Clustering and Financial Performance Analysis of Indonesian Coal Mining Industry Stock Prices

Achmad Naufal*, Muhammad Fadhil†, Wisudanto‡

* Departemen Manajemen Teknologi, Fakultas Sekolah Interdisiplin Manajemen dan Teknologi, Institut Teknologi Sepuluh Nopember, 6032231182@students.its.ac.id, 6032231222@students.its.ac.id
† Departemen Manajemen, Fakultas Ekonomi dan Bisnis, Universitas Airlangga, wisudanto@feb.unair.ac.id

Submitted: 27-10-2023, Reviewed: 07-11-2023, Accepted 25-11-2023
https://doi.org/10.47233/jteksis.v6i1.1081

Abstract
This study addresses the challenge of understanding the financial dynamics and market responses of Indonesian coal mining companies, particularly in the context of post-pandemic market fluctuations during 2022 and 2023. The primary aim was to analyze historical stock price data and financial metrics from coal companies listed for over a decade, utilizing data sourced from Yahoo Finance. The research employed a range of robust clustering techniques, including K-means, Hierarchical, and Correlation Clustering, to systematically evaluate stock price movements and financial trends. The results, derived from detailed graphical analyses, revealed distinct patterns in stock market behavior and financial performance among the companies. The study identified three major clusters representing different responses to market conditions, as visually illustrated in the t-SNE visualizations and financial performance charts. These clusters were further examined through financial metrics such as Return on Assets (ROA), Net Profit Margin (NPM), and Earnings Per Share (EPS), uncovering unique financial characteristics among companies like PTIS, IATA, and AIM5. Key findings include the superior performance of IATA in favorable market conditions, the challenges faced by AIM5 during market downturns, and the consistent performance of PTIS across varying scenarios. These insights, grounded in comprehensive data analysis and visual representation through informative charts, provide a nuanced understanding of the coal industry's financial landscape, highlighting significant divergences in company performances. Overall, this research not only sheds light on the financial dynamics of the Indonesian coal mining industry in a post-pandemic era but also introduces an innovative analytical methodology, enhancing financial analytics within the sector.

Keywords: Coal mining industry, Data clustering, Financial metrics, Financial trends, Stock price

INTRODUCTION
The coal mining industry of Indonesia has long been acknowledged as a vital component of the nation's economic structure. It has consistently made substantial contributions to Indonesia's Gross Domestic Product (GDP) and has been a significant employment creator across various associated sectors [1]. This pivotal industry extends its influence beyond mere economic metrics, as it significantly impacts the Indonesian stock market. Here, the fiscal health and stock valuations of coal mining firms often mirror the sector's general vitality and current tendencies [2]. Historically, stock prices within this sector have shown a strong correlation with global market movements, state policies, and environmental factors, which have all played crucial roles in steering investment approaches and valuations. The recent COVID-19 pandemic has further underscored the vulnerability and interconnectedness of global markets, with sectors like the coal mining industry feeling the ripple effects of the crisis. Siagian & Cahyono (2021) delved into the recovery strategies of small and medium-sized enterprises (UMKM) in the creative economic sector during the pandemic, highlighting the broader challenges and strategic shifts businesses had to undertake in these unprecedented times [3]. Moreover, a study by Shofa Rijalul Haq et al. (2022) revealed a tangible transformation in coal demand, production, and price in Indonesia during the pandemic, emphasizing the significant impact of COVID-19 on the coal mining industry [4].

Several research endeavours in the past have probed into the intricate dynamics of stock price evaluation in this domain, trying to untangle the complex web of financial and economic factors influencing stock returns and company valuations [5][6]. Nevertheless, a clear void is evident in the literature, primarily concerning a detailed study that emphasizes clustering these stock prices. Specifically, there is a need for a study that assesses different time frames and then proceeds with an extensive financial performance examination. The present research aspires to fill this void, aiming to discover underlying patterns via clustering. This approach is hoped to yield a structured and enlightening perspective on market operations, a viewpoint that will be invaluable for investors, policymakers, and other stakeholders in making informed choices in the intricate world of investments.
The analysis of stock prices, especially in an industry as dynamic as coal mining, is laden with challenges. Influences such as global market dynamics, government regulations, and environmental conditions can have profound effects on stock valuations [7]. Grouping or clustering these prices can reveal inherent trends, providing a coherent picture of the market's operations, and guiding investors and policymakers in their decision-making [8]. A detailed understanding of financial health within these groups can further shed light on the companies' financial stability, essential for strategic planning and risk mitigation [9]. Past investigations, like those conducted by Fianty and Gunawan [10], have emphasized the need for solid analytical techniques to unravel the complex interplay between financial ratios and stock prices in this sector. Similarly, the study by Hidayat and Prasoj [11] highlighted the importance of considering numerous internal and external factors in stock price evaluation. This research intends to journey through these intricacies, employing a robust approach to clustering and financial scrutiny, with the hope of providing a perspective that is both theoretically rich and practically insightful.

The primary objective of this study revolves around clustering stock prices from firms in the Indonesian coal mining sector and subsequently assessing the financial health within these clusters. A secondary goal is to explore clustering techniques, comparing different methods to identify the most suitable for this particular scenario [12]. Additionally, the research aims to provide insights on practical applications in areas like investment strategy development, policy-making, and future academic investigations. Factors like credit risk, liquidity risk, and operational risk, known to have significant influences on a company's financial stability, will be critically examined [13]. Furthermore, the influence of Corporate Social Responsibility (CSR) on financial metrics such as Return On Assets (ROA), Return On Equity (ROE), and Price Book Value (PBV) will be assessed to offer a comprehensive understanding of factors guiding stock prices and financial health [14].

Focused primarily on firms in the Indonesian coal mining sector, this research offers an in-depth analysis of this specific domain. The evaluated data, restricted to a specific timeframe, will involve daily stock prices to ensure relevance and detail [15]. With its geographical scope limited to Indonesia, the insights generated will be particularly valuable for both local and international stakeholders active in this region. The study will also operate within methodological confines, employing and contrasting specific clustering and analytical techniques relevant to this context [16]. By examining the internal and external factors that influence stock prices and financial health in the coal mining sector, the research aims to uncover detailed patterns that can guide diverse stakeholders in their decision-making processes.

The significance of this research spans various stakeholders. Investors can benefit from a deeper understanding of stock price dynamics and financial performance within the coal mining sector. Identifying clusters and evaluating their financial viability can aid in devising well-informed and risk-mitigated investment strategies [17]. Policymakers can also leverage the insights to frame policies aligned with the economic realities of the coal sector [18]. From an academic perspective, this study hopes to add a fresh viewpoint on stock price analysis in the specific context of the Indonesian coal industry [19]. By merging financial scrutiny with behavioural finance, the research seeks to furnish a comprehensive perspective, potentially guiding the formulation of robust investment strategies and policies attuned to both financial and psychological factors.

In pursuit of these goals, this research aims to address several crucial questions. These queries delve deep into the world of financial analysis and investment strategy within the Indonesian coal mining industry. By tackling these questions, the research aspires to bridge theoretical and practical gaps, ensuring the derived insights are both academically rigorous and practically invaluable. This research delves deep into the financial intricacies of Indonesia's coal mining industry, posing several core questions. Primarily, it examines the effectiveness of various clustering methods in analyzing annual groupings of daily stock prices, aiming to discern patterns vital for forecasting and informed investment [6]. It further investigates variations in financial health within and between these clusters, providing insights into the fiscal robustness of the involved companies [13]. On a methodological front, the study probes the impact of different clustering techniques on the derived financial insights [1]. Lastly, it contemplates the practical implications of these findings for stakeholders like investors and policymakers, while also considering how broader macroeconomic factors, such as inflation and exchange rates, intertwine with stock prices and company financial performances [22][23]. In essence, this research endeavors to bridge academic theory with real-world applications, offering valuable insights for the coal mining sector's financial landscape.
RESEARCH METHODS

2.1. Clustering and Time-Series Analysis

Clustering methods and time-series analysis have become pivotal in financial data analysis, particularly in dissecting the complex and often chaotic nature of stock prices within various sectors, including the mining industry. Clustering methods, such as hierarchical and non-hierarchical approaches, have been utilized to uncover underlying patterns and trends within financial data, providing a structured and insightful perspective into the market’s behavior [14][15]. Hierarchical clustering, for instance, involves creating a tree of clusters, which is beneficial for understanding hierarchical relationships within the data, while non-hierarchical methods, like k-means clustering, partition data into predefined clusters, often providing a more generalized view of data patterns [15].

On the other hand, time-series analysis, which encompasses methodologies like autoregressive models, moving averages, and hybrid models, focuses on analyzing time-ordered data points to decipher trends, patterns, and potential future movements in stock prices [16]. For instance, autoregressive models leverage previous data points to predict future values, moving averages smooth out price data to create a single flowing trend, and hybrid models might combine various predictive models to enhance forecasting accuracy [16]. Both clustering and time-series analysis have been instrumental in providing meaningful insights from financial data, especially in predicting stock price movements and formulating investment strategies, by identifying patterns and trends that might not be immediately apparent in raw, unstructured data [15][17].

Moreover, the efficacy, challenges, and limitations of these methodologies have been explored in various contexts, providing a foundational understanding of the methodological considerations that need to be taken into account in this research. For instance, while clustering can reveal inherent groupings within the data, the choice of clustering method, the number of clusters, and the distance measure can significantly impact the results and their interpretability [15]. Similarly, while time-series analysis can provide valuable insights into potential future stock price movements, the accuracy of these predictions can be influenced by factors such as the chosen model, the availability and quality of data, and external market influences [16].

A. Hierarchical Clustering

Hierarchical clustering creates a tree of clusters, which is beneficial for understanding nested relationships within the data [18]. It does not require pre-specification of the number of clusters and provides a multilevel hierarchy that is useful for deriving various granularities of insights from the data. It encompasses agglomerative (bottom-up) and divisive (top-down) approaches. Agglomerative starts with each data point as a single cluster and merges them based on similarity, while divisive starts with one cluster of all data points, progressively dividing them [18]. Various distance metrics, such as Euclidean and Manhattan distances, are employed to calculate the dissimilarity between data points, influencing the formation of clusters [18]. While hierarchical clustering is intuitive and can provide detailed dendrograms for visual analysis, it can be computationally intensive and sensitive to outliers.

B. Non-Hierarchical Clustering (e.g., K-Means)

K-Means clustering partitions data into a predefined number of clusters, minimizing within-cluster variance [19]. It is widely used due to its simplicity and efficiency in handling large datasets. The algorithm iteratively assigns data points to clusters and recalculates cluster centroids until the within-cluster variance is minimized or a termination condition is met [19]. Various methods, such as the elbow method, are used to determine the optimal number of clusters, balancing the trade-off between the number of clusters and within-cluster variance. K-Means is computationally efficient and simple to implement but is sensitive to the initial placement of centroids and the choice of ‘K’. Hasanah & Purnomo (2022) utilized the K-Means clustering algorithm for grouping books in a library case study, demonstrating its applicability in diverse fields [20]. Similarly, Satriatama et al. (2023) employed K-Means clustering to analyze patient data for diabetes, emphasizing its utility in healthcare for identifying patterns and patient characteristics [21].

3.2. Financial Performance Analysis

The financial performance analysis within the identified clusters will be conducted using a plethora of financial metrics and ratios, which will be calculated using data sourced from Yahoo Finance, ensuring consistency and reliability in the data used for both clustering and financial analysis. Metrics such as Return on Assets (ROA), Net Profit Margin (NPM), and Earnings Per Share (EPS) will be pivotal in evaluating the financial health and sustainability of the companies within each cluster. These metrics are widely recognized and utilized in financial analysis, providing insights into profitability, efficiency, and shareholder value respectively [22]. In a study by Masrizal et al. (2020), the importance of financial ratios such as return on equity and debt to equity ratio in influencing stock prices was highlighted, further
emphasizing the significance of these metrics in financial analysis [23]. Analytical techniques, notably ratio analysis, will be employed to scrutinize the financial health and sustainability of the companies within each cluster. Ratio analysis, which involves comparing different financial metrics and ratios, provides a comprehensive view of a company’s financial performance and position [24]. Purmiyati & Handoyo (2022) also emphasized the role of technical efficiency analysis in understanding the health of credit cooperatives, showcasing the depth and breadth of financial analysis techniques [25]. Trend analysis will also be utilized, exploring the progression of various financial metrics over time, providing insights into the stability, growth, and potential future trajectories of the companies within each cluster [26].

Moreover, a comparative analysis will be conducted to explore the disparities and similarities in financial performance across the various clusters. This analysis will not only provide a multifaceted view of the financial dynamics within the sector but also enable the identification of patterns, anomalies, or standout performances within and between clusters. Utama & Wisudanto (2023) conducted a systematic literature review analyzing dividend policies during the Covid-19 pandemic, highlighting the importance of comparative analyses in understanding financial trends and patterns [27]. This comparative approach allows for a nuanced understanding of the financial landscape of the Indonesian coal mining industry, providing tangible insights and data-driven foundations upon which investment and policy decisions can be made.

RESULT AND DISCUSSIONS

In the pursuit of investigating the dynamics of listed coal industry companies in Indonesia, a selection process was undertaken. Out of the 26 coal companies that have been listed since 2014, a total of 21 companies were selected for the study. The selection criterion was primarily based on the stock market activity of these companies during the 2022-2023 period.

To delineate the structure within this dataset, a clustering approach using the K-means algorithm was employed. However, prior to executing the clustering process, it was imperative to determine the most suitable scaler for our dataset. This decision was based on the performance of various scalers as assessed by the Silhouette scores. The Silhouette method provided a measure of how similar an object is to its own cluster compared to other clusters. Higher Silhouette scores thus indicated better-defined clusters.

After determining the best scaler, the next step was to decide upon the optimum number of clusters (‘K’) for the K-means algorithm. Again, Silhouette scores played a pivotal role in this determination. Once the optimum ‘K’ was discerned, the K-means clustering process was carried out. Beside that other methods like Hierarchical Clustering and Correlation Clustering will be done.

To provide a more granular insight, this clustering was done separately for two distinct segments. The first segment pertained to the stock market activity during the increasing trend observed in 2022. The subsequent segment focused on the decreasing trend observed in 2023. This distinction facilitated a more nuanced understanding of the coal companies’ performance across different market dynamics.

3.1. Descriptive Analysis

In the ever-changing landscape of energy economics, the coal industry has exhibited intriguing shifts in pricing trends. Figure X below encapsulates these fluctuations over recent years. Particularly after the COVID-19 pandemic, the graph reveals two salient trends.

![Historical Coal Price From 2021](image)

In 2022, there was a notable rise in coal prices, likely influenced by the aftermath of the pandemic and its repercussions on global energy markets. However, as we entered 2023, the trend shifted, showing a decline in prices. In conclusion, the post-pandemic period experienced a dynamic change in coal prices, with a surge in 2022 followed by a decrease in 2023. This fluctuation underscores the evolving landscape of energy economics in recent years.

3.2. K-means Clustering

In the clustering process, the choice of scaler is crucial as it significantly influences the final clusters obtained. The scaler essentially standardizes or normalizes the data to ensure that all features contribute equally to the clustering result. Different scalers have distinct methods for achieving this
normalization, and their efficacy can vary depending on the dataset.

Figure 2. Silhouette score per each Scaling Methods

For our clustering analysis, selecting an appropriate scaler is essential. Upon evaluation using the Silhouette score, which assesses the quality of clusters produced, the Robust Scaling technique outperformed others by achieving the highest score. This method is especially adept at handling data with outliers. After scaling, the optimal number of clusters (denoted by 'K') for the K-means algorithm was determined to be 3.

Figure 3. K-means result on Increase and Decrease segment.

After the clustering was conducted using the Robust Scaler and with K=3, the results displayed clear distinctions between the two segments. In the increasing trend of 2022, the 1st cluster predominated with 19 companies, while the 2nd and 3rd clusters each only contained a single company. On the other hand, for the decreasing trend in 2023, both the 1st and 2nd clusters individually consisted of one company, and the 3rd cluster had a major composition of 19 companies. This shift in company groupings between the two years underscores the dynamic nature of the market and suggests that companies had varied responses to the market conditions in 2022 and 2023.

3.3. Hierarchical Clustering

To further delve into the intricacies of the data and substantiate the findings from the K-means clustering, another prominent clustering method, hierarchical clustering, was applied.

Figure 4. Dendrogram (Complete Linkage)
In our study, hierarchical clustering was executed using the complete linkage method combined with Euclidean distance on log returns. This methodological choice led to results that were notably consistent with those obtained from the K-means approach, where the optimal K was 3. Specifically, the hierarchical clustering also yielded three main clusters. Interestingly, for both the increase and decrease segments, two out of these three clusters contained just a single company each. This concordance between the two clustering methods underlines the robustness of the identified groupings and suggests that irrespective of the technique, certain companies exhibit unique characteristics that set them apart from the larger group in both market conditions.

3.4. Cluster Visualization

For our dataset, t-SNE was employed to visualize the clusters identified by both the K-means and hierarchical clustering methods across the two segments.

Using t-SNE, the clusters from the K-means and hierarchical clustering methods were vividly portrayed. The visualization highlighted the spatial arrangement of the 19 companies in the predominant cluster and clearly demarcated the two individual companies that were uniquely clustered in the increase segment.

Similarly, for the decrease segment in 2023, the t-SNE visualization painted a distinct picture. It effectively differentiated the major cluster comprising 19 companies from the other two individual companies. The two techniques' clustering results were coherently aligned in this visual representation.

The t-SNE visualizations not only confirmed the consistency between the K-means and hierarchical clustering outcomes but also provided an insightful spatial representation of the relationships between companies. The visualizations effectively showcased how certain companies exhibited patterns that made them distinct from the majority, corroborating the clustering results in both the increase and decrease segments.

3.5. Correlation Clustering

In our dataset, the correlation clustering revealed intriguing patterns. Mirroring findings from our previous clustering techniques, three prominent groups emerged, underlining the inherent structure within the data.
The outcome of the correlation clustering reaffirms the tripartite division of our data, as observed in earlier techniques. The consistent emergence of these three groupings, irrespective of the clustering methodology, underscores their significance and robustness. Specifically, while the majority of entities share common patterns, there are outliers that demand individual attention due to their unique behaviors or characteristics.

3.6. Financial Performance Analysis

From the cluster analysis conducted on 21 companies, there were discernible patterns in how these companies were grouped based on their performance segments. In the increasing segment, two companies, PTIS and IATA, stood out as they did not share a cluster with the other companies. Similarly, in the decreasing segment, PTIS and AIMS were distinct in that they did not cluster with the remainder of the companies. Notably, PTIS appears to be unique in that it demonstrates a different pattern in both the increase and decrease segments compared to the rest of the companies in the analysis. This suggests that PTIS's performance or characteristics might differ significantly from its counterparts in the market.

Next, a comparison will be made between PTIS, IATA, AIMS, and the remaining companies using averages across various financial metrics.

Considering the patterns across the metrics, IATA consistently positions itself in the upper echelons, especially in the increasing segment, distinguishing itself as the top performer among the three when metrics trend positively. Conversely, in the decreasing segment, AIMS notably underperforms, manifesting the steepest declines and suggesting challenges that might be more pronounced than its peers. Amidst this, PTIS exhibits a unique consistency, often aligning closer to median values, which underscores its relatively stable financial performance in contrast to the more polarized performances of IATA and AIMS.

The boxplots emphasize distinct financial characteristics of PTIS, AIMS, and IATA when juxtaposed against the aggregated companies. In terms of Diluted EPS, all three selected companies exhibit values that are within the lower to mid-range of the aggregated group, indicating moderate earning potential. For ROA, while IATA and PTIS have positive values, AIMS falls into the negative, suggesting potential operational or management challenges. Notably, IATA's ROA is closer to the median of the aggregated group, whereas PTIS leans towards the lower quartile. AIMS's negative ROA stands out as a deviation from the aggregated group's positive range. Regarding Profit Margin, IATA showcases a value closer to the upper quartile of the aggregated companies, indicating efficient operations or a solid pricing strategy. In contrast, AIMS's negative profit margin underlines potential challenges in maintaining profitability. PTIS's profit margin hovers around the median, suggesting average profitability compared to the aggregated group. In essence, the clustering of PTIS, AIMS, and IATA reveals a combination of financial performances that, while unique among themselves, also diverge in various ways from the aggregated companies.
CONCLUSION

In the comprehensive study of the Indonesian listed coal industry, 21 out of 26 companies were rigorously selected based on their stock market activity during 2022-2023. Employing the K-means clustering algorithm, the research first determined the most apt scaler through Silhouette scores, ultimately choosing the Robust Scaling technique. The optimal number of clusters was discerned to be three, leading to distinct groupings of companies based on the stock market trends of 2022 and 2023. Supporting the K-means results, Hierarchical Clustering and Correlation Clustering further reiterated the tripartite division, accentuating the robustness of the findings. The research charts, such as the t-SNE visualizations (Figure 5 and Figure 6) and financial performance boxplots (Figure 8), were instrumental in illustrating these insights. They vividly demonstrated the spatial relationships and distinct financial behaviors of companies, particularly PTIS, IATA, and AIMS, under different market conditions. Visualization techniques like t-SNE vividly showcased the spatial relationships between companies, underscoring the uniqueness of certain entities.

In the detailed financial performance analysis, three companies, PTIS, IATA, and AIMS, emerged with distinctive financial patterns. IATA was identified as a top performer, AIMS exhibited challenges, especially during the market's decrease, and PTIS showcased a unique consistency. Overall, the research provides a deep, multifaceted understanding of the dynamics of the Indonesian coal industry, highlighting varied responses of companies to market conditions and revealing significant financial performance divergences among key players.

For future work, extending this study over a longer period could provide a more comprehensive understanding of the companies' evolution, especially when compared with other energy sectors like renewables. Additionally, assessing the impact of global economic shifts, environmental policies, and technological advances on these companies could offer deeper insights into the industry's dynamics. Furthermore, applying predictive analytics to forecast future market trends based on current and historical data could be a valuable tool for strategic planning in the coal industry. This expanded approach would not only build upon the current findings but also contribute significantly to a more nuanced understanding of the Indonesian coal industry within the global energy landscape.

BIBLIOGRAPHY


