Sosioedukasi
JURNAL ILMIAH ILMU PENDIDIKAN DAN SOSIAL
ISSN: 2086-6135 E-ISSN: 2541-612x
Volume 15 No. 2. June 2026 https://ejournal.unibabwi.ac.id/index.php/sosioedukasi/index

SEGMENTATION OF THE FINANCIAL PERFORMANCE OF PUBLIC COMPANIES ON THE INDONESIA STOCK EXCHANGE USING THE K-MEANS CLUSTERING METHOD

Ana Mayasari[1a*], Puji Harto[2]

[1,2]Master's Program in Accounting, Universitas Diponegoro, Indonesia

[a]*E-mail: mayaa1720@gmail.com*

(*) Corresponding Author
mayaa1720@gmail.com

ARTICLE HISTORY
Received : 20-05-2026
Revised : 07-06-2026
Accepted : 30-06-2026

KEYWORDS
K-Means Clustering;
Data Mining;
Financial Performance;
Indonesia Stock Exchange;
Issuer Segmentation;

ABSTRACT
This study aims to cluster the financial performance of companies listed on the Indonesia Stock Exchange (IDX) using the K-Means Clustering method. Although financial ratio analysis has been widely used to assess company conditions, research that comprehensively maps the financial characteristics of all IDX issuers using a data mining approach remains relatively limited. The study uses secondary data from 2024 financial reports with four main variables, namely Current Ratio, EBIT Margin, Leverage Ratio, and Debt-to-Equity Ratio. The determination of the optimal number of clusters was carried out using the Elbow and Silhouette methods, which produced four clusters as the best structure. The results show that the majority of issuers fall into the group with relatively healthy and balanced financial conditions, while some others have high leverage levels or show indications of financial pressure. Statistical testing shows significant differences between clusters across all variables used. This study contributes to providing a data mining based financial segmentation framework that can support investment decision making, portfolio diversification, and risk monitoring in the Indonesian capital market.

INTRODUCTIONS

The trajectory of the Indonesian capital market over the preceding decade has exhibited considerable growth, as evinced by the marked proliferation in the number of companies listed on the Indonesia Stock Exchange (IDX). As of 2024, there were more than 900 publicly listed companies actively traded on the IDX, spanning various industrial sectors ranging from consumer goods, finance and infrastructure to technology (Otoritas Jasa Keuangan, 2023). This diversity of listed companies presents a unique challenge for investors, analysts and regulators in evaluating and comparing financial performance effectively and efficiently (Saadah et al., 2025).

Conventional analysis of a company's financial performance is generally predicated upon the utilization of financial ratios, encompassing liquidity, profitability, solvency, and activity ratios (Brigham, 2004). Although this approach furnishes a useful fundamental overview, its individual application across hundreds of listed companies can

JURNAL ILMIAH ILMU PENDIDIKAN DAN SOSIAL
ISSN: 2086-6135 E-ISSN: 2541-612x
Volume 15 No. 2. June 2026 https://ejournal.unibabwi.ac.id/index.php/sosioedukasi/index

prove inefficient and tends to engender fragmented analysis. Therefore, a more systematic and data-driven approach is warranted to unearth latent patterns in the financial performance of listed companies on a large scale.

Data mining is the process of extracting useful information from large volumes of data through the utilization of statistical techniques, machine learning, and artificial intelligence (Camm et al., 2020). One of the most widely employed data mining techniques in the context of segmentation is clustering, particularly K-Means clustering. This method partitions observations into a number of groups (clusters) predicated on the similarity of their characteristics, such that objects within a single cluster exhibit greater resemblance to one another than to objects residing in other clusters (Camm et al., 2020).

K-means clustering has increasingly been adopted in financial analysis, with applications ranging from the classification of firms by their risk exposure (Xu et al., 2024) to the segmentation of investor behavior in European capital markets (Rusu et al., 2023). In the Indonesian context, this approach has also been applied to the IDX banking sector, where analyses revealed differentiated clusters reflecting varying risk–return characteristics (Aditiya & Sulistiyowati, 2026).

Based on the background delineated above, this study focuses on addressing several cardinal issues pertaining to the financial performance characteristics of issuers listed on the IDX based on liquidity, profitability, and solvency indicators, as well as ascertaining the optimal number of clusters capable of meaningfully discriminating among groups of issuers in consonance with their financial characteristics. Furthermore, this study seeks to interpret the profile of each formed cluster along with its business implications to support investment decision-making. The study aims to apply the K-Means Clustering method to the financial data of IDX-listed companies using a comprehensive secondary dataset. Four financial variables were selected as clustering attributes: Current Ratio (liquidity), EBIT Margin (operating profitability), Leverage Ratio (capital structure), and Debt-to-Equity Ratio (solvency). The selection of these variables is based on their relevance in financial risk assessment models, including the Altman Z-Score (Altman, 1968), which has been empirically proven to distinguish healthy companies from those at risk of bankruptcy.

Although clustering methods have been widely used in financial and capital market analysis, most previous studies have still focused on specific sectors or groups of issuers. Several studies applied K-Means to cluster banking companies based on financial ratios (Iriany et al., 2023; Puspaningsih et al., 2024), companies in the consumer goods sector (Dzuba & Krylov, 2021; Wisna et al., 2023), stocks in the consumer non-cyclicals sector (Ardhani et al., 2025), issuers in the agricultural product industry (Surya Febrian & Mutasowifin, 2025), as well as companies included in certain indices such as LQ45 (Hasnida & Kusumawati, 2022). Other studies have focused more on stock transaction characteristics or market behavior rather than the overall fundamental conditions of companies (Boloș et al., 2025). Thus, research examining the segmentation of financial performance of companies listed on the Indonesia Stock Exchange across sectors with a broad sample coverage remains relatively limited. In addition, most previous studies used a limited number of samples or focused only on certain financial dimensions, so they have not been able to comprehensively describe the heterogeneity of issuers' financial conditions. This research gap indicates the need for clustering analysis that covers companies from various industry sectors by integrating liquidity, profitability, and solvency indicators to produce a more representative classification of financial performance for the Indonesian capital market.

The principal contribution of this study lies in the development of an empirical taxonomy of the financial performance of IDX-listed companies, which may serve as a reference for investment decisions and capital market supervision. Through the analysis of 650 valid listed companies following the data cleaning process, this study furnishes a comprehensive overview of the financial landscape of public companies in Indonesia.

METHODS

Types and Sources of Data

This study utilizes secondary data sourced from the Indonesian capital market financial database, encompassing the financial information of 967 listed companies on the Indonesia Stock Exchange (IDX) for the 2024 financial year.

SJP Sosioedukasi
JURNAL ILMIAH ILMU PENDIDIKAN DAN SOSIAL
ISSN: 2086-6135 E-ISSN: 2541-612x
Volume 15 No. 2. June 2026 https://ejournal.unibabwi.ac.id/index.php/sosioedukasi/index

The financial variables collated include Total Current Assets, Total Current Liabilities, Total Assets, Retained Earnings, EBIT, Total Liabilities, and Total Equity. From these foundational variables, four financial ratios were extrapolated to serve as clustering attributes.

Research Variables

The research variables used as clustering attributes are:

Table 1. Operational Definitions of Research Variables

Variables	Definition	Formula	Source
Current Ratio	The company's ability to meet its short-term obligations	Current Assets / Current Liabilities	Brigham (2004)
EBIT Margin	Operational efficiency relative to revenue	EBIT / Total Revenue	Majka (2024)
Leverage Ratio	The proportion of assets financed by debt	Total Liabilities / Total Assets	Altman (1968)
Debt-to-Equity Ratio	The debt-to-equity ratio	Total Debt / Total Equity	Damodaran (2012)

Sources: Data Analysis

These four financial ratios were selected because they represent the main dimensions of a company's financial condition, namely liquidity (Current Ratio), operational profitability (EBIT Margin), asset funding structure (Leverage Ratio), and financial risk (Debt-to-Equity Ratio). The combination of these four variables provides a comprehensive picture of a company's financial characteristics, making them relevant for use as attributes in the clustering process (Altman, 1968; Damodaran, 2012; Brigham & Houston, 2019).

Stages of Analysis

The analysis was carried out using RStudio software (R version 4.3.x) in the following stages:

Preprocessing Data

Data preprocessing involves several steps. First, a check for missing values in each variable is carried out. Observations with missing values in one or more variables are excluded from the analysis. After eliminating missing values, 669 complete observations were obtained. Second, outliers were detected and removed using the Z-score method, whereby observations with a Z-score greater than 3 on any variable were eliminated. This process yielded 650 final observations used in the clustering analysis. A threshold of |Z-score| > 3 was used to identify extreme values that could potentially distort the Euclidean distance calculation in the K-Means algorithm. Outlier removal was carried out to improve cluster homogeneity, although this step could potentially reduce the representation of companies with very extreme financial characteristics. Thirdly, data standardization was performed using the Z-score standardization method (mean=0, standard deviation=1) to ensure all variables were on the same scale, so that no single variable dominated the distance calculations (Tan et al., 2016).

Determining the Optimal Number of Clusters

The optimal number of clusters (k) is determined through two complementary methods. The Elbow method examines the relationship between the number of clusters and the Total Within-Cluster Sum of Squares (WSS), whereby the inflection point on the WSS curve, at which the incorporation of further clusters no longer yields a substantive reduction in WSS, is construed as indicative of the optimal number of clusters. The Silhouette method gauges how aptly each observation is positioned within its designated cluster relative to other proximate clusters, with Silhouette values ranging from -1 to 1, wherein higher values are denotive of superior cluster quality. Both methods were applied for values of k ranging from 2 to 8 (Rousseeuw, 1987).

K-Means Clustering

K-Means Clustering is an unsupervised machine learning algorithm that partitions n observations into k clusters, whereby each observation is assigned to the cluster whose centroid is nearest in proximity (MacQueen, 1967). This algorithm operates iteratively through two principal steps: (1) Assignment Step: each observation is allocated to the cluster whose centroid is nearest in proximity based on Euclidean distance; (2) Update Step: the centroid of each

Sosioedukasi
JURNAL ILMIAH ILMU PENDIDIKAN DAN SOSIAL
ISSN: 2086-6135 E-ISSN: 2541-612x
Volume 15 No. 2. June 2026 https://ejournal.unibabwi.ac.id/index.php/sosioedukasi/index

cluster is recalibrated as the mean of all observations assigned to that cluster. This process is reiterated until the cluster assignments remain unchanged or converge to a stable configuration. In this study, the parameter nstart=25 was employed to execute the algorithm from 25 distinct initializations, thereby augmenting the probability of attaining the global optimal solution. The objective function minimized by K-Means is:

$$J = \Sigma_{j=1}^{k} \Sigma_{i \in C_j} ||x_i - \mu_j||^2$$

where k is the number of clusters, C_j is the set of observations in cluster j, x_i is the i-th observation, and μ_j is the centroid of cluster j.

Cluster Validation

The quality of the clusters is validated using two measures. First, the Silhouette Score (s(i)) for each observation is calculated as:

$$s(i) = [b(i) - a(i)] / \max\{a(i), b(i)\}$$

where a(i) denotes the average distance of observation i from all other observations within the same cluster, and b(i) represents the average minimum distance of observation i from observations in the nearest adjacent cluster (Rousseeuw, 1987). Subsequently, a one-way ANOVA test was conducted to evaluate whether the mean of each variable differed significantly between clusters ($p < 0.05$). In this study, ANOVA was used as a descriptive validation tool to assess the separation and characteristics of the clusters formed, not to test causal relationships between variables.

RESULTS

Stages of Analysis

Following preprocessing and the excision of outliers, the final dataset comprises 650 issuer observations. Table 2 presents the descriptive statistics for the four variables utilized in the clustering analysis.

Table 2. Descriptive Statistics of Research Variables (n=650)

Statistics	Current Ratio	EBIT Margin	Leverage Ratio	Debt-to-Equity
Minimum	0,042	-3,445	0,019	0,000
Q1 (25%)	1,155	0,019	0,224	0,072
Median	1,747	0,082	0,394	0,252
Mean	3,032	0,059	0,401	0,576
Q3 (75%)	3,301	0,175	0,547	0,668
Maximum	40,191	3,140	0,969	8,219
Standard Deviation	3,928	0,373	0,218	0,903

Sources: Data Analysis

Table 2 shows that the Current Ratio has a highly positively skewed distribution (mean = 3.032 > median = 1.747), indicating that the majority of listed companies have moderate liquidity ratios, although there are some with very high liquidity. The EBIT Margin has a positive median value (median = 0.082 or 8.2%), indicating that the majority of issuers are still able to generate operating profit. The Leverage Ratio, with a median of 0.394, indicates that, on average, IDX issuers finance approximately 39.4% of their assets through debt.

Determining the Optimal Number of Clusters

The optimal number of clusters was determined through the Elbow and Silhouette methods. Figure 1 illustrates the curve of the Total Within-Cluster Sum of Squares (WSS) plotted against the number of clusters using the Elbow method.

JURNAL ILMIAH ILMU PENDIDIKAN DAN SOSIAL
ISSN: 2086-6135 E-ISSN: 2541-612x
Volume 15 No. 2. June 2026 https://ejournal.unibabwi.ac.id/index.php/sosioedukasi/index

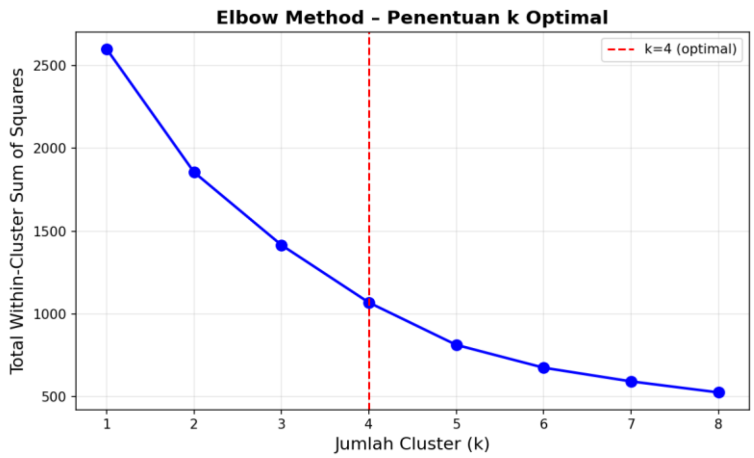

Figure 1. Elbow Method – Determining the Optimal Number of Clusters
Sources: Data Analysis

Figure 1 demonstrates that the most pronounced reduction in WSS transpires between k=1 and k=4. Beyond k=4, the curve begins to plateau gradually, indicating that the incorporation of more than four clusters does not yield a meaningful improvement in cluster quality. This finding is corroborated by the results of the Silhouette Analysis presented in Figure 2.

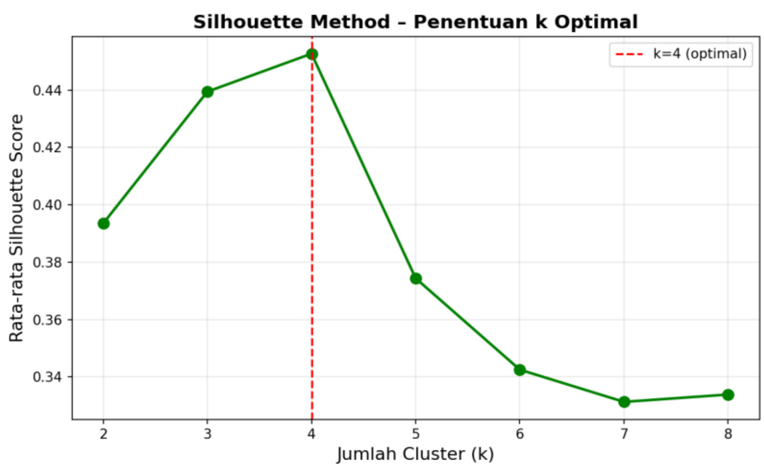

Figure 2. Silhouette Method – Determining the Optimal Number of Clusters
Sources: Data Analysis

Figure 2 shows that the Silhouette Score reached its highest value at k = 4, amounting to 0.4527, and tended to decrease as the number of clusters increased. This result indicates that the formation of four clusters provides the best balance between within-cluster homogeneity (cohesion) and between-cluster differences (separation) compared to other cluster number alternatives (Kaufman and Rousseeuw, 2009), corroborating that k=4 represents the most apposite choice for this dataset. Based on the convergence of these two methods, it is concluded that k=4 constitutes the optimal number of clusters.

Sectorally, this pattern indicates heterogeneity in company characteristics arising from differences in industry conditions, such as profitability levels, capital structure, operational risk, and growth opportunities that are not uniform

JURNAL ILMIAH ILMU PENDIDIKAN DAN SOSIAL
ISSN: 2086-6135 E-ISSN: 2541-612x
Volume 15 No. 2. June 2026 https://ejournal.unibabwi.ac.id/index.php/sosioedukasi/index

across sectors. Therefore, the formation of four clusters is able to group companies with relatively similar characteristics while maintaining meaningful differences between industry groups. In addition, increasing the number of clusters beyond four tends to split groups of companies that still have similar characteristics, thereby reducing the quality of cluster separation.

Nevertheless, the Silhouette Score value of 0.4527 needs to be interpreted carefully because it indicates cluster structure quality in the moderate-to-weak category, not a very strong cluster structure. This value indicates that there are still a number of observations located near the boundaries between clusters, so that the characteristics of several companies are similar to more than one group. According to Rousseeuw (1987), a Silhouette Score above 0.5 generally reflects a fairly good cluster structure, while values between 0.25 and 0.50 indicate an identifiable cluster structure but one whose separation level is not yet fully strong. Thus, although the Elbow and Silhouette methods consistently show that k = 4 is the most optimal number of clusters, the clustering results still need to be interpreted while considering the overlap of characteristics between companies and between industry sectors.

This finding is in line with clustering research by Meliyana et al. (2025) on economic and business data which generally have a high level of heterogeneity, so that the boundaries between groups are often not clearly formed. Under such conditions, a moderate Silhouette Score value can still be considered adequate as long as the clusters formed have clear economic and sectoral interpretations.

Results of K-Means Clustering (k=4)

The K-Means algorithm with k=4 was run using 25 different initialisations (nstart=25) to ensure the stability of the solution. The final model yielded a Between-SS/Total-SS ratio of 0.568, meaning that 56.8% of the total data variability can be explained by the differences between clusters. Table 3 summarises the distribution and average profile of each cluster.

Table 3. Average Profile of Financial Variables by Cluster

Cluster	n (%)	Current Ratio	EBIT Margin	Leverage Ratio	Debt-to-Equity	Label
1	54 (8,3%)	13,144	0,048	0,087	0,039	Liquid-Conservative
2	454 (69,8%)	2,369	0,114	0,347	0,276	Moderate-Healthy
3	131 (20,2%)	1,158	0,048	0,724	1,869	High Leverage
4	11 (1,7%)	3,080	-2,045	0,297	0,195	Distress

Sources: Data Analysis

Figure 3 illustrates the results of clustering using two principal components (Principal Component Analysis), which together account for 69.4% of the total variability in the data.

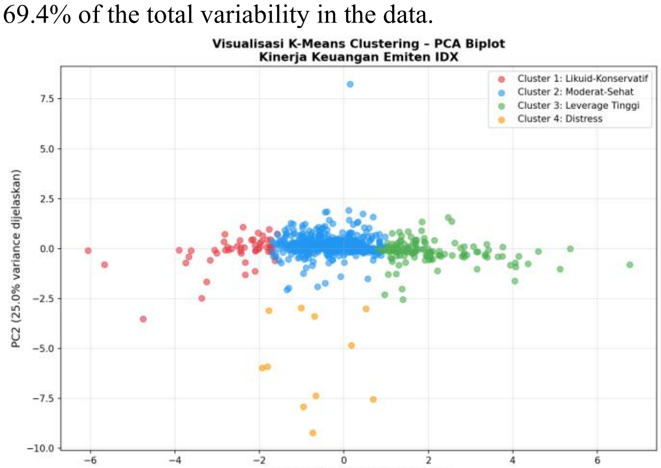

Figure 3. Visualisation of K-Means Clustering – PCA Biplot

Sosioedukasi
JURNAL ILMIAH ILMU PENDIDIKAN DAN SOSIAL
ISSN: 2086-6135 E-ISSN: 2541-612x
Volume 15 No. 2. June 2026 https://ejournal.unibabwi.ac.id/index.php/sosioedukasi/index

Sources: Data Analysis

Figure 3 shows a fairly clear separation between the four clusters, particularly between Cluster 1 (Liquid-Conservative, shown in red), located in the bottom right-hand corner; Cluster 3 (High Leverage, shown in green), clustered on the left-hand side; and Cluster 4 (Distress, shown in orange), which is isolated in a specific corner. Cluster 2 (Moderate-Healthy, blue) dominates the central part of the feature space, reflecting its characteristics as the majority group with a relatively diverse yet balanced financial profile.

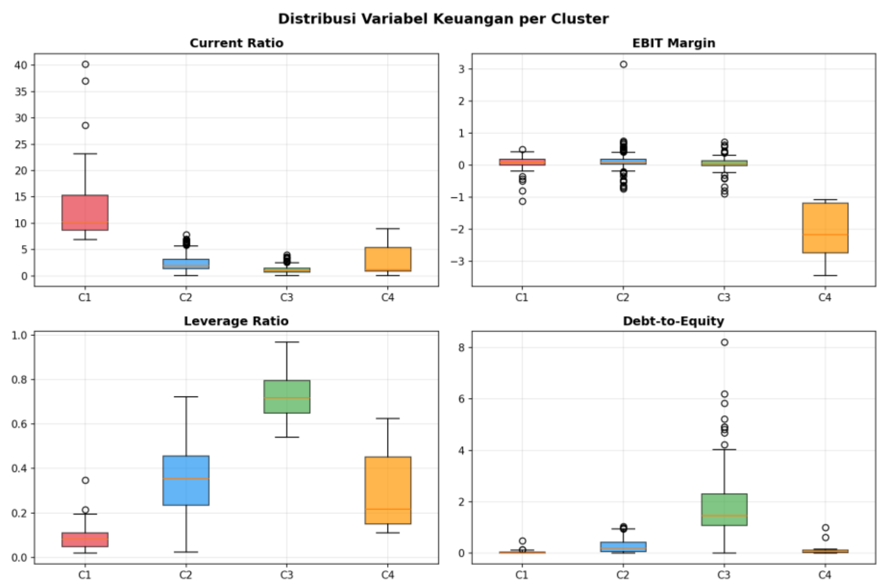

Figure 4. Distribution of Financial Variables by Cluster
Sources: Data Analysis

Figure 4 provides a more detailed overview of the distribution for each variable within each cluster. It can be seen that Cluster 1 exhibits a very wide range in the Current Ratio, whilst Cluster 3 shows a high concentration of the Leverage Ratio and Debt-to-Equity Ratio. Cluster 4 has a very low and variable EBIT Margin, confirming its unstable financial condition.

Figure 5. Average Financial Profile by Cluster (Normalised)
Sources: Data Analysis

Sosioedukasi
JURNAL ILMIAH ILMU PENDIDIKAN DAN SOSIAL
ISSN: 2086-6135 E-ISSN: 2541-612x
Volume 15 No. 2. June 2026 https://ejournal.unibabwi.ac.id/index.php/sosioedukasi/index

Interpretation and Profile of Each Cluster

Cluster 1 – Liquid-Conservative (n=54, 8,3%)

This cluster is characterized by a very high average current ratio (13.14), far exceeding the general benchmark of 2.0, which is considered healthy by industry standards (Brigham, 2004). The Leverage Ratio (0.087) and Debt-to-Equity Ratio (0.039) are very low, indicating that companies in this cluster rely very little on debt financing. The moderate EBIT Margin (0.048 or 4.8%) suggests that these companies have sufficient operational profitability, although not the highest.

This profile is typical of companies that hold substantial cash and current assets but are cautious about using leverage for expansion. Examples of listed companies in this cluster include MBSS.JK (Mitrabahtera Segara Sejati), ESSA.JK (ESSA Industries Indonesia), and DMAS.JK (Puradelta Lestari). From an investment perspective, companies in this cluster offer a very low-risk profile but with potentially limited growth due to the underutilization of leverage (Damodaran, 2022). This aligns with the capital structure trade-off theory, where excessively low debt usage may indicate suboptimal utilization of the tax shield (Modigliani & Miller, 1963 in Landi et al., 2022).

Cluster 2 – Moderate-Healthy (n=454, 69,8%)

This cluster is the largest group, accounting for almost 70% of all listed companies in the sample. Its financial profile is balanced: a Current Ratio of 2.37 (above the 2.0 threshold), the highest EBIT Margin among the four clusters (0.114 or 11.4%), a moderate Leverage Ratio (0.347), and a controlled Debt-to-Equity ratio (0.276). These characteristics describe companies that operate efficiently with an optimal capital structure.

Representative listed companies in this cluster include ICBP.JK (Indofood CBP Sukses Makmur), HRUM.JK (Harum Energy), and BUVA.JK (Bukit Uluwatu Villa). This cluster represents the ‘backbone’ of the Indonesian economy in the capital market, encompassing companies from various sectors that have successfully maintained a balance between growth and financial risk management. From a financial theory perspective, the moderate leverage profile aligns with the predictions of the Pecking Order Theory (Myers & Majluf, 1984), which states that profitable companies tend to finance their operations from retained earnings before resorting to debt.

Cluster 3 – High Leverage (n=131, 20.2%)

Cluster 3 is characterized by the highest Leverage Ratio (0.724) and a very high Debt-to-Equity Ratio (1.869), meaning that, on average, nearly 72% of the assets of companies in this cluster are financed by debt. A low Current Ratio (1.158) and a moderate EBIT Margin (0.048) add to the complexity of this cluster’s risk profile.

Representative issuers include TBIG.JK (Tower Bersama Infrastructure), KRAS.JK (Krakatau Steel), and MFMI.JK (Multifiling Mitra Indonesia). High leverage in this cluster is not necessarily negative; for infrastructure companies such as TBIG.JK, high leverage is a common business strategy as infrastructure assets generate stable cash flows and can be used as collateral (McKinsey & Company, 2021). However, under macroeconomic pressures such as interest rate hikes, companies with high leverage are more vulnerable to rising interest costs and refinancing risks (Altman, 1968). This situation is relevant to the trend of Bank Indonesia’s interest rate hikes throughout 2023–2024 (Indonesia, 2024).

Cluster 4 – Distress (n=11, 1,7%)

Cluster 4 is the smallest but most critical group, characterized by a highly negative average EBIT margin (-2,045 or -204.5%). This indicates that companies in this cluster are consistently incurring significant operating losses. Although the Current Ratio (3.080) and Leverage Ratio (0.297) appear moderate, the context of a highly negative EBIT Margin suggests fundamental issues within the business model or industry conditions that threaten the sustainability of the business.

Issuers in this cluster include NETV.JK (MDTV Media Technologies), MEDS.JK (Hetzer Medical Indonesia), and FLMC.JK (Falmaco Nonwoven Industri). This situation aligns with the characteristics identified by Altman (1968) in his model as indicators of potential bankruptcy, particularly through the variable of operational profitability. For investors, this cluster represents the ‘highly speculative’ or even ‘in default’ categories in credit rating terminology (Moody’s, 2023). Regulators and stock exchanges need to pay particular attention to issuers in this cluster to ensure transparency of information to the public.

SJP Sosioedukasi
JURNAL ILMIAH ILMU PENDIDIKAN DAN SOSIAL
ISSN: 2086-6135 E-ISSN: 2541-612x
Volume 15 No. 2. June 2026 https://ejournal.unibabwi.ac.id/index.php/sosioedukasi/index

Statistical Validation

A one-way ANOVA was conducted to affirm that the differences in means between clusters were statistically significant for each variable. Table 4 presents a summary of the ANOVA results.

Table 4. Results of the ANOVA test for differences between clusters

Variables	F-statistic	p-value	Significance
Current Ratio	347,38	< 0,001	***
EBIT Margin	269,39	< 0,001	***
Leverage Ratio	416,35	< 0,001	***
Debt-to-Equity Ratio	236,11	< 0,001	***

Note: *** = significant at p < 0.001

Sources: Data Analysis

Table 4 shows that all the financial variables used indicate statistically highly significant differences between clusters ($F > 236$, $p < 0.001$), confirming that the four clusters formed are a meaningful representation of different financial performance segments among IDX-listed companies.

DISCUSSION AND IMPLICATIONS

The dominance of the Moderate-Healthy Cluster (69.8%) indicates that most IDX-listed companies exhibit relatively balanced financial characteristics. This finding suggests that a substantial proportion of listed firms maintain financial conditions that may be viewed positively by market participants. This finding is consistent with a report by the McKinsey Global Institute (McKinsey & Company, 2021) which identifies Indonesia as one of the emerging markets with increasingly strong corporate fundamentals.

Secondly, the presence of the High Leverage Cluster (20.2%) warrants particular attention in the context of monetary policy. Given that Bank Indonesia has implemented a series of interest rate adjustments during the 2023–2024 period in response to global inflationary pressures (Indonesia, 2024), highly leveraged companies may face a significant increase in interest costs. From a theoretical perspective, companies with higher leverage tend to be more sensitive to changes in interest rates because debt servicing costs may increase when financing conditions become less favorable.

Thirdly, the identification of the Distress Cluster (1.7%, n=11) has important implications for the stock exchange and regulators. The identification of a Distress Cluster (1.7%) suggests that clustering techniques may provide useful information for monitoring financially vulnerable issuers. Therefore, such approaches could be explored as a complementary tool within existing surveillance and risk-monitoring framework (Otoritas Jasa Keuangan, 2023). The implementation of similar models has proven effective on developed stock exchanges such as the NYSE and the LSE (Duffie & Singleton, 2012).

Fourthly, the Liquid-Conservative Cluster (8.3%) is noteworthy as it indicates potential under-investment. Free cash flow theory (Jensen, 1986) states that uninvested cash surpluses can reduce long-term shareholder value. Based on free cash flow theory (Jensen, 1986), the accumulation of substantial liquid resources may indicate the possibility of underutilized funds. Consequently, further evaluation of capital allocation strategies may be relevant for firms within this cluster.

The formation of several clusters with different financial characteristics indicates that companies listed on the Indonesia Stock Exchange have a fairly high level of financial heterogeneity. This finding is in line with recent research showing that company capital structure is not uniform, but is influenced by differences in operational conditions, growth opportunities, business risk, and access to funding sources (Ai et al., 2020; Morais et al., 2022). From the trade-off theory perspective, variations in characteristics between clusters can reflect differences in company strategies in balancing the benefits of debt use with the potential costs of financial distress (da Rosa München, 2022; Ugur et al., 2022). On the other hand, the presence of companies with relatively high leverage levels also supports the view that funding decisions are often influenced by financing needs and limited internal funds, resulting in different capital structures between companies (Hardianti, 2023; Mursalini et al., 2025). In addition, the emergence of a group

SJP Sosioedukasi
JURNAL ILMIAH ILMU PENDIDIKAN DAN SOSIAL
ISSN: 2086-6135 E-ISSN: 2541-612x
Volume 15 No. 2. June 2026 https://ejournal.unibabwi.ac.id/index.php/sosioedukasi/index

of companies showing characteristics leading to financial pressure indicates that financial distress risk is a multidimensional phenomenon influenced by a combination of profitability, liquidity, corporate governance, and capital structure factors (Hu & Zheng, 2015). Therefore, the use of a clustering approach in this study provides added value because it is able to identify financial characteristic patterns simultaneously, resulting in a more comprehensive understanding compared to analysis that focuses on only one financial indicator separately (Chhillar & Lellapalli, 2022).

Overall, the results of this study show that the financial condition of companies listed on the Indonesia Stock Exchange cannot be represented by one dominant characteristic. Although most companies are in the group with relatively healthy financial conditions, the presence of high leverage, distress, and liquid-conservative clusters indicates variations in financial strategies and different risk levels between companies. This finding shows that the evaluation of public company conditions needs to consider a combination of several financial indicators simultaneously, because each company may face different challenges and opportunities even though they operate in the same market environment. Thus, a cluster-based approach not only helps identify groups of companies with similar characteristics, but also provides a more comprehensive picture of the financial dynamics of public companies in Indonesia.

CONCLUSION

This study has successfully applied the K-Means Clustering method to group 650 companies listed on the Indonesia Stock Exchange based on Current Ratio, EBIT Margin, Leverage Ratio, and Debt-to-Equity Ratio. The analysis results show that the optimal number of clusters is four groups with significantly different financial characteristics. Validation using the ANOVA test confirmed that the differences between clusters are significant, so that the resulting grouping is able to represent the heterogeneity of the financial conditions of companies listed on the Indonesia Stock Exchange. Practically, the results of this grouping can be used as a basis for identifying the risk profile and financial condition of companies, thus being useful for investors, financial analysts, and regulators in supporting more effective decision making and monitoring processes.

The limitations of this study lie in the use of single period observation data and a still limited number of financial variables. Therefore, future research is recommended to use longitudinal data, add more diverse financial and market variables, and compare results with alternative clustering methods to obtain a more comprehensive understanding of company characteristics in the Indonesian capital market.

REFERENCES

Aditiya, F. R., & Sulistiyowati, N. (2026). K-Means Clustering of Indonesian Banking Stocks Using Financial Ratios. *Jurnal Techno Nusa Mandiri*, *23*(1), 23–30. https://doi.org/10.33480/techno.v23i1.8334

Ai, H., Frank, M. Z., & Sanati, A. (2020). The trade-off theory of corporate capital structure. In A. Banerjee (Ed.), *Oxford Research Encyclopedia of Economics and Finance (New York*. Oxford research encyclopedia of economics and finance. https://doi.org/10.1093/acrefore/9780190625979.013.602

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, *23*(4), 589–609. https://doi.org/10.2307/2978933

Ardhani, N. T., Notodiputro, K. A., & Oktarina, S. D. (2025). Winsorization untuk Penanganan Pencilan dalam Penggerombolan Saham Sektor Consumer Non-Cyclical dengan K-Means dan K-Medoids. *Indonesian Journal of Statistics & Its Applications*, *9*(1), 46. https://doi.org/10.29244/ijsa.v9i1p46-60

Boloș, M.-I., Rusu, Ștefan, Leordeanu, M., Sabău-Popa, C. D., Perțicaș, D. C., & Crișan, M.-I. (2025). K-Means Clustering for Portfolio Optimization: Symmetry in Risk–Return Tradeoff, Liquidity, Profitability, and Solvency. *Symmetry*, *17*(6), 847. https://doi.org/10.3390/sym17060847

Brigham, E. F. (2004). Fundamentals of Financial Management. In *Thomson South-Western*.

Camm, J. D., Cochran, J. J., Fry, M. J., & Ohlmann, J. W. (2020). *Business Analytics*. Cengage Learning.

Chhillar, P., & Lellapalli, R. V. (2022). Role of earnings management and capital structure in signalling early stage

of financial distress: a firm life cycle perspective. *Cogent Economics & Finance*, *10*(1), 2106634. https://doi.org/10.1080/23322039.2022.2106634
da Rosa München, D. (2022). The effect of financial distress on capital structure: The case of Brazilian banks. *The Quarterly Review of Economics and Finance*, *86*, 296–304. https://doi.org/10.1016/j.qref.2022.07.010
Damodaran, A. (2012). *Investment Valuation: Tools and Techniques for Determining the Value of Any Asset*. John Wiley & Sons.
Duffie, D., & Singleton, K. J. (2012). Credit Risk: Pricing, Measurement, and Management. In *Credit risk*. Princeton University Press.
Dzuba, S., & Krylov, D. (2021). Cluster analysis of financial strategies of companies. *Mathematics*, *9*(24), 3192. https://doi.org/10.3390/math9243192
Hardianti, F. (2023). The Influence of Good Corporate Governance Mechanisms on Financial Performance with Transparency as a Moderating Variable in Companies Listed on the Indonesian Indonesia Stock Exchange in 2020–2022. *TRANSEKONOMIKA: AKUNTANSI, BISNIS DAN KEUANGAN*, *3*(5), 938–957. https://doi.org/10.55047/transekonomika.v3i5.565
Hu, D., & Zheng, H. (2015). Does ownership structure affect the degree of corporate financial distress in China? *Journal of Accounting in Emerging Economies*, *5*(1), 35–50. https://doi.org/10.1108/JAEE-09-2011-0037
Iriany, A., Putri, H. R. A., & Tua, H. M. (2023). K-means Nonhierarchical Cluster and Dbscan Outlier Detection In the Grouping of Stock Issuers. *Proof*, *3*, 21–28. https://doi.org/10.37394/232020.2023.3.4
Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *The American Economic Review*, *76*(2), 323–329. https://doi.org/10.2139/ssrn.99580
Kaufman, L., & Rousseeuw, P. J. (2009). *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley & Sons. https://doi.org/10.1002/9780470316801
Landi, G. C., Iandolo, F., Renzi, A., & Rey, A. (2022). Embedding sustainability in risk management: The impact of environmental, social, and governance ratings on corporate financial risk. *Corporate Social Responsibility and Environmental Management*, *29*(4), 1096–1107. https://doi.org/10.1002/csr.2256
MacQueen, J. B. (1967). Some Methods for Classification and Analysis of Multivariate Observations. *Berkeley, University of California Press*, *1*, 281–297.
Majka, M. (2024). EBIT: A Key Financial Performance Indicator. *ResearchGate, November*.
McKinsey & Company. (2021). *Future of Asia*. McKinsey Global Institute.
Meliyana, S. M., Dunggio, A. S. A., Muhammad, S., & Rahman, A. (2025). Comparative Performance Analysis of Two Clustering Methods for Grouping Indonesian Provinces Based on Forest Area Size. *Daengku: Journal of Humanities and Social Sciences Innovation*, *5*(4), 573–581. https://doi.org/10.35877/454RI.daengku4439
Morais, F., Serrasqueiro, Z., & Ramalho, J. J. S. (2022). Capital structure speed of adjustment heterogeneity across zero leverage and leveraged European firms. *Research in International Business and Finance*, *62*, 101682. https://doi.org/10.1016/j.ribaf.2022.101682
Mursalini, W. I., Sriyanti, E., & Putri, A. N. (2025). Factors That Affect Stock Returns in Food and Beverage Companies Listed on the Indonesian Stock Exchange. *CASHFLOW : Current Advanced Research On Sharia Finance And Economic Worldwide*, *4*(2), 109–118. https://doi.org/10.55047/cashflow.v4i2.1467
Myers, S. C., & Majluf, N. S. (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, *13*(2), 187–221. https://doi.org/10.1016/0304-405X(84)90023-0
Otoritas Jasa Keuangan. (2023). *Roadmap Pasar Modal Indonesia 2023–2027. Otoritas Jasa Keuangan (OJK)*.
Puspaningsih, E. S., Di Asih, I. M., & Tarno, T. (2024). K-Medoids Clustering dan Mean-Value at Risk untuk Optimasi Portofolio Saham Jakarta Islamic Index. *Indonesian Journal of Applied Statistics*, *6*(1), 85–95. https://doi.org/10.13057/ijas.v6i1.79231
Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*(C). https://doi.org/10.1016/0377-0427(87)90125-7

Sosioedukasi
JURNAL ILMIAH ILMU PENDIDIKAN DAN SOSIAL
ISSN: 2086-6135 E-ISSN: 2541-612x
Volume 15 No. 2. June 2026 https://ejournal.unibabwi.ac.id/index.php/sosioedukasi/index

Rusu, Ștefan, Boloș, M. I., & Leordeanu, M. (2023). K-means and agglomerative hierarchical clustering analysis of ESG scores, yearly variations, and stock returns: Insights from the energy sector in Europe and the United States. *Journal of Financial Studies*, *8*(Special-June_2023), 166–180.

Saadah, L., Hafizah, D., & Zalianti, P. M. (2025). Analysis of Factors That Influence Sticky Costs in Transportation Sub-Sector Companies. *International Journal of Accounting, Business, and Economic Policy*, *1*(1), 55–63. https://doi.org/10.66324/ijabep.v1i1.22

Surya Febrian, S., & Mutasowifin, A. (2025). Selection of agricultural industry stocks by application of K-means algorithm with Elbow method. *BIO Web of Conferences*, *171*, 4003. https://doi.org/10.1051/bioconf/202517104003

Tan, P.-N., Steinbach, M., & Kumar, V. (2016). *Introduction to Data Mining*. Pearson Education India.

Ugur, M., Solomon, E., & Zeynalov, A. (2022). Leverage, competition and financial distress hazard: Implications for capital structure in the presence of agency costs. *Economic Modelling*, *108*, 105740. https://doi.org/10.1016/j.econmod.2021.105740

Wisna, N., Fairus, L., & Kastaman, K. (2023). Analisis Rasio Profitabilitas Perusahaan Konstruksi Berdasarkan K-Means Clustering Menggunakan Metode Phyton. *Jurnal Ilmiah Edunomika*, *7*(2), 1–8. https://doi.org/10.29040/jie.v7i2.10062

Xu, J., Xu, K., Wang, Y., Shen, Q., & Li, R. (2024). A k-means algorithm for financial market risk forecasting. *ArXiv Preprint ArXiv:2405.13076*, 1–12. https://doi.org/10.48550/arXiv.2405.13076