

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

Ştefan Rusu^{1*}, Marcel Ioan Boloş², Marius Leordeanu^{3,4}

¹⁾²⁾ University of Oradea, Oradea, Romania

³⁾Polytechnic University of Bucharest, Bucharest, Romania

⁴⁾ The Institute of Mathematics of the Romanian Academy, Bucharest, Romania

Abstract

This study employs k-means clustering and agglomerative hierarchical clustering techniques to visually examine the potential relationship between Environmental Social and Governance (ESG) scores, their year-over-year variations, and annual stock returns for a sample of 34 energy sector companies operating in Europe and the United States. While the agglomerative hierarchical clustering dendrogram suggests two clusters, the elbow method of the k-means algorithm suggests 2-4 clusters. The results indicate that neither ESG scores nor their year-on-year variations had an impact on the annual returns of the stocks. The conclusion is further confirmed by the Pearson correlation coefficient. However, the ESG scores of European energy companies show a tighter dispersion and smaller year-over-year change, making them more predictable ESG score-wise and thus, potentially, more attractive to ESG-driven investors.

Keywords

Stock market, clustering, ESG, machine learning, k-means clustering, agglomerative hierarchical clustering

JEL Classification

C38; G10; G11, Q50

Introduction

Starting to gain popularity in the last two decades due to several factors such as climate changes or corporate misbehaviour, ESG, the short for Environmental (E), Social (S) and Governance (G) has been on the radar and requirements for more and more investors around the world. The three components of ESG evaluate the following: E – takes into account climate change, natural resources, pollution, waste and environmental

* Corresponding author, **Ştefan Rusu** – stefan.rusu.d@gmail.com.

opportunities to evaluate how firms take actions to protect and minimize damage to the environment; S – takes into account employee relations, working conditions, organizational diversity, human rights, employee equity and justice, inclusion, product responsibility, and community health and safety to evaluate how firms treat its employees and the communities that they serve; while G – takes into account board functions, structure, firm policies, compensation, lobbying, corruption, donation, visions and strategies to evaluate how firms' management leads and oversees their organizational authority.

ESG-driven investment strategies are becoming more and more popular. Several coalitions worldwide, such as the United Nations-backed Principles for Responsible Investment (PRI Association, n.d.) and the Climate Action 100+ (Climate Action 100+, n.d.), as well as large asset managers with tens of trillions of dollars under management, such as Blackrock (BlackRock, n.d.) or Blackstone (Blackstone, n.d.), invest with ESG principles in mind. Laying aside the obvious benefits that ESG-driven enterprises might have on the environment and community, this study focuses on the main goal of stock market investments – that of returns for the investors - and asks: how does ESG scoring impact the investment returns?

This study uses two machine learning (ML) algorithms – k-means clustering and agglomerative hierarchical clustering – to do a visual inspection as a first step to see if the ESG scores have any impact on the stock returns of companies. Due to the recent energy crisis in Europe, this study focuses on energy sector companies both in Europe, as well as the US to spot the possible differences. In the end, the Pearson correlation coefficient (Pearson, 1895) will be used to confirm or deny the visual inspection findings.

1. Review of the scientific literature

(Torre, et al., 2020) studied 46 company stocks, constituents of the Eurostoxx50 index, between 2010 and 2018. They looked at the relation between monthly stock returns and ESG indexes using a panel analysis and a multiple linear regression. The "ESG Overall" index was used from CSRHub, that considers the community, employees, environment, and governance. Their results indicate that the linear correlation between the ESG Index and stock returns is very weak and even absent. While considering the panel data analysis, they concluded that the ESG Overall index has a statistically significant and positive impact on stock returns, while the random effects model has shown that the ESG Overall index has a different impact across the analysed companies. Taking the regression analysis into account, for 7 of the 46 companies, mainly those in the utilities and energy sectors, the authors have found a significant correlation between the ESG Overall factors and stock returns but go on to explain that those results may be because ESG investments play a significant role on those companies' profitability.

(Giese, et al., 2019) analysed the link between ESG information and the valuation and performance of companies through the examination of 3 channels: a standard discounted cashflow channel, an idiosyncratic risk channel and a valuation channel. MSCI ESG ratings for the MSCI World Index universe between January 2007 and May 2017 were used along with financial variables. Industry-adjusted ESG scores were used to neutralize the results for industry exposure and size. Authors find out that ESG information has

affected the performance and valuation of the analysed companies through their idiosyncratic risk profile (higher profitability and lower exposure to tail risk), as well as through the systematic risk profile (lower costs of capital and higher valuations) concluding that the ESG characteristics transmission to financial value is a multi-channel process. Considering that the ESG ratings have a lower intensity than traditional factors such as volatility and momentum, the authors suggest that the ESG ratings have a longer-term impact and thus, suggest that they are more suitable candidates to be used in factor investing.

The effect of ESG score on stock returns in the UK's FTSE All Share Index between July 2003 and December 2020 was examined by (Luo, 2022). Thomson Reuters' ESG combined score, as well as the Environment, Social and Governance separate scores, were used. The (Fama & French, 1993) three-factor model, the (Carhart, 1997) momentum-extended Fama-French three-factor model, (Frazzini & Pedersen, 2014) betting against beta-extended Fama-French three-factor model and the (Asness, et al., 2019) quality minus junk extended Fama-French three-factor model were used to evaluate the portfolio performance. The findings support previous research that firms with lower ESG scores tend to have higher returns than firms with higher ESG scores. The study also found that the environmental and social premiums are more significant than the ESG premium, while the governance premium is insignificant. The authors suggest that the ESG premium is more significant for low-liquidity securities, but rather insignificant for high-liquidity stocks, indicating a link between ESG and stock liquidity.

(Limkriangkrai, et al., 2016) studied the impact of the ESG components: Environmental (E), Social (S) and Governance (G) and the composite score on stock returns and corporate financing decisions for the largest listed companies in the Australian market. They found no significant difference in risk-adjusted returns for portfolios based on ESG ratings. However, companies with low E and high G ratings tend to raise less debt, companies with high G ratings hold less cash, and those with low G ratings have lower dividend payouts. S ratings do not seem to have any impact on corporate financing decisions.

The impact of COVID-19 and ESG ratings on the stock performance of the US airline industry was studied by (Chen, et al., 2022) using 4 airline stocks' daily performance between January 2019 and September 2020. An autoregressive jump intensity trend model is used to evaluate the return volatility dynamics. It was found that companies with higher ESG scores experienced shorter recovery times for short-run stock return volatility after the COVID-19 shock, implying that promoting ESG offered them a defining mechanism during times of crisis. This result suggests that promoting ESG could be a suitable strategy for contemporary businesses in the airline industries and is suitable to incorporate into business operational goals.

(Consolandi, et al., 2020) examined the impact of ESG materiality on equity returns for a large sample of US-listed companies in the Russel 3000 index, from January 2008 to July 2019. Portfolios were built based on varying weights based on ESG momentum and the Gini index of materiality and based on the standard capitalization-weighted portfolio. The portfolios that were built weighting ESG momentum and the Gini index have outperformed the market capitalization-weighted portfolios, as well as the ESG

momentum-weighted portfolio, this being true especially from 2013 onwards. The results indicate that ESG rating changes have a consistent impact on equity performance, and the market rewards companies operating in industries with a high concentration of ESG materiality, as the market punishes companies with too many material targets. Additionally, the equity premium of listed companies is better explained by the concentration of material issues rather than the ESG momentum.

The relationship between ESG ratings and stock performance during the COVID-19 crash was studied by (Engelhardt, et al., 2021) between February and March 2020 considering 1452 European companies. The Thomson Reuters (actual Refinitiv) Eikon ESG ratings are used. Their findings show that higher-ESG-rated companies show higher abnormal returns and lower idiosyncratic volatility during the COVID crash, even when accounting for several multivariate specifications and robustness checks. Authors find that the social score is the main driver of good results and should be considered an important factor in investing in times of crisis. Furthermore, they suggest that the ESG score is more value-enhancing in low-trust countries and countries with weaker security regulations and disclosure standards.

A stepwise regression model and a panel regression model were used by (Trisnowati, et al., 2022) to analyse the effect of ESG score on the stock returns of 26 companies listed on the Indonesian stock exchange between 2015 and 2020. The results indicated that the ESG score did not have a significant effect on the stock returns, but it was an important variable to consider in the model, alongside financial performance variables such as Return on Assets and Debt to Equity Ratio.

(Zehir & Aybars, 2020) analysed the performance of portfolios comprised of European and Turkish companies between 2004 and 2018, built on ESG considerations. The ex-Thomson Reuters (actual Refinitiv) Eikon was used to obtain both the adjusted monthly closing prices, as well as the ESG scores. Portfolios were comprised of the top 10% ESG score companies, and the bottom 10% ESG score companies after ranking the companies using the (Kempf & Osthoff, 2007) methodology. The same was done for the individual components of the ESG score - Environment, Social and Governance. The (Fama & French, 1993) three-factor model is used to build a number of 6 total portfolios. Both the Fama and French three-factor model and Capital Asset Pricing Model (CAPM) were used to evaluate the portfolios. Both the Fama and French and CAPM models results show that the portfolios provide nearly zero abnormal returns.

2. Research methodology

This study aims to provide a first step in the form of a visual inspection that can be used by someone to pursue investigating the relationships between two variables – in this case, the ESG scores and the stock returns for companies in the energy sector. Using Machine Learning clustering algorithms provides a non-biased way to separate the data using the human eye.

The period analysed spans from the year 2013 to the end-of-year 2022 and looks at the energy-sector companies listed on the US and European stock exchanges using the companies that comprise the Energy Select SPDR Fund (XLE) for the US market and the

SPDR MSCI Europe Energy UCIS ETF (MSCI Europe Energy) as proxies. The two different markets were chosen due to their geographical as well as regulatory differences. Considering the yearly reporting frequency of the ESG scores, it makes sense to use the yearly stock return of the analysed companies along with the ESG score provided by Refinitiv Eikon, as well as the ESG delta that shows the year-to-year percentage change of the score, to examine the relationship between the ESG score and stock returns:

$$\text{ESG delta} = \text{ESG_score}_{t1} / \text{ESG_score}_{t0} \quad (1)$$

where:

ESG delta – shows the year-on-year relative changes in ESG score

ESG_score_{t1} – ESG company score in year 1

ESG_score_{t0} – ESG company score in year 0

The ESG score is the result of a data-driven assessment of companies' relative ESG performance and capacity, integrating and considering company sizes and industry materiality. The ESG performance is based on verifiable data reported in the public domain.

Refinitiv (Refinitiv, 2022) compiles and maintains the ESG data for 12500 public and private companies on a global level. Excluding special cases, this data is updated once a year to reflect the companies' own ESG disclosures. While 630 company-level ESG measures are collected, only 186 of the most comparable markers go into the final scoring process. The 186 data points make 10 ESG category scores that form the three pillars of ESG. Environmental – that considers resource use, emissions and innovation, Social – that looks at the workforce, human rights, community and product responsibility and Governance – which considers the management, shareholders and the Corporate Social Responsibility (CSR) strategy.

Due to the way ESG scores are reported, at the date of the study, May 2023, the latest available ESG information is for the year 2021. With 23 companies in the XLE and 11 in the MSCI Europe Energy, after checking for the missing data, we end up with 290 observations. The data were obtained from the Refinitiv Eikon database through the Python API.

K-means clustering and agglomerative hierarchical clustering will be used to divide the plots into clusters to inspect the relationship easily visually between them. In the end, the Pearson correlation coefficient (Pearson, 1895) is used to confirm or deny the results of the visual inspection.

First introduced by (MacQueen, 1967), k-means clustering is an unsupervised machine learning algorithm used to group observations together based on their distance to their respective cluster. To choose the right number of clusters, the widely used elbow method is utilized here as well (Bedzek, 1981). The method uses the Within Cluster Sum of Squares between the data points and the cluster centroid that they are assigned to, to plot a graph in which the point in which the line bends and resembles an elbow is the right number of clusters to be chosen.

K-means algorithm steps:

1. The k-means algorithm itself is initialized on a k number of clusters indicated by the elbow method or the visual inspection. Based on the number of clusters, k data points are chosen randomly to be the first cluster centroids.
2. Based on the Euclidean distance between the other data points and the cluster centroid, the points are assigned to a cluster.
3. After all the data points are assigned, each cluster's centroid is reassigned based on the mean of the data points.
4. Steps 2 and 3 are repeated until the changes to the cluster centroid's positions are no longer significant.
5. The k-means algorithm outputs the clusters and their corresponding data points.

Agglomerative hierarchical clustering (Nielsen, 2016) is a bottom-up clustering method which starts by treating each data point as a single cluster and merging the closest ones, based on the Euclidean distance between each cluster until one big cluster is formed. A dendrogram (Jardine & Sibson, 1968) is formed after each iteration of the clustering process. Ward's method (Ward, 1963) is used to pair the cluster to merge at each step. The right number of clusters is chosen by visualizing the dendrogram and selecting the longest vertical lines that aren't intersected by other lines.

3. Results and discussions

Plotting the MSCI Energy Europe and XLE companies ESG score and delta against the % yearly returns (Figures no. 1, 2, 3 and 4), we cannot see a pattern in the data. Although, the plots show that the European companies tend to have more grouped ESG scores ranging from the 50s to 90s, while the US companies tend to have more spread out ESG scores from low 10s to 90s. The same holds for the magnitude of the ESG score changes from year to year, measured by the ESG delta. Where European energy companies see changes from negative high single digits to positive tens, their US counterpart see changes from around -40% to +60% year-over-year.

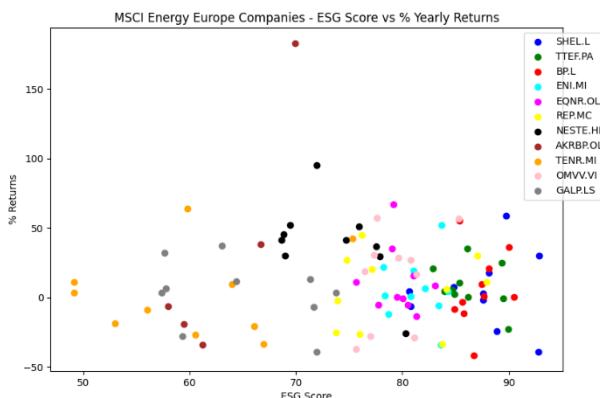


Figure no. 1: Scatterplot of the ESG scores (x-axis) and % yearly returns (y-axis) for the companies in the MSCI Europe Energy ETF

Source: Author's own work

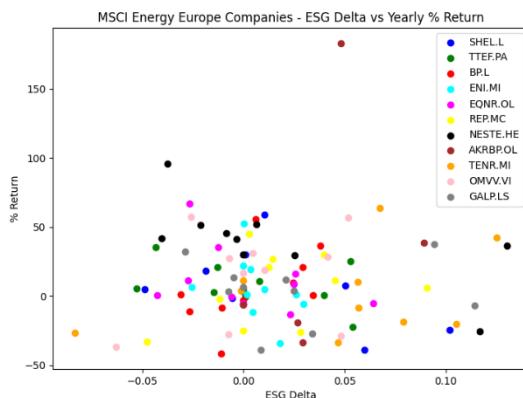


Figure no. 2. Scatterplot of the ESG Delta (x-axis) and % yearly returns (y-axis) for the companies in the MSCI Europe Energy ETF

Source: Author's own work

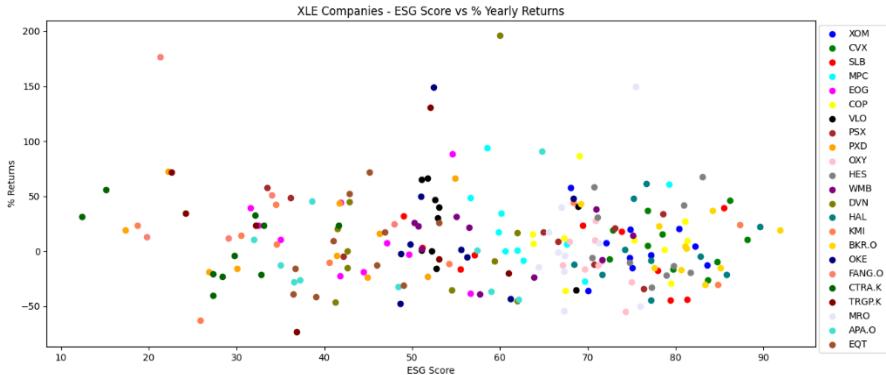


Figure no. 3: Scatterplot of the ESG scores (x-axis) and % yearly returns (y-axis) for the companies in the XLE ETF

Source: Author's own work

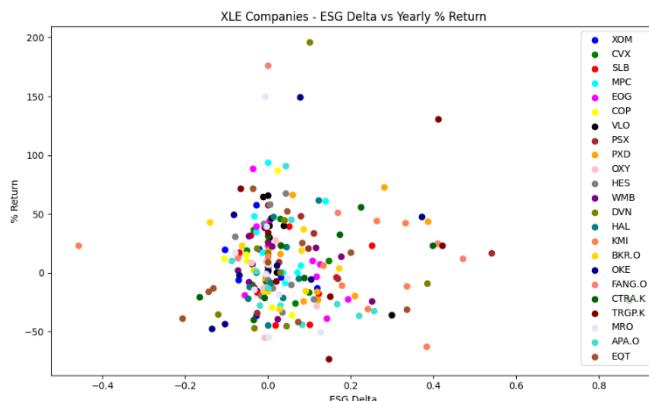


Figure no. 4: Scatterplot of the ESG Delta (x-axis) and % yearly returns (y-axis) for the companies in the XLE ETF

Source: Author's own work

Going forward to the k-means clustering, both for the MSCI Energy Europe and XLE, as well as ESG score and ESG delta, the elbow method suggests an inconclusive number of 2-4 clusters as optimal, as can be seen in the example in Figure no. 5.

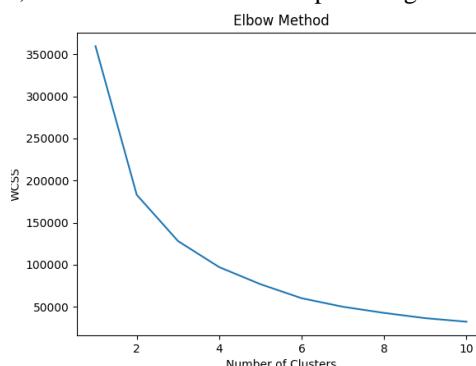


Figure no. 5. Elbow method for companies in the XLE index based on the ESG Score

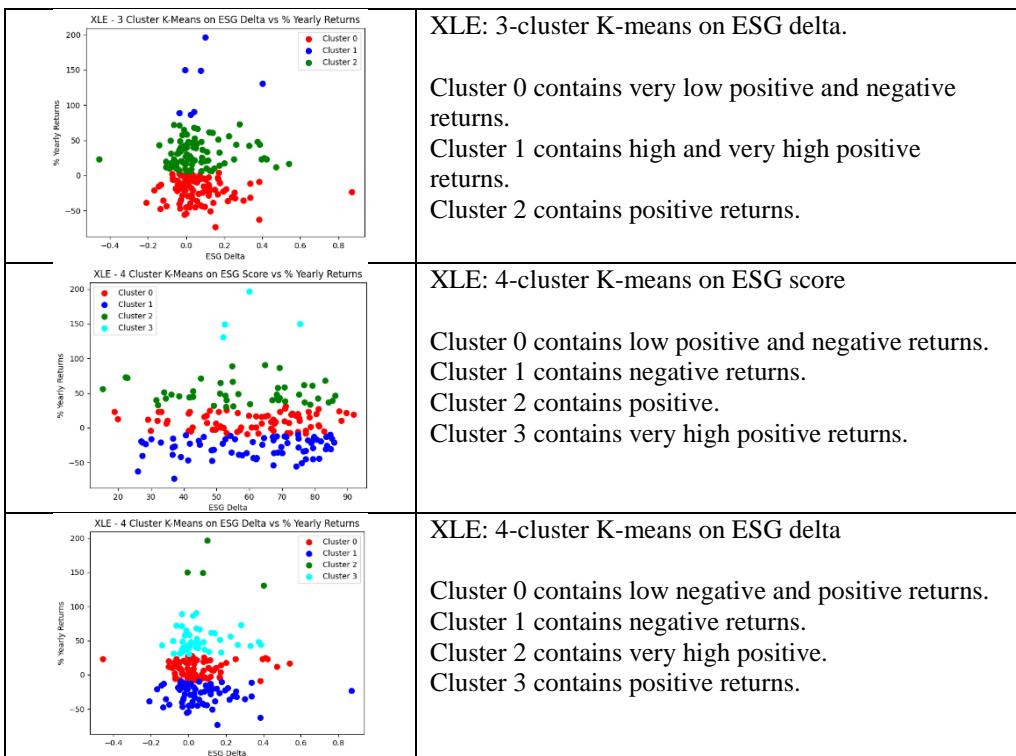
Source: Author's own work

The results of the K-means clustering, for both markets do not show groups made based on the ESG scores or ESG delta, but on the % yearly returns. The clusterings and their interpretations can be viewed in Table no. 1.

Table no. 1. ESG score, ESG delta and % yearly returns scatterplots for MSCI Energy Europe and XLE clustered in 2, 3 and 4 clusters using the k-means clustering technique

Clustering results of the k-means clustering applied on ESG scores, ESG delta and % yearly returns	
K-means scatterplots	Interpretation
	<p>MSCI Energy Europe: 2-cluster K-means on ESG score.</p> <p>Cluster 0 contains positive returns. Cluster 1 contains very low positive and negative returns.</p>
	<p>MSCI Energy Europe: 2-cluster K-means on ESG delta.</p> <p>Cluster 0 contains positive returns. Cluster 1 contains very low positive and negative returns.</p>
	<p>MSCI Energy Europe: 3-cluster K-means on ESG score.</p> <p>Cluster 0 contains low positive and negative returns. Cluster 1 contains positive returns. Cluster 2 contains very high positive returns.</p>
	<p>MSCI Energy Europe: 3-cluster K-means on ESG delta.</p> <p>Cluster 0 contains low positive and negative returns. Cluster 1 contains positive returns. Cluster 2 contains very high positive returns.</p>

	<p>MSCI Energy Europe: 4-cluster K-means on ESG delta.</p> <p>Cluster 0 contains positive returns. Cluster 1 contains negative returns. Cluster 2 contains low positive and negative returns. Cluster 3 contains very high positive returns.</p>
	<p>XLE: 2-cluster K-means on ESG score.</p> <p>Cluster 0 contains positive returns. Cluster 1 contains very low positive and negative returns.</p>
	<p>XLE: 2-cluster K-means on ESG delta.</p> <p>Cluster 0 contains positive returns. Cluster 1 contains very low positive and negative returns.</p>
	<p>XLE: 3-cluster K-means on ESG score.</p> <p>Cluster 0 contains very low positive and negative returns. Cluster 1 contains high and very high positive returns. Cluster 2 contains positive returns.</p>



Source: Author's own work

From the plots above, it can be seen that k-means clustering does not provide conclusive visual proof that ESG score or ESG delta would have a significant impact on stock prices. Going further to the agglomerative hierarchical clustering, the number of clusters suggested by the dendrograms is 2, as can be seen in Figure no. 6.

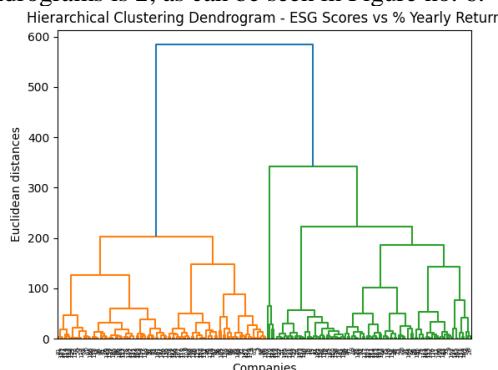
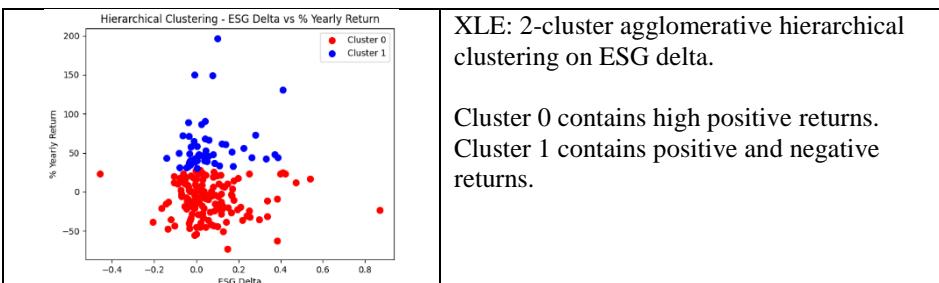


Figure no. 6. Dendrogram for companies in the XLE index based on the ESG Score

Similar results to the k-means clustering are obtained with the agglomerative hierarchical clustering as well. The results of the agglomerative hierarchical clustering, for both markets, do not show clusters made based on the ESG scores or ESG delta, but on the % yearly returns. The clusterings and their interpretations can be viewed in Table no. 2.

Table no. 2. ESG score, ESG delta vs % yearly returns scatterplots for MSCI Energy Europe and XLE clustered in 2 clusters using the agglomerative hierarchical clustering technique

Clustering results of the agglomerative hierarchical clustering applied on ESG scores, ESG delta and % yearly returns	
Agglomerative hierarchical clustering scatterplots	Interpretation
	<p>MSCI Energy Europe: 2-cluster agglomerative hierarchical clustering on ESG score.</p> <p>Cluster 0 contains positive returns. Cluster 1 contains low positive and negative returns.</p>
	<p>MSCI Energy Europe: 2-cluster agglomerative hierarchical clustering on ESG delta.</p> <p>Cluster 0 contains high positive returns. Cluster 1 contains positive and negative returns.</p>
	<p>XLE: 2-cluster agglomerative hierarchical clustering on ESG score.</p> <p>Cluster 0 contains positive returns. Cluster 1 contains very low positive and negative returns.</p>



It can be seen that agglomerative hierarchical clustering does not provide conclusive visual proof that ESG score or ESG delta would have a significant impact on stock prices.

Calculating the Pearson correlation coefficient between the % yearly return and the ESG score and ESG delta, it can be seen that if a link between ESG scores or ESG delta and yearly stock returns were to exist, it would be too low to be relevant. The results can be observed in Table no. 3.

Table no. 3. The Pearson correlation scores

Pearson correlation scores		
Returns ESG	ESG score	ESG delta
MSCI Europe's companies % yearly return	-0.02538	-0.02799
XLE's companies % yearly return	-0.04397	0.01575

Conclusions

This study used k-means clustering and agglomerative hierarchical clustering algorithms to cluster energy companies listed on the European and United States markets based on the stock yearly returns and their ESG scores. At a visual inspection level using the clustering algorithms, ESG score and ESG delta do not influence stock returns, as can be seen from the scatter plots. Both unsupervised machine learning algorithms used chose to split the data points based on the yearly returns of the stocks included in the ETFs, rather than the ESG scores. This fact is confirmed by the correlation scores as well.

Nonetheless, it is interesting to see when plotting all the companies in both the European as well as the US energy companies ETFs that the companies in the European ETF have a tighter spread of ESG scores and a smaller move of the ESG score year-on-year, as measured by ESG delta when compared to their US counterparts.

This information can be of use to investors that need to adhere to specific ESG investment directives, such as investing in stocks with a specific minimum ESG score. Choosing to allocate their money to European energy stocks could prove to be more predictable than US energy stocks, ESG score-wise, thus diminishing the risk of having to do unwanted portfolio relocations due to sharp changes in the ESG scores. The avoidance of unwanted

portfolio relocations could save those investors both on transaction costs and fees, as well as potential returns.

References

- [1] Asness, C., Frazzini, A. & Pedersen, L., 2019. Quality minus junk. *Review of Accounting Studies*, pp. 34-112.
- [2] Bedzek, J. C., 1981. *Pattern Recognition with Fuzzy Objective Function Algorithms*.
- [3] BlackRock, n.d. Sustainable investing at BlackRock. [Online] Available at: <https://www.blackrock.com/ch/individual/en/themes/sustainable-investing>
- [4] Blackstone, n.d. An Integrated Approach to ESG. [Online] Available at: <https://www.blackstone.com/our-impact/an-integrated-approach-to-esg/>
- [5] Carhart, M., 1997. On persistence in mutual fund performance. *The Journal of Finance*, pp. 57-82.
- [6] Chen, C.-D., Su, C.-H. (. & Chen, M.-H., 2022. Understanding how ESG-focused airlines reduce the impact of the COVID-19 pandemic on stock returns. *Journal of Air Transport Management*.
- [7] Climate Action 100+, n.d. *GLOBAL INVESTORS DRIVING BUSINESS TRANSITION*. [Online] Available at: <https://www.climateaction100.org/>
- [8] Consolandi, C., Eccles, R. G. & Giampaolo, 2020. How material is a material issue? Stock returns and the financial relevance and financial intensity of ESG materiality. *Journal of Sustainable Finance & Investment*.
- [9] Engelhardt, N., Ekkenga, J. & Posch, P., 2021. ESG Ratings and Stock Performance during the COVID-19 Crisis. *Sustainability*.
- [10] Fama, E. & French, K., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, pp. 3-56.
- [11] Frazzini, A. & Pedersen, L., 2014. Betting against beta. *Journal of Financial Economics*, pp. 1-25.
- [12] Giese, G. et al., 2019. Foundations of ESG Investing: How ESG Affects Equity Valuation, Risk, and Performance. *The Journal of Portfolio Management*, pp. 69-83.
- [13] Jardine, N. & Sibson, R., 1968. The Construction of Hierarchic and Non-Hierarchic Classifications. *The Computer Journal, Volume 11, Issue 2*, p. 177–184.
- [14] Kempf, A. & Osthoff, P., 2007. The effect of socially responsible investing on portfolio. *European Financial Management*, Vol. 13 No. 5., pp. 908-922.
- [15] Lee, M. T. & Suh, I., 2022. Understanding the effects of Environment, Social, and Governance conduct on financial performance: Arguments for a process and integrated modelling approach. *Sustainable Technology and Entrepreneurship*.
- [16] Limkriangkrai, M., Koh, S. & Durand, R. B., 2016. Environmental, Social, and Governance (ESG) Profiles, Stock Returns, and Financial Policy: Australian Evidence. *International Review of Finance*.
- [17] Luo, D., 2022. ESG, liquidity, and stock returns. *Journal of International Financial Markets, Institutions & Money*.
- [18] MacQueen, J., 1967. *MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations..* s.l., s.n., pp. 281-297.
- [19] Nielsen, F., 2016. *Introduction to HPC with MPI for Data Science*. s.l.:Springer.

- [20] Pearson, K., 1895. *Note on Regression and Inheritance in the Case of Two Parents.* s.l., Royal Society, pp. 240-242.
- [21] PRI Association, n.d. *What are the Principles for Responsible Investment?*. [Online] Available at: <https://www.unpri.org/about-us/what-are-the-principles-for-responsible-investment>
- [22] Refinitiv, 2022. *ENVIRONMENTAL, SOCIAL*, s.l.: London Stock Exchange Group.
- [23] Torre, M. L., Mango, F., Cafaro, A. & Leo, S., 2020. Does the ESG Index Affect Stock Return? Evidence from the Eurostoxx50. *Sustainability*, pp. 12, 6387.
- [24] Trisnowati, Y., Achsani, N. A. & Sembel, R. e. a., 2022. The effect of ESG score, financial performance, and macroeconomics on stock returns during the pandemic era in Indonesia. *International Journal of Energy Economics and Policy*, pp. 166-172.
- [25] Ward, J. H. J., 1963. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association* 58, pp. 236-244.
- [26] Zehir, E. & Aybars, A., 2020. Is there any effect of ESG scores on portfolio performance? Evidence from Europe and Turkey. *Journal of Capital Markets Studies*, pp. 129-143.