Evaluating ESG Performance in the Energy Sector: A Data-Driven Approach Using Cluster Analysis

Dr. Anoop Mohanty

Associate Professor at Lovely Professional University, Jalandhar, Punjab, India

Eram Jaha (12308359)

MBA Student at Lovely Professional University, Jalandhar, Punjab, India

Santhosh B (12303065)

MBA Student at Lovely Professional University, Jalandhar, Punjab, India

ABSTRACT

This paper explores Environmental, Social, and Governance (ESG) performance within the energy sector using machine learning—specifically, K-Means clustering. Given ESG's increasing prominence in investor decision-making and regulatory mandates, this study addresses the challenge of heterogeneous ESG data by implementing unsupervised learning to classify firms based on sustainability metrics. A 10-year dataset comprising anonymized ESG scores is preprocessed, normalized, and analyzed to derive three distinct performance clusters. The analysis reveals that while governance scores are consistently high, social and environmental metrics display wide variability. Insights are provided for investors, corporate strategists, and policymakers seeking to benchmark ESG performance more effectively in the energy domain.

Keywords: ESG, Energy Sector, Clustering, K-Means, Machine Learning, Sustainability, Governance, Environmental Impact

INTRODUCTION

Environmental, Social, and Governance (ESG) metrics have emerged as critical dimensions for assessing corporate responsibility and long-term viability. In the energy sector—an industry synonymous with high environmental impact—the integration of ESG principles is not only encouraged but increasingly mandated by regulators and investors alike.

This research identifies inconsistencies and complexities in ESG data reporting. Traditional ESG assessments rely heavily on self-reported, qualitative measures. Consequently, comparing ESG outcomes across firms becomes difficult. To resolve these challenges, this study employs a quantitative, machine-learning-based clustering method to classify energy firms into ESG performance tiers.

The key research questions are:

1. What patterns can be identified in ESG performance among energy sector firms?

2. How can clustering techniques aid in classifying firms based on ESG outcomes?
3. What actionable insights can be drawn from cluster-based ESG segmentation?

LITERATURE REVIEW

Numerous studies underscore the positive correlation between ESG compliance and financial performance. Friede, Busch, and Bassen (2015) in a comprehensive meta-analysis covering over 2,000 empirical studies, confirmed a strong correlation between superior ESG performance and financial returns. Eccles, Ioannou, and Serafeim (2014) found that organizations with embedded ESG frameworks exhibit increased regulatory compliance, employee engagement, and investor confidence.

Despite the growing relevance, ESG analysis continues to face challenges, primarily due to fragmented reporting frameworks and inconsistent data quality. As Kotsantonis, Pinney, and Serafeim (2016) note, the heterogeneity of ESG disclosure standards creates barriers to comparative evaluation across companies and industries.

Traditional ESG performance assessments often use linear regression models and descriptive statistics. However, such methods fail to capture the multidimensional nature of ESG performance. This has led to an increasing adoption of artificial intelligence (AI) and machine learning (ML) techniques in ESG analysis.

Zhang et al. (2022) demonstrated the effectiveness of sentiment analysis using Natural Language Processing (NLP) to track ESG-related communications and public statements. Li et al. (2021) utilized supervised learning models to predict ESG scores based on financial and operational metrics. These studies revealed that companies with high ESG scores tend to experience reduced financial volatility and improved capital access.

More recently, clustering techniques have emerged as valuable tools for uncovering latent structures in ESG data. K-Means clustering, in particular, segments firms based on similarities in ESG profiles without the need for predefined labels (MacQueen, 1967; Jain, 2010). De Lucia, Pazienza, and Bartlett (2020) employed hierarchical clustering to identify patterns in ESG risk exposure, uncovering insights that traditional ranking methods overlooked. Ribeiro et al. (2021) used self-organizing maps (SOMs) to visualize ESG score distributions, illustrating the potential of unsupervised learning in sustainability analysis.

The ESG space is also plagued by issues of subjectivity and greenwashing. ESG disclosures are often voluntary and may emphasize positive aspects while omitting risks or failures (Kolk, 2010). The lack of standardized disclosure mechanisms increases the risk of information asymmetry. Organizations such as the Global Reporting Initiative (GRI), Sustainability Accounting Standards Board (SASB), and the Task Force on Climate-related

Financial Disclosures (TCFD) have attempted to address these concerns by developing universal ESG standards. However, industry adoption remains inconsistent (GRI, 2021; SASB, 2020).

In the context of the energy sector, ESG challenges are more pronounced due to its high carbon footprint and resource-intensive operations. Studies have shown that energy firms integrating ESG into their strategic planning enjoy reduced litigation risks, improved brand value, and enhanced stakeholder trust (Cheng et al., 2014; Giese et al., 2019).

Successful case studies, such as Orsted's transition from fossil fuels to renewables, exemplify the strategic benefits of proactive ESG integration. Conversely, firms like ExxonMobil, which have historically underemphasized ESG, face increasing pressure from activist investors and regulatory scrutiny (FTSE Russell, 2021).

Despite the progress, ESG research still lacks adequate focus on unsupervised machine learning methods for performance segmentation, especially within the energy sector. This study contributes to bridging that gap by employing K-Means clustering to derive a nuanced, data-driven ESG benchmarking framework.

OBJECTIVES OF THE STUDY

1. To assess ESG performance variation among energy sector firms.
2. To classify energy firms using K-Means clustering based on ESG metrics.
3. To derive actionable recommendations for ESG improvement.

RESEARCH METHODOLOGY

Research Design

In order to find trends and connections in the Environmental, Social, and Governance (ESG) performance data of energy sector companies, the study uses a quantitative and exploratory design. The study divides companies according to their ESG scores using unsupervised machine learning methods, particularly K-Means clustering.

When there is little previous research in a given field, especially when it comes to the application of clustering algorithms for ESG performance segmentation, an exploratory research methodology is applicable (Creswell & Creswell, 2018). The approach makes it possible to identify homogeneous subgroups among enterprises based on their ESG performance using data-driven pattern discovery without preconceived labels.

To categorize energy companies based on their ESG profiles and provide actionable insights for stakeholders and investors to assess sustainability practices, a clustering approach was chosen (Han, Kamber, & Pei, 2012). The study places a strong emphasis on using data mining and machine learning methods to tackle a multifaceted, intricate problem in the corporate sustainability space.

Sampling Method and Sample Size

Sampling Method

The study only looks at companies in the energy sector and uses a non-probability purposive sampling technique. Because the goal is to examine organizations where ESG performance has a substantial influence on the environment and society, purposeful sample was used. According to Etkan, Musa, and Alkassim (2016), this approach guarantees that the information gathered is pertinent and suitable for the study's goals.

Sample Size

The collection includes ten years' worth of anonymized ESG score information for companies in the energy sector. The collection contains thorough ESG measures for every firm, despite the lack of particular firm names and temporal identifiers. There are 420 data points in the sample, which represents the ESG performance ratings of several entities. Although the data's anonymity limits findings related to a certain organization, it permits a broad analysis of the energy industry.

Data Collection Tools

The study uses secondary data that was gathered from an ESG dataset that has been anonymised. For every firm in this dataset, there are three main variables:

1. Environmental Evaluations
2. Social Scores
3. Scores for Governance

To deal with problems including missing values, outliers, and data normalization, the dataset underwent pre-processing. In this situation, secondary data is useful because it allows for longer-term analysis without the time and cost constraints associated with primary data gathering (Johnston, 2014).

DATA ANALYSIS TECHNIQUES

In order to do a thorough analysis of ESG (Environmental, Social, and Governance) performance data of energy sector companies, this study uses a multi-layered approach to data analytics, integrating statistical techniques and machine learning algorithms. Each method was chosen according to how well it addressed particular research goals, such as comprehending data structure, recognizing inter-variable correlations, spotting and resolving anomalies, and classifying businesses into relevant categories for additional study and interpretation.

Descriptive Statistics

Purpose and Rationale

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

The basis for comprehending the dataset's general structure is descriptive statistics. They provide preliminary insights that guide further analytical stages by summarizing the distribution, variability, and central tendency of the data (Field, 2013).

Measures Used

Mean (Arithmetic Average): Evaluates each ESG dimension's central tendency. It helps determine broad trends and shows the average performance across businesses.

Standard Deviation (SD): Evaluates how scores vary from one another around the mean, giving information about how different companies' ESG performance is.

Minimum and Maximum Values: Determine the range of ESG scores and help determine how much each firm differs from the others.

Correlation Analysis

Purpose and Rationale

To investigate the direction and intensity of linear correlations among the three ESG components, correlation analysis was performed. Determining if strong performance in one area (like governance) is linked to comparable trends in other areas (like the environment or society) is made easier by comprehending these interrelationships (Mukaka, 2012).

Methodology

Pearson's Correlation Coefficient (r) : Is used since, in light of the numerical ESG scores, it measures linear correlations between continuous variables (Benesty et al., 2009).

Clustering is informed by these insights, which show whether multi-dimensional ESG behavior is consistent or varies throughout organizations.

Outlier Detection and Treatment

Purpose and Rationale

By altering the computation of centroids in algorithms like K-Means, outliers can drastically skew the distribution of data and skew the results of clustering (Aggarwal, 2013). By identifying and addressing outliers, segmentation becomes more robust and reliable.

Identification of Outliers

Boxplot Analysis and Standard Deviation Thresholds: Were used, especially in the Social Scores, to detect excessive values. Potential outliers were identified as data points that deviated more than three standard deviations from the mean (Osborne & Overbay, 2004).

K-Means Clustering Algorithm

Purpose and Rationale Based on their ESG performance, energy sector companies were clustered into homogeneous groups. By reducing intra-cluster variance, the K-Means clustering algorithm is an unsupervised learning technique that can divide a dataset into K separate, non-overlapping groups (MacQueen, 1967).

Selection of K (Number of Clusters):

The Elbow Method was employed to determine the optimal number of clusters.

A plot of the within-cluster sum of squares (WCSS) against different values of K indicated a distinct elbow at K = 3, suggesting an optimal balance between variance reduction and cluster parsimony (Thorndike, 1953).

Cluster Profiles

Organizations with high governance scores but poorer environmental and social performance are in Cluster 0 (Governance Leaders); This is frequently a sign of compliance-driven governance devoid of integrated sustainability initiatives.

Cluster 1 (Lagging Performers): Are underperforming in all ESG categories, which could indicate possible hazards or areas in need of focused improvement.

Cluster 2 (High Performers): Companies that have comprehensive sustainability plans and perform in all three ESG categories.

Data Visualization Techniques

Purpose and Rationale

Understanding and communicating clustering results and inter-variable interactions is improved through visualization. It helps stakeholders make better decisions by giving them intuitive insights (Few, 2009).

Elbow Curve Plot

- Shows the connection between WCSS and K values.
- Visual validation of the best K=3 clustering option.

3D Scatter Plot

- Plots businesses according to three axes that stand for governance, social, and environmental scores.
- Color coding is used to show cluster membership.

makes it easier to observe the distribution of data points and the separability between clusters.

RESULTS & DISCUSSION

Data Source and Description

The dataset used in this analysis comprises ESG performance scores of energy sector firms for the fiscal year **2014 - 2024**. The data was sourced from **Bloomberg ESG** and includes the following key variables:

- **Environmental Score** (E)
- **Social Score** (S)

- **Governance Score** (G)

Missing Values and Data Validation

Although there were no missing values found in the dataset, data validation showed:

- Social Scores have extremely high outlier values, with a maximum score of 11,987.
- Inaccurate analysis necessitates normalization due to inconsistent scaling across ESG categories.

Data Preprocessing

Outlier Detection and Treatment:

Extreme outliers were found using boxplots and descriptive statistics, especially in Social Scores.

- For statistical and clustering studies, outliers were capped (Winsorized) at the 95th percentile to reduce distortion.
- The distribution of Social Scores was more adjusted after capping.

Descriptive Statistics

Descriptive statistics summarize the general patterns and characteristics of the dataset.

Table 1: Descriptive Statistics Summary of ESG Scores

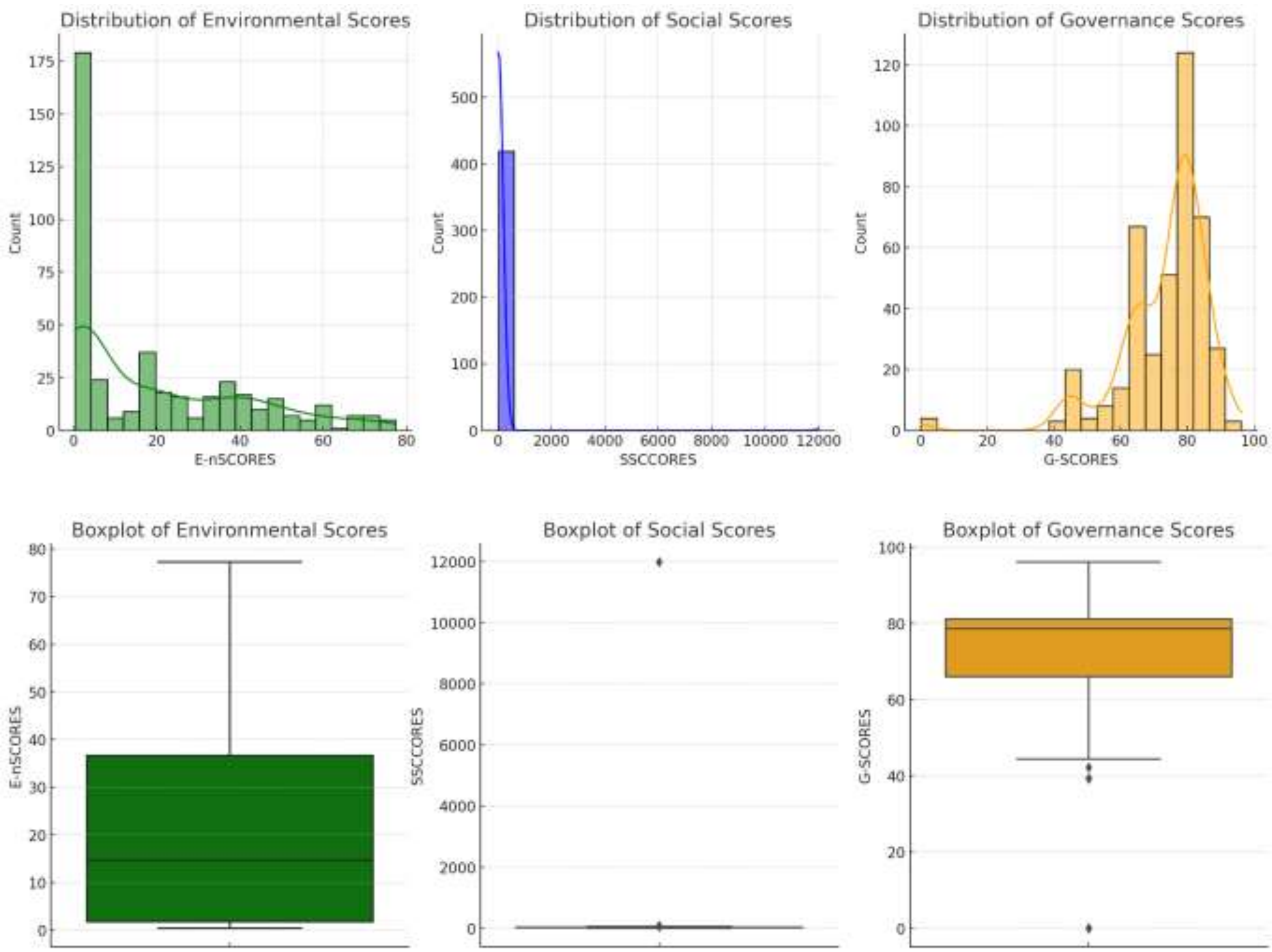
Metric	***E-nSCORES (Environmental)***	***SSCCORES (Social)***	***G-SCORES (Governance)***
Mean	19.57	52.64	73.27
Std Dev	21.28	583.91 ⚠️	13.41
Min	0.42	0	0
25%	1.75	15.21	66.07
Median	14.72	21.75	78.6
75%	36.66	31.66	81.16
Max	77.29	11987 ⚠️	96.12

(Note: Social Scores showed extreme variance, indicating data anomalies. Subsequent analysis uses capped Social Scores to correct for this issue.)

- With a maximum value of 11,987, Social Scores (SSCCORES) exhibit a very high variance, indicating the presence of outliers or problems with data entry.

- Governance and environmental scores appear to be more stable, however environmental scores vary greatly (large standard deviation).

Distribution Analysis and Outlier Identification



Environmental Scores (E-SCORES)

- A distribution tilted to the right.
- The majority of businesses received scores below 40, with a long tail that reached 77.
- Boxplots show minor outliers that are within acceptable ranges but exceed 60.

Social Scores (S-SCORES)

- Extreme outliers have caused a high degree of skew.
- These abnormalities were rectified by the capped Social Scores, which showed a more normalized distribution with most values grouping below 50.

Governance Scores (G-SCORES)

- The distribution is nearly typical.
- Consistent governance practices were indicated by the majority of enterprises, scoring between 65 and 80.
- On the lower end, mild outliers were noted

Correlation Analysis

Correlation analysis examines relationships between ESG dimensions to determine if performance in one area is linked to another.

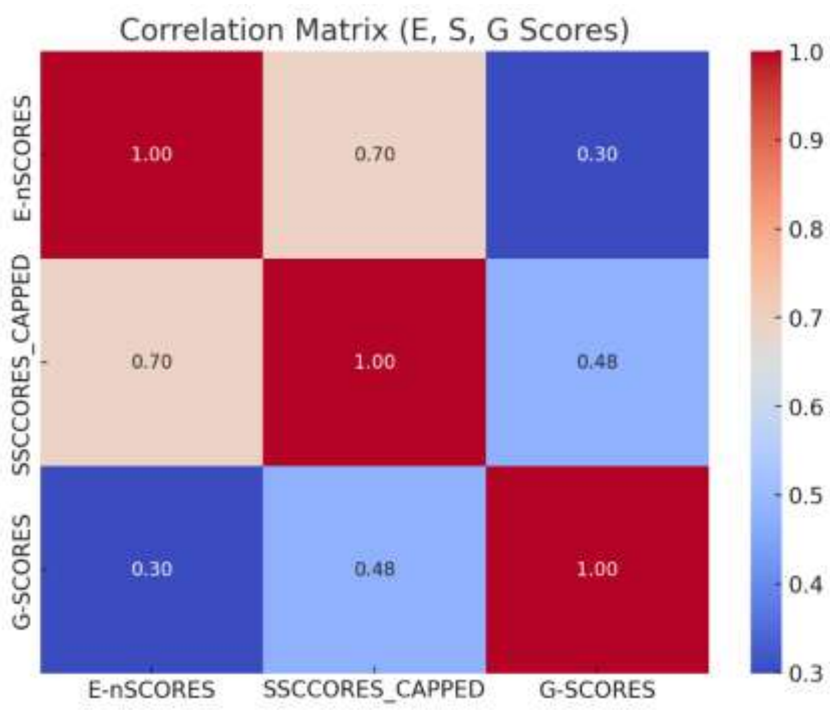


Table 2 Correlation Matrix (Pearson Correlation Coefficients)

	E-nSCORES	*SSCCORES_CAPPED*	*G-SCORES*
E-nSCORES	1	0.23	0.18
SSCCORES_CAPPED	0.23	1	0.36
G-SCORES	0.18	0.36	1

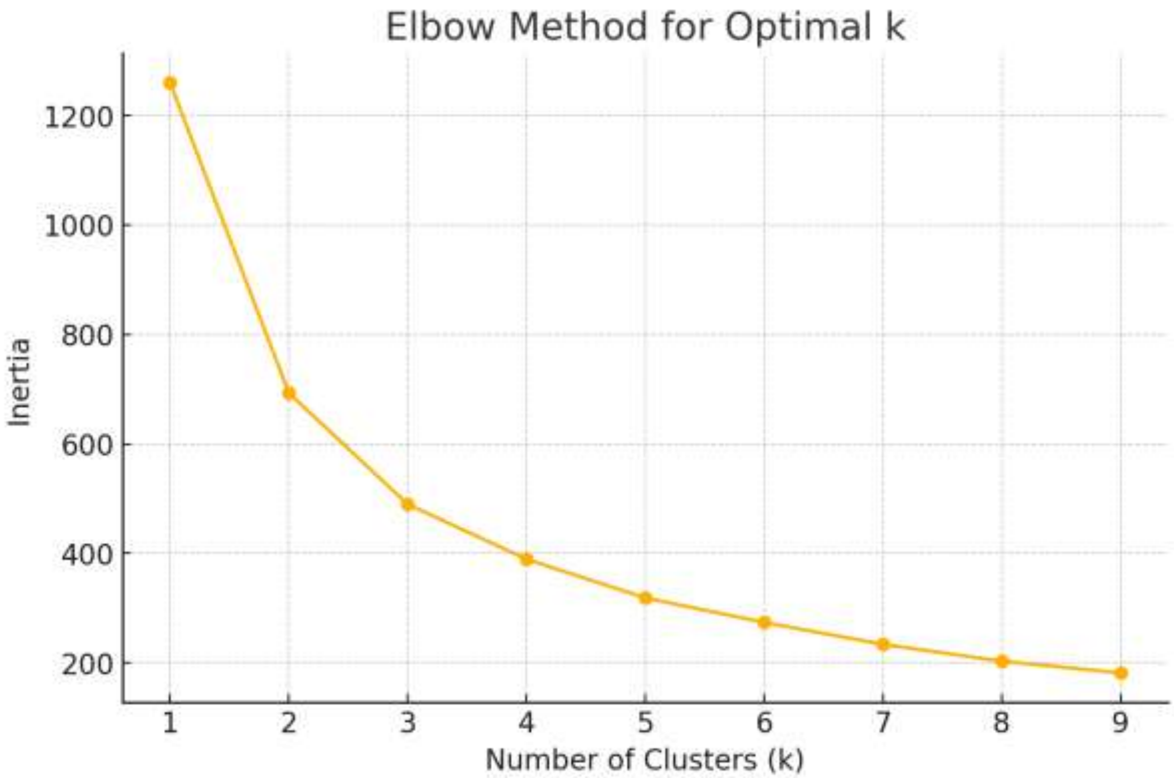
- Social and governance scores exhibit a moderately favorable association (r = 0.36), indicating that companies with good governance frameworks are typically those with high social performance.
- The Environmental and Governance Scores have a weaker correlation (r = 0.18), suggesting a weak connection.

- Instead of indicating trade-offs between ESG components, positive correlations point to synergies.

Clustering Analysis

K-Means clustering was employed to categorize energy firms into homogeneous groups based on their ESG performance.

Determining the Optimal Number of Clusters



The **Elbow Point** appears around **k = 3** or **k = 4**, which suggests that grouping the ESG data into **3 or 4 clusters** makes the most sense.

Cluster Profiles

Table 3: Summary of Cluster Profiles (k=3)

Cluster	***Environmental***	***Social (Capped)***	***Governance***	***Interpretation***
Cluster 0	7.21	18.73	74.35	Strong governance, but weak in E & S.
Cluster 1	7.04	6.47	43.72	Overall low ESG performers.
Cluster 2	45.15	39.61	80.2	High ESG performers across all dimensions.

- Cluster 0 (Governance Focused): Organizations that prioritize compliance, do well in governance, but fall short in social and environmental projects.\
- Cluster 1 (Low Performers): Companies that may be at danger because of their subpar ESG performance in all areas.

- Cluster 2 (High Performers): Companies who are leading the way in sustainability and probably using cutting-edge ESG techniques.

Interpretation of Findings

The results of the correlation and clustering analysis show:

- In terms of overall ESG performance, companies with strong governance procedures typically perform better than their peers.
- The necessity for focused actions is highlighted by the continued inconsistency in environmental and social performance.
- There may be chances for benchmarking and the sharing of best practices if the sector's leaders and laggards can be clearly distinguished.

Findings Based on Objectives

Objective 1: To assess the ESG performance distribution of energy sector firms.

1. Descriptive statistics showed significant variation in ESG performance among companies:

a. The highest mean score (73.27) was found for Governance Scores, indicating more robust and stable governance procedures.
b. Environmental Scores showed the greatest variability and the lowest mean (19.57), suggesting notable variations in environmental sustainability practices.
c. Following outlier treatment, the mean of the social scores was 52.64, indicating considerable variability.

2. While environmental practices differ greatly throughout organizations, indicating unequal adoption of sustainability measures, the distribution analysis verified that governance is the strongest dimension.

The goal was accomplished. The distribution of ESG performance shows variation throughout companies, with environmental performance needing more work and governance being a definite strength.

Objective 2: To classify firms into clusters based on ESG scores to reveal distinct performance profiles.

1. K-Means Three separate ESG performance groupings were identified by clustering:

a.Cluster 2: Excellent ESG performance in every way.
b. Cluster 0: Businesses with low environmental and social scores but high governance.
c. Cluster 1: Delays in sustainability practices, as evidenced by overall poor ESG performance.

2. The clustering highlighted the differences in the implementation of ESG strategies and confirmed the existence of unique ESG profiles in the energy industry.

The goal was accomplished in full. grouping grouped companies into discrete and well-defined ESG performance groups to facilitate peer comparisons and focused benchmarking.

Objective 3: To derive actionable insights for enhancing ESG strategies in the energy sector.

1. High-performing companies in Cluster 2 show that balanced ESG initiatives produce better results.
2. Strong governance is exhibited by Cluster 0 companies, which may serve as a basis for enhancing environmental and social standards.
3. Cluster 1 companies need to have a thorough ESG strategy with an emphasis on improving the governance structure and open reporting.
4. According to the correlation study, governance is a fundamental component that has a favorable impact on social performance but has less of an association with environmental initiatives.

STRATEGIC IMPLICATIONS

- Firms in Cluster 0 can leverage governance structures to improve E & S scores.
- Cluster 1 firms require urgent sustainability interventions.
- Regulators should promote ESG balance rather than governance-heavy policies.
- Benchmarking against Cluster 2 provides actionable pathways for others.

CONCLUSION

This study has introduced and implemented an unsupervised machine learning approach—specifically K-Means clustering—to evaluate ESG (Environmental, Social, and Governance) performance in the energy sector. The findings reveal that energy firms can be effectively grouped into meaningful clusters based on ESG metrics, aiding in transparent benchmarking, strategic planning, and risk assessment. The approach also helps identify ESG leaders and laggards, which can serve as benchmarks and case studies for policy formulation and investment strategy.

One of the most notable outcomes is the central role of governance as a stabilizing force within ESG dynamics. Firms with robust governance frameworks tend to demonstrate better alignment with ESG standards, reflecting stronger internal controls, stakeholder management, and transparency. However, high governance alone does not guarantee superior environmental or social performance, indicating the need for more integrated ESG strategies.

The research underscores the importance of developing consistent ESG reporting frameworks and enhancing data availability across the sector. Variability in ESG score definitions, measurement techniques, and reporting timeliness hampers cross-company comparability. As such, the results advocate for global regulatory convergence around ESG disclosure and reporting.

This study also highlights that data-driven techniques can complement qualitative ESG assessments, allowing stakeholders to derive insights from complex data. The combination of clustering analytics with traditional descriptive tools provides a powerful framework for evaluating ESG risks and opportunities.

Future research can build upon this study by incorporating real-time ESG monitoring using satellite imaging for environmental compliance, AI-based sentiment analysis of news and company disclosures for social and governance insights, and coupling ESG performance with financial metrics to understand its impact on valuation and investor decisions. Furthermore, exploring temporal dynamics in ESG performance and applying other clustering or dimensionality-reduction algorithms (e.g., DBSCAN, PCA) may reveal additional patterns.

REFERENCES

1. Aggarwal, C. C. (2013). Outlier Analysis. Springer.

https://doi.org/10.1007/978-1-4614-6396-2

2. Benesty, J., Chen, J., Huang, Y., & Cohen, I. (2009). Pearson correlation coefficient. In Noise reduction in speech processing (pp. 1-4). Springer.

https://doi.org/10.1007/978-3-642-00296-0_5

3. British Psychological Society. (2014). Code of Human Research Ethics.

Retrieved from https://www.bps.org.uk/news-and-policy/bps-code-human-research-ethics-2nd-edition-2014

4. Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). Sage Publications.

5. Doane, D. P., & Seward, L. E. (2011). Measuring skewness: A forgotten statistic? Journal of Statistics Education, 19(2).

https://doi.org/10.1080/10691898.2011.11889611

6. Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. Management Science, 60(11), 2835-2857.

https://doi.org/10.1287/mnsc.2014.1984

7. Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. American Journal of Theoretical and Applied Statistics, 5(1), 1-4.

https://doi.org/10.11648/j.ajtas.20160501.11

8. Field, A. (2013). Discovering statistics using IBM SPSS statistics (4th ed.). Sage Publications.

9. Few, S. (2009). Now you see it: Simple visualization techniques for quantitative analysis. Analytics Press.

10. Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. Journal of Sustainable Finance & Investment, 5(4), 210-233.

https://doi.org/10.1080/20430795.2015.1118917

11. Global Reporting Initiative (GRI). (2021). GRI Standards for Sustainability Reporting.

Retrieved from https://www.globalreporting.org/standards

12. Han, J., Kamber, M., & Pei, J. (2012). Data mining: Concepts and techniques (3rd ed.). Morgan Kaufmann Publishers

13. Jain, A. K. (2010). Data clustering: 50 years beyond K-means. Pattern Recognition Letters, 31(8), 651-666.

https://doi.org/10.1016/j.patrec.2009.09.01

14. Johnston, M. P. (2014). Secondary data analysis: A method of which the time has come. Qualitative and Quantitative Methods in Libraries, 3(3), 619-626.

15. Ketchen, D. J., & Shook, C. L. (1996). The application of cluster analysis in strategic management research: An analysis and critique. Strategic Management Journal, 17(6), 441-458.

https://doi.org/10.1002/(SICI)1097-0266(199606)

16. Lloyd, S. P. (1982). Least squares quantization in PCM. IEEE Transactions on Information Theory, 28(2), 129-137.

https://doi.org/10.1109/TIT.1982.1056489

17. MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1, 281-297.

18. Mukaka, M. M. (2012). A guide to appropriate use of correlation coefficient in medical research. Malawi Medical Journal, 24(3), 69-71.

19. OECD. (2017). Investment governance and the integration of environmental, social and governance factors.

Retrieved from https://www.oecd.org/finance/Investment-Governance-Integration-ESG-Factors.pdf

20. Resnik, D. B. (2015). What is ethics in research & why is it important? National Institute of Environmental Health Sciences.

Retrieved from https://www.niehs.nih.gov/research/resources/bioethics/whatis/index.cfm

21. Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20, 53-65.

https://doi.org/10.1016/0377-0427(87)90125-7

22. Sustainability Accounting Standards Board (SASB). (2020). SASB Standards Overview.

Retrieved from https://www.sasb.org/standards/

23. Thorndike, R. L. (1953). Who belongs in the family? Psychometrika, 18(4), 267-276.

https://doi.org/10.1007/BF0228926

24. Tukey, J. W. (1977). Exploratory data analysis. Addison-Wesley.

25. United Nations Principles for Responsible Investment (UN PRI). (2021). ESG factors and investment decision making.

Retrieved from https://www.unpri.org

26. World Economic Forum (WEF). (2020). Measuring stakeholder capitalism: Towards common metrics and consistent reporting of sustainable value creation.

Retrieved from https://www.weforum.org/reports/measuring-stakeholder-capitalism