively low standard deviation suggests moderate variation in fiscal independence across regencies, with no extreme outliers observed.

The PAD Effectiveness Ratio ($x_2$) has an average value of 90.78%, which falls within the Effective category. Although the overall average indicates good performance, there is considerable variation among regencies, as reflected by a standard deviation of 15.92%, suggesting differences in administrative capacity in managing local revenue. Sekadau (68.66%) and Bengkayang (71.78%) fall into the Less Effective category, while Sanggau


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

(113.56%) and Mempawah (110.97%), which exceed 100%, are classified as Very Effective. This wide variation indicates that, despite generally effective PAD performance at the provincial level, certain regencies still require particular attention to improve their local revenue realization capacity.

The Regional Expenditure Performance Ratio ($x_3$) has an average value of 87.50%, with the highest variability among all variables (SD = 19.50%). This is largely driven by several regencies such as Landak (59.56%), Sekadau (57.77%), and Sintang (60.68%), which exhibit relatively low expenditure realization compared to others. This condition indicates potential constraints in budget implementation, thereby requiring special attention, particularly in strengthening administrative and managerial capacity. On the other hand, Sanggau (122.00%) and Mempawah (103.98%) demonstrate relatively strong budget execution performance.

The Expenditure Harmony Ratio ($x_4$) has an average value of 15.03%, indicating that operational expenditure still significantly dominates capital expenditure across all regencies. Based on the criteria proposed by Tim Litbang Depdagri Fisipol UGM in Syukur (2021), all regencies fall into the Poor category, as the proportion of operational expenditure exceeds that of capital expenditure. This reflects a structural challenge in regional budget allocation in West Kalimantan, where spending orientation remains more focused on operational needs rather than long-term development investment.

## 4.2 Variable Redudancy Check

Before to the clustering process, a Pearson correlation matrix was computed to identify whether any pairs of variables exhibit high correlations that could potentially lead to redundancy or dominance in distance calculations. The results of the correlation analysis are presented in Table 8.

**Table 8.** Pearson Correlation Matrix

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|
| $x_1$ | 1.00 | 0.42 | 0.50 | 0.34 |
| $x_2$ | 0.42 | 1.00 | 0.44 | 0.39 |
| $x_3$ | 0.50 | 0.44 | 1.00 | 0.23 |
| $x_4$ | 0.34 | 0.38 | 0.23 | 1.00 |

The results indicate that no pair of variables exceeds the commonly accepted threshold of 0.80, which is typically used as an indicator of high redundancy among variables [24]. The remaining variable pairs show low to moderate levels of correlation. Therefore, all variables are considered suitable for inclusion in the cluster analysis, as no variable is identified as dominant or redundant in the distance calculation

## 4.3 Clustering Using the Average Linkage Method

Cluster analysis was conducted using the Hierarchical Clustering method with Average Linkage and Manhattan distance. Prior to clustering, the data were standardized using z scores to eliminate bias due to differences in variable scales. The results of determining the optimal number of clusters are presented in the Figure 1.

Based on the Elbow method graph in Figure 1, the most substantial decrease in the total within cluster sum of squares occurs from $k = 1$ to $k = 2$. After $k = 2$, the rate of decrease gradually slows without forming a very clear elbow point. Meanwhile, the Silhouette method graph in Figure 1 shows a clear peak at $k = 2$ with a value of 0.35, followed by a secondary peak at $k = 3$ with value of 0.31. Both graphs generally indicate $k = 2$ as the primary candidate. However, $k = 3$ is also worth considering based on the secondary peak observed in the Silhouette method. The results of the provisional cluster grouping are then visualized in the dendrogram as shown below.

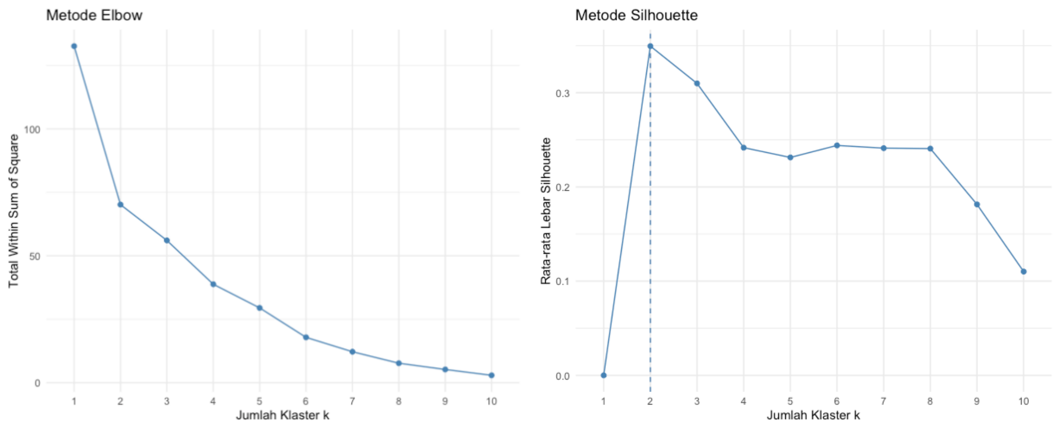


**Fig 1.** Grafik Elbow dan Silhouette

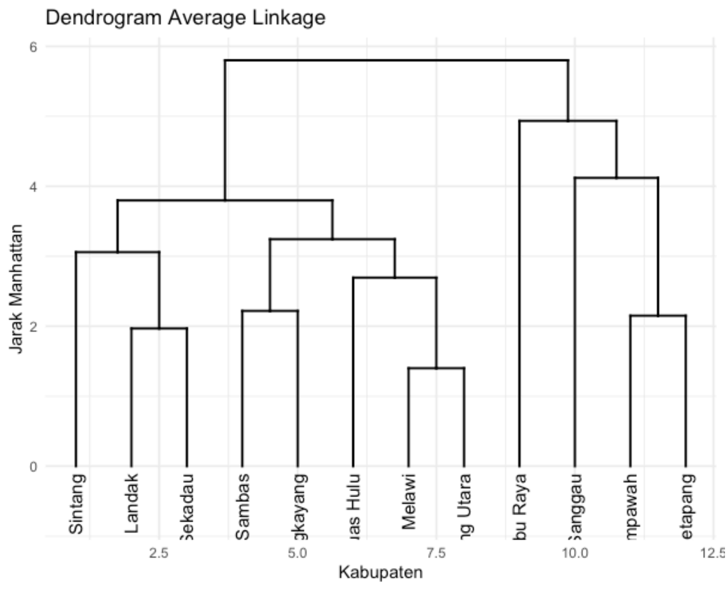


**Fig 2.** Dendrogram of Clustering Using the Average Linkage Method

Next, the cluster validation for determining the optimal number of clusters will be carried out based on the explanation in sub section 2.4.

### 4.4 Validation of the Optimal Number of Clusters


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

To confirm the clustering quality for $k = 2$ or $k = 3$, internal validity indices will be calculated, and the comparison of these values can be seen in Table 9.

**Table 9.** Comparison of Cluster Validation Results

| Number of Clusters (K) | Average Silhouette (Global) | Davies-Bouldin | Dunn Index |
|---|---|---|---|
| 2 | **0.35** | 1.09 | 0.60 |
| 3 | 0.28 | **0.73** | **0.72** |

The results indicate that the indices do not produce entirely consistent conclusions. The Silhouette value for $k = 2$ $(0.35)$ is higher than that for $k = 3$ $(0.28)$, suggesting slightly better cluster separation for the two cluster solution. However, the higher Dunn Index for $k = 2$ . However, the higher Dunn Index for $k = 3$ $(0.72)$ and the lower Davies Bouldin Index for $k = 3$ $(0.73)$ indicate that the three cluster solution exhibits better separation and compactness.

Therefore, the $k = 3$ solution is selected as the optimal number of clusters, as it provides a better balance between statistical validity and interpretability. The cluster membership for each group is presented in Table 10.

**Table 10.** Grouping of Regencies into Clusters

| ***Group*** | ***Member in the Cluster*** |
|---|---|
| 1 | Sambas, Bengkayang, Landak, Sintang, Kapuas Hulu, Melawi, Kayong Utara, dan Sekadau |
| 2 | Mempawah, Sanggau, Ketapang |
| 3 | Kubu Raya |

### 4.5 Characteristics of Each Cluster

The average value of each variable in each cluster, as presented in Table 11, can be used as a reference for interpreting the characteristics of each cluster.

**Table 11.** Average Value of each Variable in each Cluster

| Variable | Description | Cluster from k (%) | | |
|---|---|---|---|---|
| | | Cluster 1 | Cluster 2 | Cluster 3 |
| $x_1$ | Financial Independence Ratio | 7.21 | 12.14 | 16.47 |
| $x_2$ | PAD Effectiveness Ratio | 83.39 | 108.99 | 95.20 |
| $x_3$ | Regional Tax Effectiveness Ratio | 79.28 | 105.14 | 100.31 |
| $x_4$ | Harmony Ratio | 13.54 | 20.54 | 10.42 |


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

Cluster 1 consists of eight regencies, with an average $x_1$ of 7.21%. All members of this cluster fall into the Instructive category, indicating a very high dependence on central government transfer funds. This condition is consistent with the relatively weaker economic position of most regencies in this cluster. Melawi (IDR29.74 million per capita), Landak (IDR35.71 million per capita), and Bengkayang (IDR40.33 million per capita) record relatively low GRDP per capita in West Kalimantan [25], implying a limited local revenue base.

The average $x_2$ of 83.39% indicates a moderately effective PAD realization, although several regencies such as Sambas (73.36%) and Bengkayang (71.78%) still require improvement. A more prominent challenge is observed in $x_3$, where the average expenditure performance is only 79.28%, with Sekadau (57.77%), Landak (59.56%), and Sintang (60.68%) show particularly low realization levels, suggesting constraints in budget implementation. These issues are likely associated with limited administrative capacity and the geographical complexity of relatively large and remote areas. The average $x_4$ of 13.54% places all members in the Poor category, reflecting the dominance of operational expenditure and the limited fiscal space for development investment.

Cluster 2 consists of three regencies that consistently demonstrate better fiscal performance than Cluster 1 across all dimensions. The average $x_2$ of 108.99% places this cluster in the Very Effective category, with Sanggau (113.56%) and Mempawah (110.97%) significantly exceeding their PAD targets. This achievement is supported by a stronger economic base. Ketapang (IDR65.21 million per capita) ranks as the second highest GRDP per capita in West Kalimantan, while Sanggau (IDR51.97 million per capita) ranks sixth [25]. The dominance of mining and plantation sectors in these regencies provides a stronger foundation for effective local revenue generation.The average $x_3$ of 105.140% indicates excellent expenditure realization, reflecting strong budget execution capacity. The average $x_4$ of 20.54% is also the highest among all clusters, indicating a relatively better orientation toward development investment, although still categorized as Poor in absolute terms. Meanwhile, the average $x_1$ of 12.14% remains within the Instructive category, but is nearly twice that of Cluster 1, suggesting relatively lower dependence on transfers.

Kubu Raya is separated as its own cluster due to a unique combination of characteristics not shared by other regencies. It has the highest $x_1$ value in the sample (16.47%) and the second lowest $x_4$ (10.42%). From a fiscal independence perspective, Kubu Raya's position is supported by a relatively strong economic base, with a GRDP per capita of IDR64.21 million (third highest in West Kalimantan). Its proximity to Pontianak City provides greater access to trade and industrial activities, contributing to higher PAD relative to transfer revenues [25]. The values of $x_2$ (95.20%) and $x_3$ (100,31%) for Kubu Raya lie between those of Cluster 1 and Cluster 2, and are not sufficiently extreme to drive separation along these dimensions. The key distinguishing factor is the low $x_4$ (10.42%) which contrasts with its relatively high $x_1$. This presents an important policy implication: relatively stronger fiscal independence has not yet been accompanied by optimal expenditure allocation. As a singleton, Kubu Raya is more appropriately interpreted as a regency with a unique fiscal profile rather than as a representative of a broader cluster type.


[DOI: 10.14710/JSUNIMUS.14.1.2026.43-60]

Overall, the clustering results reveal a clear and multidimensional fiscal disparity among regencies in West Kalimantan. Cluster 2 consistently outperforms others across all fiscal dimensions, supported by a stronger economic base. Cluster 1 reflects moderate fiscal conditions with high transfer dependence and suboptimal expenditure performance. Cluster 3 represents a unique profile, characterized by relatively higher fiscal independence but very limited capital expenditure allocation. These findings confirm that fiscal disparities in West Kalimantan cannot be addressed through a uniform policy approach. Each group of regencies faces distinct challenges and requires differentiated strategies that take into account their respective structural economic conditions. This is aligned with the spirit of Law Number 1 of 2022 concerning Fiscal Relations between the Central Government and Regional Governments, which emphasizes the sustainable strengthening of regional fiscal independence.

## 5. CONCLUSION

This study identified three optimal clusters of regencies in West Kalimantan based on fiscal performance using hierarchical clustering with the Average Linkage method and Manhattan distance. The three cluster solution was selected based on the Dunn Index (0.72) and Davies-Bouldin Index (0.73), both of which favored $k = 3$ over $k = 2$. The results reveal significant and multidimensional fiscal disparities across regencies, particularly in terms of fiscal independence, PAD effectiveness, expenditure performance, and budget allocation orientation. Cluster 1 is characterized by high dependence on central transfers and moderate performance, Cluster 2 demonstrates consistently strong fiscal performance supported by a stronger economic base, while Cluster 3 represents a unique fiscal profile with relatively higher independence but suboptimal expenditure allocation. These disparities suggest that fiscal policy in West Kalimantan requires differentiated approaches tailored to the specific characteristics of each cluster rather than a uniform intervention.

This study is limited by the use of data from a single fiscal year (2024), which does not capture long term trends or fluctuations in fiscal performance. In addition, the analysis is based on four macro level financial ratios, which may not fully represent the complexity of regional fiscal conditions. The relatively small sample size (12 regencies) and the presence of a singleton cluster also limit the generalizability of the findings. Future studies are recommended to incorporate time series data covering multiple fiscal years to better capture dynamic fiscal patterns. Additionally, the inclusion of more detailed fiscal variables, such as personnel expenditure ratios and the contribution of regionally owned enterprises (BUMD), would provide a more comprehensive analysis. Comparing clustering results using different linkage methods is also suggested to evaluate the robustness of the clustering structure.

## REFERENCES

[1] A. N. Fadila, A. P. Sari, and K. Yunita, “Telaah Audit Refocusing APBD Kalbar Dalam Upaya Penanganan Pandemi COVID-19 (Studi Pada Inspektorat Kalimantan Barat),” *JAAKFE UNTAN (Jurnal Audit dan Akuntansi Fakultas Ekonomi Universitas Tanjungpura)*, vol. 12, no. 2, p. 176, Jun. 2023, doi: 10.26418/jaakfe.v12i2.59660.

[2] Mursalin and Khaeriyah, "Accountability and Transparency in Regional Budget Management for Good Governance," *International Journal of Economics, Commerce, and Management*, vol. 2, no. 4, pp. 140–148, Oct. 2025, doi: 10.62951/ijecm.v2i4.991.

[3] S. Islam, "A Systematic Review of Public Budgeting Strategies in Developing Economies: Tools for Transparent Fiscal Governance," *American Journal of Advanced Technology and Engineering Solutions*, vol. 01, no. 01, pp. 602–635, Apr. 2025, doi: 10.63125/wm547117.

[4] Amirullah, Giriati, and Mustaruddin, "The Impact of Fiscal Decentralization on Development Performance in West Kalimantan with Opportunistic Behavior as a Mediating Variable," *International Journal of Social Science and Human Research*, vol. 07, no. 09, pp. 6770–6785, Sep. 2024, doi: 10.47191/ijsshr/v7-i09-08.

[5] W. B. Xie, Y. L. Lee, C. Wang, D. B. Chen, and T. Zhou, "Hierarchical clustering Supported by Reciprocal Nearest Neighbors," *Inf. Sci. (N. Y).*, vol. 527, pp. 279–292, Apr. 2020, doi: 10.1016/j.ins.2020.04.016.

[6] M. N. H. T. Putra, "Klasterisasi Kabupaten/Kota di Provinsi Kalimantan Barat Berdasarkan Rasio Guru dengan Murid di Tingkat SD, SMA, dan SMA Serta IPM Tahun 2022 Menggunakan Metode K-Means Clustering," *Journal FORMASI: Forum Analisis Statistik Juni*, vol. 3, no. 1, pp. 42–50, Jun. 2023, doi: 10.xxxxx/formasi.2021.1.1.1-12.

[7] D. Saputra, A. Ardania, S. Putri, A. T. J. A. Asri, and L. Harsyah, "Analisis Cluster untuk Pengelompokkan Provinsi di Indonesia berdasarkan Tingkat Kemiskinan menggunakan Metode Average Linkage," *Indonesian Journal of Applied Statistics and Data Science*, vol. 1, no. 1, Nov. 2024, [Online]. Available: https://journal.unram.ac.id/index.php/ijasds

[8] T. Raffinot, "Asset Allocation, Economic Cycles and Machine Learning," Université Paris-Dauphine, Paris, 2018. [Online]. Available: https://theses.hal.science/tel-01872176v1

[9] S. Hadi, I. Gunawan, and J. Dalle, *Statistika Inferensial Teori dan Aplikasinya*, 2nd ed. Depok: PT RajaGrafindo Persada, 2016.

[10] X. Chen *et al.*, "Clustering Analysis for The Evolutionary Relationships of SARS-CoV-2 Strains," *Sci. Rep.*, vol. 14, Mar. 2024, doi: 10.1038/s41598-024-57001-5.

[11] D. A. Pradaningtyas, T. Margawati, and J. T. Putro, "Disparities Among Districts in Central Java Province: Cluster Analysis Based on Several Well-Being Indicators," *Jurnal Ekonomi Pembangunan*, vol. 20, no. 1, pp. 34–47, Jun. 2022.

[12] N. Thamrin and A. W. Wijayanto, "Analisis Cluster dengan Menggunakan Hard Clustering dan Soft Clustering untuk Pengelompokkan Tingkat Kesejahteraan Kabupaten/Kota di Pulau Jawa," *Indonesian Journal of Statistics and Its Applications*, vol. 5, no. 1, pp. 141–160, Mar. 2021, doi: 10.29244/ijsa.v5i1p141-160.

[13] A. T. R. Dani, S. Wahyuningsih, and nanda A. Rizki, "Penerapan Hierarchical Clustering Metode Agglomerative pada Data Runtun Waktu," *Jambura Journal of Mathematics*, vol. 1, no. 2, pp. 64–78, Jul. 2019, [Online]. Available: http://ejurnal.ung.ac.id/index.php/jjom,P-

[14] I. F. Fauzi, M. G. Resmi, and T. I. Hermanto, "Penentuan Jumlah Cluster Optimal Menggunakan Davies Bouldin Index pada Algoritma K-Means untuk Menentukan Kelompok Penyakit," vol. 7, no. 2, pp. 1–15, 2023.

[15] M. T. Jatipaningrum, S. E. Azhari, and K. Suryowati, "Pengelompokan Kabupaten dan Kota di Provinsi Jawa Timur Berdasarkan Tingkat Kesejahteraan dengan Metode K-Means dan Density-Based Spatial Clustering of Applications with Noise," *Jurnal Derivat*, vol. 9, no. 1, Jul. 2022, [Online]. Available: https://jatim.bps.go.id

[16] I. K. Indriani, Y. A. Nugroho, and W. A. Sari, “Penilaian Kondisi Keuangan Provinsi Kalimantan Barat - Proyeksi dalam Menghadapi Pandemi COVID 19,” *Journal of Economics and Business Aseanomics*, vol. 6, no. 2, pp. 30–52, 2021.

[17] N. Zukhri, “Kinerja Keuangan Provinsi Kepulauan Bangka Belitung Ditinjau Dari Derajat Kemandirian, Ketergantungan dan Desentralisasi Fiskal,” *Indonesian Treasury Review: Jurnal Perbendaharaan, Keuangan Negara dan Kebijakan Publik*, vol. 5, no. 2, pp. 143–149, 2020.

[18] Siswanto and D. A. Maylani, “Analisis Laporan Keuangan untuk Menilai Kinerja Keuangan Pemerintah Daerah,” *Jurnal Nominal Barometer Riset Akuntansi dan Manajemen*, vol. 11, no. 1, pp. 130–138, Apr. 2022, doi: 10.21831/nominal.v11i1.48423.

[19] P. Rizki and M. Erwati, “Analisis Kemandirian, Efektivitas dan Efisiensi Pada Kinerja Keuangan Daerah (Studi Kasus Hanya Pemerintah Daerah Provinsi Jambi Tahun 2019-2023 Bukan Pada Kabupaten/Kota),” *Jambi Accounting Review (JAR)*, vol. 5, no. 3, pp. 1–9, Dec. 2024, doi: 10.22437/jar.v5i3.38738.

[20] R. Hafizi and F. A. Amalia, “Kinerja Keuangan Pemerintah Daerah Kabupaten Jember,” *JATI: Jurnal Akuntansi Terapan Indonesia*, vol. 05, no. 02, pp. 116–130, Oct. 2022, doi: 10.18196/jati.v5i2.141.

[21] W. Widodo, N. M. Widodo, and A. Prihadyatama, “Rasio Efektivitas dan Kontribusi Pajak Daerah Terhadap Pendapatan Asli Daerah (Studi Kasus Kota Madiun Tahun 2018-2022),” *Jurnal Maneksi*, vol. 13, no. 1, pp. 9–5, Mar. 2024, doi: 10.31959/jm.v13i1.1898.

[22] R. Wulandari, B. A. H. Lestari, and A. B. Suryantara, “Analisis Rasio Keuangan dalam Mengukur Kinerja Keuangan Pemerintah Daerah Kota Mataram,” *Jurnal Risma*, vol. 3, no. 2, pp. 56–69, Jun. 2023.

[23] M. Dewi and A. P. Nilasari, “Analisis Laporan Keuangan Pemerintah Daerah Kabupaten Wonosobo,” *Jurnal Ekonomi Kreatif dan Manajemen Bisnis Digital*, no. 1, pp. 1–9, 2022.

[24] W. Alwi and Muh. Hasrul, “Analisis Klaster untuk Pengelompokan Kabupaten/Kota di Propinsi Sulawesi Selatan Berdasarkan Indikator Kesejahteraan Rakyat,” *Jurnal MSA (Matematika dan Statistika serta Aplikasinya)*, vol. 6, no. 1, pp. 35–42, Jun. 2018, doi: 10.24252/msa.v6i1.4782.

[25] A. D. Darmawan, “Data 2024: PDRB ADHB per Kapita Kabupaten Sekadau Rp.42,39 Juta,” databoks.katadata.co.id.