

# Analyse ESG et rendements boursiers : Clustering, exploration empirique et discussion méthodologique

## Application à la supply chain de NVIDIA (2016–2026)

Théo Linale  
Victoria Vidal  
Franck Jorel Foguefeussi

## Table des matières

<b>1</b>	<b>Introduction et contexte général</b>	<b>2</b>
1.1	Les défis de notre époque et l'émergence d'une nouvelle exigence . . . . .	2
1.2	L'évolution des attentes des investisseurs . . . . .	2
1.3	Les questions des investisseurs . . . . .	2
1.4	Les critères ESG et leurs piliers . . . . .	3
1.5	Qui utilise l'ESG aujourd'hui ? . . . . .	3
1.6	Papier de référence et objectif du projet . . . . .	3
<b>2</b>	<b>Données et construction du panel</b>	<b>4</b>
2.1	Sources de données . . . . .	4
2.2	Construction des variables . . . . .	4
2.3	Traitement des valeurs manquantes et contrôle qualité . . . . .	4
<b>3</b>	<b>Limites méthodologiques et corrections apportées</b>	<b>4</b>
3.1	Limite 1 : biais d'échelle et domination de la variance . . . . .	5
3.2	Correction : standardisation par Z-score . . . . .	5
3.3	Limite 2 : structure non gaussienne des données . . . . .	6
3.4	Limite 3 : nature panel et dépendances . . . . .	6
<b>4</b>	<b>Résultats empiriques</b>	<b>7</b>
4.1	Corrélations ESG–rendements . . . . .	7
4.2	Analyse graphique : relation ESG–rendement . . . . .	7
4.3	Choix du nombre de clusters . . . . .	8
4.4	Visualisation des clusters . . . . .	9
4.5	Profil économique des clusters . . . . .	9
4.6	Régressions linéaires . . . . .	10
<b>5</b>	<b>Conclusion</b>	<b>10</b>
5.1	Synthèse des résultats . . . . .	10
5.2	Implications et perspectives . . . . .	11

# 1 Introduction et contexte général

## 1.1 Les défis de notre époque et l'émergence d'une nouvelle exigence

Les économies contemporaines font face à des défis majeurs qui dépassent largement la seule question de la performance financière. Le changement climatique, la multiplication des catastrophes naturelles, l'urgence environnementale, l'accroissement des inégalités sociales ainsi que la répétition de scandales de gouvernance (corruption, manque de transparence, rémunérations excessives) ont profondément modifié la manière dont les entreprises sont perçues et évaluées.

Dans ce contexte, les entreprises ne sont plus jugées uniquement sur leur capacité à générer du profit à court terme, mais également sur leur impact environnemental, social et sur la qualité de leur gouvernance. Cette évolution s'accompagne de l'émergence d'une nouvelle exigence pour les investisseurs : *investir différemment*, en intégrant des critères extra-financiers dans leurs décisions d'allocation de capital.

## 1.2 L'évolution des attentes des investisseurs

Historiquement, dans les années 1990–2000, la logique dominante reposait sur la maximisation du rendement financier à court terme et sur la création de valeur pour l'actionnaire (*shareholder value*), conformément à la vision défendue notamment par Milton Friedman. L'impact social et environnemental des entreprises était largement ignoré.

À partir des années 2010, on observe un changement progressif de paradigme. Les investisseurs s'intéressent de plus en plus à une approche combinant performance financière et impact positif, dans une logique de *stakeholder capitalism*. Cette approche se traduit par la prise en compte du *triple bottom line* (People, Planet, Profit) et par le développement de l'investissement socialement responsable (ISR), qui consiste à intégrer des critères environnementaux, sociaux et de gouvernance dans les décisions d'investissement en plus des critères financiers traditionnels.

## 1.3 Les questions des investisseurs

Cette évolution soulève plusieurs questions fondamentales :

- Comment savoir si une entreprise est réellement responsable ?
- Comment comparer objectivement deux entreprises sur leur impact environnemental et social ?
- Comment mesurer quelque chose d'aussi complexe que la « responsabilité d'entreprise » ?
- Quels critères utiliser et à quelles sources de données faire confiance ?

Ces interrogations justifient le recours à des systèmes d'évaluation structurés, mesurables et comparables, parmi lesquels les scores ESG occupent aujourd'hui une place centrale.

## 1.4 Les critères ESG et leurs piliers

Les critères ESG (Environnement, Social, Gouvernance) constituent un cadre d’analyse visant à évaluer la performance extra-financière des entreprises :

- **E – Environnemental** : gestion du climat et des émissions de CO<sub>2</sub>, pollution et déchets, gestion des ressources naturelles, biodiversité, économie circulaire.
- **S – Social** : traitement des employés, diversité et inclusion, conditions de travail, santé et sécurité, droits de l’homme, relations communautaires, responsabilité produit.
- **G – Gouvernance** : transparence du conseil d’administration, lutte contre la corruption, rémunération des dirigeants, structure de gouvernance, gestion des risques, éthique des affaires.

Ces dimensions sont agrégées par des fournisseurs de données spécialisés (tels que Bloomberg, MSCI, Sustainalytics ou Refinitiv) afin de produire des scores synthétiques, largement utilisés par les investisseurs institutionnels.

## 1.5 Qui utilise l’ESG aujourd’hui ?

L’ESG est aujourd’hui intégré par de nombreux acteurs majeurs de la finance mondiale. On peut citer, par exemple, des coalitions internationales comme *Climate Action 100+* ou les *Principles for Responsible Investment* (PRI) soutenus par l’ONU, ainsi que de grands gestionnaires d’actifs tels que BlackRock ou Blackstone, qui intègrent désormais les critères ESG dans leurs processus de gestion des risques et d’investissement.

## 1.6 Papier de référence et objectif du projet

Ce travail s’inscrit dans la continuité du papier :

*“K-means et analyse de clustering hiérarchique agglomératif des scores ESG, variations annuelles et rendements boursiers : perspectives du secteur de l’énergie en Europe et aux États-Unis”,*

Ștefan Rusu, Marcel Ioan Boloș, Marius Leordeanu (2023).

L’objectif de ce projet est de reproduire et d’améliorer l’approche proposée dans ce papier, en corrigeant certaines limites méthodologiques, et en l’appliquant à un nouvel univers d’investissement : 18 entreprises de la supply chain de NVIDIA, à partir de données Bloomberg BESG sur la période 2016–2026.

Nous ne cherchons pas à établir une relation causale entre ESG et performance financière, mais à analyser la structure empirique des données, à identifier d’éventuels profils d’entreprises et à évaluer la sensibilité des résultats aux choix méthodologiques.

## 2 Données et construction du panel

### 2.1 Sources de données

À l’origine, notre base de données ne contenait que des scores ESG issus de Bloomberg (scores environnementaux, sociaux, et score ESG agrégé). En revanche, contrairement au papier de référence, nous ne disposions pas directement des rendements boursiers.

Nous avons donc :

- Collecté les données de prix boursiers via Yahoo Finance,
- Calculé les rendements à partir des prix ajustés,
- Agrégé ces rendements à la même fréquence temporelle que les données ESG,
- Puis fusionné (merge) les données ESG et financières par date et par entreprise.

L’échantillon final est composé de 18 entreprises de la supply chain de NVIDIA, observées sur la période 2016–2026.

### 2.2 Construction des variables

À partir des prix ajustés  $P_t$ , les rendements simples sont calculés comme :

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}.$$

Nous utilisons également les scores ESG et, lorsque pertinent, leurs variations temporelles. L’ensemble des données est restructuré sous forme de panel (une ligne par couple entreprise–date), afin de faciliter les analyses statistiques et les visualisations.

### 2.3 Traitement des valeurs manquantes et contrôle qualité

Les données ESG présentent des valeurs manquantes qui ne sont pas aléatoires : elles correspondent souvent à des périodes où le score n’était pas encore disponible pour certaines entreprises.

On distingue notamment :

- Les valeurs manquantes en début de série (left-censoring),
- Les valeurs manquantes sporadiques, liées à des recalculs ou à des problèmes de données.

Un contrôle qualité est effectué après la fusion ESG–Finance afin de vérifier la couverture temporelle, le taux de rendements manquants et l’hétérogénéité des manques par entreprise. Les entreprises présentant un taux de valeurs manquantes trop élevé sont exclues afin de garantir la robustesse statistique des résultats.

## 3 Limites méthodologiques et corrections apportées

L’analyse exploratoire du panel fusionné ESG–rendements met en évidence plusieurs limites méthodologiques susceptibles d’affecter la validité et l’interprétation des résultats.

Ces limites concernent principalement les différences d'échelle entre variables, la structure statistique des données ESG et la nature panel de l'échantillon.

### 3.1 Limite 1 : biais d'échelle et domination de la variance

Les scores ESG sont bornés et relativement stables dans le temps, tandis que les rendements boursiers présentent une volatilité nettement plus élevée et des distributions asymétriques avec queues épaisses.

Or, l'algorithme K-means repose sur une minimisation de l'inertie intra-cluster :

$$\sum_{k=1}^K \sum_{i \in C_k} \|x_i - \mu_k\|^2.$$

En présence de variables ayant des variances très différentes, la distance euclidienne est dominée par la variable la plus dispersée.

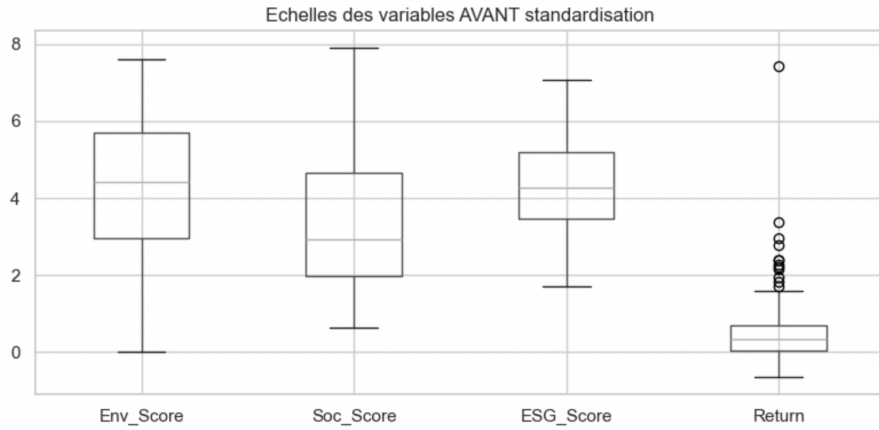


FIGURE 1 – Distribution des variables avant standardisation.

La Figure 1 met clairement en évidence que l'étendue des rendements est proportionnellement plus importante que celle des scores ESG. Sans correction, la segmentation serait presque entièrement pilotée par la volatilité des rendements, ce qui reviendrait à effectuer un clustering essentiellement financier plutôt qu'ESG-financier.

### 3.2 Correction : standardisation par Z-score

Afin d'assurer une contribution équilibrée de chaque variable, nous appliquons une standardisation :

$$Z = \frac{X - \mu}{\sigma}.$$

Cette transformation centre les variables et les ramène à une variance unitaire.

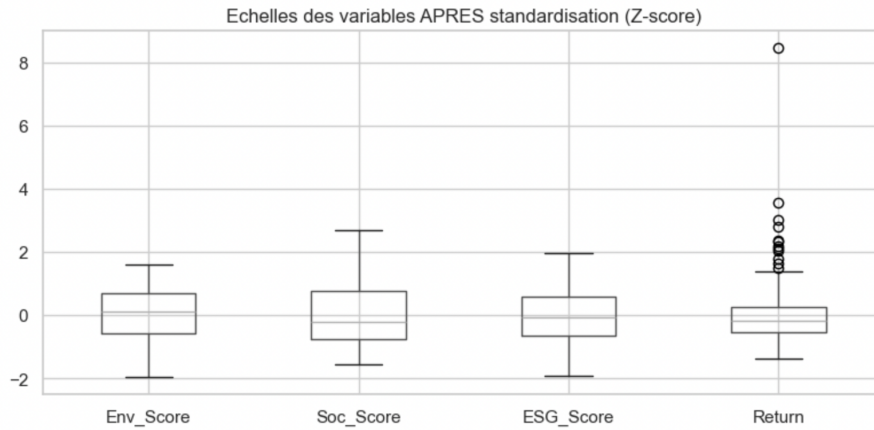


FIGURE 2 – Distribution des variables après standardisation.

La Figure 2 confirme que les échelles deviennent comparables. Cette étape est essentielle pour toute analyse multivariée basée sur des distances.

### 3.3 Limite 2 : structure non gaussienne des données

Les distributions ESG présentent des asymétries et parfois une bimodalité liée à des différences sectorielles. De plus, les rendements financiers présentent des queues épaisses.

Or, K-means suppose implicitement des clusters convexes et de dispersion comparable. Si la structure réelle des données s'écarte de cette hypothèse, les clusters obtenus peuvent refléter une approximation géométrique plutôt qu'une structure économique sous-jacente.

Nous utilisons donc la méthode du coude et le score de silhouette afin d'évaluer empiriquement la qualité de la segmentation.

### 3.4 Limite 3 : nature panel et dépendances

Les données combinent une dimension inter-entreprises et une dimension temporelle.

Une corrélation simple entre ESG et rendement agrège :

- des différences structurelles entre entreprises,
- des variations communes liées aux conditions macroéconomiques.

Cela limite toute interprétation causale. Nous introduisons donc des effets fixes temporels dans certaines régressions afin d'isoler les variations transversales.

## 4 Résultats empiriques

### 4.1 Corrélations ESG–rendements

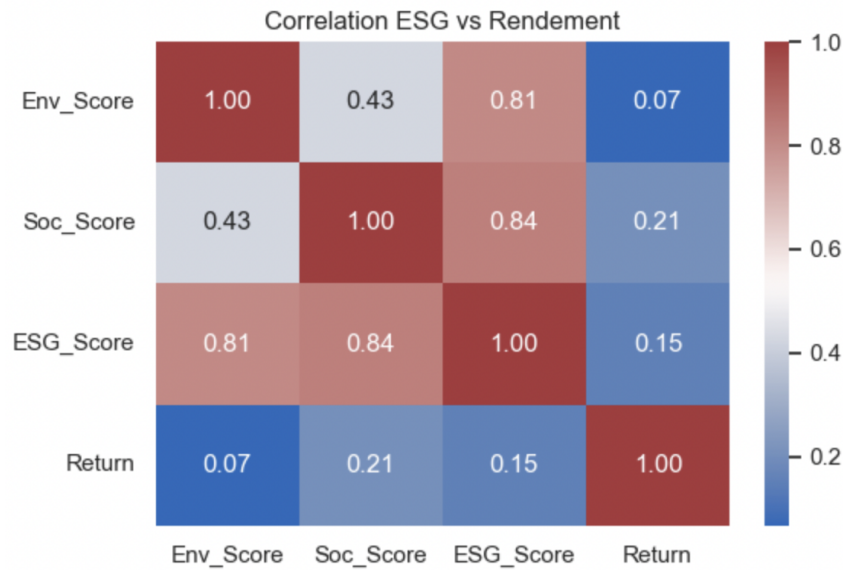


FIGURE 3 – Matrice de corrélation entre scores ESG et rendements.

La matrice de corrélation (Figure 3) met en évidence des corrélations positives mais faibles entre les scores ESG et les rendements. Les coefficients se situent dans une fourchette modérée, ce qui suggère qu’il n’existe pas de relation linéaire forte et directe dans l’échantillon agrégé.

Toutefois, cette mesure reste purement descriptive. Elle ne permet pas de distinguer les effets sectoriels, ni les dynamiques temporelles, ni les potentielles non-linéarités.

### 4.2 Analyse graphique : relation ESG–rendement

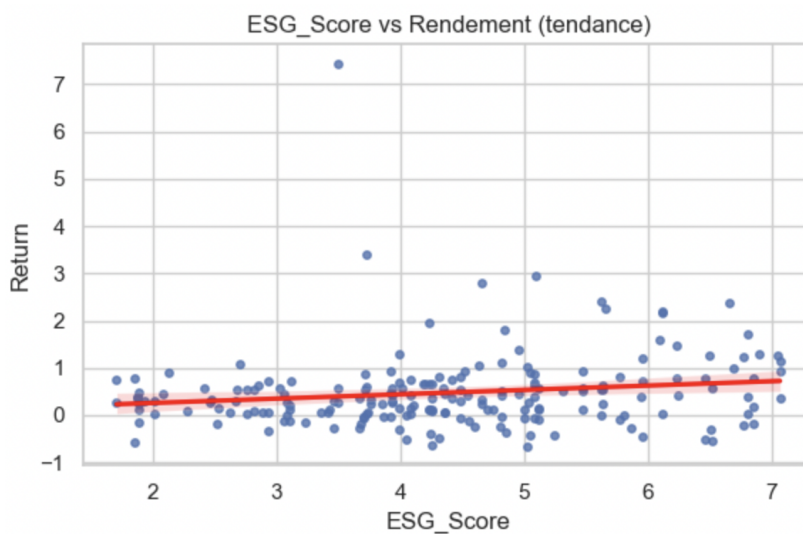


FIGURE 4 – Nuage de points ESG vs rendement avec droite de tendance.

Le nuage de points met en évidence une tendance positive mais faible. La dispersion importante indique que le score ESG n'explique qu'une part réduite de la variabilité des rendements.

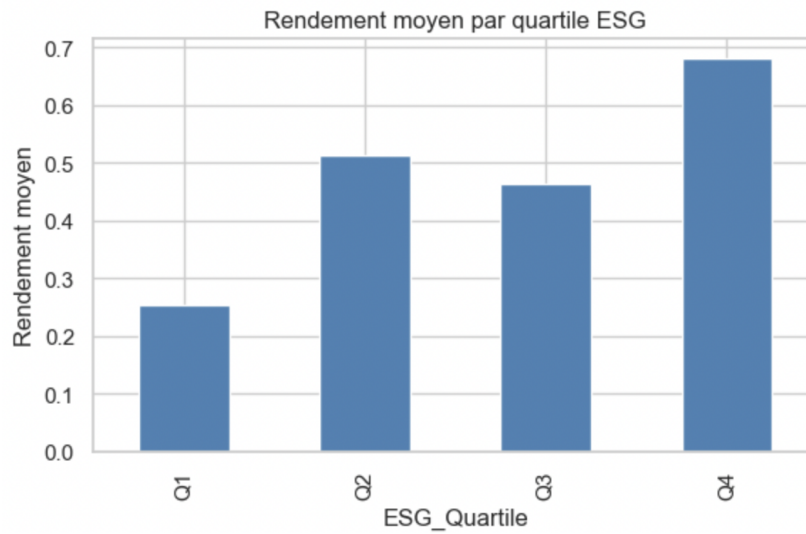


FIGURE 5 – Rendement moyen par quartile ESG.

L'analyse par quartiles suggère une relation monotone : les entreprises à ESG élevé présentent un rendement moyen supérieur. Toutefois, cette relation ne contrôle ni la taille, ni le secteur, ni le risque systématique.

### 4.3 Choix du nombre de clusters

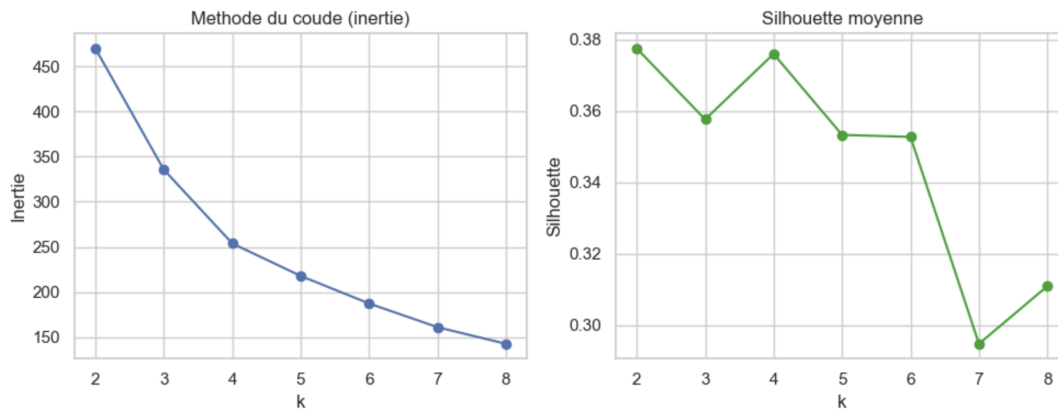


FIGURE 6 – Méthode du coude et score de silhouette.

La méthode du coude suggère une inflexion autour de  $K = 4$ , tandis que le score de silhouette est maximal pour  $K = 2$  (0.37).

Une silhouette de 0.37 indique une structure modérée : les clusters sont distincts mais partiellement chevauchants.



## 4.4 Visualisation des clusters

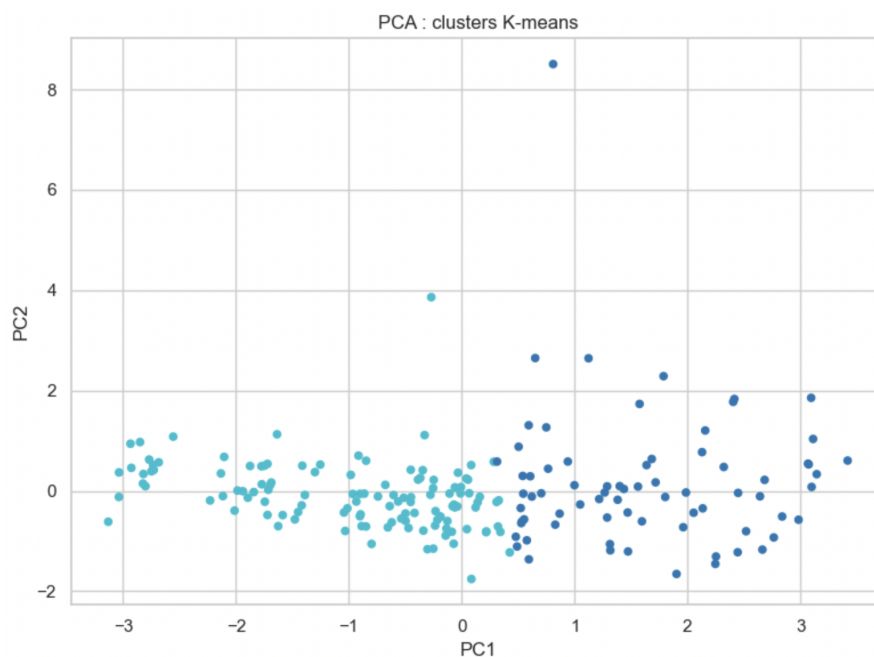


FIGURE 7 – Projection ACP colorée par clusters.

La projection ACP montre que la première composante principale capture une part importante de la variance totale. Cette composante semble correspondre à un facteur ESG commun, étant donné la forte corrélation entre les composantes Env, Soc et ESG global.

La séparation des clusters s'effectue principalement le long de cette dimension.

## 4.5 Profil économique des clusters

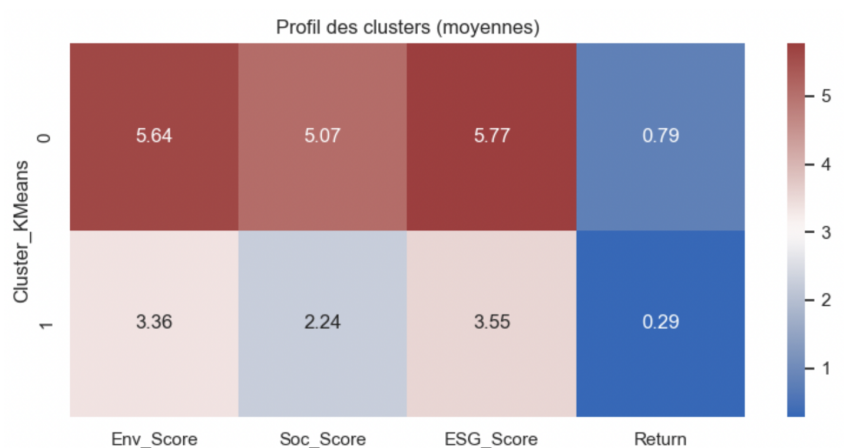


FIGURE 8 – Profil moyen des clusters.

Le cluster à ESG élevé présente un rendement moyen supérieur. Cette segmentation suggère l'existence d'un gradient ESG associé à une différence de performance moyenne.

Cependant, cette relation reste descriptive et peut refléter :

- des différences sectorielles,
- des effets taille,
- des caractéristiques fixes d'entreprise.

## 4.6 Régressions linéaires

Nous estimons trois modèles :

- Modèle 1 : relation simple ESG–rendement,
- Modèle 2 : inclusion simultanée des composantes ESG,
- Modèle 3 : ajout d'effets fixes temporels.

	Modèle 1	Modèle 2	Modèle 3
Coefficient ESG	0.092	-0.074	0.112
p-value	0.011	0.772	0.009
$R^2$	0.024	0.043	0.278

Le modèle simple montre un effet positif et significatif du score ESG sur les rendements.

Après introduction d'effets fixes temporels, le coefficient reste positif et significatif, suggérant que la relation ne s'explique pas uniquement par des tendances macroéconomiques communes.

L'inclusion simultanée des composantes ESG rend les coefficients instables, confirmant la présence de multicollinéarité.

## 5 Conclusion

### 5.1 Synthèse des résultats

Les analyses descriptives (corrélations simples, visualisation graphique et analyse par quartiles) mettent en évidence une association positive mais modérée entre le score ESG et les rendements boursiers. Cette relation s'accompagne toutefois d'une forte dispersion des observations, suggérant que le score ESG n'explique qu'une part limitée de la variabilité totale des rendements.

Le clustering K-means (avec  $k = 2$  retenu selon le critère de silhouette) identifie principalement une différenciation selon le niveau ESG. Le groupe caractérisé par des scores ESG élevés présente un rendement moyen supérieur au groupe à ESG plus faible. Néanmoins, la structure observée apparaît davantage comme un gradient ESG continu que comme une segmentation économique fortement structurée. La projection ACP confirme que la première composante principale capture essentiellement un facteur ESG commun dominant.

Les régressions linéaires confirment un effet positif et statistiquement significatif du score ESG sur les rendements. Cet effet persiste après introduction d'effets temporels annuels, suggérant qu'il ne s'explique pas uniquement par des tendances macroéconomiques communes. Toutefois, la part de variance directement expliquée par le score ESG demeure

limitée, ce qui indique que l'ESG constitue au mieux un déterminant partiel de la performance boursière moyenne.

## 5.2 Implications et perspectives

Les résultats suggèrent l'existence d'une association positive mais modérée entre ESG et performance moyenne au sein de la supply chain étudiée.

Toutefois, l'analyse ne permet pas d'établir une relation causale entre performance ESG et rendement boursier. Elle met néanmoins en évidence un signal empirique cohérent à travers plusieurs approches complémentaires — descriptives, multivariées et économétriques.

Une validation plus robuste nécessiterait un panel élargi, l'introduction d'effets fixes entreprise, ainsi que l'intégration explicite de facteurs de risque financiers standards. Une extension vers des modèles panel plus avancés ou des approches quasi-expérimentales pourrait également permettre d'approfondir l'analyse de la relation entre performance extra-financière et performance financière.